

ECS2 2018–2019

Cours, TD, TP, devoirs

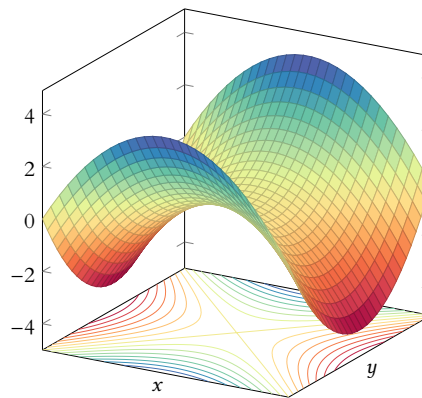
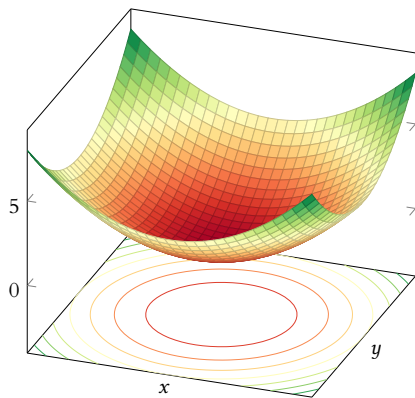
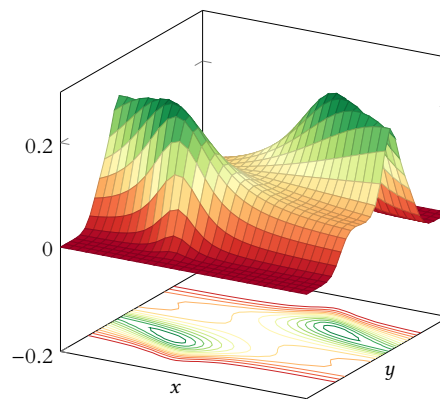
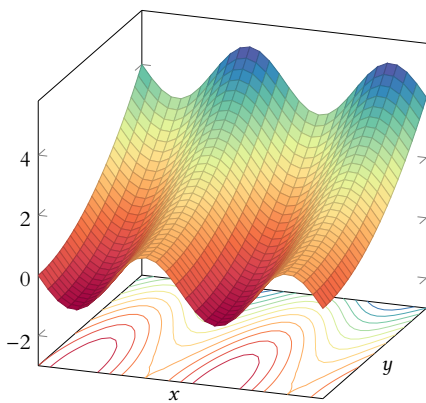


Table des matières

1 Révisions et compléments sur les suites et les séries.	9
Cours	9
Méthodes	23
Exercices	24
Correction des exercices	28
2 Variables aléatoires discrètes. Espérance conditionnelle.	45
Cours	45
Méthodes	56
Exercices	57
Correction des exercices	60
3 Révisions et compléments d'algèbre linéaire	71
Cours	71
Méthodes	87
Exercices	89
Correction des exercices	92
4 Intégrales impropres	103
Cours	103
Méthodes	123
Exercices	125
Correction des exercices	128
5 Matrices	143
Cours	143
Méthodes	158
Exercices	160
Correction des exercices	163
6 Couples de variables aléatoires discrètes.	175
Cours	175
Méthodes	190
Exercices	192
Correction des exercices	195
7 Valeurs propres, vecteurs propres	209
Cours	209
Méthodes	228
Exercices	229
Correction des exercices	232

8	Variables aléatoires à densité.	245
	Cours 245
	Méthodes 262
	Exercices 264
	Correction des exercices 267
9	Produits scalaires, espaces euclidiens	281
	Cours 281
	Méthodes 295
	Exercices 296
	Correction des exercices 299
10	Diagonalisation	309
	Cours 309
	Méthodes 324
	Exercices 325
	Correction des exercices 328
11	Vecteurs aléatoires	341
	Cours 341
	Exercices 351
	Correction des exercices 355
12	Supplémentaire orthogonal. Projecteurs orthogonaux	373
	Cours 373
	Méthodes 380
	Exercices 381
	Correction des exercices 384
13	Introduction aux fonctions de plusieurs variables	393
	Cours 393
	Méthodes 411
	Exercices 412
	Correction des exercices 415
14	Convergence des variables aléatoires	427
	Cours 427
	Méthodes 444
	Exercices 445
	Correction des exercices 448
15	Endomorphismes et matrices symétriques	461
	Cours 461
	Méthodes 469
	Exercices 470
	Correction des exercices 472
16	Estimation ponctuelle	481
	Cours 481
	Méthodes 492
	Exercices 493
	Correction des exercices 496

17	Topologie de \mathbb{R}^n, calcul différentiel d'ordre 2.	507
	Cours 507
	Exercices 527
	Correction des exercices 530
18	Estimation par intervalle de confiance	545
	Cours 545
	Exercices 552
	Correction des exercices 553
19	Optimisation sous contraintes	559
	Cours 559
	Méthodes 568
	Exercices 570
	Correction des exercices 572
20	Révisions d'analyse	583
	Exercices 584
	Correction des exercices 589
	EML 2015 609
	Correction 611
	ECRICOME 2009 617
	Correction 618
	ECRICOME 2007 622
	Correction 623
21	Révisions de probabilités	625
	Exercices 626
	Correction des exercices 631
	EML 2018 651
	Correction 653
	ECRICOME 2010 659
	Correction 662
	EDHEC 2013 669
	Correction 670
22	Révisions d'algèbre	673
	Exercices 674
	Correction des exercices 678
	EDHEC 2008 696
	Correction 697
	EDHEC 2012 699
	Correction 700
	EML 2003 702
	Correction 704
23	Devoirs à la maison	707
	Devoir maison 0 (à rendre le 03.09.18) 708
	Correction 712
	Devoir maison 1 (à rendre le 13.09.18) 724
	Correction 726
	Devoir maison 2 (à rendre le 20.09.18) 729
	Correction 730
	Devoir maison 2 sujet dur (à rendre le 20.09.18) 733
	Correction 735

Devoir maison 3 (à rendre le 04.10.18)	.742
Correction	.744
Devoir maison 4 (à rendre le 11.10.18)	.748
Correction	.750
Devoir maison 4 bis (à rendre le 11.10.18)	.755
Correction	.758
Devoir maison 4 (à rendre le 11.10.18)	.768
Correction	.771
Devoir maison 5 (à rendre le 18.10.18)	.782
Correction	.783
Devoir maison 6 (à rendre le 15.11.18)	.785
Correction	.787
Devoir maison 7 (à rendre le 22.11.18)	.793
Correction	.795
Devoir maison 8 (à rendre le 6.12.18)	.800
Correction	.801
Devoir maison 9 (à rendre le 13.12.18)	.806
Correction	.808
Devoir maison 9 (HEC) (à rendre le 20.12.18)	.812
Correction	.815
Devoir maison 10 (à rendre le 20.12.18)	.824
Correction	.825
Devoir maison 11 (à rendre le 10.01.19)	.827
Correction	.831
Devoir maison 12 (à rendre le 17.01.19)	.846
Correction	.847
Devoir maison 13 (à rendre le 25.01.19)	.850
Correction	.851
Devoir maison 14 (à rendre le 05.02.19)	.853
Correction	.855
Devoir maison 15 (SciLab) (à rendre le 14.03.19)	.862
Correction	.864
Devoir maison 16 (Maths II) (à rendre le 14.03.19)	.868
Correction	.871
Devoir maison 17 (à rendre le 21.03.19)	.883
Correction	.884
24 Devoirs en classe	887
Devoir surveillé 1 (08.09.18)	.888
Correction	.891
Devoir surveillé 2 (29.09.18)	.899
Correction	.902
Concours blanc 1 (07.11.18)	.912
Correction	.916
Devoir surveillé 4 (01.12.18)	.928
Correction	.931
Devoir surveillé 5 (12.01.19)	.939
Correction	.942
Devoir surveillé 6 (09.02.19)	.951
Correction	.955
Concours blanc 2 (05.03.19)	.965
Correction	.969
Concours blanc 2 : Maths II (08.03.19)	.981
Correction	.985
Devoir surveillé 8 (23.03.19)	.998
Correction	.1001
Devoir surveillé 8 (HEC) (23.03.19)	.1010

Correction	.1014
Devoir surveillé 8 : Maths II (23.03.19)	.1027
Correction	.1031
25 Travaux pratiques de SciLab	1045
Liste des commandes SciLab exigibles.	.1046
TP 1 : Manipulation de matrices.	.1048
Correction	.1052
TP 2 : Simulation de lois discrètes	.1056
Correction	.1060
TP 3 : Statistiques à une variable.	.1066
Correction	.1069
TP 4 : Chaînes de Markov I	.1073
Correction	.1076
TP 5 : Chaînes de Markov II	.1081
Correction	.1085
TP 6 : Simulation de lois à densité	.1087
Correction	.1091
TP 7 : Statistiques bivariées	.1094
Correction	.1098
TP 8 : Révisions	.1103
Correction	.1107
TP 9 : Fonctions de plusieurs variables.	.1115
Correction	.1118
TP 10 : Estimation, méthodes de Monte-Carlo.	.1121
Correction	.1124
Index	1127

RÉVISIONS ET COMPLÉMENTS SUR LES SUITES ET LES SÉRIES.

Cours	9
1.1 Suites convergentes	9
1.2 Comparaison de suites	10
1.3 Séries numériques	13
1.4 Familles sommables	18
Méthodes	23
Exercices	24
Correction des exercices	28

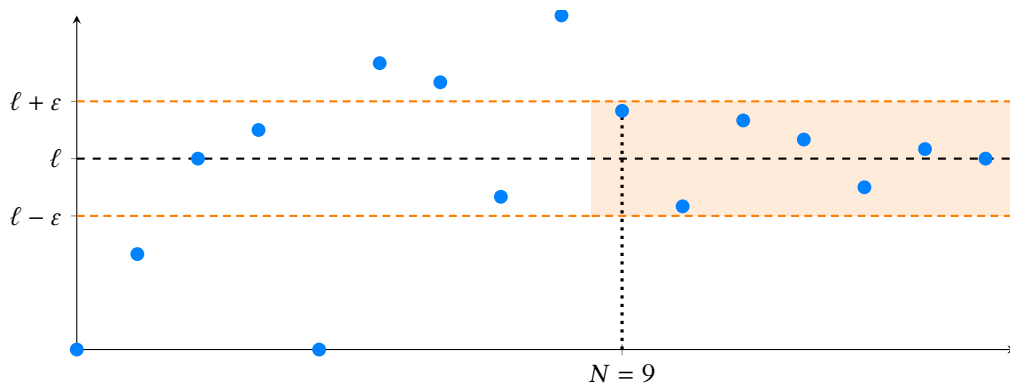
Ce chapitre est essentiellement composé de rappels de première année sur les suites et les séries. Pour cette raison, nous donnons peu de preuves, celles-ci se trouvant dans votre cours de première année.

1.1 SUITES CONVERGENTES

Définition 1.1 – Une suite $(u_n)_{n \in \mathbf{N}}$ est convergente si il existe $\ell \in \mathbf{R}$ tel que

$$\forall \varepsilon > 0, \exists N \in \mathbf{N}, n \geq N \Rightarrow |u_n - \ell| < \varepsilon.$$

Un tel réel ℓ , si il existe, est alors unique et est appelé limite de la suite (u_n) . Si (u_n) n'est pas convergente, on dit que (u_n) diverge.



Remarques

- Il n'y a pas unicité de N , par exemple ici $N = 10$ convient également, et plus généralement, tout $N \geq 9$ convient.
- N dépend de ε : si ε diminue, il faudra sûrement augmenter la valeur de N .
- Il se peut que pour (un ou plusieurs) $n < N$, on ait également $|u_n - \ell| < \varepsilon$. C'est par exemple ici le cas pour $n = 7$.

FIGURE 1.1 – La suite (u_n) converge vers ℓ : pour $n \geq N$, u_n est dans la «bande» $]l - \varepsilon, l + \varepsilon[$.

Remarque. De manière équivalente, cela signifie que pour tout intervalle ouvert I contenant ℓ , I contient tous les termes de la suite, sauf un nombre fini.

Définition 1.2 – On dit que (u_n) diverge vers $+\infty$ (respectivement $-\infty$) si

$$\forall A \in \mathbf{R}, \exists N \in \mathbf{N}, n \geq N \Rightarrow u_n \geq A \text{ (resp. } u_n \leq A).$$

Définition 1.3 – Une suite (u_n) est majorée si il existe $M \in \mathbf{R}$ tel que pour tout $n \in \mathbf{N}$, $u_n \leq M$.

Une suite (u_n) est minorée si il existe $m \in \mathbf{R}$ tel que pour tout $n \in \mathbf{N}$, $m \leq u_n$.

Une suite (u_n) est bornée si elle est à la fois majorée et minorée. C'est le cas si et seulement si il existe $M \in \mathbf{R}^+$ tel que pour tout $n \in \mathbf{N}$, $|u_n| \leq M$.

Théorème 1.4 (de la limite monotone) : • Soit (u_n) une suite croissante. Alors

1. si (u_n) est majorée, elle converge et $\lim_{n \rightarrow +\infty} u_n = \sup\{u_n, n \in \mathbf{N}\}$.

2. si (u_n) n'est pas majorée, alors $\lim_{n \rightarrow +\infty} u_n = +\infty$.

• Soit (u_n) une suite décroissante. Alors

1. si (u_n) est minorée, elle converge et $\lim_{n \rightarrow +\infty} u_n = \inf\{u_n, n \in \mathbf{N}\}$.

2. si (u_n) n'est pas minorée, alors $\lim_{n \rightarrow +\infty} u_n = -\infty$.

Proposition 1.5 : Soit (u_n) une suite. Alors (u_n) est convergente si et seulement si les deux suites $(u_{2n})_{n \in \mathbf{N}}$ et $(u_{2n+1})_{n \in \mathbf{N}}$ convergent vers une même limite.

Remarque. Il est important que les limites des deux suites (u_{2n}) et (u_{2n+1}) soient les mêmes. Par exemple, pour $u_n = (-1)^n$, les suites (u_{2n}) et (u_{2n+1}) sont constantes, donc convergentes. Mais la première converge vers 1, la seconde vers -1 et (u_n) diverge.

Démonstration. Supposons que (u_n) converge, et soit $\ell = \lim_{n \rightarrow +\infty} u_n$.

Soit I un intervalle ouvert contenant ℓ .

Alors tous les termes de la suite, sauf un nombre fini sont dans I . En particulier, tous les termes pairs, sauf un nombre fini sont dans I et donc $\lim_{n \rightarrow +\infty} u_{2n} = \ell$.

De même, tous les termes impairs, sauf un nombre fini sont dans I et donc $\lim_{n \rightarrow +\infty} u_{2n+1} = \ell$.

Inversement, supposons que les deux suites (u_{2n}) et (u_{2n+1}) convergent vers ℓ .

Alors si I est un intervalle ouvert contenant ℓ , seuls un nombre fini de termes d'ordre pair sont hors de I , et seuls un nombre fini de termes d'ordre impair sont hors de I .

Par conséquent, il n'y a qu'un nombre fini de termes de la suite (u_n) hors de I : (u_n) converge vers ℓ . \square

Définition 1.6 – Deux suites (u_n) et (v_n) sont adjacentes si :

1. (u_n) est croissante
2. (v_n) est décroissante
3. $\lim_{n \rightarrow +\infty} (u_n - v_n) = 0$.

Proposition 1.7 : Si (u_n) et (v_n) sont adjacentes, alors elles convergent vers une même limite.

Méthode

Le théorème de la limite monotone est très efficace pour montrer qu'une suite monotone converge, et c'est souvent le premier résultat auquel penser. En revanche il ne permet que rarement de déterminer la limite de la suite.

Remarque

La suite (u_{2n}) est la suite formée des termes d'ordre pair de (u_n) : u_0, u_2, u_4, \dots . Et de même, (u_{2n+1}) est la suite formée des termes d'ordre impair de (u_n) : u_1, u_3, u_5, \dots .

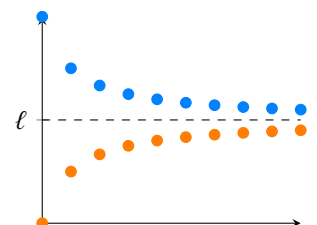


FIGURE 1.2– Deux suites adjacentes convergent vers une même limite.

1.2 COMPARAISON DE SUITES

Nous rappelons ici quelques points importants sur les notations \sim et o .

Tous ces résultats sont ici énoncés en termes de suites, mais restent valables pour des fonctions, que ce soit au voisinage de $\pm\infty$ ou d'un réel a .

1.2.1 Définitions, opérations usuelles

Définition 1.8 – Soient (u_n) et (v_n) deux suites. On dit que (u_n) est négligeable devant (v_n) et on note $u_n = o_{n \rightarrow +\infty}(v_n)$ s'il existe une suite $(\varepsilon_n)_{n \in \mathbf{N}}$ telle que $\lim_{n \rightarrow +\infty} \varepsilon_n = 0$ et $\forall n \in \mathbf{N}, u_n = \varepsilon_n v_n$.
Si pour tout $n \in \mathbf{N}, v_n \neq 0$, ceci est équivalent à $\lim_{n \rightarrow +\infty} \frac{u_n}{v_n} = 0$.

Cas particulier : on a $u_n = o_{n \rightarrow +\infty}(1)$ si et seulement si $\lim_{n \rightarrow +\infty} \frac{u_n}{1} = 0 \Leftrightarrow \lim_{n \rightarrow +\infty} u_n = 0$.

! La notation $o(u_n)$, bien que très pratique et très puissante, est aussi difficile à manipuler. En effet, $o(u_n)$ ne désigne pas une seule suite, mais n'importe quelle suite négligeable devant u_n .

Certains calculs sont alors contraires à l'intuition : par exemple $o(u_n) - o(u_n)$ n'est pas nécessairement¹ la suite nulle. On a juste $o(u_n) - o(u_n) = o(u_n)$.

Ainsi, $\frac{1}{n^2} = o_{n \rightarrow +\infty}\left(\frac{1}{n}\right)$ et $\frac{1}{n^3} = o_{n \rightarrow +\infty}\left(\frac{1}{n}\right)$, mais $\frac{1}{n^2} - \frac{1}{n^3} \neq 0$, mais est encore négligeable² devant $\frac{1}{n}$.

De même, on a $2 \times o(u_n) = o(2u_n) = o(u_n)$.

Nous verrons en TD d'autres exemples de manipulation de la notation o .

En cas de doute, on reviendra toujours à la définition à l'aide du quotient.

Définition 1.9 – Soient (u_n) et (v_n) deux suites. On dit que (u_n) est équivalente à (v_n) et on note $u_n \underset{n \rightarrow +\infty}{\sim} v_n$ si $u_n - v_n = o_{n \rightarrow +\infty}(v_n)$ (ou encore $u_n = v_n + o_{n \rightarrow +\infty}(v_n)$).
Si pour tout $n \in \mathbf{N}, v_n \neq 0$, ceci est équivalent à $\lim_{n \rightarrow +\infty} \frac{u_n}{v_n} = 1$.

Remarque. Si $u_n \xrightarrow{n \rightarrow +\infty} \ell$, avec $\ell \neq 0$, alors $u_n \underset{n \rightarrow +\infty}{\sim} \ell$.

! En revanche, ceci n'est pas vrai pour une suite qui tend vers 0. Les seules suites équivalentes à 0 sont celles qui sont nulles à partir d'un certain rang. On n'écrira donc **jamais**³ $u_n \underset{n \rightarrow +\infty}{\sim} 0$, sauf si (u_n) est la suite nulle.

Nous rappelons ici quelques règles de calcul sur les équivalents et les o . La plupart se retrouvent en revenant aux définitions.

Proposition 1.10 : Soient $(u_n), (v_n)$ et (w_n) trois suites à termes non nuls. Alors

1. si $u_n = o_{n \rightarrow +\infty}(v_n)$ et si $v_n = o_{n \rightarrow +\infty}(w_n)$, alors $u_n = o_{n \rightarrow +\infty}(w_n)$.
2. si $u_n = o_{n \rightarrow +\infty}(v_n)$, alors $u_n w_n = o_{n \rightarrow +\infty}(v_n w_n)$.
3. si $u_n = o_{n \rightarrow +\infty}(v_n)$ et si $v_n \underset{n \rightarrow +\infty}{\sim} w_n$, alors $u_n = o_{n \rightarrow +\infty}(w_n)$.
4. si $u_n \underset{n \rightarrow +\infty}{\sim} v_n$ et si $v_n \underset{n \rightarrow +\infty}{\sim} w_n$, alors $u_n \underset{n \rightarrow +\infty}{\sim} w_n$.
5. si $u_n \underset{n \rightarrow +\infty}{\sim} v_n$ et si $w_n \underset{n \rightarrow +\infty}{\sim} z_n$, alors $u_n w_n \underset{n \rightarrow +\infty}{\sim} v_n z_n$.
6. si $\lambda \in \mathbf{R}$ et si $u_n \underset{n \rightarrow +\infty}{\sim} v_n$, alors $\lambda u_n \underset{n \rightarrow +\infty}{\sim} \lambda v_n$.
7. si $u_n \underset{n \rightarrow +\infty}{\sim} v_n$, alors pour tout $k \in \mathbf{N}$ fixé, $u_n^k \underset{n \rightarrow +\infty}{\sim} v_n^k$.
Plus généralement, si u_n et v_n sont à termes positifs, alors pour tout $\alpha \in \mathbf{R}^*$ fixé, $u_n^\alpha \underset{n \rightarrow +\infty}{\sim} v_n^\alpha$.

Remarque. Si la propriété 5 nous dit qu'il est possible de multiplier les équivalents, rappelons qu'on ne peut pas ajouter des équivalents.

Si on a besoin de l'équivalent d'une somme, on utilise généralement un développement limité.

Enfin, la proposition suivante est souvent utilisée pour étudier la convergence de séries :

Proposition 1.11 : Si $u_n \underset{n \rightarrow +\infty}{\sim} v_n$ et si (v_n) est à termes strictement positifs, alors il existe $N \in \mathbf{N}$ tel que pour $n \geq N, u_n \geq 0$.

Remarque

En pratique, on utilise toujours cette seconde définition.

¹ Bien que ce soit possible.

² Pour s'en convaincre, le plus simple est de passer au quotient.

Méthode

En pratique, pour montrer que deux suites sont équivalentes, on montre que leur quotient tend vers 1.

³ Mon prof de prépa avait l'habitude de dire «on n'est jamais équivalent à 0, sauf si on est complètement nul !».

Remarque

En particulier, pour $\alpha = \frac{1}{2}$,
 $\sqrt{u_n} \underset{n \rightarrow +\infty}{\sim} \sqrt{v_n}$.



Démonstration. Puisque $\frac{u_n}{v_n} \xrightarrow[n \rightarrow +\infty]{} 1$, alors il existe⁴ $N \in \mathbf{N}$ tel que pour $n \geq N$, $\frac{u_n}{v_n} \geq 0$ et donc $u_n \geq 0$. \square

⁴ Cela revient à prendre $\varepsilon = 1$ dans la définition de la convergence :

$$\left| \frac{u_n}{v_n} - 1 \right| \leq 1 \Rightarrow \frac{u_n}{v_n} \geq 0.$$

1.2.2 Croissances comparées usuelles

Les limites suivantes sont à connaître par cœur :

Proposition 1.12 : Pour $a > 0, b > 0$ et $q > 1$, on a

$$\begin{aligned} 1. (\ln n)^a &= o_{n \rightarrow +\infty}(n^b) & 2. n^b &= o_{n \rightarrow +\infty}(q^n) \\ 3. q^n &= o_{n \rightarrow +\infty}(n!) & 4. n! &= o_{n \rightarrow +\infty}(n^n). \end{aligned}$$

Démonstration. Seule la dernière n'a peut-être pas été prouvée en première année. L'idée est de remarquer que $n! = \underbrace{1 \times 2 \times \dots \times n}_{n \text{ termes}}$ et que $n^n = \underbrace{n \times n \times \dots \times n}_{n \text{ termes}}$ et que donc

$$\frac{n!}{n^n} = \underbrace{\frac{1}{n}}_{\xrightarrow[n \rightarrow +\infty]{} 0} \underbrace{\frac{2}{n} \dots \frac{n-1}{n} \frac{n}{n}}_{\leq 1} \xrightarrow[n \rightarrow +\infty]{} 0.$$

\square

1.2.3 Développements limités et équivalents

Nous ne redonnons pas la définition ni les propriétés des développements limités. Rappelons juste deux choses fondamentales :

- on peut ajouter ou multiplier⁵ des développements limités de même ordre n ;
- si $f(x) = \lambda_0 + \lambda_1(x-a) + \dots + \lambda_n(x-a)^n + o_{x \rightarrow a}((x-a)^n)$ est le développement limité de f au voisinage de a , et si $p = \min\{k \in \llbracket 0, n \rrbracket : \lambda_k \neq 0\}$, alors $f(x) \underset{x \rightarrow a}{\sim} \lambda_p(x-a)^p$.

Les développements limités suivants sont à connaître par cœur :

Proposition 1.13 : Au voisinage de 0, on a :

$$\begin{aligned} e^x &= 1 + x + \frac{x^2}{2!} + \dots + \frac{x^n}{n!} + o(x^n) = \sum_{k=0}^n \frac{x^k}{k!} + o(x^n) \\ \cos x &= 1 - \frac{x^2}{2} + \frac{x^4}{4!} + \dots + \frac{(-1)^n}{(2n)!} x^{2n} + o(x^{2n+1}) = \sum_{k=0}^n \frac{(-1)^k}{(2k)!} x^{2k} + o(x^{2n+1}) \\ \sin x &= x - \frac{x^3}{6} + \dots + \frac{(-1)^n}{(2n+1)!} x^{2n+1} + o(x^{2n+2}) = \sum_{k=0}^n \frac{(-1)^k}{(2k+1)!} x^{2k+1} + o(x^{2n+2}) \\ \frac{1}{1-x} &= 1 + x + x^2 + \dots + x^n + o(x^n) = \sum_{k=0}^n x^k + o(x^n) \\ \frac{1}{1+x} &= 1 - x + x^2 + \dots + (-1)^n x^n + o(x^n) = \sum_{k=0}^n (-1)^k x^k + o(x^n) \\ \ln(1+x) &= x - \frac{x^2}{2} + \dots + \frac{(-1)^{n-1}}{n} x^n + o(x^n) = \sum_{k=1}^n \frac{(-1)^{k-1}}{k} x^k + o(x^n) \\ \ln(1-x) &= -x - \frac{x^2}{2} - \dots - \frac{x^n}{n} + o(x^n) = -\sum_{k=1}^n \frac{x^k}{k} + o(x^n) \\ (1+x)^\alpha &= 1 + \alpha x + \frac{\alpha(\alpha-1)}{2!} x^2 + \dots + \frac{\alpha(\alpha-1) \dots (\alpha-n+1)}{n!} x^n + o(x^n) \end{aligned}$$

Remarque. De ces développements limités découlent tous les équivalents usuels :

$$e^x - 1 \underset{x \rightarrow 0}{\sim} x, \ln(1+x) \underset{x \rightarrow 0}{\sim} x, (1+x)^\alpha - 1 \underset{x \rightarrow 0}{\sim} \alpha x, \dots$$

⁵ En ne gardant que les termes d'ordre inférieur ou égal à n .

Autrement dit

$f(x)$ est équivalent au premier terme **non nul** de son développement limité.

Astuce

Si on connaît le DL de $\frac{1}{1-x}$, celui de $\frac{1}{1+x}$ s'obtient en changeant les x en $-x$

Astuce

Ce développement limité se retrouve en intégrant celui de $\frac{1}{1+x}$ (qui est la dérivée de $\ln(1+x)$).

En particulier

Pour $\alpha = \frac{1}{2}$ et $\alpha = -\frac{1}{2}$, on retrouve les DL de $\sqrt{1+x}$ et $\frac{1}{\sqrt{1+x}}$.

Comme mentionné précédemment, les développements limités permettent d'obtenir des équivalents de sommes.

Exemple 1.14

Soit $(u_n)_{n \geq 1}$ la suite définie par $u_n = e^{-\frac{1}{2n^2}} - \cos \frac{1}{n}$.

Alors, puisque $\frac{1}{n} \xrightarrow[n \rightarrow +\infty]{} 0$, on a

$$\begin{aligned} u_n &= 1 - \frac{1}{2n^2} + \frac{1}{2} \frac{1}{(2n^2)^2} + o_{n \rightarrow +\infty} \left(\frac{1}{n^4} \right) - \left(1 - \frac{1}{2n^2} + \frac{1}{24n^4} + o_{n \rightarrow +\infty} \left(\frac{1}{n^4} \right) \right) \\ &= \frac{1}{8n^4} - \frac{1}{24n^4} + o_{n \rightarrow +\infty} \left(\frac{1}{n^4} \right) \\ &= \frac{1}{12n^4} + o_{n \rightarrow +\infty} \left(\frac{1}{n^4} \right) \\ &\underset{n \rightarrow +\infty}{\sim} \frac{1}{12n^4}. \end{aligned}$$

Enfin, donnons juste un équivalent classique, qu'il faut savoir retrouver à défaut de le connaître par cœur :

Proposition 1.15 : On a $\ln(x) \underset{x \rightarrow 1}{\sim} x - 1$.

Démonstration. Il s'agit de remarquer que $\ln(x) = \ln(1 + (x - 1))$, avec $x - 1 \xrightarrow[x \rightarrow 1]{} 0$, et donc en utilisant le classique $\ln(1 + u) \underset{u \rightarrow 0}{\sim} u$, on a donc

$$\ln(x) = \ln(1 + (x - 1)) \underset{x \rightarrow 1}{\sim} x - 1. \quad \square$$

1.3 SÉRIES NUMÉRIQUES

1.3.1 Définition, premières propriétés

Définition 1.16 – Soit (u_n) une suite. On dit que la série de terme général u_n , ou la série $\sum u_n$, converge si

$$\lim_{n \rightarrow +\infty} \sum_{k=0}^n u_k$$

existe. Dans ce cas, on note alors $\sum_{k=0}^{+\infty} u_k$ cette limite, appelée somme de la série.

Dans le cas contraire, on dit que la série $\sum u_n$ diverge.

Terminologie

Si on note $S_n = \sum_{k=0}^n u_k$, on dit que S_n est la somme partielle d'ordre n de la série $\sum u_k$. Ainsi, une série converge si et seulement si la suite de ses sommes partielles converge.

Exemple 1.17

$\sum_{k \geq 1} \frac{1}{k(k+1)}$ converge car

$$\sum_{k=1}^n \frac{1}{k(k+1)} = \sum_{k=1}^n \left(\frac{1}{k} - \frac{1}{k+1} \right) = \sum_{k=1}^n \frac{1}{k} - \sum_{k=1}^n \frac{1}{k+1} = \sum_{k=1}^n \frac{1}{k} - \sum_{k=2}^{n+1} \frac{1}{k} = 1 - \frac{1}{n+1} \xrightarrow[n \rightarrow +\infty]{} 1.$$

Terminologie

Une telle série où les termes successifs se simplifient deux à deux est appelée série télescopique.

Proposition 1.18 : Si $\sum u_n$ converge, alors $u_n \xrightarrow[n \rightarrow \infty]{} 0$.

Par contraposée, si $u_n \not\rightarrow 0$, alors $\sum u_n$ diverge, on parle alors de divergence grossière.

⚠ La réciproque est fautive ! Par exemple $\sum \frac{1}{n}$ diverge bien que $\frac{1}{n} \rightarrow 0$.

Exemple 1.19

Déterminons la nature de la série de terme général $u_n = \left(1 - \frac{1}{2n}\right)^n = e^{n \ln\left(1 - \frac{1}{2n}\right)}$

Nous savons que

$$\ln\left(1 - \frac{1}{2n}\right) \underset{n \rightarrow +\infty}{\sim} -\frac{1}{2n} \text{ et donc } n \ln\left(1 - \frac{1}{2n}\right) \underset{n \rightarrow +\infty}{\sim} -\frac{1}{2}.$$

Par continuité de l'exponentielle, on a alors $u_n \xrightarrow[n \rightarrow +\infty]{} e^{-1/2} \neq 0$.

Donc $\sum u_n$ diverge grossièrement.

⚠ Danger !

1^∞ est une forme indéterminée. Pour lever l'indétermination, on passe systématiquement par la forme exponentielle.

Proposition 1.20 : Soient (u_n) et (v_n) deux suites, $\alpha, \beta \in \mathbf{R}$. Si $\sum u_n$ et $\sum v_n$ convergent, alors $\sum (\alpha u_n + \beta v_n)$ converge et

$$\sum_{n=0}^{+\infty} (\alpha u_n + \beta v_n) = \alpha \sum_{n=0}^{+\infty} u_n + \beta \sum_{n=0}^{+\infty} v_n.$$

Définition 1.21 – Si $\sum_n u_n$ est une série convergente, alors pour tout $n \in \mathbf{N}$, la

série $\sum_{k \geq n+1} u_k$ converge, et sa somme $R_n = \sum_{k=n+1}^{+\infty} u_k$ est appelé reste d'ordre n de la série.

Remarque

On ne parle du reste d'une série que dans le cas d'une série convergente, il n'est pas défini sinon.

Proposition 1.22 : Si $\sum u_n$ converge et si $R_n = \sum_{k=n+1}^{+\infty} u_k$, alors $\lim_{n \rightarrow +\infty} R_n = 0$.

Démonstration. Notons $S = \sum_{k=0}^{+\infty} u_k$. Alors

$$R_n = \sum_{k=n+1}^{+\infty} u_k = \sum_{k=0}^{+\infty} u_k - \sum_{k=0}^n u_k = S - \sum_{k=0}^n u_k.$$

Mais puisque $\sum u_k$ converge, on a $\lim_{n \rightarrow +\infty} \sum_{k=0}^n u_k = S$ et donc $\lim_{n \rightarrow +\infty} R_n = S - S = 0$. \square

1.3.2 Séries à termes positifs

Proposition 1.23 : Soit $\sum u_n$ une série à termes positifs. Alors $\sum u_n$ converge si et seulement si la suite de ses sommes partielles est majorée, c'est-à-dire s'il existe $M \in \mathbf{R}$ tel que

$$\forall n \in \mathbf{N}, \sum_{k=0}^n u_k \leq M.$$

Démonstration. Notons $S_n = \sum_{k=0}^n u_k$ la suite des sommes partielles de la série.

On a alors $S_{n+1} - S_n = \sum_{k=0}^{n+1} u_k - \sum_{k=0}^n u_k = u_{n+1} \geq 0$, et donc $(S_n)_{n \in \mathbf{N}}$ est une suite croissante.

Par le théorème de la limite monotone, elle converge si et seulement si elle est majorée. \square

Remarque. En particulier, si $\sum u_n$ est une série à termes positifs qui diverge, la suite de ses sommes partielles n'est pas majorée, et donc, par le théorème de la limite monotone, tend vers $+\infty$.

Proposition 1.24 : Si (u_n) et (v_n) sont deux suites telles que $\forall n \in \mathbf{N}, 0 \leq u_n \leq v_n$, alors :

- si $\sum v_n$ converge, alors $\sum u_n$ converge
- si $\sum u_n$ diverge, alors $\sum v_n$ diverge.

Attention !

Si l'on sait juste dire que la plus grande diverge ou que la plus petite converge, cela ne nous apprend rien sur l'autre !

Exemples 1.25

- Soit $u_n = \frac{|\cos n|}{n^2}$. Alors $0 \leq u_n \leq \frac{1}{n^2}$, et nous savons que la série $\sum \frac{1}{n^2}$ converge (série de Riemann convergente), et donc $\sum u_n$ converge.
- Soit $v_n = \frac{1 + \ln n}{n}$, pour $n \geq 1$. Alors $0 \leq \frac{1}{n} \leq v_n$. Nous savons que la série $\sum \frac{1}{n}$ diverge (série harmonique), donc $\sum v_n$ diverge.

Proposition 1.26 : Si (u_n) et (v_n) sont deux suites à termes positifs⁶ telles que $u_n \underset{n \rightarrow +\infty}{\sim} v_n$, alors $\sum u_n$ et $\sum v_n$ sont de même nature, i.e.

- si $\sum u_n$ converge, $\sum v_n$ converge
- si $\sum u_n$ diverge, $\sum v_n$ diverge.

⁶ En fait, ce résultat reste valable si les deux séries sont à termes négatifs, l'essentiel étant qu'elles soient de signe constant. Et il suffit de prouver que l'une des deux est de signe constant, l'autre l'étant alors automatiquement par la proposition 1.12

Exemples 1.27

- Soit $u_n = \ln \left(\frac{2n}{2n+1} \right)$. Alors $\frac{2n}{2n+1} \underset{n \rightarrow +\infty}{\rightarrow} 1$ et donc

$$u_n \underset{n \rightarrow +\infty}{\sim} \frac{2n}{2n+1} - 1 = \frac{-1}{2n+1} \underset{n \rightarrow +\infty}{\sim} -\frac{1}{2n}.$$

Or, $\sum -\frac{1}{2n}$ diverge (et est de signe constant), donc $\sum u_n$ diverge.

- Soit $v_n = 1 - \cos \frac{1}{n}$. Alors $\forall n \in \mathbf{N}, v_n \geq 0$, et

$$v_n = 1 - \left(1 - \frac{1}{2n^2} + o\left(\frac{1}{n^2}\right) \right) = \frac{1}{2n^2} + o\left(\frac{1}{n^2}\right) \sim \frac{1}{2n^2}.$$

Puisque la série $\sum \frac{1}{2n^2}$ converge et est à termes positifs, il en est de même de $\sum u_n$.

1.3.3 Séries absolument convergentes

Définition 1.28 – Soit (u_n) une suite. On dit que la série $\sum u_n$ est absolument convergente si $\sum |u_n|$ est convergente.

Remarque. Pour une série à termes positifs⁷, la convergence est équivalente à la convergence absolue car $u_n = |u_n|$.

⁷ De manière plus générale, ceci est valable pour les séries dont le terme général est de signe constant.

Proposition 1.29 : Si $\sum u_n$ est absolument convergente, elle est convergente et

$$\left| \sum_{n=0}^{+\infty} u_n \right| \leq \sum_{n=0}^{+\infty} |u_n| \quad (\text{inégalité triangulaire généralisée}).$$

⚠ Réciproque fausse

Il existe des séries qui sont convergentes et non absolument convergentes, par exemple $\sum \frac{(-1)^n}{n}$.

Notons que la suite $(|u_n|)$ est toujours positive. Aussi, pour prouver la convergence absolue de $\sum u_n$, on pourra appliquer à $\sum |u_n|$ les résultats relatifs aux séries à termes positifs.

Exemple 1.30

Soit $u_n = \frac{\sin(\sqrt{n})}{n^2}$. Alors $0 \leq |u_n| \leq \frac{|\sin \sqrt{n}|}{n^2} \leq \frac{1}{n^2}$.

Puisque la série $\sum \frac{1}{n^2}$ converge, il en est de même de $\sum |u_n|$.

Ainsi $\sum u_n$ est absolument convergente, et donc convergente.

Proposition 1.31 : Soient (u_n) et (v_n) deux suites telles que

- $\sum v_n$ soit absolument convergente
- $u_n = o(v_n)$.

Alors $\sum u_n$ est absolument convergente.

Remarque

Ce résultat s'applique en particulier si $v_n \geq 0$ et $\sum v_n$ converge.

Exemple 1.32

$$u_n = e^{-(\ln n)^2}.$$

Alors $\frac{u_n}{\frac{1}{n^2}} = n^2 u_n = e^{2 \ln n - (\ln n)^2} \xrightarrow{n \rightarrow +\infty} 0$. Donc $u_n = o\left(\frac{1}{n^2}\right)$.

Puisque $\sum \frac{1}{n^2}$ est absolument convergente, alors $\sum u_n$ est absolument convergente (et par conséquent convergente).

1.3.4 Séries de référence

Les séries suivantes sont les seules dont on doit connaître par cœur la nature, et que l'on doit reconnaître immédiatement.

Proposition 1.33 (Séries de Riemann) : Soit $\alpha \in \mathbf{R}$. Alors $\sum \frac{1}{n^\alpha}$ converge si et seulement si $\alpha > 1$.

Remarque. En règle générale, lorsque $\alpha > 1$, on ne connaît pas la valeur de $\sum_{n=1}^{+\infty} \frac{1}{n^\alpha}$, on sait juste que la série converge.

Quelques valeurs sont connues, comme $\sum_{n=1}^{+\infty} \frac{1}{n^2} = \frac{\pi^2}{6}$ ou $\sum_{n=1}^{+\infty} \frac{1}{n^4} = \frac{\pi^4}{90}$.

Plus généralement, il existe des formules pour $\sum_{n=1}^{+\infty} \frac{1}{n^\alpha}$ lorsque α est un entier pair, mais si α est un entier impair, ou n'est pas entier, on ne sait (presque) rien dire, et il s'agit d'un sujet de recherche actif.

Remarque

Une preuve de la (surprenant) formule

$$\sum_{n=1}^{+\infty} \frac{1}{n^2} = \frac{\pi^2}{6}$$

se trouve par exemple dans le problème 2 du sujet de Lyon 2005.

Exemple 1.34

$\sum \frac{1}{n\sqrt{n}} = \sum \frac{1}{n^{3/2}}$ converge.

$\sum \frac{1}{n}$ diverge. Cette série est appelée série harmonique.

Corollaire 1.35 (Règle $n^\alpha u_n$) – Soit (u_n) une suite réelle. S'il existe $\alpha > 1$ tel que $n^\alpha u_n \xrightarrow[n \rightarrow \infty]{} 0$, alors $\sum u_n$ converge absolument.

Démonstration. Soit $\alpha > 1$ tel que $n^\alpha u_n \xrightarrow[n \rightarrow \infty]{} 0$.

Alors $\frac{u_n}{\frac{1}{n^\alpha}} = n^\alpha u_n \rightarrow 0$, ce qui signifie donc que $u_n = o\left(\frac{1}{n^\alpha}\right)$. Puisque la série $\sum \frac{1}{n^\alpha}$ converge absolument, il en va de même de $\sum u_n$. \square

Exemples 1.36

- Soit $u_n = \frac{\ln n}{n^2}$. Cherchons $\alpha > 1$ tel que $n^\alpha u_n \rightarrow 0$. On a $n^\alpha u_n = \frac{\ln n}{n^{2-\alpha}}$. Par croissance comparées, si $2 - \alpha > 0 \Leftrightarrow \alpha < 2$, on a bien $\lim_{n \rightarrow +\infty} n^\alpha u_n = 0$.

Prenons par exemple $\alpha = \frac{3}{2}$.

Puisque $n^{3/2} u_n \rightarrow 0$, $\sum u_n$ converge.

- Soit $u_n = \left(1 + \frac{2}{\sqrt{n}}\right)^{-n}$. Alors $n^2 u_n = e^{2 \ln n - n \ln\left(1 + \frac{2}{\sqrt{n}}\right)}$. Or,

$$2 \ln n - n \ln\left(1 + \frac{2}{\sqrt{n}}\right) = \underbrace{2 \ln n}_{=o(\sqrt{n})} - n \left(\frac{2}{\sqrt{n}} + o\left(\frac{1}{\sqrt{n}}\right)\right) = 2 \ln n - 2\sqrt{n} + o(\sqrt{n}) \sim -2\sqrt{n}.$$

On en déduit que $2 \ln n - n \ln\left(1 + \frac{2}{\sqrt{n}}\right) \xrightarrow[n \rightarrow +\infty]{} -\infty$, et donc $n^2 u_n \rightarrow 0$.

Par la règle $n^\alpha u_n$, on en déduit que $\sum u_n$ converge.

Rappel

Si $u_n = v_n + o(v_n)$, alors $u_n \sim v_n$.



Proposition 1.37 (Série exponentielle) : Soit $x \in \mathbf{R}$. Alors $\sum \frac{x^n}{n!}$ converge, et

$$\sum_{n=0}^{+\infty} \frac{x^n}{n!} = e^x.$$

Cas particulier

Pour $x = 1$, on a alors

$$e = \sum_{n=0}^{+\infty} \frac{1}{n!}.$$

Proposition 1.38 (Série géométrique et séries géométriques dérivées) : Soit $q \in \mathbf{R}$. Alors les séries $\sum_{n \geq 0} q^n$, $\sum_{n \geq 1} nq^{n-1}$ et $\sum_{n \geq 2} n(n-1)q^{n-2}$ convergent si et seulement si $|q| < 1$. Dans ce cas, on a alors

$$\begin{aligned} \sum_{n=0}^{+\infty} q^n &= \frac{1}{1-q}. \\ \sum_{n=0}^{+\infty} nq^{n-1} &= \frac{1}{(1-q)^2}. \\ \sum_{n=0}^{+\infty} n(n-1)q^{n-2} &= \frac{2}{(1-q)^3}. \end{aligned}$$

Astuce

Si l'on dérive q^n par rapport à q , on obtient nq^{n-1} . Et si l'on dérive $\frac{1}{1-q}$, on obtient $\frac{1}{(1-q)^2}$. Il suffit donc d'appliquer la première relation, et de la dériver une ou deux fois pour retrouver les deux autres.

Remarque. Vous avez pu apprendre d'autres formules en première année pour les séries géométriques dérivées, à savoir

$$\sum_{n=1}^{+\infty} nq^n = \frac{q}{(1-q)^2} \text{ et } \sum_{n=1}^{+\infty} n^2q^n = \frac{q(1+q)}{(1-q)^3}.$$

Ces formules sont strictement équivalentes à celles données plus haut. En effet, on a

$$\sum_{n \geq 1} nq^{n-1} = \frac{1}{q} \sum_{n=1}^{+\infty} nq^n = \frac{1}{(1-q)^2} \text{ et}$$

$$\begin{aligned} \sum_{n=2}^{+\infty} n(n-1)q^{n-2} &= \sum_{n=1}^{+\infty} n(n-1)q^{n-2} = \frac{1}{q^2} \sum_{n=1}^{+\infty} n^2q^n - \frac{1}{q^2} \sum_{n=1}^{+\infty} nq^n \\ &= \frac{1}{q} \frac{1+q}{(1-q)^3} - \frac{1}{q} \frac{1}{(1-q)^2} = \frac{1+q-(1-q)}{q(1-q)^3} = \frac{2}{(1-q)^3}. \end{aligned}$$

Il suffit donc d'apprendre l'une des deux formules, et d'être capable de s'adapter si dans un énoncé on a besoin de l'autre⁸.

⁸ En pratique, on a aussi souvent besoin de l'une que de l'autre.

1.4 FAMILLES SOMMABLES

1.4.1 Retour sur la convergence absolue des séries



Proposition 1.39 : Soit $\sum u_n$ une série absolument convergente. Alors, quel que soit $\varphi : \mathbf{N} \rightarrow \mathbf{N}$ bijective, la série $\sum u_{\varphi(n)}$ converge absolument, et

$$\sum_{n=0}^{+\infty} u_{\varphi(n)} = \sum_{n=0}^{+\infty} u_n.$$

Démonstration. Admis. □

Ce résultat signifie que pour une série absolument convergente, l'ordre de sommation est sans importance. C'est fondamental pour définir l'espérance d'une variable aléatoire discrète, car la somme ne dépend pas de la «numérotation» qu'on a choisie pour $X(\Omega)$ (une telle numérotation consiste à ordonner les éléments de $X(\Omega)$).

⚠ C'est faux si $\sum u_n$ est convergente mais pas absolument convergente.

Par exemple, $\sum \frac{(-1)^n}{n}$ est convergente, mais pas absolument convergente. Il est possible (mais délicat) de montrer qu'il existe $\varphi : \mathbf{N} \rightarrow \mathbf{N}$ bijective (c'est-à-dire une «renumérotation» des éléments de \mathbf{N}) telle que $\sum u_{\varphi(n)}$ diverge, ou encore telle que $\sum_{n=0}^{+\infty} u_{\varphi(n)} = a$, a étant un réel quelconque dont dépendra φ .

1.4.2 Ensembles dénombrables

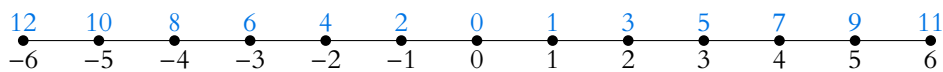
Définition 1.40 – Un ensemble I est dit dénombrable si il est fini ou s'il existe une bijection $\varphi : \mathbf{N} \rightarrow I$.

Remarque. Cela signifie que l'on peut «numéroter» les éléments de I : $\varphi(0)$ est le «premier» élément de I , $\varphi(1)$ le «second», etc.

Exemples 1.41

- \mathbf{N} est dénombrable
- \mathbf{N}^* est dénombrable, avec $\varphi : \mathbf{N} \rightarrow \mathbf{N}^*, n \mapsto n + 1$
- \mathbf{Z} est dénombrable, avec

$$\varphi : \mathbf{N} \rightarrow \mathbf{Z}, \quad n \mapsto \begin{cases} \lfloor \frac{n}{2} \rfloor + 1 & \text{si } n \text{ est impair} \\ -\frac{n}{2} & \text{si } n \text{ est pair} \end{cases}$$

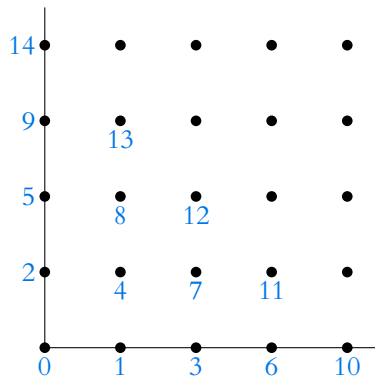


- $\mathbf{N} \times \mathbf{N}$ est dénombrable

$$\varphi : \mathbf{N} \times \mathbf{N} \rightarrow \mathbf{N}$$

$$(p, q) \mapsto \frac{(p + q + 1)(p + q)}{2} + p$$

est une bijection de $\mathbf{N} \times \mathbf{N}$ dans \mathbf{N} (difficile), donc $\varphi^{-1} : \mathbf{N} \rightarrow \mathbf{N} \times \mathbf{N}$ est une bijection.



- \mathbf{R} n'est pas dénombrable (très difficile, admis).

Proposition 1.42 :

- Si I et J sont dénombrables, alors $I \times J$ est dénombrable.
- Si $(I_j)_{j \in \mathbf{N}}$ est une suite d'ensembles dénombrables, alors $\bigcup_{j \in \mathbf{N}} I_j$ est encore dénombrable.

Notons qu'on peut désormais donner une définition plus rigoureuse de variable aléatoire discrète maintenant qu'on dispose de la notion de dénombrabilité :

Définition 1.43 – Une variable aléatoire réelle $X : \Omega \rightarrow \mathbf{R}$ est discrète si $X(\Omega)$ est dénombrable.

1.4.3 Familles sommables

Lorsqu'on manipule une série convergente $\sum u_n$, on somme une infinité d'éléments. Comme cela n'a a priori pas de sens, ce que nous faisons en réalité, c'est sommer un nombre fini de termes (ce sont les sommes partielles), puis un passage à la limite. Et comme il y a un ordre naturel sur \mathbf{N} , on somme les éléments de (u_n) dans cet ordre : d'abord u_0 , puis u_1 , puis u_2 , etc.

Imaginons maintenant que l'on ait une «grosse» famille de nombres, qui n'est plus indicée par \mathbf{N} , mais par un ensemble dénombrable plus gros. Par exemple $(u_{p,q})_{(p,q) \in \mathbf{N} \times \mathbf{N}}$. Cela signifie donc que pour tout couple $(p, q) \in \mathbf{N} \times \mathbf{N}$, on se donne un nombre $u_{p,q}$. Imaginons à présent que l'on s'intéresse à la somme de cette infinité de nombres. Comment lui donner un sens ?

Nous pouvons encore calculer des sommes finies, puis passer à la limite, mais dans quel ordre sommer les termes de la famille ? Il n'y a aucun ordre naturel sur $\mathbf{N} \times \mathbf{N}$, c'est-à-dire qu'il n'y a pas **un** plus petit élément de $\mathbf{N} \times \mathbf{N}$, puis un autre élément qui viendrait juste après, puis encore un autre, etc.

L'astuce consiste à numéroter les éléments de $\mathbf{N} \times \mathbf{N}$: il y a plein de manières de le faire, mais peu importe, numéroter les éléments de $\mathbf{N} \times \mathbf{N}$ revient à numéroter ceux de notre grosse famille de nombres. Il y en aura donc un premier, puis un deuxième, etc, et nous pourrons donc les sommer dans cet ordre.

C'est l'idée qui est à la base de la définition qui suit.

Définition 1.44 – Soit I un ensemble dénombrable, et soit $(u_i)_{i \in I}$ une famille de réels indexés par I . On dit que la famille $(u_i)_{i \in I}$ est sommable si pour une (et donc pour toute⁹) bijection $\varphi : \mathbf{N} \rightarrow I$, $\sum u_{\varphi(n)}$ est absolument convergente. La somme de cette série est appelée somme de la famille $(u_i)_{i \in I}$, et est notée

$$\sum_{i \in I} u_i.$$

⁹ C'est la proposition 1.39 : si on change l'ordre de sommation, cela ne change rien à la somme.

Remarque. Pour une suite¹⁰ $(u_n)_{n \in \mathbf{N}}$, la famille $(u_n)_n$ est sommable si et seulement si la série de terme général u_n est absolument convergente.

¹⁰ C'est-à-dire une famille indicée par \mathbf{N} .

Lemme 1.45. Comme pour les séries absolument convergentes, l'ordre de sommation n'est pas important.

Théorème 1.46 (de sommation par paquets, ou de Fubini) : Soit I un ensemble dénombrable, et $(u_i)_{i \in I}$ une famille de réels indexés par I .

On suppose que $I = \bigcup_{n \in \mathbf{N}} I_n$, où les I_n sont des ensembles dénombrables ou finis, deux à deux disjoints.

Alors la famille $(u_i)_{i \in I}$ est sommable si et seulement si :

- $\forall n \in \mathbf{N}$, la famille $(u_i)_{i \in I_n}$ est sommable (et donc $\sum_{i \in I_n} |u_i|$ existe)

- $\sum_n \left(\sum_{i \in I_n} |u_i| \right)$ est convergente.

Dans ce cas, on a alors

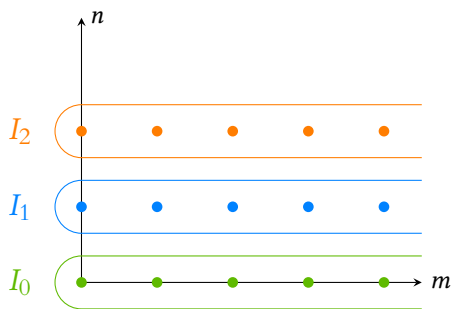
$$\sum_{i \in I} u_i = \sum_{n \in \mathbf{N}} \left(\sum_{i \in I_n} u_i \right).$$

Le programme officiel stipule que si ce théorème devait être utilisé, alors le sujet devrait en rappeler l'énoncé. Toutefois, rien ne garantit que ce soit bien le cas sur les sujets de parisiennes...

Remarque. La difficulté est souvent de trouver les bons I_n , c'est-à-dire de trouver comment découper I en «paquets» pour lesquels on sait étudier les sommes $\sum_{i \in I_n} u_i$.

Exemples 1.47

- Prenons $I = \mathbf{N} \times \mathbf{N}$, et pour $(m, n) \in \mathbf{N} \times \mathbf{N}$, $u_{m,n} = \frac{(-1)^m}{m!n!}$.



Alors I est dénombrable, et si on pose

$$I_n = \{(m, n), m \in \mathbf{N}\}$$

on a bien $I = \bigcup_{n \in \mathbf{N}} I_n$, avec les I_n deux à deux disjoints.

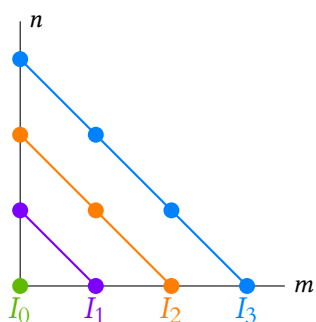
Pour $n \in \mathbf{N}$, $(u_i)_{i \in I_n}$ est sommable car $\sum_m u_{m,n} = \sum_m \frac{1}{m!n!} = \frac{1}{n!} \sum_m \frac{1}{m!}$ est une série absolument convergente.

On a alors $\sum_{m \in I_n} |u_{m,n}| = \frac{1}{n!} \sum_{m=0}^{+\infty} \frac{1}{m!} = \frac{e}{n!}$.

Mais alors $\sum_n \frac{e}{n!}$ est convergente, donc $(u_{m,n})_{(m,n) \in \mathbf{N} \times \mathbf{N}}$ est sommable, et

$$\sum_{(m,n) \in \mathbf{N} \times \mathbf{N}} u_{m,n} = \sum_{n=0}^{+\infty} \left(\sum_{m=0}^{+\infty} \frac{(-1)^m}{m!} \frac{1}{n!} \right) = \sum_{n=0}^{+\infty} \frac{e^{-1}}{n!} = e^{-1} \sum_{n=0}^{+\infty} \frac{1}{n!} = e^{-1} \times e = 1.$$

- Toujours avec $I = \mathbf{N} \times \mathbf{N}$, posons $u_{m,n} = \frac{1}{(m+n)!}$.



Posons cette fois, pour $k \in \mathbf{N}$

$$I_k = \{(m, n) \in \mathbf{N} \times \mathbf{N} : m+n = k\}$$

On a bien $I = \bigcup_{k \in \mathbf{N}} I_k$, avec les I_k deux à deux disjoints. Notons que cette fois, les I_k sont finis.

Pour $k \in \mathbf{N}$, la famille $(u_{m,n})_{(m,n) \in I_k}$ est sommable car elle est finie. De plus,

$$\sum_{(m,n) \in I_k} \frac{1}{(m+n)!} = \sum_{(m,n) \in I_k} \frac{1}{k!} = \frac{k+1}{k!}.$$

Ensuite, la série $\sum_k \frac{k+1}{k!} = \sum_k \left(\frac{1}{(k-1)!} + \frac{1}{k!} \right)$ est absolument convergente, donc la famille $(u_{m,n})_{(m,n) \in \mathbf{N} \times \mathbf{N}}$ est sommable, et

$$\sum_{(m,n) \in \mathbf{N} \times \mathbf{N}} \frac{1}{(m+n)!} = \sum_{k \in \mathbf{N}} \left(\sum_{(m,n) \in I_k} \frac{1}{(m+n)!} \right) = \sum_{k=0}^{+\infty} \left(\frac{1}{(k-1)!} + \frac{1}{k!} \right) = 2e.$$

Le $k+1$ provient du fait que $\text{Card}(I_k) = k+1$, ce dont on se convainc facilement sur le dessin.

Corollaire 1.48 – Si $I = \mathbf{N} \times \mathbf{N}$ et si $(u_{m,n})_{(m,n) \in \mathbf{N} \times \mathbf{N}}$ est sommable, alors

$$\sum_{(m,n) \in \mathbf{N} \times \mathbf{N}} u_{m,n} = \sum_{n=0}^{+\infty} \sum_{m=0}^{+\infty} u_{m,n} = \sum_{m=0}^{+\infty} \sum_{n=0}^{+\infty} u_{m,n}.$$

Démonstration. Pour la première somme, c'est le théorème de sommation par paquets, appliqué en notant que

$$\mathbf{N} \times \mathbf{N} = \bigcup_{n=0}^{+\infty} I_n, \text{ où } I_n = \{(m, n), m \in \mathbf{N}\}.$$

Autrement dit, dans chaque «paquet» se trouvent les termes possédant «le même n ». Et donc sommer sur le paquet I_n revient à faire varier m sans changer n :

$$\sum_{i \in I_n} u_i = \sum_{m=0}^{+\infty} u_{m,n}.$$

Notons que si l'on sait que la famille est sommable, toutes les séries en jeu sont automatiquement convergentes.

La seconde égalité se démontre de la même manière, en prenant des paquets pour lesquels m est constant et n varie, c'est-à-dire en notant que

$$\mathbf{N} \times \mathbf{N} = \bigcup_{m \in \mathbf{N}} J_m, \text{ où } J_m = \{(m, n), n \in \mathbf{N}\}.$$

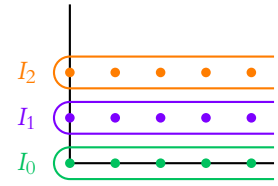


FIGURE 1.3— Les paquets I_n .

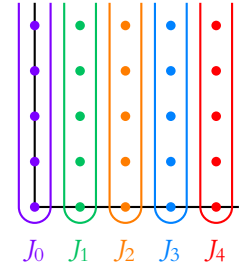


FIGURE 1.4— Les paquets J_m .

□

MÉTHODES DU CHAPITRE 1

✓ MÉTHODE 1.1 : Étudier la convergence d'une série numérique $\sum u_n$.

- En l'absence d'indication, commencer par s'assurer (au moins au brouillon) que $u_n \rightarrow 0$. Si ce n'est pas le cas, $\sum u_n$ diverge (grossièrement).
- Reconnaître une série usuelle (Riemann, géométrique, géométrique dérivée, exponentielle) ou une somme de telles séries, et utiliser alors les résultats du cours. \triangle On ne sait rien dire de la somme de deux séries divergentes, elle peut être convergente comme elle peut être divergente.
- Pour une série à termes positifs, majorer u_n par le terme général d'une série convergente (et alors $\sum u_n$ converge) ou au contraire minorer u_n par le terme général d'une série divergente (et alors $\sum u_n$ diverge.)
- Si u_n est de signe constant, chercher un équivalent, éventuellement à l'aide d'un développement limité.
- Si u_n n'est pas de signe constant, étudier $\sum |u_n|$ pour prouver une éventuelle convergence absolue. \triangle Si $\sum |u_n|$ diverge, on sait juste que $\sum u_n$ ne converge pas absolument, ça ne veut pas dire qu'elle diverge !
- Utiliser la règle $n^\alpha u_n$ (comparaison à une série de Riemann convergente.)
- Si rien d'autre ne semble marcher, essayer de calculer la suite des sommes partielles et étudier sa convergence.

Exercices de TD : 1.10, 1.11, 1.12, 1.14, 1.16, 1.17, 1.18

✓ MÉTHODE 1.2 : Calculer la somme d'une série $\sum u_n$.

- S'assurer que la série est bien convergente (en général ça a déjà été fait dans les questions précédentes.)
- $\sum u_n$ est-elle une série géométrique (ou dérivées) ou une série exponentielle ou somme de telles séries ? Dans ce cas, on sait conclure.
- Calculer les sommes partielles $\sum_{k=0}^N u_k$. En général, s'il faut procéder ainsi, c'est soit qu'on a affaire à une série télescopique (et alors la plupart des termes vont se simplifier), soit que l'énoncé nous aide par exemple à encadrer les sommes partielles. Puis faire tendre N vers $+\infty$.

Exercices de TD : 1.10, 1.11, 1.17, 1.18



MÉTHODE 1.3 : Étudier la convergence d'une série double $\sum_{(m,n)} u_{m,n}$.

On essaiera systématiquement d'utiliser le théorème de sommation par paquets, sans oublier les valeurs absolues si besoin.

Il s'agit alors de trouver les «bons paquets» pour lesquels les sommes se calculent facilement.

En général, ce sont ceux de l'exemple 1.47. On favorisera le premier type de paquets (sommer sur m , puis sur n (ou le contraire)), et si ce n'est pas possible, on choisira le second type (sommation à $m+n$ fixé). La présence de $(m+n)!$ dans $u_{m,n}$ doit faire penser à ce second type de paquets.

Exercices de TD : 1.19, 1.20, 1.21

EXERCICES DU CHAPITRE 1

EXERCICE 1.1 Soit $(u_n)_{n \geq 1}$ une suite telle que $u_1 > 0$ et pour tout $n \in \mathbf{N}^*$, $u_{n+1} = u_n \left(1 - \frac{1}{2n!}\right)$.

F

Montrer que pour tout $n \in \mathbf{N}^*$, $u_n > 0$, puis en déduire (u_n) est convergente.

EXERCICE 1.2 Déterminer les limites des suites suivantes :

PD

$$\begin{aligned}
 1. \ u_n &= \frac{2^n + (-1)^n}{3n + (-1)^{n+1}} & 2. \ u_n &= \frac{n \sin n}{1 + n^2} & 3. \ u_n &= \sqrt[n]{n} & 4. \ u_n &= \left(\frac{n-1}{n+1}\right)^{n+2} \\
 5. \ u_n &= n^2 \left(\cos \frac{1}{n} - \cos \frac{1}{n+1}\right) & 6. \ u_n &= \left(\cos \frac{1}{n}\right)^{n^2} & 7. \ u_n &= \frac{3n^2 + (-1)^n}{\sqrt{n^2 + 2} + \ln(n)}
 \end{aligned}$$

EXERCICE 1.3 Soit $k \in \mathbf{N}$ fixé. Déterminer un équivalent, lorsque $n \rightarrow +\infty$, de $\binom{n}{k}$.

PD

EXERCICE 1.4 Soit (u_n) la suite définie par $u_0 = 1$ et $\forall n \in \mathbf{N}$, $u_{n+1} = \ln(1 + u_n)$.

PD

- 1) Montrer que (u_n) est bien définie, et que pour tout $n \in \mathbf{N}$, $u_n > 0$.
- 2) Déterminer le sens de variation de la suite (u_n) .
- 3) Montrer que (u_n) converge vers un réel ℓ vérifiant $\ell = \ln(1 + \ell)$, puis déterminer la valeur de ℓ .

EXERCICE 1.5 Soit $(u_n)_{n \in \mathbf{N}^*}$ la suite définie par $u_1 > 0$ et pour tout $n \in \mathbf{N}^*$, $u_{n+1} = \sum_{k=1}^n \frac{u_k}{k}$.

PD

- 1) Montrer que pour tout $n \in \mathbf{N}^*$, $u_n \geq u_1$.
- 2) En déduire que pour tout $n \in \mathbf{N}^*$, $u_n \geq \left(\sum_{k=1}^{n-1} \frac{1}{k}\right) u_1$.
- 3) Déterminer la limite de (u_n) .

EXERCICE 1.6 Déterminer un équivalent simple des suites suivantes :

AD

$$1. \ u_n = \exp\left(\frac{1 - \sqrt{n}}{1 + n}\right) - 1 \qquad 2. \ v_n = \sqrt{n^2 + 1} - \sqrt[3]{n^3 - n} \qquad 3. \ w_n = \sin\left(\cos\left(\frac{\pi}{2} - \frac{n}{2^n}\right)\right).$$

EXERCICE 1.7 Soient (u_n) et (v_n) deux suites définies par $u_0 < v_0$, et pour tout $n \in \mathbf{N}$, $\begin{cases} u_{n+1} = \frac{u_n + v_n}{2} \\ v_{n+1} = \frac{u_n + 2v_n}{3} \end{cases}$

AD

- 1) Montrer que la suite (w_n) définie par $w_n = u_n - v_n$ est géométrique.
- 2) Montrer que les suites (u_n) et (v_n) convergent vers une même limite.

EXERCICE 1.8 Soit (u_n) une suite à termes positifs telle que $u_n = o_{n \rightarrow +\infty}(\sqrt{n})$, et soit $v_n = \left(1 + \frac{u_n}{n}\right)^n e^{-u_n}$.

AD

Déterminer un équivalent de $\ln(v_n)$, puis en déduire la limite de (v_n) .

EXERCICE 1.9 Déterminer si les affirmations suivantes sont vraies ou fausses :

AD

- 1) Si $u_n = (2n - 1)^3$, alors :
 - $u_n = o_{n \rightarrow +\infty}(n^3)$ $u_n \underset{n \rightarrow +\infty}{\sim} 8n^3$ $u_n = o_{n \rightarrow +\infty}(n^4)$ $u_n \underset{n \rightarrow +\infty}{\sim} n^3$ $u_n = o_{n \rightarrow +\infty}\left(\frac{n^4}{2}\right)$.
- 2) Si $u_n \xrightarrow[n \rightarrow +\infty]{} +\infty$, alors :
 - $\ln(u_n) = o(n)$ $\ln(u_n) = o(u_n)$ $\ln(n) = o(u_n)$
- 3) Soit (u_n) une suite réelle. Alors $u_{n+1} \underset{n \rightarrow +\infty}{\sim} u_n$.
- 4) Soient $(u_n), (v_n), (a_n), (b_n)$ quatre suites telles que $u_n = o_{n \rightarrow +\infty}(a_n)$ et $v_n = o_{n \rightarrow +\infty}(b_n)$.
 - $u_n + v_n = o_{n \rightarrow +\infty}(a_n + b_n)$ si $a_n \underset{n \rightarrow +\infty}{\sim} b_n$, alors $u_n + v_n = o_{n \rightarrow +\infty}(a_n)$.
 - si $a_n = o_{n \rightarrow +\infty}(b_n)$, alors $u_n + v_n = o_{n \rightarrow +\infty}(b_n)$.
- 5) Soit $u_n = \frac{2}{n} - \frac{3}{n\sqrt{n}} + \frac{1}{n^2}$, alors :

$$\square u_n \underset{n \rightarrow +\infty}{\sim} \frac{2}{n} \quad \square u_n = \underset{n \rightarrow +\infty}{o} \left(\frac{1}{n} \right) \quad \square \frac{1}{n} = \underset{n \rightarrow +\infty}{o} (u_n) \quad \square u_n = \frac{2}{n} + \underset{n \rightarrow +\infty}{o} \left(\frac{1}{n} \right)$$

$$\square u_n = \frac{2}{n} + \underset{n \rightarrow +\infty}{o} \left(\frac{1}{n\sqrt{n}} \right)$$

6) Si $u_n \underset{n \rightarrow +\infty}{\sim} v_n$, alors :

$$\square u_n + 1 \underset{n \rightarrow +\infty}{\sim} v_n + 1 \quad \square 2u_n \underset{n \rightarrow +\infty}{\sim} 2v_n \quad \square u_n - v_n \underset{n \rightarrow +\infty}{\sim} 0 \quad \square -u_n \underset{n \rightarrow +\infty}{\sim} -v_n$$

$$\square u_n v_n \underset{n \rightarrow +\infty}{\sim} u_n^2 \quad \square e^{u_n} \underset{n \rightarrow +\infty}{\sim} e^{v_n} \quad \square \ln(u_n) \underset{n \rightarrow +\infty}{\sim} \ln(v_n)$$

7) (★) Soit (u_n) une suite strictement positive. Laquelle des situations suivantes est équivalente au fait que $u_n \underset{n \rightarrow +\infty}{\sim} 2^n$?

$$\square \frac{u_{n+1}}{u_n} \xrightarrow{n \rightarrow +\infty} 2 \quad \square u_n - 2^n \xrightarrow{n \rightarrow +\infty} 0 \quad \text{(a) } \ln u_n \underset{n \rightarrow +\infty}{\sim} n \ln(2) \quad \text{(b) } \frac{u_n}{2^n} - 1 = \underset{n \rightarrow +\infty}{o} (1).$$

EXERCICE 1.10 Déterminer si les séries suivantes convergent, et le cas échéant, calculer leur somme :

$$1. \sum_{n \geq 0} \frac{3^{2n+1}}{n!} \quad 2. \sum_{n \geq 1} n 2^{n-1} \quad 3. \sum_{n \geq 1} \frac{(-1)^n}{n!} \quad 4. \sum_{n \geq 1} n \frac{3^n}{4^{n+1}} \quad 5. \sum_{n \geq 2} \frac{n(n-1)}{2^{n+1}}.$$

EXERCICE 1.11 Pour $n \in \mathbf{N}$, on pose $u_n = \frac{n}{(n+1)!}$.

1) Montrer que la série $\sum u_n$ converge.

2) Déterminer deux réels a et b tels que pour tout $n \in \mathbf{N}$, $u_n = \frac{a}{n!} + \frac{b}{(n+1)!}$, et en déduire la valeur de $\sum_{n=0}^{+\infty} u_n$.

3) Montrer que la série de terme général $(n^2 - 1)u_n$ converge et calculer sa somme.

EXERCICE 1.12 Dans chacun des cas suivants, déterminer la nature de la série $\sum u_n$:

$$1. u_n = \frac{n^3 + 2n}{n^4 + n^3 + 1} \quad 2. u_n = \frac{(-1)^n(n+2)}{n^3 + 1} \quad 3. u_n = \left(\frac{\ln n}{n} \right)^2 \quad 4. u_n = \left(1 - \frac{1}{n\sqrt{n}} \right)^{n\sqrt{n}}$$

$$5. u_n = \frac{1}{n^{1+\frac{1}{n}}} \quad 6. u_n = \ln(2 - e^{1/n})$$

EXERCICE 1.13 Comparaison série/intégrale

1) Montrer que la série $\sum_{k \geq 0} \frac{1}{1+k^2}$ converge.

2) En utilisant les variations de $x \mapsto \frac{1}{1+x^2}$, montrer que pour tout $k \geq 1$, $\frac{1}{1+(k+1)^2} \leq \int_k^{k+1} \frac{dt}{1+t^2} \leq \frac{1}{1+k^2}$.

3) Montrer que pour tout $n \in \mathbf{N}$, $\sum_{k=1}^n \frac{1}{1+k^2} \leq \text{Arctan}(n) \leq \sum_{k=0}^{n-1} \frac{1}{1+k^2}$.

4) En déduire que $\frac{\pi}{2} \leq \sum_{k=0}^{+\infty} \frac{1}{1+k^2} \leq \frac{\pi}{2} + 1$.

EXERCICE 1.14 Série harmonique et constante d'Euler.

Pour $n \geq 2$, on pose $u_n = \frac{1}{n} + \ln\left(\frac{n-1}{n}\right)$.

1) Montrer que la série de terme général u_n converge, puis déterminer les sommes partielles de cette série.

2) En déduire qu'il existe une constante γ telle que $\sum_{k=1}^n \frac{1}{k} = \ln(n) + \gamma + \underset{n \rightarrow +\infty}{o}(1)$, puis donner un équivalent de $\sum_{k=1}^n \frac{1}{k}$.

EXERCICE 1.15 Une série semi-convergente.

Pour tout $n \in \mathbf{N}^*$, on pose $u_n = \frac{(-1)^n}{\sqrt{n}}$ et $S_n = \sum_{k=1}^n u_k$.

1) La série de terme général u_n est-elle absolument convergente ?

2) Montrer que les suites $(S_{2n})_{n \in \mathbf{N}^*}$ et $(S_{2n+1})_{n \in \mathbf{N}^*}$ sont adjacentes.

3) En déduire que $\sum u_n$ converge.

EXERCICE 1.16 Soit $(u_n)_{n \in \mathbb{N}}$ une suite convergente. On pose alors pour tout $n \in \mathbb{N}$, $v_n = u_n - 2u_{n+1} + u_{n+2}$. Déterminer la nature de la série de terme général v_n , et en cas de convergence, calculer sa somme.

AD

EXERCICE 1.17 Pour $n \geq 1$, soit $u_n = \frac{(-1)^{n+1}n}{4n^2 - 1}$.

PD

- 1) Déterminer deux réels a et b tels que $\forall k \geq 1, \frac{k}{4k^2 - 1} = \frac{a}{2k + 1} + \frac{b}{2k - 1}$.
- 2) En déduire que la série de terme général u_n converge, et calculer sa somme.

EXERCICE 1.18 Pour tout entier $k \geq 1$, on pose $u_k = \ln\left(1 + \frac{2}{k(k+3)}\right)$.

AD

- 1) Montrer que la série de terme général u_k converge.
- 2) En factorisant $k(k+3)+2$, déterminer une suite $(v_k)_{k \geq 1}$ telle que pour tout $k \geq 1, u_k = v_{k+1} - v_k$.
- 3) En déduire la valeur de $\sum_{k=1}^{+\infty} u_k$.

EXERCICE 1.19 La famille $\left(\frac{(-1)^{n+m}2^m}{n \times m!}\right)_{\substack{n \geq 1 \\ m \geq 0}}$ est-elle sommable ?

PD

EXERCICE 1.20 Les familles suivantes sont-elles sommables ? Si oui, calculer leur somme.

AD

$$\left(\frac{1}{2^{m+n}m!}\right)_{\substack{m \geq 0 \\ n \geq 0}}, \left(\frac{(-1)^m n^m}{m!}\right)_{\substack{n \geq 0 \\ m \geq 0}} \text{ et } \left(\frac{n+m-1}{2^{m+n-2}}\right)_{\substack{n \geq 0 \\ m \geq 1}}.$$

EXERCICE 1.21 Pour quelles valeurs de α la famille $\left(\frac{1}{(m+n)^\alpha}\right)_{\substack{m \geq 1 \\ n \geq 1}}$ est-elle sommable ?

AD

EXERCICE 1.22 (QSP HEC 2008)

Représenter dans le plan l'ensemble des points de coordonnées (a, b) tels que $a > 0, b > 0$ et la série de terme général $u_n = \frac{a^n}{1+b^n}$ soit convergente.

D

EXERCICE 1.23 (QSP HEC 2014)

Pour tout entier $n \geq 1$, on pose $u_n = \ln(n) + a \ln(n+1) + b \ln(n+2)$.

D

- 1) Déterminer les réels a et b tels que la série de terme général u_n soit convergente.
- 2) Calculer alors la somme de cette série.

EXERCICE 1.24 (QSP ESCP 2006)

TD

Montrer que pour tout $n \in \mathbb{N}^*$, il existe un unique $y_n \in \mathbb{R}_+^*$ tel que $\ln(y_n) + y_n = \frac{1}{n}$.

Montrer que la suite $(y_n)_{n \in \mathbb{N}^*}$ converge et si on note ℓ sa limite, déterminer un équivalent de $y_n - \ell$.

EXERCICE 1.25 Une série dont le terme général est défini par une relation de récurrence (Oral HEC 2012)

TD

Soit $(u_n)_{n \in \mathbb{N}}$ la suite réelle définie par : $u_0 = 1$ et $\forall n \in \mathbb{N}, u_{n+1} = \frac{2n+2}{2n+5}u_n$.

- 1) Écrire une fonction Sci Lab ayant pour argument un entier n et renvoyant $\sum_{k=0}^n u_k$.
- 2) Soit $\alpha \in \mathbb{R}$. On pose pour tout $n \in \mathbb{N}^*, v_n = \frac{(n+1)^\alpha u_{n+1}}{n^\alpha u_n}$.
 - a) Rappeler le développement limité à l'ordre deux au voisinage de 0 de $x \mapsto \ln(1+x)$.
 Montrer que $\ln v_n = (\alpha+1) \ln\left(1 + \frac{1}{n}\right) - \ln\left(1 + \frac{5}{2n}\right)$.
 Pour quelle valeur α_0 du réel α la série de terme général $\ln v_n$ est-elle convergente ?
 - b) Expliciter $\sum_{k=1}^n \ln v_k$ sans signe \sum , et en déduire qu'il existe un réel strictement positif C tel que $u_n \underset{n \rightarrow +\infty}{\sim} \frac{C}{n^{\alpha_0}}$.
 Qu'en déduit-on pour la série $\sum u_n$?
 - c) Justifier l'existence d'un réel strictement positif D (indépendant de n) tel que $\forall n \in \mathbb{N}, \sum_{k=0}^n ku_k \leq D \times \sqrt{n}$.

- 3) a) Établir pour tout entier naturel n , la relation : $2 \sum_{k=1}^{n+1} k u_k + 3 \sum_{k=1}^{n+1} u_k = 2 \sum_{k=0}^n k u_k + 2 \sum_{k=0}^n u_k$.
- b) En déduire la valeur de $\sum_{n=0}^{+\infty} u_n$.

CORRECTION DES EXERCICES DU CHAPITRE 1

SOLUTION DE L'EXERCICE 1.1

Notons que pour tout $n \in \mathbf{N}^*$, on a $(2n)! \geq 2$, et donc $0 < 1 - \frac{1}{(2n)!} \leq 1$

Prouvons alors par récurrence que pour tout $n \in \mathbf{N}^*$, $u_n > 0$.

Par hypothèse, $u_1 > 0$. Supposons donc que $u_n > 0$. Alors

$$u_{n+1} = \underbrace{u_n}_{>0} \times \underbrace{\left(1 - \frac{1}{(2n)!}\right)}_{>0}.$$

Et donc par le principe de récurrence, pour tout $n \in \mathbf{N}^*$, $u_n > 0$.

Notons qu'on a alors, pour tout $n \in \mathbf{N}^*$, $\frac{u_{n+1}}{u_n} = 1 - \frac{1}{(2n)!} \leq 1$, et donc (u_n) est décroissante.

Étant minorée¹, d'après le théorème de la limite monotone, (u_n) converge.

¹ Par 0.

SOLUTION DE L'EXERCICE 1.2

1. Puisque $(-1)^n = o_{n \rightarrow +\infty}(2^n)$, il vient $2^n + (-1)^n = 2^n + o_{n \rightarrow +\infty}(2^n) \underset{n \rightarrow +\infty}{\sim} 2^n$.

De même, on a $3n + (-1)^{n+1} \underset{n \rightarrow +\infty}{\sim} 3n$ et donc $u_n \underset{n \rightarrow +\infty}{\sim} \frac{2^n}{3n} \rightarrow +\infty$.

2. Puisque $|\sin n| \leq 1$, il vient $0 \leq |u_n| \leq \frac{n}{1+n^2}$.

Mais $\frac{n}{1+n^2} \underset{n \rightarrow +\infty}{\sim} \frac{n}{n^2} \rightarrow 0$.

Ainsi, le théorème des gendarmes, on a donc $|u_n| \xrightarrow{n \rightarrow +\infty} 0$ et donc $\lim_{n \rightarrow +\infty} u_n = 0$.

3. On a $u_n = e^{\frac{1}{n} \ln(n)}$. Or, par croissance comparées, $\frac{\ln n}{n} \xrightarrow{n \rightarrow +\infty} 0$.

Et donc par continuité de l'exponentielle, $\lim_{n \rightarrow +\infty} e^{\frac{1}{n} \ln(n)} = e^0 = 1$.

4. Pour les mêmes raisons qu'à la question précédente, il faut passer par la forme exponentielle :

$$u_n = e^{(n+2) \ln\left(\frac{n-1}{n+1}\right)} = e^{(n+2) \ln\left(1 - \frac{2}{n+1}\right)}.$$

Or, $(n+2) \ln\left(1 - \frac{2}{n+1}\right) \underset{n \rightarrow +\infty}{\sim} -\frac{2(n+2)}{n+1} \underset{n \rightarrow +\infty}{\sim} -2 \xrightarrow{n \rightarrow +\infty} -2$.

Donc $u_n \xrightarrow{n \rightarrow +\infty} e^{-2}$.

Alternative : si on ne voit pas que $\frac{n-1}{n+1} = 1 - \frac{2}{n+1}$, il est possible de remarquer que

$$\frac{n-1}{n+1} \xrightarrow{n \rightarrow +\infty} 1.$$

Et donc

$$\ln\left(\frac{n-1}{n+1}\right) = \ln\left(1 + \underbrace{\left(\frac{n-1}{n+1} - 1\right)}_{\xrightarrow{n \rightarrow +\infty} 0}\right) \underset{n \rightarrow +\infty}{\sim} \frac{n-1}{n+1} - 1 = -\frac{2}{n+1}.$$

5. Un développement limité à l'ordre 2 du cosinus en 0, ce qui est légitime ici car $\frac{1}{n} \xrightarrow{n \rightarrow +\infty} 0$

nous indique que $\cos \frac{1}{n} = 1 - \frac{1}{2n^2} + o_{n \rightarrow +\infty}\left(\frac{1}{n^2}\right)$. Et de même,

$$\cos \frac{1}{n+1} = 1 - \frac{1}{2(n+1)^2} + o_{n \rightarrow +\infty}\left(\frac{1}{(n+1)^3}\right) = 1 - \frac{1}{2(n+1)^2} + o_{n \rightarrow +\infty}\left(\frac{1}{n^2}\right).$$

Par conséquent,

$$\cos \frac{1}{n} - \cos \frac{1}{n+1} = \frac{1}{2(n+1)^2} - \frac{1}{2n^2} + o_{n \rightarrow +\infty}\left(\frac{1}{n^2}\right) = \frac{-2n-1}{2n^2(n+1)^2} + o_{n \rightarrow +\infty}\left(\frac{1}{n^2}\right).$$

 **Danger !**

Il y a ici un piège très classique : 1^∞ est une forme indéterminée.

► Pour lever l'indétermination, on repassera toujours par la forme exponentielle : $a^b = e^{b \ln(a)}$.

Détails

Puisque $\frac{1}{(n+1)^2} \sim \frac{1}{n^2}$, on a

$$o\left(\frac{1}{(n+1)^2}\right) = o\left(\frac{1}{n^2}\right)$$

et donc nous pouvons regrouper les deux o en un seul : $o\left(\frac{1}{n^2}\right)$.

Après multiplication par n^2 , il vient

$$u_n = \frac{-2n^3 - n^2}{\underbrace{2n^2(n+1)^2}_{\rightarrow 0}} + \underbrace{o(1)}_{\rightarrow 0} \xrightarrow{n \rightarrow +\infty} 0.$$

6. En revenant à la forme exponentielle, on a $u_n = \exp\left(n^2 \ln\left(\cos\frac{1}{n}\right)\right)$.

Or, puisque $\cos\frac{1}{n} \xrightarrow{n \rightarrow +\infty} 1$, on a donc

$$\ln\left(\cos\frac{1}{n}\right) \underset{n \rightarrow +\infty}{\sim} \cos\frac{1}{n} - 1 \underset{n \rightarrow +\infty}{\sim} -\frac{1}{2n^2}.$$

Et donc $n^2 \ln\left(\cos\frac{1}{n}\right) \underset{n \rightarrow +\infty}{\sim} -\frac{1}{2}$.

Par continuité de l'exponentielle, on a donc $u_n \xrightarrow{n \rightarrow +\infty} e^{-\frac{1}{2}} = \frac{1}{\sqrt{e}}$.

7. On a $3n^2 + (-1)^n \underset{n \rightarrow +\infty}{\sim} 3n^2$.

D'autre part, $\sqrt{n^2 + 2} \underset{n \rightarrow +\infty}{\sim} n$ et donc $\ln(n) = \underset{n \rightarrow +\infty}{o}\left(\sqrt{n^2 + 2}\right)$.

On en déduit donc que $\sqrt{n^2 + 2} + \ln(n) \underset{n \rightarrow +\infty}{\sim} \sqrt{n^2 + 2} \underset{n \rightarrow +\infty}{\sim} n$.

Par conséquent, $u_n \underset{n \rightarrow +\infty}{\sim} \frac{3n^2}{n}$, donc $u_n \rightarrow +\infty$.

SOLUTION DE L'EXERCICE 1.3

Par définition, $\binom{n}{k} = \frac{n!}{k!(n-k)!} = \frac{n(n-1)\cdots(n-k+1)}{k!}$.

Le numérateur est alors un produit de k termes, avec k fixé. C'est un polynôme de degré k , dont le coefficient dominant² vaut 1.

Et donc³ $n(n-1)\cdots(n-k+1) \underset{n \rightarrow +\infty}{\sim} n^k$.

On en déduit que $\binom{n}{k} \underset{n \rightarrow +\infty}{\sim} \frac{n^k}{k!}$.

SOLUTION DE L'EXERCICE 1.4

1. Pour montrer que la suite (u_n) est bien définie, il s'agit de prouver que pour tout n , $1 + u_n > 0$, pour que u_{n+1} soit bien défini.

Or, si nous prouvons que $u_n > 0$, alors automatiquement on aura $1 + u_n > 0$.

Prouvons donc par récurrence sur $n \in \mathbf{N}$ que $u_n > 0$. C'est vrai pour $n = 0$.

Supposons donc que c'est vrai pour $n \in \mathbf{N}$. Alors $1 + u_n > 1$, et donc par stricte croissance du logarithme, $u_{n+1} = \ln(1 + u_n) > \ln(1) = 0$.

En vertu du principe de récurrence, pour tout $n \in \mathbf{N}$, $u_n > 0$.

2. On a $u_1 = \ln(2) < 1 = u_1$. Donc $u_2 < u_1$.

Montrons par récurrence sur $n \in \mathbf{N}$ que $u_{n+1} < u_n$.

Nous venons d'initialiser la récurrence, et si $u_{n+1} < u_n$, alors $1 + u_{n+1} < 1 + u_n$, et donc, par stricte croissance du logarithme, $u_{n+2} = \ln(1 + u_{n+1}) < \ln(1 + u_n) = u_{n+1}$.

Par le principe de récurrence, pour tout $n \in \mathbf{N}$, $u_{n+1} < u_n$, et donc (u_n) est (strictement) décroissante.

3. Puisque (u_n) est minorée par 0, par le théorème de la limite monotone, elle converge vers un réel ℓ . Et on a alors, par continuité du logarithme,

$$\ell = \lim_{n \rightarrow +\infty} u_n = \lim_{n \rightarrow +\infty} u_{n+1} = \lim_{n \rightarrow +\infty} \ln(1 + u_n) = \ln(1 + \ell).$$

Il nous reste donc à déterminer les réels $\ell \geq 0$ tels que $\ell = \ln(1 + \ell)$.

Posons pour cela $f(x) = \ln(1 + x) - x$. Alors f est dérivable sur \mathbf{R}_+ , et

$$\forall x \in \mathbf{R}_+, f'(x) = \frac{1}{1+x} - 1 = \frac{-x}{1+x} < 0.$$

Donc f' est strictement décroissante sur \mathbf{R}_+ .

Puisque $f(0) = 0$, on en déduit que pour tout $x > 0$, $f(x) < f(0) = 0$, soit encore

Détails

On a utilisé ici le fait que lorsque $u \rightarrow 1$,

$$\ln(u) = \ln(1 + (u-1)) \sim u-1.$$

Rédaction

Le fait que si $u_n \rightarrow \ell$, alors $f(u_n) \rightarrow f(\ell)$ n'est vrai que si f est continue en ℓ . Il ne faut alors pas oublier de le mentionner.

Détails

Par définition, on a $u_n \sim v_n$ si et seulement si

$$u_n = v_n + o(v_n).$$

² C'est-à-dire le terme en n^k .

³ Un polynôme est équivalent en $+\infty$ à son terme de plus haut degré.

Méthode

Montrer qu'une quantité est bien définie, c'est montrer que toutes les expressions qui la composent ont bien un sens. Autrement dit, qu'elle ne contient pas «opérations interdites» comme une division par 0, la racine d'un nombre négatif ou le logarithme d'un nombre négatif. Ici, le seul problème qui pourrait se produire serait si $1 + u_n \leq 0$.

Rédaction

Si $u_n \rightarrow \ell$, alors $f(u_n) \rightarrow f(\ell)$ est un résultat qui n'est valable que si f est continue en ℓ .

Il est donc important de préciser qu'ici le logarithme est continu.

$$\ln(1+x) < x.$$

$$\text{Et donc } \ell = \ln(1+\ell) \Leftrightarrow \ln(1+\ell) - \ell = 0 \Leftrightarrow f(\ell) = 0 \Leftrightarrow \ell = 0.$$

On en déduit donc que $\ell = \lim_{n \rightarrow +\infty} u_n = 0$.

SOLUTION DE L'EXERCICE 1.5

1. Puisque $u_{n+1} = u_1 + \frac{u_2}{2} + \dots + \frac{u_n}{n}$, il suffit de montrer que pour tout $n \in \mathbf{N}^*$, $u_n \geq 0$.

Montrons donc par récurrence forte sur $n \in \mathbf{N}^*$ que $u_n > u_1$.

Pour u_1 , c'est une hypothèse faite par l'énoncé.

Supposons donc que pour tout $k \leq n$, $u_k \geq 0$. Alors

$$u_{n+1} = u_1 + \frac{u_2}{2} + \dots + \frac{u_n}{n} \geq 0.$$

Et donc par le principe de récurrence forte, pour tout $n \in \mathbf{N}^*$, $u_n \geq 0$, de sorte que

$$u_n = u_1 + \underbrace{\frac{u_2}{2} + \dots + \frac{u_{n-1}}{n-1}}_{\geq 0} \geq u_1.$$

2. On a immédiatement, pour tout $n \geq 2$,

$$u_n = u_1 + \frac{u_2}{2} + \dots + \frac{u_{n-1}}{n-1} \geq u_1 + \frac{u_1}{2} + \dots + \frac{u_1}{n-1} = u_1 \left(\sum_{k=1}^{n-1} \frac{1}{k} \right).$$

3. Nous savons que la série harmonique, de terme général $\frac{1}{k}$, est divergente. Puisqu'il s'agit d'une série à termes positifs, la suite de ses sommes partielles est croissante.

N'étant pas convergente, elle tend donc vers $+\infty$: $\lim_{n \rightarrow +\infty} \sum_{k=1}^n \frac{1}{k} = +\infty$.

Et donc en particulier, $\lim_{n \rightarrow +\infty} \sum_{k=1}^{n-1} \frac{1}{k} = +\infty$.

Et puisque $u_1 > 0$, on en déduit que $\lim_{n \rightarrow +\infty} u_1 \left(\sum_{k=1}^{n-1} \frac{1}{k} \right) = +\infty$.

En raison de l'inégalité prouvée à la question précédente, on a donc $\lim_{n \rightarrow +\infty} u_n = +\infty$.

SOLUTION DE L'EXERCICE 1.6

1. Notons que $\frac{1-\sqrt{n}}{1+n} \underset{n \rightarrow +\infty}{\sim} \frac{-\sqrt{n}}{n} = -\frac{1}{\sqrt{n}} \xrightarrow{n \rightarrow +\infty} 0$.

Et donc il vient

$$u_n \underset{n \rightarrow +\infty}{\sim} \frac{1-\sqrt{n}}{1+n} \underset{n \rightarrow +\infty}{\sim} -\frac{1}{\sqrt{n}}.$$

2. On a $u_n = n \left(\sqrt{1 + \frac{1}{n^2}} - \sqrt[3]{1 - \frac{1}{n^2}} \right)$.

Or, nous savons que $(1+u)^\alpha = 1 + \alpha u + o(u)$. Donc

$$\begin{aligned} \sqrt{1 + \frac{1}{n^2}} - \sqrt[3]{1 - \frac{1}{n^2}} &= 1 + \frac{1}{2n^2} + o\left(\frac{1}{n^2}\right) - \left(1 - \frac{1}{3n^2} + o\left(\frac{1}{n^2}\right)\right) \\ &= \frac{5}{6n^2} + o\left(\frac{1}{n^2}\right) \underset{n \rightarrow +\infty}{\sim} \frac{5}{6n^2}. \end{aligned}$$

Et donc $u_n \underset{n \rightarrow +\infty}{\sim} \frac{5}{6n}$.

3. Commençons par noter que $\cos\left(\frac{\pi}{2} - \frac{n}{2^n}\right) = \sin\left(\frac{n}{2^n}\right)$.

Puisque $n = o(2^n)$, on a $\lim_{n \rightarrow +\infty} \frac{n}{2^n} = 0$, et donc $\lim_{n \rightarrow +\infty} \sin\left(\frac{n}{2^n}\right) = 0$.

On en déduit donc que

$$u_n = \sin\left(\sin\left(\frac{n}{2^n}\right)\right) \underset{n \rightarrow +\infty}{\sim} \sin\left(\frac{n}{2^n}\right) \underset{n \rightarrow +\infty}{\sim} \frac{n}{2^n}.$$

Réc. forte

Ici, pour prouver que $u_{n+1} \geq 0$, nous n'utilisons pas seulement $u_n \geq 0$, mais $u_1 \geq 0, u_2 \geq 0, \dots, u_n \geq 0$. Il ne s'agit donc pas d'une récurrence simple, mais d'une récurrence forte : on suppose que la propriété est vérifiée pour **tous les termes** d'ordre inférieur ou égal à n , et on la montre pour le terme d'ordre $n+1$.

C'est équivalent à une récurrence simple, où la propriété que l'on souhaite démontrer n'est pas $\mathcal{P}(n) : u_n \geq 0$, mais $\mathcal{P}(n) : \forall k \in \llbracket 1, n \rrbracket, u_k \geq 0$.

Détails

Il s'agit d'utiliser l'équivalent usuel

$$e^x - 1 \underset{x \rightarrow 0}{\sim} x.$$

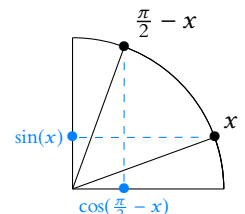
Notons que pour l'utiliser, il était indispensable de commencer par s'assurer que la quantité dans l'exponentielle tend bien vers 0.

Trigo

Comme (presque) toutes les formules de trigonométrie, la formule

$$\cos\left(\frac{\pi}{2} - x\right) = \sin(x)$$

peut se retrouver sur un dessin



SOLUTION DE L'EXERCICE 1.7

1. On a, pour tout $n \in \mathbf{N}$,

$$w_{n+1} = u_{n+1} - v_{n+1} = \frac{u_n + v_n}{2} - \frac{u_n + 2v_n}{3} = \frac{1}{6}u_n - \frac{1}{6}v_n = \frac{1}{6}w_n.$$

Donc (w_n) est géométrique de raison $\frac{1}{6}$.

2. Nous allons prouver que (u_n) et (v_n) sont deux suites adjacentes. Notons que par la question précédente, on a

$$u_n - v_n = w_n = \frac{1}{6^n} w_0 \xrightarrow{n \rightarrow +\infty} 0.$$

D'autre part,

$$u_{n+1} - u_n = \frac{u_n + v_n}{2} - u_n = \frac{v_n - u_n}{2} = -\frac{w_n}{2} = -\frac{1}{2 \times 6^n} (u_0 - v_0) > 0.$$

Et donc (u_n) est croissante.

De même, on a

$$v_{n+1} - v_n = \frac{u_n + 2v_n}{3} - v_n = \frac{1}{3}(u_n - v_n) = \frac{1}{3 \times 2^n} (u_0 - v_0) < 0.$$

Et donc (v_n) est décroissante.

Par conséquent, les deux suites (u_n) et (v_n) sont adjacentes, et donc convergent vers une même limite.

SOLUTION DE L'EXERCICE 1.8

On a $\ln(v_n) = n \ln\left(1 + \frac{u_n}{n}\right) - u_n = n\left(\ln\left(1 + \frac{u_n}{n}\right) - \frac{u_n}{n}\right)$.

Mais puisque $u_n = o(\sqrt{n})$, et $\sqrt{n} = o(n)$, alors $u_n = o(n)$ et donc $\lim_{n \rightarrow +\infty} \frac{u_n}{n} = 0$.

Un développement limité à l'ordre 2 nous donne alors

$$\ln(v_n) = n \left(\frac{u_n}{n} - \frac{u_n^2}{2n^2} + o_{n \rightarrow +\infty} \left(\frac{u_n^2}{n^2} \right) - \frac{u_n}{n} \right) = n \left(-\frac{u_n^2}{2n^2} + o_{n \rightarrow +\infty} \left(\frac{u_n^2}{n^2} \right) \right) \underset{n \rightarrow +\infty}{\sim} -n \frac{u_n^2}{2n^2} \underset{n \rightarrow +\infty}{\sim} -\frac{u_n^2}{2n}.$$

Mais puisque $u_n = o(\sqrt{n})$, alors $u_n^2 = o(n)$, et donc $\lim_{n \rightarrow +\infty} \frac{u_n^2}{n} = 0$.

On en déduit donc que $\lim_{n \rightarrow +\infty} \ln(v_n) = 0$.

Et alors, par continuité de l'exponentielle, $\lim_{n \rightarrow +\infty} v_n = e^{\ln(v_n)} = e^0 = 1$.

SOLUTION DE L'EXERCICE 1.9

1. Si $u_n = (2n+1)^2$, alors :

✗ $u_n = o_{n \rightarrow +\infty}(n^3)$: en effet, $\frac{(2n-1)^3}{n^3} \underset{n \rightarrow +\infty}{\sim} \frac{8n^3}{n^3} \xrightarrow{n \rightarrow +\infty} 8 \neq 0$.

✓ $u_n \underset{n \rightarrow +\infty}{\sim} 8n^3$: voir ci-dessus.

✓ $u_n = o(n^4)$: on a $u_n \sim 8n^3$ et donc $\frac{u_n}{n^4} \sim \frac{8n^3}{n^4} = \frac{8}{n} \xrightarrow{n \rightarrow +\infty} 0$.

✗ $u_n \sim n^3$: les constantes ont leur importance dans les équivalents.

En effet, on a $\frac{u_n}{n^3} \underset{n \rightarrow +\infty}{\sim} \frac{8n^3}{n^3} \xrightarrow{n \rightarrow +\infty} 8 \neq 1$.

✓ $u_n = o_{n \rightarrow +\infty}\left(\frac{n^4}{2}\right)$: les constantes ne servent à rien dans les o : être négligeable devant n^4 est pareil qu'être négligeable devant $\frac{n^4}{2}$.

2. Si $u_n \xrightarrow{n \rightarrow +\infty} +\infty$, alors :

✗ $\ln(u_n) = o(n)$: par exemple, si $u_n = e^n$, alors $\ln(u_n) = n$, qui n'est évidemment pas négligeable devant n .

Intuition

Notons que puisque les deux suites doivent converger vers une même limite, la suite croissante doit être celle qui possède le plus petit terme initial. Or ici $u_0 < v_0$, donc la suite croissante doit nécessairement être (u_n) et la suite décroissante doit être (v_n) .

Détails

Pour s'en convaincre, on peut revenir au quotient : on a

$$\lim_{n \rightarrow +\infty} \frac{u_n}{n^4/2} = 0$$

si et seulement si

$$\lim_{n \rightarrow +\infty} \frac{u_n}{n^4} = 0.$$

✓ $\ln(u_n) = o(u_n)$: revenons à la définition d'un o : $\lim_{n \rightarrow +\infty} \frac{\ln(u_n)}{u_n} = 0$ par composition de limites⁴.

✗ $\ln(n) = o(u_n)$: prendre par exemple $u_n = \ln(n)$.

3. ✗ Par exemple, prenons $u_n = e^n$. Alors $u_{n+1} = e^{n+1} = e \times u_n$, de sorte que $\frac{u_{n+1}}{u_n} \xrightarrow[n \rightarrow +\infty]{} e \neq 1$, et donc $u_n \neq u_{n+1}$.

Notons toutefois que pour une suite convergeant vers une limite $\ell \neq 0$, le résultat est vrai car $\frac{u_{n+1}}{u_n} \xrightarrow[n \rightarrow +\infty]{} \frac{\ell}{\ell} = 1$.

4. Soient $(u_n), (v_n), (a_n), (b_n)$ quatre suites réelles telles que : $u_n = o_{n \rightarrow +\infty}(a_n)$ et $v_n = o_{n \rightarrow +\infty}(b_n)$.

✗ $u_n + v_n = o_{n \rightarrow +\infty}(a_n + b_n)$: prenons $a_n = 1, b_n = -1$, et $u_n = \frac{1}{n}, v_n = \frac{2}{n}$.

Alors $u_n = o(a_n)$ et $v_n = o(b_n)$.

Mais $a_n + b_n = 0$ et $u_n + v_n \neq 0$, donc n'est pas négligeable devant la suite $a_n + b_n$.

✓ Si $a_n \sim b_n$, alors $u_n + v_n = o(a_n)$. En effet, on a alors $v_n = o(a_n)$ et donc $u_n + v_n = o(a_n)$. Pour s'en convaincre, on peut une fois de plus revenir aux quotients :

$$\lim_{n \rightarrow +\infty} \frac{u_n + v_n}{a_n} = \lim_{n \rightarrow +\infty} \frac{u_n}{a_n} + \frac{v_n}{a_n} = 0 + 0 = 0.$$

✓ Si $a_n = o(b_n)$, alors $u_n + v_n = o(b_n)$: si $u_n = o(a_n)$ et $a_n = o(b_n)$, alors $u_n = o(b_n)$. Et donc comme précédemment, $u_n + v_n = o(b_n)$.

5. Soit $u_n = \frac{2}{n} - \frac{3}{n\sqrt{n}} + \frac{1}{n^2}$. Alors :

✓ $u_n \underset{n \rightarrow +\infty}{\sim} \frac{2}{n}$.

6. Si $u_n \sim v_n$, alors :

✗ $u_n + 1 \sim v_n + 1$: en effet, si $u_n = -1 + \frac{1}{n}$, et $v_n = -1$, alors $u_n \sim v_n$, mais pourtant $v_n + 1 = 0$ alors que $u_n + 1 \neq 0$, et donc $u_n \neq v_n$.

✓ $2u_n \sim 2v_n$: on a bien le droit de multiplier les équivalents. Et en particulier de les multiplier par une constante.

✗ $u_n - v_n \sim 0$: prenons $u_n = \frac{1}{n} + 1$ et $v_n = 1$. Alors $u_n \sim v_n$, et $u_n - v_n = \frac{1}{n}$ n'est pas équivalent à 0.

✓ $-u_n \sim -v_n$: comme précédemment, on peut multiplier les équivalents.

✓ $u_n v_n \sim u_n^2$: on peut multiplier les équivalents.

Or, $u_n \sim v_n$ et $u_n \sim u_n$, donc par produit d'équivalents, $u_n v_n \sim u_n u_n = u_n^2$.

✗ $e^{u_n} \sim e^{v_n}$: prenons $u_n = n$ et $v_n = n + 1$. Alors $\frac{e^{u_n}}{e^{v_n}} = \frac{e^n}{e^{n+1}} = \frac{1}{e} \not\xrightarrow[n \rightarrow +\infty]{} 1$ et donc $e^{u_n} \neq e^{v_n}$.

7. ✗ $\frac{u_{n+1}}{u_n} \xrightarrow[n \rightarrow +\infty]{} 2$

Par exemple si $u_n = n2^n$, on a $\frac{u_{n+1}}{u_n} = 2 \frac{n+1}{n} \xrightarrow[n \rightarrow +\infty]{} 2$, et pourtant $\frac{u_n}{2^n} = n$, de sorte que $u_n \neq 2^n$.

✗ $u_n - 2^n \xrightarrow[n \rightarrow +\infty]{} 0$: la condition donnée est suffisante, puisque si $u_n - 2^n \xrightarrow[n \rightarrow +\infty]{} 0$, alors

$$\frac{u_n}{2^n} - 1 \xrightarrow[n \rightarrow +\infty]{} 0 \text{ et donc } \frac{u_n}{2^n} \xrightarrow[n \rightarrow +\infty]{} 1.$$

Toutefois, elle n'est pas nécessaire. Par exemple, si on pose $u_n = 2^n + n$, alors $u_n = 2^n + o_{n \rightarrow +\infty}(2^n)$, de sorte que $u_n \underset{n \rightarrow +\infty}{\sim} 2^n$.

Mais pourtant, $u_n - 2^n = n \not\xrightarrow{} 0$.

⁴ Il est bien connu que

$$\lim_{x \rightarrow +\infty} \frac{\ln(x)}{x} = 0.$$

En revanche

Si $\ell = 0$, le résultat est faux, comme le prouve $u_n = e^{-n}$.

$o(1)$

Rappelons que $u_n = o(1)$ si et seulement si $\lim_{n \rightarrow +\infty} u_n = 0$.

Équivalent à 0

Rappelons que seule la suite nulle est équivalent à 0.

Quest. subsidiaire

Montrer que $e^{u_n} \sim e^{v_n}$ si et seulement si $u_n - v_n \rightarrow 0$.

✗ $\ln u_n \sim n \ln(2)$: puisqu'on ne peut pas composer les équivalents, il n'est pas possible de passer à l'exponentielle pour en déduire que $u_n = e^u \sim e^{n \ln 2} = 2^n$.
Par exemple, si $u_n = 3 \times 2^n$, alors $\ln u_n = \ln(3) + n \ln(2) \underset{n \rightarrow +\infty}{\sim} n \ln(2)$, mais pourtant $u_n \not\sim 2^n$.

✓ $\frac{u_n}{2^n} - 1 = o_{n \rightarrow +\infty}(1)$

Ceci est équivalent à $\frac{u_n}{2^n} = 1 + o_{n \rightarrow +\infty}(1) \underset{n \rightarrow +\infty}{\sim} 1$.

Et donc, après multiplication par 2^n , à $u_n \underset{n \rightarrow +\infty}{\sim} 2^n$.

SOLUTION DE L'EXERCICE 1.10

1. On a $\frac{3^{2n+1}}{n!} = 3 \frac{9^n}{n!}$.

On reconnaît donc une série exponentielle de paramètre 9, qui converge, de sorte que

$$\sum_{n=0}^{+\infty} \frac{3^{2n+1}}{n!} = 3 \sum_{n=0}^{+\infty} \frac{9^n}{n!} = 3e^9.$$

2. Il s'agit d'une série géométrique dérivée, de raison $2 > 1$, et donc c'est une série divergente.

3. Nous reconnaissons une série exponentielle, donc convergente, de paramètre -1 . On a alors

$$\sum_{n=1}^{+\infty} \frac{(-1)^n}{n!} = \sum_{n=0}^{+\infty} \frac{(-1)^n}{n!} - 1 = e^{-1} - 1.$$

4. On a $n \frac{3^n}{4^{n+1}} = \frac{1}{4} n \left(\frac{3}{4}\right)^n = \frac{3}{16} n \left(\frac{3}{4}\right)^{n-1}$.

Nous reconnaissons une série géométrique dérivée de raison $\frac{3}{4} \in]-1, 1[$, et donc convergente. On a alors

$$\sum_{n=1}^{+\infty} n \frac{3^n}{4^{n+1}} = \frac{3}{16} \sum_{n=1}^{+\infty} n \left(\frac{3}{4}\right)^{n-1} = \frac{3}{16} \frac{1}{\left(1 - \frac{3}{4}\right)^2} = 3.$$

5. On a $\frac{n(n-1)}{2^{n+1}} = \frac{1}{8} n(n-1) \left(\frac{1}{2}\right)^{n-2}$.

Nous y voyons alors une série géométrique dérivée seconde, de raison $\frac{1}{2} \in]-1, 1[$ et donc convergente. On a alors

$$\sum_{n=2}^{+\infty} \frac{n(n-1)}{2^{n+1}} = \frac{1}{8} \sum_{n=2}^{+\infty} n(n-1) \left(\frac{1}{2}\right)^{n-2} = \frac{1}{8} \frac{2}{\left(1 - \frac{1}{2}\right)^3} = 2.$$

SOLUTION DE L'EXERCICE 1.11

1. On a $u_n = \frac{n}{n+1} \frac{1}{n!} \underset{n \rightarrow +\infty}{\sim} \frac{1}{n!}$.

Puisque la série de terme général $\frac{1}{n!}$ est une série⁵ convergente, par le critère de comparaison pour les séries à termes positifs, on en déduit que $\sum u_n$ converge.

2. Supposons que deux tels réels a et b existent. Alors pour tout $n \in \mathbf{N}$,

$$u_n = \frac{a}{n!} + \frac{b}{(n+1)!} = \frac{a(n+1) + b}{(n+1)!} = \frac{n}{(n+1)!}.$$

C'est le cas si et seulement si pour tout $n \in \mathbf{N}$, $n = a(n+1) + b = an + (a+b)$.

Puisque deux polynôme coïncident en une infinité de nombres⁶ si et seulement si ils sont

$$\text{égaux, l'égalité de départ est équivalente à } \begin{cases} a+b=0 \\ a=1 \end{cases} \Leftrightarrow \begin{cases} a=1 \\ b=-1 \end{cases}.$$

On a alors pour $N \in \mathbf{N}$,

$$\sum_{n=0}^N u_n = \sum_{n=0}^N \frac{1}{n!} - \sum_{n=0}^N \frac{1}{(n+1)!} = 1 - \frac{1}{(N+1)!} \xrightarrow{N \rightarrow +\infty} 1.$$

Attention !

Bien vérifier où commence la série : ici elle commence à $n = 1$, alors qu'une série exponentielle commence à $n = 0$, donc si on souhaite faire apparaître $\sum_{n=0}^{+\infty} \frac{(-1)^n}{n!}$, il faut ajouter le terme correspondant à $n = 0$.

⁵ Exponentielle.

⁶ Et il y a bien une infinité de nombres dans \mathbf{N} .

Et donc $\sum_{n=0}^{+\infty} u_n = 1$.

3. Puisque $n^2 - 1 = (n-1)(n+1)$, il vient donc, pour $n \geq 2$,

$$(n^2 - 1)u_n = \frac{(n-1)n(n+1)}{(n+1)!} = \frac{1}{(n-2)!}.$$

D'autre part, pour $n = 0$ et $n = 1$, on a $(n^2 - 1)u_n = 0$.

Et donc pour $N \in \mathbf{N}$, il vient

$$\sum_{n=0}^N (n^2 - 1)u_n = \sum_{n=2}^N \frac{1}{(n-2)!} = \sum_{k=0}^{N-2} \frac{1}{k!} \xrightarrow{N \rightarrow +\infty} \sum_{k=0}^{+\infty} \frac{1}{k!} = e.$$

Ceci prouve à la fois la convergence de la série de terme général $(n^2 - 1)u_n$ et le fait que

$$\sum_{n=0}^{+\infty} (n^2 - 1)u_n = e.$$

SOLUTION DE L'EXERCICE 1.12

1. On a $u_n \underset{n \rightarrow +\infty}{\sim} \frac{n^3}{n^4} = \frac{1}{n}$.

Or la série de terme général $\frac{1}{n}$ diverge, et puisque $u_n \geq 0$, d'après le critère des équivalents, $\sum_n u_n$ diverge également.

2. On a $|u_n| = \frac{n+2}{n^3+1} \underset{n \rightarrow +\infty}{\sim} \frac{1}{n^2}$.

Ainsi, par le critère des équivalents, qui s'applique évidemment car $|u_n| \geq 0$, $\sum |u_n|$ converge.

Par conséquent $\sum u_n$ est absolument convergente, et donc convergente.

3. Cherchons $\alpha > 1$ tel que $n^\alpha u_n \xrightarrow{n \rightarrow +\infty} 0$. On a

$$n^\alpha u_n = \frac{\ln^2 n}{n^{2-\alpha}}.$$

Par croissances comparées, cette quantité tend vers 0 si et seulement $2 - \alpha > 0 \Leftrightarrow \alpha < 2$.

Donc par exemple, pour $\alpha = \frac{3}{2}$, on a $n^{3/2} u_n \rightarrow 0$.

Et donc, d'après la règle $n^\alpha u_n$, $\sum u_n$ converge.

4. Commençons par déterminer la limite de u_n car il n'est pas évident que celle-ci soit nulle. On a

$$u_n = \left(1 - \frac{1}{n\sqrt{n}}\right)^{n\sqrt{n}} = e^{n\sqrt{n} \ln\left(1 - \frac{1}{n\sqrt{n}}\right)}.$$

Or, $\ln\left(1 - \frac{1}{n\sqrt{n}}\right) \underset{n \rightarrow +\infty}{\sim} -\frac{1}{n\sqrt{n}}$, et donc

$$n\sqrt{n} \ln\left(1 - \frac{1}{n\sqrt{n}}\right) \underset{n \rightarrow +\infty}{\sim} -n\sqrt{n} \frac{1}{n\sqrt{n}} = -1 \xrightarrow{n \rightarrow +\infty} -1.$$

Par continuité de l'exponentielle, on a donc $u_n \xrightarrow{n \rightarrow +\infty} e^{-1} \neq 0$.

Et donc la série de terme général u_n diverge grossièrement.

5. On a $nu_n = \frac{1}{n^{1/n}} = e^{-\frac{1}{n} \ln(n)}$.

Or, $\frac{1}{n} \ln n \xrightarrow{n \rightarrow +\infty} 0$, et donc $nu_n \xrightarrow{n \rightarrow +\infty} e^0 = 1$.

Par conséquent, $u_n \underset{n \rightarrow +\infty}{\sim} \frac{1}{n}$, et donc par critère de comparaison pour les séries à termes positifs, $\sum u_n$ diverge.

Méthode

Pour une série dont le terme général n'est pas de signe constant, on commencera toujours pas s'intéresser à la convergence absolue.

Rédaction

Si on ne veut pas faire appel à la règle $n^\alpha u_n$, il suffit de dire que $u_n = o\left(\frac{1}{n^{3/2}}\right)$.

Danger !

Il y a ici une grosse erreur à ne pas commettre, c'est de dire que $\sum u_n$ est une série de Riemann avec $\alpha = 1 + \frac{1}{n} > 1$, donc convergente. En effet, une série de Riemann est de la forme $\frac{1}{n^\alpha}$, où α est une constante, qui ne peut donc pas dépendre de n .

6. Lorsque $n \rightarrow +\infty$, $2 - e^{1/n} \rightarrow 1$, de sorte que

$$\ln(2 - e^{1/n}) \underset{n \rightarrow +\infty}{\sim} 2 - e^{1/n} - 1 \underset{n \rightarrow +\infty}{\sim} -\frac{1}{n}.$$

Et donc par critère de comparaison pour les séries de signe constant, la série de terme général u_n diverge.

SOLUTION DE L'EXERCICE 1.13

1. On a $\frac{1}{k^2} \underset{k \rightarrow +\infty}{\sim} \frac{1}{k^2}$.

Puisque la série⁷ de terme général $\frac{1}{k^2}$ converge, par critère de comparaison pour les séries à termes positifs, il en est de même de $\sum \frac{1}{1+k^2}$.

2. La fonction $x \mapsto \frac{1}{1+x^2}$ est décroissante.

En particulier, pour tout $t \in [k, k+1]$, on a

$$\frac{1}{1+(k+1)^2} \leq \frac{1}{1+t^2} \leq \frac{1}{1+k^2}.$$

Et donc par croissance de l'intégrale,

$$\frac{1}{1+(k+1)^2} = \int_k^{k+1} \frac{1}{1+(k+1)^2} dt \leq \int_k^{k+1} \frac{dt}{1+t^2} \leq \int_k^{k+1} \frac{1}{1+k^2} dt = \frac{1}{1+k^2}.$$

3. En sommant la relation établie précédemment pour k allant de 0 à $n-1$, on obtient

$$\sum_{k=0}^{n-1} \frac{1}{1+(k+1)^2} \leq \sum_{k=0}^{n-1} \int_k^{k+1} \frac{dt}{1+t^2} \leq \sum_{k=0}^{n-1} \frac{1}{1+k^2}.$$

Par la relation de Chasles, le terme du milieu est égal à

$$\int_0^n \frac{dt}{1+t^2} = [\text{Arctan}(t)]_0^n = \text{Arctan}(n) - \text{Arctan}(0) = \text{Arctan}(n).$$

D'autre part, après le changement d'indice $i = k+1$, la somme de gauche devient $\sum_{i=1}^n \frac{1}{1+i^2}$.

Et donc on a bien démontré que

$$\sum_{i=1}^n \frac{1}{1+i^2} \leq \text{Arctan}(n) \leq \sum_{k=0}^{n-1} \frac{1}{1+k^2}.$$

4. La seconde inégalité de la question précédente devient, en faisant tendre n vers $+\infty$,

$$\frac{\pi}{2} = \lim_{n \rightarrow +\infty} \text{Arctan}(n) \leq \lim_{n \rightarrow +\infty} \sum_{k=0}^{n-1} \frac{1}{1+k^2} = \sum_{k=0}^{+\infty} \frac{1}{1+k^2}.$$

Enfin, la première inégalité s'écrit encore

$$\sum_{k=0}^n \frac{1}{1+k^2} - 1 \leq \text{Arctan}(n) \Leftrightarrow \sum_{k=0}^n \frac{1}{1+k^2} \leq \text{Arctan}(n) + 1.$$

Là encore, en faisant tendre n vers $+\infty$ il vient :

$$\sum_{k=0}^{+\infty} \frac{1}{1+k^2} = \lim_{n \rightarrow +\infty} \sum_{k=0}^n \frac{1}{1+k^2} \leq \lim_{n \rightarrow +\infty} \text{Arctan}(n) + 1 = \frac{\pi}{2} + 1.$$

SOLUTION DE L'EXERCICE 1.14

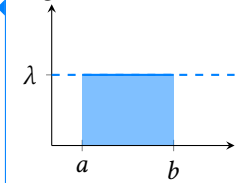
⁷ De Riemann.

Astuce

Il n'y a aucun calcul de primitive à faire pour l'intégrale d'une constante.

En effet, graphiquement

$\int_a^b \lambda dt$ est l'aire d'un rectangle : c'est $\lambda(b-a)$.



1. Notons que $\ln\left(\frac{n-1}{n}\right) = \ln\left(1 - \frac{1}{n}\right)$ et puisque $\frac{1}{n} \xrightarrow{n \rightarrow +\infty} 0$, il vient

$$\ln\left(1 - \frac{1}{n}\right) = -\frac{1}{n} - \frac{1}{2n^2} + o_{n \rightarrow +\infty}\left(\frac{1}{n^2}\right).$$

Et donc $u_n = \frac{1}{n} - \frac{1}{n} - \frac{1}{2n^2} + o_{n \rightarrow +\infty}\left(\frac{1}{n^2}\right) = -\frac{1}{2n^2} + o_{n \rightarrow +\infty}\left(\frac{1}{n^2}\right) \underset{n \rightarrow +\infty}{\sim} -\frac{1}{2n^2}$.

Mais la série de terme général $-\frac{1}{2n^2}$ est une série de Riemann convergente, et donc, d'après le critère des équivalents pour les série de signe constant⁸, $\sum u_n$ converge également.

⁸ $-\frac{1}{2n^2}$ est de signe constant, donc le critère s'applique.

2. Soit $n \in \mathbf{N}$. Alors

$$\begin{aligned} \sum_{k=2}^n u_k &= \sum_{k=2}^n \frac{1}{k} + \sum_{k=2}^n \ln\left(\frac{k-1}{k}\right) \\ &= \sum_{k=2}^n \frac{1}{k} + \sum_{k=2}^n \ln(k-1) - \sum_{k=2}^n \ln(k) \\ &= \sum_{k=2}^n \frac{1}{k} + \sum_{i=1}^{n-1} \ln(i) - \sum_{k=2}^n \ln(k) \\ &= \sum_{k=2}^n \frac{1}{k} + \underbrace{\ln(1)}_{=0} - \ln(n). \end{aligned}$$

Chgt d'indice
 $i = k - 1$.

3. On a $\sum_{k=1}^n \frac{1}{k} = \sum_{k=2}^n u_k + \frac{1}{1} + \ln(n)$.

Mais $\sum_{k=2}^n u_k + 1 \xrightarrow{n \rightarrow +\infty} \sum_{k=2}^{+\infty} u_k + 1$.

Notons donc $\gamma = 1 + \sum_{k=2}^{+\infty} u_k$, de sorte que $\sum_{k=2}^n u_k = \gamma + o_{n \rightarrow +\infty}(1)$.

Et donc $\sum_{k=1}^n \frac{1}{k} = \ln(n) + \gamma + o_{n \rightarrow +\infty}(1)$.

Enfin, puisque $\lim_{n \rightarrow +\infty} \ln(n) = +\infty$, $\gamma + o_{n \rightarrow +\infty}(1) = o_{n \rightarrow +\infty}(\ln n)$ et donc

$$\sum_{k=1}^n \frac{1}{k} \underset{n \rightarrow +\infty}{\sim} \ln n.$$

$o(1)$
Rappelons que $o(1)$ désigne toute suite de limite nulle, et donc si $v_n \rightarrow \ell$, on a $v_n - \ell \rightarrow 0 \Leftrightarrow v_n - \ell = o(1)$ et donc $v_n = \ell + o(1)$.

SOLUTION DE L'EXERCICE 1.15

1. On a $|u_n| = \frac{1}{\sqrt{n}}$, qui est le terme général d'une série de Riemann divergente. Donc $\sum |u_n|$ ne converge pas, et donc $\sum u_n$ n'est pas absolument convergente.
2. On a

$$S_{2n} - S_{2n+2} = \sum_{k=1}^{2n} \frac{(-1)^k}{\sqrt{k}} - \sum_{k=1}^{2n+2} \frac{(-1)^k}{\sqrt{k}} = \frac{1}{\sqrt{2k+1}} - \frac{1}{\sqrt{2k+2}} \geq 0.$$

Et donc $(S_{2n})_{n \in \mathbf{N}^*}$ est croissante.

De même, on a

$$S_{2n+1} - S_{2n+3} = \frac{1}{\sqrt{2n+3}} - \frac{1}{\sqrt{2n+2}} \leq 0.$$

Et donc (S_{2n+1}) est décroissante.

Enfin, $S_{2n} - S_{2n+1} = \sum_{k=1}^{2n} \frac{(-1)^k}{\sqrt{k}} - \sum_{k=1}^{2n+1} \frac{(-1)^k}{\sqrt{k}} = \frac{1}{\sqrt{2n+1}} \xrightarrow{n \rightarrow +\infty} 0$.

Par conséquent, les suites (S_{2n}) et (S_{2n+1}) sont adjacentes, et donc tendent vers une même limite ℓ .

⚠ Attention !
La suite (S_{2n}) ne contient que les termes pairs de (S_n) , et donc le terme suivant S_{2n} est $S_{2(n+1)} = S_{2n+2}$, et non S_{2n+1} , qui ne fait pas partie de la suite (S_{2n}) .

3. Puisque les suites des termes d'ordre pair⁹ et d'ordre impair de (S_n) convergent vers la même limite ℓ , $(S_n)_{n \in \mathbf{N}^*}$ converge également vers ℓ . Et donc $\sum u_n$ converge.

SOLUTION DE L'EXERCICE 1.16

Commençons par noter que si $\ell = \lim_{n \rightarrow +\infty} u_n$, alors $v_n \xrightarrow{n \rightarrow +\infty} \ell - 2\ell + \ell = 0$, donc la série $\sum v_n$ ne diverge pas grossièrement. Pour $N \in \mathbf{N}$, on a

$$\begin{aligned} \sum_{n=0}^N v_n &= \sum_{n=0}^N (u_n - 2u_{n+1} + u_{n+2}) \\ &= \sum_{n=0}^N u_n - 2 \sum_{n=0}^N u_{n+1} + \sum_{n=0}^N u_{n+2} \\ &= \sum_{n=0}^N u_n - 2 \sum_{k=1}^{N+1} u_k + \sum_{i=2}^{N+2} u_i \\ &= \sum_{n=0}^N u_n - \sum_{k=1}^{N+1} u_k + \sum_{i=2}^{N+2} u_i - \sum_{k=1}^{N+1} u_k \\ &= u_0 - u_{N+1} + u_{N+2} - u_1 \\ &\xrightarrow{N \rightarrow +\infty} u_0 - \ell + \ell - u_1 = u_0 - u_1. \end{aligned}$$

Et donc la série de terme général v_n converge, et $\sum_{n=0}^{+\infty} v_n = u_0 - u_1$.

SOLUTION DE L'EXERCICE 1.17

1. On a

$$\frac{a}{2k+1} + \frac{b}{2k-1} = \frac{a(2k-1) + b(2k+1)}{4k^2-1} = \frac{2(a+b)k + b-a}{4k^2-1}.$$

On souhaite donc avoir $\forall k \in \mathbf{N}$, $2(a+b)k + b-a = k$.

Mais deux polynômes coïncident en une infinité de valeurs si et seulement si ils sont égaux, et donc $a+b = \frac{1}{2}$ et $b-a = 0$.

La seule solution est donc $a = b = \frac{1}{4}$.

2. D'après la question précédente, on a

$$\frac{(-1)^{k+1}k}{4k^2-1} = \frac{(-1)^{k+1}}{4} \left(\frac{1}{2k+1} + \frac{1}{2k-1} \right).$$

Enfin, notons que $2(k+1) - 1 = 2k+1$, et donc nous avons affaire à une série télescopique :

$$\sum_{k=1}^N \frac{(-1)^{k+1}k}{4k^2-1} = \frac{1}{4} \left(\frac{1}{1} + \frac{1}{3} - \frac{1}{3} - \frac{1}{5} + \frac{1}{5} + \frac{1}{7} + \dots + \frac{(-1)^{N+1}}{2N-1} + \frac{(-1)^{N+1}}{2N+1} \right) = \frac{1}{4} \left(1 + \frac{(-1)^{N+1}}{2N+1} \right).$$

Puisque $\frac{(-1)^{N+1}}{2N+1} \xrightarrow{N \rightarrow \infty} 0$ (produit d'une suite de limite nulle par une suite bornée), on en déduit que

$$\sum_{n=1}^N \frac{(-1)^{n+1}n}{4n^2-1} \xrightarrow{N \rightarrow \infty} \frac{1}{4}.$$

Et donc ceci signifie que la série de terme général u_n converge et

$$\sum_{n=1}^{+\infty} \frac{(-1)^{n+1}n}{4n^2-1} = \frac{1}{4}.$$

SOLUTION DE L'EXERCICE 1.18

Rappel

Par **définition**, une série converge si et seulement si la suite de ses sommes partielles est une suite convergente.

Chgts d'indices

On a posé $k = n+1$ et $i = n+2$.

Rédaction

Pour les séries télescopiques, on procédera toujours de la même manière :

- 1) On commence par calculer les sommes partielles de la série afin de faire apparaître les simplifications (soit à l'aide d'un changement d'indice, soit en écrivant la somme avec des pointillés et en barrant les termes qui se simplifient).
- 2) Seulement après, on passe à la limite.

1. Lorsque $k \rightarrow +\infty$, on a $\frac{2}{k(k+3)} \rightarrow 0$, de sorte que $u_k \underset{k \rightarrow +\infty}{\sim} \frac{2}{k(k+3)} \underset{k \rightarrow +\infty}{\sim} \frac{2}{k^2}$.
 Mais la série de terme général $\frac{2}{k^2}$ est une série¹⁰ convergente, et donc, d'après le critère des équivalents pour les séries à termes positifs, la série $\sum_k u_k$ converge.

¹⁰ De Riemann.

2. Pour $k \geq 1$, on a

$$u_k = \ln\left(\frac{k(k+3)+2}{k(k+3)}\right) = \ln\left(\frac{k^2+3k+2}{k(k+3)}\right).$$

Mais k^2+3k+2 est un polynôme de degré 2 en k , de discriminant $\Delta = 9 - 4 \times 2 = 1$, et donc de racines $k_1 = \frac{-3+1}{2} = -1$ et $k_2 = -2$. Ainsi, $k^2+3k+2 = (k+1)(k+2)$.

Et donc $u_k = \ln\left(\frac{(k+1)(k+2)}{k(k+3)}\right) = \ln\left(\frac{k+1}{k+3}\right) - \ln\left(\frac{k}{k+2}\right)$.

Si on pose $v_k = \ln\left(\frac{k}{k+2}\right)$, alors on a bien $u_k = v_{k+1} - v_k$.

3. Pour $N \geq 1$, on a donc

$$\begin{aligned} \sum_{k=1}^N u_k &= \sum_{k=1}^N (v_{k+1} - v_k) \\ &= \sum_{k=1}^N v_{k+1} - \sum_{k=1}^N v_k \\ &= \sum_{i=2}^{N+1} v_i - \sum_{k=1}^N v_k \\ &= v_{N+1} - v_1 = \ln\left(\frac{N+1}{N+3}\right) + \ln(3) \\ &\xrightarrow{N \rightarrow +\infty} \ln(3). \end{aligned}$$

On en déduit donc que $\sum_{k=1}^{+\infty} u_k = \ln(3)$.

SOLUTION DE L'EXERCICE 1.19

On a $\sum_n \frac{(-1)^{n+m} 2^m}{n \times m!}$ qui n'est pas absolument convergente, car $\sum_n \frac{(-1)^n}{n}$ n'est pas absolument convergente.

Or, si la famille était sommable, par le théorème de sommation par paquets, $\frac{(-1)^m}{m!} \sum_n \frac{(-1)^n}{n}$

devrait converger absolument.

On en déduit que la famille n'est pas sommable.

SOLUTION DE L'EXERCICE 1.20

1. Puisque le seul outil à notre disposition pour étudier la sommabilité d'une famille est le théorème de sommation par paquets, essayons de l'utiliser avec des «paquets» simples. Les plus simples d'entre eux sont les

$$I_m = \{(n, m), n \in \mathbf{N}^*\}.$$

À m fixé, la série $\sum_{n \geq 0} \frac{1}{2^{m+n} m!} = \frac{1}{2^m m!} \sum_n \frac{1}{2^n}$ est une série géométrique convergente.

Puisqu'il s'agit d'une série à termes positifs, elle est même absolument convergente et

$$\sum_{n=0}^{+\infty} \frac{1}{2^{m+n} m!} = \frac{1}{2^m m!} \sum_{n=0}^{+\infty} \frac{1}{2^n} = \frac{2}{2^m m!}.$$

Mais alors $\sum_{m \geq 0} \frac{2}{2^m m!}$ est convergente¹¹, donc la famille est sommable et

$$\sum_{m, n \geq 0} \frac{1}{2^{m+n} m!} = \sum_{m=0}^{+\infty} \sum_{n=0}^{+\infty} \frac{1}{2^{m+n} m!} = \sum_{m=0}^{+\infty} \frac{2}{2^m m!} = 2e^{1/2}.$$

Rédaction

Si on ne souhaite pas s'embêter avec des Σ , développer les sommes avec des pointillés et montrer qu'on a bien reconnu les termes qui se simplifient devrait être tout aussi convaincant (bien qu'un peu moins rigoureux).
 Par exemple :

$$\cancel{v_2} + \cancel{v_3} + \dots + \cancel{v_N} + v_{N+1} - (v_1 + \cancel{v_2} + \dots + \cancel{v_N}) = v_{N+1} - v_1$$

Chgt d'indice

$i = k + 1$

Détails

$\frac{N+1}{N+3} \underset{N \rightarrow +\infty}{\sim} \frac{N}{N} = 1$, donc
 $\ln\left(\frac{N+1}{N+3}\right) \xrightarrow{N \rightarrow +\infty} \ln(1) = 0$.

Hypothèses

Pourtant elle est convergente ! Mais le théorème de sommation par paquets nous demande qu'elle soit absolument convergente.

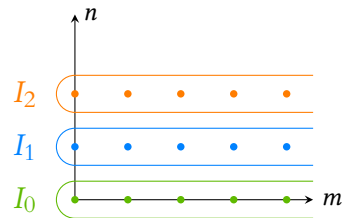


FIGURE 1.1– Les paquets I_m .

¹¹ On a reconnu une série exponentielle.

2. Pour n fixé, on a $\sum_m \frac{(-1)^m n^m}{m!}$ qui est absolument convergente, et

$$\sum_{m=0}^{+\infty} \frac{(-1)^m n^m}{m!} = e^{-n} \text{ et } \sum_{m=0}^{+\infty} \left| \frac{(-1)^m n^m}{m!} \right| = e^n$$

Mais $\sum_n e^n$ n'est pas absolument convergente (il s'agit d'une série qui diverge grossièrement), donc la famille n'est pas sommable.

3. Puisque l'expression fait clairement apparaître des $n + m$, sommons sur des paquets du type $I_k = \{(n, m) \in \mathbf{N}^* \times \mathbf{N} : n + m = k\}$.

Alors il s'agit de paquets finis, et I_k est de cardinal k , car $m \geq 1$ et $n \geq 0$.
On a alors

$$\sum_{(n,m) \in I_k} \frac{n+m-1}{2^{m+n-2}} = \sum_{(n,m) \in I_k} \frac{k-1}{2^{k-2}} = \frac{k(k-1)}{2^{k-2}}$$

En faisant varier k , nous reconnaissons là le terme général d'une série géométrique dérivée de raison $\frac{1}{2}$, donc absolument convergente. De plus, nous savons calculer sa somme, qui vaut

$$\sum_{k=0}^{+\infty} \frac{k(k-1)}{2^{k-2}} = \frac{2}{(1-\frac{1}{2})^3} = 16.$$

Donc la famille est sommable, et sa somme vaut 16.

SOLUTION DE L'EXERCICE 1.21

Sommons sur des paquets de la forme $I_k = \{(m, n) \in \mathbf{N}^* \times \mathbf{N}^* : m + n = k\}$, avec $k \geq 2$. Alors ces paquets sont finis, et de cardinal $k - 1$.

$$\sum_{(m,n) \in I_k} \frac{1}{(m+n)^\alpha} = \sum_{m+n=k} \frac{1}{k^\alpha} = \frac{k-1}{k^\alpha}.$$

Et donc la famille est sommable si et seulement si la série de terme général $\frac{k-1}{k^\alpha}$ est absolument convergente.

Mais $\frac{k-1}{k^\alpha} \underset{k \rightarrow +\infty}{\sim} \frac{k}{k^\alpha} = \frac{1}{k^{\alpha-1}}$, qui est le terme général d'une série convergente si et seulement si $\alpha - 1 > 1 \Leftrightarrow \alpha > 2$.

Et donc la famille est sommable si et seulement si $\alpha > 2$.

SOLUTION DE L'EXERCICE 1.22

- Si $b > 1$, alors au voisinage de $+\infty$, on a $u_n \underset{n \rightarrow +\infty}{\sim} \frac{a^n}{b^n} = \left(\frac{a}{b}\right)^n$.

Et donc, par application du critère des équivalents pour les séries de terme général positif, $\sum u_n$ converge si et seulement si $\frac{a}{b} < 1 \Leftrightarrow a < b$.

- Si $b = 1$, alors $u_n = \frac{a^n}{2}$, qui converge si et seulement si $a < 1$.

- Enfin, si $b < 1$, alors $1 + b^n \underset{n \rightarrow +\infty}{\sim} 1$, de sorte que $u_n \underset{n \rightarrow +\infty}{\sim} a^n$ qui converge donc également si et seulement si $a < 1$.

SOLUTION DE L'EXERCICE 1.23

1. On a

$$\begin{aligned} u_n &= \ln(n) + a \left(\ln(n) + \ln\left(1 + \frac{1}{n}\right) \right) + b \left(\ln(n) + \ln\left(1 + \frac{2}{n}\right) \right) \\ &= (1 + a + b) \ln(n) + a \left(\frac{1}{n} - \frac{1}{2n^2} + \underset{n \rightarrow +\infty}{o}\left(\frac{1}{n^2}\right) \right) + b \left(\frac{2}{n} - \frac{2}{n^2} + \underset{n \rightarrow +\infty}{o}\left(\frac{1}{n^2}\right) \right) \\ &= (1 + a + b) \ln(n) + \frac{a + 2b}{n} - \frac{a + 4b}{2n^2} + \underset{n \rightarrow +\infty}{o}\left(\frac{1}{n^2}\right). \end{aligned}$$

Danger !

Si l'on oublie les valeurs absolues, on a $\sum_n e^{-n}$ qui est absolument convergente, et l'on conclurait, de manière erronée, que la famille est sommable.

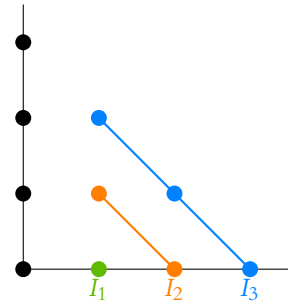


FIGURE 1.2– Les paquets I_k .

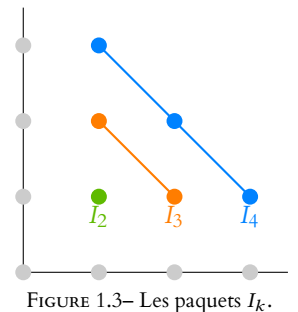


FIGURE 1.3– Les paquets I_k .

Détails

$b^n \rightarrow +\infty$, donc $1 + b^n \sim b^n$.

Si $1 + a + b \neq 0$, alors $u_n \underset{n \rightarrow +\infty}{\sim} (1 + a + b) \ln(n)$ et donc $\sum u_n$ diverge.

Une condition nécessaire pour que $\sum u_n$ converge est donc déjà $a + b = -1$.

Si c'est le cas et que $a + 2b \neq 0$, alors $u_n \underset{n \rightarrow +\infty}{\sim} \frac{a + 2b}{n}$ et donc¹² $\sum u_n$ diverge.

Il faut donc avoir de plus $a + 2b = 0$.

$$\text{Or, } \begin{cases} a + b = -1 \\ a + 2b = 0 \end{cases} \Leftrightarrow \begin{cases} a = -2 \\ b = 1 \end{cases}$$

Et alors, on a $u_n \underset{n \rightarrow +\infty}{\sim} \frac{-1}{n^2}$, de sorte¹³ que $\sum u_n$ converge.

Donc $\sum u_n$ converge si et seulement si $a = -2$ et $b = 1$.

2. On a alors, pour $N \geq 2$,

$$\begin{aligned} \sum_{k=1}^N u_k &= \sum_{k=1}^N \ln(k) - 2 \sum_{k=1}^N \ln(k+1) + \sum_{k=1}^N \ln(k+2) \\ &= \sum_{k=1}^N \ln(k) - 2 \sum_{k=2}^{N+1} \ln(k) + \sum_{k=3}^{N+2} \ln(k) \\ &= -\ln(N+1) - \ln(2) + \ln(N+2) = \ln\left(\frac{N+2}{N+1}\right) - \ln(2) \\ &\xrightarrow{N \rightarrow +\infty} -\ln(2). \end{aligned}$$

Et donc $\sum_{k=1}^{+\infty} u_k = -\ln(2)$.

SOLUTION DE L'EXERCICE 1.24

La fonction $f : x \mapsto \ln(x) + x$ est continue et strictement croissante sur \mathbf{R}_+^* .

Comme on a de plus $\lim_{x \rightarrow 0^+} f(x) = -\infty$ et $\lim_{x \rightarrow +\infty} f(x) = +\infty$, par le théorème de la bijection,

pour tout $n \in \mathbf{N}^*$, il existe un unique $y_n > 0$ tel que $f(y_n) = \frac{1}{n} \Leftrightarrow \ln(y_n) + y_n = \frac{1}{n}$.

Notons que puisque $\frac{1}{n} \leq 1 = f(1)$, nous pouvons affirmer que $y_n \in]0, 1]$.

D'autre part, on a $f(y_{n+1}) = \frac{1}{n+1} < \frac{1}{n} = f(y_n)$, donc par croissance de f , cela signifie que $y_{n+1} < y_n$: la suite $(y_n)_{n \in \mathbf{N}^*}$ est décroissante.

Étant minorée par 0, par le théorème de la limite monotone, elle converge vers un réel $\ell \in [0, 1]$.

En passant à la limite dans la relation $\ln(y_n) + y_n = \frac{1}{n}$, il vient, par continuité du logarithme

$$: \ln(\ell) + \ell = 0 \Leftrightarrow \ell = -\ln \ell.$$

On en déduit que

$$\begin{aligned} y_n - \ell &= -\ln(y_n) + \frac{1}{n} - \ell \\ &= \frac{1}{n} - \ln(y_n) + \ln(\ell) \\ &= \frac{1}{n} - \ln\left(\frac{y_n}{\ell}\right) \\ &= \frac{1}{n} - \ln\left(1 + \left(\frac{y_n}{\ell} - 1\right)\right) \\ &= \frac{1}{n} - \left(\frac{y_n}{\ell} - 1\right) + \underset{n \rightarrow +\infty}{o}\left(\frac{y_n}{\ell} - 1\right) \\ &= \frac{1}{n} - \frac{1}{\ell}(y_n - \ell) + \underset{n \rightarrow +\infty}{o}(y_n - \ell). \end{aligned}$$

Et donc en passant tous les termes en $y_n - \ell$ du même côté du signe égal, il vient

$$\frac{1}{n} = \left(1 + \frac{1}{\ell}\right)(y_n - \ell) + \underset{n \rightarrow +\infty}{o}(y_n - \ell) \Leftrightarrow \frac{1}{n} \underset{n \rightarrow +\infty}{\sim} \frac{\ell + 1}{\ell}(y_n - \ell) \Leftrightarrow y_n - \ell \underset{n \rightarrow +\infty}{\sim} \frac{\ell}{\ell + 1} \frac{1}{n}.$$

¹² Par comparaison à une série de Riemann.

¹³ Toujours par comparaison à une série de Riemann.

Remarque

Notons que ℓ ne peut pas être nul, car on aurait alors $\ln(y_n) + y_n \rightarrow -\infty$, alors que $\frac{1}{n} \rightarrow 0$.

Détails

Puisque $\frac{y_n}{\ell} - 1 \rightarrow 0$, on peut utiliser le développement limité à l'ordre 1 en 0 de $\ln(1+x)$.

o

Rappelons que les constantes ne sont pas utiles dans les o :

$$u_n = o(v_n) \Leftrightarrow u_n = o(\lambda v_n).$$

Ici, on a multiplié par ℓ .

SOLUTION DE L'EXERCICE 1.25

1. Le programme suivant convient :

```

1  fonction S = somme(n)
2      u = 1 ;
3      S = u ;
4      for i = 0 : n-1
5          u = (2*i+2)/(2*i+5)*u
6          S = S+u
7      end
8  endfunction

```

Notons que la boucle va de 0 à $n-1$, alors qu'il aurait pu sembler plus naturel d'aller de 1 à n .

La principale raison en est que nous disposons d'une formule donnant u_{n+1} en fonction de u_n . Or les u_k , pour k allant de 1 à n sont les u_{i+1} , pour i allant de 0 à $n-1$.

Si on veut tout de même utiliser une boucle allant de 1 à n , on remarquera que $u_n = \frac{2n}{2n+3}u_{n-1}$, et donc on pourra remplacer la ligne 5 par

$$5 \quad S = (2*i)/(2*i+3)*u$$

2.a. On a $\ln(1+x) = x - \frac{x^2}{2} + o_{x \rightarrow 0}(x^2)$.

De plus, pour tout $n \in \mathbf{N}^*$,

$$\begin{aligned} \ln v_n &= \ln \left(\left(\frac{n+1}{n} \right)^\alpha \right) + \ln \left(\frac{u_{n+1}}{u_n} \right) \\ &= \alpha \ln \left(\frac{n+1}{n} \right) + \ln \left(\frac{2n+2}{2n+5} \right) \\ &= \alpha \ln \left(\frac{n+1}{n} \right) + \ln \left(\frac{2(n+1)}{n(2+\frac{5}{n})} \right) \\ &= (\alpha+1) \ln \left(\frac{n+1}{n} \right) - \ln \left(\frac{2+\frac{5}{n}}{2} \right) \\ &= (\alpha+1) \ln \left(1 + \frac{1}{n} \right) - \ln \left(1 + \frac{5}{2n} \right). \end{aligned}$$

On en déduit donc que

$$\ln v_n = (\alpha+1) \left(\frac{1}{n} - \frac{1}{2n^2} + o_{n \rightarrow +\infty} \left(\frac{1}{n^2} \right) \right) - \left(\frac{5}{2n} - \frac{25}{8n^2} + o_{n \rightarrow +\infty} \left(\frac{1}{n^2} \right) \right) = \left(\alpha - \frac{3}{2} \right) + \frac{21-4\alpha}{8n^2} + o_{n \rightarrow +\infty} \left(\frac{1}{n^2} \right).$$

En particulier, si $\alpha \neq \frac{3}{2}$, alors $\ln v_n \underset{n \rightarrow +\infty}{\sim} \left(\alpha - \frac{3}{2} \right) \frac{1}{n}$.

Et donc, par comparaison à une série de Riemann, $\sum \ln v_n$ diverge.

En revanche, si $\alpha = \frac{3}{2}$, alors $\ln v_n \underset{n \rightarrow +\infty}{\sim} \frac{15}{8n^2}$, et donc, toujours par comparaison à une série de Riemann, $\sum \ln v_n$ converge.

Ainsi, $\alpha_0 = \frac{3}{2}$.

- 2.b. Puisque $\ln v_k = \ln((k+1)^\alpha u_{k+1}) - \ln(k^\alpha u_k)$, nous avons affaire à une série télescopique, de sorte que

$$\sum_{k=1}^n \ln v_k = \ln((n+1)^\alpha u_{n+1}) - \ln(u_1) = \ln((n+1)^\alpha u_{n+1}) - \ln \frac{2}{5}.$$

Pour $\alpha = \alpha_0$, la série $\sum \ln v_n$ converge et donc la suite de ses sommes partielles admet une limite finie ℓ .

DL/~

Rappelons qu'une expression est égale au premier terme **non nul** de son développement limité. Et en particulier, lorsque ce premier terme est nul, on n'écrira pas d'équivalent à 0, mais on ira chercher le terme non nul suivant.

Et donc $\ln((n+1)^{\alpha_0} u_{n+1}) - \ln \frac{2}{5} \xrightarrow{n \rightarrow +\infty} \ell$.

On en déduit, par continuité de l'exponentielle que

$$(n+1)^{\alpha_0} u_{n+1} \xrightarrow{n \rightarrow +\infty} \frac{2}{5} e^\ell.$$

Ainsi, si on pose $C = \frac{2}{5} e^\ell > 0$, il vient $u_n \underset{n \rightarrow +\infty}{\sim} \frac{C}{n^{\alpha_0}}$.

Et puisque $\alpha_0 = \frac{3}{2} > 1$, on en déduit, par comparaison à une série de Riemann que la série de terme général u_n converge.

2.c. En utilisant l'équivalent obtenu à la question précédente, on observe que $k\sqrt{k}u_k \underset{k \rightarrow +\infty}{\sim} C \xrightarrow{k \rightarrow +\infty} C$.

Et donc la suite $(k\sqrt{k}u_k)$ étant convergente, elle est bornée.

Soit donc $M \in \mathbf{R}$ tel que pour tout $k \in \mathbf{N}$, $k\sqrt{k}u_k \leq M \Leftrightarrow ku_k \leq \frac{M}{\sqrt{k}}$.

$$\text{Alors, } \sum_{k=1}^n ku_k \leq M \sum_{k=1}^n \frac{1}{\sqrt{k}}.$$

Mais, par décroissance de la fonction racine carrée sur $[k-1, k]$, on a, pour $k \geq 2$ et pour tout $t \in [k-1, k]$, $\frac{1}{\sqrt{k}} \leq \frac{1}{\sqrt{t}}$.

Par conséquent, par croissance de l'intégrale,

$$\frac{1}{\sqrt{k}} = \int_{k-1}^k \frac{dt}{\sqrt{k}} \leq \int_{k-1}^k \frac{dt}{\sqrt{t}}.$$

En sommant ces relations pour k variant de 2 à n , il vient

$$\sum_{k=2}^n \frac{1}{\sqrt{k}} \leq \sum_{k=2}^n \int_{k-1}^k \frac{dt}{\sqrt{t}} = \int_1^n \frac{dt}{\sqrt{t}}.$$

Relation de Chasles.

Mais cette dernière intégrale est facile à calculer :

$$\int_1^n \frac{dt}{\sqrt{t}} = [2\sqrt{t}]_1^n = 2\sqrt{n} - 2.$$

Et donc $\sum_{k=1}^n \frac{1}{\sqrt{k}} \leq 2\sqrt{n} - 2 + 1 \leq 2\sqrt{n}$.

En en déduit donc que

$$\sum_{k=0}^n ku_k = \sum_{k=1}^n ku_k \leq \underbrace{2M}_{=D} \sqrt{n}.$$

3.a. On a, pour tout k , $u_{k+1} = \frac{2k+2}{2k+5}u_k$ et donc $(2k+5)u_{k+1} = (2k+2)u_k$.

Ainsi, il vient

$$\begin{aligned} 2 \sum_{k=1}^{n+1} ku_k + 3 \sum_{k=1}^{n+1} u_k &= \sum_{k=1}^{n+1} (2k+3)u_k \\ &= \sum_{i=0}^n (2i+5)u_{i+1} \\ &= \sum_{i=0}^n (2i+2)u_i \\ &= 2 \sum_{k=0}^n ku_k + 2 \sum_{k=0}^n u_k. \end{aligned}$$

Chgt d'indice
 $i = k - 1$.

3.b. De la relation précédente, il vient

$$2(n+1)u_{n+1} + \sum_{k=1}^n u_k + 3u_{n+1} = 2u_0 = 2.$$

Or, $u_{n+1} \xrightarrow[n \rightarrow +\infty]{} 0$, et d'après l'équivalent obtenu à la question 2.b, $(n+1)u_{n+1} \underset{n \rightarrow +\infty}{\sim} \frac{C}{\sqrt{n+1}} \xrightarrow[n \rightarrow +\infty]{} 0$.

Ainsi, en passant à la limite, il vient $\sum_{k=1}^{+\infty} u_k = 2$.

En ajoutant $u_0 = 1$ on obtient donc enfin $\sum_{n=0}^{+\infty} u_n = 3$.

VARIABLES ALÉATOIRES DISCRÈTES.

ESPÉRANCE CONDITIONNELLE.

Cours	45
2.1 Variables aléatoires	45
2.2 Variables aléatoires réelles discrètes	48
2.3 Probabilités conditionnelles et espérance conditionnelle	51
Méthodes	56
Exercices	57
Correction des exercices	60

2.1 VARIABLES ALÉATOIRES

2.1.1 Espaces probabilisables

Définition 2.1 – Soit Ω un ensemble¹. Une partie $\mathcal{A} \subset \mathcal{P}(\Omega)$ est appelée tribu (ou σ -algèbre) si

- $\Omega \in \mathcal{A}$
- $\forall A \in \mathcal{A}, \bar{A} \in \mathcal{A}$
- si $(A_n)_{n \in \mathbb{N}}$ est une suite d'éléments de \mathcal{A} , alors $\bigcup_{n \in \mathbb{N}} A_n \in \mathcal{A}$.

On dit alors que (Ω, \mathcal{A}) est un espace probabilisable, et les éléments de \mathcal{A} sont appelés événements.

¹ Généralement appelé «l'univers».

Dans la pratique, on connaît rarement la tribu sur laquelle on travaille.

Le seul cas où on a vraiment accès à cette tribu est celui où Ω est un ensemble fini (c'est-à-dire lorsqu'on considère une expérience n'ayant qu'un nombre fini d'issues possibles, par exemple le lancer d'un ou plusieurs dés), et alors $\mathcal{A} = \mathcal{P}(\Omega)$ est l'ensemble des parties de Ω . Par exemple, si on lance deux dés, alors $\Omega = \llbracket 1, 6 \rrbracket^2$. Autrement dit, un élément de Ω est un couple de la forme (i, j) , où i représente le résultat du premier dé et j celui du second. Et alors \mathcal{A} est l'ensemble des parties de Ω .

Par exemple, l'événement «le premier dé vaut 6» est la partie

$$\{(6, 1), (6, 2), (6, 3), (6, 4), (6, 5), (6, 6)\} \subset \Omega.$$

L'événement «la somme des deux dés est paire» est

$$\{(1, 1), (1, 3), \dots, (2, 2), \dots, (6, 4), (6, 6)\} \subset \Omega.$$



Proposition 2.2 : Si $S \subset \mathcal{P}(\Omega)$, alors il existe une plus petite (au sens de l'inclusion) tribu contenant S . On la note $\mathcal{A}(S)$, et on l'appelle tribu engendrée par S . Autrement dit, si \mathcal{A} est une tribu telle que $S \subset \mathcal{A}$, alors $\mathcal{A}(S) \subset \mathcal{A}$.

2.1.2 Espace probabilisé

Définition 2.3 – Soit (Ω, \mathcal{A}) un espace probabilisable. Une application $P : \mathcal{A} \rightarrow \mathbf{R}_+$ est appelée probabilité si :

- $P(\Omega) = 1$
- si $(A_n)_{n \in \mathbf{N}}$ est une suite d'événements deux à deux incompatibles, alors

$$P\left(\bigcup_{n \in \mathbf{N}} A_n\right) = \sum_{n \in \mathbf{N}} P(A_n).$$

On dit alors que (Ω, \mathcal{A}, P) est un espace probabilisé.

Rappel

Deux événements A et B sont dits incompatibles si $A \cap B = \emptyset$.

Définition 2.4 – Soit $(A_i)_{i \in I}$ une famille d'événements deux à deux incompatibles indexée par I . On dit que :

- $(A_i)_{i \in I}$ est un système complet d'événements si $\bigcup_{i \in I} A_i = \Omega$.
- $(A_i)_{i \in I}$ est un système quasi-complet d'événements si $P\left(\bigcup_{i \in I} A_i\right) = 1$.

Exemple 2.5

On lance une pièce équilibrée jusqu'à faire face, et pour $i \in \mathbf{N}$, on note A_i l'événement «il faut faire i lancers avant de faire face pour la première fois».

Alors ces événements sont deux à deux incompatibles, et on sait que $P(A_i) = \left(\frac{1}{2}\right)^i$.

Donc

$$P\left(\bigcup_{i=1}^n A_i\right) = \sum_{i=1}^n P(A_i) = \sum_{i=1}^n \left(\frac{1}{2}\right)^i = 1 - \frac{1}{2^n}.$$

Par le théorème de la limite monotone, on a alors

$$P\left(\bigcup_{i=1}^{\infty} A_i\right) = \lim_{n \rightarrow \infty} P\left(\bigcup_{i=1}^n A_i\right) = \lim_{n \rightarrow \infty} 1 - \frac{1}{2^n} = 1.$$

Ainsi, $(A_i)_{i \in \mathbf{N}^*}$ est un système quasi-complet d'événements. Ce n'est pas un système complet d'événements, car l'événement «ne jamais faire face» n'est pas inclus dans

$\bigcup_{i=1}^{\infty} A_i$. En revanche, cet événement est de probabilité nulle.

Limite monotone

Nous parlons ici du théorème de la limite monotone relatif aux probabilités : si (B_n) est une suite croissante d'événements, c'est-à-dire telle que $B_n \subset B_{n+1}$, on a

$$P\left(\bigcup_{n=0}^{+\infty} B_n\right) = \lim_{n \rightarrow +\infty} P(B_n).$$

Il s'applique ici en posant

$$B_n = \bigcup_{i=0}^n A_i$$

car alors (B_n) est une suite croissante d'événements.



Proposition 2.6 : Si $(A_i)_{i \in I}$ est un système quasi-complet d'événements tel que $\bigcup_{i \in I} A_i \neq \Omega$, alors $(A_i)_{i \in I} \cup \left(\overline{\bigcup_{i \in I} A_i}\right)$ est un système complet d'événements.

Démonstration. Par définition, $\bigcup_{i \in I} A_i \cup \overline{\bigcup_{i \in I} A_i} = \Omega$. De plus, si $i \neq j \in I$, alors A_i et A_j sont incompatibles par définition, et si $i \in I$, alors

$$A_i \cap \overline{\bigcup_{i \in I} A_i} = \emptyset.$$

□

Dire que $\bigcup_{i \in I} A_i \neq \Omega$ revient à dire que (A_i) n'est pas un système complet d'événements.

Définition 2.7 – Un événement A est dit

- (quasi)-certain si $P(A) = 1$
- (quasi)-impossible si $P(A) = 0$.

2.1.3 Variables aléatoires réelles



Définition 2.8 – On appelle tribu des boréliens la tribu de \mathbf{R} engendrée par les intervalles $] -\infty; a]$, $a \in \mathbf{R}$. On la note $\mathcal{B}(\mathbf{R})$.

Remarque. Tous les intervalles sont dans $\mathcal{B}(\mathbf{R})$ car :

- $]a; +\infty[$ est le complémentaire de $] -\infty; a]$, donc appartient à $\mathcal{B}(\mathbf{R})$
- $] -\infty; a[= \bigcup_{n \in \mathbf{N}^*}] -\infty; a - \frac{1}{n}]$
- $[a; +\infty[$ est le complémentaire de $] -\infty; a]$.
- si $a < b$, alors $[a; b] = [a; +\infty[\cap] -\infty; b]$
- etc

Définition 2.9 – Soit (Ω, \mathcal{A}) un espace probabilisable. Une variable aléatoire réelle sur Ω est une application $X : \Omega \rightarrow \mathbf{R}$ telle que

$$\forall x \in \mathbf{R}, [X \leq x] = \{\omega \in \Omega : X(\omega) \leq x\} \in \mathcal{A}.$$

Proposition 2.10 : Soit X une variable aléatoire réelle définie sur (Ω, \mathcal{A}) . Alors

$$\forall B \in \mathcal{B}(\mathbf{R}), [X \in B] = \{\omega \in \Omega : X(\omega) \in B\} \in \mathcal{A}.$$

Si on prend $B =] -\infty, x],] -\infty, x[, [x, +\infty[$, cela signifie notamment que si $x \in \mathbf{R}$, alors les ensembles $[X \leq x], [X < x], [X \geq x]$, etc, sont bien des événements².

Proposition 2.11 : Soient X et Y deux variables aléatoires réelles définies sur (Ω, \mathcal{A}) , et soit $\lambda \in \mathbf{R}$. Alors $X + Y, XY$ et λX sont également des variables aléatoires réelles sur (Ω, \mathcal{A}) .

Démonstration. Admis. □

Définition 2.12 – Si X est une variable aléatoire réelle sur un espace probabilisé (Ω, \mathcal{A}, P) , on appelle fonction de répartition (ou loi) de X la fonction

$$F_X : \mathbf{R} \rightarrow \mathbf{R} \\ x \mapsto P(X \leq x)$$

Proposition 2.13 : F_X est croissante, continue à droite en tout point et

$$\lim_{x \rightarrow -\infty} F_X(x) = 0 \text{ et } \lim_{x \rightarrow +\infty} F_X(x) = 1.$$

Terminologie

Les boréliens sont ainsi nommés en l'honneur d'Émile Borel (1871–1956), mathématicien français qui contribua à définir le formalisme actuel des probabilités.

En pratique, on a très rarement besoin de revenir à cette définition, sauf peut-être dans des sujets de Maths II.

² et donc on a bien le droit de s'intéresser à la probabilité de ces événements.

Vigilance !

Ne pas oublier qu'il s'agit d'une inégalité large. Si cela n'a pas d'importance pour une variable à densité, ce n'est pas le cas pour des variables discrètes.

Inversement, si une fonction f vérifie ces propriétés, alors il existe une variable aléatoire X telle que $f = F_X$.

Définition 2.14 – Soit (Ω, \mathcal{A}) un espace probabilisable, et soit $A \in \mathcal{A}$. On appelle variable indicatrice de A , et on note $\mathbb{1}_A$ la fonction $\mathbb{1}_A : \Omega \rightarrow \mathbf{R}$ définie par

$$\mathbb{1}_A(\omega) = \begin{cases} 0 & \text{si } \omega \notin A \\ 1 & \text{si } \omega \in A \end{cases}$$

Proposition 2.15 : Si $A \in \mathcal{A}$, alors $\mathbb{1}_A$ est une variable aléatoire réelle.

Démonstration. Soit $x \in \mathbf{R}$. Si $x < 0$, alors $[\mathbb{1}_A \leq x] = \emptyset \in \mathcal{A}$.
Si $0 \leq x < 1$, alors

$$[\mathbb{1}_A \leq x] = [\mathbb{1}_A = 0] = \{\omega \in \Omega : \omega \notin A\} = \bar{A} \in \mathcal{A}.$$

Enfin, si $x \geq 1$, alors $[\mathbb{1}_A \leq x] = [\mathbb{1}_A \leq 1] = \Omega \in \mathcal{A}$.

Ainsi, $\forall x \in \mathbf{R}, [\mathbb{1}_A \leq x] \in \mathcal{A}$: $\mathbb{1}_A$ est bien une variable aléatoire réelle sur (Ω, \mathcal{A}) . \square

Complément

$\mathbb{1}_A$ est une variable qui ne prend que les valeurs 0 et 1 : elle suit donc une loi de Bernoulli. De plus, $P(\mathbb{1}_A = 1) = P(A)$, donc le paramètre de la loi de Bernoulli est $P(A)$. Cela permet notamment de calculer l'espérance et la variance de $\mathbb{1}_A$.

Exemple 2.16

On lance n fois une pièce équilibrée. On note A_i l'événement «on fait pile au $i^{\text{ème}}$ lancer». Alors

- $\sum_{i=1}^n \mathbb{1}_{A_i}$ est le nombre de piles obtenu au cours des n lancers.
- $1 - \prod_{i=1}^n \mathbb{1}_{A_i}$ est la variable indicatrice de l'événement «faire au moins un face lors des n lancers».

2.2 VARIABLES ALÉATOIRES RÉELLES DISCRÈTES

Dans toute la suite, (Ω, \mathcal{A}, P) désigne un espace probabilisé.

2.2.1 Définition

Définition 2.17 – Une variable aléatoire réelle X définie sur Ω est dite discrète si l'ensemble $X(\Omega)$ des valeurs prises par X est dénombrable.

Remarque. Autrement dit, cela signifie qu'il existe un ensemble $I \subset \mathbf{N}$, fini ou infini, tel qu'on puisse «numéroter» les éléments du support de X à l'aide des éléments de I : $X(\Omega) = \{x_i, i \in \mathbf{N}\}$.

Exemples 2.18

- Si $X \hookrightarrow \mathcal{G}(p)$, alors X est discrète car $X(\Omega) = \mathbf{N}^*$ est dénombrable.
- Si $X \hookrightarrow \mathcal{N}(0; 1)$, alors X n'est pas discrète car $X(\Omega) = \mathbf{R}$, et on ne peut pas «numéroter» les éléments de \mathbf{R} .

Proposition 2.19 : Soit X une variable aléatoire discrète. Alors la loi de X est entièrement caractérisée par la donnée :

1. du support $X(\Omega)$ de X .
2. de tous les $P(X = x)$, pour $x \in X(\Omega)$.

Autrement dit

Pour une variable aléatoire discrète, on ne donne généralement pas la fonction de répartition, mais seulement les $P(X = x)$, pour $x \in X(\Omega)$, ce qui est plus simple à manipuler.

2.2.2 Espérance

Définition 2.20 – Une variable aléatoire réelle discrète X avec $X(\Omega) = \{x_i, i \in I\}$ admet une espérance si la série $\sum_i x_i P(X = x_i)$ converge **absolument**. Dans ce cas, on note $E(X)$ la somme de cette série.

Les résultats qui suivent, bien qu'intuitifs, sont en réalité très difficiles à démontrer, et nous les admettons.

Proposition 2.21 : Si X et Y sont deux variables aléatoires discrètes sur (Ω, \mathcal{A}, P) qui admettent une espérance et si $(\lambda, \mu) \in \mathbf{R}^2$, alors $\lambda X + \mu Y$ admet une espérance et

$$E(\lambda X + \mu Y) = \lambda E(X) + \mu E(Y).$$

Théorème 2.22 (de transfert) : Soit X une variable aléatoire réelle discrète sur Ω , avec $X(\Omega) = \{x_i, i \in I\}, I \subset \mathbf{N}$. Soit $g : X(\Omega) \rightarrow \mathbf{R}$. Alors la fonction $g(X) : \Omega \rightarrow \mathbf{R}$ est une variable aléatoire discrète, qui admet une espérance si et seulement si la série $\sum_i g(x_i)P(X = x_i)$ converge absolument. Dans ce cas, on a alors

$$E(g(X)) = \sum_{i \in I} g(x_i)P(X = x_i).$$

Proposition 2.23 : Si X et Y sont deux variables aléatoires discrètes telles que $0 \leq |X| \leq Y$ presque sûrement (i.e. si $P(0 \leq |X| \leq Y) = 1$), et si Y admet une espérance, alors X admet une espérance, et dans ce cas $|E(X)| \leq E(Y)$.

Corollaire 2.24 (Croissance de l'espérance) – Si X et Y sont deux variables aléatoires discrètes admettant une espérance avec $X \leq Y$ presque sûrement, alors $E(X) \leq E(Y)$.

Exemple 2.25

Soit X une variable aléatoire suivant une loi géométrique, et soit $Y = \sqrt{X}$. Alors, $\forall \omega \in \Omega, 0 \leq Y(\omega) \leq X(\omega)$, car $\forall n \in \mathbf{N}^*, \sqrt{n} \leq n$. Par conséquent, l'événement $0 \leq Y \leq X$ est certain. Puisque X admet une espérance, il en est de même de Y , et $E(Y) \leq E(X)$.

2.2.3 Variance

Définition 2.26 – Soit X une variable aléatoire discrète X admettant une espérance. On dit que X admet une variance si $(X - E(X))^2$ admet une espérance, et on note alors $V(X) = E((X - E(X))^2)$.

Remarque. Puisque $(X - E(X))^2 \geq 0$, en cas d'existence, on a

$$V(X) = E((X - E(X))^2) \geq E(0) = 0.$$

Remarque

En termes savants : X admet une espérance si et seulement si la famille

$$(xP(X = x))_{x \in X(\Omega)}$$

est sommable.

En d'autres termes, l'espérance est linéaire.

Cas particulier

Si on prend $g(x) = |x|$, on remarque alors que $|X|$ admet une espérance si et seulement si X admet une espérance.

Proposition 2.27 : Soit X une variable aléatoire discrète. Alors :

- X admet une variance si et seulement si X^2 admet une espérance (i.e. si et seulement si X admet un moment d'ordre 2).
Dans ce cas, $V(X) = E(X^2) - E(X)^2$ (formule de Huygens).
- Si X n'admet pas d'espérance, alors X n'admet pas de moment d'ordre 2.

2.2.4 Lois usuelles

Dans toute la suite, on note X une variable aléatoire sur un espace probabilisé (Ω, \mathcal{A}, P) .

Nom	Paramètres	$X(\Omega)$	Loi	$E(X)$	$V(X)$
Loi de Bernoulli $\mathcal{B}(p)$	$p \in [0, 1]$	$\{0, 1\}$	$P(X = 0) = 1 - p$ $P(X = 1) = p$	p	$p(1 - p)$
Loi uniforme sur $\llbracket 1, n \rrbracket$	$n \in \mathbf{N}^*$	$\llbracket 1, n \rrbracket$	$P(X = k) = \frac{1}{n}$	$\frac{n+1}{2}$	$\frac{n^2-1}{12}$
Loi uniforme sur $\llbracket a, b \rrbracket$	$(a, b) \in \mathbf{Z}^2, a < b$	$\llbracket a, b \rrbracket$	$P(X = k) = \frac{1}{b-a+1}$	$\frac{a+b}{2}$	$\frac{(b-a+1)^2-1}{12}$
Loi binomiale $\mathcal{B}(n, p)$	$n \in \mathbf{N}^*$ $p \in [0; 1]$	$\llbracket 0, n \rrbracket$	$P(X = k) = \binom{n}{k} p^k (1-p)^{n-k}$	np	$np(1-p)$
Loi de Poisson $\mathcal{P}(\lambda)$	$\lambda > 0$	\mathbf{N}	$P(X = k) = e^{-\lambda} \frac{\lambda^k}{k!}$	λ	λ
Loi géométrique $\mathcal{G}(p)$	$p \in]0; 1[$	\mathbf{N}^*	$P(X = k) = p(1-p)^{k-1}$	$\frac{1}{p}$	$\frac{1-p}{p^2}$

Situations typiques :

- X suit la loi $\mathcal{B}(p)$ lorsqu'on considère une expérience à deux issues : succès avec probabilité p et échec (avec probabilité $1-p$), et si X vaut 1 en cas de succès et 0 sinon (par exemple : un pile ou face).
- Un nombre tiré au hasard dans $\llbracket 1, n \rrbracket$ suit la loi $\mathcal{U}(\llbracket 1, n \rrbracket)$
- Le nombre de succès lors de n répétitions indépendantes d'un processus de Bernoulli de paramètre p suit la loi $\mathcal{B}(n, p)$ (par exemple : on lance n fois une pièce).
- Si on répète de manières indépendantes un processus de Bernoulli de paramètre p jusqu'à l'obtention d'un succès, alors le nombre d'essais suit la loi $\mathcal{G}(p)$ (par exemple : on lance une pièce jusqu'à faire pile).
- La loi de Poisson ne correspond pas à une situation typique, et lorsqu'il faut l'utiliser, l'énoncé l'indique explicitement.

Notation

Pour une loi géométrique, on note souvent $q = 1 - p$ la probabilité d'échec, de sorte que $P(X = k) = pq^{k-1}$.

Simulation en Scilab

La commande `grand(q, r, loi)` permet de simuler une matrice $q \times r$ dont les entrées suivent la loi spécifiée. Le paramètre **loi** doit prendre l'une des valeurs suivantes :

- 'uin', a, b pour $\mathcal{U}(\llbracket a, b \rrbracket)$.
- 'bin', n, p pour $\mathcal{B}(n, p)$.
- 'poi', λ pour $\mathcal{P}(\lambda)$.
- 'geom', p pour $\mathcal{G}(p)$.

Par exemple, `grand(10, 10, 'bin', 20, 0.1)` retourne une matrice 10×10 dont les nombres ont été tirés aléatoirement suivant $\mathcal{B}(20, 0.1)$.

Il n'y a pas de commande spécifique pour la loi de Bernoulli de paramètre p , mais notons que $\mathcal{B}(p) = \mathcal{B}(1, p)$ et donc nous pouvons utiliser `grand(q, r, 'bin', 1, p)`.

2.3 PROBABILITÉS CONDITIONNELLES ET ESPÉRANCE CONDITIONNELLE

Dans toute cette partie, (Ω, \mathcal{A}, P) désigne un espace probabilisé.

2.3.1 Rappels sur les probabilités conditionnelles

Définition 2.28 – Soit $A \in \mathcal{A}$ un événement tel que $P(A) \neq 0$. Alors pour tout événement B , on appelle probabilité de B sachant A le réel noté $P_A(B)$ (ou parfois $P(B|A)$) défini par

$$P_A(B) = \frac{P(A \cap B)}{P(A)}.$$

Proposition 2.29 : P_A est une probabilité sur (Ω, \mathcal{A}) .

Remarque. Cela implique que tous les théorèmes généraux relatifs aux probabilités s'appliquent encore pour P_A .

Par exemple $P_A(\bar{B}) = 1 - P_A(B)$ et $P_A(B \cup C) = P_A(B) + P_A(C) - P_A(B \cap C)$.

Proposition 2.30 (Formule des probabilités composées) : Soient A_1, A_2, \dots, A_n des événements tels que $P(A_1 \cap \dots \cap A_{n-1}) \neq 0$. Alors

$$P\left(\bigcap_{i=1}^n A_i\right) = P(A_1)P_{A_1}(A_2)P_{A_1 \cap A_2}(A_3) \cdots P_{A_1 \cap \dots \cap A_{n-1}}(A_n).$$

Exemple 2.31

Une urne contient $n - 1$ boules rouges et une boule verte. On effectue des tirages sans remise jusqu'à obtenir la boule verte, et on note X la variable aléatoire égale au rang d'apparition de la boule verte.

Il est évident que $X(\Omega) = \llbracket 1, n \rrbracket$.

Notons R_i : «on obtient une boule rouge au $i^{\text{ème}}$ tirage», de sorte que pour tout $k \in \llbracket 1, n \rrbracket$,

$$[X = k] = R_1 \cap R_2 \cap \dots \cap R_{k-1} \cap \bar{R}_k.$$

Et donc par la formule des probabilités composées,

$$\begin{aligned} P(X = k) &= P(R_1)P_{R_1}(R_2) \cdots P_{R_1 \cap \dots \cap R_{k-2}}(R_{k-1})P_{R_1 \cap \dots \cap R_{k-1}}(\bar{R}_k) \\ &= \frac{n-1}{n} \frac{n-2}{n-1} \cdots \frac{n-1-(k-2)}{n-(k-2)} \frac{1}{n-(k-1)} = \frac{1}{n}. \end{aligned}$$

Et donc $X \leftrightarrow \mathcal{U}(\llbracket 1, n \rrbracket)$.

Proposition 2.32 (Formule des probabilités totales) : Si $(A_i)_{i \in I}$ est un système complet d'événements, avec I fini ou dénombrable, et si on note $J = \{i \in I : P(A_i) \neq 0\}$, alors

$$\forall B \in \mathcal{A}, P(B) = \sum_{i \in I} P(A_i \cap B) = \sum_{i \in J} P(A_i)P_{A_i}(B).$$

Remarques. • Notons qu'il y a en fait deux formes de la formule des probabilités totales : l'une avec les $P(A_i \cap B)$, et l'autre avec les $P_{A_i}(B)$. On choisira l'une ou l'autre suivant que l'on connaisse plutôt les probabilités conditionnelles $P_{A_i}(B)$ ou les $P(A_i \cap B)$.

• La formule reste valable si $(A_i)_{i \in I}$ est un système quasi-complet d'événements.

Danger !

On ne parle pas d'événement conditionnel, cela n'a pas de sens. Ce qu'on aurait envie d'appeler « A sachant B » est en fait $A \cap B$ (A et B sont tous deux réalisés).

Méthode

Pour calculer la probabilité d'une intersection, sauf si l'on sait que ces événements sont indépendants, on fera appel à cette formule.

Proposition 2.33 (Formules de Bayes) :

- Si A et B sont deux événements de probabilités non nulles,

$$P_B(A) = \frac{P_A(B)P(A)}{P(B)}.$$

- Si $(A_i)_{i \in I}$ est un système complet d'événements fini ou dénombrable, si $J = \{i \in I : P(A_i) \neq 0\}$, alors pour tout événement B de probabilité non nulle, et pour tout $i \in J$

$$P_B(A_i) = \frac{P_{A_i}(B)P(A_i)}{\sum_{j \in J} P_{A_j}(B)P(A_j)}.$$

Remarque. La formule de Bayes et la formule des probabilités totales restent valables pour des systèmes quasi-complets d'événements.

2.3.2 Loi conditionnelle, espérance conditionnelle

Définition 2.34 – Si X est une variable aléatoire réelle discrète sur (Ω, \mathcal{A}, P) , on appelle loi de X sachant A , ou loi de X conditionnellement à A , la loi de X pour la probabilité P_A , c'est-à-dire la donnée de tous les couples $(x, P_A(X = x))$, pour $x \in X(\Omega)$.

Exemple 2.35

Le nombre de clients qui entrent dans un restaurant un soir de semaine est une variable aléatoire X qui suit une loi de Poisson de paramètre λ .

Chaque client choisit, indépendamment du choix de ses voisins, de la viande avec probabilité p et du poisson avec la probabilité $1 - p$. On note Y le nombre de clients qui choisissent de la viande.

Alors $\{[X = n], n \in \mathbf{N}\}$ est un système complet d'événements, donc

$$\forall k \in \mathbf{N}, P(Y = k) = \sum_{n \in \mathbf{N}} P(X = n)P_{[X=n]}(Y = k).$$

Mais la loi de Y sachant $[X = n]$ est la loi binomiale $\mathcal{B}(n, p)$, donc

$$P_{[X=n]}(Y = k) = \begin{cases} \binom{n}{k} p^k (1-p)^{n-k} & \text{si } 0 \leq k \leq n \\ 0 & \text{sinon} \end{cases}$$

Ainsi, $\forall k \in \mathbf{N}$,

$$\begin{aligned} P(Y = k) &= \sum_{n \geq k} e^{-\lambda} \frac{\lambda^n}{n!} \binom{n}{k} p^k (1-p)^{n-k} \\ &= \sum_{n \geq k} e^{-\lambda} \frac{\lambda^n}{n!} \frac{n!}{k!(n-k)!} p^k (1-p)^{n-k} \\ &= e^{-\lambda} \frac{p^k \lambda^k}{k!} \sum_{n \geq k} \frac{\lambda^{n-k}}{(n-k)!} (1-p)^{n-k} \\ &= e^{-\lambda} \frac{p^k \lambda^k}{k!} \sum_{i \geq 0} \frac{\lambda^i (1-p)^i}{i!} \\ &= e^{-\lambda} \frac{p^k \lambda^k}{k!} e^{\lambda(1-p)} \\ &= e^{-\lambda p} \frac{(p\lambda)^k}{k!}. \end{aligned}$$

Ainsi, Y suit une loi de Poisson de paramètre λp .

Méthode

La formule de Bayes est utile lorsqu'on cherche une probabilité conditionnelle $P_A(B)$ et que l'on connaît $P_B(A)$. En général, cela suffit à faire penser à la formule de Bayes.

Exemple 2.35

Le nombre de clients qui entrent dans un restaurant un soir de semaine est une variable aléatoire X qui suit une loi de Poisson de paramètre λ .

Chaque client choisit, indépendamment du choix de ses voisins, de la viande avec probabilité p et du poisson avec la probabilité $1 - p$. On note Y le nombre de clients qui choisissent de la viande.

Alors $\{[X = n], n \in \mathbf{N}\}$ est un système complet d'événements, donc

$$\forall k \in \mathbf{N}, P(Y = k) = \sum_{n \in \mathbf{N}} P(X = n)P_{[X=n]}(Y = k).$$

Mais la loi de Y sachant $[X = n]$ est la loi binomiale $\mathcal{B}(n, p)$, donc

$$P_{[X=n]}(Y = k) = \begin{cases} \binom{n}{k} p^k (1-p)^{n-k} & \text{si } 0 \leq k \leq n \\ 0 & \text{sinon} \end{cases}$$

Ainsi, $\forall k \in \mathbf{N}$,

$$\begin{aligned} P(Y = k) &= \sum_{n \geq k} e^{-\lambda} \frac{\lambda^n}{n!} \binom{n}{k} p^k (1-p)^{n-k} \\ &= \sum_{n \geq k} e^{-\lambda} \frac{\lambda^n}{n!} \frac{n!}{k!(n-k)!} p^k (1-p)^{n-k} \\ &= e^{-\lambda} \frac{p^k \lambda^k}{k!} \sum_{n \geq k} \frac{\lambda^{n-k}}{(n-k)!} (1-p)^{n-k} \\ &= e^{-\lambda} \frac{p^k \lambda^k}{k!} \sum_{i \geq 0} \frac{\lambda^i (1-p)^i}{i!} \\ &= e^{-\lambda} \frac{p^k \lambda^k}{k!} e^{\lambda(1-p)} \\ &= e^{-\lambda p} \frac{(p\lambda)^k}{k!}. \end{aligned}$$

Ainsi, Y suit une loi de Poisson de paramètre λp .

Explication

En effet, une fois le nombre de clients fixé égal à n , on reconnaît une situation typique d'une loi binomiale (répétition d'épreuves de Bernoulli indépendantes).

Les termes avec $n < k$ sont nuls.

Chgt d'indice

$i = n - k$.

Définition 2.36 – Soit X une variable aléatoire réelle discrète sur (Ω, \mathcal{A}, P) , avec $X(\Omega) = \{x_i, i \in I\}$. Soit $A \in \mathcal{A}$ tel que $P(A) \neq 0$. On dit que X admet une espérance sachant A (ou espérance conditionnellement à A) si X admet une espérance pour la probabilité P_A , autrement dit si la famille $(x_i P_A(X = x_i))_{i \in I}$ est sommable.

Dans ce cas, on note

$$E(X|A) = \sum_{i \in I} x_i P_A(X = x_i)$$

l'espérance de X sachant A .

Exemple 2.37

Reprenons l'exemple des clients au restaurant. Alors pour $n \in \mathbf{N}$,

$$\begin{aligned} E(Y|X = n) &= \sum_{k \in \mathbf{N}} k P_{[X=n]}(Y = k) \\ &= \sum_{k=0}^n k \binom{n}{k} p^k (1-p)^{n-k} \\ &= \sum_{k=1}^n k \binom{n}{k} p^k (1-p)^{n-k} \\ &= \sum_{k=1}^n np \binom{n-1}{k-1} p^{k-1} (1-p)^{n-k} \\ &= np \sum_{\ell=0}^{n-1} \binom{n-1}{\ell} p^\ell (1-p)^{n-1-\ell} \\ &= np(p + (1-p))^{n-1} = np \end{aligned}$$

Sous réserve de convergence.

La somme est finie donc converge bien (et donc $E(Y|X = n)$ existe).

Rappel

$$k \binom{n}{k} = n \binom{n-1}{k-1}.$$

Si cette formule est au programme, il est aisé de la retrouver en revenant aux factorielles.

Notons que ce résultat est bien cohérent à ce qu'on aurait pu attendre, vu que la loi de Y sachant $[X = n]$ est la loi binomiale $\mathcal{B}(n, p)$, d'espérance np . D'ailleurs, la somme qui apparaît à la seconde ligne du calcul précédent est exactement celle qui définit l'espérance d'une loi binomiale $\mathcal{B}(n, p)$, dont on sait qu'elle vaut np .

En pratique

On ne fera jamais un tel calcul : la loi conditionnelle de Y sachant $[X = n]$ est une loi binomiale, donc son espérance conditionnelle est l'espérance d'une loi binomiale.

2.3.3 La formule de l'espérance totale

À l'instar de la formule des probabilités totales, qui permet de déterminer une probabilité à l'aide de probabilités conditionnelles, nous voyons à présent une formule permettant de déterminer une espérance à l'aide d'espérances conditionnelles.

Commençons par un cas facile : celui où X est une variable aléatoire finie, avec $X(\Omega) = \{x_1, \dots, x_n\}$ et où A_1, \dots, A_p est un système complet d'événements. Alors

$$\begin{aligned} E(X) &= \sum_{i=1}^n x_i P(X = x_i) \\ &= \sum_{i=1}^n x_i \sum_{j=1}^p P(A_j) P_{A_j}(X = x_i) \\ &= \sum_{i=1}^n \sum_{j=1}^p P(A_j) x_i P_{A_j}(X = x_i) \\ &= \sum_{j=1}^p P(A_j) \sum_{i=1}^n x_i P_{A_j}(X = x_i) \\ &= \sum_{j=1}^p P(A_j) E(X|A_j) \end{aligned}$$

Par définition de l'espérance.

Formule des probabilités totales.

Définition de l'espérance conditionnelle.

Cette formule se généralise au cas où X n'est pas nécessairement à support fini, et où le système complet d'événements n'est pas nécessairement fini. Toutefois, il y a alors un certain nombre de précautions à prendre, puisqu'au lieu de deux sommes finies, nous sommes confrontées à deux sommes infinies, donc potentiellement divergentes, et qu'il n'est pas possible d'invertir sans prendre de précautions³.

³ C'est le théorème de sommation par paquets.

Proposition 2.38 (Formule de l'espérance totale) : Soit X une variable aléatoire discrète sur (Ω, \mathcal{A}, P) , et soit $(A_i)_{i \in I}$ un système complet d'événements, avec I dénombrable ou fini, et notons $J = \{i \in I : P(A_i) \neq 0\}$. Alors X admet une espérance si et seulement si :

- $\forall i \in J, E(X|A_i)$ existe
- $\sum_{i \in J} P(A_i)E(|X||A_i)$ est sommable.

Dans ce cas, on a alors
$$E(X) = \sum_{i \in J} P(A_i)E(X|A_i) = \sum_{(x,i) \in X(\Omega) \times J} x P_{A_i}(X = x) P(A_i).$$



Démonstration. Supposons, pour simplifier⁴ que $J = \emptyset$, c'est-à-dire que $\forall i \in I, P(A_i) \neq 0$. Numérotons les éléments du support de X , de sorte que $X(\Omega) = \{x_n, n \in \mathbf{N}\}$. Notons, pour $n \in \mathbf{N}$ et $i \in I, u_{i,n} = x_n P(A_i) P_{A_i}(X = x_n) = x_n P([X = x_n] \cap A_i)$. Par la formule des probabilités totales, pour tout $n \in \mathbf{N}, \sum_i u_{i,n}$ converge absolument et

⁴ On ne perd pas de généralité : enlever les termes nuls ne change rien à la somme.

$$\sum_{i \in I} |u_{i,n}| = \sum_{i \in I} |x_n| P(A_i) P_{A_i}(X = x_n) = |x_n| P(X = x_n).$$

\Rightarrow : supposons que $E(X)$ existe. Alors pour tout $i \in I$, et tout $n \in \mathbf{N}$, on a

$$0 \leq |x_n| P_{A_i}(X = x_n) = \frac{1}{P(A_i)} |x_n| P([X = x_n] \cap A_i) \leq \frac{1}{P(A_i)} |x_n| P(X = x_n)$$

et donc par le théorème de comparaison⁵, $E(X|A_i)$ existe.

⁵ La série de terme général $|x_n| P(X = x_n)$ converge par hypothèse.

De plus, $\sum_n \left(\sum_i |u_{i,n}| \right) = \sum_n |x_n| P(X = x_n)$ converge par hypothèse, et donc la famille $(u_{i,n})$ est sommable.

Par le théorème de sommation par paquets, on a donc

$$\begin{aligned} E(X) &= \sum_{n \in \mathbf{N}} x_n P(X = x_n) \\ &= \sum_{n \in \mathbf{N}} x_n \sum_{i \in I} P_{A_i}(X = x_n) P(A_i) = \sum_{n \in \mathbf{N}} \sum_{i \in I} u_{i,n} \\ &= \sum_{i \in I} \sum_{n \in \mathbf{N}} u_{i,n} = \sum_{i \in I} P(A_i) \sum_{n \in \mathbf{N}} x_n P_{A_i}(X = x_n) \\ &= \sum_{i \in I} P(A_i) E(X|A_i). \end{aligned}$$

La famille étant sommable, il est possible d'invertir les deux sommes.

\Leftarrow : supposons que les espérances conditionnelles existent et que $\sum_i P(A_i) E(|X||A_i)$ converge.

Alors par hypothèse⁶, pour tout $i \in I, \sum_n |u_{i,n}|$ converge et vaut $E(|X||A_i)$ et donc $\sum_i \left(\sum_n |u_{i,n}| \right)$ converge, de sorte que la famille $(u_{i,n})$ est sommable.

⁶ C'est la définition de l'espérance conditionnelle.

En particulier, $\sum_{n \in \mathbf{N}} \sum_{i \in I} |u_{i,n}| = \sum_n |x_n| P(X = x_n)$ converge, donc $E(X)$ existe, et alors, par le même calcul que précédemment⁷, on a encore l'égalité $E(X) = \sum_{i \in I} P(A_i) E(X|A_i)$. \square

⁷ C'est-à-dire l'inversion des deux sommes.

Remarques. • En pratique, les variables X considérées seront toujours positives. Et donc on peut remplacer les $E(|X||A_i)$ par des $E(X|A_i)$.

- Les systèmes complets d'événements que nous considérons sont toujours soit finis (et alors la sommabilité de la famille $P(A_i)E(|X| | A_i)$ est automatique du moment que les espérances conditionnelles existent, puisqu'il s'agit d'une somme finie), soit indicés par \mathbf{N} ou \mathbf{N}^* , et alors étudier la sommabilité de la famille $P(A_i)E(|X| | A_i)$ revient à étudier la convergence d'une série.

En pratique, il suffira donc de connaître l'énoncé suivant :

Corollaire 2.39 (Formule de l'espérance totale version 2) : Si X est une variable aléatoire discrète à valeurs positives et si $(A_n)_{n \in \mathbf{N}}$ est un système complet d'événements, alors X admet une espérance si et seulement si :

- $\forall n \in \mathbf{N}, E(X|A_n)$ existe
- $\sum_n P(A_n)E(X|A_n)$ converge.

Dans ce cas, $E(X) = \sum_{n=0}^{+\infty} P(A_n)E(X|A_n)$.

Démonstration. Si X est positive, alors $|X| = X$ et donc $E(|X| | A_i) = E(X | A_i)$. □

Exemple 2.40

Reprenons l'exemple du restaurant. On sait déjà que $E(Y|X = n) = np$.

Mais, pour $n \geq 1$, $P(X = n)E(Y|X = n) = npe^{-\lambda} \frac{\lambda^n}{n!} = p\lambda e^{-\lambda} \frac{\lambda^{n-1}}{(n-1)!}$.

Or, la série $\sum_{n \geq 1} p\lambda e^{-\lambda} \frac{\lambda^{n-1}}{(n-1)!}$ converge absolument.

Donc Y admet une espérance, et

$$E(Y) = \sum_{n=0}^{\infty} P(X = n)E(Y|X = n) = \sum_{n=1}^{\infty} p\lambda e^{-\lambda} \frac{\lambda^{n-1}}{(n-1)!} = p\lambda e^{-\lambda} e^{\lambda} = p\lambda.$$

Notons que là encore, ce résultat était prévisible puisque nous avons prouvé précédemment que Y suit une loi de Poisson de paramètre $p\lambda$.

Mais si on ne s'intéresse qu'à l'espérance de Y , il est bien plus rapide d'utiliser la formule de l'espérance totale comme on vient de la faire, sans nécessiter le calcul fastidieux de la loi de Y via la formule des probabilités totales.

MÉTHODES DU CHAPITRE 2

✓ MÉTHODE 2.1 : Reconnaître et utiliser une loi usuelle.

Il faut être capable de reconnaître les lois suivantes :

- Loi de Bernoulli $\mathcal{B}(p)$: c'est le cas de toute variable aléatoire qui ne prend que les valeurs 0 et 1 (en particulier une variable indicatrice). Alors $p = P(X = 1)$.
- Loi binomiale $\mathcal{B}(n, p)$: nombre de succès lors d'une répétition de n épreuves de Bernoulli **indépendantes** dont la proba de succès vaut p .
- Loi géométrique $\mathcal{G}(p)$: nombre de tentatives nécessaires à l'obtention du premier succès lors d'une répétition d'épreuves de Bernoulli **indépendantes** dont la proba de succès vaut p .

En particulier, les deux dernières peuvent apparaître lors de tirages **avec** remise, rarement lors de tirages sans remises (car on perd alors l'indépendance des répétitions).

Il faut connaître sans hésitation les espérances et variances des lois usuelles.

Exercices de TD : 2.1, 2.3, 2.11. Cas de lois conditionnelles : 2.13, 2.15

✓ MÉTHODE 2.2 : Étudier l'existence/calculer la valeur de l'espérance d'une variable discrète X .

Rappelons qu'une variable finie admet **toujours** une espérance.

Sinon, on pourra étudier l'existence de $E(X)$:

- soit en revenant à la définition et en montrant qu'une série converge **absolument**.
- soit en reconnaissant une loi usuelle ou une combinaison linéaire de telles lois.
- soit en utilisant le théorème de transfert si on reconnu que $X = f(Y)$, ou Y est une variable aléatoire. Dans ce cas encore, on souhaite une convergence **absolue**.
- soit en utilisant la proposition 2.2.2 qui prouve l'existence d'une espérance à l'aide d'une majoration (mais ne donne alors pas la valeur de cette espérance !).
- soit à l'aide de la formule de l'espérance totale si l'on dispose d'espérances conditionnelles (voir ci-dessous).

Exercices de TD : 2.4, 2.6, 2.7, 2.8, 2.9, 2.10, 2.11

✓ MÉTHODE 2.3 : Utiliser un système complet d'événements.

Lorsqu'on a envie de faire une disjonction de cas, un système complet d'événements se cache souvent derrière.

Les plus couramment utilisés sont $\{A, \bar{A}\}$ où A est un événement et $\{[X = x], x \in X(\Omega)\}$, où X est une variable aléatoire discrète.

Suivant le contexte, on utilisera alors la formule des probabilités totales ou la formule de Bayes (cas des probabilités conditionnelles).

Exercices de TD : 2.2

Sujets : EDHEC 2007 Problème

✓ MÉTHODE 2.4 : Utiliser la formule de l'espérance totale.

Lorsqu'on utilise la formule de l'espérance totale, on précisera **toujours** le système complet d'événements qu'on utilise. Pour calculer les espérances conditionnelles, il s'agit en général de remarquer que la loi conditionnelle est une loi usuelle. Si ce n'est pas le cas, il faudra revenir à la définition.

Si l'existence de $E(X)$ a déjà été prouvée, on ne se souciera pas de problèmes de convergence. Sinon, il faudra le prouver rigoureusement.

Exercices de TD : 2.12, 2.13, 2.14, 2.15, 2.16, 2.18

EXERCICES DU CHAPITRE 2

EXERCICE 2.1 Soit $(u_n)_{n \geq 1}$ la suite définie pour $n \geq 1$ par $u_n = \frac{1}{n(n+1)}$.

F

1) Montrer que $\sum_{n \geq 1} u_n$ converge et que $\sum_{n=1}^{+\infty} u_n = 1$.

2) Soit X une variable aléatoire à valeurs dans \mathbf{N}^* telle que pour tout $n \geq 1$, $P(X = n) = u_n$.
 X admet-elle une espérance ? Une variance ? Mêmes questions pour la variable aléatoire \sqrt{X} .

EXERCICE 2.2 On lance simultanément deux dés, jusqu'à ce que les deux dés donnent le même nombre. On note X le nombre de lancers nécessaires. Déterminer la loi de X , son espérance et sa variance.

F

EXERCICE 2.3 On dispose de deux dés équilibrés : le dé A possède quatre faces rouges et deux faces noires, le dé B possède quatre faces noires et deux faces rouges.

PD

On lance une pièce de monnaie truquée, qui tombe sur pile avec probabilité $\frac{1}{3}$. Si la pièce tombe sur pile, alors on ne joue qu'avec le dé A, si la pièce de monnaie tombe sur face, on ne joue qu'avec le dé B.

On note R_i l'événement «le $i^{\text{ème}}$ lancer de dé donne une face rouge».

1) Calculer $P(R_1)$, $P(R_2)$, puis $P(R_1 \cap R_2)$. Les événements R_1 et R_2 sont-ils indépendants ?

2) On a obtenu «rouge» aux deux premiers lancers. Calculer la probabilité d'obtenir «rouge» au troisième.

3) On a obtenu «rouge» aux n premiers lancers. Calculer la probabilité qu'on joue avec le dé A.

EXERCICE 2.4 Absence de mémoire de la loi géométrique

PD

Soit X une variable aléatoire suivant la loi géométrique de paramètre $p \in]0, 1[$.

Montrer que pour $(m, n) \in (\mathbf{N}^*)^2$, $P_{[X > m]}(X > m + n) = P(X > n)$. Interpréter ce résultat.

EXERCICE 2.5 On dispose de n sacs numérotés de 1 à n . Chaque sac contient $n + 1$ jetons. Dans le $k^{\text{ème}}$ sac se trouvent k jetons gagnants, les autres étant perdants. Un joueur choisit au hasard un sac et y pioche un jeton.

PD

1) Quelle est la probabilité que le jeton tiré soit gagnant ?

2) Pour $k \in \llbracket 1, n \rrbracket$, sachant que le jeton tiré est gagnant, quelle est la probabilité que le tirage ait eu lieu dans le sac numéro k ?

EXERCICE 2.6 Soit X une variable aléatoire suivant la loi géométrique $\mathcal{G}(p)$, et soit $Y = \frac{1}{X}$. Y admet-elle une espérance ?

F

EXERCICE 2.7 Soit X une variable aléatoire suivant une loi de Poisson $\mathcal{P}(\lambda)$.

PD

Montrer que $E\left(\frac{1}{X+1}\right)$ existe et la calculer.

EXERCICE 2.8 Propriétés des variables indicatrices

AD

Soit (Ω, \mathcal{A}, P) un espace probabilisé.

1) Montrer que si A, B sont deux événements, alors

$$\mathbb{1}_{\bar{A}} = 1 - \mathbb{1}_A, \mathbb{1}_{A \cap B} = \mathbb{1}_A \mathbb{1}_B, \mathbb{1}_{A \cup B} = \mathbb{1}_A + \mathbb{1}_B - \mathbb{1}_A \mathbb{1}_B.$$

2) Démontrer que $E(\mathbb{1}_A) = P(A)$.

3) Soient $(A_k)_{1 \leq k \leq n}$ n événements distincts. Interpréter la variable $\sum_{k=1}^n \mathbb{1}_{A_k}$.

4) **Application** : un ascenseur dessert N étages. Au rez-de-chaussée, k personnes montent dans l'ascenseur et choisissent un étage, indépendamment des autres.

a) Pour $i \in \llbracket 1, N \rrbracket$, soit A_i l'événement «l'ascenseur s'arrête à l'étage numéro i ». Déterminer $P(A_i)$.

b) Soit X la variable aléatoire égale au nombre d'étages auxquels l'ascenseur s'est arrêté. Exprimer X en fonction des $\mathbb{1}_{A_i}$ et en déduire l'espérance de X .

EXERCICE 2.9 Inégalité de Cauchy-Schwarz pour l'espérance

AD

Soient X et Y deux variables aléatoires discrètes sur (Ω, \mathcal{A}, P) admettant une variance.

On souhaite prouver que $E(XY)^2 \leq E(X^2)E(Y^2)$.

1) Montrer que pour tous réels x et y , on a $|xy| \leq \frac{1}{2}(x^2 + y^2)$. (Indication : développer $(|x| - |y|)^2$)

- 2) Montrer que $E(X^2)$ et $E(Y^2)$ existent, puis que $E(XY)$ existe.
- 3) Montrer que $\forall t \in \mathbf{R}, E(X^2)t^2 + 2tE(XY) + E(Y^2) \geq 0$.
- 4) En distinguant les cas $E(X^2) = 0$ et $E(X^2) \neq 0$, prouver l'inégalité recherchée.

EXERCICE 2.10 On dispose d'une pièce qui tombe sur «pile» avec probabilité p . On la lance jusqu'à faire «face», et on note X le nombre de lancers nécessaires. Quelle est la probabilité que ce nombre de lancers soit pair ?

PD

EXERCICE 2.11 Soit X une variable aléatoire à valeurs dans \mathbf{N} admettant une espérance.

AD

- 1) Pour tout entier k , donner une relation liant $P(X = k)$, $P(X > k)$ et $P(X > k - 1)$.

$$\text{En déduire que pour tout } n \in \mathbf{N}^*, \sum_{k=0}^n kP(X = k) = \sum_{k=0}^{n-1} P(X > k) - nP(X > n).$$

- 2) Montrer que pour tout $n \geq 1$, $nP(X > n) \leq \sum_{k=n+1}^{+\infty} kP(X = k)$.

- 3) En déduire que $E(X) = \sum_{k=0}^{+\infty} P(X > k)$.

EXERCICE 2.12 Le concierge alcoolique

PD

Un concierge possède 10 clés sur son trousseau.

Lorsqu'il souhaite ouvrir une porte, il choisit une clé au hasard dans son trousseau jusqu'à obtenir la bonne. Soit X la variable aléatoire désignant le nombre de clés à essayer par le concierge pour ouvrir une porte.

- 1) Déterminer la loi de X si le concierge essaie les clés sans remise, puis la loi de X s'il les essaie avec remise.
- 2) Le concierge essaie les clés sans remise s'il est sobre et avec remise s'il est ivre. Le concierge est ivre un jour sur 3.
 - a) Montrer que X admet une espérance, et la déterminer.
 - b) Aujourd'hui, le concierge a eu besoin de 6 essais pour ouvrir sa porte. Quelle est la probabilité qu'il soit ivre ?
 - c) Même question avec 11 essais.

EXERCICE 2.13 On lance un dé équilibré jusqu'à obtenir un 6, et on note X le nombre de lancers nécessaires. S'il a fallu lancer k fois le dé, on place une boule blanche et k boules noires dans une urne, et on effectue des tirages avec remise jusqu'à obtenir la boule blanche. On note alors Y le nombre de tirages nécessaires.

AD

- 1) Déterminer la loi de X .
- 2) Déterminer la loi de Y sachant $[X = k]$.
- 3) Montrer que Y admet une espérance, et la déterminer.
- 4) Que se passe-t-il si les tirages ont lieu sans remise ?

EXERCICE 2.14 Une urne contient n boules portant des numéros deux à deux distincts.

AD

Un joueur effectue des tirages sans remise dans cette urne jusqu'à obtenir la boule portant le plus grand numéro. On note X_1 le nombre de tirages nécessaires.

S'il reste des boules dans l'urne, un second joueur effectue lui aussi des tirages sans remise jusqu'à obtenir la boule portant le plus grand numéro parmi celles restant dans l'urne. On note X_2 le nombre de tirages effectués par le seconde joueur (avec $X_2 = 0$ si le premier joueur a vidé l'urne).

- 1) Déterminer la loi de X_1 .
- 2) Pour $k \in \llbracket 1, n \rrbracket$, déterminer la loi de X_2 sachant que $[X_1 = k]$.
- 3) Montrer que X_2 admet une espérance et la calculer.

EXERCICE 2.15 On dispose d'une pièce qui tombe sur pile avec probabilité $p \in]0, 1[$.

AD

On la lance jusqu'à obtenir pour la deuxième fois pile, et on note X le nombre de lancers nécessaires.

S'il a fallu n lancers, on effectue une seconde série de n lancers et on note Y le nombre de faces obtenus lors de cette seconde série de lancers.

- 1) Déterminer la loi de X .
- 2) Montrer que Y admet une espérance et la calculer.

EXERCICE 2.16 (D'après oral ESCP 1999)

D

Au casino, un croupier mélange trois cartes : as de cœur, roi de cœur et valet de pique. Il les présente face cachée, et un joueur choisit l'une des cartes au hasard. Si c'est un cœur, il gagne deux euros pour le roi ou un euro pour l'as, et le jeu recommence. Si c'est le valet de pique, le jeu s'arrête.

On note N le nombre de cartes tirées par le joueur, et X la somme qu'il a gagnée à la fin de la partie.

- 1) Déterminer la loi de N .
- 2) Déterminer la loi de X sachant que $N = n$.
- 3) Quel prix minimum le casino doit-il faire payer les parties pour que ce jeu soit rentable ?

EXERCICE 2.17 Somme d'un nombre aléatoire de variables aléatoires (Un classique de l'oral)

Soit $(X_n)_{n \geq 1}$ une suite de variables aléatoires sur un espace probabilisable (Ω, \mathcal{A}) et soit N une variable aléatoire sur (Ω, \mathcal{A}) , à valeurs dans \mathbf{N}^* .

On définit une application X sur Ω par $X = \sum_{k=1}^N X_k$, c'est-à-dire que

$$\forall \omega \in \Omega, X(\omega) = \sum_{i=1}^{N(\omega)} X_i(\omega).$$

Montrer que X est une variable aléatoire sur (Ω, \mathcal{A}) .

EXERCICE 2.18 (QSP HEC ECE 2014)

On dispose de n urnes U_1, U_2, \dots, U_n , et on dispose 3 boules dans chaque urne.

Dans l'ensemble des $3n$ boules, une seule est bleue, les autres sont rouges.

Sachant que l'on a tiré sans remise deux boules rouges dans l'urne U_1 , quelle est la probabilité que la boule bleue se trouve dans l'urne U_2 ?

EXERCICE 2.19 (QSP HEC 2016)

Soit X une variable aléatoire discrète admettant des moments jusqu'à l'ordre 4 et telle que
$$\begin{cases} E(X) = \alpha \\ E(X^2) = E(X^4) = 1 \end{cases}$$

- 1) Montrer que $\alpha \in [-1, 1]$.
- 2) Déterminer la loi de X .

CORRECTION DES EXERCICES DU CHAPITRE 2

SOLUTION DE L'EXERCICE 2.1

- 1.a. Pour tout n , on a $\frac{1}{n} - \frac{1}{n+1} = \frac{n+1-n}{n(n+1)} = \frac{1}{n(n+1)}$.
Et donc pour $N \geq 1$, il vient

$$\sum_{k=1}^N u_k = \sum_{k=1}^N \left(\frac{1}{k} - \frac{1}{k+1} \right) = \sum_{k=1}^N \frac{1}{k} - \sum_{k=1}^N \frac{1}{k+1}.$$

Soit encore

$$\sum_{k=1}^N u_k = \sum_{k=1}^N \frac{1}{k} - \sum_{i=2}^{N+1} \frac{1}{i} = \frac{1}{1} - \frac{1}{N+1}.$$

Et donc $\lim_{N \rightarrow +\infty} \sum_{k=1}^N u_k = 1$. Ceci prouve donc que la série de terme général u_k converge et

que $\sum_{k=1}^{+\infty} u_k = 1$.

2. Par définition, X admet une espérance si et seulement si la série de terme général

$$kP(X = k) = \frac{k}{k(k+1)} = \frac{1}{k+1} \text{ converge absolument.}$$

Puisqu'il s'agit d'une série à termes positifs, la convergence est équivalente à la convergence absolue.

Et $\frac{1}{k+1} \underset{k \rightarrow +\infty}{\sim} \frac{1}{k}$. Puisque la série de terme général $\frac{1}{k}$ diverge, d'après le critère des équivalents pour les séries à termes positifs, la série de terme général $kP(X = k)$ diverge. Ainsi, X n'admet pas d'espérance.

De plus, X admet une variance si et seulement si elle admet un moment d'ordre 2. Mais si c'était le cas, elle admettrait également une espérance.

Ce n'est pas le cas, donc X n'admet pas de moment d'ordre 2, et donc pas de variance.

Par le théorème de transfert, \sqrt{X} admet une espérance si et seulement si $\sum_{n \geq 1} \sqrt{n}P(X = n)$ converge¹

$$\text{Or, } \sqrt{n}P(X = n) = \frac{\sqrt{n}}{n(n+1)} \underset{n \rightarrow +\infty}{\sim} \frac{\sqrt{n}}{n^2} = \frac{1}{n^{3/2}}.$$

On reconnaît alors le terme général d'une série de Riemann convergente, et donc par le critère de comparaison pour les séries à termes positifs, $\sum_n \sqrt{n}P(X = n)$ converge, de sorte

que \sqrt{X} admet une espérance.

En revanche, $(\sqrt{X})^2 = X$ n'admet pas d'espérance, donc \sqrt{X} n'admet pas de moment d'ordre 2, et donc pas de variance.

SOLUTION DE L'EXERCICE 2.2

Lors d'un lancer de deux dés, il y a $6^2 = 36$ issues possibles. Parmi celles-ci, 6 ont leurs deux dés identiques.

Donc à chaque lancer, la probabilité d'avoir deux dés identiques est $\frac{6}{36} = \frac{1}{6}$.

Les lancers étant indépendants, on répète donc l'expérience jusqu'à obtention d'un premier succès², et X désigne le nombre d'essais nécessaires.

On reconnaît donc une loi géométrique : $X \hookrightarrow \mathcal{G}\left(\frac{1}{6}\right)$.

$$\text{Et donc } E(X) = \frac{1}{\frac{1}{6}} = 6 \text{ et } V(X) = \frac{1 - \frac{1}{6}}{\left(\frac{1}{6}\right)^2} = 30.$$

SOLUTION DE L'EXERCICE 2.3

Notons A l'événement : «on joue avec le dé A» et B l'événement «on joue avec le dé B».

On a donc $P(A) = \frac{1}{3}$ et $P(B) = \frac{2}{3}$.

Ultra-classique

Cette décomposition intervient régulièrement dans des énoncés de concours. Parfois elle y est donnée, et parfois non, donc mieux vaut connaître cette relation par cœur.

Chgt d'indice

On a posé $i = k + 1$.

Convergence

Rappelons que par définition, une série converge si et seulement si la suite de ses sommes partielles converge, et que la somme de la série est alors la limite de la suite des sommes partielles.

Rédaction

Le critère des équivalents ne fonctionne que pour les séries à termes positifs, donc si on souhaite l'utiliser, on prendra soin de vérifier (et d'écrire explicitement) la positivité des séries concernées.

¹ Absolument, mais il s'agit d'une série à termes positifs.

² Ici le succès est «obtenir deux dés identiques».

1. Par la formule des probabilités totales appliquée avec le système complet d'événements $\{A, B\}$, on a

$$P(R_1) = P_A(R_1)P(A) + P_B(R_1)P(B) = \frac{2}{3} \frac{1}{3} + \frac{2}{3} \frac{1}{3} = \frac{4}{9}.$$

Le même calcul nous donne également $P(R_2) = \frac{4}{9}$.

Enfin, toujours par la formule des probabilités totales,

$$\begin{aligned} P(R_1 \cap R_2) &= P_A(R_1 \cap R_2)P(A) + P_B(R_1 \cap R_2)P(B) = P(A)P_A(R_1)P_A(R_2) + P(B)P_B(R_1)P_B(R_2) \\ &= \frac{1}{3} \frac{2}{3} \frac{2}{3} + \frac{2}{3} \frac{1}{3} \frac{1}{3} = \frac{6}{27}. \end{aligned}$$

On constate que $P(R_1 \cap R_2) \neq P(R_1)P(R_2)$, et donc R_1 et R_2 ne sont pas indépendants.

Bien que ceci soit plutôt surprenant au premier abord, il y a une explication simple : si on obtient une face rouge lors du premier lancer, il est plus probable que nous soyons en train de jouer avec le dé A .

Et donc il est plus probable d'obtenir encore une face rouge au second lancer.

Alors qu'au contraire, si le premier lancer a donné une face noire, il est plus probable qu'on joue avec B , et donc qu'on obtienne encore une face noire au second lancer.

Il en serait autrement si on choisissait (toujours au hasard) un nouveau dé à chaque lancer.

2. Nous cherchons $P_{R_1 \cap R_2}(R_3)$. Par définition, on a

$$P_{R_1 \cap R_2}(R_3) = \frac{P(R_1 \cap R_2 \cap R_3)}{P(R_1 \cap R_2)}$$

Or, nous savons déjà que $P(R_1 \cap R_2) = \frac{6}{27}$.

Comme à la question précédente, on prouve³ que $P(R_1 \cap R_2 \cap R_3) = \frac{1}{3} \left(\frac{2}{3}\right)^3 + \frac{2}{3} \left(\frac{1}{3}\right)^3 = \frac{10}{81}$.

On en déduit que $P_{R_1 \cap R_2}(R_3) = \frac{5}{9}$.

3. Par la formule de Bayes, on a

$$\begin{aligned} P_{R_1 \cap R_2 \cap \dots \cap R_n}(A) &= \frac{P_A(R_1 \cap \dots \cap R_n)P(A)}{P(R_1 \cap \dots \cap R_n)} = \frac{P_A(R_1 \cap \dots \cap R_n)P(A)}{P(A)P_A(R_1 \cap \dots \cap R_n) + P(B)P_B(R_1 \cap \dots \cap R_n)} \\ &= \frac{\frac{1}{3} \left(\frac{2}{3}\right)^n}{\frac{1}{3} \left(\frac{2}{3}\right)^n + \frac{2}{3} \left(\frac{1}{3}\right)^n} = \frac{\frac{2^n}{3^{n+1}}}{\frac{2^n}{3^{n+1}} + \frac{2^{n+2}}{3^{n+1}}} = \frac{2^n}{2^n + 2}. \end{aligned}$$

SOLUTION DE L'EXERCICE 2.4

Pour $\ell \in \mathbf{N}$, on a

$$\begin{aligned} P(X > \ell) &= \sum_{i=\ell+1}^{+\infty} P(X = i) = \sum_{i=\ell+1}^{+\infty} p(1-p)^{i-1} \\ &= p(1-p)^\ell \sum_{j=0}^{+\infty} (1-p)^j = p(1-p)^\ell \frac{1}{1-(1-p)} = (1-p)^\ell. \end{aligned}$$

Par définition de la probabilité conditionnelle, on a

$$P_{[X > n]}(X > n+k) = \frac{P([X > n] \cap [X > n+k])}{P(X > n)}.$$

Mais puisque $[X > n+k] \subset [X > n]$, on a $[X > n] \cap [X > n+k] = [X > n+k]$, et donc

$$P_{[X > n]}(X > n+k) = \frac{P(X > n+k)}{P(X > n)} = \frac{(1-p)^{n+k}}{(1-p)^n} = (1-p)^k = P(X > k).$$

Cette propriété s'interprète bien si on pense en termes de nombres d'essais avant un premier succès : elle traduit le fait que de savoir qu'on a déjà eu n échecs ne change pas la probabilité de n'avoir que des échecs lors des k répétitions suivantes.

SOLUTION DE L'EXERCICE 2.5

Indépendance

Les lancers sont indépendants, une fois le dé choisi. Autrement dit, R_1 et R_2 sont indépendants pour les probabilités P_A et P_B . Cela ne signifie pas pour autant qu'ils sont indépendants pour P , comme nous le verrons ci-dessous.

³ À l'aide de la formule des probabilités totales.

Intuition

Ce résultat était largement prévisible si on se rappelle qu'une loi géométrique correspond au nombre d'essais avant l'obtention d'un premier succès lors de répétitions d'épreuves de Bernoulli. Alors $[X > \ell]$ est réalisé si et seulement si les ℓ premiers essais ont été des échecs, ce qui se produit avec probabilité $(1-p)^\ell$.

1. Pour $k \in \llbracket 1, n \rrbracket$, notons S_k l'événement «le jeton est tiré dans le $k^{\text{ème}}$ sac» et A l'événement «le jeton tiré est gagnant».
Alors $\{S_1, \dots, S_k\}$ est un système complet d'événements, de sorte que, par la formule des probabilités totales,

$$P(A) = \sum_{k=1}^n P(S_k)P_{S_k}(A) = \sum_{k=1}^n \frac{1}{n} \frac{k}{n+1} = \frac{1}{n(n+1)} \sum_{k=1}^n k = \frac{1}{n(n+1)} \frac{n(n+1)}{2} = \frac{1}{2}.$$

2. Nous cherchons donc $P_A(S_k)$. Or nous connaissons $P_{S_k}(A)$, ce qui doit nous faire penser à la formule de Bayes. On a

$$P_A(S_k) = \frac{P_{S_k}(A)P(S_k)}{P(A)} = \frac{\frac{k}{n+1} \frac{1}{n}}{\frac{1}{2}} = \frac{2k}{n(n+1)}.$$

SOLUTION DE L'EXERCICE 2.6

On a $X \geq 1$, et donc $0 \leq \frac{1}{X} \leq 1$. Puisque la variable certaine égale à 1 admet une espérance, il en est de même de $\frac{1}{X}$.

Autre solution : d'après le théorème de transfert, $\frac{1}{X}$ admet une espérance si et seulement si la série de terme général $\frac{1}{n}P(X = n)$ converge.⁴

Or, $\frac{1}{n}P(X = n) = \frac{p(1-p)^{n-1}}{n}$ de sorte que

$$\forall n \geq 1, 0 \leq \frac{1}{n}P(X = n) \leq p(1-p)^{n-1}.$$

Puisque la série de terme général $p(1-p)^{n-1}$ converge, il en est de même de la série de terme général $\frac{1}{n}P(X = n)$. Et donc $\frac{1}{X}$ admet une espérance.

SOLUTION DE L'EXERCICE 2.7

Notons que $0 \leq \frac{1}{X+1} \leq 1$, et donc $\frac{1}{X+1}$ admet bien une espérance.

De plus, par le théorème de transfert, on a

$$\begin{aligned} E(X) &= \sum_{k=0}^{+\infty} \frac{1}{k+1} P(X = k) \\ &= \sum_{k=0}^{+\infty} \frac{1}{k+1} e^{-\lambda} \frac{\lambda^k}{k!} \\ &= \frac{e^{-\lambda}}{\lambda} \sum_{k=0}^{+\infty} \frac{\lambda^{k+1}}{(k+1)!} \\ &= \frac{e^{-\lambda}}{\lambda} \sum_{i=1}^{+\infty} \frac{\lambda^i}{i!} \\ &= \frac{e^{-\lambda}}{\lambda} \left(\sum_{i=0}^{+\infty} \frac{\lambda^i}{i!} - 1 \right) \\ &= \frac{e^{-\lambda}}{\lambda} (e^\lambda - 1) \\ &= \frac{1 - e^{-\lambda}}{\lambda}. \end{aligned}$$

SOLUTION DE L'EXERCICE 2.8

1. Il est clair que toutes les variables concernées ne peuvent prendre que les valeurs 0 ou 1, et donc il s'agit de variables indicatrices.
On a $1 - \mathbb{1}_A(\omega) = 1$ si et seulement si $\mathbb{1}_A(\omega) = 0$ et donc si et seulement si $\omega \in \bar{A}$. Donc $1 - \mathbb{1}_A$ est bien l'indicatrice de \bar{A} .
De même, $\mathbb{1}_A(\omega)\mathbb{1}_B(\omega)$ ne peut prendre la valeur 1 que si on a à la fois $\mathbb{1}_B(\omega) = 1$ et

Rédaction

Ne pas oublier de vérifier la positivité si l'on souhaite montrer l'existence d'une espérance par domination, le résultat n'étant pas valable pour des variables aléatoires négatives.

⁴ Absolument, mais puisqu'il s'agit d'une série à termes positifs, la convergence est équivalente à la convergence absolue.

Mieux

En déduire que $E\left(\frac{1}{X}\right) \leq 1$.

Remarque

Le fait d'avoir déjà mentionné l'existence d'une espérance nous garantit que cette série converge (absolument), il n'y aura donc pas à le vérifier, ni même à le mentionner.

Chgt d'indice

$i = k + 1$.

$\mathbb{1}_A(\omega) = 1$. C'est-à-dire si et seulement si $\omega \in A \cap B$.

On en déduit que $\mathbb{1}_A \mathbb{1}_B = \mathbb{1}_{A \cap B}$.

Si $\omega \in A$ et $\omega \in B$, alors $\mathbb{1}_A(\omega) = \mathbb{1}_B(\omega) = \mathbb{1}_{A \cap B}(\omega) = 1$.

Et donc $\mathbb{1}_A(\omega) + \mathbb{1}_B(\omega) - \mathbb{1}_{A \cap B}(\omega) = 2 - 1 = 1$.

Si $\omega \in A$ et $\omega \notin B$, alors $\mathbb{1}_A(\omega) = 1$ et $\mathbb{1}_B(\omega) = \mathbb{1}_{A \cap B}(\omega) = 0$, de sorte que

$$\mathbb{1}_A(\omega) + \mathbb{1}_B(\omega) - \mathbb{1}_{A \cap B}(\omega) = 1.$$

De même si $\omega \in B$ et $\omega \notin A$. Enfin, si $\omega \notin A$ et $\omega \notin B$, alors

$$\mathbb{1}_A(\omega) + \mathbb{1}_B(\omega) - \mathbb{1}_{A \cap B}(\omega) = 0 + 0 - 0 = 0.$$

Donc $\mathbb{1}_A(\omega) + \mathbb{1}_B(\omega) - \mathbb{1}_{A \cap B}(\omega)$ vaut 1 si et seulement si $\omega \in A \cup B$: c'est donc égal à $\mathbb{1}_{A \cup B}(\omega)$.

2. Par définition,

$$E(\mathbb{1}_A) = 1 \times P(\mathbb{1}_A = 1) + (-1) \times P(\mathbb{1}_A = -1) = 1 \times P(A) + 0 \times P(\bar{A}) = P(A).$$

3. $\sum_{k=1}^n A_k$ est la variable aléatoire qui compte le nombre d'événements réalisés parmi A_1, \dots, A_n .

4.a. L'ascenseur ne s'arrête pas à l'étage i si et seulement si toutes les personnes présentes à l'intérieur ont choisi un étage différent du $i^{\text{ème}}$.

Or la probabilité qu'une personne choisisse un autre étage que le $i^{\text{ème}}$ est $\frac{N-1}{N}$ et par indépendance des choix des différentes personnes, la probabilité que personne ne choisisse l'étage i est

$$P(\bar{A}_i) = \left(\frac{N-1}{N}\right)^k.$$

En passant à l'événement contraire $P(A_i) = 1 - \left(\frac{N-1}{N}\right)^k$.

4.b. D'après la question 3, $X = \sum_{i=1}^N \mathbb{1}_{A_i}$ de sorte que par linéarité de l'espérance,

$$E(X) = \sum_{i=1}^N E(\mathbb{1}_{A_i}) = \sum_{i=1}^N \left(1 - \left(\frac{N-1}{N}\right)^k\right) = N \left(1 - \left(\frac{N-1}{N}\right)^k\right).$$

SOLUTION DE L'EXERCICE 2.9

1. Pour tout $x, y \in \mathbf{R}$, on a $(|x| - |y|)^2 \geq 0$.

Or $(|x| - |y|)^2 = |x|^2 - 2|xy| + |y|^2 = x^2 - 2|xy| + y^2$.

Et donc

$$x^2 + y^2 - 2|xy| \geq 0 \Leftrightarrow |xy| \leq \frac{1}{2}(x^2 + y^2).$$

2. Puisque X et Y admettent une variance, elles admettent des moments d'ordre 2. Donc $E(X^2)$ et $E(Y^2)$ existent. Et alors $\frac{1}{2}(X^2 + Y^2)$ admet également une espérance.

Enfin, $0 \leq |XY| \leq \frac{1}{2}(X^2 + Y^2)$, on en déduit donc que XY admet une espérance.

3. Soit $t \in \mathbf{R}$. Nous savons que $(tX + Y)^2 = t^2X^2 + 2tXY + Y^2$, et les trois variables aléatoires X^2, XY et Y^2 admettent une espérance, donc il en est de même de leur somme.

Puisque la variable aléatoire $(tX + Y)^2$ est positive, on a alors

$$E((tX + Y)^2) \geq 0.$$

Par linéarité de l'espérance,

$$0 \leq E((tX + Y)^2) = E(t^2X^2 + 2tXY + Y^2) = t^2E(X^2) + 2tE(XY) + E(Y^2).$$

Détail

On a une somme de 0 et de 1, et il y a autant de 1 que d'événements réalisés.

Classique

Cette inégalité est à connaître et à savoir redémontrer si besoin.

4. Si $E(X^2) = 0$, alors la fonction affine $t \mapsto 2tE(XY) + E(Y^2)$ est toujours positive, et donc nécessairement son coefficient directeur est nul. Soit ici $E(XY) = 0$.
Mais puisque $E(X^2) \geq 0$ et $E(Y^2) \geq 0$, il est évident que

$$E(XY)^2 \leq E(X^2)E(Y^2).$$

Si $E(X^2) \neq 0$, alors $t \mapsto E(X^2)t^2 + 2tE(XY) + E(Y^2)$ est une fonction polynôme du second degré qui ne change pas de signe, puisque toujours positive. Donc son discriminant est négatif ou nul.

Or ce discriminant est $\Delta = 4E(XY)^2 - 4E(X^2)E(Y^2)$.

On en déduit que $E(XY)^2 \leq E(X^2)E(Y^2)$.

SOLUTION DE L'EXERCICE 2.10

Il est clair que X suit une loi géométrique de paramètre $q = 1 - p$.

Mais alors la probabilité que X soit pair est

$$\sum_{n=1}^{+\infty} P(X = 2n) = \sum_{n=1}^{+\infty} qp^{2n-1} = \frac{q}{p} \sum_{n=1}^{+\infty} (p^2)^n = \frac{q}{p} \left(\frac{1}{1-p^2} - 1 \right) = \frac{q}{p} \underbrace{\frac{p^2}{(1-p)(1+p)}}_{=q} = \frac{p}{1+p}.$$

SOLUTION DE L'EXERCICE 2.11

1. Puisque X ne prend que des valeurs entières, on a $[X > k - 1] = [X > k] \cup [X = k]$, et ces deux événements sont incompatibles, de sorte que

$$P(X > k - 1) = P(X > k) + P(X = k) \Leftrightarrow P(X = k) = P(X > k - 1) - P(X > k).$$

On en déduit que pour $n \in \mathbf{N}^*$,

$$\begin{aligned} \sum_{k=0}^n kP(X = k) &= \sum_{k=1}^n kP(X = k) = \sum_{k=1}^n k(P(X > k - 1) - P(X > k)) \\ &= \sum_{k=1}^n kP(X > k - 1) - \sum_{k=1}^n kP(X > k) \\ &= \sum_{i=0}^{n-1} (i+1)P(X > i) - \sum_{k=1}^n kP(X > k) \\ &= \sum_{i=1}^{n-1} iP(X > i) + \sum_{i=0}^{n-1} P(X > i) - \sum_{k=1}^{n-1} kP(X > k) - nP(X > n) \\ &= \sum_{k=0}^{n-1} P(X > k) - nP(X > n). \end{aligned}$$

2. On a $[X > n] = \bigcup_{k=n+1}^{+\infty} [X = k]$ et ces événements sont deux à deux incompatibles, de sorte que

$$P(X > n) = \sum_{k=n+1}^{+\infty} P(X = k).$$

Et donc $nP(X > n) = \sum_{k=n+1}^{+\infty} nP(X = k)$.

Or pour $k \geq n + 1$, on a $kP(X = k) \geq nP(X = k)$ et donc

$$0 \leq \underbrace{\sum_{k=n+1}^{+\infty} nP(X = k)}_{=nP(X > n)} \leq \sum_{k=n+1}^{+\infty} kP(X = k).$$

3. Puisque X admet une espérance, la série $\sum_k kP(X = k)$ converge, et donc $\sum_{k=n+1}^{+\infty} kP(X = k)$ est le reste d'une série convergente, qui tend donc vers 0 lorsque n tend vers $+\infty$.

Détails

Il est archi-connu qu'un polynôme du second degré qui possède deux racines réelles distinctes (donc de discriminant strictement positif) n'est pas du même signe entre ses racines qu'à l'extérieur.

Parité

Un entier positif p est pair si et seulement si $p = 2n$, avec $n \in \mathbf{N}$.

Classique

Cette relation sert régulièrement. S'il n'est pas nécessaire de l'apprendre par cœur, il faut la comprendre et savoir la retrouver rapidement.

Le terme correspondant à $k = 0$ est nul.

Chgt d'indice

$i = k - 1 \Leftrightarrow k = i + 1$

Par le théorème des gendarmes, on en déduit que $nP(X > n) \xrightarrow[n \rightarrow +\infty]{} 0$.

Et donc, en reprenant l'égalité de la question 1, la série de terme général $P(X > k)$ converge et

$$\sum_{k=0}^{+\infty} P(X > k) = \lim_{n \rightarrow +\infty} \sum_{k=0}^{n-1} P(X > k) = \lim_{n \rightarrow +\infty} \left(\sum_{k=0}^n kP(X = k) + nP(X > n) \right) = \lim_{n \rightarrow +\infty} \sum_{k=0}^n kP(X = k) = E(X).$$

SOLUTION DE L'EXERCICE 2.12

1. Si le concierge essaie les clés sans remise, alors X prend ses valeurs dans $\llbracket 1, 10 \rrbracket$. Pour $p \in \llbracket 1, 10 \rrbracket$, soit M_p l'événement «la $m^{\text{ème}}$ clé est mauvaise».

Alors $[X = k] = M_1 \cap M_2 \cap \dots \cap M_{k-1} \cap \overline{M_k}$.

Par la formule des probabilités composées,

$$\begin{aligned} P(X = k) &= P(M_1)P_{M_1}(M_2) \cdots P_{M_1 \cap \dots \cap M_{k-2}}(M_{k-1})P_{M_1 \cap \dots \cap M_{k-1}}(\overline{M_k}) \\ &= \frac{9}{10} \frac{8}{9} \cdots \frac{10 - (k-2) - 1}{10 - (k-2)} \frac{1}{10 - (k-1)} = \frac{1}{10}. \end{aligned}$$

Ainsi, X suit une loi uniforme⁵ sur $\llbracket 1, 10 \rrbracket$.

Dans le cas où le concierge est ivre, X suit une loi géométrique de paramètre $\frac{1}{10}$, puisque X compte le nombre d'essais nécessaires à l'obtention d'un succès lors de répétitions indépendantes d'une épreuve de Bernoulli de paramètre $\frac{1}{10}$.

Ainsi, $\forall k \in \mathbf{N}^*$,

$$P(X = k) = \frac{1}{10} \frac{9^{k-1}}{10^{k-1}} = \frac{9^{k-1}}{10^k}.$$

- 2.a. Notons I l'événement «le concierge est ivre» et S l'événement «le concierge est sobre», de sorte que $\{I, S\}$ est un système complet d'événements. Alors nous savons que la loi de X conditionnellement à S est la loi uniforme $\mathcal{U}(\llbracket 1, 10 \rrbracket)$ et que la loi de X conditionnellement à I est la loi géométrique $\mathcal{G}(\frac{1}{10})$.

Ainsi, $E(X|S) = \frac{11}{2}$ et $E(X|I) = 10$.

Par la formule de l'espérance totale⁶, X admet une espérance, et

$$E(X) = P(I)E(X|I) + P(S)E(X|S) = \frac{1}{3} \times 10 + \frac{2}{3} \frac{11}{2} = \frac{21}{3} = 7.$$

- 2.b. Nous cherchons à calculer $P_{[X=6]}(I)$. Par la formule de Bayes,

$$\begin{aligned} P_{[X=6]}(I) &= \frac{P_I(X = 6)P(I)}{P_I(X = 6)P(I) + P_S(X = 6)P(S)} \\ &= \frac{\frac{9^5}{10^6} \frac{1}{3}}{\frac{9^5}{10^6} \frac{1}{3} + \frac{1}{10} \frac{2}{3}} \\ &= \frac{9^5}{9^5 + 2 \times 10^5} \approx 0.228 \end{aligned}$$

Dans le cas où il a besoin de 11 essais, un calcul comme précédemment avec la formule de Bayes permettrait d'arriver à $P_{[X=11]}(I) = 1$, en notant que $P_S(X = 11) = 0$. Mais bien entendu, le bon sens⁷ nous permet de conclure sans calcul que le concierge est ivre.

SOLUTION DE L'EXERCICE 2.13

1. Puisque X compte le nombre d'essais nécessaires à l'obtention du premiers succès lors d'une répétition d'épreuves de Bernoulli indépendantes de paramètre $\frac{1}{6}$, $X \hookrightarrow \mathcal{G}\left(\frac{1}{6}\right)$.
2. La loi de Y conditionnellement à l'événement $[X = k]$ est une loi géométrique de paramètre $\frac{1}{k+1}$.
3. D'après la question précédente, $E(Y|X = k) = k + 1$. Et par la formule de l'espérance totale appliquée au système complet d'événements $\{[X = k], k \in \mathbf{N}^*\}$, Y admet une espérance si et seulement si

$$\sum_{k \geq 1} P(X = k)E(Y|X = k) = \sum_{k \geq 1} (k+1) \frac{5^{k-1}}{6^k}$$

Méthode

On cherche à calculer la probabilité d'une intersection, il faut penser à la formule des probabilités composées.

⁵ Surpris(e) ? Essayez de retrouver ce résultat par dénombrement, vous verrez qu'il était prévisible.

⁶ Il n'y a pas de problème de convergence, puisque le système complet d'événements est fini.

⁷ Argument à ne pas évoquer trop souvent dans vos copies ! Ce qui ne vous empêche pas de l'utiliser.

Rédaction

Il n'y a pas de manière abrégée de noter ceci, on fera toujours une phrase et on évitera les notations du type

$$Y \hookrightarrow_{[X=k]} \mathcal{G}\left(\frac{1}{k+1}\right)$$

qui n'est pas du tout une notation standard.

converge absolument. Mais il s'agit de la somme d'une série géométrique et d'une série géométrique dérivée, toutes deux de raison $\frac{5}{6}$ donc convergentes. Alors

$$E(Y) = \frac{1}{6} \sum_{k=1}^{+\infty} k \left(\frac{5}{6}\right)^{k-1} + \frac{1}{6} \sum_{k=0}^{+\infty} \left(\frac{5}{6}\right)^k = \frac{1}{6} (36 + 6) = 7.$$

4. Si les tirages ont lieu sans remise, alors la loi de Y conditionnellement à l'événement $[X = k]$ est⁸ une loi uniforme sur $\llbracket 1, k + 1 \rrbracket$. Et donc $E(Y|X = k) = \frac{k+2}{2}$, de sorte que, toujours par la formule de l'espérance totale,

$$\begin{aligned} E(Y) &= \sum_{k=1}^{+\infty} P(X = k)E(Y|X = k) \\ &= \sum_{k=1}^{+\infty} \frac{1}{6} \left(\frac{5}{6}\right)^{k-1} \frac{k+2}{2} \\ &= \frac{1}{12} \sum_{k=1}^{+\infty} k \left(\frac{5}{6}\right)^{k-1} + \frac{1}{6} \sum_{k=1}^{+\infty} \left(\frac{5}{6}\right)^{k-1} \\ &= \frac{1}{12} \sum_{k=1}^{+\infty} k \left(\frac{5}{6}\right)^{k-1} + \frac{1}{6} \sum_{i=1}^{+\infty} \left(\frac{5}{6}\right)^i \\ &= \frac{1}{12} \frac{1}{\left(1 - \frac{5}{6}\right)^2} + \frac{1}{6} \frac{1}{1 - \frac{5}{6}} \\ &= 3 + 1 = 4. \end{aligned}$$

⁸ Voir l'exercice précédent, dans le cas où les clés étaient tirées sans remise.

SOLUTION DE L'EXERCICE 2.14

1. Notons A_k l'événement «la boule portant le plus grand numéro est obtenue lors du $k^{\text{ème}}$ tirage».

On a alors $[X_1 = 1] = A_1$ et donc $P(X_1 = 1) = P(A_1) = \frac{1}{n}$.

Et pour $k \in \llbracket 2, n \rrbracket$, on a

$$[X_1 = k] = \overline{A_1} \cap \overline{A_2} \cap \dots \cap \overline{A_{k-1}} \cap A_k.$$

Par la formule des probabilités composées, il vient donc

$$\begin{aligned} P(X_1 = k) &= P(\overline{A_1})P_{\overline{A_1}}(\overline{A_2}) \dots P_{\overline{A_1} \cap \dots \cap \overline{A_{k-2}}}(\overline{A_{k-1}})P_{\overline{A_1} \cap \dots \cap \overline{A_{k-1}}}(A_k) \\ &= \frac{n-1}{n} \frac{n-2}{n-1} \dots \frac{n-(k-2)-1}{n-(k-2)} \frac{1}{n-(k-1)} = \frac{1}{n}. \end{aligned}$$

Ainsi, pour tout $k \in \llbracket 1, n \rrbracket$, $P(X_1 = k) = \frac{1}{n}$: X_1 suit la loi uniforme sur $\llbracket 1, n \rrbracket$.

2. Sachant que $[X_1 = k]$, il reste $n - k$ boules dans l'urne, portant des numéros deux à deux distincts.

Et alors, comme dans la question 1, la loi de X_2 sachant $[X_1 = k]$ est une loi uniforme sur $\llbracket 1, n - k \rrbracket$.

Notons que si $k = n$, alors il ne reste pas de boule dans l'urne, et donc X_2 est la variable certaine égale à 0.

3. On a pour tout $k \in \llbracket 1, n - 1 \rrbracket$, $E(X_2|[X_1 = k]) = \frac{1+n-k}{2}$ et $E(X_2|[X_1 = n]) = 0$.

Puisque $X_2(\Omega) = \llbracket 1, n - 1 \rrbracket$, c'est une variable finie, qui admet donc une espérance, et par la formule de l'espérance totale appliquée au système complet d'événements $\{[X_1 = k], k \in \llbracket 1, n \rrbracket\}$, il vient

$$\begin{aligned} E(X_2) &= \sum_{k=1}^n P(X_1 = k)E(X_2|[X_1 = k]) = \sum_{k=1}^{n-1} \frac{1}{n} \frac{1+n-k}{2} \\ &= \frac{1}{n} \sum_{k=1}^{n-1} \frac{n-1}{2} - \frac{1}{2n} \sum_{k=1}^{n-1} k = \frac{(n+1)(n-1)}{2n} - \frac{n(n-1)}{4n} \\ &= \frac{(n-1)(2n+2-n)}{4n} = \frac{(n-1)(n+2)}{4n}. \end{aligned}$$

Chgt d'indice

$i = k - 1$.

Convergence

On a reconnu des séries géométriques et géométriques dérivées donc convergentes.

Analogie

En probabilités, il peut être utile d'avoir en tête quelques exemples simples et de remarquer quand une situation ressemble à l'un de ces exemples. Ici, cela revient par exemple à tirer des boules sans remise dans une urne comportant une noire et $n - 1$ blanches, jusqu'à obtenir la boule noire. La numérotation des boules n'est pas vraiment importante (à part que les numéros sont deux à deux distincts), et donc inutile de se compliquer la vie avec des événements prenant en compte le numéro des boules (à part «le plus grand numéro» et son contraire «pas le plus grand numéro»).

Si on n'a pas mentionné l'existence de $E(X_2)$ précédemment, on peut le faire ici : la somme est finie, donc converge.

SOLUTION DE L'EXERCICE 2.15

1. Commençons par remarquer que la variable aléatoire X prend ses valeurs dans $\mathbf{N} \setminus \{0, 1\}$. Notons P_i l'événement «on a obtenu un pile lors du $i^{\text{ème}}$ lancer» et F_i l'événement contraire. Alors, pour $n \geq 2$, on a

$$[X = n] = \bigcup_{i=1}^{n-1} F_1 \cap \cdots \cap F_{i-1} \cap P_i \cap F_{i+1} \cap \cdots \cap F_{n-1} \cap P_n.$$

Ces événements sont deux à deux incompatibles, de sorte que :

$$\begin{aligned} P(X = n) &= \sum_{i=1}^{n-1} P(F_1 \cap \cdots \cap F_{i-1} \cap P_i \cap F_{i+1} \cap \cdots \cap F_{n-1} \cap P_n) \\ &= \sum_{i=1}^{n-1} P(F_1) \cdots P(F_{i-1})P(P_i)P(F_{i+1}) \cdots P(F_{n-1})P(P_n) \\ &= \sum_{i=1}^{n-1} p^2(1-p)^{n-2} = (n-1)p^2(1-p)^{n-2}. \end{aligned}$$

2. Sachant que $[X = n]$, Y suit une loi binomiale $\mathcal{B}(n, 1-p)$. Et donc $E(Y|[X = n]) = n(1-p)$. D'après la formule de l'espérance totale appliquée au système complet d'événements $\{[X = n], n \geq 2\}$, on a donc, sous réserve de convergence

$$\begin{aligned} E(Y) &= \sum_{n=2}^{+\infty} P(X = n)E(Y|[X = n]) \\ &= \sum_{n=2}^{+\infty} n(1-p)(n-1)p^2(1-p)^{n-2} \\ &= p^2(1-p) \sum_{n=2}^{+\infty} n(n-1)(1-p)^{n-2} \\ &= p^2(1-p) \frac{2}{(1-(1-p))^3} \\ &= p^2(1-p) \frac{2}{p^3} = \frac{2(1-p)}{p}. \end{aligned}$$

Détails

On doit nécessairement avoir un pile au $n^{\text{ème}}$ lancer, et un seul autre lors des $n-1$ lancers précédents.

Indépendance des lancers.

On a reconnu une série géométrique dérivée seconde convergente car de raison $1-p \in]-1, 1[$, donc $E(Y)$ existe.

SOLUTION DE L'EXERCICE 2.16

1. À chaque tour, la probabilité que le jeu s'arrête est de $1/3$, et les parties étant indépendantes, le nombre N de parties suit une loi géométrique $\mathcal{G}(\frac{1}{3})$.

$$\forall n \in \mathbf{N}^*, P(N = n) = \frac{1}{3} \left(\frac{2}{3}\right)^{n-1}.$$

2. Si $[N = n]$ est réalisé, c'est que le joueur a gagné les $n-1$ premières parties, et a perdu la dernière.

Notons R le nombre de rois obtenus lors de ces $n-1$ premières parties. La loi de R sachant $N = n$ est une loi binomiale $\mathcal{B}(n-1, 1/2)$.

De plus, le gain est donné par $Y = 2R + (n-1-R) = R + (n-1)$.

On a alors, pour $k \in \llbracket n-1, 2(n-1) \rrbracket$,

$$P_{[N=n]}(Y = k) = P_{[N=n]}(R + n - 1 = k) = P_{[N=n]}(R = k - n + 1) = \binom{n-1}{k-n+1} \frac{1}{2^{n-1}}.$$

3. Pour que le jeu soit rentable, il faut faire payer la partie au minimum $E(Y)$ euros. Or nous avons, pour tout $n \in \mathbf{N}^*$,

$$E(Y|N = n) = E(R|X = n) + (n-1) = \frac{n-1}{2} + n-1 = \frac{3(n-1)}{2}.$$

Précision

Le second paramètre est $1/2$ car nous savons que la carte tirée est soit l'as soit le roi, (et jamais le valet sinon la partie s'arrêterait), les deux étant équiprobables.

Support

Il est aisé de voir que le support de Y sachant $[N = n]$ est $\llbracket n-1, 2n-2 \rrbracket$: au pire on a n as, et au mieux on a n rois.

En appliquant la formule de l'espérance totale au système complet d'événement $\{[N = n], n \in \mathbf{N}^*\}$, il vient (sous réserve de convergence)

$$\begin{aligned} E(Y) &= \sum_{n=1}^{+\infty} E(Y|N = n)P(N = n) \\ &= \frac{3}{2} \sum_{n=1}^{+\infty} (n-1) \frac{1}{3} \left(\frac{2}{3}\right)^{n-1} \\ &= \frac{1}{2} \sum_{n=1}^{+\infty} n \left(\frac{2}{3}\right)^{n-1} - \frac{1}{2} \sum_{n=0}^{+\infty} \left(\frac{2}{3}\right)^n \\ &= \frac{1}{2} \frac{1}{\frac{1}{3^2}} - \frac{1}{2} \frac{1}{\frac{1}{3}} \\ &= \frac{9}{2} - \frac{3}{2} = 3 \end{aligned}$$

Convergence

On reconnaît des séries géométriques et géométriques dérivées absolument convergentes, donc $E(Y)$ existe.

SOLUTION DE L'EXERCICE 2.17

Il s'agit de revenir à la définition d'une variable aléatoire : X est une variable aléatoire si pour tout $x \in \mathbf{R}$, $[X \leq x] \in \mathcal{A}$.

Or, N étant une variable aléatoire à valeur dans \mathbf{N}^* , on a

$$\Omega = \bigcup_{n \in \mathbf{N}^*} [N = n].$$

Ainsi,

$$[X \leq x] = \Omega \cap [X \leq x] = \bigcup_{n \in \mathbf{N}^*} [X \leq x] \cap [N = n].$$

Mais si $N = n$, alors $X = \sum_{i=1}^n X_i$ est bien une variable aléatoire car somme d'un nombre fini de variables aléatoires.

En particulier, $\left[\sum_{i=1}^n X_i \leq x \right] \in \mathcal{A}$.

Et donc

$$[X \leq x] = \bigcup_{n \in \mathbf{N}^*} \underbrace{\left[\sum_{i=1}^n X_i \leq x \right]}_{\in \mathcal{A}} \cap \underbrace{[N = n]}_{\in \mathcal{A}} \in \mathcal{A}.$$

car une tribu est stable par intersections et par unions dénombrables.

SOLUTION DE L'EXERCICE 2.18

Notons B_i l'événement «la boule bleue est dans l'urne U_i .»

Notons R_1 (resp. R_2) l'événement «la première (resp. deuxième) boule tirée dans l'urne U_1 est rouge».

La probabilité demandée est alors $P_{R_1 \cap R_2}(B_2)$.

Par la formule de Bayes, on a

$$P_{R_1 \cap R_2}(B_2) = \frac{P_{B_2}(R_1 \cap R_2)P(B_2)}{P(R_1 \cap R_2)}.$$

Mais, sachant que la boule bleue est dans l'urne U_2 , l'urne U_1 ne contient que des boules rouges, et donc $P_{B_2}(R_1 \cap R_2) = 1$.

D'autre part, on a $P(B_2) = \frac{1}{n}$, puisque la boule bleue a autant de chances de se trouver dans chaque urne.

Reste donc à calculer $P(R_1 \cap R_2)$.

Pour cela, appliquons la formule des probabilités totales au système complet d'événements $\{B_1, B_2, \dots, B_n\}$:

$$P(R_1 \cap R_2) = \sum_{i=1}^n P(B_i)P_{B_i}(R_1 \cap R_2) = \frac{1}{n} \sum_{i=1}^n P_{B_i}(R_1 \cap R_2) = \frac{1}{n} (P_{B_1}(R_1 \cap R_2) + n - 1).$$

Méthode

Cette question est très rare dans les sujets «usuels», mais classique à l'oral et dans les sujets de parisiennes. Il s'agit donc de prouver que $[X \leq x]$ est bien un événement (i.e. un élément de la tribu \mathcal{A}), et pour cela, il faut utiliser les propriétés des tribus et le fait que les X_i et N sont des variables aléatoires.

Mais on a alors

$$P_{B_1}(R_1 \cap R_2) = \frac{P(B_1 \cap R_1 \cap R_2)}{P(B_1)} = \frac{P(B_1)P_{B_1}(R_1)P_{B_1 \cap R_1}(R_2)}{P(B_1)} = P_{B_1}(R_1)P_{B_1 \cap R_1}(R_2) = \frac{2}{3} \frac{1}{2} = \frac{1}{3}.$$

Et donc on en déduit que

$$P(R_1 \cap R_2) = \frac{1}{n} (P_{B_1}(R_1)P_{B_1 \cap R_1}(R_2) + n - 1) = \frac{1}{n} \left(\frac{2}{3} \frac{1}{2} + n - 1 \right) = 1 - \frac{3n - 2}{3n}.$$

Il vient alors

$$P_{R_1 \cap R_2}(B_2) = \frac{1}{n} \frac{3n}{3n - 2} = \frac{3}{3n - 2}.$$

SOLUTION DE L'EXERCICE 2.19

1. Nous savons que $V(X) \geq 0$, et par la formule de Huygens, $V(X) = E(X^2) - E(X)^2 = 1 - \alpha^2$.
Et donc $1 - \alpha^2 \geq 0 \Leftrightarrow \alpha \in [-1, 1]$.
2. La variance de X^2 est donnée par $V(X^2) = E(X^4) - E(X^2)^2 = 1 - 1 = 0$.
Or, une variable de variance nulle est nécessairement une variable certaine, qui est alors égale à son espérance : $X^2 = 1$ presque sûrement.
Ainsi, X ne prend que les valeurs 1 et -1 .
On a alors $\alpha = E(X) = P(X = 1) - P(X = -1) = P(X = 1) - (1 - P(X = 1)) = 2P(X = 1) - 1$.
Et donc $P(X = 1) = \frac{1 + \alpha}{2}$ et donc $P(X = -1) = \frac{1 - \alpha}{2}$.

Remarque

Nous venons de prouver que

$$P_{B_1}(R_1 \cap R_2) = P_{B_1}(R_1)P_{B_1 \cap R_1}(R_2).$$

Il s'agit en fait de la formule des probabilités composées, mais appliquée à la probabilité P_{B_1} et non à la probabilité P .

En effet, on a alors

$$(P_{R_1})_{B_1}(R_2) = P_{R_1 \cap B_1}(R_2),$$

ce qui est relativement intuitif : la probabilité sachant B_1 , sachant R_1 est la probabilité sachant $R_1 \cap B_1$.

RÉVISIONS ET COMPLÉMENTS D'ALGÈBRE LINÉAIRE

Cours	71
3.1 Espace vectoriel, sous-espace vectoriel	71
3.2 Familles, bases, dimension	73
3.3 Somme de sous-espaces vectoriels	77
3.4 Applications linéaires	81
Méthodes	87
Exercices	89
Correction des exercices	92

Dans tout ce chapitre, sans plus de précision, la lettre **K** désignera indifféremment **R** ou **C**. La grande majorité des résultats de ce chapitre ont été vus en première année, les preuves qui ne sont pas données ici figurent donc dans le cours de première année.

3.1 ESPACE VECTORIEL, SOUS-ESPACE VECTORIEL

3.1.1 Définition

Définition 3.1 – Un **K**-espace vectoriel est un ensemble E muni d'une addition $+ : E \times E \rightarrow E$ et d'une multiplication externe $\cdot : \mathbf{K} \times E \rightarrow E$ vérifiant les axiomes suivants :

- 1) $\forall (u, v) \in E^2, u + v = v + u$ (commutativité de l'addition)
- 2) $\forall (u, v, w) \in E^3, (u + v) + w = u + (v + w)$ (associativité de l'addition)
- 3) $\exists ! 0_E \in E : \forall u \in E, u + 0_E = 0_E + u = u$ (élément neutre)
- 4) $\forall u \in E, u + (-1) \cdot u = 0_E$
- 5) $\forall u \in E, 1 \cdot u = u$
- 6) $\forall (\alpha, \beta) \in \mathbf{K}^2, (\alpha\beta) \cdot u = \alpha \cdot (\beta \cdot u)$
- 7) $\forall (u, v) \in E^2, \forall \lambda \in \mathbf{K}, \lambda \cdot (u + v) = \lambda \cdot u + \lambda \cdot v$
- 8) $\forall (\lambda, \mu) \in \mathbf{K}^2, \forall u \in E, (\alpha + \beta) \cdot u = \alpha \cdot u + \beta \cdot u$

Les éléments de **K** sont appelés les scalaires, et les éléments de E sont appelés vecteurs. L'élément 0_E est appelé vecteur nul.

En pratique

Il n'y a jamais besoin de connaître par cœur ces huit axiomes.

Notation

Généralement, on note 0 au lieu de 0_E .
En général, il n'y a pas d'ambiguïté, mais on prendre bien garde au fait que, suivant le contexte, 0 peut désigner le nombre 0 , une application linéaire nulle, la matrice nulle, la fonction nulle, etc.

Remarque. Quand il n'y a pas d'ambiguïté, on note λu au lieu de $\lambda \cdot u$.

3.1.2 Espaces vectoriels usuels

Il a été prouvé en première année que les ensembles suivants, munis de leurs lois usuelles sont des espaces vectoriels sur **K**.

- \mathbf{K}^n , $\mathbf{K} = \mathbf{R}$ ou $\mathbf{K} = \mathbf{C}$
- $\mathcal{M}_{n,p}(\mathbf{K})$: l'ensemble des matrices à n lignes et p colonnes à coefficients dans **K**, $\mathbf{K} = \mathbf{R}$ ou $\mathbf{K} = \mathbf{C}$
- $\mathbf{K}_n[X]$ et $\mathbf{K}[X]$, $\mathbf{K} = \mathbf{R}$ ou $\mathbf{K} = \mathbf{C}$
- $\mathbf{R}^{\mathbf{N}}$ l'ensemble des suites réelles, $\mathbf{K} = \mathbf{R}$
- si I est un intervalle de **R**, $\mathcal{F}(I, \mathbf{R})$ l'ensemble des fonctions de I dans **R**, $\mathbf{K} = \mathbf{R}$.

- $\mathcal{C}^0(I, \mathbf{R})$ l'ensemble des fonctions continues de I dans \mathbf{R} , et plus généralement, $\mathcal{C}^n(I, \mathbf{R})$ (respectivement $\mathcal{C}^\infty(I, \mathbf{R})$) l'ensemble des fonctions de classe \mathcal{C}^n (resp. de classe \mathcal{C}^∞) de I dans \mathbf{R} , $\mathbf{K} = \mathbf{R}$

3.1.3 Sous-espaces vectoriels

Définition 3.2 – Soit E un \mathbf{K} -espace vectoriel, et F un sous-ensemble de E . On dit que F est un sous-espace vectoriel de E si :

- F est non vide
- $\forall (x, y) \in F, x + y \in F$
- $\forall \lambda \in \mathbf{K}, \forall x \in F, \lambda \cdot x \in F$.

Méthode

En général, pour prouver que F est non vide, on prouve que $0_E \in F$.

Proposition 3.3 : Soit F une partie non vide d'un espace vectoriel E . Alors F est un sous-espace vectoriel de E si et seulement si $\forall (x, y) \in F^2, \forall \lambda \in \mathbf{K}, \lambda x + y \in F$.

Exemples 3.4

- $\{0_E\}$ et E sont deux sous-espaces vectoriels de E .
- $\mathbf{R}_n[X]$ est un sous-espace vectoriel de $\mathbf{R}[X]$.
- $\mathcal{C}^\infty(\mathbf{R}, \mathbf{R})$ est un sous-espace vectoriel de $\mathcal{C}^0(\mathbf{R}, \mathbf{R})$.
- $F = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathcal{M}_2(\mathbf{R}) : a + d = 0 \right\}$ est un sous-espace vectoriel de $\mathcal{M}_2(\mathbf{R})$.

En effet, F est non vide car la matrice nulle est dans F , et si $M = \begin{pmatrix} a_1 & b_1 \\ c_1 & d_1 \end{pmatrix}$ et

$N = \begin{pmatrix} a_2 & b_2 \\ c_2 & d_2 \end{pmatrix}$ sont dans F et $\lambda \in \mathbf{R}$, alors $\lambda M + N = \begin{pmatrix} \lambda a_1 + a_2 & \lambda b_1 + b_2 \\ \lambda c_1 + c_2 & \lambda d_1 + d_2 \end{pmatrix}$ et

$$(\lambda a_1 + a_2) + (\lambda c_1 + c_2) = \lambda(a_1 + c_1) + (a_2 + c_2) = 0.$$

et donc $\lambda M + N \in F$.

Alternative

Puisque nous connaissons les applications linéaires, on pourrait aussi remarquer que $\begin{pmatrix} a & b \\ c & d \end{pmatrix} \mapsto a + d$ est une application linéaire de $\mathcal{M}_2(\mathbf{R})$ dans \mathbf{R} . Par conséquent, son noyau est un sous-espace vectoriel de $\mathcal{M}_2(\mathbf{R})$. Mais ce noyau n'est autre que F .

Proposition 3.5 : Soit E un espace vectoriel, F et G deux sous-espaces vectoriels de E . Alors $F \cap G$ est un sous-espace vectoriel de E .

Danger !

$F \cup G$ est en revanche rarement un sous-espace vectoriel de E .

3.1.4 Combinaisons linéaires, sous-espace vectoriel engendré par une partie

Soit E un \mathbf{K} -espace vectoriel.

Définition 3.6 – Soient $(x_1, \dots, x_n) \in E^n$. Un vecteur $x \in E$ est combinaison linéaire de (x_1, \dots, x_n) s'il existe des scalaires $(\lambda_1, \dots, \lambda_n) \in \mathbf{K}^n$ tels que

$$x = \lambda_1 x_1 + \dots + \lambda_n x_n.$$

Définition 3.7 – Soient $(x_1, \dots, x_n) \in E^n$. On appelle sous-espace vectoriel engendré par la famille (x_1, \dots, x_n) , et on note $\text{Vect}(x_1, \dots, x_n)$ l'ensemble des combinaisons linéaires de (x_1, \dots, x_n) .

Autrement dit, $\text{Vect}(x_1, \dots, x_n) = \{\lambda_1 x_1 + \dots + \lambda_n x_n, (\lambda_1, \dots, \lambda_n) \in \mathbf{K}^n\}$.

Exemple 3.8

Dans \mathbf{R}^3 , soit $x_1 = (1, 0, 2)$ et $x_2 = (-1, 1, 0)$. Alors

$$\begin{aligned} \text{Vect}(x_1, x_2) &= \{\lambda_1 x_1 + \lambda_2 x_2, (\lambda_1, \lambda_2) \in \mathbf{R}^2\} \\ &= \{\lambda_1(1, 0, 2) + \lambda_2(-1, 1, 0), (\lambda_1, \lambda_2) \in \mathbf{R}^2\} \\ &= \{(\lambda_1 - \lambda_2, \lambda_2, 2\lambda_1), (\lambda_1, \lambda_2) \in \mathbf{R}^2\}. \end{aligned}$$

Proposition 3.9 : Soit $(x_1, \dots, x_n) \in E^n$. Alors :

- $\text{Vect}(x_1, \dots, x_n)$ est un sous-espace vectoriel de E
- si F est un sous-espace vectoriel de E contenant x_1, \dots, x_n , alors

$$\text{Vect}(x_1, \dots, x_n) \subset F.$$

Remarque

Cela signifie que $\text{Vect}(x_1, \dots, x_n)$ est le plus petit (au sens de l'inclusion) sous-espace vectoriel contenant à la fois x_1, x_2, \dots, x_n .

Démonstration. Prouvons uniquement le second point, et soit F contenant tous les x_i , $i \in \llbracket 1, n \rrbracket$. Soit $x \in \text{Vect}(x_1, \dots, x_n)$. Par définition, il existe des scalaires $\lambda_1, \dots, \lambda_n$ tels que

$$x = \lambda_1 x_1 + \dots + \lambda_n x_n.$$

Mais puisque $x_1, \dots, x_n \in F$, $x \in F$. Par conséquent, $\text{Vect}(x_1, \dots, x_n) \subset F$. □

3.2 FAMILLES, BASES, DIMENSION

3.2.1 Familles libres, familles liées

Définition 3.10 – Soit E un espace vectoriel, $(x_1, \dots, x_n) \in E^n$. La famille (x_1, \dots, x_n) est dite libre si

$$\forall (\lambda_1, \dots, \lambda_n) \in \mathbf{K}^n, \sum_{i=1}^n \lambda_i x_i = 0_E \Rightarrow \lambda_1 = \lambda_2 = \dots = \lambda_n = 0.$$

Remarque. Une famille de deux vecteurs est libre si et seulement si les deux vecteurs ne sont pas colinéaires¹, mais il n'existe pas de généralisation à des familles de plus de 2 vecteurs.

¹ Pour u, v non nuls, cela signifie tout simplement qu'il n'existe pas de scalaire $\lambda \in \mathbf{K}^*$ tel que $u = \lambda v$.

Exemple 3.11

Dans \mathbf{R}^3 , prouvons que la famille $((1, 1, -1); (-2, -1, 4); (3, 3, -2))$ est libre. Soient $(\lambda_1, \lambda_2, \lambda_3) \in \mathbf{R}^3$ tels que

$$\lambda_1(1, 1, -1) + \lambda_2(-2, -1, 4) + \lambda_3(3, 3, -2) = (0, 0, 0).$$

Ceci équivaut à $(\lambda_1 - 2\lambda_2 + 3\lambda_3, \lambda_1 - \lambda_2 + 3\lambda_3, -\lambda_1 + 4\lambda_2 - 2\lambda_3) = (0, 0, 0)$ et donc à

$$\begin{cases} \lambda_1 - 2\lambda_2 + 3\lambda_3 = 0 \\ \lambda_1 - \lambda_2 + 3\lambda_3 = 0 \\ -\lambda_1 + 4\lambda_2 - 2\lambda_3 = 0 \end{cases} \begin{matrix} \longleftrightarrow \\ \xleftrightarrow{L_2 \leftarrow L_2 - L_1} \\ \xleftrightarrow{L_3 \leftarrow L_3 + L_1} \end{matrix} \begin{cases} \lambda_1 + 3\lambda_3 = 0 \\ \lambda_2 = 0 \\ \lambda_3 = 0 \end{cases} \longleftrightarrow \begin{cases} \lambda_1 = 0 \\ \lambda_2 = 0 \\ \lambda_3 = 0 \end{cases}$$

Ainsi, $\lambda_1 = \lambda_2 = \lambda_3 = 0$, et la famille est donc libre.

Définition 3.12 – Une famille (x_1, \dots, x_n) qui n'est pas libre est dite liée. Cela signifie qu'il existe des scalaires $\lambda_1, \dots, \lambda_n$ **non tous nuls** tels que

$$\lambda_1 x_1 + \dots + \lambda_n x_n = 0_E.$$

Terminologie

Ne pas confondre non tous nuls avec tous non nuls : non tous nuls signifie que l'un au moins des λ_i n'est pas nul, alors que tous non nuls signifie que tous les λ_i sont non nuls.

$$\sum_{\substack{1 \leq j \leq n \\ j \neq i}} \lambda_j x_j$$

Remarque. Une telle relation existe, avec $\lambda_i \neq 0$ si et seulement si $x_i = -\frac{\sum_{\substack{1 \leq j \leq n \\ j \neq i}} \lambda_j x_j}{\lambda_i}$.

Autrement dit, une famille est liée si et seulement si l'un des vecteurs est combinaison linéaire des autres.

Proposition 3.13 : Si une sous-famille de $\mathcal{F} = (x_1, \dots, x_n)$ est liée, alors nécessairement \mathcal{F} est liée.

Démonstration. Quitte à renuméroter les vecteurs de \mathcal{F} , on peut supposer que la sous-famille (x_1, \dots, x_p) , $p \leq n$ est liée. Alors, par définition, il existe des scalaires $\lambda_1, \dots, \lambda_p$ non tous nuls tels que $\lambda_1 x_1 + \dots + \lambda_p x_p = 0_E$.

Posons alors $\lambda_{p+1} = \dots = \lambda_n = 0$. Alors les scalaires $\lambda_1, \dots, \lambda_n$ sont non tous nuls, et

$$\lambda_1 x_1 + \dots + \lambda_n x_n = \lambda_1 x_1 + \dots + \lambda_p x_p = 0_E$$

de sorte que \mathcal{F} est liée. \square

Corollaire 3.14 – • Une famille contenant le vecteur nul est liée.
• Une sous-famille d'une famille libre est libre.

Démonstration. • La famille formée du seul vecteur nul est liée, par exemple car $1 \cdot 0_E = 0_E$. Par conséquent, toute famille contenant le vecteur nul est liée.

• Soit (x_1, \dots, x_n) une famille libre et (x_1, \dots, x_p) une sous-famille, avec $p \neq n$.

Si la famille (x_1, \dots, x_p) était liée, alors ce serait également le cas de (x_1, \dots, x_n) , ce qui est absurde. \square

3.2.2 Familles génératrices

Définition 3.15 – Une famille (x_1, \dots, x_n) est dite génératrice de E si

$$E = \text{Vect}(x_1, \dots, x_n).$$

De manière équivalente, (x_1, \dots, x_n) est génératrice si tout vecteur de E peut s'écrire comme combinaison linéaire de (x_1, \dots, x_n) .

Exemple 3.16

Montrons que la famille $((1, 2); (2, -1))$ est génératrice de \mathbb{C}^2 .

Soit $(x, y) \in \mathbb{C}^2$. Nous cherchons à prouver qu'il existe $(\lambda_1, \lambda_2) \in \mathbb{C}^2$ tels que

$$(x, y) = \lambda_1(1, 2) + \lambda_2(2, -1) = (\lambda_1 + 2\lambda_2, 2\lambda_1 - \lambda_2)$$

c'est-à-dire à prouver qu'il existe des² solutions au système

$$\begin{cases} \lambda_1 + 2\lambda_2 = x \\ 2\lambda_1 - \lambda_2 = y \end{cases} \xleftrightarrow{L_2 \leftarrow L_2 - 2L_1} \begin{cases} \lambda_1 + 2\lambda_2 = x \\ -5\lambda_2 = y - 2x \end{cases} \longleftrightarrow \begin{cases} \lambda_1 = \frac{9x-2y}{5} \\ \lambda_2 = \frac{2x-y}{5} \end{cases}$$

Ainsi, tout vecteur de \mathbb{C}^2 est combinaison linéaire de $(1, 2)$ et $(2, -1)$: il s'agit d'une famille génératrice de \mathbb{C}^2 .

Proposition 3.17 : Une famille contenant une famille génératrice de E est génératrice de E .

Méthode

Pour prouver qu'une famille (x_1, \dots, x_n) est génératrice, il faut donc prouver que tout vecteur de E est combinaison linéaire de x_1, \dots, x_n . On a toutefois des méthodes plus efficaces lorsqu'on connaît la dimension de E .

² Peu importe qu'il y ait une ou plusieurs solutions, l'important étant qu'il en existe.

Démonstration. Soit $\mathcal{F} = (x_1, \dots, x_n)$ une famille de vecteurs de E , et (quitte à renuméroter les vecteurs) soit $\mathcal{F}' = (x_1, \dots, x_p)$, $p \leq n$ une sous-famille de \mathcal{F} qui soit génératrice de E . Puisque \mathcal{F}' est génératrice de E , pour tout vecteur $x \in E$, il existe des scalaires $\lambda_1, \dots, \lambda_p$ tels que

$$x = \lambda_1 x_1 + \dots + \lambda_p x_p = \lambda_1 x_1 + \dots + \lambda_p x_p + 0x_{p+1} + \dots + 0x_n.$$

Ainsi, x est combinaison linéaire de vecteurs de \mathcal{F} : \mathcal{F} est génératrice de E . \square

Remarque. Si (x_1, \dots, x_n) est génératrice de E , alors tout vecteur de E peut s'écrire comme combinaison linéaire de x_1, \dots, x_n , mais pas nécessairement de manière unique : il peut y avoir plusieurs combinaisons linéaires égales au même vecteur.

Exemple 3.18

$(1, 2), (2, -1)$ est génératrice de \mathbb{C}^2 , donc il en est de même de $((1, 2), (2, -1), (3, 1))$.
Alors $(3, 1) = (1, 2) + (2, -1) + 0(3, 1) = 0(1, 2) + 0(2, -1) + 1(3, 1)$.

3.2.3 Bases, dimension

Définition 3.19 – Une famille de vecteurs de E qui est à la fois libre et génératrice est appelée une base de E .

Remarque. Notons qu'en vertu de la proposition 3.13, une sous-famille d'une base est encore une famille libre.

Et de même, par la proposition 3.17, si on ajoute des vecteurs à une base, on obtient encore une famille génératrice.

Proposition 3.20 : Une famille (e_1, \dots, e_n) est une base de E si et seulement si tout vecteur de E peut s'écrire de manière unique comme combinaison linéaire de (e_1, \dots, e_n) . C'est à dire que

$$\forall x \in E, \exists!(\lambda_1, \dots, \lambda_n) \in \mathbf{K}^n : x = \sum_{i=1}^n \lambda_i e_i.$$

Démonstration. \Rightarrow Soit $\mathcal{B} = (e_1, \dots, e_n)$ une base de E . Alors \mathcal{B} est génératrice de E , donc si $x \in E$, $\exists(\lambda_1, \dots, \lambda_n) \in \mathbf{K}^n$ tels que $x = \lambda_1 e_1 + \dots + \lambda_n e_n$.

Supposons qu'il existe deux telles écritures

$$x = \lambda_1 e_1 + \dots + \lambda_n e_n = \mu_1 e_1 + \dots + \mu_n e_n.$$

Alors $0_E = x - x = (\lambda_1 - \mu_1)e_1 + \dots + (\lambda_n - \mu_n)e_n$.

Mais puisque \mathcal{B} est libre, nécessairement

$$\lambda_1 - \mu_1 = \dots = \lambda_n - \mu_n = 0$$

et donc $\lambda_1 = \mu_1, \dots, \lambda_n = \mu_n$: l'écriture est unique.

\Leftarrow Si tout vecteur de E s'écrit de manière unique comme combinaison linéaire des e_i , en particulier tout vecteur est combinaison linéaire des vecteurs de \mathcal{B} , donc \mathcal{B} est génératrice. De plus, si il existe des scalaires $\lambda_1, \dots, \lambda_n$ tels que

$$0_E = \lambda_1 e_1 + \dots + \lambda_n e_n = 0e_1 + \dots + 0e_n$$

par unicité de l'écriture de 0_E comme combinaison linéaire des e_i , il vient $\lambda_1 = \dots = \lambda_n = 0$. Donc \mathcal{B} est libre : c'est une base de E . \square

Remarque

Le côté générateur apporte l'existence d'une écriture de x comme combinaison linéaire des e_i , et c'est le côté libre qui apporte l'unicité de cette écriture.

Définition 3.21 – Si $\mathcal{B} = (e_1, \dots, e_n)$ est une base de E et si $x \in E$, alors il existe un unique n -uplet $(\lambda_1, \dots, \lambda_n)$ de scalaires tel que $x = \sum_{i=1}^n \lambda_i e_i$. On dit que $(\lambda_1, \dots, \lambda_n)$ sont les coordonnées de x dans la base \mathcal{B} , et on note

$$x_{\mathcal{B}} = \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix} \in \mathcal{M}_{n,1}(\mathbf{R}).$$

$x_{\mathcal{B}}$ est appelé vecteur colonne des coordonnées de x dans la base \mathcal{B} .

Pour $E = \mathbf{R}^2$ ou \mathbf{R}^3 muni de sa base canonique, on retrouve les coordonnées d'un point dans le plan ou dans l'espace, telles que vous les manipulez depuis que vous êtes petits.

Définition 3.22 – On dit qu'un espace vectoriel E est de dimension finie s'il existe une famille génératrice finie de E .

Théorème 3.23 (de la dimension) : Soit E un espace vectoriel de dimension finie. Alors :

- il existe une base de E
- toutes les bases de E ont le même cardinal.

Le cardinal d'une base est alors appelé dimension de E et noté $\dim E$.

Remarques. • Par convention $\dim\{0_E\} = 0$.

- Un sous-espace vectoriel de dimension 1 est appelé une droite vectorielle (ou plus simplement une droite), et un sous-espace de dimension 2 est appelé un plan vectoriel.

Exemples 3.24 Bases canoniques

Les exemples suivants sont classiques et à connaître.

- Une base de \mathbf{K}^n est $((1, 0, \dots, 0), (0, 1, \dots, 0), \dots, (0, \dots, 0, 1))$, appelée base canonique de \mathbf{K}^n . Par conséquent, $\dim \mathbf{K}^n = n$.
- Une base de $\mathbf{K}_n[X]$ est $(1, X, \dots, X^n)$, appelée base canonique. Par conséquent, $\dim \mathbf{K}_n[X] = n + 1$.
- La base canonique de $\mathcal{M}_{n,p}(\mathbf{K})$ est $\{E_{i,j}, 1 \leq i \leq n, 1 \leq j \leq p\}$. Par conséquent, $\dim \mathcal{M}_{n,p}(\mathbf{K}) = n \times p$.

Théorème 3.25 (de la base incomplète) : Soit E un espace vectoriel de dimension finie, soit \mathcal{L} une famille libre de E et \mathcal{G} une famille génératrice de E . Alors on peut compléter \mathcal{L} à l'aide de vecteurs de \mathcal{G} pour former une base de E .

Corollaire 3.26 – Soit E un espace vectoriel de dimension finie, et F un sous-espace vectoriel de E . Alors F est de dimension finie et $\dim F \leq \dim E$. De plus, $F = E$ si et seulement si $\dim F = \dim E$.

Proposition 3.27 : Soit E un espace vectoriel de dimension finie n . Alors :

- toute famille libre de E est de cardinal inférieur ou égal à n
- une famille libre de cardinal n est une base de E
- toute famille génératrice de E est de cardinal supérieur ou égal à n
- une famille génératrice de E de cardinal n est une base de E .

Convention

Il n'y a pas de définition mathématique au terme «canonique», qui est un synonyme de «standard». Donc ces trois espaces ont des bases dites «canoniques», mais on ne parlera pas de base canonique pour un autre espace vectoriel.

Rappel

La matrice $E_{i,j}$ est la matrice de $\mathcal{M}_{n,p}(\mathbf{K})$ dont tous les coefficients sont nuls, sauf situé sur la $i^{\text{ème}}$ ligne et $j^{\text{ème}}$ colonne qui vaut 1.

Méthode

Il est souvent pénible de montrer qu'une famille est génératrice, alors que la liberté est souvent plus simple à montrer. Aussi, autant que possible, lorsqu'il s'agit de montrer qu'une famille est une base d'un espace vectoriel E , on montrera qu'elle est libre et contient $\dim E$ vecteurs. Ceci nécessite bien entendu de connaître la dimension de E .

3.3 SOMME DE SOUS-ESPACES VECTORIELS

Dans cette partie, on note E un espace vectoriel, F, G, F_1, \dots, F_n des sous-espaces vectoriels de E .

3.3.1 Somme de deux sous-espaces vectoriels

Définition 3.28 – On appelle somme de F et G , et on note $F + G$ l'ensemble

$$F + G = \{x + y, x \in F, y \in G\}.$$

Proposition 3.29 : $F + G$ est un sous-espace vectoriel de E , et si H est un sous-espace vectoriel de E tel que $F \subset H, G \subset H$, alors nécessairement $F + G \subset H$.

Démonstration. Prouvons uniquement le second point : $\forall x \in F, \forall y \in G, x \in H$ (car $F \subset H$) et $y \in H$ (car $G \subset H$), donc $x + y \in H$.

Ainsi, $\forall z = x + y \in F + G, z \in H$, ceci prouve que $F + G \subset H$. \square

Proposition 3.30 : Si F et G sont de dimension finie, que (x_1, \dots, x_n) est une famille génératrice de F et que (y_1, \dots, y_p) est une famille génératrice de G , alors $(x_1, \dots, x_n, y_1, \dots, y_p)$ est une famille génératrice de $F + G$.

Démonstration. Soit $z \in F + G$ et soient $x \in F, y \in G$ tels que $z = x + y$.

Puisque (x_1, \dots, x_n) est une famille génératrice de F , il existe des scalaires $\lambda_1, \dots, \lambda_n$ tels

que $x = \sum_{i=1}^n \lambda_i x_i$. De même, il existe des scalaires (μ_1, \dots, μ_p) tels que $y = \sum_{j=1}^p \mu_j y_j$.

Alors $z = \sum_{i=1}^n \lambda_i x_i + \sum_{j=1}^p \mu_j y_j$, de sorte que z est bien combinaison linéaire de $(x_1, \dots, x_n, y_1, \dots, y_p)$.

Ceci étant vrai pour tout $z \in F + G$, il s'agit bien d'une famille génératrice de $F + G$. \square

Corollaire 3.31 – Si F et G sont de dimension finie, alors $F + G$ est encore de dimension finie et

$$\dim(F + G) \leq \dim F + \dim G.$$

Démonstration. Soit (x_1, \dots, x_n) une base³ de F et (y_1, \dots, y_p) une base de G .

Alors $(x_1, \dots, x_n, y_1, \dots, y_p)$ est une famille génératrice de $F + G$, qui est finie, de sorte que $F + G$ est de dimension finie.

Elle est donc⁴ de cardinal supérieur ou égal à $\dim(F + G)$.

Mais ce cardinal n'est autre que $n + p$, avec $n = \dim F$ et $p = \dim G$. Et donc

$$\dim(F + G) \leq \dim F + \dim G.$$

\square

3.3.2 Somme directe de deux sous-espaces vectoriels

Définition 3.32 – On dit que F et G sont en somme directe si

$$\forall x \in F, \forall y \in G, x + y = 0_E \Rightarrow x = y = 0_E.$$

Dans ce cas, on note $F \oplus G$ au lieu de $F + G$.

Autrement dit

$F + G$ est le plus petit (au sens de l'inclusion) sous-espace vectoriel de E contenant à la fois F et G .

³ En particulier, c'est une famille génératrice de F .

⁴ C'est le troisième point de la proposition 3.27.

Notation

$F \oplus G$ n'est pas un nouvel objet : c'est le sous-espace vectoriel $F + G$. Mais la notation $F \oplus G$ nous donne une information supplémentaire : F et G sont en somme directe.

Proposition 3.33 : Soit E de dimension finie, \mathcal{B}_F une base de F et \mathcal{B}_G une base de G . Alors il y a équivalence entre

- 1) F et G sont en somme directe
- 2) $F \cap G = \{0_E\}$
- 3) la concaténation de \mathcal{B}_F et \mathcal{B}_G est une base de $F + G$
- 4) $\dim F + \dim G = \dim(F + G)$.

Démonstration. Notons $\mathcal{B}_F = (e_1, \dots, e_n)$ et $\mathcal{B}_G = (f_1, \dots, f_p)$.

Nous allons prouver que 1) \Rightarrow 2) \Rightarrow 3) \Rightarrow 1), puis que 3) \Leftrightarrow 4).

1) \Rightarrow 2) : si $x \in F \cap G$, alors $-x \in G$, et $0_E = x + (-x)$, avec $x \in F$ et $-x \in G$, de sorte que $x = -x = 0_E$ (par définition de la somme directe).

Ainsi, $F \cap G \subset \{0_E\}$, et puisque $\{0_E\} \subset F \cap G$, alors $F \cap G = \{0_E\}$.

2) \Rightarrow 3) : Nous savons déjà que $(e_1, \dots, e_n, f_1, \dots, f_p)$ est génératrice de $F + G$. Prouvons qu'elle est libre : soit $(\lambda_1, \dots, \lambda_n, \mu_1, \dots, \mu_p) \in \mathbf{K}^{n+p}$ tel que

$$0 = \lambda_1 e_1 + \dots + \lambda_n e_n + \mu_1 f_1 + \dots + \mu_p f_p.$$

Posons $x = \sum_{i=1}^n \lambda_i e_i$ et $y = \sum_{j=1}^p \mu_j f_j$, de sorte que $x \in F$ et $y \in G$. De plus, $x = -y$, donc

$x \in G$, et donc $x \in F \cap G$.

Puisqu'on a supposé que $F \cap G = \{0_E\}$, nécessairement, $x = 0_E$. Mais alors, puisque (e_1, \dots, e_n) est une base de F , cela signifie que $\lambda_1 = \dots = \lambda_n = 0$.

On prouve de même que $y \in F \cap G$, et donc que $\mu_1 = \dots = \mu_p = 0$.

Ainsi, la famille $(e_1, \dots, e_n, f_1, \dots, f_p)$ est libre : c'est une base de $F + G$.

3) \Rightarrow 1) : soient $x \in F$ et $y \in G$ tels que $0_E = x + y$. Par hypothèse $(e_1, \dots, e_n, f_1, \dots, f_p)$ est une base de $F + G$.

Soient donc des scalaires $\lambda_1, \dots, \lambda_n, \mu_1, \dots, \mu_p$ tels que $x = \sum_{i=1}^n \lambda_i e_i$ et $y = \sum_{j=1}^p \mu_j f_j$.

Alors $0_E = \sum_{i=1}^n \lambda_i e_i + \sum_{j=1}^p \mu_j f_j$, et par liberté de la famille $(e_1, \dots, e_n, f_1, \dots, f_p)$ on a

$$\lambda_1 = \dots = \lambda_n = \mu_1 = \dots = \mu_p = 0_E$$

et donc $x = y = 0$: F et G sont en somme directe.

3) \Rightarrow 4) : évident par cardinalité des bases.

4) \Rightarrow 3) : $(e_1, \dots, e_n, f_1, \dots, f_p)$ est une famille génératrice de $F + G$, de cardinal $\dim F + \dim G = \dim(F + G)$: c'est donc une base de $F + G$. \square

Définition 3.34 – Si F et G sont en somme directe, une base de $F + G$ qui est obtenue par concaténation d'une base de F et d'une base de G est appelée base adaptée à la somme directe $F \oplus G$.

⚠ Dimension infinie

Dans un espace de dimension infinie (par exemple l'ensemble des polynômes ou l'ensemble des fonctions continues), les deux premiers points restent équivalents. Mais les deux derniers n'ont pas toujours un sens.

🚫 Danger !

Si nous venons de prouver que la concaténation de toute base de F et de toute base de G est une base de $F + G$, la réciproque est fautive ; il existe des bases de $F + G$ qui ne sont pas obtenues ainsi (et ne sont donc pas adaptées à la somme directe).

3.3.3 Sous-espaces supplémentaires

Définition 3.35 – Soit E un espace vectoriel, F et G deux sous-espaces vectoriels de E . On dit que F et G sont supplémentaires dans E si F et G sont en somme directe et que $F \oplus G = E$.

... sont supplémentaires. Notons qu'on a toujours $\{0_E\} \subset F \cap G$ et donc qu'il suffit de prouver que $F \cap G \subset \{0_E\}$, c'est-à-dire que si un vecteur x est à la fois dans F et dans G , alors il est nul.

Proposition 3.36 : Soit E un espace vectoriel de dimension finie. Deux sous-espaces vectoriels F et G sont supplémentaires dans E si et seulement si $\begin{cases} F \cap G = \{0_E\} \\ \dim F + \dim G = \dim E \end{cases}$

Démonstration. Si $E = F \oplus G$, alors nous avons déjà prouvé que $F \cap G = \{0_E\}$, et que $\dim(F + G) = \dim F + \dim G$. Puisque $E = F + G$, $\dim E = \dim F + \dim G$.

Inversement, si $F \cap G = \{0_E\}$ et $\dim F + \dim G = \dim E$, alors F et G sont en somme directe (car $F \cap G = \{0_E\}$), et donc $\dim(F + G) = \dim F + \dim G = \dim E$. Puisque $F \oplus G \subset E$, par égalité des dimensions, $F \oplus G = E$. \square

Exemple 3.37

Dans \mathbf{R}^3 , considérons $e_1 = (1, 1, -3)$, $e_2 = (-1, 0, 4)$ et $e_3 = (1, 4, 1)$, et posons $F = \text{Vect}(e_1, e_2)$, $G = \text{Vect}(e_3)$.

e_1 et e_2 ne sont pas colinéaires, donc (e_1, e_2) est libre : c'est donc une base de F , et $\dim F = 2$.

Trivialement, $\dim G = 1$.

Soit $x \in F \cap G$. Alors $\exists(\lambda, \mu) \in \mathbf{R}^2$ tels que $x = \lambda e_1 + \mu e_2$ et $\exists v \in \mathbf{R} : x = v e_3$.

Donc $(\lambda - \mu, \mu, -3\lambda + 4\mu) = (v, 4v, v)$

$$\begin{aligned} \iff \begin{cases} \lambda - \mu = v \\ \lambda = 4v \\ -3\lambda + 4\mu = v \end{cases} &\iff \begin{cases} \lambda - \mu - v = 0 \\ \lambda - 4v = 0 \\ -3\lambda + 4\mu - v = 0 \end{cases} &\xleftrightarrow{L_2 \leftrightarrow L_1} \begin{cases} \lambda - \mu - v = 0 \\ -3\lambda + 4\mu - v = 0 \\ \lambda - 4v = 0 \end{cases} \\ &\xleftrightarrow{\begin{matrix} L_2 \leftarrow L_2 + 3L_1 \\ L_3 \leftarrow L_3 - L_1 \end{matrix}} \begin{cases} \lambda - \mu - v = 0 \\ \mu - 4v = 0 \\ \mu - 3v = 0 \end{cases} &\iff \begin{cases} \lambda - \mu - v = 0 \\ \mu - 4v = 0 \\ v = 0 \end{cases} &\iff \begin{cases} \lambda = 0 \\ \mu = 0 \\ v = 0 \end{cases} \end{aligned}$$

Donc $x = \mu(1, 4, 1) = (0, 0, 0)$. Ainsi, $F \cap G = \{0_{\mathbf{R}^3}\}$, et par ce qui a été dit précédemment sur les dimensions, $\mathbf{R}^3 = F \oplus G$.

On montrerait de même que $\text{Vect}((0, 0, 1))$ est un autre supplémentaire de F .

Théorème 3.38 : Soit E un espace vectoriel de dimension finie et F un sous-espace vectoriel de E . Alors il existe un supplémentaire de F dans E .

Terminologie
Attention à ne pas dire le supplémentaire de F dans E , mais bien un supplémentaire de F dans E , car il existe une infinité de tels supplémentaires.

Démonstration. Notons $n = \dim F$ et $p = \dim E$, avec $n \leq p$. Soit (e_1, \dots, e_n) une base de F , et \mathcal{B} une base de E . Par le théorème de la base incomplète, il existe des vecteurs (e_{n+1}, \dots, e_p) de \mathcal{B} tels que (e_1, \dots, e_p) soit une base de E . Posons alors $G = \text{Vect}(e_{n+1}, \dots, e_p)$. G est alors un sous-espace vectoriel de E dont une base est (e_{n+1}, \dots, e_p) . Par construction, $E = F + G$, et la concaténation d'une base de F et d'une base de G est une base de E , donc la somme est directe : $E = F \oplus G$. \square

3.3.4 Somme de n sous-espaces vectoriels

Définition 3.39 – On appelle somme de F_1, \dots, F_n , et on note $\sum_{i=1}^n F_i$ l'ensemble
$$\sum_{i=1}^n F_i = \{x_1 + \dots + x_n, x_i \in F_i\}.$$

Autrement dit
Un vecteur de E est dans $\sum F_i$ s'il s'écrit comme la somme d'un élément de F_1 , un élément de F_2, \dots , un élément de F_n .

La proposition suivante se prouve de manière analogue au cas $n = 2$:

Proposition 3.40 : $\sum_{i=1}^n F_i$ est un sous-espace vectoriel de E , et si H est un sous-espace vectoriel de E tel que $\forall i \in \llbracket 1, n \rrbracket, F_i \subset H$, alors $\sum_{i=1}^n F_i \subset H$.

Proposition 3.41 : Si les F_i sont de dimension finie, alors $\sum_{i=1}^n F_i$ est de dimension finie et

$$\dim\left(\sum_{i=1}^n F_i\right) \leq \sum_{i=1}^n \dim F_i.$$

Démonstration. On prouve comme à la proposition 3.30 que la concaténation de bases des F_i est génératrice de $\sum F_i$, et on conclut comme au corollaire 3.31. \square

3.3.5 Somme directe de n sous-espaces vectoriels

Définition 3.42 – On dit que la somme $\sum_{i=1}^n F_i$ est directe si

$$\forall (x_1, \dots, x_n) \in F_1 \times \dots \times F_n, 0_E = x_1 + \dots + x_n \Rightarrow x_1 = \dots = x_n = 0.$$

Dans ce cas, on note

$$\sum_{i=1}^n F_i = \bigoplus_{i=1}^n F_i = F_1 \oplus \dots \oplus F_n.$$

Proposition 3.43 : Il y a équivalence entre :

- 1) les F_i sont en somme directe
- 2) tout élément de $\sum_{i=1}^n F_i$ s'écrit de manière unique comme somme d'éléments des F_i .

Démonstration. 1) \Rightarrow 2) : si $x = x_1 + \dots + x_n = y_1 + \dots + y_n$, avec $\forall i \in \llbracket 1, n \rrbracket, (x_i, y_i) \in F_i^2$, alors

$$0 = (x_1 - y_1) + \dots + (x_n - y_n), \text{ avec } x_i - y_i \in F_i.$$

Par conséquent, $\forall i \in \llbracket 1, n \rrbracket, x_i - y_i = 0$ et donc $x_i = y_i$: l'écriture de x comme somme d'éléments des F_i est unique.

2) \Rightarrow 1) : Supposons que $0 = x_1 + \dots + x_n$, avec $x_i \in F_i$. Puisque $0 = 0 + \dots + 0$, avec $\forall i \in \llbracket 1, n \rrbracket, 0 \in F_i$, par unicité de l'écriture, $\forall i \in \llbracket 1, n \rrbracket, x_i = 0$, et donc les F_i sont en somme directe. \square

Corollaire 3.44 – Soit E un espace vectoriel et F et G deux sous-espaces vectoriels de E . Alors F et G sont supplémentaires dans E si et seulement si tout vecteur de E s'écrit de manière unique comme somme d'un élément de F et d'un élément de G .

Démonstration. Si $E = F \oplus G$. Alors tout vecteur de $F + G = E$ s'écrit de manière unique comme somme d'un vecteur de F et d'un vecteur de G .

Méthode

Lorsqu'on veut prouver que F et G sont supplémentaires, et qu'on n'a pas accès à leurs dimensions, on utilisera cette propriété et un raisonnement par analyse-synthèse.

Inversement, si tout vecteur de E s'écrit de manière unique comme somme d'un vecteur de F et d'un vecteur de G , en particulier, tout vecteur de E est somme d'un vecteur de F et d'un vecteur de G , donc est dans $F + G$. Et donc $E = F + G$. Si l'écriture est unique, alors la somme est directe, et donc $E = F \oplus G$. \square

Proposition 3.45 : Si F_1, \dots, F_n sont de dimension finie, et si \mathcal{B}_1 est une base de $F_1, \dots, \mathcal{B}_n$ est une base de F_n , alors il y a équivalence entre :

- 1) la somme $\sum_{i=1}^n F_i$ est directe
- 2) la concaténation des \mathcal{B}_i est une base de $\sum_{i=1}^n F_i$
- 3) $\dim\left(\sum_{i=1}^n F_i\right) = \sum_{i=1}^n \dim F_i$.

 **Danger !**

Pour plus de deux sous-espaces vectoriels, il n'existe pas d'analogue à la condition $F \cap G = \{0_E\}$.

Démonstration. 1) \Rightarrow 2) : si la somme est directe, alors tout élément de $\sum_{i=1}^n F_i$ s'écrit de manière unique comme somme d'éléments des F_i . Mais pour tout i , tout vecteur de F_i s'écrit de manière unique comme combinaison linéaire des éléments de \mathcal{B}_i .

Ainsi, tout vecteur de $\sum_{i=1}^n F_i$ s'écrit de manière unique comme combinaison linéaire de $\mathcal{B}_1, \dots, \mathcal{B}_n$: c'est une base de $\sum_{i=1}^n F_i$.

2) \Rightarrow 1) : inversement, si $\mathcal{B}_1 \cup \dots \cup \mathcal{B}_n$ est une base de $F_1 + \dots + F_n$, et si $0 = x_1 + \dots + x_n$, avec $x_i \in F_i$.

Chacun des x_i s'écrit comme combinaison linéaire d'éléments de \mathcal{B}_i , et donc 0 est combinaison linéaire des vecteurs de $\mathcal{B}_1 \cup \dots \cup \mathcal{B}_n$. Mais cette famille est libre, et donc tous les coefficients de cette combinaison linéaire sont nuls. On en déduit que $x_1 = \dots = x_n = 0$.

2) \Rightarrow 3) : évident puisque le cardinal de la concaténation des \mathcal{B}_i est $\sum_{i=1}^n \dim F_i$.

3) \Rightarrow 2) : si la dimension de la somme est la somme des dimensions. Nous savons que la concaténation des \mathcal{B}_i est une famille génératrice de $\sum_{i=1}^n F_i$. Elle est de cardinal

$$\text{Card}(\mathcal{B}_1) + \dots + \text{Card}(\mathcal{B}_n) = \dim F_1 + \dots + \dim F_n = \dim\left(\sum_{i=1}^n F_i\right).$$

C'est donc une base de $\sum_{i=1}^n F_i$. \square

3.4 APPLICATIONS LINÉAIRES

3.4.1 Définition

Définition 3.46 – Soient E, F deux \mathbf{K} -espaces vectoriels. Une application $f : E \rightarrow F$ est dite \mathbf{K} -linéaire si

$$\forall (x, y) \in E^2, \forall \lambda \in \mathbf{K}, f(\lambda x + y) = \lambda f(x) + f(y).$$

On note $\mathcal{L}(E, F)$ l'ensemble des applications linéaires de E dans F .

Proposition 3.47 : $\mathcal{L}(E, F)$ est un \mathbf{K} -espace vectoriel.

Si E et F sont tous deux de dimension finie, alors $\mathcal{L}(E, F)$ est aussi de dimension finie et

$$\dim \mathcal{L}(E, F) = \dim E \times \dim F.$$

Proposition 3.48 : Soit $f \in \mathcal{L}(E, F)$, $(x_1, \dots, x_n) \in E^n$, $(\lambda_1, \dots, \lambda_n) \in \mathbf{K}^n$. Alors

$$f\left(\sum_{i=1}^n \lambda_i x_i\right) = \sum_{i=1}^n \lambda_i f(x_i).$$

En particulier, $f(0_E) = 0_F$.

Définition 3.49 – Si f est une application linéaire de E dans lui-même, on dit que f est un endomorphisme de E , et on note $\mathcal{L}(E) = \mathcal{L}(E, E)$ l'ensemble des endomorphismes de E .

3.4.2 Noyau, image

Soient E et F deux \mathbf{K} -espaces vectoriels, $f \in \mathcal{L}(E, F)$.

Définition 3.50 – On appelle noyau de f et on note $\text{Ker } f$ l'ensemble

$$\text{Ker } f = \{x \in E : f(x) = 0_F\}.$$

Proposition 3.51 : Soit $f \in \mathcal{L}(E, F)$. Alors

- $\text{Ker } f$ est un sous-espace vectoriel de E .
- f est injectif si et seulement si $\text{Ker } f = \{0_E\}$.

Définition 3.52 – On appelle image de f , et on note $\text{Im } f$ l'ensemble

$$\text{Im } f = \{y \in F : \exists x \in E : f(x) = y\}.$$

Proposition 3.53 : Soit $f \in \mathcal{L}(E, F)$. Alors

- $\text{Im } f$ est un sous-espace vectoriel de F .
- f est surjectif si et seulement si $\text{Im } f = F$.

3.4.3 Rang

Définition 3.54 – Soient E et F deux espaces vectoriels tels que E soit de dimension finie et soit $f \in \mathcal{L}(E, F)$. Alors $\text{Im } f$ est de dimension finie. On appelle rang de f et on note $\text{rg}(f) = \dim \text{Im } f$.

Théorème 3.55 (du rang) : Soit E un espace vectoriel de dimension finie, F un espace vectoriel et $f \in \mathcal{L}(E, F)$. Alors

$$\dim E = \dim \text{Ker } f + \text{rg } f = \dim \text{Ker } f + \dim \text{Im } f.$$

 **Danger !**

Ceci ne vaut que pour les applications linéaires !
Si f est linéaire, on essaiera toujours de caractériser l'injectivité à l'aide du noyau plutôt que de revenir à la définition de l'injectivité.

En particulier, pour tout $x \in E$, $f(x) \in \text{Im } f$, par définition.

Ce second point est juste la définition de la surjectivité, et n'utilise pas la linéarité de f .

Corollaire 3.56 – Soient E et F deux espaces vectoriels de dimension finie avec $\dim E = \dim F$, et soit $f \in \mathcal{L}(E, F)$. Alors il y a équivalence entre :

- 1) f est bijectif
- 2) f est injectif
- 3) f est surjectif.

En particulier

Ce résultat s'applique notamment lorsque f est un endomorphisme de E , car alors l'espace de départ et d'arrivée étant les mêmes, ils ont même dimension.

Démonstration. 1) \Rightarrow 2) : évident.

2) \Rightarrow 3) Si f est injectif, alors $\text{Ker } f = \{0_E\}$ et donc $\dim \text{Ker } f = 0$.

Par le théorème du rang, $\dim \text{Im } f = \dim E = \dim F$.

Or, $\text{Im } f \subset F$, et donc, par égalité des dimensions, $\text{Im } f = F$. Ainsi, f est surjectif.

3) \Rightarrow 1) Si f est surjectif, alors $\dim \text{Im } f = \dim F = \dim E$, et par le théorème du rang, $\dim \text{Ker } f = 0$. Ainsi, $\text{Ker } f = \{0_E\}$, donc f est injectif, et donc bijectif. \square

Exemple 3.57

Soit $f : \mathbf{R}^3 \rightarrow \mathbf{R}^3$ définie par $f(x, y, z) = (-2x - y + 2z, -3x - 2y + 2z, -2x)$.

Soit alors $(x, y, z) \in \text{Ker } f$.

$$\begin{cases} -2x - y + 2z = 0 \\ -3x - 2y + 2z = 0 \\ -2x = 0 \end{cases} \Leftrightarrow \begin{cases} 2z = y \\ 2y = 2z \\ x = 0 \end{cases} \Leftrightarrow \begin{cases} z = 0 \\ y = 0 \\ x = 0 \end{cases}$$

Donc $\text{Ker } f = \{0_E\}$, et f est donc injectif. Par conséquent, f est bijectif.

3.4.4 Isomorphismes

Définition 3.58 – Soient E et F deux \mathbf{K} -espaces vectoriels, et $f \in \mathcal{L}(E, F)$. Si f est bijectif, on dit que f est un isomorphisme.

Dans le cas où $E = F$, et si f est bijectif, on dit que f est un automorphisme de E .

Ainsi, un automorphisme est un endomorphisme bijectif.

Proposition 3.59 : Si $f \in \mathcal{L}(E, F)$ est un isomorphisme, alors $f^{-1} \in \mathcal{L}(F, E)$.

Proposition 3.60 : Soit E un \mathbf{K} -espace vectoriel de dimension finie n . Alors il existe un isomorphisme $f : \mathbf{K}^n \rightarrow E$.

Démonstration. Soit (e_1, \dots, e_n) une base de E , et soit $f : \mathbf{K}^n \rightarrow E$ définie par

$$f(\lambda_1, \dots, \lambda_n) = \sum_{i=1}^n \lambda_i e_i.$$

Il est facile de vérifier que f est linéaire. Soit $x \in E$. Alors

$$\exists! (\lambda_1, \dots, \lambda_n) \in \mathbf{K}^n : x = \sum_{i=1}^n \lambda_i e_i = f(\lambda_1, \dots, \lambda_n).$$

Autrement dit, tout élément de E admet un unique antécédent par f : f est un isomorphisme. \square

Proposition 3.61 : Soient E et F deux \mathbf{K} -espaces vectoriels de dimension finie. Alors il existe un isomorphisme $f : E \rightarrow F$ si et seulement si $\dim E = \dim F$.

Démonstration. S'il existe un isomorphisme $f : E \rightarrow F$, alors par le théorème du rang, $\dim E = \dim \text{Im } f$. Mais si f est surjectif, $\text{Im } f = F$ et donc $\dim \text{Im } f = \dim F$. Donc nécessairement $\dim E = \dim F$.

Inversement, si $\dim E = \dim F = n$. Alors par la proposition précédente, il existe $f : \mathbf{K}^n \rightarrow E$ et $g : \mathbf{K}^n \rightarrow F$ qui soient des isomorphismes. Mais alors $g \circ f^{-1}$ est un isomorphisme de E sur F . \square

Proposition 3.62 : Soit E un espace vectoriel de dimension finie, (e_1, \dots, e_n) une base de E , F un espace vectoriel et $f \in \mathcal{L}(E, F)$. Alors f est un isomorphisme si et seulement si $(f(e_1), \dots, f(e_n))$ est une base de F .

Démonstration. Si f est un isomorphisme, alors $\dim F = \dim \text{Im } f = \dim E$.

Soit $y \in F$, et soit $x = f^{-1}(y)$. Alors il existe des scalaires $(\lambda_1, \dots, \lambda_n)$ tels que $x = \sum_{i=1}^n \lambda_i e_i$,

et donc

$$y = f(x) = f\left(\sum_{i=1}^n \lambda_i e_i\right) = \sum_{i=1}^n \lambda_i f(e_i).$$

Ainsi, $(f(e_1), \dots, f(e_n))$ est une famille génératrice de F .

Puisqu'elle possède $n = \dim F$ éléments, c'est une base de F .

Inversement, supposons que $(f(e_1), \dots, f(e_n))$ est une base de F .

Soit $y \in F$: il existe des scalaires $\lambda_1, \dots, \lambda_n$ tels que $y = \sum_{i=1}^n \lambda_i f(e_i) = f\left(\sum_{i=1}^n \lambda_i e_i\right) \in \text{Im } f$.

Ainsi, f est surjectif, et comme E et F ont même dimension, il s'agit d'un isomorphisme. \square

3.4.5 Projecteurs

Définition 3.63 – Soit E un espace vectoriel, F et G deux sous-espaces vectoriels supplémentaires dans E . Alors tout élément $x \in E$ s'écrit de manière unique $x = x_1 + x_2, x_1 \in F, x_2 \in G$. On appelle projection sur F parallèlement à G l'endomorphisme $p \in \mathcal{L}(E)$ qui à x associe x_1 , sa composante suivant F .

Définition 3.64 – On appelle projecteur de E tout $p \in \mathcal{L}(E)$ tel que $p \circ p = p$.

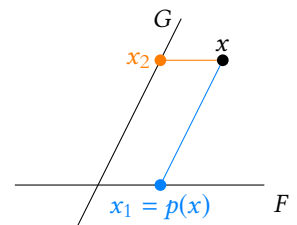


FIGURE 3.1– La projection sur F parallèlement à G .

Exemple 3.65

L'application $f : \mathcal{M}_n(\mathbf{R}) \rightarrow \mathcal{M}_n(\mathbf{R}), M \mapsto \frac{M + {}^t M}{2}$ est un projecteur de $\mathcal{M}_n(\mathbf{R})$.

En effet, elle est clairement⁵ linéaire, et

$$f^2(M) = \frac{f(M) + {}^t f(M)}{2} = \frac{1}{4} (M + {}^t M + {}^t (M + {}^t M)) = \frac{M + {}^t M}{2} = f(M).$$

Ceci étant vrai pour tout $M \in \mathcal{M}_n(\mathbf{R}), f^2 = f$.

A priori, projecteurs et projections sont deux notions différentes. Le théorème suivant prouve qu'il s'agit en fait de la même chose :

Théorème 3.66 : Soit E un espace vectoriel et $p \in \mathcal{L}(E)$. Alors p est un projecteur si et seulement si p est une projection. Dans ce cas, p est la projection sur $\text{Im } p$ parallèlement à $\text{Ker } p$.

Exercice
Par un raisonnement analogue, prouver que pour tout $f \in \mathcal{L}(E, F)$, l'image d'une famille génératrice de E par f est toujours une famille génératrice de $\text{Im } f$.

⁵ C'est une combinaison linéaire des deux applications linéaires $M \mapsto {}^t M$ et $M \mapsto M$.

Remarque
En particulier, si p est un projecteur, $\text{Im } p$ et $\text{Ker } p$ sont supplémentaires dans E .

Démonstration. Supposons que $p \circ p = p$. Montrons que $\text{Im } p$ et $\text{Ker } p$ sont supplémentaires dans E .

Nous allons pour cela prouver que tout vecteur de E s'écrit de manière unique comme somme d'un vecteur de $\text{Im } p$ et d'un vecteur de $\text{Ker } p$.

Procédons par analyse-synthèse. Soit $x \in E$ et supposons que $x = x_1 + x_2$, $x_1 \in \text{Im } p$, $x_2 \in \text{Ker } p$. Alors

$$p(x) = p(x_1) + p(x_2) = p(x_1).$$

Mais $x_1 \in \text{Im } p$, donc il existe $y_1 \in E$ tel que $x_1 = p(y_1)$, et alors $p(x_1) = p^2(y_1) = p(y_1) = x_1$. Donc $p(x) = x_1$, et alors $x_2 = x - x_1 = x - p(x)$.

Or, pour tout $x \in E$, $x = p(x) + (x - p(x))$.

Mais $p(x) \in \text{Im } p$ par définition et $x - p(x) \in \text{Ker } p$ car

$$p(x - p(x)) = p(x) - p^2(x) = p(x) - p(x) = 0.$$

Donc tout vecteur de E s'écrit bien d'au moins une manière comme somme d'un élément de $\text{Im } p$ et d'un élément de $\text{Ker } p$.

Ainsi, tout vecteur de E s'écrit de manière unique comme somme d'un élément de $\text{Im } p$ et d'un élément de $\text{Ker } p$: $\text{Im } p$ et $\text{Ker } p$ sont supplémentaires dans E .

Enfin, si $x \in E$, $x = \underbrace{p(x)}_{\in \text{Im } p} + \underbrace{x - p(x)}_{\in \text{Ker } p}$ et $p(x) = p(x)$.

Donc p est l'application qui à x associe sa composante $p(x)$ suivant $\text{Im } p$: c'est donc la projection sur $\text{Im } p$ parallèlement à $\text{Ker } p$.

Inversement, si p est la projection sur F parallèlement à G , avec $E = F \oplus G$.

Soit alors $x \in E$. De manière unique, $x = x_1 + x_2$, avec $x_1 \in F$, $x_2 \in G$.

Alors $p(x) = x_1 = x_1 + 0$, avec $x_1 \in F$ et $0 \in G$. Donc $p^2(x) = p(x_1) = x_1 = p(x)$. Ceci prouve que $p^2 = p$, et donc p est un projecteur.

De plus, $\text{Ker } p = \{x_1 + x_2, x_1 \in F, x_2 \in G : x_1 = 0\} = \{x_2, x_2 \in G\} = G$.

De même, il est clair que $\text{Im } p \subset F$. Soit $x \in F$. Alors $p(x) = x$, et donc $x \in \text{Im } p$. Ceci prouve que $F \subset \text{Im } p$, et donc $\text{Im } p = F$. \square

Exemple 3.67

Si f est le projecteur $M \mapsto \frac{M + {}^t M}{2}$, alors

$$M \in \text{Ker}(f) \Leftrightarrow \frac{M + {}^t M}{2} = 0 \Leftrightarrow M = -{}^t M$$

et donc $\text{Ker}(f)$ l'ensemble $\mathcal{A}_n(\mathbf{R})$ des matrices antisymétriques de $\mathcal{M}_n(\mathbf{R})$.

D'autre part, il est clair que pour tout $M \in \mathcal{M}_n(\mathbf{R})$, $f(M)$ est une matrice symétrique,

puisque ${}^t f(M) = {}^t \left(\frac{M + {}^t M}{2} \right) = \frac{M + {}^t M}{2} = f(M)$.

Donc déjà $\text{Im}(f) \subset \mathcal{S}_n(\mathbf{R})$. De plus, pour tout $M \in \mathcal{S}_n(\mathbf{R})$, on a $f(M) = M$, et donc $M \in \text{Im}(f)$.

Ainsi, $\text{Im}(f) = \mathcal{S}_n(\mathbf{R})$.

Puisque nous savons déjà que f est un projecteur, il vient donc $\mathcal{M}_n(\mathbf{R}) = \mathcal{S}_n(\mathbf{R}) \oplus \mathcal{A}_n(\mathbf{R})$, et f est la projection sur $\mathcal{S}_n(\mathbf{R})$ parallèlement à $\mathcal{A}_n(\mathbf{R})$.

3.4.6 Formes linéaires et hyperplans

Définition 3.68 – Une application linéaire d'un \mathbf{K} -espace vectoriel E vers \mathbf{K} est appelée forme linéaire sur E .

Analyse-synthèse

Nous venons de procéder à l'analyse : si une écriture $x = x_1 + x_2$ existe, elle est unique et nous savons exprimer x_1 et x_2 en fonction de x .

Il reste à procéder à la synthèse en vérifiant que les formules obtenues précédemment définissent bien des vecteurs $x_1 \in \text{Im } p$ et $x_2 \in \text{Ker } p$ tels que $x = x_1 + x_2$.

Remarque

Le fait que $\mathcal{S}_n(\mathbf{R})$ et $\mathcal{A}_n(\mathbf{R})$ soient supplémentaires dans $\mathcal{M}_n(\mathbf{R})$ implique que toute matrice de $M \in \mathcal{M}_n(\mathbf{R})$ s'écrit de manière unique $M = S + A$, avec $S \in \mathcal{S}_n(\mathbf{R})$ et $A \in \mathcal{A}_n(\mathbf{R})$.

Nous savons alors que $S = f(M) = \frac{M + {}^t M}{2}$, et donc $A = M - f(M) = \frac{M - {}^t M}{2}$.

Définition 3.69 – Soit E un espace vectoriel de dimension finie non réduit à $\{0_E\}$. Un sous-espace vectoriel F de E est appelé hyperplan de E si $\dim(F) = \dim(E) - 1$.

Proposition 3.70 : Soit E un espace vectoriel de dimension finie. Alors un sous-espace vectoriel F de E est un hyperplan de E si et seulement si il existe une forme linéaire **non nulle** $\varphi \in \mathcal{L}(E, \mathbf{K})$ telle que $F = \text{Ker}\varphi$.

Bien évidemment, le noyau de la forme linéaire nulle n'est pas un hyperplan puisque c'est E tout entier.

Exemple 3.71

Étudions la dimension de $F = \{(x, y, z, t) \in \mathbf{R}^4 : 2x - 3y + 2t = 0\}$.
Soit φ l'application de \mathbf{R}^4 dans \mathbf{R} définie par $\varphi(x, y, z, t) = 2x - 3y + 2t$.
Il est aisé de prouver que φ est une forme linéaire sur \mathbf{R}^4 , non nulle (par exemple car $\varphi(1, 0, 0, 0) = 1 \neq 0$).
Ainsi $F = \text{Ker}\varphi$ est un hyperplan de \mathbf{R}^4 , et donc $\dim F = \dim \mathbf{R}^4 - 1 = 4 - 1 = 3$.

3.4.7 Sous-espaces stables par un endomorphisme

Définition 3.72 – Soit E un \mathbf{K} -espace vectoriel et $f \in \mathcal{L}(E)$. Un sous-espace vectoriel F de E est dit stable par f si

$$\forall x \in F, f(x) \in F.$$

Exemples 3.73

- $\{0_E\}$ et E sont toujours des sous-espaces vectoriels stables, quel que soit $f \in \mathcal{L}(E)$.
- Soit $E = \mathbf{R}[X]$, $f : P \mapsto XP'$, et $F = \mathbf{R}_n[X]$.
Si $\deg P \leq n$, alors $\deg P' \leq n - 1$ et donc $\deg(XP') \leq n$.
Ainsi, $\forall P \in \mathbf{R}_n[X], f(P) \in \mathbf{R}_n[X]$. Donc $\mathbf{R}_n[X]$ est stable par f .
- Soit $f : \mathbf{R}^2 \rightarrow \mathbf{R}^2$ définie par $f(x, y) = (5x + 3y, -6x - 4y)$. Alors $\text{Vect}((1, -2))$ est stable par f .
En effet, soit $(\lambda, -2\lambda) \in \text{Vect}((1, -2))$. Alors

$$f(\lambda, -2\lambda) = (5\lambda - 6\lambda, -6\lambda + 8\lambda) = (-\lambda, 2\lambda) = -\lambda(1, -2) \in \text{Vect}((1, -2)).$$



Remarque. Si $f \in \mathcal{L}(E)$ et si F est un sous-espace vectoriel de E , alors on peut considérer la restriction $f|_F$ de f à F , c'est-à-dire l'application de F dans E , qui à $x \in F$ associe $f(x)$. Si F est stable par f , alors pour tout $x \in F$, $f|_F(x) = f(x) \in F$, et donc $f|_F$ va de F dans F : c'est un endomorphisme de F .

Autrement dit

$f|_F$ n'est rien d'autre que f , sauf qu'au lieu de le voir comme une application définie sur E tout entier, on le voit comme une application définie uniquement sur F .

Proposition 3.74 : Soit $f \in \mathcal{L}(E)$. Alors : $\text{Im} f$ et $\text{Ker} f$ sont stables par f .

Démonstration. Soit $x \in \text{Ker} f$. Alors $f(x) = 0_E \in \text{Ker} f$, de sorte que $\text{Ker} f$ est stable par f .

Soit $x \in \text{Im} f$. Alors $f(x) \in \text{Im} f$ par définition de l'image, donc $\text{Im} f$ est stable par f .

□

MÉTHODES DU CHAPITRE 3

✓ MÉTHODE 3.1 : Déterminer la dimension d'un sous-espace vectoriel F de E .

Puisque la dimension est le cardinal d'une base, il faudra déterminer une base de F . Sauf dans quelques cas particuliers :

- Si E est défini comme le noyau d'une forme linéaire non nulle sur E . Alors F est un hyperplan et $\dim F = \dim E - 1$.
- Si on connaît un supplémentaire G de F dans E , alors $\dim F = \dim E - \dim G$.

Exercices de TD : 3.2, 3.3, 3.5

✓ MÉTHODE 3.2 : Montrer qu'une famille de vecteurs (x_1, \dots, x_n) d'un espace vectoriel E est libre

La plupart du temps, on revient à la définition : on prend $\lambda_1, \dots, \lambda_n$ des réels que l'on suppose vérifier $\sum_{i=1}^n \lambda_i x_i = 0$ et on se débrouille (au cas par cas, il n'y a pas de méthode générale) pour montrer que les λ_i sont tous nuls.

Cas particuliers :

- Pour une famille de deux vecteurs non nuls, il suffit de montrer que ces deux vecteurs ne sont pas colinéaires.
- Si on connaît une famille libre contenant x_1, \dots, x_n , le résultat est automatique.
- Pour une famille de polynômes, s'ils sont de degrés deux à deux distincts, alors la famille est libre (\triangleleft réciproque fausse !)

Exercices de TD : 3.1

✓ MÉTHODE 3.3 : Montrer qu'une famille (x_1, \dots, x_n) est une base d'un espace vectoriel E .

Il faut prouver que la famille est à la fois libre et génératrice. Pour la liberté, on procède comme d'habitude. En revanche, l'aspect générateur est rarement facile à prouver. Si on connaît la dimension de E , on préférera prouver que la famille est libre et possède $\dim E$ éléments.

Exercices de TD : 3.2, 3.3, 3.11, 3.20

✓ MÉTHODE 3.4 : Montrer qu'un sous-espace vectoriel F de E est stable par $f \in \mathcal{L}(E)$.

Il s'agit de montrer que pour tout $x \in F$, $f(x)$ est encore dans F .

Pour cela, on prend donc $x \in F$ (sans faire d'autre hypothèses sur x), et on montre que $f(x)$ est encore dans F .

Exercices de TD : 3.16, 3.17

✓ MÉTHODE 3.5 : Montrer que $F \subset E$ est un sous-espace vectoriel de E .

Ne pas oublier de vérifier que $0 \in F$ afin de s'assurer que F est non vide.

Puis prouver que si x, y sont deux éléments quelconques de F et λ est un scalaire quelconque dans \mathbf{K} , alors $\lambda x + y \in F$.

Parfois on arrive à écrire directement F comme noyau d'une application linéaire définie sur E , ce qui suffit à prouver que F est un sous-espace vectoriel de E .

Exercices de TD : 3.2, 3.2

✓ MÉTHODE 3.6 : Étudier l'injectivité d'une application linéaire $f \in \mathcal{L}(E, F)$.

On utilisera toujours la caractérisation par le noyau plutôt que de revenir à la définition de l'injectivité. Il s'agit donc de prouver que si $x \in E$ vérifie $f(x) = 0_F$, alors nécessairement $x = 0_E$.

Exercices de TD : 3.8, 3.13, 3.20

✓ MÉTHODE 3.7 : Étudier la surjectivité d'une application linéaire $f \in \mathcal{L}(E, F)$.

Cela revient à montrer que $\text{rg } f = \dim \text{Im } f = \dim F$.

- En premier lieu, si on a accès à la dimension de $\text{Ker } f$, on pensera au théorème du rang.
- Si on connaît la matrice M de f dans des bases de E et F , alors le rang de f est égal au rang de M .
- Si rien d'autre ne fonctionne, il faut revenir à la définition de la surjectivité, et pour tout $y \in F$, essayer de trouver un antécédent de y par f , c'est-à-dire $x \in E$ tel que $y = f(x)$.

Exercices de TD : 3.13, 3.20

✓ **MÉTHODE 3.8 : Montrer que deux sous-espaces vectoriels $F, G \subset E$ sont en somme directe**

Il suffit de prouver que $F \cap G = \{0_E\}$. Pour cela, on prendra x un vecteur que l'on suppose appartenir à la fois à F et à G , et on montrera que nécessairement $x = 0_E$.

✓ **MÉTHODE 3.9 : Montrer que deux sous-espaces vectoriels $F, G \subset E$ sont supplémentaires dans E .**

• Si l'on a accès aux dimensions de F et G et à la dimension de E , on montrera que

- $\dim F + \dim G = \dim E$
- $F \cap G = \{0_E\}$ (i.e. $F \cap G$ ne contient que le vecteur nul).

• Sinon on montrera que tout élément de E peut s'écrire de manière unique comme somme d'un élément de F et d'un élément de G (souvent via un raisonnement par analyse-synthèse).

Exercices de TD : 3.3, 3.5, 3.6, 3.12, 3.19

✓ **MÉTHODE 3.10 : Montrer que n sous-espaces vectoriels $F_1, \dots, F_n \subset E$ sont en somme directe**

On revient systématiquement à la définition d'une somme directe, en prenant $x_i \in F_i$ tel que $x_1 + \dots + x_n = 0$, et on montre que nécessairement $x_1 = \dots = x_n = 0$.

⚠ Il n'y a pas de caractérisation de la somme directe à base d'intersection et/ou de dimension, celle ci est réservée au cas de deux sous-espaces vectoriels.

Exercices de TD : 3.18

EXERCICES DU CHAPITRE 3

EXERCICE 3.1

F

- 1) Soit (e_1, e_2, e_3) une famille génératrice de \mathbf{R}^3 . Montrer que (e_1, e_3) est libre.
- 2) Soit E un espace vectoriel et soient u, v, w trois vecteurs deux à deux non colinéaires. La famille (u, v, w) est-elle nécessairement libre ?
- 3) Soit (e_1, e_2, e_3) une famille libre d'un espace vectoriel E . Les familles $(e_1, e_2 + e_1, e_2 + e_3)$ et $(e_1, e_2, 2e_3)$ sont-elles libres ?

EXERCICE 3.2 Montrer que les ensembles suivants sont des espaces vectoriels, en déterminer une base et la dimension.

F

$$E = \{(x, y, z) \in \mathbf{R}^3 : x + 2y = 0\} \quad F = \{(x, y, z) \in \mathbf{R}^3 : x + 5y - 3z = 0 \text{ et } -x - 4y + 2z = 0\}$$

$$G = \{P \in \mathbf{R}_2[X] : (X-1)P' - XP'' = 2P\} \quad H = \{M \in \mathcal{M}_2(\mathbf{R}) : MN = NM\} \text{ où } N = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}, \text{ avec } \lambda_1, \lambda_2 \text{ deux réels fixés distincts}$$

EXERCICE 3.3 Dans $\mathcal{M}_2(\mathbf{R})$, on considère les deux sous-ensembles suivants :

PD

$$F = \left\{ \begin{pmatrix} a & 2a+b \\ -b & -a \end{pmatrix}, (a, b) \in \mathbf{R}^2 \right\} \text{ et } G = \left\{ \begin{pmatrix} a & 3a+b \\ -b & -2a+b \end{pmatrix}, (a, b) \in \mathbf{R}^2 \right\}.$$

- 1) Montrer que F et G sont des sous-espaces vectoriels de $\mathcal{M}_2(\mathbf{R})$ et en déterminer une base.
- 2) Montrer que F et G sont supplémentaires dans $\mathcal{M}_2(\mathbf{R})$.

EXERCICE 3.4 Soit E un espace vectoriel, $f \in \mathcal{L}(E)$ et $\lambda \in \mathbf{R}$. Montrer que $E_\lambda(f) = \{x \in E : f(x) = \lambda x\}$ est un sous-espace vectoriel de E .

F

EXERCICE 3.5 Soit $F = \{(x, y, z, t) \in \mathbf{R}^4 : x + y + 2z - t = 0\}$ et $G = \text{Vect}(1, 1, 1, 1)$. Montrer que F et G sont deux sous-espaces supplémentaires de \mathbf{R}^4 .

F

EXERCICE 3.6 Soit $F = \{f \in \mathcal{C}(\mathbf{R}, \mathbf{R}) \text{ telles que } f(0) = f(1)\}$.

AD

- 1) Montrer que F est un espace vectoriel.
- 2) Soit $g : x \mapsto x$. Montrer que F et $\text{Vect}(g)$ sont supplémentaires dans $\mathcal{C}(\mathbf{R}, \mathbf{R})$.

EXERCICE 3.7 Un résultat à connaître

PD

Soient E, F, G trois espaces vectoriels, et soient $f \in \mathcal{L}(E, F), g \in \mathcal{L}(F, G)$. Montrer que $g \circ f = 0 \iff \text{Im}(f) \subset \text{Ker}(g)$.

EXERCICE 3.8 Soit p l'application $\mathbf{C}^2 \rightarrow \mathbf{C}^2$ définie par $p(x, y) = \frac{1}{5}(x + 2y, 2x + 4y)$.

PD

- 1) Montrer que p est un endomorphisme.
- 2) Déterminer une base de $\text{Ker} p$. L'application p est-elle injective ?
- 3) Déterminer une base de Imp .
- 4) Montrer que $p \circ p = p$. Les sous-espaces $\text{Ker} p$ et Imp sont-ils supplémentaires dans \mathbf{C}^2 ?

EXERCICE 3.9 Projecteurs et symétries

AD

Soit E un espace vectoriel de dimension finie. On dit qu'un endomorphisme $f \in \mathcal{L}(E)$ est une symétrie si $f^2 = \text{id}_E$.

- 1) Soit p un projecteur de E .
 - a) Montrer que $\text{Im}(p) = \text{Ker}(p - \text{id}_E)$.
 - b) Montrer que $2p - \text{id}_E$ est une symétrie.
- 2) On suppose dans cette question que $g \in \mathcal{L}(E)$ est une symétrie.
 - a) Montrer que $\frac{1}{2}(g + \text{id}_E)$ est un projecteur.
 - b) En déduire que $\text{Ker}(g - \text{id})$ et $\text{Ker}(g + \text{id}_E)$ sont supplémentaires dans E .

EXERCICE 3.10 Projecteurs associés

AD

Soient p et q deux projecteurs d'un espace vectoriel E tels que $p + q = \text{id}_E$. Montrer que $E = \text{Im} p \oplus \text{Im} q$.

EXERCICE 3.11 Dans $\mathbf{R}_4[X]$, on pose $F = \{P \in \mathbf{R}_4[X] : P(0) = 0\}$, $G = \{P \in \mathbf{R}_4[X] : P(4) = 0\}$ et $H = F \cap G$.

PD

- 1) Montrer que F est un sous-espace vectoriel de $\mathbf{R}_4[X]$ et en donner une base.
- 2) Montrer que G est un sous-espace vectoriel de $\mathbf{R}_4[X]$ et en donner une base formée de puissances de $(X - 4)$.
- 3) Montrer que $H = \{X(X - 4) \times Q, Q \in \mathbf{R}_2[X]\}$. En déduire une base de H .

EXERCICE 3.12 Soit E un espace vectoriel de dimension finie n et soit φ une forme linéaire non nulle sur E . Montrer que pour tout $u \notin \text{Ker}(\varphi)$, $E = \text{Ker}(\varphi) \oplus \text{Vect}(u)$. PD

EXERCICE 3.13 Polynômes de Lagrange PD

Soit $n \in \mathbf{N}$, et soient a_0, \dots, a_n $n + 1$ réels deux à deux distincts. On considère l'application $\varphi : \begin{cases} \mathbf{R}_n[X] \rightarrow \mathbf{R}^{n+1} \\ P \mapsto (P(a_0), \dots, P(a_n)) \end{cases}$

- 1) Montrer que φ est linéaire.
- 2) Montrer que φ est un isomorphisme.
- 3) Pour $i \in \llbracket 0, n \rrbracket$, on pose $L_i = \prod_{\substack{j=0 \\ j \neq i}}^n \frac{X - a_j}{a_i - a_j}$.
 - a) Calculer $\varphi(L_i)$.
 - b) En déduire que (L_0, \dots, L_n) est une base de $\mathbf{R}_n[X]$.

EXERCICE 3.14 Soit E, F, G trois espaces vectoriels de dimension finie et soient $f \in \mathcal{L}(E, F)$ et $g \in \mathcal{L}(F, G)$. AD

- 1) Montrer que $\text{rg}(g \circ f) \leq \text{rg}(g)$.
- 2) Soit (e_1, \dots, e_p) une base de $\text{Im } f$. Montrer que $g(e_1), \dots, g(e_p)$ est une famille génératrice de $\text{Im}(g \circ f)$. En déduire que $\text{rg}(g \circ f) \leq \text{rg}(f)$.
- 3) Montrer que $\text{rg}(g \circ f) \leq \min(\text{rg}(g), \text{rg}(f))$.

EXERCICE 3.15 Soit E un espace vectoriel de dimension finie et $f \in \mathcal{L}(E)$ tel que $f^2 + f - 6 \text{id}_E = 0$. On pose alors $u = f + 3 \text{id}_E$ et $v = f - 2 \text{id}_E$. AD

- 1) Calculer $u - v$. En déduire que $E = \text{Im } u + \text{Im } v$.
- 2) Calculer $u \circ v$ et $v \circ u$. En déduire que $\text{Im } u \subset \text{Ker } v$ et $\text{Im } v \subset \text{Ker } u$.
- 3) Montrer que $E = \text{Ker } u \oplus \text{Ker } v$.

EXERCICE 3.16 Soit E un espace vectoriel, et soient $f, g \in \mathcal{L}(E)$ deux endomorphismes qui commutent (i.e. $f \circ g = g \circ f$). PD

- 1) Montrer que $\text{Ker } f$ est stable par g .
- 2) Montrer que $\text{Im } f$ est stable par g .

EXERCICE 3.17 Soit E un espace vectoriel, $p \in \mathcal{L}(E)$ un projecteur et $f \in \mathcal{L}(E)$. Montrer que p et f commutent si et seulement si $\text{Ker } p$ et $\text{Im } p$ sont stables par f . AD

EXERCICE 3.18 Soit $n \geq 1$ fixé. Pour tout $i \in \llbracket 0, n \rrbracket$, on pose $F_i = \{P \in \mathbf{R}_n[X] : \forall j \in \llbracket 0, n \rrbracket \setminus \{i\}, P(j) = 0\}$. D

- 1) Montrer que les F_i sont des sous-espaces vectoriels de $\mathbf{R}_n[X]$, et que pour $P \in \mathbf{R}_n[X]$, on a $P \in F_i$ si et seulement si il existe $\lambda \in \mathbf{R}$ tel que $P = \lambda \prod_{\substack{j=0 \\ j \neq i}}^n (X - j)$.
- 2) Montrer que la somme $F_0 + F_1 + \dots + F_n$ est directe.
- 3) En déduire que $\mathbf{R}_n[X] = \bigoplus_{i=0}^n F_i$.

EXERCICE 3.19 (QSP ESCP 2004) D

Soit E un espace vectoriel de dimension finie, et soient f et g deux endomorphismes de E tels que $E = \text{Im } f + \text{Im } g$ et $E = \text{Ker } f + \text{Ker } g$. Montrer que ces sommes sont directes.

EXERCICE 3.20 Polynômes d'interpolation de Lagrange (Oral ESCP 2012) TD

Dans tout cet exercice, n désigne un entier naturel non nul et (a_1, \dots, a_n) une famille de nombres réels distincts. Pour tous $(i, j) \in \mathbf{N}^2$, on note $\delta_{i,j} = 1$ si $i = j$ et $\delta_{i,j} = 0$ si $i \neq j$.

- 1) a) Soit $i \in \llbracket 1, n \rrbracket$. Montrer qu'il existe un unique polynôme $L_i \in \mathbf{R}_{n-1}[X]$ tel que pour tout $j \in \llbracket 1, n \rrbracket$, $L_i(a_j) = \delta_{i,j}$.
 b) Montrer que la famille $(L_i)_{i \in \llbracket 1, n \rrbracket}$ est une base de $\mathbf{R}_{n-1}[X]$.
- 2) Soit $\pi : \mathbf{R}[X] \rightarrow \mathbf{R}[X]$ définie par $\forall P \in \mathbf{R}[X], \pi(P) = \sum_{i=1}^n P(a_i)L_i$.

- a) Montrer que π est un projecteur de $\mathbf{R}[X]$.
- b) Déterminer le noyau et l'image de π .
- c) On note $F = \left\{ Q \prod_{i=1}^n (X - a_i), Q \in \mathbf{R}[X] \right\}$. Montrer que $F \oplus \mathbf{R}_{n-1}[X] = \mathbf{R}[X]$.
- d) Soit $P \in \mathbf{R}_{n-1}[X]$. Déterminer les coordonnées de P dans la base $(L_i)_{i \in \llbracket 1, n \rrbracket}$.
- 3) Soit $\varepsilon : \mathbf{R}_{n-1}[X] \rightarrow \mathbf{R}^n, P \mapsto (P(a_i))_{i \in \llbracket 1, n \rrbracket}$.
- a) Montrer que ε est un isomorphisme.
- b) Soit $f \in \mathcal{F}(\mathbf{R}, \mathbf{R})$. Montrer qu'il existe un unique polynôme $P \in \mathbf{R}_{n-1}[X]$ tel que $P(a_i) = f(a_i)$, pour tout $i \in \llbracket 1, n \rrbracket$.
Ce polynôme s'appelle le polynôme d'interpolation de Lagrange associé à la fonction f aux points (a_1, \dots, a_n) .
- 4) Soient $a, b \in \mathbf{R}$ tels que $a < b$. Soient $f \in \mathcal{C}^n([a, b], \mathbf{R})$, a_1, a_2, \dots, a_n tels que $a \leq a_1 < \dots < a_n \leq b$ et P le polynôme d'interpolation de Lagrange associé à f et aux points (a_1, \dots, a_n) .
- a) Soit $x \in [a, b] \setminus \{a_1, \dots, a_n\}$ et K réel. On définit la fonction φ par :

$$\varphi : [a, b] \rightarrow \mathbf{R}, t \mapsto f(t) - P(t) - K \prod_{i=1}^n (t - a_i).$$

Montrer qu'il existe K tel que $\varphi(x) = 0$.

- b) Montrer que pour cette valeur de K , il existe $\zeta \in [a, b]$ tel que $\varphi^{(n)}(\zeta) = 0$.
- c) Montrer que pour tout $x \in [a, b]$, on a :

$$|f(x) - P(x)| \leq \frac{\prod_{i=1}^n |x - a_i|}{n!} \sup_{[a, b]} |f^{(n)}|.$$

CORRECTION DES EXERCICES DU CHAPITRE 3

SOLUTION DE L'EXERCICE 3.1

1. Puisque (e_1, e_2, e_3) est une famille génératrice de cardinal $3 = \dim \mathbf{R}^3$, c'est une base de \mathbf{R}^3 . Elle est donc libre, et par conséquent, la sous-famille (e_1, e_3) est également libre.
2. Non, par exemple si u et v sont non nuls et que $w = u + v$, alors les vecteurs u, v, w sont deux à deux non colinéaires, mais pourtant ne forment pas une famille libre puisque $u + v - w = 0_E$.
3. La première famille est libre. En effet, soient $\lambda_1, \lambda_2, \lambda_3$ des scalaires tels que

$$\lambda_1 e_1 + \lambda_2(e_2 + e_1) + \lambda_3(e_2 + e_3) = 0_E.$$

Alors $(\lambda_1 + \lambda_2)e_1 + (\lambda_2 + \lambda_3)e_2 + \lambda_3 e_3 = 0_E$.

Puisque (e_1, e_2, e_3) est libre, on en déduit que $\begin{cases} \lambda_1 + \lambda_2 = 0 \\ \lambda_2 + \lambda_3 = 0 \\ \lambda_3 = 0 \end{cases}$ et donc $\lambda_1 = \lambda_2 = \lambda_3 = 0$.

On prouve de même que la seconde famille est libre.

SOLUTION DE L'EXERCICE 3.2

$$E = \{(x, y, z) \in \mathbf{R}^3 : x = -2y\} = \{(-2y, y, z), (y, z) \in \mathbf{R}^2\} = \{y(-2, 1, 0) + z(0, 0, 1), (y, z) \in \mathbf{R}^2\} = \text{Vect}((-2, 1, 0), (0, 0, 1)).$$

Ainsi, la famille $(-2, 1, 0), (0, 0, 1)$ est génératrice de E . Elle est libre car formée de deux vecteurs non colinéaires : c'est une base de E , qui est donc de dimension 2.

$$\begin{aligned} F &= \left\{ (x, y, z) \in \mathbf{R}^3 : \begin{cases} x + 5y - 3z = 0 \\ -x - 4y + 2z = 0 \end{cases} \right\} \\ &= \left\{ (x, y, z) \in \mathbf{R}^3 : \begin{cases} x + 5y - 3z = 0 \\ y = z \end{cases} \right\} \\ &= \left\{ (x, y, z) \in \mathbf{R}^3 : \begin{cases} x = -2z \\ y = z \end{cases} \right\} \\ &= \{(-2z, z, z), z \in \mathbf{R}\} = \{z(-2, 1, 1), z \in \mathbf{R}\} = \text{Vect}(-2, 1, 1). \end{aligned}$$

La famille $(-2, 1, 1)$ est donc génératrice de F . Elle est libre car formée d'un seul vecteur non nul¹ : c'est donc une base de F et $\dim F = 1$.

¹ Une famille formée d'un seul vecteur est libre... à condition que ce vecteur soit non nul !

Soit $P = aX^2 + bX + c \in \mathbf{R}_2[X]$. Alors $P \in G$ si et seulement si

$$(X - 1)(2aX + b) - 2aX = aX^2 + bX + c \Leftrightarrow 2aX^2 + bX - 2aX - b = aX^2 + bX + c.$$

Par identification des coefficients, c'est le cas si et seulement si

$$\begin{cases} 2a = a \\ b - 2a = b \\ -b = c \end{cases} \Leftrightarrow \begin{cases} a = 0 \\ c = 0 \end{cases}$$

Et donc $P \in G \Leftrightarrow P = bX$, de sorte que $G = \text{Vect}(X)$.

En particulier, G est bien un sous-espace vectoriel de $\mathbf{R}_2[X]$, de base X , de sorte que $\dim G = 1$.

Soit $M = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathcal{M}_2(\mathbf{R})$. Alors

$$MN = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} = \begin{pmatrix} a\lambda_1 & b\lambda_2 \\ c\lambda_2 & d\lambda_2 \end{pmatrix} \text{ et } NM = \begin{pmatrix} a\lambda_1 & b\lambda_1 \\ c\lambda_2 & d\lambda_2 \end{pmatrix}.$$

Ainsi, $M \in H$ si et seulement si

$$\begin{cases} a\lambda_1 = a\lambda_1 \\ b\lambda_1 = b\lambda_2 \\ c\lambda_1 = c\lambda_2 \\ d\lambda_2 = d\lambda_2 \end{cases} \Leftrightarrow \begin{cases} b(\lambda_1 - \lambda_2) = 0 \\ c(\lambda_1 - \lambda_2) = 0 \end{cases}$$

Remarque

Il n'y a donc aucune contrainte sur b .

Astuce

Si on arrive à écrire un ensemble sous forme d'un Vect, c'est automatiquement un sous-espace vectoriel.

Et puisque $\lambda_1 \neq \lambda_2$, ceci équivaut à $b = c = 0$.

Ainsi,

$$H = \left\{ \begin{pmatrix} a & 0 \\ 0 & d \end{pmatrix}, (a, d) \in \mathbf{R}^2 \right\} = \left\{ a \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + d \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}, (a, d) \in \mathbf{R}^2 \right\} = \text{Vect} \left(\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \right).$$

Donc H est un sous-espace vectoriel de $\mathcal{M}_2(\mathbf{R})$, dont une famille génératrice est $\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$.

Puisqu'il s'agit d'une famille de deux vecteurs de $\mathcal{M}_2(\mathbf{R})$ non colinéaires, elle est libre, et donc c'est une base de H .

SOLUTION DE L'EXERCICE 3.3

1. Il est clair que $F \subset \mathcal{M}_2(\mathbf{R})$, et la matrice nulle est dans F , pour $a = b = 0$.

Soient $M = \begin{pmatrix} a & 2a+b \\ -b & -a \end{pmatrix}$ et $N = \begin{pmatrix} a' & 2a'+b' \\ -b' & -a' \end{pmatrix}$ deux éléments de F , et soit $\lambda \in \mathbf{R}$. Alors

$$\lambda M + N = \lambda \begin{pmatrix} a & 2a+b \\ -b & -a \end{pmatrix} + \begin{pmatrix} a' & 2a'+b' \\ -b' & -a' \end{pmatrix} = \begin{pmatrix} \lambda a + a' & 2(\lambda a + a') + (\lambda b + b') \\ -(\lambda b + b') & -(\lambda a + a') \end{pmatrix} \in F.$$

Et donc F est un sous-espace vectoriel de $\mathcal{M}_2(\mathbf{R})$.

De plus, pour $\begin{pmatrix} a & 2a+b \\ -b & -a \end{pmatrix} \in F$, on a

$$\begin{pmatrix} a & 2a+b \\ -b & -a \end{pmatrix} = a \begin{pmatrix} 1 & 2 \\ 0 & -1 \end{pmatrix} + b \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \in \text{Vect} \left(\begin{pmatrix} 1 & 2 \\ 0 & -1 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \right).$$

Donc déjà, la famille $\left(\begin{pmatrix} 1 & 2 \\ 0 & -1 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \right)$ est génératrice de F . Puisqu'elle est formée de deux vecteurs non colinéaires², elle est libre. C'est donc une base de F .

² Il est clair que ces matrices ne sont pas proportionnelles.

On montre de même que G est un sous-espace vectoriel de $\mathcal{M}_2(\mathbf{R})$, qui possède pour base $\left(\begin{pmatrix} 1 & 3 \\ 0 & -2 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ -1 & 1 \end{pmatrix} \right)$.

2. D'après ce qui précède, on a $\dim F = \dim G = 2$, et donc

$$\dim \mathcal{M}_2(\mathbf{R}) = 4 = 2 + 2 = \dim F + \dim G.$$

Montrons donc que $F \cap G = \{0\}$.

Soit $M \in F \cap G$.

Alors $M \in F$, donc il existe deux réels a et b tels que $M = \begin{pmatrix} a & 2a+b \\ -b & -a \end{pmatrix}$, et de même, $M \in G$

donc il existe deux réels c et d tels que $M = \begin{pmatrix} c & 3c+d \\ -d & -2c+d \end{pmatrix}$.

Par identification des coefficients, il vient donc

$$\begin{cases} a = c \\ 2a + b = 3c + d \\ -b = -d \\ -a = -2c + d \end{cases} \iff \begin{cases} a = c \\ 2a + b = 3a + b \\ b = d \\ a = -2a + b \end{cases} \iff a = b = c = d = 0.$$

Et donc $M = 0$.

On en déduit que $F \cap G = \{0\}$, et donc F et G sont supplémentaires dans $\mathcal{M}_2(\mathbf{R})$.

SOLUTION DE L'EXERCICE 3.4

Notons que $E_\lambda(f)$ est non vide car il contient le vecteur nul : $f(0) = 0 = \lambda \cdot 0$.

Soient $u, v \in E_\lambda(f)$ et $\alpha \in \mathbf{R}$. Alors

$$f(\alpha u + v) = \alpha f(u) + f(v) = \alpha \lambda u + \lambda v = \lambda(\alpha u + v)$$

et donc $\alpha u + v \in E_\lambda(f)$.

Ainsi, $E_\lambda(f)$ est un sous-espace vectoriel de E .

Remarque

Nous ne sommes même pas obligés de résoudre entièrement le système : si on arrive à prouver que $a = b = 0$, alors nécessairement $M = 0$.

Alternative

On peut aussi remarquer que $E_\lambda(f) = \text{Ker}(f - \lambda \text{id}_E)$ et donc est automatiquement un sous-espace vectoriel de E .

SOLUTION DE L'EXERCICE 3.5

Notons que F est un sous-espace vectoriel de \mathbf{R}^4 car c'est le noyau de la forme linéaire $\varphi : (x, y, z, t) \mapsto x + y + 2z - t$.

Ceci prouve du même coup qu'il s'agit d'un hyperplan de \mathbf{R}^4 , et donc qu'il est de dimension 3. De plus, $(1, 1, 1, 1)$ est non nul donc forme une famille libre de G . Puisqu'elle est évidemment génératrice, c'est une base de G qui est donc de dimension 1.

Donc déjà, $\dim F + \dim G = 4 = \dim \mathbf{R}^4$.

Soit $x \in F \cap G$. Puisque $x \in G$, il existe $\lambda \in \mathbf{R}$ tel que $x = \lambda(1, 1, 1, 1) = (\lambda, \lambda, \lambda, \lambda)$.

Mais $x \in F$, donc $\lambda + \lambda + 2\lambda - \lambda = 0 \Leftrightarrow 3\lambda = 0 \Leftrightarrow \lambda = 0$.

Et donc $x = 0$, de sorte que $F \cap G = \{0\}$.

Par conséquent, F et G sont supplémentaires dans \mathbf{R}^4 .

Sinon

Si on n'a pas vu qu'une forme linéaire se cachait derrière F , on cherche une base à la main. C'est plus long et plus susceptible de mener à des erreurs de calcul, mais tout aussi juste.

SOLUTION DE L'EXERCICE 3.6

1. Prouvons que F est un sous-espace vectoriel de $\mathcal{C}(\mathbf{R}, \mathbf{R})$.

Par définition, $F \subset \mathcal{C}(\mathbf{R}, \mathbf{R})$, et la fonction nulle est dans F , puisqu'elle vaut 0 en $x = 0$ et en $x = 1$.

Soient f et g deux éléments de F , et soit $\lambda \in \mathbf{R}$. Alors $\lambda f + g$ est continue et

$$(\lambda f + g)(0) = \lambda f(0) + g(0) = \lambda f(1) + g(1) = (\lambda f + g)(1).$$

Ainsi, $\lambda f + g \in F$, et donc F est un sous-espace vectoriel de $\mathcal{C}(\mathbf{R}, \mathbf{R})$.

2. Nous allons prouver que toute fonction continue s'écrit de manière unique comme un élément de F plus un élément de $\text{Vect}(g)$.

Procédons par analyse-synthèse : soit $f \in \mathcal{C}(\mathbf{R}, \mathbf{R})$, et supposons que $f = f_1 + f_2$, avec $f_1 \in F$ et $f_2 \in \text{Vect}(g)$.

Il existe donc $\lambda \in \mathbf{R}$ tel que $f_1 = \lambda g$.

On a alors $f(0) = f_1(0) + \lambda \cdot 0 = f_1(0)$ et $f(1) = f_1(1) + \lambda = f_1(0) + \lambda$.

Et donc $f(1) = f(0) + \lambda \Leftrightarrow \lambda = f(1) - f(0)$.

On en déduit que $f_1 = f - f_2 = f - (f(0) - f(1)) \cdot g$.

Inversement, posons $f_2 = (f(1) - f(0)) \cdot g$ et $f_1 = f - f_2$.

Alors il est évident que $f_2 \in \text{Vect}(g)$, et $f_1 \in F$ car f_1 est continue car différence de fonctions continues, et

$$f_1(0) = f(0) - (f(1) - f(0)) \underbrace{g(0)}_{=0} = f(0) \text{ et } f_1(1) = f(1) - (f(1) - f(0)) \underbrace{g(1)}_{=1} = f(0) = f_1(0).$$

Enfin, on a bien

$$f_1 + f_2 = f - f_2 + f_2 = f.$$

Ainsi, f s'écrit de manière unique comme une fonction de F plus une fonction de $\text{Vect}(g)$, et donc $\mathcal{C}(\mathbf{R}, \mathbf{R}) = F \oplus \text{Vect}(g)$.

SOLUTION DE L'EXERCICE 3.7

Supposons que $g \circ f = 0$, et soit $x \in \text{Im } f$.

Par définition de $\text{Im } f$, il existe $y \in E$ tel que $x = f(y)$. Et alors

$$g(x) = g(f(y)) = \underbrace{(g \circ f)}_{=0}(y) = 0,$$

ce qui prouve donc que $x \in \text{Ker } g$.

Ainsi, tout élément de $\text{Im } f$ est dans $\text{Ker } g : \text{Im } f \subset \text{Ker } g$.

Supposons à présent que $\text{Im } f \subset \text{Ker } g$. Alors pour $x \in E$, on a $f(x) \in \text{Im } f$, et donc $f(x) \in \text{Ker } g$. On en déduit donc que $g(f(x)) = 0$.

Ainsi, pour tout $x \in E$, $(g \circ f)(x) = 0$, ce qui prouve que $g \circ f = 0$.

SOLUTION DE L'EXERCICE 3.8

Danger

Puisque $\mathcal{C}(\mathbf{R}, \mathbf{R})$ n'est pas de dimension finie (cf exercice précédent), il n'est pas question d'utiliser un critère à base de dimensions.

Détails

On a utilisé ici le fait que $f_1 \in F$, et donc $f_1(0) = f_1(1)$.

Explications

Nous venons de prouver que si f était somme d'un élément f_1 de F et d'un élément f_2 de $\text{Vect}(g)$, alors il n'y avait qu'une seule décomposition possible. Autrement dit, que f s'écrit **d'au plus** une manière comme un élément de F et un élément de $\text{Vect}(g)$. Il reste à vérifier l'existence d'une telle décomposition.

Méthode

Pour montrer une équivalence, sauf cas faciles, il est préférable de raisonner par double implication.

Notations

En général, on note plutôt $y = g(x)$. Tout dépend de comment vous avez appelé l'élément que vous choisissez dans $\text{Im } f$... L'essentiel est de retenir que $x \in \text{Im } f$ si et seulement si x est l'image par f d'un vecteur de E . Donc $x = f(y)$, $x = f(u)$, etc, peu importe, tant qu'on n'utilise pas deux fois la même lettre pour désigner deux objets différents.

1. p est clairement définie de \mathbb{C}^2 dans lui-même, donc il reste à prouver que p est linéaire. Soient $(x, y), (x', y') \in \mathbb{C}^2$, et soit $\lambda \in \mathbb{C}$. Alors

$$\begin{aligned} p(\lambda(x, y) + (x', y')) &= p(\lambda x + x', \lambda y + y') = \frac{1}{5}(\lambda x + x' + 2(\lambda y + y'), 2(\lambda x + x') + 4(\lambda y + y')) = \\ &= \lambda \frac{1}{5}(x + 2y, 2x + 4y) + \frac{1}{5}(x' + 2y', 2x' + 4y') = \lambda p(x, y) + p(x', y'). \end{aligned}$$

Ainsi p est linéaire : c'est un endomorphisme de \mathbb{C}^2 .

2. Par définition,

$$\text{Ker } p = \{(x, y) \in \mathbb{C}^2 : p(x, y) = (0, 0)\} = \left\{ (x, y) \in \mathbb{C}^2 : \begin{cases} x + 2y = 0 \\ 2x + 4y = 0 \end{cases} \right\} = \{(-2y, y), y \in \mathbb{C}\} = \text{Vect}(-2, 1).$$

On en déduit qu'une base de $\text{Ker } p$ est donnée par $(-2, 1)$. En particulier, $\text{Ker } p$ est de dimension 1, et donc p n'est pas injective.

3. D'après le théorème du rang, $\text{Im } p$ est de dimension 1, donc est engendré par n'importe lequel de ses vecteur non nuls. Par exemple, $p(5, 0) = (1, 2) \in \text{Im } p$, donc une base de $\text{Im } p$ est formée par $(1, 2)$.
4. Soit $(x, y) \in \mathbb{C}^2$. On a

$$\begin{aligned} p(p(x, y)) &= p\left(\frac{1}{5}(x + 2y, 2x + 4y)\right) \\ &= \frac{1}{5}p(x + 2y, 2x + 4y) = \frac{1}{25}(x + 2y + 2(2x + 4y), 2(x + 2y) + 4(2x + 4y)) \\ &= \frac{1}{25}(5x + 10y, 10x + 20y) = \frac{1}{5}(x + 2y, 2x + 4y) = p(x, y). \end{aligned}$$

Nous avons donc bien $p \circ p = p$.

p est donc un projecteur, de sorte que $\text{Ker } p$ et $\text{Im } p$ sont supplémentaires dans \mathbb{C}^2 .

SOLUTION DE L'EXERCICE 3.9

- 1.a. On a $x \in \text{Ker}(p - \text{id}_E) \Leftrightarrow p(x) - x = 0_E \Leftrightarrow p(x) = x$. Or p est la projection sur $\text{Im}(p)$ parallèlement à $\text{Ker}(p)$: pour $x = x_1 + x_2 \in E$, avec $x_1 \in \text{Im}(p)$ et $x_2 \in \text{Ker}(p)$, on a $p(x) = x_1$. Et donc $p(x) = x$ si et seulement si $x = x_1 \in \text{Im}(p)$. Ainsi, $\text{Im}(p) = \text{Ker}(p - \text{id}_E)$.

- 1.b. On a

$$(2p - \text{id}_E)^2 = 4p^2 - 4p \circ \text{id}_E + \text{id}_E = 4p - 4p + \text{id}_E = \text{id}_E.$$

Donc $2p - \text{id}_E$ est une symétrie.

- 2.a. De même, on a

$$\left(\frac{1}{2}(g + \text{id}_E)\right)^2 = \frac{1}{4}(g^2 + 2g + \text{id}_E) = \frac{1}{4}(\text{id}_E + 2g + \text{id}_E) = \frac{1}{2}(g + \text{id}_E).$$

Et donc $\frac{1}{2}(g + \text{id}_E)$ est un projecteur.

- 2.b. Puisque $\frac{1}{2}(g + \text{id}_E)$ est un projecteur, son image et son noyau sont supplémentaires dans E .

$$\text{Or } \text{Ker}\left(\frac{1}{2}(g + \text{id}_E)\right) = \text{Ker}(g + \text{id}_E).$$

D'autre part, d'après la question 1.a, on a

$$\text{Im}\left(\frac{1}{2}(g + \text{id}_E)\right) = \text{Ker}\left(\frac{1}{2}(g + \text{id}_E) - \text{id}_E\right) = \text{Ker}\left(\frac{1}{2}(g - \text{id}_E)\right) = \text{Ker}(g - \text{id}_E).$$

Et donc $E = \text{Ker}(g - \text{id}_E) \oplus \text{Ker}(g + \text{id}_E)$.

SOLUTION DE L'EXERCICE 3.10

Nous savons que $E = \text{Im } p \oplus \text{Ker } p$ et que p est la projection sur $\text{Im } p$ parallèlement à $\text{Ker } p$. Autrement dit, si $x \in E$, alors de manière unique, $x = x_1 + x_2$, avec $x_1 \in \text{Im } p$ et $x_2 \in \text{Ker } p$ et $p(x) = x_1$.

Alors $q(x) = \text{id}(x) - p(x) = x - x_1 = x_2$.

Ainsi, q est la projection sur $\text{Ker } p$ parallèlement à $\text{Im } p$, et donc $\text{Im } q = \text{Ker } p$.

On en déduit donc que $E = \text{Im } p \oplus \text{Ker } p = \text{Im } p \oplus \text{Im } q$.

Alternative

Nous pourrions également retrouver ce résultat en prouvant par exemple que la juxtaposition d'une base de $\text{Ker } p$ et d'une base de $\text{Im } p$ est une base de \mathbb{C}^2 .

Carré

On a développé $(2p - \text{id}_E)^2$ à l'aide du binôme de Newton, ce qui est légitime car id_E et p commutent (id_E commute à tout endomorphisme).

Notons qu'on aurait aussi pu écrire

$$\begin{aligned} (2p - \text{id}_E)^2 &= \\ (2p - \text{id}_E) \circ (2p - \text{id}_E) & \end{aligned}$$

et développer cette expression.

Détails

Pour tout $\lambda \in \mathbb{R}^*$ et tout $f \in \mathcal{L}(E)$, on a $\text{Ker}(\lambda f) = \text{Ker}(f)$ car $\lambda f(x) = 0 \Leftrightarrow f(x) = 0$.

Plus généralement

De la même manière, on prouve que si p est la projection sur F parallèlement à G , alors $\text{id} - p$ est la projection sur G parallèlement à F .

SOLUTION DE L'EXERCICE 3.11

1. On a $F \subset \mathbf{R}_4[X]$ par définition et le polynôme nul est dans F car il s'annule en 0. Soit $P, Q \in F$ et $\lambda \in \mathbf{R}$. Alors

$$(\lambda P + Q)(0) = \lambda P(0) + Q(0) = 0.$$

Ainsi, $\lambda P + Q \in F$ et donc F est un sous-espace vectoriel de $\mathbf{R}_4[X]$.

De plus, on a $P = a_0 + a_1X + a_2X^2 + a_3X^3 + a_4X^4 \in F$ si et seulement si $a_0 = 0$, c'est-à-dire si et seulement si $P = a_1X + a_2X^2 + a_3X^3 + a_4X^4 \in \text{Vect}(X, X^2, X^3, X^4)$.

Puisque la famille (X, X^2, X^3, X^4) est libre³, c'est donc une base de F .

³ C'est une sous-famille de la base canonique de $\mathbf{R}_4[X]$

2. On a $G \subset \mathbf{R}_4[X]$ par définition et le polynôme nul est dans G car il s'annule en 4. Soit $P, Q \in G$ et $\lambda \in \mathbf{R}$. Alors

$$(\lambda P + Q)(4) = \lambda P(4) + Q(4) = 0.$$

Ainsi, $\lambda P + Q \in G$ et donc G est un sous-espace vectoriel de $\mathbf{R}_4[X]$.

De plus, d'après la formule de Taylor pour les polynômes, nous savons que tout polynôme P de $\mathbf{R}_4[X]$ s'écrit

$$P = P(4) + P'(4)(X - 4) + \frac{P''(4)}{2}(X - 4)^2 + \frac{P^{(3)}(4)}{3}(X - 4)^3 + \frac{P^{(4)}(4)}{24}(X - 4)^4.$$

En particulier, pour un polynôme $P \in G$, on a

$$P = P(4) + \frac{P'(4)}{1}(X-4) + \frac{P''(4)}{2}(X-4)^2 + \frac{P^{(3)}(4)}{3}(X-4)^3 + \frac{P^{(4)}(4)}{24}(X-4)^4 \in \text{Vect}(X-4, (X-4)^2, (X-4)^3, (X-4)^4).$$

La famille $(X - 4, (X - 4)^2, (X - 4)^3, (X - 4)^4)$ est donc génératrice de G , et elle est libre car formée de polynômes de degrés deux à deux distincts : c'est une base de G .

3. Un polynôme P est dans H si et seulement si il possède 0 et 4 comme racines, c'est-à-dire si et seulement si il est divisible par $(X - 0)(X - 4)$, et donc de la forme $P = X(X - 4)Q$, $Q \in \mathbf{R}[X]$. Et puisque $\deg P = 2 + \deg Q \leq 4$, nécessairement $\deg Q \leq 2$: $Q \in \mathbf{R}_2[X]$. Ainsi, $H = \{X(X - 4) \times Q, Q \in \mathbf{R}_2[X]\}$.

En particulier, si $P \in H$, alors il existe $a_0, a_1, a_2 \in \mathbf{R}$ tels que

$$P = X(X - 4)(a_0 + a_1X + a_2X^2) = a_0X(X - 4) + a_1X^2(X - 4) + a_2X^3(X - 4) \in \text{Vect}(X(X - 4), X^2(X - 4), X^3(X - 4)).$$

Donc $(X(X - 4), X^2(X - 4), X^3(X - 4))$ est une famille génératrice de H , elle est libre car formée de polynômes de degrés deux à deux distincts : c'est une base de H .

Rappel

Une famille de polynômes non nuls de degrés deux à deux distincts est une famille libre de $\mathbf{R}[X]$.

dim Vect(u) = 1

u est une famille d'un seul vecteur non nul, donc libre. Comme elle est évidemment génératrice de $\text{Vect}(u)$, c'est une base de $\text{Vect}(u)$, qui est alors de dimension 1.

SOLUTION DE L'EXERCICE 3.12

Soit $u \notin \text{Ker } \varphi$. En particulier, $u \neq 0$, et donc $\text{Vect}(u) = \{\lambda u, \lambda \in \mathbf{R}\}$ est de dimension 1.

Puisque φ est une forme linéaire non nulle, son noyau⁴ est de dimension $n - 1$.

Ainsi, $\dim \text{Vect}(u) + \dim \text{Ker } \varphi = 1 + n - 1 = n = \dim E$.

Soit $x \in \text{Vect}(u) \cap \text{Ker } \varphi$. Alors il existe $\lambda \in \mathbf{R}$ tel que $x = \lambda u$. Mais $\varphi(x) = 0 \Leftrightarrow \varphi(\lambda u) = \lambda \varphi(u) = 0$.

Par hypothèse, $u \notin \text{Ker } \varphi$, donc $\varphi(u) \neq 0$, et donc $\lambda = 0$. On en déduit que $x = 0 \cdot u = 0_E$.

Ainsi $\text{Ker } \varphi \cap \text{Vect}(u) = \{0\}$.

On en déduit que $\text{Ker } \varphi$ et $\text{Vect}(u)$ sont supplémentaires dans E .

⁴ ce noyau est un hyperplan.

SOLUTION DE L'EXERCICE 3.13

1. Soient $P, Q \in \mathbf{R}_n[X]$ et $\lambda \in \mathbf{R}$. Alors

$$\begin{aligned} \varphi(\lambda P + Q) &= ((\lambda P + Q)(a_0), \dots, (\lambda P + Q)(a_n)) = (\lambda P(a_0) + Q(a_0), \dots, \lambda P(a_n) + Q(a_n)) = \\ &= \lambda(P(a_0), \dots, P(a_n)) + (Q(a_0), \dots, Q(a_n)) = \lambda \varphi(P) + \varphi(Q). \end{aligned}$$

Donc φ est bien linéaire.

2. On a $\text{Ker } \varphi = \{P \in \mathbf{R}_n[X] : P(a_0) = \dots = P(a_n) = 0\}$. Or, un polynôme de degré au plus n qui admet $n + 1$ racines (car les a_i sont tous distincts) est nécessairement le polynôme nul. Par conséquent, $\text{Ker } \varphi = \{0\}$, et φ est injectif. Mais $\dim \mathbf{R}_n[X] = n + 1 = \dim \mathbf{R}^{n+1}$. Par conséquent, φ est bijectif car injectif : c'est donc un isomorphisme.

3.a. On a $L_i(a_k) = \prod_{\substack{j=0 \\ j \neq i}}^n \frac{(a_k - a_j)}{a_i - a_j}$.

Si $k = i$, alors tous les $a_k - a_j$ sont non nuls⁵ et donc $L_i(a_i) = \prod_{\substack{j=0 \\ j \neq i}}^n \frac{(a_i - a_j)}{a_i - a_j} = 1$.

Au contraire si $k \neq i$, alors le produit contient un facteur $a_k - a_k = 0$, et donc $L_i(a_k) = 0$. Ainsi, $\varphi(L_i) = (L_i(a_0), \dots, L_i(a_n)) = (0, \dots, 0, 1, 0, \dots, 0)$, où le coefficient 1 se trouve en $i + 1$ -ème position.

- 3.b. Si on note e_i le i -ème vecteur de la base canonique de \mathbf{R}^{n+1} , alors $\varphi(L_i) = e_{i+1}$, soit encore $L_i = \varphi^{-1}(e_{i+1})$.
Et donc (L_0, \dots, L_n) est l'image de la base (e_1, \dots, e_{n+1}) par l'isomorphisme φ^{-1} : c'est une base de $\mathbf{R}_n[X]$.

SOLUTION DE L'EXERCICE 3.14

1. Montrons que $\text{Im}(g \circ f) \subset \text{Im } g$. Soit donc $y \in \text{Im}(g \circ f)$. Alors il existe $x \in E$ tel que $y = g(f(x))$. Et donc en particulier, $y \in \text{Im } g$.
Ainsi, $\text{Im}(g \circ f) \subset \text{Im } g$ et en passant aux dimensions, il vient

$$\dim \text{Im}(g \circ f) \leq \dim \text{Im } g \Leftrightarrow \text{rg}(g \circ f) \leq \text{rg}(g).$$

2. Soit $x \in \text{Im}(g \circ f)$. Alors il existe $y \in E$ tel que $x = g(f(y))$. Mais $f(y) \in \text{Im } f$ s'écrit (de manière unique) comme combinaison linéaire de e_1, \dots, e_p : il existe des scalaires $\lambda_1, \dots, \lambda_p$ tels que $f(y) = \sum_{i=1}^p \lambda_i e_i$. Et alors

$$x = g(f(y)) = g\left(\sum_{i=1}^p \lambda_i e_i\right) = \sum_{i=1}^p \lambda_i g(e_i).$$

Donc x est bien combinaison linéaire de $g(e_1), \dots, g(e_p)$.

On en déduit que $g(e_1), \dots, g(e_p)$ est une famille génératrice de $\text{Im}(g \circ f)$.

Et puisque le cardinal d'une famille génératrice est toujours supérieur ou égal à la dimension, $\text{rg}(g \circ f) = \dim \text{Im}(g \circ f) \leq p = \dim \text{Im } f = \text{rg}(f)$.

3. On a à la fois $\text{rg}(g \circ f) \leq \text{rg}(f)$ et $\text{rg}(g \circ f) \leq \text{rg}(g)$, donc

$$\text{rg}(g \circ f) \leq \min(\text{rg}(g), \text{rg}(f)).$$

SOLUTION DE L'EXERCICE 3.15

1. On a $u - v = f + 3 \text{id}_E - (f - 2 \text{id}_E) = 5 \text{id}_E$.
Ainsi, pour $x \in E$, on a

$$x = \text{id}_E(x) = \frac{1}{5}(u - v)(x) = \underbrace{\frac{1}{5}u(x)}_{\in \text{Im } u} + \underbrace{\left(-\frac{1}{5}v(x)\right)}_{\in \text{Im } v} \in \text{Im } u + \text{Im } v.$$

Donc $E \subset \text{Im } u + \text{Im } v$, et l'inclusion réciproque étant évidente⁶, $E = \text{Im } u + \text{Im } v$.

2. On a $u \circ v = (f + 3 \text{id}_E) \circ (f - 2 \text{id}_E) = f^2 + f - 6 \text{id}_E = 0$ et de même $v \circ u = (f - 2 \text{id}_E) \circ (f + 3 \text{id}_E) = f^2 + f - 6 \text{id}_E = 0$.

Alors en réutilisant le résultat de l'exercice 7, de $u \circ v = 0$ on déduit $\text{Im } v \subset \text{Ker } u$ et de $v \circ u = 0$, on déduit $\text{Im } u \subset \text{Ker } v$.

3. Il est évident que $\text{Ker } u + \text{Ker } v \subset E$ car $\text{Ker } u$ et $\text{Ker } v$ sont deux sous-espaces vectoriels de E .

Par la première question, si $x \in E$, alors $x = x_1 + x_2$, avec $x_1 \in \text{Im } u$ et $x_2 \in \text{Im } v$. Mais d'après la seconde question, on a alors $x_1 \in \text{Ker } v$ et $x_2 \in \text{Ker } u$. Donc $x \in \text{Ker } v + \text{Ker } u$, et alors $E \subset \text{Ker } u + \text{Ker } v$, de sorte que $E = \text{Ker } u + \text{Ker } v$.

Il reste donc à montrer que cette somme est directe : soit $x \in \text{Ker } u \cap \text{Ker } v$. Alors comme dans la question 1, $x = \underbrace{u(x)}_{=0} - \underbrace{v(x)}_{=0} = 0$. Donc $\text{Ker } u \cap \text{Ker } v = \{0\}$, et donc la somme est

directe. Ainsi

$$E = \text{Ker } u \oplus \text{Ker } v.$$

⁵ Car les a_i sont deux à deux distincts.

Rappel

L'image d'une base par un isomorphisme est une base de l'espace d'arrivée.

Méthode

Montrer qu'une famille est génératrice de E c'est prouver que tout vecteur de E peut s'écrire comme combinaison linéaire de vecteurs de \mathcal{F} .

Rappel

Être plus petit que le minimum de n nombres, c'est être plus petit que chacun de ces n nombres.

⁶ Car $\text{Im } u$ et $\text{Im } v$ sont deux sous-espaces vectoriels de E .

SOLUTION DE L'EXERCICE 3.16

1. Soit $x \in \text{Ker } f$. Alors $f(x) = 0$, et par conséquent $g(f(x)) = 0$. Mais

$$f(g(x)) = (f \circ g)(x) = (g \circ f)(x) = g(f(x)) = g(0) = 0.$$

On a donc $f(g(x)) = 0$, ce qui signifie que $g(x) \in \text{Ker } f$. On en déduit que $\text{Ker } f$ est stable par g .

2. Soit $y \in \text{Im } f$. Alors $\exists x \in E : y = f(x)$. Et donc

$$g(y) = g(f(x)) = (g \circ f)(x) = (f \circ g)(x) = f(g(x)).$$

Ceci prouve que $g(y) \in \text{Im } f$, et donc $\text{Im } f$ est stable par g .

SOLUTION DE L'EXERCICE 3.17

Il est classique⁷ que si f et p commutent, alors $\text{Ker } p$ et $\text{Im } p$ sont stables par f . Inversement, supposons que ces deux espaces soient stables par f . Alors puisque p est un projecteur, on a $E = \text{Ker } p \oplus \text{Im } p$.

Soit $x \in E$. De manière unique, $x = x_1 + x_2$, avec $x_1 \in \text{Im } p$ et $x_2 \in \text{Ker } p$.

Alors $f \circ p(x) = f(p(x_1)) = f(x_1)$ et $p \circ f(x) = p(f(x_1)) + p(f(x_2))$.

Mais $\text{Ker } p$ est stable par f , de sorte que $f(x_2) \in \text{Ker } p$, et donc $p(f(x_2)) = 0$.

De même, $\text{Im } p$ est stable, donc $f(x_1) \in \text{Im } p$. Et nous savons que pour tout élément y dans $\text{Im } p$, $p(y) = y$. Donc $p \circ f(x) = p(f(x_1)) = f(x_1)$.

Ainsi, $\forall x \in E, p \circ f(x) = f \circ p(x)$, donc $p \circ f = f \circ p : f$ et p commutent.

SOLUTION DE L'EXERCICE 3.18

1. Pour $k \in \llbracket 0, n \rrbracket$, notons $f_k : \mathbf{R}_n[X] \rightarrow \mathbf{R}$ la forme linéaire définie par $f_k(P) = P(k)$. Alors $\text{Ker } f_k$ est un sous-espace vectoriel de $\mathbf{R}_n[X]$, et

$$F_i = \bigcap_{\substack{k=0 \\ k \neq i}}^n \text{Ker } f_k$$

est donc également un sous-espace vectoriel de $\mathbf{R}_n[X]$.

Si $P \in F_i$, alors les $j \in \llbracket 0, n \rrbracket, j \neq i$ sont tous racines de P , donc P est divisible par $\prod_{\substack{j=0 \\ j \neq i}}^n (X - j)$.

Soit encore $P = Q \times \prod_{\substack{j=0 \\ j \neq i}}^n (X - j)$, avec $Q \in \mathbf{R}[X]$.

Mais $\deg P \leq n$ et $\deg \left(\prod_{\substack{j=0 \\ j \neq i}}^n (X - j) \right) = n$, de sorte que, $\deg Q = 0$.

Et donc Q est une constante $\lambda \in \mathbf{R} : P = \lambda \prod_{\substack{j=0 \\ j \neq i}}^n (X - j)$.

Inversement, puisque les $j \in \llbracket 0, n \rrbracket, j \neq i$ sont racines de $\prod_{\substack{j=0 \\ j \neq i}}^n (X - j)$, si $P = \lambda \prod_{\substack{j=0 \\ j \neq i}}^n (X - j)$, alors

$P \in F_i$.

2. Supposons que $P_0 + P_1 + \dots + P_n = 0$, avec $P_0 \in F_0, P_1 \in F_1, \dots, P_n \in F_n$. Alors, pour $i \in \llbracket 0, n \rrbracket$, en évaluant cette relation en $X = i$, il vient

$$P_0(i) + P_1(i) + \dots + P_n(i) = 0 \Rightarrow P_i(i) = 0.$$

Or, P_i est de la forme $P_i = \lambda_i \prod_{\substack{j=0 \\ j \neq i}}^n (X - j)$, de sorte que $P_i(i) = \lambda_i \underbrace{\prod_{\substack{j=0 \\ j \neq i}}^n (i - j)}_{\neq 0}$.

Et donc $\lambda_i = 0$, de sorte que $P_i = 0$.

Donc $P_0 = P_1 = \dots = P_n = 0 : la somme F_0 + \dots + F_n est directe.$

Rédaction

Il s'agit de montrer que quel que soit $x \in F, f(x) \in F$. Lorsqu'on écrit soit $x \in F$, cela signifie que l'on prend $x \in F$, quelconque. Si nous montrons que $f(x) \in F$, on aura bien montré que quel que soit $x \in F, f(x) \in F$.

⁷ Voir l'exercice précédent.

Astuce

Le noyau d'une application linéaire est toujours un sous-espace vectoriel de l'ensemble de départ. Notons que pour être parfaitement rigoureux, il faudrait prouver la linéarité des f_k , ce qui n'est pas bien dur, mais presque aussi long que d'utiliser le théorème de caractérisation pour montrer que F_i est un sev de $\mathbf{R}_n[X]$.

Méthode

Pour prouver que n sous-espaces vectoriels ($n > 2$) sont en somme directe, nous disposons de peu d'outils, et le plus simple est probablement d'utiliser la définition de somme directe.

Alternative

Il est également possible de remarquer que P_i possède déjà n racines (les $i \in \llbracket 0, n \rrbracket, i \neq j$) car il est dans F_i , et nous venons de montrer qu'il possède également i comme racine. Il est donc de degré au plus n et possède $n + 1$ racines : $P_i = 0$.

3. Nous avons prouvé à la question 1 que chacun des F_i est de dimension 1. Puisque la somme est directe, il vient

$$\dim \left(\bigoplus_{i=0}^n F_i \right) = \sum_{i=0}^n \dim F_i = n + 1.$$

Or, $\bigoplus_{i=0}^n F_i$ est un sous-espace vectoriel de $\mathbf{R}_n[X]$, de dimension $n + 1$: c'est $\mathbf{R}_n[X]$ tout entier :

$$\mathbf{R}_n[X] = \bigoplus_{i=0}^n F_i.$$

SOLUTION DE L'EXERCICE 3.19

On a $\dim E \leq \dim \operatorname{Im} f + \dim \operatorname{Im} g$ et de même, $\dim E \leq \dim \operatorname{Ker} f + \dim \operatorname{Ker} g$. Or, d'après le théorème du rang, nous savons que

$$\dim \operatorname{Ker} f = \dim E - \dim \operatorname{Im} f \text{ et } \dim \operatorname{Ker} g = \dim E - \dim \operatorname{Im} g.$$

Ainsi, il vient

$$\dim E \leq \dim E - \dim \operatorname{Im} f + \dim E - \dim \operatorname{Im} g \iff \dim \operatorname{Im} f + \dim \operatorname{Im} g \leq \dim E.$$

Puisque nous avons déjà l'inégalité inverse, on en déduit que $\dim E = \dim \operatorname{Im} f + \dim \operatorname{Im} g$, et donc $\dim(\operatorname{Im} f + \operatorname{Im} g) = \dim \operatorname{Im} f + \dim \operatorname{Im} g$.

Et donc $\operatorname{Im} f$ et $\operatorname{Im} g$ sont en somme directe.

Toujours avec le théorème du rang, on en déduit donc que $\dim \operatorname{Ker} f + \dim \operatorname{Ker} g = \dim E$, et donc que $\operatorname{Ker} f$ et $\operatorname{Ker} g$ sont en somme directe.

SOLUTION DE L'EXERCICE 3.20

- 1.a. Supposons qu'un tel polynôme L_i existe. Alors les $a_j, 1 \leq j \leq n, j \neq i$ sont racines de L_i , et donc L_i est divisible par $\prod_{\substack{j=1 \\ j \neq i}}^n (X - a_j)$: il existe $Q_i \in \mathbf{R}[X]$ tel que $L_i = Q_i \prod_{\substack{j=1 \\ j \neq i}}^n (X - a_j)$.

Mais alors $\deg L_i = \deg Q_i + (n - 1)$, et puisque $\deg L_i \leq n - 1, \deg Q_i \leq 0$.

Autrement dit, Q_i est une constante λ .

Et alors

$$1 = L_i(a_i) = \lambda \prod_{\substack{j=1 \\ j \neq i}}^n (a_i - a_j) \iff \lambda = \frac{1}{\prod_{\substack{j=1 \\ j \neq i}}^n (a_i - a_j)}.$$

$$\text{Et donc } L_i = \prod_{\substack{j=1 \\ j \neq i}}^n \frac{X - a_j}{a_i - a_j}.$$

Inversement, le polynôme $L_i = \prod_{\substack{j=1 \\ j \neq i}}^n \frac{X - a_j}{a_i - a_j}$ est bien de degré inférieur ou égal⁸ à $n - 1$, et

vérifie $\forall j \in \llbracket 1, n \rrbracket, L_i(a_j) = \delta_{i,j}$.

Donc il existe bien un unique polynôme de $\mathbf{R}_{n-1}[X]$ vérifiant ces conditions.

- 1.b. Soient $\lambda_1, \dots, \lambda_n$ des réels tels que $\sum_{i=1}^n \lambda_i L_i = 0$.

Alors pour $j \in \llbracket 1, n \rrbracket$, en évaluant cette relation en a_j , il vient

$$\sum_{i=1}^n \lambda_i L_i(a_j) = 0 \iff \sum_{i=1}^n \lambda_i \delta_{i,j} = 0 \iff \lambda_j = 0.$$

Et donc $\lambda_1 = \dots = \lambda_n = 0$: la famille (L_1, \dots, L_n) est libre.

Puisqu'elle est de cardinal $n = \dim \mathbf{R}_{n-1}[X]$, c'est une base de $\mathbf{R}_{n-1}[X]$.

Rappel

On a toujours

$$\dim(F + G) \leq \dim F + \dim G.$$

Et on égalité si et seulement si la somme est directe.

Remarque

Le fait que les a_k soient deux à deux distincts garantit que

$$\prod_{\substack{j=1 \\ j \neq i}}^n (a_i - a_j) \neq 0.$$

⁸ En réalité, il est de degré égal à $n - 1$.

2.a. Notons que pour tout $j \in \llbracket 1, n \rrbracket$, on a

$$\pi(P)(a_j) = \sum_{i=1}^n P(a_i)L_i(a_j) = \sum_{i=1}^n P(a_i)\delta_{i,j} = P(a_j).$$

Et donc

$$\pi(\pi(P)) = \sum_{i=1}^n \pi(P)(a_i)L_i = \sum_{i=1}^n P(a_i)L_i = \pi(P).$$

Puisque d'autre part la linéarité de π est évidente⁹, π est un projecteur de $\mathbf{R}[X]$.

⁹ L'évaluation des polynômes en a_i est linéaire.

2.b. Pour $P \in \mathbf{R}[X]$, on a $P \in \text{Ker}(\pi) \Leftrightarrow \sum_{i=1}^n P(a_i)L_i = 0$, et par liberté de la famille (L_1, \dots, L_n) ,

c'est le cas si et seulement si $P(a_1) = \dots = P(a_n) = 0$.

Autrement dit, $\text{Ker}(\pi)$ est formé des polynômes qui possèdent a_1, \dots, a_n comme racines.

C'est donc l'ensemble des polynômes divisibles par $\prod_{i=1}^n (X - a_i)$: c'est F .

D'autre part, il est évident que $\text{Im}(\pi) \subset \mathbf{R}_{n-1}[X]$.

Comme on a $\pi(L_i) = \sum_{j=1}^n L_i(a_j)L_j = L_i$, les L_i sont tous dans $\text{Im}(\pi)$ et donc $\mathbf{R}_{n-1}[X] =$

$\text{Vect}(L_1, \dots, L_n) \subset \text{Im}(\pi)$.

Et donc $\text{Im}(\pi) = \mathbf{R}_{n-1}[X]$.

2.c. Il serait possible de prouver le résultat souhaité par analyse-synthèse, en prouvant que tout polynôme de $\mathbf{R}[X]$ s'écrit de manière unique comme un élément de F plus un polynôme de $\mathbf{R}_{n-1}[X]$.

Mais puisque nous savons que π est un projecteur, on a

$$\mathbf{R}[X] = \text{Im}(\pi) \oplus \text{Ker}(\pi) = \mathbf{R}_{n-1}[X] \oplus F.$$

2.d. Si $P \in \mathbf{R}_{n-1}[X] = \text{Im}(\pi)$, alors $P = \pi(P) = \sum_{i=1}^n P(a_i)L_i$.

Et donc les coordonnées de P dans la base (L_1, \dots, L_n) sont $(P(a_1), \dots, P(a_n))$.

3.a. La linéarité de ε découle immédiatement de la linéarité de l'évaluation des polynômes.

Puisque $\mathbf{R}_{n-1}[X]$ et \mathbf{R}^n sont tous deux de dimension n , il suffit de montrer que ε est injective.

Soit donc $P \in \text{Ker } \varepsilon$. Alors

$$P(a_1) = P(a_2) = \dots = P(a_n) = 0.$$

Donc a_1, \dots, a_n sont n racines distinctes¹⁰ de P . Mais P est de degré au plus $n - 1$, et donc est le polynôme nul.

Ainsi $\text{Ker } \varepsilon = \{0\}$ et donc ε est un isomorphisme.

Alternative : puisque $\varepsilon(L_i) = (0, \dots, 0, 1, 0, \dots, 0)$, l'image de la base (L_0, L_1, \dots, L_n) est la base canonique de \mathbf{R}^n .

Et donc l'image d'une base de $\mathbf{R}_{n-1}[X]$ par ε est une base de \mathbf{R}^n , ce qui suffit à affirmer que ε est un isomorphisme.

3.b. Il s'agit de prouver que $(f(a_1), \dots, f(a_n))$ possède un unique antécédent par ε .

Mais puisque ε est bijectif, il existe un unique $P \in \mathbf{R}_{n-1}[X]$ tel que

$$\varepsilon(P) = (f(a_1), \dots, f(a_n)) \Leftrightarrow \forall i \in \llbracket 1, n \rrbracket, P(a_i) = f(a_i).$$

4.a. On a $\varphi(x) = f(x) - P(x) - K \prod_{i=1}^n (x - a_i)$.

Et donc $\varphi(x) = 0 \Leftrightarrow K = \frac{f(x) - P(x)}{\prod_{i=1}^n (x - a_i)}$.

¹⁰ Car les a_i le sont.

Rappel

Le seul polynôme qui possède strictement plus de racines que son degré est le polynôme nul.

4.b. Pour tout $i \in \llbracket 1, n \rrbracket$, $\varphi(a_i) = f(a_i) - P(a_i) - \prod_{j=1}^n (a_i - a_j) = f(a_i) - P(a_i) = 0$.

Donc φ s'annule en a_1, \dots, a_n et en x , qui sont bien $n + 1$ réels distincts.

Par le théorème de Rolle, il existe donc n réels distincts qui annulent φ' .

Puis, toujours par le théorème de Rolle, il existe $n - 1$ réels distincts qui annulent φ'' .

Et ainsi de suite, par application répétées du théorème de Rolle, il existe $\zeta \in]a, b[$ tel que $\varphi^{(n)}(\zeta) = 0$.

4.c. Puisque P est de degré au plus $n - 1$, $P^{(n)} = 0$.

D'autre part, $\prod_{i=1}^n (X - a_i)$ étant de degré n , sa dérivée $n^{\text{ème}}$ est égale à $n!$ fois son coefficient

dominant, c'est-à-dire $n! \times K$.

Et donc $\varphi^{(n)} = f^{(n)} - n! \times K$.

Par conséquent, $K = \frac{f^{(n)}(\zeta)}{n!}$.

Et donc

$$\begin{aligned} |f(x) - P(x)| &= \left| \underbrace{\varphi(x)}_{=0} + K \prod_{i=1}^n (x - a_i) \right| \\ &= |K| \prod_{i=1}^n |x - a_i| \\ &= \frac{|f^{(n)}(\zeta)|}{n!} \prod_{i=1}^n |x - a_i| \\ &\leq \sup_{x \in [a, b]} |f^{(n)}| \frac{\prod_{i=1}^n |x - a_i|}{n!}. \end{aligned}$$

Dérivée $n^{\text{ème}}$

Si besoin, dérivez n fois

$$a_n X^n + \dots + a_1 + a_0$$

pour vous convaincre que la dérivée $n^{\text{ème}}$ d'un polynôme de degré n vaut $n!$ fois son coefficient dominant.

Attention : ceci ne vaut que si on dérive autant de fois que le degré.

INTÉGRALES IMPROPRES

Cours	103
4.1 Rappels d'intégration sur un segment	103
4.2 Rappels sur les intégrales convergentes	106
4.3 Techniques de calcul d'intégrales	113
4.4 Une nouvelle fonction : la fonction Γ	117
4.5 Compléments	118
Méthodes	123
Exercices	125
Correction des exercices	128

4.1 RAPPELS D'INTÉGRATION SUR UN SEGMENT

Rappel : un segment de \mathbf{R} est un intervalle de la forme $[a, b]$, avec $a, b \in \mathbf{R}$.

4.1.1 Définition, premières propriétés

En particulier

Un segment est un intervalle fermé des deux côtés. Par exemple, $[1, +\infty[$ et $]1, 2]$ ne sont pas des segments.

Proposition 4.1 : Soit f une fonction continue sur un segment $[a, b]$. Alors f admet des primitives sur $[a, b]$.

De plus si F et G sont deux primitives de f , alors il existe une constante $\lambda \in \mathbf{R}$:
 $\forall x \in [a, b], F(x) = G(x) + \lambda$.

Définition 4.2 – Soit f une fonction continue sur un segment I . Alors si F est une primitive de f , on pose, pour tout $a, b \in I$,

$$\int_a^b f(t) dt = [F(t)]_a^b = F(b) - F(a).$$

Cette quantité ne dépend pas de la primitive F choisie, et est appelée intégrale de f sur le segment $[a, b]$.

Proposition 4.3 : Soient f et g deux fonctions continues sur I et soit $\lambda \in \mathbf{R}$. Alors, pour tout $a, b \in I$:

- $\int_a^b (\lambda f(t) + g(t)) dt = \lambda \int_a^b f(t) dt + \int_a^b g(t) dt$ (linéarité de l'intégrale)
- $\int_a^b f(t) dt = - \int_b^a f(t) dt$
- pour tout $c \in I$, $\int_a^b f(t) dt = \int_a^c f(t) dt + \int_c^b f(t) dt$ (relation de Chasles).
- si $a \leq b$ et si $\forall t \in [a, b], f(t) \geq 0$, alors $\int_a^b f(t) dt \geq 0$ (positivité de l'intégrale).
- si $a \leq b$ et si $\forall t \in [a, b], f(t) \leq g(t)$, alors $\int_a^b f(t) dt \leq \int_a^b g(t) dt$ (croissance de l'intégrale).

Théorème 4.4 (Théorème fondamental de l'analyse) : Soit I un intervalle de \mathbf{R} , et soit $f : I \rightarrow \mathbf{R}$ une fonction continue sur I .

Alors l'application $F : x \mapsto \int_a^x f(t) dt$ est de classe \mathcal{C}^1 sur I , et

$$\forall x \in I, F'(x) = f(x).$$

Plus précisément, F est l'unique primitive de f sur I qui s'annule en a .

Démonstration. Nous savons¹ que f admet une primitive G , qui est donc de classe \mathcal{C}^1 sur I .

Et on a alors $F(x) = \int_a^x f(t) dt = [G(t)]_a^x = G(x) - G(a)$.

Et donc F est \mathcal{C}^1 puisque G l'est, et on a $F'(x) = G'(x) = f(x)$, de sorte que F est une primitive de f .

De plus, $F(a) = \int_a^a f(t) dt = 0$, et comme toute autre primitive de f diffère de F par une constante, on peut bien affirmer que F est l'unique primitive de f qui s'annule en a . \square

¹ Même s'il s'agit d'un résultat qui a été admis.

Rappel

Deux primitives de f qui coïncident en un point sont égales.

Proposition 4.5 : Soit $a < b$, et soit f continue sur $[a, b]$. Si f est positive sur $[a, b]$, alors

$$\int_a^b f(t) dt = 0 \Leftrightarrow \forall t \in [a, b], f(t) = 0.$$

Démonstration. Il est évident que si f est nulle sur $[a, b]$, alors $\int_a^b f(t) dt = 0$.

Raisonnons par contraposée, et prouvons que sous les hypothèses de l'énoncé, s'il existe

$x_0 \in]a, b[$ tel que $f(x_0) > 0$, alors $\int_a^b f(t) dt > 0$.

Soit donc $\varepsilon = \frac{f(x_0)}{2} > 0$.

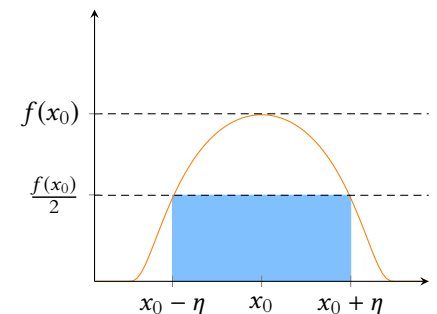
Par continuité de f en x_0 , il existe $\eta > 0$ tel que

$$\forall x \in [a, b] \cap [x_0 - \eta, x_0 + \eta], |f(x) - f(x_0)| \leq \varepsilon = \frac{f(x_0)}{2}.$$

En particulier, $\forall x \in [a, b] \cap [x_0 - \eta, x_0 + \eta], f(x) \geq \frac{f(x_0)}{2} > 0$.

Donc $\int_{x_0-\eta}^{x_0+\eta} f(t) dt \geq \int_{x_0-\eta}^{x_0+\eta} \frac{f(x_0)}{2} dt = \eta f(x_0) > 0$.

Et puisque f est positive sur $[a, b]$, $\int_a^b f(t) dt \geq \int_{x_0-\eta}^{x_0+\eta} f(t) dt > 0$. \square



L'aire du rectangle vaut $\eta f(x_0) > 0$.

Hypothèses !

On vérifiera bien **toutes** les hypothèses : ce théorème ne fonctionne plus si f est seulement continue par morceaux, et encore moins si on ne suppose pas f positive.

4.1.2 Primitives usuelles

Proposition 4.6 : Soit u une fonction de classe \mathcal{C}^1 sur un intervalle I . Alors

- une primitive de $u'u^n$, $n \in \mathbf{N}$ est $\frac{1}{n+1}u^{n+1}$
- une primitive de $u'e^u$ est e^u
- si u ne s'annule pas sur I , une primitive de $\frac{u'}{u}$ est $\ln(|u|)$
- si $u > 0$ sur I , et si $\alpha \neq -1$ alors une primitive de $u'u^\alpha$ est $\frac{1}{\alpha+1}u^{\alpha+1}$

Les primitives suivantes doivent être connues sans hésitations :

Fonction $x \mapsto \dots$	Primitive $x \mapsto \dots$	Intervalle
$x^n, n \in \mathbf{N}$	$\frac{1}{n+1}x^{n+1}$	\mathbf{R}
$\frac{1}{x}$	$\ln(x)$	\mathbf{R}_+^* ou \mathbf{R}_-^*
$x^n, n \in \mathbf{Z}, n \leq -2$	$\frac{1}{n+1}x^{n+1}$	\mathbf{R}_+^* ou \mathbf{R}_-^*
$\frac{1}{\sqrt{x}}$	$2\sqrt{x}$	\mathbf{R}_+^*
$x^\alpha, \alpha \in \mathbf{R} - \{-1\}$	$\frac{1}{\alpha+1}x^{\alpha+1}$	\mathbf{R}_+^*
e^x	e^x	\mathbf{R}
$\sin(x)$	$-\cos(x)$	\mathbf{R}
$\cos(x)$	$\sin(x)$	\mathbf{R}
$\tan(x)$	$-\ln(\cos(x))$	$]-\frac{\pi}{2} + k\pi, \frac{\pi}{2} + k\pi[, k \in \mathbf{Z}$
$\frac{1}{1+x^2}$	$\text{Arctan}(x)$	\mathbf{R}
$\ln(x)$	$x \ln(x) - x$	\mathbf{R}_+^*

▲ Dérivée/primitive

Toutes ces formules sont en fait des formules de dérivation. On notera que les formules pour la dérivée d'un produit et d'une composée ont leur interprétation en terme d'intégration : ce sont respectivement l'intégration par parties et le changement de variable. En revanche, on notera que bien que la dérivation soit quelque chose de totalement algorithmique (on peut dériver toute fonction obtenue à l'aide de fonctions usuelles et les dérivées des fonctions usuelles), la recherche de primitive n'est pas toujours facile et demande bien plus de pratique afin de reconnaître les formes les plus classiques.

— Astuce

Plutôt que d'apprendre cette formule par cœur, il suffit d'être capable de la retrouver à l'aide de la figure 4.2.

4.1.3 Sommes de Riemann

Proposition 4.7 : Soit f une fonction continue sur $[a, b]$. Alors

$$\lim_{n \rightarrow +\infty} \frac{b-a}{n} \sum_{k=0}^{n-1} f\left(a + k \frac{b-a}{n}\right) = \int_a^b f(t) dt.$$

Remarques. C'est ce résultat qui justifie l'interprétation de l'intégrale comme une aire dans le cas où f est positive.

En effet, $\frac{b-a}{n} f\left(a + k \frac{b-a}{n}\right)$ est l'aire d'un rectangle de largeur $\frac{b-a}{n}$ et de hauteur $f\left(a + k \frac{b-a}{n}\right)$.

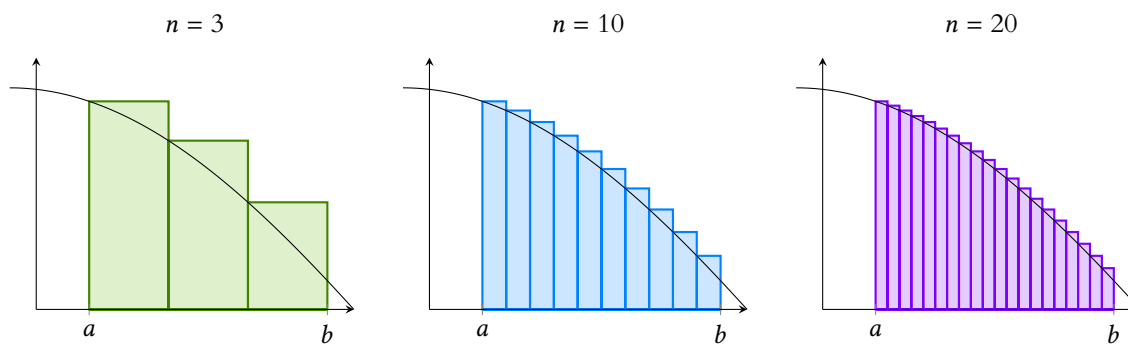


FIGURE 4.1 – Illustration de la convergence des sommes de Riemann vers l'intégrale. On a ici représenté les sommes de Riemann pour $n = 3, n = 10$ et $n = 20$.

Le même résultat reste valable si l'on considère plutôt les sommes²

$$\frac{b-a}{n} \sum_{k=1}^n f\left(a + k \frac{b-a}{n}\right).$$

² parfois appelées sommes de Riemann «à droite».

Ces rectangles sont en fait ceux représentés à droite sur la figure 4.2.

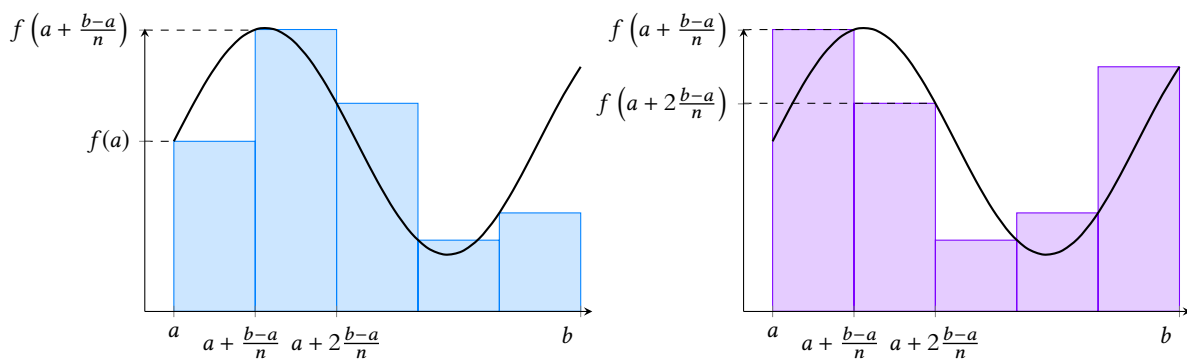


FIGURE 4.2 – Les deux types de sommes de Riemann.

Exemple 4.8

Pour $n \geq 1$, posons $u_n = \sum_{k=1}^n \frac{n}{n^2 + k^2}$. Alors,

$$u_n = \frac{1}{n^2} \sum_{k=1}^n \frac{n}{1 + \left(\frac{k}{n}\right)^2} = \frac{1-0}{n} \sum_{k=1}^n \frac{1}{1 + \left(\frac{k(1-0)}{n}\right)^2}.$$

Ainsi, en posant $f(t) = \frac{1}{1+t^2}$, on a

$$u_n = \frac{1}{n} \sum_{k=1}^n f\left(0 + k \frac{(1-0)}{n}\right)$$

et on reconnaît là des sommes de Riemann³, de sorte que

$$\lim_{n \rightarrow \infty} u_n = \int_0^1 f(t) dt = [\text{Arctan}(t)]_0^1 = \frac{\pi}{4}.$$

Méthode

La plupart du temps, la première difficulté si l'on n'a pas d'indication est de reconnaître qu'on a affaire à une somme de Riemann, et de savoir pour quels a et b . Si on ne les devine pas au premier abord, on essaiera $a = 0$ et $b = 1$.

³ Avec $a = 0$ et $b = 1$.

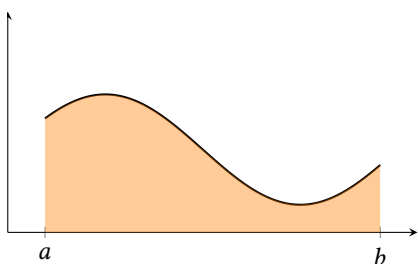
4.2 RAPPELS SUR LES INTÉGRALES CONVERGENTES

Si f est une fonction continue sur $[a, b]$, l'intégrale $\int_a^b f(t) dt$ est toujours bien définie, ce qui revient à dire que l'aire de la partie située sous la courbe représentative de f a bien un sens.

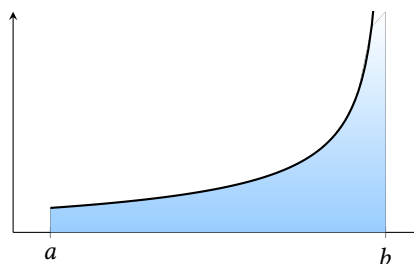
En revanche, si f n'est continue que sur $[a, b[$, il se peut que $\lim_{t \rightarrow b^-} f(t) = \pm\infty$. Et alors l'aire de la partie située sous la courbe, qui est une partie «infinie», n'est pas forcément bien définie.

Si c'est le cas, on dira que $\int_a^b f(t) dt$ est une intégrale convergente.

f continue sur $[a, b]$.



f continue sur $[a, b[$



La partie bleue est infinie, et donc son aire n'est pas forcément finie.

Notons que puisqu'il est impossible de représenter entièrement cette partie, un dessin ne permet pas de dire si cette aire est finie ou non, il faut nécessairement passer par le calcul.

4.2.1 Définition

Définition 4.9 – Soit f une fonction continue sur $[a, b]$, $a \in \mathbf{R}, b \in \mathbf{R} \cup \{+\infty\}$.

Si $\lim_{x \rightarrow b^-} \int_a^x f(t) dt$ existe, on dit que l'intégrale $\int_a^b f(t) dt$ converge, et on note

$$\int_a^b f(t) dt = \lim_{x \rightarrow b^-} \int_a^x f(t) dt.$$

Exemple 4.10

Soit $f(t) = \frac{1}{t^2}$. Alors $\forall x \geq 1$,

$$\int_1^x \frac{dt}{t^2} = \left[-\frac{1}{t} \right]_1^x = 1 - \frac{1}{x} \xrightarrow{x \rightarrow \infty} 1.$$

Donc $\int_1^{+\infty} \frac{dt}{t^2}$ existe et vaut 1.

Remarque. Si $b \in \mathbf{R}$ et si f se prolonge sur $[a, b]$ en une fonction continue \tilde{f} , alors $\forall x \in [a, b]$,

$$\int_a^x f(t) dt = \int_a^x \tilde{f}(t) dt \xrightarrow{x \rightarrow b^-} \int_a^b \tilde{f}(t) dt$$

Donc $\int_a^b f(t) dt$ existe⁴ et vaut $\int_a^b \tilde{f}(t) dt$.

On parle alors d'intégrale faussement impropre.

Exemple 4.11

La fonction $t \mapsto \frac{\sin t}{t}$ continue sur \mathbf{R}_+^* est prolongeable en une fonction continue en 0.

Donc $\int_{-\pi}^0 \frac{\sin t}{t} dt$ est une intégrale faussement impropre, qui est donc convergente.

De même, on définit les intégrales sur un segment de la forme $]a, b]$, $a \in \mathbf{R} \cup \{-\infty\}$ comme limites (si elles existent) lorsque x tend vers a par valeurs supérieures de $\int_x^b f(t) dt$.

Enfin, si f est une fonction continue sur un intervalle $]a, b[$, $a, b \in \overline{\mathbf{R}}$, l'intégrale $\int_a^b f(t) dt$ est dite convergente si pour un (et donc pour tous) $c \in]a, b[$, les deux intégrales

$$\int_a^c f(t) dt \text{ et } \int_c^b f(t) dt$$

convergent. On pose alors $\int_a^b f(t) dt = \int_a^c f(t) dt + \int_c^b f(t) dt$.

Exemple 4.12 Convergence et calcul de $\int_{-\infty}^{+\infty} \frac{dt}{1+t^2}$.

Pour tout $x > 0$, on a

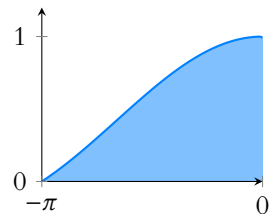
$$\int_{-x}^0 \frac{dt}{1+t^2} = [\text{Arctant}]_{-x}^0 = -\text{Arctan}(-x) \text{ et } \int_0^x \frac{dt}{1+t^2} = [\text{Arctan}]_0^x = \text{Arctan}(x).$$

Donc, en faisant tendre x vers $+\infty$, les deux intégrales $\int_{-\infty}^0 \frac{dt}{1+t^2}$ et $\int_0^{+\infty} \frac{dt}{1+t^2}$

Danger !

On ne parle de fonction prolongeable par continuité en b que si $b \in \mathbf{R}$, tout ceci ne peut pas marcher en $\pm\infty$.

⁴ L'intégrale de \tilde{f} est l'intégrale d'une fonction continue sur un segment, donc elle existe.



$\frac{\sin t}{t}$ est prolongeable par continuité en 0, donc l'aire sous la courbe existe.

Précision

On a choisi ici de couper l'intégrale en 0 ce qui semble plus logique, mais totalement arbitraire, on aurait pu choisir de la couper en -1 ou en $\sqrt{\pi}$.

Les problèmes de convergence éventuels sont en $-\infty$ et $+\infty$, partout ailleurs, nous savons que les intégrales convergent, car la fonction étant continue, on est ramené à des intégrales sur un segment !

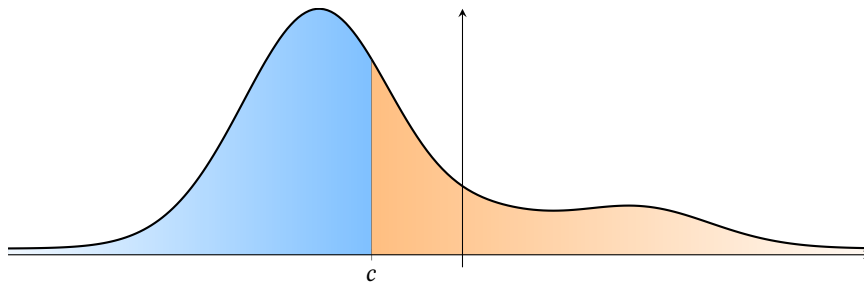


FIGURE 4.3 – Par définition, l'intégrale entre a et b converge si les deux intégrales entre a et c et entre c et b convergent. Graphiquement, cela signifie que les deux aires représentées ci-dessus sont finies.

On constate notamment que ceci ne dépend pas du choix de $c \in]a, b[$.

convergent et

$$\int_{-\infty}^0 \frac{dt}{1+t^2} = \int_0^{+\infty} \frac{dt}{1+t^2} = \frac{\pi}{2}.$$

Par conséquent, $\int_{-\infty}^{+\infty} \frac{dt}{1+t^2}$ converge et

$$\int_{-\infty}^{+\infty} \frac{dt}{1+t^2} = \int_{-\infty}^0 \frac{dt}{1+t^2} + \int_0^{+\infty} \frac{dt}{1+t^2} = \pi.$$

⚠ Bien que les intégrales ressemblent beaucoup aux séries, il se peut que $\int_1^{+\infty} f(t) dt$ converge et que f n'admette pas de limite en $+\infty$. Par conséquent, il n'y a pas pour les intégrales d'analogue à la divergence grossière des séries.

Plus généralement, nous pouvons être amenés à considérer des fonctions qui ne sont pas continues sur un intervalle, mais sur un intervalle privé d'un nombre fini de points.

Définition 4.13 – Soient $a < b \in \overline{\mathbf{R}}$, et soit $f :]a, b[\rightarrow \mathbf{R}$. On suppose qu'il existe des réels a_1, \dots, a_n tels que les restrictions de f à $]a, a_1[$, $]a_1, a_2[$, \dots , $]a_n, b[$ soient continues et que les intégrales

$$\int_a^{a_1} f(t) dt, \int_{a_1}^{a_2} f(t) dt, \dots, \int_{a_n}^b f(t) dt$$

soient convergentes.

On dit alors que $\int_a^b f(t) dt$ est convergente, et on pose

$$\int_a^b f(t) dt = \int_a^{a_1} f(t) dt + \int_{a_1}^{a_2} f(t) dt + \dots + \int_{a_n}^b f(t) dt.$$

Exemple 4.14 Densité de la loi exponentielle

Soit $f : \mathbf{R} \rightarrow \mathbf{R}$ définie par $f(t) = \begin{cases} 0 & \text{si } t < 0 \\ e^{-t} & \text{si } t \geq 0 \end{cases}$.

Alors f est continue sur $\mathbf{R}_-^* =]-\infty, 0[$ et sur $\mathbf{R}_+^* =]0, +\infty[$. Les deux intégrales

$\int_{-\infty}^0 f(t) dt$ et $\int_0^{+\infty} f(t) dt$ sont convergentes, donc $\int_{-\infty}^{+\infty} f(t) dt$ est convergente

$$\int_{-\infty}^{+\infty} f(t) dt = \int_{-\infty}^0 f(t) dt + \int_0^{+\infty} f(t) dt = 1.$$

4.2.2 Intégrale des fonctions positives

Proposition 4.15 : Soient $f, g :]a, b[\rightarrow \mathbf{R}$ deux fonctions continues sur $]a, b[$, ($a, b \in \overline{\mathbf{R}}$) telles que

$$\forall t \in]a, b[, 0 \leq f(t) \leq g(t).$$

Alors :

- Si $\int_a^b g(t) dt$ converge, alors $\int_a^b f(t) dt$ converge.
- Si $\int_a^b f(t) dt$ diverge, alors $\int_a^b g(t) dt$ diverge.

Proposition 4.16 : Soit f, g deux fonctions continues sur $]a, b[, b \in \overline{\mathbf{R}}$, positives, telles que $f(t) \underset{t \rightarrow b}{\sim} g(t)$. Alors les intégrales

$$\int_a^b f(t) dt \text{ et } \int_a^b g(t) dt$$

sont de même nature.

Remarque. Ce résultat reste en fait valable pour f et g de signe constant sur $]a, b[$.

Exemples 4.17

- Étude de $\int_0^1 \frac{dt}{e^t - 1}$.

La fonction $t \mapsto \frac{1}{e^t - 1}$ est continue sur $]0, 1[$, donc il faut étudier la nature de l'intégrale au voisinage de 0.

On a $e^t - 1 \underset{t \rightarrow 0}{\sim} t$, donc $f(t) \underset{t \rightarrow 0}{\sim} \frac{1}{t}$, et puisqu'il s'agit de fonctions positives, on peut utiliser le critère des équivalents.

Nous savons que $\int_0^1 \frac{dt}{t}$ diverge, donc il en est de même de $\int_0^1 \frac{dt}{e^t - 1}$.

- Étude de $\int_1^{+\infty} \frac{dt}{t\sqrt{t-1}}$.

La fonction $f : t \mapsto \frac{1}{t\sqrt{t-1}}$ est continue et positive sur $]1; +\infty[$.

Au voisinage de 1, on a $f(t) \underset{t \rightarrow 1}{\sim} \frac{1}{\sqrt{t-1}}$, et $\int_1^2 \frac{dt}{\sqrt{t-1}}$ est une intégrale de Riemann convergente, donc $\int_1^2 f(t) dt$ est convergente.

Au voisinage de $+\infty$, on a $f(t) \underset{t \rightarrow +\infty}{\sim} \frac{1}{t\sqrt{t}} = \frac{1}{t^{3/2}}$. Or l'intégrale $\int_2^{+\infty} \frac{1}{t^{3/2}} dt$ est convergente, donc par comparaison, il en est de même de $\int_2^{+\infty} f(t) dt$.

En conclusion, $\int_1^{+\infty} \frac{dt}{t\sqrt{t-1}}$ est convergente.

Proposition 4.18 : Soit f une fonction continue et positive sur $]a, b[$ telle que $\int_a^b f(t) dt$ converge. Alors $\int_a^b f(t) dt = 0$ si et seulement si $\forall x \in]a, b[, f(x) = 0$.

Hypothèses

Ce résultat n'est plus valable pour des fonctions qui changent de signe, donc lorsqu'on l'utilise, on indiquera bien qu'on le sait, et qu'on a vu que les fonctions en question sont positives.

Et même...

Ce résultat reste valable pour f et g de signe constant au voisinage de b . En effet, le problème de convergence est en b , ce qui se passe «loin de b » importe peu.

Méthode

En écrivant précisément où la fonction est continue, on identifie les points où il faut étudier la convergence de l'intégrale : ce sont ceux où le domaine de convergence est ouvert.

Démonstration. Il est évident que si f est nulle, alors l'intégrale est nulle. Supposons à présent que $\int_a^b f(t) dt = 0$. Soit $x \in [a, b]$. Alors

$$0 \leq \int_a^x f(t) dt \leq \int_a^b f(t) dt = 0$$

Donc $\int_a^x f(t) dt = 0$, et f étant continue et positive sur le segment $[a, x]$, on en déduit que f est nulle sur $[a, x]$. Ceci étant vrai pour tout $x \in [a, b]$, f est nulle sur $[a, b]$. \square

4.2.3 Intégrales de référence

Les intégrales suivantes sont les intégrales de référence, à connaître par cœur.

Théorème 4.19 (Intégrales de Riemann) : Soit $\alpha \in \mathbf{R}$. Alors

• $\int_1^{+\infty} \frac{dt}{t^\alpha}$ converge si et seulement si $\alpha > 1$, et on a alors

$$\int_1^{+\infty} \frac{dt}{t^\alpha} = \frac{1}{\alpha - 1}.$$

• $\int_0^1 \frac{dt}{t^\alpha}$ converge si et seulement si $\alpha < 1$, on a alors

$$\int_0^1 \frac{dt}{t^\alpha} = \frac{1}{1 - \alpha}.$$

• Plus généralement, pour $a < b$, $\int_a^b \frac{dt}{(t-a)^\alpha}$ et $\int_a^b \frac{dt}{(b-t)^\alpha}$ convergent si et seulement si $\alpha < 1$.

Valeurs

◀ S'il faut être capable de dire sans hésitations si une intégrale de Riemann converge ou diverge, il n'est pas nécessaire de connaître par cœur la valeur, mieux vaut être capable de la retrouver (la primitive est facile à calculer) plutôt que de prendre le risque de se tromper.

Démonstration. Les deux premiers points ont été prouvés en première année, prouvons le troisième.

• Si $\alpha \neq 1$, pour $x \in]a, b]$, on a

$$\int_x^b \frac{dt}{(t-a)^\alpha} = \left[\frac{1}{1-\alpha} (t-a)^{1-\alpha} \right]_x^b = \frac{1}{1-\alpha} \left(\frac{1}{(b-a)^{\alpha-1}} - \frac{1}{(x-a)^{\alpha-1}} \right).$$

Si $\alpha > 1$, alors $\alpha - 1 > 0$, et donc $\frac{1}{(x-a)^{\alpha-1}} \xrightarrow{x \rightarrow a^+} +\infty$, et donc $\int_x^b \frac{dt}{(t-a)^\alpha} \xrightarrow{x \rightarrow a^+} +\infty$.

Si $\alpha < 1$, alors $\alpha - 1 < 0$ et donc $\frac{1}{(x-a)^{\alpha-1}} = (x-a)^{1-\alpha} \xrightarrow{x \rightarrow a^+} 0$, de sorte que

$$\int_x^b \frac{dt}{(t-a)^\alpha} \xrightarrow{x \rightarrow a^+} \frac{(b-a)^{1-\alpha}}{1-\alpha}.$$

Enfin, pour $\alpha = 1$, on a

$$\int_x^b \frac{dt}{t-a} = [\ln(t-a)]_x^b = \ln(b-a) - \ln(x-a) \xrightarrow{x \rightarrow a^+} +\infty.$$

Et donc l'intégrale converge si et seulement si $\alpha < 1$. \square

Exemples 4.20

- $\int_0^1 \frac{dt}{t}$ et $\int_1^{+\infty} \frac{dt}{t}$ sont toutes deux divergentes.
- $\int_1^{+\infty} \frac{dt}{t\sqrt{t}}$ est convergente ($\alpha = \frac{3}{2} > 1$).
- $\int_1^3 \frac{dt}{\sqrt{t-1}}$ converge, car il s'agit d'une intégrale de Riemann convergente (avec $\alpha = \frac{1}{2} < 1$).

Proposition 4.21 : Soit $a \in \mathbf{R}$. Alors l'intégrale $\int_0^{+\infty} e^{-at} dt$ converge si et seulement si $a > 0$. Dans ce cas, on a alors

$$\int_0^{+\infty} e^{-at} dt = \frac{1}{a}.$$

Valeur

Là encore, nul besoin d'apprendre par cœur cette valeur, puisqu'une primitive de $t \mapsto e^{-at}$ est facile à retrouver.

4.2.4 Intégrales absolument convergentes

Rappelons que la fonction $x \mapsto |x|$ est continue sur \mathbf{R} , de sorte que si f est continue sur $]a, b[$, alors $|f|$ est également continue sur $]a, b[$.

Définition 4.22 – Soit $f :]a, b[\rightarrow \mathbf{R}$, avec $a, b \in \overline{\mathbf{R}}$. Si $\int_a^b |f(t)| dt$ est convergente, on dit que l'intégrale $\int_a^b f(t) dt$ est absolument convergente.

Remarque. Si f est une fonction positive sur $]a, b[$, alors l'intégrale $\int_a^b f(t) dt$ est absolument convergente si et seulement si elle est convergente.

Proposition 4.23 : Soit $f :]a, b[\rightarrow \mathbf{R}$ continue.

Si $\int_a^b f(t) dt$ est absolument convergente, alors elle est convergente et on a alors

$$\left| \int_a^b f(t) dt \right| \leq \int_a^b |f(t)| dt.$$

Terminologie

Cette inégalité est appelée l'inégalité triangulaire pour les intégrales. Si l'on voit l'intégrale comme un analogue continu de la somme, alors c'est bien similaire à l'inégalité triangulaire «classique» :

$$\left| \sum a_i \right| \leq \sum |a_i|.$$

Exemple 4.24

Soit $f : [1; +\infty[$ définie par $f(t) = \frac{\sin t}{t\sqrt{t}}$.

Alors on a $\left| \frac{\sin t}{t\sqrt{t}} \right| \leq \frac{1}{t\sqrt{t}} = \frac{1}{t^{3/2}}$.

Or, l'intégrale $\int_1^{+\infty} \frac{1}{t^{3/2}} dt$ est une intégrale de Riemann convergente, donc par comparaison de fonctions positives, $\int_1^{+\infty} \left| \frac{\sin t}{t\sqrt{t}} \right| dt$ converge, et donc $\int_1^{+\infty} \frac{\sin t}{t\sqrt{t}} dt$ est absolument convergente, donc convergente.

Proposition 4.25 : Soient f, g deux fonctions continues sur $[a, b[$, avec g positive sur $[a, b[$ et $f(t) = o_{t \rightarrow b}(g(t))$.

Si $\int_a^b g(t) dt$ est convergente, alors $\int_a^b f(t) dt$ est absolument convergente (et donc convergente).

Démonstration. Si $f(t) = o_{t \rightarrow b}(g(t))$, alors il existe $A > 0$ tel que pour tout $t \in [A, b[$, $0 \leq |f(t)| \leq g(t)$.

Et puisque $\int_A^b g(t) dt$ converge, il en est de même de $\int_A^b |f(t)| dt$.

Et donc, f étant continue sur $[a, A]$, $\int_a^A |f(t)| dt$ converge⁵ et ainsi $\int_a^b |f(t)| dt$ converge. □

⁵ C'est l'intégrale de la fonction continue $|f|$ sur un segment, donc il n'y a pas de problème de convergence.

Exemple 4.26 Intégrale de Gauss

Soit $f(t) = e^{-t^2}$. La fonction f est continue sur \mathbf{R}_+ , et on a $e^{-t^2} = o_{+\infty}(e^{-t})$.

En effet, $\lim_{t \rightarrow +\infty} \frac{e^{-t^2}}{e^{-t}} = \lim_{t \rightarrow +\infty} e^{-t^2+t} = 0$.

Puisque $\int_0^{+\infty} e^{-t} dt$ converge, on en déduit que $\int_0^{+\infty} e^{-t^2} dt$ converge.

De plus, on admettra que $\int_0^{+\infty} e^{-t^2} dt = \frac{\sqrt{\pi}}{2}$.

Astuce

Ce résultat est admis, et n'a pas été prouvé non plus en première année. Il n'est en général par utile de connaître par cœur la valeur de cette intégrale. Si toutefois on en a besoin, on se rappellera que l'intégrale de la densité d'une loi $\mathcal{N}(0, 1)$ sur \mathbf{R} est égale à 1. Mais cela nécessite de connaître l'expression de cette densité !

Exemple 4.27

Soit $f(t) = \frac{\ln t}{t^2}$. La fonction f est continue et positive sur $[1; +\infty[$. De plus, on a

$$\lim_{t \rightarrow +\infty} \frac{t^{3/2} \ln t}{t^2} = \lim_{t \rightarrow +\infty} \frac{\ln t}{\sqrt{t}} = 0$$

donc $f(t) = o_{t \rightarrow +\infty}\left(\frac{1}{t^{3/2}}\right)$. Mais l'intégrale $\int_1^{+\infty} \frac{dt}{t^{3/2}}$ est convergente (intégrale de Riemann), de sorte que $\int_1^{+\infty} \frac{\ln t}{t^2} dt$ converge.

4.2.5 Reste d'une intégrale convergente

Définition 4.28 – Soit $f : [a, b[\rightarrow \mathbf{R}$ une fonction continue telle que $\int_a^b f(t) dt$ soit convergente. On appelle reste de l'intégrale la fonction $R : [a, b[\rightarrow \mathbf{R}$ définie pour tout $x \in [a, b[$ par

$$R(x) = \int_x^b f(t) dt.$$

Proposition 4.29 : Si $\int_a^b f(t) dt$ est convergente, alors

$$\lim_{x \rightarrow b^-} R(x) = \lim_{x \rightarrow b^-} \int_x^b f(t) dt = 0.$$

Ce résultat est similaire au résultat sur le reste d'une série convergente.

Démonstration. Soit $x \in [a, b]$. Alors

$$R(x) = \int_x^b f(t) dt = \int_a^b f(t) dt - \int_a^x f(t) dt \xrightarrow{x \rightarrow b^-} \int_a^b f(t) dt - \int_a^b f(t) dt = 0.$$

□

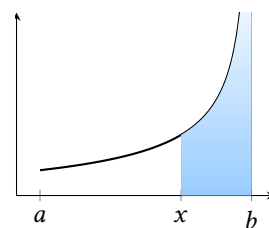


FIGURE 4.4– Graphiquement : le reste est l'aire sous la courbe entre x et b : il tend vers 0 lorsque $x \rightarrow b$.

Exemple 4.30

On a prouvé que $\int_1^{+\infty} \frac{\sin t}{t\sqrt{t}} dt$ est convergente. Donc $\lim_{x \rightarrow +\infty} \int_x^{+\infty} \frac{\sin t}{t\sqrt{t}} dt = 0$.

4.3 TECHNIQUES DE CALCUL D'INTÉGRALES

4.3.1 Intégration par parties

Proposition 4.31 : Soient f, g deux fonctions de classe \mathcal{C}^1 sur $[a, b]$. Alors

$$\int_a^b f'(t)g(t) dt = [f(t)g(t)]_a^b - \int_a^b f(t)g'(t) dt.$$

Dans le cas d'une intégrale impropre sur $[a, b]$, lorsqu'on souhaite réaliser une intégration par parties, on commencera par effectuer cette intégration par parties sur un segment $[a, x]$, avec $x \in [a, b]$, puis on fera tendre x vers b par valeurs inférieures. Dans ce cas, on veillera bien à justifier que toutes les limites considérées existent.

Ce théorème n'est pas nouveau, on parle bien ici d'intégration sur un **segment**, c'est le résultat de première année.

Exemple 4.32

Calculons $\int_0^{+\infty} e^{-t} \sin(t) dt$.

La fonction $t \mapsto e^{-t} \sin(t)$ est continue sur $[0, +\infty[$.

Pour tout $t \geq 0$, on a $0 \leq |\sin t| \leq 1$ et donc $0 \leq |e^{-t} \sin t| \leq e^{-t}$.

Puisque $\int_0^{+\infty} e^{-t} dt$ converge, $\int_0^{+\infty} e^{-t} \sin(t) dt$ converge absolument.

Soit donc $A > 0$. Alors les fonctions $u(t) = e^{-t}$ et $v(t) = -\cos(t)$ sont \mathcal{C}^1 sur $[0, A]$, avec $u'(t) = -e^{-t}$ et $v'(t) = \sin t$. Une intégration par parties donne alors

$$\int_0^A e^{-t} \sin(t) dt = [-e^{-t} \cos(t)]_0^A - \int_0^A e^{-t} \cos(t) dt = 1 - e^{-A} \cos(A).$$

Mais une seconde intégration par parties nous donne alors

$$\int_0^A e^{-t} \cos(t) dt = [e^{-t} \sin(t)]_0^A + \int_0^A e^{-t} \sin(t) dt.$$

Ainsi,

$$\int_0^A e^{-t} \sin(t) dt = 1 - e^{-A} \cos(A) - e^{-A} \sin(A) + \int_0^A e^{-t} \sin(t) dt.$$

Soit encore

$$\int_0^A e^{-t} \sin(t) dt = \frac{1}{2} - e^{-A} \frac{\cos(A) + \sin(A)}{2}.$$

Mais pour tout $A > 0$, $|\cos(A) + \sin(A)| \leq |\cos(A)| + |\sin(A)| \leq 2$, et donc la fonction $A \mapsto \cos(A) + \sin(A)$ est bornée sur \mathbf{R}_+ .

Par conséquent⁶,

$$\int_0^{+\infty} e^{-t} \sin(t) dt = \lim_{A \rightarrow +\infty} \int_0^A e^{-t} \sin(t) dt = \frac{1}{2}.$$

⁶ Le produit d'une fonction bornée par une fonction de limite nulle tend vers 0

4.3.2 Changement de variable

Théorème 4.33 : Soit f une fonction continue sur $] \alpha, \beta[$, $\alpha, \beta \in \overline{\mathbf{R}}$ et soit φ une fonction \mathcal{C}^1 sur $]a, b[\subset \mathbf{R}$, strictement monotone et telle que $\lim_{t \rightarrow a^+} \varphi(t) = \alpha$ et $\lim_{t \rightarrow b^-} \varphi(t) = \beta$. Alors les intégrales $\int_{\alpha}^{\beta} f(x) dx$ et $\int_a^b f(\varphi(t))\varphi'(t) dt$ sont de même nature, et en cas de convergence, sont égales.



Démonstration. Traitons d'abord le cas où φ est croissante.

Commençons par noter que d'après le théorème de la bijection, φ réalise une bijection de $]a, b[$ sur $] \alpha, \beta[$.

Notons F une primitive de f sur $] \alpha, \beta[$, et coupons en deux notre intégrale de départ, en prenant $c \in]a, b[$.

Alors, pour $y \in]c, b[$, on a

$$\int_c^y f(\varphi(t))\varphi'(t) dt = [F(\varphi(t))]_c^y = F(\varphi(y)) - F(\varphi(c)).$$

D'autre part, pour $u \in]\varphi(c), \beta[$, on a $\int_{\varphi(c)}^u f(x) dx = F(u) - F(\varphi(c))$.

Mais alors

$$\begin{aligned} \int_c^b f(\varphi(t))\varphi'(t) dt \text{ converge} &\iff \lim_{y \rightarrow b} \int_c^y f(\varphi(t))\varphi'(t) dt \text{ existe} \\ &\iff \lim_{y \rightarrow b} F(\varphi(y)) \text{ existe} \\ &\iff \lim_{u \rightarrow \beta} F(u) \text{ existe} \\ &\iff \lim_{u \rightarrow \beta} \int_{\varphi(c)}^u f(x) dx \text{ existe} \\ &\iff \int_{\varphi(c)}^{\beta} f(x) dx \text{ converge.} \end{aligned}$$

De plus, dans le cas où ces deux intégrales convergent, on a alors $\lim_{y \rightarrow b} F(\varphi(y)) = \lim_{u \rightarrow \beta} F(u)$, et donc les deux intégrales sont égales.

De même, on prouve que $\int_a^c f(\varphi(t))\varphi'(t) dt$ et $\int_{\alpha}^{\varphi(c)} f(x) dx$ sont de même nature, et en cas de convergence sont égales.

Or, par définition, $\int_a^b f(\varphi(t))\varphi'(t) dt$ converge si et seulement si les deux intégrales $\int_a^c f(\varphi(t))\varphi'(t) dt$ et $\int_c^b f(\varphi(t))\varphi'(t) dt$ convergent.

Donc si et seulement si les deux intégrales $\int_{\alpha}^{\varphi(c)} f(x) dx$ et $\int_{\varphi(c)}^{\beta} f(x) dx$ convergent.

Soit si et seulement si $\int_{\alpha}^{\beta} f(x) dx$ converge.

Et en cas de convergence, on a alors

$$\begin{aligned} \int_a^b f(\varphi(t))\varphi'(t) dt &= \int_a^c f(\varphi(t))\varphi'(t) dt + \int_c^b f(\varphi(t))\varphi'(t) dt \\ &= \int_{\alpha}^{\varphi(c)} f(x) dx + \int_{\varphi(c)}^{\beta} f(x) dx \\ &= \int_{\alpha}^{\beta} f(x) dx. \end{aligned}$$

Hypothèses

Les hypothèses sont un peu plus contraignantes qu'en première année : il faut non seulement vérifier que le changement de variable est \mathcal{C}^1 , mais en plus s'assurer de sa stricte monotonie.

Équivalence

Ici, on a une équivalence car φ est une bijection. Dans le sens \Leftarrow , il n'est pas nécessaire de supposer ceci, c'est juste une composition de limite. Mais pour obtenir l'implication réciproque, il faut composer des limites avec φ^{-1} , qui existe si φ est bijective.

L'idée de la preuve est la même lorsque φ est décroissante, sauf que cette fois, φ réalise une bijection de $]a, b[$ sur $] \beta, \alpha[$.

On a alors $\int_c^b f(\varphi(t))\varphi'(t) dt$ qui «correspond» à $\int_\alpha^{\varphi(c)} f(t) dt$. \square

Exemple 4.34 Calcul de $\int_0^{+\infty} \frac{1}{e^t + e^{-t}} dt$ via le changement de variable $x = e^t$

Notons que lorsqu'on donne un changement de variable de cette forme⁷, cela signifie que l'intégrale de départ est $\int_a^b f(\varphi(t))\varphi'(t) dt$, et donc il faut déjà trouver la fonction f pour laquelle ceci est vrai.

Notons donc $\varphi(t) = e^t$. Cette fonction est de classe \mathcal{C}^1 sur $[0, +\infty[$, strictement croissante, et on a $\lim_{t \rightarrow 0} \varphi(t) = 1$ et $\lim_{t \rightarrow +\infty} \varphi(t) = +\infty$.

De plus, $\varphi'(t) = e^t$ et on a alors

$$\frac{1}{e^t + e^{-t}} = \frac{1}{e^t + e^{-t}} \frac{e^t}{e^t} = \frac{1}{\underbrace{(e^t)^2 + 1}_{=f(\varphi(t))}} \overbrace{e^t}^{=\varphi'(t)}.$$

Et donc on doit avoir $f(\varphi(t)) = \frac{1}{(e^t)^2 + 1}$, et donc on peut prendre $f : x \mapsto \frac{1}{x^2 + 1}$.

Alors, par le théorème de changement de variable, les intégrales $\int_0^{+\infty} \frac{1}{e^t + e^{-t}} dt$


et $\int_1^{+\infty} \frac{1}{x^2 + 1} dx$ sont de même nature, et en cas de convergence sont égales.

Mais pour $A > 0$, on a

$$\int_1^A \frac{1}{x^2 + 1} dx = [\text{Arctan}(x)]_1^A = \text{Arctan}(A) - \text{Arctan}(1) \xrightarrow{A \rightarrow +\infty} \frac{\pi}{2} - \frac{\pi}{4} = \frac{\pi}{4}.$$

On en déduit que $\int_1^{+\infty} \frac{dx}{x^2 + 1}$ converge et vaut $\frac{\pi}{4}$.

Et donc $\int_0^{+\infty} \frac{1}{e^t + e^{-t}} dt$ converge également et vaut aussi $\frac{\pi}{4}$.

 **Une rédaction bien plus simple.**

En pratique, cette rédaction est un peu lourde. Voici une rédaction plus facile, et qui avec un peu d'entraînement sera bien plus simple à utiliser.

Soit $\varphi(t) = e^t$. Cette fonction est de classe \mathcal{C}^1 sur $[0, +\infty[$, strictement croissante, et on a $\lim_{t \rightarrow 0} \varphi(t) = 1$ et $\lim_{t \rightarrow +\infty} \varphi(t) = +\infty$.

On a alors $x = e^t \Leftrightarrow t = \ln(x)$, de sorte que $dt = \frac{dx}{x}$.

Et donc, par le théorème de changement de variable, les intégrales $\int_0^{+\infty} \frac{1}{e^t + e^{-t}} dt$

et $\int_1^{+\infty} \frac{1}{x + \frac{1}{x}} \left(\frac{dx}{x}\right) = \int_1^{+\infty} \frac{dx}{x^2 + 1}$ sont de même nature et en cas de convergence sont égales.

On conclut comme dans la rédaction précédente, en étudiant la nature de la seconde intégrale et en calculant sa valeur.

Pour une explication détaillée de la raison pour laquelle cette méthode correspond bien au théorème énoncé précédemment, on pourra consulter le complément 4.5.3

⁷ Où l'on donne la «nouvelle» variable en fonction de «l'ancienne».

Hypothèses

Cette étape est la même que dans la rédaction précédente et est inévitable : il s'agit de vérifier les hypothèses du théorème de changement de variable.

Détails

Il s'agit ici d'une notation : dx et dt n'ont pas de signification mathématique précise pour nous.

L'idée est de partir de l'égalité $t = \ln(x)$, et de la dériver des deux côtés.

À gauche, on dérive par rapport à t , et donc on note dt pour s'en souvenir, à droite on dérive par rapport à x et on note dx pour s'en souvenir.

Remarques. • Comme pour l'intégration par parties, on pourrait choisir de procéder à un changement de variable sur un segment (comme fait en première année), puis prendre des limites. L'inconvénient de cette méthode est qu'il faut alors exprimer la bijection réciproque de φ , afin d'être capable d'explicitier correctement les bornes. Ce théorème permet donc de gagner un peu de temps, et surtout d'alléger les notations.

• Le théorème affirme que les deux intégrales avant et après changement de variable sont

de même nature. Il suffit donc d'étudier la convergence d'une seule des deux intégrales.

• Le théorème s'applique notamment pour un changement de variable affine, c'est-à-dire de la forme $\varphi(t) = at + b$, avec $a \neq 0$.

Le programme officiel précise qu'un changement de variable qui ne serait pas affine doit être donné dans l'énoncé.

Méthode : Calculer une intégrale par changement de variable $x = \varphi(t)$

Le changement de variable doit être indiqué dans l'énoncé, sauf éventuellement s'il est affine.

On suppose ici que t est la variable de l'intégrale de départ, et qu'on nous donne x en fonction de t (par exemple : $x = e^t$ ou $x = t^2$).

On doit alors :

- Vérifier que φ est \mathcal{C}^1 et vérifier sa **stricte** monotonie (si φ est affine, ces conditions sont automatiquement vérifiées).
- Déterminer les limites de φ aux bornes de l'intervalle afin de déterminer les bornes de l'intégrale obtenue après changement de variable.
- Exprimer t en fonction de x , dériver la relation ainsi obtenue de manière à obtenir dt en fonction de dx .
- Effectuer le changement de variable, c'est-à-dire remplacer les t de l'intégrale de départ par leur expression en fonction de x , et remplacer dt par l'expression obtenue à l'étape précédente.
- Étudier la nature de **l'une ou l'autre** des intégrales avant changement de variable ou après.

Proposition 4.35 (Intégrale des fonctions paires et impaires) : Soit f une fonction paire (respectivement impaire) et continue sur $] -a, a[$, $a \in \mathbf{R} \cup \{+\infty\}$;

Alors $\int_{-a}^a f(t) dt$ converge si et seulement si $\int_0^a f(t) dt$ converge, et dans ce cas

$$\int_{-a}^a f(t) dt = 2 \int_0^a f(t) dt. \left(\text{resp. } \int_{-a}^a f(t) dt = 0 \right)$$

Démonstration. Dans le cas où f est paire, le changement de variable $u = -t$ nous indique que

$$\int_0^a f(t) dt \text{ et } - \int_0^{-a} f(-u) du = \int_{-a}^0 f(u) du$$

sont de même nature, et qu'en cas de convergence elles sont égales, d'où le résultat.

Dans le cas où f est impaire, on a alors

$$\int_0^a f(t) dt \text{ et } - \int_0^{-a} f(-u) du = - \int_{-a}^0 f(u) du$$

qui sont de même nature et égales en cas de convergence.

Mais dans ce dernier cas, on a alors

$$\int_{-a}^a f(t) dt = \int_{-a}^0 f(t) dt + \int_0^a f(t) dt = 0.$$

□

⚠ Cette proposition ne dispense pas totalement d'une étude de convergence : il est indispensable de déterminer la nature de $\int_0^a f(t) dt$.

En revanche, il n'y a alors qu'un seul problème éventuel de convergence à étudier (au voisinage de a), alors qu'a priori, pour déterminer la nature de $\int_{-a}^a f(t) dt$, il y a d'éventuels problèmes de convergence en a et en $-a$.

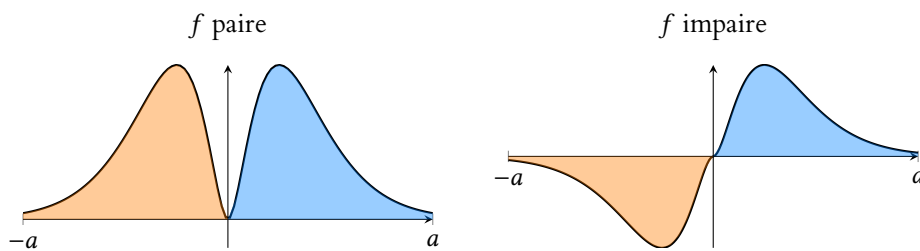


FIGURE 4.5 – Dans les deux cas, les deux aires sont égales. Dans le cas d'une fonction impaire, les deux intégrales correspondantes sont alors de signe opposé.

⚠ Même pour une fonction impaire, il faut étudier la convergence, par exemple $\int_{-\infty}^{+\infty} t^3 dt$ ne converge pas, car $\int_0^{+\infty} t^3 dt$ diverge.
 Il n'est donc pas question de dire que $\int_{-\infty}^{+\infty} t^3 dt = 0$.

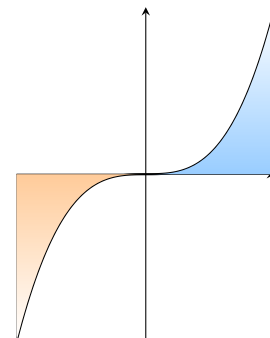


FIGURE 4.6– L'intégrale de t^3 entre $-\infty$ et $+\infty$ diverge : aucune des deux aires n'est finie.

4.4 UNE NOUVELLE FONCTION : LA FONCTION Γ

Proposition 4.36 : Soit $x \in \mathbf{R}_+^*$. Alors l'intégrale $\int_0^{+\infty} t^{x-1} e^{-t} dt$ est convergente.

Démonstration. La fonction $t \mapsto t^{x-1} e^{-t}$ est positive et continue sur $]0; +\infty[$.
 Au voisinage de 0, $e^{-t} \sim 1$, de sorte que $t^{x-1} e^{-t} \underset{t \rightarrow 0}{\sim} t^{x-1} = \frac{1}{t^{1-x}}$.
 Or $x > 0 \Leftrightarrow 1 - x < 1$, et donc $\int_0^1 \frac{1}{t^{1-x}} dt$ est une intégrale de Riemann convergente.
 Par comparaison des fonctions positives, on en déduit que $\int_0^1 t^{x-1} e^{-t} dt$ est convergente.
 Au voisinage de $+\infty$, on a $t^2 t^{x-1} e^{-t} = t^{x+1} e^{-t} \xrightarrow[t \rightarrow \infty]{} 0$ par croissance comparée.
 Donc $t^{x-1} e^{-t} = o\left(\frac{1}{t^2}\right)$. Mais, $\int_1^{+\infty} \frac{1}{t^2} dt$ est une intégrale de Riemann convergente, et donc par comparaison des fonctions positives, $\int_1^{+\infty} t^{x-1} e^{-t} dt$ est convergente.
 On en déduit donc que

$$\int_0^{+\infty} t^{x-1} e^{-t} dt \text{ est convergente.}$$

□

Définition 4.37 – Pour $x \in \mathbf{R}_+^*$, on pose

$$\Gamma(x) = \int_0^{+\infty} t^{x-1} e^{-t} dt.$$

Proposition 4.38 : $\forall x \in \mathbf{R}_+^*$, on a $\Gamma(x + 1) = x\Gamma(x)$.

Démonstration. Procédons à une intégration par parties : soit $[a, b]$ un segment de \mathbf{R}_+^* . Alors

$$\int_a^b t^x e^{-t} dt = [-t^x e^{-t}]_a^b + x \int_a^b t^{x-1} e^{-t} dt = a^x e^{-a} - b^x e^{-b} + x \int_a^b t^{x-1} e^{-t} dt.$$

En passant à la limite lorsque a tend vers 0, on a

$$\int_0^b t^x e^{-t} dt = -b^x e^{-b} + x \int_0^b t^{x-1} e^{-t} dt.$$

Puis, lorsque $b \rightarrow +\infty$, par croissances comparées, il vient

$$\Gamma(x + 1) = \int_0^{+\infty} t^x e^{-t} dt = x \int_0^{+\infty} t^{x-1} e^{-t} dt = x\Gamma(x).$$

□

Corollaire 4.39 – $\forall n \in \mathbf{N}, \Gamma(n + 1) = n!$

Démonstration. Prouvons le résultat par récurrence sur n . Pour $n = 0$, on a

$$\Gamma(1) = \int_0^{+\infty} e^{-t} dt = \lim_{a \rightarrow 0^+} \int_a^1 e^{-t} dt + \lim_{b \rightarrow \infty} \int_1^b e^{-t} dt = \lim_{a \rightarrow 0^+} (e^{-a} - 1) - \lim_{b \rightarrow \infty} (1 - e^{-b}) = 1 = 0!.$$

Supposons que $\Gamma(n) = (n - 1)!$. Alors $\Gamma(n + 1) = n\Gamma(n) = n(n - 1)! = n!$
Ainsi, par le principe de récurrence,

$$\forall n \in \mathbf{N}, \Gamma(n + 1) = n!$$

□

Exemple 4.40

On a par exemple $\Gamma(4) = 6$ et $\Gamma(6) = 120$.

Remarque. Ainsi, la fonction Γ est un prolongement de la factorielle (définie uniquement pour les entiers positifs) à l'ensemble des nombres réels.

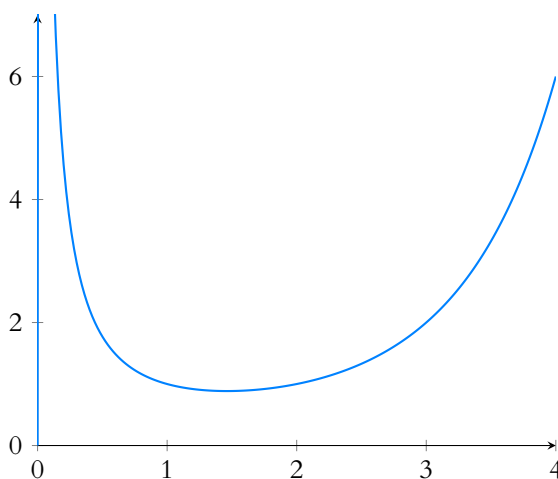


FIGURE 4.7 – La fonction Γ .

Un peu d'histoire

Historiquement, c'est cette formule qui a justifié qu'on s'intéresse à la fonction Γ : les mathématiciens du XVIII^e siècle cherchaient un moyen de prolonger la factorielle, a priori définie uniquement pour les nombres entiers à tous les réels positifs. On s'est aperçus plus tard que cette fonction était importante dans de nombreux domaines des mathématiques. En ce qui nous concerne, nous l'utiliserons surtout en probabilités, où elle nous permettra de définir une nouvelle densité, qui intervient pas exemple lorsqu'on s'intéresse à la somme de plusieurs variables aléatoires suivant une même loi exponentielle.

▲ Croissances

On sait que $e^n = o(n!)$. Une étude un peu plus poussée de la fonction Γ nous permettrait de prouver que

$$e^x = o_{x \rightarrow +\infty}(\Gamma(x)).$$

Autrement dit, la fonction Γ tend vers $+\infty$ «encore plus vite» que la fonction exponentielle.

4.5 COMPLÉMENTS

4.5.1 Sur l'histoire du théorème fondamental de l'analyse.

La preuve que nous avons donnée du théorème fondamental de l'analyse (4.4) repose sur un résultat admis en première année, qui est l'existence de primitives pour toute fonction

continue. Mais comment ce résultat se démontre-t-il ?

Historiquement, les choses ont été faites dans un autre ordre : l'intégrale a été définie avant de parler de primitive. Ceci peut paraître surprenant lorsqu'on regarde notre définition de l'intégrale : $\int_a^b f(t) dt = [F(t)]_a^b$, qui nécessite l'existence de primitives.

Mais on peut également définir l'intégrale en terme d'aire : pour une fonction positive, $\int_a^b f(t) dt$ est l'aire de la partie du plan délimitée par la courbe de f , l'axe des abscisses et les droites verticales d'équations $x = a$ et $x = b$.

Le problème est que bien qu'étant une notion intuitive, l'aire n'est pas bien définie en termes mathématiques.

Il est facile de définir l'aire d'un carré : un carré de côté 1cm possède une aire de 1cm^2 , et plus généralement, un carré de côté $a\text{cm}$ possède une aire de $a^2\text{cm}^2$.

À partir de là, il est facile de définir l'aire d'un rectangle de longueur a et de largeur b : il suffit de le découper en $a \times b$ carrés de côté 1, et donc son aire est $a \times b$.

Mais une fois définie l'aire des rectangles, comment définir l'aire d'autres parties du plan, par exemple l'aire d'un cercle ? Même en cherchant bien, on n'arrivera jamais à découper un cercle en petits carrés⁸.

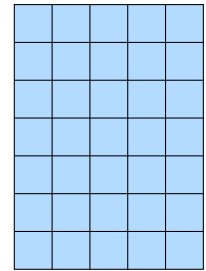


FIGURE 4.8– Découpage d'un rectangle de côtés a et b en $a \times b$ carrés de côté 1.

⁸ En tous cas pas en un nombre fini de carrés.

L'idée de Riemann pour définir l'aire sous la courbe d'une fonction a alors été celle qui se cache derrière les sommes de Riemann : approcher cette aire par celle de n rectangles, quitte à commettre une petite erreur, puis faire tendre le nombre de rectangles vers $+\infty$ afin de rendre l'erreur de plus en plus petite.

On définit alors $\int_a^b f(t) dt$ comme étant égale à $\lim_{n \rightarrow +\infty} \frac{b-a}{n} \sum_{k=0}^{n-1} f\left(a + k \frac{b-a}{n}\right)$.

Il y a alors un peu de travail pour montrer que cette limite existe bien, mais une fois ce travail fait, on peut démontrer sans grandes difficultés toutes les propriétés classiques de l'intégrale : la linéarité, la croissance, la relation de Chasles, etc.

Et une fois tout ce travail effectué, on peut passer à la démonstration du théorème fondamental de l'analyse, et prouver que $F : x \mapsto \int_a^x f(t) dt$ est dérivable⁹.

Pour cela, il faut revenir à la définition d'une dérivée et montrer que $\lim_{h \rightarrow 0} \frac{F(x+h) - F(x)}{h}$ existe et vaut $f(x)$.

Or on a

$$\frac{F(x+h) - F(x)}{h} = \frac{1}{h} \int_x^{x+h} f(t) dt.$$

Lorsque h est «petit», sur l'intervalle $[x, x+h]$, f est proche de $f(x)$ et donc $\int_x^{x+h} f(t) dt \simeq f(x)h$.

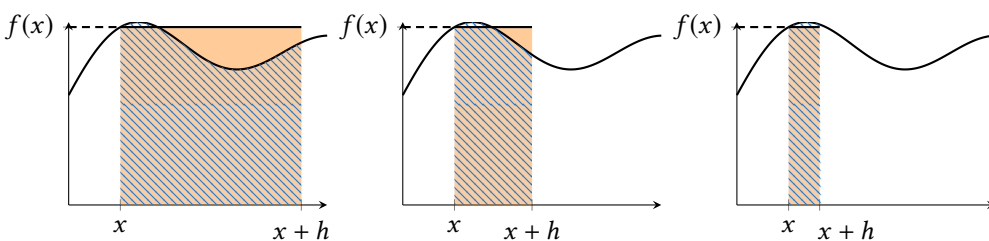


FIGURE 4.9 – Plus h est petit, plus l'aire $\int_x^{x+h} f(t) dt$ est proche de celle d'un rectangle de hauteur $f(x)$ et de largeur h .

Et donc $\frac{F(x+h) - F(x)}{h} \simeq f(x)$.

Bien entendu, ceci ne constitue pas une preuve, mais en étant un peu plus rigoureux, c'est ainsi qu'on prouve que $F' = f$ et donc que f admet au moins une primitive.

⁹ La dérivabilité de F est le seul point vraiment délicat dans le théorème.

4.5.2 De la nécessité de faire les intégrations par parties sur un segment

Nous avons insisté précédemment sur le fait qu'une intégration par parties dans une intégrale impropre doit nécessairement être réalisée sur un segment.

Si on ne le fait pas, on pourrait être tentés d'écrire des choses qui n'ont pas de sens, par exemple

$$\int_0^1 t \ln t \, dt = \left[\frac{t^2}{2} \ln(t) \right]_0^1 - \int_0^1 \frac{t}{2} \, dt.$$

En effet, le crochet de gauche n'est pas bien défini : la fonction $t \mapsto \frac{t^2}{2} \ln t$ n'est pas définie en 0, et ne peut donc pas être évaluée en 0.

Bien sûr, on peut remplacer l'évaluation en 0 par une limite et considérer que lorsque qu'on écrit le crochet ainsi, on a

$$\left[\frac{t^2}{2} \ln(t) \right]_0^1 = \frac{1^2}{2} \ln(1) - \lim_{t \rightarrow 0^+} \frac{t^2}{2} \ln(t),$$

mais cela suppose d'avoir préalablement étudié l'existence de cette limite.

Nous ne nous autorisons donc pas à écrire $[F(t)]_a^b$ si F n'est pas définie en a ou en b .

Des cas encore plus défavorables peuvent se produire. Considérons par exemple l'intégrale

$$\int_0^1 \frac{\sin t}{t} \, dt.$$

Puisque $\frac{\sin t}{t} \underset{t \rightarrow 0}{\sim} \frac{t}{t} = 1$, il s'agit d'une intégrale faussement impropre et donc convergente,

mais il ne s'agit pas pour autant d'une intégrale sur un segment, la fonction $t \mapsto \frac{\sin t}{t}$ n'étant pas continue en 0.

Une intégration par parties sans précautions nous conduirait à écrire

$$\int_0^1 \frac{\sin t}{t} \, dt = \left[-\frac{\cos t}{t} \right]_0^1 - \int_0^1 \frac{\cos t}{t^2} \, dt.$$

Cette fois, on ne sait pas trop quelle signification donner au crochet, puisque $\frac{\cos t}{t} \underset{t \rightarrow 0}{\sim} \frac{1}{t}$ qui n'a pas de limite.

Et pire encore, puisque $\frac{\cos t}{t^2} \underset{t \rightarrow 0}{\sim} \frac{1}{t^2}$, l'intégrale $\int_0^1 \frac{\cos t}{t^2} \, dt$ est une intégrale divergente, et cela n'a donc aucun sens de l'utiliser dans un calcul !

4.5.3 Sur la rédaction des changements de variable.

Nous avons expliqué à l'exemple 4.34 comment réaliser aisément un changement de variable, sans nécessairement connaître par cœur la formule de changement de variable, ni même sans se soucier de la fonction à laquelle nous cherchons à l'appliquer.

Voyons sur deux exemples comment procéder, et quel est le lien avec l'énoncé du théorème.

1^{er} cas : on donne la «nouvelle» variable en fonction de l'ancienne.

Calculons $\int_0^1 \frac{\ln t}{\sqrt{1-t}} \, dt$ à l'aide du changement de variable $x = \sqrt{1-t}$.

La fonction $t \mapsto \sqrt{1-t}$ est \mathcal{C}^1 et strictement croissante sur $]0, 1[$, donc l'emploi du théorème de changement de variable est légitime.

On a alors $\lim_{t \rightarrow 0^+} \sqrt{1-t} = 1$ et $\lim_{t \rightarrow 1^-} \sqrt{1-t} = 0$.

De plus, on a $x = \sqrt{1-t} \Leftrightarrow t = 1 - x^2$ et donc $dt = -2x \, dx$.

Donc, sous réserve de convergence de l'une de ces deux intégrales,

$$\int_0^1 \frac{\ln t}{\sqrt{1-t}} \, dt = \int_1^0 \frac{\ln(1-x^2)}{x} (-2x \, dx) = 2 \int_0^1 \ln(1-x^2) \, dx.$$

Or, $\ln(1-x^2) = \ln(1+x) + \ln(1-x)$, et donc pour $A \in [0, 1[$,

$$\int_0^A \ln(1-x^2) \, dx = \int_0^A \ln(1+x) \, dx + \int_0^A \ln(1-x) \, dx$$

Divergente

Rappelons qu'il n'y a rien d'intéressant à dire à propos d'une intégrale divergente, si ce n'est qu'elle diverge.

La quantité $\int_0^1 \frac{\cos t}{t^2} \, dt$ n'est pas définie, ce n'est pas un nombre réel, et donc on **ne peut pas** l'utiliser dans un calcul.

$$\begin{aligned}
 &= \int_1^{A+1} \ln(u) du + \int_{1-A}^1 \ln(v) dv = \int_{1-A}^{A+1} \ln(t) dt \\
 &= [t \ln t - t]_{1-A}^{A+1} = (1+A) \ln(1+A) - (1+A) - (1-A) \ln(1-A) + (1-A) \\
 &\xrightarrow{A \rightarrow 1^-} 2 \ln(2) - 2.
 \end{aligned}$$

Détail
 On a ici réalisé deux changements de variable affines sur des segments, qui se passent d'explications.

Ainsi, $\int_0^1 \ln(1-x^2) dx$ converge, et donc il en est de même de $\int_0^1 \frac{\ln t}{\sqrt{1-t}} dt$, et cette intégrale vaut $4 \ln(2) - 4$.

En rédigeant ainsi, il n'a pas été nécessaire de s'interroger sur ce qu'étaient vraiment les fonctions f et φ du théorème de changement de variable : $\int_a^b f(\varphi(t))\varphi'(t) dt = \int_\alpha^\beta f(x) dx$.

En réalité, la fonction f est la fonction qui figure dans la deuxième intégrale (donc ici $x \mapsto -2 \ln(1-x^2)$), et la fonction φ est celle qui donne x (la nouvelle variable) en fonction de t (l'ancienne variable). Donc ici, $\varphi : t \mapsto \sqrt{1-t}$.

Notons la présence du signe $-$, qui avait disparu puisque nous avons « remis les bornes dans le bon sens ».

On vérifie alors aisément que $f(\varphi(t))\varphi'(t) = \frac{\ln t}{\sqrt{1-t^2}}$.

Afin de bien comprendre le lien entre ce que nous avons fait et l'énoncé du théorème de changement de variable, notons $g : t \mapsto \frac{\ln(t)}{\sqrt{1-t}}$ la fonction que nous cherchions à intégrer au départ.

On veut donc avoir $g(t) = f(\varphi(t))\varphi'(t)$, soit $f \circ \varphi(t) = \frac{g(t)}{\varphi'(t)}$.

Et donc en composant par φ^{-1} pour isoler f , on a

$$f(t) = \frac{g(\varphi^{-1}(t))}{\varphi' \circ \varphi^{-1}(t)} = g(\varphi^{-1}(t)) \times (\varphi^{-1})'(t).$$

Rappel
 Il existe une formule pour la dérivée d'une bijection réciproque, c'est $(\varphi^{-1})'(t) = \frac{1}{\varphi' \circ \varphi^{-1}(t)}$.

Or, lorsque nous passons de $x = \sqrt{1-t}$ à $t = 1-x^2$, nous calculons la bijection réciproque de φ , qui est donc $\varphi^{-1} : t \mapsto 1-t^2$.

Et lorsqu'on dérive cette relation pour en déduire $dt = -2x dx$, on ne fait rien d'autre que dire que $(\varphi^{-1})'(x) = -2x$, soit encore $(\varphi^{-1})'(t) = -2t$.

Et donc

$$f(t) = \frac{g(\varphi^{-1}(t))}{(\varphi^{-1})'(t)} = \frac{\ln(1-t^2)}{\sqrt{1-\sqrt{1-t^2}}^2} (-2t) = -2 \ln(1-t^2).$$

Autrement dit, f est la fonction qu'on obtient lorsqu'on compose g par φ^{-1} (c'est ce que nous avons fait en remplaçant les t de l'intégrale de départ par des $1-x^2$), puis en multipliant par $(\varphi^{-1})'$ (et c'est ce que nous avons fait en remplaçant le dt par $-2x dx = (\varphi^{-1})'(x)$).

2^{ème} cas : on donne l'«ancienne» variable en fonction de la nouvelle.

Considérons l'intégrale $\int_0^1 \frac{1}{\sqrt{1-t^2}} dt$ et calculons la à l'aide du changement de variable $t = \sin(x)$.

Il y a là une petite subtilité pour calculer les bornes. On a $t = 0$ pour $x = 0$ et $t = 1$ pour $x = \frac{\pi}{2}$.

La fonction $x \mapsto \sin(x)$ est \mathcal{C}^1 et strictement croissante sur $]0, \frac{\pi}{2}[$ donc le changement de variable est légitime.

De plus, $t = \sin(x)$ implique que $dt = \cos(x) dx$.

Et donc, sous réserve de convergence,

$$\int_0^1 \frac{dt}{\sqrt{1-t^2}} = \int_0^{\pi/2} \frac{1}{\sqrt{1-(\sin x)^2}} \cos x dx = \int_0^{\pi/2} \frac{1}{\sqrt{\cos^2(x)}} \cos x dx = \int_0^{\pi/2} 1 dx = \frac{\pi}{2}.$$

Puisque la dernière intégrale converge, l'intégrale de départ est également convergente.

Bornes
 Il faut un peu d'intuition pour trouver ces bornes, vu qu'on n'a pas comme précédemment une expression de x en fonction de t .
 Notons qu'on aurait aussi pu choisir de prendre 2π et $2\pi + \frac{\pi}{2}$ comme nouvelles bornes.

Détails
 Sur $[0, \pi/2]$, le cosinus est positif, donc on a bien $\sqrt{\cos^2(x)} = |\cos(x)| = \cos(x)$.

Expliquons là aussi le lien avec le théorème de changement de variable. Puisqu'on nous donne l'expression de la nouvelle variable en fonction de l'ancienne, on a ici φ^{-1} et non φ ,

à savoir $\varphi^{-1}(t) = \sin t$.

Nous n'avons alors pas nécessairement besoin de déterminer φ pour aller plus loin.

En effet, nous n'avons jamais eu besoin d'exprimer x en fonction de t pour aller plus loin, et c'est tant mieux, car nous ne savons pas exprimer la bijection réciproque de la fonction sinus.

Comment a-t-on alors «trouvé» la fonction f , qui ici est la fonction constante $x \mapsto 1$.

Notons encore une fois g la fonction de départ : $g(t) = \frac{1}{\sqrt{1-t^2}}$.

Nous cherchons donc f telle que $g(t) = f(\varphi(t))\varphi'(t)$.

Soit encore $\frac{g(t)}{\varphi'(t)} = f(\varphi(t))$.

En composant à droite par φ^{-1} , on a donc

$$f(t) = \frac{g(\varphi^{-1}(t))}{\varphi' \circ \varphi^{-1}(t)} = g(\varphi^{-1}(t)) (\varphi^{-1})'(t).$$

Et donc autrement, pour trouver la fonction f de la seconde intégrale¹⁰, il faut composer la fonction de départ g par φ^{-1} (et c'est ce que l'on fait lorsqu'on remplace les t par des $\sin x$), puis multiplier par $(\varphi^{-1})'(t)$ (c'est ce qu'on fait en remplaçant dt par $\cos(x) dx = (\varphi^{-1})'(x) dx$).

Il reste alors un tout petit problème : nous avons mentionné la stricte monotonie et le caractère \mathcal{C}^1 de $x \mapsto \sin(x)$, c'est-à-dire de φ^{-1} et pas de φ ...

Ceci n'est pas important, puisque si une fonction est \mathcal{C}^1 et strictement monotone, sa bijection réciproque le sera aussi, avec le même sens de variation.

Donc dans les deux cas, notre méthode nous permet donc d'appliquer correctement le théorème de changement de variable, sans avoir besoin de savoir dès le départ à quelle fonction f on souhaite l'appliquer.

Cela étant dit, l'unique but de ce paragraphe était de vous convaincre du bien fondé de cette méthode, et nous n'aurons plus jamais besoin de nous poser autant de questions, il suffira d'appliquer correctement la méthode, sans chercher à expliciter¹¹ la fonction f à laquelle on applique le théorème de changement de variable.

Remarque

Cette fonction est bien connue et s'appelle Arcsin, mais elle ne figure pas au programme d'ECS.

¹⁰ Celle que l'on saura calculer.

¹¹ Nous venons de voir que dans les deux cas, la fonction f était celle qui figurait dans l'intégrale après changement de variable.

MÉTHODES DU CHAPITRE 4

✓ **MÉTHODE 4.1** : Étudier la convergence d'une intégrale $\int_a^b f(t) dt$

• Étudier précisément le domaine de continuité de f , repérer les problèmes éventuels (là où le domaine est ouvert), et à l'aide de la relation de Chasles, couper l'intégrale en morceaux ne contenant qu'un seul problème à la fois.

Si l'un de ces morceaux diverge, l'intégrale est divergente, sinon on doit prouver que **tous** les morceaux sont convergents. Ensuite, on peut :

- Utiliser une primitive de f si on en trouve une, calculer l'intégrale sur un segment, puis faire un passage à la limite.
- En une borne **finie**, étudier l'existence d'une éventuelle limite de f . Si on trouve une limite finie (= si f est prolongeable par continuité), l'intégrale est faussement impropre, donc converge.
- Pour une fonction **positive**, utiliser des majorations, des minoration, des équivalents pour se ramener à une intégrale qu'on sait étudier aisément (le plus souvent une intégrale de Riemann, éventuellement une intégrale étudiée plus tôt au cours de l'exercice).
- Montrer que t est négligeable devant une intégrale de Riemann convergente en montrant que $t^\alpha f(t) \rightarrow 0$ pour α bien choisi.

Exercices de TD : 4.1, 4.3, 4.4, 4.5, 4.6, 4.8, 4.9, 4.10, 4.11, 4.13, 4.14, 4.15, 4.16

✓ **MÉTHODE 4.2** : Calculer $\int_a^b f(t) dt$ lorsqu'on connaît une primitive de f .

On calcule alors l'intégrale sur un **segment** $[c, d]$, puis on fait tendre c vers a puis d vers b .

Notons que si le domaine de continuité de f est fermé d'un des deux côtés (i.e. du type $[a, b[$ ou $]a, b]$) il suffit de prendre une seule limite.

Exercices de TD : 4.1

✓ **MÉTHODE 4.3** : Calculer une intégrale $\int_a^b f(t) dt$ par intégration par parties.

On se place **toujours** sur un segment $[c, d] \subset]a, b[$, sur lequel on réalise une intégration par parties classique. Puisqu'il est plus difficile d'intégrer que de dériver, on commence en général par chercher comment écrire f comme un produit dont on sait intégrer un des deux termes.

Après avoir réalisé l'intégration par parties sur un segment, on fait tendre les bornes c et d vers a et b .

Exercices de TD : 4.11, 4.13, 4.14

✓ **MÉTHODE 4.4** : Calculer une intégrale par changement de variable $t = \varphi(x)$

Le changement de variable doit être indiqué dans l'énoncé, sauf éventuellement s'il est affine. On doit alors :

- Vérifier que φ est \mathcal{C}^1 et vérifier sa **stricte** monotonie (si φ est affine, ces conditions sont automatiquement vérifiées).
- Déterminer les limites de φ aux bornes de l'intervalle afin de déterminer les bornes de l'intégrale obtenue après changement de variable.
- Exprimer x en fonction de t , dériver la relation ainsi obtenue de manière à obtenir dt en fonction de dx .
- Appliquer correctement la formule de changement de variable.
- Étudier la nature de **l'une ou l'autre** des intégrales avant changement de variable ou après.

Exercices de TD : 4.6, 4.8, 4.10, 4.12, 4.15, 4.16

✓ **MÉTHODE 4.5** : Reconnaître et utiliser la fonction Γ


On devra connaître

- la formule définissant Γ : $\forall x > 0, \Gamma(x) = \int_0^{+\infty} t^{x-1} e^{-t} dt$,

- la formule $\Gamma(n) = (n - 1)!$ $n \in \mathbf{N}^*$.

Exercices de TD : 4.7, 4.8, 4.12

Des sujets de révision intéressants :

- EDHEC 2015 exercice 1
- EDHEC 2003, exercice 2 (DS2 2016/2017) : un sujet mêlant intégrales et séries, thème assez classique.
- ECRICOME 2009, exercice 2 : technique, mais si vous savez tout faire dans cet exercice, vous êtes au point sur les intégrales impropres !
- EM Lyon 2013, problème 1.
-  ESSEC 2013 : très calculatoire, mais plutôt abordable pour un sujet de parisienne.

EXERCICES DU CHAPITRE 4

EXERCICE 4.1 Prouver la convergence des intégrales suivantes et calculer leur valeur :

$$\int_0^{+\infty} t e^{-\frac{t^2}{\sigma^2}} dt \ (\sigma \neq 0), \int_0^1 \ln t dt, \int_{-\infty}^{+\infty} e^{-x} e^{-e^{-x}} dx, \int_0^{+\infty} e^{-2t} \cos t dt, \int_1^{+\infty} \frac{dt}{t(t+1)}.$$

EXERCICE 4.2 Des études de convergence

Déterminer la nature des intégrales suivantes.

$$\int_0^1 \frac{1-x}{\ln x} dx, \int_0^{+\infty} \frac{dt}{\sqrt{1+e^t}}, \int_0^{+\infty} \frac{u \ln u}{(1+u^2)^{3/2}} du, \int_0^{+\infty} \frac{dt}{(1+2t)\sqrt{t}}.$$

EXERCICE 4.3 Un classique des problèmes d'algèbre bilinéaire

- 1) Montrer que pour tout polynôme $P \in \mathbf{R}[X]$, $\int_0^{+\infty} P(t)^2 e^{-t^2} dt$ est convergente. Montrer qu'alors elle est positive.
- 2) À quelle condition sur P a-t-on $\int_0^{+\infty} P(t)^2 e^{-t^2} dt = 0$?

EXERCICE 4.4 Des sommes de Riemann

Déterminer les limites suivantes :

$$\lim_{n \rightarrow +\infty} \frac{1}{n} \sum_{k=0}^{n-1} \frac{1}{1 + \frac{k}{n}}, \lim_{n \rightarrow +\infty} \frac{1}{n} \sum_{k=1}^n \cos^2\left(\frac{k\pi}{n}\right), \lim_{n \rightarrow +\infty} \sum_{k=0}^n \frac{1}{\sqrt{n^2 + 2kn}}.$$

EXERCICE 4.5 Intégrale de Fresnel

À l'aide du changement de variable $u = t^2$, puis d'une intégration par parties, montrer que $\int_0^{+\infty} \sin(t^2) dt$ converge.

EXERCICE 4.6 Déterminer la nature, et le cas échéant, calculer la valeur des intégrales suivantes :

$$\int_{-\infty}^{+\infty} \frac{t \ln(t^2)}{1+t^4} dt, \int_{-\infty}^{+\infty} \ln\left(\frac{1}{t^2}\right) \frac{1}{t} dt, \int_{-\infty}^{+\infty} t e^{-t^2} dt.$$

EXERCICE 4.7 Soit α un réel strictement positif et $n \in \mathbf{N}$. Montrer la convergence et calculer la valeur de $\int_0^{+\infty} t^n e^{-\alpha t} dt$.

EXERCICE 4.8 (Extrait d'Éricome 93)

Justifier l'existence et calculer la valeur de $\int_1^{+\infty} \frac{\ln(t)}{t^{3/2}} dt$.

EXERCICE 4.9 Soit a un réel positif.

- 1) Déterminer deux constantes λ et μ telles que $\forall t > 0, \frac{1}{t(t+a)} = \frac{\lambda}{t} + \frac{\mu}{t+a}$.
- 2) À l'aide du changement de variable $x = e^t$, prouver la convergence de $\int_0^{+\infty} \frac{dt}{e^t + a}$ et calculer sa valeur.

EXERCICE 4.10 Changements de variable

Déterminer la nature, et le cas échéant, calculer la valeur des intégrales suivantes grâce au changement de variable indiqué.

$$\int_1^{+\infty} \frac{\text{Arctan}\left(\frac{1}{t}\right)}{t^2} dt \left(u = \frac{1}{t}\right), \int_0^{+\infty} e^{-\sqrt{t}} dt \left(u = \sqrt{t}\right), \int_0^{+\infty} \frac{\ln t}{1+t^2} dt, \left(x = \frac{1}{t}\right), \int_0^{+\infty} \frac{dx}{e^x + e^{-x}}, t = e^x.$$

EXERCICE 4.11 Justifier l'existence de $\int_0^{+\infty} \ln\left(1 + \frac{1}{x^2}\right) dx$ et calculer sa valeur par une intégration par parties.

EXERCICE 4.12 (Extrait ECRICOME 94)

Pour $k \in \mathbf{N}$, à l'aide du changement de variable $x = -(k+1) \ln t$, montrer que $\int_0^1 t^k (\ln t)^k dt$ converge et donner sa valeur.

EXERCICE 4.13 Valeur de Γ aux demi-entiers.

AD

On rappelle que l'intégrale de Gauss est $\int_0^{+\infty} e^{-t^2/2} dt = \sqrt{\frac{\pi}{2}}$.

- 1) En réalisant le changement de variable $u = \frac{t^2}{2}$ dans l'intégrale ci-dessus, déterminer la valeur de $\Gamma(\frac{1}{2})$.
- 2) En déduire que pour tout entier naturel n , $\Gamma\left(n + \frac{1}{2}\right) = \frac{(2n)!\sqrt{\pi}}{4^n n!}$.

EXERCICE 4.14 Fonction définie par une intégrale

AD

Pour $x \in \mathbf{R}$, on note, sous réserve de convergence, $F(x) = \int_1^{+\infty} \frac{e^{-tx^2}}{1+t^3} dt$.

- 1) Montrer que, pour tout $x \in \mathbf{R}$, l'intégrale définissant $F(x)$ converge. Ainsi, F est une fonction définie sur \mathbf{R} .
- 2) Montrer que F est paire.
- 3) Étudier les variations de F sur \mathbf{R}_+ , et calculer sa limite en $+\infty$. En déduire le tableau de variation de F (on ne demande pas la valeur de $F(0)$).
- 4) a) Montrer que pour tout $(a, b) \in (\mathbf{R}_+)^2$, on a $|e^{-a} - e^{-b}| \leq |a - b|$.
b) En déduire que pour tous $(x, y) \in \mathbf{R}^2$, on a $|F(x) - F(y)| \leq |x^2 - y^2|$.
c) Montrer que F est continue sur \mathbf{R} .

EXERCICE 4.15

D

- 1) Soit $\alpha > 0$. Montrer que $I = \int_1^{+\infty} \frac{\cos(t)}{t} dt$ et $J = \int_0^{+\infty} \frac{\cos(t) - \cos(\alpha t)}{t} dt$ convergent.
- 2) Montrer que pour tout $\varepsilon > 0$, $\int_\varepsilon^{+\infty} \frac{\cos(t) - \cos(\alpha t)}{t} dt = \int_\varepsilon^{\alpha\varepsilon} \frac{\cos t}{t} dt$.
- 3) En déduire que $J = \ln(\alpha)$.

EXERCICE 4.16 Deux expressions intégrales de $\zeta(2)$.

D

On admet dans cet exercice que $\sum_{k=1}^{+\infty} \frac{1}{k^2} = \frac{\pi^2}{6}$.

Soit f la fonction définie sur \mathbf{R}_+^* par $f(x) = \frac{x}{e^x - 1}$. Pour tout $n \in \mathbf{N}$, on pose $I_n = \int_0^{+\infty} f(x)e^{-nx} dx$.

- 1) Montrer que pour tout $n \geq 0$, I_n est convergente et que $0 \leq I_n \leq \frac{1}{n}$.
- 2) Montrer que pour $n \geq 1$, on a $f(x) = f(x)e^{-nx} + \sum_{k=1}^n xe^{-kx}$.
- 3) En déduire que $\int_0^{+\infty} f(x) dx = \frac{\pi^2}{6}$.
- 4) À l'aide du changement de variable $x = -\ln(t)$, prouver la convergence de $\int_0^1 \frac{\ln t}{1-t} dt$ et calculer sa valeur.

EXERCICE 4.17 Montrer que $\int_0^{+\infty} \ln(t)e^{-t} dt$ converge et est négative.

AD

EXERCICE 4.18 (QSP ESCP 2016)

TD

Soit f une fonction continue sur $[1, +\infty[$ telle que $\int_1^{+\infty} f(t) dt$ converge. Montrer que $\int_1^{+\infty} \frac{f(t)}{t} dt$ converge également.

EXERCICE 4.19 (Oral ESCP 2013)

TD

- 1) Montrer que pour tout $x > 0$, l'intégrale $\int_0^{+\infty} \frac{t}{1+xe^t} dt$ converge.
On note f l'application définie sur \mathbf{R}_+^* par : pour tout $x > 0$, $f(x) = \int_0^{+\infty} \frac{t}{1+xe^t} dt$.
- 2) Montrer que f est strictement décroissante sur \mathbf{R}_+^* .
- 3) a) Déterminer $\lim_{x \rightarrow +\infty} f(x)$, puis un équivalent de $f(x)$ pour x au voisinage de $+\infty$.
b) Déterminer $\lim_{x \rightarrow 0} f(x)$.

- 4) a) En utilisant le changement de variable $u = x \cdot e^t$ que l'on justifiera, montrer que :

$$f(x) = \ln(x)(\ln(x) - \ln(1+x)) + \int_x^{+\infty} \frac{\ln u}{u(1+u)} du.$$

- b) En déduire que f est dérivable sur \mathbf{R}_+^* et calculer $f'(x)$. Retrouver ainsi le sens de variation de f .

CORRECTION DES EXERCICES DU CHAPITRE 4

SOLUTION DE L'EXERCICE 4.1

1. La fonction $t \mapsto te^{-\frac{t^2}{\sigma^2}}$ est continue sur $[0, +\infty[$. De plus, on a, pour $A > 0$,

$$\int_0^A te^{-\frac{t^2}{\sigma^2}} = \left[\frac{-\sigma^2}{2} e^{-\frac{t^2}{\sigma^2}} \right]_0^A = \frac{\sigma^2}{2} \left(1 - e^{-\frac{A^2}{\sigma^2}} \right) \xrightarrow{A \rightarrow +\infty} \frac{\sigma^2}{2}.$$

Donc $\int_0^{+\infty} te^{-\frac{t^2}{\sigma^2}} dt$ converge et vaut $\frac{\sigma^2}{2}$.

2. La fonction $t \mapsto \ln t$ est continue sur $]0, 1]$, et pour $A \in]0, 1]$, on a

$$\int_A^1 \ln t dt = [t \ln t - t]_A^1 = -1 - A \ln(A) + A \xrightarrow{A \rightarrow 0^+} -1.$$

Donc $\int_0^1 \ln t dt$ converge et vaut -1 .

3. La fonction $t \mapsto e^{-x} e^{-e^{-x}}$ est continue sur \mathbf{R} .

Pour $A > 0$, on a

$$\int_0^A e^{-t} e^{-e^{-t}} dt = [e^{-e^{-t}}]_0^A = e^{-e^{-A}} - e^{-1} \xrightarrow{A \rightarrow +\infty} 1 - e^{-1}.$$

Donc $\int_0^{+\infty} e^{-t} e^{-e^{-t}} dt$ converge et vaut $1 - e^{-1}$.

De même, pour $B < 0$,

$$\int_B^0 e^{-t} e^{-e^{-t}} dt = [e^{-e^{-t}}]_B^0 = e^{-1} - e^{-e^{-B}} \xrightarrow{B \rightarrow -\infty} e^{-1}.$$

Et donc $\int_{-\infty}^0 e^{-t} e^{-e^{-t}} dt$ converge et vaut e^{-1} .

On en déduit donc que $\int_{-\infty}^{+\infty} e^{-t} e^{-e^{-t}} dt$ converge et vaut $1 - e^{-1} + e^{-1} = 1$.

4. La fonction $t \mapsto e^{-2t} \cos t$ est continue sur $[0, +\infty[$.

De plus, pour $t \geq 0$, on a $|e^{-2t} \cos t| \leq e^{-2t}$ et $\int_0^{+\infty} e^{-2t} dt$ converge, donc $\int_0^{+\infty} e^{-2t} \cos t dt$

converge absolument.

Pour $A > 0$, une double intégration par parties nous donne

$$\begin{aligned} \int_0^A e^{-2t} \cos t dt &= [e^{-2t} \sin t]_0^A + \int_0^A 2e^{-2t} \sin t dt \\ &= e^{-2A} \sin(A) + [-2e^{-2t} \cos t]_0^A - \int_0^A 4e^{-2t} \cos t dt. \end{aligned}$$

Et donc $\int_0^A e^{-2t} \cos t dt = \frac{1}{5} (2 - 2e^{-2A} \cos(A) + e^{-2A} \sin(A))$.

Mais les fonctions \cos et \sin étant bornées, on a

$$\lim_{A \rightarrow +\infty} e^{-2A} \cos(A) = \lim_{A \rightarrow +\infty} e^{-2A} \sin(A) = 0.$$

Donc $\int_0^{+\infty} e^{-2t} \cos(t) dt$ converge et vaut $\frac{2}{5}$.

5. La fonction $t \mapsto \frac{1}{t(t+1)}$ est continue sur $[1, +\infty[$, et pour tout $t \geq 1$, on a $\frac{1}{t(t+1)} = \frac{1}{t} - \frac{1}{t+1}$.

Donc pour $A > 1$,

$$\int_1^A \frac{dt}{t(t+1)} = \int_1^A \frac{dt}{t} - \int_1^A \frac{dt}{t+1} = [\ln t]_1^A - [\ln(1+t)]_1^A = \ln(2) - \ln\left(1 + \frac{1}{A}\right) \xrightarrow{A \rightarrow +\infty} \ln(2).$$

On en déduit que $\int_1^{+\infty} \frac{dt}{t(t+1)}$ converge et vaut $\ln(2)$.

Signe

Ne soyons pas surpris par le signe : sur $]0, 1]$, $\ln(t)$ est négatif !

Méthode

Si l'on doit procéder à une double intégration par parties, on prendra soin de dériver deux fois le même terme, et pas d'intégrer la deuxième fois le terme intégré la première fois, ce qui aurait pour effet de revenir à l'intégrale de départ. Ici, nous avons dérivé l'exponentielle la première fois, il faut donc dériver de nouveau l'exponentielle la seconde fois.

SOLUTION DE L'EXERCICE 4.2

1. La fonction $t \mapsto \frac{1-t}{\ln t}$ est continue sur $]0, 1[$.

Au voisinage de 0, on a $\frac{1-t}{\ln t} \xrightarrow{t \rightarrow 0} 0$.

Et donc $\int_0^{1/2} \frac{1-t}{\ln t} dt$ est faussement impropre, donc convergente.

Au voisinage de 1, on a $\ln t = \ln(1 + (t-1))$ avec $t-1 \rightarrow 0$, et donc $\ln(t) \underset{t \rightarrow 1}{\sim} t-1$.

Et donc $\frac{1-t}{\ln t} \underset{t \rightarrow 1^-}{\sim} \frac{1-t}{t-1} = -1$.

En particulier, la fonction $t \mapsto \frac{1-t}{\ln t}$ est prolongeable par continuité en 1, donc $\int_{1/2}^1 \frac{1-t}{\ln t} dt$ est faussement impropre, donc convergente.

Et donc $\int_0^1 \frac{1-t}{\ln t} dt$ converge.

2. La fonction $t \mapsto \frac{1}{\sqrt{1+e^t}}$ est continue et positive sur $[0; +\infty[$.

Au voisinage de $+\infty$, on a $e^t + 1 \underset{t \rightarrow +\infty}{\sim} e^t$ et donc $\frac{1}{\sqrt{e^t + 1}} \underset{t \rightarrow +\infty}{\sim} \frac{1}{\sqrt{e^t}} = e^{-1/2t}$.

Mais $\int_0^{+\infty} e^{-1/2t} dt$ converge, et ces fonctions sont positives, donc d'après le critère des équivalents, $\int_0^{+\infty} \frac{dt}{\sqrt{e^t + 1}}$ converge.

3. La fonction $u \mapsto \frac{u \ln u}{(1+u^2)^{3/2}}$ est continue sur $]0; +\infty[$.

Au voisinage de 0, on a $\frac{u \ln u}{(1+u^2)^{3/2}} \underset{u \rightarrow 0^+}{\sim} u \ln u \xrightarrow{u \rightarrow 0^+} 0$ par croissances comparées. Et donc l'intégrande est prolongeable par continuité en 0 : $\int_0^1 \frac{u \ln u}{(1+u^2)^{3/2}} du$ est faussement impropre.

Au voisinage de $+\infty$, on a $1+u^2 \underset{u \rightarrow +\infty}{\sim} u^2$ et donc $\frac{u \ln u}{(1+u^2)^{3/2}} \underset{u \rightarrow +\infty}{\sim} \frac{u \ln u}{(u^2)^{3/2}} = \frac{\ln u}{u^2}$.

Puisque les fonctions sont positives, d'après le critère des équivalents, $\int_1^{+\infty} \frac{u \ln u}{(1+u^2)^{3/2}} du$

et $\int_1^{+\infty} \frac{\ln u}{u^2} du$ sont de même nature.

Mais $u^{3/2} \frac{\ln u}{u^2} = \frac{\ln u}{\sqrt{u}} \xrightarrow{u \rightarrow +\infty} 0$, de sorte que $\frac{\ln u}{u^2} = \underset{u \rightarrow +\infty}{o} \left(\frac{1}{u\sqrt{u}} \right)$.

Puisque $\int_1^{+\infty} \frac{du}{u^{3/2}}$ converge, il en est de même de $\int_1^{+\infty} \frac{\ln u}{u^2} du$ et donc $\int_1^{+\infty} \frac{u \ln u}{(1+u^2)^{3/2}} du$.

Ainsi, $\int_0^{+\infty} \frac{u \ln u}{(1+u^2)^{3/2}} du$ converge.

4. La fonction $t \mapsto \frac{1}{(1+2t)\sqrt{t}}$ est continue sur $]0; +\infty[$, donc il y a d'éventuels problèmes de convergence en 0 et en $+\infty$.

Au voisinage de 0, $\frac{1}{(1+2t)\sqrt{t}} \underset{t \rightarrow 0}{\sim} \frac{1}{\sqrt{t}}$. Or, l'intégrale de Riemann $\int_0^1 \frac{dt}{\sqrt{t}}$ est convergente, donc par critère de comparaison pour les fonctions positives, il en est de même de $\int_0^1 \frac{dt}{(1+2t)\sqrt{t}}$.

Au voisinage de $+\infty$, on a $\frac{1}{(1+2t)\sqrt{t}} \underset{t \rightarrow +\infty}{\sim} \frac{1}{2t^{3/2}}$. Puisque l'intégrale $\int_1^{+\infty} \frac{dt}{t^{3/2}}$, il en est

de même de $\int_1^{+\infty} \frac{dt}{(1+2t)\sqrt{t}}$.

Ainsi, $\int_0^{+\infty} \frac{dt}{(1+2t)\sqrt{t}}$ converge.

Astuce

Si cet équivalent n'est pas à connaître par cœur, il est bon de savoir le retrouver ! Notons qu'il peut également s'obtenir à l'aide de la formule de Taylor-Young appliquée à la fonction \ln au voisinage de 1.

Rédaction

Commencer systématiquement par étudier le domaine de continuité de l'intégrande permet de repérer les problèmes qui vont nécessiter une étude. Ici, l'intervalle est ouvert en 0 et en $+\infty$: il va falloir étudier séparément ces deux problèmes. Et pour cela on coupe l'intervalle en deux.

SOLUTION DE L'EXERCICE 4.3

1. La fonction $t \mapsto P(t)^2 e^{-t^2}$ est continue sur $[0; +\infty[$, donc le seul éventuel problème de convergence est au voisinage de $+\infty$.
 Mais si on note $a_n t^n$ le terme de plus haut degré de $P(t)$, alors au voisinage de $+\infty$, $P(t)^2 e^{-t^2} \underset{t \rightarrow +\infty}{\sim} a_n^2 t^{2n} e^{-t^2}$. En particulier, $t^2 P(t)^2 e^{-t^2} \underset{t \rightarrow +\infty}{\sim} a_n^2 t^{2n+2} e^{-t^2}$.
 Posons alors $T = t^2$, de sorte que $T \xrightarrow[t \rightarrow +\infty]{} +\infty$, et alors par croissances comparées

$$t^2 a_n^2 t^{2n} e^{-t^2} = T a_n^2 T^n e^{-T} = a_n^2 T^{n+1} e^{-T} \xrightarrow[T \rightarrow +\infty]{} 0.$$

Donc $P(t)^2 e^{-t^2} = o_{+\infty} \left(\frac{1}{t^2} \right)$. Puisque l'intégrale $\int_1^{+\infty} \frac{dt}{t^2}$ converge, par critère de comparaison pour les fonctions positives, il en est de même de $\int_1^{+\infty} P(t)^2 e^{-t^2} dt$ et donc¹ de $\int_0^{+\infty} P(t)^2 e^{-t^2} dt$.

De plus, la fonction $t \mapsto P(t)^2 e^{-t^2}$ est continue sur $[0; +\infty[$, positive et donc

$$\int_0^{+\infty} P(t)^2 e^{-t^2} dt \geq 0.$$

2. La fonction $t \mapsto P(t)^2 e^{-t^2} dt$ est continue et positive sur $]0; +\infty[$. Donc

$$\int_0^{+\infty} P(t)^2 e^{-t^2} dt = 0 \Leftrightarrow \forall t \in [0; +\infty[, P(t)^2 e^{-t^2} = 0 \Leftrightarrow \forall t \in [0, +\infty[, P(t)^2 = 0.$$

Donc si l'intégrale est nulle, le polynôme P possède une infinité² de racines : tous les éléments de \mathbf{R}_+ . Mais le seul polynôme qui possède une infinité de racines est le polynôme nul, donc $P = 0$.

Inversement, il est évident que si P est le polynôme nul, alors $\int_0^{+\infty} P(t)^2 e^{-t^2} dt = 0$.

SOLUTION DE L'EXERCICE 4.4

Notons qu'ici l'énoncé indique qu'il s'agit de sommes de Riemann, mais le plus dur est d'y penser lorsqu'il s'agit de calculer de telles limites sans indications.

1. Pour la première, on reconnaît directement une somme de Riemann, pour la fonction $t \mapsto \frac{1}{1+x}$ sur le segment $[0, 1]$. Donc

$$\lim_{n \rightarrow +\infty} \frac{1}{n} \sum_{k=0}^{n-1} \frac{1}{1 + \frac{k}{n}} = \int_0^1 \frac{dt}{1+t} = [\ln(1+t)]_0^1 = \ln(2).$$

2. Posons $f(t) = \cos^2(t)$. Alors

$$\frac{1}{n} \sum_{k=1}^n \cos^2 \left(\frac{k\pi}{n} \right) = \frac{1}{\pi} \frac{\pi - 0}{n} \sum_{k=1}^n f \left(\frac{k(\pi - 0)}{n} \right) \xrightarrow[n \rightarrow +\infty]{} \frac{1}{\pi} \int_0^\pi f(t) dt.$$

Or, $\cos^2(t) = \frac{1 + \cos(2t)}{2}$ et donc

$$\int_0^\pi \cos^2(t) dt = \frac{1}{2} \left[t + \frac{\sin(2t)}{2} \right]_0^\pi = \frac{\pi}{2}.$$

On en déduit que

$$\lim_{n \rightarrow +\infty} \frac{1}{n} \sum_{k=1}^n \cos^2 \left(\frac{k\pi}{n} \right) = \frac{1}{\pi} \frac{\pi}{2} = \frac{1}{2}.$$

3. Faisons apparaître une somme de Riemann en factorisant le dénominateur par n :

$$\sum_{k=0}^n \frac{1}{\sqrt{n^2 + 2kn}} = \frac{1}{n} \sum_{k=0}^n \frac{1}{\sqrt{1 + 2\frac{k}{n}}}.$$

Rédaction

La tentation est grande de dire

$$t^2 t^{2n} e^{-t^2} \xrightarrow[t \rightarrow +\infty]{} 0$$

«par croissances comparées.» Pourtant, aucune croissance comparée du cours ne fait apparaître de e^{-t^2} . Si on veut se ramener à une croissance comparée usuelle, on ne peut faire l'économie du changement de variable $T = t^2$.

¹ La fonction intégrée est continue sur le segment $[0, 1]$ et donc $\int_0^1 P(t)^2 e^{-t^2} dt$ converge automatiquement.

Une exponentielle n'est jamais nulle !

² Tous les nombres de \mathbf{R}_+^* .

Racines

L'argument de racines est indispensable ici : le fait que l'intégrale soit nulle garantit que P est nul sur $[0; +\infty[$, mais ne suffit pas à assurer que $P = 0$. Le fait qu'un polynôme nul sur \mathbf{R}_+^* soit nécessairement le polynôme nul provient de l'argument de racines donné ici.

Autre option

Si on ne réalise pas qu'on peut faire apparaître artificiellement la constante π qui manque devant la somme, on peut aussi prendre $f(t) = \cos^2(\pi t)$ et l'intégrer entre 0 et 1.

Méthode

Pour intégrer $\cos^k x$ ou $\sin^k x$, on commence par linéariser, par exemple à l'aide des formules d'Euler.

On reconnaît alors presque une somme de Riemann, à l'exception du fait que la somme va de 0 à n . Or, le théorème du cours est valable pour des sommes allant de 0 à $n-1$ ou de 1 à n . Mais on a

$$\frac{1}{n} \sum_{k=0}^n \frac{1}{\sqrt{1+2\frac{k}{n}}} = \frac{1}{n} \sum_{k=0}^{n-1} \frac{1}{\sqrt{1+2\frac{k}{n}}} + \frac{1}{n\sqrt{3}}$$

avec $\lim_{n \rightarrow +\infty} \frac{1}{n\sqrt{3}} = 0$.

Posons $f(t) = \frac{1}{\sqrt{1+2t}}$, fonction qui est continue sur le segment $[0, 1]$. Alors

$$\frac{1}{n} \sum_{k=0}^{n-1} \frac{1}{\sqrt{1+2\frac{k}{n}}} = \frac{1}{n} \sum_{k=0}^{n-1} f\left(0 + (1-0)\frac{k}{n}\right) \xrightarrow{n \rightarrow +\infty} \int_0^1 f(t) dt = [\sqrt{1+2t}]_0^1 = \sqrt{3} - 1.$$

SOLUTION DE L'EXERCICE 4.5

Comme indiqué, procédons au changement de variable $u = t^2$. Ce changement de variable est bien légitime car \mathcal{C}^1 sur \mathbf{R}_+^* et strictement croissant.

Alors $du = 2t dt \Leftrightarrow dt = \frac{du}{2\sqrt{u}}$, et $\lim_{t \rightarrow 0} t^2 = 0$ et $\lim_{t \rightarrow +\infty} t^2 = +\infty$.

Par conséquent, les intégrales $\int_0^{+\infty} \sin(t^2) dt$ et $\int_0^{+\infty} \frac{\sin u}{2\sqrt{u}} du$ sont de même nature³.

La fonction $u \mapsto \frac{\sin u}{2\sqrt{u}}$ est continue sur $]0; +\infty[$, donc il y a d'éventuels problèmes de convergence en 0 et en $+\infty$.

Or, au voisinage de 0, on a $\frac{\sin u}{2\sqrt{u}} \underset{u \rightarrow 0^+}{\sim} \frac{u}{2\sqrt{u}} = 2\sqrt{u} \xrightarrow{u \rightarrow 0^+} 0$.

Donc $u \mapsto \frac{\sin u}{2\sqrt{u}}$ se prolonge par continuité en 0 et donc $\int_0^1 \frac{\sin u}{2\sqrt{u}} du$ est faussement impropre et donc convergente.

Pour $A > 0$, exprimons $\int_1^A \frac{\sin u}{2\sqrt{u}} du$ grâce à une intégration par parties : on a

$$\int_1^A \frac{\sin u}{2\sqrt{u}} du = \left[\frac{-\cos u}{2\sqrt{u}} \right]_1^A - \int_1^A \frac{\cos u}{4u\sqrt{u}} du = \frac{\cos 1}{2} - \frac{\cos A}{2\sqrt{A}} - \int_1^A \frac{\cos u}{4u\sqrt{u}} du.$$

Lorsque $A \rightarrow +\infty$, on a $\frac{\cos A}{\sqrt{A}} \rightarrow 0$ car $\cos(A)$ est borné et $\frac{1}{\sqrt{A}} \rightarrow 0$.

Et donc il reste à déterminer si $\lim_{A \rightarrow +\infty} \int_1^A \frac{\cos u}{4u\sqrt{u}} du$ existe, c'est-à-dire à étudier la nature de l'intégrale $\int_1^{+\infty} \frac{\cos u}{u\sqrt{u}} du$.

Or, pour tout $u \geq 1$, on a $\left| \frac{\cos u}{u\sqrt{u}} \right| \leq \frac{1}{u\sqrt{u}}$ et $\int_1^{+\infty} \frac{du}{u\sqrt{u}}$ est une intégrale de Riemann⁴ convergente.

Par domination, $\int_1^{+\infty} \left| \frac{\cos u}{u\sqrt{u}} \right| du$ converge et donc $\int_1^{+\infty} \frac{\cos u}{u\sqrt{u}} du$ converge absolument, donc converge.

Et donc,

$$\lim_{A \rightarrow +\infty} \int_1^A \frac{\sin u}{2\sqrt{u}} du = \frac{\cos 1}{2} - \int_1^{+\infty} \frac{\cos u}{4u\sqrt{u}} du.$$

Puisque cette limite existe, on en déduit que $\int_0^{+\infty} \frac{\sin u}{2\sqrt{u}} du$ converge, et donc il en est de même de $\int_0^{+\infty} \sin(t^2) dt$.

SOLUTION DE L'EXERCICE 4.6

Remarque

Ces sommes correspondent au fait que l'on coupe l'intervalle en n morceaux.

³ Et donc nous devons étudier la nature de la seconde intégrale.

Rédaction

En étudiant le comportement de la fonction au voisinage de 0, on s'intéresse au problème éventuel de convergence en 0, et pas en $+\infty$.

Donc on peut écrire que

$\int_0^1 \frac{\sin u}{2\sqrt{u}} du$ (ou plus généralement que pour tout $a > 0$, $\int_0^a \frac{\sin u}{2\sqrt{u}} du$) converge, mais pas $\int_0^{+\infty} \frac{\sin u}{2\sqrt{u}} du$.

⁴ Avec $\alpha = 3/2$.

1. La fonction $t \mapsto \frac{t \ln(t^2)}{1+t^4}$ est continue sur \mathbf{R}^* et impaire. Donc $\int_{-\infty}^{+\infty} \frac{t \ln(t^2)}{1+t^4} dt$ converge si et seulement si $\int_0^{+\infty} \frac{t \ln(t^2)}{1+t^4} dt$ converge, et si c'est le cas, $\int_{-\infty}^{+\infty} \frac{t \ln(t^2)}{1+t^4} dt = 0$.
Il y a des problèmes de convergence en 0 et en $+\infty$.
Au voisinage de 0, on a

$$\frac{t \ln(t^2)}{1+t^4} = \frac{2t \ln(t)}{1+t^4} \underset{t \rightarrow 0}{\sim} t \ln t \xrightarrow{t \rightarrow 0} 0.$$

Donc $\int_0^1 \frac{t \ln(t^2)}{1+t^4} dt$ est faussement impropre, donc converge.
Au voisinage de $+\infty$, on a

$$\frac{t \ln(t^2)}{1+t^4} = \frac{2t \ln t}{1+t^4} \underset{t \rightarrow +\infty}{\sim} \frac{2 \ln t}{t^3}.$$

On a alors $t^2 \frac{t \ln(t^2)}{1+t^4} \underset{t \rightarrow +\infty}{\sim} \frac{2 \ln t}{t} \xrightarrow{t \rightarrow +\infty} 0$.
Et donc $\frac{t \ln(t^2)}{1+t^4} = o\left(\frac{1}{t^2}\right)$, de sorte que $\int_1^{+\infty} \frac{t \ln(t^2)}{1+t^4} dt$ converge.
On en déduit donc que $\int_{-\infty}^{+\infty} \frac{t \ln(t^2)}{1+t^4} dt$ converge et vaut 0.

2. La fonction $t \mapsto \ln\left(\frac{1}{t^2}\right) \frac{1}{t}$ est continue sur \mathbf{R}^* , et elle y est impaire.
Donc $\int_{-\infty}^{+\infty} \frac{1}{t} \ln\left(\frac{1}{t^2}\right) dt$ converge si et seulement si $\int_{-\infty}^{+\infty} \frac{1}{t} \ln\left(\frac{1}{t^2}\right) dt$ converge.
Au voisinage de 0, on a $\frac{1}{t} \ln\left(\frac{1}{t^2}\right) = -\frac{2 \ln t}{t}$.
Mais pour $x \in]0, 1]$,

$$\int_x^1 \frac{\ln t}{t} dt = [\ln(t)^2]_A^1 = -(\ln A)^2 \xrightarrow{A \rightarrow 0^+} +\infty.$$

Et donc $\int_0^1 \frac{\ln t}{t} dt$ diverge. On en déduit donc que $\int_{-\infty}^{+\infty} \frac{1}{t} \ln\left(\frac{1}{t^2}\right) dt$ diverge.

3. La fonction $t \mapsto te^{-t^2}$ est continue sur \mathbf{R} , et elle y est impaire. Donc si $\int_0^{+\infty} te^{-t^2} dt$ converge, alors il en est de même de $\int_{-\infty}^{+\infty} te^{-t^2} dt$, qui est alors nulle.
Or, $t^3 te^{-t^2} = t^4 e^{-t^2}$. Si on pose $X = t^2$, on a alors $X^2 e^{-X} \xrightarrow{X \rightarrow +\infty} 0$ par croissance comparée.
Et donc $te^{-t^2} = o\left(\frac{1}{t^3}\right)$. Puisque $\int_1^{+\infty} \frac{dt}{t^3}$ converge, il en est de même de $\int_0^{+\infty} te^{-t^2} dt$.
Ainsi, $\int_{-\infty}^{+\infty} te^{-t^2} dt$ converge et est nulle.

SOLUTION DE L'EXERCICE 4.7

Procédons au changement de variable $x = at$, qui réalise une bijection \mathcal{C}^1 strictement croissante de \mathbf{R}_+^* sur \mathbf{R}_+^* .

On a alors $dx = a dt \Leftrightarrow dt = \frac{1}{a} dx$. Les intégrales $\int_0^{+\infty} t^n e^{-at} dt$ et $\frac{1}{a^{n+1}} \int_0^{+\infty} x^n e^{-x} dx$ sont de même nature, et en cas de convergence, sont égales.

Or, cette seconde intégrale est précisément $\Gamma(n+1)$, donc elle converge et vaut $n!$
On en déduit que l'intégrale de départ converge et

$$\int_0^{+\infty} t^n e^{-at} dt = \frac{n!}{a^{n+1}}.$$

SOLUTION DE L'EXERCICE 4.8

☠ Fonction impaire

L'imparité d'une fonction n'implique pas automatiquement que son intégrale est nulle : encore faut-il que cette intégrale converge. Et donc il faut d'abord étudier sa nature.

⚠ Attention !

On ne sort pas le carré du ln avant de s'être placé sur \mathbf{R}_+^* , ou alors avec une valeur absolue :

$$\ln(t^2) = 2 \ln(|t|)$$

sous peine d'écrire le logarithme d'un nombre négatif.

Intervalle

Attention : \mathbf{R}^* n'est pas un intervalle, mais l'union de deux intervalles : \mathbf{R}_+^* et \mathbf{R}_-^* . Et donc il y a aussi un éventuel problème de convergence au voisinage de 0.

Chgt de variable

Ce changement de variable est affine, l'énoncé n'est donc pas tenu de vous le préciser, vous devez être capable d'y penser seul.

La fonction $t \mapsto \frac{\ln(t)}{t^{3/2}}$ est continue sur $[1; +\infty[$, donc le seul éventuel problème de convergence est au voisinage de $+\infty$.

Or, on a

$$t^{5/4} \frac{\ln t}{t^{3/2}} = \frac{\ln t}{t^{1/4}} \xrightarrow{t \rightarrow +\infty} 0$$

de sorte que $\frac{\ln t}{t^{3/2}} = o_{t \rightarrow +\infty} \left(\frac{1}{t^{5/4}} \right)$. Puisque $\int_1^{+\infty} \frac{1}{t^{5/4}} dt$ est absolument convergente, il en est de même de $\int_1^{+\infty} \frac{\ln t}{t^{3/2}} dt$.

Pour calculer la valeur de cette intégrale, procédons à une intégration par parties sur un segment de la forme $[1, A]$, $A \geq 1$. Posons donc $u(t) = \ln t$ et $v' = \frac{1}{t\sqrt{t}}$, soit $u'(t) = \frac{1}{t}$ et $v(t) = -\frac{2}{\sqrt{t}}$. Alors u et v sont de classe \mathcal{C}^1 sur $[1, A]$ et

$$\begin{aligned} \int_1^A \frac{\ln t}{t\sqrt{t}} dt &= \left[-\frac{2\ln(t)}{\sqrt{t}} \right]_1^A + \int_1^A \frac{2}{t\sqrt{t}} dt \\ &= -\frac{2\ln(A)}{\sqrt{A}} + \left[-\frac{4}{\sqrt{t}} \right]_1^A \\ &= -\frac{2\ln(A)}{\sqrt{A}} - \frac{4}{\sqrt{A}} + 4 \\ &\xrightarrow{A \rightarrow +\infty} 4 \end{aligned}$$

On en déduit que $\int_1^{+\infty} \frac{\ln t}{t\sqrt{t}} dt = 4$.

SOLUTION DE L'EXERCICE 4.9

1. On a $\frac{\lambda}{t} + \frac{\mu}{t+a} = \frac{(\lambda+\mu)t + a\lambda}{t(t+a)}$.

Donc, on a $\frac{(\lambda+\mu)t + a\lambda}{t(t+a)} = \frac{1}{t(t+a)}$ pour tout $t > 0$ si et seulement si $(\lambda+\mu)t + a\lambda = 1$, pour tout $t > 0$.

Or deux polynômes coïncident en une infinité⁵ de valeurs si et seulement si ils ont les mêmes coefficients.

Et donc $\frac{(\lambda+\mu)t + a\lambda}{t(t+a)} = \frac{1}{t(t+a)}$ pour tout $t > 0$ si et seulement si

$$\begin{cases} \lambda = \frac{1}{a} \\ \lambda + \mu = 0 \end{cases} \Leftrightarrow \begin{cases} \lambda = \frac{1}{a} \\ \mu = -\frac{1}{a} \end{cases}$$

Et donc

$$\forall t > 0, \frac{1}{t(t+a)} = \frac{1}{a} \left(\frac{1}{t} - \frac{1}{t+a} \right).$$

2. La fonction $t \mapsto e^t$ est de classe \mathcal{C}^1 sur $[0; +\infty[$, y est strictement croissante, et $\lim_{t \rightarrow 0} e^t = 1$, $\lim_{t \rightarrow +\infty} e^t = +\infty$.

De plus, on a $x = e^t \Rightarrow dx = e^t dt = x dt \Leftrightarrow dt = \frac{dx}{x}$. Donc par changement de variable, on a, sous réserve de convergence de l'une des deux intégrales,

$$\int_0^{+\infty} \frac{dt}{e^t + a} = \int_1^{+\infty} \frac{dx}{x(x+a)}.$$

Or, pour $A > 1$, on a

$$\begin{aligned} \int_1^A \frac{dx}{x(x+a)} &= \frac{1}{a} \int_1^A \frac{dx}{x} - \frac{1}{a} \int_1^A \frac{dx}{x+a} \\ &= \frac{1}{a} [\ln(x)]_1^A - \frac{1}{a} [\ln(x+a)]_1^A \end{aligned}$$

Méthode

C'est la «règle $t^\alpha f(t)$ », analogue pour les intégrales de la règle $n^\alpha u_n$: on essaie de se ramener à une intégrale de Riemann, toute la difficulté est de trouver laquelle...

⁵ Il y a une infinité de réels strictement positifs.

Rédaction

Lors d'un changement de variable, on n'oubliera pas :

- de vérifier le caractère \mathcal{C}^1 du changement de variable
- de vérifier sa stricte monotonie
- de changer les bornes de l'intégrale en conséquence.

$$\begin{aligned}
&= \frac{1}{a} (\ln(A) - \ln(a + A) + \ln(1 + a)) \\
&= \frac{1}{a} \left(\ln\left(\frac{A}{a + A}\right) + \ln(1 + a) \right) \\
&\xrightarrow{A \rightarrow +\infty} \frac{\ln(1 + a)}{a}.
\end{aligned}$$

Ainsi, $\int_1^{+\infty} \frac{dx}{x(x+a)}$ converge et vaut $\frac{\ln(1+a)}{a}$, de sorte que $\int_0^{+\infty} \frac{dt}{e^t + a}$ converge également et vaut $\frac{\ln(1+a)}{a}$.

SOLUTION DE L'EXERCICE 4.10

1. La fonction $t \mapsto \frac{\text{Arctan}\left(\frac{1}{t}\right)}{t^2}$ est continue et positive sur $[1; +\infty[$.
Or, pour tout $t \geq 1$, on a

$$0 \leq \frac{\text{Arctan}\left(\frac{1}{t}\right)}{t^2} \leq \frac{\pi}{2} \frac{1}{t^2}.$$

Par comparaison à une intégrale de Riemann convergente, $\int_1^{+\infty} \frac{\text{Arctan}\left(\frac{1}{t}\right)}{t^2} dt$ converge.

Et alors, la fonction $t \mapsto \frac{1}{t}$ réalisant une bijection \mathcal{C}^1 strictement décroissante de $[1, +\infty[$ sur $]0, 1]$, avec $dt = -\frac{du}{u^2}$, il vient

$$\int_1^{+\infty} \frac{\text{Arctan}\left(\frac{1}{t}\right)}{t^2} dt = - \int_1^0 \frac{\text{Arctan}(u)}{\frac{1}{u^2}} \frac{du}{u^2} = \int_0^1 \text{Arctan}(u) du.$$

Afin de calculer cette dernière intégrale, procédons à une intégration par parties sur le segment $[0, 1]$:

$$\int_0^1 \text{Arctan}(u) du = [t \text{Arctan}(t)]_0^1 - \int_0^1 \frac{u}{1+u^2} du = \text{Arctan}(1) - \left[\frac{1}{2} \ln(1+u^2) \right]_0^1 = \frac{\pi}{4} - \frac{\ln(2)}{2}.$$

2. Le changement de variable $u = \sqrt{t}$ est de classe \mathcal{C}^1 sur $[0, +\infty[$, et strictement croissant. Donc le changement de variable est justifié.
On a alors $u^2 = t$ et donc $2u du = dt$, et, sous réserve de convergence de l'une des deux intégrales, on a

$$\int_0^{+\infty} e^{-\sqrt{t}} dt = 2 \int_0^{+\infty} e^{-u} u du.$$

Mais la seconde intégrale n'est autre que $\Gamma(1)$: elle converge et vaut 1, de sorte que $\int_0^{+\infty} e^{-\sqrt{t}} dt$ converge et vaut 2.

3. La fonction $t \mapsto \frac{1}{t}$ est \mathcal{C}^1 sur $]0, +\infty[$ et elle y est strictement décroissante. Le changement de variable est donc légitime.
De plus, si $x = \frac{1}{t}$, alors $t = \frac{1}{x}$ et donc $dt = -\frac{dx}{x^2}$.
Sous réserve de convergence, on a alors

$$\int_0^{+\infty} \frac{\ln(t)}{1+t^2} dt = \int_{+\infty}^0 \frac{\ln(1/x)}{1+\frac{1}{x^2}} \left(-\frac{1}{x^2}\right) dx = \int_0^{+\infty} \frac{-\ln(x)}{x^2+1} dx = - \int_0^{+\infty} \frac{\ln x}{1+x^2} dx.$$

Et donc, $\int_0^{+\infty} \frac{\ln x}{1+x^2} dx$, si elle existe, est égale à son propre opposé, et donc est nulle.

Il reste donc à étudier la convergence de cette intégrale.

La fonction $t \mapsto \frac{\ln t}{1+t^2}$ est continue sur $]0; +\infty[$.

Au voisinage de 0, on a $\frac{\ln t}{1+t^2} \underset{t \rightarrow 0^+}{\sim} \ln(t)$, et nous avons prouvé à l'exercice 1 que $\int_0^1 \ln t dt$

Méthode

On n'oublie pas de vérifier la **stricte monotonie** du changement de variable.

Arctan

Inutile d'apprendre par cœur une primitive de Arctan, mais il est bon de savoir qu'on peut en obtenir une par intégration par parties

Rappel

$\tan \frac{\pi}{4} = 1$ donc

$$\text{Arctan}(1) = \frac{\pi}{4}.$$

Convergence

D'après le théorème de changement de variable, il faut étudier au moins l'une des deux intégrales : celle avant changement de variable, ou celle obtenue après le changement de variable. Ici elles sont égales.

converge.

Par critère de comparaison pour les fonctions de signe constant⁶, on en déduit que

$$\int_0^1 \frac{\ln t}{1+t^2} dt \text{ converge.}$$

Au voisinage de $+\infty$, on a $\frac{\ln t}{1+t^2} \underset{t \rightarrow +\infty}{\sim} \frac{\ln t}{t^2}$.

Mais $t^{3/2} \frac{\ln t}{t^2} = \frac{\ln t}{\sqrt{t}}$ et donc $\frac{\ln t}{t^2} = \underset{t \rightarrow +\infty}{o} \left(\frac{1}{t\sqrt{t}} \right)$.

Puisque $\int_1^{+\infty} \frac{1}{t\sqrt{t}} dt$ converge⁷, il en est de même de $\int_1^{+\infty} \frac{\ln t}{t^2} dt$, et donc, par critère de

comparaison pour les fonctions positives⁸, il en est de même de $\int_1^{+\infty} \frac{\ln t}{1+t^2} dt$, et donc de

$$\int_0^{+\infty} \frac{\ln t}{1+t^2} dt.$$

On en déduit donc que $\int_0^{+\infty} \frac{\ln t}{1+t^2} dt = 0$.

SOLUTION DE L'EXERCICE 4.11

La fonction $x \mapsto \ln \left(1 + \frac{1}{x^2} \right)$ est continue sur $]0; +\infty[$.

Au voisinage de $+\infty$, on a $\ln \left(1 + \frac{1}{x^2} \right) \underset{x \rightarrow +\infty}{\sim} \frac{1}{x^2}$, donc $\int_1^{+\infty} \ln \left(1 + \frac{1}{x^2} \right) dx$ converge.

Au voisinage de 0, $\ln \left(1 + \frac{1}{x^2} \right) \underset{x \rightarrow 0}{\sim} \ln \left(\frac{1}{x^2} \right) \underset{x \rightarrow 0}{\sim} -2 \ln(x)$, donc $\int_0^1 \ln \left(1 + \frac{1}{x^2} \right) dx$ converge⁹.

Soit $[a; b]$ un segment inclus dans $]0; \infty[$.

A l'aide d'une intégration par parties, en posant $u(x) = x$ et $v(x) = \ln \left(1 + \frac{1}{x^2} \right)$, qui sont

deux fonctions de classe \mathcal{C}^1 sur $[a, b]$ avec $u'(x) = 1$ et $v'(x) = \frac{-2 \frac{1}{x^3}}{1 + \frac{1}{x^2}} = \frac{-2}{1 + x^2} \frac{1}{x}$.

$$\begin{aligned} \int_a^b \ln \left(1 + \frac{1}{x^2} \right) dx &= \left[x \ln \left(1 + \frac{1}{x^2} \right) \right]_a^b + \int_a^b \frac{2}{1+x^2} dx \\ &= b \ln \left(1 + \frac{1}{b^2} \right) - a \ln \left(1 + \frac{1}{a^2} \right) + 2 \operatorname{Arctan} b - 2 \operatorname{Arctan} a. \end{aligned}$$

Or, lorsque $a \rightarrow 0^+$, on a $a \ln \left(1 + \frac{1}{a^2} \right) \underset{a \rightarrow 0^+}{\sim} a \ln \frac{1}{a^2} \underset{a \rightarrow 0^+}{\sim} -2a \ln(a) \rightarrow 0$ et $-2 \operatorname{Arctan} a \rightarrow 0$.

Et lorsque $b \rightarrow +\infty$, $b \ln \left(1 + \frac{1}{b^2} \right) \underset{b \rightarrow +\infty}{\sim} \frac{b}{b^2} \rightarrow 0$ et $2 \operatorname{Arctan} b \rightarrow \pi$.

En conclusion, $\int_0^{+\infty} \ln \left(1 + \frac{1}{x^2} \right) dx = \pi$.

SOLUTION DE L'EXERCICE 4.12

Notons que pour $k = 0$, on a $t^k (\ln t)^k = 1$, et donc $\int_0^1 t^k (\ln t)^k dt = \int_0^1 dt$ converge et vaut 1.

Nous supposons donc $k > 0$.

La fonction $t \mapsto t^k (\ln t)^k$ est continue sur $]0, 1]$, et par croissances comparées,

$\lim_{t \rightarrow 0^+} t^k (\ln t)^k = 0$. Ainsi, $\int_0^1 t^k (\ln t)^k dt$ est faussement impropre, et donc convergente.

L'application $t \mapsto -(k+1) \ln(t)$ réalise une bijection \mathcal{C}^1 strictement décroissante de $]0, 1]$ sur $[0, +\infty[$. Alors, en posant $x = -(k+1) \ln t$, il vient $t = e^{-x/(k+1)}$, de sorte que

$dt = -\frac{1}{k+1} e^{-x/(k+1)} dx$. Et donc

$$\begin{aligned} \int_0^1 t^k (\ln t)^k dt &= - \int_{+\infty}^0 e^{-kx/(k+1)} \left(-\frac{x}{k+1} \right)^k \left(-\frac{1}{k+1} e^{-x/(k+1)} \right) dx \\ &= (-1)^k \frac{1}{(k+1)^{k+1}} \int_0^{+\infty} e^{-x} x^k dx. \end{aligned}$$

⁶ Ici les fonctions sont **néga-** tives sur $]0, 1]$.

⁷ C'est une intégrale de Riemann convergente : $\alpha = 3/2$.

⁸ $\frac{\ln t}{1+t^2}$ est positive sur $[1; +\infty[$.

⁹ Par exemple car on a prouvé la convergence de $\int_0^1 \ln x dx$ à l'exercice 1, ou parce que l'on remarque qu'il s'agit d'une fonction négligeable devant $\frac{1}{\sqrt{x}}$ au voisinage de 0.

On reconnaît alors l'intégrale $\Gamma(k + 1)$, et donc

$$\int_0^1 t^k (\ln t)^k dt = \frac{(-1)^k}{(k + 1)^{k+1}} k!$$

SOLUTION DE L'EXERCICE 4.13

1. La fonction $t \mapsto e^{-t^2/2}$ étant paire, on a

$$\int_0^{+\infty} e^{-t^2/2} dt = \frac{1}{2} \int_{-\infty}^{+\infty} e^{-t^2/2} dt = \sqrt{\frac{\pi}{2}}.$$

L'application $t \mapsto \frac{t^2}{2}$ est \mathcal{C}^1 sur \mathbf{R}_+^* , strictement croissante, et $\lim_{t \rightarrow 0} \frac{t^2}{2} = 0$ et $\lim_{t \rightarrow +\infty} \frac{t^2}{2} = +\infty$.

Donc par le théorème de changement de variable¹⁰ en posant $u = \frac{t^2}{2} \Leftrightarrow t = \sqrt{2u}$, on a $dt = \frac{1}{\sqrt{2u}} du$ et donc

$$\int_0^{+\infty} e^{-t^2/2} dt = \int_0^{+\infty} \frac{1}{\sqrt{2u}} e^{-u} du = \frac{1}{\sqrt{2}} \int_0^{+\infty} u^{1/2-1} e^{-u} du = \frac{1}{\sqrt{2}} \Gamma\left(\frac{1}{2}\right).$$

On en déduit que $\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}$.

2. Par récurrence sur n , prouvons que

$$\Gamma\left(n + \frac{1}{2}\right) = \frac{(2n)! \sqrt{\pi}}{4^n n!}.$$

Pour $n = 0$, c'est le résultat de la question précédente. Supposons le résultat vrai pour n . Alors

$$\begin{aligned} \Gamma\left(n + 1 + \frac{1}{2}\right) &= \left(n + \frac{1}{2}\right) \Gamma\left(n + \frac{1}{2}\right) = \frac{2n + 1}{2} \frac{(2n)! \sqrt{\pi}}{4^n n!} \\ &= \frac{2n + 2}{2n + 2} \frac{2n + 1}{2} \frac{(2n)! \sqrt{\pi}}{4^n n!} \\ &= \frac{(2n + 2)(2n + 1)}{4(n + 1)} \frac{(2n)! \sqrt{\pi}}{4^n n!} \\ &= \frac{(2n + 2)! \sqrt{\pi}}{4^{n+1} (n + 1)!}. \end{aligned}$$

Ainsi, la formule est vraie au rang $n + 1$, et par le principe de récurrence,

$$\forall n \in \mathbf{N}, \Gamma\left(n + \frac{1}{2}\right) = \frac{(2n)! \sqrt{\pi}}{4^n n!}.$$

SOLUTION DE L'EXERCICE 4.14

1. La fonction $t \mapsto \frac{e^{-tx^2}}{1 + t^3}$ est continue sur $[1; +\infty[$.

Au voisinage de $+\infty$, on a $t^2 \frac{e^{-tx^2}}{1 + t^3} \underset{t \rightarrow +\infty}{\sim} \frac{e^{-tx^2}}{t} \underset{t \rightarrow +\infty}{\rightarrow} 0$.

Et donc $\frac{e^{-tx^2}}{1 + t^3} = o_{t \rightarrow +\infty}\left(\frac{1}{t^2}\right)$.

Puisque $\int_1^{+\infty} \frac{dt}{t^2}$ converge, il en est de même de $\int_1^{+\infty} \frac{e^{-xt^2}}{1 + t^3} dt$.

2. Soit $x \in \mathbf{R}$. Alors, pour tout $t \geq 1$, il vient $\frac{e^{-t(-x)^2}}{1 + t^3} = \frac{e^{-tx^2}}{1 + t^3}$ et donc

$$F(-x) = \int_1^{+\infty} \frac{e^{-t(-x)^2}}{1 + t^3} dt = \int_1^{+\infty} \frac{e^{-tx^2}}{1 + t^3} dt = F(x).$$

Et donc F est paire.

Rappel

On a $\Gamma(n) = (n - 1)!$ et $\int_0^{+\infty} t^n e^{-t} dt = \Gamma(n+1) = n!$

¹⁰ L'intégrale avant changement de variable est déjà convergente, donc il n'y a aucune étude de convergence à réaliser.

Méthode

Si on a anticipé le résultat, on sait qu'il faudra à la fin obtenir du $(2n + 2)!$ au numérateur. Or on a déjà $(2n)! \times (2n + 1) = (2n + 1)!$. Il ne manque donc plus que le terme en $2n + 2$. Pour le faire apparaître, il suffit de multiplier numérateur et dénominateur par $(2n + 2)$.

3. Soient $x < y$ deux réels strictement positifs. Alors, pour tout $t \geq 0$, on a

$$\frac{e^{-tx^2}}{1+t^3} \geq \frac{e^{-ty^2}}{1+t^3}$$

de sorte que par croissance de l'intégrale,

$$F(x) = \int_1^{+\infty} \frac{e^{-xt^2}}{1+t^3} dt \geq \int_1^{+\infty} \frac{e^{-yt^2}}{1+t^3} dt = F(y).$$

Donc F est décroissante sur \mathbf{R}_+ .

D'autre part, pour $x \in \mathbf{R}_+$, on a, par croissance de l'intégrale,

$$0 \leq F(x) = \int_1^{+\infty} \frac{e^{-xt^2}}{1+t^3} dt \leq \int_1^{+\infty} \frac{e^{-xt}}{1+t^3} dt \leq \int_1^{+\infty} \frac{e^{-xt}}{1+t^3} dt \leq \int_1^{+\infty} e^{-xt} dt \leq \int_0^{+\infty} e^{-xt} dt = \frac{1}{x}.$$

Et donc en particulier, lorsque $x \rightarrow +\infty$, par le théorème des gendarmes, $\lim_{x \rightarrow +\infty} F(x) = 0$.

Par parité de F , son tableau de variation est donc donné par :

x	$-\infty$	0	$+\infty$
F	0	$F(0)$	0

4.a. La fonction $t \mapsto e^{-t}$ est de classe \mathcal{C}^1 sur \mathbf{R}_+ , et sa dérivée est majorée en valeur absolue par 1.

D'après l'inégalité des accroissements finis¹¹, on a alors, pour tout $(a, b) \in \mathbf{R}_+$,

$$|e^{-a} - e^{-b}| \leq |a - b|.$$

¹¹ Rappelons qu'il s'agit en fait de l'inégalité de Taylor-Lagrange à l'ordre 0.

4.b. Soient $(x, y) \in \mathbf{R}^2$. Alors, par la question précédente, pour tout $t \in [1, +\infty[$, on a

$$|e^{-tx^2} - e^{-ty^2}| \leq |tx^2 - ty^2| = t|x^2 - y^2|.$$

Et donc

$$\left| \frac{e^{-tx^2} - e^{-ty^2}}{1+t^3} \right| \leq \frac{t}{1+t^3} |x^2 - y^2| \leq \frac{1}{t^2} |x^2 - y^2|.$$

Alors, par croissance de l'intégrale, on a

$$\int_1^{+\infty} \left| \frac{e^{-tx^2} - e^{-ty^2}}{1+t^3} \right| dt \leq |x^2 - y^2| \int_1^{+\infty} \frac{dt}{t^2} \leq |x^2 - y^2|.$$

Enfin, on a

$$\begin{aligned} |F(x) - F(y)| &= \left| \int_1^{+\infty} \frac{e^{-tx^2} - e^{-ty^2}}{1+t^3} dt \right| \\ &\leq \int_1^{+\infty} \left| \frac{e^{-tx^2} - e^{-ty^2}}{1+t^3} \right| dt \leq |x^2 - y^2|. \end{aligned}$$

4.c. Soit $x \in \mathbf{R}$ fixé. Alors, lorsque $y \rightarrow x$, on $|x^2 - y^2| \rightarrow 0$ et donc, par le théorème des gendarmes,

$$\lim_{y \rightarrow x} |F(x) - F(y)| = 0 \Leftrightarrow \lim_{y \rightarrow x} F(y) = F(x).$$

Et donc F est continue en x . Ceci étant vrai pour tout $x \in \mathbf{R}$, on en déduit que F est continue sur \mathbf{R} .

⚠ Danger

Attention à ne pas affirmer que la valeur absolue de l'intégrale est égale à l'intégrale de la valeur absolue.
 En revanche, l'inégalité triangulaire pour les intégrales nous affirme que la valeur absolue de l'intégrale est **inférieure ou égale** à l'intégrale de la valeur absolue.

SOLUTION DE L'EXERCICE 4.15

1. Soit $A \geq 1$. Procédons à une intégration par parties sur le segment $[1, A]$ en posant $u(t) = \frac{1}{t}$ et $v(t) = \sin(t)$, qui sont deux fonctions de classe \mathcal{C}^1 avec $u'(t) = -\frac{1}{t^2}$ et $v'(t) = \cos(t)$. On a alors

$$\begin{aligned} \int_1^A \frac{\cos t}{t} dt &= \left[\frac{\sin(t)}{t} \right]_1^A + \int_1^A \frac{\cos t}{t^2} dt \\ &= \frac{\sin(A)}{A} - \sin(1) + \int_1^A \frac{\cos t}{t^2} dt. \end{aligned}$$

Puisque $\forall t \geq 1, \left| \frac{\cos t}{t^2} \right| \leq \frac{1}{t^2}$, $\int_1^{+\infty} \frac{\cos t}{t^2} dt$ est absolument convergente, donc convergente, de sorte que $\lim_{A \rightarrow +\infty} \int_1^A \frac{\cos t}{t^2} dt = \int_1^{+\infty} \frac{\cos t}{t^2} dt$.

D'autre part, \sin est bornée, et $\frac{1}{A} \xrightarrow{A \rightarrow +\infty} 0$, de sorte que $\frac{\sin(A)}{A} \xrightarrow{A \rightarrow +\infty} 0$.

Et donc

$$\int_1^A \frac{\cos t}{t} dt \xrightarrow{A \rightarrow +\infty} -\sin(1) + \int_1^{+\infty} \frac{\cos t}{t^2} dt.$$

Ceci prouve donc la convergence de I .

La fonction $t \mapsto \frac{\cos(t) - \cos(\alpha t)}{t}$ est continue sur $]0, +\infty[$.

En $+\infty$, on peut prouver comme précédemment la convergence de $\int_1^{+\infty} \frac{\cos(\alpha t)}{t} dt$, ce qui, couplé à la convergence de I , nous donne la convergence de $\int_0^{+\infty} \frac{\cos t - \cos(\alpha t)}{t} dt$.

En 0, on a

$$\cos(t) - \cos(\alpha t) = 1 - \frac{t^2}{2} + o_{t \rightarrow 0}(t^2) - 1 + \frac{\alpha^2 t^2}{2} + o_{t \rightarrow 0}(t^2) = (1 - \alpha^2) \frac{t^2}{2} + o_{t \rightarrow 0}(t^2) \sim_{t \rightarrow 0} \frac{(1 - \alpha^2)t^2}{2}.$$

Et donc $\frac{\cos t - \cos(\alpha t)}{t} \sim_{t \rightarrow 0} \frac{(1 - \alpha^2)t}{2} \xrightarrow{t \rightarrow 0} 0$.

Ainsi, $\int_0^1 \frac{\cos t - \cos(\alpha t)}{t} dt$ est faussement impropre, donc convergente.

2. On a $\int_{\varepsilon}^{+\infty} \frac{\cos t - \cos(\alpha t)}{t} dt = \int_{\varepsilon}^{+\infty} \frac{\cos t}{t} dt - \int_{\varepsilon}^{+\infty} \frac{\cos(\alpha t)}{t} dt$.

Dans cette seconde intégrale, procédons au changement de variable¹² $x = \alpha t$.

Alors

$$\int_{\varepsilon}^{+\infty} \frac{\cos(\alpha t)}{t} dt = \int_{\alpha\varepsilon}^{+\infty} \frac{\cos u}{\frac{u}{\alpha}} \frac{du}{\alpha} = \int_{\alpha\varepsilon}^{+\infty} \frac{\cos u}{u} du.$$

Et donc par la relation de Chasles,

$$J = \int_{\varepsilon}^{+\infty} \frac{\cos t}{t} dt - \int_{\alpha\varepsilon}^{+\infty} \frac{\cos u}{u} du = \int_{\varepsilon}^{\alpha\varepsilon} \frac{\cos t}{t} dt.$$

3. Soit $\varepsilon \leq \frac{\pi}{\alpha}$, de sorte que $[\varepsilon, \alpha\varepsilon] \subset [0, \pi]$.

Alors sur l'intervalle $[\varepsilon, \alpha\varepsilon]$, la fonction \cos est décroissante, de sorte que $\cos(\alpha\varepsilon) \leq \cos(t) \leq \cos(\varepsilon)$.

Par croissance de l'intégrale, on a donc

$$\int_{\varepsilon}^{\alpha\varepsilon} \frac{\cos(\alpha\varepsilon)}{t} dt \leq \int_{\varepsilon}^{\alpha\varepsilon} \frac{\cos t}{t} dt \leq \int_{\varepsilon}^{\alpha\varepsilon} \frac{\cos \varepsilon}{t} dt.$$

Soit encore

$$\cos(\alpha\varepsilon) \int_{\varepsilon}^{\alpha\varepsilon} \frac{dt}{t} \leq \int_{\varepsilon}^{\alpha\varepsilon} \frac{\cos t}{t} dt \leq \cos(\varepsilon) \int_{\varepsilon}^{\alpha\varepsilon} \frac{dt}{t} \Leftrightarrow \cos(\varepsilon) \ln(\alpha) \leq \int_{\varepsilon}^{\alpha\varepsilon} \frac{\cos t}{t} dt \leq \cos(\varepsilon) \ln(\alpha).$$

En faisant tendre ε vers 0, on a $\lim_{\varepsilon \rightarrow 0} \cos(\varepsilon) = \lim_{\varepsilon \rightarrow 0} \cos(\alpha\varepsilon) = 1$ et donc par le théorème des gendarmes,

$$J = \lim_{\varepsilon \rightarrow 0} \int_{\varepsilon}^{+\infty} \frac{\cos(t) - \cos(\alpha t)}{t} dt = \lim_{\varepsilon \rightarrow 0} \int_{\varepsilon}^{\alpha\varepsilon} \frac{\cos t}{t} dt = \ln(\alpha).$$

Remarque

Si $\alpha = 1$, on trouve un équivalent nul, mais la fonction de départ était nulle, donc il n'y a pas de problème.

¹² Légitime car affine.

Intuition

L'idée ici est que si ε est petit, l'intervalle $[\varepsilon, \alpha\varepsilon]$ est petit, et sur cet intervalle, $\cos(t)$ est proche de 1, de sorte que $\frac{\cos t}{t}$ est proche de $\frac{1}{t}$, ce qui va faire apparaître le \ln attendu.

SOLUTION DE L'EXERCICE 4.16

1. La fonction $x \mapsto \frac{xe^{-nx}}{e^x - 1}$ est continue sur $]0; +\infty[$.

De plus, pour $x \geq 0$, on a $e^x - 1 \geq x$. En effet, posons $g(x) = e^x - 1 - x$. Alors g est dérivable sur \mathbf{R}_+ , de dérivée égale à $g'(x) = e^x - 1 \geq 0$.

Donc g est croissante sur \mathbf{R}_+ , et $g(0) = 0$, de sorte que pour tout $x \geq 0$

$$g(x) \geq 0 \Leftrightarrow e^x - 1 \geq x.$$

Et donc pour tout $x \geq 0$, $0 \leq \frac{x}{e^x - 1} \leq 1$.

On en déduit que, quel que soit $n \in \mathbf{N}$, $0 \leq \frac{xe^{-nx}}{e^x - 1} \leq e^{-nx}$.

Mais $\int_0^{+\infty} e^{-nx} dx$ converge et vaut $\frac{1}{n}$, de sorte que I_n converge et $0 \leq I_n \leq \frac{1}{n}$.

2. Notons que pour tout $x > 0$,

$$\sum_{k=1}^n xe^{-kx} = x \sum_{k=1}^n (e^{-x})^k = xe^{-x} \frac{1 - e^{-nx}}{1 - e^{-x}} = \frac{x(1 - e^{-nx})}{e^x - 1} = f(x) - f(x)e^{-nx}.$$

3. En intégrant la relation précédente, on a donc

$$\int_0^{+\infty} f(x) dx = I_n + \sum_{k=1}^n \int_0^{+\infty} xe^{-kx} dx.$$

Procédons au changement de variable $t = kx$ dans l'intégrale $\int_0^{+\infty} xe^{-kx} dx$. Ce changement de variable est légitime car affine, et on a donc

$$\int_0^{+\infty} xe^{-kx} dx = \int_0^{+\infty} \frac{t}{k} e^{-t} \frac{dt}{k} = \frac{1}{k^2} \int_0^{+\infty} te^{-t} dt = \frac{1}{k^2} \Gamma(2) = \frac{1}{k^2}.$$

Ainsi, pour tout $n \in \mathbf{N}^*$, on a

$$\int_0^{+\infty} f(t) dt = I_n + \sum_{k=1}^n \frac{1}{k^2}.$$

D'après la question 1, on a $\lim_{n \rightarrow +\infty} I_n = 0$ et donc

$$\int_0^{+\infty} f(t) dt = \lim_{n \rightarrow +\infty} \sum_{k=1}^n \frac{1}{k^2} = \sum_{k=1}^{+\infty} \frac{1}{k^2} = \frac{\pi^2}{6}.$$

4. La fonction $t \mapsto -\ln(t)$ est de classe \mathcal{C}^1 sur $]0, 1]$ et y est strictement décroissante, donc le changement de variable est légitime.

Lorsque $t \rightarrow 0$, alors $-\ln(t) \rightarrow +\infty$ et lorsque $t \rightarrow 1$, alors $-\ln(t) \rightarrow 0$.

En posant $x = -\ln(t)$, il vient $t = e^{-x}$ et donc $dt = -e^{-x} dx$. Ainsi,

$$\int_0^1 \frac{\ln t}{1-t} dt = \int_{+\infty}^0 \frac{-x}{1 - e^{-x}} (-e^{-x} dx) = - \int_0^{+\infty} \frac{x}{e^x - 1} dx = -\frac{\pi^2}{6}.$$

SOLUTION DE L'EXERCICE 4.17

La fonction $t \mapsto \ln te^{-t}$ est continue sur $]0; +\infty[$.

Au voisinage de 0, on a $\ln te^{-t} \underset{t \rightarrow 0^+}{\sim} \ln t$.

Puisque ces deux fonctions sont de signe constant sur $]0, 1]$, les intégrales $\int_0^1 \ln t dt$ et

$\int_0^1 \ln te^{-t} dt$ sont de même nature.

Or, pour $A \in]0, 1]$, on a

$$\int_A^1 \ln t dt = [t \ln t - t]_A^1 = -1 - A \ln(A) + A \xrightarrow{A \rightarrow 0^+} -1.$$

Remarque

$e^x \geq x + 1$ est aussi une inégalité de convexité : l'exponentielle étant convexe, sa courbe est au dessus de toutes ses tangentes. En particulier, sa tangente en $x = 0$ est la droite d'équation $y = x + 1$.

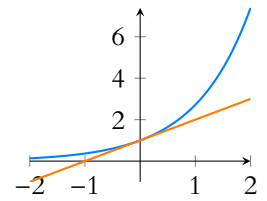


FIGURE 4.1– La convexité de e^x .

Remarque

La convergence de l'intégrale obtenue après changement de variable avait déjà été établie, donc l'intégrale avant changement de variable converge.

Donc $\int_0^1 \ln t \, dt$ converge, et par conséquent il en est de même de $\int_0^1 \ln te^{-t} \, dt$.
 Au voisinage de $+\infty$, on a

$$t^2 \ln te^{-t} \frac{\ln t}{t} \times t^3 e^{-t} \xrightarrow{t \rightarrow +\infty} 0.$$

Et donc $\ln te^{-t} = o_{t \rightarrow +\infty} \left(\frac{1}{t^2} \right)$.

Puisque $\int_1^{+\infty} \frac{1}{t^2} \, dt$ converge, il en est de même de $\int_1^{+\infty} \ln te^{-t} \, dt$ et donc de $\int_0^{+\infty} \ln te^{-t} \, dt$.

La fonction \ln est concave sur \mathbf{R}_+^* , et donc sa courbe représentative est située sous ses tangentes.

En particulier, la tangente en $x = 1$ est la droite d'équation $y = x - \ln$ de sorte que pour tout $t > 0$, $\ln t \leq t - 1$.

Et donc pour tout $t > 0$, $\ln te^{-t} \leq (t - 1)e^{-t} \leq te^{-t} - e^{-t}$.

On en déduit¹³ que

$$\int_0^{+\infty} \ln te^t \, dt \leq \int_0^{+\infty} te^{-t} \, dt - \int_0^{+\infty} e^{-t} \, dt = \Gamma(2) - \Gamma(1) = 1! - 0! = 0.$$

SOLUTION DE L'EXERCICE 4.18

Un premier réflexe serait de dire que pour tout $t \geq 1$, $\frac{f(t)}{t} \leq f(t)$ et de penser que ceci suffit à conclure.

Malheureusement cette inégalité n'est vraie que si $f(t) \geq 0$, et de toutes façons, pour prouver la convergence d'une intégrale par domination, il faut la positivité de tous les termes.

Une autre erreur à ne pas faire serait de dire que puisque $\int_1^{+\infty} f(t) \, dt$ converge, alors $\lim_{x \rightarrow +\infty} f(x) = 0$, et s'en tirer à l'aide de majorations.

En effet, nous connaissons un résultat similaire pour les séries, mais il n'en existe pas pour les intégrales : il existe des fonctions f positives, telles que $\int_0^{+\infty} f(t) \, dt$ converge, mais qui n'admettent pas de limite en $+\infty$.

Notons F la fonction définie sur $[1, +\infty[$ par $F(x) = \int_1^x f(t) \, dt$.

Alors F est la primitive de f qui s'annule en $x = 1$. En particulier, c'est une fonction de classe \mathcal{C}^1 .

Une intégration par parties nous donne alors, pour tout $x \geq 1$,

$$\int_1^x \frac{f(t)}{t} \, dt = \left[\frac{F(t)}{t} \right]_1^x + \int_1^x \frac{F(t)}{t^2} \, dt = \frac{F(x)}{x} + \int_1^x \frac{F(t)}{t^2} \, dt.$$

Mais puisque $\int_1^{+\infty} f(t) \, dt$ converge, $F(x)$ admet une limite ℓ lorsque $x \rightarrow +\infty$.

Et alors, par quotient de limites, $\frac{F(x)}{x} \xrightarrow{x \rightarrow +\infty} 0$.

D'autre part, il existe¹⁴ $A > 1$ tel que pour $x \geq A$, $\ell - 1 \leq F(x) \leq \ell + 1$.

Notons alors $M = \max(|\ell - 1|, |\ell + 1|)$, de sorte que pour tout $t \geq A$, $|F(t)| \leq M$.

Par conséquent, pour tout $t \geq A$, $\left| \frac{F(t)}{t^2} \right| \leq \frac{M}{t^2}$.

Et donc par comparaison à une intégrale de Riemann, $\int_A^{+\infty} \frac{F(t)}{t^2} \, dt$ converge absolument et donc converge.

Il en est donc de même de $\int_1^{+\infty} \frac{F(t)}{t^2} \, dt$. Et donc $\int_1^x \frac{F(t)}{t^2} \, dt$ admet une limite finie lorsque $x \rightarrow +\infty$.

Et donc $\int_1^x \frac{f(t)}{t} \, dt$ admet une limite lorsque $x \rightarrow +\infty$ de sorte que $\int_1^{+\infty} \frac{f(t)}{t} \, dt$ converge.

SOLUTION DE L'EXERCICE 4.19

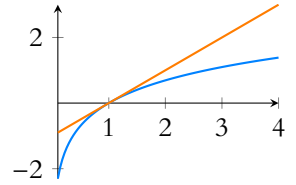


FIGURE 4.2– $\ln(x) \leq x - 1$.

¹³ Toutes les intégrales en jeu sont convergentes.

Détails

On peut par exemple regarder la fonction qui vaut k sur $\left[k, k + \frac{1}{k^3} \right]$ ($k \in \mathbf{N}^*$) et 0 ailleurs.

Nous ne détaillons pas tout, mais en essayant de la représenter et de calculer l'aire sous la courbe, vous devriez voir apparaître le lien avec une série de Riemann convergente.

Notons qu'il est également possible de trouver des fonctions continues d'intégrale convergente et sans limite en $+\infty$.

¹⁴ C'est la définition de limite, dans laquelle on a pris $\varepsilon = 1$.

1. La fonction $t \mapsto \frac{t}{1+xe^t}$ est continue sur $[0, +\infty[$, donc le seul éventuel problème de convergence est au voisinage de $+\infty$.

Or, au voisinage de $+\infty$, on a $\frac{t}{1+xe^t} \underset{t \rightarrow +\infty}{\sim} \frac{t}{xe^t} = \frac{1}{x}te^{-t}$.

Puisque $\int_0^{+\infty} te^{-t} dt$ converge¹⁵, par critère de comparaison pour les fonctions positives,

$\int_0^{+\infty} \frac{t}{1+xe^t} dt$ converge.

2. Soient $x < y$ deux réels strictement positifs.

Alors pour tout $t \geq 0$, on a $\frac{t}{1+xe^t} \geq \frac{t}{1+ye^t}$.

D'autre part, pour $t = 1$, on a $\frac{1}{1+xe} > \frac{1}{1+ye}$.

Autrement dit, la fonction $t \mapsto \frac{t}{1+xe^t} - \frac{t}{1+ye^t}$ est continue sur \mathbf{R}_+^* , positive, et non nulle.

Donc son intégrale, qui est $f(x) - f(y)$ est positive et non nulle, donc strictement positive. Ainsi, $f(x) > f(y)$: f est strictement décroissante sur \mathbf{R}_+^* .

- 3.a. Pour tout $x \in \mathbf{R}_+^*$ et tout $t \in \mathbf{R}_+$, on a

$$0 \leq \frac{t}{1+xe^t} \leq \frac{t}{xe^t} = \frac{1}{x}te^{-t}.$$

Et donc par croissance de l'intégrale,

$$0 \leq f(x) \leq \frac{1}{x} \int_0^{+\infty} te^{-t} dt \leq \frac{1}{x} \Gamma(2) = \frac{1}{x}.$$

Et donc par le théorème des gendarmes, $\lim_{x \rightarrow +\infty} f(x) = 0$.

On a alors

$$\begin{aligned} |xf(x) - 1| &= \left| \int_0^{+\infty} \frac{xt}{1+xe^t} dt - \int_0^{+\infty} te^{-t} dt \right| \\ &= \left| \int_0^{+\infty} t \frac{x - e^{-t} - x}{1+xe^t} dt \right| \\ &= \left| \int_0^{+\infty} \frac{-te^{-t}}{1+xe^t} dt \right| \\ &\leq \int_0^{+\infty} \frac{te^{-t}}{1+xe^t} dt \\ &\leq \frac{1}{x} \int_0^{+\infty} te^{-2t} dt. \end{aligned}$$

Et donc en particulier¹⁶, lorsque $x \rightarrow +\infty$, $xf(x) - 1 \xrightarrow{x \rightarrow +\infty} 0$.

Soit encore $xf(x) \xrightarrow{x \rightarrow +\infty} 1$, de sorte que $f(x) \underset{x \rightarrow +\infty}{\sim} \frac{1}{x}$.

- 3.b. Comme nous savons que f est décroissante, par le théorème de la limite monotone, soit elle admet une limite finie ℓ en 0, soit $\lim_{x \rightarrow 0^+} f(x) = +\infty$.

Supposons donc que $\lim_{x \rightarrow 0^+} f(x) = \ell$, avec $\ell \in \mathbf{R}$.

Puisque la fonction intégrée est positive, pour tout $A \geq 0$, on a

$$f(x) = \int_0^{+\infty} \frac{t}{1+xe^t} dt \geq \int_0^A \frac{t}{1+xe^t} dt.$$

Mais pour $t \in [0, A]$, $e^t \leq e^A$ et donc

$$\frac{t}{1+xe^A} \leq \frac{t}{1+xe^t}.$$

¹⁵ C'est $\Gamma(2)$.

⚠ Danger !

Une grosse erreur ici serait de dériver (par rapport à x) ce qui se trouve sous l'intégrale et donc d'affirmer que

$$f'(x) = - \int_0^{+\infty} \frac{te^t}{(1+xe^t)^2} dt.$$

En effet, nous ne disposons d'aucun résultat nous assurant que nous avons bien le droit de faire ceci, autrement dit, la dérivée d'une intégrale n'est pas toujours l'intégrale de la dérivée.

Remarque

C'est l'inégalité triangulaire qui nous permet d'affirmer ceci, mais puisque la fonction intégrée est toujours négative, l'intégrale est négative, et donc sa valeur absolue est égale à son opposée. On pourrait donc affirmer qu'il y a égalité, et non juste une inégalité.

¹⁶ L'intégrale qui apparaît dans la majoration ci-dessus est une constante, qui ne dépend pas de x . Il s'agit alors juste d'appliquer le théorème des gendarmes.

Limite monotone

Attention au fait qu'on est ici en 0, et donc à la borne de gauche de l'intervalle de définition de f . Si elle n'admet pas de limite finie, elle tend vers $+\infty$ (alors que si on s'intéressait à la limite en $+\infty$, elle serait soit finie, soit égale à $-\infty$). En cas de doute, dessinez au brouillon une fonction décroissante qui n'admet pas de limite finie en 0 !

Par croissance de l'intégrale, on en déduit que

$$\int_0^A \frac{t}{1+xe^t} dt \geq \int_0^A \frac{t}{1+xe^A} dt = \frac{A^2}{2} \frac{1}{1+xe^A}.$$

Nous venons donc de prouver que pour tout $x \in \mathbf{R}_+^*$ et tout $A \in \mathbf{R}_+$, $f(x) \geq \frac{A^2}{2(1+xe^A)}$.

Et donc en faisant tendre x vers 0, il vient $\ell \geq \frac{A^2}{2}$.

Or cette inégalité ne peut pas être vérifiée pour tout $A \geq 0$, puisque $\frac{A^2}{2} \xrightarrow{A \rightarrow +\infty} +\infty$.

C'est donc que notre hypothèse de départ est fautive : f n'admet pas de limite finie en 0 et donc $\lim_{x \rightarrow 0^+} f(x) = +\infty$.

4.a. La fonction $t \mapsto xe^t$ est \mathcal{C}^1 et strictement croissante sur \mathbf{R}_+ , donc le changement de variable est bien légitime.

Posons alors $u = xe^t$, de sorte que $t = \ln\left(\frac{u}{x}\right) = \ln(u) - \ln(x)$ et donc $dt = \frac{du}{u}$.

On a donc

$$\begin{aligned} f(x) &= \int_0^{+\infty} \frac{t}{1+xe^t} dt \\ &= \int_x^{+\infty} \frac{\ln(u) - \ln(x)}{1+u} \frac{du}{u} \\ &= \int_x^{+\infty} (\ln(u) - \ln(x)) \frac{1}{u(1+u)} du \\ &= \int_x^{+\infty} \frac{\ln u}{u(1+u)} du - \ln(x) \int_x^{+\infty} \left(\frac{1}{u} - \frac{1}{1+u}\right) du. \end{aligned}$$

Or, pour $A \geq x$, on a

$$\begin{aligned} \int_x^A \left(\frac{1}{u} - \frac{1}{1+u}\right) du &= [\ln(u) - \ln(1+u)]_x^A = \ln(A) - \ln(1+A) + \ln(x) - \ln(1+x) \\ &= \ln\left(\underbrace{\frac{A}{1+A}}_{\xrightarrow{A \rightarrow +\infty} 1}\right) + \ln(x) - \ln(1+x) \xrightarrow{A \rightarrow +\infty} \ln(x) - \ln(1+x). \end{aligned}$$

Et donc on a bien prouvé que

$$f(x) = \ln(x) (\ln(x) - \ln(1+x)) + \int_x^{+\infty} \frac{\ln(u)}{u(1+u)} du.$$

4.b. Notons que $\int_x^{+\infty} \frac{\ln u}{u(1+u)} du = \int_1^{+\infty} \frac{\ln u}{u(1+u)} du - \int_1^x \frac{\ln u}{u(1+u)} du$.

Or, par le théorème fondamental de l'analyse, $x \mapsto \int_1^x \frac{\ln u}{u(1+u)} du$ est \mathcal{C}^1 sur \mathbf{R}_+^* , de dérivée égale à $x \mapsto \frac{\ln(x)}{x(1+x)}$.

Et donc $x \mapsto \int_x^{+\infty} \frac{\ln u}{u(1+u)} du$ est également \mathcal{C}^1 , de dérivée égale à $x \mapsto -\frac{\ln(x)}{x(1+x)}$.

On en déduit que f est dérivable sur \mathbf{R}_+^* car somme et produit de fonctions dérivables sur \mathbf{R}_+^* , avec

$$\begin{aligned} f'(x) &= \frac{1}{x} (\ln(x) - \ln(1+x)) + \ln(x) \left(\frac{1}{x} - \frac{1}{1+x}\right) - \frac{\ln(x)}{x(1+x)} \\ &= \frac{1}{x} \ln\left(\frac{x}{1+x}\right). \end{aligned}$$

Or, pour $x > 0$, $\frac{x}{1+x} < 1$, de sorte que $\ln\left(\frac{x}{1+x}\right) < 0$, et donc $f'(x) < 0$.

On retrouve alors que f est strictement décroissante sur \mathbf{R}_+^* .

Convergence

La convergence de cette seconde intégrale est garantie par le théorème de changement de variable puisque l'intégrale de départ, $f(x)$, est convergente.

Astuce

$$\frac{1}{u(1+u)} = \frac{1}{u} - \frac{1}{1+u}.$$

Convergence

Disons tout de même deux mots de la convergence de cette dernière intégrale : au voisinage de $+\infty$,

$$\frac{\ln u}{u(1+u)} \sim \frac{\ln u}{u^2} = o\left(\frac{1}{u^{3/2}}\right).$$

MATRICES

Cours	143
5.1 Rappels.	143
5.2 Matrices et applications linéaires	147
5.3 Trace d'une matrice carrée	154
5.4 Hors programme : suites de matrices	156
Méthodes	158
Exercices.	160
Correction des exercices	163

Dans tout ce chapitre, \mathbf{K} désignera indifféremment \mathbf{R} ou \mathbf{C} . De plus, on supposera que tous les espaces vectoriels que l'on manipule sont de dimension finie. Les propositions non démontrées ici l'ont été en première année.

5.1 RAPPELS

5.1.1 Définition, opérations usuelles

Définition 5.1 – On appelle matrice à n lignes et p colonnes à coefficients dans \mathbf{K} un tableau de $n \times p$ éléments de \mathbf{K} .
L'ensemble des matrices à n lignes et p colonnes à coefficients dans \mathbf{K} est noté $\mathcal{M}_{n,p}(\mathbf{K})$. Dans le cas où $n = p$, on note $\mathcal{M}_n(\mathbf{K})$ au lieu de $\mathcal{M}_{n,n}(\mathbf{K})$.

Définition 5.2 – Si $n \in \mathbf{N}^*$, on note $I_n = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & & \vdots \\ 0 & 0 & \ddots & \vdots \\ 0 & \dots & 0 & 1 \end{pmatrix}$ et $0_n = \begin{pmatrix} 0 & \dots & 0 \\ \vdots & & \vdots \\ 0 & \dots & 0 \end{pmatrix}$.

Définition 5.3 –

- Soit $A = (a_{i,j}) \in \mathcal{M}_{n,p}(\mathbf{K})$ et $\lambda \in \mathbf{K}$. On note $\lambda A = (\lambda a_{i,j}) \in \mathcal{M}_{n,p}(\mathbf{K})$.
- Si $A = (a_{i,j}), B = (b_{i,j}) \in \mathcal{M}_{n,p}(\mathbf{K})$, on pose $A + B = (a_{i,j} + b_{i,j}) \in \mathcal{M}_{n,p}(\mathbf{K})$.

Proposition 5.4 : Muni de ces opérations, $\mathcal{M}_{n,p}(\mathbf{K})$ est un \mathbf{K} -espace vectoriel. De plus, il est de dimension finie, et $\dim \mathcal{M}_{n,p}(\mathbf{K}) = n \times p$. En particulier, $\dim \mathcal{M}_n(\mathbf{K}) = n^2$.

Définition 5.5 – Soit $A = (a_{i,j}) \in \mathcal{M}_{n,p}(\mathbf{K})$ et $B = (b_{i,j}) \in \mathcal{M}_{p,q}(\mathbf{K})$. Alors on note $A \times B = (c_{i,j}) \in \mathcal{M}_{n,q}(\mathbf{K})$ la matrice définie par

$$c_{i,j} = \sum_{k=1}^p a_{i,k} b_{k,j}.$$

Formule

C'est cette formule qui justifie la manière dont on fait le produit de deux matrices : on fait le produit des coefficients de la $i^{\text{ème}}$ ligne de A par la $j^{\text{ème}}$ colonne de B .
Si on ne revient en pratique jamais à la formule pour faire le produit de deux matrices 3×3 , il est tout de même bon de la connaître car elle sert régulièrement dans des problèmes théoriques.

Proposition 5.6 :

- Si $A, B \in \mathcal{M}_{n,p}(\mathbf{K})$ et si $C \in \mathcal{M}_{p,q}(\mathbf{K})$, alors $(A+B)C = AC + BC$.
- si $A \in \mathcal{M}_{n,p}(\mathbf{K})$, et si $B, C \in \mathcal{M}_{p,q}(\mathbf{K})$, alors $A(B+C) = AB + AC$.
- si $A \in \mathcal{M}_n(\mathbf{K})$, alors $AI_n = I_nA = A$ et $A0_n = 0_nA = 0_n$.

Remarques. • En général, $AB \neq BA$. Par exemple, si $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ et $B = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$, alors

$$AB = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \text{ et } BA = \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}.$$

- Si $A, B \in \mathcal{M}_n(\mathbf{K})$ vérifient $AB = BA$, on dit que A et B commutent.
- La matrice I_n , et de manière plus générale, la matrice $\lambda I_n, \lambda \in \mathbf{K}$ commute à toutes les matrices de $\mathcal{M}_n(\mathbf{K})$.

Définition 5.7 – Si $A \in \mathcal{M}_{n,p}(\mathbf{K})$, on appelle transposée de A , et on note ${}^tA = (b_{i,j})$ la matrice de $\mathcal{M}_{p,n}(\mathbf{K})$ définie par $b_{i,j} = a_{j,i}$.

Proposition 5.8 :

- Si $A, B \in \mathcal{M}_{n,p}(\mathbf{K})$, et si $\lambda \in \mathbf{K}$, alors

$${}^t(\lambda A + B) = \lambda {}^tA + {}^tB.$$

- Si $A \in \mathcal{M}_{n,p}(\mathbf{K})$ et $B \in \mathcal{M}_{p,q}(\mathbf{K})$, alors ${}^t(AB) = {}^tB {}^tA$.

Par cœur

Transposer une matrice, c'est échanger lignes et colonnes, donc on échange i et j . Il est important de connaître (ou de savoir retrouver) cette formule, et pas seulement de savoir calculer la transposée d'une matrice donnée.

En Scilab, la transposée d'une matrice A est obtenue avec la commande A' .

5.1.2 Matrices carrées

Définition 5.9 – Soit $A = (a_{i,j}) \in \mathcal{M}_n(\mathbf{K})$. On dit que A est :

- **scalaire** s'il existe $\lambda \in \mathbf{K}$ tel que $A = \lambda I_n$.
- **diagonale** si $a_{i,j} = 0$ pour $i \neq j$.

Dans ce cas, si $A = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$, on note $A = \text{diag}(\lambda_1, \dots, \lambda_n)$.

- **triangulaire supérieure (respectivement inférieure)** si $a_{i,j} = 0$ pour $i > j$ (resp. $j > i$).
- **symétrique** si $A = {}^tA$. On note $S_n(\mathbf{K})$ l'ensemble des matrices symétriques de $\mathcal{M}_n(\mathbf{K})$. C'est un sous-espace vectoriel car c'est le noyau de l'application linéaire $\varphi : \mathcal{M}_n(\mathbf{K}) \rightarrow \mathcal{M}_n(\mathbf{K})$ définie par $\varphi(A) = A - {}^tA$.
- **antisymétrique** si $A = -{}^tA$. On note $A_n(\mathbf{K})$ l'ensemble des matrices antisymétriques de $\mathcal{M}_n(\mathbf{K})$. C'est un sous-espace vectoriel de $\mathcal{M}_n(\mathbf{K})$ car c'est le noyau de $\psi : \mathcal{M}_n(\mathbf{K}) \rightarrow \mathcal{M}_n(\mathbf{K})$, l'application linéaire définie par $\psi(A) = A + {}^tA$.

Par cœur

Il ne suffit pas de savoir dire si une matrice 3×3 ou 4×4 est symétrique, il faut connaître la définition en terme de la transposée.

En particulier

Si A est antisymétrique, on a alors $a_{i,i} = -a_{i,i}$, et donc les coefficients diagonaux de A sont tous nuls.

Définition 5.10 – Une matrice $A \in \mathcal{M}_n(\mathbf{K})$ est dite inversible s'il existe $B \in \mathcal{M}_n(\mathbf{K})$ telle que $AB = BA = I_n$. Une telle matrice B est alors unique et s'appelle l'inverse de A , et on la note A^{-1} .

On note $GL_n(\mathbf{K})$ (et on appelle groupe linéaire d'ordre n sur \mathbf{K}) l'ensemble des matrices inversibles de $\mathcal{M}_n(\mathbf{K})$.

En Scilab, l'inverse d'une matrice A s'obtient par la commande $\text{inv}(A)$.



Remarque. En fait, si on prouve l'existence d'une matrice $B \in \mathcal{M}_n(\mathbf{K})$ telle que $AB = I_n$ ou $BA = I_n$, alors automatiquement l'autre égalité est vérifiée, et donc A est inversible et $B = A^{-1}$.

Pour prouver que A est inversible, on échelonne A à l'aide d'opérations sur les lignes. Les opérations permises sont :

- l'échange de deux lignes
- la multiplication d'une ligne par un scalaire **non nul**
- l'ajout à une ligne d'une combinaison linéaire des autres.

Alors A est inversible si et seulement si tous les coefficients diagonaux de la matrice échelonnée sont non nuls.

De plus, cette méthode permet d'obtenir facilement l'inverse de A lorsqu'il existe.

Méthode de calcul de A^{-1} : on écrit la matrice I_n à droite de la matrice A , et toujours à l'aide d'opérations sur les lignes, on transforme A en l'identité, tout en réalisant les mêmes opérations sur la matrice de droite.

À la fin du processus, la matrice de droite est A^{-1} .

Exemple 5.11

$$A = \begin{pmatrix} 3 & 0 & 4 \\ 3 & 1 & 3 \\ -1 & 0 & -1 \end{pmatrix}.$$

$$\begin{aligned} \left(\begin{array}{ccc|ccc} 3 & 0 & 4 & 1 & 0 & 0 \\ 3 & 1 & 3 & 0 & 1 & 0 \\ -1 & 0 & -1 & 0 & 0 & 1 \end{array} \right) & \xleftrightarrow{L_3 \leftrightarrow L_1} \left(\begin{array}{ccc|ccc} -1 & 0 & -1 & 0 & 0 & 1 \\ 3 & 1 & 3 & 0 & 1 & 0 \\ 3 & 0 & 4 & 1 & 0 & 0 \end{array} \right) \\ & \xleftrightarrow{\substack{L_2 \leftarrow L_2 + 3L_1 \\ L_3 \leftarrow L_3 + 3L_1}} \left(\begin{array}{ccc|ccc} -1 & 0 & -1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 1 & 3 \\ 0 & 0 & 1 & 1 & 0 & 3 \end{array} \right) \\ & \xleftrightarrow{L_1 \leftarrow L_1 + L_3} \left(\begin{array}{ccc|ccc} -1 & 0 & 0 & 1 & 0 & 4 \\ 0 & 1 & 0 & 0 & 1 & 3 \\ 0 & 0 & 1 & 1 & 0 & 3 \end{array} \right) \\ & \xleftrightarrow{L_1 \leftarrow -L_1} \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & -1 & 0 & -4 \\ 0 & 1 & 0 & 0 & 1 & 3 \\ 0 & 0 & 1 & 1 & 0 & 3 \end{array} \right) \end{aligned}$$

$$\text{Ainsi, } A^{-1} = \begin{pmatrix} -1 & 0 & -4 \\ 0 & 1 & 3 \\ 1 & 0 & 3 \end{pmatrix}.$$

Méthode

Le début du calcul de l'inverse d'une matrice nécessite les mêmes opérations que celles nécessaires au calcul du rang.

Dans une question du type «déterminer si A est inversible, et si oui, calculer A^{-1} », il y a alors deux options :
1) calculer le rang par un pivot. Puis si A est inversible, calculer A^{-1} en commençant par faire une seconde fois les opérations nécessaires au calcul du rang.

2) être malin, et commencer directement un calcul d'inverse (par la méthode de son choix). Ce calcul d'inverse va nécessiter d'échelonner la matrice (ce qui permettra d'obtenir le rang et donc de savoir si l'on continue ou non les calculs). Si la matrice n'est pas inversible, on s'arrête là (et le calcul d'inverse n'a servi à rien), sinon on continue jusqu'à obtention de A^{-1} .

Proposition 5.12 : Si $A, B \in \mathcal{M}_n(\mathbf{K})$ sont inversibles, alors AB est inversible, et

$$(AB)^{-1} = B^{-1}A^{-1}.$$

Définition 5.13 – Si $A \in \mathcal{M}_n(\mathbf{K})$, alors on pose

- $A^0 = I_n$
- $\forall k \in \mathbf{N}, A^{k+1} = A^k A$

Si de plus A est inversible, alors pour $k \in \mathbf{N}$ on pose $A^{-k} = (A^{-1})^k$.

Remarque. Les puissances de A ne sont définies que si A est carrée.

De plus, toutes les puissances de A commutent entre elles. Plus généralement, si $P, Q \in \mathbf{K}[X]$, $P(A)$ et $Q(A)$ commutent.

5.1.3 Rang d'une matrice

Définition 5.14 – Soit E un \mathbf{K} -espace vectoriel et (e_1, \dots, e_n) une famille de vecteurs de E . Le rang de cette famille est la dimension de $\text{Vect}(e_1, \dots, e_n)$, que l'on note $\text{rg}(e_1, \dots, e_n)$.

Remarque. $\text{rg}(e_1, \dots, e_n)$ est le cardinal maximal d'une sous-famille libre de (e_1, \dots, e_n) . En particulier, une famille est libre si et seulement si son rang est égal à son cardinal.

Définition 5.15 – Si $M \in \mathcal{M}_{n,p}(\mathbf{K})$, on appelle rang de M , et on note $\text{rg}(M)$ le rang de la famille de ses vecteurs colonnes dans $\mathcal{M}_{n,1}(\mathbf{K})$.

Méthode : pour calculer le rang d'une matrice A , on échelonne la matrice par opérations sur les lignes. Le rang de A est alors le nombre de pivots non nuls de la matrice bien échelonnée obtenue.

Rappelons qu'une matrice échelonnée est une matrice telle que chaque ligne commence par strictement plus de zéros que la précédente. Une matrice triangulaire n'est donc pas

toujours échelonnée, par exemple $\begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$ n'est pas échelonnée.

Exemple 5.16

$$A = \begin{pmatrix} 1 & -1 & 0 & 0 \\ 0 & 1 & -2 & 1 \\ 1 & 1 & -4 & 2 \end{pmatrix}.$$

$$\begin{pmatrix} 1 & -1 & 0 & 0 \\ 0 & 1 & -2 & 1 \\ 1 & 1 & -4 & 2 \end{pmatrix} \xleftrightarrow{L_3 \leftarrow L_3 - L_1} \begin{pmatrix} 1 & -1 & 0 & 0 \\ 0 & 1 & -2 & 1 \\ 0 & 2 & -4 & 2 \end{pmatrix} \xleftrightarrow{L_3 \leftarrow L_3 - 2L_2} \begin{pmatrix} 1 & -1 & 0 & 0 \\ 0 & 1 & -2 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

Ainsi, $\text{rg}(A) = 2$.

Ceci signifie que la famille de ses vecteurs colonnes engendre un espace vectoriel de dimension 2, ce qui était prévisible car C_3 et C_4 sont combinaisons linéaires de C_1 et C_2 ($C_3 = -2C_1 - 2C_2$ et $C_4 = C_1 + C_2$).

Proposition 5.17 : Si $M \in \mathcal{M}_{n,p}(\mathbf{K})$, alors $\text{rg}(M) = \text{rg}({}^t M)$.

Remarque. Puisque la transposée échange lignes et colonnes et que $\text{rg}({}^t A) = \text{rg}(A)$:

- le rang d'une matrice est aussi le rang de la famille de ses vecteurs lignes
- on a le droit de procéder à des opérations élémentaires sur les colonnes.

Proposition 5.18 : Soit $A \in \mathcal{M}_{n,p}(\mathbf{K})$ une matrice non nulle. Alors A est de rang 1 si et seulement si ses colonnes sont deux à deux proportionnelles.

Démonstration. Notons A_1, \dots, A_p les colonnes de A .

Supposons qu'elles sont deux à deux proportionnelles.

Puisque A est non nulle, il existe $i \in \llbracket 1, p \rrbracket$ tel que $A_i \neq 0$.

Par proportionnalité des colonnes, pour tout $j \in \llbracket 1, p \rrbracket$, il existe λ_j tel que $A_j = \lambda_j A_i$.

Et donc $\text{Vect}(A_1, \dots, A_p) = \text{Vect}(A_i)$, qui est donc de dimension 1, car $A_i \neq 0$.

Par conséquent, $\text{rg}(A) = 1$.

Inversement, supposons que A soit de rang 1. Alors $\text{Vect}(A_1, \dots, A_n)$ est de dimension 1. L'une des colonnes A_i est non nulle, et donc forme une famille libre de $\text{Vect}(A_1, \dots, A_n)$. Puisque ce dernier est de dimension 1, la famille formée de la seule colonne A_i en est une base.

Et donc pour tout $j \in \llbracket 1, n \rrbracket$, $A_j \in \text{Vect}(A_1, \dots, A_n) = \text{Vect}(A_i)$.

Il existe donc $\lambda_j \in \mathbf{K}$ tel que $A_j = \lambda_j A_i$: toutes les colonnes de A sont proportionnelles à A_i , et donc sont deux à deux proportionnelles. \square

Remarque. Puisque le rang de A est aussi le rang de la famille de ses vecteurs lignes, les lignes de A sont aussi deux à deux proportionnelles si $\text{rg}(A) = 1$.

En Scilab le rang d'une matrice A s'obtient à l'aide de $\text{rank}(A)$.

Rédaction

Lorsqu'on échelonne une matrice, on prendra bien soin d'utiliser des symboles \leftrightarrow et non des égalités : la matrice échelonnée n'est bien entendu **pas égale** à la matrice de départ puisqu'on a changé ses coefficients. En revanche, leurs rangs sont égaux.

En particulier

Ce résultat sert notamment si toutes les colonnes de A sont égales.

Exemple 5.19

$$\text{Soit } A = \begin{pmatrix} 1 & 2 & \dots & n \\ 1 & 2 & \dots & n \\ \vdots & \vdots & & \vdots \\ 1 & 2 & \dots & n \end{pmatrix}$$

Alors les colonnes de A sont deux à deux proportionnelles¹, et donc $\text{rg}(A) = 1$.

¹ Et non toutes nulles.

5.2 MATRICES ET APPLICATIONS LINÉAIRES**5.2.1 Matrice d'une famille de vecteurs**

Définition 5.20 – Soit E un espace vectoriel de dimension finie, et soit $\mathcal{B} = (u_1, \dots, u_n)$ une base de E .

Si (e_1, \dots, e_p) est une famille de vecteurs, on appelle matrice de (e_1, \dots, e_p) dans la base \mathcal{B} , et on note $\text{Mat}_{\mathcal{B}}(e_1, \dots, e_p)$ la matrice de $\mathcal{M}_{n,p}(\mathbf{K})$ dont les colonnes sont $(e_1)_{\mathcal{B}}, \dots, (e_p)_{\mathcal{B}}$.

Autrement dit $\text{Mat}_{\mathcal{B}}(e_1, \dots, e_p) = (a_{i,j})_{\substack{1 \leq i \leq n \\ 1 \leq j \leq p}}$ où

$$\forall j \in \llbracket 1, p \rrbracket, e_j = \sum_{i=1}^n a_{i,j} u_i.$$

Exemple 5.21

Dans $\mathbf{R}_2[X]$, posons $P_1 = -X^2 + 3X + 2$, $P_2 = X$ et $P_3 = -X^2 + 3X + 4$.
Alors la matrice de la famille (P_1, P_2, P_3) dans la base canonique de $\mathbf{R}_2[X]$ est :

$$\text{Mat}_{\mathcal{B}_{can}}(P_1, P_2, P_3) = \begin{pmatrix} P_1 & P_2 & P_3 \\ 3 & 0 & 4 \\ 3 & 1 & 3 \\ -1 & 0 & -1 \end{pmatrix} \begin{matrix} 1 \\ X \\ X^2 \end{matrix}.$$

Proposition 5.22 : Si (e_1, \dots, e_p) est une famille de vecteurs d'un espace vectoriel de dimension finie n , et si \mathcal{B} est une base de E , alors le rang de (e_1, \dots, e_p) est égal au rang de la matrice de (e_1, \dots, e_p) dans la base \mathcal{B} .

Rappel

La matrice de (e_1, \dots, e_n) dans la base \mathcal{B} est la matrice dont la $i^{\text{ème}}$ colonne est le vecteur colonne des coordonnées de e_i dans \mathcal{B} .

Démonstration. L'idée est de remarquer que toute relation entre e_1, \dots, e_p se traduit par une relation sur les colonnes de $\text{Mat}_{\mathcal{B}}(e_1, \dots, e_p)$. \square

Exemple 5.23

Si on reprend la famille de l'exemple précédent, la matrice $\text{Mat}_{\mathcal{B}_{can}}(P_1, P_2, P_3)$ est de rang 3, et donc $\text{rg}(P_1, P_2, P_3) = 3$.

Par conséquent, (P_1, P_2, P_3) est une famille libre.

5.2.2 Matrice d'une application linéaire

Définition 5.24 – Soient E et F deux espaces vectoriels de dimension finie, $\mathcal{B} = (e_1, \dots, e_p)$ une base de E et $\mathcal{B}' = (u_1, \dots, u_n)$ une base de F . Alors on appelle matrice de $f \in \mathcal{L}(E, F)$ dans les bases \mathcal{B} et \mathcal{B}' et on note $\text{Mat}_{\mathcal{B}', \mathcal{B}}(f)$ la matrice de la famille $f((e_1), \dots, f(e_p))$ dans la base \mathcal{B}' .

Autrement dit, on a $\text{Mat}_{\mathcal{B}', \mathcal{B}}(f) = (a_{i,j})_{\substack{1 \leq i \leq n \\ 1 \leq j \leq p}}$ où

$$\forall j \in \llbracket 1, p \rrbracket, f(e_j) = \sum_{i=1}^n a_{i,j} u_i.$$

Soit encore

$$\text{Mat}_{\mathcal{B}', \mathcal{B}}(f) = \begin{pmatrix} f(e_1) & f(e_2) & \dots & f(e_p) \\ a_{1,1} & a_{1,2} & \dots & a_{1,p} \\ a_{2,1} & a_{2,2} & \dots & a_{2,p} \\ \vdots & \vdots & & \vdots \\ a_{n,1} & a_{n,2} & \dots & a_{n,p} \end{pmatrix} \begin{matrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{matrix}.$$

Dimensions

Attention à la taille de la matrice : elle a autant de lignes que la dimension de l'espace d'arrivée F et autant de colonne que la dimension de l'espace de départ E .

Proposition 5.25 : Soit $f \in \mathcal{L}(E, F)$, \mathcal{B} une base de E et \mathcal{B}' une base de F . Soit $x \in E$,

et soit $y = f(x)$. Si l'on note $x_{\mathcal{B}} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ le vecteur des coordonnées de x dans la base \mathcal{B} et

$y_{\mathcal{B}'} = \begin{pmatrix} y_1 \\ \vdots \\ y_p \end{pmatrix}$ le vecteur colonne des coordonnées de y dans la base \mathcal{B}' alors

$$y_{\mathcal{B}'} = \text{Mat}_{\mathcal{B}', \mathcal{B}}(f) x_{\mathcal{B}}.$$

Ainsi, la matrice de f permet d'obtenir les coordonnées de $f(x)$ (dans la base \mathcal{B}') à partir de celles de x (dans la base \mathcal{B}).

Proposition 5.26 : Soient E, F, G trois espaces vectoriels de dimension finie, de bases respectives $\mathcal{B}, \mathcal{B}'$ et \mathcal{B}'' . Alors :

- si $f, g \in \mathcal{L}(E, F)$, $\lambda \in \mathbf{K}$, alors

$$\text{Mat}_{\mathcal{B}', \mathcal{B}}(\lambda f + g) = \lambda \text{Mat}_{\mathcal{B}', \mathcal{B}}(f) + \text{Mat}_{\mathcal{B}', \mathcal{B}}(g).$$

- si $f \in \mathcal{L}(E, F)$, $g \in \mathcal{L}(F, G)$, alors

$$\text{Mat}_{\mathcal{B}'', \mathcal{B}}(g \circ f) = \text{Mat}_{\mathcal{B}'', \mathcal{B}'}(g) \times \text{Mat}_{\mathcal{B}', \mathcal{B}}(f).$$

Autrement dit

L'application qui à $f \in \mathcal{L}(E, F)$ associe sa matrice dans les bases \mathcal{B} et \mathcal{B}' est linéaire.

Remarques. • Ce sont ces formules qui justifient que le produit de matrices est défini de manière «si compliquée» : on voulait que le produit de matrices corresponde à la composition des applications linéaires. On aurait tout aussi bien pu définir le produit de deux matrices en faisant le produit des coefficients de chaque matrice, ce serait bien plus simple. Mais alors on ne disposerait plus de l'interprétation en termes d'applications linéaires.

• Une fois fixées deux bases \mathcal{B} et \mathcal{B}' de E et F , on a un isomorphisme de $\mathcal{L}(E, F)$ dans $\mathcal{M}_{n,p}(\mathbf{K})$, défini par $f \mapsto \text{Mat}_{\mathcal{B}', \mathcal{B}}(f)$.

En particulier, on a $\dim \mathcal{L}(E, F) = \dim \mathcal{M}_{n,p}(\mathbf{K}) = n \times p = \dim E \times \dim F$.

Autrement dit

Une fois des bases de E et de F fixées, on a une correspondance entre matrices et applications linéaires.

Corollaire 5.27 – $f \in \mathcal{L}(E, F)$ est un isomorphisme si et seulement si $\text{Mat}_{\mathcal{B}', \mathcal{B}}(f)$ est inversible, et dans ce cas

$$\text{Mat}_{\mathcal{B}, \mathcal{B}'}(f^{-1}) = (\text{Mat}_{\mathcal{B}', \mathcal{B}}(f))^{-1}.$$

Précision

Si $\dim E \neq \dim F$, alors la matrice de f ne peut être inversible car elle n'est pas carrée. Ceci est cohérent avec le fait que s'il existe un isomorphisme $E \rightarrow F$, alors nécessairement $\dim E = \dim F$.

Démonstration. Si $f \in \mathcal{L}(E, F)$ est un isomorphisme, alors

$$\text{Mat}_{\mathcal{B}, \mathcal{B}'}(f^{-1}) \times \text{Mat}_{\mathcal{B}', \mathcal{B}}(f) = \text{Mat}_{\mathcal{B}, \mathcal{B}}(\text{id}_E) = I_n$$

Donc $\text{Mat}_{\mathcal{B}', \mathcal{B}}(f)$ est inversible, et son inverse est $\text{Mat}_{\mathcal{B}, \mathcal{B}'}(f^{-1})$.

Inversement, si $\text{Mat}_{\mathcal{B}', \mathcal{B}}(f)$ est inversible, soit $g \in \mathcal{L}(F, E)$ tel que $\text{Mat}_{\mathcal{B}, \mathcal{B}'}(g) = (\text{Mat}_{\mathcal{B}', \mathcal{B}}(f))^{-1}$.
Alors

$$\text{Mat}_{\mathcal{B}, \mathcal{B}}(g \circ f) = I_n = \text{Mat}_{\mathcal{B}, \mathcal{B}}(\text{id}_E).$$

Donc $g \circ f = \text{id}_E$, et donc $g = f^{-1}$. □

Cas des endomorphismes

Définition 5.28 – Si E est un \mathbf{K} -espace vectoriel de dimension finie n , si \mathcal{B} est une base de E et si $f \in \mathcal{L}(E)$, alors on note $\text{Mat}_{\mathcal{B}}(f) = \text{Mat}_{\mathcal{B}, \mathcal{B}}(f)$.

On a alors, $\forall k \in \mathbf{N}, \text{Mat}_{\mathcal{B}}(f^k) = (\text{Mat}_{\mathcal{B}}(f))^k$.
Plus généralement, si $P \in \mathbf{K}[X], \text{Mat}_{\mathcal{B}}(P(f)) = P(\text{Mat}_{\mathcal{B}}(f))$.

5.2.3 Retour sur le rang d'une matrice

Proposition 5.29 : Soit $f \in \mathcal{L}(E, F), \mathcal{B}_E$ une base de E , et \mathcal{B}_F une base de F .
Alors le rang de la matrice $\text{Mat}_{\mathcal{B}_F, \mathcal{B}_E}(f)$ est égal au rang de l'application linéaire f .

Démonstration. On a

$$\begin{aligned} \text{rg}(f) &= \dim \text{Im}(f) = \dim \text{Vect}(f(e_1), \dots, f(e_p)) \\ &= \text{rg}(f(e_1), \dots, f(e_p)) = \text{rg} \text{Mat}_{\mathcal{B}_F}(f(e_1), \dots, f(e_p)) = \text{rg} \text{Mat}_{\mathcal{B}_F, \mathcal{B}_E}(f). \end{aligned}$$

□

Détails

L'image par f d'une base de E est toujours une famille génératrice de $\text{Im}(f)$.

Proposition 5.30 : Soit $A \in \mathcal{M}_n(\mathbf{K})$. Alors il y a équivalence entre :

1. A est inversible
2. $\text{rg}(A) = n$
3. pour tout $X \in \mathcal{M}_{n,1}(\mathbf{K}), AX = 0 \Rightarrow X = 0$.

Démonstration. Notons f l'unique endomorphisme de \mathbf{K}^n dont la matrice dans la base canonique est A .

On a alors

$$\text{rg}(A) = n \Leftrightarrow \text{rg}(f) = n \Leftrightarrow f \text{ est bijective} \Leftrightarrow A \text{ est inversible.}$$

Et donc 1) \Leftrightarrow 2).

Supposons que 2) soit vérifié, et soit $X \in \mathcal{M}_{n,1}(\mathbf{K})$ tel que $AX = 0$.

Soit x l'unique vecteur de \mathbf{K}^n tel que $X = (x)_{\mathcal{B}_{can}}$. Alors $AX = 0 \Leftrightarrow f(x) = 0$.

Mais f est bijective, et donc $\text{Ker}(f) = \{0\}$. On en déduit que $x = 0$ et donc $X = 0$.

Ainsi, 2) \Rightarrow 3).

Inversement, supposons que 3) soit vérifiée. Si $x \in \text{Ker } f$, soit $X = (x)_{\mathcal{B}_{can}} \in \mathcal{M}_{n,1}(\mathbf{K})$.

Alors $f(x) = 0 \Leftrightarrow AX = 0$, et donc $X = 0 \Leftrightarrow x = 0$.

Ainsi, $\text{Ker}(f) = \{0\}$, et donc f est injective, donc bijective : A est inversible. Ainsi, 3) \Rightarrow 1). □

5.2.4 Matrices de passage

Soit $f \in \mathcal{L}(E)$. Si \mathcal{B} et \mathcal{B}' sont deux bases de E , alors $\text{Mat}_{\mathcal{B}}(f)$ et $\text{Mat}_{\mathcal{B}'}(f)$ sont a priori deux matrices différentes.

Nous expliquons dans cette partie comment passer de l'une à l'autre.

Proposition 5.31 : Soit E un espace vectoriel de dimension finie n , et soit \mathcal{B} une base de E . Alors une famille (e_1, \dots, e_n) de n vecteurs de E est une base de E si et seulement si $\text{Mat}_{\mathcal{B}}(e_1, \dots, e_n)$ est inversible.

Remarque

Bien évidemment, cette proposition ne s'applique que pour une famille de cardinal n , puisqu'une famille de cardinal différent de n ne peut en aucun cas être une base de E .

Démonstration. Puisque (e_1, \dots, e_n) est de cardinal $n = \dim E$, on a

$$\begin{aligned} (e_1, \dots, e_n) \text{ est une base de } E &\Leftrightarrow (e_1, \dots, e_n) \text{ est libre} \\ &\Leftrightarrow \operatorname{rg}(e_1, \dots, e_n) = n \\ &\Leftrightarrow \operatorname{rg}(\operatorname{Mat}_{\mathcal{B}}(e_1, \dots, e_n)) = n \\ &\Leftrightarrow \operatorname{Mat}_{\mathcal{B}}(e_1, \dots, e_n) \text{ est inversible} \end{aligned}$$

□

Définition 5.32 – Soient $\mathcal{B} = (e_1, \dots, e_n)$ et $\mathcal{B}' = (f_1, \dots, f_n)$ deux bases d'un espace vectoriel E de dimension n . Alors la matrice de passage de \mathcal{B} à \mathcal{B}' , notée $P_{\mathcal{B}, \mathcal{B}'}$ est la matrice de la famille (f_1, \dots, f_n) dans la base \mathcal{B} . Autrement dit, si $P_{\mathcal{B}, \mathcal{B}'} = (a_{i,j})_{1 \leq i, j \leq n} \in \mathcal{M}_n(\mathbf{K})$, alors

$$\forall j \in \llbracket 1, n \rrbracket, f_j = \sum_{i=1}^n a_{i,j} e_i.$$

Remarque. $P_{\mathcal{B}, \mathcal{B}'}$ est la matrice des coordonnées des vecteurs de la nouvelle base (\mathcal{B}') dans l'ancienne base (\mathcal{B}).

Proposition 5.33 : Soient \mathcal{B} et \mathcal{B}' deux bases d'un espace vectoriel E de dimension finie.

Alors :

- $P_{\mathcal{B}, \mathcal{B}'} = \operatorname{Mat}_{\mathcal{B}, \mathcal{B}'}(\operatorname{id}_E)$
- $P_{\mathcal{B}, \mathcal{B}'}$ est inversible et

$$(P_{\mathcal{B}, \mathcal{B}'})^{-1} = P_{\mathcal{B}', \mathcal{B}}.$$

Méthode

En général, quand on a deux bases différentes de E , on a l'expression des vecteurs de l'une en fonction de ceux de l'autre. On a alors une des deux matrices de passage. Pour obtenir l'autre matrice de passage, il suffit d'inverser celle dont on dispose.

Démonstration. Pour le premier point, remarquons que la $j^{\text{ème}}$ colonne de $\operatorname{Mat}_{\mathcal{B}, \mathcal{B}'}$ est formée des coordonnées de $f_j = \operatorname{id}_E(f_j)$ dans la base \mathcal{B} .

Donc $P_{\mathcal{B}, \mathcal{B}'} = \operatorname{Mat}_{\mathcal{B}, \mathcal{B}'}(\operatorname{id}_E)$.

Pour le second point, nous savons que l'identité est un isomorphisme, égal à son propre inverse. Donc $P_{\mathcal{B}, \mathcal{B}'}$ est inversible, et

$$P_{\mathcal{B}, \mathcal{B}'}^{-1} = (\operatorname{Mat}_{\mathcal{B}, \mathcal{B}'}(\operatorname{id}_E))^{-1} = \operatorname{Mat}_{\mathcal{B}', \mathcal{B}}(\operatorname{id}_E^{-1}) = \operatorname{Mat}_{\mathcal{B}', \mathcal{B}}(\operatorname{id}_E) = P_{\mathcal{B}', \mathcal{B}}.$$

□

Exemple 5.34 Retour sur l'exemple 5.21

Reprenons $\mathcal{B} = (P_1, P_2, P_3) = (-X^2 + 3X + 3, X, -X^2 + 3X + 4)$.

$$\text{Alors } P_{\mathcal{B}_{can}, \mathcal{B}} = \begin{pmatrix} P_1 & P_2 & P_3 \\ 3 & 0 & 4 \\ 3 & 1 & 3 \\ -1 & 0 & -1 \end{pmatrix} \begin{matrix} 1 \\ X \\ X^2 \end{matrix}.$$

Et donc

$$P_{\mathcal{B}, \mathcal{B}_{can}} = \begin{pmatrix} 3 & 0 & 4 \\ 3 & 1 & 3 \\ -1 & 0 & -1 \end{pmatrix}^{-1} = \begin{pmatrix} 1 & X & X^2 \\ -1 & 0 & -4 \\ 0 & 1 & 3 \\ 1 & 0 & 3 \end{pmatrix} \begin{matrix} P_1 \\ P_2 \\ P_3 \end{matrix}.$$

5.2.5 Formules de changement de base

Proposition 5.35 : Soit E un \mathbf{K} -espace vectoriel de dimension finie n , \mathcal{B} et \mathcal{B}' deux bases de E . Soit $x \in E$. Alors :

$$x_{\mathcal{B}} = P_{\mathcal{B}, \mathcal{B}'} x_{\mathcal{B}'}$$

Démonstration. Notons $\mathcal{B} = (e_1, \dots, e_n)$ et $\mathcal{B}' = (f_1, \dots, f_n)$.

Notons également $x_{\mathcal{B}} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$, $x_{\mathcal{B}'} = \begin{pmatrix} x'_1 \\ \vdots \\ x'_n \end{pmatrix}$ de sorte que

$$x = \sum_{i=1}^n x_i e_i = \sum_{i=1}^n x'_i f_i$$

Enfin, notons $P_{\mathcal{B}, \mathcal{B}'} = (a_{i,j})_{1 \leq i, j \leq n}$ de telle sorte que $\forall j \in \llbracket 1, n \rrbracket$, $f_j = \sum_{i=1}^n a_{i,j} e_i$. Alors

$$x = \sum_{j=1}^n x'_j f_j = \sum_{j=1}^n x'_j \sum_{i=1}^n a_{i,j} e_i = \sum_{i=1}^n \left(\sum_{j=1}^n a_{i,j} x'_j \right) e_i$$

Ainsi², $x_i = \sum_{j=1}^n a_{i,j} x'_j$.

On reconnaît là l'expression du coefficient de la $i^{\text{ème}}$ ligne du produit $P_{\mathcal{B}, \mathcal{B}'} \begin{pmatrix} x'_1 \\ \vdots \\ x'_n \end{pmatrix}$. □

Autrement dit, la matrice de passage de \mathcal{B} à \mathcal{B}' permet d'obtenir les coordonnées d'un vecteur dans la base \mathcal{B} à partir de ses coordonnées dans la base \mathcal{B}' .

Théorème 5.36 (Formule de changement de base) : Soit E un \mathbf{K} -espace vectoriel de dimension finie n , \mathcal{B} et \mathcal{B}' deux bases de E . Soit $f \in \mathcal{L}(E)$. Alors

$$\begin{aligned} \text{Mat}_{\mathcal{B}'}(f) &= P_{\mathcal{B}', \mathcal{B}} \text{Mat}_{\mathcal{B}}(f) P_{\mathcal{B}, \mathcal{B}'} \\ &= P_{\mathcal{B}, \mathcal{B}'}^{-1} \text{Mat}_{\mathcal{B}}(f) P_{\mathcal{B}, \mathcal{B}'} \end{aligned}$$

Démonstration. Nous savons que $f = \text{id}_E \circ f \circ \text{id}_E$, et donc

$$\text{Mat}_{\mathcal{B}'}(f) = \text{Mat}_{\mathcal{B}', \mathcal{B}}(\text{id}_E) \times \text{Mat}_{\mathcal{B}, \mathcal{B}}(f) \times \text{Mat}_{\mathcal{B}, \mathcal{B}'}(\text{id}_E) = P_{\mathcal{B}', \mathcal{B}} \text{Mat}_{\mathcal{B}}(f) P_{\mathcal{B}, \mathcal{B}'}$$

□

Exemple 5.37

Soit f l'endomorphisme de $\mathbf{R}_2[X]$ défini par $f(P) = P + P'$.

$$\text{Alors } \text{Mat}_{\mathcal{B}_{can}}(f) = \begin{pmatrix} f(1) & f(X) & f(X^2) \\ 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 2 \end{pmatrix} \begin{matrix} 1 \\ X \\ X^2 \end{matrix}$$

Mais alors,

$$\text{Mat}_{\mathcal{B}}(f) = P_{\mathcal{B}, \mathcal{B}_{can}} \times \text{Mat}_{\mathcal{B}_{can}}(f) \times P_{\mathcal{B}_{can}, \mathcal{B}}$$

Danger !

La terminologie est trompeuse : la matrice de passage de \mathcal{B} à \mathcal{B}' permet d'obtenir les coordonnées dans \mathcal{B} à partir de celles dans \mathcal{B}' , et non l'inverse (comme le nom le laisserait penser).

Plus simplement

$P_{\mathcal{B}, \mathcal{B}'} = \text{Mat}_{\mathcal{B}, \mathcal{B}'}(\text{id})$.
Et alors ce résultat est une application directe de la proposition 5.25.

² Par unicité de la décomposition de x dans la base \mathcal{B}' .

Méthode

Cette formule permet, à partir de la matrice de f dans une base \mathcal{B} , et de la connaissance de la matrice de passage de \mathcal{B} à $\mathcal{B}' = (e_1, \dots, e_n)$, d'obtenir la matrice de f dans la base \mathcal{B}' sans avoir besoin de décomposer les $f(e_i)$ dans \mathcal{B}' .

$$\begin{aligned}
&= \begin{pmatrix} -1 & 0 & -4 \\ 0 & 1 & 3 \\ 1 & 0 & 3 \end{pmatrix} \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 3 & 0 & 4 \\ 3 & 1 & 3 \\ -1 & 0 & -1 \end{pmatrix} \\
&= \begin{pmatrix} f(P_1) & f(P_2) & f(P_3) \\ -2 & -1 & -3 \\ -2 & 1 & -2 \\ 3 & 1 & 4 \end{pmatrix} \begin{matrix} P_1 \\ P_2 \\ P_3 \end{matrix}.
\end{aligned}$$

Notons que sans la formule de changement de base, nous aurions pu obtenir cette matrice en calculant $f(P_1)$, $f(P_2)$ et $f(P_3)$, puis en déterminant leurs coordonnées dans la base (P_1, P_2, P_3) en résolvant trois systèmes.

5.2.6 Matrices semblables

Définition 5.38 – Deux matrices A et B de $\mathcal{M}_n(\mathbf{K})$ sont dites semblables s'il existe $P \in GL_n(\mathbf{K})$ telle que

$$B = P^{-1}AP.$$

En particulier, si \mathcal{B} et \mathcal{B}' sont deux bases d'un espace vectoriel E , et si $f \in \mathcal{L}(E)$, alors les matrices de f dans les bases \mathcal{B} et \mathcal{B}' sont semblables.

En effet, par la formule de changement de base, on a

$$\text{Mat}_{\mathcal{B}'} = P_{\mathcal{B}, \mathcal{B}'}^{-1} \text{Mat}_{\mathcal{B}}(f) P_{\mathcal{B}, \mathcal{B}'}$$

Remarques. • A est semblable à I_n si et seulement si

$$\exists P \in GL_n(\mathbf{K}) : A = P^{-1}I_n P = I_n.$$

Donc I_n est la seule matrice semblable à I_n .

- De même, la seule matrice semblable à la matrice nulle est la matrice nulle.

Proposition 5.39 : Si $A, B \in \mathcal{M}_n(\mathbf{K})$ sont semblables, alors $\text{rg}(A) = \text{rg}(B)$.

Démonstration. Supposons donc qu'il existe $P \in GL_n(\mathbf{K})$ telle que $B = P^{-1}AP$.

Soit $f \in \mathcal{L}(\mathbf{K}^n)$ l'unique endomorphisme de \mathbf{K}^n dont la matrice dans la base canonique est A .

Notons C_1, \dots, C_n les colonnes de A , et soit e_1, \dots, e_n l'unique famille de vecteurs de \mathbf{K}^n telle que $C_i = (e_i)_{\mathcal{B}_{can}}$.

Alors la matrice de (e_1, \dots, e_n) dans la base canonique est, par définition, la matrice dont les colonnes sont $(e_1)_{\mathcal{B}_{can}}, \dots, (e_n)_{\mathcal{B}_{can}}$, c'est-à-dire C_1, \dots, C_n . Autrement dit, cette matrice est la matrice P , et donc elle de rang n , de sorte (e_1, \dots, e_n) est une base de \mathbf{K}^n . Notons \mathcal{B}' cette base.

La matrice de passage de \mathcal{B}_{can} à \mathcal{B}' est alors la matrice P .

Et donc par la formule de changement de base,

$$B = P^{-1}AP = P_{\mathcal{B}_{can}, \mathcal{B}'}^{-1} \text{Mat}_{\mathcal{B}_{can}}(f) P_{\mathcal{B}_{can}, \mathcal{B}'} = \text{Mat}_{\mathcal{B}'}(f).$$

En particulier, puisque A et B représentent le même endomorphisme f dans deux bases différentes, on a $\text{rg}(A) = \text{rg}(f) = \text{rg}(B)$. \square

\triangleleft La réciproque est fautive. Par exemple si $A = I_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ et $B = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$, alors A et B ont le même rang, mais ne sont pas semblables car la seule matrice semblable à I_2 est I_2 .

Proposition 5.40 : Si A et $B \in \mathcal{M}_n(\mathbf{K})$ sont semblables, et si A est inversible, alors B est inversible et A^{-1} et B^{-1} sont semblables.

Plus généralement

Une matrice semblable à λI_n est nécessairement égale à λI_n .

Méthode

On a prouvé que si A et B étaient semblables, alors elles représentent le même endomorphisme de \mathbf{K}^n dans deux bases différentes. Autrement dit, deux matrices sont semblables si et seulement si elles représentent le même endomorphisme dans deux bases différentes. En particulier, pour montrer que deux matrices A et B sont semblables, on essaiera de trouver un endomorphisme f de \mathbf{K}^n (ou de tout autre espace de dimension n) et deux bases de \mathbf{K}^n dans lesquelles les matrices de f sont respectivement A et B .

Démonstration. Si A est inversible, alors elle est de rang n , et donc B est également de rang n , donc inversible.

De plus soit P inversible telle que $B = P^{-1}AP$. Alors

$$B^{-1} = (P^{-1}AP)^{-1} = P^{-1}A^{-1}(P^{-1})^{-1} = P^{-1}A^{-1}P$$

et donc A^{-1} et B^{-1} sont semblables. □

Proposition 5.41 : Si A et B sont semblables, alors pour tout $P \in \mathbf{K}[X]$, $P(A)$ et $P(B)$ sont semblables.

Démonstration. Puisque A et B sont semblables, il existe $Q \in GL_n(\mathbf{K})$ telle que $B = Q^{-1}AQ$. Commençons par remarquer que $\forall k \in \mathbf{N}, B^k = Q^{-1}A^kQ$. En effet

$$B^k = (Q^{-1}AQ)(Q^{-1}AQ) \cdots (Q^{-1}AQ) = Q^{-1}A(QQ^{-1})AQ \cdots (QQ^{-1})AQ = Q^{-1}A^kQ.$$

Alors soit $P = \sum_{k=0}^d a_k X^k \in \mathbf{K}[X]$. On a

$$P(B) = \sum_{k=0}^d a_k B^k = \sum_{k=0}^d a_k Q^{-1}A^kQ = Q^{-1} \left(\sum_{k=0}^d a_k A^k \right) Q = Q^{-1}P(A)Q.$$

□

Ceci prouve en fait que si A et B sont semblables, alors A^k et B^k sont semblables, pour tout $k \in \mathbf{N}$.

5.2.7 Retour sur les projecteurs

Soit E un \mathbf{K} -espace vectoriel de dimension n .

Proposition 5.42 : Soit $f \in \mathcal{L}(E)$, $\mathcal{B} = (e_1, \dots, e_n)$ une base de E . Alors les assertions suivantes sont équivalentes :

1. f est un projecteur de rang r ,
2. il existe une base \mathcal{B}' de E telle que,

$$\text{Mat}_{\mathcal{B}'}(f) = \begin{pmatrix} 1 & & 0 & \dots & 0 \\ & \ddots & \vdots & & \vdots \\ & & 1 & 0 & \dots & 0 \\ 0 & \dots & 0 & 0 & \dots & 0 \\ \vdots & & \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & 0 & 0 & \dots & 0 \end{pmatrix} = \text{Diag}(\underbrace{1, \dots, 1}_{r \text{ fois}}, \underbrace{0, \dots, 0}_{n-r \text{ fois}}) = A_r$$

3. $\text{Mat}_{\mathcal{B}}(f)$ est semblable à A_r .

⚠ Attention !
Nous avons bien dit qu'il existe une telle base. Ne pas en déduire que dans toute base de E , la matrice d'un projecteur est diagonale avec des 0 et des 1 sur la diagonale !

Démonstration. 1) \Rightarrow 2) : supposons que f soit un projecteur.

Nous savons que $E = \text{Im} f \oplus \text{Ker} f$, et si $x = x_1 + x_2$ avec $x_1 \in \text{Im} f$ et $x_2 \in \text{Ker} f$, alors $f(x) = x_1$.

Soit (e_1, \dots, e_r) une base de $\text{Im} f$, et soit (e_{r+1}, \dots, e_n) une base de $\text{Ker} f$.

Alors $\mathcal{B}' = (e_1, \dots, e_r, e_{r+1}, \dots, e_n)$ est une base de E , adaptée à la somme directe $E = \text{Im} f \oplus \text{Ker} f$, et

$$\forall i \in \llbracket 1, r \rrbracket, f(e_i) = e_i, \forall i \in \llbracket r+1, n \rrbracket, f(e_i) = 0$$

Donc la matrice de f dans la base \mathcal{B}' est :

$$\text{Mat}_{\mathcal{B}'}(f) = \begin{pmatrix} f(e_1) & \dots & f(e_r) & f(e_{r+1}) & \dots & f(e_n) \\ 1 & & & & & 0 \\ & \ddots & & & & \vdots \\ & & 1 & & & \vdots \\ & & & 0 & & \vdots \\ & & & & \ddots & \vdots \\ & & & & & 0 \end{pmatrix} = A_r.$$

Base adaptée
La preuve est constructive : une base \mathcal{B}' dans laquelle la matrice de f est de la forme souhaitée est une base obtenue par concaténation d'une base de $\text{Im} f$ et d'une base de $\text{Ker} f$.

Remarque
Notons que le nombre de 1 sur la diagonale de A est égal à $r = \dim \text{Im}(f) = \text{rg}(f)$.
Autrement dit,
 $A = \text{Diag}(\underbrace{1, \dots, 1}_{r \text{ fois}}, \underbrace{0, \dots, 0}_{(n-r) \text{ fois}})$

2) ⇒ 3) $\text{Mat}_{\mathcal{B}}(f)$ et $\text{Mat}_{\mathcal{B}'}(f)$ sont semblables d'après la formule de changement de base.

3) ⇒ 1) Notons que $A_r^2 = A_r$. Et donc si $\text{Mat}_{\mathcal{B}}(f)$ est semblable à la matrice A_r , alors il existe $P \in GL_n(\mathbb{K})$ telle que $\text{Mat}_{\mathcal{B}}(f) = P^{-1}A_rP$ et donc

$$\text{Mat}_{\mathcal{B}}(f) = (P^{-1}A_rP)^2 = P^{-1}A_r^2P = \text{Mat}_{\mathcal{B}}(f).$$

On en déduit que $f^2 = f$ et donc que f est un projecteur.

Et alors $\text{rg}(f) = J(\text{Mat}_{\mathcal{B}}(f)) = \text{rg}(A_r) = r$. □

La principale utilité de ce résultat est de déterminer la matrice d'un projecteur dans une base : dans «une bonne base», c'est-à-dire une matrice adaptée à la somme directe $\text{Im } p \oplus \text{Ker } p$, nous connaissons cette matrice. Si on souhaite obtenir la matrice de p dans une autre base, il suffit d'appliquer la formule de changement de base.

5.2.8 Retour sur les sous-espaces stables



Soit E un espace vectoriel de dimension finie, soit $f \in \mathcal{L}(E)$, soit F un sous-espace stable par f et H un supplémentaire de F dans E .

Soit alors $f|_F$ la restriction de f à F , qui va de F dans E .

Puisque F est stable par f , l'image d'un vecteur de F par f (et donc par $f|_F$) est dans F , et on peut ainsi voir $f|_F$ comme une application de F dans F : on l'appelle **l'endomorphisme induit par f sur F** .

En particulier, soit (e_1, \dots, e_r) une base de F , et (e_{r+1}, \dots, e_n) une base de G , de sorte que $\mathcal{B} = (e_1, \dots, e_n)$ est une base de E adaptée à la somme directe $F \oplus G$.

Alors, pour tout $j \in \llbracket 1, r \rrbracket$, $f(e_j) \in F$ et donc $f(e_j)$ s'écrit de manière unique $f(e_j) = \sum_{i=1}^r a_{i,j}e_i$.

Et donc la matrice de f dans la base \mathcal{B} est

$$\text{Mat}_{\mathcal{B}}(f) = \begin{pmatrix} f(e_1) & \dots & f(e_r) & f(e_{r+1}) & \dots & f(e_n) \\ a_{1,1} & \dots & a_{1,r} & \star & \dots & \star \\ \vdots & & \vdots & \vdots & & \vdots \\ a_{r,1} & \dots & a_{r,r} & \star & \dots & \star \\ 0 & \dots & 0 & \star & \dots & \star \\ \vdots & & \vdots & \vdots & & \vdots \\ 0 & \dots & 0 & \star & \dots & \star \end{pmatrix} \begin{matrix} e_1 \\ \vdots \\ e_r \\ e_{r+1} \\ \vdots \\ e_n \end{matrix}.$$

En particulier, la matrice $(a_{i,j})_{1 \leq i,j \leq r}$ est la matrice de $f|_F$ dans la base (e_1, \dots, e_r) .

Inversement, s'il existe une base $(e_1, \dots, e_r, e_{r+1}, \dots, e_n)$ dans laquelle la matrice de f est de cette forme³, alors $F = \text{Vect}(e_1, \dots, e_r)$ est stable par f .

Si de plus G est également stable par f , alors on peut tenir le même raisonnement avec les $f(e_j), r + 1 \leq j \leq n$, qui sont tous dans G , et montrer que la matrice de f est de la forme

$$\text{Mat}_{\mathcal{B}}(f) = \begin{pmatrix} f(e_1) & \dots & f(e_r) & f(e_{r+1}) & \dots & f(e_n) \\ a_{1,1} & \dots & a_{1,p} & 0 & \dots & 0 \\ \vdots & & \vdots & \vdots & & \vdots \\ a_{p,1} & \dots & a_{p,p} & 0 & \dots & 0 \\ 0 & \dots & 0 & \star & \dots & \star \\ \vdots & & \vdots & \vdots & & \vdots \\ 0 & \dots & 0 & \star & \dots & \star \end{pmatrix} \begin{matrix} e_1 \\ \vdots \\ e_r \\ e_{r+1} \\ \vdots \\ e_n \end{matrix}.$$

Restriction
Rappelons que l'application $f|_F$ est l'application de F dans E définie par $x \mapsto f(x)$. C'est donc f , mais son ensemble de départ est F et non E .

Terminologie
Une telle matrice est dite triangulaire par blocs.

³ C'est-à-dire avec $a_{i,j} = 0$ pour $j \leq r$ et $i > r$.

Terminologie
Une telle matrice est dite diagonale par blocs.

5.3 TRACE D'UNE MATRICE CARRÉE

5.3.1 Définition, linéarité

Définition 5.43 – Soit $A = (a_{i,j}) \in \mathcal{M}_n(\mathbf{K})$. Alors on appelle trace de A , et on note

$$\text{tr}(A) = \sum_{i=1}^n a_{i,i} \in \mathbf{K}. \text{ C'est la somme des coefficients diagonaux de } A.$$

Exemples 5.44

On a

- $\text{tr}(0) = 0$ et $\text{tr}(I_n) = n$.
- Plus généralement, $\text{tr}(\text{diag}(\lambda_1, \dots, \lambda_n)) = \lambda_1 + \dots + \lambda_n$.
- Si A est antisymétrique, alors⁴ $\text{tr}(A) = 0$.

⁴ En effet, tous les coefficients diagonaux de A sont alors nuls.

Proposition 5.45 : L'application $\text{tr} : \mathcal{M}_n(\mathbf{K}) \rightarrow \mathbf{K}$ est une forme linéaire.

Démonstration. Puisque tr est à valeurs dans \mathbf{K} , il s'agit de prouver que tr est linéaire. Soient donc $A = (a_{i,j}), B = (b_{i,j}) \in \mathcal{M}_n(\mathbf{K})$, et soit $\lambda \in \mathbf{K}$. Alors

$$\text{tr}(\lambda A + B) = \sum_{i=1}^n (\lambda A + B)_{i,i} = \sum_{i=1}^n (\lambda a_{i,i} + b_{i,i}) = \lambda \sum_{i=1}^n a_{i,i} + \sum_{i=1}^n b_{i,i} = \lambda \text{tr}(A) + \text{tr}(B).$$

□

Proposition 5.46 : Si $A \in \mathcal{M}_n(\mathbf{K})$, alors $\text{tr}({}^t A) = \text{tr}(A)$.

Démonstration. C'est évident car les coefficients diagonaux de ${}^t A$ sont les mêmes que ceux de A . □

5.3.2 Invariance de la trace par similitude

Proposition 5.47 : Soient $A, B \in \mathcal{M}_n(\mathbf{K})$. Alors $\text{tr}(AB) = \text{tr}(BA)$.

Démonstration. Par définition,

$$\text{tr}(AB) = \sum_{i=1}^n (AB)_{i,i} = \sum_{i=1}^n \sum_{k=1}^n a_{i,k} b_{k,i} = \sum_{k=1}^n \sum_{i=1}^n b_{k,i} a_{i,k} = \sum_{k=1}^n (BA)_{k,k}$$

□

Remarque. Ce résultat est en fait encore valable pour $A \in \mathcal{M}_{n,p}(\mathbf{K})$ et $B \in \mathcal{M}_{p,n}(\mathbf{K})$, c'est-à-dire dès que les deux produits AB et BA sont tous deux définis et carrés.

Corollaire 5.48 – Deux matrices semblables ont même trace.

Démonstration. Supposons que A et B soient semblables. Alors soit $P \in GL_n(\mathbf{K})$ telle que $B = P^{-1}AP$. Alors

$$\text{tr}(B) = \text{tr}(P^{-1}AP) = \text{tr}(P^{-1}(AP)) = \text{tr}(APP^{-1}) = \text{tr}(A).$$

□

⚠ Danger !

Pour autant, on n'a pas $\text{tr}(AB) = \text{tr}(A) \times \text{tr}(B)$.
Par exemple,

$$\text{tr}(I_n^2) = n \neq \text{tr}(I_n)^2 = n^2.$$

⚠ Réciproque fausse

Deux matrices peuvent en revanche avoir la même trace et ne pas être semblables. Il suffit de reprendre les matrices de la remarque suivant la proposition 5.39.

Exemple 5.49

Soit $A \in \mathcal{M}_n(\mathbf{K})$ telle que $A^2 = A$. Alors A est la matrice dans la base canonique de \mathbf{K}^n d'un projecteur f de \mathbf{K}^n . Par l'étude des projecteurs réalisée précédemment, il existe une base dans laquelle la matrice de f est diagonale, avec $\text{rg}(f) = \text{rg}(A)$ coefficients diagonaux égaux à 1 et les autres à zéro.

Mais alors A est semblable à cette matrice, dont la trace vaut $\text{rg}(A)$.

On en déduit que $\text{tr}(A) = \text{rg}(A)$.

5.4 HORS PROGRAMME : SUITES DE MATRICES

Les résultats qui suivent sont hors programme, et ne sont donc pas à connaître. Toutefois, les sujets introduisant des suites de matrices ne sont pas rares⁵.

Les résultats qui suivent sont alors nécessairement fournis par l'énoncé, mais mieux vaut les avoir déjà rencontré avant afin de bien les comprendre.

Définition 5.50 – Une suite de $(A(n))_{n \in \mathbf{N}}$ de matrices de $\mathcal{M}_{p,q}(\mathbf{K})$ est la donnée, pour tout entier n , d'une matrice $A(n)$ de $\mathcal{M}_{p,q}(\mathbf{R})$.

Pour tout $(i, j) \in \llbracket 1, p \rrbracket \times \llbracket 1, q \rrbracket$, on note alors $a_{i,j}(n)$ le coefficient à la $i^{\text{ème}}$ et $j^{\text{ème}}$ colonne de $A(n)$.

⁵ Les sujets de parisiennes, mais pas seulement : on trouve par exemple des suites de matrices dans Lyon 2010 et dans EDHEC 2016

Notation

On pourrait noter A_n au lieu de $A(n)$, mais alors il serait compliqué de noter le coefficient (i, j) de A_n , il faudrait un triple indice.

Exemple 5.51

$$A(n) = \begin{pmatrix} 1 + \frac{(-1)^n}{2^n} & 3 + \frac{1}{n} \\ 0 & 1 - \frac{1}{4^n} \end{pmatrix}.$$

Définition 5.52 – On dit qu'une suite $(A(n))_{n \in \mathbf{N}}$ de matrices de $\mathcal{M}_{p,q}(\mathbf{R})$ converge vers $A = (a_{i,j}) \in \mathcal{M}_{p,q}(\mathbf{R})$ si pour tout $(i, j) \in \llbracket 1, p \rrbracket \times \llbracket 1, q \rrbracket$, $\lim_{n \rightarrow +\infty} a_{i,j}(n) = a_{i,j}$.

Exemple 5.53

En reprenant l'exemple précédent, la suite $(A(n))_{n \in \mathbf{N}}$ converge vers la matrice

$$\begin{pmatrix} 1 & 3 \\ 0 & 1 \end{pmatrix}.$$

Proposition 5.54 :

- Soit $(A(n))_{n \in \mathbf{N}}$ une suite de matrices de $\mathcal{M}_{p,q}(\mathbf{R})$ qui converge vers $A \in \mathcal{M}_{p,q}(\mathbf{R})$. Alors pour tout $\lambda \in \mathbf{R}$, la suite $(\lambda A(n))_{n \in \mathbf{N}}$ converge vers λA .
- Soient $(A(n))_{n \in \mathbf{N}}$ et $(B(n))_{n \in \mathbf{N}}$ deux suites de matrices de $\mathcal{M}_{p,q}(\mathbf{R})$, qui convergent respectivement vers A et B . Alors $(A(n) + B(n))_{n \in \mathbf{N}}$ converge vers $A + B$.
- Si $(A(n))_{n \in \mathbf{N}}$ est une suite de matrices de $\mathcal{M}_{p,q}(\mathbf{R})$ qui converge vers A et soit $(B(n))_{n \in \mathbf{N}}$ une suite de matrices de $\mathcal{M}_{q,r}(\mathbf{R})$ qui converge vers B . Alors la suite $(A(n)B(n))_{n \in \mathbf{N}}$ converge vers $AB \in \mathcal{M}_{p,r}(\mathbf{R})$.

Démonstration. Prouvons par exemple le dernier point.

Pour tout $(i, j) \in \llbracket 1, p \rrbracket \times \llbracket 1, r \rrbracket$, le coefficient (i, j) de $A(n)B(n)$ est

$$\sum_{k=1}^q a_{i,k}(n)b_{k,j}(n) \xrightarrow{n \rightarrow +\infty} \sum_{k=1}^q a_{i,k}b_{k,j}.$$

On reconnaît là le coefficient (i, j) du produit AB . □

Exemple 5.55

$$\text{Soit } X(n) = \begin{pmatrix} e^{-n} \\ 1 + \frac{\sin n}{n} \end{pmatrix}.$$

Alors la suite $(X(n))_{n \in \mathbf{N}}$ converge vers $X = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ et donc $(A(n)X(n))_{n \in \mathbf{N}}$ converge

$$\text{vers } AX = \begin{pmatrix} 1 & 3 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

MÉTHODES DU CHAPITRE 5

✓ MÉTHODE 5.1 : Déterminer le rang d'une matrice M .

On utilise la méthode du pivot en effectuant des opérations sur les lignes jusqu'à obtenir une matrice bien échelonnée. Le rang de M est alors le rang de la matrice échelonnée ainsi obtenue.

- Si toutes les colonnes de M sont proportionnelles (et en particulier si elles sont égales) et que $M \neq 0$, alors $\text{rg}(M) = 1$.
- Pour une matrice de taille n quelconque, se souvenir que par définition, le rang de la matrice est le rang de la famille des vecteurs colonnes.

Exercices de TD : 5.2, 5.3

✓ MÉTHODE 5.2 : Déterminer si une matrice M est inversible, et si c'est le cas, calculer son inverse M^{-1} .

Seule une matrice carrée peut être inversible !

Si M est une matrice carrée de taille n , elle est inversible si et seulement si elle est de rang n .

Avant même de savoir si la matrice est inversible, on pourra procéder comme si l'on cherchait son inverse, en ajoutant

un second membre (qui sera au choix $\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ ou I_n). En effet, le calcul de l'inverse nécessite à un moment d'échelonner la

matrice, ce qui permet de déterminer son rang. Et dans le cas où elle est inversible, une fois la matrice échelonnée, on a déjà effectué la moitié du calcul de l'inverse, évitant ainsi de refaire deux fois les mêmes opérations.

Enfin, notons que si M est la matrice d'une application linéaire f , M est inversible si et seulement si f est bijectif (et M^{-1} est la matrice de f^{-1}).

Exercices de TD : 5.2, 5.7

✓ MÉTHODE 5.3 : Déterminer la matrice d'un projecteur $p \in \mathcal{L}(E)$ dans une base \mathcal{B} .

Rappelons que si p est la projection sur F parallèlement à G , alors $F = \text{Im}(p)$ et $G = \text{Ker}(p)$.

On connaît la matrice de p dans une base adaptée à la somme directe $E = \text{Im}(p) \oplus \text{Ker}(p)$: elle est diagonale, ne possède que des 0 et des 1 sur sa diagonale, et le nombre de 1 est égal à $\dim \text{Im } p$.

Fort de ces deux résultats, on procède ainsi :

1. Déterminer une base de $\text{Im}(p)$ et une base de $\text{Ker}(p)$.
2. La concaténation de ces deux bases est une base \mathcal{B}' de E adaptée à la somme directe $E = \text{Im}(p) \oplus \text{Ker}(p)$.
3. Déterminer la matrice de passage de \mathcal{B} à \mathcal{B}' , ainsi que son inverse.
4. La matrice de p dans la base \mathcal{B}' est donc diagonale avec des 0 et des 1 sur la diagonale. Et donc, par la formule de changement de base,

$$\text{Mat}_{\mathcal{B}}(p) = P_{\mathcal{B}, \mathcal{B}'} \text{Mat}_{\mathcal{B}'}(p) P_{\mathcal{B}, \mathcal{B}'}^{-1}.$$

Exercices de TD : 5.4

✓ MÉTHODE 5.4 : Utiliser des matrices de passage.

• Savoir écrire la matrice d'une famille de vecteurs dans une base et l'utiliser pour déterminer si une famille de vecteurs est une base ou non.

• Savoir différencier $P_{\mathcal{B}, \mathcal{B}'}$ de $P_{\mathcal{B}', \mathcal{B}}$ et connaître le lien entre les deux.

• Utiliser les matrices de passage dans les formules de changement de base.

Exercices de TD : 5.4, 5.5, 5.8, 5.10

✓ MÉTHODE 5.5 : Montrer que deux matrices A et B ne sont pas semblables

• Montrer que A et B n'ont pas le même rang ou pas la même trace.

• Si A ou B est une matrice scalaire, vérifier que l'autre ne l'est pas (la seule matrice semblable à une matrice scalaire est elle-même.)

Exercices de TD : 5.12



MÉTHODE 5.6 : Montrer que deux matrices A et B sont semblables

Déterminer un endomorphisme f dont la matrice dans deux bases différentes sont A et B . En général pour cela, on définit f comme étant l'unique endomorphisme de \mathbf{K}^n dont la matrice dans la base canonique est A . Et alors on montre qu'il existe une base de \mathbf{K}^n dans laquelle la matrice de f est B . Pour cela, on essaiera de chercher ce que doit vérifier une telle base.

Exercices de TD : 5.11, 5.12, 5.16, 5.19

EXERCICES DU CHAPITRE 5

EXERCICE 5.1 Un grand classique

Soit $A \in \mathcal{M}_n(\mathbf{R})$. Montrer que $\text{tr}({}^tAA) = 0$ si et seulement si $A = 0$.

PD

EXERCICE 5.2 Déterminer le rang des matrices suivantes, et donner leur inverse s'il existe.

F

$$A = \begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & 4 \\ -1 & 0 & -3 \end{pmatrix}, B = \begin{pmatrix} 1 & 0 & -3 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \\ -2 & -1 & 2 \end{pmatrix}, C = \begin{pmatrix} 0 & -1 & 1 \\ 2 & -3 & 4 \\ 1 & -1 & 2 \end{pmatrix}, D = \begin{pmatrix} 1 & -1 & -2 & 2 \\ -1 & -1 & 0 & 2 \\ 0 & 2 & 2 & -4 \\ 0 & 1 & 1 & -2 \end{pmatrix}.$$

EXERCICE 5.3 Soit $(x_1, \dots, x_n) \in \mathbf{R}^n$ non nul. Déterminer le rang des matrices

PD

$$M = \begin{pmatrix} x_1 & x_2 & \dots & x_n & 0 \\ 0 & 0 & & 0 & x_1 \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \dots & 0 & x_n \end{pmatrix} \in \mathcal{M}_{n+1}(\mathbf{R}) \text{ et } N = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} (x_1 \ \dots \ x_n) \in \mathcal{M}_n(\mathbf{R}).$$

EXERCICE 5.4 Matrice d'une projection

PD

Dans \mathbf{R}^3 , on pose $F = \{(x, y, z) \in \mathbf{R}^3 : 2x + 2y - z = 0\}$ et $G = \text{Vect}((-1, 1, 1))$.

- 1) Montrer que F et G sont supplémentaires dans \mathbf{R}^3 . Déterminer une base de F et en déduire une base \mathcal{B} de \mathbf{R}^3 adaptée à la somme directe $F \oplus G$.
- 2) On note p la projection sur F parallèlement à G .
 - a) Déterminer la matrice de p dans la base \mathcal{B} .
 - b) En déduire la matrice de p dans la base canonique de \mathbf{R}^3 .
 - c) Soit q la projection sur G parallèlement à F . Montrer que $q = \text{id}_E - p$ et en déduire la matrice de q dans la base canonique.

EXERCICE 5.5 Formule de changement de base

F

Soit $\mathcal{B} = (e_1, e_2, e_3)$ une base de \mathbf{R}^3 , et soit $f \in \mathcal{L}(\mathbf{R}^3)$ tel que $\text{Mat}_{\mathcal{B}}(f) = A = \begin{pmatrix} -1 & 13 & -16 \\ 0 & 7 & -10 \\ 0 & 5 & -7 \end{pmatrix}$.

- 1) Montrer que $\mathcal{B}' = (e_1, 5e_1 + 3e_2 + 2e_3, 2e_1 + e_2 + e_3)$ est une base de \mathbf{R}^3 .
- 2) Déterminer la matrice de f dans la base \mathcal{B}' . En déduire la valeur de A^4 .

EXERCICE 5.6 Montrer qu'il n'existe pas deux matrices $A, B \in \mathcal{M}_n(\mathbf{R})$ telles que $AB - BA = I_n$.

PD

EXERCICE 5.7 Soit φ l'endomorphisme de $\mathbf{R}_n[X]$ défini par $\varphi(P) = P(X + 1)$.

PD

- 1) Déterminer la matrice A de φ dans la base canonique de $\mathbf{R}_n[X]$.
- 2) Justifier que A est inversible, et calculer A^{-1} .

EXERCICE 5.8 Soit φ l'application définie sur $\mathbf{R}_2[X]$ et à valeurs dans $\mathbf{R}[X]$, qui à un polynôme P associe $\varphi(P) = (X^2 - 1)P' - 2XP$.

PD

- 1) Montrer que φ est un endomorphisme de $\mathbf{R}_2[X]$.
- 2) Déterminer la matrice A de φ dans la base canonique, notée \mathcal{B}_{can} .
- 3) On pose $P_1 = (X + 1)^2, P_2 = X^2 - 1$ et $P_3 = (X - 1)^2$. Montrer que $\mathcal{B} = (P_1, P_2, P_3)$ est une base de $\mathbf{R}_2[X]$, puis donner la matrice de passage de la base canonique à \mathcal{B} .
- 4) En déduire la matrice de φ dans la base \mathcal{B} .
- 5) Pour $n \in \mathbf{N}$, déterminer la matrice de φ^n dans la base \mathcal{B} . En déduire la matrice φ^n dans la base canonique.

EXERCICE 5.9 Soient $A, B \in \mathcal{M}_3(\mathbf{R})$ telles que $AB = 0$. Montrer que l'une au moins de ces deux matrices est de rang inférieur ou égal à 1.

PD

EXERCICE 5.10 On considère les matrices $A = \begin{pmatrix} 0 & -2 & -2 \\ 2 & 5 & 3 \\ -2 & -4 & -2 \end{pmatrix}$ et $D = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}$.

AD

Soit $f \in \mathcal{L}(\mathbf{R}^3)$ l'endomorphisme de \mathbf{R}^3 dont la matrice dans la base canonique \mathcal{B} est A .

- 1) Montrer qu'il existe une base $\mathcal{C} = (e_1, e_2, e_3)$ de \mathbf{R}^3 telle que la matrice de f dans \mathcal{C} soit D .
- 2) Déterminer une matrice P de $GL_3(\mathbf{R})$ telle que $A = PDP^{-1}$. Calculer P^{-1} .
- 3) Calculer $A^n, \forall n \in \mathbf{N}^*$.

EXERCICE 5.11 Soit $A = \begin{pmatrix} 3 & -3 & -1 \\ 1 & 1 & 0 \\ -1 & -3 & 0 \end{pmatrix}$ et $T = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$.

- 1) Montrer que A est semblable à T , et déterminer $P \in GL_3(\mathbf{R})$ telle que $A = P^{-1}TP$.
- 2) Pour $n \in \mathbf{N}$, calculer T^n , puis en déduire A^n .

EXERCICE 5.12 Dans chaque cas, déterminer si les matrices A et B sont semblables :

1. $A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}, B = \begin{pmatrix} -1 & 0 & 3 \\ 0 & 4 & 0 \\ 0 & 0 & 1 \end{pmatrix}$
2. $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}, B = \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix}$
3. $A = \begin{pmatrix} 2 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \end{pmatrix}, B = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$
4. $A = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}, B = \begin{pmatrix} 3 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}$

EXERCICE 5.13 Soit $A \in \mathcal{M}_n(\mathbf{C})$. Déterminer l'ensemble des matrices $M \in \mathcal{M}_n(\mathbf{C})$ vérifiant $M + {}^tM = \text{tr}(M)A$.
On pourra distinguer les cas $\text{tr}(A) = 2$ et $\text{tr}(A) \neq 2$.

EXERCICE 5.14 Soient A et B deux matrices semblables de $\mathcal{M}_n(\mathbf{K})$, et soit $P \in \mathbf{K}[X]$. Montrer que P est un polynôme annulateur de A si et seulement si il est annulateur de B .

EXERCICE 5.15 Soit $F = \{M \in \mathcal{M}_n(\mathbf{R}) : \text{tr}(M) = 0\}$. Montrer que F est un sous-espace vectoriel de $\mathcal{M}_n(\mathbf{R})$, et montrer que si A est une matrice de $\mathcal{M}_n(\mathbf{R})$ de trace non nulle, alors $\mathcal{M}_n(\mathbf{R}) = F \oplus \text{Vect}(A)$.

EXERCICE 5.16 Soit $M \in \mathcal{M}_n(\mathbf{R})$ une matrice vérifiant $\text{rg}(M) = \text{tr}(M) = 1$.

- 1) Montrer que M est semblable à une matrice dont les $n - 1$ premières colonnes sont nulles.
- 2) Soit f l'endomorphisme de \mathbf{R}^n dont la matrice dans la base canonique est M . Montrer que f est un projecteur.

EXERCICE 5.17 Théorème d'Hadamard pour les matrices à diagonale strictement dominante

- 1) Soit $A \in \mathcal{M}_n(\mathbf{R})$. Montrer que A est inversible si et seulement si : $\forall X \in \mathcal{M}_{n,1}(\mathbf{R}), AX = 0 \Rightarrow X = 0$.
- 2) Soit $A = (a_{i,j}) \in \mathcal{M}_n(\mathbf{R})$ telle que $\forall i \in \llbracket 1, n \rrbracket, |a_{i,i}| > \sum_{\substack{j=1 \\ j \neq i}}^n |a_{i,j}|$. Soit $X \in \mathcal{M}_{n,1}(\mathbf{R})$ tel que $AX = 0$, et soit $i_0 \in \llbracket 1, n \rrbracket$ tel que $|x_{i_0}| = \max_i |x_i|$.
Montrer que $x_{i_0} = 0$, et en déduire que A est inversible.

EXERCICE 5.18

- 1) Soit $A = (a_{i,j})_{1 \leq i, j \leq n} \in \mathcal{M}_n(\mathbf{R})$. Montrer que $\text{tr}({}^tAA) = \sum_{i=1}^n \sum_{j=1}^n a_{i,j}^2$.
- 2) Pour $A \in \mathcal{M}_n(\mathbf{R})$, on définit une application φ_A sur $\mathcal{M}_n(\mathbf{R})$ par $\varphi : M \mapsto \text{tr}({}^tAM)$.
 - a) Montrer que φ_A est une forme linéaire sur $\mathcal{M}_n(\mathbf{R})$.
 - b) Prouver que $A \mapsto \varphi_A$ est un isomorphisme de $\mathcal{M}_n(\mathbf{R})$ sur $\mathcal{L}(\mathcal{M}_n(\mathbf{R}), \mathbf{R})$.
 - c) Soit φ une forme linéaire sur $\mathcal{M}_n(\mathbf{R})$. Prouver qu'il existe un unique $A \in \mathcal{M}_n(\mathbf{R})$ tel que $\varphi = \varphi_A$.

EXERCICE 5.19 (QSP HEC 2014)

Soit $A \in \mathcal{M}_3(\mathbf{R})$ une matrice non nulle telle que $A^2 = 0$. Montrer que A est semblable à $\begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$.

EXERCICE 5.20 (QSP HEC 2016)

Soit $n \geq 2$, et soit $M = (m_{i,j})_{1 \leq i, j \leq n} \in \mathcal{M}_n(\mathbf{R})$ la matrice définie par $\forall (i, j) \in \llbracket 1, n \rrbracket^2, m_{i,j} = i^{j-1}$.
Montrer que M est inversible

EXERCICE 5.21 Hyperplans de $\mathcal{M}_n(\mathbf{R})$ stables pour le produit matriciel (Oral ESCP 2008)

Soit n un entier supérieur ou égal à 2. Soit F un sous-espace vectoriel de $\mathcal{M}_n(\mathbf{R})$, de dimension $n^2 - 1$, stable par la multiplication matricielle : $\forall (M, M') \in F^2, MM' \in F$. On suppose que $I_n \notin F$, où I_n désigne la matrice identité de $\mathcal{M}_n(\mathbf{R})$.

- 1) Montrer que $\mathcal{M}_n(\mathbf{R}) = F \oplus \text{Vect}(I_n)$.

- 2) a) Soit p le projecteur de $\mathcal{M}_n(\mathbf{R})$ sur $\text{Vect}(I_n)$ parallèlement à F .
Montrer que $\forall (M, M') \in (\mathcal{M}_n(\mathbf{R}))^2$, $p(MM') = p(M)p(M')$.
- b) Montrer que pour toute matrice M de $\mathcal{M}_n(\mathbf{R})$ telle que $M^2 \in F$, alors $M \in F$.
- c) Soit $(E_{i,j})_{1 \leq i, j \leq n}$ la base canonique de $\mathcal{M}_n(\mathbf{R})$.
Calculer $E_{i,j} \times E_{k,\ell}$ pour $(i, j, k, \ell) \in \llbracket 1, n \rrbracket^2$.
- d) Montrer que F contient la base canonique de $\mathcal{M}_n(\mathbf{R})$.
- e) Conclure.

CORRECTION DES EXERCICES DU CHAPITRE 5

SOLUTION DE L'EXERCICE 5.1

Il est évident que si $A = 0$, alors ${}^tAA = 0$ et donc $\text{tr}({}^tAA) = 0$.
Inversement, supposons que $A = (a_{i,j})_{1 \leq i, j \leq n}$ soit telle que $\text{tr}({}^tAA) = 0$.
Alors

$$\begin{aligned}\text{tr}({}^tAA) &= \sum_{i=1}^n ({}^tAA)_{i,i} \\ &= \sum_{i=1}^n \sum_{k=1}^n ({}^tA)_{i,k} a_{k,i} \\ &= \sum_{i=1}^n \sum_{k=1}^n a_{k,i} a_{k,i} \\ &= \sum_{i=1}^n \sum_{k=1}^n a_{k,i}^2.\end{aligned}$$

Mais une somme de nombres positifs est nulle si et seulement si chacun de ces nombres est nul : $\text{tr}({}^tAA) = 0$ si et seulement si pour tout $(i, k) \in \llbracket 1, n \rrbracket^2$, $a_{k,i}^2 = 0 \Leftrightarrow a_{k,i} = 0$.
Et donc $\text{tr}({}^tAA) = 0$ si et seulement si $A = 0$.

SOLUTION DE L'EXERCICE 5.2

Il s'agit de calculer le rang par la méthode du pivot, et, le cas échéant¹, de calculer l'inverse toujours par la même méthode.

A est de rang 3, donc inversible, et $A^{-1} = \begin{pmatrix} -3 & -3 & -4 \\ -4 & -3 & -4 \\ 1 & 1 & 1 \end{pmatrix}$.

B est de rang 3. N'étant pas carrée, elle ne peut être inversible.

C est de rang 3, donc inversible, et $C^{-1} = \begin{pmatrix} -2 & 1 & -1 \\ 0 & -1 & 2 \\ 1 & -1 & 2 \end{pmatrix}$.

D est de rang 2. Notons que cela se trouve aussi sans pivot si l'on observe qu'il existe des relations entre les colonnes, C_1, \dots, C_4 de D .
En effet, on a $C_4 = -2C_2$ et $C_3 = C_2 - C_1$. Par conséquent, $\text{Vect}(C_1, C_2, C_3, C_4) = \text{Vect}(C_1, C_2)$.
Puisque ces deux colonnes ne sont pas colinéaires, elles forment une famille libre de $\text{Vect}(C_1, C_2)$, qui est donc de dimension 2.

SOLUTION DE L'EXERCICE 5.3

Notons C_1, \dots, C_{n+1} les colonnes de M .

Puisque $(x_1, \dots, x_n) \in \mathbf{R}^n$ est non nul, il existe $i \in \llbracket 1, n \rrbracket$ tel que $x_i \neq 0$.

Alors l'espace engendré par les n premières colonnes de M est une droite vectorielle de $\mathcal{M}_{n+1,1}(\mathbf{R})$, engendrée par C_i . En effet, pour tout $j \in \llbracket 1, n \rrbracket$, on a $C_j = \frac{x_j}{x_i} C_i$.

Par conséquent, l'espace vectoriel engendré par les n premières colonnes de M est de dimension 1.

De plus, la dernière colonne n'est pas dans cet espace vectoriel car elle n'est pas colinéaire à C_i .

On en déduit que la $i^{\text{ème}}$ colonne de M et sa dernière colonne forment une famille libre, et donc une base de l'espace vectoriel engendré par les colonnes de M . Ainsi, $\text{rg}(M) = 2$.

Notons que $N = \begin{pmatrix} x_1x_1 & x_1x_2 & \dots & x_1x_n \\ x_2x_1 & x_2x_2 & \dots & x_2x_n \\ \vdots & \vdots & & \vdots \\ x_nx_1 & x_nx_2 & \dots & x_nx_n \end{pmatrix}$.

En particulier, toutes les colonnes de N sont colinéaires à $\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ qui n'est pas nul. Alors N est non nulle à son tour, car si $x_i \neq 0$, alors le coefficient (i, i) de N , qui est x_i^2 est non nul.

SOLUTION DE L'EXERCICE 5.4

Détail

Par définition, le coefficient (i, k) de tA est le coefficient (k, i) de A .

¹ Rappelons qu'une matrice carrée de taille n est inversible si et seulement si elle est de rang n .

Rappel

Par définition, le rang d'une matrice est la dimension de l'espace engendré par ses colonnes.

Définition

Ici pas de pivot, mais un retour à la définition du rang de M : c'est le rang de la famille des vecteurs colonnes de M .

Rédaction

Rappelons qu'une matrice dont toutes les colonnes sont proportionnelles est de rang 1 si elle est non nulle. Il n'est donc pas inutile de préciser que N est non nulle.

1. Notons que H est de dimension 2 car c'est le noyau de la forme linéaire $(x, y, z) \mapsto 2x + 2y - z$, et donc c'est un hyperplan de \mathbf{R}^3 .
Pour $(x, y, z) \in \mathbf{R}^3$, on a

$$(x, y, z) \in F \Leftrightarrow (x, y, z) = (x, y, 2x + 2y) \Leftrightarrow (x, y, z) = x(1, 0, 2) + y(0, 1, 2) \\ \Leftrightarrow (x, y, z) \in \text{Vect}((1, 0, 2), (0, 1, 2)).$$

Et donc la famille $(1, 0, 2), (0, 1, 2)$ est génératrice de F . Étant formée de deux vecteurs non colinéaires elle est libre et donc c'est une base de F .

La famille $(1, 0, 2), (0, 1, 2), (-1, 1, 1)$ est donc la concaténation d'une base de F et d'une base de G .

Sa matrice dans la base canonique est $P = \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & 1 \\ 2 & 2 & 1 \end{pmatrix}$.

Un calcul de pivot nous prouve que cette matrice est inversible et que $P^{-1} = \begin{pmatrix} -1 & -2 & 1 \\ 2 & 3 & -1 \\ -2 & -2 & 1 \end{pmatrix}$.

En particulier, $\mathcal{B} = (1, 0, 2), (0, 1, 2), (-1, 1, 1)$ est une base de \mathbf{R}^3 et donc la somme $F + G$ est directe. Puisque $\dim(F \oplus G) = 3 = \dim \mathbf{R}^3$, on alors $\mathbf{R}^3 = F \oplus G$. Et alors, \mathcal{B} , qui est la concaténation d'une base de F et d'une base de G , est bien une base adaptée à la somme directe.

- 2.a. Puisque \mathcal{B} est une base de \mathbf{R}^3 adaptée à la somme directe $F \oplus G$.
De plus $\text{rg}(p) = \dim \text{Im}(p) = \dim F = 2$ et donc la matrice de p dans la base \mathcal{B} est

$$\text{Mat}_{\mathcal{B}}(p) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

- 2.b. Par la formule de changement de base, on a

$$\text{Mat}_{\mathcal{B}_{can}}(p) = P_{\mathcal{B}_{can}, \mathcal{B}} \text{Mat}_{\mathcal{B}}(p) P_{\mathcal{B}, \mathcal{B}_{can}}.$$

Or, par définition, la matrice de passage de la base canonique à \mathcal{B} est précisément la matrice P de la question 1, donc

$$A = \text{Mat}_{\mathcal{B}_{can}}(p) = \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & 1 \\ 2 & 2 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} -1 & -2 & 1 \\ 2 & 3 & -1 \\ -2 & -2 & 1 \end{pmatrix} = \begin{pmatrix} -1 & -2 & 1 \\ 2 & 3 & -1 \\ 2 & 2 & 0 \end{pmatrix}.$$

SOLUTION DE L'EXERCICE 5.5

1. Écrivons la matrice de la famille \mathcal{B}' dans la base \mathcal{B} . Il s'agit de $P = \begin{pmatrix} 1 & 5 & 2 \\ 0 & 3 & 1 \\ 0 & 2 & 1 \end{pmatrix} \begin{matrix} e_1 \\ e_2 \\ e_3 \end{matrix}$.

On vérifie aisément que cette matrice est inversible, donc \mathcal{B}' est une base et P est la matrice de passage de \mathcal{B} à \mathcal{B}' .

De plus, $P_{\mathcal{B}', \mathcal{B}} = P^{-1} = \begin{pmatrix} 1 & -1 & -1 \\ 0 & 1 & -1 \\ 0 & -2 & 3 \end{pmatrix}$.

2. D'après la formule de changement de base, la matrice de f dans la base \mathcal{B}' est

$$M = \text{Mat}_{\mathcal{B}'}(f) = P_{\mathcal{B}', \mathcal{B}} \text{Mat}_{\mathcal{B}}(f) P_{\mathcal{B}, \mathcal{B}'} = P^{-1} A P = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}.$$

On constate que $M^2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$ et donc $M^4 = I_3$.

On en déduit² que $A^4 = P M^4 P^{-1} = P P^{-1} = I_3$.

SOLUTION DE L'EXERCICE 5.6

Raisonnons par l'absurde et supposons au contraire que deux telles matrices existent. Alors, par linéarité de la trace, on aurait

$$\text{tr}(AB) - \text{tr}(BA) = \text{tr}(I_n) = n.$$

Remarque

Si on n'a pas vu précédemment que F était un hyperplan, on le retrouve ici : $\dim F = 2$.

Somme directe

Rappelons que F et G sont supplémentaires si et seulement si la concaténation d'une base de F et d'une base de G est une base de \mathbf{R}^3 .
Ce n'est bien évidemment pas la seule méthode de prouver que F et G sont supplémentaires dans \mathbf{R}^3 .

Vérification

Un bon moyen de vérifier son résultat est de s'assurer que cette matrice vérifie $A^2 = A$, signe que c'est bien une matrice de projecteur !

Danger !

Attention au sens : on a écrit les vecteurs de \mathcal{B}' dans la base \mathcal{B} , on a donc la matrice de passage de \mathcal{B} à \mathcal{B}' .

² Sans faire le calcul : on sait que A^4 est semblable à $M^4 = I_3$, et la seule matrice semblable à I_3 est I_3 elle-même.

Or, nous savons que $\text{tr}(AB) = \text{tr}(BA)$, et donc l'égalité précédente s'écrit alors $0 = n$, d'où une contradiction : il n'est pas possible de trouver un couple de matrices $(A, B) \in \mathcal{M}_n(\mathbf{R})^2$ vérifiant $AB - BA = I_n$.

SOLUTION DE L'EXERCICE 5.7

1. Rappelons que la base canonique de $\mathbf{R}_n[X]$ est $(1, X, \dots, X^n)$.
On a $\varphi(1) = 1$, et pour $k \geq 1$,

$$\varphi(X^k) = (X + 1)^k = \sum_{i=0}^k \binom{k}{i} X^i.$$

On en déduit que la matrice A est égale à

$$A = \begin{pmatrix} \varphi(1) & \varphi(X) & \varphi(X^2) & \varphi(X^3) & \dots & \varphi(X^n) \\ 1 & 1 & 1 & 1 & \dots & 1 \\ 0 & 1 & \binom{2}{1} & \binom{3}{1} & \dots & \binom{n}{1} \\ 0 & 0 & \binom{2}{2} & \binom{3}{2} & \dots & \binom{n}{2} \\ \vdots & \vdots & & \ddots & & \vdots \\ 0 & 0 & 0 & \dots & \dots & \binom{n}{n} \end{pmatrix} \begin{matrix} 1 \\ X \\ X^2 \\ \vdots \\ X^n \end{matrix} \in \mathcal{M}_{n+1}(\mathbf{R}).$$

2. Nous voyons directement que A est inversible, puisqu'elle est triangulaire supérieure, avec des coefficients tous non nuls³ sur la diagonale, donc de rang $n + 1$.
Toutefois, pour calculer A^{-1} , il est préférable de remarquer que φ est un isomorphisme, et que $\varphi^{-1} : P \mapsto P(X - 1)$.
Et par conséquent, la matrice de φ^{-1} dans la base canonique est l'inverse de $A = \text{Mat}_{\mathcal{B}_{can}}(\varphi)$.
Or, pour tout $k \in \llbracket 1, n \rrbracket$,

$$\varphi^{-1}(X^k) = (X - 1)^k = \sum_{i=0}^k \binom{k}{i} X^i (-1)^{k-i}.$$

On en déduit que

$$A^{-1} = \text{Mat}_{\mathcal{B}_{can}}(\varphi^{-1}) = \begin{pmatrix} \varphi^{-1}(1) & \varphi^{-1}(X) & \varphi^{-1}(X^2) & \dots & \varphi^{-1}(X^n) \\ 1 & -1 & 1 & \dots & (-1)^n \\ 0 & 1 & -\binom{2}{1} & \dots & (-1)^{n-1} \binom{n}{1} \\ 0 & 0 & 1 & \dots & (-1)^{n-2} \binom{n}{2} \\ \vdots & \vdots & & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 \end{pmatrix} \begin{matrix} 1 \\ X \\ X^2 \\ \vdots \\ X^n \end{matrix}.$$

SOLUTION DE L'EXERCICE 5.8

1. Commençons par prouver que φ est bien à valeurs dans $\mathbf{R}_2[X]$.
Pour $P = aX^2 + bX + c \in \mathbf{R}_2[X]$, on a

$$\varphi(P) = (X^2 - 1)(2aX + b) - 2X(aX^2 + bX + c) = -bX^2 - 2(a + c)X - b \in \mathbf{R}_2[X].$$

De plus, pour $P, Q \in \mathbf{R}_2[X]$ et $\lambda \in \mathbf{R}$, on a

$$\begin{aligned} \varphi(\lambda P + Q) &= (X^2 - 1)(\lambda P + Q)' - 2X(\lambda P + Q) \\ &= (X^2 - 1)\lambda P' + (X^2 - 1)Q' - 2\lambda XP - 2XQ \\ &= \lambda((X^2 - 1)P' - 2XP) + (X^2 - 1)Q' - 2XQ = \lambda\varphi(P) + \varphi(Q). \end{aligned}$$

Donc φ est linéaire : c'est bien un endomorphisme de $\mathbf{R}_2[X]$.

2. On a $\varphi(1) = -2X$, $\varphi(X) = X^2 - 1 - 2X^2 = -X^2 - 1$ et $\varphi(X^2) = 2X(X^2 - 1) - 2X^3 = -2X$.
On en déduit que

$$\text{Mat}_{\mathcal{B}_{can}}(\varphi) = \begin{pmatrix} \varphi(1) & \varphi(X) & \varphi(X^2) \\ 0 & -1 & 0 \\ -2 & 0 & -2 \\ 0 & -1 & 0 \end{pmatrix} \begin{matrix} 1 \\ X \\ X^2 \end{matrix}.$$

Décalage

Attention : lorsqu'on calcule $\varphi(X^k)$, on obtient en fait les coefficients de la $(k + 1)$ ^{ème} colonne de A , et non ceux de la k ^{ème}.

Ce décalage est usuel lorsqu'on travaille avec des endomorphismes de $\mathbf{R}_n[X]$, car le k ^{ème} élément de la base canonique est X^{k-1} .

Le même problème se pose avec les lignes.

³ Et en fait tous égaux à 1.

Remarque

Au premier abord, on peut seulement affirmer que les deux polynômes $(X^2 - 1)P'$ et $2XP'$ sont de degré au plus 3. Il s'agit donc de prouver que leurs termes de degré 3 sont les mêmes, afin que $\varphi(P)$ soit dans $\mathbf{R}_2[X]$.

3. On a $P_1 = X^2 + 2X + 1$, $P_2 = X^2 - 1$ et $P_3 = X^2 - 2X + 1$.
Et donc la matrice de (P_1, P_2, P_3) dans la base \mathcal{B}_{can} est

$$\text{Mat}_{\mathcal{B}_{can}}(P_1, P_2, P_3) = \begin{pmatrix} P_1 & P_2 & P_3 \\ 1 & -1 & 1 \\ 2 & 0 & -2 \\ 1 & 1 & 1 \end{pmatrix} \begin{matrix} 1 \\ X \\ X^2 \end{matrix}.$$

Un calcul de pivot prouve que cette matrice est inversible, et donc que (P_1, P_2, P_3) est une base de $\mathbf{R}_2[X]$.

$$\text{Alors } P_{\mathcal{B}_{can}, \mathcal{B}} = \begin{pmatrix} 1 & -1 & 1 \\ 2 & 0 & -2 \\ 1 & 1 & 1 \end{pmatrix} \text{ et } P_{\mathcal{B}, \mathcal{B}_{can}} = P_{\mathcal{B}_{can}, \mathcal{B}}^{-1} = \frac{1}{4} \begin{pmatrix} 1 & 1 & 1 \\ -2 & 0 & 2 \\ 1 & -1 & 1 \end{pmatrix}.$$

4. D'après la formule de changement de base, on a

$$\text{Mat}_{\mathcal{B}}(\varphi) = P_{\mathcal{B}, \mathcal{B}_{can}} \text{Mat}_{\mathcal{B}_{can}}(\varphi) P_{\mathcal{B}_{can}, \mathcal{B}} = \begin{pmatrix} -2 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 2 \end{pmatrix}.$$

Remarque

Le calcul de $\varphi(P_1)$ nous aurait permis facilement de constater que $\varphi(P_1) = -2P_1$.

5. Puisque pour élever une matrice diagonale à la puissance n , il suffit d'élever chacun de ses coefficients à la puissance n , on a $\text{Mat}_{\mathcal{B}}(\varphi^n) = (\text{Mat}_{\mathcal{B}}(\varphi))^n = \begin{pmatrix} (-2)^n & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 2^n \end{pmatrix}$.

Toujours d'après la formule de changement de base, on a alors

$$\begin{aligned} \text{Mat}_{\mathcal{B}_{can}}(\varphi^n) &= P_{\mathcal{B}_{can}, \mathcal{B}} \text{Mat}_{\mathcal{B}}(\varphi^n) P_{\mathcal{B}, \mathcal{B}_{can}} \\ &= \begin{pmatrix} 1 & -1 & 1 \\ 2 & 0 & -2 \\ 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} (-2)^n & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 2^n \end{pmatrix} \frac{1}{4} \begin{pmatrix} 1 & 1 & 1 \\ -2 & 0 & 2 \\ 1 & -1 & 1 \end{pmatrix} \\ &= \frac{1}{4} \begin{pmatrix} (-2)^n & 0 & 2^n \\ 2(-2)^n & 0 & -2^{n+1} \\ (-2)^n & 0 & 2^n \end{pmatrix} \begin{pmatrix} 1 & 1 & 1 \\ -2 & 0 & 2 \\ 1 & -1 & 1 \end{pmatrix} \\ &= \frac{1}{4} \begin{pmatrix} (-2)^n + 2^n & (-2)^n - 2^n & (-2)^n + 2^n \\ -(-2)^{n+1} - 2^{n+1} & -(-2)^{n+1} + 2^{n+1} & -(-2)^{n+1} - 2^{n+1} \\ (-2)^n + 2^n & (-2)^n - 2^n & (-2)^n + 2^n \end{pmatrix}. \end{aligned}$$

Notons qu'on peut donner un résultat un peu moins dur à manipuler en distinguant suivant la parité de n :

$$\text{Mat}_{\mathcal{B}_{can}}(\varphi^n) = \begin{pmatrix} 2^{n-1} & 0 & 2^{n-1} \\ 0 & 2^n & 0 \\ 2^{n-1} & 0 & 2^{n-1} \end{pmatrix} \text{ si } n \text{ est pair, } \quad \text{Mat}_{\mathcal{B}_{can}}(\varphi^n) = \begin{pmatrix} 0 & -2^{n-1} & 0 \\ -2^n & 0 & -2^n \\ 0 & -2^{n-1} & 0 \end{pmatrix} \text{ si } n \text{ est impair.}$$

SOLUTION DE L'EXERCICE 5.9

Notons f et g les endomorphismes de \mathbf{R}^3 dont les matrices dans la base canonique de \mathbf{R}^3 sont A et B .

Alors $AB = 0 \Leftrightarrow f \circ g = 0$. Ainsi⁴, nous avons $\text{Im}(g) \subset \text{Ker } f$.

Supposons par l'absurde que $\text{rg}(f) \geq 2$ et $\text{rg}(g) \geq 2$.

Alors $\dim \text{Ker } f = 3 - \text{rg}(f) \leq 1$.

Mais $\dim \text{Im } g \geq 2$, et donc il n'est pas possible d'avoir $\text{Im } g \subset \text{Ker } f$, d'où une contradiction.

On en déduit que $\text{rg}(A) = \text{rg}(f) \leq 1$ ou $\text{rg}(B) = \text{rg}(g) \leq 1$.

SOLUTION DE L'EXERCICE 5.10

1. Notons que si une telle base existe, alors on doit avoir

$$\text{Mat}_{\mathcal{E}}(f) = \begin{pmatrix} f(e_1) & f(e_2) & f(e_3) \\ 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix} \begin{matrix} e_1 \\ e_2 \\ e_3 \end{matrix}.$$

⁴ Voir l'exercice 7 du TD3.

Méthode

Cette partie n'a pas besoin d'être rédigée et on peut se contenter de la faire au brouillon : elle permet de clarifier les conditions que doivent vérifier les vecteurs (e_1, e_2, e_3) et nous indique donc quels systèmes résoudre.

Nous allons donc prouver qu'il existe une base (e_1, e_2, e_3) vérifiant $f(e_1) = 0$, $f(e_2) = e_2$ et $f(e_3) = e_3$.

Soit donc $(x, y, z) \in \mathbb{R}^3$. On a $f(x, y, z) = 0$ si et seulement si

$$A \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \Leftrightarrow \begin{cases} -2y - 2z = 0 \\ 2x + 5y + 3z = 0 \\ -2x - 4y - 2z = 0 \end{cases} \Leftrightarrow \begin{cases} y = -z \\ 2x - 2z = 0 \\ -2x + 2z = 0 \end{cases} \Leftrightarrow \begin{cases} y = -z \\ x = z \end{cases}$$

Donc l'ensemble des solutions est $\text{Vect}((1, -1, 1))$.

Prenons donc $e_1 = (1, -1, 1)$.

De même, on a $f(x, y, z) = (x, y, z)$ si et seulement si

$$A \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x \\ y \\ z \end{pmatrix} \Leftrightarrow \begin{cases} -2x - 2z = x \\ 2x + y + 3z = y \\ -2x - 4y - 2z = z \end{cases} \Leftrightarrow \begin{cases} x = -2y \\ z = 0 \end{cases}$$

On peut donc prendre $e_2 = (-2, 1, 0)$.

De même, on a $f(x, y, z) = 2(x, y, z)$ si et seulement si

$$A \begin{pmatrix} x \\ y \\ z \end{pmatrix} = 2 \begin{pmatrix} x \\ y \\ z \end{pmatrix} \Leftrightarrow \begin{cases} -2x - 2z = 2x \\ 2x + y + 3z = 2y \\ -2x - 4y - 2z = 2z \end{cases} \Leftrightarrow \begin{cases} x = 0 \\ y = -z \end{cases}$$

On peut donc prendre $e_3 = (0, 1, -1)$.

Vérifions à présent que la famille (e_1, e_2, e_3) ainsi choisie est bien une base de \mathbb{R}^3 .

La matrice de cette famille dans la base canonique est

$$P = \text{Mat}_{\mathcal{B}}(e_1, e_2, e_3) = \begin{pmatrix} e_1 & e_2 & e_3 \\ 1 & -2 & 0 \\ -1 & 1 & 1 \\ 1 & 0 & -1 \end{pmatrix}.$$

Un pivot montre que cette matrice est inversible, et donc que (e_1, e_2, e_3) est une base de \mathbb{R}^3 .

Et alors, nécessairement,

$$\text{Mat}_{\mathcal{E}}(f) = \begin{pmatrix} f(e_1) & f(e_2) & f(e_3) \\ 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix} \begin{matrix} e_1 \\ e_2 \\ e_3 \end{matrix} = D.$$

2. Puisque (e_1, e_2, e_3) est une base de \mathbb{R}^3 , par définition d'une matrice de passage, on a $P_{\mathcal{B}, \mathcal{E}} = \text{Mat}_{\mathcal{B}}(e_1, e_2, e_3) = P$.
Et alors, par la formule de changement de base,

$$A = \text{Mat}_{\mathcal{B}}(f) = P_{\mathcal{B}, \mathcal{E}} \text{Mat}_{\mathcal{E}}(f) P_{\mathcal{B}, \mathcal{E}}^{-1} = PDP^{-1}.$$

De plus, on a alors $P^{-1} = \begin{pmatrix} 1 & 2 & 2 \\ 0 & 1 & 1 \\ 1 & 2 & 1 \end{pmatrix}$.

3. On a alors, pour tout $n \in \mathbb{N}^*$,

$$\begin{aligned} A^n &= (PDP^{-1})^n = PD^n P^{-1} \\ &= \begin{pmatrix} 1 & -2 & 0 \\ -1 & 1 & 1 \\ 1 & 0 & -1 \end{pmatrix} \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2^n \end{pmatrix} \begin{pmatrix} 1 & 2 & 2 \\ 0 & 1 & 1 \\ 1 & 2 & 1 \end{pmatrix} \\ &= \begin{pmatrix} 0 & -2 & 0 \\ 0 & 1 & 2^n \\ 0 & 0 & -2^n \end{pmatrix} \begin{pmatrix} 1 & 2 & 2 \\ 0 & 1 & 1 \\ 1 & 2 & 1 \end{pmatrix} = \begin{pmatrix} 0 & -2 & -2 \\ 2^n & 1 + 2^{n+1} & 2^n + 1 \\ -2^n & -2^{n+1} & -2^n \end{pmatrix}. \end{aligned}$$

Méthode

On peut prendre pour e_1 n'importe quelle solution non nulle de $f(x, y, z) = (x, y, z)$. En revanche, prendre $e_1 = 0$ risquerait de poser problème plus tard : une famille contenant le vecteur nul n'est jamais libre, et donc ne peut constituer une base de \mathbb{R}^3 .

Remarque

Nous disposerons plus tard d'un argument nous dispensant de prouver la liberté de cette famille car elle est automatique (pour les cubes : il s'agit d'une famille de trois vecteurs propres associés à des valeurs propres distinctes).

Remarque

Si on choisit une autre base vérifiant les conditions de la question 1, on obtient une autre matrice P (et donc une autre matrice P^{-1}) qui convient tout à fait : il n'y a pas unicité d'une telle matrice P !

Vérification

Pour $n = 1$, on doit retrouver la matrice A !

SOLUTION DE L'EXERCICE 5.11

1. Soit f l'endomorphisme de \mathbf{R}^3 dont la matrice dans la base canonique est A . Alors si $\mathcal{B} = (e_1, e_2, e_3)$ est une base de \mathbf{R}^3 telle que $\text{Mat}_{\mathcal{B}}(f) = D$, on doit avoir

$$\text{Mat}_{\mathcal{B}}(f) = \begin{pmatrix} f(e_1) & f(e_2) & f(e_3) \\ 2 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix} \begin{matrix} e_1 \\ e_2 \\ e_3 \end{matrix}$$

Il nous faut donc trouver trois vecteurs e_1, e_2 et e_3 tels que $f(e_1) = 2e_1, f(e_2) = e_2$ et $f(e_3) = e_2 + e_3$.

Or, $f(x, y, z) = 2(x, y, z)$ si et seulement si

$$A \begin{pmatrix} x \\ y \\ z \end{pmatrix} = 2 \begin{pmatrix} x \\ y \\ z \end{pmatrix} \Leftrightarrow \begin{cases} 3x - 3y - z = 2x \\ x + y = 2y \\ -x - 3y = 2z \end{cases} \Leftrightarrow \begin{cases} x = 3y + z \\ x = y \\ x + 3y + 2z = 0 \end{cases} \Leftrightarrow \begin{cases} x = y \\ z = -2y \end{cases}$$

Ainsi, on peut prendre $e_1 = (1, 1, -2)$.

De même, la résolution de $A \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$ nous montre qu'on peut prendre $e_2 = (0, 1, -3)$.

Et alors, on a $f(x, y, z) = (x, y, z) + e_2$ si et seulement si

$$A \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x \\ y \\ z \end{pmatrix} + \begin{pmatrix} 0 \\ 1 \\ -3 \end{pmatrix} \Leftrightarrow \begin{cases} 2x - 3y - z = 0 \\ x + y = y + 1 \\ -x - 3y = z - 3 \end{cases} \Leftrightarrow \begin{cases} x = 1 \\ z = 2 - 3y \end{cases}$$

On peut par exemple prendre $e_3 = (1, 1, -1)$.

Il reste donc à vérifier que (e_1, e_2, e_3) forme bien une base de \mathbf{R}^3 . Pour cela, utilisons la matrice de (e_1, e_2, e_3) dans la base canonique, qui est

$$\text{Mat}_{\mathcal{B}_{can}}(e_1, e_2, e_3) = \begin{pmatrix} e_1 & e_2 & e_3 \\ 1 & 0 & 1 \\ 1 & 1 & 1 \\ -2 & -3 & -1 \end{pmatrix}.$$

Un calcul de pivot prouve que cette matrice est inversible, d'inverse $\begin{pmatrix} 2 & -3 & -1 \\ -1 & 1 & 0 \\ -1 & 3 & 1 \end{pmatrix}$.

En particulier, (e_1, e_2, e_3) est une base \mathcal{B} , et alors, par construction

$$\text{Mat}_{\mathcal{B}}(f) = \begin{pmatrix} f(e_1) & f(e_2) & f(e_3) \\ 2 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix} \begin{matrix} e_1 \\ e_2 \\ e_3 \end{matrix}.$$

Les matrices A et T représentant toutes les deux l'endomorphisme f dans deux bases de \mathbf{R}^3 , elles sont semblables.

Plus précisément, par la formule de changement de base,

$$A = \text{Mat}_{\mathcal{B}_{can}}(f) = P_{\mathcal{B}_{can}, \mathcal{B}} \text{Mat}_{\mathcal{B}_{can}}(f) P_{\mathcal{B}_{can}, \mathcal{B}}^{-1}.$$

De plus, $P_{\mathcal{B}_{can}, \mathcal{B}} = \begin{pmatrix} 1 & 0 & 1 \\ 1 & 1 & 1 \\ -2 & -3 & -1 \end{pmatrix}$ et $P_{\mathcal{B}_{can}, \mathcal{B}}^{-1} = \begin{pmatrix} 2 & -3 & -1 \\ -1 & 1 & 0 \\ -1 & 3 & 1 \end{pmatrix}$.

2. Notons que $T = D + N$ où $D = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ et $N = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$.

$$\text{On a alors } DN = ND = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}.$$

Puisque D et N commutent, on peut appliquer la formule du binôme de Newton :

$$T^n = (N + D)^n = \sum_{k=0}^n \binom{n}{k} N^k D^{n-k}.$$

Or on a $D^2 = 0$, et donc $D^k = 0$ pour tout $k \geq 2$. Il ne reste donc que les termes correspondants à $k = 0$ et $k = 1$ dans la somme, de sorte que

$$T^n = \binom{n}{0} N^0 D^n + \binom{n}{1} N D^{n-1} = \begin{pmatrix} 2^n & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} + n \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 2^n & 0 & 0 \\ 0 & 1 & n \\ 0 & 0 & 1 \end{pmatrix}.$$

Et alors

$$\begin{aligned} A^n &= (PTP^{-1})^n = PT^n P^{-1} \\ &= \begin{pmatrix} 1 & 0 & 1 \\ 1 & 1 & 1 \\ -2 & -3 & -1 \end{pmatrix} \begin{pmatrix} 2^n & 0 & 0 \\ 0 & 1 & n \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 2 & -3 & -1 \\ -1 & 1 & 0 \\ -1 & 3 & 1 \end{pmatrix} \\ &= \begin{pmatrix} 2^{n+1} - 1 & 3(1 - 2^n) & 1 - 2^n \\ 2^{n+1} - n - 2 & 3(n + 1 - 2^n) & n + 1 - 2^n \\ -2^{n+2} + 4 + 3n & 3(2^{n+1} - 3n - 2) & 2^{n+1} - 3n - 1 \end{pmatrix}. \end{aligned}$$

SOLUTION DE L'EXERCICE 5.12

- A et B n'ont pas même trace, donc ne peuvent être semblables.
- A et B ont même trace (2) et même rang (2). Soit donc f l'endomorphisme de \mathbf{R}^2 dont la matrice dans la base canonique est A .
Posons alors $e_1 = (1/2, 0)$ et $e_2 = (0, 1)$. La famille (e_1, e_2) est libre (car formée de deux éléments non colinéaires) et donc est une base de \mathbf{R}^2 .
On a $f(e_1) = f(1/2, 0) = \frac{1}{2}f(1, 0) = \frac{1}{2}(1, 0) = e_1$. Et $f(e_2) = f(0, 1) = (1, 1) = 2e_1 + e_2$.
Donc la matrice de f dans la base (e_1, e_2) est B .
Par conséquent, A et B représentent toutes deux f dans deux bases différentes : elles sont semblables.
- A et B ont même rang, mais n'ont pas la même trace : elles ne sont donc pas semblables.
- A et B ont même trace et même rang. Toutefois, $A = 2I_3$, et donc la seule matrice semblable à A est A . On en déduit que A et B ne sont pas semblables.

SOLUTION DE L'EXERCICE 5.13

Notons que si $\text{tr}(M) = 0$, alors $M + {}^t M = 0$, et donc M est antisymétrique.
Inversement, une matrice antisymétrique M est toujours de trace nulle, et donc on a toujours $M + {}^t M = \text{tr}(M)A$.
Donc l'ensemble $\mathcal{A}_n(\mathbf{C})$ des matrices antisymétriques est toujours inclus dans l'ensemble des solutions de l'équation.

Notons que si une matrice M est solution de l'équation, alors

$$2\text{tr}(M) = \text{tr}(M + {}^t M) = 2\text{tr}(\text{tr}(M)A) = \text{tr}(M)\text{tr}(A).$$

- Si $\text{tr}(A) \neq 2$, nécessairement $\text{tr}(M) = 0$ et donc $M + {}^t M = 0$: M est antisymétrique. Donc l'ensemble des solutions est l'ensemble $\mathcal{A}_n(\mathbf{C})$ des matrices antisymétriques.

- Si $\text{tr}(A) = 2$.

Alors, puisque $M + {}^t M$ est toujours symétrique, si M est solution, alors $\text{tr}(M)A$ est symétrique.

Si $\text{tr}(M) \neq 0$, alors c'est que A est symétrique.

⚠ Attention !

Deux matrices semblables ont même rang et même trace, mais si elles ont même rang et même trace, cela ne suffit pas à garantir qu'elles sont semblables !

Précision

Si P est la matrice de passage de la base canonique à la base (e_1, e_2) , alors, par la formule de changement de base, on a

$$B = P^{-1}AP.$$

Remarque

Nous avons dit précédemment que si M est une solution de trace nulle, alors M est nécessairement antisymétrique.

Donc si A n'est pas symétrique, il n'y a pas de solutions de trace non nulle, et donc l'ensemble des solutions est $\mathcal{S}_n(\mathbf{C})$.

Enfin, dans le cas où A est symétrique et où $\text{tr}(A) = 2$, soit M une solution de l'équation. Alors M s'écrit de manière unique $M = M_1 + M_2$, où M_1 est symétrique et M_2 est antisymétrique.

Et donc $M + {}^tM = 2M_1$, de sorte que $2M_1 = \text{tr}(M)A$, et donc $M_1 \in \text{Vect}(A)$.
Donc toute solution de l'équation est dans $\text{Vect}(A) \oplus \mathcal{S}_n(\mathbf{C})$.

Inversement, si $M = \lambda A + M_2$, avec $M_2 \in \mathcal{S}_n(\mathbf{C})$, alors $\text{tr}(M) = \lambda \text{tr}(A) = 2\lambda$. Et donc

$$M + {}^tM = 2\lambda A = 2\text{tr}(M)A.$$

Ainsi, l'ensemble des solutions de l'équation est $\text{Vect}(A) \oplus \mathcal{S}_n(\mathbf{C})$.

SOLUTION DE L'EXERCICE 5.14

Si A et B sont semblables, alors $P(A)$ et $P(B)$ sont semblables. Or, si $P(A) = 0$, alors $P(B)$ est semblable à la matrice nulle, et la seule matrice semblable à la matrice nulle est la matrice nulle. Donc $P(B) = 0$.

Inversement, si $P(B) = 0$, alors $P(A) = 0$ et donc P est annulateur de A .

SOLUTION DE L'EXERCICE 5.15

D'après le cours, tr est une forme linéaire sur $\mathcal{M}_n(\mathbf{R})$, non nulle car $\text{tr}(I_n) = n$ et $F = \text{Ker } \text{tr}$ est donc un hyperplan de $\mathcal{M}_n(\mathbf{R})$: il est de dimension $\dim \mathcal{M}_n(\mathbf{R}) - 1 = n^2 - 1$.

De plus, $A \neq 0$, et donc $\dim \text{Vect}(A) = 1$.

Soit $M \in F \cap \text{Vect}(A)$. Puisque $M \in \text{Vect}(A)$, il existe $\lambda \in \mathbf{R}$ tel que $M = \lambda A$, et alors

$$\text{tr}(M) = \lambda \text{tr}(A).$$

Mais d'autre part, si $M \in F$, alors $\text{tr}(M) = 0$, donc $\lambda = 0$.

Ainsi, $F \cap \text{Vect}(A) = \{0\}$ et par conséquent, on a

$$\begin{cases} \dim F + \dim \text{Vect}(A) = n^2 = \dim \mathcal{M}_n(\mathbf{R}) \\ F \cap \text{Vect}(A) = \{0\} \end{cases}$$

donc $\mathcal{M}_n(\mathbf{R}) = F \oplus \text{Vect}(A)$.

SOLUTION DE L'EXERCICE 5.16

1. Soit f l'endomorphisme de \mathbf{R}^n canoniquement associé à M .

Montrer que M est semblable à une matrice dont les $n-1$ premières colonnes sont nulles, c'est montrer qu'il existe une base (e_1, \dots, e_n) de \mathbf{R}^n telle que $f(e_1) = 0, f(e_2) = 0, \dots, f(e_{n-1}) = 0$.

Mais f est de rang 1, et donc $\text{Ker } f$ est de dimension $n - 1$.

Soit $D = \text{Vect}(e_n)$ un supplémentaire⁵ de $\text{Ker } f$ et (e_1, \dots, e_{n-1}) une base de $\text{Ker } f$.

Alors D et $\text{Ker } f$ sont en somme directe, de sorte que $\mathcal{B} = (e_1, \dots, e_n)$, qui est la concaténation d'une base $\text{Ker}(f)$ et d'une base de D , est une base de \mathbf{R}^n .

Si on note c_1, \dots, c_n les réels tels que $f(e_n) = \sum_{i=1}^n c_i e_i$, on a alors

$$A = \text{Mat}_{\mathcal{B}} = \begin{pmatrix} f(e_1) & \dots & f(e_{n-1}) & f(e_n) \\ 0 & \dots & 0 & c_1 \\ 0 & \dots & 0 & c_2 \\ \vdots & & \vdots & \vdots \\ 0 & \dots & 0 & c_n \end{pmatrix} \begin{matrix} e_1 \\ e_2 \\ \vdots \\ e_n \end{matrix}.$$

La matrice A possède alors bien ses $n - 1$ premières colonnes nulles, et est semblable à M .⁶ De manière plus précise, si $P \in GL_n(\mathbf{R})$ est la matrice de passage de la base canonique à la base (e_1, \dots, e_n) , alors $A = P^{-1}MP$.

Rappel

$$\mathcal{M}_n(\mathbf{C}) = \mathcal{S}_n(\mathbf{C}) \oplus \mathcal{A}_n(\mathbf{C}).$$

Remarque

On peut aussi faire le calcul en introduisant une matrice inversible Q telle que $A = Q^{-1}BQ$.

Non nulle

Dire qu'une forme linéaire est non nulle, c'est dire qu'elle ne prend pas toujours la valeur 0. Ce n'est pas pareil que de dire qu'elle ne s'annule jamais !

Remarque

On aurait également pu procéder par analyse-synthèse et prouver que tout élément de $\mathcal{M}_n(\mathbf{R})$ s'écrit de manière unique comme une matrice de trace nulle plus un multiple de A , mais c'est plus laborieux : si on a accès aux dimensions, mieux vaut raisonner comme nous venons de le faire.

⁵ Nous savons qu'il en existe.

⁶ Puisqu'elles représentent les deux l'endomorphisme f dans deux bases différentes.

2. On a $M^2 = (PAP^{-1})^2 = PA^2P^{-1}$.
Or, $\text{tr}(A) = c_n = \text{tr}(M)$.
De plus, on a

$$A^2 = \begin{pmatrix} 0 & \dots & 0 & c_1 \\ 0 & \dots & 0 & c_2 \\ \vdots & & \vdots & \vdots \\ 0 & \dots & 0 & c_n \end{pmatrix} \begin{pmatrix} 0 & \dots & 0 & c_1 \\ 0 & \dots & 0 & c_2 \\ \vdots & & \vdots & \vdots \\ 0 & \dots & 0 & c_n \end{pmatrix} = \begin{pmatrix} 0 & \dots & 0 & c_n c_1 \\ 0 & \dots & 0 & c_n c_2 \\ \vdots & & \vdots & \vdots \\ 0 & \dots & 0 & c_n c_n \end{pmatrix} = c_n A = \text{tr}(M)A.$$

Donc $M^2 = P \text{tr}(M) A P^{-1} = \text{tr}(M) \cdot M = M$.
Et donc $f^2 = f$: f est un projecteur.

SOLUTION DE L'EXERCICE 5.17

1. Si A est inversible, en multipliant l'égalité $AX = 0$ par A^{-1} , il vient $X = A^{-1}0 = 0$.
Inversement, supposons que $AX = 0 \Rightarrow X = 0$.
Soit alors f l'unique endomorphisme de \mathbf{R}^n dont la matrice dans la base canonique est A . Alors la propriété précédente s'écrit $f(x) = 0 \Rightarrow x = 0$, autrement dit $\text{Ker } f = \{0\}$ et donc f est injectif. Mais un endomorphisme de \mathbf{R}^n est injectif si et seulement si c'est un isomorphisme. On en déduit que f est un isomorphisme, et donc que A est inversible.
2. Soit donc i_0 comme dans l'énoncé. En considérant les coefficients de la $i_0^{\text{ème}}$ ligne de l'égalité $AX = X$, il vient

$$0 = \sum_{j=1}^n a_{i_0, j} x_j = 0.$$

En particulier, $a_{i_0, i_0} x_{i_0} = - \sum_{\substack{j=1 \\ j \neq i_0}}^n a_{i_0, j} x_j$.

D'après l'inégalité triangulaire, on a alors

$$|x_{i_0}| |a_{i_0, i_0}| \leq \sum_{\substack{j=1 \\ j \neq i_0}}^n |a_{i_0, j}| |x_j| \leq \sum_{\substack{j=1 \\ j \neq i_0}}^n |a_{i_0, j}| \cdot |x_{i_0}|.$$

$$\forall j \in \llbracket 1, n \rrbracket, |x_j| \leq |x_{i_0}|.$$

Si $x_{i_0} \neq 0$, alors l'inégalité ainsi obtenue devient

$$|a_{i_0, i_0}| \leq \sum_{\substack{j=1 \\ j \neq i_0}}^n |a_{i_0, j}|$$

ce qui contredit l'hypothèse de l'énoncé. On en déduit que $|x_{i_0}| = 0$, ce qui implique que $\forall j \in \llbracket 1, n \rrbracket, 0 \leq |x_j| \leq |x_{i_0}| = 0$, et donc que $X = 0$.

Nous venons donc de prouver que $AX = 0 \Rightarrow X = 0$, ce qui d'après la question précédente implique que A est inversible.

SOLUTION DE L'EXERCICE 5.18

1. On a

$$\begin{aligned} \text{tr}({}^t AA) &= \sum_{i=1}^n ({}^t AA)_{i, i} \\ &= \sum_{i=1}^n \sum_{k=1}^n ({}^t A)_{i, k} a_{k, i} \\ &= \sum_{i=1}^n \sum_{k=1}^n a_{k, i} a_{k, i} = \sum_{i=1}^n \sum_{k=1}^n a_{k, i}^2. \end{aligned}$$

2. L'application φ_A est bien à valeurs dans \mathbf{R} .
Pour $M, N \in \mathcal{M}_n(\mathbf{R})$ et $\lambda \in \mathbf{R}$, on a

$$\varphi_A(\lambda M + N) = \text{tr}({}^t A(\lambda M + N)) = \text{tr}(\lambda {}^t AM + {}^t AN) = \lambda \text{tr}({}^t AM) + \text{tr}({}^t AN) = \lambda \varphi_A(M) + \varphi_A(N).$$

Donc φ_A est bien une forme linéaire sur $\mathcal{M}_n(\mathbf{R})$.

Rappel
Deux matrices semblables ont même trace.

- 2.b. Notons $f : \mathcal{M}_n(\mathbf{R}) \rightarrow \mathcal{L}(\mathcal{M}_n(\mathbf{R}), \mathbf{R})$ l'application qui à A associe φ_A .
Il s'agit de prouver que f est linéaire. Soient donc $A, B \in \mathcal{M}_n(\mathbf{R})$ et $\lambda \in \mathbf{R}$.
Alors, pour tout $M \in \mathcal{M}_n(\mathbf{R})$, on a

$$\begin{aligned} f(\lambda A + B)(M) &= \varphi(\lambda A + B)(M) = \text{tr}({}^t(\lambda A + B)M) = \lambda \text{tr}({}^tAM) + \text{tr}({}^tBM) \\ &= \lambda \varphi_A(M) + \varphi_B(M) = \lambda f(A)(M) + f(B)(M). \end{aligned}$$

Ceci étant vrai pour tout $M \in \mathcal{M}_n(\mathbf{R})$, on a donc $f(\lambda A + B) = \lambda f(A) + f(B)$.
Donc f est bien linéaire.

Soit $A = (a_{i,j}) \in \text{Ker } f$. Alors la forme linéaire φ_A est nulle.

Et en particulier, $\varphi_A(A) = 0$. D'après la question 1, c'est donc que $\sum_{i=1}^n \sum_{j=1}^n a_{i,j}^2 = 0$.

Or, une somme de réels positifs est nulle si et seulement si chacun de ces nombres est nul, et donc $\forall (i, j) \in \llbracket 1, n \rrbracket^2, a_{i,j}^2 = 0 \Leftrightarrow a_{i,j} = 0$.

Et donc A est la matrice nulle.

Ainsi, $\text{Ker } f = \{0\}$, de sorte que f est injective.

Mais $\dim \mathcal{M}_n(\mathbf{R}) = \dim \mathcal{L}(\mathcal{M}_n(\mathbf{R}), \mathbf{R})$, donc f est bijective : c'est un isomorphisme de $\mathcal{M}_n(\mathbf{R})$ sur $\mathcal{L}(\mathcal{M}_n(\mathbf{R}), \mathbf{R})$.

- 2.c. Si $\varphi \in \mathcal{L}(\mathcal{M}_n(\mathbf{R}), \mathbf{R})$, alors φ admet un unique antécédent A par f : il existe un unique $A \in \mathcal{M}_n(\mathbf{R})$ tel que $\varphi = \varphi_A$.

SOLUTION DE L'EXERCICE 5.19

Soit f l'endomorphisme de \mathbf{R}^3 dont la matrice dans la base canonique est A .

Alors nous savons que $f \neq 0$ (car $A \neq 0$) et $f^2 = 0$ (car $A^2 = 0$).

A ne peut être inversible car $A^2 = 0$, et donc f n'est pas bijective.

En particulier, $\dim \text{Ker } f \geq 1$, et f étant non nulle, $\dim \text{Ker } f \leq 2$.

Enfin, $f \circ f = 0$, donc $\text{Im } f \subset \text{Ker } f$. Ceci impose alors $\dim \text{Im } f \leq \dim \text{Ker } f$.

Comme le théorème du rang nous indique que $\dim \text{Ker } f + \dim \text{Im } f = 3$, il n'y a pas d'autre solution que $\dim \text{Im } f = 1$ et $\dim \text{Ker } f = 2$.

Soit alors $x \in \mathbf{R}^3$ tel que $f(x) \neq 0$.

On a $f(x) \in \text{Ker } f$ car $f^2(x) = 0$. Autrement dit, la famille formée du seul vecteur $f(x)$ est une famille libre de $\text{Ker } f$, qui peut être complétée en une base $(f(x), y)$ de $\text{Ker } f$.

Enfin, $x \notin \text{Ker } f$ par hypothèse, et donc $\text{Ker } f \cap \text{Vect}(x) = \{0\}$.

On en déduit que $\mathbf{R}^3 = \text{Ker } f \oplus \text{Vect}(x)$, et donc $(f(x), y, x)$ est une base de \mathbf{R}^3 obtenue par concaténation de bases de $\text{Ker } f$ et de $\text{Vect}(x)$.

Dans cette base, la matrice de f est $\begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$, qui est donc semblable à A .

SOLUTION DE L'EXERCICE 5.20

On a donc

$$M = \begin{pmatrix} 1 & 1 & \dots & 1 & 1 \\ 1 & 2 & \dots & 2^{n-2} & 2^{n-1} \\ \vdots & \vdots & & \vdots & \vdots \\ \vdots & \vdots & & \vdots & \vdots \\ 1 & n & \dots & n^{n-2} & n^{n-1} \end{pmatrix}.$$

Soit $X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathcal{M}_{n,1}(\mathbf{R})$ tel que $MX = 0$.

Alors pour tout $i \in \llbracket 1, n \rrbracket$, en considérant le coefficient à la $i^{\text{ème}}$ ligne du produit AX , il

vient $\sum_{j=1}^n x_j i^{j-1} = 0$.

Si on note $P = \sum_{j=1}^n x_j X^{j-1} \in \mathbf{R}_{n-1}[X]$, cela signifie donc que pour tout $i \in \llbracket 1, n \rrbracket, P(i) = 0$.

Ainsi, P possède n racines distinctes, et étant de degré inférieur ou égal à $n - 1$, il est donc nul : $\forall j \in \llbracket 1, n \rrbracket, x_j = 0$.

Méthode

Deux applications sont égales, si elles coïncident en tout point de leur ensemble de définition.

C'est ce qui nous oblige à considérer ici $M \in \mathcal{M}_n(\mathbf{R})$ en plus de A et B .

Méthode

Montrer que deux matrices de $\mathcal{M}_n(\mathbf{R})$ sont semblables revient à montrer qu'elles représentent le même endomorphisme de \mathbf{R}^n dans deux bases différentes. Si, comme ici, il n'y a pas d'endomorphisme dans l'énoncé, on peut toujours en introduire un : l'endomorphisme de \mathbf{R}^n dont la matrice dans la base canonique (ou en fait n'importe quelle autre base) est A .

Notons que si $f(x) \neq 0$, nécessairement $x \neq 0$.

Donc nécessairement $X = 0$, et donc M est inversible.

Solution alternative : on peut remarquer que M est la matrice dans les bases canoniques de l'application linéaire $\varphi : \mathbf{R}_{n-1}[X] \rightarrow \mathbf{R}^n$ définie par $P \mapsto (P(1), P(2), \dots, P(n))$. En effet, on a alors, en notant (e_1, \dots, e_n) la base canonique de \mathbf{R}^n :

$$\text{Mat}(\varphi) = \begin{pmatrix} \varphi(1) & \varphi(X) & \dots & \varphi(X^{n-1}) \\ 1 & 1 & \dots & 1 \\ 1 & 2 & \dots & 2^{n-1} \\ \vdots & \vdots & & \vdots \\ 1 & n & \dots & n^{n-1} \end{pmatrix} \begin{matrix} e_1 \\ e_2 \\ \vdots \\ e_n \end{matrix} = M.$$

Mais il est classique⁷ que φ est un isomorphisme, et donc M est inversible.

SOLUTION DE L'EXERCICE 5.21

1. On a déjà $\dim \text{Vect}(I_n) + \dim F = 1 + n^2 - 1 = n^2 = \dim \mathcal{M}_n(\mathbf{R})$.

Soit $M \in \text{Vect}(I_n) \cap F$. Alors il existe $\lambda \in \mathbf{R}$ tel que $M = \lambda I_n$.

Supposons que $M \neq 0 \Leftrightarrow \lambda \neq 0$. Alors $I_n = \frac{1}{\lambda} M \in F$, contredisant les hypothèses de l'énoncé. Donc $M = 0$ et par conséquent $F \cap \text{Vect}(I_n) = \{0\}$.

On en déduit que $\mathcal{M}_n(\mathbf{R}) = F \oplus \text{Vect}(I_n)$.

- 2.a. Soient $M, M' \in \mathcal{M}_n(\mathbf{R})$.

De manière unique, on a $M = M_1 + \lambda_1 I_n$, avec $M_1 \in F$ et $\lambda_1 \in \mathbf{R}$, et de même, $M' = M_2 + \lambda_2 I_n$.

Nous savons qu'alors $p(M) = \lambda_1 I_n$ et $p(M') = \lambda_2 I_n$.

De plus,

$$MM' = (M_1 + \lambda_1 I_n)(M_2 + \lambda_2 I_n) = M_1 M_2 + \lambda_1 M_2 + \lambda_2 M_1 + \lambda_1 \lambda_2 I_n.$$

Or, puisque $M_1 \in F$ et $M_2 \in F$, $M_1 M_2 \in F$ de sorte que $\lambda_1 M_2 + \lambda_2 M_1 + M_1 M_2 \in F$. De plus $\lambda_1 \lambda_2 I_n \in \text{Vect}(I_n)$.

Donc $p(MM') = \lambda_1 \lambda_2 I_n = p(M)p(M')$.

- 2.b. Notons que $F = \text{Ker}(p)$.

Et donc $M^2 \in F \Leftrightarrow p(M^2) = 0 \Leftrightarrow p(M)p(M) = 0$.

Or, $p(M)$ est de la forme λI_n , et donc $p(M)^2 = \lambda^2 I_n$, qui est nul si et seulement si $\lambda = 0$.

Donc $p(M) = 0$, de sorte que $M \in \text{Ker}(p) = F$.

- 2.c. Pour toute matrice $A = (a_{i,j})_{1 \leq i, j \leq n}$, on a

$$(AE_{k,\ell})_{u,v} = \sum_{j=1}^n A_{u,j} (E_{k,\ell})_{j,v}.$$

Mais $(E_{k,\ell})_{j,v} = 0$, sauf si $j = k$ et $v = \ell$, auquel cas il vaut 1.

Donc déjà, si $v \neq \ell$, $(AE_{k,\ell})_{u,v} = 0$.

Et si $v = \ell$, dans la somme, seul le coefficient correspondant à $j = k$ est non nul, et on obtient donc $(AE_{k,\ell})_{u,\ell} = a_{u,k}$.

$$\text{Ainsi, } (AE_{k,\ell})_{u,v} = \begin{cases} 0 & \text{si } v \neq \ell \\ a_{u,k} & \text{si } v = \ell \end{cases}$$

$$\text{En particulier, } (E_{i,j} E_{k,\ell})_{u,v} = \begin{cases} 0 & \text{si } v \neq \ell \\ (E_{i,j})_{u,\ell} & \text{sinon} \end{cases}$$

$$\text{Mais puisque } (E_{i,j})_{u,\ell} = \begin{cases} 1 & \text{si } u = i \text{ et } j = \ell \\ 0 & \text{sinon} \end{cases} \text{ il vient}$$

$$E_{i,j} E_{k,\ell} = \begin{cases} 0 & \text{si } j \neq k \\ E_{i,\ell} & \text{si } j = k \end{cases}$$

- 2.d. Si $i \neq j$, $E_{i,j} \times E_{i,j} = 0 \in F$.

Donc d'après 2.b, $E_{i,j} \in F$.

Ce raisonnement ne s'applique plus si $i = j$, mais alors $E_{i,i} = E_{i,k} E_{k,i}$ pour tout $k \neq i$, de sorte que $E_{i,i}$ est dans F car produit de deux éléments de F .

Ainsi, tous les $E_{i,j}$, $1 \leq i, j \leq n$ sont dans F .

Rappel

Une matrice $M \in \mathcal{M}_n(\mathbf{K})$ est inversible si et seulement si pour tout $X \in \mathcal{M}_{n,1}(\mathbf{K})$,

$$MX = 0 \Rightarrow X = 0.$$

⁷ Voir par exemple l'exercice 13 du TD3.

Rappel

Si on a trouvé l'unique décomposition de MM' comme un élément de F plus un élément de $\text{Vect}(I_n)$, on a directement $p(MM')$: c'est la composante suivant $\text{Vect}(I_n)$.

Autrement dit

Multiplier A à droite par $E_{k,\ell}$ donne une matrice dont toutes les colonnes sont nulles, sauf la $\ell^{\text{ème}}$, qui est égale à la $k^{\text{ème}}$ colonne de A .

Remarque

Si l'on prend en compte la remarque ci-dessus, c'est évident : toutes les colonnes de $E_{i,j}$ sont nulles, sauf la $j^{\text{ème}}$.

Puisque $E_{i,j} E_{k,\ell}$ ne « garde » que la $k^{\text{ème}}$ colonne de $E_{i,j}$, elle est nulle si $j \neq k$, et contient un 1 en $i^{\text{ème}}$ position sinon.

- 2.e. Puisque F est un sous-espace vectoriel qui contient une base de $\mathcal{M}_n(\mathbf{R})$, il contient $\mathcal{M}_n(\mathbf{R})$ tout entier : $F = \mathcal{M}_n(\mathbf{R})$.
Ceci contredit le fait que F soit un hyperplan de $\mathcal{M}_n(\mathbf{R})$.
Par conséquent, il n'existe pas d'hyperplans de $\mathcal{M}_n(\mathbf{R})$ stables par le produit matriciel et ne contenant pas l'identité.

COUPLES DE VARIABLES ALÉATOIRES

DISCRÈTES.

Cours	175
6.1 Indépendance de variables aléatoires	175
6.2 Couples de variables aléatoires	177
6.3 Covariance, corrélation linéaire	184
Méthodes	190
Exercices	192
Correction des exercices	195

6.1 INDÉPENDANCE DE VARIABLES ALÉATOIRES

6.1.1 Indépendance d'événements

Définition 6.1 – Soit (Ω, \mathcal{A}, P) un espace probabilisé. Deux événements A et $B \in \mathcal{A}$ sont dits indépendants si

$$P(A \cap B) = P(A)P(B).$$

Remarque. Si A et B sont incompatibles avec $P(A)P(B) \neq 0$, alors A et B ne sont pas indépendants.

Exemple 6.2

On lance deux fois une pièce qui tombe sur pile avec probabilité $p \in]0, 1[$.

L'univers Ω est alors $\{P, F\}^2$.

On note A «les deux lancers ne donnent pas le même résultat», et B : «le second lancer donne face».

On a alors, $P(B) = 1 - p$, et par indépendance des lancers,

$$P(A) = P(F_1 \cap P_2) + P(P_1 \cap F_2) = P(F_1)P(P_2) + P(P_1)P(F_2) = (1-p)p + p(1-p) = 2p(1-p).$$

D'autre part, $P(A \cap B) = P(P_1 \cap F_2) = P(P_1)P(F_2) = p(1-p)$.

Donc $P(A \cap B) - P(A)P(B) = p(1-p) - 2p^2(1-p) = p(1-p)(1-2p)$, qui est nul si et seulement si $p = \frac{1}{2}$.

Donc A et B sont indépendants si et seulement si $p = \frac{1}{2}$.

Proposition 6.3 : Si $P(A) \neq 0$, alors A et B sont indépendants si et seulement si $P_A(B) = P(B)$.

Proposition 6.4 : Si A et B sont indépendants, alors A et \bar{B} sont indépendants.

Remarque

On a donc également \bar{A} et B qui sont indépendants, de même que \bar{A} et \bar{B} .

Démonstration. Par la formule des probabilités totales, on a $P(A) = P(A \cap B) + P(A \cap \bar{B})$. Et donc

$$P(A \cap \bar{B}) = P(A) - P(A \cap B) = P(A) - P(A)P(B) = P(A)(1 - P(B)) = P(A)P(\bar{B}).$$

Et donc A et \bar{B} sont indépendants. \square

6.1.2 Indépendance de variables aléatoires



Définition 6.5 – Soit X une variable aléatoire réelle sur (Ω, \mathcal{A}, P) . Alors la tribu engendrée par X , notée \mathcal{A}_X , est la plus petite tribu contenant tous les événements $[X \leq x]$, $x \in \mathbf{R}$.

En particulier, puisque $\forall x \in \mathbf{R}, [X \leq x] \in \mathcal{A}$, on a $\mathcal{A}_X \subset \mathcal{A}$.

Remarque

\mathcal{A}_X est la plus petite tribu contenant les informations utiles à l'étude de X . Elle est formée de tous les événements obtenus à partir de X . Par exemple $[1 < X \leq 2]$ et $[e^{X^2} \geq 6]$ sont dans \mathcal{A}_X .

Définition 6.6 – Soient X et Y deux variables aléatoires réelles définies sur un même espace probabilisé (Ω, \mathcal{A}, P) . Alors X et Y sont dites indépendantes si

$$\forall (x, y) \in \mathbf{R}^2, P([X \leq x] \cap [Y \leq y]) = P(X \leq x)P(Y \leq y).$$

Proposition 6.7 : Soient X et Y deux variables aléatoires réelles définies sur un même espace probabilisé (Ω, \mathcal{A}, P) . Alors il y a équivalence entre

1. X et Y sont indépendantes
2. pour tous intervalles $I, J \subset \mathbf{R}$, alors

$$P([X \in I] \cap [Y \in J]) = P(X \in I)P(Y \in J)$$

3. pour tous événements $A_1 \in \mathcal{A}_X, A_2 \in \mathcal{A}_Y$,

$$P(A_1 \cap A_2) = P(A_1)P(A_2).$$

En pratique

On n'aura jamais besoin de revenir aux tribus. Il faut surtout retenir que l'indépendance de X et de Y signifie que tout événement formé à partir de X est indépendant de tout événement formé à partir de Y .

Exemple 6.8

On lance deux dés, on note X la variable aléatoire somme des deux dés, et Y la variable aléatoire égale au produit des deux dés.

Si $[Y = 6]$, alors l'une des variables vaut 2 et l'autre vaut 3 ou l'une vaut 1 et l'autre vaut 6. Donc $[X = 5]$ ou $[X = 7]$.

Par conséquent, $P([X = 6] \cap [Y = 6]) = 0 \neq P(X = 6)P(Y = 6)$.

Et donc X et Y ne sont indépendantes.

Méthode

Pour prouver que deux variables ne sont pas indépendantes, le plus simple est souvent de trouver deux événements (l'un dépendant de la première variable, l'autre de la seconde), de probabilité non nulle et incompatibles.

6.1.3 Espérance conditionnelle dans le cas de l'indépendance



Proposition 6.9 : Soient X et Y deux variables aléatoires indépendantes sur (Ω, \mathcal{A}, P) , avec X discrète, et soit $A \in \mathcal{A}_Y$ est tel que $P(A) \neq 0$. Alors X admet une espérance si et seulement si $E(X|A)$ existe, et alors

$$E(X) = E(X|A).$$

Démonstration. Notons $X(\Omega) = \{x_k, k \in \mathbf{N}\}$.

Pour tout k , on a $x_k P(X = x_k) = x_k P_A(X = x_k)$, et donc $\sum_k x_k P(X = x_k)$ converge absolument si et seulement si $\sum_k P_A(X = x_k)$ converge absolument et alors

$$E(X) = \sum_{k=0}^{+\infty} x_k P(X = x_k) = \sum_{k=0}^{+\infty} x_k P_A(X = x_k) = E(X|A).$$

\square

Remarque. En particulier, ceci est valable pour A de la forme $[Y \leq x]$ ou $[Y = x]$.

Explication

X est indépendante de Y , et $A \in \mathcal{A}_Y$ signifie que A est un événement formé à l'aide de Y . Savoir qu'il est réalisé n'influe donc pas sur la loi de X .

Exemple 6.10

Prenons un exemple simple : on lance deux dés, et on note X (resp. Y) le résultat du premier (resp. du second) dé.

Il est évident que savoir que le premier dé a fait 1 ne change rien à la valeur moyenne du second.

Autrement dit, $E(Y|[X = 1]) = E(Y)$ car X et Y sont indépendantes.

En revanche, si on connaît la somme des deux dés, c'est-à-dire la valeur de $X + Y$ (qui n'est pas indépendante de Y), alors cela influe la valeur moyenne du second.

En particulier, on a $E(Y|[X + Y = 2]) = 1 \neq E(Y)$.

Détails

Il n'y a pas de calcul à faire pour en arriver là : si $X + Y = 2$, alors $X = Y = 1$.

6.2 COUPLES DE VARIABLES ALÉATOIRES**6.2.1 Définition**

Définition 6.11 – Si X, Y sont deux variables aléatoires réelles sur un même espace probabilisé (Ω, \mathcal{A}, P) , on dit que (X, Y) est un couple de variables aléatoires sur (Ω, \mathcal{A}, P) .

Définition 6.12 – Si (X, Y) est un couple de variables aléatoires réelles sur (Ω, \mathcal{A}, P) , on appelle loi du couple (X, Y) la fonction

$$\begin{aligned} F_{(X,Y)} : \mathbf{R}^2 &\rightarrow \mathbf{R} \\ (x, y) &\mapsto P([X \leq x] \cap [Y \leq y]) \end{aligned}$$

$F_{(X,Y)}$ est en quelque sorte l'analogue de la fonction de répartition, mais pour un couple de variables (et non pour une seule variable).

Remarque. X et Y sont indépendantes si et seulement si

$$\forall (x, y) \in \mathbf{R}^2, F_{(X,Y)}(x, y) = F_X(x)F_Y(y).$$

6.2.2 Cas des variables aléatoires discrètes

On suppose à présent que X et Y sont deux variables aléatoires discrètes, avec

$$X(\Omega) = \{x_i, i \in \mathbf{N}\} \text{ et } Y(\Omega) = \{y_j, j \in \mathbf{N}\}.$$

Alors la loi du couple (X, Y) est entièrement déterminée par la donnée de l'ensemble des

$$p_{i,j} = P([X = x_i] \cap [Y = y_j]), (i, j) \in \mathbf{N} \times \mathbf{N}.$$

On dit alors que la donnée de ces réels est la **loi conjointe** du couple (X, Y) , que la loi de X est la **première loi marginale** du couple (X, Y) , et que la loi de Y est la **seconde loi marginale** du couple (X, Y) .

Remarques. • En particulier, connaissant la loi du couple (X, Y) , on retrouve la loi de X grâce à la formule des probabilités totales appliquée au système complet d'événements $\{[Y = y], y \in Y(\Omega)\}$:

$$\forall x \in X(\Omega), P(X = x) = \sum_{y \in Y(\Omega)} P([X = x] \cap [Y = y]).$$

De même, la loi de Y s'obtient à l'aide du système complet d'événements $\{[X = x], x \in X(\Omega)\}$. Ainsi, la donnée de la loi conjointe permet de retrouver les lois marginales, mais la réciproque est fautive !

• $\{[X = x_i] \cap [Y = y_j], (i, j) \in \mathbf{N} \times \mathbf{N}\}$ est un système complet d'événements.



Proposition 6.13 : Une suite double $(p_{i,j})_{i \in I, j \in J}$ avec I et J dénombrables est la loi conjointe d'un couple de variables aléatoires discrètes si et seulement si

- 1) $\forall (i, j) \in I \times J, p_{i,j} \geq 0$
- 2) la famille $(p_{i,j})$ est sommable et $\sum_{(i,j) \in I \times J} p_{i,j} = 1$

Remarque

Cette proposition possède un analogue pour une seule variable : si I est un ensemble dénombrable, et si $(p_i)_{i \in I}$ est une famille de réels, alors il existe une variable aléatoire X telle que

$$\forall i \in I, P(X = i) = p_i$$

si et seulement si :

- $\forall i \in I, p_i \geq 0$
- $(p_i)_{i \in I}$ est sommable, de somme 1 (ce dernier point étant la traduction de «la somme des probabilités vaut toujours 1»).

Exemple 6.14

Soit $p \in]0, 1[$, et soit la famille $(p_{m,n})_{(m,n) \in \mathbf{N} \times \mathbf{N}}$ définie par

$$\forall (m, n) \in \mathbf{N} \times \mathbf{N}, p_{m,n} = \begin{cases} \frac{e^{-2}}{m!(n-m)!} & \text{si } n \geq m \\ 0 & \text{sinon} \end{cases}$$

Alors, il est clair que les $p_{m,n}$ sont positifs.

Pour n fixé, on a $\sum_{m=0}^{+\infty} p_{m,n} = \sum_{m=0}^n p_{m,n}$ qui converge et qui vaut

$$\sum_{m=0}^n \frac{e^{-2}}{m!(n-m)!} = \frac{e^{-2}}{n!} \sum_{m=0}^n \binom{n}{m} = \frac{e^{-2}}{n!} 2^n.$$

Et alors

$$\sum_{n=0}^{+\infty} \sum_{m=0}^{+\infty} p_{m,n} = \sum_{n=0}^{+\infty} \frac{e^{-2} 2^n}{n!} = e^{-2} \sum_{n=0}^{+\infty} \frac{2^n}{n!} = e^{-2} e^2 = 1.$$

Donc la famille $(p_{m,n})$ est sommable et de somme 1. Elle définit donc bien la loi d'un couple de variables aléatoires.

Soit (X, Y) un couple de variables aléatoires à valeurs dans \mathbf{N} telles que

$$\forall (m, n) \in \mathbf{N} \times \mathbf{N}, P([X = m] \cap [Y = n]) = p_{m,n}.$$

Alors la loi de X , la première loi marginale de ce couple est donnée par

$$\begin{aligned} P(X = m) &= \sum_{n=0}^{+\infty} P([X = m] \cap [Y = n]) \\ &= \sum_{n=m}^{+\infty} \frac{e^{-2}}{m!(n-m)!} \\ &= \frac{e^{-2}}{m!} \sum_{n=m}^{+\infty} \frac{1}{(n-m)!} \\ &= \frac{e^{-2}}{m!} \sum_{i=0}^{+\infty} \frac{1}{i!} \\ &= \frac{e^{-2}}{m!} e = \frac{e^{-1}}{m!}. \end{aligned}$$

Détails

C'est la formule des probabilités totales appliquées à $\{[Y = n], n \in \mathbf{N}\}$.

$$p_{m,n} = 0 \text{ pour } n < m.$$

Chgt d'indice

$$i = n - m.$$

Ainsi, X suit une loi de Poisson de paramètre 1.

De même, pour $m \in \mathbf{N}$, on a

$$P(Y = n) = \sum_{m=0}^{+\infty} p_{m,n} = \sum_{m=0}^n \frac{e^{-2}}{m!(n-m)!} = \frac{e^{-2} 2^n}{n!}.$$

Ces sommes ont déjà été calculées précédemment.

Et donc $Y \hookrightarrow \mathcal{P}(2)$.

De plus, X et Y ne sont pas indépendantes car

$$P([X = 1] \cap [Y = 2]) = 0 \neq P(X = 1)P(Y = 2).$$

Proposition 6.15 : X et Y sont indépendantes si et seulement

$$\forall (i, j) \in \mathbf{N} \times \mathbf{N}, P([X = x_i] \cap [Y = y_j]) = P(X = x_i)P(Y = y_j).$$

Exemple 6.16

Soient X et Y deux variables aléatoires suivant une loi de Bernoulli. Alors $[X = 0] = \overline{[X = 1]}$ et $[Y = 0] = \overline{[Y = 1]}$. Donc si $[X = 1]$ et $[Y = 1]$ sont indépendants, alors pour tout $(i, j) \in \{0, 1\}^2$, $[X = i]$ et $[Y = j]$ sont indépendants. Et donc X et Y sont indépendantes. Ainsi, X et Y sont indépendantes si et seulement si $[X = 1]$ et $[Y = 1]$ sont indépendants.

Détails
C'est la proposition 6.4.



Définition 6.17 – Si (X, Y) est un couple de variables aléatoires discrètes, alors la tribu engendrée par (X, Y) , notée $\mathcal{A}_{(X, Y)}$ est la plus petite tribu contenant tous les $[X = x_i] \cap [Y = y_j]$.

6.2.3 Fonction d'un couple de variables aléatoires discrètes.

Proposition 6.18 : Si (X, Y) est un couple de variables aléatoires discrètes et si

$$\begin{aligned} g : X(\Omega) \times Y(\Omega) &\rightarrow \mathbf{R} \\ (x, y) &\mapsto g(x, y) \end{aligned}$$

est une fonction, alors $g(X, Y)$ est une variable aléatoire discrète sur (Ω, \mathcal{A}) .

Démonstration. On admettra que $g(X, Y)$ est une variable aléatoire, montrons uniquement qu'elle est discrète.

On a $g(X, Y)(\Omega) = \{g(x_i, y_j), (i, j) \in I \times J\}$, qui est dénombrable car $I \times J$ l'est. □

En particulier, si X et Y sont deux variables aléatoires discrètes, alors $X+Y, XY, \min(X, Y), \max(X, Y)$ sont encore des variables aléatoires discrètes.



Remarque. Si $Z = g(X, Y)$, on a alors $\mathcal{A}_Z \subset \mathcal{A}_{(X, Y)}$. En effet, si $z \in Z(\Omega)$, alors

$$\{(x, y) \in X(\Omega) \times Y(\Omega) : g(x, y) = z\}$$

est dénombrable, et donc

$$[Z = z] = \bigcup_{(x, y) \in g^{-1}(z)} ([X = x] \cap [Y = y]) \in \mathcal{A}_{(X, Y)}$$

car $\mathcal{A}_{(X, Y)}$ est stable par union dénombrable. Donc tous les $[Z = z]$ sont dans $\mathcal{A}_{(X, Y)}$. Mais \mathcal{A}_Z est la plus petite tribu contenant tous les $[Z = z]$, donc $\mathcal{A}_Z \subset \mathcal{A}_{(X, Y)}$.

Ce résultat signifie en fait qu'un événement «formé à partir de Z » peut aussi être «formé à partir de X et de Y .»

Proposition 6.19 : Soit (X, Y) un couple de variables aléatoires discrètes sur (Ω, \mathcal{A}) et $g : X(\Omega) \times Y(\Omega) \rightarrow \mathbf{R}$. Alors, en notant $Z = g(X, Y)$, la loi de Z est donnée par

$$\forall z \in Z(\Omega), P(Z = z) = \sum_{\substack{(x, y) \in X(\Omega) \times Y(\Omega) \\ z = g(x, y)}} P([X = x] \cap [Y = y]).$$

Explication
On regarde les différentes possibilités d'avoir $g(x, y) = z$, et on somme les probabilités associées.

Remarque. Il se peut que cette somme soit infinie, mais alors sa convergence est automatique.

Démonstration. Par la formule des probabilités totales appliquée au système complet d'événements $\{[X = x] \cap [Y = y], x \in X(\Omega), y \in Y(\Omega)\}$, on a

$$P(Z = z) = \sum_{(x,y) \in X(\Omega) \times Y(\Omega)} P([Z = z] \cap [X = x] \cap [Y = y]).$$

Mais si $g(x, y) \neq z$, alors on ne peut avoir à la fois $[X = x]$, $[Y = y]$ et $[Z = z]$, puisque $Z = g(X, Y) = g(x, y)$.

Et donc dans ce cas, $P([Z = z] \cap [X = x] \cap [Y = y]) = 0$.

Au contraire, si $g(x, y) = z$, alors on a $[X = x] \cap [Y = y] \subset [Z = z]$ et donc

$$P([Z = z] \cap [X = x] \cap [Y = y]) = P([X = x] \cap [Y = y]).$$

Et donc il reste

$$P(Z = z) = \sum_{\substack{(x,y) \in X(\Omega) \times Y(\Omega) \\ g(x,y)=z}} P([X = x] \cap [Y = y]).$$

□

Détails

Si $A \subset B$, alors $A \cap B = A$.

Corollaire 6.20 (Théorème de transfert) – Sous les hypothèses précédentes, alors $Z = g(X, Y)$ admet une espérance si et seulement si la famille $(g(x, y)P([X = x] \cap [Y = y]))_{(x,y) \in X(\Omega) \times Y(\Omega)}$ est sommable, et alors

$$E(Z) = \sum_{\substack{x \in X(\Omega) \\ y \in Y(\Omega)}} g(x, y)P([X = x] \cap [Y = y]).$$

Démonstration. Admis. □

Exemple 6.21

Une urne contient n boules numérotées de 1 à n .

On tire deux boules sans remise, et on note X le numéro de la première boule tirée, et Y le numéro de la seconde. On pose alors $Z = \min(X, Y)$ le numéro de la plus petite des deux boules.

Alors, par le théorème de transfert, $E(Z) = \sum_{k=1}^n \sum_{\ell=1}^n \min(k, \ell)P([X = k] \cap [Y = \ell])$.

De plus, il y a $n(n-1)$ tirages possibles, dont un seul réalise $[X = k] \cap [Y = \ell]$, si $k \neq \ell$.

Autrement dit, $P([X = k] \cap [Y = \ell]) = \begin{cases} \frac{1}{n(n-1)} & \text{si } k \neq \ell \\ 0 & \text{sinon} \end{cases}$.

Et donc

$$E(Z) = \sum_{k=1}^n \sum_{\substack{\ell=1 \\ \ell \neq k}}^n \frac{\min(k, \ell)}{n(n-1)}.$$

Or on a, à k fixé,

$$\sum_{\substack{\ell=1 \\ \ell \neq k}}^n \min(k, \ell) = \sum_{\ell=1}^{k-1} \ell + \sum_{\ell=k+1}^n k = \frac{k(k-1)}{2} + k(n-k).$$

Et donc

$$\begin{aligned} \sum_{k=1}^n \sum_{\substack{\ell=1 \\ \ell \neq k}}^n \min(k, \ell) &= \sum_{k=1}^n \left(\sum_{\ell=1}^{k-1} \min(k, \ell) + \sum_{\ell=k+1}^n \min(k, \ell) \right) \\ &= \sum_{k=1}^n \left(\sum_{\ell=1}^{k-1} \ell + \sum_{\ell=k+1}^n k \right) \end{aligned}$$

$$\begin{aligned}
&= \sum_{k=1}^n \left(\frac{k(k-1)}{2} + k(n-k) \right) \\
&= -\frac{1}{2} \sum_{k=1}^n k^2 - \frac{1}{2} \sum_{k=1}^n k + n \sum_{k=1}^n k \\
&= \frac{n^2(n+1)}{2} - \frac{n(n+1)(2n+1)}{12} - \frac{n(n+1)}{4} = \frac{n(n+1)(n-1)}{3}.
\end{aligned}$$

On en déduit que $E(Z) = \frac{n+1}{3}$.

6.2.4 Somme de deux variables indépendantes : le produit de convolution

Proposition 6.22 (Produit de convolution discret) : Si X et Y sont deux variables aléatoires sur (Ω, \mathcal{A}, P) , alors la loi de $Z = X + Y$ est donnée par

$$\forall z \in (X + Y)(\Omega), P(Z = z) = \sum_{x \in X(\Omega)} P([X = x] \cap [Y = z - x]).$$

Démonstration. Appliquons la formule des probabilités totales au système complet d'événements $\{[X = x], x \in X(\Omega)\}$. Alors

$$P(Z = z) = \sum_{x \in X(\Omega)} P([Z = z] \cap [X = x]).$$

Mais $[Z = z] \cap [X = x] = [X + Y = z] \cap [X = x] = [X = x] \cap [Y = z - x]$.
Notons que si jamais $z - x$ n'appartient pas au support de Y , alors $P([X = x] \cap [Y = z - x]) = 0$.
On peut donc enlever un certain nombre de termes de la somme. \square

Corollaire 6.23 : Si X et Y sont deux variables aléatoires indépendantes, à valeurs dans \mathbf{N} , alors

$$\forall n \in \mathbf{N}, P(X + Y = n) = \sum_{k=0}^n P(X = k)P(Y = n - k).$$

Démonstration. D'après la proposition précédente, on a

$$\forall n \in \mathbf{N}, P(X + Y = n) = \sum_{k=0}^{+\infty} P([X = k] \cap [Y = n - k]).$$

Par indépendance de X et Y , on a $P([X = k] \cap [Y = n - k]) = P(X = k)P(Y = n - k)$.
De plus, si $k > n$, alors $n - k \leq 0$, et Y étant à valeurs dans \mathbf{N} , $P(Y = n - k) = 0$. Et donc

$$P(X + Y = n) = \sum_{k=0}^n P(X = k)P(Y = n - k).$$

\square

Exemple 6.24

On lance deux dés à 6 faces de manière indépendante, et on note X (resp. Y) le résultat du premier (resp. du deuxième) lancer.
Alors $X \hookrightarrow \mathcal{U}(\llbracket 1, 6 \rrbracket)$ et $Y \hookrightarrow \mathcal{U}(\llbracket 1, 6 \rrbracket)$.

Bornes

Il ne s'agit pas d'apprendre cette formule par cœur, en tous cas en ce qui concerne les bornes de sommation. De toutes façons, dans 99,9% des cas, c'est le corollaire suivant qui servira. Et dans les 0,1% restants, fiez-vous à votre intuition !

Détail

Si on connaît les valeurs de X et de $X + Y$, alors on connaît celle de Y .

De plus, $(X + Y)(\Omega) = \llbracket 2, 12 \rrbracket$.

$$\forall \ell \in \llbracket 2, 12 \rrbracket, P(X + Y = \ell) = \sum_{k=0}^{\ell} P(X = k)P(Y = \ell - k).$$

$$\text{Mais } P(X = k) = \begin{cases} \frac{1}{6} & \text{si } 1 \leq k \leq 6 \\ 0 & \text{sinon} \end{cases} \text{ et } P(Y = \ell - k) = \begin{cases} \frac{1}{6} & \text{si } 1 \leq \ell - k \leq 6 \\ 0 & \text{sinon} \end{cases}.$$

On peut donc se contenter de sommer sur l'ensemble des k tels que $P(X = k) \neq 0$ et $P(Y = \ell - k) \neq 0$, c'est-à-dire

$$\begin{cases} 1 \leq k \leq 6 \\ 1 \leq \ell - k \leq 6 \end{cases} \Leftrightarrow \begin{cases} 1 \leq k \leq 6 \\ \ell - 6 \leq k \leq \ell - 1 \end{cases} \Leftrightarrow \max(1, \ell - 6) \leq k \leq \min(6, \ell - 1)$$

Donc

$$P(X + Y = \ell) = \frac{1}{36} (\min(\ell - 1, 6) - \max(1, \ell - 6) + 1)$$

Pour obtenir des formules un peu plus explicites, on peut remarquer que si $\ell \in \llbracket 2, 7 \rrbracket$, alors $\min(\ell - 1, 6) = \ell - 1$ et $\max(1, \ell - 6) = 1$, de sorte que

$$P(X + Y = \ell) = \frac{1}{36} (\ell - 1 - 1 + 1) = \frac{\ell - 1}{36}.$$

Et si $\ell \in \llbracket 7, 12 \rrbracket$, alors $\min(\ell - 1, 6) = 6$ et $\max(1, \ell - 6) = \ell - 6$, de sorte que

$$P(X + Y = \ell) = \frac{1}{36} (6 - (\ell - 6) + 1) = \frac{13 - \ell}{36}.$$

Par exemple, $P(X + Y = 5) = \frac{4}{36}$ et $P(X + Y = 12) = \frac{1}{36}$.

⚠ Attention !

Beaucoup reste à faire une fois qu'on a écrit cette formule, car la somme peut encore contenir bon nombre de termes nuls. Par exemple, pour $k = 0$, puisque X ne peut prendre la valeur 0, on a $P(X = 0) = 0$. Et donc il n'est pas question pour l'instant de remplacer toutes les probas par $\frac{1}{6}$!

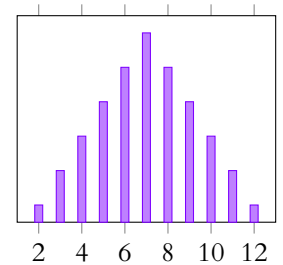


FIGURE 6.1— Loi de la somme de deux uniformes indépendantes. Cette loi est appelée loi triangulaire

6.2.5 Minimum et maximum de deux variables

Si X et Y sont deux variables aléatoires discrètes, alors $Z = \max(X, Y)$ aussi et $Z(\Omega) \subset X(\Omega) \cup Y(\Omega)$. De plus, $\forall k \in \mathbf{R}$, on a

$$[Z \leq k] = [X \leq k] \cap [Y \leq k]$$

En effet, $\forall \omega \in \Omega$

$$\max(X(\omega), Y(\omega)) \leq k \Leftrightarrow X(\omega) \leq k \text{ et } Y(\omega) \leq k \Leftrightarrow \omega \in [X \leq k] \cap [Y \leq k]$$

Ainsi, pour déterminer la loi de $\max(X, Y)$, en l'absence d'indications explicites de la part de l'énoncé, on utilisera la fonction de répartition.

Exemple 6.25

Soient X et Y deux variables aléatoires discrètes indépendantes, et $Z = \max(X, Y)$. Alors, $\forall x \in \mathbf{R}$,

$$F_Z(x) = P(Z \leq x) = P([X \leq x] \cap [Y \leq x]) = P(X \leq x)P(Y \leq x) = F_X(x)F_Y(x).$$

Par exemple, si X et Y sont indépendantes, suivant toutes deux la loi uniforme sur $\llbracket 1, n \rrbracket$, alors, $\forall k \in \mathbf{N}$,

$$F_X(k) = F_Y(k) = P(X \leq k) = \sum_{i=1}^k P(X = i) = \begin{cases} 0 & \text{si } k \leq 0 \\ \frac{k}{n} & \text{si } 0 \leq k \leq n \\ 1 & \text{sinon} \end{cases}$$

En effet

Être plus grand que $\min(x, y)$, c'est être à la fois plus grand que x et que y .

$$\text{et donc } F_Z(k) = P(Z \leq k) = \begin{cases} 0 & \text{si } k \leq 0 \\ \frac{k^2}{n^2} & \text{si } 0 \leq k \leq n \\ 1 & \text{sinon} \end{cases}$$

En particulier, pour $k \in \llbracket 1, n \rrbracket$, il vient

$$P(Z = k) = P(Z \leq k) - P(Z \leq k - 1) = \frac{k^2}{n^2} - \frac{(k - 1)^2}{n^2} = \frac{2k - 1}{n^2}.$$

Si X et Y sont deux variables aléatoires discrètes, alors $Z = \min(X, Y)$ est encore une variable aléatoire discrète avec $Z(\Omega) \subset X(\Omega) \cup Y(\Omega)$. De plus, pour $k \in \mathbf{R}$, on a

$$[Z \geq k] = [X \geq k] \cap [Y \geq k]$$

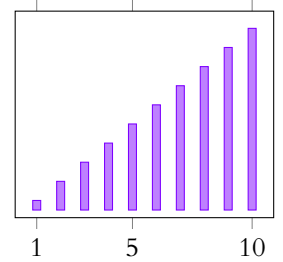


FIGURE 6.2– Loi du maximum de deux uniformes sur $\llbracket 1, 10 \rrbracket$

6.2.6 Stabilité des lois binomiales et de Poisson

Lemme 6.26 (Formule de Vandermonde). $\forall n_1, n_2 \in \mathbf{N}, \forall \ell \in \mathbf{N}$, on a

$$\sum_{k=0}^{\ell} \binom{n_1}{k} \binom{n_2}{\ell - k} = \binom{n_1 + n_2}{\ell}.$$

Remarque. Notons que si $\ell > n_1$, alors pour $k > n_1$, on a $\binom{n_1}{k} = 0$, et donc il est possible de retirer les derniers termes de la somme.

De même, si $\ell \geq \ell_2$, alors pour $k < \ell - n_2 \Leftrightarrow \ell - k > n_2$, $\binom{n_2}{\ell - k} = 0$, et donc il est possible de retirer les premiers termes de la somme.

Démonstration. Première preuve : par dénombrement.

Considérons une urne dans laquelle se trouvent n_1 boules blanches et n_2 boules rouges. Alors il y a $\binom{n_1 + n_2}{\ell}$ manières de choisir ℓ boules dans cette urne. Choisir ℓ boules, c'est en choisir k blanches et $\ell - k$ rouges, avec $0 \leq k \leq \ell$.

Or, il y a $\binom{n_1}{k}$ manières de choisir k boules blanches, et pour chaque choix de k boules blanches, il y a $\binom{n_2}{\ell - k}$ manières d'en choisir $\ell - k$ manières de choisir $\ell - k$ boules noires.

Donc au total, il y a $\sum_{k=0}^{n_1} \binom{n_1}{k} \binom{n_2}{\ell - k}$ manières de choisir ℓ boules. On en déduit que

$$\sum_{k=0}^{n_1} \binom{n_1}{k} \binom{n_2}{\ell - k} = \binom{n_1 + n_2}{\ell}.$$

Seconde preuve : avec des polynômes.

On a $(1 + X)^{n_1} (1 + X)^{n_2} = (1 + X)^{n_1 + n_2}$.

Mais, d'après la formule du binôme de Newton, on a

$$\begin{aligned} (1 + X)^{n_1} &= \sum_{k=0}^{n_1} \binom{n_1}{k} X^k \\ (1 + X)^{n_2} &= \sum_{k=0}^{n_2} \binom{n_2}{k} X^k \\ (1 + X)^{n_1 + n_2} &= \sum_{k=0}^{n_1 + n_2} \binom{n_1 + n_2}{k} X^k. \end{aligned}$$

Et donc

$$(1 + X)^{n_1} (1 + X)^{n_2} = \sum_{k=0}^{n_1 + n_2} \left(\sum_{\ell=0}^{n_1} \binom{n_1}{\ell} \binom{n_2}{k - \ell} \right) X^k.$$

Ainsi, par identification des coefficients, pour tout $k \leq n_1 + n_2$, on a

$$\binom{n_1 + n_2}{k} = \sum_{\ell=0}^{n_1} \binom{n_1}{\ell} \binom{n_2}{k - \ell}.$$

□

Hors programme

Cette formule nous sera nécessaire dans la preuve de la proposition 6.27, mais n'est pas au programme, il n'est donc pas nécessaire de la connaître.

Détail

C'est la **définition** du produit de deux polynômes !

Proposition 6.27 (Stabilité des lois binomiales) : Soit $p \in [0, 1]$, et X_1, X_2 deux variables aléatoires réelles indépendantes sur (Ω, \mathcal{A}, P) telles que $X_1 \hookrightarrow \mathcal{B}(n_1, p), X_2 \hookrightarrow \mathcal{B}(n_2, p)$.
Alors $X_1 + X_2 \hookrightarrow \mathcal{B}(n_1 + n_2, p)$.

Hypothèses

Il est important que les variables soient indépendantes, et que le p soit le même pour les deux variables.

Démonstration. On sait que $X_1(\Omega) = \llbracket 0, n_1 \rrbracket$ et $X_2(\Omega) = \llbracket 0, n_2 \rrbracket$.

Donc $(X_1 + X_2)(\Omega) \subset \llbracket 0, n_1 + n_2 \rrbracket$. De plus, $\forall \ell \in \llbracket 0, n_1 + n_2 \rrbracket$,

$$\begin{aligned} P(X_1 + X_2 = \ell) &= \sum_{k=0}^{n_1} P(X = k)P(Y = \ell - k) \\ &= \sum_{k=0}^{n_1} \binom{n_1}{k} p^k (1-p)^{n_1-k} \binom{n_2}{\ell-k} p^{\ell-k} (1-p)^{n_2-(\ell-k)} \\ &= p^\ell (1-p)^{n_1+n_2-\ell} \sum_{k=0}^{n_1} \binom{n_1}{k} \binom{n_2}{\ell-k} \\ &= p^\ell (1-p)^{n_1+n_2-\ell} \binom{n_1+n_2}{\ell} \end{aligned}$$

ce qui prouve bien le résultat annoncé. \square

Proposition 6.28 (Stabilité des lois de Poisson) : Soient X, Y deux variables aléatoires indépendantes sur (Ω, \mathcal{A}, P) , telles que $X \hookrightarrow \mathcal{P}(\lambda)$ et $Y \hookrightarrow \mathcal{P}(\mu)$. Alors

$$X + Y \hookrightarrow \mathcal{P}(\lambda + \mu).$$

Hypothèses

On n'oublie pas l'indépendance, sans laquelle cette proposition n'est plus vraie.

Démonstration. On a $(X + Y)(\Omega) = \mathbf{N}$. De plus, pour $\ell \in \mathbf{N}$,

$$\begin{aligned} P(X + Y = \ell) &= \sum_{k=0}^{\ell} P(X = k)P(Y = \ell - k) = \sum_{k=0}^{\ell} e^{-\lambda} \frac{\lambda^k}{k!} e^{-\mu} \frac{\mu^{\ell-k}}{(\ell-k)!} \\ &= e^{-(\lambda+\mu)} \sum_{k=0}^{\ell} \frac{1}{k!(\ell-k)!} \lambda^k \mu^{\ell-k} = e^{-(\lambda+\mu)} \sum_{k=0}^{\ell} \frac{1}{\ell!} \binom{\ell}{k} \lambda^k \mu^{\ell-k} \\ &= e^{-(\lambda+\mu)} \frac{1}{\ell!} \sum_{k=0}^{\ell} \binom{\ell}{k} \lambda^k \mu^{\ell-k} = e^{-(\lambda+\mu)} \frac{1}{\ell!} (\lambda + \mu)^\ell \end{aligned}$$

On reconnaît là une loi de Poisson de paramètre $\lambda + \mu$. \square



Nous ne donnons pas de résultat concernant la stabilité des lois géométriques, et pour cause : la somme de deux lois géométriques indépendantes ne suit pas une loi géométrique !

Une bonne raison à cela est que si $X_1 \hookrightarrow \mathcal{G}(p_1)$ et $X_2 \hookrightarrow \mathcal{G}(p_2)$, alors $X_1 + X_2$ ne prend que des valeurs supérieures ou égales à 2. Et donc ne peut suivre une loi géométrique, puisque son support n'est pas \mathbf{N}^* .

En revanche

On peut quand même déterminer la loi de la somme de deux variables géométriques indépendantes à l'aide du produit de convolution. Cf exercice 6.2.

6.3 COVARIANCE, CORRÉLATION LINÉAIRE

6.3.1 Espérance d'un produit

Proposition 6.29 : Si X et Y sont deux variables aléatoires réelles discrètes indépendantes sur (Ω, \mathcal{A}, P) , admettant une espérance, alors

$$E(XY) = E(X)E(Y).$$



Démonstration. On sait par le théorème de transfert¹ que

$$\begin{aligned} E(XY) &= \sum_{(x,y) \in X(\Omega) \times Y(\Omega)} xyP([X = x] \cap [Y = y]) \\ &= \sum_{(x,y) \in X(\Omega) \times Y(\Omega)} xyP(X = x)P(Y = y) \\ &= \sum_{x \in X(\Omega)} xP(X = x) \sum_{y \in Y(\Omega)} yP(Y = y) \\ &= \sum_{x \in X(\Omega)} xP(X = x)E(Y) = E(X)E(Y) \end{aligned}$$

□

Sans l'hypothèse d'indépendance, ce résultat n'est plus vrai. Toutefois, nous pouvons tout de même garantir l'existence de $E(XY)$ lorsque les deux variables X et Y admettent une variance.

Lemme 6.30. Soient $x, y \in \mathbf{R}$. Alors

$$|xy| \leq \frac{1}{2}(x^2 + y^2).$$

Démonstration. On a $(|x| - |y|)^2 \geq 0$.
 Mais $(|x| - |y|)^2 = |x|^2 - 2|x||y| + |y|^2 = x^2 - 2|xy| + y^2$.
 Et donc

$$2|xy| \leq x^2 + y^2 \Leftrightarrow |xy| \leq \frac{1}{2}(x^2 + y^2).$$

□

¹ En vérité, il faudrait prouver la sommabilité de cette famille. Mais si on remplace les x par des $|x|$ et les y par des $|y|$, alors le même calcul prouve que la famille est bien sommable.

Classique !
 Pas forcément à connaître par cœur, mais à savoir retrouver (et redémontrer) au besoin.

Proposition 6.31 : Soient X, Y deux variables aléatoires discrètes sur (Ω, \mathcal{A}, P) admettant une variance. Alors XY admet une espérance, et

$$E(XY) = \sum_{\substack{x \in X(\Omega) \\ y \in Y(\Omega)}} xyP([X = x] \cap [Y = y]).$$

Démonstration. Commençons par prouver que l'espérance existe : on a

$$|XY| \leq \frac{1}{2}(X^2 + Y^2)$$

Mais X^2 et Y^2 admettent toutes deux une espérance car X et Y admettent une variance, donc $E(XY)$ existe par domination.
 De plus, par le théorème de transfert appliqué à la fonction $g(x, y) = xy$, cette espérance vaut

$$E(XY) = \sum_{(x,y) \in X(\Omega) \times Y(\Omega)} xyP([X = x] \cap [Y = y]).$$

□

6.3.2 Covariance

Définition 6.32 – Soient X et Y deux variables aléatoires discrètes sur (Ω, \mathcal{A}, P) , admettant un moment d'ordre 2. La covariance de X et Y est alors définie par

$$\text{Cov}(X, Y) = E((X - E(X))(Y - E(Y))).$$

Remarque. $\text{Cov}(X, Y)$ existe bien car $X - E(X)$ et $Y - E(Y)$ admettent toutes deux une variance, et donc leur produit admet une espérance.

Proposition 6.33 (Formule de Huygens) : Si X et Y sont deux variables aléatoires réelles discrètes admettant un moment d'ordre 2, alors

$$\text{Cov}(X, Y) = E(XY) - E(X)E(Y).$$

Démonstration.

$$(X - E(X))(Y - E(Y)) = XY - X \underbrace{E(Y)}_{\in \mathbf{R}} - Y \underbrace{E(X)}_{\in \mathbf{R}} + E(X)E(Y).$$

Donc par linéarité de l'espérance, on a

$$\text{Cov}(X, Y) = E(XY) - E(X)E(Y) - E(X)E(Y) + E(X)E(Y) = E(XY) - E(X)E(Y).$$

□

Remarque. Il est possible d'interpréter le signe de la covariance. En effet, s'il est positif, cela signifie qu'en moyenne, $X - E(X)$ et $Y - E(Y)$ sont de même signe, donc que lorsque X augmente, en moyenne Y augmente aussi.

En revanche s'il est négatif, cela signifie qu'en moyenne, $X - E(X)$ et $Y - E(Y)$ sont de signes opposés, donc que lorsque X augmente, en moyenne Y diminue.

Proposition 6.34 : Soient X, Y, Z trois variables aléatoires discrètes sur (Ω, \mathcal{A}, P) , admettant un moment d'ordre 2 (ou de manière équivalente, admettant une variance). Alors

1. $\text{Cov}(X, Y) = \text{Cov}(Y, X)$
2. $\forall \lambda \in \mathbf{R}, \text{Cov}(\lambda X + Y, Z) = \lambda \text{Cov}(X, Z) + \text{Cov}(Y, Z)$
3. $\forall \lambda \in \mathbf{R}, \text{Cov}(X, \lambda Y + Z) = \lambda \text{Cov}(X, Y) + \text{Cov}(X, Z)$
4. $\text{Cov}(X, X) = V(X) \geq 0$

Démonstration. 1. Évident.

2. Soit $\lambda \in \mathbf{R}$. Alors

$$\begin{aligned} \text{Cov}(\lambda X + Y, Z) &= E((\lambda X + Y)Z) - E(\lambda X + Y)E(Z) \\ &= \lambda E(XZ) + E(YZ) - (\lambda E(X) + E(Y))E(Z) \\ &= \lambda (E(XZ) - E(X)E(Z)) + E(YZ) - E(Y)E(Z) \\ &= \lambda \text{Cov}(X, Z) + \text{Cov}(Y, Z). \end{aligned}$$

3. Pour $\lambda \in \mathbf{R}$, on a, en utilisant les deux premières propriétés

$$\text{Cov}(X, \lambda Y + Z) = \text{Cov}(\lambda Y + Z, X) = \lambda \text{Cov}(Y, X) + \text{Cov}(Z, X) = \lambda \text{Cov}(X, Y) + \text{Cov}(X, Z).$$

4. $\text{Cov}(X, X) = E((X - E(X))^2) = V(X) \geq 0$.

□

Corollaire 6.35 – Si X et Y sont deux variables aléatoires discrètes indépendantes admettant un moment d'ordre 2, alors $\text{Cov}(X, Y) = 0$.

Démonstration. Puisque X et Y sont indépendantes, alors $E(XY) = E(X)E(Y)$, et alors par la formule de Huygens, $\text{Cov}(X, Y) = E(XY) - E(X)E(Y) = 0$. □

Remarque. Si $\text{Cov}(X, Y) = 0$, on dit que X et Y sont non corrélées. Ainsi, deux variables indépendantes sont non corrélées, mais la réciproque est fautive.

Remarque

Pour $X = Y$, on retrouve

$$V(X) = \text{Cov}(X, X) = E(X^2) - E(X)^2.$$

Terminologie

Nous dirons bientôt que la covariance est une forme bilinéaire symétrique, symétrique signifiant qu'elle vérifie la première propriété, bilinéaire signifiant qu'elle vérifie à la fois la seconde et la troisième.

⚠ Danger !

La réciproque est **FAUSSE** ! Deux variables aléatoires peuvent avoir une covariance nulle sans être indépendantes.

Exemple 6.36

Soit X une variable aléatoire telle que²

$$X(\Omega) = \{-1, 1\} \text{ et } P(X = 1) = P(X = -1) = \frac{1}{2}.$$

Soit Y une variable aléatoire sur le même espace qui suit la loi de Bernoulli de paramètre p , avec $p \in]0, 1[$.

On suppose que X et Y sont indépendantes, et on note $Z = XY$.

Par indépendance de X et Y , on a $E(Z) = E(X)E(Y) = 0$ car $E(X) = 0$.

D'autre part, on a $E(ZY) = E(XY^2) = E(XY) = E(X)E(Y) = 0$.

Et donc, $\text{Cov}(Z, Y) = E(ZY) - E(Z)E(Y) = 0$.

Pourtant Z et Y ne sont pas indépendantes, car

$$P([Z = 0] \cap [Y = 1]) = 0 \text{ et } P(Z = 0)P(Y = 1) = P(Y = 0)P(Y = 1) = p(1 - p) \neq 0.$$

² Une telle variable est parfois appelée variable de Rademacher

Détails

Puisque Y ne prend que les valeurs 0 et 1, $Y^2 = Y$.

6.3.3 Variance d'une somme de variables aléatoires discrètes

Proposition 6.37 : Soient X et Y deux variables aléatoires discrètes sur (Ω, \mathcal{A}, P) admettant une variance. Alors $X + Y$ admet une variance, et

$$V(X + Y) = V(X) + V(Y) + 2 \text{Cov}(X, Y).$$

Démonstration. Notons que

$$(X + Y - E(X + Y))^2 = ((X - E(X)) + (Y - E(Y)))^2 = (X - E(X))^2 + (Y - E(Y))^2 + 2(X - E(X))(Y - E(Y))$$

et que ces trois variables aléatoires admettent une espérance. Par linéarité de l'espérance, il vient alors

$$V(X + Y) = V(X) + V(Y) + 2 \text{Cov}(X, Y).$$

□

Corollaire 6.38 – Si X et Y sont indépendantes et admettent une variance, alors

$$V(X + Y) = V(X) + V(Y).$$

Écart-type

Il n'y a pas de formule analogue pour l'écart-type. En particulier, on n'a jamais (ou presque jamais) $\sigma_{X+Y} = \sigma_X + \sigma_Y$.

Exemple 6.39

Si X et Y sont indépendantes, avec $X \hookrightarrow \mathcal{B}(10, 1/4)$ et $Y \hookrightarrow \mathcal{P}(3)$, alors $Z = X + Y$ admet une variance, égale à $V(Z) = V(X) + V(Y) = \frac{15}{8} + 3 = \frac{39}{8}$.

6.3.4 Coefficient de corrélation linéaire

Proposition 6.40 : Soit X une variable aléatoire discrète admettant une variance, telle que $V(X) = 0$.

Alors X est presque sûrement constante égale à $E(X)$.

$$P(X = E(X)) = 1.$$

Démonstration. Notons $X(\Omega) = \{x_n, n \in \mathbf{N}\}$. Alors, par le théorème de transfert,

$$V(X) = \sum_{n=0}^{+\infty} (x_n - E(X))^2 P(X = x_n).$$

Mais chacun des $(x_n - E(X))^2 P(X = x_n)$ est positif.

Or une somme de termes positifs est nulle si et seulement si chacun des termes est nul.

Et donc

$$\forall n \in \mathbf{N}, (x_n - E(X))^2 P(X = x_n) = 0.$$

En particulier, si $x_n \neq E(X)$, $P(X = x_n) = 0$. Et puisqu'on a

$$\sum_{n=0}^{+\infty} P(X = x_n) = 1$$

«La somme des probas vaut 1», c'est la formule des probabilités totales.

nécessairement, $P(X = E(X)) = 1$. □

Définition 6.41 – Soient X, Y deux variables aléatoires réelles discrètes sur (Ω, \mathcal{A}, P) admettant une variance non nulle³. Alors le coefficient de corrélation linéaire de X et Y est

$$\rho(X, Y) = \frac{\text{Cov}(X, Y)}{\sqrt{V(X)}\sqrt{V(Y)}}.$$

³ Cette hypothèse est en fait presque toujours vérifiée d'après la proposition suivante.

Proposition 6.42 : Sous les hypothèses précédentes, on a

$$|\rho(X, Y)| \leq 1.$$

De plus, $\rho(X, Y) = \pm 1$ si et seulement si il existe un couple (a, b) de réels, avec $a \neq 0$, tels que $P(Y = aX + b) = 1$. Dans ce cas

- si $\rho(X, Y) = 1$, alors $a > 0$: X et Y varient dans le même sens.
- si $\rho(X, Y) = -1$, alors $a < 0$: X et Y varient dans des sens opposés.

Démonstration. Considérons la fonction $f(t) = V(tX - Y)$.

Alors $\forall t \in \mathbf{R}, f(t) \geq 0$ car une variance est toujours positive.

De plus, $f(t) = V(tX) + V(Y) - 2 \text{Cov}(tX, Y) = t^2 V(X) + V(Y) - 2t \text{Cov}(X, Y)$.

Ainsi, f est une fonction polynôme du second degré, et puisqu'elle est toujours positive, c'est qu'elle ne possède qu'au plus une racine, et son discriminant est donc négatif ou nul.

Ainsi, $4 \text{Cov}(X, Y)^2 - 4V(X)V(Y) \leq 0$.

On en déduit que $\text{Cov}(X, Y)^2 \leq V(X)V(Y)$ et donc

$$|\text{Cov}(X, Y)| \leq \sqrt{V(X)}\sqrt{V(Y)} \Leftrightarrow \left| \frac{\text{Cov}(X, Y)}{\sqrt{V(X)}\sqrt{V(Y)}} \right| \leq 1.$$

De plus, on a l'égalité si et seulement si le discriminant est nul, c'est-à-dire si et seulement si f s'annule en un certain réel a .

Mais alors $V(aX - Y) = 0$, et donc $aX - Y$ est une variable aléatoire presque certaine :

$$\exists b \in \mathbf{R} : P(aX - Y = -b) = 1 \Leftrightarrow P(Y = aX + b) = 1.$$

Dans ce cas, on a alors $V(Y) = V(aX + b) = a^2 V(X)$ et donc $\sqrt{V(Y)} = |a|\sqrt{V(X)}$. De même, on a

$$\text{Cov}(X, Y) = \text{Cov}(X, aX + b) = a \text{Cov}(X, X) + \text{Cov}(X, b) = aV(X) + \text{Cov}(X, b).$$

Mais, $V(b) = 0$ et donc $|\text{Cov}(X, b)| \leq \sqrt{V(X)}\sqrt{V(b)} = 0$.

Ainsi $\text{Cov}(X, Y) = aV(X)$. On en déduit que

$$\rho(X, Y) = \frac{\text{Cov}(X, Y)}{\sqrt{V(X)}\sqrt{V(Y)}} = \frac{a}{|a|}.$$

Et donc $\rho(X, Y)$ est du signe de a . □

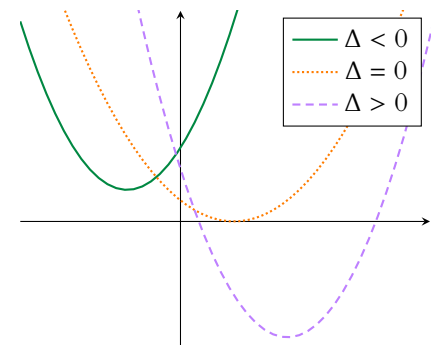


FIGURE 6.3— Un polynôme de degré 2 de signe constant est de discriminant négatif.

Précision

a est non nul car le polynôme f a un terme constant non nul, donc ne s'annule pas en 0.

Valeur absolue

$\frac{a}{|a|}$ vaut 1 si $a > 0$ et -1 si $a < 0$.

Corollaire 6.43 – Si X et Y sont indépendantes, et si X et Y admettent une variance non nulle, alors $\rho(X, Y) = 0$.

Remarque. Comme pour la covariance, la réciproque est fausse.

MÉTHODES DU CHAPITRE 6

✓ MÉTHODE 6.1 : Obtenir les lois marginales d'un couple (X, Y) à partir de la loi conjointe

Si on veut la loi de X , on applique la formule des probabilités totales au système complet d'événements $\{[Y = y], y \in Y(\Omega)\}$, et si on cherche la loi de Y , on applique la formule des probabilités totales au système complet d'événements $\{[X = x], x \in X(\Omega)\}$

Exercices de TD : 6.5, 6.7, 6.10, 6.14, 6.16

✓ MÉTHODE 6.2 : Déterminer la loi de $X + Y$: produit de convolution discret

La formule donnée ne vaut que pour des variables indépendantes, on prend donc bien soin de vérifier que c'est le cas. Lorsque c'est possible, on commence par déterminer le support de $X + Y$ afin d'être sûr de ne pas se lancer dans des calculs inutiles par la suite. Ensuite, on commencera par écrire la formule générale :

✓ Si X et Y sont à valeurs dans \mathbf{N} (ce qui est le plus souvent le cas)

$$\forall n \in (X + Y)(\Omega), P(X + Y = n) = \sum_{k=0}^n P(X = k)P(Y = n - k).$$



Dans le cas général,

$$\forall z \in (X + Y)(\Omega), P(X + Y = z) = \sum_{\substack{x \in X(\Omega) \\ z - x \in Y(\Omega)}} P(X = x)P(Y = z - x).$$

Avant de remplacer les probabilités par leur expression, on élimine les éventuels termes nuls que la somme comporte (par exemple, si $X \hookrightarrow \mathcal{G}(p)$, la somme peut commencer à $k = 1$ car $P(X = 0) = 0$).

Enfin, on calcule la somme.

Exercices de TD : 6.2

✓ MÉTHODE 6.3 : Déterminer la loi de $\min(X, Y)$ et de $\max(X, Y)$

- Commencer par déterminer le support de $\min(X, Y)$ ou de $\max(X, Y)$.
- Calculer $P(\min(X, Y) \geq k)$ et/ou $P(\max(X, Y) \leq k)$, en notant que

$$P(\min(X, Y) \geq k) = P([X \geq k] \cap [Y \geq k]) \text{ et } P(\max(X, Y) \leq k) = P([X \leq k] \cap [Y \leq k]).$$

Si les variables sont indépendantes, alors la probabilité de l'intersection est le produit des probabilités (sinon, utiliser la loi conjointe de X et de Y).

- On retrouve alors $P(\min(X, Y) = k)$ et $P(\max(X, Y) = k)$ en notant que pour une variable à valeurs entières, on a

$$P(X = k) = P(X \leq k) - P(X \leq k - 1) = P(X \geq k) - P(X \geq k + 1).$$

Exercices de TD : 6.6, 6.12

✓ MÉTHODE 6.4 : Calculer une covariance

- Utiliser les propriétés de la covariance (symétrie, bilinéarité, $\text{Cov}(X, X) = V(X)$), et le fait que deux variables indépendantes sont de covariance nulle et les variances usuelles.
- Utiliser la formule de Huygens, qui nécessite d'avoir accès à $E(XY)$. Cette dernière s'obtient principalement grâce au théorème de transfert.
- Si on connaît à la fois $V(X + Y)$, $V(X)$ et $V(Y)$, alors on peut isoler $\text{Cov}(X, Y)$ dans la relation $V(X + Y) = V(X) + 2\text{Cov}(X, Y) + V(Y)$.

Exercices de TD : 6.7, 6.8, 6.9, 6.10, 6.11, 6.12, 6.13, 6.14

✓ MÉTHODE 6.5 : Calculer la variance d'une somme

- Si les variables sont indépendantes, $V(X + Y) = V(X) + V(Y)$.
- Sinon, calculer la covariance de X et Y , et utiliser

$$V(X + Y) = V(X) + V(Y) + 2 \operatorname{Cov}(X, Y).$$

- Si on a la loi de $X + Y$, revenir aux fondamentaux : calculer l'espérance, le moment d'ordre 2, puis utiliser la formule de Huygens pour la variance.

Exercices de TD : 6.2, 6.10

✓ MÉTHODE 6.6 : Étudier l'indépendance des deux variables aléatoires discrètes.

- Pour prouver que X et Y ne sont pas indépendantes, on peut essayer de trouver $x \in X(\Omega), y \in Y(\Omega)$ tels que $P(X = x) \neq 0$, $P(Y = y) \neq 0$ et $P([X = x] \cap [Y = y]) = 0$.
- Si on a réussi à obtenir $\operatorname{Cov}(X, Y)$, et que cette covariance est non nulle, alors X et Y ne sont pas indépendantes.
- Si au contraire on cherche à prouver que deux variables sont indépendantes, on n'a guère d'autre solution que de revenir à la définition : prouver que

$$\forall x \in X(\Omega), y \in Y(\Omega), P([X = x] \cap [Y = y]) = P(X = x)P(Y = y).$$

Exercices de TD : 6.1, 6.5, 6.7, 6.8, 6.10, 6.11, 6.13

EXERCICES DU CHAPITRE 6

EXERCICE 6.1 Soient X et Y deux variables aléatoires indépendantes suivant la loi de Poisson de paramètre λ . On pose $Z = \min(X, Y)$ et $T = \max(X, Y)$. Les variables Z et T sont-elles indépendantes ?

PD

EXERCICE 6.2 Soient X et Y deux variables aléatoires indépendantes suivant la même loi géométrique $\mathcal{G}(p)$.

AD

- 1) Déterminer la loi de $X + Y$ et $V(X + Y)$.
- 2) Déterminer la loi de X sachant $[X + Y = n]$.

EXERCICE 6.3 Soient X et Y deux variables aléatoires indépendantes telles que $X \hookrightarrow \mathcal{U}(\llbracket 0, 3 \rrbracket)$ et $Y \hookrightarrow \mathcal{G}(\frac{3}{4})$. Montrer que X^Y admet une espérance, et la calculer.

PD

EXERCICE 6.4 Matrices aléatoires

AD

Soient X et Y deux variables aléatoires indépendantes sur (Ω, \mathcal{A}, P) suivant la même loi géométrique de paramètre $p \in]0, 1[$.

Pour tout $\omega \in \Omega$, on pose $M(\omega) = \begin{pmatrix} X(\omega) & Y(\omega) \\ Y(\omega) & X(\omega) \end{pmatrix}$. Déterminer la probabilité que $M(\omega)$ soit inversible.

EXERCICE 6.5 Soit $a > 0$ et soient X et Y deux variables aléatoires sur (Ω, \mathcal{A}, P) , à support dans \mathbf{N} , telles que la loi conjointe du couple (X, Y) soit donnée par

F

$$\forall (i, j) \in \mathbf{N}^2, P([X = i] \cap [Y = j]) = \lambda \frac{a^{i+j}}{i!j!}.$$

- 1) Déterminer la valeur de λ .
- 2) Donner les lois marginales du couple (X, Y) .
- 3) X et Y sont-elles indépendantes ?

EXERCICE 6.6 Soient X et Y deux variables aléatoires indépendantes suivant une loi géométrique $\mathcal{G}(p), p \in]0, 1[$.

PD

- 1) Pour $k \in \mathbf{N}^*$, calculer $P(Y \geq k)$.
- 2) Déterminer la loi de $U = \min(X, Y)$. Montrer que U suit une loi géométrique dont on déterminera le paramètre.
- 3) Déterminer la loi de $V = \max(X, Y)$.

EXERCICE 6.7 Soit $p \in]0, 1[$, et soit (X, Y) un couple de variables aléatoires à valeurs dans \mathbf{N} dont la loi conjointe est donnée par

PD

$$\forall (n, k) \in \mathbf{N}^2, P([X = n] \cap [Y = k]) = \begin{cases} \lambda(1-p)^k & \text{si } k \geq n \\ 0 & \text{sinon} \end{cases}$$

- 1) Déterminer la valeur de λ .
- 2) Déterminer les lois marginales du couple (X, Y) . Quelle est la loi de $X + 1$? En déduire $E(X)$ et $V(X)$.
- 3) Montrer que X et $Y - X$ suivent la même loi.
- 4) Montrer que X et $Y - X$ sont indépendantes. En déduire $\text{Cov}(X, Y)$. X et Y sont-elles indépendantes ?

EXERCICE 6.8 Soit $p \in]0, 1[$ et $n \geq 2$. On considère n joueurs de basket-ball qui tirent chacun deux lancers francs. On considère qu'à chaque lancer, un joueur a une probabilité p de marquer, et que les deux lancers sont indépendants. On note X la variable aléatoire égale au nombre de joueurs ayant marqué leur premier lancer franc, et Z la variable aléatoire égale au nombre de joueurs ayant marqué au moins un lancer franc.

PD

- 1) Déterminer la loi de X .
- 2) Montrer que Z suit une loi binomiale, donner son espérance et sa variance.
- 3) On pose $Y = Z - X$. Que représente la variable aléatoire Y ? Déterminer sa loi.
- 4) Les variables X et Y sont-elles indépendantes ? Calculer $\text{Cov}(X, Y)$.

EXERCICE 6.9 On suppose que le nombre de personnes qui se présentent à l'entrée d'un cinéma en une heure est une variable aléatoire X qui suit une loi de Poisson de paramètre $\lambda > 0$.

AD

Le cinéma comporte $N \geq 3$ caisses, et on suppose que chaque personne choisit au hasard sa caisse parmi les N .

Pour $i \in \llbracket 1, N \rrbracket$, on note X_i le nombre de personnes ayant choisi la caisse numéro i .

- 1) Déterminer la loi de X_i conditionnellement à l'événement $[X = n]$, puis la loi de X_i .
- 2) Déterminer, sans nouveaux calculs la loi de $X_1 + X_2$.

- 3) En déduire la covariance de X_1 et X_2 .
- 4) Déterminer le coefficient de corrélation linéaire entre X_1 et X_2 .
- 5) (★) Les variables aléatoires X_1 et X_2 sont-elles indépendantes ?

EXERCICE 6.10 Une urne contient n boules numérotées de 1 à n , dans laquelle on effectue deux tirages sans remise. On note X le numéro porté par la première boule et Y le numéro porté par la seconde boule.

AD

- 1) Déterminer la loi conjointe de (X, Y) , puis les lois de X et de Y .
- 2) Les variables aléatoires X et Y sont-elles indépendantes ?
- 3) Déterminer l'espérance de XY , puis en déduire $\text{Cov}(X, Y)$. Retrouver alors le résultat de la question précédente.
- 4) Calculer $V(X + Y)$, puis $\rho_{X, Y}$.

EXERCICE 6.11 Soit X et Y deux variables aléatoires indépendantes, suivant une même loi géométrique $\mathcal{G}(p)$. On pose $S = X + Y$ et $U = X - Y$.

PD

- 1) Sans déterminer les lois de S et U , montrer que $\text{Cov}(S, U) = 0$.
- 2) Les variables S et U sont-elles indépendantes ?

EXERCICE 6.12 Minimum, maximum et somme de lois uniformes

AD

On considère une urne contenant n boules numérotées de 1 à n , et on effectue deux tirages avec remise dans cette urne. On note X le numéro de la première boule extraite et Y le numéro de la seconde. On note également I le plus petit des deux numéros tirés et S le plus grand. On rappelle (cf cours) que la loi de S est donnée par $\forall k \in \llbracket 1, n \rrbracket, P(S = k) = \frac{2k - 1}{n^2}$.

- 1) Donner la loi de I .
- 2) Calculer $E(S)$, $E(I)$, puis $V(S)$ et $V(I)$.
- 3) Déterminer une relation liant X, Y, I et S . En déduire $V(I + S)$, puis le coefficient de corrélation linéaire de I et S .

EXERCICE 6.13 Soient X et Y deux variables aléatoires sur le même espace probabilisé, suivant des lois de Bernoulli de paramètres respectifs p_1 et p_2 .

AD

- 1) Montrer que $|\text{Cov}(X, Y)| \leq \frac{1}{4}$.
- 2) Montrer que X et Y sont indépendantes si et seulement si $\text{Cov}(X, Y) = 0$.

EXERCICE 6.14 Un exercice technique

D

Soit $n \in \mathbf{N}^*$. On considère une urne contenant 1 boule numérotée 1, 2 boules numérotées 2, etc, n boules numérotées n . On tire deux boules, sans remise dans cette urne, et on note X_1 le numéro de la première boule et X_2 celui de la seconde.

- 1) Déterminer la loi de X_1 et calculer son espérance.
- 2) Déterminer la loi conjointe du couple (X_1, X_2) et en déduire la loi de X_2 .
- 3) X_1 et X_2 sont-elles indépendantes ?
- 4) Calculer ρ_{X_1, X_2} .

EXERCICE 6.15 (QSP ESCP 2014)

D

Soient X et Y deux variables aléatoires discrètes sur (Ω, \mathcal{A}, P) , indépendantes, admettant une espérance $M \neq 0$ et une variance $V \neq 0$.

- 1) Calculer l'espérance et la variance de XY en fonction de M et V .
- 2) Les variables $X + Y$ et XY sont-elles indépendantes ?

EXERCICE 6.16 (Oral HEC 2010)

TD

Soit p un réel donné de $]0, 1[$. On pose $q = 1 - p$. On considère un couple (U, T) de variables aléatoires discrètes définies sur un espace probabilisé (Ω, \mathcal{A}, P) dont la loi de probabilité est donnée par : pour tout entier $n \geq 2$, pour tout $t \in \mathbf{Z}$,

$$P([U = n] \cap [T = t]) = \begin{cases} p^2 q^{n-2} & \text{si } |t| \leq n - 2 \text{ et si } n, |t| \text{ de même parité} \\ 0 & \text{sinon} \end{cases}$$

- 1) Vérifier que $\sum_{n=2}^{+\infty} \sum_{\substack{|t| \leq n-2 \\ n, |t| \text{ de même parité}}} p^2 q^{n-2} = 1$.
- 2) a) Déterminer la loi marginale de U .

b) En distinguant les trois cas $t = 0$, $t > 0$ et $t < 0$, montrer que la loi marginale de T est donnée par :

$$\forall t \in \mathbf{Z}, P(T = t) = \frac{pq^{|t|}}{1+q}.$$

- c) Calculer $E(T)$.
- 3) Soit n un entier supérieur ou égal à 2.
- Déterminer la loi conditionnelle de T sachant $[U = n]$.
 - Calculer l'espérance conditionnelle $E(T|U = n)$ de T sachant $[U = n]$.
- 4) a) Justifier l'existence de $E(U)$ et de $E(UT)$. Calculer $\text{Cov}(U, T)$.
- b) Les variables U et T sont-elles indépendantes ?

CORRECTION DES EXERCICES DU CHAPITRE 6

SOLUTION DE L'EXERCICE 6.1

On a $Z(\Omega) = T(\Omega) = \mathbf{N}$, par exemple car si $[X = k]$ et $[Y = k]$, alors $Z = T = k$, et donc $[Z = k]$ et $[T = k]$ sont de probabilité non nulle.

En particulier, $P(Z = 1) \neq 0$ et $P(T = 0) \neq 0$.

Mais on doit toujours avoir $Z \leq T$ et donc $P([Z = 1] \cap [T = 0]) = 0 \neq P(Z = 1)P(T = 0)$.

Et donc Z et T ne sont pas indépendantes.

Question subsidiaire

Que vaut $P(T = 0)$?

SOLUTION DE L'EXERCICE 6.2

1. Puisque les variables sont indépendantes et à valeurs dans \mathbf{N}^* , leur somme est à valeurs dans $\mathbf{N} \setminus \{1, 2\}$, et sa loi est donnée par

$$\begin{aligned} P(X + Y = n) &= \sum_{k=0}^n P(X = k)P(Y = n - k) \\ &= \sum_{k=1}^{n-1} P(X = k)P(Y = n - k) \\ &= \sum_{k=1}^{n-1} p(1-p)^{k-1} p(1-p)^{n-k-1} \\ &= \sum_{k=2}^n p^2(1-p)^{n-2} = p^2(n-1)(1-p)^{n-2} \end{aligned}$$

Ce résultat s'interprète bien si l'on se dit que c'est le nombre d'essais nécessaires pour obtenir 2 succès dans une répétition d'épreuves de Bernoulli de paramètre p : il faut 2 succès (d'où le p^2), $n-2$ échecs (d'où le $(1-p)^{n-2}$), et nous avons $(n-1)$ choix pour la position du premier succès (mais par contre aucun pour le deuxième, qui a nécessairement lieu au dernier essai).

Puisque X et Y sont indépendantes, nous savons que

$$V(X + Y) = V(X) + V(Y) = 2V(X) = \frac{2-2p}{p^2}$$

ce qui pourrait bien entendu se retrouver par un calcul direct à partir de la loi de $X + Y$.

2. Pour $k \in \mathbf{N}^*$, on a $[X + Y = n] \cap [X = k] = [Y = n - k] \cap [X = k]$, et donc par indépendance de X et Y ,

$$P_{[X+Y=n]}(X = k) = \frac{P([X = k] \cap [X + Y = n])}{P(X + Y = n)} = \frac{P([X = k] \cap [Y = n - k])}{P(X + Y = n)} = \frac{P([X = k])P([Y = n - k])}{P(X + Y = n)}.$$

En particulier, pour $k \geq n$, alors $P(Y = n - k) = 0$ et donc $P_{[X+Y=n]}(X = k) = 0$.

En revanche, pour $k \in \llbracket 1, n-1 \rrbracket$, on a

$$P_{[X+Y=n]}(X = k) = \frac{P(Y = n - k)P(X = k)}{P(X + Y = n)} = \frac{p(1-p)^{n-k-1} p(1-p)^{k-1}}{p^2(n-1)(1-p)^{n-2}} = \frac{1}{n-1}.$$

Et donc la loi de X sachant $[X + Y = n]$ est une loi uniforme sur $\llbracket 1, n-1 \rrbracket$.

SOLUTION DE L'EXERCICE 6.3

Par le théorème de transfert, sous réserve de convergence ,

$$\begin{aligned} E(X^Y) &= \sum_{i=0}^3 \sum_{j=1}^{+\infty} i^j P([X = i] \cap [Y = j]) \\ &= \sum_{i=1}^3 \sum_{j=1}^{+\infty} i^j P(X = i)P(Y = j) \\ &= \sum_{i=1}^3 \sum_{j=1}^{+\infty} i^j \frac{1}{4} \frac{3}{4} \left(\frac{1}{4}\right)^{j-1} \end{aligned}$$

Méthode

Pour $k = 0$ $P(X = k) = 0$ et pour $k = n$, on a $P(Y = n - k) = P(Y = 0) = 0$. Il faut impérativement s'intéresser aux bornes avant de remplacer les probabilités par leur expression. C'est du bon sens : puisque X et Y prennent des valeurs supérieures à 1, pour que $X + Y$ soit égale à n , il faut que X prenne une valeur entre 1 et $n - 1$.

Remarque

Notons que tous les termes étant positifs, la convergence équivaut à la convergence absolue (=sommabilité) requise par le théorème de transfert.

Si $i = 0$, alors pour $j \geq 1$, $i^j = 0$.

$$\begin{aligned}
&= \frac{3}{16} \sum_{i=1}^3 i \sum_{j=1}^{+\infty} \left(\frac{i}{4}\right)^{j-1} \\
&= \frac{3}{16} \sum_{i=1}^3 \frac{i}{1 - \frac{i}{4}} \\
&= \frac{3}{16} \sum_{i=1}^3 \frac{4i}{4-i} \\
&= \frac{3}{16} \left(\frac{4}{3} + 4 + 12 \right) = \frac{13}{4}.
\end{aligned}$$

Convergence

On a reconnu des séries géométriques de raison dans $] -1, 1[$, donc convergentes. Et puisqu'on a la somme d'un nombre fini de séries convergentes, leur somme converge également, de sorte que $E(X^Y)$ existe bien.

SOLUTION DE L'EXERCICE 6.4

Notons A l'événement $\{\omega \in \Omega : M(\omega) \text{ est inversible}\}$.

Puisque $M(\omega)$ est une matrice 2×2 , elle est inversible si et seulement si son déterminant est non nul, soit si et seulement si $X^2(\omega) - Y^2(\omega) \neq 0$.

Autrement dit, on a $\bar{A} = [X^2 = Y^2]$, et puisque X et Y sont à valeurs positives, on a $\bar{A} = [X = Y]$.

En appliquant la formule des probabilités totales au système complet d'événements $\{[X = k], k \in \mathbf{N}^*\}$, il vient

$$P(\bar{A}) = P(X = Y) = \sum_{k=1}^{+\infty} P([X = Y] \cap [X = k]) = \sum_{k=1}^{+\infty} P([X = k] \cap [Y = k]).$$

Par indépendance de X et Y , il vient donc

$$P(\bar{A}) = \sum_{k=1}^{+\infty} P(X = k)P(Y = k) = \sum_{k=1}^{+\infty} p^2(1-p)^{2(k-1)} = p^2 \sum_{i=0}^{+\infty} ((1-p)^2)^i = p^2 \frac{1}{1 - (1-p)^2} = \frac{p^2}{2p - p^2} = \frac{p}{2-p}.$$

Et donc

$$P(A) = 1 - P(\bar{A}) = 1 - \frac{p}{2-p} = \frac{2-2p}{2-p}.$$

SOLUTION DE L'EXERCICE 6.5

1. Nous savons¹ que la famille $(P([X = i] \cap [Y = j]))_{(i,j) \in \mathbf{N}^2}$ est sommable, et de somme 1. Or, pour $j \in \mathbf{N}$ fixé, on a

$$\sum_{i=1}^{+\infty} \lambda \frac{a^{i+j}}{i!j!} = \lambda \frac{a^j}{j!} \sum_{i=0}^{+\infty} \frac{a^i}{i!} = \lambda \frac{a^j}{j!} e^a.$$

Et alors, en sommant à présent sur j , il vient

$$\sum_{j=0}^{+\infty} \sum_{i=0}^{+\infty} \lambda \frac{a^{i+j}}{i!j!} = \lambda e^a \sum_{j=1}^{+\infty} \frac{a^j}{j!} = \lambda e^{2a}.$$

On en déduit donc que $1 = \lambda e^{2a} \Leftrightarrow \lambda = e^{-2a}$.

2. Par la formule des probabilités totales appliquée au système complet d'événements $\{[X = i], i \in \mathbf{N}\}$, on a

$$\forall j \in \mathbf{N}, P(Y = j) = \sum_{i=0}^{+\infty} P([X = i] \cap [Y = j]) = e^{-2a} e^a \frac{a^j}{j!} = e^{-a} \frac{a^j}{j!}$$

On reconnaît alors une loi de Poisson de paramètre a . De même, tous les calculs étant totalement symétriques en i et en j , on montrerait que X suit également une loi de Poisson de paramètre a .

3. On a $P(X = i) = e^{-\lambda} \frac{\lambda^i}{i!}$ et $P(Y = j) = e^{-\lambda} \frac{\lambda^j}{j!}$ et donc

$$\forall (i, j) \in \mathbf{N}^2, P(X = i)P(Y = j) = e^{-2\lambda} \frac{\lambda^{i+j}}{i!j!} = P([X = i] \cap [Y = j]),$$

de sorte que X et Y sont indépendantes.

¹ L'énoncé nous dit que nous sommes en présence d'une loi conjointe.

Astuce

Le calcul a déjà été fait précédemment : en étudiant la somme de la loi conjointe, on a en fait déterminé (sans le dire) une des deux lois marginales.

SOLUTION DE L'EXERCICE 6.6

1. Pour tout $k \in \mathbf{N}^*$,

$$P(Y \geq k) = \sum_{i=k}^{+\infty} P(Y = i) = \sum_{i=k}^{+\infty} p(1-p)^{i-1} = p(1-p)^{k-1} \sum_{j=0}^{+\infty} (1-p)^j = p(1-p)^{k-1} \frac{1}{p} = (1-p)^{k-1}.$$

Interprétation : ce résultat se comprend bien si on voit une géométrie comme le nombre d'essais avant l'obtention du premier succès lors d'une répétition d'épreuves de Bernoulli. En effet, $[X \geq k]$ est alors l'événement «avoir besoin d'au moins k essais», ce qui se dit encore «les $k-1$ premiers essais ont été des échecs». Et donc on comprend bien le $(1-p)^{k-1}$.

2. Pour tout $k \in \mathbf{N}^*$, on a $[U \geq k] = [X \geq k] \cap [Y \geq k]$.

Et donc par indépendance de X et Y , il vient $P(U \geq k) = P(X \geq k)P(Y \geq k)$. Et X et Y ayant même loi, on a donc également $P(Y \geq k) = (1-p)^{k-1}$.

Ainsi $P(U \geq k) = (1-p)^{2k-2}$. Puisque U est à valeurs entières, on en déduit que, pour tout $k \in \mathbf{N}^*$,

$$\begin{aligned} P(U = k) &= P(U \geq k) - P(U \geq k+1) = (1-p)^{2k-2} - (1-p)^{2k} = (1-p)^{2k-2}(1 - (1-p)^2) \\ &= (1 - (1-p)^2) \left((1-p)^2 \right)^{k-1}. \end{aligned}$$

Ainsi, U suit une loi géométrique, de paramètre $(1 - (1-p)^2)$.

3. De même $V(\Omega) = \mathbf{N}^*$, et pour tout $k \in \mathbf{N}^*$,

$$[V \leq k] = [\max(X, Y) \leq k] = [X \leq k] \cap [Y \leq k].$$

Par indépendance de X et Y , on en déduit que

$$P(Y \leq k) = P(X \leq k)P(Y \leq k) = P(X \leq k)^2$$

Mais on a

$$P(X \leq k) = 1 - P(X > k) = 1 - P(X \geq k+1) = 1 - (1-p)^k.$$

On en déduit que $P(V \leq k) = (1 - (1-p)^k)^2$.

Et puisque V est à valeurs entières, on en déduit que $\forall k \in \mathbf{N}^*$,

$$\begin{aligned} P(V = k) &= P(V \leq k) - P(V \leq k-1) = \left(1 - (1-p)^k\right)^2 - \left(1 - (1-p)^{k-1}\right)^2 \\ &= \left(2 - (2-p)(1-p)^{k-1}\right) p(1-p)^{k-1}. \end{aligned}$$

SOLUTION DE L'EXERCICE 6.7

1. Puisqu'on a affaire à une loi conjointe, on doit avoir $\sum_{k=0}^{+\infty} \sum_{n=0}^{+\infty} P([X = n] \cap [Y = k]) = 1$.

Or, pour $k \in \mathbf{N}$ fixé², on a

$$\sum_{n=0}^{+\infty} P([X = n] \cap [Y = k]) = \sum_{n=0}^k \lambda(1-p)^k = \lambda(k+1)(1-p)^k.$$

Et alors

$$\sum_{k=0}^{+\infty} \sum_{n=0}^{+\infty} P([X = n] \cap [Y = k]) = \lambda \sum_{k=0}^{+\infty} (k+1)(1-p)^k = \lambda \sum_{i=1}^{+\infty} i(1-p)^{i-1} = \frac{\lambda}{p^2}.$$

On a donc $\frac{\lambda}{p^2} = 1 \Leftrightarrow \lambda = p^2$.

2. Appliquons la formule des probabilités totales au système complet d'événements $\{[Y = k], k \in \mathbf{N}\}$. Alors

$$\forall n \in \mathbf{N}, P(X = n) = \sum_{k=0}^{+\infty} P([X = n] \cap [Y = k]) = p^2 \sum_{k=n}^{+\infty} (1-p)^k = p^2(1-p)^n \sum_{i=0}^{+\infty} (1-p)^i = p(1-p)^n.$$

Rappel

Être plus petit que le minimum de deux nombres, c'est être plus petit que chacun des deux nombres.

Support

X et Y sont à support inclus dans \mathbf{N}^* , donc il en est de même de U .

Rappel

Être plus grand que le maximum de deux nombres, c'est être plus grand que chacun de ces deux nombres.

X et Y ont même loi, donc

$$P(X \leq k) = P(Y \leq k).$$

Sommabilité

Puisqu'on a une loi conjointe, cette famille est automatiquement sommable.

² C'est le théorème de sommation par paquets : pour calculer la somme de la famille, on commence par calculer les sommes à k fixé (mais on aurait pu choisir de les calculer à n fixé), puis on calcule la somme de ces sommes.

Chgt d'indice

$i = k - n$

Notons dès à présent que $(X + 1)(\Omega) = \mathbf{N}^*$, et que pour tout $k \in \mathbf{N}^*$, on a

$$P(X + 1 = k) = P(X = k - 1) = p(1 - p)^{k-1}$$

de sorte que $X + 1 \hookrightarrow \mathcal{G}(p)$. On en déduit que

$$E(X) = E(X + 1) - E(1) = \frac{1}{p} - 1 \text{ et } V(X) = V(X + 1) = \frac{1 - p}{p^2}.$$

Pour la seconde loi marginale, appliquons la formule des probabilités totales au système complet d'événements $\{[X = n], n \in \mathbf{N}\}$. Alors pour tout $k \in \mathbf{N}$, on a

$$P(Y = k) = \sum_{n=0}^{+\infty} P([X = n] \cap [Y = k]) = p^2(k + 1)(1 - p)^k.$$

Précision

Cette somme a déjà été calculée dans la question 1.

3. Il s'agit de déterminer la loi de $Y - X$. Or, pour $k \in \mathbf{N}$, par la formule des probabilités totales, on a

$$\begin{aligned} P(Y - X = k) &= \sum_{n=0}^{+\infty} P([Y - X = k] \cap [X = n]) = \sum_{n=0}^{+\infty} P([X = n] \cap [Y = n + k]) \\ &= p^2 \sum_{n=0}^{+\infty} (1 - p)^{n+k} = p^2(1 - p)^k \sum_{n=0}^{+\infty} (1 - p)^n = p(1 - p)^k. \end{aligned}$$

On constate alors que $Y - X$ suit la même loi que X .

4. Pour tout couple $(n, k) \in \mathbf{N}^2$, on a d'une part

$$P([X = n] \cap [Y - X = k]) = P([X = n] \cap [Y = n + k]) = p^2(1 - p)^{n+k}$$

et d'autre part

$$P(X = k)P(Y - X = n) = p(1 - p)^n p(1 - p)^k = p^2(1 - p)^{n+k}.$$

Et donc pour tout couple $(n, k) \in \mathbf{N}^2$,

$$P([X = k] \cap [Y - X = n]) = P(X = k)P(Y - X = n).$$

Donc les variables X et $Y - X$ sont indépendantes.

On en déduit que $\text{Cov}(X, Y - X) = 0$. Or,

$$\text{Cov}(X, Y - X) = \text{Cov}(X, Y) - \text{Cov}(X, X) = \text{Cov}(X, Y) - V(X).$$

Et donc $\text{Cov}(X, Y) = V(X) = \frac{1 - p}{p^2}$.

SOLUTION DE L'EXERCICE 6.8

- Chaque joueur marque (ou non) indépendamment des autres, et donc on compte le nombre de succès au cours d'une répétition d'épreuves de Bernoulli indépendantes de paramètre p . Et alors $X \hookrightarrow \mathcal{B}(n, p)$.
- La probabilité qu'un joueur rate ses deux lancers est $(1 - p)^2$ (par indépendance des deux lancers), et donc la probabilité qu'il marque au moins un des deux lancers est $1 - (1 - p)^2$. On en déduit que $Z \hookrightarrow \mathcal{B}(n, 1 - (1 - p)^2)$.
- Y représente le nombre de joueurs ayant marqué uniquement leur second lancer franc. Comme pour chaque joueur, ceci se produit avec probabilité $p(1 - p)$, on en déduit que $Y \hookrightarrow \mathcal{B}(n, p(1 - p))$.
- X et Y ne sont pas indépendantes car on a

$$P(X = n) \neq 0, P(Y = n) \neq 0, \text{ et pourtant } P([X = n] \cap [Z = n]) = 0.$$

De plus, on a $V(Z) = V(X + Y) = V(X) + V(Y) + 2\text{Cov}(X, Y)$, ce qui peut encore se réécrire sous la forme

$$\text{Cov}(X, Y) = \frac{1}{2}(V(X + Y) - V(X) - V(Y)).$$

Remarque

Un calcul pourrait prouver que pour $k < 0$, $P(Y - X = k) = 0$, mais on peut également remarquer que la somme des $P(Y - X = n)$, $n \in \mathbf{N}$ vaut déjà 1, et donc qu'il n'y a «plus de place» pour d'autres probabilités non nulles.

³ En effet, on ne peut avoir à la fois les n joueurs ayant marqué leur premier lancer franc et ayant manqué le premier et réussi le second.

Astuce

Cette formule est toujours vraie, et facile à retrouver. Elle permet de calculer la covariance si l'on connaît $V(X + Y)$, $V(X)$ et $V(Y)$.

Soit encore

$$\begin{aligned}\text{Cov}(X, Y) &= \frac{n}{2} \left((1-p)^2(1 - (1-p)^2) - p(1-p) - p(1-p)(1 - p(1-p)) \right) \\ &= \frac{n}{2}(1-p) \left((1-p)(2p - p^2) - p - p(1-p + p^2) \right) \\ &= \frac{n(1-p)}{2} (2p - p^2 - 2p^2 + p^3 - p - p + p^2 - p^3) \\ &= -np^2(1-p)\end{aligned}$$

Notons qu'on aurait pu commencer par calculer la covariance, constater qu'elle est non nulle, et en déduire que X et Y ne sont pas indépendantes.

SOLUTION DE L'EXERCICE 6.9

1. Chaque client se présente à la caisse i avec probabilité $\frac{1}{N}$, et ce, indépendamment du choix des autres clients.

Puisqu'il y a n clients, nous reconnaissons alors une loi usuelle : la loi de X_i sachant $[X = n]$ est une loi binomiale $\mathcal{B}\left(n, \frac{1}{N}\right)$. Ainsi, on a

$$P_{[X=n]}(X_i = k) = \begin{cases} 0 & \text{si } k > n \\ \binom{n}{k} \frac{1}{N^k} \left(1 - \frac{1}{N}\right)^{n-k} & \text{sinon} \end{cases}$$

Puisque $\{[X = n], n \in \mathbf{N}\}$ est un système complet d'événements, par la formule des probabilités totales, on a

$$\begin{aligned}P(X_i = k) &= \sum_{n=0}^{+\infty} P_{[X=n]}(X_i = k)P(X = n) \\ &= \sum_{n=0}^{+\infty} \binom{n}{k} \frac{1}{N^k} \left(\frac{N-1}{N}\right)^{n-k} \frac{\lambda^n}{n!} e^{-\lambda} \\ &= \frac{e^{-\lambda}}{k!} \lambda^k \frac{1}{N^k} \sum_{n=k}^{+\infty} \frac{1}{(n-k)!} \lambda^{n-k} \left(\frac{N-1}{N}\right)^{n-k} \\ &= \frac{e^{-\lambda}}{k!} \lambda^k \frac{1}{N^k} \sum_{i=0}^{+\infty} \frac{\lambda^i}{i!} \left(\frac{N-1}{N}\right)^i \\ &= \frac{e^{-\lambda}}{k!} \lambda^k \frac{1}{N^k} e^{\lambda \frac{N-1}{N}} \\ &= \left(\frac{\lambda}{N}\right)^k \frac{1}{k!} e^{-\lambda/N}\end{aligned}$$

On en déduit que X_i suit une loi de Poisson de paramètre $\frac{\lambda}{N}$.

2. On peut remarquer que la loi de $X_1 + X_2$ sachant $X = n$ est une loi binomiale $\mathcal{B}\left(n, \frac{2}{N}\right)$. Le même calcul que dans la question précédente montre alors que $X_1 + X_2$ suit une loi de Poisson de paramètre $\frac{2\lambda}{N}$.
3. Puisque X_1 et X_2 admettent une variance, il en est de même de $X_1 + X_2$, et

$$V(X_1 + X_2) = V(X_1) + V(X_2) + 2 \text{Cov}(X_1, X_2).$$

Mais puisque nous avons prouvé que X_1, X_2 et $X_1 + X_2$ suivent toutes trois des lois de Poisson, nous connaissons leurs variances : on a

$$V(X_1) = V(X_2) = \frac{\lambda}{N} \text{ et } V(X_1 + X_2) = \frac{2\lambda}{N}.$$

On en déduit alors que

$$\text{Cov}(X_1, X_2) = \frac{1}{2} (V(X_1 + X_2) - V(X_1) - V(X_2)) = 0.$$

Signe

Notons qu'on a une covariance négative. Cela signifie qu'en moyenne, quand X augmente, Y a tendance à diminuer, et vice-versa. Cela semble conforme à l'intuition : plus le nombre de joueurs marquant leur premier panier est important, plus le nombre de joueurs manquant le premier et réussissant le second est faible.

Si $n < k$, $\binom{n}{k} = 0$.

Chgt d'indice

$i = n - k$.

Explication

$X_1 + X_2$ est le nombre de clients ayant choisi l'une des deux premières caisses. Or, pour chaque client, la probabilité qu'il choisisse l'une des deux premières caisses est $\frac{2}{N}$.

4. Notons que $X = X_1 + \dots + X_n$. On a donc

$$\text{Cov}(X_1, X) = \text{Cov}\left(X_1, \sum_{i=1}^n X_i\right) = \sum_{i=1}^n \text{Cov}(X_1, X_i) = \text{Cov}(X_1, X_1) = V(X_1) = \frac{\lambda}{N}.$$

On en déduit que

$$\rho(X_1, X) = \frac{\text{Cov}(X_1, X)}{\sqrt{V(X)V(X_1)}} = \frac{\lambda\sqrt{N}}{N\sqrt{\lambda}\sqrt{\lambda}} = \frac{1}{\sqrt{N}}.$$

5. La covariance est nulle, mais ceci n'implique pas que les variables soient indépendantes : c'est la réciproque qui est vraie ! Nous devons donc faire le calcul.

On a, en appliquant la formule des probabilités totales au système complet d'événements $\{[X = n], n \in \mathbf{N}\}$

$$\begin{aligned} P([X_1 = i] \cap [X_2 = j]) &= \sum_{n=0}^{+\infty} P_{[X=N]}([X_1 = i] \cap [X_2 = j])P(X = n) \\ &= \sum_{n=i+j}^{+\infty} \binom{n}{i} \binom{n-i}{j} \frac{1}{N^{i+j}} \left(1 - \frac{2}{N}\right)^{n-i-j} e^{-\lambda} \frac{\lambda^n}{n!} \\ &= \frac{e^{-\lambda} \lambda^{i+j}}{N^{i+j}} \sum_{n=i+j}^{+\infty} \frac{n!}{i!(n-i)!j!(n-i-j)!} \left(1 - \frac{2}{N}\right)^{n-i-j} \frac{\lambda^{n-i-j}}{n!} \\ &= \frac{e^{-\lambda} \lambda^{i+j}}{i!j!N^{i+j}} \sum_{n=i+j}^{+\infty} \left(1 - \frac{2}{N}\right)^{n-i-j} \frac{\lambda^{n-i-j}}{(n-i-j)!} \\ &= e^{-\lambda} \frac{\lambda^{i+j}}{i!j!N^{i+j}} \sum_{k=0}^{+\infty} \frac{\lambda^k \left(1 - \frac{2}{N}\right)^k}{k!} \\ &= e^{-\lambda} \frac{\lambda^{i+j}}{i!j!N^{i+j}} e^{\lambda\left(1 - \frac{2}{N}\right)} \\ &= e^{-\frac{2\lambda}{N}} \frac{\lambda^{i+j}}{i!j!N^{i+j}} \\ &= \frac{e^{-\frac{\lambda}{N}}}{i!} \left(\frac{\lambda}{N}\right)^i \frac{e^{-\frac{\lambda}{N}}}{j!} \left(\frac{\lambda}{N}\right)^j = P(X_1 = i)P(X_2 = j). \end{aligned}$$

Donc les variables aléatoires X_i et X_j sont indépendantes.

SOLUTION DE L'EXERCICE 6.10

1. Il y a en tout $n(n-1)$ tirages possibles.

Pour $k \neq \ell$, alors il y a un unique⁴ tirage réalisant $[X = k] \cap [Y = \ell]$, et donc

$$P([X = k] \cap [Y = \ell]) = \frac{1}{n(n-1)}.$$

Si $k = \ell$, alors $[X = k]$ et $[Y = \ell]$ sont incompatibles, donc $P([X = k] \cap [Y = \ell]) = 0$.

Il est évident que X suit la loi uniforme $\mathcal{U}(\llbracket 1, n \rrbracket)$.

Pour $\ell \in \llbracket 1, n \rrbracket$ fixé, on a, d'après la formule des probabilités totales appliquée au système complet d'événements $\{[X = k], k \in \llbracket 1, n \rrbracket\}$,

$$\begin{aligned} P(Y = \ell) &= \sum_{k=1}^n P([X = k] \cap [Y = \ell]) \\ &= \sum_{\substack{k=1 \\ k \neq \ell}}^n \frac{1}{n(n-1)} \\ &= \frac{1}{n(n-1)} \sum_{\substack{k=1 \\ k \neq \ell}}^n 1 = \frac{n-1}{n(n-1)} = \frac{1}{n}. \end{aligned}$$

Et donc nous constatons⁵ que Y suit également la loi uniforme sur $\llbracket 1, n \rrbracket$.

Détails

Les $\text{Cov}(X_1, X_i)$ sont nuls pour $i \neq 1$.
Nous venons de le prouver pour $\text{Cov}(X_1, X_2)$, le raisonnement est le même pour $i \in \llbracket 3, n \rrbracket$.

Explication

Il faut choisir i personnes parmi n prenant la caisse 1, et donc j parmi $n-i$ prenant la caisse 2.
Et alors les $n-i-j$ autres personnes choisissent une caisse parmi les $N-2$ autres.

Chgt d'indice

$k = n - i - j$.

Surprenant

Ce résultat n'est pas vraiment celui que l'on attend, on aurait tendance à penser que le nombre de personnes se présentant en caisse 1 dépend du nombre de personnes qui se présentent en caisse 2.

⁴ C'est le tirage où on obtient, dans cet ordre, les boules k et ℓ .

⚠ Danger !

Attention à bien enlever le terme $k = \ell$ de la somme, car il vaut 0, et non $\frac{1}{n(n-1)}$.

⁵ Sans grande surprise !

2. On a $P(X = 1)P(Y = 1) = \frac{1}{n^2} \neq 0$, alors que $P([X = 1] \cap [Y = 1]) = 0$, donc X et Y ne sont pas indépendantes.
3. Notons que XY est une variable aléatoire finie, et donc admet une espérance. Par le théorème de transfert, on a donc

$$\begin{aligned}
 E(XY) &= \sum_{k=1}^n \sum_{\ell=1}^n k\ell P([X = k] \cap [Y = \ell]) \\
 &= \sum_{k=1}^n \sum_{\substack{\ell=1 \\ \ell \neq k}}^n \frac{k\ell}{n(n-1)} \\
 &= \frac{1}{n(n-1)} \sum_{k=1}^n k \sum_{\substack{\ell=1 \\ \ell \neq k}}^n \ell \\
 &= \frac{1}{n(n-1)} \sum_{k=1}^n k \left(\sum_{\ell=1}^n \ell - k \right) \\
 &= \frac{1}{n(n-1)} \sum_{k=1}^n k \left(\frac{n(n+1)}{2} - k \right) \\
 &= \frac{1}{n(n-1)} \left(\frac{n(n+1)}{2} \sum_{k=1}^n k - \sum_{k=1}^n k^2 \right) \\
 &= \frac{1}{n(n-1)} \left(\left(\frac{n(n+1)}{2} \right)^2 - \frac{n(n+1)(2n+1)}{6} \right) \\
 &= \frac{1}{n(n-1)} \frac{n(n+1)}{12} (3n(n+1) - 2(2n+1)) \\
 &= \frac{1}{n(n-1)} \frac{n(n+1)}{12} (3n^2 - n - 2) \\
 &= \frac{n+1}{12(n-1)} (n-1)(3n+2) = \frac{(n+1)(3n+2)}{12}.
 \end{aligned}$$

On en déduit donc que

$$\begin{aligned}
 \text{Cov}(X, Y) &= E(XY) - E(X)E(Y) = \frac{(n+1)(3n+2)}{12} - \frac{n+1}{2} \frac{n+1}{2} \\
 &= \frac{(n+1)(3n+2)}{12} - \frac{(n+1)^2}{4} = \frac{n+1}{12} (3n+2 - 3(n+1)) \\
 &= -\frac{n+1}{12}.
 \end{aligned}$$

Puisque cette covariance n'est pas nulle, X et Y ne sont pas indépendantes.

4. On a

$$V(X + Y) = V(X) + 2 \text{Cov}(X, Y) + V(Y) = \frac{n^2 - 1}{12} - \frac{2(n+1)}{12} + \frac{n^2 - 1}{12} = \frac{(n+1)(n-2)}{6}.$$

Et d'autre part,

$$\rho_{X,Y} = \frac{\text{Cov}(X, Y)}{\sqrt{V(X)}\sqrt{V(Y)}} = \frac{\text{Cov}(X, Y)}{V(X)} = -\frac{n+1}{12} \frac{12}{n^2-1} = -\frac{1}{n-1}.$$

SOLUTION DE L'EXERCICE 6.11

1. En utilisant successivement la linéarité à gauche puis la linéarité à droite de la covariance, il vient

$$\begin{aligned}
 \text{Cov}(S, U) &= \text{Cov}(X + Y, X - Y) = \text{Cov}(X, X - Y) + \text{Cov}(Y, X - Y) \\
 &= \text{Cov}(X, X) - \text{Cov}(X, Y) + \text{Cov}(Y, X) - \text{Cov}(Y, Y) \\
 &= V(X) - V(Y) = 0.
 \end{aligned}$$

2. Ceci est bien évidemment une question piège : on a prouvé que la covariance était nulle, ceci ne saurait en aucun cas suffire à assurer l'indépendance de U et de S . Notons toutefois qu'on ne peut jamais avoir $U \geq S$. Et donc, on a par exemple

$$P(U = 2) \neq 0, P(S = 2) \neq 0 \text{ et } P([U = 2] \cap [S = 2]) = 0.$$

SOLUTION DE L'EXERCICE 6.12

1. Rappelons que X et Y sont indépendantes (car les tirages se font avec remise) et que X et Y suivent toutes deux la loi uniforme $\mathcal{U}(\llbracket 1, n \rrbracket)$. Alors $S = \max(X, Y)$ et $I = \min(X, Y)$. On a alors $I(\Omega) = \llbracket 1, n \rrbracket$, et pour tout $k \in \llbracket 1, n \rrbracket$,

$$P(I \geq k) = P([X \geq k] \cap [Y \geq k]) = P(X \geq k)P(Y \geq k) = P(X \geq k)^2.$$

$$\text{Mais } P(X \geq k) = \sum_{i=k}^n P(X = i) = \frac{1}{n} \sum_{i=k}^n 1 = \frac{n-k+1}{n}. \text{ Et donc}$$

$$P(I \geq k) = \left(\frac{n-k+1}{n}\right)^2 = \frac{(n-k+1)^2}{n^2}.$$

Mais puisque I est à valeurs entières, on a $P(I = k) = P(I \geq k) - P(I \geq k+1)$, et donc

$$P(I = k) = \frac{1}{n^2} \left((n-k+1)^2 - (n-k)^2 \right) = \frac{1}{n^2} (n-k+1+n-k)(n-k+1-n+k) = \frac{2n-2k+1}{n^2}.$$

2. Notons que les espérances et variances en jeu existent toutes car I et S sont des variables finies. On a

$$\begin{aligned} E(S) &= \sum_{k=1}^n kP(S = k) = \sum_{k=1}^n k \frac{2k-1}{n^2} \\ &= \frac{1}{n^2} \left(2 \sum_{k=1}^n k^2 - \sum_{k=1}^n k \right) = \frac{1}{n^2} \left(\frac{n(n+1)(2n+1)}{3} - \frac{n(n+1)}{2} \right) \\ &= \frac{1}{n^2} n(n+1) \frac{4n-1}{6} = \frac{(n+1)(4n-1)}{6n}. \end{aligned}$$

De même, on a

$$\begin{aligned} E(S^2) &= \sum_{k=1}^n k^2 P(S = k) = \frac{1}{n^2} \left(2 \sum_{k=1}^n k^3 - \sum_{k=1}^n k^2 \right) \\ &= \frac{1}{n^2} \left(\frac{n^2(n+1)^2}{2} - \frac{n(n+1)(2n+1)}{6} \right) = \frac{1}{n^2} n(n+1) \frac{3n^2+3n-2n-1}{6} \\ &= \frac{n+1}{n} \frac{3n^2+n-1}{6}. \end{aligned}$$

Enfin, par la formule de Huygens, il vient

$$\begin{aligned} V(S) &= E(S^2) - E(S)^2 = \frac{n+1}{n} \frac{3n^2+n-1}{6} - \frac{(n+1)^2(4n-1)^2}{36n^2} \\ &= (n+1) \frac{6n(3n^2+n-1) - (n+1)(16n^2-16n+1)}{36n^2} \\ &= (n+1) \frac{2n^3-2n^2+n-1}{36n^2} = \frac{(n+1)(n-1)(2n^2+1)}{36n^2}. \end{aligned}$$

De même, on a

$$\begin{aligned} E(I) &= \sum_{k=1}^n kP(I = k) = \sum_{k=1}^n k \frac{2n-2k+1}{n^2} \\ &= \frac{2n+1}{n^2} \frac{n(n+1)}{2} - \frac{2}{n^2} \frac{n(n+1)(2n+1)}{6} \\ &= \frac{(n+1)(2n+1)}{2n} - \frac{(n+1)(2n+1)}{3n} \end{aligned}$$

Précision

Il est facile de se convaincre que $P(U = 2) \neq 0$. Mais si on souhaite le prouver, il n'est pas nécessaire de calculer exactement cette probabilité, on peut juste remarquer que

$$[X = 3] \cap [Y = 1] \subset [U = 2]$$

et donc

$$\begin{aligned} P(U = 2) \\ \geq P(X = 3)P(Y = 1) \neq 0. \end{aligned}$$

Calcul

$$a^2 - b^2 = (a+b)(a-b).$$

Rappel

Pour tout $n \in \mathbf{N}$,

$$\sum_{k=1}^n k^3 = \frac{n^2(n+1)^2}{4}.$$

$$=(n+1)\frac{3(2n+1)-2(2n+1)}{6n}=\frac{(n+1)(2n+1)}{6n}.$$

$$\begin{aligned} E(I^2) &= \sum_{k=1}^n k^2 P(I=k) = \sum_{k=1}^n k^2 \frac{2n-2k+1}{n^2} \\ &= \frac{2n+1}{n^2} \frac{n(n+1)(2n+1)}{6} - \frac{2}{n^2} \frac{n^2(n+1)^2}{4} \\ &= \frac{(n+1)(2n+1)^2}{6n} - \frac{(n+1)^2}{2} = \frac{(n+1)(2n+1)^2}{6n} - \frac{3n(n+1)^2}{6n} \\ &=(n+1)\frac{4n^2+4n+1-3n^2-3n}{6n} = \frac{(n+1)(n^2+n+1)}{6n}. \end{aligned}$$

On en déduit alors que

$$\begin{aligned} V(I) = E(I^2) - E(I)^2 &= \frac{(n+1)(n^2+n+1)}{6n} - \frac{(n+1)^2(2n+1)^2}{36n^2} \\ &= \frac{6n(n+1)(n^2+n+1) - (n+1)^2(2n+1)^2}{36n^2} = (n+1)\frac{6n^3+6n^2+6n-4n^3-8n^2-5n-1}{36n^2} \\ &=(n+1)\frac{2n^3-2n^2+n-1}{36n^2} = \frac{(n+1)(n-1)(2n^2+1)}{36n^2}. \end{aligned}$$

3. Notons que $I+S=X+Y$. En effet, dans les deux cas, on considère la somme des numéros portés par les deux boules.

Donc $V(I+S) = V(X+Y)$, et par indépendance de X et Y , il vient

$$V(I+S) = V(X) + V(Y) = \frac{n^2-1}{12} + \frac{n^2-1}{12} = \frac{n^2-1}{6}.$$

D'autre part, on a

$$V(I+S) = V(I) + V(S) + 2 \operatorname{Cov}(I, S) \Leftrightarrow \operatorname{Cov}(I, S) = \frac{V(I+S) - V(I) - V(S)}{2}.$$

Et donc

$$\operatorname{Cov}(I, S) = \frac{1}{2} \left(\frac{n^2-1}{6} - \frac{(n+1)(n-1)(2n^2+1)}{18n^2} \right) = \frac{(n^2-1)^2}{36n^2}.$$

On en déduit que

$$\rho(I, S) = \frac{\operatorname{Cov}(I, S)}{\sqrt{V(I)V(S)}} = \frac{\operatorname{Cov}(I, S)}{V(I)} = \frac{1}{2n^2+1}.$$

SOLUTION DE L'EXERCICE 6.13

1. Nous savons qu'on aura toujours

$$|\rho_{X,Y}| \leq 1 \Leftrightarrow |\operatorname{Cov}(X, Y)| \leq \sqrt{V(X)}\sqrt{V(Y)}.$$

Mais nous savons également que

$$V(X) = p_1(1-p_1) \text{ et } V(X_2) = p_2(1-p_2).$$

Or, la fonction $p \mapsto p(1-p)$ admet sur $[0, 1]$ un maximum⁶ en $p = \frac{1}{2}$, et ce maximum vaut $\frac{1}{4}$.

On a donc

$$|\operatorname{Cov}(X, Y)| \leq \sqrt{\frac{1}{4}}\sqrt{\frac{1}{4}} = \frac{1}{4}.$$

2. Nous savons déjà que si X et Y sont indépendantes, alors $\operatorname{Cov}(X, Y) = 0$.

Supposons au contraire que $\operatorname{Cov}(X, Y) = 0$.

Alors, par la formule de Huygens, $E(XY) = E(X)E(Y) = p_1p_2$.

Mais d'autre part, d'après le théorème de transfert,

$$E(XY) = \sum_{i=0}^1 \sum_{j=0}^1 i \times j P([X=i] \cap [Y=j]) = P([X=1] \cap [Y=1]).$$

Interprétation

Plus n est grand, plus la corrélation entre I et S est faible.

Terminologie

Cette inégalité est parfois appelée inégalité de Cauchy-Schwarz pour la variance, nom qui sera justifié dans le chapitre 9.

⁶ Pour le voir, on peut soit dériver, soit remarquer qu'on a affaire à un polynôme de degré 2 et appliquer une formule bien connue.

Et donc $P([X = 1] \cap [Y = 1]) = p_1 p_2 = P(X = 1)P(Y = 1)$.

Donc les événements $[X = 1]$ et $[Y = 1]$ sont indépendants, ce qui⁷ suffit à garantir l'indépendance de X et Y .

⁷ Pour des lois de Bernoulli seulement ! C'est l'exemple 6.15 du cours.

SOLUTION DE L'EXERCICE 6.14

1. Il y a en tout $1 + 2 + \dots + n = \frac{n(n+1)}{2}$ boules dans l'urne, dont k portent le numéro k .

Et donc, pour tout $k \in \llbracket 1, n \rrbracket$, $P(X_1 = k) = \frac{2k}{n(n+1)}$.

Alors, on a

$$E(X_1) = \sum_{k=1}^n kP(X_1 = k) = \sum_{k=1}^n \frac{2k^2}{n(n+1)} = \frac{2}{n(n+1)} \frac{n(n+1)(2n+1)}{6} = \frac{2n+1}{3}.$$

2. Procédons là aussi par dénombrement : il y a en tout

$$\frac{n(n+1)}{2} \left(\frac{n(n+1)}{2} - 1 \right) = \frac{n(n+1)}{2} \frac{n^2 + n - 2}{2} = \frac{(n-1)n(n+1)(n+2)}{4}$$

tirage possibles.

Si $k \neq \ell$, il y a $k \times \ell$ tirages réalisant $[X_1 = k] \cap [X_2 = \ell]$.

Et si $k = \ell$, il y en a $k \times (k-1)$.

Et donc la loi conjointe du couple (X_1, X_2) est donnée par

$$P([X_1 = k] \cap [X_2 = \ell]) = \begin{cases} \frac{4k(k-1)}{(n-1)n(n+1)(n+2)} & \text{si } k = \ell \\ \frac{4k\ell}{(n-1)n(n+1)(n+2)} & \text{sinon} \end{cases}$$

D'après la formule des probabilités totales appliquée au système complet d'événements $\{[X_1 = k], 1 \leq k \leq n\}$, on a alors, pour tout $\ell \in \llbracket 1, n \rrbracket$,

$$\begin{aligned} P(X_2 = \ell) &= \sum_{k=1}^n P([X_1 = k] \cap [X_2 = \ell]) \\ &= \sum_{\substack{k=1 \\ k \neq \ell}}^n \frac{4k\ell}{(n-1)n(n+1)(n+2)} + \frac{4\ell(\ell-1)}{(n-1)n(n+1)(n+2)} \\ &= \frac{4}{(n-1)n(n+1)(n+2)} \left(\ell \sum_{k=1}^n k - \ell^2 + \ell^2 - \ell \right) \\ &= \frac{4}{(n-1)n(n+1)(n+2)} \left(\ell \frac{n(n+1)}{2} - \ell \right) \\ &= \frac{4}{(n-1)n(n+1)(n+2)} \frac{\ell(n-1)(n+2)}{2} = \frac{2\ell}{n(n+1)}. \end{aligned}$$

3. X_1 et X_2 ne sont pas indépendantes car $P(X_1 = 1) \neq 0$, $P(X_2 = 1) \neq 0$ et pourtant $P([X_1 = 1] \cap [X_2 = 1]) = 0$.

4. Utilisons la formule de Huygens : $\text{Cov}(X_1, X_2) = E(X_1 X_2) - E(X_1)E(X_2)$.
Commençons par noter que X_1 et X_2 ayant même loi, elles ont la même espérance. Et donc $E(X_2) = \frac{2n+1}{3}$.

Ensuite, pour obtenir l'espérance de $X_1 X_2$, on peut utiliser le théorème de transfert :

$$\begin{aligned} E(X_1 X_2) &= \sum_{k=1}^n \sum_{\ell=1}^n k\ell P([X_1 = k] \cap [X_2 = \ell]) \\ &= \sum_{k=1}^n \sum_{\substack{\ell=1 \\ \ell \neq k}}^n \frac{4k^2\ell^2}{(n-1)n(n+1)(n+2)} + \sum_{k=1}^n \frac{4k^3(k-1)}{(n-1)n(n+1)(n+2)} \\ &= \frac{4}{(n-1)n(n+1)(n+2)} \left(\sum_{k=1}^n \sum_{\ell=1}^n k^2\ell^2 - \sum_{k=1}^n k^4 + \sum_{k=1}^n k^4 - \sum_{k=1}^n k^3 \right) \end{aligned}$$

Espérance

Notons que X_1 est une variable finie, donc elle admet automatiquement une espérance, il n'y a pas à se poser de questions de convergence.

Détail

Si $k = 1$, il n'est pas possible d'avoir une boule numéro 1 lors des deux tirages, mais alors $k(k-1) = 0$.

Remarque

On obtient la même loi que X_1 !

$$\begin{aligned}
&= \frac{4}{(n-1)n(n+1)(n+2)} \left(\sum_{k=1}^n k^2 \sum_{\ell=1}^n \ell^2 - \frac{n^2(n+1)^2}{4} \right) \\
&= \frac{4}{(n-1)n(n+1)(n+2)} \left(\frac{n^2(n+1)^2(2n+1)^2}{36} - \frac{n^2(n+1)^2}{4} \right) \\
&= \frac{4}{(n-1)n(n+1)(n+2)} \left(\frac{n^2(n+1)^2((2n+1)^2 - 9)}{36} \right) \\
&= \frac{4}{(n-1)n(n+1)(n+2)} \frac{n^2(n+1)^2(n-1)(n+2)}{9} \\
&= \frac{4n(n+1)}{9}.
\end{aligned}$$

On en déduit donc que

$$\begin{aligned}
\text{Cov}(X_1, X_2) &= E(X_1 X_2) - E(X_1)E(X_2) \\
&= \frac{4n(n+1)}{9} - \frac{(2n+1)^2}{9} = -\frac{1}{9}.
\end{aligned}$$

Enfin, pour obtenir le coefficient de corrélation linéaire, il nous faut calculer les variances de X_1 et de X_2 , qui sont égales car ces variables sont de même loi :

$$E(X_1^2) = \sum_{k=1}^n k^2 P(X_1 = k) = \frac{2}{n(n+1)} \sum_{k=1}^n k^3 = \frac{n(n+1)}{2}.$$

Et alors, par la formule de Huygens,

$$V(X_1) = E(X_1^2) - E(X_1)^2 = \frac{n(n+1)}{2} - \frac{(2n+1)^2}{9} = \frac{(n-1)(n+2)}{18}.$$

Et donc, il vient enfin

$$\rho_{X_1, X_2} = \frac{\text{Cov}(X_1, X_2)}{\sqrt{V(X_1)}\sqrt{V(X_2)}} = -\frac{2}{(n-1)(n+2)}.$$

SOLUTION DE L'EXERCICE 6.15

1. Puisque X et Y sont indépendantes, on a $E(XY) = E(X)E(Y) = M^2$. De plus, par la formule de Huygens, on a $E(X^2) = V(X) + E(X)^2 = V + M^2$ et de même $E(Y^2) = V + M^2$. Et alors, X^2 et Y^2 étant indépendantes⁸, il vient

$$E((XY)^2) = E(X^2)E(Y^2) = (V + M^2)^2.$$

Alors, toujours par la formule de Huygens,

$$V(XY) = E((XY)^2) - E(XY)^2 = (V + M^2)^2 - (M^2)^2 = V^2 + 2VM^2 + M^4 - M^4 = V^2 + 2VM^2.$$

2. D'après la formule de Huygens⁹, on a

$$\begin{aligned}
\text{Cov}(X + Y, XY) &= E(X + Y(XY)) - E(X + Y)E(XY) \\
&= E(X^2 Y) + E(XY^2) - (E(X) + E(Y))E(XY) \\
&= E(X^2)E(Y) + E(X)E(Y^2) - E(X)E(X)E(Y) - E(Y)E(X)E(Y) \\
&= (V + M^2)M + M(V + M^2) - M^3 - M^3 = 2MV.
\end{aligned}$$

Et donc $\text{Cov}(X + Y, XY) \neq 0$, de sorte que XY et $X + Y$ ne sont pas indépendantes.

SOLUTION DE L'EXERCICE 6.16

1. Si $n \geq 2$ est pair, $n = 2k$, alors les entiers pairs compris entre $-(n-2)$ et $n-2$ sont les $2i$, $-(k-1) \leq i \leq k-1$, de sorte que

$$\sum_{\substack{|i| \leq n-2 \\ |i| \text{ pair}}} p^2 q^{n-2} = \sum_{i=-(k-1)}^{k-1} p^2 q^{n-2} = p^2 q^{n-2} (2k-2+1) = (n-1)p^2 q^{n-2}.$$

⁸ Car X et Y le sont.

⁹ Celle pour la covariance.

X^2 et Y sont indépendantes, de même que X et Y^2 .

Tous les termes de la somme sont égaux, il suffit de compter le nombre de termes.

De même, si $n = 2k + 1$, alors les entiers impairs entre $-2k + 1$ et $2k - 1$ sont les $2i + 1, -k \leq i \leq k - 1$, de sorte que

$$\sum_{\substack{|t| \leq n-2 \\ t \text{ impair}}} p^2 q^{n-2} = \sum_{i=-k}^{k-1} p^2 q^{n-2} = p^2 q^{n-2} (k - 1 + k + 1) = (2k) p^2 q^{n-2} = (n - 1) p^2 q^{n-2}.$$

Et donc

$$\sum_{n=2}^{+\infty} \sum_{\substack{|t| \leq n-2 \\ n, |t| \text{ de même parité}}} p^2 q^{n-2} = \sum_{n=2}^{+\infty} (n - 1) p^2 q^{n-2} = p^2 \sum_{m=1}^{+\infty} m q^{m-1} = p^2 \frac{1}{(1 - q)^2} = 1.$$

Chgt d'indice
 $m = n - 1.$

2.a. D'après la formule des probabilités totales appliquée au système complet d'événements $\{[T = t], t \in \mathbf{Z}\}$, on a

$$\begin{aligned} P([U = n]) &= \sum_{t \in \mathbf{Z}} P([U = n] \cap [T = t]) \\ &= \sum_{\substack{|t| \leq n-2 \\ |t|, n \text{ de même parité}}} p^2 q^{n-2} \\ &= (n - 1) p^2 q^{n-2}. \end{aligned}$$

Il ne reste qu'un nombre fini de termes dans la somme : les autres sont nuls !

2.b. Dans les trois cas, nous allons utiliser la formule des probabilités totales appliquée au système complet d'événements $\{[U = n], n \geq 2\}$.

Si $t = 0$, alors

$$\begin{aligned} P(T = 0) &= \sum_{n=2}^{+\infty} P([U = n] \cap [T = 0]) \\ &= \sum_{\substack{n=2 \\ n \text{ pair}}}^{+\infty} p^2 q^{n-2} \\ &= \sum_{i=1}^{+\infty} p^2 q^{2i-2} \\ &= p^2 \sum_{j=0}^{+\infty} (q^2)^j \\ &= p^2 \frac{1}{1 - q^2} = \frac{p^2}{2p - p^2} = \frac{p}{2 - p} = \frac{pq^0}{1 + q}. \end{aligned}$$

Chgt d'indice
 $j = i - 1.$

Si $t > 0$, alors $|t| \leq n - 2 \Leftrightarrow n \geq t + 2$. Et donc

$$P(T = t) = \sum_{\substack{n \geq t+2 \\ t, n \text{ de même parité}}}^{+\infty} p^2 q^{n-2}$$

Si t est pair, $t = 2k$, alors on obtient

$$P(T = t) = \sum_{i=k+1}^{+\infty} p^2 q^{2i-2} = p^2 q^{2k} \sum_{j=0}^{+\infty} (q^2)^j = p^2 q^{2k} \frac{1}{1 - q^2} = \frac{pq^t}{1 + q}.$$

Détails
Les entiers pairs supérieurs à $t + 2 = 2k + 2$ sont les $2i$, avec $i \geq k + 1$.

Et si t est impair, $t = 2k + 1$, alors

$$P(T = t) = \sum_{i=k+1}^{+\infty} p^2 q^{2i+1-2} = p^2 q^{2k+1} \sum_{j=0}^{+\infty} (q^2)^j = p^2 q^{2k+1} \frac{1}{1 - q^2} = \frac{pq^t}{1 + q}.$$

Enfn, si t est négatif, alors

$$P(T = t) = \sum_{\substack{|t| \leq n-2 \\ |t|, n \text{ de même parité}}} p^2 q^{n-2} = P(T = |t|) = \frac{pq^{|t|}}{1 + q}.$$

2.c. On a, sous réserve de convergence absolue,

$$E(T) = \sum_{t < 0} tP(T = t) + 0 \cdot P(T = 0) + \sum_{t > 0} tP(T = t) = \sum_{t=1}^{+\infty} -t \frac{pq^t}{1+q} + \sum_{t=1}^{+\infty} t \frac{pq^{-t}}{1+q}.$$

Les deux séries en jeu convergent absolument car $t \frac{pq^t}{1+q}$ est le terme général d'une série géométrique dérivée de raison $q \in]0, 1[$, donc convergente.

Et alors on a

$$E(T) = \sum_{t=1}^{+\infty} (-t) \frac{pq^t}{1+q} + \sum_{t=1}^{+\infty} t \frac{pq^{-t}}{1+q} = 0.$$

3.a. Notons que $P_{[U=n]}(T = t)$ est non nul si et seulement si $|t| \leq n - 2$ et si n et t ont même parité.

Et donc, sachant que $[U = n]$, T prend ses valeurs dans $I_n = \{-(n-2), -(n-4), \dots, n-4, n-2\}$.

Et pour tout t dans I_n , on $P([U = n] \cap [T = t]) = p^2 q^{n-2}$ de sorte que les $P_{[U=n]}(T = t)$ ont tous la même valeur : la loi de T sachant $[U = n]$ est une loi uniforme sur $I_n = \{-(n-2), -(n-4), \dots, n-4, n-2\}$.

En particulier, la loi de T sachant $[U = n]$ est finie, donc $E(T|U = n)$ existe.

De plus, si $t \in I_n$, alors $-t \in I_n$, et donc la loi de $-T$ sachant $[U = n]$ est la même que la loi de T sachant $[U = n]$.

Et donc $2E(T|U = n) = E(T|U = n) + E(-T|U = n) = E(0|U = n) = 0$.

On en déduit donc que $E(T|U = n) = 0$.

4.a. U admet une espérance si et seulement si la série de terme général $n(n-1)p^2q^{n-2}$ converge. Or, il s'agit là d'une série géométrique dérivée seconde, convergente car de raison $q \in]0, 1[$. Donc $E(U)$ existe.

Pour l'existence de $E(UT)$, notons que pour tout $n \in \mathbf{N}$, la loi de UT sachant $[U = n]$ est la loi de nT sachant $[U = n]$. Et en particulier, $E(UT|U = n) = E(nT|U = n) = nE(T|U = n) = 0$. D'après la formule de l'espérance totale appliquée au système complet d'événements $\{[U = n], n \geq 2\}$, on a

$$E(UT) = \sum_{n=2}^{+\infty} P(U = n)E(UT|U = n) = 0.$$

Et donc, par la formule de Huygens, $\text{Cov}(U, T)$ existe et vaut

$$\text{Cov}(U, T) = \underbrace{E(UT)}_{=0} - \underbrace{E(U)}_{=0} \underbrace{E(T)}_{=0} = 0.$$

4.b. On a $P(U = 2) \neq 0$ et $P(T = 1) \neq 0$, mais $P([U = 2] \cap [T = 1]) = 0$ car 1 et 2 ne sont pas de même parité.

Donc U et T ne sont pas indépendantes.

Détails

Cette famille est bien sommable : on a séparé \mathbf{Z} en trois «paquets» : les entiers strictement positifs, les entiers strictement négatifs, et 0. Chacune des trois sommes converge absolument (voir ci-dessous) et donc la famille est sommable.

Remarque

Si n est pair, alors $0 \in I_n$, et si n est impair, $0 \notin I_n$.

⚠ Danger !

Il ne s'agit pas d'une loi uniforme sur $\llbracket -(n-2), n-2 \rrbracket$, car T ne prend que des valeurs de même parité que n .

Remarque

On pourrait en fait montrer que si X et Y sont deux lois géométriques indépendantes de même paramètre p , alors le couple $(X + Y, X - Y)$ a même loi que le couple (U, T) .

VALEURS PROPRES, VECTEURS PROPRES

Cours	209
7.1 Valeurs propres d'une matrice carrée	210
7.2 Détermination pratique des valeurs propres et des sous-espaces propres.	213
7.3 Valeurs propres et vecteurs propres d'un endomorphisme	216
7.4 Complément sur le rang et les matrices échelonnées	222
Méthodes	228
Exercices	229
Correction des exercices	232

INTRODUCTION

Un endomorphisme de \mathbf{R}^n est généralement donné par des formules, ou encore par sa matrice dans une base.

Par exemple, considérons l'application $f : \mathbf{R}^2 \rightarrow \mathbf{R}^2$ définie par

$$f(x, y) = (y, x) \text{ soit encore } \text{Mat}_{\mathcal{B}_{can}}(f) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

De même, soit g l'application $\mathbf{R}^3 \rightarrow \mathbf{R}^3$ définie par

$$g(x, y, z) = \frac{1}{2} (x - \sqrt{2}y + z, \sqrt{2}x - \sqrt{2}z, x\sqrt{2}y + z) \text{ soit encore } \text{Mat}_{\mathcal{B}_{can}}(g) = \frac{1}{2} \begin{pmatrix} 1 & -\sqrt{2} & 1 \\ \sqrt{2} & 0 & -\sqrt{2} \\ 1 & \sqrt{2} & 1 \end{pmatrix}.$$

Ni les formule ni les matrices ne nous aident vraiment à comprendre ce que sont f et g .

Et pourtant, ces deux applications ont une interprétation géométrique simple.

Ainsi, f est la symétrie par rapport à la droite d'équation $y = x$.

Et g est en fait une rotation dans l'espace : c'est la rotation qui a pour axe la droite dirigée par le vecteur $(1, 0, 1)$, et dont l'angle vaut $\frac{\pi}{2}$.

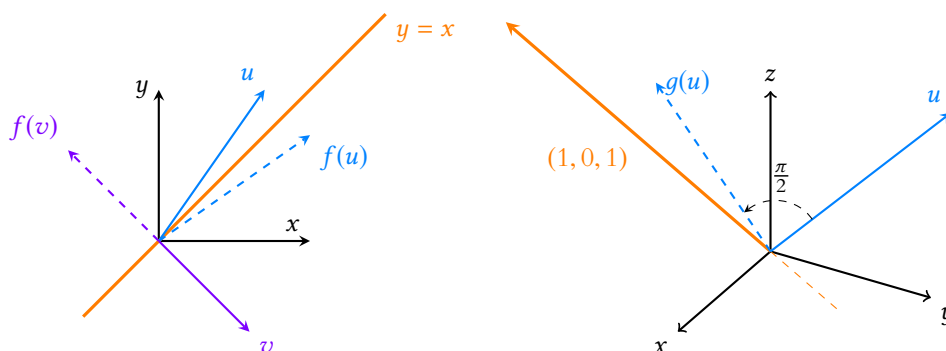


FIGURE 7.1 – Dans les deux cas, on a représenté un ou des vecteurs ainsi que leur image par l'endomorphisme.

g est une rotation autour de la droite (en orange) dirigée par $(1, 0, 1)$. L'image d'un vecteur u est obtenue en tournant ce vecteur d'un quart de tour autour de l'axe.

Dans les deux cas, il est intéressant de noter qu'il existe des droites qui sont invariantes par la transformation. Par exemple, dans le cas de f , la droite d'équation $y = x$ (dirigée par le vecteur $(1, 1)$) n'est pas affectée par la symétrie.

De même, la droite d'équation $y = -x$ (dirigée par $(1, -1)$) n'est globalement pas affectée par la symétrie. Pourtant le vecteur $(1, -1)$ est envoyé sur son opposé, mais la droite elle-même ne «bouge» pas.

Dans le cas de g , toute droite est «tournée» d'un angle $\frac{\pi}{2}$ par la rotation, sauf l'axe de la rotation, qui est invariant.

Le but principal de ce chapitre sera, étant donné un endomorphisme, d'essayer de trouver les droites qui sont invariantes par cet endomorphisme.

Notons que lorsqu'on dit qu'une droite $\text{Vect}(u)$ est invariante par f , en termes d'algèbre linéaire, cela signifie que $\text{Vect}(u)$ est stable par f . Et en particulier, $f(u) \in \text{Vect}(u)$, de sorte qu'il existe un scalaire λ tel que $f(u) = \lambda u$.

Les vecteurs pour lesquels il existe un tel scalaire, appelés vecteurs propres, vont jouer un rôle central dans ce chapitre.

Dans toute la suite, \mathbf{K} désigne indifféremment \mathbf{R} ou \mathbf{C} .

7.1 VALEURS PROPRES D'UNE MATRICE CARRÉE

7.1.1 Rappels et compléments sur l'utilisation des vecteurs colonnes

Soit $A = (a_{i,j}) \in \mathcal{M}_{n,1}(\mathbf{K})$. Alors l'application $f_A : \mathcal{M}_{n,1}(\mathbf{K}) \rightarrow \mathcal{M}_{n,1}(\mathbf{K})$ définie par $f_A(X) = AX$ est linéaire et donc est un endomorphisme de $\mathcal{M}_{n,1}(\mathbf{K})$.

Soit $\mathcal{B} = (E_1, \dots, E_n) = \left(\begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \dots, \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix} \right)$ la base canonique de $\mathcal{M}_{n,1}(\mathbf{K})$.

Alors on a

$$f_A(E_1) = AE_1 = \begin{pmatrix} a_{1,1} & \dots & a_{1,n} \\ \vdots & & \vdots \\ \vdots & & \vdots \\ a_{n,1} & \dots & a_{n,n} \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} = \begin{pmatrix} a_{1,1} \\ a_{2,1} \\ \vdots \\ a_{n,1} \end{pmatrix} = a_{1,1}E_1 + a_{2,1}E_2 + \dots + a_{n,1}E_n.$$

Et de même, on prouve que $\forall j \in \llbracket 1, n \rrbracket$, $f_A(E_j) = \sum_{i=1}^n a_{i,j}E_i$.

Et donc, la matrice de f_A dans la base \mathcal{B} est

$$\text{Mat}_{\mathcal{B}}(f_A) = \begin{pmatrix} f_A(E_1) & f_A(E_2) & \dots & f_A(E_n) \\ a_{1,1} & a_{1,2} & \dots & a_{1,n} \\ a_{2,1} & a_{2,2} & \dots & a_{2,n} \\ \vdots & \vdots & & \vdots \\ a_{n,1} & a_{n,2} & \dots & a_{n,n} \end{pmatrix} \begin{matrix} E_1 \\ E_2 \\ \vdots \\ E_n \end{matrix} = A.$$

On dira que f_A est l'endomorphisme de $\mathcal{M}_{n,1}(\mathbf{K})$ canoniquement associé à A .

7.1.2 Valeurs propres, vecteurs propres, sous-espaces propres

Définition 7.1 – Soit $A \in \mathcal{M}_n(\mathbf{K})$, et soit $\lambda \in \mathbf{K}$. On dit que λ est valeur propre de A s'il existe un vecteur colonne $X \in \mathcal{M}_{n,1}(\mathbf{K})$ **non nul** tel que $AX = \lambda X$.

Remarque. Il est important de demander à ce que $X \neq 0$, car

$$\forall \lambda \in \mathbf{K}, A \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix} = \lambda \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}.$$

Et donc si on ne demandait pas à X d'être non nul, tout élément de \mathbf{K} serait valeur propre de A , et donc notre définition serait peu intéressante puisqu'une valeur propre de A ne

Calcul

Si l'on voulait être vraiment rigoureux, il faudrait utiliser la formule définissant le produit matriciel, mais ici E_1 est vraiment simple et on se convainc facilement de ce résultat en faisant «à la main» le produit de matrices, même si on l'écrit avec des pointillés.

Détail

f_A est donc l'unique endomorphisme de $\mathcal{M}_{n,1}(\mathbf{K})$ dont la matrice dans la base canonique (de $\mathcal{M}_{n,1}(\mathbf{K})$) est A .

serait rien d'autre qu'un scalaire¹.

Dans ce cas, autant continuer à dire comme d'habitude «soit λ un scalaire» plutôt que «soit λ une valeur propre de A » !

¹ C'est-à-dire un élément de \mathbf{K} .

Définition 7.2 – Soit $A \in \mathcal{M}_n(\mathbf{K})$ et soit $\lambda \in \mathbf{K}$. Un vecteur colonne $X \in \mathcal{M}_{n,1}(\mathbf{R})$ **non nul** tel que $AX = \lambda X$ est appelé **vecteur propre de A associé à la valeur propre λ** .

Remarque. Avec les définitions précédentes, on a

λ est valeur propre de $A \Leftrightarrow$ il existe un vecteur propre associé à A pour la valeur propre λ .

Exemple 7.3

$$\text{Soit } A = \begin{pmatrix} 1 & -2 & 2 \\ -2 & 1 & 2 \\ -2 & -2 & 5 \end{pmatrix}.$$

Alors on a $A \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 3 \\ 0 \\ 3 \end{pmatrix} = 3 \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$, et donc 3 est valeur propre de A , et $\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$ est un vecteur propre de A associé à la valeur propre 3.

Définition 7.4 – Si $A \in \mathcal{M}_n(\mathbf{K})$, on appelle spectre de A et on note $\text{Spec}(A)$ (ou parfois $\text{Sp}(A)$) l'ensemble des valeurs propres de A .

En Sci Lab, la commande `spec(A)` permet d'obtenir le spectre de A .

Notons que Sci Lab utilise des algorithmes de calcul approché pour obtenir les valeurs propres, et qu'il se peut que les valeurs données ne soient pas exactes.

Définition 7.5 – Si $A \in \mathcal{M}_n(\mathbf{K})$, et si $\lambda \in \mathbf{K}$ est valeur propre de A (i.e. si $\lambda \in \text{Spec}(A)$), on appelle sous-espace propre de A associé à la valeur propre λ , et on note $E_\lambda(A)$ l'ensemble des vecteurs colonnes $X \in \mathcal{M}_{n,1}(\mathbf{R})$ tels que $AX = \lambda X$. Autrement dit

$$E_\lambda(A) = \{X \in \mathcal{M}_{n,1}(\mathbf{R}) : AX = \lambda X\}.$$

Vecteur nul

Le vecteur nul est toujours dans $E_\lambda(A)$ car

$$A \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix} = \lambda \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}.$$

Ainsi, $E_\lambda(A)$ est formé de l'ensemble des vecteurs propres de A pour la valeur propre λ et du vecteur nul.

7.1.3 Structure et dimension des sous-espaces propres

Proposition 7.6 : Soit $A \in \mathcal{M}_n(\mathbf{K})$, et soit $\lambda \in \text{Spec}(A)$. Alors $E_\lambda(A)$ est un sous-espace vectoriel de $\mathcal{M}_{n,1}(\mathbf{K})$. De plus, on a toujours

$$1 \leq \dim E_\lambda(A) \leq n.$$

Démonstration. On a $E_\lambda(A) \subset \mathcal{M}_{n,1}(\mathbf{K})$ par définition, et nous avons déjà prouvé que le vecteur nul de $\mathcal{M}_{n,1}(\mathbf{K})$ est toujours dans $E_\lambda(A)$.

Montrons que $E_\lambda(A)$ est stable par combinaison linéaire : soient $X, Y \in \mathcal{M}_{n,1}(\mathbf{K})$ tels que $AX = \lambda X$ et $AY = \lambda Y$, et soit $\alpha \in \mathbf{K}$. Alors

$$A(\alpha X + Y) = \alpha AX + AY = \alpha \lambda X + \lambda Y = \lambda(\alpha X + Y).$$

Ainsi, $\alpha X + Y \in E_\lambda(A)$, et donc $E_\lambda(A)$ est bien un sous-espace vectoriel de $\mathcal{M}_{n,1}(\mathbf{K})$.

De plus, puisque $\dim \mathcal{M}_{n,1}(\mathbf{K}) = n$, on a nécessairement $\dim E_\lambda(A) \leq n$.

Enfin, puisque $\lambda \in \text{Spec}(A)$, il existe $X \in \mathcal{M}_{n,1}(\mathbf{K})$ **non nul** tel que $AX = \lambda X \Leftrightarrow X \in E_\lambda(A)$.

En particulier, $E_\lambda(A)$ n'est pas réduit au vecteur nul, et donc $\dim E_\lambda(A) \geq 1$. □

$\dim E_\lambda(A) \geq 1$

En fait, $E_\lambda(A)$ est toujours un sous-espace vectoriel de $\mathcal{M}_{n,1}(\mathbf{K})$, même si λ n'est pas valeur propre de A . Mais alors il ne contient que le vecteur nul et donc est de dimension 0.

Cas particulier : la valeur propre 0

Proposition 7.7 : Soit $A \in \mathcal{M}_n(\mathbf{K})$. Alors $0 \in \text{Spec}(A)$ si et seulement si $\text{rg}(A) < n$.
De plus, on a alors

$$\dim E_0(A) = n - \text{rg}(A).$$

Démonstration.

$$\begin{aligned} 0 \text{ est valeur propre de } A &\iff \text{il existe } X \neq 0 \text{ tel que } AX = 0 \\ &\iff A \text{ n'est pas inversible} \\ &\iff \text{rg}(A) < n. \end{aligned}$$

Soit f l'endomorphisme de $\mathcal{M}_{n,1}(\mathbf{K})$ canoniquement associé à A . Alors

$$E_0(A) = \{X \in \mathcal{M}_{n,1}(\mathbf{K}) : AX = 0\} = \{X \in \mathcal{M}_{n,1}(\mathbf{K}) : f(X) = 0\} = \text{Ker}(f).$$

D'autre part, la matrice de f dans la base canonique de $\mathcal{M}_{n,1}(\mathbf{K})$ est A , donc $\text{rg}(f) = \text{rg}(A)$. Par le théorème du rang appliqué à f , on a alors

$$\dim E_0(A) = \dim \text{Ker}(f) = \dim \mathcal{M}_{n,1}(\mathbf{K}) - \text{rg}(f) = n - \text{rg}(A).$$

□

Cas général

Lemme 7.8. Soit $A \in \mathcal{M}_n(\mathbf{K})$ et soit $\lambda \in \mathbf{K}$. Alors λ est valeur propre de A si et seulement si 0 est valeur propre de $A - \lambda I_n$. Dans ce cas, on a alors $E_\lambda(A) = E_0(A - \lambda I_n)$.

Démonstration.

$$\begin{aligned} \lambda \text{ est valeur propre de } A &\iff \text{il existe } X \in \mathcal{M}_{n,1}(\mathbf{K}) \text{ non nul tel que } AX = \lambda X \\ &\iff \text{il existe } X \in \mathcal{M}_{n,1}(\mathbf{K}) \text{ non nul tel que } (A - \lambda I_n)X = 0 \\ &\iff 0 \text{ est valeur propre de } A - \lambda I_n \end{aligned}$$

De plus, $X \in E_\lambda(A) \iff AX = \lambda X \iff (A - \lambda I_n)X = 0 \iff X \in E_0(A - \lambda I_n)$.

□

Théorème 7.9 : Soit $A \in \mathcal{M}_n(\mathbf{K})$ et soit $\lambda \in \mathbf{K}$. Alors $\lambda \in \text{Spec}(A)$ si et seulement si $\text{rg}(A - \lambda I_n) < n$, c'est-à-dire si et seulement si $A - \lambda I_n$ n'est pas inversible. Dans ce cas, $\dim E_\lambda(A) = n - \text{rg}(A - \lambda I_n)$.

Démonstration. En combinant le lemme 7.8 et la proposition 7.7, on a

$$\lambda \in \text{Spec}(A) \iff 0 \in \text{Spec}(A - \lambda I_n) \iff \text{rg}(A - \lambda I_n) < n.$$

De même

$$\dim E_\lambda(A) = \dim E_0(A - \lambda I_n) = n - \text{rg}(A - \lambda I_n).$$

□

Exemple 7.10

Nous avons déjà vu que 3 est valeur propre de $A = \begin{pmatrix} 1 & -2 & 2 \\ -2 & 1 & 2 \\ -2 & -2 & 5 \end{pmatrix}$.

Nous avons alors $\dim E_3(A) = \underbrace{3}_{=n} - \text{rg}(A - \underbrace{3}_{=\lambda} I_3) = 3 - \text{rg} \begin{pmatrix} -2 & -2 & 2 \\ -2 & -2 & 2 \\ -2 & -2 & 2 \end{pmatrix}$.

Détails

C'est la proposition 5.27

Rappel

Nous avons défini précédemment cet endomorphisme par $X \mapsto AX$.

Mais cette dernière matrice a toutes ses colonnes proportionnelles, donc est de rang 1 : $\dim E_3(A) = 3 - 1 = 2$.

Corollaire 7.11 – Soit $A \in \mathcal{M}_n(\mathbf{K})$.

Alors A et ${}^t A$ ont les mêmes valeurs propres : $\text{Spec}(A) = \text{Spec}({}^t A)$.

De plus, les sous-espaces propres de A et de ${}^t A$ associés à une valeur propre λ ont même dimension : $\dim E_\lambda(A) = \dim E_\lambda({}^t A)$.

Danger !

En revanche, les sous-espaces propres de A et de ${}^t A$ n'ont aucune raison d'être **égaux** : un vecteur X vérifiant $AX = \lambda X$ ne vérifiera pas nécessairement ${}^t AX = \lambda X$.

Démonstration. Notons que pour tout $\lambda \in \mathbf{K}$, on a

$$\text{rg}({}^t A - \lambda I_n) = \text{rg}({}^t(A - \lambda I_n)) = \text{rg}(A - \lambda I_n).$$

Et donc $\lambda \in \mathbf{K}$ est valeur propre de A si et seulement si $\text{rg}(A - \lambda I_n) < n \Leftrightarrow \text{rg}({}^t A - \lambda I_n) < n$, soit si et seulement si λ est valeur propre de ${}^t A$. On a alors

$$\dim E_\lambda(A) = n - \text{rg}(A - \lambda I_n) = n - \text{rg}({}^t A - \lambda I_n) = \dim E_\lambda({}^t A).$$

□

7.2 DÉTERMINATION PRATIQUE DES VALEURS PROPRES ET DES SOUS-ESPACES PROPRES

7.2.1 Cas général

En pratique, pour déterminer les valeurs propres de $A \in \mathcal{M}_n(\mathbf{K})$, il faut déterminer pour quelles valeurs de λ on a $\text{rg}(A - \lambda I_n) < n$.

Pour cela :

- on écrit la matrice $A - \lambda I_n$, où λ est une indéterminée.
- Puis, à l'aide de l'algorithme du pivot, on transforme $A - \lambda I_n$ en une matrice triangulaire. **Autant que possible, on échangera deux lignes afin de ne plus avoir de λ sur la diagonale, sauf sur le dernier coefficient en bas à droite.** Si ce n'est pas possible, on sera amenés à prendre pour pivot une expression dépendant de λ , qui peut être nulle pour certaines valeurs de λ . Dans ce cas, on distinguera plusieurs cas, suivant les valeurs de λ , afin de ne pas diviser une ligne par 0.
- Une fois la matrice rendue triangulaire, on utilise le fait qu'une matrice triangulaire n'est pas inversible si et seulement si l'un de ses coefficients diagonaux est nul. Et donc la recherche des valeurs de λ telles que $A - \lambda I_n$ ne soit pas inversible se ramène à la résolution d'une (ou plusieurs) équation polynomiale en λ . Si elle est de degré supérieur ou égal à 3, on commencera par chercher des racines évidentes (dans le cas de valeurs propres compliquées, l'énoncé fournira des indications pour trouver au moins une valeur propre).

Remarque

Nous cherchons à calculer le rang d'une matrice, donc le but est de l'échelonner, pas d'obtenir une diagonale de 1.

Exemples 7.12

• $A = \begin{pmatrix} 3 & 4 & 4 \\ 2 & -1 & -2 \\ -2 & 0 & 1 \end{pmatrix} \in \mathcal{M}_3(\mathbf{R})$. Soit $\lambda \in \mathbf{R}$. Alors

$$\begin{aligned} A - \lambda I_3 &= \begin{pmatrix} 3-\lambda & 4 & 4 \\ 2 & -1-\lambda & -2 \\ -2 & 0 & 1-\lambda \end{pmatrix} \xleftrightarrow{L_3 \leftrightarrow L_1} \begin{pmatrix} -2 & 0 & 1-\lambda \\ 2 & -1-\lambda & -2 \\ 3-\lambda & 4 & 4 \end{pmatrix} \\ &\xleftrightarrow{L_2 \leftarrow L_2 + L_1} \begin{pmatrix} -2 & 0 & 1-\lambda \\ 0 & -1-\lambda & -1-\lambda \\ 3-\lambda & 4 & 4 \end{pmatrix} \xleftrightarrow{L_3 \leftarrow 2L_3 + (3-\lambda)L_1} \begin{pmatrix} -2 & 0 & 1-\lambda \\ 0 & -1-\lambda & -1-\lambda \\ 0 & 8 & \lambda^2 - 4\lambda + 11 \end{pmatrix} \\ &\xleftrightarrow{L_3 \leftrightarrow L_2} \begin{pmatrix} -2 & 0 & 1-\lambda \\ 0 & 8 & \lambda^2 - 4\lambda + 11 \\ 0 & -1-\lambda & -1-\lambda \end{pmatrix} \xleftrightarrow{L_3 \leftarrow 8L_3 + (1+\lambda)L_2} \begin{pmatrix} -2 & 0 & 1-\lambda \\ 0 & 8 & \lambda^2 - 4\lambda + 11 \\ 0 & 0 & (\lambda+1)(\lambda^2 - 4\lambda + 3) \end{pmatrix} \end{aligned}$$

Donc $\lambda \in \text{Spec}(A)$ si et seulement si $\text{rg}(A - \lambda I_3) < 3$, donc si et seulement si $\lambda = -1$ ou $\lambda^2 - 4\lambda + 3 = 0 \Leftrightarrow (\lambda - 1)(\lambda - 3) = 0$. Ainsi, $\text{Spec}(A) = \{-1, 1, 3\}$.

• $B = \begin{pmatrix} 4 & -2 & 1 \\ 2 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \in \mathcal{M}_3(\mathbf{R})$. Alors, pour $\lambda \in \mathbf{R}$,

$$B - \lambda I_3 = \begin{pmatrix} 4 - \lambda & -2 & 1 \\ 2 & -\lambda & 0 \\ 0 & 0 & 1 - \lambda \end{pmatrix} \xleftrightarrow{L_1 \leftrightarrow L_2} \begin{pmatrix} 2 & -\lambda & 0 \\ 4 - \lambda & -2 & 1 \\ 0 & 0 & 1 - \lambda \end{pmatrix} \\ \xleftrightarrow{L_2 \leftarrow 2L_2 - (4 - \lambda)L_1} \begin{pmatrix} 2 & -\lambda & 0 \\ 0 & -\lambda^2 + 4\lambda - 4 & 2 \\ 0 & 0 & 1 - \lambda \end{pmatrix}$$

Ainsi, $\lambda \in \text{Spec}(B)$ si et seulement si $\text{rg}(B - \lambda I_3) < 3$, donc si et seulement si $-\lambda^2 + 4\lambda - 4 = 0 \Leftrightarrow (\lambda - 2)^2 = 0$ ou $1 - \lambda = 0$. Donc $\text{Spec}(B) = \{1, 2\}$.

• $C = \begin{pmatrix} 2 & 1 & 0 \\ 0 & 1 & 0 \\ 1 & 1 & 1 \end{pmatrix}$. Pour $\lambda \in \mathbf{R}$, on a

$$C - \lambda I_3 = \begin{pmatrix} 2 - \lambda & 1 & 0 \\ 0 & 1 - \lambda & 0 \\ 1 & 1 & 1 - \lambda \end{pmatrix} \xleftrightarrow{L_1 \leftrightarrow L_3} \begin{pmatrix} 1 & 1 & 1 - \lambda \\ 0 & 1 - \lambda & 0 \\ 2 - \lambda & 1 & 0 \end{pmatrix} \xleftrightarrow{L_3 \leftarrow L_3 - (2 - \lambda)L_1} \begin{pmatrix} 1 & 1 & 1 - \lambda \\ 0 & 1 - \lambda & 0 \\ 0 & -1 + \lambda & (\lambda - 1)(2 - \lambda) \end{pmatrix}$$

Si $\lambda = 1$, alors $\text{rg}(C - I_3) = \text{rg} \begin{pmatrix} 1 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} = 1$. Ainsi, $\lambda = 1$ est valeur propre de C .

Si $\lambda \neq 1$, alors divisons² L_2 et L_3 par $1 - \lambda$, de sorte que

$$C - \lambda I_3 \xleftrightarrow{} \begin{pmatrix} 1 & 1 & 1 - \lambda \\ 0 & 1 & 0 \\ 0 & -1 & \lambda - 2 \end{pmatrix} \xleftrightarrow{L_3 \leftarrow L_3 + L_2} \begin{pmatrix} 1 & 1 & 1 - \lambda \\ 0 & 1 & 0 \\ 0 & 0 & \lambda - 2 \end{pmatrix}.$$

Donc $\text{rg}(C - \lambda I_3) < 3 \Leftrightarrow \lambda = 2$. Ainsi, $\text{Spec}(C) = \{1, 2\}$.

Remarque : notons que si on voulait éviter la distinction de cas, il était possible là où nous avons distingué les deux cas, de réaliser l'opération $L_3 \leftarrow L_3 + L_2 \dots$

Il y a presque toujours une solution évitant la distinction de cas, qu'on n'utilisera qu'en dernier recours !

Rédaction

Cet argument de rang doit apparaître systématiquement dans votre rédaction : c'est lui qui justifie qu'un calcul de pivot permet d'obtenir les valeurs propres.

² C'est ici qu'il est important d'avoir distingué deux cas, sinon on diviserait par 0 !

De plus, cette méthode permet d'obtenir facilement la dimension des sous-espaces propres à l'aide de la formule $\dim E_\lambda(A) = n - \text{rg}(A - \lambda I_n)$.

Dans le premier exemple, si $\lambda \in \{-1, 1, 3\}$, alors $\text{rg}(A - \lambda I_3) = 2$, et donc $\dim E_\lambda(A) = 3 - 2 = 1$.

Dans le second exemple, on a $\dim E_1(A) = 3 - \text{rg}(A - I_3) = 1$ et $\dim E_2(A) = 3 - \text{rg}(A - 2I_3) = 3 - 2 = 1$.

Si $\lambda \in \text{Spec}(A)$, alors déterminer $E_\lambda(A)$ revient à résoudre le système

$$AX = \lambda X \Leftrightarrow (A - \lambda I_n)X = 0.$$

Par exemple, dans le cas de $A = \begin{pmatrix} 4 & -2 & 1 \\ 2 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$, on a

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} \in E_2(A) \Leftrightarrow (A - 2I_3) \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \Leftrightarrow \begin{pmatrix} 2 & -2 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \\ \Leftrightarrow \begin{cases} 2x - 2y = 0 \\ 2z = 0 \\ -z = 0 \end{cases} \Leftrightarrow \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x \\ x \\ 0 \end{pmatrix}$$

Ainsi, $E_2(A) = \text{Vect} \left(\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} \right)$. Notons qu'on retrouve bien $\dim E_2(A) = 1$.

Inconnue

L'inconnue est alors $X \in \mathcal{M}_{n,1}(\mathbf{K})$.

7.2.2 Cas particuliers : matrices 2 × 2, matrices triangulaires et diagonales

Rappel : Si $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathcal{M}_2(\mathbf{K})$, alors A est inversible si et seulement si $ad - bc \neq 0$.

Cette proposition permet d'éviter le recours à un pivot pour déterminer les valeurs propres d'une matrice 2 × 2. En effet, pour $\lambda \in \mathbf{K}$, on a alors

$$\lambda \in \text{Spec}(A) \Leftrightarrow A - \lambda I_2 \text{ n'est pas inversible} \Leftrightarrow \det(A - \lambda I_2) = 0.$$

Exemple 7.13

Soit $A = \begin{pmatrix} 5 & -1 \\ 6 & 0 \end{pmatrix} \in \mathcal{M}_2(\mathbf{R})$.

Alors λ est valeur propre de A si et seulement si

$$\det(A - \lambda I_2) = \det \begin{pmatrix} 5 - \lambda & -1 \\ 6 & -\lambda \end{pmatrix} = 0 \Leftrightarrow \lambda(\lambda - 5) + 6 = 0 \Leftrightarrow \lambda^2 - 5\lambda + 6 = 0.$$

On a alors affaire à un polynôme de degré 2 dont les racines sont³ 2 et 3. Et donc $\text{Spec}(A) = \{2, 3\}$.

Terminologie

Le scalaire $ad - bc$ est appelé déterminant de A et est noté $\det(A)$.

³ Après un calcul de discriminant.

Proposition 7.14 : Si $A \in \mathcal{M}_n(\mathbf{K})$ est une matrice triangulaire, alors ses valeurs propres sont exactement ses coefficients diagonaux.

Démonstration. Notons $A = \begin{pmatrix} \lambda_1 & \star & \dots & \star \\ 0 & \lambda_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \star \\ 0 & \dots & 0 & \lambda_n \end{pmatrix}$ une matrice triangulaire supérieure. Alors,

pour $\lambda \in \mathbf{K}$, on a

$$A - \lambda I_n = \begin{pmatrix} \lambda_1 - \lambda & \star & \dots & \star \\ 0 & \lambda_2 - \lambda & \ddots & \vdots \\ \vdots & \ddots & \ddots & \star \\ 0 & \dots & 0 & \lambda_n - \lambda \end{pmatrix}.$$

Ainsi, $\text{rg}(A - \lambda I_n) < n$ si et seulement si $\lambda \in \{\lambda_1, \dots, \lambda_n\}$.

Par conséquent, $\text{Spec}(A) = \{\lambda_1, \dots, \lambda_n\}$.

Si A est une matrice triangulaire inférieure, alors A et ${}^t A$ ont les mêmes valeurs propres.

Mais ${}^t A$ est une matrice triangulaire supérieure, donc ses valeurs propres sont ses coefficients diagonaux, qui sont les mêmes que ceux de A . □

Notation : lorsqu'on écrit \star , cela signifie que le coefficient en question peut prendre n'importe quelle valeur.

Remarque

Il se peut que plusieurs des λ_i soient égaux, donc il peut y avoir strictement moins de n valeurs propres (et même une seule si $\lambda_1 = \dots = \lambda_n$).

Exemple 7.15

Si $A = \begin{pmatrix} -2 & 1 & 3 \\ 0 & 0 & 4 \\ 0 & 0 & 5 \end{pmatrix}$, alors les valeurs propres de A sont $-2, 0$ et 5 .

Proposition 7.16 : Si $A \in \mathcal{M}_n(\mathbf{K})$ est triangulaire, et si $\lambda \in \text{Spec}(A)$, alors la dimension du sous-espace propre $E_\lambda(A)$ est inférieure ou égale au nombre de fois où λ apparaît sur la diagonale de A .

Danger !

On a bien une inégalité, et non une égalité !

Démonstration. Soit $\lambda \in \text{Spec}(A)$. Supposons que λ apparaisse k fois sur la diagonale de A . Alors $A - \lambda I_n$ possède exactement k coefficients diagonaux nuls, et donc $n - k$ coefficients diagonaux non nuls. Ainsi, $\text{rg}(A - \lambda I_n) \geq n - k$. On en déduit que

$$\dim E_\lambda(A) = n - \text{rg}(A - \lambda I_n) \leq n - (n - k) = k.$$

□

Exemple 7.17

Si $A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 2 \\ 0 & 0 & 2 \end{pmatrix}$, alors les valeurs propres de A sont 1 et 2.

De plus, $\dim E_1(A) = 1$ et $1 \leq \dim E_2(A) \leq 2$. Plus précisément, on a $\dim E_2(A) =$

$$3 - \text{rg}(A - 2I_3) = 3 - \text{rg} \begin{pmatrix} -1 & 0 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{pmatrix} = 3 - 2 = 1.$$

Proposition 7.18 : Si A est diagonale, alors les valeurs propres de A sont exactement les coefficients diagonaux de A . De plus, si $\lambda \in \text{Spec}(A)$, alors la dimension de $E_\lambda(A)$ est égale au nombre de fois où λ apparaît sur la diagonale de A .

Démonstration. Le fait que les valeurs propres soient les coefficients diagonaux découle de la proposition précédente.

Si λ apparaît k fois sur la diagonale de A , alors $A - \lambda I_n$ est diagonale, avec exactement $n - k$ coefficients non nuls : elle est de rang $n - k$, et donc $\dim E_\lambda(A) = n - (n - k) = k$. □

Exemple 7.19

Si $A = \begin{pmatrix} 2 & 0 & 0 & 0 \\ 0 & -3 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -3 \end{pmatrix}$, alors $\text{Spec}(A) = \{-3, 1, 2\}$ et

$$\dim E_1(A) = \dim E_2(A) = 1 \text{ et } \dim E_{-3}(A) = 2.$$

7.3 VALEURS PROPRES ET VECTEURS PROPRES D'UN ENDOMORPHISME

Dans toute la suite, E désigne un espace vectoriel sur \mathbf{K} de dimension finie n , et f est un endomorphisme de E .

7.3.1 Définition

Définition 7.20 – Un scalaire $\lambda \in \mathbf{K}$ est appelé valeur propre de f s'il existe $x \in E$ non nul tel que $f(x) = \lambda x$.

On note $\text{Spec}(f)$ l'ensemble des valeurs propres de f .

Définition 7.21 – Si $\lambda \in \text{Spec}(f)$, un vecteur non nul $x \in E$ tel que $f(x) = \lambda x$ est appelé vecteur propre de f associé à la valeur propre λ .

Le sous-espace propre associé à la valeur propre λ est $E_\lambda(f) = \{x \in E : f(x) = \lambda x\}$. Il est formé du vecteur nul et de l'ensemble des vecteurs propres de f associés à la valeur propre λ .

Remarque

On a même directement les sous-espaces propres, car on remarque que E_1 est un vecteur propre pour la valeur propre 2, et puisque $\dim E_1(A) = 1$, il vient alors

$$E_1(A) = \text{Vect} \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}.$$

De même, E_2 et E_4 sont des vecteurs propres de A pour la valeur propre -3 et $\dim E_{-3}(A) = 2$ donc

$$\begin{aligned} E_{-3}(A) &= \text{Vect}(E_2, E_4) \\ &= \text{Vect} \left(\begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} \right). \end{aligned}$$

De manière générale, si la j -ème colonne de A est nulle, sauf peut-être son coefficient diagonal noté λ , alors λ est valeur propre de A et $E_j \in E_\lambda(A)$.

Hypothèses

Un vecteur propre est non nul par définition. Notons qu'on a toujours

$$f(0_E) = 0_E = \lambda \cdot 0_E.$$

Exemple 7.22

Soit f l'endomorphisme de $\mathbf{R}_2[X]$ défini par $f(P) = P'$.

Alors 0 est valeur propre de f car $f(1) = 0 = 0 \times 1$.

De plus, si $\lambda \neq 0$, alors $\lambda \notin \text{Spec}(f)$.

En effet, supposons qu'il existe $P \neq 0$ tel que $f(P) = P' = \lambda P$.

Alors $\deg P' = \deg(\lambda P) = \deg P$. Ceci est impossible si $P \neq 0$, donc $\lambda \notin \text{Spec}(f)$.

7.3.2 Étude des sous-espaces propres

Proposition 7.23 : Soit $\lambda \in \mathbf{K}$. Alors il y a équivalence entre

1. λ est valeur propre de f
2. $f - \lambda \text{id}_E$ n'est pas injectif
3. $f - \lambda \text{id}_E$ n'est pas bijectif
4. $\text{rg}(f - \lambda \text{id}_E) < \dim E$

Démonstration. λ est valeur propre de f si et seulement si il existe $x \in E$ non nul tel que $f(x) = \lambda x \Leftrightarrow (f - \lambda \text{id}_E)(x) = 0$.

Ainsi, λ est valeur propre de f si et seulement si $\text{Ker}(f - \lambda \text{id}_E) \neq \{0\}$, soit si et seulement si $f - \lambda \text{id}_E$ n'est pas injectif. Donc 1) \Leftrightarrow 2).

Puisque $f - \lambda \text{id}_E$ est un endomorphisme de E , qui est de dimension finie, $f - \lambda \text{id}_E$ est injectif si et seulement si il est bijectif. Donc 2) \Leftrightarrow 3).

Enfin, par le théorème du rang appliqué à $f - \lambda \text{id}_E$, $f - \lambda \text{id}_E$ est injectif si et seulement si $\text{Ker}(f - \lambda \text{id}_E) = \{0\} \Leftrightarrow \text{rg}(f - \lambda \text{id}_E) = \dim E$. Donc par contraposition, 2) \Leftrightarrow 4). \square

Corollaire 7.24 – En particulier, 0 est valeur propre de f si et seulement si f n'est pas injectif.

Proposition 7.25 : Soit $\lambda \in \text{Spec}(f)$. Alors $E_\lambda(f) = \text{Ker}(f - \lambda \text{id}_E)$.
Par conséquent⁴, $E_\lambda(f)$ est un sous-espace vectoriel de E .

⁴ Le noyau d'un endomorphisme de E est toujours un sous-espace vectoriel de E .

Démonstration. Soit $x \in E$. Alors

$$x \in E_\lambda(f) \Leftrightarrow f(x) = \lambda x \Leftrightarrow (f - \lambda \text{id}_E)(x) = 0 \Leftrightarrow x \in \text{Ker}(f - \lambda \text{id}_E).$$

\square

Remarque. En particulier, si 0 est valeur propre de f , alors $E_0(f) = \text{Ker}(f)$.

Proposition 7.26 : Si $\lambda \in \text{Spec}(f)$, alors $E_\lambda(f)$ est stable par f .

Démonstration. Soit $x \in E_\lambda(f)$. Alors $f(x) = \lambda \underbrace{x}_{\in E_\lambda(f)} \in E_\lambda(f)$.

Donc $E_\lambda(f)$ est stable par f . \square

Proposition 7.27 : Si $\lambda \in \text{Spec}(f)$, alors $1 \leq \dim E_\lambda(f) \leq \dim E$.
Plus précisément, on a

$$\dim E_\lambda(f) = \dim E - \text{rg}(f - \lambda \text{id}_E).$$

Démonstration. Commençons par remarquer que si $\lambda \in \text{Spec}(f)$, alors il existe x non nul dans $E_\lambda(f)$. Donc $E_\lambda(f)$ n'est pas réduit au vecteur nul, et par conséquent, $\dim E_\lambda(f) \geq 1$. Puisque $E_\lambda(f)$ est un sous-espace vectoriel de E , il est évident que $\dim E_\lambda(f) \leq \dim E$. Enfin, par le théorème du rang appliqué à $f - \lambda \text{id}_E$, notons que

$$\dim E_\lambda(f) = \dim \text{Ker}(f - \lambda \text{id}_E) = \dim E - \text{rg}(f - \lambda \text{id}_E) = \dim E - \text{rg}(f - \lambda \text{id}_E).$$

□

7.3.3 Lien entre matrices et endomorphismes

Soit $f \in \mathcal{L}(E)$, soit \mathcal{B} une base de E et soit $A = \text{Mat}_{\mathcal{B}}(f) \in \mathcal{M}_n(\mathbf{K})$.

Alors un scalaire $\lambda \in \mathbf{K}$ est valeur propre de f si et seulement si λ est valeur propre de A . En effet, $\lambda \in \text{Spec}(f)$ si et seulement si $\text{rg}(f - \lambda \text{id}_E) < n$. Or, la matrice de l'endomorphisme $f - \lambda \text{id}_E$ dans la base \mathcal{B} est $A - \lambda I_n$. Donc

$$\lambda \in \text{Spec}(f) \Leftrightarrow \text{rg}(f - \lambda \text{id}_E) < n \Leftrightarrow \text{rg}(A - \lambda I_n) < n \Leftrightarrow \lambda \in \text{Spec}(A).$$

De plus, un vecteur $x \in E$ est dans $E_\lambda(f)$ si et seulement si le vecteur colonne de ses coordonnées dans la base \mathcal{B} est vecteur propre de A . En effet, si on note $X = x_{\mathcal{B}}$, alors

$$x \in E_\lambda(f) \Leftrightarrow f(x) = \lambda x \Leftrightarrow AX = \lambda X \Leftrightarrow X \in E_\lambda(A).$$

Exemple 7.28

Soit $f : \mathbf{R}_2[X] \rightarrow \mathbf{R}_2[X]$ défini par $f(P) = P'$, et soit $\mathcal{B} = (1, X, X^2)$ la base canonique de $\mathbf{R}_2[X]$. Alors

$$A = \text{Mat}_{\mathcal{B}}(f) = \begin{pmatrix} f(1) & f(X) & f(X^2) \\ 0 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{pmatrix} \begin{matrix} 1 \\ X \\ X^2 \end{matrix}$$

A est triangulaire, donc ses valeurs propres sont exactement ses coefficients diagonaux. Ainsi, $\text{Spec}(A) = \{0\}$, et donc $\text{Spec}(f) = \{0\}$.

De plus, on a $\dim E_0(A) = 3 - \text{rg}(A - 0 \text{ cot } I_3) = 1$.

Enfin, $X = \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in E_0(A)$ si et seulement si $AX = 0 \Leftrightarrow \begin{cases} y = 0 \\ 2z = 0 \end{cases}$ de sorte que

$$E_0(A) = \text{Vect} \left(\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \right).$$

Or, l'unique polynôme de $\mathbf{R}_2[X]$ dont les coordonnées dans la base \mathcal{B} sont $(1, 0, 0)$ est $P = 1$, donc $E_0(f) = \text{Vect}(1) = \mathbf{R}_0[X]$.

Remarque

Nous avons déjà prouvé un peu plus tôt que l'unique valeur propre de f était 0, mais cela avait nécessité un petit argument de degré. Si l'on passe par la matrice, on retrouve directement ce résultat en calculant les valeurs propres de A , ce qui ne nécessite aucune astuce.

Remarquons que si $\mathcal{B} = (e_1, \dots, e_n)$ est une base de E , alors e_i est vecteur propre de f associé à la valeur propre λ si et seulement si la i -ème colonne de $\text{Mat}_{\mathcal{B}}(f)$ est nulle, à l'exception du coefficient diagonal qui vaut λ .

Exemple 7.29

Si $\text{Mat}_{\mathcal{B}}(f) = \begin{pmatrix} f(e_1) & f(e_2) & f(e_3) \\ 2 & 0 & 3 \\ -1 & 4 & 8 \\ 4 & 0 & 0 \end{pmatrix} \begin{matrix} e_1 \\ e_2 \\ e_3 \end{matrix}$, alors $4 \in \text{Spec}(f)$, et $e_2 \in E_4(f)$.

En effet, on a $e_2 \neq 0$ et $f(e_2) = 4e_2$.

Proposition 7.30 : Deux matrices semblables ont les mêmes valeurs propres et des sous-espaces propres de même dimension.

Démonstration. Si A et B sont semblables, alors⁵ elle représentent un même endomorphisme $f \in \mathcal{L}(\mathbf{R}^n)$ dans deux bases différentes. Mais alors les valeurs propres de A sont les valeurs propres de f , et si $\lambda \in \text{Spec}(A)$, alors $\dim E_\lambda(A) = \dim E_\lambda(f)$. De même pour les valeurs propres de B . □

Exemple 7.31

Soient $A = \begin{pmatrix} 2 & 1 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix}$ et $B = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix}$.

Alors A et B ont même rang, même trace, et ont toutes deux une unique⁶ valeur propre 2. Mais on a

$$\dim E_2(A) = 3 - \text{rg}(A - 2I_3) = 3 - \text{rg} \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} = 3 - 2 = 1$$

alors que

$$\dim E_2(B) = 3 - \text{rg}(B - 2I_3) = 3 - \text{rg} \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} = 3 - 1 = 2.$$

Donc A et B ne sont pas semblables.

Corollaire 7.32 – Soit p un projecteur de $\mathcal{L}(E)$. Alors $\text{Spec}(p) \subset \{0, 1\}$. De plus, si $p \neq 0$ et $p \neq \text{id}$, alors $\text{Spec}(p) = \{0, 1\}$.

Démonstration. Notons $n = \dim E$, et soit \mathcal{B} une base de E . Alors nous savons⁷ que $A = \text{Mat}_{\mathcal{B}}(p)$ est semblable à une matrice diagonale D ne comportant que des 0 et des 1 sur la diagonale.

Et donc les valeurs propres de p sont les valeurs propres de D : ce sont 0 et/ou 1. De plus, si $p \neq 0$ et $p \neq \text{id}$, alors $0 < \text{rg}(p) < n$, et donc D possède exactement $\text{rg}(p)$ fois 1 sur sa diagonale et $n - \text{rg}(p)$ fois 0. Donc 0 et 1 sont tous les deux valeurs propres de D , et donc de A . □

7.3.4 Somme directe de sous-espaces propres et conséquences

Dans toute la suite, nous énoncerons les propriétés des sous-espaces propres en termes d'endomorphismes, mais toutes se traduisent en termes de matrices car pour $A \in \mathcal{M}_n(\mathbf{K})$, les sous-espaces propres de A sont exactement ceux de l'endomorphisme $f_A \in \mathcal{L}(\mathcal{M}_{n,1}(\mathbf{K}))$.

Lemme 7.33. Soit $p \in \mathbf{N}^*$, soient $\lambda_1, \dots, \lambda_p$ des éléments de \mathbf{K} deux à deux distincts. Alors pour tout $i \in \llbracket 1, p \rrbracket$ il existe un (unique) polynôme $P_i \in \mathbf{K}_{p-1}[X]$ tel que

$$\forall j \in \llbracket 1, p \rrbracket, P_i(\lambda_j) = \begin{cases} 1 & \text{si } i = j \\ 0 & \text{sinon} \end{cases}$$

Démonstration. Soit $\varphi : \mathbf{K}_{p-1}[X] \rightarrow \mathbf{K}^p$ l'application définie par

$$\varphi(P) = (P(\lambda_1), \dots, P(\lambda_p)).$$

Il n'est pas difficile de prouver que φ est linéaire.

⚠ Réciproque fautive
 Une fois de plus la réciproque n'est pas vraie. Il est possible de trouver deux matrices A et B avec :

- même rang
- même trace
- mêmes valeurs propres
- des sous-espaces propres de même dimension

mais telles que A et B ne soient pas semblables.

⁵ Voir par exemple la preuve de la proposition 5.36.

⁶ A et B sont triangulaires, donc leurs valeurs propres sont leurs coefficients diagonaux.

⁷ C'est la proposition 5.38.

Proj. dégénérés
 Si $p = 0$, alors p est la projection sur $\{0\}$ parallèlement à E et si $p = \text{id}$, alors p est la projection sur E parallèlement à $\{0\}$. Ce sont bien deux projecteurs, mais pas nécessairement des plus intéressants.

Soit $P \in \text{Ker}(\varphi)$. Alors $(P(\lambda_1), \dots, P(\lambda_p)) = (0, \dots, 0)$ et donc P possède p racines distinctes $\lambda_1, \dots, \lambda_p$.

Mais P étant de degré au plus $p - 1$, c'est nécessairement le polynôme nul.

Ainsi, φ est injective, et puisque $\dim \mathbf{K}_{p-1}[X] = \dim \mathbf{K}^p = p$, φ est bijective. Par conséquent il existe un unique polynôme $P_i \in \mathbf{K}_{p-1}[X]$ tel que $\varphi(P_i) = (0, \dots, 0, 1, 0, \dots, 0)$, et donc un unique polynôme P_i tel que

$$P_i(\lambda_j) = \begin{cases} 1 & \text{si } i = j \\ 0 & \text{sinon} \end{cases}$$

□

Lemme 7.34. Soit $f \in \mathcal{L}(E)$, soit $\lambda \in \text{Spec}(f)$, et soit $P \in \mathbf{K}[X]$. Alors,

$$\forall x \in E_\lambda(f), P(f)(x) = P(\lambda)x.$$

Démonstration. Commençons par montrer le résultat par récurrence sur i pour les monômes de la forme X^i , autrement dit, que pour tout $i \in \mathbf{N}$, $f^i(x) = \lambda^i x$.

Il est clair que pour $i = 0$, $f^0(x) = \text{id}(x) = x = \lambda^0 x$.

Supposons à présent que pour $f^i(x) = \lambda^i x$. Alors

$$f^{i+1}(x) = f(f^i(x)) = f(\lambda^i x) = \lambda^i f(x) = \lambda^i \lambda x = \lambda^{i+1}(x).$$

Par le principe de récurrence, la propriété est vraie pour tout $i \in \mathbf{N}$.

Soit à présent $P = \sum_{i=0}^d a_i X^i$. Alors

$$P(f)(x) = \sum_{i=0}^d a_i f^i(x) = \sum_{i=0}^d a_i \lambda^i x = \left(\sum_{i=0}^d a_i \lambda^i \right) x = P(\lambda)x.$$

□

Le résultat qui suit est fondamental, notamment en raison de ses nombreux corollaires, souvent très utiles.

Proposition 7.35 : Soit $f \in \mathcal{L}(E)$, et soient $\lambda_1, \dots, \lambda_p \in \text{Spec}(f)$ des valeurs propres deux à deux distinctes de f . Alors la somme $\sum_{i=1}^p E_{\lambda_i}(f)$ est directe.

Démonstration. Soient $x_1 \in E_{\lambda_1}(f), \dots, x_p \in E_{\lambda_p}(f)$ tels que $x_1 + \dots + x_p = 0_E$.

Soit $i \in \llbracket 1, p \rrbracket$, et soit $P_i \in \mathbf{K}_{p-1}[X]$ tel que

$$\forall j \in \llbracket 1, p \rrbracket, P_i(\lambda_j) = \begin{cases} 1 & \text{si } i = j \\ 0 & \text{sinon} \end{cases}$$

Alors, en appliquant l'endomorphisme $P_i(f)$ à la relation $x_1 + \dots + x_p = 0_E$, il vient

$$\begin{aligned} P_i(f)(x_1 + \dots + x_p) = 0 &\Leftrightarrow P_i \left(\sum_{j=1}^p x_j \right) = 0 \Rightarrow \sum_{j=1}^p P_i(f)(x_j) = 0 \\ &\Rightarrow \sum_{j=1}^p P_i(\lambda_j) x_j = 0 \Rightarrow \underbrace{P_i(\lambda_i)}_{=1} x_i = 0 \\ &\Rightarrow x_i = 0 \end{aligned}$$

Ainsi, $\forall i \in \llbracket 1, p \rrbracket, x_i = 0$: la somme est bien directe. □

Corollaire 7.36 – Soit E un espace vectoriel de dimension finie n , et soit $f \in \mathcal{L}(E)$. Alors f possède au plus $\dim E$ valeurs propres.

Remarque

Cette preuve est non constructive : elle garantit l'existence d'un tel polynôme mais ne dit absolument pas comment le trouver. Il existe néanmoins des formules pour construire les P_i . On appelle ces polynômes les polynômes de Lagrange.

Remarque

Ainsi, si λ est une valeur propre de f et que x est un vecteur propre associé, alors $P(\lambda)$ est une valeur propre de $P(f)$ et x est un vecteur propre de $P(f)$ pour la valeur propre $P(\lambda)$.

⚠ Attention !

La somme directe des sous-espaces propres de f n'est pas nécessairement égale à E tout entier

Les $P_i(\lambda_j)$ sont nuls si $i \neq j$, donc il ne reste qu'un seul terme.

Démonstration. Raisonnons par l'absurde en supposant que f possède $n + 1$ valeurs propres distinctes $\lambda_1, \dots, \lambda_{n+1}$. Posons alors

$$F = \bigoplus_{i=1}^{n+1} E_{\lambda_i}(f).$$

F est un sous-espace vectoriel de E , et⁸

$$\dim F = \sum_{i=1}^{n+1} \dim E_{\lambda_i}(f) \geq \sum_{i=1}^{n+1} 1 = n + 1.$$

Ceci est impossible car E est de dimension n , donc f possède au maximum n valeurs propres distinctes. □

Corollaire 7.37 – Soit $A \in \mathcal{M}_n(\mathbf{K})$. Alors A possède au plus n valeurs propres.

Remarque. Il se peut bien évidemment que A possède strictement moins de n valeurs propres. Par exemple $A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ possède une seule valeur propre 1.

Et $B = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \in \mathcal{M}_2(\mathbf{R})$ ne possède aucune valeur propre !

En effet, $\lambda \in \mathbf{R}$ est valeur propre de B si et seulement si

$$\det(B - \lambda I_2) = 0 \Leftrightarrow \det \begin{pmatrix} -\lambda & -1 \\ 1 & -\lambda \end{pmatrix} = 0 \Leftrightarrow \lambda^2 + 1 = 0.$$

Cette équation ne possède pas de solutions réelles, donc B ne possède pas de valeurs propres réelles.

▲ Réels/complexes

Notons que si l'on voit B comme matrice de $\mathcal{M}_2(\mathbf{C})$, alors B possède deux valeurs propres (complexes) qui sont i et $-i$.

Corollaire 7.38 – Soit E un espace vectoriel de dimension finie et soit $f \in \mathcal{L}(E)$. Alors

$$\sum_{\lambda \in \text{Spec}(f)} \dim E_{\lambda}(f) \leq \dim E.$$

Démonstration. $\bigoplus_{\lambda \in \text{Spec}(f)} E_{\lambda}(f)$ est un sous-espace vectoriel de dimension $\sum_{\lambda \in \text{Spec}(f)} \dim E_{\lambda}(f)$. □

Corollaire 7.39 – Soit $A \in \mathcal{M}_n(\mathbf{K})$. Alors $\sum_{\lambda \in \text{Spec}(A)} \dim E_{\lambda}(A) \leq n$.

Démonstration. C'est le corollaire précédent, appliqué à l'endomorphisme $f_A \in \mathcal{L}(\mathcal{M}_{n,1}(\mathbf{K}))$. □

Proposition 7.40 : Soit $f \in \mathcal{L}(E)$, et soient $\lambda_1, \dots, \lambda_p$ des valeurs propres de f deux à deux distinctes, et pour tout $i \in \llbracket 1, p \rrbracket$ soit x_i un vecteur propre associé à la valeur propre λ_i .

Alors (x_1, \dots, x_p) est une famille libre de E .

Démonstration. Soient $\mu_1, \dots, \mu_p \in \mathbf{K}$ tels que $\mu_1 x_1 + \dots + \mu_p x_p = 0$.

Alors $\mu_1 x_1 \in E_{\lambda_1}(f), \dots, \mu_p x_p \in E_{\lambda_p}(f)$.

Puisque les sous-espaces propres sont en somme directe, nécessairement

$$\mu_1 x_1 = \dots = \mu_p x_p = 0.$$

Mais les x_i étant des vecteurs propres, ils sont non nuls, donc $\mu_1 = 0, \dots, \mu_p = 0$: la famille est libre. \square

Proposition 7.41 : Si f possède n (où $n = \dim E$) valeurs propres distinctes, alors les sous-espaces propres associés sont tous de dimension 1. Dans ce cas, la concaténation de bases des sous-espaces propres forme une base de E .

Démonstration. Notons $\lambda_1, \dots, \lambda_n$ les valeurs propres de f . Nous savons que $\forall i \in \llbracket 1, n \rrbracket$, $\dim E_{\lambda_i}(f) \geq 1$. Si l'un des sous-espaces propres était de dimension supérieure strictement à 1, alors $F = \bigoplus_{i=1}^n E_{\lambda_i}(f)$ serait un sous-espace vectoriel de E vérifiant

$$\dim F = \sum_{i=1}^n \dim E_{\lambda_i}(f) \geq n + 1$$

ce qui est impossible. Donc tous les λ_i sont de dimension 1.

Par la proposition précédente, la concaténation de bases des sous-espaces propres forme une famille libre de E . Mais cette famille est de cardinal $n = \dim E$: c'est donc une base de E . \square

Corollaire 7.42 – Si $A \in \mathcal{M}_n(\mathbf{K})$ possède n valeurs propres distinctes, alors tous ses sous-espaces propres sont de dimension 1. En particulier, il existe une base de $\mathcal{M}_{n,1}(\mathbf{K})$ formée de vecteurs propres de A .

Exemple 7.43

Soit $A = \begin{pmatrix} 2 & 1 & -4 \\ 0 & 0 & 4 \\ 0 & 0 & 5 \end{pmatrix}$. Alors A possède trois valeurs propres distinctes : 0, 2 et 5.

Par conséquent, les trois sous-espaces propres associés sont de dimension un.

Remarque

Dans une base de E formée de vecteurs propres, la matrice de f est alors diagonale, et possède les valeurs propres de f sur sa diagonale.

Remarque

La matrice de f_A dans cette base de vecteurs propres est alors diagonale, et donc A est semblable à une matrice diagonale. Nous consacrerons un chapitre ultérieur à essayer de donner d'autres conditions suffisantes pour qu'une matrice soit semblable à une matrice diagonale.

7.4 COMPLÉMENT SUR LE RANG ET LES MATRICES ÉCHELONNÉES

Rappelons ici quelques faits (vus en première année) sur les matrices échelonnées, les matrices triangulaires et le calcul du rang. Il s'agit surtout d'éclairer ces résultats à la lumière de ce que vous avez depuis appris en algèbre linéaire.

7.4.1 Matrices échelonnées

Définition 7.44 – Soit $A \in \mathcal{M}_{n,p}(\mathbf{K})$. On dit que A est **échelonnée** si chaque ligne non nulle commence par strictement plus de zéros que la précédente.

Exemple 7.45

Les matrices $A = \begin{pmatrix} 1 & 2 & -1 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$ et $B = \begin{pmatrix} 1 & 0 & 2 & -4 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}$ sont échelonnées, la matrice

$B = \begin{pmatrix} 2 & 1 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \end{pmatrix}$ ne l'est pas.

Remarques. • En particulier, si une matrice échelonnée possède des lignes nulles, alors celles-ci se trouvent tout en bas de la matrice, car une ligne non nulle ne peut se trouver sous une ligne nulle.

- Si la matrice est rectangulaire avec $n > p$, et si A est échelonnée, alors nécessairement les $n - p$ dernières lignes de A sont nulles. En effet, la première ligne a au maximum p coefficients non nuls, la deuxième au maximum $p - 1$, etc, et la p -ème ligne a au plus un coefficient non nul, et donc les suivantes sont nécessairement nulles.

Il se peut toutefois que davantage de lignes soient nulles, c'est par exemple le cas de

la matrice échelonnée $\begin{pmatrix} 1 & 2 \\ 0 & 0 \\ 0 & 0 \end{pmatrix} \in \mathcal{M}_{3,2}(\mathbf{R})$.

Définition 7.46 – Si $A \in \mathcal{M}_{n,p}(\mathbf{K})$ est échelonnée, alors on appelle **pivot** le premier terme non nul d'une ligne non nulle.

Exemples 7.47

Sur les matrices échelonnées suivantes, les pivots sont entourés :

$$\begin{pmatrix} 0 & \boxed{2} & 1 \\ 0 & 0 & \boxed{1} \\ 0 & 0 & 0 \end{pmatrix}, \begin{pmatrix} \boxed{2} & 0 & 0 \\ 0 & \boxed{1} & 0 \\ 0 & 0 & \boxed{-4} \\ 0 & 0 & 0 \end{pmatrix}, \begin{pmatrix} \boxed{2} & 0 & 0 \\ 0 & 0 & \boxed{-1} \\ 0 & 0 & 0 \end{pmatrix}$$

Théorème 7.48 (Méthode du pivot) : Toute matrice peut être transformée en matrice échelonnée à l'aide d'opérations sur les lignes.

Nous ne donnons pas de preuve de ce fait, mais la preuve est algorithmique, et vous connaissez l'algorithme du pivot : étant donnée une matrice, vous disposez d'une méthode pour la transformer en matrice échelonnée à l'aide d'opérations sur les lignes.

7.4.2 Rang d'une matrice

Rappelons que le rang d'une matrice est défini comme le rang de la famille de ses vecteurs colonnes, mais nous savons également⁹ qu'il s'agit du rang de la famille de ses vecteurs lignes (résultat admis en ECS, mais dont la démonstration serait tout de même abordable, bien qu'un peu longue).

⁹ C'est une conséquence du fait que $\text{rg}(A) = \text{rg}({}^t A)$.

Rang d'une matrice échelonnée

Proposition 7.49 : Soit $A \in \mathcal{M}_{n,p}(\mathbf{K})$ une matrice échelonnée. Alors le rang de A est égal au nombre de pivots de A .

Démonstration. Notons r le nombre de pivots de A . D'après les définitions qui précèdent, cela signifie que A possède r lignes non nulles (les r premières), et que chacune commence par strictement plus de zéros que la précédente.

Notons L_1, \dots, L_r ces lignes (qui sont dans \mathbf{K}^p). Alors le rang de A est la dimension de $\text{Vect}(L_1, \dots, L_r) = \text{Vect}(L_1, \dots, L_r)$, et est nécessairement inférieur ou égal à r .

Il reste à voir que (L_1, \dots, L_r) est une famille libre : soient donc $\lambda_1, \dots, \lambda_r$ des scalaires tels que

$$\lambda_1 L_1 + \dots + \lambda_r L_r = 0.$$

Alors en regardant le premier coefficient non nul de L_1 , on voit que nécessairement $\lambda_1 = 0$ (car les lignes suivantes commencent par, rappelons-le, strictement plus de zéros que L_1). Puis de proche en proche, on prouve que $\lambda_2 = \dots = \lambda_r = 0$.

Alors (L_1, \dots, L_r) est une base de $\text{Vect}(L_1, \dots, L_r)$, et donc le rang de A est exactement r . \square

Vect

Enlever les lignes nulles ne change pas l'espace vectoriel qu'elles engendrent.

Rang d'une matrice quelconque : la méthode du pivot

Commençons par rappeler un fait important : composer une application linéaire à gauche par une application linéaire bijective ne change pas son rang.

Proposition 7.50 : Soient E et F deux espaces vectoriels de dimension finie, $f \in \mathcal{L}(E, F)$ et $g \in \mathcal{L}(F)$ un isomorphisme.
Alors $\text{rg}(g \circ f) = \text{rg}(f)$.

Démonstration. Par définition, $\text{rg}(f)$ est la dimension de $\text{Im } f$.

Soit donc $f(e_1), \dots, f(e_p)$ une base de $\text{Im}(f)$, et montrons que $(g \circ f)(e_1), \dots, (g \circ f)(e_p)$ est une base de $\text{Im}(g \circ f)$.

Soit $y \in \text{Im}(g \circ f)$. Alors, il existe $x \in E$ tel que $y = (g \circ f)(x)$.

Mais alors $f(x) \in \text{Im } f$ et donc il existe des scalaires $\lambda_1, \dots, \lambda_p$ tels que $f(x) = \sum_{i=1}^p \lambda_i f(e_i)$.

Et donc $y = g(f(x)) = \sum_{i=1}^p \lambda_i (g \circ f)(e_i)$.

Ainsi, $(g \circ f)(e_1), \dots, (g \circ f)(e_p)$ est génératrice de $\text{Im}(g \circ f)$.

Prouvons qu'elle est libre : soient $\lambda_1, \dots, \lambda_p$ des scalaires tels que $\sum_{i=1}^p \lambda_i (g \circ f)(e_i) = 0$.

Alors $g\left(\sum_{i=1}^p \lambda_i f(e_i)\right) = 0$, et puisque g est injective, on a donc $\sum_{i=1}^p \lambda_i f(e_i) = 0$.

La famille $f(e_1), \dots, f(e_p)$ étant libre, il vient donc $\lambda_1 = \dots = \lambda_p = 0$.

Et donc $(g \circ f)(e_1), \dots, (g \circ f)(e_p)$ est libre : c'est une base de $\text{Im}(g \circ f)$.

On a donc bien

$$\text{rg}(g \circ f) = \dim \text{Im}(g \circ f) = p = \dim \text{Im}(f) = \text{rg}(f).$$

□

La traduction matricielle de cet énoncé est alors la suivante :

Proposition 7.51 : Soit $A \in \mathcal{M}_{n,p}(\mathbf{K})$, et soit $P \in GL_n(\mathbf{K})$ une matrice inversible. Alors $\text{rg}(PA) = \text{rg}(A)$.

Démonstration. Soit $f \in \mathcal{L}(\mathbf{K}^p, \mathbf{K}^n)$ l'unique application linéaire dont la matrice dans les bases canoniques de \mathbf{K}^p et \mathbf{K}^n est A .

Soit $g \in \mathcal{L}(\mathbf{K}^n)$ l'unique endomorphisme de \mathbf{K}^n dont la matrice dans la base canonique est P . Puisque P est inversible, g est un isomorphisme.

Et alors

$$\text{rg}(PA) = \text{rg}(g \circ f) = \text{rg}(f) = \text{rg}(A).$$

□

Théorème 7.52 : Soit $A \in \mathcal{M}_{n,p}(\mathbf{K})$. Si $B \in \mathcal{M}_{n,p}(\mathbf{K})$ est une matrice échelonnée obtenue à partir de A à l'aide d'opérations élémentaires sur les lignes, alors $\text{rg}(A) = \text{rg}(B)$.

Démonstration. Nous allons en fait prouver que chaque opération élémentaire sur les lignes de A correspond à la multiplication (à gauche) de A par une matrice inversible. Le résultat est alors une conséquence de la proposition précédente.

Rappelons que les opérations élémentaires sur les lignes sont de trois types :

- la multiplication d'une ligne par un scalaire **non nul**
- l'ajout à une ligne d'une autre ligne
- l'échange de deux lignes

Expliquons sur des exemples¹⁰ comment ces trois opérations consistent en fait en des multiplications à gauche par des matrices inversibles.

Soit donc $A \in \mathcal{M}_3(\mathbf{K})$.

- Multiplier la deuxième ligne de A par $\lambda \neq 0$ revient à considérer $\begin{pmatrix} 1 & 0 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & 1 \end{pmatrix} \times A$ (faire le calcul «à la main» pour s'en convaincre).

De même, multiplier la troisième ligne de A par λ revient à considérer $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \lambda \end{pmatrix} \times A$.

Il est clair que dans ces deux cas, on a bien multiplié A à gauche par une matrice **inversible**¹¹.

- Ajouter L_2 à L_3 revient à considérer $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix} \times A$.

De même, ajouter L_3 à L_1 revient à considérer $\begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \times A$.

Encore une fois, on a bien multiplié A par des matrices **inversibles**.

- L'échange de lignes est un petit peu plus complexe. Échanger L_1 et L_2 revient à multiplier

A par $\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$, et échanger L_1 et L_3 revient à multiplier A par $\begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}$.

A priori, nous n'avons aucune raison de penser que ces matrices sont inversibles, mais elles sont en fait égales à leur propre inverse : en effet, si on échange deux lignes, puis qu'on échange de nouveau ces deux mêmes lignes, alors on retrouve la matrice de départ... □

Corollaire 7.53 : Si $A \in \mathcal{M}_{n,p}(\mathbf{K})$, alors $\text{rg}(A)$ est égal au nombre de pivots d'une matrice échelonnée obtenue à partir de A à l'aide de la méthode du pivot.

C'est ce résultat qu'on utilise tout le temps pour calculer le rang d'une matrice.

Rang des matrices diagonales et triangulaires

Lorsque nous avons déterminé les valeurs propres des matrices triangulaires et diagonales, nous avons utilisé des résultats sur le rang de telles matrices que nous détaillons ici.

Proposition 7.54 : Le rang d'une matrice diagonale $A \in \mathcal{M}_n(\mathbf{K})$ est égal au nombre de ses coefficients diagonaux non nuls.

Démonstration. Commençons par remarquer qu'une matrice diagonale n'est pas forcément échelonnée, par exemple $\begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}$ n'est pas échelonnée car elle commence par une ligne nulle.

Toutefois, par échanges successifs de lignes, on peut descendre les lignes nulles en bas de la matrice et remonter progressivement les lignes non nulles, de manière à se ramener à une matrice échelonnée¹². Les pivots sont alors les coefficients diagonaux non nuls de A . □

Exemple 7.55

Soit $A = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$. Alors $A \xleftrightarrow{L_1 \leftrightarrow L_2} \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \xleftrightarrow{L_2 \leftrightarrow L_3} \begin{pmatrix} 0 & \boxed{1} & 0 & 0 \\ 0 & 0 & \boxed{2} & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$.

Donc $\text{rg}(A) = 2$.

¹⁰ Si vous comprenez ces trois exemples, vous devriez pouvoir comprendre directement le cas général sans qu'il soit nécessaire d'écrire tous les détails (qui sont assez fastidieux à écrire).

¹¹ Puisqu'échelonnée avec 3 pivots.

Faites-le !
Prenez vraiment une minute pour faire ce calcul et vous convaincre que ceci se généralise facilement à des matrices de taille quelconque.

Exercice
En fait, l'échange de lignes peut se réaliser uniquement à l'aide des deux premières opérations. Essayez par exemple d'échanger L_1 et L_2 uniquement en utilisant les deux premiers types d'opérations.
Indication : commencer par faire
 $L_1 \leftarrow L_1 + L_2$.

¹² Par exemple pour la matrice A ci-dessus, on obtiendrait comme matrice échelonnée la matrice

$$\begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{pmatrix}.$$

Proposition 7.56 : Soit $A \in \mathcal{M}_n(\mathbf{K})$ une matrice triangulaire. Alors son rang est supérieur ou égal au nombre de coefficients diagonaux non nuls.

Démonstration. Faisons la preuve dans le cas de matrices triangulaires supérieures, pour les matrices triangulaires inférieures, il suffit de transposer¹³.

Il est tentant de penser qu'une matrice triangulaire est échelonnée. C'est vrai si tous les coefficients diagonaux sont non nuls, mais pas forcément si un ou plusieurs coefficients diagonaux sont nuls.

Par exemple $\begin{pmatrix} 2 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ est triangulaire, mais elle n'est pas échelonnée.

Supposons que A possède r coefficients diagonaux non nuls.

Quitte à échanger des lignes, il est possible de « remonter » les r lignes correspondant à des coefficients diagonaux non nuls en haut de la matrice, et de « descendre » les $n - r$ lignes correspondant à des coefficients diagonaux nuls en bas de la matrice. Par exemple (nous avons entouré les coefficients diagonaux non nuls de la matrice de départ) :

$$\begin{pmatrix} 0 & 1 & 1 & 1 \\ 0 & \boxed{2} & 0 & 1 \\ 0 & 0 & 0 & 5 \\ 0 & 0 & 0 & \boxed{-1} \end{pmatrix} \begin{matrix} \xleftrightarrow{L_1 \leftrightarrow L_2} \\ \xleftrightarrow{L_3 \leftrightarrow L_4} \end{matrix} \begin{pmatrix} 0 & \boxed{2} & 0 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & \boxed{-1} \\ 0 & 0 & 0 & 5 \end{pmatrix} \xleftrightarrow{L_2 \leftrightarrow L_3} \begin{pmatrix} 0 & \boxed{2} & 0 & 1 \\ 0 & 0 & 0 & \boxed{-1} \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 5 \end{pmatrix}.$$

On peut ainsi obtenir une matrice dont les r premières lignes forment une matrice échelonnée, avec r pivots, et donc de rang r .

Ainsi, les r premières lignes de cette « sous-matrice » forment une famille libre (de \mathbf{K}^n), et donc l'espace vectoriel engendré par les lignes de A est de dimension au moins r . Ainsi, $\text{rg}(A) \geq r$. \square

Remarque. Il se peut qu'on ait l'égalité : par exemple $\begin{pmatrix} 1 & 2 & -1 \\ 0 & 0 & 1 \\ 0 & 0 & -2 \end{pmatrix}$ possède deux coefficients

diagonaux non nuls et est de rang 2.

Mais il se peut aussi qu'il s'agisse d'une inégalité stricte, c'est-à-dire que le rang de A soit strictement plus grand que le nombre de coefficients diagonaux non nuls. C'est par exemple

le cas de $\begin{pmatrix} -2 & 1 & 3 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$, qui est de rang 2 alors qu'elle ne possède qu'un seul coefficient diagonal non nul.

7.4.3 Quelques mots des systèmes linéaires homogènes

Résolution des systèmes homogènes

Considérons dans la suite un système de n équations à p inconnues, de la forme

$$\begin{cases} a_{1,1}x_1 + a_{1,2}x_2 + \cdots + a_{1,p}x_p = 0 \\ \vdots \\ a_{n,1}x_1 + a_{n,2}x_2 + \cdots + a_{n,p}x_p = 0 \end{cases}$$

Définition 7.57 – On dit que le système est échelonné si la matrice $A = (a_{i,j}) \in \mathcal{M}_{n,p}(\mathbf{K})$ est échelonnée.

Pour résoudre un tel système, à l'aide d'opérations sur les lignes, on commence par l'échelonner (ce qui est toujours possible).

Supposons que la matrice échelonnée ainsi obtenue possède r pivots.

Alors on choisit comme inconnues principales les inconnues correspondant aux colonnes où apparaissent les pivots, et on cherche à les exprimer en fonction des autres inconnues.

¹³ Rappelons que le rang de A est égal au rang de sa transposée.

Exemple 7.58

Sur le système échelonné suivant :

$$\begin{cases} \boxed{2x} + 3y + 4z + 2t = 0 \\ \boxed{z} + 2t = 0 \end{cases}$$

on va exprimer x et z en fonction de y et t , ces deux dernières pouvant prendre toutes les valeurs réelles possibles.

Alors l'ensemble des solutions de ce système sera un espace vectoriel de dimension 2 (car il y a 2 degrés de liberté, correspondant à y et t).

Dans le cas général, s'il y a r pivots, il y aura $p - r$ variables libres, et alors l'ensemble des solutions du système sera de dimension $p - r$.

Une autre explication du rang d'une matrice

Nous donnons ici une explication en termes de systèmes linéaires du fait que le rang d'une matrice A est égal au nombre de pivots d'une matrice échelonnée obtenue à partir de A par opérations sur les lignes.

À présent, considérons une matrice $A \in \mathcal{M}_{n,p}(\mathbf{K})$, et notons f l'application linéaire de \mathbf{K}^p dans \mathbf{K}^n dont la matrice dans les bases canoniques est A . Alors $\text{Ker } f$ est l'ensemble des $(x_1, \dots, x_p) \in \mathbf{K}^p$ tels que

$$A \begin{pmatrix} x_1 \\ \vdots \\ x_p \end{pmatrix} = 0.$$

Notons A' la matrice échelonnée du système (échelonné) obtenu par opérations élémentaires (c'est également la matrice qu'on obtiendrait en faisant sur les lignes de A les opérations faites sur les équations du système).

Alors si A' possède r pivots, par ce qui a été dit précédemment, l'ensemble des solutions du système est de dimension $p - r$.

Ainsi, $\text{Ker } f$ est de dimension $p - r$.

Par le théorème du rang, on a alors $\text{rg}(f) = \dim \mathbf{K}^p - \dim \text{Ker } f = p - (p - r) = r$.

Nous venons donc de prouver que le rang de f (et par conséquent le rang de A) est égal au nombre de pivots d'une matrice échelonnée obtenue à partir de A à l'aide d'opérations sur les lignes.

MÉTHODES DU CHAPITRE 7

✓ MÉTHODE 7.1 : Déterminer les valeurs propres d’une matrice $A \in \mathcal{M}_n(\mathbf{K})$

- Dans le cas très particulier des matrices triangulaires ou diagonales, les valeurs propres s’obtiennent sans calculs : ce sont les coefficients diagonaux.
- Pour la valeur propre 0, il suffit d’étudier l’inversibilité de A , ou encore de prouver que son rang est strictement inférieur à n .
- Pour une matrice 2×2 , on pourra utiliser le fait que λ est valeur propre si et seulement si $\det(A - \lambda I_2) = 0$, ce qui évite de recourir à une méthode du pivot.
- Dans le cas général, on utilise le fait que λ est valeur propre de A si et seulement si $\text{rg}(A - \lambda I_n) < n$. On peut alors procéder à un calcul de pivot sur la matrice $A - \lambda I_n$, avec $\lambda \in \mathbf{K}$.

Il faut être très soigneux sur les opérations élémentaires, et en particulier, ne pas multiplier une ligne par 0, quitte à distinguer plusieurs cas.

- L’énoncé d’un exercice peut parfois aider à trouver des valeurs propres sans avoir recours à un pivot, avec des indications comme «calculer $\text{rg}(A - 2I_n)$ » (et alors si ce rang est inférieur strictement à n , 2 est valeur propre de A), ou encore «calculer AV », où $V \in \mathcal{M}_{n,1}(\mathbf{K})$ est un vecteur non nul donné par l’énoncé (et alors s’il existe $\lambda \in \mathbf{K}$ tel que $AV = \lambda V$, λ est valeur propre de A).

Exercices de TD : 7.1, 7.2, 7.4

✓ MÉTHODE 7.2 : Déterminer le sous-espace propre d’une matrice $A \in \mathcal{M}_n(\mathbf{K})$ pour la valeur propre λ .

Rappelons que la dimension de $E_\lambda(A)$ est donnée par $\dim E_\lambda(A) = n - \text{rg}(A - \lambda I_n)$.

Sauf si on dispose déjà d’un vecteur propre, le plus simple est de résoudre le système $AX = \lambda X$, avec $X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathcal{M}_{n,1}(\mathbf{K})$,

puis de mettre l’ensemble des solutions sous la forme d’un Vect afin d’obtenir une base de $E_\lambda(A)$.

Exercices de TD : 7.1

✓ MÉTHODE 7.3 : Déterminer les valeurs propres et les vecteurs propres d’un endomorphisme f

Faute d’indication, le plus simple est d’écrire la matrice de f dans une base \mathcal{B} , et de déterminer les valeurs propres de cette matrice.

Si on résout un système pour obtenir les vecteurs propres de cette matrice, on n’oubliera pas que ces vecteurs propres ne sont pas des vecteurs propres de f (ils ne vivent pas dans le bon ensemble !), mais les coordonnées des vecteurs propres de f dans la base \mathcal{B} . Si on souhaite obtenir des vecteurs propres de f , il faut alors «retraduire» les vecteurs colonnes ainsi obtenus en vecteurs de E .

Exercices de TD : 7.1, 7.8, 7.10, 7.13, 7.15, 7.16

✓ MÉTHODE 7.4 : Utiliser $\sum_{\lambda \in \text{Spec}(f)} \dim E_\lambda(f) \leq \dim E$

Ce résultat permet, lorsqu’on dispose de plusieurs valeurs propres, ou d’une valeur propre avec un «gros» sous-espace propre (de dimension proche de $\dim E$), de borner le nombre d’autres valeurs propres et/ou leur dimension.

Enfin, si on a déjà des sous-espaces propres dont la dimension vaut n , alors ce résultat garantit qu’on a toutes les valeurs propres.

Exercices de TD : 7.4, 7.8, 7.9, 7.14

EXERCICES DU CHAPITRE 7

EXERCICE 7.1 Déterminer les valeurs propres et une base des sous-espaces propres des matrices suivantes :

$$A = \begin{pmatrix} -1 & 4 \\ 6 & 1 \end{pmatrix}, B = \begin{pmatrix} 3 & 1 & 2 \\ 1 & 1 & 0 \\ -1 & 1 & 2 \end{pmatrix}, C = \begin{pmatrix} -1 & 0 & 2 \\ -1 & 2 & -1 \\ -1 & 1 & 0 \end{pmatrix}, D = \begin{pmatrix} 1+2i & 2i & 1+i \\ -1-i & -i & -1-i \\ -2i & -2i & -i \end{pmatrix}, E = \begin{pmatrix} 1 & -1 & 1 \\ 2 & -5 & 4 \\ 2 & -7 & 6 \end{pmatrix}, F = \begin{pmatrix} 7 & -1 & 2 & -3 \\ 1 & 5 & 2 & -3 \\ 0 & 0 & 6 & 0 \\ 1 & -1 & 2 & 3 \end{pmatrix}.$$

Pour la matrice D , il est conseillé de commencer par l'opération $L_1 \leftarrow L_1 + L_2 + L_3$ sur la matrice $D - \lambda I_3$ puis de distinguer deux cas suivant que $\lambda = -i$ ou $\lambda \neq -i$.

EXERCICE 7.2 Vers la diagonalisation

- 1) Soit $A = \begin{pmatrix} -4 & -2 \\ 1 & -1 \end{pmatrix}$. Déterminer les valeurs propres de A et les dimensions des sous-espaces propres associés.

Montrer qu'il existe une base de $\mathcal{M}_{2,1}(\mathbf{R})$ formée de vecteurs propres de A et déterminer la matrice de f_A dans cette base.

- 2) Soit $B = \begin{pmatrix} 3 & 0 & 1 \\ 0 & 2 & -4 \\ 0 & 0 & 5 \end{pmatrix} \in \mathcal{M}_3(\mathbf{R})$. Montrer que B est semblable à une matrice diagonale que l'on précisera.

EXERCICE 7.3 Soit f l'endomorphisme de $\mathcal{M}_2(\mathbf{R})$ défini par $f(M) = {}^t M$. Déterminer les valeurs propres et les sous-espaces propres de f .

EXERCICE 7.4 Un grand classique : la matrice J

Soit $n \geq 2$ et soit $J_n \in \mathcal{M}_n(\mathbf{R})$ la matrice dont tous les coefficients valent 1. On note $V \in \mathcal{M}_{n,1}(\mathbf{R})$ le vecteur colonne dont tous les coefficients sont égaux à 1.

- Calculer le rang de J_n . En déduire que $0 \in \text{Spec}(J_n)$ et déterminer $\dim E_0(J_n)$.
- Calculer $J_n V$ et en déduire une valeur propre non nulle de J_n ainsi que la dimension du sous-espace propre associé.
- Montrer que J_n ne possède que deux valeurs propres.
- Montrer que J_n est semblable à une matrice diagonale dont on précisera les coefficients diagonaux.

EXERCICE 7.5 Soient $f, g \in \mathcal{L}(E)$ tels que $f \circ g = g \circ f$. Montrer que tout sous-espace propre de f est stable par g .

EXERCICE 7.6 Soit $f \in \mathcal{L}(E)$ et soit $x \in E$ non nul. Montrer que x est un vecteur propre de f si et seulement si $\text{Vect}(x)$ est stable par f .

EXERCICE 7.7 Soit $A = (a_{i,j})_{1 \leq i, j \leq n} \in \mathcal{M}_n(\mathbf{R})$.

- On suppose qu'il existe $\lambda \in \mathbf{R}$ tel que la somme des coefficients de chaque ligne de A soit égale à λ . Montrer que λ est valeur propre de A et que le vecteur $V \in \mathcal{M}_{n,1}(\mathbf{R})$ dont tous les coefficients sont égaux à 1 est un vecteur propre de A associé à la valeur propre λ .
- On suppose à présent que la somme des coefficients de chaque colonne de A est égale à λ . Montrer que $\lambda \in \text{Spec}(A)$.

EXERCICE 7.8 D'après EDHEC 2009

Soit $n \in \mathbf{N}^*$. Soit $f : \mathbf{R}_{2n+1}[X] \rightarrow \mathbf{R}_{2n+1}[X]$ l'application qui à un polynôme P associe $X^{2n+1}P\left(\frac{1}{X}\right)$.

- Montrer que f est un endomorphisme de $\mathbf{R}_{2n+1}[X]$.
- Calculer $f(1 + X^{2n+1})$ et en déduire que 1 est valeur propre de f . De même, calculer $f(1 - X^{2n+1})$ et en déduire une valeur propre λ de f .
- Soit $P = \sum_{k=0}^{2n+1} a_k X^k \in \mathbf{R}_{2n+1}[X]$.
 - Montrer que $P \in E_1(f)$ si et seulement si pour tout $k \in \llbracket 0, 2n+1 \rrbracket$, $a_k = a_{2n+1-k}$.
 - En déduire une base et la dimension de $E_1(f)$.
- De même déterminer une base et la dimension de $E_\lambda(f)$, où λ est la valeur propre trouvée à la question 2.
- Déterminer $\text{Spec}(f)$.

EXERCICE 7.9 Soit f un endomorphisme de \mathbf{R}^n de rang 1. Montrer que l'un au moins des endomorphismes $f - \text{id}$ et $f + \text{id}$ est bijectif.

AD

EXERCICE 7.10 Soit E un espace vectoriel de dimension finie, et soit $u \in \mathcal{L}(E)$. On suppose que u est nilpotent, c'est-à-dire qu'il existe $p \in \mathbf{N}$ tel que $u^p = 0$. Montrer que $\text{Spec}(u) = \{0\}$.

AD

EXERCICE 7.11 Soit f un endomorphisme de E et λ une valeur propre non nulle de f . Montrer que $E_\lambda(f) \subset \text{Im } f$.

PD

EXERCICE 7.12 Des matrices non semblables

AD

1) Soient $A = \begin{pmatrix} 2 & 1 & 0 & 0 \\ 0 & 2 & 1 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 2 \end{pmatrix}$ et $B = \begin{pmatrix} 2 & 1 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 2 \end{pmatrix} \in \mathcal{M}_4(\mathbf{R})$.

- a) Montrer que A et B ont même rang, même trace, mêmes valeurs propres et des sous-espaces propres de même dimension.
- b) Calculer $(A - 2I_4)^2$ et $(B - 2I_4)^2$. En déduire que A et B ne sont pas semblables.

2) Soient $A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ et $B = \begin{pmatrix} 0 & -2 \\ 2 & 0 \end{pmatrix} \in \mathcal{M}_2(\mathbf{R})$.

- a) Montrer que A et B ont même rang, même trace, et n'ont pas de valeurs propres réelles.
- b) Déterminer les valeurs propres de A et B vues comme matrices de $\mathcal{M}_2(\mathbf{C})$. En déduire que A et B ne sont pas semblables dans $\mathcal{M}_2(\mathbf{R})$, c'est-à-dire qu'il n'existe pas de matrice $P \in \text{GL}_2(\mathbf{R})$ telle que $A = P^{-1}BP$.

EXERCICE 7.13 Polynômes d'Hermite (D'après ESSEC 2002)

AD

Soit $n \geq 2$, et soit $\varphi : \mathbf{R}_n[X] \rightarrow \mathbf{R}[X]$ défini par $\varphi(P) = 2XP' - P''$.

- 1) Montrer que φ est un endomorphisme de $\mathbf{R}_n[X]$.
- 2) Écrire la matrice de φ dans la base canonique de $\mathbf{R}_n[X]$. En déduire les valeurs propres de φ et la dimension de ses sous-espaces propres.
- 3) Montrer que pour tout $p \in \llbracket 0, n \rrbracket$, il existe un unique polynôme unitaire H_p tel que $H_p'' - 2XH_p' + 2pH_p = 0$.
- 4) Montrer que pour tout $p \in \llbracket 0, n \rrbracket$, H_p est de degré p .

EXERCICE 7.14 Soit $A \in \mathcal{M}_n(\mathbf{R})$ telle que $\text{tr}(A) \neq 0$, et soit $f : \mathcal{M}_n(\mathbf{R}) \rightarrow \mathcal{M}_n(\mathbf{R})$ défini par $f(M) = \text{tr}(A)M - \text{tr}(M)A$.

AD

- 1) Montrer que f est un endomorphisme de $\mathcal{M}_n(\mathbf{R})$.
- 2) Déterminer les valeurs propres de f et la dimension des sous-espaces propres associés.

EXERCICE 7.15 On définit une application f sur $\mathbf{R}_n[X]$ ($n \geq 2$) par $f(P) = (X + 1)(X + 3)P' - nXP$.

D

- 1) Montrer que f est un endomorphisme de $\mathbf{R}_n[X]$.
- 2) Soit P un vecteur propre de f associé à la valeur propre λ .
 - a) Montrer que $\deg(P) = n$.
 - b) Montrer que les seules racines complexes de P sont -1 et -3 . En déduire que P est de la forme $a(X + 1)^k(X + 3)^\ell$, avec $a \in \mathbf{R}^*$ et $(k, \ell) \in \llbracket 0, n \rrbracket$, $k + \ell = n$.
- 3) Pour $(k, \ell) \in \llbracket 0, n \rrbracket^2$, $k + \ell = n$, on pose $P_{k,\ell} = (X + 1)^k(X + 3)^\ell$. Calculer $f(P_{k,\ell})$. En déduire les valeurs propres de f et la dimension de ses sous-espaces propres.

EXERCICE 7.16 (QSP ESCP 2007)

D

Soit φ une forme linéaire non nulle d'un espace vectoriel E de dimension finie, et soit u un vecteur non nul de E . On définit un endomorphisme de E par $f(x) = x + \varphi(x) \cdot u$. Déterminer les valeurs propres et les vecteurs propres de f .

EXERCICE 7.17 (Oral HEC 2010)

D

Soit $n \geq 2$, et soit M_n la matrice de $\mathcal{M}_n(\mathbf{R})$ définie par $M_n = \begin{pmatrix} 1 & 1 & \dots & \dots & 1 \\ 1 & 0 & \dots & \dots & 0 \\ 0 & 1 & \ddots & & 0 \\ \vdots & & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & 1 & 0 \end{pmatrix}$.

- 1) Montrer que le réel a est valeur propre de M_n si et seulement si a est racine de $P_n(X) = X^n - X^{n-1} - X^{n-2} - \dots - X - 1$.
- 2) Déterminer alors le sous-espace propre associé à la valeur propre a .

EXERCICE 7.18 (D'après oral ESCP 2012)

TD

Soient $f, g \in \mathcal{L}(\mathbf{R}^n)$. On suppose que $f \circ g = g \circ f$ et que f admet n valeurs propres distinctes.

- 1) Montrer que tout sous-espace propre de f est stable par g .
- 2) Montrer que tout vecteur propre de f est vecteur propre de g .
- 3) Montrer qu'il existe une base de E formée à la fois de vecteurs propres de f et de vecteurs propres de g . Que dire des matrices de f et g dans une telle base ?
- 4) Soit A la matrice de f dans la base canonique de \mathbf{R}^n . Déterminer le nombre de matrices $B \in \mathcal{M}_n(\mathbf{R})$ telles que $B^2 = A$ (i.e. le nombre de «racines carrées» de A).

CORRECTION DES EXERCICES DU CHAPITRE 7

SOLUTION DE L'EXERCICE 7.1

1. Un réel λ est valeur propre de A si et seulement si $A - \lambda I_2$ n'est pas inversible. Soit si et seulement si $\det(A - \lambda I_2) = 0$. Or

$$\det(A - \lambda I_2) = \det \begin{pmatrix} -1 - \lambda & 4 \\ 6 & 1 - \lambda \end{pmatrix} = (-1 - \lambda)(1 - \lambda) - 24 = \lambda^2 - 25.$$

Donc $\lambda \in \text{Spec}(A) \Leftrightarrow \lambda^2 - 25 = 0 \Leftrightarrow \lambda \in \{-5, 5\}$.

Ainsi, A possède deux valeurs propres, 5 et -5 .

$$\text{On a } \begin{pmatrix} x \\ y \end{pmatrix} \in E_{-5}(A) \Leftrightarrow \begin{pmatrix} -1 & 4 \\ 6 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = -5 \begin{pmatrix} x \\ y \end{pmatrix} \Leftrightarrow \begin{cases} -x + 4y = -5x \\ 6x + y = -5y \end{cases} \Leftrightarrow x = -y.$$

$$\text{Donc } E_{-5}(A) = \left\{ \begin{pmatrix} x \\ -x \end{pmatrix}, x \in \mathbf{R} \right\} = \text{Vect} \begin{pmatrix} 1 \\ -1 \end{pmatrix}.$$

De même, on a $\begin{pmatrix} x \\ y \end{pmatrix} \in E_5(A)$ si et seulement si

$$\begin{pmatrix} -1 & 4 \\ 6 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = 5 \begin{pmatrix} x \\ y \end{pmatrix} \Leftrightarrow \begin{cases} -x + 4y = 5x \\ 6x + y = 5y \end{cases} \Leftrightarrow y = \frac{3}{2}x.$$

$$\text{Et donc } E_5(A) = \text{Vect} \begin{pmatrix} 1 \\ \frac{3}{2} \end{pmatrix} = \text{Vect} \begin{pmatrix} 2 \\ 3 \end{pmatrix}.$$

2. $\lambda \in \mathbf{R}$ est valeur propre de B si et seulement si $\text{rg}(B - \lambda I_3) < 3$. Or

$$\begin{aligned} B - \lambda I_3 &= \begin{pmatrix} 3 - \lambda & 1 & 2 \\ 1 & 1 - \lambda & 0 \\ -1 & 1 & 2 - \lambda \end{pmatrix} \xrightarrow{L_1 \leftrightarrow L_3} \begin{pmatrix} -1 & 1 & 2 - \lambda \\ 1 & 1 - \lambda & 0 \\ 3 - \lambda & 1 & 2 \end{pmatrix} \\ &\xrightarrow{L_2 \leftarrow L_2 + L_1, L_3 \leftarrow L_3 + (3 - \lambda)L_1} \begin{pmatrix} -1 & 1 & 2 - \lambda \\ 0 & 2 - \lambda & 2 - \lambda \\ 0 & 4 - \lambda & \lambda^2 - 5\lambda + 8 \end{pmatrix} \xrightarrow{L_2 \leftarrow L_2 - L_3} \begin{pmatrix} -1 & 1 & 2 - \lambda \\ 0 & -2 & -\lambda^2 + 4\lambda - 6 \\ 0 & 4 - \lambda & \lambda^2 - 5\lambda + 8 \end{pmatrix} \xrightarrow{L_3 \leftarrow 2L_3 + (4 - \lambda)L_2} \begin{pmatrix} -1 & 1 & 2 - \lambda \\ 0 & -2 & -\lambda^2 + 4\lambda - 6 \\ 0 & 0 & \lambda^3 - 6\lambda^2 + 12\lambda - 8 \end{pmatrix} \end{aligned}$$

Ainsi, λ est valeur propre de B si et seulement si $\lambda^3 - 6\lambda^2 + 12\lambda - 8 = 0 \Leftrightarrow (\lambda - 2)^3 = 0 \Leftrightarrow \lambda = 2$.

Et donc 2 est la seule valeur propre de B : $\text{Spec}(B) = \{2\}$.

Autre méthode : si on n'a pas pensé après la deuxième étape à faire l'opération $L_2 \leftarrow L_2 - L_3$ pour faire disparaître les λ , on peut diviser la ligne 2 par $2 - \lambda$. Mais il faut¹ alors distinguer deux cas : $2 - \lambda \neq 0$ et $2 - \lambda = 0$.

Si $2 - \lambda = 0$: alors

$$B - \lambda I_3 \xrightarrow{L_2 \leftarrow L_2 / (2 - \lambda)} \begin{pmatrix} -1 & 1 & 2 - \lambda \\ 0 & 1 & 1 \\ 0 & 4 - \lambda & \lambda^2 - 5\lambda + 8 \end{pmatrix} \xrightarrow{L_3 \leftarrow L_3 - (4 - \lambda)L_2} \begin{pmatrix} -1 & 1 & 2 - \lambda \\ 0 & 1 & 1 \\ 0 & 0 & \lambda^2 - 4\lambda + 4 \end{pmatrix}.$$

Et cette matrice est de rang inférieur ou égal à 2 si et seulement si $\lambda^2 - 4\lambda + 4 = 0 \Leftrightarrow (\lambda - 2)^2 = 0 \Leftrightarrow \lambda = 2$.

Puisque nous avons supposé $\lambda \neq 2$, ceci ne nous donne pas de valeur propre.

Si $\lambda = 2$: alors

$$B - 2I_3 \xrightarrow{L_2 \leftarrow L_2 / 0} \begin{pmatrix} -1 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 2 & 2 \end{pmatrix}.$$

Cette matrice est de rang au plus 2 car elle possède une ligne nulle, et donc 2 valeur propre de B .

$$\text{On a } \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in E_2(B) \text{ si et seulement si } B \begin{pmatrix} x \\ y \\ z \end{pmatrix} = 2 \begin{pmatrix} x \\ y \\ z \end{pmatrix} \text{ soit}$$

$$\begin{cases} 3x + y + 2z = 2x \\ x + y = 2y \\ -x + y + 2z = 2z \end{cases} \Leftrightarrow \begin{cases} x + y + 2z = 0 \\ x = y \end{cases} \Leftrightarrow \begin{cases} z = -x \\ y = x \end{cases}$$

Rédaction

Il est important de prendre le temps d'expliquer ce que l'on fait : λ est valeur propre si et seulement si $A - \lambda I_2$ n'est pas inversible etc.

En particulier, en cas d'erreur de calcul, un raisonnement juste peut permettre d'espérer quelques points. Un calcul faux sans explications n'en rapportera aucun.

Rédaction

Ici aussi, il est important d'expliquer la démarche avant de se lancer dans un calcul de pivot.

¹ Pour éviter de diviser par 0.

Méthode

Nous ne pouvons pas conclure que 2 est valeur propre car nous avons supposé $\lambda \neq 2$, mais il faut alors encore traiter le cas $\lambda = 2$!

$$\text{Ainsi, } E_2(B) = \text{Vect} \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix}.$$

$$3. \text{ Spec}(C) = \{1\}, E_1(C) = \text{Vect} \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}.$$

4. $\lambda \in \mathbb{C}$ est valeur propre de D si et seulement si $\text{rg}(D - \lambda I_3) < 3$. Or,

$$D - \lambda I_3 = \begin{pmatrix} 1+2i-\lambda & 2i & 1+i \\ -1-i & -i-\lambda & -1-i \\ -2i & -2i & -i-\lambda \end{pmatrix} \xleftrightarrow{L_1 \leftarrow L_1 + L_2 + L_3} \begin{pmatrix} -i-\lambda & -i-\lambda & -i-\lambda \\ -1-i & -i-\lambda & -1-i \\ -2i & -2i & -i-\lambda \end{pmatrix}.$$

Si $\lambda \neq -i \Leftrightarrow -i-\lambda \neq 0$, alors il est possible de diviser la première ligne par $-i-\lambda$ et donc

$$\begin{pmatrix} 1 & 1 & 1 \\ -1-i & -i-\lambda & -1-i \\ -2i & -2i & -i-\lambda \end{pmatrix} \xleftrightarrow{L_2 \leftarrow L_2 + (1+i)L_1, L_3 \leftarrow L_3 + 2iL_1} \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1-\lambda & 0 \\ 0 & 0 & \lambda_i \end{pmatrix}$$

Cette matrice n'est pas inversible si et seulement si $\lambda = 1$ ou $\lambda = i$, qui sont donc deux valeurs propres de D .

Si $\lambda = -i$, alors $D - \lambda I_3 = D + iI_3 \xleftrightarrow{} \begin{pmatrix} 0 & 0 & 0 \\ -1-i & 0 & -1-i \\ -2i & -2i & 0 \end{pmatrix}$ qui n'est pas de rang 3 car sa première ligne est nulle. Donc $-i$ est valeur propre de D .

On a $\begin{pmatrix} x \\ y \\ z \end{pmatrix} \in E_i(D)$ si et seulement si $D \begin{pmatrix} x \\ y \\ z \end{pmatrix} = i \begin{pmatrix} x \\ y \\ z \end{pmatrix}$ soit

$$\begin{cases} (1+2i)x + 2iy + (1+i)z = ix \\ -(1+i)x - iy - (1+i)z = iy \\ -2ix - 2iy - iz = iz \end{cases} \Leftrightarrow \begin{cases} (1+i)x + 2iy + (1+i)z = 0 \\ i(x+y+z) = 0 \end{cases} \Leftrightarrow \begin{cases} z = -x - y \\ (1+i)x + 2iy - (1+i)x - (1+i)y = 0 \end{cases} \Leftrightarrow \begin{cases} z = -x \\ y = 0 \end{cases}$$

Donc $E_i(D) = \text{Vect} \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}$. De même, on a $\begin{pmatrix} x \\ y \\ z \end{pmatrix} \in E_{-i}(D)$ si et seulement si $D \begin{pmatrix} x \\ y \\ z \end{pmatrix} = -i \begin{pmatrix} x \\ y \\ z \end{pmatrix}$

soit

$$\begin{cases} (1+2i)x + 2iy + (1+i)z = -ix \\ -(1+i)x - iy - (1+i)z = -iy \\ -2ix - 2iy - iz = -iz \end{cases} \Leftrightarrow \begin{cases} (1+3i)x + 2iy + (1+i)z = 0 \\ x = -z \\ x = -y \end{cases} \Leftrightarrow \begin{cases} x = -z \\ x = -y \end{cases}$$

Et donc $E_{-i}(D) = \text{Vect} \begin{pmatrix} 1 \\ -1 \\ -1 \end{pmatrix}$.

Enfin, $\begin{pmatrix} x \\ y \\ z \end{pmatrix} \in E_1(D)$ si et seulement si $D \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$, soit

$$\begin{cases} (1+2i)x + 2iy + (1+i)z = x \\ -(1+i)x - iy - (1+i)z = y \\ -2ix - 2iy - iz = z \end{cases} \Leftrightarrow \begin{cases} 2ix + 2iy + (1+i)z = 0 \\ x + y + z = 0 \\ -2ix - 2iy - (1+i)z = 0 \end{cases} \Leftrightarrow \begin{cases} z = 0 \\ x = -y \end{cases}$$

Et donc $E_1(D) = \text{Vect} \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}$.

$$5. \text{ Spec}(E) = \{-1, 1, 2\}. E_1(E) = \text{Vect} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, E_{-1}(E) = \text{Vect} \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}, E_2(E) = \text{Vect} \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}.$$

$$6. \text{ Spec}(F) = \{3, 6\}. E_3(F) = \text{Vect} \begin{pmatrix} 1 \\ 1 \\ 0 \\ 1 \end{pmatrix}, E_6(F) = \text{Vect} \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 \\ 0 \\ -1 \\ 0 \end{pmatrix}, \begin{pmatrix} 3 \\ 0 \\ 0 \\ 1 \end{pmatrix}.$$

SOLUTION DE L'EXERCICE 7.2

1. λ est valeur propre de A si et seulement si $A - \lambda I_2$ n'est pas inversible, soit si et seulement si $\det(A - \lambda I_2) = 0$, c'est-à-dire si et seulement si

$$(-4 - \lambda)(-1 - \lambda) + 2 = 0 \Leftrightarrow \lambda^2 + 5\lambda + 6 = 0 \Leftrightarrow \lambda \in \{-2, -3\}.$$

Donc $\text{Spec}(A) = \{-2, -3\}$.

Puisqu'on a deux valeurs propres distinctes, alors nécessairement les sous-espaces propres sont tous deux de dimension 1.

On a alors $E_{-2}(A) \oplus E_{-3}(A) \subset \mathcal{M}_{2,1}(\mathbf{R})$, et de dimension 2. Autrement dit,

$$\mathcal{M}_{2,1}(\mathbf{R}) = E_{-2}(A) \oplus E_{-3}(A)$$

et donc la concaténation de bases des sous-espaces propres forme une base (e_1, e_2) de $\mathcal{M}_{2,1}(\mathbf{R})$, qui est nécessairement formée de vecteurs propres de A .

On a alors $f_A(e_1) = Ae_1 = -2e_1$ et $f_A(e_2) = Ae_2 = -3e_2$. Donc la matrice de f_A dans la base

$$(e_1, e_2) \text{ est } D = \begin{pmatrix} f(e_1) & f(e_2) \\ -2 & 0 \\ 0 & -3 \end{pmatrix} \begin{matrix} e_1 \\ e_2 \end{matrix}.$$

Puisque la matrice de f_A dans la base canonique de $\mathcal{M}_{2,1}(\mathbf{R})$ est A , on en déduit que A est semblable à la matrice D .

2. B est triangulaire, donc ses valeurs propres sont ses coefficients diagonaux. On en déduit que B possède trois valeurs propres qui sont 3, 2 et -5 .

Puisqu'on a trois valeurs propres distinctes, alors les trois sous-espaces propres sont de dimension 1.

Alors, la concaténation de bases des sous-espaces propres de B est une base de $\mathcal{M}_{3,1}(\mathbf{R})$ formée de vecteurs propres de B .

Et alors, comme précédemment, la matrice de f_B dans cette base est diagonale, et possède les valeurs propres de B sur sa diagonale. Donc B est semblable à une matrice diagonale dont les coefficients diagonaux sont les valeurs propres de B .

SOLUTION DE L'EXERCICE 7.3

Écrivons la matrice de f dans la base canonique $\mathcal{B} = (E_{1,1}, E_{1,2}, E_{2,1}, E_{2,2})$ de $\mathcal{M}_2(\mathbf{R})$.

On a

$$E_{1,1} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, E_{1,2} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, E_{2,1} = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, E_{2,2} = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}.$$

Et donc $f(E_{1,1}) = E_{1,1}$, $f(E_{1,2}) = E_{2,1}$, $f(E_{2,1}) = E_{1,2}$ et $f(E_{2,2}) = E_{2,2}$.

Ainsi, la matrice de f dans la base \mathcal{B} est

$$A = \begin{pmatrix} f(E_{1,1}) & f(E_{1,2}) & f(E_{2,1}) & f(E_{2,2}) \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{matrix} E_{1,1} \\ E_{1,2} \\ E_{2,1} \\ E_{2,2} \end{matrix}.$$

Pour $\lambda \in \mathbf{R}$, on a donc $\lambda \in \text{Spec}(A)$ si et seulement si $\text{rg}(A - \lambda I_4) < 4$. Or

$$A - \lambda I_4 = \begin{pmatrix} 1 - \lambda & 0 & 0 & 0 \\ 0 & -\lambda & 1 & 0 \\ 0 & 1 & -\lambda & 0 \\ 0 & 0 & 0 & 1 - \lambda \end{pmatrix} \xleftrightarrow{L_3 \leftrightarrow L_2} \begin{pmatrix} 1 - \lambda & 0 & 0 & 0 \\ 0 & 1 & -\lambda & 0 \\ 0 & -\lambda & 1 & 0 \\ 0 & 0 & 0 & 1 - \lambda \end{pmatrix} \xleftrightarrow{L_3 \leftarrow L_3 + \lambda L_2} \begin{pmatrix} 1 - \lambda & 0 & 0 & 0 \\ 0 & 1 & -\lambda & 0 \\ 0 & 0 & 1 - \lambda^2 & 0 \\ 0 & 0 & 0 & 1 - \lambda \end{pmatrix}.$$

Et donc, $A - \lambda I_4$ est de rang inférieur ou égal à 4 si et seulement si $1 - \lambda = 0$ ou $1 - \lambda^2 = 0$, soit $\lambda = 1$ ou $\lambda = -1$.

On en déduit que $\text{Spec}(A) = \{-1, 1\}$, et donc les valeurs propres de f sont 1 et -1 .

On a alors $M \in E_1(f) \Leftrightarrow {}^t M = M \Leftrightarrow M \in \mathcal{S}_2(\mathbf{R})$. Et donc $E_1(f)$ est l'ensemble des matrices symétriques de $\mathcal{M}_2(\mathbf{R})$.

De même, $M \in E_{-1}(f)$ si et seulement si $M = -{}^t M$ et donc $E_{-1}(f)$ est l'ensemble des matrices antisymétriques de $\mathcal{M}_2(\mathbf{R})$.

Astuce

Pour une matrice 2×2 , le déterminant peut permettre d'obtenir les valeurs propres sans utiliser de pivot.

Précision

On a alors $e_1 \in E_{-2}(A)$ et $e_2 \in E_{-3}(A)$.

Remarque

Il n'y a pas unicité de la matrice diagonale : tout dépend dans quel ordre on a choisi de concaténer les bases des sous-espaces propres pour obtenir une base de $\mathcal{M}_{3,1}(\mathbf{R})$ formée de vecteurs propres de B .

Méthode

Pour déterminer les valeurs propres d'un endomorphisme, le plus simple est (lorsque c'est possible) d'écrire la matrice de cet endomorphisme dans une base (n'importe laquelle), puis de déterminer les valeurs propres de cette matrice.

Autre méthode : si on n'a pas remarqué que nous connaissons déjà les sous-espaces propres de f , il est possible de déterminer les sous-espaces propres de A .

On trouve alors

$$E_1(A) = \text{Vect} \left(\begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} \right) \text{ et } E_{-1}(A) = \left(\begin{pmatrix} 0 \\ 1 \\ -1 \\ 0 \end{pmatrix} \right).$$

Et donc, on a

$$E_1(f) = \text{Vect}(E_{1,1}, E_{1,2} + E_{2,1}, E_{2,2}) \text{ et } E_{-1}(f) = \text{Vect}(E_{1,2} - E_{2,1}).$$

SOLUTION DE L'EXERCICE 7.4

1. Toutes les colonnes de J_n sont égales, et J_n est non nulle, donc $\text{rg}(J_n) = 1$. Ainsi, J_n n'est pas inversible, et donc 0 est valeur propre de J_n , et alors

$$\dim E_0(J_n) = n - \text{rg}(J_n) = n - 1.$$

2. On a

$$J_n V = \begin{pmatrix} 1 & \dots & 1 \\ \vdots & & \vdots \\ 1 & \dots & 1 \end{pmatrix} \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} = \begin{pmatrix} n \\ \vdots \\ n \end{pmatrix} = nV.$$

Ainsi, n est valeur propre de J_n , et V est un vecteur propre de J_n pour la valeur propre n . De plus, on a $\dim E_n(J_n) \geq 1$, et on doit avoir

$$\dim E_0(J_n) + \dim E_n(J_n) \leq n \Leftrightarrow \dim E_n(J_n) \leq n - \dim E_0(J_n) \leq 1.$$

Ainsi, $\dim E_n(J_n) = 1$.

3. On a $\dim E_0(J_n) + \dim E_n(J_n) = n$, donc il ne peut y avoir d'autres valeurs propres de J_n car on a nécessairement

$$\sum_{\lambda \in \text{Spec}(J_n)} \dim E_\lambda(J_n) \leq n.$$

4. Notons f l'endomorphisme de $\mathcal{M}_{n,1}(\mathbf{R})$ canoniquement associé à J_n . On a $\mathcal{M}_{n,1}(\mathbf{R}) = E_0(J_n) \oplus E_n(J_n)$, donc une base de $\mathcal{M}_{n,1}(\mathbf{R})$ est obtenue par concaténation d'une base de $E_0(J_n)$ et d'une base² de $E_n(J_n)$. Notons donc (e_1, \dots, e_{n-1}) une base de $E_0(J_n)$ et e_n une base de $E_n(J_n)$. Alors on a

$$f(e_1) = J_n e_1 = 0 \cdot e_1 = 0, \dots, f(e_{n-1}) = 0$$

et $f(e_n) = J_n e_n = n e_n$.

Donc la matrice de f dans la base (e_1, \dots, e_n) est

$$\begin{pmatrix} f(e_1) & f(e_2) & \dots & f(e_{n-1}) & f(e_n) \\ 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & & \vdots & \vdots \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & \dots & \dots & 0 & 0 \\ 0 & \dots & \dots & 0 & n \end{pmatrix} \begin{matrix} e_1 \\ e_2 \\ \vdots \\ e_{n-1} \\ e_n \end{matrix} = \text{Diag}(\underbrace{0, \dots, 0}_{n-1 \text{ fois}}, n).$$

Et donc J_n est semblable à cette matrice diagonale.

SOLUTION DE L'EXERCICE 7.5

Soit λ une valeur propre de f , et soit $x \in E_\lambda(f)$. Alors $f(x) = \lambda x$.

Il vient alors

$$f(g(x)) = g(f(x)) = g(\lambda x) = \lambda g(x).$$

Ainsi $g(x) \in E_\lambda(f)$. Donc $E_\lambda(f)$ est stable par g .

⚠ Danger !

Ne pas oublier que les vecteurs propres de A sont les coordonnées dans la base \mathcal{B} des vecteurs propres de f . Il faut donc «traduire» ces vecteurs de $\mathcal{M}_{4,1}(\mathbf{R})$ en vecteurs propres de f , c'est-à-dire en matrices de $\mathcal{M}_2(\mathbf{R})$.

Rappel

0 est valeur propre de A si et seulement si A n'est pas inversible.

Rappel

L'endomorphisme f est défini par

$$f(X) = J_n X.$$

Sa matrice dans la base canonique de $\mathcal{M}_{n,1}(\mathbf{R})$ est J_n .

² Notons que grâce à la question 2, on sait qu'une telle base est par exemple formée du vecteur V .

Alternative

Une autre solution est de remarquer que g et $f - \lambda \text{id}_E$ commutent. Et donc, d'après le résultat prouvé dans l'exercice 3.16,

$$E_\lambda(f) = \text{Ker}(f - \lambda \text{id}_E)$$

est stable par g .

SOLUTION DE L'EXERCICE 7.6

Si x est un vecteur propre de f , associé à la valeur propre λ , alors pour tout $u \in \text{Vect}(x)$, il existe $\alpha \in \mathbf{K}$ tel que $u = \alpha x$.

Et alors $f(u) = f(\alpha x) = \alpha f(x) = \alpha \lambda x \in \text{Vect}(x)$.

Ceci prouve donc que $\text{Vect}(x)$ est stable par f .

Inversement, supposons que $\text{Vect}(x)$ soit stable par f .

Alors en particulier, $f(x) \in \text{Vect}(x)$, et donc il existe $\lambda \in \mathbf{K}$ tel que $f(x) = \lambda x$.

Et puisque $x \neq 0$, c'est donc que x est un vecteur propre de f associé à la valeur propre λ .

SOLUTION DE L'EXERCICE 7.7

1. On a

$$AV = \begin{pmatrix} a_{1,1} & \dots & a_{1,n} \\ \vdots & & \vdots \\ a_{n,1} & \dots & a_{n,n} \end{pmatrix} \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} = \begin{pmatrix} a_{1,1} + \dots + a_{1,n} \\ \vdots \\ a_{n,1} + \dots + a_{n,n} \end{pmatrix} = \begin{pmatrix} \lambda \\ \vdots \\ \lambda \end{pmatrix} = \lambda \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} = \lambda V.$$

Et puisque $V \neq 0$, ceci prouve à la fois que $\lambda \in \text{Spec}(A)$ et que V est un vecteur propre de A associé à la valeur propre λ .

2. Comme la transposition échange lignes et colonnes, la somme de chaque ligne de tA est égale à λ .

Et donc d'après la première question, λ est valeur propre de tA .

Mais A et tA ont les mêmes valeurs propres, et donc $\lambda \in \text{Spec}(A)$.

En revanche, A et tA n'ont pas les mêmes vecteurs propres³, et on ne peut cette fois pas affirmer que V est un vecteur propre de A , c'est uniquement un vecteur propre de tA .

Méthode

Puisque le vecteur V nous est donné, vérifier que V est vecteur propre, c'est juste prouver que $AV = \lambda V$.

³ Bien que les sous-espaces propres aient même dimension.

SOLUTION DE L'EXERCICE 7.8

1. Soit $P = \sum_{k=0}^{2n+1} a_k X^k$. Alors

$$f(P) = X^{2n+1} \sum_{k=0}^{2n+1} a_k \frac{1}{X^k} = \sum_{k=0}^{2n+1} a_k X^{2n+1-k} \in \mathbf{R}_{2n+1}[X].$$

Donc f est bien à valeurs dans $\mathbf{R}_{2n+1}[X]$.

Soient $P, Q \in \mathbf{R}_{2n+1}[X]$, et soit $\lambda \in \mathbf{R}$. Alors

$$f(\lambda P + Q) = X^{2n+1} \left(\lambda P \left(\frac{1}{X} \right) + Q \left(\frac{1}{X} \right) \right) = \lambda X^{2n+1} P \left(\frac{1}{X} \right) + X^{2n+1} Q \left(\frac{1}{X} \right) = \lambda f(P) + f(Q).$$

Donc f est linéaire : c'est bien un endomorphisme de $\mathbf{R}_{2n+1}[X]$.

2. On a

$$f(1 + X^{2n+1}) = X^{2n+1} \left(1 + \frac{1}{X^{2n+1}} \right) = X^{2n+1} + 1 = 1 + X^{2n+1} = 1 \cdot (1 + X^{2n+1}).$$

On en déduit que 1 est valeur propre de f et que $1 + X^{2n+1}$ est un vecteur propre de f pour la valeur propre 1.

De même, on a $f(1 - X^{2n+1}) = X^{2n+1} - 1 = -1 \cdot (1 - X^{2n+1})$.

On en déduit que -1 est valeur propre de f .

3.a. Nous avons déjà calculé $f(P)$ à la question 1, et on a donc $P \in E_1(f)$ si et seulement si $f(P) = P$, c'est-à-dire si et seulement si

$$\sum_{k=0}^{2n+1} a_k X^{2n+1-k} = \sum_{k=0}^{2n+1} a_k X^k.$$

Dans la première somme, procédons au changement d'indice $i = 2n + 1 - k$, de sorte que l'équation précédente devient

$$\sum_{i=0}^{2n+1} a_{2n+1-i} X^i = \sum_{k=0}^{2n+1} a_k X^k.$$

Et donc, par identification des coefficients, on a $f(P) = P$ si et seulement si

$$\forall k \in \llbracket 0, 2n + 1 \rrbracket, a_{2n+1-k} = a_k.$$

Détails

Par définition, une valeur propre λ est un scalaire tel qu'il existe $x \neq 0$ tel que $f(x) = \lambda x$.

Ici, en posant $P = 1 + X^{2n+1}$, on a $f(P) = P$, donc on a à la fois une valeur propre (1) et un vecteur propre associé : P .

3.b. On a donc un système de $2n + 2$ équations, d'inconnues $a_0, a_1, \dots, a_{2n+1}$. Or ce système est

$$\begin{cases} a_0 & = a_{2n+1} \\ a_1 & = a_{2n} \\ & \vdots \\ a_{2n} & = a_1 \\ a_{2n+1} & = a_0 \end{cases}$$

On remarque qu'alors, chaque équation apparaît deux fois, et donc (a_0, \dots, a_{2n+1}) est solution du système si et seulement si

$$\begin{cases} a_0 & = a_{2n+1} \\ a_1 & = a_{2n} \\ & \vdots \\ a_n & = a_{n+1} \end{cases}$$

Autrement dit, $f(P) = P$ si et seulement si

$$P = a_0(1 + X^{2n+1}) + a_1(X + X^{2n}) + \dots + a_n(X^n + X^{n+1}) = \sum_{k=0}^n a_k(X^k + X^{2n+1-k}).$$

Et donc si et seulement si $P \in \text{Vect}(1 + X^{2n+1}, X + X^{2n}, \dots, X^n + X^{n+1})$.

Ainsi, une famille génératrice de $E_1(f)$ est $\{X^k + X^{2n+1-k}, k \in \llbracket 0, n \rrbracket\}$. Cette famille est libre car formée de polynômes de degrés deux à deux distincts : c'est une base de $E_1(f)$, qui est donc de dimension $n + 1$.

4. De même, on a $f(P) = -P$ si et seulement si

$$\sum_{k=0}^{2n+1} a_k X^{2n+1-k} = - \sum_{k=0}^{2n+1} a_k X^k \Leftrightarrow \sum_{i=0}^{2n+1} a_{2n+1-i} X^i = - \sum_{k=0}^{2n+1} a_k X^k.$$

Et donc⁴ si et seulement si pour tout $k \in \llbracket 0, 2n + 1 \rrbracket$, $a_k = -a_{2n+1-k}$.

Là encore, c'est le cas si et seulement si

$$\begin{cases} a_0 & = -a_{2n+1} \\ a_1 & = -a_{2n} \\ & \vdots \\ a_n & = -a_{n+1} \end{cases}$$

Et donc une base de $E_{-1}(f)$ est $\{X^k - X^{2n+1-k}, k \in \llbracket 0, n \rrbracket\}$ de sorte que $E_{-1}(f)$ est de dimension $n + 1$.

5. Nous avons déjà deux valeurs propres, 1 et -1 , avec $\dim E_1(f) + \dim E_{-1}(f) = 2n + 2$.

Puisque d'autre part on a

$$\sum_{\lambda \in \text{Spec}(f)} \dim E_\lambda(f) \leq \dim \mathbf{R}_{2n+1}[X] = 2n + 2,$$

f ne peut posséder d'autre valeur propre. Et donc $\text{Spec}(f) = \{-1, 1\}$.

SOLUTION DE L'EXERCICE 7.9

Si f est de rang 1, alors $\dim \text{Ker } f = n - 1$, et donc $\dim E_0(f) = n - 1$.

Puisqu'on doit avoir $\sum_{\lambda \in \text{Spec}(f)} \dim E_\lambda(f) \leq n$, f possède au plus une valeur propre non nulle.

Et donc l'un au moins des deux nombres 1 et -1 n'est pas valeur propre de f , de sorte que l'un au moins des endomorphismes $f - \text{id}$ et $f + \text{id}$ est bijectif.

SOLUTION DE L'EXERCICE 7.10

Puisque $u^p = 0$, u n'est pas inversible, et donc 0 est bien valeur propre de u .

Soit λ une valeur propre de u , et soit x un vecteur propre associé. Alors

$$u^2(x) = u(u(x)) = u(\lambda x) = \lambda u(x) = \lambda^2 x.$$

⁴ Toujours pas identification des coefficients.

Rappel
 $\lambda \in \text{Spec}(f)$ si et seulement si $f - \lambda \text{id}$ n'est pas bijectif.

Puis, de proche en proche, on prouve que $u^3(x) = \lambda^3 x, \dots, u^p(x) = \lambda^p x$.
Or, $u^p(x) = 0$, et x étant non nul⁵, nécessairement $\lambda^p = 0$. Et donc $\lambda = 0$.
Ainsi, 0 est la seule valeur propre de u : $\text{Spec}(u) = \{0\}$.

SOLUTION DE L'EXERCICE 7.11

Soit $x \in E_\lambda(f)$. Alors $f(x) = \lambda x$, et λ étant non nul, on a $x = \frac{1}{\lambda} f(x) = f\left(\frac{1}{\lambda} x\right) \in \text{Im } f$.

Donc $E_\lambda(f) \subset \text{Im } f$.

SOLUTION DE L'EXERCICE 7.12

- 1.a. Puisque A et B sont triangulaires et que leurs coefficients diagonaux sont tous non nuls, on a évidemment $\text{rg}(A) = \text{rg}(B) = 4$.
De plus, $\text{tr}(A) = \text{tr}(B) = 8$.
Les valeurs propres de A et de B sont égales à leurs coefficients diagonaux, donc $\text{Spec}(A) = \text{Spec}(B) = \{2\}$.
Enfin, on a

$$\dim E_2(A) = 4 - \text{rg}(A - 2I_4) = 4 - 2 = 2 \text{ et } \dim E_2(B) = 4 - \text{rg}(B - 2I_4) = 4 - 2 = 2.$$

- 1.b. On a

$$A - 2I_4 = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \text{ et donc } (A - 2I_4)^2 = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

D'autre part,

$$B - 2I_4 = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix} \text{ et donc } (B - 2I_4)^2 = 0.$$

Puisque seule la matrice nulle est semblable à elle-même, $(A - 2I_4)^2$ et $(B - 2I_4)^2$ ne sont pas semblables.

Posons $P = (X - 2)^2 \in \mathbf{R}_2[X]$. Alors si A et B étaient semblables, $P(A) = (A - 2I_4)^2$ et $P(B) = (B - 2I_4)^2$ devraient être semblables.

On en déduit que A et B ne sont pas semblables.

- 2.a. A et B sont toutes deux de rang 2, et de trace nulle.
Pour $\lambda \in \mathbf{R}$, on a $\lambda \in \text{Spec}(A)$ si et seulement si $\text{rg}(A - \lambda I_2) < 2$. Or

$$A - \lambda I_2 = \begin{pmatrix} -\lambda & -1 \\ 1 & -\lambda \end{pmatrix} \xrightarrow{L_1 \leftrightarrow L_2} \begin{pmatrix} 1 & -\lambda \\ -\lambda & -1 \end{pmatrix} \xrightarrow{L_2 \leftarrow L_2 + \lambda L_1} \begin{pmatrix} 1 & -\lambda \\ 0 & -1 - \lambda^2 \end{pmatrix}$$

Donc $\text{rg}(A - \lambda I_2) < 2$ si et seulement si $1 + \lambda^2 = 0$, ce qui n'est pas possible pour $\lambda \in \mathbf{R}$.

Donc A ne possède pas de valeurs propres.

De même, on prouverait que $\text{rg}(B - \lambda I_2) < 2$ si et seulement si $4 + \lambda^2 = 0$, ce qui n'est pas non plus possible, donc B ne possède pas de valeurs propres réelles.

- 2.b. Si l'on voit A et B comme des matrices complexes⁶, alors les calculs précédemment réalisés restent valables : $\lambda \in \mathbf{C}$ est valeur propre de A si et seulement si $1 + \lambda^2 = 0$, c'est-à-dire si et seulement si $\lambda = \pm i$. Donc $\text{Spec}(A) = \{i, -i\}$.

De même, $\lambda \in \mathbf{C}$ est valeur propre de B si et seulement si $\lambda^4 + 4 = 0$, c'est-à-dire si et seulement si $\lambda = \pm 2i$. Donc $\text{Spec}(B) = \{2i, -2i\}$.

Si A et B étaient semblables dans $\mathcal{M}_2(\mathbf{R})$, alors il existerait une matrice $P \in GL_2(\mathbf{R})$ telle que $A = P^{-1}BP$. Mais P est alors également une matrice de $GL_2(\mathbf{C})$, et donc A et B seraient semblables dans $\mathcal{M}_2(\mathbf{C})$. Et par conséquent, A et B devraient avoir les mêmes valeurs propres complexes.

Ce n'est pas le cas, donc on en déduit que A et B ne sont pas semblables dans $\mathcal{M}_2(\mathbf{R})$.

SOLUTION DE L'EXERCICE 7.13

1. Commençons par montrer que φ est à valeurs dans $\mathbf{R}_n[X]$. Il est clair que φ est à valeurs dans $\mathbf{R}[X]$.

Si $P \in \mathbf{R}_n[X]$, alors $\deg P' \leq n - 1$ et $\deg 2XP' \leq n$. Puisque $\deg P'' \leq n - 2$, on a donc $\deg \varphi(P) \leq n$.

Si $P, Q \in \mathbf{R}_n[X]$ et si $\lambda \in \mathbf{R}$, alors

$$\varphi(\lambda P + Q) = 2X(\lambda P + Q)' - (\lambda P + Q)'' = 2X(\lambda P' + Q') - \lambda P'' - Q'' = \lambda(2XP' - P'') + (2XQ' - Q'') = \lambda\varphi(P) + \varphi(Q).$$

Ainsi, φ est bien linéaire, et donc est un endomorphisme de $\mathbf{R}_n[X]$.

⁵ C'est un vecteur propre, donc il est non nul par définition.

Rappel

Si A et B sont semblables, alors, pour tout $P \in \mathbf{K}[X]$, $P(A)$ et $P(B)$ sont semblables.

⁶ Un élément de $\mathcal{M}_2(\mathbf{R})$ est aussi un élément de $\mathcal{M}_2(\mathbf{C})$, donc on peut chercher ses valeurs propres complexes : les $\lambda \in \mathbf{C}$ tels que $A - \lambda I_2$ n'est pas inversible.

2. On a $\varphi(1) = 0$, $\varphi(X) = 2X$ et pour $k \in \llbracket 2, n \rrbracket$,

$$\varphi(X^k) = 2kX^k - k(k-1)X^{k-2}.$$

Et donc la matrice de φ dans la base canonique de $\mathbf{R}_n[X]$ est

$$\begin{pmatrix} \varphi(1) & \varphi(X) & \varphi(X^2) & \dots & \varphi(X^{n-1}) & \varphi(X^n) \\ 0 & 0 & 2 & & & 0 \\ 0 & 2 & 0 & & & \vdots \\ \vdots & 0 & 4 & \ddots & & \vdots \\ \vdots & & \ddots & \ddots & \ddots & \vdots \\ \vdots & & & \ddots & 2(n-1) & 0 \\ 0 & \dots & \dots & \dots & 0 & 2n \end{pmatrix} \begin{matrix} 1 \\ X \\ X^2 \\ \vdots \\ X^{n-1} \\ X^n \end{matrix}.$$

Cette matrice est triangulaire, donc ses valeurs propres sont ses coefficients diagonaux, à savoir : $0, 2, \dots, 2n$.

Ainsi, $\text{Spec}(\varphi) = \{2p, p \in \llbracket 0, n \rrbracket\}$.

Puisque de plus, chacune des valeurs propres n'apparaît qu'une seule fois sur la diagonale, les sous-espaces propres sont tous de dimension 1.

3. On a $P'' - 2XP' + 2pP = 0 \Leftrightarrow \varphi(P) = 2pP \Leftrightarrow P \in E_{2p}(\varphi)$.

Or, $E_{2p}(\varphi)$ est de dimension 1, donc engendré par un unique polynôme P_{2p} .

Et donc les éléments de $E_{2p}(\varphi)$ sont tous de la forme λP_{2p} , avec $\lambda \in \mathbf{R}$.

Un et un seul de ces éléments est un polynôme de coefficient dominant égal à 1, et donc $E_{2p}(\varphi)$ contient un unique polynôme unitaire.

Il existe donc un unique polynôme unitaire H_p tel que $H_p'' - 2XH_p' + 2pH_p = 0$.

4. Soit d_p le degré de H_p . Alors son coefficient dominant est 1.

Alors $2XH_p' - H_p'' = 2d_pX^{d_p} + Q_p$, avec $\deg(Q_p) < d_p$.

De même $2pH_p = 2pX^{d_p} + R_p$, avec $\deg(R_p) < d_p$.

Par identification des coefficients de degré d_p , $2p = 2d_p$ et donc $d_p = p$.



Une autre solution, plus matricielle : nous expliquons ici comment «lire» sur la matrice le fait qu'un vecteur propre associé à la valeur propre $2k$ est forcément de degré k .

L'explication est un peu longue à écrire, mais ce type de raisonnement peut tout à fait être utilisé à l'oral sans écrire tous les détails.

Puisque la matrice de f est triangulaire, et que le coefficient diagonal k se trouve sur la colonne correspondant à $f(X^k)$, on a envie de penser que le degré d'un vecteur propre associé à la valeur propre k est égal à k . C'est bien le cas, mais il y a un peu de travail pour en arriver là.

Commençons par remarquer que pour tout $k \in \llbracket 0, n \rrbracket$, $\mathbf{R}_k[X]$ est stable par f .

En effet, si $\deg P \leq k$, alors $\deg f(P) = \deg(2XP' - P'') = \deg(2XP') \leq k$.

Notons alors f_k la restriction de f à $\mathbf{R}_k[X]$, c'est-à-dire l'endomorphisme de $\mathbf{R}_k[X]$ défini par $f_k(P) = f(P)$.

Alors la matrice de f_k dans la base canonique est la «sous-matrice» de $\text{Mat}_{\mathcal{B}_{can}}(f)$ que l'on obtient en ne gardant que les $k+1$ premières lignes et les $k+1$ premières colonnes.

Et de même pour la matrice de f_{k-1} .

$$\text{Mat}_{\mathcal{B}_{can}}(f) = \begin{pmatrix} 0 & 0 & 2 & \dots & 0 & \dots & 0 \\ 0 & 2 & 0 & & & & \vdots \\ \vdots & \ddots & \ddots & & 0 & & \vdots \\ \vdots & & \ddots & \ddots & 2(k-1) & 0 & \vdots \\ \vdots & & & \ddots & 2k & & \vdots \\ \vdots & & & & \ddots & \ddots & 0 \\ 0 & \dots & \dots & \dots & \dots & 0 & 2n \end{pmatrix}$$

→ $\text{Mat}_{\mathcal{B}_{can}}(f_{k-1})$
→ $\text{Mat}_{\mathcal{B}_{can}}(f_k)$

Détails

Si le coefficient dominant de P_{2p} est a , alors le coefficient dominant de λP est λa .

Et donc λP est unitaire si et seulement si

$$\lambda a = 1 \Leftrightarrow \lambda = \frac{1}{a}.$$

Plus généralement, tout sous-espace vectoriel de $\mathbf{R}[X]$ de dimension 1 contient un unique polynôme unitaire.

Matriciellement

La stabilité se voit aussi car la matrice de f dans la base canonique est triangulaire supérieure, de sorte que $\deg f(X^k) \leq k$.

En particulier, k n'est pas valeur propre de $\text{Mat}_{\mathcal{B}_{can}}(f_{k-1})$, et donc il n'existe pas de polynôme non nul P , de degré inférieur ou égal à $k - 1$ tel que $f_{k-1}(P) = kP \Leftrightarrow f(P) = kP$.
 En revanche, k est valeur propre de f_k , et donc il existe un polynôme P non nul, de degré au plus k , tel que $f_k(P) = kP \Leftrightarrow f(P) = kP$.
 Et donc un tel P n'étant pas de degré au plus $k - 1$, il est de degré **exactement** k .
 Enfin, puisque les sous-espaces propres de f sont tous de dimension 1, $E_k(f) = \text{Vect}(P)$, qui ne contient que le polynôme nul et des polynômes de degré k .
 En particulier, tous les vecteurs propres de f associés à la valeur propre k sont de degré k .

SOLUTION DE L'EXERCICE 7.14

1. f est bien à valeurs dans $\mathcal{M}_n(\mathbf{R})$. Soient donc $M, N \in \mathcal{M}_n(\mathbf{R})$ et $\lambda \in \mathbf{R}$. Alors, par linéarité de la trace, il vient

$$\begin{aligned} f(\lambda M + N) &= \text{tr}(A)(\lambda M + N) - \text{tr}(\lambda M + N)A \\ &= \lambda(\text{tr}(A)M - \text{tr}(M)A) + (\text{tr}(A)N - \text{tr}(N)A) = \lambda f(M) + f(N). \end{aligned}$$

Donc f est linéaire et donc c'est un endomorphisme de $\mathcal{M}_n(\mathbf{R})$.

2. On a $f(A) = \text{tr}(A)A - \text{tr}(A)A = 0$.

Donc 0 est valeur propre⁷ de f et $A \in E_0(f) = \text{Ker } f$.

De plus, $M \in \text{Ker } f \Leftrightarrow \text{tr}(A)M = \text{tr}(M)A \Leftrightarrow M = \frac{\text{tr}(M)}{\text{tr}(A)}A$. Donc $\text{Ker } f \subset \text{Vect}(A)$.

On en déduit que $E_0(f) = \text{Ker}(f) = \text{Vect}(A)$.

De plus, si $\text{tr}(M) = 0$, on remarque que $f(M) = \text{tr}(A)M$. Puisqu'il existe des matrices non nulles de trace nulle, cela signifie que $\text{tr}(A)$ est valeur propre de f .

On a alors

$$M \in E_{\text{tr}(A)}(f) \Leftrightarrow \text{tr}(A)M - \text{tr}(M)A = \text{tr}(A)M \Leftrightarrow \text{tr}(M)A = 0.$$

Or A est non nulle, donc cette dernière condition est réalisée si et seulement si $\text{tr}(M) = 0$. Ainsi, $E_{\text{tr}(A)}(f) = \{M \in \mathcal{M}_n(\mathbf{R}) : \text{tr}(M) = 0\} = \text{Ker}(\text{tr})$.

Mais la trace est une forme linéaire non nulle sur $\mathcal{M}_n(\mathbf{R})$ et donc son noyau est un hyperplan de $\mathcal{M}_n(\mathbf{R})$.

Donc $\dim E_{\text{tr}(A)}(f) = \dim \mathcal{M}_n(\mathbf{R}) - 1 = n^2 - 1$.

On a donc

$$\dim E_0(f) + \dim E_{\text{tr}(A)}(f) = 1 + n^2 - 1 = n^2 = \dim \mathcal{M}_n(\mathbf{R}).$$

Donc f ne peut avoir d'autres valeurs propres : $\text{Spec}(f) = \{0, \text{tr}(A)\}$.

SOLUTION DE L'EXERCICE 7.15

1. Commençons par prouver que f est bien à valeurs dans $\mathbf{R}_n[X]$. Soit $P = \sum_{k=0}^n a_k X^k \in \mathbf{R}_n[X]$.

Alors

$$f(P) = (X + 1)(X + 3) \left(\sum_{k=0}^n k a_k X^{k-1} \right) - \sum_{k=0}^n n a_k X^{k+1} \in \mathbf{R}_{n+1}[X].$$

Alors le coefficient de degré $n + 1$ de $f(P)$ est $n a_n - n a_n = 0$. Donc $f(P)$ est de degré au plus n et donc est bien dans $\mathbf{R}_n[X]$.

Soient $P, Q \in \mathbf{R}_n[X]$ et $\lambda \in \mathbf{R}$. Alors

$$\begin{aligned} f(\lambda P + Q) &= (X + 1)(X + 3)(\lambda P + Q)' - nX(\lambda P + Q) \\ &= (X + 1)(X + 3)(\lambda P' + Q') - nX(\lambda P + Q) \\ &= \lambda((X + 1)(X + 3)P' - nXP) + ((X + 1)(X + 3)Q' - nXQ) \\ &= \lambda f(P) + f(Q). \end{aligned}$$

Et donc f est linéaire : c'est un endomorphisme de $\mathbf{R}_n[X]$.

- 2.a. Notons $P = \sum_{k=0}^d a_k X^k$, avec $a_d \neq 0$. Alors

$$f(P) = (X + 1)(X + 3) \left(\sum_{k=1}^d k a_k X^{k-1} \right) - n \sum_{k=0}^d a_k X^{k+1} = \lambda \sum_{k=0}^d a_k X^k.$$

En identifiant les coefficients des termes de degré $d + 1$, il vient $d a_d - n a_d = 0$. Et puisque $a_d \neq 0$ par hypothèse, nécessairement $d = n$: P est de degré n .

⁷ A est non nulle car de trace non nulle. Donc A est un vecteur propre de f associé à la valeur propre 0.

Intuition

Si on n'a pas vu la forme linéaire qui se cache ici, on peut voir que pour choisir une matrice de trace nulle, on a le choix de tous les coefficients, sauf du dernier en bas à droite qui doit être choisi pour que la somme des coefficients diagonaux soit égale à 1. On a donc $n^2 - 1$ «degrés de liberté», ce qui doit laisser penser que $\text{Ker}(\text{tr})$ est de dimension $n^2 - 1$, et donc un hyperplan. Et donc le noyau d'une forme linéaire.

Remarque

Demander à a_d d'être non nul revient à dire que d est précisément le degré de P . On peut bien entendu toujours écrire $P = \sum_{k=0}^n a_k X^k$, mais alors a_n, a_{n-1} , etc peuvent être nuls.

2.b. Soit z une racine complexe de P . Alors, en évaluant la relation $f(P) = \lambda P$ en $X = z$, il vient

$$(z+1)(z+3)P'(z) - zP(z) = \lambda P(z) \Leftrightarrow (z+1)(z+3)P'(z) = 0.$$

Donc soit $(z+1)(z+3) = 0 \Leftrightarrow z \in \{-1, -3\}$, soit $P'(z) = 0$.

Supposons donc que $z \notin \{-1, -3\}$, de sorte que $P'(z) = 0$. Alors z est une racine double⁸ de P .

Notons r l'ordre de multiplicité de z , de sorte que $P(X) = (X-z)^r Q(z)$, avec $Q(z) \neq 0$.

Dans ce cas, nous savons que z est racine de P' d'ordre de multiplicité $r-1$:

$P'(X) = (X-z)^{r-1} Q_1(X)$, avec $Q_1(z) \neq 0$. Alors

$$f(P) = \lambda P \Leftrightarrow (X+1)(X+3)P'(X) = (\lambda - nX)P(X) \Leftrightarrow (X-1)(X+3)(X-z)^{r-1} Q_1(X) = (\lambda - nX)(X-z)^r Q(X).$$

En divisant par $(X-z)^{r-1}$, il vient $(X+1)(X+3)Q_1(z) = (\lambda - nX)(X-z)Q(X)$.

Mais z est racine du membre de gauche et pas du membre de droite, d'où une contradiction.

On en déduit que $z \in \{-1, -3\}$.

Puisque tout polynôme de $\mathbf{C}[X]$ se décompose en produit de facteurs de degré 1, il existe un complexe $a \neq 0$ et des complexes z_1, \dots, z_n tels que $P(X) = a(X-z_1) \cdots (X-z_n)$.

Mais les z_i sont tous égaux à -1 ou -3 d'après ce qui précède, et a est le coefficient dominant de P , donc est réel : $P = a(X+1)^k(X+3)^\ell$, avec $k+\ell = \deg(P) = n$.

3. On a

$$\begin{aligned} f(P_{k,\ell}) &= (X+1)(X+3) \left(k(X+1)^{k-1}(X+3)^\ell + (X+1)^k \ell (X+3)^{\ell-1} \right) - nX(X+1)^k(X+3)^\ell \\ &= (X+1)^k(X+3)^\ell (k(X+3) + \ell(X+1) - nX) \\ &= (X+1)^k(X+3)^\ell (3k + \ell) = (3k + \ell)P_{k,\ell}. \end{aligned}$$

On en déduit que $P_{k,\ell}$ est vecteur propre de f associé à la valeur propre $3k + \ell = 3k + n - k = 2k + n$.

Et donc f possède $n+1$ valeurs propres qui sont $n, n+2, n+4, \dots, 3n$.

Puisque $\dim \mathbf{R}_{n+1}[X] = n+1$, f ne possède pas d'autres valeurs propres, et ses sous-espaces propres sont tous de dimension 1 : $\text{Spec}(f) = \{n+2k, k \in \llbracket 0, n \rrbracket\}$ et pour tout $k \in \llbracket 0, n \rrbracket$, $E_{n+2k}(f) = \text{Vect}(P_{k,n-k})$.

SOLUTION DE L'EXERCICE 7.16

Pour tout $x \in \text{Ker}(\varphi)$, on a $f(x) = x$. Donc 1 est valeur propre de f et $\text{Ker}(\varphi) \subset E_1(f)$, qui est donc de dimension au moins égale⁹ à $\dim E - 1$.

Notons que $f(u) = (1 + \varphi(u))u$.

Donc si $u \notin \text{Ker}(\varphi)$, alors $1 + \varphi(u)$ est une valeur propre de f , et $\text{Vect}(u) \subset E_{1+\varphi(u)}(f)$.

On a alors $\dim E_1(f) + \dim E_{1+\varphi(u)} \geq \dim E$, et donc cette inégalité est une égalité, de sorte que $\dim E_1(f) = \dim E - 1$ et $\dim E_{1+\varphi(u)} = 1$. Ainsi, f ne possède pas d'autres valeurs propres : $\text{Spec}(f) = \{1, 1 + \varphi(u)\}$ et $E_1(f) = \text{Ker}(\varphi)$ et $E_{1+\varphi(u)} = \text{Vect}(u)$.

Au contraire, si $u \notin \text{Ker}(\varphi)$, montrons que 1 est la seule valeur propre de f .

Supposons que λ soit une valeur propre de f différente de 1, et soit x un vecteur propre associé.

Alors $\varphi(x) \neq 0$, et donc $f(x) = x + \varphi(x)u = \lambda x$. Soit encore $x = \frac{\varphi(x)}{\lambda - 1}u$.

Mais puisque $\varphi(u) = 0$, on a alors $\varphi(x) = \frac{\varphi(x)}{\lambda - 1}\varphi(u) = 0$. D'où une contradiction : f ne possède qu'une seule valeur propre qui est 1, et $E_1(f) = \text{Ker}(\varphi)$.

SOLUTION DE L'EXERCICE 7.17

1. Cette matrice est difficile à aborder car de taille n quelconque. Faisons toutefois à la main les premières opérations d'un pivot. On a, pour a réel,

$$M_n - aI_n = \begin{pmatrix} 1-a & 1 & \dots & \dots & 1 \\ 1 & -a & \dots & \dots & 0 \\ 0 & 1 & \ddots & & 0 \\ \vdots & & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & 1 & -a \end{pmatrix} \xleftrightarrow{L_2 \leftrightarrow L_1} \begin{pmatrix} 1 & -a & \dots & \dots & 0 \\ 1-a & 1 & \dots & \dots & 1 \\ 0 & 1 & \ddots & & 0 \\ \vdots & & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & 1 & -a \end{pmatrix}$$

⁸ Car est à la fois racine de P et de P' .

Coeff. dominant

Ne pas oublier la constante a qui est le coefficient dominant de P , c'est-à-dire le coefficient du terme de degré n .

⁹ $\text{Ker}(\varphi)$ est un hyperplan de E .

Méthode

On peut aussi faire un «vrai» pivot avec $n = 2$ et $n = 3$, et regarder si un motif transposable à une matrice de taille n émerge des calculs que l'on réalise.

$$\begin{aligned}
 & \xleftrightarrow{L_2 \leftarrow L_2 - (1-a)L_1} \begin{pmatrix} 1 & -a & \dots & \dots & 0 \\ 0 & -P_2(a) & 1 & \dots & 1 \\ 0 & 1 & -a & & 0 \\ \vdots & & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & 1 & -a \end{pmatrix} \xleftrightarrow{L_3 \leftrightarrow L_2} \begin{pmatrix} 1 & -a & \dots & \dots & 0 \\ 0 & 1 & -a & \dots & 0 \\ 0 & -P_2(a) & 1 & & 1 \\ \vdots & & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & 1 & -a \end{pmatrix} \\
 & \xleftrightarrow{L_3 \leftarrow L_3 + P_2(a)L_2} \begin{pmatrix} 1 & -a & \dots & \dots & 0 \\ 0 & 1 & -a & \dots & 0 \\ 0 & 0 & -P_3(a) & & 1 \\ \vdots & & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & 1 & -a \end{pmatrix}
 \end{aligned}$$

$1 + a(1 - a) = 1 + a - a^2$. On reconnaît alors $-P_2(a)$.

De même, ici, après calculs, on reconnaît $-P_3(a)$.

Remarquons que pour tout i ,

$$1 - aP_i(a) = 1 - a(a^i - a^{i-1} - \dots - a - 1) = 1 - a^{i+1} + a^i + \dots + a^2 + a = -P_{i+1}(a).$$

En procédant de proche en proche, en permutant à chaque fois la ligne i avec la ligne $i + 1$, puis en procédant à l'opération $L_{i+1} \leftarrow L_{i+1} + P_i(a)L_i$, on obtient la matrice échelonnée

$$\begin{pmatrix} 1 & -a & 0 & \dots & 0 \\ 0 & 1 & \ddots & & \vdots \\ \vdots & & \ddots & \ddots & \vdots \\ \vdots & & & 1 & -a \\ 0 & \dots & \dots & 0 & -P_n(a) \end{pmatrix}$$

Ainsi, a est valeur propre si et seulement si cette matrice est de rang strictement inférieur à n , c'est-à-dire si et seulement si $P_n(a) = 0$.

2. Un vecteur $X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ est dans $E_a(M_n)$ si et seulement si $M_n X = aX \Leftrightarrow (M_n - aI_n)X = 0$. Le système associé est alors

$$\begin{cases} (1-a)x_1 + x_2 + \dots + x_n = 0 \\ x_1 - ax_2 = 0 \\ x_2 - ax_3 = 0 \\ \vdots \\ x_{n-2} - ax_{n-1} = 0 \\ x_{n-1} - ax_n = 0 \end{cases} \Leftrightarrow \begin{cases} (1-a)x_1 + x_2 + \dots + x_n = 0 \\ x_1 = ax_2 \\ x_2 = ax_3 \\ \vdots \\ x_{n-2} = ax_{n-1} \\ x_{n-1} = ax_n \end{cases}$$

Alors en remplaçant de proche en proche, on a $x_{n-2} = ax_{n-1} = a^2x_n$, puis $x_{n-3} = a^3x_n, \dots, x_1 = a^{n-1}x_n$.

La première équation s'écrit alors

$$(1 - a)a^{n-1}x_n + a^{n-2}x_n + \dots + x_n = 0 \Leftrightarrow (-a^n + a^{n-1} + \dots + a + 1)x_n = 0 \Leftrightarrow P_n(a)x_1 = 0.$$

Si a est valeur propre, alors cette équation est en fait $0 = 0$.

Ainsi, on a $X = \begin{pmatrix} a^{n-1}x_1 \\ a^{n-2}x_2 \\ \vdots \\ ax_1 \\ x_1 \end{pmatrix}$ de sorte que $E_a(M_n) = \text{Vect} \left(\begin{pmatrix} a^{n-1} \\ a^{n-2} \\ \vdots \\ a \\ 1 \end{pmatrix} \right)$.

Logique !

Si a est valeur propre, il existe nécessairement une infinité de solutions au système, qui n'est donc pas de Cramer. Et donc l'une (au moins) des équations doit être inutile.

SOLUTION DE L'EXERCICE 7.18

1. Soit $\lambda \in \text{Spec}(f)$, et soit $x \in E_\lambda(f)$. Alors $f(g(x)) = g(f(x)) = g(\lambda x) = \lambda g(x)$, de sorte que $g(x) \in E_\lambda(f)$. On en déduit que $E_\lambda(f)$ est stable par g , et donc tous les sous-espaces propres de f sont stables par g .

2. Soit $x \in E$ un vecteur propre de f , associé à la valeur propre λ . Puisque $E_\lambda(f)$ est de dimension 1, x est une base de $E_\lambda(f)$ (car un vecteur propre est toujours non nul par définition), et $E_\lambda(f)$ étant stable par g , on en déduit qu'il existe $\mu \in \mathbf{R}$ tel que $g(x) = \mu x$. Ainsi, x est bien vecteur propre de g .
3. Si f possède n valeurs propres distinctes, alors tous les sous-espaces propres associés sont de dimension un. De plus, la somme des sous-espaces propres est égale à E tout entier, et donc la concaténation de bases des sous-espaces propres est une base de E formée de vecteurs propres de f .

Mais d'après la question précédente, il s'agit également de vecteurs propres de g . Par conséquent, les matrices de f et g sont toutes deux diagonales dans cette base.

4. D'après ce qui a été dit précédemment, il existe une base de \mathbf{R}^n dans laquelle la matrice de f est diagonale. Notons D cette matrice. Alors A et B sont semblables car elles représentent le même endomorphisme. Par conséquent, il existe $P \in GL_n(\mathbf{R})$ telle que $A = P^{-1}DP$. Si $B^2 = A$, alors B commute avec A car $AB = B^3 = BA$. Soit alors g l'endomorphisme de \mathbf{R}^n dont la matrice dans la base canonique de \mathbf{R}^n est B . Alors f et g commutent. D'après ce qui a été dit dans les questions précédentes, dans une base de \mathbf{R}^n formée de vecteurs propres de f , la matrice de g doit être diagonale. En particulier, $\Delta = PBP^{-1}$ doit être diagonale, et vérifier $\Delta^2 = D$.

Nous sommes donc ramenés à trouver le nombre de matrices diagonales Δ vérifiant $\Delta^2 = D = \text{Diag}(\lambda_1, \dots, \lambda_n)$, où les λ_i sont les valeurs propres de f .

Mais rappelons que si $\Delta = \begin{pmatrix} \delta_1 & & \\ & \ddots & \\ & & \delta_n \end{pmatrix}$, alors $\Delta^2 = \begin{pmatrix} \delta_1^2 & & \\ & \ddots & \\ & & \delta_n^2 \end{pmatrix}$.

Ainsi, les coefficients diagonaux de Δ doivent être les racines carrées de ceux de D , ou les opposées de ces racines carrées.

- Si D (et donc A) possède au moins une valeur propre strictement négative, alors il n'existe pas de matrice Δ vérifiant $\Delta^2 = D$, et donc pas de matrice B vérifiant $B^2 = A$.
- Si D (et donc A) a toutes ses valeurs propres strictement positives, alors il existe 2^n matrices Δ vérifiant $\Delta^2 = D$, et donc 2^n matrices vérifiant $B^2 = A$.

Notons que si $\Delta^2 = D$, alors en posant $B = P^{-1}\Delta P$, on a bien $B^2 = A$, et que si Δ_1, Δ_2 sont deux matrices distinctes vérifiant $\Delta_1^2 = \Delta_2^2 = D$, alors les matrices $P^{-1}\Delta_1 P$ et $P^{-1}\Delta_2 P$ sont également distinctes, tout simplement car elles n'ont pas les mêmes valeurs propres.

- Enfin, si 0 est valeur propre de D et que toutes les autres valeurs propres sont strictement positives, alors il existe 2^{n-1} matrices Δ , et donc autant de B .

Plus généralement

Tout sous-espace stable de dimension 1 est nécessairement engendré par un vecteur propre.

Détails

Pour chaque coefficient δ_i , il y a deux choix : $\delta_i = \pm\sqrt{\lambda_i}$.

Détails

Si $\lambda_i = 0$, alors il n'y a qu'une option pour δ_i : $\delta_i = 0$.

VARIABLES ALÉATOIRES À DENSITÉ.

Cours 245
8.1 Espérance, variance, moments 245
8.2 Lois à densité usuelles 250
8.3 Somme de variables aléatoires à densité 255
8.4 Quelques mots de la simulation en Scilab 259
8.5 Compléments 259
Méthodes 262
Exercices 264
Correction des exercices 267

Dans tout ce chapitre, (Ω, \mathcal{A}, P) désigne un espace probabilisé.

8.1 ESPÉRANCE, VARIANCE, MOMENTS

8.1.1 Quelques rappels

Définition 8.1 – On dit qu’une variable aléatoire réelle sur (Ω, \mathcal{A}, P) est une variable à densité si sa fonction de répartition F_X est :

- continue sur \mathbf{R}
- de classe \mathcal{C}^1 sur \mathbf{R} sauf éventuellement en un nombre fini de points.

Dans ce cas, on appelle densité de X toute fonction positive sur \mathbf{R} ne différant de F'_X qu’en un nombre fini¹ de points.

¹ Et en particulier en les points où F'_X n’est pas définie.

Proposition 8.2 : Soit f une fonction de \mathbf{R} dans \mathbf{R} . Alors f est la densité d’une variable aléatoire si et seulement si :

- f est continue sur \mathbf{R} sauf éventuellement en un nombre fini de points
- f est positive sur \mathbf{R} .
- $\int_{-\infty}^{+\infty} f(t) dt$ converge et vaut 1.

Proposition 8.3 : Si X est une variable aléatoire à densité sur (Ω, \mathcal{A}, P) , et si f_X est une densité de X , alors

$$\forall x \in \mathbf{R}, F_X(x) = P(X \leq x) = \int_{-\infty}^x f_X(t) dt.$$

Plus généralement, $\forall a, b \in \mathbf{R}$ avec $a \leq b$,

$$P(a \leq X \leq b) = \int_a^b f(t) dt.$$

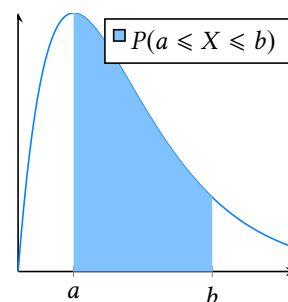


FIGURE 8.1– Lien entre probabilité et intégrale.

Remarque. En particulier, $\forall a \in \mathbf{R}, P(X = a) = \int_a^a f(t) dt = 0$.

Ceci implique notamment que X prend une infinité de valeurs.

⚠ Il existe² des variables aléatoires qui ne sont ni discrètes, ni à densité, même si cette année nous nous limiterons à ces deux types de variables aléatoires.

² **Exercice :** trouver une fonction de répartition qui n’est ni celle d’une variable discrète, ni celle d’une variable à densité.

8.1.2 Espérance

Définition 8.4 – Soit X une variable aléatoire à densité sur (Ω, \mathcal{A}, P) . On dit que X admet une espérance si $\int_{-\infty}^{+\infty} xf(x) dx$ converge absolument. On appelle alors espérance de X , et on note $E(X)$ le réel défini par

$$E(X) = \int_{-\infty}^{+\infty} xf(x) dx.$$

Remarque. En fait, puisque $xf(x)$ est de signe constant sur $] - \infty, 0[$ et sur $]0, \infty[$, cette intégrale est absolument convergente si et seulement si elle est convergente.

Discret/densité
Si l'on remplace la somme par une intégrale et les $P(X = k)$ par $f(x)$, cette formule ressemble beaucoup à celle définissant l'espérance des variables discrètes.

Exemple 8.5 Loi de Cauchy

Soit f la fonction définie sur \mathbf{R} par $f(x) = \frac{1}{\pi} \frac{1}{1+x^2}$. Alors f est continue sur \mathbf{R} , positive et $f(x) \underset{x \rightarrow +\infty}{\sim} \frac{1}{\pi x^2}$, donc par comparaison à une intégrale de Riemann, $\int_0^{+\infty} f(x) dx$ converge. Par parité de f , on a donc $\int_{-\infty}^{+\infty} f(x) dx$ qui converge et

$$\int_{-\infty}^{+\infty} f(x) dx = 2 \int_0^{+\infty} f(x) dx$$

Pour $A > 0$, on a

$$\int_0^A \frac{1}{\pi} \frac{1}{1+x^2} dx = \frac{1}{\pi} \text{Arctan}(A) \xrightarrow{A \rightarrow +\infty} \frac{1}{2}.$$

On en déduit que

$$\int_{-\infty}^{+\infty} f(x) dx = 1,$$

de sorte que f est une densité. Soit X une variable aléatoire de densité f .

On a $xf(x) \underset{x \rightarrow +\infty}{\sim} \frac{1}{\pi} \frac{x}{1+x^2} = \frac{1}{\pi x}$, de sorte que $\int_0^{+\infty} xf(x) dx$ diverge, et donc X n'admet pas d'espérance.

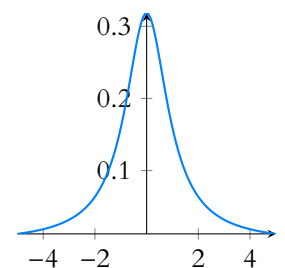


FIGURE 8.2– Densité de la loi de Cauchy.

Proposition 8.6 (Linéarité de l'espérance) : Soient X et Y deux variables aléatoires à densité admettant une espérance. Alors pour tous $\lambda, \mu \in \mathbf{R}$, la variable aléatoire $\lambda X + \mu Y$ admet une espérance et

$$E(\lambda X + \mu Y) = \lambda E(X) + \mu E(Y).$$

Proposition 8.7 : Soient X, Y deux variables aléatoires à densité sur (Ω, \mathcal{A}, P) , admettant une espérance et telles que $0 \leq |X| \leq Y$ presque sûrement. Alors X admet une espérance et $|E(X)| \leq E(Y)$.

Précision
Ce résultat, bien qu'intuitif, est loin d'être trivial : déjà il faudrait prouver que $\lambda X + \mu Y$ est une variable aléatoire, puis en déterminer une densité (si elle en admet une !). Nous admettons donc ce résultat.

8.1.3 Fonction d'une variable aléatoire

Rappel : Si X est une variable aléatoire de densité f_X , alors pour tous $a \in \mathbf{R}^*, b \in \mathbf{R}$, $aX + b$ est encore une variable à densité dont une densité est donnée par

$$x \mapsto \frac{1}{|a|} f\left(\frac{x-b}{a}\right).$$

Exemple 8.8

Soit X une variable aléatoire suivant la loi $\mathcal{E}(2)$. Alors une densité de X est donnée

$$\text{par } f_X(x) = \begin{cases} 2e^{-2x} & \text{si } x \geq 0 \\ 0 & \text{sinon} \end{cases}$$

Si l'on pose $Y = -2X + 3$, alors Y est une variable à densité, dont une densité est donnée par

$$f_Y(x) = \frac{1}{2}f_X\left(\frac{3-x}{2}\right) = \begin{cases} e^{-(3-x)} & \text{si } \frac{3-x}{2} \geq 0 \\ 0 & \text{sinon} \end{cases} = \begin{cases} e^{x-3} & \text{si } x \leq 3 \\ 0 & \text{sinon} \end{cases}$$

Proposition 8.9 : Soit X une variable aléatoire réelle sur (Ω, \mathcal{A}, P) , et soit $g : X(\Omega) \rightarrow \mathbf{R}$ une fonction continue. Alors $g(X)$ est encore une variable aléatoire réelle sur (Ω, \mathcal{A}, P) .

Exemple 8.10

Si X est une variable aléatoire, alors $X^2, e^X, aX+b$ sont encore des variables aléatoires.

Ce résultat ne dit pas si $g(X)$ est encore une variable à densité, ni comment trouver une telle densité. Si g est \mathcal{C}^1 sur $X(\Omega)$ et strictement monotone, alors on pourrait montrer «à la main» que $g(X)$ est encore à densité.

Nous ne donnons aucun résultat général allant dans ce sens, et pour déterminer si $g(X)$ est ou non à densité et en déterminer (sous réserve d'existence) une densité, on reviendra à la définition en vérifiant si sa fonction de répartition est continue sur \mathbf{R} et \mathcal{C}^1 sauf en un nombre fini de points.

Exemple 8.11

Soit X une variable aléatoire suivant la loi uniforme $\mathcal{U}([0, 1])$, et soit $Y = -\ln(X)$.

Alors $Y(\Omega) = [0, +\infty[$, et donc pour $x < 0$, $P(Y \leq x) = 0$.

Pour $x \geq 0$, on a, par croissance de la fonction exponentielle,

$$P(Y \leq x) = P(-\ln X \leq x) = P(\ln X \geq -x) = P(X \geq e^{-x}).$$

Or $e^{-x} \in [0, 1]$ et donc $P(X \geq e^{-x}) = 1 - F_X(e^{-x}) = 1 - e^{-x}$.

Ainsi, la fonction de répartition de Y est

$$F_Y(x) = \begin{cases} 0 & \text{si } x < 0 \\ 1 - e^{-x} & \text{si } x \geq 0 \end{cases}$$

On pourrait prouver qu'il s'agit bien de la fonction de répartition d'une variable à densité et en trouver une densité en dérivant F_X là où c'est possible, mais il est plus simple de reconnaître la fonction de répartition d'une loi exponentielle $\mathcal{E}(1)$. Et donc $Y \leftrightarrow \mathcal{E}(1)$.

Théorème 8.12 (Théorème de transfert) : Soit X une variable aléatoire admettant une densité f_X nulle en dehors de $]a, b[$ ($a, b \in \mathbf{R}$), et soit $\varphi :]a, b[\rightarrow \mathbf{R}$ une fonction continue sauf éventuellement en un nombre fini de points. Alors $E(\varphi(X))$ existe si et seulement si $\int_a^b \varphi(x)f_X(x) dx$ converge absolument. En cas de convergence, on a

$$E(\varphi(X)) = \int_a^b \varphi(x)f_X(x) dx.$$

ln(0)...

La variable X peut prendre la valeur 0, ce qui est alors problématique pour passer au logarithme. Toutefois, ceci ne se produit qu'avec probabilité nulle, et donc n'a pas de réelle influence. Si l'on souhaite être exact, on n'a qu'à dire que Y prend la valeur 0 si $X = 0$, ce qui ne changera pas la fonction de répartition de Y et donc pas sa loi.

Loi

Rappelons que, par définition, la loi d'une variable aléatoire est déterminée par sa fonction de répartition : deux variables ont même loi si et seulement si elles ont même fonction de répartition. Il n'y a donc aucun besoin de repasser par les densités.

Démonstration. Admis. Voir le complément 8.5.2 pour une preuve dans un cas particulier. \square

Exemple 8.13

Soit X une variable aléatoire suivant la loi exponentielle de paramètre λ , et soit Y la variable aléatoire définie par $Y = e^X$. Alors Y admet une espérance si et seulement si l'intégrale $\int_{-\infty}^{+\infty} e^t f_X(t) dt = \int_0^{+\infty} e^t \lambda e^{-\lambda t} dt$ converge absolument.

Or, $\int_0^{+\infty} e^{-(\lambda-1)t} dt$ converge si et seulement si $\lambda - 1 > 0 \Leftrightarrow \lambda > 1$.

De plus, dans ce cas, on a alors $E(Y) = \frac{\lambda}{\lambda - 1}$.

Corollaire 8.14 – Si X est une variable aléatoire à densité admettant une espérance, alors pour $a \neq 0$ et $b \in \mathbf{R}$, $aX + b$ admet une espérance, et

$$E(aX + b) = aE(X) + b.$$

8.1.4 Moments d'ordre supérieur, variance

Définition 8.15 – Soit X une variable aléatoire à densité sur (Ω, \mathcal{A}, P) , de densité f_X , et soit $r \in \mathbf{N}^*$. On dit que X admet un moment d'ordre r si X^r admet une espérance.

Par le théorème de transfert, c'est le cas si et seulement si

$$\int_{-\infty}^{+\infty} x^r f_X(x) dx \text{ converge absolument}^3.$$

Dans ce cas, on note $m_r(X) = E(X^r)$.

³ Comme pour l'espérance, cette intégrale converge si et seulement si elle converge absolument.

Proposition 8.16 : Soit X une variable aléatoire à densité sur (Ω, \mathcal{A}, P) et soit $r, q \in \mathbf{N}^*$, avec $q \leq r$. Si X admet un moment d'ordre r , alors X admet un moment d'ordre q .

Démonstration. On a, pour tout $t \in \mathbf{R}$, $|t|^q \leq 1 + |t|^r$, et donc

$$|t|^q f(t) \leq f(t) + |t|^r f(t).$$

Mais $\int_{-\infty}^{+\infty} f(t) dt$ converge⁴, et donc si $\int_{-\infty}^{+\infty} |t|^r f(t) dt$ converge, il en est de même de $\int_{-\infty}^{+\infty} |t|^q f(t) dt$. \square

Détails

Si $|t| \leq 1$, c'est évident car $|t|^q \leq 1$ et sinon $|t|^q \leq |t|^r$.

⁴ Et vaut 1 !

Remarque. En particulier, une variable aléatoire à densité qui admet un moment d'ordre 2 admet une espérance.

Définition 8.17 – Soit X une variable aléatoire réelle de densité f_X admet une espérance. On dit que X admet une variance si $(X - E(X))^2$ admet une espérance. D'après le théorème de transfert, c'est le cas si et seulement si

$$\int_{-\infty}^{+\infty} (x - E(X))^2 f_X(x) dx \text{ converge absolument.}$$

Dans ce cas, la variance de X est

$$V(X) = \int_{-\infty}^{+\infty} (x - E(X))^2 f_X(x) dx = E((X - E(X))^2)$$

Proposition 8.18 : Si X est une variable aléatoire réelle de densité f admettant une variance, alors $V(X) > 0$.

On appelle alors écart-type de X , et on note $\sigma(X)$ le nombre défini

$$\sigma(X) = \sqrt{V(X)}.$$

Démonstration. Notons que la fonction $x \mapsto (x - E(X))^2 f(x)$ est positive sur \mathbf{R} , et donc son intégrale est positive ou nulle, de sorte que $V(X) \geq 0$.

De plus, si $V(X) = 0$ alors $(x - E(X))^2 f(x)$ est nulle en tous les points de continuité de f . Puisque $x - E(X)$ ne s'annule que pour $x = E(X)$, cela signifie que f est nulle sauf éventuellement en un nombre fini de points, ce qui implique alors $\int_{-\infty}^{+\infty} f(t) dt = 0$, ce qui contredit le fait que $\int_{-\infty}^{+\infty} f(t) dt = 1$. \square

Proposition 8.19 (Formule de Huygens) : Soit X une variable aléatoire à densité admettant une espérance. Alors X admet une variance si et seulement si X admet un moment d'ordre 2. Lorsque c'est le cas, on a alors

$$V(X) = E(X^2) - E(X)^2.$$

Démonstration. On a, $\forall x \in \mathbf{R}$,

$$(x - E(X))^2 f_X(x) = (x^2 - 2xE(X) + E(X)^2) f_X(x) = x^2 f_X(x) - 2xE(X) f_X(x) + E(X)^2 f_X(x).$$

Mais $\int_{-\infty}^{+\infty} 2xE(X) f_X(x) dx$ converge et vaut $2E(X)^2$, et il en est de même de

$$\int_{-\infty}^{+\infty} E(X)^2 f_X(x) dx = E(X)^2 \int_{-\infty}^{+\infty} f(t) dt = E(X)^2.$$

Ainsi, $\int_{-\infty}^{+\infty} (x - E(X))^2 f_X(x) dx$ converge si et seulement si $\int_{-\infty}^{+\infty} x^2 f_X(x) dx$ converge.

En cas de convergence, on a alors

$$V(X) = \int_{-\infty}^{+\infty} (x - E(X))^2 f_X(x) dx = \int_{-\infty}^{+\infty} x^2 f_X(x) dx - 2E(X)^2 + E(X)^2 = E(X^2) - E(X)^2.$$

\square

Proposition 8.20 : Soit X une variable aléatoire à densité admettant une variance et soient $a, b \in \mathbf{R}$. Alors $aX + b$ admet une variance et $V(aX + b) = a^2 V(X)$.

Détails

On a une égalité reliant quatre intégrales, dont deux convergent. Si la troisième converge, alors la quatrième aussi (une somme d'intégrales convergentes est convergente). Et inversement si la troisième diverge, nécessairement la quatrième aussi.

Démonstration. Puisque $(aX + b)^2 = a^2X^2 + 2abX + b^2$ et que $E(X^2)$ et $E(X)$ existent, $E((aX + b)^2)$ existe, et donc $aX + b$ admet un moment d'ordre deux, et donc une variance. Par la formule de Huygens, on a alors :

$$\begin{aligned} V(aX + b) &= E((aX + b)^2) - E(aX + b)^2 = a^2E(X^2) + 2abE(X) + b^2 - (aE(X) + b)^2 \\ &= a^2E(X^2) - a^2E(X)^2 = a^2V(X). \end{aligned}$$

□

8.1.5 Variables centrées et centrées réduites

Définition 8.21 – Une variable aléatoire réelle sur (Ω, \mathcal{A}, P) est dite centrée si X admet une espérance et si $E(X) = 0$.

Une variable aléatoire réelle est dite centrée réduite si X admet une variance et que

$$E(X) = 0 \text{ et } V(X) = 1.$$

Proposition 8.22 : Si X est une variable aléatoire réelle discrète ou à densité admettant une espérance, alors $X - E(X)$ est centrée.

Si X est une variable aléatoire réelle discrète ou à densité admettant une variance non nulle⁵, alors $\frac{X - E(X)}{\sigma(X)}$ est centrée réduite.

⁵ Cette hypothèse est en fait équivalente à dire que X n'est pas presque certaine.

Démonstration. C'est une conséquence directe des formules classiques $E(X + a) = E(X) + a$ et $V(aX + b) = a^2V(X)$. □

8.2 LOIS À DENSITÉ USUELLES

8.2.1 Lois uniformes

Loi uniforme sur $[0, 1]$

Définition 8.23 – Une variable aléatoire réelle sur (Ω, \mathcal{A}, P) suit la loi uniforme sur $[0, 1]$ si elle admet pour densité la fonction f_X définie par

$$f_X(x) = \begin{cases} 1 & \text{si } x \in [0, 1] \\ 0 & \text{sinon} \end{cases}.$$

On note alors $X \hookrightarrow \mathcal{U}([0, 1])$.

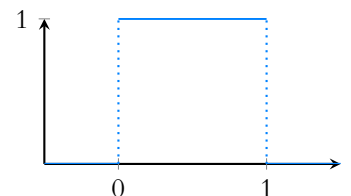


FIGURE 8.3– Densité de la loi uniforme sur $[0, 1]$

Proposition 8.24 : Si $X \hookrightarrow \mathcal{U}([0, 1])$, alors sa fonction de répartition est donnée par

$$f_X(x) = \begin{cases} 0 & \text{si } x \leq 0 \\ x & \text{si } 0 \leq x \leq 1. \\ 1 & \text{si } x \geq 1 \end{cases}$$

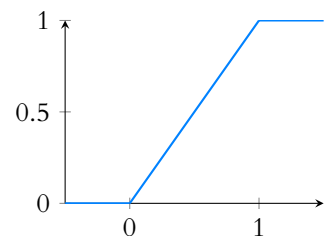


FIGURE 8.4– La fonction de répartition de $\mathcal{U}([0, 1])$.

Proposition 8.25 : Si $X \hookrightarrow \mathcal{U}([0, 1])$, alors $E(X) = \frac{1}{2}$ et $V(X) = \frac{1}{12}$.

Démonstration. On a (sous réserve de convergence de l'intégrale) :

$$V(X) = \int_{-\infty}^{+\infty} \left(x - \frac{1}{2}\right)^2 f_X(x) dx = \int_0^1 \left(x - \frac{1}{2}\right)^2 dx = \int_{-\frac{1}{2}}^{\frac{1}{2}} u^2 du = \left[\frac{u^3}{3}\right]_{-1/2}^{1/2} = \frac{1}{12}.$$

□

Loi uniforme sur $[a, b]$

Définition 8.26 – Soient $a < b$. Une variable aléatoire réelle X suit la loi uniforme sur $[a, b]$ si elle admet pour densité la fonction f_X définie par

$$f_X(x) = \begin{cases} \frac{1}{b-a} & \text{si } a \leq x \leq b \\ 0 & \text{sinon} \end{cases}$$

On note alors $X \leftrightarrow \mathcal{U}([a, b])$.

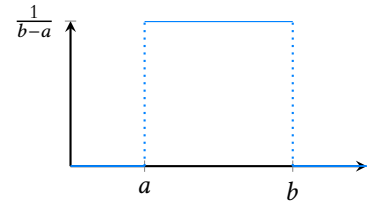


FIGURE 8.5– Densité de la loi uniforme sur $[a, b]$

Remarque. On parle aussi parfois de loi uniforme sur $[a, b]$, sur $[a, b[$ ou sur $]a, b]$. Il s’agit bien de la même loi, une loi uniforme prenant les valeurs a ou b avec probabilité nulle.

Proposition 8.27 : $X \leftrightarrow \mathcal{U}([a, b])$ si et seulement si $\frac{1}{b-a}(X-a) \leftrightarrow \mathcal{U}([0, 1])$.

Démonstration. Il s’agit juste d’une transformation affine : une densité de $\frac{1}{b-a}(X-a)$ est

$$x \mapsto (b-a)f_X((b-a)x+a) = \begin{cases} 1 & \text{si } a \leq (b-a)x+a \leq b \\ 0 & \text{sinon} \end{cases} = \begin{cases} 1 & \text{si } 0 \leq x \leq 1 \\ 0 & \text{sinon} \end{cases}$$

On reconnaît alors la densité d’une loi $\mathcal{U}([0, 1])$ et donc $\frac{1}{(b-a)}(X-a) \leftrightarrow \mathcal{U}([0, 1])$.

La réciproque se prouve de la même manière. □

Remarque. Cette formule, associée à la connaissance de la fonction de répartition de $\mathcal{U}([0, 1])$ devrait permettre de déterminer la fonction de répartition de X . Mais si l’on se souvient que cette fonction est affine par morceaux, vaut 0 en a et 1 en b , il n’est pas très dur de retrouver son expression à partir d’un dessin.

Corollaire 8.28 – Si $X \leftrightarrow \mathcal{U}([a, b])$, alors

$$E(X) = \frac{a+b}{2} \text{ et } V(X) = \frac{(b-a)^2}{12}.$$

Démonstration. Puisque $\frac{1}{b-a}(X-a) \leftrightarrow \mathcal{U}([0, 1])$, on a $\frac{1}{2} = E\left(\frac{1}{b-a}(X-a)\right) = \frac{1}{b-a}E(X) - \frac{a}{b-a}$ donc

$$E(X) = (b-a)\left(\frac{1}{2} + \frac{a}{b-a}\right) = \frac{b-a}{2} + a = \frac{a+b}{2}.$$

De même, on a $\frac{1}{12} = V\left(\frac{1}{b-a}(X-a)\right) = \frac{1}{(b-a)^2}V(X)$ donc $V(X) = \frac{(b-a)^2}{12}$. □

Intuition
Il s’agit surtout de se rappeler qu’une transformée affine d’une loi uniforme suit encore une loi uniforme. Et alors, si X prend ses valeurs dans $[a, b]$, alors $X-a$ prend ses valeurs dans $[0, b-a]$ et donc $\frac{X-a}{b-a}$ prend ses valeurs dans $[0, 1]$.

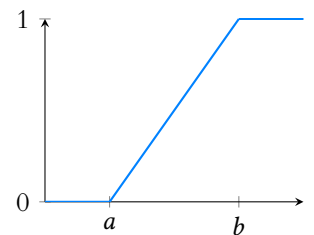


FIGURE 8.6– La fonction de répartition de $\mathcal{U}([a, b])$.

8.2.2 Lois exponentielles

Définition 8.29 – Soit X une variable aléatoire réelle sur (Ω, \mathcal{A}, P) , et soit $\lambda > 0$. On dit que X suit la loi exponentielle de paramètre λ si elle admet pour densité la fonction f_X définie par

$$f_X(x) = \begin{cases} 0 & \text{si } x < 0 \\ \lambda e^{-\lambda x} & \text{si } x \geq 0 \end{cases}$$

On note alors $X \leftrightarrow \mathcal{E}(\lambda)$.

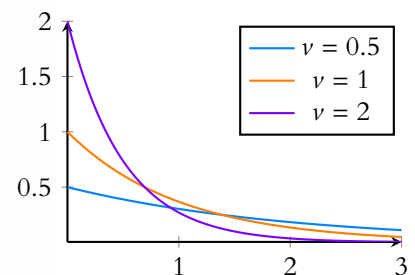


FIGURE 8.7– Densités de différentes lois exponentielles

Proposition 8.30 : Si $X \hookrightarrow \mathcal{E}(\lambda)$, alors sa fonction de répartition est donnée par

$$F_X(x) = \begin{cases} 0 & \text{si } x < 0 \\ 1 - e^{-\lambda x} & \text{si } x \geq 0 \end{cases}$$

De plus, on a $E(X) = \frac{1}{\lambda}$ et $V(X) = \frac{1}{\lambda^2}$.

Fonction de rép.

La loi exponentielle est une des rares lois (avec la loi uniforme) dont on sait expliciter la fonction de répartition, il faut savoir en profiter.

Démonstration. La fonction de répartition et l'espérance ont déjà été vues en première année, attardons nous sur la variance.

Sous réserve de convergence, on a

$$E(X^2) = \int_0^{+\infty} x^2 \lambda e^{-\lambda x} dx.$$

Effectuons le changement de variable $t = \lambda x$. Alors

$$\int_0^{+\infty} x^2 \lambda e^{-\lambda x} dx = \lambda \int_0^{+\infty} \left(\frac{t}{\lambda}\right) e^{-t} \frac{dt}{\lambda} = \frac{1}{\lambda^2} \int_0^{+\infty} t^2 e^{-t} dt = \frac{\Gamma(3)}{\lambda^2} = \frac{2}{\lambda^2}.$$

On en déduit que $V(X) = E(X^2) - E(X)^2 = \frac{2}{\lambda^2} - \frac{1}{\lambda^2} = \frac{1}{\lambda^2}$. \square

Proposition 8.31 : Soit $\lambda > 0$. Alors $X \hookrightarrow \mathcal{E}(\lambda) \Leftrightarrow \lambda X \hookrightarrow \mathcal{E}(1)$.

Démonstration. Pour $x \in \mathbf{R}$, on a

$$P(\lambda X \leq x) = P\left(X \leq \frac{x}{\lambda}\right) = \begin{cases} 1 - e^{-\lambda \frac{x}{\lambda}} & \text{si } x \geq 0 \\ 0 & \text{sinon} \end{cases} = \begin{cases} 1 - e^{-x} & \text{si } x \geq 0 \\ 0 & \text{sinon} \end{cases}$$

On reconnaît alors la fonction de répartition d'une loi exponentielle de paramètre 1, donc $\lambda X \hookrightarrow \mathcal{E}(1)$.

Inversement, on montre de la même manière que si $\lambda X \hookrightarrow \mathcal{E}(1)$, alors $X \hookrightarrow \mathcal{E}(\lambda)$. \square

Astuce

L'important est de savoir qu'une telle relation existe. Si on ne sait plus où se trouve le λ dans la formule, il est facile de le retrouver en utilisant les espérances : l'espérance d'une loi $\mathcal{E}(\lambda)$ est $\frac{1}{\lambda}$, et celle d'une loi $\mathcal{E}(1)$ est 1. Pour passer de la première à la seconde, il faut multiplier par λ .

8.2.3 Lois gamma

Proposition 8.32 : Soit $\nu > 0$. Alors la fonction

$$f : x \mapsto \begin{cases} 0 & \text{si } x \leq 0 \\ \frac{1}{\Gamma(\nu)} x^{\nu-1} e^{-x} & \text{si } x > 0 \end{cases}$$

est la densité d'une variable aléatoire.

Si X est une variable aléatoire réelle admettant pour densité f , on dit que X suit la loi γ de paramètre ν , et on note $X \hookrightarrow \gamma(\nu)$.

Remarque

La constante $\frac{1}{\Gamma(\nu)}$ sert à garantir que l'intégrale sur \mathbf{R} vaut 1. Sur le même principe, on peut créer une densité à partir de toute fonction f positive sur \mathbf{R} , continue sauf en un nombre fini de points et d'intégrale convergente en divisant f par son intégrale sur \mathbf{R} .

Démonstration. f est continue en tout point de \mathbf{R} , sauf peut-être⁶ en 0, et il est clair que f est positive. De plus,

$$\int_{-\infty}^{+\infty} f(x) dx = \frac{1}{\Gamma(\nu)} \int_0^{+\infty} t^{\nu-1} e^{-x} dx = 1.$$

Donc f est bien une densité. \square

⁶ En fait, pour $\nu > 1$, f est continue en 0. Mais ne perdons pas de temps à le vérifier : une densité à le droit d'avoir un nombre fini de points de discontinuité.

Remarque. Puisque $\Gamma(1) = 1$, la loi $\gamma(1)$ n'est autre que la loi $\mathcal{E}(1)$.

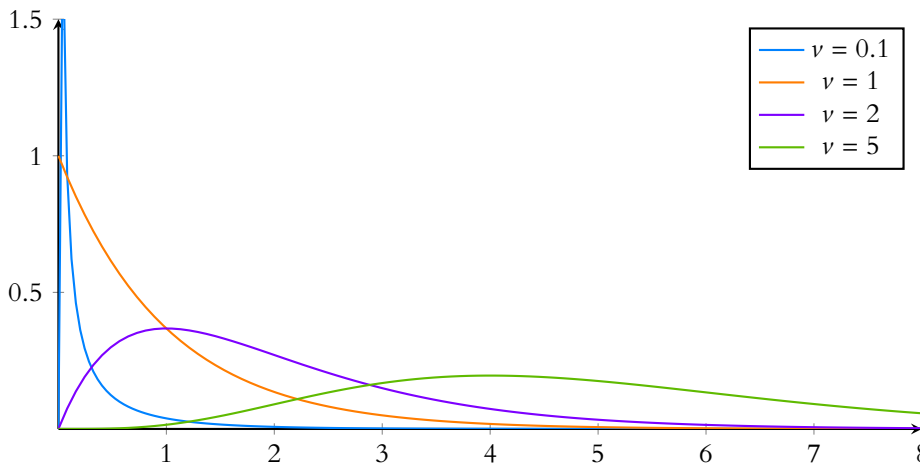


FIGURE 8.8 – Densités de différentes lois γ .

Remarque : notons que la densité d'une loi $\gamma(v)$ est bornée si et seulement si $v \geq 1$.

Proposition 8.33 : Si $X \leftrightarrow \gamma(v)$, alors $E(X) = v$ et $V(X) = v$.

Démonstration. Sous réserve de convergence, on a

$$E(X) = \int_{-\infty}^{+\infty} xf(x) dx = \frac{1}{\Gamma(v)} \int_0^{+\infty} xx^{v-1} e^{-x} dx = \frac{\Gamma(v+1)}{\Gamma(v)} = v.$$

De même,

$$E(X^2) = \frac{1}{\Gamma(v)} \int_0^{+\infty} x^{v+1} e^{-x} dx = \frac{\Gamma(v+2)}{\Gamma(v)} = \frac{(v+1)\Gamma(v+1)}{\Gamma(v)} = v(v+1)$$

On en déduit que

$$V(X) = E(X^2) - E(X)^2 = v(v+1) - v^2 = v.$$

□

Rappel
 $\Gamma(v+2) = (v+1)\Gamma(v+1)$.

8.2.4 Lois normales

Définition 8.34 – Soit $\mu \in \mathbf{R}$ et $\sigma > 0$. Alors une variable aléatoire réelle X suit la loi normale $\mathcal{N}(\mu, \sigma^2)$ si elle admet pour densité la fonction

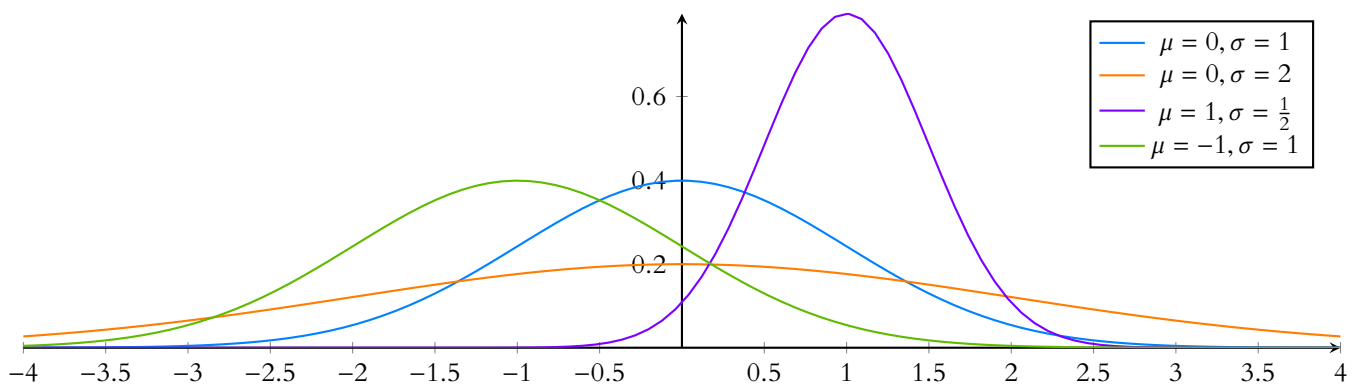
$$f_X : x \mapsto \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}.$$


FIGURE 8.9 – Quelques lois normales

Proposition 8.35 : Si $X \hookrightarrow \mathcal{N}(\mu, \sigma^2)$, alors pour $a, b \in \mathbf{R}$, $a \neq 0$, on a

$$aX + b \hookrightarrow \mathcal{N}(a\mu + b, a^2\sigma^2).$$

En particulier, $X \hookrightarrow \mathcal{N}(\mu, \sigma^2) \Leftrightarrow \frac{X - \mu}{\sigma} \hookrightarrow \mathcal{N}(0, 1)$.

Démonstration. Si f désigne la densité de $\mathcal{N}(\mu, \sigma^2)$, alors une densité de $aX + b$ est donnée par

$$x \mapsto \frac{1}{|a|} f\left(\frac{x-b}{a}\right) = \frac{1}{|a|\sigma\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{x-(a\mu+b)}{a\sigma}\right)^2\right).$$

On reconnaît alors une densité de la loi $\mathcal{N}(a\mu + b, a^2\sigma^2)$. \square

Proposition 8.36 : Si $X \hookrightarrow \mathcal{N}(0, 1)$, alors $E(X) = 0$ et $V(X) = 1$. Autrement dit X est centrée réduite.

Démonstration. L'espérance a déjà été calculée en première année, intéressons-nous au moment d'ordre 2.

Sous réserve d'existence,

$$E(X^2) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} x^2 e^{-x^2/2} dx = \frac{2}{\sqrt{2\pi}} \int_0^{+\infty} x^2 e^{-x^2/2} dx.$$

Soit $A > 0$. Alors par intégration par parties, avec $u = x$, $v' = xe^{-x^2/2}$

$$\int_0^A x^2 e^{-x^2/2} dx = [-xe^{-x^2/2}]_0^A + \int_0^A e^{-x^2/2} dx \xrightarrow{A \rightarrow \infty} \int_0^{+\infty} e^{-x^2/2} dx = \sqrt{\frac{\pi}{2}}.$$

Donc $E(X^2) = \frac{2}{\sqrt{2\pi}} \sqrt{\frac{\pi}{2}} = 1$, et donc $V(X) = E(X^2) - E(X)^2 = E(X^2) = 1$. \square

Corollaire 8.37 – Si $X \hookrightarrow \mathcal{N}(\mu, \sigma^2)$, alors

$$E(X) = \mu \text{ et } V(X) = \sigma^2.$$

Démonstration. On sait que $\frac{X - \mu}{\sigma} \hookrightarrow \mathcal{N}(0, 1)$, donc $\frac{1}{\sigma}(E(X) - \mu) = 0$, de sorte que $E(X) = \mu$.

De même, $\frac{1}{\sigma^2}V(X) = 1$, donc $V(X) = \sigma^2$. \square

Définition 8.38 – On note $\Phi : \mathbf{R} \rightarrow \mathbf{R}$ la fonction de répartition d'une variable aléatoire X suivant la loi normale centrée réduite $\mathcal{N}(0, 1)$. C'est-à-dire

$$\forall x \in \mathbf{R}, \Phi(x) = \int_{-\infty}^x \frac{1}{\sqrt{2\pi}} e^{-t^2/2} dt.$$

Proposition 8.39 : $\forall x \in \mathbf{R}, \Phi(-x) = 1 - \Phi(x)$. En particulier, $\Phi(0) = \frac{1}{2}$.

Démonstration. On a, par le changement de variable $u = -t$

$$\Phi(-x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{-x} e^{-t^2/2} dt = \frac{1}{\sqrt{2\pi}} \int_x^{+\infty} e^{-u^2/2} du = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-u^2/2} du - \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-u^2/2} du = 1 - \Phi(x).$$

\square

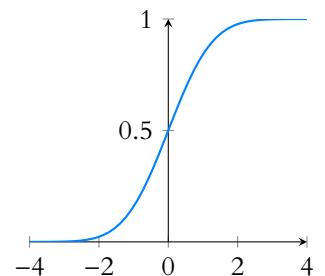


FIGURE 8.10– La fonction Φ .

Remarques. • On ne sait pas exprimer Φ à l'aide de fonctions usuelles (résultat difficile prouvé par Liouville vers 1840). En revanche, on sait qu'il s'agit d'une primitive de la densité d'une loi $\mathcal{N}(0, 1)$.

• Pour déterminer la fonction de répartition d'une loi normale quelconque, on se ramènera systématiquement à une loi normale centrée réduite.

Plus précisément : si $X \leftrightarrow \mathcal{N}(\mu, \sigma^2)$ alors $\frac{X - \mu}{\sigma} \leftrightarrow \mathcal{N}(0, 1)$ et donc la fonction de répartition de X est donnée par

$$\forall x \in \mathbf{R}, \quad F_X(x) = P(X \leq x) = P\left(\frac{X - \mu}{\sigma} \leq \frac{x - \mu}{\sigma}\right) = \Phi\left(\frac{x - \mu}{\sigma}\right).$$

Primitive

Et donc on connaît Φ' , qui vaut

$$\Phi'(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}.$$

8.2.5 Tableau récapitulatif

Nom	Paramètres	Densité f_X	Fonction de répartition	$E(X)$	$V(X)$
Loi uniforme sur $[0, 1]$	-	$f_X(t) = \begin{cases} 1 & \text{si } t \in [0, 1] \\ 0 & \text{sinon} \end{cases}$	$F_X(x) = \begin{cases} 0 & \text{si } x < 0 \\ x & \text{si } 0 \leq x \leq 1 \\ 1 & \text{si } x > 1 \end{cases}$	$\frac{1}{2}$	$\frac{1}{12}$
Loi uniforme sur $[a, b]$	$a < b$	$f_X(t) = \begin{cases} \frac{1}{b-a} & \text{si } t \in [a, b] \\ 0 & \text{sinon} \end{cases}$	$F_X(x) = \begin{cases} 0 & \text{si } x < a \\ \frac{x-a}{b-a} & \text{si } a \leq x \leq b \\ 1 & \text{si } x > b \end{cases}$	$\frac{a+b}{2}$	$\frac{(b-a)^2}{12}$
Loi exponentielle $\mathcal{E}(\lambda)$	$\lambda > 0$	$f_X(t) = \begin{cases} 0 & \text{si } t < 0 \\ \lambda e^{-\lambda t} & \text{si } t \geq 0 \end{cases}$	$F_X(x) = \begin{cases} 0 & \text{si } x < 0 \\ 1 - e^{-\lambda x} & \text{si } x \geq 0 \end{cases}$	$\frac{1}{\lambda}$	$\frac{1}{\lambda^2}$
Loi gamma $\gamma(\nu)$	$\nu > 0$	$f_X(t) = \begin{cases} 0 & \text{si } t \leq 0 \\ \frac{1}{\Gamma(\nu)} t^{\nu-1} e^{-t} & \text{si } t > 0 \end{cases}$	Pas au programme	ν	ν
Loi normale $\mathcal{N}(0, 1)$	-	$\frac{1}{\sqrt{2\pi}} e^{-t^2/2}$	$F_X(x) = \Phi(x)$	0	1
Loi normale $\mathcal{N}(\mu, \sigma^2)$	$\mu \in \mathbf{R}$ $\sigma > 0$	$f_X(t) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{t-\mu}{\sigma}\right)^2}$	$F_X(x) = \Phi\left(\frac{x-\mu}{\sigma}\right)$	μ	σ^2

8.3 SOMME DE VARIABLES ALÉATOIRES À DENSITÉ

8.3.1 Produit de convolution

Théorème 8.40 (Produit de convolution des variables à densité) : Soient X et Y deux variables aléatoires réelles indépendantes sur (Ω, \mathcal{A}, P) , de densités respectives f_X et f_Y . Notons $Z = X + Y$. Si la fonction h définie par la relation

$$h(x) = \int_{-\infty}^{+\infty} f_X(t) f_Y(x - t) dt$$

est définie et continue sauf peut-être en un nombre fini de points, alors c'est une densité de Z .

Discret/densité

Ce théorème ressemble beaucoup au produit de convolution discret : on a remplacé la somme par une intégrale et les $P(X = k)$ par les densités.

En pratique

Lorsqu'on aura à utiliser le produit de convolution, il n'y aura jamais aucune subtilité sur la continuité de h , et on pourra toujours dire sans précaution (autre que l'indépendance !) que h est une densité de $X + Y$.

Remarques. • Si f_X ou f_Y sont bornées (ce qui est le cas pour une loi uniforme, une loi exponentielle, une loi normale ou pour une loi $\gamma(\nu)$ avec $\nu \geq 1$), alors h est automatiquement définie et continue sauf en un nombre fini de points.

• Ce résultat est faux sans l'hypothèse d'indépendance !

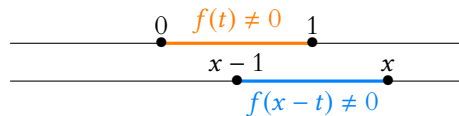
Exemple 8.41

Soient X et Y deux variables aléatoires indépendantes suivant la loi uniforme sur $[0, 1]$, et soit $Z = X + Y$. Notons f la densité commune de X et Y , et soit

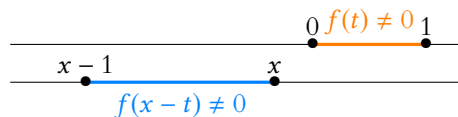
$$h(x) = \int_{-\infty}^{+\infty} f(t)f(x-t) dt$$

Mais $f(t)f(x-t) \neq 0$ si et seulement si

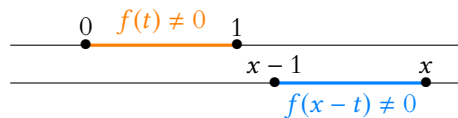
$$\begin{cases} 0 \leq t \leq 1 \\ 0 \leq x-t \leq 1 \end{cases} \Leftrightarrow \begin{cases} 0 \leq t \leq 1 \\ x-1 \leq t \leq x \end{cases} \Leftrightarrow \max(0, x-1) \leq t \leq \min(1, x).$$



Donc si $x < 0$, pour $t \leq 0$, $f(t) = 0$, et pour $t \geq 0$, $x-t < 0$ de sorte que $f(x-t) = 0$. Ainsi, pour tout $t \in \mathbf{R}$, $f(t)f(x-t) = 0$ et donc $h(x) = 0$.

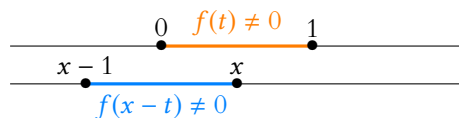


De même, si $x-1 > 1 \Leftrightarrow x > 2$, alors $h(x) = 0$.



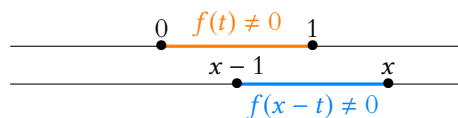
Pour $x \in [0, 1]$, on a

$$h(x) = \int_0^x 1 \times 1 dt = x.$$



Et pour $x \in]1, 2]$,

$$h(x) = \int_{x-1}^1 1 \times 1 dt = 1 - (x-1) = 2-x.$$



$$\text{Ainsi } h(x) = \begin{cases} x & \text{si } x \in [0, 1] \\ 2-x & \text{si } x \in]1, 2] \\ 0 & \text{sinon} \end{cases}$$

Cette fonction est définie et continue sur \mathbf{R} , donc Z est une variable aléatoire à densité, et une densité de Z est h .

Méthode

Un dessin permet de se représenter les valeurs de t pour lesquelles l'intégrande est non nul, étape indispensable avant de se lancer dans le calcul de l'intégrale. Il faut alors bien comprendre que les deux dessins représentent l'axe des t , et imaginer que x se déplace. Il s'agit alors de distinguer (suivant les valeurs de x) pour quelles valeurs de t on a à la fois $f(t) \neq 0$ et $f(x-t) \neq 0$.

Méthode

Il n'y a aucun besoin de réaliser 4 dessins différents (en cas de besoin, le faire au brouillon suffit largement !), le premier dessin doit permettre de distinguer tous ces cas, il suffit d'imaginer que x se déplace.

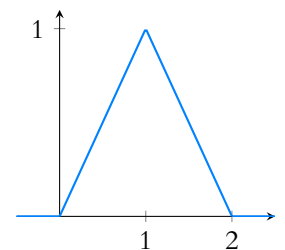


FIGURE 8.11– La densité h

Proposition 8.42 : Si X et Y sont deux variables aléatoires réelles à densité indépendantes admettant une variance, alors XY admet une espérance, et $E(XY) = E(X)E(Y)$.

Proposition 8.43 : Si X et Y sont deux variables aléatoires réelles à densité indépendantes admettant une variance, alors $X + Y$ admet une variance et

$$V(X + Y) = V(X) + V(Y).$$

Remarque
La notion de covariance se généralise aux variables à densité, mais cela dépasse le cadre du programme d'ECS.

8.3.2 Stabilité des lois γ

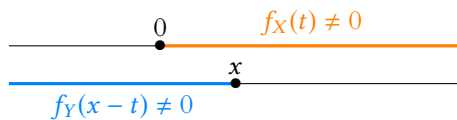
Proposition 8.44 : Soit $v_1, v_2 > 0$, et soient X, Y deux variables aléatoires réelles à densité indépendantes sur le même espace probabilisé (Ω, \mathcal{A}, P) telles que $X \hookrightarrow \gamma(v_1)$ et $Y \hookrightarrow \gamma(v_2)$. Alors $X + Y \hookrightarrow \gamma(v_1 + v_2)$.

Démonstration. Soit f_X une densité de X et f_Y une densité de Y . Soit alors h la fonction définie, sous réserve de convergence, par

$$h(x) = \int_{-\infty}^{+\infty} f_X(t)f_Y(x - t) dt.$$

D'après le théorème 8.40, si h est définie et continue sauf en un nombre fini de points, c'est une densité de $X + Y$.

On a alors $f_X(t) \neq 0 \Leftrightarrow t > 0$ et $f_Y(x - t) \neq 0 \Leftrightarrow x - t > 0 \Leftrightarrow t < x$.



Si $x \leq 0$, alors $h(x) = 0$, car on ne peut avoir à la fois $t \geq 0$ et $x - t \geq 0$ (f_X et f_Y sont nulles sur \mathbf{R}_-).

Si $x > 0$, alors $f_X(t)f_Y(x - t) = 0$ pour $t \leq 0$ et $x - t \leq 0$, soit $t \notin]0, x[$, de sorte que

$$h(x) = \frac{1}{\Gamma(v_1)\Gamma(v_2)} \int_0^x t^{v_1-1}(x - t)^{v_2-1} e^{-t-(x-t)} dt = \frac{1}{\Gamma(v_1)\Gamma(v_2)} \int_0^x t^{v_1-1}(x - t)^{v_2-1} e^{-x} dt.$$

Procédons au changement de variable $u = \frac{t}{x}$. On a alors⁷

$$\int_0^x t^{v_1-1}(x - t)^{v_2-1} e^{-x} dt = e^{-x} x^{v_1+v_2-1} \int_0^1 u^{v_1-1}(1 - u)^{v_2-1} du.$$

Posons $B(v_1, v_2) = \int_0^1 u^{v_1-1}(1 - u)^{v_2-1} du \in \mathbf{R}$.

Il s'agit d'une intégrale convergente, car au voisinage de 0, $u^{v_1-1}(1 - u)^{v_2-1} \underset{u \rightarrow 0^+}{\sim} u^{v_1-1}$,

avec $v_1 - 1 > -1$, et donc, par comparaison avec l'intégrale de Riemann convergente

$$\int_0^{1/2} \frac{du}{u^{1-v_1}} du, \int_0^{1/2} u^{v_1-1}(1 - u)^{v_2-1} du \text{ converge.}$$

On conclut de même au voisinage de 1 en notant que $u^{v_1-1}(1 - u)^{v_2-1} \underset{u \rightarrow 1^-}{\sim} (1 - u)^{v_2-1}$, avec $v_2 - 1 > -1$.

On a donc

$$h(x) = \begin{cases} \frac{B(v_1, v_2)}{\Gamma(v_1)\Gamma(v_2)} x^{v_1+v_2-1} e^{-x} & \text{si } x > 0 \\ 0 & \text{si } x \leq 0 \end{cases}$$

Cette fonction est bien définie, et elle est continue sauf peut-être en 0 : c'est une densité de $X + Y$.

En particulier, $\int_{-\infty}^{+\infty} h(x) dx = 1$ et donc

$$1 = \frac{B(v_1, v_2)}{\Gamma(v_1)\Gamma(v_2)} \int_0^{+\infty} x^{v_1+v_2-1} e^{-x} dx = \frac{B(v_1, v_2)\Gamma(v_1 + v_2)}{\Gamma(v_1)\Gamma(v_2)}.$$

Remarque
D'après la remarque suivant le théorème 8.40, c'est automatique si l'une des deux densités est bornée, ce qui est le cas si $v_1 \geq 1$ ou $v_2 \geq 1$.

⁷ Sous réserve de convergence de ces intégrales.

Ainsi, $B(v_1, v_2) = \frac{\Gamma(v_1)\Gamma(v_2)}{\Gamma(v_1 + v_2)}$. On en déduit que

$$\forall x > 0, h(x) = \frac{B(v_1, v_2)x^{v_1+v_2-1}e^{-x}}{\Gamma(v_1)\Gamma(v_2)} = \frac{1}{\Gamma(v_1 + v_2)}x^{v_1+v_2-1}e^{-x}.$$

On reconnaît alors la densité de la loi $\gamma(v_1 + v_2)$, donc $Z \hookrightarrow \gamma(v_1 + v_2)$. \square

Remarque. Si l'on se souvient que la somme de deux lois γ est encore une loi γ , il est facile de retrouver le résultat par linéarité de l'espérance : une loi γ est caractérisée par son espérance, et si $X + Y$ suit une loi γ , son paramètre est nécessairement $E(X) + E(Y) = v_1 + v_2$.

Corollaire 8.45 – Si X et Y sont indépendantes et suivent la loi $\mathcal{E}(1)$, alors $X + Y \hookrightarrow \gamma(2)$.

Démonstration. $X \hookrightarrow \mathcal{E}(1) \Leftrightarrow X \hookrightarrow \gamma(1)$. \square

8.3.3 Stabilité des lois normales

Proposition 8.46 : Soient X, Y deux variables aléatoires réelles indépendantes sur (Ω, \mathcal{A}, P) telles que $X \hookrightarrow \mathcal{N}(\mu_1, \sigma_1^2)$ et $Y \hookrightarrow \mathcal{N}(\mu_2, \sigma_2^2)$. Alors

$$X + Y \hookrightarrow \mathcal{N}(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2).$$

Démonstration. Commençons par le cas où X et Y sont toutes deux centrées, c'est-à-dire si $\mu_1 = \mu_2 = 0$. Alors une densité de $Z = X + Y$ est donnée par

$$h(x) = \frac{1}{\sigma_1\sigma_2 2\pi} \int_{-\infty}^{+\infty} e^{-\frac{1}{2}\left(\frac{t}{\sigma_1}\right)^2} e^{-\frac{1}{2}\left(\frac{x-t}{\sigma_2}\right)^2} dt = \frac{e^{-\frac{x^2}{2\sigma_2^2}}}{2\pi\sigma_1\sigma_2} \int_{-\infty}^{+\infty} e^{-\frac{1}{2}\left(\frac{t^2}{\sigma_1^2} + \frac{t^2}{\sigma_2^2} - \frac{2xt}{\sigma_2^2}\right)} dt$$

Mais

$$\frac{t^2}{\sigma_1^2} + \frac{t^2}{\sigma_2^2} - \frac{2xt}{\sigma_2^2} = \frac{\sigma_1^2 + \sigma_2^2}{\sigma_1^2\sigma_2^2}t^2 - 2t\frac{x}{\sigma_2^2}.$$

En posant $\sigma = \sqrt{\sigma_1^2 + \sigma_2^2}$, on a

$$\frac{t^2}{\sigma_1^2} + \frac{t^2}{\sigma_2^2} - \frac{2xt}{\sigma_2^2} = \frac{\sigma_1^2 + \sigma_2^2}{\sigma_1^2\sigma_2^2}t^2 - 2t\frac{x}{\sigma_2^2} = \left(\frac{\sigma t}{\sigma_1\sigma_2} - \frac{x\sigma_1}{\sigma_2\sigma}\right)^2 - \frac{x^2\sigma_1^2}{\sigma_2^2\sigma^2}$$

de sorte que

$$h(x) = \frac{1}{2\pi\sigma_1\sigma_2} e^{-\frac{x^2}{2\sigma_2^2}\left(1 - \frac{\sigma_1^2}{\sigma^2}\right)} \int_{-\infty}^{+\infty} e^{-\frac{1}{2}\left(\frac{\sigma t}{\sigma_1\sigma_2} - \frac{x\sigma_1}{\sigma_2\sigma}\right)^2} dt.$$

Réalisons le changement de variable affine $u = \frac{\sigma}{\sigma_1\sigma_2}t - \frac{x\sigma_1}{\sigma_2\sigma}$, soit $dt = \frac{\sigma_1\sigma_2}{\sigma}du$. Alors

$$h(x) = \frac{1}{2\pi\sigma_1\sigma_2} e^{-\frac{x^2}{2\sigma_2^2}\left(1 - \frac{\sigma_1^2}{\sigma^2}\right)} \int_{-\infty}^{+\infty} e^{-\frac{u^2}{2}} \frac{\sigma_1\sigma_2}{\sigma} du = \frac{1}{2\pi\sigma} e^{-\frac{x^2}{2\sigma^2}} \int_{-\infty}^{+\infty} e^{-u^2/2} du = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{x^2}{2\sigma^2}}.$$

On a bien là la densité de la loi normale $\mathcal{N}(0, \sigma^2)$.

Dans le cas général, si $X \hookrightarrow \mathcal{N}(\mu_1, \sigma_1^2)$, alors $X - \mu_1 \hookrightarrow \mathcal{N}(0, \sigma_1^2)$.

De même, $Y - \mu_2 \hookrightarrow \mathcal{N}(0, \sigma_2^2)$.

On a alors $(X + Y) - (\mu_1 + \mu_2) \hookrightarrow \mathcal{N}(0, \sigma_1^2 + \sigma_2^2)$ et donc

$$X + Y \hookrightarrow \mathcal{N}(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2).$$

\square

⚠ On n'ajoute pas les écart-types, mais bien les variances. Par conséquent, l'écart-type de $X + Y$ est $\sqrt{\sigma_1^2 + \sigma_2^2}$.

Remarque. Comme pour les lois γ , il suffit en fait de se rappeler que la somme de deux lois normales indépendantes est encore une loi normale.

Alors son espérance sera forcément la somme des espérances (par linéarité de l'espérance) et sa variance sera la somme des variances (car X et Y sont indépendantes).

8.4 QUELQUES MOTS DE LA SIMULATION EN Scilab

La commande `grand` permet également de simuler des variables à densité, et Scilab connaît toutes les lois à densité que nous venons d'étudier. On utilisera pour cela

- `grand(n,p,'unf',a,b)` pour simuler une loi uniforme sur $[a,b]$
- `grand(n,p,'nor',mu,sigma)` pour simuler une loi normale de moyenne μ et d'écart-type σ
- `grand(n,p,'exp',lambda)` pour simuler une loi exponentielle d'espérance $\frac{1}{\lambda}$.
- `grand(n,p,'gam',nu,1)` pour simuler une loi γ d'espérance ν .

Uniformes

Attention à ne pas confondre les uniformes discrètes et à densité : `uin` (pour «uniform integer») pour les uniformes discrètes, et `unf` pour les uniformes à densité.

Attention

Il faut entrer $1/\lambda$ comme paramètre pour une loi $\mathcal{E}(\lambda)$.

8.5 COMPLÉMENTS

8.5.1 Analogie entre variables discrètes et variables à densité.

Nous savons déjà que les séries et les intégrales partagent de nombreuses propriétés : la linéarité, la croissance, des critères de comparaison semblables, etc.

Ceci explique que les formules vues sur les variables à densité soient similaires à des formules rencontrées précédemment sur les variables discrètes : il suffit généralement de remplacer les $P(X = x)$ par la densité f_X et les sommes par des intégrales.

Le tableau suivant résume ces analogies :

Variables discrètes	Variables à densité
$P(X = x)$	$f_X(t)$
$\sum_{x \in X(\Omega)} P(X = k) = 1$	$\int_{-\infty}^{+\infty} f_X(t) dt = 1$
$P(X \leq x) = \sum_{\substack{y \in X(\Omega) \\ y \leq x}} P(X = y)$	$P(X \leq x) = \int_{-\infty}^x f_X(t) dt$
$E(X) = \sum_{x \in X(\Omega)} xP(X = x)$	$E(X) = \int_{-\infty}^{+\infty} t f_X(t) dt$
Sous réserve de convergence absolue $E(g(X)) = \sum_{x \in X(\Omega)} g(x)P(X = x)$	Sous réserve de convergence absolue $E(g(X)) = \int_{-\infty}^{+\infty} g(t)f_X(t) dt$
Si X et Y sont indépendantes $P(X + Y = n) = \sum_{k \in Y(\Omega)} P(X = k)P(Y = n - k)$	Si X et Y sont indépendantes $f_{X+Y}(x) = \int_{-\infty}^{+\infty} f_X(t)f_Y(x - t) dt$

Notons en particulier que ceci explique bien l'origine de la formule du produit de convolution pour les variables à densité si on a compris celle pour les variables discrètes⁸.

8.5.2 Sur le théorème de transfert

Nous donnons ici une preuve du théorème de transfert dans le cas où φ est une fonction strictement monotone de classe \mathcal{C}^1 .

Commençons par traiter un exemple à la main : supposons que $X \leftrightarrow \mathcal{E}(1)$ et que $Y = \frac{1}{X}$.

Comment étudier l'espérance de Y si l'on ne dispose pas du théorème de transfert ?

Puisque X est à valeurs dans \mathbf{R}_+^* , il en est de même de Y .

⁸ Qui n'est qu'une application de la formule des probabilités totales.

Donc pour $x \leq 0$, $P(Y \leq x) = 0$.

Pour $x > 0$, on a

$$P(Y \leq x) = P\left(\frac{1}{X} \leq x\right) = P\left(X \geq \frac{1}{x}\right) = 1 - \left(1 - e^{-\frac{1}{x}}\right) = e^{-\frac{1}{x}}.$$

Donc la fonction de répartition de Y est $F_Y(x) = \begin{cases} 0 & \text{si } x \leq 0 \\ e^{-\frac{1}{x}} & \text{si } x > 0 \end{cases}$.

Il est facile de vérifier que cette fonction est continue sur \mathbf{R} et \mathcal{C}^1 sauf éventuellement en

0, de sorte que Y est une variable à densité de densité $f_Y : x \mapsto \begin{cases} 0 & \text{si } x \leq 0 \\ \frac{1}{x^2} e^{-\frac{1}{x}} & \text{si } x > 0 \end{cases}$.

Donc Y admet une espérance si et seulement si l'intégrale $\int_0^{+\infty} x f_Y(x) dx = \int_0^{+\infty} \frac{1}{x} e^{-\frac{1}{x}} dx$ converge.

Un changement de variable $t = \frac{1}{x}$ dans cette intégrale nous donne alors

$$\int_0^{+\infty} t e^{-t} \frac{t}{t^2} = \int_0^{+\infty} \frac{1}{t} e^{-t} dt.$$

On reconnaît alors exactement l'intégrale qu'aurait fait apparaître le théorème de transfert.

Passons à présent à un cas bien plus général : soit X une variable aléatoire de densité f_X , nulle en dehors de $]a, b[$ (avec éventuellement $a = -\infty$ ou $b = +\infty$).

Ceci signifie donc $\int_a^b f_X(t) dt = 1$, et par conséquent X est⁹ à valeurs dans $]a, b[$: $P(X \in]a, b[) = 1$.

Soit alors φ une fonction de classe \mathcal{C}^1 , strictement monotone sur $]a, b[$. Dans la suite nous allons supposer que φ est strictement croissante, mais la preuve serait la même dans le cas d'une fonction décroissante, il faudrait juste penser à changer le sens des inégalités en appliquant φ .

Notons $\alpha = \lim_{x \rightarrow a^+} \varphi(x)$ et $\beta = \lim_{x \rightarrow b^-} \varphi(x)$, de sorte que $Y = \varphi(X)$ est à valeurs dans $]\alpha, \beta[$.

Ceci implique déjà que pour $x \leq \alpha$, $F_Y(x) = 0$ et pour $x \geq \beta$, $F_Y(x) = 1$.

Pour $x \in]\alpha, \beta[$, on a alors

$$F_Y(x) = P(Y \leq x) = P(\varphi(X) \leq x) = P(X \leq \varphi^{-1}(x)) = F_X(\varphi^{-1}(x)).$$

Donc la fonction de répartition de Y est donnée par

$$F_Y(x) = \begin{cases} 0 & \text{si } x \leq \alpha \\ F_X(\varphi^{-1}(x)) & \text{si } \alpha < x < \beta \\ 1 & \text{si } x \geq \beta \end{cases}$$

Par composition de fonctions de classe \mathcal{C}^1 , il est donc évident que F_Y est \mathcal{C}^1 , sauf peut-être en α et en β .

D'autre part, on a $\lim_{x \rightarrow \alpha^+} \varphi^{-1}(x) = a$ et puisque F_X est continue en a , $\lim_{y \rightarrow a} F_X(t) = F_X(a) = 0$.

Et donc par composition de limite, on a $\lim_{x \rightarrow \alpha^+} F_Y(x) = 0$. Et donc F_Y est continue en a .

On prouve de la même manière que F_Y est continue en b , et donc continue sur \mathbf{R} .

Ainsi, Y est une variable à densité, et l'une de ses densités est par exemple

$$f_Y : x \mapsto \begin{cases} (\varphi^{-1})'(x) F_X'(\varphi^{-1}(x)) & \text{si } \alpha < x < \beta \\ 0 & \text{sinon} \end{cases} = \begin{cases} (\varphi^{-1})'(x) f_X(\varphi^{-1}(x)) & \text{si } \alpha < x < \beta \\ 0 & \text{sinon} \end{cases}$$

Et donc Y admet une espérance si et seulement si $\int_{\alpha}^{\beta} x f_Y(x) dx = \int_{\alpha}^{\beta} x (\varphi^{-1})'(x) f_X(\varphi^{-1}(x)) dx$ converge.

Notons que cette intégrale est encore égale à $\int_{\alpha}^{\beta} \varphi(\varphi^{-1}(x)) f_X(\varphi^{-1}(x)) (\varphi^{-1})'(x) dx$.

Détails

La fonction inverse étant décroissante sur \mathbf{R}_+^* , on a bien pris soin de changer le sens des inégalités.

⁹ Presque sûrement.

Détails

Puisque φ est croissante, il en est de même de sa bijection réciproque φ^{-1} et donc on a conservé le sens des inégalités.

Et donc par le théorème de changement de variable, qui s'applique bien puisque φ^{-1} est de classe \mathcal{C}^1 et strictement monotone, sous réserve de convergence

$$E(Y) = \int_{\alpha}^{\beta} \varphi(\varphi^{-1}(x)) (\varphi^{-1})'(x) f_X(\varphi^{-1}(x)) dx = \int_a^b \varphi(t) f_X(t) dt.$$

Et donc nous retrouvons bien l'intégrale qui apparaît dans le théorème de transfert.

Détails

Le changement de variable que l'on effectue est $t = \varphi^{-1}(x)$.

MÉTHODES DU CHAPITRE 8

✓ MÉTHODE 8.1 : Montrer qu'une fonction f est une densité de probabilité

On revient à la définition en prouvant que f est

- positive sur \mathbf{R} (généralement sans difficultés)
- continue sur \mathbf{R} sauf éventuellement en un nombre fini de points. On précisera alors quels sont les points où f pourrait ne pas être continue. On ne perdra pas de temps à déterminer si oui ou non f y est continue, cela n'est pas nécessaire.
- $\int_{-\infty}^{+\infty} f(t) dt$ converge et vaut 1.

Exercices de TD : 8.2, 8.5, 8.17, 8.21

✓ MÉTHODE 8.2 : Déterminer la fonction de répartition d'une variable aléatoire $Y = \varphi(X)$

Déterminer le support de Y peut permettre de déterminer sans calculs à quels endroits F_Y vaut 0 ou 1.

Pour $x \in Y(\Omega)$, on revient à la définition : $F_Y(x) = P(Y \leq x)$, puis on transformera cette inégalité pour l'exprimer en fonction de X , c'est-à-dire la mettre sous la forme $P(X \leq \star)$ ou $P(X \geq \star)$.

Lors des différentes étapes du calcul, on veillera bien au sens des inégalités, et à justifier pourquoi ce sens change ou non.

⚠ On demande rarement de «déterminer la fonction de répartition», mais plus souvent de «déterminer la loi», ce qui signifie la même chose.

Exercices de TD : 8.1, 8.3, 8.4, 8.5, 8.6, 8.9, 8.10, 8.21

✓ MÉTHODE 8.3 : Prouver qu'une variable aléatoire X est à densité

Si la fonction de répartition de X n'a pas été déterminée, on commencera par la déterminer.

Il faut ensuite prouver que cette fonction est

- continue sur \mathbf{R} tout entier. En général, seuls un ou deux points peuvent poser problème. En ces points, il faudra alors revenir à la définition de la continuité, avec des limites.
- de classe \mathcal{C}^1 , sauf éventuellement en un nombre fini de points. Là encore, on se contentera de préciser quels sont ces points, mais on ne poussera pas l'étude plus loin.

Exercices de TD : 8.3, 8.5, 8.9, 8.10

✓ MÉTHODE 8.4 : Déterminer la densité d'une variable aléatoire X

Il faudra bien entendu avoir déjà prouvé que la variable X est à densité (cf ci-dessus).

Une densité de X est alors toute fonction coïncidant avec F'_X sauf éventuellement en un nombre fini de points.

- Si F_X est dérivable sur \mathbf{R} tout entier, alors on peut prendre F'_X comme densité de X .
- Sinon, on ne dira pas que F'_X est une densité (puisque'elle n'est pas définie partout), et on prendra comme densité une fonction qui coïncide avec F'_X là où cette dernière existe, et n'importe quelle valeur (0 est le choix standard) en les points (nécessairement en nombre fini) où F'_X n'est pas définie.

Exercices de TD : 8.3, 8.5, 8.9, 8.10, 8.11

✓ MÉTHODE 8.5 : Utiliser le théorème de transfert pour calculer une espérance

Si on y pense généralement pour calculer $E(e^X)$ ou $E(\ln(X))$, on n'oubliera pas que le théorème de transfert permet également de calculer les moments d'une variable aléatoire en l'appliquant aux fonctions $x \mapsto x^r$.

Attention aux bornes : s'il est facile d'écrire

$$E(g(X)) = \int_{-\infty}^{+\infty} g(x)f_X(x) dx$$

on n'oubliera pas de changer les bornes de l'intégrale si la densité est nulle sur une partie de \mathbf{R} .

⚠ Le théorème de transfert ne peut que permettre de déterminer une espérance, jamais une loi !

Exercices de TD : 8.2, 8.3, 8.10

✓ **MÉTHODE 8.6 : Déterminer la densité d'une somme de variables aléatoires X et Y**

Dans tous les cas, les seuls résultats dont on dispose nécessitent l'**indépendance** de X et de Y , que l'on n'oubliera pas de vérifier.

- Si X et Y suivent toutes deux des lois γ ou toutes deux des lois normales, on pourra utiliser les résultats de stabilité du cours.
- Sinon, on fera appel au produit de convolution : une densité de $X + Y$ est

$$h(x) = \int_{-\infty}^{+\infty} f_X(t)f_Y(x-t) dt$$

Les bornes de l'intégrale à calculer sont susceptibles de changer suivant la valeur de x donc :

- on commencera par déterminer, à x fixé, à quelle condition **sur** t les densités $f_X(t)$ et $f_Y(x-t)$ sont non nulles
- on représentera sur un dessin les valeurs de t pour lesquelles $f_X(t)$ et $f_Y(x-t)$ sont non nulles
- après avoir éventuellement distingué plusieurs cas à l'aide du dessin, déterminer les bornes des intégrales qu'il faut calculer, puis calculer ces intégrales.

Exercices de TD : 8.12, 8.13, 8.14, 8.15, 8.17, 8.18, 8.19, 8.21

EXERCICES DU CHAPITRE 8

EXERCICE 8.1 Soit U une variable aléatoire suivant la loi uniforme sur $[0, 1]$, et soit $\lambda > 0$. On pose $X = -\frac{1}{\lambda} \ln(1 - U)$. Déterminer la loi de X . F

EXERCICE 8.2 Soit X une variable aléatoire suivant la loi $\gamma(\nu)$, et soit $Y = \frac{1}{X}$. Déterminer à quelle(s) condition(s) sur ν la variable Y admet une variance et le cas échéant, déterminer $V(Y)$. PD

EXERCICE 8.3 En utilisant des lois normales bien choisies, calculer les valeurs des intégrales suivantes : PD

$$I = \int_{-\infty}^{+\infty} e^{-\frac{1}{8}(t^2+2t+1)} dt, J = \int_{-\infty}^{+\infty} t e^{-\frac{1}{8}(t^2+2t+1)} dt, K = \int_{-\infty}^{+\infty} e^{-2t^2+12t-16} dt, L = \int_{-\infty}^{+\infty} t^2 e^{-2t^2+12t-16} dt.$$

EXERCICE 8.4 (D'après ECRICOME 2015) PD

1) Déterminer un réel a tel que la fonction f définie par : $f(x) = \begin{cases} 0 & \text{si } x < 2 \\ \frac{a}{x \ln^2(x)} & \text{si } x \geq 2 \end{cases}$ soit une densité.

Dans la suite, X est une variable aléatoire admettant f comme densité.

- 2) Déterminer la fonction de répartition F de X .
- 3) La variable aléatoire X admet-elle une espérance ?

EXERCICE 8.5 Soit X une variable aléatoire qui suit la loi uniforme $\mathcal{U}([-1, 1])$. Montrer que $Y = X^2$ est une variable à densité et en déterminer une densité. PD
Mêmes questions si $X \hookrightarrow \mathcal{E}(\lambda)$.

EXERCICE 8.6 Soit X une variable aléatoire qui suit la loi exponentielle de paramètre λ . Montrer que X admet des moments de tout ordre et les calculer. PD

EXERCICE 8.7 Loi logistique PD

Soit F la fonction définie sur \mathbf{R} par $F(x) = \frac{1}{1 + e^{-x}}$.

- 1) Montrer que F est la fonction de répartition d'une variable à densité X . Déterminer une densité de X
- 2) Soient X et Y deux variables aléatoires indépendantes possédant F comme fonction de répartition. On pose $Z = \max(X, Y)$
 - a) Montrer que Z est une variable à densité.
 - b) Déterminer une densité de Z .

EXERCICE 8.8 Loi de Pareto PD

Soit $a > 0$ et soit $f_a : t \mapsto \begin{cases} at^{-a-1} & \text{si } t \in [1; +\infty[\\ 0 & \text{sinon} \end{cases}$.

- 1) Montrer que f_a est une densité.
- 2) Soit X une variable aléatoire admettant f_a pour densité.
 - a) Déterminer la fonction de répartition F_X de X .
 - b) X possède-t-elle une espérance ? Si oui, la calculer. Même question pour la variance.
 - c) On pose $Y = \ln X$. Déterminer la fonction de répartition de Y , puis reconnaître la loi de Y

EXERCICE 8.9 Soit $n \in \mathbf{N}^*$ et soit X une variable aléatoire qui suit la loi uniforme sur $[0, n]$. Déterminer la loi de $\lfloor X \rfloor$. F

EXERCICE 8.10 Produit de deux uniformes PD

Soient X et Y deux variables aléatoires sur (Ω, \mathcal{A}, P) , indépendantes et suivant toutes deux la loi uniforme $\mathcal{U}([0, 1])$.

- 1) Déterminer la loi de $Z = -\ln X$.
- 2) On admettra (temporairement) que Z et $T = -\ln Y$ sont indépendantes. Déterminer la loi de $Z + T$, puis la fonction de répartition de XY et enfin une densité de XY .

EXERCICE 8.11 Soient X et Y deux variables aléatoires indépendantes suivant la loi normale $\mathcal{N}(-1, 2)$. Déterminer la loi de $3X - 2Y$. F

EXERCICE 8.12 Soient X et Y deux variables aléatoires indépendantes telles que $X \hookrightarrow \mathcal{E}(2)$ et $Y \hookrightarrow \mathcal{E}\left(\frac{1}{3}\right)$. Déterminer sans calculs la loi de $2X + \frac{Y}{3}$. F

EXERCICE 8.13 Soit X une variable aléatoire suivant la loi $\mathcal{N}(0, 1)$, et soit $n \in \mathbf{N}^*$. Quelle est la valeur de $a \in \mathbf{R}_+$ maximisant $P(a \leq X < na)$? AD

EXERCICE 8.14 Minimum de deux lois exponentielles PD

Soient X et Y deux variables aléatoires indépendantes sur (Ω, \mathcal{A}, P) telles que $X \hookrightarrow \mathcal{E}(\lambda)$ et $Y \hookrightarrow \mathcal{E}(\mu)$. Déterminer la loi de $Z = \min(X, Y)$.

EXERCICE 8.15 Somme de deux lois exponentielles PD

Soient X et Y deux variables aléatoires indépendantes suivant la même loi exponentielle $\mathcal{E}(\lambda)$. On pose $Z = X + Y$.

- 1) Déterminer la loi de λX .
- 2) En déduire la loi de λZ , puis une densité de Z .

EXERCICE 8.16 Soit $a > 0$. PD

- 1) Déterminer la valeur de la constante λ pour laquelle $f : t \mapsto \begin{cases} \lambda t e^{-at} & \text{si } t \geq 0 \\ 0 & \text{sinon} \end{cases}$ est la densité d'une variable aléatoire X .
- 2) Déterminer la fonction de répartition de X .
- 3) La variable X admet-elle une espérance ? Si oui, la calculer.

EXERCICE 8.17 Soit Φ la fonction de répartition d'une variable aléatoire suivant la loi normale centrée réduite. Étudier l'existence, et le cas échéant calculer la valeur de AD

$$\int_0^{+\infty} (1 - \Phi(t)) dt.$$

EXERCICE 8.18 (D'après ECRICOME 2001) AD

Soient X et Y deux variables aléatoires à densité définies sur le même espace probabilisé (Ω, \mathcal{A}, P) , indépendantes, et telles que $X \hookrightarrow \mathcal{E}(a)$ et $Y \hookrightarrow \mathcal{E}(b)$, où a et b sont deux réels strictement positifs.

- 1) Déterminer une densité de la variable aléatoire $-X$.

- 2) Montrer que $Y - X$ admet pour densité la fonction h définie par $h(t) = \begin{cases} \frac{ab}{a+b} e^{-bt} & \text{si } t > 0 \\ \frac{ab}{a+b} e^{at} & \text{si } t \leq 0 \end{cases}$

EXERCICE 8.19 Différence de deux lois uniformes : la loi de Xenakis (D'après EDHEC 2010) AD

Soit $a > 0$. On considère deux variables aléatoires X et Y , définies sur un espace probabilisé (Ω, \mathcal{A}, P) , indépendantes, et suivant toutes deux la loi uniforme sur $[0, a[$.

On pose alors $Z = |X - Y|$.

- 1) a) Déterminer une densité de $-Y$.
- b) En déduire que la variable aléatoire $X - Y$ admet pour densité la fonction g définie par

$$g(x) = \begin{cases} \frac{a - |x|}{a^2} & \text{si } x \in [-a, a] \\ 0 & \text{sinon} \end{cases}$$

On note G la fonction de répartition de $X - Y$.

- 2) a) Exprimer la fonction de répartition H de la variable aléatoire Z en fonction de G .
- b) En déduire qu'une densité de Z est la fonction h définie par

$$h(x) = \begin{cases} \frac{2(a-x)}{a^2} & \text{si } x \in [0, a] \\ 0 & \text{sinon} \end{cases}$$

EXERCICE 8.20 Soit X une variable aléatoire qui suit la loi normale $\mathcal{N}(0, 4)$. On pose $M = \begin{pmatrix} 2X & 1 \\ -4 & X \end{pmatrix}$. Déterminer la probabilité que M admette deux valeurs propres distinctes. AD

EXERCICE 8.21 Loi exponentielle bilatérale ou loi de Laplace

AD

Soit $\lambda > 0$, et soit $f : x \mapsto \frac{\lambda}{2} e^{-\lambda|x|}$.

- 1) Montrer que f est une densité de probabilité.
- 2) Soient X et Y deux variables aléatoires indépendantes suivant la loi $\mathcal{E}(\lambda)$. Montrer que $Z = X - Y$ admet f pour densité. En déduire la loi de $|Z|$.
- 3) Soient X_1 et X_2 deux variables aléatoires indépendantes admettant f pour densité. Déterminer une densité de $X_1 + X_2$.

EXERCICE 8.22 (QSP ESCP 2013)

D

Soient X et Y deux variables aléatoires indépendantes sur (Ω, \mathcal{A}, P) telles que $X \hookrightarrow \mathcal{N}(\mu, \sigma^2)$ et $Y \hookrightarrow \mathcal{N}(\nu, \sigma^2)$. Déterminer une condition nécessaire et suffisante sur μ et ν pour que $P(X \leq Y) \geq \frac{1}{2}$.

EXERCICE 8.23 (QSP HEC 2009)

D

Soit U une variable aléatoire définie sur un espace probabilisé (Ω, \mathcal{A}, P) , de loi uniforme sur $]0, 1]$, et soit $q \in]0, 1]$. Déterminer la loi de la variable aléatoire $X = 1 + \left\lfloor \frac{\ln U}{\ln q} \right\rfloor$, où $\lfloor x \rfloor$ désigne la partie entière de x .

EXERCICE 8.24 Distance entre deux points aléatoires (Oral HEC 2015)

TD

- 1) Soient a, b, α trois réels strictement positifs vérifiant $0 < \alpha < a^2 \leq b^2$.
 - a) Établir la convergence de l'intégrale $\int_0^\alpha \left(\frac{a}{\sqrt{t}} - 1 \right) \left(\frac{b}{\sqrt{\alpha-t}} - 1 \right) dt$. Cette intégrale est notée $I_{a,b}(\alpha)$.
 - b) Calculer $I_{a,b}(\alpha)$ à l'aide du changement de variable $t = \alpha \cos^2 u$.

Dans le plan rapporté à un repère orthonormé, on place deux points M et N tels que leurs abscisses respectives X_M et X_N suivent la loi uniforme sur $]0, a[$ et leurs ordonnées Y_M et Y_N suivent la loi uniforme sur $]0, b[$.

On suppose que les quatre variables aléatoires X_M, X_N, Y_M et Y_N sont définies sur un espace probabilisé (Ω, \mathcal{A}, P) et sont indépendantes.

On note D la variable aléatoire égale à la longueur du segment $[M, N]$: $D^2 = (X_M - X_N)^2 + (Y_M - Y_N)^2$.

- 2)
 - a) Quelle est la loi suivie par $-X_M$?
 - b) On pose : $Z_a = (X_N - X_M)$ et $Z_b = (Y_N - Y_M)$. Déterminer les lois de probabilité de Z_a et Z_b respectivement.
 - c) Montrer qu'une densité $f_{Z_a^2}$ de Z_a^2 est donnée par $f_{Z_a^2}(x) = \begin{cases} \frac{1}{a^2} \left(\frac{a}{\sqrt{x}} - 1 \right) & \text{si } 0 < x < a^2 \\ 0 & \text{sinon} \end{cases}$

- 3) Soit $\theta < a$. Calculer $P(D \leq \theta)$.

CORRECTION DES EXERCICES DU CHAPITRE 8

SOLUTION DE L'EXERCICE 8.1

Puisque U est à valeurs dans $[0, 1[$ X est à valeurs dans \mathbf{R}_+ . Et donc pour $x < 0$, $F_X(x) = 0$. Soit $x \geq 0$. Alors

$$F_X(x) = P\left(-\frac{1}{\lambda} \ln(1 - U) \leq x\right) = P(\ln(1 - U) \geq -\lambda x).$$

Par croissance de la fonction exponentielle, on a alors

$$F_X(x) = P(1 - U \geq e^{-\lambda x}) = P(U \leq 1 - e^{-\lambda x}) = F_U(1 - e^{-\lambda x}).$$

Mais $1 - e^{-\lambda x} \in [0, 1]$, et nous connaissons la fonction de répartition de U qui est donnée par

$$F_U(x) = \begin{cases} 0 & \text{si } x < 0 \\ x & \text{si } 0 \leq x \leq 1 \\ 1 & \text{si } x > 1 \end{cases}$$

On a donc $F_X(x) = 1 - e^{-\lambda x}$. Donc la fonction de répartition de X est

$$F_X(x) = \begin{cases} 0 & \text{si } x < 0 \\ 1 - e^{-\lambda x} & \text{si } x \geq 0 \end{cases}$$

Nous reconnaissons là la fonction de répartition d'une variable aléatoire suivant la loi exponentielle $\mathcal{E}(\lambda)$ et donc $X \hookrightarrow \mathcal{E}(\lambda)$.

SOLUTION DE L'EXERCICE 8.2

D'après le théorème de transfert, sous réserve de convergence, on a

$$E(Y) = E\left(\frac{1}{X}\right) = \int_0^{+\infty} \frac{1}{t} \frac{1}{\Gamma(v)} t^{v-1} e^{-t} dt = \frac{1}{\Gamma(v)} \int_0^{+\infty} t^{v-2} e^{-t} dt.$$

Cette dernière intégrale converge si et seulement si $v - 1 > 0 \Leftrightarrow v > 1$, et dans ce cas,

$$E(Y) = \frac{1}{\Gamma(v)} \int_0^{+\infty} t^{v-2} e^{-t} dt = \frac{1}{\Gamma(v)} \Gamma(v-1) = \frac{1}{v-1}.$$

De même, toujours par le théorème de transfert, sous réserve de convergence,

$$E(Y^2) = E\left(\frac{1}{X^2}\right) = \int_0^{+\infty} \frac{1}{t^2} \frac{1}{\Gamma(v)} t^{v-1} e^{-t} dt = \frac{1}{\Gamma(v)} \int_0^{+\infty} t^{v-3} e^{-t} dt.$$

Cette intégrale converge si et seulement si $v - 2 > 0 \Leftrightarrow v > 2$.

$$\text{Et alors } E(Y^2) = \frac{1}{\Gamma(v-2)} \Gamma(v) = \frac{1}{(v-1)(v-2)}.$$

On en déduit donc que Y admet une variance si et seulement si $v > 3$, et qu'alors

$$V(Y) = E(Y^2) - E(Y)^2 = \frac{1}{(v-1)(v-2)} - \frac{1}{(v-1)^2} = \frac{1}{(v-1)^2(v-2)}.$$

SOLUTION DE L'EXERCICE 8.3

1. Comme indiqué, il s'agit de faire apparaître la densité d'une loi normale¹. Pour cela, nous aimerions, si possible trouver μ et σ tels que

$$e^{-\frac{1}{8}(t^2+2t+1)} = e^{-\frac{1}{2}\left(\frac{t-\mu}{\sigma}\right)^2} \Leftrightarrow \frac{1}{4}(t^2+2t+1) = \left(\frac{t-\mu}{\sigma}\right)^2 = \frac{1}{\sigma^2}(t-\mu)^2.$$

Mais $t^2 + 2t + 1 = (t+1)^2$, et donc $\mu = -1$ et $\sigma = 2$ conviennent.

Soit donc X une variable aléatoire suivant la loi $\mathcal{N}(-1, 2^2)$.

Alors une densité de X est $f_X : t \mapsto \frac{1}{2\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{t+1}{2}\right)^2} = \frac{1}{2\sqrt{2\pi}} e^{-\frac{1}{8}(t^2+2t+1)}$.

Or nous savons que

$$\int_{-\infty}^{+\infty} f_X(t) dt = 1 \Leftrightarrow \int_{-\infty}^{+\infty} e^{-\frac{1}{8}(t^2+2t+1)} dt = 2\sqrt{2\pi}.$$

Support

En fait, U peut également prendre la valeur 1, mais ceci n'arrivant qu'avec probabilité 0, on se permet de le négliger.

¹ Toute la difficulté étant de trouver quelle loi normale.

Remarque

Notons que ceci prouve au passage que $\int_{-\infty}^{+\infty} e^{-\frac{1}{8}(t^2+2t+1)} dt$ converge.

2. La présence du t devant l'exponentielle nous fait penser à l'espérance d'une loi normale. Et en effet, si $X \hookrightarrow \mathcal{N}(-1, 2)$, alors

$$E(X) = -1 \Leftrightarrow \int_{-\infty}^{+\infty} \frac{1}{2\sqrt{\pi}} t e^{-\frac{1}{8}(t^2+2t+1)} dt = -1 \Leftrightarrow J = -2\sqrt{2\pi}.$$

3. Là encore, nous allons essayer de faire apparaître la densité d'une loi normale. On a

$$-2t^2 + 12t - 16 = -\frac{1}{2}(4t^2 - 24t + 32).$$

Il s'agit alors de mettre sous forme canonique le polynôme $4t^2 - 24t + 32$. On a

$$4t^2 - 24t + 32 = 4(t^2 - 6t + 8) = 4((t-3)^2 - 9 + 8) = 4((t-3)^2 - 1).$$

Et donc il vient, sous réserve de convergence,

$$K = \int_{-\infty}^{+\infty} e^{-\frac{1}{2}4(t-3)^2} e^2 dt = e \int_{-\infty}^{+\infty} e^{-\frac{1}{2}4(t-3)^2} dt.$$

Or, si X est une variable aléatoire suivant la loi normale $\mathcal{N}\left(3, \frac{1}{4}\right)$, une densité de X est

$$f_X : t \mapsto \frac{2}{\sqrt{2\pi}} e^{-\frac{1}{2}4(t-3)^2}.$$

En particulier,

$$\int_{-\infty}^{+\infty} f_X(t) dt = 1 \Leftrightarrow \int_{-\infty}^{+\infty} e^{-\frac{1}{2}4(t-3)^2} dt = \frac{\sqrt{2\pi}}{2} \Leftrightarrow K = e^2 \frac{\sqrt{2\pi}}{2}.$$

4. L'exponentielle n'ayant pas changé par rapport à la question précédente, nous allons considérer la même loi normale, mais cette fois le facteur t^2 devant l'exponentielle nous laisse penser à un moment d'ordre 2.

Si $X \hookrightarrow \mathcal{N}\left(3, \frac{1}{4}\right)$, alors par la formule de Huygens,

$$E(X^2) = V(X) + E(X)^2 = \frac{1}{4} + 9 = \frac{37}{4}.$$

Soit encore

$$\int_{-\infty}^{+\infty} \frac{2}{\sqrt{2\pi}} t^2 e^{-\frac{1}{2}4(t-3)^2} dt = \frac{37}{4} \Leftrightarrow \int_{-\infty}^{+\infty} t^2 e^{-\frac{1}{2}4(t-3)^2} dt = \frac{37\sqrt{2\pi}}{8}.$$

Et donc il vient

$$L = e^2 \int_{-\infty}^{+\infty} t^2 e^{-\frac{1}{2}4(t-3)^2} dt = e^2 \frac{37\sqrt{2\pi}}{8}.$$

SOLUTION DE L'EXERCICE 8.4

1. Pour $a \geq 0$, la fonction f est positive et continue sauf en $x = 2$. De plus, pour $A > 2$, on a

$$\int_{-\infty}^A f(t) dt = \int_2^A \frac{a}{t \ln^2(t)} dt = a \left[-\frac{1}{\ln t} \right]_2^A = a \left(\frac{1}{\ln 2} - \frac{1}{\ln(A)} \right) \xrightarrow{A \rightarrow +\infty} \frac{a}{\ln(2)}.$$

Donc $\int_{-\infty}^{+\infty} f(t) dt = 1 \Leftrightarrow a = \ln(2)$.

On en déduit que pour $a = \ln(2)$ (qui est bien positif), f est une densité de probabilité.

2. Pour $x \leq 2$, on a

$$F(x) = P(X \leq x) = \int_{-\infty}^x f(t) dt = \int_{-\infty}^x 0 dt = 0.$$

Méthode

Il s'agit de la mise sous forme canonique vue au lycée.

On commence par factoriser par le coefficient dominant, puis une fois que l'on a $t^2 + bt + c$, il s'agit de faire apparaître un carré contenant tous les termes en t (en l'occurrence, $\left(t + \frac{b}{2}\right)^2$), et enfin de s'occuper des constantes.

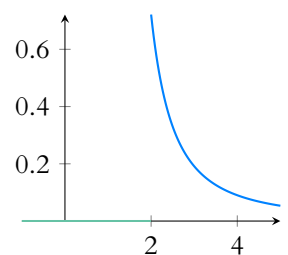


FIGURE 8.1– La densité f .

Et pour $x > 2$, alors

$$F(x) = \int_2^x \frac{\ln 2}{t \ln^2 t} dt = \ln 2 \left(\frac{1}{\ln 2} - \frac{1}{\ln x} \right) = 1 - \frac{\ln 2}{\ln x}.$$

Ainsi, la fonction de répartition F de X est donnée par

$$F(x) = \begin{cases} 0 & \text{si } x \leq 2 \\ 1 - \frac{\ln 2}{\ln x} & \text{si } x > 2 \end{cases}$$

3. X admet une espérance si et seulement si $\int_2^{+\infty} \frac{1}{\ln^2 t} dt$ converge.

Or, au voisinage de $+\infty$, on a

$$\frac{\frac{1}{t}}{\frac{1}{\ln^2 t}} = \frac{\ln^2 t}{t} \xrightarrow{t \rightarrow +\infty} 0$$

de sorte que $\frac{1}{t} = o\left(\frac{1}{\ln^2 t}\right)$.

Ainsi, si $\int_2^{+\infty} \frac{1}{\ln^2 t} dt$ converge, il en est de même de $\int_2^{+\infty} \frac{dt}{t}$, ce qui est absurde.

Par conséquent, X n'admet pas d'espérance.

SOLUTION DE L'EXERCICE 8.5

Notons que Y prend ses valeurs dans $[0, 1]$ et donc $F_Y(x) = 0$ si $x < 0$ et $F_X(x) = 1$ si $x > 1$.

Pour $x \in [0, 1]$, on a $[Y \leq x] = [X^2 \leq x] = [-\sqrt{x} \leq X \leq \sqrt{x}]$.

Et donc $F_Y(x) = F_X(\sqrt{x}) - F_X(-\sqrt{x})$.

Mais nous savons que $F_X(t) = \begin{cases} 0 & \text{si } t \leq -1 \\ \frac{t+1}{2} & \text{si } -1 \leq t \leq 1 \\ 1 & \text{si } t \geq 1 \end{cases}$.

Et donc $F_Y(x) = \frac{\sqrt{x} - (-\sqrt{x})}{2} = \sqrt{x}$.

Ainsi, la fonction de répartition de Y est donnée par

$$F_Y(x) = \begin{cases} 0 & \text{si } x < 0 \\ \sqrt{x} & \text{si } 0 \leq x \leq 1 \\ 1 & \text{si } x \geq 1 \end{cases}$$

Cette fonction est continue sur $]-\infty, 0[$ et sur $]1, +\infty[$ car elle y est constante, et elle est continue sur $[0, 1]$ car la fonction racine carrée l'est.

De plus, elle est continue en 0 car

$$\lim_{x \rightarrow 0^-} F_Y(x) = 0 = \lim_{x \rightarrow 0^+} F_Y(x) = F_Y(0).$$

De même, F_Y est continue en 1 car

$$\lim_{x \rightarrow 1^-} F_Y(x) = 1 = \lim_{x \rightarrow 1^+} F_Y(x) = F_Y(1).$$

Et donc F_Y est continue sur \mathbf{R} . Elle est \mathcal{C}^1 sauf éventuellement en 0 et en 1, et donc Y est une variable à densité.

On peut alors prendre pour densité de Y toute fonction qui coïncide avec F'_Y sur $\mathbf{R} \setminus \{0, 1\}$, par exemple

$$f_Y : x \mapsto \begin{cases} \frac{1}{2\sqrt{x}} & \text{si } 0 < x < 1 \\ 0 & \text{sinon} \end{cases}$$

Dans le cas où $X \leftrightarrow \mathcal{E}(\lambda)$, on a toujours pour $x \geq 0$,

$$F_Y(x) = P(X^2 \leq x) = P(-\sqrt{x} \leq X \leq \sqrt{x}) = F_X(\sqrt{x}) - F_X(-\sqrt{x}).$$

La différence est que cette fois, X ne prend que des valeurs positives, donc $F_X(-\sqrt{x}) = 0$.

Et donc

$$F_X(x) = F_X(\sqrt{x}) = 1 - e^{-\lambda\sqrt{x}}.$$

Méthode

Commencer par systématiquement s'interroger sur le support d'une variable aléatoire permet de restreindre le calcul de $F_Y(x)$ aux valeurs de x pour lesquelles le résultat est non trivial. Ceci évite notamment de pénibles distinctions de cas dans les calculs.

Astuce

En fait, lorsqu'on dit que F_Y est continue sur $[0, 1]$, cela prouve qu'elle est continue à droite en 0, et donc il suffit de vérifier qu'elle est continue à gauche.

Donc la fonction de répartition de Y est $x \mapsto \begin{cases} 0 & \text{si } x < 0 \\ 1 - e^{-\lambda\sqrt{x}} & \text{si } x \geq 0 \end{cases}$

Il est clair que F_Y est \mathcal{C}^1 sauf peut-être en 0. De plus

$$\lim_{x \rightarrow 0^-} F_Y(x) = 0 = \lim_{x \rightarrow 0^+} F_Y(x)$$

donc F_Y est continue en 0 et donc sur \mathbf{R} .

Ainsi, F_Y est continue sur \mathbf{R} et \mathcal{C}^1 sauf peut-être en 0 : Y est une variable à densité.

Une densité en est alors toute fonction qui coïncide avec F'_Y sauf en un nombre fini de points, par exemple on peut prendre

$$f_Y : x \mapsto \begin{cases} 0 & \text{si } x \leq 0 \\ \frac{\lambda}{2\sqrt{x}} e^{-\lambda\sqrt{x}} & \text{si } x > 0 \end{cases}$$

SOLUTION DE L'EXERCICE 8.6

Soit $k \in \mathbf{N}^*$. D'après le théorème de transfert, X admet un moment d'ordre k si et seulement

si $\int_0^{+\infty} \lambda t^k e^{-\lambda t} dt$ converge².

Or, le changement de variable³ $x = \lambda t$ nous dit que cette intégrale converge si et seulement

si $\int_0^{+\infty} \lambda \left(\frac{x}{\lambda}\right)^k e^{-x} \frac{dx}{\lambda}$ converge, et en cas de convergence elles sont égales.

Or, cette dernière intégrale n'est autre que $\frac{1}{\lambda^k} \Gamma(k+1)$, qui converge et vaut $\frac{k!}{\lambda^k}$.

On en déduit que X admet un moment d'ordre k et que

$$E(X^k) = \int_0^{+\infty} \lambda t^k e^{-\lambda t} t = \frac{k!}{\lambda^k}.$$

SOLUTION DE L'EXERCICE 8.7

1. La fonction F est clairement de classe \mathcal{C}^1 sur \mathbf{R} , et elle y est croissante.

De plus, on a $\lim_{x \rightarrow -\infty} F(x) = 0$ et $\lim_{x \rightarrow +\infty} F(x) = 1$.

Il s'agit donc bien de la fonction de répartition d'une variable aléatoire.

Et F étant \mathcal{C}^1 sur \mathbf{R} , elle y est continue, et bien entendu \mathcal{C}^1 sauf en un nombre fini de points : c'est la fonction de répartition d'une variable à densité X .

Et puisque F est \mathcal{C}^1 sur \mathbf{R} tout entier, une densité de X est $F' : x \mapsto \frac{e^{-x}}{(1+e^{-x})^2}$.

2.a. Pour tout $x \in \mathbf{R}$, on a $[Z \leq x] = [X \leq x] \cap [Y \leq x]$.

Puisque X et Y sont indépendantes, on a donc

$$F_Z(x) = P(Z \leq x) = P(X \leq x)P(Y \leq x) = F(x)^2 = \frac{1}{(1+e^{-x})^2}.$$

Cette fonction est évidemment \mathcal{C}^1 sur \mathbf{R} , et donc⁴ Z est une variable à densité.

2.b. Une densité de Z est alors

$$F'_Z : x \mapsto \frac{2e^{-x}}{(1+e^{-x})^2}.$$

SOLUTION DE L'EXERCICE 8.8

1. La fonction f_a est continue sur \mathbf{R} sauf en a , et elle y est clairement positive car $a > 0$.

De plus, puisque $a > 0$, $a+1 > 1$ et donc $\int_1^{+\infty} at^{-a-1} dt$ est convergente car il s'agit d'une intégrale de Riemann convergente.

Enfin, pour $x > 1$, on a

$$\int_{-\infty}^x f_a(t) dt = \int_1^x f_a(t) dt = \int_1^x \frac{a}{t^{a+1}} dt = \left[\frac{1}{t^a} \right]_1^x = 1 - \frac{1}{x^a}.$$

En passant à la limite lorsque $x \rightarrow \infty$, on a

$$\int_{-\infty}^{+\infty} f_a(t) dt = \int_1^{+\infty} f_a(t) dt = 1.$$

Ainsi, f_a est bien une densité.

² Absolument, mais il s'agit d'une fonction positive.

³ Affine.

Méthode

Attention à la formulation des questions : si on dispose (comme à l'exercice 2) d'une variable aléatoire et que l'on souhaite prouver qu'elle est à densité, il suffit de prouver que sa fonction de répartition est continue et \mathcal{C}^1 sauf éventuellement en un nombre fini de points.

Mais ici, nous ne savons pas que F est une fonction de répartition, il s'agit donc de prouver que c'est le cas en étudiant les limites, la croissance, et en prouvant qu'elle est continue à droite en tout point (ce qui est facilité par le fait qu'ici, elle est continue).

⁴ Puisqu'elle est \mathcal{C}^1 , elle est en particulier continue.

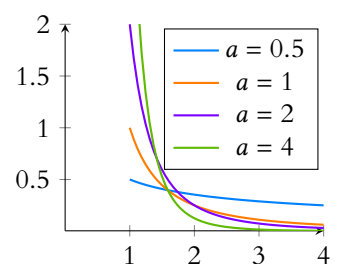


FIGURE 8.2– La densité f_a .

M. VIENNEY

2.a. Le calcul réalisé précédemment prouve que

$$\forall x \in \mathbf{R}, F_X(x) = \begin{cases} 0 & \text{si } x \leq 1 \\ 1 - \frac{1}{x^a} & \text{si } x > 1 \end{cases}$$

2.b. Par définition, X possède une espérance si et seulement si $\int_{-\infty}^{+\infty} t f_a(t) dt$ converge.

Or, $t \mapsto t f_a(t)$ est nulle sur $] -\infty, 1[$, donc il suffit de vérifier que $\int_1^{+\infty} \frac{a}{t^a} dt$ converge.

C'est le cas si et seulement si $a > 1$, et alors

$$E(X) = \int_{-\infty}^{+\infty} t f_a(t) dt = \int_1^{+\infty} \frac{a}{t^a} dt = \frac{a}{a-1}.$$

De même, on montre que X possède un moment d'ordre 2 (et donc une variance) si et seulement si $a > 2$, et qu'alors $E(X^2) = \frac{a}{a-2}$. Par la formule de Huygens, on a alors

$$V(X) = E(X^2) - E(X)^2 = \frac{a}{(a-2)} - \frac{a^2}{(a-1)^2} = \frac{a(a-1)^2 - a^2(a-2)}{(a-1)^2(a-2)} = \frac{a}{(a-1)^2(a-2)}.$$

2.c. On a, pour $y \in \mathbf{R}$,

$$\begin{aligned} F_Y(y) &= P(Y \leq y) = P(\ln X \leq y) = P(X \leq e^y) = F_X(e^y) \\ &= \begin{cases} 0 & \text{si } e^y \leq 1 \\ 1 - \frac{1}{e^{ay}} & \text{si } e^y > 1 \end{cases} = \begin{cases} 0 & \text{si } y \leq 0 \\ 1 - e^{-ay} & \text{si } y > 0 \end{cases} \end{aligned}$$

La fonction exp est croissante, donc on préserve le sens des inégalités.

On reconnaît là la fonction de répartition d'une loi exponentielle $\mathcal{E}(a)$, donc $Y \hookrightarrow \mathcal{E}(a)$.

SOLUTION DE L'EXERCICE 8.9

Notons que $\lfloor X \rfloor$ ne prend que des valeurs entières, entre 0 et n .

De plus, on a $P(\lfloor X \rfloor = n) = P(n \leq X < n+1) = P(X = n) = 0$.

Pour $k \in \llbracket 0, n-1 \rrbracket$, on a

$$P(\lfloor X \rfloor = k) = P(k \leq X < k+1) = \frac{(k+1) - k}{n} = \frac{1}{n}.$$

Ainsi, $\lfloor X \rfloor$ suit la loi uniforme sur $\llbracket 0, n-1 \rrbracket$.

SOLUTION DE L'EXERCICE 8.10

1. Puisque X prend ses valeurs dans $]0, 1]$, $\ln X$ prend ses valeurs dans $] -\infty, 0]$, et donc Z prend ses valeurs dans \mathbf{R}_+^* .

Donc pour $x \leq 0$, $F_Z(x) = 0$.

Soit $x > 0$. Alors

$$F_Z(x) = P(Z \leq x) = P(-\ln X \leq x) = P(\ln X \geq -x) = P(X \geq e^{-x}) = 1 - e^{-x}.$$

La fonction de répartition de Z est alors

$$F_Z : x \mapsto \begin{cases} 0 & \text{si } x \leq 0 \\ 1 - e^{-x} & \text{si } x > 0 \end{cases}$$

On reconnaît là la fonction de répartition d'une loi exponentielle de paramètre 1, et donc $Z \hookrightarrow \mathcal{E}(1)$.

2. On prouve de même que $T \hookrightarrow \mathcal{E}(1)$, et donc, par indépendance de Z et T , $Z + T \hookrightarrow \gamma(2)$. La variable aléatoire XY ne prend que des valeurs dans $[0, 1]$, donc

$$F_{XY}(x) = 0 \text{ si } x < 0 \text{ et } F_{XY}(x) = 1 \text{ si } x > 1.$$

Pour $x \in [0, 1]$, on a, par croissance du logarithme,

$$F_{XY}(x) = P(XY \leq x) = P(\ln X + \ln Y \leq \ln x) = P(Z + T \geq -\ln x) = \int_{-\ln x}^{+\infty} t e^{-t} dt.$$

La fonction exp est croissante, donc on préserve le sens des inégalités.

Densité

On pourrait bien entendu dériver cette fonction de répartition afin de retrouver la densité d'une loi exponentielle, mais rappelons que la fonction de répartition suffit à caractériser la loi.

Support

En effet, X et Y sont toutes deux à support dans $[0, 1]$, et donc il en est de même de XY .

Pour calculer cette intégrale, procédons à une intégration par parties sur un segment $[-\ln x, A]$. Alors

$$\int_{-\ln x}^A te^{-t} dt = [-te^{-t}]_{-\ln x}^A + \int_{-\ln x}^A e^{-t} dt = Ae^{-A} - x \ln x + x - e^{-A} \xrightarrow{A \rightarrow +\infty} x - x \ln x$$

Par conséquent, on a

$$F_{XY}(x) = \begin{cases} 0 & \text{si } x \leq 0 \\ x - x \ln(x) & \text{si } 0 < x \leq 1 \\ 1 & \text{si } x > 1 \end{cases}$$

On en déduit que F_{XY} est continue sur $]-\infty, 0]$, sur $]0, 1]$ et sur $]1, +\infty[$. De plus, elle vérifie

$$\lim_{x \rightarrow 0^+} F_{XY}(x) = 0 = F_{XY}(0) \text{ et } \lim_{x \rightarrow 1^-} F_{XY}(x) = 1 = F_{XY}(1)$$

de sorte que F_{XY} est continue en 0 et en 1 et donc sur \mathbf{R} .

De plus, elle est dérivable, de dérivée nulle en dehors de $[0, 1]$ et de dérivée égale à $-\ln x$ sur $]0, 1[$.

F_{XY} est donc de classe \mathcal{C}^1 sur \mathbf{R} privé de 0 et de 1.

Ainsi, XY est une variable à densité, et toute fonction égale à sa dérivée sur $\mathbf{R} - \{0, 1\}$ en est une densité. Par exemple, on peut prendre

$$x \mapsto \begin{cases} -\ln x & \text{si } 0 < x < 1 \\ 0 & \text{sinon} \end{cases}.$$

SOLUTION DE L'EXERCICE 8.11

Par transformation affine de lois normales, $3X$ suit la loi normale $\mathcal{N}(-3, 18)$ et $-2Y$ suit la loi $\mathcal{N}(2, 8)$.

Par somme de lois normales indépendantes, on en déduit que $3X - 2Y$ suit la loi $\mathcal{N}(-1, 26)$.

SOLUTION DE L'EXERCICE 8.12

Puisque $X \leftrightarrow \mathcal{E}(2)$, alors $2X \leftrightarrow \mathcal{E}(1) = \gamma(1)$. De même, $\frac{Y}{3} \leftrightarrow \gamma(1)$.

Or $2X$ et $\frac{Y}{3}$ sont indépendantes, de sorte que $2X + \frac{Y}{3} \leftrightarrow \gamma(2)$.

SOLUTION DE L'EXERCICE 8.13

Pour $a \geq 0$, posons $g(a) = P(a \leq X < na) = \Phi(na) - \Phi(a)$.

Or, la fonction Φ est dérivable sur \mathbf{R} , avec $\Phi'(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$.

Donc g est dérivable, de dérivée

$$g'(a) = \frac{n}{\sqrt{2\pi}} e^{-n^2 a^2/2} - \frac{1}{\sqrt{2\pi}} e^{-a^2/2} = \frac{1}{\sqrt{2\pi}} (ne^{-n^2 a^2/2} - e^{-a^2/2}) = \frac{1}{\sqrt{2\pi}} e^{-a^2/2} (ne^{-(n^2-1)a^2/2} - 1).$$

Cette fonction s'annule pour $ne^{-(n^2-1)a^2/2} = 1$, soit

$$-(n^2 - 1)\frac{a^2}{2} = -\ln(n) \Leftrightarrow a = \sqrt{\frac{2 \ln n}{n^2 - 1}}.$$

Il est facile de voir que cette valeur correspond à un maximum de g , et donc que la valeur de a maximisant $P(a \leq X < na)$ est

$$a = \sqrt{\frac{2 \ln n}{n^2 - 1}}.$$

SOLUTION DE L'EXERCICE 8.14

Pour tout $x \in \mathbf{R}$, on a

$$\begin{aligned} F_Z(x) &= P(Z \leq x) = 1 - P(Z > x) \\ &= 1 - P([X > x] \cap [Y > x]) \end{aligned}$$

Alternative

Ici, nous profitons du fait qu'il est possible de calculer cette intégrale, mais ce n'est pas obligatoire. En effet, on peut se contenter de dire que

$$F_{XY}(x) = 1 - F_{Z+T}(\ln x).$$

On sait alors que F_{Z+T} est continue et \mathcal{C}^1 sauf en 0. De plus on connaît la dérivée de F_{Z+T} , qui est la densité d'une loi $\gamma(2)$. Tous ces éléments suffisent à prouver que F_{XY} est continue, \mathcal{C}^1 sur en 0 et en 1, et à calculer F'_{XY} .

Rédaction

Attention à ne pas dire que la dérivée de F_{XY} est une densité de XY . Ceci n'est vrai que si F_{XY} est dérivable sur \mathbf{R} tout entier, ce qui n'est pas le cas ici.

Rappel

$X \leftrightarrow \mathcal{E}(\lambda)$ si et seulement si $\lambda X \leftrightarrow \mathcal{E}(1)$.

Inégalités

Pour l'instant, rien ne nous indique que Z est une variable à densité, donc on ne peut remplacer les inégalités strictes par des inégalités larges au prétexte que $P(Z = x) = 0$.

$$\begin{aligned}
&= 1 - P(X > x)P(Y > x) \\
&= 1 - (1 - P(X \leq x))(1 - P(Y \leq x)) \\
&= 1 - (1 - F_X(x))(1 - F_Y(x)).
\end{aligned}$$

X et Y sont indépendantes.

Si $x < 0$, on a $F_X(x) = F_Y(x) = 0$ et donc $F_Z(x) = 1 - 1 \times 1 = 0$.

Si $x \geq 0$, alors $F_X(x) = 1 - e^{-\lambda x}$ et $F_Y(x) = 1 - e^{-\mu x}$. Donc $F_Z(x) = 1 - e^{-\lambda x} e^{-\mu x} = 1 - e^{-(\lambda+\mu)x}$.

$$\text{Ainsi, on a } F_Z(x) = \begin{cases} 0 & \text{si } x < 0 \\ 1 - e^{-(\lambda+\mu)x} & \text{si } x \geq 0 \end{cases}$$

On reconnaît là la fonction de répartition d'une variable aléatoire suivant la loi exponentielle de paramètre $\lambda + \mu$: $Z \hookrightarrow \mathcal{E}(\lambda + \mu)$.

SOLUTION DE L'EXERCICE 8.15

- La loi de λX et la loi de λY est une loi $\gamma(1)$.
- Par indépendance de λX et λY , $\lambda Z = \lambda X + \lambda Y$ suit une loi $\gamma(2)$. Donc une densité en est donnée par

$$f(x) = \begin{cases} xe^{-x} & \text{si } x > 0 \\ 0 & \text{sinon} \end{cases}$$

Par transformation affine de densité, une densité de $Z = \frac{1}{\lambda}(\lambda Z)$ est

$$\lambda f(\lambda x) = \begin{cases} \lambda^2 x e^{-\lambda x} & \text{si } x > 0 \\ 0 & \text{sinon} \end{cases}$$

SOLUTION DE L'EXERCICE 8.16

- f est continue, sauf peut-être⁵ en 0, et positive si $\lambda \geq 0$.
De plus, on a, sous réserve de convergence,

$$\int_{-\infty}^{+\infty} f(t) dt = \int_0^{+\infty} \lambda t e^{-at} dt = \lambda \int_0^{+\infty} t e^{-at} dt.$$

À l'aide du changement de variable $x = at$, il vient

$$\lambda \int_0^{+\infty} \frac{x}{a} e^{-x} \frac{dx}{a} = \frac{\lambda}{a^2} \Gamma(2) = \frac{\lambda}{a^2}.$$

Et donc cette intégrale vaut 1 si et seulement si $\lambda = a^2$.

Ainsi, f est une densité si et seulement si $\lambda = a^2$.

- Puisque f est nulle sur $]-\infty, 0]$, il est clair que pour $x \leq 0$, $F_X(x) = 0$.
Pour $x \geq 0$, on a

$$F_X(x) = \int_0^x a^2 t e^{-at} dt = [-at e^{-at}]_0^x + \int_0^x a e^{-at} dt = -axe^{-ax} + [-e^{-at}]_0^x = 1 - e^{-ax} - axe^{-ax}.$$

$$\text{Et donc } F_X(x) = \begin{cases} 0 & \text{si } x \leq 0 \\ 1 - e^{-ax} - axe^{-ax} & \text{sinon} \end{cases}$$

- X admet une espérance si et seulement si $\int_0^{+\infty} a^2 t^2 e^{-at} dt$ converge.

Mais par le même changement de variable $x = at$ qu'à la question 1, on obtient

$$E(X) = \frac{1}{a} \int_0^{+\infty} x^2 e^{-x} dx = \frac{\Gamma(3)}{a} = \frac{2}{a}.$$

SOLUTION DE L'EXERCICE 8.17

Notons φ la⁶ densité de la loi normale centrée réduite, de sorte que $\Phi' = \varphi$.

Procédons alors à une intégration par parties sur $[0, A]$ en posant $u(t) = t$ et $v(t) = 1 - \Phi(t)$.

Ces deux fonctions sont bien de classe \mathcal{C}^1 sur \mathbf{R}_+ et $u'(t) = 1$ et $v'(t) = -\varphi(t) = -\frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}}$.

Alors

$$\int_0^A (1 - \Phi(t)) dt = [t(1 - \Phi(t))]_0^A + \int_0^A t\varphi(t) dt$$

⁵ En fait elle est continue en 0, mais pourquoi perdre du temps à le vérifier ?

⁶ Il y a une petite imprécision quand on dit la densité, puisqu'il existe une infinité de densités. Lorsqu'on dit cela pour une loi normale, on entend par là l'unique densité **continue**, qui est celle donnée dans le cours (changer la valeur en un nombre fini de points nous ferait perdre le caractère continu).

$$\begin{aligned}
&= A(1 - \Phi(A)) + \frac{1}{\sqrt{2\pi}} \int_0^A te^{-t^2/2} dt \\
&= A(1 - \Phi(A)) + \left[-\frac{1}{\sqrt{2\pi}} e^{-t^2/2} \right]_0^A \\
&= A(1 - \Phi(A)) + \frac{1}{\sqrt{2\pi}} - \frac{1}{\sqrt{2\pi}} e^{-A^2/2}.
\end{aligned}$$

Mais, on a

$$A(1 - \Phi(A)) = A \int_A^{+\infty} \frac{1}{\sqrt{2\pi}} e^{-t^2/2} dt = \int_A^{+\infty} \frac{1}{\sqrt{2\pi}} e^{-t^2/2} dt.$$

Or, pour $t \geq A$, on a $Ae^{-t^2/2} \leq te^{-t^2/2}$, de sorte que

$$\int_A^{+\infty} e^{-t^2/2} dt \leq \int_A^{+\infty} te^{-t^2/2} dt = \lim_{B \rightarrow +\infty} \int_A^B te^{-t^2/2} dt = \lim_{B \rightarrow +\infty} [-e^{-t^2/2}]_A^B = e^{-A^2/2}.$$

Et donc pour $A \geq 0$, on a $0 \leq A(1 - \Phi(A)) \leq e^{-A^2/2}$.

D'après le théorème des gendarmes, on a donc $\lim_{A \rightarrow +\infty} A(1 - \Phi(A)) = 0$. On en déduit que

$$\lim_{A \rightarrow +\infty} \int_0^A (1 - \Phi(t)) dt = \frac{1}{\sqrt{2\pi}}.$$

Et donc $\int_0^{+\infty} (1 - \Phi(t)) dt$ converge et vaut $\frac{1}{\sqrt{2\pi}}$.

SOLUTION DE L'EXERCICE 8.18

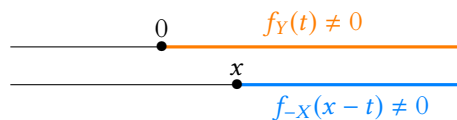
1. Par transformation affine de densités, une densité de $-X$ est

$$x \mapsto f_X(-x) = \begin{cases} ae^{ax} & \text{si } x \leq 0 \\ 0 & \text{si } x > 0 \end{cases}$$

2. Les variables Y et $-X$ sont indépendantes, et la densité de Y est bornée, donc une densité de $Y - X$ est donnée par le produit de convolution :

$$h(x) = \int_{-\infty}^{+\infty} f_Y(t)f_{-X}(x-t) dt = \int_0^{+\infty} f_Y(t)f_{-X}(x-t) dt.$$

On a $f_Y(t) \neq 0 \Leftrightarrow t \geq 0$ et $f_{-X}(x-t) \neq 0 \Leftrightarrow x-t \leq 0 \Leftrightarrow t \geq x$.



Ainsi, on a $f_Y(t)f_{-X}(x-t) \neq 0 \Leftrightarrow \begin{cases} t \geq 0 \\ t \geq x \end{cases} \Leftrightarrow t \geq \max(x, 0)$.

Si $x > 0$, alors

$$h(x) = \int_x^{+\infty} abe^{-bt} e^{a(x-t)} dt = abe^{ax} \int_x^{+\infty} e^{-(a+b)t} dt.$$

Mais pour $A > x$, on a

$$\int_x^A e^{-(a+b)t} dt = \left[\frac{-1}{a+b} e^{-(a+b)t} \right]_x^A = \frac{1}{a+b} (e^{-(a+b)x} - e^{-(a+b)A}) \xrightarrow{A \rightarrow +\infty} \frac{e^{-(a+b)x}}{a+b}.$$

Et donc $h(x) = abe^{ax} \frac{e^{-(a+b)x}}{a+b} = \frac{abe^{-bx}}{a+b}$.

Si $x \leq 0$, alors

$$h(x) = abe^{ax} \int_0^{+\infty} e^{-(a+b)t} dt = \frac{e^{ax}}{a+b}.$$

SOLUTION DE L'EXERCICE 8.19

Détails

Si X est une variable aléatoire suivant la loi normale centrée réduite, alors $1 - \Phi(A) = P(X > A)$.

Domaine

Ce calcul nous donne ce qu'on pouvait déjà voir sur le dessin : il va falloir distinguer les deux cas $x \geq 0$ et $x < 0$. Si vous y arrivez sur le dessin, il est inutile de réécrire cette ligne de calcul et vice-versa.

- 1.a. Une densité de Y est $f_Y : x \mapsto \begin{cases} \frac{1}{a} & \text{si } x \in [0, a] \\ 0 & \text{sinon} \end{cases}$

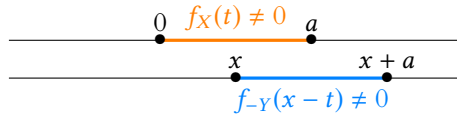
Donc par transformation affine, une densité de $-Y$ est

$$f_{-Y} : x \mapsto \frac{1}{|-1|} f_Y(-x) = \begin{cases} \frac{1}{a} & \text{si } -x \in [0, a] \\ 0 & \text{sinon} \end{cases} = \begin{cases} \frac{1}{a} & \text{si } -a \leq x \leq 0 \\ 0 & \text{sinon} \end{cases}$$

- 1.b. Les variables X et $-Y$ sont indépendantes, donc par convolution, une densité de $X - Y = X + (-Y)$ est donnée par

$$g : x \mapsto \int_{-\infty}^{+\infty} f_X(t) f_{-Y}(x-t) dt.$$

On a $f_X(t) \neq 0 \Leftrightarrow 0 \leq t \leq a$ et $f_{-Y}(x-t) \neq 0 \Leftrightarrow -a \leq x-t \leq 0 \Leftrightarrow x \leq t \leq x+a$.



Ainsi, si $x \leq -a$, on a $g(x) = 0$. Et de même si $x \geq a$.

Pour $x+a \in [0, a] \Leftrightarrow x \in [-a, 0]$, on a

$$g(x) = \int_0^{x+a} f_X(t) f_{-Y}(x-t) dt = \int_0^{x+a} \frac{1}{a^2} dt = \frac{x+a}{a^2}.$$

Et si $0 \leq x \leq a$, alors

$$g(x) = \int_x^a f_X(t) f_{-Y}(x-t) dt = \frac{a-x}{a^2}.$$

Notons que pour $-a \leq x \leq 0$, on a bien $x+a = a+x = a-(-x) = a-|x|$. Et donc une densité de $X - Y$ est donnée par

$$g(x) = \begin{cases} \frac{a-|x|}{a^2} & \text{si } x \in [-a, a] \\ 0 & \text{sinon} \end{cases}$$

- 2.a. Il est évident que Z ne prend que des valeurs positives. Et donc, pour $x < 0$, on a

$$H(x) = F_Z(x) = P(Z \leq x) = 0.$$

Pour $x \geq 0$, alors on a

$$\begin{aligned} H(x) = F_Z(x) &= P(Z \leq x) = P(|X - Y| \leq x) = P(-x \leq X - Y \leq x) \\ &= P(X - Y \leq x) - P(X - Y < -x) = G(x) - G(-x). \end{aligned}$$

- 2.b. G est continue sur \mathbf{R} en tant que fonction de répartition d'une variable à densité. Donc H est continue sur $[0; +\infty[$, et elle l'est évidemment sur \mathbf{R}^* .

Comme on a de plus $H(0) = 0 = \lim_{x \rightarrow 0^+} H(x)$, alors H est continue en 0, et donc sur \mathbf{R} .

Il est évident que F_Z est dérivable sur \mathbf{R}^* , de dérivée nulle.

La densité g que nous avons obtenue précédemment est continue sur $] -\infty, a[$, sur $] -a, a[$ et sur $]a; +\infty[$, donc G est dérivable sur ces intervalles, avec $G'(x) = g(x)$.

Ainsi, F_Z est dérivable sur \mathbf{R} , sauf peut-être en $-a$ en 0 et en a . On a alors

$$\forall x \notin \{-a, 0, a\}, F_Z'(x) = G'(x) + G'(-x) = g(x) + g(-x).$$

Toute fonction qui coïncide avec F_Z' là où celle-ci est définie est une densité de Z . On peut donc prendre la fonction

$$x \mapsto \begin{cases} g(x) + g(-x) & \text{si } x \geq 0 \\ 0 & \text{sinon} \end{cases}$$

Notons alors que g est paire, de sorte que $g(x) + g(-x) = 2g(x)$. Et donc une densité de Z est

$$h(x) = \begin{cases} 2g(x) & \text{si } x \geq 0 \\ 0 & \text{sinon} \end{cases} = \begin{cases} \frac{2(a-x)}{a^2} & \text{si } x \in [0, a] \\ 0 & \text{sinon} \end{cases}$$

Autrement dit

$-Y$ suit la loi uniforme sur $[-a, 0]$, et on pouvait obtenir ce résultat par transformation affine de loi uniforme.

Support

Ce résultat est prévisible si on s'intéresse au support : $X - Y$ prend ses valeurs dans $[-a; a]$.

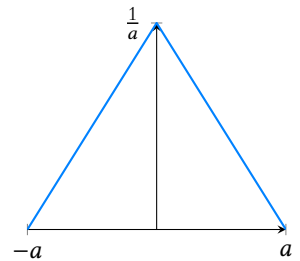


FIGURE 8.3- La densité g .

Densité

On a

$$\begin{aligned} &P(X - Y < -x) \\ &= P(X - Y \leq -x) = G(-x) \end{aligned}$$

car $X - Y$ est une variable à densité.

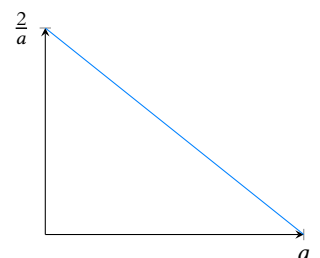


FIGURE 8.4- La densité h .

SOLUTION DE L'EXERCICE 8.20

Un réel λ est valeur propre de M si et seulement si $\det(M - \lambda I_2) = 0$, soit si et seulement si

$$(2X - \lambda)(X - \lambda) + 4 = 0 \Leftrightarrow \lambda^2 - 3X\lambda + 2X^2 + 4 = 0.$$

Cette équation possède deux racines distinctes si et seulement si son discriminant est strictement positif, c'est-à-dire si et seulement si $X^2 - 16 > 0$.

C'est le cas si et seulement si $X > 4$ ou $X < -4$.

Autrement dit, l'événement « M possède deux valeurs propres distinctes» est l'événement $[X > 4] \cup [X < -4]$.

Par incompatibilité de $[X > 4]$ et $[X < -4]$, la probabilité cherchée est donc $P(X > 4) + P(X < -4)$.

Mais X suivant une loi $\mathcal{N}(0, 4)$, $\frac{X}{2}$ suit une loi normale centrée réduite, de sorte que

$$P(X > 4) = P\left(\frac{X}{2} > 2\right) = 1 - P\left(\frac{X}{2} \leq 2\right) = 1 - \Phi(2).$$

De même, $P(X < -4) = \Phi(-2) = 1 - \Phi(2)$.

Et donc la probabilité que M possède deux valeurs propres distinctes est $2 - 2\Phi(2) \simeq 0.045$.

SOLUTION DE L'EXERCICE 8.21

1. Il est facile de voir que f est continue et positive sur \mathbf{R} .

De plus, f est paire, donc $\int_{-\infty}^{+\infty} f(t) dt$ converge si et seulement si $\int_0^{+\infty} f(t) dt$ converge.

Or

$$\int_0^{+\infty} f(t) dt = \frac{\lambda}{2} \int_0^{+\infty} e^{-\lambda t} dt = \frac{1}{2}.$$

Donc $\int_{-\infty}^{+\infty} f(t) dt$ converge et vaut $2 \times \frac{1}{2} = 1$.

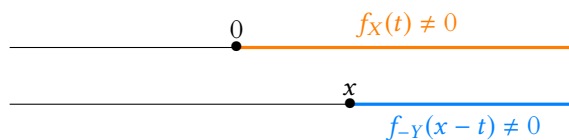
Donc f est bien une densité.

2. Une densité de $-Y$ est⁷ est $t \mapsto \begin{cases} \lambda e^{\lambda t} & \text{si } t \leq 0 \\ 0 & \text{sinon} \end{cases}$

Puisque X et Y sont indépendantes, il en est de même de X et $-Y$. Comme de plus la densité de X est bornée, on peut utiliser le produit de convolution : une densité de $X - Y$ est

$$h : x \mapsto \int_{-\infty}^{+\infty} f_X(t) f_{-Y}(x-t) dt.$$

Or, $f_X(t) \neq 0$ si et seulement si $t \geq 0$ et $f_{-Y}(x-t) \neq 0 \Leftrightarrow x-t \leq 0 \Leftrightarrow t \geq x$.



Pour $x \leq 0$, on a alors

$$h(x) = \int_0^{+\infty} \lambda^2 e^{-\lambda t} e^{\lambda(x-t)} dt = \frac{\lambda}{2} e^{\lambda x} \int_0^{+\infty} 2\lambda e^{-2\lambda t} dt = \frac{\lambda}{2} e^{\lambda x}.$$

Et pour $x \geq 0$, il vient

$$h(x) = \int_x^{+\infty} \lambda^2 e^{-\lambda t} e^{\lambda(x-t)} dt = \frac{\lambda}{2} e^{\lambda x} \int_x^{+\infty} 2\lambda e^{-2\lambda t} dt.$$

Notons que là encore, nous connaissons directement la valeur de l'intégrale : si V suit une

loi exponentielle $\mathcal{E}(2\lambda)$, alors $\int_x^{+\infty} 2\lambda e^{-\lambda t} dt = P(V \geq x) = 1 - F_V(x) = e^{-2\lambda x}$.

Et donc $h(x) = \frac{\lambda}{2} e^{\lambda x} e^{-2\lambda x} = \frac{\lambda}{2} e^{-\lambda x}$.

Au final, on a donc, pour tout $x \in \mathbf{R}$, $h(x) = \frac{\lambda}{2} e^{-\lambda|x|} = f(x)$.

Méthode

Pour calculer des probabilités avec une loi normale, on n'a d'autre solution que de les exprimer à l'aide de Φ en faisant apparaître une loi normale centrée réduite.

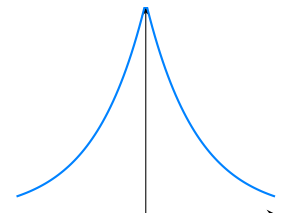


FIGURE 8.5— La densité f .

⁷ Par transformation affine.

Astuce

En faisant apparaître

$$\int_0^{+\infty} 2\lambda e^{-2\lambda t} dt$$

on s'est ramené à une intégrale connue : celle de la densité d'une loi exponentielle de paramètre 2λ , et on a donc évité un calcul de primitive.

La variable aléatoire $Z = |X - Y|$ ne prend que des valeurs positives, donc pour $x < 0$, $P(Z \leq x) = 0$.

Et pour $x \geq 0$, on a $P(Z \leq x) = P(-x \leq X - Y \leq x) = \int_{-x}^x \frac{\lambda}{2} e^{-\lambda|t|} dt$.

On reconnaît alors l'intégrale d'une fonction paire, de sorte que

$$P(Z \leq x) = 2 \int_0^x \frac{\lambda}{2} e^{-\lambda|t|} dt = \int_0^x \lambda e^{-\lambda t} dt = 1 - e^{-\lambda x}.$$

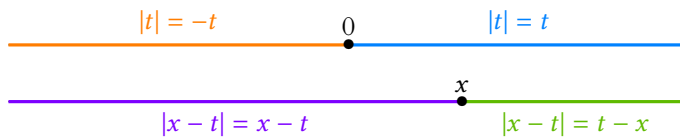
On reconnaît donc la fonction de répartition d'une loi exponentielle de paramètre λ : $Z \hookrightarrow \mathcal{E}(\lambda)$.

3. f est bornée par $\frac{1}{2}$, donc, X et Y étant indépendantes, $S = X + Y$ admet une densité, et une densité de S est donnée par

$$h(x) = \int_{-\infty}^{+\infty} f(t)f(x-t)dt = \frac{\lambda^2}{4} \int_{-\infty}^{+\infty} e^{-\lambda|t|} e^{-\lambda|x-t|} dt.$$

L'intégrande n'est jamais nulle, mais il y a le problème des valeurs absolues.

Pour cela, adaptons le dessin usuel : $|t| = t \Leftrightarrow t \geq 0$ et $|x-t| = x-t \Leftrightarrow x-t \geq 0 \Leftrightarrow x \geq t$.



Si $x \geq 0$, alors

$$\begin{aligned} h(x) &= \frac{\lambda^2}{4} \left(\int_{-\infty}^0 e^{\lambda t} e^{\lambda(t-x)} dt + \int_0^x e^{-\lambda t} e^{\lambda(t-x)} dt + \int_x^{+\infty} e^{-\lambda t} e^{\lambda(x-t)} dt \right) \\ &= \frac{\lambda^2}{4} \left(e^{-\lambda x} \int_{-\infty}^0 e^{2\lambda t} dt + e^{-\lambda x} \int_0^x dt + e^{\lambda x} \int_x^{+\infty} e^{-2\lambda t} dt \right) \\ &= \frac{\lambda^2}{4} \left(\frac{1}{2\lambda} e^{-\lambda x} + x e^{-\lambda x} + \frac{1}{2\lambda} e^{\lambda x} e^{-2\lambda x} \right) = \frac{\lambda e^{-\lambda x}}{4} (\lambda x + 1). \end{aligned}$$

Si $x \leq 0$, alors en procédant de même, on a

$$h(x) = \frac{\lambda^2}{4} \left(\int_{-\infty}^x e^{\lambda t} e^{\lambda(t-x)} dt + \int_x^0 e^{\lambda t} e^{-\lambda(t-x)} dt + \int_0^{+\infty} e^{-\lambda t} e^{\lambda(x-t)} dt \right) = \frac{\lambda^2}{4} \left(\frac{1}{2\lambda} e^{\lambda x} - x e^{\lambda x} + \frac{1}{2\lambda} e^{\lambda x} \right)$$

Notons qu'il est possible de regrouper les deux formules obtenues sous une seule forme

$$\forall x \in \mathbf{R}, h(x) = \frac{\lambda}{4} e^{-\lambda|x|} (1 + \lambda|x|).$$

SOLUTION DE L'EXERCICE 8.22

Soit $Z = X - Y$. Puisque $-Y \hookrightarrow \mathcal{N}(-v, \sigma^2)$ et que X et $-Y$ sont indépendantes, par stabilité des lois normales $Z \hookrightarrow \mathcal{N}(\mu - v, 2\sigma^2)$.

Mais $P(X \leq Y) = P(Z \leq 0)$.

Or, nous savons que $P(Z \leq \mu - v) = \frac{1}{2}$.

Donc si $\mu \leq v$, $P(Z \leq 0) \geq P(Z \leq \mu - v) = \frac{1}{2}$.

Inversement, si $v < \mu$, alors

$$P(Z \leq 0) < P(Z \leq \mu - v) = \frac{1}{2}.$$

Au final, on a

$$P(X \leq Y) \geq \frac{1}{2} \Leftrightarrow \mu \leq v.$$

SOLUTION DE L'EXERCICE 8.23

Méthode

On ne sait à peu près jamais intégrer une quantité avec des valeurs absolues. Pour calculer une telle intégrale, il faut donc la «découper» (à l'aide de Chasles) en différents morceaux sur lesquels on peut se passer de la valeur absolue, en revenant à la définition :

$$|x| = \begin{cases} x & \text{si } x \geq 0 \\ -x & \text{sinon} \end{cases}$$

Rédaction

Ne pas oublier de vérifier l'hypothèse d'indépendance pour utiliser la stabilité des lois normales (ou tout autre résultat de stabilité).

Il est facile de remarquer que $X(\Omega) = \mathbf{N}^*$, et donc que X est une variable aléatoire discrète. On a, pour $k \in \mathbf{N}^*$,

$$P(X = k) = P\left(\left\lfloor \frac{\ln U}{\ln q} \right\rfloor = k - 1\right) = P\left(k - 1 \leq \frac{\ln U}{\ln q} < k\right) = P(k \ln q < \ln U \leq (k - 1) \ln q).$$

Par croissance de l'exponentielle, on en déduit que

$$P(X = k) = P(q^k < U \leq q^{k-1}) = q^{k-1} - q^k$$

car U suit la loi uniforme sur $[0, 1]$.

Et donc $P(X = k) = q^{k-1}(1 - q)$: X suit la loi géométrique de paramètre $1 - q$.

SOLUTION DE L'EXERCICE 8.24

1.a. La fonction $f : t \mapsto \left(\frac{a}{\sqrt{t}} - 1\right) \left(\frac{b}{\sqrt{\alpha - t}} - 1\right)$ est continue sur $]0, \alpha[$, donc il y a d'éventuels problèmes de convergence au voisinage de 0 et de α .

Notons qu'il s'agit d'une fonction positive, donc nous pouvons utiliser le critère des équivalents.

Au voisinage de 0, $f(t) \sim \frac{a}{\sqrt{t}} \left(\frac{b}{\sqrt{\alpha}} - 1\right)$.

Or, $\left(\frac{b}{\sqrt{\alpha}} - 1\right)$ est une constante, et $\int_0^{\alpha/2} \frac{a}{\sqrt{t}} dt$ est une intégrale de Riemann convergente.

Donc $\int_0^{\alpha/2} \left(\frac{a}{\sqrt{t}} - 1\right) \left(\frac{b}{\sqrt{\alpha - t}} - 1\right) dt$ converge.

Au voisinage de α , $\left(\frac{a}{\sqrt{t}} - 1\right) \left(\frac{b}{\sqrt{\alpha - t}} - 1\right) \underset{t \rightarrow \alpha}{\sim} \left(\frac{a}{\sqrt{\alpha}} - 1\right) \frac{b}{\sqrt{\alpha - t}}$.

Mais $\int_{\alpha/2}^{\alpha} \frac{b}{\sqrt{\alpha - t}}$ est une intégrale de Riemann convergente, donc $\int_{\alpha/2}^{\alpha} \left(\frac{a}{\sqrt{t}} - 1\right) \left(\frac{b}{\sqrt{\alpha - t}} - 1\right) dt$ converge.

1.b. La fonction $t \mapsto \alpha \cos^2 t$ est de classe \mathcal{C}^1 sur $[0, \frac{\pi}{2}]$, strictement décroissante et à valeurs dans $[0, \alpha]$.

Si on pose $u = \alpha \cos^2 t$, alors $dt = -2\alpha \sin u \cos u du$ et donc

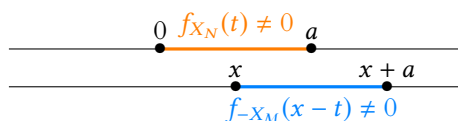
$$\begin{aligned} \int_0^{\alpha} \left(\frac{a}{\sqrt{t}} - 1\right) \left(\frac{b}{\sqrt{\alpha - t}} - 1\right) dt &= -2\alpha \int_{\pi/2}^0 \left(\frac{a}{\sqrt{\alpha \cos^2 u}} - 1\right) \left(\frac{b}{\sqrt{\alpha(1 - \cos^2 u)}} - 1\right) \sin u \cos u du \\ &= 2 \int_0^{\pi/2} (a - \sqrt{\alpha} \cos u)(b - \sqrt{\alpha} \sin u) du \\ &= 2 \int_0^{\pi/2} ab du - 2a\sqrt{\alpha} \int_0^{\pi/2} \sin u du - 2b\sqrt{\alpha} \int_0^{\pi/2} \cos u du + 2\alpha \int_0^{\pi/2} \sin u \cos u du \\ &= \pi ab - 2\sqrt{\alpha}(a + b) + 2\alpha \left[\frac{\sin^2 u}{2}\right]_0^{\pi/2} \\ &= \pi ab - 2\sqrt{\alpha}(a + b) + \alpha. \end{aligned}$$

2.a. Par transformation affine de loi uniforme, $-X_M$ suit la loi uniforme sur $[-a, 0]$.

2.b. Les variables X_N et $-X_M$ étant indépendantes, une densité de Z_a est donnée par

$$f_{Z_a}(x) = \int_{-\infty}^{+\infty} f_{X_N}(t) f_{-X_M}(x - t) dt.$$

Or, $f_{X_N}(t) \neq 0 \Leftrightarrow 0 \leq t \leq a$ et $f_{-X_M}(x - t) \neq 0 \Leftrightarrow -a \leq x - t \leq 0 \Leftrightarrow x \leq t \leq x + a$.



Donc si $x \leq -a$, $f_{Z_a}(x) = 0$ et de même, si $x \geq a$, $f_{Z_a}(x) = 0$.

Si $-a \leq x \leq 0$, alors

$$f_{Z_a}(x) = \int_0^{x+a} \frac{1}{a} \frac{1}{a} dt = \frac{x + a}{a^2}.$$

Danger !
 $q \in]0, 1[$, et donc $\ln(q) < 0$: il ne faut donc pas oublier de changer le sens des inégalités en multipliant par $\ln(q)$.

Détails
 -1 est négligeable devant $\frac{a}{\sqrt{t}}$ (qui tend vers $+\infty$), donc la somme des deux est équivalente au terme prépondérant : $\frac{a}{\sqrt{t}}$.

Et si $0 \leq x \leq a$, alors

$$f_{Z_a}(x) = \int_x^a \frac{1}{a} \frac{1}{a} dt = \frac{a-x}{a^2}.$$

De même⁸, on prouve qu'une densité de Z_b est donnée par

⁸ En changeant les a en b .

$$f_{Z_b}(x) = \begin{cases} \frac{x+b}{b^2} & \text{si } -b \leq x \leq 0 \\ \frac{b-x}{b^2} & \text{si } 0 \leq x \leq b \\ 0 & \text{sinon} \end{cases} = \begin{cases} \frac{b-|x|}{b^2} & \text{si } -b \leq x \leq b \\ 0 & \text{sinon} \end{cases}$$

- 2.c. Puisque $Z_a(\Omega) = [-a, a]$, il vient $Z_a^2(\Omega) = [0, a^2]$.
Et donc $P(Z_a^2 \leq x) = 0$ si $x < 0$ et $P(Z_a^2 \leq x) = 1$ si $x \geq a^2$.
Si $x \in [0, a^2]$, alors

$$F_{Z_a^2}(x) = P(Z_a^2 \leq x) = P(-\sqrt{x} \leq Z_a \leq \sqrt{x}) = \int_{-\sqrt{x}}^{\sqrt{x}} f_{Z_a}(t) dt = 2 \int_0^{\sqrt{x}} \frac{a-t}{a^2} dt = 2 \left[\frac{t}{a} - \frac{t^2}{2a^2} \right]_0^{\sqrt{x}} = \frac{2\sqrt{x}}{a} - \frac{x}{a^2}.$$

Il est alors facile de prouver que $F_{Z_a^2}$ est continue en 0 et en a^2 , et donc que $F_{Z_a^2}$ est une fonction continue sur \mathbf{R} et \mathcal{C}^1 sauf peut-être en 0 et en a^2 .

Par conséquent, Z_a^2 est une variable à densité, et une densité en est donnée par

$$f_{Z_a^2}(x) = \begin{cases} \frac{1}{a\sqrt{x}} - \frac{1}{a^2} = \frac{1}{a^2} \left(\frac{a}{\sqrt{x}} - 1 \right) & \text{si } 0 < x < a^2 \\ 0 & \text{sinon} \end{cases}$$

On en déduit de même qu'une densité de Z_b^2 est

$$f_{Z_b^2}(x) = \begin{cases} \frac{1}{b^2} \left(\frac{b}{\sqrt{x}} - 1 \right) & \text{si } 0 < x < b^2 \\ 0 & \text{sinon} \end{cases}$$

3. On a $D^2 = Z_a^2 + Z_b^2$. Par le lemme des coalitions⁹, les variables Z_a^2 et Z_b^2 sont indépendantes car X_N, X_M, Y_N et Y_M le sont. Une densité de D^2 est alors donnée par

$$f_{D^2}(x) = \int_{-\infty}^{+\infty} f_{Z_a^2}(t) f_{Z_b^2}(x-t) dt = \int_0^{a^2} f_{Z_a^2}(t) f_{Z_b^2}(x-t) dt.$$

Notons que pour calculer $P(D \leq \theta) = P(0 \leq D^2 \leq \theta^2)$, nous n'avons besoin de connaître une densité de D^2 que sur l'intervalle $[0, \theta^2] \subset [0, a^2]$.

Or, pour $0 \leq x < a$, on a $f_{Z_a^2}(t) = \frac{1}{a^2} \left(\frac{a}{\sqrt{t}} - 1 \right)$ et

$$f_{Z_b^2}(x-t) \neq 0 \Leftrightarrow 0 \leq x-t \leq b^2 \Leftrightarrow x-b^2 \leq t \leq x.$$

Donc, pour $x \in [0, a^2]$, on a

$$f_{D^2}(x) = \int_0^x \frac{1}{a^2} \left(\frac{a}{\sqrt{t}} - 1 \right) \frac{1}{b^2} \left(\frac{b}{\sqrt{x-t}} - 1 \right) dt = \frac{1}{a^2 b^2} I_{a,b}(x).$$

Ainsi, il vient¹⁰

$$P(D \leq \theta) = P(D^2 \leq \theta^2) = \int_{-\infty}^{\theta^2} f_{D^2}(x) dx = \int_0^{\theta^2} \frac{1}{a^2 b^2} I_{a,b}(x) dx.$$

En utilisant le résultat de la question 1, on a donc

$$\begin{aligned} P(D \leq \theta) &= \frac{1}{a^2 b^2} \int_0^{\theta^2} (\pi ab - 2\sqrt{x}(a+b) + x) dx \\ &= \frac{\pi \theta^2}{ab} - \frac{4}{3} \frac{a+b}{a^2 b^2} \theta^3 + \frac{\theta^4}{2a^2 b^2}. \end{aligned}$$

⁹ Nous verrons ce résultat plus tard, mais il est très intuitif : Z_a est formé à partir de X_N et X_M , et Z_b à partir de Y_N et Y_M . Comme ces variables sont indépendantes, il doit en être de même de Z_a et Z_b .

¹⁰ D ne prend que des valeurs positives, donc

$$[D \leq \theta] = [D^2 \leq \theta^2].$$

PRODUITS SCALAIRES, ESPACES

EUCLIDIENS

Cours	281
9.1 Produits scalaires	281
9.2 Orthogonalité	285
9.3 Espaces euclidiens	290
Méthodes	295
Exercices	296
Correction des exercices	299

Dans ce chapitre, tous les espaces vectoriels considérés sont des espaces vectoriels sur \mathbf{R} .

9.1 PRODUITS SCALAIRES

9.1.1 Formes bilinéaires symétriques

Définition 9.1 – Une forme bilinéaire sur un espace vectoriel E est une application $\varphi : E \times E \rightarrow \mathbf{R}$ vérifiant

$$1) \forall y \in E, \varphi_y : \begin{cases} E & \longrightarrow \mathbf{R} \\ x & \longmapsto \varphi(x, y) \end{cases} \text{ est linéaire, c'est-à-dire}$$

$$\forall (x, z) \in E^2, \forall \lambda \in \mathbf{R}, \varphi(\lambda x + z, y) = \lambda \varphi(x, y) + \varphi(z, y)$$

$$2) \forall x \in E, \varphi_x : \begin{cases} E & \longrightarrow \mathbf{R} \\ y & \longmapsto \varphi(x, y) \end{cases} \text{ est linéaire, c'est-à-dire}$$

$$\forall (y, z) \in E^2, \forall \lambda \in \mathbf{R}, \varphi(x, \lambda y + z) = \lambda \varphi(x, y) + \varphi(x, z).$$

Terminologie

On dit alors que φ est linéaire par rapport à la première variable, ou linéaire à gauche.

On dit alors que φ est linéaire par rapport à la seconde variable, ou linéaire à droite.

⚠ Ceci ne signifie pas que φ est linéaire de l'espace vectoriel $E \times E$ dans \mathbf{R} .

Exemples 9.2

$$\bullet \varphi : \begin{cases} \mathbf{R}^n \times \mathbf{R}^n & \longrightarrow \mathbf{R} \\ (x_1, \dots, x_n), (y_1, \dots, y_n) & \longmapsto \sum_{i=1}^n x_i y_i \end{cases} \text{ est bilinéaire.}$$

En effet, soient $(x_1, \dots, x_n), (y_1, \dots, y_n), (z_1, \dots, z_n) \in \mathbf{R}^n$ et soit $\lambda \in \mathbf{R}$. Alors

$$\begin{aligned} \varphi((\lambda x_1, \dots, \lambda x_n) + (y_1, \dots, y_n), (z_1, \dots, z_n)) &= \varphi((\lambda x_1 + y_1, \dots, \lambda x_n + y_n), (z_1, \dots, z_n)) \\ &= \sum_{i=1}^n (\lambda x_i + y_i) z_i = \lambda \sum_{i=1}^n x_i z_i + \sum_{i=1}^n y_i z_i \\ &= \lambda \varphi((x_1, \dots, x_n), (z_1, \dots, z_n)) + \varphi((y_1, \dots, y_n), (z_1, \dots, z_n)) \end{aligned}$$

On prouverait de même la linéarité par rapport à la seconde variable.

$$\bullet \det_2 : \begin{cases} \mathbf{R}^2 \times \mathbf{R}^2 & \longrightarrow \mathbf{R} \\ (a \ b), (c \ d) & \longmapsto ad - bc \end{cases} \text{ est une forme bilinéaire sur } \mathbf{R}^2. \text{ En effet,}$$

pour $(a, b), (a', b')$ et $(c, d) \in \mathbf{R}^2, \lambda \in \mathbf{R}$, on a

$$\det_2(\lambda(a, b) + (a', b'), (c, d)) = (\lambda a + a')d - (\lambda b + b')c$$

Remarque

En fait, il est souvent inutile de prouver cette linéarité par rapport à la seconde variable par le calcul, voir page suivante.

$$\begin{aligned}
 &= \lambda(ad - bc) + (a'd - b'c) \\
 &= \lambda \det_2((a, b), (c, d)) + \det_2((a', b'), (c, d)).
 \end{aligned}$$

Définition 9.3 – Une forme bilinéaire $\varphi : E \times E \rightarrow \mathbf{R}$ est dite symétrique si

$$\forall (x, y) \in E^2, \varphi(x, y) = \varphi(y, x).$$

Exemples 9.4

• $\varphi : \left. \begin{array}{l} \mathbf{R}^n \times \mathbf{R}^n \longrightarrow \mathbf{R} \\ (x_1, \dots, x_n), (y_1, \dots, y_n) \longmapsto \sum_{i=1}^n x_i y_i \end{array} \right\} \text{ est symétrique car}$

$$\varphi((x_1, \dots, x_n), (y_1, \dots, y_n)) = \sum_{i=1}^n x_i y_i = \sum_{i=1}^n y_i x_i = \varphi((y_1, \dots, y_n), (x_1, \dots, x_n))$$

• \det_2 n'est pas symétrique car

$$\det_2((1 \ 0), (0 \ 1)) = 1 \text{ et } \det_2((0 \ 1), (1 \ 0)) = -1$$

Remarque. Pour vérifier que φ est bilinéaire symétrique, il suffit de vérifier que :

- $\forall (x, y, z) \in E^3, \forall \lambda \in \mathbf{R}, \varphi(\lambda x + y, z) = \lambda \varphi(x, z) + \varphi(y, z)$: φ est linéaire par rapport à sa première variable
- $\forall (x, y) \in E^2, \varphi(x, y) = \varphi(y, x)$: φ est symétrique.

En effet, on a alors

$$\forall (x, y, z) \in E^3, \varphi(x, \lambda y + z) = \varphi(\lambda y + z, x) = \lambda \varphi(y, x) + \varphi(z, x) = \lambda \varphi(x, y) + \varphi(x, z).$$

Donc φ est nécessairement linéaire par rapport à sa seconde variable.

Exemples 9.5

• Si $E = \mathcal{C}^0([0, 1], \mathbf{R})$, alors $\varphi : \left. \begin{array}{l} E \times E \longrightarrow \mathbf{R} \\ (f, g) \longmapsto \int_0^1 f(t)g(t) dt \end{array} \right\} \text{ est une forme bilinéaire symétrique car}$

$$\forall (f, g, h) \in \mathcal{C}^0([0, 1], \mathbf{R})^2, \forall \lambda \in \mathbf{R}, \int_0^1 (\lambda f(t) + g(t))h(t) dt = \lambda \int_0^1 f(t)h(t) dt + \int_0^1 g(t)h(t) dt$$

et $\forall f, g \in \mathcal{C}^0([0, 1], \mathbf{R}), \int_0^1 f(t)g(t) dt = \int_0^1 g(t)f(t) dt.$

• Si E est l'ensemble des variables aléatoires sur (Ω, \mathcal{A}, P) qui admettent une variance, alors $\text{Cov} : (X, Y) \mapsto \text{Cov}(X, Y)$ est une forme bilinéaire symétrique.

Méthode

On procédera toujours ainsi pour prouver qu'une application est bilinéaire symétrique : on commencera par prouver la symétrie, puis on prouvera la linéarité par rapport à l'une des deux variables. On en déduira qu'alors automatiquement la bilinéarité. Attention à bien expliquer le raisonnement : « φ est symétrique, linéaire par rapport à la première variable donc bilinéaire».

Proposition 9.6 : Si $\varphi : E \times E \rightarrow \mathbf{R}$ est une forme bilinéaire symétrique, alors :

1. $\forall x \in E, \varphi(x, 0_E) = \varphi(0_E, x) = 0.$
2. $\forall (x, y) \in E^2, \varphi(x + y, x + y) = \varphi(x, x) + 2\varphi(x, y) + \varphi(y, y).$

Démonstration. 1) Pour $x \in E$, l'application $y \mapsto \varphi(x, y)$ est linéaire, donc en particulier envoie le vecteur nul sur 0 : $\varphi(x, 0_E) = 0.$

Et par symétrie, on a alors $\varphi(0_E, x) = \varphi(x, 0_E) = 0.$

2) Soient $x, y \in E$. Alors

$$\begin{aligned}\varphi(x+y, x+y) &= \varphi(x, x+y) + \varphi(y, x+y) \\ &= \varphi(x, x) + \varphi(x, y) + \varphi(y, x) + \varphi(y, y) \\ &= \varphi(x, x) + \varphi(x, y) + \varphi(x, y) + \varphi(y, y) \\ &= \varphi(x, x) + 2\varphi(x, y) + \varphi(y, y).\end{aligned}$$

Linéarité à gauche.

Linéarité à droite.

Symétrie.

□

9.1.2 Produit scalaire

Définition 9.7 – Une application $\varphi : E \times E \rightarrow \mathbf{R}$ est un produit scalaire si :

1. φ est une forme bilinéaire symétrique sur E
2. $\forall x \in E, \varphi(x, x) \geq 0$ (on dit que φ est positive)
3. $\forall x \in E, \varphi(x, x) = 0 \Leftrightarrow x = 0_E$ (on dit que φ est définie).

Un produit scalaire est donc une forme bilinéaire symétrique définie positive.

Remarques. • Pour prouver que $\varphi : E \times E \rightarrow \mathbf{R}$ est un produit scalaire, il faut vérifier quatre propriétés :

- 1) φ est linéaire par rapport à sa première variable.
- 2) φ est symétrique (et donc 1) et 2) impliquent la linéarité par rapport à la seconde variable).
- 3) Pour tout $x \in E, \varphi(x, x) \geq 0$.
- 4) Si $\varphi(x, x) = 0$, alors $x = 0_E$ (comme indiqué précédemment, l'implication réciproque est toujours vraie et n'a donc pas besoin d'être vérifiée.)

• En général, on note $\langle x, y \rangle$ ou $(x|y)$ au lieu de $\varphi(x, y)$.

Remarque

L'implication

$$x = 0_E \Rightarrow \varphi(x, x) = 0$$

est toujours vérifiée pour une forme bilinéaire, donc il s'agit surtout de vérifier que

$$\varphi(x, x) = 0 \Rightarrow x = 0_E.$$

Rédaction

Ce dernier point est souvent celui qui est le plus difficile à prouver, et qui demande une vraie justification. Attention à ne pas le négliger.

Exemples 9.8

$$\bullet \varphi : \begin{array}{l} \mathbf{R}^n \times \mathbf{R}^n \longrightarrow \mathbf{R} \\ (x_1, \dots, x_n), (y_1, \dots, y_n) \longmapsto \sum_{i=1}^n x_i y_i \end{array}$$

On a déjà prouvé qu'il s'agissait d'une forme bilinéaire symétrique.

Soit $(x_1, \dots, x_n) \in \mathbf{R}^n$. Alors

$$\varphi((x_1, \dots, x_n), (x_1, \dots, x_n)) = \sum_{i=1}^n x_i^2 \geq 0.$$

De plus, une somme de nombres positifs est nulle si et seulement si tous ces nombres sont nuls, donc

$$\begin{aligned}\varphi((x_1, \dots, x_n), (x_1, \dots, x_n)) = 0 &\Leftrightarrow \sum_{i=1}^n x_i^2 = 0 \Leftrightarrow \forall i \in \llbracket 1, n \rrbracket, x_i^2 = 0 \\ &\Leftrightarrow \forall i \in \llbracket 1, n \rrbracket, x_i = 0 \\ &\Leftrightarrow (x_1, \dots, x_n) = (0, \dots, 0).\end{aligned}$$

Donc φ est un produit scalaire, appelé **produit scalaire canonique sur \mathbf{R}^n** .

• $\langle f, g \rangle = \int_0^1 f(t)g(t) dt$ est un produit scalaire sur $\mathcal{C}^0([0, 1], \mathbf{R})$ car :

- nous avons déjà prouvé sa bilinéarité.
- $\forall f \in \mathcal{C}^0([0, 1], \mathbf{R}), \int_0^1 f^2(t) dt \geq 0$ (positivité de l'intégrale).
- si $\int_0^1 f^2(t) dt = 0$, alors puisque f^2 est continue et positive sur $[0, 1]$, c'est

Cas particuliers

Pour $n = 2$ ou $n = 3$, on retrouve les produits scalaires du plan et de l'espace étudiés au lycée.

que $\forall t \in [0, 1], f^2(t) = 0$, et donc $\forall t \in [0, 1], f(t) = 0$: f est la fonction nulle sur $[0, 1]$.

9.1.3 Inégalité de Cauchy-Schwarz

Théorème 9.9 (Inégalité de Cauchy-Schwarz) : Soit φ un produit scalaire sur E . Alors,

$$\forall (x, y) \in E^2, |\varphi(x, y)|^2 \leq \varphi(x, x) \times \varphi(y, y)$$

De manière équivalente, $|\varphi(x, y)| \leq \sqrt{\varphi(x, x)}\sqrt{\varphi(y, y)}$.

De plus, on a l'inégalité ci-dessus est une égalité si et seulement si x et y sont colinéaires.

Démonstration. Pour $x = 0_E$ ou $y = 0_E$, c'est évident, donc nous supposons $x \neq 0_E$ et $y \neq 0_E$. Définissons une fonction f sur \mathbf{R} par $f(t) = \varphi(tx + y, tx + y)$. Alors $\forall t \in \mathbf{R}, f(t) \geq 0$ car φ est positive. De plus, pour $t \in \mathbf{R}$, on a, par bilinéarité de φ ,

$$f(t) = \varphi(tx + y, tx + y) = \varphi(tx, tx) + 2\varphi(tx, y) + \varphi(y, y) = t^2\varphi(x, x) + 2t\varphi(x, y) + \varphi(y, y)$$

et donc f est un polynôme du second degré en la variable t . Puisqu'elle est toujours positive, c'est que son discriminant est négatif ou nul. Mais

$$\Delta = 4\varphi(x, y)^2 - 4\varphi(x, x)\varphi(y, y).$$

On en déduit que

$$\varphi(x, y)^2 \leq \varphi(x, x)\varphi(y, y).$$

D'autre part, si on a égalité alors $\Delta = 0$, et donc f possède alors une racine a . Mais alors $\varphi(ax + y, ax + y) = f(ax + y) = 0$, et donc $ax + y = 0_E$: x et y sont colinéaires.

Inversement, si x et y sont colinéaires, alors soit $x = 0_E$ (auquel cas l'inégalité est bien une égalité : $0 = 0$), soit il existe $\lambda \in \mathbf{R}$ tel que $y = \lambda x$. Mais alors

$$\varphi(x, y)^2 = \varphi(x, \lambda x)^2 = (\lambda\varphi(x, x))^2 = \lambda^2\varphi(x, x)^2 = \varphi(x, x)(\lambda^2\varphi(x, x)) = \varphi(x, x)\varphi(y, y).$$

L'inégalité de Cauchy-Schwarz est donc bien une égalité. □

Corollaire 9.10 (Inégalité de Cauchy-Schwarz dans \mathbf{R}^n) – Pour tous n -uplets $(x_1, \dots, x_n), (y_1, \dots, y_n)$ de réels, on a

$$\left| \sum_{i=1}^n x_i y_i \right| \leq \sqrt{\sum_{i=1}^n x_i^2} \sqrt{\sum_{i=1}^n y_i^2}.$$

Démonstration. Il s'agit simplement de l'inégalité de Cauchy-Schwarz pour le produit scalaire canonique de \mathbf{R}^n . □

Corollaire 9.11 (Inégalité de Cauchy-Schwarz pour les intégrales) – Si f et g sont des fonctions continues sur $[a, b], a < b$, alors

$$\left| \int_a^b f(t)g(t) dt \right| \leq \sqrt{\int_a^b f^2(t) dt} \sqrt{\int_a^b g^2(t) dt}.$$

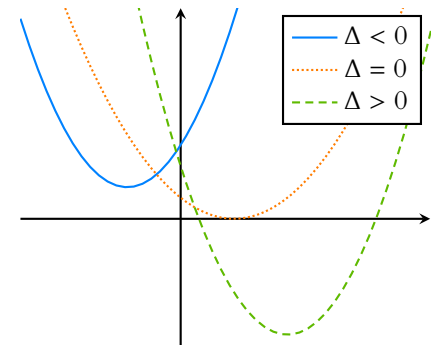


FIGURE 9.1– Un polynôme de degré 2 de signe constant est de discriminant négatif.

Remarque : cette preuve est la même que celle montrant que

$$\|\text{Cov}(X, Y)\| \leq \sqrt{V(X)}\sqrt{V(Y)}.$$

C'est normal, car on n'utilise pas ici

$$\varphi(x, x) = 0 \Leftrightarrow x = 0$$

et la covariance vérifie toutes les autres propriétés d'un produit scalaire.

9.1.4 Norme associée à un produit scalaire

Définition 9.12 – Si φ est un produit scalaire sur E , on appelle norme associée à φ l'application

$$\|\cdot\| : \begin{cases} E & \longrightarrow \mathbf{R}^+ \\ x & \longmapsto \sqrt{\varphi(x, x)} \end{cases}$$

En particulier, $\|x\|^2 = \varphi(x, x)$.

On appelle alors norme du vecteur $x \in E$ le **réel positif** $\|x\|$.

Remarque. En particulier, l'inégalité de Cauchy-Schwarz se reformule en termes de normes de la façon suivante :

$$\forall (x, y) \in E^2, |\varphi(x, y)| \leq \|x\| \cdot \|y\|.$$

Proposition 9.13 : Si φ est un produit scalaire sur E , et si $\|\cdot\|$ est la norme associée, alors

- $\forall \lambda \in \mathbf{R}, \forall x \in E, \|\lambda x\| = |\lambda| \cdot \|x\|$
- $\forall x \in E, \|x\| = 0 \Leftrightarrow x = 0_E$
- $\forall (x, y) \in E^2, \|x + y\| \leq \|x\| + \|y\|$. Cette inégalité s'appelle l'inégalité triangulaire¹.

Démonstration. • $\|\lambda x\| = \sqrt{\varphi(\lambda x, \lambda x)} = \sqrt{\lambda^2 \varphi(x, x)} = |\lambda| \sqrt{\varphi(x, x)} = |\lambda| \cdot \|x\|$.

- Il est clair que $\|0_E\| = 0$.
Inversement, si $\|x\| = 0$, alors $\varphi(x, x) = \|x\|^2 = 0$, et donc $x = 0_E$ car φ est définie.
- Soient $x, y \in E$. Alors

$$\begin{aligned} \|x + y\|^2 &= \varphi(x + y, x + y) = \varphi(x, x) + 2\varphi(x, y) + \varphi(y, y) \\ &= \|x\|^2 + 2\varphi(x, y) + \|y\|^2 \\ &\leq \|x\|^2 + 2\|x\| \cdot \|y\| + \|y\|^2 \\ &\leq (\|x\| + \|y\|)^2. \end{aligned}$$

En prenant la racine carrée, il vient $\|x + y\| \leq \|x\| + \|y\|$.

□

Exemple 9.14

La norme associée au produit scalaire canonique de \mathbf{R}^n est

$$\|(x_1, \dots, x_n)\| = \sqrt{\sum_{i=1}^n x_i^2}.$$

Pour $n = 2$ ou $n = 3$, on retrouve les normes d'un vecteur du plan ou de l'espace étudiées au lycée, où la norme d'un vecteur était définie comme étant sa «longueur».

Définition 9.15 – Un vecteur x est dit unitaire si $\|x\| = 1$.

Si x est non nul, alors $\frac{x}{\|x\|} = \frac{1}{\|x\|}x$ est unitaire. En effet,

$$\left\| \frac{1}{\|x\|}x \right\| = \frac{1}{\|x\|} \|x\| = 1.$$

9.2 ORTHOGONALITÉ

Dans toute cette partie, on considère un espace vectoriel E muni d'un produit scalaire que l'on notera $\langle x, y \rangle$ plutôt que $\varphi(x, y)$.

On a donc pour tout $x \in E, \|x\| = \sqrt{\langle x, x \rangle}$.

Remarque

Dans le cas du produit scalaire canonique de \mathbf{R}^2 , étudié au lycée, cette propriété était déjà connue car :

$$\vec{x} \cdot \vec{y} = \|\vec{x}\| \cdot \|\vec{y}\| \cos(\vec{x}, \vec{y}).$$

et puisque $|\cos(\vec{x}, \vec{y})| \leq 1$

$$|\vec{x} \cdot \vec{y}| \leq \|\vec{x}\| \cdot \|\vec{y}\|.$$

¹ En particulier, pour $\mathbf{R} = \mathbf{R}^1$ muni de son produit scalaire canonique, on retrouve l'inégalité triangulaire que l'on connaît bien pour des réels.

C'est l'inégalité de Cauchy-Schwarz.

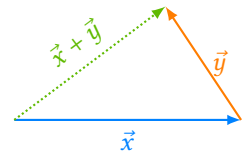


FIGURE 9.2– L'inégalité triangulaire : le plus court chemin entre deux points est la ligne droite.

Astuce

Cette remarque permet facilement, pour tout vecteur non nul x de fabriquer un vecteur de norme 1 colinéaire à x : il suffit de le diviser par sa norme.

9.2.1 Définition

Définition 9.16 – Deux vecteurs x et y de E sont dits orthogonaux si $\langle x, y \rangle = 0$.

Exemples 9.17

- Le vecteur nul est orthogonal à tout vecteur de E .
- Dans \mathbf{R}^3 muni du produit scalaire canonique, $(1, 0, 1)$ et $(-1, -5, 1)$ sont orthogonaux car $1 \times (-1) + 0 \times (-5) + 1 \times 1 = 0$.
En revanche, $(1, 0, 0)$ et $(2, 0, 1)$ ne sont pas orthogonaux.
- Le seul vecteur orthogonal à lui-même est le vecteur nul.
En effet, on a $\langle x, x \rangle = 0 \Leftrightarrow x = 0_E$.
- Si (x, y) est une famille liée², alors x et y sont orthogonaux si et seulement $x = 0_E$ ou $y = 0_E$.

² C'est-à-dire si x et y sont colinéaires.

Définition 9.18 – Soient F et G deux sous-espaces vectoriels de E .
On dit que F et G sont orthogonaux si

$$\forall x \in F, \forall y \in G, \langle x, y \rangle = 0.$$

De manière équivalente, ceci signifie que **tout** vecteur de F est orthogonal à **tout** vecteur de G .

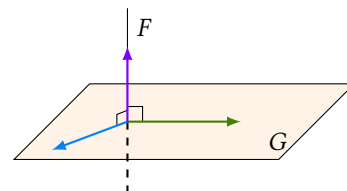


FIGURE 9.3– Tout vecteur de F est orthogonal à tout vecteur de G .

Exemple 9.19

Dans \mathbf{R}^3 , soit $F = \{(x, y, z) \in \mathbf{R}^3 : x + y = 0\}$ et $G = \{(x, y, z) \in \mathbf{R}^3 : x = y \text{ et } z = 0\}$.
Alors F et G sont orthogonaux. En effet, soit $(x, y, z) \in F : (x, y, z) = (x, -x, z)$.
Soit $(x', y', z') \in G$, alors $(x', y', z') = (x', x', 0)$.
Et alors $\langle (x, y, z), (x', y', z') \rangle = xx' + yy' + zz' = xx' - xx' = 0$.

Remarque. x et y sont deux vecteurs orthogonaux si et seulement si $\text{Vect}(x)$ et $\text{Vect}(y)$ sont deux sous-espaces orthogonaux.

Proposition 9.20 : Si F et G sont deux sous-espaces vectoriels orthogonaux, alors $F \cap G = \{0_E\}$. Par conséquent, F et G sont en somme directe.

Démonstration. Soit $x \in F \cap G$. Alors on a à la fois $x \in F$ et $x \in G$, de sorte que $\langle \underbrace{x}_{\in F}, \underbrace{x}_{\in G} \rangle = 0$, et donc $x = 0_E$.

Ainsi, on a bien $F \cap G = \{0_E\}$. □

⚠ Pour autant, deux sous-espaces orthogonaux ne sont pas forcément supplémentaires dans E . Par exemple, dans \mathbf{R}^3 , $F = \text{Vect}(1, 0, 0)$ et $G = \text{Vect}(0, 1, 0)$ sont orthogonaux, mais $\dim F + \dim G = 2 \neq \dim \mathbf{R}^3$.

Définition 9.21 – Une famille (x_1, \dots, x_n) de vecteurs de E est dite orthogonale si

$$\forall i \neq j, x_i \text{ et } x_j \text{ sont orthogonaux.}$$

On dit qu'une famille (x_1, \dots, x_n) est orthonormée si elle est orthogonale, et formée de vecteurs unitaires, c'est-à-dire si $\forall i \in \llbracket 1, n \rrbracket, \|x_i\| = 1$.

Autrement dit, (x_1, \dots, x_n) est orthonormée si et seulement si

$$\forall (i, j) \in \llbracket 1, n \rrbracket^2, \langle x_i, x_j \rangle = \begin{cases} 1 & \text{si } i = j \\ 0 & \text{si } i \neq j \end{cases}$$

Remarque. À partir d'une famille orthogonale (x_1, \dots, x_n) ne contenant pas le vecteur nul, il est facile d'obtenir une famille orthonormée : il suffit de prendre $\left(\frac{x_1}{\|x_1\|}, \dots, \frac{x_n}{\|x_n\|}\right)$.

Proposition 9.22 : Une famille orthogonale (x_1, \dots, x_n) ne contenant pas le vecteur nul est libre.

Démonstration. Soit (x_1, \dots, x_n) une famille orthogonale dont tous les vecteurs sont non nuls.

Soient $\lambda_1, \dots, \lambda_n \in \mathbf{R}$ tels que $\sum_{i=1}^n \lambda_i x_i = 0$. Alors,

$$\forall j \in \llbracket 1, n \rrbracket, 0 = \left\langle \sum_{i=1}^n \lambda_i x_i, x_j \right\rangle = \sum_{i=1}^n \lambda_i \langle x_i, x_j \rangle = \lambda_j \langle x_j, x_j \rangle = \lambda_j \|x_j\|^2.$$

Mais $x_j \neq 0_E$, et donc $\|x_j\| \neq 0$. On en déduit que $\lambda_j = 0$. Ceci étant vrai pour tout $j \in \llbracket 1, n \rrbracket$, la famille (x_1, \dots, x_n) est libre. □

9.2.2 Théorème de Pythagore

Proposition 9.23 : Soient x, y deux vecteurs de E . Alors x et y sont orthogonaux si et seulement si $\|x + y\|^2 = \|x\|^2 + \|y\|^2$.

Démonstration. On a

$$\|x + y\|^2 = \langle x + y, x + y \rangle = \langle x, x \rangle + 2\langle x, y \rangle + \langle y, y \rangle = \|x\|^2 + \|y\|^2 + 2\langle x, y \rangle.$$

Et donc $\|x + y\|^2 = \|x\|^2 + \|y\|^2$ si et seulement si $\langle x, y \rangle = 0$. □

Remarques. • Bien entendu, dans le cas de \mathbf{R}^2 muni du produit scalaire canonique, on retrouve le théorème de Pythagore appris à la maternelle.

• Puisque l'énoncé est une équivalence, on a à la fois le théorème de Pythagore usuel, et sa réciproque.

Corollaire 9.24 – Soit (x_1, \dots, x_n) une famille orthogonale de vecteurs de E . Alors

$$\left\| \sum_{i=1}^n x_i \right\|^2 = \sum_{i=1}^n \|x_i\|^2.$$

Démonstration. Utilisons la bilinéarité du produit scalaire :

$$\left\| \sum_{i=1}^n x_i \right\|^2 = \left\langle \sum_{i=1}^n x_i, \sum_{j=1}^n x_j \right\rangle$$

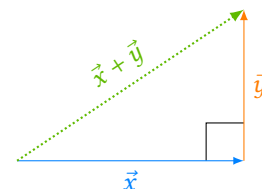


FIGURE 9.4– Théorème de Pythagore : pas de surprise...

Intuition

En divisant un vecteur par sa norme, on modifie sa norme, mais pas sa direction. Or le fait que deux vecteurs soient ou non orthogonaux ne dépend que de leurs directions, et pas de leurs normes.

$$\begin{aligned}
 &= \sum_{i=1}^n \left\langle x_i, \sum_{j=1}^n x_j \right\rangle && \text{Linéarité à gauche.} \\
 &= \sum_{i=1}^n \sum_{j=1}^n \langle x_i, x_j \rangle && \text{Linéarité à droite.} \\
 &= \sum_{i=1}^n \langle x_i, x_i \rangle && \text{Si } i \neq j, \text{ alors } \langle x_i, x_j \rangle = 0. \\
 &= \sum_{i=1}^n \|x_i\|^2.
 \end{aligned}$$

□

9.2.3 Procédé d'orthonormalisation de Gram-Schmidt

Nous décrivons ici un algorithme permettant de construire une famille orthonormée de vecteurs à partir d'une famille libre.

Proposition 9.25 : Soit (e_1, \dots, e_n) une famille libre de vecteurs de E .

Définissons par récurrence une famille (x_1, \dots, x_n) en posant $x_1 = \frac{e_1}{\|e_1\|}$ et

$$\forall k \in \llbracket 2, n \rrbracket, x_k = \frac{1}{\|x_k^*\|} x_k^* \text{ où } x_k^* = e_k - \sum_{i=1}^{k-1} \langle e_k, x_i \rangle x_i.$$

Alors :

- la famille (x_1, \dots, x_n) est orthonormée
- $\forall k \in \llbracket 1, n \rrbracket, \text{Vect}(e_1, \dots, e_k) = \text{Vect}(x_1, \dots, x_k)$.

Explications

Nous réinterpréterons cette formule dans le chapitre 12, et son origine sera alors plus claire.



Démonstration. Montrons par récurrence sur k que la famille (x_1, \dots, x_k) est orthonormée et que $\text{Vect}(e_1, \dots, e_k) = \text{Vect}(x_1, \dots, x_k)$.

Pour $k = 1$, e_1 est non nul³ et donc $\frac{e_1}{\|e_1\|}$ est de norme 1.

Supposons que (x_1, \dots, x_k) soit une famille orthonormée et que $\text{Vect}(e_1, \dots, e_k) = \text{Vect}(x_1, \dots, x_k)$. Alors, pour tout $i \in \llbracket 1, k \rrbracket$, on a

$$\begin{aligned}
 \langle x_{k+1}^*, x_i \rangle &= \left\langle e_{k+1} - \sum_{j=1}^k \langle e_{k+1}, x_j \rangle x_j, x_i \right\rangle \\
 &= \langle e_{k+1}, x_i \rangle - \sum_{j=1}^k \langle \langle e_{k+1}, x_j \rangle x_j, x_i \rangle \\
 &= \langle e_{k+1}, x_i \rangle - \sum_{j=1}^k \langle e_{k+1}, x_j \rangle \langle x_j, x_i \rangle \\
 &= \langle e_{k+1}, x_i \rangle - \langle e_{k+1}, x_i \rangle \\
 &= 0.
 \end{aligned}$$

Ainsi, x_{k+1}^* est orthogonal à tous les $x_i, i \leq k$, et donc il en est de même de x_{k+1} .

Notons que x_{k+1}^* est non nul⁴, car si on avait $x_{k+1}^* = 0_E$, alors

$$e_{k+1} = \sum_{i=1}^k \langle e_{k+1}, x_i \rangle x_i \in \text{Vect}(x_1, \dots, x_k) = \text{Vect}(e_1, \dots, e_k),$$

contredisant le fait que (e_1, \dots, e_{k+1}) est libre.

Comme de plus les $x_i, i \in \llbracket 1, k+1 \rrbracket$ sont tous de norme 1 par construction, (x_1, \dots, x_{k+1}) est une famille orthonormée de E .

Il est clair que $x_{k+1} \in \text{Vect}(x_1, \dots, x_k, e_{k+1}) = \text{Vect}(e_1, \dots, e_k, e_{k+1})$ et donc, $\text{Vect}(x_1, \dots, x_{k+1}) \subset \text{Vect}(e_1, \dots, e_{k+1})$.

³ Car la famille (e_1, \dots, e_n) est libre et ne contient donc pas le vecteur nul.

Détails

$\langle e_{k+1}, x_j \rangle$ est un réel, donc on peut le «sortir» du produit scalaire par linéarité à gauche.

$$\langle x_i, x_j \rangle = \begin{cases} 1 & \text{si } i = j \\ 0 & \text{sinon} \end{cases}$$

⁴ Et donc il est légitime de le diviser par sa norme car celle-ci est non nulle.

Conséquence

Puisque $x_k^* \neq 0_E$, alors x_k est bien défini.

Mais $\text{Vect}(e_1, \dots, e_{k+1})$ est de dimension $k + 1$, et x_1, \dots, x_{k+1} est libre⁵ et génératrice de $\text{Vect}(x_1, \dots, x_{k+1})$: c'est une base de $\text{Vect}(x_1, \dots, x_{k+1})$.
On en déduit que $\dim \text{Vect}(x_1, \dots, x_{k+1}) = \dim \text{Vect}(e_1, \dots, e_{k+1})$ et donc que ces deux espaces sont égaux. \square

⁵ Car orthonormée.

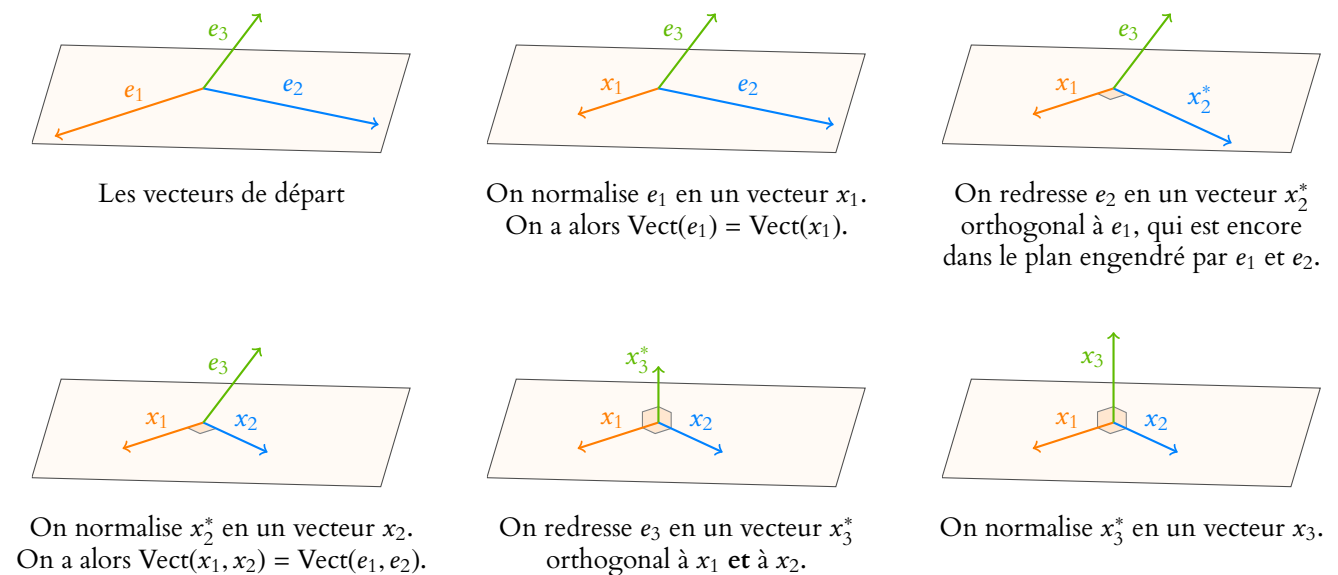


Figure 9.5– Un exemple d'orthonormalisation de Gram-Schmidt sur une famille de 3 vecteurs de \mathbf{R}^3 .

Remarques. • En particulier, lorsque la famille (e_1, \dots, e_n) est une base de E , alors la famille (x_1, \dots, x_n) est une base orthonormée de E , car elle est libre⁶ et génératrice grâce au second point de la proposition.

• Si (e_1, \dots, e_n) est déjà orthogonale, alors $(x_1, \dots, x_n) = \left(\frac{e_1}{\|e_1\|}, \dots, \frac{e_n}{\|e_n\|} \right)$.

Et si (e_1, \dots, e_n) est orthonormée, alors $(x_1, \dots, x_n) = (e_1, \dots, e_n)$.

⁶ Car orthogonale et ne contenant pas le vecteur nul (puisque formée de vecteurs unitaires).

Autrement dit

Si on applique Gram-Schmidt à une famille orthonormée, alors celle-ci ne change pas.

Exemple 9.26

Dans $\mathbf{R}_2[X]$ muni du produit scalaire $\langle P, Q \rangle = \int_0^1 P(t)Q(t) dt$, appliquons le procédé d'orthonormalisation de Gram-Schmidt à la base canonique $(1, X, X^2)$.

Notons donc $P_1 = 1, P_2 = X$ et $P_3 = X^2$, et construisons une famille orthonormée (Q_1, Q_2, Q_3) à partir de (P_1, P_2, P_3) .

On a $\|P_1\|^2 = \int_0^1 1 dt = 1$ et donc $\|P_1\| = 1$. Soit donc $Q_1 = P_1 = 1$. Posons ensuite

$$Q_2^* = X - \langle X, 1 \rangle 1 = X - \int_0^1 t dt = X - \frac{1}{2}.$$

$$\text{Alors } \|Q_2^*\|^2 = \int_0^1 \left(t - \frac{1}{2} \right)^2 dt = \frac{1}{12}.$$

Posons donc $Q_2 = \frac{1}{\|Q_2^*\|} Q_2^* = \sqrt{12} \left(X - \frac{1}{2} \right)$. Soit ensuite Q_3^* défini par

$$\begin{aligned} Q_3^* &= X^2 - \langle X^2, Q_1 \rangle Q_1 - \langle X^2, Q_2 \rangle Q_2 \\ &= X^2 - \int_0^1 t^2 dt - \sqrt{12} \left(\int_0^1 t^2 \left(t - \frac{1}{2} \right) dt \right) Q_2 \\ &= X^2 - \frac{1}{3} - \left(X - \frac{1}{2} \right) = X^2 - X + \frac{1}{6} \end{aligned}$$

On a alors

$$\begin{aligned}\|Q_3^*\|^2 &= \int_0^1 \left(t^2 - t + \frac{1}{6}\right)^2 dt \\ &= \int_0^1 \left(t^4 + t^2 + \frac{1}{36} + \frac{t^2}{3} - \frac{t}{3} - 2t^3\right) dt \\ &= \frac{1}{5} + \frac{1}{3} + \frac{1}{36} + \frac{1}{9} - \frac{1}{6} - \frac{1}{2} = \frac{1}{180}.\end{aligned}$$

Posons alors $Q_3 = \frac{Q_3^*}{\|Q_3^*\|} = \sqrt{180}Q_3^* = \sqrt{180}\left(X^2 - X + \frac{1}{6}\right)$.

La famille (Q_1, Q_2, Q_3) est alors une famille orthonormée de $\mathbf{R}_2[X]$, donc libre. Elle est de cardinal $3 = \dim \mathbf{R}_2[X]$: il s'agit d'une base orthonormée de $\mathbf{R}_2[X]$.



Astuce : notons qu'il est possible de gagner un peu de temps dans les calculs, en remarquant

que $x_i^* = e_i - \sum_{k=1}^{i-1} \langle e_i, x_k \rangle x_k$, s'écrit encore $e_i = x_i^* + \sum_{k=1}^{i-1} \langle e_i, x_k \rangle x_k$ où x_i^* et les x_k sont deux à deux orthogonaux. Donc d'après le théorème de Pythagore,

$$\|e_i\|^2 = \|x_i^*\|^2 + \sum_{k=1}^{i-1} \langle e_i, x_k \rangle^2 \underbrace{\|x_k\|^2}_{=1} \Leftrightarrow \|x_i^*\|^2 = \|e_i\|^2 - \sum_{k=1}^{i-1} \langle e_i, x_k \rangle^2.$$

Ainsi, dans l'exemple précédent, on a avait

$$\|Q_3^*\|^2 = \|X^2\|^2 - \langle X^2, Q_1 \rangle^2 - \langle X^2, Q_2 \rangle^2 = \frac{1}{5} - \frac{1}{9} - \frac{1}{12} = \frac{1}{180}.$$

Et donc le calcul de la norme de Q_3^* ne nous a pas demandé de nouveau calcul d'intégrale (hormis celle définissant $\|P_3\|^2$, qui était somme toute assez simple à calculer).

9.3 ESPACES EUCLIDIENS

9.3.1 Définition

Définition 9.27 – On appelle espace euclidien un espace vectoriel E de dimension finie sur \mathbf{R} muni d'un produit scalaire $\langle \cdot, \cdot \rangle$.

Exemples 9.28

\mathbf{R}^n muni de son produit scalaire usuel est un espace euclidien.

De même, $\mathbf{R}_n[X]$ muni du produit scalaire $\langle P, Q \rangle = \int_0^1 P(t)Q(t) dt$ est un espace euclidien.

Définition 9.29 – Une famille (e_1, \dots, e_n) est une base orthonormée d'un espace euclidien E si c'est une base et qu'il s'agit d'une famille orthonormée. Autrement dit, c'est une base telle que

$$\forall (i, j) \in \llbracket 1, n \rrbracket^2, \langle e_i, e_j \rangle = \begin{cases} 0 & \text{si } i \neq j \\ 1 & \text{si } i = j \end{cases}$$

Remarque. Une famille orthonormée de cardinal $\dim E$ est automatiquement une base orthonormée, puisqu'elle est libre⁷ et de cardinal $\dim E$.

Remarque

Dans \mathbf{R}^2 muni de son produit scalaire canonique, une base orthonormée est ce qu'on appelait au lycée un repère orthonormé du plan.

⁷ Car orthogonale, et ne contenant pas le vecteur nul car formée de vecteurs unitaires.

Exemple 9.30

La base canonique de \mathbf{R}^n est une base orthonormée de \mathbf{R}^n muni de son produit scalaire canonique.

Proposition 9.31 : Si E est un espace euclidien, alors il existe une base orthonormée de E .

Démonstration. Soit (e_1, \dots, e_n) une base de E , et soit (x_1, \dots, x_n) la famille obtenue en appliquant le procédé d'orthonormalisation de Gram-Schmidt à (e_1, \dots, e_n) .

Il s'agit d'une famille orthonormée, de cardinal $n = \dim E$: c'est donc une base⁸ de E . \square

Remarque. Ce théorème ne dit pas comment trouver une telle base (même si la preuve nous éclaire un peu : il suffit d'appliquer Gram-Schmidt à une base quelconque... à condition d'en connaître une !).

9.3.2 Calculs dans une base orthonormée

Dans la suite, notons E un espace euclidien muni d'une base **orthonormée** $\mathcal{B} = (e_1, \dots, e_n)$.

Dans une base orthonormée, il est aisé d'obtenir les coordonnées d'un vecteur, sans avoir besoin de résoudre un système⁹.

Proposition 9.32 (Coordonnées d'un vecteur dans une base orthonormée) :

Pour tout vecteur $x \in E$, on a

$$x = \sum_{i=1}^n \langle x, e_i \rangle e_i.$$

Autrement dit, les coordonnées de x dans la base \mathcal{B} sont données par $x_{\mathcal{B}} = \begin{pmatrix} \langle x, e_1 \rangle \\ \vdots \\ \langle x, e_n \rangle \end{pmatrix}$.

Démonstration. Soit $x \in E$. Puisque \mathcal{B} est une base, il existe des scalaires uniquement

déterminés x_1, \dots, x_n tels que $x = \sum_{i=1}^n x_i e_i$.

Mais alors

$$\forall i \in \llbracket 1, n \rrbracket, \langle x, e_i \rangle = \left\langle \sum_{j=1}^n x_j e_j, e_i \right\rangle = \sum_{j=1}^n x_j \underbrace{\langle e_j, e_i \rangle}_{=0 \text{ si } i \neq j} = x_i \langle e_i, e_i \rangle = x_i.$$

Et donc $x = \sum_{i=1}^n \langle x, e_i \rangle e_i$. \square

Exemple 9.33

- Dans \mathbf{R}^n muni de sa structure euclidienne usuelle, et si $\mathcal{B} = (e_1, \dots, e_n)$ est la base canonique c'est évident car $e_i = (0, \dots, 0, 1, 0, \dots, 0)$ et donc

$$\forall x = (x_1, \dots, x_n) \in \mathbf{R}^n, \langle x, e_i \rangle = x_i$$

Terminologie

Il n'existe pas qu'une seule base orthonormée de E , et on dira bien **une** base orthonormée, et non **la** base orthonormée de E .

⁸ Et donc une base orthonormée.

⁹ Ni d'utiliser une matrice de passage, ce qui souvent nécessite d'inverser une autre matrice de passage, ce qui demande autant de calculs que la résolution d'un système.

Proposition 9.34 (Calcul de la norme à l'aide d'une base orthonormée) :

$$\forall x \in E, \|x\|^2 = \sum_{i=1}^n \langle x, e_i \rangle^2.$$

Démonstration. Soit $x \in E$. Par la proposition précédente, $x = \sum_{i=1}^n x_i e_i$, avec $x_i = \langle x, e_i \rangle$.

Mais alors

$$\|x\|^2 = \left\langle \sum_{i=1}^n x_i e_i, \sum_{j=1}^n x_j e_j \right\rangle = \sum_{i=1}^n \sum_{j=1}^n x_i x_j \langle e_i, e_j \rangle = \sum_{i=1}^n x_i^2 = \sum_{i=1}^n \langle x, e_i \rangle^2.$$

□

Proposition 9.35 : L'application $\varphi : (X, Y) \mapsto {}^tXY$ est un produit scalaire sur $\mathcal{M}_{n,1}(\mathbf{R})$. On l'appelle produit scalaire canonique de $\mathcal{M}_{n,1}(\mathbf{R})$.

Démonstration. La bilinéarité est triviale.

Si $X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ et $Y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$ sont deux éléments de $\mathcal{M}_{n,1}(\mathbf{R})$, alors

$${}^tXY = (x_1 \quad \dots \quad x_n) \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = \sum_{i=1}^n x_i y_i.$$

En particulier, on a ${}^tYX = \sum_{i=1}^n y_i x_i = {}^tXY$, donc φ est symétrique.

Si $X \in \mathcal{M}_{n,1}(\mathbf{R})$, alors $\varphi(X, X) = \sum_{i=1}^n x_i^2 \geq 0$, et cette quantité est nulle si et seulement si tous les x_i sont nuls, c'est-à-dire si et seulement si $X = 0$. □

Remarque. Notons que ce produit scalaire ressemble beaucoup au produit scalaire canonique

de \mathbf{R}^n : le produit scalaire de $\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ par $\begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$ vaut $\sum_{i=1}^n x_i y_i$.

Cela n'a rien de surprenant si l'on se rappelle que \mathbf{R}^n et $\mathcal{M}_{n,1}(\mathbf{R})$ sont très semblables : un vecteur de \mathbf{R}^n est de la forme (x_1, \dots, x_n) alors qu'un vecteur de $\mathcal{M}_{n,1}(\mathbf{R})$ est de la forme

$$\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}.$$

Rappelons que si $x = \sum_{i=1}^n x_i e_i \in E$, on note $x_{\mathcal{B}} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ le vecteur des coordonnées de x dans la base \mathcal{B} .

Proposition 9.36 : Soient $x, y \in E$, et soit $x_{\mathcal{B}}$ (resp. $y_{\mathcal{B}}$) le vecteur colonne des coordonnées de x (resp. y) dans la base \mathcal{B} . Alors :

- $\langle x, y \rangle = {}^t x_{\mathcal{B}} y_{\mathcal{B}}$
- $\|x\|^2 = {}^t x_{\mathcal{B}} x_{\mathcal{B}}$.

⚠ Attention !

Ceci ne vaut que si l'on travaille dans une base orthonormée !

Encore une fois, dans \mathbf{R}^2 ou \mathbf{R}^3 , muni du produit scalaire usuel, on retrouve les classiques définitions du lycée :

$$\|(x, y)\| = \sqrt{x^2 + y^2}$$

et

$$\|(x, y, z)\| = \sqrt{x^2 + y^2 + z^2}.$$

Autrement dit

Pour le dire de manière plus pompeuse, l'application qui à un vecteur $x \in \mathbf{R}^n$ associe sa transposée est un isomorphisme entre \mathbf{R}^n et $\mathcal{M}_{n,1}(\mathbf{R})$.

Remarque

Le produit d'un vecteur ligne par un vecteur colonne est une matrice 1×1 , c'est-à-dire un réel.

Démonstration. Notons $x_{\mathcal{B}} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ et $y_{\mathcal{B}} = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$, de sorte que

$$x = \sum_{i=1}^n x_i e_i \text{ et } y = \sum_{i=1}^n y_i e_i.$$

Alors

$${}^t x_{\mathcal{B}} y_{\mathcal{B}} = (x_1 \quad \dots \quad x_n) \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = \sum_{i=1}^n x_i y_i$$

Mais

$$\langle x, y \rangle = \left\langle \sum_{i=1}^n x_i e_i, \sum_{j=1}^n y_j e_j \right\rangle = \sum_{i=1}^n \sum_{j=1}^n x_i y_j \langle e_i, e_j \rangle = \sum_{i=1}^n x_i y_i.$$

Donc $\langle x, y \rangle = {}^t x_{\mathcal{B}} y_{\mathcal{B}}$.

En particulier, pour $y = x$, on obtient

$$\|x\|^2 = \langle x, x \rangle = {}^t x_{\mathcal{B}} x_{\mathcal{B}}.$$

□

Cette proposition explique l'intérêt des bases orthonormées : si on travaille dans une base orthonormée, tous les calculs de produits scalaires et de normes se font comme dans $\mathcal{M}_{n,1}(\mathbf{R})$ muni de son produit scalaire canonique (et donc comme dans \mathbf{R}^n muni de son produit scalaire canonique) : $\langle x, y \rangle = \sum_{i=1}^n x_i y_i$ où $x = \sum_{i=1}^n x_i e_i$ et $y = \sum_{i=1}^n y_i e_i$.

Définition 9.37 – Une matrice $M \in \mathcal{M}_n(\mathbf{R})$ est dite orthogonale si ${}^t M M = I_n$, c'est-à-dire si M est inversible, et que son inverse est égale à sa transposée.

Exemples 9.38

- I_n est orthogonale car ${}^t I_n I_n = I_n I_n = I_n$.
- Soit $D = \text{Diag}(\lambda_1, \dots, \lambda_n)$ une matrice diagonale. Alors

$${}^t D D = D^2 = \text{Diag}(\lambda_1^2, \dots, \lambda_n^2).$$

Donc D est orthogonale si et seulement si $D^2 = I_n$, soit si et seulement si $\forall i \in \llbracket 1, n \rrbracket, \lambda_i^2 = 1$.

Ainsi, une matrice diagonale est orthogonale si et seulement si tous ses coefficients diagonaux sont égaux à 1 ou à -1 .

- $A = \begin{pmatrix} \frac{4}{5} & \frac{3}{5} \\ -\frac{3}{5} & \frac{4}{5} \end{pmatrix}$ est orthogonale. En effet,

$${}^t A A = \begin{pmatrix} \frac{4}{5} & -\frac{3}{5} \\ \frac{3}{5} & \frac{4}{5} \end{pmatrix} \begin{pmatrix} \frac{4}{5} & \frac{3}{5} \\ -\frac{3}{5} & \frac{4}{5} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Remarque. Si A est une matrice orthogonale, alors A^{-1} est encore orthogonale, car $A^{-1} = {}^t A$, et donc ${}^t (A^{-1}) A^{-1} = A^t A = I_n$.

De même, si A, B sont deux matrices orthogonales de $\mathcal{M}_n(\mathbf{R})$, alors AB est encore orthogonale.

Détails

Une matrice diagonale est toujours symétrique, donc ${}^t D = D$.

Proposition 9.39 : Soit E un espace euclidien de dimension n , et soit $\mathcal{B} = (e_1, \dots, e_n)$ une base orthonormée de E .

Alors une base $\mathcal{B}' = (f_1, \dots, f_n)$ de E est orthonormée si et seulement si la matrice de passage $P_{\mathcal{B}, \mathcal{B}'}$ est orthogonale.

Démonstration. Notons Y_1, \dots, Y_n les colonnes de $P_{\mathcal{B}, \mathcal{B}'}$. Par définition d'une matrice de passage, Y_i est le vecteur colonne des coordonnées de f_i dans la base \mathcal{B} . Notons que le coefficient (i, j) de ${}^t P_{\mathcal{B}, \mathcal{B}'} P_{\mathcal{B}, \mathcal{B}'}$ est égal à ${}^t Y_i Y_j = \langle f_i, f_j \rangle$

Si \mathcal{B}' est orthonormée, alors ${}^t Y_i Y_j = \langle f_i, f_j \rangle = \begin{cases} 0 & \text{si } i \neq j \\ 1 & \text{sinon} \end{cases}$.

Et donc la matrice ${}^t P_{\mathcal{B}, \mathcal{B}'} P_{\mathcal{B}, \mathcal{B}'}$ a tous ses coefficients nuls, sauf les coefficients diagonaux qui valent 1 : c'est donc la matrice identité. Par conséquent $P_{\mathcal{B}, \mathcal{B}'}$ est une matrice orthogonale.

Inversement, si $P_{\mathcal{B}, \mathcal{B}'}$ est orthogonale, alors

$$\forall (i, j) \in \llbracket 1, n \rrbracket^2, \langle f_i, f_j \rangle = {}^t Y_i Y_j = \begin{cases} 1 & \text{si } i = j \\ 0 & \text{sinon} \end{cases}$$

Donc \mathcal{B}' est orthogonale. □

Remarques. • L'intérêt de ce résultat réside dans l'application de la formule de changement de base entre deux bases orthonormées \mathcal{B} et \mathcal{B}' : une fois que l'on connaît une matrice de passage ($P_{\mathcal{B}, \mathcal{B}'}$ ou $P_{\mathcal{B}', \mathcal{B}}$), pas besoin de calculer son inverse, il suffit de transposer pour l'obtenir.

• La même preuve montrerait qu'une matrice $M \in \mathcal{M}_n(\mathbf{R})$ est orthogonale si et seulement si la famille de ses vecteurs colonnes est une base orthonormée de $\mathcal{M}_{n,1}(\mathbf{R})$.

En effet, si M est une matrice inversible¹⁰, alors si on note (C_1, \dots, C_n) la famille de ses vecteurs colonnes, M est la matrice de passage de la base canonique de $\mathcal{M}_{n,1}(\mathbf{R})$ à la base (C_1, \dots, C_n) .

Et la base canonique de $\mathcal{M}_{n,1}(\mathbf{R})$ étant orthonormée¹¹, M est orthogonale si et seulement si (C_1, \dots, C_n) est orthonormée.

Détail

Ce coefficient s'obtient à l'aide la i -ème ligne de ${}^t P_{\mathcal{B}, \mathcal{B}'}$, et de la j -ème colonne de $P_{\mathcal{B}, \mathcal{B}'}$.



¹⁰ Une matrice non inversible ne saurait être orthogonale.

¹¹ Pour le produit scalaire canonique de $\mathcal{M}_{n,1}(\mathbf{R})$.

MÉTHODES DU CHAPITRE 9

✓ MÉTHODE 9.1 : Montrer qu'une application $\varphi : E \times E \rightarrow \mathbf{R}$ est un produit scalaire

IL faut pour cela se rappeler qu'il y a quatre points à prouver. Si on se contente de prouver que φ est symétrique et linéaire à gauche, alors on expliquera bien que cela suffit à la bilinéarité.

Le point le plus délicat est souvent $\varphi(x, x) = 0 \Rightarrow x = 0$. Les principaux arguments sont :

- une somme de nombres positifs est nulle si et seulement si chacun des nombres est nul
- l'intégrale sur $]a, b[$ d'une fonction continue positive est nulle ssi la fonction est nulle sur $]a, b[$
- un polynôme de degré n est nul ssi il possède au moins $n + 1$ racines distinctes.

Exercices de TD : 9.3, 9.4, 9.5, 9.9, 9.13, 9.15

✓ MÉTHODE 9.2 : Utiliser l'inégalité de Cauchy-Schwarz

Le plus dur est de penser à l'utiliser dans \mathbf{R}^n muni du produit scalaire canonique lorsque rien ne l'indique. Des inégalités avec des sommes et des carrés (ou des racines) doivent être un indice. On choisira alors soigneusement les vecteurs auxquels on l'applique.

Exercices de TD : 9.8, 9.11

✓ MÉTHODE 9.3 : Construire des bases orthonormées à l'aide de Gram-Schmidt

En l'absence d'indication, on partira de n'importe quelle base (dans les espaces vectoriels usuels, on utilisera de préférence la base canonique). Le reste est algorithmique, mais les calculs demandent souvent beaucoup de soin. On simplifiera autant que possible les racines carrées qui pourraient apparaître afin de ne pas alourdir les calculs.

Exercices de TD : 9.2, 9.12, 9.15

✓ MÉTHODE 9.4 : Montrer l'orthogonalité de deux sous-espaces vectoriels F et G .

Il s'agit de prouver que **tout** vecteur de F est orthogonal à **tout** vecteur de G . Pour cela on prendra $x \in F$ et $y \in G$ quelconques, et on prouvera que $\langle x, y \rangle = 0$.

Exercices de TD : 9.4, 9.15

EXERCICES DU CHAPITRE 9

EXERCICE 9.1 Identités de polarisation, ou comment calculer un produit scalaire à partir de normes

Soit E un espace vectoriel muni d'un produit scalaire $\langle \cdot, \cdot \rangle$. Montrer que pour tous $x, y \in E$, on a

$$\langle x, y \rangle = \frac{1}{2} (\|x + y\|^2 - \|x\|^2 - \|y\|^2) = \frac{1}{4} (\|x + y\|^2 - \|x - y\|^2).$$

F

EXERCICE 9.2 La famille $(-1, 1, 1), (1, -1, 1), (1, 1, -1)$ est-elle une base orthonormée de l'espace vectoriel euclidien \mathbf{R}^3 muni de son produit scalaire canonique ? Si non, l'orthonormaliser à l'aide du procédé de Gram-Schmidt. Donner la décomposition de $(1, 1, 1)$ dans la base ainsi obtenue.

PD

EXERCICE 9.3 (Vu dans EML 2007)

Montrer qu'on définit bien un produit scalaire sur $\mathbf{R}_2[X]$ en posant $\langle P, Q \rangle = \int_{-1}^1 P(t)Q(t)(1 - t^2) dt$.

PD

EXERCICE 9.4 Produit scalaire canonique de $\mathcal{M}_n(\mathbf{R})$ (Vu dans Ecricome 2007, EDHEC 2007)

On définit une application $\varphi : \mathcal{M}_n(\mathbf{R}) \times \mathcal{M}_n(\mathbf{R}) \rightarrow \mathbf{R}$ en posant $\forall (A, B) \in \mathcal{M}_n(\mathbf{R})^2, \varphi(A, B) = \text{tr}({}^t AB)$.

AD

- 1) Montrer que si $A = (a_{i,j})_{1 \leq i,j \leq n}$ et $B = (b_{i,j})_{1 \leq i,j \leq n}$, alors $\varphi(A, B) = \sum_{i=1}^n \sum_{j=1}^n a_{i,j} b_{i,j}$ et en déduire que φ est un produit scalaire sur $\mathcal{M}_n(\mathbf{R})$.
- 2) Montrer que $\mathcal{S}_n(\mathbf{R})$ et $\mathcal{A}_n(\mathbf{R})$ sont orthogonaux pour ce produit scalaire.

EXERCICE 9.5 (D'après EML 2014)

Pour $P, Q \in \mathbf{R}_4[X]$, on pose $\langle P, Q \rangle = \sum_{i=0}^4 P(i)Q(i)$.

PD

- 1) Montrer que φ est un produit scalaire sur $\mathbf{R}_4[X]$.
- 2) On considère à présent l'ensemble $E = \{P \in \mathbf{R}_4[X] : P(0) = P(4) = 0\}$.
 - a) Montrer que E est un sous-espace vectoriel de $\mathbf{R}_4[X]$.
 - b) Montrer que $E = \{X(X - 4)Q, Q \in \mathbf{R}_2[X]\}$. En déduire une base et la dimension de E .
- 3) On définit une famille (M_1, M_2, M_3) de $\mathbf{R}_4[X]$ par $P_1 = (X - 2)(X - 3), P_2 = (X - 1)(X - 3), P_3 = (X - 1)(X - 2)$. On pose alors $M_i = X(X - 4)P_i$.
 - a) Montrer que pour tout $i \in \llbracket 1, 3 \rrbracket, M_i(i) \neq 0$. On pose alors $N_i = \frac{1}{M_i(i)} M_i$.
 - b) Montrer que (N_1, N_2, N_3) est une base orthonormée de E .
 - c) En déduire que pour $P \in E$, on a $P = P(1)N_1 + P(2)N_2 + P(3)N_3$.

EXERCICE 9.6 Vers les projecteurs orthogonaux

Soit E un espace euclidien, et soit F un sous-espace vectoriel de E . Soit (e_1, \dots, e_p) une base orthonormée de F .

PD

On pose alors, pour tout $x \in E, q(x) = x - \sum_{i=1}^p \langle x, e_i \rangle e_i$.

- 1) Montrer que q est un endomorphisme de E , puis que q est un projecteur.
- 2) Montrer que $F \subset \text{Ker}(q)$.

EXERCICE 9.7 Soit E un espace euclidien. Soient F_1, \dots, F_n des sous-espaces vectoriels deux à deux orthogonaux de E . Montrer que la somme $F_1 + \dots + F_n$ est directe

AD

EXERCICE 9.8 Soient x et y deux vecteurs non nuls d'un espace euclidien E . Montrer que x et y sont orthogonaux si et seulement si $\forall \lambda \in \mathbf{R}, \|x + \lambda y\| \geq \|x\|$. Pour le sens réciproque, on pourra étudier la fonction $f : \lambda \mapsto \|x + \lambda y\|^2 - \|x\|^2$.

PD

EXERCICE 9.9 (Extrait de EDHEC 2011)

Soit n un entier naturel supérieur ou égal à 2. On note $\mathbf{R}_n[X]$ l'espace vectoriel des polynômes à coefficients réels, de degré inférieur ou égal à n .

AD

- 1) Montrer que, pour tout couple (P, Q) d'éléments de $\mathbf{R}_n[X]$, l'intégrale $\int_0^{+\infty} P(t)Q(t) e^{-t} dt$ est convergente.

2) Montrer que l'application de $\mathbf{R}_n[X] \times \mathbf{R}_n[X] \rightarrow \mathbf{R}$ définie par $(P, Q) \mapsto \langle P, Q \rangle = \int_0^{+\infty} P(t)Q(t)e^{-t} dt$ est un produit scalaire sur $\mathbf{R}_n[X]$. On note alors $\|\cdot\|$ la norme associée.

3) a) Soient P et Q deux éléments de $\mathbf{R}_n[X]$, P' et Q' leurs polynômes dérivés respectifs. Établir la relation suivante :

$$\langle P', Q \rangle + \langle P, Q' \rangle = \langle P, Q \rangle - P(0)Q(0).$$

b) En déduire que si P est un polynôme non constant de $\mathbf{R}_n[X]$, orthogonal à tout polynôme de degré strictement inférieur, alors on a $\|P\| = |P(0)|$.

EXERCICE 9.10 Caractérisation des matrices orthogonales

Dans cet exercice, $\mathcal{M}_{n,1}(\mathbf{R})$ est muni de son produit scalaire canonique $\langle X, Y \rangle = {}^tXY$. Soit $P \in \mathcal{M}_n(\mathbf{R})$.

1) Montrer que si P est orthogonale, alors $\forall X \in \mathcal{M}_{n,1}(\mathbf{R})$, $\|PX\| = \|X\|$.

2) Inversement, on suppose que pour tout $X \in \mathcal{M}_{n,1}(\mathbf{R})$, $\|PX\| = \|X\|$.

Montrer que $\forall (X, Y) \in (\mathcal{M}_{n,1}(\mathbf{R}))^2$, $\langle PX, PY \rangle = \langle X, Y \rangle$.

En déduire que P est orthogonale.

EXERCICE 9.11 Montrer que $\forall n \in \mathbf{N}^*$, $n^2 \leq \left(\sum_{i=1}^n \sqrt{i}\right)^2 \leq \frac{n^2(n+1)}{2}$.

EXERCICE 9.12 Procédé d'orthonormalisation de Gram-Schmidt

Dans chacun des cas suivants, donner une base orthonormée du sous espace vectoriel F de E , muni du produit scalaire $\langle \cdot, \cdot \rangle$.

1. $E = \mathbf{R}^3$, $F = \text{Vect}((1, 0, -2), (1, 1, 1))$ muni du produit scalaire canonique.

2. $E = F = \mathbf{R}_2[X]$, $\langle P, Q \rangle = \int_{-1}^1 P(t)Q(t) dt$

3. $E = \mathcal{M}_2(\mathbf{R})$, $F = \text{Vect}\left(\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}, \begin{pmatrix} 1 & 1 \\ 1 & -2 \end{pmatrix}\right)$, $\langle A, B \rangle = \text{tr}({}^tAB)$.

4. $E = \mathbf{R}^4$, $F = \text{Vect}((1, 0, 1, 0), (1, 1, 0, 1), (2, 0, 1, 1))$, produit scalaire canonique.

5. $E = \mathbf{R}[X]$, $F = \text{Vect}(X^2 + 1, X^3 + 1)$, $\langle P, Q \rangle = \int_0^1 P(t)Q(t) dt$

EXERCICE 9.13

1) Soient $(a_0, \dots, a_n) \in \mathbf{R}^{n+1}$. Montrer que $(P|Q) = \sum_{k=0}^n P^{(k)}(a_k)Q^{(k)}(a_k)$ est un produit scalaire sur $\mathbf{R}_n[X]$.

2) Pour $n = 2$, $a_0 = -1$, $a_1 = 0$ et $a_2 = 1$, donner une base orthonormée de $\mathbf{R}_2[X]$.

EXERCICE 9.14 Soit \mathbf{R}^4 muni du produit scalaire canonique.

1) Montrer que la base canonique \mathcal{B} de \mathbf{R}^4 est orthogonale.

2) On pose $e_1 = (\frac{3}{5}, \frac{4}{5}, 0, 0)$, $e_2 = (\frac{-4}{5}, \frac{3}{5}, 0, 0)$, $e_3 = (0, 0, \frac{5}{13}, \frac{-12}{13})$, $e_4 = (0, 0, \frac{12}{13}, \frac{5}{13})$. Montrer que $\mathcal{C} = (e_1, e_2, e_3, e_4)$ est une base orthogonale de \mathbf{R}^4 .

3) Déterminer $P_{\mathcal{C}, \mathcal{B}}$.

EXERCICE 9.15 Soit $E = \mathbf{R}_2[X]$. Pour tous $P, Q \in E$, on pose $\langle P, Q \rangle = P(0)Q(0) + P'(0)Q'(0) + P''(0)Q''(0)$.

1) Montrer que $\langle \cdot, \cdot \rangle$ est un produit scalaire sur E .

2) Vérifier que la base canonique de $\mathbf{R}_2[X]$ est orthogonale. En déduire une base orthonormée de $\mathbf{R}_2[X]$.

3) Montrer que $F = \text{Vect}(1 + X, X^2)$ et $G = \text{Vect}(1 - X)$ sont deux sous-espaces vectoriels orthogonaux de E .

EXERCICE 9.16 (QSP ESCP 2001)

Soit E un espace euclidien de dimension n , et soit (e_1, \dots, e_n) une famille de n vecteurs de norme 1.

On suppose que $\forall (i, j) \in \llbracket 1, n \rrbracket^2$, $i \neq j \Rightarrow \|e_i - e_j\| = 1$. Montrer que (e_1, \dots, e_n) est une base de E .

EXERCICE 9.17 (Oral ESCP 2016)

Soit $\mu \geq 1$, et soit E un espace euclidien de produit scalaire $\langle \cdot, \cdot \rangle$ et de norme $\|\cdot\|$.

Une famille (u_1, \dots, u_n) de vecteurs de E est dite μ -presque orthogonale si :

• pour tout $i \in \llbracket 1, n \rrbracket$, $\|u_i\| = 1$

• pour tout $(x_1, \dots, x_n) \in \mathbf{R}^n$, $\frac{1}{\mu} \sum_{i=1}^n x_i^2 \leq \left\| \sum_{i=1}^n x_i u_i \right\|^2 \leq \mu \sum_{i=1}^n x_i^2$.

1) Montrer qu'il famille μ -presque orthogonale est libre.

- 2) Soit (u_1, \dots, u_n) une famille de vecteurs de E . Montrer que (u_1, \dots, u_n) est 1-presque orthogonale si et seulement si elle est orthonormée.
(Pour l'une des implications, on pourra considérer la combinaison linéaire $u_i + u_j$ avec $i \neq j$).
- 3) Soit f un endomorphisme de E .
- Montrer qu'il existe un réel k tel que $\forall x \in E, \|f(x)\| \leq k\|x\|$.
 - Montrer que si f est un automorphisme de E , alors il existe un réel $\lambda \geq 1$ tel que $\forall x \in E, \frac{1}{\lambda}\|x\| \leq \|f(x)\| \leq \lambda\|x\|$.
- 4) Soit $n \in \mathbf{N}^*$ et soit (u_1, \dots, u_n) une famille libre de vecteurs unitaires de E .
Montrer l'existence d'un réel $\mu \geq 1$ tel que (u_1, \dots, u_n) soit μ -presque orthogonale.

CORRECTION DES EXERCICES DU CHAPITRE 9

SOLUTION DE L'EXERCICE 9.1

Par bilinéarité du produit scalaire, on a

$$\begin{aligned}\|x + y\|^2 &= \langle x + y, x + y \rangle = \langle x, x + y \rangle + \langle y, x + y \rangle \\ &= \langle x, x \rangle + \langle x, y \rangle + \langle y, x \rangle + \langle y, y \rangle \\ &= \|x\|^2 + \langle x, y \rangle + \langle x, y \rangle + \|y\|^2 \\ &= \|x\|^2 + 2\langle x, y \rangle + \|y\|^2.\end{aligned}$$

On en déduit donc que

$$\langle x, y \rangle = \frac{1}{2} (\|x + y\|^2 - \|x\|^2 - \|y\|^2).$$

De même, on a¹

$$\|x - y\|^2 = \|x\|^2 - 2\langle x, y \rangle + \|y\|^2.$$

On a alors $\|x + y\|^2 - \|x - y\|^2 = 4\langle x, y \rangle$. Et donc

$$\langle x, y \rangle = \frac{1}{4} (\|x + y\|^2 - \|x - y\|^2).$$

SOLUTION DE L'EXERCICE 9.2

Détaillons le calcul étape par étape. Notons $e_1 = (-1, 1, 1)$, $e_2 = (1, -1, 1)$, $e_3 = (1, 1, -1)$ la

base de départ. Il s'agit bien d'une base car $\begin{pmatrix} -1 & 1 & 1 \\ 1 & -1 & 1 \\ 1 & 1 & -1 \end{pmatrix}$ est une matrice de rang 3, donc

inversible.

On a alors

$$\|e_1\|^2 = (-1)^2 + 1^2 + 1^2 = 3 \text{ et donc } \|e_1\| = \sqrt{3}.$$

Posons alors $x_1 = \frac{e_1}{\|e_1\|} = \frac{1}{\sqrt{3}}(-1, 1, 1)$.

Soit alors $x_2^* = e_2 - \langle e_2, x_1 \rangle x_1$. On a

$$x_2^* = (1, -1, 1) - \frac{1}{\sqrt{3}} \frac{1}{\sqrt{3}} \langle (-1, 1, 1), (1, -1, 1) \rangle = (1, -1, 1) - \frac{1}{3}(-1-1+1)(-1, 1, 1) = \left(\frac{2}{3}, -\frac{2}{3}, \frac{4}{3}\right).$$

$$\text{Alors } \|x_2^*\|^2 = \left(\frac{2}{3}\right)^2 + \left(-\frac{2}{3}\right)^2 + \left(\frac{4}{3}\right)^2 = \frac{24}{9}.$$

$$\text{Et donc } \|x_2^*\| = \sqrt{\frac{24}{9}} = \frac{2\sqrt{6}}{3}. \text{ Posons alors } x_2 = \frac{x_2^*}{\|x_2^*\|} = \frac{1}{\sqrt{6}}(1, -1, 2).$$

Enfin, posons $x_3^* = e_3 - \langle e_3, x_1 \rangle x_1 - \langle e_3, x_2 \rangle x_2$.

$$\begin{aligned}x_3^* &= (1, 1, -1) - \frac{1}{3} \langle (1, 1, -1), (-1, 1, 1) \rangle (-1, 1, 1) - \frac{1}{6} \langle (1, 1, -1), (1, -1, 2) \rangle (1, -1, 2) \\ &= (1, 1, -1) - \frac{1}{3}(-1 + 1 - 1)(-1, 1, 1) - \frac{1}{6}(1 - 1 - 2)(1, -1, 2) \\ &= (1, 1, 0).\end{aligned}$$

Alors $\|x_3^*\| = \sqrt{1 + 1} = \sqrt{2}$. Posons donc $x_3 = \frac{x_3^*}{\|x_3^*\|} = \frac{1}{\sqrt{2}}(1, 1, 0)$.

Alors, la famille $(x_1, x_2, x_3) = \left(\frac{1}{\sqrt{3}}(-1, 1, 1), \frac{1}{\sqrt{6}}(1, -1, 2), \frac{1}{\sqrt{2}}(1, 1, 0)\right)$ est orthonormée. Elle est donc libre, et étant de cardinal 3, c'est une base de \mathbf{R}^3 .

La décomposition de $(1, 1, 1)$ dans cette base orthonormée est donc :

$$(1, 1, 1) = \left\langle (1, 1, 1), \frac{1}{\sqrt{3}}(-1, 1, 1) \right\rangle \frac{1}{\sqrt{3}}(-1, 1, 1) + \left\langle (1, 1, 1), \frac{1}{\sqrt{6}}(1, -1, 2) \right\rangle \frac{1}{\sqrt{6}}(1, -1, 2) + \left\langle (1, 1, 1), \frac{1}{\sqrt{2}}(1, 1, 0) \right\rangle \frac{1}{\sqrt{2}}(1, 1, 0)$$

¹ On peut le reprouver en utilisant la bilinéarité, mais il est plus rapide d'utiliser ce qu'on vient de faire et de remplacer y par $-y$.

Covariance

Puisqu'on n'utilise ici que la bilinéarité du produit scalaire, les mêmes résultats valent en remplaçant produit scalaire par covariance et norme par variance. D'ailleurs, nous avons fait exactement le même calcul dans l'exercice 8 du TD6 : la connaissance des variances permet de retrouver la covariance.

Méthode

Soyons efficaces dans les calculs. Les racines éventuelles qui sont apparues en prenant les normes se simplifient généralement bien car apparaissent en fait, comme ici, au carré. Mieux vaut donc se débarasser tout de suite de ces racines en les élevant au carré plutôt que de garder deux racines à deux endroits différents et attendre la fin pour les simplifier.

Rappel

Dans une base orthonormée (e_1, \dots, e_n) , on a

$$x = \sum_{i=1}^n \langle x, e_i \rangle e_i.$$

$$\begin{aligned}
&= \langle (1, 1, 1), (-1, 1, 1) \rangle \frac{1}{3}(-1, 1, 1) + \langle (1, 1, 1), (1, -1, 2) \rangle \frac{1}{6}(1, -1, 2) + \langle (1, 1, 1), (1, 1, 0) \rangle \frac{1}{2}(1, 1, 0) \\
&= \frac{1}{3}(-1, 1, 1) + \frac{1}{3}(1, -1, 2) + (1, 1, 0).
\end{aligned}$$

SOLUTION DE L'EXERCICE 9.3

Soient $P, Q \in \mathbf{R}_2[X]$. Alors

$$\langle Q, P \rangle = \int_{-1}^1 Q(t)P(t)(1-t^2) dt = \int_{-1}^1 P(t)Q(t)(1-t^2) dt = \langle P, Q \rangle.$$

Donc $\langle \cdot, \cdot \rangle$ est symétrique.

Soient $P, Q, R \in \mathbf{R}_2[X]$ et $\lambda \in \mathbf{R}$. Alors

$$\begin{aligned}
\langle \lambda P + Q, R \rangle &= \int_{-1}^1 (\lambda P(t) + Q(t))R(t)(1-t^2) dt \\
&= \lambda \int_{-1}^1 P(t)R(t)(1-t^2) dt + \int_{-1}^1 Q(t)R(t)(1-t^2) dt \\
&= \lambda \langle P, R \rangle + \langle Q, R \rangle
\end{aligned}$$

Ainsi, $\langle \cdot, \cdot \rangle$ est linéaire à gauche, et donc bilinéaire symétrique.

Si $P \in \mathbf{R}_2[X]$, alors

$$\langle P, P \rangle = \int_{-1}^1 P(t)^2(1-t^2) dt.$$

Puisque la fonction $t \mapsto P(t)^2(1-t^2)$ est positive sur $[-1, 1]$, $\langle P, P \rangle \geq 0$.

Enfin, si $\langle P, P \rangle = 0$, alors $\int_{-1}^1 P(t)^2(1-t^2) dt = 0$. Mais la fonction $t \mapsto P(t)^2(1-t^2)$ est continue² et positive sur $[-1, 1]$. C'est donc la fonction nulle, et ainsi,

$$\forall t \in [-1, 1], P(t)^2(1-t^2) = 0.$$

Mais pour $t \in]-1, 1[$, $1-t^2 \neq 0$ donc $P(t)^2 = 0$ et donc $P(t) = 0$.
 P possède donc une infinité de racines : P est le polynôme nul.

Ainsi, $\langle \cdot, \cdot \rangle$ est bien un produit scalaire sur $\mathbf{R}_2[X]$.

SOLUTION DE L'EXERCICE 9.4

1. Notons $A = (a_{i,j})$ et $B = (b_{i,j})$. Alors ${}^tA = (a_{j,i})$ et donc les coefficients du produit tAB sont donnés par

$$({}^tAB)_{i,j} = \sum_{k=1}^n ({}^tA)_{i,k} b_{k,j} = \sum_{k=1}^n a_{k,i} b_{k,j}.$$

Il vient alors

$$\text{tr}({}^tAB) = \sum_{i=1}^n ({}^tAB)_{i,i} = \sum_{i=1}^n \sum_{k=1}^n a_{k,i} b_{k,i}.$$

Soient $A, B \in \mathcal{M}_n(\mathbf{R})$. Alors

$$\varphi(A, B) = \text{tr}({}^tAB) = \text{tr}({}^t({}^tAB)) = \text{tr}({}^tBA) = \varphi(B, A).$$

Donc φ est symétrique.

Soient $A, B, C \in \mathcal{M}_n(\mathbf{R})$ et $\lambda \in \mathbf{R}$. Alors

$$\varphi(A+B, C) = \text{tr}({}^t(\lambda A + B)C) = \text{tr}((\lambda {}^tA + {}^tB)C) = \lambda \text{tr}({}^tAC) + \text{tr}({}^tBC) = \lambda \varphi(A, C) + \varphi(B, C).$$

Ainsi, φ est linéaire à gauche, et étant symétrique, elle est bilinéaire.

Si $A = (a_{i,j}) \in \mathcal{M}_n(\mathbf{R})$, alors $\varphi(A, A) = \sum_{i=1}^n \sum_{k=1}^n a_{i,k}^2 \geq 0$.

Enfin, si $\varphi(A, A) = 0$, alors $\sum_{i=1}^n \sum_{k=1}^n a_{i,k}^2 = 0$.

Rédaction

Pour conclure que linéaire à gauche implique bilinéaire, il faut impérativement avoir prouvé auparavant la symétrie.

² Si l'hypothèse de continuité n'était pas nécessaire pour prouver la positivité de l'intégrale, elle l'est ici.

Rappel

Pour toute matrice $A \in \mathcal{M}_n(\mathbf{R})$,

$$\text{tr}({}^tA) = \text{tr}(A).$$

Cela est dû au fait que les coefficients diagonaux de tA sont les mêmes que ceux de A .

Or une somme de nombres positifs est nulle si et seulement si chacun des nombres est nul, donc

$$\forall (i, k) \in \llbracket 1, n \rrbracket^2, a_{i,k}^2 = 0 \Rightarrow a_{i,k} = 0.$$

Et donc A est la matrice nulle.

Ainsi, φ est un produit scalaire sur $\mathcal{M}_n(\mathbf{R})$.

2. Il s'agit de montrer que si $S \in \mathcal{S}_n(\mathbf{R})$ et si $A \in \mathcal{A}_n(\mathbf{R})$, alors $\varphi(A, B) = 0$.
Mais on a d'une part, puisque A est symétrique,

$$\varphi(A, B) = \text{tr}({}^t AB) = \text{tr}(AB).$$

D'autre part, B étant antisymétrique, on a

$$\varphi(A, B) = \varphi(B, A) = \text{tr}({}^t BA) = \text{tr}(-BA) = -\text{tr}(BA) = -\text{tr}(AB).$$

Donc $\varphi(A, B) = -\varphi(A, B)$ et ainsi $\varphi(A, B) = 0$.

On en déduit que $\mathcal{S}_n(\mathbf{R})$ et $\mathcal{A}_n(\mathbf{R})$ sont orthogonaux.

SOLUTION DE L'EXERCICE 9.5

1. Soient $P, Q, R \in \mathbf{R}_4[X]$ et soit $\lambda \in \mathbf{R}$. Alors

$$\begin{aligned} \langle \lambda P + Q, R \rangle &= \sum_{i=0}^4 (\lambda P + Q)(i)R(i) \\ &= \lambda \sum_{i=0}^4 P(i)R(i) + \sum_{i=0}^4 Q(i)R(i) \\ &= \lambda \langle P, R \rangle + \langle Q, R \rangle. \end{aligned}$$

Donc $\langle \cdot, \cdot \rangle$ est linéaire à gauche.

De plus, pour tous $P, Q \in \mathbf{R}_4[X]$, on a

$$\langle Q, P \rangle = \sum_{i=0}^4 Q(i)P(i) = \sum_{i=0}^4 P(i)Q(i) = \langle P, Q \rangle.$$

Donc $\langle \cdot, \cdot \rangle$ est symétrique, étant déjà linéaire à gauche, elle est donc bilinéaire.

Pour $P \in \mathbf{R}_4[X]$, on a

$$\langle P, P \rangle = \sum_{i=0}^4 P(i)^2 \geq 0.$$

Enfin, $\langle P, P \rangle = 0$ si et seulement si $\forall i \in \llbracket 0, 4 \rrbracket, P(i)^2 = 0 \Leftrightarrow P(i) = 0$.

Mais un polynôme de degré au plus 4 qui possède 5 racines³ est nécessairement le polynôme nul, donc $\langle P, P \rangle = 0 \Rightarrow P = 0$.

Et donc $\langle \cdot, \cdot \rangle$ est un produit scalaire sur $\mathbf{R}_4[X]$.

- 2.a. Nous pourrions bien entendu le prouver «à la main» comme d'habitude.
Préférons tout de même noter que $E_1 = \{P \in \mathbf{R}_4[X] : P(0) = 0\}$ est un hyperplan de $\mathbf{R}_4[X]$, car noyau de la forme linéaire $P \mapsto P(0)$. C'est donc en particulier un sous-espace vectoriel de $\mathbf{R}_4[X]$.
De même, $E_2 = \{P \in \mathbf{R}_4[X] : P(4) = 0\}$ est un hyperplan.
Et donc $E = E_1 \cap E_2$ est un sous-espace vectoriel de $\mathbf{R}_4[X]$ car intersection de deux sous-espaces vectoriels.
- 2.b. Soit $P \in \mathbf{R}_4[X]$. Alors $P \in E$ si et seulement si 0 et 4 sont des racines de P , donc si et seulement si P est divisible par $X(X-4)$.
Donc si et seulement si il existe $Q \in \mathbf{R}[X]$ tel que $P = X(X-4)Q$. Et P étant de degré au plus 4, nécessairement Q est de degré au plus 2, donc $E = \{X(X-4)Q, Q \in \mathbf{R}_2[X]\}$.
Puisque tout polynôme de $\mathbf{R}_2[X]$ s'écrit de manière unique $Q = aX^2 + bX + c$, tout polynôme de E s'écrit de manière unique $P = aX^3(X-4) + bX^2(X-4) + cX(X-4)$.
Et donc $X^3(X-4), X^2(X-4), X(X-4)$ est une base de E , qui est donc de dimension 3.
- 3.a. Il est clair que les racines de M_i sont tous les entiers de 0 à 4, sauf i , et donc i n'étant pas racine de M_i , $M_i(i) \neq 0$.

Orthogonaux

Deux sous-espaces F et G sont orthogonaux si tout vecteur de F est orthogonal à tout vecteur de G .

³ Il y a bien 5 nombres dans $\llbracket 0, 4 \rrbracket$.

3.b. Pour $(i, j) \in \llbracket 1, 3 \rrbracket^2$, on a

$$\langle N_i, N_j \rangle = \left\langle \frac{1}{M_i(i)} M_i, \frac{1}{M_j(j)} M_j \right\rangle = \frac{1}{M_i(i)M_j(j)} \langle M_i, M_j \rangle.$$

$$\text{Or, } \langle M_i, M_j \rangle = \sum_{k=0}^4 M_i(k)M_j(k).$$

Mais $M_i(k) = 0$ si $k \neq i$, donc $\langle M_i, M_j \rangle = M_i(i)M_j(i)$.

Si $i \neq j$, on a alors $M_j(i) = 0$, donc $\langle N_i, N_j \rangle = 0$.

Et si $i = j$, alors $\langle M_i, M_i \rangle = \frac{1}{M_i(i)^2} M_i(i)^2 = 1$.

Et donc (M_1, M_2, M_3) est une famille orthonormée de E . En particulier, elle est libre. Étant de cardinal $3 = \dim E$, c'est une base de E .

3.c. Puisque (N_1, N_2, N_3) est une base orthonormée de E , tout polynôme P s'écrit

$$P = \langle P, N_1 \rangle N_1 + \langle P, N_2 \rangle N_2 + \langle P, N_3 \rangle N_3.$$

$$\text{Or, pour } i \in \llbracket 1, 3 \rrbracket, \langle P, N_i \rangle = \sum_{k=0}^4 P(k)N_i(k) = P(i)N_i(i) = P(i) \frac{M_i(i)}{M_i(i)} = P(i).$$

Et donc $P = P(1)N_1 + P(2)N_2 + P(3)N_3$.

SOLUTION DE L'EXERCICE 9.6

1. Il est clair que q est à valeurs dans E , montrons qu'il s'agit bien d'une application linéaire. Soient $x, y \in E$ et $\lambda \in \mathbf{R}$. Alors

$$\begin{aligned} q(\lambda x + y) &= \lambda x + y - \sum_{i=1}^p \langle \lambda x + y, e_i \rangle e_i \\ \lambda x + y - \sum_{i=1}^p (\lambda \langle x, e_i \rangle + \langle y, e_i \rangle) e_i & \\ &= \lambda x + y - \lambda \sum_{i=1}^p \langle x, e_i \rangle e_i - \sum_{i=1}^p \langle y, e_i \rangle e_i \\ &= \lambda \left(x - \sum_{i=1}^p \langle x, e_i \rangle e_i \right) + \left(y - \sum_{i=1}^p \langle y, e_i \rangle e_i \right) \\ &= \lambda q(x) + q(y). \end{aligned}$$

Donc q est bien un endomorphisme de E .

On a alors, pour $x \in E$,

$$q(q(x)) = q \left(x - \sum_{i=1}^p \langle x, e_i \rangle e_i \right) = q(x) - \sum_{i=1}^p \langle x, e_i \rangle q(e_i).$$

Or, pour $i \in \llbracket 1, p \rrbracket$, on a

$$q(e_i) = e_i - \sum_{k=1}^p \langle e_i, e_k \rangle e_k = e_i - \langle e_i, e_i \rangle e_i = e_i - \underbrace{\|e_i\|^2}_{=1} = 0_E.$$

Et donc $q(q(x)) = q(x)$: q est un projecteur.

2. Nous venons de prouver que $q(e_1) = \dots = q(e_p) = 0_E$.

Et donc $\text{Ker}(q)$ est un sous-espace vectoriel de E , qui contient e_1, \dots, e_p . Il contient donc $\text{Vect}(e_1, \dots, e_p) = F$.

Notons que ceci pourrait se prouver à la main : si $x \in F$, alors il existe des réels $\lambda_1, \dots, \lambda_p$

tels que $x = \sum_{i=1}^p \lambda_i e_i$. Et alors $q(x) = \sum_{i=1}^p \lambda_i q(e_i) = 0_E$, de sorte que $x \in \text{Ker } q$.

Notons qu'on peut faire mieux, et prouver que $F = \text{Ker } q$.

En effet, si $x \in \text{Ker } q$, alors $x - \sum_{i=1}^p \langle x, e_i \rangle e_i = 0_E \Leftrightarrow x = \sum_{i=1}^p \langle x, e_i \rangle e_i$.

Et donc $x \in \text{Vect}(e_1, \dots, e_p) = F$, de sorte que $\text{Ker}(q) \subset F$, et donc par double inclusion, $F = \text{Ker}(q)$.

Détails

Nous avons (encore) utilisé le fait que $N_i(k) = 0$ si $k \neq i$.

Linéarité à gauche.

Détails

Les $\langle x, e_i \rangle$ sont des réels, on peut donc (par bilinéarité), les sortir du produit scalaire.

Rappel

$\text{Vect}(e_1, \dots, e_p)$ est le plus petit sous-espace vectoriel de E qui contient e_1, \dots, e_p .

SOLUTION DE L'EXERCICE 9.7

Soient $x_1 \in F_1, x_2 \in F_2, \dots, x_n \in F_n$ tels que

$$x_1 + x_2 + \dots + x_n = 0_E.$$

Alors, pour $i \in \llbracket 1, n \rrbracket$, on a

$$\langle x_i, x_1 + x_2 + \dots + x_n \rangle = \langle x_i, 0_E \rangle = 0.$$

Or, par linéarité à droite du produit scalaire, on a également

$$\langle x_i, x_1 \rangle + \dots + \langle x_i, x_n \rangle = 0.$$

Mais les F_i étant deux à deux orthogonaux par hypothèse, on a $\langle x_i, x_j \rangle = 0$ si $i \neq j$.

On en déduit que $\|x_i\|^2 = \langle x_i, x_i \rangle = 0$, et donc nécessairement $x_i = 0_E$.

Ceci étant vrai pour tout i , on a $x_1 = x_2 = \dots = x_n = 0$: les F_i sont en somme directe.

SOLUTION DE L'EXERCICE 9.8

Si x et y sont orthogonaux, alors x et λy le sont encore⁴, et par le théorème de Pythagore,

$$\|x + \lambda y\|^2 = \|x\|^2 + \|\lambda y\|^2 = \|x\|^2 + \lambda^2 \|y\|^2 \geq \|x\|^2.$$

Et une norme étant positive, il vient donc⁵ $\|x + \lambda y\| \geq \|x\|$.

Inversement, supposons que pour tout $\lambda \in \mathbf{R}$, $\|x + \lambda y\| \geq \|x\|$.

Alors $f(\lambda) = \|x + \lambda y\|^2 - \|x\|^2 \geq 0$. Mais

$$\|x + \lambda y\|^2 - \|x\|^2 = \|x\|^2 + 2\lambda \langle x, y \rangle + \lambda^2 \|y\|^2 - \|x\|^2 = \lambda^2 \|y\|^2 + 2\lambda \langle x, y \rangle = \lambda (\lambda \|y\|^2 + 2\langle x, y \rangle).$$

La fonction f étant toujours positive, cela signifie que $\lambda + 2\langle x, y \rangle$ est toujours du signe de λ .

Or, $\lambda + 2\langle x, y \rangle$ change de signe en $\lambda = -2\langle x, y \rangle$, et donc $\langle x, y \rangle = 0$.

SOLUTION DE L'EXERCICE 9.9

1. On sait que pour tout $k \in \mathbf{N}$, l'intégrale $\int_0^{+\infty} t^k e^{-t} dt$ converge⁶.

Or PQ est un polynôme de $\mathbf{R}_{2n}[X]$, donc de la forme $PQ = \sum_{k=0}^{2n} a_k X^k$, de sorte que

$$\int_0^{+\infty} P(t)Q(t)e^{-t} dt = \sum_{k=0}^{2n} a_k \int_0^{+\infty} t^k e^{-t} dt$$

converge car combinaison linéaire d'intégrales convergentes.

2. Soient $P, Q \in \mathbf{R}_n[X]$. Alors

$$\langle P, Q \rangle = \int_0^{+\infty} P(t)Q(t)e^{-t} dt = \int_0^{+\infty} Q(t)P(t)e^{-t} dt = \langle Q, P \rangle.$$

Donc $\langle \cdot, \cdot \rangle$ est symétrique.

Soient $P, Q, R \in \mathbf{R}_n[X]$ et $\lambda \in \mathbf{R}$. Alors

$$\begin{aligned} \langle \lambda P + Q, R \rangle &= \int_0^{+\infty} (\lambda P(t) + Q(t))R(t)e^{-t} dt \\ &= \lambda \int_0^{+\infty} P(t)R(t)e^{-t} dt + \int_0^{+\infty} Q(t)R(t)e^{-t} dt = \lambda \langle P, R \rangle + \langle Q, R \rangle. \end{aligned}$$

Donc $\langle \cdot, \cdot \rangle$ est linéaire à gauche. Puisqu'elle est symétrique, elle est également linéaire à droite et donc bilinéaire symétrique.

Soit $P \in \mathbf{R}_n[X]$. Alors la fonction $t \mapsto P(t)^2 e^{-t}$ est continue et positive sur $[0, +\infty[$, donc par croissance de l'intégrale,

$$\langle P, P \rangle = \int_0^{+\infty} P(t)^2 e^{-t} dt \geq 0.$$

De même, si $\langle P, P \rangle = 0$ alors la fonction $t \mapsto P(t)^2 e^{-t}$ est nulle sur \mathbf{R}_+ . Or, pour tout $t \geq 0$, $e^{-t} \neq 0$ et donc alors $P(t)^2 = 0$ et donc $P(t) = 0$.

Alors P possède une infinité de racines : tous les nombres positifs. Et donc P est nécessairement le polynôme nul.

Ainsi, $\langle \cdot, \cdot \rangle$ est un produit scalaire sur $\mathbf{R}_n[X]$.

⁴ Car $\langle x, \lambda y \rangle = \lambda \langle x, y \rangle = 0$.

⁵ En passant à la racine.

⁶ Et cette intégrale vaut $\Gamma(k+1) = k!$

Classique !

Ce type de produit scalaire à base d'intégrale est ultra-classique, il est indispensable de savoir répondre à une telle question.

Rédaction

Si l'on veut se passer de prouver qu'elle est linéaire à droite, il faut expliquer que c'est automatique car on a déjà prouvé symétrique, et ne pas se contenter de dire qu'elle est linéaire à gauche : la définition de produit scalaire nécessite la bilinéarité, et pas seulement la linéarité à gauche.

3.a. Soit $A > 0$. Alors procédons à une intégration par parties sur le segment $[0, A]$, en posant $u(t) = P(t)$ et $v(t) = Q(t)e^{-t}$, de sorte que u et v sont de classe \mathcal{C}^1 sur $[0, A]$. Il vient

$$\begin{aligned} \int_0^A P'(t)Q(t)e^{-t} dt &= [P(t)Q(t)e^{-t}]_0^A - \int_0^A P(t)Q'(t)e^{-t} dt + \int_0^A P(t)Q(t)e^{-t} dt \\ &= P(A)Q(A)e^{-A} - P(0)Q(0) - \int_0^A P(t)Q'(t)e^{-t} dt + \int_0^A P(t)Q(t)e^{-t} dt \\ &\xrightarrow{A \rightarrow +\infty} -P(0)Q(0) - \int_0^{+\infty} P(t)Q'(t)e^{-t} dt + \int_0^{+\infty} P(t)Q(t)e^{-t} dt \end{aligned}$$

Donc on en déduit⁷ que

$$\langle P', Q \rangle = -P(0)Q(0) + \langle P, Q \rangle - \langle P, Q' \rangle.$$

Soit encore

$$\langle P', Q \rangle + \langle P, Q' \rangle = \langle P, Q \rangle - P(0)Q(0).$$

3.b. Notons $d = \deg P$. Puisque P est non constant, on a $\deg P \geq 1$, et donc $\deg P' = \deg P - 1 < \deg P$.

Et alors, la relation de la question précédente nous donne :

$$\|P\|^2 = \langle P, P \rangle = 2\langle P', P \rangle + P(0)^2.$$

Mais par hypothèse, P et P' sont orthogonaux, donc $\|P\|^2 = P(0)^2$.

Et $\|P\|$ étant positif, il vient

$$\|P\| = \sqrt{\|P\|^2} = \sqrt{P(0)^2} = |P(0)|.$$

SOLUTION DE L'EXERCICE 9.10

1. Soit $X \in \mathcal{M}_{n,1}(\mathbf{R})$. Alors on a $\|PX\| = {}^t(PX)PX = {}^tX \underbrace{{}^tPP}_{=I_n} X = {}^tXX = \|X\|^2$.

Et puisqu'une norme est toujours positive, on en déduit que $\|PX\| = \|X\|$.

2. Soient X et $Y \in \mathcal{M}_{n,1}(\mathbf{R})$. Nous savons que⁸

$$\|P(X + Y)\|^2 = \|PX\|^2 + 2\langle PX, PY \rangle + \|PY\|^2 = \|X\|^2 + 2\langle PX, PY \rangle + \|Y\|^2.$$

D'autre part, on a également

$$\|X + Y\|^2 = \|X\|^2 + 2\langle X, Y \rangle + \|Y\|^2$$

et donc $\|P(X + Y)\|^2 = \|X + Y\|^2 \Rightarrow \langle PX, PY \rangle = \langle X, Y \rangle$.

En particulier, si (X_1, \dots, X_n) sont les vecteurs de la base canonique de $\mathcal{M}_{n,1}(\mathbf{R})$, alors pour

$$\text{tous } (i, j) \in \llbracket 1, n \rrbracket^2, \langle X_i, X_j \rangle = \begin{cases} 0 & \text{si } i \neq j \\ 1 & \text{sinon} \end{cases} \text{ et donc } \langle PX_i, PX_j \rangle = \begin{cases} 0 & \text{si } i \neq j \\ 1 & \text{sinon} \end{cases}.$$

Autrement dit (PX_1, \dots, PX_n) est une famille orthonormée de $\mathcal{M}_{n,1}(\mathbf{R})$, de cardinal n , donc⁹ c'est une base orthonormée \mathcal{B} de $\mathcal{M}_{n,1}(\mathbf{R})$.

Et alors P est la matrice de passage de \mathcal{B} à la base canonique : elle est alors orthogonale car matrice de passage entre deux bases orthonormées.

Autre méthode : montrons que $A = {}^tPP - I_n$ est la matrice nulle, ce qui prouvera que ${}^tPP = I_n$.

Pour tout $X, Y \in \mathcal{M}_{n,1}(\mathbf{R})$, on a ${}^tXAY = {}^tX{}^tPPY - {}^tXY = \langle PX, PY \rangle - \langle PX, PY \rangle$.

En particulier, pour $X = AY$, il vient ${}^t(AY)AY = 0 \Leftrightarrow \|AY\|^2 = 0$.

Et donc, pour tout $Y \in \mathcal{M}_{n,1}(\mathbf{R})$, $\|AY\| = 0 \Rightarrow \forall Y \in \mathcal{M}_{n,1}(\mathbf{R}), AY = 0$.

Ceci implique nécessairement que¹⁰ $A = 0$, et donc P est orthogonale.

SOLUTION DE L'EXERCICE 9.11

Appliquons l'inégalité de Cauchy-Schwarz dans \mathbf{R}^n à $x = (\sqrt{1}, \dots, \sqrt{n})$ et $y = (1, \dots, 1)$. Il vient alors

$$\left(\sum_{i=1}^n \sqrt{i} \right)^2 \leq \left(\sum_{i=1}^n \sqrt{i^2} \right) \left(\sum_{i=1}^n 1 \right) = \frac{n(n+1)}{2} n = \frac{n^2(n+1)}{2}.$$

Rédaction

Attention à ne pas écrire

$$\int_0^{+\infty} = \int_0^A.$$

Dans un premier temps on travaille donc avec des intégrales entre 0 et A . Les intégrales entre 0 et $+\infty$ n'apparaîtront qu'après un passage à la limite $A \rightarrow +\infty$.

⁷ Car

$$\langle P, Q \rangle = \lim_{A \rightarrow +\infty} \int_0^A P(t)Q(t)e^{-t} dt.$$

⁸ C'est l'identité (dite «de polarisation») de l'exercice 1.

⁹ Toute famille orthonormée est libre.

Méthode

La difficulté est de choisir les bons vecteurs. Ici, on voulait, d'une part que $x_i y_i = \sqrt{i}$ et d'autre part que $x_i^2 = i$ car on avait reconnu

$$\frac{n(n+1)}{2} = \sum_{i=1}^n i$$

dans le membre de droite.

De plus, $\forall i \in \llbracket 1, n \rrbracket$, $\sqrt{i} \geq 1$, de sorte que $\sum_{i=1}^n \sqrt{i} \geq n$ et donc $\left(\sum_{i=1}^n \sqrt{i}\right)^2 \geq n^2$.

SOLUTION DE L'EXERCICE 9.12

- $\frac{1}{\sqrt{5}}(1, 0, -2), \frac{1}{\sqrt{70}}(6, 5, 3).$ 2. $\frac{1}{\sqrt{2}}, \sqrt{\frac{3}{2}}X, \sqrt{\frac{45}{8}}\left(X^2 - \frac{1}{3}\right).$
- $\frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \frac{1}{\sqrt{2}}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$ 4. $\frac{1}{\sqrt{2}}(1, 0, 1, 0), \frac{1}{\sqrt{10}}(1, 2, -1, 2), \frac{1}{\sqrt{15}}(1, -3, -1, 2)$
- $\sqrt{\frac{15}{28}}(X^2 + 1), \frac{\sqrt{7}}{2}(16X^3 - 15X^2 + 1)$

SOLUTION DE L'EXERCICE 9.13

- Comme souvent, le seul point délicat est de prouver que si $\langle P|P \rangle = 0$, alors $P = 0$.

Soit donc $P \in \mathbf{R}_n[X]$ tel que $\sum_{k=0}^n P^{(k)}(a_k)^2 = 0$.

Alors la somme de carrés est nulle si et seulement si chacun des termes est nul, donc $\forall k \in \llbracket 0, n \rrbracket, P^{(k)}(a_k) = 0$.

On a $P = \sum_{i=0}^n b_i X^i$. Donc $P^{(n)} = b_n n!$

Ainsi, $P^{(n)}$ est le polynôme constant $b_n n!$. Puisque $P^{(n)}(a_n) = 0$, $b_n n! = 0$, et donc $b_n = 0$. Ainsi $\deg P \leq n - 1$.

Mais alors $P^{(n-1)}$ est une constante, et puisque $P^{(n-1)}(a_{n-1}) = 0$, alors $P^{(n-1)}$ est nul, donc P est de degré au plus $n - 2$, etc.

On prouve ainsi que P est nécessairement nul.

- Il s'agit d'appliquer le procédé de Gram-Schmidt à la base canonique $(P_0, P_1, P_2) = (1, X, X^2)$ de $\mathbf{R}_2[X]$.

Posons $P_0 = 1$, et $Q_0 = \frac{P_0}{\|P_0\|}$. Puisque $\|P_0\| = \sqrt{\langle P_0|P_0 \rangle} = 1$, alors $Q_0 = 1$.

Soit $Q_1^* = P_1 - \langle P_1, Q_0 \rangle Q_0$. Alors

$$\langle P_1, Q_0 \rangle = P_1(-1)Q_0(-1) + \underbrace{P_1'(0)Q_0'(0)}_{=0} + \underbrace{P_1''(1)Q_0''(1)}_{=0} = -1.$$

Et donc $Q_1^* = X + 1$ et

$$\|Q_1^*\|^2 = Q_1^*(-1)^2 + (Q_1^*)'(0)^2 + (Q_1^*)''(1)^2 = 1.$$

Posons donc $Q_1 = Q_1^* = X + 1$. Soit enfin $Q_2^* = P_2 - \langle P_2, Q_0 \rangle Q_0 - \langle P_2, Q_1 \rangle Q_1$. On vérifie que $\langle P_2, Q_0 \rangle = 1$ et $\langle P_2, Q_1 \rangle = 0$ et donc $Q_2^* = X^2 - 1$.

Alors $\|Q_2^*\|^2 = Q_2^*(-1)^2 + (Q_2^*)'(0)^2 + (Q_2^*)''(0)^2 = 0 + 0 + 4$.

Et donc posons $Q_2 = \frac{Q_2^*}{\|Q_2^*\|} = \frac{1}{2}(X^2 - 1)$. Alors $(Q_0, Q_1, Q_2) = \left(1, X + 1, \frac{1}{2}(X^2 - 1)\right)$ est une base orthonormée de $\mathbf{R}_2[X]$.

SOLUTION DE L'EXERCICE 9.14

- C'est du cours.
- Il s'agit d'une simple formalité, mais il faut calculer tous les produits scalaire $\langle e_i, e_j \rangle$.
- Nous savons que

$$P_{\mathcal{B}, \mathcal{C}} = \begin{pmatrix} \frac{3}{5} & \frac{4}{5} & 0 & 0 \\ \frac{4}{5} & \frac{3}{5} & 0 & 0 \\ 0 & 0 & \frac{5}{13} & \frac{12}{13} \\ 0 & 0 & \frac{-12}{13} & \frac{5}{13} \end{pmatrix}.$$

Mais les deux bases étant orthonormées, la matrice de passage de \mathcal{B} à \mathcal{C} est une matrice orthogonale, et donc son inverse est facile à calculer : il s'agit de sa transposée :

$$P_{\mathcal{C}, \mathcal{B}} = \begin{pmatrix} \frac{3}{5} & \frac{4}{5} & 0 & 0 \\ \frac{4}{5} & \frac{3}{5} & 0 & 0 \\ 0 & 0 & \frac{5}{13} & \frac{-12}{13} \\ 0 & 0 & \frac{12}{13} & \frac{5}{13} \end{pmatrix}.$$

Détails

Si l'on dérive n fois P , tous les termes de degré inférieur ou égal à n s'annulent, et il ne reste que la dérivée n -ième de $b_n X^n$. Or, les dérivées successives de X^n sont

$$nX^{n-1}, n(n-1)X^{n-2}, \dots, n!$$

Méthode

Pour obtenir une base orthonormée, il suffit d'appliquer le procédé de Gram-Schmidt à **n'importe quelle** base de E . Ici, on a choisi la base canonique, mais on aurait pu prendre toute autre base de $\mathbf{R}_2[X]$.

SOLUTION DE L'EXERCICE 9.15

1. La bilinéarité, la symétrie et $\langle P, P \rangle \geq 0$ ne présentent aucune difficulté.
Soit $P \in \mathbf{R}_2[X]$ tel que $\langle P, P \rangle = 0 \Leftrightarrow P(0)^2 + P'(0)^2 + P''(0)^2 = 0$.
Une somme de nombres positifs est nulle si et seulement si tous ces nombres sont nuls, donc $P(0) = P'(0) = P''(0) = 0$.
Et donc 0 est une racine de P de multiplicité au moins égale à 3 : puisque P est de degré au plus deux, c'est donc que $P = 0$.
Ainsi, P est bien un produit scalaire sur $\mathbf{R}_2[X]$.
Alternative : on peut aussi écrire $P = aX^2 + bX + c$, et alors $P(0) = c$, $P'(0) = b$ et $P''(0) = 2a$, de sorte que

$$\langle P, P \rangle = 0 \Leftrightarrow c^2 + b^2 + 4a^2 = 0 \Leftrightarrow a = b = c = 0 \Leftrightarrow P = 0.$$

2. Il s'agit de prouver que deux vecteurs distincts de la base canonique $(1, X, X^2)$ sont orthogonaux. Or, on a

$$\begin{aligned}\langle 1, X \rangle &= 1 \times 0 + 0 \times 1 + 0 \times 0 = 0 \\ \langle 1, X^2 \rangle &= 1 \times 0 + 0 \times 0 + 0 \times 2 = 0 \\ \langle X, X^2 \rangle &= 0 \times 0 + 1 \times 0 + 0 \times 2 = 0\end{aligned}$$

Donc la base canonique est bien orthogonale.

Pour obtenir une base orthonormée, il suffit de diviser chacun des vecteurs de cette famille par sa norme.

Mais $\|1\| = 1$, $\|X\| = 1$ et $\|X^2\| = 2$. Donc $\left(1, X, \frac{X^2}{2}\right)$ est une base orthonormée de $\mathbf{R}_2[X]$.

3. On a, par bilinéarité,

$$\langle 1 + X, 1 - X \rangle = \langle 1, 1 \rangle + \langle X, 1 \rangle - \langle 1, X \rangle - \langle X, X \rangle = 1 - 1 = 0.$$

Et de même, $\langle X^2, 1 - X \rangle = \langle X^2, 1 \rangle - \langle X^2, X \rangle = 0$.

Donc si $P = \lambda(1 + X) + \mu X^2 \in F$ et $Q = \nu(1 - X) \in G$, alors

$$\langle P, Q \rangle = \langle \lambda(1 + X) + \mu X^2, \nu(1 - X) \rangle = \lambda\nu\langle 1 + X, 1 - X \rangle + \mu\nu\langle X^2, 1 - X \rangle = 0.$$

Et donc tout vecteur de F est orthogonal à tout vecteur de G : F et G sont orthogonaux.

SOLUTION DE L'EXERCICE 9.16

D'après l'identité de polarisation, pour $i \neq j$, on a

$$\langle e_i, e_j \rangle = \frac{1}{2} (\|e_i\|^2 + \|e_j\|^2 - \|e_i - e_j\|^2) = \frac{1}{2}.$$

Soient $\lambda_1, \dots, \lambda_n$ des réels tels que $\sum_{i=1}^n \lambda_i e_i = 0$.

Alors en prenant le produit scalaire avec e_j , il vient

$$\forall j \in \llbracket 1, n \rrbracket, \lambda_j \langle e_j, e_j \rangle + \sum_{\substack{i=1 \\ i \neq j}}^n \lambda_i \langle e_i, e_j \rangle = 0$$

soit encore

$$\forall j \in \llbracket 1, n \rrbracket, \lambda_j + \frac{1}{2} \sum_{\substack{i=1 \\ i \neq j}}^n \lambda_i = 0 \Leftrightarrow \lambda_j = -2 \sum_{i=1}^n \lambda_i.$$

En particulier, on a $\lambda_1 = \lambda_2 = \dots = \lambda_n$, et donc

$$\forall j \in \llbracket 1, n \rrbracket, \lambda_j = -2n\lambda_j$$

ce qui implique nécessairement $\lambda_1 = \dots = \lambda_n = 0$.

Nous venons de prouver que la famille (e_1, \dots, e_n) est libre, et donc étant de cardinal $n = \dim E$, c'est une base de E .

Orthogonalité

Dire que u et v sont orthogonaux équivaut à dire que $\text{Vect}(u)$ et $\text{Vect}(v)$ sont deux sous-espaces orthogonaux. En particulier, diviser les vecteurs par leur norme ne change rien à l'orthogonalité.

Polarisation

Les identités qu'on appelle identité de polarisation sont les égalités prouvées à l'exercice 1. Il n'est pas nécessaire de les connaître par cœur, mais il est intéressant de savoir qu'elles existent. Il est alors facile de les retrouver.

Précision

On vient de montrer que λ_j est égale à une somme qui ne dépend pas de j . Et donc tous les λ_j sont égaux.

SOLUTION DE L'EXERCICE 9.17

1. Soit (u_1, \dots, u_n) une famille μ -presque orthogonale, et soient x_1, \dots, x_n des réels tels que $\sum_{i=1}^n x_i u_i = 0_E$. Alors, on a

$$0 \leq \frac{1}{\mu} \sum_{i=1}^n x_i^2 \leq \underbrace{\left\| \sum_{i=1}^n x_i u_i \right\|^2}_{= \|0_E\|^2 = 0}.$$

On en déduit que $\sum_{i=1}^n x_i^2 = 0$, et donc pour tout $i \in \llbracket 1, n \rrbracket$, $x_i = 0$.

Ainsi, la famille (u_1, \dots, u_n) est libre.

2. Soit (u_1, \dots, u_n) une famille orthonormée. Alors, pour tout $(x_1, \dots, x_n) \in \mathbf{R}^n$, on a

$$\left\| \sum_{i=1}^n x_i u_i \right\|^2 = \sum_{i=1}^n x_i^2.$$

Et donc (u_1, \dots, u_n) est 1-presque orthogonale.

Inversement, supposons que (u_1, \dots, u_n) est 1-presque orthogonale. Alors par définition, les u_i sont tous unitaires.

Soient alors $(i, j) \in \llbracket 1, n \rrbracket^2$, avec $i \neq j$. On a alors

$$\|u_i + u_j\|^2 = \|u_i\|^2 + \|u_j\|^2 + 2\langle u_i, u_j \rangle = 2(1 + \langle u_i, u_j \rangle).$$

Mais $u_i + u_j$ n'est autre que la combinaison linéaire $\sum_{k=1}^n x_k u_k$ où les x_k sont tous nuls, sauf x_i et x_j qui valent tous les deux 1.

Et donc on a $\sum_{k=1}^n x_k^2 = 2$.

Or, la famille étant 1-presque orthogonale, il vient

$$\frac{1}{1} \sum_{k=1}^n x_k^2 \leq \|u_i + u_j\|^2 \leq \sum_{k=1}^n x_k^2 \Leftrightarrow 2 \leq 2(1 + \langle u_i, u_j \rangle) \leq 2.$$

Et donc $2(1 + \langle u_i, u_j \rangle) = 0 \Leftrightarrow \langle u_i, u_j \rangle = 0$.

Ceci prouve donc que (u_1, \dots, u_n) est orthogonale, et donc orthonormée.

- 3.a. Considérons une base orthonormée (e_1, \dots, e_n) de E . Alors pour tout $x \in E$, il existe des réels x_1, \dots, x_n tels que $x = \sum_{i=1}^n x_i e_i$, et donc $\|x\|^2 = \sum_{i=1}^n x_i^2$.

D'autre part, on a $f(x) = \sum_{i=1}^n x_i f(e_i)$.

Et donc

$$\|f(x)\|^2 = \left\| \sum_{i=1}^n x_i f(e_i) \right\|^2 \leq \left(\sum_{i=1}^n |x_i| \cdot \|f(e_i)\| \right)^2.$$

Appliquons alors l'inégalité de Cauchy-Schwarz dans \mathbf{R}^n :

$$\left(\sum_{i=1}^n |x_i| \cdot \|f(e_i)\| \right)^2 \leq \left(\sum_{i=1}^n |x_i|^2 \right) \left(\sum_{i=1}^n \|f(e_i)\|^2 \right).$$

Et donc $\|f(x)\|^2 \leq \|x\| \times \left(\sum_{i=1}^n \|f(e_i)\|^2 \right)$.

Ainsi, en posant $k = \sqrt{\sum_{i=1}^n \|f(e_i)\|^2}$, qui est bien une constante indépendante de x , on a $\|f(x)\| \leq k\|x\|$.

Détails

C'est le corollaire du théorème de Pythagore.

Cauchy-Schwarz

Il s'agit de l'inégalité de Cauchy-Schwarz dans \mathbf{R}^n , pour le produit scalaire canonique, appliquée aux deux vecteurs

$$x = (|x_1|, \dots, |x_n|), \\ y = (\|f(e_1)\|, \dots, \|f(e_n)\|).$$

- 3.b. Si f est un automorphisme, alors le résultat de la question précédente s'applique en particulier à f^{-1} : il existe $k \in \mathbf{R}$ tel que pour tout $x \in E$, $\|f^{-1}(x)\| \leq k\|x\|$.
 En particulier, si l'on applique ceci à $f(x)$ au lieu de x , il vient

$$\|f^{-1}(f(x))\| \leq k\|f(x)\| \Leftrightarrow \|x\| \leq k\|f(x)\|.$$

Notons que l'expression de k trouvée à la question précédente prouve qu'on peut supposer k strictement positif, puisque, f^{-1} étant un automorphisme, les $f^{-1}(e_i)$ sont non nuls, et

donc $\sqrt{\sum_{i=1}^n \|f^{-1}(e_i)\|^2} > 0$.

On a donc $\frac{1}{k}\|x\| \leq \|f(x)\|$.

D'autre part, il existe une constante M telle que $\forall x \in E, \|f(x)\| \leq M\|x\|$.

Notons alors $\lambda = \max(M, k)$, de sorte que $M \leq \lambda$ et $\frac{1}{\lambda} \leq \frac{1}{k}$.

On a alors bien, pour tout $x \in E$,

$$\frac{1}{\lambda}\|x\| \leq \|f(x)\| \leq \lambda\|x\|.$$

Remarque

Puisqu'il vient $\frac{1}{\lambda} \leq \lambda$, nécessairement $\lambda \geq 1$.

4. Soit (u_1, \dots, u_n) une famille libre de vecteurs unitaires, et soit $F = \text{Vect}(u_1, \dots, u_n)$.
 Soit alors (e_1, \dots, e_n) une base orthonormée de F , et soit f l'unique endomorphisme de F tel que $f(e_i) = u_i$.
 Alors f est un automorphisme puisque l'image de la base (e_1, \dots, e_n) est une famille libre de F , et donc une base de F .
 Par la question 3.b, il existe donc une constante $\lambda \geq 1$ telle que

$$\forall x \in F, \frac{1}{\lambda}\|x\| \leq \|f(x)\| \leq \lambda\|x\|.$$

En particulier, pour $(x_1, \dots, x_n) \in \mathbf{R}^n$, soit $x = \sum_{i=1}^n x_i e_i$.

Alors $\|x\|^2 = \sum_{i=1}^n x_i^2$ et $\|f(x)\|^2 = \left\| \sum_{i=1}^n x_i u_i \right\|^2$.

Et donc

$$\frac{1}{\lambda^2}\|x\|^2 \leq \|f(x)\|^2 \leq \lambda^2\|x\|^2 \Leftrightarrow \frac{1}{\lambda^2} \sum_{i=1}^n x_i^2 \leq \left\| \sum_{i=1}^n x_i u_i \right\|^2 \leq \lambda^2 \sum_{i=1}^n x_i^2.$$

DIAGONALISATION

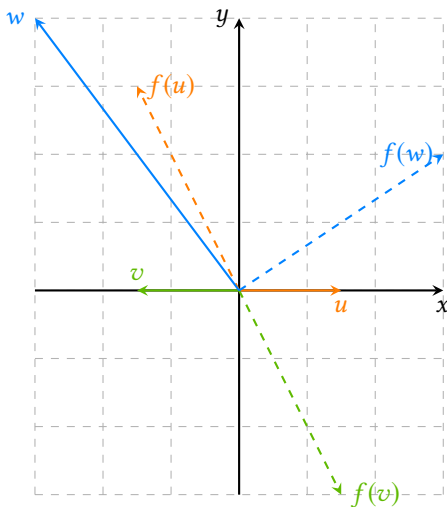
Cours	309
10.1 Polynômes annulateurs et valeurs propres	310
10.2 Endomorphismes et matrices diagonalisables	314
10.3 Applications de la diagonalisation	319
Méthodes	324
Exercices	325
Correction des exercices	328

Dans ce chapitre, nous allons nous intéresser aux endomorphismes f d'un espace vectoriel E pour lesquels il existe une base de E formée de vecteurs propres de f .

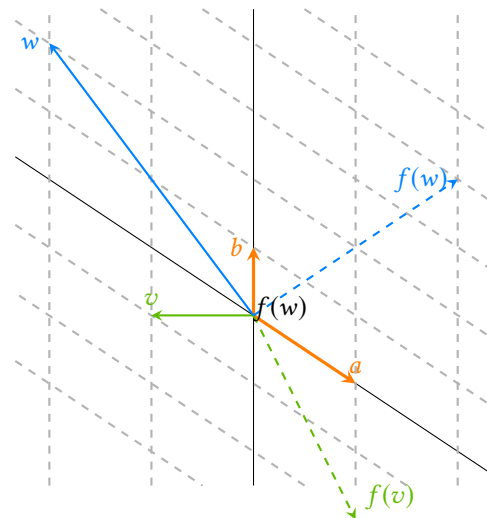
Avant de commencer, essayons d'expliquer sur un exemple pourquoi il s'agit d'une notion intéressante.

Considérons l'endomorphisme de \mathbf{R}^2 défini par $f(x, y) = (2x, -2x - y)$, et essayons de comprendre géométriquement à quoi il correspond.

Sur la figure de gauche ci-dessous, nous avons représenté trois vecteurs u, v et w ainsi que leurs images par f .



Dans la base canonique, il semble difficile d'interpréter géométriquement f .



En revanche, dans une base de vecteurs propres, tout semble plus clair.

A priori, il semble bien difficile de comprendre «géométriquement» quelle est la transformation qui se cache derrière f .

Un calcul de valeurs propres et de vecteurs propres nous indiquerait que $\text{Spec}(f) = \{-1, 2\}$, et que $E_{-1}(f) = \text{Vect}\left(\frac{3}{2}, -1\right)$ et $E_2(f) = \text{Vect}(0, 1)$.

Les deux vecteurs propres $a = \left(\frac{3}{2}, -1\right)$ et $b = (0, 1)$ n'étant pas colinéaires, ils forment une famille libre, et donc une base \mathcal{B} de \mathbf{R}^2 .

Il est alors plus judicieux de travailler dans cette base : tout vecteur $u \in \mathbf{R}^2$ s'écrit de manière unique $u = \lambda a + \mu v$, et alors par linéarité de f ,

$$f(u) = \lambda f(a) + \mu f(v) = -\lambda a + 2\mu b.$$

Autrement dit, dans la base (a, b) , on comprend bien la transformation réalisée par f : calculer l'image d'un vecteur par f , c'est multiplier par -1 sa composante suivant a et par 2 sa composante suivant b .

Or, changer de base, c'est tout simplement changer de repère !

Prenons par exemple le vecteur w . Comme le montre la figure de droite, il se décompose sous la forme $w = -2a + 2b$, et donc ses coordonnées dans la base \mathcal{B} sont $(-2, 2)$.

Et donc son image par f est $-2a + 4b$.

De même, $v = -a - b$, et donc $f(v) = a - 2b$.

Ainsi, lorsqu'elles existent¹, des bases de vecteurs propres nous permettent de mieux comprendre f .

¹ Et nous verrons que ce n'est malheureusement pas toujours le cas.

Dans tout ce chapitre, \mathbf{K} désigne indifféremment \mathbf{R} ou \mathbf{C} , et E un \mathbf{K} -espace vectoriel de dimension finie n .

Si $A \in \mathcal{M}_n(\mathbf{K})$, alors f_A désigne l'endomorphisme de $\mathcal{M}_{n,1}(\mathbf{K})$ défini par $f_A(X) = AX$. On rappelle que la matrice de f_A dans la base canonique de $\mathcal{M}_{n,1}(\mathbf{K})$ est alors A .

10.1 POLYNÔMES ANNULATEURS ET VALEURS PROPRES

10.1.1 Polynômes d'endomorphismes et de matrices

Définition 10.1 – Soit $f \in \mathcal{L}(E)$, et soit $P = \sum_{k=0}^d a_k X^k \in \mathbf{K}[X]$.

On note alors $P(f)$ l'endomorphisme de E défini par

$$P(f) = \sum_{k=0}^d a_k f^k = a_0 \text{id}_E + a_1 f + a_2 f^2 + \cdots + a_d f^d.$$

De même, si $A \in \mathcal{M}_n(\mathbf{K})$, on note $P(A)$ la matrice de $\mathcal{M}_n(\mathbf{K})$ définie par

$$P(A) = \sum_{k=0}^d a_k A^k = a_0 I_n + a_1 A + \cdots + a_d A^d.$$

Précaution

Attention au terme constant, qui devient id_E (pour un endomorphisme) ou I_n (pour une matrice).

Remarque. En particulier, si $P = X$, $P(f) = f$.

Proposition 10.2 : Soit $f \in \mathcal{L}(E)$, soit \mathcal{B} une base de E et soit $A = \text{Mat}_{\mathcal{B}}(f) \in \mathcal{M}_n(\mathbf{K})$. Alors pour tout $P \in \mathbf{K}[X]$, $\text{Mat}_{\mathcal{B}}(P(f)) = P(A)$.

Proposition 10.3 : Soit $f \in \mathcal{L}(E)$ (respectivement $A \in \mathcal{M}_n(\mathbf{K})$), et soient $P, Q \in \mathbf{K}[X]$ et $\lambda \in \mathbf{K}$. Alors

- $(\lambda P + Q)(f) = \lambda P(f) + Q(f)$ (resp. $(\lambda P + Q)(A) = \lambda P(A) + Q(A)$)
- $(PQ)(f) = P(f) \circ Q(f)$ (resp. $(PQ)(A) = P(A) \times Q(A)$)

Corollaire 10.4 – Soit $f \in \mathcal{L}(E)$, $P, Q \in \mathbf{K}[X]$. Alors $P(f)$ et $Q(f)$ commutent. En particulier, f commute avec $P(f)$.

Démonstration. Le produit des polynômes est commutatif, c'est-à-dire $PQ = QP$.

Donc $P(f) \circ Q(f) = (PQ)(f) = (QP)(f) = Q(f) \circ P(f)$.

En particulier, pour $Q = X$, on obtient $P(f) \circ f = f \circ P(f)$. □

De même, si $A \in \mathcal{M}_n(\mathbf{K})$, alors $P(A)$ et $Q(A)$ commutent.

Proposition 10.5 : Soient $f, g \in \mathcal{L}(E)$ deux endomorphismes qui commutent. Alors tout sous-espace propre de f est stable par g .

En particulier
On a déjà prouvé que pour tout $P \in \mathbf{K}[X]$, $P(f)$ commute avec f , donc tout sous-espace propre de f est stable par $P(f)$, et vice-versa.

Démonstration. Soit λ une valeur propre de f , et soit $E_\lambda(f)$ le sous-espace propre associé. Alors pour $x \in E_\lambda(f)$, on a

$$f(g(x)) = g(f(x)) = g(\lambda x) = \lambda g(x).$$

Donc $g(x) \in E_\lambda(f)$, et ainsi, $E_\lambda(f)$ est stable par g . □

10.1.2 Polynômes annulateurs

Définition 10.6 – Soit $f \in \mathcal{L}(E)$ (respectivement $A \in \mathcal{M}_n(\mathbf{K})$), et soit $P \in \mathbf{K}[X]$. P est un polynôme annulateur de f (resp. de A) si $P(f) = 0_E$ (resp. $P(A) = 0_n$).

Remarque. Notons que si $A = \text{Mat}_{\mathcal{B}}(f)$ est la matrice de f dans une base, alors $P \in \mathbf{K}[X]$ est annulateur de f si et seulement si P est annulateur de A .

Exemples 10.7

- Soit p un projecteur de $\mathcal{L}(E)$. Alors $p^2 = p$, donc $p^2 - p = 0$. Ainsi, $X^2 - X$ est un polynôme annulateur de p .

- Soit $A = \begin{pmatrix} 1 & 0 & -1 \\ 3 & 1 & 0 \\ -3 & -1 & 1 \end{pmatrix}$, et soit $P = X^3 - 3X^2 - 1$.

On a

$$A^2 = \begin{pmatrix} 4 & 1 & -2 \\ 6 & 1 & -3 \\ -9 & -2 & 4 \end{pmatrix}, A^3 = \begin{pmatrix} 13 & 3 & -6 \\ 18 & 4 & -9 \\ -27 & -6 & 13 \end{pmatrix}$$

de sorte que

$$P(A) = A^3 - 3A^2 - I_3 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Ainsi P est un polynôme annulateur de A . • Soit $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ une matrice 2×2 .

Alors $X^2 - \text{tr}(A)X + \det(A)$ est annulateur de A . En effet, on a

$$A^2 - (a+d)A + (ad-bc)I_2 = \begin{pmatrix} a^2 + bc & ab + bd \\ ca + cd & cb + d^2 \end{pmatrix} - \begin{pmatrix} a^2 + ad & ab + bd \\ ac + cd & ad + d^2 \end{pmatrix} + \begin{pmatrix} ad - bc & 0 \\ 0 & ad - bc \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

Proposition 10.8 : Soit $f \in \mathcal{L}(E)$. Alors il existe un polynôme annulateur **non nul** de f , de degré au plus n^2 .

Mieux
En fait, il est possible de trouver un polynôme annulateur de degré au plus n , mais c'est bien plus dur (voir la partie II de HEC 2013).

Démonstration. Rappelons que $\mathcal{L}(E)$ est de dimension $\dim E \times \dim E = n^2$.

La famille $(\text{id}_E, f, f^2, \dots, f^{n^2})$ est de cardinal $n^2 + 1$: elle est nécessairement liée.

Il existe donc des scalaires $(a_0, a_1, \dots, a_{n^2})$ non tous nuls tels que

$$a_0 \text{id}_E + a_1 f + \dots + a_{n^2} f^{n^2} = 0.$$

Si on note $P = \sum_{k=0}^{n^2} a_k X^k \in \mathbf{K}_{n^2}[X]$, alors P est non nul² et la relation précédente s'écrit $P(f) = 0$. □

² Puisque les a_i ne sont pas tous nuls.

Corollaire 10.9 – Soit $A \in \mathcal{M}_n(\mathbf{K})$. Alors il existe un polynôme non nul annulateur de A .

Démonstration. C'est la même preuve, en utilisant $\dim \mathcal{M}_n(\mathbf{K}) = n^2$ et la famille $(I_n, A, A^2, \dots, A^{n^2})$. □

Proposition 10.10 (Polynômes annulateurs des matrices diagonales) : Soit

$$D = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \ddots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & \lambda_n \end{pmatrix} \in \mathcal{M}_n(\mathbf{K}) \text{ une matrice diagonale.}$$

Alors $P \in \mathbf{K}[X]$ est annulateur de D si et seulement si tous les λ_i sont racines de P .

En particulier

Il est facile de trouver un polynôme annulateur de D : on peut prendre

$$P(X) = (X - \lambda_1) \cdots (X - \lambda_n).$$

Et même mieux, si des coefficients diagonaux apparaissent plusieurs fois sur la diagonale, il suffit de les prendre comme racines simples de P , c'est-à-dire de prendre

$$P = \prod_{\lambda \in \text{Spec}(D)} (X - \lambda).$$

Démonstration. Nous savons que

$$\forall k \in \mathbf{N}, D^k = \begin{pmatrix} \lambda_1^k & 0 & \dots & 0 \\ 0 & \lambda_2^k & \ddots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & \lambda_n^k \end{pmatrix}.$$

Ainsi, si $P = \sum_{k=0}^d a_k X^k \in \mathbf{K}[X]$, alors

$$P(D) = \sum_{k=0}^d a_k D^k = \begin{pmatrix} \sum_{k=0}^d a_k \lambda_1^k & 0 & \dots & 0 \\ 0 & \sum_{k=0}^d a_k \lambda_2^k & \ddots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & \sum_{k=0}^d a_k \lambda_n^k \end{pmatrix} = \begin{pmatrix} P(\lambda_1) & 0 & \dots & 0 \\ 0 & P(\lambda_2) & \ddots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & P(\lambda_n) \end{pmatrix}.$$

Donc $P(D) = 0$ si et seulement si $P(\lambda_i) = 0, \forall i \in \llbracket 1, n \rrbracket$. □

⚠ Ce résultat n'est plus valable pour les matrices triangulaires. Par exemple si $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$, alors $P = X - 1$ s'annule en 1 (qui est le seul coefficient diagonal de A), et pourtant

$$P(A) = A - I_2 = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \neq \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}.$$

10.1.3 Polynômes annulateurs et valeurs propres

Rappelons le résultat suivant, prouvé dans le chapitre 7 :

Proposition 10.11 : Soit $f \in \mathcal{L}(E), \lambda \in \text{Spec}(f)$ et $x \in E_\lambda(f)$. Alors pour $P \in \mathbf{K}[X]$,

$$[P(f)](x) = P(\lambda) \cdot x.$$

En particulier, ceci implique que si x est un vecteur propre de f pour la valeur propre λ , alors x est également un vecteur propre de $P(f)$, associé à la valeur propre $P(\lambda)$.

Proposition 10.12 : Soit $f \in \mathcal{L}(E)$ et $P \in \mathbf{K}[X]$ un polynôme annulateur de f . Alors toute valeur propre de f est racine de P .

Remarque

Les mêmes résultats restent valables pour une matrice : si P est annulateur de $A \in \mathcal{M}_n(\mathbf{K})$, alors les valeurs propres de A sont parmi les racines de P .

Démonstration. Soit $\lambda \in \text{Spec}(f)$ et soit x un vecteur propre associé à la valeur propre λ . Alors $[P(f)](x) = P(\lambda)x$. Mais $P(f) = 0$ car P est annulateur de f , donc $P(\lambda) \cdot x = 0$. Puisque $x \neq 0$ (c'est un vecteur propre), alors $P(\lambda) = 0 : \lambda$ est racine de P . \square

\triangle La réciproque est fautive : il se peut que certaines racines de P ne soient pas valeurs propres de f . Une des raisons à cela est que si P est un polynôme annulateur de f et si $\lambda \in \mathbf{K}$, alors $(X - \lambda)P$ est encore annulateur de f , et possède λ comme racine. Ainsi, on peut facilement créer un polynôme annulateur avec davantage de racines que P , alors que le spectre de f ne change pas.
Donc les racines d'un polynôme annulateur sont les valeurs propres **possibles** de f .

Exemples 10.13

- Si p est un projecteur de E , alors les seules valeurs propres possibles de p sont les racines de $X^2 - X$ (qui est annulateur). On a donc $\text{Spec}(p) \subset \{0, 1\}$.
- Si $A \in \mathcal{M}_n(\mathbf{R})$ est annihilée par $X^2 + 1$, alors A ne possède pas de valeurs propres réelles (mais peut avoir i et/ou $-i$ comme valeurs propres complexes).

Proposition 10.14 (Cas des matrices 2×2) : Si $A \in \mathcal{M}_2(\mathbf{K})$, alors $\lambda \in \mathbf{K}$ est valeur propre de A si et seulement si λ est racine de $X^2 - \text{tr}(A)X + \det(A)$.

Danger
Ce résultat est spécifique aux matrices 2×2 et à ce polynôme annulateur particulier. Pour un autre polynôme annulateur, on ne peut dire mieux que : si λ est valeur propre, alors λ est racine de P .

Démonstration. Notons $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$. Alors λ est valeur propre de A si et seulement si $A - \lambda I_2$ n'est pas inversible, c'est-à-dire si et seulement si $\det(A - \lambda I_2) = 0$. Or,

$$\det(A - \lambda I_2) = \det \begin{pmatrix} a - \lambda & b \\ c & d - \lambda \end{pmatrix} = (a - \lambda)(d - \lambda) - bc = \lambda^2 - \lambda(a + d) + (ad - bc) = \lambda^2 - \text{tr}(A)\lambda + \det(A).$$

\square

Exemple 10.15

Soit $A = \begin{pmatrix} 10 & 12 \\ -6 & -8 \end{pmatrix}$. Alors $\text{tr}(A) = 2$ et $\det(A) = -8$, donc les valeurs propres de A sont les racines de $X^2 - 2X - 8 = (X + 2)(X - 4) : \text{Spec}(A) = \{-2, 4\}$.

10.1.4 Application des polynômes annulateurs

Proposition 10.16 : Soit $f \in \mathcal{L}(E)$, et soit $P = \sum_{k=0}^d a_k X^k \in \mathbf{K}[X]$ un polynôme annulateur de f , tel que $P(0) \neq 0$ (i.e. tel que $a_0 \neq 0$). Alors f est bijectif et

$$f^{-1} = -\frac{1}{a_0} \sum_{k=1}^d a_k f^{k-1}.$$

En particulier, f^{-1} est un polynôme en f .

De même
Un énoncé similaire est valable pour les matrices.

Démonstration. Si $P(0) \neq 0$, alors 0 n'est pas racine de P , et donc n'est pas valeur propre de f . Ceci signifie donc que f est bijectif. De plus

$$P(f) = 0 \Leftrightarrow -a_0 \text{id}_E = \sum_{k=1}^d a_k f^k = f \circ \left(\sum_{k=1}^d a_k f^{k-1} \right) \Leftrightarrow \text{id}_E = f \circ \left(-\frac{1}{a_0} \sum_{k=1}^d a_k f^{k-1} \right).$$

En composant à gauche par f^{-1} , il vient

$$f^{-1} = -\frac{1}{a_0} \sum_{k=1}^d a_k f^{k-1}.$$

□

Un polynôme annulateur permet également de calculer les puissances d'une matrice en utilisant la division euclidienne de X^n par ce polynôme annulateur.

Exemple 10.17

$$\text{Soit } B = \begin{pmatrix} 1 & 0 & -1 \\ 0 & 0 & 0 \\ 0 & 1 & 1 \end{pmatrix}.$$

On vérifie que $P(X) = X^3 - 2X^2 + X = X(X-1)^2$ est un polynôme annulateur de B . Soit donc $n \in \mathbf{N}$. Alors de manière unique, $X^n = PQ_n + R_n$, où R_n est un polynôme de degré au plus 2.

Notons $R_n = a_n X^2 + b_n X + c$.

Puisque $P(0) = 0$, on a $0^n = P(0)Q_n(0) + R_n(0) = R_n(0)$. Ainsi, $c_n = 0$.

De même, $P(1) = 0$, donc $1^n = a_n + b_n$.

Enfin, $P'(1) = 0$ et donc $nX^{n-1} = P'Q_n + PQ'_n + R'_n$, ce qui nous donne

$$n = \underbrace{P'_n(1)}_{=0} Q_n(1) + \underbrace{P_n(1)}_{=0} Q'_n(1) + R'_n(1) = 2a_n + b_n.$$

On en déduit que $a_n = n - 1$, $b_n = 2 - n$ et $c_n = 0$, de sorte que

$$R_n(X) = (n-1)X^2 + (2-n)X.$$

Et donc

$$B^n = \underbrace{P(B)}_{=0} Q_n(B) + R_n(B) = R_n(B) = (n-1)B^2 + (2-n)B.$$

Puisque $B^2 = \begin{pmatrix} 1 & -1 & -2 \\ 0 & 0 & 0 \\ 0 & 1 & 1 \end{pmatrix}$, il vient donc

$$B^n = \begin{pmatrix} n-1+2-n & 1-n & -2(n-1)+n-2 \\ 0 & 0 & 0 \\ 0 & (n-1)+(2-n) & n-1+2-n \end{pmatrix} = \begin{pmatrix} 1 & 1-n & -n \\ 0 & 0 & 0 \\ 0 & 1 & 1 \end{pmatrix}.$$

10.2 ENDOMORPHISMES ET MATRICES DIAGONALISABLES

10.2.1 Définitions

Définition 10.18 – Soit $f \in \mathcal{L}(E)$. On dit que f est diagonalisable s'il existe une base \mathcal{B} de E telle que $\text{Mat}_{\mathcal{B}}(f)$ soit diagonale.

Une matrice $A \in \mathcal{M}_n(\mathbf{K})$ est dite diagonalisable si elle est semblable à une matrice diagonale.

En particulier

Une matrice diagonale est semblable à elle-même, donc diagonalisable.

Remarque. La matrice de $f \in \mathcal{L}(E)$ dans une base $\mathcal{B} = (e_1, \dots, e_n)$ est diagonale si et seulement si \mathcal{B} est formée de vecteurs propres de f . En effet, on a

$$\text{Mat}_{\mathcal{B}}(f) = \begin{pmatrix} f(e_1) & f(e_2) & \dots & f(e_n) \\ \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_n \end{pmatrix} \begin{matrix} e_1 \\ e_2 \\ \vdots \\ e_n \end{matrix}$$

si et seulement si pour tout $i \in \llbracket 1, n \rrbracket$, $f(e_i) = \lambda_i e_i$.

Exemple 10.19 Diagonalisabilité des projecteurs

On a déjà prouvé précédemment que si p est un projecteur, alors il existe une base dans laquelle la matrice de p est diagonale, avec des 0 et des 1 sur la diagonale. Autrement dit, tout projecteur est diagonalisable.

Nous savons de plus³ que $E_0(p) = \text{Ker } p$ et que $E_1(p) = \{x \in E : p(x) = x\}$.

Mais pour $x = x_1 + x_2$, avec $x_1 \in \text{Im } p$ et $x_2 \in \text{Ker } p$, on a $p(x) = x_1$ et donc $x = p(x)$ si et seulement si $x = x_1 \Leftrightarrow x_2 = 0 \Leftrightarrow x \in \text{Im}(p)$.

Donc le sous-espace propre de p associé à la valeur propre 1 est $\text{Im}(p)$.

En particulier, si p est la projection sur F parallèlement à G , alors $E_0(p) = G$ et $E_1(p) = F$.

³ Comme pour tout endomorphisme.

Proposition 10.20 : Soit $f \in \mathcal{L}(E)$ et soit \mathcal{B} une base de E , $A = \text{Mat}_{\mathcal{B}}(f) \in \mathcal{M}_n(\mathbf{K})$. Alors f est diagonalisable si et seulement si A est diagonalisable.

Démonstration. Supposons que f soit diagonalisable, et soit \mathcal{C} une base de E telle que $D = \text{Mat}_{\mathcal{C}}(f)$ soit diagonale.

Alors si P est la matrice de passage de \mathcal{C} à \mathcal{B} , on a par la formule de changement de base

$$A = P^{-1}DP$$

et donc A est semblable à D qui est diagonale : A est diagonalisable.

Inversement, supposons que A soit diagonalisable, et soient $P \in GL_n(\mathbf{K})$, et $D \in \mathcal{M}_n(\mathbf{K})$ diagonale telles que $A = P^{-1}DP$.

Soit g l'endomorphisme de E dont la matrice dans la base \mathcal{B} est P^{-1} . Si $\mathcal{B} = (e_1, \dots, e_n)$ posons $f_i = g(e_i)$, de sorte que P^{-1} soit la matrice de la famille (f_1, \dots, f_n) dans la base \mathcal{B} . En effet, on a

$$P^{-1} = \begin{pmatrix} g(e_1) & \dots & g(e_n) \\ \star & \dots & \star \\ \vdots & & \vdots \\ \star & \dots & \star \end{pmatrix} \begin{matrix} e_1 \\ \vdots \\ e_n \end{matrix} = \begin{pmatrix} f_1 & \dots & f_n \\ \star & \dots & \star \\ \vdots & & \vdots \\ \star & \dots & \star \end{pmatrix} \begin{matrix} e_1 \\ \vdots \\ e_n \end{matrix} = \text{Mat}_{\mathcal{B}}(f_1, \dots, f_n).$$

Mais P^{-1} étant inversible, la famille (f_1, \dots, f_n) est une base \mathcal{C} , et par définition $P^{-1} = P_{\mathcal{B}, \mathcal{C}}$. Alors par la formule de changement de base, $D = \text{Mat}_{\mathcal{C}}(f)$, et donc f est diagonalisable car sa matrice dans la base \mathcal{C} est diagonale. \square

Remarque. Si f est diagonalisable, et si $D = \text{Mat}_{\mathcal{B}}(f)$ est diagonale, alors les coefficients diagonaux de D sont égaux à ses valeurs propres, et donc aux valeurs propres de f .

Mieux : si $\lambda \in \text{Spec}(f)$, alors $\dim E_{\lambda}(D) = \dim E_{\lambda}(f)$, et donc λ apparaît nécessairement $\dim E_{\lambda}(f)$ fois sur la diagonale de D .

De même, si D est une matrice diagonale semblable à une matrice diagonalisable A , alors les coefficients diagonaux de D sont les valeurs propres de A , chacun apparaissant autant de fois que la dimension du sous-espace propre correspondant de A .

Exemple 10.21

Toute matrice A vérifiant $A^2 = A$ est diagonalisable. En effet, f_A est alors un projecteur de $\mathcal{M}_{n,1}(\mathbf{K})$, donc diagonalisable.

10.2.2 Critères de diagonalisabilité

Théorème 10.22 : Soit $f \in \mathcal{L}(E)$. Alors les assertions suivantes sont équivalentes :

1. f est diagonalisable
2. $\sum_{\lambda \in \text{Spec}(f)} \dim E_\lambda(f) = \dim E$
3. E est somme directe des sous-espaces propres de f
4. Il existe une base de E formée de vecteurs propres de f .

Démonstration. Notons $n = \dim E$. 1) \Rightarrow 2) Soit \mathcal{B} une base de E dans laquelle la matrice D de f est diagonale.

Alors les valeurs propres de f sont les coefficients diagonaux de D , et $\forall \lambda \in \text{Spec}(f)$, la dimension de $E_\lambda(f)$ est égale au nombre de fois où λ apparaît sur la diagonale de D .

En particulier, puisque D possède n coefficients diagonaux, il vient

$$n = \sum_{\lambda \in \text{Spec}(f)} \dim E_\lambda(f).$$

2) \Rightarrow 3) Nous savons déjà que les sous-espaces propres de f sont en somme directe. Mais

$$\dim \left(\bigoplus_{\lambda \in \text{Spec}(f)} E_\lambda(f) \right) = \sum_{\lambda \in \text{Spec}(f)} \dim E_\lambda(f) = \dim E$$

Puisque $\bigoplus_{\lambda \in \text{Spec}(f)} E_\lambda(f) \subset E$, alors

$$E = \bigoplus_{\lambda \in \text{Spec}(f)} E_\lambda(f)$$

3) \Rightarrow 4) Si E est la somme directe des sous-espaces propres de f , alors en concaténant des bases des sous-espaces propres de f , on obtient une base de E . Mais alors une telle base est formée de vecteurs propres de f .

4) \Rightarrow 1) Soit $\mathcal{B} = (e_1, \dots, e_n)$ une base de E formée de vecteurs propres de f .

Notons $\forall i \in \llbracket 1, n \rrbracket$, $f(e_i) = \lambda_i e_i$. Alors la matrice de f dans la base \mathcal{B} est

$$D = \begin{pmatrix} f(e_1) & f(e_2) & \dots & f(e_n) \\ \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \ddots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & \lambda_n \end{pmatrix} \begin{matrix} e_1 \\ e_2 \\ \vdots \\ e_n \end{matrix},$$

et donc f est diagonalisable. □

Corollaire 10.23 : Soit $A \in \mathcal{M}_n(\mathbb{K})$. Alors les assertions suivantes sont équivalentes :

1. A est diagonalisable
2. $\sum_{\lambda \in \text{Spec}(A)} \dim E_\lambda(A) = n$
3. $\mathcal{M}_{n,1}(\mathbb{K})$ est somme directe des sous-espaces propres de A
4. Il existe une base de $\mathcal{M}_{n,1}(\mathbb{K})$ formée de vecteurs propres de A .

Démonstration. Si f_A est l'endomorphisme de $\mathcal{M}_{n,1}(\mathbb{K})$ canoniquement associé⁴ à A , alors les valeurs propres de A sont les valeurs propres de f_A , et les vecteurs propres de A sont ceux de f_A .

Et donc le corollaire s'obtient en appliquant le théorème précédent à f_A . □

Précision

Pour $\lambda \in \text{Spec}(f)$, $E_\lambda(f)$ contient des vecteurs propres, et le vecteur nul. Puisqu'une base ne peut contenir le vecteur nul, une base de $E_\lambda(f)$ est nécessairement formée de vecteurs propres de f (associés à la valeur propre λ).

⁴ **Rappel :** il s'agit de l'endomorphisme $X \mapsto AX$.

Exemple 10.24

Soit $A = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix} \in \mathcal{M}_3(\mathbf{R})$. Alors $A^2 = 3A$, et donc $X^2 - 3X = X(X - 3)$ est un

polynôme annulateur de A .

On en déduit que les seules valeurs propres possibles de A sont 0 et 3.

De plus, $\text{rg}(A) = 1 < 3$ et $\text{rg}(A - 3I_3) = 2 < 3$, de sorte que 0 et 3 sont bien valeurs propres de A : $\text{Spec}(A) = \{0, 3\}$ et

$$\dim E_0(A) = 3 - 1 = 2, \dim E_3(A) = 3 - 2 = 1.$$

Ainsi, A est diagonalisable et donc semblable à une⁵ matrice diagonale.

⁵ Et en fait à plusieurs.

Corollaire 10.25 – Soit $f \in \mathcal{L}(E)$ (respectivement $A \in \mathcal{M}_n(\mathbf{K})$) admettant $\dim E$ (resp. n) valeurs propres distinctes. Alors f (resp. A) est diagonalisable.

Démonstration. Si f admet n valeurs propres distinctes, alors les sous-espaces propres de f sont tous de dimension 1, et donc la somme de leurs dimensions vaut n . f est donc diagonalisable. \square

10.2.3 Pratique de la diagonalisation

En pratique, si une matrice $A \in \mathcal{M}_n(\mathbf{K})$ est diagonalisable, il est intéressant non seulement de connaître une matrice diagonale D à laquelle elle est semblable, mais également une matrice inversible P telle que $A = P^{-1}DP$.

Pour D , nous avons déjà expliqué dans la remarque suivant la proposition 10.20 que ses coefficients diagonaux doivent être les valeurs propres de A , chacun apparaissant autant de fois que la dimension du sous-espace propre associé.

Pour déterminer D , on utilise le fait suivant⁶ : la matrice de f_A dans une base \mathcal{B} est diagonale si et seulement si \mathcal{B} est une base de vecteurs propres de f_A . Ainsi, en notant \mathcal{B}_{can} la base canonique de $\mathcal{M}_{n,1}(\mathbf{K})$ on a

$$A = P_{\mathcal{B}_{can}, \mathcal{B}} \text{Mat}_{\mathcal{B}}(f_A) P_{\mathcal{B}, \mathcal{B}_{can}}.$$

Ainsi, P doit être la matrice de passage d'une base de vecteurs propres à la base canonique.

Méthode : diagonalisation d'une matrice $A \in \mathcal{M}_n(\mathbf{K})$

- Déterminer les valeurs propres de A , à l'aide d'un pivot, ou en utilisant un polynôme annulateur si on en a un.
- Déterminer une base de chacun des sous-espaces propres.
- La somme des dimensions des sous-espaces propres est-elle égale à n ? Si oui, alors A est diagonalisable, sinon ce n'est pas le cas.
- Si A est diagonalisable, soient $\lambda_1, \dots, \lambda_p$ les valeurs propres distinctes. La concaténation d'une base de E_{λ_1} , d'une base de E_{λ_2} , ..., d'une base de E_{λ_p} est alors une base de $\mathcal{M}_{n,1}(\mathbf{K})$ formée de vecteurs propres de A . Si on note P^{-1} la matrice de cette base dans la base canonique (c'est-à-dire la matrice dont les colonnes sont les vecteurs propres de A), on a alors

$$A = P^{-1}DP \text{ avec } D = \text{diag}(\lambda_1, \dots, \lambda_1, \lambda_2, \dots, \lambda_2, \dots, \lambda_p)$$

où chaque valeur propre apparaît autant de fois que la dimension du sous-espace propre associé.

⚠ Danger

Il n'y a pas de réciproque ! Une matrice diagonalisable peut avoir strictement moins de n valeurs propres. Par exemple

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$$

ne possède que deux valeurs propres (-1 et 1), mais est diagonalisable (car diagonale).

⁶ Voir la remarque suivant la définition 10.18.

⚠ Danger !

Attention à ne pas se tromper dans le sens des matrices de passage : $P_{\mathcal{B}_{can}, \mathcal{B}}$ est facile à obtenir si on connaît les vecteurs de \mathcal{B} , alors que $P_{\mathcal{B}, \mathcal{B}_{can}}$ l'est beaucoup moins !

⚠ Certains énoncés demandent de trouver P telle que $A = P^{-1}DP$ (et dans ce cas, P^{-1} est la matrice de passage de la base canonique à une base de vecteurs propres), alors que d'autres notent $A = PDP^{-1}$ (et alors c'est P qui est la matrice de passage de la base canonique

à une base de vecteurs propres). Le plus simple afin d'éviter de les confondre est de réfléchir systématiquement en termes de formule de changement de base.

Exemple 10.26

$$\text{Soit } A = \begin{pmatrix} 4 & 6 & -6 \\ -3 & -5 & 3 \\ 0 & 0 & -2 \end{pmatrix} \in \mathcal{M}_3(\mathbf{R}).$$

$$\text{On a alors } A^2 = \begin{pmatrix} -2 & -6 & 6 \\ 3 & 7 & -3 \\ 0 & 0 & 4 \end{pmatrix}, \text{ de sorte que}$$

$$A^2 + A = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix} = 2I_3.$$

Un polynôme annulateur de A est donc $X^2 + X - 2 = (X + 2)(X - 1)$, de sorte que les seules valeurs propres possibles de A sont -2 et 1 .

Il reste à vérifier si -2 et 1 sont ou non des valeurs propres de A .

$$\text{Mais } A - I_3 = \begin{pmatrix} 3 & 6 & -6 \\ -3 & -6 & 3 \\ 0 & 0 & -3 \end{pmatrix}. \text{ On constate alors que les colonnes de } A - I_3 \text{ sont}$$

liées par la relation

$$C_2 = 2C_1 \Leftrightarrow 2 \cdot C_1 + (-1) \cdot C_2 + 0 \cdot C_3 = 0.$$

Ceci signifie que

$$(A - I_3) \begin{pmatrix} 2 \\ -1 \\ 0 \end{pmatrix} = \begin{pmatrix} 2 \times 3 + (-1) \times 6 + 0 \times (-6) \\ 2 \times (-3) + (-1) \times (-6) + 0 \times 3 \\ 2 \times 0 + (-1) \times 0 + 0 \times (-3) \end{pmatrix} = 2 \cdot C_1 + (-1) \cdot C_2 + 0 \cdot C_3 = 0.$$

Et donc $\text{rg}(A - I_3) = 2$ (car ses colonnes sont liées, et ne sont pas toutes proportionnelles), de sorte que 1 est valeur propre de A , avec $\dim E_1(A) = 3 - 2 = 1$, et que

$\begin{pmatrix} 2 \\ -1 \\ 0 \end{pmatrix}$ est un vecteur propre de A associé à la valeur propre 1 , et donc

$$E_1(A) = \text{Vect} \begin{pmatrix} 2 \\ -1 \\ 0 \end{pmatrix}.$$

De même, on a $A + 2I_3 = \begin{pmatrix} 6 & 6 & -6 \\ -3 & -3 & 3 \\ 0 & 0 & 0 \end{pmatrix}$, qui est de rang 1 car toutes ses colonnes

sont proportionnelles.

On a alors les deux relations $C_1 = C_2 \Leftrightarrow 1 \cdot C_1 + (-1) \cdot C_2 + 0 \cdot C_3 = 0$ et $C_1 = -C_3 \Leftrightarrow 1 \cdot C_1 + 0 \cdot C_2 + 1 \cdot C_3 = 0$, qui nous donnent

$$(A + 2I_3) \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \text{ et } (A + 2I_3) \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

Et donc⁷ $E_{-2}(A) = \text{Vect} \left(\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} \right)$.

Posons alors $P^{-1} = \begin{pmatrix} 1 & 0 & 2 \\ 0 & 1 & -1 \\ 1 & 1 & 0 \end{pmatrix}$ et $D = \begin{pmatrix} -2 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & 1 \end{pmatrix}$. Alors

$$A = P^{-1}DP.$$

Astuce

Nous avons déjà vu au chapitre 7 que si on note E_i le $i^{\text{ème}}$ vecteur de la base canonique de $\mathcal{M}_{n,1}(\mathbf{K})$, alors multiplier une matrice M par E_i donne la $i^{\text{ème}}$ colonne de M .

Et donc ici, si on multiplie

$A - I_3$ par $\begin{pmatrix} 2 \\ -1 \\ 0 \end{pmatrix}$, on obtient le vecteur colonne

$$2 \cdot C_1 - C_2 + 0 \cdot C_3.$$

Cette astuce permet de trouver rapidement des vecteurs propres associés à la valeur propre λ si on arrive à trouver des relations entre les colonnes de $A - \lambda I_n$.

Si de telles relations ne sont pas évidentes, on peut toujours résoudre un système comme nous l'avons fait jusqu'à présent.

⁷ Les deux vecteurs propres obtenus ne sont pas colinéaires, et donc forment une famille libre de E_{-2} , de cardinal égal à $\dim E_{-2}$.

Ordre

Puisqu'on a choisi de commencer par mettre dans P^{-1} les vecteurs propres de A associés à la valeur propre -2 , on commence par mettre les -2 sur la diagonale de D .

On pourrait vérifier ce résultat en calculant P (qui est l'inverse de P^{-1}) et en calculant le produit $P^{-1}AP$.

Remarques. • Il n'y a pas unicité de la matrice diagonale : si on concatène les bases des sous-espaces propres dans un autre ordre, on obtient une autre matrice diagonale.

• Même une fois la matrice D fixée, il n'y a pas unicité de la matrice P : il y en a une infinité⁸ qui vérifient bien $A = P^{-1}DP$.

• Il se peut qu'une matrice à coefficients réels ne soit pas diagonalisable dans $\mathcal{M}_n(\mathbf{R})$, mais qu'elle le soit dans $\mathcal{M}_n(\mathbf{C})$. C'est par exemple le cas de $A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$, qui ne possède pas de valeurs propres réelles, mais qui possède deux valeurs propres complexes i et $-i$, et qui par conséquent est diagonalisable dans $\mathcal{M}_2(\mathbf{C})$.

Autrement dit, il existe une matrice $P \in GL_2(\mathbf{C})$ telle que $A = P^{-1} \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix} P$.

⁸ En effet, il existe une infinité de bases de chaque sous-espace propre, donc une infinité de bases de $\mathcal{M}_{n,1}(\mathbf{K})$ formées de vecteurs propres de A .

10.3 APPLICATIONS DE LA DIAGONALISATION

Nous donnons ici quelques exemples d'applications de la diagonalisation d'une matrice. Si la plupart de ces exemples restent cantonnés à l'algèbre, la diagonalisation a aussi de nombreuses applications dans d'autres domaines. Nous verrons notamment en TP comment l'utiliser⁹ pour étudier des phénomènes probabilistes (chaînes de Markov), et plus tard dans l'année, nous utiliserons la diagonalisation pour étudier la nature des points critiques d'une fonction de plusieurs variables.

Aucun de ces exemples n'est à savoir refaire sans indication, mais tous ces thèmes sont susceptibles d'être abordés dans des sujets de concours.

⁹ En pratique, nous n'aurons besoin que de la notion de vecteur propre.

10.3.1 Inverse d'une matrice diagonalisable

Soit $A \in \mathcal{M}_n(\mathbf{K})$ une matrice diagonalisable inversible. Alors 0 n'est pas valeur propre de A . Soient donc $D = \text{Diag}(\lambda_1, \dots, \lambda_n)$ une matrice diagonale et P une matrice inversible telles que $A = P^{-1}DP$.

Alors $D^{-1} = \text{Diag}\left(\frac{1}{\lambda_1}, \dots, \frac{1}{\lambda_n}\right)$, de sorte que

$$A^{-1} = P^{-1}D^{-1}(P^{-1})^{-1} = P^{-1}\text{Diag}\left(\frac{1}{\lambda_1}, \dots, \frac{1}{\lambda_n}\right)P.$$

Exemple 10.27

$$\text{Soit } A = \begin{pmatrix} 4 & 6 & -6 \\ -3 & -5 & 3 \\ 0 & 0 & -2 \end{pmatrix}.$$

$$\text{Alors } A = P^{-1}\text{Diag}(-2, -2, 1)P, \text{ avec } P^{-1} = \begin{pmatrix} 1 & 0 & 2 \\ 0 & 1 & -1 \\ 1 & 1 & 0 \end{pmatrix}.$$

En particulier, 0 n'est pas valeur propre de A , donc A est inversible, et

$$A^{-1} = P^{-1}\text{Diag}\left(-\frac{1}{2}, -\frac{1}{2}, 1\right)P = \frac{1}{2} \begin{pmatrix} 5 & 6 & -6 \\ -3 & -4 & 3 \\ 0 & 0 & -1 \end{pmatrix}.$$

Remarque

Si l'unique but est de calculer l'inverse de A , cette méthode est bien plus longue qu'un calcul classique d'inverse. En revanche, si pour une raison ou une autre A a déjà été diagonalisée, cette méthode est plus rapide (notamment pour des matrices de grande taille).

Si A est une matrice diagonalisable inversible, telle que $\text{Spec}(A) = \{\lambda_1, \dots, \lambda_p\}$, alors il existe une matrice diagonale D et une matrice inversible P telles que $A = P^{-1}DP$.

Un polynôme annulateur de D est alors $Q(X) = \prod_{i=1}^p (X - \lambda_i)$.

Mais alors Q est également un polynôme annulateur de A car

$$Q(A) = Q(P^{-1}DP) = P^{-1} \underbrace{Q(D)}_{=0} P = 0.$$

Si on écrit Q sous la forme $Q = \sum_{i=0}^p a_i X^i$, alors $a_0 = (-1)^p \prod_{i=1}^p \lambda_i \neq 0$ et donc

$$A^{-1} = -\frac{1}{a_0} \sum_{i=1}^p a_i A^{i-1}.$$

Exemple 10.28

Soit $A = \begin{pmatrix} -1 & 2 & 4 \\ 0 & -2 & 3 \\ 0 & 0 & \frac{1}{2} \end{pmatrix}$. Alors $\text{Spec}(A) = \{\frac{1}{2}, -1, -2\}$, et donc A est diagonalisable

car elle possède trois valeurs propres distinctes.

Un polynôme annulateur de A est alors

$$P = \left(X - \frac{1}{2}\right)(X + 2)(X + 1) = X^3 + \frac{5}{2}X^2 + \frac{1}{2}X - 1.$$

Ainsi, l'inverse de A est donné par

$$A^{-1} = A^2 + \frac{5}{2}A + \frac{1}{2}I_3 = \begin{pmatrix} -1 & -1 & 14 \\ 0 & -\frac{1}{2} & 3 \\ 0 & 0 & 2 \end{pmatrix}.$$

Détails

Pour obtenir l'expression du terme constant a_0 en fonction des λ_i , il faut remarquer que lorsqu'on développe le produit $\prod (X - \lambda_i)$, le seul terme qui ne contient pas de X est celui obtenu par

$$(-\lambda_1) \times (-\lambda_2) \cdots (-\lambda_p).$$

Remarque

Notons que cette méthode ne nécessite pas de calculer P ou P^{-1} .

Elle est plus rapide qu'une inversion par la méthode du pivot uniquement si les valeurs propres sont faciles à trouver (par exemple si l'on dispose d'un polynôme annulateur ou si la matrice est triangulaire).

10.3.2 Calcul des puissances d'une matrice diagonalisable

Soit $A \in \mathcal{M}_n(\mathbf{K})$ une matrice diagonalisable. Alors il existe $P \in GL_n(\mathbf{K})$ telle que $A = P^{-1}DP$,

avec $D = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \ddots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & \lambda_n \end{pmatrix}$ une matrice diagonale.

Alors pour $k \in \mathbf{N}^*$, on a

$$A^k = (P^{-1}DP)^k = P^{-1}D^kP.$$

Mais D étant diagonale, il est très facile de calculer ses puissances :

$$\forall k \in \mathbf{N}, D^k = \begin{pmatrix} \lambda_1^k & 0 & \dots & 0 \\ 0 & \lambda_2^k & \ddots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & \lambda_n^k \end{pmatrix}$$

et donc calculer A^k .

Exemple 10.29

Soit $A = \begin{pmatrix} 4 & 6 & -6 \\ -3 & -5 & 3 \\ 0 & 0 & -2 \end{pmatrix}$. On sait déjà que $A = P^{-1}\text{Diag}(-2, -2, 1)P$, avec

$$P = \begin{pmatrix} -1 & -2 & 2 \\ 1 & 2 & -1 \\ 1 & 1 & -1 \end{pmatrix} \text{ et } P^{-1} = \begin{pmatrix} 1 & 0 & 2 \\ 0 & 1 & -1 \\ 1 & 1 & 0 \end{pmatrix}.$$

Alors $\forall n \in \mathbf{N}$,

$$A^n = P^{-1} \begin{pmatrix} (-2)^n & 0 & 0 \\ 0 & (-2)^n & 0 \\ 0 & 0 & 1^n \end{pmatrix} P$$

$$\begin{aligned}
&= \begin{pmatrix} (-2)^n & 0 & 2 \\ 0 & (-2)^n & -1 \\ (-2)^n & (-2)^n & 0 \end{pmatrix} \begin{pmatrix} -1 & -2 & 2 \\ 1 & 2 & -1 \\ 1 & 1 & -1 \end{pmatrix} \\
&= \begin{pmatrix} -(-2)^n + 2 & (-2)^{n+1} + 2 & -(-2)^{n+1} - 2 \\ (-2)^n - 1 & -(-2)^{n+1} - 1 & -(-2)^n + 1 \\ 0 & 0 & (-2)^n \end{pmatrix}
\end{aligned}$$

Si de plus A est inversible (c'est-à-dire si 0 n'est pas valeur propre de A), alors $A^{-1} = P^{-1}D^{-1}P$, avec

$$D^{-1} = \begin{pmatrix} \frac{1}{\lambda_1} & 0 & \dots & 0 \\ 0 & \frac{1}{\lambda_2} & \ddots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & \frac{1}{\lambda_n} \end{pmatrix}.$$

Ainsi, la formule précédente reste valable si $k \leq 0$:

$$\forall k \in \mathbf{Z}, A^k = P^{-1} \begin{pmatrix} \lambda_1^k & 0 & \dots & 0 \\ 0 & \lambda_2^k & \ddots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & \lambda_n^k \end{pmatrix} P.$$

10.3.3 Matrices diagonalisables semblables

Si A est diagonalisable, alors elle est semblable à toute matrice diagonale D dont les coefficients diagonaux sont les valeurs propres de A , λ apparaissant $\dim E_\lambda(A)$ fois sur la diagonale de D : il suffit d'appliquer la méthode décrite précédemment en faisant varier l'ordre dans lequel on concatène les bases des sous-espaces propres.

Proposition 10.30 : Soient A et B deux matrices **diagonalisables**. Alors A est semblable à B si et seulement si A et B ont les mêmes valeurs propres et des sous-espaces propres de mêmes dimensions.

Démonstration. Nous avons déjà vu que si A et B sont semblables, alors elles ont même valeurs propres et des sous-espaces propres de mêmes dimensions.

Inversement, supposons que $\text{Spec}(A) = \text{Spec}(B) = \{\lambda_1, \dots, \lambda_p\}$ et que

$$\forall i \in \llbracket 1, p \rrbracket, \dim E_{\lambda_i}(A) = \dim E_{\lambda_i}(B) = r_i.$$

Posons $D = \text{diag}(\underbrace{\lambda_1, \dots, \lambda_1}_{r_1 \text{ fois}}, \underbrace{\lambda_2, \dots, \lambda_2}_{r_2 \text{ fois}}, \dots, \underbrace{\lambda_p, \dots, \lambda_p}_{r_p \text{ fois}})$.

Alors il existe $P \in GL_n(\mathbf{K})$ telle que $A = P^{-1}DP$ et il existe $Q \in GL_n(\mathbf{K})$ telle que $B = Q^{-1}DQ$. On a alors

$$A = P^{-1}DP = P^{-1}(QBQ^{-1})P = (Q^{-1}P)^{-1}B(Q^{-1}P), \text{ avec } Q^{-1}P \in GL_n(\mathbf{K}),$$

et donc A et B sont semblables. \square

\triangleleft Nous avons déjà vu que ce résultat est faux si A ou B n'est pas diagonalisable. Pour des matrices non diagonalisables, il est très dur de donner une condition nécessaire et suffisante de similitude, et nous disposons seulement des conditions nécessaires vues précédemment (mêmes rangs, traces, valeurs propres et dimensions des sous-espaces propres).

Exemple 10.31

$$\text{Soit } A = \begin{pmatrix} 1 & 2 & -4 \\ 0 & 2 & -1 \\ 0 & 0 & -3 \end{pmatrix} \text{ et } B = \begin{pmatrix} 2 & -5 & 0 \\ 0 & -3 & 1 \\ 0 & 0 & 1 \end{pmatrix}.$$

Alors A et B ont toutes les deux trois valeurs propres, qui sont 2, 3 et 1, avec des sous-espaces propres de dimension un. Elles sont donc diagonalisables, et ayant mêmes valeurs propres et mêmes dimensions des sous-espaces propres, elles sont semblables.

10.3.4 Suites récurrentes linéaires

En première année, vous avez appris à déterminer le terme général d'une suite récurrente linéaire d'ordre 2, c'est-à-dire vérifiant une relation de récurrence du type

$$u_{n+2} = au_{n+1} + bu_n.$$

La diagonalisation permet, au moins dans certains cas, de généraliser ce résultat à des suites vérifiant des relations de récurrence plus complexes.

Par exemple, soit (u_n) une suite vérifiant :

$$\forall n \in \mathbf{N}, u_{n+3} = u_{n+2} + 4u_{n+1} - 4u_n.$$

Posons alors, pour tout $n \in \mathbf{N}$, $X_n = \begin{pmatrix} u_n \\ u_{n+1} \\ u_{n+2} \end{pmatrix}$. On vérifie aisément que pour tout $n \in \mathbf{N}$,

$$\begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -4 & 4 & 1 \end{pmatrix} X_n = \begin{pmatrix} u_{n+1} \\ u_{n+2} \\ u_{n+2} + u_{n+1} - 4u_n \end{pmatrix} = \begin{pmatrix} u_{n+1} \\ u_{n+2} \\ u_{n+3} \end{pmatrix} = X_{n+1}.$$

Et donc par récurrence, $\forall n \in \mathbf{N}$, $X_n = A^n X_0$, où $A = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -4 & 4 & 1 \end{pmatrix}$.

Un réel λ est valeur propre de A si et seulement si $\text{rg}(A - \lambda I_3) < 3$. Or

$$\begin{aligned} A - \lambda I_3 &= \begin{pmatrix} -\lambda & 1 & 0 \\ 0 & -\lambda & 1 \\ -4 & 4 & 1 - \lambda \end{pmatrix} \xrightarrow{L_1 \leftrightarrow L_3} \begin{pmatrix} -4 & 4 & 1 - \lambda \\ 0 & -\lambda & 1 \\ -\lambda & 1 & 0 \end{pmatrix} \xrightarrow{L_3 \leftarrow 4L_3 - \lambda L_1} \begin{pmatrix} -4 & 4 & 1 - \lambda \\ 0 & -\lambda & 1 \\ 0 & 4(1 - \lambda) & -\lambda(1 - \lambda) \end{pmatrix} \\ &\xrightarrow{L_2 \leftarrow 4L_2 - \lambda L_3} \begin{pmatrix} -4 & 4 & 1 - \lambda \\ 0 & -4 & 4 + \lambda(1 - \lambda) \\ 0 & 4(1 - \lambda) & -\lambda(1 - \lambda) \end{pmatrix} \xrightarrow{L_3 \leftarrow L_3 + (1 - \lambda)L_2} \begin{pmatrix} -4 & 4 & 1 - \lambda \\ 0 & -4 & 4 + \lambda(1 - \lambda) \\ 0 & 0 & -(\lambda^3 - \lambda^2 - 4\lambda + 4) \end{pmatrix}. \end{aligned}$$

Ainsi, les valeurs propres de A sont les solutions de $\lambda^3 - \lambda^2 - 4\lambda + 4 = 0$.

Ce sont donc¹⁰ 1, 2 et -2 . Puisque A possède trois valeurs propres et est une matrice 3×3 , elle est diagonalisable.

¹⁰ 1 est une racine «évidente», et on obtient les deux autres après factorisation par $(\lambda - 1)$.

Et donc il existe une matrice $P \in GL_3(\mathbf{R})$ telle que $A = P^{-1} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & -2 \end{pmatrix} P$.

Alors, pour tout $n \in \mathbf{N}$, $X_n = A^n X_0 = P^{-1} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2^n & 0 \\ 0 & 0 & (-2)^n \end{pmatrix} P X_0$.

Posons $Y_n = P X_n$, de sorte que $Y_n = P A^n X_0 = D^n P X_0 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2^n & 0 \\ 0 & 0 & (-2)^n \end{pmatrix} Y_0$.

Si $Y_0 = \begin{pmatrix} \alpha \\ \beta \\ \gamma \end{pmatrix}$, alors $Y_n = \begin{pmatrix} \alpha \\ 2^n \beta \\ (-2)^n \gamma \end{pmatrix}$.

Et donc $X_n = P^{-1} Y_n$.

Détails

Puisque $A = P^{-1} D P$, on a

$$A^n = P^{-1} D^n P$$

et donc $P A^n = D^n P$.

En notant $P^{-1} = \begin{pmatrix} a & b & c \\ \star & \star & \star \\ \star & \star & \star \end{pmatrix}$, il vient donc

$$\begin{pmatrix} u_n \\ u_{n+1} \\ u_{n+2} \end{pmatrix} = X_n = \begin{pmatrix} a & b & c \\ \star & \star & \star \\ \star & \star & \star \end{pmatrix} \begin{pmatrix} \alpha \\ 2^n \beta \\ (-2)^n \gamma \end{pmatrix} = \begin{pmatrix} a\alpha + 2^n \beta b + (-2)^n c\gamma \\ \star \\ \star \end{pmatrix}.$$

En particulier, il existe trois constantes λ, μ et ν telles que, $\forall n \in \mathbf{N}$, $u_n = \lambda + \mu 2^n + \nu (-1)^n$. Pour calculer ces trois constantes, on pourrait calculer P et P^{-1} afin de connaître les valeurs exactes de a, b, c et α, β, γ .

Mais une fois leur existence acquise, il est plus simple de les calculer en résolvant un système faisant intervenir les 3 premiers termes de la suite, à savoir u_0, u_1 et u_2 .



Exercice : procéder de la même manière pour une suite récurrente linéaire d'ordre 2 vérifiant $u_{n+2} = au_{n+1} + bu_n$ afin de retrouver le théorème vu en première année sur les suites récurrentes linéaire d'ordre 2.

On pourra à cet effet introduire la matrice $A = \begin{pmatrix} 0 & 1 \\ b & a \end{pmatrix}$.

Dans le cas où A ne possède qu'une valeur propre r , on pourra montrer qu'elle est semblable à $B = \begin{pmatrix} r & 1 \\ 0 & r \end{pmatrix}$.

Pour une étude plus générale des suites récurrentes linéaires à l'aide d'outils d'algèbre linéaire, on pourra consulter le¹¹ sujet d'HEC 2009.

Pol. caractéristique

Comme pour les suites récurrentes linéaires d'ordre 2, on pourrait faire intervenir un polynôme caractéristique de la suite.

Ici, ce polynôme serait

$$X^3 - X^2 - 4X + 4.$$

Il possède trois racines qui sont 1, 2 et -2.

Le résultat ci-contre ressemble alors beaucoup au cas des suites récurrentes linéaires d'ordre 2 dont le polynôme caractéristique possède deux racines réelles distinctes.

¹¹ Très beau et très instructif.

MÉTHODES DU CHAPITRE 10

✓ MÉTHODE 10.1 : Exploiter un polynôme annulateur

L'énoncé guidera systématiquement la recherche d'un polynôme annulateur.

Il faudra alors commencer par déterminer ses racines. Les valeurs propres sont alors nécessairement parmi ces racines, mais il se peut que certaines racines ne soient pas des valeurs propres. Il faudra alors le vérifier «à la main» :

- dans le cas d'une matrice, par exemple en calculant $\text{rg}(A - \lambda I_n)$ ou en utilisant toute autre indication
- dans le cas d'un endomorphisme, en trouvant un vecteur propre.

Exercices de TD : 10.6, 10.8, 10.11, 10.17

✓ MÉTHODE 10.2 : Étudier la diagonalisabilité d'une matrice $A \in \mathcal{M}_n(\mathbf{K})$

1. Déterminer les valeurs propres à l'aide de la méthode du pivot, ou en utilisant toute indication fournie par l'énoncé (il peut notamment s'agir d'une question donnant un vecteur $X \in \mathcal{M}_{n,1}(\mathbf{K})$ et demandant de calculer AX . Si $AX = \lambda X$, où $\lambda \in \mathbf{K}$, alors λ est une valeur propre de A .)
2. Calculer la dimension de chacun des sous-espaces propres. Pour cela, on privilégiera la formule $\dim E_\lambda(A) = n - \text{rg}(A - \lambda I_n)$ plutôt qu'une résolution de système.
3. Vérifier si la somme des dimensions des sous-espaces propres vaut n ou non pour conclure.

Exercices de TD : 10.1, 10.2, 10.4, 10.10, 10.11, 10.15, 10.16

✓ MÉTHODE 10.3 : Étudier la diagonalisabilité d'un endomorphisme f

- On peut essayer d'écrire la matrice de f dans une base bien choisie. f est alors diagonalisable si et seulement si sa matrice l'est.
- Si l'énoncé permet d'obtenir un polynôme P annulateur de f , exploiter ce polynôme afin de déterminer les valeurs propres possibles, puis déterminer les sous-espaces propres (éventuellement réduits au vecteur nul dans le cas de racines de P qui ne seraient pas valeurs propres de f) afin d'en obtenir la dimension. On conclut comme pour une matrice en comparant la somme des dimensions des sous-espaces propres à la dimension de l'espace tout entier.

Exercices de TD : 10.7, 10.8, 10.9, 10.12, 10.17, 10.19

EXERCICES DU CHAPITRE 10

EXERCICE 10.1 Déterminer si les matrices suivantes sont diagonalisables, et le cas échéant, les diagonaliser. F

$$A = \begin{pmatrix} -1 & 4 \\ -1 & 3 \end{pmatrix}, B = \begin{pmatrix} -5 & 12 \\ -4 & 9 \end{pmatrix}, C = \begin{pmatrix} 1 & 2 & 3 \\ 0 & 4 & 5 \\ 0 & 0 & 2 \end{pmatrix}, D = \begin{pmatrix} -2 & 1 & 1 \\ 0 & 2 & 4 \\ 0 & 0 & -2 \end{pmatrix}, E = \begin{pmatrix} i+1 & 5-4i & -2i \\ 0 & i & 0 \\ 0 & 2-2i & 1 \end{pmatrix}, F = \begin{pmatrix} -2 & 6 & 2 \\ -2 & 5 & -1 \\ -2 & 3 & 3 \end{pmatrix}, G = \begin{pmatrix} 3 & 1 & -1 \\ 0 & 4 & 0 \\ -1 & 1 & 3 \end{pmatrix}$$

EXERCICE 10.2 Montrer que $A = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}$ n'est pas diagonalisable en tant que matrice de $\mathcal{M}_3(\mathbf{R})$ mais est diagonalisable en tant que matrice de $\mathcal{M}_3(\mathbf{C})$. F

EXERCICE 10.3 Soit $A \in \mathcal{M}_n(\mathbf{K})$ ne possédant qu'une seule valeur propre. Donner une condition nécessaire et suffisante sur A pour que A soit diagonalisable. F

EXERCICE 10.4 Soit la matrice $A = \begin{pmatrix} 1 & 1 & a \\ 0 & 2 & 0 \\ 0 & 0 & a \end{pmatrix} \in \mathcal{M}_3(\mathbf{R})$, avec $a \in \mathbf{R}$. PD

- 1) Quel est le rang de A ? (On pourra distinguer plusieurs cas suivant les valeurs de a).
- 2) A est-elle diagonalisable ?

EXERCICE 10.5 Soit $A = \begin{pmatrix} -7 & 10 \\ -5 & 8 \end{pmatrix}$. PD

- 1) Montrer que A est diagonalisable, et déterminer D diagonale et P inversible telles que $A = P^{-1}DP$.
- 2) En déduire la valeur de A^n , $\forall n \in \mathbf{N}$.

EXERCICE 10.6 Un endomorphisme nilpotent PD

Soit $n \geq 2$, et soit f l'endomorphisme de $\mathbf{R}_n[X]$ défini par $f(P) = P'$.

- 1) Montrer que $f^{n+1} = 0$ et $f^n \neq 0$ (on dit que f est nilpotent d'indice $n+1$).
- 2) En déduire les valeurs propres de f . Est-il diagonalisable ?
- 3) Soit $g : P \mapsto P - P'$. Exprimer g en fonction de f , et en déduire que g est bijectif.
- 4) Montrer que $g^{-1} = \text{id} + f + f^2 + \dots + f^n$.

EXERCICE 10.7 Soit $\alpha \in \mathbf{R}^*$, et soit f l'application définie sur $\mathbf{R}_n[X]$ par $f(P) = (X - \alpha)P'$. PD

- 1) Montrer que f est un endomorphisme de $\mathbf{R}_n[X]$.
- 2) f est-il diagonalisable ?

EXERCICE 10.8 Soient a, b deux réels, avec $a \neq 0$ et soit u l'endomorphisme de $\mathbf{R}_n[X]$ défini par $u(P) = P(aX + b)$. AD

- 1) Déterminer les valeurs propres de u .
- 2) Montrer que si $a \notin \{-1, 1\}$, alors u est diagonalisable.
- 3) Si $a = 1$, quelles sont les valeurs de b pour lesquelles u est diagonalisable ?
- 4) On suppose à présent que $a = -1$.
 - a) Pour $P \in \mathbf{R}_n[X]$, calculer $u^2(P)$. En déduire un polynôme annulateur de u , et retrouver alors les valeurs propres de u .
 - b) Montrer que $p = \frac{\text{id} + u}{2}$ est un projecteur de $\mathbf{R}_n[X]$. En déduire que u est diagonalisable.

EXERCICE 10.9 (D'après EDHEC 2005) PD

- 1) Déterminer la dimension de $\text{Ker}(\text{tr}) = \{M \in \mathcal{M}_n(\mathbf{R}) : \text{tr}(M) = 0\}$. En déduire que $\mathcal{M}_n(\mathbf{R}) = \text{Ker}(\text{tr}) \oplus \text{Vect}(I_n)$.
- 2) Soit f l'application définie sur $\mathcal{M}_n(\mathbf{R})$ par $f(M) = M + \text{tr}(M)I_n$.
 - a) Montrer que f est un endomorphisme de $\mathcal{M}_n(\mathbf{R})$.
 - b) En utilisant la question 1, montrer que f est diagonalisable et déterminer ses valeurs propres.

EXERCICE 10.10 Une suite récurrente linéaire d'ordre 3

Soit (u_n) la suite définie par $u_0 = 1, u_1 = 0, u_2 = 2$ et $\forall n \in \mathbf{N}, u_{n+3} = 2u_{n+2} + u_{n+1} - 2u_n$.

On pose, pour $n \in \mathbf{N}, X_n = \begin{pmatrix} u_n \\ u_{n+1} \\ u_{n+2} \end{pmatrix}$.

- 1) Montrer qu'il existe une matrice $A \in \mathcal{M}_3(\mathbf{R})$ telle que $\forall n \in \mathbf{N}, X_{n+1} = AX_n$.
- 2) En déduire que $\forall n \in \mathbf{N}, X_n = A^n X_0$.
- 3) Montrer que A est diagonalisable.
- 4) En déduire qu'il existe trois réels λ, μ, ν tels que $\forall n \in \mathbf{N}, u_n = \lambda + \mu(-1)^n + \nu 2^n$.
- 5) Déterminer le terme général de (u_n) .

EXERCICE 10.11 Soit $A = \begin{pmatrix} -5 & 6 & 6 \\ -6 & 7 & 6 \\ 3 & -3 & -2 \end{pmatrix}$.

- 1) Calculer A^2 . En déduire un polynôme annulateur de A de degré 2.
- 2) Déterminer le reste de la division euclidienne de X^n par $X^2 + X - 2$, et en déduire A^n , pour tout $n \in \mathbf{N}$.
- 3) Montrer que A est diagonalisable, et déterminer D diagonale et P inversible telles que $A = P^{-1}DP$. Retrouver la valeur de A^n .

EXERCICE 10.12 Diagonalisation des endomorphismes de rang 1

Soit E un espace vectoriel de dimension finie n , et soit $f \in \mathcal{L}(E)$ un endomorphisme de rang 1. Soit \mathcal{B} une base de E et $A = \text{Mat}_{\mathcal{B}}(f)$.

- 1) Montrer qu'il existe une base de E dans laquelle la matrice de f est de la forme $M = \begin{pmatrix} 0 & \dots & 0 & \alpha_1 \\ \vdots & & \vdots & a_2 \\ \vdots & & \vdots & \vdots \\ 0 & \dots & 0 & \alpha_n \end{pmatrix}$.
- 2) Montrer que M est diagonalisable si et seulement si $\alpha_n \neq 0$.
- 3) En déduire que f est diagonalisable si et seulement si $\text{tr}(A) \neq 0$.

EXERCICE 10.13 Racines carrées d'une matrice diagonalisable

Soit $C = \begin{pmatrix} 9 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 1 \end{pmatrix}$.

- 1) Montrer que si $D \in \mathcal{M}_3(\mathbf{R})$ commute avec C , alors D est diagonale. En déduire que si $D^2 = C$, alors D est diagonale.
- 2) Déterminer le nombre d'éléments $D \in \mathcal{M}_3(\mathbf{R})$ tels que $D^2 = C$.

- 3) Soit $A = \begin{pmatrix} 9 & 0 & 0 \\ -5 & 4 & 0 \\ 8 & 0 & 1 \end{pmatrix}$. Déterminer le nombre d'éléments B de $\mathcal{M}_3(\mathbf{R})$ tels que $B^2 = A$.

EXERCICE 10.14 Soient $a, b \in \mathbf{R}$. Dans $\mathcal{M}_n(\mathbf{R})$ on pose

$$M = \begin{pmatrix} a & b & \dots & b \\ b & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & b \\ b & \dots & b & a \end{pmatrix} \text{ et } J = \begin{pmatrix} 1 & 1 & \dots & 1 \\ 1 & \ddots & \ddots & 1 \\ \vdots & \ddots & \ddots & 1 \\ 1 & \dots & 1 & 1 \end{pmatrix}.$$

- 1) Calculer JV , où $V \in \mathcal{M}_{n,1}(\mathbf{R})$ est le vecteur colonne dont tous les coefficients valent 1. En déduire que J est diagonalisable et déterminer une matrice diagonale semblable à J .
- 2) Écrire M comme combinaison linéaire de I_n et J . En déduire que M est diagonalisable, et déterminer une matrice diagonale semblable à M .

EXERCICE 10.15 Résolution d'une équation matricielle

Soit $A = \begin{pmatrix} 4 & 4 & -4 \\ 2 & 2 & 4 \\ 4 & 4 & 2 \end{pmatrix}$.

On cherche dans cet exercice à trouver toutes les matrices $M \in \mathcal{M}_3(\mathbf{R})$ telles que $M^2 + M = A$ (★).

- 1) Déterminer les valeurs propres et les vecteurs propres de A .
- 2) Soit M une solution de (\star) .
 - a) Vérifier que $AM = MA$.
 - b) Soit λ une valeur propre de A et X un vecteur propre associé. Montrer que $MX \in E_\lambda(A)$, et en déduire qu'il existe $\mu \in \mathbf{R}$ tel que $MX = \mu X$. Donner une relation entre λ et μ .
- 3) Déterminer toutes les solutions de (\star) , et donner l'unique solution dont toutes les valeurs propres sont positives.

EXERCICE 10.16 Soient X et Y deux variables aléatoires indépendantes sur (Ω, \mathcal{A}, P) suivant la loi géométrique $\mathcal{G}(p)$, $p \in]0, 1[$.

Déterminer la probabilité que la matrice $\begin{pmatrix} X & 0 \\ X+Y & Y \end{pmatrix}$ soit diagonalisable.

EXERCICE 10.17 (QSP ESCP 2015)

Soit f un endomorphisme d'un espace vectoriel E de dimension finie $n \geq 1$, et a et b deux réels distincts. On suppose que

$$(f - \text{aid}_E)^3 \circ (f - \text{bid}_E) = 0 \text{ et } (f - \text{aid}_E)^2 \circ (f - \text{bid}_E) \neq 0.$$

Étudier la diagonalisabilité de f .

EXERCICE 10.18 Le sujet d'oral d'un ancien de Fauriel (QSP HEC 2016)

Soit f l'endomorphisme de $\mathcal{M}_n(\mathbf{R})$ défini par $f(M) = M + 2^t M + 3\text{tr}(M)I_n$. f est-il diagonalisable ?

EXERCICE 10.19 (QSP HEC 2016)

Soit A une matrice de $\mathcal{M}_n(\mathbf{R})$. On considère l'application φ_A qui à toute matrice $M \in \mathcal{M}_n(\mathbf{R})$ associe la matrice $\varphi_A(M) = AM$.

- 1) Comparer les spectres de A et φ_A .
- 2) On suppose que A est diagonalisable. Montrer que φ_A est diagonalisable.

EXERCICE 10.20 Diagonalisation simultanée de deux endomorphismes (Oral ESCP 2016)

Soit E un \mathbf{R} -espace vectoriel de dimension finie et f un endomorphisme de E diagonalisable. On note $\{\lambda_1, \dots, \lambda_p\}$ l'ensemble de ses valeurs propres et E_1, \dots, E_p les sous-espaces propres associés.

Soit F un sous-espace vectoriel de E stable par f , tel que $F \neq \{0\}$ et $F \neq E$. Soit x un vecteur de F .

- 1) Montrer qu'il existe un unique p -uplet $(x_1, \dots, x_p) \in E_1 \times \dots \times E_p$ tel que $x = x_1 + \dots + x_p$.
- 2) On suppose désormais $x \neq 0$. Montrer que, quitte à modifier l'ordre, on peut supposer qu'il existe $r \in \llbracket 1, p \rrbracket$ tel que $x_i = 0$ pour $i > r$ et $x_i \neq 0$ pour $i \leq r$. On a alors $x = x_1 + \dots + x_r$.
On note V_x le sous-espace vectoriel engendré par (x_1, \dots, x_r) .
- 3)
 - a) Montrer que (x_1, \dots, x_r) est une base de V_x .
 - b) Montrer que pour tout $j \geq 0$, $f^j(x) \in V_x$.
 - c) Déterminer la matrice A de la famille $(x, f(x), \dots, f^{r-1}(x))$ dans la base (x_1, \dots, x_r) de V_x .

Notons C_1, \dots, C_r les colonnes de A et $\alpha_1, \dots, \alpha_r$ des réels tels que $\sum_{j=1}^r \alpha_j C_j = 0$. Montrer que le polynôme

$$P = \sum_{j=1}^r \alpha_j X^{j-1} \text{ est le polynôme nul. En déduire que } A \text{ est inversible.}$$

- d) Montrer que pour tout $i \in \llbracket 1, p \rrbracket$, $x_i \in F$, puis que $F = \bigoplus_{i=1}^p (F \cap E_i)$.

- 4) Soit g un endomorphisme de E , diagonalisable et commutant avec f (i.e. tel que $f \circ g = g \circ f$). Montrer qu'il existe une base de E formée de vecteurs propres communs à f et g .

CORRECTION DES EXERCICES DU CHAPITRE 10

SOLUTION DE L'EXERCICE 10.1

1. Un réel λ est valeur propre de A si et seulement si $A - \lambda I_2$ n'est pas inversible, soit si et seulement si $\det(A - \lambda I_2) = 0 \Leftrightarrow \lambda^2 - 2\lambda + 1 = 0 \Leftrightarrow (\lambda - 1)^2 = 0$.

Donc A possède une unique valeur propre 1. Puisque $\dim E_1(A) = 2 - \text{rg}(A - I_2) = 2 - \text{rg} \begin{pmatrix} -2 & 4 \\ -1 & 2 \end{pmatrix} = 1$, elle n'est pas diagonalisable.

2. Les valeurs propres de B sont les racines de $X^2 - 4X + 3 = (X - 1)(X - 3)$. Donc B possède deux valeurs propres, et par conséquent est diagonalisable. On a, par exemple¹,

$$B = \begin{pmatrix} -3 & 2 \\ -2 & 1 \end{pmatrix} \begin{pmatrix} 3 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & -2 \\ 2 & -3 \end{pmatrix}$$

3. C est triangulaire et possède trois coefficients diagonaux distincts. Donc elle possède trois valeurs propres distinctes : 1, 2 et 4, et donc est diagonalisable. On a, par exemple

$$C = P^{-1} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 4 \end{pmatrix} P \text{ avec } P^{-1} = \begin{pmatrix} 1 & 4 & 2 \\ 0 & 5 & 3 \\ 0 & -2 & 0 \end{pmatrix}$$

4. D étant triangulaire, ses valeurs propres sont sur sa diagonale : on a $\text{Spec}(D) = \{2, -2\}$, et nécessairement $\dim E_2(D) = 1$.

De plus, on a alors $\dim E_{-2}(D) = 3 - \text{rg}(D + 2I_3) = 2$.

Donc D est diagonalisable car $\dim E_{-2}(D) + \dim E_2(D) = 3$. On a alors $E_{-2}(D) = \text{Vect} \left(\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix} \right)$

$$\text{et } E_2(D) = \text{Vect} \left(\begin{pmatrix} 1 \\ 4 \\ 0 \end{pmatrix} \right).$$

5. Un complexe λ est valeur propre de E si et seulement si $\text{rg}(E - \lambda I_3) < 3$.

$$E - \lambda I_3 = \begin{pmatrix} i+1-\lambda & 5-4i & -2i \\ 0 & i-\lambda & 0 \\ 0 & 2-2i & 1-\lambda \end{pmatrix} \xleftrightarrow{L_2 \leftrightarrow L_3} \begin{pmatrix} i+1-\lambda & 5-4i & -2i \\ 0 & 2-2i & 1-\lambda \\ 0 & i-\lambda & 0 \end{pmatrix} \xleftrightarrow{L_3 \leftarrow (2-2i)L_3 - (i-\lambda)L_2} \begin{pmatrix} i+1-\lambda & 5-4i & -2i \\ 0 & 2-2i & 1-\lambda \\ 0 & 0 & -(i-\lambda)(1-\lambda) \end{pmatrix}.$$

Ainsi, F possède trois valeurs propres qui sont $1+i$, i et 1.

En particulier, elle est diagonalisable.

$$\text{On a alors } E_{1+i}(E) = \text{Vect} \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, E_i(E) = \text{Vect} \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix} \text{ et } E_1(E) = \text{Vect} \begin{pmatrix} -2 \\ 0 \\ 1 \end{pmatrix}.$$

$$\text{Donc } E = P^{-1}DP \text{ avec } P^{-1} = \begin{pmatrix} 1 & 1 & -2 \\ 0 & 1 & 0 \\ 0 & 2 & 1 \end{pmatrix} \text{ et } D = \begin{pmatrix} 1+i & 0 & 0 \\ 0 & i & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

6. F n'est pas diagonalisable car possède une unique valeur propre² 2 et $\dim E_2(F) = 3 - \text{rg}(F - 2I_3) = 3 - 2 = 1$.

7. G est diagonalisable avec $\text{Spec}(G) = \{2, 4\}$ et $E_2(G) = \text{Vect} \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$ et $E_4(G) = \text{Vect} \left(\begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix} \right)$.

On a donc $G = P^{-1}DP$ avec

$$P^{-1} = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 0 & 2 \\ 1 & -1 & 1 \end{pmatrix}, P = \frac{1}{2} \begin{pmatrix} 1 & -1 & 1 \\ 1 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}, D = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 4 \end{pmatrix}$$

SOLUTION DE L'EXERCICE 10.2

$\lambda \in \mathbb{C}$ est valeur propre de A si et seulement si $\text{rg}(A - \lambda I_3) < 3$.

$$A - \lambda I_3 = \begin{pmatrix} 2-\lambda & 0 & 0 \\ 0 & -\lambda & -1 \\ 0 & 1 & -\lambda \end{pmatrix} \xleftrightarrow{L_2 \leftrightarrow L_3} \begin{pmatrix} 2-\lambda & 0 & 0 \\ 0 & 1 & -\lambda \\ 0 & -\lambda & -1 \end{pmatrix} \xleftrightarrow{L_3 \leftarrow L_3 + \lambda L_2} \begin{pmatrix} 2-\lambda & 0 & 0 \\ 0 & 1 & -\lambda \\ 0 & 0 & -1-\lambda^2 \end{pmatrix}.$$

¹ Si on choisit une autre base de vecteurs propres de B , on obtient une autre matrice P , mais qui vérifie également

$$B = P^{-1}DP.$$

On pourrait également changer l'ordre des deux valeurs propres dans la matrice diagonale.

² Réelle. Il y a en fait deux valeurs propres complexes conjuguées.

Donc les valeurs propres de A sont les complexes λ vérifiant $2 - \lambda = 0$ ou $-1 - \lambda^2 = 0$, soit $\lambda = 2$ ou $\lambda = i$ ou $\lambda = -i$.

Donc si on voit A comme une matrice complexe, elle possède trois valeurs propres, et donc est diagonalisable.

En revanche, elle ne possède qu'une valeur propre réelle, qui est 2, et

$$\dim E_2(A) = 3 - \operatorname{rg}(A - 2I_3) = 3 - \operatorname{rg} \begin{pmatrix} 0 & 0 & 0 \\ 0 & -2 & -1 \\ 0 & -1 & -2 \end{pmatrix} = 3 - 2 = 1.$$

Donc $\sum_{\lambda \in \operatorname{Spec}(A)} \dim E_\lambda(A) = 1 < 3$: A n'est pas diagonalisable dans $\mathcal{M}_3(\mathbf{R})$.

SOLUTION DE L'EXERCICE 10.3

Si A possède λ comme unique valeur propre, alors $\dim E_\lambda(A) = n$, et alors A est semblable à une matrice diagonale, qui possède n fois λ sur sa diagonale. Autrement dit, A est semblable à λI_n . Mais la seule matrice semblable à λI_n est λI_n . On en déduit que A est diagonalisable si et seulement si $A = \lambda I_n$.

SOLUTION DE L'EXERCICE 10.4

1. Si $a \neq 0$, alors A est échelonnée, car triangulaire et à coefficients diagonaux non nuls. Donc elle est de rang 3.

Si $a = 0$, alors A est encore échelonnée, mais cette fois ne possède que deux pivots. Donc elle est de rang 2.

2. Si $a \notin \{1, 2\}$, alors A possède trois valeurs propres distinctes³, et donc est diagonalisable.

Si $a = 1$, alors $A = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{pmatrix}$. On a donc $\operatorname{Spec}(A) = \{1, 2\}$ et $\dim E_2(A) = 1$ et $\dim E_1(A) =$

$3 - \operatorname{rg}(A - I_3) = 3 - 2 = 1$. Donc A n'est pas diagonalisable car

$$\dim E_1(A) + \dim E_2(A) = 2 \neq 3.$$

Enfin, dans le cas où $a = 2$, alors $A = \begin{pmatrix} 1 & 1 & 2 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}$.

Alors $\dim E_1(A) = 1$, et $\dim E_2(A) = 3 - \operatorname{rg}(A - 2I_3) = 2$ et donc A est diagonalisable car

$$\dim E_1(A) + \dim E_2(A) = 3.$$

SOLUTION DE L'EXERCICE 10.5

1. Un réel λ est valeur propre de A si et seulement si $A - \lambda I_2$ n'est pas inversible, soit si et seulement si $\det(A - \lambda I_2) = 0 \Leftrightarrow \lambda^2 - \lambda - 6 = 0 \Leftrightarrow (\lambda + 2)(\lambda - 3) = 0$.

Ainsi, A possède deux valeurs propres distinctes, qui sont 2 et -3 , et par conséquent, est diagonalisable.

On a alors $E_{-2}(A) = \operatorname{Vect} \left(\begin{pmatrix} 2 \\ 1 \end{pmatrix} \right)$ et $E_3(A) = \operatorname{Vect} \left(\begin{pmatrix} 1 \\ 1 \end{pmatrix} \right)$. Donc si on pose $P^{-1} = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}$ et

$D = \begin{pmatrix} -2 & 0 \\ 0 & 3 \end{pmatrix}$, alors

$$A = P^{-1}DP \text{ avec } P = \begin{pmatrix} 1 & -1 \\ -1 & 2 \end{pmatrix}.$$

2. Pour tout entier n , on a

$$\begin{aligned} A^n &= P^{-1}D^nP = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} (-2)^n & 0 \\ 0 & 3^n \end{pmatrix} \begin{pmatrix} 1 & -1 \\ -1 & 2 \end{pmatrix} = \begin{pmatrix} -(-2)^{n+1} & 3^n \\ (-2)^n & 3^n \end{pmatrix} \begin{pmatrix} 1 & -1 \\ -1 & 2 \end{pmatrix} \\ &= \begin{pmatrix} -(-2)^{n+1} - 3^n & (-2)^{n+1} + 2 \times 3^n \\ (-2)^n - 3^n & -(-2)^n + 2 \times 3^n \end{pmatrix} \end{aligned}$$

SOLUTION DE L'EXERCICE 10.6

1. Il est clair que pour tout k , f^k est l'application $P \mapsto P^{(k)}$.

Or, nous savons que pour $P \in \mathbf{R}_n[X]$, $\deg P' \leq \deg P - 1$, $\deg P'' \leq \deg P - 2$, etc, $\deg P^{(n+1)} \leq \deg P - n - 1 \leq n - n - 1 \leq -1$.

Et donc⁴ $f^{(n+1)}(P) = 0$.

Ceci étant vrai pour tout $P \in \mathbf{R}_n[X]$, on a donc $f^{n+1} = 0$.

En revanche, $f^n(X^n) = n! \neq 0$, et donc $f^n \neq 0$.

Détails

Cela signifie qu'il existe une matrice diagonale D à coefficients complexes et une matrice $P \in GL_3(\mathbf{C})$ tels que $A = P^{-1}DP$.

Complexes/réels

Lorsqu'on considère une matrice réelle, sans plus de précisions, on s'intéressera aux valeurs propres réelles et au fait de savoir si elle est diagonalisable dans $\mathcal{M}_n(\mathbf{R})$, c'est-à-dire de chercher D diagonale et P inversible, à coefficients réels tels que $A = P^{-1}DP$.

Rappel

Il n'y a pas de calculs à faire pour obtenir la dimension de $E_2(A)$: il n'y a qu'un deux sur la diagonale, donc le sous-espace propre E_2 est de dimension 1

Astuce

Il est facile de vérifier le résultat en prenant $n = 0$ (on doit alors trouver I_2) et $n = 1$ (on doit alors retrouver A). Si cela marche pour $n = 0$ et $n = 1$, il est fort probable que l'expression trouvée soit la bonne.

⁴ Un polynôme de degré négatif est nécessairement le polynôme nul, qui, par convention, est de degré égal à $-\infty$.

2. Nous venons de prouver que X^{n+1} est un polynôme annulateur de f . Et donc 0 est la seule valeur propre possible de f .
 D'autre part, nous savons que f n'est pas bijectif, car sinon f^{n+1} le serait également⁵, ce qui n'est clairement pas le cas.
 Donc $\text{Spec}(f) = \{0\}$.
 Notons que $E_0(f) = \text{Ker } f = \{P \in \mathbf{R}_n[X] : P' = 0\} = \mathbf{R}_0[X]$, l'ensemble des polynômes constants.
 Donc $\dim E_0(f) < \dim \mathbf{R}_{n+1}[X] : f$ n'est pas diagonalisable.
3. On a $g = \text{id} - f$. Puisque 1 n'est pas valeur propre de f , $\text{id} - f$ est bijectif.
 On a alors

$$\begin{aligned} g \circ (\text{id} + f + f^2 + \dots + f^n) &= (\text{id} - f) \circ (\text{id} + f + \dots + f^n) \\ &= \text{id} + f + f^2 + \dots + f^n - f - f^2 - \dots - f^n - \underbrace{f^{n+1}}_{=0} \\ &= \text{id}. \end{aligned}$$

Et donc $\text{id} + f + \dots + f^n$ est bien l'inverse de g .

SOLUTION DE L'EXERCICE 10.7

1. Pour $P \in \mathbf{R}_n[X]$, on a $\deg P' \leq n - 1$ et donc $\deg f(P) \leq n$, de sorte que $f(P) \in \mathbf{R}_n[X]$.
 Soient $P, Q \in \mathbf{R}_n[X]$ et $\lambda \in \mathbf{R}$. Alors

$$f(\lambda P + Q) = (X - \alpha)(\lambda P + Q)' = \lambda(X - \alpha)P' + (X - \alpha)Q' = \lambda f(P) + f(Q).$$

Donc f est linéaire, c'est un endomorphisme de $\mathbf{R}_n[X]$.

2. On a $f(1) = 0$ et pour $k \geq 1$, $f(X^k) = k(X - \alpha)X^{k-1}$.
 Ainsi, la matrice de f dans la base canonique de $\mathbf{R}_n[X]$ est

$$\text{Mat}_{\mathcal{B}_{can}}(f) = \begin{pmatrix} f(1) & f(X) & f(X^2) & \dots & f(X^{n-1}) & f(X^n) \\ 0 & -\alpha & 0 & \dots & \dots & 0 \\ \vdots & 1 & -2\alpha & & & \vdots \\ \vdots & \ddots & 2 & \ddots & & \vdots \\ \vdots & & \ddots & \ddots & \ddots & \vdots \\ \vdots & & & \ddots & n-1 & -n\alpha \\ 0 & \dots & \dots & \dots & 0 & n \end{pmatrix} \begin{matrix} 1 \\ X \\ X^2 \\ \vdots \\ X^{n-1} \\ X^n \end{matrix}.$$

Cette matrice est triangulaire, donc ses valeurs propres sont sur sa diagonale. On en déduit que $\text{Spec}(f) = \llbracket 0, n \rrbracket$.

En particulier, f possède $n + 1$ valeurs propres, et $n + 1 = \dim \mathbf{R}_n[X]$. Donc f est diagonalisable.

SOLUTION DE L'EXERCICE 10.8

1. On a $u(1) = 1$, $u(X) = aX + b$, et plus généralement,

$$u(X^k) = (aX + b)^k = \sum_{i=0}^k \binom{k}{i} a^i X^i b^{k-i} = a^k X^k + k a^{k-1} b X^{k-1} + \dots + b^k.$$

Et donc la matrice de u dans la base canonique de $\mathbf{R}_n[X]$ est

$$\text{Mat}_{\mathcal{B}_{can}}(u) = \begin{pmatrix} f(1) & f(X) & f(X^2) & \dots & f(X^{n-1}) & f(X^n) \\ 1 & b & b^2 & \dots & \dots & b^n \\ \vdots & a & 2ab & & & \vdots \\ \vdots & \ddots & a^2 & \ddots & & \vdots \\ \vdots & & \ddots & \ddots & \ddots & \vdots \\ \vdots & & & \ddots & a^{n-1} & na^{n-1}b \\ 0 & \dots & \dots & \dots & 0 & a^n \end{pmatrix} \begin{matrix} 1 \\ X \\ X^2 \\ \vdots \\ X^{n-1} \\ X^n \end{matrix}.$$

⁵ Comme composée d'applications bijectives.

Remarque

Ce résultat est à mettre en parallèle de l'identité classique

$$1 - a^p = (1 - a) \times (1 + a + \dots + a^{p-1})$$

qui reste valable pour un endomorphisme :

$$\text{id} - f^p = (\text{id} - f) \circ (\text{id} + \dots + f^{p-1}).$$

En particulier, lorsque $f^p = 0$, alors on retrouve l'inverse de $(\text{id} - f)$.

⚠ Danger !

Bien penser en rédigeant une telle question qu'un polynôme de $\mathbf{R}_n[X]$ n'est pas nécessairement de degré n , mais de degré au plus n .

Cette matrice étant triangulaire, ses valeurs propres sont ses coefficients diagonaux : $\text{Spec}(u) = \{a^k, k \in \llbracket 0, n \rrbracket\}$.

2. Si $a \notin \{-1, 1\}$, alors les a^k sont deux à deux distincts, et donc u possède $n + 1 = \dim \mathbf{R}_n[X]$ valeurs propres distinctes : elle est diagonalisable.

3. Si $a = 1$, alors u ne possède qu'une valeur propre, qui est 1.

Il est donc diagonalisable si et seulement si $\dim E_1(u) = n + 1$, soit encore si et seulement si $E_1(u) = \mathbf{R}_n[X]$.

C'est le cas si et seulement si pour tout $P \in \mathbf{R}_n[X]$, $u(P) = P$. En particulier, pour $P = X$, il faut alors que $X + b = X$, donc que $b = 0$.

Et inversement, si $b = 0$, alors u est l'identité, qui est bien entendu⁶ diagonalisable.

Donc si $a = 1$, u est diagonalisable si et seulement si $b = 0$.

4.a. Si $a = -1$, alors u possède deux valeurs propres qui sont 1 et -1 .

Commençons par calculer $u^2(P)$ sur un exemple pour bien comprendre ce qu'il se passe.

Prenons donc $P(X) = X^2 - 2X + 3$. Il vient alors $u(P) = (-X + b)^2 - 2(-X + b) + 3$.

Et donc en appliquant encore une fois P , ce qui revient à remplacer tous les X par des $-X + b$, il vient

$$u^2(P) = (-(-X + b) + b)^2 - 2(-(-X + b) + b) + 3 = P.$$

Et plus généralement, pour $P \in \mathbf{R}_n[X]$, on a $u^2(P) = u(u(P)) = P(-(-X + b) + b) = P(X - b + b) = P(X)$.

Autrement dit, $u^2 = \text{id}_{\mathbf{R}_n[X]}$. Et donc un polynôme annulateur de u est $X^2 - 1$, qui possède 1 et -1 comme racines, qui sont donc bien les deux seules valeurs propres possibles de u .

4.b. On a

$$p^2 = \left(\frac{\text{id} + u}{2}\right)^2 = \frac{1}{4}(\text{id} + u)^2 = \frac{1}{4}(u^2 + 2u + \text{id}) = \frac{1}{4}(2\text{id} + 2u) = p.$$

Et donc p est un projecteur.

En particulier, p est diagonalisable : il existe une base \mathcal{B} de $\mathbf{R}_n[X]$ dans laquelle $\text{Mat}_{\mathcal{B}}(p)$ est diagonale.

Et alors, dans cette même base, il vient $\text{Mat}_{\mathcal{B}}(u) = 2\text{Mat}_{\mathcal{B}}(p) - I_{n+1}$, qui est diagonale.

Et donc u est diagonalisable.

SOLUTION DE L'EXERCICE 10.9

1. Nous savons que la trace est une forme linéaire sur $\mathcal{M}_n(\mathbf{R})$, non nulle car $\text{tr}(I_n) = n \neq 0$. Donc son noyau est un hyperplan de $\mathcal{M}_n(\mathbf{R})$, de dimension $n^2 - 1$.

2.a. Soient $M, N \in \mathcal{M}_n(\mathbf{R})$ et $\lambda \in \mathbf{R}$. Alors

$$f(\lambda M + N) = \lambda M + N + \text{tr}(\lambda M + N)I_n = \lambda(M + \text{tr}(M)I_n) + (N + \text{tr}(N)I_n) = \lambda f(M) + f(N).$$

Donc f est linéaire : c'est un endomorphisme de $\mathcal{M}_n(\mathbf{R})$.

2.b. Si $M \in \text{Ker}(\text{tr})$, alors $f(M) = M$.

Donc 1 est valeur propre de f et $\text{Ker}(\text{tr}) \subset E_1(f)$, de sorte que $\dim E_1(f) \geq n^2 - 1$.

De plus, $f(I_n) = I_n + nI_n = (n + 1)I_n$.

Donc $n + 1$ est valeur propre de f , et ainsi, $\dim E_{n+1}(f) \geq 1$.

Mais alors $\dim E_1(f) + \dim E_{n+1}(f) \geq n^2 - 1 + 1 = n^2 = \dim \mathcal{M}_n(\mathbf{R})$.

Puisque l'inégalité inverse est toujours vérifiée, alors $\dim E_1(f) + \dim E_{n+1}(f) = n^2$, et donc f est diagonalisable, et ses valeurs propres sont 1 et $n + 1$.

Une autre solution : prouvons qu'il existe une base de $\mathcal{M}_n(\mathbf{R})$ formée de vecteurs propres de f .

Soit donc (e_1, \dots, e_{n^2-1}) une base de Ker tr . Alors pour tout $i \in \llbracket 1, n^2 - 1 \rrbracket$,

$$f(e_i) = e_i + \underbrace{\text{tr}(e_i)}_{=0} I_n = e_i.$$

Autrement dit, e_1, \dots, e_{n^2-1} sont des vecteurs propres de f associés à la valeur propre 1.

D'autre part, I_n est une base de $\text{Vect}(I_n)$, et $f(I_n) = I_n + \text{tr}(I_n)I_n = (n + 1)I_n$.

Donc I_n est un vecteur propre de f associé à la valeur propre $n + 1$.

On en déduit donc que la famille $(e_1, \dots, e_{n^2-1}, I_n)$, obtenue par concaténation de bases de $\text{Ker}(\text{tr})$ et de $\text{Vect}(I_n)$ est une base⁷ de $\mathcal{M}_n(\mathbf{R})$, formée de vecteurs propres de f .

⁶ Sa matrice dans toute base est I_{n+1} , qui est diagonale.

Danger !

Il n'est pas vrai en général pour deux endomorphismes que

$$(f + g)^2 = f^2 + 2f \circ g + g^2.$$

Ce n'est vrai que s'ils commutent (c'est la formule du binôme !).

C'est bien le cas ici de u et id , puisque l'identité commute à tout endomorphisme.

Remarque

Le même raisonnement prouverait que tout endomorphisme f vérifiant $f^2 = \text{id}$ (un tel endomorphisme s'appelle une symétrie) est diagonalisable, car $\frac{\text{id}+f}{2}$ est un projecteur donc diagonalisable.

Remarque

A priori, f pourrait avoir d'autres valeurs propres. Mais puisque la somme des dimensions des deux sous-espaces propres est déjà égale à la dimension de $\mathcal{M}_n(\mathbf{R})$, il n'y a « pas de place » pour d'autres valeurs propres.

Détails

Les e_i sont non nuls, car ils forment une famille libre.

⁷ C'est ici qu'on utilise que $\text{Ker}(\text{tr})$ et $\text{Vect}(I_n)$ sont supplémentaires dans $\mathcal{M}_n(\mathbf{R})$.

Et donc f est diagonalisable.

La matrice de f dans cette base sera alors

$$\begin{pmatrix} f(e_1) & \dots & f(e_{n^2-1}) & f(I_n) \\ 1 & & 0 & 0 \\ 0 & \ddots & \vdots & \vdots \\ \vdots & \ddots & 1 & 0 \\ 0 & \dots & 0 & n+1 \end{pmatrix} \begin{matrix} e_1 \\ \vdots \\ e_{n^2-1} \\ I_n \end{matrix}.$$

Elle est diagonale, donc ses valeurs propres, qui sont celles de f , sont 1 et $n+1$.

SOLUTION DE L'EXERCICE 10.10

1. On a

$$X_{n+1} = \begin{pmatrix} u_{n+1} \\ u_{n+2} \\ u_{n+3} \end{pmatrix} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -2 & 1 & 2 \end{pmatrix} \begin{pmatrix} u_n \\ u_{n+1} \\ u_{n+2} \end{pmatrix}.$$

Ainsi, si on pose $A = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -2 & 1 & 2 \end{pmatrix}$, alors $\forall n \in \mathbf{N}, X_{n+1} = AX_n$.

2. Prouvons le résultat par récurrence sur n .

Si $n = 0$, alors $A^0 = I_n$, et donc le résultat est évident.

Supposons que $X_n = A^n X_0$. Alors

$$X_{n+1} = AX_n = AA^n X_0 = A^{n+1} X_0.$$

Ainsi, la propriété est vraie au rang $n+1$, et par le principe de récurrence, $\forall n \in \mathbf{N}, X_n = A^n X_0$.

3. Soit $\lambda \in \mathbf{R}$. Alors λ est valeur propre de A si et seulement si $\text{rg}(A - \lambda I_3) < 3$. Mais on a

$$\begin{aligned} A - \lambda I_3 &= \begin{pmatrix} -\lambda & 1 & 0 \\ 0 & -\lambda & 1 \\ -2 & 1 & 2-\lambda \end{pmatrix} \xrightarrow{L_1 \leftrightarrow L_3} \begin{pmatrix} -2 & 1 & 2-\lambda \\ 0 & -\lambda & 1 \\ -\lambda & 1 & 0 \end{pmatrix} \xrightarrow{L_3 \leftarrow 2L_3 - \lambda L_1} \begin{pmatrix} -2 & 1 & 2-\lambda \\ 0 & -\lambda & 1 \\ 0 & 2-\lambda & \lambda^2 - 2\lambda \end{pmatrix} \\ &\xrightarrow{L_2 \leftarrow L_2 - L_3} \begin{pmatrix} -2 & 1 & 2-\lambda \\ 0 & -2 & 2\lambda - \lambda^2 + 1 \\ 0 & 2-\lambda & \lambda^2 - 2\lambda \end{pmatrix} \xrightarrow{L_3 \leftarrow 2L_3 + (2-\lambda)L_2} \begin{pmatrix} -2 & 1 & 2-\lambda \\ 0 & -2 & 2\lambda - \lambda^2 + 1 \\ 0 & 0 & \lambda^3 - 2\lambda^2 - \lambda + 2 \end{pmatrix} \end{aligned}$$

Ainsi, $\text{rg}(A - \lambda I_3) < 3 \Leftrightarrow \lambda^3 - 2\lambda^2 - \lambda + 2 = 0$.

Une racine évidente de ce polynôme est 1, de sorte qu'on peut le factoriser par $(\lambda - 1)$:

$$\lambda^3 - 2\lambda^2 - \lambda + 2 = (\lambda - 1)(\lambda^2 - \lambda - 2) = (\lambda - 1)(\lambda + 1)(\lambda - 2).$$

Ainsi, $\text{Spec}(A) = \{-1, 1, 2\}$.

Puisqu'on a trois valeurs propres distinctes, A est automatiquement diagonalisable.

4. Il existe une matrice inversible P telle que

$$A = P^{-1} \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 2 \end{pmatrix} P^{-1}$$

et donc

$$\forall n \in \mathbf{N}, X_n = P^{-1} \begin{pmatrix} 1 & 0 & 0 \\ 0 & (-1)^n & 0 \\ 0 & 0 & 2^n \end{pmatrix} P X_0.$$

Le calcul montre que u_n est alors une combinaison linéaire⁸ de 1^n , $(-1)^n$ et 2^n .

Il existe donc trois réels λ, μ, ν tels que

$$\forall n \in \mathbf{N}, u_n = \lambda + \mu(-1)^n + \nu 2^n.$$

5. Il reste à déterminer ces trois constantes, et pour cela, nous allons utiliser les termes initiaux :

$$\begin{cases} 1 = \lambda + \mu + \nu \\ 0 = \lambda - \mu + 2\nu \\ 2 = \lambda + \mu + 4\nu \end{cases} \Leftrightarrow \begin{cases} \lambda = 0 \\ \mu = \frac{2}{3} \\ \nu = \frac{1}{3} \end{cases}$$

Méthode

Pour trouver les racines d'un polynôme de degré 3, si l'énoncé n'indique rien, il faudra chercher une racine «évidente», souvent 0, 1, -1, 2 ou -2.

⁸ Dont les coefficients ne dépendent pas de n , mais seulement des coefficients de P^{-1} et de PX_0 .

Et donc

$$\forall n \in \mathbf{N}, u_n = \frac{2}{3}(-1)^n + \frac{2^n}{3}.$$

SOLUTION DE L'EXERCICE 10.11

1. On a $A^2 = \begin{pmatrix} 7 & -6 & -6 \\ 6 & -5 & -6 \\ -3 & 3 & 4 \end{pmatrix} = 2I_3 - A$.

On a donc $A^2 + A - 2I_3 = 0$: $X^2 + X - 2$ est un polynôme annulateur de A .

2. Notons $R_n = a_n + b_n X \in \mathbf{R}_1[X]$ le reste de la division euclidienne de X^n par $X^2 + X - 2$, et notons $Q_n \in \mathbf{R}[X]$ le quotient, de sorte que

$$X^n = (X^2 + X - 2)Q_n + R_n.$$

Les racines de $X^2 + X - 2$ sont 1 et -2 . Ainsi, on a $1^n = Q_n(1) \times 0 + a_n + b_n$ et de même $(-2)^n = R_n(-2) = a_n - 2b_n$.

Après résolution du système $\begin{cases} 1 = a_n + b_n \\ (-2)^n = a_n - 2b_n \end{cases}$ on obtient $b_n = \frac{1 - (-2)^n}{3}$ et $a_n = \frac{2 + (-2)^n}{3}$.

On en déduit que

$$\begin{aligned} A^n &= Q_n(A) \times \underbrace{(A^2 + A - 2I_n)}_{=0} + a_n I_n + b_n A \\ &= a_n I_n + b_n A \\ &= \begin{pmatrix} 2(-2)^n - 1 & 2(1 - (-2)^n) & 2(1 - (-2)^n) \\ 2((-2)^n - 1) & 3 + (-2)^{n+1} & 2(1 - (-2)^n) \\ 1 - (-2)^n & (-2)^n - 1 & (-2)^n \end{pmatrix}. \end{aligned}$$

3. Les valeurs propres de A sont parmi les racines de $X^2 + X - 2$, donc ne peuvent être que -2 et 1 . Or, on a

$$A \begin{pmatrix} x \\ y \\ z \end{pmatrix} = -2 \begin{pmatrix} x \\ y \\ z \end{pmatrix} \Leftrightarrow \begin{cases} -5x + 6y + 6z = -2x \\ -6x + 7y + 6z = -2y \\ 3x - 3y - 2z = -2z \end{cases} \Leftrightarrow \begin{cases} -x + 2y + 2z = 0 \\ -2x + 3y + 2z = 0 \\ x = y \end{cases} \Leftrightarrow x = -2z, y = z$$

Et donc une base de $E_{-2}(A)$ est $\begin{pmatrix} -2 \\ -2 \\ 1 \end{pmatrix}$.

De même, on a

$$A \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x \\ y \\ z \end{pmatrix} \Leftrightarrow \begin{cases} -5x + 6y + 6z = x \\ -6x + 7y + 6z = y \\ 3x - 3y - 2z = z \end{cases} \Leftrightarrow -x + y + z = 0.$$

Et donc une base de $E_1(A)$ est $\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$.

En particulier, on a $\dim E_1(A) + \dim E_{-2}(A) = 2 + 1 = 3$, donc A est diagonalisable. On a alors $A = P^{-1}DP$, avec

$$D = \begin{pmatrix} -2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, P^{-1} = \begin{pmatrix} -2 & 1 & 1 \\ -2 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix}, P = \begin{pmatrix} -1 & 1 & 1 \\ -2 & 3 & 2 \\ 1 & -1 & 0 \end{pmatrix}.$$

Pour tout entier n , on a alors

$$A^n = P^{-1}D^n P = P^{-1} \begin{pmatrix} (-2)^n & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} P.$$

Après calculs, on retrouve bien le résultat obtenu à la question précédente.

⚠ Attention !

Comme d'habitude, rien ne nous permet à ce stade d'affirmer que -2 et 1 sont des valeurs propres de A , il faut le vérifier, soit en cherchant une base de vecteurs propres (si on trouve 0 comme unique solution à $AX = \lambda X$, c'est que λ n'est pas valeur propre), soit en calculant le rang de $A - \lambda I_3$, pour $\lambda = -2$ et $\lambda = 1$.

SOLUTION DE L'EXERCICE 10.12

- Par le théorème du rang, si f est de rang égal à 1, alors $\dim \text{Ker } f = n - 1$. Soit donc (e_1, \dots, e_{n-1}) une base de $\text{Ker } f$. Par le théorème de la base incomplète, il est possible de compléter cette base en une base (e_1, \dots, e_n) de E . Et alors nécessairement, les $n - 1$ premières colonnes de la matrice de f dans cette base seront nulles.
- M est triangulaire supérieure, donc ses valeurs propres sont ses coefficients diagonaux : 0 et α_n .
Si $\alpha_n \neq 0$, alors M possède deux valeurs propres distinctes, et $\dim E_{\alpha_n}(f) = 1$.
On a alors $\dim E_0(f) + \dim E_{\alpha_n}(f) = \dim \text{Ker } f + 1 = n$. Et donc f est diagonalisable.
Si $\alpha_n = 0$, alors f possède 0 comme seule valeur propre, avec

$$\dim E_n(f) = \dim \text{Ker } f = n - 1 \neq n.$$

Donc f n'est pas diagonalisable.

- Puisque $\text{tr}(M) = \alpha_n$, nous venons de prouver que f est diagonalisable si et seulement si $\text{tr}(M) \neq 0$.
Or A et M représentent toutes les deux l'endomorphisme f dans deux bases différentes de E , donc sont semblables. Et deux matrices semblables ayant même trace, f est donc diagonalisable si et seulement si $\text{tr}(A) \neq 0$.

SOLUTION DE L'EXERCICE 10.13

- Soit $A = \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix}$ une matrice commutant avec C . Alors

$$CA = \begin{pmatrix} 9 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix} = \begin{pmatrix} 9a & 9b & 9c \\ 4d & 4e & 4f \\ g & h & i \end{pmatrix} \text{ et } AC = \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix} \begin{pmatrix} 9 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 9a & 4b & c \\ 9d & 4e & f \\ 9g & 4h & i \end{pmatrix}.$$

On a donc $AC = CA$ si et seulement si $b = c = d = f = g = h = 0$, c'est-à-dire si et seulement si A est diagonale.

Si $D^2 = C$, alors D et C commutent car C est alors un polynôme en D . Et par ce qui précède, D est alors diagonale.

- Soit $D = \begin{pmatrix} a & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & c \end{pmatrix}$ une matrice diagonale. Alors $D^2 = \begin{pmatrix} a^2 & 0 & 0 \\ 0 & b^2 & 0 \\ 0 & 0 & c^2 \end{pmatrix}$.

$$\text{Et donc } D^2 = C \text{ si et seulement si } \begin{cases} a^2 = 9 \\ b^2 = 4 \\ c^2 = 1 \end{cases}$$

Il y a donc $2^3 = 8$ matrices telles que $D^2 = C$, ce sont celles avec $a = \pm 3, b = \pm 4$ et $c = \pm 1$.

- A est triangulaire, et possède trois valeurs propres distinctes : elle est diagonalisable, semblable à C . Soit $P \in GL_3(\mathbf{R})$ telle que $A = P^{-1}CP$.
Alors si B vérifie $B^2 = A$, on a

$$(PBP^{-1})^2 = PB^2P^{-1} = PAP^{-1} = C.$$

Donc PBP^{-1} est une des 8 matrices obtenues à la question précédente, et donc $B = P^{-1}D$, où $D^2 = C$.

Inversement, si $B = P^{-1}DP$, avec $D^2 = C$, alors $B^2 = P^{-1}CP = A$.

Enfin, notons que les 8 matrices ainsi obtenues sont bien distinctes, puisqu'elles possèdent des valeurs propres⁹ distinctes.

SOLUTION DE L'EXERCICE 10.14

- On a $JV = \begin{pmatrix} n \\ \vdots \\ n \end{pmatrix} = nV$ et donc¹⁰, n est une valeur propre de J .

De plus, J est de rang 1, donc 0 est valeur propre de J et $\dim E_0(J) = n - \text{rg}(J) = n - 1$.

On a alors $1 \leq \dim E_n(J) \leq n - \dim E_0(J) = 1$, et donc $\dim E_0(J) = 1$.

Ainsi, $\dim E_0(J) + \dim E_n(J) = n$, et donc J est diagonalisable.

Il existe donc P inversible et $\Delta = \text{Diag}(n, 0, \dots, 0)$ telles que $J = P^{-1}\Delta P$.

Méthode

Une telle base, si elle existe, possède ses $n - 1$ premiers vecteurs dans le noyau de f . Il s'agit donc de montrer qu'une telle base existe.

Matrice triangulaires

Rappelons que pour une matrice triangulaire, la dimension d'un sous-espace propre est inférieure ou égale au nombre de fois où le coefficient apparaît sur la diagonale. En particulier, un coefficient qui n'apparaît qu'une seule fois correspond nécessairement à un sous-espace propre de dimension 1.

Généralisation

Le même raisonnement tiendrait pour toute matrice diagonale donc les coefficients diagonaux sont deux à deux distincts.

⁹ Qui sont celles des matrices D telles que $D^2 = C$.

¹⁰ Car $V \neq 0$.

2. On a $M = bJ + (a - b)I_n$. Et donc

$$M = P^{-1}b\Delta P + P^{-1}(a - b)I_n P = P^{-1}(b\Delta + (a - b)I_n)P = P^{-1}\text{Diag}(bn + a - b, a - b, \dots, a - b)P.$$

Donc M est semblable à une matrice diagonale, et par conséquent diagonalisable.

On en déduit même que $\text{Spec}(M) = \{a - b, a + (n - 1)b\}$, avec $\dim E_{a-b}(J) = n - 1$ et $\dim E_{a+(n-1)b}(J) = 1$.

SOLUTION DE L'EXERCICE 10.15

1. Soit $\lambda \in \mathbf{R}$. Alors λ est valeur propre de A si et seulement si $\text{rg}(A - \lambda I_3) < 3$. Or,

$$\begin{aligned} A - \lambda I_3 &= \begin{pmatrix} 4 - \lambda & 4 & -4 \\ 2 & 2 - \lambda & 4 \\ 4 & 4 & 2 - \lambda \end{pmatrix} \xleftrightarrow{L_1 \leftrightarrow L_3} \begin{pmatrix} 4 & 4 & 2 - \lambda \\ 2 & 2 - \lambda & 4 \\ 4 - \lambda & 4 & -4 \end{pmatrix} \xleftrightarrow{\substack{L_2 \leftarrow 2L_2 - L_1 \\ L_3 \leftarrow 4L_3 - (4 - \lambda)L_1}} \begin{pmatrix} 4 & 4 & 2 - \lambda \\ 0 & -2\lambda & 6 + \lambda \\ 0 & 4\lambda & -\lambda^2 + 6\lambda - 24 \end{pmatrix} \\ &\xleftrightarrow{L_3 \leftarrow L_3 + 2L_2} \begin{pmatrix} 4 & 4 & 2 - \lambda \\ 0 & -2\lambda & 6 + \lambda \\ 0 & 0 & -\lambda^2 + 8\lambda - 12 \end{pmatrix}. \end{aligned}$$

Et donc $\text{rg}(A - \lambda I_3) < 3$ si et seulement si $\lambda = 0$ ou $\lambda^2 - 8\lambda + 12 = 0 \Leftrightarrow (\lambda - 2)(\lambda - 6) = 0$. Ainsi, $\text{Spec}(A) = \{0, 2, 6\}$.

Notons dès à présent que puisque nous avons trois valeurs propres, les sous-espaces propres seront tous de dimension 1.

Puisque les colonnes de A vérifient $C_1 - C_2 + 0 \cdot C_3 = 0$, un vecteur propre de A pour la

valeur propre 0 est $\begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}$. Et donc¹¹ $E_0(A) = \text{Vect}\left(\begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}\right)$.

De même, $A - 2I_3 = \begin{pmatrix} 2 & 4 & -4 \\ 2 & 0 & 4 \\ 4 & 4 & 0 \end{pmatrix}$ a ses colonnes liées par la relation $2C_1 - 2C_2 - C_3 = 0$,

donc $E_2(A) = \text{Vect}\left(\begin{pmatrix} 2 \\ -2 \\ -1 \end{pmatrix}\right)$.

Enfin, $A - 6I_3 = \begin{pmatrix} -2 & 4 & -4 \\ 2 & -4 & 4 \\ 4 & 4 & -4 \end{pmatrix}$ a ses colonnes liées par la relation $0C_1 + C_2 + C_3 = 0$, et

donc $E_6(A) = \text{Vect}\left(\begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}\right)$.

2.a. Puisque A est un polynôme¹² en M , A et M commutent.

2.b. On a $AMX = MAX = M\lambda X = \lambda MX$.

Donc $MX \in E_\lambda(A)$. Mais nous avons vu précédemment que les $E_\lambda(A)$ sont tous de dimension 1, et donc X , qui est un vecteur non nul de $E_\lambda(A)$ est une base de $E_\lambda(A)$.

On en déduit qu'il existe un réel μ tel que $MX = \mu X$.

Autrement dit, que X est un vecteur propre de M associé à une valeur propre μ .

Mais nous savons alors que $\lambda X = AX = (M^2 + M)X = (\mu^2 + \mu)X$, et donc¹³, $\lambda = \mu + \mu^2$.

3. Nous venons de prouver que les vecteurs propres de A sont des vecteurs propres de M .

Et donc en particulier, si \mathcal{B} est une base de $\mathcal{M}_{3,1}(\mathbf{R})$ formée de vecteurs propres de A , c'est aussi une base formée de vecteurs propres de M .

Autrement dit, la matrice de f_M dans la base \mathcal{B} doit être diagonale : $P_{\mathcal{B}, \mathcal{B}_{can}} M P_{\mathcal{B}, \mathcal{B}_{can}}^{-1}$ est diagonale.

Mais nous connaissons déjà une base de vecteurs propres de A : il suffit de concaténer les

bases des sous-espaces propres de A : on peut par exemple prendre $\mathcal{B} = \left(\begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 \\ -2 \\ -1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} \right)$.

Dans ce cas, on a donc $P_{\mathcal{B}, \mathcal{B}_{can}} = \begin{pmatrix} 1 & 2 & 0 \\ -1 & -2 & 1 \\ 0 & -1 & 1 \end{pmatrix}$.

Si on note P cette matrice, on a donc PMP^{-1} qui est une matrice diagonale $D = \begin{pmatrix} \mu_1 & 0 & 0 \\ 0 & \mu_2 & 0 \\ 0 & 0 & \mu_3 \end{pmatrix}$,

Astuce

Pour toute matrice inversible P , $P^{-1}I_n P = I_n$.
Et donc si on ajoute un multiple de I_n à une matrice diagonalisable, alors on obtient encore une matrice diagonalisable.

¹¹ $E_0(A)$ est de dimension 1, donc il suffit d'en trouver un vecteur non nul pour en avoir une base.

Relations

Les relations entre les colonnes ne sont pas toujours faciles à trouver : si on ne les voit pas, mieux vaut résoudre un système pour trouver des vecteurs propres.

¹² $A = M + M^2$, soit encore $A = P(M)$, avec $P(X) = X + X^2$.

¹³ X étant non nul.

avec $\mu_1^2 + \mu_1 = 0$, $\mu_2^2 + \mu_2 = 2$, $\mu_3^2 + \mu_3 = 6$.

Les solutions de ces équations sont $\mu_1 = 0$ et $\mu_1 = -1$, $\mu_2 = 1$ et $\mu_2 = -1$, $\mu_3 = 2$ et $\mu_3 = -3$.
Autrement dit, toutes les solutions de l'équation sont les

$$P \begin{pmatrix} \mu_1 & 0 & 0 \\ 0 & \mu_2 & 0 \\ 0 & 0 & \mu_3 \end{pmatrix} P^{-1} \text{ avec } (\mu_1, \mu_2, \mu_3) \in \{0, 1\} \times \{1, -1\} \times \{2, -3\}.$$

Notons que ces solutions sont bien distinctes, puisqu'elles ont à chaque fois des valeurs propres différentes : il y a donc 8 solutions à l'équation.

En particulier, l'unique solution dont toutes les valeurs propres sont positives, est celle avec $\mu_1, \mu_2, \mu_3 \geq 0$, c'est :

$$P \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix} P^{-1} = \begin{pmatrix} 2 & 2 & -2 \\ 0 & 0 & 2 \\ 1 & 1 & 1 \end{pmatrix}.$$

SOLUTION DE L'EXERCICE 10.16

Notons M la matrice $\begin{pmatrix} X & 0 \\ X + Y & Y \end{pmatrix}$.

Alors M est triangulaire, et donc ses valeurs propres sont ses coefficients diagonaux.

Si ils sont différents, alors M est diagonalisable car elle possède alors deux valeurs propres.

S'ils sont égaux, alors M possède une seule valeur propre λ .

Donc elle est diagonalisable si et seulement si¹⁴ elle est égale à λI_2 , ce qui n'est pas le cas ici car $X + Y$ ne s'annule jamais¹⁵.

Ainsi, M est diagonalisable si et seulement si $X \neq Y$.

Or, on a

$$P(X \neq Y) = 1 - P(X = Y) = 1 - \sum_{k=1}^{+\infty} P([X = k] \cap [Y = k]).$$

Par indépendance de X et Y , on a

$$P([X = k] \cap [Y = k]) = P(X = k)P(Y = k) = P(X = k)^2 = p^2(1 - p)^{2(k-1)}.$$

Et donc

$$\begin{aligned} \sum_{k=1}^{+\infty} P([X = k] \cap [Y = k]) &= \sum_{k=1}^{+\infty} p^2(1 - p)^{2(k-1)} \\ &= p^2 \sum_{i=0}^{+\infty} (1 - p)^{2i} \\ &= p^2 \frac{1}{1 - (1 - p)^2}. \end{aligned}$$

On en déduit que la probabilité que M soit diagonalisable est

$$1 - p^2 \frac{1}{1 - (1 - p)^2} = \frac{2p - p^2 - p^2}{2p - p^2} = 2 \frac{p(1 - p)}{p(2 - p)} = 2 \frac{1 - p}{2 - p}.$$

SOLUTION DE L'EXERCICE 10.17

Les valeurs propres de f sont parmi les racines du polynôme annulateur $(X - a)^3(X - b)$, et donc $\text{Spec}(f) \subset \{a, b\}$.

Si f est diagonalisable, alors un polynôme P dont toutes les valeurs propres de f sont racines doit être annulateur de f .

En particulier, $(X - a)^2(X - b)$, qui possède a et b pour racines devrait être annulateur de f . Or, ce n'est pas le cas, donc f n'est pas diagonalisable.

SOLUTION DE L'EXERCICE 10.18

Commençons par remarquer qu'on a $f(I_n) = I_n + 2I_n + 3nI_n = 3(n + 1)I_n$.

Et donc $3(n + 1)$ est une valeur propre de f .

Si M est antisymétrique, alors ses coefficients diagonaux sont nuls, et donc $\text{tr}(M) = 0$, de sorte que $f(M) = M - 2M = -M$.

Et donc -1 est valeur propre de f , et $E_{-1}(f)$ contient $\mathcal{A}_n(\mathbf{R})$, l'ensemble des matrices antisymétriques.

Autrement dit

Ce que nous venons de prouver est que puisque A et M possèdent une base commune de vecteurs propres alors il existe une matrice P telle que PAP^{-1} et PMP^{-1} soient simultanément diagonales. Ou encore, qu'elles sont diagonalisables, avec la même matrice de passage. En règle générale, si A et B sont deux matrices diagonalisables, alors il existe une matrice P_1 telle que $P_1AP_1^{-1}$ soit diagonale, il existe une matrice P_2 telle que $P_2BP_2^{-1}$ soit diagonale, mais il n'y a a priori aucun lien entre P_1 et P_2 .

¹⁴ Voir l'exercice 3.

¹⁵ Ou plutôt, ne s'annule presque jamais : $P(X + Y = 0) = 0$.

Probas totales

Cette formule est intuitive, mais si l'on veut être rigoureux, c'est la formule des probabilités totales appliquée au s.c.e

$$\{[X = k], k \in \mathbf{N}^*\}.$$

Explication

En effet, si D est la matrice de f dans une base de vecteurs propres, alors $P(D) = 0$ si et seulement si P possède toutes les valeurs propres de f comme racines.

En particulier, si f est diagonalisable, alors

$$\prod_{\lambda \in \text{Spec}(f)} (X - \lambda)$$

est annulateur de f .

Enfin, si M est symétrique et de trace nulle, alors $f(M) = M + 2M = 3M$.
 Puisqu'il existe bien des matrices non nulles qui sont symétriques et de trace nulle¹⁶, c'est donc que 3 est valeur propre de f .
 Notons F l'ensemble des matrices symétriques de trace nulle, de sorte que $F \subset E_3(f)$.
 Alors F est le noyau de la forme linéaire $\varphi : \mathcal{S}_n(\mathbf{R}) \rightarrow \mathbf{R}$ définie par $\varphi(M) = \text{tr}(M)$.
 Cette forme linéaire étant non nulle, on en déduit que F est un hyperplan de $\mathcal{S}_n(\mathbf{R})$, et donc de dimension $\dim \mathcal{S}_n(\mathbf{R}) - 1$.

Première méthode : si l'on sait¹⁷ que $\mathcal{M}_n(\mathbf{R}) = \mathcal{S}_n(\mathbf{R}) \oplus \mathcal{A}_n(\mathbf{R})$.
 Alors $\dim \mathcal{S}_n(\mathbf{R}) + \dim \mathcal{A}_n(\mathbf{R}) = \dim \mathcal{M}_n(\mathbf{R})$.
 Par ce qui a été dit précédemment, on a

$$\dim E_{3(n+1)}(f) \geq 1, \dim E_{-1}(f) \geq \dim \mathcal{A}_n(\mathbf{R}) \text{ et } \dim E_3(f) \geq \dim \mathcal{S}_n(\mathbf{R}) - 1.$$

Et donc

$$\dim E_{-1}(f) + \dim E_3(f) + \dim E_{3(n+1)}(f) \geq \dim \mathcal{A}_n(\mathbf{R}) + \dim \mathcal{S}_n(\mathbf{R}) = \dim \mathcal{M}_n(\mathbf{R}).$$

Puisque l'inégalité réciproque est toujours vraie, on a donc

$$\dim E_{-1}(f) + \dim E_3(f) + \dim E_{3(n+1)}(f) = \dim \mathcal{M}_n(\mathbf{R})$$

et par conséquent f est diagonalisable.

Seconde méthode : si on ne sait pas que $\mathcal{A}_n(\mathbf{R})$ et $\mathcal{S}_n(\mathbf{R})$ sont supplémentaires dans $\mathcal{M}_n(\mathbf{R})$, on peut montrer «à la main» que $\mathcal{M}_n(\mathbf{R})$ est la somme directe de ses sous-espaces propres.

Notons qu'à ce stade, nous ne sommes pas sûrs d'avoir toutes les valeurs propres de f , mais comme nous n'en trouvons pas d'autres que les trois trouvées précédemment, nous allons essayer de prouver que $\mathcal{M}_n(\mathbf{R}) = \mathcal{A}_n(\mathbf{R}) \oplus F \oplus \text{Vect}(I_n)$.

Raisonnons par analyse synthèse : soit $M \in \mathcal{M}_n(\mathbf{R})$, et supposons que $M = A + B + C$, avec $A \in \mathcal{A}_n(\mathbf{R})$, $B \in F$ et $C \in \text{Vect}(I_n)$.

Il existe alors $\lambda \in \mathbf{R}$ tel que $C = \lambda I_n$.

Et alors $\text{tr}(M) = \text{tr}(A) + \text{tr}(B) + \lambda \text{tr}(I_n) = 0 + 0 + \lambda n$.

Et donc $\lambda = \frac{\text{tr}(M)}{n}$.

De plus, on a ${}^t M = {}^t A + {}^t B + \lambda I_n = -A + B + \lambda I_n$.

Et donc $A = \frac{M - {}^t M}{2}$ et $B = \frac{M + {}^t M}{2} - \lambda I_n = \frac{M + {}^t M}{2} - \frac{\text{tr}(M)}{n} I_n$.

Ainsi, si une telle décomposition $M = A + B + C$ existe, elle est unique.

Inversement, pour $M \in \mathcal{M}_n(\mathbf{R})$, posons

$$A = \frac{M - {}^t M}{2}, B = \frac{M + {}^t M}{2} - \frac{\text{tr}(M)}{n} I_n \text{ et } C = \frac{\text{tr}(M)}{n} I_n.$$

Alors il est clair que $C \in \text{Vect}(I_n)$, A est antisymétrique car ${}^t A = \frac{{}^t M - M}{2} = -A$.

De plus, B est symétrique car ${}^t B = \frac{{}^t M + M}{2} - \frac{\text{tr}(M)}{n} I_n = B$, et elle est de trace nulle car

$$\text{tr}(B) = \frac{1}{2} \text{tr}({}^t M) + \frac{1}{2} \text{tr}(M) - \frac{\text{tr}(M)}{n} \text{tr}(I_n) = \frac{1}{2} \text{tr}(M) + \frac{1}{2} \text{tr}(M) - \text{tr}(M) = 0.$$

Donc $B \in F$.

Ainsi, tout élément de $\mathcal{M}_n(\mathbf{R})$ s'écrit de manière unique comme un élément de $\mathcal{A}_n(\mathbf{R})$ plus un élément de F , plus un élément de $\text{Vect}(I_n)$.

En particulier, toute matrice de $\mathcal{M}_n(\mathbf{R})$ s'écrit comme un élément de $E_{-1}(f)$ plus un élément de $E_{-3}(f)$ plus un élément de $E_{3(n+1)}(f)$.

Et donc $\mathcal{M}_n(\mathbf{R}) = E_{-1}(f) + E_{-3}(f) + E_{3(n+1)}(f)$.

Mais comme nous savons déjà que les sous-espaces propres sont en somme directe, on a donc $\mathcal{M}_n(\mathbf{R}) = E_{-1}(f) \oplus E_{-3}(f) \oplus E_{3(n+1)}(f)$.

Et donc $\mathcal{M}_n(\mathbf{R})$ est somme directe de sous-espaces propres de f : f est diagonalisable.

Notons alors au passage que $E_{-1}(f) = \mathcal{A}_n(\mathbf{R})$, $E_{-3}(f) = F$ et $E_{3(n+1)}(f) = \text{Vect}(I_n)$, alors que nous n'avions jusqu'à présent que des inclusions.

¹⁶ Prendre par exemple $D = \text{Diag}(1, \dots, 1, 1 - n)$.

¹⁷ Ce point n'est pas explicitement au programme, mais est suffisamment classique pour qu'on puisse l'utiliser à l'oral sans le redémontrer.

Remarque

Il n'y a pas besoin de les connaître ici, mais les dimensions de $\mathcal{A}_n(\mathbf{R})$ et $\mathcal{S}_n(\mathbf{R})$ sont connues :

$$\dim \mathcal{A}_n(\mathbf{R}) = \frac{n(n-1)}{2}$$

$$\dim \mathcal{S}_n(\mathbf{R}) = \frac{n(n+1)}{2}.$$

SOLUTION DE L'EXERCICE 10.19

1. Soit λ une valeur propre de A , et soit $X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathcal{M}_{n,1}(\mathbf{R})$ un vecteur propre associé.

$$\text{Soit alors } M = \begin{pmatrix} x_1 & 0 & \dots & 0 \\ x_2 & 0 & \dots & 0 \\ \vdots & \vdots & & \vdots \\ x_n & 0 & \dots & 0 \end{pmatrix} = \left(\begin{array}{c|c|c|c} x_1 & 0 & & 0 \\ \vdots & \vdots & & \vdots \\ x_n & 0 & & 0 \end{array} \right) \in \mathcal{M}_n(\mathbf{R}).$$

On a alors

$$\varphi_A(M) = AM = \left(\begin{array}{c|c|c|c} AX & 0 & \dots & 0 \end{array} \right) = \left(\begin{array}{c|c|c|c} \lambda X & 0 & \dots & 0 \end{array} \right) = \lambda M.$$

Et puisque $M \neq 0$ car $X \neq 0$, on en déduit que λ est valeur propre de φ_A .
Donc déjà $\text{Spec}(A) \subset \text{Spec}(\varphi_A)$.

Inversement, soit $M = \left(\begin{array}{c|c|c|c} M_1 & M_2 & \dots & M_n \end{array} \right)$ un vecteur propre de φ_A associé à la valeur

propre λ .

On a alors

$$\lambda M = AM = \left(\begin{array}{c|c|c|c} AM_1 & AM_2 & \dots & AM_n \end{array} \right).$$

Mais puisque $M \neq 0$, l'un au moins des M_i est non nul, et alors $AM_i = \lambda M_i$, de sorte que M_i est un vecteur propre de A associé à la valeur propre λ .

Et donc $\text{Spec}(\varphi_A) \subset \text{Spec}(A)$.

On a donc prouvé que A et φ_A possèdent les mêmes valeurs propres.

2. Si A est diagonalisable, alors il existe une base (X_1, \dots, X_n) de $\mathcal{M}_{n,1}(\mathbf{R})$ formée de vecteurs propres.

Pour $(i, j) \in \llbracket 1, n \rrbracket^2$, notons alors $M_{i,j} = \left(\begin{array}{c|c|c|c} 0 & \dots & 0 & X_i \\ \vdots & & \vdots & \vdots \\ 0 & \dots & 0 & X_i \end{array} \right) \in \mathcal{M}_n(\mathbf{R})$ où X_i est

placé dans la $j^{\text{ème}}$ colonne de $M_{i,j}$.

Alors d'après ce qui a été dit à la première question, $M_{i,j}$ est un vecteur propre de φ_A , associé à la même valeur propre que X_i .

Prouvons que la famille $(M_{i,j})_{1 \leq i, j \leq n}$ est une famille libre de $\mathcal{M}_n(\mathbf{R})$.

Soient donc $(\lambda_{i,j})_{1 \leq i, j \leq n}$ des réels tels que $\sum_{i=1}^n \sum_{j=1}^n \lambda_{i,j} M_{i,j} = 0$. Alors

$$0 = \left(\begin{array}{c|c|c|c} \sum_{i=1}^n \lambda_{i,1} X_i & \sum_{i=1}^n \lambda_{i,2} X_i & \dots & \sum_{i=1}^n \lambda_{i,n} X_i \end{array} \right).$$

Par identification des colonnes¹⁸, on a donc

$$\forall j \in \llbracket 1, n \rrbracket \sum_{i=1}^n \lambda_{i,j} X_i = 0.$$

Mais la famille (X_1, \dots, X_n) étant libre, il vient donc $\forall (i, j) \in \llbracket 1, n \rrbracket^2, \lambda_{i,j} = 0$.

Ainsi, la famille $(M_{i,j})_{1 \leq i, j \leq n}$ est libre, et étant de cardinal $n^2 = \dim \mathcal{M}_n(\mathbf{R})$, c'est une base de $\mathcal{M}_n(\mathbf{R})$.

Nous venons donc de prouver qu'il existe une base de $\mathcal{M}_n(\mathbf{R})$ formée de vecteurs propres de φ_A : l'endomorphisme φ_A est diagonalisable.

SOLUTION DE L'EXERCICE 10.20

Produit matriciel

Lorsqu'on fait le produit AB de deux matrices (carrées ou non), la $j^{\text{ème}}$ colonne de AB ne dépend que de A et de la $j^{\text{ème}}$ colonne de B .

Et même, en y regardant de plus près, elle est égale au produit de A par la $j^{\text{ème}}$ colonne de B .

¹⁸ Deux matrices sont égales si et seulement si leurs vecteurs colonnes sont égaux.

Remarque

On a même la dimension des sous-espaces propres de φ_A : si $\lambda \in \text{Spec}(A)$, alors

$$\dim E_\lambda(\varphi_A) = n \times \dim E_\lambda(A).$$

1. Puisque f est diagonalisable, on a $E = \bigoplus_{i=1}^p E_i$ et donc tout vecteur de E s'écrit de manière unique comme somme de vecteurs des E_i . C'est en particulier le cas de x .
2. Puisque $x \neq 0$, les x_i ne sont pas tous nuls. Donc quitte à changer la numérotation des λ_i et des E_i , on peut supposer que $x_1, \dots, x_r \neq 0$ et $x_{r+1} = \dots = x_p = 0$.
- 3.a. Pour $i \in \llbracket 1, r \rrbracket$, $x_i \in E_i$ et $x_i \neq 0$.
Donc x_i est un vecteur propre de f , associé à la valeur propre λ_i .
En particulier, (x_1, \dots, x_r) est une famille formée de vecteurs propres associés à des valeurs propres distinctes : c'est donc une famille libre.
Par définition de V_x , cette famille est génératrice de V_x , et donc c'est une base de V_x .
- 3.b. Pour $j \geq 0$, on a

$$f^j(x) = f^j\left(\sum_{i=1}^r x_i\right) = \sum_{i=1}^r f^j(x_i) = \sum_{i=1}^r \lambda_i^j x_i \in V_x.$$

- 3.c. D'après le calcul qui a été réalisé à la question précédente, on a

$$\text{Mat}_{(x_1, \dots, x_r)}(x, f(x), \dots, f^{r-1}(x)) = \begin{pmatrix} x & f(x) & \dots & f^{r-1}(x) \\ 1 & \lambda_1 & \dots & \lambda_1^{r-1} \\ 1 & \lambda_2 & \dots & \lambda_2^{r-1} \\ \vdots & \vdots & \dots & \vdots \\ 1 & \lambda_r & \dots & \lambda_r^{r-1} \end{pmatrix} \begin{matrix} x_1 \\ x_2 \\ \vdots \\ x_r \end{matrix}.$$

Notons que $C_j = \begin{pmatrix} \lambda_1^{j-1} \\ \lambda_2^{j-1} \\ \vdots \\ \lambda_r^{j-1} \end{pmatrix}$.

Et donc

$$\sum_{j=1}^r \alpha_j C_j = 0 \iff \begin{pmatrix} \sum_{j=1}^r \alpha_j \lambda_1^{j-1} \\ \sum_{j=1}^r \alpha_j \lambda_2^{j-1} \\ \vdots \\ \sum_{j=1}^r \alpha_j \lambda_r^{j-1} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \iff \begin{pmatrix} P(\lambda_1) \\ P(\lambda_2) \\ \vdots \\ P(\lambda_r) \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$$

Et donc $P(\lambda_1) = \dots = P(\lambda_r) = 0$.

Or P est de degré au plus $r-1$ et possède r racines, c'est donc qu'il est nul.

On en déduit¹⁹ que $\alpha_1 = \dots = \alpha_r = 0$.

Et donc la famille (C_1, \dots, C_r) est libre. Elle est donc de rang r , de sorte que $\text{rg}(A) = r$.

Et donc A est une matrice inversible.

- 3.d. Puisque A est inversible, $(x, f(x), \dots, f^{r-1}(x))$ est une base de V_x .
En particulier, pour tout $i \in \llbracket 1, r \rrbracket$, x_i s'écrit comme combinaison linéaire de $x, f(x), \dots, f^{r-1}(x)$.
Mais puisque F est stable par f , alors $f(x) \in F$. Et donc $f(f(x)) = f^2(x) \in F$, etc, $f^{j-1}(x) \in F$.
Ainsi, x_i est combinaison linéaire d'éléments de F : $x_i \in F$.
Prouvons à présent que les $F \cap E_i$ sont en somme directe.
Supposons donc que $0_E = y_1 + \dots + y_p$, avec $y_i \in F \cap E_i$.
Alors en particulier, $y_i \in E_i$. Et puisque les E_i sont en somme directe, on en déduit que $y_1 = \dots = y_p = 0$.

De plus, nous venons de montrer que si $x \in F$, alors $x = x_1 + \dots + x_p$, avec $x_i \in F \cap E_i$.

Donc $F = \sum_{i=1}^p (F \cap E_i)$ et puisque la somme est directe, $F = \bigoplus_{i=1}^p (F \cap E_i)$.

4. Soient F_1, \dots, F_q les sous-espaces propres de g .
Puisque f et g commutent, les F_j sont stables par g .
Et donc pour tout $j \in \llbracket 1, q \rrbracket$, $F_j = \bigoplus_{i=1}^p (F_j \cap E_i)$.

Rappel

Le sous-espace propre est formé des vecteurs propres de f associés à la valeur propre λ_i et du vecteur nul.

¹⁹ Par identification des coefficients.

Rappel

Rappelons que par définition, le rang d'une matrice est le rang de la famille de ses vecteurs colonnes.

Remarque

Nous avons montré que x_1, \dots, x_r sont dans F . Pour x_{r+1}, \dots, x_p , c'est trivial car ils sont nuls et que le vecteur nul est dans F .

Et donc une base de F_j obtenue par concaténation de bases de $F_j \cap E_i$ sera formée de vecteurs qui sont à la fois vecteurs propres de f , car dans E_i , et de vecteurs propres de g , car dans F_j .

Et puisque $E = \bigoplus_{j=1}^q F_j$, la concaténation des bases des F_j précédemment obtenues est une

base de E formée de vecteurs qui sont à la fois des vecteurs propres de f et des vecteurs propres de g .

Autrement dit, il existe une base \mathcal{B} de E telle que $\text{Mat}_{\mathcal{B}}(f)$ et $\text{Mat}_{\mathcal{B}}(g)$ soient toutes les deux diagonales.

VECTEURS ALÉATOIRES

Cours	341
11.1 Généralités	341
11.2 Variables aléatoires indépendantes	343
11.3 Lois de sommes, de max, de min	347
Exercices	351
Correction des exercices	355

Dans tout le chapitre, (Ω, \mathcal{A}, P) désigne un espace probabilisé.

11.1 GÉNÉRALITÉS

11.1.1 Définitions

Définition 11.1 – Un vecteur aléatoire sur (Ω, \mathcal{A}, P) à valeurs dans \mathbf{R}^n est la donnée d'un n -uplet de variables aléatoires réelles (X_1, \dots, X_n) définies sur (Ω, \mathcal{A}, P) .

Exemple 11.2

Par exemple, un couple de variables aléatoires est un vecteur aléatoire à valeurs dans \mathbf{R}^2 .

Définition 11.3 – La loi d'un vecteur aléatoire (X_1, \dots, X_n) est la fonction $F_{(X_1, \dots, X_n)}$ définie sur \mathbf{R}^n par $\forall (x_1, \dots, x_n) \in \mathbf{R}^n$

$$F_{(X_1, \dots, X_n)}(x_1, \dots, x_n) = P\left(\bigcap_{i=1}^n [X_i \leq x_i]\right) = P([X_1 \leq x_1] \cap \dots \cap [X_n \leq x_n]).$$

Remarque

Notons que cette fonction est une généralisation à n variables aléatoires de la fonction de répartition d'une variable aléatoire.

Définition 11.4 – Soit (X_1, \dots, X_n) un vecteur aléatoire à valeurs dans \mathbf{R}^n . La loi de X_i est appelée $i^{\text{ème}}$ loi marginale de (X_1, \dots, X_n) .

⚠ La connaissance de la loi du vecteur (X_1, \dots, X_n) permet de retrouver les lois marginales, mais l'inverse n'est pas vrai : connaître les n lois marginales d'un vecteur ne permet pas de retrouver la loi du vecteur. Cela dépend des éventuels «liens» entre les X_i (indépendance ou non).

Par exemple, si X_1 et X_2 sont deux variables aléatoires indépendantes de même loi $\mathcal{P}(\lambda)$, alors les couples (X_1, X_2) et (X_1, X_1) n'ont pas même loi, bien qu'ils aient les mêmes lois marginales.

Continue ?

Nous n'avons pas encore défini les fonctions continues sur \mathbf{R}^n (cela viendra au chapitre 13), mais on peut pour l'instant omettre cette hypothèse, puisque toutes les fonctions de \mathbf{R}^n dans \mathbf{R} que nous manipulerons cette année seront continues.

Proposition 11.5 : Si deux vecteurs aléatoires (X_1, \dots, X_n) et (Y_1, \dots, Y_n) ont même loi et si g est une fonction continue sur \mathbf{R}^n à valeurs dans \mathbf{R} , alors les variables aléatoires $g(X_1, \dots, X_n)$ et $g(Y_1, \dots, Y_n)$ ont même loi.

Exemple 11.6

Supposons que $X_1, Y_1 \leftrightarrow \gamma(v_1)$, que $X_2, Y_2 \leftrightarrow \gamma(v_2)$, et que X_1 et X_2 sont indépendantes, de même que Y_1 et Y_2 .

Alors les couples (X_1, X_2) et (Y_1, Y_2) ont même loi puisque pour $(x, y) \in \mathbf{R}^2$, on a

$$\begin{aligned} F_{(X_1, X_2)}(x, y) &= P([X_1 \leq x] \cap [X_2 \leq y]) = P(X_1 \leq x)P(X_2 \leq y) \\ &= P(Y_1 \leq x)P(Y_2 \leq y) \\ &= P([Y_1 \leq x] \cap [Y_2 \leq y]) = F_{(Y_1, Y_2)}(x, y). \end{aligned}$$

Alors $Z = \max(X_1, X_2)$ et $Z' = \max(Y_1, Y_2)$ ont même loi.

Notons que pour cela, il faut remarquer que la fonction $(x, y) \mapsto \max(x, y)$ est continue sur \mathbf{R}^2 , ce qui sera détaillé plus tard.

X_1 et X_2 sont indépendantes.

X_1 et X_2 ont même loi, de même que Y_1 et Y_2 .

Y_1 et Y_2 sont indépendantes.

Évident ?

Bien que le résultat semble très intuitif sur cet exemple, la proposition précédente est loin d'être une tautologie.

11.1.2 Vecteurs aléatoires discrets

Définition 11.7 – Un vecteur aléatoire (X_1, \dots, X_n) sur (Ω, \mathcal{A}, P) est discret si toutes les variables X_1, \dots, X_n sont des variables aléatoires discrètes.

Remarque. Si (X_1, \dots, X_n) est un vecteur aléatoire discret, alors $X_1(\Omega) \times X_2(\Omega) \times \dots \times X_n(\Omega)$ est dénombrable car il s'agit d'un produit d'ensembles dénombrables.

Proposition 11.8 : Si (X_1, \dots, X_n) est un vecteur aléatoire discret, alors la loi de (X_1, \dots, X_n) est entièrement caractérisée par les

$$P([X_1 = x_1] \cap \dots \cap [X_n = x_n]), x_1 \in X_1(\Omega), \dots, x_n \in X_n(\Omega).$$

Remarques. • Il s'agit du prolongement de ce que nous savons déjà pour une variable discrètes¹ : on ne manipule jamais la fonction de répartition d'une variable discrète, et on lui préfère les $P(X = x)$, $x \in X(\Omega)$.

• Ce résultat est bien entendu faux dès que l'une des variables aléatoires X_i est à densité, car alors $P(X_i = x) = 0$ pour tout x réel.

¹ Et pour les couples de variables aléatoires discrètes.

Proposition 11.9 : Soit (X_1, \dots, X_n) un vecteur aléatoire discret. Alors la $i^{\text{ème}}$ loi marginale de (X_1, \dots, X_n) (c'est-à-dire la loi de X_i) est donnée par : $\forall x_i \in X_i(\Omega)$

$$P(X_i = x_i) = \sum_{(x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n) \in X_1(\Omega) \times \dots \times X_{i-1}(\Omega) \times X_{i+1}(\Omega) \times \dots \times X_n(\Omega)} P\left(\bigcap_{j=1}^n [X_j = x_j]\right).$$

Autrement dit, on omet le terme correspondant à X_i pour la somme.

Démonstration. Il s'agit de la formule des probabilités totales appliquée au système complet d'événements

$$\{(x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n) \in X_1(\Omega) \times \dots \times X_{i-1}(\Omega) \times X_{i+1}(\Omega) \times \dots \times X_n(\Omega)\}.$$

□

Proposition 11.10 : Si g est une fonction définie sur \mathbf{R}^n et si (X_1, \dots, X_n) est un vecteur aléatoire discret sur (Ω, \mathcal{A}, P) , alors $Y = g(X_1, \dots, X_n)$ est une variable aléatoire discrète, et

$$\forall y \in Y(\Omega), P(Y = y) = \sum_{\substack{(x_1, \dots, x_n) \in X_1(\Omega) \times \dots \times X_n(\Omega) \\ g(x_1, \dots, x_n) = y}} P([X_1 = x_1] \cap \dots \cap [X_n = x_n]).$$

Remarque. En particulier, si (X_1, \dots, X_n) est un vecteur aléatoire discret, alors

$$\sum_{i=1}^n X_i, \min(X_1, \dots, X_n), \max(X_1, \dots, X_n) \text{ et } \prod_{i=1}^n X_i$$

sont des variables aléatoires discrètes.

11.2 VARIABLES ALÉATOIRES INDÉPENDANTES

11.2.1 Définition

Définition 11.11 – Soient (X_1, \dots, X_n) des variables aléatoires sur (Ω, \mathcal{A}, P) . On dit que X_1, \dots, X_n sont mutuellement indépendantes (ou indépendantes) si

$$\forall (x_1, \dots, x_n) \in \mathbf{R}^n, P\left(\prod_{i=1}^n [X_i \leq x_i]\right) = \prod_{i=1}^n P(X_i \leq x_i).$$

Remarques. • De manière équivalente, cela signifie que la loi du vecteur aléatoire (X_1, \dots, X_n) est donnée par

$$\forall (x_1, \dots, x_n) \in \mathbf{R}^n, F_{(X_1, \dots, X_n)}(x_1, \dots, x_n) = \prod_{i=1}^n F_{X_i}(x_i).$$

Dans ce cas, la connaissance des lois marginales suffit à déterminer la loi du vecteur (X_1, \dots, X_n) .

• Dire que n variables aléatoires sont mutuellement indépendantes n'est pas équivalent à dire qu'elles sont deux à deux indépendantes.

Par exemple, lançons deux dés, et notons X_1 la variable aléatoire qui vaut 1 si le résultat du premier dé est pair et 0 sinon, X_2 la variable aléatoire qui vaut 1 si le résultat du second dé est pair et 0 sinon et enfin X_3 la variable aléatoire qui vaut 1 si la somme des deux dés est paire et 0 sinon.

Il est évident que X_1 et X_2 suivent une loi de Bernoulli de paramètre $\frac{1}{2}$, et sont indépendantes.

D'autre part, on a

$$\begin{aligned} P(X_3 = 1) &= P([X_1 = 0] \cap [X_2 = 0] \cup ([X_1 = 1] \cap [X_2 = 1])) \\ &= P(X_1 = 0)P(X_2 = 0) + P(X_1 = 1)P(X_2 = 1) \\ &= \frac{1}{2^2} + \frac{1}{2^2} = \frac{1}{2}. \end{aligned}$$

Et donc X_3 suit également la loi $\mathcal{B}(\frac{1}{2})$.

On a alors $P([X_1 = 1] \cap [X_3 = 1]) = P([X_1 = 1] \cap [X_2 = 1]) = \frac{1}{4}$ et donc X_1 et X_3 sont indépendantes.

De même, X_2 et X_3 sont indépendantes. Et pourtant on a

$$P([X_1 = 1] \cap [X_2 = 1] \cap [X_3 = 0]) = 0 \neq P(X_1 = 1)P(X_2 = 1)P(X_3 = 0)$$

de sorte que (X_1, X_2, X_3) ne sont pas mutuellement indépendantes.

Définition 11.12 – Soit $(X_n)_{n \geq 0}$ une suite **infinie** de variables aléatoires définies sur le même espace probabilisé (Ω, \mathcal{A}, P) . On dit que les variables $(X_n)_{n \geq 0}$ sont indépendantes si pour toute partie **finie** $I \subset \mathbf{N}$, les variables $(X_i)_{i \in I}$ sont mutuellement indépendantes.

11.2.2 Diverses caractérisations de l'indépendance



Rappel

Deux variables aléatoires X et Y suivant des lois de Bernoulli sont indépendantes si et seulement si les événements $[X = 1]$ et $[Y = 1]$ sont indépendants.

Remarque

Nous sommes obligés d'adapter la définition utilisée pour une famille finie, car nous ne pouvons faire de produit infini de probabilités.



Proposition 11.13 : Soient X_1, \dots, X_n des variables aléatoires sur (Ω, \mathcal{A}, P) . Alors les assertions suivantes sont équivalentes :

- X_1, \dots, X_n sont mutuellement indépendantes
- pour tous I_1, \dots, I_n intervalles de \mathbf{R} , on a

$$P([X_1 \in I_1] \cap \dots \cap [X_n \in I_n]) = \prod_{i=1}^n P(X_i \in I_i)$$

- pour tout $A_1 \in \mathcal{A}_{X_1}, \dots, A_n \in \mathcal{A}_{X_n}$, les événements A_1, \dots, A_n sont mutuellement indépendants.

Remarques. • Rappelons que \mathcal{A}_{X_1} , est la tribu engendrée par X_1 , c'est-à-dire la plus petite tribu contenant tous les événements «formés à partir de X_1 ». Autrement dit, le dernier point signifie que si (A_1, \dots, A_n) est une famille d'événements telle que « A_i dépend de X_i », alors (A_1, \dots, A_n) sont mutuellement indépendants.

• En particulier, en prenant $I_3 = \dots = I_n = \mathbf{R}$, on constate que si X_1, \dots, X_n sont indépendantes, alors X_1 et X_2 le sont également.

On généralise facilement : toute sous-famille d'une famille de variables aléatoires mutuellement indépendantes est encore une famille de variables mutuellement indépendantes.

Dans le cas de variables aléatoires discrètes, nous savons que la loi de (X_1, \dots, X_n) est entièrement déterminée par les $P\left(\bigcap_{i=1}^n [X_i = x_i]\right)$, et donc

Proposition 11.14 : Soit (X_1, \dots, X_n) un vecteur aléatoire discret. Alors les variables X_1, \dots, X_n sont mutuellement indépendantes si et seulement si

$$\forall (x_1, \dots, x_n) \in X_1(\Omega) \times \dots \times X_n(\Omega), P\left(\bigcap_{i=1}^n [X_i = x_i]\right) = \prod_{i=1}^n P(X_i = x_i).$$

Détails

On a $P(X_i \in \mathbf{R}) = 1$.

11.2.3 Le lemme des coalitions

Proposition 11.15 (Lemme des coalitions) : Soit (X_1, \dots, X_n) une famille de variables aléatoires mutuellement indépendantes sur (Ω, \mathcal{A}, P) . Alors toute variable aléatoire fonction de k d'entre elles est indépendante de toute variable aléatoire fonction des $n - k$ autres.

Démonstration. Ne donnons pas une preuve, mais plutôt une idée de la preuve : soient donc $f : \mathbf{R}^k \rightarrow \mathbf{R}$ et $g : \mathbf{R}^{n-k} \rightarrow \mathbf{R}$ deux fonctions, et soient

$$X = f(X_1, \dots, X_k), Y = g(X_{k+1}, \dots, X_n).$$

Par définition, X et Y sont indépendantes si et seulement si pour tout événement $A \in \mathcal{A}_X$ et $B \in \mathcal{A}_Y$, A et B sont indépendants.

Mais un événement de \mathcal{A}_X est un événement «dépendant de X_1, \dots, X_p », et un événement dépendant de Y est un événement «dépendant de X_{p+1}, \dots, X_n ».

Par définition de l'indépendance de X_1, \dots, X_n , de tels événements doivent être indépendants. \square

Remarques. • En particulier, si X et Y sont indépendantes, alors $f(X)$ et $g(Y)$ sont indépendantes, quelles que soient les fonctions f et g . Par exemple, X^2 est indépendante de e^{Y^2} .

• Ce résultat peut être généralisé de la manière suivante : si (X_1, \dots, X_n) sont mutuellement indépendantes, et si $1 \leq k_1 < k_2 < \dots < k_p = n$, alors

$$Y_1 = f_1(X_1, \dots, X_{k_1}), Y_2 = f_2(X_{k_1+1}, \dots, X_{k_2}), \dots, Y_p = f_p(X_{k_{p-1}+1}, \dots, X_n)$$

sont mutuellement indépendantes.

Évident ?

Ce résultat est très intuitif, et semble évident, mais il est en fait assez délicat à prouver. On n'oubliera pas de le mentionner lorsqu'on l'utilise, même si c'est l'intuition qui nous dicte l'indépendance.



Exemples 11.16

- Si (X_1, X_2, X_3) sont mutuellement indépendantes, alors $Y = X_1 + X_2$ est indépendante de $Z = e^{X_3}$.
- Si (X_1, \dots, X_{2n}) sont mutuellement indépendantes, alors $Y = \max(X_1, \dots, X_n)$ est indépendante de $\min(X_{n+1}, \dots, X_{2n})$.

Exemple 11.17 D'après Oral HEC

Soient X_1, \dots, X_n des variables aléatoires indépendantes, suivant toutes la loi uniforme sur $[0, 1]$.

On souhaite calculer $P([X_n \geq X_1] \cap [X_n \geq X_2] \cap \dots \cap [X_n \geq X_{n-1}])$.

Notons que $[X_n \geq X_i] = [X_n - X_i \geq 0]$. Malheureusement, les variables aléatoires $X_n - X_i$ ne sont a priori pas indépendantes² et donc les événements $[X_n \geq X_1], \dots, [X_n \geq X_{n-1}]$ ne sont pas indépendants.

Mais on a également $[X_n \geq X_1] \cap \dots \cap [X_n \geq X_{n-1}] = [X_n \geq \max(X_1, \dots, X_{n-1})]$. Ceci s'écrit encore $[X_n - \max(X_1, \dots, X_{n-1}) \geq 0]$, et les deux variables X_n et $\max(X_1, \dots, X_{n-1})$ sont indépendantes par le lemme des coalitions, et donc nous pouvons utiliser le produit de convolution.

Pour cela, commençons par déterminer la fonction de répartition de $\max(X_1, \dots, X_{n-1})$.

Pour $x \in \mathbf{R}$, on a $P(\max(X_1, \dots, X_{n-1}) \leq x) = P([X_1 \leq x] \cap \dots \cap [X_{n-1} \leq x])$, et par indépendance des X_i ,

$$P(\max(X_1, \dots, X_{n-1}) \leq x) = P(X_1 \leq x) \cdots P(X_{n-1} \leq x) = P(X_1 \leq x)^{n-1} = F_{X_1}(x)^{n-1}$$

$$= \begin{cases} 0 & \text{si } x < 0 \\ x^{n-1} & \text{si } 0 \leq x \leq 1 \\ 1 & \text{si } x > 1 \end{cases}$$

Cette fonction est continue³ sur \mathbf{R} , et \mathcal{C}^1 sauf peut-être en 0 et en 1.

Donc $\max(X_1, \dots, X_{n-1})$ est une variable à densité, et une densité en est donnée par

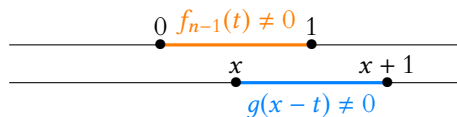
$$f_{n-1}(x) = \begin{cases} (n-1)x^{n-2} & \text{si } 0 \leq x \leq 1 \\ 0 & \text{sinon} \end{cases}$$

De plus, une densité de $-X_n$ est donnée par $g(t) = \begin{cases} 1 & \text{si } -1 \leq t \leq 0 \\ 0 & \text{sinon} \end{cases}$.

Par indépendance de X_n et $\max(X_1, \dots, X_{n-1})$, une densité de $\max(X_1, \dots, X_{n-1}) - X_n$ est donnée⁴ par

$$h_n(x) = \int_{-\infty}^{+\infty} f_{n-1}(t)g(x-t) dt.$$

On a alors $g(x-t) \neq 0 \Leftrightarrow -1 \leq x-t \leq 0 \Leftrightarrow x \leq t \leq x+1$.



Si $x \geq 1$ ou $x+1 \leq 0 \Leftrightarrow x \leq -1$, alors $h_n(x) = 0$.

Si $0 \leq x \leq 1$, alors

$$h_n(x) = \int_x^1 f_{n-1}(t)g(x-t) dt = \int_x^1 (n-1)t^{n-2} dt = [t^{n-1}]_x^1 = 1 - x^{n-1}.$$

Si $-1 \leq x \leq 0$, alors

$$h_n(x) = \int_0^{x+1} f_{n-1}(t)g(x-t) dt = \int_0^{x+1} (n-1)t^{n-2} dt = [t^{n-1}]_0^{x+1} = (x+1)^{n-1}.$$

² En tous cas, le lemme des coalitions ne nous dit rien à ce sujet.

³ Le plus simple pour le voir est probablement de remarquer que F_{X_1} est continue car fonction de répartition d'une variable à densité, et donc $F_{X_1}^{n-1}$ est également continue.

⁴ Notons que g est bornée, et donc h_n sera bien une densité.

Donc une densité de $Z_n = \max(X_1, \dots, X_{n-1}) - X_n$ est

$$h_n : x \mapsto \begin{cases} (x+1)^{n-1} & \text{si } -1 \leq x \leq 0 \\ 1-x^{n-1} & \text{si } 0 < x \leq 1 \\ 0 & \text{sinon} \end{cases}$$

Et alors, la probabilité initialement cherchée est

$$\begin{aligned} P([X_n \geq X_1] \cap \dots \cap [X_n \geq X_{n-1}]) &= P(\max(X_1, \dots, X_{n-1}) \leq X_n) \\ &= P(\max(X_1, \dots, X_{n-1}) - X_n \leq 0) \\ &= P(Z_n \leq 0) \\ &= \int_{-\infty}^0 h_n(x) dx \\ &= \int_{-1}^0 (x+1)^{n-1} dx = \left[\frac{(x+1)^n}{n} \right]_{-1}^0 \\ &= \frac{1}{n}. \end{aligned}$$

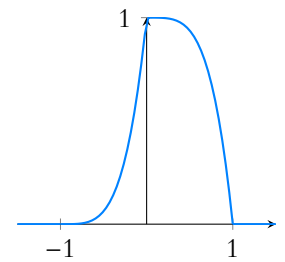


FIGURE 11.1– La densité h pour $n = 5$.

Logique ?

Ce résultat est largement intuitif : les X_i jouant des rôles symétriques, la probabilité que X_1 soit la plus grande soit être la même que celle que X_2 soit la plus grande, etc. Et donc chacune d'entre elles possède une probabilité égale à $\frac{1}{n}$ d'être la plus grande.

11.2.4 Espérance, variance

Proposition 11.18 : Soient (X_1, \dots, X_n) des variables aléatoires mutuellement indépendantes et admettant toutes une espérance. Alors la variable aléatoire $X_1 \cdots X_n$ admet une espérance, et

$$E(X_1 \cdots X_n) = E(X_1) \cdots E(X_n).$$

Démonstration. La preuve se fait par récurrence sur n .

En effet, nous avons déjà prouvé⁵ ce résultat pour $n = 2$.

Supposons donc qu'il est vrai pour n variables indépendantes, et soient X_1, \dots, X_{n+1} des variables aléatoires indépendantes. Alors $X_1 \cdots X_n$ et X_{n+1} sont indépendantes par le lemme des coalitions, et admettent toutes deux une espérance⁶.

Donc $X_1 \cdots X_n X_{n+1}$ admet une espérance et

$$E(X_1 \cdots X_n X_{n+1}) = E(X_1 \cdots X_n)E(X_{n+1}) = E(X_1) \cdots E(X_n)E(X_{n+1}).$$

□

Proposition 11.19 : Soient X_1, \dots, X_n n variables aléatoires mutuellement indépendantes sur (Ω, \mathcal{A}, P) admettant une variance. Alors $X_1 + \dots + X_n$ admet une variance, et

$$V\left(\sum_{i=1}^n X_i\right) = \sum_{i=1}^n V(X_i).$$

Plus généralement, pour tout $(\lambda_1, \dots, \lambda_n) \in \mathbf{R}^n$, $\lambda_1 X_1 + \dots + \lambda_n X_n$ admet une variance, et

$$V\left(\sum_{i=1}^n \lambda_i X_i\right) = \sum_{i=1}^n \lambda_i^2 V(X_i).$$

Démonstration. Prouvons le résultat par récurrence sur n .

Pour $n = 2$, le résultat est connu.

Supposons le résultat vrai pour n variables aléatoires, et soient (X_1, \dots, X_{n+1}) $n + 1$ variables aléatoires mutuellement indépendantes. Alors X_1, \dots, X_n sont mutuellement indépendantes, et donc par hypothèse de récurrence,

$$V(X_1 + \dots + X_n) = \sum_{i=1}^n V(X_i).$$

⁵ Au moins pour les variables discrètes, et nous avons admis qu'il reste valable pour des variables à densité.

⁶ Pour la première des deux, c'est précisément l'hypothèse de récurrence.

Remarque

Pour n variables discrètes qui ne sont pas indépendantes, il existe une formule pour calculer la variance de la somme en utilisant la covariance, mais elle n'est plus au programme. Voir l'exercice 11.4.

Par le lemme des coalitions, $X_1 + \dots + X_n$ et X_{n+1} sont indépendantes, et donc

$$V(X_1 + \dots + X_n + X_{n+1}) = V(X_1 + \dots + X_n) + V(X_{n+1}) = \sum_{i=1}^n V(X_i) + V(X_{n+1}) = \sum_{i=1}^{n+1} V(X_i).$$

Ainsi, la propriété est vraie au rang $n + 1$, et donc par le principe de récurrence, elle est vraie pour tout $n \in \mathbf{N}$.

Pour le second point, notons que quitte à enlever les termes avec $\lambda_i = 0$, on peut supposer que tous les λ_i sont non nuls.

Et alors $\lambda_1 X_1, \dots, \lambda_n X_n$ sont mutuellement indépendantes, admettent une variance, de sorte que

$$V\left(\sum_{i=1}^n \lambda_i X_i\right) = \sum_{i=1}^n V(\lambda_i X_i) = \sum_{i=1}^n \lambda_i^2 V(X_i).$$

□

11.3 LOIS DE SOMMES, DE MAX, DE MIN

11.3.1 Stabilité des lois usuelles

Proposition 11.20 : Soit $p \in [0, 1]$, et soient X_1, \dots, X_n des variables aléatoires sur (Ω, \mathcal{A}, P) , mutuellement indépendantes telles que $X_i \hookrightarrow \mathcal{B}(n_i, p)$. Alors

$$\sum_{i=1}^n X_i \hookrightarrow \mathcal{B}\left(\sum_{i=1}^n n_i, p\right).$$

⚠ Attention !

Ce résultat n'est plus vrai si on enlève l'indépendance, ni si le second paramètre n'est pas le même pour les différentes lois binomiales.

Démonstration. Le résultat a déjà été démontré pour $n = 2$.

Procédons par récurrence sur n , et supposons le résultat vrai pour n variables aléatoires indépendantes, et soient X_1, \dots, X_{n+1} $n + 1$ variables aléatoires indépendantes sur (Ω, \mathcal{A}, P) telles que $X_i \hookrightarrow \mathcal{B}(n_i, p)$.

Par le lemme des coalitions, $X_1 + \dots + X_n$ est indépendante de X_{n+1} . De plus, par hypothèse de récurrence,

$$X_1 + \dots + X_n \hookrightarrow \mathcal{B}\left(\sum_{i=1}^n n_i, p\right).$$

Ainsi,

$$(X_1 + \dots + X_n) + X_{n+1} \hookrightarrow \mathcal{B}\left(\sum_{i=1}^n n_i + n_{n+1}, p\right) = \mathcal{B}\left(\sum_{i=1}^{n+1} n_i, p\right).$$

Par le principe de récurrence, la propriété est donc vraie pour tout n . □

Corollaire 11.21 – En particulier, si X_1, \dots, X_n sont n variables aléatoires indépendantes suivant une loi de Bernoulli de paramètre p , alors $\sum_{i=1}^n X_i$ suit la loi binomiale $\mathcal{B}(n, p)$.

Démonstration. Il suffit de remarquer que la loi de Bernoulli de paramètre p est également la loi binomiale de paramètres 1 et p . □

Remarques. • Cela permet de justifier un principe connu depuis longtemps : une loi binomiale $\mathcal{B}(n, p)$ correspond au nombre de succès lors d'une répétition de n épreuves de Bernoulli indépendantes de même paramètre p .

• On retrouve ainsi sans calculs le fait que l'espérance d'une loi binomiale est np , et que sa variance est $np(1 - p)$, car la variance d'une loi de Bernoulli de paramètre p est $p(1 - p)$.

Proposition 11.22 : Soient X_1, \dots, X_n des variables aléatoires indépendantes sur (Ω, \mathcal{A}, P) telles que $X_i \hookrightarrow \mathcal{P}(\lambda_i)$. Alors

$$\sum_{i=1}^n X_i \hookrightarrow \mathcal{P}\left(\sum_{i=1}^n \lambda_i\right).$$

Démonstration. Par récurrence sur n . □

Proposition 11.23 : Soient X_1, \dots, X_n n variables aléatoires indépendantes sur (Ω, \mathcal{A}, P) telles que $X_i \hookrightarrow \gamma(v_i)$. Alors

$$\sum_{i=1}^n X_i \hookrightarrow \gamma(v_1 + \dots + v_n).$$

Démonstration. Par récurrence sur n . □

Corollaire 11.24 – Soient X_1, \dots, X_n des variables aléatoires mutuellement indépendantes sur (Ω, \mathcal{A}, P) suivant toutes la loi exponentielle de paramètre 1. Alors $X_1 + \dots + X_n$ suit la loi $\gamma(n)$.

Pour déterminer la loi d'une somme de variables aléatoires indépendantes suivant une loi exponentielle de même paramètre, on se ramènera toujours à des variables de loi $\mathcal{E}(1)$ en utilisant la propriété suivante : $\lambda X \hookrightarrow \mathcal{E}(1) \Leftrightarrow X \hookrightarrow \mathcal{E}(\lambda)$.

Exemple 11.25 Somme de lois exponentielles : à savoir refaire

Soient X_1, \dots, X_n n variables aléatoires suivant la loi $\mathcal{E}(\lambda)$, et soit $Y = X_1 + \dots + X_n$. Alors $Y = \frac{1}{\lambda}(\lambda X_1 + \dots + \lambda X_n)$.

Mais les λX_i sont mutuellement indépendantes (par le lemme des coalitions), et suivent toutes la loi $\mathcal{E}(1)$, donc⁷

$$\lambda X_1 + \dots + \lambda X_n = \lambda Y \hookrightarrow \gamma(n).$$

Ainsi, une densité de λY est donnée par

$$f : t \mapsto \begin{cases} 0 & \text{si } t \leq 0 \\ \frac{1}{\Gamma(n)} t^{n-1} e^{-t} = \frac{1}{(n-1)!} t^{n-1} e^{-t} & \text{si } t > 0 \end{cases}$$

Alors par transformation affine, $Y = \frac{1}{\lambda}(\lambda Y)$ est encore une variable à densité, et l'une de ses densités est donnée par

$$t \mapsto \lambda f(\lambda t) = \begin{cases} 0 & \text{si } t \leq 0 \\ \frac{\lambda^n t^{n-1} e^{-\lambda t}}{(n-1)!} & \text{si } t > 0 \end{cases}$$

On dit alors que Y suit la loi $\Gamma\left(\frac{1}{\lambda}, n\right)$ (hors programme).

⁷ **Rappel** : la loi exponentielle $\mathcal{E}(1)$ est également la loi gamma $\gamma(1)$.

Proposition 11.26 : Soient X_1, \dots, X_n des variables aléatoires indépendantes sur (Ω, \mathcal{A}, P) telles que $X_i \hookrightarrow \mathcal{N}(\mu_i, \sigma_i^2)$. Alors

$$\sum_{i=1}^n X_i \hookrightarrow \mathcal{N}\left(\sum_{i=1}^n \mu_i, \sum_{i=1}^n \sigma_i^2\right).$$

Démonstration. Par récurrence sur n . □

En particulier, si X_1, \dots, X_n sont n variables aléatoires indépendantes suivant toutes la loi normale centrée réduite, alors $X_1 + \dots + X_n$ suit la loi $\mathcal{N}(0, n)$, et donc

$$\frac{X_1 + \dots + X_n}{\sqrt{n}} \hookrightarrow \mathcal{N}(0, 1).$$

11.3.2 Maximum, minimum

Nous généralisons ici les résultats déjà connus sur le maximum et le minimum de deux variables aléatoires.

Définition 11.27 – Soient X_1, \dots, X_n des variables aléatoires définies sur (Ω, \mathcal{A}, P) . On définit deux variables aléatoires $Y = \min(X_1, \dots, X_n)$ et $Z = \max(X_1, \dots, X_n)$ par

$$\forall \omega \in \Omega, Y(\omega) = \min(X_1(\omega), \dots, X_n(\omega)) \text{ et } Z(\omega) = \max(X_1(\omega), \dots, X_n(\omega)).$$

Remarque. On rencontre aussi parfois les notations $\sup(X_1, \dots, X_n)$ au lieu de $\max(X_1, \dots, X_n)$ et $\inf(X_1, \dots, X_n)$ au lieu de $\min(X_1, \dots, X_n)$.

Proposition 11.28 : Soient X_1, \dots, X_n n variables aléatoires sur (Ω, \mathcal{A}, P) . Alors si $Z = \max(X_1, \dots, X_n)$, on a, pour tout $x \in \mathbf{R}$

$$[\max(X_1, \dots, X_n) \leq x] = \bigcap_{i=1}^n [X_i \leq x].$$

En particulier, si X_1, \dots, X_n sont mutuellement indépendantes, alors

$$F_Z(x) = \prod_{i=1}^n F_{X_i}(x).$$

Démonstration. Il s'agit essentiellement de bon sens : le plus grand des X_i est inférieur ou égal à x si et seulement si tous les X_i sont inférieurs à x . Donc

$$[\max(X_1, \dots, X_n) \leq x] = \bigcap_{i=1}^n [X_i \leq x].$$

□

En particulier, si les X_i sont des variables à densité indépendantes, alors les F_{X_i} sont toutes continues sur \mathbf{R} et \mathcal{C}^1 sauf éventuellement en un nombre fini de points. Et donc il en est de même de F_Z , de sorte que Z est une variable à densité.

Exemple 11.29

Soient X_1, \dots, X_n n variables aléatoires indépendantes sur (Ω, \mathcal{A}, P) suivant toutes la Bernoulli de paramètre p , et soit $Z = \max(X_1, \dots, X_n)$. Alors Z ne peut prendre que les valeurs 0 et 1 : c'est donc une variable de Bernoulli.

Attention !

L'écart type de $\sum_{i=1}^n X_i$ est $\sqrt{\sum_{i=1}^n \sigma_i^2}$, et non $\sum_{i=1}^n \sigma_i$.

Remarque

Nous n'en faisons toutefois pas un résultat général, s'il est demandé de prouver qu'un minimum de variables à densité est encore une variable à densité, il faudra être capable de tenir ce raisonnement, ou de calculer explicitement la fonction F_Z et montrer qu'elle est bien continue et \mathcal{C}^1 sauf en un nombre fini de points.

De plus, on a

$$P(Z = 0) = P(Z \leq 0) = P\left(\bigcap_{i=1}^n [X_i \leq 0]\right) = P\left(\bigcap_{i=1}^n [X_i = 0]\right) = (1-p)^n.$$

Donc Z suit la loi de Bernoulli $\mathcal{B}(1 - (1-p)^n)$.

Remarque. Rappelons que pour une variable aléatoire X à valeurs dans \mathbf{Z} , on obtient $P(X = k)$ à partir de la fonction de répartition grâce à la formule

$$P(X = k) = P(X \leq k) - P(X \leq k-1) = F_X(k) - F_X(k-1).$$

Proposition 11.30 : Soient X_1, \dots, X_n n variables aléatoires définies sur (Ω, \mathcal{A}, P) . Alors, en notant $Y = \min(X_1, \dots, X_n)$, on a, pour tout $x \in \mathbf{R}$

$$[\min(X_1, \dots, X_n) > x] = \bigcap_{i=1}^n [X_i > x].$$

En particulier, si X_1, \dots, X_n sont mutuellement indépendantes, alors

$$1 - F_Y(x) = \prod_{i=1}^n (1 - F_{X_i}(x)).$$

Comme pour le cas d'un maximum, remarquons que si les X_i sont indépendantes et sont toutes des variables à densité, alors F_Y est continue sur \mathbf{R} et \mathcal{C}^1 sauf éventuellement en un nombre fini de points, de sorte que Y est encore une variable à densité.

Remarque

On peut également remplacer les inégalités strictes par des inégalités larges. On préférera toutefois les inégalités strictes, car le contraire de $[X > x]$ est $[X \leq x]$, ce qui permet de faire apparaître directement la fonction de répartition (bien entendu, si les X_i sont à densité, on peut également utiliser les inégalités larges).

Exemple 11.31

Soient X_1, \dots, X_n n variables aléatoires indépendantes de même loi géométrique $\mathcal{G}(p)$. Alors pour tout entier $k \geq 1$, on a

$$\begin{aligned} P(X_i \geq k) &= \sum_{m=k}^{+\infty} P(X_i = m) = \sum_{m=k}^{+\infty} p(1-p)^{m-1} \\ &= p(1-p)^{k-1} \sum_{i=0}^{+\infty} (1-p)^i = p \frac{(1-p)^{k-1}}{1-(1-p)} = (1-p)^{k-1}. \end{aligned}$$

Donc si $Y = \min(X_1, \dots, X_n)$, il vient

$$\forall k \in \mathbf{N}^*, P(Y \geq k) = \prod_{i=1}^n P(X_i \geq k) = (1-p)^{n(k-1)}.$$

En particulier, on en déduit que

$$\forall k \geq 1, P(Y = k) = P(Y \geq k) - P(Y \geq k+1) = (1-p)^{n(k-1)} - (1-p)^{nk} = (1-p)^{n(k-1)}(1-(1-p)^n)$$

En posant $q = 1 - (1-p)^n$, on a alors $P(Y = k) = q(1-q)^{k-1}$, et donc $Y \hookrightarrow \mathcal{G}(q)$.

Intuition

Si on voit X_i comme le nombre d'essais nécessaires à l'obtention d'un premier succès lors d'une répétition d'épreuves de Bernoulli, alors $[X_i \geq k]$ est l'événement «il faut au moins k essais avant le premier succès», qui n'est autre que l'événement «les $k-1$ premiers essais sont des échecs». Et alors il est évident que la probabilité de ce dernier événement est $(1-p)^{k-1}$.

EXERCICES DU CHAPITRE 11

EXERCICE 11.1 Minimum de lois exponentielles

Soit $\lambda > 0$ et soient X_1, \dots, X_n des variables aléatoires mutuellement indépendantes, telles que X_i suit la loi exponentielle de paramètre $\lambda_i > 0$. On pose $Y = \min(X_1, \dots, X_n)$. Déterminer la loi de Y , et en déduire $E(Y)$.

F

EXERCICE 11.2 Minimum et maximum de lois uniformes

Soit $a > 0$ et soient X_1, \dots, X_n n variables aléatoires indépendantes suivant la loi uniforme $\mathcal{U}([0, a])$. On pose

$$Y = \max(X_1, \dots, X_n) \text{ et } Z = \min(X_1, \dots, X_n).$$

Montrer que Y et Z sont des variables à densité et en déterminer une densité.

PD

EXERCICE 11.3 Somme de lois exponentielles

Soient X_1, X_2, X_3 des variables indépendantes de loi $\mathcal{E}(5)$.

F

- 1) Déterminer la loi de $5(X_1 + X_2 + X_3)$.
- 2) En déduire une densité de $X_1 + X_2 + X_3$.

EXERCICE 11.4 Soient X, Y, Z trois variables aléatoires indépendantes suivant toutes la loi de Poisson $\mathcal{P}(\lambda)$. Montrer que $\frac{X+Y}{1+Z}$ admet une espérance, et la déterminer.

PD

EXERCICE 11.5 Soient (X_1, \dots, X_n) des variables aléatoires indépendantes suivant la loi de Bernoulli $\mathcal{B}(p)$, $p \in]0, 1[$. On pose, pour tout $i \in \llbracket 1, n-1 \rrbracket$, $Y_i = X_i X_{i+1}$.

AD

Montrer que Y_i et Y_j sont indépendantes si et seulement si $|i - j| > 1$.

EXERCICE 11.6 Un sac contient N boules numérotées de 1 à N . On effectue dans ce sac n tirages d'une boule avec remise et on note Z_n le plus petit numéro obtenu. Déterminer la loi de Z_n .

AD

EXERCICE 11.7 Soient X_1, \dots, X_n des variables aléatoires indépendantes, de même loi donnée par

PD

$$\forall i \in \llbracket 1, n \rrbracket, P(X_i = 1) = \frac{1}{3}, P(X_i = 2) = \frac{2}{3}.$$

On pose $S_n = \sum_{k=1}^n X_k$.

- 1) Déterminer la loi de $Y_i = X_i - 1$.
- 2) Quelle est la loi de $Y_1 + \dots + Y_n$?
- 3) En déduire la loi de S_n .

EXERCICE 11.8 Une urne contient n boules, numérotées de 1 à n ($n \geq 3$).

AD

On tire simultanément trois boules dans l'urne, et on note X_1 le plus petit des trois numéros, X_3 le plus grand et X_2 le troisième.

- 1) Déterminer la loi du vecteur (X_1, X_2, X_3) .
- 2) En déduire la loi de X_2 , puis son espérance.

EXERCICE 11.9 Ecrimage 2015

PD

Soit f la fonction définie par $f(x) = \begin{cases} 0 & \text{si } x \leq 0 \\ \alpha & \text{si } x > 0 \end{cases}$

- 1) Montrer qu'il existe un unique $\alpha \in \mathbf{R}$ tel que f soit une densité de probabilités. Dans la suite, on considère que α prend cette valeur.
- 2) Soit X une variable aléatoire de densité f . Déterminer la fonction de répartition F_X de X . X admet-elle une espérance ?

Dans la suite, n désigne un entier supérieur ou égal à 2, et X_1, \dots, X_n sont des variables aléatoires indépendantes, admettant toutes f pour densité. On pose alors, pour tout $k \in \llbracket 1, n \rrbracket$ $Y_k = \min(X_1, \dots, X_k)$.

- 3) Déterminer la fonction de répartition de Y_2 , et en déduire que Y_2 est une variable à densité, dont on donnera une densité f_2 .

- 4) Montrer que pour tout $x \in \mathbf{R}_+^*$, on a $\text{Arctan}(x) + \text{Arctan}\left(\frac{1}{x}\right) = \frac{\pi}{2}$.
- 5) En déduire un équivalent de f_2 en $+\infty$.
- 6) Montrer que Y_2 admet une espérance. En déduire que pour tout $k \geq 2$, Y_k admet une espérance.

EXERCICE 11.10 Loi du χ^2

AD

Soient X_1, \dots, X_n des variables aléatoires indépendantes suivant toutes la loi normale centrée réduite. On pose $Y_n = \sum_{i=1}^n X_i^2$.

- 1) Montrer que la fonction de répartition de $Y = X_1^2$ est égale à $F_Y(x) = \begin{cases} 0 & \text{si } x \leq 0 \\ 2\Phi(\sqrt{x}) - 1 & \text{sinon} \end{cases}$
- 2) En déduire que Y est une variable à densité et en déterminer une densité.
- 3) Déterminer une densité de $\frac{Y}{2}$, et reconnaître sa loi.
- 4) En déduire la loi de $\frac{Y_n}{2}$, puis une densité de Y_n .

EXERCICE 11.11 (D'après Ecricome 1997)

AD

Soient X_1, X_2, X_3 trois variables aléatoires mutuellement indépendantes sur un même espace probabilisé, suivant toutes trois la loi uniforme sur $[0, 1]$.

- 1) Déterminer la fonction de répartition de $Y = \min(X_1, X_2)$.
- 2) En déduire que Y admet une densité, et déterminer l'une de ses densités.
- 3) Déterminer une densité de la variable aléatoire $Z = Y + X_3$.
- 4) En déduire $E(Z)$ et $V(Z)$.

Une situation concrète où une telle variable apparaît est la suivante : trois personnes entrent simultanément dans une poste ne comportant que deux guichets. On note X_i le temps passé par la i -ème personne au guichet. Alors $\min(X_1, X_2)$ est le temps que doit attendre la troisième personne pour qu'un guichet se libère, et Z est le temps total qu'elle passe dans le bureau de poste.

EXERCICE 11.12 Soit $p \in]0, 1[$. On considère une suite $(X_n)_{n \in \mathbf{N}^*}$ de variables aléatoires mutuellement indépendantes,

AD

de même loi, et telles que $P(X_1 = -1) = 1 - p$ et $P(X_1 = 1) = p$. On pose $S_n = \sum_{k=1}^n X_k$.

- 1) Déterminer deux réels a et b tels que $Y_n = aX_n + b$ suive une loi de Bernoulli de paramètre p .
- 2) En utilisant la question précédente, déterminer la loi de S_n et calculer son espérance.
- 3) **Application** : un ivrogne sort du bar, et chaque seconde avance d'un mètre vers la droite avec probabilité p , ou d'un mètre vers la gauche avec probabilité $1 - p$.
 - a) En moyenne, où se trouve-t-il après 10 secondes ?
 - b) Proposer une ligne de code Sci Lab simulant la marche d'un tel ivrogne pendant 1000 secondes.
 - c) Proposer un code Sci Lab permettant de tracer la trajectoire suivie par notre ivrogne (on tracera le temps en abscisse et sa position en ordonnée).

EXERCICE 11.13 La loi Γ à deux paramètres

AD

Soient b et ν deux réels strictement positifs. On définit une fonction

$$f : t \mapsto \begin{cases} 0 & \text{si } t \leq 0 \\ \frac{t^{\nu-1} e^{-t/b}}{b^\nu \Gamma(\nu)} & \text{si } t > 0 \end{cases}$$

- 1) Montrer que f est une densité de probabilité.

Dans la suite, on considère une variable aléatoire X admettant f pour densité, et on dit que X suit la loi Γ de paramètres b et ν , et on écrit $X \hookrightarrow \Gamma(b, \nu)$.

- 2) Quelle est la loi de $\frac{1}{b}X$? Quelle est la loi de X si $\nu = 1$?
- 3) Reconnaitre la loi $\Gamma(1, \nu)$.
- 4) Montrer que $\forall k \in \mathbf{N}^*$, X admet un moment d'ordre k , et le calculer. Vérifier que $E(X) = b\nu$ et $V(X) = b^2\nu$.
- 5) Soient X_1, X_2 deux variables aléatoires réelles indépendantes sur (Ω, \mathcal{A}, P) telles que $X_1 \hookrightarrow \Gamma(b, \nu_1)$ et $X_2 \hookrightarrow \Gamma(b, \nu_2)$. Montrer que $X_1 + X_2 \hookrightarrow \Gamma(b, \nu_1 + \nu_2)$. On pourra commencer par considérer la loi de $\frac{X_1}{b} + \frac{X_2}{b}$.

- 6) Plus généralement, si X_1, \dots, X_n sont n variables aléatoires mutuellement indépendantes telles que $X_i \hookrightarrow \Gamma(b, v_i)$, montrer que $X_1 + \dots + X_n \hookrightarrow \Gamma(b, v_1 + \dots + v_n)$.
- 7) Soient X_1, \dots, X_n des variables aléatoires mutuellement indépendantes sur (Ω, \mathcal{A}, P) suivant la même loi exponentielle $\mathcal{E}(\lambda)$. Montrer que $X_1 + \dots + X_n$ suit la loi $\Gamma(\frac{1}{\lambda}, n)$.
- 8) Soit Y une variable aléatoire suivant la loi $\mathcal{N}(0, 1)$. Montrer que Y^2 suit une loi Γ dont on précisera les paramètres. On admet que $\Gamma(1/2) = \sqrt{\pi}$ (*Question facultative : expliquer comment retrouver ce résultat grâce à la loi de Y^2*).

EXERCICE 11.14 Lois uniformes et loi de Pareto

Soient Z_1, \dots, Z_n des variables aléatoires indépendantes sur (Ω, \mathcal{A}, P) suivant toutes la loi uniforme sur $]0; a], a > 0$. On pose alors $X_i = \frac{1}{Z_i}$.

- 1) Déterminer la loi des X_i , et montrer que ce sont des variables à densité, dont on déterminera une densité.
- 2) Soit $X = \min(X_1, \dots, X_n)$. Montrer que X est une variable aléatoire à densité, et déterminer l'une de ses densités.
- 3) En déduire l'espérance et la variance de X .

EXERCICE 11.15 Loi binomiale négative et séries géométriques dérivées d'ordre supérieur

Soit $(X_k)_{k \geq \mathbf{N}^*}$ une suite de variables aléatoires indépendantes, suivant toutes la loi géométrique de paramètre $p \in]0, 1[$.

Pour tout $n \geq 1$, on pose $S_n = \sum_{k=1}^n X_k$.

- 1) Déterminer le support de S_n .
- 2) Montrer par récurrence sur n que pour tout $n \in \mathbf{N}^*$ et tout $k \in S_n(\Omega)$, $P(S_n = k) = \binom{k-1}{n-1} p^n (1-p)^{k-n}$. Savez-vous interpréter ce résultat ?
On dit alors que S_n suit la loi binomiale négative $\mathcal{B}\mathcal{N}(n, p)$.
- 3) En considérant l'espérance de S_n , montrer que $\sum_{k=n}^{+\infty} \binom{k}{n} (1-p)^{k-n} = \frac{1}{p^{n+1}}$.
- 4) En déduire que pour tout $x \in]0, 1[$, $\sum_{k=n}^{+\infty} k(k-1) \dots (k-n+1) x^{k-n} = \frac{n!}{(1-x)^{n+1}}$. Que retrouvez-vous pour $n = 1$ et $n = 2$?
- 5) (★) Soient X et Y deux variables aléatoires indépendantes telles que $X \hookrightarrow \mathcal{B}\mathcal{N}(n_1, p)$ et $Y \hookrightarrow \mathcal{B}\mathcal{N}(n_2, p)$. Quelle est la loi de $X + Y$?

EXERCICE 11.16 Soit $(X_n)_{n \in \mathbf{N}}$ une suite de variables aléatoires indépendantes sur (Ω, \mathcal{A}, P) , suivant toutes une loi uniforme sur $[0, 1]$.

- 1) Pour tout entier $n \in \mathbf{N}$ on pose $Y_n = \min(X_0, \dots, X_n)$. Montrer que Y_n est une variable à densité, et déterminer l'une de ses densités.
- 2) Soit à présent N une variable aléatoire suivant une loi de Poisson de paramètre λ , et indépendante des (X_n) . On pose $Z = Y_N = \min(X_0, \dots, X_N)$, c'est-à-dire

$$\forall \omega \in \Omega, Z(\omega) = \min(X_0(\omega), \dots, X_{N(\omega)}(\omega)).$$

- a) Montrer que Z est une variable aléatoire.
- b) Montrer que Z admet une densité, et déterminer l'une de ses densités.

EXERCICE 11.17 (Oral ESCP 2014)

Soit $(X_n)_{n \geq 1}$ une suite de variables aléatoires indépendantes, identiquement distribuées, définies sur un même espace probabilisé (Ω, \mathcal{A}, P) et suivant la loi exponentielle de paramètre $\lambda > 0$.

Pour tout entier $n \geq 1$, on pose $S_n = \sum_{k=1}^n X_k$ et $U_n = \min_{1 \leq k \leq n} X_k$.

- 1) Soit n un entier naturel non nul.
 - a) Déterminer la loi de U_n .
 - b) Donner une densité de S_n . Montrer que pour tout $x > 0$

$$P(S_n > x) = e^{-\lambda x} \sum_{k=0}^{n-1} \frac{\lambda^k x^k}{k!}.$$

- 2) Soit N une variable aléatoire définie sur (Ω, \mathcal{A}, P) , indépendante des (X_i) et qui suit la loi géométrique de paramètre p , avec $0 < p < 1$. On définit S et U par :

$$\text{pour tout } \omega \in \Omega, S(\omega) = S_{N(\omega)}(\omega) \text{ et } U(\omega) = U_{N(\omega)}(\omega).$$

On admet que S et U sont des variables aléatoires.

- a) Montrer que U est une variable aléatoire à densité et déterminer une densité de U .
- b) Déterminer la loi de S .

EXERCICE 11.18 (QSP ESCP 2014)

Soit $(X_i)_{i \geq 1}$ une suite de variables aléatoires sur un espace probabilisé (Ω, \mathcal{A}, P) , mutuellement indépendantes et suivant toutes la loi géométrique de paramètre p .

Pour $n \in \mathbf{N}^*$, on pose $S_n = \sum_{k=1}^n X_k$.

- 1) Soit $n \in \mathbf{N}^*$. Montrer que la variable aléatoire $\frac{1}{S_n}$ admet une espérance, que l'on notera par la suite m_n .
- 2) Soient $(k, n) \in \mathbf{N} \times \mathbf{N}^*$. Exprimer l'espérance de $\frac{S_k}{S_n}$ en fonction de m_n .

TD

CORRECTION DES EXERCICES DU CHAPITRE 11

SOLUTION DE L'EXERCICE 11.1

Pour tout $x \in \mathbf{R}$, on a

$$[Y > x] = \bigcap_{i=1}^n [X_i > x].$$

Et par indépendance des X_i , on a alors

$$P(Y > x) = P\left(\bigcap_{i=1}^n [X_i > x]\right) = \prod_{i=1}^n P(X_i > x).$$

Comme de plus les X_i suivent des lois exponentielles, on a

$$\forall i \in \llbracket 1, n \rrbracket, P(X_i > x) = \begin{cases} 1 & \text{si } x \leq 0 \\ e^{-\lambda_i x} & \text{si } x > 0 \end{cases}$$

On en déduit donc que

$$F_Y(x) = P(Y \leq x) = 1 - P(Y > x) = 1 - \begin{cases} 1^n & \text{si } x \leq 0 \\ \prod_{i=1}^n e^{-\lambda_i x} & \text{si } x > 0 \end{cases} = \begin{cases} 0 & \text{si } x \leq 0 \\ 1 - e^{-(\lambda_1 + \dots + \lambda_n)x} & \text{si } x > 0 \end{cases}$$

On en déduit donc que $Y \hookrightarrow \mathcal{E}(\lambda_1 + \dots + \lambda_n)$ et donc que $E(Y) = \frac{1}{\lambda_1 + \dots + \lambda_n}$.

Remarque

Le contraire de l'événement $[Y \leq x]$ est $[Y > x]$ et non $[Y \geq x]$. Ces deux événements ont la même probabilité si Y est à densité. Pour l'instant nous ne savons pas si Y est ou non à densité, donc gardons $[Y > x]$.

SOLUTION DE L'EXERCICE 11.2

Rappelons que la fonction de répartition des X_i est donnée par

$$\forall i \in \llbracket 1, n \rrbracket, F_{X_i}(x) = \begin{cases} 0 & \text{si } x \leq 0 \\ \frac{x}{a} & \text{si } 0 < x < a \\ 1 & \text{si } x \geq a \end{cases}$$

Or, pour tout $x \in \mathbf{R}$, on a

$$[Y \leq x] = \bigcap_{i=1}^n [X_i \leq x] \text{ et } [Z > x] = \bigcap_{i=1}^n [X_i > x].$$

Par indépendance des X_i , il vient alors

$$F_Y(x) = \prod_{i=1}^n P(X_i \leq x) = \prod_{i=1}^n F_{X_i}(x) = \begin{cases} 0 & \text{si } x \leq 0 \\ \frac{x^n}{a^n} & \text{si } 0 < x < a \\ 1 & \text{si } x \geq a \end{cases}$$

et de même

$$F_Z(x) = P(Z \leq x) = 1 - P(Z > x) = 1 - \prod_{i=1}^n P(X_i > x) = 1 - \prod_{i=1}^n (1 - F_{X_i}(x)) = \begin{cases} 0 & \text{si } x \leq 0 \\ 1 - \left(1 - \frac{x}{a}\right)^n & \text{si } 0 < x < a \\ 1 & \text{si } x \geq a \end{cases}$$

Puisque les X_i sont à densité, les F_{X_i} sont continues sur \mathbf{R} et de classe \mathcal{C}^1 sauf en un nombre fini de points¹.

Les formules

$$F_Y = \prod_{i=1}^n F_{X_i} \text{ et } F_Z = 1 - \prod_{i=1}^n (1 - F_{X_i})$$

permettent alors d'affirmer² que F_Y et F_Z sont également continues sur \mathbf{R} et \mathcal{C}^1 sauf éventuellement en 0 et en a . Donc Y et Z sont des variables à densité.

Rédaction

On commencera toujours par justifier les égalités d'événements, puis on n'oubliera pas de mentionner l'indépendance, qui permet de passer de la probabilité d'une intersection au produit des probabilités.

¹ En fait, la connaissance de densités de X_i nous permet d'affirmer que les F_{X_i} sont de classe \mathcal{C}^1 sauf en 0 et en a .

² Par opérations sur les fonctions continues/ \mathcal{C}^1 .

Une densité en est alors toute fonction coïncidant avec la dérivée de leur fonction de répartition là où cette dérivée existe. Par exemple, on peut prendre

$$f_Z(x) = \begin{cases} 0 & \text{si } x \leq 0 \\ \frac{n}{a^n} x^{n-1} & \text{si } 0 < x < a \\ 0 & \text{si } x \geq a \end{cases} \text{ et } f_Y(x) = \begin{cases} 0 & \text{si } x < 0 \\ \frac{n}{a} \left(1 - \frac{x}{a}\right)^{n-1} & \text{si } 0 < x < a \\ 0 & \text{si } x \geq a \end{cases}$$

SOLUTION DE L'EXERCICE 11.3

1. Les variables $5X_1$, $5X_2$ et $5X_3$ suivent toutes une loi $\mathcal{E}(1) = \gamma(1)$, et puisqu'elles sont indépendantes, $5(X_1 + X_2 + X_3) \leftrightarrow \gamma(3)$.
2. Une densité de $5(X_1 + X_2 + X_3)$ est

$$f : x \mapsto \begin{cases} 0 & \text{si } x \leq 0 \\ \frac{1}{\Gamma(3)} x^2 e^{-x} & \text{sinon} \end{cases}^3$$

³ Et on sait que

$$\Gamma(3) = 2! = 2.$$

Par transformation affine, une densité de $X_1 + X_2 + X_3 = \frac{1}{5}5(X_1 + X_2 + X_3)$ est donc

$$x \mapsto 5f(5x) = \begin{cases} 0 & \text{si } x \leq 0 \\ \frac{5}{2}(5x)^2 e^{-5x} & \text{sinon} \end{cases} = \begin{cases} 0 & \text{si } x \leq 0 \\ \frac{125}{2} x^2 e^{-5x} & \text{sinon} \end{cases}.$$

SOLUTION DE L'EXERCICE 11.4

Commençons par prouver que $\frac{1}{1+Z}$ admet une espérance. Par le théorème de transfert, on a, sous réserve de convergence,

$$\begin{aligned} E\left(\frac{1}{1+Z}\right) &= \sum_{k=0}^{+\infty} \frac{1}{k+1} P(Z=k) \\ &= \sum_{k=0}^{+\infty} \frac{1}{k+1} e^{-\lambda} \frac{\lambda^k}{k!} \\ &= \frac{e^{-\lambda}}{\lambda} \sum_{k=0}^{+\infty} \frac{\lambda^{k+1}}{(k+1)!} \\ &= \frac{e^{-\lambda}}{\lambda} \sum_{i=1}^{+\infty} \frac{\lambda^i}{i!} \\ &= \frac{e^{-\lambda}}{\lambda} \left(\sum_{i=0}^{+\infty} \frac{\lambda^i}{i!} - 1 \right) \\ &= \frac{e^{-\lambda}}{\lambda} (e^\lambda - 1) \\ &= \frac{1 - e^{-\lambda}}{\lambda}. \end{aligned}$$

Chgt d'indice

$$i = k + 1.$$

On a reconnu une série exponentielle, donc convergente, et donc $E\left(\frac{1}{1+Z}\right)$ existe bien.

D'autre part, $E(X + Y) = E(X) + E(Y) = 2\lambda$.

Mais par le lemme des coalitions, $X + Y$ et $\frac{1}{1+Z}$ sont indépendantes, et donc leur produit admet une espérance et

$$E\left(\frac{X+Y}{1+Z}\right) = E(X+Y)E\left(\frac{1}{1+Z}\right) = 2(1 - e^{-\lambda}).$$

SOLUTION DE L'EXERCICE 11.5

Notons que la condition $|i - j| > 1$ signifie $i \neq j$ (mais il est évident que si $i = j$, alors Y_i et Y_j ne sont pas indépendantes⁴) et i et j non consécutifs.

Commençons par supposer que $|i - j| > 1$. Alors, les variables X_i, X_{i+1}, X_j et X_{j+1} étant mutuellement indépendantes, par le lemme des coalitions, il en est de même de $Y_i = X_i X_{i+1}$ et $Y_j = X_j X_{j+1}$.

⁴ car elles sont égales !

Supposons au contraire que $|i - j| \leq 1$. Comme on a déjà traité le cas $i = j$, on peut⁵ supposer que $j = i + 1$.

Alors $Y_i = X_i X_{i+1}$ et $Y_j = X_{i+1} X_{i+2}$.

Y_i et Y_j suivent des lois de Bernoulli puisqu'elles ne prennent que les valeurs 0 et 1, et de plus, on a

$$P(Y_i = 1) = P([X_i = 1] \cap [X_{i+1} = 1]).$$

Par indépendance de X_i et X_{i+1} , il vient alors

$$P(Y_i = 1) = P(X_i = 1)P(X_{i+1} = 1) = p^2$$

et donc Y_i suit la loi de Bernoulli de paramètre p^2 . On prouverait le même résultat pour Y_j . De plus, $Y_i Y_j = X_i X_{i+1}^2 X_{i+2} = X_i X_{i+1} X_{i+2}$.

Ainsi, $Y_i Y_j$ suit une loi de Bernoulli car elle ne prend que les valeurs 0 et 1, et comme précédemment,

$$P(Y_i Y_j = 1) = P([X_i = 1] \cap [X_{i+1} = 1] \cap [X_{i+2} = 1]) = P(X_i = 1)P(X_{i+1} = 1)P(X_{i+2} = 1) = p^3.$$

Alors, par la formule de Huygens, on a

$$\text{Cov}(Y_i, Y_j) = E(Y_i Y_j) - E(Y_i)E(Y_j) = p^3 - p^4 = p^3(1 - p).$$

Cette covariance étant non nulle, on en déduit que Y_i et Y_j ne sont pas indépendantes.

Ainsi, on a bien prouvé que Y_i et Y_j sont indépendantes si et seulement si $|i - j| > 1$.

SOLUTION DE L'EXERCICE 11.6

Commençons par remarquer que $Z_n(\Omega) = \llbracket 1, N \rrbracket$.

Notons X_i le numéro de la boule obtenue lors du i -ème tirage, de sorte que $X_i \leftrightarrow \mathcal{U}(\llbracket 1, N \rrbracket)$.

De plus, les tirages ayant lieu avec remise, X_1, \dots, X_n sont mutuellement indépendantes.

On a alors $Z_n = \min(X_1, \dots, X_n)$.

Pour tout $k \in \llbracket 1, N \rrbracket$, on a donc $[Z_n \geq k] = \bigcap_{i=1}^n [X_i \geq k]$, et par indépendance des X_i ,

$$P(Z_n \geq k) = \prod_{i=1}^n P(X_i \geq k) = P(X_1 \geq k)^n.$$

$$\text{Or } P(X_1 \geq k) = \sum_{\ell=k}^N P(X_1 = \ell) = \sum_{\ell=k}^N \frac{1}{N} = \frac{N - k + 1}{N}.$$

$$\text{Ainsi, il vient } P(Z_n \geq k) = \left(\frac{N - k + 1}{N} \right)^n.$$

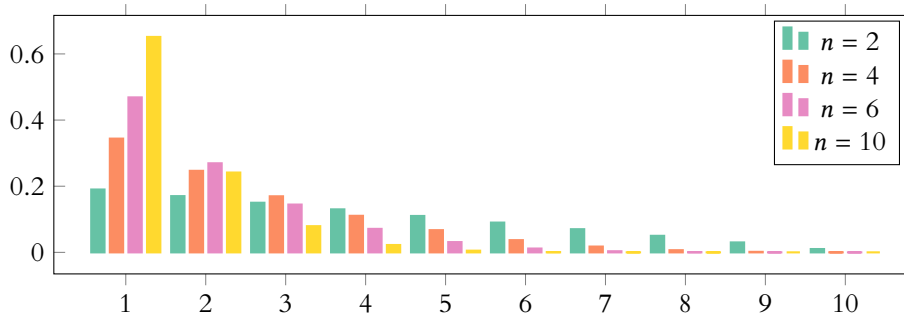
Enfin, puisque Z_n est à valeurs entières, on a $[Z_n \geq k] = [Z_n = k] \cup [Z_n \geq k + 1]$, de sorte que⁶

$$P(Z_n \geq k) = P(Z_n = k) + P(Z_n \geq k + 1) \Leftrightarrow P(Z_n = k) = P(Z_n \geq k) - P(Z_n \geq k + 1).$$

Et donc, pour tout $k \in \llbracket 1, N \rrbracket$,

$$P(Z_n = k) = \left(\frac{N - k + 1}{N} \right)^n - \left(\frac{N - k}{N} \right)^n.$$

Notons que pour $k = N$, on obtient $P(Z_n = N) = \frac{1}{N^n}$, ce qui était prévisible, puisque le seul moyen que la plus petite des boules porte le numéro N est d'avoir tiré n fois la boule N .



⁵ Quitte à échanger i et j .

Carré

Une variable de Bernoulli élevée au carré (ou à n'importe quelle puissance) est égale à elle-même car $0^2 = 0$ et $1^2 = 1$!

Rappel

Pour une loi de Bernoulli, l'espérance est égale au paramètre.

⁶ Les événements $[Z_n = k]$ et $[Z_n \geq k + 1]$ sont incompatibles.

Représentation de la loi de Z_n pour $N = 10$ et quelques valeurs de n . Sans surprise : plus n est grand plus la probabilité de prendre une valeur proche de 1 est grande.

SOLUTION DE L'EXERCICE 11.7

1. Y_i ne prend que les valeurs 0 et 1, et $P(Y_i = 1) = P(X_i = 2) = \frac{2}{3}$.

Donc Y_i suit une loi de Bernoulli $\mathcal{B}(\frac{2}{3})$.

2. Les X_i étant indépendantes, il en est de même des Y_i .

On en déduit que $Y_1 + \dots + Y_n \hookrightarrow \mathcal{B}(n, \frac{2}{3})$.

3. On a $S_n = \sum_{i=1}^n X_i = \sum_{i=1}^n ((X_i - 1) + 1) = n + \sum_{i=1}^n Y_i$.

Par conséquent, S_n prend ses valeurs dans $\llbracket n, 2n \rrbracket$, et pour tout $k \in \llbracket n, 2n \rrbracket$, on a

$$P(S_n = k) = P(Y_1 + \dots + Y_n = k - n) = \binom{n}{k-n} \left(\frac{2}{3}\right)^{k-n} \left(\frac{1}{3}\right)^{n-(k-n)} = \binom{n}{k-n} \frac{2^{k-n}}{3^n}.$$

SOLUTION DE L'EXERCICE 11.8

1. Notons que par définition, on a toujours $X_1 < X_2 < X_3$. Donc si (i, j, k) ne vérifie pas $i < j < k$, alors $P([X_1 = i] \cap [X_2 = j] \cap [X_3 = k]) = 0$.

Au contraire, si $i < j < k$, procédons par dénombrement.

Il y a $\binom{n}{3} = \frac{n(n-1)(n-2)}{6}$ manières de choisir 3 boules parmi n .

Et pour $i < j < k$, il y a un seul tirage vérifiant $[X_1 = i] \cap [X_2 = j] \cap [X_3 = k]$. Donc on a

$$P([X_1 = i] \cap [X_2 = j] \cap [X_3 = k]) = \begin{cases} \frac{6}{n(n-1)(n-2)} & \text{si } 1 \leq i < j < k \leq n \\ 0 & \text{sinon} \end{cases}$$

2. Notons dès à présent qu'on a $X_2(\Omega) = \llbracket 2, n-1 \rrbracket$.

Par la formule des probabilités totales appliquée au système complet d'événements

$$\{[X_1 = i] \cap [X_3 = k], 1 \leq i \leq n-2, 3 \leq k \leq n\}$$

on a, pour tout $j \in \llbracket 2, n-1 \rrbracket$

$$\begin{aligned} P(X_2 = j) &= \sum_{i=1}^{n-2} \sum_{k=j+1}^n P([X_1 = i] \cap [X_2 = j] \cap [X_3 = k]) \\ &= \sum_{i=1}^{j-1} \sum_{k=j+1}^n \frac{6}{n(n-1)(n-2)} = \frac{6}{n(n-1)(n-2)} \sum_{i=1}^{j-1} \sum_{k=j+1}^n 1 \\ &= \frac{6}{n(n-1)(n-2)} \sum_{i=1}^{j-1} (n-j) \\ &= \frac{6(n-j)(j-1)}{n(n-1)(n-2)}. \end{aligned}$$

On a alors

$$\begin{aligned} E(X_2) &= \sum_{j=2}^{n-1} jP(X_2 = j) \\ &= \sum_{j=2}^{n-1} j \frac{(n-j)(j-1)}{n(n-1)(n-2)} \\ &= \frac{6}{n(n-1)(n-2)} \sum_{j=1}^n j(n-j)(j-1) \\ &= \frac{6}{n(n-1)(n-2)} \left(\sum_{j=1}^n (n+1)j^2 - \sum_{j=1}^n j^3 - n \sum_{j=1}^n j \right) \\ &= \frac{6}{n(n-1)(n-2)} \left(\frac{n(n+1)^2(2n+1)}{6} - \frac{n^2(n+1)^2}{4} - \frac{n^2(n+1)}{2} \right) \end{aligned}$$

Évident !

Si l'on connaît le numéro de la plus grande boule, celui de la plus petite et celui de la troisième boule, alors on connaît parfaitement le tirage, il n'y en a qu'un possible.

⚠ Danger !

Au moment de remplacer les probabilités par leur expression, on n'oubliera pas d'enlever de la somme les termes nuls !

Les termes correspondant à $j = 1$ et $j = n$ sont nuls.

$$\begin{aligned}
 &= \frac{6}{n(n-1)(n-2)} n(n+1) \frac{2(n+1)(2n+1) - 3n(n+1) - 6n}{12} \\
 &= \frac{6}{n(n-1)(n-2)} n(n+1) \frac{4n^2 + 6n + 2 - 3n^2 - 3n - 6n}{12} \\
 &= \frac{6}{n(n-1)(n-2)} n(n+1) \frac{(n-2)(n-1)}{12} = \frac{n+1}{2}.
 \end{aligned}$$

SOLUTION DE L'EXERCICE 11.9

1. La fonction f est continue, sauf éventuellement en 0, et elle est positive si et seulement si $\alpha \geq 0$. De plus, pour $A > 0$, on a

$$\int_{-\infty}^A f(t) dt = \int_0^A f(t) dt = \alpha \int_0^A \frac{1}{1+t^2} dt = \alpha \operatorname{Arctan}(A) \xrightarrow{A \rightarrow +\infty} \alpha \frac{\pi}{2}.$$

Donc si X admet f pour densité, nécessairement $\alpha = \frac{2}{\pi}$.

2. Pour $x \geq 0$, nous venons de prouver que

$$F(x) = \int_{-\infty}^x f(t) dt = \alpha \operatorname{Arctan}(x) = \frac{2}{\pi} \operatorname{Arctan}(x).$$

Et pour $x < 0$, alors

$$F(x) = P(X \leq x) = \int_{-\infty}^x f(t) dt = \int_{-\infty}^x 0 dt = 0.$$

Donc la fonction de répartition de X est donnée par

$$F(x) = \begin{cases} 0 & \text{si } x \leq 0 \\ \frac{2}{\pi} \operatorname{Arctan}(x) & \text{si } x > 0 \end{cases}$$

Notons que X n'admet pas d'espérance puisque $tf(t) \underset{t \rightarrow +\infty}{\sim} \frac{2}{\pi} \frac{1}{t}$, et que $\int_1^{+\infty} \frac{dt}{t}$ diverge.

3. Pour $x \in \mathbf{R}$, on a

$$P(Y_2 \leq x) = 1 - P(Y_2 > x) = 1 - P([X_1 > x] \cap [X_2 > x]) = 1 - P(X_1 > x)P(X_2 > x) = 1 - (1 - P(X_1 \leq x))^2.$$

Et donc

$$P(Y_2 \leq x) = \begin{cases} 1 - \left(1 - \frac{2}{\pi} \operatorname{Arctan}(x)\right)^2 & \text{si } x \geq 0 \\ 0 & \text{sinon} \end{cases}$$

La fonction F_{Y_2} est \mathcal{C}^1 sauf éventuellement en 0. De plus, elle est continue en 0 car

$$\lim_{x \rightarrow 0^-} F_{Y_2}(x) = 0 \text{ et que } \lim_{x \rightarrow 0^+} F_{Y_2}(x) = 1 - 1^2 = 0.$$

Donc F_{Y_2} est continue sur \mathbf{R} , et \mathcal{C}^1 sauf éventuellement en 0.

Et donc Y_2 est une variable à densité, dont une densité est donnée par

$$f_2(x) = \begin{cases} 0 & \text{si } x \leq 0 \\ \frac{1}{\pi} \frac{1}{1+x^2} \left(1 - \frac{2}{\pi} \operatorname{Arctan}(x)\right) & \text{si } x > 0 \end{cases}$$

4. Soit g la fonction définie sur $]0, +\infty[$ par $g(x) = \operatorname{Arctan}(x) + \operatorname{Arctan}\left(\frac{1}{x}\right)$. Alors g est dérivable, et on a

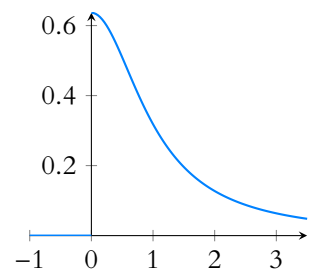
$$g'(x) = \frac{1}{1+x^2} - \frac{1}{x^2} \frac{1}{1+\frac{1}{x^2}} = \frac{1}{1+x^2} - \frac{1}{x^2+1} = 0.$$

Ainsi, g est constante sur $]0, +\infty[$. De plus, on a

$$\lim_{x \rightarrow +\infty} g(x) = \frac{\pi}{2} + \operatorname{Arctan}(0) = \frac{\pi}{2}$$

de sorte que g est la fonction constante, égale à $\frac{\pi}{2}$. Ainsi

$$\forall x > 0, \operatorname{Arctan}(x) + \operatorname{Arctan}\left(\frac{1}{x}\right) = \frac{\pi}{2}.$$

FIGURE 11.1– La densité f **Plus généralement**

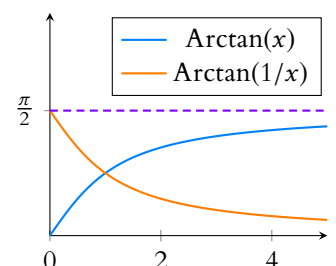
Si f est une fonction positive, continue sauf en un nombre fini de points, et telle que

$$\int_{-\infty}^{+\infty} f(t) dt$$

converge, alors il existe une seule constante k telle que kf soit une densité. Cette constante vérifie

$$\frac{1}{k} = \int_{-\infty}^{+\infty} f(t) dt.$$

C'est la seule constante qui rende l'aire totale sous la courbe de kf égale à 1.



5. D'après la question précédente, on a

$$\forall x > 0, \frac{\pi}{2} - \text{Arctan}(x) = \text{Arctan}\left(\frac{1}{x}\right).$$

Or, pour $x > 0$, on a

$$f_2(x) = \frac{8}{\pi^2} \frac{1}{1+x^2} \left(\frac{\pi}{2} - \text{Arctan}(x) \right) = \frac{8}{\pi^2} \frac{1}{1+x^2} \text{Arctan}\left(\frac{1}{x}\right).$$

Au voisinage de 0, on a $\text{Arctan}(x) \sim x$, et donc au voisinage de $+\infty$, $\text{Arctan}\left(\frac{1}{x}\right) \sim \frac{1}{x}$. Et alors

$$f_2(x) \underset{x \rightarrow +\infty}{\sim} \frac{8}{\pi^2} \frac{1}{x+x^3} \quad \boxed{\underset{x \rightarrow +\infty}{\sim} \frac{8}{\pi^2} \frac{1}{x^3}}.$$

6. La fonction $t \mapsto t f_2(t)$ est continue et positive sur $[0; +\infty[$, et au voisinage de $+\infty$, $t f_2(t) \underset{t \rightarrow +\infty}{\sim} \frac{8}{\pi t^2}$, de sorte que $\int_1^{+\infty} t f_2(t) dt$ converge (par comparaison à une intégrale de Riemann convergente). Et donc elle converge absolument car $t f_2(t)$ est positive. Donc Y_2 admet une espérance. Pour $n \geq 2$, on a

$$0 \leq \min(X_1, \dots, X_n) \leq \min(X_1, X_2) \Leftrightarrow 0 \leq Y_n \leq Y_2.$$

Et puisque Y_2 admet une espérance, alors Y_n admet aussi une espérance.

SOLUTION DE L'EXERCICE 11.10

1. Y ne prend que des valeurs positives, donc pour $x < 0$, on a $F_Y(x) = P(Y \leq x) = 0$. Si $x \geq 0$, alors

$$\begin{aligned} F_Y(x) &= P(Y \leq x) = P(X_1^2 \leq x) = P(-\sqrt{x} \leq X_1 \leq \sqrt{x}) \\ &= F_{X_1}(\sqrt{x}) - F_{X_1}(-\sqrt{x}) = \Phi(\sqrt{x}) - \Phi(-\sqrt{x}). \end{aligned}$$

Mais nous savons que pour tout $t \in \mathbf{R}$, $\Phi(-t) = 1 - \Phi(t)$, de sorte que

$$F_Y(x) = \Phi(\sqrt{x}) - (1 - \Phi(\sqrt{x})) = 2\Phi(\sqrt{x}) - 1.$$

Ainsi, la fonction de répartition de Y est bien donnée par

$$F_Y(x) = \begin{cases} 0 & \text{si } x < 0 \\ 2\Phi(\sqrt{x}) - 1 & \text{sinon} \end{cases}$$

2. F_Y est clairement \mathcal{C}^1 sur $] -\infty, 0[$. De plus, Φ est de classe \mathcal{C}^1 sur \mathbf{R} , donc par composition de fonctions \mathcal{C}^1 , F_Y est de classe \mathcal{C}^1 sur $]0, +\infty[$. Il reste juste à vérifier que F_Y est continue en 0. On a

$$\lim_{x \rightarrow 0^-} F_Y(x) = 0 \text{ et } \lim_{x \rightarrow 0^+} 2\Phi(\sqrt{x}) - 1 = 2\Phi(0) - 1 = 0 = F_Y(0).$$

Donc F_Y est continue en 0, et donc continue sur \mathbf{R} et \mathcal{C}^1 sauf éventuellement en 0. Donc Y admet une densité, et une densité est toute fonction qui coïncide avec F'_Y là où celle-ci est définie. Par exemple, on peut prendre

$$g : x \mapsto \begin{cases} 0 & \text{si } x \leq 0 \\ \frac{1}{\sqrt{x}} \Phi'(\sqrt{x}) & \text{sinon} \end{cases} = \begin{cases} 0 & \text{si } x \leq 0 \\ \frac{1}{\sqrt{x}\sqrt{2\pi}} e^{-x/2} & \text{sinon} \end{cases}$$

3. Par transformation affine, une densité de $\frac{Y}{2}$ est $h : x \mapsto 2g(2x) = \begin{cases} 0 & \text{si } x \leq 0 \\ \frac{1}{\sqrt{\pi}} \frac{1}{\sqrt{x}} e^{-x} & \text{sinon} \end{cases}$

On reconnaît alors une densité qui ressemble fort à $\gamma(1/2)$, sauf pour la constante, qui est

Équivalent

On sait qu'au voisinage de 0, $\text{Arctan}(t) \sim t$. Et puisque $\frac{1}{x} \rightarrow 0$ lorsque $x \rightarrow +\infty$, on a bien l'équivalent indiqué.

Densité

Notons qu'on a $P(a \leq X_1 \leq b) = F_{X_1}(b) - F_{X_1}(a)$ uniquement car X_1 est à densité. Dans le cas contraire, on aurait seulement $P(a < X_1 \leq b) = F_{X_1}(b) - F_{X_1}(a)$.

Remarque

On retrouve bien l'expression de l'énoncé pour $x = 0$ car $\Phi(0) = \frac{1}{2}$.

Dérivée de Φ

On sait que la fonction Φ est dérivable sur \mathbf{R} et que sa dérivée est la densité usuelle de la loi $\mathcal{N}(0, 1)$, notée φ .

$\frac{1}{\sqrt{\pi}}$, alors qu'on aimerait avoir $\Gamma(1/2)$.

Mais puisque h est une densité, on a $\int_{-\infty}^{+\infty} h(t) dt$ soit encore

$$\int_0^{+\infty} \frac{1}{\sqrt{\pi}} t^{1/2-1} e^{-t} dt = 1 \Leftrightarrow \frac{1}{\sqrt{\pi}} \Gamma\left(\frac{1}{2}\right) = 1 \Leftrightarrow \Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}.$$

Et donc $\frac{Y}{2}$ suit la loi $\gamma(1/2)$.

4. Par stabilité des lois gamma, et par indépendance des X_i^2 , $\frac{Y_n}{2}$ suit une loi $\gamma(n/2)$.
Et donc une densité en est

$$f : x \mapsto \begin{cases} 0 & \text{si } x \leq 0 \\ \frac{1}{\Gamma(n/2)} x^{n/2-1} e^{-x} & \text{si } x > 0 \end{cases}$$

Par transformation affine, une densité de Y est

$$x \mapsto \frac{1}{2} f\left(\frac{x}{2}\right) = \begin{cases} 0 & \text{si } x \leq 0 \\ \frac{1}{2^{n/2} \Gamma(n/2)} x^{n/2-1} e^{-x/2} & \text{si } x > 0 \end{cases}$$

SOLUTION DE L'EXERCICE 11.11

1. Pour tout $x \in \mathbf{R}$, on a $[Y > x] = [X_1 > x] \cap [X_2 > x]$.
Et donc, par indépendance de X_1 et X_2 , il vient

$$F_Y(x) = P(Y \leq x) = 1 - P(Y > x) = 1 - P([X_1 > x] \cap [X_2 > x])$$

$$= 1 - P(X_1 > x)P(X_2 > x) = 1 - (1 - F_{X_1}(x))(1 - F_{X_2}(x)) = \begin{cases} 0 & \text{si } x \leq 0 \\ 1 - (1-x)^2 & \text{si } 0 < x < 1 \\ 1 & \text{si } x \geq 1 \end{cases}$$

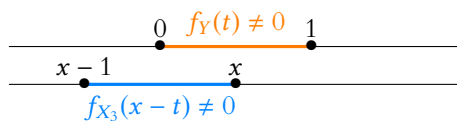
2. Puisque F_{X_1} et F_{X_2} sont continues sur \mathbf{R} et de classe \mathcal{C}^1 sauf éventuellement en un nombre fini de points⁷, il en est de même de F_Y .
Une densité de Y est alors toute fonction coïncidant avec F'_Y là où F_Y est dérivable. Par exemple, on peut prendre

$$f_Y(x) = \begin{cases} 0 & \text{si } x \leq 0 \\ 2(1-x) & \text{si } 0 < x < 1 \\ 0 & \text{si } x \geq 1 \end{cases}$$

3. Par le lemme des coalitions, $Y = \min(X_1, X_2)$ et X_3 sont indépendantes.
Puisque la densité de X_3 est bornée, par produit de convolution, une densité de $Y + X_3$ est alors donnée par

$$h(x) = \int_{-\infty}^{+\infty} f_Y(t) f_{X_3}(x-t) dt.$$

Mais on a $f_Y(t) \neq 0 \Leftrightarrow 0 \leq t \leq 1$ et $f_{X_3}(x-t) \neq 0 \Leftrightarrow 0 \leq x-t \leq 1 \Leftrightarrow x-1 \leq t \leq x$.



Ainsi, pour $x \leq 0$, on a $f_Z(x) = 0$.

De même, pour $x-1 \geq 1 \Leftrightarrow x \geq 2$, on a $f_Z(x) = 0$.

Si $x \in [0, 1]$, il vient

$$f_Z(x) = \int_0^x f_Y(t) dt = \int_0^x 2(1-t) dt = -(1-x)^2 + 1 = 2x - x^2.$$

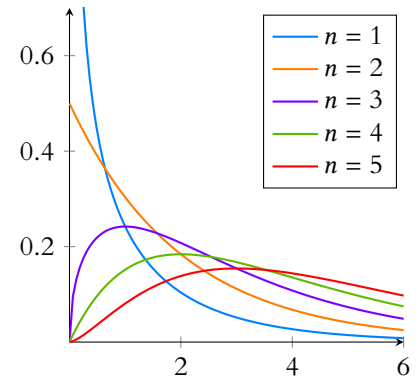


FIGURE 11.2– La densité de la loi du χ^2 pour différentes valeurs de n .

⁷ En l'occurrence, puisqu'il s'agit de lois usuelles, nous savons quels sont ces points : ce sont 0 et 1.

Rappel

Une densité de X_3 est

$$f_{X_3} : x \mapsto \begin{cases} 1 & \text{si } 0 \leq x \leq 1 \\ 0 & \text{sinon} \end{cases}$$

Enfin, pour $1 \leq x \leq 2$, on a

$$f_Z(x) = \int_{x-1}^1 2(1-t) dt = (2-x)^2.$$

Une densité de Z est donc

$$f_Z(x) = \begin{cases} 2x - x^2 & \text{si } 0 \leq x \leq 1 \\ 4 - 4x + x^2 & \text{si } 1 < x \leq 2 \\ 0 & \text{sinon} \end{cases}$$

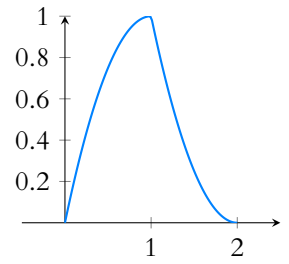


FIGURE 11.3– La densité de Z .

4. Notons que $0 \leq Z \leq 2$, de sorte que l'existence de l'espérance de Z est garantie. On a alors

$$E(Z) = \int_{-\infty}^{+\infty} t f_Z(t) dt = \int_0^1 t(2t - t^2) dt + \int_1^2 t(2-t)^2 dt = \frac{5}{6}.$$

De même, le moment d'ordre 2 de Z est donné⁸ par

$$E(Z^2) = \int_{-\infty}^{+\infty} t^2 f_Z(t) dt = \int_0^1 t^2(2t - t^2) dt + \int_1^2 t^2(2-t)^2 dt = \frac{5}{6}.$$

Et donc, par la formule de Huygens, la variance de V vaut

$$V(Z) = E(Z^2) - E(Z)^2 = \frac{5}{6} - \frac{25}{36} = \frac{5}{36}.$$

⁸ Il existe car $0 \leq Z^2 \leq 4$

SOLUTION DE L'EXERCICE 11.12

1. $Y_n = \frac{1}{2}(X_n + 1)$ ne prend que les valeurs 0 et 1, avec $P(Y_n = 1) = P(X_n = 1) = p$, donc suit une loi de Bernoulli $\mathcal{B}(p)$.
2. On a

$$S_n = \sum_{k=1}^n X_k = \sum_{k=1}^n (2Y_k - 1) = 2 \sum_{k=1}^n Y_k - n$$

et $\sum_{k=1}^n Y_k$ suit une loi binomiale $\mathcal{B}(n, p)$ car les Y_i sont indépendantes.

De plus, S_n prend ses valeurs dans $\llbracket -n, n \rrbracket$, et donc pour tout $k \in \llbracket -n, n \rrbracket$,

$$P(S_n = k) = P\left(2 \sum_{k=1}^n Y_k = k + n\right) = P\left(\sum_{k=1}^n Y_k = \frac{k+n}{2}\right) = \begin{cases} 0 & \text{si } k+n \text{ est impair} \\ \binom{n}{\frac{k+n}{2}} p^{\frac{k+n}{2}} (1-p)^{\frac{2-k-n}{2}} & \text{si } k+n \text{ est pair} \end{cases}$$

On a alors, par linéarité de l'espérance,

$$E(S_n) = \sum_{i=1}^n E(X_i) = \sum_{i=1}^n ((-1)(1-p) + p) = n(2p - 1).$$

- 3.a. Si on note X_i la variable aléatoire qui vaut 1 si l'ivrogne se déplace vers la gauche et -1 s'il se déplace d'un mètre sur la droite, alors S_n désigne sa position à l'instant n . En particulier, on a

$$E(S_{10}) = 2 \times 10 \times p - 10 = 20p - 10.$$

Remarquons que si $p = 0.5$, alors en moyenne il est revenu à son point de départ, que si $p > 0.5$, alors il se sera en moyenne déplacé vers la gauche, et que si $p < 0.5$, alors il se sera en moyenne déplacé vers la droite.

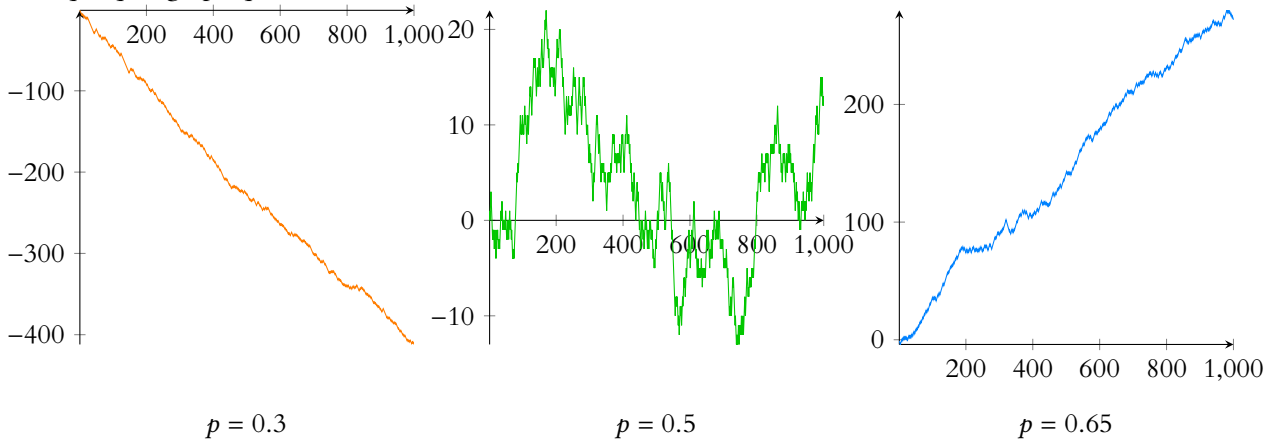
- 3.b. Nous savons que pour simuler une des X_i , il suffit de simuler une loi de Bernoulli de paramètre p , de la multiplier par 2 et de retirer 1.

```
A = 2*grand(1,1000,'bin',1,p)-ones(1,1000);
```

- 3.c. Simulons donc 1000 telles lois, puis traçons $\text{cumsum}(A)$, qui représente successivement $S_1, S_2, \dots, S_{1000}$.

`plot(1 :1000, cumsum(A))`

Voici quelques graphiques obtenus avec ces commandes :



SOLUTION DE L'EXERCICE 11.13

1. La fonction f est continue sur $] -\infty, 0[$ et sur $] 0, +\infty[$. De plus, f est positive, et par le changement de variable $x = \frac{t}{b}$, on a

$$\int_0^{\infty} f(t) dt = \frac{1}{b^v \Gamma(v)} \int_0^{\infty} (bx)^{v-1} e^{-x} b dx = \frac{1}{\Gamma(v)} \int_0^{\infty} x^{v-1} e^{-x} dx = \frac{\Gamma(v)}{\Gamma(v)} = 1.$$

Donc f est bien une densité de probabilité.

2. Par transformation affine, $\frac{1}{b}X$ admet pour densité $t \mapsto bf(bt) = \begin{cases} 0 & \text{si } t \leq 0 \\ \frac{1}{\Gamma(v)} t^{v-1} e^{-t} & \text{si } t > 0 \end{cases}$

On reconnaît une densité d'une loi $\gamma(v)$ et donc $\frac{1}{b}X$ suit la loi $\gamma(v)$.

Si $v = 1$, alors

$$f(t) = \frac{1}{b} e^{-\frac{t}{b}}.$$

On reconnaît la densité de la loi exponentielle de paramètre $\frac{1}{b}$, donc la loi $\Gamma(b, 1)$ est en fait la loi $\mathcal{E}\left(\frac{1}{b}\right)$.

3. Si $b = 1$, on reconnaît la densité usuelle de la loi $\gamma(v)$.
4. X admet un moment d'ordre k si et seulement si $\int_0^{+\infty} t^k \frac{t^{v-1} e^{-\frac{t}{b}}}{b^v \Gamma(v)} dt$ converge.

Or, le changement de variable⁹ $u = \frac{t}{b}$ donne

⁹ Affine.

$$\int_0^{\infty} \frac{1}{b^v \Gamma(v)} t^{k+v-1} e^{-\frac{t}{b}} dt = \frac{1}{b^v \Gamma(v)} \int_0^{+\infty} (bu)^{v+k-1} e^{-u} du = \frac{b^k}{\Gamma(v)} \int_0^{+\infty} u^{v+k-1} e^{-u} du = b^k \frac{\Gamma(v+k)}{\Gamma(v)}.$$

Ainsi, $E(X^k)$ existe et

$$E(X^k) = b^k \frac{\Gamma(v+k)}{\Gamma(v)} = b^k \frac{(v+k-1)\Gamma(v+k-1)}{\Gamma(v)} = \dots = b^k (v+k-1)(v+k-2) \dots v.$$

En particulier, $E(X) = bv$ et

$$V(X) = E(X^2) - E(X)^2 = b^2(v+1)v - b^2v^2 = b^2v.$$

5. Notons que par transformation affine, si X_i suit la loi $\Gamma(b, v_i)$, alors $\frac{X_i}{b}$ suit la loi $\gamma(v_i)$. On en déduit que $\frac{1}{b}(X_1 + X_2) = \frac{X_1}{b} + \frac{X_2}{b}$ suit la loi $\gamma(v_1 + v_2)$. Et donc $X_1 + X_2 = b\left(\frac{X_1}{b} + \frac{X_2}{b}\right)$ suit la loi $\Gamma(b, v_1 + v_2)$.

Convergence

L'égalité entre intégrales est valable sous réserve de convergence, mais puisque la seconde converge, alors automatiquement la première converge aussi.

Détail

On utilise en fait k fois la formule

$$\Gamma(v+1) = v\Gamma(v).$$

6. Par récurrence sur n , on déduit de la question précédente que $X_1 + \dots + X_n$ suit la loi $\Gamma(b, v_1 + \dots + v_n)$.
7. La loi exponentielle de paramètre λ est la loi $\Gamma(\frac{1}{\lambda}, 1)$. Donc d'après la question précédente, les variables aléatoires X_i suivent la loi $\Gamma(\frac{1}{\lambda}, 1)$, et donc

$$X_1 + \dots + X_n \hookrightarrow \Gamma\left(\frac{1}{b}, n\right).$$

8. Commençons par déterminer la loi de $Z = Y^2$.
Si $x \leq 0$, alors $F_Z(x) = P(Y^2 \leq x) = 0$.
Si $x \geq 0$, alors

$$F_Z(x) = P(Y^2 \leq x) = P(-\sqrt{x} \leq Y \leq \sqrt{x}) = \Phi(\sqrt{x}) - \Phi(-\sqrt{x}).$$

Ainsi, F_Z est continue sur $] -\infty, 0[$ et sur $]0, +\infty[$. De plus, étant donné que Φ est continue en 0, on a

$$\lim_{x \rightarrow 0^-} F_Z(x) = 0 = \lim_{x \rightarrow 0^+} F_Z(x) = \Phi(0) - \Phi(0).$$

Par conséquent, F_Z est continue sur \mathbf{R} , et elle est \mathcal{C}^1 sur $] -\infty, 0[$ et sur $]0, +\infty[$, avec

$$\forall x \neq 0, F'_Z(x) = \begin{cases} 0 & \text{si } x < 0 \\ \frac{1}{\sqrt{2\pi}2\sqrt{x}} e^{-\frac{x}{2}} + \frac{1}{2\sqrt{2\pi}\sqrt{x}} e^{-\frac{x}{2}} = \frac{1}{\sqrt{2}\sqrt{\pi}} \frac{1}{\sqrt{x}} e^{-\frac{x}{2}} & \text{si } x > 0 \end{cases}$$

Une densité de Z est donc toute fonction qui coïncide avec F'_Z sur \mathbf{R}^* . Par exemple, on peut prendre :

$$f : x \mapsto \begin{cases} 0 & \text{si } x \leq 0 \\ \frac{1}{\sqrt{2}\sqrt{\pi}} \frac{1}{\sqrt{x}} e^{-\frac{x}{2}} & \text{si } x > 0 \end{cases}$$

Étant donné que $\Gamma(\frac{1}{2}) = \sqrt{\pi}$, on peut encore écrire la densité de Z sous la forme

$$x \mapsto \begin{cases} 0 & \text{si } x \leq 0 \\ \frac{1}{2^{1/2}} \frac{1}{\Gamma(1/2)} x^{\frac{1}{2}-1} e^{-\frac{x}{2}} & \text{si } x > 0 \end{cases}$$

Il s'agit de la densité de la loi $\Gamma(2, \frac{1}{2})$, donc $Y^2 \hookrightarrow \Gamma(2, \frac{1}{2})$.

On pourrait comme l'indique l'énoncé retrouver que $\Gamma(1/2) = \sqrt{\pi}$. En effet, nous avons trouvé une densité de Y^2 , qui doit donc être d'intégrale égale à 1. Or, en procédant au changement de variable $u = \frac{t}{2}$ lors du calcul de $\int_0^\infty \frac{1}{\sqrt{2\pi}\sqrt{x}} e^{-\frac{t}{2}} dt = 1$, on obtient $\frac{\Gamma(1/2)}{\sqrt{\pi}} = 1$, et donc $\Gamma(1/2) = \sqrt{\pi}$.

SOLUTION DE L'EXERCICE 11.14

1. Puisque Z_i est à valeurs dans $]0, a]$, X_i est à valeurs dans $\left[\frac{1}{a}, +\infty\right[$.

De plus, pour $x \in \left[\frac{1}{a}, +\infty\right[$, on a

$$F_{X_i}(x) = P(X_i \leq x) = P\left(Z_i \geq \frac{1}{x}\right) = 1 - P\left(Z_i \leq \frac{1}{x}\right) = 1 - \frac{1}{ax}.$$

Cette fonction de répartition est continue sur $] -\infty, \frac{1}{a}[$ et sur $\left[\frac{1}{a}, +\infty\right[$. De plus, elle est continue en $x = \frac{1}{a}$ car

$$\lim_{x \rightarrow \frac{1}{a}^-} F_{X_i}(x) = 0 = \lim_{x \rightarrow \frac{1}{a}^+} F_{X_i}(x).$$

De plus, F_{X_i} est de classe \mathcal{C}^1 sur \mathbf{R} privé de $\frac{1}{a}$, et donc X_i est une variable à densité, dont une densité est donnée par

$$F'_{X_i}(x) = \begin{cases} 0 & \text{si } x \leq \frac{1}{a} \\ \frac{1}{ax^2} & \text{sinon} \end{cases}$$

2. Puisque les X_i sont indépendantes (car les Z_i le sont), alors

$$\forall x \in \mathbf{R}, P(X > x) = \prod_{i=1}^n P(X_i > x) = \begin{cases} 1 & \text{si } x \leq \frac{1}{a} \\ \frac{1}{a^n x^n} & \text{si } x > \frac{1}{a} \end{cases}$$

On en déduit que

$$\forall x \in \mathbf{R}, F_X(x) = \begin{cases} 0 & \text{si } x \leq \frac{1}{a} \\ 1 - \frac{1}{a^n x^n} & \text{sinon} \end{cases}$$

Cette fonction est continue sur \mathbf{R} et de classe \mathcal{C}^1 sur \mathbf{R} privé de $\frac{1}{a}$, donc X est une variable à densité, et l'une de ses densités est

$$F'_X(x) = \begin{cases} 0 & \text{si } x \leq \frac{1}{a} \\ \frac{n}{a^n x^{n+1}} & \text{si } x > \frac{1}{a} \end{cases}$$

3. Par définition, et sous réserve de convergence¹⁰, on a

$$E(X) = \int_{1/a}^{+\infty} \frac{n}{a^n x^n} dx = \lim_{A \rightarrow +\infty} \left[-\frac{n}{(n-1)a^n} \frac{1}{x^{n-1}} \right]_{1/a}^A = \frac{n}{(n-1)a} \text{ si } n \geq 2.$$

De même, (et toujours sous réserve de convergence¹¹), on a

$$E(X^2) = \int_{1/a}^{+\infty} \frac{n}{a^n x^{n-1}} dx = \left[-\frac{n}{(n-2)a^n} \frac{1}{x^{n-2}} \right]_{1/a}^{\infty} = \frac{n}{(n-2)a^2}$$

et cette fois la convergence a lieu si et seulement si $n \geq 3$.

Alors, pour $n \geq 3$, on a

$$V(X) = E(X^2) - E(X)^2 = \frac{n}{(n-2)a^2} - \frac{n^2}{(n-1)^2 a^2} = \frac{n(n-1)^2 - (n-2)n^2}{(n-1)^2(n-2)a^2} = \frac{n}{(n-1)^2(n-2)a^2}.$$

SOLUTION DE L'EXERCICE 11.15

- Puisque les X_k sont toutes à valeurs plus grandes que 1, S_n prend des valeurs plus grandes que n , et donc $S_n(\Omega) = \llbracket n, +\infty \rrbracket$.
- Montrons par récurrence sur $n \in \mathbf{N}^*$ la propriété

$$\mathcal{P}(n) : \forall k \geq n, P(S_n = k) = \binom{k-1}{n-1} p^n (1-p)^{k-n}.$$

Pour $n = 1$, on a $S_n = X_1$, et donc

$$\forall k \geq 1, P(S_n = k) = P(X_1 = k) = p(1-p)^{k-1} = \binom{k-1}{0} p^1 (1-p)^{k-1}.$$

Supposons donc que $\forall k \geq n, P(S_n = k) = \binom{k-1}{n-1} p^n (1-p)^{k-n}$.

On a alors $S_{n+1} = \sum_{i=1}^n X_i + X_{n+1} = S_n + X_{n+1}$.

Or, par le lemme des coalitions, S_n et X_{n+1} sont indépendantes. On peut donc utiliser le produit de convolution discret : pour tout $k \geq n+1$,

$$\begin{aligned} P(S_{n+1} = k) &= \sum_{i=0}^k P(S_n = i) P(X_{n+1} = k-i) \\ &= \sum_{i=n}^{k-1} P(S_n = i) P(X_{n+1} = k-i) \\ &= \sum_{i=n}^{k-1} \binom{i-1}{n-1} p^n (1-p)^{i-n} p(1-p)^{k-i-1} \end{aligned}$$

Pour $i < n, P(S_n = i) = 0$ et pour $i = k,$

$P(X_{n+1} = k-i) = P(X_{n+1} = 0) = 0.$

$$= p^{n+1}(1-p)^{k-(n+1)} \sum_{i=n}^{k-1} \binom{i-1}{n-1}.$$

Il s'agit donc de prouver que $\sum_{i=n}^{k-1} \binom{i-1}{n-1} = \binom{k-1}{n}$.

Or, nous savons que $\binom{i-1}{n-1} + \binom{i-1}{n} = \binom{i}{n}$ et donc

$$\begin{aligned} \sum_{i=n}^{k-1} \binom{i-1}{n-1} &= \sum_{i=n}^{k-1} \left(\binom{i}{n} - \binom{i-1}{n} \right) \\ &= \sum_{i=n}^{k-1} \binom{i}{n} - \sum_{i=n}^{k-1} \binom{i-1}{n} \\ &= \sum_{i=n}^{k-1} \binom{i}{n} - \sum_{i=n-1}^{k-2} \binom{i}{n} \\ &= \binom{k-1}{n} - \underbrace{\binom{n-1}{n}}_{=0} \\ &= \binom{k-1}{n}. \end{aligned}$$

Somme télescopique.

Ainsi, on a donc bien,

$$\forall k \geq n+1, P(S_{n+1} = k) = p^{n+1}(1-p)^{k-(n+1)} \sum_{i=n}^{k-1} \binom{i-1}{n-1} = \binom{k-1}{n} p^{n+1}(1-p)^{k-(n+1)}.$$

Et donc par le principe de récurrence, pour tout $n \in \mathbf{N}^*$ et tout $k \geq n$,

$$P(S_n = k) = \binom{k-1}{n-1} p^n (1-p)^{k-n}.$$

Ce résultat s'explique de manière intuitive si on voit une loi géométrique de paramètre p comme le nombre d'essais avant l'obtention d'un premier succès lors d'une répétition d'épreuves de Bernoulli indépendantes de probabilité de succès égale à p .

En effet, S_n représente alors le nombre d'essais avant l'obtention du $n^{\text{ème}}$ succès.

S'il faut k essais, il faut donc qu'il y ait eu n succès et $k-n$ échecs lors de ces essais, ce qui explique la provenance du $p^n(1-p)^{k-n}$.

Le $n^{\text{ème}}$ essai est nécessairement un succès, et il reste donc à «choisir» la position des $n-1$ premiers succès parmi les $k-1$ premiers essais. Il y a $\binom{k-1}{n-1}$ manières de le faire.

3. Nous savons, par linéarité de l'espérance, que $E(S_n) = \sum_{i=1}^n E(X_i) = \frac{n}{p}$.

Mais d'autre part, par définition de l'espérance,

$$E(S_n) = \sum_{k=n}^{+\infty} k P(S_n = k) = \sum_{k=n}^{+\infty} k \binom{k-1}{n-1} p^n (1-p)^{k-n}.$$

En notant que $k \binom{k-1}{n-1} = n \binom{k}{n}$, on a donc

$$\sum_{k=n}^{+\infty} n \binom{k}{n} p^n (1-p)^{k-n} = \frac{n}{p} \Leftrightarrow \sum_{k=n}^{+\infty} \binom{k}{n} (1-p)^{k-n} = \frac{1}{p^{n+1}}.$$

4. Commençons par poser $p = 1 - x$. Puisque $x \in]0, 1[$, alors $p \in]0, 1[$, et donc le résultat de la question précédente s'applique :

$$\sum_{k=n}^{+\infty} \binom{k}{n} x^{k-n} = \frac{1}{(1-x)^{n+1}}.$$

Il suffit alors de se souvenir que $\binom{k}{n} = \frac{k!}{n!(k-n)!} = \frac{1}{n!} k(k-1) \cdots (k-n+1)$ et on a alors, après multiplication par $n!$

$$\sum_{k=n}^{+\infty} k(k-1) \cdots (k-n+1)x^{k-n} = \frac{n!}{(1-x)^{n+1}}.$$

Pour $n = 1$, on retrouve la somme d'une série géométrique dérivée : $\sum_{k=1}^{+\infty} kx^{k-1} = \frac{1}{(1-x)^2}$.

Et pour $n = 2$, on retrouve la somme d'une série géométrique dérivée seconde : $\sum_{k=2}^{+\infty} k(k-1)x^{k-2} = \frac{2}{(1-x)^3}$.

Remarque : nous venons de prouver ici une formule pour la somme des séries géométriques dérivées $n^{\text{ème}}$. Notons que $k(k-1) \cdots (k-n+1)x^{k-n}$ est la dérivée $n^{\text{ème}}$ de x^k et que $\frac{n!}{(1-x)^{n+1}}$ est la dérivée $n^{\text{ème}}$ de $\frac{1}{1-x} = \sum_{k=0}^{+\infty} x^k$.

5. Nous avons envie de voir X comme la somme de n_1 géométriques indépendantes, et Y comme la somme de n_2 autres géométriques indépendantes. Malheureusement, ceci n'est pas directement possible, et on voit mal comment définir ces géométriques uniquement à partir de X et Y .

Considérons plutôt $X_1, \dots, X_{n_1}, X_{n_1+1}, \dots, X_{n_1+n_2}$ des variables aléatoires mutuellement indépendantes, suivant toutes la loi géométrique $\mathcal{G}(p)$, et posons alors $S_1 = \sum_{i=1}^{n_1} X_i$ et

$$S_2 = \sum_{i=n_1+1}^{n_1+n_2} X_i.$$

Alors $S_1 \hookrightarrow \mathcal{B.N}(n_1, p)$, $S_2 \hookrightarrow \mathcal{B.N}(n_2, p)$ et $S_1 + S_2 = \sum_{i=1}^{n_1+n_2} X_i \hookrightarrow \mathcal{B.N}(n_1 + n_2, p)$.

De plus, par le lemme des coalitions, S_1 et S_2 sont indépendantes.

Autrement dit, S_1 et S_2 suivent les mêmes lois que X et Y , et sont indépendantes, comme le sont X et Y . Donc les couples (S_1, S_2) et (X, Y) ont même loi conjointe (car dans le cas de variables indépendantes, la connaissance des lois marginales détermine entièrement la loi du couple).

Et donc, la fonction $(x, y) \mapsto x + y$ étant continue sur \mathbf{R}^2 , $S_1 + S_2$ et $X + Y$ ont même loi, de sorte que $X + Y \hookrightarrow \mathcal{B.N}(n_1 + n_2, p)$.

Nous venons donc de prouver un résultat de stabilité pour les lois binomiales négatives.

SOLUTION DE L'EXERCICE 11.16

1. Puisque toutes les X_i sont à support dans $[0, 1]$, il est évident que Y_n est également à support dans $[0, 1]$, et donc

$$F_{Y_n}(x) = 0 \text{ si } x < 0 \text{ et } F_{Y_n}(x) = 1 \text{ si } x \geq 1.$$

Ensuite, pour $x \in [0, 1]$, on a

$$1 - F_{Y_n}(x) = P(Y_n > x) = P\left(\bigcap_{i=0}^n [X_i > x]\right)$$

Mais par indépendance des X_i , il vient alors

$$1 - F_{Y_n}(x) = \prod_{i=0}^n P(X_i > x) = (1-x)^{n+1}.$$

Par conséquent, $F_{Y_n}(x) = \begin{cases} 0 & \text{si } x < 0 \\ 1 - (1-x)^{n+1} & \text{si } 0 \leq x \leq 1 \\ 1 & \text{si } x > 1 \end{cases}$

Il est évident que F_{Y_n} est continue sur $]-\infty, 0[$, sur $]0, 1[$ et sur $]1, +\infty[$.

De plus, on a

$$\lim_{x \rightarrow 0^-} F_{Y_n}(x) = 0 = \lim_{x \rightarrow 0^+} F_{Y_n}(x)$$

Danger !

En règle générale, il n'est pas permis de dériver terme à terme une série. Nous venons de prouver que dans le cas très particulier des séries géométriques, ceci est possible.

Explication

Prenons un exemple très simple : celui d'une variable X suivant la loi binomiale $\mathcal{B}(2, 0.5)$.

Nous avons bien entendu envie de voir X comme le nombre de succès lors de deux répétitions d'expériences de Bernoulli de paramètre 0.5.

Pourtant, il n'est pas possible, uniquement à partir de X (c'est-à-dire uniquement avec les événements $[X = 0]$, $[X = 1]$ et $[X = 2]$), deux variables de Bernoulli indépendantes dont la somme vaut X . C'est le même type de raisonnement ici.

et de même

$$\lim_{x \rightarrow 1^-} F_{Y_n}(x) = 1 = \lim_{x \rightarrow 1^+} F_{Y_n}(x).$$

Ceci prouve bien que F_{Y_n} est continue sur \mathbf{R} . De plus, F_{Y_n} est \mathcal{C}^1 sur chacun des intervalles $] -\infty, 0[$, $]0, 1[$ et $]1, +\infty[$. On a alors

$$F'_{Y_n}(x) = \begin{cases} 0 & \text{si } x < 0 \\ (n+1)(1-x)^n & \text{si } 0 < x < 1 \\ 0 & \text{si } x > 1 \end{cases}$$

On constate alors que F_{Y_n} est de classe \mathcal{C}^1 sur $] -\infty, 0[\cup]0, 1[\cup]1, +\infty[$.

Ceci prouve donc que Y_n est une variable à densité, et l'une de ses densités est donnée par

$$f(x) = \begin{cases} 0 & \text{si } x \leq 0 \text{ ou } x \geq 1 \\ (n+1)(1-x)^n & \text{si } 0 < x < 1 \end{cases}$$

2.a. Il s'agit de prouver que pour tout $x \in \mathbf{R}$, $[Z \leq x] \in \mathcal{A}$.

Pour cela, nous allons prouver que $[Z > x] \in \mathcal{A}$, \mathcal{A} étant stable par passage au complémentaire, on aura donc bien $[Z \leq x] \in \mathcal{A}$.

Or, on a $\Omega = \bigcup_{n=0}^{+\infty} [N = n]$ et donc

$$[Z > x] = [Z > x] \cap \Omega = \bigcup_{n=0}^{+\infty} [Z > x] \cap [N = n].$$

Or, on a, par définition d'un minimum¹²

$$[Z > x] \cap [N = n] = [Y_n > x] \cap [N = n] = [N = n] \cap [\min(X_0, \dots, X_n) > x] = [N = n] \cap \bigcap_{k=0}^n [X_k > x].$$

¹² Et non par indépendance, l'indépendance n'intervient que lorsqu'on considère les probabilités de ces événements.

Puisque les X_i sont des variables aléatoires, pour tout i , par définition, $[X_i \leq x] \in \mathcal{A}$. Et donc par passage au complémentaire, $[X_i > x] \in \mathcal{A}$.

De même, $[N = n] \in \mathcal{A}$, et donc, par intersection¹³ d'éléments de \mathcal{A} , on a

$$[N = n] \cap [Z > x] = [N = n] \cap \bigcap_{i=0}^n [X_i > x] \in \mathcal{A}.$$

¹³ finie

Alors, par union dénombrable d'éléments de \mathcal{A} , on a bien

$$[Z > x] = \bigcup_{n=0}^{+\infty} \underbrace{[Z > x] \cap [N = n]}_{\in \mathcal{A}} \in \mathcal{A}.$$

Et donc Z est bien une variable aléatoire.

2.b. Soit $x \in \mathbf{R}$. D'après la formule des probabilités totales appliquée au système complet d'événements $\{[N = k], k \in \mathbf{N}\}$, on a

$$P(Z \leq x) = \sum_{k=0}^{\infty} P([Z \leq x] \cap [N = k]) = \sum_{k=0}^{\infty} P([Y_k \leq x] \cap [N = k])$$

Mais N étant indépendante de la suite (X_n) , pour tout entier n on a (X_0, \dots, X_n, N) mutuellement indépendantes, et donc par le lemme des coalitions, N et Y_k sont indépendantes pour tout k , de sorte que

$$P(Z \leq x) = \sum_{k=0}^{\infty} P(Y_k \leq x)P(N = k).$$

Si $x \leq 0$, alors $P(Z \leq x) = 0$ par la question précédente.

De même, si $x \geq 1$, alors $P(Z \leq x) = 1$.

Pour $x \in]0, 1[$, il vient

$$F_Z(x) = \sum_{k=0}^{\infty} (1 - (1-x)^{k+1}) e^{-\lambda} \frac{\lambda^k}{k!}$$

Détail

Si $N = k$, alors $Z = Y_k$, par définition.

$$\begin{aligned}
&= \sum_{k=0}^{\infty} e^{-\lambda} \frac{\lambda^k}{k!} - (1-x)e^{-\lambda} \sum_{k=0}^{\infty} \frac{(1-x)^k \lambda^k}{k!} \\
&= 1 - (1-x)e^{-\lambda} e^{(1-x)\lambda} = 1 - (1-x)e^{-\lambda x}.
\end{aligned}$$

Il est facile de vérifier que F_Z est continue sur \mathbf{R} (il suffit en fait de le vérifier en 0 et en 1), et qu'elle est dérivable sur $]-\infty, 0[$, sur $]0, 1[$ et sur $]1; +\infty[$, avec

$$F'_Z(x) = \begin{cases} 0 & \text{si } x < 0 \text{ ou } x > 1 \\ (1 + \lambda - \lambda x)e^{-\lambda x} & \text{si } x \in]0, 1[\end{cases}$$

Cette fonction étant continue sur \mathbf{R} sauf en 0 et en 1, on en déduit que Z admet une densité, et que l'une de ses densités est donnée par

$$f(x) = \begin{cases} 0 & \text{si } x < 0 \text{ ou } x > 1 \\ (1 + \lambda - \lambda x)e^{-\lambda x} & \text{si } x \in [0, 1] \end{cases}$$

SOLUTION DE L'EXERCICE 11.17

1.a. Pour $x \in \mathbf{R}$, on a

$$[U_n > x] = \bigcap_{i=1}^n [X_i > x].$$

Par indépendance des X_i , il vient alors

$$P(U_n > x) = P\left(\bigcap_{i=1}^n [X_i > x]\right) = \prod_{i=1}^n P(X_i > x) = (1 - F_{X_1}(x))^n = \begin{cases} 1 & \text{si } x \leq 0 \\ e^{-n\lambda x} & \text{si } x > 0 \end{cases}$$

On en déduit que

$$F_{U_n}(x) = 1 - P(U_n > x) = \begin{cases} 0 & \text{si } x \leq 0 \\ 1 - e^{-n\lambda x} & \text{si } x > 0 \end{cases}$$

et donc U_n suit la loi exponentielle de paramètre $n\lambda$.

1.b. Les λX_i sont mutuellement indépendantes et suivent toutes la loi $\mathcal{E}(1) = \gamma(1)$, de sorte que $\lambda S_n = \lambda(X_1 + \dots + X_n) \leftrightarrow \gamma(n)$.

Par transformation affine, une densité de $S_n = \frac{1}{\lambda} \lambda S_n$ est donc

$$x \mapsto \begin{cases} 0 & \text{si } x \leq 0 \\ \frac{\lambda^n}{(n-1)!} x^{n-1} e^{-\lambda x} & \text{si } x > 0 \end{cases}$$

En particulier, pour $x > 0$, on a

$$P(S_n > x) = \int_x^{+\infty} \frac{\lambda^n}{(n-1)!} t^{n-1} e^{-\lambda t} dt.$$

Procédons alors à une intégration par parties sur un segment de la forme $[x, A]$, en posant $u = t^{n-1}$ et $v' = e^{-\lambda t}$. Alors

$$\frac{\lambda^n}{(n-1)!} \int_x^A t^{n-1} e^{-\lambda t} dt = \frac{\lambda^{n-1}}{(n-1)!} [-t^{n-1} e^{-\lambda t}]_x^A + \frac{\lambda^{n-1}}{(n-2)!} \int_x^A t^{n-2} e^{-\lambda t} dt.$$

En passant à la limite lorsque $A \rightarrow +\infty$, il vient, par croissance comparée,

$$\int_x^{+\infty} \frac{\lambda^n}{(n-1)!} t^{n-1} e^{-\lambda t} dt = \frac{\lambda^{n-1}}{(n-1)!} x^{n-1} e^{-\lambda x} + \frac{\lambda^{n-1}}{(n-2)!} \int_x^{+\infty} t^{n-2} e^{-\lambda t} dt.$$

En procédant par intégration par parties successives, on prouve que

$$\int_x^{+\infty} \frac{\lambda^n}{(n-1)!} t^{n-1} e^{-\lambda t} dt = e^{-\lambda x} \sum_{k=1}^{n-1} \frac{\lambda^k x^k}{k!} + \lambda \int_x^{+\infty} e^{-\lambda t} dt.$$

Il est alors facile¹⁴ de calculer cette dernière intégrale, qui vaut $\frac{e^{-\lambda x}}{\lambda}$. Ainsi, on a bien

$$P(U_n > x) = e^{-\lambda x} \sum_{k=1}^{n-1} \frac{\lambda^k x^k}{k!} + e^{-\lambda x} = e^{-\lambda x} \sum_{k=0}^{n-1} \frac{\lambda^k x^k}{k!}.$$

Convergence

Notons que la convergence de la seconde intégrale est automatique car celle de la première l'était étant donné qu'il s'agit d'une probabilité.

¹⁴ On connaît une primitive de l'intégrande.

2.a. Notons que U est à valeurs dans \mathbf{R}_+ , de sorte que $F_U(x) = 0$ si $x < 0$. Soit donc $x \in \mathbf{R}_+$. En appliquant la formule des probabilités totales au système complet d'événements $\{[N = k], k \in \mathbf{N}^*\}$, on a

$$P(U \leq x) = \sum_{k=1}^{+\infty} P([U \leq x] \cap [N = k]) = \sum_{k=1}^{+\infty} P([U_k \leq x] \cap [N = k]).$$

Mais N étant indépendante des X_i , par le lemme des coalitions, elle est indépendante de chacune des U_k . Et donc

$$\begin{aligned} P(U \leq x) &= \sum_{k=1}^{+\infty} P(U_k \leq x)P(N = k) \\ &= \sum_{k=1}^{+\infty} (1 - e^{-k\lambda x})p(1-p)^{k-1} \\ &= \sum_{k=1}^{+\infty} p(1-p)^{k-1} - pe^{-\lambda x} \sum_{\ell=0}^{+\infty} e^{-\ell\lambda x}(1-p)^\ell \\ &= 1 - pe^{-\lambda x} \sum_{\ell=0}^{+\infty} (e^{-\lambda x}(1-p))^\ell \\ &= 1 - pe^{-\lambda x} \frac{1}{1 - (1-p)e^{-\lambda x}} \\ &= 1 - \frac{p}{e^{\lambda x} - (1-p)}. \end{aligned}$$

Détail
Si $N = k$, alors $U = U_k$.

Chgt d'indice
 $\ell = k - 1$

Il est alors évident que F_U est \mathcal{C}^1 sur \mathbf{R}_+^* ainsi que sur \mathbf{R}_+^* . De plus, on a

$$\lim_{x \rightarrow 0^+} F_U(x) = 1 - \frac{p}{1 - (1-p)} = 0 = \lim_{x \rightarrow 0^-} F_U(x).$$

Par conséquent, F_U est continue en 0 et donc sur \mathbf{R} , et elle est \mathcal{C}^1 sauf peut-être en 0. Ceci prouve donc que U est une variable à densité. Une densité est alors toute fonction coïncidant avec F'_U sur \mathbf{R}_+^* . On peut par exemple prendre

$$f_U : x \mapsto \begin{cases} 0 & \text{si } x \leq 0 \\ \frac{p\lambda}{(e^{\lambda x} - (1-p))^2} & \text{si } x > 0 \end{cases}$$

2.b. De même, S est à valeurs dans \mathbf{R}_+ , et pour $x \geq 0$, on a

$$P(S > x) = \sum_{k=1}^{+\infty} P(S_k > x)P(N = k) = \sum_{k=1}^{+\infty} e^{-\lambda x} \sum_{n=0}^{k-1} \frac{\lambda^n x^n}{n!} p(1-p)^{k-1}.$$

Bien que la première somme soit infinie, on peut permuter les deux sommes. En effet, la somme telle que nous l'avons écrite converge (c'est une conséquence de la formule des probabilités totales), et comme il s'agit de nombres positifs, le théorème de sommation par paquets affirme que si l'on somme d'abord sur k , puis sur n , la somme ainsi obtenue est également convergente, et égale à la première. Donc

$$P(S > x) = \sum_{n=0}^{+\infty} \sum_{k=n+1}^{+\infty} e^{-\lambda x} \frac{\lambda^n x^n}{n!} p(1-p)^{k-1}.$$

Mais on a

$$\sum_{k=n+1}^{+\infty} p(1-p)^{k-1} = p(1-p)^n \sum_{\ell=0}^{+\infty} (1-p)^\ell = (1-p)^n.$$

Et donc

$$P(S > x) = e^{-\lambda x} \sum_{n=0}^{+\infty} \frac{(1-p)^n \lambda^n x^n}{n!} = e^{-\lambda x} e^{(1-p)\lambda x} = e^{-p\lambda x}.$$

On en déduit que $F_S(x) = \begin{cases} 0 & \text{si } x \leq 0 \\ 1 - e^{-p\lambda x} & \text{si } x > 0 \end{cases}$ et donc S suit la loi exponentielle $\mathcal{E}(p\lambda)$.

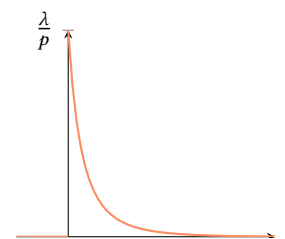


FIGURE 11.4— La densité f_U .

Méthode
Attention aux bornes lorsqu'on permute deux sommes (qu'elles soient finies ou infinies). On sommat pour $0 \leq n \leq k - 1$ donc pour $k \geq n + 1$. Dans la nouvelle somme, une fois établi le fait que n peut prendre toutes les valeurs entières, on somme toujours pour $k \geq n + 1$, c'est-à-dire k allant de $n + 1$ à $+\infty$.

SOLUTION DE L'EXERCICE 11.18

1. Puisque les X_i sont à valeurs dans \mathbf{N}^* , S_n est à valeurs dans $\llbracket n, +\infty \llbracket$.

Et donc $0 \leq \frac{1}{S_n} \leq 1$.

Puisque la variable constante égale à 1 admet une espérance, il en est de même de $\frac{1}{S_n}$.

2. Commençons par le cas trivial¹⁵ : $n = k$. Alors $\frac{S_k}{S_n} = 1$, et donc $E\left(\frac{S_k}{S_n}\right) = 1$.

• Si $k \geq n$, alors

$$\frac{S_k}{S_n} = \frac{X_1 + \dots + X_n + X_{n+1} + \dots + X_k}{S_k} = 1 + \frac{X_{n+1}}{S_n} + \dots + \frac{X_k}{S_n}.$$

Par le lemme des coalitions, X_i est indépendante de S_n , pour $i \geq n + 1$. Et donc X_i est indépendante de $\frac{1}{S_n}$.

On a alors $E\left(\frac{X_i}{S_n}\right) = E(X_i) \times E\left(\frac{1}{S_n}\right) = \frac{m_n}{p}$.

Et donc $E\left(\frac{S_k}{S_n}\right) = 1 + (k - n) \frac{m_n}{p}$.

• Dans le cas où $k < n$, alors $\frac{S_k}{S_n} = \sum_{i=1}^k \frac{X_i}{S_n}$, et les $\frac{X_i}{S_n}$ admettent une espérance car sont dominées par 1.

De plus, par indépendance des X_i , pour tout $i \in \llbracket 1, n \llbracket$, $\frac{X_i}{X_1 + \dots + X_n}$ a même loi que

$\frac{X_1}{X_1 + \dots + X_n}$ et donc même espérance.

Notons $A = E\left(\frac{X_1}{X_1 + \dots + X_n}\right)$. Alors on a

$$1 = E(1) = E\left(\frac{S_n}{S_n}\right) = \sum_{i=1}^n E\left(\frac{X_i}{S_n}\right) = nA.$$

Et donc $A = \frac{1}{n}$, de sorte que $E\left(\frac{S_k}{S_n}\right) = \frac{k}{n}$.

¹⁵ Même s'il n'est pas nécessaire de distinguer ce cas, qui recoupe les deux cas suivants.

Détails

Les X_i étant indépendantes et de même loi, il est évident que les vecteurs (X_1, \dots, X_n) et

$$(X_i, X_2, \dots, X_{i-1}, X_1, X_{i+1}, \dots, X_n)$$

ont même loi.

Et alors $\frac{X_1}{S_n}$ a même loi que

$\frac{X_i}{S_n}$: c'est la proposition 11.5 du cours appliquée à la fonction

$$(x_1, \dots, x_n) \mapsto \frac{x_1}{x_1 + \dots + x_n}.$$

SUPPLÉMENTAIRE ORTHOGONAL. PROJECTEURS ORTHOGONAUX

Cours 373
12.1 Supplémentaire orthogonal. 373
12.2 Projecteurs orthogonaux 376
Méthodes 380
Exercices. 381
Correction des exercices 384

Dans tout le chapitre, E désigne un espace euclidien de dimension n , dont le produit scalaire est noté $\langle \cdot, \cdot \rangle$ et la norme associée $\| \cdot \|$.

12.1 SUPPLÉMENTAIRE ORTHOGONAL

12.1.1 Définition

Définition 12.1 – Si F est un sous-espace vectoriel d'un espace euclidien E , on note

$$F^\perp = \{x \in E : \forall y \in F, \langle x, y \rangle = 0\}$$

Autrement dit, F^\perp est l'ensemble des vecteurs de E qui sont orthogonaux à tous les vecteurs de F .

Danger !

Sur un même espace vectoriel, on peut mettre plusieurs produits scalaires (par exemple, on a déjà rencontré différents produits scalaires sur $\mathbf{R}_n[X]$). Or si on change le produit scalaire, on change F^\perp .

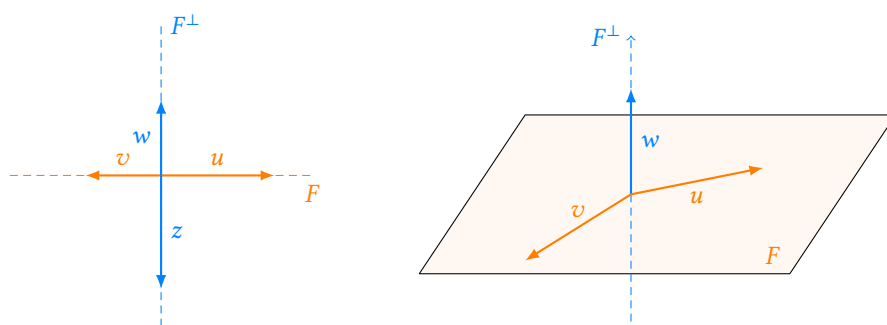


FIGURE 12.1 – Deux exemples dans \mathbf{R}^2 et \mathbf{R}^3 : F^\perp est formé de tous les vecteurs orthogonaux à tous les vecteurs de F .

Proposition 12.2 : Soit F un sous-espace vectoriel de E . Alors :

1. F^\perp est un sous-espace vectoriel de E .
2. si G est un sous-espace vectoriel de E orthogonal¹ à F , alors $G \subset F^\perp$.

¹ **Rappel** : cela signifie que tout vecteur de F est orthogonal à tout vecteur de G .

Autrement dit

F^\perp est le plus grand (au sens de l'inclusion) sous-espace vectoriel de E orthogonal à F .

Démonstration. 1) $F^\perp \subset E$ par définition, et $0_E \in F^\perp$ car pour tout $x \in F$, $\langle 0_E, x \rangle = 0$. Soient $y_1, y_2 \in F^\perp$ et $\lambda \in \mathbf{R}$. Alors, quel que soit $x \in F$, on a

$$\langle \lambda y_1 + y_2, x \rangle = \lambda \langle y_1, x \rangle + \langle y_2, x \rangle = 0.$$

Et donc $\lambda y_1 + y_2$ est orthogonal à tout vecteur de F : $\lambda y_1 + y_2 \in F^\perp$.

Donc F^\perp est stable par combinaison linéaire, et donc est un sous-espace vectoriel de E .

2) Soit G un sous-espace vectoriel de E , orthogonal à F . Soit alors $y \in G$. Pour tout $x \in F$, on a $\langle x, y \rangle = 0$ et donc $y \in F^\perp$. On en déduit que $G \subset F^\perp$. \square

12.1.2 Propriétés de l'orthogonal

Proposition 12.3 : Soit F un sous-espace vectoriel de E , et soit (e_1, \dots, e_p) une base de F . Alors pour $x \in E$, on a

$$x \in F^\perp \Leftrightarrow \forall i \in \llbracket 1, p \rrbracket, \langle x, e_i \rangle = 0.$$

Démonstration. Soit $x \in E$. Si $x \in F^\perp$, alors x est orthogonal à tout vecteur de F , et en particulier à e_1, \dots, e_p .

Inversement, supposons que x soit orthogonal à e_1, \dots, e_p , et soit $y \in F$. Alors il existe des réels $\lambda_1, \dots, \lambda_p$ tels que $y = \sum_{i=1}^p \lambda_i e_i$. Et alors

$$\langle x, y \rangle = \langle x, \sum_{i=1}^p \lambda_i e_i \rangle = \sum_{i=1}^p \lambda_i \langle x, e_i \rangle = 0.$$

Et donc, $\forall y \in F, \langle x, y \rangle = 0$. Ceci prouve que $x \in F^\perp$. \square

Autrement dit

Un vecteur est dans F^\perp si et seulement si il est orthogonal à tous les vecteurs d'une base de F . Notons que la base en question n'a pas besoin d'être orthonormée.

Exemples 12.4

- Dans \mathbf{R}^3 muni du produit scalaire canonique, soit $F = \{(x, y, z) \in \mathbf{R}^3 : x - y + z = 0\}$. Alors on a

$$(x, y, z) \in F \Leftrightarrow (x, y, z) = (x, y, y - x) \Leftrightarrow (x, y, z) = x(1, 0, -1) + y(0, 1, 1) \Leftrightarrow (x, y, z) \in \text{Vect}((1, 0, -1), (0, 1, 1)).$$

Donc une base de F est $(1, 0, -1), (0, 1, 1)$.

Un vecteur (x, y, z) de \mathbf{R}^3 est dans F^\perp si et seulement si

$$\begin{cases} \langle (x, y, z), (1, 0, -1) \rangle = 0 \\ \langle (x, y, z), (0, 1, 1) \rangle = 0 \end{cases} \Leftrightarrow \begin{cases} x - z = 0 \\ y + z = 0 \end{cases} \Leftrightarrow \begin{cases} x = z \\ y = -z \end{cases}$$

Donc $F^\perp = \text{Vect}((1, -1, 1))$.

- Sur $\mathbf{R}_n[X]$, posons $\langle P, Q \rangle = \sum_{i=0}^n P^{(i)}(0)Q^{(i)}(0)$. Alors, il est facile de voir que $\langle \cdot, \cdot \rangle$ est bilinéaire symétrique. De plus, si $P \in \mathbf{R}_n[X]$, alors

$$\langle P, P \rangle = \sum_{i=0}^n P^{(i)}(0)^2 \geq 0.$$

Enfin, $\langle P, P \rangle = 0$ si et seulement si

$$\forall i \in \llbracket 0, 1n \rrbracket, P^{(i)}(0) = 0.$$

Notons alors $P = \sum_{i=0}^n a_i X^i$. Par dérivations successives, on prouve que

$$P^{(0)}(0) = P(0) = a_0, P^{(1)}(0) = a_1, P^{(2)}(0) = a_2 \times 2!, \dots, P^{(n)}(0) = a_n \times n!.$$

Et donc, si $\langle P, P \rangle = 0$, alors $a_0 = a_1 = \dots = a_n = 0$, et donc $P = 0$.

Donc $\langle \cdot, \cdot \rangle$ est un produit scalaire sur $\mathbf{R}_n[X]$.

Alternative

On peut aussi utiliser un argument de racines, et prouver que si $\langle P, P \rangle = 0$, alors

$$P(0) = P'(0) = \dots = P^{(n)}(0).$$

Ainsi, 0 est racine de P d'ordre au moins $n + 1$. Mais P étant de degré au plus n , cela signifie que $P = 0$.

Soit donc $F = \text{Vect}(1 + X + \dots + X^n)$. Alors $P = \sum_{i=0}^n a_i X^i \in F^\perp$ si et seulement si

$$\langle P, 1 + X + \dots + X^n \rangle = 0 \Leftrightarrow \sum_{i=0}^n a_i i!^2 = 0 \Leftrightarrow a_0 = -a_1 - a_2 2^2 - \dots - a_n n!^2.$$

Ainsi, F^\perp possède pour base

$$(X - 1, X^2 - 2!^2, \dots, X^n - n!^2).$$

Lemme 12.5. Soit F un sous-espace vectoriel d'un espace euclidien E . Soit $\mathcal{B}_F = (e_1, \dots, e_p)$ une base orthonormée de F . Alors il est possible de compléter \mathcal{B}_F en une base orthonormée de E .

Démonstration. La famille (e_1, \dots, e_p) est libre et peut donc² être complétée en une base $(e_1, \dots, e_p, e_{p+1}, \dots, e_n)$ de E .

Appliquons alors le procédé d'orthonormalisation de Gram-Schmidt à la famille (e_1, \dots, e_n) , et notons (e'_1, \dots, e'_n) la base orthonormée de E ainsi obtenue.

Prouvons par récurrence sur $i \leq p$ que $e'_i = e_i$. Par construction, on a, $e'_1 = \frac{e_1}{\|e_1\|} = e_1$.

Supposons³ donc que $(e'_1, \dots, e'_i) = (e_1, \dots, e_i)$, avec $i < p$.

Alors par construction, $e'_{i+1} = \frac{1}{\|e'_{i+1}\|} e^*_{i+1}$, avec

$$e^*_{i+1} = e_{i+1} - \sum_{k=1}^i \langle e_{i+1}, e'_k \rangle e'_k = e_{i+1} - \sum_{k=1}^i \underbrace{\langle e_{i+1}, e_k \rangle}_{=0} e_k = e_{i+1}$$

En particulier, $\|e^*_{i+1}\| = \|e_{i+1}\| = 1$. Et donc $e'_{i+1} = e_i$.

Ainsi, par le principe de récurrence, pour tout $i \in \llbracket 1, n \rrbracket$, $e_i = e'_i$.

Et donc $(e'_1, \dots, e'_p, e'_{p+1}, \dots, e'_n) = (e_1, \dots, e_p, e'_{p+1}, \dots, e'_n)$ est une base orthonormée de E , qui complète la base (e_1, \dots, e_p) de F . □

² C'est le théorème de la base incomplète.

³ Notons qu'il s'agit d'une récurrence forte.

Autrement dit

Nous venons de prouver que si on applique Gram-Schmidt à une famille orthonormée, alors l'algorithme laissera la famille inchangée (i.e. renvoie la même famille).

Proposition 12.6 : Soit E un espace euclidien et F un sous-espace vectoriel de E . Alors :

- $\dim F^\perp = \dim E - \dim F$.
- F et F^\perp sont supplémentaires dans E . On dit que F^\perp est le supplémentaire orthogonal de F .
- $(F^\perp)^\perp = F$
- la concaténation d'une base orthonormée (f_1, \dots, f_p) de F et d'une base orthonormée (f_{p+1}, \dots, f_n) de F^\perp est une base orthonormée de E .



Démonstration. • Soit (e_1, \dots, e_p) une base orthonormée de F . Alors nous pouvons la compléter en une base orthonormée de E , (e_1, \dots, e_n) .

Soit $x = \sum_{i=1}^n \lambda_i e_i \in E$. Alors

$$\forall i \in \llbracket 1, n \rrbracket, \lambda_i = \langle x, e_i \rangle.$$

Et donc $x \in F^\perp \Leftrightarrow \forall i \in \llbracket 1, p \rrbracket, \langle x, e_i \rangle = 0 \Leftrightarrow \forall i \in \llbracket 1, p \rrbracket, \lambda_i = 0 \Leftrightarrow x \in \text{Vect}(e_{p+1}, \dots, e_n)$.

On en déduit que $F^\perp = \text{Vect}(e_{p+1}, \dots, e_n)$. En particulier, $\dim F^\perp = n - p = \dim E - \dim F$.

• Par définition, F et F^\perp sont deux sous-espaces vectoriels orthogonaux. Et donc ils sont en somme directe : $F \cap F^\perp = \{0_E\}$.

De plus on a $\dim F^\perp + \dim F = \dim E - \dim F + \dim F = \dim E$.

Et donc F et F^\perp sont supplémentaires dans E .

• Soit $x \in F$. Alors $\forall y \in F^\perp, \langle x, y \rangle = 0$, et donc $x \in (F^\perp)^\perp$. On en déduit que $F \subset (F^\perp)^\perp$. Mais $\dim (F^\perp)^\perp = \dim E - \dim F^\perp = \dim E - \dim E + \dim F = \dim F$.

Alternative

Nous venons de prouver que (e_{p+1}, \dots, e_n) est une base de F^\perp , donc la concaténation d'une base de F et d'une base de F^\perp est une base de E : ils sont donc supplémentaires dans E .

Nécessairement, cela signifie que $(F^\perp)^\perp = F$.

- Nous savons déjà que F et F^\perp étant supplémentaires dans E , la concaténation d'une base de F et d'une base de F^\perp forme une base de E . Il reste à voir qu'elle est orthonormée. Le seul point à prouver est que si $i \in \llbracket 1, p \rrbracket$ et $j \in \llbracket p+1, n \rrbracket$, alors $\langle f_i, f_j \rangle = 0$. Mais alors $f_i \in F$ et $f_j \in F^\perp$, donc $\langle f_i, f_j \rangle = 0$. \square

12.2 PROJECTEURS ORTHOGONAUX

12.2.1 Définition

Définition 12.7 – Si F est un sous-espace vectoriel de E , alors on appelle projection orthogonale sur F et on note p_F la projection sur F parallèlement à F^\perp . Tout vecteur $x \in E$ s'écrit de manière unique $x = x_F + x_{F^\perp}$, avec $x_F \in F$ et $x_{F^\perp} \in F^\perp$, et on a alors

$$p_F(x) = x_F.$$

On dit alors que $p_F(x)$ est le projeté orthogonal de x sur F .

Remarques. • En particulier, on a toujours $x - p_F(x) = (x_F + x_{F^\perp}) - x_F = x_{F^\perp} \in F^\perp$.

• On a donc $E_1(p) = \text{Im}(p) = F$ et $E_0(p) = \text{Ker}(p) = F^\perp$.

• On sait qu'un projecteur est toujours la projection sur $\text{Im } p$ parallèlement à $\text{Ker } p$. Autrement dit, un projecteur p est orthogonal si et seulement si $(\text{Im } p)^\perp = \text{Ker } p$.

Terminologie

En général, pour un projecteur, il faut préciser sur quel espace on projette, mais également parallèlement à quel espace.

Pour un projecteur orthogonal, il suffit de préciser F , la projection étant alors par définition parallèlement à F^\perp .

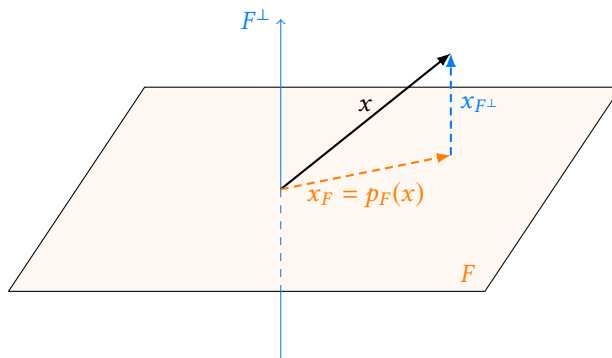


FIGURE 12.2 – Projection orthogonale sur un plan F .

Méthode : calculer un projeté orthogonal

Pour calculer le projeté orthogonal de x sur F , on procède ainsi :

- Déterminer une base (e_1, \dots, e_p) de F , non nécessairement orthonormée.
- Écrire le fait que $p_F(x) \in F$: $p_F(x)$ est combinaison linéaire de e_1, \dots, e_p et donc il existe des réels $\lambda_1, \dots, \lambda_p$ tels que $p_F(x) = \sum_{i=1}^p \lambda_i e_i$.
- Utiliser le fait que $x - p_F(x) \in F^\perp$. En particulier, $\langle x - p_F(x), e_i \rangle = 0, \forall i \in \llbracket 1, p \rrbracket$. Ceci fournit p équations d'inconnues $\lambda_1, \dots, \lambda_p$.
- Résoudre le système pour obtenir $\lambda_1, \dots, \lambda_p$, et donc $p_F(x)$.

Exemple 12.8

Dans $E = \mathbf{R}^3$, soit $F = \left\{ (x, y, z) \in \mathbf{R}^3 : x - \frac{y}{2} + z = 0 \right\}$.

Alors $F = \text{Vect}(e_1, e_2)$, où $e_1 = (1, 2, 0)$ et $e_2 = (1, 0, -1)$.

Soit $x = (2, 2, 2)$. Alors $p_F(x) \in F$ et donc il existe deux réels λ_1 et λ_2 tels que $p_F(x) = \lambda_1 e_1 + \lambda_2 e_2$ et $x - p_F(x) \in F^\perp$.

Or, $x - p_F(x) = (2, 2, 2) - \lambda_1(1, 2, 0) - \lambda_2(1, 0, -1) = (2 - \lambda_1 - \lambda_2, 2 - 2\lambda_1, 2 + \lambda_2)$.
Ainsi

$$\begin{cases} \langle x - p_F(x), e_1 \rangle = 2 - \lambda_1 - \lambda_2 + 2(2 - 2\lambda_1) = 0 \\ \langle x - p_F(x), e_2 \rangle = 2 - \lambda_1 - \lambda_2 - (2 + \lambda_2) = 0 \end{cases} \Leftrightarrow \begin{cases} 5\lambda_1 + \lambda_2 = 6 \\ \lambda_1 + 2\lambda_2 = 0 \end{cases} \Leftrightarrow \begin{cases} \lambda_1 = \frac{4}{3} \\ \lambda_2 = -\frac{2}{3} \end{cases}$$

Donc $p_F(x) = \frac{4}{3}e_1 - \frac{2}{3}e_2 = \frac{1}{3}(2, 8, 2)$.

12.2.2 Propriétés

Proposition 12.9 : Soit (e_1, \dots, e_p) une base orthonormée de F . Alors, pour tout $x \in E$,

$$p_F(x) = \sum_{i=1}^p \langle x, e_i \rangle e_i.$$

Exemple 12.10

Soit $e \in F$ un vecteur non nul, et soit $F = \text{Vect}(e)$ la droite engendrée par e . Alors une base orthonormée de F est $\frac{e}{\|e\|}$, et la projection orthogonale sur F est alors

$$x \mapsto \left\langle x, \frac{e}{\|e\|} \right\rangle \frac{e}{\|e\|} = \frac{\langle x, e \rangle}{\|e\|^2} e.$$

Remarque. En particulier, si e est un vecteur de norme 1, le projeté orthogonal de x sur $\text{Vect}(e)$ est $x \mapsto \langle x, e \rangle e$.

Démonstration. Soit (e_{p+1}, \dots, e_n) une base orthonormée de F^\perp .

Alors $(e_1, \dots, e_p, e_{p+1}, \dots, e_n)$ est une base orthonormée de E , et donc pour tout $x \in E$,

$$x = \sum_{i=1}^n \langle x, e_i \rangle e_i = \underbrace{\sum_{i=1}^p \langle x, e_i \rangle e_i}_{\in F} + \underbrace{\sum_{i=p+1}^n \langle x, e_i \rangle e_i}_{\in F^\perp}.$$

Par définition de p_F , on a alors $p_F(x) = \sum_{i=1}^p \langle x, e_i \rangle e_i$. \square

Maintenant que nous avons cette formule, revenons un instant sur le procédé d'orthonormalisation de Gram-Schmidt.

Soit (e_1, \dots, e_n) une famille libre, et soit (x_1, \dots, x_n) la famille orthonormée obtenue par Gram-Schmidt. Celle-ci était définie par récurrence par

$$\forall k \in \llbracket 2, n \rrbracket, x_k = \frac{x_k^*}{\|x_k^*\|} \text{ où } x_k^* = e_k - \sum_{i=1}^{k-1} \langle e_k, x_i \rangle x_i.$$

Si on note $F_{k-1} = \text{Vect}(x_1, \dots, x_{k-1})$, alors (x_1, \dots, x_{k-1}) est une base orthonormée de F_{k-1}

par construction, et donc $p_{F_{k-1}}(e_k) = \sum_{i=1}^{k-1} \langle e_k, x_i \rangle x_i$.

Autrement dit, on a $x_k^* = e_k - p_{F_{k-1}}(e_k) \in F_{k-1}^\perp$.

Et en particulier, x_k^* est orthogonal à tous les vecteurs x_1, \dots, x_{k-1} précédemment construits (ce que nous avons prouvé par récurrence dans la preuve de Gram-Schmidt donnée dans le chapitre 9).

Remarque

Si on dispose d'une base orthonormée de F , alors il est plus rapide d'utiliser cette formule que la méthode décrite précédemment. En revanche, si ce n'est pas le cas, il sera sûrement plus long de commencer par utiliser Gram-Schmidt afin d'obtenir une base orthonormée.

Proposition 12.11 : Soit \mathcal{B} une base orthonormée de E , et soit (u_1, \dots, u_p) une base orthonormée de F . Soient U_1, \dots, U_p les vecteurs colonnes des coordonnées de u_1, \dots, u_p dans la base \mathcal{B} . Alors la matrice de p_F dans la base \mathcal{B} est donnée par

$$\text{Mat}_{\mathcal{B}}(p_F) = \sum_{i=1}^p U_i {}^t U_i.$$

Démonstration. Notons $F_i = \text{Vect}(u_i)$. Alors $p_F = \sum_{i=1}^p p_{F_i}$.

En effet, on a $\forall x \in E$,

$$p_F(x) = \sum_{i=1}^p \langle u_i, x \rangle u_i = \sum_{i=1}^p p_{F_i}(x).$$

Il suffit donc de déterminer la matrice de p_{F_i} dans la base \mathcal{B} .

Mais si $x \in E$, soit $X \in \mathcal{M}_{n,1}(\mathbf{R})$ le vecteur colonne des coordonnées de x dans la base \mathcal{B} . De même, notons $Y \in \mathcal{M}_{n,1}(\mathbf{R})$ le vecteur colonne des coordonnées de $p_{F_i}(x)$ dans la base \mathcal{B} de sorte que $Y = \langle u_i, x \rangle U_i$.

Alors puisque \mathcal{B} est orthonormée, $\langle u_i, x \rangle = {}^t U_i X \in \mathbf{R}$.

Alors $Y = \underbrace{{}^t U_i X}_{\in \mathbf{R}} U_i = U_i {}^t U_i X$.

Ceci étant vrai pour tout $x \in E$, on en déduit que $\text{Mat}_{\mathcal{B}}(p_{F_i}) = U_i {}^t U_i$. Et alors

$$\text{Mat}_{\mathcal{B}}(p_F) = \text{Mat}_{\mathcal{B}}\left(\sum_{i=1}^p p_{F_i}\right) = \sum_{i=1}^p \text{Mat}_{\mathcal{B}}(p_{F_i}) = \sum_{i=1}^p U_i {}^t U_i.$$

□

Remarque. En particulier, $\text{Mat}_{\mathcal{B}}(p_F)$ est une matrice symétrique car

$${}^t \left(\sum_{i=1}^p U_i {}^t U_i \right) = \sum_{i=1}^p {}^t (U_i {}^t U_i) = \sum_{i=1}^p U_i {}^t U_i.$$

Exemple 12.12

Soit $E = \mathbf{R}^3$ muni de son produit scalaire canonique.

Posons $e_1 = (1, 0, 0)$ et $e_2 = \frac{1}{\sqrt{2}}(0, 1, 1)$, et $F = \text{Vect}(e_1, e_2)$.

Alors (e_1, e_2) est une base orthonormée de F , et donc par la proposition précédente, la matrice de p_F dans la base canonique de \mathbf{R}^3 (qui est une base orthonormée) est

$$\text{Mat}(p_F) = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} (1 \ 0 \ 0) + \frac{1}{2} \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} (0 \ 1 \ 1) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{2} & \frac{1}{2} \\ 0 & \frac{1}{2} & \frac{1}{2} \end{pmatrix}.$$

12.2.3 Distance d'un point à un sous-espace vectoriel

Proposition 12.13 : Soit $x \in E$, et soit F un sous-espace vectoriel de E . Alors l'ensemble $\{\|x - u\|, u \in F\}$ admet un minimum, et ce minimum est atteint en un unique vecteur $v \in F$, défini par $v = p_F(x)$.

Remarque. Le minimum de $\{\|x - u\|, u \in F\}$ est la «distance de x à F ». La proposition précédente affirme que cette distance est atteinte pour $u = p_F(x)$.

Démonstration. D'après le théorème de Pythagore, pour tout $u \in F$, on a

$$\|x - u\|^2 = \underbrace{\|x - p_F(x)\|}_{\in F^\perp}^2 + \underbrace{\|p_F(x) - u\|}_{\in F}^2 = \|x - p_F(x)\|^2 + \|p_F(x) - u\|^2 \geq \|x - p_F(x)\|^2.$$

Hypothèses !

On prendra bien garde au fait qu'il faut **deux** bases orthonormées : une de E et une de F .

Dimensions

En réalité, ${}^t U_i X$ n'est pas un réel, mais une matrice 1×1 . Donc multiplier U_i à gauche par ${}^t U_i X$ n'est pas un produit de matrices qui a du sens, car le nombre de colonnes de ${}^t U_i X$ n'est pas égal au nombre de lignes de U_i .

Donc on entend par là la multiplication par le scalaire $\langle u_i, x \rangle$.

En revanche, multiplier U_i à droite par ${}^t U_i X$ est bien un produit de matrices qui a du sens. Et il revient à multiplier chaque composante de U_i par $\langle u_i, x \rangle$.

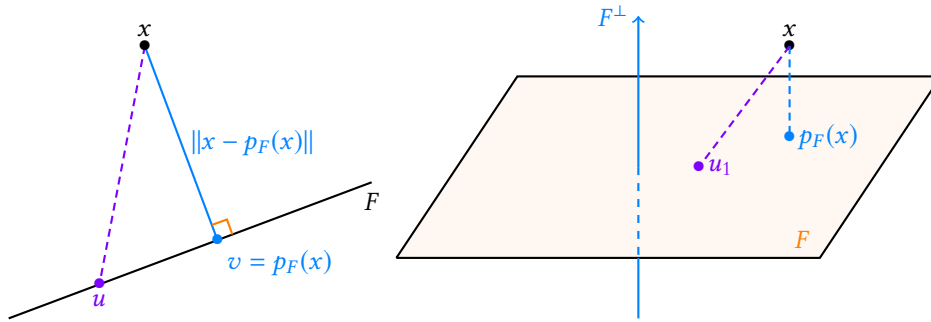


FIGURE 12.3 – En bleu : la distance d'un point x à une droite ou à un plan.

En particulier, pour tout $u \in F$,

$$\|x - p_F(x)\| \leq \|x - u\|.$$

De plus, on a égalité si et seulement si $\|p_F(x) - u\| = 0 \Leftrightarrow p_F(x) = u$.

Donc le minimum de $\{\|x - u\|, u \in F\}$ est $\|x - p_F(x)\|$, atteint uniquement en $u = p_F(x)$. \square

Exemple 12.14

Montrons que $\min_{(x,y) \in \mathbb{R}^2} (x-1)^2 + (x-2y+2)^2 + y^2$ existe, et calculons ce minimum.

On a $(x-1)^2 + (x-2y+2)^2 + y^2 = \|(x-1, x-2y+2, y)\|^2 = \|(x, x-2y, y) - (1, -2, 0)\|^2$.

Soit $F = \{(x, x-2y, y), (x, y) \in \mathbb{R}^2\} \subset \mathbb{R}^3$.

Alors le minimum cherché est le minimum de la distance entre $u = (1, -2, 0)$ et un vecteur de F . Ce minimum existe et est atteint en un unique vecteur $v \in F$, et $v = p_F(u)$.

12.2.4 Théorème des pseudo-solutions

Considérons $\mathcal{M}_{n,1}(\mathbb{R})$ muni de son produit scalaire canonique défini par $\langle X, Y \rangle = {}^tXY$.

Si $X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$, on a alors $\|X\|^2 = \sum_{i=1}^n x_i^2$, et de plus la base canonique de $\mathcal{M}_{n,1}(\mathbb{R})$ est orthonormée pour ce produit scalaire.

Proposition 12.15 : Soit $A \in \mathcal{M}_{n,p}(\mathbb{R})$ de rang p , et soit $B \in \mathcal{M}_{n,1}(\mathbb{R})$. Alors il existe un unique $X_0 \in \mathcal{M}_{p,1}(\mathbb{R})$ minimisant $\{\|AX - B\|, X \in \mathcal{M}_{p,1}(\mathbb{R})\}$.

Démonstration. Soit f l'application de $\mathcal{M}_{p,1}(\mathbb{R})$ dans $\mathcal{M}_{n,1}(\mathbb{R})$ définie par $X \mapsto AX$, et soit $F = \text{Im}f$. Alors l'ensemble $\{\|Y - B\|, Y \in F\}$ possède un minimum atteint en un unique point $Y_0 = p_F(B)$.

De plus, f est injective car $\dim \text{Ker}(f) = \dim \mathcal{M}_{p,1}(\mathbb{R}) - \text{rg}(f) = p - \text{rg}(A) = p - p = 0$. Donc il existe un unique $X_0 \in \mathcal{M}_{p,1}(\mathbb{R})$ tel que $Y_0 = AX_0$. \square

Remarques. • Si A est de rang p , alors nécessairement $n \geq p$.

- A est de rang p signifie que le rang de A est le plus grand possible.
- Il n'existe pas forcément de solution à l'équation $AX = B$, mais ce que nous dit le théorème est que X_0 est «le plus proche possible» d'une solution de $AX = B$: on ne peut rendre $\|AX - B\|$ plus petit.

Rappel

Ce produit scalaire est appelé produit scalaire canonique sur $\mathcal{M}_{n,1}(\mathbb{R})$, et c'est l'analogue du produit scalaire canonique de \mathbb{R}^n si l'on identifie les vecteurs lignes (\mathbb{R}^n) avec les vecteurs colonnes ($\mathcal{M}_{n,1}(\mathbb{R})$).

Remarque : en réalité, la preuve reste valable pour n'importe quel produit scalaire sur $\mathcal{M}_{n,1}(\mathbb{R})$. Le résultat n'est au programme que pour le produit scalaire canonique, et nous verrons en TP l'utilité de ce théorème pour approximer un nuage de points par une droite «la meilleure» possible.

Plus loin

Pour en apprendre plus au sujet des pseudo-solutions d'une équation du type $AX = B$, on pourra consulter le sujet d'ESSEC 2012.

MÉTHODES DU CHAPITRE 12

✓ MÉTHODE 12.1 : Déterminer l'orthogonal d'un sous-espace vectoriel F .

Le point fondamental à retenir est que $x \in F^\perp$ si et seulement si x est orthogonal à tous les éléments d'une base de F . Ainsi, si on a déjà une base de F , on a directement une caractérisation de F^\perp .

Exercices de TD : 12.1, 12.4, 12.8

✓ MÉTHODE 12.2 : Utiliser la supplémentarité de F et F^\perp

On retiendra notamment la formule donnant la dimension de F^\perp en fonction de celle de F , qui sert souvent dans des problèmes de recherche de valeurs propres/sous-espaces propres.

Ceci permet également d'obtenir une base orthonormée de E à partir de deux bases orthonormées de F et de F^\perp (dans le cas où on doit appliquer Gram-Schmidt pour obtenir de telles bases, on préférera donc appliquer deux fois Gram-Schmidt à de «petites» familles (les bases de F et F^\perp) plutôt qu'à une seule «grosse» famille (une base de E)).

Exercices de TD : 12.2, 12.4, 12.8, 12.10, 12.13

✓ MÉTHODE 12.3 : Déterminer la matrice dans une base \mathcal{B} du projecteur orthogonal sur F .

Si la base \mathcal{B} est orthonormée, et si on dispose d'une base orthonormée de F , alors le cours donne une formule.

Sinon, on utilisera toute indication donnée par l'énoncé. Dans le pire des cas, on pourra calculer les images des vecteurs de la base \mathcal{B} , et les décomposer dans la base \mathcal{B} (comme pour trouver la matrice de n'importe quel endomorphisme !)

Exercices de TD : 12.8

✓ MÉTHODE 12.4 : Calculer le projeté orthogonal de x sur F .

Les deux points fondamentaux sont que :

1. $p_F(x) \in F$.
2. $x - p_F(x) \in F^\perp$.

En décomposant $p_F(x)$ dans une base de F et en utilisant la caractérisation de F^\perp rappelée dans la méthode 12.1, on obtient un système de $\dim F$ équations à $\dim F$ inconnues, qui a nécessairement une seule solution. Sa résolution permet de déterminer $p_F(x)$.

Si F^\perp est de «petite» dimension (en pratique, si $\dim F^\perp < \dim F$), on préférera calculer le projeté orthogonal sur F^\perp , puis utiliser le fait que $p_F(x) = x - p_{F^\perp}(x)$.

Exercices de TD : 12.4, 12.7, 12.9, 12.11, 12.12, 12.12, 12.13

✓ MÉTHODE 12.5 : Utiliser une projection orthogonale pour minimiser une quantité.

Des indications, plus ou moins précises, permettent souvent de ramener la recherche d'un minimum à la recherche de $\min\{\|x - u\|, u \in F\}$, où x est un vecteur fixé et F un sous-espace vectoriel.

On sait alors qu'un tel minimum existe, et est uniquement atteint en $u = p_F(x)$. Il s'agit alors de calculer ce projeté orthogonal, éventuellement en utilisant les indications fournies par l'énoncé.

Exercices de TD : 12.4, 12.9, 12.11, 12.12

EXERCICES DU CHAPITRE 12

EXERCICE 12.1 Dans \mathbf{R}^4 muni du produit scalaire canonique, déterminer une base de F^\perp dans les deux cas suivants :

1. $F = \text{Vect}((1, 1, 0, -2), (0, 1, 3, 2))$
2. $F = \{(x, y, z, t) \in \mathbf{R}^4 : x + y - z + t = 0 \text{ et } x - y - t = 0\}$

EXERCICE 12.2 Soient F et G deux sous-espaces supplémentaires d'un espace euclidien E . Montrer que $E = F^\perp \oplus G^\perp$.

EXERCICE 12.3 Soit E un espace euclidien de base orthonormée (e_1, \dots, e_n) , soit F un sous-espace vectoriel de E de dimension 1, et soit p le projecteur orthogonal sur F . Montrer que $\sum_{i=1}^n \|p(e_i)\|^2 = 1$.

EXERCICE 12.4 Vecteur normal à un hyperplan

- 1) Soit F un hyperplan d'un espace euclidien E . Montrer qu'il existe un vecteur $u \in E$ tel que $\forall x \in E, x \in F \Leftrightarrow \langle x, u \rangle = 0$.
- 2) **Application** : dans \mathbf{R}^4 muni du produit scalaire canonique, on pose $F = \{(x, y, z, t) \in \mathbf{R}^4 : 2x + 3y - z + t = 0\}$. Montrer que F est un hyperplan de \mathbf{R}^4 et déterminer un vecteur $u \in \mathbf{R}^4$ tel que $x \in F \Leftrightarrow \langle x, u \rangle = 0$.

EXERCICE 12.5 Sur $\mathcal{M}_n(\mathbf{R})$, on pose $\langle A, B \rangle = \text{tr}({}^tAB)$. Soit p l'application définie sur $\mathcal{M}_n(\mathbf{R})$ par $p(M) = \frac{M + {}^tM}{2}$.

- 1) Montrer que $\langle \cdot, \cdot \rangle$ est un produit scalaire sur $\mathcal{M}_n(\mathbf{R})$.
- 2) Montrer que p est un projecteur de $\mathcal{M}_n(\mathbf{R})$, et déterminer son image et son noyau.
- 3) Montrer que p est le projecteur orthogonal sur $\mathcal{S}_n(\mathbf{R})$, l'ensemble des matrices symétriques de $\mathcal{M}_n(\mathbf{R})$.

- 4) Soit $M = (m_{i,j})_{1 \leq i, j \leq n} \in \mathcal{M}_n(\mathbf{R})$ la matrice définie par $m_{i,j} = \begin{cases} 1 & \text{si } i = 1 \\ 0 & \text{sinon} \end{cases}$.

Calculer $\min_{A \in \mathcal{A}_n(\mathbf{R})} \|A - M\|$. Même question si M est la matrice dont tous les coefficients valent 1.

EXERCICE 12.6 Soit E un espace euclidien et soit p un projecteur orthogonal de E . Montrer que pour tout $x \in E$, on a $\langle p(x), x \rangle = \|p(x)\|^2$.

EXERCICE 12.7 Calcul de projetés orthogonaux

Soit $E = \mathbf{R}_3[X]$ muni du produit scalaire $\langle P, Q \rangle = \int_0^1 P(t)Q(t) dt$.

- 1) Déterminer le projeté orthogonal de $X^2 + X + 1$ sur $F = \mathbf{R}_1[X]$.
- 2) Déterminer le projeté orthogonal de $X^3 + X^2 + X + 1$ sur $F = \text{Vect}(X^3 + X, X^2, 1)$.
- 3) Déterminer le projeté orthogonal de $X^2 - 1$ sur $F = \text{Vect}(1 + X, X^2 - X)$.

EXERCICE 12.8 Matrice d'un projecteur orthogonal

Dans \mathbf{R}^4 muni du produit scalaire canonique on considère le sous-espace vectoriel F défini par

$$F = \{(x, y, z, t) \in \mathbf{R}^4 : x + y - z = 0\}.$$

On note alors p le projecteur orthogonal sur F . Le but de l'exercice est de déterminer la matrice de p dans la base canonique \mathcal{B} par deux méthodes.

- 1) **Première méthode** :
 - a) Déterminer une base de F . En déduire une base orthonormée de F .
 - b) Déterminer $\text{Mat}_{\mathcal{B}}(p)$.
- 2) **Seconde méthode** :
 - a) Déterminer une base de F^\perp . En déduire une base orthonormée de F^\perp .
 - b) On note q la projection orthogonale sur F^\perp . Déterminer la matrice de q dans la base \mathcal{B} .
 - c) Donner une relation liant p et q . En déduire $\text{Mat}_{\mathcal{B}}(p)$.

EXERCICE 12.9 Pour $P = a_0 + a_1X + a_2X^2$ et $Q = b_0 + b_1X + b_2X^2$ dans $\mathbf{R}_2[X]$, on pose $\langle P, Q \rangle = \sum_{i=0}^2 a_i b_i$.

Soit $F = \{P \in \mathbf{R}_2[X] : P(1) = 0\}$.

- 1) Montrer que F est un sous-espace vectoriel de $\mathbf{R}_2[X]$ et en déterminer une base.
- 2) Montrer que $\langle \cdot, \cdot \rangle$ est un produit scalaire sur $\mathbf{R}_2[X]$.
- 3) Déterminer la distance entre le polynôme X^2 et F .

EXERCICE 12.10 Soit E un espace euclidien et soit f un endomorphisme de E tel que $\forall x \in E, \langle f(x), x \rangle = 0$.

- 1) Montrer que $\forall (x, y) \in E^2, \langle f(x), y \rangle = -\langle x, f(y) \rangle$.
- 2) Montrer que $\text{Ker}(f) = (\text{Im}(f))^\perp$.
- 3) Montrer que si λ est valeur propre de f , alors $\lambda = 0$. L'endomorphisme f est-il diagonalisable ?

AD

EXERCICE 12.11 Soit $E = \mathbf{R}_{2n}[X]$. On pose, pour tous $P, Q \in E, \langle P, Q \rangle = \sum_{k=-n}^n P(k)Q(k)$.

- 1) Montrer que $\langle \cdot, \cdot \rangle$ est un produit scalaire sur E .
- 2) Montrer que si P est pair (cela signifie que la fonction polynomiale associée est paire) et que si Q est impair, alors P et Q sont orthogonaux.
- 3) Déterminer, en fonction de n , les valeurs de a et b qui rendent minimale l'expression $\|X^2 - aX + b\|^2$.

AD

EXERCICE 12.12 Le but de cet exercice est de prouver l'existence et de calculer la valeur de

AD

$$\Delta = \inf_{(a,b) \in \mathbf{R}^2} \int_0^1 (e^t - at - b)^2 dt.$$

Sur $E = \mathcal{C}([0, 1], \mathbf{R})$, on définit un produit scalaire par $\forall (f, g) \in E^2, \langle f, g \rangle = \int_0^1 f(t)g(t) dt$.

Soient f_1, f_2 et g les éléments de E définis par $\forall t \in \mathbf{R}, f_1(t) = 1, f_2(t) = t, g(t) = e^t$.

On pose alors $F = \text{Vect}(f_1, f_2)$. Soit $Q = af_2 + bf_1 \in F$.

- 1) Donner sous forme d'intégrale $\|g - Q\|^2$.
- 2) En déduire qu'il existe un unique $Q_0 \in F$ minimisant $\|g - Q_0\|^2$.
- 3) Déterminer sans calcul les valeurs de $\langle g - Q_0, f_1 \rangle$ et $\langle g - Q_0, f_2 \rangle$.
- 4) En déduire la valeur de Δ .

EXERCICE 12.13 (QSP ESCP 2014)

Soit E un espace euclidien de dimension $n \geq 1$, et soit (a, b) une famille orthonormée de E .

Soit f l'application définie sur E par $f(x) = \langle x, a \rangle b - \langle x, b \rangle a$.

- 1) Montrer que $\text{Im}(f) = (\text{Ker}(f))^\perp$.
- 2) L'application f est-elle diagonalisable ?

D

EXERCICE 12.14 Pour $A = (a_{i,j})_{1 \leq i,j \leq n}$ et $B = (b_{i,j})_{1 \leq i,j \leq n} \in \mathcal{M}_n(\mathbf{R})$, on pose $\langle A, B \rangle = \sum_{i=1}^n \sum_{j=1}^n a_{i,j}b_{i,j}$.

D

- 1) Montrer que $\langle \cdot, \cdot \rangle$ est un produit scalaire sur $\mathcal{M}_n(\mathbf{R})$.
- 2) Soit $F = \left\{ M = (m_{i,j})_{1 \leq i,j \leq n} \in \mathcal{M}_n(\mathbf{R}), \sum_{1 \leq i,j \leq n} m_{i,j} = 0 \right\}$. Montrer que F est un sous-espace vectoriel de $\mathcal{M}_n(\mathbf{R})$, et que si J_n désigne la matrice de $\mathcal{M}_n(\mathbf{R})$ dont tous les coefficients valent 1, alors $F^\perp = \text{Vect}(J_n)$.
- 3) Soit $A = (a_{i,j})_{1 \leq i,j \leq n} \in \mathcal{M}_n(\mathbf{R})$. Calculer la valeur de $\delta = \min_{M \in F} \sum_{1 \leq i,j \leq n} (m_{i,j} - a_{i,j})^2$.

EXERCICE 12.15 (Oral ESCP 2016)

Soit $(E, \langle \cdot, \cdot \rangle)$ un espace euclidien muni de la norme $\| \cdot \|$ associée au produit scalaire, et $u \in \mathcal{L}(E)$ vérifiant :

TD

$$\forall x \in E, \|u(x)\| \leq \|x\|.$$

- 1) a) Soit x un vecteur appartenant à $\text{Ker}(u - \text{id}) \cap \text{Im}(u - \text{id})$. Justifier qu'il existe $y \in E$ tel que $\forall n \in \mathbf{N}, nx = u^n(y) - y$.
b) En déduire que $E = \text{Ker}(u - \text{id}) \oplus \text{Im}(u - \text{id})$.

2) On pose $\forall p \in \mathbf{N}, v_p = \frac{1}{p+1} \sum_{k=0}^p u^k$, et on note w le projecteur sur $\text{Ker}(u - \text{id})$ parallèlement à $\text{Im}(u - \text{id})$.

Montrer que pour tout $x \in E$, la suite $(v_p(x))_{p \in \mathbf{N}}$ converge vers $w(x)$, c'est-à-dire que

$$\forall x \in E, \lim_{p \rightarrow +\infty} \|v_p(x) - w(x)\| = 0.$$

- 3) Soit Q un projecteur de E , distinct de l'application nulle.
- Montrer que si $\text{Ker}(Q)$ et $\text{Im}(Q)$ sont orthogonaux, alors $\forall x \in E, \|Q(x)\| \leq \|x\|$.
 - Réciproquement, on suppose que $\forall x \in E, \|Q(x)\| \leq \|x\|$. Soit $x \in \text{Im}(Q)$ et $y \in \text{Ker}(Q)$. En considérant les vecteurs $z = x + \lambda y$ pour $\lambda \in \mathbf{R}$, montrer que $\langle x, y \rangle = 0$.
 - En déduire que projecteur Q non nul est orthogonal si et seulement si $\forall x \in E, \|Q(x)\| \leq \|x\|$.
- 4) En déduire que w est un projecteur orthogonal.

CORRECTION DES EXERCICES DU CHAPITRE 12

SOLUTION DE L'EXERCICE 12.1

1. On sait que $(x, y, z, t) \in F^\perp$ si et seulement si $\begin{cases} \langle (x, y, z, t), e_1 \rangle = 0 \\ \langle (x, y, z, t), e_2 \rangle = 0 \end{cases}$. Ainsi, on a

$$F^\perp = \{(x, y, z, t) \in \mathbf{R}^4 : x + y - 2t = 0 \text{ et } y + 3z + 2t = 0\}.$$

Nous avons donc répondu à la question, mais allons un peu plus loin en déterminant une base de F^\perp : $(x, y, z, t) \in F^\perp$ si et seulement si

$$\begin{cases} x + y - 2t = 0 \\ y + 3z + 2t = 0 \end{cases} \Leftrightarrow \begin{cases} x = 2t - y \\ y = -3z - 2t \end{cases} \Leftrightarrow \begin{cases} x = 3z + 4t \\ y = -2z - 2t \end{cases}$$

Et donc,

$$(x, y, z, t) \in F^\perp \Leftrightarrow (x, y, z, t) = (3z + 4t, -2z - 2t, z, t) \Leftrightarrow (x, y, z, t) \in \text{Vect}((3, -2, 1, 0), (4, -2, 0, 1)).$$

Ainsi, $F^\perp = \text{Vect}((3, -2, 1, 0), (4, -2, 0, 1))$.

2. Commençons par déterminer une base de E . On a

$$(x, y, z, t) \in F \Leftrightarrow \begin{cases} x + y - z + t = 0 \\ x - y - t = 0 \end{cases} \Leftrightarrow \begin{cases} z = 2x \\ t = x - y \end{cases} \Leftrightarrow (x, y, z, t) = (x, y, 2x, x - y) \in \text{Vect}((1, 0, 2, 1), (0, 1, 0, -1)).$$

Cette famille génératrice de F étant formée de deux vecteurs non colinéaires, elle est libre et donc il s'agit d'une base de F .

Ainsi, un vecteur $(x, y, z, t) \in \mathbf{R}^4$ est dans F^\perp si et seulement si

$$\begin{cases} \langle (x, y, z, t), (1, 0, 2, 1) \rangle = 0 \\ \langle (x, y, z, t), (0, 1, 0, -1) \rangle = 0 \end{cases} \Leftrightarrow \begin{cases} x + 2z + t = 0 \\ y - t = 0 \end{cases} \Leftrightarrow \begin{cases} x = -t - 2z \\ y = t \end{cases}$$

On en déduit que

$$(x, y, z, t) \in F^\perp \Leftrightarrow (x, y, z, t) = (-t - 2z, t, z, t) \Leftrightarrow (x, y, z, t) \in \text{Vect}((-2, 0, 1, 0), (-1, 1, 0, 1)).$$

On vérifie aisément que cette famille est libre, et donc que c'est une base de F^\perp .

Notons qu'on obtient alors F^\perp de dimension 2, ce qui est cohérent avec

$$\dim F^\perp = \dim \mathbf{R}^4 - \dim F.$$

SOLUTION DE L'EXERCICE 12.2

Puisque F et G sont supplémentaires, on a $\dim F + \dim G = \dim E$. Et donc

$$\dim F^\perp + \dim G^\perp = \dim E - \dim F + \dim E - \dim G = 2 \dim E - \dim E = \dim E.$$

Soit $x \in F^\perp \cap G^\perp$. Puisque $E = F \oplus G$, x s'écrit de manière unique $x = x_F + x_G$, $x_F \in F$ et $x_G \in G$.

Mais alors

$$\|x\|^2 = \langle x, x \rangle = \langle x_F + x_G, x \rangle = \langle x_F, x \rangle + \langle x_G, x \rangle.$$

Mais puisque $x \in F^\perp$ et $x_F \in F$, il vient $\langle x_F, x \rangle = 0$. Et de même $\langle x_G, x \rangle = 0$.

On en déduit que $\|x\| = 0$ et donc $x = 0_E$.

Ainsi, nous venons de prouver que $F^\perp \cap G^\perp = \{0_E\}$.

Combiné à l'égalité précédemment prouvée sur les dimensions, cela prouve que $E = F^\perp \oplus G^\perp$.

SOLUTION DE L'EXERCICE 12.3

Soit u une base orthonormée de F .

On a alors pour tout $i \in \llbracket 1, n \rrbracket$, $p(e_i) = \langle e_i, u \rangle u$. Donc

$$\sum_{i=1}^n \|p(e_i)\|^2 = \sum_{i=1}^n \langle e_i, u \rangle^2 \underbrace{\|u\|^2}_{=1} = \sum_{i=1}^n \langle e_i, u \rangle^2 = \|u\|^2 = 1.$$

Remarque

On a obtenu ici une base de F^\perp parmi d'autres, il y a, comme d'habitude, une infinité de bases de F^\perp .

Astuce

Si on veut gagner du temps, on peut noter qu'un vecteur est dans F si et seulement si il est orthogonal à $(1, 1 - 1, 1)$ et à $(1, -1, 0, -1)$: ceci se voit directement sur les deux équations définissant F . Autrement dit, en notant G le sous-espace engendré par ces deux vecteurs, $F = G^\perp$. Et donc $F^\perp = (G^\perp)^\perp = G$!

Rappel

Il suffit de prendre n'importe quelle base e de F et de poser $u = \frac{e}{\|e\|}$.

Norme

On reconnaît l'expression de $\|u\|^2$ car (e_1, \dots, e_n) est une base orthonormée de E .

SOLUTION DE L'EXERCICE 12.4

1. Nous savons que $F = (F^\perp)^\perp$.
Or, $\dim F^\perp = \dim E - \dim F = \dim E - (\dim E - 1) = 1$.
Soit donc u une base de F^\perp . Alors¹
Pour $x \in E$, on a donc
- $$x \in F \Leftrightarrow x \in (F^\perp)^\perp \Leftrightarrow \langle x, u \rangle = 0.$$
2. Puisque F est le noyau de la forme linéaire $\varphi : (x, y, z, t) \mapsto 2x + 3y - z + t$.
Et donc étant le noyau d'une forme linéaire, c'est un hyperplan.
D'après la question précédente, nous pouvons prendre pour u une base de F^\perp .
Or, ce F^\perp étant de dimension 1, il suffit de trouver un vecteur non nul de F^\perp : ce sera automatiquement une base de F^\perp .
Mais pour $(x, y, z, t) \in \mathbf{R}^4$, on a

$$(x, y, z, t) \in F \Leftrightarrow 2x + 3y - z + t = 0 \Leftrightarrow \langle (x, y, z, t), (2, 3, -1, 1) \rangle = 0.$$

Et donc $u = (2, 3, -1, 1)$ est dans F^\perp , et donc est une base de F^\perp .

Remarque : notons que dans le cas d'un hyperplan de \mathbf{R}^3 , cette notion a déjà été rencontrée : il s'agit de celle de vecteur normal.

Par exemple, si $F = \{(x, y, z) \in \mathbf{R}^3 : x + y - 2z = 0\}$, alors $u = (1, 1, -2)$ est précisément ce qu'on appelait vecteur normal à F en terminale. Autrement dit, pour se donner un plan² dans l'espace, il suffit de se donner un vecteur normal au plan.

¹ Un vecteur est dans $(F^\perp)^\perp$ si et seulement si il est orthogonal à tous les éléments d'une base de F^\perp .

Mais ici, une telle base ne contient qu'un vecteur : le vecteur u .

² Ici on parle de plan vectoriel, c'est-à-dire de sous-espace vectoriel, donc qui passe nécessairement par $(0, 0, 0)$.

SOLUTION DE L'EXERCICE 12.5

1. Soient $A, B, C \in \mathcal{M}_n(\mathbf{R})$ et $\lambda \in \mathbf{R}$. Alors

$$\langle \lambda A + B, C \rangle = \text{tr}({}^t(\lambda A + B)C) = \text{tr}(\lambda {}^tAC + {}^tBC) = \lambda \text{tr}({}^tAC) + \text{tr}({}^tBC) = \lambda \langle A, C \rangle + \langle B, C \rangle.$$

Donc $\langle \cdot, \cdot \rangle$ est linéaire à gauche.

Soient $A, B \in \mathcal{M}_n(\mathbf{R})$. Alors

$$\langle A, B \rangle = \text{tr}({}^tAB) = \text{tr}({}^t({}^tAB)) = \text{tr}({}^tBA) = \langle B, A \rangle.$$

Donc $\langle \cdot, \cdot \rangle$ est symétrique, donc bilinéaire symétrique.

Soit $A = (a_{i,j})_{1 \leq i, j \leq n} \in \mathcal{M}_n(\mathbf{R})$. Alors

$$\langle A, A \rangle = \text{tr}({}^tAA) = \sum_{i=1}^n ({}^tAA)_{i,i} = \sum_{i=1}^n \sum_{j=1}^n ({}^tA)_{i,j} A_{j,i} = \sum_{i=1}^n \sum_{j=1}^n a_{j,i}^2 \geq 0.$$

Et une somme de carrés étant nulle si et seulement si chacun de ces nombres est nul, il vient alors

$$\langle A, A \rangle = 0 \Leftrightarrow \forall (i, j) \in \llbracket 1, n \rrbracket^2, a_{j,i}^2 = 0 \Leftrightarrow \forall (i, j) \in \llbracket 1, n \rrbracket^2, a_{j,i} = 0 \Leftrightarrow A = 0.$$

Et donc $\langle \cdot, \cdot \rangle$ est bien un produit scalaire sur $\mathcal{M}_n(\mathbf{R})$.

2. Il est évident que p est un endomorphisme de $\mathcal{M}_n(\mathbf{R})$. Et on a alors, pour $M \in \mathcal{M}_n(\mathbf{R})$,

$$p^2(M) = p\left(\frac{M + {}^tM}{2}\right) = \frac{1}{2}\left(\frac{M + {}^tM}{2} + {}^t\left(\frac{M + {}^tM}{2}\right)\right) = \frac{1}{2}\left(\frac{M + {}^tM + {}^tM + M}{2}\right) = \frac{M + {}^tM}{2} = p(M).$$

Ceci étant vrai pour tout $M \in \mathcal{M}_n(\mathbf{R})$, il vient $p^2 = p$, et donc p est un projecteur.

On a alors $M \in \text{Ker}(p) \Leftrightarrow \frac{M + {}^tM}{2} = 0 \Leftrightarrow {}^tM = -M$.

Et donc $\text{Ker}(p)$ est l'ensemble $\mathcal{A}_n(\mathbf{R})$ des matrices antisymétriques.

D'autre part, nous savons que pour tout projecteur, $\text{Im}(p) = E_1(p)$.

Et donc $M \in \text{Im}(p) \Leftrightarrow p(M) = M \Leftrightarrow \frac{M + {}^tM}{2} = M \Leftrightarrow M = {}^tM$.

Et donc $\text{Im}(p)$ est l'ensemble $\mathcal{S}_n(\mathbf{R})$ des matrices symétriques.

3. Il s'agit de prouver que $\mathcal{A}_n(\mathbf{R}) = \mathcal{S}_n(\mathbf{R})^\perp$.

Or, puisque nous savons déjà que $\mathcal{M}_n(\mathbf{R}) = \mathcal{S}_n(\mathbf{R}) \oplus \mathcal{A}_n(\mathbf{R})$, alors

$$\dim \mathcal{S}_n(\mathbf{R})^\perp = \dim \mathcal{M}_n(\mathbf{R}) - \dim \mathcal{S}_n(\mathbf{R}) = \dim \mathcal{A}_n(\mathbf{R}).$$

Remarque

Ceci prouve donc que $\mathcal{S}_n(\mathbf{R})$ et $\mathcal{A}_n(\mathbf{R})$ sont supplémentaires dans $\mathcal{M}_n(\mathbf{R})$, puisque l'image et le noyau d'un projecteur sont toujours supplémentaires.

D'autre part, pour $A \in \mathcal{A}_n(\mathbf{R})$ et $S \in \mathcal{S}_n(\mathbf{R})$, on a

$$\langle A, S \rangle = \text{tr}({}^t AS) = -\text{tr}(AS) = -\text{tr}(SA) = -\text{tr}({}^S A) = -\langle S, A \rangle = -\langle A, S \rangle.$$

On en déduit que $\langle A, S \rangle = 0$. Et donc toute matrice de $\mathcal{A}_n(\mathbf{R})$ est orthogonale à toute matrice de $\mathcal{S}_n(\mathbf{R})$: $\mathcal{A}_n(\mathbf{R}) \subset \mathcal{S}_n(\mathbf{R})^\perp$.

Ces deux espaces étant de même dimension : ils sont égaux, et donc p est la projection orthogonale sur $\mathcal{S}_n(\mathbf{R})$.

4. Nous savons que le minimum cherché existe, et est égal à la distance entre M et $\mathcal{A}_n(\mathbf{R})$. Ce minimum est $\|q(M) - M\|$, où q est le projecteur orthogonal sur $\mathcal{A}_n(\mathbf{R})$.

Mais $q = \text{id} - p$. Et donc $q(M) - M = -p(M) = -\frac{M + {}^t M}{2}$.

On en déduit que le minimum cherché est $\left\| \frac{M + {}^t M}{2} \right\| = \frac{1}{2} \|M + {}^t M\|$.

$$\text{Or, } M + {}^t M = \begin{pmatrix} 2 & 1 & \dots & 1 \\ 1 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 0 & \dots & 0 \end{pmatrix}, \text{ de sorte que}$$

$$\|M + {}^t M\|^2 = 4 + 2(n-1) = 2n + 2.$$

Et donc le minimum cherché est $\sqrt{\frac{n+1}{2}}$.

Si M est la matrice dont tous les coefficients valent 1, alors $M \in \mathcal{S}_n(\mathbf{R}) = \text{Ker}(q)$, et donc $q(M) = 0$.

On en déduit que le minimum cherché est $\|M\| = \sqrt{n^2} = n$.

SOLUTION DE L'EXERCICE 12.6

Soit $x \in E$. De manière unique, $x = x_F + x_{F^\perp}$, avec $x_F \in F$ et $x_{F^\perp} \in F^\perp$.

On a alors, par définition de p , $p(x) = x_F$. Il vient donc

$$\langle p(x), x \rangle = \langle x_F, x_F + x_{F^\perp} \rangle = \langle x_F, x_F \rangle + \underbrace{\langle x_F, x_{F^\perp} \rangle}_{=0} = \langle x_F, x_F \rangle = \|x_F\|^2 = \|p(x)\|^2.$$

SOLUTION DE L'EXERCICE 12.7

1. Une base de $\mathbf{R}_1[X]$ est $(1, X)$.

Le projeté orthogonal de $P = X^2 + X + 1$ sur $F = \mathbf{R}_1[X]$ s'écrit $aX + b$, et on doit avoir $X^2 + X + 1 - p_F(P) \in F^\perp$ donc

$$\langle X^2 + X + 1 - aX - b, 1 \rangle = \int_0^1 (t^2 + (1-a)t + (1-b))dt = \frac{1}{3} + \frac{1-a}{2} + 1 - b = 0$$

De même

$$\langle X^2 + X + 1 - p_F(P), X \rangle = \frac{1}{4} + \frac{1-a}{3} + \frac{1-b}{2} = 0$$

et donc a et b sont solutions du système

$$\begin{cases} 2 + 3(1-a) + 6(1-b) = 0 \\ 3 + 4(1-a) + 6(1-b) = 0 \end{cases} \Leftrightarrow \begin{cases} 3a + 6b = 11 \\ 4a + 6b = 13 \end{cases} \Leftrightarrow \begin{cases} a = 2 \\ b = \frac{5}{6} \end{cases}$$

et donc

$$p_F(P) = 2X + \frac{5}{6}.$$

2. Notons que $X^3 + X^2 + X + 1 = (X^3 + X) + X^2 + 1 \in F$, et donc $p_F(X^3 + X^2 + X + 1) = X^3 + X^2 + X + 1$.
3. Notons $P = X^2 - 1$, et soit $p_F(P) = a(1+X) + b(X^2 - X)$. Alors comme précédemment, on obtient le système

$$\begin{cases} 4(1-b) - 6(a-b) - 12(1+a) + 3(1-b) - 4(a-b) - 6(1+a) = 0 \\ 12(1-b) - 15(a-b) - 20(1+a) - 15(1-b) + 20(a-b) + 30(1+a) = 0 \end{cases} \Leftrightarrow \begin{cases} -28a + 3b = 11 \\ 15a - 2b = -7 \end{cases} \Leftrightarrow \begin{cases} a = -\frac{1}{11} \\ b = \frac{31}{11} \end{cases}$$

On en déduit que $p_F(x) = -\frac{20}{11}X^2 + \frac{32}{11}x - \frac{10}{11}$.

Astuce

La formule prouvée dans la question 1 pour $\langle A, A \rangle$ permet d'obtenir facilement $\|A\|^2$ en fonction des coefficients de A : c'est la somme des carrés des coefficients.

Astuce

Comme pour tous les projecteurs, si $x \in \text{Im } p$, alors $p(x) = x$.

SOLUTION DE L'EXERCICE 12.8

- 1.a. On $(x, y, z, t) \in F \Leftrightarrow (x, y, z, t) = (x, y, x + y, t)$, donc une base de F est

$$(1, 0, 1, 0), (0, 1, 1, 0), (0, 0, 0, 1).$$

Une base orthonormée de F est alors obtenue par le procédé d'orthonormalisation de Gram-Schmidt.

$$\text{Posons } e_1 = \frac{(1, 0, 1, 0)}{\|(1, 0, 1, 0)\|} = \frac{1}{\sqrt{2}}(1, 0, 1, 0).$$

$$\text{Soit à présent } e_2^* = (0, 1, 1, 0) - \frac{1}{2}\langle(1, 0, 1, 0), (0, 1, 1, 0)\rangle(1, 0, 1, 0) = \left(-\frac{1}{2}, 1, \frac{1}{2}, 0\right).$$

$$\text{Alors } \|e_2^*\| = \sqrt{\frac{3}{2}}. \text{ Posons alors } e_2 = \sqrt{\frac{2}{3}}\left(-\frac{1}{2}, 1, \frac{1}{2}, 0\right).$$

Puisque le troisième vecteur de la base est orthogonal à e_1 et à e_2 , on peut directement³ poser $e_3^* = (0, 0, 0, 1)$. Ce vecteur est alors unitaire, et donc $e_3 = (0, 0, 0, 1)$.

Ainsi, (e_1, e_2, e_3) est une base orthonormée de F .

- 1.b. Une formule du cours nous donne directement la matrice de p dans la base \mathcal{B} . Notons que cette formule est valable car \mathcal{B} est une base orthonormée de \mathbf{R}^4 .

$$\text{Mat}_{\mathcal{B}}(p) = \frac{1}{2} \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix} (1 \ 0 \ 1 \ 0) + \frac{2}{3} \begin{pmatrix} -\frac{1}{2} \\ 1 \\ \frac{1}{2} \\ 0 \end{pmatrix} \left(-\frac{1}{2} \ 1 \ \frac{1}{2} \ 0\right) + \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} (0 \ 0 \ 0 \ 1) = \frac{1}{3} \begin{pmatrix} 2 & -1 & 1 & 0 \\ -1 & 2 & 1 & 0 \\ 1 & 1 & 2 & 0 \\ 0 & 0 & 0 & 3 \end{pmatrix}.$$

- 2.a. Notons que F est un hyperplan de \mathbf{R}^4 , car nous avons prouvé à la question 1 qu'il est de dimension 3, mais également car on remarque directement qu'il s'agit du noyau de la forme linéaire non nulle $(x, y, z, t) \mapsto x + y - z$.

Donc son orthogonal est de dimension $4 - 3 = 1$, et donc engendré par un seul vecteur.

Or, si $(x, y, z, t) \in F$, alors $\langle(x, y, z, t), (1, 1, -1, 0)\rangle = x + y - z = 0$.

Ainsi, $(1, 1, -1, 0) \in F^\perp$, et, grâce à ce qui a été dit sur la dimension, $F^\perp = \text{Vect}((1, 1, -1, 0))$.

- 2.b. La base \mathcal{B} est orthonormée, et une base orthonormée de F^\perp est formée du seul vecteur $\frac{1}{\sqrt{3}}(1, 1, -1, 0)$. Donc la matrice de q dans la base \mathcal{B} est donnée par

$$\text{Mat}_{\mathcal{B}}(q) = \frac{1}{3} \begin{pmatrix} 1 \\ 1 \\ -1 \\ 0 \end{pmatrix} (1 \ 1 \ -1 \ 0) = \frac{1}{3} \begin{pmatrix} 1 & 1 & -1 & 0 \\ 1 & 1 & -1 & 0 \\ -1 & -1 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

- 2.c. Si $x = x_F + x_{F^\perp} \in \mathbf{R}^4$, alors $p(x) + q(x) = x_F + x_{F^\perp} = x$ et donc $p + q = \text{id}_E$. On en déduit que

$$\text{Mat}_{\mathcal{B}}(p) = I_4 - \text{Mat}_{\mathcal{B}}(q) = I_4 - \frac{1}{3} \begin{pmatrix} 1 \\ 1 \\ -1 \\ 0 \end{pmatrix} (1 \ 1 \ -1 \ 0) = \frac{1}{3} \begin{pmatrix} 1 & 1 & -1 & 0 \\ 1 & 1 & -1 & 0 \\ -1 & -1 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} = \frac{1}{3} \begin{pmatrix} 2 & -1 & 1 & 0 \\ -1 & 2 & 1 & 0 \\ 1 & 1 & 2 & 0 \\ 0 & 0 & 0 & 3 \end{pmatrix}.$$

SOLUTION DE L'EXERCICE 12.9

1. L'application $P \mapsto P(1)$ est une application linéaire de $\mathbf{R}_2[X]$ dans \mathbf{R} , non nulle⁴ : c'est donc une forme linéaire non nulle sur $\mathbf{R}_2[X]$. Et alors, F est son noyau, donc est un hyperplan de $\mathbf{R}_2[X]$, donc de dimension $3 - 1 = 2$.

Les polynômes $X - 1$ et $X^2 - 1$ sont dans F , et forment une famille libre car ils sont de degré différents. Puisque $\dim F = 2$, $(X - 1, X^2 - 1)$ est donc une base de F .

2. Il est évident⁵ que $\langle \cdot, \cdot \rangle$ est bilinéaire symétrique.

Si $P = aX^2 + bX + c \in \mathbf{R}_2[X]$, alors $\langle P, P \rangle = a^2 + b^2 + c^2 = 0$, et donc $\langle P, P \rangle = 0$ si et seulement si $a = b = c$, c'est-à-dire si et seulement si $P = 0$.

Ainsi, $\langle \cdot, \cdot \rangle$ est un produit scalaire sur $\mathbf{R}_2[X]$.

3. La distance entre X^2 et F est $\min\{\|X^2 - P\|, P \in F\}$. On sait que ce minimum est atteint pour $P_0 = p_F(X^2)$.

³ En fait, il faudrait faire le calcul comme d'habitude, pour réaliser que les deux produits scalaires apparaissant dans la définition de e_3^* sont nuls.

Alternative

On peut procéder comme d'habitude afin de déterminer F^\perp , en utilisant la base de F déterminée à la question 1.

⁴ Par exemple, l'image du polynôme constant égal à 1 est non nulle.

⁵ Bien entendu, une telle rédaction n'est pas acceptable en devoir, il n'y a que moi qui y soit autorisé !

Or on a $p_F(X^2) = a(X-1) + b(X^2-1) \in F$. Soit encore $p_F(X^2) = bX^2 + aX - (a+b)$.

Puisque $X^2 - p_F(X^2) \in F^\perp$, il vient

$$\begin{cases} \langle X-1, X^2 - p_F(X^2) \rangle = \langle X-1, (1-b)X^2 - aX + (a+b) \rangle = 0 \\ \langle X^2-1, X^2 - p_F(x) \rangle = \langle X^2-1, (1-b)X^2 - aX + (a+b) \rangle = 0 \end{cases} \Leftrightarrow \begin{cases} -a - (a+b) = 0 \\ (1-b) - (a+b) = 0 \end{cases} \Leftrightarrow \begin{cases} a = -\frac{1}{3} \\ b = \frac{2}{3} \end{cases}$$

Donc $p_F(X^2) = \frac{1}{3}(1-X + 2(X^2-1)) = \frac{1}{3}(2X^2 - X - 1)$.

On en déduit que la distance cherchée est

$$\left\| X^2 - \frac{1}{3}(2X^2 - X - 1) \right\| = \left\| \frac{1}{3}(X^2 + X + 1) \right\| = \frac{1}{3} \|(X^2 + X + 1)\| \frac{\sqrt{3}}{3} = \frac{1}{\sqrt{3}}.$$

SOLUTION DE L'EXERCICE 12.10

1. Soient $(x, y) \in E^2$. Alors $\langle f(x+y), x+y \rangle = 0$. Mais

$$\langle f(x+y), x+y \rangle = \langle f(x) + f(y), x+y \rangle = \underbrace{\langle f(x), x \rangle}_{=0} + \langle f(x), y \rangle + \langle f(y), x \rangle + \underbrace{\langle f(y), y \rangle}_{=0}.$$

On en déduit donc que

$$\langle f(x), y \rangle + \langle f(y), x \rangle = 0 \Leftrightarrow \langle f(x), y \rangle = -\langle f(y), x \rangle = -\langle x, f(y) \rangle.$$

2. Soit $x \in \text{Ker } f$ et $y \in \text{Im } f$. Alors il existe $z \in E$ tel que $y = f(z)$. Et alors

$$\langle x, z \rangle = \langle x, f(y) \rangle = -\langle f(x), z \rangle = -\langle 0, z \rangle = 0.$$

Ainsi, $\text{Ker } f \subset (\text{Im } f)^\perp$.

Mais par le théorème du rang, on a $\dim \text{Ker } f = n - \dim \text{Im } f$.

D'autre part, on a $\dim(\text{Im } f)^\perp = \dim E - \dim \text{Im } f = \dim \text{Ker } f$.

Or, deux espaces de même dimensions tels que l'un soit inclus dans l'autre sont nécessairement égaux, donc $\text{Ker } f = (\text{Im } f)^\perp$.

3. Soit λ une valeur propre de f , et soit x un vecteur propre associé à la valeur propre λ . Alors

$$\langle x, f(x) \rangle = \langle x, \lambda x \rangle = \lambda \|x\|^2.$$

Mais par hypothèse, ce produit scalaire est nul, et $\|x\|^2$ étant non nul⁶, c'est que nécessairement $\lambda = 0$.

Ainsi, la seule valeur propre possible de f est 0.

On en déduit que f est diagonalisable si et seulement si $E_0(f) = E$, c'est-à-dire si et seulement si f est l'endomorphisme nul.

SOLUTION DE L'EXERCICE 12.11

1. Soient $P, Q \in E$. Alors

$$\langle P, Q \rangle = \sum_{k=-n}^n P(k)Q(k) = \sum_{k=-n}^n Q(k)P(k) = \langle Q, P \rangle.$$

Donc $\langle \cdot, \cdot \rangle$ est symétrique.

Soient $P, Q, R \in E$ et $\lambda \in \mathbf{R}$. Alors

$$\langle \lambda P + Q, R \rangle = \sum_{k=-n}^n (\lambda P(k) + Q(k))R(k) = \lambda \sum_{k=-n}^n P(k)R(k) + \sum_{k=-n}^n Q(k)R(k) = \lambda \langle P, R \rangle + \langle Q, R \rangle.$$

Donc $\langle \cdot, \cdot \rangle$ est linéaire à gauche, et étant symétrique, c'est une forme bilinéaire symétrique.

Soit $P \in E$. Alors

$$\langle P, P \rangle = \sum_{k=-n}^n P(k)^2 \geq 0.$$

De plus, si $\langle P, P \rangle = 0$ alors $\forall k \in \llbracket -n, n \rrbracket, P(k)^2 = 0 \Leftrightarrow \forall k \in \llbracket -n, n \rrbracket, P(k) = 0$.

Mais alors P possède $2n+1$ racines, qui sont les entiers compris entre $-n$ et n . Mais P étant de degré au plus $2n$, cela signifie que P est le polynôme nul.

Ainsi, $\langle \cdot, \cdot \rangle$ est un produit scalaire sur E .

⁶ Un vecteur propre est non nul par définition

Rédaction

On pensera bien à écrire

$$\forall k \in \llbracket -n, n \rrbracket, P(k) = 0$$

et à ne pas se contenter de $P(k) = 0$, sans préciser de quel k il est question.

2. On a

$$\langle P, Q \rangle = P(0)Q(0) + \sum_{k=1}^n (P(k)Q(k) + P(-k)Q(-k)).$$

Si P est pair et Q est impair, alors

$$\langle P, Q \rangle = P(0)Q(0) + \sum_{k=1}^n (P(k)Q(k) - P(k)Q(k)) = P(0)Q(0).$$

Mais si Q est impair, alors $Q(0) = -Q(0) = 0$. Et donc $\langle P, Q \rangle = 0$: P et Q sont orthogonaux.

3. Lorsque a et b parcourent \mathbf{R} , alors $aX - b$ parcourt $\mathbf{R}_1[X]$. Autrement dit, on cherche à minimiser

$$\{\|X^2 - P\|^2, P \in \mathbf{R}_1[X]\}.$$

Alors, un théorème du cours nous assure que ce minimum existe et est atteint uniquement lorsque P est le projeté orthogonal de X^2 sur $\mathbf{R}_1[X]$.

Calculons donc ce projeté orthogonal, noté P .

On sait que $P \in \mathbf{R}_1[X]$, donc il existe $\lambda, \mu \in \mathbf{R}$ tels que $P = \lambda X + \mu$.

De plus, $X^2 - P \in \mathbf{R}_1[X]^\perp$, et une base de $\mathbf{R}_1[X]^\perp$ est $(1, X)$, donc

$$\begin{cases} \langle X^2 - P, 1 \rangle = 0 \\ \langle X^2 - P, X \rangle = 0 \end{cases}$$

Ce système s'écrit alors

$$\begin{cases} \langle X^2 - \lambda X - \mu, 1 \rangle = 0 \\ \langle X^2 - \lambda X - \mu, X \rangle = 0 \end{cases} \Leftrightarrow \begin{cases} \mu \langle 1, 1 \rangle = \langle X^2, 1 \rangle \\ \lambda \langle X, X \rangle = 0 \end{cases}$$

Or, on a $\langle 1, 1 \rangle = \sum_{k=-n}^n 1 = 2n + 1$ et

$$\langle 1, X^2 \rangle = \sum_{k=-n}^n k^2 = 2 \sum_{k=1}^n k^2 = \frac{n(n+1)(2n+1)}{3}.$$

On a donc $\lambda = 0$ et $\mu = \frac{n(n+1)}{3}$.

Par conséquent, le minimum cherché est atteint pour $a = 0$ et $b = \frac{n(n+1)}{3}$.

SOLUTION DE L'EXERCICE 12.12

1. Par définition de la norme, on a

$$\|g - Q\|^2 = \langle g - Q, g - Q \rangle = \int_0^1 (e^t - a - bt)^2 dt.$$

2. C'est un théorème du cours : il existe un unique Q_0 minimisant la distance entre $Q \in F$ et g . Plus précisément : Q_0 est le projeté orthogonal de g sur F .

3. Puisque Q_0 est le projeté orthogonal de g sur F , $g - Q_0 \in F^\perp$, et donc $\langle g - Q_0, f_1 \rangle = \langle g - Q_0, f_2 \rangle = 0$.

4. Notons $Q_0 = a_0 f_2 + b_0 f_1$, de sorte que

$$\langle g - Q_0, f_1 \rangle = \int_0^1 (e^t - a_0 t - b_0) t dt \text{ et } \langle g - Q_0, f_2 \rangle = \int_0^1 (e^t - a_0 t - b_0) dt.$$

Or,

$$\int_0^1 (te^t - a_0 t^2 - b_0 t) dt = \int_0^1 te^t dt - \frac{a_0}{3} - \frac{b_0}{2}.$$

Et une intégration par parties nous donne

$$\int_0^1 te^t dt = [te^t]_0^1 - \int_0^1 e^t dt = e^1 - e^1 + 1 = 1.$$

Parité

On pourrait en fait prouver qu'un polynôme est pair (resp. impair) si et seulement il ne comporte que des termes de degré pair (resp. impair).

Parité

On utilise ici la question précédente, qui nous donne directement

$$\langle 1, X \rangle = \langle X^2, X \rangle = 0.$$

Remarque

Le théorème du cours dit que

$$\|g - Q_0\| = \min_{Q \in F} \|g - Q\|.$$

Mais la fonction carré étant croissante sur \mathbf{R}_+ , minimiser $\|g - Q\|$ ou minimiser son carré revient au même.

Détails

f_1 et f_2 sont tous deux dans F .

Donc $\langle g - Q_0, f_1 \rangle = 0 \Leftrightarrow 1 - \frac{a_0}{3} - \frac{b_0}{2} = 0 \Leftrightarrow 2a_0 + 3b_0 = 6$.

De même, on a

$$\int_0^1 (e^t - a_0 t - b_0) dt = e - 1 - \frac{a_0}{2} - b_0$$

et donc $\langle g - Q_0, f_2 \rangle = 0 \Leftrightarrow a_0 + 2b_0 = 2e - 2$.

Pour trouver a_0 et b_0 , il nous reste à résoudre le système

$$\begin{cases} 2a_0 + 3b_0 = 6 \\ a_0 + 2b_0 = 2e \end{cases} \Leftrightarrow \begin{cases} b_0 = 4e - 10 \\ a_0 = 2e - 2b_0 \end{cases} \Leftrightarrow \begin{cases} a_0 = 18 - 6e \\ b_0 = 4e - 10 \end{cases}$$

On a alors

$$\begin{aligned} \Delta &= \int_0^1 (e^t - a_0 t - b_0)^2 dt \\ &= \int_0^1 e^{2t} dt + a_0^2 \int_0^1 t^2 dt + b_0^2 - 2a_0 \int_0^1 t e^t dt - 2b_0 \int_0^1 e^t dt + 2a_0 b_0 \int_0^1 t dt \\ &= \frac{1}{2}(e^2 - 1) + \frac{a_0^2}{3} + b_0^2 - 2a_0 - 2b_0(e - 1) + a_0 b_0 \\ &= \frac{1}{6} (3e^2 - 3 + 2(18 - 6e)^2 + 6(4e - 10)^2 - 12(18 - 6e) - 12(4e - 10)(e - 1) + 6(4e - 10)(18 - 6e)) \\ &= 20e - \frac{7}{2}e^2 - \frac{57}{2}. \end{aligned}$$

SOLUTION DE L'EXERCICE 12.13

1. On a $f(a) = b$ et $f(b) = -a$, de sorte que $\text{Vect}(a, b) \subset \text{Im}(f)$, et l'inclusion réciproque est évidente⁷.

On en déduit donc que $\text{Im } f = \text{Vect}(a, b)$.

De plus, $x \in \text{Ker } f \Leftrightarrow \langle x, a \rangle b - \langle x, b \rangle a = 0$.

Mais (a, b) est une famille libre de sorte que

$$f(x) = 0 \Leftrightarrow \langle x, a \rangle = \langle x, b \rangle = 0.$$

Puisque (a, b) est une base de $\text{Vect}(a, b)$, on en déduit que $x \in \text{Ker } f \Leftrightarrow x \in \text{Vect}(a, b)^\perp$.

On a donc bien $\text{Ker } f = (\text{Im } f)^\perp$.

Et en prenant l'orthogonal des deux membres de cette égalité, on obtient $(\text{Ker } f)^\perp = ((\text{Im } f)^\perp)^\perp = \text{Im } f$.

2. On sait déjà que 0 est valeur propre de f , et $\dim E_0(f) = \dim \text{Ker } f = n - 2$. Soit (e_2, \dots, e_n) une base orthonormée de $\text{Ker } f$. Alors (a, b, e_2, \dots, e_n) est une base orthonormée de E , car obtenue par concaténation d'une base orthonormée de $\text{Im } f$ et d'une base orthonormée de son supplémentaire orthogonal.

La matrice de f dans cette base est

$$A = \begin{pmatrix} 0 & -1 & 0 & \dots & 0 \\ 1 & 0 & 0 & \dots & 0 \\ 0 & \dots & \dots & \dots & \vdots \\ \vdots & & & & \vdots \\ 0 & \dots & \dots & \dots & 0 \end{pmatrix}.$$

Alors λ est valeur propre de f si et seulement si $\text{rg}(A - \lambda I_n) < n$. Mais

$$A - \lambda I_n = \begin{pmatrix} -\lambda & -1 & 0 & \dots & 0 \\ 1 & -\lambda & 0 & \dots & 0 \\ 0 & \dots & -\lambda & \dots & \vdots \\ \vdots & & & \ddots & \vdots \\ 0 & \dots & \dots & \dots & -\lambda \end{pmatrix} \xleftrightarrow{L_1 \leftrightarrow L_2} \begin{pmatrix} 1 & -\lambda & 0 & \dots & 0 \\ -\lambda & -1 & 0 & \dots & 0 \\ 0 & \dots & -\lambda & \dots & \vdots \\ \vdots & & & \ddots & \vdots \\ 0 & \dots & \dots & \dots & -\lambda \end{pmatrix} \xleftrightarrow{L_2 \leftarrow L_2 + \lambda L_1} \begin{pmatrix} -1 & -\lambda & 0 & \dots & 0 \\ 0 & -\lambda^2 - 1 & 0 & \dots & 0 \\ 0 & \dots & -\lambda & \dots & \vdots \\ \vdots & & & \ddots & \vdots \\ 0 & \dots & \dots & \dots & -\lambda \end{pmatrix}$$

Ainsi, λ est valeur propre de A si et seulement si $\lambda = 0$ ou $-\lambda^2 - 1 = 0$.

Cette deuxième équation n'a pas de solution, donc 0 est l'unique valeur propre de f .

On sait déjà que $\dim E_0(f) = \dim \text{Ker } f = n - \dim(\text{Vect}(a, b)) = n - 2 < \dim E$ et donc f n'est pas diagonalisable.

⁷ Car $f(x)$ est toujours combinaison linéaire de a et b .

Rappel

Une famille orthogonale est toujours libre.

Rappel

On a toujours

$$(F^\perp)^\perp = F.$$

SOLUTION DE L'EXERCICE 12.14

- Classique, voir l'exercice 5.
- Notons qu'étant donné le résultat donné dans l'énoncé, F^\perp doit être de dimension 1, et donc F doit être de dimension $\dim \mathcal{M}_n(\mathbf{R}) - 1$, c'est-à-dire un hyperplan de $\mathcal{M}_n(\mathbf{R})$. Il doit donc y avoir une forme linéaire quelque part...

En effet, notons que F est le noyau de la forme linéaire $\varphi : (m_{i,j})_{1 \leq i,j \leq n} \mapsto \sum_{1 \leq i,j \leq n} m_{i,j}$. Cette

forme linéaire étant non nulle, F est un hyperplan de $\mathcal{M}_n(\mathbf{R})$, et donc de dimension $n^2 - 1$. On en déduit que $\dim F^\perp = \dim \mathcal{M}_n(\mathbf{R}) - (n^2 - 1) = 1$.

De plus, pour $M \in F$, on a $\langle M, J_n \rangle = \sum_{i=1}^n \sum_{j=1}^n m_{i,j} \times 1 = 0$. Et donc $J_n \in F^\perp$.

Puisque $\dim F^\perp = 1$, on en déduit que J_n est une base de F^\perp .

- On a $\delta = \min_{M \in F} \|A - M\|^2$.

La quantité cherchée est donc le carré de la distance de A à F , et nous savons que c'est $\|A - p_F(A)\|^2$, où p_F désigne le projecteur orthogonal sur F .

Mais on a aussi $A - p_F(A) = p_{F^\perp}(A)$, où p_{F^\perp} désigne le projecteur orthogonal sur F^\perp .

Puisque J_n est une base de F^\perp , $\frac{J_n}{\|J_n\|} = \frac{J_n}{n}$ est une base orthonormée de F^\perp . Et donc

$$p_{F^\perp}(A) = \left\langle A, \frac{J_n}{n} \right\rangle \frac{J_n}{n} = \langle A, J_n \rangle \frac{J_n}{n^2}.$$

Et donc

$$\delta = \|p_{F^\perp}(A)\|^2 = \frac{1}{n^4} (\langle A, J_n \rangle)^2 \|J_n\|^2 = \frac{1}{n^2} \left(\sum_{1 \leq i,j \leq n} a_{i,j} \right)^2.$$

SOLUTION DE L'EXERCICE 12.15

- Puisque $x \in \text{Ker}(u - \text{id}) = E_1(u)$, on a $u(x) = x$.
D'autre part, $x \in \text{Im}(u - \text{id})$, et donc il existe $y \in E$ tel que $x = u(y) - y$.
Et alors $x = u(x) = u(u(y) - y) = u^2(y) - u(y)$.
Puis en appliquant de nouveau u , $x = u(x) = u^3(y) - u^2(y)$.
De proche en proche, on a donc, pour tout $k \in \mathbf{N}^*$, $x = u^k(y) - u^{k-1}(y)$.
Et alors pour $n \in \mathbf{N}^*$,

$$u^n(y) - y = \sum_{k=1}^n (u^k(y) - u^{k-1}(y)) = \sum_{k=1}^n x = n \cdot x.$$

Ce résultat est évidemment encore valable pour $n = 0$ car alors $u^n(y) = y$.

- Par le théorème du rang, il est évident que $\dim E = \dim \text{Ker}(u - \text{id}) + \dim \text{Im}(u - \text{id})$.
Il suffit donc de prouver que $\text{Ker}(u - \text{id}) \cap \text{Im}(u - \text{id}) = \{0_E\}$.
Soit donc $x \in \text{Ker}(u - \text{id}) \cap \text{Im}(u - \text{id})$, et soit y comme dans la question précédente.
On a alors pour tout $n \in \mathbf{N}$,

$$\|nx\| = n\|x\| = \|u^n(y) - y\| \leq \|u^n(y)\| + \|y\|.$$

Mais nous savons⁸ que $\|u(y)\| \leq \|y\|$.

Et donc $\|u^2(y)\| = \|u(u(y))\| \leq \|u(y)\| \leq \|y\|$.

Une récurrence rapide prouve donc que pour tout $n \in \mathbf{N}$, $\|u^n(y)\| \leq \|y\|$.

Et donc $n\|x\| \leq \|y\| + \|y\| \leq 2\|y\|$.

Autrement dit, la suite $(n\|x\|)_{n \in \mathbf{N}}$ est bornée, ce qui n'est possible que pour $\|x\| = 0 \Leftrightarrow x = 0_E$.

Ceci prouve donc que $\text{Ker}(u - \text{id}) \cap \text{Im}(u - \text{id}) = \{0_E\}$ et donc que $E = \text{Ker}(u - \text{id}) \oplus \text{Im}(u - \text{id})$.

- Soit $x \in E$. Alors, de manière unique, $x = x_1 + x_2$, avec $x_1 \in \text{Ker}(u - \text{id})$ et $x_2 \in \text{Im}(u - \text{id})$.
On a alors $w(x) = x_1$.
D'autre part, $u(x) = u(x_1) + u(x_2) = x_1 + u(x_2)$, et plus généralement, $u^k(x) = x_1 + u^k(x_2)$,

de sorte que $v_p(x) = x_1 + \frac{1}{p+1} \sum_{k=0}^p u^k(x_2)$.

Et donc $v_p(x) - w(x) = \frac{1}{p+1} \sum_{k=0}^p u^k(x_2)$.

⁸ C'est l'hypothèse faite sur u .

Puisque $x_2 \in \text{Im}(u - \text{id})$, il existe $y \in E$ tel que $x_2 = u(y) - y$.
 Par conséquent, $u^k(x_2) = u^{k+1}(y) - u^k(y)$, de sorte que

$$v_p(x_2) = \frac{1}{p+1} \sum_{k=0}^p (u^{k+1}(y) - u^k(y)) = \frac{1}{p+1} (u^{p+1}(y) - y).$$

Et donc

$$\begin{aligned} \|v_p(x) - w(x)\| &= \|v_p(x_2)\| \\ &= \frac{1}{p+1} \|u^p(y) - y\| \\ &\leq \frac{1}{p+1} (\|u^p(y)\| + \|y\|) && \text{Inégalité triangulaire.} \\ &\leq \frac{1}{p+1} (\|y\| + \|y\|) \\ &\leq \frac{2\|y\|}{p+1} \xrightarrow{p \rightarrow +\infty} 0. \end{aligned}$$

Et donc on a bien $v_p(x)$ qui converge vers $w(x)$.

- 3.a. Si $\text{Im}(Q)$ et $\text{Ker}(Q)$ sont orthogonaux⁹, alors pour tout $x \in E$, il existe $x_1 \in \text{Im}(Q)$ et $x_2 \in \text{Ker}(Q)$ tels que $x = x_1 + x_2$, et $Q(x) = x_1$.
 Et donc, par le théorème de Pythagore,

⁹ C'est-à-dire si Q est un projecteur orthogonal.

$$\|x\|^2 = \|x_1 + x_2\|^2 = \|x_1\|^2 + \|x_2\|^2 \geq \|x_1\|^2 \geq \|Q(x)\|^2.$$

Et donc on a bien $\|Q(x)\| \leq \|x\|$.

- 3.b. Par hypothèse, pour tout $\lambda \in \mathbf{R}$, on a $\|Q(x + \lambda y)\| \leq \|x + \lambda y\| \Leftrightarrow \|x + \lambda y\|^2 - \|x\|^2 \leq 0$.
 Or, $\|x + \lambda y\|^2 = \|x\|^2 + 2\lambda \langle x, y \rangle + \lambda^2 \|y\|^2$.
 Et donc, la fonction $\lambda \mapsto \lambda^2 \|y\|^2 + 2\lambda \langle x, y \rangle$ est de signe constant.
 Ce n'est possible que si $\langle x, y \rangle = 0$.

- 3.c. Nous savons déjà d'après la question 3.a que si Q est un projecteur orthogonal, alors pour tout $x \in E$, $\|Q(x)\| \leq \|x\|$.

Inversement, si Q vérifie pour tout $x \in E$, $\|Q(x)\| \leq \|x\|$, alors nous venons de prouver que tout vecteur de $\text{Ker}(Q)$ est orthogonal à tout vecteur de $\text{Im}(Q)$. Et donc $\text{Ker}(Q) \subset (\text{Im } Q)^\perp$.
 Comme de plus on a $\dim(\text{Im } Q)^\perp = \dim E - \dim \text{Im } Q = \dim \text{Ker } Q$, il vient $\text{Ker } Q = (\text{Im } Q)^\perp$.
 Et donc Q est le projecteur orthogonal sur $\text{Im } Q$.

Danger !

Nous n'avons pas prouvé l'égalité entre $\text{Ker}(Q)$ et $(\text{Im } Q)^\perp$.
 Il se pourrait encore qu'il y ait d'autres vecteurs dans $(\text{Im } Q)^\perp$ que ceux de $\text{Ker } Q$.

4. Par définition, w est un projecteur, et donc il s'agit de prouver que pour tout $x \in E$, $\|w(x)\| \leq \|x\|$.
 Or, pour tout $p \in \mathbf{N}$, $\|w(x)\| = \|w(x) - v_p(x) + v_p(x)\| \leq \|w(x) - v_p(x)\| + \|v_p(x)\|$.
 Et

$$\|v_p(x)\| \leq \frac{1}{p+1} \sum_{k=0}^p \|u^k(x)\| \leq \frac{1}{p+1} \sum_{k=0}^p \|x\| \leq \|x\|.$$

Puisque $\lim_{p \rightarrow +\infty} \|w(x) - v_p(x)\| = 0$, il vient, par passage à la limite,

$$\|w(x)\| \leq \|x\|.$$

Et donc w est un projecteur orthogonal.

INTRODUCTION AUX FONCTIONS DE PLUSIEURS VARIABLES

Cours	393
13.1 Fonctions de plusieurs variables	393
13.2 Fonctions continues	396
13.3 Calcul différentiel du premier ordre	398
13.4 Extrema globaux	406
Méthodes	411
Exercices	412
Correction des exercices	415

INTRODUCTION

L'utilisation de la dérivée et du tableau de variation permet facilement de trouver les extrema d'une fonction de \mathbf{R} dans \mathbf{R} .

Toutefois, dans «la vraie vie», les fonctions que l'on considère dépendent rarement d'un unique paramètre, mais plutôt de plusieurs facteurs.

Par exemple, considérons un horloger de luxe, qui souhaite lancer un nouveau modèle en France et en Suisse, et qui s'autorise des prix différents dans les deux pays.

En France, une étude de marché a montré que le nombre q_1 d'exemplaires qu'il pouvait espérer vendre est relié au prix de vente p_1 (en dizaine de milliers d'euros) de la manière suivante : $q_1 = 120 - 4p_1$.

De même, en Suisse, on a $q_2 = 140 - 2p_2$.

On considère que les coûts de production sont donnés par $C = 1000 + 2q^2$, où $q = q_1 + q_2$ est le nombre total de montres produites.

Le profit de l'entreprise est alors

$$\begin{aligned} G(p_1, p_2) &= p_1 q_1 + p_2 q_2 - C \\ &= p_1(120 - 4p_1) + p_2(140 - 2p_2) - 2(p_1 + p_2)^2 - 1000 \\ &= -6p_1^2 - 4p_2^2 - 4p_1 p_2 + 120p_1 + 140p_2 - 1000. \end{aligned}$$

Quelles sont les valeurs de p_1 et p_2 pour lesquelles ce profit est maximal ?

Autrement dit, quelles sont les $(p_1, p_2) \in \mathbf{R}^2$ maximisant $G(p_1, p_2)$?

Interprétation

Plus le prix de vente est élevé, plus la demande est faible.

Interprétation

Le terme 1000 correspond à des frais fixes de l'entreprise, indépendants de sa production.

13.1 FONCTIONS DE PLUSIEURS VARIABLES

Dans tout ce chapitre, on notera $\langle \cdot, \cdot \rangle$ le produit scalaire canonique sur \mathbf{R}^n , et $\| \cdot \|$ la norme associée. Ainsi

$$\forall x = (x_1, \dots, x_n), y = (y_1, \dots, y_n) \in \mathbf{R}^n, \langle x, y \rangle = \sum_{i=1}^n x_i y_i \text{ et } \|x\| = \sqrt{\sum_{i=1}^n x_i^2}.$$

13.1.1 Quelques précisions sur la norme euclidienne

Rappelons que la norme $\| \cdot \|$ sur \mathbf{R}^n vérifie les trois propriétés suivantes :

- $\forall x \in \mathbf{R}^n, \forall \lambda \in \mathbf{R}, \|\lambda x\| = |\lambda| \cdot \|x\|$.
- $\forall x \in \mathbf{R}^n, \|x\| = 0 \Leftrightarrow x = 0$.
- $\forall x, y \in \mathbf{R}^n, \|x + y\| \leq \|x\| + \|y\|$ (inégalité triangulaire).

La norme d'un vecteur mesure sa «longueur». Si $n = 2$ ou $n = 3$, on retrouve les expressions déjà connues dans le plan et l'espace.

De plus, $\|x - y\|$ représente la distance entre deux n -uplets x et y : si $\|x - y\|$ est petit, alors x et y sont proches.

Ainsi, la norme est en quelque sorte l'analogie pour \mathbf{R}^n de la valeur absolue pour \mathbf{R} .

13.1.2 Premières définitions

Définition 13.1 – Si D est une partie de \mathbf{R}^n , une application f de $D \rightarrow \mathbf{R}$, définie par

$$(x_1, \dots, x_n) \mapsto f(x_1, \dots, x_n)$$

est appelée fonction (numérique¹) à n variables.

¹ Numérique signifie que la fonction est à valeurs dans \mathbf{R} (elle prend pour valeurs des nombres).

Exemples 13.2

- $f : \mathbf{R}^2 \rightarrow \mathbf{R}$ définie par $f(x, y) = \ln(1 + x^2 + y^2)$.
- $g : \mathbf{R}^3 \rightarrow \mathbf{R}$ définie par $g(x, y, z) = \begin{cases} 0 & \text{si } (x, y, z) = 0 \\ \frac{e^{x+y-2z}}{x^2 + y^2 + z^2} & \text{sinon} \end{cases}$
- $h : \mathbf{R}^2 \rightarrow \mathbf{R}$ définie par $h(x, y) = y^2(2x + 3y) - x(xy)$
- $i : \mathbf{R}^3 \rightarrow \mathbf{R}$ définie par $i(x, y, z) = xyz(x + y + z - x^2)$
- $j(x, y) = \frac{2xy}{x - y}$ est définie sur $\{(x, y) \in \mathbf{R}^2 : x \neq y\}$.

Terminologie

On dit que h et i sont des fonctions polynomiales.

13.1.3 Graphe, courbes de niveau

Définition 13.3 – Soit $f : \mathbf{R}^n \rightarrow \mathbf{R}$. On appelle graphe de f l'ensemble de \mathbf{R}^{n+1} défini par

$$\{(x_1, \dots, x_{n+1}) \in \mathbf{R}^{n+1} : x_{n+1} = f(x_1, \dots, x_n)\}.$$

Si $n = 1$, on retrouve la courbe d'une fonction d'une variable réelle : $\mathcal{G}_f = \{(x, y) \in \mathbf{R}^2 : y = f(x)\}$.

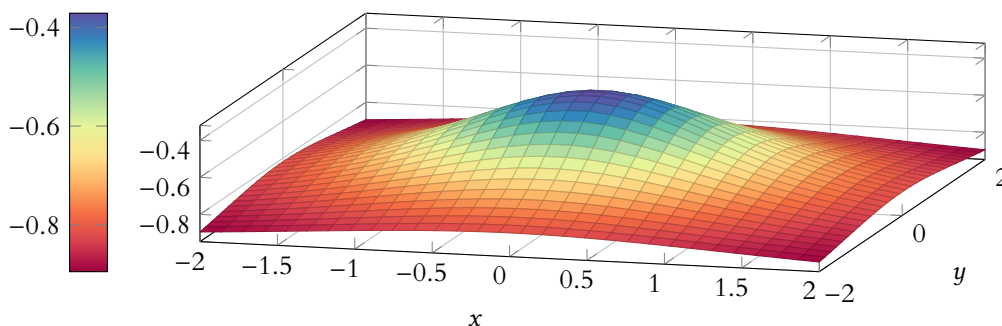
Si $n = 2$, alors la courbe de f est une surface tracée dans l'espace.

Remarque

Si $n \geq 3$, alors il est très dur de se représenter le graphe de f , qui vit alors dans \mathbf{R}^4 , espace «trop gros» pour qu'on se le représente facilement.

Exemples 13.4

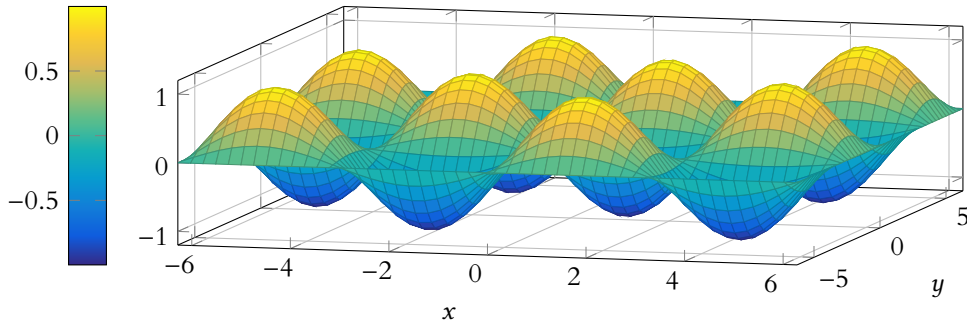
- Soit $f(x, y) = -e^{-\frac{1}{x^2+y^2+1}}$. Alors le graphe de f est



- Soit $g(x, y) = \sin(x) \sin(y)$. Alors le graphe de g est

2D/3D

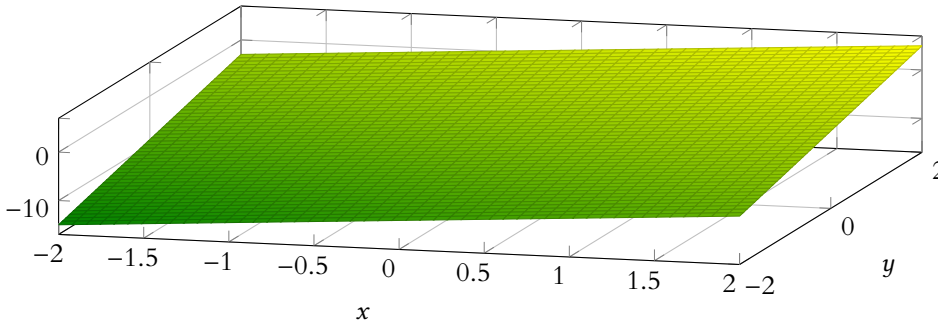
Bien qu'il s'agisse d'une représentation «en 3 dimensions», ces figures sont toujours représentées sur le plan de la feuille ou de l'écran. Les couleurs aident alors à se représenter une troisième dimension, les variations de couleurs représentant les variations d'altitude.



- Si l'on considère une partie de la surface du globe, et qu'on la suppose plate. Alors tout point de la surface est représenté par deux coordonnées x et y (qui sont la latitude et la longitude). Alors si l'on note $h(x, y)$ l'altitude en ce point, le graphe de la fonction h représente la surface de la Terre.

Cas particulier : si $f(x, y) = ax + by + c$, alors le graphe de f est un plan de l'espace, car c'est l'ensemble des points (x, y, z) vérifiant

$$z = ax + by + c \Leftrightarrow z - ax - by - c = 0.$$

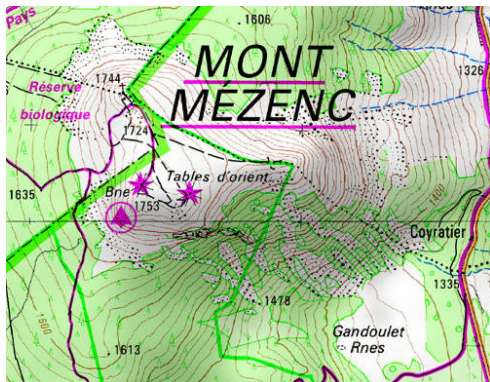


Remarque
 Les plans «verticaux», c'est-à-dire ceux qui possèdent une équation de la forme $ax + by = c$ ne sont pas des graphes de fonctions (de même que les droites de la forme $x = a$ ne sont pas des représentations de fonctions affines). De manière générale, le graphe d'une fonction ne possède qu'un seul point d'abscisse et d'ordonnée données.

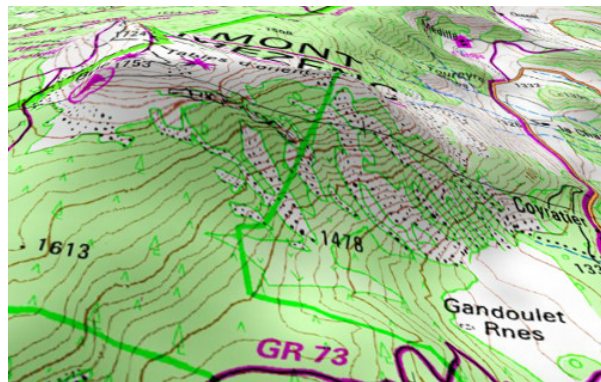
Définition 13.5 – Soit $f : \mathbb{R}^n \rightarrow \mathbb{R}$, et soit $\lambda \in \mathbb{R}$. On appelle ligne de niveau λ de f l'ensemble des points $(x_1, \dots, x_n) \in \mathbb{R}^n$ qui vérifient l'équation $f(x_1, \dots, x_n) = \lambda$.

Exemples 13.6

Si f est la fonction qui à la longitude et la latitude associe l'altitude, alors les lignes de niveau représentent les points qui sont à la même altitude : si on se promène sur une ligne de niveau, on ne monte ni ne descend. Ce sont bien les lignes de niveau représentées sur les cartes topographiques.



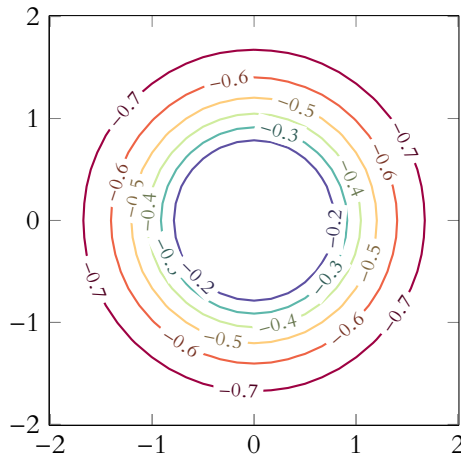
Des lignes de niveau en Haute-Loire



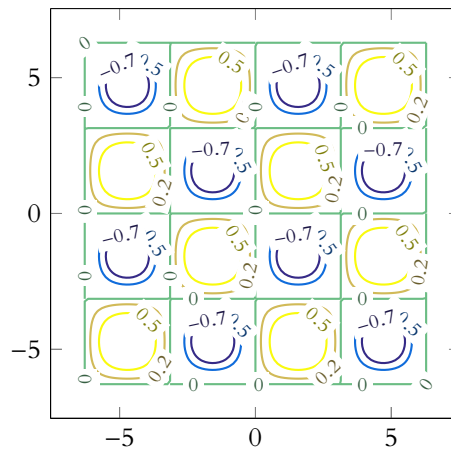
Les mêmes lignes de niveau en 3D.

Considérons une fonction $f : \mathbf{R}^3 \rightarrow \mathbf{R}$ qui en tout point de l'espace donne la température en ce point. Alors la ligne de niveau 0 est appelée en météorologie l'isotherme 0. C'est l'ensemble des points de l'espace en lesquels la température est égale à 0°C . C'est une surface de l'espace.

Des courbes de niveau des deux fonctions f et g données précédemment sont



Des lignes de niveau de f .



Des lignes de niveau de g .

Interprétation

Il n'est pas surprenant que les lignes de niveau de f soient des cercles de centre O . En effet, sur un tel cercle, $x^2 + y^2$ est constant, et donc f l'est également.

Quand aux lignes de niveau de g , on y voit clairement que g est nulle si et seulement si $x = k\pi$ ou $y = \ell\pi$, $(k, \ell) \in \mathbf{N}^2$.

13.1.4 Utilisation de Scilab

Scilab permet de tracer des fonctions de deux variables à l'aide de la commande `fplot3d`. Pour cela il faut d'abord créer la fonction f à l'aide de l'instruction `function`.

Ensuite, on se donne l'intervalle sur lequel on veut représenter f à l'aide de la commande `linspace`, de manière à créer deux vecteur x et y .

Enfin, l'instruction `fplot3d(x,y,f)` permet d'obtenir le graphique désiré.

Exemple 13.7

Pour obtenir le graphe de $f(x,y) = -e^{-\frac{1}{x^2+y^2}}$, on tape

```
1 function z=f(x,y)
2   z = -exp(-1/(x^2 + y^2))
3 endfunction
4 x = linspace(-2,2,100)
5 y = linspace(-2,2,100)
6 fplot3d(x,y,f)
```

De même, Scilab permet d'obtenir des lignes de niveau de f à l'aide la commande

`contour(x,y,f,10)`

Le 10 sert à préciser le nombre de lignes de niveau souhaité. Notons qu'il n'est pas possible de demander directement à obtenir telle ou telle ligne de niveau.

Nous reviendrons sur toutes ces commandes lors d'un TP en fin d'année.

13.2 FONCTIONS CONTINUES

Définition 13.8 – Soit $f : \mathbf{R}^n \rightarrow \mathbf{R}$. On dit que f est continue en $a = (a_1, \dots, a_n) \in \mathbf{R}^n$ si

$$\forall \varepsilon > 0, \exists \eta > 0 : \forall x = (x_1, \dots, x_n) \in \mathbf{R}^n, \|x - a\| \leq \eta \Rightarrow |f(x) - f(a)| \leq \varepsilon.$$

f est continue sur \mathbf{R}^n si pour tout point $a \in \mathbf{R}^n$, f est continue en a .

Détails

En pratique, les deux vecteurs x et y permettent à Scilab de créer un maillage de \mathbf{R}^2 . Les images de ces points sont alors calculées, puis reliées par des morceaux de plans. L'image ainsi obtenue est généralement assez fade, et il faut utiliser d'autres commandes (dont nous parlerons pas) pour obtenir notamment des couleurs plus lisibles.

Analogie

Cette définition est très similaire à celle de la continuité d'une fonction de \mathbf{R} dans \mathbf{R} : si la distance entre x et a est «suffisamment petite», alors $f(x)$ et $f(a)$ sont proches.

Exemple 13.9 Continuité des applications coordonnées

Soit $i \in \llbracket 1, n \rrbracket$, et soit $p_i : \mathbf{R}^n \rightarrow \mathbf{R}$ définie par $p_i(x_1, \dots, x_n) = x_i$. Alors p_i est continue sur \mathbf{R}^n .

En effet, soit $(a_1, \dots, a_n) \in \mathbf{R}^n$, et soit $\varepsilon > 0$.

Alors, si $\|x - a\| \leq \varepsilon$, on a

$$|p_i(x) - p_i(a)| = |x_i - a_i| \leq \sqrt{\sum_{j=1}^n (x_j - a_j)^2} = \|x - a\| \leq \varepsilon.$$

Ainsi, en posant $\eta = \varepsilon$, on a bien

$$\forall \varepsilon > 0, \|x - a\| \leq \eta \Rightarrow |p_i(x) - p_i(a)| \leq \varepsilon.$$

Proposition 13.10 : Soient f, g deux fonctions continues sur \mathbf{R}^n et à valeurs dans \mathbf{R} . Alors

- si $\lambda \in \mathbf{R}$, λf est continue sur \mathbf{R}^n
- $f + g$ est continue sur \mathbf{R}^n
- $f \times g$ est continue sur \mathbf{R}^n
- si g ne s'annule pas sur \mathbf{R}^n , $\frac{f}{g}$ est continue sur \mathbf{R}^n .

Démonstration. Les preuves sont en fait exactement les mêmes que pour les propriétés analogues pour les fonctions d'une variable, en remplaçant les valeurs absolues par les normes. \square

Corollaire 13.11 – Les fonctions polynomiales sont continues sur \mathbf{R}^n .

Démonstration. Si $i \in \llbracket 1, n \rrbracket$ et si $r \in \mathbf{N}$, alors $(x_1, \dots, x_n) \mapsto x_i^r = p_i(x)^r$ est continue sur \mathbf{R}^n comme produit de fonctions continues.

De même, les fonctions $(x_1, \dots, x_n) \mapsto x_1^{r_1} \cdots x_n^{r_n}$ sont continues car produits de fonctions continues.

Les fonctions polynomiales sont des sommes de telles fonctions, donc sont continues. \square

Proposition 13.12 : Soit I un intervalle de \mathbf{R} , soit $\varphi : I \rightarrow \mathbf{R}$ une fonction continue, et soit $f : \mathbf{R}^n \rightarrow \mathbf{R}$ une fonction continue sur \mathbf{R}^n telle que

$$\forall (x_1, \dots, x_n) \in \mathbf{R}^n, f(x_1, \dots, x_n) \in I.$$

Alors $\varphi \circ f : \mathbf{R}^n \rightarrow \mathbf{R}$ est continue sur \mathbf{R}^n .



Démonstration. L'idée principale est que, par continuité de f , si $x \in \mathbf{R}^n$ et $y \in \mathbf{R}^n$ sont suffisamment proches, alors $f(x)$ et $f(y)$ sont suffisamment proches. Mais alors, par continuité de φ , $\varphi(f(x))$ et $\varphi(f(y))$ sont également suffisamment proches.

Plus précisément : soit $x = (x_1, \dots, x_n) \in \mathbf{R}^n$, et soit $\varepsilon > 0$.

La fonction φ est continue en $f(x)$, et donc il existe $\mu > 0$ tel que

$$\forall t \in I, |t - f(x)| \leq \mu \Rightarrow |\varphi(t) - \varphi(f(x))| \leq \varepsilon.$$

Puisque f est continue en x , il existe $\eta > 0$ tel que

$$\forall y \in \mathbf{R}^n, \|y - x\| \leq \eta \Rightarrow |f(y) - f(x)| \leq \mu.$$

Définition

On ne définira pas formellement les fonctions polynomiales à plusieurs variables, mais ce sont les sommes et produits de fonctions de la forme $x_1^{k_1}, \dots, x_n^{k_n}$, où les k_i sont des entiers positifs. Par exemple $x^2y - y^3z^2$ est polynomiale, $\frac{x}{y}, y\sqrt{x}$, $\cos(x^2)y^3$ ne le sont pas.

Remarque

L'hypothèse $f(x_1, \dots, x_n) \in I$ sert à garantir que la fonction composée $\varphi \circ f$ est bien définie. Par exemple $\ln(1 + x^2 + 2y^2)$ est bien définie sur \mathbf{R}^2 tout entier, alors que ce n'est pas le cas de $\sqrt{1 - x^2 - y^2}$.

Et donc, pour $y \in \mathbf{R}^n$ tel que $\|x - y\| \leq \eta$, on a $\|f(y) - f(x)\| \leq \mu$ et donc $|\varphi(f(y)) - \varphi(f(x))| \leq \varepsilon$.
Autrement dit, nous venons de prouver que pour tout $\varepsilon > 0$, il existe $\eta > 0$ tel que

$$\forall y \in \mathbf{R}^n, \|y - x\| \leq \eta \Rightarrow |\varphi(f(y)) - \varphi(f(x))| \leq \varepsilon.$$

Et donc $\varphi \circ f$ est continue en x , et ce quel que soit $x \in \mathbf{R}^n$: elle est continue sur \mathbf{R}^n . \square

Déjà vu ?

Cette preuve est la même que celle de la proposition analogue pour les fonctions d'une variable, nous avons seulement remplacé les valeurs absolues par des normes.

Exemples 13.13

- Soit $f(x, y) = x \ln(1 + x^2 + y^2)$.

Alors f est continue sur \mathbf{R}^2 car :

- $(x, y) \mapsto 1 + x^2 + y^2$ est continue sur \mathbf{R}^2 (car polynomiale) et à valeurs dans \mathbf{R}_+^*
- $t \mapsto \ln(t)$ est continue sur \mathbf{R}_+^* . Donc $(x, y) \mapsto \ln(1 + x^2 + y^2)$ est continue sur \mathbf{R}^2 par composition de fonctions continues.
- $(x, y) \mapsto x$ est continue sur \mathbf{R}^2 (car polynomiale).

Donc f est continue comme produit de fonctions continues.

- Soit $g : (x, y) \mapsto \max(x, y)$.

Alors on a $g(x, y) = \frac{1}{2}(x + y + |x - y|)$.

Or, la fonction $(x, y) \mapsto x + y$ est continue car polynomiale, de même que la fonction $(x, y) \mapsto x - y$.

Puisque la fonction $t \mapsto |t|$ est continue sur \mathbf{R} , par composition de fonctions continues, la fonction $(x, y) \mapsto |x - y|$ est continue sur \mathbf{R}^2 .

Et donc g est continue sur \mathbf{R}^2 .

Exercice

Prouver de même que la fonction $(x, y) \mapsto \min(x, y)$ est continue sur \mathbf{R}^2 .

13.3 CALCUL DIFFÉRENTIEL DU PREMIER ORDRE

13.3.1 Dérivées partielles d'ordre 1, gradient

Définition 13.14 – Soit $f : \mathbf{R}^n \rightarrow \mathbf{R}$, et soit $a = (a_1, \dots, a_n) \in \mathbf{R}^n$. Alors pour $i \in \llbracket 1, n \rrbracket$, la i -ème fonction partielle de f en a est la fonction $f_i : \mathbf{R} \rightarrow \mathbf{R}$ définie par

$$f_i(t) = f(a_1, \dots, a_{i-1}, t, a_{i+1}, \dots, a_n).$$

Explication

f_i est la fonction d'une seule variable obtenue à partir de f en fixant toutes les variables, sauf la i -ème que l'on fait varier.

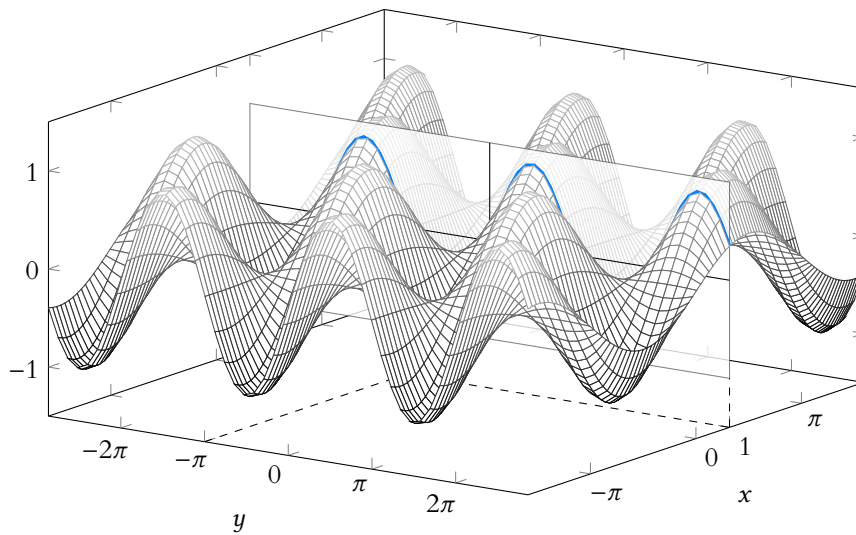
Remarque. Si f est continue sur \mathbf{R}^n , alors toutes ses fonctions partielles sont continues².

² En tant que fonctions d'une variable.

Exemple 13.15

Dans le cas où $n = 2$, et $a = (a_1, a_2)$. f_2 est alors la fonction $t \mapsto f(a_1, t)$.

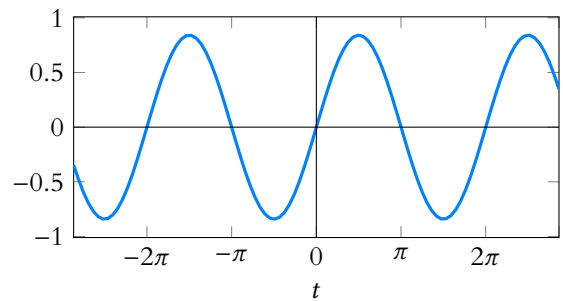
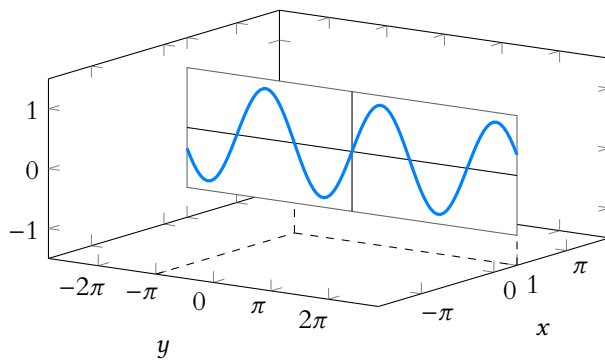
La courbe de f_2 est alors obtenue en faisant l'intersection du graphe de f avec le plan d'équation $x = a_1$.



Ici, on considère la fonction

$$f(x, y) = \sin(x)\sin(y)$$

On souhaite tracer la courbe de la seconde fonction partielle de f en $a = (1, -\pi)$. Pour cela, on s'intéresse à l'intersection du graphe de f et du plan d'équation $x = 1$.



Remarquons que la i -ème fonction partielle en $a = (a_1, \dots, a_n)$ ne dépend que des $a_j, j \neq i$, et pas de a_i . Ici, la seconde fonction partielle en $(1, -\pi)$ est la même qu'en $(1, 0)$ (dans les deux cas, on coupe par le plan d'équation $x = 1$.)

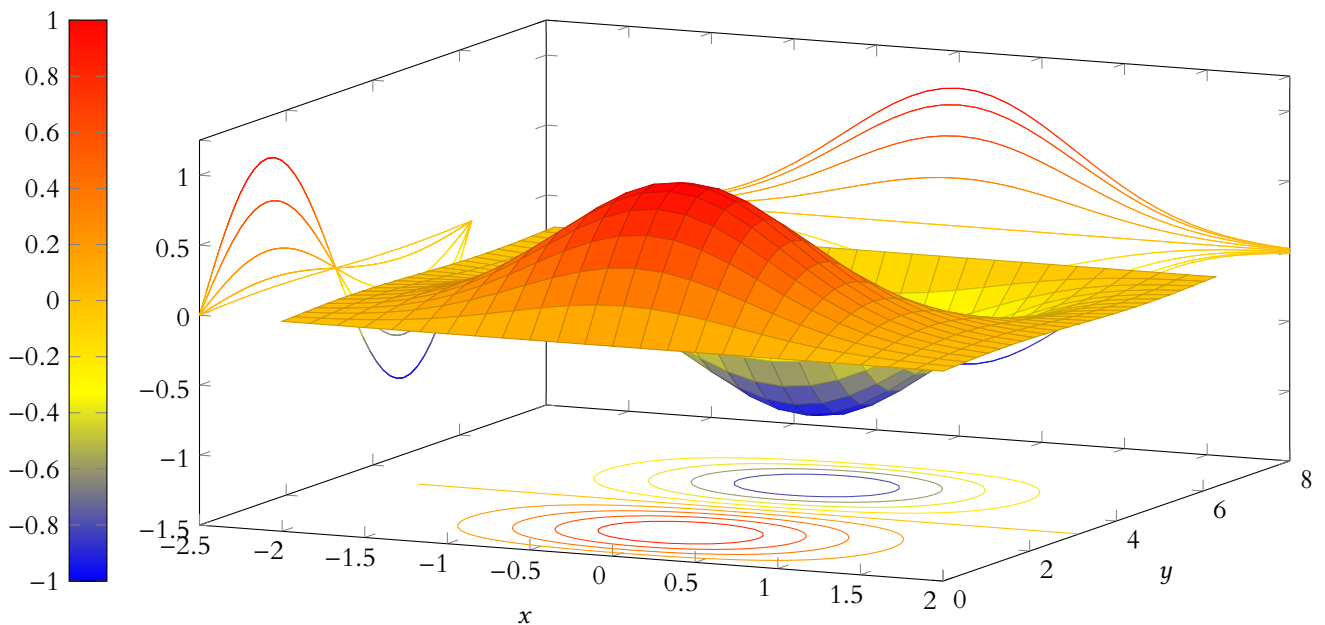


FIGURE 13.1 – Sur le plan du fond sont représentées des premières fonctions partielles (sections par les plans $y = \lambda$), sur le plan de gauche des secondes fonctions partielles, et sur le plan du bas, des lignes de niveau.

Définition 13.16 – Soit $f : \mathbf{R}^n \rightarrow \mathbf{R}$, et soit $a = (a_1, \dots, a_n) \in \mathbf{R}^n$. On dit que f admet une i -ème dérivée partielle en a si f_i est dérivable en a_i , c'est-à-dire si

$$\lim_{h \rightarrow 0} \frac{f(a_1, \dots, a_{i-1}, a_i + h, a_{i+1}, \dots, a_n) - f(a_1, \dots, a_n)}{h} \text{ existe.}$$

On note alors $\partial_i f(a)$ cette limite.

Si pour tout $a \in \mathbf{R}^n$, f admet une i -ème dérivée partielle en a , on note $\partial_i f$ la fonction de \mathbf{R}^n dans \mathbf{R} qui à $a \in \mathbf{R}^n$ associe $\partial_i f(a)$, et on l'appelle i -ème dérivée partielle de f .

En pratique

Pour étudier l'existence, et le cas échéant calculer $\partial_i f$, il suffit de fixer toutes les variables sauf la i -ème, et d'étudier la fonction $x_i \mapsto f(x_1, \dots, x_n)$, qui est une fonction de la seule variable x_i . Toutes les règles apprises en première année sur les fonctions d'une variable sont valables, et il n'est pas nécessaire de revenir au taux d'accroissement.

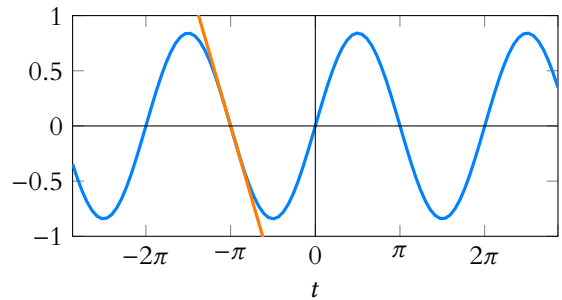
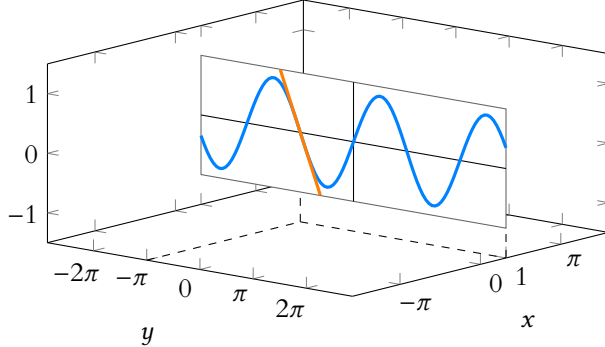
Exemple 13.17

Reprenons le cas de $f(x, y) = \sin(x) \sin(y)$ et $a = (1, -\pi)$.

Alors la seconde fonction partielle de f en a est $f_2 : t \mapsto f(1, t) = \sin(1) \sin(t)$.

Sa dérivée est $t \mapsto \sin(1) \cos(t)$ et donc $\partial_2 f(1, -\pi) = \sin(1) \cos(\pi) = -\sin(1)$.

Graphiquement, cela signifie que $-\sin(1)$ est le coefficient directeur de la tangente à la courbe de f_2 en $t = -\pi$.



Exemples 13.18

• Soit $f : \mathbf{R}^3 \rightarrow \mathbf{R}$ définie par $f(x, y, z) = xyz + 2x^3 + yz$.

Alors f admet des dérivées partielles, et

$$\partial_1 f(x, y, z) = yz + 6x^2, \partial_2 f(x, y, z) = xz + z, \partial_3 f(x, y, z) = xy + y.$$

• Soit $g : \mathbf{R}^2 \rightarrow \mathbf{R}$ définie par $g(x, y) = xe^{x^2+y^2}$. Alors g admet des dérivées partielles

$$\partial_1 g(x, y) = e^{x^2+y^2} + 2xe^{x^2+y^2}, \partial_2 g(x, y) = 2xye^{x^2+y^2}.$$

Définition 13.19 – Soit $f : \mathbf{R}^n \rightarrow \mathbf{R}$ et $a \in \mathbf{R}^n$ tels que $\forall i \in \llbracket 1, n \rrbracket$, f admette une i -ème dérivée partielle en a .

Alors on appelle gradient de f en a et on note $\nabla f(a)$ le vecteur de \mathbf{R}^n défini par

$$\nabla f(a) = (\partial_1 f(a), \dots, \partial_n f(a)).$$

⚠ Attention !

Le gradient est un vecteur, c'est donc n nombres, et non un seul !

Exemple 13.20

Soit $f(x, y) = x^2(1 + \ln(1 + y^2))$. Alors en tout point (x, y) , on a

$$\partial_1 f(x, y) = 2x(1 + \ln(y^2)), \partial_2 f(x, y) = \frac{2yx^2}{1 + y^2}, \nabla f(x, y) = \left(2x(1 + \ln(y^2)), \frac{2yx^2}{1 + y^2} \right) \in \mathbf{R}^2.$$

13.3.2 Fonctions de classe \mathcal{C}^1 .

Définition 13.21 – Une fonction $f : \mathbf{R}^n \rightarrow \mathbf{R}$ est de classe \mathcal{C}^1 sur \mathbf{R}^n si pour tout $i \in \llbracket 1, n \rrbracket$, f_i admet une i -ème dérivée partielle, et si $\partial_i f$ est continue sur \mathbf{R}^n .

Proposition 13.22 : Soient f, g deux fonctions de classe \mathcal{C}^1 sur \mathbf{R}^n . Alors

- $\forall \lambda \in \mathbf{R}, \lambda f$ est de classe \mathcal{C}^1 sur \mathbf{R}^n .
- $f + g$ est de classe \mathcal{C}^1 sur \mathbf{R}^n .
- $f \times g$ est de classe \mathcal{C}^1 sur \mathbf{R}^n .
- Si g ne s'annule pas sur \mathbf{R}^n , alors $\frac{f}{g}$ est de classe \mathcal{C}^1 sur \mathbf{R}^n .
- Si $\varphi : I \rightarrow \mathbf{R}$ est de classe \mathcal{C}^1 sur I et si f prend ses valeurs dans I , alors $\varphi \circ f$ est de classe \mathcal{C}^1 sur \mathbf{R}^n .

Remarque. Les règles de calcul de la dérivée d'une somme, d'un produit, etc permettent de calculer les dérivées partielles d'une somme, d'un produit, d'une composée, etc.

Proposition 13.23 : Les fonctions polynomiales sont de classe \mathcal{C}^1 sur \mathbf{R}^n .

Démonstration. Soit comme précédemment p_i l'application définie par $p_i(x_1, \dots, x_n) = x_i$. Alors $\partial_i p_i = 1$ et $\forall j \in \llbracket 1, n \rrbracket, j \neq i, \partial_j p_i = 0$.

Ces dérivées partielles sont continues, donc p_i est de classe \mathcal{C}^1 .

Par somme et produit de fonctions continues, toute fonction polynomiale est \mathcal{C}^1 sur \mathbf{R}^n . \square

Exemples 13.24

- Soit $f(x, y) = xy e^{-x+y}$.
La fonction $(x, y) \mapsto -x + y$ est \mathcal{C}^1 car polynomiale. Puisque $t \mapsto e^t$ est \mathcal{C}^1 sur \mathbf{R} , la fonction $(x, y) \mapsto e^{-x+y}$ est \mathcal{C}^1 sur \mathbf{R}^n .
Enfin, $(x, y) \mapsto xy$ est \mathcal{C}^1 sur \mathbf{R}^n car polynomiale, donc par produit de fonctions \mathcal{C}^1 , f est \mathcal{C}^1 sur \mathbf{R}^n .

- La fonction norme $(x_1, \dots, x_n) \mapsto \sqrt{\sum_{i=1}^n x_i^2}$ n'est pas \mathcal{C}^1 sur \mathbf{R}^n .

En effet, si $(x_1, \dots, x_{n-1}) = (0, \dots, 0)$, alors $\partial_n f(0, \dots, 0)$ n'existe pas car $t \mapsto \sqrt{t}$ n'est pas dérivable en 0.

13.3.3 Formule de Taylor à l'ordre 1

Théorème 13.25 (Formule de Taylor à l'ordre 1) : Soit f une fonction de classe \mathcal{C}^1 sur \mathbf{R}^n , et soit $a \in \mathbf{R}^n$.

Alors il existe une fonction $\varepsilon : \mathbf{R}^n \rightarrow \mathbf{R}$, continue en 0 et vérifiant $\varepsilon(0) = 0$ telle que

$$\forall h \in \mathbf{R}^n, f(a+h) = f(a) + \langle \nabla f(a), h \rangle + \|h\| \varepsilon(h).$$

Ceci peut également s'écrire :

$$\forall h = (h_1, \dots, h_n) \in \mathbf{R}^n, f(a+h) = f(a) + \sum_{i=1}^n h_i \partial_i f(a_1, \dots, a_n) + \sqrt{\sum_{i=1}^n h_i^2} \varepsilon(h_1, \dots, h_n).$$

Analogie

Cette formule est à voir comme un analogue à la formule de Taylor-Young à l'ordre 1 en dimension supérieur ou égale à 2. En effet, le terme $\langle \nabla f(a), h \rangle$ est un terme de degré 1 : il s'agit d'une combinaison linéaire des h_i , et c'est donc l'analogue du terme en $x - a$ de Taylor-Young. Quant à la fonction $\|h\| \varepsilon(h)$, c'est une fonction qui est «petite quand h est petit», c'est l'analogue du $o(x - a)$.

Remarques. • Pour $n = 1$, et en posant $h = x - a$, on retrouve la formule de Taylor-Young à l'ordre 1 :

$$f(x) = f(a) + f'(a)(x - a) + o(x - a).$$

Si l'on néglige le terme en $o(x - a)$, on retrouve l'équation de la tangente à f en a : c'est la droite d'équation $y = f(a) + f'(a)(x - a)$.

• Pour $n = 2$, en négligeant le terme $\|h\|\varepsilon(h)$, on obtient, pour h suffisamment petit,

$$f(a + h) \simeq f(a) + \langle \nabla f(a), h \rangle.$$

Soit, en posant $a + h = (x, y)$,

$$f(a + h) \simeq f(a) + \partial_1 f(a)(x - a_1) + \partial_2 f(a)(y - a_2).$$

Or l'ensemble des points $(x, y, z) \in \mathbb{R}^3$ vérifiant

$$z = f(a) + \partial_1 f(a)(x - a_1) + \partial_2 f(a)(y - a_2)$$

est un plan. On l'appelle le plan tangent au graphe de f au point a .

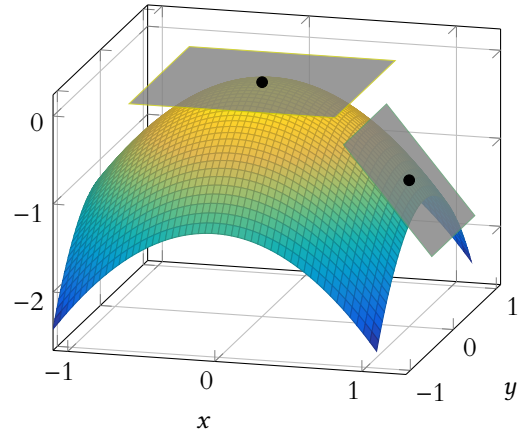


FIGURE 13.2- Deux plans tangents

13.3.4 Dérivées directionnelles

Définition 13.26 – Soit $x, u \in \mathbb{R}^n$, avec u non nul. Alors la droite passant par x et de direction u est

$$\{x + tu, t \in \mathbb{R}\}.$$

Remarque. Si v est un vecteur colinéaire à u , alors la droite passant par x et de direction v est la même que la droite passant par x et de direction u . La norme d'un vecteur et son sens n'ont aucune influence, seule sa direction compte. On peut toujours, quitte à diviser u par $\|u\|$ supposer que u est un vecteur unitaire.

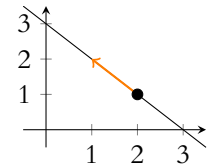


FIGURE 13.3- La droite passant par $(2, 1)$ et de direction $u = (-1, 1)$.

Proposition 13.27 : Soit $f : \mathbb{R}^n \rightarrow \mathbb{R}$ de classe \mathcal{C}^1 sur \mathbb{R}^n , et soient $a, u \in \mathbb{R}^n$. Alors la fonction d'une variable

$$g : t \mapsto g(a + tu)$$

est de classe \mathcal{C}^1 sur \mathbb{R} , et

$$\forall t \in \mathbb{R}, g'(t) = \langle \nabla f(a + tu), u \rangle = \sum_{i=1}^n u_i \partial_i f(a + tu).$$



Démonstration. Soient $t, h \in \mathbb{R}$. Prouvons que $\lim_{h \rightarrow 0} \frac{g(t+h) - g(t)}{h}$ existe. On a

$$g(t + h) = f(a + (t + h)u) = f((a + tu) + hu).$$

Par la formule de Taylor, il vient

$$g(t + h) = f(a + tu) + \langle \nabla f(a + tu), hu \rangle + \|hu\|\varepsilon(hu) = g(t) + h\langle \nabla f(a + tu), u \rangle + |h| \cdot \|u\|\varepsilon(hu).$$

On en déduit que

$$\frac{g(t + h) - g(t)}{h} = \langle \nabla f(a + tu), u \rangle + \frac{|h|}{h} \|u\|\varepsilon(hu).$$

Mais ε est continue en 0, et $\varepsilon(0) = 0$, donc

$$\lim_{h \rightarrow 0} \frac{|h|}{h} \varepsilon(hu) = 0.$$

Ainsi,

$$\lim_{h \rightarrow 0} \frac{g(t + h) - g(t)}{h} = \langle \nabla f(a + tu), u \rangle.$$

Remarque

$\frac{|h|}{h}$ vaut 1 si $h > 0$ et -1 si $h < 0$.

On en déduit que g est dérivable en t et

$$g'(t) = \langle \nabla f(a + tu), u \rangle = \sum_{i=1}^n \partial_i f(a + tu) u_i.$$

Enfin, puisque $\partial_i f$ est continue, $t \mapsto \partial_i f(a + tu)$ est continue sur \mathbf{R} , et donc g' est continue sur \mathbf{R} . Par conséquent, g est de classe \mathcal{C}^1 sur \mathbf{R} . \square

Remarques. Pour $n = 2$, le graphe de la fonction g est la section du graphe de f par le plan vertical contenant la droite D .

Par exemple, si $a = (0.5, -0.5)$, $u = (-1, 1)$, pour $f(x, y) = x^3 - 4xy^2$, on obtient la figure 9.4.

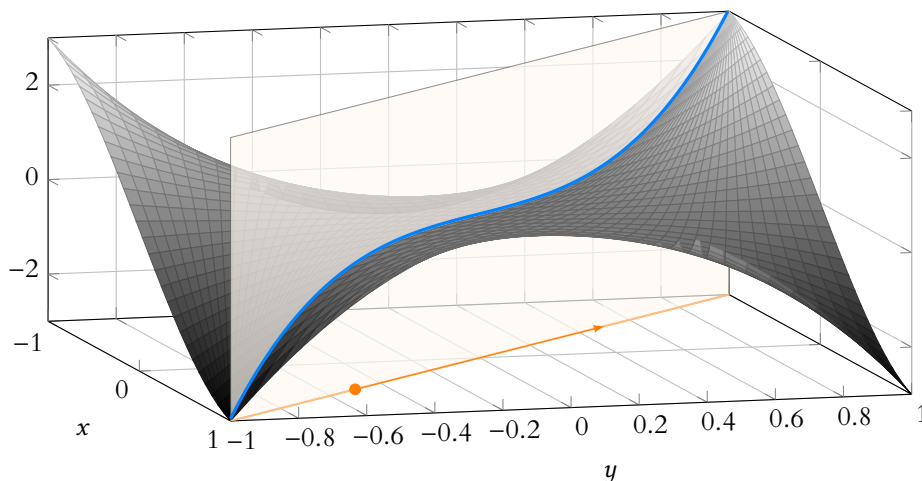


FIGURE 13.4 – En orange la droite D passant par a et de direction u et en bleu, la section du graphe de f par le plan vertical contenant D .

À l'image de ce que nous avons précédemment fait pour les dérivées partielles, il semble légitime de vouloir donner un sens à la dérivée de g en 0, qui semble correspondre à «la dérivée de f en a dans la direction de u ».

L'inconvénient de la formule que nous venons de prouver est que la dérivée de g dépend de la norme de u . En effet, plus la norme de u est grande, et plus on parcourt «rapidement» la droite D , et donc plus les dérivées de g sont «grandes» (au moins en valeur absolue).

Exemple

Si on considère la fonction $f(x, y) = yx$, on a

$$\nabla f(x, y) = (y, x).$$

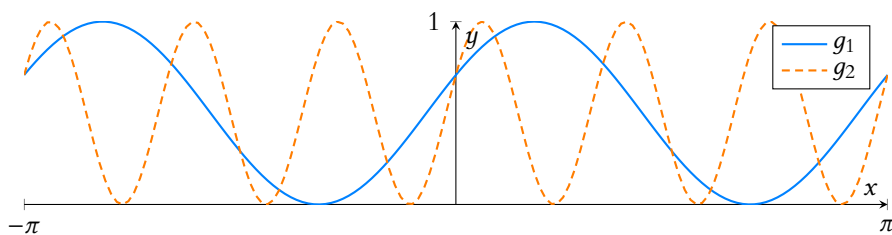
Considérons alors $a = (1, 1)$ et les vecteurs $u_1 = (1, 1)$ et $v_1 = (2, 2)$, qui ont même direction. Avec le premier, on obtient $g'_1(0) = 2$, alors qu'avec le second, $g'_2(0) = 4$.

Exemple 13.28

Prenons encore pour exemple la fonction $f(x, y) = \sin(x) \sin(y)$. Soient alors $u_1 = (1, 1)$ et $u_2 = (3, 3)$, qui ont bien évidemment même direction. Posons $a = (1, 1)$ et soient

$$g_1 : t \mapsto f(a + tu_1) = \sin^2(1 + t) \text{ et } g_2 : t \mapsto f(a + tu_2) = \sin^2(1 + 3t).$$

Les courbes représentatives de g_1 et g_2 sont données par :



Ce dessin peut à première vue sembler surprenant, car ces deux courbes sont obtenues par intersection du graphe de f par le plan vertical passant en $(0, 0)$ et contenant

la droite dirigée par u_1 .

C'est bien le cas, ce qui explique que ces deux courbes soient tout de même d'apparence similaire. La différence est la «vitesse» à laquelle on parcourt ces courbes : on va trois fois «plus vite» pour u_2 que pour u_1 puisque $g_2(t) = g_1(3t)$.

On constate sur le dessin³ que $g'_2(0) \neq g'_1(0)$.

Pourtant, on souhaiterait tout de même que cette dérivée corresponde à «la dérivée de f dans la direction de u_1 ». Comment choisir entre u_1 et u_2 ?

³ mais il est très facile de le voir par le calcul !

⁴ C'est-à-dire de norme 1.

Pour pallier à ce problème, on choisit de prendre pour u un vecteur unitaire⁴.

Définition 13.29 – Soit $f : \mathbf{R}^n \rightarrow \mathbf{R}$, soit $a \in \mathbf{R}^n$, et soit $u \in \mathbf{R}^n$ un vecteur unitaire.

Si la fonction d'une variable réelle $g(t) = f(a + tu)$ admet une dérivée en $t = 0$, on appelle dérivée directionnelle de f en a dans la direction de u le nombre $g'(0)$.

Rappel

Pour toute direction u non nulle, un vecteur unitaire de même direction que u est obtenu via $\frac{u}{\|u\|}$.

Subtilité : la dérivée directionnelle dépend encore du sens de u et pas uniquement de sa direction.

Exemple 13.30

Si $u = e_i = (0, \dots, 0, 1, 0, \dots, 0)$, alors la dérivée directionnelle de f en a suivant la direction u est $\partial_i f(a)$.

En effet, on a alors $g(t) = g(a + tu) = f(a_1, \dots, a_{i-1}, a_i + t, a_{i+1}, \dots, a_n)$, et donc

$$g'(0) = \lim_{h \rightarrow 0} \frac{f(a + he_i) - f(a)}{h} = \lim_{h \rightarrow 0} \frac{f(a_1, \dots, a_{i-1}, a_i + h, a_{i+1}, \dots, a_n) - f(a_1, \dots, a_n)}{h} = \partial_i f(a).$$

Proposition 13.31 : Soit $f : \mathbf{R}^n \rightarrow \mathbf{R}$ de classe \mathcal{C}^1 , et soient $a, u \in \mathbf{R}^n$, avec u unitaire. Alors f admet une dérivée directionnelle en a dans la direction u et celle-ci vaut

$$\langle \nabla f(a), u \rangle = \sum_{i=1}^n \partial_i f(a) u_i.$$

Démonstration. C'est la proposition précédente pour u de norme 1 et $t = 0$. □

Remarque. Soit u un vecteur unitaire. Alors par l'inégalité de Cauchy-Schwarz, on a

$$|\langle \nabla f(a), u \rangle| \leq \|\nabla f(a)\| \cdot \underbrace{\|u\|}_{=1} = \|\nabla f(a)\|.$$

De plus, on a égalité si et seulement u et $\nabla f(a)$ sont colinéaires⁵.

Ainsi, le gradient représente la direction dans laquelle la variation de f est la plus forte au voisinage de a , autrement dit, il indique la direction «de plus grande pente»⁶.

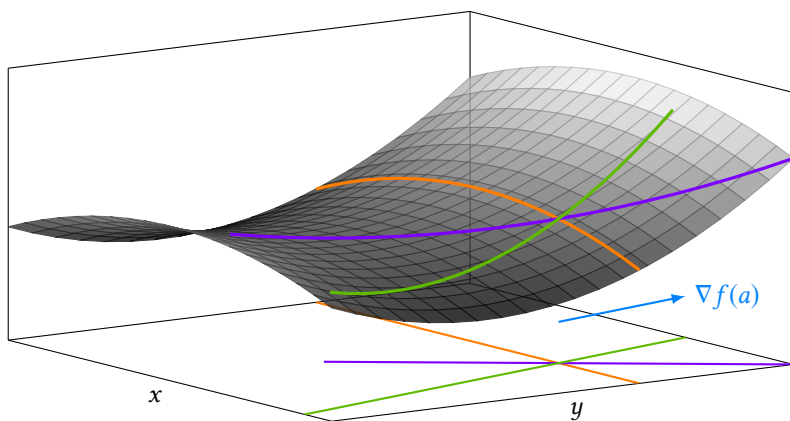
⁵ Notons qu'il n'y a que deux vecteurs unitaires de même direction que $\nabla f(a)$: ce sont $\frac{\nabla f(a)}{\|\nabla f(a)\|}$ et son opposé.

Exemple 13.32

Pour la fonction suivante sont représentées plusieurs courbes dans différentes directions, passant toutes par le même point.

Celle de ces courbes qui possède la pente la plus importante est la verte, qui a même direction que le gradient (représenté en bleu).

⁶ Si on pose une goutte d'eau sur la surface de f au point a , cette goutte se dirigera (par gravité) dans la direction de $\nabla f(a)$.



Remarque : sur cet exemple, on constate que le sens du gradient indique la direction de la «montée», et non de la «descente». Ce fait est toujours vérifié.

13.3.5 Gradient et lignes de niveau

Commençons par un exemple, et étudions la fonction $f(x, y) = 1 - x^2 - y^2$. Elle est \mathcal{C}^1 car polynomiale, et son gradient est

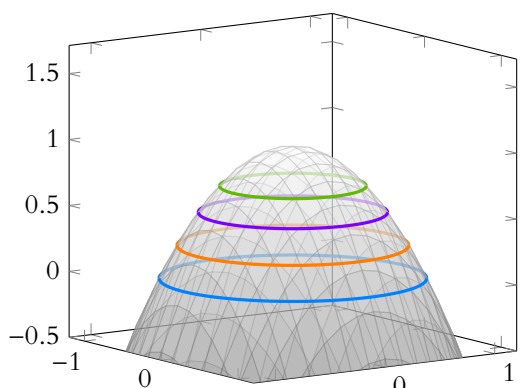
$$\nabla f(x, y) = (-2x, -2y).$$

Pour $\lambda \leq 1$, la ligne de niveau λ est

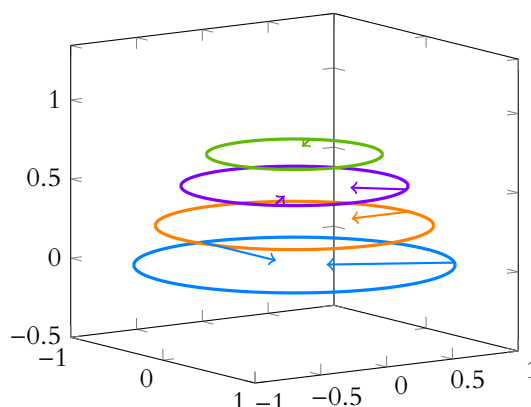
$$\mathcal{C}_\lambda = \{(x, y) \in \mathbf{R}^2 : 1 - x^2 - y^2 = \lambda \Leftrightarrow x^2 + y^2 = 1 - \lambda\}.$$

Ainsi, \mathcal{C}_λ est un cercle de centre $(0, 0)$ et de rayon $\sqrt{1 - \lambda}$.

On prend $\lambda \leq 1$ car f ne prend que des valeurs inférieures ou égales à 1, donc pour $\lambda > 1$, la ligne de niveau λ est vide.



Le graphe de f et quelques lignes de niveau



Quelques lignes de niveau et le gradient en quelques points, vus dans \mathbf{R}^3

La partie de \mathcal{C}_λ correspondant à $y \geq 0$ est alors la courbe représentative de la fonction $f_\lambda : [-\sqrt{1 - \lambda}, \sqrt{1 - \lambda}] \rightarrow \mathbf{R}$ définie par $f_\lambda(x) = \sqrt{1 - \lambda - x^2}$.

Cette fonction est dérivable sur $]-\sqrt{1 - \lambda}, \sqrt{1 - \lambda}[$, de dérivée $f'_\lambda(x) = -\frac{x}{\sqrt{1 - \lambda - x^2}}$.

Et donc, au point $(x, y) \in \mathcal{C}_\lambda$, la tangente à \mathcal{C}_λ est une droite de vecteur directeur

$$u_{(x,y)} = \left(1, -\frac{x}{\sqrt{1 - x^2 - \lambda}}\right) = \left(1, -\frac{x}{y}\right).$$

On a alors

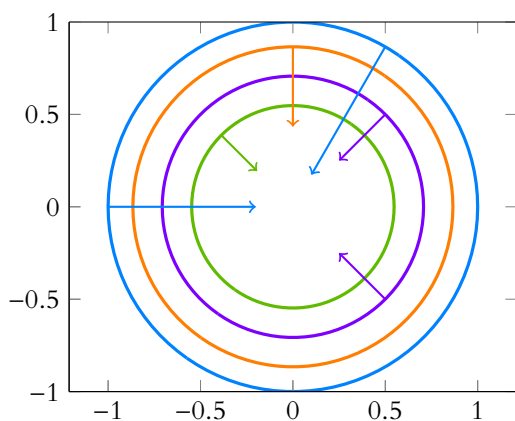
$$\langle u_{(x,y)}, \nabla f(x, y) \rangle = \left\langle \left(1, -\frac{x}{y}\right), (-2x, -2y) \right\rangle = -2x + 2y \frac{x}{y} = 0.$$

Et donc la tangente à \mathcal{C}_λ est orthogonale à $\nabla f(x, y)$.

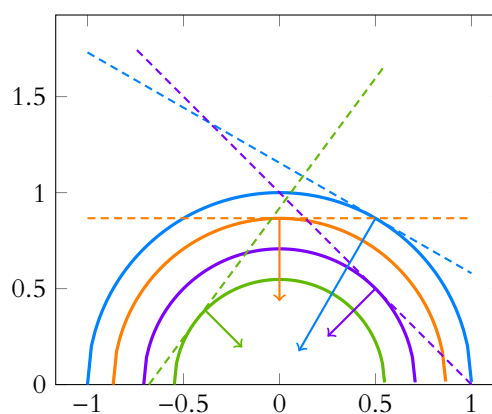
Explication

Pour $y \geq 0$, on a

$$x^2 + y^2 = 1 - \lambda \Leftrightarrow y = \sqrt{1 - \lambda - x^2}.$$



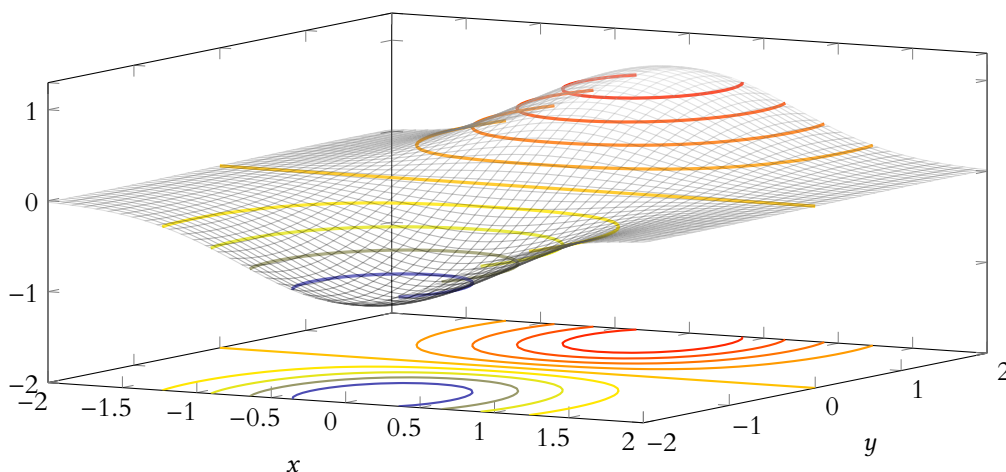
Lignes de niveau et gradients
vus dans \mathbf{R}^2 .



Les tangentes aux lignes de niveau sont
orthogonales au gradient.

Cette propriété est en fait toujours vérifiée : on dit généralement que «les courbes de niveau sont orthogonales⁷ au gradient».

Par exemple, le graphique ci-dessous représente quelques lignes de niveau pour la fonction $f(x, y) = e^{-x^2} \sin(y)$. On «voit» alors que ces lignes de niveau sont orthogonales à la direction de plus grande pente, et donc au gradient.



⁷ Les lignes de niveau étant des courbes (et non des droites), cette terminologie n'est pas totalement correcte. Il faudrait plutôt dire que la tangente à la courbe de niveau est orthogonale au gradient.

13.4 EXTREMA GLOBAUX

13.4.1 Définition

Définition 13.33 – Soit $f : \mathbf{R}^n \rightarrow \mathbf{R}$. On dit que f admet un maximum (resp. un minimum) en $a \in \mathbf{R}^n$ si

$$\forall x \in \mathbf{R}^n, f(x) \leq f(a) \text{ (resp. } f(x) \geq f(a)).$$

Dans ce cas, on dit que $f(a)$ est le maximum (resp. le minimum) de f sur \mathbf{R}^n .

Terminologie

Comme pour les fonctions d'une variable, ce qu'on appelle maximum/minimum est la valeur prise par la fonction, et non le/les point(s) où elle atteint cette valeur.

Exemple 13.34

- La fonction $f(x, y) = x^2 + y^4$ admet un minimum en $(0, 0)$. En effet, on a toujours $f(x, y) \geq 0$ et $f(0, 0) = 0$. En revanche, elle n'admet pas de maximum car pour x fixé, $\lim_{y \rightarrow +\infty} f(x, y) = +\infty$.
- La fonction $g(x, y) = \sin(x) \sin(y)$ a un maximum égal à 1 et un minimum égal à -1 .

Explication

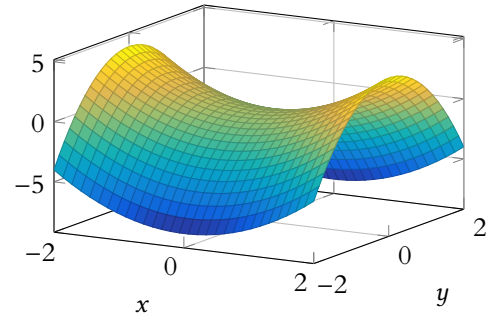
Puisque cette limite est infinie, f peut prendre des valeurs arbitrairement grandes, et donc n'a pas de maximum.

De plus, le maximum est atteint si et seulement si $\sin(x)$ et $\sin(y)$ sont simultanément égaux à 1 ou à -1 . Il est donc atteint une infinité de fois, en les

$$\left(\frac{\pi}{2} + k\pi, \frac{\pi}{2} + \ell\pi\right), (k, \ell) \in \mathbf{N}^2, k \text{ et } \ell \text{ de même parité.}$$

De même, le minimum est atteint une infinité de fois.

- La fonction $h(x, y) = x^2 - 2y^2$ n'admet ni maximum ni minimum. En effet, $\lim_{x \rightarrow +\infty} f(x, 0) = +\infty$ et $\lim_{y \rightarrow +\infty} f(0, y) = -\infty$.



13.4.2 Condition nécessaire

Théorème 13.35 : Soit $f : \mathbf{R}^n \rightarrow \mathbf{R}$ une fonction de classe \mathcal{C}^1 . Si f admet un maximum (resp. un minimum) en $a \in \mathbf{R}^n$, alors $\nabla f(a) = 0$.
De manière équivalente,

$$\forall i \in \llbracket 1, n \rrbracket, \partial_i f(a) = 0.$$

Démonstration. Supposons que f admette un maximum⁸ en a . Soit e_i le $i^{\text{ème}}$ vecteur de la base canonique de \mathbf{R}^n . Notons alors f_i la fonction d'une variable $f_i : t \mapsto f(a + te_i)$. Alors f_i est dérivable, de dérivée $f'_i(t) = \partial_i f(a + te_i)$. Pour tout $t \in \mathbf{R}$, on a

$$f_i(t) = f(a_1, \dots, a_{i-1}, a_i + t, a_{i+1}, \dots, a_n) \leq f(a_1, \dots, a_n) = f_i(0).$$

Donc f_i admet un extremum en $t = 0$, et par conséquent $f'_i(0) = \partial_i f(a) = 0$. Ceci étant vrai pour tout $i \in \llbracket 1, n \rrbracket$, on en déduit que $\nabla f(a) = 0$. □

⁸ La preuve est la même si f admet un minimum, en renversant le sens des inégalités.

Détail

On utilise ici un résultat bien connu sur les fonctions d'une variable.

Définition 13.36 – Un point $a \in \mathbf{R}^n$ tel que $\nabla f(a) = 0$ est appelé point critique de f .

Remarques. • La proposition précédente dit qu'un extremum d'une fonction \mathcal{C}^1 est nécessairement atteint en un point critique.

- On généralise ainsi le résultat connu pour $n = 1$: si f admet un extremum en a , alors $f'(a) = 0$.
- \triangleleft Comme⁹ pour $n = 1$, il s'agit d'une condition nécessaire, mais pas d'une condition suffisante : a peut être un point critique de f sans que pour autant f atteigne un extremum en a .

Par exemple $f(x, y) = x^2 - 2y^2$ admet un point critique en $a = (0, 0)$. En effet, $\partial_1 f(x, y) = 2x$ et $\partial_2 f(x, y) = 4y$. Donc (x, y) est un point critique de f si et seulement si

$$\begin{cases} 2x = 0 \\ 4y = 0 \end{cases} \Leftrightarrow (x, y) = (0, 0).$$

Mais $f(0, 0) = 0$ et $f(1, 0) = 1$ et $f(0, 1) = -2$. Donc $f(0, 0)$ n'est ni un maximum ni un minimum de f .

- Pour $n = 2$, a est un point critique de f si et seulement si le plan tangent de f en a a pour équation $z = f(a)$, c'est-à-dire si et seulement si c'est un plan horizontal.

Nous ne disposons pas d'outils généraux plus élaborés pour déterminer si un point critique est ou non un extremum de f .

La détermination de la nature d'un point critique est en général un problème difficile, pour lequel il existe de nombreuses¹⁰ techniques différentes, mais aucune qui puisse fonctionner pour toutes les fonctions, il faudra donc être capable de s'adapter et se montrer réactif aux indications que pourrait fournir l'énoncé d'un exercice.

⁹ Par exemple, la fonction $t \mapsto t^3$ admet un point critique en 0 qui n'est ni un maximum, ni un minimum.

¹⁰ Certaines d'entre elles seront étudiées en TD.

Exemple 13.37

Reprenons l'exemple donné dans l'introduction du chapitre, où

$$G(x, y) = -6x^2 - 4y^2 - 4xy + 120x + 140y - 1000.$$

Alors G est \mathcal{C}^1 car polynomiale, et ses dérivées partielles sont données par $\partial_1 G(x, y) = -12x - 4y + 120$ et $\partial_2 G(x, y) = -8y - 4x + 140$.

Ainsi, (x, y) est un point critique de G si et seulement si

$$\begin{cases} -12x - 4y + 120 = 0 \\ -8y - 4x + 140 = 0 \end{cases} \Leftrightarrow \begin{cases} x = 5 \\ y = 15 \end{cases}$$

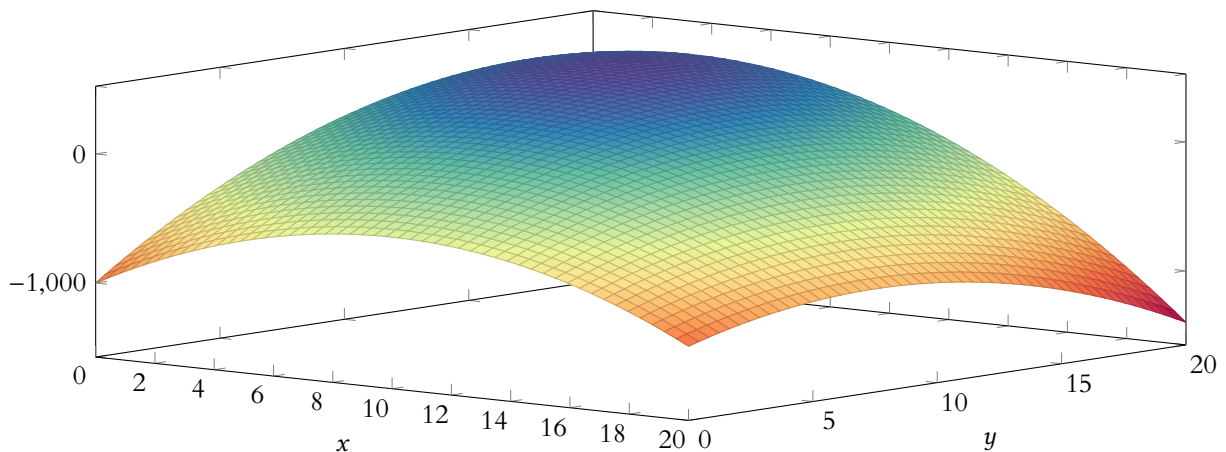
Pour déterminer si cet unique point critique correspond à un extremum, essayons de déterminer le signe de $G(x, y) - G(5, 15)$: alors G admet un extremum¹¹ en $(5, 15)$ si et seulement ce signe est constant¹².

On a $G(5, 15) = 350$ et donc

$$\begin{aligned} G(x, y) - G(5, 15) &= -6x^2 - 4y^2 - 4xy + 120x + 140y - 1000 - 350 \\ &= -6\left(x^2 + \frac{2}{3}xy - 20x\right) - 4y^2 + 140y - 1350 \\ &= -6\left(x^2 + 2x\left(\frac{1}{3}y - 10\right)\right) - 4y^2 + 140y - 1350 \\ &= -6\left[\left(x + \frac{1}{3}y - 10\right)^2 - \left(\frac{1}{3}y - 10\right)^2\right] - 4y^2 + 140y - 1350 \\ &= -6\left(x + \frac{1}{3}y - 10\right)^2 - \frac{10}{3}y^2 + 100y - 750 \\ &= -6\left(x + \frac{1}{3}y - 10\right)^2 - \frac{10}{3}(y^2 - 30y) - 750 \\ &= -6\left(x + \frac{1}{3}y - 10\right)^2 - \frac{10}{3}\left((y - 15)^2 - 15^2\right) - 750 \\ &= -6\left(x + \frac{1}{3}y - 10\right)^2 - \frac{10}{3}(y - 15)^2. \end{aligned}$$

Et donc, pour tout $(x, y) \in \mathbf{R}^2$, $G(x, y) - G(5, 15) \leq 0 \Leftrightarrow G(x, y) \leq G(5, 15)$: G admet un maximum en $(5, 15)$.

Et donc, pour revenir à notre problème initial, les montres doivent être vendues 50 000 euros en France, 150 000 en Suisse.



¹¹ Plus précisément, si $G(x, y) - G(5, 15)$ est positif pour tout $(x, y) \in \mathbf{R}^2$, alors on est en présence d'un minimum, et si cette différence est négative pour tout $(x, y) \in \mathbf{R}^2$, alors on est en présence d'un maximum. Et si jamais le signe de $G(x, y) - G(5, 15)$ n'est pas constant, G n'admet ni maximum, ni minimum.

¹² C'est-à-dire est indépendant de x et y .

On a isolé tous les termes avec des x .

Méthode

On a cherché à regrouper tous les termes en x au sein d'une expression de la forme $(x + a)^2$. Or nous savons que

$$(x + a)^2 = x^2 + 2ax + a^2.$$

Pour bien faire apparaître les termes en x^2 et en x , il faut donc prendre $a = \frac{1}{3}y - 10$. Il ne reste alors qu'à retrancher le terme a^2 qui apparaît et qui ne figurait pas dans notre expression de départ.

Exemple 13.38

Soit $f(x, y) = x^4 + y^4 - 2(x - y)^2$.
 Alors f est de classe \mathcal{C}^1 sur \mathbf{R}^2 car il s'agit d'une fonction polynomiale.
 De plus, on a

$$\partial_1 f(x, y) = 4x^3 - 4(x - y) \text{ et } \partial_2 f(x, y) = 4y^3 + 4(x - y).$$

Ainsi, (x, y) est un point critique de f si et seulement si

$$\begin{cases} 4x^3 - 4(x - y) = 0 \\ 4y^3 + 4(x - y) = 0 \end{cases} \Leftrightarrow \begin{cases} x^3 - x + y = 0 \\ y^3 + x - y = 0 \end{cases} \Leftrightarrow \begin{cases} x^3 + y^3 = 0 \\ x^3 - x + y = 0 \end{cases} \Leftrightarrow \begin{cases} x = -y \\ x^3 - 2x = 0 \end{cases}$$

Or, les solutions de $x^3 - 2x = 0$ sont $\{0, -\sqrt{2}, \sqrt{2}\}$.
 f admet donc trois points critiques qui sont $(0, 0)$, $(-\sqrt{2}, \sqrt{2})$ et $(\sqrt{2}, -\sqrt{2})$.

$$f(0, 0) = 0 \text{ et } f(-\sqrt{2}, \sqrt{2}) = f(\sqrt{2}, -\sqrt{2}) = -8.$$

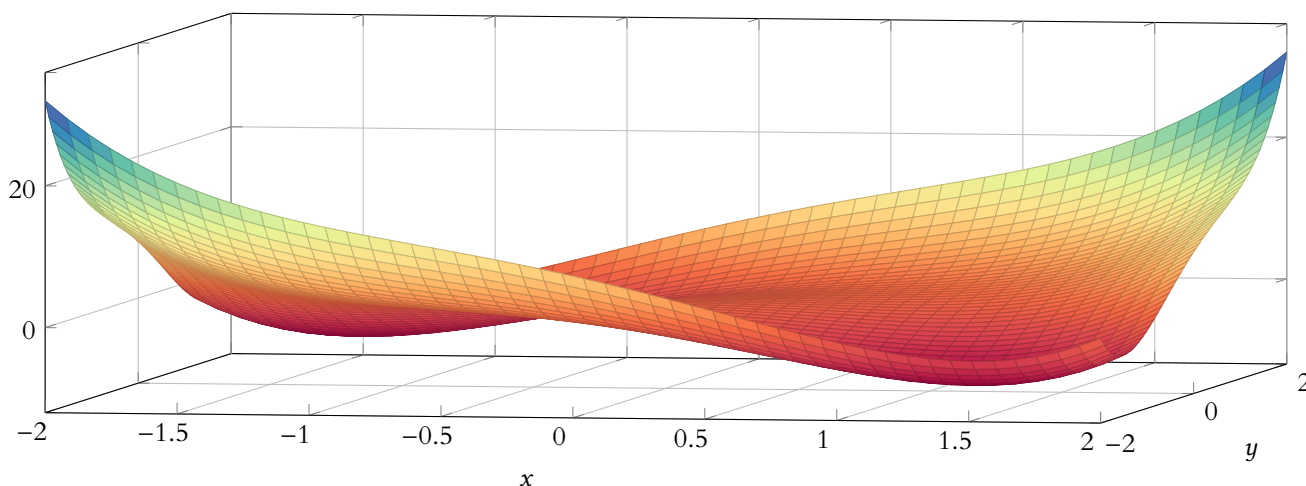
Éliminons tout de suite $(0, 0)$ qui n'est pas un minimum, et qui n'est pas non plus un maximum car $f(2, 0) = 8 > 0$.

De plus, pour tout $(x, y) \in \mathbf{R}^2$, on a

$$f(x, y) - f(\sqrt{2}, -\sqrt{2}) = x^4 + y^4 - 2x^2 + 4xy - 2y^2 + 8 = (x^2 - 2)^2 + (y^2 - 2)^2 + 2(x + y)^2 \geq 0.$$

Ainsi, $\forall (x, y) \in \mathbf{R}^2, f(x, y) \geq f(-\sqrt{2}, \sqrt{2})$, donc f atteint bien un minimum en $(-\sqrt{2}, \sqrt{2})$ et de même en $(\sqrt{2}, -\sqrt{2})$. De plus, ce minimum vaut -8 .

⚠ Danger !
 À ce stade là, nous ne savons aucunement si ces points critiques correspondent à des maxima, des minima... ou rien du tout. Le théorème précédent stipule qu'un extremum est atteint en un point critique, mais il n'y a pas de réciproque. Tout le travail reste donc à faire.



Exemple 13.39 D'après EDHEC 2005

Soit $f : \mathbf{R}^n \rightarrow \mathbf{R}$ définie par

$$\forall (x_1, \dots, x_n) \in \mathbf{R}^n, f(x_1, \dots, x_n) = \sum_{k=1}^n x_k^2 + \left(\sum_{k=1}^n x_k \right)^2 - \sum_{k=1}^n x_k.$$

Alors f est de classe \mathcal{C}^1 sur \mathbf{R}^n car il s'agit d'une fonction polynomiale.
 De plus, $\forall i \in \llbracket 1, n \rrbracket$,

$$\partial_i f(x_1, \dots, x_n) = 2x_i + 2 \left(\sum_{k=1}^n x_k \right) - 1.$$

Donc (x_1, \dots, x_n) est un point critique de f si et seulement si

$$\begin{cases} 2x_1 + 2\left(\sum_{k=1}^n x_k\right) = 1 \\ \vdots \\ 2x_n + 2\left(\sum_{k=1}^n x_k\right) = 1 \end{cases}$$

En réalisant l'opération élémentaire $L_i - L_1, \forall i \in \llbracket 2, n \rrbracket$, on obtient

$$\begin{cases} x_1 = \frac{1}{2} - \sum_{k=1}^n x_k \\ 2x_2 - 2x_1 = 0 \\ \vdots \\ 2x_n - 2x_1 = 0 \end{cases} \Leftrightarrow \begin{cases} x_1 = \frac{1}{2} - \sum_{k=1}^n x_k \\ x_2 = x_1 \\ \vdots \\ x_n = x_1 \end{cases} \Leftrightarrow \begin{cases} x_1 = \frac{1}{2} - nx_1 \\ x_2 = \dots = x_n = x_1 \end{cases}$$

Au final, f possède un unique point critique $\left(\frac{1}{2n+2}, \dots, \frac{1}{2n+2}\right)$.

Soit $(x_1, \dots, x_n) \in \mathbf{R}^n$ et notons $(h_1, \dots, h_n) = \left(x_1 - \frac{1}{2(n+1)}, \dots, x_n - \frac{1}{2(n+1)}\right)$.

Alors

$$\begin{aligned} f(x_1, \dots, x_n) - f\left(\frac{1}{2n+2}, \dots, \frac{1}{2n+2}\right) &= f\left(\frac{1}{2(n+1)} + h_1, \dots, \frac{1}{2(n+1)} + h_n\right) - f\left(\frac{1}{2n+2}, \dots, \frac{1}{2n+2}\right) \\ &= \sum_{i=1}^n \left(\frac{1}{2(n+1)} + h_i\right)^2 + \left(\sum_{i=1}^n \frac{1}{2(n+1)} + h_i\right)^2 - \sum_{i=1}^n \left(\frac{1}{2(n+1)} + h_i\right) - \sum_{i=1}^n \left(\frac{1}{2(n+1)}\right)^2 - \left(\sum_{i=1}^n \frac{1}{2(n+1)}\right)^2 + \sum_{i=1}^n \frac{1}{2(n+1)} \\ &= \sum_{i=1}^n \left(\frac{1}{2(n+1)}\right)^2 + \sum_{i=1}^n \frac{h_i}{n+1} + \sum_{i=1}^n h_i^2 + \left(\frac{n}{2(n+1)} + \sum_{i=1}^n h_i\right)^2 - \frac{n}{2(n+1)} - \sum_{i=1}^n h_i - \frac{n}{(2(n+1))^2} - \left(\frac{n}{2(n+1)}\right)^2 + \frac{n}{2(n+1)} \\ &= \frac{n}{(2(n+1))^2} + \frac{1}{n+1} \sum_{i=1}^n h_i + \sum_{i=1}^n h_i^2 + \left(\frac{n}{2(n+1)} + \sum_{i=1}^n h_i\right)^2 + \frac{n}{n+1} \left(\sum_{i=1}^n h_i\right) + \left(\sum_{i=1}^n h_i\right)^2 - \sum_{i=1}^n h_i - \frac{n}{(2(n+1))^2} - \left(\frac{n}{2(n+1)}\right)^2 \\ &= \sum_{i=1}^n h_i^2 + \left(\sum_{i=1}^n h_i\right)^2 \geq 0. \end{aligned}$$

Et donc f admet un minimum en son unique point critique.

MÉTHODES DU CHAPITRE 13

✓ MÉTHODE 13.1 : Montrer qu'une fonction f est continue ou de classe \mathcal{C}^1

On se ramènera toujours aux théorèmes sur les opérations élémentaires et aux fonctions polynomiales.

On prendra bien garde à expliquer quelles sont les fonctions de base que l'on utilise et quelles sont les opérations effectuées (somme, produit, composition par une fonction d'une variable)

Exercices de TD : 13.2, 13.2, 13.3, 13.9, 13.10, 13.11

✓ MÉTHODE 13.2 : Calculer les dérivées partielles ou le gradient d'une fonction de classe \mathcal{C}^1

Pour calculer la i -ème dérivée partielle, il suffit de dériver par rapport à x_i , en considérant que les autres variables sont fixées. Toutes les formules de dérivation vues en première année sont alors valables.

Exercices de TD : 13.2, 13.2, 13.3, 13.9, 13.10, 13.11, 13.18

✓ MÉTHODE 13.3 : Déterminer les points critiques d'une fonction f

Par définition, un point critique est un point où le gradient s'annule, c'est-à-dire où toutes les dérivées s'annulent. On a donc un système de n équations à n inconnues.

Sauf si ce système est linéaire (auquel cas on dispose de la méthode du pivot), il n'existe pas de méthode générale pour le résoudre.

On cherchera tout de même à procéder par substitution, c'est-à-dire à exprimer toutes les variables à l'aide de l'une d'entre elles et on essaiera de faire apparaître une équation en cette seule variable.

Afin d'éviter toute erreur de logique, il est conseillé de ne pas procéder par équivalence mais uniquement par implication, et de vérifier à la fin de la résolution que les n -uplets obtenus sont bien des points critiques.

Exercices de TD : 13.7, 13.9, 13.10, 13.11, 13.12, 13.13, 13.14, 13.15, 13.16, 13.17, 13.18

✓ MÉTHODE 13.4 : Montrer qu'une fonction n'admet pas d'extremum

- On peut par exemple trouver des limites infinies pour montrer que f prend des valeurs arbitrairement grandes (resp. arbitrairement petites). Dans ce cas, f n'admet pas de maximum (resp. de minimum). Les limites que nous connaissons n'ont de sens que quand on fait varier une seule variable. Il faut donc fixer (intelligemment) toutes les variables sauf une, que l'on fait ensuite tendre vers $\pm\infty$.

- Déterminer les points critiques, qui sont les seuls points où f peut avoir un extremum.

Si a est un point critique, on peut alors calculer la valeur de a et essayer de trouver des points où f prend des valeurs strictement supérieures (resp. inférieures) à $f(a)$ et alors f pas un maximum (resp. un minimum).

Exercices de TD : 13.7, 13.12, 13.15

✓ MÉTHODE 13.5 : Montrer qu'une fonction f admet un maximum/un minimum

On sait que les extrema sont nécessairement atteints en des points critiques de f , donc sauf indication contraire, on commence par déterminer ces points critiques, et à calculer les valeurs qu'y prend f .

Cela permet parfois d'éviter d'avoir à étudier certains de ces points critiques, prouvant directement qu'ils ne correspondent ni à un maximum ni à un minimum.

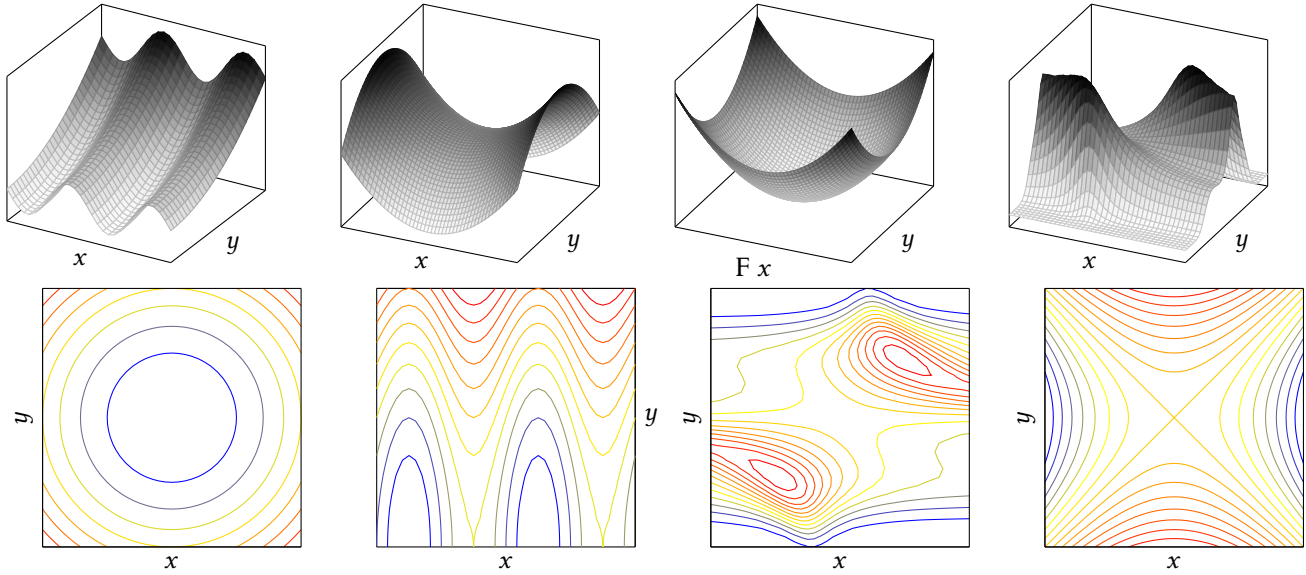
Pour le reste, il n'existe pas de méthode «clés en mains» pour déterminer si un point critique a correspond à un extrema. De manière générale, il faut montrer que si $f(x) - f(a)$ est de signe constant pour tout $x \in \mathbf{R}^n$. On peut pour cela essayer de faire apparaître des carrés. Un exercice ou un sujet de concours devrait normalement fournir suffisamment d'indications pour conduire ce type de calculs.

Exercices de TD : 13.8, 13.9, 13.10, 13.11, 13.16, 13.17

EXERCICES DU CHAPITRE 13

EXERCICE 13.1 Lecture graphique

Pour chacune des fonctions tracées ci-dessous, lui associer ses lignes de niveau.



EXERCICE 13.2 Pour chacune des fonctions suivantes, montrer qu'elle est \mathcal{C}^1 sur son ensemble de définition, et calculer son gradient.

- $f : \mathbf{R}^2 \rightarrow \mathbf{R}, (x, y) \mapsto xye^{-x^2+2y}$
- $g : \mathbf{R}^3 \rightarrow \mathbf{R}, (x, y, z) \mapsto xye^z + xze^y + yxe^x$
- $h : \mathbf{R}^n \rightarrow \mathbf{R}, (x_1, \dots, x_n) \mapsto (x_1^2 + \dots + x_n^2)e^{-(x_1^2 + \dots + x_n^2)}$.

EXERCICE 13.3 Soit f une fonction continue sur \mathbf{R} et soit $g : \mathbf{R}^2 \rightarrow \mathbf{R}$ définie par $g(x, y) = \int_x^{xy} f(t) dt$.
Montrer que g est de classe \mathcal{C}^1 sur \mathbf{R} et déterminer ses dérivées partielles.

EXERCICE 13.4 Soit $f : \mathbf{R}^3 \rightarrow \mathbf{R}$ définie par $f(x, y, z) = \begin{cases} \frac{xyz}{\sqrt{x^2 + y^2 + z^2}} & \text{si } (x, y, z) \neq (0, 0, 0) \\ 0 & \text{sinon} \end{cases}$

- Montrer que $\forall (x, y, z) \in \mathbf{R}^3, |f(x, y, z)| \leq |yz|$.
- En déduire que f est continue sur \mathbf{R}^3 .

EXERCICE 13.5 Soit f la fonction définie sur \mathbf{R}^2 par $f(x, y) = x^2 + 2xy + y^2 + 2x + 3y - 1$.

- Déterminer la dérivée directionnelle de f en $a = (-1, 1)$ dans la direction $(\frac{4}{5}, \frac{-3}{5})$.
- Même question dans la direction $\frac{\nabla f(a)}{\|\nabla f(a)\|}$.

EXERCICE 13.6 Soit $f(x, y) = (y - 2x^2)^3 + 1$.

- Montrer que les lignes de niveau de f sont des paraboles.
- Soit $\lambda \in \mathbf{R}$ et (x_0, y_0) un point de \mathcal{C}_λ , la courbe de niveau λ de f . Déterminer un vecteur directeur de la tangente à \mathcal{C}_λ en (x_0, y_0) .
Montrer que ce vecteur est orthogonal à $\nabla f(x_0, y_0)$.

EXERCICE 13.7 Soit f la fonction définie sur \mathbf{R}^3 par $f(x, y, z) = \frac{x^2}{2} + xyz + y - z$. Déterminer les points critiques de f et déterminer ses éventuels extrema.

EXERCICE 13.8 Étude d'un extremum par variation de fonctions

Soit f la fonction définie sur \mathbf{R}^2 par $f(x, y) = x^4 + y^4 - 4xy$.

- Montrer que f n'admet pas de maximum.
- On se propose de montrer que f possède un minimum.
 - En considérant $f(-x, -y)$, montrer qu'on peut se restreindre à $y \geq 0$.

- b) Pour $y \geq 0$ fixé, montrer que la fonction $x \mapsto f(x, y)$ admet un minimum noté $g(y)$.
- c) Étudier les variations de $y \mapsto g(y)$ et en déduire que f admet un minimum, et préciser le(s) point(s) où ce minimum est atteint.

EXERCICE 13.9 Soit f la fonction définie sur \mathbf{R}^2 par $f(x, y) = \frac{-3y}{x^2 + y^2 + 1}$.

PD

- 1) Montrer que f est de classe \mathcal{C}^1 sur \mathbf{R}^2 et calculer ses dérivées partielles.
- 2) Déterminer les points critiques de f .
- 3) Déterminer les extrema de f .

EXERCICE 13.10 (Ecricome ECE 2006)

PD

Soit $g : \mathbf{R}^2 \rightarrow \mathbf{R}$ la fonction définie par $g(x, y) = e^x(x + y^2 + e^x)$. On considère également la fonction f , définie sur \mathbf{R} par $f(t) = 1 + t + 2e^t$.

- 1) Montrer que g est \mathcal{C}^1 sur \mathbf{R}^2 , et calculer ses dérivées partielles.
- 2) Montrer qu'il existe un unique réel α tel que $f(\alpha) = 0$. Justifier que $\alpha \in [-2, -1]$.
- 3) En déduire que g possède un unique point critique, que l'on exprimera en fonction de α .
- 4) En remarquant que pour tout $y \in \mathbf{R}$, $g(x, y) \geq g(x, 0)$, montrer que g possède un minimum global.

EXERCICE 13.11 Utilisation de polynômes de degré deux

AD

Soit $f(x, y) = x^4 + 2x^2 + y^2 - 4xy + 8x - 4y + 2$.

- 1) Montrer que f est \mathcal{C}^1 et calculer ses dérivées partielles.
- 2) Montrer que f admet exactement trois points critiques : $(0, 2)$, et deux autres points notés α et β . Justifier que f n'admet pas d'extremum en $(0, 2)$ et que $f(\alpha) = f(\beta)$.
- 3) On souhaite étudier la nature des deux points critiques. Pour cela, on cherche à déterminer le signe de $g(x, y) = f(x, y) - f(\alpha)$.
 - a) Soit $x \in \mathbf{R}$ fixé. Montrer que $y \mapsto g(x, y)$ est un polynôme de degré 2, et calculer son discriminant, noté $\Delta(x)$.
 - b) Montrer que la fonction Δ est de signe constant.
 - c) Conclure.

EXERCICE 13.12

AD

- 1) Montrer que l'équation $x + e^{x-\frac{1}{x}} = 0$ admet -1 comme unique solution sur \mathbf{R}^* .
- 2) On pose dans la suite $f(x, y) = xe^y + ye^x$.
 - a) Montrer que f possède un unique point critique (a, b) que l'on déterminera.
 - b) f atteint-elle un extremum en ce point critique ?

EXERCICE 13.13 Montrer que la fonction définie sur \mathbf{R}^2 par $f(x, y) = \sin x + y^2 - 2y + 1$ est \mathcal{C}^1 et qu'elle possède une infinité de points critiques. Ces points critiques correspondent-ils à des extrema de f ?

AD

EXERCICE 13.14 Soit $f : \mathbf{R}^3 \rightarrow \mathbf{R}$ définie par $f(x, y, z) = xyz + xy + yz + xz$.

PD

Montrer que f est de classe \mathcal{C}^1 et déterminer ses points critiques. Correspondent-ils à des extrema de f ?

EXERCICE 13.15 Déterminer si les fonctions suivantes admettent ou non des extrema et déterminer l'ensemble des points où ces extrema sont atteints.

AD

1. $f : \mathbf{R}^2 \rightarrow \mathbf{R}$, $f(x, y) = x^4 - 4x^2y^2 + 4y^4$.
2. $g : \mathbf{R}^2 \rightarrow \mathbf{R}$, $g(x, y) = x^2 + xy + y^2$.
3. $h : \mathbf{R}^3 \rightarrow \mathbf{R}$, $h(x, y, z) = \frac{x^2}{2} + y^2 + 2z^2 - xy - xz + x^4$.
4. $i : \mathbf{R}^2 \rightarrow \mathbf{R}$, $i(x, y) = -2x^2 - 2xy - 3y^2 - 4$.
5. $j : \mathbf{R}^3 \rightarrow \mathbf{R}$, $j(x, y, z) = x^2 + y^2 + z^2 - 2xyz$.

EXERCICE 13.16 Soit f la fonction définie sur \mathbf{R}^3 par $f(x, y, z) = (x + y + z)e^{-\frac{1}{6}(x^2 + y^2 + z^2)}$.

AD

- 1) Montrer que f est \mathcal{C}^1 sur \mathbf{R}^3 et déterminer ses points critiques.
- 2) Montrer que pour tout $(x, y, z) \in \mathbf{R}^3$, $|x + y + z| \leq \sqrt{3}\sqrt{x^2 + y^2 + z^2}$.
- 3) En étudiant la fonction définie sur \mathbf{R}^+ par $g(t) = \sqrt{t}e^{-t/6}$, déterminer la nature des points critiques de f .

EXERCICE 13.17 Soit $f : \mathbf{R}^3 \rightarrow \mathbf{R}$ définie par $f(x, y, z) = (x + z^2)e^{x(y^2 + z^2 + 1)}$.

D

- 1) Montrer que f est de classe \mathcal{C}^1 sur \mathbf{R}^3 .
- 2) Montrer que f possède un unique point critique $A \in \mathbf{R}^3$ que l'on déterminera. Calculer $f(A)$.
- 3) Montrer que si $x \geq 0$, alors $f(x, y, z) \geq 0$, et que si $x < 0$, alors $f(x, y, z) \geq xe^x$.
- 4) Déterminer le minimum de la fonction $x \mapsto xe^x$, et en déduire que f atteint son minimum en A .

EXERCICE 13.18 Soit $n \geq 1$ et soit f la fonction définie sur \mathbf{R}^n par

$$f(x_1, \dots, x_n) = \sum_{k=1}^n x_k^2 + \left(1 - \sum_{k=1}^n x_k\right)^2.$$

- 1) Montrer que f est de classe \mathcal{C}^1 sur \mathbf{R}^n et qu'elle possède un unique point critique $a = (a_1, \dots, a_n)$.
- 2) Pour $(h_1, \dots, h_n) \in \mathbf{R}^n$, expliciter $f(a_1 + h_1, \dots, a_n + h_n) - f(a_1, \dots, a_n)$, et en déduire la nature du point critique a .

EXERCICE 13.19 Fonctions convexes sur \mathbf{R}^n (Oral ESCP 2001)

Soit $n \in \mathbf{N}$ tel que $n \geq 2$. L'espace vectoriel \mathbf{R}^n est muni de sa structure euclidienne canonique, le produit scalaire étant noté $\langle \cdot, \cdot \rangle$ et la norme associée $\|\cdot\|$.

Soit f une application de classe \mathcal{C}^1 définie sur \mathbf{R}^n , à valeurs dans \mathbf{R} , convexe, c'est-à-dire vérifiant pour tout $(x, y) \in (\mathbf{R}^n)^2$ et pour tout réel $\lambda \in [0, 1]$:

$$f((1 - \lambda)x + \lambda y) \leq (1 - \lambda)f(x) + \lambda f(y).$$

Pour tout $(h, x) \in (\mathbf{R}^n)^2$ fixé, on définit la fonction $\varphi_{h,x}$ de la variable réelle t par :

$$\varphi_{h,x}(t) = f(x + th).$$

- 1) a) Montrer que $\varphi_{h,x}$ est une fonction convexe de \mathbf{R} dans \mathbf{R} .
b) En déduire que $\varphi'_{h,x}(0) \leq \varphi_{h,x}(1) - \varphi_{h,x}(0)$.
- 2) Montrer que pour tout $(x, y) \in (\mathbf{R}^n)^2$, on a

$$\langle \nabla f(x), y - x \rangle \leq f(y) - f(x).$$

En déduire que si x est un point critique de f , alors f possède un minimum, atteint en x .

- 3) On suppose dans cette question que $f(0) = 0$ et que $\nabla f(0) = 0$. On suppose également que f est strictement convexe, c'est-à-dire qu'elle vérifie pour tout $(x, y) \in (\mathbf{R}^n)^2$, tels que $x \neq y$, pour tout réel $\lambda \in]0, 1[$:

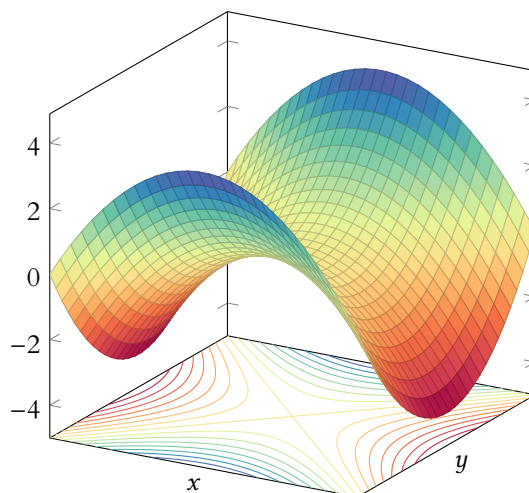
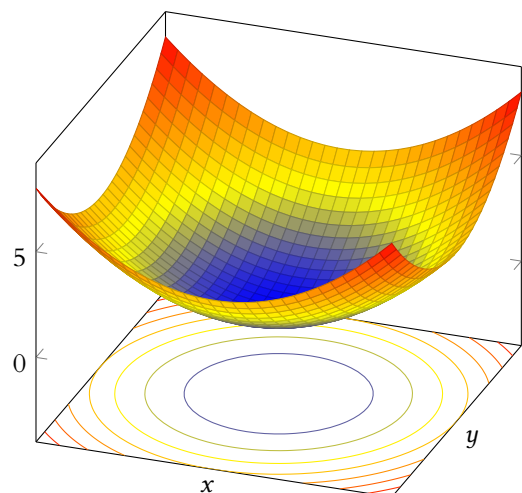
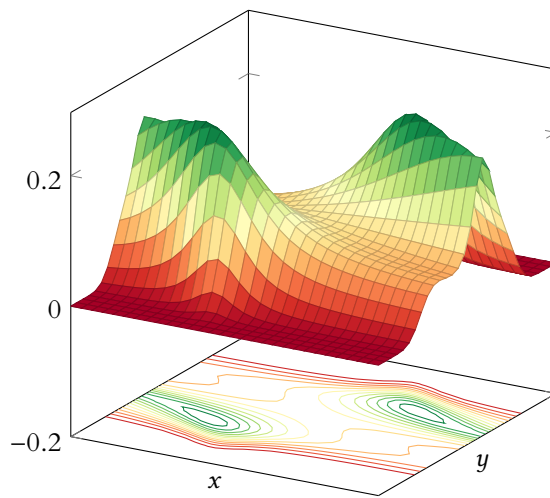
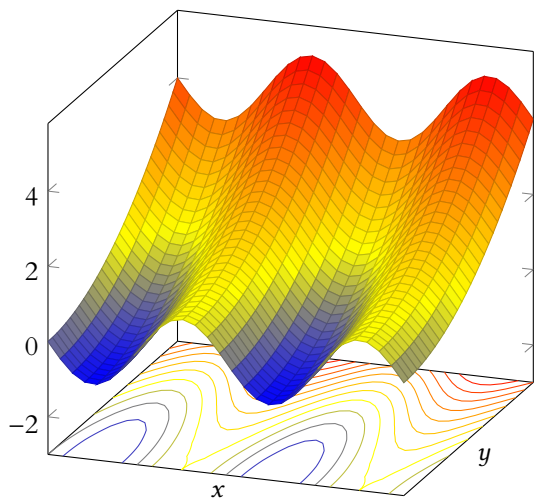
$$f((1 - \lambda)x + \lambda y) < (1 - \lambda)f(x) + \lambda f(y).$$

Montrer que pour tout $x \in \mathbf{R}^n$, $f(0) \leq f(x)$, puis que si $x \neq 0$, alors $f(x) > 0$.

CORRECTION DES EXERCICES DU CHAPITRE 13

SOLUTION DE L'EXERCICE 13.1

Un dessin vaut mieux qu'une longue explication :



SOLUTION DE L'EXERCICE 13.2

- La fonction $(x, y) \mapsto -x^2 + 2y$ est polynomiale donc \mathcal{C}^1 sur \mathbf{R}^2 .
Par composition avec la fonction exponentielle, qui est \mathcal{C}^1 sur \mathbf{R} , $(x, y) \mapsto e^{-x^2+2y}$ est \mathcal{C}^1 sur \mathbf{R}^2 .
De plus, $(x, y) \mapsto xy$ est polynomiale sur \mathbf{R}^2 donc \mathcal{C}^1 , et donc par produit de fonctions \mathcal{C}^1 , f est \mathcal{C}^1 sur \mathbf{R}^2 . On a alors

$$\forall (x, y) \in \mathbf{R}^2, \partial_1 f(x, y) = e^{-x^2+2y} (y - 2x^2y), \partial_2 f(x, y) = e^{-x^2+2} (x + 2xy).$$

Et donc

$$\nabla f(x, y) = (e^{-x^2+2y} (y - 2x^2y), e^{-x^2+2} (x + 2xy)).$$

- La fonction $(x, y, z) \mapsto xy$ est polynomiale, donc \mathcal{C}^1 sur \mathbf{R}^3 . De plus, $(x, y, z) \mapsto z$ est \mathcal{C}^1 sur \mathbf{R}^3 , et comme la fonction exponentielle est \mathcal{C}^1 sur \mathbf{R} , alors $(x, y, z) \mapsto e^z$ est \mathcal{C}^1 sur \mathbf{R}^3 .
Par produit de fonctions \mathcal{C}^1 , $(x, y, z) \mapsto xye^z$ est \mathcal{C}^1 sur \mathbf{R}^3 .
De même, on prouve que $(x, y, z) \mapsto xze^y$ et $(x, y, z) \mapsto yxe^x$ sont \mathcal{C}^1 , donc g est \mathcal{C}^1 sur \mathbf{R}^3 par somme de fonctions \mathcal{C}^1 .
On a alors

$$\partial_1 f(x, y, z) = ye^z + ze^y + ye^x + yxe^x, \partial_2 f(x, y, z) = xe^z + xze^y + xe^x, \partial_3 f(x, y, z) = xye^z + xe^y.$$

Méthode

Pour calculer la k -ième dérivée partielle d'une fonction, on fixe toutes les variables sauf x_k (c'est-à-dire qu'on les considère comme des constantes) et on dérive alors la fonction de la seule variable x_k .
Toutes les règles usuelles de dérivation s'appliquent : dérivée d'un produit, d'un quotient, d'une composée de fonctions, etc.

3. La fonction $(x_1, \dots, x_n) \mapsto x_1^2 + \dots + x_n^2$ est \mathcal{C}^1 car polynomiale. Donc, par composition avec la fonction $t \mapsto te^{-t}$, qui est \mathcal{C}^1 sur \mathbf{R} , la fonction h est polynomiale sur \mathbf{R}^n . On a alors, pour $i \in \llbracket 1, n \rrbracket$,

$$\partial_i h(x_1, \dots, x_n) = 2x_i e^{-(x_1^2 + \dots + x_n^2)} - 2x_i(x_1^2 + \dots + x_n^2) e^{-(x_1^2 + \dots + x_n^2)}.$$

SOLUTION DE L'EXERCICE 13.3

Soit $F : x \mapsto \int_1^x f(t) dt$. Alors nous savons¹ que F est de classe \mathcal{C}^1 , et que sa dérivée est f . On a de plus, pour tout $(x, y) \in \mathbf{R}^2$,

$$g(x, y) = \int_1^{xy} f(t) dt - \int_1^x f(t) dt = F(xy) - F(x).$$

Puisque les fonctions $(x, y) \mapsto xy$ et $(x, y) \mapsto x$ sont de classe \mathcal{C}^1 sur \mathbf{R}^2 car polynomiales, la fonction g y est également \mathcal{C}^1 car somme et composée de fonctions \mathcal{C}^1 .

Et on a alors

$$\partial_1 g(x, y) = yF'(xy) - F'(x) = yf(xy) - f(x) \text{ et } \partial_2 g(x, y) = xF'(xy) = xf(xy).$$

SOLUTION DE L'EXERCICE 13.4

1. Si $(x, y, z) = (0, 0, 0)$, le résultat est évident. Sinon, on a $\sqrt{x^2 + y^2 + z^2} \geq \sqrt{x^2} = |x|$, de sorte que

$$|f(x, y, z)| = \frac{|xyz|}{\sqrt{x^2 + y^2 + z^2}} \leq \frac{|x||yz|}{|x|} = |yz|.$$

2. Les fonctions $(x, y, z) \mapsto xyz$ et $(x, y, z) \mapsto x^2 + y^2 + z^2$ sont continues sur \mathbf{R}^3 , et la fonction $t \mapsto \sqrt{t}$ est continue sur \mathbf{R}_+ . Ainsi, $(x, y, z) \mapsto \sqrt{x^2 + y^2 + z^2}$ est continue sur \mathbf{R}^3 , et ne s'annule pas sur $\mathbf{R}^3 - \{(0, 0, 0)\}$, donc f est continue sur $\mathbf{R}^3 - \{(0, 0, 0)\}$. De plus, notons que pour $\eta > 0$, si $\|(x, y, z)\| \leq \eta$, alors $|y| \leq \eta$ et $|z| \leq \eta$. Soit donc $\varepsilon > 0$, et soit $(x, y, z) \in \mathbf{R}^3$ tel que $\|(x, y, z)\| \leq \sqrt{\varepsilon}$. Alors

$$|f(x, y, z) - f(0, 0, 0)| = |f(x, y, z)| \leq |y||z| \leq (\sqrt{\varepsilon})^2 = \varepsilon.$$

Ainsi, en posant $\eta = \sqrt{\varepsilon}$, on a bien prouvé que

$$\forall \varepsilon > 0, \exists \eta > 0 : \|(x, y, z)\| \leq \eta \Rightarrow |f(x, y, z) - f(0, 0, 0)| \leq \varepsilon.$$

La fonction f est donc continue en $(0, 0, 0)$, et elle est ainsi continue sur \mathbf{R}^3 tout entier.

SOLUTION DE L'EXERCICE 13.5

La fonction f est \mathcal{C}^1 car polynomiale. De plus

$$\forall (x, y) \in \mathbf{R}^2, \partial_1 f(x, y) = 2x + 2y + 2 \text{ et } \partial_2 f(x, y) = 2x + 2y + 3$$

1. Notons que le vecteur en question est bien de norme 1. D'après le théorème du cours, puisque f est \mathcal{C}^1 , sa dérivée directionnelle en a dans la direction de $(\frac{4}{5}, \frac{-3}{5})$ est

$$\left\langle \nabla f(-1, 1), \left(\frac{4}{5}, \frac{-3}{5} \right) \right\rangle = \frac{4}{5} \partial_1 f(-1, 1) - \frac{3}{5} \partial_2 f(-1, 1) = \frac{8}{5} - \frac{9}{5} = -\frac{1}{5}.$$

2. D'après le cours, nous savons que la dérivée directionnelle est égale à

$$\frac{1}{\|\nabla f(a)\|} \langle \nabla f(a), \nabla f(a) \rangle = \|\nabla f(a)\| = \sqrt{13}.$$

Rappel : la dérivée directionnelle est la plus grande dans la direction de $\nabla f(a)$, donc en tout autre point, la dérivée directionnelle sera, en valeur absolue, inférieure ou égale à $\sqrt{13} = \|\nabla f(a)\|$.

SOLUTION DE L'EXERCICE 13.6

¹ C'est le théorème fondamental de l'analyse.

Détails

Il s'agit du même raisonnement qu'à la question précédente : $|y| \leq \|(x, y, z)\|$ et $|z| \leq \|(x, y, z)\|$.

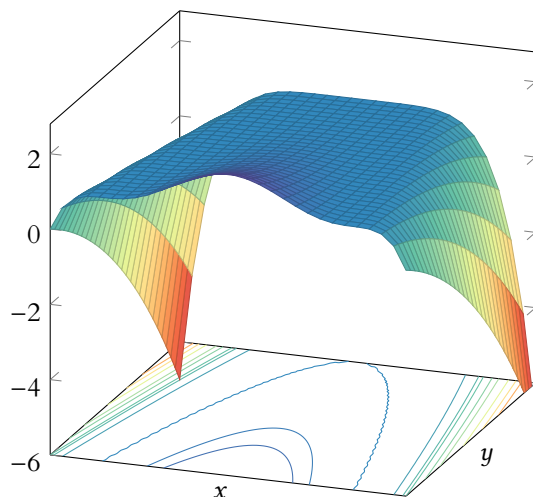


FIGURE 13.1 – La fonction f et des lignes de niveau (constatons que ce sont bien des paraboles).

1. Soit $\lambda \in \mathbf{R}$. Alors la courbe de niveau λ de f est

$$\begin{aligned} \mathcal{C}_\lambda(f) &= \{(x, y) \in \mathbf{R}^2 : f(x, y) = \lambda\} \\ &= \{(x, y) \in \mathbf{R}^2 : (y - 2x^2)^3 = \lambda - 1\} = \{(x, y) \in \mathbf{R}^2 : y = 2x^2 + \sqrt[3]{\lambda - 1}\}. \end{aligned}$$

On reconnaît alors l'équation d'une parabole.

2. D'après ce que nous venons de dire, la courbe de niveau λ de f n'est autre que la courbe représentative de $g_\lambda : x \mapsto 2x^2 + \sqrt[3]{\lambda - 1}$.
Un vecteur directeur de la tangente en (x_0, y_0) est alors donné par $u = (1, g'_\lambda(x_0)) = (1, 4x_0)$.
D'autre part, f est \mathcal{C}^1 sur \mathbf{R}^2 car polynomiale, et ses dérivées partielles sont données par

$$\partial_1 f(x, y) = -12x(y - 2x^2)^2, \quad \partial_2 f(x, y) = 3(y - 2x^2)^2.$$

Et donc $\nabla f(x_0, y_0) = (-12x_0(y_0 - 2x_0^2)^2, 3(y_0 - 2x_0^2)^2)$.

Et donc on a

$$\langle u, \nabla f(x_0, y_0) \rangle = -12x_0(y_0 - 2x_0^2)^2 + 4x_0(y_0 - 2x_0^2)^2 = 0.$$

Et donc le gradient est bien orthogonal au vecteur directeur de la tangente.

SOLUTION DE L'EXERCICE 13.7

La fonction f est bien évidemment de classe \mathcal{C}^1 sur \mathbf{R}^3 car polynomiale.
De plus on a

$$\partial_1 f(x, y, z) = x + yz, \quad \partial_2 f(x, y, z) = xz + 1, \quad \partial_3 f(x, y, z) = xy - 1$$

de sorte que $\nabla f(x, y, z) = (x + yz, xz + 1, xy - 1)$.

Un triplet $(x, y, z) \in \mathbf{R}^3$ est un point critique de f si et seulement si

$$\begin{cases} x + yz = 0 \\ xz + 1 = 0 \\ xy - 1 = 0 \end{cases}$$

En particulier, on a $xy = 1$, et donc $x \neq 0$ et $y \neq 0$. On a alors

$$y = \frac{1}{x} \text{ et de même, } z \neq 0 \text{ et } z = -\frac{1}{x}$$

de sorte que la première équation devient $x - \frac{1}{x^2} = 0 \Leftrightarrow x^3 = 1$.

Cette équation possède une unique solution $x = 1$.

Et donc la seule solution possible au système est $(1, 1, -1)$.

Racine cubique

Rappelons que la fonction $\varphi : t \mapsto t^3$ est une bijection de \mathbf{R} dans \mathbf{R} (car elle est strictement croissante) et donc que tout réel x admet un unique antécédent par φ , noté $\sqrt[3]{x}$.
En particulier, notons que la fonction racine cubique est défini sur \mathbf{R} tout entier, y compris sur \mathbf{R}_- (à la différence de la fonction racine carrée).

Remarque

Notons qu'il n'y avait pas unicité du vecteur directeur de la tangente, mais que tout autre vecteur directeur est de la forme λu avec $\lambda \neq 0$, ce qui ne change pas l'orthogonalité.

Méthode

Pour résoudre un système qui n'est pas linéaire, la seule option est d'essayer de procéder par substitution, et donc de commencer par exprimer toutes les variables en fonction d'une seule (ici on a choisi x). À la fin, il reste alors à résoudre une équation (plus ou moins complexe) en cette seule variable.

Inversement, on vérifie que $(1, 1, -1)$ est bien une solution, et donc l'unique point critique de f est $(1, 1, -1)$.

Ce point critique ne saurait être un extremum de f car on a

$$\lim_{x \rightarrow +\infty} f(x, 0, 0) = \lim_{x \rightarrow +\infty} \frac{x^2}{2} = +\infty$$

et donc f n'admet pas de maximum.

De même, on a

$$\lim_{z \rightarrow +\infty} f(0, z, 0) = \lim_{z \rightarrow +\infty} -z = -\infty$$

et donc f n'admet pas de minimum.

SOLUTION DE L'EXERCICE 13.8

1. Fixons par exemple $y = 0$. Alors $f(x, 0) = x^4 \xrightarrow{x \rightarrow +\infty} +\infty$.

Donc f peut prendre des valeurs arbitrairement grandes : elle ne possède pas de maximum.

2.a. On a, pour tout $(x, y) \in \mathbb{R}^2$, $f(-x, -y) = f(x, y)$.

Notons alors $\mathcal{H} = \{(x, y) \in \mathbb{R}^2 : y \geq 0\}$ le demi-plan formé des points d'ordonnée positive. Il est évident que si f possède un minimum sur \mathbb{R}^2 , alors elle possède un minimum sur \mathcal{H} . Inversement, supposons que f possède un minimum sur \mathcal{H} , c'est-à-dire qu'il existe $(x_0, y_0) \in \mathcal{H}$ tel que

$$\forall (x, y) \in \mathcal{H}, f(x, y) \geq f(x_0, y_0).$$

Montrons qu'il s'agit également d'un minimum de f sur \mathbb{R}^2 . Soit donc $(x, y) \in \mathbb{R}^2$.

- Si $y \geq 0$, alors $(x, y) \in \mathcal{H}$ et donc $f(x, y) \geq f(x_0, y_0)$.
- Si $y < 0$, alors $(-x, -y) \in \mathcal{H}$ et donc $f(x, y) = f(-x, -y) \geq f(x_0, y_0)$.

Ainsi, on a prouvé que pour tout $(x, y) \in \mathbb{R}^2$, $f(x, y) \geq f(x_0, y_0)$: f possède un minimum en (x_0, y_0) .

Dans la suite, on restreint donc la recherche de minimum à \mathcal{H} .

Limites

Le calcul de ces limites prouve que f peut prendre des valeurs arbitrairement grandes ou arbitrairement petites, et donc n'admet pas d'extremum. Une telle limite ne pourrait jamais justifier l'existence d'un extremum, mais uniquement la réfuter.

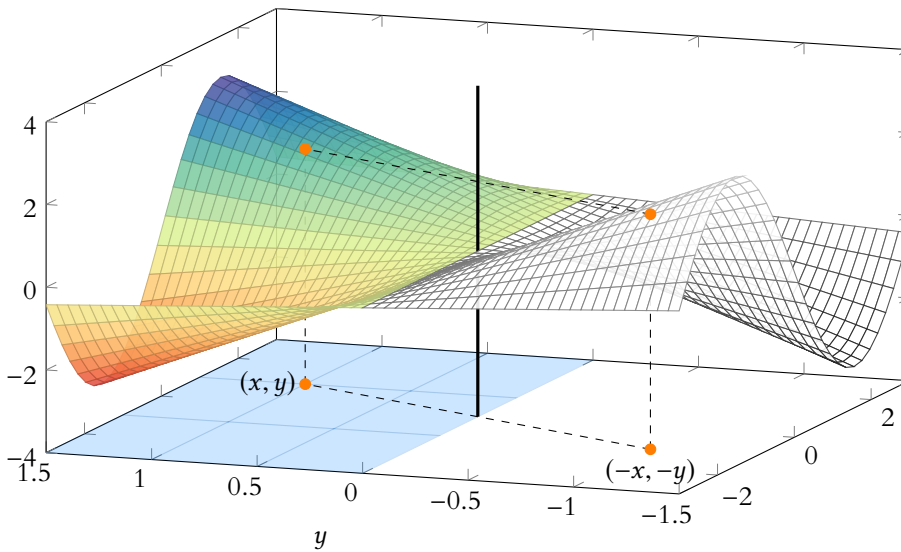


FIGURE 13.2 – Un exemple de fonction vérifiant $f(x, y) = f(-x, -y)$ (il ne s'agit pas de la fonction f de l'exercice).

2.b. Pour y fixé, soit $g_y : x \mapsto x^4 + y^4 - 4xy$.

Elle est dérivable² sur \mathbb{R} , de dérivée égale à $g'_y(x) = 4(x^3 - y)$. Son tableau de variations est alors le suivant :

x	$-\infty$	$\sqrt[3]{y}$	$+\infty$
$g'_y(x)$		0	
g_y	$+\infty$	$f(\sqrt[3]{y}, y)$	$+\infty$

Sur le plan du bas, on a représenté \mathcal{H} en bleu. La restriction de f à \mathcal{H} est donc la partie en couleur de la surface. Il suffit de connaître cette partie pour en déduire le graphe de f sur \mathbb{R}^2 : l'autre partie est obtenue par symétrie par rapport à l'axe des z (résultat à mettre en parallèle de la symétrie de la courbe représentative d'une fonction paire d'une seule variable). En particulier, si la restriction de f à \mathcal{H} possède un minimum, f possède un minimum sur \mathbb{R}^2 tout entier.

² Notons que g_y est une fonction d'une seule variable, tout ce qu'il y a de plus classique !

et donc g_y admet un minimum en $x = \sqrt[3]{y}$.

Et ce minimum vaut alors

$$g(y) = g_y(\sqrt[3]{y}) = (\sqrt[3]{y})^4 + y^4 - 4\sqrt[3]{y}y = y^4 - 3y^{4/3}.$$

2.c. La fonction g est dérivable et $g'(y) = 4(y^3 - y^{1/3})$.

On a alors $g'(y) = 0 \Leftrightarrow y^3 = y^{1/3} \Leftrightarrow y^9 = y \Leftrightarrow y = 0$ ou $y = 1$.

Le tableau de variation de g est donc

y	0		1		$+\infty$	
$g'(y)$	0	-	0	+		
g	0	↘		-2	↗	
						$+\infty$

Elle admet donc un minimum en $y = 1$, et ce minimum vaut $g(1) = -2$. On en déduit que

$$\forall (x, y) \in \mathbf{R} \times [0, +\infty[, f(x, y) \geq g(y) \geq -2.$$

De plus, on a $f(x, y) = -2$ si et seulement si les inégalités ci-dessus sont des égalités.

En particulier, il faut avoir $g(y) = -2 \Leftrightarrow y = 1$.

Et il faut alors également que $f(x, 1) = g_1(x) \Leftrightarrow x = \sqrt[3]{1} = 1$.

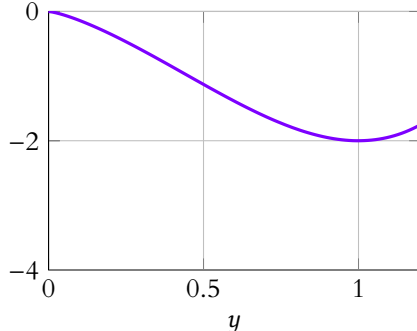
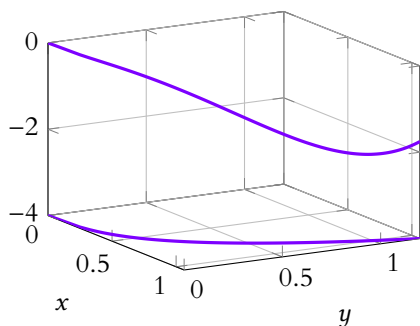
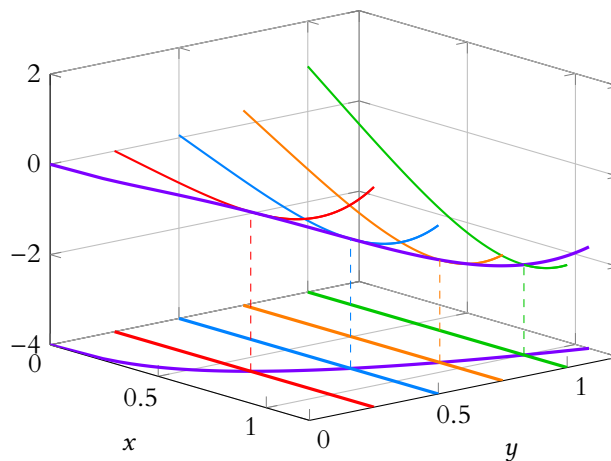
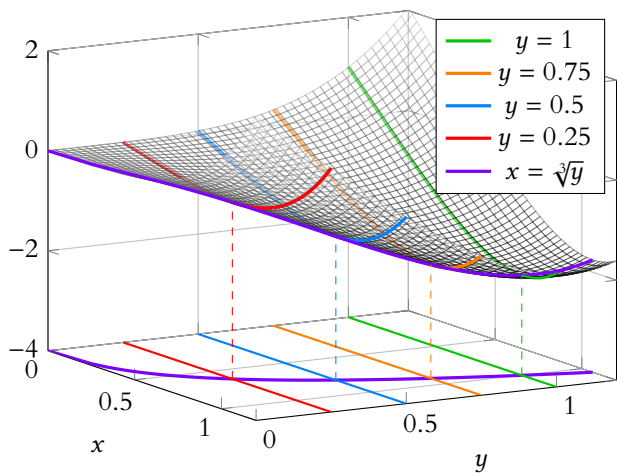
Enfin, si $y \leq 0$ on a

$$f(x, y) = f(-x, -y) \geq -2$$

avec égalité si et seulement si $(-x, -y) = (-1, -1)$.

Donc $\forall (x, y) \in \mathbf{R}^2, f(x, y) \geq -2$, avec égalité si et seulement si $(x, y) = (1, 1)$ ou $(x, y) = (-1, -1)$.

f possède donc un minimum atteint en deux points qui sont $(1, 1)$ et $(-1, -1)$.



Pour y fixé, $x \mapsto f(x, y)$ admet un minimum en $x = \sqrt[3]{y}$. En violet, la courbe reliant ces différents minimums. Il s'agit d'une courbe tracée dans \mathbf{R}^3 , mais en la regardant «de côté», c'est-à-dire comme le graphe d'une fonction ne dépendant que de la variable y (c'est le dernier dessin), on obtient le graphe de la fonction g . Celle-ci atteint son minimum en $y = 1$. Si on le voit de nouveau dans l'espace, il s'agit du minimum de la fonction f .

Remarque : a priori, cette méthode pourrait fonctionner pour toutes les fonctions de deux variables (et pourrait se généraliser aux fonctions de n variables). Toutefois, il faut pour cela être capable de donner l'expression exacte du (ou des) point(s) en le(s)quel(s) $x \mapsto f(x, y)$ admet un extremum, ce qui est en fait rarement possible.

SOLUTION DE L'EXERCICE 13.9

1. La fonction $(x, y) \mapsto -3y$ est de classe \mathcal{C}^1 sur \mathbf{R}^2 car polynomiale. Il en est de même de $(x, y) \mapsto 1 + x^2 + y^2$, qui ne s'annule alors pas³ sur \mathbf{R}^2 . Et donc f est de classe \mathcal{C}^1 sur \mathbf{R}^2 car quotient de deux fonctions \mathcal{C}^1 . On a alors, pour $(x, y) \in \mathbf{R}^2$:

$$\partial_1 f(x, y) = \frac{6xy}{(1 + x^2 + y^2)^2}, \quad \partial_2 f(x, y) = \frac{-3(1 + x^2 + y^2) + 3y(2y)}{(1 + x^2 + y^2)^2} = \frac{3y^2 - 3x^2 - 3}{(1 + x^2 + y^2)^2}.$$

2. $(x, y) \in \mathbf{R}^2$ est un point critique de f si et seulement si

$$\begin{cases} \partial_1 f(x, y) = 0 \\ \partial_2 f(x, y) = 0 \end{cases} \Leftrightarrow \begin{cases} 6xy = 0 \\ y^2 - x^2 - 1 = 0 \end{cases}$$

Or, on a $xy = 0$ si et seulement si $x = 0$ ou $y = 0$.

Si $x = 0$, alors $y^2 - 1 = 0 \Leftrightarrow y = \pm 1$.

Et si $y = 0$, alors $-x^2 - 1 = 0 \Leftrightarrow x^2 = -1$ ne possède pas de solution.

Donc les seuls points critiques possibles sont $(0, 1)$ et $(0, -1)$, on vérifie aisément que ce sont bien deux points critiques.

3. On a $f(0, 1) = -\frac{3}{2}$ et $f(0, -1) = \frac{3}{2}$.
On a, quel que soit $(x, y) \in \mathbf{R}^2$,

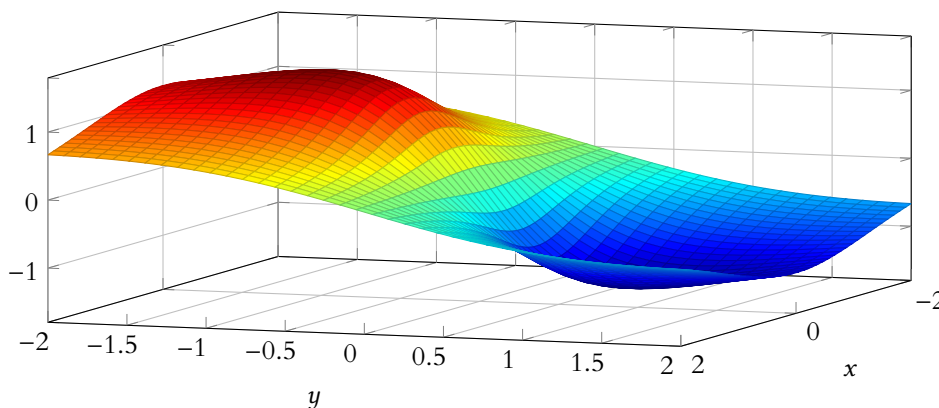
$$\begin{aligned} f(x, y) - f(0, -1) &= f(x, y) - \frac{3}{2} = \frac{-6y}{2(1 + x^2 + y^2)} - \frac{3(1 + x^2 + y^2)}{2(1 + x^2 + y^2)} \\ &= -3 \frac{x^2 + y^2 + 2y + 1}{x^2 + y^2 + 1} \\ &= -3 \frac{x^2 + (y + 1)^2}{x^2 + y^2 + 1} \leq 0. \end{aligned}$$

Et donc $f(x, y) \leq f(0, -1)$: f admet un maximum en $(0, -1)$.

De même, on a, pour tout $(x, y) \in \mathbf{R}^2$,

$$f(x, y) - f(0, 1) = f(x, y) + \frac{3}{2} = 3 \frac{x^2 + y^2 - 2y + 1}{x^2 + y^2 + 1} = 3 \frac{x^2 + (y - 1)^2}{x^2 + y^2 + 1} \geq 0.$$

Et donc, $f(x, y) \geq f(0, 1)$: f admet un minimum en $(0, 1)$.



SOLUTION DE L'EXERCICE 13.10

1. La fonction $(x, y) \mapsto x$ est de classe \mathcal{C}^1 sur \mathbf{R}^2 car polynomiale. Et donc par composition avec la fonction $t \mapsto e^t$, $(x, y) \mapsto e^x$ est également de classe \mathcal{C}^1 . D'autre part, $(x, y) \mapsto x + y^2$ est polynomiale, donc de classe \mathcal{C}^1 , et donc par somme et produit de fonctions \mathcal{C}^1 , g est une fonction de classe \mathcal{C}^1 sur \mathbf{R}^2 . On a alors

$$\partial_1 g(x, y) = e^x(x + y^2 + e^x) + e^x(1 + e^x) = e^x(1 + x + y^2 + 2e^x), \quad \partial_2 g(x, y) = 2ye^x.$$

³ Elle est clairement supérieure à 1.

En particulier

Ceci prouve déjà que f n'a pas de maximum en $(0, 1)$ et pas de minimum en $(0, -1)$.

2. La fonction f est clairement dérivable sur \mathbf{R} , et $f'(t) = 1 + 2e^t > 0$.
Donc f est strictement croissante sur \mathbf{R} . De plus, $\lim_{t \rightarrow +\infty} f(t) = +\infty$ et $\lim_{t \rightarrow -\infty} f(t) = -\infty$, donc par le théorème de la bijection il existe un unique $\alpha \in \mathbf{R}$ tel que $f(\alpha) = 0$.

De plus, on a $f(-2) = -2 + 2e^{-2} = 2\left(\frac{1}{e^2} - 1\right) < 0$ et $f(-1) = 2e^{-1} > 0$.

Et donc par croissance de f , on a nécessairement $\alpha \in]-2, -1[$.

3. Un point $(x, y) \in \mathbf{R}^2$ est un point critique de g si et seulement si

$$\begin{cases} \partial_1 g(x, y) = 0 \\ \partial_2 g(x, y) = 0 \end{cases} \Leftrightarrow \begin{cases} e^x(1+x+y^2+2e^x) = 0 \\ 2ye^x = 0 \end{cases} \Leftrightarrow \begin{cases} 1+x+y^2+2e^x = 0 \\ y = 0 \end{cases} \Leftrightarrow \begin{cases} f(x) = 0 \\ y = 0 \end{cases}$$

Et donc $(\alpha, 0)$ est l'unique point critique de g .

4. Puisque $e^x > 0$ et $y^2 > 0$, il est évident que pour tout $(x, y) \in \mathbf{R}^2$,

$$g(x, y) = e^x(1+x+y^2+e^x) \geq e^x(1+x+e^x) = g(x, 0).$$

Notons donc $h : x \mapsto g(x, 0) = e^x(1+x+e^x)$. Alors h est dérivable sur \mathbf{R} , et sa dérivée est $h' : x \mapsto e^x(1+x+2e^x) = e^x f(x)$.

Donc cette dérivée est du signe de f . et donc h possède un minimum en α .

On en déduit donc que pour tout $x \in \mathbf{R}$, $g(x, 0) = h(x) \geq h(\alpha) = g(\alpha, 0)$.

Et donc, pour tout $(x, y) \in \mathbf{R}^2$, $g(x, y) \geq g(\alpha, 0) : g$ admet un minimum en $(\alpha, 0)$.

Notons que puisque g ne possède qu'un point critique, ce minimum n'est atteint qu'en un point.

SOLUTION DE L'EXERCICE 13.11

1. f est polynomiale, donc elle est évidemment \mathcal{C}^1 sur \mathbf{R}^2 .

On a alors, pour $(x, y) \in \mathbf{R}^2$,

$$\partial_1 f(x, y) = 4x^3 + 4x - 4y + 8, \quad \partial_2 f(x, y) = 2y - 4x - 4.$$

2. Un point $(x, y) \in \mathbf{R}^2$ est un point critique de f si et seulement si

$$\begin{cases} 4x^3 + 4x - 4y + 8 = 0 \\ 2y - 4x - 4 = 0 \end{cases} \Leftrightarrow \begin{cases} x^3 + x - y + 2 = 0 \\ y = 2x + 2 \end{cases} \Leftrightarrow \begin{cases} x^3 - x = 0 \\ y = 2x + 2 \end{cases} \Leftrightarrow \begin{cases} x(x^2 - 1) = 0 \\ y = 2x + 2 \end{cases}$$

Mais $x(x^2 - 1) = 0$ si et seulement si $x = 0$, $x = 1$ ou $x = -1$. Et on trouve donc trois points critiques qui sont $(0, 2)$, $(-1, 0)$ et $(1, 4)$.

On a alors $f(0, 2) = -2$, $f(-1, 0) = f(1, 4) = -3$.

Donc déjà f n'admet pas de minimum en $(0, 2)$.

De plus, f n'admet pas de maximum car $\lim_{x \rightarrow +\infty} f(x, 0) = \lim_{x \rightarrow +\infty} x^4 - 2x^2 + 8x + 2 = +\infty$.

Posons $\alpha = (-1, 0)$ et $\beta = (1, 4)$. Puisque $f(\alpha) = f(\beta)$, il reste donc à déterminer si f admet un minimum en α , et alors l'autre point critique sera de même nature.

- 3.a. La fonction g est donnée par

$$g(x, y) = x^4 + 2x^2 + y^2 - 4xy + 8x - 4y + 5 = y^2 - (4x+4)y + (x^4 + 2x^2 + 8x + 5).$$

Lorsque seul y varie, il s'agit d'une fonction polynomiale de degré 2, dont le discriminant est

$$\Delta(x) = (4x+4)^2 - 4(x^4 + 2x^2 + 8x + 5) = 16x^2 + 32x + 16 - 4x^4 - 8x^2 - 32x - 20 = -4x^4 + 8x^2 - 4.$$

- 3.b. On a $\Delta(x) = -4(x^4 - 2x^2 + 1) = -4(x^2 - 1)^2$ qui est donc toujours négatif (et même strictement négatif). Cela signifie alors que $y \mapsto g(x, y)$ ne change pas de signe.

Puisque le coefficient de degré 2 de $g(x, y)$ est positif, elle ne prend que des valeurs positives :

$$\forall y \in \mathbf{R}, g(x, y) \geq 0 \Leftrightarrow f(x, y) \geq f(\alpha).$$

Et le résultat qui vient d'être prouvé est valable pour tout $x \in \mathbf{R}$, de sorte que

$$\forall (x, y) \in \mathbf{R}^2, f(x, y) \geq f(\alpha).$$

f admet donc un minimum en $(-1, 0)$, et puisque $f(-1, 0) = f(1, 4)$, elle admet également un minimum en $(1, 4)$.

Rappel

Le théorème de la bijection s'applique car f est strictement monotone, continue (car dérivable), et que 0 est bien compris entre $-\infty$ et $+\infty$, les deux limites que nous venons d'obtenir.

Détails

L'étude de f faite à la question 2 nous donne son signe : elle est croissante, et s'annule en α . Elle est donc négative avant α et positive après α . Ce qui nous donne donc le sens de variation de h .

Remarque

Ce discriminant s'annule pour $x = \pm 1$. Ceci est cohérent avec le fait que pour $x = \pm 1$, la fonction $y \mapsto f(x, y)$ s'annule au moins une fois (lorsque $y = 0$ ou $y = 4$).

SOLUTION DE L'EXERCICE 13.12

- Il est facile de vérifier que -1 est solution. De plus, il est clair qu'une telle solution vérifiant $x = -e^{x-1/x}$ est nécessairement négative. Sur \mathbf{R}_-^* , la fonction $x \mapsto x - \frac{1}{x}$ est strictement croissante, donc il en est de même de $x \mapsto x + e^{x-1/x}$ comme somme et composée de fonctions strictement croissantes. Cette fonction est donc injective sur \mathbf{R}_-^* : de sorte que 0 y possède un unique antécédent, qui est -1 .
- Les fonctions $(x, y) \mapsto x$ et $(x, y) \mapsto y$ sont \mathcal{C}^1 sur \mathbf{R}^2 car polynomiales, et la fonction $t \mapsto e^t$ est \mathcal{C}^1 sur \mathbf{R} . Ainsi, les fonctions $(x, y) \mapsto e^x$ et $(x, y) \mapsto e^y$ sont \mathcal{C}^1 sur \mathbf{R}^2 . Et alors f est \mathcal{C}^1 sur \mathbf{R}^2 car somme et produit de fonctions \mathcal{C}^1 . De plus, on a

$$\forall (x, y) \in \mathbf{R}^2, \partial_1 f(x, y) = ye^x + e^y \text{ et } \partial_2 f(x, y) = xe^y + e^x.$$

Donc (a, b) est un point critique de f si et seulement si

$$\begin{cases} be^a + e^b = 0 \\ ae^b + e^a = 0 \end{cases} \Leftrightarrow \begin{cases} e^b = -be^a \\ -abe^a + e^a = 0 \end{cases} \Leftrightarrow \begin{cases} e^b = -be^a \\ e^a(1 - ab) = 0 \end{cases}$$

En utilisant la seconde ligne⁴, on a alors $1 - ab = 0 \Leftrightarrow ab = 1$, et donc $a \neq 0$ et $b = \frac{1}{a}$. Alors l'équation $ae^b + e^a = 0$ devient

$$ae^{1/a} + e^a = 0 \Leftrightarrow e^{1/a} (a + e^{a-1/a}) = 0.$$

On en déduit, à l'aide de la question précédente que $a = -1$, et donc $b = -1$. Inversement, on vérifie aisément qu'il s'agit bien là d'un point critique de f .

- On a $f(-1, -1) = -2e^{-1}$. Or, $f(x, 0) = x$, et en faisant tendre x vers $+\infty$ (resp. $-\infty$), on remarque que f n'est pas majorée (resp. pas minorée).

SOLUTION DE L'EXERCICE 13.13

La fonction $(x, y) \mapsto x$ est \mathcal{C}^1 car polynomiale, et donc, par composition avec la fonction sin, qui est \mathcal{C}^1 sur \mathbf{R} , la fonction $(x, y) \mapsto \sin(x)$ est de classe \mathcal{C}^1 sur \mathbf{R}^2 . De plus, $(x, y) \mapsto y^2 - 2y + 1$ est polynomiale sur \mathbf{R}^2 , donc \mathcal{C}^1 et ainsi, f est \mathcal{C}^1 car somme de fonctions \mathcal{C}^1 .

On a alors, pour $(x, y) \in \mathbf{R}^2$,

$$\partial_1 f(x, y) = \cos x, \partial_2 f(x, y) = 2y - 2.$$

(x, y) est donc un point critique de f si et seulement si $\begin{cases} \cos x = 0 \\ 2y - 2 = 0 \end{cases}$

Ce système possède une infinité de solutions, qui sont les $(\frac{\pi}{2} + k\pi, 1)$, $k \in \mathbf{Z}$.

On a alors $f(\frac{\pi}{2} + k\pi, 1) = \begin{cases} 1 & \text{si } k \text{ est pair} \\ -1 & \text{sinon} \end{cases}$

Notons que f ne possède pas de maximum car $\lim_{y \rightarrow +\infty} f(0, y) = +\infty$. Et donc les $(\frac{\pi}{2} + k\pi, 1)$ avec k pair ne correspondent pas à des extrema de f .

Enfin, pour $(x, y) \in \mathbf{R}^2$, on a

$$f(x, y) = \sin(x) + (y - 1)^2 \geq -1 + 0 = -1.$$

Et donc tous les $(\frac{\pi}{2} + k\pi, 1)$ avec k impair correspondent à des minima de f .

SOLUTION DE L'EXERCICE 13.14

La fonction f est polynomiale sur \mathbf{R}^3 , donc de classe \mathcal{C}^1 . De plus, on a

$$\forall (x, y, z) \in \mathbf{R}^3, \partial_1 f(x, y, z) = yz + y + z, \partial_2 f(x, y, z) = xz + x + z \text{ et } \partial_3 f(x, y, z) = xy + x + y.$$

Ainsi, $(x, y, z) \in \mathbf{R}^3$ est un point critique de f si et seulement si

$$\begin{cases} yz + y + z = 0 \\ xy + x + y = 0 \\ xz + x + z = 0 \end{cases} \xLeftrightarrow_{L_1 \mapsto L_1 - L_2} \begin{cases} y(z - x) + (x - z) = 0 \\ xy + x + y = 0 \\ xz + x + z = 0 \end{cases} \Leftrightarrow \begin{cases} (y + 1)(z - x) = 0 \\ xy + x + y = 0 \\ xz + x + z = 0 \end{cases}$$

⚠ Danger !

Si on ne voit pas qu'une solution est nécessairement dans \mathbf{R}_-^* , il est également possible de tenir le même raisonnement sur \mathbf{R}_+^* . Mais on prendra garde à ne pas affirmer que f est strictement croissante sur \mathbf{R}^* . En effet, \mathbf{R}^* n'est pas un intervalle, et $f' > 0$ sur \mathbf{R}^* ne suffit pas à garantir que f soit croissante sur \mathbf{R}^* . Elle l'est sur \mathbf{R}_-^* et sur \mathbf{R}_+^* , mais pas sur \mathbf{R}^* .

⁴ Une exponentielle n'est jamais nulle.

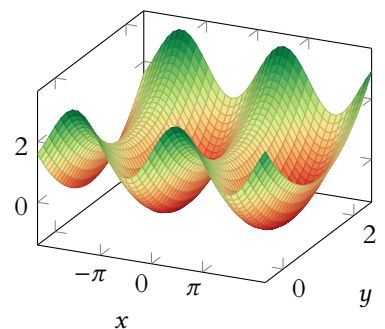


FIGURE 13.3– La fonction f .

En un point critique, on a donc soit $y + 1 = 0 \Leftrightarrow y = -1$, soit $z - x = 0 \Leftrightarrow z = x$.

Mais si $y = -1$, alors la seconde ligne devient $-1 = 0$, ce qui est impossible.

Donc nécessairement $x = z$, et alors la dernière ligne devient $x^2 + 2x = 0$. Donc $x = 0$ ou $x = 2$.

Enfin, dans les deux cas, en soustrayant L_2 à L_3 , il vient $(x + 1)(z - y) = 0$, et donc $z = y$.

On vérifie aisément qu'on a ainsi deux points critiques $(0, 0, 0)$ et $(2, 2, 2)$, qui sont les seuls points critiques de f .

Ces points critiques ne sont pas des extrema de f car $f(x, 1, 1) = 3x$ qui n'est ni majorée ni minorée.

SOLUTION DE L'EXERCICE 13.15

1. On a $f(x, y) = (x^2 - 2y^2)^2 \geq 0$.

De plus, $f(x, y) = 0$ si et seulement si $x^2 = 2y^2 \Leftrightarrow x = \pm\sqrt{2}y$.

f possède donc un minimum égal à 0, atteint en tout point de la forme $(\sqrt{2}y, y)$, $y \in \mathbf{R}$.

Et f ne possède pas de maximum car $\lim_{x \rightarrow +\infty} f(x, 0) = +\infty$.

2. Pour tout $(x, y) \in \mathbf{R}^2$, on a

$$g(x, y) = \left(x + \frac{y}{2}\right)^2 + \frac{3y^2}{4} \geq 0.$$

Comme on a $g(0, 0) = 0$, alors 0 est le minimum de g .

Comme de plus $g(x, y) = 0$ si et seulement si $y = 0$ et $x + \frac{y}{2} = 0$, alors $(0, 0)$ est le seul point où ce minimum est atteint.

3. h est polynomiale, donc de classe \mathcal{C}^1 . On a alors

$$\partial_1 h(x, y, z) = x - y - z + 4x^3, \partial_2 h(x, y, z) = 2y - x, \partial_3 h(x, y, z) = 4z - x.$$

(x, y, z) est un point critique de h si et seulement si

$$\begin{cases} x - y - z + 4x^3 = 0 \\ 2y - x = 0 \\ 4z - x = 0 \end{cases} \Leftrightarrow \begin{cases} x - y - z + 4x^3 = 0 \\ y = \frac{x}{2} \\ z = \frac{x}{4} \end{cases} \Leftrightarrow \begin{cases} x - \frac{x}{2} - \frac{x}{4} + 4x^3 = 0 \\ y = \frac{x}{2} \\ z = \frac{x}{4} \end{cases}$$

Cette dernière équation est $\frac{x}{4} + 4x^3 = 0 \Leftrightarrow x(\frac{1}{4} + 4x^2) = 0$, qui a pour seule solution $x = 0$. Et donc h admet un unique point critique $(0, 0, 0)$.

Remarquons qu'on a $h(x, y, z) = \left(z - \frac{x}{2}\right)^2 + \left(y - \frac{x}{2}\right)^2 + z^2 + x^4 \geq 0 = f(0, 0, 0)$.

Donc h admet un minimum en $(0, 0, 0)$.

4. i est polynomiale, donc de classe \mathcal{C}^1 sur \mathbf{R}^2 . On a alors

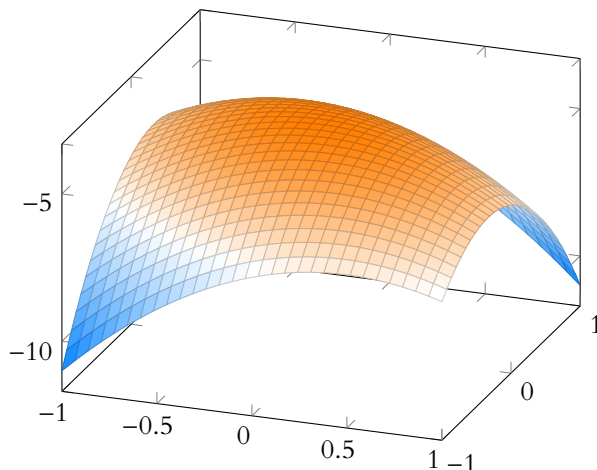
$$\nabla i(x, y) = (-4x - 2y, -2x - 6y)$$

Donc i admet un unique point critique $(0, 0)$, et on a alors $i(0, 0) = -4$.

On a $i(x, y) - i(0, 0) = i(x, y) + 4 = -2x^2 - 2xy - 3y^2$. Mais on a alors

$$i(x, y) + 4 = -(x + y)^2 - x^2 - 2y^2 \leq 0.$$

Et donc i admet un maximum en $(0, 0)$.



Autrement dit

f peut prendre des valeurs arbitrairement grandes et arbitrairement petites. Donc f n'admet ni maximum ni minimum.

Remarque

On pourrait aussi chercher les points critiques de f et vérifier qu'ils correspondent bien à des minima, mais lorsqu'on a la possibilité de s'en passer, autant aller plus vite !

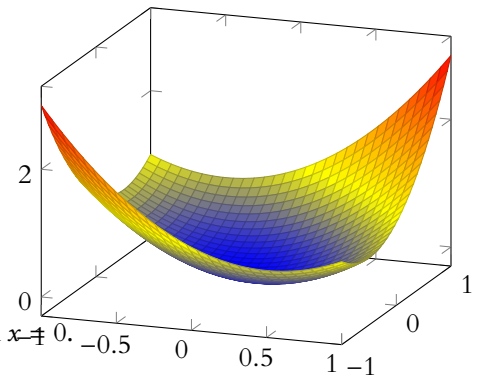


FIGURE 13.4— Le graphe de g .

Méthode

Lorsqu'on cherche à écrire une quantité comme somme de carrés, il faut commencer par s'intéresser aux termes dont le signe n'est pas constant : ici $2xy$, et essayer de les faire apparaître comme double produit (le $2ab$ de $(a + b)^2$).

5. On a $\lim_{x \rightarrow +\infty} j(x, 0, 0) = +\infty$ et $\lim_{x \rightarrow +\infty} j(x, x, x) = \lim_{x \rightarrow +\infty} 3x^2 - 2x^3 = -\infty$.
Et donc j n'admet ni maximum, ni minimum.

SOLUTION DE L'EXERCICE 13.16

1. Les fonctions $(x, y, z) \mapsto x + y + z$ et $(x, y, z) \mapsto x^2 + y^2 + z^2$ sont \mathcal{C}^1 sur \mathbf{R}^3 car polynomiales. $t \mapsto e^{-t/6}$ est \mathcal{C}^1 sur \mathbf{R} . Donc par composition et produit de fonctions de classe \mathcal{C}^1 , f est également \mathcal{C}^1 sur \mathbf{R}^3 . On a alors

$$\begin{aligned} \partial_1 f(x, y, z) &= -\frac{x}{3}(x+y+z)e^{-1/6(x^2+y^2+z^2)} + e^{-1/6(x^2+y^2+z^2)}, \quad \partial_2 f(x, y, z) = -\frac{y}{3}(x+y+z)e^{-1/6(x^2+y^2+z^2)} + e^{-1/6(x^2+y^2+z^2)} \\ \partial_3 f(x, y, z) &= -\frac{z}{3}(x+y+z)e^{-1/6(x^2+y^2+z^2)} + e^{-1/6(x^2+y^2+z^2)}. \end{aligned}$$

(x, y, z) est un point critique de f si et seulement si

$$\begin{cases} e^{-1/6(x^2+y^2+z^2)}(x(x+y+z)-3) = 0 \\ e^{-1/6(x^2+y^2+z^2)}(y(x+y+z)-3) = 0 \\ e^{-1/6(x^2+y^2+z^2)}(z(x+y+z)-3) = 0 \end{cases} \Leftrightarrow \begin{cases} x(x+y+z)-3 = 0 \\ y(x+y+z)-3 = 0 \\ z(x+y+z)-3 = 0 \end{cases} \Leftrightarrow \begin{cases} x = y = z \\ x(x+y+z) = 3 \end{cases} \Leftrightarrow \begin{cases} x = y = z \\ x^2 = 1 \end{cases}$$

Donc f admet deux points critiques qui sont $(1, 1, 1)$ et $(-1, -1, -1)$.

2. L'inégalité de Cauchy-Schwarz appliquée aux vecteurs (x, y, z) et $(1, 1, 1)$ nous donne

$$|x + y + z| \leq \sqrt{x^2 + y^2 + z^2} \sqrt{1^2 + 1^2 + 1^2} = \sqrt{3} \sqrt{x^2 + y^2 + z^2}.$$

3. La fonction $g : t \mapsto \sqrt{t}e^{-t/6}$ est de classe \mathcal{C}^1 sur \mathbf{R}_+^* , avec

$$g'(t) = -\frac{1}{6}\sqrt{t}e^{-t} + \frac{1}{2\sqrt{t}}e^{-t}.$$

On a alors

$$g'(t) \geq 0 \Leftrightarrow \frac{1}{2\sqrt{t}} - \frac{1}{6}\sqrt{t} \geq 0 \Leftrightarrow \frac{1}{2\sqrt{t}} \geq \frac{1}{6}\sqrt{t} \Leftrightarrow t \leq 3.$$

Et donc g admet un maximum en $t = 3$, et ce maximum vaut $g(3) = \sqrt{3}e^{-1/2}$.
Alors, pour tout $(x, y, z) \in \mathbf{R}^3$, on a

$$\begin{aligned} |f(x, y, z)| &= |x + y + z|e^{-1/6(x^2+y^2+z^2)} \\ &\leq \sqrt{3}\sqrt{x^2 + y^2 + z^2}e^{-1/6(x^2+y^2+z^2)} \\ &\leq \sqrt{3}g(x^2 + y^2 + z^2) \leq \sqrt{3}\sqrt{3}e^{-1/2} = 3e^{-1/2} = f(1, 1, 1). \end{aligned}$$

Donc f admet un maximum en $(1, 1, 1)$.

De même, on a $f(x, y, z) \geq -3e^{-2} = f(-1, -1, -1)$, donc f admet un minimum en $(-1, -1, -1)$.

SOLUTION DE L'EXERCICE 13.17

1. Les fonctions $(x, y, z) \mapsto x + z^2$ et $(x, y, z) \mapsto x(y^2 + z^2 + 1)$ sont de classe \mathcal{C}^1 sur \mathbf{R}^3 car polynomiales.
De plus, la fonction $t \mapsto e^t$ est \mathcal{C}^1 sur \mathbf{R} , et donc f est \mathcal{C}^1 sur \mathbf{R}^3 .
2. On a, $\forall (x, y, z) \in \mathbf{R}^3$,

$$\partial_1 f(x, y, z) = e^{x(y^2+z^2+1)}(1 + (x+z^2)(y^2+z^2+1)), \quad \partial_2 f(x, y, z) = 2xy(x+z^2)e^{x(y^2+z^2+1)}$$

et $\partial_3 f(x, y, z) = e^{x(y^2+z^2+1)}(2z + 2zx(x+z^2))$.

Donc $(a, b, c) \in \mathbf{R}^3$ est un point critique de f si et seulement si⁵

$$\begin{cases} 1 + (a + c^2)(b^2 + c^2 + 1) = 0 \\ 2ab(a + c^2) = 0 \\ 2c(1 + a(a + c^2)) = 0 \end{cases}$$

En utilisant la seconde équation, il vient soit $a = 0$, soit $b = 0$ soit $a + c^2 = 0$.

- Si $a + c^2 = 0$, alors la première équation devient $1 = 0$, ce qui n'est pas possible.

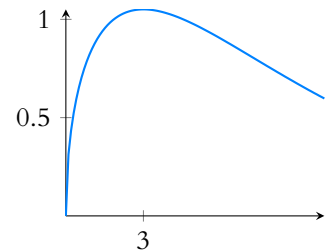


FIGURE 13.5– La fonction g .

Remarque

Ces inégalités sont des égalités si et seulement si $x^2 + y^2 + z^2 = 3$ (car g admet un unique minimum) et si il y a égalité dans le résultat de la question 2. Or il y a égalité dans Cauchy-Schwarz si et seulement si les vecteurs $(1, 1, 1)$ et (x, y, z) sont égaux, c'est-à-dire si et seulement si $x = y = z$.
On retrouve alors les deux points critiques : $(1, 1, 1)$ et $(-1, -1, -1)$.

⁵ Une exponentielle ne s'annulant jamais, on peut tout de suite enlever les exponentielles.

- Si $a = 0$, alors la dernière équation n'est satisfaite que si $c = 0$, et alors la première équation est de nouveau $1 = 0$, qui est impossible.
 - Si $b = 0$, la dernière équation nous indique que $c = 0$ ou $1 + a(a + c^2) = 0$.
- Si $1 + a(a + c^2) = 0$, notons qu'on ne peut avoir $a \geq 0$, car sinon $-1 = a(a + c^2) \geq 0$.

Donc $a < 0$, et donc $a + c^2 = -\frac{1}{a}$, de sorte que la première équation devient $1 - \frac{c^2 + 1}{a} = 0 \Leftrightarrow a = c^2 + 1 \geq 0$, d'où une contradiction.

On a donc nécessairement $c = 0$, et puisque b est déjà nul, la première équation est $a + 1 = 0 \Leftrightarrow a = -1$.

Et donc le seul point critique éventuel est donc $A = (-1, 0, 0)$. On a alors $f(A) = -e^{-1}$.

3. Si $x \geq 0$, alors $x + z^2 \geq 0$ et $e^{x(y^2+z^2+1)} \geq 0$, donc $f(x, y, z) \geq 0$.
 Si $x \leq 0$, on a $y^2 + z^2 + 1 \geq 1$ et donc après multiplication⁶ par x , $x(y^2 + z^2 + 1) \leq x$, de sorte que $e^x \geq e^{x(y^2+z^2+1)}$.
 En multipliant de nouveau par x , on a donc $xe^x \leq xe^{x(y^2+z^2+1)}$. Et donc, $f(x, y, z) = (x + z^2)e^{x(y^2+z^2+1)} \geq xe^{x(y^2+z^2+1)} \geq xe^x$.
4. Une rapide étude de la fonction $x \mapsto xe^x$ prouve qu'elle atteint son minimum sur \mathbf{R} en $x = -1$, et que ce minimum vaut $-e^{-1} = f(A)$.
 Donc f atteint bien son minimum en A , et ce minimum vaut $f(A) = -e^{-1}$.

⁶ Puisque $x \leq 0$, le sens de l'inégalité change !

SOLUTION DE L'EXERCICE 13.18

1. f est \mathcal{C}^1 car elle est polynomiale. On a alors

$$\forall i \in \llbracket 1, n \rrbracket, \partial_i f(x_1, \dots, x_n) = 2x_i - 2 \left(1 - \sum_{k=1}^n x_k \right).$$

Ainsi, (x_1, \dots, x_n) est un point critique de f si et seulement si

$$\begin{cases} x_1 - (1 - \sum_{k=1}^n x_k) = 0 \\ x_2 - (1 - \sum_{k=1}^n x_k) = 0 \\ \vdots \\ x_n - (1 - \sum_{k=1}^n x_k) = 0 \end{cases} \Leftrightarrow \begin{cases} x_1 = (1 - \sum_{k=1}^n x_k) \\ x_2 = (1 - \sum_{k=1}^n x_k) \\ \vdots \\ 2x_n = (1 - \sum_{k=1}^n x_k) \end{cases} \Leftrightarrow \begin{cases} x_1 = (1 - \sum_{k=1}^n x_k) \\ x_2 = x_1 \\ \vdots \\ x_n = x_1 \end{cases}$$

Mais alors la première équation devient alors, par substitution,

$$x_1 = 1 - nx_1 \Leftrightarrow (n+1)x_1 = 1 \Leftrightarrow x_1 = \frac{1}{n+1}.$$

Donc f admet un unique point critique $(a_1, \dots, a_n) = \left(\frac{1}{n+1}, \dots, \frac{1}{n+1}\right)$.

2. Soit $(h_1, \dots, h_n) \in \mathbf{R}^n$. Alors on a

$$\begin{aligned} f(a_1 + h_1, \dots, a_n + h_n) - f(a_1, \dots, a_n) &= \sum_{k=1}^n (a_k + h_k)^2 + \left(1 - \sum_{k=1}^n (a_k + h_k)\right)^2 - f(a_1, \dots, a_n) \\ &= \sum_{k=1}^n a_k^2 + 2 \sum_{k=1}^n a_k h_k + \sum_{k=1}^n h_k^2 - \left(1 - \sum_{k=1}^n (a_k + h_k)\right)^2 - \sum_{k=1}^n a_k^2 + \left(1 - \sum_{k=1}^n a_k\right)^2 \\ &= 2 \sum_{k=1}^n a_k h_k + \sum_{k=1}^n h_k^2 - \left(2 - 2 \sum_{k=1}^n a_k - \sum_{k=1}^n h_k\right) \left(\sum_{k=1}^n h_k\right) \\ &= \frac{2}{n+1} \sum_{k=1}^n h_k + \sum_{k=1}^n h_k^2 - \left(\frac{2}{n+1} - \sum_{k=1}^n h_k\right) \left(\sum_{k=1}^n h_k\right) \\ &= \sum_{k=1}^n h_k^2 + \left(\sum_{k=1}^n h_k\right)^2 \geq 0 \end{aligned}$$

Méthode

Tout élément de \mathbf{R}^n peut s'écrire sous la forme

$$(x_1, \dots, x_n) = (a_1 + h_1, \dots, a_n + h_n).$$

Il suffit en effet de poser $h_i = x_i - a_i$.

Somme de carrés.

Et donc f admet un minimum en (a_1, \dots, a_n) .

SOLUTION DE L'EXERCICE 13.19

1.a. Il s'agit de revenir à la définition : soient $t_1, t_2 \in \mathbf{R}$ et $\lambda \in [0, 1]$. Alors

$$\begin{aligned}\varphi_{h,x}(\lambda t_1 + (1-\lambda)t_2) &= f(x + (\lambda t_1 + (1-\lambda)t_2)h) \\ &= f(\lambda(x + t_1 h) + (1-\lambda)(x + t_2 h)) \\ &\leq \lambda f(x + t_1 h) + (1-\lambda)f(x + t_2 h) = \lambda \varphi_{x,h}(t_1) + (1-\lambda)\varphi_{x,h}(t_2).\end{aligned}$$

Donc $\varphi_{x,h}$ est convexe.

1.b. $\varphi_{h,x}$ étant convexe, ses cordes sont au dessus de ses tangentes.
En particulier, la tangente en 0 est la droite d'équation $y = \varphi'_{h,x}(0)t + \varphi_{h,x}(0)$.
Donc pour $t = 1$, on a $\varphi_{h,x}(1) \geq \varphi'_{h,x}(0) + \varphi_{h,x}(0)$ soit encore

$$\varphi'_{h,x}(0) \leq \varphi_{h,x}(1) - \varphi_{h,x}(0).$$

2. On sait que $\varphi'_{h,x}(0) = \langle \nabla f(x), h \rangle$. En particulier, si on prend $h = y - x$, alors il vient

$$\langle \nabla f(x), y - x \rangle \leq \varphi_{h,x}(0) - \varphi_{h,x}(1) = f(x) - f(x + (y - x)) = f(x) - f(y).$$

En particulier, si x est un point critique de f , $\nabla f(x) = 0$ et donc pour tout $y \in \mathbf{R}^n$,
 $0 \leq f(y) - f(x) \Leftrightarrow f(x) \leq f(y)$.
 f possède alors un minimum en x .

3. Pour $x = 0$, la relation précédente donne $\forall y \in \mathbf{R}^n$, $0 \leq f(y)$. S'il existait un $x \neq 0$ tel que $f(x) = 0$, alors il viendrait

$$f\left(\frac{1}{2} \times 0 + \frac{1}{2}x\right) < \frac{1}{2}f(0) + \frac{1}{2}f(x) \leq 0.$$

Ceci est contradictoire avec $f(y) \geq 0, \forall y \in \mathbf{R}^n$.

Plus généralement

Pour une fonction strictement convexe, le minimum, s'il existe, ne peut être atteint qu'en un unique point.

CONVERGENCE DES VARIABLES ALÉATOIRES

Cours 427
14.1 Convergence en probabilité 427
14.2 Convergence en loi 432
14.3 Théorème central limite et approximations 438
Méthodes 444
Exercices 445
Correction des exercices 448

14.1 CONVERGENCE EN PROBABILITÉ

14.1.1 Inégalités de concentration

Proposition 14.1 (Inégalité de Markov) : Soit X une variable aléatoire **positive** sur (Ω, \mathcal{A}, P) , discrète ou à densité admettant une espérance. Alors,

$$\forall a > 0, P(X \geq a) \leq \frac{E(X)}{a}.$$

Démonstration. Ce résultat a déjà été démontré en première année dans le cas des variables discrètes. Supposons donc que X admet une densité f . Puisque X est à valeurs positives, on peut supposer¹ que $f(x) = 0$ pour $x < 0$. On a alors

$$E(X) = \int_0^{+\infty} t f(t) dt = \int_0^a t f(t) dt + \int_a^{+\infty} t f(t) dt \geq \int_0^a t f(t) dt + \int_a^{+\infty} a f(t) dt$$

car pour tout $t \geq a$, $t f(t) \geq a f(t)$.

De plus, pour tout $t \in [0, a]$, $t f(t) \geq 0$, et donc $\int_0^a t f(t) dt \geq 0$.

On en déduit que

$$E(X) \geq a \int_a^{+\infty} f(t) dt = a P(X \geq a).$$

Et donc

$$P(X \geq a) \leq \frac{E(X)}{a}.$$

□

Notons que cette inégalité est plutôt grossière. Par exemple, si X suit une loi exponentielle $\mathcal{E}(\lambda)$, alors grâce à l'inégalité de Markov, on a

$$P(X \geq a) \leq \frac{E(X)}{a} = \frac{1}{\lambda a}.$$

Mais nous savons que

$$P(X \geq a) = 1 - P(X \leq a) = 1 - (1 - e^{-\lambda a}) = e^{-\lambda a}$$

▲ Pour la culture

Cette inégalité, et celles qui vont suivre sont souvent appelées des «inégalités de concentration». En effet, elles donnent une idée de la manière dont les probabilités «se concentrent» autour de telle ou telle partie de \mathbf{R} . L'inégalité de Markov quantifie la manière dont les valeurs prises par X «se concentrent» autour de 0. Et l'inégalité de Bienaymé-Tchebychev, qui va suivre, explique comment les valeurs prises par X se concentrent autour de $E(X)$.

¹ Il existe une telle densité nulle sur \mathbf{R}_- , même si ce n'est pas forcément le cas de toute densité de X .

et que $e^{-\lambda a} = \underset{a \rightarrow +\infty}{o} \left(\frac{1}{\lambda a} \right)$.

Il y a donc un écart conséquent entre la borne donnée par l'inégalité de Markov et la valeur exacte de $P(X \geq a)$, cette dernière étant bien plus petite que $\frac{E(X)}{a}$.

Toutefois, cette inégalité a l'avantage d'être très générale et de s'appliquer pour des variables dont on ne connaîtrait que l'espérance, et pour lesquelles il est difficile d'estimer la valeur exacte de $P(X \geq a)$.

Corollaire 14.2 – Soit X une variable aléatoire (discrète ou à densité) sur (Ω, \mathcal{A}, P) admettant un moment d'ordre 2. Alors

$$\forall a > 0, P(|X| \geq a) \leq \frac{E(X^2)}{a^2}.$$

Démonstration. La variable X^2 est positive et admet une espérance, et donc l'inégalité de Markov s'applique

$$\forall a > 0, P(X^2 \geq a^2) \leq \frac{E(X^2)}{a^2}.$$

Mais $P(|X| \geq a) = P(X^2 \geq a^2)$, et donc

$$P(|X| \geq a) \leq \frac{E(X^2)}{a^2}.$$

□

Corollaire 14.3 – De manière plus générale, si X possède un moment d'ordre r , alors

$$\forall a > 0, P(|X| \geq a) \leq \frac{E(|X|^r)}{a^r}.$$

Preuve

La preuve est la même que celle du corollaire précédent, en remplaçant 2 par r .

Proposition 14.4 (Inégalité de Bienaymé-Tchebychev) : Soit X une variable aléatoire sur (Ω, \mathcal{A}, P) admettant une variance. Alors,

$$\forall \varepsilon > 0, P(|X - E(X)| \geq \varepsilon) \leq \frac{V(X)}{\varepsilon^2}.$$

Interprétation

Cette inégalité permet de quantifier la probabilité qu'une variable X prenne une valeur éloignée de la moyenne : plus $V(X)$ est petit, ou plus ε est grand, plus cette probabilité est faible.

Démonstration. La variable aléatoire $|X - E(X)|$ admet un moment d'ordre 2 car

$$E(|X - E(X)|^2) = E((X - E(X))^2) = V(X).$$

Par conséquent, le corollaire de l'inégalité de Markov s'écrit

$$\forall \varepsilon > 0, P(|X - E(X)| \geq \varepsilon) \leq \frac{E((X - E(X))^2)}{\varepsilon^2} = \frac{V(X)}{\varepsilon^2}.$$

□

Remarques. • Comme l'inégalité de Markov, l'inégalité de Bienaymé-Tchebychev est en fait assez grossière², mais a le mérite d'être très générale et de ne pas nécessiter de calculs précis sur la loi de X : son espérance et sa variance suffisent.

• L'inégalité de Markov et l'inégalité de Bienaymé-Tchebychev restent valables si on remplace les inégalités larges dans les probabilités par des inégalités strictes. En effet, $[X > a] \subset [X \geq a]$ et donc $P(X > a) \leq P(X \geq a)$. Et de même $P(|X - E(X)| > \varepsilon) \leq P(|X - E(X)| \geq \varepsilon)$.

² Au sens où le majorant est souvent bien plus grand que la valeur de $P(|X - E(X)| \geq \varepsilon)$.

Exemple 14.5

Soit X une variable aléatoire suivant la loi normale centrée réduite. Alors, pour tout $x > 0$,

$$\begin{aligned} p(|X - E(X)| \geq x) &= P(|X| \geq x) = P([X \leq -x] \cup [X \geq x]) \\ &= P(X \leq -x) + P(X \geq x) = \Phi(-x) + 1 - \Phi(x) = 2 - 2\Phi(x). \end{aligned}$$

D'après l'inégalité de Bienaymé-Tchebychev, on a alors

$$2 - 2\Phi(x) \leq \frac{1}{x^2} \Leftrightarrow 1 - \Phi(x) \leq \frac{1}{2x^2}.$$

Exemple 14.6

Combien de fois faut-il lancer un dé équilibré pour garantir avec moins de 5% d'erreur que la fréquence d'apparition du 1 sera $\frac{1}{6} \pm 0.01$?

Soit n le nombre de lancers, et soit X le nombre de lancers dont le résultat est 1.

La fréquence d'apparition du 1 est $Y = \frac{X}{n}$.

Nous cherchons donc quelle valeur donner à n pour avoir

$$P\left(\left|Y - \frac{1}{6}\right| < \frac{1}{100}\right) > 0.95.$$

Autrement dit, nous voulons $P\left(\left|Y - \frac{1}{6}\right| \geq 0.01\right) \leq 0.05$.

X suit une loi binomiale $\mathcal{B}(n, \frac{1}{6})$, de sorte que $E(X) = \frac{n}{6}$ et $V(X) = n \frac{5}{36}$.

Et donc $E(Y) = \frac{1}{6}$ et $V(Y) = \frac{1}{n^2} V(X) = \frac{5}{36n}$.

D'après l'inégalité de Bienaymé-Tchebychev,

$$P\left(\left|Y - \frac{1}{6}\right| \geq 0.01\right) \leq \frac{50000}{36n}.$$

Ainsi, pour $n \geq \frac{50000}{36 \times 0.05} \Leftrightarrow n \geq 27\,778$, on a l'inégalité souhaitée.

Cette inégalité est loin d'être optimale, et un calcul détaillé avec Sci Lab, en utilisant la loi binomiale, prouve qu'en fait $n \geq 5395$ suffit.

14.1.2 Convergence en probabilité

Définition 14.7 – Soit $(X_n)_{n \in \mathbb{N}^*}$ une suite de variables aléatoires définies sur (Ω, \mathcal{A}, P) et soit X une variable aléatoire sur (Ω, \mathcal{A}, P) . On dit que la suite (X_n) converge en probabilité vers la variable aléatoire X si

$$\forall \varepsilon > 0, \lim_{n \rightarrow +\infty} P(|X_n - X| \geq \varepsilon) = 0.$$

On note alors $X_n \xrightarrow{P} X$.

Interprétation

Cette définition très formelle n'empêche pas de laisser un peu de place à l'intuition : elle signifie que quel que soit $\varepsilon > 0$ la «marge d'erreur» que l'on se donne, la probabilité que X_n et X prennent des valeurs différant de plus de ε tend vers 0 (i.e. devient très petite) lorsque n tend vers $+\infty$.

Remarque. On peut aussi bien remplacer les inégalités larges ($\geq \varepsilon$) par des inégalités strictes ($> \varepsilon$).

Exercice : le prouver.

Exemple 14.8 Maximum de lois uniformes

Soit (X_n) une suite de variables aléatoires indépendantes suivant toutes la même loi uniforme sur $[0, 1]$, et soit $Y_n = \max(X_1, \dots, X_n)$.

On a alors : $\forall x \in [0, 1]$,

$$P(Y_n \leq x) = P\left(\bigcap_{i=1}^n [X_i \leq x]\right) = \prod_{i=1}^n P(X_i \leq x) = x^n.$$

Et donc pour $0 < \varepsilon < 1$,

$$P(|Y_n - 1| \geq \varepsilon) = P(1 - Y_n \geq \varepsilon) = P(Y_n \leq 1 - \varepsilon) = (1 - \varepsilon)^n.$$

Mais pour $0 < \varepsilon < 1$, $0 < 1 - \varepsilon < 1$ et donc $P(|Y_n - 1| \geq \varepsilon) \xrightarrow[n \rightarrow +\infty]{} 0$.

Ainsi (Y_n) converge en probabilité vers la variable aléatoire certaine égale à 1.

Ce résultat était prévisible : plus n est grand, plus les chances qu'au moins l'un des n nombres X_1, \dots, X_n soit proche de 1 sont élevées.

Valeur absolue

Y_n est toujours inférieure ou égale à 1, de sorte que $1 - Y_n \geq 0$ et donc

$$|Y_n - 1| = 1 - Y_n.$$

Remarques. • Nous avons donné le détail uniquement pour $\varepsilon < 1$ car nous évitons ainsi une distinction de cas. Ceci est largement suffisant, la partie importante du $\forall \varepsilon > 0$ est en fait «pour tout ε suffisamment petit».

Le fait que la définition soit vérifiée pour $\varepsilon = 1000$, $\varepsilon = 10$ ou même $\varepsilon = 0.01$ n'est pas très importante, ce qui compte c'est qu'elle le soit pour les ε dans un petit intervalle de la forme $]0, \alpha]$, $\alpha > 0$.

• Nous avons là un exemple de variables aléatoires à densité qui convergent en probabilité vers une variable aléatoire discrète, le contraire étant également possible.

Proposition 14.9 : Soit (X_n) une suite de variables aléatoires sur (Ω, \mathcal{A}, P) qui converge en probabilité vers une variable aléatoire X . Si f est une fonction continue sur un intervalle I tel que $P(X \in I) = 1$, alors la suite $(f(X_n))_n$ converge en probabilité vers la variable aléatoire $f(X)$.

Démonstration. Admis. □

Une version imprécise de ce résultat mais souvent suffisante est la suivante : si $X_n \xrightarrow{P} X$ et si f est continue, alors $f(X_n) \xrightarrow{P} f(X)$.

14.1.3 La loi faible des grands nombres

Théorème 14.10 (Loi faible des grands nombres) : Soit (X_n) une suite de variables aléatoires *indépendantes* définies sur un même espace probabilisé (Ω, \mathcal{A}, P) , admettant une *même espérance* m et une *même variance* σ^2 . Si on pose $\bar{X}_n = \frac{X_1 + \dots + X_n}{n}$, alors (\bar{X}_n) converge en probabilité vers la variable aléatoire certaine égale à m . Autrement dit

$$\bar{X}_n \xrightarrow{P} m.$$

Cas particulier

Le plus souvent, on applique ce théorème dans le cas où les X_i suivent la même loi (et ont un moment d'ordre 2).

Démonstration. Par linéarité de l'espérance, \bar{X}_n admet une espérance, et

$$E(\bar{X}_n) = \frac{1}{n} \sum_{i=1}^n E(X_i) = \frac{1}{n} \sum_{i=1}^n m = m.$$

De plus, les X_i étant indépendantes, et de variance σ^2 , alors \bar{X}_n admet une variance, et

$$V(\bar{X}_n) = \frac{1}{n^2} \sum_{i=1}^n V(X_i) = \frac{\sigma^2}{n}.$$

Par l'inégalité de Bienaymé-Tchebychev, on a alors

$$\forall \varepsilon > 0, 0 \leq P(|\bar{X}_n - E(\bar{X}_n)| \geq \varepsilon) = P(|\bar{X}_n - m| \geq \varepsilon) \leq \frac{V(\bar{X}_n)}{\varepsilon^2} = \frac{\sigma^2}{n\varepsilon^2} \xrightarrow[n \rightarrow +\infty]{} 0.$$

Ceci prouve bien que $\bar{X}_n \xrightarrow{P} m$. □

Remarque. Le résultat est en fait assez intuitif, et nous l'avons déjà utilisé de nombreuses fois, notamment lors de simulation avec SciLab : si on souhaite estimer la valeur de l'espérance d'une variable X , on simule n réalisations de cette variable (ce sont les X_1, \dots, X_n de l'énoncé) et on calcule leur moyenne (c'est le $\overline{X_n}$). Nous savons que l'espérance n'est pas exactement égale à cette moyenne, mais que plus n est grand, moins on a de chances³ de se tromper.

En particulier, lorsque les X_i sont des variables de Bernoulli, on a le résultat suivant :

Corollaire 14.11 (Théorème de Bernoulli) – Soit (X_n) une suite de variables de Bernoulli indépendantes de même paramètre p . Alors si on note $\overline{X_n} = \frac{X_1 + \dots + X_n}{n}$, alors $\overline{X_n} \xrightarrow{P} p$.

Remarques. • Ce théorème est en fait fondamental : c'est lui qui justifie que le formalisme que nous employons correspond bien à l'intuition que l'on se fait d'une probabilité : la probabilité d'un événement est proche de la fréquence de réalisation de cet événement lorsqu'on répète un grand nombre de fois l'expérience.

• De plus, la preuve de ce théorème via l'inégalité de Bienaymé-Tchebychev nous donne un ordre de grandeur de la vitesse de convergence :

$$\forall \varepsilon > 0, P(|\overline{X_n} - p| \geq \varepsilon) \leq \frac{p(1-p)}{n\varepsilon^2} \leq \frac{1}{4n\varepsilon^2}.$$

Ce dernier point vient du fait que la fonction $x \mapsto x(1-x)$ a pour maximum $\frac{1}{4}$ sur $[0, 1]$. Au final, cette inégalité nous permet de donner un ordre de grandeur du nombre de fois où il faudrait répéter une expérience pour approcher p avec une marge d'erreur fixée.

³ C'est exactement le sens de la convergence en probabilité : plus n est grand, plus la possibilité d'avoir un écart donné avec l'espérance est faible.

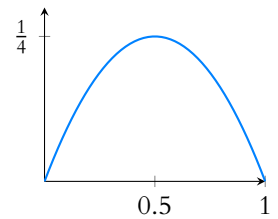


FIGURE 14.1– La fonction $x \mapsto x(1-x)$.

Exemple 14.12

On dispose d'une pièce truquée, et on souhaite évaluer la probabilité p qu'elle tombe sur pile.

La première idée est de la lancer n fois, et si l'on note X_i la variable de Bernoulli valant 1 si la pièce tombe sur pile au i -ème lancer, alors on prendra $\overline{X_n} = \frac{X_1 + \dots + X_n}{n}$

comme approximation de p .

Cette intuition est confirmée par la loi des grands nombres, qui nous dit qu'on a bien $\overline{X_n} \xrightarrow{P} p$, mais on n'est jamais totalement à l'abri d'une série de lancers où on « n'aurait pas eu de chance », c'est-à-dire telle que $\overline{X_n}$ prenne une valeur éloignée de la valeur de p .

On peut par exemple souhaiter approcher p à 0.1 près, tout en étant sûr à 99 % du résultat⁴.

Autrement dit, on souhaite avoir

$$P(|\overline{X_n} - p| < 0.1) \geq 0.99.$$

Soit encore

$$P(|\overline{X_n} - p| \geq 0.1) < 0.01.$$

On ne peut utiliser la majoration $\frac{p(1-p)}{n \times 0.1^2}$, car on ne connaît pas p . Donc utilisons plutôt

$$P(|\overline{X_n} - p| \geq 0.1) \leq \frac{1}{4n \times 0.1^2} = \frac{25}{n}.$$

Si $\frac{25}{n} \leq 0.01 \Leftrightarrow n \geq 2500$, alors on a bien

$$P(|\overline{X_n} - p| \geq 0.1) \leq 0.01.$$

Si on réalise plus de 2500 lancers, on peut donc être certain à 99% que la valeur de $\overline{X_n}$ approche celle de p à 0.1 près.

⁴ On ne peut jamais être sûr à 100% du résultat, il existe toujours un risque de jouer de malchance lors des lancers !

Incertitude
Cela signifie qu'il y a toujours au plus 1% de chance que la valeur obtenue ne soit pas la bonne.

14.1.4 La méthode de Monte-Carlo

La loi faible des grands nombres peut nous aider à estimer des valeurs approchées d'intégrales que nous ne savons pas calculer.

Soit $f : [0, 1] \rightarrow \mathbf{R}$ une fonction continue, et soit $(U_i)_{i \in \mathbf{N}}$ une suite de variables aléatoires indépendantes suivant la loi uniforme sur $[0, 1]$. Notons alors $X_i = f(U_i)$, de sorte que les X_i sont indépendantes et de même loi.

Par le théorème de transfert, on a alors $E(X_i) = \int_0^1 f(t) dt$ et $E(X_i^2) = \int_0^1 f(t^2) dt$. En

particulier, les X_i possèdent un moment d'ordre 2 et donc une variance.

Par la loi faible des grands nombres, on a donc

$$\frac{X_1 + \cdots + X_n}{n} \xrightarrow{P} E(X_1) = \int_0^1 f(t) dt.$$

Autrement dit, pour calculer une valeur approchée de $\int_0^1 f(t) dt$, on peut simuler (par exemple à l'aide de Scilab) un grand nombre de fois les variables U_i , calculer les $f(U_i)$ correspondants et faire leur moyenne.

Exemple 14.13

Le code suivant nous permet de calculer une valeur approchée de

$$\int_0^1 \frac{1}{\sqrt{2\pi}} e^{-t^2/2} dt = \Phi(1) - \frac{1}{2}.$$

```

1 function y = f(x)
2     y = 1/sqrt(2*pi)*exp(-x^2/2)
3 endfunction
4 U = grand(1000,1,'unf',0,1);
5 S = 0;
6 for i = 1 :1000
7     S = S+f(U(i))
8 end
9 S = S/1000;
```

On obtient alors à peu près 0.34.

⚠ Comme toujours lors de simulations, nous ne sommes pas à l'abri d'un «mauvais» tirage, qui donnerait alors un résultat assez éloigné de la valeur réelle de l'intégrale. Toutefois, plus n est grand moins ceci a de chances de se produire.

Toutes les méthodes de ce type qui permettent d'approcher une intégrale à l'aide de la loi faible des grands nombres sont appelées méthodes de Monte-Carlo⁵.

14.2 CONVERGENCE EN LOI

14.2.1 Définition, premières propriétés

Définition 14.14 – Soit (X_n) une suite de variables aléatoires sur (Ω, \mathcal{A}, P) , et soit X une variable aléatoire définie sur (Ω, \mathcal{A}, P) . On dit que la suite (X_n) converge en loi vers la variable aléatoire X si pour tout réel x où F_X est continue, on a

$$\lim_{n \rightarrow +\infty} F_{X_n}(x) = F_X(x).$$

On note alors $X_n \xrightarrow{\mathcal{L}} X$.

Remarque

Notons qu'il s'agit d'un analogue aléatoire aux sommes de Riemann : dans une somme de Riemann, on calcule les valeurs de f en n points répartis régulièrement dans le segment $[0, 1]$, puis on moyenne ces n valeurs. C'est la même chose ici, sauf que nos n points sont choisis aléatoirement.

⁵ Ces méthodes sont appelées ainsi car elles font appel au hasard, et sont donc un clin d'œil aux casinos monégasques. Notons qu'il existe d'autres algorithmes probabilistes dits de Las Vegas et d'Atlantic City.

Exemple 14.15

Soit $(X_n)_{n \in \mathbf{N}^*}$ une suite de variables aléatoires telles que X_n suit la loi exponentielle de paramètre n .

Alors, pour tout $x \in \mathbf{R}$, $F_{X_n}(x) = \begin{cases} 0 & \text{si } x \leq 0 \\ 1 - e^{-nx} & \text{si } x > 0 \end{cases}$.

Ainsi, $\lim_{n \rightarrow +\infty} F_{X_n}(x) = \begin{cases} 0 & \text{si } x \leq 0 \\ 1 & \text{si } x > 0 \end{cases}$.

Notons que ceci n'est pas une fonction de répartition, car elle n'est pas continue à droite en 0.

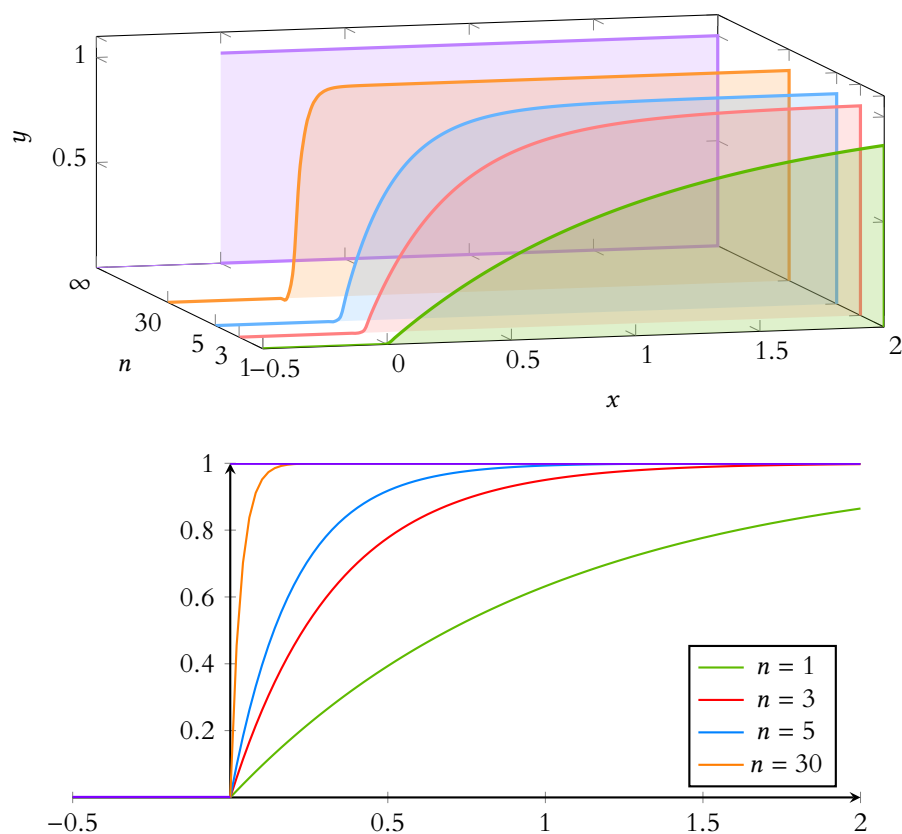
Mais $F_X : x \mapsto \begin{cases} 0 & \text{si } x < 0 \\ 1 & \text{si } x \geq 0 \end{cases}$ est bien une fonction de répartition : c'est la fonction

de répartition d'une variable aléatoire constante égale à 0.

Puisque cette fonction est continue sur \mathbf{R}^* , on a bien

$$\forall x \in \mathbf{R}^*, \lim_{n \rightarrow +\infty} F_{X_n}(x) = F_X(x).$$

Autrement dit, en tout point de continuité de F_X , $\lim_{n \rightarrow +\infty} F_{X_n}(x) = F(x)$ et donc (X_n) converge en loi vers la variable constante égale à 0.



Remarques. • Même si nous avons demandé aux X_n et à X d'être définies sur le même espace probabilisé, cela n'est pas indispensable, puisque le passage à la limite se fait dans tous les cas sur les fonctions de répartition, qui sont définies sur \mathbf{R} .

• Il n'y a pas unicité de la limite. D'ailleurs, la variable aléatoire X n'a pas d'importance, tout ce qui compte est sa loi. Et donc si $X_n \xrightarrow{\mathcal{L}} X$ et si Y est une autre variable de même loi que X , alors $X_n \xrightarrow{\mathcal{L}} Y$.

Ceci donne lieu à un certain nombre de curiosités : par exemple, si $X \leftrightarrow \mathcal{N}(0, 1)$, alors $-X$ suit aussi la même loi.

Et donc si $X_n \xrightarrow{\mathcal{L}} X$, alors on a également $X_n \xrightarrow{\mathcal{L}} -X$, et ce bien que $X \neq 0$.

Méthode : Montrer une convergence en loi

Pour étudier la convergence en loi d'une suite (X_n) de variables aléatoires :

- Déterminer les fonctions de répartition des X_n .
- Étudier, à $x \in \mathbf{R}$ fixé la convergence de la suite (de réels) $(F_{X_n}(x))_n$.
- Si on trouve une limite $F(x)$ pour tout x ou presque, vérifier que la fonction F ainsi définie est bien la fonction de répartition d'une variable aléatoire : en particulier, on doit avoir F continue à droite en tout point et

$$\lim_{x \rightarrow -\infty} F(x) = 0 \text{ et } \lim_{x \rightarrow +\infty} F(x) = 1.$$

- Si ces conditions sont vérifiées et si la convergence a bien lieu en tout point de continuité de F , alors (X_n) converge en loi vers toute variable aléatoire de fonction de répartition F .

Exemples 14.16

- Soit (X_n) une suite de variables aléatoires discrètes telles que X_n suit la loi uniforme sur $\llbracket 1, n \rrbracket$.

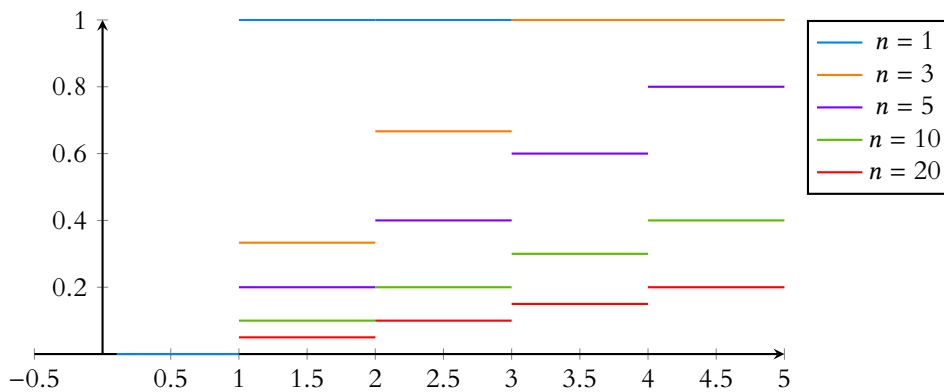
$$\text{Alors } F_{X_n}(x) = P(X_n \leq x) = P(X_n \leq \lfloor x \rfloor) = \begin{cases} 0 & \text{si } x < 1 \\ \frac{\lfloor x \rfloor}{n} & \text{si } x \in [1, n] \\ 1 & \text{si } x > n \end{cases}$$

Alors pour $x < 1$, on a $\lim_{n \rightarrow +\infty} F_{X_n}(x) = 0$.

Pour $x > 1$ et pour $n \geq \lfloor x \rfloor + 1$, on a $F_{X_n}(x) = \frac{\lfloor x \rfloor}{n} \rightarrow 0$.

Donc pour tout $x \in \mathbf{R}$, $\lim_{n \rightarrow +\infty} F_{X_n}(x) = 0$.

Mais la fonction nulle n'est pas la fonction de répartition d'une variable aléatoire⁶, et donc la suite $(X_n)_n$ ne converge pas en loi.



- Pour tout $n \in \mathbf{N}^*$, soit X_n une variable aléatoire de fonction de répartition

$$F_{X_n}(x) = \frac{1}{(1 + e^{-x - \ln n})^n}.$$

Alors pour x fixé, on a

$$(1 + e^{-x - \ln n})^n = e^{n \ln(1 + \frac{e^{-x}}{n})}$$

Or $\frac{e^{-x}}{n} \rightarrow 0$, de sorte que $\ln(1 + \frac{e^{-x}}{n}) \underset{n \rightarrow +\infty}{\sim} \frac{e^{-x}}{n}$, et donc

$$n \ln\left(1 + \frac{e^{-x}}{n}\right) \underset{n \rightarrow +\infty}{\sim} n \frac{e^{-x}}{n} = e^{-x}.$$

Limite

L'expression donnée n'est valable que pour $n \geq \lfloor x \rfloor + 1$, c'est-à-dire pour n suffisamment grand. C'est suffisant pour déterminer la limite, car celle-ci ne dépend pas des «petites» valeurs de n , mais bien uniquement de ce qui se passe pour n «suffisamment grand».

⁶ Et contrairement à l'exemple précédent, changer sa valeur en un nombre fini de points ne suffira pas à en faire une fonction de répartition, on aura toujours une limite nulle en $+\infty$.

Précision

Il s'agit bien d'une fonction de répartition car :

- $\lim_{x \rightarrow -\infty} F_n(x) = 0$
- $\lim_{x \rightarrow +\infty} F_n(x) = 1$
- F_n est croissante
- F_n est continue sur \mathbf{R} et donc en particulier continue à droite en tout point.

De plus, F_n est \mathcal{C}^1 sur \mathbf{R} , donc c'est la fonction de répartition d'une variable à densité.

Alors $F_{X_n}(x) \xrightarrow{n \rightarrow +\infty} e^{-e^{-x}}$.

La fonction $x \mapsto e^{e^{-x}}$ est \mathcal{C}^1 sur \mathbf{R} , elle y est croissante, et

$$\lim_{x \rightarrow -\infty} e^{e^{-x}} = 0 \text{ et } \lim_{x \rightarrow +\infty} e^{-e^{-x}} = 1.$$

C'est donc bien la fonction de répartition d'une variable (à densité) et $(X_n)_n$ converge en loi vers toute variable suivant cette loi⁷.

⁷ Cette loi est appelée loi de Gumbel.

Proposition 14.17 : Si $X_n \xrightarrow{\mathcal{L}} X$, alors pour tous $a < b \in \mathbf{R}^2$, si a et b sont des points de continuité de F_X , alors

$$\lim_{n \rightarrow +\infty} P(a < X_n \leq b) = P(a < X \leq b).$$

Démonstration. On a

$$P(a < X_n \leq b) = F_{X_n}(b) - F_{X_n}(a) \xrightarrow{n \rightarrow +\infty} F_X(b) - F_X(a) = P(a < X \leq b).$$

□



Proposition 14.18 : Soient $(X_n)_{n \in \mathbf{N}^*}$ et X des variables aléatoires définies sur (Ω, \mathcal{A}, P) . Si $X_n \xrightarrow{P} X$, alors $X_n \xrightarrow{\mathcal{L}} X$.

Remarques. • Avoir ce résultat en tête permet toutefois de se faire une idée du résultat que l'on attend lorsqu'on cherche à prouver une convergence en probabilité sans aucune idée du résultat attendu. En effet, si $X_n \xrightarrow{P} X$, alors $X_n \xrightarrow{\mathcal{L}} X$. Et la convergence en loi étant plus facile à étudier (il s'agit d'un problème de limites), on peut commencer par déterminer si la suite (X_n) converge en loi, ce qui permet au moins de déterminer la loi de l'éventuelle limite en probabilité.

• La réciproque est fautive : il se peut que $X_n \xrightarrow{\mathcal{L}} X$ et que (X_n) ne converge pas en probabilité vers X . Par exemple, considérons une variable aléatoire X suivant la loi $\mathcal{N}(0, 1)$, et pour tout n , posons $X_n = X$. Alors la suite $X_n \xrightarrow{\mathcal{L}} X$. Mais si $Y = -X$, alors $Y \hookrightarrow \mathcal{N}(0, 1)$, et donc on a également $X_n \xrightarrow{\mathcal{L}} Y$.

Mais (X_n) ne converge pas en probabilité vers Y car $X_n - Y = 2X$ et donc pour $\varepsilon = 2$, on a

$$\begin{aligned} P(|X_n - Y| \geq 2) &= P(|2X| \geq 2) = P(2X \leq -2) + P(2X \geq 2) \\ &= P(X \leq -1) + P(X \geq 1) = \Phi(-1) + 1 - \Phi(1) = 2 - 2\Phi(1) \not\rightarrow 0. \end{aligned}$$

14.2.2 Cas des variables aléatoires à valeurs entières

Proposition 14.19 : Soient $(X_n)_{n \in \mathbf{N}}$ et X des variables aléatoires sur (Ω, \mathcal{A}, P) , à valeurs dans \mathbf{Z} . Alors $X_n \xrightarrow{\mathcal{L}} X$ si et seulement si

$$\forall k \in \mathbf{Z}, \lim_{n \rightarrow +\infty} P(X_n = k) = P(X = k).$$



Démonstration. Supposons que $X_n \xrightarrow{\mathcal{L}} X$. Alors pour tout $k \in \mathbf{Z}$, F_X est continue en $k + \frac{1}{2}$ et $k - \frac{1}{2}$. Donc

$$\begin{aligned} \lim_{n \rightarrow \infty} P(X_n = k) &= \lim_{n \rightarrow +\infty} P\left(k - \frac{1}{2} < X_n \leq k + \frac{1}{2}\right) = \lim_{n \rightarrow +\infty} \left(F_{X_n}\left(k + \frac{1}{2}\right) - F_{X_n}\left(k - \frac{1}{2}\right)\right) \\ &= F_X\left(k + \frac{1}{2}\right) - F_X\left(k - \frac{1}{2}\right) = P(X = k). \end{aligned}$$

Densité

Ce résultat est particulièrement intéressant si X est à densité, car alors on n'a plus à se préoccuper de savoir si a et b sont des points de continuité de F_X .

Hors programme

Ce résultat n'est pas au programme, et n'est donc pas utilisable tel quel. On en donnera toutefois la preuve en TD, exercice issu d'un oral d'HEC

Hypothèses

Il faut bien que les X_i soient à valeurs dans \mathbf{Z} , et pas uniquement discrètes.

Par exemple, si $X \hookrightarrow \mathcal{U}([1, n])$ et si $Y = \frac{X}{n}$ alors

$$X_n(\Omega) = \left\{ \frac{1}{n}, \frac{2}{n}, \dots, \frac{n}{n} \right\}$$

est fini, donc Y est discrète, mais n'est pas à valeurs dans \mathbf{Z} .

Il se peut qu'une suite de variables discrètes converge en loi vers une variable à densité, auquel cas ce résultat ne s'applique plus.

Inversement, supposons que pour tout $k \in \mathbf{Z}$, $P(X_n = k) \xrightarrow{n \rightarrow +\infty} P(X = k)$.

Soit donc $x \in \mathbf{R}$. Il s'agit de prouver que $\lim_{n \rightarrow +\infty} F_{X_n}(x) = F_X(x)$.

Cas particulier : si X et les X_n sont à valeurs dans \mathbf{N} .

Pour $x < 0$, on a évidemment $F_{X_n}(x) = 0 \xrightarrow{n \rightarrow +\infty} 0 = F_X(x)$.

Et pour $x \geq 0$, alors

$$F_{X_n}(x) = P(X_n \leq x) = P(X_n \leq \lfloor x \rfloor) = \sum_{k=0}^{\lfloor x \rfloor} P(X_n = k)$$

$$\xrightarrow{n \rightarrow +\infty} \sum_{k=0}^{\lfloor x \rfloor} P(X = k) = P(X \leq \lfloor x \rfloor) = P(X \leq x) = F_X(x).$$

Cas général : on souhaiterait raisonner comme précédemment en remarquant que

$P(X_n \leq x) = \sum_{k=-\infty}^{\lfloor x \rfloor} P(X_n = k)$, mais malheureusement, s'il était facile de passer à la limite dans le cas précédent⁸, nous ne disposons pas de résultat pour sommer un nombre infini de limites.

La preuve est donc bien plus compliquée dans ce cas général.

⁸ On faisait alors une somme finie de limites.

Soit $\varepsilon > 0$. Puisque F_X est une fonction de répartition, elle est croissante et

$$F_X(x) \xrightarrow{x \rightarrow -\infty} 0 \text{ et } F_X(x) \xrightarrow{x \rightarrow +\infty} 1$$

Il existe donc deux réels A et B tels que $A < x < B$ tels que $F_X(A) \leq \varepsilon$ et $F_X(B) \geq 1 - \varepsilon$.

$F_{X_n}(x) - F_{X_n}(A) = P(A < X_n \leq x) = \sum_{k \in \mathbf{Z} \cap]A, x]} P(X_n = k)$. Notons qu'il s'agit là d'une

somme finie : il n'y a qu'un nombre fini d'entiers entre x et A . Et donc

$$\lim_{n \rightarrow +\infty} F_{X_n}(x) - F_{X_n}(A) = \sum_{k \in \mathbf{Z} \cap]A, x]} P(X = k) = F_X(x) - F_X(A).$$

En particulier, il existe n_0 tel que pour $x \geq n_0$,

$$F_{X_n}(x) - F_{X_n}(A) \geq F_X(x) - F_X(A) - \varepsilon.$$

On en déduit que

$$F_{X_n}(x) \geq \underbrace{F_{X_n}(A)}_{\geq 0} + \underbrace{F_X(x) - F_X(A)}_{\geq -\varepsilon} - \varepsilon \geq F_X(x) - 2\varepsilon.$$

De même, on a

$$F_{X_n}(x) - F_{X_n}(B) = - \sum_{k \in \mathbf{Z} \cap]x, B]} P(X_n = k) \xrightarrow{n \rightarrow +\infty} F_X(x) - F_X(B).$$

Donc il existe un entier $n_1 \in \mathbf{N}$ tel que pour $n \geq n_1$,

$$F_{X_n}(x) - F_{X_n}(B) \leq F_X(x) - F_X(B) + \varepsilon.$$

Alors

$$F_{X_n}(x) \leq F_X(x) + \underbrace{-F_X(B)}_{\leq -1} + \underbrace{F_{X_n}(B)}_{\leq 1} + \varepsilon \leq F_X(x) + 2\varepsilon.$$

Ainsi, pour $n \geq N = \max(n_0, n_1)$, on a

$$F_X(x) - 2\varepsilon \leq F_{X_n}(x) \leq F_X(x) + 2\varepsilon$$

et donc $\lim_{n \rightarrow +\infty} F_{X_n}(x) = F_X(x)$, soit $X_n \xrightarrow{\mathcal{L}} X$. □

Proposition 14.20 : Soit (X_n) une suite de variables aléatoires sur (Ω, \mathcal{A}, P) telles que X_n suit la loi binomiale $\mathcal{B}(n, p_n)$ avec $np_n \xrightarrow{n \rightarrow +\infty} \lambda > 0$.

Alors $X_n \xrightarrow{\mathcal{L}} X$ où X est une variable aléatoire suivant une loi de Poisson $\mathcal{P}(\lambda)$.

Remarque

La condition $np_n \rightarrow \lambda$ est équivalente à

$$p_n \underset{n \rightarrow +\infty}{\sim} \frac{\lambda}{n}.$$

Démonstration. X et les X_n étant à valeurs dans \mathbf{N} , on peut appliquer le résultat précédent. Soit $k \in \mathbf{N}$. On a alors

$$P(X_n = k) = \binom{n}{k} (p_n)^k (1 - p_n)^{n-k} = \frac{n!}{k!(n-k)!} p_n^k (1-p_n)^{n-k} = \frac{n!}{k!(n-k)!} \left(\frac{p_n}{1-p_n}\right)^k (1-p_n)^n.$$

Mais $p_n \underset{n \rightarrow +\infty}{\sim} \frac{\lambda}{n} \rightarrow 0$, de sorte que

$$\left(\frac{p_n}{1-p_n}\right)^k \underset{n \rightarrow +\infty}{\sim} p_n^k \underset{n \rightarrow +\infty}{\sim} \frac{\lambda^k}{n^k}.$$

De plus, à k fixé, $\frac{n!}{(n-k)!} = n(n-1)\cdots(n-k+1)$ est un polynôme en n , équivalent à son terme dominant qui est n^k .

Enfin, $(1-p_n)^n = e^{n \ln(1-p_n)}$ avec $\ln(1-p_n) \underset{n \rightarrow +\infty}{\sim} -p_n \underset{n \rightarrow +\infty}{\sim} -\frac{\lambda}{n}$, de sorte que

$$(1-p_n)^n \xrightarrow{n \rightarrow +\infty} e^{-\lambda}.$$

Au final, on a alors

$$\lim_{n \rightarrow +\infty} P(X_n = k) = \lim_{n \rightarrow +\infty} \frac{n^k \lambda^k}{k! n^k} e^{-\lambda} = e^{-\lambda} \frac{\lambda^k}{k!}.$$

Donc X_n converge bien en loi vers une variable suivant une loi de Poisson $\mathcal{P}(\lambda)$. □

Explication

Ici on ne représente pas les fonctions de répartition, car les X_n et X étant discrètes, on peut se contenter d'étudier les

$$\lim_{n \rightarrow +\infty} P(X_n = k)$$

pour $k \in \mathbf{Z}$.

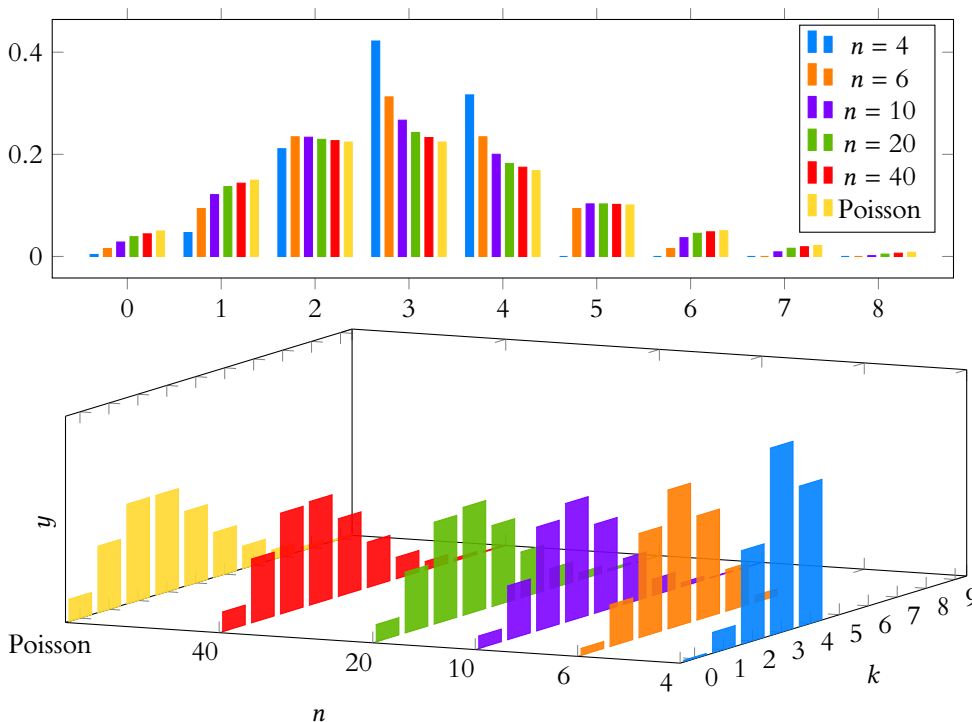


FIGURE 14.2 – Convergence en loi de $\mathcal{B}(n, 3/n)$ vers $\mathcal{P}(3)$

Corollaire 14.21 – Pour $p \leq 0.1$, $n \geq 30$ et $np \leq 15$, on peut approcher la loi binomiale $\mathcal{B}(n, p)$ par la loi de Poisson $\mathcal{P}(np)$.

14.2.3 Convergence en loi et opérations

Attention : si on dispose de deux suites (X_n) et (Y_n) de variables aléatoires convergeant en loi respectivement vers X et Y , alors on ne peut pas, sans plus d'hypothèses, affirmer que $X_n + Y_n \xrightarrow{\mathcal{L}} X + Y$. En revanche, on dispose des deux résultats suivants :

Proposition 14.22 : Soit (X_n) une suite de variables aléatoires sur (Ω, \mathcal{A}, P) qui converge en loi vers une variable aléatoire X . Si f est une fonction continue sur un intervalle I tel que $P(X \in I) = 1$, alors la suite $(f(X_n))_n$ converge en loi vers la variable aléatoire $f(X)$.

Démonstration. La preuve de ce résultat dans le cas général est hors programme, mais nous pouvons tout de même en donner une preuve dans le cas où f est une bijection strictement croissante⁹. En effet, pour $x \in \mathbf{R}$, on a alors

$$P(f(X_n) \leq x) = P(X_n \leq f^{-1}(x)) \xrightarrow{n \rightarrow +\infty} P(X \leq f^{-1}(x)) = P(f(X) \leq x).$$

Ce qui prouve bien que $f(X_n) \xrightarrow{\mathcal{L}} f(X)$. □

Explication

La convergence en loi vers X ne dépend que de la loi de X , c'est-à-dire que si $X_n \xrightarrow{\mathcal{L}} X$ et si X' a même loi que X , alors $X_n \xrightarrow{\mathcal{L}} X'$. En revanche, la loi de $X + Y$ ne dépend pas que des lois de X et de Y , mais de leur loi conjointe (la loi de $X + Y$ ne sera pas la même si X et Y sont indépendantes ou non). Et donc il n'y a aucune raison pour que $X_n + Y_n$ converge vers $X + Y$.

⁹ Par exemple si f est le logarithme ou l'exponentielle.

Proposition 14.23 (Théorème de Slutsky) : Si $X_n \xrightarrow{\mathcal{L}} X$ et si (Y_n) converge en probabilité vers une constante c , alors :

- $(X_n + Y_n)$ converge en loi vers $X + c$
- $(X_n Y_n)$ converge en loi vers cX .

Démonstration. Admis, voir en TD pour une preuve. □

Remarque. En réalité, $(X_n)_n$ converge en probabilité vers une constante c si et seulement si $(X_n)_n$ converge en loi vers c , mais ce résultat n'est pas au programme.

14.3 THÉORÈME CENTRAL LIMITE ET APPROXIMATIONS

14.3.1 Le théorème central limite

Définition 14.24 – Les variables aléatoires $(X_n)_{n \in \mathbf{N}^*}$ sont dites **indépendantes identiquement distribuées** (en abrégé **i.i.d.**) si elles sont indépendantes et suivent la même loi.

Nous avons prouvé précédemment que si les X_n sont i.i.d, d'espérance m et de variance σ^2 , alors $E(\bar{X}_n) = m$ et $V(\bar{X}_n) = \frac{\sigma^2}{n}$.

Et donc si on note \bar{X}_n^* la variable centrée réduite associée à \bar{X}_n , on a

$$\bar{X}_n^* = \frac{\bar{X}_n - m}{\frac{\sigma}{\sqrt{n}}} = \sqrt{n} \frac{\bar{X}_n - m}{\sigma}.$$

D'autre part, si on note $S_n = \sum_{i=1}^n X_i$, alors $E(S_n) = nm$ et $V(S_n) = n\sigma^2$, de sorte que la variable centrée réduite associée à S_n est

$$\frac{S_n - nm}{\sigma \sqrt{n}} = \frac{n\bar{X}_n - nm}{\sigma \sqrt{n}} = \sqrt{n} \frac{\bar{X}_n - m}{\sigma} = \bar{X}_n^*.$$

Autrement dit, la variable centrée réduite associée à S_n est également la variable centrée réduite associée à \bar{X}_n .

Théorème 14.25 (Théorème central limite) : Soit $(X_n)_{n \in \mathbb{N}^*}$ une suite de variables aléatoires indépendantes et de **même loi**, admettant une espérance m et une variance σ^2 non nulle¹⁰. Alors si $\bar{X}_n = \frac{X_1 + \dots + X_n}{n}$, et si \bar{X}_n^* est la variable centrée réduite associée à \bar{X}_n , alors $(\bar{X}_n^*)_n$ converge en loi vers une variable aléatoire X suivant une loi normale centrée réduite.

¹⁰ Rappelons encore une fois que seules les variables certaines ont une variance nulle.

Démonstration. Admis, très difficile. □

Remarque. Ce théorème est aussi parfois appelé théorème de la limite centrée, ou de la limite centrale. Ces dénominations ne sont pas historiquement correctes¹¹ mais communément utilisées, et on les rencontre parfois dans des sujets de concours.

Ce théorème est en fait un raffinement de la loi faible des grands nombres.

En effet, celle-ci nous indique que $\frac{X_1 + \dots + X_n}{n}$ converge en probabilité vers m .

Autrement dit, pour n grand, \bar{X}_n est «proche» de m .

La question qu'on peut se poser est de savoir à quel point la loi de \bar{X}_n «ressemble» à celle de la variable certaine égale à m .

Le théorème central limite affirme que pour n grand $\frac{\sqrt{n}}{\sigma}(\bar{X}_n - m)$ est «proche» d'une variable X suivant la loi normale centrée réduite.

Ainsi, $\bar{X}_n - m \approx \frac{\sigma}{\sqrt{n}}X \Leftrightarrow \mathcal{N}\left(0, \frac{\sigma^2}{n}\right)$, ou encore $\bar{X}_n \approx \mathcal{N}\left(m, \frac{\sigma}{n}\right)$.

¹¹ Ce théorème a été nommé ainsi dans un article en allemand publié par George Pólya où il est très clair que l'adjectif central s'applique au théorème (et pas à la limite), qui est central en théorie des probabilités.

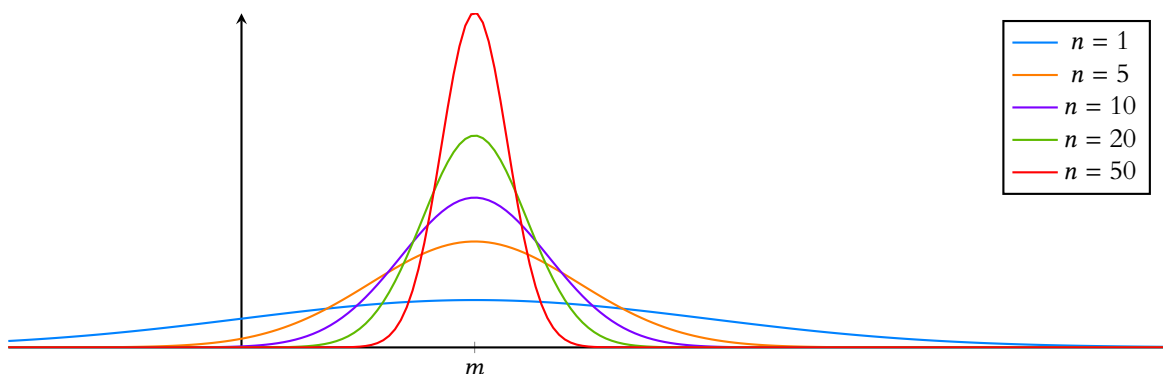


FIGURE 14.3 – La densité de $\mathcal{N}\left(m, \frac{\sigma}{n}\right)$ pour différentes valeurs de n .

Plus n est grand, plus l'écart-type est faible, et plus les valeurs prises par \bar{X}_n sont «resserrées» autour de m .

Cela signifie aussi que $X_1 + \dots + X_n$ possède, pour n grand, une loi proche de $\mathcal{N}(nm, n\sigma)$. Ce résultat explique pourquoi on rencontre si souvent la loi normale «dans la nature» : si l'on somme des variables aléatoires indépendantes et de même loi un grand nombre de fois, la variable aléatoire obtenue est «presque» normale.

Remarque
Pour illustrer la convergence en loi, nous aurions dû regarder les fonctions de répartition et non les densités, mais le résultat se «voit» mieux sur les densités.

Par exemple, nous représentons ci-dessous la densité de $S_n = X_1 + \dots + X_n$ où les X_i sont i.i.d. et suivent la loi uniforme $\mathcal{U}([0, 1])$. On constate que plus n est grand, plus la densité obtenue est proche d'une loi normale¹².

À droite, on a représenté les densités des variables centrées réduites associées aux S_n , qui convergent vers la densité d'une loi normale.

Enfin, notons que dans le cas très particulier où les X_i suivent une loi normale, alors, par stabilité des lois normales, $S_n = X_1 + \dots + X_n$ suit encore une loi normale.

Et donc, par transformation affine de lois normales, la variable centrée réduite associée à S_n suit également une loi normale, qui est donc nécessairement la loi $\mathcal{N}(0, 1)$.

Le théorème central limite n'est donc d'aucun intérêt dans ce cas : les S_n^* suivent déjà une loi normale centrée réduite, inutile de parler de convergence en loi !

¹² Pour le dire vulgairement : «proche d'une courbe en cloche».

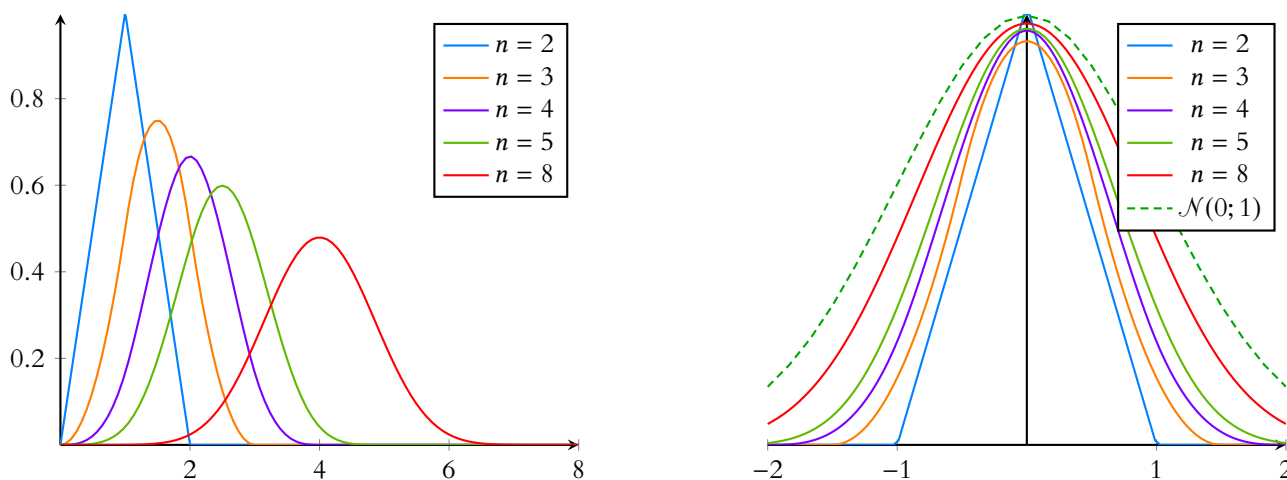


FIGURE 14.4 – Densités de la somme de n variables aléatoires indépendantes suivant la loi uniforme sur $[0, 1]$ (à gauche), et des variables centrées réduites associées (à droite).

14.3.2 Approximations classiques

Soit (X_n) une suite de variables aléatoires indépendantes de même loi de Bernoulli $\mathcal{B}(p)$.

Alors $n\bar{X}_n = \sum_{i=1}^n X_i \hookrightarrow \mathcal{B}(n, p)$, et $\bar{X}_n^* = \frac{\bar{X}_n - np}{\sqrt{np(1-p)}}$.

Nous savons que, d'après le théorème central limite, (\bar{X}_n^*) converge en loi vers $\mathcal{N}(0, 1)$.

Ainsi, pour n suffisamment grand, on peut considérer que \bar{X}_n^* suit à peu près une loi $\mathcal{N}(0, 1)$.

Et alors, par transformation affine de loi normale, $\bar{X}_n = \sqrt{np(1-p)}\bar{X}_n^* + np$ suit approximativement une loi $\mathcal{N}(np, np(1-p))$.

Proposition 14.26 : Lorsque $n \geq 30$, $np \geq 15$ et $n(1-p) \geq 15$, on approche la loi $\mathcal{B}(n, p)$ par la loi $\mathcal{N}(np, np(1-p))$.

Hypothèses

Le programme officiel stipule que les conditions sous lesquelles l'approximation est valable doivent systématiquement être précisées.

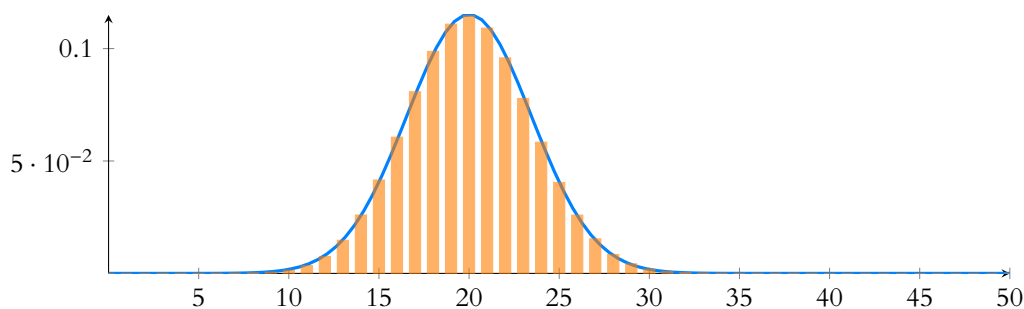


FIGURE 14.5 – Approximation de la loi $\mathcal{B}(50, 0.4)$ par la loi $\mathcal{N}(20, 12)$.

Remarques. • On approche $\mathcal{B}(n, p)$ par la loi normale de même espérance et de même variance.

- Il n'est pas question de «prouver» que les bonnes hypothèses sont $n \geq 30$, $np \geq 15$ et $n(1-p) \geq 15$. Ces valeurs ont été obtenues numériquement : on a constaté qu'elles fournissaient une bonne approximation. D'ailleurs suivant le niveau de précision souhaité pour l'approximation, il peut-être préférable d'utiliser des hypothèses plus contraignantes.
- Lorsque np est trop petit, on approchera plutôt la loi $\mathcal{B}(n, p)$ par la loi $\mathcal{P}(np)$ (en pratique, on demande souvent $np \leq 15$ et $p \leq 0.1$).

Exemple 14.27

Revenons sur l'exemple du nombre de lancers d'un dé nécessaires pour que la fréquence d'apparition du 1 soit égale à $\frac{1}{6} \pm 0.01$.

Si n est suffisamment grand, on peut approcher la loi de X par la loi normale $\mathcal{N}\left(\frac{n}{6}, \frac{5n}{36}\right)$, et donc approcher la loi de $Y = \frac{X}{n}$ par la loi normale $\mathcal{N}\left(\frac{1}{6}, \frac{5}{36n}\right)$.

On souhaitait alors

$$P\left(\left|Y - \frac{1}{6}\right| \geq 0.01\right) \leq 0.05$$

soit $P\left(\frac{1}{6} - 0.01 \leq Y \leq \frac{1}{6} + 0.01\right) \geq 0.95$. Mais

$$P\left(\frac{1}{6} - 0.01 \leq Y \leq \frac{1}{6} + 0.01\right) = P\left(-\frac{0.01\sqrt{36n}}{\sqrt{5}} \leq \frac{Y - \frac{1}{6}}{\sqrt{\frac{5}{36n}}} \leq \frac{0.01\sqrt{36n}}{\sqrt{5}}\right)$$

et $\frac{Y - \frac{1}{6}}{\sqrt{\frac{5}{36n}}} \leftrightarrow \mathcal{N}(0, 1)$, de sorte que l'on souhaite

$$2\Phi\left(\frac{0.01\sqrt{36n}}{\sqrt{5}}\right) - 1 \geq 0.95.$$

Une table de la loi normale centrée réduite nous indique qu'il faut alors

$$\frac{0.01\sqrt{36n}}{\sqrt{5}} \geq 1.96 \Leftrightarrow n \geq \frac{5 \times 1.96^2}{36 \times 0.01^2} \Leftrightarrow n \geq 5336.$$

Notons que ce résultat est à la fois bien mieux que celui obtenu à l'aide de Bienaymé-Tchebychev (ce qui vient du fait qu'on a utilisé au moins en partie la connaissance de la loi de X ici), et ... faux car comme mentionnée précédemment, un calcul détaillé à l'aide de la loi binomiale prouve qu'il faut $n \geq 5395$. L'erreur est relativement petite, et vient du fait qu'on a réalisé une approximation de la loi de X en utilisant une loi normale plutôt qu'une loi binomiale.

De même, soit (X_n) une suite de variables aléatoires indépendantes, de loi de Poisson $\mathcal{P}(\lambda)$.

Alors $Y_n = X_1 + \dots + X_n \leftrightarrow \mathcal{P}(n\lambda)$.

Et alors $\bar{Y}_n^* = \frac{X_1 + \dots + X_n - n\lambda}{\sqrt{\lambda}\sqrt{n}}$ converge en loi vers une variable suivant une loi normale centrée réduite.

On peut donc considérer que \bar{Y}_n^* suit approximativement une loi $\mathcal{N}(0, 1)$, et alors

$$X_1 + \dots + X_n = \sqrt{n}\sqrt{\lambda}\bar{Y}_n^* + n\lambda$$

suit approximativement une loi $\mathcal{N}(n\lambda, n\lambda)$.

Ainsi, pour n suffisamment grand, la loi $\mathcal{P}(n\lambda)$ est approximativement la loi $\mathcal{N}(n\lambda, n\lambda)$.

Autrement dit, pour λ assez grand, on approxime la loi $\mathcal{P}(\lambda)$ par la loi $\mathcal{N}(\lambda, \lambda)$.

Proposition 14.28 : Pour $\lambda \geq 15$, on approche la loi $\mathcal{P}(\lambda)$ par la loi $\mathcal{N}(\lambda, \lambda)$.

14.3.3 Correction de continuité

Dans les deux exemples précédents, on approche une variable discrète X par une variable à densité Y .

Comment utiliser au mieux ceci pour calculer les $P(X = k)$? Il n'est pas question de dire que $P(X = k) \simeq P(Y = k)$, car Y étant à densité, on a $P(Y = k) = 0$.

La règle communément acceptée est d'utiliser le fait que

$$P(X = k) = P(k - 0.5 \leq X \leq k + 0.5)$$

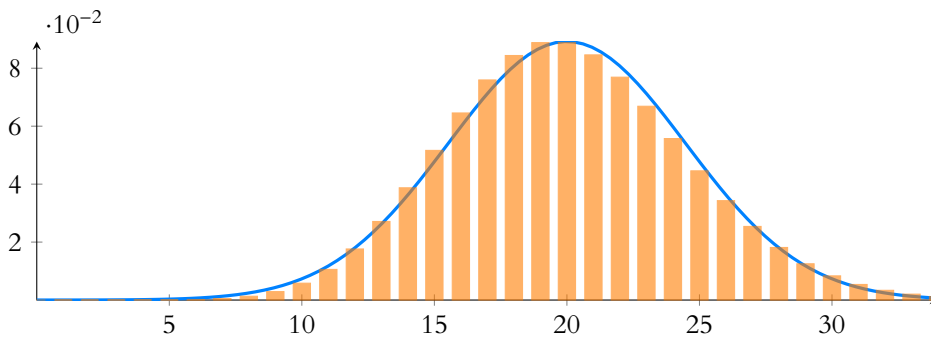


FIGURE 14.6 – Approximation de la loi $\mathcal{P}(20)$ par la loi $\mathcal{N}(20, 20)$.

et donc d'approcher $P(X = k)$ par $P(k - 0.5 \leq Y \leq k + 0.5)$.

De plus, afin de garantir que la somme des probabilités reste égale à 1, on remplacera $P(X = 0)$ par $P(Y \leq 0.5)$.

Et dans le cas de l'approximation d'une loi binomiale $\mathcal{B}(n, p)$ par une loi normale, on remplacera $P(X = n)$ par $P(n - 0.5 \leq Y)$.

Pas de preuve !

Ces résultats sont des résultats empiriques, c'est-à-dire qu'on a constaté que généralement, ils fonctionnaient bien. Il n'est pas question de les prouver, juste d'accepter que procéder ainsi fournit de « bonnes » approximations. Tout dépend alors du sens que l'on accorde à « bonnes » : généralement elles suffisent, d'autre fois elles ne vont pas être assez précises pour ce que l'on souhaite faire.

Exemple 14.29

Soit $X \hookrightarrow \mathcal{B}(100, 0.4)$. Les conditions sont réalisées pour que l'on puisse approcher X par une variable aléatoire Y de loi $\mathcal{N}(40, 24)$. Alors $P(X = 35) \simeq P(34.5 \leq Y \leq 35.5)$. Mais

$$\begin{aligned} P(34.5 \leq Y \leq 35.5) &= P\left(\frac{34.5 - 40}{\sqrt{24}} \leq \frac{Y - 40}{\sqrt{24}} \leq \frac{35.5 - 40}{\sqrt{24}}\right) \\ &= P(-1.12 \leq \bar{Y}^* \leq -0.92) = \Phi(-0.92) - \Phi(-1.12) = \Phi(1.12) - \Phi(0.92) \\ &= 0.8686 - 0.8212 = 0.0474. \end{aligned}$$

Un calcul direct donne $P(X = 35) = 0.0491$.

14.3.4 Rappels sur l'utilisation d'une table de la loi normale centrée réduite

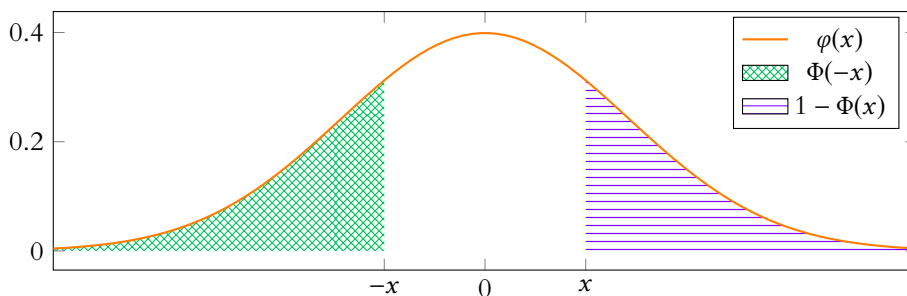
Le tableau qui suit comporte les valeurs de Φ , la fonction de répartition de la loi normale centrée réduite.

Utilisation : on lit les décimales dans les lignes, et les centièmes en colonnes.

Par exemple, la valeur de $\Phi(1.65)$ se trouve à l'intersection de la ligne 1.6 et de la colonne 0.05. On trouve $\Phi(1.65) = 0.9505$, à 10^{-4} près.

Rappelons rapidement quelques propriétés de Φ :

- Φ est continue et strictement croissante
- $\Phi(0) = 0.5$ et $\Phi(-x) = 1 - \Phi(x)$ (cette relation se retrouve graphiquement à l'aide de la figure ci-dessous).



Conséquence

Le théorème des valeurs intermédiaires prouve que pour tout réel $\alpha \in]0, 1[$, il existe un unique $t_\alpha \in \mathbf{R}$ tel que $\Phi(t_\alpha) = \alpha$.

Pour toute autre loi normale, on se rappellera que $X \hookrightarrow \mathcal{N}(\mu, \sigma^2) \Leftrightarrow \frac{X - \mu}{\sigma} \hookrightarrow \mathcal{N}(0, 1)$.

	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7703	0.7734	0.7764	0.7793	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8906	0.8925	0.8943	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.9980	0.9981
2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986
3.0	0.9986	0.9987	0.9987	0.9988	0.9988	0.9989	0.9989	0.9989	0.9990	0.9990
3.1	0.9990	0.9991	0.9991	0.9991	0.9992	0.9992	0.9992	0.9992	0.9993	0.9993
3.2	0.9993	0.9993	0.9994	0.9994	0.9994	0.9994	0.9994	0.9995	0.9995	0.9995
3.3	0.9995	0.9995	0.9995	0.9996	0.9996	0.9996	0.9996	0.9996	0.9996	0.9997
3.4	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9998
3.5	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998
3.6	0.9998	0.9998	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999
3.7	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999
3.8	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999
3.9	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

MÉTHODES DU CHAPITRE 14

✓ MÉTHODE 14.1 : Prouver une convergence en probabilité

- Commencer par s'assurer si ce n'est pas directement une conséquence de la loi faible des grands nombres.
 - Sinon, on pourra essayer d'utiliser l'inégalité de Bienaymé-Tchebychev.
 - Si aucune des deux méthodes ne fonctionne, il faudra calculer directement les $P(|X_n - X| \geq \varepsilon)$, puis passer à la limite.
- Exercices de TD : 14.4, 14.5, 14.8, 14.9, 14.10, 14.12

✓ MÉTHODE 14.2 : Étudier la convergence en loi d'une suite (X_n) de variables aléatoires

1. Commencer par calculer la fonction de répartition F_n de chacune des X_n .
2. Pour $x \in \mathbf{R}$ fixé, calculer la limite de $F_n(x)$ lorsque $n \rightarrow +\infty$.
3. Si cette limite existe et est notée $F(x)$, vérifier que F est la fonction de répartition d'une variable aléatoire, ou si du moins, elle coïncide avec la fonction de répartition d'une variable aléatoire en tous ses points de continuité.

Notons que le théorème central limite fournit une convergence en loi, et que dans le cas où l'on doit prouver la convergence en loi vers une loi normale, on pensera d'abord au théorème central limite !

Dans le cas où les X_n sont à valeurs dans \mathbf{Z} et où la limite en loi (fournie par l'énoncé) est également à valeurs dans \mathbf{Z} , on se contentera de calculer les $\lim_{n \rightarrow +\infty} P(X_n = k), k \in \mathbf{Z}$.

Exercices de TD : 14.6, 14.7, 14.11, 14.17

✓ MÉTHODE 14.3 : Utiliser le théorème central limite

Si les X_n sont i.i.d. et admettent une variance, on retiendra que la suite des variables centrées réduites associées à $X_1 + \dots + X_n$ converge en loi vers une variable suivant la loi $\mathcal{N}(0, 1)$.

La variable centrée réduite associée à $X_1 + \dots + X_n$ est la même que la variable centrée réduite associée à $\overline{X}_n = \frac{X_1 + \dots + X_n}{n}$.

Suivant le contexte, on choisira donc de l'exprimer en fonction de $X_1 + \dots + X_n$ ou de \overline{X}_n .

La convergence en loi ne peut en aucun cas permettre de calculer des probabilités, mais uniquement des limites de probabilités : on n'a pas dit que \overline{X}_n^* suivait une loi normale centrée réduite, mais converge en loi vers une telle loi !

Exercices de TD : 14.13, 14.14

✓ MÉTHODE 14.4 : Approximer une loi par une autre

On ne procédera à une telle approximation que si ceci possède un réel intérêt, c'est-à-dire si cela nous permet de faire un calcul qu'on ne pourrait faire sinon.

- Pour une loi binomiale, on a le choix entre l'approximation par une loi de Poisson et l'approximation par une loi normale. La première est valable pour p «petit», alors que la seconde convient pour des p ni trop proches de 0, ni trop proches de 1. Dans les deux cas, on se souviendra qu'il s'agit d'approcher par une loi de même espérance que la binomiale d'origine, et dans le cas de l'approximation par une normale, la variance doit être la même que celle de la binomiale.
- Pour l'approximation d'une loi de Poisson par une normale, on se souviendra que cette dernière doit avoir même espérance et même variance que la loi de départ.

Exercices de TD : 14.15, 14.16

EXERCICES DU CHAPITRE 14

EXERCICE 14.1 Soit $(X_n)_{n \geq 1}$ une suite de variables aléatoires telles que $X_n \hookrightarrow \mathcal{N}\left(m, \frac{\sigma^2}{n}\right)$.

PD

- 1) Déterminer la fonction de répartition de X_n .
- 2) Étudier la convergence en loi de la suite (X_n) .

EXERCICE 14.2 On note Φ la fonction de répartition d'une variable aléatoire suivant la loi $\mathcal{N}(0, 1)$.

PD

- 1) Montrer à l'aide de l'inégalité de Bienaymé-Tchebychev que $\forall x \in \mathbf{R}_+^*$, $1 - \Phi(x) \leq \frac{1}{2x^2}$.
- 2) En déduire que $\int_0^{+\infty} (1 - \Phi(t)) dt$ converge et calculer sa valeur à l'aide d'une intégration par parties.

EXERCICE 14.3

PD

- 1) Soit (Y_n) une suite de variables aléatoires mutuellement indépendantes, de même loi et admettant un moment d'ordre 2. On pose $S_n = \sum_{k=1}^n Y_k$.

Prouver que : $\forall a \in]0, +\infty[$, $P\left(\left|\frac{S_n}{n} - E(Y_1)\right| \geq a\right) \leq \frac{V(Y_1)}{na^2}$.

- 2) **Application** : On effectue des tirages successifs, avec remise, d'une boule dans une urne contenant 2 boules rouges et 3 boules noires. À partir de quel nombre de tirages peut-on garantir à plus de 95% que la proportion de boules rouges obtenues restera comprise entre 0.35 et 0.45 ?

Indication : considérer la suite (Y_i) de variables aléatoires de Bernoulli où Y_i mesure l'issue du i -ième tirage.

EXERCICE 14.4 Méthode de Monte-Carlo

PD

Soit $(X_n)_{n \in \mathbf{N}^*}$ une suite de variables aléatoires i.i.d. suivant la loi exponentielle de paramètre 1, et soit $x > \frac{1}{2}$.

On pose alors $U_n = \frac{1}{n} \sum_{k=1}^n X_k^{x-1}$.

- 1) Montrer que $U_n \xrightarrow{P} \Gamma(x)$.
- 2) En déduire une fonction Sci Lab qui prend en paramètre un réel x strictement positif et qui retourne une valeur approchée de $\Gamma(x)$.
- 3) (★) Comment adapter ce programme pour obtenir une valeur approchée de $\Gamma(x)$, pour $0 < x \leq \frac{1}{2}$?

EXERCICE 14.5 Soit (X_n) une suite de variables aléatoires à valeurs positives, ayant toutes une espérance et une variance. On suppose que $\lim_{n \rightarrow +\infty} E(X_n) = m$ et $\lim_{n \rightarrow +\infty} V(X_n) = 0$.

PD

En appliquant l'inégalité de Markov à $(X_n - m)^2$, montrer que $X_n \xrightarrow{P} m$.

EXERCICE 14.6 Soit (X_n) une suite variables aléatoires indépendantes suivant toutes la loi géométrique de paramètre $p \in]0, 1[$. On pose alors $Y_n = \min(X_1, \dots, X_n)$.

PD

Étudier la convergence en loi de $(Y_n)_{n \geq 1}$.

EXERCICE 14.7 Soit $f : x \mapsto \frac{e^x}{(1 + e^x)^2}$.

AD

- 1) Montrer que f est une densité de probabilité.
- 2) Soit à présent $(X_n)_{n \in \mathbf{N}}$ une suite de variables i.i.d. de densité f .
 - a) Déterminer la fonction de répartition des X_i .
 - b) En déduire la fonction de répartition de $M_n = \max(X_1, \dots, X_n)$.
 - c) Étudier la convergence en loi de la suite $(M_n - \ln(n))_{n \in \mathbf{N}}$.

EXERCICE 14.8 Soit (X_n) une suite de variables aléatoires indépendantes sur (Ω, \mathcal{A}, P) , suivant toutes la loi de Bernoulli de paramètre p . On note $Y_n = X_n + X_{n+1}$, et on pose $S_n = \frac{Y_1 + \dots + Y_n}{n}$.

PD

Montrer que la suite (S_n) converge en probabilités vers la variable certaine égale à $2p$.

EXERCICE 14.9 Minimum de lois de Rayleigh

PD

Soit $\sigma > 0$. On définit une fonction f sur \mathbf{R} par $f(x) = \begin{cases} \frac{x}{\sigma^2} e^{-\frac{x^2}{2\sigma^2}} & \text{si } x \geq 0 \\ 0 & \text{sinon} \end{cases}$

- 1) Montrer que f est une densité de probabilités.
- 2) Soit X une variable aléatoire de densité f . Déterminer la fonction de répartition de X .
- 3) Soient $(X_i)_{i \geq 1}$ des variables aléatoires indépendantes, de densité f . On pose alors

$$\forall n \in \mathbf{N}^*, Y_n = \min(X_1, \dots, X_n).$$

Montrer que $\forall \varepsilon > 0, \lim_{n \rightarrow +\infty} P(|Y_n| \leq \varepsilon) = 1$. Qu'en déduire sur la suite (Y_n) ?

EXERCICE 14.10 Soit $(X_n)_{n \in \mathbf{N}^*}$ une suite de variables aléatoires indépendantes sur (Ω, \mathcal{A}, P) suivant toutes une loi uniforme sur $[a, b]$.

On note $Y_n = \min(X_1, \dots, X_n)$. Montrer que (Y_n) converge en probabilités vers la variable aléatoire certaine égale à a .

EXERCICE 14.11 Minimum de lois discrètes

AD

Soit $N \geq 2$, et soit $(X_n)_{n \in \mathbf{N}^*}$ une suite de variables aléatoires indépendantes suivant la loi uniforme sur $\llbracket 1, N \rrbracket$. On pose alors $Y_n = \min(X_1, \dots, X_n)$.

Montrer que la suite (Y_n) converge en loi vers une variable que l'on précisera.

Question subsidiaire : montrer qu'il s'agit également d'une convergence en probabilités.

EXERCICE 14.12 Soit X une variable aléatoire et (Y_n) une suite de variables aléatoires telles que $Y_n \xrightarrow{P} \mathcal{B}\left(\frac{1}{n}\right)$.

AD

On pose alors $X_n = X + Y_n$. Montrer que $X_n \xrightarrow{P} X$.

Même question si $Y_n \xrightarrow{P} \mathcal{N}\left(\frac{1}{n}, \frac{1}{n^2}\right)$.

EXERCICE 14.13 (D'après EDHEC 2007)

AD

On considère une suite $(X_n)_{n \geq 1}$ de variables aléatoires définies sur le même espace probabilisé (Ω, \mathcal{A}, P) , mutuellement indépendantes, et qui suivent toutes la loi exponentielle de paramètre 1.

On pose $S_n = \sum_{i=1}^n X_i$.

- 1) Rappeler quelle est la loi suivie par S_n . Donner l'espérance et la variance de S_n .
- 2) À l'aide du théorème de la limite centrée, établir que $\lim_{n \rightarrow \infty} P(S_n \leq n) = \frac{1}{2}$.
- 3) En déduire un équivalent lorsque n tend vers $+\infty$ de $\int_0^n t^{n-1} e^{-t} dt$.

EXERCICE 14.14 Soit (S_n) une suite de variables aléatoires telles que $S_n \xrightarrow{P} \mathcal{B}(n, p)$. Déterminer $\lim_{n \rightarrow +\infty} P(S_n \leq np)$.

AD

Indication : on pourra introduire pour cela une suite de variables aléatoires indépendantes suivant une même loi de Bernoulli $\mathcal{B}(p)$.

EXERCICE 14.15 Dans le désert, une Renault 4L crève en moyenne tous les 4000 km. On considère donc qu'à chaque kilomètre, la probabilité de crever est de $\frac{1}{4000}$.

PD

Un équipage s'inscrit au 4L Trophy, rallye de 6000 km dans le désert. Il aimerait savoir combien de roues de secours emporter pour avoir moins de 10 % de chances de manquer de roues de secours.

Deux roues de secours sont-elles suffisantes ? Trois ? On donne $e^{-3/2} \simeq 0.22$.

EXERCICE 14.16 Une entreprise compte 300 employés. Chacun d'eux téléphone en moyenne 6 minutes par heure. Combien de lignes téléphonique doit faire installer l'entreprise afin que la probabilité que toutes les lignes soient occupées en même temps soit inférieure ou égale à 0.025 ?

PD

EXERCICE 14.17 Soit $(U_n)_{n \in \mathbf{N}^*}$ une suite de variables aléatoires indépendantes suivant toutes la loi uniforme sur $[0, 1]$. On note M_n et X_n les variables aléatoires définies par

AD

$$M_n = \max(U_1, \dots, U_n) \text{ et } X_n = n(1 - M_n).$$

Étudier la convergence en loi de la suite de variables aléatoires $(X_n)_{n \in \mathbf{N}^*}$.

EXERCICE 14.18 Une utilisation de Slutsky

D

Soit $(u_n)_n$ une suite de réels convergente, de limite ℓ . Soit (X_n) une suite de variables aléatoires convergeant en loi vers une variable aléatoire à densité X .

- 1) Montrer que $u_n \xrightarrow{P} \ell$.
- 2) Étudier la convergence en loi de la suite $(X_n - u_n)_n$.
- 3) Montrer que $\lim_{n \rightarrow +\infty} P(X_n \leq u_n) = P(X \leq \ell)$.

EXERCICE 14.19 (QSP ESCP 2016)

On considère une variable aléatoire X définie sur un espace probabilisé (Ω, \mathcal{A}, P) qui suit une loi de Poisson de paramètre $\lambda > 0$.

- 1) Soit t un réel strictement positif. Montrer que la variable aléatoire e^{tX} admet une espérance et la calculer.
- 2) Soit $n \in \mathbf{N}$. Prouver que $P(X \geq n) \leq e^{-tn} E(e^{tX})$.
- 3) En déduire que : $\sum_{k=n}^{+\infty} \frac{\lambda^k}{k!} \leq e^n \left(\frac{\lambda}{n}\right)^n$.

EXERCICE 14.20 (QSP HEC 2007)

Déterminer une suite $(X_n)_{n \in \mathbf{N}^*}$ de variables aléatoires telles que :

- chacune des variables X_n ne prenne que deux valeurs
- $(X_n)_{n \in \mathbf{N}^*}$ converge en probabilités vers la variable nulle
- $\lim_{n \rightarrow +\infty} E(X_n) = 1$
- $\lim_{n \rightarrow +\infty} V(X_n) = +\infty$.

EXERCICE 14.21 (Oral ESCP 2017)

Dans cet exercice, toutes les variables aléatoires sont définies sur le même espace probabilisé (Ω, \mathcal{A}, P) . Soit $(X_n)_n, (Y_n)_n$ deux suites de variables aléatoires et X, Y deux variables aléatoires.

On suppose que $(X_n)_n$ converge en probabilité vers X et que $(Y_n)_n$ converge en probabilité vers Y .

- 1) Vérifier que $X_n Y_n - XY = (X_n - X)Y + (Y_n - Y)X + (X_n - X)(Y_n - Y)$.

Soit $\varepsilon > 0$. On pose :

$$A_n = [|X_n Y_n - XY| > \varepsilon], B_n = [|X_n - X||Y| > \varepsilon/3], C_n = [|Y_n - Y||X| > \varepsilon/3], D_n = [|X_n - X||Y_n - Y| > \varepsilon/3].$$

- 2) Montrer que $P(A_n) \leq P(B_n) + P(C_n) + P(D_n)$.
- 3) Soit $\delta > 0$.

- a) Montrer qu'il existe un réel $A > 0$ tel que pour tout $t > A$, $P(|Y| > t) < \frac{\delta}{2}$.

- b) Montrer que pour tout $t > 0$, $P(B_n) \leq P\left(|X_n - X| > \frac{\varepsilon}{3t}\right) + P(|Y| > t)$.

- c) Soit $t > A$. Montrer qu'il existe un entier N_0 tel que pour tout $n \geq N_0$, $P\left(|X_n - X| > \frac{\varepsilon}{3t}\right) < \frac{\delta}{2}$.

- d) En déduire que $\lim_{n \rightarrow +\infty} P(B_n) = 0$.

On montrerait de même et on admet ici que $\lim_{n \rightarrow +\infty} P(C_n) = 0$.

- 4) Montrer que $P(D_n) \leq P\left(|X_n - X| > \sqrt{\frac{\varepsilon}{3}}\right) + P\left(|Y_n - Y| > \sqrt{\frac{\varepsilon}{3}}\right)$.

En déduire que $\lim_{n \rightarrow +\infty} P(D_n) = 0$.

- 5) Montrer que la suite $(X_n Y_n)$ converge en probabilité vers XY .

EXERCICE 14.22 La convergence en probabilités implique la convergence en loi (Oral HEC)

Soient $(X_n)_{n \geq 1}$ et X des variables aléatoires définies sur un espace probabilisé (Ω, \mathcal{A}, P) telles que $X_n \xrightarrow{P} X$. On note F_n la fonction de répartition de X_n et F la fonction de répartition de X .

Soit x un point de continuité de F , et soit $\delta > 0$ fixé.

- 1) Montrer qu'on peut choisir ε tel que $F(x - \varepsilon) > F(x) - \delta$ et $F(x + \varepsilon) < F(x) + \delta$.
- 2) Montrer que

$$[X_n \leq x] \subset [X \leq x + \varepsilon] \cup [|X_n - X| > \varepsilon].$$

Montrer de même que

$$[X \leq x - \varepsilon] \subset [X_n \leq x] \cup [|X_n - X| > \varepsilon].$$

- 3) En déduire que

$$F_n(x) \leq F(x) + \delta + P(|X_n - X| > \varepsilon) \text{ et } F(x) - \delta \leq F_n(x) + P(|X_n - X| > \varepsilon).$$

- 4) Montrer que $X_n \xrightarrow{\mathcal{L}} X$.

CORRECTION DES EXERCICES DU CHAPITRE 14

SOLUTION DE L'EXERCICE 14.1

1. Comme d'habitude, la fonction de répartition d'une loi normale s'exprime à l'aide de Φ en faisant apparaître une loi normale centrée réduite. Plus précisément, pour $x \in \mathbf{R}$, on a

$$P(X_n \leq x) = P(X_n - m \leq x - m) = P\left(\frac{X_n - m}{\frac{\sigma}{\sqrt{n}}} \leq \frac{x - m}{\frac{\sigma}{\sqrt{n}}}\right) = \Phi\left(\sqrt{n}\frac{x - m}{\sigma}\right).$$

Donc pour $x < m$, on a $\frac{x - m}{\sigma} < 0$, et donc $\sqrt{n}\frac{x - m}{\sigma} \xrightarrow{n \rightarrow +\infty} -\infty$, de sorte que $P(X_n \leq x) \xrightarrow{n \rightarrow +\infty} 0$.

Au contraire, pour $x > m$, alors $\frac{x - m}{\sigma} > 0$, et donc $\sqrt{n}\frac{x - m}{\sigma} \xrightarrow{n \rightarrow +\infty} +\infty$. On en déduit que $P(X_n \leq x) \xrightarrow{n \rightarrow +\infty} 1$.

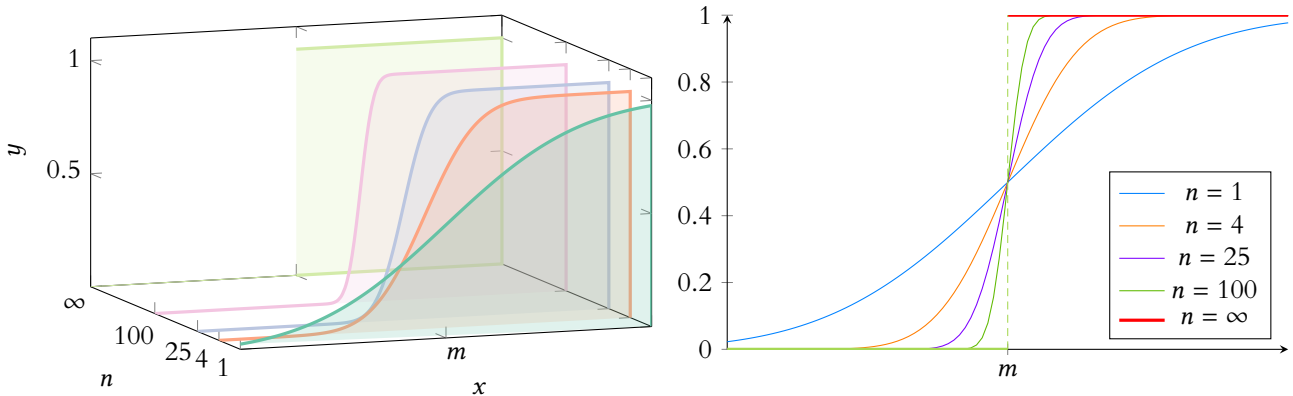
Ainsi, pour $x \neq m$, $F_n(x)$ converge vers $F_X(x)$, où X est la variable constante égale à m . Puisque F_X n'est pas continue en $x = m$, il n'est pas utile de s'intéresser à la limite des $F_n(0)$, nous en savons déjà assez pour en déduire que $X_n \xrightarrow{\mathcal{L}} X$.

Lois normales

Puisque $X_n \hookrightarrow \mathcal{N}\left(m, \frac{\sigma^2}{n}\right)$ alors $\sqrt{n}\frac{X - m}{\sigma} \hookrightarrow \mathcal{N}(0, 1)$.

Limites

On a utilisé ici le fait que Φ , comme n'importe quelle fonction de répartition, est de limite nulle en $-\infty$ et égale à 1 en $+\infty$.



SOLUTION DE L'EXERCICE 14.2

1. Soit X une variable aléatoire suivant la loi $\mathcal{N}(0, 1)$. Alors on a $1 - \Phi(x) = 1 - P(X \leq x) = P(X > x)$. Mais, par symétrie de Φ , on a $P(X > x) = P(X \geq x) = P(X \leq -x)$, et alors

$$1 - \Phi(x) = \frac{1}{2}P(|X| \geq x).$$

Par l'inégalité de Bienaymé-Tchebychev, il vient

$$P(|X| \geq x) \leq \frac{V(X)}{x^2} = \frac{1}{x^2}.$$

On en déduit que $1 - \Phi(x) \leq \frac{1}{2x^2}$.

2. La fonction $t \mapsto 1 - \Phi$ est continue et positive sur \mathbf{R}_+ . Comme $\int_1^{+\infty} \frac{dt}{2t^2}$ est une intégrale de Riemann convergente, il en est de même de $\int_1^{+\infty} (1 - \Phi(t)) dt$. Et donc $\int_0^{+\infty} (1 - \Phi(t)) dt$ converge.

Soit $A > 0$. Procédons à une intégration par parties sur le segment $[0, A]$ en posant $u(t) = t$ et $v(t) = 1 - \Phi(t)$, qui sont deux fonctions \mathcal{C}^1 , avec $v'(t) = -\varphi(t) = -\frac{1}{\sqrt{2\pi}}e^{-t^2/2}$. Alors

$$\int_0^A (1 - \Phi(t)) dt = [t(1 - \Phi(t))]_0^A + \frac{1}{\sqrt{2\pi}} \int_0^A t e^{-t^2/2} dt = A(1 - \Phi(A)) + \frac{1}{\sqrt{2\pi}} [-e^{-t^2/2}]_0^A.$$

Détails

On peut décomposer l'événement $|X| \geq x$ en $[X \geq x] \cup [X \leq -x]$ et ces deux événements sont incompatibles.

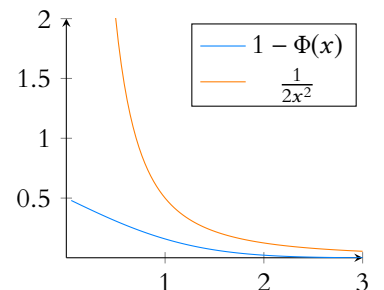


FIGURE 14.1- L'inégalité est loin d'être optimale.

Rappel

Si on ne connaît pas d'expression « simple » de Φ , on connaît sa dérivée qui est la densité usuelle de la loi normale.

Lorsque $A \rightarrow +\infty$, on a

$$0 \leq A(1 - \Phi(A)) \leq \frac{1}{2A} \xrightarrow{A \rightarrow +\infty} 0$$

$$\text{et } \left[-e^{-t^2/2} \right]_0^A = 1 - e^{-A^2/2} \xrightarrow{A \rightarrow +\infty} 1.$$

On en déduit que

$$\int_0^{+\infty} (1 - \Phi(t)) dt = \frac{1}{\sqrt{2\pi}}.$$

SOLUTION DE L'EXERCICE 14.3

1. S_n possède une variance car les Y_i possèdent une variance, et par indépendance des Y_i , il vient

$$V(S_n) = \sum_{i=1}^n V(Y_i) = nV(Y_1) \text{ et donc } V\left(\frac{S_n}{n}\right) = \frac{1}{n^2}V(S_n) = \frac{V(Y_1)}{n}.$$

$$\text{De plus, } E\left(\frac{S_n}{n}\right) = \frac{E(S_n)}{n} = E(Y_1).$$

Alors l'inégalité de Bienaymé-Tchebychev appliquée à la variable aléatoire $\frac{S_n}{n}$ nous donne

$$\forall a > 0, P\left(\left|\frac{S_n}{n} - E\left(\frac{S_n}{n}\right)\right| \geq a\right) = P\left(\left|\frac{S_n}{n} - E(Y_1)\right| \geq a\right) \leq \frac{V\left(\frac{S_n}{n}\right)}{a^2} = \frac{V(Y_1)}{na^2}.$$

2. Notons comme indiqué dans l'énoncé, Y_i la variable aléatoire qui vaut 1 si la boule obtenue au i -ème tirage est rouge et 0 sinon.

Alors les Y_i sont indépendantes¹, et suivent la même loi de Bernoulli de paramètre $\frac{2}{5}$.

$$\text{Donc } E(Y_1) = \frac{2}{5} \text{ et } V(Y_1) = \frac{2}{5} \frac{3}{5} = \frac{6}{25}.$$

En notant $S_n = \sum_{i=1}^n Y_i$, $\frac{S_n}{n}$ représente la proportion de boules rouges obtenues au cours des n premiers tirages.

Et donc par la question précédente, il vient

$$\forall a > 0, P\left(\left|\frac{S_n}{n} - 0.4\right| \geq a\right) \leq \frac{6}{25na^2}.$$

En particulier, pour $a = 0.05$, on a

$$P\left(\left|\frac{S_n}{n} - 0.4\right| \geq 0.05\right) = P\left(\frac{S_n}{n} \notin]0.35, 0.45[\right) \leq \frac{6}{25 \times 0.05^2 \times n} = \frac{96}{n}.$$

Et donc, en passant à l'événement contraire,

$$P\left(\frac{S_n}{n} \in [0.35, 0.45]\right) \geq 1 - \frac{96}{n}.$$

Pour s'assurer que cette probabilité est bien supérieure à 0.95, on peut faire en sorte que

$$\frac{96}{n} \leq 0.05 \Leftrightarrow n \geq \frac{96}{0.05} = 1920.$$

Ainsi, nous sommes certains que pour plus de 1920 tirages, il y a plus de 95% de chances d'obtenir une proportion de rouges comprise entre 0.35 et 0.45.

SOLUTION DE L'EXERCICE 14.4

1. Les X_k^{x-1} sont indépendantes et identiquement distribuées car les X_k le sont. De plus, par le théorème de transfert,

$$E(X_k^{x-1}) = \int_0^{+\infty} t^{x-1} e^{-t} dt = \Gamma(x).$$

Et de même,

$$E\left(\left(X_k^{x-1}\right)^2\right) = \int_0^{+\infty} t^{2(x-1)} e^{-t} dt = \Gamma(2x-1),$$

¹ Car les tirages ont lieu avec remise.

En revanche

Rien ne nous permet de garantir que 1920 est optimal, l'inégalité de Bienaymé-Tchebychev étant probablement trop imprécise.

On a bien $2x - 1 > 0$, donc l'intégrale converge.

de sorte que les X_k^{x-1} ont une variance.
Par la loi faible des grands nombres, on a donc

$$U_n = \frac{1}{n} \sum_{k=1}^n X_k^{x-1} \xrightarrow{P} E(X_1^{x-1}) = \Gamma(x).$$

2. Il s'agit de simuler la variable U_n pour une «grande» valeur de n . Prenons par exemple $n = 1000$.

```

1  function y = gamma(x)
2      X = grand(1,1000,'exp',1)
3      y = 0
4      for i=1 :1000
5          y = y+X(i)^(x-1)
6      end
7      y = y/1000
8  endfunction

```

Notons que ce programme est probabiliste, et donc pourrait éventuellement renvoyer une valeur très éloignée de la valeur réelle de $\Gamma(x)$. Plus n est grand moins cela est probable, mais nous ne pouvons jamais complètement l'exclure.

3. Pour $x \in \left]0, \frac{1}{2}\right[$, ce qui a été écrit précédemment ne fonctionne pas, puisque les X_k^{x-1} n'ont pas de variance, et donc la loi faible des grands nombres ne s'applique plus.

En revanche, nous savons dans ce cas que $\Gamma(x+1) = x\Gamma(x)$, et puisque $x+1 > \frac{1}{2}$, ce qui précède s'applique.

Donc nous pouvons utiliser le programme précédent pour approcher $\Gamma(x+1)$, puis diviser le résultat par x .

SOLUTION DE L'EXERCICE 14.5

Notons que $(X_n - m)^2$ est positive, et admet une espérance car

$$(X_n - m)^2 = X_n^2 - 2mX_n + m^2$$

et que X_n admet un moment d'ordre 2.

En particulier,

$$E((X_n - m)^2) = E(X_n^2) - 2mE(X_n) + m^2 = V(X_n) + E(X_n)^2 - 2mE(X_n) + m^2.$$

Et donc, pour tout $\varepsilon > 0$, on a, d'après l'inégalité de Markov,

$$P(|X_n - m| \geq \varepsilon) = P((X_n - m)^2 \geq \varepsilon^2) \leq \frac{E((X_n - m)^2)}{\varepsilon^2} = \frac{V(X_n) + E(X_n)^2 - 2mE(X_n) + m^2}{\varepsilon^2}.$$

Mais lorsque $n \rightarrow +\infty$, $E(X_n)^2 - 2mE(X_n) + m^2 \rightarrow m^2 - 2m^2 + m^2 = 0$, de sorte que

$$\lim_{n \rightarrow +\infty} P(|X_n - m| \geq \varepsilon) = 0.$$

Et donc $X_n \xrightarrow{P} m$.

SOLUTION DE L'EXERCICE 14.6

On sait que pour $k \in \mathbf{N}^*$, $P(Y_n \geq k) = P\left(\bigcap_{i=1}^n [X_i \geq k]\right)$ et par indépendance des X_i , il vient alors

$$P(Y_n \geq k) = \prod_{i=1}^n P(X_i \geq k) = P(X_1 \geq k)^n.$$

Or, nous savons que

$$P(X_1 \geq k) = \sum_{i=k}^{+\infty} P(X_1 = i) = \sum_{i=k}^{+\infty} p(1-p)^{i-1} = p(1-p)^{k-1} \sum_{j=0}^{+\infty} (1-p)^j = (1-p)^{k-1}.$$

Remarque

Notons que la valeur de cette variance ne nous sera pas utile pour appliquer la loi faible des grands nombres, et donc ne perdons pas de temps à la calculer.

Intuition

Si on voit X_1 comme le nombre d'essais avant le premier succès lors d'une répétition d'épreuves de Bernoulli de paramètre p , $[X_1 \geq k]$ est l'événement «les $k-1$ premiers essais ont été des échecs. Et donc on retrouve ainsi $(1-p)^{k-1}$ ».

Et donc $P(Y_n \geq k) = (1-p)^{n(k-1)}$.

On en déduit alors que

$$P(Y_n = k) = P(Y_n \geq k) - P(Y_n \geq k+1) = (1-p)^{n(k-1)} (1 - (1-p)^n).$$

Puisque les Y_n sont à valeurs dans \mathbf{Z} , il suffit d'étudier la limite de $P(Y_n = k)$.

Mais puisque $(1-p) \in]0, 1[$ on a $\lim_{n \rightarrow +\infty} (1-p)^n = 0$ et donc

$$\lim_{n \rightarrow +\infty} P(Y_n = k) = \begin{cases} 1 & \text{si } k = 1 \\ 0 & \text{sinon} \end{cases}$$

Et donc $Y_n \xrightarrow{\mathcal{L}} X$, où X est une variable aléatoire certaine égale à 1.

SOLUTION DE L'EXERCICE 14.7

1. Il est évident que f est positive et continue sur \mathbf{R} . De plus, pour $A < B$, on a

$$\int_A^B f(t) dt = \int_A^B \frac{e^t}{(1+e^t)^2} dt = \left[-\frac{1}{1+e^t} \right]_A^B = \frac{1}{1+e^B} - \frac{1}{1+e^A} \xrightarrow[A \rightarrow +\infty]{B \rightarrow +\infty} 1.$$

Et donc $\int_{-\infty}^{+\infty} f(t) dt$ converge et vaut 1, de sorte que f est bien une densité de probabilité.

- 2.a. Pour $x \in \mathbf{R}$, on a

$$F_{X_i}(x) = \int_{-\infty}^x f(t) dt = \lim_{A \rightarrow -\infty} \int_A^x \frac{e^t}{(1+e^t)^2} dt = \lim_{A \rightarrow -\infty} \left[-\frac{1}{1+e^t} \right]_A^x = 1 - \frac{1}{1+e^x}.$$

- 2.b. Pour $x \in \mathbf{R}$, on a

$$\begin{aligned} F_{M_n}(x) &= P(M_n \leq x) = P(\max(X_1, \dots, X_n) \leq x) = P\left(\bigcap_{i=1}^n [X_i \leq x]\right) \\ &= \prod_{i=1}^n P(X_i \leq x) = F_{X_1}(x)^n = \left(1 - \frac{1}{1+e^x}\right)^n. \end{aligned}$$

Les X_i sont indépendantes.

- 2.c. Pour $x \in \mathbf{R}$, on a

$$P(M_n - \ln(n) \leq x) = P(M_n \leq x + \ln(n)) = \left(1 - \frac{1}{1+e^{x+\ln(n)}}\right)^n = \left(1 - \frac{1}{1+ne^x}\right)^n = \exp\left(n \ln\left(1 - \frac{1}{1+ne^x}\right)\right).$$

Mais, lorsque $n \rightarrow +\infty$, $\ln\left(1 - \frac{1}{1+ne^x}\right) \underset{n \rightarrow +\infty}{\sim} -\frac{1}{1+ne^x}$.

Et donc $n \ln\left(1 - \frac{1}{1+ne^x}\right) \underset{n \rightarrow +\infty}{\sim} -\frac{n}{1+ne^x} \underset{n \rightarrow +\infty}{\sim} -\frac{1}{e^x} \xrightarrow{n \rightarrow +\infty} -e^{-x}$.

Et donc, la fonction exponentielle étant continue, par composition de limites

$$\left(1 - \frac{1}{1+e^{x+\ln n}}\right)^n \xrightarrow{n \rightarrow +\infty} e^{-e^{-x}}.$$

Posons alors $F(x) = e^{-e^{-x}}$. Alors F est \mathcal{C}^1 sur \mathbf{R} , croissante car sa dérivée est $e^{-x-e^{-x}} > 0$, et

$$\lim_{x \rightarrow -\infty} F(x) = 0 \text{ et } \lim_{x \rightarrow +\infty} F(x) = 1.$$

Donc F est la fonction de répartition d'une variable à densité.

Et si X est une variable aléatoire de fonction de répartition F , $M_n - \ln n \xrightarrow{\mathcal{L}} X$.

SOLUTION DE L'EXERCICE 14.8

Notons que bien que les Y_i soit identiquement distribuées, il n'est pas question d'appliquer la loi faible des grands nombres, car elles ne sont pas indépendantes².

En revanche, on peut³ essayer d'utiliser l'inégalité de Bienaymé-Tchebychev. On a

$$E(S_n) = \frac{1}{n} \sum_{i=1}^n E(Y_i) = 2p.$$

Interprétation

Ce résultat est assez logique : si on répète suffisamment de fois l'expérience, on va finir par obtenir un succès lors du premier essai, même s'il faudra sans doute beaucoup de répétitions pour en arriver là.

Composition

Rappelons que si on peut composer les limites, on ne peut pas en toute généralité composer les équivalents, ce qui explique l'usage des limites ici.

² En tous cas, aucun argument facile ne permet d'affirmer qu'elles le sont.

Le calcul de $\text{Cov}(Y_i, Y_{i+1})$ permettrait de prouver qu'elles ne sont pas indépendantes.

³ Comme dans la démonstration de la loi faible des grands nombres.

Ensuite, on a $S_n = \frac{X_1 + 2X_2 + \dots + 2X_n + X_{n+1}}{n}$.

Or, $X_1, 2X_2, \dots, 2X_n, X_{n+1}$ sont mutuellement indépendantes, et possèdent toutes une variance, à savoir

$$V(X_1) = V(X_{n+1}) = p(1-p) \text{ et } V(2X_2) = \dots = V(2X_n) = 4p(1-p).$$

Et donc S_n admet une variance égale à

$$\begin{aligned} V(S_n) &= \frac{1}{n^2} (V(X_1) + V(2X_2) + \dots + V(2X_n) + V(X_{n+1})) \\ &= \frac{1}{n^2} (V(X_1) + 4V(X_2) + \dots + 4V(X_n) + V(X_{n+1})) \\ &= \frac{p(1-p)}{n^2} (1 + 4 + \dots + 4 + 1) \\ &= \frac{1}{n^2} (4n + 2)p(1-p). \end{aligned}$$

Par l'inégalité de Bienaymé-Tchebychev, on a alors $\forall \epsilon > 0$,

$$P(|S_n - 2p| \geq \epsilon) \leq \frac{1}{n^2 \epsilon} (4n + 2)p(1-p) \xrightarrow{n \rightarrow +\infty} 0.$$

Par le théorème des gendarmes⁴, on a donc $P(|S_n - 2p| \geq \epsilon) \xrightarrow{n \rightarrow +\infty} 0$ et donc $S_n \xrightarrow{P} 2p$.

⁴ Une probabilité est toujours positive.

SOLUTION DE L'EXERCICE 14.9

- Il est clair que f est continue sauf peut être en 0 et positive. De plus, pour $A > 0$, on a

$$\int_{-\infty}^A f(t) dt = \int_0^A \frac{x}{\sigma^2} e^{-\frac{x^2}{2\sigma^2}} dx = \left[-e^{-\frac{x^2}{2\sigma^2}} \right]_0^A = 1 - e^{-\frac{A^2}{2\sigma^2}} \xrightarrow{A \rightarrow +\infty} 1.$$

Et donc $\int_{-\infty}^{+\infty} f(t) dt$ converge et vaut 1 : f est une densité de probabilités.

- Le calcul a en fait été fait dans la question précédente :

$$F_X(x) = \int_{-\infty}^x f(t) dt = \begin{cases} 0 & \text{si } x < 0 \\ 1 - e^{-\frac{x^2}{2\sigma^2}} & \text{si } x \geq 0 \end{cases}$$

- Notons que f est nulle sur \mathbf{R}_*^* , et donc les X_i prennent des valeurs positives. Il en est donc de même de Y_n , de sorte que $|Y_n| = Y_n$. Et donc

$$P(|Y_n| \geq \epsilon) = P(Y_n \geq \epsilon) = P\left(\bigcap_{i=1}^n [X_i \geq \epsilon]\right) = \prod_{i=1}^n P(X_i \geq \epsilon) = (1 - F_{X_1}(\epsilon))^n = e^{-n \frac{\epsilon^2}{2\sigma^2}}.$$

En particulier, $P(|Y_n| \geq \epsilon) \xrightarrow{n \rightarrow +\infty} 0$ et donc $Y_n \xrightarrow{P} 0$.

SOLUTION DE L'EXERCICE 14.10

Soit $\epsilon > 0$. Alors, puisque Y_n ne prend que des valeurs supérieures ou égales à a , on a $|Y_n - a| = Y_n - a$ et donc

$$P(|Y_n - a| \geq \epsilon) = P(Y_n - a \geq \epsilon) = P(Y_n \geq a + \epsilon).$$

Si $\epsilon > b - a$, alors $a + \epsilon > b$ et donc

$$P(Y_n \geq a + \epsilon) = 0 \xrightarrow{n \rightarrow \infty} 0.$$

Si $\epsilon \leq b - a$, alors, par indépendance des X_i

$$P(Y_n \geq a + \epsilon) = P\left(\bigcap_{i=1}^n [X_i \geq a + \epsilon]\right) = \prod_{i=1}^n P(X_i \geq a + \epsilon) = \left(1 - \frac{a + \epsilon - a}{b - a}\right)^n = \left(\frac{b - a - \epsilon}{b - a}\right)^n \xrightarrow{n \rightarrow \infty} 0.$$

Dans tous les cas, on a

$$\lim_{n \rightarrow \infty} P(|Y_n - a| \geq \epsilon) = 0$$

et donc $Y_n \xrightarrow{P} a$.

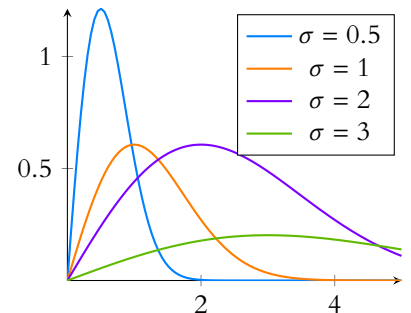


FIGURE 14.2- La densité f pour différentes valeurs de σ .

Facultatif
Montrer que la convergence a également lieu en loi.

Rappel
La fonction de répartition de X_i est
$$x \mapsto \begin{cases} 0 & \text{si } x \leq a \\ \frac{x-a}{b-a} & \text{si } a < x < b \\ 1 & \text{sinon} \end{cases}$$

SOLUTION DE L'EXERCICE 14.11

Notons tout de suite que les Y_n sont à valeurs dans $\llbracket 1, N \rrbracket$, et donc dans \mathbf{Z} .

Pour étudier la convergence en loi, il suffit donc d'étudier les $\lim_{n \rightarrow +\infty} P(Y_n = k)$, $k \in \llbracket 1, N \rrbracket$.

Mais pour $k \in \llbracket 1, N \rrbracket$, on a

$$P(Y_n \geq k) = P\left(\bigcap_{i=1}^n [X_i \geq k]\right) = P(X_1 \geq k)^n.$$

$$\text{Or, } P(X_1 \geq k) = \sum_{i=k}^N P(X_1 = i) = \sum_{i=k}^N \frac{1}{N} = \frac{N-k+1}{N}. \text{ Et donc } P(Y_n \geq k) = \left(\frac{N-k+1}{N}\right)^n.$$

Puisque Y_n est à valeurs entières,

$$P(Y_n = k) = P(Y_n \geq k) - P(Y_n \geq k+1) = \left(\frac{N-k+1}{N}\right)^n - \left(\frac{N-k}{N}\right)^n.$$

Or, k et N étant fixés, on a

$$\lim_{n \rightarrow +\infty} \left(\frac{N-k+1}{N}\right)^n = \begin{cases} 1 & \text{si } k = 1 \\ 0 & \text{sinon} \end{cases} \text{ et } \lim_{n \rightarrow +\infty} \left(\frac{N-k}{N}\right)^n = 0.$$

$$\text{Et donc } \lim_{n \rightarrow +\infty} P(Y_n = k) = \begin{cases} 1 & \text{si } k = 1 \\ 0 & \text{sinon} \end{cases}$$

On en déduit donc que (Y_n) converge en loi vers la variable certaine égale à 1.

Cette convergence a également lieu en probabilités car pour $\varepsilon > 0$, on a

$$P(|Y_n - 1| \geq \varepsilon) = P(Y_n - 1 \geq \varepsilon) = P(Y_n \geq 1 + \varepsilon) \leq P(Y_n \geq 2) = \left(\frac{N-1}{N}\right)^n \xrightarrow{n \rightarrow +\infty} 0.$$

Et donc $Y_n \xrightarrow{P} 1$.

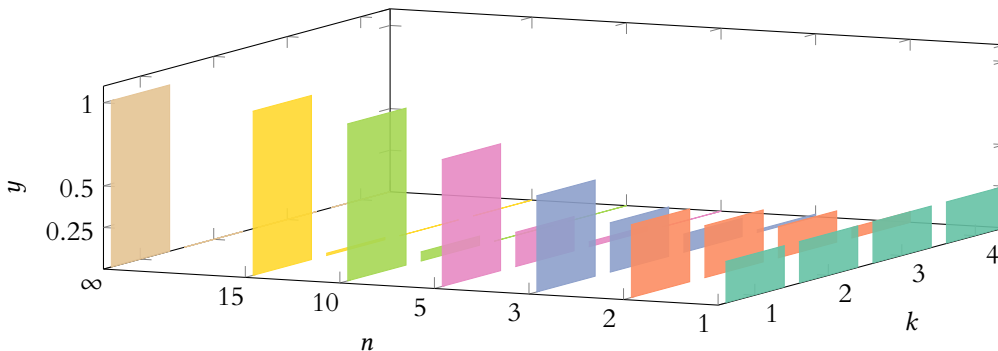


FIGURE 14.3 – La convergence en loi des Y_n vers 1 pour $N = 4$.

SOLUTION DE L'EXERCICE 14.12

Soit $\varepsilon > 0$. Alors on a $P(|X_n - X| \geq \varepsilon) = P(|Y_n| \geq \varepsilon)$.

Cette probabilité vaut 0 si $\varepsilon > 1$ et $\frac{1}{n}$ sinon.

Dans tous les cas elle tend vers 0 et donc $X_n \xrightarrow{P} X$.

Dans le cas où $X_n \hookrightarrow \mathcal{N}\left(\frac{1}{n}, \frac{1}{n^2}\right)$, on a

$$\begin{aligned} P(|Y_n| \geq \varepsilon) &= P(Y_n \geq \varepsilon) + P(Y_n \leq -\varepsilon) = P(nY_n - 1 \geq n\varepsilon - 1) + P(nY_n - 1 \leq -n\varepsilon - 1) \\ &= 1 - \Phi(n\varepsilon - 1) + \Phi(-n\varepsilon - 1) \xrightarrow{n \rightarrow +\infty} 0. \end{aligned}$$

Et donc $X_n \xrightarrow{P} X$.

Limite

Puisque N et k sont fixés, on a des suites géométriques. On distingue le cas $k = 1$, car alors $\frac{N-k+1}{N} = 1$.

Détails

Puisque Y_n ne prend que des valeurs entières, $Y_n \geq 1 + \varepsilon$ est réalisé dès que $Y_n \geq 2$. Si $\varepsilon \leq 1$ on a même

$$[Y_n \geq 1 + \varepsilon] = [Y_n \geq 2].$$

Limite

Lorsque $n \rightarrow +\infty$, alors $n\varepsilon - 1 \rightarrow +\infty$ et donc

$$\Phi(n\varepsilon - 1) \xrightarrow{n \rightarrow +\infty} 1$$

car $\Phi(x) \xrightarrow{x \rightarrow +\infty} 1$.

SOLUTION DE L'EXERCICE 14.13

1. Les X_i suivent la loi $\mathcal{E}(1)$, qui est également la loi $\gamma(1)$, et sont indépendantes, donc $S_n = \sum_{i=1}^n X_i$ suit la loi $\gamma(n)$. En particulier, on a

$$E(S_n) = V(S_n) = n.$$

2. Notons que la variable centrée réduite associée à S_n est $\frac{S_n - E(S_n)}{\sqrt{V(S_n)}} = \frac{S_n - n}{\sqrt{n}}$.

D'après le théorème central limite (qui s'applique car les X_n sont i.i.d et admettent une variance), $\left(\frac{S_n - n}{\sqrt{n}}\right)_n$ converge en loi vers une variable X suivant la loi normale centrée réduite.

En particulier, on a

$$\lim_{n \rightarrow \infty} P\left(\frac{S_n - n}{\sqrt{n}} \leq 0\right) = P(X \leq 0) = \frac{1}{2}.$$

Mais

$$P\left(\frac{S_n - n}{\sqrt{n}} \leq 0\right) = P(S_n - n \leq 0) = P(S_n \leq n).$$

Et donc

$$\lim_{n \rightarrow \infty} P(S_n \leq n) = \frac{1}{2}.$$

3. Rappelons qu'une densité de S_n est

$$f_n : t \mapsto \begin{cases} 0 & \text{si } t \leq 0 \\ \frac{t^{n-1}}{\Gamma(n)} e^{-t} & \text{si } t > 0 \end{cases} = \begin{cases} 0 & \text{si } t \leq 0 \\ \frac{t^{n-1}}{(n-1)!} e^{-t} & \text{si } t > 0 \end{cases}$$

Ainsi, pour tout n , on a

$$P(S_n \leq n) = \int_{-\infty}^n f_n(t) dt = \int_0^n \frac{t^{n-1}}{(n-1)!} e^{-t} dt$$

et alors le résultat de la question 2 nous indique que

$$\lim_{n \rightarrow \infty} \int_0^n \frac{t^{n-1}}{(n-1)!} e^{-t} dt = \frac{1}{2}.$$

SOLUTION DE L'EXERCICE 14.14

Soit $(X_n)_n$ une suite de variables aléatoires i.i.d. suivant la loi de Bernoulli de paramètre p .

Alors, par stabilité des lois binomiales, $\sum_{i=1}^n X_i$ suit la même loi que S_n .

Mais par le théorème central limite, si on note S_n^* la variable centrée réduite associée à $\sum_{i=1}^n X_i$, alors $S_n^* \xrightarrow{\mathcal{L}} X$, où $X \hookrightarrow \mathcal{N}(0, 1)$.

Et en particulier,

$$\lim_{n \rightarrow +\infty} P(S_n^* \leq 0) = P(X \leq 0) = \Phi(0) = \frac{1}{2}.$$

Mais

$$P(S_n^* \leq 0) = P\left(\frac{S_n - E(S_n)}{\sqrt{V(S_n)}} \leq 0\right) = P(S_n - np \leq 0) = P(S_n \leq np).$$

Et donc on en déduit que $\lim_{n \rightarrow +\infty} P(S_n \leq np) = \frac{1}{2}$.

Interprétation : nous savons que np est la moyenne de S_n .

Dire que $P(S_n \leq np)$ est environ égal à $\frac{1}{2}$ revient à dire que la médiane de S_n est environ égale à np .

Autrement dit, pour une loi binomiale, lorsque n tend vers l'infini, la moyenne et la médiane coïncident.

Rappel

L'espérance et la variance d'une loi $\gamma(\nu)$ sont toutes deux égales, et valent ν .

⚠ Danger

Attention à ne pas dire que

$$P\left(\frac{S_n - n}{\sqrt{n}} \leq 0\right)$$

est égale à $P(X \leq 0)$.

Pour n grand, ces probabilités sont sans doute proches, mais on ne peut faire l'économie de la limite.

Rappel

Pour tout entier $n > 0$, on a

$$\Gamma(n) = (n-1)!$$

Médiane

La médiane d'une variable aléatoire X est un nombre m , si un tel nombre existe, tel que $P(X \leq m) = \frac{1}{2}$.

SOLUTION DE L'EXERCICE 14.15

Notons X le nombre de crevaisons. Puisqu'on considère qu'à chaque kilomètre, la probabilité de crever vaut $\frac{1}{4000}$, le nombre X de crevaisons durant la course suit une loi binomiale $\mathcal{B}\left(6000, \frac{1}{4000}\right)$.

Les conditions d'approximation de X par une variable suivant une loi de Poisson de paramètre $\frac{6000}{4000} = \frac{3}{2}$ sont donc réunies.

Dans la suite, on considère donc que $X \hookrightarrow \mathcal{P}\left(\frac{3}{2}\right)$.

La probabilité que deux roues de secours suffisent est la probabilité d'avoir au plus deux crevaisons, soit $P(X \leq 2)$. Or

$$P(X \leq 2) = P(X = 0) + P(X = 1) + P(X = 2) = e^{-\frac{3}{2}} \left(1 + \frac{3}{2} + \frac{\left(\frac{3}{2}\right)^2}{2}\right) = e^{-\frac{3}{2}} \frac{29}{8} \approx 0.8.$$

Donc deux roues de secours ne suffisent pas.

En revanche, on a

$$P(X \leq 3) = e^{-\frac{3}{2}} \left(1 + \frac{3}{2} + \frac{\left(\frac{3}{2}\right)^2}{2} + \frac{\left(\frac{3}{2}\right)^3}{6}\right) = e^{-\frac{3}{2}} \frac{13}{3} \approx 0.96$$

donc trois roues de secours offrent une marge de sécurité suffisante.

SOLUTION DE L'EXERCICE 14.16

À un instant donné, la probabilité qu'un employé soit au téléphone est égale à $\frac{6}{60} = \frac{1}{10}$. Ainsi, le nombre d'employés en train de téléphoner est une variable aléatoire X qui suit une loi binomiale $\mathcal{B}(300, 1/10)$.

Les conditions d'approximation de la binomiale par la loi normale sont réunies⁵, donc on peut approximer X par une variable Y suivant la loi normale $\mathcal{N}(30, 27)$.

On cherche k tel que $P(Y > k) \leq 0.025$.

$$\text{Or, on a } P(Y > k) = 1 - P(Y \leq k) = 1 - \Phi\left(\frac{k - 30}{\sqrt{27}}\right).$$

On veut donc

$$\Phi\left(\frac{k - 30}{\sqrt{27}}\right) \geq 0.975.$$

Une table de loi normale nous indique qu'il faut

$$\frac{k - 30}{\sqrt{27}} \geq 1.96 \Leftrightarrow k \geq \sqrt{27} \times 1.96 + 30 \geq 40.1$$

Ainsi, il faut avoir au minimum 41 lignes de téléphone.

SOLUTION DE L'EXERCICE 14.17

Il est classique que la fonction de répartition de M_n est

$$F_{M_n}(x) = \begin{cases} 0 & \text{si } x \leq 0 \\ x^n & \text{si } x \in [0, 1] \\ 1 & \text{si } x \geq 1 \end{cases}$$

Alors, pour tout $x \in \mathbf{R}$, on a

$$F_{X_n}(x) = P(X_n \leq x) = P(n(1 - M_n) \leq x) = P\left(M_n \geq \frac{n - x}{n}\right) = 1 - F_{M_n}\left(1 - \frac{x}{n}\right) = 1 - F_{M_n}\left(1 - \frac{x}{n}\right).$$

De plus, on a $\frac{n-x}{n} \leq 0$ si et seulement si $x \geq n$ et $\frac{n-x}{n} \geq 1 \Leftrightarrow x \leq 0$. On en déduit que

$$F_{X_n}(x) = \begin{cases} 0 & \text{si } x \leq 0 \\ 1 - \left(1 - \frac{x}{n}\right)^n & \text{si } x \in [0, n] \\ 1 & \text{sinon} \end{cases}$$

⁵ En effet, on a $300 \geq 30$, $300 \times 1/10 \geq 15$ et $300 \times 9/10 \geq 15$.

Explication

On cherche la probabilité que le nombre de personnes téléphonant en même temps soit supérieur au nombre k de lignes.

Il est clair que si $x \leq 0$, alors $F_{M_n}(x) \xrightarrow{n \rightarrow \infty} 0$.

Si $x > 0$, alors pour n suffisamment grand, on a $x \leq n$, et donc

$$F_{X_n}(x) = 1 - \left(1 - \frac{x}{n}\right)^n.$$

Or, nous savons que

$$\left(1 - \frac{x}{n}\right)^n = e^{n \ln\left(1 - \frac{x}{n}\right)}.$$

Puisque x est fixé, $\frac{x}{n} \xrightarrow{n \rightarrow \infty} 0$, et donc $\ln\left(1 - \frac{x}{n}\right) \sim -\frac{x}{n}$, de sorte que

$$n \ln\left(1 - \frac{x}{n}\right) \sim -n \frac{x}{n} = -x \xrightarrow{n \rightarrow \infty} -x.$$

Par continuité de l'exponentielle, on en déduit que

$$F_{X_n}(x) \xrightarrow{n \rightarrow \infty} 1 - e^{-x}.$$

On reconnaît là la fonction de répartition d'une variable aléatoire de loi exponentielle $\mathcal{E}(1)$. Donc (X_n) converge en loi vers une variable aléatoire de loi exponentielle $\mathcal{E}(1)$.

SOLUTION DE L'EXERCICE 14.18

- Il s'agit quasiment d'une tautologie : soit $\varepsilon > 0$ fixé. Alors l'événement $[|u_n - \ell| \geq \varepsilon]$ n'est pas un événement aléatoire : il est réalisé, ou il ne l'est pas. Or, nous savons⁶, qu'il existe $n_0 \in \mathbf{N}$ tel que pour $n \geq n_0$, $|u_n - \ell| < \varepsilon$. Et donc pour $n \geq n_0$, on a $P(|u_n - \ell| \geq \varepsilon) = 0$, de sorte que

$$\lim_{n \rightarrow +\infty} P(|u_n - \ell| \geq \varepsilon) = 0 \text{ et donc } u_n \xrightarrow{P} \ell.$$

- $X_n \xrightarrow{\mathcal{L}} X$ et $u_n \xrightarrow{P} \ell$ où ℓ est une constante. Et donc, par le lemme de Slutsky, on a $X_n - u_n \xrightarrow{\mathcal{L}} X - \ell$.
- On a $P(X_n \leq u_n) = P(X_n - u_n \leq 0)$. Puisque X est à densité, il en est de même⁷ de $X - \ell$, et donc la fonction de répartition de $X - \ell$ est continue en 0, de sorte que, par définition de la convergence en loi,

$$\lim_{n \rightarrow +\infty} P(X_n \leq u_n) = \lim_{n \rightarrow +\infty} P(X_n - u_n \leq 0) = P(X - \ell \leq 0) = P(X \leq \ell).$$

SOLUTION DE L'EXERCICE 14.19

- Par le théorème de transfert, on a, sous réserve de convergence⁸

$$\begin{aligned} E(e^{tX}) &= \sum_{n=0}^{+\infty} e^{tn} P(X = n) \\ &= \sum_{n=0}^{+\infty} e^{tn} \frac{\lambda^n}{n!} e^{-\lambda} \\ &= e^{-\lambda} \sum_{n=0}^{+\infty} (e^t)^n \frac{\lambda^n}{n!} \\ &= e^{-\lambda} \exp(e^t \lambda) \\ &= \exp(\lambda(e^t - 1)). \end{aligned}$$

- Notons que $[X \geq n] = [e^{tX} \geq e^{tn}]$ et donc par l'inégalité de Markov appliquée à e^{tX} , qui est bien positive et admet une espérance,

$$P(X \geq n) = P(e^{tX} \geq e^{tn}) \leq \frac{E(e^{tX})}{e^{tn}} \leq e^{-tn} E(e^{tX}).$$

Suffisamment gd.

Puisque nous calculons une limite, seuls les n suffisamment grands ont une importance.

Et donc à x fixé, on finit par avoir $n \geq x$.

⁶ C'est la définition de la convergence d'une suite.

Précision

L'énoncé de Slutsky dirait plutôt que $X_n + u_n \xrightarrow{\mathcal{L}} X + \ell$, mais puisque $u_n \xrightarrow{P} \ell$, on a également $-u_n \xrightarrow{P} -\ell$.

⁷ Une transformée affine de variable à densité est à densité.

⁸ Absolue, mais il s'agit d'une série à termes positifs.

On a reconnu une série exponentielle, donc convergente, donc $E(e^{tX})$ existe.

3. En utilisant ce qui précède, il vient

$$P(X \geq n) \leq e^{-tn} \exp(\lambda(e^t - 1)).$$

Ceci étant vrai pour tout $t \geq 0$, cherchons la valeur de t qui minimise

$$e^{-tn} \exp(\lambda(e^t - 1)) = \exp(-tn + \lambda(e^t - 1)).$$

Soit donc $f(t) = -tn + \lambda(e^t - 1)$. Alors f est dérivable, et $f'(t) = -n + \lambda e^t$, de sorte que $f'(t) = 0 \Leftrightarrow t = \ln\left(\frac{n}{\lambda}\right)$.

Et donc f est décroissante sur $\left[0, \ln\left(\frac{n}{\lambda}\right)\right]$ et croissante sur $\left[\ln\left(\frac{n}{\lambda}\right), +\infty\right[$.

Elle admet donc un minimum en $\ln\left(\frac{n}{\lambda}\right)$, qui vaut

$$f\left(\ln\left(\frac{n}{\lambda}\right)\right) = -n \ln\left(\frac{n}{\lambda}\right) + \lambda\left(\frac{n}{\lambda} - 1\right) = -n \ln\left(\frac{n}{\lambda}\right) + n - \lambda.$$

On en déduit, pour $t = \ln\left(\frac{n}{\lambda}\right)$ que

$$P(X \geq n) \leq \exp\left(-n \ln\left(\frac{n}{\lambda}\right) + n - \lambda\right) \leq \left(\frac{\lambda}{n}\right)^n e^n e^{-\lambda}.$$

Mais puisque d'autre part, X est à valeurs entières, on a

$$P(X \geq n) = \sum_{k=n}^{+\infty} P(X = k) = e^{-\lambda} \sum_{k=n}^{+\infty} \frac{\lambda^k}{k!}.$$

Et donc il vient, après division par $e^{-\lambda}$,

$$\sum_{k=n}^{+\infty} \frac{\lambda^k}{k!} \leq \left(\frac{\lambda}{n}\right)^n e^n.$$

SOLUTION DE L'EXERCICE 14.20

Puisque nous voulons que $X_n \xrightarrow{P} 0$, il semble légitime d'essayer d'utiliser des variables aléatoires qui prennent la valeur 0, avec une probabilité qui tend vers 1.

Essayons par exemple $P(X_n = 0) = 1 - \frac{1}{n}$.

On a alors, pour $\varepsilon > 0$,

$$P(|X_n - 0| \geq \varepsilon) \leq P(X_n \neq 0) \leq \frac{1}{n} \xrightarrow{n \rightarrow +\infty} 0.$$

Et donc $X_n \xrightarrow{P} 0$.

Notons alors a_n l'autre valeur prise par X_n , de sorte que $P(X_n = a_n) = \frac{1}{n}$.

On a alors $E(X_n) = \frac{a_n}{n}$ et $E(X_n^2) = \frac{a_n^2}{n}$.

Par la formule de Huygens, on a $V(X_n) = E(X_n^2) - E(X_n)^2$, et donc si on arrive à avoir $E(X_n) \xrightarrow{n \rightarrow +\infty} 1$ et $E(X_n^2) \xrightarrow{n \rightarrow +\infty} +\infty$, alors automatiquement, on aura $V(X_n) \xrightarrow{n \rightarrow +\infty} +\infty$.

La condition $E(X_n) \xrightarrow{n \rightarrow +\infty} 1$ est satisfaite si et seulement si $a_n \sim_{n \rightarrow +\infty} n$.

Prenons donc $a_n = n$. On a alors $E(X_n^2) = n \xrightarrow{n \rightarrow +\infty} +\infty$.

Ainsi, la suite de variables aléatoires définies par

$$X_n(\Omega) = \{0, n\}, P(X_n = 0) = 1 - \frac{1}{n}, P(X_n = n) = \frac{1}{n}$$

vérifie bien toutes les conditions requises par l'énoncé.

Commentaires : • les calculs n'ont rien de difficile, toute la difficulté réside ici dans la prise d'initiative : comment démarrer un tel exercice ? Il faut ici un peu d'intuition pour penser à chercher des variables avec $P(X_n = 0) = 1 - \frac{1}{n}$.

• Il n'y a pas du tout unicité de la solution, par exemple, en gardant $P(X_n = 0) = 1 - \frac{1}{n}$, prendre n'importe quelle suite (a_n) équivalente à n pour la seconde valeur de X_n convenait. De même, nous aurions pu prendre $P(X_n = 0) = 1 - \frac{1}{n^2}$ ou même plus généralement $P(X_n = u_n) = v_n$ avec (u_n) n'importe quelle suite de limite nulle et (v_n) n'importe quelle suite de limite égale à 1.

SOLUTION DE L'EXERCICE 14.21

1. C'est un simple calcul..
 2.a. Si on a à la fois $|X_n - X||Y| \leq \frac{\varepsilon}{3}$, $|Y_n - Y||X| \leq \frac{\varepsilon}{3}$ et $|X_n - X| \cdot |Y_n - Y| \leq \frac{\varepsilon}{3}$.
 Alors, par l'inégalité triangulaire, on a

$$\begin{aligned} |X_n Y_n - XY| &= |(X_n - X)Y + (Y_n - Y)X + (X_n - X)(Y_n - Y)| \\ &\leq |X_n - X||Y| + |Y_n - Y||X| + |X_n - X| \cdot |Y_n - Y| \\ &\leq \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} \leq \varepsilon. \end{aligned}$$

Autrement dit, nous venons de prouver que $\overline{B_n} \cap \overline{C_n} \cap \overline{D_n} \subset [|X_n Y_n - XY| \leq \varepsilon]$.
 En passant aux événements contraires, il vient

$$|X_n Y_n - XY| > \varepsilon \subset B_n \cup C_n \cup D_n.$$

Et donc

$$P(|X_n Y_n - XY| > \varepsilon) \leq P(B_n \cup C_n \cup D_n) \leq P(B_n) + P(C_n) + P(D_n).$$

- 3.a.i. On a $P(|Y| > t) = 1 - P(|Y| \leq t)$.
 Or, $|Y|$ est une variable aléatoire sur (Ω, \mathcal{A}, P) , donc sa fonction de répartition tend vers 1 en $+\infty$ et donc $\lim_{t \rightarrow +\infty} P(|Y| > t) = 0$. Et donc⁹ il existe $A > 0$ tel que pour $t \geq A$, $g(t) < \frac{\delta}{2}$.
 Et donc pour $t \geq A$, $0 \leq P(|Y| > t) \leq g(t) < \frac{\delta}{2}$.

⁹ C'est la **définition** d'une limite.

- 3.a.ii. Comme à la question 2, on a

$$[|Y| \leq t] \cap \left[|X_n - X| \leq \frac{\varepsilon}{3t} \right] \subset \underbrace{\left[|X_n - X||Y| \leq \frac{\varepsilon}{3} \right]}_{= \overline{B_n}}.$$

Et donc

$$P(B_n) \leq P\left(|X_n - X| > \frac{\varepsilon}{3t}\right) + P(|Y| > t).$$

- 3.a.iii. C'est la définition de $X_n \xrightarrow{P} X$: puisque $\lim_{n \rightarrow +\infty} P\left(|X_n - X| > \frac{\varepsilon}{3t}\right) = 0$, il existe donc $N_0 \in \mathbf{N}$ tel pour $n \geq N_0$, $P\left(|X_n - X| > \frac{\varepsilon}{3t}\right) < \frac{\delta}{2}$.

- 3.a.iv. Soit $\delta > 0$ fixé. Et soit alors A comme à la question 3.a, et $t > A$.

On a alors, pour tout $n \in \mathbf{N}$, $P(B_n) \leq P\left(|X_n - X| > \frac{\varepsilon}{3t}\right) + \frac{\delta}{2}$.

Et en particulier, si N_0 est comme à la question 3.c, pour $n \geq N_0$, on a $0 \leq P(B_n) < \delta$.

Autrement dit, nous venons de prouver que pour tout $\delta > 0$, il existe $N_0 \in \mathbf{N}$ tel que pour $n \geq N_0$, $P(B_n) < \delta$.

Nous reconnaissons là la définition de $\lim_{n \rightarrow +\infty} P(B_n) = 0$.

- 3.b. Toujours comme à la question 2, on a, par passage aux événements contraires,

$$\left[|X_n - X| \leq \sqrt{\frac{\varepsilon}{3}} \right] \cap \left[|Y_n - Y| \leq \sqrt{\frac{\varepsilon}{3}} \right] \subset \underbrace{\left[|X_n - X||Y_n - Y| \leq \frac{\varepsilon}{3} \right]}_{= \overline{D_n}}.$$

Et donc

$$P(D_n) \leq P\left(\left[|X_n - X| > \sqrt{\frac{\varepsilon}{3}}\right] \cup \left[|Y_n - Y| > \sqrt{\frac{\varepsilon}{3}}\right]\right) \leq P\left(\left[|X_n - X| > \sqrt{\frac{\varepsilon}{3}}\right]\right) + P\left(\left[|Y_n - Y| > \sqrt{\frac{\varepsilon}{3}}\right]\right).$$

Puisque $X_n \xrightarrow{P} X$, on a $\lim_{n \rightarrow +\infty} P\left(\left[|X_n - X| > \sqrt{\frac{\varepsilon}{3}}\right]\right) = 0$.

Et de même, $Y_n \xrightarrow{P} Y$ et donc $\lim_{n \rightarrow +\infty} P\left(\left[|Y_n - Y| > \sqrt{\frac{\varepsilon}{3}}\right]\right) = 0$.

On en déduit que $\lim_{n \rightarrow +\infty} P(D_n) = 0$.

3.c. D'après la question 2, et les résultats de 3.d et 4, on a

$$0 \leq P(A_n) \leq \underbrace{P(B_n) + P(C_n) + P(D_n)}_{\xrightarrow[n \rightarrow +\infty]{} 0}.$$

Et donc par le théorème des gendarmes, $\lim_{n \rightarrow +\infty} P(A_n) = 0$.

Autrement dit, $X_n Y_n \xrightarrow{P} XY$.

SOLUTION DE L'EXERCICE 14.22

1. F étant continue en x , il existe $\eta > 0$ tel que pour $y \in]x - \eta, x + \eta[$, $|F(y) - F(x)| < \delta$.

Si l'on prend par exemple $\varepsilon = \frac{\eta}{2}$, alors on a bien

$$|F(x - \varepsilon) - F(x)| < \delta \text{ et } |F(x + \varepsilon) - F(x)| < \delta$$

et donc $F(x - \varepsilon) > F(x) - \delta$ et $F(x + \varepsilon) < F(x) + \delta$.

2. Si $X_n \leq x$, alors soit $|X_n - X| > \varepsilon$, soit $|X_n - X| \leq \varepsilon$, et alors

$$X \leq |X| \leq |X - X_n + X_n| \leq |X - X_n| + |X_n| \leq \varepsilon + x.$$

On a donc bien

$$[X_n \leq x] \subset [X \leq x + \varepsilon] \cup [|X_n - X| > \varepsilon].$$

En remplaçant x par $x - \varepsilon$ et en inversant le rôle de X_n et X , on obtient alors la seconde inclusion demandée.

3. En passant aux probabilités on a

$$F(x) = P(X_n \leq x) \leq F(x + \varepsilon) + P(|X_n - X| > \varepsilon).$$

Et de même,

$$F(x - \varepsilon) \leq F_n(x) + P(|X_n - X| > \varepsilon).$$

En combinant ces inégalités au résultat de la question 1, on obtient alors

$$F_n(x) \leq F(x) + \delta + P(|X_n - X| > \varepsilon) \text{ et } F(x) - \delta \leq F_n(x) + P(|X_n - X| > \varepsilon).$$

4. Puisque $X_n \xrightarrow{P} X$, alors il existe $N \in \mathbf{N}$ tel que pour $n \geq N$, $P(|X_n - X| > \varepsilon) \leq \delta$. Et donc pour $n \geq N$, il vient

$$F_n(x) \leq F(x) + 2\delta \text{ et } F(x) - 2\delta \leq F_n(x)$$

soit encore $|F_n(x) - F(x)| \leq 2\delta$.

Ceci étant valable pour tout $\delta > 0$, on en déduit que $\lim_{n \rightarrow +\infty} F_n(x) = F(x)$.

Et comme ce résultat est vrai en tout point de continuité de F , on a bien $X_n \xrightarrow{\mathcal{L}} X$.

Rappel

La probabilité d'une union est inférieure ou égale à la somme des probabilités.

ENDOMORPHISMES ET MATRICES SYMÉTRIQUES

Cours	461
15.1 Endomorphismes symétriques et matrices symétriques	461
15.2 Diagonalisation des endomorphismes et des matrices symétriques	462
15.3 Forme quadratique associée à une matrice symétrique	467
Méthodes	469
Exercices	470
Correction des exercices	472

Dans tout ce chapitre, E désigne un espace euclidien de dimension n , dont on note le produit scalaire $\langle \cdot, \cdot \rangle$ et la norme associée $\| \cdot \|$.
On munit $\mathcal{M}_{n,1}(\mathbf{R})$ de son produit scalaire canonique, c'est-à-dire $\langle X, Y \rangle = {}^tXY$, pour lequel la norme est donnée par

$$\left\| \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \right\|^2 = \sum_{i=1}^n x_i^2.$$

Rappelons que la base canonique de $\mathcal{M}_{n,1}(\mathbf{R})$ est orthonormée pour ce produit scalaire.

15.1 ENDOMORPHISMES SYMÉTRIQUES ET MATRICES SYMÉTRIQUES

15.1.1 Définition

Définition 15.1 – Soit $f \in \mathcal{L}(E)$. f est un endomorphisme symétrique si

$$\forall (x, y) \in E^2, \langle f(x), y \rangle = \langle x, f(y) \rangle.$$

Exemple 15.2 Un projecteur orthogonal est symétrique

Soit p_F un projecteur orthogonal. Alors p_F est un endomorphisme symétrique.
En effet, soient $x, y \in E$. Alors avec les notations habituelles¹,

$$\langle p_F(x), y \rangle = \langle x_F, y_F + y_{F^\perp} \rangle = \langle x_F, y_F \rangle + \underbrace{\langle x_F, y_{F^\perp} \rangle}_{=0} = \langle x_F, y_F \rangle$$

et de même

$$\langle x, p_F(y) \rangle = \langle x_F + x_{F^\perp}, y_F \rangle = \langle x_F, y_F \rangle + \underbrace{\langle x_{F^\perp}, y_F \rangle}_{=0} = \langle x_F, y_F \rangle$$

et donc $\langle p_F(x), y \rangle = \langle x, p_F(y) \rangle$.

¹ C'est-à-dire $x = x_F + x_{F^\perp}$, avec $x_F \in F$ et $x_{F^\perp} \in F^\perp$ et de même pour y .

15.1.2 Propriétés

Proposition 15.3 : Soit $f \in \mathcal{L}(E)$ et soit $\mathcal{B} = (e_1, \dots, e_n)$ une base *orthonormée* de E . Alors f est un endomorphisme symétrique si et seulement si $\text{Mat}_{\mathcal{B}}(f)$ est symétrique.

Démonstration. Supposons que $A = \text{Mat}_{\mathcal{B}}(f)$ soit symétrique. Alors, $\forall (x, y) \in E^2$, si $x_{\mathcal{B}}$ et $y_{\mathcal{B}}$ désignent les vecteurs colonnes des coordonnées de x et y dans la base \mathcal{B} , on a

$$\langle x, f(y) \rangle = {}^t x_{\mathcal{B}} A y_{\mathcal{B}} \text{ et } \langle f(x), y \rangle = {}^t (A x_{\mathcal{B}}) y_{\mathcal{B}} = {}^t x_{\mathcal{B}} {}^t A y_{\mathcal{B}} = {}^t x_{\mathcal{B}} A y_{\mathcal{B}}.$$

Donc f est un endomorphisme symétrique.

Inversement, soit f un endomorphisme symétrique, et soit $A = (a_{i,j}) = \text{Mat}_{\mathcal{B}}(f) \in \mathcal{M}_n(\mathbf{R})$.

Alors $f(e_j) = \sum_{i=1}^n a_{i,j} e_i$, et puisque \mathcal{B} est orthonormée, $\langle f(e_j), e_k \rangle = \sum_{i=1}^n a_{i,j} \langle e_i, e_k \rangle = a_{k,j}$.

De même, $\langle e_j, f(e_k) \rangle = a_{j,k}$.

Mais $\langle f(e_j), e_k \rangle = \langle e_j, f(e_k) \rangle$ et donc $a_{k,j} = a_{j,k}$: A est symétrique. \square

Proposition 15.4 : Soit $p \in \mathcal{L}(E)$ un projecteur. Alors p est un projecteur orthogonal si et seulement si p est un endomorphisme symétrique.
Dans ce cas, p est alors la projection orthogonale sur $\text{Im}(p)$, et $(\text{Im}(p))^{\perp} = \text{Ker } p$.

Démonstration. Nous avons déjà vu précédemment qu'un projecteur orthogonal est un endomorphisme symétrique, mais donnons-en tout de même² une seconde preuve.

Soit p un projecteur orthogonal, et soit $F = \text{Im } p$.

Considérons alors (e_1, \dots, e_p) une base orthonormée de F et (e_{p+1}, \dots, e_n) une base orthonormée de F^{\perp} , de sorte que $\mathcal{B} = (e_1, \dots, e_n)$ est une base orthonormée de E .

Alors nous savons³ que la matrice de p dans cette base est $\text{Diag}(\underbrace{1, \dots, 1}_{p \text{ fois}}, \underbrace{0, \dots, 0}_{n-p \text{ fois}})$.

Cette matrice étant symétrique, et la base \mathcal{B} étant orthonormée, p est un endomorphisme symétrique.

Inversement, soit p un projecteur symétrique.

Notons $F = \text{Im}(p)$ et $G = \text{Ker}(p)$, de sorte que p est la projection sur F parallèlement à G .

Soit $x \in F$. On a alors $p(x) = x$ et donc pour tout $y \in G$,

$$\langle x, y \rangle = \langle p(x), y \rangle = \langle x, p(y) \rangle = \langle x, 0_E \rangle = 0.$$

Cela étant vrai pour tout $x \in F$ et tout $y \in G$, alors $F \subset G^{\perp}$.

Mais $\dim G^{\perp} = \dim E - \dim G = \dim E - \dim \text{Ker}(p) = \dim \text{Im}(p) = \dim F$, et donc $F = G^{\perp}$.

Ainsi, p est la projection sur F , parallèlement à F^{\perp} : c'est le projecteur orthogonal sur $F = \text{Im } p$, et $(\text{Im } p)^{\perp} = F^{\perp} = G = \text{Ker } p$. \square

Proposition 15.5 : Soit f un endomorphisme symétrique, et soit F un sous-espace vectoriel de E , stable par f . Alors F^{\perp} est également stable par f .

Démonstration. Soit $y \in F^{\perp}$. Alors, $\forall x \in F$,

$$\langle f(y), x \rangle = \langle y, \underbrace{f(x)}_{\in F} \rangle = 0.$$

On en déduit que $f(y) \in F^{\perp}$, et donc F^{\perp} est stable par f . \square

15.2 DIAGONALISATION DES ENDOMORPHISMES ET DES MATRICES SYMÉTRIQUES

15.2.1 Valeurs propres d'un endomorphisme symétrique

Rappelons qu'une matrice $A \in \mathcal{M}_n(\mathbf{R})$ peut être vue comme une matrice de $\mathcal{M}_n(\mathbf{C})$. Par conséquent, on peut aussi bien s'intéresser à ses valeurs propres réelles qu'à ses valeurs propres complexes.

Plus précisément, si $A \in \mathcal{M}_n(\mathbf{R})$, alors $\lambda \in \mathbf{C}$ est valeur propre de A si et seulement si il existe $X \in \mathcal{M}_{n,1}(\mathbf{C})$ non nul tel que $AX = \lambda X$.

Cela n'implique pas nécessairement qu'il existe un tel X à coefficients réels.

Rappel

Le calcul du produit scalaire à l'aide des vecteurs colonnes des coordonnées n'est possible que si la base \mathcal{B} est orthonormée.

² Pour le plaisir !

³ C'est la proposition 5.39.

Méthode

Si on prouve que $\forall x \in F, \forall y \in G, \langle x, y \rangle = 0$, alors on prouve seulement que $F \subset G^{\perp}$ (ou de manière équivalente, que $G \subset F^{\perp}$). Mais pas que $F = G^{\perp}$: il pourrait a priori y avoir d'autres vecteurs dans G^{\perp} que les vecteurs de F .

Cas particuliers

$(\text{Im } f)^{\perp}$ et $(\text{Ker } f)^{\perp}$ sont stables par f .

Par exemple, $A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ est une matrice réelle qui n'a pas de valeurs propres dans \mathbf{R} car le polynôme $X^2 - \text{tr}(A)X + \det(A) = X^2 + 1$ n'a pas de racines réelles. En revanche, ce polynôme possède deux racines complexes qui sont i et $-i$, donc A possède i et $-i$ comme valeurs propres complexes. Pour trouver les vecteurs propres complexes associés à la valeur propre i , il faut résoudre le système $AX = iX$, où $X = \begin{pmatrix} x \\ y \end{pmatrix}$ est un vecteur complexe.

On trouve alors, après calcul, $E_i(A) = \text{Vect} \left(\begin{pmatrix} i \\ -1 \end{pmatrix} \right) = \left\{ \lambda \begin{pmatrix} -i \\ 1 \end{pmatrix}, \lambda \in \mathbf{C} \right\}$.

Proposition 15.6 : Soit $A \in \mathcal{M}_n(\mathbf{C})$. Alors A possède au moins une valeur propre complexe.

Démonstration. Nous savons que P possède un polynôme annulateur non nul $P \in \mathbf{C}[X]$. Mais par le théorème de d'Alembert-Gauss, il existe des complexes $\lambda_1, \dots, \lambda_p$ et $\alpha \neq 0$ tels que

$$P = \alpha \prod_{i=1}^p (X - \lambda_i).$$

Alors P étant annulateur de A , les éventuelles valeurs propres de A sont parmi $\lambda_1, \dots, \lambda_d$.

Supposons par l'absurde qu'aucun des λ_i ne soit valeur propre de A . Alors

$$P(A) = \alpha \prod_{i=1}^p (A - \lambda_i I_n) = 0.$$

Les λ_i n'étant pas valeurs propres de A , les $A - \lambda_i I_n$ sont toutes inversibles.

Mais alors $0 = \prod_{i=1}^p (A - \lambda_i I_n)$ est également inversible⁴, ce qui est absurde.

Par conséquent, l'un⁵ des λ_i est valeur propre de A . □

Corollaire 15.7 – Soit $A \in \mathcal{M}_n(\mathbf{R})$. Alors A possède une valeur propre complexe.

Proposition 15.8 : Soit $A \in \mathcal{M}_n(\mathbf{R})$ une matrice symétrique **réelle**. Alors A possède au moins une valeur propre réelle, et toutes les valeurs propres de A sont réelles.

Démonstration. A possède au moins une valeur propre complexe. Soit $\lambda \in \mathbf{C}$ une valeur propre complexe de A , et soit $X \in \mathcal{M}_{n,1}(\mathbf{C})$ un vecteur propre de A associé à la valeur propre λ .

Montrons que λ est en fait réel en prouvant que $\bar{\lambda} = \lambda$.

D'une part, on a $AX = \lambda X$, donc $A\bar{X} = \bar{A}\bar{X} = \bar{A}\bar{X} = \bar{\lambda}\bar{X} = \bar{\lambda}\bar{X}$.

Et alors ${}^t X A \bar{X} = {}^t X \bar{\lambda} \bar{X} = \bar{\lambda} {}^t X \bar{X}$.

Mais d'autre part, $AX = \lambda X \Leftrightarrow {}^t (AX) = {}^t X A = \lambda {}^t X$, de sorte que

$${}^t X A \bar{X} = \lambda {}^t X \bar{X}.$$

On en déduit que $\lambda \underbrace{{}^t X \bar{X}}_{\in \mathbf{C}} = \bar{\lambda} \underbrace{{}^t X \bar{X}}_{\in \mathbf{C}}$, soit encore $(\lambda - \bar{\lambda}) {}^t X \bar{X} = 0$.

Mais si $X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$, alors ${}^t X \bar{X} = (x_1 \quad \dots \quad x_n) \begin{pmatrix} \bar{x}_1 \\ \vdots \\ \bar{x}_n \end{pmatrix} = \sum_{i=1}^n x_i \bar{x}_i = \sum_{i=1}^n |x_i|^2 \neq 0$ car $X \neq 0$.

C'est donc $\lambda - \bar{\lambda} = 0 \Leftrightarrow \lambda = \bar{\lambda}$, et donc $\lambda \in \mathbf{R}$.

Ainsi, toutes les valeurs propres de A sont des réels. □

Danger !
Ce résultat n'est pas vrai si on parle de valeurs propres réelles. Par exemple, nous venons de voir que $\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ n'admet pas de valeurs propres réelles.

⁴ Car un produit de matrices inversibles est inversible.

⁵ Au moins.

Conjugaison
Lorsqu'on note \bar{X} , on entend par là le vecteur colonne dont les composantes sont les conjugués des composantes de X .
De même lorsqu'on note \bar{A} . Ici, A est à coefficients réels, donc $\bar{A} = A$.

Conjugué/module
Rappelons que pour tout complexe z , on a
 $z \cdot \bar{z} = |z|^2$.

Corollaire 15.9 – Si $f \in \mathcal{L}(E)$ est un endomorphisme symétrique, alors f possède au moins une valeur propre.

Démonstration. Soit \mathcal{B} une base orthonormée de E . Alors $\text{Mat}_{\mathcal{B}}(f)$ est symétrique, à coefficients réels, et donc possède une valeur propre réelle λ .
Et donc λ est également une valeur propre de f . \square

Proposition 15.10 : Soit $f \in \mathcal{L}(E)$ un endomorphisme symétrique. Alors les sous-espaces propres de f sont deux à deux orthogonaux.
Plus précisément : si e_1, \dots, e_p sont des vecteurs propres de f associés à des valeurs propres distinctes, alors la famille (e_1, \dots, e_p) est orthogonale.

Démonstration. Notons $\lambda_i \in \mathbf{R}$ tel que $f(x_i) = \lambda_i x_i$, et prouvons que si $i \neq j$, alors e_i et e_j sont orthogonaux.

Puisque f est symétrique, on a $\langle f(e_i), e_j \rangle = \langle e_i, f(e_j) \rangle$.

Or, $f(e_i) = \lambda_i e_i$ et donc $\langle f(e_i), e_j \rangle = \langle \lambda_i e_i, e_j \rangle = \lambda_i \langle e_i, e_j \rangle$.

De même, $\langle e_i, f(e_j) \rangle = \langle e_i, \lambda_j e_j \rangle = \lambda_j \langle e_i, e_j \rangle$.

Et donc $\lambda_i \langle e_i, e_j \rangle = \lambda_j \langle e_i, e_j \rangle \Leftrightarrow (\lambda_i - \lambda_j) \langle e_i, e_j \rangle = 0$.

Puisque $\lambda_i \neq \lambda_j$, on a donc $\langle e_i, e_j \rangle = 0$, et donc (e_1, \dots, e_p) est orthogonale. \square

⁶ Par linéarité à gauche du produit scalaire.

Corollaire 15.11 – Soit $A \in \mathcal{M}_n(\mathbf{R})$ une matrice symétrique. Alors les sous-espaces propres de A sont deux à deux orthogonaux.

Démonstration. Soit f_A l'endomorphisme de $\mathcal{M}_{n,1}(\mathbf{R})$ canoniquement associé⁷ à A .

Puisque la base canonique de $\mathcal{M}_{n,1}(\mathbf{R})$ est orthonormée, f_A est un endomorphisme symétrique, et donc la proposition précédente s'y applique. \square

⁷ c'est-à-dire dont la matrice dans la base canonique est A . Rappelons que $f_A(X) = AX$.

15.2.2 Diagonalisation des endomorphismes symétriques

Théorème 15.12 : Soit $f \in \mathcal{L}(E)$ un endomorphisme symétrique. Alors f est diagonalisable, à valeurs propres réelles, et il existe une base orthonormée de E formée de vecteurs propres de f .



Démonstration. (Hors programme)

Prouvons la proposition par récurrence sur $n = \dim E$.

Plus précisément, notons $\mathcal{P}(n)$ la propriété : «pour tout espace euclidien E de dimension inférieure ou égale à n , et pour tout $f \in \mathcal{L}(E)$ symétrique, alors il existe une base orthonormée de E formée de vecteurs propres de f .»

Si $n = 1$, alors f possède une valeur propre réelle λ , donc est diagonalisable. Si x est un vecteur propre de f associé à la valeur propre λ , alors $\frac{x}{\|x\|}$ est une base orthonormée de E , formée de vecteurs propres de f .

Supposons que $\mathcal{P}(n)$ soit vraie, et soit E un espace euclidien de dimension $n + 1$ et $f \in \mathcal{L}(E)$ un endomorphisme symétrique.

f possède une valeur propre λ . Soit alors $F = E_\lambda(f)$.

Si $F = E$, alors $E = \text{Ker}(f - \lambda \text{id}_E)$, de sorte que $f - \lambda \text{id}_E = 0$: $f = \lambda \text{id}_E$ est diagonalisable.

Et alors toute base orthonormée de E est formée⁸ de vecteurs propres de f .

Si $p = \dim F < \dim E$. Alors F est stable par f , et donc $G = F^\perp$ est stable par f .

Soit $g = f|_G$ la restriction de f à G . Puisque G est stable par f , g est un endomorphisme de G . Montrons qu'en fait, g est un endomorphisme symétrique de G .

$$\forall x, y \in G, \langle g(x), y \rangle = \langle f(x), y \rangle = \langle x, f(y) \rangle = \langle x, g(y) \rangle.$$

Remarque
En particulier, si une telle base existe, il existe une base de E formée de vecteurs propres de f et donc f est diagonalisable.

⁸ Puisque dans ce cas, tout vecteur non nul de E est un vecteur propre de f , associé à la valeur propre λ .

Donc g est un endomorphisme symétrique de G . Par hypothèse de récurrence⁹, il existe une base orthonormée de G formée de vecteurs propres de g .
 Ces vecteurs propres de g sont alors des vecteurs propres de f .
 Soit (e_1, \dots, e_p) une base orthonormée de F . Elle est automatiquement formée de vecteurs propres de f car $F = E_\lambda(f)$ ne contient que des vecteurs propres et le vecteur nul (et que ce dernier ne peut faire partie d'une base).
 Soit alors $(e_{p+1}, \dots, e_{n+1})$ une base orthonormée de G formée de vecteurs propres de f .
 Alors $(e_1, \dots, e_p, e_{p+1}, \dots, e_n)$ est une base orthonormée¹⁰ de E , formée de vecteurs propres de f .
 Par conséquent, f est diagonalisable, et donc $\mathcal{P}(n+1)$ est vraie.
 Par le principe de récurrence, la propriété est vraie pour tout entier $n \geq 1$. \square

⁹ Qui s'applique car $\dim G < \dim E$.

¹⁰ Car concaténation d'une base orthonormée de F et d'une base orthonormée de F^\perp .

15.2.3 Diagonalisation des matrices symétriques

Théorème 15.13 : Soit $A \in \mathcal{M}_n(\mathbf{R})$ une matrice symétrique. Alors A est diagonalisable, à valeurs propres réelles, et il existe une matrice orthogonale P et une matrice diagonale D telles que

$$A = P^{-1}DP = {}^tPDP.$$

Démonstration. Soit f l'endomorphisme de \mathbf{R}^n dont la matrice dans la base canonique \mathcal{B} est A .
 Alors f est symétrique¹¹, et donc il existe une base orthonormée \mathcal{B}' de \mathbf{R}^n formée de vecteurs propres de f .
 Si $D = \text{Mat}_{\mathcal{B}'}(f)$ alors D est diagonale et

$$A = P_{\mathcal{B}', \mathcal{B}}^{-1}DP_{\mathcal{B}', \mathcal{B}}.$$

Et puisque $P_{\mathcal{B}', \mathcal{B}}$ est une matrice de passage entre deux bases orthonormées de \mathbf{R}^n , elle est orthogonale et donc $P_{\mathcal{B}', \mathcal{B}}^{-1} = {}^tP_{\mathcal{B}', \mathcal{B}}$. \square

\triangle Ce résultat est spécifique aux matrices symétriques réelles : une matrice symétrique complexe n'est pas nécessairement diagonalisable.
 Pour s'en convaincre, on peut par exemple chercher les valeurs propres et les sous-espaces propres de la matrice symétrique $\begin{pmatrix} 1 & i \\ i & -1 \end{pmatrix}$.

Terminologie
 Ce théorème est parfois appelé «théorème spectral».

¹¹ Car la base canonique de \mathbf{R}^n est orthonormée pour le produit scalaire canonique.

Rédaction
 En particulier, quand on voudra justifier qu'une matrice est diagonalisable car symétrique, on prendra bien soin de dire « A est symétrique, à coefficient réels, donc diagonalisable.»

15.2.4 Diagonalisation pratique d'une matrice symétrique

Proposition 15.14 : Soit $A \in \mathcal{M}_n(\mathbf{R})$ une matrice diagonalisable, soient $\lambda_1, \dots, \lambda_p$ les valeurs propres de A , et pour $i \in \llbracket 1, p \rrbracket$ soit $r_i = \dim E_{\lambda_i}(A)$. Alors

$$\sum_{i=1}^p \lambda_i r_i = \text{tr}(A).$$

Démonstration. La matrice A est semblable à

$$D = \text{diag}(\underbrace{\lambda_1, \dots, \lambda_1}_{r_1 \text{ fois}}, \dots, \underbrace{\lambda_p, \dots, \lambda_p}_{r_p \text{ fois}}).$$

Mais deux matrices semblables ont même trace, donc

$$\text{tr}(A) = \text{tr}(D) = \underbrace{\lambda_1 + \dots + \lambda_1}_{r_1 \text{ fois}} + \dots + \underbrace{\lambda_p + \dots + \lambda_p}_{r_p \text{ fois}} = \sum_{i=1}^p \lambda_i r_i.$$

\square

Remarque. L'intérêt de cette proposition est de trouver «la dernière» valeur propre lorsqu'on connaît déjà les autres.

\triangle **Attention !**
 Cette proposition n'est valable que si on sait déjà que A est diagonalisable. En pratique, on ne pourra l'utiliser que pour les matrices symétriques.

Exemple 15.15

Soit $A = \begin{pmatrix} 1 & -2 & -2 \\ -2 & 1 & -2 \\ -2 & -2 & 1 \end{pmatrix}$. A est symétrique, à coefficients réels, et donc diagonalisable.

On a $\text{rg}(A - 3I_3) = \text{rg} \begin{pmatrix} -2 & -2 & -2 \\ -2 & -2 & -2 \\ -2 & -2 & -2 \end{pmatrix} = 1$, de sorte que 3 est valeur propre de A et

$$\dim E_3(A) = 3 - 1 = 2.$$

Ainsi, A possède au plus une autre valeur propre λ .

Puisque $\text{tr}(A) = 3 = 3 + 3 + \lambda$, on en déduit que $\lambda = -3$ est la dernière valeur propre de A .

Méthode : diagonaliser une matrice symétrique en base orthonormée

Soit à présent $A \in \mathcal{M}_n(\mathbf{R})$ une matrice symétrique. Il existe une matrice diagonale $D \in \mathcal{M}_n(\mathbf{R})$ et une matrice orthogonale P telle que $A = {}^tPDP$. Pour les déterminer, nous procéderons ainsi.

- Déterminer les valeurs propres $\lambda_1, \dots, \lambda_p$ de A , soit à l'aide d'un pivot, soit en utilisant toute indication fournie par l'énoncé.
- Déterminer une base de chacun des sous-espaces propres.
- À l'aide du procédé de Gram-Schmidt, construire une base orthonormée de chacun des sous-espaces propres.
- La concaténation des bases de $E_{\lambda_1}, \dots, E_{\lambda_p}$ ainsi obtenues est une base orthonormée \mathcal{B} de $\mathcal{M}_{n,1}(\mathbf{R})$ car nous savons que les sous-espaces propres sont deux à deux orthogonaux.
- La matrice de passage de la base canonique de $\mathcal{M}_{n,1}(\mathbf{R})$ à la base \mathcal{B} est alors une matrice orthogonale (car matrice de passage entre deux bases orthonormées) : c'est $P^{-1} = {}^tP$. Pour obtenir P , il suffit de transposer cette matrice.
- Comme d'habitude, la matrice D est

$$D = \text{diag}(\underbrace{\lambda_1, \dots, \lambda_1}_{\dim E_{\lambda_1}(A)}, \dots, \underbrace{\lambda_p, \dots, \lambda_p}_{\dim E_{\lambda_p}(A)}).$$

Exemple 15.16

Soit $A = \begin{pmatrix} 1 & 0 & 2 \\ 0 & -1 & 0 \\ 2 & 0 & 1 \end{pmatrix}$. A est symétrique, à coefficients réels, donc diagonalisable.

Notons que $\text{rg}(A + I_3) = \text{rg} \begin{pmatrix} 2 & 0 & 2 \\ 0 & 0 & 0 \\ 2 & 0 & 2 \end{pmatrix} = 1$, et donc -1 est valeur propre de A , avec

$$\dim E_{-1}(A) = 2.$$

Si λ est la dernière valeur propre de A , alors $\text{tr}(A) = 1 = -1 - 1 + \lambda$, et donc $\lambda = 3$, avec $\dim E_3(A) = 1$.

Une résolution de systèmes¹² montre que

$$E_{-1}(A) = \text{Vect} \left(\begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \right) \text{ et } E_3(A) = \text{Vect} \left(\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \right).$$

Alors une base orthonormée¹³ de $E_{-1}(A)$ est : $\frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$.

De même, une base orthonormée de $E_3(A)$ est $\frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$.

¹² Ou l'utilisation de relations sur les colonnes.

¹³ La famille de départ étant déjà orthogonale, il n'est pas vraiment nécessaire de parler de Gram-Schmidt : il suffit de diviser chaque vecteur par sa norme pour obtenir une famille orthonormée.

Par conséquent, on a $A = {}^tPDP$ avec

$$D = \text{diag}(-1, -1, 3) \text{ et } {}^tP = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 0 & 1 \\ 0 & \sqrt{2} & 0 \\ -1 & 0 & 1 \end{pmatrix}, \text{ donc } P = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 0 & -1 \\ 0 & \sqrt{2} & 0 \\ 1 & 0 & 1 \end{pmatrix}$$

15.3 FORME QUADRATIQUE ASSOCIÉE À UNE MATRICE SYMÉTRIQUE

15.3.1 Définition

Définition 15.17 – Soit $A \in \mathcal{M}_n(\mathbf{R})$ une matrice symétrique. On appelle forme quadratique associée à A , et on note q_A l'application de \mathbf{R}^n dans \mathbf{R} définie par

$$\forall (x_1, \dots, x_n) \in \mathbf{R}^n, q_A(x) = (x_1 \quad \dots \quad x_n) A \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}.$$

Exemples 15.18

- Soit $A = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}$ une matrice diagonale. Alors $\forall (x, y) \in \mathbf{R}$,

$$q_A(x, y) = (x \quad y) \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = (x \quad y) \begin{pmatrix} \lambda_1 x \\ \lambda_2 y \end{pmatrix} = \lambda_1 x^2 + \lambda_2 y^2.$$

- Soit $A = \begin{pmatrix} 1 & 3 & -1 \\ 3 & 2 & 0 \\ -1 & 0 & 1 \end{pmatrix}$. Alors

$$\forall (x, y, z) \in \mathbf{R}^3, q_A(x, y, z) = (x \quad y \quad z) \begin{pmatrix} x + 3y - z \\ 3x + 2y \\ -x + z \end{pmatrix} = x^2 + 6xy - 2xz + 2y^2 + z^2.$$

Remarque

De manière générale, on constate que le terme en x_i^2 est le coefficient $a_{i,i}$ et que le terme en $x_i x_j$ est deux fois le coefficient $a_{i,j}$.

Remarques. • De manière générale, lorsqu'on exprime une forme quadratique en fonction de x_1, \dots, x_n , on ne doit obtenir que des termes «de degré 2», c'est-à-dire en x_i^2 et en $x_i x_j$. Il n'y aura jamais de termes en x_i ou de produits de trois variables ou plus.

• Une forme quadratique est ainsi une fonction polynomiale sur \mathbf{R}^n , et en particulier est de classe \mathcal{C}^1 .

⚠ Attention, une forme quadratique n'est pas une application linéaire !

Par exemple, si $\lambda \in \mathbf{R}$, alors $\forall x \in \mathbf{R}^n$, on a

$$q_A(\lambda x) = \lambda (x_1 \quad \dots \quad x_n) A \lambda \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \lambda^2 q_A(x) \neq \lambda q_A(x).$$

15.3.2 Signe d'une forme quadratique

Lemme 15.19. Soit A une matrice symétrique, et soit f l'endomorphisme de \mathbf{R}^n dont la matrice dans la base canonique est A . Alors

$$\forall x \in \mathbf{R}^n, q_A(x) = \langle x, f(x) \rangle.$$

Démonstration. Soit $x \in \mathbf{R}^n$, et soit $X \in \mathcal{M}_{n,1}(\mathbf{R})$ le vecteur colonne de ses coordonnées dans la base canonique.

Le vecteur colonne des coordonnées de $f(x)$ dans la base canonique est $Y = A \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$.

Et alors, la base canonique étant une base orthonormée, on a

$$\langle x, f(x) \rangle = {}^tXY = {}^tXAX = q_A(x).$$

□

Proposition 15.20 : Soit $A \in \mathcal{M}_n(\mathbf{R})$ une matrice symétrique, et soit q_A la forme quadratique associée.

- Si toutes les valeurs propres de A sont positives, alors $\forall x \in \mathbf{R}^n, q_A(x) \geq 0$.
- Si toutes les valeurs propres de A sont négatives, alors $\forall x \in \mathbf{R}^n, q_A(x) \leq 0$.
- Si toutes les valeurs propres de A sont strictement positives, alors $\forall x \in \mathbf{R}^n - \{0\}, q_A(x) > 0$.
- Si toutes les valeurs propres de A sont strictement négatives, alors $\forall x \in \mathbf{R}^n - \{0\}, q_A(x) < 0$.
- Si A possède deux valeurs propres non nulles de signes contraires, alors q_A n'est pas de signe constant.

Démonstration. Soit f l'endomorphisme de \mathbf{R}^n dont la matrice dans la base canonique est A . Alors f est symétrique¹⁴, et donc il existe une base \mathcal{B} de \mathbf{R}^n , orthonormée et formée de vecteurs propres de f .

Notons $\mathcal{B} = (e_1, \dots, e_n)$ et λ_i la valeur propre de f associée à e_i .

Alors si $x \in \mathbf{R}^n$ s'écrit $x = \sum_{i=1}^n \mu_i e_i$, on a

$$q_A(x) = \langle x, f(x) \rangle = \left\langle \sum_{i=1}^n \mu_i e_i, \sum_{j=1}^n \mu_j f(e_j) \right\rangle = \sum_{i=1}^n \sum_{j=1}^n \mu_i \mu_j \langle e_i, \lambda_j e_j \rangle = \sum_{i=1}^n \lambda_i \mu_i^2.$$

En particulier, si toutes les valeurs propres de f sont positives, alors pour tout $x \in \mathbf{R}^n$, de coordonnées (μ_1, \dots, μ_n) dans la base \mathcal{B} , alors

$$q_A(x) = \sum_{i=1}^n \lambda_i \mu_i^2 \geq 0.$$

De même si les λ_i sont tous négatifs, on a $\forall x \in \mathbf{R}^n, q_A(x) \leq 0$.

Si les λ_i sont tous strictement positifs : soit $x \in \mathbf{R}^n$ non nul. Si l'on note μ_1, \dots, μ_n les coordonnées de x dans la base \mathcal{B} , alors l'un au moins des μ_i est non nul. Notons-le μ_{i_0} . Alors

$$q_A(x) = \sum_{i=1}^n \lambda_i \mu_i^2 \geq \lambda_{i_0} \mu_{i_0}^2 > 0.$$

De même, si les λ_i sont tous strictement négatifs, alors $\forall x \neq 0, q_A(x) < 0$.

Enfin, si A possède une valeur propre strictement positive (appelons-la λ_i) et une valeur propre strictement négative (appelons-la λ_j), alors on a

$$q_A(e_i) = \lambda_i > 0 \text{ et } q_A(e_j) = \lambda_j < 0$$

donc q_A n'est pas de signe constant.

□

Réciproque

Notons que nous avons traité ici tous les cas possibles, et que nous pourrions en même temps énoncer la réciproque à ce résultat : la connaissance du signe de q_A permet de retrouver le signe des valeurs propres de A . Par exemple q_A est positive **si et seulement si** toutes les valeurs propres de A sont positives.

¹⁴ Car la base canonique de \mathbf{R}^n est orthonormée pour le produit scalaire canonique.

Détail

Si $i \neq j$, alors $\langle e_i, e_j \rangle = 0$ car \mathcal{B} est orthonormée. Et si $i = j$, alors $\langle e_i, e_j \rangle = 1$.

Méthode

La preuve fournit en fait une méthode pour trouver un x tel que $q(x) > 0$: il «suffit» de trouver un vecteur propre associé à une valeur propre strictement positive. Et de même, si x est un vecteur propre associé à une valeur propre strictement négative, $q(x) < 0$.

MÉTHODES DU CHAPITRE 15

✓ MÉTHODE 15.1 : Montrer qu'un endomorphisme f est symétrique

- En revenant à la définition : $\langle f(x), y \rangle = \langle x, f(y) \rangle$.

Pour cela, on peut partir par exemple de $\langle f(x), y \rangle$ et le transformer jusqu'à obtenir une expression dans laquelle x et y jouent des rôles symétriques.

L'expression obtenue étant valable pour tout x et pour tout y , elle l'est en particulier en échangeant x et y et donc $\langle x, f(y) \rangle = \langle f(x), y \rangle$.

- Si la matrice de f dans une base **orthonormée** est symétrique, alors f est symétrique.

Exercices de TD : 15.5, 15.8, 15.9

✓ MÉTHODE 15.2 : Diagonaliser une matrice symétrique A à l'aide d'une matrice de passage orthogonale

Voir cours.

Exercices de TD : 15.1, 15.2

EXERCICES DU CHAPITRE 15

EXERCICE 15.1 Soit $A = \begin{pmatrix} 3 & 2 & 4 \\ 2 & 0 & 2 \\ 4 & 2 & 3 \end{pmatrix} \in \mathcal{M}_3(\mathbf{R})$.

F

- 1) Montrer que A est diagonalisable.
- 2) Déterminer le rang de $A + I_3$. En déduire que -1 est valeur propre de A et déterminer $\dim E_{-1}(A)$. En déduire les autres valeurs propres de A .
- 3) Déterminer une base des sous-espaces propres de A .
- 4) Déterminer une matrice orthogonale P et une matrice diagonale D telle que $A = {}^tPDP$.

EXERCICE 15.2 Pour chacune des matrices symétriques suivantes, déterminer D diagonale et P orthogonale telles que $A = {}^tPDP$.

F

$$1. A = \begin{pmatrix} 4 & 3 \\ 3 & -4 \end{pmatrix} \quad 2. A = \begin{pmatrix} 1 & 2 & -1 \\ 2 & 4 & -2 \\ -1 & -2 & 1 \end{pmatrix} \quad 3. A = \begin{pmatrix} -1 & 0 & 2 \\ 0 & 1 & 2 \\ 2 & 2 & 0 \end{pmatrix} \quad 4. A = \begin{pmatrix} 0 & 3 & 2 & 3 \\ 3 & 0 & 2 & 3 \\ 2 & 2 & 2 & 2 \\ 3 & 3 & 2 & 0 \end{pmatrix}$$

EXERCICE 15.3

PD

- 1) Soient $A, B \in \mathcal{M}_n(\mathbf{R})$ deux matrices symétriques. Donner une condition nécessaire et suffisante pour que AB soit symétrique.
- 2) Soient f et g deux endomorphismes symétriques d'un espace euclidien E . Déduire de la question précédente une condition nécessaire et suffisante pour que $f \circ g$ soit un endomorphisme symétrique.

EXERCICE 15.4 Soit f un endomorphisme symétrique d'un espace euclidien E tel que $\forall x \in E, \langle x, f(x) \rangle = 0$. Montrer que f est l'endomorphisme nul.

F

EXERCICE 15.5 Soit E un espace euclidien de dimension $n \geq 3$, et soit (a, b) une famille orthonormée de E . Soit $f : E \rightarrow E$ l'application définie par $f(x) = \langle x, a \rangle b + \langle x, b \rangle a$.

PD

- 1) Montrer que f est un endomorphisme symétrique de E .
- 2) Soit $v \in \text{Vect}(a, b)^\perp$. Calculer $f(v)$. En déduire une valeur propre de f .
- 3) Calculer $f(a + b)$ et $f(a - b)$. En déduire les valeurs propres et les sous-espaces propres de f .

EXERCICE 15.6 Soit $A = (a_{i,j}) \in \mathcal{S}_n(\mathbf{R})$, et soient $\lambda_1, \dots, \lambda_p$ les valeurs propres de A . Montrer que

PD

$$\sum_{i=1}^n \sum_{j=1}^n a_{i,j}^2 = \sum_{i=1}^p \lambda_i^2 \times \dim E_{\lambda_i}(A).$$

EXERCICE 15.7 Soit $A \in \mathcal{M}_n(\mathbf{R})$, et soit $B = {}^tAA$.

AD

- 1) Montrer que B est symétrique, et que toutes ses valeurs propres sont positives.
- 2) Prouver que pour $X \in \mathcal{M}_{n,1}(\mathbf{R})$, on a $BX = 0 \Leftrightarrow AX = 0$. En déduire que $\text{rg}(A) = \text{rg}(B)$.

EXERCICE 15.8 (D'après EML 2013)

AD

Soit $X \in \mathcal{M}_{n,1}(\mathbf{R})$ avec $\|X\| = 1$, et soit $S = X^tX$. On munit $\mathcal{M}_n(\mathbf{R})$ du produit scalaire $\langle M, N \rangle = \text{tr}({}^tMN)$.

- 1) Montrer que S est symétrique et vérifie $S^2 = S$.
- 2) Soit $\Phi : \mathcal{M}_n(\mathbf{R}) \rightarrow \mathcal{M}_n(\mathbf{R})$ défini par $\Phi(M) = SM$. Vérifier que Φ est un endomorphisme symétrique de $\mathcal{M}_n(\mathbf{R})$.
- 3) Montrer que $\Phi^2 = \Phi$. En déduire les valeurs propres de Φ .
- 4) Montrer que $\text{Ker}(\Phi)$ et $\text{Ker}(\Phi - \text{id}_{\mathcal{M}_n(\mathbf{R})})$ sont supplémentaires orthogonaux dans $\mathcal{M}_n(\mathbf{R})$.

EXERCICE 15.9 Polynômes de Tchebychev de seconde espèce (D'après EML 2005)

AD

Soit $f : \mathbf{R}_n[X] \rightarrow \mathbf{R}_n[X]$ défini par $f(P) = (X^2 - 1)P'' + 3XP'$.

- 1) Montrer que $(P|Q) = \int_{-1}^1 \sqrt{1-t^2} P(t)Q(t) dt$ est un produit scalaire sur $\mathbf{R}_n[X]$.
- 2) a) Montrer que f est un endomorphisme de $\mathbf{R}_n[X]$, et qu'il est symétrique pour le produit scalaire $(\cdot|\cdot)$.

Indication : on pourra commencer par montrer que $(f(P)|Q) = \int_{-1}^1 (1-t^2)^{\frac{3}{2}} P'(t)Q'(t) dt$.

b) Déterminer les valeurs propres de f .

EXERCICE 15.10 Soit $A \in \mathcal{M}_n(\mathbf{R})$ une matrice symétrique dont toutes les valeurs propres sont positives. Montrer qu'il existe une matrice $B \in \mathcal{M}_n(\mathbf{R})$, symétrique, à valeurs propres positives, telle que $B^2 = A$.

PD

EXERCICE 15.11 Racine carrée d'un endomorphisme symétrique défini positif

Soit E un espace euclidien. Un endomorphisme symétrique $f \in \mathcal{L}(E)$ est dit défini positif si pour tout $x \in E \setminus \{0\}$, $\langle f(x), x \rangle > 0$.

AD

- 1) Soit $f \in \mathcal{L}(E)$ tel qu'il existe $g \in \mathcal{L}(E)$, symétrique tel que $g^2 = f$. Montrer que f est symétrique et que $\forall x \in E, \langle f(x), x \rangle \geq 0$.
- 2) Montrer qu'un endomorphisme symétrique est défini positif si et seulement si $\text{Spec}(f) \subset \mathbf{R}_+^*$.
- 3) En déduire que si f est symétrique défini positif, alors f est inversible et f^{-1} est encore symétrique défini positif.
- 4) Soit f un endomorphisme symétrique défini positif de E . Montrer qu'il existe $g \in \mathcal{L}(E)$, symétrique, défini positif, tel que $g^2 = f$. *Indication* : on pourra s'intéresser à la matrice de f dans une base orthonormée de E formée de vecteurs propres de f .

EXERCICE 15.12 Soit $A \in \mathcal{M}_n(\mathbf{R})$ une matrice symétrique. On suppose qu'il existe $k \in \mathbf{N}^*$ tel que $A^k = I_n$. Montrer que $A^2 = I_n$.

AD

EXERCICE 15.13 Quotients de Rayleigh

Soit E un espace euclidien et f un endomorphisme symétrique de E . Soit λ la plus petite valeur propre de f et μ la plus grande valeur propre.

AD

- 1) Montrer que pour tout $x \in E - \{0\}$, on a $\lambda \leq \frac{\langle f(x), x \rangle}{\|x\|^2} \leq \mu$.
- 2) En déduire que $\lambda = \min_{x \neq 0} \frac{\langle f(x), x \rangle}{\|x\|^2}$ et que $\mu = \max_{x \neq 0} \frac{\langle x, f(x) \rangle}{\|x\|^2}$.

EXERCICE 15.14 Soit $A = \begin{pmatrix} 2 & 2 & -2 \\ 2 & 5 & -4 \\ -2 & -4 & 5 \end{pmatrix}$. Calculer $A \begin{pmatrix} 2 \\ -1 \\ 0 \end{pmatrix}$, puis en déduire avec le moins de calculs possible les valeurs propres de A . Quel est le signe de la forme quadratique associée à A ?

PD

EXERCICE 15.15 Soit $A = \begin{pmatrix} 2 & 0 & -36 \\ 0 & -25 & 0 \\ -36 & 0 & 23 \end{pmatrix}$, et soit q la forme quadratique associée à A .

PD

Déterminer les valeurs propres de A , puis déterminer un vecteur x tel que $q(x) > 0$ et un vecteur y tel que $q(y) < 0$.

EXERCICE 15.16 (QSP ESCP 2015)

Soit A une matrice de $\mathcal{M}_n(\mathbf{R})$ telle que ${}^tAA = A^tA$.

On suppose qu'il existe $p \in \mathbf{N}$ tel que $A^p = 0$. Montrer que $A = 0$.

D

EXERCICE 15.17 (QSP HEC 2009)

Soient f et g deux endomorphismes symétriques d'un espace euclidien E , dont toutes les valeurs propres sont positives ou nulles.

TD

- 1) Montrer qu'il existe un endomorphisme symétrique ϕ tel que $f = \phi^2$.
- 2) Montrer que $\text{Ker}(f + g) = \text{Ker}(f) \cap \text{Ker}(g)$.

EXERCICE 15.18 (Oral ESCP 2013)

Soit $n \geq 2$. On note \mathcal{S}_n l'ensemble des matrices symétriques de $\mathcal{M}_n(\mathbf{R})$. On note \mathcal{S}_n^+ l'ensemble des matrices de \mathcal{S}_n dont toutes les valeurs propres sont réelles et positives.

TD

- 1) a) Soit $S \in \mathcal{S}_n$. Montrer que $S \in \mathcal{S}_n^+ \Leftrightarrow \forall X \in \mathcal{M}_{n,1}(\mathbf{R}), {}^tXSX \geq 0$.
 b) Soit $A \in \mathcal{M}_n(\mathbf{R})$. Montrer que $S = {}^tAA \in \mathcal{S}_n^+$.
 c) Réciproquement, montrer que pour toute matrice $S \in \mathcal{S}_n^+$, il existe $A \in \mathcal{M}_n(\mathbf{R})$ telle que $S = {}^tAA$.
- 2) Soient U et V deux matrices de $\mathcal{M}_n(\mathbf{R})$.
 a) Montrer que si 0 est valeur propre de UV , alors 0 est aussi valeur propre de VU .
 b) Montrer que les matrices UV et VU ont les mêmes valeurs propres complexes.
- 3) a) Soient S et T deux matrices de \mathcal{S}_n^+ . Montrer que $S + T \in \mathcal{S}_n^+$.
 b) \mathcal{S}_n^+ est-il un sous-espace vectoriel de \mathcal{S}_n ?
- 4) Soient S et T deux matrices de \mathcal{S}_n .
 a) À quelle condition nécessaire et suffisante sur S et T a-t-on ST symétrique ?
 b) On suppose que S et T appartiennent à \mathcal{S}_n^+ . En utilisant les questions 1.c et 2.b, montrer que toutes les valeurs propres de la matrice ST sont réelles et positives.

CORRECTION DES EXERCICES DU CHAPITRE 15

SOLUTION DE L'EXERCICE 15.1

1. A est symétrique à coefficients réels, donc diagonalisable.
2. Les trois lignes de $A + I_3$ sont colinéaires, donc $\text{rg}(A + I_3) = 1$, et sorte que -1 est valeur propre de A et $\dim E_{-1}(A) = 2$.
3. La trace de A est égale à 6, mais si λ est la dernière valeur propre de A , alors $\text{tr}(A) = -1 \times 2 + \lambda$. Donc $\lambda = 8$.
Ainsi, $\text{Spec}(A) = \{-1, 8\}$.
4. En remarquant que les colonnes de $A + I_3$ sont liées par les relations $C_1 - 2C_2 = 0$ et $-2C_2 + C_3 = 0$, et que celles de $A - 8I_3$ sont liées par la relation $2C_1 + C_2 + 2C_3 = 0$, on obtient

$$E_{-1}(A) = \text{Vect} \left(\begin{pmatrix} 1 \\ -2 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ -2 \\ 1 \end{pmatrix} \right) \text{ et } E_8(A) = \text{Vect} \begin{pmatrix} 2 \\ 1 \\ 2 \end{pmatrix}$$

5. Le procédé d'orthonormalisation de Gram-Schmidt donne comme base orthonormée de $E_{-1}(A)$ la famille $\frac{1}{\sqrt{5}} \begin{pmatrix} 1 \\ -2 \\ 0 \end{pmatrix}, \frac{1}{3\sqrt{5}} \begin{pmatrix} -4 \\ -2 \\ 5 \end{pmatrix}$, et comme base orthonormée de $E_8(A)$ la famille

$$\frac{1}{3} \begin{pmatrix} 2 \\ 1 \\ 2 \end{pmatrix}.$$

$$\text{Alors en posant } {}^tP = \begin{pmatrix} \frac{1}{\sqrt{5}} & -\frac{4}{3\sqrt{5}} & \frac{2}{3} \\ -\frac{2}{\sqrt{5}} & -\frac{2}{3\sqrt{5}} & \frac{1}{3} \\ 0 & \frac{5}{3\sqrt{5}} & \frac{2}{3} \end{pmatrix} \text{ et } D = \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 8 \end{pmatrix}, \text{ on a } A = {}^tPDP.$$

SOLUTION DE L'EXERCICE 15.2

1. $\text{Spec}(A) = \{-5, 5\}$, et on peut prendre $P = \frac{1}{\sqrt{10}} \begin{pmatrix} 1 & -3 \\ -3 & -1 \end{pmatrix}, D = \begin{pmatrix} -5 & 0 \\ 0 & 5 \end{pmatrix}$.
2. $\text{Spec}(A) = \{0, 6\}$, avec $\dim E_0(A) = 2$.
On peut prendre

$$P = \begin{pmatrix} \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{6}} & \frac{2}{\sqrt{6}} & -\frac{1}{\sqrt{6}} \end{pmatrix}, D = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 6 \end{pmatrix}$$

3. $\text{Spec}(A) = \{-3, 0, 3\}$. On peut prendre

$$P = \frac{1}{3} \begin{pmatrix} 2 & 1 & -2 \\ 2 & -2 & 1 \\ 1 & 2 & 2 \end{pmatrix} \text{ et } D = \begin{pmatrix} -3 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 3 \end{pmatrix}.$$

4. $\text{Spec}(A) = \{-3, 0, 8\}$ avec $\dim E_{-3}(A) = 2$. On peut alors prendre

$$P = \begin{pmatrix} \frac{1}{2\sqrt{3}} & \frac{1}{2\sqrt{3}} & \frac{-\sqrt{3}}{2} & \frac{1}{2\sqrt{3}} \\ \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} & 0 & \frac{-2}{\sqrt{6}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 & 0 \\ \frac{1}{2} & \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \end{pmatrix}, D = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & -3 & 0 & 0 \\ 0 & 0 & -3 & 0 \\ 0 & 0 & 0 & 8 \end{pmatrix}.$$

SOLUTION DE L'EXERCICE 15.3

1. Si A et B sont symétriques, alors ${}^t(AB) = {}^tB{}^tA = BA$.
Et donc AB est symétrique si et seulement si ${}^t(AB) = AB \Leftrightarrow BA = AB$.
Autrement dit, le produit AB est symétrique si et seulement si A et B commutent.
2. Soit \mathcal{B} une base orthonormée de E . Alors $\text{Mat}_{\mathcal{B}}(f)$ et $\text{Mat}_{\mathcal{B}}(g)$ sont des matrices symétriques.

Rédaction

Une matrice symétrique à coefficients complexes n'est pas toujours diagonalisable, donc on prendra bien soin de préciser symétrique à coefficients réels.

⚠ Attention !

Comme d'habitude sur les exercices de diagonalisation, en partant d'une autre base de vecteurs propres, on a une autre matrice de passage ! Il n'y a pas unicité de P (ni même de D car on peut choisir d'écrire les valeurs propres dans un autre ordre).

Rappel

Un endomorphisme est symétrique si et seulement si sa matrice dans une base orthonormée est symétrique.

Donc $f \circ g$ est symétrique si et seulement si $\text{Mat}_{\mathcal{B}}(f \circ g) = \text{Mat}_{\mathcal{B}}(f) \times \text{Mat}_{\mathcal{B}}(g)$ est symétrique. Et d'après la question précédente, c'est le cas si et seulement si ces deux matrices commutent. Donc si et seulement si f et g commutent.

Notons que ce résultat peut également se retrouver sans passer par les matrices, mais c'est alors plus fastidieux.

En effet, pour $(x, y) \in E$, la symétrie de f et g nous permet d'obtenir

$$\langle (f \circ g)(x), y \rangle = \langle f(g(x)), y \rangle = \langle g(x), f(y) \rangle = \langle x, g(f(y)) \rangle = \langle x, (g \circ f)(y) \rangle.$$

Donc il est clair que si $f \circ g = g \circ f$, ceci prouve la symétrie de $f \circ g$.

Inversement, si $f \circ g$ est symétrique, alors le calcul que nous venons de faire prouve que pour tous $x, y \in E$,

$$\langle x, (f \circ g)(y) \rangle = \langle x, (g \circ f)(y) \rangle \Leftrightarrow \langle x, (f \circ g - g \circ f)(y) \rangle = 0.$$

En particulier, pour $y \in E$ et pour $x = (f \circ g - g \circ f)(y)$, alors $\langle (f \circ g - g \circ f)(y), (f \circ g - g \circ f)(y) \rangle = 0$.

Soit encore $\|(f \circ g - g \circ f)(y)\|^2 = 0 \Leftrightarrow (f \circ g - g \circ f)(y) = 0$.

Ceci étant vrai pour tout $y \in E$, on en déduit que $f \circ g = g \circ f$.

SOLUTION DE L'EXERCICE 15.4

f est diagonalisable car symétrique. Il suffit donc de prouver que la seule valeur propre de f est 0.

Soit donc $\lambda \in \text{Spec}(f)$ et x un vecteur propre associé. On a alors

$$0 = \langle f(x), x \rangle = \langle \lambda x, x \rangle = \lambda \langle x, x \rangle = \lambda \|x\|^2.$$

Mais $x \neq 0$ par hypothèse¹ et donc $\lambda = 0$.

Ainsi, la seule valeur propre de f est 0 et donc $f = 0$.

SOLUTION DE L'EXERCICE 15.5

1. La linéarité de f découle de la linéarité à gauche du produit scalaire et ne présente par de difficulté particulière.

De plus, pour $x \in E$, $\langle x, a \rangle$ et $\langle x, b \rangle \in \mathbf{R}$ et donc on a bien $f(x) \in E$, de sorte que f est un endomorphisme de E .

Soient alors $x, y \in E$. Alors

$$\langle f(x), y \rangle = \langle \langle x, a \rangle b + \langle x, b \rangle a, y \rangle = \langle x, a \rangle \langle y, b \rangle + \langle x, b \rangle \langle y, a \rangle.$$

Et de même,

$$\langle x, f(y) \rangle = \langle y, a \rangle \langle x, b \rangle + \langle y, b \rangle \langle x, a \rangle = \langle f(x), y \rangle.$$

Et donc f est symétrique.

2. Si $v \in \text{Vect}(a, b)^\perp$, alors $\langle x, a \rangle = \langle x, b \rangle = 0$, et donc $f(v) = 0$.

Autrement dit, $\text{Vect}(a, b)^\perp \subset \text{Ker } f$.

Mais puisque $\dim E \geq 3$, alors $\dim \text{Vect}(a, b)^\perp = \dim E - 2 \geq 1$, et donc $\text{Ker } f \neq \{0\}$.

On en déduit que 0 est valeur propre de f , avec un sous-espace propre de dimension supérieure ou égale à $\dim E - 2$.

3. On a

$$f(a + b) = f(a) + f(b) = \underbrace{\|a\|^2}_{=1} b + \underbrace{\langle a, b \rangle}_{=0} a + \underbrace{\|b\|^2}_{=1} a + \langle a, b \rangle b = a + b.$$

Puisque $a + b \neq 0$, on en déduit que 1 est une valeur propre de f .

De même, on a

$$f(a - b) = f(a) - f(b) = \|a\|^2 b + \langle a, b \rangle a - \|b\|^2 a - \langle a, b \rangle b = -(a - b).$$

Et donc $a - b$ est un vecteur propre de f pour la valeur propre -1 .

On a donc $\dim E_0(f) + \dim E_1(f) + \dim E_{-1}(f) \geq \dim E$.

Puisque l'inégalité contraire est toujours vraie, on en déduit que $\dim E_0(f) + \dim E_1(f) + \dim E_{-1}(f) = \dim E$.

Donc f ne possède pas d'autres valeurs propres que 0, 1 et -1 , et $\dim E_0(f) = \dim E - 2$, $\dim E_1(f) = \dim E_{-1}(f) = 1$.

Et grâce à ce qui a été dit précédemment², on a donc

$$E_0(f) = \text{Vect}(a, b)^\perp, E_1(f) = \text{Vect}(a + b), E_{-1}(f) = \text{Vect}(a - b).$$

Détails

En effet, la matrice de f dans une base de vecteurs propres sera alors nulle, et donc $f = 0$.

¹ Un vecteur propre est toujours non nul.

² On a déjà des inclusions :

$$\text{Vect}(a, b)^\perp \subset E_0(f)$$

$$\text{Vect}(a + b) \subset E_1(f)$$

$$\text{Vect}(a - b) \subset E_{-1}(f).$$

SOLUTION DE L'EXERCICE 15.6

A est symétrique réelle, donc diagonalisable. Elle est alors semblable à

$$\text{Diag}(\underbrace{\lambda_1, \dots, \lambda_1}_{\dim E_{\lambda_1}(A) \text{ fois}}, \dots, \underbrace{\lambda_p, \dots, \lambda_p}_{\dim E_{\lambda_p}(A) \text{ fois}}).$$

Alors on a $A^2 = {}^tAA$ qui est semblable à $D^2 = \text{Diag}(\underbrace{\lambda_1^2, \dots, \lambda_1^2}_{\dim E_{\lambda_1}(A) \text{ fois}}, \dots, \underbrace{\lambda_p^2, \dots, \lambda_p^2}_{\dim E_{\lambda_p}(A) \text{ fois}})$.

Mais deux matrices semblables ont même trace, donc

$$\text{tr}({}^tAA) = \sum_{i=1}^p \lambda_i^2 \dim E_{\lambda_i}(A).$$

D'autre part, on peut prouver que $\text{tr}({}^tAA) = \sum_{i=1}^n \sum_{j=1}^n a_{i,j}^2$.

Classique

Nous avons en fait déjà rencontré plusieurs fois ce résultat lorsque nous avons prouvé que

$$(A, B) \mapsto \text{tr}({}^tAB)$$

est un produit scalaire sur $\mathcal{M}_n(\mathbf{R})$. Nous ne le reproduisons donc pas ici mais renvoyons par exemple à l'exercice 3 du TD 8 pour une preuve.

SOLUTION DE L'EXERCICE 15.7

1. B est symétrique car ${}^tB = {}^t({}^tAA) = {}^tA({}^tA) = {}^tAA = B$.
Soit à présent λ une valeur propre de tAA et soit $X \in \mathcal{M}_{n,1}(\mathbf{R})$ un vecteur propre associé. On a alors ${}^tAAX = \lambda X$ et donc ${}^tX{}^tAAX = {}^tX\lambda X = \lambda {}^tXX = \lambda \|X\|^2$.
Mais d'autre part, ${}^tX{}^tAAX = {}^t(AX)(AX) = \|AX\|^2 \geq 0$.
Puisque X est un vecteur propre, il est non nul, et donc $\|X\|^2 > 0$, de sorte que $\lambda = \frac{\|AX\|^2}{\|X\|^2} \geq 0$.
Nous avons donc bien prouvé que toutes les valeurs propres de B sont positives.
2. Si $AX = 0$, alors évidemment $BX = {}^tAAX = 0$.
Inversement, si $BX = 0$, alors ${}^tAAX = 0$, et donc ${}^tX{}^tAAX = 0$.
Mais ${}^tX{}^tAAX = {}^t(AX)(AX) = \|AX\|^2$.
Et donc $\|AX\| = 0 \Leftrightarrow AX = 0$.

Astuce

Multiplier à gauche par tX permet de faire apparaître une norme (pour le produit scalaire canonique de $\mathcal{M}_{n,n,1}(\mathbf{R})$). Ce type de questions est assez classique, et l'astuce n'est pas toujours indiquée par l'énoncé !

En d'autres termes, nous venons de prouver que

$$E_0(A) = \{X \in \mathcal{M}_{n,1}(\mathbf{R}) : AX = 0\} = \{X \in \mathcal{M}_{n,1}(\mathbf{R}) : BX = 0\} = E_0(B).$$

Et donc

$$\text{rg}(A) = n - \dim E_0(A) = n - \dim E_0(B) = \text{rg}(B).$$

SOLUTION DE L'EXERCICE 15.8

1. On a ${}^tS = {}^t(X{}^tX) = X{}^tX = S$. Donc S est symétrique.
De plus, on a
$$S^2 = (X{}^tX)X{}^tX = X \underbrace{({}^tXX)}_{=1} {}^tX = X{}^tX = S.$$
2. Il est facile de vérifier que Φ est linéaire, et donc est un endomorphisme de $\mathcal{M}_n(\mathbf{R})$.
De plus, pour $U, V \in \mathcal{M}_n(\mathbf{R})$, on a
$$\langle U, \Phi(V) \rangle = \text{tr}({}^tUX{}^tXV) = \text{tr}({}^t({}^tXU)XV) = \langle {}^tXU, {}^tXV \rangle.$$

De même, on a $\langle \Phi(U), V \rangle = \langle V, \Phi(U) \rangle = \langle {}^tXV, {}^tXU \rangle = \langle {}^tXU, {}^tXV \rangle = \langle U, \Phi(V) \rangle$.
Et donc Φ est un endomorphisme symétrique de $\mathcal{M}_n(\mathbf{R})$.
3. Pour $M \in \mathcal{M}_n(\mathbf{R})$, on a $\Phi^2(M) = \Phi(SM) = S^2M = SM = \Phi(M)$.
Et donc $\Phi^2 = \Phi$.
On en déduit que Φ est un projecteur et donc que $\text{Spec}(\Phi) \subset \{0, 1\}$.
De plus, Φ n'est pas l'endomorphisme nul, car $S \neq 0$ et $\Phi(S) = S$, et Φ n'est pas l'endomorphisme identité car si $Y \in \mathcal{M}_{n,1}(\mathbf{R})$ est orthogonal³ à X, alors $\Phi(Y) = X{}^tXY = X\langle X, Y \rangle = 0$.
Et donc $\text{Spec}(\Phi) = \{0, 1\}$.
4. Nous venons de prouver que Φ est un projecteur, et qu'il est symétrique : c'est donc un projecteur orthogonal.
Et alors $\text{Ker}(\Phi) = \text{Im}(\Phi)^\perp$.
Mais comme pour tout projecteur, $\text{Im}(\Phi) = E_1(\Phi) = \text{Ker}(\Phi - \text{id}_{\mathcal{M}_n(\mathbf{R})})$.
Et donc $\text{Ker}(\Phi) = \text{Ker}(\Phi - \text{id}_{\mathcal{M}_n(\mathbf{R})})^\perp$.

De même

Ici, quand on dit «de même», il ne s'agit pas de dire que le même type de calcul, que nous ne faisons pas, nous mènerait au même résultat. Mais plutôt de remarquer que la formule prouvée précédemment est valable pour tout U et pour tout V. Et donc en particulier si on échange le rôle de U et V !

Rappel

Tout projecteur distinct de 0 et de id possède 0 et 1 comme valeurs propres.

³ Pour le produit scalaire canonique de $\mathcal{M}_{n,1}(\mathbf{R})$.

SOLUTION DE L'EXERCICE 15.9

1. Notons que l'intégrale définissant $(P|Q)$ est une intégrale sur un segment, donc convergente. Il est évident que $(P|Q) = (Q|P)$, et la linéarité à gauche n'est pas très difficile par linéarité de l'intégrale.

De plus, la fonction $t \mapsto P(t)^2\sqrt{1-t^2}$ est positive et donc par positivité de l'intégrale, $(P|P) \geq 0$.

Enfin, $t \mapsto P(t)^2\sqrt{1-t^2}$ étant continue et positive sur $[-1, 1]$, $\int_{-1}^1 P(t)^2 e^{-t} dt = 0$ si et seulement si pour tout $t \in [-1, 1]$, $P(t)^2\sqrt{1-t^2} = 0$.

Soit si et seulement si pour tout $t \in]-1, 1[, P(t) = 0$. Mais un polynôme possédant une infinité de racines est nécessairement le polynôme nul, et donc $(P|P) = 0 \Leftrightarrow P = 0$.

Donc $(\cdot|\cdot)$ est un produit scalaire sur $\mathbf{R}_n[X]$.

- 2.a. D'après l'énoncé, f est à valeurs dans $\mathbf{R}_n[X]$, ce que l'on pourrait vérifier aisément. De plus, f est linéaire car si $P, Q \in \mathbf{R}_n[X]$ et si $\lambda \in \mathbf{R}$, alors

$$f(\lambda P + Q) = (X^2 - 1)(\lambda P + Q)'' + 3X(\lambda P + Q)' = \lambda((X^2 - 1)P'' + 3XP') + (X^2 - 1)Q'' + 3XQ' = \lambda f(P) + f(Q).$$

De plus, on a

$$(f(P)|Q) = \int_{-1}^1 \sqrt{1-t^2} \left((t^2 - 1)P''(t) + 3tP'(t) \right) Q(t) dt = \int_{-1}^1 \left(3t\sqrt{1-t^2}P'(t) - (1-t^2)^{3/2}P''(t) \right) Q(t) dt.$$

Partons du résultat indiqué dans l'énoncé, et procédons à une intégration par parties sur $[-1, 1]$, en posant $u(t) = (1-t^2)^{3/2}P'(t)$ et $v(t) = Q(t)$, qui sont deux fonctions de classe \mathcal{C}^1 sur $[-1, 1]$, avec $u'(t) = -3t\sqrt{1-t^2}P'(t) + (1-t^2)^{3/2}P''(t)$ et $v'(t) = Q'(t)$.

Alors

$$\begin{aligned} & \int_{-1}^1 (1-t^2)^{3/2}P'(t)Q'(t) dt \\ &= \underbrace{\left[(1-t^2)^{3/2}P'(t)Q(t) \right]_{-1}^1}_{=0} + \int_{-1}^1 \left(3t\sqrt{1-t^2}P'(t) - (1-t^2)^{3/2}P''(t) \right) Q(t) dt \\ &= (f(P)|Q). \end{aligned}$$

Ceci étant valable pour tout P et pour tout Q , on a alors

$$(P|f(Q)) = (f(Q)|P) = \int_{-1}^1 (1-t^2)Q'(t)P'(t) dt = (f(P)|Q).$$

Et donc f est un endomorphisme symétrique.

- 2.b. On a $f(1) = 0$, $f(X) = 3X$ et pour $k \in \llbracket 2, n \rrbracket$,

$$f(X^k) = (X^2 - 1)k(k-1)X^{k-2} + 3kX^k = (k^2 + 2k)X^k - k(k-1)X^{k-2}.$$

Ainsi, la matrice de f dans la base canonique de $\mathbf{R}_n[X]$ est

$$\text{Mat}_{\mathcal{B}_{can}}(f) = \begin{pmatrix} f(1) & f(X) & f(X^2) & f(X^3) & \dots & \dots & f(X^{n-1}) & f(X^n) \\ 0 & 0 & -2 & 0 & \dots & \dots & \dots & 0 \\ 0 & 3 & 0 & -6 & \ddots & & & \vdots \\ 0 & 0 & 8 & 0 & \ddots & \ddots & & \vdots \\ \vdots & & \ddots & 15 & \ddots & \ddots & \ddots & \vdots \\ \vdots & & & \ddots & \ddots & \ddots & \ddots & 0 \\ \vdots & & & & \ddots & \ddots & \ddots & -n(n-1) \\ \vdots & & & & & \ddots & n^2 - 1 & 0 \\ 0 & \dots & \dots & \dots & \dots & \dots & 0 & n^2 + 2 \end{pmatrix} \begin{matrix} 1 \\ X \\ X^2 \\ X^3 \\ \vdots \\ X^{n-2} \\ X^{n-1} \\ X^n \end{matrix}.$$

Cette matrice est triangulaire supérieure, donc ses valeurs propres sont ses coefficients diagonaux :

$$\text{Spec}(f) = \{k^2 + 2k, k \in \llbracket 0, n \rrbracket\}.$$

Rédaction 

Pour montrer qu'un endomorphisme est symétrique, on procède en général comme on vient de le faire : on transforme l'expression de $\langle f(x), y \rangle$ jusqu'à obtenir une expression où x et y jouent des rôles symétriques. On pourrait ensuite refaire le même type de calcul pour $\langle x, f(y) \rangle$, mais on peut s'en passer en remarquant que le premier calcul est valable si on échange les rôles de x et y .

SOLUTION DE L'EXERCICE 15.10

Puisque A est symétrique, elle est semblable à une matrice D diagonale. Plus précisément : il existe une matrice P orthogonale et une matrice D diagonale telles que $A = P^{-1}DP = {}^tPDP$.

Notons que si $D = \text{Diag}(\lambda_1, \dots, \lambda_n)$, alors les λ_i sont les valeurs propres de A , et donc par hypothèse sont positifs.

Soit alors $C = \text{Diag}(\sqrt{\lambda_1}, \dots, \sqrt{\lambda_n})$. Il est évident que $C^2 = D$.

Posons alors $B = {}^tPCP = P^{-1}CP$, de sorte que

$$B^2 = {}^tPC \underbrace{P^tP}_{=I_n} CP = {}^tPC^2P = {}^tPDP = A.$$

De plus, ${}^tB = {}^t({}^tPCP) = {}^tP^tCP = {}^tPCP = B$, donc B est symétrique.

Enfin, puisque B est semblable à C , $\text{Spec}(B) = \text{Spec}(C)$. Or les valeurs propres de C sont les $\sqrt{\lambda_i} \geq 0$, et donc toutes les valeurs propres de B sont positives.

SOLUTION DE L'EXERCICE 15.11

- Supposons donc qu'il existe g symétrique tel que $f = g^2$. Alors, pour tous $x, y \in E$, on a

$$\langle f(x), y \rangle = \langle g^2(x), y \rangle = \langle g(x), g(y) \rangle = \langle x, g^2(y) \rangle = \langle x, f(y) \rangle.$$

Et donc f est symétrique.

D'autre part, pour $x \in E, x \neq 0$, on a

$$\langle f(x), x \rangle = \langle g^2(x), x \rangle = \langle g(x), g(x) \rangle = \|g(x)\|^2.$$

Donc déjà, $\langle f(x), x \rangle \geq 0$, il suffit de voir que ce produit scalaire est non nul. Mais c'est le cas si et seulement si $g(x) = 0_E$. Mais dans ce cas, on aurait alors $\langle g(x), x \rangle = \langle 0_E, x \rangle = 0$, contredisant le fait que g est défini positif.

- Supposons que toutes les valeurs propres de f soient strictement positives, et soit (e_1, \dots, e_n) une base orthonormée de E formée de vecteurs propres de f , et soit λ_i la valeur propre associée à e_i .

Alors, tout vecteur non nul x de E s'écrit de manière unique $x = \sum_{i=1}^n x_i e_i$, où les x_i ne sont pas tous nuls⁴. Et alors

$$\langle f(x), x \rangle = \left\langle \sum_{i=1}^n x_i e_i, f \left(\sum_{j=1}^n x_j e_j \right) \right\rangle = \left\langle \sum_{i=1}^n x_i e_i, \sum_{j=1}^n x_j \lambda_j e_j \right\rangle = \sum_{i=1}^n \sum_{j=1}^n x_i x_j \langle e_i, e_j \rangle = \sum_{i=1}^n x_i^2 \lambda_i.$$

Puisque les $x_i^2 \lambda_i$ sont positifs, $\langle x, f(x) \rangle \geq 0$, et n'étant pas tous nuls, $\langle x, f(x) \rangle > 0$, de sorte que f est défini positif.

Inversement, si f est défini positif, soit λ une valeur propre de f et x un vecteur propre associé. Alors

$$\langle x, f(x) \rangle = \langle x, \lambda x \rangle = \lambda \|x\|^2.$$

Mais par hypothèse, $\langle x, f(x) \rangle > 0$, et x étant non nul⁵, $\|x\| > 0$, de sorte que $\lambda = \frac{\langle x, f(x) \rangle}{\|x\|^2} > 0$.

Donc toutes les valeurs propres de f sont strictement positives.

- Si f est défini positif, alors par la question précédente, ses valeurs propres sont toutes non nulles, et donc f est inversible.

Puisque f est diagonalisable, sa matrice A dans une base orthonormée \mathcal{B} est semblable à une matrice diagonale $D = \text{Diag}(\lambda_1, \dots, \lambda_n)$: il existe P orthogonale telle que $A = {}^tPDP$. Et alors⁶ $A^{-1} = {}^tP \text{Diag}(\lambda_1^{-1}, \dots, \lambda_n^{-1}) P$ ne possède que des valeurs propres strictement positives.

Mais puisque $A^{-1} = \text{Mat}_{\mathcal{B}}(f^{-1})$, les valeurs propres de f^{-1} sont toutes strictement positives.

Reste à prouver que f est symétrique. Mais A étant symétrique, A^{-1} l'est également, et donc⁷, f^{-1} est symétrique.

Rappel

Une matrice orthogonale est une matrice inversible telle que $P^{-1} = {}^tP$.

▶ Dans la suite on écrira donc indifféremment P^{-1} ou tP , puisqu'il s'agit de la même matrice.

Astuce

▶ Ce type de raisonnement est assez classique.

On retiendra que pour une matrice diagonale D à coefficients positifs, il est facile de trouver une «racine carrée de D » : on peut prendre une matrice diagonale dont les coefficients diagonaux sont les racines de ceux de D .

Et pour une matrice qui n'est plus diagonale, mais diagonalisable, on utilisera le fait qu'elle est semblable à une matrice diagonale, dont on sait trouver une «racine carrée».

⁴ Car $x \neq 0_E$.

⁵ C'est un vecteur propre !

⁶ Rappelons que P étant orthogonale, $P^{-1} = {}^tP$.

⁷ La base \mathcal{B} est orthonormée.

4. Puisque f est symétrique, il existe une base orthonormée \mathcal{B} de E formée de vecteurs propres de f .

Et alors la matrice de f dans cette base est diagonale : $A = \text{Diag}(\lambda_1, \dots, \lambda_n)$.

On a alors $A = (\text{Diag}(\sqrt{\lambda_1}, \dots, \sqrt{\lambda_n}))^2$.

Soit donc g l'unique endomorphisme de E dont la matrice dans la base \mathcal{B} est $B = \text{Diag}(\sqrt{\lambda_1}, \dots, \sqrt{\lambda_n})$.

Puisque \mathcal{B} est orthonormée, et que B est une matrice symétrique, g est un endomorphisme symétrique.

De plus, $B^2 = A$ donc $g^2 = f$.

Enfin, les valeurs propres de g sont les $\sqrt{\lambda_i}$, qui sont tous strictement positifs, donc par la question 2, g est défini positif.

Et donc g satisfait bien aux conditions de l'énoncé.

SOLUTION DE L'EXERCICE 15.12

Puisque A est symétrique réelle, elle est diagonalisable, semblable à une matrice diagonale $D = \text{Diag}(\lambda_1, \dots, \lambda_n)$.

D'autre part, on a supposé que A est annihilée par le polynôme $X^p - 1$, et donc les λ_i sont tous des racines de $X^p - 1$.

Mais les seules racines réelles de $X^p - 1$ sont 1 et, dans le cas où p est pair, -1 .

Donc $\text{Spec}(A) \subset \{-1, 1\}$.

En particulier, pour tout $i \in \llbracket 1, n \rrbracket$, $\lambda_i^2 = 1$.

Mais A^2 est semblable à $D^2 = \text{Diag}(\lambda_1^2, \dots, \lambda_n^2) = I_n$.

La seule matrice semblable à I_n étant I_n , on en déduit que $A^2 = I_n$.

SOLUTION DE L'EXERCICE 15.13

1. Soit (e_1, \dots, e_n) une base orthonormée de E formée de vecteurs propres de f , et soit λ_i la valeur propre associée à e_i .

Quitte à renuméroter les e_i , on peut supposer que $\lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_n$, de sorte que $\lambda = \lambda_1$ et $\mu = \lambda_n$.

Alors si $x \in E$ est un vecteur non nul, de manière unique, $x = \sum_{i=1}^n \mu_i e_i$ et donc

$$f(x) = \sum_{i=1}^n \mu_i f(e_i) = \sum_{i=1}^n \mu_i \lambda_i e_i.$$

Il vient alors

$$\langle f(x), x \rangle = \left\langle \sum_{i=1}^n \mu_i \lambda_i e_i, \sum_{j=1}^n \mu_j e_j \right\rangle = \sum_{i=1}^n \sum_{j=1}^n \mu_i \mu_j \lambda_i \underbrace{\langle e_i, e_j \rangle}_{=0 \text{ si } i \neq j} = \sum_{i=1}^n \mu_i^2 \lambda_i \underbrace{\|e_i\|^2}_{=1}.$$

Puisque pour tout $i \in \llbracket 1, n \rrbracket$, on a $\lambda \leq \lambda_i \leq \mu$, il vient

$$\lambda \sum_{i=1}^n \mu_i^2 \leq \langle f(x), x \rangle \leq \mu \sum_{i=1}^n \mu_i^2.$$

D'autre part, la base (e_1, \dots, e_n) étant orthonormée, on a

$$\|x\|^2 = \langle x, x \rangle = \left\langle \sum_{i=1}^n \mu_i e_i, \sum_{j=1}^n \mu_j e_j \right\rangle = \sum_{i=1}^n \sum_{j=1}^n \mu_i \mu_j \langle e_i, e_j \rangle = \sum_{i=1}^n \mu_i^2.$$

On en déduit que pour tout $x \neq 0$, on a

$$\lambda \leq \frac{\langle f(x), x \rangle}{\|x\|^2} \leq \mu.$$

2. Avec les notations précédentes, on a $f(e_1) = \lambda_1 e_1 = \lambda e_1$, de sorte que

$$\frac{\langle f(e_1), e_1 \rangle}{\|e_1\|^2} = \frac{\lambda_1 \langle e_1, e_1 \rangle}{1} = \lambda_1.$$

Ainsi, lorsque x parcourt $E - \{0\}$, la plus petite valeur prise par $\frac{\langle f(x), x \rangle}{\|x\|^2}$ est $\lambda = \lambda_1$.

De même, on a $\frac{\langle f(e_n), e_n \rangle}{\|e_n\|^2} = \lambda_n = \mu$.

Remarque

Avec un peu plus de travail, on pourrait prouver que g est l'unique endomorphisme satisfaisant à ces conditions.

Méthode

Dans le cas où les matrices semblables vous poseraient des problèmes, n'hésitez pas à dire «il existe D diagonale et P inversible telles que $A = P^{-1}DP$ » et à faire tous les calculs avec P et P^{-1} .

Déjà vu

Nous venons ici de reprouver une formule qui est en fait du cours : si (e_1, \dots, e_n) est une base orthonormée de E et si

$$x = \sum_{i=1}^n x_i e_i$$

alors

$$\|x\|^2 = \sum_{i=1}^n x_i^2.$$

On a donc bien prouvé que :

$$\lambda = \min_{x \neq 0} \frac{\langle f(x), x \rangle}{\|x\|^2} \text{ et que } \mu = \max_{x \neq 0} \frac{\langle x, f(x) \rangle}{\|x\|^2}.$$

SOLUTION DE L'EXERCICE 15.14

$$\text{On a } A \begin{pmatrix} 2 \\ -1 \\ 0 \end{pmatrix} = \begin{pmatrix} 2 \\ -1 \\ 0 \end{pmatrix}.$$

Donc 1 est valeur propre de A et $\begin{pmatrix} 2 \\ -1 \\ 0 \end{pmatrix}$ est un vecteur propre associé.

$$\text{De plus, on a } \text{rg}(A - I_3) = \text{rg} \begin{pmatrix} 1 & 2 & -2 \\ 2 & 4 & -4 \\ -1 & -4 & 4 \end{pmatrix}.$$

Toutes les colonnes de cette matrice sont colinéaires⁸, donc elle est de rang 1.

Et donc $\dim E_1(A) = 3 - \text{rg}(A - I_3) = 2$.

De plus, A est symétrique, à coefficients réels, donc diagonalisable, et elle admet donc une autre valeur propre λ et $12 = \text{tr}(A) = 1 + 1 + \lambda \Leftrightarrow \lambda = 10$.

On en déduit que toutes les valeurs propres de A sont positives, et donc que $q(x) \geq 0$ pour tout $x \in \mathbf{R}^n$.

SOLUTION DE L'EXERCICE 15.15

Notons que, d'après la deuxième colonne de A , -25 est valeur propre de A .

$$\text{De plus, } \dim E_{-25}(A) = 3 - \text{rg}(A + 25I_3) = 3 - \text{rg} \begin{pmatrix} 27 & 0 & -36 \\ 0 & 0 & 0 \\ -36 & 0 & 48 \end{pmatrix}.$$

Mais la première et la dernière colonne sont colinéaires ($L_3 = -4/3L_1$) et donc $A + 25I_3$ est de rang 1.

On en déduit que $E_{-25}(A)$ est de dimension 2.

De plus, A est symétrique réelle, donc diagonalisable. Elle possède donc une seule autre valeur propre λ , avec un sous-espace propre de dimension 1.

Et donc on a $0 = \text{tr}(A) = -25 \times 2 + \lambda \Leftrightarrow \lambda = 50$.

Ainsi, A possède une valeur propre strictement positive et une valeur propre strictement négative et donc n'est pas de signe constant.

De plus, nous savons que si $X \in \mathcal{M}_{n,1}(\mathbf{R})$ est un vecteur propre de A associé à une valeur propre strictement positive (resp. négative), alors $q(X) > 0$ (resp. < 0).

Un calcul permet de prouver que $\begin{pmatrix} -3 \\ 0 \\ 4 \end{pmatrix} \in E_{50}(A)$ et donc si on pose $x = (-3, 0, 4)$, $q(x) > 0$.

De même, $\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \in E_{-25}(A)$ et donc si $y = (0, 1, 0)$, alors $q(y) < 0$.

SOLUTION DE L'EXERCICE 15.16

La matrice tAA est symétrique car sa transposée est ${}^tA({}^tA) = {}^tAA$.

Si $p \in \mathbf{N}^*$ est tel que $A^p = 0$, alors, puisque A et sa transposée commutent, il vient

$$({}^tAA)^p = {}^tA^pA^p = 0.$$

Ainsi, tAA est annihilée par le polynôme X^p , qui possède 0 comme unique racine. Donc la seule valeur propre de tAA est 0.

Puisque cette matrice est symétrique, elle est diagonalisable, et donc semblable à $\text{Diag}(0, \dots, 0) = 0$. Et donc ${}^tAA = 0$.

D'autre part, si on note $A = (a_{i,j})$, alors $0 = \text{tr}({}^tAA) = \sum_{i=1}^n \sum_{j=1}^n a_{i,j}^2$.

On en déduit que les $a_{i,j}$ sont tous nuls, et donc $A = 0$.

SOLUTION DE L'EXERCICE 15.17

- Puisque f est symétrique, il existe une base orthonormée de E formée de vecteur propres de f . Notons \mathcal{B} une telle base.

Max/min

Un minimum est un mino- rant qui est une valeur at- teinte. Donc dans la première question, nous avons prouvé que λ est un mino- rant de l'ensemble des $\frac{\langle f(x), x \rangle}{\|x\|^2}$, et puisque cette valeur est at- teinte pour $x = e_1$, c'est bien le minimum.

⁸ La seconde est égale à deux fois la première et la troi- sième est -2 fois la première.

Transposée

Si X est un vecteur colonne, tX est un vecteur ligne, donc un élément de \mathbf{R}^n . Nous transposons seulement pour être en cohérence avec le fait qu'une forme quadratique est définie sur \mathbf{R}^n et non sur $\mathcal{M}_{n,1}(\mathbf{R})$ (le passage de l'un à l'autre se faisant sans difficulté par transposition).

⚠ Attention !

Une matrice annihilée par X^p n'est pas nécessairement nulle ! Regarder par exemple $A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$. Ce que nous affirmons ici c'est que si cette matrice est de plus diagonalisable, alors elle est nulle.

Alors la matrice de f dans cette base est $D = \text{Diag}(\lambda_1, \dots, \lambda_n)$ où les λ_i sont les valeurs propres de f (et donc sont positifs).

Soit alors $A = \text{Diag}(\sqrt{\lambda_1}, \dots, \sqrt{\lambda_n})$, et soit ϕ l'unique endomorphisme de E dont la matrice dans la base \mathcal{B} est A .

Alors $A^2 = D$, et donc $\phi^2 = f$.

D'autre part, \mathcal{B} est orthonormée, et A est symétrique, donc ϕ est un endomorphisme symétrique.

2. Il est évident que si $x \in \text{Ker } f \cap \text{Ker } g$, alors $(f + g)(x) = 0$ et donc $x \in \text{Ker}(f + g)$.
 Considérons à présent deux endomorphismes symétriques ϕ et ψ tels que $f = \phi^2$ et $g = \psi^2$.
 Soit alors $x \in \text{Ker}(f + g)$, de sorte que $f(x) = -g(x)$.
 Alors $\langle x, f(x) \rangle = -\langle x, g(x) \rangle$. Soit encore

$$\langle x, \phi^2(x) \rangle = -\langle x, \psi^2(x) \rangle \Leftrightarrow \langle \phi(x), \phi(x) \rangle = -\langle \psi(x), \psi(x) \rangle \Leftrightarrow \|\phi(x)\|^2 = -\|\psi(x)\|^2.$$

Un carré est toujours positif, donc $\psi(x) = \phi(x) = 0$, et donc en particulier, $f(x) = g(x) = 0$.

Ainsi, $x \in \text{Ker}(f) \cap \text{Ker}(g)$ et donc $\text{Ker}(f + g) \subset \text{Ker}(f) \cap \text{Ker}(g)$.

Puisque l'inclusion réciproque a été prouvée précédemment, on a bien $\text{Ker}(f + g) = \text{Ker}(f) \cap \text{Ker}(g)$.

SOLUTION DE L'EXERCICE 15.18

- 1.a. Supposons que $S \in \mathcal{S}_n^+$. Alors puisque S est symétrique réelle, elle est diagonalisable en base orthonormée : soit X_1, \dots, X_n une base orthonormée⁹ de $\mathcal{M}_{n,1}(\mathbf{R})$ formée de vecteurs propres de S , et soit λ_i la valeur propre de S associée à X_i .

Pour $X \in \mathcal{M}_{n,1}(\mathbf{R})$, il existe donc $\alpha_1, \dots, \alpha_n$ des réels tels que $X = \sum_{i=1}^n \alpha_i X_i$. Alors

$$\begin{aligned} {}^t X S X &= \left(\sum_{j=1}^n \alpha_j {}^t X_j \right) S \left(\sum_{i=1}^n \alpha_i X_i \right) \\ &= \left(\sum_{j=1}^n \alpha_j {}^t X_j \right) \left(\sum_{i=1}^n \alpha_i S X_i \right) \\ &= \left(\sum_{j=1}^n \alpha_j {}^t X_j \right) \left(\sum_{i=1}^n \alpha_i \lambda_i X_i \right) \\ &= \sum_{i=1}^n \sum_{j=1}^n \alpha_j \alpha_i \lambda_i {}^t X_j X_i \\ &= \sum_{i=1}^n \alpha_i^2 \lambda_i \underbrace{\|X_i\|^2}_{=1} \\ &\geq 0. \end{aligned}$$

Inversement, supposons que pour tout $X \in \mathcal{M}_{n,1}(\mathbf{R})$, ${}^t X S X \geq 0$. Soit λ une valeur propre de S et soit $X \in \mathcal{M}_{n,1}(\mathbf{R})$ un vecteur propre associé.

Alors ${}^t X S X = {}^t X \lambda X = \lambda \|X\|^2$.

Et donc¹⁰ $\lambda = \frac{{}^t X S X}{\|X\|^2} \geq 0$.

Ainsi, toutes les valeurs propres de S sont positives, et donc $S \in \mathcal{S}_n^+$.

- 1.b. Il est clair que S est symétrique car ${}^t S = {}^t ({}^t A A) = {}^t A A = S$.
 Soit $X \in \mathcal{M}_{n,1}(\mathbf{R})$. Alors ${}^t X S X = {}^t X {}^t A A X = {}^t (A X) A X = \|A X\|^2 \geq 0$.
 D'après la question 1.a, S est donc dans \mathcal{S}_n^+ .
- 1.c. Soit $S \in \mathcal{S}_n^+$. Alors S est diagonalisable en base orthonormée : il existe une matrice diagonale $D = \text{Diag}(\lambda_1, \dots, \lambda_n)$ et une matrice orthogonale P telles que $S = {}^t P D P$. Notons que les λ_i sont tous positifs, puisque $S \in \mathcal{S}_n^+$.
 Soit alors $D_1 = \text{Diag}(\sqrt{\lambda_1}, \dots, \sqrt{\lambda_n})$, de sorte que ${}^t D_1 D_1 = D_1^2 = D$. Et donc en posant $A = D_1 P$, il vient

$${}^t A A = {}^t P {}^t D_1 D_1 P = {}^t P D P = S.$$

- 2.a. Si 0 est valeur propre de UV , alors UV n'est pas inversible, et donc l'une¹¹ des deux matrices U ou V n'est pas inversible.
 Notons alors u et v les endomorphismes de \mathbf{R}^n dont les matrices respectives dans la base

Rappel

La matrice d'un endomorphisme dans une base formée de vecteurs propres est diagonale.

⁹ Pour le produit scalaire canonique de $\mathcal{M}_{n,1}(\mathbf{R})$, qui est défini par

$$\langle X, Y \rangle = {}^t X Y.$$

Les X_i étant deux à deux orthogonaux, ${}^t X_j X_i = 0$ si $i \neq j$.

¹⁰ $\|X\|$ est non nul car il s'agit d'un vecteur propre.

¹¹ Au moins.

canonique sont égales à U et V .

Il est classique que $\text{Im}(v \circ u) \subset \text{Im}(v)$ et $\text{Ker}(u) \subset \text{Ker}(v \circ u)$.

Donc si v n'est pas bijectif, il n'est pas surjectif, et donc $\text{Im}(v) \neq \mathbf{R}^n$. On en déduit que $\text{Im}(v \circ u) \neq \mathbf{R}^n$ et donc $v \circ u$ n'est pas bijectif.

Et si u n'est pas bijectif, $\text{Ker}(u) \neq \{0\}$ et donc $\text{Ker}(v \circ u) \neq \{0\}$, de sorte que $v \circ u$ n'est pas bijectif.

Dans tous les cas, $v \circ u$ n'est pas bijectif : VU n'est pas inversible, et donc 0 est valeur propre de VU .

- 2.b. Le cas de la valeur propre 0 vient d'être traité, et 0 est valeur propre de UV si et seulement si 0 est valeur propre de VU .

Soit à présent $\lambda \in \mathbf{C}^*$ une valeur propre complexe non nulle de UV . Alors il existe $X \in \mathcal{M}_{n,1}(\mathbf{C})$ non nul tel que $UVX = \lambda X$. Et donc $VUVX = \lambda VX$.

Ainsi, si $VX \neq 0$, c'est un vecteur propre de VU associé à la valeur propre λ .

Mais VX ne peut être nul, car sinon on aurait $\lambda X = UVX = 0$, et puisque $\lambda \neq 0$, cela impliquerait $X = 0$, ce qui n'est pas le cas puisque X est un vecteur propre.

Ainsi, toute valeur propre complexe de UV est une valeur propre de VU . Et en échangeant le rôle de U et V , on prouve la réciproque, de sorte que UV et VU ont exactement les mêmes valeurs propres complexes.

- 3.a. Puisque \mathcal{S}_n est un sous-espace vectoriel de $\mathcal{M}_n(\mathbf{R})$, $S + T$ est encore symétrique. Et si $X \in \mathcal{M}_{n,1}(\mathbf{R})$, alors

$${}^tX(T + S)X = {}^tXTX + {}^tXSX \geq 0.$$

Et donc $S + T \in \mathcal{S}_n^+$.

- 3.b. Nous venons de prouver la stabilité de \mathcal{S}_n^+ pour la somme, mais pas la stabilité par multiplication externe.

Et de fait : $I_n \in \mathcal{S}_n^+$, mais $-I_n \notin \mathcal{S}_n^+$, puisque sa seule valeur propre est $-1 < 0$.

- 4.a. On a ${}^t(ST) = {}^tT^tS = TS$, et donc ST est symétrique si et seulement si $ST = TS$, c'est-à-dire si et seulement si S et T commutent.

- 4.b. D'après la question 1.c, il existe deux matrices A et B de $\mathcal{M}_n(\mathbf{R})$ telles que $S = {}^tAA$ et $T = {}^tBB$.

Et donc $ST = {}^tAA^tBB$. Puisque d'après 2.b, les valeurs propres d'un produit de matrices ne dépendent pas de l'ordre dans lequel on fait le produit, les valeurs propres de ST sont donc celles de ${}^tB^tAAB = {}^t(AB)(AB)$.

Mais d'après la question 1.b, cette matrice est dans \mathcal{S}_n^+ , et donc toutes ses valeurs propres complexes sont en fait réelles et positives.

Et donc toutes les valeurs propres complexes de ST sont réelles et positives.

Rappel

Pour un endomorphisme d'un espace de dimension finie, injectivité, bijectivité et surjectivité sont équivalentes.

Réels/complexes

Une matrice à coefficients réels peut être vue comme une matrice à coefficients complexes (les réels sont les complexes de partie imaginaire nulle), et donc il est légitime de s'intéresser à ses valeurs propres complexes.

Sev de $\mathcal{M}_n(\mathbf{R})$

Un bon moyen de prouver que \mathcal{S}_n est un sous-espace vectoriel de $\mathcal{M}_n(\mathbf{R})$ est de remarquer qu'il s'agit du noyau de l'endomorphisme de $\mathcal{M}_n(\mathbf{R})$ qui à M associe $M - {}^tM$. Et donc comme tout noyau, c'est un sous-espace vectoriel.

⚠ Attention !

On n'a pas dit pour autant que $ST \in \mathcal{S}_n^+$: ce n'est le cas que si ST est symétrique, c'est-à-dire si S et T commutent.

ESTIMATION PONCTUELLE

Cours	481
16.1 Vocabulaire de l'estimation	481
16.2 Biais d'un estimateur	483
16.3 Risque quadratique et estimateurs convergents	487
Méthodes	492
Exercices.	493
Correction des exercices	496

16.1 VOCABULAIRE DE L'ESTIMATION

16.1.1 Cadre général

Imaginons qu'on dispose d'une famille de variables aléatoires $(X_n)_{n \in \mathbf{N}^*}$ i.i.d dont on connaît «le type de loi», mais dont un ou plusieurs paramètres restent à déterminer.

Exemples 16.1

- On sait que les X_n suivent une loi exponentielle $\mathcal{E}(\lambda)$, mais on ne connaît pas la valeur de λ .
- Les X_n suivent une loi normale $\mathcal{N}(\mu, \sigma^2)$ mais on ne connaît ni μ ni σ .

Une telle situation est fréquente, le type de loi pouvant être connu pour de multiples raisons, aussi bien théoriques qu'empiriques.

Par exemple, on sait que la durée de vie d'un composant électronique suit en général une loi exponentielle, que le nombre de voitures se présentant à un péage en une heure suit une loi de Poisson, etc.

Le but de ce chapitre est d'étudier les moyens d'évaluer au mieux les paramètres inconnus à l'aide de la connaissance d'un échantillon des X_i , c'est-à-dire à l'aide d'observations de valeurs prises par les X_i (dans les exemples précédents, cela signifie à partir de l'observation de la durée de vie de n composants électroniques, ou du nombre de voitures se présentant au péage en n heures).

16.1.2 Espace des paramètres

Définition 16.2 – On appelle espace des paramètres l'ensemble $\Theta \subset \mathbf{R}^n$ des valeurs que peut prendre le paramètre θ .

Exemples 16.3

- Dans le cas où l'on sait que les X_i suivent une loi $\mathcal{E}(\lambda)$, alors $\theta = \lambda$ et $\Theta = \mathbf{R}_+^*$.
- Si les X_i suivent une loi normale $\mathcal{N}(\mu, \sigma^2)$ dont on ne connaît ni l'espérance ni la variance, alors $\theta = (\mu, \sigma)$, et $\Theta = \mathbf{R} \times \mathbf{R}_+^*$.
- Si l'on sait que les X_i suivent une loi normale centrée dont on ne connaît pas l'écart-type, alors $\theta = \sigma$, et $\Theta = \mathbf{R}_+^*$.

16.1.3 Échantillons

Définition 16.4 – On appelle n -échantillon un n -uplet (X_1, \dots, X_n) de variables aléatoires i.i.d. Leur loi commune est alors appelée loi-mère de l'échantillon.



On admet dans toute la suite qu'il existe un espace probabilisable (Ω, \mathcal{A}) sur lequel sont définies les X_i , et tel que pour tout $\theta \in \Theta$, il existe une loi de probabilité P_θ , tel que les X_i suivent la loi de paramètre θ .

Dans toute la suite, on se place sur un tel espace, et on suppose que pour tout $\theta \in \Theta$, les X_i sont i.i.d pour P_θ .

Exemple 16.5

Dans le cas où les X_i suivent une loi exponentielle de paramètre $\lambda = \theta$, alors on a

$$\forall \theta \in \mathbf{R}_+^*, \forall x > 0, P_\theta(X_i \leq x) = 1 - e^{-\theta x}.$$

△ Rappelons que la définition d'une variable aléatoire nécessite juste la donnée d'un espace probabilisable (Ω, \mathcal{A}) , mais que pour parler de sa loi, il faut une probabilité P . Ici nous disposons de plusieurs probabilités sur le même espace, et la loi des X_i n'est pas la même pour différentes valeurs de θ .

Définition 16.6 – Si X est une variable aléatoire sur (Ω, \mathcal{A}) , on notera, si elle existe, $E_\theta(X)$ l'espérance de X pour la probabilité P_θ . De même, si X admet une variance pour la probabilité P_θ , on note $V_\theta(X)$ cette variance.

Exemple 16.7

Si X_i suit une loi exponentielle de paramètre θ , alors $E_\theta(X_i) = \frac{1}{\theta}$.

16.1.4 Estimateur

Définition 16.8 – On appelle estimateur (d'ordre n) de θ toute variable aléatoire de la forme $\varphi(X_1, \dots, X_n)$, où (X_1, \dots, X_n) est un n -échantillon et où φ est une fonction de \mathbf{R}^n dans \mathbf{R} .

Plus généralement, si $g : \Theta \rightarrow \mathbf{R}$ est une fonction, on appelle estimateur de $g(\theta)$ toute variable aléatoire de la forme $\varphi(X_1, \dots, X_n)$ où (X_1, \dots, X_n) est un n -échantillon.

Remarques. • L'idée générale est que si (x_1, \dots, x_n) est un n -uplet de valeurs prises par les X_i (c'est-à-dire des valeurs réellement observées par l'expérience), alors $\varphi(x_1, \dots, x_n)$ est une estimation de $g(\theta)$.

• La terminologie est trompeuse : si $\varphi(X_1, \dots, X_n)$ est un estimateur de $g(\theta)$, cela ne signifie absolument pas que $\varphi(x_1, \dots, x_n)$ donnera une estimation convenable de $g(\theta)$. Le but de ce chapitre est justement de comparer les différents estimateurs de $g(\theta)$ et d'essayer d'isoler les plus pertinents.

Exemple 16.9

Le nombre de voitures se présentant à un péage en une journée est généralement bien modélisé par une loi de Poisson $\mathcal{P}(\lambda)$.

Pour un péage donné, on peut se demander quelle est la valeur de λ .

Soient donc (X_1, \dots, X_n) un n -échantillon de variables aléatoires de loi mère $\mathcal{P}(\lambda)$, $\lambda \in \mathbf{R}_+^*$.

Alors $T = \frac{X_1 + \dots + X_n}{n}$ est un estimateur de λ .

Interprétation

Un estimateur est une variable aléatoire qui dépend uniquement de X_1, \dots, X_n et surtout pas de θ . En effet, nous cherchons à estimer la valeur de θ , donc il serait absurde de vouloir utiliser sa valeur, que nous ne connaissons pas !

Cela signifie que si l'on observe pendant n heures le nombre de voitures franchissant le péage et qu'on note x_1, \dots, x_n les valeurs observées¹, alors on pourra prendre comme estimation de λ le nombre $\frac{x_1 + \dots + x_n}{n}$.

De même $V = \min(X_1, \dots, X_n)$ est un autre estimateur de λ . Ce qui signifie que dans ce cas, on peut prendre $\min(x_1, \dots, x_n)$ comme estimation de λ .

Ce qu'on essaye de faire dans la suite, c'est de décider :

- s'il est pertinent de prendre ces deux nombres comme estimation de λ ou non
- si l'un de ces deux nombres est plus pertinent² que l'autre.

Un troisième estimateur de λ est par exemple $W = -(X_1 + X_n)$.

Intuitivement, il n'est absolument pas pertinent³, puisqu'il s'agit d'observer ce qui se passe pendant n heures, et de ne garder que les observations faites à la première et à la dernière heure (pourquoi ne pas utiliser les autres ?), et surtout car il prendra toujours une valeur négative, alors que le λ que l'on cherche à estimer est positif.

¹ Autrement dit, x_1, \dots, x_n sont des réalisations des variables aléatoires X_i .

² En gros : «laquelle de ces deux valeurs est la plus proche de la véritable valeur de λ » ?

³ Et pourtant, il répond bien à la définition d'estimateur.

Exemples 16.10

- Si la loi mère est une loi uniforme $\mathcal{U}([0, \theta])$, alors $X_1, \frac{X_1 + X_2}{2}, \dots, \frac{X_1 + \dots + X_n}{n}$ sont des estimateurs de θ . De même, $X_1 + X_2$ est un estimateur de θ .
- Si les X_i suivent une loi normale $\mathcal{N}(\mu, \sigma^2)$, alors $\frac{X_1 + \dots + X_n}{n}$ est un estimateur de μ . Ici $\mu = p(\theta)$, où $\theta = (\mu, \sigma)$ et $p(x, y) = x$. C'est également un estimateur de σ , probablement peu pertinent...
- La raison pour laquelle on veut également pouvoir estimer $g(\theta)$ et non seulement θ est qu'il est parfois plus facile d'avoir $g(\theta)$ que θ . Par exemple, dans le cadre d'une loi-mère exponentielle de paramètre θ , il est plus facile d'évaluer $\frac{1}{\theta}$, qui est l'espérance des X_i que θ . De même, dans le cas d'une loi uniforme sur $[a, b]$, on peut vouloir estimer l'étendue de la loi, c'est-à-dire $b - a$.

Définition 16.11 – On appelle moyenne empirique l'estimateur

$$\bar{X}_n = \frac{X_1 + \dots + X_n}{n}.$$

16.2 BIAIS D'UN ESTIMATEUR

16.2.1 Définition

Définition 16.12 – Soit T_n un estimateur de $g(\theta)$. Si pour tout $\theta \in \Theta$, T_n admet une espérance pour la probabilité P_θ , alors on appelle biais de T_n en $g(\theta)$ le réel

$$b_\theta(T_n) = E_\theta(T_n) - g(\theta).$$

Si pour tout $\theta \in \Theta$ on a $b_\theta(T_n) = 0$ (ou de manière équivalente $E_\theta(T_n) = g(\theta)$), on dit que T_n est un estimateur sans biais.

Dans le cas contraire, on dit que T_n est un estimateur biaisé de $g(\theta)$.

Remarque

Le biais peut tout à fait dépendre de θ , ce sera même fréquemment le cas.

Exemple 16.13

Dans le cas où la loi-mère est une loi uniforme sur $[0, \theta]$. Soit $T_n = \max(X_1, \dots, X_n)$.

Il est classique que pour $x \in \mathbf{R}$, par indépendance des X_i , on a

$$P(T_n \leq x) = \prod_{i=1}^n P(X_i \leq x) = P(X_1 \leq x)^n = \begin{cases} 0 & \text{si } x \leq 0 \\ \left(\frac{x}{\theta}\right)^n & \text{si } 0 < x \leq \theta \\ 1 & \text{si } x > \theta \end{cases}$$

De plus, puisque X_1 est une variable à densité, F_{X_1} est continue sur \mathbf{R} et \mathcal{C}^1 sauf éventuellement en un nombre fini de points, et donc il en est de même de F_{T_n} .

T_n est donc une variable à densité, et une densité en est donnée par

$$f_{\theta}(x) = \begin{cases} n \frac{x^{n-1}}{\theta^n} & \text{si } x \in [0, \theta] \\ 0 & \text{sinon} \end{cases}$$

Ainsi,

$$E_{\theta}(T_n) = \int_0^{\theta} x f_{\theta}(x) dx = \int_0^{\theta} n \frac{x^n}{\theta^n} dx = \frac{n}{n+1} \theta.$$

Donc T_n est un estimateur biaisé de θ et le biais de T_n en θ est

$$b_{\theta}(T_n) = \frac{n}{n+1} \theta - \theta = -\frac{\theta}{n+1}.$$

En revanche, $\frac{n+1}{n} T_n$ est un estimateur sans biais de θ car

$$E_{\theta} \left(\frac{n+1}{n} T_n \right) = \frac{n+1}{n} E_{\theta}(T_n) = \frac{n+1}{n} \frac{n}{n+1} \theta = \theta.$$

Essayons de simuler cette situation en SciLab .

Commençons par tirer un $\theta > 0$ au hasard. Par exemple avec

```
theta = grand(1,1,'exp',10) ;
```

Ainsi, θ est un nombre positif, enregistré dans la mémoire de SciLab , nous pouvons y avoir accès facilement⁴, mais nous ne le connaissons pas !

Pour simuler un n -échantillon de la loi uniforme sur $[0, \theta]$, nous pouvons utiliser

```
X = grand(1,100,'unf',0,theta) ;
```

Ainsi, X est un vecteur ligne, à 100 colonnes, chacune de ces colonnes correspondant à une simulation de la loi uniforme $\mathcal{U}([0, \theta])$.

Afin d'estimer la valeur de θ , on peut donc faire

```
T = max(X)
```

Cette valeur sera nécessairement inférieure à θ , et en moyenne sera légèrement biaisée.

Si l'on souhaite une valeur plus «fiable», on peut utiliser

```
101/100*max(X)
```

Par exemple, en réalisant une de ces simulations, j'ai obtenu $T_{100} = 1.516$, $\frac{101}{100} T_{100} = 1.531$, alors que la valeur de θ était 1.553.

Notons qu'on n'est **jamais** à l'abri d'un «mauvais» échantillon : il se peut⁵ que tous les X_i soient inférieurs à $\frac{\theta}{2}$ et donc que T_n et $\frac{n+1}{n} T_n$ retourne une estimation de θ assez éloignée de sa valeur réelle.

Si l'on souhaite réellement voir le biais comme une moyenne, il faut simuler un grand nombre d'échantillons. Par exemple

```
1 X = grand(1000,100,'unf',0,theta) ;
2 B = zeros(1000,1) ;
3 for i = 1 :1000
4     B(i) = max(X(i, :))-theta ;
5 end
6 mean(B)
```

Remarque

On a choisi ici de tirer θ suivant une loi exponentielle de paramètre 1/10, afin qu'il soit positif, mais la loi que nous avons choisie n'a aucune importance !

⁴ En tapant par exemple `disp(theta)`.

Moyenne

Cela signifie que si on répète un grand nombre de fois l'expérience, c'est-à-dire que l'on simule de nouveau X et que l'on calcule `max(X)`, alors la moyenne des valeurs observées sera légèrement inférieure à θ .

⁵ Même si ceci est assez peu probable, surtout pour de grandes valeurs de n .

Dans ce programme, chaque ligne de X correspond à une simulation de 100 variables aléatoires X_1, \dots, X_{100} , à l'aide desquelles on estime θ .

Pour chaque estimation de θ ainsi obtenue, on calcule $T_{100} - \theta$, qui est l'écart entre la vraie valeur de θ et celle obtenue sur l'échantillon.

Si on fait la moyenne de ces écarts, on obtient le biais de T_{100} .

Exemple 16.14

Dans le cas où la loi-mère possède une espérance m et une variance σ^2 (qu'une fois encore, nous devrions noter $m(\theta)$ et $\sigma^2(\theta)$), posons

$$\tilde{S}_n = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X}_n)^2.$$

Alors

$$\begin{aligned} \tilde{S}_n &= \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X}_n)^2 \\ &= \frac{1}{n} \sum_{i=1}^n \sum_{i=1}^n (X_i^2 - 2X_i\bar{X}_n + \bar{X}_n^2) \\ &= \frac{1}{n} \sum_{i=1}^n X_i^2 - 2\bar{X}_n \underbrace{\frac{1}{n} \sum_{i=1}^n X_i}_{=\bar{X}_n} + \frac{1}{n} \sum_{i=1}^n \bar{X}_n^2 \\ &= \frac{1}{n} \sum_{i=1}^n X_i^2 - \bar{X}_n^2. \end{aligned}$$

Et donc par linéarité de l'espérance,

$$E(\tilde{S}_n) = \frac{1}{n} \sum_{i=1}^n E(X_i^2) - E(\bar{X}_n^2) = E(X_1^2) - E(\bar{X}_n^2).$$

Une double application de la formule de Huygens nous donne alors

$$E(X_1^2) = V(X_1) + E(X_1)^2 = \sigma^2 + m^2$$

et

$$E(\bar{X}_n^2) = V(\bar{X}_n) + E(\bar{X}_n)^2 = \frac{1}{n^2} \sum_{i=1}^n V(X_i) + m^2 = \frac{\sigma^2}{n} + m^2.$$

Et donc enfin

$$E(\tilde{S}_n) = \frac{n-1}{n} \sigma^2.$$

Par conséquent, le biais de \tilde{S}_n en tant qu'estimateur de σ^2 est

$$b_\theta(\tilde{S}_n) = E_\theta(\tilde{S}_n) - \sigma^2 = \frac{n-1}{n} \sigma^2 - \sigma^2 = -\frac{\sigma^2}{n}.$$

Ainsi, \tilde{S}_n est un estimateur biaisé de σ^2 , alors que $\frac{n}{n-1} \tilde{S}_n$ est un estimateur sans biais de σ^2 puisque

$$E_\theta\left(\frac{n}{n-1} \tilde{S}_n\right) = \frac{n}{n-1} E_\theta(\tilde{S}_n) = \sigma^2.$$

Scilab

C'est ce résultat qui explique que lorsqu'on calcule un écart-type avec Scilab en utilisant la commande `stdev`, le résultat retourné est

$$\sqrt{\frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2}$$

alors qu'on attendrait plutôt

$$\sqrt{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2}.$$

Proposition 16.15 : Si la loi-mère possède une espérance m (que l'on devrait en fait noter $m(\theta)$) pour tout $\theta \in \Theta$, alors

$$\bar{X}_n = \frac{1}{n} \sum_{i=1}^n X_i$$

est un estimateur sans biais de m .

Interprétation

Ce résultat justifie le fait que pour estimer une espérance, on puisse faire la moyenne des valeurs observées. Ceci ne signifie pas que pour chaque échantillon, la valeur de \bar{X}_n soit exactement égale à m , mais que si on répète ce procédé de nombreuses fois, alors l'erreur moyenne sera nulle.

Démonstration. Pour tout $i \in \llbracket 1, n \rrbracket$, on a $E_\theta(X_i) = m$, et donc par linéarité de E_θ ,

$$E_\theta(\bar{X}_n) = \frac{1}{n} \sum_{i=1}^n E_\theta(X_i) = m.$$

□

Exemple 16.16

Si les X_i suivent la loi exponentielle de paramètre λ , alors $\bar{X}_n = \frac{1}{n} \sum_{i=1}^n X_i$ est un estimateur sans biais de $\frac{1}{\lambda}$.

⚠ Pour autant, $\frac{1}{\bar{X}_n}$ n'est pas un estimateur sans biais de λ , car $E\left(\frac{1}{\bar{X}_n}\right) \neq \frac{1}{E(\bar{X}_n)}$.

16.2.2 Estimateurs asymptotiquement sans biais

Définition 16.17 – Une suite d'estimateurs de $g(\theta)$ est la donnée pour tout $n \geq 1$ d'un estimateur T_n d'ordre n de $g(\theta)$.
En particulier, chaque T_n est de la forme $\varphi_n(X_1, \dots, X_n)$.

Exemple 16.18

Dans le cas où les X_i suivent une loi uniforme sur $[0, \theta]$, en posant $T_n = \max(X_1, \dots, X_n)$, alors $(T_n)_{n \geq 1}$ est une suite d'estimateurs de θ .

Définition 16.19 – Une suite d'estimateurs de $g(\theta)$ est dite asymptotiquement sans biais si

$$\lim_{n \rightarrow +\infty} b_\theta(T_n) = 0.$$

Par abus de langage, on dit parfois que T_n est un estimateur asymptotiquement sans biais de $g(\theta)$.

Interprétation

Cela signifie que les T_n peuvent être des estimateurs biaisés, mais que plus n est grand, plus le biais est petit.

Exemples 16.20

- Nous avons déjà vu que pour tout n , \bar{X}_n est un estimateur sans biais de l'espérance m , donc $(T_n)_{n \geq 1}$ est une suite d'estimateurs de m , asymptotiquement sans biais.
- Dans le cas où la loi-mère est une loi exponentielle $\mathcal{E}(\theta)$, alors $(\bar{X}_n)_{n \geq 1}$ est une suite d'estimateurs asymptotiquement sans biais de $\frac{1}{\theta}$.
- Dans le cas où la loi-mère est une loi uniforme sur $[0, \theta]$, et si $T_n = \max(X_1, \dots, X_n)$,

alors nous savons que les T_n sont des estimateurs de θ , et que leur biais en θ est

$$b_{\theta}(T_n) = -\frac{\theta}{n+1} \xrightarrow{n \rightarrow \infty} 0.$$

Donc $(T_n)_{n \geq 1}$ est une suite d'estimateurs de θ , asymptotiquement sans biais.

16.3 RISQUE QUADRATIQUE ET ESTIMATEURS CONVERGENTS

Nous avons étudié précédemment le biais d'un estimateur, et en particulier la notion d'estimateur sans biais.

Montrons sur un exemple qu'un estimateur sans biais n'est pas nécessairement pertinent. Prenons l'exemple d'un sondage, où une proportion p de la population se dit prête à voter pour un candidat A .

De manière plus formelle, nous disposons de variables aléatoires X_i suivant une loi de Bernoulli de paramètre p , et l'on souhaite estimer p .

Alors X_1 est un estimateur sans biais de p , car

$$E(X_1) - p = p - p = 0.$$

Toutefois, utiliser X_1 comme estimation de p est une hérésie : cela revient à demander à un seul sondé s'il compte voter pour le candidat. Si oui, on estime que A sera élu à l'unanimité, sinon on estime que A n'aura aucune voix !

Il semble donc important de dégager d'autres critères de pertinence d'un estimateur que son biais. Le risque quadratique est un de ces critères (parmi d'autres).

16.3.1 Estimateurs convergents

Définition 16.21 – Une suite $(T_n)_{n \geq 1}$ d'estimateurs de $g(\theta)$ est dite convergente si pour tout $\theta \in \Theta$, la suite (T_n) converge en probabilité vers $g(\theta)$. C'est-à-dire que

$$\forall \theta \in \Theta, \forall \varepsilon > 0, P_{\theta}(|T_n - g(\theta)| \geq \varepsilon) \xrightarrow{n \rightarrow +\infty} 0.$$

Par abus de langage, on dit parfois que T_n est un estimateur convergent de $g(\theta)$.

Remarque. Comme d'habitude, les premiers réflexes pour prouver une convergence en probabilité doivent être Bienaymé-Tchebychev et/ou Markov.

Exemples 16.22

Dans le cas où la loi-mère est la loi uniforme sur $[0, \theta]$.

- Pour tout n , $2X_n$ est un estimateur sans biais de θ , car

$$E_{\theta}(2X_n) = 2E_{\theta}(X_n) = 2 \frac{\theta}{2} = \theta.$$

Pourtant $(2X_n)_{n \geq 1}$ n'est pas une suite convergente d'estimateurs de θ , car pour $0 < \varepsilon < \theta$, on a

$$\begin{aligned} P_{\theta}(|2X_n - \theta| \geq \varepsilon) &= 1 - P_{\theta}(-\varepsilon \leq 2X_n - \theta \leq \varepsilon) = 1 - P_{\theta}\left(\frac{\theta - \varepsilon}{2} \leq X_n \leq \frac{\theta + \varepsilon}{2}\right) \\ &= 1 - \frac{1}{\theta} \left(\frac{\theta + \varepsilon}{2} - \frac{\theta - \varepsilon}{2}\right) = 1 - \frac{\varepsilon}{\theta} \not\xrightarrow{n \rightarrow \infty} 0. \end{aligned}$$

- Si $T_n = \max(X_1, \dots, X_n)$, alors (T_n) est un d'estimateur convergent de θ . En effet, pour $0 < \varepsilon < \theta$, on a

$$P_{\theta}(|T_n - \theta| \geq \varepsilon) = P_{\theta}(\theta - T_n \geq \varepsilon) = P_{\theta}\left(\bigcap_{i=1}^n [X_i \leq \theta - \varepsilon]\right) = \prod_{i=1}^n P_{\theta}(X_i \leq \theta - \varepsilon) = \left(\frac{\theta - \varepsilon}{\theta}\right)^n \xrightarrow{n \rightarrow \infty} 0.$$

Interprétation

Concrètement, cela signifie que pour estimer θ , vous observez une réalisation de X_n , et la multipliez par deux. En moyenne, X_n prend la valeur $\theta/2$, donc en moyenne (c'est-à-dire si vous faites cette expérience un grand nombre de fois) vous ne faites pas d'erreur : c'est un estimateur sans biais. Par contre ponctuellement, vous avez des chances que votre estimation soit très mauvaise, et augmenter n n'y changera rien (puisque de toutes façons vous n'utilisez à chaque fois que le résultat de la n -ème observation).

Remarque. Dans le premier cas, on a une suite d'estimateurs sans biais, mais qui ne converge pas, dans le second, on a une suite d'estimateurs biaisés, mais qui converge. Il est souvent préférable d'avoir le deuxième type d'estimateurs : le premier nous dit qu'en moyenne on ne fait pas d'erreur, mais ponctuellement, l'erreur peut être très grande. Le second nous dit que plus notre échantillon est grand, plus la probabilité de faire une erreur supérieure à ε est petite.

⚠ Même si (T_n) est un estimateur convergent, la convergence en probabilité peut-être assez lente. Par exemple, la figure ci-contre donne, en fonction de n , la probabilité $P_\theta(|T_n - \theta| \geq 0.1)$ pour $\theta = 20$. On constate que même pour $n = 200$, la probabilité que l'écart entre θ et T_n soit supérieur à 0.1 est de l'ordre de 0.4. Ainsi, savoir qu'un estimateur est convergent ne suffit pas à dire quelle valeur de n est «raisonnable» pour obtenir une bonne estimation de $g(\theta)$.

Proposition 16.23 : Si la loi mère admet une espérance m et une variance σ^2 , alors

$$\bar{X}_n = \frac{X_1 + \dots + X_n}{n} \text{ est un estimateur sans biais convergent de } m.$$

Démonstration. Nous avons déjà prouvé qu'il s'agissait d'un estimateur sans biais, et le fait qu'il soit convergent est une reformulation de la loi faible des grands nombres. \square

Exemple 16.24 Estimation du paramètre d'une loi de Bernoulli : les sondages

Supposons que dans une population, la proportion d'individus prêts à voter pour le candidat A soit p , et cherchons à estimer p .

On interroge des individus au hasard en leur posant la question «êtes-vous prêts à voter pour A ?», et leur réponse est assimilée à une variable aléatoire X_i de loi $\mathcal{B}(p)$. Le sondage le plus simple que l'on puisse envisager est d'interroger n personnes, puis de regarder quelle proportion d'individus se sont déclarés favorables au candidat A .

Autrement dit, on estime p par $\bar{X}_n = \frac{1}{n} \sum_{i=1}^n X_i$.

Le résultat énoncé précédemment prouve que \bar{X}_n est un estimateur convergent et sans biais de $E(X_1) = p$, et donc qu'il semble pertinent d'utiliser \bar{X}_n pour estimer la valeur de p .

Proposition 16.25 : Soit $(T_n)_{n \geq 1}$ un estimateur convergent de $g(\theta)$. Alors pour toute fonction continue f , $(f(T_n))_{n \geq 1}$ est un estimateur convergent de $f(g(\theta))$.

Démonstration. Nous savons que si $T_n \xrightarrow{P} g(\theta)$ et si f est continue, alors $f(T_n) \xrightarrow{P} f(g(\theta))$. \square

Exemple 16.26 Sondage avec inversion aléatoire des réponses

Pour s'assurer que les personnes interrogées répondent honnêtement à une question sensible lors d'un sondage, on propose le protocole suivant : la personne interrogée commence par lancer un dé, répond franchement si le dé tombe sur 1, et dit le contraire de ce qu'elle pense sinon.

On note p la proportion de personnes pensant «Oui».

La probabilité qu'une personne réponde «Oui» est $q = \frac{1}{6}p + \frac{5}{6}(1-p)$.

On note F_n la fréquence des personnes répondant «Oui».

Alors F_n est un estimateur sans biais et convergent de q , puisqu'il s'agit de la moyenne empirique.

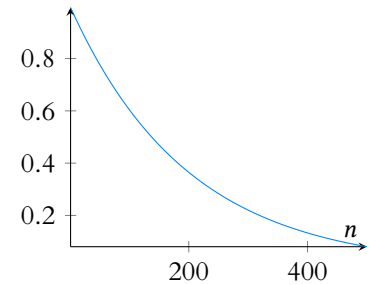


FIGURE 16.1— $P(|M_n - \theta| \geq 0.1)$ en fonction de n .

Mais nous avons

$$q = -\frac{2}{3}p + \frac{5}{6} \Leftrightarrow p = \frac{\frac{5}{6} - q}{\frac{2}{3}} = \frac{5}{4} - \frac{3}{2}q.$$

Donc $E\left(\frac{5}{4} - \frac{3}{2}F_n\right) = \frac{5}{4} - \frac{3}{2}E(F_n) = \frac{5}{4} - \frac{3}{2}q = p$.

Ainsi, $T_n = \frac{5}{4} - \frac{3}{2}F_n$ est un estimateur sans biais de p .

De plus, (F_n) étant un estimateur convergent de q , T_n est un estimateur convergent de p , car c'est l'image de F_n par la fonction continue $x \mapsto \frac{5}{4} - \frac{3}{2}x$.

16.3.2 Risque quadratique

Définition 16.27 – Soit T_n un estimateur de $g(\theta)$. On suppose que pour tout $\theta \in \Theta$, T_n admet un moment d'ordre 2 pour P_θ .

On appelle alors risque quadratique de T_n en θ le nombre

$$r_\theta(T_n) = E_\theta \left((T_n - g(\theta))^2 \right).$$

Exemple 16.28

Supposons que la loi-mère soit une loi de Poisson $\mathcal{P}(\lambda)$.

Alors $\bar{X}_n = \frac{X_1 + \dots + X_n}{n}$ est un estimateur sans biais convergent de λ . De plus,

$$E_\theta((\bar{X}_n - \lambda)^2) = V_\theta(\bar{X}_n) = \frac{1}{n^2} \sum_{i=1}^n V(X_i) = \frac{\lambda}{n^2}.$$

Remarque. Tout comme la variance d'une variable aléatoire, le risque quadratique mesure la «dispersion» de T_n par rapport à $g(\theta)$.

Par exemple, un estimateur peut être sans biais, mais avoir un risque quadratique élevé, s'il prend en moyenne la valeur $g(\theta)$, mais ne prend que des valeurs relativement éloignées de $g(\theta)$. En revanche, un autre estimateur peut s'avérer biaisé, mais prendre des valeurs relativement proches de $g(\theta)$, auquel cas son risque quadratique sera plus faible.

En règle générale, si on dispose de deux estimateurs de $g(\theta)$, on jugera que le meilleur des deux est celui de plus faible risque quadratique.

16.3.3 Décomposition biais-variance du risque quadratique

Proposition 16.29 : Soit T_n un estimateur de $g(\theta)$ admettant un risque quadratique. Alors

$$r_\theta(T_n) = b_\theta(T_n)^2 + V_\theta(T_n).$$

Démonstration. Appliquons la formule de Huygens à la variable aléatoire $T_n - g(\theta)$. Alors

$$V_\theta(T_n - g(\theta)) = \underbrace{E_\theta((T_n - g(\theta))^2)}_{=r_\theta(T_n)} - E_\theta(T_n - g(\theta))^2.$$

Mais $E_\theta(T_n - g(\theta)) = E_\theta(T_n) - g(\theta) = b_\theta(T_n)$.

D'autre part, $g(\theta)$ étant une constante, on a $V_\theta(T_n - g(\theta)) = V_\theta(T_n)$, de sorte que

$$V_\theta(T_n) = r_\theta(T_n) - b_\theta(T_n)^2 \Leftrightarrow r_\theta(T_n) = b_\theta(T_n)^2 + V_\theta(T_n).$$

□

Remarques. • Si le risque quadratique est faible, alors à la fois le biais et la variance sont faibles, de sorte que même si l'estimateur est biaisé, il ne donne en moyenne pas des valeurs

très éloignées de $g(\theta)$.

• Si deux estimateurs ont le même biais, alors celui dont le risque quadratique est le plus faible est celui de plus petite variance. Toutefois, il n'est pas garanti que ce risque soit le plus petit pour tout θ ...

Exemple 16.30

Reprenons une fois de plus l'exemple où les X_i suivent la loi uniforme sur $[0, \theta]$, et soit $M_n = \max(X_1, \dots, X_n)$.

Nous avons alors donné précédemment une densité de M_n , qui est

$$f_\theta : x \mapsto \begin{cases} n \frac{x^{n-1}}{\theta^n} & \text{si } x \in [0, \theta] \\ 0 & \text{sinon} \end{cases}$$

$$\text{Alors } E(M_n^2) = \int_0^\theta x^2 f_\theta(x) dx = \int_0^\theta n \frac{x^{n+1}}{\theta^n} dx = \frac{n}{n+2} \theta^2.$$

$$\text{Et donc } V(M_n) = E(M_n^2) - E(M_n)^2 = \frac{n}{n+2} \theta^2 - \left(\frac{n}{n+1} \theta\right)^2 = \frac{n}{(n+1)^2(n+2)} \theta^2.$$

$$\text{Ainsi, } r(M_n) = b(M_n)^2 + V(M_n) = \frac{\theta^2}{(n+1)^2} + \frac{n}{(n+1)^2(n+2)} \theta^2 = \frac{2}{(n+1)(n+2)} \theta^2.$$

D'autre part, nous savons que $M'_n = \frac{n+1}{n} M_n$ est un estimateur sans biais de θ et donc

$$r(M'_n) = V(M'_n) = \frac{(n+1)^2}{n^2} V(M_n) = \frac{1}{n(n+2)} \theta^2.$$

Et donc $r(M'_n) < r(M_n)$: M'_n est un meilleur estimateur que M_n .

Si l'on souhaite vérifier cela sur une simulation, procédons en tirant 1000 échantillons de 10 nombre tirés sur une loi uniforme.

```
X = grand(1000, 10, 'unf', 0, theta);
```

Puis, pour chacun de ces échantillons, calculons le carré de l'écart entre la valeur estimée de θ (avec M_n et avec M'_n) et la «vraie» valeur de θ .

```
1 r=0;
2 s=0;
3 for i = 1 : 1000
4     r(i) = (max(X(i, :))-theta)^2;
5     s(i) = (11/10*max(X(i, :))-theta)^2;
6 end
```

r et s sont donc deux matrices à 1000 colonnes, chacune d'entre elles comprenant le carré de l'écart à θ sur **un** échantillon.

La moyenne de r , obtenue à l'aide de `mean(r)` doit correspondre à peu près à $r(M_n)$ et la moyenne de s doit correspondre à peu près à $r(M'_n)$.

On constate alors bien que $r(M_n) < r(M'_n)$.

Enfin, soit $T_n = 2\overline{X}_n$.

$$\text{On sait que } b(T_n) = 0 \text{ et } V(T_n) = 4V(\overline{X}_n) = 4 \frac{V(X_1)}{n} = 4 \frac{\theta^2}{12n} = \frac{\theta^2}{3n}.$$

On vérifie aisément que pour $n \geq 3$, $r(M_n) < r(T_n)$ et donc M_n (et a fortiori M'_n) est un meilleur estimateur que T_n .

Gardons à l'esprit que nous ne faisons que des moyennes, nous ne sommes pas à l'abri que, ponctuellement, sur un échantillon donné, T_n donne une bien meilleure estimation de θ que M_n et M'_n .

Par exemple, si $\theta = 10$, $n = 10$ et que nous disposons de l'échantillon : 4, 4, 4.5, 4.5, 5, 5, 5.5, 5.5, 6, 6.

Alors T_{10} va nous donner comme estimation de θ la valeur 10 (soit exactement la

Remarque

Nous avons donc ici deux estimateurs de θ : M_n et T_n . Et bien que M_n soit biaisé (mais asymptotiquement sans biais) alors que T_n est sans biais, M_n est un meilleur estimateur que T_n . Le biais n'est donc pas forcément un critère de choix entre deux estimateurs.



bonne valeur), alors que M_{10} va donner 6 et $M'_{10} = \frac{11}{10}M_{10}$ va nous donner 6.6.
Le risque quadratique ne «voit» donc que l'erreur moyenne.

Remarque

Bien entendu l'échantillon choisi dans cet exemple doit être «rare», et peu représentatif puisque les 10 nombres sont tous éloignés de θ . Mais puisque nous ne connaissons pas θ , nous n'avons aucun moyen de décider si cet échantillon est «bon» ou non. Et donc on utilisera quand même M'_n pour estimer θ , puisque c'est l'estimateur qui en moyenne commet la plus petite erreur.

Exemple 16.31

Cherchons à estimer le paramètre p d'une loi de Bernoulli (par exemple une proportion lors d'un sondage). Notons $S_n = \sum_{i=1}^n X_i$. On sait que $\frac{S_n}{n}$ est un estimateur sans biais convergent de p .

De plus, puisque $S_n \leftrightarrow \mathcal{B}(n, p)$, nous savons que $V(S_n) = np(1 - p)$, et donc

$$V\left(\frac{S_n}{n}\right) = \frac{p(1-p)}{n} \text{ et } r\left(\frac{S_n}{n}\right) = \frac{p(1-p)}{n}.$$

Soit à présent $T_n = \frac{S_n + 1}{n + 2}$.

Alors $E(T_n) = \frac{E(S_n) + 1}{n + 2} = \frac{np + 1}{n + 2}$, et $V(T_n) = \frac{1}{(n + 2)^2}V(S_n) = \frac{np(1-p)}{(n + 2)^2}$.

On en déduit que

$$r(T_n) = b(T_n)^2 + V(T_n) = \left(\frac{np + 1}{n + 2} - p\right)^2 + \frac{np(1-p)}{(n + 2)^2}.$$

Pour p proche de $\frac{1}{2}$, on vérifie que

$$r(T_n) < r\left(\frac{S_n}{n}\right).$$

En revanche, si p est éloigné de $\frac{1}{2}$, alors $r\left(\frac{S_n}{n}\right) < r(T_n)$.

Le risque quadratique n'est donc pas un critère permettant de choisir entre ces deux estimateurs.

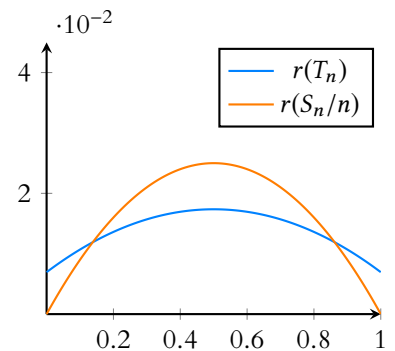


FIGURE 16.2– Comparaison des risques quadratiques de T_n et S_n/n pour $n = 10$.

16.3.4 Une condition suffisante de convergence

Proposition 16.32 : Soit $(T_n)_{n \geq 1}$ une suite d'estimateurs de $g(\theta)$ admettant un risque quadratique. Si pour tout $\theta \in \Theta$,

$$\lim_{n \rightarrow \infty} r_\theta(T_n) = 0$$

alors (T_n) est un estimateur convergent de $g(\theta)$.

Danger !

La réciproque est fautive. De plus, un estimateur convergent n'est pas nécessairement sans biais, ni même asymptotiquement sans biais.

Démonstration. Il s'agit de prouver que T_n converge en probabilité vers $g(\theta)$.

Puisque $T_n - g(\theta)$ admet un moment d'ordre deux, par le corollaire de l'inégalité de Markov, on a pour tout $\varepsilon > 0$

$$P_\theta(|T_n - g(\theta)| \geq \varepsilon) \leq \frac{E((T_n - g(\theta))^2)}{\varepsilon^2} = \frac{r_\theta(T_n)}{\varepsilon^2} \xrightarrow{n \rightarrow \infty} 0.$$

□

Exemple 16.33

En reprenant l'exemple précédent, $r(T_n) \xrightarrow{n \rightarrow \infty} 0$, donc T_n est un estimateur convergent de p .

MÉTHODES DU CHAPITRE 16

✓ MÉTHODE 16.1 : Calculer le biais d'un estimateur T_n de $g(\theta)$

Il suffit de calculer l'espérance de T_n , puis de revenir à la définition du biais.

Si besoin, on pensera à utiliser un théorème de transfert pour calculer cette espérance.

Exercices de TD : 16.1, 16.2, 16.3, 16.4, 16.6, 16.7, 16.9

✓ MÉTHODE 16.2 : Montrer qu'un estimateur T_n de $g(\theta)$ est sans biais

On peut calculer le biais de T_n en $g(\theta)$ et montrer que celui-ci est nul. Ou se contenter de calculer l'espérance de T_n et remarquer que celle-ci vaut $g(\theta)$.

Exercices de TD : 16.2, 16.4, 16.5, 16.6, 16.7

✓ MÉTHODE 16.3 : Proposer un estimateur sans biais de $g(\theta)$ à partir d'un estimateur T_n

Commencer par calculer l'espérance de T_n , puis déterminer une constante λ telle que $\lambda E(T_n) = g(\theta)$. Si on pose alors $U_n = \lambda T_n$, alors $E(U_n) = E(\lambda T_n) = g(\theta)$.

⚠ La constante λ ne peut dépendre de θ car dans la définition d'un estimateur, on ne peut utiliser la valeur de θ : nous cherchons à l'estimer, et ne pouvons donc la supposer connue !

Exercices de TD : 16.3, 16.8, 16.9

✓ MÉTHODE 16.4 : Montrer qu'un estimateur T_n de $g(\theta)$ est asymptotiquement sans biais

On gardera à l'esprit qu'on a alors affaire non pas à un estimateur, mais à une suite $(T_n)_{n \in \mathbb{N}}$ d'estimateurs de $g(\theta)$.

Il faut alors calculer le biais de T_n puis montrer que ce biais tend vers 0 lorsque $n \rightarrow +\infty$.

Exercices de TD : 16.2, 16.3, 16.8

✓ MÉTHODE 16.5 : Calculer le risque quadratique d'un estimateur T_n de $g(\theta)$.

On ne reviendra jamais à la définition, et on utilisera toujours la décomposition biais-variance du risque quadratique. Il s'agit donc de calculer l'espérance de T_n , d'en déduire le biais, puis sa variance. Alors $r_\theta(T_n) = b_\theta(T_n)^2 + V_\theta(T_n)$.

Exercices de TD : 16.1, 16.7, 16.8, 16.11

✓ MÉTHODE 16.6 : Comparer deux estimateurs

Si, parmi plusieurs estimateurs de $g(\theta)$, on doit décider lequel est meilleur, on calcule le risque quadratique de chacun des estimateurs, et on garde celui dont le risque quadratique est le plus faible.

Pour décider que T_n est meilleur que U_n , il faut que $r_\theta(T_n) \leq r_\theta(U_n)$ pour tout $\theta \in \Theta$.

Si jamais le risque quadratique de T_n est plus faible pour certaines valeurs de θ et qu'au contraire c'est le risque quadratique de U_n qui est le plus faible pour d'autres valeurs de θ , il n'est pas possible de déclarer que l'un des deux est meilleur que l'autre.

Exercices de TD : 16.1, 16.7, 16.8, 16.11

✓ MÉTHODE 16.7 : Montrer qu'un estimateur T_n de $g(\theta)$ est convergent

La méthode la plus simple, est de calculer le risque quadratique de T_n puis de montrer qu'il tend vers 0 lorsque $n \rightarrow +\infty$.

Si jamais ce n'est pas le cas, il faut prouver que $T_n \xrightarrow{P} g(\theta)$, soit en utilisant les inégalités de Markov et/ou Bienaymé-Tchebychev, soit en calculant explicitement $P_\theta(|T_n - g(\theta)| \geq \varepsilon)$ pour $\varepsilon > 0$.

Exercices de TD : 16.3, 16.4, 16.5, 16.6, 16.8

EXERCICES DU CHAPITRE 16

EXERCICE 16.1 Soit $(X_n)_{n \geq 1}$ une suite de variables aléatoires indépendantes suivant la loi exponentielle $\mathcal{E}(\lambda)$. On pose

F

$$\bar{X}_n = \frac{1}{n} \sum_{i=1}^n X_i \text{ et } T_n = \frac{1}{n+1} \sum_{i=1}^n X_i.$$

- 1) Calculer le biais de chacun de ces estimateurs de $\frac{1}{\lambda}$.
- 2) Comparer les risques quadratiques de \bar{X}_n et T_n en tant qu'estimateurs de $\frac{1}{\lambda}$.

EXERCICE 16.2 Soit $(X_n)_{n \in \mathbb{N}^*}$ un échantillon de la loi de Poisson de paramètre λ .

PD

Pour tout $n \geq 1$, on pose $S_n = \sum_{k=1}^n X_k$ et $Y_n = e^{-\frac{S_n}{n}}$.

Y_n est-il un estimateur sans biais de $e^{-\lambda}$? Est-il asymptotiquement sans biais ?

EXERCICE 16.3 Soit (X_n) une suite de variables aléatoires indépendantes suivant la loi uniforme sur $[a, b]$, où b est connu, et où l'on cherche à déterminer a . Pour tout $n \geq 1$, on pose $Y_n = \min(X_1, \dots, X_n)$

PD

- 1) Montrer que Y_n est une variable à densité, et en donner une densité.
- 2) Déterminer $E(Y_n)$, puis le biais de Y_n en a . Comment interpréter le signe de ce biais ?
- 3) En utilisant la question précédente, proposer un estimateur sans biais de a .
- 4) Montrer que Y_n est un estimateur asymptotiquement sans biais et convergent de a .

EXERCICE 16.4 Soit $(X_n)_{n \geq 1}$ une suite i.i.d de variables aléatoires suivant la loi $\mathcal{B}(p)$, où p est inconnu. On pose

PD

$$\bar{X}_n = \frac{1}{n} \sum_{i=1}^n X_i \text{ et } T_n = \frac{2}{n(n+1)} \sum_{i=1}^n iX_i.$$

- 1) Montrer que \bar{X}_n et T_n sont deux estimateurs sans biais de p .
- 2) Calculer le risque quadratique de \bar{X}_n et de T_n . Lequel des deux est le meilleur estimateur de p ?
- 3) Montrer que \bar{X}_n et T_n sont deux estimateurs convergents de p .

EXERCICE 16.5 Soient X_1, \dots, X_n des variables aléatoires i.i.d. telles que

PD

$$P(X_1 = -1) = (1-p)^2, P(X_1 = 0) = 2p(1-p), P(X_1 = 1) = p^2.$$

On cherche dans la suite à estimer $p \in]0, 1[$.

Montrer que $Z_n = \sum_{k=1}^n \frac{1+X_k}{2n}$ est un estimateur sans biais et convergent de p .

EXERCICE 16.6 Soit X une variable aléatoire suivant la loi normale centrée de variance σ^2 , avec σ inconnu.

AD

Soit $(X_n)_{n \in \mathbb{N}^*}$ une suite de variables indépendantes, de même loi que X . On note alors $S_n = \frac{1}{n} \sum_{i=1}^n X_i^2$.

Montrer que S_n est un estimateur sans biais et convergent de σ^2 .

EXERCICE 16.7 Comparaison de deux estimateurs du paramètre d'une loi uniforme

AD

Soit θ un réel strictement positif, et soient $(X_k)_{k \in \mathbb{N}^*}$ une suite de variables i.i.d. suivant la loi uniforme sur $[0, 2\theta]$.

On pose $M_n = \max(X_1, \dots, X_n)$ et $\bar{X}_n = \frac{1}{n} \sum_{i=1}^n X_i$.

- 1) Déterminer la loi de M_n , calculer son espérance et sa variance.
- 2) En déduire que $U_n = \frac{n+1}{2n} M_n$ est un estimateur sans biais de θ .
- 3) De \bar{X}_n et U_n , lequel est un meilleur estimateur de θ ?

EXERCICE 16.8 Soit θ un réel strictement positif, et soit f_θ la fonction définie sur \mathbf{R} par

AD

$$f_\theta(x) = \begin{cases} 0 & \text{si } x \leq \theta \\ e^{-(x-\theta)} & \text{si } x > \theta \end{cases}$$

- 1) Montrer que f est une densité de probabilités (savez-vous le faire sans calcul ?).
- 2) Soit X une variable aléatoire dont la densité est f_θ . Calculer $E(X)$ et $V(X)$.
- 3) Dans toute la suite, $(X_n)_{n \geq 1}$ est une suite de variables aléatoires indépendantes sur (Ω, \mathcal{A}) , de densité f_θ .
Pour tout $n \geq 1$, on pose

$$\overline{X}_n = \frac{X_1 + \dots + X_n}{n} \text{ et } Y_n = \min(X_1, \dots, X_n).$$

- a) À partir de \overline{X}_n , construire un estimateur T_n sans biais et convergent de θ .
- b) Montrer que Y_n est un estimateur asymptotiquement sans biais et convergent de θ (on pourra commencer par étudier $Y_n - \theta$).
- c) Construire à partir de Y_n un estimateur U_n sans biais et convergent de θ .
- d) Comparer les deux estimateurs T_n et U_n .

EXERCICE 16.9 Estimation du paramètre d'une loi exponentielle

Soit $(X_n)_{n \geq 1}$ une suite de variables aléatoires indépendantes suivant la loi exponentielle de paramètre λ . On pose $S_n = \sum_{i=1}^n X_i$.

- 1) Déterminer une densité de S_n .
- 2) Calculer l'espérance de $\frac{1}{S_n}$.
- 3) En déduire un estimateur sans biais de λ .

EXERCICE 16.10 Maximum de vraisemblance (D'après ESSEC ECE 2007)

Pour $a > 0$, on note $f_a : t \mapsto \begin{cases} 0 & \text{si } t < 1 \\ \frac{a}{t^{a+1}} & \text{si } t \geq 1 \end{cases}$. Nous avons prouvé dans l'exercice 8 du TD8 que f_a est une densité.

On considère alors une suite $(X_n)_{n \in \mathbf{N}^*}$ de variables aléatoires i.i.d. de densité f_a , où a est inconnu.

Pour $n \in \mathbf{N}^*$, on définit une fonction L_n (appelée fonction de vraisemblance) :

$$L_n : [1; +\infty[^n \times \mathbf{R}_+^*, (x_1, \dots, x_n, a) \mapsto \prod_{i=1}^n f_a(x_i).$$

- 1) Pour (x_1, \dots, x_n) fixés, montrer que $a \mapsto L_n(x_1, \dots, x_n, a)$ possède un maximum atteint en un unique réel a que l'on exprimera en fonction de x_1, \dots, x_n . On pourra à cet effet étudier les variations de la fonction $a \mapsto \ln(L_n(x_1, \dots, x_n, a))$.

On a donc trouvé une expression de $a = \varphi_n(x_1, \dots, x_n)$ où φ_n est une fonction définie sur $[1; +\infty[^n$. On pose alors $T_n = \varphi_n(X_1, \dots, X_n)$, et on appelle T_n estimateur du maximum de vraisemblance.

- 2) Montrer que pour $k \in \llbracket 1, n \rrbracket$, $\ln(X_k)$ suit une loi exponentielle dont on précisera le paramètre. En déduire une densité de $S_n = \ln(X_1) + \dots + \ln(X_n)$.
- 3) Exprimer T_n en fonction de S_n . En déduire $E(T_n)$ et $V(T_n)$.
- 4) Montrer que T_n est un estimateur asymptotiquement sans biais et convergent de a .

EXERCICE 16.11 (QSP ESCP 2005)

Soient T_1 et T_2 deux estimateurs de θ , sans biais et indépendants. Pour tout a réel, on pose $\Theta_a = aT_1 + (1-a)T_2$.

- 1) Θ_a est-il un estimateur sans biais de θ ?
- 2) Parmi tous les Θ_a , lequel a le plus petit risque quadratique ?

EXERCICE 16.12 (D'après oral ESCP 2016)

Soit (Ω, \mathcal{A}, P) un espace probabilisé et $(X_i)_{i \in \mathbf{N}^*}$ une suite de variables aléatoires définies sur cet espace, mutuellement indépendantes, de même loi et ayant pour densité la fonction

$$f : t \mapsto \begin{cases} \frac{2t}{\theta^2} & \text{si } 0 \leq t \leq \theta \\ 0 & \text{sinon} \end{cases}$$

Pour tout $n \in \mathbf{N}^*$, on pose $\overline{X}_n = \frac{1}{n} \sum_{i=1}^n X_i$.

- 1) Vérifier que f est bien une densité de probabilités.
- 2) a) Calculer $E(X_1)$ et en déduire que pour tout $n > 0$, la variable aléatoire $T_n = \frac{3}{2} \overline{X}_n$ est un estimateur sans biais du paramètre θ .

- b) Calculer son risque quadratique $r_\theta(T_n)$ et étudier la convergence en probabilités de la suite d'estimateurs $(T_n)_{n \in \mathbf{N}^*}$.
- 3) Soit la variable aléatoire $M_n = \max(X_1, \dots, X_n)$.
- a) Déterminer la fonction de répartition G de M_n .
- b) Calculer $E(M_n)$ et étudier les propriétés de la variable aléatoire $M'_n = \frac{2n+1}{2n}M_n$ en tant qu'estimateur du paramètre θ .

EXERCICE 16.13 (QSP ESCP 2015)

Soient (X_1, \dots, X_n) des variables aléatoires i.i.d. suivant la même loi d'espérance μ et de variance σ^2 . On définit deux estimateurs de μ en posant

$$T_n = \frac{X_1 + X_2}{2} \text{ et } V_n = \frac{1}{n} \sum_{i=1}^n X_i - \frac{X_1 - X_2}{2}.$$

Lequel de ces deux estimateurs est le meilleur ?

EXERCICE 16.14 Soit X_1, \dots, X_n un n -échantillon de la loi de Poisson de paramètre $\lambda > 0$ inconnu. On cherche à estimer $e^{-\lambda}$.

On définit des variables de Bernoulli Y_1, \dots, Y_n par $Y_k = \begin{cases} 1 & \text{si } X_k = 0 \\ 0 & \text{sinon} \end{cases}$. On pose alors

$$\bar{Y}_n = \frac{1}{n} \sum_{k=1}^n Y_k \text{ et } S_n = \sum_{k=1}^n X_k.$$

- 1) a) Montrer que \bar{Y}_n est un estimateur sans biais de $e^{-\lambda}$?
b) Calculer $V(\bar{Y}_n)$ et en déduire que \bar{Y}_n est un estimateur convergent de $e^{-\lambda}$.
- 2) Pour $j \in \mathbf{N}$, on pose $\varphi(j) = P_{[S_n=j]}(X_1 = 0)$. Calculer $\varphi(j)$.
- 3) On pose à présent $T_n = \varphi(S_n)$.
a) Prouver que T_n est un estimateur sans biais de $e^{-\lambda}$.
b) Calculer $V(T_n)$ et en déduire que T_n est un estimateur convergent de $e^{-\lambda}$.
- 4) Comparer les risques quadratiques des deux estimateurs T_n et \bar{Y}_n .

EXERCICE 16.15 Estimation par capture-recapture (Oral ESCP 2012)

On cherche à évaluer le nombre N de lions d'Asie, espèce en voie de disparition, encore en vie dans la forêt de Gir. Pour cela, on capture d'abord en une seule fois m lions ($m \in \mathbf{N}^*$), que l'on tatoue avant de les relâcher dans la nature, et on admet que pendant toute la durée de l'étude, il n'y a ni naissance ni décès, puis l'on utilise l'une des deux méthodes suivantes.

- 1) **Première méthode** On capture successivement au hasard (donc avec équiprobabilité) et avec remise en liberté après observation du sujet, n lions. Soit Y_n le nombre de lions tatoués parmi eux.
 - a) Déterminer la loi de Y_n . En déduire que $\frac{1}{nm}Y_n$ est un estimateur sans biais et convergent de $\frac{1}{N}$.
 - b) Pourquoi ne peut-on pas prendre $\frac{nm}{Y_n}$ comme estimateur de N ?
 - c) On pose $B_n = \frac{m(n+1)}{Y_n+1}$. Calculer l'espérance de B_n , et montrer que B_n est un estimateur asymptotiquement sans biais de N .
- 2) **Deuxième méthode** On se donne $n \in \mathbf{N}^*$. On capture également, un par un, et avec remise en liberté après observation du sujet, n lions de Gir. On note X_n la variable aléatoire égale au nombre de lions qu'il a été juste nécessaire de capturer pour en obtenir n tatoués. On pose $D_1 = X_1$, et pour tout $i \in \llbracket 2, n \rrbracket$, $D_i = X_i - X_{i-1}$. On admet que les D_i sont mutuellement indépendantes.
 - a) Pour tout $i \in \llbracket 2, n \rrbracket$, que représente concrètement D_i ?
 - b) Déterminer, pour tout $i \in \llbracket 2, n \rrbracket$, la loi de D_i , son espérance et sa variance. En déduire l'espérance et la variance de X_n .
 - c) On pose $A_n = \frac{m}{n}X_n$. Montrer que A_n est un estimateur sans biais et convergent de N , et déterminer son risque quadratique.

CORRECTION DES EXERCICES DU CHAPITRE 16

SOLUTION DE L'EXERCICE 16.1

1. Nous savons déjà que \overline{X}_n est sans biais et puisque $T_n = \frac{n}{n+1}\overline{X}_n$, alors

$$E(T_n) = \frac{n}{n+1}E(\overline{X}_n) = \frac{n}{n+1} \frac{1}{\lambda}.$$

On en déduit que le biais de T_n en $\frac{1}{\lambda}$ est $b(T_n) = \frac{n}{n+1} \frac{1}{\lambda} - \frac{1}{\lambda} = -\frac{1}{\lambda(n+1)}$.

2. Par indépendance des X_i , on a

$$V(\overline{X}_n) = \frac{1}{n^2}V\left(\sum_{i=1}^n X_i\right) = \frac{1}{n}V(X_1) = \frac{1}{n} \frac{1}{\lambda^2}$$

de sorte que

$$r(\overline{X}_n) = 0 + \frac{1}{n\lambda^2}.$$

De plus, $V(T_n) = \frac{n^2}{(n+1)^2}V(\overline{X}_n) = \frac{n}{(n+1)^2\lambda^2}$.

On en déduit que le risque quadratique de T_n en $\frac{1}{\lambda}$ est

$$r(T_n) = \frac{1}{(n+1)^2\lambda^2} + V(T_n) = \frac{1}{(n+1)^2\lambda^2} + \frac{n}{(n+1)^2\lambda^2} = \frac{1}{(n+1)\lambda^2}.$$

Et donc, bien que T_n soit biaisé, contrairement à \overline{X}_n , il possède un risque quadratique plus faible.

SOLUTION DE L'EXERCICE 16.2

1. Notons que par stabilité des lois de Poisson, S_n suit une loi de Poisson de paramètre $n\lambda$. Et donc par le théorème de transfert, on a, sous réserve de convergence

$$\begin{aligned} E(Y_n) &= E\left(e^{-\frac{S_n}{n}}\right) = \sum_{k=0}^{+\infty} e^{-\frac{k}{n}} P(S_n = k) \\ &= \sum_{k=0}^{+\infty} e^{-\frac{k}{n}} e^{-n\lambda} \frac{(n\lambda)^k}{k!} \\ &= e^{-n\lambda} \sum_{k=0}^{+\infty} \left(e^{-1/n} n\lambda\right)^k \frac{1}{k!} = e^{-n\lambda} e^{n\lambda e^{-1/n}}. \end{aligned}$$

Et donc le biais de Y_n en tant qu'estimateur de $e^{-\lambda}$ est

$$b(Y_n) = E(Y_n) - e^{-\lambda} = e^{n\lambda(e^{-1/n}-1)} - e^{-\lambda} \neq 0.$$

2. Lorsque $n \rightarrow +\infty$, $e^{-1/n} - 1 \underset{n \rightarrow +\infty}{\sim} -\frac{1}{n}$ et donc

$$n\lambda \left(e^{1/n} - 1\right) \underset{n \rightarrow +\infty}{\sim} -\lambda \underset{n \rightarrow +\infty}{\longrightarrow} -\lambda.$$

Par continuité de l'exponentielle, on peut donc composer les limites : $e^{n\lambda(e^{-1/n}-1)} \underset{n \rightarrow +\infty}{\longrightarrow} e^{-\lambda}$ et donc $b(Y_n) \underset{n \rightarrow +\infty}{\longrightarrow} 0$.

Ainsi, Y_n est un estimateur asymptotiquement sans biais de $e^{-\lambda}$.

SOLUTION DE L'EXERCICE 16.3

Rédaction 

Puisque T_n pourrait également être considéré comme un estimateur de λ ou de λ^2 , mieux vaut préciser (par une phrase) qu'on note $r(T_n)$ le risque quadratique de T_n en tant qu'estimateur de $\frac{1}{\lambda}$ et non d'autre chose.

Convergence

Il s'agit a priori d'une convergence **absolue**, mais il s'agit d'une série de signe constant, donc la convergence absolue est équivalente à la convergence.

On a reconnu une série exponentielle, donc convergente.

1. Pour $x \in \mathbf{R}$, on a, par indépendance des X_i

$$P(Y_n > x) = P\left(\bigcap_{i=1}^n [X_i > x]\right) = \prod_{i=1}^n P(X_i > x) = \begin{cases} 0 & \text{si } x \geq b \\ \left(\frac{b-x}{b-a}\right)^n & \text{si } a \leq x \leq b \\ 1 & \text{si } x \leq a \end{cases}$$

On en déduit que la fonction de répartition de Y_n est donnée par

$$F_{Y_n}(x) = 1 - P(Y_n > x) = \begin{cases} 0 & \text{si } x \leq a \\ 1 - \left(\frac{b-x}{b-a}\right)^n & \text{si } a \leq x \leq b \\ 1 & \text{si } x \geq b \end{cases}$$

Cette fonction est évidemment de classe \mathcal{C}^1 sur $] -\infty, a[$, sur $]a, b[$ et sur $]b, +\infty[$.

De plus, on a

$$\lim_{x \rightarrow a^-} F_{Y_n}(x) = 0 = \lim_{x \rightarrow a^+} F_{Y_n}(x) = F_{Y_n}(a)$$

et donc elle est continue en a . On prouve de même qu'elle est continue en b et donc continue sur \mathbf{R} .

Et donc F_{Y_n} est continue sur \mathbf{R} et \mathcal{C}^1 sauf éventuellement en a et en b : Y_n est une variable à densité.

Une densité de Y_n est alors donnée par

$$f_n(x) = \begin{cases} 0 & \text{si } x \notin]a, b[\\ \frac{n(b-x)^{n-1}}{(b-a)^n} & \text{si } x \in]a, b[\end{cases}$$

2. Par définition de l'espérance, on a

$$E(Y_n) = \int_{-\infty}^{\infty} x f_n(x) dx = \frac{1}{(b-a)^n} \int_a^b nx(b-x)^{n-1} dx$$

Par une intégration par parties, on obtient alors

$$E(Y_n) = \frac{1}{(b-a)^n} \left([x(b-x)^n]_a^b - \int_a^b (b-x)^n dx \right) = \frac{1}{(b-a)^n} \left(a(b-a)^n + \frac{(b-a)^{n+1}}{n+1} \right) = a + \frac{b-a}{n+1}.$$

Ainsi, le biais de Y_n en a est

$$b(Y_n) = E(Y_n) - a = \frac{b-a}{n+1} \geq 0.$$

Le fait que ce biais soit positif n'a rien de surprenant : Y_n est toujours supérieur à a , car les X_i le sont.

3. Posons $T_n = \frac{n+1}{n} Y_n - \frac{b}{n}$. Alors

$$E(T_n) = -\frac{b}{n} + \frac{n+1}{n} E(Y_n) = \frac{b}{n} + \frac{na+b}{n} = a.$$

Donc T_n est un estimateur sans biais de a .

4. Nous savons déjà que

$$b(T_n) = \frac{b-a}{n+1} \xrightarrow{n \rightarrow \infty} 0$$

dont Y_n est un estimateur asymptotiquement sans biais de a .

De plus, pour $a > \varepsilon > 0$, on a

$$P(|Y_n - a| \geq \varepsilon) = P(Y_n - a \geq \varepsilon) = P(Y_n \geq a + \varepsilon) = \left(\frac{b-\varepsilon}{b-a}\right)^n \xrightarrow{n \rightarrow \infty} 0.$$

Donc (Y_n) converge en probabilité vers a : Y_n est un estimateur convergent de a .

Remarque : pour étudier la convergence en probabilités, il suffit en fait de s'intéresser à $\lim_{n \rightarrow +\infty} P(|Y_n - a| \geq \varepsilon)$ pour ε suffisamment petit, ce qui explique qu'on ne s'intéresse qu'à $\varepsilon < a$. On pourrait toutefois faire le calcul pour $\varepsilon \geq a$.

Astuce

En fait, F_{Y_n} est continue sur $[a, b]$, donc pour prouver qu'elle est continue en a , il suffit de prouver que la limite à gauche en a est égale à $F_{Y_n}(a)$.

Estimateur

Notons qu'on n'aurait pas le droit, par exemple, d'ajouter $\frac{a-b}{n+1}$ à T_n en vue d'obtenir un estimateur sans biais de a . En effet, a est le paramètre inconnu que nous cherchons à estimer, et donc nous ne pouvons utiliser sa valeur. Ceci serait d'ailleurs contraire à la définition d'estimateur ; lorsqu'on écrit $T_n = \varphi(X_1, \dots, X_n)$, il faut que la fonction φ ne dépende pas de θ ! En revanche, puisque la valeur de b est supposée connue, il n'y a pas de problème à l'utiliser dans la définition de T_n .

SOLUTION DE L'EXERCICE 16.4

1. Le fait que $E(\overline{X}_n)$ soit un estimateur sans biais de p est un point de cours.
On a

$$E(T_n) = \frac{2}{n(n+1)} \sum_{i=1}^n iE(X_i) = \frac{2}{n(n+1)} \sum_{i=1}^n ip = p \frac{2}{n(n+1)} \sum_{i=1}^n i = p.$$

Donc le biais de T_n en p est

$$b(T_n) = E(T_n) - p = p - p = 0.$$

Ainsi, T_n est un estimateur sans biais de p .

2. Puisque \overline{X}_n est sans biais, son risque quadratique vaut sa variance et, par indépendance des X_i ,

$$r(\overline{X}_n) = V(\overline{X}_n) = \frac{1}{n^2} \sum_{i=1}^n V(X_i) = \frac{p(1-p)}{n}.$$

De même, on a

$$\begin{aligned} V(T_n) &= \frac{4}{n^2(n+1)^2} \sum_{i=1}^n i^2 V(X_i) = p(1-p) \frac{4}{n^2(n+1)^2} \sum_{i=1}^n i^2 \\ &= p(1-p) \frac{4}{n^2(n+1)^2} \frac{n(n+1)(2n+1)}{6} = p(1-p) \frac{2}{3} \frac{2n+1}{n(n+1)}. \end{aligned}$$

Par conséquent, il vient

$$r(T_n) = V(T_n) = \frac{4}{3} p(1-p) \frac{2n+1}{n(n+1)}.$$

On a alors $r(\overline{X}_n) < r(T_n)$, donc \overline{X}_n est un meilleur estimateur que T_n .

3. On a à la fois $r(\overline{X}_n) \xrightarrow{n \rightarrow +\infty} 0$ et $r(T_n) \xrightarrow{n \rightarrow +\infty} 0$ donc \overline{X}_n et T_n sont deux estimateurs convergents de p .

SOLUTION DE L'EXERCICE 16.5

On a $E(X_1) = p^2 - (1-p)^2 = (p + (1-p))(p - (1-p)) = 2p - 1$.

De plus, X_1 est à support fini, donc admet une variance σ^2 . On a donc

$$E(Z_n) = \frac{1}{2n} \sum_{i=1}^n E(X_i + 1) = \frac{1}{2n} \sum_{i=1}^n 2p = p.$$

Ainsi, le biais de Z_n en tant qu'estimateur de p est

$$b(Z_n) = E(Z_n) - p = p - p = 0.$$

Donc Z_n est un estimateur sans biais de p .

Les X_i étant indépendantes, on a

$$V(Z_n) = \frac{1}{4n^2} \sum_{i=1}^n V(X_i + 1) = \frac{1}{4n^2} \sum_{i=1}^n V(X_i) = \frac{\sigma^2}{2n}.$$

Donc le risque quadratique de Z_n en p est

$$r(Z_n) = b(Z_n) + V(Z_n) = V(Z_n) = \frac{\sigma^2}{4n} \xrightarrow{n \rightarrow +\infty} 0.$$

Ainsi, Z_n est un estimateur convergent de p .

Alternative : on peut aussi remarquer que les $\frac{X_i + 1}{2}$ sont i.i.d., d'espérance p et de variance finie. Puisque Z_n est leur moyenne, la loi faible des grands nombres garantit que $Z_n \xrightarrow{P} p$.

Indépendance

L'hypothèse d'indépendance est indispensable pour garantir que la variance de la somme vaut la somme des variances, attention à ne pas l'oublier.

Variance

Il serait tout à fait possible de calculer cette variance en remarquant que X_1^2 suit une loi de Bernoulli, ce qui permet aisément d'obtenir le moment d'ordre 2. En pratique, nous n'aurons pas besoin de cette valeur.

SOLUTION DE L'EXERCICE 16.6

Les X_i^2 sont indépendantes et identiquement distribuées. De plus, par la formule de Huygens, on a

$$E(X_i^2) = V(X_i) + E(X_i)^2 = \sigma^2 + 0 = \sigma^2.$$

Ainsi, par indépendance des X_i^2 , il vient

$$E(S_n) = \frac{1}{n} \sum_{i=1}^n E(X_i^2) = \sigma^2.$$

Donc S_n est un estimateur sans biais de σ^2 .

De plus, X_i^2 admet un moment d'ordre 2 si et seulement si X_i admet un moment d'ordre 4.

Par le théorème de transfert, c'est le cas si et seulement si $\int_{-\infty}^{\infty} x^4 \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{x^2}{2\sigma^2}} dx$ converge¹.

Mais la fonction $x \mapsto x^4 \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{x^2}{2\sigma^2}}$ est continue sur \mathbf{R} , et au voisinage de $\pm\infty$, on a

$$x^4 \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{x^2}{2\sigma^2}} = o\left(\frac{1}{x^2}\right), \text{ de sorte que l'intégrale converge.}$$

Ainsi, X_i^2 admet un moment d'ordre 2, et alors par la loi faible des grands nombres, S_n converge en probabilité vers σ^2 : S_n est un estimateur convergent de σ^2 .

SOLUTION DE L'EXERCICE 16.7

1. On a, pour tout $x \in \mathbf{R}$, par indépendance des X_i ,

$$P(M_n \leq x) = P\left(\bigcap_{i=1}^n [X_i \leq x]\right) = \prod_{i=1}^n P(X_i \leq x) = P(X_1 \leq x)^n = \begin{cases} 0 & \text{si } x \leq 0 \\ \left(\frac{x}{2\theta}\right)^n & \text{si } 0 < x < 2\theta \\ 1 & \text{si } x \geq 2\theta \end{cases}$$

De plus, $x \mapsto P(X_1 \leq x)$ est continue sur \mathbf{R} et \mathcal{C}^1 sauf en un nombre fini de points² car X_1 est à densité, et donc il en est de même de la fonction de répartition de M_n .

Ainsi, M_n est une variable à densité, dont une densité est

$$f : t \mapsto \begin{cases} n \frac{x^{n-1}}{(2\theta)^n} & \text{si } 0 \leq x \leq 2\theta \\ 0 & \text{sinon} \end{cases}$$

On a alors

$$E(M_n) = \int_0^{2\theta} t f(t) dt = \frac{2\theta n}{n+1}.$$

De même, on a

$$E(M_n^2) = \int_0^{2\theta} t^2 f(t) dt = \frac{4\theta^2 n}{n+2}.$$

Par la formule de Huygens, il vient alors

$$V(M_n) = E(M_n^2) - E(M_n)^2 = \frac{4\theta^2 n}{n+2} - \frac{4n^2\theta^2}{(n+1)^2} = \frac{4\theta n((n+1)^2 - n(n+2))}{(n+2)(n+1)^2} = \frac{4n\theta^2}{(n+1)^2(n+2)}.$$

2. On a $E(U_n) = \frac{n+1}{2n} E(M_n) = \frac{n+1}{2n} \frac{2\theta n}{n+1} = \theta$.
Et donc M_n est un estimateur sans biais de θ .
3. \overline{X}_n est un estimateur sans biais de l'espérance des X_i , qui vaut θ . Donc $b(\overline{X}_n) = 0$.
De plus, par indépendance des X_i , on a

$$V(\overline{X}_n) = \frac{1}{n^2} \sum_{i=1}^n V(X_i) = \frac{1}{n} V(X_1) = \frac{4\theta^2}{12n} = \frac{\theta^2}{3n}.$$

Et alors le risque quadratique de \overline{X}_n en θ vaut

$$r(\overline{X}_n) = b(\overline{X}_n)^2 + V(\overline{X}_n) = \frac{\theta^2}{3n}.$$

Huygens

Rappelons que la formule de Huygens, souvent utilisée pour obtenir la variance à partir du moment d'ordre 2 peut également être utilisée dans l'autre sens.

En particulier, pour les lois usuelles dont nous connaissons espérance et variance, elle permet de retrouver le moment d'ordre 2.

¹ Absolument, mais il s'agit de l'intégrale d'une fonction positive.

Hypothèses

Ne pas oublier de vérifier l'existence d'un moment d'ordre 2 pour appliquer la loi faible des grands nombres !

² En l'occurrence, ici 0 et 2θ .

Sans biais

Pour montrer qu'un estimateur est sans biais de $g(\theta)$, on n'est pas obligé de calculer son biais, il suffit de constater que son espérance vaut $g(\theta)$.

De même, on a

$$r(U_n) = V(U_n) = \frac{(n+1)^2}{4n^2} V(M_n) = \frac{\theta^2}{n(n+2)}.$$

Afin de comparer les risques de U_n et \overline{X}_n , on peut calculer leur quotient :

$$\frac{r(U_n)}{r(\overline{X}_n)} = \frac{3}{n+2} \leq 1 \Leftrightarrow r(U_n) \leq r(\overline{X}_n).$$

On en déduit que U_n est un meilleur estimateur que \overline{X}_n .

SOLUTION DE L'EXERCICE 16.8

1. On peut soit faire le calcul, soit reconnaître directement la densité de $Y + \theta$, où Y suit une loi exponentielle de paramètre 1 (cette densité s'obtient, une fois n'est pas coutume, par transformation affine de densités!).
2. Utilisons le raisonnement de la question précédente : $E(X) = E(Y) + \theta = 1 + \theta$.
De même, $V(X) = V(Y + \theta) = V(Y) = 1$.
- 3.a. Puisque les X_i admettent une variance, par la loi faible des grands nombres, \overline{X}_n est un estimateur sans biais et convergent de $E(X) = \theta + 1$.
Donc si on pose $T_n = \overline{X}_n - 1$, alors

$$E_\theta(T_n) = E_\theta(\overline{X}_n - 1) = \theta + 1 - 1 = \theta.$$

Donc T_n est un estimateur sans biais de θ .

Pour $\varepsilon > 0$, on a

$$P(|T_n - \theta| \geq \varepsilon) = P(|\overline{X}_n - 1 - \theta| \geq \varepsilon) = P(|\overline{X}_n - (\theta + 1)| \geq \varepsilon) \xrightarrow{n \rightarrow +\infty} 0$$

car $\overline{X}_n \xrightarrow{P} \theta + 1$.

Et donc T_n est un estimateur convergent de θ .

- 3.b. $Y_n - \theta = \min(X_1, \dots, X_n) - \theta = \min(X_1 - \theta, \dots, X_n - \theta)$.
Mais les $X_i - \theta$ sont indépendantes, et suivent toutes la même loi exponentielle $\mathcal{E}(1)$. Ainsi, la fonction de répartition de $Y_n - \theta$ est donnée par

$$\begin{aligned} F_{Y_n - \theta}(x) &= P(Y_n \leq x + \theta) \\ &= 1 - P(Y_n > x + \theta) = 1 - P\left(\bigcap_{i=1}^n [X_i > x + \theta]\right) \\ &= 1 - \prod_{i=1}^n P(X_i > x + \theta) = 1 - \prod_{i=1}^n (1 - F_{X_i - \theta}(x)) \\ &= \begin{cases} 0 & \text{si } x \leq 0 \\ 1 - e^{-nx} & \text{si } x > 0 \end{cases} \end{aligned}$$

Donc $Y_n - \theta$ suit une loi exponentielle de paramètre n .

Et alors le biais de Y_n en θ est

$$b_\theta(Y_n) = E_\theta(Y_n) - \theta = E_\theta(Y_n - \theta) = \frac{1}{n}.$$

Ce biais tend vers 0 lorsque n tend vers $+\infty$, et donc Y_n est un estimateur asymptotiquement sans biais de θ .

Enfin, pour $\varepsilon > 0$, on a

$$P_\theta(\underbrace{|Y_n - \theta|}_{\geq 0} \geq \varepsilon) = P_\theta(Y_n - \theta \geq \varepsilon) = e^{-n\varepsilon} \xrightarrow{n \rightarrow \infty} 0.$$

Ainsi, $Y_n \xrightarrow{P} \theta$ et donc Y_n est un estimateur convergent de θ .

- 3.c. Nous avons prouvé précédemment que $E_\theta(Y_n) = \theta + \frac{1}{n}$, donc si l'on pose $U_n = Y_n - \frac{1}{n}$, alors $E_\theta(U_n) = \theta$, et donc U_n est un estimateur sans biais de θ .
De plus,

$$V_\theta(U_n) = V_\theta(Y_n) = V_\theta(Y_n) = V(Y_n - \theta) = \frac{1}{n^2}.$$

Astuce

Pour comparer des risques quadratiques, on peut toujours étudier le signe de la différence, mais il est fréquent qu'ils partagent des facteurs communs, auquel cas l'étude du quotient est plus simple car beaucoup de termes se simplifient.

Remarque

On aurait aussi pu utiliser la continuité de la fonction $t \mapsto t - 1$ et le fait que si $X_n \xrightarrow{P} X$, alors $f(X_n) \xrightarrow{P} f(X)$, mais ici c'est utiliser un marteau piqueur pour écraser une mouche !

Par indépendance des X_i .

Positivité

Les X_i prennent tous des valeurs supérieures à θ , donc il en est de même de Y_n .

Rappel

Nous venons de prouver que $Y_n - \theta$ suit la loi exponentielle de paramètre n .

Ainsi, le risque quadratique de U_n en θ est $r_\theta(U_n) = \frac{1}{n^2} \xrightarrow[n \rightarrow \infty]{} 0$, et donc U_n est un estimateur convergent de θ .

3.d. Les deux estimateurs étant sans biais, leur risque quadratique est égal à leur variance.

$$\text{On a, } V_\theta(T_n) = V_\theta(\bar{X}_n - 1) = V_\theta(\bar{X}_n) = \frac{1}{n^2}.$$

$$\text{Ainsi, } r_\theta(T_n) = V_\theta(T_n) = V_\theta(\bar{X}_n) = \frac{1}{n^2}.$$

Donc $r(U_n) < r(T_n)$, de sorte que U_n est un meilleur estimateur de θ que T_n .

SOLUTION DE L'EXERCICE 16.9

1. Les λX_i sont indépendantes et suivent une loi $\gamma(1)$, donc par stabilité des lois γ , $\lambda S_n \hookrightarrow \gamma(n)$.

Une densité de $S_n = \frac{1}{\lambda}(\lambda S_n)$ est alors obtenue par transformation affine de la densité de la loi γ : il s'agit de

$$g_n : t \mapsto \begin{cases} 0 & \text{si } t \leq 0 \\ \frac{\lambda^n}{(n-1)!} t^{n-1} e^{-\lambda t} & \text{si } t > 0 \end{cases}$$

2. D'après le théorème de transfert, on a, sous réserve de convergence,

$$E\left(\frac{1}{S_n}\right) = \int_0^{+\infty} \frac{1}{t} f_n(t) dt = \int_0^{+\infty} \frac{\lambda^n}{(n-1)!} t^{n-2} e^{-\lambda t} dt.$$

Procédons alors au changement de variable $x = \lambda t$, de sorte que $dt = \frac{dx}{\lambda}$. Il vient

$$\int_0^{+\infty} \frac{1}{t} f_n(t) dt = \frac{\lambda}{(n-1)!} \int_0^{+\infty} x^{n-2} e^{-x} dx = \frac{\lambda}{(n-1)!} \Gamma(n-1) = \frac{\lambda}{n-1}.$$

3. On en déduit que $U_n = \frac{n-1}{S_n}$ est un estimateur sans biais de λ puisque

$$E(U_n) = (n-1)E\left(\frac{1}{S_n}\right) = \lambda.$$

Remarque

On sait que $\bar{X}_n = \frac{S_n}{n}$ est un estimateur sans biais de $\frac{1}{\lambda}$. On serait alors tentés de penser qu'alors $\frac{1}{\bar{X}_n} = \frac{n}{S_n}$ est un estimateur sans biais de λ . Il n'en est rien, il faut en fait prendre $\frac{n-1}{S_n}$.

SOLUTION DE L'EXERCICE 16.10

1. On a $L_n(x_1, \dots, x_n, a) = a^n \prod_{i=1}^n \frac{1}{x_i^{a+1}}$, de sorte que

$$\ln(L_n(x_1, \dots, x_n, a)) = n \ln a - (a+1) \sum_{i=1}^n \ln(x_i).$$

En particulier, la fonction $a \mapsto \ln(L_n(x_1, \dots, x_n, a))$ est dérivable, de dérivée égale

$$\frac{n}{a} - \sum_{i=1}^n \ln(x_i).$$

Cette dérivée est positive si et seulement si $a \leq \frac{n}{\sum_{i=1}^n \ln(x_i)}$.

Et donc $a \mapsto \ln(L_n(x_1, \dots, x_n, a))$ admet un maximum en $a = \frac{n}{\sum_{i=1}^n \ln(x_i)}$.

Et par croissance de la fonction exponentielle, il en est de même de la fonction $a \mapsto L_n(x_1, \dots, x_n, a)$.

Dans la suite, on pose donc $T_n = \frac{n}{\sum_{i=1}^n \ln(X_i)}$.

2. Puisque X_k ne prend que des valeurs supérieures à 1, $\ln(X_k)$ ne prend que des valeurs positives, de sorte que $P(\ln(X_k) \leq x) = 0$ si $x < 0$.
Si $x \geq 0$, alors

$$P(\ln(X_k) \leq x) = P(X_k \leq e^x) = \int_1^{e^x} \frac{a}{t^{a+1}} dt = \left[-\frac{1}{t^a} \right]_1^{e^x} = 1 - e^{-ax}.$$

Et donc $\ln(X_k) \hookrightarrow \mathcal{E}(a)$.

On en déduit que $a \ln(X_k)$ suit une loi exponentielle de paramètre 1, et donc que $aS_n = a \ln(X_1) + \dots + a \ln(X_n)$ suit la loi $\gamma(n)$, par stabilité³ des lois gammas.

Et donc une densité de aS_n est $g_n : t \mapsto \begin{cases} 0 & \text{si } t \leq 0 \\ \frac{1}{(n-1)!} t^{n-1} e^{-t} & \text{si } t > 0 \end{cases}$

³ Qui s'applique car les $a \ln(X_k)$ sont indépendantes car les X_k le sont.

Par transformation affine de densités, une densité de $S_n = \frac{1}{a} \times aS_n$ est alors

$$t \mapsto ag_n(at) = \begin{cases} 0 & \text{si } t \leq 0 \\ \frac{a^n}{(n-1)!} t^{n-1} e^{-at} & \text{si } t > 0 \end{cases}$$

3. On a $T_n = \frac{n}{S_n}$. Donc par le théorème de transfert, sous réserve de convergence, l'espérance de T_n est donnée par

$$E(T_n) = \int_0^{+\infty} \frac{n}{t} \frac{a^n}{(n-1)!} t^{n-1} e^{-at} dt = \frac{na^2}{(n-1)!} \int_0^{+\infty} (at)^{n-2} e^{-at} dt.$$

⁴ Affine donc légitime.

Un changement de variable⁴ $x = at$ donne alors

$$E(T_n) = \frac{na^2}{(n-1)!} \int_0^{+\infty} x^{n-2} e^{-x} \frac{dx}{a} = \frac{an}{(n-1)!} (n-2)! \Gamma(n-1) = \frac{an}{(n-1)!} (n-2)! = \frac{an}{n-1}.$$

Donc déjà, on a $b(T_n) = E(T_n) - a = \frac{a}{n-1} \xrightarrow{n \rightarrow +\infty} 0$, et donc T_n est un estimateur asymptotiquement sans biais de a .

De même, on a, toujours sous réserve de convergence :

$$E(T_n^2) = \int_0^{+\infty} \frac{n^2}{t^2} \frac{a^n}{(n-1)!} t^{n-1} e^{-at} dt = \frac{n^2 a^3}{(n-1)!} \int_0^{+\infty} (at)^{n-3} e^{-at} dt.$$

Et le même changement de variable $x = at$ nous donne

$$E(T_n^2) = \frac{n^2 a^3}{(n-1)!} \int_0^{+\infty} x^{n-3} e^{-x} \frac{dx}{a} = \frac{n^2 a^2}{(n-1)!} (n-3)! = \frac{n^2 a^2}{(n-1)(n-2)}.$$

La formule de Huygens nous donne alors

$$V(T_n) = E(T_n^2) - E(T_n)^2 = \frac{n^2 a^2}{(n-1)(n-2)} - \frac{n^2 a^2}{(n-1)^2} = n^2 a^2 \frac{n-1 - (n-2)}{(n-1)^2(n-2)} = \frac{n^2 a^2}{(n-1)^2(n-2)}.$$

Et donc le risque quadratique de T_n en a est donné par

$$r(T_n) = b(T_n)^2 + V(T_n) = \frac{a^2}{(n-1)^2} + \frac{n^2 a^2}{(n-1)^2(n-2)} = \frac{a^2(n^2 + n - 2)}{(n-1)^2(n-2)} \xrightarrow{n \rightarrow +\infty} 0.$$

Puisque le risque quadratique tend vers 0, on en déduit que T_n est un estimateur convergent de a .

SOLUTION DE L'EXERCICE 16.11

1. On a

$$E(\Theta_a) = aE(T_1) + (1-a)E(T_2) = a\theta + (1-a)\theta = \theta$$

et donc Θ_a est un estimateur sans biais de θ .

Convergence

Rappelons que l'intégrale avant changement de variable converge si et seulement si l'intégrale après changement de variable converge. Et donc nous n'étudions pas la convergence de la première intégrale, celle-ci étant garantie si on constate que la seconde converge (ce qui est le cas ici car nous reconnaissons une intégrale usuelle).

2. T_1 et T_2 étant indépendants, on a

$$V(\Theta_a) = a^2 V(T_1) + (1-a)^2 V(T_2) = a^2(V(T_1) + V(T_2)) - 2aV(T_2) + V(T_2).$$

Si $V(T_1) + V(T_2) = 0$ (c'est-à-dire si T_1 et T_2 sont tous deux constants, et donc nécessairement égaux à θ), alors pour tout a , Θ_a est constant, égal à θ .

Au contraire, si $V(T_1) + V(T_2) > 0$, alors $V(\Theta_a)$ est un polynôme de degré deux en a , qui atteint son minimum⁵ en $a = \frac{V(T_2)}{V(T_1) + V(T_2)}$.

Le biais de Θ_a étant nul, son risque quadratique est égal à sa variance, et donc est minimal pour $a = \frac{V(T_2)}{V(T_1) + V(T_2)}$.

SOLUTION DE L'EXERCICE 16.12

1. La fonction f est positive sur \mathbf{R} , et continue sauf éventuellement en 0 et en θ . De plus, on a

$$\int_{-\infty}^{+\infty} f(t) dt = \int_0^\theta \frac{2t}{\theta^2} dt = \frac{\theta^2}{\theta^2} = 1.$$

Et donc f est bien une densité de probabilités.

2.a. On a

$$E(X_1) = \int_{-\infty}^{+\infty} t f(t) dt = \frac{2}{\theta^2} \int_0^\theta t^3 dt = \frac{2}{\theta^2} \frac{\theta^3}{3} = \frac{2}{3}\theta.$$

On en déduit que $E(\overline{X}_n) = \frac{2}{3}\theta$ et donc

$$E(T_n) = \frac{3}{2}E(\overline{X}_n) = \theta.$$

Et donc T_n est un estimateur sans biais de θ .

2.b. Commençons par calculer $V(X_1)$.

On a $E(X_1^2) = \frac{2}{\theta^2} \int_0^\theta t^3 dt = \frac{\theta^2}{2}$, et donc par la formule de Huygens,

$$V(X_1) = E(X_1^2) - E(X_1)^2 = \frac{\theta^2}{2} - \frac{4}{9}\theta^2 = \theta^2 \frac{1}{18}.$$

Et donc, par indépendance des X_i , $V(\overline{X}_n) = \frac{1}{n^2} \sum_{k=1}^n V(X_k) = \frac{V(X_1)}{n} = \frac{1}{18} \frac{\theta^2}{n}$.

Enfin, il vient

$$r(T_n) = V(T_n) = \frac{9}{4}V(\overline{X}_n) = \frac{\theta^2}{8n}.$$

En particulier, ce risque quadratique tendant vers 0 lorsque $n \rightarrow +\infty$, on en déduit que T_n est un estimateur convergent de θ .

3.a. Pour $x \in \mathbf{R}$, on a

$$P(X_1 \leq x) = \int_{-\infty}^x f(t) dt = \begin{cases} 0 & \text{si } x \leq 0 \\ \frac{x^2}{\theta^2} & \text{si } x \in [0, \theta] \\ 1 & \text{si } x > \theta \end{cases}$$

Et donc

$$G(x) = P(M_n \leq x) = P\left(\bigcap_{i=1}^n [X_i \leq x]\right) = \prod_{i=1}^n P(X_i \leq x) = \begin{cases} 0 & \text{si } x \leq 0 \\ \left(\frac{x}{\theta}\right)^{2n} & \text{si } x \in [0, \theta] \\ 1 & \text{si } x > \theta \end{cases}$$

3.b. Il est aisé de prouver que G est continue sur \mathbf{R} et \mathcal{C}^1 sauf peut-être en 0 et en θ , et donc que M_n est une variable à densité, dont une densité est donnée par

$$g : x \mapsto \begin{cases} 0 & \text{si } x \leq 0 \\ 2n \frac{x^{2n-1}}{\theta^{2n}} & \text{si } 0 < x < \theta \\ 0 & \text{si } x \geq \theta \end{cases}$$

Rappel

Une variable aléatoire est de variance nulle si et seulement si c'est une variable certaine.

⁵ Une formule bien connue nous explique que le sommet d'une parabole d'équation $y = ax^2 + bx + c$ a pour abscisse $x = -\frac{b}{2a}$.

Déjà vu ?

Un œil exercé aura peut-être reconnu que f est la densité du maximum de deux variables aléatoires indépendantes suivant la loi uniforme sur $[0, \theta]$.

On a alors,

$$E(M_n) = \int_0^\theta t g(t) dt = \frac{2n}{\theta^{2n}} \int_0^\theta t^{2n} dt = \frac{2n}{2n+1} \theta.$$

En particulier, $E(M'_n) = \theta$, et donc M'_n est un estimateur sans biais de θ . De même, on a

$$E(M_n^2) = \frac{2n}{\theta^{2n}} \int_0^\theta t^{2n+1} dt = \frac{2n}{2n+2} \theta^2$$

de sorte que $V(M_n) = E(M_n^2) - E(M_n)^2 = \frac{2n}{(2n+2)(2n+1)^2} \theta^2$.

On en déduit que $V(M'_n) = \frac{(2n+1)^2}{4n^2} V(M_n) = \frac{\theta^2}{2n(2n+2)}$.

Et donc le risque quadratique de M'_n en tant qu'estimateur de θ vaut

$$r(M'_n) = V(M'_n) = \frac{\theta^2}{2n(2n+2)}.$$

En particulier, il est aisé de vérifier que $r(M'_n) \leq r(T_n)$ et donc que M'_n est un meilleur estimateur de θ que T_n .

SOLUTION DE L'EXERCICE 16.13

Notons qu'il s'agit de deux estimateurs sans biais de μ car

$$E(T_n) = \frac{E(X_1) + E(X_2)}{2} = \mu \text{ et } E(V_n) = \frac{1}{n} \sum_{i=1}^n E(X_i) - \frac{E(X_1) - E(X_2)}{2} = \mu - \frac{\mu - \mu}{2}.$$

Afin de calculer leurs risques quadratiques, il faut calculer leurs variances. On a (grâce à l'indépendance des X_i),

$$V(T_n) = \frac{V(X_1) + V(X_2)}{4} = \frac{\sigma^2}{2}, \text{ donc } r(T_n) = \frac{\sigma^2}{2}.$$

De même, on a

$$\begin{aligned} V(V_n) &= V\left(\frac{2-n}{2n} X_1 + \frac{2+n}{2n} X_2 + \sum_{k=3}^n X_k\right) \\ &= \frac{(2-n)^2}{4n^2} V(X_1) + \frac{(2+n)^2}{4n^2} V(X_2) + \frac{1}{n^2} \sum_{i=3}^n V(X_i) \\ &= \sigma^2 \left(\frac{(2-n)^2}{4n^2} + \frac{(2+n)^2}{4n^2} + \frac{n-2}{n^2} \right) = \sigma^2 \frac{n+2}{2n}. \end{aligned}$$

Donc $r(V_n) = V(V_n) \geq r(T_n)$, de sorte T_n est un meilleur estimateur que V_n .

SOLUTION DE L'EXERCICE 16.14

1.a. On a $P(Y_k = 1) = P(X_k = 0) = e^{-\lambda} \frac{\lambda^0}{0!} = e^{-\lambda}$.

Donc les Y_k suivent la loi de Bernoulli de paramètre $e^{-\lambda}$.

Et donc $E(\bar{Y}_n) = e^{-\lambda}$, donc \bar{Y}_n est un estimateur sans biais de $e^{-\lambda}$.

1.b. Puisque les X_k sont mutuellement indépendantes, il en est de même des Y_k . Et donc

$$V(\bar{Y}_n) = \frac{1}{n^2} \sum_{k=1}^n V(Y_k) = \frac{1}{n^2} \sum_{k=1}^n e^{-\lambda}(1 - e^{-\lambda}) = \frac{e^{-\lambda}(1 - e^{-\lambda})}{n} \xrightarrow{n \rightarrow +\infty} 0.$$

Et donc en particulier, \bar{Y}_n est un estimateur convergent de $e^{-\lambda}$.

2. Par définition, on a $\varphi(j) = \frac{P([X_1 = 0] \cap [S_n = j])}{P(S_n = j)}$.

Mais $[X_1 = 0] \cap [S_n = j] = [X_1 = 0] \cap \left[\sum_{k=2}^n X_k = j \right]$.

Par stabilité⁶ des lois de Poisson, $\sum_{k=2}^n X_k$ suit une loi de Poisson de paramètre $(n-1)\lambda$ et

Intuition

Au premier abord, on pourrait penser que V_n est meilleur que T_n puisqu'il prend bien en compte tout l'échantillon. Malheureusement, le facteur $\frac{X_1 - X_2}{2}$ l'empêche d'avoir une variance «petite».

Détails

Des variables de Bernoulli X et Y sont indépendantes si et seulement si $[X = 1]$ et $[Y = 1]$ sont indépendantes. Or ici l'événement $[Y_k = 1]$ est l'événement $[X_k = 0]$, et donc les $[X_k = 0]$ sont mutuellement indépendants.

⁶ Les X_k sont indépendantes.

par le lemme des coalitions, est indépendante de X_1 .

Par conséquent,

$$\begin{aligned} P([X_1 = 0] \cap [S_n = j]) &= P\left([X_1 = 0] \cap \left[\sum_{k=2}^n X_k = j\right]\right) = P(X_1 = 0)P\left(\sum_{k=2}^n X_k = j\right) \\ &= e^{-\lambda} e^{-(n-1)\lambda} \frac{((n-1)\lambda)^j}{j!}. \end{aligned}$$

D'autre part, S_n suit une loi de Poisson de paramètre $n\lambda$ et donc $P(S_n = j) = e^{-n\lambda} \frac{(n\lambda)^j}{j!}$.

Et donc

$$\varphi(j) = \frac{e^{-\lambda} e^{-(n-1)\lambda} \frac{((n-1)\lambda)^j}{j!}}{e^{-n\lambda} \frac{(n\lambda)^j}{j!}} = \left(\frac{n-1}{n}\right)^j.$$

3.a. Utilisons le théorème de transfert pour calculer $E(T_n)$. On a, sous réserve de convergence,

$$\begin{aligned} E(T_n) &= \sum_{j=0}^{+\infty} \varphi(j)P(S_n = j) = \sum_{j=0}^{+\infty} \left(\frac{n-1}{n}\right)^j e^{-n\lambda} \frac{(n\lambda)^j}{j!} \\ &= e^{-n\lambda} \sum_{j=0}^{+\infty} \frac{((n-1)\lambda)^j}{j!} \\ &= e^{-n\lambda} e^{(n-1)\lambda} = e^{-\lambda}. \end{aligned}$$

On a reconnu une série exponentielle, donc convergente.

Et donc T_n est un estimateur sans biais de $e^{-\lambda}$.

3.b. D'après la formule de Huygens, on a $V(T_n) = E(T_n^2) - e^{-2\lambda}$.
Et par le théorème de transfert,

$$\begin{aligned} E(T_n^2) &= \sum_{j=0}^{+\infty} \varphi(j)^2 P(S_n = j) \\ &= \sum_{j=0}^{+\infty} \left(\frac{n-1}{n}\right)^{2j} e^{-n\lambda} \frac{(n\lambda)^j}{j!} \\ &= e^{-n\lambda} \sum_{j=0}^{+\infty} \left(\frac{(n-1)^2 \lambda}{n}\right)^j \frac{1}{j!} \\ &= e^{-n\lambda} e^{(n-1)^2 \lambda / n} \\ &= e^{-\frac{2n+1}{n} \lambda}. \end{aligned}$$

On a reconnu une série exponentielle de paramètre $\frac{(n-1)^2 \lambda}{n}$.

Et donc $V(T_n) = e^{-\frac{2n+1}{n} \lambda} - e^{-2\lambda} = e^{-2\lambda} \left(e^{\frac{\lambda}{n}} - 1\right) \xrightarrow{n \rightarrow +\infty} 0$.

Et donc en particulier, T_n est un estimateur convergent de $e^{-\lambda}$.

4. Puisque \bar{Y}_n et T_n sont des estimateurs sans biais de $e^{-\lambda}$, leurs risques quadratiques sont égaux à leurs variances.

On a alors

$$\frac{V(\bar{Y}_n)}{V(T_n)} = \frac{e^{-\lambda}(1 - e^{-\lambda})}{e^{-2\lambda}(e^{\lambda/n} - 1)} = \frac{e^{\lambda} - 1}{e^{\lambda/n} - 1}.$$

Mais $0 < e^{\lambda/n} - 1 \leq e^{\lambda-1}$ et donc $\frac{V(\bar{Y}_n)}{V(T_n)} \geq 1 \Leftrightarrow V(\bar{Y}_n) \geq V(T_n)$.

On en déduit que T_n est un meilleur estimateur que \bar{Y}_n .

SOLUTION DE L'EXERCICE 16.15

1. Première méthode

1.a. Y_n suit une loi binomiale $\mathcal{B}\left(n, \frac{m}{N}\right)$.

Ainsi, $E(Y_n) = \frac{mn}{N}$, et donc $E\left(\frac{Y_n}{mn}\right) = \frac{1}{N}$. Donc $\frac{Y_n}{mn}$ est un estimateur sans biais de $\frac{1}{N}$.

Le fait que Y_n soit un estimateur convergent est en fait la loi faible des grands nombres :

Y_n est la somme de n variables de Bernoulli indépendantes de paramètre $\frac{m}{N}$, donc⁷ $\frac{Y_n}{n}$ converge en probabilités vers la constante $\frac{m}{N}$: c'est un estimateur convergent de $\frac{m}{N}$.
Et la fonction $t \mapsto \frac{t}{m}$ étant continue, $\frac{Y_n}{nm}$ est un estimateur convergent de $\frac{1}{N}$.

⁷ La loi faible des grands nombres s'applique car une variable de Bernoulli possède une variance.

- 1.b. Le fait que $\frac{Y_n}{mn}$ soit un estimateur sans biais de $\frac{1}{N}$ ne garantit pas en général que $\frac{mn}{Y_n}$ soit un estimateur sans biais de N .
De manière plus terre à terre, $\frac{mn}{Y_n}$ a une probabilité non nulle de ne pas être défini : il suffit pour cela que Y_n soit nul, ce qui arrive avec une probabilité égale à $(1 - \frac{m}{N})^n \neq 0$.

- 1.c. D'après le théorème de transfert, on a⁸

$$\begin{aligned} E(B_n) &= \sum_{k=0}^n \frac{m(n+1)}{k+1} P(Y_n = k) \\ &= \sum_{k=0}^n \frac{m(n+1)}{k+1} \binom{n}{k} \left(\frac{m}{N}\right)^k \left(1 - \frac{m}{N}\right)^{n-k} \\ &= m(n+1) \sum_{k=0}^n \frac{1}{n+1} \binom{n+1}{k+1} \left(\frac{m}{N}\right)^k \left(1 - \frac{m}{N}\right)^{n-k} \\ &= N \sum_{k=1}^{n+1} \binom{n+1}{k} \left(\frac{m}{N}\right)^{k-1} \left(\frac{m}{N}\right)^k \left(1 - \frac{m}{N}\right)^{n+1-k} \\ &= N \left(1^{n+1} - \left(1 - \frac{m}{N}\right)^{n+1}\right) \end{aligned}$$

⁸ La question de la convergence ne se pose pas ici, puisqu'on a affaire à une somme finie (Y_n et donc B_n sont des variables à support fini).

On en déduit que le biais de $\frac{m(n+1)}{Y_n+1}$ en N est

$$b(B_n) = E(B_n) - N = N \left(1 - \frac{m}{N}\right)^{n+1} \xrightarrow{n \rightarrow \infty} 0.$$

Donc B_n est un estimateur asymptotiquement sans biais de N (au moins si $m < N$).

2. Deuxième méthode

- 2.a. D_i représente le nombre de lions à capturer entre la capture du $(i - 1)$ ème lion (exclu) et celle du i ème lion (inclus).
2.b. Puisqu'on relâche les lions déjà capturés, à chaque capture de lion, la probabilité d'en avoir un qui est tatoué est $\frac{m}{N}$. Et donc D_i compte le nombre d'essais nécessaires à l'obtention d'un succès lors d'une répétition d'épreuves de Bernoulli indépendantes : D_i suit la loi géométrique de paramètre $\frac{m}{N}$.
Par conséquent, on a

$$E(D_i) = \frac{N}{m} \text{ et } V(D_i) = \frac{1 - \frac{m}{N}}{\left(\frac{m}{N}\right)^2} = \frac{N(N - m)}{m^2}.$$

- 2.c. D'après la question précédente, on a

$$E(X_n) = \sum_{i=1}^n E(D_i) = \frac{nN}{m}$$

et donc $E(A_n) = N$, de sorte que A_n est un estimateur sans biais de N .
De plus, les D_i étant indépendantes, on a

$$V(X_n) = \sum_{i=1}^n V(D_i) = n \frac{N(N - m)}{m^2}$$

et donc

$$r(A_n) = V(A_n) = \frac{m^2}{n^2} V(X_n) = \frac{N(N - m)}{n} \xrightarrow{n \rightarrow \infty} 0.$$

Puisque $r(A_n) \xrightarrow{n \rightarrow \infty} 0$, A_n est un estimateur convergent de N .

TOPOLOGIE DE \mathbf{R}^n .

CALCUL DIFFÉRENTIEL D'ORDRE 2.

Cours 507
17.1 Éléments de topologie de \mathbf{R}^n 508
17.2 Fonctions continues sur une partie de \mathbf{R}^n 511
17.3 Fonctions de classe \mathcal{C}^1 , extrema locaux. 513
17.4 Fonctions de classe \mathcal{C}^2 516
17.5 Nature locale des points critiques. 519
17.6 À propos des extrema globaux sur un fermé borné. 526
Exercices 527
Correction des exercices 530

Dans tout ce chapitre, \mathbf{R}^n sera muni de sa norme euclidienne

$$\forall (x_1, \dots, x_n) \in \mathbf{R}^n, \|(x_1, \dots, x_n)\| = \sqrt{\sum_{i=1}^n x_i^2}.$$

MOTIVATIONS

Dans ce chapitre, nous considérons de nouveau des fonctions dépendant de plusieurs variables, mais nous autorisons à présent des fonctions définies non plus sur \mathbf{R}^n tout entier, mais sur une partie de \mathbf{R}^n .

Par exemple, la fonction $f(x, y) = \sqrt{y} \ln(x)$ n'est définie que pour $x > 0$ et $y \geq 0$, c'est-à-dire sur $\mathbf{R}_+^* \times \mathbf{R}^+$.

Comme précédemment, nous allons principalement nous intéresser à la recherche des extrema de f .

Prenons pour commencer un exemple très simple d'une fonction d'une seule variable : la fonction f définie sur $[2, 5]$ par $f(x) = x^2$.

Il est aisé de dresser son tableau de variation et de constater qu'elle admet un unique minimum en $x = 2$ et un unique maximum en $x = 5$.

Toutefois, ni 2 ni 5 ne sont des points critiques de f puisque $f'(2) = 4$ et $f'(5) = 10$.

En effet, il a été vu en première année qu'un extremum d'une fonction dérivable f est atteint en un point critique de f si f est définie sur un **intervalle ouvert**, ce qui n'est pas le cas de $[2, 5]$.

Question : par quoi remplacer «intervalle ouvert» pour que ce résultat reste valable pour des fonctions de plusieurs variables ?

D'autre part, sans avoir besoin de faire le tableau de variations, nous pouvons garantir l'existence d'un minimum et d'un maximum de f sur $[2, 5]$ car une fonction continue sur un **segment** admet un maximum et un minimum.

Question : peut-on généraliser ce résultat aux fonctions de plusieurs variables en remplaçant segment par autre chose ?

Enfin, nous généraliserons la notion d'extremum local. Pour les fonctions d'une variable, il est possible d'étudier si un point critique est un extremum local à l'aide du signe des dérivées secondes.

Question : par quoi remplacer des dérivées secondes pour obtenir une généralisation de ce résultat ?

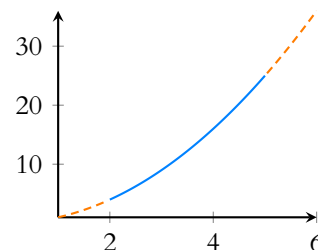


FIGURE 17.1— 2 et 5 ne sont pas des points critiques de f , car on peut la prolonger à gauche de 2 et à droite de 5 en une fonction ne possédant pas de points critiques.

17.1 ÉLÉMENTS DE TOPOLOGIE DE \mathbf{R}^n

17.1.1 Boules ouvertes, boules fermées

Définition 17.1 – Soit $a \in \mathbf{R}^n$ et soit $r > 0$.

On appelle boule fermée de centre a et de rayon r l'ensemble noté $B_f(a, r)$ défini par

$$B_f(a, r) = \{x \in \mathbf{R}^n : \|x - a\| \leq r\}.$$

On appelle boule ouverte de centre a et de rayon r l'ensemble noté $B_o(a, r)$ défini par

$$B_o(a, r) = \{x \in \mathbf{R}^n : \|x - a\| < r\}.$$

Remarque. On a toujours $B_o(a, r) \subset B_f(a, r)$, et le complémentaire de $B_o(a, r)$ dans $B_f(a, r)$ est formé de l'ensemble des points x vérifiant $\|x - a\| = r$.

Intuition

La boule fermée est l'ensemble des points à distance au plus r de a .
La boule ouverte est l'ensemble des points dont la distance à a est inférieure strictement à r .

Exemples 17.2

- Dans \mathbf{R} , $B_f(a, r) = [a - r, a + r]$ et $B_o(a, r) =]a - r, a + r[$.
- Dans \mathbf{R}^2 , $B_f(r, a)$ est un disque de centre a et de rayon r , $B_o(a, r)$ est le disque privé du cercle de centre a et de rayon r .
- Dans \mathbf{R}^3 , $B_f(a, r)$ est la boule de centre a et de rayon r , $B_o(a, r)$ est la même boule privée de la sphère de centre a et de rayon r .

Et si $n > 3$?

En dimension 4 ou plus, il est beaucoup plus dur de se représenter mentalement une boule, et donc pour l'intuition, on s'appuiera essentiellement sur les cas $n = 2$ et $n = 3$.

Intuition

Un ouvert est un ensemble tel que si $x \in A$, alors les points suffisamment proches de x sont également dans A . Graphiquement, un ouvert est une partie qui ne contient pas son «bord».

17.1.2 Image réciproque d'un ouvert, d'un fermé

Définition 17.3 – Une partie A de \mathbf{R}^n est un ouvert de \mathbf{R}^n si

$$\forall x \in A, \exists r > 0 : B_o(x, r) \subset A.$$

Exemples 17.4

- Soit $a \in \mathbf{R}^n$, et soit $r > 0$. Alors $B_o(a, r)$ est un ouvert. En effet, soit $x \in B_o(a, r)$. Alors $\|a - x\| < r$. Soit $r_1 > 0$ tel que $\|a - x\| + r_1 < r$ (notons qu'il existe toujours un tel r_1 , par exemple $r_1 = \frac{r - \|a - x\|}{2}$). Alors $B_o(x, r_1) \subset B_o(a, r)$. En effet, pour $y \in B_o(x, r_1)$, on a

$$\|y - a\| = \|(y - x) + (x - a)\| \leq \|y - x\| + \|x - a\| \leq r_1 + \|x - a\| < r,$$

de sorte que $y \in B_o(a, r)$.

- La boule fermée $B_f(0, r)$ n'est en revanche pas un ouvert. Si x est tel que $\|x\| = r$, alors quel que soit $r_1 > 0$, la boule $B_o(x, r_1)$ n'est pas incluse dans $B_f(0, r)$.

En effet, soit $y = \left(1 + \frac{r_1}{2r}\right)x$. Alors $y \in B_o(x, r_1)$ car

$$\|x - y\| = \left\|x - \left(1 + \frac{r_1}{2r}\right)x\right\| = \frac{r_1}{2r}\|x\| = \frac{r_1}{2} < r_1,$$

mais $\|0 - y\| = \left\|0 - \left(1 + \frac{r_1}{2r}\right)x\right\| = \left(1 + \frac{r_1}{2r}\right)\underbrace{\|x\|}_{=r} > r$, de sorte que $y \notin B_f(0, r)$.

Plus généralement, une boule fermée n'est jamais un ouvert.

- Dans \mathbf{R} , un intervalle ouvert $]a, b[$ est ouvert. En effet, soit $x \in]a, b[$. Alors pour r

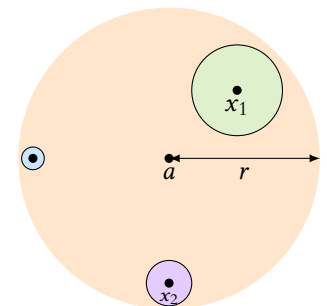


FIGURE 17.2– Le rayon r_1 de la «petite» boule dépend du point x .

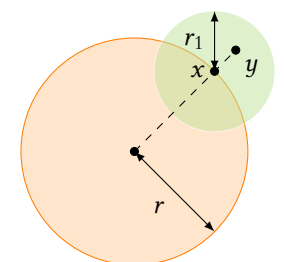


FIGURE 17.3– $B_o(x, r_1)$ n'est pas inclus dans $B_f(0, r)$. On constate que x et y sont colinéaires.

suffisamment petit, on a $B_o(x, r) \subset]x - r, x + r[\subset]a, b[$.

Notons que $]a, b[$ est la boule ouverte de centre $\frac{a+b}{2}$ et de rayon $\frac{b-a}{2}$, donc si on sait qu'une boule ouverte est un ouvert, alors il est évident que $]a, b[$ est un ouvert de \mathbf{R} .

De même, $] - \infty, a[$ et $]a, +\infty[$ sont des ouverts.

• \mathbf{R}^* est un ouvert de \mathbf{R} . Soit $x \in \mathbf{R}^*$, et posons $r = \frac{|x|}{2}$.

Alors $B_o(x, r) =]x - r, x + r[$ ne contient pas 0, et donc est inclus dans \mathbf{R}^* .

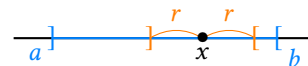


FIGURE 17.4— Pour tout $x \in]a, b[$, il existe r tel que $]x - r, x + r[\subset]a, b[$.

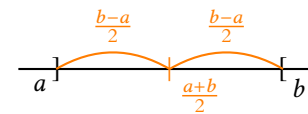


FIGURE 17.5— Un intervalle ouvert est une boule ouverte (de \mathbf{R}) dont le centre est le milieu de l'intervalle.

Proposition 17.5 :

- \mathbf{R}^n est un ouvert de \mathbf{R}^n .
- une intersection finie d'ouverts de \mathbf{R}^n est un ouvert de \mathbf{R}^n .
- une union (éventuellement infinie) d'ouverts de \mathbf{R}^n est un ouvert de \mathbf{R}^n .

Exemple 17.6

$\mathbf{R}^* =] - \infty, 0[\cup]0, +\infty[$ est un ouvert de \mathbf{R} car $] - \infty, 0[$ et $]0, +\infty[$ le sont.

Définition 17.7 – Un ensemble $A \subset \mathbf{R}^n$ est dit fermé si son complémentaire est un ouvert.

Intuition

Une partie fermée est une partie qui contient son «bord».

Exemples 17.8

- Un segment $[a, b]$ de \mathbf{R} est fermé. En effet, son complémentaire est $] - \infty, a[\cup]b, +\infty[$, qui est réunion de deux ouverts de \mathbf{R}^n .
- \mathbf{R}_+ et \mathbf{R}_- sont deux fermés de \mathbf{R} . En effet, leurs complémentaires respectifs sont $] - \infty, 0[$ et $]0, +\infty[$ qui sont ouverts.
- Une boule fermée $B_f(a, r)$ est un fermé de \mathbf{R}^n .
Soit $x \notin B_f(a, r)$. Alors $\|a - x\| > r$, et donc il existe $r_1 > 0$ tel que $\|a - x\| - r_1 > r$.
Et alors pour $y \in B_o(x, r_1)$, on a

$$\|a - x\| = \|(a - y) + (y - x)\| \leq \|a - y\| + \|y - x\| \Leftrightarrow \|a - x\| - \|y - x\| \leq \|a - y\|.$$

Mais alors $\|a - y\| > r$, et donc $y \notin B_f(a, r)$. On a donc prouvé que $B_o(x, r_1)$ est inclus dans le complémentaire de $B_f(a, r)$, et donc ce complémentaire est ouvert : $B_f(a, r)$ est fermé.

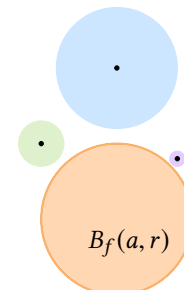


FIGURE 17.6— Pour tout point hors de la boule fermée $B_f(a, r)$, il existe une boule centrée en ce point qui ne rencontre pas $B_f(a, r)$.

Proposition 17.9 :

- \mathbf{R}^n est un fermé de \mathbf{R}^n
- Toute réunion finie de fermés de \mathbf{R}^n est encore un fermé de \mathbf{R}^n .
- Toute intersection (éventuellement infinie) de fermés de \mathbf{R}^n est encore un fermé de \mathbf{R}^n .

Exemple 17.10

Si $a \in \mathbf{R}$, alors $\{a\}$ est un fermé de \mathbf{R} . En effet, $\{a\} = \bigcap_{n \in \mathbf{N}} \left[a - \frac{1}{n}, a + \frac{1}{n} \right]$.

Par conséquent, son complémentaire $\mathbf{R} - \{a\}$ est un ouvert de \mathbf{R} .

Proposition 17.11 : • Un produit cartésien d'ensembles ouverts est ouvert.
 • Un produit cartésien d'ensembles fermés est fermé.

Exemples 17.12

- $\{(x_1, \dots, x_n) \in \mathbf{R}^n : \forall i \in \llbracket 1, n \rrbracket, x_i > 0\}$ est un ouvert de \mathbf{R}^n : c'est $]0, +\infty[\times]0, +\infty[\times \dots \times]0, +\infty[$
- $(\mathbf{R}^*)^n$ est un ouvert de \mathbf{R}^n .
 - $\{(x, y) \in \mathbf{R}^2 : x \geq 0 \text{ et } y \geq 0\} = (\mathbf{R}_+)^2$ est un fermé de \mathbf{R}^2 .
 - $\{(x, y) \in \mathbf{R}^2 : a < x < b \text{ et } c < y < d\} =]a, b[\times]c, d[$ est un ouvert de \mathbf{R}^2
 - De même, $[a, b] \times [c, d]$ est un fermé de \mathbf{R}^2 .

17.1.3 Parties bornées de \mathbf{R}^n

Définition 17.13 – Une partie $A \subset \mathbf{R}^n$ est dite bornée s'il existe $M > 0$ tel que

$$\forall x \in A, \|x\| \leq M.$$

Autrement dit, il existe $M > 0$ tel que $A \subset B_f(0, M)$.

Intuition

Cela signifie que A est inclus dans une boule (possiblement de rayon très grand).

Exemples 17.14

- Un segment de \mathbf{R} est une partie bornée de \mathbf{R} .
- Un intervalle de \mathbf{R} est borné si et seulement si ses deux bornes sont finies.
- Une boule de \mathbf{R}^n (ouverte ou fermée) est bornée : elle est toujours incluse dans la boule fermée de même rayon.

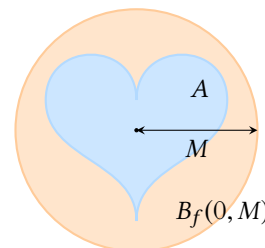


FIGURE 17.7– Un exemple de partie bornée de \mathbf{R}^2 .

Proposition 17.15 : Une partie $A \subset \mathbf{R}^n$ est bornée si et seulement si il existe $N > 0$ tel que

$$\forall (x_1, \dots, x_n) \in A, |x_1| \leq N, |x_2| \leq N, \dots, |x_n| \leq N.$$

Démonstration. Supposons que A soit bornée. Alors il existe $M > 0$ tel que

$$\forall (x_1, \dots, x_n) \in A, \|(x_1, \dots, x_n)\| \leq M.$$

Mais $\forall i \in \llbracket 1, n \rrbracket, x_i^2 \leq \sum_{j=1}^n x_j^2 = \|(x_1, \dots, x_n)\|^2 \leq M^2$.

Donc $|x_i| \leq M$.

Inversement, si il existe $N > 0$ tel que

$$\forall (x_1, \dots, x_n) \in A, \forall i \in \llbracket 1, n \rrbracket, |x_i| \leq N$$

alors

$$\forall (x_1, \dots, x_n) \in A, \|(x_1, \dots, x_n)\|^2 = \sum_{i=1}^n x_i^2 \leq \sum_{i=1}^n N^2 = nN^2$$

et donc

$$\forall (x_1, \dots, x_n) \in A, \|(x_1, \dots, x_n)\| \leq \sqrt{n}N.$$

Donc A est une partie bornée de \mathbf{R}^n . □

Intuition

Ce résultat affirme qu'une partie bornée est en fait une partie incluse dans un «cube» centré en l'origine, et de côté $2N$.

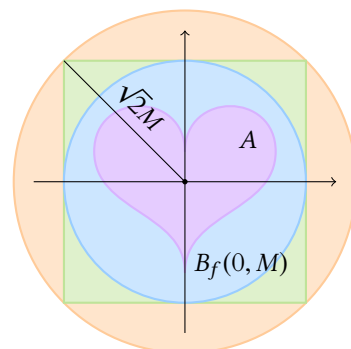


FIGURE 17.8– L'idée de la preuve est essentiellement de dire qu'une boule est incluse dans un cube, lui-même inclus dans une boule plus grande.

Exemples 17.16

- Dans \mathbf{R}^2 , l'ensemble $\mathcal{E} = \left\{ (x, y) \in \mathbf{R}^2 : \frac{x^2}{a^2} + \frac{y^2}{b^2} = 1 \right\}$ est borné.

En effet, pour $(x, y) \in \mathcal{E}$, on a alors

$$\frac{x^2}{a^2} \leq 1 \Leftrightarrow |x| \leq |a|$$

et de même $|y| \leq |b|$, de sorte que pour $N = \max(|a|, |b|)$, $|x| \leq N$ et $|y| \leq N$.

- Dans \mathbf{R}^3 , considérons l'ensemble $\mathcal{H} = \left\{ (x, y, z) \in \mathbf{R}^3 : \frac{x}{yz} = 1 \right\}$.

Il n'est pas borné car pour tout $x > 0$, $(x, \sqrt{x}, \sqrt{x}) \in \mathcal{H}$, et donc x peut prendre des valeurs arbitrairement grandes.

17.2 FONCTIONS CONTINUES SUR UNE PARTIE DE \mathbf{R}^n

Une fonction de plusieurs variables peut n'être définie que sur une partie de \mathbf{R}^n . Par exemple, la fonction

$$f : (x_1, \dots, x_n) \mapsto \sum_{i=1}^n \frac{1}{x_i}$$

est définie sur $(\mathbf{R}^*)^n$, qui est un ouvert de \mathbf{R}^n en tant que produit d'ouverts de \mathbf{R} (\mathbf{R}^* est un ouvert de \mathbf{R}).

De même la fonction g définie par

$$g(x_1, \dots, x_n) = \frac{1}{\sum_{i=1}^n x_i^2}$$

est définie sur $\mathbf{R}^n - \{(0, \dots, 0)\}$ qui est un ouvert de \mathbf{R}^n car il s'agit du complémentaire du singleton $\{(0, \dots, 0)\}$.

17.2.1 Définition, opération sur les fonctions continues

Définition 17.17 – Soit $\Omega \subset \mathbf{R}^n$. Une fonction $f : \Omega \rightarrow \mathbf{R}$ est dite continue en $x_0 \in \Omega$ si

$$\forall \varepsilon > 0, \exists \eta > 0 : \forall x \in B_f(x_0, \eta) \cap \Omega, |f(x) - f(x_0)| \leq \varepsilon.$$

f est continue sur Ω si elle est continue en tout point de Ω .

Intuition

L'intuition est inchangée : pour tout $x \in \Omega$ «suffisamment proche» de x_0 , $f(x)$ et $f(x_0)$ sont proches.

Remarque. En revenant à la définition d'une boule ouverte, cela signifie encore :

$$\forall \varepsilon > 0, \exists \eta > 0 : \forall x \in \Omega, \|x - x_0\| \leq \eta \Rightarrow |f(x) - f(x_0)| \leq \varepsilon.$$

Dans le cas où $\Omega = \mathbf{R}^n$, on retrouve bien la définition vue précédemment d'une fonction continue en x_0 .

Tous les résultats vus précédemment sur les fonctions continues de \mathbf{R}^n dans \mathbf{R} se transposent aux fonctions continues de Ω dans \mathbf{R} . En particulier, les fonctions polynomiales sont continues, la somme de fonctions continues est continue, etc.

Exemple 17.18

Soit $f : \mathbf{R}^* \times]0; +\infty[\rightarrow \mathbf{R}$ définie par $f(x, y) = \ln(x^2 + y)$.

La fonction $(x, y) \mapsto x^2 + y$ est continue (car polynomiale) sur l'ouvert $\mathbf{R}^* \times]0; +\infty[$, et à valeurs dans \mathbf{R}_+^* . La fonction \ln étant continue sur \mathbf{R}_+^* , f est continue sur (l'ouvert) $\mathbf{R}^* \times]0; +\infty[$.

17.2.2 Image réciproque des parties ouvertes/fermées par une fonction continue

Proposition 17.19 : Soit f une fonction continue sur \mathbf{R}^n et à valeurs dans \mathbf{R} . Alors pour tout ouvert $\mathcal{O} \subset \mathbf{R}$, $\{x \in \mathbf{R}^n : f(x) \in \mathcal{O}\}$ est un ouvert de \mathbf{R}^n .
De même, pour tout fermé F de \mathbf{R} , $\{x \in \mathbf{R}^n : f(x) \in F\}$ est un fermé de \mathbf{R}^n .



Démonstration. Soit \mathcal{O} un ouvert de \mathbf{R} , et soit $x_0 \in \{x \in \mathbf{R}^n : f(x) \in \mathcal{O}\}$, c'est-à-dire tel que $f(x_0) \in \mathcal{O}$.

Puisque \mathcal{O} est ouvert, il existe $r > 0$ tel que $]f(x_0) - r, f(x_0) + r[\subset \mathcal{O}$.

Étant donné que f est continue en x_0 , il existe $\eta > 0$ tel que

$$\forall x \in B_o(x_0, \eta), |f(x) - f(x_0)| < r.$$

Mais alors pour $x \in B_o(x_0, \eta)$, on a

$$|f(x) - f(x_0)| < r \Leftrightarrow f(x) \in]f(x_0) - r, f(x_0) + r[$$

Donc pour $x \in B_o(x_0, \eta)$, on a $f(x) \in \mathcal{O}$, et donc $x \in \{x \in \mathbf{R}^n : f(x) \in \mathcal{O}\}$.

Ainsi, $B_o(x_0, \eta) \subset \{x \in \mathbf{R}^n : f(x) \in \mathcal{O}\}$, et donc $\{x \in \mathbf{R}^n : f(x) \in \mathcal{O}\}$ est ouvert.

Si F est un fermé de \mathbf{R} , son complémentaire \bar{F} est ouvert.

Mais alors $\{x \in \mathbf{R}^n : f(x) \in F\}$ est le complémentaire de $\{x \in \mathbf{R}^n : f(x) \in \bar{F}\}$ qui est ouvert par ce qui précède.

Puisque le complémentaire d'un ouvert est un fermé, $\{x \in \mathbf{R}^n : f(x) \in F\}$ est fermé dans \mathbf{R}^n . \square

Remarque. Si A est une partie de \mathbf{R} , alors $\{x \in \mathbf{R}^n : f(x) \in A\}$ est appelé l'image réciproque de A . C'est l'ensemble de tous les antécédents par f d'éléments de A .

Corollaire 17.20 – Soit f une fonction continue de \mathbf{R}^n dans \mathbf{R} . Alors

- $\forall a \in \mathbf{R}, \{x \in \mathbf{R}^n : f(x) < a\}$ et $\{x \in \mathbf{R}^n : f(x) > a\}$ sont des ouverts de \mathbf{R}^n .
- $\forall a \in \mathbf{R}, \{x \in \mathbf{R}^n : f(x) \leq a\}, \{x \in \mathbf{R}^n : f(x) \geq a\}$ et $\{x \in \mathbf{R}^n : f(x) = a\}$ sont des fermés de \mathbf{R}^n .

Démonstration. On a $\{x \in \mathbf{R}^n : f(x) < a\} = \{x \in \mathbf{R}^n : f(x) \in]-\infty, a[\}$, et $] - \infty, a[$ est un ouvert de \mathbf{R} , donc son image réciproque est un ouvert de \mathbf{R}^n . De même pour $\{x \in \mathbf{R}^n : f(x) > a\} = \{x \in \mathbf{R}^n : f(x) \in]a, +\infty[\}$.

On remarque de même que $\{x \in \mathbf{R}^n : f(x) \leq a\} = \{x \in \mathbf{R}^n : f(x) \in]-\infty, a] \}$, et $] - \infty, a]$ est un fermé de \mathbf{R} . Enfin $\{x \in \mathbf{R}^n : f(x) = a\}$ est l'image réciproque de $\{a\}$ et $\{a\}$ est un fermé de \mathbf{R} . \square

Exemples 17.21

- Soit $f(x, y) = x + y + \ln(1 + x^2 + y^2)$. Alors f est continue sur \mathbf{R}^2 , et donc $\{(x, y) \in \mathbf{R}^2 : x + y + \ln(1 + x^2 + y^2) \leq 3\}$ est un fermé de \mathbf{R}^2 .
- $E = \{(x, y, z) \in \mathbf{R}^3 : x^3 + 2xy + 5 = 2\}$ est un fermé de \mathbf{R}^3 car $f(x, y, z) = x^3 + 2xy + 5$ est une fonction continue de \mathbf{R}^3 dans \mathbf{R} .
- Soit $D = \{(x, y) \in \mathbf{R}^2 : x \geq 0, y \geq 0 \text{ et } x + y \leq 1\}$ est un fermé de \mathbf{R}^2 .
En effet, on a $D = \{(x, y) \in \mathbf{R}^2 : x \geq 0\} \cap \{(x, y) \in \mathbf{R}^2 : y \geq 0\} \cap \{(x, y) \in \mathbf{R}^2 : x + y \leq 1\}$, et ces trois ensembles sont fermés, donc leur intersection est encore fermée.
- Soit f la fonction définie par $f(x_1, \dots, x_n) = \frac{1}{\sum_{i=1}^n x_i}$.

Alors son ensemble de définition est $D = \{(x_1, \dots, x_n) \in \mathbf{R}^n : x_1 + \dots + x_n \neq 0\}$.

Son complémentaire est $\{(x_1, \dots, x_n) \in \mathbf{R}^n : x_1 + \dots + x_n = 0\}$, et c'est un fermé car la fonction $(x_1, \dots, x_n) \mapsto \sum_{i=1}^n x_i$ est continue sur \mathbf{R}^n . Donc D est un ouvert

Intuitivement

Un ensemble défini par une inégalité stricte ne contient pas son bord (qui correspond au cas où $f(x) = a$).

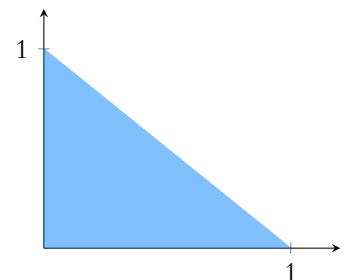


FIGURE 17.9– $D = \{(x, y) \in \mathbf{R}^2 : x \geq 0, y \geq 0 \text{ et } x + y \leq 1\}$.

de \mathbf{R}^n .

17.2.3 Fonctions continues sur un fermé borné

Rappel : Soit $I = [a, b]$ un segment de \mathbf{R} , et soit $f : I \rightarrow \mathbf{R}$ une fonction continue. Alors f est bornée et atteint ses bornes. Autrement dit, f admet un maximum et un minimum sur I .

Notons qu'un segment est une partie de \mathbf{R} qui est à la fois fermée et bornée¹. Le théorème précédent se généralise de la manière suivante :

Théorème 17.22 : Soit f une fonction continue sur un fermé borné $A \subset \mathbf{R}^n$, et à valeurs dans \mathbf{R} . Alors f admet un maximum global et un minimum global sur A .

Démonstration. Admis. □

Remarque. Ce théorème garantit l'existence des extrema, mais ne dit pas du tout comment les trouver !

Exemple 17.23

Soit $D = \{(x, y) \in \mathbf{R}^2 : x \geq 0, y \geq 0 \text{ et } x + y \leq 1\}$.
 Alors nous avons prouvé précédemment que D était fermé.
 De plus, $\forall (x, y) \in D, |x| \leq 1$ et $|y| \leq 1$, donc D est borné.
 Ainsi, la fonction $f(x, y) = xy(1 - x - y)$, qui est continue sur D admet un maximum global et un minimum global sur D .

¹ Bien que la réciproque soit fautive : $[0, 1] \cup [2, 3]$ est une partie fermée bornée de \mathbf{R} , mais ce n'est pas un segment.

17.3 FONCTIONS DE CLASSE \mathcal{C}^1 , EXTREMA LOCAUX

17.3.1 Fonctions de classe \mathcal{C}^1 sur un ouvert de \mathbf{R}^n

Rappel : soit I un intervalle de \mathbf{R} . Pour qu'une fonction $f : I \rightarrow \mathbf{R}$ soit dérivable en $a \in I$, il faut qu'elle possède une dérivée à gauche et une dérivée à droite en a , autrement dit qu'il existe un intervalle **ouvert** contenant a et inclus dans I .

Lemme 17.24. Soit $\Omega \subset \mathbf{R}^n$ un ouvert de \mathbf{R}^n , et soit $a = (a_1, \dots, a_n) \in \mathbf{R}^n$. Alors, $\forall i \in \llbracket 1, n \rrbracket$, il existe $r_i > 0$ tel que

$$\{a_1\} \times \dots \times \{a_{i-1}\} \times]a_i - r_i, a_i + r_i[\times \{a_{i+1}\} \times \dots \times \{a_n\} \subset \Omega.$$

Démonstration. Soit $r > 0$ tel que $B_o(a, r) \subset \Omega$. Alors $r_i = r$ convient : pour tout $t \in]-r, r[$, on a

$$\|(a_1, \dots, a_{i-1}, a_i + t, a_{i+1}, \dots, a_n) - (a_1, \dots, a_n)\| = \|(0, \dots, 0, t, 0, \dots, 0)\| = |t| < r$$

et donc $(a_1, \dots, a_{i-1}, a_i + t, a_{i+1}, \dots, a_n) \in B_o(a, r) \subset \Omega$.

Notons que les r_i obtenus dans la preuve sont tous égaux, et ne sont pas forcément les plus grands possibles, comme on le voit en comparant cette figure avec celle donnée plus haut. □

Ceci implique que si f est une fonction définie sur Ω et à valeurs dans \mathbf{R} , alors pour tout $a = (a_1, \dots, a_n) \in \mathbf{R}^n$, la $i^{\text{ème}}$ fonction partielle de f en a est définie sur un intervalle ouvert $]a_i - r_i, a_i + r_i[$ contenant a_i . Et alors nous pouvons étudier sa dérivée en a_i .

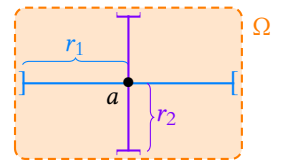


FIGURE 17.10– Illustration du lemme 17.24.

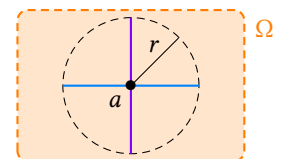


FIGURE 17.11– Illustration du principe de la preuve.

Définition 17.25 – Soit $\Omega \subset \mathbf{R}^n$ un ouvert de \mathbf{R}^n , soit $f : \Omega \rightarrow \mathbf{R}$, et soit $a \in \Omega$. Soit $i \in \llbracket 1, n \rrbracket$. Si la fonction

$$f_i : t \mapsto f(a_1, \dots, a_{i-1}, t, a_{i+1}, \dots, a_n)$$

est dérivable en a_i , on note $\partial_i f(a_1, \dots, a_n)$ le nombre dérivée $f'_i(a_i)$, et on l'appelle $i^{\text{ème}}$ dérivée partielle de f en a .

Lorsque toutes les dérivées partielles de f en a existent, on appelle gradient de f en a , et on note $\nabla f(a)$ le vecteur de \mathbf{R}^n défini par

$$\nabla f(a) = (\partial_1 f(a), \dots, \partial_n f(a)).$$

Si toutes les applications $\partial_i f$ sont définies et continues sur Ω , on dit que f est de classe \mathcal{C}^1 sur Ω .

Tous les théorèmes sur les fonctions \mathcal{C}^1 sur \mathbf{R}^n restent valables sur un ouvert Ω de \mathbf{R}^n . En particulier, les fonctions polynomiales sont encore de classe \mathcal{C}^1 , la somme de fonctions \mathcal{C}^1 est \mathcal{C}^1 , etc.

La formule de Taylor à l'ordre 1 est également valable sur un ouvert de \mathbf{R}^n .

Précision

La proposition précédente garantit que f_i est définie sur un intervalle ouvert contenant a_i , et donc il est légitime de s'intéresser à sa dérivabilité en a_i .

17.3.2 Extrema locaux

Définition 17.26 – Soit $f : \Omega \rightarrow \mathbf{R}$, avec Ω une partie de \mathbf{R}^n . f admet un maximum local en $a \in \Omega$ si il existe $r > 0$ tel que

$$\forall x \in B_o(a, r) \cap \Omega, f(x) \leq f(a).$$

De même, f admet un minimum local en $a \in \Omega$ si il existe $r > 0$ tel que

$$\forall x \in B_o(a, r) \cap \Omega, f(x) \geq f(a).$$

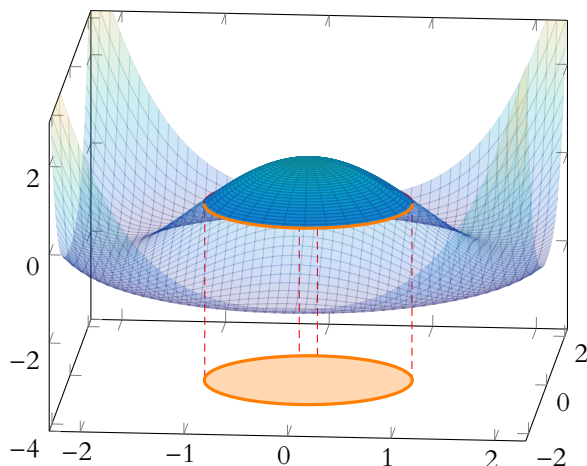


FIGURE 17.12 – Un maximum local en $(0, 0)$: pour tout $x \in B_o((0, 0), 1)$, on a $f(x) \leq f(0, 0)$. Notons qu'il ne s'agit pas d'un maximum global : en dehors de $B_o((0, 0), 1)$, f prend des valeurs supérieures à $f(0, 0)$.

Proposition 17.27 : Si f admet un maximum (respectivement un minimum) global en a , alors f admet un maximum (respectivement un minimum) local en a .

Démonstration. Si f admet un maximum global en a , alors pour tout $r > 0$, on a

$$\forall x \in B_o(a, r) \cap \Omega, f(x) \leq f(a)$$

et donc f admet un maximum local en a . \square

⚠ La réciproque est fautive, comme le montre l'exemple de la figure précédente.

Proposition 17.28 (Condition nécessaire à l'existence d'un extremum local) :
 Soit $f : \Omega \rightarrow \mathbf{R}^n$ une fonction de classe \mathcal{C}^1 sur un ouvert Ω de \mathbf{R}^n . Si f admet un extremum local en $a \in \mathbf{R}^n$, alors $\nabla f(a) = 0$, i.e. a est un point critique de f .

Démonstration. Supposons que f possède un maximum local en $a \in \Omega$, et soit $r > 0$ tel que $B_o(a, r) \subset \Omega$ et

$$\forall x \in B_o(a, r), f(x) \leq f(a).$$

Alors la $i^{\text{ème}}$ fonction partielle de f en a , notée f_i , possède un maximum local en $t = a_i$ car

$$\forall t \in]a_i - r, a_i + r[, f_i(t) = f(a_1, \dots, a_{i-1}, t, a_{i+1}, \dots, a_n) \leq f(a).$$

Par conséquent, sa dérivée s'annule en a_i : $\partial_i f(a) = 0$.

Ceci étant vrai pour tout $i \in \llbracket 1, n \rrbracket$, on en déduit que $\nabla f(a) = 0$. □

⚠ Réciproque fautive ! Par exemple considérons la fonction $f : \mathbf{R}^2 \rightarrow \mathbf{R}$ définie par $f(x, y) = (x - y)^2 + x^3 + y^3$.

Alors f est de classe \mathcal{C}^1 sur \mathbf{R}^2 , et

$$\nabla f(x, y) = (2(x - y) + 3x^2, -2(x - y) + 3y^2).$$

Ainsi, (x, y) est un point critique de f si et seulement si

$$\begin{cases} 2(x - y) + 3x^2 = 0 \\ -2(x - y) + 3y^2 = 0 \end{cases} \Rightarrow 3x^2 + 3y^2 = 0 \Rightarrow x = y = 0.$$

Ainsi, $(0, 0)$ est l'unique point critique de f , et on a $f(0, 0) = 0$.

Pourtant, $f(x, x) = 2x^3$, de sorte que toute boule ouverte de centre $(0, 0)$ contiendra des couples (x, x) avec $x < 0$ et donc $f(x, x) < 0$ et des couples (x, x) avec $x > 0$ et donc $f(x, x) > 0$. Donc f n'a pas d'extremum local en $(0, 0)$.

Remarque
 Calculer $f(x, x)$ revient à chercher l'image d'un point situé sur la droite d'équation $y = x$.

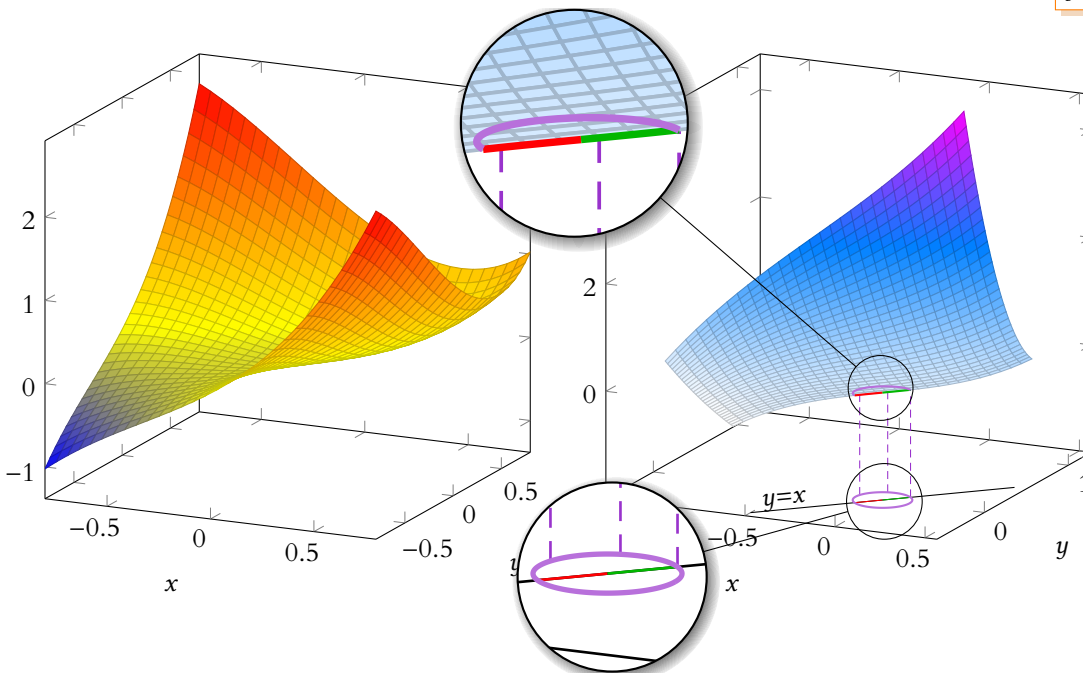


FIGURE 17.13 – La fonction $f(x, y) = (x - y)^2 + x^3 + y^3$. Le second dessin est un zoom sur le point critique. On y constate que dans toute boule de centre $(0, 0)$, des points de la droite d'équation $y = x$ ont une image positive et d'autres ont une image négative.

Définition 17.29 – Soit $f : \Omega \rightarrow \mathbf{R}$ une fonction de classe \mathcal{C}^1 sur un ouvert de \mathbf{R}^n . Un point critique a de f en lequel f n'atteint ni maximum ni minimum est appelé point selle (ou point col) de f .

Terminologie

L'exemple précédent montre qu'un point selle ne ressemble pas toujours à une selle. Nous verrons plus tard que c'est quand même souvent le cas.

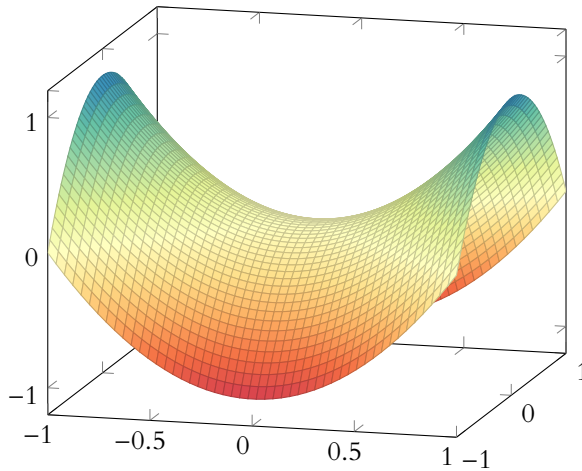


FIGURE 17.14 – L'exemple typique de point selle : la fonction $f(x, y) = x^2 - y^2$.

17.4 FONCTIONS DE CLASSE \mathcal{C}^2

17.4.1 Dérivées partielles d'ordre 2

Définition 17.30 – Soit f une fonction de classe \mathcal{C}^1 sur un ouvert $\Omega \subset \mathbf{R}^n$. Soient $(i, j) \in \llbracket 1, n \rrbracket^2$. La fonction $\partial_j f$ est une fonction continue de Ω dans \mathbf{R} . Si elle admet une dérivée partielle par rapport à la i -ième variable, on note

$$\partial_{i,j}^2 f = \partial_i(\partial_j f)$$

cette dérivée partielle, et on l'appelle dérivée partielle d'ordre 2 d'indice (i, j) .

Exemples 17.31

- Soit $f :]0, +\infty[\times]0, +\infty[$ définie par $f(x, y) = \ln(x) + \ln(y) - xy^2$. Alors f est de classe \mathcal{C}^1 et

$$\partial_1 f(x, y) = \frac{1}{x} - y^2, \partial_2 f(x, y) = \frac{1}{y} - 2xy.$$

Alors f admet des dérivées partielles d'ordre 2 et

$$\partial_{1,1}^2 f(x, y) = -\frac{1}{x^2}, \partial_{2,2}^2 f(x, y) = -\frac{1}{y^2} - 2x, \partial_{1,2}^2 f(x, y) = -2y, \partial_{2,1}^2 f(x, y) = -2y.$$

- $f(x, y) = \begin{cases} \frac{xy^3}{x^2 + y^2} & \text{si } (x, y) \neq (0, 0) \\ 0 & \text{si } (x, y) = (0, 0) \end{cases}$

Pour $(x, y) \neq (0, 0)$, on a

$$\partial_1 f(x, y) = \frac{y^3(x^2 + y^2) - xy^3 2x}{(x^2 + y^2)^2} = \frac{y^3(y^2 - x^2)}{(x^2 + y^2)^2}.$$



De même, on a

$$\partial_2 f(x, y) = \frac{3xy^2(x^2 + y^2) - xy^3(2y)}{(x^2 + y^2)^2} = \frac{xy^2(3x^2 + y^2)}{(x^2 + y^2)^2}.$$

Ensuite, par définition d'une dérivée partielle, on a

$$\partial_1 f(0, 0) = \lim_{x \rightarrow 0} \frac{f(x, 0) - f(0, 0)}{x} = \lim_{x \rightarrow 0} \frac{0}{x} = 0.$$

De même,

$$\partial_2 f(0, 0) = \lim_{y \rightarrow 0} \frac{f(0, y) - f(0, 0)}{y} = 0.$$

Enfin, on a

$$\partial_{1,2}^2 f(0, 0) = \lim_{x \rightarrow 0} \frac{\partial_2 f(x, 0) - \partial_2 f(0, 0)}{x} = 0$$

et

$$\partial_{2,1}^2 f(0, 0) = \lim_{y \rightarrow 0} \frac{\partial_1 f(0, y) - \partial_1 f(0, 0)}{y} = \lim_{y \rightarrow 0} \frac{y^5}{y(y^2)^2} = 1.$$

17.4.2 Fonctions de classe \mathcal{C}^2

Définition 17.32 – Soit f une fonction de classe \mathcal{C}^1 sur un ouvert $\Omega \subset \mathbf{R}^n$. Si f admet des dérivées partielles d'ordre 2 d'indice (i, j) , pour tout $(i, j) \in \llbracket 1, n \rrbracket^2$ et que ces dérivées partielles sont des fonctions continues de Ω dans \mathbf{R} , on dit que f est de classe \mathcal{C}^2 sur Ω .

Exemple 17.33

$f(x, y) = \ln(x) + \ln(y) - xy^2$ est de classe \mathcal{C}^2 sur $]0, +\infty[^2$, puisque ses dérivées partielles d'ordre 2, calculées plus haut, sont toutes continues sur $]0, +\infty[^2$.

Proposition 17.34 : Les fonctions polynomiales sont de classe \mathcal{C}^2 sur \mathbf{R}^n .

Démonstration. Remarquons que si f est une fonction polynomiale sur \mathbf{R}^n , alors pour tout $j \in \llbracket 1, n \rrbracket$, $\partial_j f$ est encore une fonction polynomiale sur \mathbf{R}^n .

En particulier, elle est \mathcal{C}^1 : $\forall i \in \llbracket 1, n \rrbracket$, $\partial_i(\partial_j f)$ existe et est continue.

Par conséquent, f admet des dérivées partielles d'ordre 2 d'indice (i, j) pour tout $(i, j) \in \llbracket 1, n \rrbracket^2$, et ces dérivées partielles sont continues, donc f est de classe \mathcal{C}^2 . \square

Proposition 17.35 (Opérations sur les fonctions \mathcal{C}^2) : Soient f, g deux fonctions de classe \mathcal{C}^2 sur un ouvert $\Omega \subset \mathbf{R}^n$. Alors

- $f + g$ est de classe \mathcal{C}^2 sur Ω .
- pour tout $\alpha \in \mathbf{R}$, αf est de classe \mathcal{C}^2 sur Ω .
- $f \times g$ est de classe \mathcal{C}^2 sur Ω .
- Si g ne s'annule pas sur Ω , alors $\frac{f}{g}$ est de classe \mathcal{C}^2 sur Ω .
- Si f prend ses valeurs dans un intervalle $I \subset \mathbf{R}$, et si $\varphi : I \rightarrow \mathbf{R}$ est une fonction de classe \mathcal{C}^2 , alors $\varphi \circ f : (x_1, \dots, x_n) \mapsto \varphi(f(x_1, \dots, x_n))$ est une fonction de classe \mathcal{C}^2 sur Ω .

Exemple 17.36

La fonction $f : (x_1, \dots, x_n) \mapsto e^{-(x_1^2 + \dots + x_n^2)}$ est de classe \mathcal{C}^2 car $(x_1, \dots, x_n) \mapsto x_1^2 + \dots + x_n^2$ l'est en tant que fonction polynomiale, et car la fonction $t \mapsto e^{-t}$ est une fonction de classe \mathcal{C}^2 de \mathbf{R} dans \mathbf{R} .

Définition 17.37 – Soit f une fonction de classe \mathcal{C}^2 sur Ω , et soit $x \in \Omega$. On appelle matrice hessienne² de f en x la matrice

$$\nabla^2(f)(x) = (\partial_{i,j}^2 f(x))_{1 \leq i, j \leq n} \in \mathcal{M}_n(\mathbf{R}).$$

² Du nom de Ludwig Otto HESSE, mathématicien allemand, 1811-1874.

⚠ La hessienne $\nabla^2(f)(x)$ est une matrice carrée, à ne pas confondre avec le gradient $\nabla f(x)$, qui lui est un vecteur de \mathbf{R}^n .

Comme pour le gradient, il existe une matrice hessienne en **chaque point**, et ces différentes hessiennes ne sont a priori par reliées.

Exemple 17.38

- Si $f : \mathbf{R}^2 \rightarrow \mathbf{R}$ est \mathcal{C}^2 , alors

$$\forall x \in \mathbf{R}^2, \nabla^2(f)(x) = \begin{pmatrix} \partial_{1,1}^2 f(x) & \partial_{1,2}^2 f(x) \\ \partial_{2,1}^2 f(x) & \partial_{2,2}^2 f(x) \end{pmatrix}$$

- Soit $f : \mathbf{R}^n \rightarrow \mathbf{R}$ définie par

$$f(x_1, \dots, x_n) = \sum_{k=1}^n x_k^2 + \left(\sum_{k=1}^n x_k \right)^2 - \sum_{k=1}^n x_k.$$

Alors f est de classe \mathcal{C}^2 car polynomiale, et on a

$$\forall i \in \llbracket 1, n \rrbracket, \partial_i f(x_1, \dots, x_n) = 2x_i + 2 \sum_{k=1}^n x_k - 1.$$

et on en déduit que

$$\partial_{i,j}^2 f(x_1, \dots, x_n) = \begin{cases} 4 & \text{si } i = j \\ 2 & \text{si } i \neq j \end{cases}$$

Par conséquent, la matrice hessienne de f en un point (x_1, \dots, x_n) est

$$\nabla^2(f)(x_1, \dots, x_n) = \begin{pmatrix} 4 & 2 & \dots & 2 \\ 2 & 4 & & \vdots \\ \vdots & & \ddots & \vdots \\ 2 & \dots & 2 & 4 \end{pmatrix}.$$

Remarque

Ici on a un exemple de fonction dont la hessienne est la même en tout point.

17.4.3 Théorème de Schwarz

Proposition 17.39 (Théorème de Schwarz³) : Soit f une fonction admettant des dérivées partielles d'ordre 2 $\partial_{i,j}^2 f$ et $\partial_{j,i}^2 f$ continues en x . Alors

$$\partial_{i,j}^2 f(x) = \partial_{j,i}^2 f(x).$$

En particulier, si f est de classe \mathcal{C}^2 sur un ouvert Ω de \mathbf{R}^n , alors

$$\forall x \in \Omega, \forall (i, j) \in \llbracket 1, n \rrbracket^2, \partial_{i,j}^2 f(x) = \partial_{j,i}^2 f(x).$$

³ Hermann SCHWARZ, mathématicien allemand, 1843-1921. C'est le même que celui de Cauchy-Schwarz.

Démonstration. Admis. □

Exemple 17.40

$f(x, y) = \begin{cases} \frac{xy^3}{x^2+y^2} & \text{si } (x, y) \neq (0, 0) \\ 0 & \text{sinon} \end{cases}$ n'est pas de classe \mathcal{C}^2 sur \mathbf{R}^2 car on a prouvé que $\partial_{2,1}^2 f(0, 0) \neq \partial_{1,2}^2 f(0, 0)$.
 En revanche, elle l'est sur $\mathbf{R}^2 - \{(0, 0)\}$ qui est un ouvert de \mathbf{R}^2 .

Corollaire 17.41 – Si f est une fonction de classe \mathcal{C}^2 sur Ω , alors pour tout $x \in \Omega$, $\nabla^2(f)(x)$ est une matrice symétrique.

Démonstration. Pour tout $(i, j) \in \llbracket 1, n \rrbracket^2$, le coefficient sur la $i^{\text{ème}}$ ligne et $j^{\text{ème}}$ colonne de $\nabla^2(f)(x)$ est $\partial_{i,j}^2 f(x) = \partial_{j,i}^2 f(x)$, et donc est égal au coefficient situé à la $j^{\text{ème}}$ ligne et $i^{\text{ème}}$ colonne, donc $\nabla^2(f)(x)$ est symétrique. □

Corollaire 17.42 – Si f est de classe \mathcal{C}^2 sur Ω , alors pour tout $x \in \Omega$, $\nabla^2(f)(x)$ est diagonalisable.

17.5 NATURE LOCALE DES POINTS CRITIQUES

17.5.1 Quelques rappels et compléments sur les dérivées secondes

Soit $f : \mathbf{R} \rightarrow \mathbf{R}$ une fonction de classe \mathcal{C}^2 . Alors nous savons que le signe de la dérivée seconde indique la convexité de la courbe : si elle est positive, la fonction est convexe, et sinon la fonction est concave.

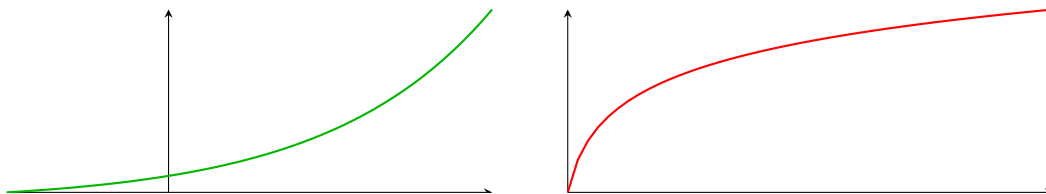


FIGURE 17.15 – Une fonction convexe et une fonction concave.

Le signe de cette dérivée seconde est susceptible de changer, dans ce cas, le signe de $f''(x)$ indique la convexité de la courbe au voisinage de x . Par exemple, sur la fonction suivante, les parties en vert correspondent à une dérivée seconde positive (la courbe y est convexe) alors que les parties en rouge indiquent une dérivée seconde négative (la courbe y est concave).

D'autre part, si a est un point critique de f (i.e. si $f'(a) = 0$) avec $f''(a) \neq 0$, alors, par la formule de Taylor-Young, on a, au voisinage de a ,

$$f(x) - f(a) = \underbrace{f'(a)}_{=0}(x - a) + \frac{f''(a)}{2}(x - a)^2 + o((x - a)^2) \underset{x \rightarrow a}{\sim} \frac{f''(a)}{2}(x - a)^2.$$

Si $f''(a) > 0$, alors $\frac{f''(a)}{2}(x - a)^2$ est positif. Et donc, au voisinage⁴ de a , $f(x) - f(a) \geq 0$ et donc $f(x) \geq f(a)$: f possède en a un minimum local.

De même, si $f''(a) < 0$, alors f possède en a un maximum local.

Notons que nous n'avons rien dit à propos du cas où $f''(a) = 0$. La formule de Taylor à l'ordre 2 ne suffit pas à conclure et il faudrait par exemple utiliser la formule de Taylor à

⁴ C'est-à-dire pour x suffisamment proche de a : dans un petit segment de la forme $]a - r, a + r[$.

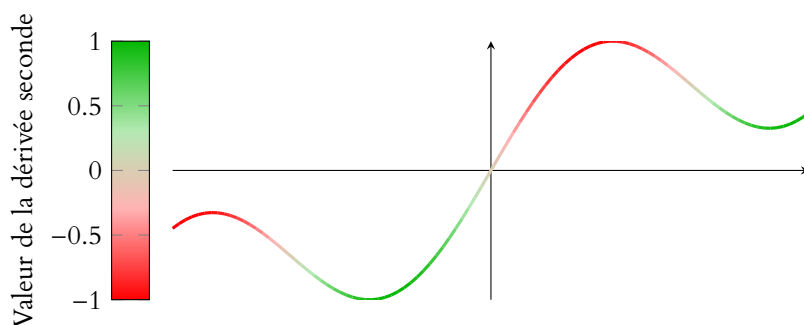


FIGURE 17.16 – Convexité et signe de la dérivée seconde : plus la couleur est proche du rouge, plus la dérivée seconde est négative, et donc plus la concavité est prononcée. À l'inverse, plus la dérivée seconde est positive, plus la convexité est prononcée.

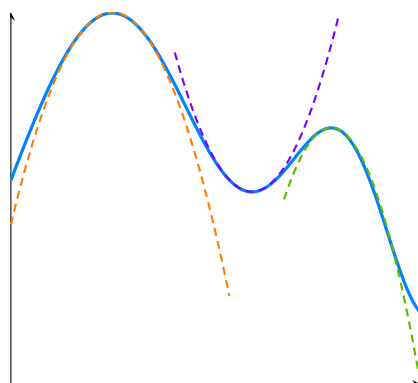
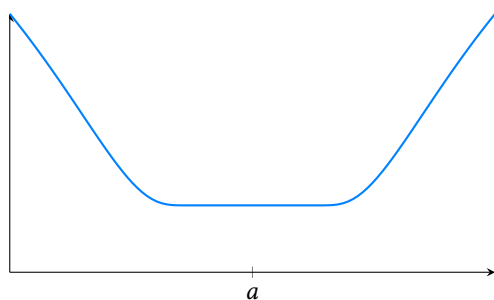


FIGURE 17.17 – Pour chaque point critique, on a tracé l'équivalent donné par la formule de Taylor à l'ordre 2. On constate que «localement» l'approximation par un polynôme de degré 2 est très bonne, et le «sens» de ce polynôme de degré 2 permet de déterminer la nature local du point critique.

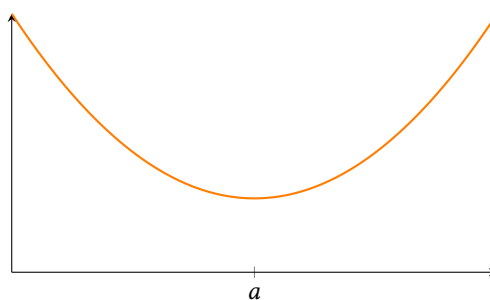
un ordre supérieur⁵ pour décider si oui ou non il y a un extremum local.

Plus précisément si $f''(a) > 0$, on a même un minimum local strict : dans un voisinage de a , pour tout $x \neq a$, on a $f(x) > f(a)$.

⁵ Jusqu'à trouver une dérivée qui ne s'annule pas : si la dérivée troisième est non nulle, on s'arrête, sinon on continue.



Un minimum non strict



Un minimum strict

Proposition 17.43 : Soit $f : I \rightarrow \mathbb{R}$ une fonction de classe \mathcal{C}^2 sur un intervalle I . Si a est un point critique de f tel que $f''(a) > 0$ (respectivement $f''(a) < 0$), alors il existe $r > 0$ tel que pour tout $x \in]a - r, a + r[\cap I$, $x \neq a$, $f(x) > f(a)$ (resp. $f(x) < f(a)$).

Démonstration. Traitons le cas où $f''(a) > 0$, la preuve étant similaire si $f''(a) < 0$. Puisque f est \mathcal{C}^2 , f'' est continue, et donc il existe $r > 0$ tel que pour $t \in]a - r, a + r[\cap I$, $f''(t) > 0$.

Par la formule de Taylor avec reste intégral à l'ordre 1, on a alors, pour $x \in]a - r, a + r[\cap I$,

$$f(x) = f(a) + \underbrace{f'(a)}_{=0}(x - a) + \int_a^x f''(t)(x - t) dt.$$

- Si $x \in]a, a + r[\cap I$, alors pour tout $t \in]a, x[$, on a $f''(t)(x - t) > 0$ et donc⁶ $\int_a^x f''(t)(x - t) dt > 0$.

Par conséquent, $f(x) > f(a)$.

- Si $x \in]a - r, a[\cap I$, alors

$$\int_a^x f''(t)(x - t) dt = - \int_x^a f''(t)(x - t) dt = \int_x^a f''(t)(t - x) dt.$$

Or, pour tout $t \in]x, a[$, on a $f''(t)(t - x) > 0$ et donc $\int_x^a f''(t)(t - x) dt > 0$.

Et ainsi, pour tout $x \in]a - r, a + r[$, $f(x) > f(a)$. □

⁶ L'intégrale d'une fonction positive et continue est positive, et n'est nulle que si cette fonction est nulle.

Bornes

Lorsqu'on souhaite utiliser la positivité de l'intégrale, il faut s'assurer que les bornes sont dans le bon sens. Si ces bornes sont des nombres, c'est généralement évident, mais dans les cas comme ici où ce sont deux inconnues, il faut être vigilant. Ici, on a changé le sens des bornes afin de les remettre dans le bon sens.

17.5.2 Dérivées directionnelles secondes

Définition 17.44 – Soit f une fonction de classe \mathcal{C}^2 sur Ω . On note alors q_x la forme quadratique associée à la matrice hessienne $\nabla^2(f)(x)$.

Remarquons qu'alors, on a

$$q_x(h) = (h_1 \quad \dots \quad h_n) \nabla^2(f)(x) \begin{pmatrix} h_1 \\ \vdots \\ h_n \end{pmatrix} = (h_1 \quad \dots \quad h_n) \begin{pmatrix} \sum_{j=1}^n \partial_{1,j}^2 f(x) h_j \\ \vdots \\ \sum_{j=1}^n \partial_{n,j}^2 f(x) h_j \end{pmatrix} = \sum_{i=1}^n \sum_{j=1}^n \partial_{i,j}^2 f(x) h_i h_j.$$

Proposition 17.45 : Soit f une fonction de classe \mathcal{C}^2 sur Ω , soit $x \in \Omega$ et soit $h \in \mathbf{R}^n$ non nul.

Alors la fonction $g : t \mapsto f(x + th)$ est définie sur un voisinage de 0, sur lequel elle est deux fois dérivable et

$$g''(t) = q_{x+th}(h).$$

En particulier, $g''(0) = q_x(h)$.

Démonstration. Nous savons déjà que g est de classe \mathcal{C}^1 au voisinage de 0, et que sa dérivée est donnée par

$$g'(t) = \langle \nabla f(x + th), h \rangle = \sum_{i=1}^n h_i \partial_i f(x + th).$$

Mais $\partial_i f$ est de classe \mathcal{C}^1 (car f est \mathcal{C}^2) et donc $t \mapsto \partial_i f(x + th)$ est dérivable, de dérivée

$$\sum_{j=1}^n \partial_{j,i}^2 f(x + th) h_j.$$

On en déduit que g' est dérivable et que

$$g''(t) = \sum_{i=1}^n h_i \sum_{j=1}^n \partial_{j,i}^2 f(x + th) h_j = \sum_{i=1}^n \sum_{j=1}^n \partial_{i,j}^2 f(x + th) h_i h_j = q_{x+th}(h).$$

□

Définition 17.46 – Si f est de classe \mathcal{C}^2 sur Ω , si $x \in \Omega$ et si $h \in \mathbf{R}^n$ est un vecteur unitaire, on appelle dérivée directionnelle seconde de f en x dans la direction de h le nombre $q_x(h)$.

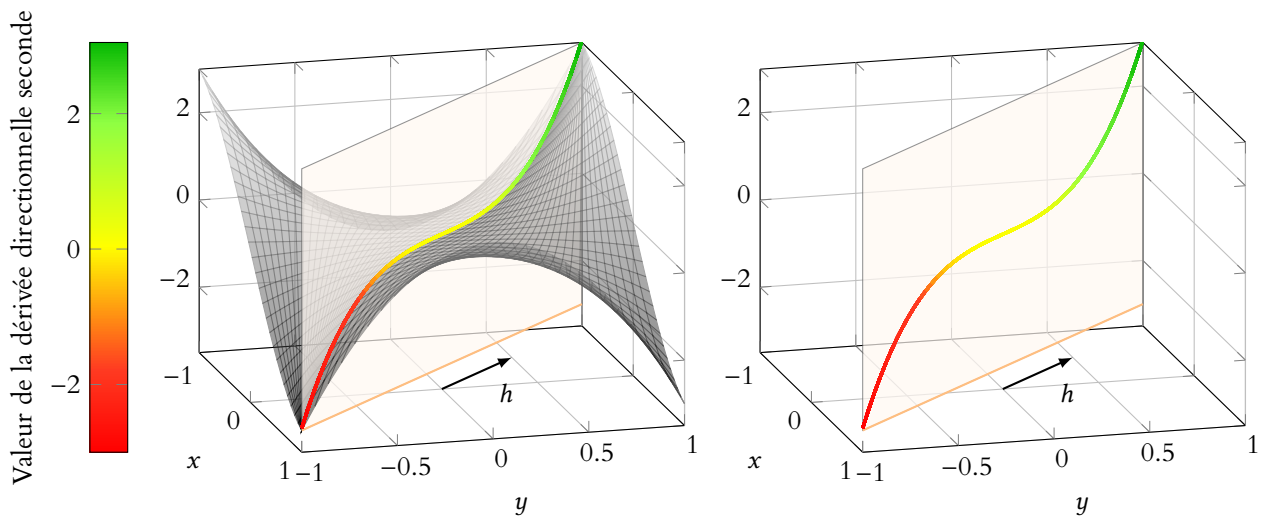


FIGURE 17.18 – Interprétation graphique de la dérivée directionnelle seconde dans la direction de $h = (-1, 1)$: son signe indique la convexité de la courbe de la fonction $t \mapsto f(a + th)$.

17.5.3 Formule de Taylor à l'ordre 2

Proposition 17.47 : Soit f une fonction de classe \mathcal{C}^2 sur un ouvert $\Omega \subset \mathbf{R}^n$. Alors pour tout $a \in \Omega$, il existe une fonction ε définie au voisinage de 0, continue en 0 et vérifiant $\varepsilon(0) = 0$ telle que pour tout $h \in \mathbf{R}^n$ «suffisamment petit»,

$$f(a + h) = f(a) + \langle \nabla f(a), h \rangle + \frac{1}{2} q_a(h) + \|h\|^2 \varepsilon(h).$$

Détails

Ω étant ouvert, il existe $r > 0$ tel que $B_o(a, r) \subset \Omega$. Par h «petit» on entend $\|h\| < r$, de sorte que $a + h \in B_o(a, r)$.

Démonstration. Admis. □

Remarques. • Pour $n = 1$, on retrouve la formule de Taylor usuelle à l'ordre 2 : le gradient est alors $f'(a)$ et la forme quadratique n'est rien d'autre que $h \mapsto f''(a)h^2$.

• Cette formule est à mettre en parallèle avec la formule de Taylor-Young à l'ordre 2 pour une fonction f de classe \mathcal{C}^2 :

$$f(a + h) = f(a) + f'(a)h + \frac{f''(a)}{2}h^2 + o_{h \rightarrow 0}(h^2).$$

Le terme $\langle \nabla f(a), h \rangle$ est à interpréter comme étant le terme de degré 1 : c'est une forme linéaire en h , qui s'écrit donc $\lambda_1 h_1 + \dots + \lambda_n h_n$, et $\frac{1}{2} q_a(h)$ est le terme de degré 2 : c'est une forme quadratique, qui s'écrit comme un «polynôme de degré 2» en les h_i .

Enfin, le terme $\|h\|^2 \varepsilon(h)$ est l'équivalent du $o(h^2)$.

• En particulier, si a est un point critique de f , alors $f(a + h) = f(a) + \frac{1}{2} q_a(h) + \|h\|^2 \varepsilon(h)$.

Cela revient à dire que pour h «petit»⁷, $f(a + h) - f(a)$ se comporte «comme» $q_a(h)$.

⁷ C'est-à-dire pour $a + h$ proche de a .

17.5.4 Condition suffisante d'existence d'extrema locaux.

Lemme 17.48. Soit A une matrice symétrique, et soit q_A la forme quadratique associée. Alors il existe deux réels α et β tels que

$$\forall x \in \mathbf{R}^n, \alpha \|x\|^2 \leq q_A(x) \leq \beta \|x\|^2.$$

De plus, on peut prendre

$$\alpha = \min\{q(x), \|x\| = 1\} \text{ et } \beta = \max\{q(x), \|x\| = 1\}.$$

Démonstration. La forme quadratique q_A est continue sur \mathbf{R}^n , car il s'agit d'une application polynomiale.

De plus, $B = \{x \in \mathbf{R}^n : \|x\| = 1\}$ est un fermé de \mathbf{R}^n car la fonction

$x \mapsto \|x\| = \sqrt{x_1^2 + \dots + x_n^2}$ y est continue.

Et B est borné car inclus dans la boule fermée de centre $0_{\mathbf{R}^n}$ et de rayon 1.

Comme B est fermé bornée et q_A continue, q_A admet sur B un minimum α et un maximum β .

Autrement dit, si $\|x\| \leq 1$, on a $\alpha \leq q_A(x) \leq \beta$.

Soit alors $x \in \mathbf{R}^n$ non nul. Le vecteur $\frac{x}{\|x\|}$ est de norme 1, donc dans B . Et donc

$$\alpha \leq q_A\left(\frac{x}{\|x\|}\right) \leq \beta.$$


Mais on sait également que $q_A(\lambda x) = \lambda^2 q_A(x)$, et donc pour $\lambda = \frac{1}{\|x\|}$, il vient

$$\alpha \leq \frac{1}{\|x\|^2} q_A(x) \leq \beta \Leftrightarrow \alpha \|x\|^2 \leq q_A(x) \leq \beta \|x\|^2.$$

L'inégalité est trivialement vérifiée pour $x = 0$ car $q_A(0) = 0$. □

Théorème 17.49 : Soit f une fonction de classe \mathcal{C}^2 sur Ω , et soit $x = (x_1, \dots, x_n)$ un point critique de f .

- Si $\text{Spec}(\nabla^2(f)(x)) \subset \mathbf{R}_+^*$, alors f admet un minimum local en x .
- Si $\text{Spec}(\nabla^2(f)(x)) \subset \mathbf{R}_-^*$, alors f admet un maximum local en x .
- Si $\nabla^2(f)(x)$ admet une valeur propre strictement positive et une valeur propre strictement négative, alors x est un point selle de f .

 *Démonstration.* Supposons que toutes les valeurs propres de f soient strictement positives. Alors x étant un point critique de f , la formule de Taylor s'écrit, pour tout h non nul,

$$f(x+h) = f(x) + \underbrace{\langle \nabla f(x), h \rangle}_{=0} + \frac{1}{2} q_x(h) + \|h\|^2 \varepsilon(h) = f(x) + \|h\|^2 \left(\frac{1}{2} \frac{1}{\|h\|^2} q_x(h) + \varepsilon(h) \right).$$

Puisque toutes les valeurs propres de $\nabla^2 f(x)$ sont strictement positives, la forme quadratique q_x ne prend que des valeurs strictement positives : $\forall u \in \mathbf{R}^n \setminus \{0\}, q_x(u) > 0$.

Et donc en particulier, $\alpha = \min_{\|u\|=1} q_x(u) > 0$.

Ainsi⁸, $\forall h \in \mathbf{R}^n$ non nul,

$$\frac{1}{2} \frac{1}{\|h\|^2} q_x(h) + \varepsilon(h) \geq \frac{\alpha}{2} + \varepsilon(h)$$

Mais puisque $\varepsilon(0) = 0$ et car ε est continue en 0, il existe $r > 0$ tel que pour $h \in B_o(0, r)$, $|\varepsilon(h)| \leq \frac{\alpha}{2}$.

Et alors, pour $y \in B_o(x, r)$, différent de x , on a $y = x + h$, où $h = y - x \in B_o(0, r)$, de sorte

$$f(y) = f(x+h) = f(x) + \frac{1}{\|h\|^2} \left(\frac{1}{2} q_x(h) + \varepsilon(h) \right) \geq f(x) + \frac{\alpha}{2} - \frac{\alpha}{2} \geq f(x).$$

Ceci prouve donc que f admet un minimum local en x .

On raisonnerait de même si toutes les valeurs propres de $\nabla^2(f)(x)$ sont négatives, prouvant que pour tout y dans une petite boule de centre x , $f(y) \leq f(x)$.

Si $\nabla^2 f(x)$ possède deux valeurs propres λ_1, λ_2 , avec $\lambda_1 > 0$ et $\lambda_2 < 0$.

Soient alors $h_1, h_2 \in \mathbf{R}^n$ tels que ${}^t h_1$ et ${}^t h_2$ soient deux vecteurs propres de $\nabla^2(f)(x)$ associés respectivement à λ_1 et λ_2 . Quitte à les diviser par leur norme, nous pouvons supposer ces vecteurs de norme 1.

Alors, la fonction $g_1 : t \mapsto f(x + th_1)$ est de classe \mathcal{C}^2 au voisinage de 0, et

$$g_1'(0) = \langle \nabla f(x), h_1 \rangle = 0 \text{ et } g_1''(0) = q_x(h_1) = h_1 \nabla^2 f(x) {}^t h_1 = h_1 \lambda_1 {}^t h_1 = \lambda_1 \|h_1\|^2.$$

Intuition

B est l'ensemble des points exactement à distance 1 de l'origine.
En dimension 2, c'est le cercle de centre 0 et de rayon 1.
En dimension 3 c'est la sphère de centre 0 et de rayon 1.

Analogie

Ce théorème est la généralisation, en dimension n de ce que nous avons dit plus tôt concernant le lien entre le signe de la dérivée seconde et la nature locale des points critiques pour les fonctions d'une variable.

⁸ C'est le lemme 17.48.

Transposée

Rappelons que si $h \in \mathbf{R}^n$, alors ${}^t h \in \mathcal{M}_{n,1}(\mathbf{R})$ et donc ${}^t h$ est bien susceptible d'être un vecteur propre de $\nabla^2 f(x)$.

Puisque $\lambda_1 > 0$, g_1 admet un minimum local en 0 : il existe $r_1 > 0$ tel que pour tout $t \in]-r_1, r_1[$, $g_1(t) \geq g_1(0) \Leftrightarrow f(x + th_1) \geq f(x)$. Et même⁹, pour t non nul : $f(x + th_1) > f(x)$.

⁹ C'est la proposition 17.43 sur les extrema locaux stricts des fonctions d'une variable.

De même, $t \mapsto f(x + th_2)$ admet un maximum local en 0 : il existe $r_2 > 0$ tel que pour tout $t \in]-r_2, r_2[$, $f(x + th_2) \leq f(x)$. Et même, pour $t \neq 0$, $f(x + th_2) < f(x)$.

Pour $r > 0$, soit $u \neq 0$ tel que $|u| < \min(r, r_1, r_2)$.

Alors $x + uh_1 \in B_o(x, r)$ car $\|x + uh_1 - x\| = \|uh_1\| = |u| < r$. Mais puisque $|u| < r_1$, il vient $f(x + uh_1) > f(x)$.

Autrement dit, dans toute boule centrée en x , f prend des valeurs strictement supérieures à $f(x)$, et donc f n'admet pas de maximum local en x .

De même, $x + uh_2 \in B_o(x, r)$ et $f(x + uh_2) < f(x)$, donc f n'admet pas de minimum local en x .

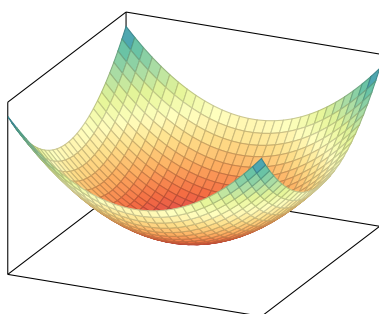
Par conséquent, f possède un point selle en x . \square

Remarques. • La formule de Taylor à l'ordre 2 nous dit qu'au voisinage d'un point critique a , on a $f(x) \approx f(a) + \frac{1}{2}q_a(x - a)$.

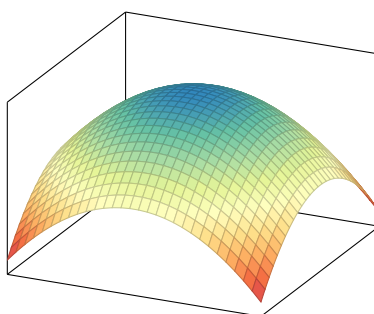
Il s'agit de l'analogue en dimension n de l'approximation locale d'une fonction d'une variable par une parabole.

Lorsque $n = 2$, le terme $f(a) + \frac{1}{2}q_a(x - a)$ possède l'allure suivante, suivant le signe des valeurs propres de $\nabla^2 f(a)$ (et donc suivant le signe de la forme quadratique q_a) :

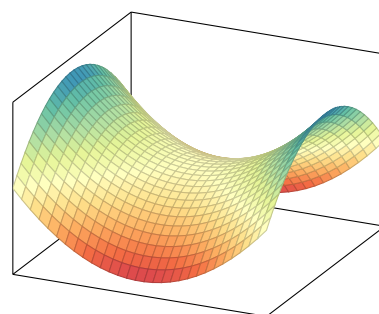
Valeurs propres positives



Valeurs propres négatives



Valeurs propres de signes opposés



• Si toutes les valeurs propres de $\nabla^2 f(x)$ sont de même signe, et que l'une d'entre elles est nulle, alors on ne peut pas conclure.

⚠ «On ne peut pas conclure» signifie juste que l'étude des valeurs propres de $\nabla^2(f)(x)$ ne suffit pas à conclure, mais s'il s'agit de déterminer la nature d'un point critique, alors on cherchera d'autres moyens de conclure !

• Notons que dans le cas où $\nabla^2(f)(x)$ possède deux valeurs propres de signe contraire, les vecteurs propres associés à ces valeurs propres nous donnent deux directions dans lesquelles f a tendance à augmenter et à diminuer.

De plus, ces deux directions sont orthogonales car les sous-espaces propres d'une matrice symétrique sont orthogonaux.

Analogie

Ce cas est l'analogue du cas où la dérivée seconde s'annule pour les fonctions d'une variable : la dérivée seconde ne suffit pas à déterminer la nature du point critique.

Méthode : Recherche d'extrema locaux

Pour déterminer les extrema locaux d'une fonction f :

1. S'assurer que l'ensemble de définition est un ouvert (indication généralement fournie par l'énoncé) : si l'on n'est pas sur un ouvert, on ne fait pas de calcul différentiel (mais on sait que si l'on est sur un fermé borné, il existe un maximum et un minimum (globaux) dès que f est continue).
2. Vérifier que f est de classe \mathcal{C}^2 (si l'énoncé ne le dit pas déjà).
3. Calculer le gradient ∇f de f .
4. Chercher les points critiques de f .
5. Calculer les dérivées partielles secondes de f .
6. Pour **chaque point critique** x , écrire la matrice hessienne $\nabla^2 f(x)$, déterminer ses valeurs propres et identifier la nature du point critique x à l'aide du signe de ces valeurs propres.

Exemple 17.50

Soit $f(x, y) = x(x + 1)^2 - y^2$. Alors f est \mathcal{C}^2 car polynomiale et

$$\partial_1 f(x, y) = (x + 1)^2 + 2x(x + 1) = 3x^2 + 4x + 1 \text{ et } \partial_2 f(x, y) = -2y.$$

Les points critiques de f sont donc $(-1, 0)$ et $(-\frac{1}{3}, 0)$.

De plus, on a

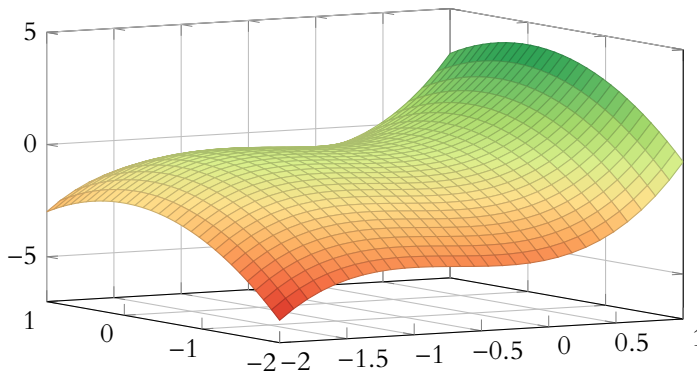
$$\partial_{1,1}^2 f(x, y) = 6x + 4, \partial_{2,2}^2 f(x, y) = -2 \text{ et } \partial_{1,2}^2 f(x, y) = \partial_{2,1}^2 f(x, y) = 0.$$

Ainsi, la hessienne au point $(-1, 0)$ est $\nabla^2 f(-1, 0) = \begin{pmatrix} -2 & 0 \\ 0 & -2 \end{pmatrix}$. Toutes ses valeurs propres sont strictement négatives, donc f possède un maximum local en $(-1, 0)$.

La hessienne au point $(-\frac{1}{3}, 0)$ est $\nabla^2 f(-\frac{1}{3}, 0) = \begin{pmatrix} 2 & 0 \\ 0 & -2 \end{pmatrix}$, qui possède deux valeurs propres de signe contraire. On en déduit que $(-\frac{1}{3}, 0)$ est un point selle de f .

De plus, on a alors $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ qui est un vecteur propre pour la valeur propre 2 et $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$ qui est un vecteur propre pour la valeur propre -2 .

Cela signifie que si depuis le point selle $(-\frac{1}{3}, 0)$, on se déplace dans la direction de $(1, 0)$, alors f augmente, et si l'on se déplace dans la direction de $(0, 1)$, f diminue.



Orthogonalité

Notons que ces deux vecteurs propres sont orthogonaux, car les sous-espaces propres de la hessienne sont deux à deux orthogonaux. Ce fait garantit que lorsqu'on a une valeur propre λ_1 strictement positive et une valeur propre λ_2 strictement négative, alors on a un **vrai** point selle : dans la direction associée à un vecteur propre pour λ_1 , f est «croissante». Et dans la direction d'un vecteur propre associée à λ_2 , qui est perpendiculaire à la première, f est «décroissante».

Exemple 17.51

Soit $f : \mathbf{R}^n \rightarrow \mathbf{R}$ définie par

$$f(x_1, \dots, x_n) = \sum_{k=1}^n x_k^2 + \left(\sum_{k=1}^n x_k \right)^2 - \sum_{k=1}^n x_k.$$

Nous avons déjà prouvé qu'en tout point $x \in \mathbf{R}^n$, $\nabla^2(f)(x) = 2J_n + 2I_n$, où $J_n \in \mathcal{M}_n(\mathbf{R})$ est la matrice dont tous les coefficients valent 1.

J_n est de rang 1, donc 0 est valeur propre de J_n et $\dim E_0(J_n) = n - 1$.

De plus, J_n est diagonalisable car symétrique, et elle admet donc une autre valeur propre λ , et on a

$$n = \text{tr}(A) = (n - 1) \times 0 + \lambda.$$

donc $\lambda = n$ est valeur propre de J_n , avec $\dim E_n(J_n) = 1$.

Ainsi, il existe P inversible telle que $J_n = P^{-1} \text{Diag}(\underbrace{0, \dots, 0}_{n-1 \text{ fois}}, n)P$. Donc

$$\begin{aligned} \nabla^2(f)(x) &= 2 \left(P^{-1} \text{Diag}(0, \dots, 0, n)P \right) + 2I_n \\ &= 2 \left(P^{-1} \text{Diag}(0, \dots, 0, n)P \right) + 2P^{-1}I_nP \end{aligned}$$

Astuce

Pour toute matrice inversible P ,

$$I_n = P^{-1}I_nP.$$

$$\begin{aligned}
 &= P^{-1}(2\text{Diag}(0, \dots, 0, n) + 2I_n)P \\
 &= P^{-1}\text{Diag}(\underbrace{2, \dots, 2}_{n-1 \text{ fois}}, 2n+2)P.
 \end{aligned}$$

Ainsi, les valeurs propres de $\nabla^2(f)(x)$ sont 2 et $2n+2$, qui sont strictement positives (et ce quel que soit x), de sorte qu'en tout point critique¹⁰, f admet un minimum local.

¹⁰ Une résolution de système prouverait qu'il existe un unique point critique.

17.6 À PROPOS DES EXTREMA GLOBAUX SUR UN FERMÉ BORNÉ

Sur un **fermé borné**, nous avons un théorème d'existence d'extrema globaux (et donc locaux).

Sur un **ouvert**, nous savons qu'un extremum local (et donc également un extremum global) est atteint en un point critique.

Malheureusement, une partie non vide de \mathbf{R}^n ne peut être à la fois fermée, bornée et ouverte...

Prenons l'exemple de la fonction $f(x, y) = x^2 + y^2$, définie sur $B_f(0, 1)$.

Il s'agit d'une fonction continue (car polynomiale), et $B_f(0, 1)$ est un fermé borné. Donc f atteint un maximum global (sur $B_f(0, 1)$), et un minimum global.

Sur la boule ouverte $B_o(0, 1)$ (qui est le plus grand ouvert que l'on puisse inclure dans $B_f(0, 1)$), f est de classe \mathcal{C}^1 , et

$$\nabla f(x, y) = (2x, 2y).$$

Ainsi, f admet sur $B_o(0, 1)$ un unique point critique en $(0, 0)$, dont il est facile de se convaincre qu'il s'agit du minimum de f .

Mais alors où est le maximum de f ? Il est dans $B_f(0, 1) - B_o(0, 1)$, c'est-à-dire dans

$$\{(x, y) \in \mathbf{R}^2 : \|(x, y)\| = 1\}.$$

Il est facile de voir que sur cet ensemble, f est constante égale à 1, et c'est là le maximum de f .

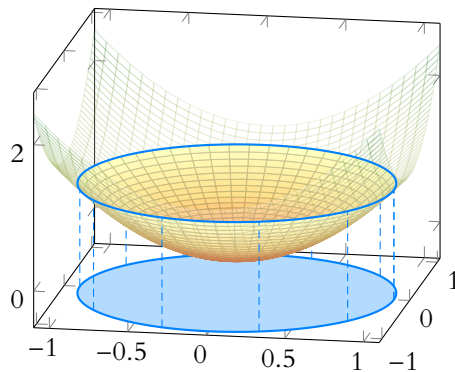


FIGURE 17.19 – Le maximum de $f(x, y) = x^2 + y^2$ sur $B_f(0, 1)$ n'est pas atteint en un point critique. En effet, la fonction admet un maximum uniquement car on considère sa restriction à un fermé borné, mais elle est définie sur un ouvert plus grand.

De manière générale, pour étudier les extrema globaux d'une fonction continue sur un fermé borné F :

- commencer par chercher le plus grand ouvert possible Ω inclus dans F (en général, il s'agit de remplacer les inégalités larges par des inégalités strictes).
- si f est \mathcal{C}^1 sur Ω , alors on peut chercher ses points critiques, et étudier¹¹ s'il s'agit ou non d'extrema de f .
- dans tous les cas, il faut ensuite étudier «à la main» le comportement de f sur le «bord» de F , c'est-à-dire sur le complémentaire de Ω dans F . Le chapitre 19 nous donnera des outils pour mener cette étude.

¹¹ Rappelons que déterminer si un point critique est un extremum global est toujours délicat et que nous ne disposons pas de méthode générale pour répondre à cette question.

EXERCICES DU CHAPITRE 17

EXERCICE 17.1 Déterminer si les ensembles suivants sont ouverts, fermés, bornés. On pourra s'aider d'un dessin lorsque c'est possible (c'est-à-dire pour les sous-ensembles de \mathbf{R}^2).

1. $A = \{(x, y) \in \mathbf{R}^2 : x^2 + y^2 \leq 3\}$
2. $B = \{(x, y, z) \in \mathbf{R}^3 : 1 < ze^{x+y} < 2\}$
3. $C = \{(x, y, z) \in \mathbf{R}^3 : xyz \neq 1\}$
4. $D = \{(x, y) \in \mathbf{R}^2 : x^2 - 1 \leq y \leq 1 - x^2\}$
5. $F = \{(x, y) \in \mathbf{R}^2 : |x + 2y| = 1 \text{ ou } |y + 2x| \geq 4\}$.

EXERCICE 17.2 Soit f la fonction définie sur \mathbf{R}^2 par $\forall (x, y) \in \mathbf{R}^2, f(x, y) = x^2(1 + y)^3 + y^2$.

- 1) Montrer que f possède un seul point critique.
- 2) Vérifier que si $\|(x, y)\| \leq 1$, alors $f(x, y) \geq 0$.
- 3) En déduire que f possède un minimum local en son point critique. Est-ce un minimum global ?

EXERCICE 17.3 Soit f la fonction définie sur l'ouvert $\Omega = \mathbf{R}^2 \setminus \{(0, 0)\}$ par $f(x, y) = (x^2 + y^2)^x$.

- 1) Justifier que f est de classe \mathcal{C}^1 sur Ω et calculer ses dérivées partielles.
- 2) Montrer que f admet 4 points critiques.
- 3) Soit g la fonction définie sur \mathbf{R}_+^* par $g(x) = f(x, 0)$.
 - a) Montrer que pour tout $(x, y) \in \mathbf{R}_+^* \times \mathbf{R}, f(x, y) \geq g(x)$.
 - b) Dresser le tableau de variation de g .
 - c) En déduire que f admet un minimum local en $(e^{-1}, 0)$.
- 4) Soit $h : x \mapsto f(x, 1) - f(0, 1)$.
 - a) Montrer que $h(x)$ est du signe de x .
 - b) En déduire que f n'admet pas d'extremum local en $(0, 1)$.

EXERCICE 17.4 Soit f la fonction définie sur $\Omega =]0, 1[\times]0, 1[$ par $f(x, y) = \frac{1}{1-x} + \frac{1}{1-y} + \frac{1}{x+y}$.

- 1) Montrer que f est \mathcal{C}^2 sur l'ouvert Ω , et calculer ses dérivées partielles premières et secondes.
- 2) Déterminer les points critiques de f .
- 3) Montrer que f admet un extremum local sur Ω .

EXERCICE 17.5 Déterminer les extrema locaux des fonctions suivantes :

1. $f : \mathbf{R}^2 \rightarrow \mathbf{R}, f(x, y) = x^3 + xy + y^3$.
2. $g : (\mathbf{R}^*)^2 \rightarrow \mathbf{R}, g(x, y) = 4xy + \frac{1}{x} + \frac{1}{y}$.
3. $h : (\mathbf{R}_+^*)^2 \rightarrow \mathbf{R}, h(x, y) = x^2 + y^2 + \frac{1}{x+y}$.

EXERCICE 17.6 Soit $f : \mathbf{R}^3 \rightarrow \mathbf{R}$ définie par $f(x, y, z) = x^3 + y^2 - xy + z^2$.

- 1) Déterminer les points critiques de f .
- 2) En déduire les extrema locaux de f . Sont-ils globaux ?

EXERCICE 17.7 Soit $f(x, y) = x^2 + y^2 + 2xy + xy^3$.

- 1) Montrer que f est de classe \mathcal{C}^2 sur \mathbf{R}^2 et qu'elle admet un unique point critique.
- 2) Déterminer les valeurs propres de la hessienne en ce point critique. Peut-on conclure quant à la nature de ce point critique ?
- 3) Étudier les signes de $f(x, x)$ et $f(x, -x)$, et conclure.

EXERCICE 17.8 (Ecrimage 2016)

Soit $\mathcal{D} = \{(x, y) \in \mathbf{R}^2 : x < y\}$, et soit f la fonction définie sur \mathcal{D} par $f(x, y) = x^2 + y^2 - \ln(y - x)$.

- 1) Justifier que f est \mathcal{C}^2 sur \mathcal{D} et calculer ses dérivées partielles premières et secondes.
- 2) Montrer que f admet un unique point critique sur \mathcal{D} .
- 3) Déterminer la nature locale de ce point critique.

EXERCICE 17.9 Soit f la fonction définie sur \mathbf{R}^3 par $f(x, y, z) = x^2 + y^2 + z^2 - 2xyz$.

- 1) Justifier que f est de classe \mathcal{C}^2 , et calculer ses dérivées partielles d'ordre 1 et 2.
- 2) Montrer que f admet exactement cinq points critiques, dont le point $(0, 0, 0)$.

- 3) Déterminer la matrice hessienne de f en $(0, 0, 0)$, et en déduire que f possède un minimum local en $(0, 0, 0)$. Est-ce un minimum global de f ?
- 4) Pour chacun des autres points critiques, vérifier que 4 est valeur propre de la matrice hessienne, et déterminer si f admet ou non un extremum local en ce point.

EXERCICE 17.10 Soit f la fonction définie sur \mathbf{R}^n par $f(x_1, \dots, x_n) = \sum_{k=1}^n x_k^2 + \left(1 - \sum_{k=1}^n x_k\right)^2$.

- 1) Montrer que f est de classe \mathcal{C}^2 sur \mathbf{R}^n et calculer ses dérivées partielles d'ordre 1 et 2.
- 2) Montrer que f admet un unique point critique $a = (a_1, \dots, a_n)$. On note $A = \nabla^2 f(a_1, \dots, a_n) \in \mathcal{M}_n(\mathbf{R})$.
- 3) Calculer $\text{rg}(A - 2I_n)$, puis déterminer les valeurs propres de A . En déduire que f admet un extremum local en (a_1, \dots, a_n) .
- 4) (★) Soit $x = (x_1, \dots, x_n) \in \mathbf{R}^n$ et posons $h = x - a$.
En appliquant la formule de Taylor avec reste intégral à la fonction $g : t \mapsto f(a + th)$, montrer que $f(x) \geq f(a)$. En déduire que f possède un extremum global.

AD

EXERCICE 17.11 Soit f la fonction définie sur $(\mathbf{R}_+^*)^2$ par $f(x, y) = x \ln y - y \ln x$.

- 1) Montrer que f est \mathcal{C}^1 et déterminer son gradient.
- 2) Soit φ la fonction définie sur \mathbf{R}_+^* par $\varphi(t) = \ln t - t + \frac{1}{t}$.
Montrer que si (x, y) est un point critique de f , alors $\varphi(\ln y) = 0$.
- 3) En étudiant le sens de variations de φ , montrer que f possède un unique point critique (x_0, y_0) .
- 4) Montrer que pour $\alpha \in [0, 1]$, $f(x_0 + \alpha, y_0 - \alpha) = -f(x_0 - \alpha, y_0 + \alpha)$. f admet-elle un extremum local sur $(\mathbf{R}_+^*)^2$?

AD

EXERCICE 17.12 Soit $\mathcal{D} = \{(x, y) \in \mathbf{R}^2 : x^2 - 1 \leq y \leq 1 - x^2\}$, et soit $\mathcal{D}_0 = \{(x, y) \in \mathbf{R}^2 : x^2 - 1 < y < 1 - x^2\}$.
On admet (cf exercice 1) que \mathcal{D} est fermé, que \mathcal{D}_0 est ouvert, et que tous les deux sont bornés.
Soit f la fonction définie sur \mathcal{D} par $f(x, y) = y^2 - x^2 y + x^2$.

- 1) Justifier que f possède un maximum M et un minimum m .
- 2) Déterminer les points critiques de f sur \mathcal{D}_0 .
- 3) Étudier les fonctions $x \mapsto f(x, x^2 - 1)$ et $x \mapsto f(x, 1 - x^2)$, et en déduire les valeurs de m et M .

AD

EXERCICE 17.13 Soit $\mathcal{D} = \{(x, y) \in \mathbf{R}^2 : x^2 + y^2 \leq 1\}$ et soit f la fonction définie sur \mathcal{D} par $f(x, y) = x^3 - 3x(1 + y^2)$.

- 1) Justifier que f admet un minimum m et un maximum M sur \mathcal{D} .
- 2) Montrer que sur $B_0(0, 1)$, f n'admet pas de point critique. Que peut-on en déduire à propos de m et M ?
- 3) En étudiant la fonction $t \mapsto f(\cos t, \sin t)$, déterminer les valeurs de m et M .

AD

EXERCICE 17.14 Soit $f : \mathbf{R}^2 \rightarrow \mathbf{R}$ définie par $f(x, y) = \left(1 + y + xy - \frac{x^2}{2}\right) e^y$.

- 1) Montrer que f est \mathcal{C}^2 sur \mathbf{R}^2 , déterminer ses dérivées partielles premières et secondes.
- 2) Montrer que f admet un unique point critique (α, β) .
- 3) Vérifier que la détermination des valeurs propres de $\nabla^2 f(\alpha, \beta)$ ne suffit pas à déterminer la nature de ce point critique.
- 4) Déterminer un vecteur propre u de $\nabla^2 f(\alpha, \beta)$ associé à la valeur propre 0.
- 5) En étudiant la fonction $t \mapsto f((\alpha, \beta) + tu)$, déterminer la nature du point critique (α, β) .

AD

EXERCICE 17.15 (Oral HEC 2018)

Soient X_1, \dots, X_n des variables aléatoires mutuellement indépendantes, de même loi, admettant chacune une espérance et un écart-type notés respectivement μ et $\sigma \neq 0$.

Pour tout $(x_1, \dots, x_n) \in \mathbf{R}^n$, on note : $f(x_1, \dots, x_n) = E\left(\left(\sum_{i=1}^n x_i X_i - \mu\right)^2\right)$.

- 1) a) Justifier l'égalité : $f(x_1, \dots, x_n) = \sigma^2 \sum_{i=1}^n x_i^2 + \mu^2 \left(\sum_{i=1}^n x_i - 1\right)^2$.
- b) Justifier, pour tout $(x_1, \dots, x_n) \in \mathbf{R}^n$, l'inégalité : $\left(\sum_{i=1}^n x_i\right)^2 \leq n \sum_{i=1}^n x_i^2$.
- c) En déduire le minimum de f sur l'ensemble $\mathcal{H} = \{(x_1, \dots, x_n) \in \mathbf{R}^n : x_1 + \dots + x_n = 1\}$.

TD

- 2) a) Justifier que f est de classe \mathcal{C}^2 sur \mathbf{R}^n et trouver son unique point critique a .
- b) Justifier que pour tout $h = (h_1, \dots, h_n) \in \mathbf{R}^n$: $f(a+h) = f(a) + 2 \int_0^1 (1-t) \left(\mu^2 \sum_{i=1}^n \sum_{j=1}^n h_i h_j + \sigma^2 \sum_{i=1}^n h_i^2 \right) dt$.
- c) En déduire que f admet un minimum global en a .

CORRECTION DES EXERCICES DU CHAPITRE 17

SOLUTION DE L'EXERCICE 17.1

Notons qu'il n'est possible de faire des dessins que dans le cas de sous-ensembles de \mathbb{R}^2 .

- Notons tout simplement que $A = B_f((0, 0), \sqrt{3})$, et donc il s'agit d'un fermé (car une boule fermée est fermée), borné (car une boule est toujours bornée).
- La fonction $(x, y, z) \mapsto ze^{x+y}$ est continue sur \mathbb{R}^3 . Donc $B_1 = \{(x, y, z) \in \mathbb{R}^3 : ze^{x+y} > 1\}$ et $B_2 = \{(x, y, z) \in \mathbb{R}^3 : ze^{x+y} < 2\}$ sont deux ouverts. Et alors $B = B_1 \cap B_2$ est ouvert car intersection de deux ouverts. B n'est pas borné, car quel que soit $x \in \mathbb{R}$, $(x, -x, 1.5) \in B$.
- La fonction $(x, y, z) \mapsto xyz$ est continue sur \mathbb{R}^3 car polynomiale, donc $\bar{C} = \{(x, y, z) \in \mathbb{R}^3 : xyz = 1\}$ est fermé. Et par conséquent, son complémentaire, qui est C , est ouvert. Il n'est pas borné car quel que soit $z \in \mathbb{R}$, $(0, 0, z) \in C$.
- Les fonctions $(x, y, z) \mapsto x + 2y$ et $(x, y, z) \mapsto z - 2x$ sont continues sur \mathbb{R}^3 car polynomiales. Donc $D_1 = \{(x, y, z) \in \mathbb{R}^3 : x + 2y < 1\}$ et $D_2 = \{(x, y, z) \in \mathbb{R}^3 : z - 2x > 3\}$ sont des ouverts de \mathbb{R}^3 . Et alors $D = D_1 \cap D_2$ est ouvert car intersection de deux ouverts. D n'est pas borné car quel que soit $x \in \mathbb{R}$, $(x, -\frac{x}{2}, 2x + 4) \in D$.
- Les fonctions $(x, y) \mapsto y - x^2 + 1$ et $(x, y) \mapsto y + x^2 - 1$ sont continues sur \mathbb{R}^2 car polynomiales, de sorte que les ensembles $\mathcal{D}_1 = \{(x, y) \in \mathbb{R}^2 : y - x^2 + 1 \geq 0\} = \{(x, y) \in \mathbb{R}^2 : x - 2 - 1 \leq y\}$ et $\mathcal{D}_2 = \{(x, y) \in \mathbb{R}^2 : y + x^2 - 1 \leq 0\} = \{(x, y) \in \mathbb{R}^2 : y \leq 1 - x^2\}$ sont fermés. Et donc $\mathcal{D} = \mathcal{D}_1 \cap \mathcal{D}_2$ est un fermé de \mathbb{R}^2 car intersection de deux fermés. De plus, pour $(x, y) \in \mathcal{D}$, on a

$$x^2 - 1 \leq y \leq 1 - x^2 \Rightarrow -1 \leq y \leq 1.$$

On a alors $x^2 \leq 1 + y \leq 2$, de sorte que $|x| \leq \sqrt{2}$.

On en déduit donc que \mathcal{D} est une partie bornée de \mathbb{R}^2 .

- Les fonctions $(x, y) \mapsto |x + 2y|$ et $(x, y) \mapsto |y + 2x|$ sont continues sur \mathbb{R}^2 , donc les deux ensembles

$$F_1 = \{(x, y) \in \mathbb{R}^2 : |x + 2y| = 1\} \text{ et } F_2 = \{(x, y) \in \mathbb{R}^2 : |y + 2x| \geq 4\}$$

sont des fermés de \mathbb{R}^2 . Alors $F = F_1 \cup F_2$ est un fermé de \mathbb{R}^2 .

F n'est pas borné car F_1 ne l'est pas : pour tout $y \in \mathbb{R}$, $(1 - 2y, y) \in F_1$.

SOLUTION DE L'EXERCICE 17.2

- La fonction f est \mathcal{C}^1 sur \mathbb{R}^2 car polynomiale. Son gradient est donné par

$$\nabla f(x, y) = (2x(1+y)^3, 3x^2(1+y)^2 + 2y).$$

Ainsi, (x, y) est un point critique de f si et seulement si

$$\begin{cases} 2x(1+y)^3 = 0 \\ 3x^2(1+y)^2 + 2y = 0 \end{cases}$$

Soit $x = 0$, et alors $y = 0$, soit $y = -1$, et alors la seconde équation n'est pas vérifiée.

Donc $(0, 0)$ est l'unique point critique de f .

- Si $\|(x, y)\| \leq 1$, alors $-1 \leq y \leq 1$ (car on a toujours $|y| \leq \|(x, y)\|$), et donc $0 \leq 1 + y \leq 2$. On en déduit aisément que $f(x, y) \geq 0$.
- Nous venons de prouver que pour $\|(x, y)\| \leq 1$, c'est-à-dire pour $(x, y) \in B_f((0, 0), 1)$, on a $f(x, y) \geq f(0, 0)$, donc f admet un minimum local en $(0, 0)$. Ce n'est pas un minimum global car $f(1, y) = (1 + y)^3 + y^2 \underset{y \rightarrow -\infty}{\sim} y^3 \underset{y \rightarrow -\infty}{\rightarrow} -\infty$.

SOLUTION DE L'EXERCICE 17.3

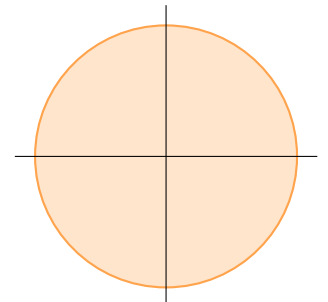


FIGURE 17.1— A est une boule fermée.

Rappel

Un et se traduit par une intersection alors qu'un ou se traduirait par une union.

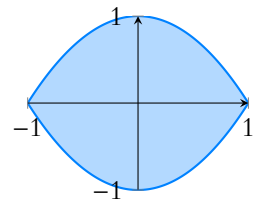


FIGURE 17.2— \mathcal{D} est la partie du plan située entre les deux paraboles d'équations $y = x^2 - 1$ et $y = 1 - x^2$.

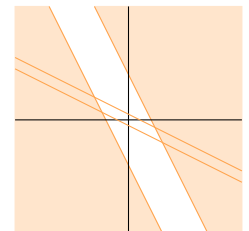


FIGURE 17.3— F est formé des points au dessus de la droite $2x + y = 4$, de ceux en dessous de $2x + y = -4$, et des droites $x + 2y = 1$ et $x + 2y = -1$.

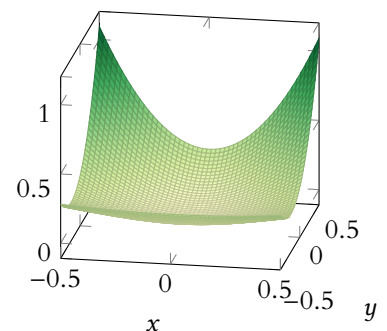


FIGURE 17.4— f admet un minimum local en $(0, 0)$.

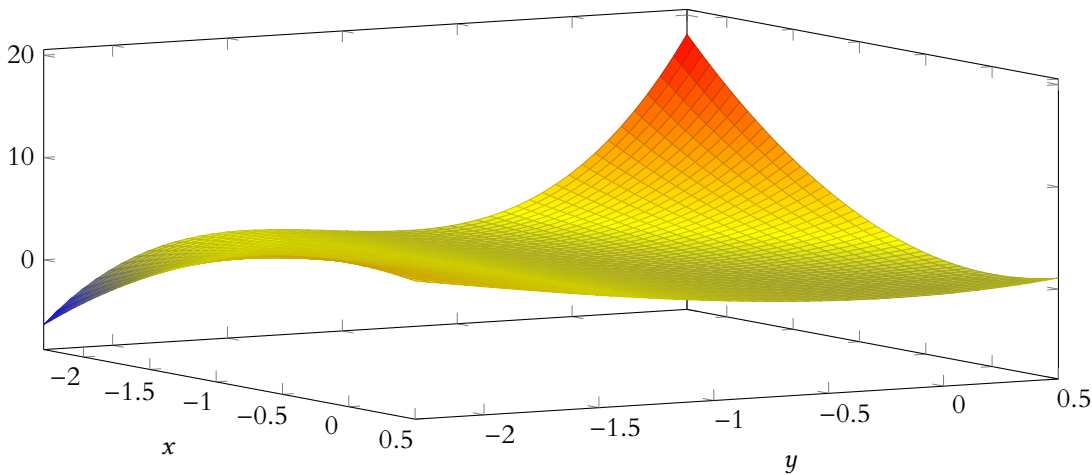


FIGURE 17.5 – Le minimum n'est pas global.

1. On a $f(x, y) = e^{x \ln(x^2 + y^2)}$.
 La fonction $(x, y) \mapsto x^2 + y^2$ est \mathcal{C}^1 sur Ω , car polynomiale, et à valeurs dans \mathbf{R}_+^* .
 Par composition avec la fonction $t \mapsto \ln t$, qui est \mathcal{C}^1 sur \mathbf{R}_+^* , $(x, y) \mapsto \ln(x^2 + y^2)$ est \mathcal{C}^1 sur Ω .
 La fonction $(x, y) \mapsto x$ est \mathcal{C}^1 sur Ω car polynomiale.
 Donc par produit de fonctions \mathcal{C}^1 , $(x, y) \mapsto x \ln(x^2 + y^2)$ est \mathcal{C}^1 sur Ω .
 Par composition avec la fonction exponentielle, \mathcal{C}^1 sur \mathbf{R} , f est une fonction de classe \mathcal{C}^1 sur Ω .

On a alors, pour $(x, y) \in \Omega$,

$$\partial_1 f(x, y) = \left(\ln(x^2 + y^2) + \frac{2x^2}{x^2 + y^2} \right) e^{x \ln(x^2 + y^2)}, \quad \partial_2 f(x, y) = \frac{2xy}{x^2 + y^2} e^{x \ln(x^2 + y^2)}.$$

2. Un point $(x, y) \in \Omega$ est un point critique de f si et seulement si

$$\begin{cases} \partial_1 f(x, y) = 0 \\ \partial_2 f(x, y) = 0 \end{cases} \Leftrightarrow \begin{cases} \ln(x^2 + y^2) + \frac{2x^2}{x^2 + y^2} = 0 \\ \frac{2xy}{x^2 + y^2} = 0 \end{cases}$$

De la deuxième équation il vient $x = 0$ ou $y = 0$.

• Si $x = 0$, alors la première équation devient $\ln(y^2) = 0 \Leftrightarrow y^2 = 1 \Leftrightarrow y = \pm 1$.
 Donc on a déjà deux points critiques : $(0, 1)$ et $(0, -1)$.

• Si $y = 0$, alors la première équation devient $\ln(x^2) + 2 = 0 \Leftrightarrow x^2 = e^{-2} \Leftrightarrow x = \pm e^{-1}$.
 Et donc nous obtenons deux autres points critiques qui sont $(e^{-1}, 0)$ et $(-e^{-1}, 0)$.

- 3.a. Puisque $y^2 \geq 0$, on a $x^2 + y^2 \geq x^2$, et donc pour $x > 0$, $(x^2 + y^2)^x \geq (x^2)^x = f(x, 0) = g(x)$.
 3.b. La fonction $g : x \mapsto (x^2)^x = e^{x \ln(x^2)} = e^{2x \ln(x)}$ est dérivable sur \mathbf{R}_+^* , et on a

$$\forall x \in \mathbf{R}_+^*, g'(x) = \left(\frac{2x}{x} + 2 \ln(x) \right) e^{2x \ln(x)} = 2(1 + \ln(x)) e^{2x \ln(x)}.$$

Et donc le tableau de variation de g est donné par :

x	0	e^{-1}	$+\infty$
$g'(x)$		-	0
			+
$g(x)$	1		$+\infty$
		$e^{-2e^{-1}}$	

Méthode

Lorsqu'on arrive à factoriser l'une des équations, on sait que l'un des facteurs du produit doit être nul. Le plus simple est alors d'arrêter d'écrire un système, de distinguer plusieurs cas (ici $x = 0$ ou $y = 0$), et dans chacun des cas de regarder ce que deviennent les autres équations (ici la première).

- 3.c. On constate que g admet un minimum¹ en e^{-1} .
 Autrement dit, pour tout $x > 0$, $g(x) \geq g(e^{-1})$.
 Et donc si $(x, y) \in \mathbb{R}_+^* \times \mathbb{R}$, alors $f(x, y) \geq g(x) \geq g(e^{-1}) = f(e^{-1}, 0)$.
 En particulier, si $r > 0$ est tel que $B_o((e^{-1}, 0)) \subset \mathbb{R}_+^* \times \mathbb{R}$, alors pour tout $(x, y) \in B_o((e^{-1}, 0), r)$,
 $f(x, y) \geq f(e^{-1}, 0)$. Ainsi, f possède un minimum local en $(e^{-1}, 0)$.

Remarque : une étude similaire de la fonction $x \mapsto f(-e^{-1}, 0)$ sur \mathbb{R}_-^* permettrait de prouver que f admet un maximum local en $(-e^{-1}, 0)$.

4. On a $h(x) = f(x, 1) - f(0, 1) = (x^2 + 1)^x - 1 = e^{x \ln(x^2 + 1)} - 1$.
 Et donc $h(x) \geq 0 \Leftrightarrow e^{x \ln(x^2 + 1)} \geq e^0 \Leftrightarrow \underbrace{x \ln(x^2 + 1)}_{\geq 0} \geq 0 \Leftrightarrow x \geq 0$.

Ceci prouve donc bien que $h(x)$ est du signe de x .

5. Soit $r > 0$. Alors pour $|x| < r$, $(x, 1) \in B_o((0, 1), r)$.
 Autrement dit, toute boule ouverte centrée en $(0, 1)$ contient des points de la forme $(x, 1)$, avec $x > 0$, mais aussi avec $x < 0$.
 Or, nous venons de prouver que si $x > 0$, alors $f(x, 1) > f(0, 1)$, et si $x < 0$, $f(x, 1) < f(0, 1)$.
 Donc dans toute boule centrée en $(0, 1)$, f prend des valeurs strictement supérieures à $f(0, 1)$ mais aussi des valeurs strictement inférieures à $f(0, 1)$: f n'admet pas d'extremum local en $(0, 1)$.

Remarque : on prouverait de manière analogue que f ne possède pas non plus d'extremum local en $(0, -1)$.

¹ Global.

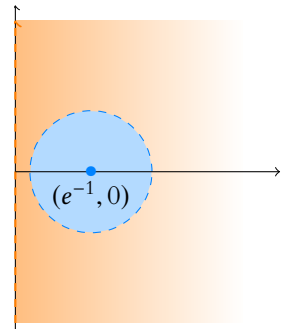


FIGURE 17.6— Puisque $\mathbb{R}_+^* \times \mathbb{R}$ est un ouvert, il contient au moins une boule centrée en $(e^{-1}, 0)$.

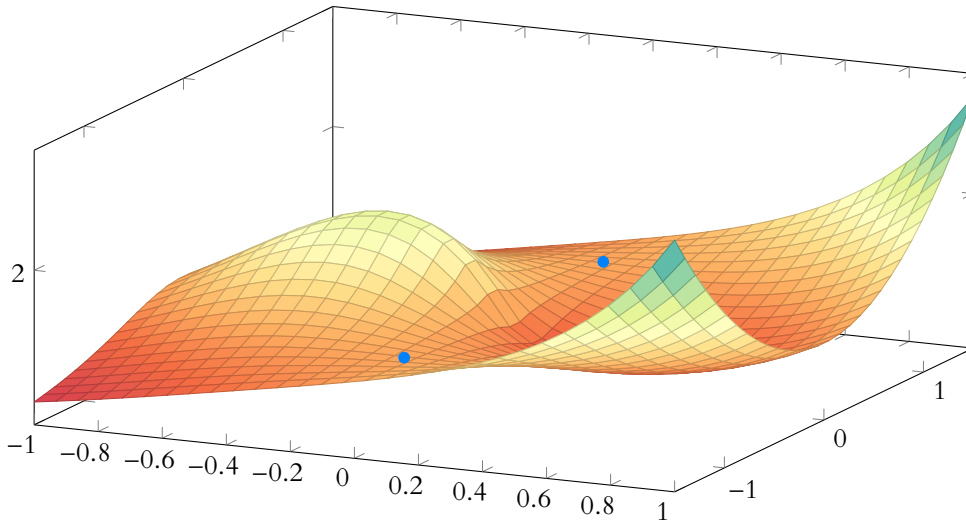


FIGURE 17.7 – La fonction f . Les deux points selles sont représentés en bleu.

SOLUTION DE L'EXERCICE 17.4

1. Les fonctions $(x, y) \mapsto 1 - x$, $(x, y) \mapsto 1 - y$, $(x, y) \mapsto x + y$ sont \mathcal{C}^2 sur Ω , car polynomiales, et ne s'annulent pas sur Ω .
 Donc leurs inverses sont \mathcal{C}^2 , et alors f est \mathcal{C}^2 sur Ω car somme de fonctions \mathcal{C}^2 . On a alors

$$\nabla f(x, y) = \left(\frac{1}{(1-x)^2} - \frac{1}{(x+y)^2}, \frac{1}{(1-y)^2} - \frac{1}{(x+y)^2} \right).$$

De même, il vient $\partial_{1,1}^2 f(x, y) = \frac{2}{(1-x)^3} + \frac{2}{(x+y)^3}$, $\partial_{2,2}^2 f(x, y) = \frac{2}{(1-y)^3} + \frac{2}{(x+y)^3}$ et

$$\partial_{1,2}^2 f(x, y) = \partial_{2,1}^2 f(x, y) = \frac{2}{(x+y)^3}.$$

2. $(x, y) \in \Omega$ est un point critique de f si et seulement si

$$\begin{cases} \frac{1}{(1-x)^2} = \frac{1}{(x+y)^2} \\ \frac{1}{(1-y)^2} = \frac{1}{(x+y)^2} \end{cases} \Leftrightarrow \begin{cases} (1-x)^2 = (x+y)^2 \\ (1-x)^2 = (1-y)^2 \end{cases}$$

Thm. de Schwarz
 Toutes les fonctions rencontrées en exercice sont \mathcal{C}^2 , donc n'oublions pas que le théorème de Schwarz nous permet d'économiser le calcul de certaines dérivées secondes.

Puisque $x \in]0, 1[$ et $y \in]0, 1[$, on a $1 - x \geq 0$, $1 - y \geq 0$ et $x + y \geq 0$. Et donc (x, y) est un point critique si et seulement si

$$\begin{cases} 1 - x = x + y \\ 1 - x = 1 - y \end{cases} \Leftrightarrow \begin{cases} 1 - x = 2x \\ x = y \end{cases} \Leftrightarrow \begin{cases} x = \frac{1}{3} \\ y = \frac{1}{3} \end{cases}$$

Ainsi, f possède sur Ω un unique point critique qui est $\left(\frac{1}{3}, \frac{1}{3}\right)$.

3. La matrice hessienne de f en $\left(\frac{1}{3}, \frac{1}{3}\right)$ est

$$\nabla^2 f\left(\frac{1}{3}, \frac{1}{3}\right) = \begin{pmatrix} \frac{27}{2} & \frac{27}{4} \\ \frac{27}{4} & \frac{27}{2} \end{pmatrix} = \frac{27}{4} \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}.$$

Or, les valeurs propres de $\begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$ sont les racines de $X^2 - 4X + 3$, c'est-à-dire 1 et 3.

Et alors, les valeurs propres de la hessienne sont $\frac{27}{4}$ et $3 \times \frac{27}{4}$.

Ces deux valeurs propres étant strictement positives, f admet en $\left(\frac{1}{3}, \frac{1}{3}\right)$ un minimum local.

SOLUTION DE L'EXERCICE 17.5

Notons que les ensembles de définition sont systématiquement ouverts, et les fonctions de classe \mathcal{C}^2 , ce qui nous permet d'utiliser les points critiques.

1. On a $\nabla f(x, y) = (3x^2 + y, x + 3y^2)$. Donc (x, y) est un point critique de f si et seulement si

$$\begin{cases} 3x^2 + y = 0 \\ x + 3y^2 = 0 \end{cases} \Leftrightarrow \begin{cases} y = -3x^2 \\ x + 27x^4 = 0 \end{cases} \Leftrightarrow \begin{cases} y = -3x^2 \\ x(1 + 27x^3) = 0 \end{cases}$$

L'équation $1 + 27x^3 = 0$ admet une unique solution $x = -\frac{1}{\sqrt[3]{27}} = -\frac{1}{3}$.

f admet donc deux points critiques $(0, 0)$ et $\left(-\frac{1}{3}, -\frac{1}{3}\right)$.

Les dérivées partielles secondes de f sont

$$\partial_{1,1}^2 f(x, y) = 6x, \quad \partial_{2,2}^2 f(x, y) = 6y, \quad \partial_{1,2}^2 f(x, y) = \partial_{2,1}^2 f(x, y) = 1.$$

Donc au point critique $(0, 0)$, on a $\nabla^2 f(0, 0) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$.

Ses valeurs propres sont les racines de $X^2 - 1$: ce sont 1 et -1 . Ces deux valeurs propres sont non nulles et de signes opposés : $(0, 0)$ est un point selle de f .

De même, on a $\nabla^2 f\left(-\frac{1}{3}, -\frac{1}{3}\right) = \begin{pmatrix} -2 & 1 \\ 1 & -2 \end{pmatrix}$.

Ses valeurs propres sont les racines de $X^2 + 4X + 3$: ce sont -3 et -1 . Elles sont toutes deux négatives, donc f admet un maximum local en $\left(-\frac{1}{3}, -\frac{1}{3}\right)$.

Enfin, f n'admet pas d'extrema globaux car

$$\lim_{x \rightarrow +\infty} f(x, 0) = \lim_{x \rightarrow +\infty} x^3 = +\infty \quad \text{et} \quad \lim_{x \rightarrow -\infty} f(0, x) = \lim_{x \rightarrow -\infty} x^3 = -\infty.$$

2. Le gradient de g est

$$\nabla g(x, y) = \left(4y - \frac{1}{x^2}, 4x - \frac{1}{y^2}\right).$$

Les dérivées partielles secondes sont données par

$$\partial_{1,1}^2 g(x, y) = \frac{2}{x^3}, \quad \partial_{2,2}^2 g(x, y) = \frac{2}{y^3}, \quad \partial_{1,2}^2 g(x, y) = \partial_{2,1}^2 g(x, y) = 4.$$

⚠ Carrés

Deux nombres dont les carrés sont égaux ne sont pas nécessairement égaux ! Ils peuvent être opposés. Ici ces deux nombres sont positifs, donc ils sont égaux, mais sans davantage d'information, on aurait juste pu affirmer que

$$|1 - x| = |1 - y|.$$

Valeurs propres

Si on multiplie une matrice par un réel λ , on multiplie ses valeurs propres par λ .

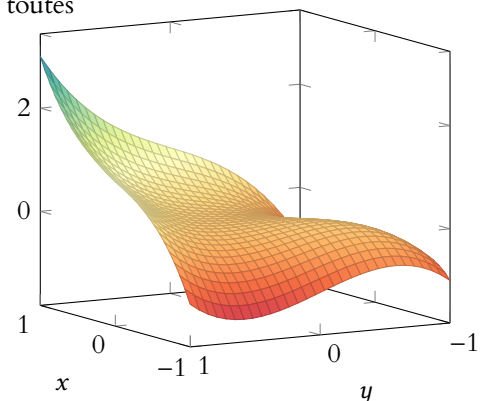


FIGURE 17.8— Le maximum local de f est à droite, le point selle est au centre.

(x, y) est un point critique de g si et seulement si

$$\begin{cases} 4x = \frac{1}{y^2} \\ 4y = \frac{1}{x^2} \end{cases} \Leftrightarrow \begin{cases} x = \frac{1}{4y^2} \\ 4y = 16y^4 \end{cases}$$

Puisque y est non nul par hypothèse, on en déduit que $y^3 = \frac{1}{4}$ et donc $y = \frac{1}{\sqrt[3]{4}}$.

$$\text{Et alors } x = \frac{1}{4} \frac{1}{4^{-2/3}} = 4^{-1/3} = \frac{1}{\sqrt[3]{4}}.$$

La matrice hessienne de g en son point critique est alors $\begin{pmatrix} 8 & 4 \\ 4 & 8 \end{pmatrix}$.

Ses valeurs propres sont les racines de $X^2 - 16X + 48 = (X - 4)(X - 12)$.

Les deux valeurs propres étant positives, on en déduit que g admet un minimum local en son unique point critique.

Ce minimum n'est pas global car $g(x, 1) = 4x + \frac{1}{x} + 1 \xrightarrow{x \rightarrow 0^-} -\infty$.

$$3. \text{ On a } \nabla h(x, y) = \left(2x - \frac{1}{(x+y)^2}, 2y - \frac{1}{(x+y)^2} \right).$$

Ainsi, (x, y) est un point critique de h si et seulement si

$$\begin{cases} 2x - \frac{1}{(x+y)^2} = 0 \\ 2y - \frac{1}{(x+y)^2} = 0 \end{cases} \Leftrightarrow \begin{cases} x = y \\ 2x = \frac{1}{4x^2} \end{cases} \Leftrightarrow \begin{cases} x = y \\ x^3 = \frac{1}{8} \end{cases} \Leftrightarrow x = y = \frac{1}{2}.$$

De plus, les dérivées partielles secondes de h sont données par

$$\partial_{1,1}^2 h(x, y) = 2 + \frac{2}{(x+y)^3}, \partial_{2,2}^2 h(x, y) = 2 + \frac{2}{(x+y)^3}, \partial_{1,2}^2 h(x, y) = \partial_{2,1}^2 h(x, y) = \frac{2}{(x+y)^3}.$$

La hessienne de h au point critique est donc $\nabla^2 h(x, y) = \begin{pmatrix} 4 & 2 \\ 2 & 4 \end{pmatrix}$.

Ses valeurs propres sont les racines de $X^2 - 8X + 12$: ce sont 6 et 2.

Ces deux valeurs propres sont strictement positives, donc h possède un minimum local en $\left(\frac{1}{2}, \frac{1}{2}\right)$.

SOLUTION DE L'EXERCICE 17.6

1. f est de classe \mathcal{C}^1 car elle est polynomiale. On a alors

$$\partial_1 f(x, y, z) = 3x^2 - y, \partial_2 f(x, y, z) = 2y - x, \partial_3 f(x, y, z) = 2z.$$

Donc (x, y, z) est un point critique de f si et seulement si

$$\begin{cases} 3x^2 - y = 0 \\ 2y - x = 0 \\ 2z = 0 \end{cases} \Leftrightarrow \begin{cases} y = 3x^2 \\ 6x^2 - x = 0 \\ z = 0 \end{cases} \Leftrightarrow \begin{cases} y = 3x^2 \\ x(6x - 1) = 0 \\ z = 0 \end{cases}$$

Ainsi, f possède deux points critiques qui sont $(0, 0, 0)$ et $\left(\frac{1}{6}, \frac{1}{12}, 0\right)$.

2. Les dérivées partielles secondes de f sont données par

$$\partial_{1,1}^2 f(x, y, z) = 6x, \partial_{2,2}^2 f(x, y, z) = 2, \partial_{3,3}^2 f(x, y, z) = 2, \partial_{1,2}^2 f(x, y, z) = -1, \partial_{1,3}^2 f(x, y, z) = \partial_{2,3}^2 f(x, y, z) = 0.$$

Ainsi, au point critique $(0, 0, 0)$, la matrice hessienne est donnée par

$$\nabla^2 f(0, 0, 0) = \begin{pmatrix} 0 & -1 & 0 \\ -1 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}.$$

Astuce

Pour manipuler des racines autres que les racines carrées, il est conseillé de repasser aux puissances.

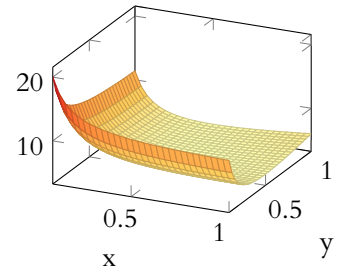


FIGURE 17.9- La fonction g et son minimum.

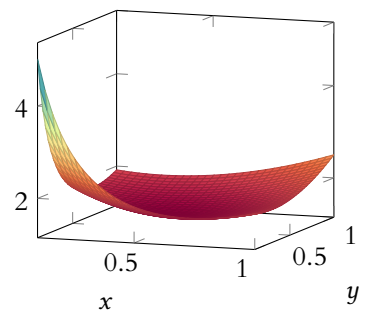


FIGURE 17.10- Le maximum local de h est à droite, le point selle est au centre.

Un pivot montre que les valeurs propres de cette matrice sont $2, 1 + \sqrt{2}$ et $1 - \sqrt{2}$.
Puisqu'il y a une valeur propre strictement positive et une valeur propre strictement négative, $(0, 0, 0)$ est un point selle de f .

De même, on a $\nabla^2 f\left(\frac{1}{6}, \frac{1}{12}, 0\right) = \begin{pmatrix} 1 & -1 & 0 \\ -1 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}$.

Cette fois, les valeurs propres de la hessienne sont $2, \frac{3 + \sqrt{5}}{2}$ et $\frac{3 - \sqrt{5}}{2}$, qui sont toutes strictement positives.

Donc f admet un minimum local en $\left(\frac{1}{6}, \frac{1}{12}, 0\right)$.

Puisqu'on a $\lim_{x \rightarrow -\infty} f(x, 0, 0) = -\infty$, il ne s'agit pas d'un minimum global.

SOLUTION DE L'EXERCICE 17.7

1. f est de classe \mathcal{C}^2 car polynomiale.

On a $\nabla f(x, y) = (2x + 2y + y^3, 2y + 2x + 3xy^2)$. Donc (x, y) est un point critique de f si et seulement si

$$\begin{cases} 2x + 2y + y^3 = 0 \\ 2y + 2x + 3xy^2 = 0 \end{cases} \xleftrightarrow{L_2 \leftarrow L_1 - L_2} \begin{cases} 2x + 2y + y^3 = 0 \\ y^2(y - 3x) = 0 \end{cases}$$

De la deuxième équation, on tire $y = 0$ (et alors $x = 0$) ou $y = 3x$.

Dans ce second cas, la première équation devient alors $8x + 27x^3 = 0 \Leftrightarrow x(8 + 27x^2) = 0$.

Donc soit $x = 0$ (et alors on retrouve $y = 0$), soit $8 + 27x^2 = 0$, ce qui est impossible.

Donc $(0, 0)$ est l'unique point critique de f .

2. On a

$$\partial_{1,1}^2 f(x, y) = 2, \partial_{2,2}^2 f(x, y) = 2 + 2xy \text{ et } \partial_{2,1}^2 f(x, y) = \partial_{1,2}^2 f(x, y) = 2 + 3y^2.$$

$$\text{Donc } \nabla^2 f(0, 0) = \begin{pmatrix} 2 & 2 \\ 2 & 2 \end{pmatrix}.$$

Cette matrice est de rang 1 car ses deux colonnes sont identiques, et donc possède 0 comme valeur propre. On ne peut donc pas utiliser le signe des valeurs propres pour déterminer la nature du point critique $(0, 0)$.

3. On a $f(x, x) = x^4 + 4x^2 = x^2(x^2 + 4) \geq 0$ et $f(x, -x) = -x^4 \leq 0$.

Supposons que f possède un minimum local en $(0, 0)$. Alors il existe $r > 0$ tel que pour $(x, y) \in B_o(0, r)$, $f(0, 0) \leq f(x, y)$.

Mais $\left\| \begin{pmatrix} r \\ 2, -\frac{r}{2} \end{pmatrix} \right\| = \frac{r}{\sqrt{2}} < r$, de sorte que $\begin{pmatrix} r \\ 2, -\frac{r}{2} \end{pmatrix} \in B_o(0, r)$.

Or, $f\left(\frac{r}{2}, -\frac{r}{2}\right) \leq 0 = f(0, 0)$, contredisant le fait que f admet un minimum local en $(0, 0)$.

On montrerait de même, à l'aide de $f\left(\frac{r}{2}, \frac{r}{2}\right)$, que f n'admet pas de maximum local en $(0, 0)$.

SOLUTION DE L'EXERCICE 17.8

1. La fonction $(x, y) \mapsto x^2 + y^2$ est de classe \mathcal{C}^2 car polynomiale.

De même, $(x, y) \mapsto y - x$ est \mathcal{C}^2 sur \mathcal{D} , à valeurs² dans \mathbf{R}_+^* . Puisque la fonction logarithme est \mathcal{C}^2 sur \mathbf{R}_+^* , on en déduit que, par somme et composition de fonction \mathcal{C}^2 , f est \mathcal{C}^2 sur \mathcal{D} . Et donc elle admet des dérivées partielles premières et secondes.

On a alors, pour tout $(x, y) \in \mathcal{D}$,

$$\partial_1 f(x, y) = 2x + \frac{1}{y-x}, \quad \partial_2 f(x, y) = 2y - \frac{1}{y-x}.$$

De même, on a

$$\partial_{1,1}^2 f(x, y) = 2 + \frac{1}{(y-x)^2}, \quad \partial_{1,2}^2 f(x, y) = \partial_{2,1}^2 f(x, y) = -\frac{1}{(x-y)^2}, \quad \partial_{2,2}^2 f(x, y) = 2 + \frac{1}{(y-x)^2}.$$

2. $(x, y) \in \mathcal{D}$ est un point critique de f si et seulement si

$$\begin{cases} 2x + \frac{1}{y-x} = 0 \\ 2y - \frac{1}{y-x} = 0 \end{cases} \Leftrightarrow \begin{cases} 2x + 2y = 0 \\ 2y = \frac{1}{y-x} \end{cases} \Leftrightarrow \begin{cases} x = -y \\ 2y = \frac{1}{2y} \end{cases} \Leftrightarrow \begin{cases} x = -y \\ 4y^2 = 1 \end{cases}$$

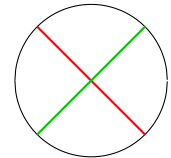


FIGURE 17.11– Toute boule centrée en $(0, 0)$ contient des points (x, x) (en vert), où f prend des valeurs positives et des points $(x, -x)$ (en rouge).

² Notons que cela tient au fait que si $(x, y) \in \mathcal{D}$, alors $x < y$, donc $y - x > 0$.

Puisque x et y sont de signes opposés et que l'on doit³ avoir $y > x$, nécessairement $y > 0$.

Et donc la seconde équation implique que $y = \frac{1}{2}$.

Ainsi, l'unique point critique de f est $\left(-\frac{1}{2}, \frac{1}{2}\right)$.

3. Notons $A = \nabla^2 f\left(-\frac{1}{2}, \frac{1}{2}\right)$ la matrice hessienne de f au point critique $\left(-\frac{1}{2}, \frac{1}{2}\right)$.

Alors $A = \begin{pmatrix} 3 & -1 \\ -1 & 3 \end{pmatrix}$. Un réel λ est valeur propre de A si et seulement si $A - \lambda I_2$ n'est pas inversible, soit si et seulement si $\det(A - \lambda I_2) = 0$.

Or, $\det(A - \lambda I_2) = (3 - \lambda)^2 - 1 = \lambda^2 - 6\lambda + 8 = (\lambda - 2)(\lambda - 4)$. Et donc $\text{Spec}(A) = \{2, 4\}$.

Puisque toutes les valeurs propres de A sont strictement positives, on en déduit que f possède un minimum local en $\left(-\frac{1}{2}, \frac{1}{2}\right)$.

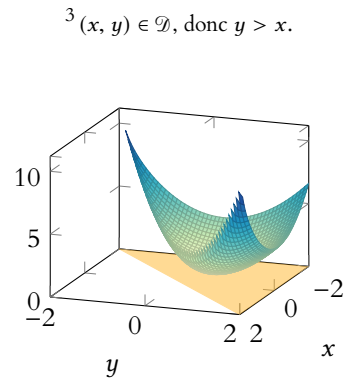


FIGURE 17.12- Le graphe de la fonction f sur l'ouvert \mathcal{D} (représenté en orange).

SOLUTION DE L'EXERCICE 17.9

1. f est polynomiale, donc de classe \mathcal{C}^2 .

On a alors $\nabla f(x, y, z) = 2(x - yz, z - xy, z - xy)$ et

$$\partial_{1,1}^2 f(x, y, z) = \partial_{2,2}^2 f(x, y, z) = \partial_{3,3}^2 f(x, y, z) = 2.$$

$$\partial_{1,2}^2 f(x, y, z) = \partial_{2,1}^2 f(x, y, z) = -2z, \quad \partial_{1,3}^2 f(x, y, z) = \partial_{3,1}^2 f(x, y, z) = -2y, \quad \partial_{2,3}^2 f(x, y, z) = \partial_{3,2}^2 f(x, y, z) = -2x.$$

2. (x, y, z) est un point critique de f si et seulement si

$$\begin{cases} x - yz = 0 \\ y - xz = 0 \\ z - yx = 0 \end{cases} \Leftrightarrow \begin{cases} x = yz \\ y = xz \\ z = yx \end{cases}$$

Si $x = 0$, on a alors $y = z = 0$.

Si $x \neq 0$, alors on a $y = xz = x(xy) = x^2y$, de sorte que $x^2 = 1$, et donc $x = \pm 1$.

Si $x = 1$, alors $y = z$, et $1 = x = yz = y^2$, de sorte que $y = \pm 1$. Et alors $z = xy$.

Les solutions du système sont donc parmi $(0, 0, 0)$, $(1, 1, 1)$, $(-1, 1, -1)$, $(-1, 1, -1)$, $(1, -1, -1)$.

On vérifie que ces cinq triplets sont bien des solutions du système, et donc que f possède cinq points critiques.

3. La matrice hessienne de f en $(0, 0, 0)$ est $\begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}$, qui possède 2 comme unique valeur

propre. Et donc, toutes les valeurs propres de la hessienne étant strictement positives, on en déduit que f admet un minimum local en $(0, 0, 0)$.

Il ne s'agit pas d'un minimum global car $\lim_{x \rightarrow +\infty} f(x, x, x) = \lim_{x \rightarrow +\infty} 3x^2 - x^3 = -\infty$.

4. Les matrices hessiennes aux autres points critiques sont

$$\nabla^2 f(1, 1, 1) = \begin{pmatrix} 2 & -2 & -2 \\ -2 & 2 & -2 \\ -2 & -2 & 2 \end{pmatrix}, \quad \nabla^2 f(1, -1, -1) = \begin{pmatrix} 2 & 2 & 2 \\ 2 & 2 & -2 \\ 2 & -2 & 2 \end{pmatrix}$$

$$\nabla^2 f(-1, -1, 1) = \begin{pmatrix} 2 & -2 & 2 \\ -2 & 2 & 2 \\ 2 & 2 & 2 \end{pmatrix}, \quad \nabla^2 f(-1, 1, -1) = \begin{pmatrix} 2 & 2 & -2 \\ 2 & 2 & 2 \\ -2 & 2 & 2 \end{pmatrix}.$$

On constate que $\nabla^2 f(1, 1, 1) - 4I_3$ est une matrice de rang 1, et donc 4 est valeur propre de $\nabla^2 f(1, 1, 1)$, avec un sous-espace propre de dimension 2.

Mais alors $6 = \text{tr}(\nabla^2 f(1, 1, 1)) = 4 \times 2 + \lambda$, où λ est la dernière valeur propre. On obtient alors $\lambda = -2$.

Ainsi, $\nabla^2 f(1, 1, 1)$ possède au moins une valeur propre strictement positive et une valeur propre strictement négative : f n'admet pas d'extremum local en $(1, 1, 1)$.

Le raisonnement est exactement le même pour chacun des autres points critiques, car 4 est alors encore valeur propre de la hessienne, avec un sous-espace propre de dimension 2, et la trace est encore égale à 6, de sorte que -2 est la dernière valeur propre.

Rang 1

Rappelons qu'une matrice de rang 1 est une matrice dont toutes les colonnes sont colinéaires.

SOLUTION DE L'EXERCICE 17.10

1. f est une fonction polynomiale sur \mathbf{R}^n , donc de classe \mathcal{C}^2 .
On a alors, pour tout $i \in \llbracket 1, n \rrbracket$,

$$\partial_i f(x_1, \dots, x_n) = 2x_i - 2 \left(1 - \sum_{k=1}^n x_k \right).$$

Puis, pour tout $(i, j) \in \llbracket 1, n \rrbracket^2$,

$$\partial_{i,j}^2 f(x_1, \dots, x_n) = \begin{cases} 4 & \text{si } i = j \\ 2 & \text{si } i \neq j \end{cases}$$

2. (x_1, \dots, x_n) est un point critique de f si et seulement si

$$\begin{cases} 2x_1 + 2 \sum_{i=1}^n x_i = 2 \\ 2x_2 + 2 \sum_{i=1}^n x_i = 2 \\ \vdots \\ 2x_n + 2 \sum_{i=1}^n x_i = 2 \end{cases} \xleftrightarrow{L_2 \leftarrow L_2 - L_1} \begin{cases} 2x_1 + 2 \sum_{i=1}^n x_i = 2 \\ x_2 = x_1 \\ \vdots \\ x_n = x_1 \end{cases} \iff \begin{cases} (2n+2)x_1 = 2 \\ x_2 = x_1 \\ \vdots \\ x_n = x_1 \end{cases}$$

Et donc f possède un unique point critique $(a_1, \dots, a_n) = \left(\frac{1}{n+1}, \dots, \frac{1}{n+1} \right)$.

3. On a $A = \begin{pmatrix} 4 & 2 & \dots & 2 \\ 2 & 4 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 2 \\ 2 & \dots & 2 & 4 \end{pmatrix}$. Et donc $A - 2I_n = \begin{pmatrix} 2 & 2 & \dots & 2 \\ 2 & 2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 2 \\ 2 & \dots & 2 & 2 \end{pmatrix}$.

En particulier, $A - 2I_n$ est de rang 1 car toutes ses colonnes sont identiques. On en déduit que 2 est valeur propre de A , et que $\dim E_2(A) = n - 1$.

De plus, A étant symétrique, elle est diagonalisable, et possède donc une autre valeur propre λ , avec un sous-espace propre de dimension 1.

Et alors $4n = \text{tr}(A) = 2(n-1) + \lambda \Leftrightarrow \lambda = 2(n+1)$.

Ainsi, toutes les valeurs propres de A sont strictement positives : f possède un minimum local en a .

4. De plus, f étant \mathcal{C}^2 , g l'est aussi, et on a, pour tout $t \geq 0$, $g'(t) = \langle \nabla f(a + th), h \rangle$ et $g''(t) = q_{a+th}(h)$.

Par la formule de Taylor avec reste intégral, on a alors

$$f(x) = g(1) = g(0) + \underbrace{g'(0)}_{=(\nabla f(a), h)=0} + \int_0^1 g''(t)(1-t) dt = f(a) + \int_0^1 q_{a+th}(h) dt.$$

Mais à la question 1, nous avons constaté que la matrice hessienne de f est la même en tout point. Et donc, quel que soit $t \in [0, 1]$, on a $q_{a+th}(h) \geq 0$.

Alors, par croissance de l'intégrale, $\int_0^1 q_{a+th}(h)(1-t) dt \geq 0$.

On en déduit que $f(a) \leq f(x)$, et donc que f possède un minimum global en a .

SOLUTION DE L'EXERCICE 17.11

1. Les fonctions $(x, y) \mapsto x$ et $(x, y) \mapsto y$ sont \mathcal{C}^1 sur $(\mathbf{R}_+^*)^2$ car polynomiales, et à valeurs dans \mathbf{R}_+^* .

Par composition par la fonction logarithme, qui est \mathcal{C}^1 sur \mathbf{R}_+^* , puis par produit et par somme, la fonction f est donc \mathcal{C}^1 sur $(\mathbf{R}_+^*)^2$.

On a alors

$$\nabla f(x, y) = \left(\ln y - \frac{y}{x}, \frac{x}{y} - \ln x \right).$$

2. (x, y) est un point critique de f si et seulement si

$$\begin{cases} \ln y = \frac{y}{x} \\ \ln x = \frac{x}{y} \end{cases}.$$

Explication

Le vecteur h n'a pas été choisi au hasard : c'est un vecteur directeur de la droite reliant h à a .

En multipliant les deux équations, on obtient $\ln x \ln y = 1$, soit encore $\ln x = \frac{1}{\ln y}$.

En particulier, $\ln x$ et $\ln y$ sont tous deux positifs. En passant au logarithme dans la première équation, on obtient $\ln(\ln y) = \ln(y) - \ln(x) = \ln(y) - \frac{1}{\ln y} \Leftrightarrow \varphi(\ln(y)) = 0$.

3. φ est dérivable et vérifie

$$\varphi'(t) = \frac{1}{t} - 1 - \frac{1}{t^2} = \frac{t - t^2 - 1}{t^2} < 0.$$

Ainsi, φ est strictement décroissante sur \mathbb{R}_+^* . Elle s'annule en $t = 1$.

Donc si (x, y) est un point critique de f , alors $\ln y = 1 \Leftrightarrow y = e$.

Et alors $\ln x = \frac{1}{\ln y} = 1 \Leftrightarrow x = e$.

On en déduit que f possède un unique point critique (e, e) .

4. On a

$$f(e + \alpha, e - \alpha) = (e + \alpha) \ln(e - \alpha) - (e - \alpha) \ln(e + \alpha) = -f(e - \alpha, e + \alpha).$$

Pour α proche de 0, on a

$$\begin{aligned} f(e + \alpha, e - \alpha) &= (e + \alpha) \ln(e - \alpha) - (e - \alpha) \ln(e + \alpha) \\ &= (e + \alpha) \left(1 + \ln \left(1 - \frac{\alpha}{e} \right) \right) - (e - \alpha) \left(1 + \ln \left(1 + \frac{\alpha}{e} \right) \right) \\ &= (e + \alpha) \left(1 - \frac{\alpha}{e} - \frac{\alpha^2}{2e^2} - \frac{\alpha^3}{3e^3} + o(\alpha^3) \right) - (e - \alpha) \left(1 + \frac{\alpha}{e} - \frac{\alpha^2}{2e^2} + \frac{\alpha^3}{3e^3} + o(\alpha^3) \right) \\ &= -\frac{5}{3e^2} \alpha^3 + o(\alpha^3). \end{aligned}$$

Ceci prouve qu'au voisinage de 0, $f(e + \alpha, e - \alpha)$ prend des valeurs non nulles. Et en particulier, pour $\alpha > 0$ suffisamment proche de 0, $f(e + \alpha, e - \alpha) < 0$.

Or, dans une boule centrée en (e, e) , il existe des points de la forme $(e + \alpha, e - \alpha)$ et des points de la forme $(e - \alpha, e + \alpha)$. En ces points, f prend des valeurs respectivement strictement négatives⁴ et strictement positives.

Donc f n'admet pas d'extremum local en (e, e) : (e, e) est un point selle de f .

SOLUTION DE L'EXERCICE 17.12

- La fonction f est continue sur \mathcal{D} car elle y est polynomiale. Puisque \mathcal{D} est un fermé borné, f y admet donc automatiquement un maximum et un minimum.
- Sur l'ouvert \mathcal{D}_0 , f est de classe \mathcal{C}^1 , toujours car elle est polynomiale. Et on a alors

$$\partial_1 f(x, y) = -2xy + 2x, \quad \partial_2 f(x, y) = 2y - x^2.$$

Et donc $(x, y) \in \mathcal{D}_0$ est un point critique de f si et seulement si

$$\begin{cases} -2xy + 2x = 0 \\ 2y - x^2 = 0 \end{cases} \Leftrightarrow \begin{cases} x(1 - y) = 0 \\ 2y - x^2 = 0 \end{cases}$$

De la première équation, il vient $x = 0$ ou $y = 1$.

Si $x = 0$, alors $y = 0$, ce qui nous donne un premier point critique $(0, 0)$.

Si $y = 1$, alors $x = \pm\sqrt{2}$.

Or les points $(\sqrt{2}, 1)$ et $(-\sqrt{2}, 1)$ ne sont pas dans l'ouvert \mathcal{D}_0 .

Donc sur \mathcal{D}_0 , f ne possède qu'un seul point critique en $(0, 0)$.

3. Soit g la fonction définie sur $[-1, 1]$ par

$$g(x) = f(x, x^2 - 1) = (x^2 - 1)^2 - x^2(x^2 - 1) + x^2 = 1.$$

Donc g est constante sur $[-1, 1]$.

Notons h la fonction définie sur $[-1, 1]$ par

$$h(x) = f(x, 1 - x^2) = (1 - x^2)^2 - x^2(1 - x^2) + x^2 = 2x^4 - 2x^2 + 1.$$

Alors h est dérivable sur $[-1, 1]$, avec

$$h'(x) = 8x^3 - 4x = 4x(2x^2 - 1).$$

Le tableau de variation de h est donc donné par

Remarque

En particulier, aucun des deux logarithmes n'est nul.

Détail

Un calcul de discriminant prouve facilement que le numérateur ne s'annule pas sur \mathbb{R} .

⁴ Et donc strictement inférieures à $f(e, e) = 0$.

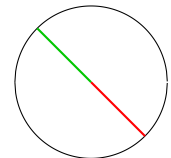


FIGURE 17.13– Toute boule centrée en (e, e) contient des points $(e + \alpha, e - \alpha)$ (en rouge) et des points $(e - \alpha, e + \alpha)$ (en vert).

x	-1	$-\frac{1}{\sqrt{2}}$	0	$\frac{1}{\sqrt{2}}$	1
x	-	-	0	+	+
$2x^2 - 1$	+	0	-	0	+
$g'(x)$	-	0	+	0	+
$g(x)$	1	$\frac{1}{2}$	1	$\frac{1}{2}$	1

Ainsi, h admet un maximum sur $[-1, 1]$, qui vaut 1 et qui est atteint en -1 , en 1 et en 0.

Et h atteint un minimum en $x = \pm \frac{1}{\sqrt{2}}$, et ce minimum vaut $\frac{1}{2}$.

Le minimum m de f est soit atteint sur l'ouvert \mathcal{D}_0 , soit sur $\mathcal{D} \setminus \mathcal{D}_0$.

Or, les points de $\mathcal{D} \setminus \mathcal{D}_0$ sont ceux de la forme $(x, x^2 - 1)$ ou de la forme $(x, 1 - x^2)$.

Nous venons de prouver que sur $\mathcal{D} \setminus \mathcal{D}_0$, f prend des valeurs comprises entre $\frac{1}{2}$ et 1.

Alors que $f(0, 0) = 0$. Donc le minimum de f sur \mathcal{D} est atteint sur \mathcal{D}_0 , et donc nécessairement en $(0, 0)$: $mf(0, 0) = 0$.

De la même manière, le maximum M de f sur \mathcal{D}_0 est nécessairement atteint soit en un point critique sur \mathcal{D}_0 , soit en un point du bord de \mathcal{D} .

Puisque $(0, 0)$ est l'unique point critique, on en déduit que M est atteint sur $\mathcal{D} \setminus \mathcal{D}_0$, et donc les études de fonctions précédemment effectuées prouvent que $M = 1$.

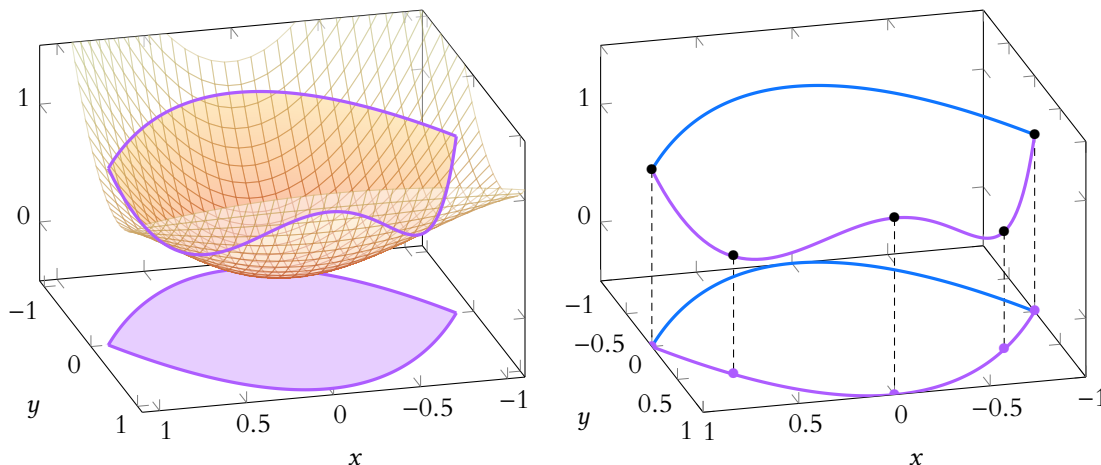


FIGURE 17.14 – En bleu la fonction $x \mapsto f(x, 1 - x^2)$ et en violet la fonction (constante) $x \mapsto f(x, x^2 - 1)$.

SOLUTION DE L'EXERCICE 17.13

- Notons que \mathcal{D} n'est autre que la boule fermée de centre $(0, 0)$ et de rayon 1. Et donc il s'agit d'un fermé borné. Puisque f , qui est polynomiale sur \mathcal{D} est continue, elle y admet donc un maximum et un minimum.
- Notons que $B_o(0, 1)$ est un ouvert, sur lequel f est de classe \mathcal{C}^1 . On a alors $\partial_1 f(x, y) = 3x^2 - 3(1 + y^2)$ et $\partial_2 f(x, y) = -xy$. Ainsi, (x, y) est un point critique de f sur $B_o(0, 1)$ si et seulement si

$$\begin{cases} 3x^2 - 3(1 + y^2) = 0 \\ -6xy = 0 \end{cases} \Leftrightarrow \begin{cases} x^2 = 1 + y^2 \\ xy = 0 \end{cases}$$

De la première équation, il vient $x^2 > 0$, et donc $x \neq 0$, de sorte que la seconde équation nous donne nécessairement $y = 0$.

Et donc $x^2 = 1$. Mais alors $x^2 + y^2 = 1$, et donc $(x, y) \notin B_o(0, 1)$.

Ainsi, f n'admet pas de point critique sur l'ouvert $B_o(0, 1)$.

Et par conséquent, le maximum et le minimum de f sur \mathcal{D} ne peuvent être atteints sur $B_o(0, 1)$.

Ainsi, f atteint forcément son maximum et son minimum sur des points de $\mathcal{D} \setminus B_o(0, 1)$, c'est-à-dire en des points du cercle de centre $(0, 0)$ et de rayon 1 (autrement dit, sur le cercle trigonométrique).

3. Soit g la fonction définie sur \mathbf{R} par $g(t) = f(\cos(t), \sin(t))$.

Lorsque t parcourt \mathbf{R} , $(\cos t, \sin t)$ parcourt le cercle trigonométrique.

Et par 2π -périodicité des fonctions trigonométriques, on peut se limiter à étudier g sur un intervalle de longueur 2π , par exemple $[-\pi, \pi]$.

Et donc l'ensemble $\{f(x, y), x^2 + y^2 = 1\}$ des valeurs prises par f sur le cercle trigonométrique est égal à $\{f(\cos t, \sin t), t \in [-\pi, \pi]\}$, c'est-à-dire à l'image de g .

On a alors

$$g(t) = \cos^3 t - 3 \cos t(1 + \sin^2 t) = \cos t(1 - 3 - 3(1 - \cos^2 t)) = \cos t(-6 + 4 \cos^2 t).$$

En particulier, on remarque que g est paire⁵, et donc il suffit de l'étudier sur $[0, \pi]$.

Ainsi, g est dérivable sur $[0, \pi]$ et

$$g'(t) = -\sin t(-6 + 4 \cos^2 t) + \cos t(-8 \sin t \cos t) = 6 \sin t(1 - 2 \cos^2 t).$$

Pour $t \in [0, \pi]$, on a

$$1 - 2 \cos^2 t \geq 0 \Leftrightarrow \cos^2 t \leq \frac{1}{2} \Leftrightarrow -\frac{\sqrt{2}}{2} \leq \cos t \leq \frac{\sqrt{2}}{2} \Leftrightarrow \frac{\pi}{4} \leq t \leq \frac{3\pi}{4}.$$

Et donc le tableau de variations de g est donné par

t	0	$\pi/4$	$3\pi/4$	π			
$\sin t$	0	+	+	0			
$1 - 2 \cos^2 t$	-	0	+	0	-		
$g'(t)$	0	-	0	+	0	-	0
$g(t)$	-2		$2\sqrt{2}$		2		

\swarrow \nearrow \searrow
 $-2\sqrt{2}$

Ainsi, g admet pour minimum $-2\sqrt{2}$ et pour maximum $2\sqrt{2}$.

On en déduit donc que $m = -2\sqrt{2}$ et $M = 2\sqrt{2}$.

De plus, le minimum de f est alors atteint en deux points de \mathcal{D} qui sont $\left(\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2}\right)$ et

$$\left(\frac{\sqrt{2}}{2}, -\frac{\sqrt{2}}{2}\right).$$

De même, le maximum M de f est atteint en $\left(-\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2}\right)$ et $\left(-\frac{\sqrt{2}}{2}, -\frac{\sqrt{2}}{2}\right)$.

SOLUTION DE L'EXERCICE 17.14

1. La fonction $(x, y) \mapsto 1 + y + xy - \frac{x^2}{2}$ est \mathcal{C}^2 sur \mathbf{R}^2 car polynomiale.

De même, $(x, y) \mapsto y$ est \mathcal{C}^2 sur \mathbf{R}^2 , et la fonction exponentielle est \mathcal{C}^2 sur \mathbf{R} . Par composition, $(x, y) \mapsto e^y$ est \mathcal{C}^2 sur \mathbf{R}^2 .

Et donc f est de classe \mathcal{C}^2 sur \mathbf{R}^2 car produit de deux fonctions de classe \mathcal{C}^2 .

On a alors

$$\partial_1 f(x, y) = e^y(y - x) \text{ et } \partial_2 f(x, y) = e^y \left(1 + y + xy - \frac{x^2}{2} + 1 + x\right).$$

Remarque

Si f possède en $x \in B_o(0, 1)$ un maximum sur \mathcal{D} , alors il s'agit évidemment d'un maximum de f sur $B_o(0, 1)$, et donc le résultat usuel reliant extremum sur un ouvert et points critiques s'applique.

⁵ Rappelons que

$$\cos(-t) = \cos(t).$$

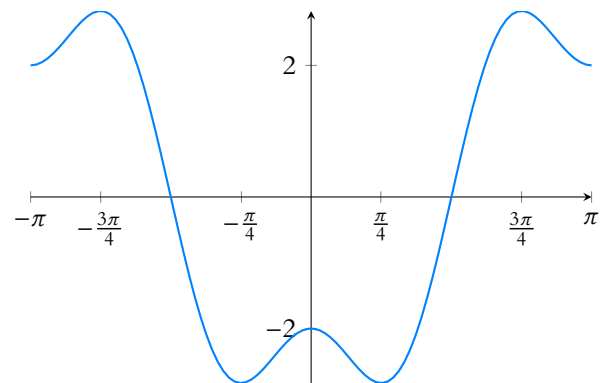


FIGURE 17.15- La fonction g .

Explications

Pour $t = \frac{\pi}{4}$, on trouve

$$(\cos t, \sin t) = \left(\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2}\right).$$

On procède de même pour chacune des valeurs de t donnant un extremum local de g .

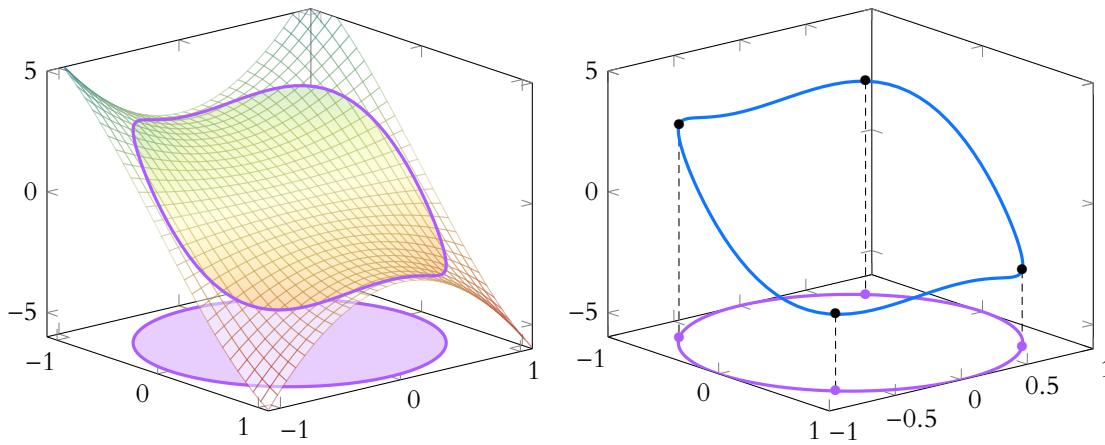


FIGURE 17.16 – En violet : l'ensemble \mathcal{D} . En bleu : les valeurs prises par f sur le bord de cet ensemble, ainsi que les quatre extrema.

En dérivant une seconde fois, il vient

$$\partial_{1,1}^2 f(x, y) = -e^y, \partial_{2,2}^2 f(x, y) = e^y \left(3 + 2x + y + xy - \frac{x^2}{2} \right), \partial_{1,2}^2 f(x, y) = \partial_{2,1}^2 f(x, y) = e^y (y - x + 1).$$

2. (x, y) est un point critique de f si et seulement si

$$\begin{cases} e^y (y - x) = 0 \\ e^y \left(2 + x + y + xy - \frac{x^2}{2} \right) = 0 \end{cases} \Leftrightarrow \begin{cases} y = x \\ 2 + x + y + xy - \frac{x^2}{2} = 0 \end{cases} \Leftrightarrow \begin{cases} y = x \\ 2 + 2x + \frac{x^2}{2} = 0 \end{cases} \Leftrightarrow \begin{cases} y = x \\ 4 + 4x + x^2 = 0 \end{cases}$$

Or $4 + 4x + x^2 = (x + 2)^2$, donc f admet un unique point critique qui est $(-2, -2)$.

3. La matrice hessienne de f en $(-2, -2)$ est $\nabla^2 f(-2, -2) = \begin{pmatrix} -e^{-2} & e^{-2} \\ e^{-2} & -e^{-2} \end{pmatrix}$.

Les deux colonnes sont colinéaires, et donc 0 est valeur propre de $\nabla^2 f(x, y)$.

Ainsi, le signe des valeurs propres de la hessienne ne suffit pas à déterminer la nature du point critique $(-2, -2)$.

4. Les deux premières colonnes sont l'opposée l'une de l'autre, donc $(1, 1)$ est vecteur propre de la hessienne pour la valeur propre 0.
5. La fonction indiquée par l'énoncé est

$$g : t \mapsto f(-2 + t, -2 + t) = e^{-2+t} \left(1 - t + \frac{t^2}{2} \right).$$

Un développement limité nous donne alors

$$g(t) = e^{-2} e^t \left(1 - t + \frac{t^2}{2} \right) = \left(1 + t + \frac{t^2}{2} + \frac{t^3}{6} + o(t^3) \right) \left(1 - t + \frac{t^2}{2} \right) = e^{-2} \left(1 + \frac{1}{6} t^3 + o(t^3) \right).$$

Au voisinage de 0, $g(t) - g(0) \sim \frac{e^{-2}}{6} t^3$ qui est du signe de t .

Ainsi, dans tout voisinage de 0, $f(-2 + t, -2 + t) - f(-2, -2)$ prend à la fois des valeurs positives et des valeurs négatives.

Donc f n'admet pas d'extremum local en $(-2, 2)$.

SOLUTION DE L'EXERCICE 17.15

- 1.a. On a

$$\left(\sum_{i=1}^n x_i X_i - \mu \right)^2 = \left(\sum_{i=1}^n x_i X_i \right)^2 - 2\mu \sum_{i=1}^n x_i X_i + \mu^2.$$

Et donc, par linéarité de l'espérance,

$$f(x_1, \dots, x_n) = E \left(\left(\sum_{i=1}^n x_i X_i \right)^2 \right) - 2\mu \sum_{i=1}^n x_i E(X_i) + \mu^2 = E \left(\left(\sum_{i=1}^n x_i X_i \right)^2 \right) - 2\mu^2 \sum_{i=1}^n x_i + \mu^2.$$

Remarque

Dans le cas d'une fonction d'une variable, si la dérivée et la dérivée seconde en un point sont nulles, on pousse le développement limité à un ordre supérieur pour étudier la nature locale du point critique.

Ici, le vecteur propre associé à la valeur propre 0 nous indique une direction dans laquelle la formule de Taylor à l'ordre 2 ne peut conclure, et nous étudions ce qu'il se passe dans cette direction, à l'aide des outils de première année.

Afin de calculer le moment d'ordre 2 de $\sum_{i=1}^n x_i X_i$, utilisons la formule de Huygens :

$$E\left(\left(\sum_{i=1}^n x_i X_i\right)^2\right) = V\left(\sum_{i=1}^n x_i X_i\right) + E\left(\sum_{i=1}^n x_i X_i\right)^2.$$

Puisque les X_i sont indépendantes, on a

$$V\left(\sum_{i=1}^n x_i X_i\right) = \sum_{i=1}^n x_i^2 V(X_i) = \sigma^2 \sum_{i=1}^n x_i^2.$$

Et donc $E\left(\left(\sum_{i=1}^n x_i X_i\right)^2\right) = \sigma^2 \sum_{i=1}^n x_i^2 + \mu^2 \left(\sum_{i=1}^n x_i\right)^2.$

Au final, on a

$$\begin{aligned} f(x_1, \dots, x_n) &= \sigma^2 \sum_{i=1}^n x_i^2 + \mu^2 \left(\left(\sum_{i=1}^n x_i\right)^2 - 2 \sum_{i=1}^n x_i + 1 \right) \\ &= \sigma^2 \sum_{i=1}^n x_i^2 + \mu^2 \left(\sum_{i=1}^n x_i - 1 \right)^2. \end{aligned}$$

- 1.b. Il s'agit de l'inégalité de Cauchy-Schwarz, appliquée au produit scalaire canonique de \mathbf{R}^n , et aux vecteurs $x = (x_1, \dots, x_n)$ et $y = (1, \dots, 1)$:

$$\langle x, y \rangle^2 \leq \|x\|^2 \cdot \|y\|^2 \Leftrightarrow \left(\sum_{i=1}^n x_i\right)^2 \leq \left(\sum_{i=1}^n x_i^2\right) \underbrace{\left(\sum_{i=1}^n 1^2\right)}_{=n}.$$

- 1.c. Si $(x_1, \dots, x_n) \in \mathcal{H}$, alors on a

$$f(x_1, \dots, x_n) = \sigma^2 \left(\sum_{i=1}^n x_i\right)^2 + \mu^2 \left(\sum_{i=1}^n x_i - 1\right)^2 = \sigma^2 \left(\sum_{i=1}^n x_i\right)^2.$$

Et donc d'après la question précédente, $f(x_1, \dots, x_n) \geq \frac{\sigma^2}{n} \left(\sum_{i=1}^n x_i\right)^2 = \frac{\sigma^2}{n}.$

Or, on a précisément

$$f\left(\frac{1}{n}, \dots, \frac{1}{n}\right) = \sigma^2 \sum_{i=1}^n \frac{1}{n^2} = \frac{\sigma^2}{n}.$$

Puisque $\left(\frac{1}{n}, \dots, \frac{1}{n}\right) \in \mathcal{H}$, $\frac{\sigma^2}{n}$ est bien le minimum de f sur \mathcal{H} .

Remarque : bien que l'énoncé ne le demande pas, il est facile de constater que ce minimum est atteint exactement deux points de \mathcal{H} . En effet, il y a égalité dans l'inégalité de Cauchy-Schwarz de la question précédente si et seulement si $(1, \dots, 1)$ et (x_1, \dots, x_n) sont colinéaires.

Or, il existe exactement deux vecteurs de \mathcal{H} colinéaires à $(1, \dots, 1)$: il s'agit de $\left(\frac{1}{n}, \dots, \frac{1}{n}\right)$ et son opposé.

- 2.a. L'expression obtenue à la question 1.a prouve que f est polynomiale, et donc de classe \mathcal{C}^2 . On a alors, pour $j \in \llbracket 1, n \rrbracket$,

$$\partial_j f(x_1, \dots, x_n) = 2\sigma^2 x_j + 2\mu^2 \left(\sum_{i=1}^n x_i - 1\right).$$

Et donc en particulier, $(x_1, \dots, x_n) \in \mathbf{R}^n$ est un point critique de f si et seulement si

$$\forall j \in \llbracket 1, n \rrbracket, \sigma^2 x_j = -\mu^2 \left(\sum_{i=1}^n x_i - 1\right).$$

⚠ Attention !

À ce stade on a juste prouvé que $\frac{\sigma^2}{n}$ est un minorant de f sur \mathcal{H} , pas qu'il s'agit d'une valeur prise par f sur \mathcal{H} .

En particulier, on a doit avoir $x_1 = x_2 = \dots = x_n = -\frac{\mu}{\sigma^2} \left(\sum_{i=1}^n x_i - 1 \right)$.

Et donc $\sum_{i=1}^n x_i = nx_1$.

La première équation devient alors

$$\sigma^2 x_1 + \mu^2 (nx_1 - 1) = 0 \Leftrightarrow x_1 = \frac{\mu^2}{\sigma^2 + n\mu^2}.$$

Et donc l'unique point critique de f est $a = \left(\frac{\mu^2}{\sigma^2 + n\mu^2}, \dots, \frac{\mu^2}{\sigma^2 + n\mu^2} \right)$.

2.b. Il s'agit d'appliquer la formule de Taylor avec reste intégral à l'ordre 1 à la fonction⁶ $g : t \mapsto f(a + th)$.

⁶ D'une variable.

En effet, nous savons que g est de classe \mathcal{C}^2 , avec $g'(t) = \langle \nabla f(a + th), h \rangle$ et $g''(t) = q_{a+th}(h)$, où q_{a+th} est la forme quadratique associée à la matrice symétrique $\nabla^2 f(a + th)$.

On a donc

$$g(1) = g(0) + g'(0) \times 1 + \int_0^1 (1-t)g''(t) dt.$$

Soit encore

$$f(a+h) = f(a) + \underbrace{\langle \nabla f(a), h \rangle}_{=0} + \int_0^1 (1-t)q_{a+th}(h) dt.$$

Il nous faut donc déterminer la matrice hessienne de f .

Pour $i \in \llbracket 1, n \rrbracket$, on a

$$\partial_{i,i}^2 f(x_1, \dots, x_n) = 2(\sigma^2 + \mu^2).$$

Et pour $i \neq j$, $\partial_{i,j}^2 f(x_1, \dots, x_n) = 2\mu^2$.

Autrement dit, la matrice hessienne de f est constante, égale à

$$A = 2 \begin{pmatrix} \sigma^2 + \mu^2 & \mu^2 & \dots & \mu^2 \\ \mu^2 & \sigma^2 + \mu^2 & \dots & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ \mu^2 & \dots & \mu^2 & \mu^2 + \sigma^2 \end{pmatrix}.$$

Et donc quel que soit $t \in [0, 1]$,

$$\begin{aligned} q_{a+th}(h) &= (h_1 \quad \dots \quad h_n) A \begin{pmatrix} h_1 \\ \vdots \\ h_n \end{pmatrix} = 2(h_1 \quad \dots \quad h_n) \begin{pmatrix} (\sigma^2 + \mu^2)h_1 + \mu^2 \sum_{j=2}^n h_j \\ \vdots \\ \mu^2 \sum_{j=1}^{n-1} h_j + (\mu^2 + \sigma^2)h_n \end{pmatrix} \\ &= 2 \left(\mu^2 \sum_{i=1}^n \sum_{j=1}^n h_i h_j + \sigma^2 \sum_{i=1}^n h_i^2 \right). \end{aligned}$$

2.c. Quel que soit $h \in \mathbf{R}^n$, on a

$$\sum_{i=1}^n \sum_{j=1}^n h_i h_j = \sum_{i=1}^n h_i \left(\sum_{j=1}^n h_j \right) = \left(\sum_{i=1}^n h_i \right) \left(\sum_{j=1}^n h_j \right) = \left(\sum_{i=1}^n h_i \right)^2 \geq 0.$$

Et donc $\mu^2 \sum_{i=1}^n \sum_{j=1}^n h_i h_j + \sigma^2 \sum_{i=1}^n h_i^2 \geq 0$.

On en déduit, par positivité de l'intégrale que $\int_0^1 (1-t) \left(\mu^2 \sum_{i=1}^n \sum_{j=1}^n h_i h_j + \sigma^2 \sum_{i=1}^n h_i^2 \right) dt \geq 0$.

Et donc que $f(a+h) \geq f(a)$, de sorte que f admet un minimum global en a .

Interprétation : bien que l'énoncé ne dise rien à ce sujet, vous avez sûrement reconnu que

$f(x_1, \dots, x_n)$ est le risque quadratique de l'estimateur $\sum_{i=1}^n x_i X_i$ en tant qu'estimateur de μ .

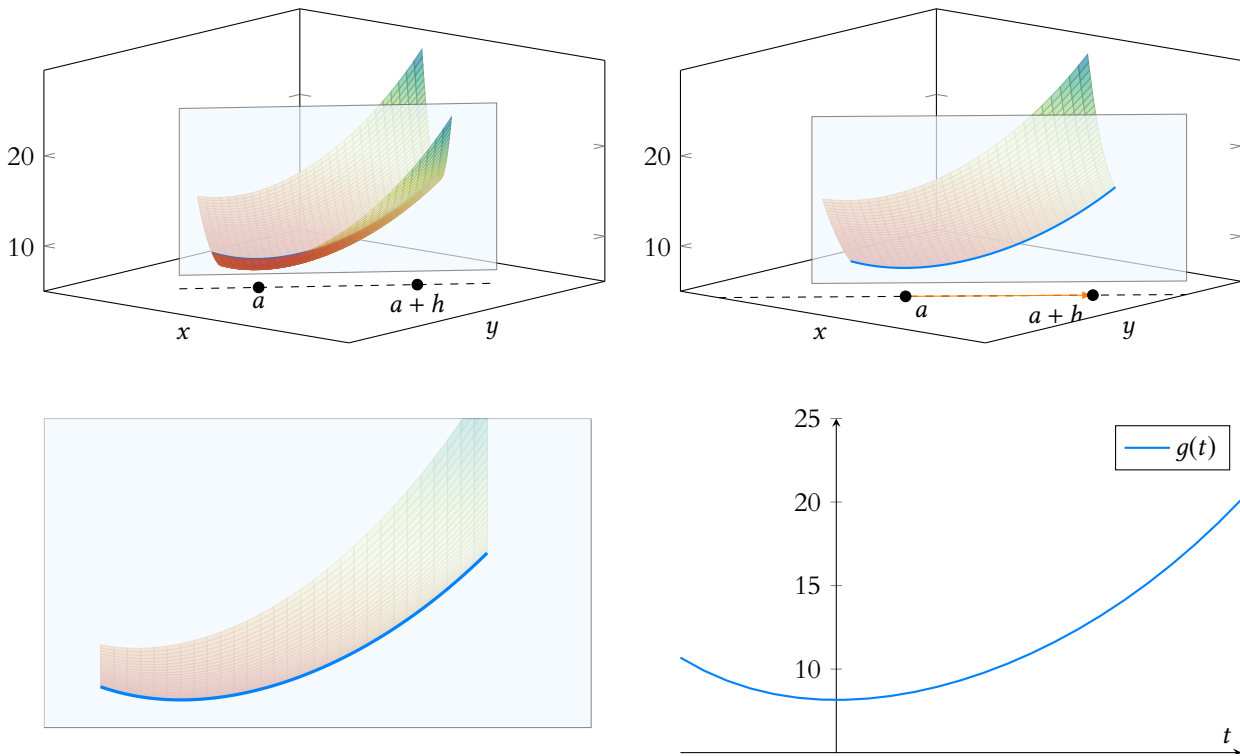


FIGURE 17.17 – Le fait que les valeurs propres de la hessienne en tout point soient positives se traduit par le fait que quel que soit le plan vertical par lequel on coupe le graphe de f , la fonction obtenue (ici g) est convexe. On dit alors que f est une fonction convexe.

Et donc nous venons de prouver que parmi tous les estimateurs de μ qui sont combinaisons linéaires des X_i , il en existe un unique de risque quadratique minimal. De plus, nous savons qu'il est convergent.

D'autre part, la contrainte $x_1 + \dots + x_n = 1$ de la question 1.c revient en réalité à s'intéresser aux estimateurs sans biais de μ qui sont combinaisons linéaire de X_1, \dots, X_n .

Nous avons donc prouvé que parmi ces estimateurs, il en existe deux qui sont de risque quadratique minimal.

Sans biais

$\sum_{i=1}^n x_i X_i$ est un estimateur sans biais de μ si et seulement si son espérance vaut μ , mais par linéarité de l'espérance, c'est le cas si et seulement si $\sum_{i=1}^n x_i = 1$.

ESTIMATION PAR INTERVALLE DE CONFIANCE

Cours	545
18.1 Introduction	545
18.2 Définitions et premiers exemples	545
18.3 Intervalle de confiance asymptotique du paramètre d'une loi de Bernoulli	547
18.4 Intervalle de confiance asymptotique de l'espérance	549
Exercices	552
Correction des exercices	553

18.1 INTRODUCTION

Si nous avons défini un certain nombre de critères permettant de mesurer la qualité d'un estimateur (estimateur sans biais, asymptotiquement sans biais, estimateur convergent), aucun de ces critères ne permet de garantir que la valeur prise par un estimateur T_n à partir d'un échantillon (x_1, \dots, x_n) sera «proche» de la valeur θ du paramètre à estimer.

Par exemple, dans le cas de l'estimation du paramètre p d'une loi de Bernoulli (autrement dit, le cas d'un sondage simple), nous avons étudié $\bar{X}_n = \frac{X_1 + \dots + X_n}{n}$ en tant qu'estimateur de p .

- \bar{X}_n est sans biais : en moyenne il prend la valeur p . Cela signifie que si on réalise un grand nombre de sondages, en moyenne on aura correctement estimé p . Mais lorsqu'on réalise un seul sondage, la moyenne n'a pas d'importance, on souhaite tirer le maximum d'information de ce sondage.

- \bar{X}_n est un estimateur convergent. Par exemple, si pour $\varepsilon = 0.01$, nous savons que

$$\lim_{n \rightarrow +\infty} P(|\bar{X}_n - p| \geq 0.01) = 0.$$

Autrement dit, pour n assez grand, la probabilité que \bar{X}_n diffère de p de plus de 0.01 devient très petite. Mais si on souhaite que cette probabilité soit inférieure à 0.05 (autrement dit, si l'on souhaite que le risque de se tromper de plus de 0.01 soit inférieur à 5%), doit-on prendre $n = 10$, $n = 100$ ou $n = 10^6$? La notion d'estimateur convergent ne nous dit rien à ce sujet, et pour le savoir, il va falloir faire machine arrière et mettre les mains dans le cambouis...

Incertitude

Être sûr à 100 % qu'une valeur estimée est proche de la valeur réelle du paramètre est impossible, on n'est jamais totalement à l'abri d'un «mauvais» échantillon. Ce que nous allons faire ici, c'est essayer de contrôler l'incertitude.

18.2 DÉFINITIONS ET PREMIERS EXEMPLES

Comme dans le chapitre d'estimation ponctuelle, on désigne par (X_n) une suite de variables aléatoires i.i.d. de même loi mère de paramètre inconnu $\theta \in \Theta$, et $g(\theta)$ désigne une fonction de θ que l'on cherche à estimer.

Dans toute la suite, U_n et V_n désignent des estimateurs¹ d'ordre n de $g(\theta)$ tels que $U_n \leq V_n$.

¹ Ou éventuellement deux suites d'estimateurs.

18.2.1 Définition

Définition 18.1 – Soit $\alpha \in]0, 1[$. On dit que $[U_n, V_n]$ est un intervalle de confiance de $g(\theta)$ au niveau de risque α , ou au niveau de confiance $1 - \alpha$ si pour tout $\theta \in \Theta$,

$$P_\theta(U_n \leq g(\theta) \leq V_n) \geq 1 - \alpha.$$

Risque/confiance

Niveau de risque α signifie que la probabilité que $g(\theta)$ ne soit pas dans l'intervalle $[U_n, V_n]$ est au plus égale à α .

Niveau de confiance $1 - \alpha$ signifie que la probabilité que $g(\theta)$ soit dans l'intervalle est au moins égale à $1 - \alpha$. Notons qu'il s'agit bien de la même chose.

Notons que U_n et V_n sont des estimateurs, donc des variables aléatoires fonctions de X_1, \dots, X_n .

L'estimation ponctuelle de $g(\theta)$ par un estimateur $T_n = \varphi(X_1, \dots, X_n)$ consistait à prendre un échantillon (x_1, \dots, x_n) , et à décider d'estimer $g(\theta)$ par $\varphi(x_1, \dots, x_n)$. Autrement dit, nous donnions une valeur précise à $g(\theta)$, au risque de n'avoir aucun contrôle sur l'erreur ainsi réalisée.

Dans l'estimation par intervalle de confiance, nous ne donnons pas une valeur précise à $g(\theta)$, mais l'encadrons entre deux valeurs obtenues à l'aide des estimateurs U_n et V_n (et bien entendu des valeurs (x_1, \dots, x_n) observées sur un échantillon).

Si l'on souhaite par exemple que notre résultat soit certain à 95%, cela revient à prendre $1 - \alpha = 0.95$, et donc un niveau de confiance de 0.95, ou de manière équivalente, un niveau de risque de 0.05.

La plupart du temps, c'est ce niveau de risque qui est utilisé, et qui est communément accepté par exemple en sciences humaines.

Mais dans des domaines plus sensibles où l'on n'a pas vraiment le droit à l'erreur (aérospatiale, physique nucléaire, etc), il arrive de travailler avec des niveaux de risque de 0.01, voire moins.

18.2.2 Premier exemple

Reprenons l'exemple de l'estimation du paramètre p d'une loi de Bernoulli.

Essayons de déterminer, à l'aide de \bar{X}_n , un intervalle de confiance de p au niveau de risque 0.05.

Nous savons par l'inégalité de Bienaymé-Tchebychev que

$$\forall \varepsilon > 0, P(|\bar{X}_n - p| \geq \varepsilon) \leq \frac{V(\bar{X}_n)}{\varepsilon^2}.$$

De plus, par indépendance des X_i , $V(\bar{X}_n) = \frac{1}{n^2} \sum_{i=1}^n V(X_i) = \frac{p(1-p)}{n}$.

Nous ne connaissons pas la valeur de p , mais la fonction $p \mapsto p(1-p)$ atteint sur $[0, 1]$ un maximum en $p = \frac{1}{2}$, et ce maximum vaut $\frac{1}{4}$, donc il est possible de majorer $p(1-p)$ par $\frac{1}{4}$, de sorte que

$$P(|\bar{X}_n - p| \geq \varepsilon) \leq \frac{1}{4n\varepsilon^2}.$$

Et donc

$$P(\bar{X}_n - \varepsilon \leq p \leq \bar{X}_n + \varepsilon) = P(|\bar{X}_n - p| \leq \varepsilon) \geq 1 - \frac{1}{4n\varepsilon^2}.$$

Pour avoir un niveau de risque de 0.05, il suffit que $\frac{1}{4n\varepsilon^2} \leq 0.05$. Soit encore $\varepsilon \geq \frac{1}{2\sqrt{0.05n}}$.

Pour $n = 100$, on obtient $\varepsilon \geq 0.22$. Donc $[\bar{X}_n - 0.22, \bar{X}_n + 0.22]$ est un intervalle de confiance de p au niveau de risque 0.05.

Cela signifie que nous pouvons affirmer qu'il y a 95% de chances que p soit compris entre $\bar{X}_n - 0.22$ et $\bar{X}_n + 0.22$.

L'amplitude de cet intervalle est énorme (rappelons que p est compris entre 0 et 1) !

Pour $n = 1000$ (taille de l'échantillon généralement utilisé par les instituts de sondage), on a encore $\varepsilon \geq 0.07$, soit un intervalle de confiance d'amplitude 0.14.

À l'inverse, on peut fixer ε , et s'intéresser à la taille de l'échantillon (la valeur de n) nécessaire pour garantir que $[\bar{X}_n - \varepsilon, \bar{X}_n + \varepsilon]$ soit un intervalle de confiance de p au niveau de risque 0.05.

Pour $\varepsilon = 0.01$ (on souhaite garantir une erreur inférieure à 1%), il faut

$$n \geq \frac{1}{4 \times 0.05 \times \varepsilon^2} = 50000.$$

Ainsi, pour garantir une erreur inférieure ou égale à 0.01 (avec une certitude de 0.95), il faut interroger 50000 personnes, ce qui est matériellement difficile.

Rappel

Il s'agit du cas d'un sondage où il n'y a que deux réponses possibles.

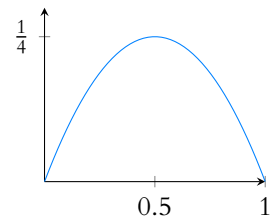


FIGURE 18.1– La fonction $p \mapsto p(1-p)$. Son maximum se trouve soit en dressant son tableau de variations, soit à l'aide d'une formule bien connue donnant l'extremum d'un polynôme de degré 2.

Amplitude

Imaginez un sondage politique où l'on vous dirait «Édouard Philippe» récolte entre 25 et 69% d'opinions favorables»...

Remarques. • Rappelons que nous n'aurons jamais aucune certitude : dans au plus 5% des cas, p sera en dehors de l'intervalle donné. Ces 5% correspondent aux «mauvais» échantillons : ce sont par exemple les échantillons où sur 1000 personnes interrogées, vous avez interrogés les seuls 300 français qui souhaitaient voter pour Skippy, le candidat du Parti pour une Totale Liberté de Pensée Cosmique vers un Nouvel Âge Réminiscent². Un tel échantillon est peu probable, mais nous ne pouvons le négliger.

- Sur ce premier exemple, nous pouvons déjà voir deux résultats plutôt intuitifs : plus le niveau de risque souhaité est petit, plus l'intervalle devra être grand, et plus n est grand (c'est-à-dire plus l'échantillon est important), plus l'intervalle de confiance sera petit.
- Nous avons déjà vu que l'inégalité de Bienaymé-Tchebychev est trop générale pour donner des résultats précis : elle ne prend pas en compte la loi de \overline{X}_n , seules son espérance et sa variance sont importantes. Nous donnerons dans la suite des intervalles de confiance plus petits en utilisant plus finement la loi de \overline{X}_n .

² Les Inconnus, www.youtube.com/watch?v=k_3hIoUGbLs

Exemple 18.2

Par exemple, simulons 10000 échantillons d'une loi de Bernoulli avec $n = 100$ à l'aide du code suivant.

```

1 p = rand(); // p est tire au hasard
2 A = grand(10000,100,'bin',1,p);
3 s = 0;
4 for i=1 : 10000
5     x = mean(A(i, :));
6     if (p<x+0.22) & (p>x-0.22) then
7         s = s+1;
8     end
9 end
10 disp(s)

```

On constate alors que $s = 10000$ (ou si vraiment on n'a pas de chance, $s = 9999$), c'est-à-dire que pour **tous** nos échantillons, p est dans l'intervalle de confiance.

En particulier, au moins 95% d'entre eux sont dans $[\overline{X}_n - 0.22, \overline{X}_n + 0.22]$, mais c'est aussi le signe que l'intervalle de confiance est «trop grand», et qu'on doit pouvoir le réduire tout en garantissant que pour 95% des échantillons, p est dans l'intervalle $[\overline{X}_n - 0.22, \overline{X}_n + 0.22]$.

18.2.3 Intervalle de confiance asymptotique

Définition 18.3 – On dit que $[U_n, V_n]$ est un intervalle de confiance asymptotique de $g(\theta)$ au niveau de confiance $1 - \alpha$ (ou niveau de risque α) si pour tout $\theta \in \Theta$, il existe une suite (α_n) à valeurs dans $[0, 1]$, de limite α , telle que

$$P_\theta(U_n \leq g(\theta) \leq V_n) \geq 1 - \alpha_n.$$

Notons que là encore, on perd de la précision, car l'intervalle obtenu sera bon, pour n assez grand, sans précision sur ce que signifie n assez grand...

En pratique, on n'utilisera pas cette définition, mais on se contentera généralement de prouver que $P_\theta(U_n \leq g(\theta) \leq V_n) \xrightarrow{n \rightarrow +\infty} 1 - \alpha$ ce qui implique³ que $[U_n, V_n]$ est un intervalle de confiance asymptotique de $g(\theta)$ au niveau de risque α .

³ Mais n'est pas équivalent !

18.3 INTERVALLE DE CONFIANCE ASYMPTOTIQUE DU PARAMÈTRE D'UNE LOI DE BERNOULLI

Soit (X_n) une suite de variables i.i.d. suivant la même loi de Bernoulli de paramètre p inconnu. On pose, comme toujours, $\overline{X}_n = \frac{X_1 + \dots + X_n}{n}$.

Alors $X_1 + \dots + X_n$ suit la loi binomiale $\mathcal{B}(n, p)$.

De plus, nous savons⁴ que $\frac{X_1 + \dots + X_n - np}{\sqrt{np(1-p)}} = \sqrt{n} \frac{\overline{X}_n - p}{\sqrt{p(1-p)}}$ converge en loi vers une

⁴ C'est le théorème central limite.

variable X suivant une loi $\mathcal{N}(0, 1)$.

Donc pour tous $a < b \in \mathbf{R}$,

$$\lim_{n \rightarrow +\infty} P\left(a \leq \sqrt{n} \frac{\bar{X}_n - p}{\sqrt{p(1-p)}} \leq b\right) = P(a \leq X \leq b) = \Phi(b) - \Phi(a).$$

Pour obtenir un intervalle de confiance asymptotique au niveau de confiance $1 - \alpha$, il faut donc $\Phi(b) - \Phi(a) \geq 1 - \alpha$.

Il existe une infinité de manières de choisir de tels a et b .

La plus courante est de prendre $a = -b$, et donc $\Phi(b) - \Phi(a) = 2\Phi(b) - 1$.

La fonction Φ réalisant une bijection strictement croissante de \mathbf{R} sur $]0, 1[$, il existe un unique réel t_α tel que $\Phi(t_\alpha) = 1 - \frac{\alpha}{2}$.

Alors $2\Phi(t_\alpha) - 1 = 1 - \alpha$.

Notons qu'on retrouve ainsi, pour $\alpha = 0.05$, $1 - \frac{\alpha}{2} = 0.975$ et donc $t_{0.05} \approx 1.96$.

Alors on a

$$\begin{aligned} & \lim_{n \rightarrow +\infty} P\left(-t_\alpha \leq \sqrt{n} \frac{\bar{X}_n - p}{\sqrt{p(1-p)}} \leq t_\alpha\right) = 1 - \alpha \\ \Leftrightarrow & \lim_{n \rightarrow +\infty} P\left(-t_\alpha \frac{\sqrt{p(1-p)}}{\sqrt{n}} \leq \bar{X}_n - p \leq t_\alpha \frac{\sqrt{p(1-p)}}{\sqrt{n}}\right) = 1 - \alpha \\ \Leftrightarrow & \lim_{n \rightarrow +\infty} P\left(\bar{X}_n - t_\alpha \frac{\sqrt{p(1-p)}}{\sqrt{n}} \leq p \leq \bar{X}_n + t_\alpha \frac{\sqrt{p(1-p)}}{\sqrt{n}}\right) = 1 - \alpha. \end{aligned}$$

Il est tentant de s'arrêter ici. Malheureusement, ceci n'est pas un intervalle de confiance asymptotique.

En effet, $\bar{X}_n \pm \frac{t_\alpha \sqrt{p(1-p)}}{\sqrt{n}}$ dépendent du paramètre inconnu p , et donc il ne s'agit pas d'estimateurs de p .

Pour pallier à ce problème, utilisons, comme précédemment le fait que $0 \leq p(1-p) \leq \frac{1}{4}$, de sorte que

$$\left[\bar{X}_n - t_\alpha \frac{\sqrt{p(1-p)}}{\sqrt{n}} \leq p \leq \bar{X}_n + t_\alpha \frac{\sqrt{p(1-p)}}{\sqrt{n}}\right] \subset \left[\bar{X}_n - \frac{t_\alpha}{2\sqrt{n}} \leq p \leq \bar{X}_n + \frac{t_\alpha}{2\sqrt{n}}\right]$$

et donc

$$\lim_{n \rightarrow +\infty} P\left(\bar{X}_n - \frac{t_\alpha}{2\sqrt{n}} \leq p \leq \bar{X}_n + \frac{t_\alpha}{2\sqrt{n}}\right) \geq 1 - \alpha.$$

Proposition 18.4 : Un intervalle de confiance asymptotique de p au niveau de confiance $1 - \alpha$ est

$$\left[\bar{X}_n - \frac{t_\alpha}{2\sqrt{n}}, \bar{X}_n + \frac{t_\alpha}{2\sqrt{n}}\right].$$

Exemple 18.5

Pour $\alpha = 0.05$, on obtient comme intervalle de confiance $\left[\bar{X}_n - \frac{0.98}{\sqrt{n}}, \bar{X}_n + \frac{0.98}{\sqrt{n}}\right]$,

alors que Bienaymé-Tchebychev nous avait donné $\left[\bar{X}_n - \frac{2.36}{\sqrt{n}}, \bar{X}_n + \frac{2.36}{\sqrt{n}}\right]$.

Par exemple, pour $n = 1000$, on a $\frac{0.98}{\sqrt{1000}} \approx 0.03$. L'intervalle de confiance obtenu est d'amplitude 0.06, contre 0.14 pour celui de Bienaymé-Tchebychev.

De même, si l'on souhaite garantir une précision de 1% avec un niveau de risque

Déjà vu ?

Ce 1.96 ne doit pas vous surprendre plus que ça, vous l'avez rencontré en terminale : si X suit une loi normale $\mathcal{N}(\mu, \sigma^2)$, alors

$$P(\mu - 1.96\sigma \leq X \leq \mu + 1.96\sigma) = 0.95$$

Faites le calcul si vous n'êtes pas convaincus, vous verrez que la raison en est que $\Phi(1.96) \approx 0.975$.

De même, une autre valeur classique est $\Phi(2.57) \approx 0.995$, que l'on rencontre lorsqu'on travaille au niveau de confiance 0.99.

0.05, alors il faut avoir

$$\frac{0.98}{\sqrt{n}} \leq 0.01 \Leftrightarrow n \geq \left(\frac{0.98}{0.01}\right)^2 = 9604.$$

Si ce chiffre reste élevé, il est tout de même meilleur que le 50000 obtenu à l'aide de l'inégalité de Bienaymé-Tchebychev.

Remarques. • Il s'agit là d'un intervalle de confiance **asymptotique**, donc le niveau de risque est bien le bon lorsque n est grand... Avons-nous n suffisamment grand ?

• Remarquons que les calculs réalisés précédemment sont exactement ceux que nous aurions fait en approchant la loi de $X_1 + \dots + X_n$ par la loi normale $\mathcal{N}(np, np(1-p))$. Donc sous les conditions d'approximation de la loi binomiale par la loi normale ($n \geq 30$, p ni trop petit ni trop grand), cet intervalle de confiance est considéré comme fiable.

• Il est bon de savoir que $t_{0.05} \approx 1.96$ et $t_{0.01} \approx 2.57$ (pour un niveau de risque de 0.01).

Exemple 18.6

Un exemple très classique est le premier tour de l'élection présidentielle de 2002. Quelques jours avant les élections, des sondages réalisés auprès de 1000 personnes donnaient 14.5% d'intentions de vote à Jean-Marie Le Pen et 17% à Lionel Jospin. Les intervalles de confiance que nous venons d'obtenir nous indiquent qu'avec une certitude de 95%, le score final de Le Pen serait entre 11.5% et 17.5% et celui de Jospin entre 14% et 20%. L'intersection de ces intervalles est loin d'être vide, et donc il était douteux de conclure uniquement sur la base d'une estimation ponctuelle. Notons que les scores finaux ont été de 16.83% pour Le Pen et 16,18% pour Jospin, donc bien dans les intervalles de confiance que nous venons de calculer.

Estimation ponctuelle

Les nombres ainsi donnés sont des estimations ponctuelles, c'est-à-dire la valeur brute de \bar{X}_n .

Exemple 18.7 Simulation

Simulons 1000 échantillons de taille 1000, et comptons pour combien d'entre eux p est bien dans l'intervalle de confiance⁵ précédent.

```

1  function [y,p] = asymptotique(n,N,alpha)
2      p = rand();
3      A = grand(N,n,'bin',1,p);
4      t = cdfnor('X',0,1,1-alpha/2,alpha/2); //calcul de
      t_alpha
5      y = 0;
6      for i=1 :N
7          x = mean(A(i, :));
8          if (p<x+t/(2*sqrt(n))) & (p>x-t/(2*sqrt(n))) then
9              y = y+1;
10         end
11     end
12     y = y/N;
13 endfunction

```

Si on appelle par exemple `asymptotique(1000,1000,0.05)`, on constate que la proportion d'échantillons pour lesquels p est dans l'intervalle $\left[\bar{X}_n - \frac{t_\alpha}{2\sqrt{n}}, \bar{X}_n + \frac{t_\alpha}{2\sqrt{n}}\right]$ est généralement proche de 95%.

⁵ Asymptotique

Remarque

On constate que parfois, le nombre d'échantillons «corrects» est proche de 1000. Une étude plus poussée prouve que c'est pour les valeurs de p proches de 0 ou de 1, pour lesquelles la majoration $p(1-p) \leq \frac{1}{4}$ était bien trop grossière.

18.4 INTERVALLE DE CONFIANCE ASYMPTOTIQUE DE L'ESPÉRANCE

Le raisonnement de la section précédente peut être utilisé pour toute loi-mère admettant une variance, car le théorème central limite s'applique encore.

Toutefois, notons que la variance va apparaître dans les calculs (c'était le $p(1-p)$ dans le

cas de variables de Bernoulli). Nous avons réussi à nous en «débarrasser» précédemment en la majorant par $\frac{1}{4}$, mais cela était spécifique à la loi de Bernoulli. Nous donnons ici une méthode générale, qui ne nécessite pas d'avoir davantage d'information sur la variance.



Soit donc (X_n) une suite de variables i.i.d., que nous supposons posséder un moment d'ordre 4. On note m l'espérance commune des X_i et σ^2 leur variance.

On note comme d'habitude $\bar{X}_n = \frac{X_1 + \dots + X_n}{n}$ la moyenne empirique, et

$\tilde{S}_n = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X}_n)^2$ la variance empirique. Notons que nous avons déjà étudié \tilde{S}_n dans

le chapitre 16, et avons prouvé qu'il s'agit d'un estimateur biaisé, mais asymptotiquement sans biais de σ^2 .

Nous avons également prouvé que $\tilde{S}_n = \left(\frac{1}{n} \sum_{i=1}^n X_i^2 \right) - \bar{X}_n^2$.

Puisque les X_i possèdent un moment d'ordre 4, les X_i^2 possèdent un moment d'ordre 2, et donc par la loi faible des grands nombres,

$$\frac{1}{n} \sum_{i=1}^n X_i^2 \xrightarrow{P} E(X_1^2).$$

Par la formule de Huygens, on a $E(X_1^2) = V(X_1) + E(X_1)^2 = m^2 + \sigma^2$.

De plus, $\bar{X}_n \xrightarrow{P} m$, et donc par composition par la fonction carré (qui est continue),

$$\bar{X}_n^2 \xrightarrow{P} m^2.$$

Alors (moyennant un lemme dont la preuve se trouve en fin de chapitre)

$$\tilde{S}_n = \frac{1}{n} \sum_{i=1}^n X_i^2 - \bar{X}_n^2 \xrightarrow{P} (m^2 + \sigma^2) - m^2 = \sigma^2.$$

Ainsi, $\frac{1}{\sqrt{\tilde{S}_n}} \xrightarrow{P} \frac{1}{\sigma}$, et donc $\frac{\sigma}{\sqrt{\tilde{S}_n}} \xrightarrow{P} 1$.

Par le théorème central limite, $\sqrt{n} \frac{\bar{X}_n - m}{\sigma} \xrightarrow{\mathcal{L}} X$, où X est une variable aléatoire suivant la loi normale centrée réduite.

Mais alors par le théorème de Slutsky,

$$\sqrt{n} \frac{\bar{X}_n - m}{\sqrt{\tilde{S}_n}} = \sqrt{n} \frac{\bar{X}_n - m}{\sigma} \times \frac{\sigma}{\sqrt{\tilde{S}_n}} \xrightarrow{\mathcal{L}} X \times 1 = X.$$

Soit donc à présent $\alpha \in]0, 1[$ fixé, et soit t_α tel que $\Phi(t_\alpha) = 1 - \frac{\alpha}{2}$. Alors

$$\lim_{n \rightarrow +\infty} P \left(-t_\alpha \leq \sqrt{n} \frac{\bar{X}_n - m}{\sqrt{\tilde{S}_n}} \leq t_\alpha \right) = P(-t_\alpha \leq X \leq t_\alpha) = 1 - \alpha.$$

Comme dans le cas de variables de Bernoulli, on obtient

$$\lim_{n \rightarrow +\infty} P \left(\bar{X}_n - \frac{t_\alpha \sqrt{\tilde{S}_n}}{\sqrt{n}} \leq m \leq \bar{X}_n + \frac{t_\alpha \sqrt{\tilde{S}_n}}{\sqrt{n}} \right) = 1 - \alpha.$$

Nous avons donc prouvé le résultat suivant :

Proposition 18.8 : Si les (X_n) possèdent un moment d'ordre 4, alors un intervalle de confiance asymptotique de m au niveau de risque α est

$$\left[\bar{X}_n - \frac{t_\alpha \sqrt{\tilde{S}_n}}{\sqrt{n}}, \bar{X}_n + \frac{t_\alpha \sqrt{\tilde{S}_n}}{\sqrt{n}} \right].$$

Remarque

Notons que nous prouvons ainsi que bien que \tilde{S}_n soit un estimateur biaisé de σ^2 , il est tout de même convergent.

Slutsky

C'est là le principal usage que nous ferons du théorème de Slutsky : la détermination d'intervalles de confiance asymptotiques.

Remarque. Notons que sans plus d'information, nous ne savons pas à quel point n doit être grand pour que cet intervalle de confiance possède un niveau de risque proche de α . Cela provient du fait que dans le théorème central limite, nous n'avons aucun contrôle sur n : on ne sait pas à partir de quelle valeur de n l'approximation de $\overline{X_n^*}$ par une loi normale est «bonne».

Prouvons enfin le lemme mentionné précédemment :

Lemme 18.9. Soient (X_n) et (Y_n) deux suites de variables aléatoires telles que $X_n \xrightarrow{P} x$ et $Y_n \xrightarrow{P} y$, avec x, y des constantes.
Alors $X_n - Y_n \xrightarrow{P} x - y$.

Démonstration. Soit $\varepsilon > 0$. Alors, si on a à la fois $|X_n - x| < \frac{\varepsilon}{2}$ et $|Y_n - y| < \frac{\varepsilon}{2}$, d'après l'inégalité triangulaire,

$$|(X_n - Y_n) - (x - y)| = |(X_n - x) + (y - Y_n)| \leq |X_n - x| + |Y_n - y| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

Autrement dit, nous avons prouvé l'inclusion d'événements

$$\left[|X_n - x| < \frac{\varepsilon}{2}\right] \cap \left[|Y_n - y| < \frac{\varepsilon}{2}\right] \subset \left[|(X_n - Y_n) - (x - y)| < \varepsilon\right].$$

En passant à l'événement contraire

$$\left[|(X_n - Y_n) - (x - y)| \geq \varepsilon\right] \subset \left[|X_n - x| \geq \frac{\varepsilon}{2}\right] \cup \left[|Y_n - y| \geq \frac{\varepsilon}{2}\right]$$

et donc

$$P\left(\left|(X_n - Y_n) - (x - y)\right| \geq \varepsilon\right) \leq P\left(|X_n - x| \geq \frac{\varepsilon}{2}\right) + P\left(|Y_n - y| \geq \frac{\varepsilon}{2}\right) \xrightarrow{n \rightarrow +\infty} 0.$$

Et donc on a bien $X_n - Y_n \xrightarrow{P} x - y$. □

Rappels

$\overline{A \cap B} = \overline{A} \cup \overline{B}$
et si $A \subset B$, alors $\overline{B} \subset \overline{A}$.

EXERCICES DU CHAPITRE 18

EXERCICE 18.1 Estimation de l'espérance d'une loi normale de variance connue

PD

Soient X_1, \dots, X_n des variables aléatoires i.i.d. suivant la loi normale d'espérance θ inconnue et de variance 1.

On pose $\bar{X}_n = \frac{1}{n} \sum_{i=1}^n X_i$.

- 1) Quelle est la loi de \bar{X}_n ?
- 2) Soit $\alpha \in]0, 1[$, et soit t_α l'unique réel tel que $\Phi(t_\alpha) = 1 - \frac{\alpha}{2}$.

Montrer que $\left[\bar{X}_n - \frac{t_\alpha}{\sqrt{n}}, \bar{X}_n + \frac{t_\alpha}{\sqrt{n}} \right]$ est un intervalle de confiance de θ au niveau de confiance $1 - \alpha$.

- 3) On suppose à présent que les X_i suivent la loi $\mathcal{N}(\theta, \sigma^2)$, où σ^2 est connu. Sur le même principe, proposer un intervalle de confiance de θ au niveau de risque α .

EXERCICE 18.2 On suppose que la probabilité qu'un individu contagieux transmette un virus à un individu sain est $p \in]0, 1[$ inconnu, et que l'on cherche à évaluer.

PD

Soit $(Y_n)_{n \in \mathbb{N}^*}$ une suite de variables indépendantes et identiquement distribuées suivant la loi de Bernoulli de paramètre p .

- 1) On pose $\bar{Y}_n = \frac{1}{n} \sum_{i=1}^n Y_i$.

- a) Montrer que \bar{Y}_n est un estimateur sans biais de p .
- b) Déterminer le risque quadratique de \bar{Y}_n en p .

- 2) À l'aide de l'inégalité de Bienaymé-Tchebychev, montrer que $\left[\bar{Y}_n - \sqrt{\frac{5}{n}}, \bar{Y}_n + \sqrt{\frac{5}{n}} \right]$ est un intervalle de confiance de p au niveau de confiance 0.95.

EXERCICE 18.3 Soit $(X_i)_{i \in \mathbb{N}^*}$ une suite de variables i.i.d. d'espérance μ inconnue et de variance σ^2 connue.

PD

Déterminer à l'aide du théorème central limite un intervalle de confiance asymptotique de μ au niveau de confiance $1 - \alpha$.

EXERCICE 18.4 Soient $(X_i)_{i \geq 1}$ des variables aléatoires i.i.d. suivant la loi exponentielle $\mathcal{E}(\lambda)$ où $\lambda > 0$ est inconnu.

AD

On pose $\bar{X}_n = \frac{1}{n} \sum_{i=1}^n X_i$.

- 1) Montrer que $\lambda\sqrt{n}\bar{X}_n - \sqrt{n}$ converge en loi vers une variable suivant la loi normale centrée réduite.
- 2) Soit $\alpha \in]0, 1[$, et soit t_α l'unique réel tel que $\Phi(t_\alpha) = 1 - \frac{\alpha}{2}$.

Montrer que $\left[\left(1 - \frac{t_\alpha}{\sqrt{n}}\right) \frac{1}{\bar{X}_n}, \left(1 + \frac{t_\alpha}{\sqrt{n}}\right) \frac{1}{\bar{X}_n} \right]$ est un intervalle de confiance asymptotique de λ au niveau de risque α .

EXERCICE 18.5 Soit $(X_i)_{i \in \mathbb{N}^*}$ une suite de variables aléatoires i.i.d. suivant la loi uniforme sur $[0, \theta]$, avec θ inconnu.

AD

On pose $M_n = \max(X_1, \dots, X_n)$.

- 1) Déterminer la fonction de répartition de M_n .
- 2) Montrer que pour $\theta > 0$ fixé, il existe deux uniques réels x_1 et x_2 , que l'on exprimera en fonction de θ , tels que $F_{M_n}(x_1) = 0.025$ et $F_{M_n}(x_2) = 0.975$.
- 3) En déduire un intervalle de confiance de θ au niveau de confiance 0.95.
- 4) Proposer un autre intervalle de θ au niveau de confiance 0.95, de la forme $[M_n, U_n]$, où U_n est à déterminer. *Plus difficile* : lequel de ces deux intervalles possède l'étendue la plus faible ?

EXERCICE 18.6 Soit $(X_i)_{i \in \mathbb{N}^*}$ une suite de variables aléatoires i.i.d. suivant la loi uniforme $\mathcal{U}([0, \theta])$, θ inconnu.

D

On pose $M_n = \max(X_1, \dots, X_n)$.

- 1) Montrer que $\left(n \left(1 - \frac{M_n}{\theta}\right)\right)$ converge en loi vers une variable aléatoire X suivant la loi exponentielle $\mathcal{E}(1)$.
- 2) Déterminer un intervalle de confiance asymptotique de θ au niveau de risque α , sous la forme $[M_n, U_n]$.

EXERCICE 18.7 (QSP ESCP 2007)

TD

Soient $n \geq 1$ et (X_1, \dots, X_n) un échantillon identiquement distribué indépendant de loi de Poisson de paramètre $\lambda > 0$ inconnu.

On pose $\bar{X}_n = \frac{1}{n} \sum_{i=1}^n X_i$ et $T_n = \sqrt{n} \frac{\bar{X}_n - \lambda}{\sqrt{\lambda}}$.

En utilisant T_n , déterminer un intervalle de confiance asymptotique de λ au niveau de risque α .

CORRECTION DES EXERCICES DU CHAPITRE 18

SOLUTION DE L'EXERCICE 18.1

1. Par stabilité des lois normales, et par indépendance des X_i , $\sum_{i=1}^n X_i$ suit la loi normale $\mathcal{N}(n\theta, n)$.

Et donc, par transformation affine de lois normales, \bar{X}_n suit la loi $\mathcal{N}\left(\theta, \frac{1}{n}\right)$.

2. Calculons $P\left(\bar{X}_n - \frac{t_\alpha}{\sqrt{n}} \leq \theta \leq \bar{X}_n + \frac{t_\alpha}{\sqrt{n}}\right)$. On a

$$\begin{aligned} P\left(\bar{X}_n - \frac{t_\alpha}{\sqrt{n}} \leq \theta \leq \bar{X}_n + \frac{t_\alpha}{\sqrt{n}}\right) &= P\left(-\frac{t_\alpha}{\sqrt{n}} \leq \theta - \bar{X}_n \leq \frac{t_\alpha}{\sqrt{n}}\right) \\ &= P(-t_\alpha \leq \sqrt{n}(\theta - \bar{X}_n) \leq t_\alpha) \end{aligned}$$

Mais la variable $\sqrt{n}(\bar{X}_n - \theta)$ suit la loi normale centrée réduite, et donc, par transformation affine, il en est de même de $\sqrt{n}(\theta - \bar{X}_n) = -\sqrt{n}(\bar{X}_n - \theta)$.

On en déduit que la probabilité cherchée est

$$\Phi(t_\alpha) - \Phi(-t_\alpha) = 2\Phi(t_\alpha) - 1 = 2 - \alpha - 1 = 1 - \alpha.$$

Ainsi, on a bien

$$P\left(\bar{X}_n - \frac{t_\alpha}{\sqrt{n}} \leq \theta \leq \bar{X}_n + \frac{t_\alpha}{\sqrt{n}}\right) = 1 - \alpha$$

et donc $\left[\bar{X}_n - \frac{t_\alpha}{\sqrt{n}}, \bar{X}_n + \frac{t_\alpha}{\sqrt{n}}\right]$ est un intervalle de confiance de θ au niveau de confiance¹ $1 - \alpha$.

¹ Ou au niveau de risque α , il s'agit de la même chose.

3. Sur le même principe, on a $\bar{X}_n \hookrightarrow \mathcal{N}\left(\theta, \frac{\sigma^2}{n}\right)$ par stabilité des lois normales².

Et donc, par transformation affine de lois normales, $\frac{\bar{X}_n - \theta}{\sigma} \sqrt{n} \hookrightarrow \mathcal{N}(0, 1)$.

On en déduit donc que

$$P\left(-t_\alpha \leq \sqrt{n} \frac{\bar{X}_n - \theta}{\sigma} \leq t_\alpha\right) = \Phi(t_\alpha) - \Phi(-t_\alpha) = 1 - \alpha.$$

Mais d'autre part, on a

$$P\left(-t_\alpha \leq \sqrt{n} \frac{\bar{X}_n - \theta}{\sigma} \leq t_\alpha\right) = P\left(-\sigma \frac{t_\alpha}{\sqrt{n}} \leq \bar{X}_n - \theta \leq \sigma \frac{t_\alpha}{\sqrt{n}}\right) = P\left(\bar{X}_n - \sigma \frac{t_\alpha}{\sqrt{n}} \leq \theta \leq \bar{X}_n + \sigma \frac{t_\alpha}{\sqrt{n}}\right).$$

Et donc $P\left(\bar{X}_n - \sigma \frac{t_\alpha}{\sqrt{n}} \leq \theta \leq \bar{X}_n + \sigma \frac{t_\alpha}{\sqrt{n}}\right) = 1 - \alpha$, de sorte que $\left[\bar{X}_n - \sigma \frac{t_\alpha}{\sqrt{n}}, \bar{X}_n + \sigma \frac{t_\alpha}{\sqrt{n}}\right]$ est un intervalle de confiance de θ au niveau de confiance $1 - \alpha$.

Connu/inconnu

Notons que si σ était également un paramètre inconnu, alors il ne s'agirait pas d'un intervalle de confiance car alors $\bar{X}_n - \sigma \frac{t_\alpha}{\sqrt{n}}$ ne vérifie pas la définition d'estimateur.

SOLUTION DE L'EXERCICE 18.2

- 1.a. On a $E(\bar{Y}_n) = \sum_{i=1}^n E(Y_i) = p$, donc \bar{Y}_n est un estimateur sans biais de p .

- 1.b. Puisque \bar{Y}_n est sans biais, $r(\bar{Y}_n) = V(\bar{Y}_n)$. Et donc, par indépendance des Y_i ,

$$r(\bar{Y}_n) = \frac{1}{n^2} \sum_{i=1}^n V(Y_i) = \frac{p(1-p)}{n}.$$

2. D'après l'inégalité de Bienaymé-Tchebychev, on a

$$\forall \varepsilon > 0, P(|\bar{Y}_n - p| \geq \varepsilon) \leq \frac{p(1-p)}{n\varepsilon^2}.$$

Donc si l'on souhaite obtenir un intervalle de confiance de Y_n au niveau de confiance 0.95, il faut

$$P(\overline{Y}_n - \varepsilon \leq p \leq \overline{Y}_n + \varepsilon) \geq 0.95 \Leftrightarrow P(|\overline{Y}_n - p| \geq \varepsilon) \leq 0.05.$$

Pour cela, on peut choisir d'avoir $\frac{p(1-p)}{n\varepsilon^2} \leq 0.05$.

Nous ne connaissons pas p , mais $p(1-p) \leq \frac{1}{4}$, de sorte que

$$\frac{1}{4n\varepsilon^2} \leq 0.05 \Leftrightarrow \varepsilon \geq \frac{1}{\sqrt{4 \times 0.05n}} = \frac{1}{\sqrt{\frac{1}{5}n}} = \sqrt{\frac{5}{n}}.$$

Donc pour $\varepsilon = \sqrt{\frac{5}{n}}$, il vient

$$P\left(\overline{Y}_n - \sqrt{\frac{5}{n}} \leq p \leq \overline{Y}_n + \sqrt{\frac{5}{n}}\right) \geq 0.95.$$

Ainsi, $\left[\overline{Y}_n - \sqrt{\frac{5}{n}}, \overline{Y}_n + \sqrt{\frac{5}{n}}\right]$ est un intervalle de confiance de p au niveau de confiance 0.95.

SOLUTION DE L'EXERCICE 18.3

Puisque les X_i sont i.i.d. et possèdent une variance, le théorème central limite s'applique, et donc

$$\sqrt{n} \frac{\overline{X}_n - \mu}{\sigma} \xrightarrow{\mathcal{L}} X$$

où $\overline{X}_n = \frac{1}{n} \sum_{i=1}^n X_i$ et $X \hookrightarrow \mathcal{N}(0, 1)$. Soit alors t_α le réel tel que $\Phi(t_\alpha) = 1 - \frac{\alpha}{2}$.

Il vient alors

$$\lim_{n \rightarrow \infty} P\left(-t_\alpha \leq \sqrt{n} \frac{\overline{X}_n - \mu}{\sigma} \leq t_\alpha\right) = 1 - \alpha$$

soit encore

$$\lim_{n \rightarrow \infty} P\left(\overline{X}_n - t_\alpha \frac{\sigma}{\sqrt{n}} \leq \mu \leq \overline{X}_n + t_\alpha \frac{\sigma}{\sqrt{n}}\right) = 1 - \alpha.$$

Ainsi, un intervalle de confiance asymptotique de μ au niveau de confiance $1 - \alpha$ est

$$\left[\overline{X}_n - t_\alpha \frac{\sigma}{\sqrt{n}}, \overline{X}_n + t_\alpha \frac{\sigma}{\sqrt{n}}\right].$$

SOLUTION DE L'EXERCICE 18.4

1. C'est le théorème central limite ! En effet, les X_i sont i.i.d. et on a

$$E(\overline{X}_n) = \frac{1}{\lambda} \text{ et } V(\overline{X}_n) = \frac{1}{n\lambda^2}.$$

Donc la variable centrée réduite associée à \overline{X}_n est

$$\overline{X}_n^* = \sqrt{n}\lambda \left(\overline{X}_n - \frac{1}{\lambda}\right) = \lambda\sqrt{n}\overline{X}_n - \sqrt{n}.$$

2. On a, par définition de la convergence en loi,

$$\lim_{n \rightarrow +\infty} P\left(-t_\alpha \leq \lambda\sqrt{n}\overline{X}_n - \sqrt{n} \leq t_\alpha\right) = \Phi(t_\alpha) - \Phi(-t_\alpha) = 2\Phi(t_\alpha) - 1 = 1 - \alpha.$$

Mais on a également

$$\left[-t_\alpha \leq \lambda\sqrt{n}\overline{X}_n - \sqrt{n} \leq t_\alpha\right] = \left[-t_\alpha + \sqrt{n} \leq \lambda\sqrt{n}\overline{X}_n \leq t_\alpha + \sqrt{n}\right] = \left[\left(1 - \frac{t_\alpha}{\sqrt{n}}\right) \leq \lambda\overline{X}_n \leq \left(1 + \frac{t_\alpha}{\sqrt{n}}\right)\right].$$

En divisant par \overline{X}_n , qui est bien à valeurs strictement positives, il vient

$$\lim_{n \rightarrow +\infty} P\left(\left(1 - \frac{t_\alpha}{\sqrt{n}}\right) \frac{1}{\overline{X}_n} \leq \lambda \leq \left(1 + \frac{t_\alpha}{\sqrt{n}}\right) \frac{1}{\overline{X}_n}\right) = 1 - \alpha.$$

Et donc $\left[\left(1 - \frac{t_\alpha}{\sqrt{n}}\right) \frac{1}{\overline{X}_n}, \left(1 + \frac{t_\alpha}{\sqrt{n}}\right) \frac{1}{\overline{X}_n}\right]$ est un intervalle de confiance asymptotique de λ au niveau de risque α .

Astuce

Très classique en estimation, lorsqu'on a affaire à un $p(1-p)$ et que l'on ne connaît pas la valeur de p , on majore $p(1-p)$ par $\frac{1}{4}$ (qui se retrouve en calculant le maximum de la fonction $x \mapsto x(1-x)$).

Remarque

Il s'agit du même intervalle de confiance qu'à la question 3 de l'exercice 1, sauf qu'il est cette fois asymptotique : \overline{X}_n ne suit pas une loi normale, mais converge en loi vers une loi normale.

Méthode

Une convergence en loi vers une variable suivant une loi normale ne laisse pas trop de doutes : le TCL se cache **très probablement** derrière, et ce doit être notre premier réflexe.

SOLUTION DE L'EXERCICE 18.5

1. Nous savons que la fonction de répartition des X_i est

$$F_{X_i}(x) = \begin{cases} 0 & \text{si } x \leq 0 \\ \frac{x}{\theta} & \text{si } 0 \leq x \leq \theta \\ 1 & \text{si } x \geq \theta \end{cases}$$

Or, les X_i étant indépendantes, on a, pour $x \in \mathbf{R}$,

$$F_{M_n}(x) = P(M_n \leq x) = P\left(\bigcap_{i=1}^n [X_i \leq x]\right) = \prod_{i=1}^n P(X_i \leq x) = \begin{cases} 0 & \text{si } x \leq 0 \\ \left(\frac{x}{\theta}\right)^n & \text{si } 0 \leq x \leq \theta \\ 1 & \text{si } x \geq \theta \end{cases}$$

2. La fonction F_{M_n} réalise une bijection strictement croissante de $[0, \theta]$ sur $[0, 1]$. Donc il existe un unique x_1 (resp. x_2) tel que $F_{M_n}(x_1) = 0.025$ (resp. $F_{M_n}(x_2) = 0.975$). De manière explicite, on a

$$F_{M_n}(x_1) = 0.025 \Leftrightarrow \left(\frac{x_1}{\theta}\right)^n = 0.025 \Leftrightarrow x_1 = \theta(0.025)^{1/n}$$

et de même, $x_2 = \theta(0.975)^{1/n}$.

3. Nous avons, par définition de la fonction de répartition :

$$P(x_1 \leq M_n \leq x_2) = F_{M_n}(x_2) - F_{M_n}(x_1) = 0.975 - 0.025 = 0.95.$$

Mais

$$P(x_1 \leq M_n \leq x_2) = P(\theta(0.025)^{1/n} \leq M_n \leq \theta(0.975)^{1/n}) = P\left(\frac{M_n}{(0.975)^{1/n}} \leq \theta \leq \frac{M_n}{(0.025)^{1/n}}\right).$$

Donc un intervalle de confiance de θ au niveau de confiance 0.95 est

$$\left[\frac{M_n}{(0.975)^{1/n}}, \frac{M_n}{(0.025)^{1/n}} \right].$$

4. On a également

$$P(\theta(0.05)^{1/n} \leq M_n \leq \theta) = 1 - 0.05 = 0.95.$$

$$\text{Mais } P(\theta(0.05)^{1/n} \leq M_n \leq \theta) = P\left(\frac{M_n}{(0.05)^{1/n}} \leq \theta \leq M_n\right).$$

Et donc $\left[M_n, \frac{M_n}{(0.05)^{1/n}} \right]$ est un intervalle de confiance de θ au niveau de confiance 0.95.

L'étendue du premier intervalle est $M_n \left(\frac{1}{(0.025)^{1/n}} - \frac{1}{(0.975)^{1/n}} \right)$ alors que celle du second est $M_n \left(\frac{1}{(0.025)^{1/n}} - 1 \right)$.

Utilisons un argument de convexité : la fonction $f : x \mapsto \frac{1}{x^{1/n}}$ est convexe sur \mathbf{R}_+^* , car sa dérivée est $x \mapsto -\frac{1}{n} \frac{1}{x^{1/n+1}}$, qui est croissante.

En appliquant deux fois l'inégalité des pentes, on a alors

$$\frac{f(0.975) - f(0.025)}{0.95} \leq \frac{f(0.975) - f(0.05)}{0.925} \text{ et } \frac{f(0.975) - f(0.05)}{0.925} \leq \frac{f(1) - f(0.05)}{0.95}.$$

On en déduit donc que

$$f(1) - f(0.05) \geq f(0.975) - f(0.025) \Leftrightarrow f(0.025) - f(0.975) \geq f(0.05) - f(1),$$

c'est-à-dire que l'étendue du second intervalle est plus petite.

SOLUTION DE L'EXERCICE 18.6

Remarque

Nous avons choisi 0.975 et 0.025 mais le même raisonnement peut se tenir avec tous réels $0 < a < b < 1$ tels que $b - a = 0.95$.

On obtiendrait alors un autre intervalle de confiance, et rien ne nous permet de dire qu'il serait meilleur ou moins bien que celui que nous obtenons.

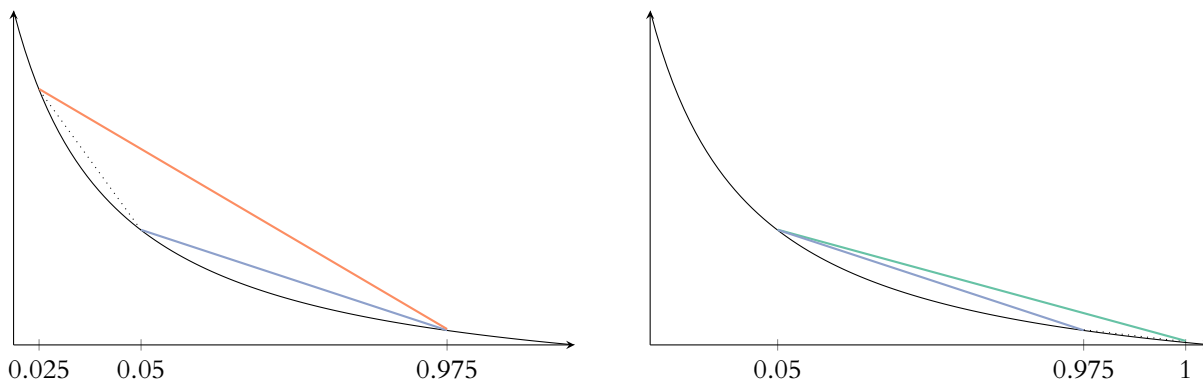


FIGURE 18.1 – Les deux inégalités des pentes : la pente en orange est plus petite⁴ que la pente violette, elle même plus petite que la pente verte.

⁴ Attention : toutes ces pentes sont négatives, et donc « plus petite » signifie « plus pentue ».

1. Commençons par déterminer la fonction de répartition de M_n . Par indépendance des X_i , on a, pour $x \in \mathbf{R}$,

$$F_{M_n}(x) = P(M_n \leq x) = P\left(\bigcap_{i=1}^n [X_i \leq x]\right) = \prod_{i=1}^n P(X_i \leq x) = F_1(x)^n = \begin{cases} 0 & \text{si } x \leq 0 \\ \left(\frac{x}{\theta}\right)^n & \text{si } x \in [0, \theta] \\ 1 & \text{si } x \geq \theta \end{cases}$$

Et donc si on pose $Y_n = n\left(1 - \frac{M_n}{\theta}\right)$, il vient

$$F_{Y_n}(x) = P\left(n\left(1 - \frac{M_n}{\theta}\right) \leq x\right) = P\left(1 - \frac{M_n}{\theta} \leq \frac{x}{n}\right) = P\left(M_n \geq \theta\left(1 - \frac{x}{n}\right)\right) = 1 - F_{M_n}\left(\theta\left(1 - \frac{x}{n}\right)\right).$$

Soit encore

$$F_{Y_n}(x) = \begin{cases} 1 & \text{si } \theta\left(1 - \frac{x}{n}\right) \leq 0 \\ 1 - \left(1 - \frac{x}{n}\right)^n & \text{si } \theta\left(1 - \frac{x}{n}\right) \in [0, \theta] \\ 0 & \text{si } \theta\left(1 - \frac{x}{n}\right) \geq \theta \end{cases} = \begin{cases} 1 & \text{si } x \geq n \\ \left(1 - \frac{x}{n}\right)^n & \text{si } 0 \leq x \leq n \\ 0 & \text{si } x \leq 0 \end{cases}$$

Soit donc $x \in \mathbf{R}$. Si $x \leq 0$, alors $F_{Y_n}(x) = 0 \xrightarrow[n \rightarrow \infty]{} 0$.

Si $x > 0$, alors pour n suffisamment grand, on a $x \leq n$ et donc $F_{Y_n}(x) = 1 - \left(1 - \frac{x}{n}\right)^n$. Mais

$$\left(1 - \frac{x}{n}\right)^n = e^{n \ln\left(1 - \frac{x}{n}\right)}$$

et puisque $\frac{x}{n} \xrightarrow[n \rightarrow \infty]{} 0$, on en déduit que

$$\ln\left(1 - \frac{x}{n}\right) \sim -\frac{x}{n}$$

et donc

$$n \ln\left(1 - \frac{x}{n}\right) \sim -x \xrightarrow[n \rightarrow \infty]{} -x$$

de sorte que

$$\left(1 - \frac{x}{n}\right)^n = e^{n \ln\left(1 - \frac{x}{n}\right)} \xrightarrow[n \rightarrow \infty]{} e^{-x}.$$

Ainsi, il vient $F_{Y_n}(x) \xrightarrow[n \rightarrow \infty]{} 1 - e^{-x}$.

$$\text{Soit } F(x) = \begin{cases} 0 & \text{si } x \leq 0 \\ 1 - e^{-x} & \text{si } x \geq 0 \end{cases}$$

Alors F est la fonction de répartition d'une variable X suivant la loi $\mathcal{E}(1)$ et nous avons prouvé que

$$\forall x \in \mathbf{R}, \lim_{n \rightarrow \infty} F_{Y_n}(x) = F(x).$$

Autrement dit $Y_n \xrightarrow{\mathcal{L}} X$.

***n* «grand»**
 Nous nous intéressons à une limite, donc à n grand. En effet, les premiers (= les 10 premiers, les 100 premiers, les 10^8 premiers, ...) termes d'une suite n'ont pas d'importance sur sa limite, tout ce qui compte est ce qui se passe «à l'infini».

2. Nous venons de prouver que pour tous réels a, b , on a

$$\lim_{n \rightarrow \infty} P(a \leq Y_n \leq b) = P(a \leq X \leq b) = e^{-a} - e^{-b}.$$

Mais

$$\begin{aligned} P(a \leq Y_n \leq b) &= P\left(a \leq n\left(1 - \frac{M_n}{\theta}\right) \leq b\right) \\ &= P\left(\frac{a}{n} \leq 1 - \frac{M_n}{\theta} \leq \frac{b}{n}\right) \\ &= P\left(1 - \frac{b}{n} \leq \frac{M_n}{\theta} \leq 1 - \frac{a}{n}\right) \\ &= P\left(\frac{M_n}{1 - \frac{a}{n}} \leq \theta \leq \frac{M_n}{1 - \frac{b}{n}}\right). \end{aligned}$$

Pour $a = 0$, on a $\frac{M_n}{1 - \frac{a}{n}} = M_n$. Il vient donc

$$\lim_{n \rightarrow \infty} P\left(M_n \leq \theta \leq \frac{M_n}{1 - \frac{b}{n}}\right) = 1 - e^{-b}.$$

Soit donc b tel que $1 - e^{-b} = 1 - \alpha \Leftrightarrow b = -\ln \alpha$: on a

$$\lim_{n \rightarrow \infty} P\left(M_n \leq \theta \leq \frac{M_n}{1 - \frac{-\ln \alpha}{n}}\right) = 1 - \alpha.$$

Un intervalle de confiance asymptotique de θ au niveau de confiance $1 - \alpha$ est donc

$$\left[M_n, \frac{M_n}{1 + \frac{\ln \alpha}{n}} \right].$$

SOLUTION DE L'EXERCICE 18.7

Par le théorème central limite, T_n converge en loi vers une variable aléatoire X suivant la loi normale centrée réduite.

Et donc pour tous réels $a < b$, on a

$$\lim_{n \rightarrow \infty} P(a \leq T_n \leq b) = P(a \leq X \leq b) = \Phi(b) - \Phi(a).$$

En particulier, soit t_α l'unique réel tel que $\Phi(t_\alpha) = 1 - \frac{\alpha}{2}$.

Alors

$$\lim_{n \rightarrow \infty} P\left(-t_\alpha \leq \sqrt{n} \frac{\bar{X}_n - \lambda}{\sqrt{\lambda}} \leq t_\alpha\right) = 1 - \alpha$$

soit encore

$$\lim_{n \rightarrow \infty} P\left(\lambda - \frac{t_\alpha}{\sqrt{n}} \sqrt{\lambda} \leq X_n \leq \lambda + \frac{t_\alpha}{\sqrt{n}} \sqrt{\lambda}\right) = 1 - \alpha.$$

Mais $\lambda - \frac{t_\alpha}{\sqrt{n}} \sqrt{\lambda} = \left(\sqrt{\lambda} - \frac{t_\alpha}{2\sqrt{n}}\right)^2 - \frac{t_\alpha^2}{4n}$, et de même, $\lambda + \frac{t_\alpha}{\sqrt{n}} \sqrt{\lambda} = \left(\sqrt{\lambda} + \frac{t_\alpha}{2\sqrt{n}}\right)^2 - \frac{t_\alpha^2}{4n}$.

Donc on a

$$\lim_{n \rightarrow \infty} P\left(\left(\sqrt{\lambda} - \frac{t_\alpha}{2\sqrt{n}}\right)^2 - \frac{t_\alpha^2}{4n} \leq \bar{X}_n \leq \left(\sqrt{\lambda} + \frac{t_\alpha}{2\sqrt{n}}\right)^2 - \frac{t_\alpha^2}{4n}\right) = 1 - \alpha.$$

Mais on a, au moins pour n suffisamment grand pour que $\sqrt{\lambda} - \frac{t_\alpha}{2\sqrt{n}} \geq 0$,

$$\left[\left(\sqrt{\lambda} - \frac{t_\alpha}{2\sqrt{n}}\right)^2 - \frac{t_\alpha^2}{4n} \leq \bar{X}_n \leq \left(\sqrt{\lambda} + \frac{t_\alpha}{2\sqrt{n}}\right)^2 - \frac{t_\alpha^2}{4n}\right] = \left[\sqrt{\lambda} - \frac{t_\alpha}{2\sqrt{n}} \leq \sqrt{\bar{X}_n + \frac{t_\alpha^2}{4n}}\right] \cap \left[\sqrt{\lambda} + \frac{t_\alpha}{2\sqrt{n}} \geq \sqrt{\bar{X}_n + \frac{t_\alpha^2}{4n}}\right]$$

Intuition

Le fait de choisir $a = 0$ nous est dicté par l'énoncé : on cherche un intervalle dont la borne inférieure est M_n , donc on remarque que pour cela, il faut nécessairement avoir $a = 0$.

Notons que d'autres choix de a nous conduiraient à d'autres intervalles, tout aussi valables (mais pas de la forme imposée par l'énoncé).

Égalité

Notons que n'importe quel b tel que $1 - e^{-b} \geq 1 - \alpha$ convient, il n'est pas nécessaire d'avoir une égalité. L'égalité nous permet toutefois d'avoir l'intervalle le plus petit possible.

Détails

Il s'agit de la mise sous forme canonique de polynômes de degré 2 en $\sqrt{\lambda}$.

$$\begin{aligned}
 &= \left[\sqrt{\overline{X}_n + \frac{t_\alpha^2}{4n}} - \frac{t_\alpha}{2\sqrt{n}} \leq \sqrt{\lambda} \leq \sqrt{\overline{X}_n + \frac{t_\alpha^2}{4n}} + \frac{t_\alpha}{2\sqrt{n}} \right] \\
 &= \left[\left(\sqrt{\overline{X}_n + \frac{t_\alpha^2}{4n}} - \frac{t_\alpha}{2\sqrt{n}} \right)^2 \leq \lambda \leq \left(\sqrt{\overline{X}_n + \frac{t_\alpha^2}{4n}} + \frac{t_\alpha}{2\sqrt{n}} \right)^2 \right].
 \end{aligned}$$

Et donc on a

$$\lim_{n \rightarrow \infty} P \left(\left(\sqrt{\overline{X}_n + \frac{t_\alpha^2}{4n}} - \frac{t_\alpha}{2\sqrt{n}} \right)^2 \leq \lambda \leq \left(\sqrt{\overline{X}_n + \frac{t_\alpha^2}{4n}} + \frac{t_\alpha}{2\sqrt{n}} \right)^2 \right) = 1 - \alpha.$$

Méthode

Le but de ce calcul est d'isoler λ , afin d'obtenir un intervalle de confiance pour λ .

Ainsi, un intervalle de confiance asymptotique de λ au niveau de confiance $1 - \alpha$ est

$$\left[\left(\sqrt{\overline{X}_n + \frac{t_\alpha^2}{4n}} - \frac{t_\alpha}{2\sqrt{n}} \right)^2, \left(\sqrt{\overline{X}_n + \frac{t_\alpha^2}{4n}} + \frac{t_\alpha}{2\sqrt{n}} \right)^2 \right].$$

OPTIMISATION SOUS CONTRAINTES

Cours	559
19.1 Optimisation sous une contrainte non critique	559
19.2 Optimisation sous contrainte(s) d'égalité(s) linéaire(s)	564
Méthodes	568
Exercices	570
Correction des exercices	572

EXEMPLE INTRODUCTIF

Un consommateur dispose de 48 euros, qu'il souhaite dépenser en achetant deux types de biens.

Le premier de ces deux biens coûte 2 euros le kilo, le second coûte 3 euros le kilo. Si on note x (resp. y) la quantité du premier (resp. du second) bien qu'il achète, alors sa fonction d'utilité est donnée par $f(x, y) = -x^2 - 2y^2 + 2xy$.

Afin de maximiser son utilité, il doit donc trouver le maximum de $f(x, y)$ non pas sur \mathbf{R}^2 , mais parmi tous les couples (x, y) de \mathbf{R}^2 vérifiant $2x + 3y = 48$.

Notons qu'une méthode «naïve¹» est de remarquer qu'un tel couple est alors de la forme $\left(x, \frac{48 - 2x}{3}\right)$ et donc il suffit de maximiser la fonction² $x \mapsto f\left(x, \frac{48 - 2x}{3}\right)$.

Nous cherchons dans ce chapitre à développer des méthodes valables dans un contexte plus général, avec des contraintes éventuellement plus complexes que $2x + 3y = 48$, ou avec plusieurs contraintes simultanées.

Dans toute la suite, f est une fonction de classe \mathcal{C}^1 définie sur un ouvert $\Omega \subset \mathbf{R}^n$.

On s'intéresse alors aux extrema (locaux ou globaux) de f sur une partie (généralement fermée) \mathcal{C} de Ω uniquement.

Notons qu'il serait possible de se contenter d'étudier f sur \mathcal{C} , mais nous allons essayer de tirer profit du fait que f soit définie sur Ω .

Terminologie

Ici notre contrainte est une contrainte budgétaire.

¹ Mais néanmoins efficace !

² D'une seule variable !

19.1 OPTIMISATION SOUS UNE CONTRAINTES NON CRITIQUE

Dans tout ce qui suit, nous travaillerons avec des contraintes d'égalité³ : soit φ une fonction de classe \mathcal{C}^1 sur Ω , et soit $c \in \mathbf{R}$. On note alors $\mathcal{C} = \{x \in \Omega : \varphi(x) = c\}$.

Nous cherchons les extrema locaux ou globaux de f sur \mathcal{C} .

Notons que \mathcal{C} est un fermé⁴, et qu'étudier f sur \mathcal{C} ne nous permettrait pas d'utiliser les outils du calcul différentiel, qui nécessitent un ouvert.

Notons également que \mathcal{C} est la ligne de niveau c de la fonction g .

³ Il serait aussi possible de s'intéresser à des contraintes du type $2x + 3y \leq 48$.

⁴ Car défini par une égalité.

19.1.1 Définitions

Définition 19.1 – Soit f une fonction définie sur Ω , et soit $\mathcal{C} = \{x \in \Omega : \varphi(x) = c\}$. On dit que f admet un maximum local (resp. un minimum local) en x_0 sous la contrainte $\varphi(x) = c$ si il existe $r > 0$ tel que

$$\forall x \in B_o(x_0, r) \cap \mathcal{C}, f(x) \leq f(x_0) \text{ (resp. } f(x) \geq f(x_0)).$$

On dit que f admet un maximum global (resp. un minimum global) en x_0 sous la contrainte $\varphi(x) = c$ si f admet un maximum global (resp. un minimum global) sur \mathcal{C} .

Remarque

Dire que f admet un extremum sous la contrainte \mathcal{C} revient à dire que f admet un extremum sur \mathcal{C} .

Exemple 19.2 Recherche d'extremum sous contrainte par substitution

Soit $f(x, y) = e^{x-2y^2}$, et soit $\mathcal{C} = \{(x, y) \in \mathbb{R}^2 : y = x^2\}$.

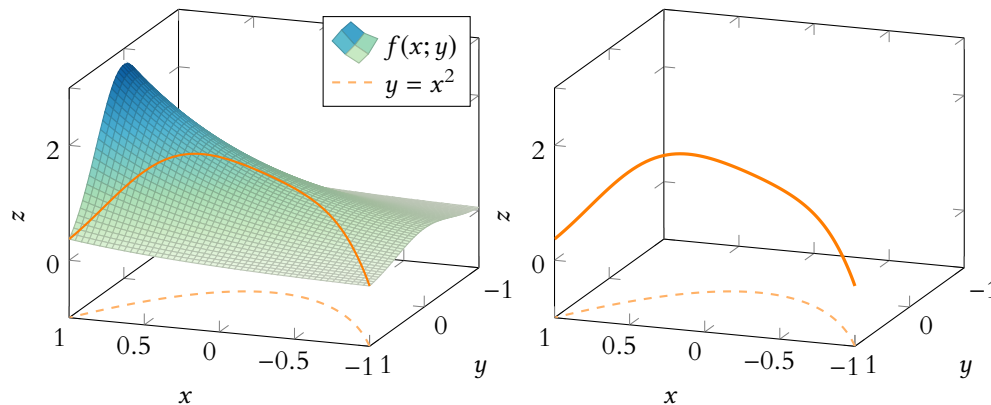
Si $(x, y) \in \mathcal{C}$, alors $y = x^2$, de sorte que $f(x, y) = e^{x-2x^4}$.

Cette fonction ne dépend plus que de la variable x . Une rapide étude de $x \mapsto x - 2x^4$ montre qu'elle admet un maximum global en $x = \frac{1}{2}$.

Donc le maximum de f sous la contrainte $y = x^2$ est atteint en $\left(\frac{1}{2}, \frac{1}{4}\right)$. Notons qu'il ne s'agit pas d'un maximum de f sur \mathbb{R}^2 tout entier.

Plus généralement

On peut adapter cette méthode aux cas où la forme de la contrainte permet d'exprimer l'une des variables (ici y) en fonction des autres, et donc de se ramener à une fonction de strictement moins de variables.



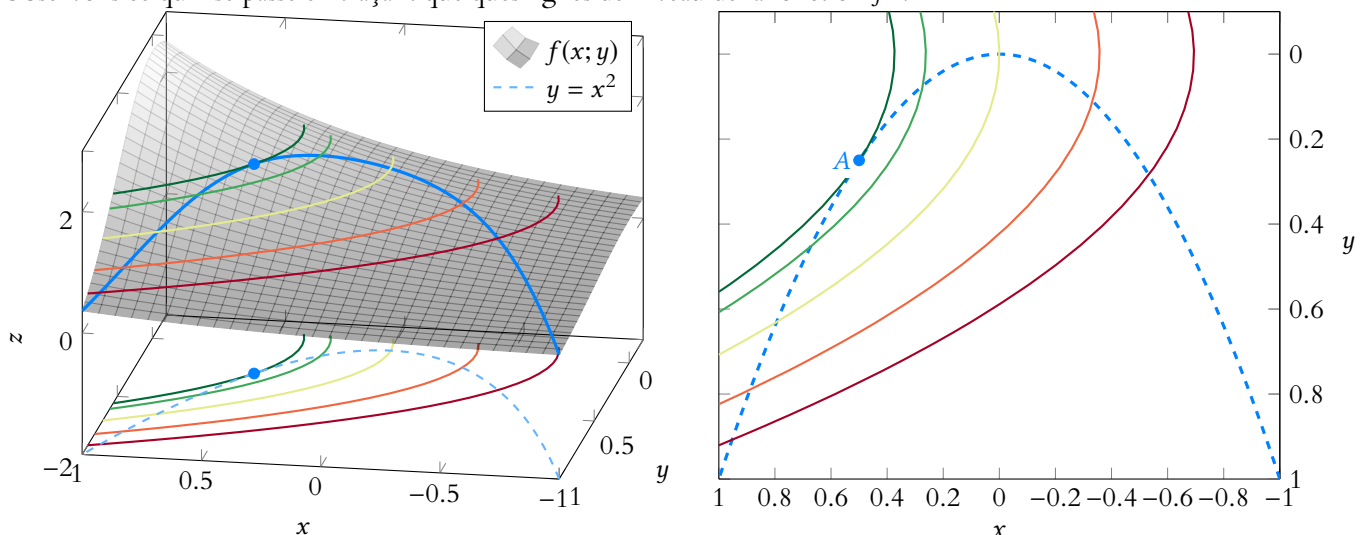
Graphiquement, trouver les extrema locaux de f sous la contrainte \mathcal{C} revient à trouver les points les plus hauts et les plus bas de la courbe orange tracée sur la surface.

19.1.2 Un peu d'intuition

Commençons par reprendre l'exemple précédent.

Notons que la contrainte \mathcal{C} s'écrit encore $y - x^2 = 0$, et donc est la ligne de niveau 0 de la fonction $g(x, y) = y - x^2$.

Observons ce qu'il se passe en traçant quelques lignes de niveau de la fonction f :



Nous constatons qu'au point A où f atteint son maximum sous la contrainte \mathcal{C} , la contrainte \mathcal{C} est tangente à la ligne de niveau de f passant par A . Autrement dit, elle intersecte la ligne de niveau de f , mais ne la «traverse» pas.

Ceci s'explique assez bien si on se souvient qu'une ligne de niveau est formée de points de même altitude.

D'un côté de la ligne de niveau se trouvent des points d'altitude strictement plus élevée, et de l'autre se trouvent des points d'altitude strictement moins élevée⁵.

Donc lorsque la contrainte \mathcal{C} traverse une ligne de niveau en un point (x, y) , cela signifie

⁵ Pour les lignes de niveau de notre exemple, les points plus élevés se trouvent à gauche d'une ligne de niveau.

qu'il existe au voisinage de (x, y) des points de \mathcal{C} en lesquels f prend des valeurs supérieures à $f(x, y)$ et des valeurs inférieures à $f(x, y)$.
Et donc f n'admet pas d'extremum local en (x, y) sous la contrainte \mathcal{C} .

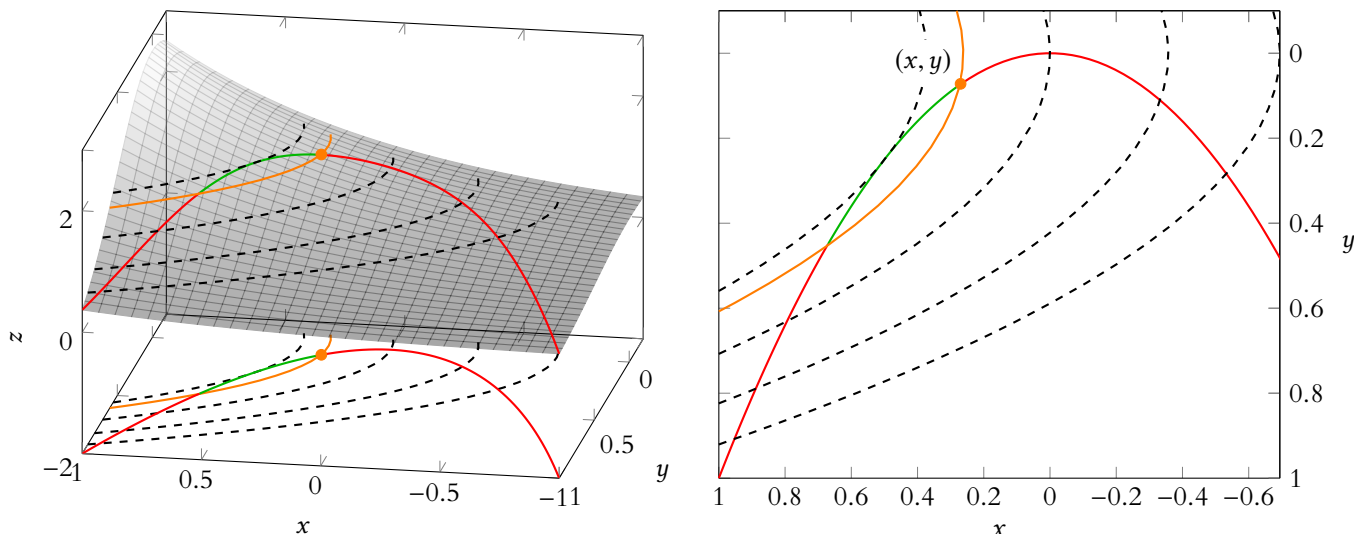


FIGURE 19.1 – En un point où la contrainte \mathcal{C} «traverse» une ligne de niveau de f , il n'y a pas d'extremum local. Ici la ligne de niveau est en orange, les points de \mathcal{C} où f prend des valeurs supérieures à $f(x, y)$ sont en vert, ceux où f prend des valeurs inférieures sont en rouge.

Ainsi, si f possède en (x, y) un extremum local sous la contrainte \mathcal{C} , la courbe de niveau de f passant par (x, y) et \mathcal{C} sont tangentes en (x, y) .

Mais rappelons nous que le gradient est orthogonal aux lignes de niveau, et que cela vaut aussi bien pour f que pour g .

Donc $\nabla f(x, y)$ et $\nabla g(x, y)$ sont tous les deux orthogonaux à une même droite (la tangente commune aux deux lignes de niveau) : ils ont donc même direction⁶, c'est-à-dire sont colinéaires.

⁶ Mais pas nécessairement même sens !

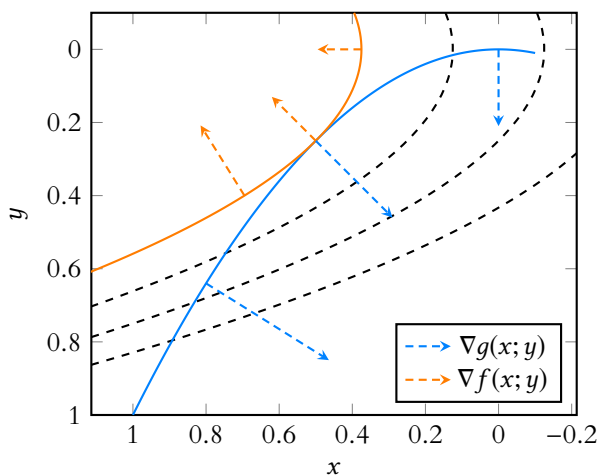


FIGURE 19.2 – En un extremum local de f sous la contrainte \mathcal{C} , les gradients de f et g sont colinéaires.

Le théorème qui va suivre, que nous admettrons, généralise ce que nous venons de voir sur cet exemple, et nous aidera (sous certaines conditions) à trouver des extrema sous contrainte.

19.1.3 Théorème des extrema liés

Définition 19.3 – On dit que $\mathcal{C} = \{x \in \Omega : \varphi(x) = c\}$ est une contrainte non critique si $\forall x \in \mathcal{C}, \nabla\varphi(x) \neq 0$.

Exemple 19.4

Soit $\mathcal{C} = \{(x_1, \dots, x_n) \in \mathbf{R}^n : x_1^2 + \dots + x_n^2 = 1\}$.
 Alors, si on note $\varphi : (x_1, \dots, x_n) \mapsto x_1^2 + \dots + x_n^2$, alors φ est \mathcal{C}^1 sur \mathbf{R}^n , avec $\nabla\varphi(x_1, \dots, x_n) = 2(x_1, \dots, x_n)$.
 En particulier, $\nabla\varphi(x_1, \dots, x_n)$ ne s'annule que lorsque $(x_1, \dots, x_n) = 0$, et donc ne s'annule pas sur \mathcal{C} .
 Par conséquent, \mathcal{C} est une contrainte non critique.

Remarque

\mathcal{C} est appelé sphère de centre 0 et de rayon 1 : pour $n = 2$ c'est un cercle, pour $n = 3$, c'est une sphère au sens usuel du terme.

Théorème 19.5 (Théorème des extrema liés) : Soit f une fonction de classe \mathcal{C}^1 sur un ouvert Ω , et soit $\mathcal{C} = \{x \in \Omega : \varphi(x) = c\}$ une contrainte non critique. Si f admet un extremum local sous la contrainte $\varphi(x) = c$ en x_0 , alors il existe un réel λ tel que

$$\begin{cases} \varphi(x_0) = c \\ \nabla f(x_0) = \lambda \nabla \varphi(x_0) \end{cases}$$

Autrement dit, les gradients de f et de φ en x_0 doivent être colinéaires. Un point x_0 vérifiant ces conditions est appelé point critique de f sous la contrainte \mathcal{C} .

Système

Notons que l'équation $\nabla f(x_0) = \lambda \nabla \varphi(x_0)$ est une égalité entre deux vecteurs de \mathbf{R}^n , et correspond donc à n équations. De plus, λ est une inconnue du système, de sorte qu'on a un système de $n + 1$ équations à $n + 1$ inconnues (les n coordonnées de x_0 et λ).

⚠ Il s'agit d'une condition **nécessaire**. Autrement dit, l'existence d'un point critique sous la contrainte \mathcal{C} ne garantit pas que f atteigne un extremum local sous la contrainte \mathcal{C} en ce point critique. Il faut alors faire une étude plus poussée.

Exemple 19.6

Reprenons l'exemple précédent : $f(x, y) = e^{x-2y^2}$ et $\mathcal{C} = \{(x, y) \in \mathbf{R}^2 : y = x^2\}$.
 Si $\varphi(x, y) = y - x^2$, alors φ est \mathcal{C}^1 et $\mathcal{C} = \{(x, y) \in \mathbf{R}^2 : \varphi(x, y) = 0\}$.
 On a alors $\nabla\varphi = (-2x, 1)$ qui ne s'annule pas sur \mathcal{C} , donc \mathcal{C} est une contrainte non critique.
 De plus, $\nabla f(x, y) = e^{x-2y^2} (1, -4y)$.
 Alors (x, y) est un point critique de f sous la contrainte \mathcal{C} si et seulement si il existe $\lambda \in \mathbf{R}$ tel que

$$\begin{cases} y = x^2 \\ \nabla f(x, y) = \lambda \nabla \varphi(x, y) \end{cases} \Leftrightarrow \begin{cases} y = x^2 \\ e^{x-2y^2} = -2\lambda x \\ -4ye^{x-2y^2} = \lambda \end{cases} \Leftrightarrow \begin{cases} y = x^2 \\ e^{x-2y^2} = -2\lambda x \\ 8\lambda xy = \lambda \end{cases}$$

La seconde équation garantit que $\lambda \neq 0$, et alors on a $8xy = 1$.
 Mais $y = x^2$, de sorte $8xy = 1 \Leftrightarrow 8x^3 = 1 \Leftrightarrow x = \frac{1}{2}$.

Ainsi, l'unique point critique de f sous la contrainte \mathcal{C} est $(\frac{1}{2}, \frac{1}{4})$.

Méthode

La valeur de λ ne nous importe pas au final, l'important est de réussir à déterminer les valeurs de x et de y . On cherchera donc à procéder par substitution sur λ afin de ne garder que des équations en x et en y .

19.1.4 Application à la recherche d'extrema sur un fermé borné

On peut être amenés à chercher les extrema d'une fonction f , continue, sur un ensemble fermé borné \mathcal{D} . Nous savons alors que f admet un maximum global et un minimum global sur \mathcal{D} .

Si \mathcal{D} est de la forme $\{x \in \mathbf{R}^n : \varphi(x) \leq c\}$.

Alors $\mathcal{D}_0 = \{x \in \mathbf{R}^n : \varphi(x) < c\}$ est un ouvert⁸, sur lequel on peut appliquer les théorèmes usuels pour déterminer les points critiques et les extrema locaux⁹ de f .

Il reste alors à étudier f sur le «bord» de \mathcal{D} , c'est-à-dire $\mathcal{C} = \{x \in \mathbf{R}^n : \varphi(x) = c\}$, et on peut alors utiliser le théorème des extrema liés.

⁷ Tous les fermés bornés ne sont pas de cette forme !

⁸ L'«intérieur» de \mathcal{D} .

⁹ Les extrema globaux de f sur \mathcal{D} sont aussi des extrema locaux de f sur \mathcal{D}_0 .

Exemple 19.7

Soit $f(x, y) = \sqrt{1+x^2+y^2} + (x+y)^2 - 1$, et soit $\mathcal{D} = \{(x, y) \in \mathbf{R}^2 : x^2 + y^2 \leq 3\}$. Alors f est de classe \mathcal{C}^1 sur \mathbf{R}^2 , et \mathcal{D} est la boule fermée de centre $(0, 0)$ et de rayon $\sqrt{3}$: c'est un fermé borné.

Par conséquent, f admet un maximum et un minimum sur \mathcal{D} .

Soit $\mathcal{D}_0 = \{(x, y) \in \mathbf{R}^2 : x^2 + y^2 < 3\}$. Alors \mathcal{D}_0 est un ouvert.

On a

$$\partial_1 f(x, y) = \frac{x}{\sqrt{1+x^2+y^2}} + 2(x+y), \quad \partial_2 f(x, y) = \frac{y}{\sqrt{1+x^2+y^2}} + 2(x+y).$$

(x, y) est un point critique de f sur \mathcal{D}_0 si et seulement si

$$\begin{aligned} \begin{cases} \frac{x}{\sqrt{1+x^2+y^2}} + 2(x+y) = 0 \\ \frac{y}{\sqrt{1+x^2+y^2}} + 2(x+y) = 0 \end{cases} &\Leftrightarrow \begin{cases} \frac{x}{\sqrt{1+x^2+y^2}} + 2(x+y) = 0 \\ \frac{x}{\sqrt{1+x^2+y^2}} = \frac{y}{\sqrt{1+x^2+y^2}} \end{cases} &\Leftrightarrow \begin{cases} \frac{x}{\sqrt{1+x^2+y^2}} + 2(x+y) = 0 \\ x = y \end{cases} \\ &\Leftrightarrow \begin{cases} \frac{x}{\sqrt{1+2x^2}} + 4x = 0 \\ x = y \end{cases} &\Leftrightarrow \begin{cases} x \left(\frac{1}{\sqrt{1+2x^2}} + 4 \right) = 0 \\ x = y \end{cases} \end{aligned}$$

De la première équation, on tire facilement $x = 0$ et ainsi, $(0, 0)$ est le seul point critique de f , et on a $f(0, 0) = 0$.

Soit à présent $\mathcal{C} = \{(x, y) \in \mathbf{R}^2 : x^2 + y^2 = 3\}$, le bord de \mathcal{D} .

Alors, si $\varphi(x, y) = x^2 + y^2$, on a $\mathcal{C} = \{(x, y) \in \mathbf{R}^2 : \varphi(x, y) = 3\}$.

Comme $\nabla\varphi(x, y) = (2x, 2y)$ ne s'annule pas sur \mathcal{C} , \mathcal{C} est une contrainte non critique.

Et donc $(x, y) \in \mathcal{C}$ est un point critique de f sous la contrainte \mathcal{C} si et seulement si il existe $\lambda \in \mathbf{R}$ tel que

$$\begin{cases} x^2 + y^2 = 3 \\ \frac{x}{\sqrt{1+x^2+y^2}} + 2(x+y) = 2\lambda x \\ \frac{y}{\sqrt{1+x^2+y^2}} + 2(x+y) = 2\lambda y \end{cases} \Leftrightarrow \begin{cases} x^2 + y^2 = 3 \\ \frac{x}{2} + 2(x+y) = 2\lambda x \\ \frac{y}{2} + 2(x+y) = 2\lambda y \end{cases}$$

En soustrayant les deux dernières équations, il vient

$$\frac{1}{2}(x-y) = 2\lambda(x-y) \Leftrightarrow (x-y)(1-4\lambda) = 0.$$

Donc soit $x = y$, auquel cas la première équation devient $2x^2 = 3 \Leftrightarrow x = \pm\sqrt{\frac{3}{2}}$.

Soit $\lambda = \frac{1}{4}$, auquel cas les deux dernières équations nous donnent $x+y = 0 \Leftrightarrow x = -y$.

Par substitution dans la première équation, on a alors $x = \pm\sqrt{\frac{3}{2}}$.

Ainsi, f admet quatre points critiques sous la contrainte \mathcal{C} :

$$\left(\sqrt{\frac{3}{2}}, \sqrt{\frac{3}{2}}\right), \left(-\sqrt{\frac{3}{2}}, -\sqrt{\frac{3}{2}}\right), \left(\sqrt{\frac{3}{2}}, -\sqrt{\frac{3}{2}}\right) \text{ et } \left(-\sqrt{\frac{3}{2}}, \sqrt{\frac{3}{2}}\right).$$

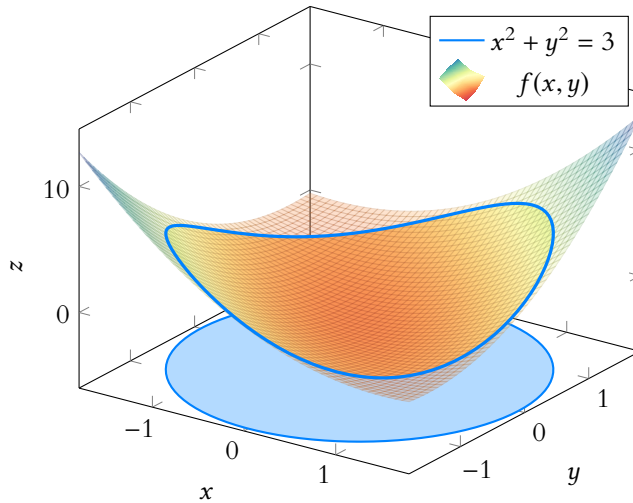
On a alors

$$f\left(\sqrt{\frac{3}{2}}, \sqrt{\frac{3}{2}}\right) = f\left(-\sqrt{\frac{3}{2}}, -\sqrt{\frac{3}{2}}\right) = 7 \text{ et } f\left(\sqrt{\frac{3}{2}}, -\sqrt{\frac{3}{2}}\right) = f\left(-\sqrt{\frac{3}{2}}, \sqrt{\frac{3}{2}}\right) = 1.$$

Ainsi, le maximum de f sur \mathcal{D} est 7 et son minimum est 0.

Astuce

Notons que sur \mathcal{C} , $\sqrt{1+x^2+y^2} = 2$, et donc il serait suffisant d'étudier le comportement de $(x+y)^2$ sur \mathcal{C} .



19.2 OPTIMISATION SOUS CONTRAINTÉ(S) D'ÉGALITÉ(S) LINÉAIRE(S)

19.2.1 Contexte

Dans cette partie, on suppose que la contrainte \mathcal{C} est l'ensemble des solutions d'un système linéaire, de la forme

$$\begin{cases} g_1(x) = b_1 \\ \vdots \\ g_p(x) = b_p \end{cases}$$

où les g_i sont de la forme $g_i : x \mapsto \lambda_{i,1}x_1 + \dots + \lambda_{i,n}x_n$, c'est-à-dire sont des formes linéaires sur \mathbf{R}^n et les b_i sont des réels.

\mathcal{C} est un fermé, car c'est l'intersection des fermés $\{x \in \mathbf{R}^n : g_i(x) = b_i\}$.

On note alors \mathcal{H} l'ensemble des solutions du système homogène associé, c'est-à-dire de

$$\begin{cases} g_1(x) = 0 \\ \vdots \\ g_p(x) = 0 \end{cases}$$

Alors \mathcal{H} est un sous-espace vectoriel de \mathbf{R}^n , car c'est $\bigcap_{i=1}^p \text{Ker } g_i$.

De plus, si $x_0 \in \mathbf{R}^n$ est une solution fixée de \mathcal{C} , alors $y \in \mathbf{R}^n$ est dans \mathcal{C} si et seulement si

$$\begin{cases} g_1(y) = b_1 \\ \vdots \\ g_p(y) = b_p \end{cases} \Leftrightarrow \begin{cases} g_1(y) = g_1(x_0) \\ \vdots \\ g_p(y) = g_p(x_0) \end{cases} \Leftrightarrow \begin{cases} g_1(y - x_0) = 0 \\ \vdots \\ g_p(y - x_0) = 0 \end{cases} \Leftrightarrow y - x_0 \in \bigcap_{i=1}^p \text{Ker } g_i = \mathcal{H}.$$

Autrement dit, si on a une solution x_0 de \mathcal{C} , toutes les autres solutions sont de la forme x_0 plus une solution du système homogène \mathcal{H} .

Soit encore $\mathcal{C} = \{x_0 + h, h \in \mathcal{H}\}$.

19.2.2 Extrema locaux sous la contrainte \mathcal{C}

Proposition 19.8 : Soit f une fonction de classe \mathcal{C}^1 sur Ω . Si f admet un extremum local en x_0 sous la contrainte \mathcal{C} , alors $\nabla f(x_0) \in \mathcal{H}^\perp$.
Un $x_0 \in \mathcal{C}$ vérifiant $\nabla f(x_0) \in \mathcal{H}^\perp$ est appelé point critique de f sous la contrainte \mathcal{C} .

Démonstration. Supposons que f possède un maximum local en x_0 . Alors il existe $r > 0$ tel que pour $x \in B_o(x_0, r) \cap \mathcal{C}$, $f(x) \leq f(x_0)$.

Formes linéaires

Une forme linéaire sur \mathbf{R}^n est toujours une application de la forme

$$(x_1, \dots, x_n) \mapsto a_1x_1 + \dots + a_nx_n.$$

En particulier, c'est une application polynomiale.

Soit $h \in \mathcal{H}$ non nul¹⁰, et soit g la fonction définie au voisinage de 0 par $g(t) \mapsto f(x_0 + th)$.
Alors pour $t < \frac{r}{\|h\|}$, on a

$$\|x_0 + th - x_0\| = \|th\| = |t| \cdot \|h\| < \frac{r}{\|h\|} \|h\| = r$$

de sorte que $x_0 + th \in B_o(x_0, r)$.

De plus, on a alors $x_0 + th \in \mathcal{C}$ car $th \in \mathcal{H}$, et donc $g(t) = f(x_0 + th) \leq f(x_0) = g(0)$.

Ainsi, g admet un maximum local en 0.

Puisque nous savons que g est dérivable, de dérivée égale à $g'(t) = \langle \nabla f(x_0 + th), h \rangle$, cela signifie donc que $g'(0) = 0 \Leftrightarrow \langle \nabla f(x_0), h \rangle = 0$.

Et donc $\nabla f(x_0)$ est orthogonal à h .

Ceci étant vrai pour tout $h \in \mathcal{H}$, on a bien $\nabla f(x_0) \in \mathcal{H}^\perp$. \square

Proposition 19.9 : Avec les notations précédentes, on a $\mathcal{H}^\perp = \text{Vect}(\nabla g_1, \dots, \nabla g_p)$.

Remarque. Les g_i sont des fonctions de la forme $(x_1, \dots, x_n) \mapsto \lambda_{i,1}x_1 + \dots + \lambda_{i,n}x_n$, de sorte que leur gradient est constant, égal à $\nabla g_i = (\lambda_{i,1}, \dots, \lambda_{i,n})$.

Par exemple, si $g(x, y, z) = -2x + 3y + z$, alors $\nabla g = (-2, 3, 1)$.

Démonstration. Commençons par remarquer que pour tout $i \in \llbracket 1, p \rrbracket$, g_i est constante¹¹ sur \mathcal{C} , et donc admet un extremum local, de sorte que $\nabla g_i \in \mathcal{H}^\perp$.

Donc $\text{Vect}(\nabla g_1, \dots, \nabla g_p) \subset \mathcal{H}^\perp$, et en particulier, $\dim \text{Vect}(\nabla g_1, \dots, \nabla g_p) \leq n - \dim \mathcal{H}$.

Inversement, soit $x = (x_1, \dots, x_n) \in (\text{Vect}(\nabla g_1, \dots, \nabla g_n))^\perp$.

Alors, avec les notations précédentes,

$$\forall i \in \llbracket 1, p \rrbracket, 0 = \langle \nabla g_i, x \rangle = \sum_{j=1}^n \lambda_{i,j} x_j = g_i(x).$$

Et donc $\forall i \in \llbracket 1, p \rrbracket, g_i(x) = 0$: c'est la définition de $x \in \mathcal{H}$.

Ainsi, $\text{Vect}(\nabla g_1, \dots, \nabla g_p)^\perp \subset \mathcal{H}$.

On en déduit que $n - \dim \text{Vect}(\nabla g_1, \dots, \nabla g_p) \leq \dim \mathcal{H}$.

Alors $\dim \text{Vect}(\nabla g_1, \dots, \nabla g_p) = n - \dim \mathcal{H} = \dim \mathcal{H}^\perp$, de sorte que $\text{Vect}(\nabla g_1, \dots, \nabla g_p) = \mathcal{H}^\perp$. \square

△ Détails

Puisque Ω est un ouvert, f est définie au moins sur une petite boule autour de x_0 .
Et pour t «suffisamment petit», $x_0 + th$ est dans cette boule, et donc $f(x_0 + th)$ est bien défini.

¹¹ C'est la définition de \mathcal{C} : sur \mathcal{C} , $g_i(x) = b_i$.

Exemple 19.10

Ce résultat permet d'étudier la nature d'un point critique à l'aide du signe sur \mathcal{H} de la forme quadratique associée à la hessienne.

Soit $f(x, y, z) = x^2 + y^2 + 2z^2$, et soit \mathcal{C} la contrainte $2x - y + z = 3$.

Alors \mathcal{H} est l'ensemble des solutions de $2x - y + z = 0$, et $\mathcal{H}^\perp = \text{Vect}((2, -1, 1))$.

Ainsi, les points critiques de f sous la contrainte \mathcal{C} sont les (x_0, y_0, z_0) vérifiant

$$\begin{cases} 2x_0 - y_0 + z_0 = 3 \\ (2x_0, 2y_0, 4z_0) = \lambda(2, -1, 1) \end{cases} \Leftrightarrow \begin{cases} 2x_0 - y_0 + z_0 = 3 \\ 2x_0 = 2\lambda \\ 2y_0 = -\lambda \\ 4z_0 = \lambda \end{cases} \Leftrightarrow \begin{cases} 2x_0 - y_0 + z_0 = 3 \\ x_0 = \lambda \\ y_0 = -\frac{\lambda}{2} \\ z_0 = \frac{\lambda}{4} \end{cases}$$

On en déduit par substitution dans la première équation que

$$2\lambda + \frac{\lambda}{2} + \frac{\lambda}{4} = 3$$

et donc $\lambda = \frac{12}{11}$, soit

$$a_0 = (x_0, y_0, z_0) = \left(\frac{12}{11}, -\frac{6}{11}, \frac{3}{11} \right).$$

Donc f possède un unique point critique sous la contrainte \mathcal{C} .

De plus, par la formule de Taylor, pour $a_0 + h \in \mathcal{C}$,

$$f(a_0 + h) = f(a_0) + \langle \nabla f(a_0), h \rangle + \frac{1}{2} q_{a_0}(h) + \|h\|^2 \varepsilon(h)$$

où $\varepsilon(0) = 0$ et ε est continue en 0.

Mais $a_0 + h \in \mathcal{C} \Leftrightarrow h \in \mathcal{H}$, et donc $\langle \nabla f(a_0), h \rangle = 0$ car $\nabla f(a_0) \in \mathcal{H}^\perp$.

Il suffit donc¹² d'étudier le signe de $q_{a_0}(h)$ pour $h \in \mathcal{H}$.

On a

$$\nabla^2(f)(a_0) = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 4 \end{pmatrix}$$

et donc q_{a_0} est toujours de signe positif. Donc f atteint un minimum local sous la contrainte \mathcal{C} en a_0 .

En fait ce minimum est également un minimum global sous la contrainte \mathcal{C} .

En effet, soit $x \in \mathcal{C}$, et soit $h = x - a_0 \in \mathcal{H}$.

Posons alors $g : t \mapsto f(a_0 + th)$, qui est une fonction \mathcal{C}^2 sur \mathbf{R} , avec $g(0) = f(a_0)$ et $g(1) = f(x)$.

On a alors $g'(t) = \langle \nabla f(a_0 + th), h \rangle$ et $g''(t) = q_{a_0+th}(h)$.

Mais pour tout t , $\nabla^2 g(a_0 + th) = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 4 \end{pmatrix}$, de sorte que $q_{a_0+th}(h) \geq 0$.

Ainsi, g est une fonction convexe, de sorte que g' est croissante sur $[0, 1]$.

Mais $g'(0) = \langle \nabla f(a_0), h \rangle = 0$ car $\nabla f(a_0) \in \mathcal{H}^\perp$ et $h \in \mathcal{H}$.

Et donc, pour tout $t \geq 0$, $g'(t) \geq 0$, ce qui implique que g est croissante sur \mathbf{R}_+ . Et en particulier, on a $g(1) \geq g(0) \Leftrightarrow f(x) \geq f(a_0)$.

Ceci étant valable pour tout $x \in \mathcal{C}$, f admet un minimum global sous la contrainte \mathcal{C} en a_0 .

Afin de déterminer si un point critique (a_1, \dots, a_n) sous la contrainte \mathcal{C} est un extremum global de f sur \mathcal{C} , on peut également étudier le signe de

$$f(x_1, \dots, x_n) - f(a_1, \dots, a_n), (x_1, \dots, x_n) \in \mathcal{C}.$$

Dans ce cas, on pourra utiliser le fait que $(x_1, \dots, x_n) \in \mathcal{C}$ si et seulement si il existe $(h_1, \dots, h_n) \in \mathcal{H}$ tel que $(x_1, \dots, x_n) = (a_1 + h_1, \dots, a_n + h_n)$.

Exemple 19.11

Cherchons le minimum de la somme des carrés de n nombres réels dont la somme vaut n .

Soit donc $f(x_1, \dots, x_n) = \sum_{i=1}^n x_i^2$.

On cherche le minimum de f sous la contrainte $\mathcal{C} : x_1 + \dots + x_n = n$.

On a alors $\nabla f(x_1, \dots, x_n) = (2x_1, \dots, 2x_n)$ et $\mathcal{H}^\perp = \text{Vect}(1, \dots, 1)$.

Un minimum éventuel sera atteint en un point tel qu'il existe $\lambda \in \mathbf{R}$ vérifiant :

$$\begin{cases} \sum_{i=1}^n x_i = n \\ (2x_1, \dots, 2x_n) = \lambda(1, \dots, 1) \end{cases} \Leftrightarrow \begin{cases} \sum_{i=1}^n x_i = n \\ x_1 = \dots = x_n = \frac{\lambda}{2} \end{cases}$$

L'unique solution est donc $x_1 = \dots = x_n = 1$.

Pour vérifier qu'il s'agit d'un minimum global sous la contrainte \mathcal{C} , écrivons :

$$f(1 + h_1, \dots, 1 + h_n) - f(1, \dots, 1) = \sum_{i=1}^n (1 + h_i)^2 - n = 2 \sum_{i=1}^n h_i + \sum_{i=1}^n h_i^2.$$

Mais si $(1 + h_1, \dots, 1 + h_n) \in \mathcal{C}$, alors $(h_1, \dots, h_n) \in \mathcal{H}$, et donc $\sum_{i=1}^n h_i = 0$.

Méthode

Cette méthode n'est vraiment efficace que si toutes les valeurs propres de la hessienne sont de même signe. Dans ce cas, on sait que la forme quadratique q_{a_0} est de signe constant sur \mathbf{R}^n , et a fortiori, sur \mathcal{H} .

¹² Sur le même principe que ce que nous avons fait lors de l'étude de la nature des points critiques, nous pourrions montrer que localement,

$$f(x + h) - f(x)$$

est du signe de $q_x(h)$ si celui-ci est non nul.

Nous n'écrivons pas tous les détails ici, mais la preuve serait la même que celle du théorème 17.49.

Notons qu'en lui-même, ce théorème n'est plus valable, puisque a_0 n'est plus un point critique de f , mais seulement un point critique de f sous contrainte.

Toutefois, le reste de la preuve s'adapte bien, puisque $\langle \nabla f(a_0), h \rangle$ est tout de même nul, bien que ce soit pour des raisons différentes.

Ainsi, on a $f(1 + h_1, \dots, 1 + h_n) - f(1, \dots, 1) = \sum_{i=1}^n h_i^2 \geq 0$.

Donc f atteint bien un minimum global sous la contrainte $\sum_{i=1}^n x_i = n$ en $(1, \dots, 1)$.

19.2.3 Cas particulier : $p = 1$ (contrainte formée d'une seule équation linéaire)

Dans le cas où on cherche les extrema d'une fonction $f : \mathbf{R}^n \rightarrow \mathbf{R}$ sous **une seule** contrainte d'égalité linéaire $\mathcal{C} : a_1x_1 + a_2x_2 + \dots + a_nx_n = 0$, on dispose de trois méthodes différentes :

- on peut utiliser ce que l'on vient de faire pour une contrainte linéaire.
- on peut exprimer l'une des variables en fonction des autres, et étudier alors une fonction $g : \mathbf{R}^{n-1} \rightarrow \mathbf{R}$, et utiliser alors tout ce que l'on sait sur les extrema des fonctions définies sur un ouvert
- on peut utiliser le fait que la contrainte \mathcal{H} est non critique et appliquer le théorème des extrema liés.

Exemple 19.12

Reprenons l'exemple $f(x, y, z) = x^2 + y^2 + 2z^2$ et $\mathcal{C} = 2x - y + z = 3$.

Nous avons déjà prouvé, à l'aide de la première méthode que $\left(\frac{12}{11}, \frac{-6}{11}, \frac{3}{11}\right)$ est le seul point critique de f sous la contrainte \mathcal{C} .

Si l'on applique la seconde méthode, notons que $(x, y, z) \in \mathcal{C}$ si et seulement si $y = 2x + z - 3$.

Soit alors $g : \mathbf{R}^2 \rightarrow \mathbf{R}$ définie par $g(x, z) = f(x, 2x + z - 3, z) = x^2 + (2x + z - 3)^2 + 2z^2$. Alors f admet un extremum local sous la contrainte \mathcal{C} en (x_0, y_0, z_0) si et seulement si g admet un extremum local¹³ en (x_0, z_0) .

Or g est \mathcal{C}^1 sur \mathbf{R}^2 car polynomiale, et on a

$$\partial_1 g(x, y) = 2x + 4(2x + z - 3), \partial_2 g(x, y) = 2(2x + z - 3) + 4z.$$

Ainsi, (x_0, z_0) est un point critique de g si et seulement si

$$\begin{cases} x + 2(2x + z - 3) = 0 \\ 2x + z - 3 + 2z = 0 \end{cases} \Leftrightarrow \begin{cases} x - 4z = 0 \\ 2x + 3z - 3 = 0 \end{cases} \Leftrightarrow \begin{cases} x = \frac{12}{11} \\ z = \frac{3}{11} \end{cases}$$

Ainsi, g admet pour point critique $\left(\frac{12}{11}, \frac{3}{11}\right)$.

De plus, on a alors $\nabla^2 g(x, z) = \begin{pmatrix} 10 & 4 \\ 4 & 6 \end{pmatrix}$ dont les valeurs propres sont $8 + \sqrt{20}$ et $8 - \sqrt{20}$ qui sont toutes deux positives.

Donc g admet en $\left(\frac{12}{11}, \frac{3}{11}\right)$ un minimum local, de sorte que f admet un minimum local en $\left(\frac{12}{11}, 2\frac{12}{11} + \frac{3}{11} - 3, \frac{3}{11}\right) = \left(\frac{12}{11}, -\frac{6}{11}, \frac{3}{11}\right)$.

Enfin, si l'on souhaite utiliser la troisième méthode, notons que la contrainte \mathcal{C} s'écrit $\varphi(x, y, z) = 3$ où $\varphi(x, y, z) = 2x - y + z$. C'est une contrainte non critique car $\nabla\varphi(x, y, z) = (2, -1, 1) \neq 0$.

Alors (x, y, z) est un point critique de f sous la contrainte \mathcal{C} si il existe $\lambda \in \mathbf{R}$ tel que

$$\begin{cases} 2x - y + z = 3 \\ 2x - y + z = 3 \\ \nabla f(x, y, z) = \lambda \nabla \varphi(x, y, z) \end{cases} \Leftrightarrow \begin{cases} 2x - y + z = 3 \\ 2x = 2\lambda \\ 2y = -\lambda \\ 4z = \lambda \end{cases}$$

On retrouve alors le même système qu'avec la première méthode, qui mène bien sûr à la même solution.

¹³ Sans contrainte !

Remarque

L'inconvénient de cette méthode est qu'elle ne dit rien quant à la nature du point critique obtenu, il faut alors l'étudier «à la main».

MÉTHODES DU CHAPITRE 19

✓ MÉTHODE 19.1 : Identifier la nature d'une contrainte

Avant de commencer les calculs, il faut déterminer la situation dans laquelle on se trouve : contrainte non critique ou contrainte d'égalité linéaire ?

- Si la contrainte est définie par plusieurs équations, celles-ci doivent être linéaires (car nous ne disposons pas d'outils adaptés sinon)
- Si la contrainte est formée d'une seule équation qui n'est pas de la forme $a_1x_1 + \dots + a_nx_n = d$, l'équation n'est pas linéaire, et on utilisera donc le théorème des extrema liés (sous réserve que la contrainte soit non critique).
- Dans le cas d'une seule contrainte linéaire, on peut utiliser indifféremment l'une ou l'autre des méthodes.

✓ MÉTHODE 19.2 : Déterminer les points critiques de f sous une contrainte non critique de la forme $\mathcal{C} = \{x \in \mathbf{R}^n : \varphi(x) = c\}$.

- Commencer par justifier soigneusement que f est \mathcal{C}^1 et calculer son gradient.
- Expliciter la fonction φ si elle n'est pas donnée par l'énoncé, vérifier qu'elle est \mathcal{C}^1 et calculer son gradient.
- Montrer que le gradient de φ ne s'annule pas sur \mathcal{C} afin de prouver que la contrainte est non critique.
- Écrire la définition d'un point critique sous contrainte. On obtient alors un système à $n + 1$ équations (dont la première est la contrainte) et à $n + 1$ inconnues : x_1, \dots, x_n et λ .
- Résoudre le système, en gardant en tête que la valeur de λ ne nous intéresse pas.

Si on souhaite étudier la nature (locale ou globale) des points critiques ainsi obtenus, on utilisera toute indication fournie par l'énoncé.

Exercices de TD : 19.2, 19.3, 19.6, 19.9, 19.10, 19.11

✓ MÉTHODE 19.3 : Étudier les extrema d'une fonction continue f sur un fermé borné \mathcal{D}

- Généralement, si l'énoncé demande de prouver que f admet un maximum et un minimum sur \mathcal{D} , il s'agit de vérifier que \mathcal{D} est fermé borné et que f y est continue.
- Déterminer les points critique sur l'«intérieur» de \mathcal{D} , le plus grand ouvert inclus dans \mathcal{D} .
- Si le bord de \mathcal{D} correspond à une contrainte non critique \mathcal{C} , utiliser le théorème des extrema liés pour déterminer les points critiques de f sous la contrainte \mathcal{C} .
- Calculer les images par f de tous les points critiques ainsi obtenus (les «vrais» points critiques se trouvant dans l'intérieur de \mathcal{D} et les points critiques sous la contrainte \mathcal{C}).
- La plus grande image obtenue est le maximum de f , et la plus petite est le minimum de f .

Exercices de TD : 19.3, 19.11

✓ MÉTHODE 19.4 : Déterminer les extrema locaux d'une fonction f sous un système \mathcal{C} de contraintes d'égalités linéaires : $g_1(x) = b_1, \dots, g_p(x) = b_p$.

1. Définir \mathcal{H} l'ensemble des solutions du système homogène associé.
2. Calculer les gradients des g_i . Rappelons qu'il s'agit de former un vecteur avec les coefficients de g_i .
3. Déterminer les points critiques de f sous la contrainte \mathcal{C} . On obtient alors un système formé de $n + p$ équations : les p équations formant la contrainte, ainsi que n équations provenant de $\nabla f(x_1, \dots, x_n) \in \mathcal{H}^\perp$. Notons qu'il y a alors $n + p$ inconnues qui sont x_1, \dots, x_n , ainsi que p nouvelles inconnues $\lambda_1, \dots, \lambda_p$.
4. Pour déterminer la nature locale d'un point critique a , on peut écrire la formule de Taylor à l'ordre 2 au voisinage de a . On n'écrira cette formule que pour $x \in \mathcal{C}$, c'est-à-dire x de la forme $a + h$ avec $h \in \mathcal{H}$. Le terme linéaire ($\langle \nabla f(a), h \rangle$) doit alors s'annuler car $\nabla f(a) \in \mathcal{H}^\perp$.
5. Il reste alors à étudier le signe de $q_a(h)$, pour $h \in \mathcal{H}$. Si on a de la chance, q_a est de signe constant sur \mathbf{R}^n car toutes les valeurs propres de $\nabla^2 f(a)$ sont de même signe. Sinon, il faut faire une étude plus poussée de q_a pour savoir si elle est ou non de signe constant sur \mathcal{H} .

Exercices de TD : 19.1, 19.4, 19.5, 19.7, 19.8

**MÉTHODE 19.5 : Déterminer la nature globale d'un point critique de f sous une contrainte d'égalité linéaire \mathcal{C} .**

Pour déterminer si un point critique $a = (a_1, \dots, a_n)$ sous la contrainte \mathcal{C} est un extremum global, on peut :

- Étudier le signe de $(x_1, \dots, x_n) - f(a_1, \dots, a_n)$. Il est alors utile d'écrire $(x_1, \dots, x_n) = (a_1 + h_1, \dots, a_n + h_n)$, avec $(h_1, \dots, h_n) \in \mathcal{H}$.

En utilisant astucieusement la (ou les) équations définissant \mathcal{H} , on peut essayer de prouver que $f(x_1, \dots, x_n) - f(a_1, \dots, a_n)$ est (ou non) de signe constant.

- Pour $x \in \mathbf{R}^n$, poser $h = x - a$ et étudier la fonction d'une variable $g : t \mapsto f(a + th)$. Cette méthode est particulièrement efficace lorsqu'en tout point de \mathcal{C} , la hessienne ne possède que des valeurs propres positives.

Exercices de TD : 19.1, 19.7, 19.8

EXERCICES DU CHAPITRE 19

EXERCICE 19.1 Déterminer les extrema locaux de $f(x, y) = (x + y)^2 + (x + z)^2 + (y + z)^2$ sous la contrainte $x + y + z = 3$. Ces extrema sont-ils globaux ?

PD

EXERCICE 19.2 Soit $f(x, y) = 2x + y$, et soit \mathcal{C} la contrainte $x^2 + y^2 = 5$.

F

- 1) Montrer que f possède un maximum et un minimum global sur \mathcal{C} .
- 2) Déterminer les points critiques de f sous la contrainte \mathcal{C} .
- 3) Conclure

EXERCICE 19.3 Soit f la fonction définie sur \mathbf{R}^2 par $f(x, y) = xy$.

PD

- 1) Déterminer les points critiques de f sous la contrainte $\mathcal{D} = \{(x, y) \in \mathbf{R}^2 : x^2 + 4y^2 = 8\}$.
- 2) Montrer que $\mathcal{D} = \{(x, y) \in \mathbf{R}^2 : x^2 + 4y^2 = 8\}$ est une partie fermée bornée de \mathbf{R}^2 .
- 3) Montrer que f admet un maximum et un minimum sur \mathcal{D} et les déterminer.
- 4) La fonction f admet-elle un maximum et un minimum sur $\mathcal{D}' = \{(x, y) \in \mathbf{R}^2 : x^2 + 4y^2 \leq 8\}$. Si oui, que valent-ils ?

EXERCICE 19.4 Soit f la fonction de \mathbf{R}^4 dans \mathbf{R} définie par $f(x, y, z, t) = x^2 + y^2 + z^2 + t^2$, et soit \mathcal{C} la contrainte définie par les équations $\begin{cases} x + y = 2 \\ z + t = 0 \end{cases}$.

PD

- 1) Déterminer l'ensemble des points critiques de f sous la contrainte \mathcal{C} .
- 2) Déterminer la nature de ces points critiques.

EXERCICE 19.5 (D'après EDHEC 2001)

AD

Soit $n \geq 2$. Soit $f :]0, 1[^n \rightarrow \mathbf{R}$ définie par $f(x_1, \dots, x_n) = \sum_{i=1}^n (1 - x_i)^n$.

Déterminer les extrema locaux de f sous la contrainte $\mathcal{C} : x_1 + x_2 + \dots + x_n = 1$.

EXERCICE 19.6 Soit $A = (a_{i,j})_{1 \leq i, j \leq n} \in \mathcal{M}_n(\mathbf{R})$ une matrice symétrique, et soit q_A la forme quadratique associée.

AD

On rappelle que $q_A(x_1, \dots, x_n) = \sum_{i=1}^n \sum_{j=1}^n a_{i,j} x_i x_j$.

On note $S = \{x \in \mathbf{R}^n : \|x\|^2 = 1\}$.

- 1) Montrer que q_A est \mathcal{C}^1 sur \mathbf{R}^n , et calculer ses dérivées partielles.

Montrer en particulier que ${}^t \nabla q_A(x_1, \dots, x_n) = 2A \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$.

- 2) Justifier que q_A admet un maximum M et un minimum m sur \mathcal{S} .
- 3) Prouver que la contrainte \mathcal{S} est non critique.

- 4) Montrer que $q_A(x_1, \dots, x_n) = M$ si et seulement si $\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ est un vecteur propre de A associé à la valeur propre M .

- 5) En déduire que M (resp. m) est la plus grande (resp. plus petite) valeur propre de A , et que pour tout $(x_1, \dots, x_n) \in \mathbf{R}^n$,

$$m\|(x_1, \dots, x_n)\|^2 \leq q_A(x_1, \dots, x_n) \leq M\|(x_1, \dots, x_n)\|^2.$$

EXERCICE 19.7 (D'après ECRICOME 2008)

AD

Soit f la fonction définie sur $]0; +\infty[^3$ par $f(x_1, x_2, x_3) = \frac{1}{4x_1} + \frac{1}{x_2} + \frac{1}{9x_3}$.

- 1) Justifier que f est de classe \mathcal{C}^2 , et calculer ses dérivées partielles d'ordre 1 et 2.
- 2) On note $\nabla^2 f(A) = [\partial_{i,j}^2 f(A)]_{1 \leq i, j \leq 3}$ la matrice hessienne de f au point $A = (a_1, a_2, a_3)$. Justifier que pour tout $A \in]0, +\infty[^3$ et pour toute matrice colonne à trois lignes H non nulle, on a ${}^t H \nabla^2 f(A) H > 0$.
- 3) Montrer que f admet un unique point critique a sous la contrainte $\mathcal{C} : x_1 + x_2 + x_3 = 110$.
- 4) On cherche désormais à prouver que a est un extremum global de f sous la contrainte \mathcal{C} . Pour cela, on considère $x \in \mathcal{C}$, et on note alors $h = x - a$ et $g : [0, 1] \rightarrow \mathbf{R}$ la fonction définie par $g(t) = f(a + th)$.

- Exprimer $g(0)$ et $g(1)$ en fonction de f .
- Montrer que $g'(0) = 0$.
- Prouver que pour tout $t \in [0, 1]$, $g''(t) \geq 0$, puis que $g'(t) \geq 0$.
- En déduire que $f(x) \geq f(a)$ et conclure.

EXERCICE 19.8 Soit f la fonction définie sur \mathbf{R}^3 par $f(x, y, z) = xy + yz + xz$, et soit \mathcal{C} la contrainte définie par $x + y + z = 2$.

On note \mathcal{H} l'ensemble des solutions de $x + y + z = 0$.

- Montrer que f admet un unique point critique $a_0 = (x_0, y_0, z_0)$ sous la contrainte \mathcal{C} .
- Déterminer une base (e_1, e_2) de \mathcal{H} .
- Montrer que $\nabla^2 f(x, y, z)$ ne dépend pas de $(x, y, z) \in \mathbf{R}^3$. On note q la forme quadratique associée à cette hessienne.
- La forme quadratique q est-elle de signe constant sur \mathbf{R}^3 ?
- Pour $(\lambda, \mu) \in \mathbf{R}^2$, calculer $q(\lambda e_1 + \mu e_2)$ et en déduire que pour tout $h \in \mathcal{H}$, $q(h) \leq 0$.
- Soit $a \in \mathcal{C}$, soit $h = a - a_0$, et soit $g : t \mapsto f(a_0 + th)$.
 - Rappeler pourquoi g est de classe \mathcal{C}^2 sur \mathbf{R} et donner l'expression de $g'(t)$ et $g''(t)$.
 - À l'aide de la formule de Taylor avec reste intégral, montrer que g admet un maximum en 0.
 - En déduire que f admet un maximum global en a_0 sous la contrainte \mathcal{C} .

EXERCICE 19.9 On souhaite minimiser la surface d'une boîte rectangulaire dont le volume doit être égal à 8 dm^3 . Autrement dit, cela revient à déterminer le minimum de la fonction $f(x, y, z) = 2(xy + yz + xz)$ sur $(\mathbf{R}_+^*)^3$ sous la contrainte $\mathcal{C} : xyz = 8$.

- Montrer que la contrainte \mathcal{C} est non critique.
- Déterminer les points critiques de f sous la contrainte \mathcal{C} .
- Montrer que $\forall (a, b, c) \in (\mathbf{R}_+^*)^3$, $\frac{1}{3}(\ln a + \ln b + \ln c) \leq \ln \left(\frac{a+b+c}{3} \right)$.
- Conclure.

EXERCICE 19.10 Soit $\mathcal{D} = \{(x, y) \in \mathbf{R}^2 : x^2 + y^2 \leq 2\}$, et soit f la fonction définie sur \mathcal{D} par

$$f(x, y) = x^4 + y^4 - 2(x - y)^2.$$

- Montrer que f admet sur \mathcal{D} un maximum M et un minimum m .
- Déterminer l'ensemble des points critiques de f sur l'ouvert $\mathcal{D}' = \{(x, y) \in \mathbf{R}^2 : x^2 + y^2 < 2\}$.
- Déterminer les points critiques de f sous la contrainte $\mathcal{C} = \{(x, y) \in \mathbf{R}^2 : x^2 + y^2 = 2\}$.
- Déterminer m et M .

EXERCICE 19.11 (Oral ESCP 2015)

Soit $n \in \mathbf{N}^*$, et a_1, \dots, a_n des réels non tous nuls. On considère les fonctions f, g et h définies sur \mathbf{R}^n par

$$f(x_1, \dots, x_n) = \prod_{k=1}^n x_k^2, \quad g(x_1, \dots, x_n) = \sum_{k=1}^n x_k^2 \quad \text{et} \quad h(x_1, \dots, x_n) = \sum_{k=1}^n a_k x_k.$$

- Justifier que f est de classe \mathcal{C}^1 sur \mathbf{R}^n . Déterminer son gradient en tout point.
 - Prouver que f admet un maximum sur la sphère unité S de \mathbf{R}^n définie par

$$S = \left\{ (x_1, \dots, x_n) \in \mathbf{R}^n : \sum_{k=1}^n x_k^2 = 1 \right\}.$$

- Déterminer le maximum de f sur S .

- En déduire que pour tout point $u = (u_1, \dots, u_n)$ de \mathbf{R}^n , on a $\left| \prod_{k=1}^n u_k \right| \leq \left(\frac{\|u\|}{\sqrt{n}} \right)^n$ où $\|\cdot\|$ désigne la norme usuelle de \mathbf{R}^n .

- Soit $i \in \llbracket 1, n \rrbracket$ tel que $a_i \neq 0$. On pose

$$H = \{(x_1, \dots, x_n) \in \mathbf{R}^n : h(x_1, \dots, x_n) = 1\} \quad \text{et} \quad B = \left\{ x \in \mathbf{R}^n : \|x\| \leq \frac{1}{|a_i|} \right\}$$

Prouver que l'ensemble $B \cap H$ est non vide, puis qu'il est fermé borné.

- Justifier que la fonction g admet un minimum sur $B \cap H$. Prouver que ce minimum est aussi le minimum de g sous la contrainte $h(x_1, \dots, x_n) = 1$.
- Déterminer le minimum de g sur H .

CORRECTION DES EXERCICES DU CHAPITRE 19

SOLUTION DE L'EXERCICE 19.1

La fonction f est de classe \mathcal{C}^2 sur \mathbf{R}^2 car polynomiale, et la contrainte $x + y + z = 3$ est linéaire. De plus, on a

$$\nabla f(x, y, z) = (4x + 2y + 2z, 2x + 4y + 2z, 2x + 2y + 4z).$$

Notons $\mathcal{H} = \{(x, y, z) \in \mathbf{R}^3 : x + y + z = 0\}$.

Alors $(x, y, z) \in \mathcal{C}$ est un point critique de f sous la contrainte \mathcal{C} si et seulement si $\nabla f(x, y, z) \in \mathcal{H}^\perp = \text{Vect}((1, 1, 1))$. Soit si et seulement si il existe $\lambda \in \mathbf{R}$ tel que

$$\begin{cases} x + y + z = 3 \\ \nabla f(x, y, z) = \lambda(1, 1, 1) \end{cases} \Leftrightarrow \begin{cases} x + y + z = 3 \\ 4x + 2y + 2z = \lambda \\ 2x + 4y + 2z = \lambda \\ 2x + 2y + 4z = \lambda \end{cases} \Leftrightarrow \begin{cases} x + y + z = 3 \\ 4x + 2y + 2z = \lambda \\ x = y \\ y = z \end{cases} \Leftrightarrow x = y = z = 1$$

Ainsi, le seul point critique de f sous la contrainte \mathcal{C} est $a = (1, 1, 1)$.

Pour $(x, y, z) \in \mathcal{C} \Leftrightarrow (x, y, z) = (1, 1, 1) + h, h \in \mathcal{H}$, par la formule de Taylor à l'ordre 2 :

$$f(x, y, z) - f(1, 1, 1) = \underbrace{\langle \nabla f(1, 1, 1), h \rangle}_{=0} + \frac{1}{2} q_a(h) + \|h\|^2 \varepsilon(h).$$

Et donc, pour h au voisinage de 0, c'est-à-dire pour $(x, y, z) \in \mathcal{C}$ «proche» de $(1, 1, 1)$, le signe de $f(x, y, z) - f(1, 1, 1)$ est celui de $q_a(h)$.

$$\text{Or, } \nabla^2 f(1, 1, 1) = \begin{pmatrix} 4 & 2 & 2 \\ 2 & 4 & 2 \\ 2 & 2 & 4 \end{pmatrix} = A.$$

Un réel λ est valeur propre de $\nabla^2 f(1, 1, 1)$ si et seulement si $\begin{pmatrix} 4 - \lambda & 2 & 2 \\ 2 & 4 - \lambda & 2 \\ 2 & 2 & 4 - \lambda \end{pmatrix}$ n'est pas inversible. Or

$$\begin{pmatrix} 4 - \lambda & 2 & 2 \\ 2 & 4 - \lambda & 2 \\ 2 & 2 & 4 - \lambda \end{pmatrix} \xleftrightarrow{L_1 \leftrightarrow L_3} \begin{pmatrix} 2 & 2 & 4 - \lambda \\ 2 & 4 - \lambda & 2 \\ 4 - \lambda & 2 & 2 \end{pmatrix} \xleftrightarrow{L_2 \leftarrow L_2 - L_1, L_3 \leftarrow L_3 - L_1} \begin{pmatrix} 2 & 4 - \lambda & 2 \\ 0 & 2 - \lambda & \lambda - 2 \\ 0 & -4 + 2\lambda & -\lambda^2 + 10\lambda - 16 \end{pmatrix} \xleftrightarrow{L_3 \leftarrow L_3 + 2L_2} \begin{pmatrix} 2 & 4 - \lambda & 2 \\ 0 & 2 - \lambda & \lambda - 2 \\ 0 & 0 & -\lambda^2 + 8\lambda - 12 \end{pmatrix}$$

Le rang de $\nabla^2 f(1, 1, 1) - \lambda I_3$ est donc inférieur à 3 si et seulement si la matrice échelonnée que nous venons d'obtenir est de rang inférieur à 3.

Ainsi, λ est valeur propre de $\nabla^2 f(1, 1, 1)$ si et seulement si $\lambda = 2$ ou si $\lambda^2 - 10\lambda + 16 = 0 \Leftrightarrow \lambda = 8$ ou $\lambda = 2$.

Les valeurs propres de $\nabla^2 f(1, 1, 1)$ sont toutes positives, de sorte que, au voisinage de $(1, 1, 1)$, le signe de $f(x, y, z) - f(1, 1, 1)$ est le même que celui de $q_a(h)$, qui est positif. Ainsi, f admet un minimum local en $(1, 1, 1)$ sous la contrainte \mathcal{C} .

Si $(x, y, z) \in \mathcal{C}$, alors $(x, y, z) = (1 + h_1, 1 + h_2, 1 + h_3)$, avec $(h_1, h_2, h_3) \in \mathcal{H} \Leftrightarrow h_1 + h_2 + h_3 = 0$. On a alors

$$\begin{aligned} f(x, y, z) - f(1, 1, 1) &= (2 + h_1 + h_2)^2 + (2 + h_1 + h_3)^2 + (2 + h_3 + h_2)^2 - 12 \\ &= 2h_1^2 + 2h_2^2 + 2h_3^2 + 2h_1h_2 + 2h_1h_3 + 2h_2h_3 \\ &= h_1(h_1 + h_2 + h_3) + h_2(h_1 + h_2 + h_3) + h_3(h_1 + h_2 + h_3) + h_1^2 + h_2^2 + h_3^2 \\ &= h_1^2 + h_2^2 + h_3^2 \geq 0 \end{aligned}$$

Ainsi, f atteint en $(1, 1, 1)$ un minimum global sous la contrainte \mathcal{C} .

Remarque

On a choisi ici d'utiliser les résultats du cours, mais il aurait également été possible de réduire l'étude de f sur \mathcal{C} à l'étude sur \mathbf{R}^2 de la fonction de deux variables

$$g(x, y) = f(x, y, 3 - x - y).$$

 \mathcal{H}^\perp

Si on reprend les notations du cours, on a \mathcal{H}^\perp qui est engendré par le gradient de la forme linéaire

$$g_1(x, y, z) = x + y + z.$$

Et ce gradient vaut $(1, 1, 1)$.

Produit scalaire

Par définition d'un point critique sous contrainte, on a

$$\nabla f(1, 1, 1) \in \mathcal{H}^\perp.$$

Puisque $h \in \mathcal{H}$, on a donc

$$\langle \nabla f(1, 1, 1), h \rangle = 0.$$

Astuce

Si on veut éviter un pivot, il est facile de voir que $\lambda = 2$ est valeur propre de A car $A - 2I_3$ est de rang 1. Et donc $\dim E_2(A) = 2$, un argument de trace nous permet alors de trouver que l'autre valeur propre de A est 8.

Remarque

Notons que ceci nous dispense d'une étude locale, et notamment du calcul de la hessienne, pour établir un résultat plus fort !

SOLUTION DE L'EXERCICE 19.2

- \mathcal{C} est un fermé borné¹, et f est une fonction continue (car polynomiale) sur \mathcal{C} . Elle y admet donc un maximum et un minimum.
- Soit $\varphi(x, y) = x^2 + y^2$. Alors φ est de classe \mathcal{C}^1 , avec $\nabla\varphi(x, y) = (2x, 2y)$, qui ne s'annule pas sur \mathcal{C} . Donc $\mathcal{C} = \{(x, y) \in \mathbf{R}^2 : \varphi(x, y) = 5\}$ est une contrainte non critique.

De plus, f est de classe \mathcal{C}^1 , avec $\nabla f(x, y) = (2, 1)$.

Donc (x, y) est un point critique de f sous la contrainte \mathcal{C} si et seulement si il existe un réel λ tel que

$$\begin{cases} x^2 + y^2 = 5 \\ 2\lambda x = -2 \\ 2\lambda y = -1 \end{cases}$$

En ajoutant les carrés des deux dernières équations, il vient $4\lambda^2(x^2 + y^2) = 5$, et puisque $x^2 + y^2 = 5$, il vient $4\lambda^2 = 1$.

Et donc $\lambda = \pm \frac{1}{2}$.

• Si $\lambda = \frac{1}{2}$, alors $x = -2$ et $y = -1$.

• Si $\lambda = -\frac{1}{2}$, alors $x = 2$ et $y = 1$.

Donc f possède deux points critiques sous la contrainte \mathcal{C} .

- On a $f(2, 1) = 5$ et $f(-2, -1) = -5$, qui sont donc respectivement le maximum et le minimum de f sous la contrainte \mathcal{C} .

¹ C'est le cercle de centre $(0, 0)$ et de rayon $\sqrt{5}$.

Non critique

Le gradient de φ s'annule en $(0, 0)$. Mais $(0, 0)$ n'est pas dans \mathcal{C} , donc $\nabla\varphi$ ne s'annule pas sur \mathcal{C} : c'est donc une contrainte non critique.

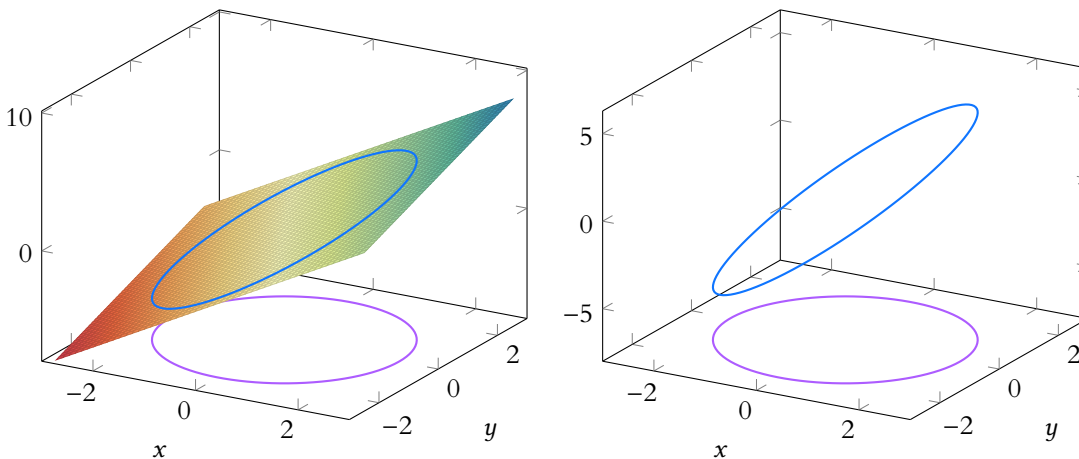


FIGURE 19.1 – En violet : la contrainte \mathcal{C} . En bleu : les valeurs prises par f sur \mathcal{C} .

SOLUTION DE L'EXERCICE 19.3

- Notons que f est de classe \mathcal{C}^1 sur l'ouvert \mathbf{R}^2 , avec $\nabla f(x, y) = (y, x)$. La contrainte \mathcal{D} est de la forme $\varphi(x, y) = 8$, où $\varphi(x, y) = x^2 + 4y^2$. Or $\nabla\varphi(x, y) = (2x, 8y)$ ne s'annule pas sur \mathcal{D} , donc il s'agit d'une contrainte non critique. (x, y) est un point critique de f sous la contrainte \mathcal{D} si et seulement si il existe $\lambda \in \mathbf{R}$ tel que

$$\begin{cases} x^2 + 4y^2 = 8 \\ y = 2\lambda x \\ x = 8\lambda y \end{cases} \Leftrightarrow \begin{cases} x^2 + 4y^2 = 8 \\ y = 2\lambda x \\ x = 16\lambda^2 x \end{cases}$$

De la dernière équation, il vient $x = 0$ ou $\lambda^2 = \frac{1}{16}$, soit $\lambda = \pm \frac{1}{4}$.

Si $x = 0$, alors $y = 0$, et la première équation n'est pas vérifiée. Donc nécessairement, $\lambda^2 = \frac{1}{16}$ et donc $\lambda = \pm \frac{1}{4}$.

Dans ce cas, on a alors $y = \pm \frac{1}{2}x$ et donc $y^2 = \frac{1}{4}x^2$.

Par substitution dans la première équation, on obtient $2x^2 = 8 \Leftrightarrow x = \pm 2$. Ainsi, f admet quatre points critiques sous la contrainte \mathcal{D} : $(2, 1)$, $(2, -1)$, $(-2, 1)$ et $(-2, -1)$.

- \mathcal{D} est un fermé car défini par $\varphi(x, y) = 8$, où φ est une fonction continue.
Il est borné car si $(x, y) \in \mathcal{C}$, alors $x^2 \leq 8 \Leftrightarrow |x| \leq 2\sqrt{2}$ et $y^2 \leq 2 \Leftrightarrow |y| \leq \sqrt{2} \leq 2\sqrt{2}$.
- f est continue sur le fermé borné \mathcal{D} , donc elle y admet un maximum M et un minimum m .
Ces extrema sont alors des extrema de f sous la contrainte \mathcal{D} , et sont donc nécessairement atteints en des points critiques de f sous la contrainte \mathcal{D} .
Or on a $f(2, 1) = f(-2, -1) = 2$ et $f(2, -1) = f(-1, 2) = -2$, donc $m = -2$ et $M = 2$.

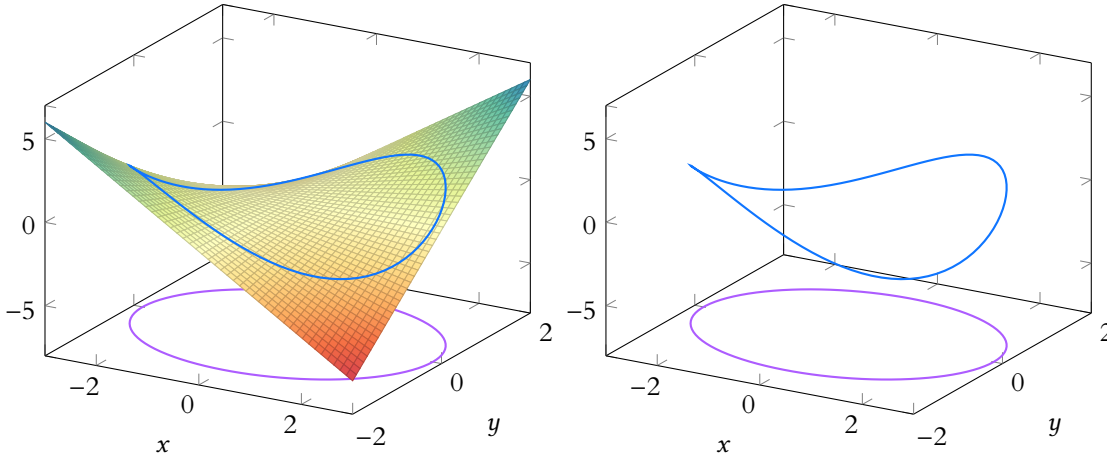


FIGURE 19.2 – En violet : la contrainte³ \mathcal{D} . En bleu : les valeurs prises par f sur \mathcal{D} .

³ Notons que géométriquement, il s'agit d'une ellipse.

- L'ensemble \mathcal{D}' est également un fermé borné sur lequel f est continue, donc elle y admet également un maximum et un minimum.
Pour les déterminer, notons que $\mathcal{D}' = \mathcal{D} \cup \mathcal{D}_0$, où $\mathcal{D}_0 = \{(x, y) \in \mathbb{R}^2 : x^2 + 4y^2 < 8\}$ est un ouvert.
Il est assez facile de constater que le seul point critique de f sur \mathcal{D}_0 est $(0, 0)$, et que $f(0, 0) = 0$.
Or, le maximum de f sur \mathcal{D}' est soit atteint en un point de \mathcal{D}_0 , auquel cas c'est en un point critique, soit en un point de \mathcal{D} , auquel cas c'est en un point critique sous la contrainte \mathcal{D} .
Puisque le maximum de f sur \mathcal{D} est supérieur à $f(0, 0)$, le maximum de f sur \mathcal{D}' vaut également 2, et il est atteint uniquement en $(2, 1)$ et en $(-2, -1)$.
Et de même, le minimum de f sur \mathcal{D}' vaut -2 .

SOLUTION DE L'EXERCICE 19.4

- Notons \mathcal{H} l'ensemble des solutions de $\begin{cases} x + y = 0 \\ z + t = 0 \end{cases}$.

La fonction f est de classe \mathcal{C}^2 sur \mathbb{R}^4 car polynomiale, et

$$\forall (x, y, z, t) \in \mathbb{R}^4, \nabla f(x, y, z, t) = 2(x, y, z, t).$$

$(x, y, z, t) \in \mathcal{C}$ est un point critique de f sous la contrainte \mathcal{C} si et seulement si $\nabla f(x, y, z, t) \in \mathcal{H}^\perp$.

Or, nous savons que $\mathcal{H}^\perp = \text{Vect}((1, 1, 0, 0), (0, 0, 1, 1))$.

Donc (x, y, z, t) est un point critique de f sous la contrainte \mathcal{C} si et seulement si il existe deux réels λ et μ tels que

$$\begin{cases} x + y = 2 \\ z + t = 0 \\ \nabla f(x, y, z, t) = \lambda(1, 1, 0, 0) + \mu(0, 0, 1, 1) \end{cases} \Leftrightarrow \begin{cases} x + y = 2 \\ z + t = 0 \\ 2x = \lambda \\ 2y = \lambda \\ 2z = \mu \\ 2t = \mu \end{cases}$$

On obtient alors $x = y = 1$ et $z = t = 0$.

Donc le seul point critique de f sous la contrainte \mathcal{C} est $a = (1, 1, 0, 0)$.

Méthode
Ce système possède six équations, pour six inconnues, qui sont x, y, z, t et λ, μ . Les valeurs de λ et μ ne nous intéressent pas : il est possible qu'on les obtienne lors de la résolution, mais les déterminer n'est pas un but en soi, ce sont les coordonnées (x, y, z, t) d'un éventuel point critique que nous cherchons.

2. En tout point $(x, y, z, t) \in \mathbf{R}^4$, la hessienne de f est donnée par

$$\nabla^2 f(x, y, z, t) = \begin{pmatrix} 2 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 2 \end{pmatrix} = 2I_4$$

de sorte que la seule valeur propre de $\nabla^2 f(1, 1, 0, 0)$ est 2.

Par conséquent, la forme quadratique q_a ne prend que des valeurs positives sur \mathbf{R}^4 , et donc sur \mathcal{H} (mieux, elle ne prend que des valeurs strictement positives sur $\mathbf{R}^n - \{0\}$).

Mais par la formule de Taylor à l'ordre 2, il existe une fonction ε (continue en 0 et vérifiant $\varepsilon(0) = 0$) telle que

$$\forall h \in \mathbf{R}^n, f(a+h) = f(a) + \langle \nabla f(a), h \rangle + \frac{1}{2} q_a(h) + \|h\|^2 \varepsilon(h).$$

En particulier, si $a+h \in \mathcal{C} \Leftrightarrow h \in \mathcal{H}$, alors $\langle \nabla f(a), h \rangle = 0$ car $\nabla f(a) \in \mathcal{H}^\perp$.

Donc au voisinage⁴ de a , le signe de $f(a+h) - f(a)$ est donné par le signe de la forme quadratique q_a .

Nous avons déjà dit que celle-ci est strictement positive pour $h \neq 0$, et donc f possède un minimum local sous la contrainte \mathcal{C} en $a = (1, 1, 0, 0)$.

SOLUTION DE L'EXERCICE 19.5

Notons que f est polynomiale et donc de classe \mathcal{C}^2 sur l'ouvert $]0, 1[^n$, avec

$$\nabla f(x_1, \dots, x_n) = -n \left((1-x_1)^{n-1}, \dots, (1-x_n)^{n-1} \right).$$

$(x_1, \dots, x_n) \in \mathcal{C}$ est un point critique de f sous la contrainte \mathcal{C} si et seulement si il existe $\lambda \in \mathbf{R}$ tel que

$$\begin{cases} x_1 + \dots + x_n = 1 \\ \nabla f(x_1, \dots, x_n) = \lambda(1, \dots, 1) \end{cases} \Leftrightarrow \begin{cases} x_1 + \dots + x_n = 1 \\ -n(1-x_1)^{n-1} = \lambda \\ \vdots \\ -n(1-x_n)^{n-1} = \lambda \end{cases}$$

En particulier, on doit avoir, pour tout $i \in \llbracket 2, n \rrbracket$, $-n(1-x_i)^{n-1} = -n(1-x_1)^{n-1}$.

Puisque les $1-x_i$ sont strictement positifs, on a donc $1-x_i = 1-x_1 \Leftrightarrow x_i = x_1$.

Et donc tous les x_i sont égaux, de sorte que $x_1 + \dots + x_n = 1 \Leftrightarrow nx_1 = 1 \Leftrightarrow x_1 = \frac{1}{n}$.

Ainsi f admet un unique point critique sous la contrainte \mathcal{C} , qui est $a = \left(\frac{1}{n}, \dots, \frac{1}{n} \right)$.

Pour déterminer la nature du point critique, notons

$$\mathcal{H} = \{(x_1, \dots, x_n) \in \mathbf{R}^n : x_1 + \dots + x_n = 1\}.$$

Alors, nous savons d'une part que $\nabla f(a) \in \mathcal{H}^\perp$, et d'autre part qu'un vecteur (x_1, \dots, x_n) est dans \mathcal{C} si et seulement si

$$(x_1, \dots, x_n) = \left(\frac{1}{n} + h_1, \dots, \frac{1}{n} + h_n \right) = a + h \text{ avec } h = (h_1, \dots, h_n) \in \mathcal{H}.$$

Par la formule de Taylor à l'ordre 2, il existe une fonction ε telle que

$$\forall (x_1, \dots, x_n) \in \mathcal{C}, f(x_1, \dots, x_n) = f(a+(h_1, \dots, h_n)) = f(a) + \underbrace{\langle \nabla f(a), h \rangle}_{=0} + \frac{1}{2} q_a(h) + \|h\|^2 \varepsilon(h).$$

En particulier, au voisinage de a (c'est-à-dire pour h proche de 0), on a $f(x_1, \dots, x_n) - f(a)$ du signe de $q_a(h)$.

Mais on a, pour tout $i \in \llbracket 1, n \rrbracket$

$$\partial_{i,i}^2 f(x_1, \dots, x_n) = n(n-1)(1-x_i)^{n-2}$$

Méthode

Si l'on souhaite justifier la nature du point critique sous la contrainte \mathcal{C} , il est préférable de réécrire la formule de Taylor à l'ordre 2, afin de bien mettre en évidence le fait que, bien que $\nabla f(a)$ ne soit pas nul, le terme $\langle \nabla f(a), h \rangle$ l'est, et que donc c'est le signe de la forme quadratique q_a qui déterminera la nature du point critique.

⁴ C'est-à-dire pour $a+h \in \mathcal{C}$ «proche de a ».

Remarque

Il serait même possible de prouver que ce minimum est global, par exemple en utilisant une méthode semblable à celle de la question 4 de l'exercice 7.

et pour $i \neq j$, $\partial_{i,j}^2 f(x_1, \dots, x_n) = 0$.

Autrement dit, $\nabla^2 f(a)$ est diagonale, et ses valeurs propres⁵ sont toutes strictement positives.

⁵ Les $\partial_{i,i}^2 f(a)$.

Donc la forme quadratique q_a est positive sur \mathbf{R}^n , et a fortiori sur \mathcal{H} .

On en déduit qu'au voisinage de a , $f(x_1, \dots, x_n) - f(a) \geq 0$, et donc f admet un minimum local en a sous la contrainte \mathcal{C} .

SOLUTION DE L'EXERCICE 19.6

1. q_A est polynomiale sur \mathbf{R}^n , et donc y est \mathcal{C}^1 .
Pour $(x_1, \dots, x_n) \in \mathbf{R}^n$, et $k \in \llbracket 1, n \rrbracket$, on a alors

$$q_A(x_1, \dots, x_n) = a_{k,k}x_k^2 + \sum_{j \neq k} a_{k,j}x_kx_j + \sum_{i \neq k} a_{i,k}x_ix_k + \sum_{i \neq j \neq k} a_{i,j}x_ix_j.$$

Il vient alors en dérivant par rapport à x_k :

$$\partial_k q_A(x_1, \dots, x_n) = 2a_{k,k}x_k + \sum_{j \neq k} a_{k,j}x_j + \sum_{i \neq k} a_{i,k}x_i = 2 \sum_{i=1}^n a_{k,i}x_i.$$

En particulier, on reconnaît là le coefficient à la $k^{\text{ème}}$ ligne de $2A \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$.

$$\text{Et donc } {}^t \nabla q_A(x_1, \dots, x_n) = 2A \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}.$$

2. S est un fermé car défini par une égalité, et borné car inclus dans la boule fermée de centre 0 et de rayon 1.
Et donc q_A étant continue sur S , elle y admet un maximum M et un minimum m .

3. Notons $\varphi(x_1, \dots, x_n) = \|(x_1, \dots, x_n)\|^2 = \sum_{i=1}^n x_i^2$, de sorte que $S = \{x \in \mathbf{R}^n : \varphi(x) = 1\}$.

Alors φ est \mathcal{C}^1 sur \mathbf{R}^n , car polynomiale, et $\nabla \varphi(x_1, \dots, x_n) = 2(x_1, \dots, x_n)$.

En particulier, $\nabla \varphi$ ne s'annule qu'en $(0, \dots, 0)$ et donc ne peut pas s'annuler sur S . Ainsi, S est une contrainte non critique.

4. Si $q_A(x_1, \dots, x_n) = M$, alors q_A admet en (x_1, \dots, x_n) un extremum global (donc local) sous la contrainte S , et donc (x_1, \dots, x_n) est un point critique de q_A sous la contrainte S .
Par conséquent, il existe un réel $\lambda \in \mathbf{R}$ tel que

$$\nabla f(x_1, \dots, x_n) = \lambda \nabla \varphi(x_1, \dots, x_n) \Leftrightarrow 2A \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = 2\lambda \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}.$$

Et donc

$$q_A(x_1, \dots, x_n) = (x_1 \quad \dots \quad x_n) A \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = (x_1 \quad \dots \quad x_n) \lambda \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \lambda \|(x_1, \dots, x_n)\|^2 = \lambda.$$

On a alors nécessairement $\lambda = M : \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ est un vecteur propre de A associé à la valeur propre λ .

5. On montrerait de même que $q_A(x_1, \dots, x_n) = m$ si et seulement si ${}^t(x_1 \quad \dots \quad x_n)$ est un vecteur propre de A associé à la valeur propre m .

De plus, si λ est une valeur propre de A , et si $\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ est un vecteur propre unitaire⁶ de A

⁶ C'est-à-dire dans S .

Détails

Puisque A est symétrique, on a $a_{i,k} = a_{k,i}$.

Remarque

On dit que S est la sphère de centre 0 et de rayon 1 : c'est l'ensemble des points dont la distance à $0_{\mathbf{R}^n}$ est égale à 1. Autrement dit, c'est le bord de la boule de centre $0_{\mathbf{R}^n}$ et de rayon 1.

associé à la valeur propre λ , alors $q_A(x) = \lambda \|(x_1, \dots, x_n)\|^2 = \lambda$.

Et donc $m \leq \lambda \leq M$, ce qui prouve bien que m est la plus petite valeur propre de A et M la plus grande.

De plus, si $x = (x_1, \dots, x_n) \in \mathbf{R}^n$ est non nul⁷, alors $\frac{x}{\|x\|} \in S$, de sorte que

$$m \leq q_A\left(\frac{x}{\|x\|}\right) \leq M \Leftrightarrow m \leq \frac{1}{\|x\|^2} q_A(x) \leq M \Leftrightarrow m\|x\|^2 \leq q_A(x) \leq M\|x\|^2.$$

SOLUTION DE L'EXERCICE 19.7

1. Les fonctions $(x_1, x_2, x_3) \mapsto 4x_1$, $(x_1, x_2, x_3) \mapsto x_2$ et $(x_1, x_2, x_3) \mapsto 9x_3$ sont polynomiales, et donc de classe \mathcal{C}^2 sur $]0, +\infty[$ et ne s'y annulent pas, de sorte que leurs inverses sont \mathcal{C}^2 . Et alors f est de classe \mathcal{C}^2 comme somme de fonctions \mathcal{C}^2 .

On a alors

$$\partial_1 f(x_1, x_2, x_3) = -\frac{1}{4x_1^2}, \partial_2 f(x_1, x_2, x_3) = -\frac{1}{x_2^2} \text{ et } \partial_3 f(x_1, x_2, x_3) = -\frac{1}{9x_3^2}.$$

Puis,

$$\partial_{1,1}^2 f(x_1, x_2, x_3) = \frac{1}{2x_1^3}, \partial_{2,2}^2 f(x_1, x_2, x_3) = \frac{2}{x_2^2} \text{ et } \partial_{3,3}^2 f(x_1, x_2, x_3) = \frac{2}{9x_3^2}.$$

Et pour $i \neq j$, $\partial_{i,j}^2 f(x_1, x_2, x_3) = 0$.

2. Les calculs de la question précédente prouvent que pour $A = (a_1, a_2, a_3)$, on a

$$\nabla^2 f(A) = \begin{pmatrix} \frac{1}{2a_1^2} & 0 & 0 \\ 0 & \frac{2}{a_2^2} & 0 \\ 0 & 0 & \frac{2}{9a_3^2} \end{pmatrix}$$

Cette matrice est diagonale, et ses coefficients diagonaux sont strictement positifs⁸, donc elle est diagonalisable, à valeurs propres strictement positives.

Et donc pour tout $H \in \mathcal{M}_{3,1}(\mathbf{R})$ non nul, ${}^t H \nabla^2 f(A) H > 0$.

Remarque : il n'est pas nécessaire de connaître le résultat du cours donnant le signe d'une forme quadratique en fonction du signe des valeurs propres pour en arriver là, il suffit de

remarquer que si $H = \begin{pmatrix} h_1 \\ h_2 \\ h_3 \end{pmatrix}$, alors

$${}^t H \nabla^2 f(A) H = \frac{h_1^2}{2a_1^2} + \frac{2h_2^2}{a_2^2} + \frac{2h_3^2}{9a_3^2} > 0.$$

3. La contrainte est une contrainte d'égalité linéaire. Donc si on note \mathcal{H} l'ensemble des solutions de $x_1 + x_2 + x_3 = 0$, nous savons qu'un point critique de f sous la contrainte \mathcal{C} est un point A tel que $\nabla f(A) \in \mathcal{H}^\perp = \text{Vect}(1, 1, 1)$.

Ainsi, $A = (a_1, a_2, a_3)$ est un point critique de f sous la contrainte \mathcal{C} si et seulement si il existe $\lambda \in \mathbf{R}$ tel que

$$\begin{cases} a_1 + a_2 + a_3 = 110 \\ \nabla f(A) = \lambda(1, 1, 1) \end{cases} \Leftrightarrow \begin{cases} a_1 + a_2 + a_3 = 110 \\ \frac{1}{4a_1^2} = \frac{1}{a_2^2} = \frac{1}{9a_3^2} = \lambda \end{cases}$$

Soit encore $a_2^2 = 4a_1^2$ et $a_3^2 = \frac{4a_1^2}{9}$.

Puisque a_1, a_2, a_3 positifs, on en déduit que $a_2 = 2a_1$ et $a_3 = \frac{2a_1}{3}$.

La première équation devient donc, après substitution,

$$a_1 + 2a_1 + \frac{2a_1}{3} = 110 \Leftrightarrow a_1 = 30.$$

Et donc le seul point critique de f sous la contrainte \mathcal{C} est $(30, 60, 20)$.

⁷ Le cas du vecteur nul est trivial car $q_A(0) = 0$.

Remarque

Ce résultat précise en fait le lemme 17.48 du cours, dans lequel cette formule avait été établie, mais sans faire le lien avec les valeurs propres de A .

⁸ Car a_1, a_2 et a_3 le sont.

- 4.a. On a $g(0) = f(a)$ et $g(1) = f(a + h) = f(a + (x - a)) = f(x)$.
- 4.b. Nous savons que $g'(0) = \langle \nabla f(a), h \rangle$.
Mais $h = x - a \in \mathcal{H}$ et $\nabla f(a) \in \mathcal{H}^\perp$, de sorte que le produit scalaire est nul, et donc $g'(0) = 0$.
- 4.c. Nous savons que $g''(t) = q_{a+th}(h)$. Mais d'après le résultat de la question 3, $q_{a+th}(h) \geq 0$, et ce pour tout $t \in [0, 1]$.
Ainsi, $g''(t) \geq 0$, ce qui implique la croissance de g' sur $[0, 1]$.
Et donc pour tout $t \in [0, 1]$, $g'(t) \geq g'(0) = 0$.
- 4.d. Puisque sa dérivée est positive, g est donc croissante sur $[0, 1]$.
Et en particulier, $f(x) = g(1) \geq g(0) = f(a)$.
Ceci étant vrai quel que soit $x \in \mathcal{C}$, f admet en a un minimum global sous la contrainte \mathcal{C} .

Remarque

Nous venons en fait de prouver qu'une fonction convexe qui possède un point critique admet nécessairement un minimum **global** en ce point critique.

SOLUTION DE L'EXERCICE 19.8

- 1. La fonction f est de classe \mathcal{C}^1 car polynomiale, et on a

$$\forall (x, y, z) \in \mathbf{R}^3, \nabla f(x, y, z) = (y + z, x + z, x + y).$$

De plus, $\mathcal{H}^\perp = \text{Vect}(1, 1, 1)$.

(x_0, y_0, z_0) est un point critique de f sous la contrainte \mathcal{C} si et seulement si il existe $\lambda \in \mathbf{R}$ tel que

$$\begin{cases} x_0 + y_0 + z_0 = 2 \\ x_0 + y_0 + z_0 = 2 \\ \nabla f(x_0, y_0, z_0) = \lambda(1, 1, 1) \end{cases} \Leftrightarrow \begin{cases} x_0 + y_0 + z_0 = 2 \\ y_0 + z_0 = \lambda \\ x_0 + z_0 = \lambda \\ x_0 + y_0 = \lambda \end{cases} \Leftrightarrow \begin{cases} x_0 = y_0 = z_0 = \frac{\lambda}{2} \\ \lambda = \frac{4}{3} \end{cases}$$

Ainsi, l'unique point critique de f sous la contrainte \mathcal{C} est $a_0 = (\frac{2}{3}, \frac{2}{3}, \frac{2}{3})$.

- 2. On a $\mathcal{H} = \{(x, y, -x - y), (x, y) \in \mathbf{R}^2\} = \{x(1, 0, -1) + y(0, 1, -1), (x, y) \in \mathbf{R}^2\} = \text{Vect}(e_1, e_2)$, avec $e_1 = (1, 0, -1), (0, 1, -1)$.

- 3. La hessienne de f en un point $(x, y, z) \in \mathbf{R}^3$ est donnée par $\nabla^2 f(x, y, z) = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}$, qui

ne dépend donc pas de (x, y, z) .

- 4. Notons A la matrice hessienne obtenue précédemment.

Alors $A + I_3 = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}$, qui est de rang 1.

Et donc $-1 \in \text{Spec}(A)$, avec $\dim E_{-1}(A) = 2$.

Puisque A est symétrique, elle est diagonalisable, et possède donc une autre valeur propre λ avec $\dim E_\lambda(A) = 1$.

Mais alors $0 = \text{tr}(A) = -1 \times 2 + \lambda \Leftrightarrow \lambda = 2$.

Ainsi, $\text{Spec}(A) = \{-1, 2\}$. Les valeurs propres étant de signes opposés, q n'est pas de signe constant sur \mathbf{R}^3 .

- 5. Si $h = \lambda e_1 + \mu e_2 = (\lambda, \mu, -\lambda - \mu)$, on a

$$\begin{aligned} q(h) &= (\lambda \quad \mu \quad -\lambda - \mu) \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix} \begin{pmatrix} \lambda \\ \mu \\ -\lambda - \mu \end{pmatrix} = 2\lambda\mu + 2\lambda(-\lambda - \mu) + 2\mu(-\lambda - \mu) \\ &= -2(\lambda^2 + \mu^2 + \lambda\mu) = -\lambda^2 - \mu^2 - (\lambda + \mu)^2 \leq 0. \end{aligned}$$

- 6.a. La fonction g est de classe \mathcal{C}^2 car f l'est, et on a, pour tout $t \in \mathbf{R}$,

$$g'(t) = \langle \nabla f(a_0 + th), h \rangle, g''(t) = q_{a_0+th}(h) = q(h).$$

- 6.b. Par la formule de Taylor avec reste intégral à l'ordre 1, appliquée en 0, on a, pour tout $t \in \mathbf{R}$,

$$g(t) = g(0) + g'(0)t + \int_0^t g''(x)(t - x) dx$$

Signes opposés

Puisque q n'est pas de signe constant sur \mathbf{R}^3 , un éventuel point critique (sans contrainte) de f serait nécessairement un point selle. En revanche, q est de signe constant sur \mathcal{H} , et nous allons prouver dans la suite que ceci impose que tout point critique de f sous la contrainte \mathcal{C} est un extremum local **sous la contrainte** \mathcal{C} .

$$= g(0) + \langle \nabla f(a_0), h \rangle t + \int_0^t q(h)(t-x) dx.$$

Mais puisque $a \in \mathcal{C}$, $h = a - a_0 \in \mathcal{H}$, et donc $\langle \nabla f(a_0), h \rangle = 0$.

D'autre part, $\int_0^t (t-x) dx = \left[-\frac{(t-x)^2}{2} \right]_0^t = \frac{t^2}{2} \geq 0$.

On en déduit donc que pour tout $t \in \mathbf{R}$,

$$g(t) = g(0) + q(h) \frac{t^2}{2} \leq g(0).$$

Et donc g admet en 0 un maximum global.

- 6.c. Pour $a \in \mathcal{C}$, nous venons de prouver que $f(a) = g(1) \leq g(0) = f(a_0)$ et donc f admet en a_0 un maximum global sous la contrainte \mathcal{C} .

SOLUTION DE L'EXERCICE 19.9

1. Posons $\varphi(x, y, z) = xyz$, de sorte que la contrainte \mathcal{C} est définie par

$$\mathcal{C} = \{(x, y, z) \in \mathbf{R}^3 : \varphi(x, y, z) = 8\}.$$

On a alors $\nabla \varphi(x, y, z) = (yz, xz, xy) \neq (0, 0, 0)$ si $(x, y, z) \in (\mathbf{R}_+^*)^3$.
Donc \mathcal{C} est une contrainte non critique.

2. (x, y, z) est un point critique sous la contrainte \mathcal{C} si et seulement si il existe un réel λ tel que

$$\begin{cases} \varphi(x, y, z) = 8 \\ \nabla f(x, y, z) = \lambda \nabla \varphi(x, y, z) \end{cases} \Leftrightarrow \begin{cases} xyz = 8 \\ 2(y+z, x+z, x+y) = \lambda(yz, xz, xy) \end{cases} \Leftrightarrow \begin{cases} xyz = 8 \\ 2y + 2z = \lambda yz \\ 2x + 2z = \lambda xz \\ 2x + 2y = \lambda xy \end{cases}$$

En soustrayant la ligne 3 à la ligne 2, il vient $2(y-x) = \lambda z(y-x)$, soit encore $(y-x)(2-\lambda z) = 0$.
Si $\lambda z = 2$, alors la seconde équation devient $2(y+z) = 2y$, soit $z = 0$, ce qui est impossible.
Donc $y = x$.

De même, on prouve que $y = z$.

Alors nécessairement, en utilisant la première équation, $x^3 = 8$, soit $x = 2$.

Notons qu'on a alors $\lambda = 2$, et donc $(2, 2, 2)$ est bien un point critique (et donc l'unique point critique) de f sous la contrainte \mathcal{C} .

3. Rappelons que la fonction $-\ln$ est convexe sur \mathbf{R}_+^* , de sorte que par l'inégalité de convexité généralisée,

$$-\ln\left(\frac{1}{3}a + \frac{1}{3}b + \frac{1}{3}c\right) \leq -\frac{1}{3}\ln(a) - \frac{1}{3}\ln(b) - \frac{1}{3}\ln(c).$$

Soit encore

$$\frac{1}{3}(\ln a + \ln b + \ln c) \leq \ln\left(\frac{a+b+c}{3}\right).$$

4. Il s'agit à présent de prouver que $f(2, 2, 2)$ est le minimum global de f sous la contrainte \mathcal{C} , autrement dit que

$$\forall (x, y, z) \in \mathcal{C}, f(x, y, z) \geq f(2, 2, 2).$$

Pour tout $(x, y, z) \in \mathcal{C}$, on a, d'après l'inégalité de la question précédente,

$$\ln\left(\sqrt[3]{x^2y^2z^2}\right) = \frac{1}{3}(\ln(xy) + \ln(xz) + \ln(yz)) \leq \ln\left(\frac{xy + yz + xz}{3}\right).$$

Puis par croissance de l'exponentielle,

$$\sqrt[3]{x^2y^2z^2} \leq \frac{1}{3}(xy + yz + xz).$$

Mais si $(x, y, z) \in \mathcal{C}$, alors $xyz = 8$, et donc $\sqrt[3]{x^2y^2z^2} = \sqrt[3]{64} = \sqrt[3]{2^6} = 4$.

Ainsi, $(xy + yz + xz) \geq 12$ et donc $f(x, y, z) \geq 24 = f(2, 2, 2)$.

Ainsi, la boîte rectangulaire de volume 8 ayant la plus petite surface est cubique, de côté 2.

Convexité

Rappelons que pour une fonction convexe f , si les λ_i sont des nombres positifs dont la somme vaut 1, alors

$$f\left(\sum_{i=1}^n \lambda_i x_i\right) \leq \sum_{i=1}^n \lambda_i f(x_i).$$

SOLUTION DE L'EXERCICE 19.10

1. \mathcal{D} est un ensemble fermé car si $\varphi(x, y) = x^2 + y^2$, alors φ est continue sur \mathbf{R}^2 et $\mathcal{D} = \{(x, y) \in \mathbf{R}^2 : \varphi(x, y) \leq 2\}$ est un fermé.
De plus, \mathcal{D} est borné car si $(x, y) \in \mathcal{D}$, alors $x^2 \leq 2$, soit $|x| \leq \sqrt{2}$, et de même $|y| \leq \sqrt{2}$.
Enfin, f est continue sur \mathcal{D} car polynomiale, et donc, comme toute fonction continue sur un fermé borné, f admet un maximum M et un minimum m sur \mathcal{D} .
2. f est de classe \mathcal{C}^2 sur \mathcal{D}' , avec

$$\partial_1 f(x, y) = 4x^3 - 4(x - y), \quad \partial_2 f(x, y) = 4y^3 + 4(x - y).$$

Donc (x, y) est un point critique de f si et seulement si

$$\begin{cases} 4x^3 = 4(x - y) \\ 4y^3 = 4(y - x) \end{cases} \Leftrightarrow \begin{cases} x^3 = x - y \\ y^3 = y - x \end{cases} \xleftrightarrow{L_1 \leftarrow L_1 + L_2} \begin{cases} x^3 + y^3 = 0 \\ y^3 = y - x \end{cases}$$

En particulier, on en déduit que $x = -y$. Et alors, la seconde équation devient $y^3 = 2y$.
Donc soit $y = 0$ (et alors $x = 0$), soit $y = \sqrt{2}$ (et alors $x = -\sqrt{2}$), soit $y = -\sqrt{2}$ (et alors $x = \sqrt{2}$).

Notons que ces deux derniers points ne sont pas dans \mathcal{D}' car

$$(\sqrt{2})^2 + (-\sqrt{2})^2 = 4.$$

Donc f possède un seul point critique sur \mathcal{D}' qui est $(0, 0)$.

3. Commençons par remarquer que \mathcal{C} est une contrainte non critique. En effet, si on pose $\varphi(x, y) = x^2 + y^2$, alors $\nabla\varphi(x, y) = (2x, 2y)$, qui ne peut s'annuler sur \mathcal{C} car si $(x, y) \in \mathcal{C}$, on a toujours soit $x \neq 0$ soit $y \neq 0$.
Ainsi, (x, y) est un point critique de sous la contrainte \mathcal{C} si et seulement si il existe $\lambda \in \mathbf{R}$ tel que

$$\begin{cases} \varphi(x, y) = 2 \\ \nabla f(x, y) = \lambda \nabla \varphi(x, y) \end{cases} \Leftrightarrow \begin{cases} x^2 + y^2 = 2 \\ 4x^3 - 4(x - y) = \lambda 2x \\ 4y^3 + 4(x - y) = \lambda 2y \end{cases} \xleftrightarrow{L_2 \leftarrow (L_2 - L_3)/2} \begin{cases} x^2 + y^2 = 2 \\ 2(x^3 - y^3) - 4(x - y) = \lambda(x - y) \\ 2y^3 + 2(x - y) = \lambda y \end{cases}$$

Mais nous savons que $x^3 - y^3 = (x - y)(x^2 + xy + y^2)$, de sorte que la seconde équation s'écrit encore : $2(x^2 + y^2 + xy)(x - y) - 4(x - y) = \lambda(x - y)$.

• Si $x = y$, cette équation est vérifiée. Mais pour que la première le soit, il faut soit que $x = y = 1$, soit que $x = y = -1$. Dans les deux cas, les deux dernières équations sont vérifiées avec $\lambda = 2$.

Donc nous avons déjà deux points critiques sous la contrainte \mathcal{C} qui sont $(1, 1)$ et $(-1, -1)$.

• Au contraire, si $x \neq y$, alors la seconde équation devient, après division par $x - y$: $2(x^2 + y^2 + xy) - 4 = \lambda$. En remplaçant $x^2 + y^2$ par 2, il vient $2(2 + xy) - 4 = \lambda$, soit $\lambda = 2xy$. En remplaçant dans la troisième équation, on a

$$2y^3 + 2x - 2y = 2xy^2 \Leftrightarrow 2y^2(y - x) = 2(y - x).$$

Comme nous avons supposé que $x \neq y$, alors $y^2 = 1$, donc $y = \pm 1$. Et alors⁹ $y = -x$.

⁹ Car $x^2 + y^2 = 2$ et $x \neq y$.

Nous avons deux autres points critiques de f sous la contrainte \mathcal{C} , qui sont $(-1, 1)$ et $(1, -1)$.

4. Afin de conclure, notons que :

$$f(0, 0) = 0, f(1, 1) = 2, f(-1, -1) = 2, f(1, -1) = -6 = f(-1, 1) = -6.$$

Nous savons que M est le maximum global de f , donc c'est en particulier un maximum local de f .

Si M est atteint en un point $x_0 \in \mathcal{D}'$, alors x_0 est un point critique de f . Et si M est atteint en un point de \mathcal{C} , alors x_0 est un maximum local de f sous la contrainte \mathcal{C} , et donc x_0 est un point critique de f sous la contrainte \mathcal{C} .

Dans tous les cas, M est atteint en l'un des points critiques trouvés dans les questions précédentes.

Le calcul nous montre alors que $M = 2$, atteint en $(1, 1)$ et $(-1, -1)$.

De même, $m = -6$, atteint en $(1, -1)$ et $(-1, 1)$.

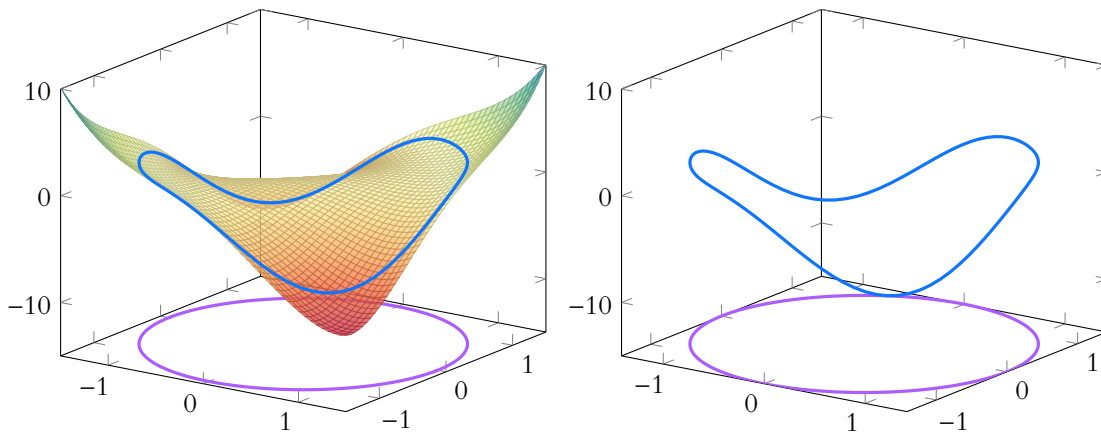


FIGURE 19.3 – En rose : la contrainte \mathcal{C} . En bleu : les valeurs prises par f sur \mathcal{C} .

SOLUTION DE L'EXERCICE 19.11

1.a. f est de classe \mathcal{C}^1 sur \mathbf{R}^n car elle est polynomiale. On a alors

$$\nabla f(x_1, \dots, x_n) = (2x_1x_2^2 \cdots x_n^2, 2x_1^2x_2 \cdots x_n^2, 2x_1^2 \cdots x_{n-1}^2x_n).$$

Autrement dit, pour tout $i \in \llbracket 1, n \rrbracket$, on a

$$\partial_i f(x_1, \dots, x_n) = 2x_i \prod_{\substack{k=1 \\ k \neq i}}^n x_k^2.$$

1.b. S est un fermé car g est continue¹⁰ et que $S = \{(x_1, \dots, x_n) \in \mathbf{R}^n : g(x_1, \dots, x_n) = 1\}$.
Si $(x_1, \dots, x_n) \in S$, alors pour tout $i \in \llbracket 1, n \rrbracket$, on a

¹⁰ Car polynomiale.

$$x_i^2 = 1 - \sum_{\substack{j=1 \\ j \neq i}}^n x_j^2 \leq 1 \Rightarrow |x_i| \leq 1.$$

Puisque f est continue sur le fermé borné S , elle y admet un maximum.

1.c. La contrainte S est non critique car $\nabla g(x_1, \dots, x_n) = (2x_1, \dots, 2x_n)$ ne s'annule pas sur S . Par conséquent, (x_1, \dots, x_n) est un point critique de f sous la contrainte S si et seulement si il existe $\lambda \in \mathbf{R}$ tel que

$$\begin{cases} x_1^2 + \cdots + x_n^2 = 1 \\ \nabla f(x_1, \dots, x_n) = \lambda \nabla g(x_1, \dots, x_n) \end{cases} \Leftrightarrow \begin{cases} x_1^2 + \cdots + x_n^2 = 1 \\ 2x_1x_2^2 \cdots x_n^2 = 2\lambda x_1 \\ \vdots \\ 2x_1^2x_2 \cdots x_n = 2\lambda x_n \end{cases}$$

Soit l'un des x_i est nul, soit on a $x_2^2 \cdots x_n^2 = \lambda = x_1^2x_3^2 \cdots x_n^2$, et donc $x_1^2 = x_2^2$.
De même, on prouve que $x_1^2 = x_2^2 = \cdots = x_n^2$.

Et donc la première équation devient $nx_1^2 = 1 \Leftrightarrow x_1 = \pm \sqrt{\frac{1}{n}}$.

Dans ce cas, on a alors $f(x_1, \dots, x_n) = \frac{1}{n^n}$.

En revanche, si l'un des x_i est nul, on a $f(x_1, \dots, x_n) = 0$.

Et donc le maximum de f est $\frac{1}{n^n}$.

1.d. Si $u \in \mathbf{R}^n$ est non nul, alors $\frac{u}{\|u\|} \in S$. Et donc

$$f(u_1, \dots, u_n) = \|u\|^{2n} f\left(\frac{u_1}{\|u\|}, \dots, \frac{u_n}{\|u\|}\right) \leq \|u\|^{2n} \frac{1}{n^n} = \left(\frac{\|u\|}{\sqrt{n}}\right)^{2n}.$$

Points critiques

Nous n'avons pas réellement déterminé les points critiques. Mais dit qu'il ne pouvaient qu'être de deux types : l'un des x_i est nul, et donc $f(x_1, \dots, x_n) = 0$. Soit tous les x_i^2 valent $\frac{1}{n}$ (et alors $f(x_1, \dots, x_n) = \frac{1}{n^n}$).
Le maximum de f sur S n'étant clairement pas 0, et devant être atteint en un point critique (sous la contrainte S), il vaut nécessairement $\frac{1}{n^n}$.

En prenant la racine carrée, il vient

$$\sqrt{f(u_1, \dots, u_n)} = \prod_{i=1}^n \sqrt{u_i^2} = \prod_{i=1}^n |u_i| = \left| \prod_{i=1}^n u_i \right| \leq \left(\frac{\|u\|}{\sqrt{n}} \right)^n.$$

2.a. Le vecteur (x_1, \dots, x_n) où $x_j = \begin{cases} 0 & \text{si } i \neq j \\ \frac{1}{a_i} & \text{si } i = j \end{cases}$ vérifie

$$h(x_1, \dots, x_n) = a_i \frac{1}{a_i} = 1 \text{ et } \|(x_1, \dots, x_n)\| = \sqrt{\sum_{i=1}^n x_i^2} = \sqrt{\frac{1}{a_i^2}} = \frac{1}{|a_i|} \leq \frac{1}{|a_i|}.$$

Il est donc à la fois dans B et dans H , et donc dans $B \cap H$. On en déduit que $B \cap H$ est non vide.

C'est un fermé car intersection de B et de H qui sont fermés. En effet h est continue et donc $H = \{x \in \mathbf{R}^n : h(x) = 1\}$ est fermé. De même, $x \mapsto \|x\|$ est continue sur \mathbf{R}^n et donc $B = \{x \in \mathbf{R}^n : \|x\| \leq \frac{1}{|a_i|}\}$ est fermé.

Enfin, $B \cap H$ est borné car inclus dans B qui est borné, puisqu'il s'agit de la boule fermée de centre $(0, 0)$ et de rayon $\frac{1}{|a_i|}$.

2.b. g est continue sur le fermé borné $B \cap H$, donc y admet un minimum.

Nous avons prouvé que $(0, \dots, 0, \frac{1}{a_i}, 0, \dots, 0)$ est dans $B \cap H$.

Or $g(0, \dots, 0, \frac{1}{a_i}, 0, \dots, 0) = \frac{1}{a_i^2}$. On en déduit que le minimum de g sur $B \cap H$ est inférieur ou égal à $\frac{1}{a_i^2}$.

Si $x \in \mathbf{R}^n$ est dans H mais pas dans B , alors $\|x\| > \frac{1}{a_i}$ et donc $g(x) = \|x\|^2 > \frac{1}{a_i^2}$.

Ainsi le minimum de g sur $B \cap H$ est également le minimum de g sur H tout entier.

2.c. Déterminons les points critiques de g sous la contrainte $H : (x_1, \dots, x_n)$ est un point critique de g sous la contrainte H si et seulement si il existe $\lambda \in \mathbf{R}$ tel que

$$\begin{cases} h(x_1, \dots, x_n) = 1 \\ \nabla g(x_1, \dots, x_n) = \lambda(a_1, \dots, a_n) \end{cases} \Leftrightarrow \begin{cases} h(x_1, \dots, x_n) = 1 \\ 2x_1 = \lambda a_1 \\ \vdots \\ 2x_n = \lambda a_n \end{cases}$$

Par substitution dans la première équation, il vient alors

$$\frac{\lambda}{2} \sum_{i=1}^n a_i^2 = 1 \Leftrightarrow \lambda = \frac{2}{\sum_{i=1}^n a_i^2}.$$

Et donc $\forall i \in \llbracket 1, n \rrbracket, x_i = \frac{a_i}{\sum_{j=1}^n a_j^2}$.

Ainsi, g admet un unique point critique sous la contrainte H . Celui-ci correspond nécessairement au minimum de g sous la contrainte H , qui vaut donc

$$g\left(\frac{a_1}{\sum_{i=1}^n a_i^2}, \dots, \frac{a_n}{\sum_{i=1}^n a_i^2}\right) = \frac{1}{\sum_{i=1}^n a_i^2}.$$

RÉVISIONS D'ANALYSE

Exercices 584
Correction des exercices 589
EML 2015 609
Correction 611
ECRICOME 2009 617
Correction 618
ECRICOME 2007 622
Correction 623

EXERCICES DU CHAPITRE 20

FONCTIONS D'UNE VARIABLE RÉELLE

EXERCICE 20.1 Prolongement \mathcal{C}^1

PD

Soit f la fonction définie sur $\left[0, \frac{\pi}{2}\right]$ par $f(x) = \begin{cases} \frac{1}{x} - \frac{1}{\sin x} & \text{si } x \neq 0 \\ 0 & \text{si } x = 0 \end{cases}$

- 1) Montrer que f est continue sur $\left[0, \frac{\pi}{2}\right]$.
- 2) Montrer que f est \mathcal{C}^1 sur $\left]0, \frac{\pi}{2}\right[$, et calculer sa dérivée.
- 3) Déterminer $\lim_{x \rightarrow 0^+} f'(x)$.
- 4) Montrer que f est de classe \mathcal{C}^1 sur $\left[0, \frac{\pi}{2}\right]$ et préciser $f'(0)$.

EXERCICE 20.2 Inégalités de convexité

PD

- 1) Montrer que pour tout $u > -1$, $\ln(1+u) \leq u$.
- 2) Montrer que pour tout $x \in \mathbf{R}$, $e^x \geq 1+x$.
- 3) Montrer que la fonction sin est concave sur $\left[0, \frac{\pi}{2}\right]$. En déduire que pour tout $x \in \left[0, \frac{\pi}{2}\right]$, on a $\frac{2}{\pi}x \leq \sin x \leq x$.

EXERCICE 20.3 Soit $f : \mathbf{R} \rightarrow \mathbf{R}$ une fonction convexe dérivable.

PD

- 1) Montrer que si $f'(a) = 0$, alors f possède un minimum global en a .
- 2) (★) On suppose à présent que f est deux fois dérivable et qu'il existe $\alpha > 0$ tel que pour tout $x \in \mathbf{R}$, $f''(x) \geq \alpha$.
On pose alors $g : x \mapsto f'(x) - \alpha x$.
 - a) Déterminer le sens de variation de g .
 - b) Montrer que pour $x \geq 0$, on a $f'(x) \geq f'(0) + \alpha x$ et que pour $x \leq 0$, $f'(x) \leq f'(0) + \alpha x$.
 - c) En déduire que f admet un minimum sur \mathbf{R} .

EXERCICE 20.4 (Ecricome 2004)

AD

On considère la fonction f définie sur l'intervalle $I = \left[0, \frac{\pi}{4}\right]$ par : $f(x) = \frac{1}{\cos x}$.

- 1) Montrer que f réalise une bijection de I dans un intervalle J que l'on précisera. On note f^{-1} la bijection réciproque.
- 2) Donner sur le même graphique l'allure des courbes représentatives de f et de f^{-1} .
- 3) Justifier que :

$$\forall x \in J, \quad \begin{cases} \cos(f^{-1}(x)) = \frac{1}{x} \\ \sin(f^{-1}(x)) = \sqrt{1 - \frac{1}{x^2}} \end{cases}$$

- 4) Montrer que f^{-1} est dérivable sur $J \setminus \{1\}$ et montrer que :

$$\forall x \in J \setminus \{1\}, \quad (f^{-1})'(x) = \frac{1}{x\sqrt{x^2-1}}$$

- 5) En déduire le développement limité en $\sqrt{2}$ de f^{-1} à l'ordre 1.

EXERCICE 20.5 Soit $P \in \mathbf{R}[X]$, de degré $n \geq 1$ et admettant n racines réelles distinctes. Montrer que P' possède $n-1$ racines réelles.

AD

EXERCICE 20.6 Soit $n \in \mathbf{N}^*$ et soit f la fonction définie par $f(x) = x^{n-1} \ln x$. Calculer la dérivée $n^{\text{ième}}$ de f .

AD

EXERCICE 20.7 Soient ω_1, ω_2 deux réels non nuls distincts, et soient a_1, a_2 deux réels non nuls.

TD

On note f la fonction définie sur \mathbf{R} par $f(x) = a_1 \cos(\omega_1 x) + a_2 \cos(\omega_2 x)$

- 1) Montrer que si $|a_1| \neq |a_2|$, alors f s'annule une infinité de fois sur \mathbf{R} .
- 2) Montrer que ce résultat reste valable si $|a_1| = |a_2|$.

FORMULES DE TAYLOR, DÉVELOPPEMENTS LIMITÉS

EXERCICE 20.8 Montrer, à l'aide de la formule de Taylor avec reste intégral, que

PD

$$\forall x \in \left[0, \frac{\pi}{2}\right], x - \frac{x^3}{3!} \leq \sin x \leq x.$$

EXERCICE 20.9 Montrer que pour tout $x \in \mathbf{R}$, $|e^x - 1 - x| \leq \frac{x^2}{2} e^{|x|}$.

PD

EXERCICE 20.10

AD

1) Soit Φ la fonction définie sur \mathbf{R}_+^* par $\Phi(x) = \ln(x)$. Montrer que pour tout $k \geq 1$ et pour tout $x > 0$,

$$\Phi^{(k)}(x) = \frac{(-1)^{k-1}(k-1)!}{x^k}.$$

2) Montrer que pour tout $t \in [0, 1]$, $\forall n \in \mathbf{N}^*$,

$$\left| \ln(1+t) + \sum_{k=1}^n \frac{(-1)^k}{k} t^k \right| \leq \frac{t^{n+1}}{n+1}.$$

3) En déduire que la série de terme général $\frac{(-1)^n}{n}$ converge et donner sa somme.

EXERCICE 20.11 Déterminer les développements limités suivants :

PD

1. DL₃(0) de $\frac{\cos x}{1+x}$ 2. DL₃(0) de $x \mapsto e^x \sqrt{1-x}$ 3. DL₅(0) de $x \mapsto \sin(x) \cos(x)$

EXERCICE 20.12 Déterminer, en utilisant des développements limités, les limites suivantes :

AD

1. $\lim_{x \rightarrow 0} \frac{e^x - \sqrt{1+2x}}{x^2}$ 2. $\lim_{x \rightarrow \infty} \sqrt[3]{x^3 + x^2} - \sqrt[3]{x^3 - x^2}$ 3. $\lim_{n \rightarrow \infty} n^2 \sin\left(\frac{1}{n}\right) \left(\sqrt[n]{e} - \frac{n+1}{n}\right)$
 4. $\lim_{x \rightarrow 0} \frac{\cos x - \sqrt{1-x^2}}{x^4}$

EXERCICE 20.13 Calcul de l'intégrale de Gauss à l'aide d'une fonction définie par une intégrale

D

1) Montrer que $\int_0^{+\infty} e^{-t^2} dt$ est convergente. On note A sa valeur.

2) Soit f la fonction définie sur \mathbf{R} par $f(x) = \int_0^1 \frac{e^{-x(1+t^2)}}{1+t^2} dt$.

a) En utilisant l'inégalité de Taylor-Lagrange, montrer qu'il existe une constante $M > 0$ telle que

$$\forall h \in [-1, 1], \forall t \in [0, 1], |e^{-h(1+t^2)} - 1 + h(1+t^2)| \leq Mh^2.$$

b) En déduire que f est dérivable sur \mathbf{R} et que, $\forall x \in \mathbf{R}$, $f'(x) = -\int_0^1 e^{-x(1+t^2)} dt$.

3) On pose $g(x) = f(x^2) + \left(\int_0^x e^{-t^2} dt\right)^2$. Montrer que g est constante sur \mathbf{R} .

4) Montrer que pour tout $x \geq 0$, on a $0 \leq f(x) \leq \frac{\pi}{4} e^{-x}$ et en déduire $\lim_{x \rightarrow +\infty} f(x)$.

5) En déduire la valeur de A .

INTÉGRALES

EXERCICE 20.14 Déterminer la nature des intégrales suivantes :

F

1. $\int_1^{+\infty} \frac{1}{x + \ln(x)} dx$. 2. $\int_0^{+\infty} \sin t \sin\left(\frac{1}{t^2}\right) dt$. 3. $\int_0^1 \frac{dt}{(1-t)\sqrt{t}}$. 4. $\int_0^{+\infty} (\ln t)e^{-t} dt$.
 5. (*) $\int_0^{+\infty} \frac{\ln(1+t)}{t^{3/2}} dt$. 6. (***) $\int_1^{+\infty} \frac{\ln(t^2-t)}{(1+t)^2} dt$. 7. (*) $\int_0^{+\infty} \frac{u^{\alpha-1}}{e^u-1} du$, $\alpha > 1$.

EXERCICE 20.15

PD

- 1) Étudier la nature de l'intégrale $I = \int_0^{+\infty} \frac{\ln t}{t^2 + 1} dt$.
- 2) En utilisant le changement de variable $u = \frac{1}{t}$, montrer que $I = 0$.
- 3) Soit $a > 0$. Montrer que

$$\int_0^{+\infty} \frac{\ln t}{t^2 + a^2} dt = \frac{\pi \ln a}{2a}.$$

EXERCICE 20.16 Vu dans EML 2005

Soit f une fonction de classe \mathcal{C}^1 sur un segment $[a, b]$. Montrer, à l'aide d'une intégration par parties, que

$$\lim_{n \rightarrow +\infty} \int_a^b f(t) \sin(nt) dt = 0.$$

EXERCICE 20.17 Soit (u_n) la suite définie par : $\forall n \geq 1, u_n = \int_0^{+\infty} \frac{\cos t}{t^2 + n} dt$.

- 1) Montrer que (u_n) est bien définie et que pour tout $n \geq 1, |u_n| \leq \int_0^{+\infty} \frac{dt}{t^2 + n}$.
- 2) En déduire que (u_n) est convergente, et déterminer sa limite.

EXERCICE 20.18 (Extrait de EML 2006)

- 1) Montrer que pour tout $n \in \mathbf{N}, I_n = \int_{-\infty}^{+\infty} t^n e^{-t^2} dt$ converge.
- 2) Montrer, à l'aide d'une intégration par parties que $\forall n \in \mathbf{N}, I_{n+2} = \frac{n+1}{2} I_n$.
- 3) Montrer que

$$I_n = \begin{cases} 0 & \text{si } n \text{ est impair} \\ \frac{(2p)!}{2^{2p} p!} \sqrt{\pi} & \text{si } n = 2p \text{ est pair} \end{cases}$$

- 4) En utilisant le changement de variable $x = \sqrt{t}$, déterminer la valeur de $\Gamma\left(p + \frac{1}{2}\right)$ pour $p \in \mathbf{N}$.

EXERCICE 20.19 (QSP ESCP 2010)

Étudier la nature de l'intégrale

$$\int_1^{+\infty} \left(\sqrt[3]{x^3 - 1} - \sqrt{x^2 - 1} \right) dx.$$

SUITES ET SÉRIES**EXERCICE 20.20** Étude d'une suite récurrente

Soit f la fonction définie sur $]0, 1[$ par $f(x) = x \left(1 + \frac{1}{\ln x} \right)$.

On considère alors la suite (u_n) définie par $\begin{cases} u_0 \in]0, e^{-1}[\\ \forall n \in \mathbf{N}, u_{n+1} = f(u_n) \end{cases}$

- 1) Montrer que $\forall n \in \mathbf{N}, 0 < u_n < e^{-1}$.
- 2) Montrer que la suite (u_n) est convergente.
- 3) En raisonnant par l'absurde, prouver que $\lim_{n \rightarrow +\infty} u_n = 0$.

EXERCICE 20.21 Déterminer si les séries suivantes convergent, et en cas de convergence, en calculer la somme.

$$1. \sum \frac{3^{2n+1}}{n!} \quad 2. \sum n \frac{4^n}{3^{2n+1}} \quad 3. \sum \frac{n^2}{3^n}$$

EXERCICE 20.22 Série harmonique et constante d'Euler

- 1) Montrer que pour tout $x > 0, \frac{1}{x+1} \leq \ln(x+1) - \ln(x) \leq \frac{1}{x}$.

2) On définit deux suites $(u_n)_{n \geq 1}$ et $(v_n)_{n \geq 1}$ par $u_n = \sum_{k=1}^n \frac{1}{k} - \ln(n)$ et $v_n = \sum_{k=1}^n \frac{1}{k} - \ln(n+1)$.

Montrer que (u_n) et (v_n) sont adjacentes.

3) En déduire que $\sum_{k=1}^n \frac{1}{k} \underset{n \rightarrow +\infty}{\sim} \ln(n)$.

EXERCICE 20.23 Fonction ζ (Extrait de EML 2010)

PD

1) Justifier la convergence des séries $\sum_{n \geq 1} \frac{1}{n^2}$, $\sum_{k \geq 0} \frac{1}{(2k+1)^2}$, $\sum_{n \geq 1} \frac{(-1)^n}{n^2}$.

2) En admettant que $\sum_{n=1}^{+\infty} \frac{1}{n^2} = \frac{\pi^2}{6}$, montrer que $\sum_{k=0}^{+\infty} \frac{1}{(2k+1)^2} = \frac{\pi^2}{8}$.

3) En déduire que $\sum_{n=1}^{+\infty} \frac{(-1)^n}{n^2} = -\frac{\pi^2}{12}$.

EXERCICE 20.24 Séries de Bertrand

AD

On s'intéresse dans cet exercice aux séries dont le terme général est donné par $u_n = \frac{1}{n^\alpha (\ln n)^\beta}$, avec $\alpha \in \mathbf{R}$, $\beta \in \mathbf{R}_+$.

1) Montrer que si $\alpha > 1$, $\sum u_n$ converge.

2) Montrer que si $\alpha < 1$, $\frac{1}{n} = o_{n \rightarrow +\infty}(u_n)$. Qu'en déduit-on sur la nature de la série $\sum u_n$.

3) On s'intéresse à présent au cas où $\alpha = 1$, et on considère la fonction f_β définie sur $]1, +\infty[$ par $f_\beta(x) = \frac{1}{x \ln(x)^\beta}$.

a) Montrer que pour tout $k \geq 2$ et pour tout $x \in [k, k+1]$, $u_{k+1} \leq \int_k^{k+1} f_\beta(x) dx \leq u_k$.

b) En déduire que pour tout $n \geq 2$,

$$\sum_{k=3}^{n+1} u_k \leq \int_2^{n+1} f_\beta(x) dx \leq \sum_{k=2}^n u_k.$$

c) Montrer que $\sum u_k$ converge si et seulement si $\beta > 1$. *Plus dur* : déterminer alors un encadrement de $\sum_{k=2}^{+\infty} u_k$.

EXERCICE 20.25 Développement en série entière de l'arctangente

AD

1) Montrer que pour tout $n \in \mathbf{N}^*$ et tout $t \in [0, 1]$, $\sum_{k=0}^n (-1)^k t^{2k} - \frac{1}{1+t^2} = (-1)^n \frac{t^{2n+2}}{1+t^2}$.

2) Montrer que $\forall n \in \mathbf{N}^*, \forall x \in [0, 1], \left| \sum_{k=0}^n (-1)^k \frac{x^{2k+1}}{2k+1} - \text{Arctan}(x) \right| \leq \frac{1}{2n+3}$.

3) En déduire, pour tout $x \in [0, 1]$, une écriture de $\text{Arctan}(x)$ comme somme d'une série.

EXERCICE 20.26 Étude d'une série alternée

AD

1) Étudier le sens de variation de la fonction $f : x \mapsto \frac{\ln x}{x}$ sur $[1; +\infty[$

2) Considérons la suite définie pour $n \geq 1$ par $u_n = \sum_{k=1}^n (-1)^k \frac{\ln k}{k}$.

a) Montrer que les suites $(u_{2n})_{n \geq 2}$ et $(u_{2n+1})_{n \geq 2}$ convergent vers une même limite.

b) En déduire que la série de terme général $\frac{(-1)^k \ln k}{k}$ est convergente.

3) (★) Soit $v_n = n^{\frac{(-1)^n}{n}} - 1$.

a) Montrer à l'aide d'un développement limité que $v_n = \frac{(-1)^n \ln n}{n} + \frac{\ln^2 n}{2n^2} + o\left(\frac{\ln^2 n}{n^2}\right)$.

b) En déduire que la série $\sum v_n$ converge.

EXERCICE 20.27 (QSP ESCP 2008)

Déterminer la nature de la série de terme général $u_n = \left(1 + \frac{1}{n}\right)^{2n} - \left(1 + \frac{2}{n}\right)^n$.



CORRECTION DES EXERCICES DU CHAPITRE 20

SOLUTION DE L'EXERCICE 20.1

1. La fonction sinus est continue sur $]0, \frac{\pi}{2}]$ et ne s'y annule pas, donc son inverse y est également continue.

De même, la fonction $x \mapsto \frac{1}{x}$ est continue sur $]0, \frac{\pi}{2}]$, et donc f est continue sur $]0, \frac{\pi}{2}]$ comme somme de fonctions continues.

Au voisinage de 0, on a

$$f(x) = \frac{1}{x} - \frac{1}{\sin x} = \frac{\sin x - x}{x \sin x}.$$

$$\text{Or, } \sin x - x = x - \frac{x^3}{6} + x + o(x^3) = -\frac{x^3}{6} + o(x^3) \underset{x \rightarrow 0}{\sim} -\frac{x^3}{6}.$$

Et sans difficultés, on a $x \sin x \underset{x \rightarrow 0}{\sim} x^2$. On en déduit que $f(x) \underset{x \rightarrow 0}{\sim} \frac{-\frac{x^3}{6}}{x^2} = -\frac{x}{6}$.

Et donc $\lim_{x \rightarrow 0^+} f(x) = 0 = f(0)$.

On en déduit que f est continue en 0, et donc sur $[0, \frac{\pi}{2}]$.

2. Comme à la question 1, on prouve que f est \mathcal{C}^1 sur $]0, \frac{\pi}{2}]$ par opérations sur les fonctions usuelles. On a alors

$$\forall x \in]0, \frac{\pi}{2}], f'(x) = -\frac{1}{x^2} + \frac{\cos x}{\sin^2 x} = \frac{x^2 \cos(x) - \sin^2(x)}{x^2 \sin^2(x)}.$$

3. On a, au voisinage de 0,

$$\begin{aligned} x^2 \cos(x) - \sin^2(x) &= x^2 \left(1 - \frac{x^2}{2} + o(x^2)\right) - \left(x - \frac{x^3}{6} + o(x^4)\right)^2 \\ &= x^2 - \frac{x^4}{2} + o(x^4) - \left(x^2 - \frac{x^3}{4} + o(x^4)\right) = -\frac{x^4}{6} + o(x^4) \underset{x \rightarrow 0}{\sim} -\frac{x^4}{6}. \end{aligned}$$

Puisque $x^2 \sin^2 x \underset{x \rightarrow 0}{\sim} x^4$, on en déduit que

$$f'(x) \underset{x \rightarrow 0}{\sim} -\frac{x^4}{6x^4} \underset{x \rightarrow 0^+}{\longrightarrow} -\frac{1}{6}.$$

4. Grâce au résultat de la question 2, il nous reste à prouver que f est dérivable en 0 et que $f'(0) = -\frac{1}{6}$.

Or on a, pour $x \in]0, \frac{\pi}{2}]$, $\frac{f(x) - f(0)}{x} = \frac{f(x)}{x}$.

Si l'on reprend les équivalents prouvés à la première question, il vient $\frac{f(x)}{x} \underset{x \rightarrow 0^+}{\sim} -\frac{1}{6}$, et donc

$$\lim_{x \rightarrow 0^+} \frac{f(x) - f(0)}{x} = -\frac{1}{6}.$$

On en déduit que f est dérivable en 0, que $f'(0) = -\frac{1}{6}$ et donc que f' est continue en 0.

Comme nous savions déjà que f' est continue sur $]0, \frac{\pi}{2}]$, f' est continue sur $[0, \frac{\pi}{2}]$ et donc f est \mathcal{C}^1 sur ce même intervalle.

SOLUTION DE L'EXERCICE 20.2

1. La fonction $x \mapsto \ln(1+x)$ est concave sur $] -1, +\infty[$, car elle est de classe \mathcal{C}^2 et que sa dérivée seconde est $x \mapsto -\frac{1}{(1+x)^2} \leq 0$.

En particulier, elle est en dessous de toutes ses tangentes.

Mais la tangente en 0 est $x \mapsto x$, et donc

$$\forall u > -1, \ln(1+u) \leq u.$$

Méthode

Rappelons que nous ne pouvons pas ajouter les équivalents, et que lorsqu'on cherche un équivalent d'une somme, on utilisera les développements limités.

Lorsqu'on est confrontés à des fractions, on commencera par les mettre sous le même dénominateur. Puis on ne cherche pas nécessairement un développement limité de la fraction complète, mais plutôt un développement limité du numérateur, qui fournit un équivalent du numérateur, ainsi qu'un développement limité du dénominateur, qui en fournit un équivalent.

Il suffira alors de diviser les équivalents afin d'obtenir un équivalent de la fraction.

Produit de D.L.

Rappelons que pour obtenir un DL_n d'un produit $f \times g$, on multiplie entre eux le DL_n et le DL_n de g , et on «oublie» tous les termes d'ordre supérieur ou égal à n .

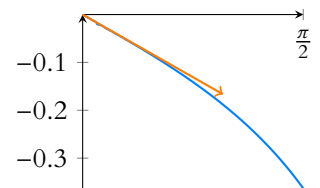


FIGURE 20.1— La courbe de f et la tangente en 0.

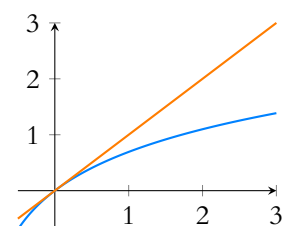


FIGURE 20.2— La fonction $x \mapsto \ln(1+x)$ et sa tangente.

2. La fonction exponentielle est convexe sur \mathbf{R} , et donc sa courbe se trouve au dessus de toutes ses tangentes.

Puisque la tangente en 0 est la droite d'équation $y = 1 + x$, on a donc

$$\forall x \in \mathbf{R}, 1 + x \leq e^x.$$

3. La fonction sin est de classe \mathcal{C}^2 sur $[0, \frac{\pi}{2}]$, et sa dérivée seconde est $x \mapsto -\sin(x)$, qui est négative sur $[0, \frac{\pi}{2}]$.

Elle est donc concave sur $[0, \frac{\pi}{2}]$.

Étant concave, elle est en dessous de toutes ses tangentes, et en particulier, de sa tangente en $x = 0$, qui est la droite d'équation $y = x$. Donc $\forall x \in [0, \frac{\pi}{2}]$, $\sin(x) \leq x$.

De plus, f est au-dessus de chacune de ses cordes.

Mais la corde joignant les points $(0, 0)$ et $(\frac{\pi}{2}, 1)$ est la droite d'équation $y = \frac{2}{\pi}x$, et donc

$$\forall x \in [0, \frac{\pi}{2}], \frac{2}{\pi}x \leq \sin x.$$

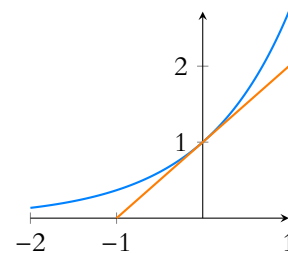
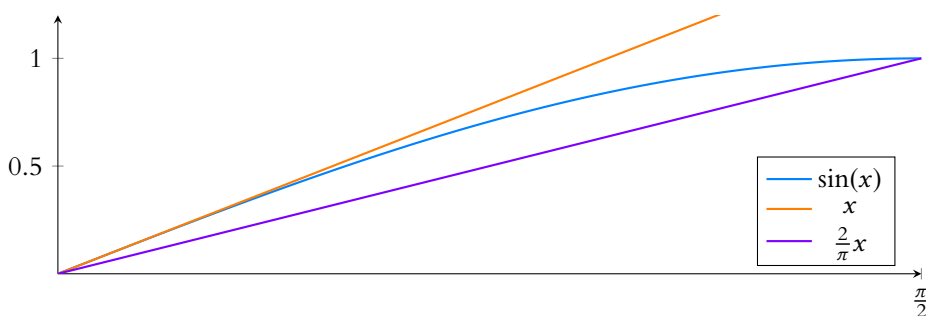


FIGURE 20.3– L'exponentielle

Corde

Rappelons qu'une corde de f est une droite joignant deux points de la courbe représentative de f .

SOLUTION DE L'EXERCICE 20.3

1. Si $f'(a) = 0$, alors l'équation de la tangente à la courbe représentative en $x = a$ est $y = f'(a)(x - a) + f(a) = f(a)$.
Et puisqu'une fonction convexe est située au dessus de ses tangentes, on a, pour tout $x \in \mathbf{R}$, $f(x) \geq f(a)$.
 f admet donc un minimum en a .
- 2.a. On a $g'(x) = f''(x) - \alpha \geq 0$, et donc g est croissante sur \mathbf{R} .
- 2.b. Puisque g est croissante, pour $x \geq 0$ on a $g(x) \geq g(0) = f'(0)$. Soit encore $f'(x) \geq f'(0) + \alpha x$.
De même, pour $x \leq 0$, on a $g(x) \leq g(0) = f'(0) \Leftrightarrow f'(x) \leq f'(0) + \alpha x$.
- 2.c. En particulier, il vient $\lim_{x \rightarrow +\infty} f'(x) = +\infty$ et $\lim_{x \rightarrow -\infty} f'(x) = -\infty$.
Puisque f' est continue, par le théorème des valeurs intermédiaires, il existe $a \in \mathbf{R}$ tel que $f'(a) = 0$.
Et alors le résultat de la question 1 s'applique : f possède un minimum global en a .

SOLUTION DE L'EXERCICE 20.4

1. La fonction f est dérivable sur $[0, \frac{\pi}{4}]$ car quotient de fonctions dérivables et on a alors $f'(x) = \frac{\sin(x)}{\cos^2(x)}$.
Mais sur $[0, \frac{\pi}{4}]$, $\sin x > 0$, de sorte que $f'(x) > 0$.
On en déduit que f est strictement croissante sur $[0, \frac{\pi}{4}]$.
D'autre part, elle y est continue¹ et donc par le théorème de la bijection, f réalise une bijection de $[0, \frac{\pi}{4}]$ sur $[f(0), f(\frac{\pi}{4})] = [1, \sqrt{2}]$.
2. Rappelons que la courbe représentative de f et la courbe représentative de f^{-1} sont symétriques par rapport à première bissectrice².
Notons que la question demande juste «l'allure» des courbes, et donc que le dessin attendu n'a pas besoin d'être extrêmement précis.

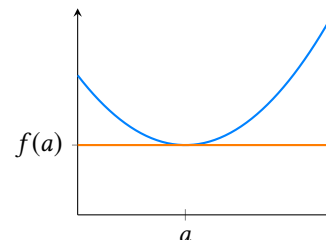
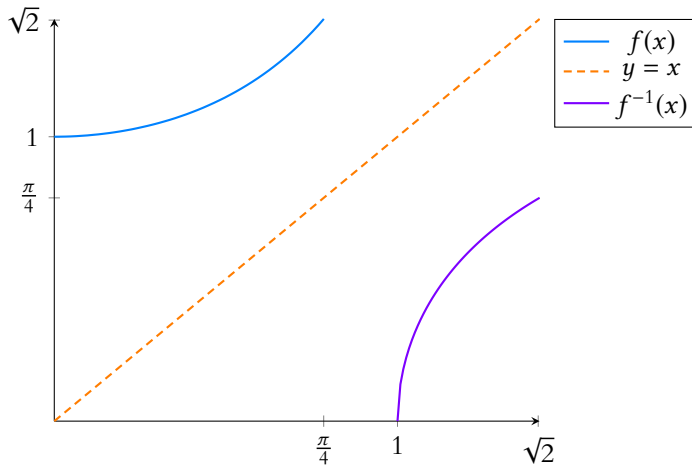


FIGURE 20.4– La tangente en a est horizontale et sous la courbe de f .

¹ Car dérivable.

² Qui, rappelons-le, est la droite d'équation $y = x$.



3. Par définition de la bijection réciproque, on a $f(f^{-1}(x)) = x$.

$$\text{Soit encore } \frac{1}{\cos(f^{-1}(x))} = x \Leftrightarrow \cos(f^{-1}(x)) = \frac{1}{x}.$$

D'autre part, on a

$$\sin^2(f^{-1}(x)) + \cos^2(f^{-1}(x)) = 1 \Leftrightarrow \sin^2(f^{-1}(x)) = 1 - \frac{1}{x^2}.$$

D'autre part, pour $x \in J$, on a $f^{-1}(x) \in \left[0, \frac{\pi}{4}\right]$, de sorte que $\sin(f^{-1}(x)) \geq 0$.

$$\text{Et donc } \sin(f^{-1}(x)) = \sqrt{1 - \frac{1}{x^2}}.$$

4. Sur J , f' s'annule uniquement en 0. Puisque $f^{-1}(x) = 0$ si et seulement si $x = 1$, f^{-1} est dérivable sur $J \setminus \{1\}$, avec

$$\begin{aligned} \forall x \in J \setminus \{1\}, (f^{-1})'(x) &= \frac{1}{f'(f^{-1}(x))} \\ &= \frac{\cos^2(f^{-1}(x))}{\sin(f^{-1}(x))} = \frac{1}{x^2} \frac{1}{\sqrt{1 - \frac{1}{x^2}}} \\ &= \frac{1}{x^2} \sqrt{\frac{x^2}{x^2 - 1}} = \frac{1}{x\sqrt{x^2 - 1}}. \end{aligned}$$

5. Par la formule de Taylor-Young à l'ordre 1, qui s'applique car f^{-1} est \mathcal{C}^1 sur $J \setminus \{1\}$, il vient

$$f^{-1}(x) = f^{-1}(\sqrt{2}) + (f^{-1})'(\sqrt{2})(x - \sqrt{2}) + o(x - \sqrt{2}).$$

Or nous savons que $f\left(\frac{\pi}{4}\right) = \frac{2}{\sqrt{2}} = \sqrt{2}$, de sorte que $f^{-1}(\sqrt{2}) = \frac{\pi}{4}$.

$$\text{Et } (f^{-1})'(\sqrt{2}) = \frac{1}{\sqrt{2}\sqrt{1}} = \frac{1}{\sqrt{2}}.$$

$$\text{Et donc } f^{-1}(x) = \frac{\pi}{4} + \frac{1}{\sqrt{2}}(x - \sqrt{2}) + o(x - \sqrt{2}).$$

SOLUTION DE L'EXERCICE 20.5

Commençons par remarquer que P' est de degré $n-1$ et donc possède au plus $n-1$ racines. Soient $a_1 < a_2 < \dots < a_n$ les racines de P .

Puisque P est dérivable sur \mathbf{R} et que $P(a_1) = P(a_2) = 0$, par le théorème de Rolle, il existe $b_1 \in]a_1, a_2[$ tel que $P'(b_1) = 0$.

Plus généralement, pour tout $i \in \llbracket 1, n-1 \rrbracket$, il existe $b_i \in]a_i, a_{i+1}[$ tel que $P'(b_i) = 0$.

Alors nécessairement les b_i sont deux à deux distincts, de sorte que P' possède au moins $n-1$ racines distinctes.

On en déduit que P' possède exactement $n-1$ racines distinctes.

SOLUTION DE L'EXERCICE 20.6

Rappel

En cas de doute sur la formule donnant la dérivée de f^{-1} , il suffit de se rappeler que $(f \circ f^{-1})(x) = x$. En dérivant cette relation, il vient alors

$$(f^{-1})'(x) \times f'(f^{-1}(x)) = 1$$

ce qui permet de retrouver $(f^{-1})'(x)$.

Ceci permet également de retrouver où f^{-1} est dérivable : c'est en tous les points x tels que $f'(f^{-1}(x)) \neq 0$, c'est-à-dire en ceux où le quotient précédent est défini.

\mathcal{C}^1

La dérivée obtenue à la question précédente est continue par opérations sur les fonctions usuelles, et donc f^{-1} est \mathcal{C}^1 .

Notons $u(x) = x^{n-1}$ et $v(x) = \ln(x)$.

Alors on a $u'(x) = (n-1)x^{n-2}$, $u''(x) = (n-1)(n-2)x^{n-3}$, etc, et une récurrence rapide prouverait que

$$\forall k \leq n-1, u^{(k)}(x) = (n-1)(n-2)\dots(n-k)x^{n-1-k} = \frac{(n-1)!}{(n-k-1)!}x^{n-1-k}.$$

Et pour $k \geq n$, on a $u^{(k)}(x) = 0$.

De même, on a

$$v'(x) = \frac{1}{x}, v''(x) = -\frac{1}{x^2}, v^{(3)}(x) = \frac{2}{x^3}, \dots$$

et par récurrence, on prouve de même que

$$\forall k \geq 1, v^{(k)}(x) = (-1)^{k-1} \frac{(k-1)!}{x^k}.$$

Par la formule de Leibniz³, on a alors

$$\begin{aligned} f^{(n)}(x) &= (u \times v)^{(n)} = \sum_{k=0}^n \binom{n}{k} v^{(k)}(x) u^{(n-k)}(x) \\ &= \sum_{k=1}^n \binom{n}{k} (-1)^{k-1} \frac{(k-1)!}{x^k} \frac{(n-1)!}{(k-1)!} x^{k-1} \\ &= \sum_{k=1}^n \binom{n}{k} (-1)^{k-1} (n-1)! \frac{1}{x} \\ &= -\frac{(n-1)!}{x} \sum_{k=1}^n \binom{n}{k} (-1)^k. \end{aligned}$$

³ Qui s'applique car u et v sont toutes deux de classe \mathcal{C}^n .

Le terme en $k=0$ s'annule car $u^{(n)}=0$.

D'autre part, nous savons, par la formule du binôme de Newton que

$$\sum_{k=0}^n \binom{n}{k} (-1)^k = (1-1)^n = 0 \Leftrightarrow \sum_{k=1}^n \binom{n}{k} (-1)^k = -1.$$

Et donc $f^{(n)}(x) = \frac{(n-1)!}{x}$.

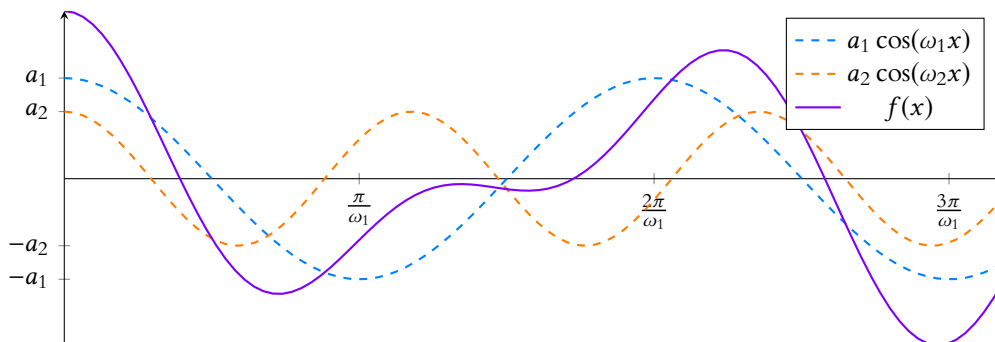
SOLUTION DE L'EXERCICE 20.7

1. Supposons (quitte à échanger a_1 et a_2), que $|a_1| > |a_2|$. Alors pour $x = \frac{2k\pi}{\omega_1}$, $k \in \mathbf{Z}$, on a $a_1 \cos(\omega_1 x) = a_1$ et $|a_2 \cos(\omega_2 x)| \leq |a_2| < |a_1|$, donc $f(x)$ est du signe de a_1 .

De même, pour $x = \frac{2k+1}{\omega_1}\pi$, $k \in \mathbf{Z}$, f est du signe de $-a_1$.

Par le théorème des valeurs intermédiaires (la fonction f est continue), sur chaque intervalle de la forme $]\frac{2k\pi}{\omega_1}, \frac{(2k+1)\pi}{\omega_1}[$, $k \in \mathbf{Z}$, il existe un réel α_k tel que $f(\alpha_k) = 0$.

Donc la fonction f s'annule une infinité de fois.



Chaque fois que $\cos(\omega_1 x) = 1$, $f(x) > 0$, et chaque fois que $\cos(\omega_1 x) = -1$, $f(x) < 0$. Entre les deux, f s'annule nécessairement au moins une fois.

2. Notons qu'une primitive de f est $F(x) = \frac{a_1}{\omega_1} \sin(\omega_1 x) + \frac{a_2}{\omega_2} \sin(\omega_2 x)$.

Supposons (quitte à les échanger) que $\omega_1 < \omega_2$. Alors $\left| \frac{a_1}{\omega_1} \right| > \left| \frac{a_2}{\omega_2} \right|$ (car nous avons supposé

que $|a_1| = |a_2|$).

Et alors il est possible de refaire le même raisonnement⁴ qu'à la question précédente : F s'annule une infinité de fois.

Mais alors, entre deux points consécutifs où F s'annule, le théorème de Rolle affirme que $F' = f$ s'annule au moins une fois.

Et donc F s'annulant une infinité de fois, il en est de même de f .

SOLUTION DE L'EXERCICE 20.8

La fonction sinus est de classe \mathcal{C}^∞ , et on a

$$\sin'(x) = \cos(x), \sin^{(2)}(x) = -\sin(x), \sin^{(3)}(x) = -\cos(x), \sin^{(4)}(x) = \sin(x).$$

Si on applique la formule de Taylor avec reste intégral à l'ordre 1 en 0, il vient

$$\sin(x) = x - \int_0^x (x-t) \sin t \, dt.$$

Or, pour $x \in \left[0, \frac{\pi}{2}\right]$, on a, quel que soit $t \in [0, x]$, $(x-t) \sin(t) \geq 0$ et donc, par croissance de l'intégrale, $\int_0^x (x-t) \sin(t) \, dt \geq 0$ et donc $\sin(x) \leq x$.

De même, en appliquant la formule de Taylor à l'ordre 3 en 0, il vient

$$\sin(x) = x - \frac{x^3}{6} + \int_0^x \frac{(x-t)^3}{6} \sin(t) \, dt.$$

Comme précédemment, on prouve que l'intégrale est positive, et donc

$$x - \frac{x^3}{6} \leq \sin(x).$$

SOLUTION DE L'EXERCICE 20.9

Notons $f(t) = e^t$, qui est de classe \mathcal{C}^∞ .

Soit $x \in \mathbf{R}$. Par l'inégalité de Taylor-Lagrange à l'ordre 1 appliquée entre 0 et x , si on note I le segment d'extrémités 0 et x et $M = \max_{t \in I} |f''(t)|$, il vient

$$|f(x) - f(0) - f'(0)x| \leq \frac{x^2}{2}M \Leftrightarrow |e^x - 1 - x| \leq \frac{x^2}{2}M.$$

Si $x \geq 0$, on a $M = e^x = e^{|x|}$, et donc l'inégalité souhaitée est démontrée.

Et si $x \leq 0$, alors le maximum de $f^{(2)}$ sur le segment $[x, 0]$ est $e^0 = 1$, et donc $M = 1 \leq e^{|x|}$, de sorte que

$$\forall x \in \mathbf{R}, |e^x - 1 - x| \leq \frac{x^2}{2}e^{|x|}.$$

SOLUTION DE L'EXERCICE 20.10

1. Prouvons le résultat par récurrence sur k .

$$\text{Pour } k = 1, \text{ on a } \Phi'(x) = \frac{1}{x} = \frac{(-1)^{1-1}(1-1)!}{x}.$$

Supposons donc que $\Phi^{(k)}(x) = \frac{(-1)^{k-1}(k-1)!}{x^k}$. Alors

$$\Phi^{(k+1)}(x) = (-1)^{k-1}(k-1)! \times \left(-\frac{k}{x^{k+1}}\right) = \frac{(-1)^k k!}{x^{k+1}}.$$

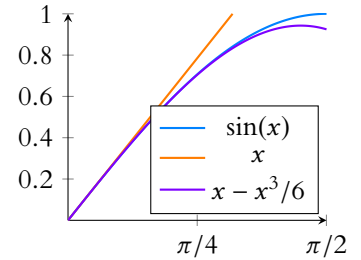
Donc la propriété est vraie au rang $k+1$, et donc par le principe de récurrence, pour tout

$$k \geq 1, \Phi^{(k)}(x) = \frac{(-1)^{k-1}(k-1)!}{x^k}.$$

2. Soit $n \in \mathbf{N}^*$ et $t \in [0, 1]$. Alors la fonction Φ est de classe \mathcal{C}^{n+1} sur $[1, 1+t]$, et pour tout $x \in [1, 1+t]$, on a

$$|\Phi^{(n+1)}(x)| = \frac{n!}{x^{n+1}} \leq n!$$

⁴ On prendra alors garde au fait que le sinus vaut 1 en $\frac{\pi}{2} + 2k\pi$ et -1 en $\frac{3\pi}{2} + 2k\pi$.



Taylor/DL

Notons que pour toutes les formules de Taylor, la partie polynomiale reste la même : c'est le développement limité. Ici on connaît le développement limité de sin en 0, inutile de le recalculer.

⚠ Danger !

Si $x \leq 0$, il ne s'agit pas réellement du segment $[0, x]$, mais plutôt du segment $[x, 0]$.

C'est l'une des subtilités de l'inégalité de Taylor-Lagrange, il faut prendre pour M un majorant de $f^{(n+1)}$ entre 0 et x .

Par l'inégalité de Taylor-Lagrange à l'ordre n appliquée entre 1 et $1 + t$, on a alors

$$\left| \Phi(1+t) - \sum_{k=0}^n \Phi^{(k)}(1) \frac{(t+1-1)^k}{k} \right| \leq \frac{n!|1+t-1|^{n+1}}{(n+1)!}$$

$$\Leftrightarrow \left| \ln(1+t) + \sum_{k=1}^n \frac{(-1)^k(k-1)!}{k!} t^k \right| \leq \frac{t^{n+1}}{n+1}$$

Notons que

$$\Phi^{(1)} = \ln(1) = 0$$

et donc le terme correspondant à $k = 0$ est nul.

3. La relation précédente appliquée en $t = 1$ donne

$$\left| \ln(2) + \sum_{k=1}^n \frac{(-1)^k}{k} \right| \leq \frac{1}{n+1}.$$

Et donc

$$-\frac{1}{n+1} \leq \ln(2) + \sum_{k=1}^n \frac{(-1)^k}{k} \leq \frac{1}{n+1} \Leftrightarrow -\ln(2) - \frac{1}{n+1} \leq \sum_{k=1}^n \frac{(-1)^k}{k} \leq -\ln(2) + \frac{1}{n+1}.$$

Par le théorème des gendarmes, on a alors

$$\lim_{n \rightarrow +\infty} \sum_{k=1}^n \frac{(-1)^k}{k} = -\ln(2).$$

Et donc, puisque la suite de ses sommes partielles converge vers $-\ln(2)$, la série de terme général $\frac{(-1)^n}{n}$ converge et sa somme vaut $-\ln(2)$.

SOLUTION DE L'EXERCICE 20.11

1. On a

$$\frac{\cos x}{1+x} = \cos(x) \times \frac{1}{1+x} = \left(1 - \frac{x^2}{2} + o(x^3)\right) \left(1 - x + \frac{x^2}{2} - \frac{x^3}{3} + o(x^3)\right) = 1 - x + \frac{1}{2}x^2 - \frac{1}{2}x^3 + o(x^3).$$

2. Nous connaissons les développements limités suivants :

$$e^x = 1 + x + \frac{x^2}{2} + \frac{x^3}{6} + o(x^3) \text{ et } \sqrt{1-x} = 1 - \frac{x}{2} - \frac{x^2}{8} - \frac{1}{16}x^3 + o(x^3)$$

donc par produit de développements limités, en «oubliant» les termes de degré supérieur à 3, il vient

$$e^x \sqrt{1-x} = \left(1 + x + \frac{x^2}{2} + \frac{x^3}{6} + o(x^3)\right) \left(1 - \frac{x}{2} - \frac{x^2}{8} - \frac{x^3}{16} + o(x^3)\right) = 1 + \frac{x}{2} - \frac{x^2}{8} - \frac{13}{48}x^3 + o(x^3)$$

3. On sait que $\sin(x) = x - \frac{x^3}{6} + \frac{x^5}{120} + o(x^5)$ et que $\cos(x) = 1 - \frac{x^2}{2} + \frac{x^4}{24} + o(x^5)$, de sorte que

$$\sin(x) \cos(x) = \left(x - \frac{x^3}{6} + \frac{x^5}{120} + o(x^5)\right) \left(1 - \frac{x^2}{2} + \frac{x^4}{24} + o(x^5)\right) = x - \frac{2}{3}x^3 + \frac{2}{15}x^5 + o(x^5).$$

SOLUTION DE L'EXERCICE 20.12

1. Réalisons un développement limité du numérateur à l'ordre 2 :

$$e^x - \sqrt{1+2x} = 1 + x + \frac{x^2}{2} - \left(1 + \frac{2x}{2} - \frac{(2x)^2}{8}\right) + o(x^2) = x^2 + o(x^2)$$

de sorte que

$$\frac{e^{2x} - \sqrt{1+x}}{x^2} = 1 + o(1).$$

On en déduit que la limite cherchée est 1.

$\sqrt{1-x}$

Pour obtenir le développement limité de $\sqrt{1-x}$, on commence par écrire celui de $\sqrt{1+x}$, puis on change les x en $-x$, ce qui a pour effet de changer le signe devant les puissances **impaires** de x (et non celui des puissances paires car $(-x)^2 = x^2$).

$o(1)$

Rappelons que $o(1)$ désigne une quantité qui tend vers 0.

2. Commençons par factoriser par le terme de plus haut degré, qui est ici $\sqrt[3]{x^3} = x$.

$$\sqrt[3]{x^3 + x^2} - \sqrt[3]{x^3 - x^2} = x \left(\sqrt[3]{1 + \frac{1}{x}} - \sqrt[3]{1 - \frac{1}{x}} \right).$$

Puisque $\frac{1}{x} \rightarrow 0$, nous pouvons utiliser le DL₁ de $(1 + u)^{\frac{1}{3}}$.

$$\sqrt[3]{1 + u} = (1 + u)^{\frac{1}{3}} = 1 + \frac{u}{3} + o(u)$$

de sorte que $\sqrt[3]{1 + \frac{1}{x}} = 1 + \frac{1}{3x} + o\left(\frac{1}{x}\right)$.

En changeant x en $-x$, il vient

$$\sqrt[3]{1 - \frac{1}{x}} = 1 - \frac{1}{3x} + o\left(\frac{1}{x}\right)$$

On en déduit que

$$\sqrt[3]{x^3 + x^2} - \sqrt[3]{x^3 - x^2} = x \left(\frac{2}{3}x + o\left(\frac{1}{x}\right) \right) = \frac{2}{3} + o(1).$$

Et donc

$$\lim_{x \rightarrow \infty} \sqrt[3]{x^3 + x^2} - \sqrt[3]{x^3 - x^2} = \frac{2}{3}.$$

3. Notons que $\sqrt[n]{e} - \frac{n+1}{n} = e^{1/n} - 1 - \frac{1}{n}$.

Puisque $\frac{1}{n} \rightarrow 0$, nous pouvons utiliser les développements limités usuels en 0 :

$$e^{1/n} - 1 - \frac{1}{n} = 1 + \frac{1}{n} + \frac{1}{2n^2} - 1 - \frac{1}{n} + o\left(\frac{1}{n^2}\right) = \frac{1}{2n^2} + o\left(\frac{1}{n^2}\right)$$

De même, $\sin \frac{1}{n} = \frac{1}{n} + o\left(\frac{1}{n^2}\right)$. Par produit de développements limités, on a

$$\sin \frac{1}{n} \left(\sqrt[n]{e} - 1 - \frac{1}{n} \right) = o\left(\frac{1}{n^2}\right)$$

On en déduit que

$$n^2 \sin \frac{1}{n} \left(\sqrt[n]{e} - \frac{n+1}{n} \right) = n^2 \times o\left(\frac{1}{n^2}\right) = o(1)$$

et donc

$$\lim_{n \rightarrow \infty} n^2 \sin \frac{1}{n} \left(\sqrt[n]{e} - \frac{n+1}{n} \right) = 0.$$

4. Calculons le développement limité en 0 du dénominateur, à l'ordre 4 :

$$\cos x - \sqrt{1 - x^2} = 1 - \frac{x^2}{2} + \frac{x^4}{24} - \left(1 - \frac{x^2}{2} - \frac{x^4}{8} \right) + o(x^4) = \frac{x^4}{6} + o(x^4)$$

On en déduit que

$$\frac{\cos(x) - \sqrt{1 - x^2}}{x^4} = \frac{1}{6} + o(1)$$

Par conséquent,

$$\lim_{x \rightarrow 0} \frac{\cos(x) - \sqrt{1 - x^2}}{x^4} = \frac{1}{6}.$$

SOLUTION DE L'EXERCICE 20.13

1. La fonction $t \mapsto e^{-t^2}$ est continue sur \mathbf{R}_+ , donc le seul éventuel problème de convergence est au voisinage de $+\infty$.

Or on a $t^2 e^{-t^2} \xrightarrow[t \rightarrow +\infty]{} 0$, de sorte que $e^{-t^2} = o\left(\frac{1}{t^2}\right)$.

Puisque $\int_1^{+\infty} \frac{1}{t^2} dt$ converge, il en est de même de $\int_1^{+\infty} e^{-t^2} dt$ et donc de $\int_0^{+\infty} e^{-t^2} dt$.

- 2.a. Notons que pour $h \in [-1, 1]$ et $t \in [0, 1]$, on a $0 \leq (1+t^2) \leq 1$ et donc $-2 \leq -h(1+t^2) \leq 2$. La fonction $x \mapsto e^x$ est continue sur le segment $[-2, 2]$ et donc y est bornée par une constante A .

Soit $h \in [-1, 1]$ fixé et $t \in [0, 1]$ fixé. Puisque la fonction exponentielle est \mathcal{C}^2 sur \mathbf{R} , et que ses dérivées successives sont toutes égales à elle-même, l'inégalité de Taylor-Lagrange appliquée entre 0 et $-h(1+t^2)$, qui appartient à $[-2, 2]$, nous donne alors

$$|e^{-h(1+t^2)} - 1 + h(1+t^2)| \leq \frac{A}{2} h^2 (1+t^2)^2 \leq 2Ah^2.$$

Si on pose alors $M = 2A$, on obtient le résultat souhaité.

- 2.b. Soit $x \in \mathbf{R}$ fixé, et soit $h \in [-1, 1]$. Pour tout $t \in [0, 1]$, en multipliant l'inégalité de la question précédente par $\frac{e^{-x(1+t^2)}}{|h|(1+t^2)}$, il vient

$$\left| \frac{1}{h} \left(\frac{e^{-(x+h)(1+t^2)}}{1+t^2} - \frac{e^{-x(1+t^2)}}{1+t^2} \right) + e^{-x(1+t^2)} \right| \leq \frac{Mh^2 e^{-x(1+t^2)}}{|h|(1+t^2)}.$$

Ceci étant valable quel que soit $t \in [0, 1]$, on en déduit par croissance de l'intégrale⁵ que

$$\int_0^1 \left| \frac{1}{h} \left(\frac{e^{-(x+h)(1+t^2)}}{1+t^2} - \frac{e^{-x(1+t^2)}}{1+t^2} \right) + e^{-x(1+t^2)} \right| dt \leq M|h| \underbrace{\int_0^1 \frac{e^{-x(1+t^2)}}{1+t^2} dt}_{=f(x)}.$$

Mais il vient alors

$$\left| \int_0^1 \left(\frac{1}{h} \left(\frac{e^{-(x+h)(1+t^2)}}{1+t^2} - \frac{e^{-x(1+t^2)}}{1+t^2} \right) + e^{-x(1+t^2)} \right) dt \right| \leq M|h|f(x).$$

Soit encore

$$\left| \frac{f(x+h) - f(x)}{h} + \int_0^1 e^{-x(1+t^2)} dt \right| \leq M|h|f(x).$$

Et donc, lorsque h tend vers 0, il vient

$$\lim_{h \rightarrow 0} \left| \frac{f(x+h) - f(x)}{h} + \int_0^1 e^{-x(1+t^2)} dt \right| = 0 \Rightarrow \lim_{h \rightarrow 0} \frac{f(x+h) - f(x)}{h} = - \int_0^1 e^{-x(1+t^2)} dt.$$

On en déduit donc que f est dérivable en x et que $f'(x) = - \int_0^1 e^{-x(1+t^2)} dt$.

3. Nous venons de prouver que f est dérivable, et $x \mapsto \int_0^x e^{-t^2} dt$ est dérivable⁶. Donc g est dérivable comme somme de fonctions dérivables, et sa dérivée vaut

$$g'(x) = 2x f'(x^2) + 2e^{-x^2} \int_0^x e^{-t^2} dt = -2x \int_0^1 e^{-x^2(1+t^2)} dt + 2e^{-x^2} \int_0^x e^{-t^2} dt.$$

Dans la seconde intégrale, procédons au changement de variable⁷ $t = xu \Leftrightarrow u = \frac{t}{x}$. Ainsi, lorsque $t = 0$, $u = 0$, et lorsque $t = x$, $u = 1$. On a de plus $dt = xdu$. Alors

$$\int_0^x e^{-t^2} dt = \int_0^1 e^{-x^2 u^2} x du$$

et donc

$$g'(x) = -2x \int_0^1 e^{-x^2(1+t^2)} dt + 2x \int_0^1 e^{-x^2(1+u^2)} du = 0.$$

On en déduit que g est une fonction constante.

Majorant

Ici, on ne demande pas de calculer A , seulement de prouver qu'il existe. Il ne coûterait pas beaucoup plus cher de prouver que $A = e^2$, mais face à une fonction plus compliquée, il est bon de se souvenir qu'une fonction continue sur un segment y est bornée, et donc qu'un majorant existe même si on ne sait pas toujours le calculer.

⁵ Les intégrales étant les intégrales de fonctions continues sur un segment, il n'y a donc pas à se préoccuper de leur convergence.

Valeur absolue

Rappelons que l'inégalité triangulaire pour les intégrales nous donne

$$\left| \int f \right| \leq \int |f|.$$

⁶ Car il s'agit de la primitive de e^{-x^2} qui s'annule en 0.

⁷ Sur un segment.

4. Pour tout $t \in [0, 1]$, on a $0 \leq \frac{e^{-x(1+t^2)}}{1+t^2} \leq \frac{e^{-x}}{1+t^2}$, et donc, par croissance de l'intégrale,

$$0 \leq \int_0^1 \frac{e^{-x(1+t^2)}}{1+t^2} dt \leq \int_0^1 \frac{e^{-x}}{1+t^2} dt = e^{-x} \int_0^1 \frac{dt}{1+t^2} = e^{-x} [\operatorname{Arctant} t]_0^1 = \frac{\pi}{4} e^{-x}.$$

On a donc bien $0 \leq f(x) \leq \frac{\pi}{4} e^{-x}$.

On en déduit, par le théorème des gendarmes, que $\lim_{x \rightarrow +\infty} f(x) = 0$.

5. g est constante, et $g(0) = f(0) = \int_0^1 \frac{dt}{1+t^2} = \frac{\pi}{4}$.

Ainsi, g est la fonction constante égale à $\frac{\pi}{4}$.

D'autre part, on a ;

$$\lim_{x \rightarrow +\infty} g(x) = \lim_{x \rightarrow +\infty} \underbrace{f(x^2)}_{=0} + \left(\int_0^{+\infty} e^{-t^2} dt \right)^2 = A^2.$$

Cette limite étant nécessairement⁸ égale à $\frac{\pi}{4}$, on en déduit que $A^2 = \frac{\pi}{4}$ et donc

$$A = \int_0^{+\infty} e^{-t^2} dt = \frac{\sqrt{\pi}}{2}.$$

⁸ Limite d'une fonction constante...

SOLUTION DE L'EXERCICE 20.14

1. La fonction $x \mapsto \frac{1}{x + \ln(x)}$ est continue sur $[1, +\infty[$, donc le seul éventuel problème de convergence est au voisinage de $+\infty$

Au voisinage de $+\infty$, on a $\ln(x) = o(x)$ et donc $x + \ln(x) \underset{x \rightarrow +\infty}{\sim} x$.

Et donc $\frac{1}{x + \ln(x)} \underset{x \rightarrow +\infty}{\sim} \frac{1}{x}$.

Puisque $\int_1^{+\infty} \frac{dx}{x}$ est une intégrale de Riemann divergente, par le critère des équivalents pour les intégrales de fonctions positives, $\int_1^{+\infty} \frac{1}{x + \ln(x)}$ diverge.

2. La fonction $t \mapsto \sin(t) \sin\left(\frac{1}{t^2}\right)$ est continue sur $]0, +\infty[$.

Au voisinage de 0, on a $\sin t \rightarrow 0$, et $t \mapsto \sin\left(\frac{1}{t^2}\right)$ est bornée, de sorte $\sin t \sin\left(\frac{1}{t^2}\right) \xrightarrow{t \rightarrow 0} 0$.

Ainsi, $\int_0^1 \sin t \sin\left(\frac{1}{t^2}\right) dt$ est faussement impropre, donc convergente.

Au voisinage de $+\infty$, $\frac{1}{t^2} \rightarrow 0$, de sorte que $\sin\left(\frac{1}{t^2}\right) \underset{t \rightarrow +\infty}{\sim} \frac{1}{t^2}$.

On en déduit que $\int_1^{+\infty} \sin\left(\frac{1}{t^2}\right) dt$ converge.

D'autre part, pour tout $t \geq 1$, on a $\left| \sin t \sin\left(\frac{1}{t^2}\right) \right| \leq \sin\left(\frac{1}{t^2}\right)$.

Et donc $\int_1^{+\infty} \sin(t) \sin\left(\frac{1}{t^2}\right) dt$ converge absolument et donc converge.

On en déduit que $\int_0^{+\infty} \sin t \sin\left(\frac{1}{t^2}\right) dt$ converge.

3. La fonction $t \mapsto \frac{1}{(1-t)\sqrt{t}}$ est continue sur $]0, 1[$.

Au voisinage de 1, $\frac{1}{(1-t)\sqrt{t}} \underset{t \rightarrow 1^-}{\sim} \frac{1}{1-t}$.

Mais $\int_{1/2}^1 \frac{dt}{1-t}$ est une intégrale de Riemann divergente, donc par critère de comparaison

pour les intégrales de fonctions positives, $\int_{1/2}^1 \frac{dt}{(1-t)\sqrt{t}}$ diverge et donc il en est de même

de $\int_0^1 \frac{dt}{(1-t)\sqrt{t}}$.

Remarque

Une valeur absolue étant toujours positive, il est bien légitime d'utiliser une majoration pour prouver une convergence absolue.

Remarque

On pourrait prouver qu'entre 0 et 1/2, l'intégrale converge. Toutefois cela ne sert à rien, puisque l'intégrale diverge en 1, elle diverge quoi qu'il se passe au voisinage de 0.

4. La fonction $t \mapsto \ln te^{-t}$ est continue sur $]0, +\infty[$.

Au voisinage de 0, on a $\ln te^{-t} \underset{t \rightarrow 0^+}{\sim} \ln t$.

D'autre part, pour $A \in]0, 1]$, on a

$$\int_A^1 \ln t \, dt = [t \ln t - t]_A^1 = -1 - A \ln(A) + A \xrightarrow{t \rightarrow 0^+} -1$$

de sorte que $\int_0^1 \ln t \, dt$ converge.

Par critère de comparaison pour les fonctions de signe constant on en déduit que

$$\int_0^1 \ln te^{-t} \, dt \text{ converge.}$$

Au voisinage de $+\infty$, on a

$$t^2 \ln te^{-t} = \frac{\ln t}{t} t^2 e^{-t} \xrightarrow{t \rightarrow +\infty} 0.$$

Et donc $\ln te^{-t} = \underset{t \rightarrow +\infty}{o} \left(\frac{1}{t^2} \right)$.

Par conséquent, $\int_1^{+\infty} \ln te^{-t} \, dt$ converge par comparaison à une intégrale de Riemann convergente.

5. La fonction $t \mapsto \frac{\ln(1+t)}{t^{3/2}}$ est continue et positive sur $]0, +\infty[$, donc il y a d'éventuels problèmes de convergence au voisinage de 0 et au voisinage de $+\infty$.

Au voisinage de 0, on a $\ln(1+t) \underset{t \rightarrow 0}{\sim} t$ et donc $\frac{\ln(1+t)}{t^{3/2}} \underset{t \rightarrow 0}{\sim} \frac{t}{t^{3/2}} = \frac{1}{\sqrt{t}}$.

Puisque $\int_0^1 \frac{1}{\sqrt{t}} \, dt$ est une intégrale de Riemann convergente, $\int_0^1 \frac{\ln(1+t)}{t^{3/2}} \, dt$ converge par le critère des équivalents.

Au voisinage de $+\infty$, on a $0 \leq \ln(1+t) \leq \ln(t^2) = 2 \ln(t)$.

Et donc $t^{5/4} \frac{\ln(1+t)}{t^{3/2}} \leq 2 \frac{\ln t}{t^{1/4}} \xrightarrow{t \rightarrow +\infty} 0$ par croissance comparée.

Et donc $\frac{\ln(1+t)}{t^{3/2}} = \underset{t \rightarrow +\infty}{o} \left(\frac{1}{t^{5/4}} \right)$.

Mais $\int_1^{+\infty} \frac{dt}{t^{5/4}}$ est une intégrale de Riemann convergente, et donc par critère de comparaison, $\int_1^{+\infty} \frac{\ln(1+t)}{t^{3/2}} \, dt$ converge.

On en déduit donc que $\int_0^{+\infty} \frac{\ln(1+t)}{t^{3/2}} \, dt$ converge.

6. La fonction $t \mapsto \ln(t^2 - t)$ est continue sur $]1, +\infty[$.

Au voisinage de 1, on a $\ln(t^2 - t) = \ln(t) + \ln(t-1) \underset{t \rightarrow 1^+}{\sim} \ln(t-1)$, et donc⁹

$$\frac{\ln(t^2 - t)}{(1+t)^2} \underset{t \rightarrow 1^+}{\sim} \frac{\ln(t-1)}{4}.$$

Mais pour $A \in]1, 2[$, on a

$$\begin{aligned} \int_A^2 \ln(t-1) \, dt &= [(t-1) \ln(t-1) - (t-1)]_A^2 \\ &= -1 + (A-1) \ln(A-1) + (A-1) \\ &\xrightarrow{A \rightarrow 1^+} -1. \end{aligned}$$

Ainsi, $\int_1^2 \ln(t-1) \, dt$ converge, et par critère de comparaison pour les intégrales de

fonctions de signe constant¹⁰, $\int_1^2 \frac{\ln(t^2 - t)}{(1+t)^2} \, dt$ converge.

Au voisinage de $+\infty$, on a $\frac{\ln(t^2 - t)}{(1+t)^2} \underset{t \rightarrow +\infty}{\sim} \frac{\ln(t^2 - t)}{t^2} = \frac{\ln(t) + \ln(t-1)}{t^2}$.

Signe ⚠

Attention à ne pas dire ici «pour les fonctions positives» car sur $]0, 1]$, $\ln(t)e^{-t} \leq 0$.

Méthode

La positivité ne pose ici aucune difficulté. La mentionner une bonne fois pour toutes nous autorisera à utiliser le critère des équivalents sans précautions dans la suite.

⁹ Le dénominateur tend vers 4.

¹⁰ Ici, les deux fonctions sont négatives sur $]1, 2]$.

D'une part, on a $t^{3/2} \frac{\ln(t)}{t^2} = \frac{\ln(t)}{\sqrt{t}} \xrightarrow{t \rightarrow +\infty} 0$ et donc $\frac{\ln(t)}{t^2} = o\left(\frac{1}{t^{3/2}}\right)$.

Or, $\int_2^{+\infty} \frac{dt}{t^{3/2}}$ converge, et donc il en est de même de $\int_2^{+\infty} \frac{\ln t}{t^2} dt$.

D'autre part, le changement de variable $u = t-1$ montre que, sous réserve de convergence¹¹,

$$\int_2^{+\infty} \frac{\ln(t-1)}{t^2} dt = \int_1^{+\infty} \frac{\ln(u)}{(u+1)^2} du.$$

Mais puisque $\frac{\ln(u)}{(u+1)^2} \underset{u \rightarrow +\infty}{\sim} \frac{\ln(u)}{u^2}$, et que $\int_1^2 \frac{\ln(u)}{u^2} du$ converge, il en est de même de

$$\int_2^{+\infty} \frac{\ln(t-1)}{t^2} dt.$$

Et donc $\int_2^{+\infty} \frac{\ln(t^2-t)}{t^2} dt$ converge, car somme d'intégrales convergentes.

Par le critère des équivalents, il en est donc de même de $\int_2^{+\infty} \frac{\ln(t^2-t)}{(1+t)^2} dt$.

Et donc $\int_1^{+\infty} \frac{\ln(t^2-t)}{(1+t)^2} dt$ converge.

7. La fonction $u \mapsto \frac{u^{\alpha-1}}{e^u-1}$ est continue sur $]0, +\infty[$.

Au voisinage de 0, on a $e^u - 1 \underset{u \rightarrow 0^+}{\sim} u$ et donc $\frac{u^{\alpha-1}}{e^u-1} \underset{u \rightarrow 0^+}{\sim} u^{\alpha-2} = \frac{1}{u^{2-\alpha}}$.

Puisqu'on a $\alpha > 1 \Leftrightarrow 2 - \alpha < 1$, $\int_0^1 \frac{du}{u^{2-\alpha}}$ converge et donc il en est de même de

$$\int_0^1 \frac{u^{\alpha-1}}{e^u-1} du.$$

Au voisinage de $+\infty$, on a $e^u - 1 \underset{u \rightarrow +\infty}{\sim} e^u$ et donc $\frac{u^{\alpha-1}}{e^u-1} \underset{u \rightarrow +\infty}{\sim} \frac{u^{\alpha-1}}{e^u} = u^{\alpha-1} e^{-u}$.

Mais nous savons que $\Gamma(\alpha) = \int_0^{+\infty} u^{\alpha-1} e^{-u} du$ converge, et donc il en est de même de

$$\int_1^{+\infty} u^{\alpha-1} e^{-u} du.$$

Par le critère des équivalents, il en est donc de même de $\int_1^{+\infty} \frac{u^{\alpha-1}}{e^u-1} du$.

Et donc $\int_0^{+\infty} \frac{u^{\alpha-1}}{e^u-1} du$ converge.

SOLUTION DE L'EXERCICE 20.15

1. La fonction $t \mapsto \frac{\ln t}{t^2+1}$ est continue sur $]0, +\infty[$, donc les éventuels problèmes de convergence sont en 0 et en $+\infty$.

Au voisinage de 0, on a $1+t^2 \sim 1$, de sorte que $\frac{\ln t}{1+t^2} \sim \ln t$.

Or, pour $A > 0$, on a

$$\int_A^1 \ln t dt = [t \ln t - t]_A^1 = A - A \ln A - 1 \xrightarrow{A \rightarrow 0^+} -1.$$

On en déduit que $\int_0^1 \ln t dt$ converge, et donc, par critère de comparaison pour les

fonctions de signe constant¹², il en est de même de $\int_0^1 \frac{\ln t}{1+t^2} dt$.

Au voisinage de $+\infty$, on a $\frac{\ln t}{1+t^2} \sim \frac{\ln t}{t^2}$.

Mais $t^{3/2} \frac{\ln t}{t^2} = \frac{\ln t}{\sqrt{t}} \xrightarrow{t \rightarrow +\infty} 0$ par croissance comparée.

On en déduit que $\frac{\ln t}{1+t^2} = o\left(\frac{1}{t^{3/2}}\right)$ et donc $\int_1^{+\infty} \frac{\ln t}{1+t^2} dt$ converge par comparaison à une intégrale de Riemann convergente.

2. La fonction $t \mapsto \frac{1}{t}$ est de classe \mathcal{C}^1 et strictement décroissante sur \mathbf{R}_+^* .

¹¹ Rappelons que lors d'un changement de variable, les deux intégrales ont même nature.

Remarque

Une erreur classique serait de dire que $t^2 - t \underset{t \rightarrow +\infty}{\sim} t^2$ et donc

$$\ln(t^2 - t) \underset{t \rightarrow +\infty}{\sim} \ln(t^2) = 2 \ln(t).$$

Nous ne pouvons rien affirmer de tel car il est interdit de composer (à gauche) les équivalents.

Toutefois, cet équivalent est bien vrai, mais si on souhaite l'utiliser ici, il faut le prouver «à la main».

¹² Ici négatif.

Chgt de variable

Pour un changement de variable dans une intégrale impropre, on n'oubliera pas de vérifier la stricte monotonie du changement de variable.

Si on pose $u = \frac{1}{t} \Leftrightarrow t = \frac{1}{u}$, alors $dt = -\frac{du}{u^2}$, et donc

$$\int_0^{+\infty} \frac{\ln t}{1+t^2} dt = \int_{+\infty}^0 \frac{\ln \frac{1}{u}}{1+\frac{1}{u^2}} \left(-\frac{du}{u^2}\right) = \int_0^{+\infty} \frac{\ln \frac{1}{u}}{1+u^2} du = - \int_0^{+\infty} \frac{\ln u}{1+u^2} du.$$

Ainsi, $I = -I$, et donc $I = 0$.

3. On a $\frac{\ln t}{t^2+a^2} = \frac{1}{a^2} \frac{\ln t}{1+(\frac{t}{a})^2}$.

Si on pose $x = \frac{t}{a}$, alors il vient

$$\int_0^{+\infty} \frac{\ln t}{1+(\frac{t}{a})^2} dt = a \int_0^{+\infty} \frac{\ln(ax)}{1+x^2} dx = a \int_0^{+\infty} \frac{\ln a}{1+x^2} dx + a \underbrace{\int_0^{+\infty} \frac{\ln x}{1+x^2} dx}_{=0}.$$

On a bien le droit de séparer l'intégrale en deux, car les deux morceaux convergent.

Or, on a

$$\int_0^{+\infty} \frac{\ln a}{1+x^2} dx = \lim_{A \rightarrow +\infty} \int_0^A \frac{\ln a}{1+x^2} dx = \lim_{A \rightarrow +\infty} [\ln(a) \operatorname{Arctan}(x)]_0^A = \ln a \frac{\pi}{2}.$$

On en déduit donc que

$$\int_0^{+\infty} \frac{\ln t}{t^2+a^2} dt = \frac{1}{a^2} \int_0^{+\infty} \frac{\ln t}{1+(\frac{t}{a})^2} dt = \frac{\pi \ln(a)}{2a}.$$

SOLUTION DE L'EXERCICE 20.16

Soit $n \in \mathbf{N}$ fixé. Procédons, comme indiqué, à une intégration par parties, ce qui est légitime car les fonctions u et $t \mapsto \cos(nt)$ sont de classe \mathcal{C}^1 . Alors

$$\int_0^\pi u(t) \sin(nt) dt = \left[-u(t) \frac{1}{n} \cos(nt) \right]_0^\pi + \frac{1}{n} \int_0^\pi u'(t) \cos(nt) dt = \frac{u(0)}{n} - \frac{u(\pi)}{n} \cos(n\pi) + \frac{1}{n} \int_0^\pi u'(t) \cos(nt) dt$$

Il est évident que lorsque $n \rightarrow +\infty$, alors $\frac{u(0)}{n} \rightarrow 0$. De même, on a

$$\left| \frac{1}{n} u(\pi) \cos(n\pi) \right| \leq \left| \frac{u(\pi)}{n} \right| \xrightarrow{n \rightarrow +\infty} 0.$$

Enfin, on a

$$\begin{aligned} \left| \frac{1}{n} \int_0^\pi u'(t) \cos(nt) dt \right| &\leq \frac{1}{|n|} \int_0^\pi |u'(t) \cos(nt)| dt \\ &\leq \frac{1}{|n|} \int_0^\pi |u'(t)| dt \\ &\xrightarrow{n \rightarrow +\infty} 0 \end{aligned}$$

Remarque

Cette intégrale ne dépend pas de n , ni d'aucune autre variable : c'est une constante !

Et donc on en déduit que $\int_0^\pi u(t) \cos(nt) dt \xrightarrow{n \rightarrow +\infty} 0$.

SOLUTION DE L'EXERCICE 20.17

1. La fonction $t \mapsto \frac{\cos t}{t^2+n}$ est continue sur $[0, +\infty[$.

Or, on a $0 \leq \left| \frac{\cos t}{t^2+n} \right| \leq \frac{1}{t^2+n} \leq \frac{1}{t^2}$.

Mais $\int_1^{+\infty} \frac{dt}{t^2}$ converge, et donc par $\int_1^{+\infty} \frac{\cos t}{t^2+n} dt$ est absolument convergente et par conséquent convergente.

Puisqu'il n'y a pas de problème de convergence au voisinage de 0, on en déduit que

$$\int_0^{+\infty} \frac{\cos t}{t^2+n} dt \text{ converge.}$$

D'autre part, par l'inégalité triangulaire, on a

$$|u_n| \leq \int_0^{+\infty} \frac{|\cos t|}{t^2+n} dt \leq \int_0^{+\infty} \frac{dt}{t^2+n}.$$

$$2. \text{ On a } \int_0^{+\infty} \frac{dt}{t^2+n} = \frac{1}{n} \int_0^{+\infty} \frac{dt}{\left(\frac{t}{\sqrt{n}}\right)^2 + 1}.$$

Procédons au changement de variable¹³ $x = \frac{t}{\sqrt{n}}$, de sorte que

$$\int_0^{+\infty} \frac{dt}{\left(\frac{t}{\sqrt{n}}\right)^2 + 1} = \int_0^{+\infty} \frac{\sqrt{n}}{x^2 + 1} dx.$$

Mais pour $A > 0$, on a

$$\int_0^A \frac{\sqrt{n}}{x^2 + 1} dx = [\sqrt{n} \operatorname{Arctan}(x)]_0^A = \sqrt{n} \operatorname{Arctan}(A) \xrightarrow{A \rightarrow +\infty} \sqrt{n} \frac{\pi}{2}.$$

Par conséquent, $|u_n| \leq \frac{\pi}{2\sqrt{n}}$ et donc par le théorème des gendarmes, $u_n \xrightarrow{n \rightarrow +\infty} 0$.

SOLUTION DE L'EXERCICE 20.18

1. Lorsque n est pair, l'intégrande est paire, et lorsque n est impair, l'intégrande est impaire.

Dans les deux cas, $\int_{-\infty}^{+\infty} t^n e^{-t^2} dt$ converge si et seulement si $\int_0^{+\infty} t^n e^{-t^2} dt$ converge.

Or, $t \mapsto t^n e^{-t^2}$ est continue sur $[0, +\infty[$, donc le seul problème de convergence est en $+\infty$. Au voisinage de $+\infty$, on a, par croissance comparée,

$$t^2 t^n e^{-t^2} = t^{n+2} e^{-t^2} \xrightarrow{t \rightarrow +\infty} 0.$$

Ainsi, $t^n e^{-t^2} = o\left(\frac{1}{t^2}\right)$. Et puisque $\int_1^{+\infty} \frac{1}{t^2} dt$ converge, il en est de même de $\int_1^{+\infty} t^n e^{-t^2} dt$ et donc de $\int_0^{+\infty} t^n e^{-t^2} dt$.

2. Soient $A < B$ deux réels. Posons alors $u(t) = -\frac{1}{2}e^{-t^2}$ et $v(t) = t^{n+1}$, qui sont deux fonctions de classe \mathcal{C}^1 sur le segment $[A, B]$, avec $u'(t) = e^{-t^2}$ et $v'(t) = (n+1)t^n$. Une intégration par parties donne alors

$$\int_A^B t^{n+2} e^{-t^2} dt = \left[-\frac{e^{-t^2}}{2} t^{n+1} \right]_A^B + \int_A^B \frac{n+1}{2} t^n e^{-t^2} dt = \frac{B^{n+1} e^{-B^2}}{2} - \frac{A^{n+1} e^{-A^2}}{2} + \frac{n+1}{2} \int_A^B t^n e^{-t^2} dt.$$

Lorsque $A \rightarrow -\infty$, il vient

$$\int_{-\infty}^B t^{n+2} e^{-t^2} dt = \frac{B^{n+1} e^{-B^2}}{2} + \frac{n+1}{2} \int_{-\infty}^B t^n e^{-t^2} dt$$

puis lorsque $B \rightarrow +\infty$,

$$\int_{-\infty}^{+\infty} t^{n+2} e^{-t^2} dt = \frac{n+1}{2} \int_{-\infty}^{+\infty} t^n e^{-t^2} dt \Leftrightarrow I_{n+2} = \frac{n+1}{2} I_n.$$

3. Si n est impair, l'intégrande étant une fonction impaire¹⁴, on en déduit que $I_n = 0$.

Si n est pair, prouvons par récurrence sur p que $I_{2p} = \frac{((2p)!)}{2^{2p} p!} \sqrt{\pi}$.

Pour $p = 0$, nous savons que la densité d'une variable aléatoire suivant la loi normale

$\mathcal{N}\left(0, \frac{1}{\sqrt{2}}\right)$ est $t \mapsto \frac{1}{\sqrt{\pi}} e^{-t^2}$, et son intégrale vaut 1. Donc

$$\int_{-\infty}^{+\infty} \frac{1}{\sqrt{\pi}} e^{-t^2} dt = 1 \Leftrightarrow I_0 = \int_{-\infty}^{+\infty} e^{-t^2} dt = \sqrt{\pi}.$$

Donc la récurrence est initialisée.

Supposons que $I_{2p} = \frac{(2p)!}{2^{2p} p!} \sqrt{\pi}$. Alors, par la question précédente,

¹³ Affine.

Détails

Si on souhaite détailler l'argument de croissance comparée, on peut poser $X = t^2$, qui tend vers $+\infty$. Et alors

$$t^{n+2} e^{-t^2} = X^{\frac{n+1}{2}} e^{-X}.$$

On peut alors utiliser une croissance comparée usuelle.

⚠ Danger !

On ne sait pas intégrer e^{-t^2} . Si vous arrivez à en exprimer une primitive à l'aide d'une fonction usuelle, celle-ci est fautive !

¹⁴ Et d'intégrale convergente !

$$I_{2(p+1)} = I_{2p+2} = \frac{2p+1}{2} \frac{(2p)!}{2^{2p}p!} \sqrt{\pi} = \frac{(2p+2)(2p+1)(2p)!}{2(2p+2)2^{2p}p!} \sqrt{\pi}$$

$$= \frac{(2p+2)!}{2^2(p+1)2^{2p}p!} \sqrt{\pi} = \frac{(2p+2)!}{2^{2p+2}(p+1)!} \sqrt{\pi}.$$

Donc la propriété est vraie au rang $p+1$, et alors, par le principe de récurrence,

$$\forall p \in \mathbf{N}, I_{2p} = \frac{(2p)!}{2^{2p}p!} \sqrt{\pi}.$$

4. Par définition, on a

$$\Gamma\left(p + \frac{1}{2}\right) = \int_0^{+\infty} t^{p-1/2} e^{-t} dt.$$

Comme indiqué, procédons au changement de variable $x = \sqrt{t}$, qui est légitime car de classe \mathcal{C}^1 et strictement croissant sur $[0, +\infty[$, avec $t = x^2$ et donc $dt = 2x dx$.

$$\Gamma\left(p + \frac{1}{2}\right) = \int_0^{+\infty} x^{2p-1} e^{-x^2} 2x dx = 2 \int_0^{+\infty} x^{2p} e^{-x^2} dx = \int_{-\infty}^{+\infty} x^{2p} e^{-x^2} dx = I_{2n}.$$

$$\text{Et donc } \Gamma\left(p + \frac{1}{2}\right) = \frac{(2p)!}{2^{2p}p!} \sqrt{\pi}.$$

SOLUTION DE L'EXERCICE 20.19

La fonction $f : x \mapsto \sqrt[3]{x^3 - 1} - \sqrt{x^2 - 1}$ est continue sur $[1, +\infty[$, donc l'intégrale est impropre en $+\infty$

$$\text{De plus, on a } f(x) = x \left(\sqrt[3]{1 - \frac{1}{x^3}} - \sqrt{1 - \frac{1}{x^2}} \right).$$

Au voisinage de $+\infty$, on a $\frac{1}{x} \rightarrow 0$ et donc

$$\sqrt[3]{1 - \frac{1}{x^3}} = 1 - \frac{1}{3x^3} + o\left(\frac{1}{x^3}\right) = 1 - o\left(\frac{1}{x^2}\right)$$

et

$$\sqrt{1 - \frac{1}{x^2}} = 1 - \frac{1}{2x^2} + o\left(\frac{1}{x^2}\right).$$

Donc

$$\sqrt[3]{1 - \frac{1}{x^3}} - \sqrt{1 - \frac{1}{x^2}} = \frac{1}{2x^2} + o\left(\frac{1}{x^2}\right) \underset{x \rightarrow +\infty}{\sim} \frac{1}{2x^2}.$$

Ainsi, $f(x) \underset{x \rightarrow +\infty}{\sim} \frac{1}{2x}$, et donc par comparaison à une intégrale de Riemann, $\int_1^{+\infty} f(x) dx$ diverge.

SOLUTION DE L'EXERCICE 20.20

1. Prouvons-le par récurrence sur n .

Par hypothèse, $0 < u_0 < e^{-1}$ et donc la récurrence est initialisée.

Supposons que $0 < u_n < \frac{1}{e}$, alors $\ln u_n < \ln(1/e) = -1$, et donc $\frac{1}{\ln u_n} > -1$.

En particulier, $1 + \frac{1}{\ln u_n} > -1 + 1 = 0$.

D'autre part, puisque $u_n < 1$, $\ln(u_n) < 0$ et donc $\frac{1}{\ln u_n} < 0$, de sorte que $1 + \frac{1}{\ln u_n} < 1$.

On en déduit donc que $0 \leq 1 + \frac{1}{\ln u_n} \leq 1$.

Et en multipliant par $0 < u_n < \frac{1}{e}$, il vient

$$0 \leq u_n \left(1 + \frac{1}{\ln u_n}\right) \leq \frac{1}{e} \Leftrightarrow 0 < u_{n+1} < e^{-1}.$$

2. On a $u_{n+1} - u_n = \frac{u_n}{\ln u_n}$, avec $u_n > 0$ et $\ln u_n < 0$, donc $u_{n+1} - u_n < 0 \Leftrightarrow u_{n+1} < u_n$.

La suite (u_n) est donc décroissante.

Puisqu'elle est minorée (par 0), elle converge.

Méthode

On sait que l'on souhaite faire apparaître au numérateur un $(2p+2)!$. Puisque nous avons déjà $(2p)!$, il ne manque que les termes $2p+2$ et $2p+1$.

Afin de les obtenir, on multiplie numérateur et dénominateur par $(2p+1)(2p+2)$.

3. Supposons que $\ell \neq 0$.

Alors $u_{n+1} \rightarrow \ell$. Mais $u_{n+1} = u_n \left(1 + \frac{1}{\ln u_n}\right) \rightarrow \ell \left(1 + \frac{1}{\ln \ell}\right)$.

Par unicité de la limite, on a donc $\ell = \ell \left(1 + \frac{1}{\ln \ell}\right)$.

Puisque $\ell \neq 0$, on peut diviser par ℓ , et il vient $1 = 1 + \frac{1}{\ln \ell}$.

Soit encore $\frac{1}{\ln \ell} = 0$, ce qui est impossible. D'où une contradiction : $\ell = 0$.

SOLUTION DE L'EXERCICE 20.21

1. Notons qu'on a $\frac{3^{2n+1}}{n!} = 3 \frac{9^n}{n!}$, et qu'on reconnaît là une série exponentielle, qui est donc convergente, et

$$\sum_{n=0}^{\infty} \frac{3^{2n+1}}{n!} = 3 \sum_{n=0}^{\infty} \frac{9^n}{n!} = 3e^9.$$

2. On a $n \frac{4^n}{3^{2n+1}} = n \frac{1}{3} \left(\frac{4}{9}\right)^n$.

Nous reconnaissons donc une série géométrique dérivée de raison $\frac{4}{9}$, donc convergente et

$$\sum_{n=0}^{\infty} n \frac{4^n}{3^{2n+1}} = \frac{4}{27} \sum_{n=0}^{\infty} n \left(\frac{4}{9}\right)^{n-1} = \frac{4}{27} \frac{1}{\left(1 - \frac{4}{9}\right)^2} = \frac{12}{25}$$

3. Nous pouvons par exemple prouver la convergence de cette série à l'aide de la règle $n^\alpha u_n$, car $n^2 u_n = \frac{n^4}{3^n} \xrightarrow{n \rightarrow \infty} 0$.

Mais afin d'en calculer la somme, on peut directement remarquer que $\frac{n^2}{3^n} = \frac{n(n-1)}{3^n} + \frac{n}{3^n}$, et que nous avons donc affaire à la somme de deux séries géométriques dérivées de raison $\frac{1}{3}$, donc toutes deux convergentes.

Au final, on a

$$\sum_{n=0}^{\infty} \frac{n^2}{3^n} = \frac{1}{9} \sum_{n=0}^{\infty} \frac{n(n-1)}{3^{n-2}} + \frac{1}{3} \sum_{n=0}^{\infty} \frac{n}{3^{n-1}} = \frac{1}{9} \frac{2}{\left(1 - \frac{1}{3}\right)^3} + \frac{1}{3} \frac{1}{\left(1 - \frac{1}{3}\right)^2} = \frac{3}{2}.$$

SOLUTION DE L'EXERCICE 20.22

1. Appliquons l'inégalité des accroissements finis à la fonction $f : t \mapsto \ln(t)$, qui est continue sur $]x, x+1]$ et dérivable sur $]x, x+1[$.

Alors sa dérivée est $f'(t) = \frac{1}{t}$, décroissante sur $]x, x+1[$, de sorte que pour tout $t \in]x, x+1[$,

$$\frac{1}{x+1} \leq f'(t) \leq \frac{1}{x}.$$

Et donc

$$\frac{1}{x+1} = \frac{1}{x+1} (x+1-x) \leq f(x+1) - f(x) \leq \frac{1}{x} (x+1-x) = \frac{1}{x}.$$

2. On a

$$\begin{aligned} u_{n+1} - u_n &= \sum_{k=1}^{n+1} \frac{1}{k} - \ln(n+1) - \sum_{k=1}^n \frac{1}{k} + \ln(n) \\ &= \frac{1}{n+1} - \ln(n+1) + \ln(n) \leq 0. \end{aligned}$$

Et donc (u_n) est décroissante.

De même, on a $v_{n+1} - v_n = \frac{1}{n+1} - \ln(n+2) + \ln(n+1) \geq 0$ et donc (v_n) est croissante.

Enfin, $u_n - v_n = \ln(n+1) - \ln(n) = \ln\left(1 + \frac{1}{n}\right) \xrightarrow{n \rightarrow +\infty} 0$.

Ainsi, (u_n) et (v_n) sont deux suites adjacentes.

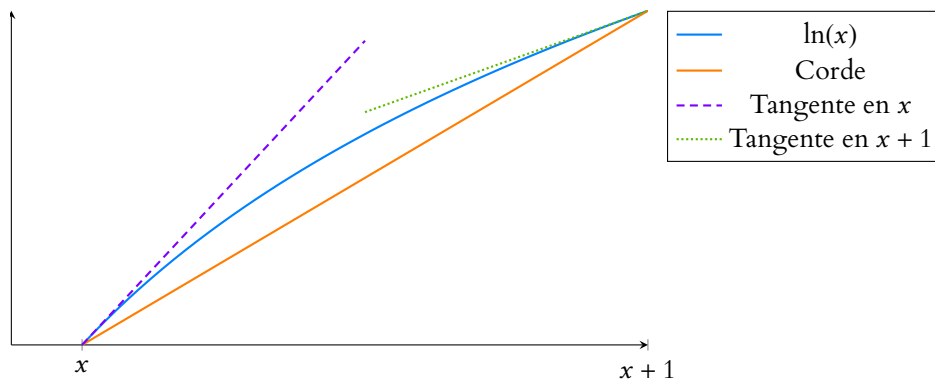


FIGURE 20.5 – Graphiquement, l'inégalité des accroissements finis dit que la pente de la corde joignant les points d'abscisses x et $x + 1$ est comprise entre la pente des tangentes en x et en $x + 1$.

3. Puisque (u_n) et (v_n) sont adjacentes, elles convergent vers une même limite γ .

$$\text{Et alors } \sum_{k=1}^n \frac{1}{k} - \ln n \xrightarrow{n \rightarrow +\infty} \gamma.$$

En particulier, $\frac{\sum_{k=1}^n \frac{1}{k}}{\ln n} - 1 \xrightarrow{n \rightarrow +\infty} 0$, de sorte que

$$\frac{\sum_{k=1}^n \frac{1}{k}}{\ln n} \xrightarrow{n \rightarrow +\infty} 1 \Leftrightarrow \sum_{k=1}^n \frac{1}{k} \underset{n \rightarrow +\infty}{\sim} \ln n.$$

Remarque : On retrouve en particulier le fait que la série harmonique diverge puisque ses sommes partielles tendent vers $+\infty$.

SOLUTION DE L'EXERCICE 20.23

1. La série $\sum \frac{1}{n^2}$ converge car il s'agit d'une série de Riemann convergente.

On a $\frac{1}{(2k+1)^2} \underset{k \rightarrow +\infty}{\sim} \frac{1}{4k^2}$, et donc par comparaison de séries à termes positifs, $\sum_k \frac{1}{(2k+1)^2}$ converge.

Enfin, la série $\sum \frac{(-1)^n}{n^2}$ converge absolument car $\left| \frac{(-1)^n}{n^2} \right| = \frac{1}{n^2}$ est le terme général d'une série de Riemann convergente.

2. Soit $N \in \mathbf{N}^*$. Alors

$$\sum_{n=1}^{2N+1} \frac{1}{n^2} = \sum_{k=0}^N \frac{1}{(2k+1)^2} + \sum_{n=1}^N \frac{1}{(2n)^2} = \sum_{k=0}^N \frac{1}{(2k+1)^2} + \frac{1}{4} \sum_{n=1}^N \frac{1}{n^2}.$$

Par passage à la limite lorsque $N \rightarrow +\infty$, il vient

$$\sum_{n=1}^{+\infty} \frac{1}{n^2} = \sum_{k=0}^{+\infty} \frac{1}{(2k+1)^2} + \frac{1}{4} \sum_{n=1}^{+\infty} \frac{1}{n^2}$$

et donc

$$\sum_{k=0}^{+\infty} \frac{1}{(2k+1)^2} = \frac{3}{4} \sum_{n=1}^{+\infty} \frac{1}{n^2} = \frac{3}{4} \frac{\pi^2}{6} = \frac{\pi^2}{8}.$$

3. Notons que $(-1)^n = \begin{cases} 1 & \text{si } n \text{ est pair} \\ -1 & \text{si } n \text{ est impair} \end{cases}$ et donc

$$\forall N \in \mathbf{N}^*, \sum_{n=1}^{2N+1} \frac{(-1)^n}{n^2} = \sum_{n=1}^N \frac{1}{(2n)^2} - \sum_{k=0}^N \frac{1}{(2k+1)^2}.$$

Ainsi, en passant à la limite lorsque $N \rightarrow \infty$, il vient

$$\sum_{n=1}^{+\infty} \frac{(-1)^n}{n^2} = \frac{1}{4} \sum_{n=1}^{+\infty} \frac{1}{n^2} - \sum_{k=0}^{+\infty} \frac{1}{(2k+1)^2} = \frac{1}{4} \frac{\pi^2}{6} - \frac{\pi^2}{8} = \frac{\pi^2}{24} - \frac{\pi^2}{8} = \frac{-\pi^2}{12}.$$

γ

Le nombre γ est appelé constante d'Euler. On ignore à peu près tout de γ , et notamment s'il est rationnel ou non.

On sait en revanche que $\gamma \approx 0.577215\dots$ (on sait en fait beaucoup mieux, les 119 milliards de premières décimales de γ ayant été calculées informatiquement).

Rédaction

Le critère des équivalents ne fonctionne que pour les séries dont le terme général est de signe constant.

Aussi, lorsqu'on utilise $u_n \sim v_n$, on n'oubliera pas de vérifier qu'au moins l'une des deux séries que l'on considère est de signe constant (l'autre l'étant alors automatiquement).

Même lorsque c'est évident, comme c'est le cas ici, mieux vaut montrer que l'on sait qu'il est indispensable de vérifier les signes.

Détails

Nous avons en fait séparé la somme en deux sommes : celle des termes d'ordre pair et celle des termes d'ordre impair, ce qui est légitime car les deux convergent.

Si c'est clair pour vous, nul besoin de repasser aux sommes partielles, il est possible de travailler directement sur les sommes de 1 à $+\infty$.

SOLUTION DE L'EXERCICE 20.24

1. Soit $\gamma \in]1, \alpha[$. Notons qu'un tel réel existe puisque $\alpha > 1$. Alors

$$\frac{u_n}{\frac{1}{n^\gamma}} = \frac{1}{n^{\alpha-\gamma} (\ln n)^\beta} \xrightarrow{n \rightarrow +\infty} 0.$$

Donc $u_n = o_{n \rightarrow +\infty} \left(\frac{1}{n^\gamma} \right)$.

Et puisque $\sum \frac{1}{n^\gamma}$ converge, alors $\sum u_n$ converge également.

2. On a

$$\frac{1}{nu_n} = \frac{\ln^\beta n}{n^{1-\alpha}} \xrightarrow{n \rightarrow +\infty} 0$$

car $1 - \alpha > 0$, et donc $\frac{1}{n} = o_{n \rightarrow +\infty} (u_n)$.

Si la série de terme général u_n convergerait, il en serait de même de la série de terme général $\frac{1}{n}$, ce qui est absurde.

Donc $\sum u_n$ diverge.

- 3.a. La fonction f_β est décroissante sur $[k, k+1]$. Et donc pour tout $x \in [k, k+1]$,

$$u_{k+1} = f_\beta(k+1) \leq f_\beta(x) \leq f_\beta(k) = u_k.$$

En intégrant cette inégalité, on a donc

$$u_{k+1} = \int_k^{k+1} u_{k+1} dx \leq \int_k^{k+1} f_\beta(x) dx \leq \int_k^{k+1} u_k dx = u_k.$$

- 3.b. Sommons les relations précédemment obtenues pour k variant de 2 à n . Il vient alors

$$\sum_{k=2}^n u_{k+1} \leq \sum_{k=2}^n \int_k^{k+1} f_\beta(x) dx \leq \sum_{k=2}^n u_k.$$

Mais par la relation de Chasles,

$$\sum_{k=3}^{n+1} u_k \leq \int_2^{n+1} f_\beta(x) dx \leq \sum_{k=2}^n u_k.$$

- 3.c. Puisque $u_k \geq 0$, la suite des sommes partielles de la série de terme général u_k est croissante.

Notons donc $S_n = \sum_{k=2}^n u_k$ ces sommes partielles.

Nous avons alors prouvé que

$$S_{n+1} - u_2 \leq \int_2^{n+1} f_\beta(x) dx \leq S_n.$$

D'autre part, pour tout $A > 2$, et pour $\beta \neq 1$,

$$\int_2^A f_\beta(x) dx = \left[\frac{1}{1-\beta} \frac{1}{(\ln x)^{\beta-1}} \right]_2^A = \frac{1}{1-\beta} \left(\frac{1}{(\ln A)^{\beta-1}} - \frac{1}{(\ln 2)^{\beta-1}} \right).$$

- Si $\beta < 1$, alors

$$\lim_{A \rightarrow +\infty} \frac{1}{(\ln A)^{\beta-1}} = (\ln A)^{1-\beta} \xrightarrow{A \rightarrow +\infty} +\infty.$$

En particulier, $\lim_{n \rightarrow +\infty} \int_2^{n+1} f_\beta(x) dx = +\infty$, et donc $\lim_{n \rightarrow +\infty} \sum_{k=2}^n u_k = +\infty$, de sorte que la série de terme général u_k diverge.

Danger !

Bien que $\ln^\beta n = o(n^\alpha)$, on n'en déduit pas que

$$\frac{1}{n^\alpha (\ln n)^\beta} \sim \frac{1}{n^\alpha}.$$

Astuce

$\frac{1}{x(\ln x)^\beta}$ est de la forme $u'u^{-\beta}$, on sait donc en calculer une primitive.

• Si $\beta > 1$: alors

$$\lim_{A \rightarrow +\infty} \frac{1}{(\ln A)^{\beta-1}} = 0$$

de sorte que $\int_2^{+\infty} f_\beta(x) dx$ converge, et vaut $\frac{1}{\beta-1} \frac{1}{(\ln 2)^{\beta-1}}$.

On a alors, pour $n \in \mathbf{N}^*$, $\int_2^{n+1} f_\beta(x) dx \leq \int_2^{+\infty} f_\beta(x) dx$.

Et donc $S_{n+1} - u_2 = \sum_{k=3}^{n+1} u_k \leq \int_2^{+\infty} f_\beta(x) dx$.

Autrement dit, la suite $(S_{n+1} - u_2)$ est majorée par $\int_2^{+\infty} f_\beta(x) dx$.

Puisqu'elle est croissante¹⁵, par le théorème de la limite monotone, elle converge, de sorte que la série de terme général u_n converge.

Et donc en passant à la limite :

$$\sum_{k=3}^{+\infty} u_k \leq \int_2^{+\infty} f_\beta(x) dx \leq \sum_{k=2}^{+\infty} u_k.$$

Soit encore

$$\sum_{k=2}^{+\infty} u_k - u_2 \leq \frac{1}{\beta-1} \frac{1}{(\ln(2))^{\beta-1}} \leq \sum_{k=2}^{+\infty} u_k.$$

On en déduit donc que

$$\frac{1}{\beta-1} \frac{1}{(\ln(2))^{\beta-1}} \leq \sum_{k=2}^{+\infty} u_k \leq \frac{1}{\beta-1} \frac{1}{(\ln(2))^{\beta-1}} + \frac{1}{2(\ln 2)^\beta}.$$

Notons enfin qu'il nous reste à traiter le cas $\beta = 1$, car alors la primitive donnée plus haut n'est plus valable.

Mais on a pour $A > 2$,

$$\int_2^A f_\beta(x) dx = \int_2^A \frac{dx}{x \ln x} = [\ln(\ln x)]_2^A = \ln(\ln A) - \ln(\ln 2) \xrightarrow{A \rightarrow +\infty} +\infty.$$

Donc on prouve alors comme pour $\beta < 1$ que $\sum u_k$ diverge.

Et au final, nous avons bien prouvé que $\sum u_k$ converge si et seulement si $\beta > 1$.

SOLUTION DE L'EXERCICE 20.25

1. Notons que $(-1)^k t^{2k} = (-t^2)^k$. Puisque $-t^2 \neq 1$, on peut donc appliquer une formule bien connue sur les sommes partielles des séries géométriques :

$$\sum_{k=0}^n (-1)^k t^{2k} = \frac{1 - (-t^2)^{n+1}}{1 + t^2} = \frac{1}{1 + t^2} + (-1)^n \frac{t^{2n+2}}{1 + t^2}.$$

Le résultat souhaité en découle alors immédiatement.

2. Intégrons l'égalité précédente entre 0 et x :

$$\sum_{k=0}^n (-1)^k \int_0^x t^{2k} dt - \int_0^x \frac{dt}{1 + t^2} = (-1)^n \int_0^x \frac{t^{2n+2}}{1 + t^2} dt.$$

Mais les intégrales de gauche se calculent aisément :

$$\int_0^x t^{2k} dt = \left[\frac{t^{2k+1}}{2k+1} \right]_0^x \text{ et } \int_0^x \frac{dt}{1 + t^2} = \text{Arctan}(x).$$

Il vient donc

$$\left| \sum_{k=0}^n (-1)^k \frac{x^{2k+1}}{2k+1} - \text{Arctan}(x) \right| = \left| \int_0^x \left(\sum_{k=0}^n (-1)^k t^{2k} - \frac{1}{1 + t^2} \right) dt \right|$$

¹⁵ Car suite des sommes partielles d'une série à termes positifs.

$$\begin{aligned}
&= \left| \int_0^x (-1)^n \frac{t^{2n+2}}{1+t^2} dt \right| \\
&\leq \int_0^x \left| (-1)^n \frac{t^{2n+2}}{1+t^2} \right| dt \\
&\leq \int_0^x \frac{t^{2n+2}}{1+t^2} dt \\
&\leq \int_0^x t^{2n+2} dt \\
&\leq \frac{x^{2n+3}}{2n+3} \\
&\leq \frac{1}{2n+3}.
\end{aligned}$$

Inégalité triangulaire.

Remarque

C'est uniquement lors de cette dernière étape que nous avons utilisé le fait que $x \in [0, 1]$.

3. On en déduit, lorsque $n \rightarrow +\infty$, que

$$\lim_{n \rightarrow +\infty} \left| \sum_{k=0}^n (-1)^k \frac{x^{2k+1}}{2k+1} - \text{Arctan}(x) \right| = 0 \Leftrightarrow \lim_{n \rightarrow +\infty} \sum_{k=0}^n (-1)^k \frac{x^{2k+1}}{2k+1} = \text{Arctan}(x).$$

Ceci prouve donc que la série de terme général $(-1)^k \frac{x^{2k+1}}{2k+1}$ converge, et que

$$\sum_{k=0}^{+\infty} (-1)^k \frac{x^{2k+1}}{2k+1} = \text{Arctan}(x).$$

Remarque

On pourrait prouver sans trop d'efforts supplémentaires (par exemple en utilisant l'imparité de Arctan que ce développement reste valable pour $x \in]-1, 1[$. En revanche, il n'est plus valable pour $|x| > 1$, car la série diverge alors grossièrement.

SOLUTION DE L'EXERCICE 20.26

1. La fonction f est dérivable, et sa dérivée est $x \mapsto \frac{1 - \ln(x)}{x^2}$, qui est négative si et seulement si $x \in [e, +\infty[$.

Donc f est croissante sur $[1, e]$ et décroissante sur $[e, +\infty[$.

2.a. On a

$$u_{2(n+1)} - u_{2n} = \sum_{k=1}^{2n+2} (-1)^k \frac{\ln k}{k} - \sum_{k=1}^{2n} (-1)^k \frac{\ln k}{k} = \frac{\ln(2n+2)}{2n+2} - \frac{\ln(2n+1)}{2n+1} = f(2n+2) - f(2n+1).$$

En particulier, cette quantité est négative d'après la première question, et donc $(u_{2n})_{n \geq 2}$ est décroissante.

De même, on prouve que

$$u_{2(n+1)+1} - u_{2n+1} = f(2n+1) - f(2n+3) \geq 0$$

et donc $(u_{2n+1})_{n \geq 2}$ est croissante.

Enfin, on a

$$u_{2n+1} - u_{2n} = \sum_{k=1}^{2n+1} (-1)^k \frac{\ln k}{k} - \sum_{k=1}^{2n} (-1)^k \frac{\ln k}{k} = -\frac{\ln(2n+1)}{2n+1} \xrightarrow{n \rightarrow +\infty} 0.$$

Et donc (u_{2n}) et (u_{2n+1}) sont deux suites adjacentes : elles convergent vers une même limite.

2.b. Puisque les suites (u_{2n}) et (u_{2n+1}) sont convergente et possèdent une même limite, alors (u_n) converge.

Puisque u_n est la somme partielle d'ordre n de la série de terme général $\frac{(-1)^k \ln k}{k}$, on en déduit que cette série converge.

3.a. On a $v_n = e^{\frac{(-1)^n}{n} \ln n} - 1$, où $\frac{(-1)^n}{n} \ln n \xrightarrow{n \rightarrow +\infty} 0$.

En utilisant le développement limité de l'exponentielle en 0, il vient

$$v_n = 1 + \frac{(-1)^n}{n} \ln n + \frac{\ln^2 n}{2n^2} + o\left(\frac{\ln^2 n}{n^2}\right) - 1 = \frac{(-1)^n}{n} \ln n + \frac{\ln^2 n}{2n^2} + o\left(\frac{\ln^2 n}{n^2}\right).$$

Méthode

Montrer que deux suites convergent vers une même limite : le premier réflexe doit être de penser à deux suites adjacentes (il se peut que ce ne soit pas le cas, mais ça vaut le coup d'essayer !).

Limite

C'est un résultat classique que (u_n) converge si et seulement si (u_{2n}) et (u_{2n+1}) convergent vers une même limite.

Ce résultat ne figure pas explicitement au programme, mais est régulièrement nécessaire au concours (par exemple il figurait à la fois dans EML 2016 et ECRI-COME 2016).

Notons bien que si (u_{2n}) et (u_{2n+1}) convergent vers des limites différentes, il ne s'applique plus, comme le montre l'exemple de la suite définie par $u_n = (-1)^n$.

3.b. \triangleleft L'erreur à ne pas faire ici : affirmer que $v_n \underset{n \rightarrow +\infty}{\sim} \frac{(-1)^n \ln n}{n}$ (ce qui est vrai !), et que puisque $\sum \frac{(-1)^n \ln n}{n}$ converge, alors il en est de même de $\sum v_n$.

En effet, le critère de comparaison ne s'applique pas ici puisque $\frac{(-1)^n \ln n}{n}$ n'est pas de signe constant !

La série de terme général $\frac{\ln^2 n}{n^2}$ converge, car $\frac{\ln^2 n}{n^2} = o\left(\frac{1}{n\sqrt{n}}\right)$. Et donc, par critère de domination¹⁶, il en est de même de la série de terme général $o\left(\frac{\ln^2 n}{n^2}\right)$.

Enfin, nous avons prouvé à la première question que $\sum \frac{(-1)^n \ln n}{n}$ converge.

Ainsi, $\sum v_n$ converge car somme de trois séries convergentes.

¹⁶ La série $\sum \frac{1}{n\sqrt{n}}$ est une série de Riemann convergente.

SOLUTION DE L'EXERCICE 20.27

On a

$$u_n = \left(1 + \frac{2}{n}\right)^n \left(\frac{\left(1 + \frac{1}{n}\right)^{2n}}{\left(1 + \frac{2}{n}\right)^n} - 1\right) = \left(1 + \frac{2}{n}\right)^n \left(e^{2n \ln\left(1 + \frac{1}{n}\right) - n \ln\left(1 + \frac{2}{n}\right)} - 1\right).$$

Notons déjà que

$$\left(1 + \frac{2}{n}\right)^n = e^{n \ln\left(1 + \frac{2}{n}\right)}$$

et que $n \ln\left(1 + \frac{2}{n}\right) \underset{n \rightarrow +\infty}{\sim} 2 \underset{n \rightarrow +\infty}{\longrightarrow} 2$ de sorte que, par continuité de l'exponentielle,

$$\left(1 + \frac{2}{n}\right)^n \underset{n \rightarrow +\infty}{\longrightarrow} e^2.$$

D'autre part,

$$2n \ln\left(1 + \frac{1}{n}\right) - n \ln\left(1 + \frac{2}{n}\right) = 2n \left(\frac{1}{n} - \frac{1}{2n^2} + o\left(\frac{1}{n^2}\right)\right) - n \left(\frac{2}{n} - \frac{2}{n^2} + o\left(\frac{1}{n^2}\right)\right) = \frac{1}{n} + o\left(\frac{1}{n}\right) \underset{n \rightarrow +\infty}{\longrightarrow} 0.$$

Et donc, à l'aide de l'équivalent classique $e^x - 1 \underset{x \rightarrow 0}{\sim} x$,

$$e^{2n \ln\left(1 + \frac{1}{n}\right) - n \ln\left(1 + \frac{2}{n}\right)} - 1 \underset{n \rightarrow +\infty}{\sim} 2n \ln\left(1 + \frac{1}{n}\right) - n \ln\left(1 + \frac{2}{n}\right) \underset{n \rightarrow +\infty}{\sim} \frac{1}{n}.$$

On en déduit que

$$u_n \underset{n \rightarrow +\infty}{\sim} \frac{e^2}{n}.$$

Et donc, d'après le critère des équivalents pour les séries à termes positifs, la série de terme général u_n diverge.

Signe

Rappelons que pour utiliser le critère des équivalents, il suffit d'avoir la positivité d'une seule des deux séries, l'autre étant alors automatique. Ici, il n'est pas évident que $u_n \geq 0$, mais puisque $\frac{e^{-2}}{n} \geq 0$, le critère des équivalents s'applique tout de même.

EML 2015

Dans tout le problème, on note E l'ensemble des fonctions $u : \mathbf{R}_+ \rightarrow \mathbf{R}$ continues sur \mathbf{R}^+ vérifiant :

$$\text{il existe } p \in \mathbf{N} \text{ tel que : } \lim_{t \rightarrow +\infty} \frac{u(t)}{t^p} = 0.$$

On remarquera que l'entier p dépend a priori de la fonction u considérée.

Partie I : Définition de la transformée de Laplace

1. Montrer que E est un sous-espace vectoriel du \mathbf{R} -espace vectoriel des fonctions définies sur \mathbf{R}^+ et à valeurs dans \mathbf{R} .

2. Soit u un élément de E .

$$\text{Montrer, pour tout } x \in]0, +\infty[: \lim_{t \rightarrow +\infty} t^2 u(t) e^{-xt} = 0.$$

En déduire que, pour tout $x \in]0, +\infty[$, l'intégrale $\int_0^{+\infty} u(t) e^{-xt} dt$ est convergente.

Dans toute la suite du problème, pour tout élément u de E , on définit la fonction $L(u)$ sur $]0; +\infty[$ par :

$$\forall x \in]0, +\infty[, L(u)(x) = \int_0^{+\infty} u(t) e^{-xt} dt.$$

3. Montrer, pour tout $\alpha \in \mathbf{R}$ et pour tout $(u, v) \in E^2 : L(\alpha u + v) = \alpha L(u) + L(v)$.

Partie II : Quelques exemples

4. Soit $a \in \mathbf{R}_+^*$ fixé. On considère, pour tout i de $\{0, 1, 2\}$ la fonction $v_i : \mathbf{R}_+ \rightarrow \mathbf{R}$ définie par

$$\forall t \in \mathbf{R}_+, v_i(t) = t^i e^{-at}.$$

Pour tout i de $\{0, 1, 2\}$, montrer que la fonction v_i appartient à E et, en utilisant par exemple des résultats sur la loi exponentielle, calculer, pour tout $x \in]0, +\infty[$, $L(v_i)(x)$.

5. Soit $n \in \mathbf{N}$ fixé. On considère la fonction $w_n : \mathbf{R}^+ \rightarrow \mathbf{R}$ définie par : $\forall t \in \mathbf{R}_+, w_n(t) = t^n$.

Montrer que la fonction w_n appartient à E et montrer, pour tout $x \in]0; +\infty[$: $L(w_n)(x) = \frac{n!}{x^{n+1}}$.

Partie III : Propriétés des transformées de Laplace

6. **Limite de $L(u)$ en $+\infty$**

Soit u un élément de E . Justifier qu'il existe $p \in \mathbf{N}$ et $A, M \in \mathbf{R}_+$ tels que

$$\forall t \in [A; +\infty[, |u(t)| \leq t^p \text{ et } \forall t \in [0, A], |u(t)| \leq M.$$

Établir : $\forall t \in \mathbf{R}_+, |u(t)| \leq M + t^p$. En déduire : $\lim_{x \rightarrow +\infty} L(u)(x) = 0$.

7. **Limite de $L(u)$ en 0**

Soit u un élément de E tel que l'intégrale $\int_0^{+\infty} u(s) ds$ converge.

On note, pour tout t de \mathbf{R}_+ , $R(t) = \int_t^{+\infty} u(s) ds$.

a. Déterminer la limite en $+\infty$ de R . Montrer que R appartient à E .

b. Montrer que R est de classe \mathcal{C}^1 sur \mathbf{R}_+ et : $\forall t \in \mathbf{R}_+, R'(t) = -u(t)$.

c. En déduire : $\forall x \in]0; +\infty[, L(u)(x) = R(0) - xL(R)(x)$.

d. Soit $\varepsilon \in]0; +\infty[$. Justifier qu'il existe $B \in [0; +\infty[$ tel que : $\forall t \in [B; +\infty[, |R(t)| \leq \frac{\varepsilon}{2}$.

$$\text{En déduire : } \forall x \in]0; +\infty[, |L(u)(x) - R(0)| \leq x \int_0^B |R(t)| dt + \frac{\varepsilon}{2}.$$

e. Conclure : $\lim_{x \rightarrow 0} L(u)(x) = \int_0^{+\infty} u(t) dt$.

8. **Transformée de Laplace d'une dérivée**

Soit $u : \mathbf{R}_+ \rightarrow \mathbf{R}$ une fonction de classe \mathcal{C}^1 sur \mathbf{R}_+ telle que $u' \in E$.

- a. Montrer qu'il existe $p \in \mathbf{N}$ et $A \in [0; +\infty[$ tels que $\forall t \in [A; +\infty[, |u(t)| \leq |u(A)| + \frac{t^{p+1}}{p+1}$.
- b. En déduire que u appartient à E .
- c. Établir $\forall x \in]0; +\infty[, L(u')(x) = -u(0) + xL(u)(x)$.

9. Dérivée puis dérivée n -ième d'une transformée de Laplace

Soit u un élément de E . On considère, pour tout n de \mathbf{N}^* , la fonction $u_n : \mathbf{R}_+ \rightarrow \mathbf{R}$ définie par

$$\forall t \in \mathbf{R}_+, u_n(t) = t^n u(t).$$

- a. Montrer que, pour tout n de \mathbf{N}^* , la fonction u_n appartient à E .
- b. Montrer $\forall a \in \mathbf{R}, |e^a - 1 - a| \leq \frac{a^2}{2} e^{|a|}$.
- c. Soient $x \in]0; +\infty[$ et $h \in \mathbf{R}^*$ tel que $|h| \leq \frac{x}{2}$.
 Montrer $\forall t \in [0; +\infty[, \left| \frac{e^{-(x+h)t} - e^{-xt}}{h} + te^{-xt} \right| \leq |h| \frac{t^2}{2} e^{-xt/2}$.
 En déduire $\left| \frac{L(u)(x+h) - L(u)(x)}{h} + L(u_1)(x) \right| \leq \frac{|h|}{2} \int_0^{+\infty} t^2 |u(t)| e^{-xt/2} dt$.
- d. Montrer que $L(u)$ est dérivable sur $]0; +\infty[$ et exprimer $(L(u))'$ à l'aide de $L(u_1)$.
- e. Montrer que $L(u)$ est de classe $\mathcal{C}^{+\infty}$ sur $]0; +\infty[$ et exprimer, pour tout n de \mathbf{N}^* , $(L(u))^{(n)}$ à l'aide de $L(u_n)$.

Partie IV : Application à la résolution d'une équation fonctionnelle

Dans cette partie, on cherche à déterminer une fonction $u : \mathbf{R}^+ \rightarrow \mathbf{R}$ de classe \mathcal{C}^2 sur \mathbf{R}_+ vérifiant

$$\forall t \in \mathbf{R}^+, u''(t) + 5u'(t) + 6u(t) = e^{-1} \text{ et } u(0) = 1 \text{ et } u'(0) = -2.$$

- 10. On suppose qu'il existe une fonction $u : \mathbf{R}^+ \rightarrow \mathbf{R}$ de classe \mathcal{C}^2 sur \mathbf{R}^+ , solution du problème et telle que $u'' \in E$.
 - a. Montrer que u appartient à E et $\forall x \in]0; +\infty[, L(u'')(x) = -xu(0) - u'(0) + x^2L(u)(x)$.
 - b. En déduire $\forall x \in]0; +\infty[, (x^2 + 5x + 6)L(u)(x) = \frac{1}{x+1} + 3 + x$,
 puis $\forall x \in]0; +\infty[, L(u)(x) = \frac{1}{2} \frac{1}{x+1} + \frac{1}{2} \frac{1}{x+3}$.
- 11. En déduire une fonction $u : \mathbf{R}^+ \rightarrow \mathbf{R}$ solution du problème posé.

EML 2015 : CORRECTION

Partie I : Définition de la transformée de Laplace

1. Soient $u, v \in E$, et soit $\lambda \in \mathbf{R}$. Par définition de E , il existe $p, q \in \mathbf{N}$ tels que

$$\lim_{t \rightarrow +\infty} \frac{u(t)}{t^p} = \lim_{t \rightarrow +\infty} \frac{v(t)}{t^q} = 0.$$

Soit alors $r = p + q \in \mathbf{N}$. On a alors

$$\lim_{t \rightarrow +\infty} \frac{\lambda u + v}{t^r} = \lim_{t \rightarrow +\infty} \lambda \frac{1}{t^q} \frac{u(t)}{t^p} + \lim_{t \rightarrow +\infty} \frac{1}{t^p} \frac{v(t)}{t^q} = 0.$$

De plus, $\lambda u + v$ est continue, donc $\lambda u + v \in E$.

Enfin, E contient la fonction nulle (pour laquelle on peut prendre par exemple $p = 0$), et donc est non vide.

Ainsi, E est un sous-espace vectoriel de l'ensemble des fonctions continues sur \mathbf{R}^+ .

2. Soit $u \in E$, et soit $p \in \mathbf{N}$ tel que $\lim_{t \rightarrow +\infty} \frac{u(t)}{t^p} = 0$. Alors

$$t^2 u(t) e^{-xt} = t^{p+2} e^{-xt} \frac{u(t)}{t^p}.$$

Par croissance comparée, $t^{p+2} e^{-xt} \xrightarrow{t \rightarrow +\infty} 0$, et par hypothèse $\frac{u(t)}{t^p} \xrightarrow{t \rightarrow +\infty} 0$.

Ainsi

$$\lim_{t \rightarrow +\infty} t^2 u(t) e^{-xt} = 0.$$

On en déduit que $u(t) e^{-xt} = o_{t \rightarrow +\infty} \left(\frac{1}{t^2} \right)$.

Puisque $\int_1^{+\infty} \frac{dt}{t^2}$ converge, il en est de même de $\int_1^{+\infty} u(t) e^{-xt} dt$.

Enfin, $t \mapsto u(t) e^{-xt}$ est continue sur \mathbf{R}^+ , de sorte qu'il n'y a pas de problème de convergence en 0. Ainsi

$$\int_0^{+\infty} u(t) e^{-xt} dt \text{ converge.}$$

3. Soient $u, v \in E$ et $\alpha \in \mathbf{R}$. Alors

$$L(\alpha u + v) = \int_0^{+\infty} (\alpha u(t) + v(t)) e^{-xt} dt = \alpha \int_0^{+\infty} u(t) e^{-xt} dt + \int_0^{+\infty} v(t) e^{-xt} dt = \alpha L(u) + L(v).$$

Partie II : Quelques exemples

4. Pour tout $i \in \{0, 1, 2\}$, v_i est continue sur \mathbf{R}^+ et $\lim_{t \rightarrow +\infty} v_i(t) = 0$ par croissance comparée.

Donc on a bien $v_i \in E$ (et on peut prendre $p = 0$).

Soit alors $x > 0$ et soit X une variable aléatoire suivant la loi exponentielle de paramètre $a + x > 0$.

Alors une densité de X est $f : t \mapsto \begin{cases} 0 & \text{si } t \leq 0 \\ (a+x)e^{-(a+x)t} & \text{si } t > 0 \end{cases}$ et nous savons que

$$\int_0^{+\infty} f(t) dt = 1, \int_0^{+\infty} t f(t) dt = E(X) = \frac{1}{a+x}, \int_0^{+\infty} t^2 f(t) dt = E(X^2) = V(X) + E(X)^2 = \frac{2}{(a+x)^2}.$$

Ainsi, il vient

$$L(v_0)(x) = \int_0^{+\infty} e^{-xt} dt = \frac{1}{x+a} \int_0^{+\infty} f(t) dt = \frac{1}{a+x}$$

$$L(v_1)(x) = \int_0^{+\infty} t e^{-xt} dt = \frac{1}{x+a} \int_0^{+\infty} t f(t) dt = \frac{1}{(a+x)^2}$$

$$L(v_2)(x) = \int_0^{+\infty} t^2 e^{-xt} dt = \frac{1}{x+a} \int_0^{+\infty} t^2 f(t) dt = \frac{2}{(a+x)^3}$$

Remarque

N'importe quel r à la fois plus grand que p et q faisait l'affaire. On aurait pu par exemple prendre $r = \max(p, q)$.

Méthode

Les variables à densité nous fournissent la valeur d'un certain nombre d'intégrales, que l'on a calculées en cours. On peut s'épargner un certain nombre de calculs d'intégrales si on reconnaît qu'une loi usuelle se cache derrière.

5. La fonction w_n est évidemment continue car polynomiale, et $\lim_{t \rightarrow +\infty} \frac{w_n(t)}{t^{n+1}} = \lim_{t \rightarrow +\infty} \frac{1}{t} = 0$.

Ainsi, $w_n \in E$.

On a alors

$$L(w_n)(x) = \int_0^{+\infty} t^n e^{-xt} dt.$$

L'application $t \mapsto xt$ est de classe \mathcal{C}^1 et réalise une bijection strictement croissante de \mathbf{R}^+ sur lui-même, de sorte que nous pouvons procéder au changement de variable $u = xt$. On a alors

$$L(w_n)(x) = \int_0^{+\infty} \left(\frac{u}{x}\right)^n e^{-u} \frac{du}{x} = \frac{1}{x^{n+1}} \int_0^{+\infty} u^n e^{-u} du = \frac{\Gamma(n+1)}{x^{n+1}} = \frac{n!}{x^{n+1}}.$$

Partie III : Propriétés des transformées de Laplace

6. Limite de $L(u)$ en $+\infty$

Par définition de E , il existe $p \in \mathbf{N}$ tel que $\lim_{t \rightarrow +\infty} \frac{u(t)}{t^p} = 0$.

Rappelons que par définition d'une limite, cela signifie que

$$\forall \varepsilon > 0, \exists A > 0 : t \geq A \Rightarrow \left| \frac{u(t)}{t^p} \right| \leq \varepsilon.$$

En particulier, si $\varepsilon = 1$, alors il existe $A > 0$ tel que

$$\forall t \in [A; +\infty[, \left| \frac{u(t)}{t^p} \right| \leq 1 \Leftrightarrow \forall t \in [A; +\infty[, |u(t)| \leq |t^p| = t^p.$$

La fonction u étant continue sur le segment $[0, A]$, elle y est bornée, et donc il existe $M \geq 0$ tel que

$$\forall t \in [0, A], |u(t)| \leq M.$$

Ainsi, si $t \in [0, A]$, $|u(t)| \leq M \leq M + t^p$ et si $t \geq A$, $|u(t)| \leq t^p \leq M + t^p$. Alors

$$\forall t \in \mathbf{R}_+, |u(t)| \leq M + t^p.$$

On en déduit que

$$\forall x > 0, |L(u)(x)| = \left| \int_0^{+\infty} u(t) e^{-xt} dt \right| \leq \int_0^{+\infty} |u(t)| dt \leq \int_0^{+\infty} M e^{-xt} dt + \int_0^{+\infty} t^p e^{-xt} dt.$$

Mais d'après la question 5, on a

$$M \int_0^{+\infty} e^{-xt} dt = \frac{M}{x} \xrightarrow{x \rightarrow +\infty} 0$$

et de même,

$$\int_0^{+\infty} t^p e^{-xt} dt = \frac{p!}{x^{p+1}} \xrightarrow{x \rightarrow +\infty} 0.$$

Et donc par le théorème des gendarmes,

$$\lim_{x \rightarrow +\infty} |L(u)(x)| = 0 \Rightarrow \lim_{x \rightarrow +\infty} L(u)(x) = 0.$$

7. Limite de $L(u)$ en 0

- 7.a. $R(t)$ est le reste d'une intégrale convergente, c'est alors du cours que $\lim_{t \rightarrow +\infty} R(t) = 0$.

De plus, pour $x \geq 0$, on a

$$R(x) = \int_x^{+\infty} u(t) dt = \int_0^{+\infty} u(t) dt - \int_0^x u(t) dt.$$

Or, $\int_0^{+\infty} u(t) dt$ est une constante, et la fonction $x \mapsto \int_0^x u(t) dt$ est continue sur \mathbf{R}_+ , d'après le théorème fondamental de l'analyse. Par conséquent R est continue sur \mathbf{R}^+ .

Puisqu'elle est de limite nulle en $+\infty$, elle est¹ dans E .

Chgt de variable

Ce changement de variable est affine, et c'est donc à vous d'y penser tout seul, l'énoncé ne dit rien à ce sujet.

L'idée est de remarquer que l'intégrale de départ ressemble beaucoup à la fonction Γ , à l'exception du xt dans l'exponentielle. Comment se ramener alors à Γ ? En posant $u = xt$!

Rappel

Une fonction continue sur un segment est bornée et atteint ses bornes (autrement dit elle admet un maximum et un minimum). Ce résultat est faux sur un intervalle ouvert.

C'est le résultat de la question 5 avec $n = 0$.

¹ On peut prendre $p = 0$ dans la définition.

- 7.b. La fonction $x \mapsto \int_0^x u(t)dt$ est de classe \mathcal{C}^1 , de dérivée égale à u . Et donc d'après le calcul effectué à la question précédente, R est de classe \mathcal{C}^1 et

$$\forall t \in \mathbf{R}^+, R'(t) = -u(t).$$

- 7.c. Soit $x > 0$, et soit $A > 0$. Posons $v(t) = e^{-xt}$ et $w(t) = -R(t)$. Alors v et w sont deux fonctions de classe \mathcal{C}^1 sur \mathbf{R}_+ , et nous pouvons procéder à une intégration par parties sur le segment $[0, A]$:

$$\int_0^A u(t)e^{-xt} dt = [-R(t)e^{-xt}]_0^A - \int_0^A xe^{-xt}R(t)dt = R(0) - R(A)e^{-xA} - x \int_0^A R(t)e^{-xt} dt.$$

En passant à la limite lorsque $A \rightarrow +\infty$, il vient

$$L(u)(x) = \int_0^{+\infty} u(t)e^{-xt} dt = R(0) - x \int_0^{+\infty} R(t)e^{-xt} dt = \boxed{R(0) - xL(R)(x)}.$$

- 7.d. Comme à la question 6, il s'agit de la définition de $\lim_{t \rightarrow +\infty} R(t) = 0$: il existe $B > 0$ tel que

$$\forall t \in [B, +\infty[, |R(t)| \leq \frac{\varepsilon}{2}.$$

On en déduit que

$$\forall x > 0, |L(u)(x) - R(0)| = |xL(R)(x)| \leq x \int_0^{+\infty} |R(t)e^{-xt}| dt \leq x \left(\int_0^B |R(t)|e^{-xt} dt + \int_B^{+\infty} |R(t)|e^{-xt} dt \right).$$

Or, pour tout $t \in [0, B]$, on a $0 \leq e^{-xt} \leq 1$, de sorte que $\int_0^B |R(t)|e^{-xt} dt \leq \int_0^B |R(t)| dt$.

Et pour $t \geq B$, $|R(t)| \leq \frac{\varepsilon}{2}$ et donc

$$\int_B^{+\infty} |R(t)|e^{-xt} dt \leq \frac{\varepsilon}{2} \int_B^{+\infty} e^{-xt} dt.$$

Mais pour $A > B$, on a

$$\int_B^A e^{-xt} dt = \left[-\frac{e^{-xt}}{x} \right]_B^A = \frac{e^{-xB}}{x} - \frac{e^{-xA}}{x} \xrightarrow{A \rightarrow +\infty} \frac{e^{-xB}}{x}.$$

On en déduit que

$$|L(u)(x) - R(0)| \leq x \int_0^B |R(t)| dt + x \frac{\varepsilon}{2} \frac{e^{-xB}}{x} \leq \boxed{x \int_0^B |R(t)| dt + \frac{\varepsilon}{2}}.$$

- 7.e. Soit $\varepsilon > 0$ fixé et soit B comme dans la question précédente (notons qu'un tel B dépend du choix de ε). Alors pour $0 \leq x < \frac{\varepsilon}{2 \int_0^B |R(t)| dt}$, on a

$$|L(u)(x) - R(0)| \leq \frac{\varepsilon}{2 \int_0^B |R(t)| dt} \int_0^B |R(t)| dt + \frac{\varepsilon}{2} = \varepsilon.$$

Ainsi, pour tout $\varepsilon > 0$, il existe $A \left(= \frac{\varepsilon}{2 \int_0^B |R(t)| dt} \right)$ tel que pour $0 \leq x < A$,

$$|L(u)(x) - R(0)| \leq \varepsilon.$$

Nous reconnaissons là la définition de

$$\lim_{x \rightarrow 0} L(u)(x) = R(0) = \int_0^{+\infty} u(t)dt.$$

Majorations

L'erreur à ne pas faire ici serait de majorer comme précédemment e^{-xt} par 1. En effet, on aurait alors

$$|B(t)e^{-xt}| \leq \frac{\varepsilon}{2}$$

mais l'intégrale de cette dernière constante sur $[B; +\infty[$ diverge !

Méthode

On sait que, par définition d'une limite, on veut arriver à : pour x « suffisamment petit »,

$$|L(u)(x) - R(0)| \leq \varepsilon.$$

Pour préciser ce que veut dire suffisamment petit, on fait le calcul à l'envers (au brouillon), et on se rend compte qu'il faut $x \int_0^B |R(t)| dt \leq \frac{\varepsilon}{2}$, ce qui implique

$$x \leq \frac{\varepsilon}{2 \int_0^B |R(t)| dt}.$$

8. Transformée de Laplace d'une dérivée

8.a. Soient, comme à la question 6, $A > 0$ et $p \in \mathbf{N}$ tels que pour $x \geq A$, $|u'(x)| \leq x^p$.

Alors pour $t \geq A$, on a $u(t) = u(A) + \int_A^t u'(x) dx$. On en déduit que

$$\begin{aligned} |u(t)| &= \left| u(A) + \int_A^t u'(x) dx \right| \\ &\leq |u(A)| + \left| \int_A^t u'(x) dx \right| \\ &\leq |u(A)| + \int_A^t x^p dx \\ &\leq |u(A)| + \frac{1}{p+1} (t^{p+1} - A^{p+1}) \\ &\leq \boxed{|u(A)| + \frac{t^{p+1}}{p+1}}. \end{aligned}$$

8.b. En conservant les notations de la question précédente, pour $t \geq A$, on a

$$\left| \frac{u(t)}{t^{p+2}} \right| \leq \frac{|u(A)|}{t^{p+2}} + \frac{1}{p+1} \frac{1}{t} \xrightarrow{t \rightarrow +\infty} 0.$$

De plus, la fonction u est continue par hypothèse, donc u appartient à E .

8.c. Soit $A > 0$. Procédons alors à une intégration par parties sur le segment $[0, A]$:

$$\int_0^A u'(t)e^{-xt} dt = [u(t)e^{-xt}]_0^A + x \int_0^A u(t)e^{-xt} dt = u(A)e^{-xA} - u(0) + x \int_0^A u(t)e^{-xt} dt.$$

En faisant tendre A vers $+\infty$ dans les deux membres de l'égalité précédente, il vient

$$\boxed{L(u')(x) = \int_0^{+\infty} u'(t)e^{-xt} dt = -u(0) + x \int_0^{+\infty} u(t)e^{-xt} dt = -u(0) + xL(u)(x).}$$

9. Dérivée puis dérivée n -ième d'une transformée de Laplace

9.a. Puisque $u \in E$, il existe un entier p tel que $\lim_{t \rightarrow +\infty} \frac{u(t)}{t^p} = 0$.

$$\text{Mais alors } \lim_{t \rightarrow +\infty} \frac{u_n(t)}{t^{n+p}} = \lim_{t \rightarrow +\infty} \frac{u(t)}{t^p} = 0.$$

Comme de plus u est continue, il en est de même de u_n comme produit de fonctions continues. Et donc $u_n \in E$.

9.b. Appliquons l'inégalité de Taylor-Lagrange à l'ordre 2 à la fonction $f : x \mapsto e^x$.

• Si $a \geq 0$, alors sur le segment $[0, a]$, f'' est majorée par e^a et donc

$$|e^a - 1 - a| = |f(a) - f(0) - af'(0)| \leq \frac{a^2}{2} e^a = \frac{a^2}{2} e^{|a|}.$$

• Si $a < 0$, alors sur le segment $[a, 0]$, f'' est majorée par $1 \leq e^{|a|}$, de sorte que

$$|e^a - 1 - a| = |f(a) - f(0) - af'(0)| \leq \frac{a^2}{2} e^{|a|}.$$

Ainsi,

$$\boxed{\forall a \in \mathbf{R}, |e^a - 1 - a| \leq \frac{a^2}{2} e^{|a|}.$$

9.c. Pour $t > 0$, appliquons l'inégalité précédente à $a = -th$. Alors il vient

$$\forall t > 0, |e^{-th} - 1 + th| \leq \frac{t^2 h^2}{2} e^{th}.$$

⚠ Danger !

Attention aux noms des variables. La variable d'intégration doit être muette, et on ne peut pas prendre comme variable d'intégration la même variable que la borne de l'intégrale. Il n'est donc pas question d'écrire

$$\int_0^t u'(t) dt.$$

En multipliant cette inégalité par $e^{-xt} > 0$, il vient

$$\forall t > 0, |e^{-(x+h)t} - e^{-xt} + t h e^{-xt}| \leq \frac{t^2 h^2}{2} e^{(h-x)t} \leq \frac{t^2 h^2}{2} e^{-xt/2}.$$

Enfin, en divisant par $|h|$, on obtient

$$\forall t > 0, \left| \frac{e^{-(x+h)t} - e^{-xt}}{h} + t e^{-xt} \right| \leq \frac{h^2 t^2}{|h| 2} e^{-xt/2} = |h| \frac{t^2}{2} e^{-xt/2}.$$

Multiplions cette inégalité par $|u(t)|$, de sorte que

$$\forall t > 0, \left| \frac{u(t)e^{-(x+h)t} - u(t)e^{-xt}}{h} + tu(t)e^{-xt} \right| \leq |h| \frac{t^2}{2} |u(t)| e^{-xt/2}.$$

Par croissance de l'intégrale, il vient alors

$$\begin{aligned} \left| \frac{L(u)(x+h) - L(u)(x)}{h} + L(u_1)(x) \right| &= \left| \int_0^{+\infty} \left(\frac{u(t)e^{-(x+h)t} - u(t)e^{-xt}}{h} + tu(t)e^{-xt} \right) dt \right| \\ &\leq \int_0^{+\infty} \left| \frac{u(t)e^{-(x+h)t} - u(t)e^{-xt}}{h} + tu(t)e^{-xt} \right| dt \\ &\leq \int_0^{+\infty} |h| \frac{t^2}{2} |u(t)| e^{-xt/2} dt \\ &\leq \boxed{\frac{|h|}{2} \int_0^{+\infty} t^2 |u(t)| e^{-xt/2} dt.} \end{aligned}$$

9.d. Soit $x \in \mathbf{R}^+$ fixé. Alors lorsque $h \rightarrow 0$, on a

$$\lim_{h \rightarrow 0} \frac{|h|}{2} \int_0^{+\infty} t^2 |u(t)| e^{-xt/2} dt = 0.$$

Et donc d'après le théorème des gendarmes,

$$\frac{L(u)(x+h) - L(u)(x)}{h} + L(u_1)(x) \xrightarrow{h \rightarrow 0} 0 \Leftrightarrow \lim_{h \rightarrow 0} \frac{L(u)(x+h) - L(u)(x)}{h} = -L(u_1)(x).$$

Ceci signifie que L est dérivable en x et que

$$L'(u)(x) = -L(u_1)(x).$$

Ceci étant vrai pour tout $x \in \mathbf{R}_+^*$, on en déduit que $\boxed{L(u) \text{ est dérivable sur } \mathbf{R}_+^* \text{ et que } L(u)' = -L(u_1).}$

9.e. Notons que pour tout $n \in \mathbf{N}$, $tu_n(t) = u_{n+1}(t)$.

Prouvons alors par récurrence sur $n \in \mathbf{N}^*$ que $L(u)$ est n -fois dérivable sur \mathbf{R}_+^* , et que $L(u)^{(n)} = (-1)^n L(u_n)$.

Pour $n = 1$, le résultat vient d'être démontré.

Supposons donc que $L(u)$ est n -fois dérivable, et que $L(u)^{(n)} = (-1)^n L(u_n)$.

Alors en appliquant le résultat de la question précédente, à la fonction u_n , $L(u_n)$ est dérivable sur \mathbf{R}_+^* , de dérivée égale à $-L(u_{n+1})$. Et donc $L(u)$ est $(n+1)$ -fois dérivable, et

$$L(u)^{(n+1)} = (L(u)^{(n)})' = (-1)^{n+1} L(u_{n+1}).$$

Ainsi, par le principe de récurrence, pour tout $n \in \mathbf{N}$, $L(u)$ est n fois dérivable et $L(u)^{(n)} = (-1)^n L(u_n)$.

En particulier, $L(u)$ est indéfiniment dérivable : $\boxed{\text{elle est de classe } \mathcal{C}^\infty \text{ sur } \mathbf{R}_+^*}.$

Partie IV : Application à la résolution d'une équation fonctionnelle

10.a. Si $u'' \in E$, alors d'après la question 8.c, on a

$$\forall x \in]0, +\infty[, L(u'')(x) = -u'(0) + xL(u')(x).$$

Et comme nous avons prouvé en 8.b que $u' \in E$, nous pouvons de nouveau appliquer le résultat de 8.c à u' :

$$\boxed{\forall x \in]0, +\infty[, L(u'')(x) = -u'(0) + x(-u(0) - xL(u)(x)) = -u'(0) - xu(0) + x^2L(u)(x).}$$

Détails

On sait $|h| \leq \frac{x}{2}$. Donc en particulier, $h \leq \frac{x}{2}$. Il vient alors

$$h - x \leq \frac{x}{2} - x = -\frac{x}{2}$$

puis en multipliant par $t \geq 0$,

$$(h - x)t \leq -t \frac{x}{2}.$$

Dérivée

Comme souvent dans ce genre de questions, si l'on veut prouver qu'une fonction est dérivable et qu'il ne s'agit pas d'une composée de fonctions usuelles, il faudra revenir à la définition de la dérivabilité : le taux d'accroissement admet une limite finie.

- 10.b. Rappelons que si v_0 désigne la fonction $t \mapsto e^{-t}$, nous avons prouvé à la question 4 que $L(v_0)(x) = \frac{1}{x+1}$.
Ainsi, si u est une solution du problème, alors par linéarité de L , on a

$$\forall x > 0, L(u'')(x) + 5L(u')(x) + 6L(u)(x) = L(v_0)(x) = \frac{1}{x+1}.$$

Et en utilisant le résultat de la question précédente ainsi que la connaissance de $u(0)$ et $u'(0)$, il vient

$$\forall x > 0, x^2 L(u)(x) + 2 - x + 5(-1 + xL(u)(x)) + 6L(u)(x) = \frac{1}{x+1} \Leftrightarrow (x^2 + 5x + 6)L(u)(x) = \frac{1}{x+1} + 3 + x.$$

Remarquons que $x^2 + 5x + 6 = (x+2)(x+3)$, qui ne s'annule pas sur \mathbf{R}^+ . Ainsi, en divisant l'égalité précédente par $x^2 + 5x + 6$, il vient

$$\forall x > 0, L(u)(x) = \frac{1}{(x+1)(x+2)(x+3)} + \frac{1}{x+2} = \frac{x^2 + 4x + 4}{(x+1)(x+2)(x+3)} = \frac{x+2}{(x+1)(x+3)}.$$

De même, on a

$$\frac{1}{2} \left(\frac{1}{x+1} + \frac{1}{x+3} \right) = \frac{1}{2} \frac{2(x+2)}{(x+1)(x+3)} = \frac{x+2}{(x+1)(x+3)}.$$

Et donc nous avons bien

$$\forall x > 0, L(u)(x) = \frac{1}{2} \frac{1}{x+1} + \frac{1}{2} \frac{1}{x+3}.$$

11. Il s'agit là de reconnaître une fonction dont la transformée de Laplace serait égale à $\frac{1}{2} \frac{1}{x+1} + \frac{1}{2} \frac{1}{x+3}$.

Mais d'après la question 4, nous savons que si $u : t \mapsto \frac{1}{2}(e^{-t} + e^{-3t})$, alors

$$\forall x > 0, L(u)(x) = \frac{1}{2} \left(\frac{1}{x+1} + \frac{1}{x+3} \right).$$

Ceci ne suffit certainement pas à garantir qu'il s'agit bien d'une solution au problème posé, mais il semble que ce soit un bon candidat. Vérifions que u vérifie bien les conditions souhaitées :

$$\forall t > 0, u'(t) = \frac{1}{2} (-e^{-t} - 3e^{-3t}) \text{ et } u''(t) = \frac{1}{2} (e^{-t} + 9e^{-3t})$$

de sorte que

$$\forall t > 0, u''(t) + 5u'(t) + 6u(t) = e^{-t}.$$

De plus, on a $u(0) = 1$ et $u'(0) = -2$, de sorte que u est bien une solution du problème posé.

Remarque

Rien ne nous permet de garantir que u soit l'unique solution du problème posé, bien que ce soit le cas.

ECRICOME 2009

Le but de l'exercice est l'étude de la fonction f définie par la formule suivante :

$$f(x) = \int_0^{+\infty} e^{-2t} \sqrt{1+x^2 e^{2t}} dt.$$

1. Domaine de définition de f :

- a. Justifier que pour tout réel $a > 0$, l'intégrale $\int_0^{+\infty} e^{-at} dt$ converge, et donner sa valeur.
- b. Soit x un réel fixé. Établir la convergence de l'intégrale $\int_0^{+\infty} e^{-2t} \sqrt{1+x^2 e^{2t}} dt$.

Par conséquent, f est définie sur \mathbf{R} , et elle est clairement paire. On va donc l'étudier sur $]0, +\infty[$.

2. Branche infinie de la courbe représentative de f :

- a. Vérifier l'encadrement suivant, pour tout réel x strictement positif et pour tout réel t positif ou nul :

$$xe^t \leq \sqrt{1+x^2 e^{2t}} \leq xe^t + \frac{e^{-t}}{2x}.$$

- b. Prouver que, pour tout réel x strictement positif, on a :

$$x \leq f(x) \leq x + \frac{1}{6x}.$$

- c. En déduire que $\lim_{x \rightarrow +\infty} |x - f(x)| = 0$.

On dit alors que la droite d'équation $y = x$ est une asymptote oblique à la courbe représentative de f .

3. Dérivabilité et monotonie de f :

- a. À l'aide du changement de variable $u = xe^t$, que l'on justifiera, prouver la formule suivante lorsque x est un réel strictement positif :

$$f(x) = x^2 \int_x^{+\infty} \frac{\sqrt{1+u^2}}{u^3} du.$$

- b. Montrer que la fonction f est de classe \mathcal{C}^1 sur $]0, +\infty[$ et que sa dérivée est donnée, pour tout réel x strictement positif par :

$$f'(x) = \frac{2f(x) - \sqrt{1+x^2}}{x}.$$

- c. Justifier, pour tout réel x strictement positif, l'égalité suivante :

$$2f(x) = \sqrt{1+x^2} + x^2 \int_x^{+\infty} \frac{du}{u\sqrt{1+u^2}}.$$

En déduire que f est strictement croissante sur $]0, +\infty[$.

4. Étude locale de f et f' en 0 :

- a. Justifier que la formule suivante est valable pour tout réel x strictement positif :

$$\int_x^{+\infty} \frac{du}{u\sqrt{1+u^2}} = -\frac{\ln(x)}{\sqrt{1+x^2}} + \int_x^{+\infty} \frac{u \ln(u)}{(1+u^2)^{3/2}} du$$

et que l'intégrale $\int_0^{+\infty} \frac{u \ln u}{(1+u^2)^{3/2}} du$ est convergente.

- b. À l'aide des questions précédentes, démontrer alors que l'on a :

$$f'(x) \underset{x \rightarrow 0^+}{\sim} -x \ln(x) \text{ et } f(x) - \frac{1}{2} \underset{x \rightarrow 0^+}{\sim} -\frac{x^2 \ln(x)}{2}.$$

- c. En déduire que f est une fonction de classe \mathcal{C}^1 sur $[0, +\infty[$ et préciser la valeur de $f'(0)$.

ECRICOME 2009 : CORRECTION

1. Domaine de définition de f

1.a. Pour $A > 0$, on a

$$\int_0^A e^{-at} dt = \left[-\frac{1}{a} e^{-at} \right]_0^A = \frac{1}{a} (1 - e^{-aA}) \xrightarrow{A \rightarrow +\infty} \frac{1}{a}.$$

Donc $\int_0^{+\infty} e^{-at} dt$ converge et vaut $\frac{1}{a}$.

1.b. La fonction $t \mapsto e^{-2t} \sqrt{1 + x^2 e^{2t}}$ est continue sur $[0, +\infty[$, donc le seul éventuel problème de convergence est au voisinage de $+\infty$. Or, lorsque $t \rightarrow +\infty$, on a $1 + x^2 e^{2t} \underset{t \rightarrow +\infty}{\sim} x^2 e^{2t}$ et

$$\text{donc } \sqrt{1 + x^2 e^{2t}} \underset{t \rightarrow +\infty}{\sim} \sqrt{x^2 e^{2t}} \underset{t \rightarrow +\infty}{\sim} |x| e^t.$$

$$\text{Et donc } e^{-2t} \sqrt{1 + x^2 e^{2t}} \underset{t \rightarrow +\infty}{\sim} e^{-2t} e^t |x| \underset{t \rightarrow +\infty}{\sim} |x| e^{-t}.$$

Puisque l'intégrale $\int_0^{+\infty} e^{-t} dt$ converge, par critère de comparaison pour les intégrales

de fonctions positives, on en déduit que $\int_0^{+\infty} e^{-2t} \sqrt{1 + x^2 e^{2t}} dt$ converge.

2. Branche infinie de la courbe représentative de f

2.a. Soit $x > 0$ et $t > 0$. Alors $x^2 e^{2t} \leq 1 + x^2 e^{2t}$, de sorte que $x e^t = \sqrt{x^2 e^{2t}} \leq \sqrt{1 + x^2 e^{2t}}$.

D'autre part, on a $\left(x e^t + \frac{e^{-t}}{2x} \right)^2 = x^2 e^{2t} + 1 + \frac{e^{-2t}}{4x^2} \geq 1 + x^2 e^{2t}$, de sorte que par passage à la racine¹, il vient

$$\sqrt{1 + x^2 e^{2t}} \leq x e^t + \frac{e^{-t}}{2x}.$$

2.b. En multipliant les trois termes de l'inégalité précédemment obtenue par e^{-2t} , il vient

$$\forall x > 0, \forall t \geq 0, x e^{-t} \leq e^{-2t} \sqrt{1 + x^2 e^{2t}} \leq x e^{-t} + \frac{e^{-3t}}{2x}.$$

Les trois termes de cet encadrement sont d'intégrale convergente sur $]0, +\infty[$. Par croissance de l'intégrale², on a donc

$$\int_0^{+\infty} x e^{-t} dt \leq \int_0^{+\infty} e^{-2t} \sqrt{1 + x^2 e^{2t}} dt \leq x \int_0^{+\infty} e^{-t} dt + \frac{1}{2x} \int_0^{+\infty} e^{-3t} dt.$$

Mais en utilisant le résultat de la question 1.a, il vient alors

$$x \leq f(x) \leq x + \frac{1}{6x}.$$

2.c. On a $0 \leq f(x) - x \leq \frac{1}{6x}$, donc par le théorème des gendarmes, $\lim_{x \rightarrow +\infty} f(x) - x = 0$ et alors

$$\lim_{x \rightarrow +\infty} |f(x) - x| = 0.$$

3. Dérivabilité et monotonie de f

3.a. La fonction $t \mapsto x e^t$ est strictement croissante sur \mathbf{R}_+ , et de classe \mathcal{C}^1 , donc le changement de variable est justifié. Notons que l'intégrale définissant f est convergente, donc l'intégrale obtenue après changement de variable l'est encore.

Lorsque $t \rightarrow 0$, alors $u \rightarrow x$, et lorsque $t \rightarrow +\infty$, alors $u \rightarrow +\infty$. On a alors $u = x e^t \Leftrightarrow t = \ln\left(\frac{u}{x}\right)$ et donc $dt = \frac{1}{x} \frac{du}{u} = \frac{du}{u}$. Il vient donc

$$f(x) = \int_0^{+\infty} e^{-2t} \sqrt{1 + x^2 e^{2t}} dt = \int_x^{+\infty} \frac{x^2}{u^2} \sqrt{1 + u^2} \frac{du}{u} = x^2 \int_x^{+\infty} \frac{\sqrt{1 + u^2}}{u^3} du.$$

Détail

$e^{2t} = (e^t)^2$ et donc

$$\sqrt{e^{2t}} = (e^{2t})^{1/2} = e^t.$$

¹ Pas besoin de valeur absolue en passant à la racine car tous les termes sont positifs.

² Les trois termes sont d'intégrale convergente.

Méthode

Ne pas oublier que si on ne travaille pas sur un segment, il est important de vérifier la stricte monotonie du changement de variable.

3.b. Notons qu'on a

$$\int_x^{+\infty} \frac{\sqrt{1+u^2}}{u^3} du = \int_1^{+\infty} \frac{\sqrt{1+u^2}}{u^3} du - \int_1^x \frac{\sqrt{1+u^2}}{u^3} du$$

Or, $x \mapsto \int_1^x \frac{\sqrt{1+u^2}}{u^3} du$ est une primitive³, sur $]0, +\infty[$, de $x \mapsto \frac{\sqrt{1+x^2}}{x^3}$, donc de classe \mathcal{C}^1 .

On en déduit⁴ donc que f est de classe \mathcal{C}^1 sur $]0, +\infty[$, et que sa dérivée est

$$f'(x) = 2x \int_x^{+\infty} \frac{\sqrt{1+u^2}}{u^3} du - x^2 \frac{\sqrt{1+x^2}}{x^3} = \boxed{\frac{2f(x) - \sqrt{1+x^2}}{x}}.$$

³ Et même la primitive qui s'annule en 1.

⁴ $\int_1^{+\infty} \frac{\sqrt{1+u^2}}{u^3} du$ est une constante, qui ne dépend pas de x , donc sa dérivée est nulle.

3.c. Soit $x > 0$ et $A > x$. Alors, en posant $g(u) = \sqrt{1+u^2}$ et $h(u) = -\frac{1}{2u^2}$, qui sont deux fonctions \mathcal{C}^1 sur $[x, A]$ avec $g'(u) = \frac{u}{\sqrt{1+u^2}}$ et $h'(u) = \frac{1}{u^3}$, une intégration par parties sur le segment $[x, A]$ nous donne

$$\int_x^A \frac{\sqrt{1+u^2}}{u^3} du = \left[-\frac{\sqrt{1+u^2}}{2u^2} \right]_x^A + \int_x^A \frac{du}{2u\sqrt{1+u^2}} du = \frac{\sqrt{1+x^2}}{2x^2} - \frac{\sqrt{1+A^2}}{2A^2} + \int_x^A \frac{du}{2u\sqrt{1+u^2}}.$$

Lorsque $A \rightarrow +\infty$, on a

$$\frac{\sqrt{1+A^2}}{2A^2} \underset{A \rightarrow +\infty}{\sim} \frac{\sqrt{A^2}}{2A^2} \underset{A \rightarrow +\infty}{\sim} \frac{1}{2A} \underset{A \rightarrow +\infty}{\rightarrow} 0.$$

D'autre part, $\frac{1}{2u\sqrt{1+u^2}} \underset{u \rightarrow +\infty}{\sim} \frac{1}{2u^{3/2}}$, de sorte que par comparaison à une intégrale de

Riemann, $\int_x^{+\infty} \frac{du}{2u\sqrt{1+u^2}}$ converge et donc

$$\lim_{A \rightarrow +\infty} \int_x^A \frac{du}{2u\sqrt{1+u^2}} = \int_x^{+\infty} \frac{du}{2u\sqrt{1+u^2}}.$$

On en déduit que

$$\int_x^{+\infty} \frac{\sqrt{1+u^2}}{u^3} du = \lim_{A \rightarrow +\infty} \int_x^A \frac{\sqrt{1+u^2}}{u^3} du = \frac{\sqrt{1+x^2}}{2x^2} + \int_x^{+\infty} \frac{du}{2u\sqrt{1+u^2}}.$$

Et donc, en multipliant par $2x^2$, $2f(x) = \sqrt{1+x^2} + x^2 \int_x^{+\infty} \frac{du}{u\sqrt{1+u^2}}$.

On en déduit que $f'(x) = \frac{1}{x} \int_x^{+\infty} \frac{du}{u\sqrt{1+u^2}}$. La fonction $u \mapsto \frac{1}{u\sqrt{1+u^2}}$ est strictement positive sur \mathbf{R}_+ , donc son intégrale sur $[x, +\infty[$ aussi. Puisqu'il en est de même de $\frac{1}{x}$, on en déduit que pour tout $x > 0$, $f'(x) > 0$, et donc f est strictement croissante sur \mathbf{R}_+^* .

4. Étude locale de f et f' en 0

4.a. Soit $x > 0$ fixé, et soit $A > x$. Alors, en posant $g(u) = \frac{1}{\sqrt{1+u^2}}$ et $h(u) = \ln u$, qui sont deux fonctions de classe \mathcal{C}^1 avec $g'(u) = -\frac{u}{(1+u^2)^{3/2}}$ et $h'(u) = \frac{1}{u}$, une intégration par parties sur le segment $[x, A]$ nous donne

$$\int_x^A \frac{du}{u\sqrt{1+u^2}} = \left[\frac{\ln(u)}{\sqrt{1+u^2}} \right]_x^A + \int_x^A \frac{u \ln u}{(1+u^2)^{3/2}} du = \frac{\ln(A)}{\sqrt{1+A^2}} - \frac{\ln(x)}{\sqrt{1+x^2}} + \int_x^A \frac{u \ln(u)}{\sqrt{1+u^2}} du.$$

Lorsque $A \rightarrow +\infty$, on a $\frac{\ln(A)}{\sqrt{1+A^2}} \underset{A \rightarrow +\infty}{\sim} \frac{\ln(A)}{A} \underset{A \rightarrow +\infty}{\rightarrow} 0$.

De plus, au voisinage de $+\infty$, on a $u^\alpha \frac{u \ln(u)}{(1+u^2)^{3/2}} \underset{u \rightarrow +\infty}{\sim} u^\alpha \frac{u \ln u}{u^3} \underset{u \rightarrow +\infty}{\sim} u^{\alpha-2} \ln u$.

Convergence

Il n'est pas immédiat que l'intégrale

$$\int_x^A \frac{du}{2u\sqrt{1+u^2}}$$

admette une limite lorsque $A \rightarrow +\infty$, et donc il est impératif de vérifier que

$$\int_x^{+\infty} \frac{du}{2u\sqrt{1+u^2}}$$

converge avant de passer à la limite.

En particulier, pour $\alpha = 3/2$, on a $u^{3/2} \frac{u \ln u}{(1+u^2)^{3/2}} \xrightarrow{u \rightarrow +\infty} 0$, de sorte que

$$\frac{u \ln u}{(1+u^2)^{3/2}} = o_{u \rightarrow +\infty} \left(\frac{1}{u^{3/2}} \right).$$

Par comparaison à une intégrale de Riemann convergente⁵, on en déduit que $\int_0^{+\infty} \frac{u \ln u}{(1+u^2)^{3/2}} du$ converge.

Il en est donc de même de $\int_x^A \frac{u \ln u}{(1+u^2)^{3/2}} du$, de sorte que lorsque $A \rightarrow +\infty$,

$$\int_x^A \frac{u \ln u}{(1+u^2)^{3/2}} du \xrightarrow{A \rightarrow +\infty} \int_x^{+\infty} \frac{u \ln u}{(1+u^2)^{3/2}} du.$$

On déduit donc de ce qui précède que

$$\int_x^{+\infty} \frac{du}{u\sqrt{1+u^2}} = \lim_{A \rightarrow +\infty} \int_x^A \frac{du}{u\sqrt{1+u^2}} = \boxed{-\frac{\ln x}{\sqrt{1+x^2}} + \int_x^{+\infty} \frac{u \ln u}{(1+u^2)^{3/2}} du.}$$

Une remarque concernant l'intégration par parties : on pourrait s'épargner une étude de convergence en remarquant que

$$\int_x^A \frac{u \ln(u)}{\sqrt{1+u^2}} du = \int_x^A \frac{du}{u\sqrt{1+u^2}} - \frac{\ln(A)}{\sqrt{1+A^2}} + \frac{\ln(x)}{\sqrt{1+x^2}}$$

et que $\frac{\ln(A)}{\sqrt{1+A^2}}$ et $\int_x^A \frac{du}{u\sqrt{1+u^2}}$ admettent tous deux⁶ une limite lorsque $A \rightarrow +\infty$.

Et donc $\int_x^A \frac{u \ln(u)}{\sqrt{1+u^2}} du$ admet une limite lorsque $A \rightarrow +\infty$, de sorte que $\int_x^{+\infty} \frac{u \ln(u)}{\sqrt{1+u^2}} du$ converge.

Le même raisonnement aurait pu être tenu à la question 3.c pour éviter de prouver la convergence de $\int_x^{+\infty} \frac{du}{2u\sqrt{1+u^2}}$.

Enfin, lorsque $u \rightarrow 0^+$, $\frac{u \ln u}{(1+u^2)^{3/2}} \underset{u \rightarrow 0^+}{\sim} u \ln u \xrightarrow{u \rightarrow 0^+} 0$ par croissances comparées.

Donc $\int_0^1 \frac{u \ln u}{(1+u^2)^{3/2}} du$ est faussement impropre en 0, donc convergente.

Et donc⁷ $\int_0^1 \frac{u \ln u}{(1+u^2)^{3/2}} du$ converge également.

4.b. On a

$$\begin{aligned} f'(x) &= x \int_x^{+\infty} \frac{du}{u\sqrt{1+u^2}} du = -\frac{x \ln x}{\sqrt{1+x^2}} + x \int_x^{+\infty} \frac{u \ln u}{(1+u^2)^{3/2}} du \\ &= x \left(-\frac{\ln x}{\sqrt{1+x^2}} + \int_x^{+\infty} \frac{u \ln u}{(1+u^2)^{3/2}} du \right). \end{aligned}$$

Ainsi, on a $\int_x^{+\infty} \frac{u \ln u}{(1+u^2)^{3/2}} du \xrightarrow{x \rightarrow 0^+} \int_0^{+\infty} \frac{u \ln u}{(1+u^2)^{3/2}} du$, qui est une constante, alors que $-\frac{\ln x}{\sqrt{1+x^2}} \underset{x \rightarrow 0^+}{\sim} -\ln(x) \rightarrow +\infty$.

On en déduit que $\int_x^{+\infty} \frac{u \ln u}{(1+u^2)^{3/2}} du = o_{x \rightarrow 0^+} \left(-\frac{\ln x}{\sqrt{1+x^2}} \right)$ et donc

$$-\frac{\ln x}{\sqrt{1+x^2}} + \int_x^{+\infty} \frac{u \ln u}{(1+u^2)^{3/2}} du \underset{x \rightarrow 0^+}{\sim} -\frac{\ln x}{\sqrt{1+x^2}} \underset{x \rightarrow 0^+}{\sim} -\ln x.$$

Et donc $\boxed{f'(x) \underset{x \rightarrow 0^+}{\sim} -x \ln x.}$

Méthode

On souhaiterait prouver la convergence par comparaison à une série de Riemann, mais laquelle ? (On a peut-être remarqué que $\alpha = 2$ ne fonctionne pas). Pour trouver le bon α , on commence par écrire le quotient avec un α quelconque, puis on choisit α tel que le quotient tende vers 0.

⁵ Car $\alpha = 3/2$.

⁶ L'intégrale admet une limite puisque nous savons déjà que $\int_x^{+\infty} \frac{du}{u\sqrt{1+u^2}}$ converge.

⁷ Nous avons déjà prouvé qu'elle convergeait au voisinage de $+\infty$.

Détails

Une fonction qui tend vers une constante est toujours négligeable devant une fonction qui tend vers $\pm\infty$ (écrire le quotient pour s'en convaincre).

o/\sim

Rappelons que $f \sim g$ si et seulement si

$$f = g + o(g).$$

De même, en combinant les résultats des questions 3.c et 4.a, on a

$$f(x) - \frac{1}{2} = \frac{\sqrt{1+x^2}}{2} - \frac{1}{2} + \frac{x^2}{2} \left(-\frac{\ln x}{\sqrt{1+x^2}} + \int_x^{+\infty} \frac{u \ln u}{(1+u^2)^{3/2}} du \right).$$

Mais d'une part, nous savons⁸ que le second terme est équivalent à $-\frac{x^2}{2} \ln x$.

⁸ C'est le calcul réalisé à la question précédente.

Soit encore $f(x) = -\frac{x^2 \ln x}{2} + o(x^2 \ln x)$.

D'autre part,

$$\sqrt{1+x^2} - 1 = 1 + \frac{1}{2}x^2 + o_{x \rightarrow 0}(x^2) - 1 = \frac{x^2}{2} + o_{x \rightarrow 0}(x^2) \underset{x \rightarrow 0}{\sim} \frac{x^2}{2}$$

de sorte que $\frac{\sqrt{1+x^2}}{2} - \frac{1}{2} \underset{x \rightarrow 0^+}{\sim} \frac{x^2}{4}$.

Mais $\frac{x^2}{4} = o_{x \rightarrow 0^+}(x^2 \ln x)$ et donc on en déduit que $\frac{\sqrt{1+x^2}}{2} - \frac{1}{2} = o_{x \rightarrow 0^+}(x^2 \ln x)$.

Et donc

$$f(x) - \frac{1}{2} = -\frac{x^2 \ln x}{2} + o_{x \rightarrow 0^+}(x^2 \ln x) + o_{x \rightarrow 0^+}(x^2 \ln x) = -\frac{x^2 \ln x}{2} + o_{x \rightarrow 0^+}(x^2 \ln x) \underset{x \rightarrow 0^+}{\sim} -\frac{x^2 \ln x}{2}.$$

4.c. On a, par la question 1.a, $f(0) = \int_0^{+\infty} e^{-2t} dt = \frac{1}{2}$ et donc

$$\lim_{x \rightarrow 0^+} \frac{f(x) - f(0)}{x} = \lim_{x \rightarrow 0^+} \frac{f(x) - \frac{1}{2}}{x} = \lim_{x \rightarrow +\infty} \frac{-x \ln x}{2} = 0.$$

Donc f est dérivable en 0 et $f'(0) = 0$.

D'autre part, $\lim_{x \rightarrow 0^+} f'(x) = \lim_{x \rightarrow 0^+} -x \ln x = 0 = f'(0)$.

Donc f' est continue en 0. Puisque nous avons déjà prouvé que f est de classe \mathcal{C}^1 sur

$]0, +\infty[$, on en déduit que f est de classe \mathcal{C}^1 sur $[0, +\infty[$.

DL

On rappelle qu'une fonction est toujours équivalente au premier terme **non nul** de son développement limité.

ECRICOME 2007

1. Montrer que lorsque x est au voisinage de 0 on a :

$$\ln(2 - e^x) = -x - x^2 + o(x^2)$$

2. a. Montrer que pour tout entier k supérieur ou égal à 2, on a :

$$2 - e^{1/k} \in]0, 1[$$

- b. En déduire le signe de $\ln(2 - e^{1/k})$, pour tout entier k supérieur ou égal à 2.
 c. Quelle est la nature de la série de terme général $\ln(2 - e^{1/k})$?
 d. Pour n entier supérieur ou égal à 2, on pose :

$$V_n = \sum_{k=2}^n \ln(2 - e^{1/k}) \text{ et } u_n = \exp V_n.$$

Déterminer $\lim_{n \rightarrow +\infty} V_n$ et $\lim_{n \rightarrow +\infty} u_n$.

3. a. Montrer que :

$$\ln(nu_n) = \sum_{k=2}^n \left[\ln(2 - e^{1/k}) - \ln\left(1 - \frac{1}{k}\right) \right]$$

- b. Déterminer un équivalent, quand k tend vers $+\infty$, de $\ln(2 - e^{1/k}) - \ln\left(1 - \frac{1}{k}\right)$.
 c. En déduire que u_n est équivalent, quand n tend vers $+\infty$, à $\frac{K}{n}$ avec $K > 0$.
 Quelle est la nature de la série de terme général u_n ?

4. On pose

$$S_n = \sum_{k=2}^n (-1)^k u_k.$$

- a. Étudier le sens de variations de la suite $(u_n)_{n \geq 2}$.
 b. Montrer que les suites $(S_{2n})_{n \geq 1}$ et $(S_{2n+1})_{n \geq 1}$ sont deux suites adjacentes.
 c. En déduire la nature de la série de terme général $(-1)^n u_n$.

ECRICOME 2007 : CORRECTION

1. La fonction $\varphi : x \mapsto \ln(2 - e^x)$ est de classe \mathcal{C}^2 au voisinage de 0, avec

$$\varphi'(x) = \frac{e^x}{e^x - 2} \text{ et } \varphi''(x) = \frac{-2e^x}{(e^x - 2)^2}.$$

D'après la formule de Taylor-Young à l'ordre 2, on a alors

$$\ln(2 - e^x) \underset{0}{=} \varphi(0) + \varphi'(0)x + \varphi''(0)\frac{x^2}{2} + o(x^2) = \boxed{-x - x^2 + o(x^2)}.$$

- 2.a. Pour $k \geq 2$, on a $\frac{1}{k} > 0$, donc $e^{\frac{1}{k}} > 1$ et donc $2 - e^{1/k} < 1$.

D'autre part, $\frac{1}{k} \leq \frac{1}{2}$ et donc $e^{1/k} \leq e^{1/2}$.

Puisque $e < 4$, $e^{1/2} < 2$, et donc $e^{1/k} < 2$, de sorte que $2 - e^{1/k} > 0$.

- 2.b. Nous savons que pour $x \in]0, 1[$, $\ln(x) < 0$, donc pour tout entier k supérieur ou égal à 2,

$$\boxed{\ln(2 - e^{1/k}) < 0}.$$

- 2.c. Lorsque $k \rightarrow +\infty$, alors $\frac{1}{k} \rightarrow 0$, donc on peut utiliser l'équivalent fourni par la première question : $\ln(2 - e^{1/k}) \underset{k \rightarrow +\infty}{\sim} -\frac{1}{k}$.

Mais nous savons que la série de terme général $\frac{1}{k}$ diverge, et donc par critère de comparaison pour les séries de signe constant¹, $\boxed{\text{la série de terme général } \ln(2 - e^{1/k}) \text{ diverge.}}$

- 2.d. (V_n) est une suite décroissante car suite des sommes partielles d'une série dont le terme général est toujours négatif. De plus, cette série diverge, donc (V_n) est divergente. Mais une suite décroissante et divergente admet nécessairement $-\infty$ comme limite : $\boxed{\lim_{n \rightarrow +\infty} V_n = -\infty}$.

Et alors, par composition de limites, $\boxed{\lim_{n \rightarrow +\infty} u_n = 0}$.

- 3.a. D'une part, on a $\ln(nu_n) = \ln(n) + \ln(u_n) = \ln(n) + V_n$.
D'autre part, notons que

$$\sum_{k=2}^n \ln\left(1 - \frac{1}{k}\right) = \sum_{k=2}^n \ln\left(\frac{k-1}{k}\right) = \sum_{k=2}^n \ln(k-1) - \ln(k).$$

Nous reconnaissons là une somme télescopique :

$$\sum_{k=2}^n \ln\left(1 - \frac{1}{k}\right) = \ln(1) - \ln(2) + \ln(2) - \ln(3) + \dots + \ln(n-2) - \ln(n-1) + \ln(n-1) - \ln(n) = -\ln(n).$$

Et donc

$$\sum_{k=2}^n \left[\ln(2 - e^{1/k}) - \ln\left(1 - \frac{1}{k}\right) \right] = V_n - (-\ln(n)) = V_n + \ln(n) = \boxed{\ln(nu_n)}.$$

- 3.b. Nous savons que $\ln(2 - e^{1/k}) = -\frac{1}{k} - \frac{1}{k^2} + o\left(\frac{1}{k^2}\right)$ et $\ln\left(1 - \frac{1}{k}\right) = -\frac{1}{k} - \frac{1}{2k^2} + o\left(\frac{1}{k^2}\right)$, donc on en déduit que

$$\ln(2 - e^{1/k}) - \ln\left(1 - \frac{1}{k}\right) = -\frac{1}{2k^2} + o\left(\frac{1}{k^2}\right) \text{ et donc } \boxed{\ln(2 - e^{1/k}) - \ln\left(1 - \frac{1}{k}\right) \underset{k \rightarrow +\infty}{\sim} -\frac{1}{2k^2}}.$$

- 3.c. Grâce à l'équivalent obtenu à la question précédente, par critère de comparaison pour les séries de signe constant (négatif ici car $-\frac{1}{2k^2} \leq 0$), on en déduit que la série de terme général $\ln(2 - e^{1/k}) - \ln\left(1 - \frac{1}{k}\right)$ converge, et donc il existe une constante A telle que

$$\lim_{n \rightarrow +\infty} \sum_{k=2}^n \left[\ln(2 - e^{1/k}) - \ln\left(1 - \frac{1}{k}\right) \right] = A.$$

DL/équivalents

Un développement limité fourni un équivalent : il suffit de prendre le premier terme **non nul** de la partie principale, donc ici

$$\ln(2 - e^x) \underset{0}{\sim} -x$$

¹ $-\frac{1}{k}$ est de signe constant !

Terminologie

Une suite qui a une limite égale à $\pm\infty$ est divergente (même si elle admet une limite) !

⚠ Danger !

On n'ajoute pas les équivalents ! Si l'on cherche un équivalent d'une somme, on passe par des développements limités.

Rappel

Une série converge si et seulement si la **suite** de ses sommes partielles converge.

Notons que cette constante A n'est autre que $\sum_{k=2}^{+\infty} \left[\ln(2 - e^{1/k}) - \ln\left(1 - \frac{1}{k}\right) \right]$.

Ainsi, $\ln(nu_n) \xrightarrow{n \rightarrow +\infty} A$, donc $nu_n \xrightarrow{n \rightarrow +\infty} e^A$.

Ceci signifie alors que $nu_n \underset{n \rightarrow +\infty}{\sim} e^A$ et donc $u_n \underset{n \rightarrow +\infty}{\sim} \frac{e^A}{n}$.

On retrouve bien le résultat cherché avec $K = e^A > 0$.

Comme la série de terme général $\frac{K}{n}$ est divergente, par critère de comparaison pour les séries à termes positifs, on en déduit que $\boxed{\text{la série de terme général } u_n \text{ est divergente.}}$

4.a. Nous avons déjà expliqué précédemment que (V_n) est décroissante, donc par croissance de l'exponentielle, $\boxed{(u_n) \text{ est décroissante.}}$

4.b. On a

$$S_{2n} - S_{2(n+1)} = \sum_{k=2}^{2n} (-1)^k u_k - \sum_{k=2}^{2n+2} (-1)^k u_k = u_{2n+1} - u_{2n+2} \geq 0$$

et donc $(S_{2n})_n$ est décroissante. De même, on a

$$S_{2n+1} - S_{2(n+1)+1} = \sum_{k=2}^{2n+1} (-1)^k u_k - \sum_{k=2}^{2n+3} (-1)^k u_k = u_{2n+3} - u_{2n+2} \leq 0$$

donc $(S_{2n+1})_n$ est croissante. Enfin, on a

$$S_{2n} - S_{2n+1} = u_{2n+1} \xrightarrow{n \rightarrow +\infty} 0.$$

Ceci achève bien de prouver que $\boxed{\text{les deux suites } (S_{2n})_n \text{ et } (S_{2n+1})_n \text{ sont adjacentes.}}$

4.c. Deux suites adjacentes sont toutes deux convergentes et convergent vers une même limite. Soit donc ℓ la limite commune aux deux suites (S_{2n}) et (S_{2n+1}) .

Ainsi, les deux suites extraites de (S_n) formées respectivement des termes d'ordre pair et des termes d'ordre impair convergent vers la même limite ℓ , donc (S_n) converge également vers ℓ .

Mais (S_n) est la suite des sommes partielles de la série de terme général $(-1)^n u_n$. Par définition, ceci signifie que $\boxed{\text{la série de terme général } (-1)^n u_n \text{ est convergente.}}$

Définition

Deux suites (u_n) et (v_n) sont adjacentes si :

- l'une est croissante
- l'autre est décroissante
- $\lim_{n \rightarrow +\infty} u_n - v_n = 0$

RÉVISIONS DE PROBABILITÉS

Exercices 626
Correction des exercices 631
EML 2018 651
Correction 653
ECRICOME 2010 659
Correction 662
EDHEC 2013 669
Correction 670

EXERCICES DU CHAPITRE 21

PROBABILITÉS DISCRÈTES

EXERCICE 21.1 On dispose de deux dés A et B . Le dé A comporte quatre faces rouges et deux faces jaunes. Le dé B comporte deux faces rouges et quatre faces jaunes. On lance une pièce de monnaie qui tombe sur pile avec probabilité $\frac{1}{3}$. Alors :

F

- si la pièce tombe sur «pile», on ne joue ensuite qu'avec le dé A
 - sinon, on ne joue ensuite qu'avec le dé B .
- 1) Déterminer la probabilité d'obtenir «rouge» au premier lancer du dé.
 - 2) On a obtenu «rouge» aux deux premiers lancers du dé. Quelle est la probabilité d'obtenir «rouge» au troisième ?
 - 3) On a obtenu «rouge» aux n premiers lancers du dé ($n \geq 1$), quelle est la probabilité p_n que la pièce soit tombée sur «pile» ?

EXERCICE 21.2 Un candidat passe chaque année 3 concours indépendants, et la probabilité de réussite à chacun de ces concours vaut $\frac{1}{3}$. Soit X le nombre d'années nécessaires à la réussite d'au moins un concours. Déterminer la loi de X , son espérance et sa variance.

PD

EXERCICE 21.3 Une urne contient au départ une boule verte et une boule rouge. On effectue des tirages successifs dans cette urne selon la procédure suivante : on tire une boule, si elle est rouge, on arrête les tirages, si elle est verte, on la remet dans l'urne en ajoutant une boule rouge. On note X le nombre de tirages effectués.

AD

- 1) Montrer que X admet une espérance et la déterminer.
- 2) Montrer que $\frac{1}{X}$ admet une espérance et la calculer.

EXERCICE 21.4 Soit $n \geq 3$ et soit $p \in]0, 1[$. On lance n fois une pièce qui tombe sur pile avec probabilité p . Pour tout $i \in \llbracket 2, n \rrbracket$, on note X_i la variable aléatoire qui vaut 1 si le $(i-1)^{\text{ème}}$ lancer donne pile et le $i^{\text{ème}}$ donne face, et 0 sinon.

AD

- 1) Déterminer la loi des X_i , leur espérance et leur variance.
- 2) Soient i et j deux entiers tels que $2 \leq i < j \leq n$. Les variables aléatoires X_i et X_j sont-elles indépendantes ? Déterminer $\text{Cov}(X_i, X_j)$.
- 3) On pose à présent $Y = \sum_{i=2}^n X_i$. Que représente Y ? Déterminer son espérance.
- 4) En utilisant la bilinéarité de la covariance, déterminer la variance de Y .

EXERCICE 21.5 On tire simultanément deux boules au hasard dans une urne contenant n boules numérotées de 1 à n . On appelle I la variable aléatoire égale au plus petit des deux numéros obtenus, et S la variable aléatoire égale au plus grand des deux numéros obtenus. Déterminer la loi du couple (I, S) , puis déterminer les lois de I et de S .

PD

EXERCICE 21.6 Une urne contient des boules rouges et des boules noires, la proportion de boules rouges étant notée p ($0 < p < 1$). On effectue une infinité de tirages avec remise dans cette urne, et on note N (respectivement R) le rang du tirage où, pour la première fois, on a obtenu une boule noire (resp. rouge).

AD

- 1) Donner la loi de N et la loi de R . Les variables aléatoires R et N sont-elles indépendantes ?
- 2) Déterminer la loi conjointe du couple (N, R) .
- 3) Calculer la covariance de N et R , puis le coefficient de corrélation linéaire.

EXERCICE 21.7 Soit $N \geq 2$. Soient X et Y deux variables aléatoires indépendantes suivant la même loi uniforme sur $\llbracket 1, N \rrbracket$. On pose $U = \max(X, Y)$ et $V = \min(X, Y)$.

AD

- 1) Déterminer la loi de U , puis celle de V . U et V sont-elles indépendantes ?
- 2) Justifier que U et V admettent une espérance et une variance, et les calculer.
- 3) Déterminer la variance de $U + V$, et en déduire $\text{Cov}(U, V)$.
- 4) Donner la loi de $X + Y$, puis celle de $U + V$.

EXERCICE 21.8 Soient X et Y deux variables aléatoires indépendantes, telles que X suit la loi $\mathcal{P}(\lambda)$ et Y suit la loi $\mathcal{P}(\mu)$. Déterminer la loi de X sachant $[X + Y = n]$. En déduire la valeur de $E(X|X + Y = n)$.

PD

EXERCICE 21.9 On lance une pièce qui tombe sur pile avec probabilité p jusqu'à obtenir pile pour la première fois. S'il a fallu n essais, on place des boules numérotées de 0 à n dans une urne, et on en tire une au hasard, et on effectue des tirages avec remise jusqu'à obtenir la boule 0. On note X le nombre de tirages nécessaires.

PD

- 1) Montrer que X admet une espérance et la déterminer.
- 2) (★) Même question, si les tirages dans l'urne ont lieu sans remise.

EXERCICE 21.10 Loi binomiale négative

AD

On considère une expérience ayant probabilité p de réussir et $1 - p$ d'échouer.

On répète l'expérience de manières indépendantes jusqu'à l'obtention de m succès, et on note X le nombre d'essais nécessaires. On note également X_i la variable aléatoire qui vaut 1 si l'expérience i a été un succès et 0 sinon.

- 1) Reconnaître la loi de X lorsque $m = 1$.
- 2) Déterminer la loi de X_i .
- 3) Écrire l'événement $[X = k]$ en fonction d'événements formés à partir des variables aléatoires $X_1 + \dots + X_{k-1}$ et X_k .
- 4) En déduire la loi de X dans le cas général.
- 5) (★) En écrivant X comme somme de m variables aléatoires suivant une loi usuelle, déterminer l'espérance de X .

EXERCICE 21.11 (QSP HEC 2016)

TD

Soit X une variable aléatoire discrète admettant un moment d'ordre 4 et vérifiant $\begin{cases} E(X) = \alpha \\ E(X^2) = E(X^4) = 1 \end{cases}$

- 1) Montrer que α est nécessairement compris entre -1 et 1 .
- 2) Déterminer la loi de X .

VARIABLES À DENSITÉ

EXERCICE 21.12 Soient X_1, \dots, X_n des variables aléatoires i.i.d. suivant la loi $\mathcal{E}(\lambda)$.

F

Déterminer la loi de $Y = \min(X_1, \dots, X_n)$, donner son espérance et sa variance.

EXERCICE 21.13 Soit X une variable aléatoire suivant la loi uniforme sur $]0, 1[$, et soit $Y = \frac{1}{X}$.

PD

Montrer que Y est une variable à densité et en déterminer une densité.

EXERCICE 21.14 Soit X une variable aléatoire qui suit la loi exponentielle de paramètre 1. On pose $Y = \ln(e^X - 1)$.

PD

- 1) Déterminer la fonction de répartition de Y .
- 2) Montrer que Y est une variable à densité, et en déterminer une densité.
- 3) Calculer $E(Y)$.

EXERCICE 21.15 Soit X une variable aléatoire admettant une densité f nulle sur $] - \infty, 0[$ et continue sur $[0, +\infty[$. On note S la fonction définie par $S(t) = P(X > t)$.

AD

- 1) À l'aide d'une intégration par parties, montrer que pour tout $A > 0$, $\int_0^A S(t) dt = AS(A) + \int_0^A t f(t) dt$.
- 2) Montrer que si X admet une espérance, alors : $\forall A > 0$, $AS(A) \leq \int_A^{+\infty} t f(t) dt$.
En déduire que l'intégrale $\int_0^{+\infty} S(t) dt$ est convergente et qu'elle est égale à $E(X)$.
- 3) (★) Inversement, montrer que si $\int_0^{+\infty} S(t) dt$ est convergente, alors X admet une espérance et que

$$E(X) = \int_0^{+\infty} S(t) dt.$$

EXERCICE 21.16 (D'après EDHEC 2008)

AD

Soient U et V deux variables aléatoires indépendantes sur le même espace probabilisé (Ω, \mathcal{A}, P) , suivant toutes les deux la loi uniforme sur $[0, 1]$.

- 1) Montrer que U^2 est une variable à densité, et en déterminer une densité.
- 2) Déterminer une densité de $-V$.

- 3) En déduire qu'une densité de $T = U^2 - V$ est la fonction $h : x \mapsto \begin{cases} \sqrt{x+1} & \text{si } -1 \leq x \leq 0 \\ 1 - \sqrt{x} & \text{si } 0 < x \leq 1 \\ 0 & \text{sinon} \end{cases}$

- 4) (★) Pour $\omega \in \Omega$, on définit un polynôme P_ω par $P_\omega(X) = X^2 - 2U(\omega)X + V(\omega)$. Déterminer la probabilité que P_ω soit de signe constant sur \mathbf{R} .

EXERCICE 21.17 Loi de Weibull

AD

Soient a et b deux réels strictement positifs. On définit une fonction F par $F(x) = \begin{cases} 0 & \text{si } x < 0 \\ 1 - e^{-(\frac{x}{b})^a} & \text{si } x \geq 0 \end{cases}$

- 1) Montrer que F est la fonction de répartition d'une variable aléatoire à densité X , et donner une densité f de X .
On dit alors que X suit la loi de Weibull de paramètres a et b , et on note $X \hookrightarrow \mathcal{W}(a, b)$.
- 2) Si $a = 1$, quelle est la loi de X ?
- 3) Déterminer la loi de X^a .
- 4) Soit $k \in \mathbf{N}^*$. Montrer que X admet un moment d'ordre k et que $E(X^k) = b^k \Gamma\left(1 + \frac{k}{a}\right)$.

On pourra utiliser le changement de variable $t = \left(\frac{x}{b}\right)^a$.

En déduire l'espérance et la variance de X .

- 5) Soient X_1, \dots, X_n des variables aléatoires i.i.d. suivant la loi $\mathcal{W}(a, b)$, et soit $Y_n = \min(X_1, \dots, X_n)$.
Montrer que Y_n suit encore une loi de Weibull dont on précisera les paramètres.

EXERCICE 21.18 (QSP HEC 2010)

D

Soit X_1, \dots, X_n des variables aléatoires définies sur un même espace probabilisé (Ω, \mathcal{A}, P) , indépendantes et suivant la loi uniforme sur $[0, a]$ ($a > 0$).

- 1) Déterminer une densité de $-\max(X_1, \dots, X_k)$, pour tout $k \in \llbracket 1, n \rrbracket$.
- 2) En déduire $P([X_n \geq X_1] \cap [X_n \geq X_2] \cap \dots \cap [X_n \geq X_{n-1}])$.

EXERCICE 21.19 (Extrait de oral ESCP 2016)

D

Dans cet exercice, les variables aléatoires sont définies sur un espace probabilisé (Ω, \mathcal{A}, P) . Pour $\alpha > 0$ donné, on dit qu'une variable aléatoire X est α -sous-gaussienne si pour tout t réel,

$$E(e^{tX}) \leq e^{\alpha^2 t^2 / 2}.$$

- 1) Soit X de loi normale centrée réduite. Montrer que X est 1-sous-gaussienne.
- 2) Soient X_1, \dots, X_n des variables aléatoires mutuellement indépendantes et α -sous-gaussiennes. Soient μ_1, \dots, μ_n des réels tels que $\sum_{i=1}^n \mu_i^2 = 1$. Montrer que $\sum_{i=1}^n \mu_i X_i$ est α -sous-gaussienne.
- 3) Soit X α -sous-gaussienne et $\lambda > 0$.
 - a) Montrer que pour tout $t > 0$, on a $P(X \geq \lambda) \leq \exp\left(\frac{\alpha^2 t^2}{2} - t\lambda\right)$.
 - b) En déduire que $P(|X| \geq \lambda) \leq 2 \exp\left(-\frac{\lambda^2}{2\alpha^2}\right)$.

EXERCICE 21.20 Quelques convolutions

AD

Déterminer une densité de $X + Y$ dans les cas suivants, où X et Y sont toujours supposées indépendantes.

1. $X \hookrightarrow \mathcal{U}([0, 1])$, $Y \hookrightarrow \mathcal{U}([0, 2])$.
2. $X \hookrightarrow \mathcal{E}(1)$ et $Y \hookrightarrow \mathcal{E}(2)$.
3. $X \hookrightarrow \mathcal{E}(1)$ et $Y \hookrightarrow \mathcal{U}([0, 1])$.
4. (★) X et Y ont pour densité $f : x \mapsto \frac{1}{2} e^{-|x|}$ (on pourra commencer par vérifier que f est bien une densité).

CONVERGENCES, ESTIMATION

EXERCICE 21.21 Soit (X_n) une suite de variables aléatoires admettant une espérance telle que $\lim_{n \rightarrow +\infty} E(|X_n|) = 0$.

F

Montrer que $X_n \xrightarrow{P} 0$.

EXERCICE 21.22 Soit θ un réel strictement positif, et soit, pour tout entier n , X_n une variable aléatoire suivant la loi de Poisson de paramètre $n\lambda$.

PD

Montrer que $Y_n = \frac{X_n - n\theta}{n}$ converge en probabilité vers 0 lorsque n tend vers $+\infty$.

EXERCICE 21.23 On considère une suite $(X_n)_{n \geq 1}$ de variables aléatoires indépendantes suivant la loi de Poisson de paramètre 1, et on pose $S_n = X_1 + \dots + X_n$.

AD

- 1) Exprimer $P(S_n \leq n)$ sous forme d'une somme.

2) En utilisant le théorème central limite, montrer que $\sum_{k=0}^n \frac{n^k}{k!} \underset{n \rightarrow +\infty}{\sim} \frac{e^n}{2}$.

EXERCICE 21.24 Loi du χ^2 et intervalle de confiance pour la variance d'une loi normale

AD

1) Soient U_1, \dots, U_n des variables aléatoires mutuellement indépendantes suivant la loi normale centrée réduite. On pose alors $W_n = \frac{1}{2} \sum_{i=1}^n U_i^2$.

a) Montrer que $\frac{1}{2}U_1^2$ est une variable aléatoire à densité, et en donner une densité. En déduire que $\Gamma(1/2) = \sqrt{\pi}$, puis reconnaître la loi usuelle suivie par $\frac{1}{2}U_1^2$.

b) Quelle est la loi de W_n ? Donner son espérance et sa variance.

c) Montrer que pour tout $\alpha \in]0, 1[$, il existe un unique réel k_α tel que $P(W_n \leq k_\alpha) = \alpha$.

2) On considère à présent X_1, \dots, X_n des variables aléatoires i.i.d. suivant la loi normale centrée de variance σ^2 inconnue, et on pose

$$S_n = \sum_{i=1}^n X_i^2 \text{ et } T_n = \frac{1}{n} \sum_{i=1}^n X_i^2.$$

a) Montrer que T_n est un estimateur sans biais de σ^2 .

b) Justifier que $\frac{n}{2\sigma^2}T_n$ suit la même loi que W_n , et en déduire que T_n est un estimateur convergent du paramètre σ^2 .

c) Montrer que $\left]0, \frac{S_n}{2k_\alpha}\right]$ est un intervalle de confiance de σ^2 au niveau de risque α .

EXERCICE 21.25 Soit $(X_n)_{n \in \mathbf{N}^*}$ une suite de variables aléatoires indépendantes suivant la loi de Bernoulli de paramètre $\frac{1}{2}$. On pose alors $\bar{X}_n = \frac{X_1 + \dots + X_n}{n}$. Montrer que $\lim_{n \rightarrow +\infty} P\left(\left|\bar{X}_n - \frac{1}{2}\right| > \frac{1}{\sqrt{n}}\right)$ existe et déterminer sa valeur.

PD

EXERCICE 21.26 Soit (X_n) une suite de variables aléatoires indépendantes suivant la loi de Bernoulli de paramètre p inconnu.

PD

Pour tout $n \in \mathbf{N}^*$, on pose $S_n = \sum_{i=1}^n X_i$ et $U_n = \frac{S_n}{n} \left(1 - \frac{S_n}{n}\right)$.

1) Calculer $E(S_n)$, $V(S_n)$ et $E(S_n^2)$. En déduire $E(U_n)$.

2) Déterminer le biais de U_n en tant qu'estimateur de $p(1-p)$. En déduire un estimateur sans biais de $p(1-p)$.

EXERCICE 21.27 Comparaison de risques quadratiques

F

Soit $(X_n)_{n \geq 1}$ une suite de variables aléatoires indépendantes suivant la loi exponentielle $\mathcal{E}(\lambda)$. On pose

$$\bar{X}_n = \frac{1}{n} \sum_{i=1}^n X_i \text{ et } T_n = \frac{1}{n+1} \sum_{i=1}^n X_i.$$

Comparer les risques quadratiques de \bar{X}_n et T_n en tant qu'estimateurs de $\frac{1}{\lambda}$.

EXERCICE 21.28 Soit $(X_n)_{n \in \mathbf{N}^*}$ une suite de variables indépendantes, de même loi et possédant un moment d'ordre 4. On note m l'espérance commune des X_i , σ^2 leur variance, et pour $n \in \mathbf{N}^*$, on pose

D

$$\bar{X}_n = \frac{1}{n} \sum_{i=1}^n X_i \text{ et } \tilde{S}_n = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X}_n)^2.$$

1) Montrer que $\tilde{S}_n = \frac{1}{n} \sum_{i=1}^n X_i^2 - \bar{X}_n^2$.

2) Montrer que $(\tilde{S}_n)_{n \geq 1}$ converge en probabilités vers σ^2 . On admettra que si $U_n \xrightarrow{P} U$ et $V_n \xrightarrow{P} V$, alors $U_n - V_n \xrightarrow{P} U - V$.

3) Étudier la convergence en loi de la suite $\frac{\sqrt{n}(\bar{X}_n - m)}{\sqrt{\tilde{S}_n}}$.

EXERCICE 21.29 (QSP ESCP 2010)

On lance une pièce qui donne face avec probabilité $p \in]0, 1[$. On note P_n (resp. F_n) la variable aléatoire égale au nombre de piles (resp. de faces) obtenus au cours des n premiers lancers.

Montrer que $\left(\frac{F_n - P_n}{n}\right)_n$ converge en probabilité vers la variable certaine égale à $2p - 1$.

EXERCICE 21.30 (QSP ESCP 2016)

On considère une variable aléatoire X définie sur un espace probabilisé (Ω, \mathcal{A}, P) qui suit une loi de Poisson de paramètre $\lambda > 0$.

- 1) Soit t un réel strictement positif. Montrer que la variable aléatoire e^{tX} admet une espérance et la calculer.
- 2) Soit $n \in \mathbf{N}$. Prouver que $P(X \geq n) \leq e^{-tn} E(e^{tX})$.
- 3) En déduire que : $\sum_{k=n}^{+\infty} \frac{\lambda^k}{k!} \leq e^n \left(\frac{\lambda}{n}\right)^n$.

D

AD

CORRECTION DES EXERCICES DU CHAPITRE 21

SOLUTION DE L'EXERCICE 21.1

Dans toute la suite, on note A (respectivement B) l'événement «on joue avec le dé A (resp. B)».

On note également R_i (resp. J_i) l'événement «le $i^{\text{ème}}$ lancer de dé donne une face rouge (resp. jaune).

1. Puisque $\{A, B\}$ est un système complet d'événements, on a, par la formule des probabilités totales,

$$P(R_1) = P(A)P_A(R_1) + P(B)P_B(R_1) = \frac{1}{3} \frac{4}{6} + \frac{2}{3} \frac{2}{6} = \frac{4}{9}.$$

2. La probabilité cherchée est donc $P_{R_1 \cap R_2}(R_3) = \frac{P(R_1 \cap R_2 \cap R_3)}{P(R_1 \cap R_2)}$. Mais, toujours par la formule des probabilités totales, on a

$$P(R_1 \cap R_2 \cap R_3) = P(A)P_A(R_1 \cap R_2 \cap R_3) + P(B)P_B(R_1 \cap R_2 \cap R_3) = \frac{1}{3} \left(\frac{4}{6}\right)^3 + \frac{2}{3} \left(\frac{2}{6}\right)^3 = \frac{10}{81}.$$

Et de même, on a

$$P(R_1 \cap R_2) = P(A)P_A(R_1 \cap R_2) + P(B)P_B(R_1 \cap R_2) = \frac{1}{3} \left(\frac{4}{6}\right)^2 + \frac{2}{3} \left(\frac{2}{6}\right)^2 = \frac{6}{27}.$$

Et donc $P_{R_1 \cap R_2}(R_3) = \frac{5}{9}$.

3. La probabilité cherchée est donc $P_{R_1 \cap R_2 \cap \dots \cap R_n}(A)$. Par la formule de Bayes, on a

$$\begin{aligned} P_{R_1 \cap \dots \cap R_n}(A) &= \frac{P_A(R_1 \cap \dots \cap R_n)P(A)}{P(R_1 \cap \dots \cap R_n)} \\ &= \frac{P_A(R_1) \cdots P_A(R_n)P(A)}{P(R_1 \cap \dots \cap R_n)} \\ &= \frac{\frac{1}{3} \left(\frac{2}{3}\right)^n}{P(R_1 \cap \dots \cap R_n)} \\ &= \frac{\frac{1}{3} \left(\frac{2}{3}\right)^n}{P(A)P_A(R_1 \cap \dots \cap R_n) + P(B)P_B(R_1 \cap \dots \cap R_n)} \\ &= \frac{\frac{1}{3} \left(\frac{2}{3}\right)^n}{\frac{1}{3} \left(\frac{2}{3}\right)^n + \frac{2}{3} \left(\frac{1}{3}\right)^n} \\ &= \frac{2^n}{2^n + 2}. \end{aligned}$$

Sachant qu'on joue avec le dé A , les lancers sont indépendants.

Formule des probabilités totales avec le s.c.e $\{A, B\}$.

SOLUTION DE L'EXERCICE 21.2

Une année donnée, le nombre Y de concours réussis par notre candidat suit une loi binomiale $\mathcal{B}\left(3, \frac{1}{3}\right)$.

En particulier, la probabilité de réussir au moins un concours vaut

$$p = P(Y > 0) = 1 - P(Y = 0) = 1 - \binom{3}{0} \left(\frac{1}{3}\right)^0 \left(\frac{2}{3}\right)^3 = 1 - \frac{8}{27} = \frac{19}{27}.$$

Et puisque les résultats d'une année à l'autre sont supposés indépendants, le nombre X d'années où le candidat doit présenter les concours avant de réussir l'un d'entre eux suit une loi géométrique de paramètre p .

En particulier, on a $E(X) = \frac{1}{p} = \frac{27}{19}$ et $V(X) = \frac{1-p}{p^2} = \frac{8}{27} \frac{27^2}{19^2} = \frac{216}{361}$.

SOLUTION DE L'EXERCICE 21.3

1. Il est clair que $X(\Omega) = \mathbf{N}^*$.

Notons B_k l'événement «la $k^{\text{ème}}$ boule tirée est rouge».

Alors pour tout $n \in \mathbf{N}^*$, on a $[X = n] = \overline{B_1} \cap \overline{B_2} \cap \dots \cap \overline{B_{n-1}} \cap B_n$.

Et donc par la formule des probabilités composées,

$$P(X = n) = P(\overline{B_1}) P_{\overline{B_1}}(\overline{B_2}) \cdots P_{\overline{B_1} \cap \dots \cap \overline{B_{n-1}}}(B_n) = \frac{1}{2} \frac{1}{3} \cdots \frac{1}{n} \frac{n}{n+1} = \frac{n}{(n+1)!}.$$

On a alors, pour $n \geq 1$,

$$nP(X = n) = \frac{n^2}{(n+1)!} = \frac{n(n+1) - n}{(n+1)!} = \frac{1}{(n-1)!} - \frac{n}{(n+1)!} = \frac{1}{(n-1)!} - \frac{n+1-1}{(n+1)!} = \frac{1}{(n-1)!} - \frac{1}{n!} + \frac{1}{(n+1)!}.$$

Et donc, sous réserve de convergence,

$$\begin{aligned} E(X) &= \sum_{n=1}^{+\infty} \frac{1}{(n-1)!} - \sum_{n=1}^{+\infty} \frac{1}{n!} + \sum_{n=1}^{+\infty} \frac{1}{(n+1)!} \\ &= \sum_{i=0}^{+\infty} \frac{1}{i!} - \sum_{n=1}^{+\infty} \frac{1}{n!} + \sum_{k=2}^{+\infty} \frac{1}{k!} \\ &= \sum_{i=1}^{+\infty} \frac{1}{i!} - 1 = e - 1. \end{aligned}$$

Nous avons reconnu des séries exponentielles, donc convergentes, et donc $E(X)$ existe.

Donc X admet bien une espérance, et celle-ci vaut $e - 1$.

2. Par le théorème de transfert, on a, sous réserve de convergence,

$$\begin{aligned} E\left(\frac{1}{X}\right) &= \sum_{n=1}^{+\infty} \frac{1}{n} P(X = n) \\ &= \sum_{n=1}^{+\infty} \frac{1}{(n+1)!} \\ &= \sum_{i=2}^{+\infty} \frac{1}{i!} \\ &= \sum_{i=0}^{+\infty} \frac{1}{i!} - \frac{1}{0!} - \frac{1}{1!} \\ &= e - 2. \end{aligned}$$

On a reconnu une série exponentielle, donc convergente. Donc l'espérance existe.

SOLUTION DE L'EXERCICE 21.4

1. Notons F_i l'événement «le $i^{\text{ème}}$ lancer donne un face». Alors on a $P(X_i = 1) = P(F_i \cap \overline{F_{i-1}})$. Et donc par indépendance des lancers, il vient

$$P(X_i = 1) = P(F_i)P(\overline{F_{i-1}}) = p(1-p).$$

Puisque X_i ne peut prendre que les valeurs 0 et 1, elle suit donc la loi de Bernoulli de paramètre $p(1-p)$.

En particulier, $E(X_i) = p(1-p)$ et $V(X_i) = p(1-p)(1-p(1-p))$.

2. Puisque X_i et X_j sont des variables de Bernoulli, elles sont indépendantes si et seulement si $[X_i = 1]$ et $[X_j = 1]$ sont indépendants.

• Si $j = i + 1$, alors, $P([X_i = 1] \cap [X_{i+1} = 1]) = P(\overline{F_{i-1}} \cap F_i \cap \overline{F_i} \cap F_{i+1}) = 0$.

Et donc X_i et X_j ne sont pas indépendantes.

De plus, on a $X_i X_j = 1$, puisque X_i et X_j ne peuvent simultanément valoir 1, de sorte que $E(X_i X_j) = 0$.

Et donc par la formule de Huygens,

$$\text{Cov}(X_i, X_j) = E(X_i X_j) - E(X_i)E(X_j) = -E(X_i)^2 = -p^2(1-p)^2.$$

• Si $j > i + 1$, alors

$$P([X_i = 1] \cap [X_j = 1]) = P(\overline{F_{i-1}} \cap F_i \cap \overline{F_{j-1}} \cap F_j) = P(\overline{F_{i-1}} \cap F_i) P(\overline{F_{j-1}} \cap F_j) = P(X_i = 1) P(X_j = 1)$$

par indépendance des lancers. Et donc X_i et X_j sont indépendantes.

Dans ce cas, on a donc $\text{Cov}(X_i, X_j) = 0$.

3. Y représente le nombre de fois où, au cours des n lancers, on a eu un pile suivi d'un face. Par linéarité de l'espérance¹, il vient donc

$$E(Y) = \sum_{i=2}^n E(X_i) = (n-1)p(1-p).$$

¹ Rappelons que l'espérance d'une somme est toujours la somme des espérances, que les variables aléatoires soient indépendantes ou non.

4. On a

$$\begin{aligned} V(Y) &= V\left(\sum_{i=2}^n X_i\right) = \text{Cov}\left(\sum_{i=2}^n X_i, \sum_{j=2}^n X_j\right) \\ &= \sum_{i=2}^n \text{Cov}\left(X_i, \sum_{j=2}^n X_j\right) \\ &= \sum_{i=2}^n \sum_{j=2}^n \text{Cov}(X_i, X_j). \end{aligned}$$

Linéarité à gauche de la covariance.

Linéarité à droite de la covariance.

Mais pour $|i-j| > 1$, nous avons prouvé à la question 2 que $\text{Cov}(X_i, X_j) = 0$, de sorte que

$$\begin{aligned} V(Y) &= \sum_{i=2}^n \text{Cov}(X_i, X_i) + \sum_{i=2}^{n-1} \text{Cov}(X_i, X_{i+1}) + \sum_{i=3}^n \text{Cov}(X_i, X_{i-1}) \\ &= (n-1)p(1-p)(1-p(1-p)) - 2(n-2)p^2(1-p)^2 \\ &= p(1-p)((n-1)(1-p(1-p)) - 2(n-2)p(1-p)). \end{aligned}$$

On ne garde que les termes avec $i=j$, $j=i+1$ ou $j=i-1$.

SOLUTION DE L'EXERCICE 21.5

Il est clair que $P([I=k] \cap [S=\ell]) = 0$ si $k \geq \ell$.
En revanche, si $k < \ell$, alors

$$P([I=k] \cap [S=\ell]) = \frac{1}{\binom{n}{2}}.$$

En effet, il y a $\binom{n}{2}$ manières distinctes de choisir deux boules dans l'urne, et une seule qui réalise l'événement $[I=k] \cap [S=\ell]$. Ainsi, la loi conjointe du couple (I, S) est donnée par

$$\forall (k, \ell) \in \llbracket 1, n \rrbracket^2, P([I=k] \cap [S=\ell]) = \begin{cases} 0 & \text{si } k \geq \ell \\ \frac{2}{n(n-1)} & \text{sinon} \end{cases}$$

Pour retrouver les lois marginales du couple, appliquons la formule des probabilités totales au système complet d'événements $\{[Y=\ell], \ell \in \llbracket 1, n \rrbracket\}$. Alors

$$P(I=k) = \sum_{\ell=1}^n P([I=k] \cap [S=\ell]) = \sum_{\ell=k+1}^n \frac{2}{n(n-1)} = \frac{2(n-k)}{n(n-1)}.$$

De même,

$$P(S=\ell) = \sum_{k=1}^n P([I=k] \cap [S=\ell]) = \sum_{k=1}^{\ell-1} \frac{2}{n(n-1)} = \frac{2(\ell-1)}{n(n-1)}.$$

SOLUTION DE L'EXERCICE 21.6

1. Par indépendance² des tirages, N suit une loi géométrique de paramètre $1-p$ et R suit une loi géométrique de paramètre p .

Il est clair que ces deux variables ne sont pas indépendantes, par exemple car $P([N=1] \cap [R=1]) = 0$ (puisque on ne peut pas obtenir au premier tirage à la fois une boule rouge et une boule noire), alors que $P(N=1) = p \neq 0$ et $P(R=1) = 1-p \neq 0$.

2. Soient $(k, \ell) \in (\mathbf{N}^*)^2$.

• Puisque le premier tirage donne soit une boule noire soit une boule rouge, l'un des deux événements $[N=1]$ et $[R=1]$ est toujours réalisé. On a donc $P([N=k] \cap [R=\ell]) = 0$ si $k > 1$ et $\ell > 1$.

• Si $k = \ell = 1$, nous avons déjà mentionné que $P([N=1] \cap [R=1]) = 0$.

² Indépendance qui est garantie par le fait que les tirages aient lieu avec remise.

- Supposons à présent que $k = 1$ et $\ell > 1$.
Notons alors R_i (resp. N_i) l'événement «la $i^{\text{ème}}$ boule tirée est rouge (resp. noire)».
Alors on a $[N = 1] \cap [R = \ell] = N_1 \cap N_2 \cap \dots \cap N_{\ell-1} \cap R_\ell$.
Et par indépendance des tirages, il vient alors

$$P([N = 1] \cap [R = \ell]) = (1-p)^{\ell-1}p.$$

De même, on prouverait que pour $\ell = 1$ et $k > 1$, on a

$$P([N = k] \cap [R = 1]) = p^{k-1}(1-p).$$

3. Par la formule de Huygens, on a $\text{Cov}(N, R) = E(NR) - E(N)E(R)$.
De plus, par le théorème de transfert³, on a, sous réserve de convergence,

$$E(NR) = \sum_{\substack{1 \leq k \leq n \\ 1 \leq \ell \leq n}} k\ell P([N = k] \cap [R = \ell]).$$

Or, tous les termes avec $k > 1$ et $\ell > 1$ sont nuls, donc il vient

$$\begin{aligned} E(NR) &= \sum_{k=2}^{+\infty} kP([N = k] \cap [R = 1]) + \sum_{\ell=2}^{+\infty} \ell P([N = 1] \cap [R = \ell]) \\ &= \sum_{k=2}^{+\infty} kP(N = k) + \sum_{\ell=2}^{+\infty} \ell P(R = \ell) \\ &= (E(N) - (1-p)) + (E(R) - p) \\ &= \frac{1}{1-p} + \frac{1}{p} - 1 \\ &= \frac{1}{p(1-p)} - 1. \end{aligned}$$

Et donc il vient

$$\text{Cov}(N, R) = E(NR) - E(N)E(R) = \frac{1}{p(1-p)} - 1 - \frac{1}{p} \frac{1}{1-p} = -1.$$

Notons que cette covariance étant non nulle, on retrouve bien le fait que R et N ne sont pas indépendantes.

4. Par définition, le coefficient de corrélation linéaire est donné par

$$\rho_{N,R} = \frac{\text{Cov}(N, R)}{\sqrt{V(R)}\sqrt{V(N)}} = -\sqrt{p(1-p)}.$$

SOLUTION DE L'EXERCICE 21.7

1. Notons que $U(\Omega) = V(\Omega) = \llbracket 1, N \rrbracket$.
Pour $k \in \llbracket 1, N \rrbracket$, on a

$$P(U \leq k) = P([X \leq k] \cap [Y \leq k]) = P(X \leq k)P(Y \leq k) = \frac{k^2}{N^2}.$$

Et donc, U étant à valeurs entières, il vient

$$P(U = k) = P(U \leq k) - P(U \leq k-1) = \frac{k^2}{N^2} - \frac{(k-1)^2}{N^2} = \frac{2k-1}{N^2}.$$

De même, on a $P(V \geq k) = P(X \geq k)P(Y \geq k) = \frac{(N-k+1)^2}{N^2}$ et donc

$$P(V = k) = P(V \geq k) - P(V \geq k+1) = \frac{(N-k+1)^2}{N^2} - \frac{(N-k)^2}{N^2} = \frac{2N-2k+1}{N^2}.$$

Les variables U et V ne sont pas indépendantes, par exemple car on a $P([U = 1] \cap [V = N]) = 0$ alors que $P(U = 1)P(V = N) \neq 0$.

Remarque

On constate que cette probabilité est la même que celle de l'événement $[R = \ell]$.
C'est logique si l'on remarque que s'il faut ℓ tirage pour obtenir une boule rouge, le premier tirage a nécessairement fourni une boule noire. Et donc

$$[N = 1] \cap [R = \ell] = [R = \ell].$$

³ Version à plusieurs variables.

Voir la remarque ci-dessus.

2. U et V sont des variables à support fini, donc admettent bien une espérance et une variance.

De plus, on a

$$E(U) = \sum_{k=1}^N kP(U=k) = \sum_{k=1}^N k \frac{2k-1}{N^2} = \frac{2}{N^2} \sum_{k=1}^N k^2 - \frac{1}{N^2} \sum_{k=1}^N k = \frac{(N+1)(2N+1)}{3} - \frac{N+1}{2} = \frac{(N+1)(4N-1)}{6N}.$$

Plutôt que de calculer $E(V)$, remarquons que $U+V = X+Y$ et donc $E(U) + E(V) = E(X) + E(Y) = N+1$.

$$\text{Et donc } E(V) = N+1 - E(U) = \frac{(N+1)(2N+1)}{6N}.$$

Par le théorème de transfert, on a

$$\begin{aligned} E(U^2) &= \sum_{k=1}^N k^2 \frac{2k-1}{N} \\ &= \frac{2}{N} \sum_{k=1}^N k^3 - \frac{1}{N} \sum_{k=1}^N k^2 \\ &= \frac{2}{N} \frac{N^2(N+1)^2}{4} - \frac{(N+1)(2N+1)}{6} \\ &= (N+1) \frac{3N(N+1) - (2N+1)}{6} = (N+1) \frac{3N^2 + N - 1}{6}. \end{aligned}$$

Et donc, par la formule de Huygens, il vient

$$\begin{aligned} V(U) &= E(U^2) - E(U)^2 = \frac{N+1}{N} \frac{3N^2 + N - 1}{6} - \frac{(N+1)^2(4N-1)^2}{36N^2} \\ &= (N+1) \frac{6N(3N^2 + N - 1) - (N+1)(16N^2 - 16N + 1)}{36N^2} \\ &= (N+1) \frac{2N^3 - 2N^2 + N - 1}{36N^2} = \frac{(N+1)(N-1)(2N^2+1)}{36N^2}. \end{aligned}$$

De même, on a

$$\begin{aligned} E(V^2) &= \sum_{k=1}^N k^2 P(V=k) = \sum_{k=1}^N k^2 \frac{2N-2k+1}{N^2} \\ &= \frac{2N+1}{N^2} \frac{N(N+1)(2N+1)}{6} - \frac{2}{N^2} \frac{N^2(N+1)^2}{4} \\ &= \frac{(N+1)(2N+1)^2}{6N} - \frac{(N+1)^2}{2} = \frac{(N+1)(2N+1)^2}{6N} - \frac{3N(N+1)^2}{6N} \\ &= (N+1) \frac{4N^2 + 4N + 1 - 3N^2 - 3N}{6N} = \frac{(N+1)(N^2 + N + 1)}{6N}. \end{aligned}$$

On en déduit alors que

$$\begin{aligned} V(V) &= E(V^2) - E(V)^2 = \frac{(N+1)(N^2 + N + 1)}{6N} - \frac{(N+1)^2(2N+1)^2}{36N^2} \\ &= \frac{6N(N+1)(N^2 + N + 1) - (N+1)^2(2N+1)^2}{36N^2} \\ &= (N+1) \frac{6N^3 + 6N^2 + 6N - 4N^3 - 8N^2 - 5N - 1}{36N^2} \\ &= (N+1) \frac{2N^3 - 2N^2 + N - 1}{36N^2} = \frac{(N+1)(N-1)(2N^2+1)}{36N^2}. \end{aligned}$$

3. Puisque $U+V = X+Y$, et que X et Y sont indépendantes, on a

$$V(U+V) = V(X) + V(Y) = \frac{N^2-1}{6}.$$

De plus, nous savons que

$$V(U+V) = V(U) + V(V) + 2\text{Cov}(U, V) \Leftrightarrow \text{Cov}(U, V) = \frac{1}{2}(V(U+V) - V(U) - V(V)).$$

Après calcul, on obtient alors

$$\text{Cov}(U, V) = \frac{(N^2 - 1)^2}{36N^2}.$$

4. Puisque X et Y sont indépendantes, par la formule du produit de convolution, on a, pour tout $k \in \llbracket 2, 2N \rrbracket$,

$$P(X + Y = k) = \sum_{i=1}^N P(X = i)P(Y = k - i).$$

Mais on a $P(Y = k - i) \neq 0 \Leftrightarrow 1 \leq k - i \leq N \Leftrightarrow k - N \leq i \leq k - 1$.

Si $k \in \llbracket 1, N \rrbracket$, alors

$$P(X + Y = k) = \sum_{i=1}^{k-1} P(X = i)P(Y = k - i) = \frac{1}{N^2} \sum_{i=1}^{k-1} 1 = \frac{k-1}{N^2}.$$

Et si $k \in \llbracket N + 1, 2N \rrbracket$, alors

$$P(X + Y = k) = \sum_{i=k-N}^N P(X = i)P(Y = k - i) = \sum_{i=k-N}^N \frac{1}{N^2} = \frac{2N - k + 1}{N^2}.$$

SOLUTION DE L'EXERCICE 21.8

Soit $n \in \mathbf{N}$ et $k \in \mathbf{N}$. Alors $[X + Y = n] \cap [X = k] = [Y = n - k] \cap [X = k]$.

Donc déjà, si $k > n$, $P_{[X+Y=n]}(X = k) = 0$ car Y est à valeurs positives.

De plus, par indépendance de X et Y , nous savons que $X + Y$ suit la loi de Poisson $\mathcal{P}(\lambda + \mu)$.

Donc pour $k \leq n$, on a

$$\begin{aligned} P_{[X+Y=n]}(X = k) &= \frac{P([X = k] \cap [Y = n - k])}{P(X + Y = n)} \\ &= \frac{P(X = k)P(Y = n - k)}{P(X + Y = n)} \\ &= e^{-\mu} \frac{\mu^{n-k}}{(n-k)!} e^{-\lambda} \frac{\lambda^k}{k!} \frac{n!}{e^{-\lambda-\mu}(\lambda + \mu)^n} \\ &= \frac{n!}{(n-k)!k!} \left(\frac{\mu}{\lambda + \mu}\right)^{n-k} \left(\frac{\lambda}{\lambda + \mu}\right)^k \\ &= \binom{n}{k} \left(\frac{\mu}{\lambda + \mu}\right)^{n-k} \left(\frac{\lambda}{\lambda + \mu}\right)^k \\ &= \binom{n}{k} \left(\frac{\lambda}{\lambda + \mu}\right)^k \left(1 - \frac{\lambda}{\lambda + \mu}\right)^{n-k}. \end{aligned}$$

On reconnaît là une loi binomiale $\mathcal{B}\left(n, \frac{\lambda}{\lambda + \mu}\right)$.

Et donc l'espérance conditionnelle $E(X|X + Y = n)$ est l'espérance de cette loi binomiale :

$$E(X|X + Y = n) = n \frac{\lambda}{\lambda + \mu}.$$

SOLUTION DE L'EXERCICE 21.9

Dans la suite, on note N la variable aléatoire égale au nombre de lancers de la pièce qui ont été nécessaires pour obtenir pile pour la première fois. Alors $N \hookrightarrow \mathcal{G}(p)$.

1. Soit $n \in \mathbf{N}^*$. Sachant que $[N = n]$ est réalisé, l'urne contient $n + 1$ boules, et donc à chaque tirage, la probabilité d'obtenir la boule 0 est $\frac{1}{n+1}$.

Les tirages étant indépendants⁴, X suit une loi géométrique de paramètre $\frac{1}{n+1}$.

En particulier, $E(X|N = n) = n + 1$.

Par la formule de l'espérance totale appliquée au système complet d'événements $\{[N = n], n \in \mathbf{N}^*\}$, sous réserve de convergence, X admet une espérance donnée par la somme de la série suivante :

$$E(X) = \sum_{n=1}^{+\infty} P(N = n)E(X|N = n)$$

Rappel

Rappelons que le produit de convolution discret n'est rien d'autre que la formule des probabilités totales appliquée avec le système complet d'événements $\{[X = i], i \in \llbracket 1, N \rrbracket\}$.

Hypothèses

L'hypothèse d'indépendance est très importante, si X et Y ne sont pas indépendantes, on ne sait pas donner la loi de $X + Y$.

X et Y sont indépendantes.

⁴ Car avec remise.

$$\begin{aligned}
&= \sum_{n=1}^{+\infty} (n+1)p(1-p)^{n-1} \\
&= p \sum_{n=1}^{+\infty} n(1-p)^{n-1} + p \sum_{n=1}^{+\infty} (1-p)^{n-1} \\
&= p \frac{1}{p^2} + \frac{p}{1-(1-p)} \\
&= \frac{1}{p} + 1.
\end{aligned}$$

Les deux séries convergent car $0 < 1-p < 1$.

2. Si les tirages ont lieu sans remise, la loi de X sachant $[N = n]$ n'est plus une loi géométrique, car les différents tirages ne sont plus indépendants.

Notons alors A_k l'événement «la boule numéro 0 est tirée lors du $k^{\text{ème}}$ tirage», de sorte que

$$[X = k] = \overline{A_1} \cap \overline{A_2} \cap \dots \cap \overline{A_{k-1}} \cap A_k.$$

Remarquons tout de suite que si $[N = n]$, alors il faut au maximum $n+1$ tirages pour vider l'urne, de sorte que X ne peut prendre que des valeurs entre 1 et $n+1$.

Par la formule des probabilités composées, on a alors

$$\begin{aligned}
P_{[N=n]}(X = k) &= P_{[N=n]}(\overline{A_1})P_{[N=n] \cap \overline{A_1}}(\overline{A_2}) \dots P_{[N=n] \cap \overline{A_1} \cap \dots \cap \overline{A_{k-1}}}(A_k) \\
&= \frac{n}{n+1} \frac{n-1}{n} \dots \frac{n-(k-2)}{n+1-(k-2)} \frac{1}{(n+1)-(k-1)} \\
&= \frac{1}{n+1}.
\end{aligned}$$

Et donc la loi de X sachant $[N = n]$ est la loi uniforme sur $[[1, n+1]]$. En particulier, on a

$$E(X|N = n) = \frac{n+2}{2}.$$

Par la formule des probabilités totales appliquée au système complet d'événements $\{[N = n], n \in \mathbf{N}^*\}$, on a alors, sous réserve de convergence,

$$\begin{aligned}
E(X) &= \sum_{n=1}^{+\infty} P(N = n)E(X|N = n) \\
&= \sum_{n=1}^{+\infty} \frac{n+2}{2} p(1-p)^{n-1} \\
&= \frac{1}{2} \sum_{n=1}^{+\infty} np(1-p)^{n-1} + \sum_{n=1}^{+\infty} p(1-p)^{n-1} \\
&= \frac{1}{2p} + 1.
\end{aligned}$$

SOLUTION DE L'EXERCICE 21.10

- Lorsque $m = 1$, X est la variable qui désigne le temps d'attente du premier succès : c'est une loi géométrique de paramètre p .
- X_i est une variable suivant la loi de Bernoulli $\mathcal{B}(p)$.
- $[X = k]$ est réalisé si et seulement si le m -ième succès est obtenu lors du k -ème essai. C'est le cas si et seulement si le k -ème essai est un succès et qu'on a obtenu $m-1$ succès lors des $k-1$ premiers essais, donc

$$[X = k] = [X_1 + \dots + X_{k-1} = m-1] \cap [X_k = 1].$$

- Il va falloir au minimum m essais avant d'obtenir m succès, et il n'y a pas de limite supérieure au nombre d'essais nécessaires, donc $X(\Omega) = [[m, +\infty[$.

Soit donc $k \geq m$, de sorte que $[X = k] = [X_1 + \dots + X_{k-1} = m-1] \cap [X_k = 1]$.

Par le lemme des coalitions⁵, la variable $X_1 + \dots + X_{k-1}$ est indépendante de X_k . Et donc

$$P(X = k) = P(X_1 + \dots + X_{k-1} = m-1)P(X_k = 1).$$

Mais X_1, \dots, X_{k-1} étant mutuellement indépendantes et suivant toutes la loi $\mathcal{B}(p)$, la variable $X_1 + \dots + X_{k-1}$ suit une loi binomiale $\mathcal{B}(k-1, p)$. Et donc

$$P(X = k) = \binom{k-1}{m-1} (1-p)^{k-1-(m-1)} p^{m-1} p = \binom{k-1}{m-1} (1-p)^{k-m} p^m.$$

Remarque

Ce résultat n'est pas forcément si facile à prévoir. On peut toutefois le voir plus simplement en considérant qu'on tire toutes les boules jusqu'à vider l'urne. Il y a alors autant de cas possibles que d'arrangements de $[[0, n]]$, soit $(n+1)!$.

Parmi ceux-ci, $n!$ ont la boule numéro 0 en $k^{\text{ème}}$ position.

Et donc la probabilité de $X = k$ vaut

$$\frac{n!}{(n+1)!} = \frac{1}{n+1}.$$

Calcul

Non seulement les deux séries convergent car on a reconnu des séries de référence (géométrique et géométrique dérivée), mais en plus nous pouvons directement reconnaître les deux sommes : la première est $E(N)$ et la seconde est

$$\sum_{n=1}^{+\infty} P(N = n) = 1.$$

Détail

$X_1 + \dots + X_{k-1}$ est une somme de 0 et de 1, et il y a autant de 1 que de succès lors des $k-1$ premiers essais. C'est donc la variable qui compte le nombre de succès lors des $k-1$ premiers essais.

⁵ Les X_i sont mutuellement indépendantes.

5. Soit T_i la variable aléatoire désignant le nombre de répétitions de l'expérience nécessaires à l'obtention du i -ème succès après l'obtention du $i - 1$ -ème succès. Alors T_i suit une loi géométrique $\mathcal{G}(p)$ et $X = T_1 + \dots + T_m$. Donc

$$E(X) = \sum_{k=1}^n E(T_k) = \sum_{k=1}^n \frac{1}{p} = \frac{m}{p}.$$

SOLUTION DE L'EXERCICE 21.11

1. Une variance étant toujours positive, on a $V(X) = E(X^2) - E(X)^2 \geq 0$, soit $1 - \alpha^2 \geq 0 \Leftrightarrow \alpha^2 \leq 1$ et donc nécessairement, $-1 \leq \alpha \leq 1$.
2. Par la formule de Huygens, on a $V(X^2) = E(X^4) - E(X^2)^2 = 1 - 1 = 0$.
Et donc X^2 est une variable certaine égale à 1 : $P(X^2 = 1) = 1$.
On en déduit que X ne peut prendre que⁶ les valeurs -1 et 1 .
Notons alors $p = P(X = 1)$, de sorte que $1 - p = P(X = -1)$.

On a alors $\alpha = E(X) = -1(1 - p) + p = 2p - 1$. Et alors $p = \frac{\alpha + 1}{2}$.

La loi de X est donc donnée par

$$P(X = 1) = \frac{\alpha + 1}{2}, P(X = -1) = 1 - \frac{\alpha + 1}{2} = \frac{1 - \alpha}{2}.$$

SOLUTION DE L'EXERCICE 21.12

Soit $x \in \mathbf{R}$. Alors on a, par indépendance des X_i ,

$$P(Y > x) = P\left(\bigcap_{i=1}^n [X_i > x]\right) = \prod_{i=1}^n P(X_i > x) = (1 - F_{X_1}(x))^n = \begin{cases} 1 & \text{si } x < 0 \\ e^{-\lambda nx} & \text{si } x \geq 0 \end{cases}$$

Et donc la fonction de répartition de Y est

$$F_Y(x) = P(Y \leq x) = 1 - P(Y > x) = \begin{cases} 0 & \text{si } x < 0 \\ 1 - e^{-\lambda nx} & \text{si } x \geq 0 \end{cases}$$

On reconnaît alors la fonction de répartition d'une loi exponentielle de paramètre $n\lambda$, de sorte que $Y \hookrightarrow \mathcal{E}(n\lambda)$.

Et alors $E(Y) = \frac{1}{n\lambda}$, $V(Y) = \frac{1}{n^2\lambda^2}$.

SOLUTION DE L'EXERCICE 21.13

Il est facile de voir que Y prend ses valeurs dans $]1, +\infty[$, et donc pour $x \leq 1$, on a $P(Y \leq x) = 0$.

Soit donc $x > 1$. Alors, par décroissance de la fonction inverse sur \mathbf{R}_+^* , il vient

$$P(Y \leq x) = P\left(\frac{1}{X} \leq x\right) = P\left(X \geq \frac{1}{x}\right) = 1 - P\left(X < \frac{1}{x}\right) = 1 - P\left(X \leq \frac{1}{x}\right) = 1 - \frac{1}{x}.$$

Ainsi, la fonction de répartition de Y est

$$x \mapsto \begin{cases} 0 & \text{si } x < 1 \\ 1 - \frac{1}{x} & \text{si } x \geq 1 \end{cases}$$

Par opérations sur les fonctions usuelles, F_Y est \mathcal{C}^1 sur $] -\infty, 1[$ et sur $[1, +\infty[$.

De plus, on a $\lim_{x \rightarrow 1^-} F_Y(x) = 0 = F_Y(1)$. Et donc F_Y est continue en 1 et donc sur \mathbf{R} , et \mathcal{C}^1 sauf peut-être en 1.

Ainsi, Y est une variable aléatoire à densité, et une densité en est toute fonction qui coïncide avec F'_Y sur $\mathbf{R} \setminus \{1\}$. Par exemple, on peut prendre

$$f_Y(x) = \begin{cases} 0 & \text{si } x \leq 1 \\ \frac{1}{x^2} & \text{si } x > 1 \end{cases}$$

SOLUTION DE L'EXERCICE 21.14

Remarque

On pourrait prouver que les T_i sont indépendantes, et donc la loi de X est aussi la loi de la convolution de m variables indépendantes suivant la même loi $\mathcal{G}(p)$.

Rappel

Une variable aléatoire de variance nulle est certaine.

⁶ En fait, X peut tout à fait prendre d'autres valeurs, mais avec probabilité nulle, ce qui ne change pas la loi de X .

1. Commençons par remarquer que $e^X - 1$ prend des valeurs positives strictement⁷, donc $\ln(e^X - 1)$ est bien définie.
Pour $x \in \mathbf{R}$, on a, en utilisant la croissance de la fonction exponentielle,

$$P(Y \leq x) = P(\ln(e^X - 1) \leq x) = P(e^X - 1 \leq e^x) = P(X \leq \ln(e^x + 1)) = 1 - e^{-\ln(e^x + 1)}.$$

$$\text{Mais } e^{-\ln(e^x + 1)} = \frac{1}{e^{\ln(e^x + 1)}} = \frac{1}{e^x + 1}.$$

$$\text{Ainsi, } \forall x \in \mathbf{R}, F_Y(x) = 1 - \frac{1}{e^x + 1}.$$

2. F_Y est une fonction de classe \mathcal{C}^1 sur \mathbf{R} , donc Y est une variable à densité, et une densité en est

$$f_Y(x) = F'_Y(x) = \frac{e^x}{(e^x + 1)^2}$$

3. Notons que la densité f_Y est paire⁸. En effet, on a

$$f_Y(-x) = \frac{e^{-x}}{(1 + e^{-x})^2} = \frac{e^{-x}}{(1 + e^{-x})^2} \frac{e^{2x}}{e^{2x}} = \frac{e^x}{(e^x + 1)^2} = f_Y(x).$$

Y admet une espérance si et seulement si $\int_{-\infty}^{+\infty} x f_Y(x) dx$ converge. Et par imparité de $x \mapsto x f_Y(x)$, c'est le cas si et seulement si $\int_0^{+\infty} x f_Y(x) dx$ converge.

La fonction intégrée étant continue sur \mathbf{R}_+ , le seul éventuel problème de convergence est au voisinage de $+\infty$.

$$\text{Mais on a } x f_Y(x) = x^3 \frac{e^x}{(1 + e^x)^2} \underset{x \rightarrow +\infty}{\sim} x^3 \frac{e^x}{e^{2x}} = x e^{-x}.$$

$$\text{Or, } \int_0^{+\infty} x e^{-x} dx = \Gamma(2) \text{ converge, et donc il en est de même de } \int_0^{+\infty} x f_Y(x) dx.$$

On en déduit que Y admet une espérance, et par imparité de $x \mapsto x f_Y(x)$,

$$E(Y) = \int_{-\infty}^{+\infty} x f_Y(x) dx = 0.$$

SOLUTION DE L'EXERCICE 21.15

1. Notons qu'on a $S(t) = P(X > t) = 1 - P(X \leq t) = 1 - F_X(t)$, où F_X est la fonction de répartition de X .

Or, X étant une variable à densité, F_X est continue sur \mathbf{R} , et étant données les hypothèses faites sur f , F_X est \mathcal{C}^1 sur \mathbf{R}_+ , avec $F'_X(t) = f(t)$. Et donc S est de classe \mathcal{C}^1 sur \mathbf{R}_+ , avec $S'(t) = -f(t)$.

Ainsi, une intégration par parties sur $[0, A]$ est légitime, et il vient

$$\begin{aligned} \int_0^A S(t) dt &= [tS(t)]_0^A + \int_0^A t f(t) dt \\ &= AS(A) + \int_0^A t f(t) dt. \end{aligned}$$

2. Si X admet une espérance, alors par définition, $\int_0^{+\infty} t f(t) dt$ converge.

Et donc $\lim_{A \rightarrow +\infty} \int_0^A t f(t) dt$ existe et vaut $E(X)$.

$$\text{On a } S(A) = P(X > A) = \int_A^{+\infty} f(t) dt, \text{ et donc } AS(A) = A \int_A^{+\infty} f(t) dt.$$

Or, pour $t \geq A$, on a $t f(t) \geq A f(t)$, de sorte que par croissance de l'intégrale,

$$\int_A^{+\infty} t f(t) dt \geq A \int_A^{+\infty} f(t) dt = AS(A).$$

De plus, on a $\lim_{A \rightarrow +\infty} \int_A^{+\infty} t f(t) dt = 0$ car il s'agit du reste d'une intégrale convergente.

Et donc par le théorème des gendarmes⁹, on a $\lim_{A \rightarrow +\infty} AS(A) = 0$.

⁷ Plus précisément, la valeur 0 est prise avec une probabilité nulle, car il s'agit de $P(X = 0)$.

Densité

Ici, F_Y est dérivable partout, donc il n'y a aucune précaution à prendre : on peut directement affirmer qu'une densité de Y est F'_Y .

⁸ Si cela semble évident sur le dessin, il n'est pas facile de le deviner sans indication.

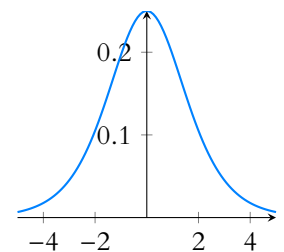


FIGURE 21.1– La densité de Y .

Remarque

Notons que la première intégrale converge automatiquement, car il s'agit du reste de l'intégrale $\int_0^{+\infty} t f(t) dt$, qui converge car X admet une espérance.

⁹ Qui s'applique car $AS(A) \geq 0$.

Et alors en passant à la limite lorsque $A \rightarrow +\infty$, il vient

$$\lim_{A \rightarrow +\infty} \int_0^A S(t) dt = \int_0^{+\infty} tf(t) dt = E(X).$$

Ceci prouve que $\int_0^{+\infty} S(t) dt$ converge et vaut $E(X)$.

3. Inversement, supposons $\int_0^{+\infty} S(t) dt$ converge.

Puisque pour tout $A > 0$, on a $AS(A) \geq 0$, il vient donc

$$\int_0^A tf(t) dt \leq \int_0^A S(t) dt \leq \int_0^{+\infty} S(t) dt.$$

Donc la fonction $A \mapsto \int_0^A tf(t) dt$ est bornée. De plus, elle est croissante par positivité de $t \mapsto tf(t)$.

Par conséquent, elle admet une limite en $+\infty$, et donc $\int_0^{+\infty} tf(t) dt$ converge. Et donc X admet une espérance.

Dans ce cas, la question 2 s'applique, et on a donc $\int_0^{+\infty} S(t) dt = E(X)$.

SOLUTION DE L'EXERCICE 21.16

1. Il est évident que U^2 prend ses valeurs dans $[0, 1]$, et donc pour $x < 0$, on a $P(U^2 \leq x) = 0$, et si $x \geq 1$, alors $P(U^2 \leq x) = 1$.
Soit à présent $x \in [0, 1]$. On a alors

$$P(U^2 \leq x) = P(-\sqrt{x} \leq U \leq \sqrt{x}) = P(U \leq \sqrt{x}) = \sqrt{x}.$$

La fonction de répartition de U^2 est donc

$$F_{U^2}(x) = \begin{cases} 0 & \text{si } x < 0 \\ \sqrt{x} & \text{si } 0 \leq x \leq 1 \\ 1 & \text{si } x > 1 \end{cases}$$

Il est clair que cette fonction est \mathcal{C}^1 sauf éventuellement en 0 et en 1.

De plus, on a $\lim_{x \rightarrow 0^-} F_{U^2}(x) = \lim_{x \rightarrow 0^+} F_{U^2}(x) = 0 = F_{U^2}(0)$, de sorte que F_{U^2} est continue en 0.

De même, on prouve que F_{U^2} est continue en 1, et donc est continue sur \mathbf{R} .
Ainsi, U^2 est une variable aléatoire à densité, dont une densité est donnée par

$$f_{U^2} : x \mapsto \begin{cases} 0 & \text{si } x \leq 0 \\ \frac{1}{2\sqrt{x}} & \text{si } 0 < x < 1 \\ 0 & \text{si } x \geq 1 \end{cases}$$

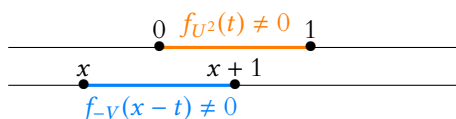
2. Il s'agit d'une transformation affine de densité : une densité de $-V$ est

$$f_{-V}(x) = f_V(-x) = \begin{cases} 1 & \text{si } -x \in [0, 1] \\ 0 & \text{sinon} \end{cases} = \begin{cases} 1 & \text{si } x \in [-1, 0] \\ 0 & \text{sinon} \end{cases}$$

3. Les variables U et V étant indépendantes, il en est de même de U^2 et de $-V$.
De plus, une densité de $-V$ est bornée donc une densité de $U^2 - V$ est donnée par le produit de convolution :

$$f_T(x) = \int_{-\infty}^{\infty} f_{U^2}(t)f_{-V}(x-t) dt.$$

On a $f_{U^2}(t) \neq 0 \Leftrightarrow 0 \leq t \leq 1$ et $f_{-V}(x-t) \neq 0 \Leftrightarrow -1 \leq x-t \leq 0 \Leftrightarrow x \leq t \leq x+1$.



Remarque

Ce résultat est à mettre en parallèle avec celui de l'exercice 1.9 qui démontre un résultat analogue pour des variables à valeurs dans \mathbf{N} .

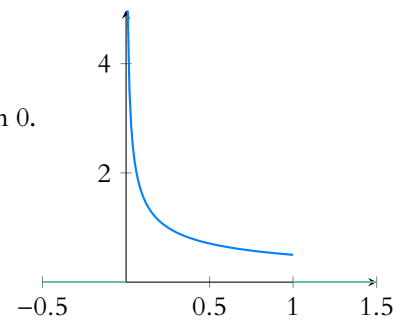


FIGURE 21.2- La densité de U^2 .

Alternative

Par transformation affine de loi uniforme, $-V \leftrightarrow \mathcal{U}([-1, 0])$.

Rappel

Le dessin sert à déterminer les valeurs de t pour lesquelles $f_{U^2}(t)f_{-V}(x-t)$ est nul, et donc à restreindre le domaine d'intégration.

Si $x > 1$, alors $f_{U^2-V}(x) = 0$.

De même, si $x \leq -1$, on a $x + 1 \leq 0$ et donc $f_T(x) = 0$.

Si $x \in [-1, 0]$, alors

$$\begin{aligned} f_T(x) &= \int_0^{x+1} \frac{1}{2\sqrt{t}} dt = \lim_{A \rightarrow 0^+} \int_A^{x+1} \frac{1}{2\sqrt{t}} dt \\ &= \lim_{A \rightarrow 0^+} [\sqrt{t}]_A^{x+1} = \lim_{A \rightarrow 0^+} (\sqrt{x+1} - \sqrt{A}) = \sqrt{x+1}. \end{aligned}$$

Enfin, si $x \in [0, 1]$, alors

$$f_T(x) = \int_x^1 \frac{1}{2\sqrt{t}} dt = [\sqrt{t}]_x^1 = 1 - \sqrt{x}.$$

En résumé, une densité de T est bien donnée par

$$f_T(x) = \begin{cases} \sqrt{x+1} & \text{si } x \in [-1, 0] \\ 1 - \sqrt{x} & \text{si } x \in [0, 1] \\ 0 & \text{sinon} \end{cases}$$

4. P_ω est un polynôme de degré 2, et donc il est de signe constant si et seulement si son discriminant est négatif.

Or, ce discriminant est $\Delta(\omega) = 4U(\omega)^2 - 4V(\omega) = 4T(\omega)$.

Et donc on a

$$P(\Delta \leq 0) = P(T \leq 0) = \int_{-\infty}^0 f_T(t) dt = \int_{-1}^0 \sqrt{t+1} dt = \left[\frac{2}{3}(t+1)^{3/2} \right]_{-1}^0 = \frac{2}{3}.$$

SOLUTION DE L'EXERCICE 21.17

1. Il est clair que F est \mathcal{C}^1 sur \mathbf{R}^+ et sur \mathbf{R}^* .

De plus, elle est continue en 0 car on a $\lim_{x \rightarrow 0^-} F(x) = 0 = \lim_{x \rightarrow 0^+} F(x) = F(0)$.

Et donc F est continue sur \mathbf{R} et \mathcal{C}^1 sauf éventuellement en 0.

Comme de plus elle est croissante sur \mathbf{R} et vérifie $\lim_{x \rightarrow -\infty} F(x) = 0$ et $\lim_{x \rightarrow +\infty} F(x) = 1$, il s'agit de la fonction de répartition d'une variable à densité X .

Une densité de X est alors toute fonction qui coïncide avec F' là où celle-ci est définie. Par exemple, on peut prendre

$$f'(x) = \begin{cases} 0 & \text{si } x < 0 \\ \frac{a}{b} \left(\frac{x}{b}\right)^{a-1} e^{-\left(\frac{x}{b}\right)^a} & \text{si } x \geq 0 \end{cases}$$

2. Si $a = 1$, la fonction de répartition de X est $x \mapsto \begin{cases} 0 & \text{si } x < 0 \\ 1 - e^{x/b} & \text{si } x \geq 0 \end{cases}$

On reconnaît alors la fonction de répartition d'une loi exponentielle de paramètre $\frac{1}{b}$ et

donc $X \hookrightarrow \mathcal{E}\left(\frac{1}{b}\right)$.

3. Puisque F est nulle sur $] -\infty, 0[$, X ne prend que des valeurs positives, et donc il en est de même de X^a .

Soit $x \geq 0$. Alors

$$P(X^a \leq x) = P(X \leq x^{1/a}) = 1 - e^{x/b^a}.$$

Autrement dit, X^a suit la loi exponentielle $\mathcal{E}\left(\frac{1}{b^a}\right)$.

4. D'après le théorème de transfert, X admet un moment d'ordre k si et seulement si

$\int_0^{+\infty} x^k \frac{a}{b} \left(\frac{x}{b}\right)^{a-1} e^{-\left(\frac{x}{b}\right)^a} dx$ converge¹⁰.

Procédons alors au changement de variable indiqué : $t = \left(\frac{x}{b}\right)^a$, qui est légitime car de classe \mathcal{C}^1 et strictement croissant.

Remarque

Notons que l'intégrale est une intégrale impropre : $\frac{1}{\sqrt{t}}$ n'est pas définie en $t = 1$.

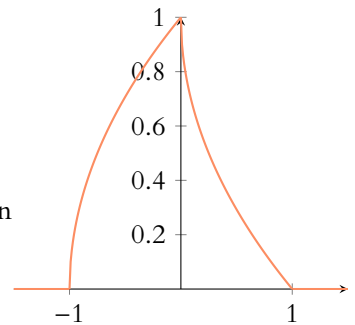


FIGURE 21.3— La densité de T .

Continuité

Notons que pour prouver que F est une fonction de répartition, il faut prouver sa continuité à droite en tout point. Mais étant F continue en tout point, elle est évidemment continue à droite en tout point.

¹⁰ Absolument, mais il s'agit de l'intégrale d'une fonction positive.

On a alors, $x = bt^{1/a}$, et donc $dx = \frac{b}{a}t^{1/a-1} dt$. Et donc, toujours sous réserve de convergence,

$$E(X^k) = \frac{a}{b} \int_0^{+\infty} b^k t^{k/a} t^{1-\frac{1}{a}} e^{-t} \frac{b}{a} t^{1/a-1} dt = b^k \int_0^{+\infty} t^{k/a+1-1} e^{-t} dt = b^k \Gamma\left(1 + \frac{k}{a}\right).$$

Notons que la convergence a bien été établie car on a reconnu une intégrale usuelle. Et donc X admet un moment d'ordre k , égal à $b^k \Gamma\left(1 + \frac{k}{a}\right)$.

En particulier, $E(X) = b\Gamma\left(1 + \frac{1}{a}\right)$ et par la formule de Huygens,

$$V(X) = E(X^2) - E(X)^2 = b^2 \left(\Gamma\left(1 + \frac{2}{a}\right) - \Gamma\left(1 + \frac{1}{a}\right)^2 \right).$$

5. Soit $x \in \mathbf{R}$. Alors, par indépendance des X_i ,

$$P(Y_n > x) = P\left(\bigcap_{i=1}^n [X_i > x]\right) = \prod_{i=1}^n P(X_i \geq x) = \begin{cases} 1 & \text{si } x < 0 \\ e^{-n(\frac{x}{b})a} & \text{si } x \geq 0 \end{cases}$$

Et donc la fonction de répartition de Y_n est

$$x \mapsto P(Y_n \leq x) = 1 - P(Y_n > x) = \begin{cases} 0 & \text{si } x < 0 \\ 1 - e^{-n(\frac{x}{b})a} & \text{si } x \geq 0 \end{cases}$$

Mais remarquons que $n\left(\frac{x}{b}\right)^a = \left(\frac{xn^{1/a}}{b}\right)^a$, de sorte que $Y_n \hookrightarrow \mathcal{W}\left(a, \frac{b}{n^{1/a}}\right)$.

SOLUTION DE L'EXERCICE 21.18

1. Pour tout $x \in \mathbf{R}$, on a, par indépendance des X_i ,

$$P(\max(X_1, \dots, X_k) \leq x) = P\left(\bigcap_{i=1}^k [X_i \leq x]\right) = P(X_1 \leq x)^k = \begin{cases} 0 & \text{si } x < 0 \\ \left(\frac{x}{a}\right)^k & \text{si } 0 \leq x \leq a \\ 1 & \text{si } x > a \end{cases}$$

Il est facile de constater que cette fonction de répartition est continue sur \mathbf{R} et \mathcal{C}^1 sauf éventuellement en 0 et en a , de sorte que $\max(X_1, \dots, X_k)$ est une variable à densité, et une densité en est donnée par

$$f_k : x \mapsto \begin{cases} k \frac{x^{k-1}}{a^k} & \text{si } 0 \leq x \leq a \\ 0 & \text{sinon} \end{cases}$$

Et donc, par transformation affine de densité, une densité de $-\max(X_1, \dots, X_k)$ est

$$g_k : x \mapsto f_k(-x) = \begin{cases} k \frac{(-x)^{k-1}}{a^k} & \text{si } -a \leq x \leq 0 \\ 0 & \text{sinon} \end{cases}$$

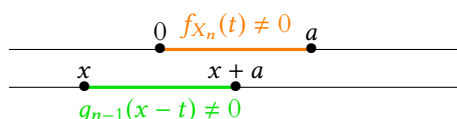
2. On a

$$P([X_n \geq X_1] \cap [X_n \geq X_2] \cap \dots \cap [X_n \geq X_{n-1}]) = P(\max(X_1, \dots, X_{n-1}) \leq X_n) = P(X_n - \max(X_1, \dots, X_{n-1}) \geq 0).$$

Par le lemme des coalitions, X_n et $-\max(X_1, \dots, X_{n-1})$ sont indépendantes, et donc une densité de $X_n - \max(X_1, \dots, X_{n-1})$ est donnée par le produit de convolution :

$$f : x \mapsto \int_{-\infty}^{+\infty} f_{X_n}(t) g_{n-1}(x-t) dt.$$

Or, on a $f_{X_n}(t) \neq 0 \Leftrightarrow 0 \leq t \leq a$ et $g_{n-1}(x-t) \neq 0 \Leftrightarrow -a \leq x-t \leq 0 \Leftrightarrow x \leq t \leq x+a$.



Ainsi, si $x + a \leq 0 \Leftrightarrow x \leq -a$ ou si $x \geq a$, on a $f(x) = 0$.

Si $-a \leq x \leq 0$, alors on a

$$\begin{aligned} f(x) &= \int_0^{x+a} \frac{1}{a} (n-1) \frac{(t-x)^{n-2}}{a^{n-1}} dt = \frac{1}{a^n} \int_0^{x+a} (n-1)(t-x)^{n-2} dt \\ &= \frac{1}{a^n} \left[(t-x)^{n-1} \right]_0^{x+a} = 1 - \frac{(-x)^{n-1}}{a^n}. \end{aligned}$$

Si $0 \leq x \leq a$, alors on a

$$f(x) = \int_x^a \frac{1}{a^n} (n-1)(t-x)^{n-2} dt = \frac{1}{a^n} \left[(t-x)^{n-1} \right]_x^a = \frac{(a-x)^{n-1}}{a^n}.$$

Et donc en particulier, il vient

$$P(X_n - \max(X_1, \dots, X_{n-1}) \geq 0) = \int_0^{+\infty} f(x) dx = \int_0^a \frac{(a-x)^{n-1}}{a^n} dx = \left[-\frac{1}{n} \frac{(a-x)^n}{a^n} \right]_0^a = \frac{1}{n}.$$

SOLUTION DE L'EXERCICE 21.19

1. D'après le théorème de transfert, on a pour tout t réel, et sous réserve de convergence¹¹

$$\begin{aligned} E(e^{tX}) &= \int_{-\infty}^{+\infty} e^{tx} \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-x^2/2+tx} dx \\ &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-\frac{1}{2}(x-t)^2+t^2/2} dx = \frac{e^{t^2/2}}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-\frac{1}{2}(x-t)^2} dx. \end{aligned}$$

Procédons alors au changement de variable affine $u = x - t$, de sorte que¹²

$$E(e^{tX}) = \frac{e^{t^2/2}}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-u^2/2} du = e^{t^2/2} \frac{1}{\sqrt{2\pi}} \sqrt{2\pi} = e^{t^2/2}.$$

Et donc X est bien 1-sous-gaussienne.

2. On a $E(e^{tX}) = E(e^{t \sum_{i=1}^n \mu_i X_i}) = E\left(\prod_{i=1}^n e^{t \mu_i X_i}\right)$.

Mais les $e^{t \mu_i X_i}$ sont mutuellement indépendantes car les X_i le sont, de sorte que

$$E\left(\prod_{i=1}^n e^{t \mu_i X_i}\right) = \prod_{i=1}^n E(e^{t \mu_i X_i}) \leq \prod_{i=1}^n e^{\alpha^2 \mu_i^2 t^2 / 2} = e^{\alpha^2 t^2 / 2 \sum_{i=1}^n \mu_i^2} = e^{\alpha^2 t^2 / 2}.$$

Et donc X est α -sous-gaussienne.

- 3.a. D'après l'inégalité de Markov, qui s'applique car e^{tX} est une variable aléatoire positive admettant une espérance, on a

$$P(X \geq \lambda) = P(e^{tX} \geq e^{t\lambda}) \leq \frac{E(e^{tX})}{e^{t\lambda}} \leq \exp\left(\frac{\alpha^2 t^2}{2} - t\lambda\right).$$

- 3.b. Notons que $P(|X| \geq \lambda) = P(X \geq \lambda) + P(X \leq -\lambda)$.

Mais d'après la question 2, avec $n = 1$ et $\lambda_1 = -1$, $-X$ est également α -sous-gaussienne.

Et donc $P(-X \geq \lambda) \leq \exp\left(\frac{\alpha^2 t^2}{2} - t\lambda\right)$.

On en déduit donc que

$$P(|X| \geq \lambda) \leq 2 \exp\left(\frac{\alpha^2 t^2}{2} - t\lambda\right).$$

Cette inégalité est valable pour tout t réel.

Soit donc φ la fonction définie sur \mathbf{R} par $t \mapsto \frac{\alpha^2 t^2}{2} - t\lambda$. Il est facile¹³ de voir que φ possède un minimum atteint en $t = \frac{\lambda}{\alpha^2}$, et ce minimum vaut alors

$$\varphi\left(\frac{\lambda}{\alpha^2}\right) = \frac{\lambda^2}{2\alpha^2} - \frac{\lambda^2}{\alpha^2} = -\frac{\lambda^2}{2\alpha^2}.$$

Remarque

Un peu d'intuition nous permet de retrouver aisément le fait que cette densité est nulle en dehors de $[-a, a]$, car il s'agit du support de $X_n - \max(X_1, \dots, X_{n-1})$.

Remarque

Ce résultat est plutôt intuitif : chacune des n variables X_1, \dots, X_n a la même probabilité d'être plus grande que les autres.

¹¹ Absolue, mais il s'agit de l'intégrale d'une fonction positive, donc la convergence absolue est équivalente à la convergence.

¹² Toujours sous réserve de convergence.

Intégrale

La valeur de cette dernière intégrale se retrouve grâce au fait que l'intégrale de la densité de la loi $\mathcal{N}(0, 1)$ est égale à 1.

Rappel

L'espérance d'un produit de variables indépendantes admettant une espérance est égale au produit des espérances.

¹³ Par exemple à l'aide d'une formule apprise au lycée donnant le minimum d'un polynôme de degré 2.

Et donc, en prenant $t = \frac{\lambda}{\alpha^2}$, l'inégalité obtenue précédemment devient

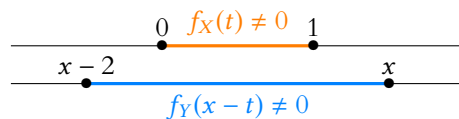
$$P(|X| \geq \lambda) \leq 2 \exp\left(-\frac{\lambda^2}{2\alpha^2}\right).$$

SOLUTION DE L'EXERCICE 21.20

Notons qu'à chaque fois, les variables sont indépendantes, et leurs densités sont bornées, donc il est légitime d'utiliser le produit de convolution : une densité de $X + Y$ est

$$f : x \mapsto \int_{-\infty}^{+\infty} f_X(t) f_Y(x-t) dt.$$

1. Ici, on a $f_X(t) \neq 0 \Leftrightarrow 0 \leq t \leq 1$, et $f_Y(x-t) \neq 0 \Leftrightarrow 0 \leq x-t \leq 2 \Leftrightarrow x-2 \leq t \leq x$.



La situation est un peu plus subtile que la plupart de celles que nous avons déjà rencontré, puisque l'intervalle du bas est plus grand que celui du haut, et donc est susceptible de le recouvrir entièrement.

Il est clair que pour $x-2 \geq 1 \Leftrightarrow x \geq 3$ et $x \leq 0$, $f(x) = 0$.

Si $0 \leq x \leq 1$, alors on a

$$f(x) = \int_0^x 1 \times \frac{1}{2} dt = \frac{x}{2}.$$

Si $1 \leq x \leq 2$, alors

$$f(x) = \int_0^1 1 \times \frac{1}{2} dt = \frac{1}{2}.$$

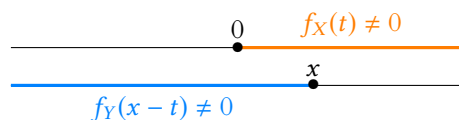
Enfin, si $2 \leq x \leq 3$, alors

$$f(x) = \int_{x-2}^1 1 \times \frac{1}{2} dt = \frac{3-x}{2}.$$

En résumé, une densité de $X + Y$ est donc

$$f : x \mapsto \begin{cases} 0 & \text{si } x \leq 0 \\ \frac{x}{2} & \text{si } 0 \leq x \leq 1 \\ \frac{1}{2} & \text{si } 1 \leq x \leq 2 \\ \frac{3-x}{2} & \text{si } 2 \leq x \leq 3 \\ 0 & \text{si } x \geq 3 \end{cases}$$

2. Cette fois on a $f_X(t) \neq 0 \Leftrightarrow t \geq 0$ et $f_Y(x-t) \neq 0 \Leftrightarrow x-t \geq 0 \Leftrightarrow t \leq x$.



Si $x < 0$, alors $f(x) = 0$.

Et si $x \geq 0$, alors

$$f(x) = \int_0^x e^{-t} 2e^{-2(x-t)} dt = 2e^{-2x} \int_0^x e^t dt = 2e^{-2x} (e^x - 1) = 2(e^{-x} - e^{-2x}).$$

3. Il est facile de voir que f est continue et positive sur \mathbf{R} .

De plus, f est paire, donc $\int_{-\infty}^{+\infty} f(t) dt$ converge si et seulement si $\int_0^{+\infty} f(t) dt$ converge.

Or

$$\int_0^{+\infty} f(t) dt = \frac{1}{2} \int_0^{+\infty} e^{-t} dt = \frac{1}{2}.$$

Détails

Puisque X suit une loi uniforme, là où sa densité n'est pas nulle, elle vaut 1. Et de même celle de Y est soit nulle, soit égale à $\frac{1}{2}$.

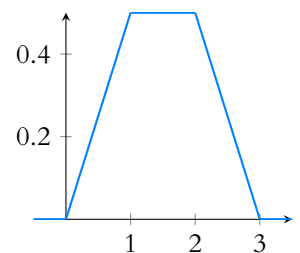


FIGURE 21.4— La densité f .

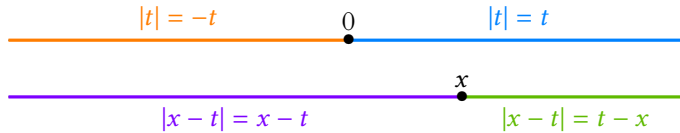
Donc $\int_{-\infty}^{+\infty} f(t) dt$ converge et vaut $2 \times \frac{1}{2} = 1$.

Donc f est bien une densité. Une densité h de $X + Y$ est alors définie par

$$h(x) = \int_{-\infty}^{+\infty} f(t)f(x-t) dt = \frac{1}{4} \int_{-\infty}^{+\infty} e^{-|t|} e^{-|x-t|} dt.$$

L'intégrande n'est jamais nulle, mais il y a le problème des valeurs absolues.

Adaptons un peu le dessin habituel : on a $|t| = t \Leftrightarrow t \geq 0$ et $|x-t| = x-t \Leftrightarrow x-t \geq 0 \Leftrightarrow x \geq t$.



Si $x \geq 0$, alors

$$\begin{aligned} h(x) &= \frac{1}{4} \left(\int_{-\infty}^0 e^t e^{t-x} dt + \int_0^x e^{-t} e^{t-x} dt + \int_x^{+\infty} e^{-t} e^{x-t} dt \right) \\ &= \frac{1}{4} \left(e^{-x} \int_{-\infty}^0 e^{2t} dt + e^{-x} \int_0^x 1 dt + e^x \int_x^{+\infty} e^{-2t} dt \right) \\ &= \frac{1}{4} \left(\frac{1}{2} e^{-x} + x e^{-x} + \frac{1}{2} e^x e^{-2x} \right) = \frac{e^{-x}}{4} (x+1). \end{aligned}$$

Si $x \leq 0$, alors en procédant de même, on a

$$\begin{aligned} h(x) &= \frac{1}{4} \left(\int_{-\infty}^x e^t e^{t-x} dt + \int_x^0 e^t e^{-(t-x)} dt + \int_0^{+\infty} e^{-t} e^{x-t} dt \right) \\ &= \frac{1}{4} \left(e^{-x} \int_{-\infty}^x e^{2t} dt + e^x \int_x^0 1 dt + e^x \int_0^{+\infty} e^{-2t} dt \right) \\ &= \frac{1}{4} \left(\frac{e^{-x}}{2} e^{2x} - x e^x + e^x \frac{1}{2} \right) \\ &= \frac{1}{4} \left(\frac{e^x}{2} - x e^x + \frac{e^x}{2} \right) \\ &= \frac{e^x}{4} \left(\frac{1}{2} - x + \frac{1}{2} \right) = \frac{e^x}{4} (1-x). \end{aligned}$$

Notons qu'il est possible de regrouper les deux formules obtenues sous une seule forme

$$\forall x \in \mathbf{R}, h(x) = \frac{1}{4} e^{-|x|} (1 + |x|).$$

SOLUTION DE L'EXERCICE 21.21

Puisque les $|X_i|$ sont positives, il est possible d'appliquer l'inégalité de Markov : pour tout $\varepsilon > 0$, on a

$$P(|X_n| \geq \varepsilon) \leq \frac{E(|X_n|)}{\varepsilon} \xrightarrow{n \rightarrow +\infty} 0.$$

Et donc par le théorème des gendarmes¹⁴, il vient donc $\lim_{n \rightarrow +\infty} P(|X_n| \geq \varepsilon) = 0$.

Ainsi, on a bien $X_n \xrightarrow{P} 0$.

SOLUTION DE L'EXERCICE 21.22

Notons que Y_n vérifie

$$E(Y_n) = \frac{E(X_n) - n\theta}{n} = \frac{n\theta - n\theta}{n} = 0 \text{ et } V(Y_n) = \frac{V(X_n)}{n^2} = \frac{n\theta}{n^2} = \frac{\theta}{n}.$$

Soit $\varepsilon > 0$. Alors par l'inégalité de Bienaymé-Tchebychev¹⁵, on a

$$P(|Y_n| \geq \varepsilon) \leq \frac{V(Y_n)}{\varepsilon^2} = \frac{\theta}{n\varepsilon^2} \xrightarrow{n \rightarrow \infty} 0.$$

Donc $Y_n \xrightarrow{P} 0$.

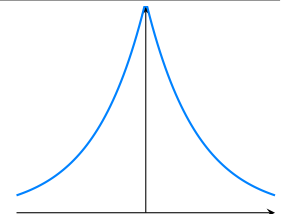


FIGURE 21.5– La densité f .

Méthode

On ne sait à peu près jamais intégrer une quantité avec des valeurs absolues. Pour calculer une telle intégrale, il faut donc la « découper » (à l'aide de Chasles) en différents morceaux sur lesquels on peut se passer de la valeur absolue, en revenant à la définition :

$$|x| = \begin{cases} x & \text{si } x \geq 0 \\ -x & \text{sinon} \end{cases}$$

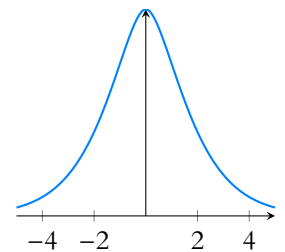


FIGURE 21.6– La densité h .

¹⁴ Une probabilité est positive !

¹⁵ Qui s'applique ici car Y_n possède une variance.

SOLUTION DE L'EXERCICE 21.23

1. Nous savons, par stabilité des lois de Poisson¹⁶, que $S_n \hookrightarrow \mathcal{P}(n)$.
On a alors

$$P(S_n \leq n) = \sum_{k=0}^n P(S_n = k) = \sum_{k=0}^n e^{-n} \frac{n^k}{k!}.$$

2. On a $E(S_n) = n$ et $V(S_n) = n$. Donc, par le théorème central limite, si on note

$$\overline{S}_n^* = \frac{S_n - n}{\sqrt{n}}$$

la variable centrée réduite associée à S_n , alors $\overline{S}_n^* \xrightarrow{\mathcal{L}} X$ où $X \hookrightarrow \mathcal{N}(0, 1)$.

De plus, on a $P(S_n \leq n) = P(\overline{S}_n^* \leq 0)$. Et la convergence en loi nous indique que

$$\lim_{n \rightarrow +\infty} P(\overline{S}_n^* \leq 0) = P(X \leq 0) = \Phi(0) = \frac{1}{2}.$$

On en déduit que

$$\lim_{n \rightarrow +\infty} e^{-n} \sum_{k=0}^n \frac{n^k}{k!} = \frac{1}{2} \Leftrightarrow \sum_{k=0}^n \frac{n^k}{k!} \underset{n \rightarrow +\infty}{\sim} \frac{e^n}{2}.$$

SOLUTION DE L'EXERCICE 21.24

- 1.a. Il est évident que $\frac{1}{2}U_1^2$ ne prend que des valeurs réelles, et donc pour $x < 0$, $P\left(\frac{1}{2}U_1^2 \leq x\right) = 0$.
Soit donc $x \geq 0$. Alors

$$P\left(\frac{1}{2}U_1^2 \leq x\right) = P(U_1^2 \leq 2x) = P(-\sqrt{2x} \leq U_1 \leq \sqrt{2x}) = \Phi(\sqrt{2x}) - \Phi(-\sqrt{2x}) = 2\Phi(\sqrt{2x}) - 1.$$

Ainsi, la fonction de répartition de $\frac{1}{2}U_1^2$ est $F : x \mapsto \begin{cases} 0 & \text{si } x < 0 \\ 2\Phi(\sqrt{2x}) - 1 & \text{si } x \geq 0 \end{cases}$

Puisque Φ est une fonction continue¹⁷, F est continue sur $[0, +\infty[$, et elle l'est évidemment sur $]-\infty, 0[$.

De plus, on a $F(0) = 2\Phi(0) - 1 = 2 \cdot \frac{1}{2} - 1 = 0 = \lim_{x \rightarrow 0^-} F(x)$. Et donc F est continue en 0 et donc continue sur \mathbf{R} .

De plus, Φ est \mathcal{C}^1 sur \mathbf{R} , de sorte que F est \mathcal{C}^1 sauf peut-être en 0, et donc $\frac{1}{2}U_1^2$ est une variable à densité.

Une densité en est alors toute fonction qui coïncide avec F' là où celle-ci est définie¹⁸. Par exemple, on peut prendre

$$f : x \mapsto \begin{cases} 0 & \text{si } x \leq 0 \\ \frac{2}{\sqrt{2x}} \Phi'(\sqrt{2x}) & \text{si } x > 0 \end{cases} = \begin{cases} 0 & \text{si } x \leq 0 \\ \frac{1}{\sqrt{\pi}\sqrt{x}} e^{-x} & \text{si } x > 0 \end{cases}$$

Mais f étant une densité, on a alors

$$\int_{-\infty}^{+\infty} f(t) dt = 1 \Leftrightarrow \int_0^{+\infty} \frac{1}{\sqrt{\pi}} t^{1/2-1} e^{-t} dt = 1 \Leftrightarrow \Gamma(1/2) = \sqrt{\pi}.$$

Et donc on a $f : x \mapsto \begin{cases} 0 & \text{si } x < 0 \\ \frac{1}{\Gamma(1/2)} x^{1/2-1} e^{-x} & \text{si } x \geq 0 \end{cases}$

On reconnaît alors la densité d'une loi γ de paramètre $1/2$, de sorte que $\frac{1}{2}U_1^2 \hookrightarrow \gamma(1/2)$.

- 1.b. Puisque les U_i sont indépendantes, il en est de même des $\frac{1}{2}U_i^2$, et donc, par stabilité des lois γ , $W_n \hookrightarrow \gamma(n/2)$.

En particulier, on a $E(W_n) = V(W_n) = \frac{n}{2}$.

¹⁶ Qui s'applique bien ici par indépendance des X_i .

Rédaction

On aurait envie de noter

$$\overline{S}_n^* \xrightarrow{\mathcal{L}} \mathcal{N}(0, 1).$$

Malheureusement, cette notation est incorrecte car on parle de convergence en loi vers une variable aléatoire. Donc on est obligé d'introduire une variable aléatoire X suivant la loi normale centrée réduite.

¹⁷ C'est la fonction de répartition d'une variable à densité.

Astuce

Lorsqu'on a précisé que F était continue sur $[0, +\infty[$, cela impliquait en particulier sa continuité à droite en 0. Et donc il suffit d'étudier la continuité à gauche en 0.

¹⁸ Ici, sur \mathbf{R}^* .

1.c. La fonction de répartition F_{W_n} de W_n est continue sur \mathbf{R}_+ , car fonction de répartition d'une variable à densité. Elle y est même strictement croissante puisque sa dérivée y est $x \mapsto \frac{1}{\Gamma(n/2)} x^{n/2-1} e^{-x} > 0$ et on a $F_{W_n}(0) = 0$ et $\lim_{x \rightarrow +\infty} F_{W_n}(x) = 1$.

Par le théorème de la bijection, il existe donc un unique $k_\alpha \geq 0$ tel que $F_{W_n}(k_\alpha) = \alpha \Leftrightarrow P(W_n \leq k_\alpha) = \alpha$.

2.a. Les X_i^2 sont indépendantes, et on a, pour tout i , $E(X_i^2) = V(X_i) + E(X_i)^2 = \sigma^2$.

Par linéarité de l'espérance, il vient donc $E(T_n) = \frac{1}{n} \sum_{i=1}^n X_i^2 = \sigma^2$.

Ainsi, T_n est un estimateur sans biais de σ^2 .

2.b. Les vecteurs aléatoires (U_1, \dots, U_n) et $\left(\frac{X_1}{\sigma}, \dots, \frac{X_n}{\sigma}\right)$ ont même loi.

Puisque la fonction $g : (x_1, \dots, x_n) \mapsto \sum_{i=1}^n x_i^2$ est continue sur \mathbf{R}^n , on en déduit que

$g(U_1, \dots, U_n)$ et $g\left(\frac{X_1}{\sigma}, \dots, \frac{X_n}{\sigma}\right)$ ont même loi.

Or, $g(U_1, \dots, U_n) = W_n$ et $g\left(\frac{X_1}{\sigma}, \dots, \frac{X_n}{\sigma}\right) = \frac{n}{2\sigma^2} T_n$.

On en déduit notamment que $V\left(\frac{n}{2\sigma^2} T_n\right) = V(W_n) = \frac{n}{2}$.

Et par conséquent,

$$V(T_n) = \frac{4\sigma^4 n}{n^2 \cdot 2} = \frac{2\sigma^4}{n}.$$

Ainsi, le risque quadratique de T_n est

$$r(T_n) = b(T_n)^2 + V(T_n) = \frac{2\sigma^4}{n} \xrightarrow{n \rightarrow +\infty} 0.$$

Et donc T_n est un estimateur convergent de σ^2 .

2.c. On a

$$\begin{aligned} P\left(0 < \sigma^2 \leq \frac{S_n}{2k_\alpha}\right) &= P\left(k_\alpha \sigma^2 \leq \frac{S_n}{2}\right) = P\left(k_\alpha \leq \frac{S_n}{2\sigma^2}\right) = 1 - P\left(\frac{S_n}{2\sigma^2} \leq k_\alpha\right) \\ &= P\left(\frac{n}{2\sigma^2} T_n \leq k_\alpha\right) = 1 - P(W_n \leq k_\alpha) = 1 - \alpha. \end{aligned}$$

Et donc $\left]0, \frac{S_n}{2k_\alpha}\right]$ est un intervalle de confiance de σ^2 au niveau de risque α .

SOLUTION DE L'EXERCICE 21.25

On a $E(\overline{X}_n) = \frac{1}{2}$ et $V(\overline{X}_n) = \frac{1}{4n}$, de sorte que la variable centrée réduite associée à \overline{X}_n est

$$\overline{X}_n^* = 2\sqrt{n}\left(\overline{X}_n - \frac{1}{2}\right).$$

Mais d'après le théorème central limite, qui s'applique car les X_i sont i.i.d. et admettent une variance, on a $\overline{X}_n^* \xrightarrow{\mathcal{L}} X$, où $X \hookrightarrow \mathcal{N}(0, 1)$.

En particulier, on a

$$P\left(\left|\overline{X}_n - \frac{1}{2}\right| > \frac{1}{\sqrt{n}}\right) = P\left(\left|\frac{\overline{X}_n^*}{2}\right| > 1\right) = 1 - P\left(\left|\frac{\overline{X}_n^*}{2}\right| \leq 1\right) = 1 - P(-2 \leq \overline{X}_n^* \leq 2).$$

Puisque $\overline{X}_n^* \xrightarrow{\mathcal{L}} X$, cette probabilité converge donc, lorsque $n \rightarrow +\infty$ vers

$$1 - P(-2 \leq X \leq 2) = 1 - \Phi(2) + \Phi(-2) = 2(1 - \Phi(2)).$$

SOLUTION DE L'EXERCICE 21.26

1. Notons que par stabilité des lois binomiales¹⁹, on a $S_n \hookrightarrow \mathcal{B}(n, p)$.

Par conséquent, il vient $E(S_n) = np$ et $V(S_n) = np(1-p)$.

Par la formule de Huygens, on a donc $E(S_n^2) = V(S_n) + E(S_n)^2 = np(1-p) + (np)^2$.

Enfin, on a, par linéarité de l'espérance

$$E(U_n) = E\left(\frac{S_n}{n} - \frac{S_n^2}{n^2}\right) = \frac{1}{n}E(S_n) - \frac{1}{n^2}E(S_n^2) = p - \frac{p(1-p)}{n} - p^2 = p(1-p) - \frac{p(1-p)}{n} = \frac{(n-1)p(1-p)}{n}.$$

Méthode

Puisqu'il s'agit de calculer la limite d'une probabilité et que cette probabilité n'est pas directement calculable, on s'orientera vers une convergence en loi.

En l'absence d'indices, et étant donné le contexte (\overline{X}_n est la moyenne de variables i.i.d.), il faut probablement chercher du côté du théorème central limite.

¹⁹ Qui s'applique car les X_i sont indépendantes.

2. Par définition, on a

$$b(U_n) = E(U_n) - p(1-p) = \frac{(n-1)p(1-p)}{n} - p(1-p) = -\frac{p(1-p)}{n}.$$

Un estimateur sans biais de $p(1-p)$ est donc $\frac{n}{n-1}U_n = \frac{S_n}{n-1} \left(1 - \frac{S_n}{n}\right)$.

SOLUTION DE L'EXERCICE 21.27

Nous savons déjà que \overline{X}_n est sans biais et que

$$V(\overline{X}_n) = \frac{1}{n}V(X_1) = \frac{1}{n} \frac{1}{\lambda^2}$$

de sorte que

$$r(\overline{X}_n) = 0 + \frac{1}{n\lambda^2}.$$

De plus, $T_n = \frac{n}{n+1}\overline{X}_n$, et donc

$$E(T_n) = \frac{n}{n+1}E(\overline{X}_n) = \frac{n}{n+1} \frac{1}{\lambda} \text{ et } V(T_n) = \frac{n^2}{(n+1)^2}V(\overline{X}_n) = \frac{n}{(n+1)^2\lambda^2}.$$

On en déduit que le biais de T_n en $\frac{1}{\lambda}$ est $b(T_n) = \frac{n}{n+1} \frac{1}{\lambda} - \frac{1}{\lambda} = -\frac{1}{\lambda(n+1)}$ et donc

$$r(T_n) = \frac{1}{(n+1)^2\lambda^2} + V(T_n) = \frac{1}{(n+1)^2\lambda^2} + \frac{n}{(n+1)^2\lambda^2} = \frac{1}{(n+1)\lambda^2}.$$

Et donc, bien que T_n soit biaisé, contrairement à \overline{X}_n , il possède un risque quadratique plus faible.

SOLUTION DE L'EXERCICE 21.28

1. On a

$$\begin{aligned} \tilde{S}_n &= \frac{1}{n} \sum_{i=1}^n (X_i^2 - 2X_i\overline{X}_n + \overline{X}_n^2) \\ &= \frac{1}{n} \sum_{i=1}^n X_i^2 - 2\overline{X}_n \underbrace{\frac{1}{n} \sum_{i=1}^n X_i}_{=\overline{X}_n} + \frac{1}{n} \sum_{i=1}^n \overline{X}_n^2 \\ &= \frac{1}{n} \sum_{i=1}^n X_i^2 - 2\overline{X}_n^2 + \overline{X}_n^2 \\ &= \frac{1}{n} \sum_{i=1}^n X_i^2 - \overline{X}_n^2. \end{aligned}$$

2. Les X_i^2 sont i.i.d., possèdent une variance²⁰, et pour tout i , $E(X_i^2) = V(X_i) + E(X_i)^2 = \sigma^2 + m^2$.

Et donc par la loi faible des grands nombres, $\frac{1}{n} \sum_{i=1}^n X_i^2 \xrightarrow{P} \sigma^2 + m^2$.

D'autre part, toujours par la loi faible des grands nombres, $\overline{X}_n \xrightarrow{P} m$.

Et puisque la fonction $t \mapsto t^2$ est continue, $\overline{X}_n^2 \xrightarrow{P} m^2$.

Grâce au résultat admis²¹,

$$\tilde{S}_n = \frac{1}{n} \sum_{i=1}^n X_i^2 - \overline{X}_n^2 \xrightarrow{P} \sigma^2 + m^2 - m^2 = \sigma^2.$$

3. Puisque les X_i sont i.i.d. et possèdent une variance, le théorème central limite s'applique.

Puisqu'on a $E(\overline{X}_n) = m$ et $V(\overline{X}_n) = \frac{\sigma^2}{n}$, il vient donc

$$\sqrt{n} \frac{\overline{X}_n - m}{\sigma} \xrightarrow{\mathcal{L}} X \text{ où } X \leftrightarrow \mathcal{N}(0, 1).$$

²⁰ Car les X_i possèdent un moment d'ordre 4.

²¹ Ce résultat est le dernier lemme du chapitre 18.

Moments

Rappelons que si le moment d'ordre r existe, alors tous les moments d'ordre p , $p \leq r$ existent.

Ici, les X_i possèdent un moment d'ordre 4 et donc un moment d'ordre 2.

Or, $\tilde{S}_n \xrightarrow{\mathcal{L}} \sigma^2$, et donc, par composition par la fonction continue $t \mapsto \sqrt{t}$, $\sqrt{\tilde{S}_n} \xrightarrow{\mathcal{L}} \sqrt{\sigma^2} = \sigma$.

Et donc $\frac{\tilde{S}_n}{\sigma} \xrightarrow{P} 1$, de sorte que²² $\frac{\sigma}{\sqrt{\tilde{S}_n}} \xrightarrow{P} 1$.

Par le théorème de Slutsky, il vient

$$\frac{\sqrt{n}(\bar{X}_n - m)}{\sqrt{\tilde{S}_n}} = \sqrt{n} \frac{\bar{X}_n - m}{\sigma} \frac{\sigma}{\sqrt{\tilde{S}_n}} \xrightarrow{\mathcal{L}} 1 \times X = X.$$

SOLUTION DE L'EXERCICE 21.29

Première méthode : à l'aide de l'inégalité de Bienaymé-Tchebychev.

Notons que F_n et P_n sont liées par la relation $P_n = n - F_n$.

De plus, P_n suit la loi binomiale $\mathcal{B}(n, p)$.

Ainsi, $\frac{F_n - P_n}{n} = \frac{2F_n - n}{n} = 2\frac{F_n}{n} - 1$, et donc

$$E\left(\frac{F_n - P_n}{n}\right) = 2p - 1 \text{ et } V\left(\frac{F_n - P_n}{n}\right) = \frac{4}{n^2} V(F_n) = \frac{4p(1-p)}{n}.$$

D'après l'inégalité de Bienaymé-Tchebychev, pour tout $\varepsilon > 0$, on a

$$P\left(\left|\frac{F_n - P_n}{n} - (2p - 1)\right| \geq \varepsilon\right) \leq \frac{4p(1-p)}{n\varepsilon^2} \xrightarrow{n \rightarrow +\infty} 0.$$

Ainsi, $\left(\frac{F_n - P_n}{n}\right)$ converge en probabilité vers la variable certaine égale à $2p - 1$.

Seconde méthode : via la loi faible des grands nombres.

Pour $i \in \mathbf{N}^*$, notons X_i la variable égale à 1 si le $i^{\text{ème}}$ lancer donne face et -1 dans le cas contraire.

Alors les X_i sont indépendantes, et suivent une même loi, qui vérifie $E(X_i) = p - (1 - p) = 2p - 1$, et qui admet une variance car il s'agit de variables à support fini.

Or, $\frac{F_n - P_n}{n} = \frac{X_1 + \dots + X_n}{n}$.

Par la loi faible des grands nombres, $\left(\frac{X_1 + \dots + X_n}{n}\right)$ converge en probabilité vers la variable certaine égale à $E(X_1) = 2p - 1$.

SOLUTION DE L'EXERCICE 21.30

1. Par le théorème de transfert, on a, sous réserve de convergence²³

$$\begin{aligned} E(e^{tX}) &= \sum_{n=0}^{+\infty} e^{tn} P(X = n) \\ &= \sum_{n=0}^{+\infty} e^{tn} \frac{\lambda^n}{n!} e^{-\lambda} \\ &= e^{-\lambda} \sum_{n=0}^{+\infty} (e^t)^n \frac{\lambda^n}{n!} \\ &= e^{-\lambda} \exp(e^t \lambda) \\ &= \exp(\lambda(e^t - 1)). \end{aligned}$$

2. Notons que $[X \geq n] = [e^{tX} \geq e^{tn}]$ et donc par l'inégalité de Markov appliquée à e^{tX} , qui est bien positive et admet une espérance,

$$P(X \geq n) = P(e^{tX} \geq e^{tn}) \leq \frac{E(e^{tX})}{e^{tn}} \leq e^{-tn} E(e^{tX}).$$

3. En utilisant ce qui précède, il vient

$$P(X \geq n) \leq e^{-tn} \exp(\lambda(e^t - 1)).$$

²² Par composition par la fonction continue $t \mapsto \frac{1}{t}$.

Remarque
Notons que F_n suit aussi une loi binomiale.

²³ Absolue, mais il s'agit d'une série à termes positifs.

On a reconnu une série exponentielle, donc convergente, donc $E(e^{tX})$ existe.

Ceci étant vrai pour tout $t \geq 0$, cherchons la valeur de t qui minimise

$$e^{-tn} \exp(\lambda(e^t - 1)) = \exp(-tn + \lambda(e^t - 1)).$$

Soit donc $f(t) = -tn + \lambda(e^t - 1)$. Alors f est dérivable, et $f'(t) = -n + \lambda e^t$, de sorte que $f'(t) = 0 \Leftrightarrow t = \ln\left(\frac{n}{\lambda}\right)$.

Et donc f est décroissante sur $\left[0, \ln\left(\frac{n}{\lambda}\right)\right]$ et croissante sur $\left[\ln\left(\frac{n}{\lambda}\right), +\infty\right[$.

Elle admet donc un minimum en $\ln\left(\frac{n}{\lambda}\right)$, qui vaut $f\left(\ln\left(\frac{n}{\lambda}\right)\right) = -n \ln\left(\frac{n}{\lambda}\right) + \lambda\left(\frac{n}{\lambda} - 1\right) = -n \ln\left(\frac{n}{\lambda}\right) + n - \lambda$.

On en déduit, pour $t = \ln\left(\frac{n}{\lambda}\right)$ que

$$P(X \geq n) \leq \exp\left(-n \ln\left(\frac{n}{\lambda}\right) + n - \lambda\right) \leq \left(\frac{\lambda}{n}\right)^n e^n e^{-\lambda}.$$

Mais puisque d'autre part, X est à valeurs entières, on a

$$P(X \geq n) = \sum_{k=n}^{+\infty} P(X = k) = e^{-\lambda} \sum_{k=n}^{+\infty} \frac{\lambda^k}{k!}.$$

Et donc il vient, après division par $e^{-\lambda}$,

$$\sum_{k=n}^{+\infty} \frac{\lambda^k}{k!} \leq \left(\frac{\lambda}{n}\right)^n e^n.$$

EML 2018

La partie I a déjà été traitée dans un DS, et n'est pas notre priorité aujourd'hui.

Vous pourrez donc vous contenter d'en lire l'énoncé et d'admettre ses résultats (qui ne servent qu'aux questions 7.d et 9.c), puis, seulement s'il vous reste du temps, vous reviendrez traiter cette première partie.

Soit n un entier naturel supérieur ou égal à 2.

On note $E = \mathbf{R}_n[X]$ l'espace vectoriel des polynômes à coefficients réels de degré inférieur ou égal à n , et $\mathcal{B} = (1, X, \dots, X^n)$ la base canonique de E .

On note, pour tout polynôme P de E :

$$\varphi(P) = \frac{1}{n}X(1-X)P' + XP.$$

Partie I : Étude d'un endomorphisme de polynômes

1.
 - a. Montrer que φ est une application linéaire.
 - b. Calculer $\varphi(X^n)$.
 - c. Montrer que φ est un endomorphisme de E .
2. Déterminer la matrice A de φ dans la base \mathcal{B} . Préciser le rang de cette matrice.
3.
 - a. L'endomorphisme φ est-il injectif ? Justifier votre réponse.
 - b. Soit P un polynôme non nul de $\text{Ker}(\varphi)$.
Montrer que P admet 1 comme unique racine (dans \mathbf{C}), et que P est de degré n .
 - c. En déduire une base de $\text{Ker}(\varphi)$.
4. Montrer que φ est diagonalisable.
5. On pose, pour tout k de $\llbracket 0, n \rrbracket$: $P_k = X^k(1-X)^{n-k}$.
 - a. Pour tout k de $\llbracket 0, n \rrbracket$, calculer $\varphi(P_k)$.
 - b. Montrer que la famille (P_0, P_1, \dots, P_n) est une base de E et expliciter la matrice de φ dans cette base.
 - c. Déterminer les sous-espaces propres de φ .

Partie II : Étude d'une suite de variables aléatoires.

On considère une urne contenant n boules numérotées de 1 à n , indiscernables au toucher. On effectue dans cette urne une suite de tirages avec remise, et on suppose que l'expérience est modélisée par un espace probabilisé (Ω, \mathcal{A}, P) .

On note alors, pour tout k de \mathbf{N}^* , Y_k la variable aléatoire prenant pour valeur le nombre de numéros distincts qui ont été tirés lors des k premiers tirages.

Par convention, on pose : $Y_0 = 0$.

6. On note, pour tout k de \mathbf{N}^* , Z_k la variable aléatoire prenant la valeur 1 si le k -ième tirage amène un numéro qui n'a pas été tiré lors des tirages précédents, et prenant la valeur 0 sinon.

On pourra remarquer que, en particulier, $Z_1 = 1$.

- a. Déterminer la loi de Z_2 .
- b. Soit $k \in \mathbf{N}^*$. Calculer, pour tout j de $\llbracket 1, k \rrbracket$, la valeur de $P_{[Y_k=j]}([Z_{k+1} = 1])$.
En déduire : $P([Z_{k+1} = 1]) = 1 - \frac{1}{n}E(Y_k)$.

- c. Soit $k \in \mathbf{N}^*$. En remarquant que $Y_k = \sum_{j=1}^k Z_j$, montrer :

$$P([Z_{k+1} = 1]) = 1 - \frac{1}{n} \sum_{j=1}^k P([Z_j = 1]).$$

- d. En déduire, pour tout k de \mathbf{N}^* : $P([Z_k = 1]) = \left(1 - \frac{1}{n}\right)^{k-1}$.

- e. Déterminer alors, pour tout k de \mathbf{N} , l'espérance de Y_k .

7. On note, pour tout k de \mathbf{N} , G_k le polynôme de $\mathbf{R}_n[X]$ défini par :

$$G_k = \sum_{i=0}^n P([Y_k = i])X^i.$$

- a. Déterminer les polynômes G_0, G_1 et G_2 .
 b. Montrer, pour tout k de \mathbf{N} et tout i de $\llbracket 0, n \rrbracket$:

$$P([Y_{k+1} = i]) = \frac{i}{n}P([Y_k = i]) + \left(1 - \frac{i-1}{n}\right)P([Y_k = i-1]).$$

- c. Montrer, pour tout k de \mathbf{N} :

$$G_{k+1} = \frac{1}{n}X(1-X)G'_k + XG_k.$$

- d. En déduire, pour tout k de \mathbf{N} :

$$G_k = \varphi^k(G_0).$$

8. a. Pour tout k de \mathbf{N} , calculer $G_k(1)$ et $G'_k(1)$.

- b. En déduire, pour tout k de \mathbf{N} :

$$E(Y_{k+1}) = \left(1 - \frac{1}{n}\right)E(Y_k) + 1.$$

- c. Retrouver alors, pour tout k de \mathbf{N} , l'expression de $E(Y_k)$ obtenue en question 6.e.

9. On rappelle que les polynômes P_0, \dots, P_n sont définis à la question 5 par :

$$\text{pour tout } j \text{ de } \llbracket 0, n \rrbracket, \quad P_j = X^j(1-X)^{n-j}.$$

- a. Calculer $\sum_{j=0}^n \binom{n}{j} P_j$.

- b. Montrer, pour tout j de $\llbracket 0, n \rrbracket$:

$$P_j = \sum_{i=j}^n \binom{n-j}{i-j} (-1)^{i-j} X^i.$$

- c. En déduire, pour tout k de \mathbf{N} :

$$\varphi^k(G_0) = \sum_{i=0}^n \left(\sum_{j=0}^i \binom{n}{j} \binom{n-j}{i-j} \left(\frac{j}{n}\right)^k (-1)^{i-j} \right) X^i.$$

- d. Montrer finalement, pour tout k de \mathbf{N} et pour tout i de $\llbracket 0, n \rrbracket$:

$$P([Y_k = i]) = \binom{n}{i} \sum_{j=0}^i \binom{i}{j} (-1)^{i-j} \left(\frac{j}{n}\right)^k.$$

EML 2018 : CORRECTION

Partie I : Étude d'un endomorphisme de polynômes.

1.a. Soient $P, Q \in \mathbf{R}_n[X]$ et $\lambda \in \mathbf{R}$. Alors, par linéarité de la dérivation,

$$\begin{aligned}\varphi(\lambda P + Q) &= \frac{1}{n}X(1-X)(\lambda P + Q)' + X(\lambda P + Q) \\ &= \frac{1}{n}X(1-X)(\lambda P' + Q') + X(\lambda P + Q) \\ &= \lambda \left(\frac{1}{n}X(1-X)P' + XP \right) + \left(\frac{1}{n}X(1-X)Q' + XQ \right) \\ &= \lambda\varphi(P) + \varphi(Q).\end{aligned}$$

Donc φ est une application linéaire de $\mathbf{R}_n[X]$ dans $\mathbf{R}[X]$.

1.b. On a

$$\varphi(X^n) = \frac{1}{n}X(1-X)nX^{n-1} + XX^n = (1-X)X^n + X^{n+1} = X^n.$$

1.c. Pour $Q \in \mathbf{R}_n[X]$, on a $\deg Q' \leq \deg Q - 1$ et donc

$$\deg(\varphi(Q)) \leq \max(\deg(Q') + \deg(X(1-X)), \deg(Q) + 1) \leq \deg(Q) + 1.$$

Or, si $P \in \mathbf{R}_n[X]$, alors en séparant le terme de degré n des autres termes, on constate qu'il existe un réel a et un polynôme $Q \in \mathbf{R}_{n-1}[X]$ tels que $P = aX^n + Q$.

Et par linéarité de φ , on a alors

$$\varphi(P) = a\varphi(X^n) + \varphi(Q) = aX^n + \varphi(Q).$$

Or, $\deg \varphi(Q) \leq \deg(Q) + 1 \leq n$, de sorte que $\varphi(P) \in \mathbf{R}_n[X]$.

Ainsi, φ est à valeurs dans E , et étant linéaire, c'est un endomorphisme de E .

2. On a $\varphi(1) = X$, $\varphi(X^n) = X^n$ et pour $k \in \llbracket 1, n-1 \rrbracket$,

$$\varphi(X^k) = \frac{k}{n}X(1-X)X^{k-1} + X^{k+1} = \frac{k}{n}X^k - \frac{k}{n}X^{k+1} + X^{k+1} = \frac{k}{n}X^k + \frac{n-k}{n}X^{k+1}.$$

Et donc la matrice A est donnée par

$$A = \text{Mat}_{\mathcal{B}}(\varphi) = \begin{pmatrix} \varphi(1) & \varphi(X) & \varphi(X^2) & \dots & \varphi(X^{n-1}) & \varphi(X^n) \\ 0 & 0 & \dots & \dots & \dots & 0 \\ 1 & \frac{1}{n} & 0 & \dots & \dots & 0 \\ 0 & \frac{n-1}{n} & \frac{2}{n} & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & \frac{n-1}{n} & 0 \\ 0 & \dots & \dots & 0 & 1 & 1 \end{pmatrix} \begin{matrix} 1 \\ X \\ X^2 \\ \vdots \\ X^{n-1} \\ X^n \end{matrix} \in \mathcal{M}_{n+1}(\mathbf{R}).$$

Puisque la première ligne de A est nulle, elle est de rang au plus n .

Puisque d'autre part elle est triangulaire¹ et possède n coefficients diagonaux non nuls, elle est de rang au moins n .

Et donc $\text{rg}(A) = 1$.

3.a. Puisque A est de rang $n < n+1$, elle n'est pas inversible.

Donc φ n'est pas un isomorphisme, et donc n'est pas injectif.

3.b. Soit $P \in \text{Ker } \varphi$. Alors $\frac{1}{n}X(1-X)P' = -XP \Leftrightarrow \frac{1}{n}(1-X)P' = -P$.

Soit a une racine complexe de P . Alors $P(a) = 0$ et donc $\frac{1}{n}(1-a)P'(a) = 0$, de sorte que $P'(a) = 0$ ou $a = 1$.

Degré

Notons que pour l'instant, nous ne pouvons rien dire du degré de $\varphi(P)$, même si on prouverait sans grandes difficultés que φ est à valeurs dans $\mathbf{R}_{n+1}[X]$.

Pour ne pas trop nous avancer, nous nous contentons pour l'instant de dire que φ est à valeurs dans $\mathbf{R}[X]$.

Notons que la question 1.c prouvera que φ est à valeurs dans $\mathbf{R}_n[X]$.

⚠ Danger !

Tout ce qu'on peut dire en toutes généralités sur le degré d'une somme est que

$$\begin{aligned}\deg(P + Q) \\ \leq \max(\deg P, \deg Q).\end{aligned}$$

Cette inégalité n'est pas nécessairement une égalité.

¹ Inférieure.

Rappel

En dimension finie, un endomorphisme est injectif si et seulement si il est bijectif.

Supposons que $a \neq 1$, et notons $m \geq 1$ la multiplicité de a comme racine de P .

Alors a est racine de P' de multiplicité $m - 1$.

Et donc a est racine de $\frac{1}{n}(1 - X)P'$ de multiplicité $m - 1$.

Mais $\frac{1}{n}(1 - X)P' = P$, et donc a en est racine de multiplicité m , d'où une contradiction.

Ainsi, $\boxed{1 \text{ est la seule racine de } P.}$

Soit $P \in \text{Ker}(\varphi)$ non nul, et notons d le degré de P , de sorte que $P = \sum_{k=0}^d a_k X^k$, avec $a_d \neq 0$.

Alors le coefficient en X^{d+1} de $\varphi(P)$ est $-\frac{d}{n}a_d + a_d$, qui est nul si et seulement si $d = n$.

Et donc si $\varphi(P) = 0$, nécessairement $d = n$: $\boxed{\text{tout polynôme non nul de } \text{Ker}(\varphi) \text{ est de degré égal à } n.}$

3.c. Nous savons que tout polynôme de $\mathbb{C}[X]$ se factorise en produit de termes de degré un. Or, un élément non nul de $\text{Ker}(\varphi)$ est nécessairement de degré n , donc produit de n termes de degré un.

Et puisque 1 en est la seule racine, nécessairement si $P \in \text{Ker}(\varphi)$, alors il existe $\lambda \in \mathbb{C}$ tel que $P = \lambda(X - 1)^n$. Et donc $\boxed{\text{Ker}(\varphi) = \text{Vect}((X - 1)^n)}$.

4. La matrice A est triangulaire inférieure, et donc ses valeurs propres sont ses coefficients diagonaux.

Ainsi, $\text{Spec}(A) = \left\{ \frac{k}{n}, k \in \llbracket 0, n \rrbracket \right\}$. Par conséquent, A possède $n+1$ valeurs propres distinctes, et donc est diagonalisable.

Et donc $\boxed{\varphi \text{ est également diagonalisable.}}$

5.a. Pour $k \in \llbracket 1, n - 1 \rrbracket$, on a $P'_k = kX^{k-1}(1 - X)^{n-k} - (n - k)X^k(1 - X)^{n-k-1}$. Et donc

$$\begin{aligned} \varphi(P_k) &= \frac{1}{n}X(1 - X) \left(kX^{k-1}(1 - X)^{n-k} - (n - k)X^k(1 - X)^{n-k-1} \right) + X^{k+1}(1 - X)^{n-k} \\ &= \frac{k}{n}X^k(1 - X)^{n-k+1} - \frac{n - k}{n}X^{k+1}(1 - X)^{n-k} + X^{k+1}(1 - X)^{n-k} = \frac{k}{n}X^k(1 - X)^{n-k} (X + 1 - X) \\ &= \boxed{\frac{k}{n}P_k}. \end{aligned}$$

D'autre part, on a $P_0 = (1 - X)^n \in \text{Ker } \varphi$ et donc $\boxed{\varphi(P_0) = 0}$.

Et enfin $P_n = X^n$ et donc $\boxed{\varphi(P_n) = X^n = P_n}$.

5.b. Nous venons de prouver que pour tout $k \in \llbracket 0, n \rrbracket$, P_k est un vecteur propre de φ pour la valeur propre $\frac{k}{n}$.

Donc (P_0, P_1, \dots, P_n) est une famille de vecteurs propres associés à des valeurs propres distinctes : elle est libre. Étant de cardinal $n + 1 = \dim E$, c'est une $\boxed{\text{base de } E}$.

Et puisqu'il s'agit d'une base de vecteurs propres de φ , la matrice de φ dans cette base est diagonale :

$$\boxed{\text{Mat}_{(P_0, \dots, P_n)}(\varphi) = \text{Diag} \left(0, \frac{1}{n}, \frac{2}{n}, \dots, \frac{n-1}{n}, 1 \right)}.$$

5.c. Puisque φ possède $n + 1 = \dim E$ valeurs propres distinctes, ses sous-espaces propres sont tous de dimension 1.

Et puisque pour tout $k \in \llbracket 0, n \rrbracket$, P_k est un vecteur propre de φ associé à la valeur propre $\frac{k}{n}$, c'est une base de $E_{\frac{k}{n}}(\varphi)$.

Ainsi, $\boxed{\text{pour tout } k \in \llbracket 0, n \rrbracket, E_{\frac{k}{n}}(\varphi) = \text{Vect}(P_k)}$.

Partie II : Étude d'une suite de variables aléatoires.

6.a. La variable Z_2 prend la valeur 0 si et seulement si les deux premiers tirages ont donné deux fois le même numéro.

Remarque

Si $m = 1$, a n'est pas racine de P' .

Degré

Lorsqu'on dit d'un polynôme qu'il est de degré d , cela signifie précisément qu'il possède un terme de degré d non nul, et aucun terme de degré plus élevé.

Réels/complexes

Nous avons ici considéré P comme un polynôme complexe.

Mais puisque nous savons qu'il s'agit d'un polynôme à coefficients réels, nous pouvons en réalité affirmer que $\lambda \in \mathbb{R}$.

Remarque

Nous traitons à part les cas $k = 0$ et $k = n$ pour éviter d'écrire des puissances négatives.

Or, il y a n^2 tirages possibles pour les deux premières boules, dont n comprenant deux fois la même boule.

$$\text{Donc } P(Z_2 = 0) = \frac{n}{n^2} = \frac{1}{n}.$$

Et par conséquent²,

$$P(Z_2 = 1) = 1 - P(Z_2 = 0) = 1 - \frac{1}{n}.$$

² Z_2 ne prend que les valeurs 0 et 1.

Ainsi, Z_2 suit la loi de Bernoulli de paramètre $1 - \frac{1}{n}$.

- 6.b. Sachant que $[Y_k = j]$, les k premiers tirages ont fait apparaître j boules différentes. Et donc pour que le $(k + 1)^{\text{ème}}$ tirage fasse apparaître une nouvelle boule, il faut que celle-ci soit tirée parmi les $n - j$ boules qui ne sont pas apparues précédemment.

$$\text{Et donc } P_{[Y_k=j]}(Z_{k+1} = 1) = \frac{n-j}{n} = 1 - \frac{j}{n}.$$

Par la formule des probabilités totales appliquée au système complet d'événements $\{[Y_k = j], j \in \llbracket 1, k \rrbracket\}$, il vient

$$\begin{aligned} P(Z_{k+1} = 1) &= \sum_{j=1}^k P(Y_k = j)P_{[Y_k=j]}(Z_{k+1} = 1) \\ &= \sum_{j=1}^k \left(1 - \frac{j}{n}\right) P(Y_k = j) = \sum_{j=1}^k P(Y_k = j) - \frac{1}{n} \sum_{j=1}^k jP(Y_k = j) \\ &= 1 - \frac{1}{n} E(Y_k). \end{aligned}$$

- 6.c. Comme indiqué dans l'énoncé, on a bien $Y_k = \sum_{j=1}^k Z_j$, puisque $\sum_{j=1}^k Z_j$ est égal au nombre de fois où l'on a obtenu une boule qui n'était pas sortie précédemment : c'est donc le nombre de boules différentes apparues lors des k premiers tirages.

Par linéarité de l'espérance, on a alors $E(Y_k) = \sum_{j=1}^k E(Z_j)$.

Mais les Z_j suivent des lois de Bernoulli, donc pour tout $j \in \llbracket 1, k \rrbracket$, $E(Z_j) = P(Z_j = 1)$.
Et donc

$$P(Z_{k+1} = 1) = 1 - \frac{1}{n} E(Y_k) = 1 - \frac{1}{n} \sum_{j=1}^k E(Z_j) = \boxed{1 - \frac{1}{n} \sum_{j=1}^k P(Z_j = 1)}.$$

- 6.d. Prouvons le résultat par récurrence forte sur k .

Pour $k = 1$, c'est évident car $P(Z_1 = 1) = 1 = \left(1 - \frac{1}{n}\right)^0$.

Supposons donc que pour tout $j \in \llbracket 1, k \rrbracket$, $P(Z_j = 1) = \left(1 - \frac{1}{n}\right)^{j-1}$. Alors

$$\begin{aligned} P(Z_{k+1} = 1) &= 1 - \frac{1}{n} \sum_{j=1}^k P(Z_j = 1) \\ &= 1 - \frac{1}{n} \sum_{j=1}^k \left(1 - \frac{1}{n}\right)^{j-1} \\ &= 1 - \frac{1}{n} \sum_{i=0}^{k-1} \left(1 - \frac{1}{n}\right)^i \\ &= 1 - \frac{1}{n} \frac{1 - \left(1 - \frac{1}{n}\right)^k}{1 - \left(1 - \frac{1}{n}\right)} \\ &= 1 - \left(1 - \left(1 - \frac{1}{n}\right)^k\right) \end{aligned}$$

Récurrence forte

Une récurrence forte s'utilise quand pour prouver la propriété au rang $n + 1$, nous n'avons pas uniquement besoin de savoir qu'elle est vraie au rang n , mais à tous les rangs $1, 2, \dots, n$.

Chgt d'indice

$i = j - 1$.

Somme des termes consécutifs d'une suite géométrique de raison $1 - \frac{1}{n} \neq 1$.

$$= \left(1 - \frac{1}{n}\right)^k.$$

Par le principe de récurrence forte, pour tout $k \in \mathbf{N}$, $P(Z_k = 1) = \left(1 - \frac{1}{n}\right)^{k-1}$.

6.e. La relation prouvée à la question 6.b nous donne

$$E(Y_k) = n - nP(Z_{k+1} = 1) = n \left(1 - \left(1 - \frac{1}{n}\right)^k\right).$$

7.a. Puisque Y_0 est certaine égale à 0 et que Y_1 est certaine égale à 1, on a $G_0 = P(Y_0 = 0) = 1$ et $G_1 = P(Y_1 = 1)X = X$. De plus

$$G_2 = P(Y_2 = 1)X + P(Y_2 = 2)X^2 = P(Y_2 = 1)X + (1 - P(Y_2 = 1))X^2 = \frac{1}{n}X + \frac{n-1}{n}X^2.$$

7.b. D'après la formule des probabilités totales appliquée au système complet d'événements $\{[Y_k = j], j \in \llbracket 1, k \rrbracket\}$, on a

$$P(Y_{k+1} = i) = \sum_{j=1}^k P([Y_{k+1} = i] \cap [Y_k = j]).$$

Or à chaque tirage, nous avons au plus un nouveau numéro, de sorte que Y_{k+1} ne peut être égal qu'à Y_k ou à $Y_k + 1$.

Autrement dit, si $j \notin \{i, i-1\}$, $P([Y_{k+1} = i] \cap [Y_k = j]) = 0$.

Et donc

$$\begin{aligned} P(Y_{k+1} = i) &= P([Y_{k+1} = i] \cap [Y_k = i]) + P([Y_{k+1} = i] \cap [Y_k = i-1]) \\ &= P(Y_k = i)P_{[Y_k=i]}(Y_{k+1} = i) + P(Y_k = i-1)P_{[Y_k=i-1]}(Y_k = i) \\ &= \frac{i}{n}P(Y_k = i) + \left(1 - \frac{i-1}{n}\right)P(Y_k = i-1). \end{aligned}$$

7.c. En utilisant le résultat de la question précédente, il vient

$$\begin{aligned} G_{k+1} &= \sum_{i=0}^n P(Y_{k+1} = i)X^i = \sum_{i=0}^n \left(\frac{i}{n}P(Y_k = i) + \left(1 - \frac{i-1}{n}\right)P(Y_k = i-1)\right)X^i \\ &= \sum_{i=0}^n \frac{i}{n}P(Y_k = i)X^i + \sum_{i=1}^{n+1} \left(1 - \frac{i-1}{n}\right)P(Y_k = i-1)X^i \\ &= \frac{1}{n} \sum_{i=0}^n iP(Y_k = i)X^i + \sum_{i=1}^n \left(1 - \frac{i}{n}\right)P(Y_k = i)X^{i+1} \\ &= \frac{1}{n} \sum_{i=0}^n iP(Y_k = i)X^i + X \sum_{i=1}^n P(Y_k = i)X^i - \frac{1}{n} \sum_{i=1}^n iP(Y_k = i)X^{i+1} \\ &= \frac{X}{n} \sum_{i=1}^n iP(Y_k = i)X^{i-1} + XG_k - \frac{X^2}{n} \sum_{i=1}^n iP(Y_k = i)X^{i-1} \\ &= \frac{1}{n}X(1-X) \sum_{i=1}^n iP(Y_k = i)X^{i-1} + XG_k \\ &= \frac{1}{n}X(1-X)G'_k + XG_k. \end{aligned}$$

7.d. Nous venons de prouver que $G_{k+1} = \varphi(G_k)$.

Une récurrence rapide prouve alors que pour tout $k \in \mathbf{N}$, $G_k = \varphi^k(G_0)$.

Détails

$\frac{i}{n}$ est la probabilité que, ayant déjà tiré i numéros distincts, on tire de nouveau l'un de ces numéros. Et au contraire, $1 - \frac{i-1}{n}$ est la probabilité de tirer un numéro distinct des $i-1$ déjà sortis.

Astuce

Pour $i = n+1$, le terme $1 - \frac{i-1}{n}$ est nul, donc on ne change rien en l'ajoutant à la somme. Et le terme en $i=0$ est nul car $P(Y_k = -1) = 0$. Donc on peut le retirer de la somme.

8.a. On a

$$G_k(1) = \sum_{i=0}^n P(Y_k = i) \times 1^k = \sum_{i=0}^n P(Y_k = i) = \boxed{1.}$$

Et de même,

$$G'_k(1) = \sum_{i=1}^n iP(Y_k = i)1^{i-1} = \sum_{i=1}^n iP(Y_k = i) = \boxed{E(Y_k).}$$

8.b. On a, en dérivant la relation de la question 7.c,

$$G'_{k+1} = \frac{1}{n}X(1-X)G'_k + \frac{1}{n}(1-X)G'_k - \frac{1}{n}XG'_k + G_k + XG'_k$$

et donc

$$E(Y_{k+1}) = G'_{k+1}(1) = -\frac{1}{n}G'_k(1) + G_k(1) + G'_k(1) = \left(1 - \frac{1}{n}\right)G'_k(1) + 1 = \boxed{\left(1 - \frac{1}{n}\right)E(Y_k) + 1.}$$

8.c. Nous venons donc de prouver que la suite $(E(Y_k))_{k \geq 0}$ est une suite arithmético-géométrique de raison $1 - \frac{1}{n}$.

Et donc nous savons qu'il existe $\lambda \in \mathbf{R}$ tel que, en posant $u_k = E(Y_k) + \lambda$, alors (u_k) est géométrique de raison $1 - \frac{1}{n}$.

Or, pour tout $k \in \mathbf{N}$,

$$\begin{aligned} u_{k+1} &= \left(1 - \frac{1}{n}\right)u_k \Leftrightarrow E(Y_{k+1}) + \lambda = \left(1 - \frac{1}{n}\right)(E(Y_k) + \lambda) \\ &\Leftrightarrow \left(1 - \frac{1}{n}\right)E(Y_k) + 1 + \lambda = \left(1 - \frac{1}{n}\right)E(Y_k) + \left(1 - \frac{1}{n}\right)\lambda \\ &\Leftrightarrow \lambda = -n. \end{aligned}$$

Autrement dit, en posant $u_k = E(Y_k) - n$, alors $(u_k)_k$ est géométrique de raison $1 - \frac{1}{n}$, et

donc $\forall k \in \mathbf{N}$, $u_k = \left(1 - \frac{1}{n}\right)^k u_0$.

Mais $u_0 = E(Y_0) - n = -n$ de sorte que

$$E(Y_k) = u_k + n = -n \left(1 - \frac{1}{n}\right)^k + n = \boxed{n \left(1 - \left(1 - \frac{1}{n}\right)^k\right).}$$

9.a. On a, d'après la formule du binôme de Newton,

$$\sum_{j=0}^n \binom{n}{j} P_j = \sum_{j=0}^n \binom{n}{j} X^j (1-X)^{n-j} = (X + (1-X))^n = 1^n = \boxed{1.}$$

9.b. Utilisons de nouveau la formule du binôme :

$$P_j = X^j (1-X)^{n-j} = X^j \sum_{k=0}^{n-j} \binom{n-j}{k} (-X)^k = X^j \sum_{i=j}^n \binom{n-j}{i-j} (-1)^{i-j} X^{i-j} = \boxed{\sum_{i=j}^n \binom{n-j}{i-j} (-1)^{i-j} X^i.}$$

9.c. D'après la question 9.a, $G_0 = \sum_{j=0}^n \binom{n}{j} P_j$.

Or, P_j étant un vecteur propre de φ pour la valeur propre $\frac{j}{n}$, on en déduit que pour tout

$k \in \mathbf{N}$, $\varphi^k(P_j) = \left(\frac{j}{n}\right)^k P_j$.

Et donc, par linéarité de φ^k ,

$$\varphi^k(G_0) = \varphi^k \left(\sum_{j=0}^n \binom{n}{j} P_j \right) = \sum_{j=0}^n \binom{n}{j} \varphi^k(P_j)$$

Méthode

Si vous connaissez directement la formule donnant λ (c'est $\lambda = \frac{b}{1-a}$ pour une suite vérifiant $u_{n+1} = au_n + b$), vous pouvez l'utiliser.

Mais si vous voulez vous éviter d'apprendre une formule de plus, il suffit de savoir qu'un tel λ existe et d'être capable de le retrouver.

Chgt d'indice

$i - j = k \Leftrightarrow i = k + j$ et donc si k varie de 0 à $n - j$, alors i varie de j à n .

$$\begin{aligned}
 &= \sum_{j=0}^n \binom{n}{j} \left(\frac{j}{n}\right)^k P_j = \sum_{j=0}^n \binom{n}{j} \left(\frac{j}{n}\right)^k \sum_{i=j}^n \binom{n-j}{i-j} (-1)^{i-j} X^i \\
 &= \boxed{\sum_{i=0}^n \sum_{j=0}^i \binom{j}{n} \binom{i-j}{n-j} \left(\frac{j}{n}\right)^k (-1)^{i-j} X^i}.
 \end{aligned}$$

9.d. Puisque $G_k = \varphi^k(G_0)$, en identifiant³ le coefficient de degré i , il vient

$$\begin{aligned}
 P(Y_k = i) &= \sum_{j=0}^i \binom{n}{j} \binom{n-j}{i-j} \left(\frac{j}{n}\right)^k (-1)^{i-j} \\
 &= \sum_{j=0}^i \frac{n!}{j!(n-j)!} \frac{(n-j)!}{(i-j)!(n-i)!} \left(\frac{j}{n}\right)^k (-1)^{i-j} \\
 &= \sum_{j=0}^i \frac{n!}{i!(n-i)! j!(i-j)!} \left(\frac{j}{n}\right)^k (-1)^{i-j} \\
 &= \boxed{\binom{n}{i} \sum_{j=0}^i \binom{i}{j} (-1)^{i-j} \left(\frac{j}{n}\right)^k}.
 \end{aligned}$$

³ Il y a unicité des coefficients d'un polynôme.

Remarque

Notons que de manière surprenante, cette somme est nulle pour $i > k$, car $P(Y_k = i)$ est évidemment nul.

Remarque : le résultat obtenu permet facilement d'en déduire une formule pour le rang du premier tirage où l'on a obtenu les n numéros. En effet, si on note X le rang de ce tirage, on a

$$\begin{aligned}
 P(X = k) &= P(Y_k = n) - P(Y_{k-1} = n) \\
 &= \sum_{j=0}^n \binom{n}{j} (-1)^{n-j} \left(\frac{j}{n}\right)^k - \sum_{j=0}^n \binom{n}{j} (-1)^{n-j} \left(\frac{j}{n}\right)^{k-1} \\
 &= \left(\frac{j}{n} - 1\right) \sum_{j=0}^n \binom{n}{j} (-1)^{n-j} \left(\frac{j}{n}\right)^{k-1}.
 \end{aligned}$$

Détails

Il faut qu'on ait eu la $n^{\text{ème}}$ boule lors du $k^{\text{ème}}$ tirage, et donc qu'on n'en ait eu que $n-1$ à l'issue du $(k-1)^{\text{ème}}$.

ECRICOME 2010

Soit r un entier naturel supérieur ou égal à 2. Une urne contient r boules numérotées $1, 2, \dots, r$. On pioche indéfiniment les boules avec remise, chaque boule pouvant être piochée de façon équiprobable.

Pour tout entier $i \in \{1, 2, \dots, r\}$, on note Y_i la variable aléatoire égale au «nombre de pioches nécessaires pour obtenir i boules distinctes». On convient que $Y_1 = 1$.

On désigne par X_r la variable aléatoire égale au «nombre de pioches nécessaires pour obtenir les r boules numérotées $1, 2, \dots, r$ ». Il est immédiat que $X_r = Y_r$.

Par exemple, en supposant que $r = 4$, si les boules piochées successivement portent les numéros

3, 3, 3, 1, 1, 1, 1, 2, 3, 2, 4, 1, ...

alors on a : $Y_1 = 1, Y_2 = 4, Y_3 = 8, Y_4 = X_4 = 11$.

La partie I établit certains résultats préliminaires qui seront utilisés dans d'autres parties.

La partie II se consacre à l'étude de la loi des variables discrètes $Y_{i+1} - Y_i$ afin d'en déduire l'espérance et la variance de la variable discrète X_r .

La partie III détermine la loi de la variable X_r puis étudie la distribution asymptotique de la variable X_r autour de sa moyenne.

On note \exp la fonction exponentielle définie par :

$$\forall x \in \mathbf{R}, \exp(x) = e^x.$$

Partie I : Résultats préliminaires.

1. Étude d'une suite.

On introduit la suite $(u_n)_{n \geq 1}$ définie par : $\forall n \geq 1, u_n = \left(\sum_{i=1}^n \frac{1}{i} \right) - \ln(n)$.

a. Écrire un programme Sci Lab permettant de calculer u_n pour un entier $n \geq 1$ donné.

b. À l'aide d'un développement limité, justifier que $u_n - u_{n+1} \underset{n \rightarrow +\infty}{\sim} \frac{1}{2n^2}$.

En déduire la nature de la série $\sum_{n \geq 1} (u_n - u_{n+1})$ puis démontrer la convergence de la suite $(u_n)_{n \geq 1}$.

c. Montrer que la suite $\left(\sum_{i=1}^n \frac{1}{i^2} \right)_{n \geq 1}$ converge (on ne demande pas le calcul de la limite).

2. Loi de Gumbel.

Soit Z une variable aléatoire. On suppose que Z suit la loi de Gumbel, c'est-à-dire que sa fonction de répartition F_Z est définie par :

$$\forall t \in \mathbf{R}, F_Z(t) = \exp(-\exp(-t)).$$

a. Vérifier que la fonction F_Z est bien une fonction de répartition puis que Z possède une densité que l'on précisera.

b. On considère la variable aléatoire $W = \exp(-Z)$.

Déterminer la fonction de répartition de la variable aléatoire W .

En déduire que la variable aléatoire W suit une loi usuelle dont on précisera le ou les paramètres.

c. Pour tout entier k , montrer que l'intégrale $\int_0^{+\infty} (\ln(x))^k e^{-x} dx$ est absolument convergente.

d. En justifiant le changement de variable $x = \exp(-t)$, démontrer que la variable Z admet un moment d'ordre k valant :

$$E(Z^k) = \int_0^{+\infty} (-\ln(x))^k e^{-x} dx.$$

Partie II : Étude de la variable X_r .

3. Étude du cas $r = 3$.

On suppose uniquement dans cette question que $r = 3$, c'est-à-dire que l'urne ne contient que trois boules numérotées respectivement 1, 2, 3, chacune pouvant être piochée avec la probabilité $\frac{1}{3}$.

- a. Soit n un entier naturel non nul.
 Comparer les événements $[Y_2 > n]$ et C_n : «les n premières pioches fournissent des boules portant toutes le même numéro».
 Calculer la probabilité $P(C_n)$. En déduire la probabilité $P(Y_2 > n)$ puis donner la loi de la variable Y_2 .

b. Justifier que :

$$\forall n \geq 1, P(Y_3 - Y_2 = n) = \sum_{k=2}^{+\infty} P([Y_3 = n + k] \cap [Y_2 = k])$$

puis que :

$$\forall n \geq 1, \forall k \geq 2, P([Y_3 = n + k] \cap [Y_2 = k]) = \frac{1}{3^{k-1}} \left(\frac{2}{3}\right)^n.$$

En déduire la loi de la variable $Y_3 - Y_2$.

Dans toute la suite du problème, r désignera un entier supérieur ou égal à 2.

4. Loi de $Y_{i+1} - Y_i$ pour $i \in \{1, 2, \dots, r - 1\}$.

a. Justifier que :

$$Y_i(\Omega) = \{i, i + 1, i + 2, \dots\} = \mathbf{N} \setminus \{0, 1, 2, \dots, i - 1\} \text{ et } (Y_{i+1} - Y_i)(\Omega) = \mathbf{N} \setminus \{0\}.$$

b. Démontrer que :

$$\forall n \geq 1, \forall k \geq i, P_{[Y_i=k]}(Y_{i+1} - Y_i = n) = \left(\frac{i}{r}\right)^{n-1} \left(1 - \frac{i}{r}\right).$$

c. En déduire que $Y_{i+1} - Y_i$ suit une loi usuelle dont on précisera le ou les paramètres puis établir que :

$$E(Y_{i+1} - Y_i) = \frac{r}{r - i} \text{ et } V(Y_{i+1} - Y_i) = \frac{r \cdot i}{(r - i)^2}.$$

5. Espérance et variance de X_r .

a. Justifier que : $X_r = 1 + \sum_{i=1}^{r-1} (Y_{r-i+1} - Y_{r-i})$.

En admettant que les variables $Y_2 - Y_1, Y_3 - Y_2, \dots, Y_r - Y_{r-1}$ sont indépendantes, vérifier que :

$$E(X_r) = r \sum_{i=1}^r \frac{1}{i} \text{ et } V(X_r) = r^2 \sum_{i=1}^r \frac{1}{i^2} - r \sum_{i=1}^r \frac{1}{i}.$$

b. À l'aide de la question 1, prouver l'existence de deux réels α et β tels que :

$$E(X_r) = r \ln(r) + \alpha r + o(r) \text{ et } V(X_r) \underset{r \rightarrow +\infty}{\sim} \beta r^2.$$

Partie III : Loi de X_r et de sa déviation asymptotique par rapport à sa moyenne.

Pour tout entier $k \in \{1, 2, \dots, r\}$ et tout entier naturel $m \geq 1$, on considère l'événement $A_{k,m}$: «le numéro k n'a pas été pioché durant les m premières pioches».

6. Loi de X_r .

Soit m un entier naturel non nul.

- a. Pour tout entier $k \in \{1, 2, \dots, r\}$, calculer la probabilité de l'événement $A_{k,m}$,
- b. On se fixe à présent k numéros $i_1 < i_2 < \dots < i_k$. Calculer la probabilité qu'aucun de ces numéros n'ait été pioché durant les m premiers tirages.
- c. Justifier que :

$$P(X_r > m) = P(A_{1,m} \cup A_{2,m} \cup \dots \cup A_{r,m}).$$

En déduire que $P(X_3 > m) = 3 \left(1 - \frac{1}{3}\right)^m - 3 \left(1 - \frac{2}{3}\right)^m$.

On admet que plus généralement que

$$P(X_r > m) = \sum_{k=1}^{r-1} (-1)^{k-1} \binom{r}{k} \left(1 - \frac{k}{r}\right)^m.$$

7. Comportement de X_r au delà de sa moyenne.

- a. À l'aide d'une récurrence sur m , montrer que, pour toute famille (D_1, \dots, D_m) d'événements, on a :

$$P(D_1 \cup D_2 \cup \dots \cup D_m) \leq P(D_1) + P(D_2) + \dots + P(D_m).$$

- b. Démontrer que pour tout réel x , on a : $\exp(x) \geq 1 + x$. En déduire que :

$$\forall m \in \mathbf{N} \setminus \{0\}, \forall k \in \{1, \dots, r\}, P(A_{k,m}) \leq \exp\left(-\frac{m}{r}\right).$$

- c. Soit $\varepsilon > 0$, on note M_r la partie entière de $(1 + \varepsilon)r \ln(r)$, c'est-à-dire l'unique entier relatif tel que :

$$M_r \leq (1 + \varepsilon)r \ln(r) < M_r + 1.$$

Comparer les événements $[X_r > M_r]$ et $[X_r > (1 + \varepsilon)r \ln(r)]$.

En déduire que :

$$P(X_r > (1 + \varepsilon)r \ln(r)) \leq \frac{e}{r^\varepsilon}.$$

Ainsi on vient d'établir que :

$$\forall \varepsilon > 0, \lim_{r \rightarrow +\infty} P(X_r > (1 + \varepsilon)r \ln(r)) = 0$$

qui peut se traduire ainsi : l'événement « X_r est significativement supérieur à sa moyenne» est un événement asymptotiquement rare.

8. Distribution de X_r autour de sa moyenne.

On introduit la suite $(Z_r)_{r \geq 2}$ de variables aléatoires définie par :

$$\forall r \geq 2, Z_r = \frac{X_r - r \ln(r)}{r}.$$

Soit t un réel fixé, on note m_r la partie entière du réel $r \ln(r) + rt$, c'est-à-dire l'unique entier relatif tel que :

$$m_r \leq r \ln(r) + rt < m_r + 1.$$

- a. Justifier l'existence d'un rang $r_0(t)$ tel que :

$$\forall r \geq r_0(t), m_r \geq 1$$

puis prouver l'égalité :

$$\forall r \geq r_0(t), P(Z_r > t) = P(X_r > m_r).$$

- b. Soit k un entier naturel. À l'aide d'un développement limité, établir que :

$$m_r \ln\left(1 - \frac{k}{r}\right) = -k \ln(r) - kt + o_{r \rightarrow +\infty}(1).$$

- c. Démontrer que, pour tout entier k , on a : $\binom{r}{k} \underset{r \rightarrow +\infty}{\sim} \frac{r^k}{k!}$.

En déduire que :

$$\forall k \in \mathbf{N}, \lim_{r \rightarrow +\infty} \binom{r}{k} \left(1 - \frac{k}{r}\right)^{m_r} = \frac{\exp(-kt)}{k!}.$$

- d. En admettant que l'on a

$$\lim_{r \rightarrow +\infty} \sum_{k=1}^{r-1} (-1)^{k-1} \binom{r}{k} \left(1 - \frac{k}{r}\right)^{m_r} = \sum_{k=1}^{+\infty} (-1)^{k-1} \frac{\exp(-kt)}{k!},$$

exprimer la valeur de la limite $\lim_{r \rightarrow +\infty} P(Z_r \leq t)$ en fonction de $F_Z(t)$ (définie à la question 2).

Quel résultat vient-on d'établir sur la suite de variables aléatoires $(Z_r)_{r \geq 2}$?

ECRICOME 2010 : CORRECTION

Partie I : Résultats préliminaires

1. Étude d'une suite.

1.a. Il s'agit de calculer la somme à l'aide d'une boucle, puis de lui retirer $\ln(n)$.

```

1 fonction u = suite(n)
2     u = 0
3     for i=1 :n
4         u = u+1/i
5     end
6     u = u-log(n)
7 endfunction

```

1.b. On a

$$\begin{aligned}
 u_{n+1} - u_n &= \sum_{i=1}^{n+1} \frac{1}{i} - \ln(n+1) - \left(\sum_{i=1}^n \frac{1}{i} + \ln(n) \right) \\
 &= \frac{1}{n+1} - \ln\left(\frac{n+1}{n}\right) = \frac{1}{n+1} - \ln\left(1 + \frac{1}{n}\right) \\
 &= \frac{1}{n} \frac{1}{1 + \frac{1}{n}} - \ln\left(1 + \frac{1}{n}\right) \\
 &= \frac{1}{n} \left(1 - \frac{1}{n} + o_{n \rightarrow +\infty}\left(\frac{1}{n}\right)\right) - \frac{1}{n} + \frac{1}{2n^2} + o_{n \rightarrow +\infty}\left(\frac{1}{n^2}\right) \\
 &= \frac{1}{n} - \frac{1}{n^2} + o_{n \rightarrow +\infty}\left(\frac{1}{n^2}\right) - \frac{1}{n} + \frac{1}{2n^2} + o_{n \rightarrow +\infty}\left(\frac{1}{n^2}\right) \\
 &= -\frac{1}{2n^2} + o_{n \rightarrow +\infty}\left(\frac{1}{n^2}\right) \underset{n \rightarrow +\infty}{\sim} -\frac{1}{2n^2}.
 \end{aligned}$$

o/\sim

Rappelons que $u_n \sim v_n$ si et seulement si

$$u_n = v_n + o(v_n).$$

Par comparaison à une série de Riemann convergente, d'après le critère de comparaison pour les séries à termes de signe constant, on en déduit que la série de terme général $u_{n+1} - u_n$ converge.

Mais pour $n \in \mathbf{N}^*$, la somme partielle d'ordre n de cette série est

$$\sum_{k=1}^n (u_{k+1} - u_k) = \sum_{k=1}^n u_{k+1} - \sum_{k=1}^n u_k = \sum_{i=2}^{n+1} u_i - \sum_{k=1}^n u_k = u_{n+1} - u_1.$$

Puisque la série converge, la suite de ses sommes partielles converge également, et donc $(u_{n+1} - u_1)_n$ converge. C'est donc que la suite (u_n) converge.

Pour la culture

La limite de cette suite est souvent notée γ et est appelée constante d'Euler.

1.c. Il s'agit de la suite des sommes partielles de la série de Riemann convergente $\sum \frac{1}{i^2}$.

Et donc¹ c'est une suite convergente.

2. Loi de Gumbel.

2.a. La fonction F_Z est continue sur \mathbf{R} par composition de fonctions continues, et donc en particulier continue à droite en tout point. Mieux : elle est même \mathcal{C}^1 .

Elle est croissante car sa dérivée, qui est $t \mapsto e^{-t} e^{-e^{-t}}$, est positive sur \mathbf{R} .

Enfin, on a $\lim_{t \rightarrow -\infty} F_Z(t) = 0$ et $\lim_{t \rightarrow +\infty} F_Z(t) = e^0 = 1$.

Donc F_Z est bien la fonction de répartition d'une variable aléatoire Z .

Et puisque nous avons déjà mentionné que F_Z est continue sur \mathbf{R} , et même \mathcal{C}^1 sur \mathbf{R} ,

Z est une variable à densité.

Une densité en est par exemple $F'_Z : t \mapsto e^{-t-e^{-t}}$.

¹ C'est la **définition** de la convergence d'une série.

- 2.b. Il est évident² que W est à valeurs strictement positives, et donc pour $x \leq 0$, $P(W \leq x) = 0$. Pour $x > 0$, on a

$$P(W \leq x) = P(e^{-Z} \leq x) = P(-Z \leq \ln(x)) = P(Z \geq -\ln(x)) = 1 - F_Z(\ln(x)) = 1 - e^{-x}.$$

Nous reconnaissons alors la fonction de répartition d'une loi exponentielle de paramètre 1 : $W \hookrightarrow \mathcal{E}(1)$.

- 2.c. La fonction $x \mapsto (\ln x)^k e^{-x}$ est continue sur $]0, +\infty[$, donc l'intégrale est impropre en 0 et en $+\infty$.

Au voisinage de $+\infty$, puisque $(\ln(x))^k = o_{x \rightarrow +\infty}(x)$, il vient $(\ln x)^k e^{-x} = o_{x \rightarrow +\infty}(x e^{-x})$.

Or, l'intégrale $\int_1^{+\infty} x e^{-x} dx$ est convergente, puisque nous savons que $\int_0^{+\infty} x e^{-x} dx = \Gamma(2)$ est convergente.

On en déduit donc que $\int_1^{+\infty} (\ln x)^k e^{-x} dx$ converge. Puisqu'il s'agit de l'intégrale d'une fonction positive, la convergence est équivalente à la convergence absolue.

Au voisinage de 0, on a $x^{1/2} (\ln x)^k e^{-x} \underset{x \rightarrow 0}{\sim} x^{1/2} (\ln(x))^k \xrightarrow{x \rightarrow 0^+} 0$ par croissances comparées.

Et donc $(\ln x)^k e^{-x} = o_{x \rightarrow 0^+} \left(\frac{1}{\sqrt{x}} \right)$.

Puisque l'intégrale de Riemann $\int_0^1 \frac{dx}{\sqrt{x}}$ converge, il en est de même de $\int_0^1 (\ln x)^k e^{-x} dx$.

D'autre part, $x \mapsto (\ln x)^k e^{-x}$ est de signe constant³ sur $]0, 1]$, et donc la convergence de l'intégrale est équivalente à la convergence absolue.

Et donc nous venons de prouver que $\int_0^{+\infty} (\ln x)^k e^{-x} dx$ est absolument convergente.

- 2.d. Par le théorème de transfert, Z admet un moment d'ordre k si et seulement si $\int_{-\infty}^{+\infty} t^k e^{-t-e^{-t}} dt$ converge absolument.

La fonction $t \mapsto e^{-t}$ est \mathcal{C}^1 sur \mathbf{R}_+ , et strictement décroissante, donc le théorème de changement de variable s'applique.

Posons alors $x = e^{-t} \Leftrightarrow t = -\ln(x)$, de sorte que $dt = -\frac{dx}{x}$. De plus, lorsque $t \rightarrow -\infty$, $x \rightarrow +\infty$ et lorsque $t \rightarrow +\infty$, $x \rightarrow 0$.

Le théorème de changement de variable nous garantit alors que $\int_{-\infty}^{+\infty} t^k e^{-t-e^{-t}} dt$ et

$\int_{+\infty}^0 (-\ln(x))^k x e^{-x} \frac{-dx}{x} = \int_0^{+\infty} (-\ln x)^k e^{-x} dx$ sont de même nature, et en cas de convergence, sont égales.

Or nous avons prouvé à la question précédente que cette seconde intégrale est convergente, et donc $E(Z^k)$ existe, et

$$E(Z^k) = \int_0^{+\infty} (-\ln x)^k e^{-x} dx.$$

Partie II : Étude de la variable X_r .

3. Étude du cas $r = 3$.
- 3.a. L'événement $[Y_2 > n]$ est réalisé si et seulement si il faut strictement plus de n tirages pour obtenir deux numéros différents. C'est le cas si et seulement si les n premiers tirages ont fourni le même numéro.

Et donc $[Y_2 > n] = C_n$.

Nommons Z_i la variable aléatoire égale au numéro de la boule obtenue lors du $i^{\text{ème}}$ tirage.

Alors $C_n = \bigcup_{k=1}^3 \bigcap_{i=1}^n [Z_i = k]$.

Et donc, par incompatibilité des événements de l'union, et indépendance des Z_i ,

$$P(C_n) = \sum_{k=1}^3 P\left(\bigcap_{i=1}^n [Z_i = k]\right) = \sum_{k=1}^3 \prod_{i=1}^n P(Z_i = k) = \sum_{k=1}^3 \left(\frac{1}{3}\right)^n = \frac{1}{3^{n-1}}.$$

² En raison de la présence de l'exponentielle.

³ Il dépend de la parité de k , mais pas de t .

CV absolue

Le théorème de changement de variable ne dit rien de la convergence absolue. En réalité, en remplaçant la fonction intégrée par sa valeur absolue, on montre aisément que l'intégrale avant changement de variable converge absolument si et seulement si l'intégrale après changement de variable converge absolument.

Enfin, puisque Y_2 est à valeur entières, pour tout $n \geq 2$,

$$P(Y_2 = n) = P(Y_2 > n - 1) - P(Y_2 > n) = \frac{1}{3^{n-2}} - \frac{1}{3^{n-1}} = \frac{2}{3} \left(\frac{1}{3}\right)^{n-2}.$$

- 3.b. Appliquons la formule des probabilités totales au système complet d'événements $\{[Y_2 = k], k \geq 2\}$. Il vient alors

$$\begin{aligned} P(Y_3 - Y_2 = n) &= \sum_{k=2}^{+\infty} P([Y_3 - Y_2 = n] \cap [Y_2 = k]) \\ &= \sum_{k=2}^{+\infty} P([Y_3 - k = n] \cap [Y_2 = k]) \\ &= \sum_{k=2}^{+\infty} P([Y_3 = n + k] \cap [Y_2 = k]). \end{aligned}$$

On a de plus, $P([Y_3 = n + k] \cap [Y_2 = k]) = P(Y_2 = k)P_{[Y_2=k]}(Y_3 = n + k)$.

Or, cette probabilité conditionnelle est la probabilité que les tirages $k + 1, k + 2, \dots, k + n - 1$ donnent une boule parmi les deux déjà sorties précédemment, ce qui se produit avec probabilité $\frac{2}{3}$, et que le tirage $n + k$ donne la troisième boule, ce qui se produit avec probabilité $\frac{1}{3}$.

Et donc⁴

$$P([Y_3 = n + k] \cap [Y_2 = k]) = P(Y_2 = k) \left(\frac{2}{3}\right)^n \frac{1}{3} = \left(\frac{1}{3}\right)^{k-2} \left(\frac{2}{3}\right)^n \frac{1}{3} = \frac{1}{3^{k-1}} \left(\frac{2}{3}\right)^n.$$

On en déduit donc que

$$P(Y_3 - Y_2 = n) = \sum_{k=2}^{+\infty} \frac{1}{3^{k-1}} \left(\frac{2}{3}\right)^n = \left(\frac{2}{3}\right)^n \frac{1}{3} \sum_{k=2}^{+\infty} \frac{1}{3^{k-1}} = \left(\frac{2}{3}\right)^n \frac{1}{3} \sum_{i=0}^{+\infty} \frac{1}{3^i} = \frac{1}{3} \left(\frac{2}{3}\right)^n \frac{1}{1 - \frac{1}{3}} = \frac{1}{3} \left(\frac{2}{3}\right)^{n-1}.$$

On reconnaît alors une loi géométrique : $Y_3 - Y_2 \hookrightarrow \mathcal{G}\left(\frac{1}{3}\right)$.

4. Loi de $Y_{i+1} - Y_i$.

- 4.a. Pour obtenir i boules portant des numéros différents, il faut au moins i essais. Et il est évident qu'il n'y a pas de limite au nombre d'essais nécessaires, puisque pour tout n , les n premiers tirages peuvent donner la boule numéro 1. Donc $Y_i(\Omega) = \mathbf{N} \setminus \{1, 2, \dots, i - 1\}$.

De plus, $Y_{i+1} - Y_i$ représente le nombre de tirages, une fois qu'on a i boules différentes, pour obtenir une boule portant un numéro non encore obtenu. Il faut au moins un tirage pour que ceci se produise, mais il n'y a pas de limite au nombre de tirages nécessaires, donc $(Y_{i+1} - Y_i)(\Omega) = \mathbf{N}^*$.

- 4.b. Commençons par noter que

$$\begin{aligned} P_{[Y_i=k]}(Y_{i+1} - Y_i = n) &= \frac{P([Y_i = k] \cap [Y_{i+1} - Y_i = n])}{P(Y_i = k)} \\ &= \frac{P([Y_i = k] \cap [Y_{i+1} = n + k])}{P(Y_i = k)} = P_{[Y_i=k]}(Y_{i+1} = n + k). \end{aligned}$$

C'est le même principe qu'à la question 3.b : sachant que $[Y_i = k]$ est réalisé, alors $[Y_{i+1} = n + k]$ est réalisé si et seulement si les tirages $k + 1, k + 2, \dots, n + k - 1$ donnent des boules qui ne sont pas parmi les i déjà précédemment obtenues, et si le tirage $n + k$ donne une nouvelle boule.

Et donc, par indépendance des tirages successifs,

$$P_{[Y_i=k]}(Y_{i+1} - Y_i = k) = \left(\frac{i}{r}\right)^{n-1} \left(1 - \frac{i}{r}\right).$$

Remarque

Cela ressemble beaucoup à une loi géométrique de paramètre $2/3$, si ce n'est qu'on a un $n - 2$ au lieu d'un $n - 1$. Ce n'est pas une grosse surprise : Y_2 compte le nombre d'essais, après le premier, avant d'obtenir une boule différente du premier, ce qui se produit avec proba $\frac{2}{3}$. Donc si on ajoute le premier essai,

$$Y_2 + 1 \hookrightarrow \mathcal{G}\left(\frac{2}{3}\right).$$

⁴ Par indépendance des tirages successifs.

Rédaction

Cette rédaction n'est pas tout à fait satisfaisante, mais on sent bien que si on commence à nommer des événements pour essayer de faire les choses correctement, on ne s'en sortira pas. Et donc il faut aussi accepter de se contenter d'un peu d'intuition probabiliste.

- 4.c. Par la formule des probabilités totales appliquée au système complet d'événements $\{[Y_i = k], k \geq i\}$, on a

$$\begin{aligned} P(Y_{i+1} - Y_i = n) &= \sum_{k=i}^{+\infty} P(Y_i = k)P_{[Y_i=k]}(Y_{i+1} - Y_i = n) \\ &= \sum_{k=i}^{+\infty} P(Y_i = k) \left(\frac{i}{r}\right)^{n-1} \left(1 - \frac{i}{r}\right) \\ &= \left(\frac{i}{r}\right)^{n-1} \left(1 - \frac{i}{r}\right) \underbrace{\sum_{k=i}^{+\infty} P(Y_i = k)}_{=1} \\ &= \left(\frac{i}{r}\right)^{n-1} \left(1 - \frac{i}{r}\right). \end{aligned}$$

Ainsi, $Y_{i+1} - Y_i$ suit la loi géométrique de paramètre $1 - \frac{i}{r}$.

Et donc en particulier,

$$E(Y_{i+1} - Y_i) = \frac{1}{1 - \frac{i}{r}} = \frac{r}{r - i} \text{ et } V(Y_{i+1} - Y_i) = \frac{\frac{i}{r}}{\left(1 - \frac{i}{r}\right)^2} = \frac{ir}{(r - i)^2}.$$

5. **Espérance et variance de X_r .**

- 5.a. On a $X_r = Y_r$. Mais d'autre part,

$$\sum_{i=1}^{r-1} (Y_{i+1} - Y_i) = \sum_{i=1}^{r-1} Y_{i+1} - \sum_{i=1}^{r-1} Y_i = Y_r - Y_1 = Y_r - 1$$

et donc $X_r = Y_r = Y_r - 1 + 1 = 1 + \sum_{i=1}^{r-1} (Y_{i+1} - Y_i)$.

Par linéarité de l'espérance, il vient donc

$$E(X_r) = 1 + \sum_{i=1}^{r-1} E(Y_{i+1} - Y_i) = 1 + \sum_{i=1}^{r-1} \frac{r}{r - i} = r \sum_{i=0}^{r-1} \frac{1}{r - i} = r \sum_{k=1}^r \frac{1}{k}.$$

Et par indépendance des $Y_{i+1} - Y_i$, il vient

$$\begin{aligned} V(Y_r) &= \underbrace{V(1)}_{=0} + \sum_{i=1}^{r-1} V(Y_{i+1} - Y_i) = \sum_{i=1}^{r-1} \frac{ri}{(r - i)^2} \\ &= \sum_{k=1}^{r-1} \frac{r(r - k)}{k^2} = \sum_{k=1}^r \frac{r(r - k)}{k^2} \\ &= r^2 \sum_{k=1}^r \frac{1}{k^2} - r \sum_{k=1}^r \frac{1}{k}. \end{aligned}$$

- 5.b. À la question 1, nous avons prouvé qu'il existe un réel α tel que

$$u_n \xrightarrow[n \rightarrow +\infty]{} \alpha \Leftrightarrow u_n = \alpha + o_{n \rightarrow +\infty}(1).$$

Et donc $E(X_r) = r(u_r + \ln(r)) = r\left(\alpha + o_{r \rightarrow +\infty}(1) + \ln(r)\right) = r \ln(r) + \alpha r + o_{r \rightarrow +\infty}(r)$.

D'autre part, on a $\sum_{i=1}^n \frac{1}{i} = u_n + \ln(n)$, et puisque (u_n) converge, $u_n = o_{n \rightarrow +\infty}(\ln n)$.

Et donc $\sum_{i=1}^n \frac{1}{i} = \ln(n) + o_{n \rightarrow +\infty}(\ln n) \underset{n \rightarrow +\infty}{\sim} \ln n$.

Chgt d'indice

On a posé $k = r - i$.

Indépendance

Bien qu'elle soit difficile à prouver, l'indépendance de $Y_{i+1} - Y_i$ est assez intuitive. En effet, $Y_{i+1} - Y_i$ représente le nombre de tirages nécessaires après l'obtention de la $i^{\text{ème}}$ boule pour obtenir la $(i + 1)^{\text{ème}}$. Mais par indépendance des tirages, le nombre d'essais nécessaires est indépendant du nombre d'essais qui ont été nécessaires pour obtenir les i premières boules.

On en déduit donc que $r \sum_{i=1}^r \frac{1}{i} \underset{r \rightarrow +\infty}{\sim} r \ln(r)$.

D'autre part, il existe un réel β tel que $\sum_{i=1}^n \frac{1}{i^2} \xrightarrow{n \rightarrow +\infty} \beta$, et donc $r^2 \sum_{i=1}^r \frac{1}{i^2} \underset{r \rightarrow +\infty}{\sim} r^2 \beta$.

Puisque $\ln r = \underset{r \rightarrow +\infty}{o}(r^2)$, on en déduit que

$$V(X_r) = r^2 \sum_{i=1}^r \frac{1}{i^2} + \underset{r \rightarrow +\infty}{o}(r^2) \underset{r \rightarrow +\infty}{\sim} r^2 \sum_{i=1}^r \frac{1}{i^2} \underset{r \rightarrow +\infty}{\sim} \boxed{\beta r^2}.$$

Partie III : Loi de X_r et de sa déviation asymptotique par rapport à sa moyenne.

6. Loi de X_r .

6.a. Avec les notations précédentes, on a

$$A_{k,m} = [Z_1 \neq k] \cap [Z_2 \neq k] \cap \dots \cap [Z_m \neq k].$$

Par indépendance des Z_i , il vient donc

$$P(A_{k,m}) = \prod_{i=1}^m P(Z_i \neq k) = \prod_{i=1}^m \left(\frac{r-1}{r} \right) = \left(\frac{r-1}{r} \right)^m.$$

6.b. À chaque tirage, la probabilité de n'obtenir aucun des k numéros choisis est $1 - \frac{k}{r}$.

Et donc par indépendance des tirages, la probabilité souhaitée est $\left(1 - \frac{k}{r} \right)^m$.

6.c. L'événement $[X_r > m]$ est réalisé si et seulement si l'un **au moins** des r numéros n'est pas sorti lors des m premiers tirages.

Donc si et seulement si l'un des événements $A_{1,m}, A_{2,m}, \dots, A_{r,m}$ est réalisé.

Et donc $\boxed{[X_r > m] = A_{1,m} \cup A_{2,m} \cup \dots \cup A_{r,m}}$.

En particulier, il vient

$$\begin{aligned} P(X_3 > m) &= P(A_{1,m} \cup A_{2,m} \cup A_{3,m}) = P((A_{1,m} \cup A_{2,m}) \cup A_{3,m}) \\ &= P(A_{1,m} \cup A_{2,m}) + P(A_{3,m}) - P((A_{1,m} \cup A_{2,m}) \cap A_{3,m}) \\ &= P(A_{1,m}) + P(A_{2,m}) - P(A_{1,m} \cap A_{2,m}) + P(A_{3,m}) - P((A_{1,m} \cap A_{3,m}) \cup (A_{2,m} \cap A_{3,m})) \\ &= P(A_{1,m}) + P(A_{2,m}) - P(A_{1,m} \cap A_{2,m}) + P(A_{3,m}) - P(A_{1,m} \cap A_{3,m}) - P(A_{2,m} \cap A_{3,m}) + P(A_{1,m} \cap A_{2,m} \cap A_{3,m}) \\ &= (P(A_{1,m}) + P(A_{2,m}) + P(A_{3,m})) - (P(A_{1,m} \cap A_{2,m}) + P(A_{1,m} \cap A_{3,m}) + P(A_{2,m} \cap A_{3,m})) + P(A_{1,m} \cap A_{2,m} \cap A_{3,m}) \\ &= 3 \left(1 - \frac{1}{3} \right)^m - 3 \left(1 - \frac{2}{3} \right)^m + 0. \end{aligned}$$

Détails

Ces probabilités sont celles qui ont été calculées en 6.b.

7. Comportement de X_r au delà de sa moyenne.

7.a. Pour $m = 1$, la formule est évidente, c'est même une égalité !

Supposons que pour toute famille (D_1, \dots, D_m) de m événements, on ait

$$P(D_1 \cup D_2 \cup \dots \cup D_m) \leq P(D_1) + \dots + P(D_m)$$

et soit (C_1, \dots, C_{m+1}) une famille de $m + 1$ événements. Alors

$$\begin{aligned} P(C_1 \cup \dots \cup C_{m+1}) &= P((C_1 \cup \dots \cup C_m) \cup C_{m+1}) \\ &= P(C_1 \cup \dots \cup C_m) + P(C_{m+1}) - P((C_1 \cup \dots \cup C_m) \cap C_{m+1}) \\ &\leq P(C_1 \cup \dots \cup C_m) + P(C_{m+1}) \\ &\leq P(C_1) + \dots + P(C_m) + P(C_{m+1}). \end{aligned}$$

$$P(A \cup B) = P(A) + P(B) - P(A \cap B).$$

Hypothèse de récurrence.

Donc la formule est encore vraie pour une famille de $m + 1$ événements, et donc est vraie pour toute famille finie d'événements.

7.b. La fonction exponentielle est convexe sur \mathbf{R} . Et donc est située sous ses tangentes. Or, la tangente en 0 est la droite d'équation $y = x + 1$, de sorte que $\boxed{\text{pour tout } x \in \mathbf{R}, e^x \leq 1 + x.}$

On a donc, pour $m \in \mathbf{N}^*$ et $k \in \llbracket 1, m \rrbracket$, $P(A_{k,m}) = \left(1 - \frac{1}{r}\right)^m$.

Appliquons l'inégalité précédente avec $x = -\frac{1}{m}$, de sorte que

$$1 + x = 1 - \frac{1}{m} \leq e^{-\frac{1}{m}}.$$

En élevant cette inégalité à la puissance r , il vient donc

$$\left(1 - \frac{1}{m}\right)^r \leq \left(e^{-\frac{1}{m}}\right)^r = e^{-\frac{r}{m}} \Leftrightarrow \boxed{P(A_{k,m}) \leq e^{-\frac{r}{m}}.}$$

7.c. On a $[X_r > M_r] = [M_r < X_r \leq (1 + \varepsilon)r \ln(n)] \cup [X_r > (1 + \varepsilon)r \ln(n)]$.
Or, X_r ne prend que des valeurs entières, et par définition de M_r , il n'y a pas d'entier dans l'intervalle $]M_r, (1 + \varepsilon)r \ln(n)[$.

Autrement dit, $[M_r < X_r \leq (1 + \varepsilon)r \ln(n)] = \emptyset$, de sorte que $\boxed{[X_r > M_r] = [X_r > (1 + \varepsilon)r \ln(n)]}$.

On en déduit que $P(X_r > (1 + \varepsilon)r \ln(n)) = P(X_r > M_r)$.

Mais d'après les questions 6.c et 7.a,

$$P(X_r > M_r) = P(A_{1,M_r} \cup A_{2,M_r} \cup \dots \cup A_{r,M_r}) \leq P(A_{1,M_r}) + \dots + P(A_{r,M_r}).$$

En utilisant la question 7.b, on a donc $P(X_r > M_r) \leq r \exp\left(-\frac{M_r}{r}\right)$.

Mais $M_r > (1 + \varepsilon)r \ln(n) - 1$, donc

$$\exp\left(-\frac{M_r}{r}\right) \leq \exp\left(-\left((1 + \varepsilon)\ln(n) + \frac{1}{r}\right)\right) = \frac{1}{r^{1+\varepsilon}} \exp\left(-\frac{1}{r}\right) \leq \frac{1}{r^{1+\varepsilon}} e^{-1}.$$

Et donc finalement

$$P(X_r > (1 + \varepsilon)r \ln(n)) = P(X_r > M_r) \leq r \frac{1}{r^{1+\varepsilon}} e^{-1} \leq \boxed{\frac{e}{r^\varepsilon}}.$$

8. Distribution de X_r autour de sa moyenne.

8.a. Puisque $r \ln(n) + rt \xrightarrow{r \rightarrow +\infty}$, il existe $r_0(t)$ tel que pour $r \geq r_0(t)$, $r \ln(n) + rt \geq 1$, et donc⁵

$$\boxed{\forall r \geq r_0(t), m_r \geq 1.}$$

On a alors, toujours pour $r \geq r_0(t)$,

$$P(Z_r > t) = P\left(\frac{X_r - r \ln(n)}{r} > t\right) = P(X_r > r \ln(n) + rt)$$

et X_r étant à valeurs entières, on prouve comme à la question 7.c que

$$P(Z_r > t) = P(X_r > r \ln(n) + rt) = P(X_r > m_r).$$

8.b. Puisque $\frac{k}{r} \xrightarrow{r \rightarrow +\infty} 0$, on a

$$\ln\left(1 - \frac{k}{r}\right) = -\frac{k}{r} - \frac{k^2}{2r^2} + o_{r \rightarrow +\infty}\left(\frac{k^2}{r^2}\right) = -\frac{k}{r} - \frac{k^2}{2r^2} + o_{r \rightarrow +\infty}\left(\frac{1}{r^2}\right).$$

Et donc $m_r \ln\left(1 - \frac{k}{r}\right) = -\frac{m_r k}{r} - \frac{m_r k^2}{2r^2} + o_{r \rightarrow +\infty}\left(\frac{m_r}{r^2}\right)$.

Mais $m_r \sim_{r \rightarrow +\infty} r \ln(n) + rt$ et $rt = o_{r \rightarrow +\infty}(r \ln(n))$, de sorte que $m_r \sim_{r \rightarrow +\infty} r \ln(n)$.

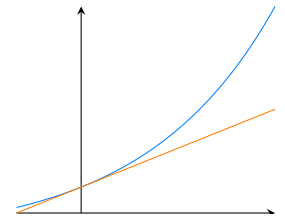


FIGURE 21.1– La fonction exponentielle et sa tangente en 0.

⁵ Un réel supérieur à 1 possède une partie entière supérieure à 1.

Détails

Lorsque $x \rightarrow +\infty$,

$$[x] \underset{x \rightarrow +\infty}{\sim} x.$$

Pour le prouver, le plus simple est sûrement de revenir à la définition de la partie entière :

$$x - 1 < [x] \leq x$$

et d'utiliser un théorème des gendarmes.

On en déduit que $\frac{m_r}{r^2} \underset{r \rightarrow +\infty}{\sim} \frac{\ln(r)}{r} \underset{r \rightarrow +\infty}{\rightarrow} 0$.

Et donc

$$m_r \ln\left(1 - \frac{k}{r}\right) = -\frac{km_r}{r} + o_{r \rightarrow +\infty}(1).$$

Enfin, par définition d'une partie entière, $0 \leq m_r - r \ln(r) - rt < 1$ donc $0 \leq \frac{m_r - r \ln(r) - rt}{r} < \frac{1}{r}$.

Par le théorème des gendarmes, on a donc $\lim_{r \rightarrow +\infty} \frac{m_r - r \ln(r) - rt}{r} = 0$ et donc $\frac{m_r}{r} = \frac{m_r - r \ln(r) - rt}{r} + \ln(r) + t = \ln(r) + t + o_{r \rightarrow +\infty}(1)$.

Nous avons donc enfin prouvé que

$$m_r \ln\left(1 - \frac{k}{r}\right) = -k \ln(r) - kt + o_{r \rightarrow +\infty}(1).$$

$o(1)$

Rappelons que la notation $o(1)$ désigne toute suite de limite nulle. En effet, on a

$$u_n = o(1) \Leftrightarrow \lim \frac{u_n}{1} = 0.$$

8.c. C'est très classique : par définition

$$\binom{r}{k} = \frac{r!}{k!(r-k)!} = \frac{r(r-1) \cdot (r-k+1)}{k!}.$$

Or le numérateur est un polynôme de degré k en r , de coefficient dominant égal à 1 : lorsque $r \rightarrow +\infty$, il est équivalent à r^k .

Et donc $\binom{r}{k} \underset{r \rightarrow +\infty}{\sim} \frac{r^k}{k!}$.

Nous avons

$$\begin{aligned} \left(1 - \frac{k}{r}\right)^{m_r} &= \exp\left(m_r \ln\left(1 - \frac{k}{r}\right)\right) = \exp(-k \ln(r) - kt) \exp(o(1)) \\ &\underset{r \rightarrow +\infty}{\sim} \exp(-k \ln(r) - kt) = \frac{1}{r^k} \exp(-kt). \end{aligned}$$

Détails

Puisque $o(1)$ tend vers 0, $\exp(o(1))$ tend vers 1.

Et par conséquent,

$$\binom{r}{k} \left(1 - \frac{k}{r}\right)^{m_r} \underset{r \rightarrow +\infty}{\sim} \frac{r^k}{k!} \frac{1}{r^k} \exp(-kt) \underset{r \rightarrow +\infty}{\rightarrow} \frac{e^{-kt}}{k!}.$$

8.d. On a donc, pour $t \in \mathbf{R}$ et pour $r \geq r_0(t)$,

$$P(Z_r \leq t) = 1 - P(Z_r > t) = 1 - P(X_r > m_r) = 1 - \sum_{k=1}^r (-1)^{k-1} \binom{r}{k} \left(1 - \frac{k}{r}\right)^{m_r} = 1 - \sum_{k=1}^{r-1} (-1)^{k-1} \binom{r}{k} \left(1 - \frac{k}{r}\right)^{m_r}.$$

Et donc

$$\begin{aligned} \lim_{r \rightarrow +\infty} P(Z_r \leq t) &= 1 - \sum_{k=1}^{+\infty} (-1)^{k-1} \frac{\exp(-kt)}{k!} \\ &= 1 + \sum_{k=0}^{+\infty} (-1)^k \frac{\exp(-kt)}{k!} - 1 \\ &= \sum_{k=0}^{+\infty} \frac{(-\exp(-t))^k}{k!} \\ &= \exp(-\exp(-t)). \end{aligned}$$

On ajoute et on retire le terme correspondant à $k = 0$.

Somme d'une série exponentielle.

Et donc nous venons de prouver que $Z_r \xrightarrow{\mathcal{L}} Z$.

EDHEC 2013

1. On considère la fonction f définie sur \mathbf{R} par

$$f(x) = \begin{cases} \frac{1}{2x^2} & \text{si } x \leq -1 \text{ ou } x \geq 1 \\ 0 & \text{sinon} \end{cases}$$

Montrer que f est une densité de probabilités.

Dans la suite, on considère une suite $(X_k)_{k \in \mathbf{N}^*}$ de variables aléatoires, toutes définies sur le même espace probabilisé (Ω, \mathcal{A}, P) , mutuellement indépendantes et admettant toutes f comme densité.

De plus, pour tout entier naturel n non nul, on pose $S_n = \text{Max}(X_1, \dots, X_n)$ et $Y_n = \frac{S_n}{n}$.

on admet que S_n et Y_n sont des variables aléatoires à densité, définies elles aussi sur (Ω, \mathcal{A}, P) .

2. Déterminer la fonction de répartition, notée F , commune aux variables aléatoires X_k .
3. On note G_n la fonction de répartition de la variable aléatoire Y_n . Déterminer explicitement $G_n(x)$ en fonction de n et de x .
4.
 - a. Montrer que pour tout réel x négatif ou nul, on a $G_n(x) \leq \frac{1}{2^n}$.
 - b. Justifier que, pour tout réel x strictement positif, il existe un entier naturel n_0 non nul, tel que, pour tout entier n supérieur ou égal à n_0 , on a $x \geq \frac{1}{n}$.
En déduire que : $\forall x > 0, \exists n_0 \in \mathbf{N}^*, \forall n \geq n_0, G_n(x) = \left(1 - \frac{1}{2nx}\right)^n$.
5.
 - a. Déterminer, pour tout réel x , la limite de $G_n(x)$ lorsque n tend vers $+\infty$. On note $G(x)$ cette limite.
 - b. Montrer que la fonction G ainsi définie est la fonction de répartition d'une variable aléatoire à densité.
 - c. En déduire que la suite $(Y_n)_{n \geq 1}$ converge en loi vers une variable aléatoire Y dont la fonction de répartition est G .
6. Vérifier que la variable aléatoire $\frac{1}{Y}$ suit une loi exponentielle dont on précisera le paramètre.

EDHEC 2013 : CORRECTION

1. La fonction f est positive sur \mathbf{R} . Elle est continue sur $]-\infty, -1[$ et sur $]1, +\infty[$ car il s'agit d'une fonction usuelle, et elle est continue sur $] - 1, 1[$ car constante. Donc f est continue sauf en -1 et en 1 .

De plus, f est paire, donc $\int_{-\infty}^{+\infty} f(t) dt$ converge si et seulement si $\int_0^{+\infty} f(t) dt = \int_1^{+\infty} f(t) dt$ converge.

Or on reconnaît alors une intégrale de Riemann qui converge. De plus, pour $A > 0$, on a

$$\int_1^A f(t) dt = \int_1^A \frac{1}{2t^2} dt = \left[-\frac{1}{2t} \right]_1^A = \frac{1}{2} - \frac{1}{2A} \xrightarrow{A \rightarrow +\infty} \frac{1}{2}.$$

On en déduit que

$$\int_{-\infty}^{+\infty} f(t) dt = 2 \int_0^{+\infty} f(t) dt = 2 \int_1^{+\infty} f(t) dt = 2 \cdot \frac{1}{2} = 1.$$

Ainsi, f est une densité de probabilité.

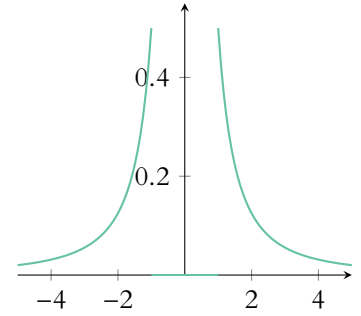


FIGURE 21.1– La densité f .

2. Soit $x \leq 1$. Alors

$$F(x) = \int_{-\infty}^x f(t) dt = \int_{-\infty}^x \frac{1}{2t^2} dt = \lim_{A \rightarrow -\infty} \int_A^x \frac{1}{2t^2} dt = \lim_{A \rightarrow -\infty} \frac{1}{2A} - \frac{1}{2x} = -\frac{1}{2x}.$$

Si $-1 \leq x \leq 1$, alors

$$F(x) = \int_{-\infty}^x f(t) dt = \int_{-\infty}^{-1} f(t) dt = F(-1) = \frac{1}{2}.$$

Enfin, pour $x \geq 1$, on a

$$F(x) = \int_{-\infty}^x f(t) dt = \underbrace{\int_{-\infty}^{-1} f(t) dt}_{=\frac{1}{2}} + \underbrace{\int_{-1}^1 f(t) dt}_{=0} + \int_1^x f(t) dt = \frac{1}{2} + \frac{1}{2} - \frac{1}{2x} = 1 - \frac{1}{2x}.$$

En résumé, on a

$$F(x) = \begin{cases} -\frac{1}{2x} & \text{si } x \leq -1 \\ \frac{1}{2} & \text{si } -1 \leq x \leq 1 \\ 1 - \frac{1}{2x} & \text{si } x \geq 1 \end{cases}$$

3. Soit $x \in \mathbf{R}$. Alors on a

$$[Y_n \leq x] = [S_n \leq nx] = \bigcap_{i=1}^n [X_i \leq nx].$$

Par indépendance des X_i , il vient alors

$$G_n(x) = P(Y_n \leq x) = \prod_{i=1}^n P(X_i \leq nx) = F(nx)^n = \begin{cases} \left(\frac{-1}{2nx}\right)^n & \text{si } nx \leq -1 \\ \frac{1}{2^n} & \text{si } -1 \leq nx \leq 1 \\ \left(1 - \frac{1}{2nx}\right)^n & \text{si } nx \geq 1 \end{cases} = \begin{cases} \left(\frac{-1}{2nx}\right)^n & \text{si } x \leq \frac{-1}{n} \\ \frac{1}{2^n} & \text{si } \frac{-1}{n} \leq x \leq \frac{1}{n} \\ \left(1 - \frac{1}{2nx}\right)^n & \text{si } x \geq \frac{1}{n} \end{cases}$$

- 4.a. G_n est croissante, comme toute fonction de répartition, et on a $G_n(0) = \frac{1}{2^n}$, donc pour tout $x \leq 0$, il vient $G_n(x) \leq G_n(0) = \frac{1}{2^n}$.

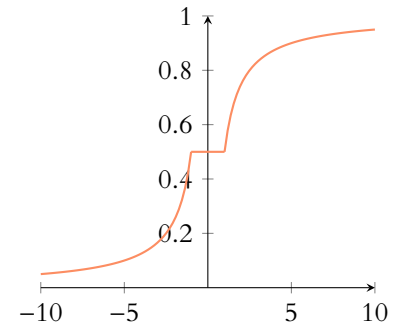


FIGURE 21.2– La fonction de répartition F .

4.b. On a $x \geq \frac{1}{n} \Leftrightarrow n \geq \frac{1}{x}$, donc si on pose $n_0 = \left\lceil \frac{1}{x} \right\rceil + 1$, alors pour $n \geq n_0$, on a bien $x \geq \frac{1}{n}$.

Et donc, d'après le résultat de la question 3, pour $n \geq n_0$, il vient $G_n(x) = \left(1 - \frac{1}{2nx}\right)^n$.

5.a. Pour $x \leq 0$, on a $0 \leq G_n(x) \leq \frac{1}{2^n}$, donc par le théorème des gendarmes, $\lim_{n \rightarrow +\infty} G_n(x) = 0$.

Pour $x > 0$, on a, pour n suffisamment grand, $G_n(x) = \left(1 - \frac{1}{2nx}\right)^n = e^{n \ln\left(1 - \frac{1}{2nx}\right)}$.

Or, $\frac{1}{2nx} \xrightarrow{n \rightarrow +\infty} 0$ et donc

$$\ln\left(1 - \frac{1}{2nx}\right) \underset{n \rightarrow +\infty}{\sim} -\frac{1}{2nx} \Rightarrow n \ln\left(1 - \frac{1}{2nx}\right) \underset{n \rightarrow +\infty}{\sim} -\frac{1}{2x} \xrightarrow{n \rightarrow +\infty} -\frac{1}{2x}.$$

Et alors, par composition de limites¹, il vient

$$\left(1 - \frac{1}{2nx}\right)^n \xrightarrow{n \rightarrow +\infty} e^{-\frac{1}{2x}}.$$

5.b. On a prouvé que $G(x) = \begin{cases} 0 & \text{si } x \leq 0 \\ e^{-\frac{1}{2x}} & \text{si } x > 0 \end{cases}$

On a bien $\lim_{x \rightarrow -\infty} G(x) = 0$ et $\lim_{x \rightarrow +\infty} G(x) = e^0 = 1$.

De plus, G est croissante, elle est évidemment \mathcal{C}^1 sur \mathbf{R}_- et sur \mathbf{R}_+ , et

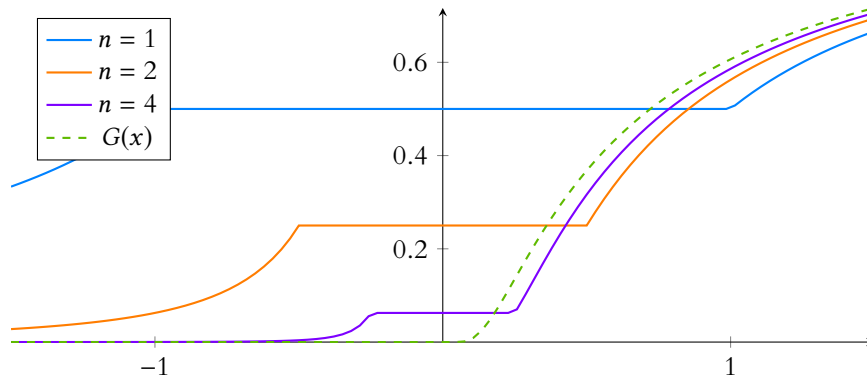
$$\lim_{x \rightarrow 0^-} G(x) = 0 = \lim_{x \rightarrow 0^+} e^{-\frac{1}{2x}}.$$

Donc G est continue² sur \mathbf{R} .

Ainsi, G est bien la fonction de répartition d'une variable à densité.

5.c. Nous avons montré que pour tout $x \in \mathbf{R}$ (et donc en tout point de continuité de G), on a

$\lim_{n \rightarrow +\infty} G_n(x) = G(x)$. Ainsi, $\boxed{Y_n \xrightarrow{\mathcal{L}} Y}$, où Y admet G comme fonction de répartition.



6. Y est à valeurs positives, et donc il en est de même de $\frac{1}{Y}$.

Ainsi, pour $x \leq 0$, $P\left(\frac{1}{Y} \leq x\right) = 0$.

Et pour $x > 0$, on a, par croissance de la fonction inverse sur \mathbf{R}_+ ,

$$P\left(\frac{1}{Y} \leq x\right) = P\left(Y \geq \frac{1}{x}\right) = 1 - P\left(Y \leq \frac{1}{x}\right) = 1 - e^{-x/2}.$$

Donc la fonction de répartition de $\frac{1}{Y}$ est $x \mapsto \begin{cases} 0 & \text{si } x \leq 0 \\ 1 - e^{-x/2} & \text{si } x > 0 \end{cases}$

On reconnaît alors la fonction de répartition d'une loi exponentielle de paramètre $\frac{1}{2}$, donc

$$\boxed{\frac{1}{Y} \hookrightarrow \mathcal{E}\left(\frac{1}{2}\right)}.$$

¹ Et surtout pas d'équivalents, on ne compose pas les équivalents à gauche !

Fct de répartition

Nous ne savons rien de la fonction G , et il est donc important de vérifier que G est bien une fonction de répartition, et donc de s'intéresser à la croissance et aux limites en $\pm\infty$, il ne suffit pas de se contenter de vérifier qu'elle est continue sur \mathbf{R} et \mathcal{C}^1 sauf en un nombre fini de points.

² En particulier, elle est continue à droite en tout point, propriété qui doit être vérifiée par toute fonction de répartition.

Remarque

La fonction de répartition G est nulle sur \mathbf{R}_- , ce qui signifie que Y est à valeurs dans \mathbf{R}_+ . Comme de plus elle est à densité, elle ne prend la valeur 0 qu'avec probabilité nulle, et donc on peut $\frac{1}{Y}$ à bien un sens.

Exercices. 674
Correction des exercices 678
EDHEC 2008 696
Correction 697
EDHEC 2012 699
Correction 700
EML 2003 702
Correction 704

EXERCICES DU CHAPITRE 22

POLYNÔMES

EXERCICE 22.1 Division euclidienne

Dans les cas suivants, déterminer le reste de la division euclidienne de P par Q .

1. $P = 3X^5 + 4X^2 + 1, Q = X^2 + 2X + 3.$
2. $P = X^n - 4X + 1, Q = (X - 1)^2.$

EXERCICE 22.2 Racines réelles d'un polynôme

Pour tout $n \in \mathbf{N}^*$, on pose $P_n = nX^{n+2} - (n+2)X^{n+1} + (n+2)X - n.$

- 1) Montrer que P_n est divisible par $(X - 1)^3.$
- 2) Quelles sont les racines réelles de $P_n'' ?$
- 3) Montrer que P_n admet au plus 4 racines réelles distinctes.

EXERCICE 22.3 (Extrait de ECRICOME 2011)

Soit n un entier naturel non nul. On considère $E = \mathbf{R}_n[X]$ l'espace vectoriel des polynômes de degré inférieur ou égal à $n.$ On définit la famille de polynômes $(P_k)_{0 \leq k \leq n}$ par

$$P_0(X) = 1 \text{ et } \forall k \in \llbracket 1, n \rrbracket, P_k(X) = \frac{X(X-k)^{k-1}}{k!}.$$

- 1) Montrer que la famille (P_0, P_1, \dots, P_n) est une base de $E.$
- 2) Montrer que pour tout entier k de $\llbracket 1, n \rrbracket,$ on a $P_k'(X+1) = P_{k-1}(X).$
Puis, pour tous les entiers (k, j) vérifiant $1 \leq j \leq k \leq n,$ donner une relation entre $P_k^{(j)}(X)$ et $P_{k-j}(X-j).$
- 3) Soit $P \in E.$ Justifier l'existence d'un $(n+1)$ -uplet $(a_0, a_1, \dots, a_n) \in \mathbf{R}^{n+1}$ tel que $P = \sum_{k=0}^n a_k P_k.$
Établir que $\forall j \in \llbracket 0, n \rrbracket, P^{(j)}(j) = a_j.$

EXERCICE 22.4 (QSP ESCP 2016)

- 1) Montrer que la suite $(u_n)_n$ définie par $u_0 = 0$ et $\forall n \in \mathbf{N}, u_{n+1} = u_n^2 + 1$ est strictement croissante.
- 2) En déduire que si P est un polynôme réel tel que $P(0) = 0$ et $P(X^2 + 1) = P(X)^2 + 1,$ alors $P = X.$

ESPACES VECTORIELS, APPLICATIONS LINÉAIRES

EXERCICE 22.5 Soit $H = \left\{ P \in \mathbf{R}_n[X] : \int_0^1 P(t) dt = 0 \right\}.$

Montrer que H est un sous-espace vectoriel de $\mathbf{R}_n[X],$ en donner une base et sa dimension.

EXERCICE 22.6 Matrice d'un projecteur

- 1) Montrer que $F = \{(x, y, z) \in \mathbf{R}^3 : x + 2y + z = 0\}$ et $G = \text{Vect}((1, 1, 1))$ sont supplémentaires dans $\mathbf{R}^3,$ et en donner des bases.
- 2) On note p la projection sur F parallèlement à $G.$
 - a) Déterminer la matrice de p dans une base adaptée à la somme directe $F \oplus G.$
 - b) En déduire la matrice de p dans la base canonique.
 - c) (★) Sans calculs supplémentaires, donner la matrice dans la base canonique de la projection sur G parallèlement à $F.$

EXERCICE 22.7 Soit E un espace vectoriel, p et q deux projecteurs de $\mathcal{L}(E).$

- 1) Montrer que $p \circ q = q \iff \text{Im } q \subset \text{Im } p.$
- 2) Montrer que $p \circ q = p \iff \text{Ker } q \subset \text{Ker } p.$

EXERCICE 22.8 (Extrait de EDHEC 2012)

On note E l'espace vectoriel des polynômes à coefficients réels, de degré inférieur ou égal à 2. On note e_0, e_1 et e_2 les polynômes de E définis par $e_0 = 1, e_1 = X$ et $e_2 = X^2.$

On rappelle que $\mathcal{B} = (e_0, e_1, e_2)$ est une base de $E.$

On considère l'application f qui, à tout polynôme P de E , associe le reste dans la division par $1 + X^3$ du polynôme $(1 - X + X^2)P$. Ainsi, il existe un unique polynôme Q tel que

$$(1 - X + X^2)P = (1 + X^3)Q + f(P), \text{ avec } \deg(f(P)) \leq 2.$$

- 1) Montrer que f est un endomorphisme de E .
- 2) Déterminer $f(e_0)$, $f(e_1)$ et $f(e_2)$ puis vérifier que $f(e_0) = -f(e_1) = f(e_2)$.
- 3) En déduire une base de $\text{Im } f$.
- 4) Donner la dimension de $\text{Ker } f$ ainsi qu'une base de $\text{Ker } f$.

EXERCICE 22.9 Soit f un endomorphisme de \mathbf{R}^n . Soit \mathcal{B} une base de \mathbf{R}^n et $M = \text{Mat}_{\mathcal{B}}(f)$. On suppose que $\text{rg}(M) = \text{tr}(M) = 1$.

$$1) \text{ Montrer qu'il existe une base } \mathcal{B}' \text{ de } \mathbf{R}^n \text{ telle que } \text{Mat}_{\mathcal{B}'}(f) = \begin{pmatrix} 0 & \dots & 0 & a_1 \\ \vdots & & \vdots & a_2 \\ \vdots & & \vdots & \vdots \\ 0 & \dots & 0 & a_n \end{pmatrix}.$$

- 2) En déduire que f est un projecteur.

EXERCICE 22.10 Sous-espaces vectoriels de $\mathbf{R}_n[X]$ stables par la dérivation

Soit $n \in \mathbf{N}^*$ et soit φ l'endomorphisme de $\mathbf{R}_n[X]$ défini par $\varphi(P) = P'$.

- 1) Montrer que pour tout $k \in \llbracket 0, n \rrbracket$, $\mathbf{R}_k[X]$ est un sous-espace vectoriel de $\mathbf{R}_n[X]$ stable par φ .
- 2) Soit F un sous-espace vectoriel de $\mathbf{R}_n[X]$, non réduit au vecteur nul et stable par φ .
 - a) Soit $P \in F$ un polynôme de degré d . Montrer que $\mathbf{R}_d[X] \subset F$.
 - b) On note $p = \max\{\deg P, P \in F\}$. Montrer que $F = \mathbf{R}_p[X]$.

VALEURS PROPRES, DIAGONALISATION.

EXERCICE 22.11 On pose $A = \begin{pmatrix} 3 & 2 \\ 2 & 3 \end{pmatrix}$ et $B = \begin{pmatrix} 0 & -1 \\ 5 & 6 \end{pmatrix}$.

- 1) A et B sont-elles diagonalisables ?
- 2) Montrer que A et B sont semblables.

EXERCICE 22.12 Soit f un endomorphisme de \mathbf{R}^n , de rang 1.

- 1) Montrer que f possède au plus une valeur propre non nulle.
- 2) En déduire qu'au moins l'un des deux endomorphismes $f - \text{id}$ et $f + \text{id}$ est bijectif.

EXERCICE 22.13 Soit $n \in \mathbf{N}^*$ et soit $f : \mathbf{R}_n[X] \rightarrow \mathbf{R}_n[X]$ l'application définie par $f(P) = P(1 - X)$.

- 1) Montrer que $f \in \mathcal{L}(\mathbf{R}_n[X])$.
- 2) Déterminer les valeurs propres de f .
- 3) Pour $k \in \llbracket 0, n \rrbracket$, calculer $f\left(\left(X - \frac{1}{2}\right)^k\right)$. En déduire la dimension des sous-espaces propres de f , et donner une base de chaque sous-espace propre.

EXERCICE 22.14

- 1) Soit $f \in \mathcal{L}(\mathbf{R}^n)$ diagonalisable. Montrer que $E = \text{Ker } f \oplus \text{Im } f$ et que $\text{Im}(f) = \bigoplus_{\substack{\lambda \in \text{Spec}(f) \\ \lambda \neq 0}} E_\lambda(f)$.
- 2) Déterminer tous les endomorphismes diagonalisables $f \in \mathcal{L}(\mathbf{R}^n)$ tels que $\text{Im } f \subset \text{Ker } f$.

EXERCICE 22.15 Endomorphismes laissant stables toutes les droites

Soit E un espace vectoriel de dimension finie n , et soit $f \in \mathcal{L}(E)$. On suppose que pour tout sous-espace vectoriel F de E vérifiant $\dim F = 1$, F est stable par f . On note $\mathcal{B} = (e_1, \dots, e_n)$ une base de E .

- 1) Montrer que tout vecteur non nul de E est un vecteur propre de f .
- 2) En déduire que la matrice de f dans la base \mathcal{B} est diagonale.
- 3) En calculant $f(e_1 + \dots + e_n)$ de deux manières différentes, montrer qu'il existe un réel $\lambda \in \mathbf{R}$ tel que $f = \lambda \text{id}_E$.

EXERCICE 22.16 Soit $A = \begin{pmatrix} 2 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 3 & -2 \end{pmatrix}$.

AD

- 1) Montrer que f est diagonalisable et déterminer une matrice D diagonale et une matrice P inversible telles que $A = P^{-1}DP$.
- 2) Soit f l'endomorphisme de \mathbf{R}^3 dont la matrice dans la base canonique est A , et soit $g \in \mathcal{L}(\mathbf{R}^3)$ tel que $g^3 + 2g = f$.
 - a) Montrer que f et g commutent. En déduire que les sous-espaces propres de f sont stables par g .
 - b) Déterminer la matrice de g dans une base de \mathbf{R}^3 formée de vecteurs propres de f .
 - c) Déterminer tous les endomorphismes $g \in \mathcal{L}(\mathbf{R}^3)$ tels que $g^3 + 2g = f$.

EXERCICE 22.17 Racines carrées d'une matrice diagonalisable (Extrait de EML 2009)

AD

- 1) Soient f et g deux endomorphismes de \mathbf{R}^n vérifiant $f \circ g = g \circ f$. On suppose de plus que f admet n valeurs propres réelles deux à deux distinctes.
 - a) Montrer que chaque sous-espace propre de f est stable par g .
 - b) En déduire que tout vecteur propre de f est vecteur propre de g .
 - c) Justifier que f est diagonalisable.
Montrer que, pour toute base \mathcal{B} de \mathbf{R}^n constituée de vecteurs propres de f , la matrice associée à g relativement à la base \mathcal{B} est diagonale. En déduire que g est diagonalisable.
- 2) Soit A une matrice de $\mathcal{M}_n(\mathbf{R})$ admettant n valeurs propres réelles strictement positives et deux à deux distinctes. On appelle racine carrée de A toute matrice $B \in \mathcal{M}_n(\mathbf{R})$ telle que $B^2 = A$.
 - a) Justifier l'existence d'une matrice inversible P de $\mathcal{M}_n(\mathbf{R})$ telle que la matrice $D = P^{-1}AP$ soit diagonale.
 - b) Donner un exemple de racine carrée de A (on l'exprimera à l'aide de P et des éléments diagonaux de D).
 - c) Soit R une racine carrée de A . Vérifier que $AR = RA$.
En déduire que la matrice $P^{-1}RP$ est diagonale.
 - d) Établir qu'il existe exactement 2^n racines carrées de A .

EXERCICE 22.18 Soit $A \in \mathcal{M}_n(\mathbf{R})$ une matrice diagonalisable possédant p valeurs propres distinctes.

D

- 1) Montrer que A possède un polynôme annulateur P de degré p , et que tout polynôme non nul annulateur de A est de degré supérieur ou égal à p .
- 2) Montrer que la dimension de $\text{Vect}(I_n, A, A^2, \dots, A^k)$ vaut $k + 1$ si $k < p$ et p sinon.

EXERCICE 22.19 (QSP HEC 2015)

D

Soit $x \in \mathbf{R}$ et soit la matrice $M_x = \begin{pmatrix} 1 & x & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$.

- 1) Pour quelles valeurs de x la matrice M_x est-elle diagonalisable ?
- 2) Montrer que lorsqu'elle n'est pas diagonalisable, M_x est semblable à la matrice $B = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$.

ESPACES EUCLIDIENS, ENDOMORPHISMES SYMÉTRIQUES

EXERCICE 22.20 (Extrait de EML 2011)

PD

Soit $n \geq 1$, et soit $E = \mathbf{R}_n[X]$.

- 1) Montrer que pour tout $(P, Q) \in E^2$, $\int_0^{+\infty} P(t)Q(t)e^{-t} dt$ converge.

On définit alors une application φ sur $E \times E$ par $\varphi(P, Q) = \int_0^{+\infty} P(t)Q(t)e^{-t} dt$.

- 2) Montrer que φ est un produit scalaire.
- 3) Soit f l'application qui à $P \in E$ associe $f(P) = XP''(X) - (X - 1)P'(X)$.
 - a) Montrer que f est un endomorphisme de E .
 - b) Montrer que pour tout $(P, Q) \in E^2$,

$$\varphi(f(P), Q) = - \int_0^{+\infty} xP'(x)Q'(x)e^{-x} dx.$$

En déduire que f est un endomorphisme symétrique pour le produit scalaire φ .

EXERCICE 22.21 Soit E un espace euclidien et soient F_1, \dots, F_n des sous-espaces vectoriels deux à deux orthogonaux de E . Montrer que la somme $F_1 + \dots + F_n$ est directe

PD

EXERCICE 22.22 Projecteurs orthogonaux

PD

Pour $A, B \in \mathcal{M}_2(\mathbf{R})$, on pose $\langle A, B \rangle = \text{tr}({}^t AB)$. On note $F = \left\{ \begin{pmatrix} a & b \\ b & -a \end{pmatrix} \in \mathcal{M}_2(\mathbf{R}), (a, b) \in \mathbf{R}^2 \right\}$.

- 1) Montrer que $\langle \cdot, \cdot \rangle$ est un produit scalaire sur $\mathcal{M}_2(\mathbf{R})$.
- 2) Montrer que F est un sous-espace vectoriel de $\mathcal{M}_2(\mathbf{R})$, en déterminer une base et la dimension.
- 3) Déterminer une base de F^\perp .
- 4) On pose $A = \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix}$. Déterminer le projeté orthogonal de A sur F et en déduire $\min_{B \in F} \|A - B\|$.

EXERCICE 22.23 Soit E un espace euclidien de dimension n , et soit (e_1, \dots, e_n) une base orthonormée de E . On note $F = \text{Vect}(e_1 + e_2 + \dots + e_n)$. Déterminer la distance de e_1 à F , c'est-à-dire $\min_{x \in F} \|e_1 - x\|$.

AD

EXERCICE 22.24 Soit E un espace euclidien et soit $f \in \mathcal{L}(E)$ un endomorphisme symétrique. Montrer qu'il y a équivalence entre

AD

- i) $\forall \lambda \in \text{Spec}(f), |\lambda| \leq 1$
- ii) $\forall x \in E, \|f(x)\| \leq \|x\|$.

EXERCICE 22.25 Soit E un espace euclidien, et soit $u \in E$ non nul.

PD

On définit une application $f : E \rightarrow E$ par $f(x) = 2 \frac{\langle x, u \rangle}{\|u\|^2} u - x$.

- 1) Montrer que f est linéaire et que $f \circ f = \text{id}_E$.
- 2) En déduire les valeurs propres et les sous-espaces propres de f . f est-il diagonalisable ?

EXERCICE 22.26

AD

- 1) Soit $A \in \mathcal{M}_n(\mathbf{R})$. On pose $S = {}^t AA$. Montrer que S est symétrique, et que toutes ses valeurs propres sont positives ou nulles.
- 2) Soit S une matrice symétrique dont toutes les valeurs propres sont positives ou nulles. Montrer qu'il existe $A \in \mathcal{M}_n(\mathbf{R})$ telle que $S = {}^t AA$.
- 3) **Application** : soit $S = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$. Déterminer une matrice $A \in \mathcal{M}_2(\mathbf{R})$ telle que $S = {}^t AA$.

EXERCICE 22.27 (Oral HEC 2016)

TD

On considère un espace euclidien $(E, (\cdot | \cdot))$ de dimension $n \geq 2$.

On note $T(E)$ l'ensemble des endomorphismes u de E qui sont symétriques, de rang inférieur ou égal à 1 et tels que pour tout $x \in E, (x|u(x)) \geq 0$.

- 1) Si $a \in E$, on note u_a l'application de E dans E qui à tout vecteur x de E associe $u_a(x) = (x|a)a$.
 - a) Montrer que pour tout $a \in E, u_a \in T(E)$.
 - b) Si a est un vecteur non nul de E , préciser les valeurs propres et sous-espaces propres de u_a .
- 2) Soit u un élément non nul de $T(E)$ et b un vecteur non nul de $\text{Im}(u)$.
 - a) Montrer que b est un vecteur propre de u associé à une valeur propre $\mu \geq 0$.
 - b) Montrer que pour tout vecteur x de $E, u(x) = \frac{\mu}{\|b\|^2} (x|b)b$.
 - c) En déduire qu'il existe un vecteur a de E tel que $u = u_a$.
 - d) L'application φ de E dans $T(E)$ qui à a associe $\varphi(a) = u_a$ est-elle surjective ? Injective ?
- 3) Soit a un vecteur non nul de E et f un endomorphisme de E .
 - a) Pour $x \in E$, expliciter $f \circ u_a(x)$.
 - b) Montrer que $f \circ u_a$ est symétrique si et seulement si a est vecteur propre de f .
 - c) Donner une condition nécessaire et suffisante pour que $f \circ u_a$ appartient à $T(E)$.
- 4) Donner une condition nécessaire et suffisante sur $(a, b) \in E^2$ pour que $u_a \circ u_b = u_b \circ u_a$.

CORRECTION DES EXERCICES DU CHAPITRE 22

SOLUTION DE L'EXERCICE 22.1

1. Posons la division euclidienne :

$$\begin{array}{r|rrrr}
 3X^5 & & & +4X^2 & +1 \\
 -3X^5 & -6X^4 & -9X^3 & & \\
 \hline
 & -6X^4 & -9X^3 & +4X^2 & +1 \\
 & 6X^4 & +12X^3 & +18X^2 & \\
 \hline
 & & 3X^3 & +22X^2 & +1 \\
 & & -3X^3 & -6X^2 & -9X \\
 \hline
 & & & 16X^2 & -9X & +1 \\
 & & & -16X^2 & -32X & -48 \\
 \hline
 & & & & -41X & -47
 \end{array}$$

2. Notons que Q possède 1 comme racine double, de sorte que $Q(1) = Q'(1) = 0$. Soient alors D et R le quotient et le reste de la division euclidienne de P par Q , de sorte que $P = DQ + R$, avec $R = aX + b \in \mathbf{R}_1[X]$. Alors $P(1) = Q(1)D(1) + a + b = a + b \Leftrightarrow a + b = -2$. De plus, on a $P' = Q'D + QD' + R'$ et donc

$$P'(1) = \underbrace{Q'(1)D(1) + D'(1)Q(1)}_{=0} + a \Leftrightarrow a = n - 4.$$

$$\text{Et donc } \begin{cases} a + b = -2 \\ a = n - 4 \end{cases} \Leftrightarrow \begin{cases} a = n - 4 \\ b = 2 - n \end{cases}.$$

On en déduit donc que le reste de la division euclidienne de P par Q est $(n - 4)X + 2 - n$.

SOLUTION DE L'EXERCICE 22.2

- 1.
- P_n
- est divisible par
- $(X - 1)^3$
- si et seulement si 1 est racine d'ordre au moins 3 de
- P_n
- , c'est-à-dire si et seulement si
- $P_n(1) = P'_n(1) = P''_n(1) = 0$
- .

Or on a $P'_n(X) = n(n + 2)X^{n+1} - (n + 1)(n + 2)X^n + n + 2$ et

$P''_n(X) = n(n + 1)(n + 2)X^n - n(n + 1)(n + 2)X^{n-1}$, de sorte que

$$P_n(1) = n - (n + 2) + (n + 2) - n = 0, \quad P'_n(1) = n(n + 2) - (n + 1)(n + 2) + n + 2 = (n + 2)(n - (n + 1) + 1) = 0$$

et $P''_n(1) = n(n + 1)(n + 2) - n(n + 1)(n + 2) = 0$.

Ainsi, 1 est bien racine de P_n d'ordre au moins 3, et donc P_n est divisible par $(X - 1)^3$.

2. On a $P'_n(X) = n(n + 1)(n + 2)X^{n-1}(X - 1)$ qui possède donc uniquement 0 et 1 comme racines réelles¹.
3. Supposons au contraire que P_n possède au moins 5 racines réelles $\lambda_1 < \lambda_2 < \lambda_3 < \lambda_4 < \lambda_5$. Puisque P_n est dérivable sur $[\lambda_1, \lambda_2]$, avec $P_n(\lambda_1) = P_n(\lambda_2) = 0$, alors, par le théorème de Rolle, P'_n possède au moins une racine dans l'intervalle $]\lambda_1, \lambda_2[$. Et de même dans les intervalles $]\lambda_2, \lambda_3[$, \dots , $]\lambda_4, \lambda_5[$. Autrement dit, P'_n possède au moins 4 racines réelles distinctes. Et donc, toujours par le théorème de Rolle, P''_n possède au moins trois racines réelles, ce qui contredit la question 2. Ainsi, P_n possède au plus 4 racines réelles distinctes.

SOLUTION DE L'EXERCICE 22.3

1. Notons que P_k est de degré égal à k . Donc (P_0, \dots, P_n) est une famille de polynômes de degrés deux à deux distincts : elle est libre. Comme elle est de cardinal $n + 1 = \dim E$, c'est donc une base de E .
2. On a $P'_1(X) = 1$, et pour $k \in \llbracket 2, n \rrbracket$, on a

$$P'_k(X) = \frac{(X - k)^{k-1} + (k - 1)X(X - k)^{k-2}}{k!} = \frac{(X - k)^{k-2}(X - k + (k - 1)X)}{k!} = \frac{(X - k)^{k-2}(k(X - 1))}{k!} = \frac{(X - 1)(X - k)^{k-2}}{(k - 1)!}.$$

Méthode

Pour éviter les erreurs de calcul, bien aligner les puissances dans la division euclidienne : une colonne pour les termes en X , une pour les termes en X^2 , etc.

Rappel

Le degré du reste est strictement inférieur au degré du polynôme par lequel on divise.

Mieux !

Il est aisé de voir que $P_n^{(3)}(1) \neq 0$, et donc 1 est racine d'ordre exactement 3 de P_n .

¹ Et même complexes !

Rolle

Ces raisonnements à l'aide du théorème de Rolle sont très classiques. C'est par exemple ce qui permet de prouver que si $P \in \mathbf{R}[X]$ est scindé à racines simples (i.e. possède autant de racines réelles distinctes que son degré), alors P' est encore scindé à racines simples. Notons que ceci ne fonctionne que pour des racines réelles, et non des racines complexes, le théorème de Rolle ne s'appliquant plus.

Et donc, $P'_k(X+1) = 1 = P_0(X)$, et pour $2 \leq k \leq n$,

$$P'_k(X+1) = \frac{X(X+1-k)^{k-2}}{(k-1)!} = \frac{X(X-(k-1))^{k-2}}{(k-1)!} = P_{k-1}(X).$$

En remplaçant X par $X-1$, il vient $P'_k(X) = P_{k-1}(X-1)$.

En dérivant cette égalité², il vient alors $P''_k(X) = P'_{k-1}(X-1) = P_{k-2}((X-1)-1) = P_{k-2}(X-2)$.

Ainsi, on prouve, de proche en proche³ que $\forall 1 \leq j \leq k \leq n$, $P_k^{(j)}(X) = P_{k-j}(X-j)$.

² Pour $k \geq 2$.

³ Une récurrence finie serait plus rigoureuse.

3. Puisque (P_0, \dots, P_n) est une base de E , tout polynôme P s'écrit de manière unique $P = \sum_{k=0}^n a_k P_k$,

avec $a_0, \dots, a_n \in \mathbf{R}$.

En évaluant cette relation en $X = 0$, il vient

$$P(0) = a_0 P_0(0) + \sum_{k=1}^n a_k P_k(0) = a_0.$$

En dérivant j fois la relation $P = \sum a_k P_k$, il vient

$$P^{(j)}(X) = \sum_{k=0}^n a_k P_k^{(j)}(X) = \sum_{k=j}^n a_k P_k^{(j)}(X) = \sum_{k=j}^n a_k P_{k-j}(X-j).$$

En évaluant en $X = j$, il vient donc $P^{(j)}(j) = \sum_{k=j}^n a_k P_{k-j}(0)$.

Mais pour $k > j$, $P_{k-j}(0) = 0$, et pour $k = j$, $P_{k-j}(0) = P_0(0) = 1$.

On a donc bien, pour tout $j \in \llbracket 1, n \rrbracket$, $a_j = P^{(j)}(j)$.

SOLUTION DE L'EXERCICE 22.4

1. Il est évident que (u_n) est une suite d'entiers positifs, et que si $n \geq 1$, alors $u_n > 0$. Montrons par récurrence sur n que $u_{n+1} \geq u_n$.
Pour $n = 0$, on a $u_0 = 0$ et $u_1 = 1 > u_0$.
Supposons donc que $u_{n+1} > u_n$. Alors

$$u_{n+2} = u_{n+1}^2 + 1 \geq u_n + 1 > u_n.$$

Et donc, par le principe de récurrence, pour tout $n \in \mathbf{N}$, $u_{n+1} > u_n$: (u_n) est strictement croissante.

2. On a $P(0) = 0$. Et donc $P(1) = P(0^2 + 1) = P(0)^2 + 1 = 1$.
Puis $P(2) = P(1^2 + 1) = P(1)^2 + 1 = 2$. Et $P(5) = P(2^2 + 1) = P(2)^2 + 1 = 2^2 + 1 = 5$, etc.
Montrons par récurrence que pour tout $n \in \mathbf{N}$, $P(u_n) = u_n$.
Pour $n = 0$, c'est l'hypothèse faite par l'énoncé : $P(0) = 0$.
Supposons donc que $P(u_n) = u_n$.
Alors $P(u_{n+1}) = P(u_n^2 + 1) = P(u_n)^2 + 1 = u_n^2 + 1 = u_{n+1}$.
Par le principe de récurrence, $P(u_n) = u_n$, pour tout $n \in \mathbf{N}$.
Soit alors $Q(X) = P(X) - X$. Alors pour tout $n \in \mathbf{N}$, $Q(u_n) = P(u_n) - u_n = 0$.
Puisque la suite (u_n) est strictement croissante, les u_n sont deux à deux distincts. Et donc Q possède une infinité de racines : tous les $u_n, n \in \mathbf{N}$.
Par conséquent, $Q = 0$, et donc $P(X) = X$.

SOLUTION DE L'EXERCICE 22.5

Plutôt que de prouver à l'aide du théorème de caractérisation que H est un sous-espace vectoriel de $\mathbf{R}_n[X]$, remarquons que $\varphi : R \mapsto \int_0^1 P(t) dt$ est une forme linéaire sur $\mathbf{R}_n[X]$, non nulle car $\varphi(1) = 1$.

Et donc $H = \text{Ker } \varphi$ est un hyperplan de $\mathbf{R}_n[X]$, ce qui prouve à la fois que H est un sous-espace vectoriel de $\mathbf{R}_n[X]$ et que $\dim H = \dim \mathbf{R}_n[X] - 1 = n$.

Soit $P = \sum_{k=0}^n a_k X^k \in \mathbf{R}_n[X]$. Alors on a

$$P \in H \Leftrightarrow \int_0^1 \left(\sum_{k=0}^n a_k t^k \right) dt = 0$$

Détails

Les P_k avec $k < j$ sont de degré strictement inférieur à j , donc leur dérivée $j^{\text{ème}}$ est nulle.

Remarque

Notons que cette méthode nous donne la dimension, mais ne nous aide pas à obtenir une base de H .

$$\begin{aligned} \Leftrightarrow \sum_{k=0}^n \frac{a_k}{k+1} &= 0 \\ \Leftrightarrow a_0 &= -\sum_{k=1}^n \frac{a_k}{k+1} \\ \Leftrightarrow P &= \sum_{k=1}^n a_k \left(X^k - \frac{1}{k+1} \right) \\ \Leftrightarrow P &\in \text{Vect} \left(X - \frac{1}{2}, X^2 - \frac{1}{3}, \dots, X^n - \frac{1}{n+1} \right). \end{aligned}$$

Classique

$$\int_0^1 t^k dt = \frac{1}{k+1}.$$

Ainsi, la famille $\left(X - \frac{1}{2}, X^2 - \frac{1}{3}, \dots, X^n - \frac{1}{n+1} \right)$ est génératrice de H , et elle est libre car formée de polynômes de degrés deux à deux distincts : c'est une base de H .

SOLUTION DE L'EXERCICE 22.6

1. On a $(x, y, z) \in F$ si et seulement si

$$x+2y+z = 0 \Leftrightarrow x = -2y-z \Leftrightarrow (x, y, z) = (-2y-z, y, z) \Leftrightarrow (x, y, z) \in \text{Vect}((-2, 1, 0), (-1, 0, 1)).$$

Et donc $F = \text{Vect}((-2, 1, 0), (-1, 0, 1))$ est un sous-espace vectoriel de \mathbf{R}^3 , et la famille $(-2, 1, 0), (-1, 0, 1)$ en est une base⁴

Il est clair que G est de dimension 1, et qu'une base en est la famille formée du seul vecteur $(1, 1, 1)$.

Donc déjà, $\dim F + \dim G = 3 = \dim \mathbf{R}^3$.

Soit donc $x \in F \cap G$.

Puisque $x \in G$, il existe $\lambda \in \mathbf{R}$ tel que $x = (\lambda, \lambda, \lambda)$.

Et puisque $x \in F$, on a donc $\lambda + 2\lambda + \lambda = 0 \Leftrightarrow \lambda = 0$.

Ainsi, $x = 0$ et donc $F \cap G = \{0\}$.

On en déduit que F et G sont supplémentaires dans \mathbf{R}^3 : $F \oplus G = \mathbf{R}^3$.

2.a. C'est du cours : la matrice de p dans une base adaptée à la somme directe $F \oplus G$ est

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

2.b. Une base adaptée à la somme directe $F \oplus G$ est, par définition, la concaténation d'une base de F et d'une base de G .

Une telle base est par exemple $\mathcal{B} = ((-2, 1, 0), (-1, 0, 1), (1, 1, 1))$.

On a alors

$$P_{\mathcal{B}_{can}, \mathcal{B}} = \begin{pmatrix} -2 & -1 & 1 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix}.$$

Et alors, par la formule de changement de base, il vient

$$\text{Mat}_{\mathcal{B}_{can}}(p) = P_{\mathcal{B}_{can}, \mathcal{B}} \text{Mat}_{\mathcal{B}}(p) P_{\mathcal{B}_{can}, \mathcal{B}} = \begin{pmatrix} 3/4 & -1/2 & -1/4 \\ -1/4 & 1/2 & -1/4 \\ -1/4 & -1/2 & 3/4 \end{pmatrix}.$$

2.c. Notons q le projecteur sur G parallèlement à F .

Il serait bien évidemment possible de faire le même type de calculs qu'à la question précédente, en notant que $\text{Mat}_{\mathcal{B}}(q) = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ et en appliquant la formule de changement

de base.

Mais notons plutôt que $p + q = \text{id}_{\mathbf{R}^3}$.

En effet, si $x \in \mathbf{R}^3$, de manière unique, $x = x_F + x_G$, avec $x_F \in F$ et $x_G \in G$.

Et alors $p(x) = x_F, q(x) = x_G$, de sorte que $p(x) + q(x) = x = \text{id}_{\mathbf{R}^3}(x)$.

Et donc $p + q = \text{id}_{\mathbf{R}^3} \Leftrightarrow \text{Mat}_{\mathcal{B}_{can}}(p) + \text{Mat}_{\mathcal{B}_{can}}(q) = I_3$.

On en déduit que

$$\text{Mat}_{\mathcal{B}_{can}}(q) = I_3 - \begin{pmatrix} 3/4 & -1/2 & -1/4 \\ -1/4 & 1/2 & -1/4 \\ -1/4 & -1/2 & 3/4 \end{pmatrix} = \begin{pmatrix} 1/4 & 1/2 & 1/4 \\ 1/4 & 1/2 & 1/4 \\ 1/4 & 1/2 & 1/4 \end{pmatrix}.$$

Remarque
Étant donné que nous savons déjà que H est de dimension n , nous pouvons aussi évoquer un argument de cardinal plutôt que de prouver la liberté.

⁴ Nous venons de prouver qu'elle est génératrice, et elle est libre car les deux vecteurs ne sont pas colinéaires.

Valeurs propres
Notons que nous savons interpréter tout ceci en termes de valeurs propres/-vecteurs propres : on a $E_0(p) = \text{Ker } p = G$ et $E_1(p) = \text{Im}(p) = F$. Et donc une base adaptée à la somme directe $F \oplus G$ est une base de vecteurs propres de p , dans laquelle la matrice de p est bien évidemment diagonale.

$$\text{On a } M = \begin{pmatrix} 3/4 & -1/2 & -1/4 \\ -1/4 & 1/2 & -1/4 \\ -1/4 & -1/2 & 3/4 \end{pmatrix}$$

SOLUTION DE L'EXERCICE 22.7

- Si $\text{Im } q \subset \text{Im } p$, alors pour tout $x \in E$, $q(x) \in \text{Im}(p)$ et donc $p(q(x)) = q(x)$, de sorte que $p \circ q = q$.
Inversement, si $p \circ q = q$, alors $\text{Im}(q) = \text{Im}(p \circ q) \subset \text{Im}(p)$.
- Si $p \circ q = p$. Soit $x \in \text{Ker } q$. Alors $q(x) = 0$ et donc $p(q(x)) = 0$. Or $p(x) = p(q(x))$, donc $x \in \text{Ker } p$.
On en déduit que $\text{Ker } q \subset \text{Ker } p$.
Inversement, si $\text{Ker } q \subset \text{Ker } p$, alors pour $x \in E$, on a $x = x_1 + x_2$, $x_1 \in \text{Im } q$, $x_2 \in \text{Ker } q$ et donc $p(q(x)) = p(x_1)$.
D'autre part, $x_2 \in \text{Ker } p$, de sorte que $p(x_2) = 0$, et donc $p(x) = p(x_1)$.
Ainsi, $p(q(x)) = p(x)$, et donc $p \circ q = p$.

SOLUTION DE L'EXERCICE 22.8

- Il est clair que f est à valeurs dans $\mathbf{R}_2[X] = E$.
Soient $P_1, P_2 \in E$ et $\lambda \in \mathbf{R}$. Alors il existe deux polynômes Q_1 et Q_2 tels que :

$$(1 - X + X^2)P_1 = (1 + X^3)Q_1 + f(P_1) \text{ et } (1 - X + X^2)P_2 = (1 + X^3)Q_2 + f(P_2).$$

Et donc $(1 - X + X^2)(\lambda P_1 + P_2) = (1 + X^3)(\lambda Q_1 + Q_2) + (\lambda f(P_1) + f(P_2))$,
avec $\deg(\lambda f(P_1) + f(P_2)) \leq 2$.

Par unicité de la division euclidienne, $\lambda f(P_1) + f(P_2)$ est alors le reste de la division euclidienne de $(1 - X + X^2)(\lambda P_1 + P_2)$ par $1 + X^3$, donc est égal à $f(\lambda P_1 + P_2)$.

Ainsi, $f(\lambda P_1 + P_2) = \lambda f(P_1) + f(P_2)$: f est linéaire, et est donc un endomorphisme de E .

- On a $(1 - X + X^2) \times 1 = (1 + X^3) \times 0 + 1 - X + X^2$, avec $\deg(1 - X + X^2) \leq 2$, donc $f(e_0) = 1 - X + X^2$.
De même, on a $(1 - X + X^2)X = X^3 - X^2 + X = (1 + X^3) \times 1 + \underbrace{(-X^2 + X - 1)}_{\in E}$, de sorte que

$$f(e_1) = -X^2 + X - 1.$$

Enfin, la division euclidienne de $(1 - X + X^2)X^2 = X^4 - X^3 + X^2$ par $1 + X^3$ est

$$\begin{array}{r|l} X^4 & -X^3 & +X^2 & & X^3 & +1 \\ X^4 & & & +X & X & -1 \\ \hline & -X^3 & +X^2 & -X & & \\ & -X^3 & & & & -1 \\ \hline & & X^2 & -X & +1 & \end{array}$$

de sorte que $f(e_2) = X^2 - X + 1$.

Notons qu'on a alors bien $f(e_0) = -f(e_1) = f(e_2)$.

- Si $P \in E$, alors il existe trois réels $\lambda_0, \lambda_1, \lambda_2$ tels que $P = \lambda_0 e_0 + \lambda_1 e_1 + \lambda_2 e_2$.
Et donc $f(P) = \lambda_0 f(e_0) + \lambda_1 f(e_1) + \lambda_2 f(e_2) = (\lambda_0 - \lambda_1 + \lambda_2)(1 - X + X^2) \in \text{Vect}(1 - X + X^2)$.
Or, si $Q \in \text{Im } f$, alors il existe $P \in E$ tel que $Q = f(P) \in \text{Vect}(1 - X + X^2)$.
Donc $\text{Im } f \subset \text{Vect}(1 - X + X^2)$.
Puisque $1 - X + X^2 = f(e_0) \in \text{Im } f$ on a $\text{Vect}(1 - X + X^2) \subset \text{Im } f$, et donc $\text{Im } f = \text{Vect}(1 - X + X^2)$.
Ainsi, une base⁵ de $\text{Im } f$ est formée du seul polynôme $1 - X + X^2$.
- D'après le théorème du rang, on a $\dim \text{Ker } f = \dim E - \dim \text{Im } f = 3 - 1 = 2$.
De plus, nous savons que $f(e_0 + e_1) = f(e_0) + f(e_1) = f(e_0) - f(e_0) = 0$ et de même $f(e_0 - e_2) = 0$. Donc $e_0 + e_1 = 1 + X$ et $e_0 - e_2 = 1 - X^2$ sont deux éléments de $\text{Ker } f$.
Puisqu'ils sont non nuls et de degré distincts, ils forment une famille libre de $\text{Ker } f$, de cardinal égal à la dimension de $\text{Ker } f$: $(1 + X, 1 - X^2)$ est une base de $\text{Ker } f$.

SOLUTION DE L'EXERCICE 22.9

- Commençons pas analyser l'énoncé. Si on note $\mathcal{B}' = (e_1, \dots, e_n)$ la base cherchée, on doit

Div. euclidienne

Rappelons que si on peut écrire $A = PQ + R$, avec $\deg(R) < \deg(A)$, alors, par unicité de la division euclidienne, Q est le quotient de la division euclidienne de A par B et R est le reste.

Plus généralement

Si $f : E \rightarrow F$ est une application linéaire, alors l'image d'une famille génératrice de E par f est une famille génératrice de $\text{Im } f$.
En revanche, rien ne garantit que l'image d'une base soit une base de l'image, il faut alors vérifier la liberté à la main.
Exercice : montrer que l'image d'une base de E est une base de $\text{Im } f$ si et seulement si f est injective.

⁵ Il s'agit évidemment d'une famille génératrice de $\text{Im } f$, et étant formée d'un seul vecteur **non nul**, elle est libre.

donc avoir

$$\text{Mat}_{\mathcal{B},\mathcal{B}'}(f) = \begin{pmatrix} f(e_1) & \dots & f(e_{n-1}) & f(e_n) \\ 0 & \dots & 0 & a_1 \\ \vdots & & \vdots & a_2 \\ \vdots & & \vdots & \vdots \\ 0 & \dots & 0 & a_n \end{pmatrix} \begin{matrix} e_1 \\ e_2 \\ \vdots \\ e_n \end{matrix}.$$

Il nous faut donc avoir $f(e_1) = \dots = f(e_{n-1}) = 0$, c'est-à-dire $e_1, \dots, e_{n-1} \in \text{Ker}(f)$.

Puisque $\text{rg}(f) = \text{rg}(M) = 1$, on a $\dim \text{Ker}(f) = n - \text{rg}(f) = n - 1$.

Soit donc (e_1, \dots, e_{n-1}) une base de $\text{Ker}(f)$.

Par le théorème de la base incomplète, il est possible de compléter cette famille en une base $(e_1, \dots, e_{n-1}, e_n)$ de E .

Et alors la matrice de f dans cette base est de la forme

$$\begin{pmatrix} f(e_1) & \dots & f(e_{n-1}) & f(e_n) \\ 0 & \dots & 0 & a_1 \\ \vdots & & \vdots & a_2 \\ \vdots & & \vdots & \vdots \\ 0 & \dots & 0 & a_n \end{pmatrix} \begin{matrix} e_1 \\ e_2 \\ \vdots \\ e_n \end{matrix}$$

où a_1, \dots, a_n sont les réels tels que $f(e_n) = a_1 e_1 + \dots + a_n e_n$.

2. Puisque $N = \text{Mat}_{\mathcal{B},\mathcal{B}'}(f)$ est semblable⁶ à M , elles ont même trace. Et donc $a_n = 1$.
On a alors

$$N^2 = \begin{pmatrix} 0 & \dots & 0 & a_1 \\ \vdots & & \vdots & a_2 \\ \vdots & & \vdots & \vdots \\ 0 & \dots & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & \dots & 0 & a_1 \\ \vdots & & \vdots & a_2 \\ \vdots & & \vdots & \vdots \\ 0 & \dots & 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & \dots & 0 & a_1 \\ \vdots & & \vdots & a_2 \\ \vdots & & \vdots & \vdots \\ 0 & \dots & 0 & 1 \end{pmatrix} = N.$$

Et donc $f^2 = f : f$ est un projecteur de \mathbf{R}^n .

SOLUTION DE L'EXERCICE 22.10

1. Soit $P \in \mathbf{R}_k[X]$. Alors $\deg P \leq k$, de sorte que $\deg \varphi(P) = \deg P' \leq k - 1$, et donc $\varphi(P) \in \mathbf{R}_k[X]$.
Ainsi, $\mathbf{R}_k[X]$ est stable par φ .
- 2.a. Puisque F est stable par φ , $P' = \varphi(P) \in F$.
Puis, $P'' = \varphi' \in F$, et de proche en proche, on prouve que pour tout $i \in \llbracket 0, d \rrbracket$, $P^{(i)} \in F$.
Mais $\deg P^{(i)} = \deg P - i = d - i$, de sorte que $(P, P', \dots, P^{(d)})$ est une famille de $\mathbf{R}_d[X]$ formée de polynômes de degrés deux à deux distincts. Elle est donc libre, et étant de cardinal $d + 1 = \dim \mathbf{R}_d[X]$, c'est une base de $\mathbf{R}_d[X]$.
Ainsi, si $Q \in \mathbf{R}_d[X]$, il existe des réels a_0, a_1, \dots, a_d tels que

$$Q = a_0 \underbrace{P}_{\in F} + a_1 \underbrace{P'}_{\in F} + \dots + a_d \underbrace{P^{(d)}}_{\in F} \in F.$$

Et donc $\mathbf{R}_d[X] \subset F$.

- 2.b. Si p est le degré maximal d'un polynôme de F , alors on a évidemment $F \subset \mathbf{R}_p[X]$.
D'autre part, par définition de p , il existe $P \in F$ tel que $\deg P = p$.
Et donc par la question précédente, $\mathbf{R}_p[X] \subset F$.
Par double inclusion, on a donc $F = \mathbf{R}_p[X]$.

SOLUTION DE L'EXERCICE 22.11

1. Soit $\lambda \in \mathbf{R}$. Alors λ est valeur propre de A si et seulement si $A - \lambda I_2$ n'est pas inversible, c'est-à-dire si et seulement si $\det(A - \lambda I_2) = 0 \Leftrightarrow \lambda^2 - 6\lambda + 5 = 0$.
Les racines de ce polynôme sont 1 et 5, et donc A possède deux valeurs propres distinctes, et donc est diagonalisable.
De même, les valeurs propres de B sont 1 et 5, et donc B est également diagonalisable.

Remarque
Et nous n'avons aucune contrainte sur e_n , puisque ma dernière colonne peut être tout à fait quelconque.

⁶ Elle représentent le même endomorphisme.

2. Puisque A est diagonalisable et possède deux valeurs propres, elle est semblable à $D = \begin{pmatrix} 1 & 0 \\ 0 & 5 \end{pmatrix}$, et B est également semblable à D .
Et donc, A et B étant semblables à la même matrice, elles sont semblables.

SOLUTION DE L'EXERCICE 22.12

1. Puisque f est de rang $1 < n$, 0 est valeur propre de f , avec $\dim E_0(f) = n - \text{rg}(f) = n - 1$.
Et puisqu'on doit avoir $\sum_{\lambda \in \text{Spec}(f)} \dim E_\lambda(f) \leq n$, f ne peut avoir au plus qu'une seule valeur propre différente de 0 .
2. Rappelons que $\lambda \in \mathbf{R}$ est valeur propre de f si et seulement si $f - \lambda \text{id}$ n'est pas bijectif.
Or, comme f ne possède au plus qu'une valeur propre non nulle, l'un au moins des deux nombres 1 et -1 n'est pas valeur propre de f .
Et donc l'un au moins des deux endomorphismes $f - \text{id}$ et $f + \text{id}$ est bijectif.

SOLUTION DE L'EXERCICE 22.13

1. Soient $P, Q \in \mathbf{R}_n[X]$ et $\lambda \in \mathbf{R}$.
Alors $f(\lambda P + Q) = (\lambda P + Q)(1 - X) = \lambda P(1 - X) + Q(1 - X) = \lambda f(P) + f(Q)$.
Donc f est linéaire.
D'autre part, pour tout $k \in \llbracket 0, n \rrbracket$, on a $f(X^k) = (1 - X)^k \in \mathbf{R}_n[X]$, de sorte que pour $P = \sum_{k=0}^n a_k X^k \in \mathbf{R}_n[X]$,

$$f(P) = \sum_{k=0}^n a_k (1 - X)^k \in \mathbf{R}_n[X].$$

Et donc f est un endomorphisme de $\mathbf{R}_n[X]$.

2. On a $f(1) = 1$, $f(X) = 1 - X$ et plus généralement, pour $k \in \llbracket 0, n \rrbracket$,

$$f(X^k) = (1 - X)^k = \sum_{i=0}^k \binom{k}{i} (-1)^i X^i.$$

Et donc la matrice de f dans la base canonique de $\mathbf{R}_n[X]$ est

$$\text{Mat}_{\mathcal{B}_{can}}(f) = \begin{pmatrix} f(1) & f(X) & f(X^2) & \dots & f(X^{n-1}) & f(X^n) \\ 1 & 1 & 1 & \dots & \dots & 1 \\ 0 & -1 & -2\alpha & \dots & -\binom{n-1}{1} & -\binom{n}{1} \\ \vdots & \ddots & 1 & \ddots & \vdots & \vdots \\ \vdots & & \ddots & \ddots & \vdots & \vdots \\ \vdots & & & \ddots & (-1)^{n-1} & (-1)^{n-1} \binom{n}{n-1} \\ 0 & \dots & \dots & \dots & 0 & (-1)^n \end{pmatrix} \begin{matrix} 1 \\ X \\ X^2 \\ \vdots \\ X^{n-1} \\ X^n \end{matrix}.$$

Cette matrice est triangulaire supérieure, donc ses valeurs propres sont ses coefficients diagonaux, à savoir -1 et 1 .

Et donc $\text{Spec}(f) = \{-1, 1\}$.

3. On a

$$f\left(\left(X - \frac{1}{2}\right)^k\right) = \left(1 - X - \frac{1}{2}\right)^k = \left(\frac{1}{2} - X\right)^k = (-1)^k \left(X - \frac{1}{2}\right)^k.$$

Puisque $\left(X - \frac{1}{2}\right)^k \neq 0$, il s'agit donc d'un vecteur propre associé à la valeur propre $(-1)^k$.

1^{er} cas : n est impair. Notons alors $n = 2p + 1$.

Par ce qui précède, les $\left(X - \frac{1}{2}\right)^{2k}$, pour $k \in \llbracket 0, p \rrbracket$ sont des vecteurs propres de f associés à la valeur propre 1 .

Et donc le sous-espace vectoriel qu'ils engendrent est inclus dans $E_1(f)$.

Puisqu'il s'agit d'une famille de polynômes de degrés deux à deux distincts, elle est libre, et

Remarque

Un résultat du cours garantit que dès que deux matrices diagonalisables ont les mêmes valeurs propres et des sous-espaces propres de même dimension, alors elles sont semblables.

donc $\dim \text{Vect} \left(1, \left(X - \frac{1}{2} \right)^2, \dots, \left(X - \frac{1}{2} \right)^{2p} \right) = p + 1$.

Et donc $\dim E_1(f) \geq p + 1$.

De même, les $\left(X - \frac{1}{2} \right)^{2k+1}$, avec $k \in \llbracket 0, p \rrbracket$ sont dans $E_{-1}(f)$, et donc $\dim E_{-1}(f) \geq p + 1$.

Or, on doit toujours avoir

$$\dim E_1(f) + \dim E_{-1}(f) \leq n = 2p + 2$$

de sorte que $\dim E_1(f) = \dim E_{-1}(f) = p + 1$.

Et donc une base de $E_1(f)$ est $1, \left(X - \frac{1}{2} \right)^2, \dots, \left(X - \frac{1}{2} \right)^{2p}$, et une base de $E_{-1}(f)$ est $\left(X - \frac{1}{2} \right), \left(X - \frac{1}{2} \right)^3, \dots, \left(X - \frac{1}{2} \right)^{2p+1}$.

2^{ème} cas : n est pair. Notons alors $n = 2p$.

On procède alors de la même manière, sauf que cette fois, on aura $\dim E_1(f) = p + 2$ et $\dim E_{-1}(f) = p + 1$.

Notons que dans les deux cas, on a $\dim E_1(f) + \dim E_{-1}(f) = n + 1 = \dim \mathbf{R}_n[X]$, et donc f est toujours diagonalisable.

SOLUTION DE L'EXERCICE 22.14

1. Nous savons que puisque f est diagonalisable, $E = \bigoplus_{\lambda \in \text{Spec}(f)} E_\lambda(f)$.

Si $0 \notin \text{Spec}(f)$, alors $\text{Ker}(f) = \{0\}$, et donc $E = \bigoplus_{\lambda \in \text{Spec}(f)} E_\lambda(f) = \bigoplus_{\substack{\lambda \in \text{Spec}(f) \\ \lambda \neq 0}} E_\lambda(f)$.

Mais dans ce cas, f est bijectif, et donc $\text{Im}(f) = \mathbf{R}^n$, ce qui prouve les deux résultats.

Dans le cas où 0 est valeur propre de f , alors nous avons $E = E_0(f) \oplus \bigoplus_{\substack{\lambda \in \text{Spec}(f) \\ \lambda \neq 0}} E_\lambda(f) =$

$\text{Ker}(f) \oplus \bigoplus_{\substack{\lambda \in \text{Spec}(f) \\ \lambda \neq 0}} E_\lambda(f)$.

Il s'agit donc de montrer que $\text{Im}(f) = \bigoplus_{\substack{\lambda \in \text{Spec}(f) \\ \lambda \neq 0}} E_\lambda(f)$.

Notons que le fait que des sous-espaces propres distincts soient en somme directe découle directement du cours.

D'autre part, on a $\dim \text{Im}(f) = n - \dim \text{Ker}(f)$ et

$$\dim \left(\bigoplus_{\substack{\lambda \in \text{Spec}(f) \\ \lambda \neq 0}} E_\lambda(f) \right) = \sum_{\substack{\lambda \in \text{Spec}(f) \\ \lambda \neq 0}} \dim E_\lambda(f) = n - \dim E_0(f) = \dim \text{Im}(f).$$

Puisque nous avons affaire à deux espaces de même dimension, montrons que l'un est inclus dans l'autre.

Soit $x \in \bigoplus_{\substack{\lambda \in \text{Spec}(f) \\ \lambda \neq 0}} E_\lambda(f)$. Alors x s'écrit⁷ $x = \sum_{\substack{\lambda \in \text{Spec}(f) \\ \lambda \neq 0}} x_\lambda$, où, pour tout $\lambda \in \text{Spec}(f) \setminus \{0\}$, $x_\lambda \in$

$E_\lambda(f)$.

Et donc

$$x = \sum_{\substack{\lambda \in \text{Spec}(f) \\ \lambda \neq 0}} \frac{\lambda x_\lambda}{\lambda} = \sum_{\substack{\lambda \in \text{Spec}(f) \\ \lambda \neq 0}} \frac{1}{\lambda} f(x) = f \left(\sum_{\substack{\lambda \in \text{Spec}(f) \\ \lambda \neq 0}} \frac{1}{\lambda} x_\lambda \right) \in \text{Im}(f).$$

Ainsi, on a

$$\bigoplus_{\substack{\lambda \in \text{Spec}(f) \\ \lambda \neq 0}} E_\lambda(f) \subset \text{Im}(f)$$

et donc, par égalité des dimensions, ces deux ensembles sont égaux.

Inégalité

Notons que pour l'instant nous n'avons qu'une inégalité sur la dimension de $E_1(f)$: certes nous avons trouvé un certain nombre de vecteurs propres, mais rien n'indique que nous les ayons tous !

Remarque

Le fait que les deux sous-espaces propres n'aient pas même dimension se « devine » sur la matrice, car il y a alors un 1 de plus que de -1 sur la diagonale (toutefois cela ne constitue en rien une preuve car pour une matrice triangulaire, la dimension du sous-espace propre n'est pas nécessairement égale au nombre de fois où la valeur propre apparaît sur la diagonale).

Rappel

La dimension d'une somme directe est la somme des dimensions.

⁷ De manière unique.

2. **Première méthode** : il est possible d'utiliser la question précédente : si $\text{Im}(f) \subset \text{Ker}(f)$, alors $\text{Im}(f) \cap \text{Ker}(f) = \text{Im}(f)$.
Mais nous venons de prouver que $\text{Im}(f)$ et $\text{Ker}(f)$ sont en somme directe, et donc $\text{Im}(f) \cap \text{Ker}(f) = \{0\}$.
On en déduit que $\text{Im}(f) = \{0\}$, et donc f est l'endomorphisme nul.

Seconde méthode : «à la main».

Si $\text{Im} f \subset \text{Ker} f$, alors pour tout $x \in \mathbf{R}^n$, on a $f(x) \in \text{Im}(f)$, et donc $f(x) \in \text{Ker}(f)$. Soit encore $f(f(x)) = 0$.

Ainsi, $f^2 = 0$. Un polynôme annulateur de f est donc X^2 , de sorte que la seule valeur propre de f est 0.

Et donc la matrice de f dans une base de vecteurs propres est $\text{Diag}(0, 0, \dots, 0) = 0$.

Et donc $f = 0$.

SOLUTION DE L'EXERCICE 22.15

- Soit $x \neq 0$ un vecteur non nul de E .
Alors $\text{Vect}(x)$ est de dimension 1, et par hypothèse, est stable par f .
Puisque $x \in \text{Vect}(x)$, on a alors $f(x) \in \text{Vect}(x)$. C'est-à-dire qu'il existe $\lambda \in \mathbf{R}$ tel que $f(x) = \lambda x$: x est donc un vecteur propre de f .
- Puisque les vecteurs d'une base sont tous non nuls, la question précédente s'applique : \mathcal{B} est une base de E formée de vecteurs propres de f . Et donc la matrice de f dans cette base est diagonale.
- Notons λ_i la valeur propre de f associée au vecteur propre e_i , c'est-à-dire le réel tel que $f(e_i) = \lambda_i e_i$.

$$\text{Alors on a } f(e_1 + \dots + e_n) = \sum_{i=1}^n f(e_i) = \sum_{i=1}^n \lambda_i e_i.$$

D'autre part, $e_1 + \dots + e_n$ est non nul, et donc la question 1 s'applique : il existe un réel $\lambda \in \mathbf{R}$ tel que $f(e_1 + \dots + e_n) = \lambda(e_1 + \dots + e_n) = \lambda e_1 + \dots + \lambda e_n$.

Et donc on a $\lambda_1 e_1 + \dots + \lambda_n e_n = \lambda e_1 + \dots + \lambda e_n$.

Par unicité de la décomposition dans la base \mathcal{B} , on a donc $\lambda_1 = \lambda, \lambda_2 = \lambda, \dots, \lambda_n = \lambda$.

Ainsi, les λ_i sont tous égaux à λ : la matrice de f dans la base \mathcal{B} est donc

$$\text{Diag}(\lambda, \dots, \lambda) = \lambda I_n.$$

Et donc⁸ $f = \lambda \text{id}_E$.

SOLUTION DE L'EXERCICE 22.16

- Soit $\lambda \in \mathbf{R}$. Alors λ est valeur propre de A si et seulement si $\text{rg}(A - \lambda I_3) < 3$. Or

$$A - \lambda I_3 = \begin{pmatrix} 2-\lambda & 1 & 1 \\ 1 & -\lambda & 1 \\ 1 & 3 & -2-\lambda \end{pmatrix} \xleftrightarrow{L_1 \leftrightarrow L_3} \begin{pmatrix} 1 & 3 & -2-\lambda \\ 1 & -\lambda & 1 \\ 2-\lambda & 1 & 1 \end{pmatrix} \xleftrightarrow{L_2 \leftarrow L_2 - L_1, L_3 \leftarrow L_3 - (2-\lambda)L_1} \begin{pmatrix} 1 & 3 & -2-\lambda \\ 0 & -3-\lambda & 3+\lambda \\ 0 & 3\lambda-5 & 5-\lambda^2 \end{pmatrix}$$

Si $3 + \lambda \neq 0 \Leftrightarrow \lambda \neq -3$, alors, en divisant L_2 par $3 + \lambda$, il vient

$$\begin{pmatrix} 1 & 3 & -2-\lambda \\ 0 & -1 & 1 \\ 0 & 3\lambda-5 & 5-\lambda^2 \end{pmatrix} \xleftrightarrow{L_3 \leftarrow L_3 + (3\lambda-5)L_2} \begin{pmatrix} 1 & 3 & -2-\lambda \\ 0 & -1 & 1 \\ 0 & 0 & -\lambda^2 + 3\lambda \end{pmatrix}.$$

Et donc $\lambda \neq -3$ est valeur propre de A si et seulement si $-\lambda^2 + 3\lambda = 0 \Leftrightarrow \lambda(3 - \lambda) = 0$, donc si et seulement si $\lambda = 0$ ou $\lambda = 3$.

D'autre part, on a

$$\text{rg}(A + 3I_3) = \text{rg} \begin{pmatrix} 1 & 3 & 1 \\ 0 & 0 & 0 \\ 0 & 4 & -4 \end{pmatrix} = 2.$$

Et donc -3 est également valeur propre de A .

On a donc $\text{Spec}(A) = \{-3, 0, 3\}$.

Remarque

On peut aussi utiliser le résultat classique suivant : $\text{Im} f \subset \text{Ker} g$ si et seulement si $g \circ f = 0$ (voir l'exercice 2.7).

Remarque

Attention au fait que le λ en question dépend du x que l'on n'a choisi, et n'a aucune raison (pour l'instant) d'être le même pour deux vecteurs x et y distincts.

⁸ λI_n est la matrice de λid_E dans toute base de E .

Astuce

Tous les calculs ayant été faits avant de supposer $\lambda \neq -3$ sont également valables pour $\lambda = -3$.

Et donc nous repartons de la matrice obtenue juste avant de supposer $\lambda \neq -3$, et remplaçons les λ par des -3 .

De plus, des résolutions de systèmes prouvent que

$$E_0(A) = \text{Vect} \begin{pmatrix} 1 \\ -1 \\ -1 \end{pmatrix}, E_3(A) = \text{Vect} \begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix} \text{ et } E_{-3}(A) = \text{Vect} \begin{pmatrix} 1 \\ 2 \\ -7 \end{pmatrix}.$$

Et donc si on pose $P = \begin{pmatrix} 1 & 1 & 2 \\ 2 & -1 & 1 \\ -7 & -1 & 1 \end{pmatrix}$, alors

$$A = P \begin{pmatrix} -3 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 3 \end{pmatrix} P^{-1}.$$

2.a. On a

$$f \circ g = (g^3 + 2g) \circ g = g^4 + 2g^2 \text{ et de même } g \circ f = g \circ (g^3 + 2g) = g^4 + 2g^2.$$

Donc f et g commutent.

Soit donc $\lambda \in \text{Spec}(f)$ et $x \in E_\lambda(f)$.

Alors $f(x) = \lambda x$. En appliquant g aux deux membres de cette égalité, il vient

$$g(f(x)) = g(\lambda x) = \lambda g(x).$$

Or f et g commutent, donc $f(g(x)) = \lambda g(x) \Leftrightarrow g(x) \in E_\lambda(f)$.

Ainsi, tous les sous-espaces propres de f sont stables par g .

2.b. Notons (e_1, e_2, e_3) la base de vecteurs propres de f obtenue à la question 1.

Alors $E_{-3}(f) = \text{Vect}(e_1)$, $E_0(f) = \text{Vect}(e_2)$ et $E_3(f) = \text{Vect}(e_3)$.

Puisque $E_{-3}(f)$ est stable par g , $g(e_1) \in E_{-3}(f) = \text{Vect}(e_1)$.

Et donc il existe un réel λ_1 tel que $g(e_1) = \lambda_1 e_1$.

De même, en utilisant le fait que $E_0(-f)$ et $E_3(f)$ sont de dimension un, on prouve que $g(e_2) = \lambda_2 e_2$ et $g(e_3) = \lambda_3 e_3$.

Ainsi, la matrice de g dans la base (e_1, e_2, e_3) est

$$B = \begin{pmatrix} g(e_1) & g(e_2) & g(e_3) \\ \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix} \begin{matrix} e_1 \\ e_2 \\ e_3 \end{matrix}.$$

Puisque $g^3 + 2g = f$, il vient donc

$$B^3 + 2B = D \Leftrightarrow \begin{pmatrix} \lambda_1^3 + 2\lambda_1 & 0 & 0 \\ 0 & \lambda_2^3 + 2\lambda_2 & 0 \\ 0 & 0 & \lambda_3^3 + 2\lambda_3 \end{pmatrix} = \begin{pmatrix} -3 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 3 \end{pmatrix}.$$

Et donc $\lambda_1^3 + 2\lambda_1 = -3$.

il est facile de constater que $\lambda_1 = -1$ convient. D'autre part, la fonction $x \mapsto x^3 + 2x$ est strictement croissante, et donc injective, de sorte que -1 est la seule solution de $\lambda_1^3 + 2\lambda_1 = -3$.

De même, on prouve que $\lambda_2 = 0$ et $\lambda_3 = 1$, de sorte que

$$B = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

2.c. Nous venons de prouver qu'il n'existe qu'un seul tel g : l'unique endomorphisme de \mathbf{R}^3 dont la matrice dans la base (e_1, e_2, e_3) est B .

Si cette réponse ne vous satisfait pas⁹, notez qu'il est également possible de déterminer la matrice de g dans la base canonique :

$$\text{Mat}_{\mathcal{B}_{can}}(g) = PBP^{-1} = \frac{1}{3} \begin{pmatrix} 2 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 3 & -2 \end{pmatrix}.$$

Remarque

Notons que nous disposons d'un argument théorique plus rapide : f est un polynôme en g , et donc commute avec g .

Méthode

Pour aborder ce type de questions, il suffit d'avoir les idées claires sur les définitions.

- 1) Écrire ce que signifie $x \in E_\lambda(f)$.
- 2) Plus bas, écrire ce que l'on veut montrer : $g(x) \in E_\lambda(f)$, et ce que cela signifie :

$$f(g(x)) = \lambda g(x).$$

- 3) Trouver les étapes qui peuvent mener de l'un à l'autre (ce qui est rarement difficile).

⁹ Personnellement, elle me suffit largement !

SOLUTION DE L'EXERCICE 22.17

- 1.a. Soit $\lambda \in \text{Spec}(f)$, et soit $x \in E_\lambda(f)$. Alors

$$f(g(x)) = (f \circ g)(x) = (g \circ f)(x) = g(f(x)) = g(\lambda x) = \lambda g(x).$$

Ainsi, tous les sous-espaces propres de f sont stables par g .

- 1.b. Notons que par hypothèse, f possède n valeurs propres distinctes, donc ses sous-espaces propres sont tous de dimension 1.

Soit x un vecteur propre de f , et soit λ la valeur propre associée. Puisque x est non nul¹⁰ et que $\dim E_\lambda(f) = 1$, on a $E_\lambda(f) = \text{Vect}(x)$.

De plus, nous avons prouvé à la question précédente que $g(x) \in E_\lambda(f) = \text{Vect}(x)$. Et donc il existe un réel μ tel que $g(x) = \mu x$: x est un vecteur propre de g pour la valeur propre μ . Ainsi, tout vecteur propre de f est vecteur propre de g .

- 1.c. f est diagonalisable car il possède n valeurs propres distinctes.

Si \mathcal{B} est une base de \mathbf{R}^n formée de vecteurs propres de f , alors par la question précédente, c'est aussi une base de vecteurs propres de g . Et donc la matrice de g dans la base \mathcal{B} est diagonale.

Puisque f est diagonalisable, une telle base \mathcal{B} existe, et donc la matrice de g dans cette base est diagonale : g est diagonalisable.

- 2.a. A possède n valeurs propres distinctes, donc elle est diagonalisable : il existe une matrice Q inversible et une matrice D diagonale telles que $A = Q^{-1}DQ$. Et si l'on pose $P = Q^{-1}$, alors $A = PDP^{-1}$ et donc $P^{-1}AP = D$ est diagonale.

- 2.b. Notons $D = \text{Diag}(\lambda_1, \dots, \lambda_n)$, de sorte que les λ_i sont les valeurs propres de A , et donc sont strictement positifs par hypothèse.

Soit $D_1 = \text{Diag}(\sqrt{\lambda_1}, \dots, \sqrt{\lambda_n})$. Alors

$$(PD_1P^{-1})^2 = PD_1^2P^{-1} = PDP^{-1} = A.$$

Ainsi, PD_1P^{-1} est une racine carrée de A .

- 2.c. Si R est une racine carrée de A , alors $A = R^2$, et donc A est un polynôme en R , qui commute donc avec R .

Soit f l'endomorphisme de \mathbf{R}^n dont la matrice dans la base canonique est A , et soit g l'endomorphisme de \mathbf{R}^n dont la matrice dans la base canonique est R .

Alors f possède n valeurs propres distinctes¹¹ et f et g commutent car A et R commutent. Soit \mathcal{B} la base de \mathbf{R}^n telle que P soit la matrice de passage de la base canonique à \mathcal{B} . Alors, par la formule de changement de base,

$$P^{-1}AP = P_{\mathcal{B}, \mathcal{B}_{can}} \text{Mat}_{\mathcal{B}_{can}}(f) P_{\mathcal{B}_{can}, \mathcal{B}} = \text{Mat}_{\mathcal{B}}(f).$$

Et donc la matrice de f dans la base \mathcal{B} est diagonale : \mathcal{B} est une base de \mathbf{R}^n formée de vecteurs propres de f . Et donc par la question 1.c, $\text{Mat}_{\mathcal{B}}(g)$ est diagonale.

Or, $\text{Mat}_{\mathcal{B}}(g) = P^{-1}RP$: $P^{-1}RP$ est une matrice diagonale.

- 2.d. Si R est une racine carrée de A , alors

$$(P^{-1}RP)^2 = P^{-1}R^2P = P^{-1}AP = D.$$

De plus, $P^{-1}RP$ est une matrice diagonale, donc de la forme $\text{Diag}(\mu_1, \dots, \mu_n)$.

On en déduit que $\text{Diag}(\mu_1, \dots, \mu_n)^2 = \text{Diag}(\mu_1^2, \dots, \mu_n^2) = \text{Diag}(\lambda_1, \dots, \lambda_n)$.

Et donc $\mu_i^2 = \lambda_i$, \dots , $\mu_n^2 = \lambda_n$. Ainsi, pour tout $i \in \llbracket 1, n \rrbracket$, $\mu_i = \pm\sqrt{\lambda_i}$.

Inversement, il est facile de vérifier que toute matrice de la forme $P\text{Diag}(a_1, \dots, a_n)P^{-1}$, avec $a_i = \pm\sqrt{\lambda_i}$ est une racine carrée de A .

Et donc A admet exactement 2^n racines carrées.

SOLUTION DE L'EXERCICE 22.18

1. Soit $D = \text{Diag}(\lambda_1, \dots, \lambda_1, \lambda_2, \dots, \lambda_2, \dots, \lambda_p, \dots, \lambda_p)$ une matrice diagonale semblable à A , où les λ_i sont les valeurs propres de A .

Si $Q \in GL_n(\mathbf{R})$ est telle que $A = Q^{-1}DQ$, alors pour tout polynôme P , on a $P(A) = Q^{-1}P(D)Q$, et donc $P(A) = 0$ si et seulement si $P(D) = 0$.

Autrement dit, P est annulateur de A si et seulement si P est annulateur de D . Montrons donc le résultat demandé pour les polynômes annulateurs de D .

Méthode

Montrer que $E_\lambda(f)$ est stable par g , c'est montrer que pour tout $x \in E_\lambda(f)$, $g(x) \in E_\lambda(f)$.

Soit encore $f(g(x)) = \lambda g(x)$. C'est donc cette égalité que l'on doit démontrer, et pas autre chose.

¹⁰ Rappelons qu'un vecteur propre est non nul par définition.

Rappel

Élever une matrice diagonale à la puissance p , c'est élever chacun de ses coefficients diagonaux à la puissance p , donc $D_1^2 = D$.

¹¹ Car A possède n valeurs propres distinctes et que $\text{Spec}(f) = \text{Spec}(A)$.

Matrices diagonales

La matrice d'un endomorphisme f dans une base \mathcal{B} est diagonale si et seulement si \mathcal{B} est formée de vecteurs propres de f .

Dénombrement

Pour chaque coefficient diagonal de $P^{-1}RP$, on a 2 choix possibles ($\sqrt{\lambda_i}$ ou $-\sqrt{\lambda_i}$), donc en tout, on a 2^n choix.

Nous avons prouvé en cours qu'un polynôme P est annulateur de D si et seulement si il possède les λ_i comme valeurs propres. Un tel polynôme doit donc posséder au moins p racines, et donc être de degré supérieur ou égal à p , ou nul.

En particulier, $\prod_{i=1}^p (X - \lambda_i)$ possède les λ_i comme racines et donc est annulateur de $D : D$

(et donc A admet un polynôme annulateur de degré p , et tous les autres sont de degré supérieur ou égal à p).

2. Commençons dans un premier temps par supposer que $k < p$, et montrons que I_n, A, \dots, A^k est une famille libre de $\mathcal{M}_n(\mathbf{R})$.

Soient donc a_0, a_1, \dots, a_k des réels tels que $a_0 I_n + a_1 A + \dots + a_k A^k = 0$.

Alors le polynôme $P = \sum_{i=0}^k a_i X^i$ est annulateur de A , et est de degré inférieur ou égal à k ,

et donc strictement inférieur à p .

D'après ce qui a été dit à la question 1, c'est donc le polynôme nul : $a_0 = a_1 = \dots = a_k = 0$.

Et donc (I_n, A, \dots, A^k) est une famille libre.

Puisqu'elle est évidemment génératrice de $\text{Vect}(I_n, A, \dots, A^k)$, c'en est une base, et donc $\dim \text{Vect}(I_n, A, \dots, A^k) = k + 1$.

Supposons à présent que $k \geq p$. Alors nous venons de prouver que $(I_n, A, A^2, \dots, A^{p-1})$ est libre.

Montrons qu'elle est en génératrice de $\text{Vect}(I_n, A, \dots, A^k)$.

Soit donc $M \in \text{Vect}(I_n, A, \dots, A^k)$. Par définition, il existe des réels a_0, \dots, a_k tels que

$$M = \sum_{i=0}^k a_i A^i.$$

Posons alors $Q(X) = \sum_{i=0}^k a_i X^i$, de sorte que $M = Q(A)$, et procédons à la division euclidienne

de Q par $P : Q = SP + R$, avec $\deg R < p$.

Puisque P est annulateur de A , on a donc $Q(A) = S(A) \underbrace{P(A)}_{=0} + R(A) = R(A)$.

Alors si $R = \sum_{i=0}^{p-1} b_i X^i$, on a donc $M = Q(A) = R(A) = \sum_{i=0}^{p-1} b_i A^i \in \text{Vect}(I_n, A, \dots, A^{p-1})$.

Ainsi, on a $\text{Vect}(I_n, A, \dots, A^k) = \text{Vect}(I_n, A, \dots, A^{p-1})$, et donc il s'agit d'un espace vectoriel de dimension p .

SOLUTION DE L'EXERCICE 22.19

1. Commençons par chercher les valeurs propres de M_x . $\lambda \in \mathbf{R}$ est une valeur propre de M_x si et seulement si $\text{rg}(M_x - \lambda I_3) < 3$. Or,

$$M_x - \lambda I_3 = \begin{pmatrix} 1 - \lambda & x & 0 \\ 0 & -\lambda & 1 \\ 0 & 1 & -\lambda \end{pmatrix} \xleftrightarrow{L_2 \leftrightarrow L_3} \begin{pmatrix} 1 - \lambda & x & 0 \\ 0 & 1 & -\lambda \\ 0 & -\lambda & 1 \end{pmatrix} \xleftrightarrow{L_3 \leftarrow L_3 + \lambda L_2} \begin{pmatrix} 1 - \lambda & x & 0 \\ 0 & 1 & -\lambda \\ 0 & 0 & 1 - \lambda^2 \end{pmatrix}.$$

Cette matrice n'est pas inversible si et seulement si $1 - \lambda = 0$ ou si $1 - \lambda^2 = 0$.

Donc les valeurs propres de M_x sont 1 et -1 .

De plus, on a $\text{rg}(M_x + I_3) = \text{rg} \left(\begin{pmatrix} 2 & x & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{pmatrix} \right) = 2$, donc $\dim E_{-1}(M_x) = 1$.

De même, on a $\text{rg}(M_x - I_3) = \text{rg} \left(\begin{pmatrix} 0 & x & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{pmatrix} \right)$. Ce rang vaut 1 si $x = 0$ et 2 sinon.

Donc $\dim E_1(M_x) = \begin{cases} 2 & \text{si } x = 0 \\ 1 & \text{sinon} \end{cases}$.

On en déduit que $\dim E_1(M_x) + \dim E_{-1}(M_x) = 3$ si et seulement si $x = 0$.

Autrement dit, M_x est diagonalisable si et seulement si $x = 0$.

2. Soit f l'endomorphisme de \mathbf{R}^3 dont la matrice dans la base canonique est M_x . Alors nous cherchons une base (e_1, e_2, e_3) de \mathbf{R}^3 dans laquelle la matrice de f est B .

Il faut donc que e_1 soit dans $E_{-1}(f)$ (et un tel vecteur non existe d'après la première question), et que e_2 soit un vecteur propre de f pour la valeur propre 1. Par exemple, nous pouvons prendre $e_2 = (1, 0, 0)$.

Nous cherchons donc $e_3 = (a, b, c)$ tel que $f(e_3) = e_3 + e_2$. Soit encore

$$M_x \begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} a \\ b \\ c \end{pmatrix} + \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \Leftrightarrow \begin{cases} a + bx = a + 1 \\ b = c \\ c = b \end{cases} \Leftrightarrow b = c = \frac{1}{x}.$$

Il y a donc une infinité de solutions au système : tous les vecteurs de la forme $\left(a, \frac{1}{x}, \frac{1}{x}\right)$.

En particulier, il existe au moins une base dans laquelle la matrice de f est $B : M_x$ et B sont semblables.

SOLUTION DE L'EXERCICE 22.20

1. La fonction $t \mapsto P(t)Q(t)e^{-t}$ est continue sur $[0, +\infty[$.

Au voisinage de $+\infty$, on a $t^2 P(t)Q(t)e^{-t} \xrightarrow[t \rightarrow +\infty]{} 0$ par croissance comparée.

Et donc $P(t)Q(t)e^{-t} = \underset{t \rightarrow +\infty}{o} \left(\frac{1}{t^2}\right)$.

Puisque $\int_1^{+\infty} \frac{dt}{t^2}$ converge, il en est de même de $\int_1^{+\infty} P(t)Q(t)e^{-t} dt$ et donc de $\int_0^{+\infty} P(t)Q(t)e^{-t} dt$.

2. La symétrie et la bilinéarité ne posent pas de difficultés.

Soit $P \in \mathbf{R}_n[X]$. Alors

$$\varphi(P, P) = \int_0^{+\infty} P(t)^2 e^{-t} dt \geq 0$$

par positivité de l'intégrale.

D'autre part, la fonction $t \mapsto P(t)^2 e^{-t}$ étant positive et continue sur \mathbf{R}_+ , si $\varphi(P, P) = 0$, alors pour tout $t \geq 0$, on a $P(t)^2 e^{-t} = 0$.

Soit encore $\forall t \geq 0, P(t) = 0$.

Et alors P possède une infinité de racines : c'est nécessairement le polynôme nul.

Ainsi, φ est bien un produit scalaire sur $\mathbf{R}_n[X]$.

- 3.a. Soit $P \in E$. Alors $\deg P'' \leq n - 2$ et donc $\deg(XP''(X)) \leq n - 1$.

De même, $\deg((X - 1)P') = \deg(X - 1) + \deg P' \leq 1 + n - 1 = n$.

Et donc $f(P) \in E$.

Soient $P, Q \in E$ et $\lambda \in \mathbf{R}$. Alors

$$\begin{aligned} f(\lambda P + Q) &= X(\lambda P + Q)'' - (X - 1)(\lambda P + Q)' \\ &= X(\lambda P'' + Q'') - (X - 1)(\lambda P' + Q') \\ &= \lambda(XP'' - (X - 1)P'(X)) + (XQ'' - (X - 1)Q'(X)) \\ &= \lambda f(P) + f(Q). \end{aligned}$$

Donc f est linéaire : c'est un endomorphisme de E .

- 3.b. Soient $P, Q \in E$.

Afin de prouver la relation indiquée par l'énoncé, procédons à une intégration par parties sur un segment de la forme $[0, A]$, $A > 0$.

Posons donc $u(x) = xP'(x)e^{-x}$ et $v(x) = Q(x)$, qui sont deux fonctions de classe \mathcal{C}^1 , avec $v'(x) = Q'(x)$ et

$$u'(x) = xP''(x)e^{-x} + P'(x)e^{-x} - P'(x)e^{-x} = xP''(x)e^{-x} - (x - 1)P'(x)e^{-x}.$$

Et donc on a

$$\begin{aligned} \int_0^A xP'(x)Q'(x)e^{-x} dx &= [xP'(x)Q(x)e^{-x}]_0^A - \int_0^A (xP''(x) - (x - 1)P'(x))Q(x)e^{-x} dx \\ &= AP'(A)Q'(A)e^{-A} - \int_0^A (xP''(x) - (x - 1)P'(x))Q(x)e^{-x} dx \\ &\xrightarrow{A \rightarrow +\infty} - \int_0^{+\infty} (xP''(x) - (x - 1)P'(x))Q(x)e^{-x} dx \\ &= -\varphi(f(P), Q). \end{aligned}$$

$E_1(f)$

Nul besoin de calculs pour trouver un vecteur propre de f pour la valeur propre 1 : $(1, 0, 0)$ se voit sur la matrice M_x .

Et puisque le sous-espace propre $E_1(M_x)$ est de dimension 1, tout autre vecteur propre sera colinéaire à $(1, 0, 0)$.

Méthode

Comme souvent dans une intégration par parties, le problème est de décider ce que l'on souhaite intégrer et ce que l'on souhaite dériver. Ici, on sait intégrer chacun des quatre termes $x, P'(x), Q'(x)$ et e^{-x} , mais on ne sait intégrer aucun produit de ces termes. Et donc il faut que $v'(x)$ soit l'un de ces quatre termes.

Puisque l'expression finale attendue ($\varphi(f(P), Q)$) fait apparaître Q et non ses dérivées, il semble plutôt logique d'essayer d'intégrer Q (même si avant d'avoir fait le calcul, rien ne nous garantit que ceci aboutira bien).

Remarque

Notons que la convergence de cette dernière intégrale est garantie par le fait que $f(P)$ et Q sont dans E , et donc la première question s'applique.

Et donc on a bien $\varphi(f(P), Q) = - \int_0^{+\infty} xP'(x)Q'(x)e^{-x} dx$.

Et alors, il vient

$$\begin{aligned}\varphi(P, f(Q)) &= \varphi(f(Q), P) \\ &= - \int_0^{+\infty} xQ'(x)P'(x)e^{-x} dx \\ &= - \int_0^{+\infty} xQ'(x)P'(x)e^{-x} dx = \varphi(f(P), Q).\end{aligned}$$

Symétrie de φ .

L'expression obtenue ci-dessus est valable pour tous P, Q .

On en déduit donc que f est un endomorphisme symétrique pour le produit scalaire φ .

SOLUTION DE L'EXERCICE 22.21

Soient $x_1 \in F_1, x_2 \in F_2, \dots, x_n \in F_n$ tels que

$$x_1 + x_2 + \dots + x_n = 0.$$

Alors, pour $i \in \llbracket 1, n \rrbracket$, on a

$$\langle x_i, x_1 + x_2 + \dots + x_n \rangle = \langle x_i, 0 \rangle = 0.$$

Or, par linéarité du produit scalaire, il vient alors

$$\langle x_i, x_1 \rangle + \dots + \langle x_i, x_n \rangle = 0.$$

Mais les F_i étant deux à deux orthogonaux par hypothèse, on a $\langle x_i, x_j \rangle = 0$ si $i \neq j$.

On en déduit que $\langle x_i, x_i \rangle = 0 = \|x_i\|^2$, et donc nécessairement $x_i = 0$.

Ceci étant vrai pour tout i , on a $x_1 = x_2 = \dots = x_n = 0$: les F_i sont en somme directe.

SOLUTION DE L'EXERCICE 22.22

1. Soient $A, B \in \mathcal{M}_2(\mathbf{R})$. Alors

$$\langle A, B \rangle = \text{tr}({}^t AB) = \text{tr}({}^t({}^t AB)) = \text{tr}({}^t BA) = \langle B, A \rangle.$$

Donc $\langle \cdot, \cdot \rangle$ est symétrique.

Soient $A, B, C \in \mathcal{M}_2(\mathbf{R})$ et $\lambda \in \mathbf{R}$. Alors

$$\langle \lambda A + B, C \rangle = \text{tr}({}^t(\lambda A + B)C) = \text{tr}(\lambda {}^t AC + {}^t BC) = \lambda \text{tr}({}^t AC) + \text{tr}({}^t BC) = \lambda \langle A, C \rangle + \langle B, C \rangle.$$

Donc $\langle \cdot, \cdot \rangle$ est linéaire à gauche, donc bilinéaire symétrique.

Soit $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathcal{M}_2(\mathbf{R})$. Alors

$$\langle A, A \rangle = \text{tr} \left(\begin{pmatrix} a & c \\ b & d \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix} \right) = \text{tr} \left(\begin{pmatrix} a^2 + c^2 & ab + cd \\ ab + cd & d^2 + b^2 \end{pmatrix} \right) = a^2 + b^2 + c^2 + d^2.$$

Donc $\langle A, A \rangle \geq 0$ et $\langle A, A \rangle = 0$ si et seulement si $a = b = c = d = 0 \Leftrightarrow A = 0$.

Ainsi, $\langle \cdot, \cdot \rangle$ est un produit scalaire sur $\mathcal{M}_2(\mathbf{R})$.

2. Il est clair que $F \subset \mathcal{M}_2(\mathbf{R})$ et que la matrice nulle est un élément de F .

Soient donc $A = \begin{pmatrix} a & b \\ b & -a \end{pmatrix}$ et $B = \begin{pmatrix} a' & b' \\ b' & -a' \end{pmatrix}$ deux éléments de F et $\lambda \in \mathbf{R}$. Alors

$$\lambda A + B = \lambda \begin{pmatrix} a & b \\ b & -a \end{pmatrix} + \begin{pmatrix} a' & b' \\ b' & -a' \end{pmatrix} = \begin{pmatrix} \lambda a + a' & \lambda b + b' \\ \lambda a + a' & -(\lambda b + b') \end{pmatrix} \in F.$$

Ainsi, F est bien un sous-espace vectoriel de $\mathcal{M}_2(\mathbf{R})$.

De plus, si $\begin{pmatrix} a & b \\ b & -a \end{pmatrix} \in F$, alors

$$\begin{pmatrix} a & b \\ b & -a \end{pmatrix} = a \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} + b \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \in \text{Vect} \left(\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \right).$$

Donc $\left(\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \right)$ est une famille génératrice de F . Puisque ces deux matrices ne sont pas colinéaires, c'est une famille libre, et donc une base de F , qui est alors de dimension 2.

Rappel

Pour une somme directe de $n \geq 2$ sous-espaces vectoriels, il n'y a pas d'analogie à $F \cap G = \{0\}$ permettant de prouver aisément que la somme est directe, il faut donc revenir à la définition d'une somme directe.

Remarque

On profite ici du fait qu'il s'agit de matrice 2×2 , mais on pourrait également utiliser la formule du produit matriciel pour montrer le même résultat sur $M_n(\mathbf{R})$ (où il n'est alors pas question de nommer tous les coefficients).

3. Soit $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathcal{M}_2(\mathbf{R})$. Alors

$$A \in F^\perp \Leftrightarrow \begin{cases} \left\langle A, \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \right\rangle = 0 \\ \left\langle A, \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \right\rangle = 0 \end{cases} \Leftrightarrow \begin{cases} a - d = 0 \\ b + c = 0 \end{cases} \Leftrightarrow \begin{cases} a = d \\ c = -b \end{cases}$$

Soit encore $A \in F^\perp \Leftrightarrow A = a \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + b \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \in \text{Vect} \left(\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \right)$.

Donc $F^\perp = \text{Vect} \left(\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \right)$.

4. Notons $p_F(A)$ le projeté orthogonal de A sur F . Alors $p_F(A) \in F$, donc il existe deux réels λ et μ tels que $p_F(A) = \begin{pmatrix} \lambda & \mu \\ \mu & -\lambda \end{pmatrix}$.

D'autre part, nous savons que $A - p_F(A) = \begin{pmatrix} 1 - \lambda & -\mu \\ 1 - \mu & \lambda \end{pmatrix} \in F^\perp$, et donc

$$\begin{cases} \left\langle \begin{pmatrix} 1 - \lambda & -\mu \\ 1 - \mu & \lambda \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \right\rangle = 0 \\ \left\langle \begin{pmatrix} 1 - \lambda & -\mu \\ 1 - \mu & \lambda \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \right\rangle = 0 \end{cases} \Leftrightarrow \begin{cases} 1 - 2\lambda = 0 \\ 1 - 2\mu = 0 \end{cases} \Leftrightarrow \lambda = \mu = \frac{1}{2}.$$

Donc $p_F(A) = \frac{1}{2} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$.

Et alors $\min_{B \in F} \|A - B\| = \|A - p_F(A)\| = \left\| \begin{pmatrix} 1/2 & -1/2 \\ 1/2 & -1/2 \end{pmatrix} \right\| = 1$.

SOLUTION DE L'EXERCICE 22.23

Posons $u = e_1 + \dots + e_n$. Puisque (e_1, \dots, e_n) est orthonormée, on a $\|e_n\| = \sqrt{1^2 + 1^2 + \dots + 1^2} = \sqrt{n}$.

Et donc une base orthonormée de F est $\frac{u}{\sqrt{n}}$.

Le projeté orthogonal de e_1 sur F est alors

$$p_F(e_1) = \left\langle \frac{u}{\sqrt{n}}, e_1 \right\rangle \frac{u}{\sqrt{n}} = \frac{u}{n}.$$

Et donc la distance de e_1 à F est

$$\|e_1 - p_F(e_1)\| = \frac{1}{n} \|(n-1)e_1 - e_2 - \dots - e_n\| = \frac{1}{n} \sqrt{(n-1)^2 + n - 1} = \sqrt{\frac{n-1}{n}} = \sqrt{1 - \frac{1}{n}}.$$

SOLUTION DE L'EXERCICE 22.24

Puisque f est symétrique, il existe une base orthonormée de E formée de vecteurs propres de f .

Notons (e_1, \dots, e_n) une telle base, et notons λ_i la valeur propre de f associée à e_i .

Supposons que $i)$ est vérifié, c'est-à-dire que toute valeur propre de f est dans $[-1, 1]$.

Soit $x \in E$. Alors il existe des réels a_1, \dots, a_n tels que $x = \sum_{i=1}^n a_i e_i$.

Et donc $f(x) = \sum_{i=1}^n a_i f(e_i) = \sum_{i=1}^n a_i \lambda_i e_i$.

Et alors

$$\|f(x)\|^2 = \sum_{i=1}^n \underbrace{\lambda_i^2}_{\leq 1} a_i^2 \leq \sum_{i=1}^n a_i^2 = \|x\|^2.$$

Puisqu'une norme est toujours positive, il vient donc $\|f(x)\| \leq \|x\|$.

Rappel

A est dans F^\perp si et seulement si A est orthogonal à tous les éléments d'une base de F .

Rappel

Si (e_1, \dots, e_p) est une base orthonormée de F , alors pour tout $x \in E$, le projeté orthogonal de x sur F est donné par

$$p_F(x) = \sum_{i=1}^p \langle x, e_i \rangle e_i.$$

Rappels

Dans une base orthonormée, on dispose d'une formule permettant de calculer aisément la norme : si

$x = \sum_{i=1}^n x_i e_i$, alors

$$\|x\|^2 = \sum_{i=1}^n x_i^2.$$

Inversement, supposons que pour tout $x \in E$, $\|f(x)\| \leq \|x\|$, et soit alors λ une valeur propre de f et x un vecteur propre associé.

On a alors $f(x) = \lambda x$ et donc $\|f(x)\| = \|\lambda x\| = |\lambda| \cdot \|x\|$.

Puisque par hypothèse, $\|f(x)\| \leq \|x\|$, on a donc $|\lambda| \cdot \|x\| \leq \|x\|$.

Mais x est un vecteur propre, donc est non nul, de sorte que $\|x\| > 0$.

Et donc en divisant l'inégalité précédente par $\|x\|$, il vient $|\lambda| \leq 1$.

SOLUTION DE L'EXERCICE 22.25

1. Soient $x, y \in E$ et $\lambda \in \mathbf{R}$. Alors on a

$$\begin{aligned} f(\lambda x + y) &= 2 \frac{\langle \lambda x + y, u \rangle}{\|u\|^2} u - (\lambda x + y) \\ &= 2 \frac{\lambda \langle x, u \rangle + \langle y, u \rangle}{\|u\|^2} u - \lambda x - y \\ &= 2\lambda \frac{\langle x, u \rangle}{\|u\|^2} u - \lambda x + 2 \frac{\langle y, u \rangle}{\|u\|^2} u - y \\ &= \lambda f(x) + f(y). \end{aligned}$$

Linéarité à gauche du produit scalaire.

Donc f est linéaire, et donc est un endomorphisme de E .

De plus, pour $x \in E$, on a

$$\begin{aligned} f(f(x)) &= f\left(2 \frac{\langle x, u \rangle}{\|u\|^2} u - x\right) \\ &= 2 \frac{\langle x, u \rangle}{\|u\|^2} f(u) - f(x) \\ &= 2 \frac{\langle x, u \rangle}{\|u\|^2} u - \left(2 \frac{\langle x, u \rangle}{\|u\|^2} u - x\right) \\ &= x. \end{aligned}$$

$f(u) = u$.

Ceci étant vrai pour tout $x \in E$, on a donc $f \circ f = \text{id}_E$.

2. Nous venons de prouver que $X^2 - 1 = (X - 1)(X + 1)$ est un polynôme annulateur de f . Donc les valeurs propres possibles de f sont -1 et 1 .

De plus, on a

$$f(x) = x \Leftrightarrow 2 \frac{\langle x, u \rangle}{\|u\|^2} u - x = x \Leftrightarrow \frac{\langle x, u \rangle}{\|u\|^2} u = x.$$

En particulier, un tel x doit être colinéaire à u , et donc $E_1(f) \subset \text{Vect}(u)$.

D'autre part, on a

$$f(u) = 2 \frac{\langle u, u \rangle}{\|u\|^2} u - u = 2u - u = u.$$

Ceci prouve que 1 est bien valeur propre de f et que $E_1(f) = \text{Vect}(u)$.

De même, on a

$$f(x) = -x \Leftrightarrow 2 \frac{\langle x, u \rangle}{\|u\|^2} u - x = -x \Leftrightarrow \langle x, u \rangle = 0 \Leftrightarrow x \in \text{Vect}(u)^\perp.$$

Et donc -1 est valeur propre¹² de f , et $E_{-1}(f) = F^\perp$.

Puisque $E = F \oplus F^\perp = E_1(f) \oplus E_{-1}(f)$, on en déduit que f est diagonalisable.

Alternative : vous avez peut-être reconnu que $\frac{\langle x, u \rangle}{\|u\|^2} u$ est le projeté orthogonal de x sur $F = \text{Vect}(u)$.

Dans ce cas, on a $f = 2p_F - \text{id}_E$, et donc $f^2 = (2p_F - \text{id}_E)^2 = 4p_F^2 - 4p_F + \text{id}_E = 4p_F - 4p_F + \text{id}_E = \text{id}_E$.

De plus, dans une «bonne» base, la matrice p_F est $D = \text{Diag}(1, 0, \dots, 0)$, et donc celle de f est $2D - I_n = \text{Diag}(1, -1, \dots, -1)$. Ceci prouve donc directement que f est diagonalisable, que ses valeurs propres sont -1 et 1 , et que $\dim E_1(f) = 1$, $\dim E_{-1}(f) = n - 1$.

¹² Au moins si $\dim E \geq 2$, car dans le cas contraire, $F^\perp = \{0\}$.

SOLUTION DE L'EXERCICE 22.26

1. On a ${}^tS = {}^t({}^tAA) = {}^tA({}^tA) = {}^tAA = S$. Donc S est symétrique.

Soit $X \in \mathcal{M}_{n,1}(\mathbf{R})$. Alors

$${}^tXSX = {}^tX{}^tAAX = {}^t(AX)(AX) = \|AX\|^2 \geq 0$$

où $\|\cdot\|$ désigne la norme pour le produit scalaire canonique de $\mathcal{M}_{n,1}(\mathbf{R})$.

En particulier, si X est un vecteur propre¹³ de S associé à une valeur propre λ , alors

$${}^tXSX = {}^tX(SX) = {}^tX(\lambda X) = \lambda{}^tXX = \lambda\|X\|^2.$$

Donc $\lambda\|X\|^2 \geq 0$, et puisque $\|X\|^2 > 0$, alors $\lambda \geq 0$.

Ainsi, toutes les valeurs propres de S sont positives ou nulles.

2. Si S est symétrique, elle est diagonalisable en base orthonormée : il existe une matrice P orthogonale telle que $S = {}^tP\text{Diag}(\lambda_1, \dots, \lambda_n)P$, où les λ_i sont les valeurs propres de S , et donc sont positifs.

Soit alors $D_1 = \text{Diag}(\sqrt{\lambda_1}, \dots, \sqrt{\lambda_n})$. Alors $D_1^2 = \text{Diag}(\lambda_1, \dots, \lambda_n)$ et ${}^tD_1 = D_1$, puisqu'une matrice diagonale est toujours symétrique.

Posons donc $A = D_1P$. On a alors

$${}^tAA = {}^tP{}^tD_1D_1P = {}^tPD_1^2P = S.$$

3. Les valeurs propres de S sont les racines de $X^2 - 4X + 3 = (X - 3)(X - 1)$. Ainsi, S est bien une matrice symétrique dont toutes les valeurs propres sont positives ou nulles.

Une résolution de système prouve que les sous-espaces propres de S sont $E_1(S) = \text{Vect}\left(\begin{pmatrix} 1 \\ -1 \end{pmatrix}\right)$

et $E_3(S) = \text{Vect}\left(\begin{pmatrix} 1 \\ 1 \end{pmatrix}\right)$.

Ainsi, une base orthonormée de $\mathcal{M}_{2,1}(\mathbf{R})$ formée de vecteurs propres de S est donc

$$\frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix}, \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

On a donc, en posant $P = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}$, $S = {}^tP \begin{pmatrix} 1 & 0 \\ 0 & 3 \end{pmatrix} P$.

D'après la question précédente, en posant $A = \begin{pmatrix} 1 & 0 \\ 0 & \sqrt{3} \end{pmatrix} P$, on a donc $S = {}^tAA$.

SOLUTION DE L'EXERCICE 22.27

- 1.a. Il est facile de voir que u_a est linéaire et donc est un endomorphisme de E .

Si $a = 0$, alors u_a est l'endomorphisme nul, qui est bien évidemment dans $T(E)$.

Supposons donc que $a \neq 0$. Si $x \in \text{Vect}(a)^\perp$, alors $u_a(x) = (x|a)a = 0$. Par conséquent, dans une base orthonormée de E obtenue par concaténation de $\frac{a}{\|a\|}$ et d'une base de $\text{Vect}(a)^\perp$,

la matrice de u_a est

$$\begin{pmatrix} \|a\|^2 & 0 & \dots & 0 \\ 0 & \vdots & & \vdots \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & 0 \end{pmatrix}.$$

Cette matrice est de rang 1 et symétrique, donc u_a est de rang 1 et est un endomorphisme symétrique¹⁴ Enfin, si $x \in E$, on a

$$(x|u_a(x)) = (x|(x|a)a) = (x|a)(x|a) = (x|a)^2 \geq 0.$$

Et donc u_a est bien un élément de $T(E)$.

- 1.b. Nous avons déjà prouvé que $\text{Vect}(a)^\perp \subset \text{Ker}(u_a)$.

Et même mieux : $x \in \text{Ker}(u_a) \Leftrightarrow (x|a) = 0 \Leftrightarrow x \in (\text{Vect}(a))^\perp$. Et donc 0 est valeur propre de u_a , avec $\dim E_0(u_a) = n - 1$.

Puisque u_a est symétrique, il est diagonalisable, et possède donc une unique autre valeur propre, avec un sous-espace propre de dimension 1.

Or $u_a(a) = \|a\|^2 a$, et donc $\|a\|^2$ est valeur propre de u_a , et $E_{\|a\|^2}(u_a) = \text{Vect}(a)$.

Méthode

Prouver que toutes les valeurs propres de S sont positives revient à prouver que la forme quadratique associée à S ne prend que des valeurs positives. Il s'agit donc de faire apparaître cette forme quadratique (sans le dire puisqu'en théorie vous ne savez rien sur les formes quadratiques) en revenant à la définition d'une forme quadratique :

$$q_S(X) = {}^tXSX.$$

¹³ Et donc non nul !

Rappel

Pour obtenir une base orthonormée de vecteurs propres, il suffit d'appliquer Gram-Schmidt à chacune des bases des sous-espaces propres. En l'occurrence ici, il suffit de diviser chaque vecteur par sa norme, il n'est pas nécessaire de se préoccuper de l'orthogonalité des deux vecteurs, qui est automatique (les sous-espaces propres d'une matrice symétrique sont deux à deux orthogonaux).

Détails

a est une base de $\text{Vect}(a)$ et donc $\frac{a}{\|a\|}$ est une base orthonormée de $\text{Vect}(a)$.

¹⁴ Car nous nous sommes placés dans une base **orthonormée**.

2.a. Si $b \in \text{Im } u$, alors il existe $x \in E$ tel que $b = u(x)$.
 Puisque u est non nul, $\text{rg}(u) \geq 1$, et donc $\text{rg}(u) = 1$.
 Ainsi, $\text{Im}(u)$ est un sous-espace vectoriel de E de dimension 1, et b en est un vecteur non nul, donc est une base de $\text{Im}(u)$: $\text{Im}(u) = \text{Vect}(b)$.
 D'autre part, on a $u(b) \in \text{Im}(u)$, et donc il existe un réel μ tel que $u(b) = \mu b$: b est un vecteur propre de u .

Et alors $(x|u(x)) = (x|\mu x) = \mu \|x\|^2$.
 Puisque $(x|u(x)) \geq 0$ par hypothèse et que $\|x\|^2 > 0$, nécessairement $\mu \geq 0$.

2.b. Notons que 0 est valeur propre de u , et $\dim E_0(u) = \dim E - \text{rg}(u) = \dim E - 1$.
 Si on avait $\mu = 0$, alors $\text{Im } u \subset \text{Ker } u$, de sorte que $u^2 = 0$, et la seule valeur propre de u serait 0.
 Mais u étant diagonalisable, cela signifierait que la matrice de u dans une base de vecteurs propres serait nulle, et donc $u = 0$.
 Ainsi, on a $\mu > 0$.
 Et donc $E = E_0(u) \oplus E_\mu(u) = E_0(u) \oplus \text{Vect}(b)$.

Soit donc $\left(\frac{b}{\|b\|}, e_2, \dots, e_n\right)$ une base orthonormée de E obtenue par concaténation d'une base orthonormée de $\text{Vect}(u)$ et d'une base orthonormée de $E_0(u) = \text{Ker}(u)$.

Si $x \in E$, nous savons que les coordonnées de x dans cette base orthonormée sont obtenues via

$$x = \left(x \left| \frac{b}{\|b\|} \right.\right) \frac{b}{\|b\|} + \sum_{i=2}^n (x|e_i) e_i.$$

Et donc

$$u(x) = \frac{(x|b)}{\|b\|^2} u(b) + \sum_{i=2}^n (x|e_i) \underbrace{u(e_i)}_{=0} = \frac{(x|b)}{\|b\|^2} \mu b.$$

2.c. Nous venons de prouver que pour tout $x \in E$,

$$u(x) = \frac{\mu}{\|b\|^2} (x|b) b = \left(x \left| \frac{\sqrt{\mu}}{\|b\|} b \right.\right) \frac{\sqrt{\mu}}{\|b\|} b.$$

Et donc, en posant $a = \frac{\sqrt{\mu}}{\|b\|} b$, qui est non nul car $b \neq 0$ et $\mu > 0$, on a $u(x) = (x|a) a = u_a(x)$.

Et donc $u = u_a$.

2.d. Nous venons de prouver que φ est surjective, puisque tout $u \in T(E)$ est dans l'image de φ .
 En revanche, φ n'est pas injective, car si $a \neq 0$, alors $\varphi(a) = \varphi(-a)$ puisque pour tout $x \in E$,

$$u_a(x) = (x|a) a = (x|(-a))(-a) = u_{-a}(x).$$

3.a. Par linéarité de f , on a

$$f \circ u_a(x) = f((x|a)a) = (x|a)f(a).$$

3.b. Notons $F = \text{Vect}(a)$, de sorte que $\frac{a}{\|a\|}$ forme une base orthonormée de F .

Posons alors $e_1 = \frac{a}{\|a\|}$, et soit (e_2, \dots, e_n) est une base orthonormée de F^\perp . Alors $\mathcal{B} = (e_1, e_2, \dots, e_n)$ est une base orthonormée de E .

Or, pour $x \in F^\perp$, on a $u_a(x) = 0$ et donc $f \circ u_a(x) = 0$.

D'autre part, $f \circ u_a \left(\frac{a}{\|a\|}\right) = f(a)$.

Et donc la matrice de $f \circ u_a$ dans la base orthonormée $\left(\frac{a}{\|a\|}, e_1, \dots, e_n\right)$ est de la forme

$$\text{Mat}_{\mathcal{B}}(f \circ u_a) = \begin{pmatrix} \star & 0 & \dots & 0 \\ \star & \vdots & & \vdots \\ \vdots & \vdots & & \vdots \\ \star & 0 & \dots & 0 \end{pmatrix} \begin{matrix} e_1 \\ e_2 \\ \vdots \\ e_n \end{matrix}.$$

Détails

u étant non nul, son rang est supérieur ou égal à 1. Et donc si $u \in T(E)$, nécessairement $\text{rg}(u) = 1$.

Détail

Nous venons de le prouver pour les $u \neq 0$, mais il est évident que $0 = \varphi(0)$.

Danger !

Une erreur grossière aurait été de dire « φ est surjective, donc injective.»
 En effet, rien n'indique que φ soit linéaire, ni même que $T(E)$ possède même dimension que E . D'ailleurs $T(E)$ n'est même pas un espace vectoriel !

Mais la base \mathcal{B} étant symétrique, $f \circ u_a$ est symétrique si et seulement si $\text{Mat}_{\mathcal{B}}(f \circ u_a)$ est une matrice symétrique. Elle est alors symétrique si et seulement si elle est diagonale, c'est-à-dire si et seulement si e_1 est un vecteur propre de $f \circ u_a$.

Et puisque a est colinéaire à e_1 , c'est le cas si et seulement si a est un vecteur propre de $f \circ u_a$.

Enfin, puisque $f \circ u_a(a) = f(a)$, c'est le cas si et seulement si a est un vecteur propre de f .

- 3.c. La matrice de $f \circ u_a$ exhibée à la question 3.b prouve que $f \circ u_a$ est de rang au plus 1. D'autre part, nous savons déjà qu'il est nécessaire que a soit un vecteur propre de f , ce que nous supposons dans la suite. Notons donc λ la valeur propre de f à laquelle a est associé. Alors tout vecteur x de E s'écrit de manière unique $x = \mu a + y$, où $a \in \mathbf{R}$ et $y \in F^\perp$.

Et alors

$$(f \circ u_a(x)|x) = (\mu\lambda a|\mu a + y) = \lambda\mu^2\|a\|^2.$$

Pour que ceci soit positif pour tout x , il faut et il suffit que $\lambda \geq 0$.

Ainsi, $f \circ u_a \in T(E)$ si et seulement si a est un vecteur propre de f , associé à une valeur propre positive ou nulle.

4. Pour $x \in E$, on a

$$u_a \circ u_b(x) = (u_b(x)|a)a = (x|b)(a|b)a \text{ et } u_b \circ u_a(x) = (x|a)(a|b)b.$$

Ainsi, si u_a et u_b commutent, alors soit $(a|b) = 0$, soit pour tout $x \in E$, $(x|b)a = (x|a)b$.

En particulier, pour $x = a$, on doit alors avoir $\underbrace{\|a\|^2}_{>0} a(a|b)b$.

Et donc a et b sont colinéaires.

Inversement, si $(a|b) = 0$, alors on a évidemment $u_a \circ u_b = u_b \circ u_a$, et si $a = \lambda b$, alors pour tout $x \in E$, $(x|b)a = (x|b)\lambda b = (x|\lambda b)b = (x|a)b$, et donc u_a et u_b commutent.

Ainsi, u_a et u_b commutent si et seulement si a et b sont orthogonaux ou colinéaires.

EDHEC 2008

Dans cet exercice, f désigne un endomorphisme d'un \mathbf{R} -espace vectoriel E de dimension finie. On se propose d'étudier quelques situations dans lesquelles on peut établir que $E = \text{Ker } f \oplus \text{Im } f$.

1. **a.** Montrer que si f est un automorphisme de E , alors on a bien $E = \text{Ker } f \oplus \text{Im } f$.
- b.** Étude d'un exemple : on considère deux sous-espaces vectoriels supplémentaires F et G , de E . Tout élément de E s'écrit donc de manière unique $x = x_F + x_G$, avec $x_F \in F$ et $x_G \in G$. On appelle alors symétrie par rapport à F parallèlement à G l'endomorphisme s de E défini par

$$s(x) = x_F - x_G.$$

Déterminer s^2 et en déduire que $E = \text{Ker } s \oplus \text{Im } s$.

2. Dans cette question, on suppose f diagonalisable et f non bijectif (le cas où f est bijectif ayant été traité dans la première question).
 - a.** Traiter le cas où f est l'endomorphisme nul.
 - b.** Dans cette question, on suppose que f n'est pas l'endomorphisme nul.
 - i.** Montrer que f a d'autres valeurs propres que la valeur propre 0.
 - ii.** Montrer que tout sous-espace propre de f associé à une valeur propre non nulle est inclus dans $\text{Im } f$.
 - iii.** En déduire que $E = \text{Ker } f \oplus \text{Im } f$.

3. Dans cette question, on considère un endomorphisme f de E dont un polynôme annulateur est de la forme

$$P = \sum_{k=1}^p a_k X^k \text{ ou encore } P = a_1 X + a_2 X^2 + \dots + a_p X^p \text{ avec } a_1 \neq 0 \text{ et } p \geq 1.$$

- a.** Soit y un élément de $\text{Im } f \cap \text{Ker } f$.
 - i.** Montrer qu'il existe un vecteur x de E tel que $y = f(x)$ et $f^2(x) = 0$.
 - ii.** En déduire que, pour tout entier k supérieur ou égal à 2, on a $f^k(x) = 0$.
- b.** Établir que $E = \text{Ker } f \oplus \text{Im } f$.

EDHEC 2008 : CORRECTION

1.a. Si f est un automorphisme de E , alors en particulier f est injectif de sorte que $\text{Ker } f = \{0\}$.
Et f est surjectif, de sorte que $\text{Im } f = E$. On a donc bien $E = \text{Ker } f \oplus \text{Im } f$.

1.b. Notons qu'avec les notations de l'énoncé, si $x = x_F + x_G$, alors $s(x) = x_F - x_G = x_F + (-x_G)$, avec $x_F \in F$ et $-x_G \in G$. Donc

$$s^2(x) = x_F - (-x_G) = x_F + x_G = x.$$

Ainsi, $s^2 = \text{id}$. Par conséquent, s est un automorphisme de E , et on a $s^{-1} = s$.

Et donc, par la question précédente, $E = \text{Im } s \oplus \text{Ker } s$.

2.a. Si f est l'endomorphisme nul, alors $\text{Ker } f = E$ et $\text{Im } f = \{0\}$, donc comme précédemment, $E = \text{Ker } f \oplus \text{Im } f$.

2.b.i. f est diagonalisable, de sorte que $\sum_{\lambda \in \text{Spec}(f)} \dim E_\lambda(f) = n$.

Si f ne possédait que 0 comme valeur propre, alors on aurait nécessairement $\dim E_0(f) = n$, c'est-à-dire $E_0(f) = E$. Or $E_0(f) = \text{Ker } f$, et donc f serait l'endomorphisme nul. Ainsi,

f possède une valeur propre non nulle.

2.b.ii. Soit λ une valeur propre non nulle de f et soit $x \in E_\lambda(f)$.

Alors $f(x) = \lambda x \Leftrightarrow x = \frac{1}{\lambda} f(x) = f\left(\frac{1}{\lambda} x\right)$. Et ainsi $x \in \text{Im } f$. On en déduit que

$$E_\lambda(f) \subset \text{Im } f.$$

2.b.iii. De la question précédente, on déduit que

$$\bigoplus_{\substack{\lambda \in \text{Spec}(f) \\ \lambda \neq 0}} E_\lambda(f) \subset \text{Im } f$$

et donc $\dim \text{Im } f \geq \sum_{\substack{\lambda \in \text{Spec}(f) \\ \lambda \neq 0}} \dim E_\lambda(f)$.

D'autre part, on a

$$\dim \text{Im } f = n - \dim \text{Ker } f = n - \dim E_0(f).$$

Enfin, f est diagonalisable, de sorte que $\sum_{\lambda \in \text{Spec}(f)} \dim E_\lambda(f) = n$.

Et donc $\dim E_0(f) = n - \sum_{\substack{\lambda \in \text{Spec}(f) \\ \lambda \neq 0}} \dim E_\lambda(f)$.

On en déduit que

$$\dim \text{Im } f \leq \sum_{\substack{\lambda \in \text{Spec}(f) \\ \lambda \neq 0}} \dim E_\lambda(f) = \dim \left(\bigoplus_{\substack{\lambda \in \text{Spec}(f) \\ \lambda \neq 0}} E_\lambda(f) \right).$$

Par double inégalité, on a alors $\dim \text{Im } f = \sum_{\substack{\lambda \in \text{Spec}(f) \\ \lambda \neq 0}} \dim E_\lambda(f)$ et donc¹ $\text{Im } f = \bigoplus_{\substack{\lambda \in \text{Spec}(f) \\ \lambda \neq 0}} E_\lambda(f)$.

Enfin, on sait que $E = \bigoplus_{\lambda \in \text{Spec}(f)} E_\lambda(f)$ et donc

$$E = E_0(f) \oplus \left(\bigoplus_{\substack{\lambda \in \text{Spec}(f) \\ \lambda \neq 0}} E_\lambda(f) \right) = \text{Ker } f \oplus \text{Im } f.$$

3.a.i. Si $y \in \text{Im } f$, alors par définition de $\text{Im } f$, il existe $x \in E$ tel que $y = f(x)$.

Mais $y \in \text{Ker } f$, de sorte que $f(y) = 0$ et donc $f(f(x)) = 0$.

Somme directe

L'unique écriture de $s(x)$ comme un élément de F et un élément de G est

$$s(x) = x_F - x_G.$$

⚠ Attention !

Nous ne venons pas de dire qu'un endomorphisme ne possédant que 0 comme valeur propre est nul, mais qu'un endomorphisme **diagonalisable** ne possédant que 0 comme valeur propre est nul.

Rappel

La dimension d'une somme directe de sous-espaces vectoriels est la somme des dimensions. Ceci est faux si la somme n'est pas directe.

¹ On a déjà une inclusion, et nous venons de prouver l'égalité des dimensions.

3.a.ii. Pour $k \geq 2$, on a

$$f^k(x) = f^{k-2}(f^2(x)) = f^{k-2}(0) = \boxed{0.}$$

3.b. Par le théorème du rang, nous savons déjà que $\dim \operatorname{Im} f + \dim \operatorname{Ker} f = n$.

De plus, si $y \in \operatorname{Ker} f \cap \operatorname{Im} f$, alors en reprenant les notations de la question précédente, on a

$$P(f)(x) = 0 \Leftrightarrow a_1 f(x) + a_2 f^2(x) + \cdots + a_p f^p(x) = 0.$$

Mais pour $k \geq 2$, $f^k(x) = 0$ et donc

$$a_1 f(x) = 0 \Leftrightarrow a_1 y = 0.$$

Puisque par hypothèse $a_1 \neq 0$, alors $y = 0$.

Nous avons donc prouvé que $\operatorname{Im} f \cap \operatorname{Ker} f = \{0\}$ et donc $\boxed{E = \operatorname{Im} f \oplus \operatorname{Ker} f.}$

EDHEC 2012

On note E l'espace vectoriel des polynômes à coefficients réels, de degré inférieur ou égal à 2. On note e_0, e_1 et e_2 les polynômes de E définis par : $e_0 = 1, e_1 = X$ et $e_2 = X^2$.

On rappelle que $\mathcal{B} = (e_0, e_1, e_2)$ est une base de E .

On considère l'application f qui, à tout polynôme P de E , associe le reste dans la division par $1 + X^3$ du polynôme $(1 - X + X^2)P$.

Ainsi, il existe un unique polynôme Q tel que

$$(1 - X + X^2)P = (1 + X^3)Q + f(P), \text{ avec } \deg(f(P)) \leq 2.$$

1. Montrer que f est un endomorphisme de E .
2.
 - a. Déterminer $f(e_0), f(e_1)$ et $f(e_2)$ puis vérifier que $f(e_0) = -f(e_1) = f(e_2)$.
 - b. En déduire une base de $\text{Im } f$.
 - c. Donner la dimension de $\text{Ker } f$ ainsi qu'une base de $\text{Ker } f$.
3.
 - a. Calculer $f(P)$ pour tout polynôme P de $\text{Im } f$, puis établir que 3 est valeur propre de f et que

$$\text{Im } f = \text{Ker}(f - 3\text{Id}).$$

- b. Montrer que f est diagonalisable.
4. On considère l'application φ de $E \times E$ dans \mathbf{R} qui, à tout couple (P, Q) de polynômes de E associe le réel $\varphi(P, Q) = \sum_{k=0}^2 a_k b_k$, où l'on a noté $P = a_0 + a_1X + a_2X^2$ et $Q = b_0 + b_1X + b_2X^2$.
 - a. Montrer que φ est un produit scalaire défini sur E .
 - b. Vérifier que $\text{Ker } f$ est le supplémentaire orthogonal de $\text{Im } f$ dans E pour ce produit scalaire.
 5.
 - a. Vérifier que \mathcal{B} est une base orthonormale pour le produit scalaire φ .
 - b. Écrire la matrice de f dans la base \mathcal{B} puis retrouver le résultat de la question 4.b.

EDHEC 2012 : CORRECTION

1. Il est clair que f est à valeurs dans $\mathbf{R}_2[X] = E$.
Soient $P_1, P_2 \in E$ et $\lambda \in \mathbf{R}$. Alors il existe deux polynômes Q_1 et Q_2 tels que :

$$(1 - X + X^2)P_1 = (1 + X^3)Q_1 + f(P_1) \text{ et } (1 - X + X^2)P_2 = (1 + X^3)Q_2 + f(P_2).$$

Et donc $(1 - X + X^2)(\lambda P_1 + P_2) = (1 + X^3)(\lambda Q_1 + Q_2) + (\lambda f(P_1) + f(P_2))$,
avec $\deg(\lambda f(P_1) + f(P_2)) \leq 2$.

Par unicité de la division euclidienne, $\lambda f(P_1) + f(P_2)$ est alors le reste de la division euclidienne de $(1 - X + X^2)(\lambda P_1 + P_2)$ par $1 + X^3$, donc est égal à $f(\lambda P_1 + P_2)$.

Ainsi, $f(\lambda P_1 + P_2) = \lambda f(P_1) + f(P_2) : f$ est linéaire, et est donc un endomorphisme de E .

- 2.a. On a $(1 - X + X^2) \times 1 = (1 + X^3) \times 0 + 1 - X + X^2$, avec $\deg(1 - X + X^2) \leq 2$, donc $f(e_0) = 1 - X + X^2$.

De même, on a $(1 - X + X^2)X = X^3 - X^2 + X = (1 + X^3) \times 1 + \underbrace{(-X^2 + X - 1)}_{\in E}$, de sorte que

$$f(e_1) = -X^2 + X - 1.$$

Enfin, la division euclidienne de $(1 - X + X^2)X^2 = X^4 - X^3 + X^2$ par $1 + X^3$ est

X^4	$-X^3$	$+X^2$	X^3	$+1$
X^4		$+X$	X	-1
	$-X^3$	$+X^2$		
	$-X^3$			-1
		X^2		$-X$
		$+1$		$+1$

de sorte que $f(e_2) = X^2 - X + 1$.

Notons qu'on a alors bien $f(e_0) = -f(e_1) = f(e_2)$.

- 2.b. Si $P \in E$, alors il existe trois réels $\lambda_0, \lambda_1, \lambda_2$ tels que $P = \lambda_0 e_0 + \lambda_1 e_1 + \lambda_2 e_2$.
Et donc $f(P) = \lambda_0 f(e_0) + \lambda_1 f(e_1) + \lambda_2 f(e_2) = (\lambda_0 - \lambda_1 + \lambda_2)(1 - X + X^2) \in \text{Vect}(1 - X + X^2)$.

Par conséquent, $\text{Im } f \subset \text{Vect}(1 - X + X^2)$.

L'inclusion réciproque est évidente car $1 - X + X^2 = f(e_0) \in \text{Im } f$.

Et donc $\text{Im } f = \text{Vect}(1 - X + X^2)$.

Donc la famille formée du seul vecteur $1 - X + X^2$ est génératrice de $\text{Im } f$, et puisqu'il s'agit d'une famille formée d'un seul vecteur non nul, elle est libre : c'est une base de $\text{Im } f$.

- 2.c. D'après le théorème du rang, on a $\dim \text{Ker } f = \dim E - \dim \text{Im } f = 3 - 1 = 2$.
De plus, nous savons que $f(e_0 + e_1) = f(e_0) + f(e_1) = f(e_0) - f(e_0) = 0$ et de même $f(e_0 - e_2) = 0$. Donc $e_0 + e_1 = 1 + X$ et $e_0 - e_2 = 1 - X^2$ sont deux éléments de $\text{Ker } f$.
Puisqu'ils sont non nuls et de degré distincts, ils forment une famille libre de $\text{Ker } f$, de cardinal égal à la dimension de $\text{Ker } f$: $(1 + X, 1 - X^2)$ est une base de $\text{Ker } f$.

- 3.a. On a $f(1 - X + X^2) = f(e_0) - f(e_1) + f(e_2) = 3f(e_0) = 3(1 - X + X^2)$.
Si $P \in \text{Im } f$, alors il existe $\lambda \in \mathbf{R}$ tel que $P = \lambda(1 - X + X^2)$, et alors

$$f(P) = \lambda f(1 - X + X^2) = 3\lambda(1 - X + X^2) = 3P.$$

Donc tout polynôme non nul¹ de $\text{Im } f$ est vecteur propre de f pour la valeur propre 3 : on a $\text{Im } f \subset \text{Ker}(f - 3\text{id}_E) = E_3(f)$.

De plus, nous savons d'après la question 2.c que 0 est valeur propre de f , avec un sous-espace propre de dimension 2.

Puisque $\sum_{\lambda \in \text{Spec}(f)} \dim E_\lambda(f) \leq \dim E = 3$, nécessairement $\dim E_3(f) \leq 1$.

On en déduit que $\dim E_3(f) = 1$, et donc² que $E_3(f) = \text{Im } f$.

- 3.b. On a $\dim E_0(f) + \dim E_3(f) = 3 = \dim E$.
Donc f est diagonalisable, et ses seules valeurs propres sont 0 et 3.

Div. euclidienne

Rappelons que si on peut écrire $A = PQ + R$, avec $\deg(R) < \deg(Q)$, alors, par unicité de la division euclidienne, Q est le quotient de la division euclidienne de A par B et R est le reste.

Plus généralement

Si $f : E \rightarrow F$ est une application linéaire, alors l'image d'une famille génératrice de E par f est une famille génératrice de $\text{Im } f$.

En revanche, rien ne garantit que l'image d'une base soit une base de l'image, il faut alors vérifier la liberté à la main.

Exercice : montrer que l'image d'une base de E est une base de $\text{Im } f$ si et seulement si f est injective.

¹ Rappelons qu'un vecteur propre est non nul par définition.

² L'un est inclus dans l'autre et ils ont même dimension.

- 4.a. Soient $P = a_0 + a_1X + a_2X^2$, $Q = b_0 + b_1X + b_2X^2$ et $R = c_0 + c_1X + c_2X^2$ trois éléments de E et $\lambda \in \mathbf{R}$. Alors

$$\begin{aligned}\varphi(\lambda P + Q, R) &= \varphi((\lambda a_0 + b_0) + (\lambda a_1 + b_1)X + (\lambda a_2 + b_2)X^2, c_0 + c_1X + c_2X^2) \\ &= \sum_{i=0}^2 (\lambda a_i + b_i)c_i = \lambda \sum_{i=0}^2 a_i c_i + \sum_{i=0}^2 b_i c_i = \lambda \varphi(P, R) + \varphi(Q, R).\end{aligned}$$

Donc φ est linéaire à gauche.
Avec les mêmes notations, on a

$$\varphi(P, Q) = \sum_{i=0}^2 a_i b_i = \sum_{i=0}^2 b_i a_i = \varphi(Q, R).$$

Donc φ est symétrique, et donc bilinéaire symétrique.

Pour $P \in E$, on a $\varphi(P, P) = \sum_{i=0}^2 a_i^2 \geq 0$ et une somme de carrés étant nulle si et seulement si chacun de ses termes est nul, il vient

$$\varphi(P, P) = 0 \Leftrightarrow \forall i \in \llbracket 0, 2 \rrbracket, a_i = 0 \Leftrightarrow P = 0.$$

Donc φ est bien un produit scalaire sur E .

- 4.b. Nous savons qu'une base de $\text{Im } f$ est $1 - X + X^2$. Or, on a

$$\varphi(1 + X, 1 - X + X^2) = 1 - 1 + 0 = 0 \text{ et } \varphi(1 - X^2, 1 - X + X^2) = 1 + 0 - 1 = 0.$$

Donc $1 + X \in (\text{Im } f)^\perp$ et $1 - X^2 \in (\text{Im } f)^\perp$. Ainsi, $\text{Ker } f = \text{Vect}(1 + X, 1 - X^2) \subset (\text{Im } f)^\perp$.
De plus, $\dim \text{Ker } f = 2 = 3 - \dim \text{Im } f = \dim(\text{Im } f)^\perp$.

Donc $\text{Ker } f = (\text{Im } f)^\perp$: $\text{Ker } f$ est le supplémentaire orthogonal de $\text{Im } f$.

- 5.a. Il est facile de voir que $\varphi(e_0, e_0) = \varphi(e_1, e_1) = \varphi(e_2, e_2) = 1$.
De même, on a $\varphi(e_0, e_1) = \varphi(e_0, e_2) = \varphi(e_1, e_2) = 0$.

Donc, pour $(i, j) \in \llbracket 0, 2 \rrbracket^2$, on a $\varphi(e_i, e_j) = \begin{cases} 1 & \text{si } i = j \\ 0 & \text{si } i \neq j \end{cases}$: \mathcal{B} est une base orthonormale de

E pour le produit scalaire φ .

- 5.b. D'après les calculs effectués à la question 1, on a

$$\text{Mat}_{\mathcal{B}}(f) = \begin{pmatrix} f(e_1) & f(e_2) & f(e_3) \\ 1 & -1 & 1 \\ -1 & 1 & -1 \\ 1 & -1 & 1 \end{pmatrix} \begin{matrix} e_1 \\ e_2 \\ e_3 \end{matrix}.$$

La matrice de f dans la base orthonormée \mathcal{B} est symétrique, donc f est un endomorphisme symétrique.

Pour $P \in E$, on a $P \in (\text{Im } f)^\perp$ si et seulement si $\forall Q \in E, \varphi(P, f(Q)) = 0$.

Mais f étant symétrique, cela est équivalent à $\forall Q \in E, \varphi(f(P), Q) = 0$.

En particulier, pour $Q = f(P)$, $\varphi(f(P), f(P)) = 0 \Leftrightarrow \|f(P)\| = 0 \Leftrightarrow P \in \text{Ker } f$.

Nous avons donc $(\text{Im } f)^\perp \subset \text{Ker } f$.

Puisque d'autre part, $\dim(\text{Im } f)^\perp = \dim E - \dim \text{Im } f = \dim \text{Ker } f$,

il vient $\text{Ker } f = (\text{Im } f)^\perp$.

Rappel

Si φ est linéaire d'un côté (à gauche ou à droite) et symétrique, alors elle est automatiquement bilinéaire.

Orthogonal

Un élément $x \in E$ est dans F^\perp si et seulement si il est orthogonal à tous les éléments d'une base de F .

Image

$\text{Im } f$ est l'ensemble des $f(Q)$, $Q \in E$.

Remarque

Ce résultat reste valable pour tout endomorphisme symétrique d'un espace euclidien.

EML 2003

Dans tout le problème, n est un entier naturel supérieur ou égal à 2, et E est un espace euclidien de dimension n dont le produit scalaire est noté $\langle \cdot, \cdot \rangle$ et la norme associée est notée $\| \cdot \|$. On note id_E l'application identique de E , et $\tilde{0}$ l'application nulle de E .

Si F est un sous-espace vectoriel de E , on note F^\perp le sous-espace vectoriel supplémentaire orthogonal de F dans E .

Le projecteur de E sur F parallèlement à F^\perp est appelé projecteur orthogonal sur F .

Pour tout endomorphisme f de E et toute valeur propre λ de f , on note $E_\lambda(f)$ le sous-espace propre de f associé à la valeur propre λ .

Partie I : Inverse généralisé d'un endomorphisme symétrique.

On considère un endomorphisme symétrique f de E , c'est-à-dire un endomorphisme f tel que :

$$\forall (x, y) \in E^2, \langle f(x), y \rangle = \langle x, f(y) \rangle.$$

On suppose de plus que f est non inversible et non nul.

1. Montrer que 0 est valeur propre de f et que f admet au moins une valeur propre non nulle.
2. a. Soient λ et μ deux valeurs propres de f . Montrer, pour tout vecteur x de $E_\lambda(f)$ et pour tout vecteur y de $E_\mu(f)$:

$$\lambda \langle x, y \rangle = \mu \langle x, y \rangle.$$

- b. En déduire que les sous-espaces propres de f sont deux à deux orthogonaux.
3. Montrer que les sous-espaces vectoriels $\text{Im } f$ et $\text{Ker } f$ sont supplémentaires orthogonaux dans E .

On suppose que f admet exactement $k + 1$ valeurs propres deux à deux distinctes $\lambda_0, \lambda_1, \dots, \lambda_k$ avec $k \geq 1$, $\lambda_0 = 0$ et $0 < |\lambda_1| \leq \dots \leq |\lambda_k|$.

Pour tout entier naturel j inférieur ou égal à k , on note p_j le projecteur orthogonal sur $E_{\lambda_j}(f)$.

4. Soit x un vecteur de E .
 - a. Montrer qu'il existe un unique $(k+1)$ -uplet (x_0, x_1, \dots, x_k) de $E_0(f) \times E_{\lambda_1}(f) \times \dots \times E_{\lambda_k}(f)$ tel que $x = x_0 + x_1 + \dots + x_k$.
 - b. Pour tout entier naturel j inférieur ou égal à k , montrer : $p_j(x) = x_j$.

Ainsi, la relation suivante est clairement vérifiée :

$$\text{id}_E = p_0 + p_1 + \dots + p_k.$$

5. a. Établir, pour tout couple (i, j) d'entiers naturels inférieurs ou égaux à k :

$$i \neq j \implies p_i \circ p_j = \tilde{0}.$$

- b. Montrer : $f = \lambda_1 p_1 + \lambda_2 p_2 + \dots + \lambda_k p_k$.
- c. Montrer que le projecteur orthogonal p sur $\text{Im } f$ vérifie :

$$p = p_1 + p_2 + \dots + p_k.$$

On note $f^\#$ l'endomorphisme de E défini par $f^\# = \frac{1}{\lambda_1} p_1 + \frac{1}{\lambda_2} p_2 + \dots + \frac{1}{\lambda_k} p_k$.

On dit que $f^\#$ est l'inverse généralisé de f .

6. a. Montrer : $f \circ f^\# = p$.
- b. En déduire : $\forall (x, y) \in E^2, (f(x) = p(y) \iff x - f^\#(y) \in \text{Ker } f)$.
7. Soit y un vecteur de E .
 - a. Montrer : $\forall x \in E, (\|f(x) - y\| = \inf_{z \in E} \|f(z) - y\| \iff x - f^\#(y) \in \text{Ker } f)$.
 - b. En déduire que $f^\#(y)$ est le vecteur x de E de plus petite norme vérifiant :

$$\|f(x) - y\| = \inf_{z \in E} \|f(z) - y\|.$$

Partie II : Application à un exemple.

Dans cette question, E est un espace euclidien de dimension 4 et $\mathcal{B} = (e_1, e_2, e_3, e_4)$ est une base orthonormale de E . On note :

$$A = \begin{pmatrix} 3 & 0 & -1 & 0 \\ 0 & 1 & 0 & -1 \\ -1 & 0 & 3 & 0 \\ 0 & -1 & 0 & 1 \end{pmatrix}.$$

Soit f l'endomorphisme de E associé à la matrice A relativement à la base \mathcal{B} .

8. Justifier que f est un endomorphisme symétrique non nul et non inversible.

9. Montrer que f admet exactement trois valeurs propres distinctes $\lambda_0, \lambda_1, \lambda_2$ avec $\lambda_0 < \lambda_1 < \lambda_2$.

On note p_1 le projecteur orthogonal sur $E_{\lambda_1}(f)$ et M_1 la matrice associée à p_1 relativement à la base \mathcal{B} .

On note p_2 le projecteur orthogonal sur $E_{\lambda_2}(f)$ et M_2 la matrice associée à p_2 relativement à la base \mathcal{B} .

10. Montrer : $A = 2M_1 + 4M_2$.

11. a. Montrer que $E_{\lambda_2}(f)$ est de dimension 1 et déterminer un vecteur v_2 de $E_{\lambda_2}(f)$ tel que $\|v_2\| = 1$.

b. Montrer : $\forall x \in E, p_2(x) = \langle x, v_2 \rangle v_2$.

c. Déterminer la matrice M_2 .

12. En déduire la matrice associée à $f^\#$ relativement à la base \mathcal{B} .

EML 2003 : CORRECTION

Partie I : Inverse généralisé d'un endomorphisme symétrique.

1. Puisque f est non inversible, 0 est une valeur propre de f .

De plus, f est diagonalisable¹. S'il possédait 0 comme unique valeur propre, sa matrice dans une base de vecteurs propres serait la matrice nulle, et donc on aurait $f = 0$, ce qui n'est pas le cas par hypothèse.

Donc f possède au moins une valeur propre non nulle.

- 2.a. On a

$$\lambda \langle x, y \rangle = \langle \lambda x, y \rangle = \langle f(x), y \rangle = \langle x, f(y) \rangle = \langle x, \mu y \rangle = \mu \langle x, y \rangle.$$

- 2.b. Si λ et μ sont deux valeurs propres distinctes de f , alors pour $x \in E_\lambda(f)$ et $y \in E_\mu(f)$, on a

$$\lambda \langle x, y \rangle = \mu \langle x, y \rangle \Leftrightarrow (\lambda - \mu) \langle x, y \rangle = 0.$$

Puisque $\lambda - \mu \neq 0$, il vient donc $\langle x, y \rangle = 0$.

Et donc tout vecteur de $E_\lambda(f)$ est orthogonal à tout vecteur de $E_\mu(f)$, de sorte que $E_\lambda(f)$ et $E_\mu(f)$ sont orthogonaux.

3. soit $x \in \text{Ker } f$ et soit $y \in \text{Im } f$. Alors $f(x) = 0_E$, et il existe $z \in E$ tel que $f(z) = y$.
On a alors

$$\langle x, y \rangle = \langle x, f(z) \rangle = \langle f(x), z \rangle = \langle 0_E, z \rangle = 0.$$

Et donc $\text{Ker } f \subset (\text{Im } f)^\perp$.

D'autre part, nous savons que

$$\dim(\text{Im } f)^\perp = \dim E - \dim \text{Im } f = \dim \text{Ker } f.$$

Et donc puisqu'on a une inclusion entre deux espaces de même dimension, il s'agit d'une égalité : $\text{Ker } f = (\text{Im } f)^\perp$.

Et donc $\text{Ker } f$ est le supplémentaire orthogonal de $\text{Im } f$.

- 4.a. Puisque f est diagonalisable, $E = \bigoplus_{i=0}^k E_{\lambda_i}(f)$.

Et donc x , comme tout vecteur de E , s'écrit de manière unique

$$x = \sum_{i=0}^k x_i, \quad (x_0, x_1, \dots, x_k) \in E_0(f) \times E_{\lambda_1}(f) \times \dots \times E_{\lambda_k}(f).$$

- 4.b. Soit $j \in \llbracket 0, k \rrbracket$. D'après la question 2, pour $i \neq j$, $E_{\lambda_i}(f)$ est orthogonal à $E_{\lambda_j}(f)$.

Ce qui s'écrit encore $E_{\lambda_i}(f) \subset E_{\lambda_j}(f)^\perp = \text{Ker } p_j$.

Et donc

$$p_j(x) = \sum_{i=0}^k p_j(x_i) = p_j(x_j) + \sum_{\substack{i=0 \\ i \neq j}}^k \underbrace{p_j(x_i)}_{=0_E} = p_j(x_j) = \boxed{x_j}.$$

- 5.a. Soit $x \in E$, et reprenons les notations précédentes : $x = x_0 + \dots + x_k$.

Alors pour $i \neq j$, on a $p_j(x) = x_j \in E_{\lambda_j}(f) \subset E_{\lambda_i}(f)^\perp = \text{Ker}(p_i)$.

Et donc $p_i(p_j(x)) = 0_E$. Ceci étant vrai pour tout $x \in E$, on en déduit que

$$\boxed{p_i \circ p_j = \tilde{0}}.$$

- 5.b. Soit $x \in E$, $x = x_0 + x_1 + \dots + x_k$.

Alors, par linéarité de f ,

$$f(x) = f(x_0) + f(x_1) + \dots + f(x_k) = 0 \cdot x_0 + \lambda_1 \cdot x_1 + \dots + \lambda_k \cdot x_k = \lambda_1 \cdot p_1(x_1) + \dots + \lambda_k \cdot p_k(x_k).$$

Et ceci étant vrai pour tout $x \in E$, on a $\boxed{f = \lambda_1 p_1 + \dots + \lambda_k p_k}$.

¹ Car symétrique.

Remarque

Notons que ce résultat est un résultat de cours. Mais si on nous le demande, c'est visiblement qu'on en attend une preuve.

Danger !

Nous n'avons pas ici prouvé une égalité, seulement une inclusion.

Notons qu'on pourrait également affirmer que $(\text{Im } f) \subset (\text{Ker } f)^\perp$, ce qui est équivalent à l'inclusion que nous avons donnée, et n'est pas l'inclusion réciproque ! Il va nous falloir utiliser un argument de dimension pour cela.

Détails

Puisque $x_j \in E_{\lambda_j}(f) = \text{Im } p_j$, on a

$$p_j(x_j) = x_j.$$

Détails

Puisque $x_i \in E_{\lambda_i}(f)$,

$$f(x_i) = \lambda_i x_i.$$

- 5.c. Soit $x \in E$. Alors de manière unique, $x = x_0 + x_1 + \dots + x_k$, avec $x_i \in E_{\lambda_i}(f)$.
En particulier, pour $i \neq 0$, on a

$$x_i \in E_{\lambda_i}(f) \Leftrightarrow f(x_i)\lambda_i x_i \Leftrightarrow x_i = f\left(\frac{1}{\lambda_i}x_i\right).$$

Et par conséquent, $x_i \in \text{Im } f$.

Puisque $\text{Im } f$ est stable par combinaisons linéaires, $x_1 + \dots + x_k \in \text{Im } f$.

Et donc $x = x_0 + (x_1 + \dots + x_k)$ est l'unique décomposition de x comme somme d'un élément de $\text{Ker } f$ plus un élément de $\text{Im } f$.

Et alors, par définition de p , $p(x) = x_1 + \dots + x_k = p_1(x) + \dots + p_k(x)$.

Ceci étant vrai pour tout $x \in E$, $p = p_1 + \dots + p_k$.

Alternative : nous savons déjà que $\text{Im } p = (\text{Ker } p)^\perp$. Et donc p est la projection sur $\text{Im } p$ parallèlement à $(\text{Im } p)^\perp = \text{Ker } p$, alors que dans le même temps, p_0 est la projection sur $\text{Ker } p$ parallèlement à $(\text{Ker } p)^\perp = \text{Im } p$.

Et donc p et p_0 sont liés par la relation $p + p_0 = \text{id}$.

Or nous savons déjà que $\text{id} = p_0 + p_1 + \dots + p_k$, donc $p = p_1 + \dots + p_k$.

- 6.a. On a

$$\begin{aligned} f \circ f^\# &= \left(\sum_{i=0}^k \lambda_i p_i \right) \circ \left(\sum_{j=1}^k \frac{1}{\lambda_j} p_j \right) \\ &= \sum_{i=0}^k \sum_{j=1}^k \frac{\lambda_i}{\lambda_j} p_i \circ p_j \\ &= \sum_{i=1}^k \frac{\lambda_i}{\lambda_i} p_i = p. \end{aligned}$$

Détails

Si vous n'êtes pas convaincus par cette relation, écrivez $x \in E$ comme somme d'un élément de $\text{Im } p$ plus un élément de $\text{Ker } p$, et constatez que

$$(p + p_0)(x) = x.$$

Si $i \neq j$, alors $p_i \circ p_j = 0$.

- 6.b. Soient $x, y \in E$. Alors

$$f(x) = p(y) \Leftrightarrow f(x) - f(f^\#(y)) = 0_E \Leftrightarrow f(x - f^\#(y)) = 0_E \Leftrightarrow x - f^\#(y) \in \text{Ker } f.$$

- 7.a. Lorsque z parcourt E , $f(z)$ parcourt $\text{Im } f$. Et donc pour y fixé,

$$\inf_{z \in E} \|f(z) - y\| = \inf_{u \in \text{Im } f} \|u - y\|.$$

Mais un théorème du cours nous garantit que cette borne inférieure est en réalité un minimum, atteint uniquement pour $u = p(y)$.

Et donc

$$\|f(x) - y\| = \inf_{z \in E} \|f(z) - y\| \Leftrightarrow f(x) = p(y) \Leftrightarrow x - f^\#(y) \in \text{Ker } f.$$

- 7.b. Pour commencer, notons que

$$\|f(f^\#(y)) - y\| = \|p(y) - y\| = \min_{u \in \text{Im } f} \|u - y\| = \inf_{z \in E} \|f(z) - y\|.$$

Il reste donc à montrer que si x est un autre vecteur vérifiant $\|f(x) - y\| = \inf_{z \in E} \|f(z) - y\|$, alors $\|x\| \geq \|f^\#(y)\|$.

Soit x un tel vecteur. D'après la question précédente, on a $x - f^\#(y) \in \text{Ker } f$.

Et donc $x - f^\#(y) \in \text{Ker } f$. Or, $f^\#(y) \in \text{Im } f = (\text{Ker } f)^\perp$.

Il vient alors, par le théorème de Pythagore,

$$\|x\|^2 = \underbrace{\|x - f^\#(y)\|}_{\in \text{Ker } f}^2 + \underbrace{\|f^\#(y)\|}_{\in (\text{Ker } f)^\perp}^2 = \|f^\#(y)\|^2 + \|x - f^\#(y)\|^2.$$

Et donc $\|x\| \geq \|f^\#(y)\|$, avec égalité si et seulement si

$$\|x - f^\#(y)\| = 0 \Leftrightarrow x = f^\#(y).$$

Et donc $f^\#(y)$ est l'unique vecteur de norme minimale parmi ceux vérifiant $\|f(x) - y\| = \inf_{z \in E} \|f(z) - y\|$.

Unicité

Si on oublie d'étudier le cas d'égalité, on aura bien prouvé que $f^\#(y)$ est de norme minimale parmi les vecteurs vérifiant la condition requise, mais il se pourrait encore tout à fait qu'il existe plusieurs vecteurs ayant tous la même norme.

Partie II : Application à un exemple.

8. La base \mathcal{B} est orthonormale, et A est symétrique, donc f est un endomorphisme symétrique.

Il est évidemment non nul puisque $A \neq 0$.

Enfin, notons que la seconde et la dernière colonne de A sont proportionnelles, de sorte que $\text{rg}(A) \leq 3 < 4$, et donc A n'est pas inversible, donc f n'est pas bijectif.

9. Nous savons déjà que 0 est valeur propre de A car A n'est pas inversible.

De plus, notons que $A - 2I_4 = \begin{pmatrix} 1 & 0 & -1 & 0 \\ 0 & -1 & 0 & -1 \\ -1 & 0 & 1 & 0 \\ 0 & -1 & 0 & -1 \end{pmatrix}$, qui est de rang 2, car la première et la

troisième colonne sont colinéaires, de même que la seconde et la dernière. Et de plus, la première et la seconde ne sont pas colinéaires.

Donc 2 est valeur propre de A et $\dim E_2(A) = 4 - \text{rg}(A - 2I_4) = 2$.

Enfin, A étant symétrique, elle est diagonalisable, et elle possède donc une dernière valeur propre λ , et donc est semblable à $D = \text{Diag}(0, 2, 2, \lambda)$.

Mais alors $\text{tr}(A) = \text{tr}(D) \Leftrightarrow 4 + \lambda = 8 \Leftrightarrow \lambda = 4$.

Et donc $\text{Spec}(A) = \{0, 2, 4\}$.

10. D'après la question 5.b, en notant que $\lambda_1 = 2$ et $\lambda_2 = 4$, on a $f = 2p_1 + 4p_2$.

Et donc matriciellement, cela se traduit en $A = 2M_1 + 4M_2$.

11.a. Puisque $\sum_{i=0}^2 \dim E_{\lambda_i}(f) = 4$, et que $\dim E_{\lambda_1}(f) = 2$ et $\dim E_{\lambda_0}(f) \geq 1$, nécessairement,

$\dim E_{\lambda_2}(f) = 1$.

On a

$$A - 4I_4 = \begin{pmatrix} -1 & 0 & -1 & 0 \\ 0 & -3 & 0 & -1 \\ -1 & 0 & -1 & 0 \\ 0 & -1 & 0 & -3 \end{pmatrix}$$

et les colonnes de cette matrice sont liées par la relation $C_1 = C_3 \Leftrightarrow C_1 - C_3 = 0$.

Donc $\begin{pmatrix} 1 \\ 0 \\ -1 \\ 0 \end{pmatrix}$ est un vecteur propre de A , associé à la valeur propre 4.

Et donc $e_1 - e_3$ est un vecteur propre de f associé à la valeur propre 4.

Pour trouver un vecteur unitaire de $E_{\lambda_2}(f)$, il suffit de diviser ce vecteur par sa norme. Or

$$\|e_1 - e_3\|^2 = 1^2 + 1^2 = 2$$

de sorte que $v_2 = \frac{1}{\sqrt{2}}(e_1 - e_3)$ est un vecteur unitaire de $E_{\lambda_2}(f)$.

11.b. Nous venons de prouver que v_2 est une base orthonormée de $E_{\lambda_2}(f)$.

Et donc une formule bien connue pour les projecteurs orthogonaux nous donne directement : $\forall x \in E, p_2(x) = \langle x, v_2 \rangle v_2$.

11.c. On a $p_2(e_1) = \langle e_1, v_2 \rangle v_2 = \frac{e_1 - e_3}{2}$, $p_2(e_3) = \langle e_3, v_2 \rangle v_2 = \frac{e_3 - e_1}{2}$ et $p_2(e_2) = p_2(e_4) = 0_E$.

Et donc

$$M_2 = \begin{pmatrix} p_2(e_1) & p_2(e_2) & p_2(e_3) & p_2(e_4) \\ \frac{1}{2} & 0 & -\frac{1}{2} & 0 \\ 0 & 0 & 0 & 0 \\ -\frac{1}{2} & 0 & \frac{1}{2} & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \begin{matrix} e_1 \\ e_2 \\ e_3 \\ e_4 \end{matrix}$$

12. On a donc $M_1 = \frac{1}{2}(A - 4M_2)$ et par conséquent,

$$\text{Mat}_{\mathcal{B}}(f^\#) = \frac{1}{2}M_1 + \frac{1}{4}M_2 = \frac{1}{4}A - \frac{3}{4}M_2 = \begin{pmatrix} 3 & 0 & 1 & 0 \\ 0 & 2 & 0 & -2 \\ 1 & 0 & 3 & 0 \\ 0 & -2 & 0 & 2 \end{pmatrix}$$

Remarque

Notons que nous n'avons pas calculé le rang de A , et que l'on pourrait encore avoir $\lambda = 0$.

Rappel

Dans une base orthonormée (e_1, \dots, e_n) , si $x = \sum_{i=1}^n x_i e_i$, alors

$$\|x\|^2 = \sum_{i=1}^n x_i^2$$

DEVOIRS À LA MAISON

Devoir maison 0 (à rendre le 03.09.18).	708
Correction	712
Devoir maison 1 (à rendre le 13.09.18).	724
Correction	726
Devoir maison 2 (à rendre le 20.09.18).	729
Correction	730
Devoir maison 2 sujet dur (à rendre le 20.09.18)	733
Correction	735
Devoir maison 3 (à rendre le 04.10.18).	742
Correction	744
Devoir maison 4 (à rendre le 11.10.18).	748
Correction	750
Devoir maison 4 bis (à rendre le 11.10.18)	755
Correction	758
Devoir maison 4 (à rendre le 11.10.18).	768
Correction	771
Devoir maison 5 (à rendre le 18.10.18).	782
Correction	783
Devoir maison 6 (à rendre le 15.11.18).	785
Correction	787
Devoir maison 7 (à rendre le 22.11.18).	793
Correction	795
Devoir maison 8 (à rendre le 6.12.18)	800
Correction	801
Devoir maison 9 (à rendre le 13.12.18).	806
Correction	808
Devoir maison 9 (HEC) (à rendre le 20.12.18).	812
Correction	815
Devoir maison 10 (à rendre le 20.12.18)	824
Correction	825
Devoir maison 11 (à rendre le 10.01.19)	827
Correction	831
Devoir maison 12 (à rendre le 17.01.19)	846
Correction	847
Devoir maison 13 (à rendre le 25.01.19)	850
Correction	851
Devoir maison 14 (à rendre le 05.02.19)	853
Correction	855
Devoir maison 15 (SciLab) (à rendre le 14.03.19).	862
Correction	864
Devoir maison 16 (Maths II) (à rendre le 14.03.19)	868
Correction	871
Devoir maison 17 (à rendre le 21.03.19)	883
Correction	884

DEVOIR MAISON 0 (À RENDRE LE 03.09.18)

Vous traiterez **au moins** un exercice d'algèbre, un exercice d'analyse et un exercice de probabilités. Pour chaque thème, le second exercice proposé est plus difficile que le premier.

Si vous avez **sérieusement** cherché une question, n'hésitez pas à demander une indication par e-mail :

vienney@ecs2-fauriel.fr

EXERCICE 1 : ALGÈBRE

Si k est un entier naturel non nul, alors pour tout endomorphisme f d'un \mathbf{R} -espace vectoriel E , on note $f^k = \underbrace{f \circ \dots \circ f}_{k \text{ termes}}$

et on pose $f^0 = \text{Id}_E$, où Id_E est l'endomorphisme identité de E .

On dit que l'endomorphisme f est nilpotent d'indice k ($k \in \mathbf{N}^*$) si l'on a :

$$f^k = 0 \text{ et } f^{k-1} \neq 0.$$

On note I_2 la matrice identité de $\mathcal{M}_2(\mathbf{R})$ et on dit qu'une matrice A de $\mathcal{M}_2(\mathbf{R})$ est nilpotente d'indice k ($k \in \mathbf{N}^*$) si l'on a $A^k = 0$ et $A^{k-1} \neq 0$ (avec la convention $A^0 = I_2$).

Partie 1

Soit $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ une matrice non nulle de $\mathcal{M}_2(\mathbf{R})$.

1. Calculer $A^2 - (a+d)A$ en fonction de I_2 .
2. On suppose dans cette question que A est nilpotente d'indice k .
 - a. Établir l'égalité : $ad - bc = 0$.
 - b. Montrer que k est supérieur ou égal à 2.
 - c. En déduire alors que $a + d = 0$.
3. Conclure que : A nilpotente $\iff A^2 = 0$.

Partie 2

On considère dans cette partie un endomorphisme f non nul d'un \mathbf{R} -espace vectoriel de E de dimension 2.

4.
 - a. Montrer que, si $\text{Ker}(f) = \text{Im}(f)$, alors on a : $f^2 = 0$.
 - b. On suppose que $f^2 = 0$. Montrer que $\text{Im}(f) \subset \text{Ker}(f)$. Établir alors que $\text{rg}(f) = 1$ puis conclure que $\text{Ker}(f) = \text{Im}(f)$.
 - c. En déduire, à l'aide de la partie 1, l'équivalence :

$$f \text{ nilpotent} \iff \text{Ker}(f) = \text{Im}(f).$$

EXERCICE 2 : ALGÈBRE

Soit $n \in \mathbf{N}^*$. Soit $E = \mathbf{R}_n[X]$. On note $E^* = \mathcal{L}(E, \mathbf{R})$ l'espace vectoriel des formes linéaires sur E .

Pour tout polynôme P , on note $P^{(k)}$ la dérivée $k^{\text{ème}}$ de P .

Soit $Q \in E$ de degré n . Pour tout $i \in \llbracket 0, n \rrbracket$, on pose $Q_i(X) = Q(X+i)$.

Le but de cet exercice est de montrer que (Q_0, Q_1, \dots, Q_n) constitue une base de E .

1. Montrer que la famille $(Q, Q', Q'', \dots, Q^{(n)})$ est une base de E .
2. Pour tout $k \in \llbracket 0, n \rrbracket$, on définit l'application linéaire f_k par :

$$f_k : E \rightarrow \mathbf{R}, P \mapsto P^{(k)}(0).$$

- a. Montrer que pour tout $k \in \llbracket 0, n \rrbracket$, f_k est dans E^* .
 - b. Pour tout couple d'entiers $(k, \ell) \in \llbracket 0, n \rrbracket^2$, calculer $f_k(X^\ell)$.
 - c. Montrer que (f_0, f_1, \dots, f_n) est une famille libre d'éléments de E^* .
3. Soit $\varphi \in E^*$.
 - a. Déterminer la dimension de E^* .

b. Montrer qu'il existe un unique $(n+1)$ -uplet $(a_0, a_1, \dots, a_n) \in \mathbf{R}^{n+1}$ tel que :

$$\forall P \in E, \varphi(P) = \sum_{i=0}^n a_i P^{(i)}(0).$$

c. Soit $\varphi \in E^*$ tel que $\varphi(Q_0) = \varphi(Q_1) = \dots = \varphi(Q_n) = 0$. Montrer que $\varphi = 0$.

Indication : on pourra commencer par montrer que le polynôme $\sum_{i=0}^n a_i Q^{(i)}$ est nul.

4. On suppose que (Q_0, Q_1, \dots, Q_n) n'est pas une base de E .

a. Montrer qu'il existe un sous-espace vectoriel H de E tel que $\dim(H) = n$ et $\text{Vect}(Q_0, Q_1, \dots, Q_n) \subset H$.

b. Montrer qu'il existe $a \in E$ tel que $E = \text{Vect}(a) \oplus H$.

c. On définit une application $\psi_a : E \rightarrow \mathbf{R}$ comme suit : pour tout $x \in E$ s'écrivant $x = \mu a + h$ avec $\mu \in \mathbf{R}$ et $h \in H$, on pose $\psi_a(x) = \mu$.

Montrer que ψ_a est une forme linéaire sur E et déterminer son noyau.

d. Aboutir à une contradiction et conclure.

EXERCICE 3 : ANALYSE

On considère la fonction f qui à tout réel associe : $f(x) = \int_0^x \ln(1+t^2) dt$.

1. a. Déterminer le signe de $f(x)$ selon le signe de x .

b. Justifier que f est de classe \mathcal{C}^1 sur \mathbf{R} et calculer $f'(x)$ pour tout réel x .

c. En déduire les variations de f sur \mathbf{R} (on ne cherchera pas à calculer les limites de f .)

2. a. Montrer que f est impaire.

b. Étudier la convexité de la fonction f et donner les coordonnées des éventuels points d'inflexion de la courbe représentative de f dans un repère orthonormé.

3. a. Déterminer les réels a et b tels que :

$$\forall t \in \mathbf{R}, \frac{t^2}{1+t^2} = a + \frac{b}{1+t^2}.$$

b. En déduire, grâce à une intégration par parties, que pour tout réel x , on a :

$$f(x) = x \left(\ln(1+x^2) - 2 \right) + 2\text{Arctan}(x).$$

4. Recherche d'un équivalent de $f(x)$ au voisinage de $+\infty$.

a. Montrer $f(x) \underset{x \rightarrow +\infty}{\sim} x \ln(1+x^2)$.

b. Vérifier que pour tout réel x strictement positif, on a $\ln(1+x^2) = 2 \ln(x) + \ln\left(1 + \frac{1}{x^2}\right)$ puis établir l'équivalent suivant : $f(x) \underset{x \rightarrow +\infty}{\sim} 2x \ln(x)$.

c. Donner sans calcul un équivalent de $f(x)$ lorsque x est au voisinage de $-\infty$.

5. Recherche d'un équivalent de $f(x)$ au voisinage de 0.

a. Justifier que f est de classe \mathcal{C}^3 , et calculer $f(0)$, $f'(0)$, $f''(0)$ et $f^{(3)}(0)$.

b. En déduire un équivalent de $f(x)$ lorsque x tend vers 0.

EXERCICE 4 : ANALYSE

1. Pour tout $n \in \mathbf{N}^*$, on pose $H_n = \sum_{k=1}^n \frac{1}{k}$ et $\gamma_n = H_n - \ln(n)$.

Montrer que la série de terme général $\gamma_{n+1} - \gamma_n$ converge.

En déduire que la suite $(\gamma_n)_{n \geq 1}$ converge. On note γ sa limite.

2. Montrer que pour tout $n \in \mathbf{N}^*$, on a : $\sum_{k=1}^{2n} \frac{(-1)^k}{k} = H_n - H_{2n}$.

En déduire que la série $\sum_{k \geq 1} \frac{(-1)^k}{k}$ converge et déterminer sa somme.

3. Dans cette question, on pose, pour tout $n \in \mathbf{N}$, $a_{3n+1} = a_{3n+2} = 1$ et $a_{3n+3} = -1$.

Déterminer $\lim_{n \rightarrow +\infty} \sum_{k=1}^{3n+3} \frac{a_k}{k}$ et en déduire la nature de la série de terme général $\frac{a_k}{k}$.

Dans la suite, on pose pour tout $n \in \mathbf{N}$, $a_{4n+1} = a_{4n+2} = 1$ et $a_{4n+3} = a_{4n+4} = -1$.

4. a. Montrer que pour tout $N \in \mathbf{N}$, on a $\sum_{n=0}^N \left(\frac{1}{4n+1} - \frac{1}{4n+3} \right) = \int_0^1 \frac{1-x^{4N+4}}{1+x^2} dx$.

b. En déduire que $\sum_{n=0}^{+\infty} \left(\frac{1}{4n+1} - \frac{1}{4n+3} \right) = \frac{\pi}{4}$.

c. Calculer de même $\sum_{n=0}^{+\infty} \left(\frac{1}{4n+2} - \frac{1}{4n+4} \right)$.

5. Montrer que $\sum_{n=1}^{+\infty} \frac{a_n}{n} = \frac{\pi}{4} + \frac{1}{2} \ln(2)$.

EXERCICE 5 : PROBABILITÉS

On considère deux jetons J_1 , et J_2 , équilibrés (c'est-à-dire tels que chaque face a une chance sur deux d'apparaître au cours d'un lancer).

Le jeton J_1 possède une face numérotée 0 et une face numérotée 1.

Le jeton J_2 possède deux faces numérotées 1.

Un joueur choisit au hasard un jeton puis effectue une série de lancers avec ce jeton.

On note E l'événement «le jeton J_1 est choisi pour le jeu» et, pour tout entier naturel k non nul, U_k l'événement «le k -ième lancer fait apparaître une face numérotée 1».

1. a. Déterminer la probabilité que le joueur obtienne n fois ($n \in \mathbf{N}^*$) une face portant le numéro 1 lors des n premiers lancers.
- b. Dans cette question, on suppose que le joueur a obtenu n fois ($n \in \mathbf{N}^*$) une face portant le numéro 1 lors des n premiers lancers. Quelle est la probabilité qu'il ait joué avec le jeton J_1 ?
Quelle est la limite de cette probabilité lorsque n tend vers $+\infty$? Interpréter ce résultat.

Dans la suite, on considère la variable aléatoire X égale au rang d'apparition de la première face portant le numéro 0 et on pose $X = 0$ si la face portant le numéro 0 n'apparaît jamais.

On considère également la variable aléatoire Y égale au rang d'apparition de la première face portant le numéro 1 et on pose $Y = 0$ si la face portant le numéro 1 n'apparaît jamais.

On suppose ces variables aléatoires définies sur le même espace probabilisé (Ω, \mathcal{A}, P) .

2. a. Calculer, pour tout entier naturel n non nul, la probabilité $P(X = n)$.
- b. En déduire que $P(X = 0) = \frac{1}{2}$. Ce résultat était-il prévisible ?
- c. Montrer que X a une espérance puis déterminer $E(X)$.
- d. Montrer que $X(X - 1)$ a une espérance, la déterminer puis vérifier que $V(X) = 2$.
3. a. Calculer, pour tout entier naturel n non nul, la probabilité $P(Y = n)$.
- b. En déduire que $P(Y = 0) = 0$.
- c. Montrer que Y a une espérance puis déterminer $E(Y)$.
- d. Montrer que $Y(Y - 1)$ a une espérance, la déterminer puis vérifier que $V(Y) = \frac{5}{4}$.
4. On définit sur (Ω, \mathcal{A}, P) la variable aléatoire S par : $\forall \omega \in \Omega, S(\omega) = \max(X(\omega), Y(\omega))$.
 - a. Déterminer $S(\Omega)$.
 - b. Montrer que $P(S = 1) = P(X = 0) = \frac{1}{2}$.
 - c. Pour tout entier n supérieur ou égal à 2, comparer d'une part $[X = n]$ et $[Y < n]$ et d'autre part $[Y = n]$ et $[X < n]$, puis en déduire que : $[S = n] = [X = n] \cup [Y = n]$.
 - d. Reconnaître alors la loi de S et préciser son espérance et sa variance.
5. On définit sur (Ω, \mathcal{A}, P) la variable aléatoire I par : $\forall \omega \in \Omega, I(\omega) = \min(X(\omega), Y(\omega))$.
 - a. Montrer que I est une variable de Bernoulli.
 - b. Déterminer $P(I = 0)$ puis donner la loi de I , ainsi que son espérance et sa variance.

EXERCICE 6 : PROBABILITÉS

Soit $n \in \mathbf{N}^*$. On considère une urne U_n contenant n boules numérotées de 1 à n . On tire au hasard une boule de U_n . Soit k le numéro de la boule tirée.

- Si $k = 1$, on arrête les tirages.
- Sinon, on retire de l'urne toutes les boules numérotées de k à n , et on effectue un nouveau tirage.

On répète ces tirages jusqu'à l'obtention de la boule numéro 1.

Soit X_n la variable aléatoire égale au nombre de tirages effectués jusqu'à l'obtention de la boule 1.

Pour tout $k \geq 1$, on note Y_k la variable aléatoire égale au numéro de la boule tirée lors du $k^{\text{ème}}$ tirage si celui-ci a lieu et 0 sinon.

Enfin, on note I_n la variable aléatoire égale au numéro de la première boule tirée dans l'urne U_n .

1. a. Quelle est la loi de la variable aléatoire I_n ?
- b. Déterminer les $P_{[I_n=1]}(X_n = k)$, $k \in \mathbf{N}^*$.
- c. Pour tout $n \geq 2$, montrer que

$$\forall j \in \mathbf{N}^*, \forall k \in \llbracket 2, n \rrbracket, P_{[I_n=k]}(X_n = j) = P(X_{k-1} = j - 1).$$

2. a. Quelle est la loi de la variable aléatoire X_1 ?
- b. Quelle est la loi de X_2 ? Calculer $E(X_2)$.
3. a. Déterminer $X_n(\Omega)$.
- b. Calculer $P(X_n = 1)$ et $P(X_n = n)$.
- c. Montrer que : $\forall n \geq 2, \forall j \geq 2, P(X_n = j) = \frac{1}{n} \sum_{k=1}^{n-1} P(X_k = j - 1)$.
- d. Pour tout $n \geq 2$ et tout $j \geq 2$, calculer $nP(X_n = j) - (n - 1)P(X_{n-1} = j)$.
En déduire que pour tout $n \geq 2$ et tout $j \geq 1$:

$$P(X_n = j) = \frac{n-1}{n} P(X_{n-1} = j) + \frac{1}{n} P(X_{n-1} = j - 1).$$

4. Montrer que : $\forall n \geq 2, E(X_n) = E(X_{n-1}) + \frac{1}{n}$. En déduire $E(X_n)$.

CORRECTION DU DEVOIR MAISON 0

EXERCICE 1 (EDHEC 2018)

Partie 1

1. On a

$$\begin{aligned} A^2 - (a+d)A &= \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix} - (a+d) \begin{pmatrix} a & b \\ c & d \end{pmatrix} \\ &= \begin{pmatrix} a^2 + bc & ab + bd \\ ac + cd & cb + d^2 \end{pmatrix} - \begin{pmatrix} (a+d)a & (a+d)b \\ (a+d)c & (a+d)d \end{pmatrix} = \begin{pmatrix} bc - ad & 0 \\ 0 & bc - ad \end{pmatrix} \\ &= \boxed{-(ad - bc)I_2}. \end{aligned}$$

Remarque

Ceci prouve que

$X^2 - (a+d)X + (ad - bc)$
est un polynôme annulateur
de A .

2.a. Nous savons que $ad - bc$ est le déterminant de A , et qu'il est nul si et seulement si A n'est pas inversible.

Or, si A était inversible, en multipliant à gauche par A^{-1} l'égalité $A^k = A \times A^{k-1} = 0$, il viendrait $A^{k-1} = 0$, ce qui est contraire aux hypothèses.

Donc A n'est pas inversible, et par conséquent, $\boxed{ad - bc = 0}$.

2.b. Une matrice nilpotente d'indice 1 est une matrice telle que $A^1 = A = 0$.

Or ici, nous avons supposé A non nulle, donc $\boxed{k \geq 2}$.

2.c. Puisque $ad - bc = 0$, d'après le résultat de la question 1, $A^2 = (a+d)A$.

Alors $A^3 = A^2A = (a+d)AA = (a+d)^2A$, et une récurrence rapide prouve que pour tout $n \geq 2$, $A^n = (a+d)^{n-1}A$.

Et en particulier $A^k = (a+d)^{k-1}A$.

Mais $A^k = 0$ et $A \neq 0$, donc nécessairement $(a+d)^{k-1} = 0 \Leftrightarrow \boxed{a+d=0}$.

3. Si A est nilpotente, nous venons de prouver que $A^2 = (a+d)A = 0$.

Et inversement, si $A^2 = 0$, alors A est nilpotente d'indice 2.

Donc $\boxed{A \text{ est nilpotente si et seulement si } A^2 = 0}$.

Autrement dit

Une matrice nilpotente non nulle de $\mathcal{M}_2(\mathbf{R})$ possède nécessairement un indice de nilpotence égal à 2.
Plus généralement, on peut montrer qu'une matrice nilpotente de $\mathcal{M}_n(\mathbf{R})$ possède un indice de nilpotence au plus égal à n .

Partie 2

4.a. Supposons que $\text{Ker}(f) = \text{Im}(f)$. Soit alors $x \in E$.

Par définition de l'image, le vecteur $f(x)$ est dans $\text{Im}(f)$, et par conséquent dans $\text{Ker}(f)$.

Et donc $f^2(x) = f(f(x)) = 0_E$. Ceci étant vrai pour tout $x \in E$, on en déduit que $\boxed{f^2 = 0}$.

4.b. Soit $y \in \text{Im } f$. Ainsi, il existe $x \in E$ tel que $y = f(x)$.

Et alors $f(y) = f^2(x) = 0_E$, de sorte que $y \in \text{Ker}(f)$.

Ceci prouve donc que $\boxed{\text{Im } f \subset \text{Ker } f}$.

On en déduit donc que $\dim \text{Ker } f \geq \dim \text{Im } f = \text{rg}(f)$.

Mais d'autre part, d'après le théorème du rang, $\dim \text{Ker } f + \text{rg}(f) = \dim E = 2$.

On a donc soit $\text{rg}(f) = 1$, soit $\text{rg}(f) = 0$.

Mais f étant non nul, ce second cas n'est pas possible¹, et donc $\boxed{\text{rg}(f) = 1}$.

Et alors, toujours par le théorème du rang, $\dim \text{Ker } f = 2 - \text{rg}(f) = 1$.

Puisque $\text{Im}(f) \subset \text{Ker}(f)$, ces deux espaces sont donc égaux : $\boxed{\text{Im}(f) = \text{Ker}(f)}$.

4.c. Puisque la composition des endomorphismes correspond au produit des matrices, un endomorphisme est nilpotent d'indice k si et seulement si sa matrice dans n'importe quelle base est nilpotente d'indice k .

Ainsi, si \mathcal{B} est une base de E et si $A = \text{Mat}_{\mathcal{B}}(f)$, alors f est nilpotent si et seulement si A est nilpotente, et donc d'après la partie 1, si et seulement si $A^2 = 0 \Leftrightarrow f^2 = 0$.

Or, les questions 4.a et 4.b prouvent que $f^2 = 0$ si et seulement si $\text{Im}(f) = \text{Ker}(f)$.

Donc $\boxed{f \text{ est nilpotent si et seulement si } \text{Im}(f) = \text{Ker}(f)}$.

Remarque

Le raisonnement reste valable si on suppose seulement $\text{Im}(f) \subset \text{Ker}(f)$.

Plus généralement

Sur le même principe, on peut prouver que $f \circ g = 0$ si et seulement si $\text{Im } g \subset \text{Ker } f$.

¹ Une application linéaire de rang 0 possède $\{0_E\}$ comme image, ce qui n'est possible que pour l'application nulle.

Détails

A^2 est la matrice de f^2 dans la base \mathcal{B} et donc elle est nulle si et seulement si f^2 est nul.

EXERCICE 2 (D'APRÈS ORAL ESCP 2015)

1. Puisque Q est de degré n , Q' est de degré $n-1$, Q'' est de degré $n-2$, et plus généralement, pour tout $k \in \llbracket 0, n \rrbracket$, $Q^{(k)}$ est de degré $n-k$.

Donc $(Q, Q', \dots, Q^{(n)})$ est une famille de polynômes de $\mathbf{R}_n[X]$ de degrés deux à deux distincts : elle est libre.

Étant de cardinal $n + 1 = \dim \mathbf{R}_n[X]$, c'est donc une base de $\mathbf{R}_n[X]$.

- 2.a. Par définition, f_k est à valeurs dans \mathbf{R} , donc il suffit de prouver qu'elle est linéaire. Soient $P, Q \in \mathbf{R}_n[X]$ et soit $\lambda \in \mathbf{R}$. Alors

$$f_k(\lambda P + Q) = (\lambda P + Q)^{(k)}(0) = (\lambda P^{(k)} + Q^{(k)})(0) = \lambda P^{(k)}(0) + Q^{(k)}(0) = \lambda f_k(P) + f_k(Q).$$

Donc f est bien une forme linéaire sur $\mathbf{R}_n[X]$ et donc un élément de E^* .

- 2.b. On a $(X^\ell)' = \ell X^{\ell-1}$, $(X^\ell)'' = \ell(\ell-1)X^{\ell-2}$, et une récurrence rapide prouve que pour tout $k \leq \ell$, on a $(X^\ell)^{(k)} = \ell(\ell-1) \cdots (\ell-k+1)X^{\ell-k}$ et pour $k > \ell$, on a $(X^\ell)^{(k)} = 0$.

Notons qu'en particulier, $(X^\ell)^{(\ell)} = \ell!$ est un polynôme constant. On en déduit que

$$f_k(X^\ell) = ((X^\ell)^{(k)})(0) = \begin{cases} 0 & \text{si } k \neq \ell \\ \ell! & \text{si } k = \ell \end{cases}$$

- 2.c. Soient $\lambda_0, \lambda_1, \dots, \lambda_n$ des réels tels que $\sum_{k=0}^n \lambda_k f_k = 0$.

Alors, pour $\ell \in \llbracket 0, n \rrbracket$, en appliquant cette forme linéaire à X^ℓ , il vient

$$\sum_{k=0}^n \lambda_k f_k(X^\ell) = 0 \Leftrightarrow \lambda_\ell \ell! = 0 \Leftrightarrow \lambda_\ell = 0.$$

Ceci étant valable pour tout ℓ , on a donc $\lambda_0 = \lambda_1 = \dots = \lambda_n = 0$.

Et donc la famille (f_0, \dots, f_n) est libre.

- 3.a. On a $\dim E^* = \dim \mathcal{L}(E, \mathbf{R}) = \dim E \times \dim \mathbf{R} = \dim E =$ $n + 1$.

- 3.b. La famille (f_0, \dots, f_n) est une famille libre de E^* , de cardinal $n + 1 = \dim E^*$. C'est donc une base de E^* . Par conséquent, tout élément de E^* s'écrit de manière unique comme combinaison linéaire de (f_0, \dots, f_n) . C'est en particulier le cas de φ : il existe un unique $(n + 1)$ -uplet (a_0, \dots, a_n) de réels tel que $\varphi = \sum_{k=0}^n a_k f_k$. Et donc pour tout $P \in E$,

$$\varphi(P) = \sum_{k=0}^n a_k f_k(P) = \sum_{k=0}^n a_k P^{(k)}(0).$$

- 3.c. Avec les notations de la question précédente, on a, pour tout $i \in \llbracket 0, n \rrbracket$,

$$0 = \varphi(Q_i) = \sum_{k=0}^n a_k Q_i^{(k)}(0).$$

Or, on a $Q_i(X) = Q(X+i)$ et donc $Q_i'(X) = Q'(X+i)$, $Q_i''(X) = Q''(X+i)$ et plus généralement, $Q_i^{(k)}(X) = Q^{(k)}(X+i)$. En particulier, on a $Q_i^{(k)}(0) = Q^{(k)}(i)$.

Les relations précédemment obtenues deviennent donc, pour tout $i \in \llbracket 0, n \rrbracket$,

$$\sum_{k=0}^n a_k Q^{(k)}(i) = 0.$$

Ainsi, le polynôme $\sum_{k=0}^n a_k Q^{(k)}$ possède $n + 1$ racines distinctes². Puisqu'il est de degré au

plus n , c'est donc le polynôme nul : $\sum_{k=0}^n a_k Q^{(k)} = 0$.

Mais à la question 1 nous avons prouvé la liberté de la famille $(Q, Q', \dots, Q^{(n)})$.

On en déduit que les a_i sont tous nuls, et donc pour tout $P \in E$,

$$\varphi(P) = \sum_{k=0}^n \underbrace{a_k}_{=0} P^{(k)}(0) = 0.$$

Par conséquent, φ est bien la forme linéaire nulle.

Rappel

Si E et F sont deux espaces vectoriels de dimension finie, alors $\mathcal{L}(E, F)$ est également de dimension finie et $\dim \mathcal{L}(E, F) = \dim E \times \dim F$.

² Les entiers de 0 à n .

4.a. Puisque (Q_0, \dots, Q_n) est de cardinal $n + 1 = \dim E$, si ce n'est pas une base, alors elle n'est pas génératrice.

En particulier, si on pose $F = \text{Vect}(Q_0, \dots, Q_n)$, alors $F \neq E$ et donc $\dim F \leq n$.

Si $\dim F = n$, alors $H = F$ convient.

Si $\dim F < n$, soit alors (e_1, \dots, e_p) une base de F , avec $p < n$.

D'après le théorème de la base incomplète, il est possible de compléter (e_1, \dots, e_p) en une base (e_1, \dots, e_{n+1}) de E . Et alors en posant $H = \text{Vect}(e_1, \dots, e_n)$, on a $\dim H = n$ car (e_1, \dots, e_n) est libre³, et $F = \text{Vect}(e_1, \dots, e_p) \subset \text{Vect}(e_1, \dots, e_n) = H$.

³ C'est une sous-famille de la famille libre (e_1, \dots, e_{n+1}) .

4.b. Soit G un supplémentaire de H dans E . Alors $\dim G = \dim E - \dim H = n + 1 - n = 1$.
Et donc toute base de G est de cardinal 1. Soit (a) une telle base : on a $G = \text{Vect}(a)$ et donc $E = G \oplus H = \boxed{\text{Vect}(a) \oplus H}$.

4.c. Par définition, ψ_a est à valeurs dans \mathbf{R} , prouvons donc qu'elle est linéaire.

Soient donc $x, y \in E$, et $\lambda \in \mathbf{R}$.

De manière unique, $x = \mu_1 a + h_1$ avec $\mu_1 \in \mathbf{R}$ et $h_1 \in H$ et de même $y = \mu_2 a + h_2$ avec $\mu_2 \in \mathbf{R}$ et $h_2 \in H$.

Alors $\lambda x + y = \underbrace{(\lambda\mu_1 + \mu_2)}_{\in \mathbf{R}} a + \underbrace{(\lambda h_1 + h_2)}_{\in H}$, et cette décomposition est unique⁴, de sorte que

⁴ Car la somme $\text{Vect}(a) + H$ est directe.

$$\psi_a(\lambda x + y) = \lambda\mu_1 + \mu_2 = \lambda\psi_a(x) + \psi_a(y).$$

Et donc ψ_a est linéaire, de E dans \mathbf{R} : c'est une forme linéaire sur E .

Pour $x = \mu a + h \in E$, on a $x \in \text{Ker}(\psi_a) \Leftrightarrow \mu = 0 \Leftrightarrow x = h$. Et donc $x \in H$, de sorte que $\text{Ker} \psi_a \subset H$.

Mais ψ_a étant une forme linéaire non nulle⁵, son noyau est un hyperplan de E et donc de dimension $n = \dim H$.

⁵ Par exemple car $\psi_a(a) = 1 \neq 0$.

Par égalité des dimensions, on en déduit que $\text{Ker} \psi_a = H$.

4.d. Pour tout $i \in \llbracket 0, n \rrbracket$, on a $Q_i = 0 \cdot a + \underbrace{Q_i}_{\in H}$. Et donc par définition de ψ_a , il vient $\psi_a(Q_i) = 0$.

Et donc d'après la question 3.c, nécessairement $\psi_a = 0$.

Or $\psi_a(a) = 1$. On aboutit donc à une contradiction, ce qui signifie que notre hypothèse de départ est fautive. On en déduit que (Q_0, \dots, Q_n) est nécessairement une base de E .

EXERCICE 3 (D'APRÈS EDHEC ECE 2018)

1.a. La fonction $t \mapsto \ln(1 + t^2)$ est positive sur \mathbf{R} , puisque $1 + t^2 \geq 1$.

Et donc si $x \geq 0$, par positivité de l'intégrale, $f(x) = \int_0^x \ln(1 + t^2) dt \geq 0$.

En revanche, si $x \leq 0$, les bornes de l'intégrale sont dans le mauvais sens, on a

$$f(x) = - \int_x^0 f(t) dt \leq 0.$$

Donc $f(x)$ est du signe de x .

1.b. D'après le théorème fondamental de l'analyse, la fonction $t \mapsto \ln(1 + t^2)$, qui est continue sur \mathbf{R} admet une primitive G , qui est donc de classe \mathcal{C}^1 .

Et alors $f(x) = [G(t)]_0^x = G(x) - G(0)$ est une fonction de classe \mathcal{C}^1 sur \mathbf{R} , avec $f'(x) = G'(x) = \boxed{\ln(1 + x^2)}$.

1.c. Puisque pour tout $x \in \mathbf{R}$, $f'(x) \geq 0$, la fonction f est donc croissante sur \mathbf{R} .

2.a. Pour $x \in \mathbf{R}$, on a $f(-x) = \int_0^{-x} \ln(1 + t^2) dt$.

Procédons alors au changement de variable $u = -t$:

$$f'(x) = \int_0^x \ln(1 + u^2)(-du) = - \int_0^x \ln(1 + t^2) dt = -f(x).$$

Et donc f est impaire.

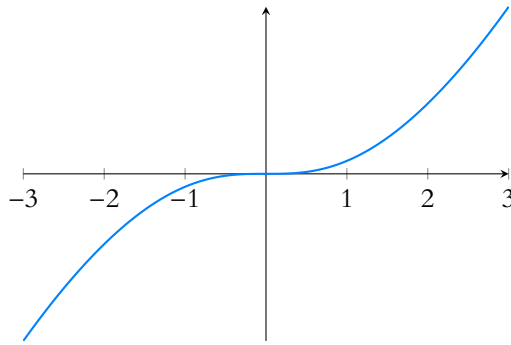
2.b. Puisque \ln est croissante, la fonction $x \mapsto \ln(1 + x^2)$ a même sens de variation que $x \mapsto 1 + x^2$, et donc est croissante sur \mathbf{R}_+ et décroissante sur \mathbf{R}_- .

⚠ Danger !

La positivité de l'intégrale qui dit que si $\forall t, f(t) \geq 0$ alors

$$\int_a^b f(t) dt \geq 0$$

ne s'applique que pour $a \leq b$.

FIGURE 23.1 – La fonction f et son point d'inflexion en $(0, 0)$.

On en déduit que f est convexe sur \mathbf{R}_+ et concave sur \mathbf{R}_- .

Elle possède donc un point d'inflexion en $x = 0$, dont les coordonnées sont $(0, f(0)) = (0, 0)$.

- 3.a. On a $a + \frac{b}{1+t^2} = \frac{a+at^2+b}{1+t^2}$, et donc $a + \frac{b}{1+t^2} = \frac{t^2}{1+t^2}$ pour tout $t \in \mathbf{R}$ si et seulement si pour tout $t \in \mathbf{R}$, $at^2 + a + b = t^2$.

Par identification des coefficients⁶, c'est donc le cas si et seulement si

$$\begin{cases} a = 1 \\ a + b = 0 \end{cases} \Leftrightarrow \boxed{\begin{cases} a = 1 \\ b = -1 \end{cases}}$$

- 3.b. Procédons à une intégration par parties en posant $u(t) = \ln(1+t^2)$ et $v(t) = t$, qui sont deux fonctions de classe \mathcal{C}^1 sur \mathbf{R} avec $u'(t) = \frac{2t}{1+t^2}$ et $v'(t) = 1$. Il vient alors

$$\begin{aligned} f(x) &= \int_0^x \ln(1+t^2) dt = [t \ln(1+t^2)]_0^x - 2 \int_0^x \frac{t^2}{1+t^2} dt \\ &= x \ln(1+x^2) - 2 \int_0^x \left(1 - \frac{1}{1+t^2}\right) dt = x \ln(1+x^2) - 2[t - \text{Arctan}(t)]_0^x \\ &= \boxed{x \ln(1+x^2) - 2x + 2\text{Arctan}(x)}. \end{aligned}$$

- 4.a. Lorsque $x \rightarrow +\infty$, $\text{Arctan}(x) \xrightarrow{x \rightarrow +\infty} \frac{\pi}{2}$, alors que $x(\ln(1+x^2) - 2) \xrightarrow{x \rightarrow +\infty} +\infty$.
Et donc $2\text{Arctan}(x) = \underset{x \rightarrow +\infty}{o}(x \ln(1+x^2) - 2)$.

$$\text{Ainsi } f(x) \underset{x \rightarrow +\infty}{\sim} x(\ln(1+x^2) - 2) \underset{x \rightarrow +\infty}{\sim} \boxed{x \ln(1+x^2)}.$$

- 4.b. On a $\ln(1+x^2) = \ln\left(x^2 \left(1 + \frac{1}{x^2}\right)\right) = \ln(x^2) + \ln\left(1 + \frac{1}{x^2}\right) = \boxed{2 \ln(x) + \ln\left(1 + \frac{1}{x^2}\right)}$.

Puisque $\ln\left(1 + \frac{1}{x^2}\right) \xrightarrow{x \rightarrow +\infty} 0$ et que $\ln(x) \xrightarrow{x \rightarrow +\infty} +\infty$, on en déduit que $\ln(1+x^2) \underset{x \rightarrow +\infty}{\sim} 2 \ln(x)$

$$\text{et donc } \boxed{f(x) \underset{x \rightarrow +\infty}{\sim} 2x \ln(x)}.$$

- 4.c. Puisque f est impaire, pour tout $x \in \mathbf{R}$, $f(x) = -f(-x)$ et donc

$$f(x) \underset{x \rightarrow -\infty}{\sim} -(-x) \ln(-x) \underset{x \rightarrow -\infty}{\sim} \boxed{x \ln(-x)}.$$

- 5.a. La fonction $f' : x \mapsto \ln(1+x^2)$ est de classe \mathcal{C}^2 car composée de fonctions usuelles de classe \mathcal{C}^2 et donc f est de classe \mathcal{C}^3 .

$$\text{On a alors } f''(x) = \frac{2x}{1+x^2} \text{ et } f^{(3)}(x) = 2 \frac{1-x^2}{(1+x^2)^2}.$$

$$\text{Et donc } \boxed{f(0) = 0, f'(0) = 0, f''(0) = 0 \text{ et } f^{(3)}(0) = 2}.$$

Convexité

Rappelons que pour une fonction dérivable, la convexité est donnée par le sens de variation de la dérivée, et qu'il n'est pas nécessaire de dériver deux fois.

⁶ Deux polynômes sont égaux si et seulement si ils ont les mêmes coefficients.

Rappel

On a $f \sim g$ si et seulement si $f = g + o(g)$.

Si on n'est pas à l'aise avec la notation o , mais qu'on a bien compris que le terme «dominant» est celui avec le \ln , on devine que l'équivalent est $x \ln(1+x^2)$.

Pour le prouver rigoureusement, il suffit alors de montrer que la limite de $\frac{f(x)}{x \ln(1+x^2)}$ est égale à 1.

5.b. Par la formule de Taylor-Young au voisinage de 0, on a donc

$$f(x) = f(0) + f'(0)x + f''(0)\frac{x^2}{2!} + f^{(3)}(0)\frac{x^3}{3!} + o_{x \rightarrow 0}(x^3) = \frac{x^3}{3} + o_{x \rightarrow 0}(x^3).$$

Et donc $f(x) \underset{x \rightarrow 0}{\sim} \frac{x^3}{3}.$

Rappel
Une fois qu'on a un développement limité d'une fonction, on en a un équivalent : il suffit de garder le premier terme **non nul** de ce développement limité.

EXERCICE 4 (D'APRÈS ORAL ESCP 2017)

1. On a

$$\begin{aligned} \gamma_{n+1} - \gamma_n &= H_{n+1} - \ln(n+1) - H_n + \ln(n) \\ &= \frac{1}{n+1} + \ln\left(\frac{n}{n+1}\right) \\ &= \frac{1}{n+1} + \ln\left(1 - \frac{1}{n+1}\right) \\ &= \frac{1}{n+1} - \frac{1}{n+1} - \frac{1}{(n+1)^2} + o_{n \rightarrow +\infty}\left(\frac{1}{(n+1)^2}\right) \\ &= -\frac{1}{(n+1)^2} + o_{n \rightarrow +\infty}\left(\frac{1}{(n+1)^2}\right) \underset{n \rightarrow +\infty}{\sim} -\frac{1}{(n+1)^2} \\ &\underset{n \rightarrow +\infty}{\sim} -\frac{1}{n^2}. \end{aligned}$$

Détails
On a utilisé le développement limité
 $\ln(1+u) = u - \frac{u^2}{2} + o_{u \rightarrow 0}(u^2)$
qui est légitime puisque $\frac{1}{n+1} \xrightarrow{n \rightarrow +\infty} 0.$

Mais puisque la série de terme général $\frac{-1}{n^2}$ converge, par critère de comparaison pour les séries de signe constant, il en est de même de $\sum_{n \geq 1} (\gamma_{n+1} - \gamma_n).$

Pour $N \in \mathbf{N}^*$, on a

$$\sum_{n=1}^N (\gamma_{n+1} - \gamma_n) = \cancel{\gamma_2} - \gamma_1 + \cancel{\gamma_3} - \cancel{\gamma_2} + \dots + \cancel{\gamma_N} - \cancel{\gamma_{N-1}} + \gamma_{N+1} - \cancel{\gamma_N} = \gamma_{N+1} - \gamma_1.$$

Puisque la série de terme général $(\gamma_{n+1} - \gamma_n)$ converge, ses sommes partielles admettent une limite lorsque $N \rightarrow +\infty$. Et donc $\lim_{N \rightarrow +\infty} \gamma_{N+1} - \gamma_1$ existe, de sorte que **la suite $(\gamma_N)_{N \geq 1}$ converge.**

Rappel
Par définition, une série converge si et seulement si la suite de ses sommes partielles converge.

2. En séparant les termes pairs des termes impairs, il vient

$$\begin{aligned} \sum_{k=1}^{2n} \frac{(-1)^k}{k} + H_{2n} &= \sum_{k=1}^n \frac{1}{2k} - \sum_{k=1}^{n-1} \frac{1}{2k+1} + \sum_{k=1}^{2n} \frac{1}{k} \\ &= \sum_{k=1}^n \frac{1}{2k} - \sum_{k=1}^{n-1} \frac{1}{2k+1} + \sum_{k=1}^n \frac{1}{2k} + \sum_{k=1}^{n-1} \frac{1}{2k+1} \\ &= 2 \sum_{k=1}^n \frac{1}{2k} = \sum_{k=1}^n \frac{1}{k} = H_n. \end{aligned}$$

Et donc $\sum_{k=1}^{2n} \frac{(-1)^k}{k} = H_n - H_{2n}.$

Lorsque $n \rightarrow +\infty$, on a $H_n = \ln(n) + \gamma + o_{n \rightarrow +\infty}(1)$, de sorte que

$$\sum_{k=1}^{2n} \frac{(-1)^k}{k} = \ln(n) + \gamma + o_{n \rightarrow +\infty}(1) - \ln(2n) - \gamma - o_{n \rightarrow +\infty}(1) = -\ln(2) + o_{n \rightarrow +\infty}(1) \xrightarrow{n \rightarrow +\infty} -\ln(2).$$

D'autre part,

$$\sum_{k=1}^{2n+1} \frac{(-1)^k}{k} = \sum_{k=1}^{2n} \frac{(-1)^k}{k} - \frac{1}{2n+1} \xrightarrow{n \rightarrow +\infty} -\ln(2).$$

$o(1)$
Rappelons que la notation $o(1)$ désigne toute suite de limite nulle.

Ainsi, si on note $S_n = \sum_{k=1}^n \frac{(-1)^k}{k}$ la somme partielle d'ordre n de la série de terme général

$\frac{(-1)^k}{k}$, nous venons de montrer que les deux suites $(S_{2n})_{n \geq 1}$ et $(S_{2n+1})_{n \geq 0}$ convergent vers une même limite égale à $-\ln(2)$, de sorte que $S_n \xrightarrow{n \rightarrow +\infty} -\ln(2)$.

Et donc la série de terme général $\frac{(-1)^k}{k}$ converge et $\sum_{k=1}^{+\infty} \frac{(-1)^k}{k} = -\ln(2)$.

3. On a,

$$\begin{aligned} \sum_{k=1}^{3n+3} \frac{a_k}{k} &= \sum_{k=1}^n \frac{1}{3k+1} + \sum_{k=1}^n \frac{1}{3k+2} - \sum_{k=1}^n \frac{1}{3k+3} \\ &= \sum_{k=1}^n \frac{1}{3k+1} + \sum_{k=1}^n \frac{1}{3k+2} + \sum_{k=1}^n \frac{1}{3k+3} - 2 \sum_{k=1}^n \frac{1}{3k+3} \\ &= H_{3n+3} - \frac{2}{3} \sum_{k=1}^n \frac{1}{k+1} \\ &= H_{3n+3} - \frac{2}{3} H_{n+1}. \end{aligned}$$

Et donc

$$\sum_{k=1}^{3n+3} \frac{a_k}{k} = \ln(3n+3) + \gamma - \frac{2}{3} \ln(n+1) - \frac{2}{3} \gamma - \frac{2}{3} \ln(n+1) - \frac{2}{3} \gamma + \frac{1}{3} \ln(n+1) + \frac{1}{3} \gamma + \frac{1}{3} \ln(n+1) + \frac{1}{3} \gamma \xrightarrow{n \rightarrow +\infty} +\infty.$$

Et donc la série de terme général $\frac{a_k}{k}$ diverge.

4.a. Notons que pour tout $k \in \mathbb{N}$, $\frac{1}{k+1} = \int_0^1 x^k dx$. Et donc

$$\begin{aligned} \sum_{n=0}^N \left(\frac{1}{4n+1} - \frac{1}{4n+3} \right) &= \sum_{n=0}^N \int_0^1 (x^{4n} - x^{4n+2}) dx \\ &= \int_0^1 (1-x^2) \sum_{n=0}^N (x^4)^n dx \\ &= \int_0^1 (1-x^2) \frac{1-(x^4)^{N+1}}{1-x^4} dx \\ &= \int_0^1 \frac{1-x^{4N+4}}{1+x^2} dx. \end{aligned}$$

4.b. Coupons en deux l'intégrale qui précède :

$$\int_0^1 \frac{1-x^{4N+4}}{1+x^2} dx = \int_0^1 \frac{dx}{1+x^2} - \int_0^1 \frac{x^{4N+4}}{1+x^2} dx.$$

La première intégrale vaut $[\text{Arctan}(x)]_0^1 = \frac{\pi}{4}$.

Pour la seconde, on peut noter que pour tout $x \in [0, 1]$, $0 \leq \frac{x^{4N+4}}{1+x^2} \leq x^{4N+4}$ et donc par croissance de l'intégrale,

$$0 \leq \int_0^1 \frac{x^{4N+4}}{1+x^2} dx \leq \int_0^1 x^{4N+4} dx = \frac{1}{4N+5}.$$

Et donc par le théorème des gendarmes, $\lim_{N \rightarrow +\infty} \int_0^1 \frac{x^{4N+4}}{1+x^2} dx = 0$.

Ainsi, $\lim_{N \rightarrow +\infty} \sum_{n=0}^N \left(\frac{1}{4n+1} - \frac{1}{4n+3} \right) dx = \frac{\pi}{4}$.

Ceci prouve donc que $\sum_{n \geq 0} \left(\frac{1}{4n+1} - \frac{1}{4n+3} \right)$ converge et que $\sum_{n=0}^{+\infty} \left(\frac{1}{4n+1} - \frac{1}{4n+3} \right) = \frac{\pi}{4}$.

Remarque

L'identité

$$\sum_{n=0}^N (x^4)^n = \frac{1-x^{4(N+1)}}{1-x^4}$$

n'est pas valable pour $x = 1$, mais elle l'est pour tout $x \in [0, 1[$, ce qui est suffisant pour l'intégrale que l'on considère.

Détails

On a une identité remarquable classique :

$$1-x^4 = (1+x^2)(1-x^2).$$

4.c. Sur le même principe, on prouve que

$$\begin{aligned}
 \sum_{n=0}^N \left(\frac{1}{4n+2} - \frac{1}{4n+4} \right) &= \sum_{n=0}^N \int_0^1 (x^{4n+1} - x^{4n+3}) dx \\
 &= \int_0^1 x(1-x^2) \sum_{n=0}^N (x^4)^n dx \\
 &= \int_0^1 \frac{x - x^{4N+5}}{1+x^2} dx \\
 &= \int_0^1 \frac{x}{1+x^2} dx - \int_0^1 \frac{x^{4N+5}}{1+x^2} dx \\
 &= \left[\frac{1}{2} \ln(1+x^2) \right]_0^1 - \int_0^1 \frac{x^{4N+5}}{1+x^2} dx \\
 &= \frac{1}{2} \ln(2) - \int_0^1 \frac{x^{4N+5}}{1+x^2} dx \\
 &\xrightarrow{n \rightarrow +\infty} \frac{1}{2} \ln(2).
 \end{aligned}$$

Et donc $\boxed{\sum_{n=0}^{+\infty} \left(\frac{1}{4n+2} - \frac{1}{4n+4} \right) = \frac{1}{2} \ln(2).}$

5. Notons $S_n = \sum_{k=1}^n \frac{a_k}{k}$. Alors

$$S_{4N+4} = \sum_{n=1}^N \left(\frac{1}{4n+1} - \frac{1}{4n+3} \right) + \sum_{n=1}^N \left(\frac{1}{4n+2} - \frac{1}{4n+4} \right) \xrightarrow{N \rightarrow +\infty} \frac{\pi}{4} + \frac{1}{2} \ln(2).$$

Toutefois, cela ne suffit pas à prouver que $\sum_{k \geq 1} \frac{a_k}{k}$ converge, nous avons prouvé que la suite

$(S_{4N+4})_{N \geq 0} = (S_{4n})_{n \geq 0}$ converge, mais cela ne suffit pas à garantir que $(S_n)_{n \geq 1}$ converge. On pourra par exemple penser au cas de la suite $u_n = (-1)^n$.

En revanche, pour $n \in \mathbf{N}$, on a $4 \lfloor \frac{n}{4} \rfloor$ est le multiple de 4 immédiatement inférieur à n , de sorte que $|n - 4 \lfloor \frac{n}{4} \rfloor| \leq 3$.

Et alors

$$\left| S_n - S_{4 \lfloor \frac{n}{4} \rfloor} \right| \leq \frac{1}{4 \lfloor \frac{n}{4} \rfloor + 1} + \frac{1}{4 \lfloor \frac{n}{4} \rfloor + 2} + \frac{1}{4 \lfloor \frac{n}{4} \rfloor + 4} \xrightarrow{n \rightarrow +\infty} 0.$$

Mais comme mentionné plus haut, $S_{4 \lfloor \frac{n}{4} \rfloor} \xrightarrow{n \rightarrow +\infty} \frac{\pi}{4} + \frac{1}{2} \ln(2)$.

On en déduit que $S_n \xrightarrow{n \rightarrow +\infty} \frac{\pi}{4} + \frac{1}{2} \ln(2)$, de sorte que

$$\boxed{\sum_{n=0}^{+\infty} \frac{a_n}{n} = \frac{\pi}{4} + \frac{1}{2} \ln(2).}$$

Remarque : on aurait également pu distinguer 4 cas, et montrer que les suites (S_{4n}) , (S_{4n+1}) , (S_{4n+2}) et (S_{4n+3}) convergent toutes vers une même limite⁷, de sorte que (S_n) converge également vers cette même limite.

⁷ Qui est $\frac{\pi}{4} + \frac{1}{2} \ln(2)$.

Autrement dit

La somme définissant S_n contient au maximum trois termes de plus que $S_{\lfloor \frac{n}{4} \rfloor}$. Nous allons utiliser dans la suite le fait que ces trois termes tendent vers 0 lorsque $n \rightarrow +\infty$.

EXERCICE 5 (EDHEC 2005)

1.a. Notons A_i l'événement «le joueur obtient la face 1 lors du $i^{\text{ème}}$ lancer».

Alors l'événement «le joueur obtient n fois face lors des n premiers lancers» est l'événement

$$A_1 \cap \dots \cap A_n = \bigcap_{i=1}^n A_i.$$

De plus, nous savons que pour tout $i \in \mathbf{N}^*$, $P_E(A_i) = \frac{1}{2}$ et $P_{\bar{E}}(A_i) = 1$.

Enfin, une fois le jeton choisi, le résultat d'un lancer n'influe pas sur le suivant, et donc les événements $(A_i)_{i \in \mathbf{N}^*}$ sont indépendants pour la probabilité P_E et pour la probabilité $P_{\bar{E}}$. Ainsi, il vient

$$P_E\left(\bigcap_{i=1}^n A_i\right) = \prod_{i=1}^n P_E(A_i) = \frac{1}{2^n}.$$

Et de même,

$$P_{\bar{E}}\left(\bigcap_{i=1}^n A_i\right) = \prod_{i=1}^n P_{\bar{E}}(A_i) = 1^n = 1.$$

Et donc, par la formule des probabilités totales appliquée au système complet d'événements $\{E, \bar{E}\}$, on a

$$P\left(\bigcap_{i=1}^n A_i\right) = P_E\left(\bigcap_{i=1}^n A_i\right)P(E) + P_{\bar{E}}\left(\bigcap_{i=1}^n A_i\right)P(\bar{E}) = \frac{1}{2} \frac{1}{2^n} + \frac{1}{2} = \boxed{\frac{2^n + 1}{2^{n+1}}}.$$

- 1.b. Nous cherchons ici à déterminer $P_{A_1 \cap A_2 \cap \dots \cap A_n}(E)$. D'après la formule de Bayes, on a

$$P_{A_1 \cap A_2 \cap \dots \cap A_n}(E) = \frac{P_E\left(\bigcap_{i=1}^n A_i\right)P(E)}{P\left(\bigcap_{i=1}^n A_i\right)} = \frac{\frac{1}{2^n} \frac{1}{2}}{\frac{2^n + 1}{2^{n+1}}} = \boxed{\frac{1}{1 + 2^n}}.$$

Cette probabilité tend vers 0 lorsque n tend vers $+\infty$.

Et donc cela signifie que plus on obtient de 1 à la suite lors des premiers lancers, plus il est probable que le jeton utilisé soit J_2 .

- 2.a. Soit $n \in \mathbf{N}^*$. Si on joue avec le jeton J_2 , alors nous n'obtiendrons jamais 0, et donc $P_{\bar{E}}(X = n) = 0$.

En revanche, si on joue avec le jeton J_1 , alors le nombre de lancers nécessaires à l'obtention du premier 0 suit une loi géométrique de paramètre $\frac{1}{2}$ et donc $P_E(X = n) = \frac{1}{2^n}$.

Par la formule des probabilités totales appliquée au système complet d'événements $\{E, \bar{E}\}$, on a donc

$$P(X = n) = P_E(X = n)P(E) + P_{\bar{E}}(X = n)P(\bar{E}) = \frac{1}{2} \times 0 + \frac{1}{2} \frac{1}{2^n} = \boxed{\frac{1}{2^{n+1}}}.$$

- 2.b. Puisque $\{[X = n], n \in \mathbf{N}\}$ est un système complet d'événements⁸, on doit avoir

$$\sum_{n=0}^{+\infty} P(X = n) = 1 \text{ et donc } P(X = 0) = 1 - \sum_{n=1}^{+\infty} P(X = n). \text{ Or}$$

$$\sum_{n=1}^{+\infty} P(X = n) = \sum_{n=1}^{+\infty} \frac{1}{2^{n+1}} = \frac{1}{4} \sum_{n=1}^{+\infty} \frac{1}{2^{n-1}} = \frac{1}{4} \sum_{i=0}^{+\infty} \frac{1}{2^i} = \frac{1}{4} \frac{1}{1 - \frac{1}{2}} = \frac{1}{2}.$$

Et donc $\boxed{P(X = 0) = \frac{1}{2}}$.

Ce résultat était prévisible : si on joue avec J_2 , alors $X = 0$ (on n'aura jamais de face 0), alors que si l'on joue avec J_1 , on finira presque sûrement par avoir une face 0.

- 2.c. La variable aléatoire X possède une espérance si et seulement si la série de terme général $nP(X = n)$ converge absolument. Puisqu'il s'agit d'une série à termes positifs, elle converge absolument si et seulement si elle converge.

Puisque $nP(X = n) = n \frac{1}{2^{n+1}} = \frac{1}{4} n \frac{1}{2^{n-1}}$, nous reconnaissons le terme général d'une série géométrique dérivée convergente⁹ et donc $E(X)$ existe et vaut

$$E(X) = \sum_{n=0}^{+\infty} nP(X = n) = \frac{1}{4} \sum_{n=1}^{+\infty} n \frac{1}{2^{n-1}} = \frac{1}{4} \frac{1}{\left(1 - \frac{1}{2}\right)^2} = \boxed{1}.$$

Indépendance

Contrairement à ce que l'on pourrait penser, les événements A_i ne sont pas indépendants, ils le sont seulement **une fois le jeton choisi**, c'est-à-dire pour P_E et $P_{\bar{E}}$.

Par exemple, si on a obtenu 1 au premier lancer, il est plus probable qu'on soit en train d'utiliser J_2 que J_1 , et donc il est plus probable d'obtenir de nouveau un 1 au second lancer.

En revanche, si on sait quel jeton nous sommes en train d'utiliser (c'est-à-dire si l'on sait que E ou \bar{E} est réalisé), alors les lancers sont bien indépendants.

Méthode

On cherche une probabilité conditionnelle $P_A(B)$ alors que l'on connaît $P_B(A)$: c'est typiquement la situation où il faut penser à la formule de Bayes.

⁸ Par définition de X : X ne peut prendre que la valeur 0 ou une valeur strictement positive.

Détails

On a reconnu une série géométrique de raison $\frac{1}{2}$.

⁹ Car de raison $\frac{1}{2} \in]-1, 1[$.

2.d. D'après le théorème de transfert¹⁰, la variable aléatoire $X(X-1)$ admet une espérance si et seulement si la série de terme général $n(n-1)P(X=n)$ converge¹¹.

Mais $n(n-1)P(X=n) = n(n-1)\frac{1}{2^{n+1}} = \frac{1}{8}n(n-1)\frac{1}{2^{n-2}}$, et on reconnaît alors le terme général d'une série géométrique dérivée seconde convergente¹².

Donc $X(X-1)$ admet une espérance et

$$E(X(X-1)) = \sum_{n=0}^{+\infty} n(n-1)P(X=n) = \frac{1}{8} \sum_{n=2}^{+\infty} n(n-1)\frac{1}{2^{n-2}} = \frac{1}{8} \frac{2}{(1-\frac{1}{2})^3} = \boxed{2.}$$

On en déduit que $E(X^2) = E(X(X-1)) + E(X) = 2 + 1 = 3$ et alors, par la formule de Huygens,

$$V(X) = E(X^2) - E(X)^2 = 3 - 1 = \boxed{2.}$$

3.a. Par le même raisonnement qu'à la question 2.a, on a $P_E(Y=n) = \frac{1}{2^n}$.

En revanche, si l'on joue avec le jeton J_2 , la première face 1 apparaît nécessairement lors du premier lancer et donc $P_{\bar{E}}(Y=1) = 1$ et $P_{\bar{E}}(Y=n) = 0$ si $n \geq 2$.

Par la formule des probabilités totales¹³, on en déduit donc que

$$P(Y=1) = P_E(Y=1)P(E) + P_{\bar{E}}(Y=1)P(\bar{E}) = \frac{1}{2} \frac{1}{2} + \frac{1}{2} = \frac{3}{4}$$

et pour $n \geq 2$,

$$P(Y=n) = P_E(Y=n)P(E) + \underbrace{P_{\bar{E}}(Y=n)}_{=0} P(\bar{E}) = \boxed{\frac{1}{2^{n+1}}}.$$

3.b. Comme à la question 2.b, on doit avoir $\sum_{n=0}^{+\infty} P(Y=n) = 1$ soit

$$P(Y=0) = 1 - \sum_{n=1}^{+\infty} P(Y=n) = 1 - \frac{3}{4} - \sum_{n=2}^{+\infty} \frac{1}{2^{n+1}} = \frac{1}{4} - \sum_{k=0}^{+\infty} \frac{1}{2^{k+3}} = \frac{1}{4} - \frac{1}{8} \sum_{k=0}^{+\infty} \frac{1}{2^k} = \frac{1}{4} - \frac{1}{8} \frac{1}{1-\frac{1}{2}} = 0.$$

Chgt d'indice
 $k = n - 2.$

3.c. La variable aléatoire Y admet une espérance si et seulement si la série de terme général $nP(Y=n)$ converge¹⁴.

Or, pour $n \geq 2$, on a $nP(Y=n) = \frac{1}{2}n\frac{1}{2^n}$, qui est le terme général d'une série géométrique dérivée convergente. Et donc

$$\begin{aligned} E(Y) &= \sum_{n=0}^{+\infty} nP(Y=n) = \frac{3}{4} + \sum_{n=2}^{+\infty} nP(Y=n) = \frac{3}{4} + \sum_{n=2}^{+\infty} nP(Y=n) \\ &= \frac{3}{4} + \frac{1}{4} \sum_{n=2}^{+\infty} n\frac{1}{2^{n-1}} = \frac{3}{4} + \frac{1}{4} \left(\sum_{n=1}^{+\infty} n\frac{1}{2^{n-1}} - 1 \right) = \frac{3}{4} + \frac{1}{4} \left(\frac{1}{(1-\frac{1}{2})^2} - 1 \right) = \frac{3}{4} + 1 - \frac{1}{4} = \boxed{\frac{3}{2}}. \end{aligned}$$

14 Absolument, mais il s'agit d'une série à termes positifs.

Convergence
Les premiers termes d'une série n'ont pas d'influence sur sa nature. Donc pour étudier la convergence, il suffit de s'intéresser à n «grand» (ici $n \geq 2$).

3.d. D'après le théorème de transfert, la variable aléatoire $Y(Y-1)$ admet une espérance si et seulement si la série de terme général $n(n-1)P(Y=n)$ converge.

Or, pour tout $n \in \mathbf{N}$, on a $n(n-1)P(Y=n) = n(n-1)\frac{1}{2^{n+1}} = \frac{1}{8}n(n-1)\frac{1}{2^{n-2}}$ qui est le terme général d'une série convergente.

Et donc $Y(Y-1)$ admet une espérance, égale à

$$E(Y(Y-1)) = \sum_{n=2}^{+\infty} n(n-1)P(Y=n) = \frac{1}{8} \sum_{n=2}^{+\infty} n(n-1)\frac{1}{2^{n-2}} = \frac{1}{8} \frac{2}{(1-\frac{1}{2})^3} = \boxed{2.}$$

On en déduit que $E(Y^2) = E(Y(Y-1)) + E(Y) = 2 + \frac{3}{2} = \frac{7}{2}$.

Et alors, par la formule de Huygens,

$$V(Y) = E(Y^2) - E(Y)^2 = \frac{7}{2} - \frac{9}{4} = \boxed{\frac{5}{4}}.$$

4.a. X et Y prennent leurs valeurs dans \mathbf{N} , donc S prend également ses valeurs dans \mathbf{N} . Nous pouvons même affirmer que S ne prend pas la valeur 0, car cela signifierait que X et Y prennent simultanément la valeur 0, et donc que le joueur n'a obtenu aucun 0 et aucun 1 lors de ses lancers. Et donc $S(\Omega) = \mathbf{N}^*$.

4.b. Notons qu'on ne peut avoir simultanément $X = 1$ et $Y = 1$, car cela signifierait que le premier lancer a donné un 1 et un 0. On a donc $[S = 1] = ([X = 0] \cap [Y = 1]) \cup ([X = 1] \cap [Y = 0])$ et ces deux événements sont incompatibles, de sorte que

$$P(S = 1) = P([X = 0] \cap [Y = 1]) + P([X = 1] \cap [Y = 0]).$$

Mais $P([X = 1] \cap [Y = 0]) \leq P(Y = 0) = 0$.

D'autre part, si $[X = 0]$, alors le joueur n'a obtenu que des faces numérotées 1. C'est en particulier le cas au premier lancer, et donc $[X = 0] \subset [Y = 1]$.

Alors $[X = 0] \cap [Y = 1] = [X = 0]$, de sorte que $P([X = 0] \cap [Y = 1]) = P(X = 0)$.

Nous avons donc bien $P(S = 1) = P(X = 0) = \frac{1}{2}$.

4.c. Si $X = n$, avec $n \geq 2$, alors le premier 0 n'est pas apparu avant le $n^{\text{ème}}$ lancer, de sorte que les lancers 1 à $n - 1$ ont nécessairement produit des 1 : $[Y < n]$ est donc réalisé.

Et donc $[X = n] \subset [Y < n]$.

De même, on $[Y = n] \subset [X < n]$.

Enfin, puisque X et Y ne peuvent prendre simultanément la même valeur, on a

$$[S = n] = ([X = n] \cap [Y < n]) \cup ([Y = n] \cap [X < n]) = [X = n] \cup [Y = n].$$

4.d. Les deux événements $[X = n]$ et $[Y = n]$ sont incompatibles, donc

$$P(S = n) = P(X = n) + P(Y = n) = \frac{1}{2^{n+1}} + \frac{1}{2^{n+1}} = \frac{1}{2^n}.$$

On a donc, pour tout $n \in \mathbf{N}^*$, $P(S = n) = \frac{1}{2^n} = \frac{1}{2} \left(1 - \frac{1}{2}\right)^{n-1}$.

Nous reconnaissons là une loi géométrique de paramètre $\frac{1}{2}$: $S \hookrightarrow \mathcal{G}\left(\frac{1}{2}\right)$.

On en déduit donc que $E(S) = \frac{1}{\frac{1}{2}} = 2$ et $V(S) = \frac{1 - \frac{1}{2}}{\frac{1}{2}^2} = 2$.

5.a. A priori, I prend ses valeurs dans \mathbf{N} . Mais le premier lancer donne toujours un 0 ou un 1, de sorte que soit $X = 1$, soit $Y = 1$. Et donc on a toujours $\min(X, Y) \leq 1$.

Ainsi, I ne peut prendre que les valeurs 0 ou 1 : c'est une variable de Bernoulli.

5.b. On a $[I = 0] = [X = 0] \cup [Y = 0]$.

Ces deux événements sont incompatibles comme expliqué à la question 4.a.

Et donc $P(I = 0) = P(X = 0) + P(Y = 0) = \frac{1}{2}$.

On en déduit que $I \hookrightarrow \mathcal{B}\left(\frac{1}{2}\right)$ et donc que $E(I) = \frac{1}{2}$ et $V(I) = \frac{1}{2} \left(1 - \frac{1}{2}\right) = \frac{1}{4}$.

EXERCICE 6 (D'APRÈS ORAL ESCP 2015)

1.a. Il est évident que I_n suit la loi uniforme sur $\llbracket 1, n \rrbracket$.

1.b. Si $I_n = 1$, alors il n'y a qu'un tirage, et le processus s'arrête.

Donc X_n vaut 1, de sorte que $P_{[I_n=1]}(X_n = k) = \begin{cases} 1 & \text{si } k = 1 \\ 0 & \text{sinon} \end{cases}$

1.c. Si $[I_n = k]$, alors le second tirage est réalisé dans une urne contenant des boules numérotées de 1 à $k - 1$.

Et donc la probabilité qu'on effectue en tout j tirages¹⁵ est celle qu'on effectue $j - 1$ tirages après le premier, c'est-à-dire en partant d'une urne contenant des boules numérotées de 1 à $k - 1$. Et donc

$$P_{[I_n=k]}(X_n = j) = P(X_{k-1} = j - 1).$$

Rappel

On a une inclusion entre événement $A \subset B$ si dès que A est réalisé, alors B l'est aussi.

Mieux

Nous pouvons même affirmer que $Y = 1$ est réalisé.

Max

Le maximum de X et Y vaut n si l'une des deux variables vaut n et si l'autre est inférieure ou égale à n . Mais X et Y ne peuvent être égale, donc il faut et il suffit que l'une vaille n et que l'autre soit strictement inférieure à n .

Autrement dit

Sachant que $I_n = 1$, X_n est une variable aléatoire certaine égale à 1.

¹⁵ En comptant le premier.

2.a. Si l'urne contient une seule boule au départ, il n'y aura nécessairement qu'un seul tirage, et donc X_1 est la variable certaine égale à 1.

2.b. Notons qu'il ne peut y avoir que deux tirages au maximum, car si on obtient la boule 2 lors du premier tirage, on aura nécessairement la boule 1 lors du second.

Or, $[X_2 = 1] = [I_2 = 1]$, de sorte que $P(X_2 = 1) = \frac{1}{2}$.

Et donc $P(X_2 = 2) = 1 - P(X_2 = 1) = \frac{1}{2}$.

Autrement dit, X_2 suit la loi uniforme sur $\llbracket 1, 2 \rrbracket$, et donc $E(X_2) = \frac{3}{2}$.

3.a. Puisque chaque tirage est effectué dans une urne contenant strictement moins de boules que lors du tirage précédent, il faudra au maximum n tirages pour avoir la boule numéro 1. Et donc $X_n(\Omega) = \llbracket 1, n \rrbracket$.

3.b. On a $P(X_n = 1) = P(I_n = 1) = \frac{1}{n}$.

Et $[X_n = n] = [Y_1 = n] \cap [Y_2 = n-1] \cap \dots \cap [Y_n = 1]$.
Par la formule des probabilités composées, on a donc

$$\begin{aligned} P(X_n = n) &= P(Y_1 = n)P_{[Y_1=n]}(Y_2 = n-1) \times \dots \times P_{[Y_1=1] \cap \dots \cap [Y_{n-1}=2]}(Y_n = 1) \\ &= \frac{1}{n} \frac{1}{n-1} \dots \frac{1}{1} \\ &= \frac{1}{n!}. \end{aligned}$$

3.c. Utilisons la formule des probabilités totales avec le système complet d'événements $\{[I_n = k], k \in \llbracket 1, n \rrbracket\}$. Alors pour tout $j \geq 2$,

$$\begin{aligned} P(X_n = j) &= \sum_{k=1}^n P(I_n = k)P_{[I_n=k]}(X_n = j) \\ &= \sum_{k=2}^n P(I_n = k)P_{[I_n=k]}(X_n = j) \\ &= \sum_{k=2}^n \frac{1}{n} P_{[I_n=k]}(X_n = j) \\ &= \sum_{k=2}^n \frac{1}{n} P(X_{k-1} = j-1) \\ &= \frac{1}{n} \sum_{\ell=1}^{n-1} P(X_\ell = j-1). \end{aligned}$$

Puisque $j \geq 2$, on ne peut avoir tiré la boule 1 au premier tirage.

C'est la question 1.c.

Chgt d'indice
 $\ell = k - 1$.

3.d. En utilisant la relation de la question précédente, on a, pour tout $j \geq 2$,

$$nP(X_n = j) - (n-1)P(X_{n-1} = j) = \sum_{k=1}^{n-1} P(X_k = j-1) - \sum_{k=1}^{n-2} P(X_k = j-1) = P(X_{n-1} = j-1).$$

Et donc, toujours pour $j \geq 2$,

$$P(X_n = j) = \frac{1}{n} (nP(X_n = j) - (n-1)P(X_{n-1} = j)) + \frac{n-1}{n} P(X_{n-1} = j) = \frac{n-1}{n} P(X_{n-1} = j) + \frac{1}{n} P(X_{n-1} = j-1).$$

Enfin, pour $j = 1$, il vient

$$P(X_n = 1) = \frac{1}{n} = \frac{n-1}{n} \frac{1}{n-1} = \frac{n-1}{n} P(X_{n-1} = 1) = \frac{n-1}{n} P(X_{n-1} = 1) + \frac{1}{n} \underbrace{P(X_{n-1} = 0)}_{=0}.$$

4. On a, par définition de l'espérance,

$$\begin{aligned}
 E(X_n) &= \sum_{j=1}^n jP(X_n = j) \\
 &= \sum_{j=1}^n \frac{n-1}{n} jP(X_{n-1} = j) + \sum_{j=1}^n \frac{1}{n} jP(X_{n-1} = j-1) \\
 &= \frac{n-1}{n} E(X_{n-1}) + \frac{1}{n} \sum_{k=2}^n jP(X_{n-1} = j-1) \\
 &= \frac{n-1}{n} E(X_{n-1}) + \frac{1}{n} \sum_{i=1}^{n-1} (i+1)P(X_{n-1} = i) \\
 &= \frac{n-1}{n} E(X_{n-1}) + \frac{1}{n} \underbrace{\sum_{i=1}^{n-1} iP(X_{n-1} = i)}_{=E(X_{n-1})} + \frac{1}{n} \underbrace{\sum_{i=1}^{n-1} P(X_{n-1} = i)}_{=1} \\
 &= \boxed{E(X_{n-1}) + \frac{1}{n}}.
 \end{aligned}$$

Le terme en $j = 1$ est nul.

Remarque

Puisque X_n est une variable finie, elle admet automatiquement une espérance.

En utilisant plusieurs fois cette relation, on obtient alors

$$E(X_n) = E(X_{n-1}) + \frac{1}{n} = E(X_{n-2}) + \frac{1}{n-1} + \frac{1}{n} = \dots = E(X_1) + \frac{1}{2} + \frac{1}{3} + \dots + \frac{1}{n-1} + \frac{1}{n}.$$

Mais X_1 étant la variable certaine égale à 1, on a $E(X_1) = 1$ et donc

$$\boxed{E(X_n) = \sum_{k=1}^n \frac{1}{k}}.$$

DEVOIR MAISON 1 (À RENDRE LE 13.09.18)

Un mobile se déplace aléatoirement sur un axe dont l'origine est le point O d'abscisse 0.

Au départ (instant 0), le mobile est situé sur le point O .

Le mobile se déplace selon la règle suivante : à l'instant n ($n \in \mathbf{N}^*$), il se place de façon équiprobable, sur l'un des points d'abscisses $0, 1, \dots, n$.

Pour tout entier naturel n , on note X_n l'abscisse de ce point à l'instant n (on a donc $X_0 = 0$).

On admet que, pour tout entier naturel n , X_n est une variable aléatoire définie sur un espace probabilisé (Ω, \mathcal{A}, P) que l'on ne cherchera pas à déterminer.

On suppose que les différents déplacements sont indépendants les uns des autres (pour les cubes : on suppose que les variables aléatoires $(X_n)_{n \in \mathbf{N}}$ sont mutuellement indépendantes).

1.
 - a. Déterminer, pour tout entier naturel n non nul, la loi de X_n .
 - b. En déduire que, pour tout entier naturel n non nul, X_n possède une espérance et une variance, puis déterminer $E(X_n)$ et $V(X_n)$.
2. On note Y le rang du premier retour à l'origine du mobile et on admet que Y est une variable aléatoire définie, elle aussi, sur (Ω, \mathcal{A}, P) .
 - a. Pour tout entier naturel n non nul, exprimer l'événement $[Y = n]$ à l'aide des variables aléatoires X_1, X_2, \dots, X_n .
 - b. En déduire que la loi de Y est définie par : $\forall n \in \mathbf{N}^*, P(Y = n) = \frac{1}{n(n+1)}$.
 - c. En remarquant que $\frac{1}{n(n+1)} = \frac{1}{n} - \frac{1}{n+1}$, vérifier par le calcul que l'on a : $\sum_{n=1}^{+\infty} P(Y = n) = 1$.
 - d. La variable Y admet-elle une espérance ?

3.
 - a. Montrer que, pour tout entier naturel k non nul, on a : $\frac{1}{k+1} \leq \ln(k+1) - \ln(k) \leq \frac{1}{k}$.

b. En déduire que : $\forall j \geq 2, \ln j \leq \sum_{k=1}^{j-1} \frac{1}{k} \leq \ln j + 1 - \frac{1}{j}$.

c. Conclure alors que : $\sum_{k=1}^{j-1} \frac{1}{k} \underset{j \rightarrow +\infty}{\sim} \ln j$.

4. On note Z le rang du deuxième retour à l'origine du mobile et on admet que Z est une variable aléatoire définie, elle aussi, sur (Ω, \mathcal{A}, P) .
 - a. Déterminer pour tout $i \geq j$, la probabilité $P_{[Y=i]}(Z = j)$.
 - b. Établir que :

$$\forall i \leq j-1, P_{[Y=i]}(Z = j) = \frac{i+1}{j(j+1)}.$$

- c. Écrire, pour tout entier naturel j supérieur ou égal à 2, la probabilité $P(Z = j)$ comme une somme finie.
- d. La variable aléatoire Z possède-t-elle une espérance ?

5. Informatique

On rappelle qu'en Scilab, l'instruction `grand(1,1,'uin',a,b)` permet de simuler une variable aléatoire suivant la loi uniforme à valeurs dans $[[a, b]]$.

- a. Écrire des commandes Scilab calculant et affichant la valeur de l'abscisse du mobile après son $n^{\text{ème}}$ déplacement lorsque la valeur de n est entrée au clavier par l'utilisateur.
- b. Compléter le script Scilab suivant pour qu'il permette d'afficher dans cet ordre les valeurs prises par les variables aléatoires Y et Z .

```

1 n = 0
2 a = 0
3 while a<2
4     n = n+1
5     if grand(1,1,'uin',0,n) == 0 then
6         a = a+1
7         if a ==1 then y=n, end
8     end

```

```
9 end
10 disp(---, 'y=')
11 disp(---, 'z=')
```

CORRECTION DU DEVOIR MAISON 1

EDHEC 2018

- 1.a. La variable X_n prend de façon équiprobable les valeurs $0, 1, \dots, n$: elle suit donc une loi uniforme sur $[[0, n]]$.
- 1.b. Il s'agit d'une loi usuelle dont l'espérance et la variance sont connues :

$$E(X) = \frac{0+n}{2} = \frac{n}{2} \text{ et } V(X) = \frac{(n+1)^2 - 1}{12}.$$

- 2.a. L'événement $[Y = n]$ est réalisé si le premier retour à l'origine a lieu à l'instant n , et donc

$$[Y = n] = [X_1 \neq 0] \cap \dots \cap [X_{n-1} \neq 0] \cap [X_n = 0] = \left(\bigcap_{i=1}^{n-1} [X_i \neq 0] \right) \cap [X_n = 0].$$

- 2.b. Notons que Y est à valeurs dans \mathbf{N}^* .
Par indépendance des déplacements successifs, il vient donc, pour $n \in \mathbf{N}^*$,

$$\begin{aligned} P(Y = n) &= \prod_{i=1}^{n-1} P(X_i \neq 0) \times P(X_n = 0) \\ &= \prod_{i=1}^{n-1} (1 - P(X_i = 0)) \times P(X_n = 0) \\ &= \prod_{i=1}^{n-1} \left(1 - \frac{1}{i+1}\right) \times \frac{1}{n+1} \\ &= \prod_{i=1}^{n-1} \frac{i}{i+1} \times \frac{1}{n+1} = \frac{1}{2} \times \frac{2}{3} \times \dots \times \frac{n-1}{n} \times \frac{1}{n+1} = \frac{1}{n(n+1)}. \end{aligned}$$

- 2.c. Pour tout $n \in \mathbf{N}^*$, on a bien $\frac{1}{n} - \frac{1}{n+1} = \frac{n+1-n}{n(n+1)} = \frac{1}{n(n+1)}$ et donc pour $N \in \mathbf{N}^*$,

$$\begin{aligned} \sum_{n=1}^N P(Y = n) &= \sum_{n=1}^N \left(\frac{1}{n} - \frac{1}{n+1} \right) = \sum_{n=1}^N \frac{1}{n} - \sum_{n=1}^N \frac{1}{n+1} \\ &= \sum_{n=1}^N \frac{1}{n} - \sum_{k=2}^{N+1} \frac{1}{k} = 1 - \frac{1}{N+1} \\ &\xrightarrow{N \rightarrow +\infty} 1. \end{aligned}$$

Et donc on a bien $\sum_{n=1}^{+\infty} P(Y = n) = 1$.

- 2.d. On a $nP(Y = n) = \frac{1}{n+1} \underset{n \rightarrow +\infty}{\sim} \frac{1}{n}$.
Et donc par comparaison de séries à termes positifs, la série de terme général $nP(Y = n)$ diverge : Y n'admet pas d'espérance.

- 3.a. Soit $k \geq 1$ fixé. La fonction $f : t \mapsto \ln(t)$ est dérivable sur $[k, k+1]$, de dérivée $f' : t \mapsto \frac{1}{t}$.

En particulier, pour tout $t \in [k, k+1]$, $\frac{1}{k+1} \leq f'(t) \leq \frac{1}{k}$.

Par l'inégalité des accroissements finis appliquée à f entre k et $k+1$, on a donc

$$\frac{1}{k+1} = \frac{1}{k+1}(k+1-k) \leq \ln(k+1) - \ln(k) \leq \frac{1}{k}(k+1-k) = \frac{1}{k}.$$

Méthode

Les sommes et produits «télescopiques» peuvent aussi bien se traiter à l'aide de pointillés comme nous l'avons fait que de manière plus rigoureuse, à l'aide de changements d'indices. Les deux rédactions sont convenables, mais si on utilise des pointillés, il faut bien s'assurer de n'avoir pas oublié de termes.

Astuce

Nous avons ici donné une indication, mais elle figure rarement dans les énoncés, et il est bon de connaître cette formule.

Chgt d'indice

On a posé $k = n+1$, et changé les bornes en conséquence :

$$1 \leq n \leq N \Leftrightarrow 2 \leq k \leq N+1.$$

Remarque

Il n'est a priori pas évident que le mobile finisse toujours par revenir à l'origine. Le fait que la somme que nous venons de calculer vaille 1 signifie que le mobile revient presque sûrement (c'est-à-dire avec probabilité 1) à l'origine. Et donc la probabilité qu'il n'y revienne jamais est nulle (et donc Y est bien définie).

3.b. Sommons les relations précédemment obtenues pour k variant de 1 à $j-1$:

$$\sum_{k=1}^{j-1} \frac{1}{k+1} \leq \sum_{k=1}^{j-1} (\ln(k+1) - \ln(k)) \leq \sum_{k=1}^{j-1} \frac{1}{k}.$$

Un changement d'indice facile dans la première somme nous donne $\sum_{k=1}^{j-1} \frac{1}{k+1} = \sum_{k=2}^j \frac{1}{k}$.

La seconde somme est une somme télescopique :

$$\sum_{k=1}^{j-1} (\ln(k+1) - \ln(k)) = (\ln(2) - \ln(1)) + (\ln(3) - \ln(2)) + \dots + (\ln(j-1) - \ln(j-2)) + (\ln(j) - \ln(j-1)) = \ln(j).$$

Et donc $\sum_{k=2}^j \frac{1}{k} \leq \ln(j) \leq \sum_{k=1}^{j-1} \frac{1}{k}$, ce qui prouve déjà l'une des deux inégalités demandées.

D'autre part, on a

$$\sum_{k=2}^j \frac{1}{k} \leq \ln(j) \Leftrightarrow \sum_{k=1}^{j-1} \frac{1}{k} - 1 + \frac{1}{j} \leq \ln(j) \Leftrightarrow \sum_{k=1}^{j-1} \frac{1}{k} \leq \ln(j) + 1 - \frac{1}{j}.$$

3.c. Divisons les inégalités de la question précédente par $\ln(j)$, pour $j \geq 3$:

$$1 \leq \frac{1}{\ln j} \sum_{k=1}^{j-1} \frac{1}{k} \leq 1 + \frac{1}{\ln j} - \frac{1}{j \ln(j)}.$$

Alors, par le théorème des gendarmes,

$$\lim_{j \rightarrow +\infty} \frac{1}{\ln j} \sum_{k=1}^{j-1} \frac{1}{k} = 1 \Leftrightarrow \sum_{k=1}^{j-1} \frac{1}{k} \underset{j \rightarrow +\infty}{\sim} \ln(j).$$

4.a. Par définition, le second retour à l'origine doit avoir lieu après le premier, et donc pour $i \geq j$, $P_{[Y=i]}(Z=j) = 0$.

4.b. Pour $i \leq j-1$, par définition d'une probabilité conditionnelle,

$$P_{[Y=i]}(Z=j) = \frac{P([Y=i] \cap [Z=j])}{P(Y=i)}.$$

Or nous avons déjà prouvé que à la question 2.a que $[Y=i] = \left(\bigcap_{k=1}^{i-1} [X_k \neq 0] \right) \cap [X_i = 0]$.

Et sur le même principe,

$$[Y=i] \cap [Z=j] = \left(\bigcap_{k=1}^{i-1} [X_k \neq 0] \right) \cap [X_i = 0] \cap \left(\bigcap_{\ell=i+1}^{j-1} [X_\ell \neq 0] \right) \cap [X_j = 0].$$

Par indépendance des déplacements successifs, il vient donc

$$\begin{aligned} P_{[Y=i]}(Z=j) &= \frac{\left(\prod_{k=1}^{i-1} P(X_k \neq 0) \right) P(X_i = 0) \left(\prod_{\ell=i+1}^{j-1} P(X_\ell \neq 0) \right) P(X_j = 0)}{\left(\prod_{k=1}^{i-1} P(X_k \neq 0) \right) P(X_i = 0)} \\ &= \left(\prod_{\ell=i+1}^{j-1} P(X_\ell \neq 0) \right) P(X_j = 0). \end{aligned}$$

Soit encore

$$P_{[Y=i]}(Z=j) = \left(\prod_{\ell=i+1}^{j-1} \frac{\ell}{\ell+1} \right) \frac{1}{j+1} = \frac{i+1}{i+2} \frac{i+2}{i+3} \cdots \frac{j-2}{j-1} \frac{j-1}{j} \frac{1}{j+1} = \frac{i+1}{j(j+1)}.$$

Soyons rigoureux !

Pour $j = 2$, on a $\ln(j) < 0$, et donc si on divise par $\ln(j)$, le sens de l'inégalité change. Cela n'a que peu d'importance puisqu'on souhaite ensuite faire tendre j vers $+\infty$.

- 4.c. D'après la formule des probabilités totales appliquée au système complet d'événements $\{[Y = i], i \in \mathbf{N}^*\}$, on a

$$P(Z = j) = \sum_{i=1}^{+\infty} P(Y = i)P_{[Y=i]}(Z = j) = \sum_{i=1}^{j-1} \frac{1}{i(i+1)} \frac{i+1}{j(j+1)} = \frac{1}{j(j+1)} \sum_{i=1}^{j-1} \frac{1}{i}.$$

- 4.d. Notons qu'il n'est pas besoin de calculs pour répondre à cette question : on a $0 \leq Y \leq Z$. Si Z admettait une espérance, alors il en serait de même de Y , ce qui n'est pas le cas d'après la question 2.d. Donc Z n'admet pas d'espérance.

Toutefois, cette méthode n'est pas forcément totalement satisfaisante¹ car elle utilise un résultat qui est bien au programme (l'existence d'une espérance par domination), mais qui a été admis et non démontré.

Notons plutôt que $jP(Z = j) = \frac{1}{j+1} \sum_{i=1}^{j-1} \frac{1}{i}$.

Et donc d'après la question 3.c, $jP(Z = j) \underset{j \rightarrow +\infty}{\sim} \frac{\ln(j)}{j}$.

Mais pour $j \geq 3$, $\frac{\ln(j)}{j} \geq \frac{1}{j} \geq 0$, et puisque la série de terme général $\frac{1}{j}$ diverge, il en est de même de la série de terme général $jP(Z = j)$, de sorte que Z n'admet pas d'espérance.

5. Informatique

- 5.a. Notons qu'il n'y a aucun besoin de calculer les déplacements successifs du mobile pour obtenir sa position après le $n^{\text{ème}}$ déplacement : indépendamment de ce qui s'est passé précédemment, lors du $n^{\text{ème}}$ déplacement, il se place de manière équiprobable sur l'un des points d'abscisses $0, 1, \dots, n$.

Et donc il suffit de tirer au hasard un nombre entre 0 et n , suivant la loi $\mathcal{U}(\llbracket 0, n \rrbracket)$.

```
1 n = input('Entrez la valeur de n :');
2 x = grand(1,1,'uin',0,n);
3 disp(x)
```

- 5.b. Le script proposé simule les différents déplacements du mobile, et augmente a de 1 en cas de retour à l'origine, jusqu'au deuxième retour à l'origine (`while a<2`).

Donc à la fin du programme a contient le rang du deuxième retour à l'origine, c'est-à-dire la valeur de Z .

Enfin, la ligne 7 permet de stocker dans la variable y le rang du premier retour à l'origine, c'est-à-dire Y .

Donc les dernières lignes doivent être complétées comme suit :

```
10 disp(y, 'y=');
11 disp(a, 'z=');
```

Subtilité

En réalité, il s'agit d'un système quasi-complet d'événements : pour avoir un système complet, il faudrait ajouter l'événement «le mobile ne revient jamais à l'origine».

Mais comme celui-ci est de probabilité nulle, comme expliqué dans la remarque de la question 2.b, il n'est pas nécessaire de le faire apparaître dans les calculs.

¹ Et n'est probablement pas celle qu'avait en tête le concepteur du sujet.

DEVOIR MAISON 2 (À RENDRE LE 20.09.18)

On pourra utiliser sans justification que $2 < e^1 < 3$.

On s'intéresse dans cet exercice à la série de terme général $u_n = (-1)^n \frac{\ln n}{n}$ pour $n \geq 1$.

1. On note : $\forall n \geq 1, w_n = \sum_{k=1}^n \frac{1}{k} - \ln(n)$.

a. Rappeler les développements limités à l'ordre 2 lorsque x tend vers 0 de $\ln(1+x)$ et $\frac{1}{1+x}$.

b. Montrer alors que : $w_{n+1} - w_n \underset{n \rightarrow +\infty}{\sim} -\frac{1}{2n^2}$.

c. Montrer que la série de terme général $(w_{n+1} - w_n)$ converge, puis que la suite (w_n) converge vers un réel γ appelé **constante d'Euler**.

2. Étudier les variations de la fonction $\varphi : t \mapsto \frac{\ln t}{t}$ sur $]0, +\infty[$. Dresser le tableau de variations de la fonction φ en précisant les limites aux bornes de son ensemble de définition.

3. On note pour tout entier $n \geq 1$,

$$S_n = \sum_{k=1}^n u_k.$$

a. Montrer que les suites $(S_{2n})_{n \geq 2}$ et $(S_{2n+1})_{n \geq 2}$ sont adjacentes.

b. Montrer que la série de terme général u_n converge. Est-elle absolument convergente ?

4. On note pour tout entier $n \geq 1, v_n = \sum_{k=1}^n \frac{\ln k}{k} - \frac{[\ln(n)]^2}{2}$.

a. Justifier que pour tout entier $n \geq 3$, on a :

$$\frac{\ln(n+1)}{n+1} \leq \int_n^{n+1} \frac{\ln(t)}{t} dt.$$

b. En déduire que la suite $(v_n)_{n \geq 3}$ est décroissante et convergente.

5. Montrer que pour tout entier $n \geq 1$,

$$S_{2n} = 2 \sum_{k=1}^n \frac{\ln(2k)}{2k} - \sum_{k=1}^{2n} \frac{\ln(k)}{k}$$

puis que :

$$S_{2n} = \ln(2) \sum_{k=1}^n \frac{1}{k} + v_n - v_{2n} - \frac{[\ln(2)]^2}{2} - \ln(2) \ln(n).$$

6. Démontrer alors que :

$$\sum_{n=1}^{+\infty} (-1)^n \frac{\ln(n)}{n} = \gamma \ln(2) - \frac{[\ln(2)]^2}{2}.$$

CORRECTION DU DEVOIR MAISON 2

ÉCRICOME 2016

1.a. On a

$$\ln(1+x) = x - \frac{x^2}{2} + o_{x \rightarrow 0}(x^2) \text{ et } \frac{1}{1+x} = 1 - x + x^2 + o_{x \rightarrow 0}(x^2).$$

1.b. Par définition de w_n , on a

$$w_{n+1} - w_n = \sum_{k=1}^{n+1} \frac{1}{k} - \ln(n+1) + \sum_{k=1}^n \frac{1}{k} - \ln(n) = \frac{1}{n+1} - \ln(n+1) + \ln(n) = \frac{1}{1+n} - \ln\left(1 + \frac{1}{n}\right).$$

Or, on a $\frac{1}{1+n} = \frac{1}{n} \frac{1}{1+\frac{1}{n}}$ avec $\frac{1}{n} \rightarrow 0$, donc

$$\begin{aligned} w_{n+1} - w_n &= \frac{1}{n} \left(1 - \frac{1}{n} + o\left(\frac{1}{n}\right)\right) - \left(\frac{1}{n} - \frac{1}{2n^2} + o\left(\frac{1}{n^2}\right)\right) \\ &= -\frac{1}{n^2} + \frac{1}{2n^2} + o\left(\frac{1}{n^2}\right) = -\frac{1}{2n^2} + o\left(\frac{1}{n^2}\right) \end{aligned} \quad \left[\underset{n \rightarrow +\infty}{\sim} -\frac{1}{2n^2} \right]$$

1.c. Par critère de comparaison de séries à terme général de signe constant¹, la série de terme général $w_{n+1} - w_n$ converge.
Or, pour $N \in \mathbf{N}^*$, on a

$$\sum_{k=1}^N (w_{k+1} - w_k) = \sum_{k=1}^N w_{k+1} - \sum_{k=1}^N w_k = w_{N+1} - w_1.$$

Et donc $w_{N+1} = \sum_{k=1}^N (w_{k+1} - w_k) + w_1 \xrightarrow{N \rightarrow +\infty} \sum_{k=1}^{+\infty} (w_{k+1} - w_k) + w_1.$

Ainsi la suite $(w_n)_{n \geq 1}$ converge. On note γ sa limite.

2. La fonction φ est dérivable sur $]0, +\infty[$, de dérivée égale à $\varphi'(t) = \frac{1 - \ln(t)}{t^2}$.

Et donc $\varphi'(t) \geq 0 \Leftrightarrow 1 - \ln(t) \geq 0 \Leftrightarrow t \leq e$.

De plus, on a $\lim_{t \rightarrow 0^+} \varphi(t) = -\infty$ et par croissances comparées, $\lim_{t \rightarrow +\infty} \varphi(t) = 0$.

Le tableau de variations de φ est donc donné par

t	0	e	$+\infty$
$\varphi'(t)$		+	0
$\varphi(t)$		$-\infty \rightarrow e^{-1}$	$\rightarrow 0$

3.a. On a

$$\begin{aligned} S_{2n} - S_{2(n+1)} &= S_{2n} - S_{2n+2} = \sum_{k=1}^{2n} \frac{(-1)^k \ln(k)}{k} - \sum_{k=1}^{2n+2} \frac{(-1)^k \ln(k)}{k} \\ &= \frac{\ln(2n+1)}{2n+1} - \frac{\ln(2n+2)}{2n+2} = \varphi(2n+1) - \varphi(2n+2). \end{aligned}$$

Mais pour $n \geq 2$, on a $2n+1 \geq 3 > e$. La fonction φ étant décroissante sur $[e, +\infty[$, il vient $\varphi(2n+1) \geq \varphi(2n+2)$ et donc $S_{2n} - S_{2n+2} \geq 0$: $(S_{2n})_{n \geq 2}$ est décroissante.

De même,

$$S_{2n+1} - S_{2(n+1)+1} = \varphi(2n+3) - \varphi(2n+2) \leq 0$$

🚫 **Danger !**

Un piège grossier dans lequel il ne faudrait pas tomber serait d'utiliser la question précédente pour obtenir des développements limités de $\frac{1}{1+n}$ et $\ln(1+n)$. On ne peut utiliser directement la question précédente car $n \rightarrow +\infty$.

↔ **Équivalents**

Si $u_n = v_n + o(v_n)$, alors $u_n \sim v_n$.

¹ On ne connaît pas le signe de $w_{n+1} - w_n$, mais $-\frac{1}{2n^2}$ est de signe constant.

— **Pour la culture**

On sait assez peu de choses de γ , si ce n'est que

$$\gamma \approx 0,5772.$$

On ignore par exemple si γ est ou non un nombre rationnel (même si l'on suspecte que ce ne soit pas le cas).

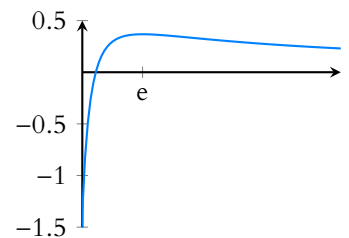


FIGURE 23.1— Représentation graphique de φ .

et donc $(S_{2n+1})_{n \geq 2}$ est croissante.

Enfin,

$$S_{2n+1} - S_{2n} = \sum_{k=2}^{2n+1} u_k - \sum_{k=2}^{2n} u_k = u_{2n} = \frac{\ln(2n)}{2n} \xrightarrow{n \rightarrow +\infty} 0.$$

Ainsi, les suites $(S_{2n})_{n \geq 2}$ et $(S_{2n+1})_{n \geq 2}$ sont adjacentes.

- 3.b. Deux suites adjacentes sont convergentes et de même limite, donc les suites $(S_{2n})_{n \geq 2}$ et $(S_{2n+1})_{n \geq 2}$ convergent vers une même limite ℓ .

Or, il s'agit des suites extraites des termes d'ordre pair (resp. impair) de la suite $(S_n)_{n \geq 4}$. Celle-ci est donc convergente, de limite ℓ .

Mais (S_n) est la suite des sommes partielles de la série de terme général u_n , qui est donc convergente.

On a $|u_n| = \frac{\ln(n)}{n}$. Et donc pour $n \geq 3$, $|u_n| \geq \frac{1}{n}$.

Puisque la série de terme général $\frac{1}{n}$ est divergente², il en est de même de $\sum |u_n|$.

Ainsi, la série de terme général u_n n'est pas absolument convergente.

- 4.a. La fonction φ est décroissante sur $[3, +\infty[$, donc pour tout $t \in [n, n+1]$, $\varphi(t) \geq \varphi(n+1)$. Alors, par croissance de l'intégrale,

$$\int_n^{n+1} \varphi(t) dt \geq \int_n^{n+1} \varphi(n+1) dt \geq \varphi(n+1)$$

et donc
$$\frac{\ln(n+1)}{n+1} \leq \int_n^{n+1} \frac{\ln(t)}{t} dt.$$

- 4.b. Pour $n \geq 3$, on a

$$\begin{aligned} v_{n+1} - v_n &= \sum_{k=1}^{n+1} \frac{\ln(k)}{k} - \frac{[\ln(n+1)]^2}{2} - \sum_{k=1}^n \frac{\ln(k)}{k} + \frac{[\ln(n)]^2}{2} \\ &= \frac{\ln(n+1)}{n+1} + \frac{[\ln(n)]^2 - [\ln(n+1)]^2}{2} \\ &\leq \int_n^{n+1} \frac{\ln t}{t} dt + \frac{[\ln(n)]^2 - [\ln(n+1)]^2}{2} \\ &\leq \left[\frac{[\ln t]^2}{2} \right]_n^{n+1} + \frac{[\ln(n)]^2 - [\ln(n+1)]^2}{2} \\ &\leq 0. \end{aligned}$$

Donc la suite $(v_n)_{n \geq 3}$ est décroissante.

Comme à la question 4.a, on prouve que pour tout entier $k \geq 3$, on a $\frac{\ln k}{k} \geq \int_k^{k+1} \frac{\ln t}{t} dt$.

Et alors, pour $n \geq 3$, on a

$$\begin{aligned} v_n &= \sum_{k=1}^n \frac{\ln k}{k} - \frac{[\ln(n)]^2}{2} \\ &= \frac{\ln 2}{2} + \sum_{k=3}^n \frac{\ln k}{k} - \frac{[\ln(n)]^2}{2} \\ &\geq \frac{\ln 2}{2} + \sum_{k=3}^n \int_k^{k+1} \frac{\ln t}{t} dt - \frac{[\ln(n)]^2}{2} \\ &\geq \frac{\ln 2}{2} + \int_3^{n+1} \frac{\ln t}{t} dt \\ &\geq \frac{\ln 2}{2} + \frac{[\ln(n+1)]^2}{2} - \frac{[\ln(3)]^2}{2} - \frac{[\ln(n)]^2}{2} \\ &\geq \frac{\ln 2}{2} - \frac{[\ln(3)]^2}{2}. \end{aligned}$$

Rappel

Une suite (u_n) converge si et seulement si les suites (u_{2n}) et (u_{2n+1}) convergent vers **une même limite**. Ce n'est plus vrai si ces limites sont différentes, comme le prouve la suite $u_n = (-1)^n$.

² Série de Riemann avec $\alpha = 1$.

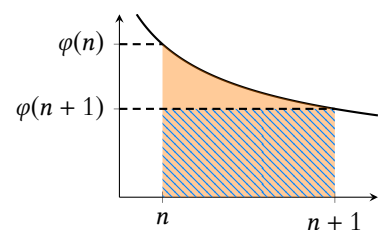


FIGURE 23.2— L'intégrale, qui est l'aire de la partie colorée est plus grande que l'aire de la partie hachurée, qui vaut $\varphi(n+1)$ (c'est l'aire d'un rectangle de largeur 1 et de hauteur $\varphi(n+1)$).

Primitive

On a

$$\frac{\ln t}{t} = \frac{1}{t} \ln t.$$

On reconnaît donc une fonction de la forme $u'v$, dont une primitive est alors donnée par $t \mapsto \frac{[\ln t]^2}{2}$.

Ainsi, la suite $(v_n)_{n \geq 3}$ est minorée, et par le théorème de la limite monotone, elle est donc convergente.

5. Partons de l'expression proposée :

$$\begin{aligned} 2 \sum_{k=1}^n \frac{\ln(2k)}{2k} - \sum_{k=1}^{2n} \frac{\ln(k)}{k} &= 2 \sum_{k=1}^n \frac{\ln(2k)}{2k} - \sum_{k=1}^n \frac{\ln(2k)}{2k} - \sum_{k=0}^{n-1} \frac{\ln(2k+1)}{2k+1} \\ &= \sum_{k=1}^n \frac{\ln(2k)}{2k} - \sum_{k=0}^{n-1} \frac{\ln(2k+1)}{2k+1} \\ &= \sum_{k=1}^{2n} \frac{(-1)^k \ln(k)}{k} = \boxed{S_{2n}}. \end{aligned}$$

On en déduit alors que

$$\begin{aligned} S_{2n} &= \sum_{k=1}^n \frac{\ln(2) + \ln(k)}{k} - \sum_{k=1}^{2n} \frac{\ln(k)}{k} \\ &= \ln(2) \sum_{k=1}^n \frac{1}{k} + \sum_{k=1}^n \frac{\ln(k)}{k} - \sum_{k=1}^{2n} \frac{\ln(k)}{k} \\ &= \ln(2) \sum_{k=1}^n \frac{1}{k} + v_n + \frac{[\ln(n)]^2}{2} - v_{2n} - \frac{[\ln(2n)]^2}{2} \\ &= \ln(2) \sum_{k=1}^n \frac{1}{k} + v_n - v_{2n} + \frac{\ln(n)^2 - (\ln(2) + \ln(n))^2}{2} \\ &= \ln(2) \sum_{k=1}^n \frac{1}{k} + v_n - v_{2n} + \frac{\ln(n)^2 - \ln(2)^2 + 2 \ln(2) \ln(n) - \ln(n)^2}{2} \\ &= \boxed{\ln(2) \sum_{k=1}^n \frac{1}{k} + v_n - v_{2n} + \ln(2) \ln(n) - \frac{[\ln 2]^2}{2}}. \end{aligned}$$

6. De la dernière égalité, il vient

$$S_{2n} = \ln(2) \left(\sum_{k=1}^n \frac{1}{k} - \ln(n) \right) + v_n - v_{2n} - \frac{[\ln(2)]^2}{2}.$$

Mais si l'on note λ la limite de la suite (v_n) , il vient alors, en passant à la limite³ dans cette égalité

$$\lim_{n \rightarrow +\infty} S_{2n} = \sum_{k=1}^{+\infty} \frac{(-1)^k \ln(k)}{k} = \ln(2)\gamma + \lambda - \lambda - \frac{[\ln(2)]^2}{2} = \boxed{\gamma \ln(2) - \frac{[\ln(2)]^2}{2}}.$$

Détails

On a séparé la seconde somme en deux sommes, formées respectivement des termes d'ordre pair et des termes d'ordre impair.

³ Tous les termes admettent une limite lorsque $n \rightarrow +\infty$, donc le passage à la limite est bien justifié.

Somme de série

La limite de (S_{2n}) est la limite de (S_n) , c'est-à-dire la limite de la suite des sommes partielles de la série de terme général u_n .

Cette limite est donc égale à

$$\sum_{k=1}^{+\infty} u_k = \sum_{k=1}^{+\infty} \frac{(-1)^k \ln(k)}{k}.$$

DEVOIR MAISON 2 SUJET DUR (À RENDRE LE 20.09.18)

Notations :

Dans tout le problème, E désigne l'espace vectoriel réel des fonctions continues sur le segment $[0, 1]$ et à valeurs réelles.

Sous réserve d'existence, on note $\varphi : x \mapsto \frac{1}{x} - \sum_{n=1}^{+\infty} \frac{2x}{n^2 - x^2}$ et $\psi : x \mapsto \sum_{n=1}^{+\infty} (-1)^{n-1} \frac{x}{n^2 - x^2}$.

Partie I : Étude de la fonction φ .

1. Montrer que, pour tout nombre réel x qui n'est pas un entier relatif, la série de terme général $u_n(x) = \frac{2x}{n^2 - x^2}$ est convergente.

Dans la suite, on notera D l'ensemble des nombres réels qui ne sont pas des entiers relatifs.

La fonction φ est donc définie sur D .

2. Imparité et périodicité de φ :

- a. Justifier que φ est impaire.

- b. Vérifier que pour x dans D : $\frac{2x}{n^2 - x^2} = \frac{1}{n - x} - \frac{1}{n + x}$.

- c. Montrer que pour x dans D : $\varphi(x + 1) = \varphi(x)$.

La fonction φ est donc périodique de période 1.

3. Continuité de φ :

- a. Justifier pour x dans l'ensemble $D \cup \{0, 1\}$ l'existence de :

$$g(x) = \sum_{n=2}^{+\infty} \frac{2x}{n^2 - x^2} = \sum_{n=2}^{+\infty} \left(\frac{1}{n - x} - \frac{1}{n + x} \right).$$

- b. Vérifier que : $\forall x \in D, \varphi(x) = \frac{1}{x} - \frac{1}{1 - x} + \frac{1}{1 + x} - g(x)$.

- c. Soit h un nombre réel de $\left] -\frac{1}{2}, \frac{1}{2} \right[$, montrer que : $\forall x \in [0, 1], |g(x + h) - g(x)| \leq C|h|$ où

$$C = \sum_{n=2}^{+\infty} \frac{2}{(n-1)\left(n - \frac{3}{2}\right)}.$$

- d. En déduire que g est continue sur $[0, 1]$ puis que φ est continue sur $]0, 1[$.

La fonction φ est donc continue sur D .

4. Étude de φ en 0 et en 1 :

- a. Montrer que : $\varphi(x) \underset{x \rightarrow 0}{\sim} \frac{1}{x}$ et que : $\lim_{x \rightarrow 0} \left(\varphi(x) - \frac{1}{x} \right) = 0$.

- b. Obtenir des résultats similaires lorsque x tend vers 1.

Partie II : Étude de l'opérateur T .

On rappelle que E désigne l'espace vectoriel réel des fonctions continues sur le segment $[0, 1]$ et à valeurs réelles.

T est l'application définie sur E par : $\forall f \in E, \forall x \in [0, 1], [T(f)](x) = f\left(\frac{x}{2}\right) + f\left(\frac{x+1}{2}\right)$.

On note, pour tout entier naturel k , e_k l'élément de E défini par : $\forall x \in [0, 1], e_k(x) = x^k$ et pour tout entier naturel n , F_n le sous-espace vectoriel de E dont une base est $\mathcal{B}_n = (e_k)_{k \in \llbracket 0, n \rrbracket}$.

5. Vérifier que T est un endomorphisme de E .

6. Étude de T sur F_n :

- a. Vérifier que : $\forall f \in F_n, T(f) \in F_n$.

On note T_n l'endomorphisme de F_n défini par : $\forall f \in F_n, T_n(f) = T(f)$.

- b. Déterminer la matrice de T_n dans la base \mathcal{B}_n .

- c. Quelles sont les valeurs propres de T_n ? T_n est-il diagonalisable ?

7. Étude du noyau de l'endomorphisme $(T - \text{Id}_E)$:

a. Montrer que $\text{Ker}(T - 2\text{id}_E)$ n'est pas réduit à $\{0_E\}$.

Soit f un élément de $\text{Ker}(T - 2\text{id}_E)$. On note $m = \min_{x \in [0,1]} f(x)$ et $M = \max_{x \in [0,1]} f(x)$. On fixe x_0 dans $[0, 1]$ tel que $m = f(x_0)$ et x_1 dans $[0, 1]$ tel que $f(x_1) = M$.

b. Montrer que : $f\left(\frac{x_0}{2}\right) = m$.

c. En déduire que : $\forall n \in \mathbf{N}, f\left(\frac{x_0}{2^n}\right) = m$.

d. En déduire que : $m = f(0)$.

e. Faire une étude similaire pour M .

f. Montrer alors que f est constante.

8. Étude de la fonction cot :

Pour tout x dans l'ensemble D , on note $\cot(x) = \pi \frac{\cos(\pi x)}{\sin(\pi x)}$.

a. Vérifier que cot est définie et continue sur D , qu'elle est impaire et périodique de période 1.

b. Montrer que : $\cot(x) \underset{x \rightarrow 0}{\sim} \frac{1}{x}$ et que : $\cot(x) - \frac{1}{x} \underset{x \rightarrow 0}{\sim} -\frac{\pi^2}{3}x$.

c. Obtenir des résultats similaires lorsque x tend vers 1.

d. Démontrer que, pour tout nombre réel x dans D , on a : $\frac{x}{2} \in D, \frac{x+1}{2} \in D$ et

$$\cot\left(\frac{x}{2}\right) + \cot\left(\frac{x+1}{2}\right) = 2\cot(x).$$

9. Calcul de φ :

a. Vérifier que, pour tout nombre réel x dans D , $\varphi\left(\frac{x}{2}\right) + \varphi\left(\frac{x+1}{2}\right) = 2\varphi(x)$.

b. Montrer que $\varphi - \cot$ se prolonge par continuité sur $[0, 1]$.

c. Démontrer alors que $\varphi = \cot$.

Autrement dit : $\forall x \in D, \pi \frac{\cos(\pi x)}{\sin(\pi x)} = \frac{1}{x} - \sum_{n=1}^{+\infty} \frac{2x}{n^2 - x^2}$.

10. Première application :

a. Déterminer $\lim_{x \rightarrow 0} \frac{1 - x \cot(x)}{2x^2}$.

Pour tout nombre réel x dans $]0, 1[$, on pose $\delta(x) = \sum_{n=1}^{+\infty} \frac{1}{n^2 - x^2} - \sum_{n=1}^{+\infty} \frac{1}{n^2}$.

b. Vérifier que $\left| \delta(x) - \frac{x^2}{1 - x^2} \right| \leq x^2 \sum_{n=2}^{+\infty} \frac{1}{n^2(n^2 - 1)}$.

c. En déduire que $\lim_{x \rightarrow 0} \sum_{n=1}^{+\infty} \frac{1}{n^2 - x^2} = \sum_{n=1}^{+\infty} \frac{1}{n^2}$.

d. Déterminer $\sum_{n=1}^{+\infty} \frac{1}{n^2}$.

CORRECTION DU DEVOIR MAISON 2

Partie I : Étude de la fonction φ .

1. Si $x \notin \mathbf{Z}$, alors pour tout $n \in \mathbf{N}^*$, $u_n(x) = \frac{2x}{n^2 - x^2}$ est bien défini.

On a alors $|u_n(x)| \underset{n \rightarrow +\infty}{\sim} \frac{2|x|}{n^2}$. Par critère de comparaison pour les séries à termes positifs, puisque $\sum_{n \geq 1} \frac{1}{n^2}$ étant convergente, il en est de même de $\sum_{n \geq 1} |u_n(x)|$.

Ainsi, la série de terme général $u_n(x)$ est absolument convergente, donc convergente.

2. Imparité et périodicité de φ .

- 2.a. Si $x \in D$, alors $x \notin \mathbf{Z}$ et donc $-x \notin \mathbf{Z}$, de sorte que $-x \in D$. Le domaine de définition de φ est donc symétrique par rapport à 0, et il est donc pertinent de s'intéresser à la parité de φ . Pour $x \in D$, on a alors

$$\varphi(-x) = \frac{1}{-x} - \sum_{n=1}^{+\infty} \frac{-2x}{n^2 - (-x)^2} = -\frac{1}{x} + \sum_{n=1}^{+\infty} \frac{2x}{n^2 - x^2} = -\varphi(x).$$

Donc φ est une fonction impaire.

- 2.b. Pour $x \in D$, $\frac{1}{n-x}$ et $\frac{1}{n+x}$ sont bien définis et

$$\frac{1}{n-x} - \frac{1}{n+x} = \frac{n+x}{n^2-x^2} - \frac{n-x}{n^2-x^2} = \frac{2x}{n^2-x^2}.$$

- 2.c. Pour $x \in D$, on a

$$\varphi(x+1) = \frac{1}{x+1} - \sum_{n=2}^{+\infty} \frac{2(x+1)}{n^2 - (x+1)^2} = \frac{1}{x+1} - \sum_{n=2}^{+\infty} \left(\frac{1}{n-(x+1)} - \frac{1}{n+(x+1)} \right).$$

Pour $N \geq 1$, on a alors

$$\begin{aligned} \frac{1}{x+1} - \sum_{n=1}^N \frac{1}{n-(x+1)} + \sum_{n=1}^N \frac{1}{n+x+1} &= \frac{1}{x+1} + \frac{1}{x} - \sum_{n=2}^N \frac{1}{(n-1)-x} + \sum_{n=1}^N \frac{1}{(n+1)+x} \\ &= \frac{1}{x} - \sum_{k=1}^{N-1} \frac{1}{k-x} + \sum_{\ell=1}^{N+1} \frac{1}{k+x} \\ &= \frac{1}{x} - \sum_{k=1}^{N-1} \left(\frac{1}{k-x} - \frac{1}{k+x} \right) + \frac{1}{N+x} + \frac{1}{N+1+x} \\ &\xrightarrow{N \rightarrow +\infty} \frac{1}{x} - \sum_{n=1}^{+\infty} \left(\frac{1}{n-x} - \frac{1}{n+x} \right) = \varphi(x). \end{aligned}$$

On en déduit que $\varphi(x+1) = \varphi(x)$.

3. Continuité de φ .

- 3.a. Pour $x \in D$, nous avons prouvé à la question 1 que la série de terme général $u_n(x)$ converge, donc en particulier, $\sum_{n \geq 2} u_n(x)$ converge.

Pour $x = 0$ ou $x = 1$, la preuve est la même, il suffit juste de remarquer que pour $n \geq 2$, $\frac{2x}{n^2 - x^2}$ est bien défini car son dénominateur ne s'annule pas.

- 3.b. Soit $x \in D$.

$$\varphi(x) = \frac{1}{x} - \sum_{n=1}^{+\infty} \left(\frac{1}{n-x} - \frac{1}{n+x} \right) = \frac{1}{x} - \frac{1}{1-x} + \frac{1}{1+x} - \sum_{n=2}^{+\infty} \left(\frac{1}{n-x} - \frac{1}{n+x} \right) = \frac{1}{x} - \frac{1}{1-x} + \frac{1}{1+x} - g(x).$$

Signes

Ne pas oublier que le critère des équivalents nécessite des séries de signe constant, ce qui n'est pas le cas de $u_n(x)$, dont le signe peut changer suivant la valeur de n . L'emploi de la valeur absolue permet d'éviter ce problème. Il serait également possible de remarquer que pour n assez grand, $u_n(x)$ est de signe constant.

⚠ Danger !

La tentation est grande de séparer la somme de la série en deux sommes, mais $\sum_{n=2}^{+\infty} \frac{1}{n+x+1}$ et $\sum_{n=2}^{+\infty} \frac{1}{n-x-1}$ ne convergent pas !

3.c. Pour $t \in \left[-\frac{1}{2}, \frac{3}{2}\right]$, posons $g_n(t) = \frac{2t}{n^2 - t^2} = \frac{1}{n-t} - \frac{1}{n+t}$.

Pour $x \in [0, 1]$ et $h \in \left[-\frac{1}{2}, \frac{1}{2}\right]$, on a

$$\begin{aligned} g_n(x+h) - g_n(x) &= \frac{1}{n-x-h} - \frac{1}{n+x+h} - \frac{1}{n-x} + \frac{1}{n+x} \\ &= \frac{1}{n-x-h} - \frac{1}{n-x} + \frac{1}{n+x} - \frac{1}{n+x+h} \\ &= \frac{h}{(n-x)(n-x-h)} + \frac{h}{(n+x+h)(n+x)}. \end{aligned}$$

Or, $n-x \geq n-1$ et $n-x-h \geq n-\frac{3}{2}$ et donc

$$0 \leq \frac{1}{(n-x)(n-x-h)} \leq \frac{1}{(n-1)(n-\frac{3}{2})}.$$

Et de même $n+x \geq n$ et $n+x+h \geq n-\frac{1}{2}$ donc

$$0 \leq \frac{1}{(n+x+h)(n+x)} \leq \frac{1}{(n-\frac{1}{2})n} \leq \frac{1}{(n-1)(n-\frac{3}{2})}.$$

On en déduit que

$$|g_n(x+h) - g_n(x)| \leq \frac{2h}{(n-1)(n-\frac{3}{2})}.$$

Il vient enfin :

$$\begin{aligned} |g(x+h) - g(x)| &= \left| \sum_{n=2}^{+\infty} g_n(x+h) - \sum_{n=2}^{+\infty} g_n(x) \right| \\ &= \left| \sum_{n=2}^{+\infty} (g_n(x+h) - g_n(x)) \right| \leq \sum_{n=2}^{+\infty} |g_n(x+h) - g_n(x)| \\ &\leq \sum_{n=2}^{+\infty} \frac{2|h|}{(n-1)(n-\frac{3}{2})} \\ &\leq |h| \sum_{n=2}^{+\infty} \frac{2}{(n-1)(n-\frac{3}{2})}. \end{aligned}$$

Sous réserve de convergence.

La convergence est établie car cette série est équivalente à une série de Riemann convergente.

3.d. Pour $x \in [0, 1]$ fixé, par le théorème des gendarmes on a $\lim_{h \rightarrow 0} |g(x+h) - g(x)| = 0$ et donc

$$\lim_{h \rightarrow 0} g(x+h) = g(x).$$

Ainsi, g est continue en x , et donc g est continue sur $[0, 1]$.

Les fonctions $x \mapsto \frac{1}{x}$, $x \mapsto \frac{1}{1-x}$ et $x \mapsto \frac{1}{1+x}$ sont continues sur $]0, 1[$, et donc φ est continue sur $]0, 1[$ car somme de fonctions continues.

4. Étude de φ en 0 et en 1.

4.a. On a $\varphi(x) = \frac{1}{x} - \frac{2x}{1-x^2} - g(x)$.

Or, au voisinage de 0, on a $\frac{2x}{1-x^2} + g(x) \xrightarrow{x \rightarrow 0} 0 + g(0) = 0$ et $\frac{1}{1-x} \xrightarrow{x \rightarrow 0^+} +\infty$, donc

$$\frac{2x}{1-x^2} + g(x) = o_{x \rightarrow 0^+} \left(\frac{1}{x} \right).$$

On en déduit que $\varphi(x) = \frac{1}{x} + o_{x \rightarrow 0^+} \left(\frac{1}{x} \right)$ et donc $\varphi(x) \sim_{x \rightarrow 0^+} \frac{1}{x}$.

D'autre part, nous venons de prouver que $\varphi(x) - \frac{1}{x} = -\frac{2x}{1-x^2} - g(x) \xrightarrow{x \rightarrow 0^+} 0$.

4.b. De même, on a

$$\varphi(x) = \frac{1}{x} - \frac{1}{1-x} + \frac{1}{1+x} - g(x).$$

Et puisque $\frac{1}{x} + \frac{1}{1+x} - g(x) \xrightarrow{x \rightarrow 1} 1 + \frac{1}{2} - g(1)$, on a $\frac{1}{x} - \frac{1}{1+x} - g(x) = o\left(\frac{1}{1-x}\right)$ et donc

$$\boxed{\varphi(x) \underset{x \rightarrow 1}{\sim} -\frac{1}{1-x}}$$

Et $\boxed{\varphi(x) - \frac{1}{1-x} \xrightarrow{x \rightarrow 1} \frac{3}{2} - g(1)}$.

Or $g(1) = \sum_{n=2}^{+\infty} \left(\frac{1}{n-1} - \frac{1}{n+1}\right)$.

Pour calculer cette somme¹, repassons par les sommes partielles : pour $N \geq 2$, on a

$$\sum_{n=2}^N \left(\frac{1}{n-1} - \frac{1}{n+1}\right) = \sum_{n=2}^N \frac{1}{n-1} - \sum_{n=2}^N \frac{1}{n+1} = 1 + \frac{1}{2} - \frac{1}{N} - \frac{1}{N+1} \xrightarrow{N \rightarrow +\infty} \frac{3}{2}.$$

Et donc $\boxed{\lim_{x \rightarrow 1} \left(\varphi(x) - \frac{1}{x-1}\right) = 0}$.

¹ On aura reconnu une somme télescopique.

Partie II : Étude de l'opérateur T .

5. Si $f \in E$, alors les fonctions $x \mapsto f\left(\frac{x}{2}\right)$ et $x \mapsto f\left(\frac{x+1}{2}\right)$ sont continues sur $[0, 1]$ par composition de fonctions continues. Et donc $T(f)$ est continue, par somme de fonctions continues, de sorte que $T(f) \in E$.

Soient $f, g \in E$ et $\lambda \in \mathbf{R}$. Alors, pour tout $x \in [0, 1]$,

$$\begin{aligned} [T(\lambda f + g)](x) &= (\lambda f + g)\left(\frac{x}{2}\right) + (\lambda f + g)\left(\frac{x+1}{2}\right) \\ &= \lambda \left(f\left(\frac{x}{2}\right) + f\left(\frac{x+1}{2}\right)\right) + \left(g\left(\frac{x}{2}\right) + g\left(\frac{x+1}{2}\right)\right) = \lambda [T(f)](x) + [T(g)](x). \end{aligned}$$

Et donc $T(\lambda f + g) = \lambda T(f) + T(g)$: T est linéaire, et donc $\boxed{\text{est un endomorphisme de } E}$.

6. Étude de T sur F_n .

- 6.a. Pour tout $k \in \llbracket 0, n \rrbracket$ et tout $x \in [0, 1]$, on a

$$[T(e_k)](x) = \left(\frac{x}{2}\right)^k + \left(\frac{1+x}{2}\right)^k = \frac{x^k}{2^k} + \sum_{i=0}^k \binom{k}{i} \frac{x^i}{2^k}$$

de sorte que $T(e_k) = \frac{e_k}{2^k} + \sum_{i=0}^k \binom{k}{i} \frac{e_i}{2^k} \in F_n$.

Et alors, si $f = \sum_{k=0}^n \lambda_k e_k \in F_n$, par linéarité de T , il vient $T(f) = \sum_{k=0}^n \lambda_k T(e_k) \in F_n$.

Ainsi, $\boxed{\forall f \in F_n, T(f) \in F_n}$.

- 6.b. D'après la question précédente, on a

$$\forall k \in \llbracket 0, n \rrbracket, T_n(e_k) = \frac{e_n}{2^{n-1}} + \sum_{i=0}^{k-1} \binom{k}{i} \frac{e_i}{2^k}.$$

La matrice de T_n dans la base \mathcal{B}_n est donc donnée par

$$\text{Mat}_{\mathcal{B}_n}(T_n) = \begin{pmatrix} T_n(e_0) & T_n(e_1) & \dots & T_n(e_{n-1}) & T_n(e_n) \\ 2 & \frac{1}{2} & \dots & \frac{1}{2^{n-1}} & \frac{1}{2^n} \\ 0 & 1 & \dots & \vdots & \frac{n}{2^n} \\ \vdots & \ddots & \ddots & \vdots & \vdots \\ \vdots & & \ddots & \frac{1}{2^{n-1}} & \frac{n}{2^n} \\ 0 & \dots & \dots & 0 & \frac{1}{2^{n-1}} \end{pmatrix} \begin{matrix} e_0 \\ e_1 \\ \vdots \\ e_{n-1} \\ e_n \end{matrix}.$$

Méthode

Prouver que deux fonctions sont égales, c'est prouver qu'elles prennent la même valeur en tout point de leur ensemble de définition.

Astuce

Pour prouver qu'un sous-espace vectoriel F est stable par un endomorphisme f , il suffit de prouver, pour une base (e_1, \dots, e_n) de F que pour tout i , $f(e_i) \in F$.

6.c. La matrice de T_n dans la base \mathcal{B}_n est triangulaire supérieure, donc ses valeurs propres² sont ses coefficients diagonaux : $\text{Spec}(T_n) = \left\{ \frac{1}{2^{k-1}}, k \in \llbracket 0, n \rrbracket \right\}$.

² Qui sont également les valeurs propres de T_n .

Ainsi, T_n possède $n + 1 = \dim F_n$ valeurs propres distinctes, donc est diagonalisable.

7. Étude du noyau de l'endomorphisme $(T - 2\text{id}_E)$.

7.a. On a $T(e_0) = 2e_0$, et donc $e_0 \in \text{Ker}(T - 2\text{id}_E)$. Puisque $e_0 \neq 0_E$, on a donc $\text{Ker}(T - 2\text{id}_E) \neq \{0_E\}$.

7.b. Puisque $\frac{x_0}{2} \in [0, 1]$, nécessairement $f\left(\frac{x_0}{2}\right) \geq m$.

Et de même, $\frac{x_0 + 1}{2} \in [0, 1]$ et donc $f\left(\frac{x_0 + 1}{2}\right) \geq m$.

De plus, $T(f) = 2f$, de sorte que $f\left(\frac{x_0}{2}\right) + f\left(\frac{x_0 + 1}{2}\right) = 2f(x_0) = 2m$.

Si on avait $f\left(\frac{x_0}{2}\right) > m$, alors nécessairement

$$2m = 2f(x_0) = f\left(\frac{x_0}{2}\right) + f\left(\frac{x_0 + 1}{2}\right) \geq f\left(\frac{x_0}{2}\right) + m > 2m.$$

Ceci est impossible, donc $f\left(\frac{x_0}{2}\right) = m$.

7.c. Prouvons par récurrence sur $n \in \mathbf{N}$ que $f\left(\frac{x_0}{2^n}\right) = m$.

Pour $n = 0$, c'est la définition de x_0 .

Supposons donc que $f\left(\frac{x_0}{2^n}\right) = m$.

Le raisonnement de la question précédente prouve que quel que soit $x \in [0, 1]$ tel que $f(x) = m$, alors $f\left(\frac{x}{2}\right) = m$.

En particulier, $f\left(\frac{1}{2} \frac{x_0}{2^n}\right) = m$, donc $f\left(\frac{x_0}{2^{n+1}}\right) = m$.

Et alors par le principe de récurrence, pour tout $n \in \mathbf{N}$, $f\left(\frac{x_0}{2^n}\right) = m$.

7.d. Lorsque $n \rightarrow +\infty$, $\frac{x_0}{2^n} \rightarrow 0$, et donc par continuité de f , on a $f\left(\frac{x_0}{2^n}\right) \xrightarrow{n \rightarrow +\infty} f(0)$.

D'autre part, $f\left(\frac{x_0}{2^n}\right) \xrightarrow{n \rightarrow +\infty} m$, donc par unicité de la limite, $m = f(0)$.

7.e. Sur le même principe, on prouve que $f\left(\frac{x_1}{2}\right) = M$, puis que $\forall n \in \mathbf{N}$, $f\left(\frac{x_1}{2^n}\right) = M$, et donc, par le même argument de limite, que $f(0) = M$.

7.f. Pour tout $x \in [0, 1]$ on a $f(0) = m \leq f(x) \leq M = f(0)$, donc $\forall x \in [0, 1]$, $f(x) = f(0)$: la fonction f est constante sur $[0, 1]$.

8. Étude de la fonction cot.

8.a. Pour $x \in D$, on a $\pi x \notin \pi\mathbf{Z}$, et donc $\sin(\pi x) \neq 0$, de sorte que $\cot(x)$ est bien défini. La fonction cosinus est paire, et la fonction sinus est impaire, de sorte que

$$\forall x \in D, \cot(-x) = \frac{\pi \cos(-\pi x)}{\sin(-\pi x)} = -\frac{\pi \cos(\pi x)}{\sin(\pi x)} = -\cot(x)$$

donc \cot est impaire.

Enfin, pour $x \in D$, on a

$$\cot(x + 1) = \frac{\pi \cos(\pi(x + 1))}{\sin(\pi(x + 1))} = \frac{\pi \cos(\pi x + \pi)}{\sin(\pi x + \pi)} = \frac{-\pi \cos(\pi x)}{-\sin(\pi x)} = \cot(x)$$

donc \cot est 1-périodique.

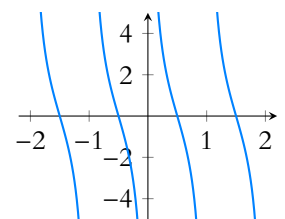


FIGURE 23.1– La fonction cot.

Périodicité

Si $x \in D$, alors $x + 1 \in D$, et donc cela a bien un sens de s'intéresser à la 1-périodicité de cot.

- 8.b. On a $\cos(\pi x) \underset{x \rightarrow 0}{\sim} 1$ et $\sin(\pi x) \underset{x \rightarrow 0}{\sim} \pi x$, de sorte que $\cot(x) \underset{x \rightarrow 0}{\sim} \frac{\pi}{\pi x} = \frac{1}{x}$.
De plus, on a, pour $x \in D$,

$$\begin{aligned} \cot(x) - \frac{1}{x} &= \frac{\pi x \cos(\pi x) - \sin(\pi x)}{x \sin(\pi x)} \\ &= \frac{\pi x \left(1 - \frac{\pi^2}{2} + o(x^2)\right) - \left(\pi x - \frac{\pi^3 x^3}{6} + o(x^3)\right)}{\pi x^2 + o(x^2)} \\ &= \frac{-\frac{\pi^3 x^3}{3} + o(x^3)}{\pi x^2 + o(x^2)} \\ &\underset{x \rightarrow 0}{\sim} \frac{-\pi^3 x^3}{3\pi x^2} = \boxed{-\frac{\pi^2 x}{3}}. \end{aligned}$$

- 8.c. En utilisant la 1-périodicité de \cot , on obtient, en posant $X = x - 1 \xrightarrow{x \rightarrow 1} 0$

$$\cot(x) = \cot(x - 1) = \cot(X) \sim \frac{1}{X} = \frac{1}{x - 1}$$

et de même $\boxed{\cot(x) - \frac{1}{x - 1} \underset{x \rightarrow 1}{\sim} -\frac{\pi^2}{3}(x - 1)}$.

- 8.d. Soit $x \in D$. Raisonnons par l'absurde et supposons que $\frac{x}{2} \notin D$. Alors $\frac{x}{2} \in \mathbf{Z}$ et donc $x = 2\frac{x}{2} \in \mathbf{Z}$, ce qui n'est pas possible. Donc $\boxed{\frac{x}{2} \in D}$.

De même, si $\frac{x+1}{2} \in \mathbf{Z}$, alors $x = 2\frac{x+1}{2} - 1 \in \mathbf{Z}$, ce qui est impossible, donc $\boxed{\frac{x+1}{2} \in D}$.

Pour $x \in D$, on a alors

$$\begin{aligned} \cot\left(\frac{x}{2}\right) + \cot\left(\frac{x+1}{2}\right) &= \frac{\pi \cos\left(\frac{\pi x}{2}\right)}{\sin\left(\frac{\pi x}{2}\right)} + \frac{\pi \cos\left(\frac{\pi x}{2} + \frac{\pi}{2}\right)}{\sin\left(\frac{\pi x}{2} + \frac{\pi}{2}\right)} \\ &= \frac{\pi \cos\left(\frac{\pi x}{2}\right)}{\sin\left(\frac{\pi x}{2}\right)} + \frac{-\pi \sin\left(\frac{\pi x}{2}\right)}{\cos\left(\frac{\pi x}{2}\right)} \\ &= \pi \frac{\cos^2\left(\frac{\pi x}{2}\right) - \sin^2\left(\frac{\pi x}{2}\right)}{\sin\left(\frac{\pi x}{2}\right) \cos\left(\frac{\pi x}{2}\right)} \\ &= \pi \frac{2 \cos(\pi x)}{\sin(\pi x)} = \boxed{2 \cot(x)}. \end{aligned}$$

9. Calcul de φ .

- 9.a. Pour $x \in D$, on a

$$\begin{aligned} \varphi\left(\frac{x}{2}\right) + \varphi\left(\frac{x+1}{2}\right) &= \frac{2}{x} - \sum_{n=1}^{+\infty} \left(\frac{1}{n - \frac{x}{2}} - \frac{1}{n + \frac{x}{2}}\right) + \frac{2}{x+1} - \sum_{n=1}^{+\infty} \left(\frac{1}{n - \frac{x+1}{2}} - \frac{1}{n + \frac{x+1}{2}}\right) \\ &= \frac{2}{x} + \frac{2}{x+1} - \sum_{n=1}^{+\infty} \left(\frac{2}{2n-x} - \frac{2}{2n+x} - \frac{2}{2n-1-x} + \frac{2}{2n+1+x}\right) \end{aligned}$$

Mais pour $N \in \mathbf{N}$, on a

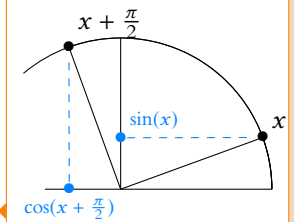
$$\begin{aligned} \sum_{n=1}^N \left(\frac{2}{2n-x} - \frac{2}{2n+x} - \frac{2}{2n-1-x} + \frac{2}{2n+1+x}\right) &= \sum_{\substack{k=2 \\ k \text{ pair}}}^{2N} \left(\frac{2}{k-x} - \frac{2}{k+x}\right) + \sum_{n=1}^N \frac{2}{2n-1-x} - \sum_{n=1}^N \frac{2}{2n+1+x} \\ &= \sum_{\substack{k=2 \\ k \text{ pair}}}^{2N} \left(\frac{2}{k-x} - \frac{2}{k+x}\right) + \sum_{\substack{k=1 \\ k \text{ impair}}}^{2N-1} \frac{2}{k-x} - \sum_{\substack{k=3 \\ k \text{ impair}}}^{2N+1} \frac{2}{k+x} \\ &= \sum_{\substack{k=2 \\ k \text{ pair}}}^{2N} \left(\frac{2}{k-x} - \frac{2}{k+x}\right) + \sum_{\substack{k=1 \\ k \text{ impair}}}^{2N+1} \frac{2}{k-x} - \frac{2}{2N+1-x} - \sum_{\substack{k=1 \\ k \text{ impair}}}^{2N+1} \frac{2}{k+x} + \frac{2}{1+x} \end{aligned}$$

Trigo

Comme (presque) toutes les formules de trigonométrie, la formule

$$\cos\left(x + \frac{\pi}{2}\right) = -\sin(x)$$

peut se retrouver sur un dessin



Trigo

$$\cos(2x) = \cos^2(x) - \sin^2(x).$$

$$\sin(2x) = 2 \cos(x) \sin(x).$$

$$\begin{aligned}
&= \sum_{k=1}^{2N+1} \left(\frac{2}{k-x} - \frac{2}{k+x} \right) - \frac{2}{2N+1-x} + \frac{2}{x+1} \\
&\xrightarrow{N \rightarrow +\infty} 2 \sum_{k=1}^{+\infty} \left(\frac{1}{k-x} - \frac{1}{k+x} \right) + \frac{2}{x+1}.
\end{aligned}$$

Et donc, par unicité de la limite³

$$\begin{aligned}
\varphi\left(\frac{x}{2}\right) + \varphi\left(\frac{x+1}{2}\right) &= \frac{2}{x} + \frac{2}{x+1} - 2 \sum_{k=1}^{+\infty} \left(\frac{1}{k-x} - \frac{1}{k+x} \right) - \frac{2}{x+1} \\
&= \frac{2}{x} - 2 \sum_{k=1}^{+\infty} \left(\frac{1}{k-x} - \frac{1}{k+x} \right) = \boxed{2\varphi(x)}.
\end{aligned}$$

³ La somme que nous venons de calculer converge également, lorsque N tend vers $+\infty$ vers

$$\varphi\left(\frac{x}{2}\right) + \varphi\left(\frac{x+1}{2}\right) - \frac{2}{x} - \frac{2}{x+1}.$$

- 9.b. La fonction $\varphi - \cot$ est continue sur $]0, 1[$ car somme de fonctions continues. De plus, en utilisant les résultats des questions 4.a et 8.b, on a

$$\lim_{x \rightarrow 0^+} (\varphi(x) - \cot(x)) = \lim_{x \rightarrow 0^+} \left(\varphi(x) - \frac{1}{x} \right) - \lim_{x \rightarrow 0^+} \left(\cot(x) - \frac{1}{x} \right) = 0 - 0 = 0.$$

Donc $\varphi - \cot$ est prolongeable par continuité en 0.

De même, on a

$$\lim_{x \rightarrow 1^-} (\varphi(x) - \cot(x)) = \lim_{x \rightarrow 1^-} \left(\varphi(x) - \frac{1}{x-1} \right) - \lim_{x \rightarrow 1^-} \left(\cot(x) - \frac{1}{x-1} \right) = 0 - 0 = 0.$$

Et donc $\varphi - \cot$ est également prolongeable par continuité en 1.

Ainsi, $\varphi - \cot$ se prolonge en une fonction continue sur $[0, 1]$.

- 9.c. Soit $f : [0, 1] \rightarrow \mathbf{R}$ la fonction obtenue par prolongement par continuité de $\varphi - \cot$. Alors, d'après les questions 8.d et 9.a, pour tout $x \in]0, 1[$,

$$f\left(\frac{x}{2}\right) + f\left(\frac{x+1}{2}\right) = \varphi\left(\frac{x}{2}\right) - \cot\left(\frac{x}{2}\right) + \varphi\left(\frac{x+1}{2}\right) - \cot\left(\frac{x+1}{2}\right) = 2\varphi(x) - 2\cot(x) = 2f(x).$$

Par continuité de f , en faisant tendre x vers 0, on obtient également

$$f(0) + f\left(\frac{1}{2}\right) = 2f(0).$$

Et en faisant tendre x vers 1, on a $f(1) + f\left(\frac{1+1}{2}\right) = 2f(1)$.

Ainsi, pour tout $x \in [0, 1]$, $f\left(\frac{x}{2}\right) + f\left(\frac{x+1}{2}\right) = 2f(x)$: la fonction f , qui est continue sur

$[0, 1]$, vérifie $T(f) = 2f$ et donc appartient à $\text{Ker}(T - 2\text{id}_E)$.

D'après la question 7.f, la fonction f est donc constante sur $[0, 1]$. Puisqu'elle s'annule en 0, il vient : $\forall x \in [0, 1], f(x) = 0$.

En particulier, pour $x \in]0, 1[$, $\varphi(x) = \cot(x)$.

Et si $x \in D$, alors $x - [x] \in]0, 1[$, de sorte que par 1-périodicité,

$$\varphi(x) = \varphi(x - [x] \times 1) = \cot(x - [x] \times 1) = \cot(x).$$

On en déduit que $\boxed{\varphi = \cot}$.

10. Première application

- 10.a. À l'aide de l'équivalent obtenu en 8.b, il vient $1 - x \cot(x) \underset{x \rightarrow 0}{\sim} \frac{\pi^2}{3} x^2$ et donc

$$\frac{1 - x \cot(x)}{2x^2} \underset{x \rightarrow 0}{\sim} \frac{\pi^2}{6} \xrightarrow{x \rightarrow 0} \boxed{\frac{\pi^2}{6}}.$$

Prolongement

D'après les calculs qui précèdent, on a donc

$$f(0) = f(1) = 0.$$

10.b. On a

$$\begin{aligned}
 \left| \delta(x) - \frac{x^2}{1-x^2} \right| &= \left| \sum_{n=1}^{+\infty} \frac{1}{n^2-x^2} - \frac{x^2}{1-x^2} - \sum_{n=1}^{+\infty} \frac{1}{n^2} \right| \\
 &= \left| \underbrace{\frac{1}{1-x^2} - \frac{x^2}{1-x^2}}_{=1} + \sum_{n=2}^{+\infty} \frac{1}{n^2-x^2} - 1 - \sum_{n=2}^{+\infty} \frac{1}{n^2} \right| \\
 &= \left| \sum_{n=2}^{+\infty} \left(\frac{1}{n^2-x^2} - \frac{1}{n^2} \right) \right| \\
 &= \sum_{n=2}^{+\infty} \frac{x^2}{(n^2-x^2)n^2} \\
 &\leq \boxed{x^2 \sum_{n=2}^{+\infty} \frac{1}{(n^2-1)n^2}}.
 \end{aligned}$$

La valeur absolue n'était pas utile car il s'agit clairement d'une quantité positive.

Puisque $x \in]0, 1[$,

$$n^2 - x^2 \geq n^2 - 1.$$

10.c. Puisque $\sum_{n=2}^{+\infty} \frac{1}{n^2(n^2-1)}$ est une constante, lorsque $x \rightarrow 0$, on a $x^2 \sum_{n=2}^{+\infty} \frac{1}{n^2(n^2-1)} \xrightarrow{x \rightarrow 0} 0$.

Par le théorème des gendarmes, on en déduit que $\lim_{x \rightarrow 0} \left(\delta(x) - \frac{x^2}{1-x^2} \right) = 0$. Et donc

$$\lim_{x \rightarrow 0} \delta(x) = \lim_{x \rightarrow 0} \left(\delta(x) - \frac{x^2}{1-x^2} \right) + \lim_{x \rightarrow 0} \frac{x^2}{1-x^2} = 0 + 0 = 0.$$

Soit $\boxed{\lim_{x \rightarrow 0} \sum_{n=1}^{+\infty} \frac{1}{n^2-x^2} = \sum_{n=1}^{+\infty} \frac{1}{n^2}}$.

10.d. En utilisant $\varphi = \cot$, il vient, pour $x \in D$, $1 - x \cot x = 1 - 1 + \sum_{n=1}^{+\infty} \frac{x^2}{n^2-x^2}$ et donc

$$\frac{1 - x \cot(x)}{2x^2} = \sum_{n=1}^{+\infty} \frac{1}{n^2-x^2}.$$

Ainsi, en utilisant 10.a et 10.c, il vient

$$\sum_{n=1}^{+\infty} \frac{1}{n^2} = \lim_{x \rightarrow 0} \sum_{n=1}^{+\infty} \frac{1}{n^2-x^2} = \lim_{x \rightarrow 0} \frac{1 - x \cot(x)}{2x^2} = \boxed{\frac{\pi^2}{6}}.$$

DEVOIR MAISON 3 (À RENDRE LE 04.10.18)

Dans ce problème, n est un entier naturel supérieur ou égal à 3 et p est un entier naturel.

Un jeu oppose n joueurs J_1, J_2, \dots, J_n .

Le jeu se déroule de la façon suivante : une pièce équilibrée est lancée $(2p + 1)$ fois. Avant les lancers, chaque joueur écrit une liste de prévisions pour ces lancers. Cette liste contient donc une suite de $(2p + 1)$ caractères P (pour *pile*) ou F (pour *face*). Les gagnants sont les joueurs ayant le plus grand nombre de prévisions correctes et ils se partagent équitablement la somme de $n!$ euros.

Par exemple, pour $p = 1$, si les lancers donnent trois fois *pile*, le joueur ayant noté (P, F, P) a deux prévisions correctes, et si les lancers donnent dans cet ordre P, F, P , le joueur ayant noté (F, P, F) n'a aucune prévision correcte.

Pour i de $\llbracket 1, n \rrbracket$, on note X_i la variable aléatoire égale au nombre de prévisions correctes du joueur J_i , on note G_i la variable aléatoire égale au gain du joueur J_i et $E(G_i)$ l'espérance de G_i .

L'objectif du problème est de déterminer l'espérance de gain du joueur J_1 selon deux stratégies présentées dans les parties 2 et 3.

Partie 1 : quelques résultats utiles pour les parties suivantes.

1. Montrer que les variables X_i suivent toutes la même loi binomiale dont on donnera les paramètres.

On pose alors, pour tout i de $\llbracket 1, n \rrbracket$ et tout k de $X_i(\Omega)$: $q_k = P(X_i = k)$ et $r_k = P(X_i \leq k)$.

2. On pose $S_p = \sum_{k=0}^p \binom{2p+1}{k}$ et $T_p = \sum_{k=p+1}^{2p+1} \binom{2p+1}{k}$.

a. Calculer $S_p + T_p$.

b. Montrer que $S_p = T_p$.

c. Dédire des deux résultats précédents la valeur de S_p puis montrer que $r_p = \frac{1}{2}$.

Partie 2 : les joueurs jouent au hasard et indépendamment les uns des autres.

Dans cette partie, chaque joueur fait ses choix indépendamment des autres joueurs.

3. Montrer que $G_1(\Omega) = \{0\} \cup \left\{ \frac{n!}{j+1} \mid j \in \llbracket 0, n-1 \rrbracket \right\}$.

4. a. Montrer que $P_{[X_1=0]} \left(G_1 = \frac{n!}{n} \right) = (q_0)^{n-1}$.

b. Montrer que, pour tout j de $\llbracket 0, n-2 \rrbracket$: $P_{[X_1=0]} \left(G_1 = \frac{n!}{j+1} \right) = 0$.

c. En déduire que l'espérance de G_1 conditionnellement à l'événement $[X_1 = 0]$ est :

$$E(G_1 | X_1 = 0) = (n-1)!(q_0)^{n-1}$$

5. a. Établir que, pour tout k non nul de $X_1(\Omega)$ et pour tout j élément de $\llbracket 0, n-1 \rrbracket$, on a :

$$P_{[X_1=k]} \left(G_1 = \frac{n!}{j+1} \right) = \binom{n-1}{j} (q_k)^j (r_{k-1})^{n-1-j}$$

- b. Établir que $\frac{1}{j+1} \binom{n-1}{j} = \frac{1}{n} \binom{n}{j+1}$ puis en déduire que, pour tout k non nul de $X_1(\Omega)$, l'espérance de G_1 conditionnellement à l'événement $[X_1 = k]$ est :

$$E(G_1 | X_1 = k) = (n-1)! \frac{(r_k)^n - (r_{k-1})^n}{q_k}$$

c. Vérifier que cette expression reste valable pour $k = 0$ en posant $r_{-1} = 0$.

6. Utiliser les questions 5.b. et 5.c. pour établir que $E(G_1) = (n-1)!$.

Partie 3 : J_1 et J_2 forment un groupe et les autres joueurs jouent comme dans la partie 2.

Dans cette partie, J_1 et J_2 adoptent la stratégie suivante : J_1 joue au hasard mais J_2 joue, pour chaque lancer, les prévisions contraires de J_1 . Par exemple, pour $p = 1$, si J_1 a choisi (F, P, P) alors J_2 choisit (P, F, F) .

On note G' le gain du groupe formé par ces deux joueurs, J_1 et J_2 décidant de partager équitablement ce gain. On a donc, en désignant par G'_1 et G'_2 les gains respectifs de J_1 et J_2 : $G' = G'_1 + G'_2$ et $G'_1 = G'_2$.

On pose, pour tout i de $\llbracket 1, \dots, n \rrbracket$ et pour tout k de $X_i(\Omega)$: $q_k = P(X_i = k)$ et $r_k = P(X_i \leq k)$. On note Y la variable aléatoire égale au nombre de prévisions correctes du meilleur de J_1 et J_2 .

7. a. Montrer que un et un seul des joueurs J_1 et J_2 a au moins $p + 1$ prévisions correctes.
b. En déduire que : $Y(\Omega) = \llbracket p + 1, 2p + 1 \rrbracket$.
8. Vérifier que, dans l'exemple donné au début de cette partie, Y prend la valeur 3 si les lancers donnent dans cet ordre F, P, P ou P, F, F et Y prend la valeur 2 sinon.
9. Pour tout k de $\llbracket p + 1, 2p + 1 \rrbracket$, montrer que : $P(Y = k) = 2q_k$.
10. Montrer que $G'(\Omega) = \{0\} \cup \left\{ \frac{n!}{j+1} \mid j \in \llbracket 0, n-2 \rrbracket \right\}$.
11. a. Établir que, pour tout k de $\llbracket p + 1, 2p + 1 \rrbracket$ et tout j de $\llbracket 0, n-1 \rrbracket$, on a :

$$P_{[Y=k]} \left(G' = \frac{n!}{j+1} \right) = \binom{n-2}{j} (q_k)^j (r_{k-1})^{n-2-j}.$$

- b. En déduire que, pour tout k de $\llbracket p + 1, 2p + 1 \rrbracket$, l'espérance de G' conditionnellement à l'événement $[Y = k]$ est :

$$E(G' | Y = k) = n(n-2)! \frac{(r_k)^{n-1} - (r_{k-1})^{n-1}}{q_k}$$

12. a. En déduire, en utilisant le résultat de la question 2.c, que :

$$E(G') = 2n(n-2)! \left(1 - \frac{1}{2^{n-1}} \right).$$

- b. Montrer par récurrence que : $\forall n \in \mathbf{N}, n \geq 3 \Rightarrow 2^{n-1} > n$.
- c. Déterminer $E(G'_1)$ et vérifier que la stratégie adoptée par les joueurs J_1 et J_2 est avantageuse pour J_1 (et donc pour J_2) du point de vue de l'espérance de leur gain.

CORRECTION DU DEVOIR MAISON 3

EDHEC 2008

Partie 1 : quelques résultats utiles pour les parties suivantes

1. Les prévisions du joueur i étant fixées, chacune d'entre elles a une probabilité égale à $\frac{1}{2}$ de se réaliser, et les lancers des dés étant indépendants, chacune se réalise ou non indépendamment des autres.

Ainsi, X_i suit la loi binomiale $\mathcal{B}\left(2p+1, \frac{1}{2}\right)$.

- 2.a. La somme $S_p + T_p$ fait clairement apparaître un binôme de Newton :

$$S_p + T_p = \sum_{k=0}^p \binom{2p+1}{k} + \sum_{k=p+1}^{2p+1} \binom{2p+1}{k} = \sum_{k=0}^{2p+1} \binom{2p+1}{k} = (1+1)^{2p+1} = \boxed{2^{2p+1}}.$$

- 2.b. Rappelons que $\forall n \in \mathbf{N}, \forall k \in \llbracket 0, n \rrbracket, \binom{n}{k} = \binom{n}{n-k}$ et donc pour $n = 2p+1$, on a

$$S_p = \sum_{k=0}^p \binom{2p+1}{k} = \sum_{k=0}^p \binom{2p+1}{2p+1-k} = \sum_{k=p+1}^{2p+1} \binom{2p+1}{k} = \boxed{T_p}.$$

- 2.c. On en déduit que $S_p + T_p = 2S_p$, et donc $2S_p = 2^{2p+1} \Leftrightarrow \boxed{S_p = 4^p}$.

Ainsi, puisque nous connaissons la loi des X_i , il vient

$$r_p = P(X_i \leq p) = \sum_{k=0}^p P(X_i = k) = \sum_{k=0}^p \binom{2p+1}{k} \frac{1}{2^k} \frac{1}{2^{2p+1-k}} = \frac{1}{2^{2p+1}} \sum_{k=0}^p \binom{2p+1}{k} = \boxed{\frac{1}{2}}.$$

Partie 2 : les joueurs jouent au hasard et indépendamment les uns des autres

3. Il est clair que $G_1(\Omega) \subset \{0\} \cup \left\{ \frac{n!}{j+1}, j \in \llbracket 0, n-1 \rrbracket \right\}$ car le gain du joueur 1 est soit nul¹, soit $n!$ divisé par le nombre de joueurs ayant le plus de prévisions correctes, et ce nombre de joueurs est nécessairement compris entre 1 et n .
De plus, 0 est bien une valeur possible pour G_1 , par exemple dans le cas où le joueur 1 n'a rien prévu correctement et que les autres ont une (ou plusieurs) prévisions exactes. Et pour $j \in \llbracket 0, n-1 \rrbracket$, le joueur 1 peut gagner exactement $\frac{n!}{j+1}$ dans le cas (par exemple) où toutes ses prévisions sont correctes et où exactement j autres joueurs ont prévu les mêmes résultats.

- 4.a. On a $P(X_1 = 0) = q_0$ (par définition de q_0), et l'événement $[X_1 = 0] \cap \left[G_1 = \frac{n!}{n}\right]$ est l'événement «tous les joueurs ont toutes leurs prévisions fausses», qui se réécrit

$$[X_1 = 0] \cap \left[G_1 = \frac{n!}{n}\right] = [X_1 = 0] \cap [X_2 = 0] \cap \dots \cap [X_n = 0].$$

Mais par indépendance des variables X_i , on a

$$P([X_1 = 0] \cap \dots \cap [X_n = 0]) = P(X_1 = 0) \dots P(X_n = 0) = q_0^n.$$

Ainsi, il vient

$$P_{[X_1=0]} \left(G_1 = \frac{n!}{n}\right) = \frac{P([X_1 = 0] \cap [G_1 = \frac{n!}{n}])}{P(X_1 = 0)} = \frac{q_0^n}{q_0} = \boxed{q_0^{n-1}}.$$

- 4.b. Si $X_1 = 0$, alors pour $j < n-1$, $P\left(\left[G_1 = \frac{n!}{j+1}\right] \cap [X_1 = 0]\right) = 0$.

En effet, si le joueur 1 n'a aucune prévision correcte, son gain ne peut être strictement positif qu'à condition que tous les autres joueurs n'aient aucune prévision correcte. Et donc G_1 ne peut prendre que la valeur 0 ou la valeur $\frac{n!}{n}$. En particulier, ce gain ne peut valoir $\frac{n!}{j+1}$.

Astuce

La somme des coefficients binomiaux est en fait $(a+b)^n$ avec $a = b = 1$.

Explication

On a utilisé ici la symétrie des coefficients binomiaux, et le fait que nous savons calculer la somme des coefficients d'une ligne du triangle de Pascal. Ce qui nous permet de calculer la somme d'une demi-ligne.

¹ si ce n'est pas lui qui a le plus de prévisions correctes.

Rédaction

Ce type de question n'est pas forcément facile à rédiger, car il faut essayer de trouver le bon équilibre entre l'intuition que l'on se fait de la situation et un raisonnement probabiliste avec des événements, des formules, etc.

4.c. On en déduit que

$$\begin{aligned} E(G_1|X_1 = 0) &= 0 \times P_{[X_1=0]}(G_1 = 0) + \sum_{j=0}^{n-1} \frac{n!}{j+1} P_{[X_1=0]} \left(G_1 = \frac{n!}{j+1} \right) \\ &= \frac{n!}{n} P_{[X_1=0]} \left(G_1 = \frac{n!}{n} \right) = (n-1)! q_0^{n-1}. \end{aligned}$$

5.a. Sachant que le joueur 1 a exactement k prévisions correctes, son gain G_1 ne peut valoir $\frac{n!}{j+1}$ que si j autres joueurs ont exactement k prévisions correctes et que tous les autres ont strictement moins de k prévisions correctes.

Or, il y a exactement $\binom{n-1}{j}$ manières de choisir les j joueurs autres² que le joueur 1 ayant k prévisions correctes.

Une fois le choix de ces j joueurs effectué, la probabilité qu'ils aient exactement k bons résultats et que les autres en aient strictement moins est, par indépendance des joueurs,

$$P(X_1 = k)^j \times P(X_1 \leq k-1)^{n-1-j} = q_k^j r_{k-1}^{n-1-j}.$$

Donc au total,

$$P_{[X_1=k]} \left(G_1 = \frac{n!}{j+1} \right) = \binom{n-1}{j} (q_k)^j (r_{k-1})^{n-1-j}.$$

5.b. On a, pour $j \leq n-1$,

$$\frac{1}{j+1} \binom{n-1}{j} = \frac{1}{j+1} \frac{(n-1)!}{j!(n-1-j)!} = \frac{(n-1)!}{(j+1)!(n-(j+1))!} = \frac{1}{n} \frac{n!}{(j+1)!(n-(j+1))!} = \frac{1}{n} \binom{n}{j+1}.$$

On en déduit que

$$\begin{aligned} E(G_1|X_1 = k) &= \sum_{j=0}^{n-1} \frac{n!}{j+1} P_{[X_1=k]} \left(G_1 = \frac{n!}{j+1} \right) = \sum_{j=0}^{n-1} \frac{n!}{j+1} \binom{n-1}{j} (q_k)^j (r_{k-1})^{n-1-j} \\ &= \sum_{j=0}^{n-1} \frac{n!}{n} \binom{n}{j+1} (q_k)^j (r_{k-1})^{n-1-j} = (n-1)! \sum_{j=0}^{n-1} \binom{n}{j+1} (q_k)^j (r_{k-1})^{n-(j+1)} \\ &= (n-1)! \sum_{i=1}^n \binom{n}{i} (q_k)^{i-1} (r_{k-1})^{n-i} \\ &= \frac{(n-1)!}{q_k} \sum_{i=0}^n \binom{n}{i} (q_k)^i (r_{k-1})^{n-i} - (n-1)! \frac{r_{k-1}^n}{q_k} \\ &= \frac{(n-1)!}{q_k} (q_k + r_{k-1})^n - \frac{(n-1)! r_{k-1}^n}{q_k} \\ &= (n-1)! \frac{r_k^n - r_{k-1}^n}{q_k} \end{aligned}$$

5.c. Pour $k = 0$, on retrouve

$$E(G_1|X_1 = 0) = (n-1)! \frac{r_0^n}{q_0} = (n-1)! \frac{q_0^n}{q_0} = (n-1)! q_0^{n-1}$$

ce qui est bien l'expression obtenue à la question 4.c.

6. D'après la formule de l'espérance totale appliquée au système complet d'événements $\{[X_1 = k], k \in \llbracket 0, 2p+1 \rrbracket\}$, on a

$$\begin{aligned} E(G_1) &= \sum_{k=0}^{2p+1} E(G_1|X_1 = k) P(X_1 = k) \\ &= \sum_{k=0}^{2p+1} (n-1)! \frac{(r_k)^n - (r_{k-1})^n}{q_k} q_k \end{aligned}$$

Espérance cond.

Par définition, l'espérance conditionnelle $E(X|A)$ est l'espérance pour la probabilité P_A , donc on a

$$E(X|A) = \sum_{x \in X(\Omega)} x P_A(X = x).$$

² Ce qui explique que l'on prenne $\binom{n-1}{j}$ et non $\binom{n}{j}$.

Classique

C'est l'égalité usuelle

$$k \binom{n}{k} = n \binom{n-1}{k-1}.$$

Chgt d'indice

$i = j + 1$.

Binôme de Newton

Car $q_k + r_{k-1} = r_k$.

Convergence

Il n'y a ici pas de problème de convergence, puisque G_1 est une variable aléatoire qui ne peut prendre qu'un nombre fini de valeurs, donc elle possède une espérance, et donc la formule de l'espérance totale s'applique.

$$\begin{aligned} &= (n-1)! \sum_{k=0}^{2p+1} ((r_k)^n - (r_{k-1})^n) \\ &= (n-1)! r_{2p+1}^n \\ &= \boxed{(n-1)!} \end{aligned}$$

Somme télescopique

$$r_{2p+1} = P(X_1 \leq 2p+1) = 1.$$

Partie 3 : J_1 et J_2 forment un groupe et les autres joueurs jouent comme dans la partie 2

- 7.a. Notons comme précédemment X_i le nombre prévisions correctes de J_i .
Alors, si $X_1 \geq p+1$, $X_2 = 2p+1 - X_1 \leq (2p+1) - (p+1) \leq p$.
Si $X_1 \leq p$, alors $X_2 = (2p+1) - k \geq p+1$.
Dans les deux cas, un et un seul des deux joueurs a au moins $p+1$ prévisions correctes.
- 7.b. Puisque $Y = \max(X_1, X_2)$, on a toujours $Y \geq p+1$, et donc $Y(\Omega) \subset \llbracket p+1, 2p+1 \rrbracket$.
De plus, on a bien l'égalité (et non juste une inclusion), car si J_1 a eu exactement $k \in \llbracket p+1, 2p+1 \rrbracket$ prévisions correctes et que tous les joueurs J_3, \dots, J_n n'ont aucune prévision correcte, alors $Y = k$.
8. Si les lancers donnent F, P, P , alors J_1 a ses trois prévisions correctes (et donc J_2 n'en a aucune), alors que si les lancers donnent P, F, F , J_2 a ses trois prévisions correctes (et donc J_1 n'en a aucune). Dans ces deux cas, Y vaut 3.
Enfin, pour tout autre résultat, l'une au moins des prévisions de J_1 s'avérera correcte et une autre fautive, donc on aura $1 \leq X_1 \leq 2$ et donc $1 \leq X_2 \leq 2$, de sorte que $Y = \max(X_1, X_2) = 2$.
9. Soit $k \in \llbracket p+1, 2p+1 \rrbracket$. Puisque $Y = \max(X_1, X_2)$, on a

$$[Y = k] = ([X_1 = k] \cap [X_2 < k]) \cup ([X_1 < k] \cap [X_2 = k]) \cup ([X_1 = k] \cap [X_2 = k]).$$

Notons que cette fois, X_1 et X_2 ne sont pas indépendantes, bien au contraire³, car on a $X_2 = 2p+1 - X_1$.
De plus, les événements $[X_1 = k]$ et $[X_2 = k]$ sont incompatibles en raison de ce qui a été prouvé à la question 7.a. Donc

$$[Y = k] = ([X_1 = k] \cap [X_2 < k]) \cup ([X_1 < k] \cap [X_2 = k]).$$

Mais $[X_1 = k] \subset [X_2 < k]$, car si $X_1 = k \geq p+1$, nécessairement $X_2 \leq p$. Ainsi, $[X_1 = k] \cap [X_2 < k] = [X_1 = k]$. De même, $[X_2 = k] \cap [X_1 < k] = [X_2 = k]$.
Donc, par incompatibilité des événements $[X_1 = k]$ et $[X_2 = k]$, on a

$$P(Y = k) = P(X_1 = k) + P(X_2 = k) = q_k + P(X_1 = 2p+1 - k) = q_k + q_{2p+1-k}.$$

Or $q_k = \binom{2p+1}{k} \frac{1}{2^{2p+1}} = \binom{2p+1}{2p+1-k} \frac{1}{2^{2p+1}} = q_{2p+1-k}$.

On en déduit que $\boxed{P(Y = k) = 2q_k}$.

10. La seule différence avec la question 3 est qu'on ne peut avoir les n joueurs qui ont tous à la fois le plus grand nombre de prévisions correctes, puisque les deux joueurs J_1 et J_2 n'ont pas le même nombre de prévisions correctes.
Donc $G'(\Omega) \subset \{0\} \cup \left\{ \frac{n!}{j+1}, j \in \llbracket 0, n-2 \rrbracket \right\}$.
Si J_1 a exactement une prévision correcte et que J_3 a toutes les prévisions correctes, alors $G' = 0$.
Enfin, si $j \in \llbracket 0, n-2 \rrbracket$, il est possible que J_1 aie toutes les prévisions correctes, de même que J_3, \dots, J_{j+2} (soit exactement j joueurs), et que tous les autres n'aient aucune prévision correcte. Dans ce cas, on a $G' = \frac{n!}{j+1}$.
En conclusion, on a bien

$$\boxed{G'(\Omega) = \{0\} \cup \left\{ \frac{n!}{j+1}, j \in \llbracket 0, n-2 \rrbracket \right\}}.$$

- 11.a. Soit $k \in \llbracket p+1, 2p+1 \rrbracket$ et soit $j \in \llbracket 0, n-2 \rrbracket$.
Sachant que l'un des deux joueurs J_1, J_2 a exactement k prévisions correctes, la probabilité

Explication

Puisque les deux joueurs jouent à chaque fois des prévisions contraires, à eux eux ils ont toujours $2p+1$ prévisions correctes et donc

$$X_1 + X_2 = 2p+1.$$

³ En fait, X_2 est une fonction affine de X_1 , donc $|\rho_{X_1, X_2}| = 1$, ce qui est le cas où elles sont «le moins indépendantes».

Support

Il est souvent facile de montrer que les valeurs possibles sont parmi telles valeurs. Mais alors on a uniquement une inclusion d'ensembles (comme ici).
Pour montrer qu'on a bien l'égalité, il faut encore prouver que G' peut bien prendre toutes ces valeurs avec une probabilité non nulle.

que G' soit égal à $\frac{n!}{j+1}$ est la probabilité que j joueurs parmi J_3, \dots, J_n aient exactement k prévisions correctes, et tous les autres aient strictement moins de k prévisions correctes, donc au plus $k-1$.

Il y a $\binom{n-2}{j}$ manières de choisir ces j joueurs. Une fois ces joueurs choisis, la probabilité qu'ils aient k prévisions correctes et que les autres en aient strictement moins de k est, par indépendance des joueurs, égale à $(q_k)^j (r_{k-1})^{n-2-j}$.

On en déduit que
$$P_{[Y=k]} \left(G' = \frac{n!}{j+1} \right) = \binom{n-2}{j} (q_k)^j (r_{k-1})^{n-2-j}.$$

11.b. Par définition de l'espérance conditionnelle, on a

$$\begin{aligned} E(G'|Y=k) &= \sum_{j=0}^{n-2} \frac{n!}{j+1} P_{[Y=k]} \left(G' = \frac{n!}{j+1} \right) \\ &= \sum_{j=0}^{n-2} \frac{n!}{j+1} \binom{n-2}{j} (q_k)^j (r_{k-1})^{n-2-j} \end{aligned}$$

$$= \frac{n!}{n-1} \sum_{j=0}^{n-2} \binom{n-1}{j+1} (q_k)^j (r_{k-1})^{n-2-j}$$

$$= n(n-2)! \sum_{i=1}^{n-1} \binom{n-1}{i} (q_k)^{i-1} (r_{k-1})^{n-1-i}$$

$$= \frac{n(n-2)!}{q_k} \sum_{i=0}^{n-1} \binom{n-1}{i} (q_k)^i (r_{k-1})^{n-1-i} - \frac{n(n-2)! r_{k-1}^{n-1}}{q_k}$$

$$= n(n-2)! (q_k + r_{k-1})^{n-1} - \frac{n(n-2)! (r_{k-1})^{n-1}}{q_k}$$

$$= n(n-2)! \frac{(r_k)^{n-1} - (r_{k-1})^{n-1}}{q_k}.$$

$$\frac{1}{j+1} \binom{n-2}{j} = \frac{1}{n-1} \binom{n-1}{j+1}$$

comme à la question 5.b.

Chgt d'indice
 $i = j + 1.$

Binôme de Newton.

$$q_k + r_{k-1} = r_k.$$

12.a. Par application de la formule de l'espérance totale au système complet d'événements $\{[Y=k], k \in \llbracket p+1, 2p+1 \rrbracket\}$

$$\begin{aligned} E(G') &= \sum_{k=p+1}^{2p+1} E(G'|Y=k) P(Y=k) \\ &= \sum_{k=p+1}^{2p+1} n(n-2)! \frac{(r_k)^{n-1} - (r_{k-1})^{n-1}}{q_k} 2q_k \end{aligned}$$

$$= 2n(n-2)! \sum_{k=p+1}^{2p+1} ((r_k)^{n-1} - (r_{k-1})^{n-1})$$

$$= 2n(n-2)! ((r_{2p+1})^{n-1} - (r_p)^{n-1})$$

$$= 2n(n-2)! \left(1 - \frac{1}{2^{n-1}} \right)$$

Convergence
Comme à la question 6, pas de problème de convergence car G' est une variable finie, donc admet une espérance.

Somme télescopique.

D'après 2.c.

12.b. Pour $n=3$, on a $2^{3-1} = 4 > 3$. Si pour $n > 3$, on a $2^{n-1} > n$, alors $2^n > 2n > n+1$. Ainsi, par le principe de récurrence,

$$\forall n \geq 3, 2^{n-1} > n.$$

12.c. Puisque $G' = 2G'_1$, on a $E(G'_1) = \frac{1}{2} E(G') = n(n-2)! \left(1 - \frac{1}{2^{n-1}} \right)$.

Or, $\frac{1}{2^{n-1}} < \frac{1}{n}$ et donc $1 - \frac{1}{2^{n-1}} > 1 - \frac{1}{n} = \frac{n-1}{n}$. On en déduit que

$$E(G'_1) > n(n-2)! \frac{n-1}{n} = (n-1)! = E(G_1).$$

Par conséquent, la seconde stratégie est en moyenne plus avantageuse pour J_1 .

DEVOIR MAISON 4 (À RENDRE LE 11.10.18)

Vous traiterez **au choix** l'un des deux exercices suivants, le second étant plus difficile que le premier.

EXERCICE 1

On désigne par α un entier strictement supérieur à 1, et on pose, pour tout entier naturel n non nul :

$$u_n(\alpha) = \int_0^{+\infty} \frac{dt}{(1+t^\alpha)^n}.$$

Dans la suite de l'exercice, on écrira u_n au lieu de $u_n(\alpha)$.

1. **a.** Vérifier que, pour tout $n \in \mathbf{N}^*$, le réel u_n est bien défini et que $u_n > 0$.
b. Étudier les variations de la suite $(u_n)_{n \geq 1}$, et en conclure qu'elle converge.
2. **a.** Montrer, grâce à une intégration par parties, que : $\forall n \in \mathbf{N}^*$, $u_n = n\alpha(u_n - u_{n+1})$.
b. En déduire que pour tout entier n supérieur ou égal à 2 on a $u_n = u_1 \prod_{k=1}^{n-1} \left(1 - \frac{1}{k\alpha}\right)$.
3. En considérant $\ln(u_n)$, montrer que $\lim_{n \rightarrow +\infty} u_n = 0$.
4. Pour tout $n \in \mathbf{N}^*$, on pose $S_n = \sum_{k=1}^n u_k$.
a. Montrer que : $\forall n \in \mathbf{N}^*$, $S_n = \frac{n\alpha}{\alpha-1} u_{n+1}$.
b. En déduire que : $\forall n \geq 2$, $\ln(S_n) = \ln(u_1) + \sum_{k=2}^n \left[\ln\left(1 - \frac{1}{k\alpha}\right) - \ln\left(1 - \frac{1}{k}\right) \right]$.
c. À l'aide d'un développement limité d'ordre 1 en $\frac{1}{k}$, donner un équivalent, lorsque k est au voisinage de $+\infty$, de $\ln\left(1 - \frac{1}{k\alpha}\right) - \ln\left(1 - \frac{1}{k}\right)$.
d. Conclure quant à la nature de la série de terme général u_n .
5. Dans cette question, on suppose que $\alpha = 2$. Compléter le programme suivant afin qu'il retourne la valeur de u_n :

```

1 fonction y=u(n)
2   y = ...
3   for k=...
4     y=...
5   end
6 endfunction

```

EXERCICE 2

Pour tout entier $n \geq 1$, on considère les intégrales :

$$u_n = \int_0^1 \frac{dt}{1+t+t^2+\dots+t^{n-1}} \text{ et } v_n = \int_0^1 \frac{u^{1/n} - u^{2/n}}{1-u} du.$$

1. **Convergence de la suite $(u_n)_{n \geq 1}$.**

a. Vérifier que :

$$\forall n \geq 1, \quad \forall t \in [0, 1], \quad \frac{1}{1+t+t^2+\dots+t^{n-1}} - (1-t) = \frac{(1-t)t^n}{1-t^n}.$$

En déduire que : $\forall n \geq 1, 0 \leq u_n - \frac{1}{2} \leq \frac{1}{n+1}$.

Quelle est la limite de la suite $(u_n)_{n \geq 1}$?

b. En utilisant le changement de variable $u = t^n$, établir que :

$$\forall n \geq 1, \quad u_n - \frac{1}{2} = \frac{v_n}{n}.$$

2. Résultats intermédiaires.

a. Pour tout entier $k \geq 1$, calculer la limite suivante : $\lim_{x \rightarrow 1} \frac{(\ln x)^k}{x-1}$.

b. Soit k un entier naturel non nul.

Prouver la convergence de l'intégrale $\int_0^1 \frac{(\ln x)^k}{x-1} dx$.

c. On introduit la fonction f définie sur \mathbf{R} par : $\forall x \in \mathbf{R}, f(x) = e^x - e^{2x}$.

À l'aide de l'inégalité de Taylor-Lagrange en 0 à l'ordre 1 appliquée à la fonction f , montrer que

$$\forall x \in]-\infty, 0], |e^x - e^{2x} + x| \leq \frac{3x^2}{2}.$$

3. Application.

a. En utilisant la question 2, démontrer que

$$\forall n \geq 1, \left| v_n + \frac{1}{n} \int_0^1 \frac{\ln u}{1-u} du \right| \leq \frac{3}{2n^2} \int_0^1 \frac{(\ln u)^2}{1-u} du.$$

b. On considère l'intégrale $I = \int_0^1 \frac{\ln u}{1-u} du$ que l'on ne cherchera pas à calculer.

Donner un équivalent de v_n puis un équivalent de $u_n - \frac{1}{2}$ en fonction de I .

CORRECTION DU DEVOIR MAISON 4

EXERCICE 1 (EDHEC 2009)

- 1.a. La fonction $t \mapsto \frac{1}{(1+t^\alpha)^n}$ est continue et positive sur $[0; +\infty[$. Donc l'éventuel problème de convergence est en $+\infty$. Or on a

$$\frac{1}{(1+t^\alpha)^n} \underset{t \rightarrow +\infty}{\sim} \frac{1}{t^{\alpha n}}$$

et puisque $\alpha > 1$ et $n \geq 1$, $\alpha n > 1$, de sorte que l'intégrale de Riemann $\int_1^{+\infty} \frac{dt}{t^{\alpha n}}$ converge.

Par critère de comparaison pour les fonctions positives, il en est de même de $\int_0^{+\infty} \frac{dt}{(1+t^\alpha)^n}$,

et donc u_n est bien défini.

Et puisque la fonction intégrée est strictement positive, par croissance de l'intégrale, on a $u_n > 0$.

- 1.b. Soit $n \in \mathbf{N}^*$ et $t \in \mathbf{R}_+$. Alors $1 \leq 1+t^\alpha$, et donc $1 \leq (1+t^\alpha)^n \leq (1+t^\alpha)^{n+1}$. Par passage à l'inverse, on a alors

$$0 \leq \frac{1}{(1+t^\alpha)^{n+1}} \leq \frac{1}{(1+t^\alpha)^n}.$$

Ceci étant vérifié pour tout $t \geq 0$, par croissance de l'intégrale, on a alors

$$0 \leq \int_0^{+\infty} \frac{dt}{(1+t^\alpha)^{n+1}} \leq \int_0^{+\infty} \frac{dt}{(1+t^\alpha)^n} \Leftrightarrow 0 \leq u_{n+1} \leq u_n.$$

Et donc la suite (u_n) est décroissante.

Alors (u_n) étant décroissante et minorée (par 0), elle converge.

- 2.a. Procédons à une intégration par parties sur un segment de la forme $[0, A]$, $A \geq 0$. En posant $u' = 1$ et $v = \frac{1}{(1+t^\alpha)^n}$, soit $u = t$ et $v' = -\frac{n\alpha t^{\alpha-1}}{(1+t^\alpha)^{n+1}}$, on définit bien deux fonctions de classe \mathcal{C}^1 sur $[0, A]$ et

$$\int_0^A \frac{dt}{(1+t^\alpha)^n} = \left[\frac{t}{(1+t^\alpha)^n} \right]_0^A + \int_0^A \frac{n\alpha t^\alpha}{(1+t^\alpha)^{n+1}} dt = \frac{A}{(1+A^\alpha)^n} + n\alpha \int_0^A \frac{t^\alpha}{(1+t^\alpha)^{n+1}} dt$$

En passant à la limite lorsque $A \rightarrow +\infty$, on a alors

$$\begin{aligned} u_n &= \int_0^{+\infty} \frac{dt}{(1+t^\alpha)^n} = n\alpha \int_0^{+\infty} \frac{t^\alpha}{(1+t^\alpha)^{n+1}} dt \\ &= n\alpha \int_0^{+\infty} \frac{t^\alpha + 1 - 1}{(1+t^\alpha)^{n+1}} dt = n\alpha \left(\int_0^{+\infty} \frac{t^\alpha + 1}{(1+t^\alpha)^{n+1}} dt - \int_0^{+\infty} \frac{dt}{(1+t^\alpha)^n} \right) \\ &= n\alpha(u_n - u_{n+1}). \end{aligned}$$

- 2.b. Procédons par récurrence sur n . D'après la question précédente, on a

$$u_1 = \alpha(u_1 - u_2) \Leftrightarrow \alpha u_2 = (\alpha + 1)u_1 \Leftrightarrow u_2 = u_1 \left(1 + \frac{1}{\alpha}\right) = \prod_{k=1}^{2-1} \left(1 + \frac{1}{k\alpha}\right).$$

Supposons que $u_n = \prod_{k=1}^{n-1} \left(1 + \frac{1}{k\alpha}\right)$. Alors, par la question précédente,

$$u_{n+1} = \frac{n\alpha - 1}{n\alpha} u_n = \left(1 - \frac{1}{n\alpha}\right) u_n = u_1 \prod_{k=1}^{n-1} \left(1 + \frac{1}{k\alpha}\right) \times \left(1 - \frac{1}{n\alpha}\right) = u_1 \prod_{k=1}^n \left(1 + \frac{1}{k\alpha}\right).$$

Méthode

Toujours commencer l'étude d'une intégrale impropre par l'étude du domaine de continuité de l'intégrande afin de repérer les éventuels problèmes de convergence de l'intégrale.

⚠ Attention !

Il est important de vérifier que $1 + t^\alpha$ est plus grand que 1. En effet, la suite $(q^n)_n$ est croissante si $q \geq 1$ et décroissante si $0 \leq q \leq 1$!

Méthode

Quand on a peu d'informations sur une suite et qu'on doit prouver qu'elle converge, il faut le plus souvent aller chercher du côté du théorème de la limite monotone.

IPP

Ici la fonction intégrée est continue en 0, donc pas la peine de prendre un segment de la forme $[A, B]$, autant inclure 0 dans le segment de départ, et conclure en prenant une seule limite.

Précaution

On peut utiliser la linéarité de l'intégrale pour séparer l'intégrale en deux, car on sait que les deux morceaux convergent.

Donc la formule donnée dans l'énoncé est encore valable au rang $n + 1$, et donc par le principe de récurrence,

$$\forall n \geq 2, u_n = u_1 \prod_{k=1}^{n-1} \left(1 - \frac{1}{k\alpha}\right).$$

3. On a

$$\ln(u_n) = \ln\left(u_1 \prod_{k=1}^{n-1} \left(1 - \frac{1}{k\alpha}\right)\right) = \ln(u_1) + \sum_{k=1}^{n-1} \ln\left(1 - \frac{1}{k\alpha}\right).$$

Mais on a

$$\ln\left(1 - \frac{1}{k\alpha}\right) \underset{k \rightarrow +\infty}{\sim} -\frac{1}{k\alpha}.$$

Par critère de comparaison pour les séries à termes de signe constant¹, on en déduit que la série de terme général $\ln\left(1 - \frac{1}{k\alpha}\right)$ diverge.

$$^1 -\frac{1}{k\alpha} \leq 0.$$

Or il s'agit d'une série à termes négatifs, donc la suite de ses sommes partielles est décroissante. Mais une suite décroissante et divergente tend nécessairement vers $-\infty$.

On en déduit que

$$\lim_{n \rightarrow +\infty} \sum_{k=1}^{n-1} \ln\left(1 - \frac{1}{k\alpha}\right) = -\infty \text{ et donc } \lim_{n \rightarrow +\infty} \ln(u_n) = -\infty.$$

Par passage à l'exponentielle, il vient alors $\lim_{n \rightarrow +\infty} u_n = 0$.

4.a. Utilisant le résultat de la question 2.a, on a

$$\begin{aligned} S_n &= \sum_{k=1}^n u_k = \alpha \sum_{k=1}^n k(u_k - u_{k+1}) \\ &= \alpha \sum_{k=1}^n k u_k - \alpha \sum_{k=1}^n k u_{k+1} = \alpha \sum_{k=1}^n k u_k - \alpha \sum_{i=2}^{n+1} (i-1) u_i \\ &= \alpha \sum_{k=1}^n k u_k - \alpha \sum_{k=2}^{n+1} k u_k + \alpha \sum_{k=2}^{n+1} u_k \\ &= \alpha (u_1 - (n+1)u_{n+1} + (S_n - u_1 + u_{n+1})) \\ &= \alpha (S_n - n u_{n+1}) \end{aligned}$$

Chgt d'indice

$$i = k + 1 \Leftrightarrow k = i - 1.$$

Et donc

$$S_n(1 - \alpha) = -\alpha n u_{n+1} \Leftrightarrow S_n = \frac{n\alpha}{\alpha - 1} u_{n+1}.$$

Alternative : notons qu'il était également possible de prouver la formule demandée par récurrence sur $n \in \mathbf{N}^*$.

4.b. On en déduit, à l'aide de la question 2.b, que pour $n \geq 2$,

$$\begin{aligned} \ln(S_n) &= \ln(n\alpha) - \ln(\alpha - 1) + \ln(u_{n+1}) \\ &= \ln(u_1) + \ln(n\alpha) - \ln(\alpha - 1) + \sum_{k=1}^n \ln\left(1 - \frac{1}{k\alpha}\right) \\ &= \ln(u_1) + \ln(n) - \ln\left(1 - \frac{1}{\alpha}\right) + \sum_{k=1}^n \ln\left(1 - \frac{1}{k\alpha}\right) \\ &= \ln(u_1) + \ln(n) + \sum_{k=2}^n \ln\left(1 - \frac{1}{k\alpha}\right) \end{aligned}$$

Il reste à constater que

$$\sum_{k=2}^n \ln\left(1 - \frac{1}{k}\right) = \sum_{k=2}^n \ln\left(\frac{k-1}{k}\right) = \sum_{k=2}^n \ln(k-1) - \sum_{k=1}^n \ln(k) = \ln(1) - \ln(n).$$

Et donc

$$\ln(S_n) = \ln(u_1) + \sum_{k=2}^n \left[\ln\left(1 - \frac{1}{k\alpha}\right) - \ln\left(1 - \frac{1}{k}\right) \right].$$

4.c. On a

$$\ln\left(1 - \frac{1}{k\alpha}\right) - \ln\left(1 - \frac{1}{k}\right) = -\frac{1}{k\alpha} + o\left(\frac{1}{k\alpha}\right) + \frac{1}{k} + o\left(\frac{1}{k}\right) = \frac{1}{k} - \frac{1}{k\alpha} + o\left(\frac{1}{k}\right) = \frac{\alpha - 1}{k\alpha} + o\left(\frac{1}{k}\right).$$

On en déduit que

$$\ln\left(1 - \frac{1}{k\alpha}\right) - \ln\left(1 - \frac{1}{k}\right) \underset{k \rightarrow +\infty}{\sim} \frac{\alpha - 1}{k\alpha}.$$

4.d. Grâce à l'équivalent obtenu à la question précédente, par critère de comparaison pour les séries de signe constant, on en déduit que la série de terme général $\ln\left(1 - \frac{1}{k\alpha}\right) - \ln\left(1 - \frac{1}{k}\right)$ diverge. Puisqu'il s'agit d'une série à termes positifs, la suite de ses sommes partielles tend vers $+\infty$, et donc $\ln(S_n) \xrightarrow[n \rightarrow +\infty]{} +\infty$.

On en déduit que $S_n \xrightarrow[n \rightarrow +\infty]{} +\infty$, et donc $\sum u_n$ diverge.

5. Il s'agit d'utiliser le résultat de la question 2.b, et notant que

$$u_0 = \int_0^{+\infty} \frac{dt}{1+t^2} = \lim_{A \rightarrow +\infty} \int_0^A \frac{dt}{1+t^2} = \lim_{A \rightarrow +\infty} \text{Arctan}(A) = \frac{\pi}{2}.$$

```

1 function y = u(n)
2   y = %pi/2
3   for k=1 :n-1
4     y = y*(1-1/(2*k)) ;
5   end
6 endfunction

```

EXERCICE 2 (ÉCRICOME 2010)

1. Convergence de la suite $(u_n)_{n \geq 1}$

1.a. Pour $t \in [0, 1[$, on a $1 + t + \dots + t^{n-1} = \frac{1-t^n}{1-t}$ et donc

$$\frac{1}{1+t+\dots+t^{n-1}} - (1-t) = \frac{1-t}{1-t^n} - \frac{(1-t)(1-t^n)}{1-t^n} = \frac{(1-t)(1-1+t^n)}{1-t^n} = \frac{(1-t)t^n}{1-t^n}.$$

On en déduit que

$$u_n - \int_0^1 (1-t) dt = \int_0^1 \frac{(1-t)t^n}{1-t^n} dt.$$

Mais $\int_0^1 (1-t) dt = \left[t - \frac{t^2}{2} \right]_0^1 = \frac{1}{2}$, de sorte que

$$u_n - \int_0^1 (1-t) dt = u_n - \frac{1}{2}.$$

Pour tout $t \in [0, 1[$, on a $t \geq t^n$ et donc $1-t \leq 1-t^n$, de sorte que $\frac{1-t}{1-t^n} \leq 1$.

En particulier, par croissance de l'intégrale,

$$\int_0^1 \frac{(1-t)t^n}{1-t^n} dt \leq \int_0^1 t^n dt = \frac{1}{n+1} \text{ et donc } u_n - \frac{1}{2} \leq \frac{1}{n+1}.$$

Négligeabilité

Puisque α est une constante, il n'y a pas de différence entre $o\left(\frac{1}{k}\right)$ et $o\left(\frac{1}{k\alpha}\right)$.

Notons que cet équivalent est positif.

Remarque

Cette dernière intégrale est impropre en 1, mais converge car différence de deux intégrales convergentes.

Puissance n

Attention au fait que pour $t \geq 1$, on a $t \leq t^n$, mais que pour $t \in [0, 1[$, on a au contraire $t^n \leq t$.

D'autre part, la fonction $t \mapsto \frac{(1-t)t^n}{1-t^n}$ étant positive sur $[0, 1[$, on a $\int_0^1 \frac{(1-t)t^n}{1-t^n} dt \geq 0$
 et donc $0 \leq u_n - \frac{1}{2}$.

Au final, il vient donc

$$0 \leq u_n - \frac{1}{2} \leq \frac{1}{n+1}.$$

Par le théorème des gendarmes, on a donc $\lim_{n \rightarrow +\infty} u_n - \frac{1}{2} = 0$ et donc $\lim_{n \rightarrow +\infty} u_n = \frac{1}{2}$.

1.b. D'après ce qui précède, on a donc $u_n - \frac{1}{2} = \int_0^1 \frac{(1-t)t^n}{1-t^n} dt$.

Le changement de variable $u = t^n$ est de classe \mathcal{C}^1 sur $[0, 1[$, et strictement croissant, donc légitime. On a alors $u = t^n \Leftrightarrow t = u^{1/n}$ et donc $dt = \frac{1}{n} u^{1/n-1} du$. Il vient donc²

$$u_n - \frac{1}{2} = \int_0^1 \frac{(1-t)t^n}{1-t^n} dt = \int_0^1 \frac{(1-u^{1/n})u}{1-u} \frac{1}{n} u^{1/n-1} du = \frac{1}{n} \int_0^1 \frac{u^{1/n} - u^{2/n}}{1-u} du = \boxed{\frac{v_n}{n}}.$$

² L'intégrale de départ est convergente, donc l'intégrale obtenue après changement de variable le sera automatiquement.

2. Résultats intermédiaires.

2.a. Lorsque $x \rightarrow 1$, on a $x-1 \rightarrow 0$ et donc

$$\ln(x) = \ln(1 + (x-1)) \underset{x \rightarrow 1}{\sim} x-1.$$

Et donc pour tout $k \geq 1$, $(\ln x)^k \underset{x \rightarrow 1}{\sim} (x-1)^k$.

On en déduit que

$$\frac{(\ln x)^k}{x-1} \underset{x \rightarrow 1}{\sim} \frac{(x-1)^k}{x-1} = (x-1)^{k-1}.$$

Par conséquent,

$$\frac{(\ln x)^k}{x-1} \underset{x \rightarrow 1}{\sim} (x-1)^{k-1} \xrightarrow{x \rightarrow 1} \begin{cases} 1 & \text{si } k = 1 \\ 0 & \text{si } k > 1 \end{cases}$$

2.b. La fonction $x \mapsto \frac{(\ln x)^k}{x-1}$ est continue sur $]0, 1[$.

Par la question précédente, elle est prolongeable par continuité en 1, donc $\int_{1/2}^1 \frac{(\ln x)^k}{x-1} dx$ converge.

Au voisinage de 0, on a $\frac{(\ln x)^k}{x-1} \underset{x \rightarrow 0}{\sim} -(\ln x)^k$.

Or, par croissances comparées, on a $\sqrt{x}(\ln x)^k \xrightarrow{x \rightarrow 0^+} 0$, de sorte que $\ln(x)^k = \underset{x \rightarrow 0^+}{o} \left(\frac{1}{\sqrt{x}} \right)$.

Puisque $\int_0^{1/2} \frac{1}{\sqrt{x}} dx$ converge, il en est de même de $\int_0^{1/2} (\ln x)^k dx$. Et alors, par le critère des équivalents pour les fonctions de signe constant $\int_0^{1/2} \frac{(\ln x)^k}{x-1} dx$ converge également.

Et donc $\int_0^1 \frac{(\ln x)^k}{x-1} dx$ converge.

Signe
 k étant fixé, la fonction $x \mapsto \frac{(\ln x)^k}{x-1}$ est de signe constant sur $]0, 1[$, bien que ce signe dépende de la parité de k .

2.c. La fonction f est de classe \mathcal{C}^2 sur $[x, 0]$, avec

$$f'(t) = e^t - 2e^{2t} \text{ et } f''(x) = e^t - 4e^{2t} = e^t(1 - 4e^t).$$

En particulier, pour tout $t \in [x, 0]$, on a

$$|f''(t)| \leq |e^t| |1 - 4e^t| \leq |1 - 4e^t|.$$

Mais la fonction $t \mapsto 1 - 4e^t$ est strictement décroissante sur \mathbf{R}_- , et on a $\lim_{t \rightarrow -\infty} 1 - 4e^t = 1$ et $1 - 4e^0 = -3$.

Majorant
 Rappelons que pour appliquer l'inégalité de Taylor-Lagrange à l'ordre n entre a et b , il faut disposer d'un majorant de $|f^{(n+1)}|$ sur le segment d'extrémités a et b , et pas nécessairement sur \mathbf{R} tout entier.

Donc pour tout $t \in \mathbb{R}_-$, $-3 \leq 1 - 4e^t \leq 1 \leq 3$, de sorte que $|1 - 4e^t| \leq 3$.
 D'après l'inégalité de Taylor-Lagrange à l'ordre 1 entre 0 et x , il vient

$$|f(x) - f'(0)x| \leq \frac{3|x|^2}{2!} \Leftrightarrow |e^x - e^{2x} + x| \leq \frac{3x^2}{2}.$$

3. Application

3.a. Pour tout $u \in]0, 1]$, on a $\frac{1}{n} \ln(u) \leq 0$, et donc, d'après la question 2.c, appliquée avec $x = \frac{1}{n} \ln u$,

$$\left| e^{\frac{1}{n} \ln u} - e^{\frac{2}{n} \ln u} + \frac{1}{n} \ln u \right| \leq \frac{3}{2n^2} (\ln u)^2 \Leftrightarrow \left| u^{1/n} - u^{2/n} + \frac{1}{n} \ln u \right| \leq \frac{3}{2n} (\ln u)^2.$$

En divisant par $1 - u$, il vient

$$\forall u \in]0, 1[, \left| \frac{u^{1/n} - u^{2/n}}{1 - u} + \frac{1}{n} \frac{\ln u}{1 - u} \right| \leq \frac{3}{2n^2} \frac{(\ln u)^2}{1 - u}.$$

Par croissance de l'intégrale³, on a alors

$$\int_0^1 \left| \frac{u^{1/n} - u^{2/n}}{1 - u} + \frac{1}{n} \frac{\ln u}{1 - u} \right| du \leq \frac{3}{2n^2} \int_0^1 \frac{(\ln u)^2}{1 - u} du.$$

Et enfin, d'après l'inégalité triangulaire pour les intégrales,

$$\left| \int_0^1 \frac{u^{1/n} - u^{2/n}}{1 - u} du + \frac{1}{n} \int_0^1 \frac{\ln u}{1 - u} du \right| \leq \int_0^1 \left| \frac{u^{1/n} - u^{2/n}}{1 - u} + \frac{1}{n} \frac{\ln u}{1 - u} \right| du \leq \frac{3}{2n^2} \int_0^1 \frac{(\ln u)^2}{1 - u} du.$$

On en déduit donc que

$$\left| v_n + \frac{1}{n} \int_0^1 \frac{\ln u}{1 - u} du \right| \leq \frac{3}{2n^2} \int_0^1 \frac{(\ln u)^2}{1 - u} du.$$

3.b. Notons $J = \int_0^1 \frac{(\ln u)^2}{1 - u} du$, qui est une constante ne dépendant pas de n .

On a alors $\left| v_n + \frac{1}{n} I \right| \leq \frac{3}{2n^2} J$.

En particulier, $\frac{n}{I} \left| v_n + \frac{1}{n} I \right| \leq \frac{3}{2n} \frac{J}{I} \xrightarrow{n \rightarrow +\infty} 0$, de sorte que

$$v_n + \frac{1}{n} I = \underset{n \rightarrow +\infty}{o} \left(\frac{I}{n} \right).$$

Et donc

$$v_n = -\frac{1}{n} I + \left(v_n + \frac{1}{n} I \right) = -\frac{I}{n} + \underset{n \rightarrow +\infty}{o} \left(\frac{I}{n} \right) \underset{n \rightarrow +\infty}{\sim} \left[-\frac{I}{n} \right].$$

Puisque $u_n - \frac{1}{2} = \frac{v_n}{n}$, on en déduit que

$$u_n - \frac{1}{2} \underset{n \rightarrow +\infty}{\sim} -\frac{I}{n^2} = -\frac{1}{n^2} \int_0^1 \frac{\ln u}{1 - u} du.$$

³ Il a été prouvé précédemment que toutes les intégrales en jeu convergent.

o/~

Par définition d'un équivalent, on a

$$u_n \sim v_n$$

si et seulement si

$$u_n = v_n + o(v_n).$$

DEVOIR MAISON 4 BIS (À RENDRE LE 11.10.18)

On pose

$$E_0 = \{f \in \mathcal{C}^0(\mathbf{R}, \mathbf{R}), \text{ bornée sur } \mathbf{R}\}.$$

Si $f \in E_0$, on notera $N_0(f) = \sup_{x \in \mathbf{R}} |f(x)|$.

Partie I : Quelques propriétés de la fonction Arctan.

1. Justifier que, pour tout $(x, y) \in \mathbf{R}^2$, $|\arctan x - \arctan y| \leq |x - y|$.
2. Montrer que, pour tout $x \in \mathbf{R}_+^*$, $\arctan x + \arctan \frac{1}{x} = \frac{\pi}{2}$.

Si $f \in E_0$, on définit, sous réserve d'existence, $\Phi(f) : x \in \mathbf{R} \mapsto \int_0^{+\infty} \arctan(tx) \frac{f(t)}{1+t^2} dt$.

L'objectif du problème est d'obtenir quelques propriétés de $\Phi(f)$ et de Φ .

Partie II : Premières propriétés de $\Phi(f)$ et de Φ .

3. Vérifier que E_0 est un sous-espace vectoriel de $\mathcal{C}^0(\mathbf{R}, \mathbf{R})$.
4. Soit $f \in E_0$. Montrer que, pour tout $x \in \mathbf{R}$, $\int_0^{+\infty} \arctan(tx) \frac{f(t)}{1+t^2} dt$ est absolument convergente.
5. Soit $f \in E_0$, montrer que $\Phi(f)$ est bornée et que

$$N_0(\Phi(f)) \leq \frac{\pi^2}{4} N_0(f).$$

6. Continuité de $\Phi(f)$ pour $f \in E_0$.

Dans cette question, f désigne un élément de E_0 et x un réel.

- a. Soit A un réel strictement positif et $h \in \mathbf{R}^*$. Vérifier que

$$|[\Phi(f)](x+h) - [\Phi(f)](x)| \leq N_0(f) \left(\int_0^A \frac{|\arctan(t(x+h)) - \arctan(tx)|}{1+t^2} dt + \int_A^{+\infty} \frac{|\arctan(t(x+h)) - \arctan(tx)|}{1+t^2} dt \right).$$

- b. En déduire que pour tout $h \in \mathbf{R}^*$, pour tout $A > 0$

$$|[\Phi(f)](x+h) - [\Phi(f)](x)| \leq N_0(f) \left(|h| \int_0^A \frac{t}{1+t^2} dt + \pi \int_A^{+\infty} \frac{dt}{1+t^2} \right).$$

- c. Soit $h \in \mathbf{R}^*$. En choisissant $A = \frac{1}{|h|}$, établir que

$$|[\Phi(f)](x+h) - [\Phi(f)](x)| \leq |h| \frac{N_0(f)}{2} \ln \left(1 + \frac{1}{h^2} \right) + \pi N_0(f) \arctan |h|.$$

- d. Montrer alors que $\Phi(f)$ est continue sur \mathbf{R} .
- e. En déduire que $\Phi : f \in E_0 \mapsto \Phi(f)$ est un endomorphisme de E_0 .

Partie III : Étude d'un exemple

Dans cette partie, on s'intéresse à l'application $g : \mathbf{R} \rightarrow \mathbf{R}, x \mapsto g(x) = \int_0^{+\infty} \frac{\arctan(tx)}{1+t^2} dt$.

g est ainsi l'image par Φ de la fonction constante égale à 1.

7. Vérifier que g est impaire.
8. **Dérivabilité de g sur $]0, +\infty[$.**

Soit x un réel strictement positif.

- a. Vérifier que, pour tout $u \in \mathbf{R}$, $|\arctan''(u)| \leq \frac{1}{1+u^2}$.
- b. Soit a, b deux réels distincts et I le segment d'extrémités a et b . Montrer que

$$\left| \arctan b - \arctan a - \frac{b-a}{1+a^2} \right| \leq \frac{(b-a)^2}{2} \max_{u \in I} \left(\frac{1}{1+u^2} \right).$$

c. Soit $h \in]-\frac{x}{2}, \frac{x}{2}[$ et t un réel positif. Établir

$$\left| \arctan(t(x+h)) - \arctan(tx) - \frac{th}{1+t^2x^2} \right| \leq \frac{t^2h^2}{2} \frac{1}{1+\frac{t^2x^2}{4}}.$$

d. Montrer alors que, pour tout $h \in]-\frac{x}{2}, \frac{x}{2}[$,

$$\left| g(x+h) - g(x) - h \int_0^{+\infty} \frac{t}{(1+t^2x^2)(1+t^2)} dt \right| \leq 2h^2 \int_0^{+\infty} \frac{t^2}{(4+t^2x^2)(1+t^2)} dt.$$

e. En déduire que g est dérivable sur $]0; +\infty[$ et justifier que, pour tout $x > 0$,

$$g'(x) = \int_0^{+\infty} \frac{t}{(1+t^2x^2)(1+t^2)} dt.$$

f. g est-elle dérivable sur $] -\infty, 0[$? Si oui, que vaut $g'(x)$ pour $x < 0$?

9. Calcul de $g'(x)$ pour $x > 0$.

a. Déterminer $g'(1)$.

b. Pour tout $x \in]0, 1[\cup]1; +\infty[$, chercher des expressions $A(x)$ et $B(x)$, indépendantes de t telles que, pour tout $t \in \mathbf{R}$,

$$\frac{t}{(1+t^2x^2)(1+t^2)} = A(x) \frac{t}{1+t^2x^2} + B(x) \frac{t}{1+t^2}.$$

c. En déduire que, pour tout $x \in]0, 1[\cup]1; +\infty[$, $g'(x) = \frac{\ln x}{x^2 - 1}$.

d. g est-elle de classe \mathcal{C}^1 sur $]0; +\infty[$?

10. Une nouvelle expression de $g(x)$ pour $x > 0$.

a. Justifier, pour tout $x > 0$, la convergence de l'intégrale $\int_0^x \frac{\ln t}{t^2 - 1} dt$.

b. Montrer que, pour tout $x > 0$,

$$\int_0^{+\infty} \frac{\arctan(tx)}{1+t^2} dt = \int_0^x \frac{\ln t}{t^2 - 1} dt.$$

11. Étude de la limite de g en $+\infty$.

a. Démontrer que, pour tout $x > 0$, $g(x) = \frac{\pi^2}{4} - \int_0^{+\infty} \arctan\left(\frac{1}{tx}\right) \frac{1}{1+t^2} dt$.

b. Écrire, pour tout $x > 0$,

$$\int_0^{+\infty} \arctan\left(\frac{1}{tx}\right) \frac{1}{1+t^2} dt = \int_0^{1/\sqrt{x}} \arctan\left(\frac{1}{tx}\right) \frac{1}{1+t^2} dt + \int_{1/\sqrt{x}}^{+\infty} \arctan\left(\frac{1}{tx}\right) \frac{1}{1+t^2} dt$$

et montrer alors que

$$\left| \int_0^{+\infty} \arctan\left(\frac{1}{tx}\right) \frac{1}{1+t^2} dt \right| \leq \frac{\pi}{2} \frac{1}{\sqrt{x}} + \frac{1}{x} \int_{1/\sqrt{x}}^{+\infty} \frac{1}{t} \frac{1}{1+t^2} dt \leq \frac{\pi}{\sqrt{x}}.$$

c. Déterminer alors $\lim_{x \rightarrow +\infty} g(x)$.

12. Application au calcul de $\sum_{n=0}^{+\infty} \frac{1}{(2n+1)^2}$.

a. Montrer que $\int_0^{+\infty} \frac{\ln t}{1-t^2} dt$ est convergente et calculer sa valeur à l'aide des questions précédentes.

b. Vérifier que $\int_0^{+\infty} \frac{\ln t}{1-t^2} dt = 2 \int_0^1 \frac{\ln t}{1-t^2} dt$.

Indication : on pourra, en le justifiant, utiliser le changement de variable $u = \frac{1}{t}$.

c. Soit $n \in \mathbf{N}$. Démontrer que

$$\int_0^1 \frac{\ln t}{1-t^2} dt = \sum_{k=0}^n \int_0^1 t^{2k} \ln t dt + \int_0^1 \frac{\ln t}{1-t^2} t^{2n+2} dt$$

(on justifiera l'existence des intégrales introduites).

d. Pour tout $n \in \mathbf{N}$, calculer $\int_0^1 t^{2n} \ln t dt$.

e. Montrer que $\lim_{n \rightarrow +\infty} \int_0^1 \frac{\ln t}{1-t^2} t^{2n+2} dt = 0$.

f. Donner alors la valeur de $\sum_{n=0}^{+\infty} \frac{1}{(2n+1)^2}$. En déduire celle de $\sum_{n=1}^{+\infty} \frac{1}{n^2}$.

CORRECTION DU DEVOIR MAISON 4

Partie I : Quelques propriétés de la fonction arctan

1. La dérivée de arctan est la fonction $t \mapsto \frac{1}{1+t^2}$, qui est bornée par 1. D'après l'inégalité des accroissements finis (qui s'applique puisque arctan est dérivable sur \mathbf{R}), on a alors, pour tous $x, y \in \mathbf{R}$,

$$|\arctan x - \arctan y| \leq |x - y|.$$

2. La fonction $\varphi : x \mapsto \arctan x + \arctan \frac{1}{x}$ est dérivable sur \mathbf{R}_+^* par composition de fonctions qui le sont, et on a

$$\varphi'(x) = \frac{1}{1+x^2} - \frac{1}{1+x^2} \frac{1}{1+\frac{1}{x^2}} = \frac{1}{x^2} - \frac{1}{x^2+1} = 0.$$

Par conséquent, la fonction est constante sur \mathbf{R}_+^* . Notons alors λ sa valeur. En prenant la limite lorsque $x \rightarrow 0^+$, on obtient

$$\lambda = \lim_{x \rightarrow 0^+} \varphi(x) = \lim_{x \rightarrow 0^+} \arctan x + \lim_{x \rightarrow 0^+} \arctan \frac{1}{x} = 0 + \frac{\pi}{2} = \frac{\pi}{2}.$$

Et donc

$$\forall x \in \mathbf{R}_+^*, \arctan x + \arctan \frac{1}{x} = \frac{\pi}{2}.$$

Partie II : Premières propriétés de $\Phi(f)$ et de Φ .

3. Il est évident que la fonction nulle est continue et bornée, et donc dans E_0 . Soient $f, g \in E_0$. Alors f et g sont continues, et

$$\forall x \in \mathbf{R}, |f(x)| \leq N_0(f) \text{ et } |g(x)| \leq N_0(g).$$

Soit alors $\lambda \in \mathbf{R}$. La fonction $\lambda f + g$ est encore continue et

$$\forall x \in \mathbf{R}, |(\lambda f + g)(x)| = |\lambda f(x) + g(x)| \leq |\lambda| |f(x)| + |g(x)| \leq |\lambda| N_0(f) + N_0(g).$$

Donc la fonction $\lambda f + g$ est bornée (par $|\lambda| N_0(f) + N_0(g)$), et ainsi $\lambda f + g \in E_0$.

Alors, E_0 est un sous-espace vectoriel de $\mathcal{C}^0(\mathbf{R}, \mathbf{R})$.

4. Soit $f \in E_0$ et soit $x \in \mathbf{R}$ fixé. Alors, pour tout $t \in [0, +\infty[$, on a

$$\left| \arctan(tx) \frac{f(t)}{1+t^2} \right| \leq \frac{\pi}{2} \frac{N_0(f)}{1+t^2}.$$

Or, $\int_0^{+\infty} \frac{dt}{1+t^2}$ est convergente, et donc il en est de même de $\int_0^{+\infty} \left| \arctan(tx) \frac{f(t)}{1+t^2} \right| dt$.

Par conséquent

$$\int_0^{+\infty} \arctan(tx) \frac{f(t)}{1+t^2} dt \text{ est absolument convergente.}$$

5. Soit $x \in \mathbf{R}$. Alors

$$\begin{aligned} |\Phi(f)(x)| &= \left| \int_0^{+\infty} \arctan(tx) \frac{f(t)}{1+t^2} dt \right| \\ &\leq \int_0^{+\infty} \left| \arctan(tx) \frac{f(t)}{1+t^2} \right| dt \\ &\leq \frac{\pi}{2} N_0(f) \int_0^{+\infty} \frac{dt}{1+t^2} \leq \frac{\pi^2}{4} N_0(f). \end{aligned}$$

Ceci prouve donc que $\Phi(f)$ est bornée et que

$$N_0(\Phi(f)) = \sup_{x \in \mathbf{R}} |\Phi(f)(x)| \leq \frac{\pi^2}{4} N_0(f).$$

Classique

Le même résultat est en fait valable sur \mathbf{R}_-^* : φ y est dérivable, de dérivée nulle, donc constante. L'étude de la limite en 0^- ou en $-\infty$ montre alors que $\forall x \in \mathbf{R}_-^*$,

$$\arctan x + \arctan \frac{1}{x} = -\frac{\pi}{2}.$$

Ceci nous rappelle qu'il faut être vigilant : si une fonction est de dérivée nulle, elle est constante sur **chaque intervalle** de son ensemble de définition, et non nécessairement sur tout son ensemble de définition (ici φ' est nulle sur \mathbf{R}^* et pourtant φ n'est pas constante sur \mathbf{R}^* .)

Explication

$N_0(f)$ est la borne supérieure des valeurs prises par f . C'est le plus petit des majorants de f .

Remarque

Ceci prouve au passage que $N_0(\lambda f + g) \leq |\lambda| N_0(f) + N_0(g)$.

Inégalité triangulaire.

Borne supérieure

Si pour tout x , $|f(x)| \leq M$, alors

$$\sup_{x \in \mathbf{R}} |f(x)| \leq M$$

car la borne supérieure est le **plus petit** des majorants.

6. Continuité de $\Phi(f)$ pour $f \in E_0$.

6.a. Par définition de Φ , on a

$$\begin{aligned} [\Phi(f)](x+h) - [\Phi(f)](x) &= \int_0^{+\infty} \arctan(t(x+h)) \frac{f(t)}{1+t^2} dt - \int_0^{+\infty} \arctan(tx) \frac{f(t)}{1+t^2} dt \\ &= \int_0^{+\infty} (\arctan(t(x+h)) - \arctan(tx)) \frac{f(t)}{1+t^2} dt. \end{aligned}$$

En prenant la valeur absolue, il vient alors

$$\begin{aligned} |[\Phi(f)](x+h) - [\Phi(f)](x)| &= \left| \int_0^{+\infty} (\arctan(t(x+h)) - \arctan(tx)) \frac{f(t)}{1+t^2} dt \right| \\ &\leq \int_0^{+\infty} \left| (\arctan(t(x+h)) - \arctan(tx)) \frac{f(t)}{1+t^2} \right| dt \\ &\leq \int_0^{+\infty} |\arctan(t(x+h)) - \arctan(tx)| \frac{N_0(f)}{1+t^2} dt \\ &\leq N_0(f) \left(\int_0^A \frac{|\arctan(t(x+h)) - \arctan(tx)|}{1+t^2} dt + \int_A^{+\infty} \frac{|\arctan(t(x+h)) - \arctan(tx)|}{1+t^2} dt \right). \end{aligned}$$

Précision

Cette intégrale est bien convergente car on a montré à la question 4 que les intégrales $\Phi(f)(x+h)$ et $\Phi(f)(x)$ sont **absolument** convergentes.

6.b. Soit $A > 0$ fixé et $h \in \mathbf{R}^*$ fixé. Majorons chacune des deux intégrales obtenues à la question précédente.

En utilisant le résultat de la question 1, on a

$$\forall t \in [0, A], |\arctan(t(x+h)) - \arctan(tx)| \leq |t(x+h) - tx| \leq |th| \leq t|h|.$$

Par croissance de l'intégrale, il vient

$$\int_0^A \frac{|\arctan(t(x+h)) - \arctan(tx)|}{1+t^2} dt \leq \int_0^A |h| \frac{t}{1+t^2} dt.$$

De même, pour tout $t \geq A$, on a

$$|\arctan(t(x+h)) - \arctan(tx)| \leq |\arctan(t(x+h))| + |\arctan(tx)| \leq \frac{\pi}{2} + \frac{\pi}{2} \leq \pi.$$

Par croissance de l'intégrale, il vient (la seconde intégrale est bien convergente par comparaison à une intégrale de Riemann)

$$\int_A^{+\infty} \frac{|\arctan(t(x+h)) - \arctan(tx)|}{1+t^2} dt \leq \pi \int_A^{+\infty} \frac{dt}{1+t^2}.$$

En utilisant le résultat de la question 6.a, on a alors

$$|[\Phi(f)](x+h) - [\Phi(f)](x)| \leq N_0(f) \left(|h| \int_0^A \frac{t}{1+t^2} dt + \pi \int_A^{+\infty} \frac{dt}{1+t^2} \right).$$

6.c. Comme indiqué dans l'énoncé, fixons h et soit $A = \frac{1}{|h|}$. Alors

$$\int_0^A \frac{t}{1+t^2} dt = \int_0^{1/|h|} \frac{t}{1+t^2} dt = \left[\frac{1}{2} \ln(1+t^2) \right]_0^{1/|h|} = \frac{1}{2} \ln \left(1 + \frac{1}{|h|^2} \right) = \frac{1}{2} \ln \left(1 + \frac{1}{t^2} \right).$$

De même, on a

$$\int_A^{+\infty} \frac{dt}{1+t^2} = \int_{1/|h|}^{+\infty} \frac{dt}{1+t^2}.$$

Pour calculer cette intégrale, soit $B > 0$. Alors

$$\begin{aligned} \int_{1/|h|}^B \frac{dt}{1+t^2} &= [\arctan(t)]_{1/|h|}^B \\ &= \arctan(B) - \arctan \frac{1}{|h|} \end{aligned}$$

$$\xrightarrow{B \rightarrow +\infty} \frac{\pi}{2} - \arctan \frac{1}{|h|}.$$

Donc $\int_{1/|h|}^{+\infty} \frac{dt}{1+t^2}$ converge et vaut $\frac{\pi}{2} - \arctan \frac{1}{|h|}$.

Utilisons à présent le résultat de la question 2, qui s'applique car $|h| \in \mathbf{R}_+^*$:

$$\frac{\pi}{2} - \arctan \frac{1}{|h|} = \arctan |h|.$$

On en déduit donc que

$$|[\Phi(f)](x+h) - [\Phi(f)](x)| \leq |h| \frac{N_0(f)}{2} \ln \left(1 + \frac{1}{h^2} \right) + \pi N_0(f) \arctan |h|.$$

6.d. Soit $x \in \mathbf{R}$ fixé. Montrons que $\Phi(f)$ est continue en x .

Lorsque $|h| \leq 1$, on a

$$0 \leq |h| \ln \left(1 + \frac{1}{h^2} \right) \leq |h| \ln \left(\frac{1+h^2}{h^2} \right) \leq |h| \ln \left(\frac{2}{h^2} \right) \leq |h| (\ln(2) - \ln(h^2)) \leq |h| \ln(2) - 2|h| \ln(|h|)$$

Par croissance comparée, ces deux termes tendent tous deux vers 0 et donc par le théorème des gendarmes

$$\lim_{h \rightarrow 0} |h| \ln \left(1 + \frac{1}{h^2} \right) = 0.$$

De même, par continuité de la fonction arctan, on a

$$\lim_{h \rightarrow 0} \arctan |h| = \arctan(0) = 0.$$

Donc en reprenant la majoration de la question 6.c, par le théorème des gendarmes, il vient

$$\lim_{h \rightarrow 0} [\Phi(f)](x+h) - [\Phi(f)](x) = 0 \Leftrightarrow \lim_{h \rightarrow 0} [\Phi(f)](x+h) = [\Phi(f)](x).$$

Ainsi, la fonction $\Phi(f)$ est continue en x , et ceci étant vrai pour tout $x \in \mathbf{R}$, Φ(f) est continue sur R.

6.e. Nous venons de prouver que si $f \in E_0$, alors $\Phi(f)$ est continue, et nous avons déjà prouvé à la question 5 qu'elle était bornée, donc $\Phi(f) \in E_0$.

Reste à voir que Φ est linéaire, ce qui est une conséquence immédiate de la linéarité de l'intégrale.

Ainsi, Φ(f) est un endomorphisme de E₀.

Partie III : Étude d'un exemple

7. Rappelons que arctan est impaire car la fonction tangente l'est. Donc si $x \in \mathbf{R}$, on a

$$g(-x) = \int_0^{+\infty} \frac{\arctan(-tx)}{1+t^2} dt = \int_0^{+\infty} -\frac{\arctan(tx)}{1+t^2} dt = -g(x).$$

Ainsi, g est impaire.

8. **Dérivabilité de g sur $]0; +\infty[$**

8.a. Nous savons déjà que $\arctan'(u) = \frac{1}{1+u^2}$ et donc $\arctan''(u) = \frac{-2u}{(1+u^2)^2}$.

Et donc $|\arctan''(u)| = \frac{2|u|}{(1+|u|^2)^2}$.

Mais pour tout $x \in \mathbf{R}$, on a

$$(1-x)^2 \geq 0 \Leftrightarrow 1 - 2x + x^2 \geq 0 \Leftrightarrow 2x \leq 1 + x^2.$$

En particulier, on a $2|u| \leq 1 + |u|^2$ et donc $0 \leq \frac{2|u|}{1+|u|^2} \leq 1$.

On en déduit alors que

$$\forall u \in \mathbf{R}, |\arctan''(u)| \leq \frac{1}{1+u^2}.$$

Rédaction

On sent bien ici que la limite de

$$|h| \ln \left(1 + \frac{1}{h^2} \right)$$

doit être nulle, non seulement parce que ça nous arrangerait bien, mais aussi par intuition des croissances comparées.

Toutefois, il ne s'agit pas d'une croissance comparée usuelle, et donc il faut réussir à s'y ramener, avec des majorations, des équivalents, etc.

D'ailleurs le rapport du jury souligne que cette question a souvent été traitée avec légèreté («par croissance comparée...»).

Rédaction

Ici, nous allons un peu vite pour la linéarité. Une copie qui traite un nombre significatif de questions pourra se permettre un tel raccourci, car la confiance du correcteur aura été gagnée grâce aux autres questions. En revanche si vous ne faite que quelques questions, mieux vaut prendre le temps d'écrire les détails.

Très classique

On a souvent besoin de cette inégalité dans les sujets de concours. C'est souvent guidé, parfois non. Mieux vaut donc s'assurer qu'on sait la retrouver.

- 8.b. La fonction \arctan est de classe \mathcal{C}^2 , donc on peut utiliser l'inégalité de Taylor-Lagrange en a à l'ordre 2 :

$$|\arctan b - \arctan a - \arctan'(a)(b-a)| \leq \frac{(b-a)^2}{2} \max_{u \in I} |\arctan''(u)|$$

Mais $\arctan'(a) = \frac{1}{1+a^2}$, et pour $u \in I$, $|\arctan''(u)| \leq \frac{1}{1+u^2}$. Donc il vient

$$\left| \arctan b - \arctan a - \frac{b-a}{1+a^2} \right| \leq \frac{(b-a)^2}{2} \max_{u \in I} \left(\frac{1}{1+u^2} \right).$$

- 8.c. Appliquons l'inégalité obtenue précédemment avec $a = t(x+h)$ et $b = tx$. Alors,

$$\left| \arctan(tx) - \arctan(t(x+h)) - \frac{tx - t(x+h)}{1+t^2x^2} \right| \leq \frac{(tx - t(x+h))^2}{2} \max_{u \in I} \left(\frac{1}{1+u^2} \right).$$

Ici, I désigne l'intervalle d'extrémités $t(x+h)$ et tx .

Mais on a $t(x - \frac{x}{2}) < t(x+h) < t(x + \frac{x}{2})$ soit $\frac{tx}{2} < t(x+h) < \frac{3tx}{2}$. Donc $I \subset \left] \frac{tx}{2}; \frac{3tx}{2} \right[$.

Puisque la fonction $u \mapsto \frac{1}{1+u^2}$ est décroissante sur I , on a alors

$$\max_{u \in I} \left(\frac{1}{1+u^2} \right) \leq \frac{1}{1 + \left(\frac{tx}{2}\right)^2}.$$

On en déduit donc que

$$\left| \arctan(tx) - \arctan(t(x+h)) - \frac{th}{1+t^2x^2} \right| \leq \frac{t^2h^2}{2} \frac{1}{1 + \frac{t^2x^2}{4}}.$$

- 8.d. Si l'on multiplie l'inégalité obtenue à la question 8.c par $\frac{1}{1+t^2} \geq 0$, il vient, pour tout $h \in]-\frac{x}{2}, \frac{x}{2}[$,

$$\left| \frac{\arctan(t(x+h))}{1+t^2} - \frac{\arctan(tx)}{1+t^2} - \frac{th}{(1+t^2)(1+t^2x^2)} \right| \leq \frac{t^2h^2}{2} \frac{1}{1+t^2} \frac{1}{1 + \frac{t^2x^2}{4}}.$$

Ceci étant valable pour tout t positif, on a alors, par croissance de l'intégrale

$$\int_0^{+\infty} \left| \frac{\arctan(t(x+h))}{1+t^2} - \frac{\arctan(tx)}{1+t^2} - \frac{th}{(1+t^2)(1+t^2x^2)} \right| dt \leq \int_0^{+\infty} \frac{t^2h^2}{2} \frac{1}{1+t^2} \frac{1}{1 + \frac{t^2x^2}{4}} dt.$$

Notons que l'intégrale de droite peut également s'écrire

$$\int_0^{+\infty} \frac{t^2h^2}{2} \frac{1}{1+t^2} \frac{1}{1 + \frac{t^2x^2}{4}} dt = 2h^2 \int_0^{+\infty} \frac{t^2}{(1+t^2)(4+t^2x^2)} dt.$$

Enfin, nous savons¹ que

$$\left| \int_0^{+\infty} \left(\frac{\arctan(t(x+h))}{1+t^2} - \frac{\arctan(tx)}{1+t^2} - \frac{th}{(1+t^2)(1+t^2x^2)} \right) dt \right| \leq \int_0^{+\infty} \left| \frac{\arctan(t(x+h))}{1+t^2} - \frac{\arctan(tx)}{1+t^2} - \frac{th}{(1+t^2)(1+t^2x^2)} \right| dt.$$

Mais par linéarité de l'intégrale², il vient

$$\int_0^{+\infty} \left(\frac{\arctan(t(x+h))}{1+t^2} - \frac{\arctan(tx)}{1+t^2} - \frac{th}{(1+t^2)(1+t^2x^2)} \right) dt = g(x+h) - g(x) - h \int_0^{+\infty} \frac{t}{(1+t^2)(1+t^2x^2)} dt.$$

Nous avons donc bien prouvé que

$$\left| g(x+h) - g(x) - h \int_0^{+\infty} \frac{t}{(1+t^2x^2)(1+t^2)} dt \right| \leq 2h^2 \int_0^{+\infty} \frac{t^2}{(4+t^2x^2)(1+t^2)} dt.$$

Méthode

Il faut être minutieux sur ce genre de question. En commençant par identifier chacun des termes (on comprend bien qu'il faut utiliser la question précédente), puis par calculer ce qui peut l'être facilement, on comprend vite qu'il va falloir montrer que le max de la question précédente est ici inférieur à $\frac{1}{1 + \frac{t^2x^2}{4}}$.

L'intégrale de droite converge bien par comparaison à une intégrale de Riemann (un équivalent suffit ici), donc l'intégrale de gauche converge par critère de comparaison pour les intégrales de fonctions positives.

¹ C'est l'inégalité triangulaire pour les intégrales absolument convergentes.

² Peut-être faut-il justifier que toutes ces intégrales convergent : pour les deux premières, c'est déjà fait, pour la troisième, c'est automatique car obtenue par différence d'intégrales convergentes.

8.e. Pour $h \neq 0$, l'inégalité précédente se réécrit

$$\left| \frac{g(x+h) - g(x)}{h} - \int_0^{+\infty} \frac{t}{(1+t^2x^2)(1+t^2)} dt \right| \leq 2h \int_0^{+\infty} \frac{t^2}{(4+t^2x^2)(1+t^2)} dt.$$

Or, lorsque $x \rightarrow 0$, le terme de droite tend vers 0. Donc par le théorème des gendarmes,

$$\begin{aligned} \lim_{h \rightarrow 0} \left| \frac{g(x+h) - g(x)}{h} - \int_0^{+\infty} \frac{t}{(1+t^2x^2)(1+t^2)} dt \right| &= 0 \\ \Leftrightarrow \lim_{h \rightarrow 0} \frac{g(x+h) - g(x)}{h} - \int_0^{+\infty} \frac{t}{(1+t^2x^2)(1+t^2)} dt &= 0. \end{aligned}$$

On en déduit que

$$\lim_{h \rightarrow 0} \frac{g(x+h) - g(x)}{h} = \int_0^{+\infty} \frac{t}{(1+t^2x^2)(1+t^2)} dt.$$

Ceci prouve que g est dérivable³ en x et que

$$g'(x) = \int_0^{+\infty} \frac{t}{(1+t^2x^2)(1+t^2)} dt.$$

8.f. g est impaire. Donc elle est dérivable sur $] -\infty, 0[$ car elle l'est sur $]0, +\infty[$, et on a alors

$$\forall x < 0, g'(x) = g'(-x) = \int_0^{+\infty} \frac{t}{(1+t^2x^2)(1+t^2)} dt.$$

9. Calcul de $g'(x)$ pour $x > 0$.

9.a. On a $g'(1) = \int_0^{+\infty} \frac{t}{(1+t^2)^2} dt$. Mais pour $A > 0$, on a

$$\begin{aligned} \int_0^A \frac{t}{(1+t^2)^2} dt &= \left[-\frac{1}{2} \frac{1}{1+t^2} \right]_0^A \\ &= \frac{1}{2} - \frac{1}{2} \frac{1}{1+A^2} \\ &\xrightarrow{A \rightarrow +\infty} \frac{1}{2} \end{aligned}$$

Et donc $g'(1) = \int_0^{+\infty} \frac{t}{(1+t^2)^2} dt = \frac{1}{2}$.

9.b. Supposons donc que de telles expressions existent, et que pour tout $t \in \mathbf{R}$,

$$\begin{aligned} \frac{t}{(1+t^2x^2)(1+t^2)} &= A(x) \frac{t}{1+t^2x^2} + B(x) \frac{t}{1+t^2} \\ \Leftrightarrow \frac{t}{(1+t^2x^2)(1+t^2)} &= \frac{A(x)t(1+t^2) + B(x)t(1+t^2x^2)}{(1+t^2x^2)(1+t^2)} \\ \Leftrightarrow \frac{t}{(1+t^2x^2)(1+t^2)} &= \frac{t(A(x) + B(x)) + t^3(A(x) + x^2B(x))}{(1+t^2x^2)(1+t^2)} \end{aligned}$$

Nous devons donc avoir pour tout $t \in \mathbf{R}$, $t = t(A(x) + B(x)) + t^3(A(x) + x^2B(x))$. Or deux polynômes sont égaux si et seulement si leurs coefficients sont égaux, c'est-à-dire si et seulement si

$$\begin{cases} A(x) + B(x) = 1 \\ A(x) + x^2B(x) = 0 \end{cases} \Leftrightarrow \begin{cases} A(x) = 1 - B(x) \\ B(x)(x^2 - 1) = -1 \end{cases} \Leftrightarrow \begin{cases} A(x) = -\frac{x^2}{1-x^2} \\ B(x) = \frac{1}{1-x^2} \end{cases}$$

Notons que ces quantités sont bien indépendantes de t .

Précision

Aussi compliquée soit l'intégrale, elle ne dépend pas de h , et nous avons prouvé qu'il s'agit bien d'une intégrale convergente, donc c'est une constante !

³ Nous revenons à la définition de la dérivée : limite du taux d'accroissement.

Parité/imparité

Pour retrouver le fait que g est dérivable sur \mathbf{R}^* , il suffit de se rappeler qu'on a $g(x) = -g(-x)$, et que donc, par composition de fonctions dérivables, g est dérivable sur \mathbf{R}^* . De plus, en dérivant cette relation, on a $g'(x) = -(-g'(-x)) = g'(-x)$.

De petits dessins, par exemple avec les fonctions carré (qui est paire) et cube (qui est impaire) peuvent être d'une aide précieuse dans ces cas là.

9.c. D'après la question précédente, on a pour $x \in]0, 1[\cup]1, +\infty[$,

$$g'(x) = \int_0^{+\infty} \left(-\frac{x^2}{1-x^2} \frac{t}{1+t^2x^2} + \frac{1}{1-x^2} \frac{t}{1+t^2} \right) dt.$$

Pour $A > 0$, on a alors

$$\int_0^A \frac{t}{1+t^2x^2} dt = \left[\frac{1}{2x^2} \ln(1+t^2x^2) \right]_0^A = \frac{1}{2x^2} \ln(1+x^2A^2)$$

et de même

$$\int_0^A \frac{t}{1+t^2} dt = \left[\frac{1}{2} \ln(1+t^2) \right]_0^A = \frac{1}{2} \ln(1+A^2).$$

Donc il vient, par linéarité de l'intégrale⁴

$$\begin{aligned} \int_0^A \left(-\frac{x^2}{1-x^2} \frac{t}{1+t^2x^2} + \frac{1}{1-x^2} \frac{t}{1+t^2} \right) dt &= -\frac{x^2}{1-x^2} \int_0^A \frac{t}{1+t^2x^2} dt + \frac{1}{1-x^2} \int_0^A \frac{t}{1+t^2} dt \\ &= \frac{-1}{2(1-x^2)} \ln(1+x^2A^2) + \frac{1}{2(1-x^2)} \ln(1+A^2) \\ &= \frac{1}{2(1-x^2)} \ln \left(\frac{1+A^2}{1+x^2A^2} \right) \\ &\xrightarrow{A \rightarrow +\infty} \frac{1}{2(1-x^2)} \ln \frac{1}{x^2} = \frac{\ln x}{x^2-1}. \end{aligned}$$

On en déduit que

$$g'(x) = \frac{\ln x}{x^2-1}.$$

9.d. Par composition de fonction usuelles, g' est continue sur $]0, 1[\cup]1, +\infty[$. Il reste donc juste à voir si elle est continue en 1. Or,

$$\frac{\ln(x)}{x^2-1} = \frac{\ln(1+(x-1))}{(x-1)(x+1)} \underset{x \rightarrow 1}{\sim} \frac{x-1}{(x-1)(x+1)} \underset{x \rightarrow 1}{\sim} \frac{1}{x+1} \underset{x \rightarrow 1}{\sim} \frac{1}{2}.$$

Donc $\lim_{x \rightarrow 1} g'(x) = \frac{1}{2} = g'(1)$: g' est continue en 1 et donc g est de classe \mathcal{C}^1 sur $]0, +\infty[$.

10. Une nouvelle expression de $g(x)$ pour $x > 0$.

10.a. Si $x < 1$, la fonction $t \mapsto \frac{\ln t}{t^2-1}$ est continue et positive⁵ sur $]0, x]$, donc l'unique éventuel problème de convergence est au voisinage de 0.

⁵ car numérateur et dénominateur sont négatifs !

Au voisinage de 0, $\frac{\ln t}{t^2-1} \underset{t \rightarrow 0}{\sim} \ln t$. Mais $\int_0^{1/2} \ln t dt$ est convergente car

$$\int_A^{1/2} \ln t dt = [t \ln t - t]_A^{1/2} = -\frac{\ln 2 + 1}{2} - A \ln A + A \xrightarrow{A \rightarrow 0^+} -\frac{\ln 2 + 1}{2}.$$

On en déduit donc que $\int_0^{1/2} \frac{\ln t}{t^2-1} dt$ converge et donc $\int_0^x \frac{\ln t}{t^2-1} dt$ converge.

Maintenant, si $x \geq 1$, la fonction $t \mapsto \frac{\ln t}{t^2-1}$ est continue sur $]0, 1[\cup]1, x]$, et donc il y a également un éventuel problème de convergence en 1. Mais nous avons prouvé précédemment que

$$\lim_{x \rightarrow 1} \frac{\ln x}{x^2-1} = \frac{1}{2}$$

de sorte que $\int_{1/2}^1 \frac{\ln t}{t^2-1} dt$ est faussement impropre, et donc convergente.

Dans tous les cas,

$$\forall x > 0, \int_0^x \frac{\ln t}{t^2-1} dt \text{ converge.}$$

⚠ Danger !

La tentation est grande ici d'utiliser la linéarité de l'intégrale, mais si on sépare en deux, les deux intégrales ainsi obtenues divergent, car

$$\frac{t}{1+t^2x^2} \underset{x \rightarrow 1}{\sim} \frac{1}{x^2t}$$

et de même

$$\frac{t}{1+t^2} \underset{x \rightarrow 1}{\sim} \frac{1}{t}.$$

⁴ Ici on parle de la linéarité de l'intégrale sur un segment, il n'y a donc pas de précautions à prendre

10.b. Puisque g' est continue sur \mathbf{R}_+ , nous savons que

$$\forall x \in \mathbf{R}_+, \forall y \in \mathbf{R}_+, g(x) - g(y) = \int_y^x g'(t) dt = \int_y^x \frac{\ln t}{t^2 - 1} dt.$$

À présent, lorsque $y \rightarrow 0^+$, par continuité de g , $g(y) \rightarrow g(0) = \int_0^{+\infty} 0 dt = 0$.

Et $\int_y^x \frac{\ln t}{t^2 - 1} dt \xrightarrow{y \rightarrow 0^+} \int_0^x \frac{\ln t}{t^2 - 1} dt$.

On a donc

$$\forall x \in \mathbf{R}_+, g(x) = \int_0^{+\infty} \frac{\arctan(tx)}{1+t^2} dt = \int_0^x \frac{\ln t}{t^2 - 1} dt.$$

11. Étude de la limite de g en $+\infty$

11.a. Nous savons que pour $x > 0$ et $t > 0$, par la question 2, $\arctan(tx) = \frac{\pi}{2} - \arctan\left(\frac{1}{tx}\right)$. Et donc

$$g(x) = \int_0^{+\infty} \left(\frac{\pi}{2} - \arctan\left(\frac{1}{tx}\right) \right) \frac{1}{1+t^2} dt.$$

Mais on a

$$\frac{\pi}{2} \int_0^{+\infty} \frac{dt}{1+t^2} = \frac{\pi}{2} \lim_{x \rightarrow +\infty} \arctan(x) = \frac{\pi^2}{4}.$$

Donc l'intégrale $\int_0^{+\infty} \arctan\left(\frac{1}{tx}\right) \frac{1}{1+t^2} dt$ qui est convergente car différence de deux intégrales convergentes, et

$$g(x) = \frac{\pi^2}{4} - \int_0^{+\infty} \arctan\left(\frac{1}{tx}\right) \frac{1}{1+t^2} dt.$$

11.b. Par la relation de Chasles, on a

$$\int_0^{+\infty} \arctan\left(\frac{1}{tx}\right) \frac{1}{1+t^2} dt = \int_0^{1/\sqrt{x}} \arctan\left(\frac{1}{tx}\right) \frac{1}{1+t^2} dt + \int_{1/\sqrt{x}}^{+\infty} \arctan\left(\frac{1}{tx}\right) \frac{1}{1+t^2} dt.$$

Il est aisé de voir que toutes ces intégrales sont positives⁶, et donc

$$\left| \int_0^{+\infty} \arctan\left(\frac{1}{tx}\right) \frac{1}{1+t^2} dt \right| = \int_0^{1/\sqrt{x}} \arctan\left(\frac{1}{tx}\right) \frac{1}{1+t^2} dt + \int_{1/\sqrt{x}}^{+\infty} \arctan\left(\frac{1}{tx}\right) \frac{1}{1+t^2} dt.$$

Mais pour $t \in \left]0, \frac{1}{\sqrt{x}}\right[$, on a

$$0 \leq \arctan\left(\frac{1}{tx}\right) \frac{1}{1+t^2} \leq \arctan\left(\frac{1}{tx}\right) \leq \frac{\pi}{2}.$$

Donc, par croissance de l'intégrale,

$$0 \leq \int_0^{1/\sqrt{x}} \arctan\left(\frac{1}{tx}\right) \frac{1}{1+t^2} dt \leq \int_0^{1/\sqrt{x}} \frac{\pi}{2} dt = \frac{\pi}{2\sqrt{x}}.$$

Pour la seconde intégrale, notons que d'après la question 1 appliquée avec $y = 0$, on a

$$\forall x > 0, \forall t \in \mathbf{R}_+, 0 \leq \arctan\left(\frac{1}{tx}\right) \leq \frac{1}{tx}.$$

Et donc, par croissance de l'intégrale

$$0 \leq \int_{1/\sqrt{x}}^{+\infty} \arctan\left(\frac{1}{tx}\right) \frac{1}{1+t^2} dt \leq \frac{1}{x} \int_{1/\sqrt{x}}^{+\infty} \frac{1}{t(1+t^2)} dt$$

Mais pour $t \in \left[\frac{1}{\sqrt{x}}, +\infty\right[$, on a

$$0 \leq \frac{1}{t} \leq \sqrt{x} \Rightarrow 0 \leq \frac{1}{t} \frac{1}{1+t^2} \leq \sqrt{x} \frac{1}{1+t^2}$$

Subtilité

Remplacer $g'(t)$ par $\frac{\ln t}{t^2-1}$ est juste... si $t \neq 1$. Mais comme deux fonctions qui sont égales sauf en un nombre fini de points ont la même intégrale, on ne fait pas d'erreur ici.

Astuce

On pourrait redémontrer que cette intégrale converge par le calcul, mais il est bien plus rapide d'utiliser ce que nous avons déjà fait !

⁶ Car intégrales de fonctions positives

On met ici la valeur absolue pour être en accord avec l'énoncé, mais elle ne sert strictement à rien !

Méthode

Avant de se lancer dans des calculs sans fin, il est conseillé de chercher au brouillon quel est la contribution de chaque terme au majorant que l'on souhaite (ici $\frac{\pi}{\sqrt{x}}$).

Ici, on souhaiterait que chacun soit inférieur à $\frac{\pi}{2\sqrt{x}}$.

Cette dernière intégrale converge bien par comparaison à une intégrale de Riemann convergente.

si bien que, par croissance de l'intégrale,

$$0 \leq \int_{1/\sqrt{x}}^{+\infty} \frac{1}{t} \frac{1}{1+t^2} dt \leq \sqrt{x} \int_{1/\sqrt{x}}^{+\infty} \frac{dt}{1+t^2} \leq \sqrt{x} \int_0^{+\infty} \frac{dt}{1+t^2} \leq \sqrt{x} \frac{\pi}{2}.$$

On en déduit que

$$\frac{1}{x} \int_{1/\sqrt{x}}^{+\infty} \frac{1}{t} \frac{1}{1+t^2} dt \leq \frac{\sqrt{x} \pi}{x} \frac{1}{2} = \frac{\pi}{2\sqrt{x}}.$$

Et donc

$$\left| \int_0^{+\infty} \arctan\left(\frac{1}{tx}\right) \frac{1}{1+t^2} dt \right| \leq \frac{\pi}{2\sqrt{x}} + \frac{\pi}{2\sqrt{x}} = \boxed{\frac{\pi}{\sqrt{x}}}.$$

11.c. Par le théorème des gendarmes, il vient alors

$$\lim_{x \rightarrow +\infty} \int_0^{+\infty} \arctan\left(\frac{1}{tx}\right) \frac{1}{1+t^2} dt = 0$$

puis

$$\boxed{\lim_{x \rightarrow +\infty} g(x) = \frac{\pi^2}{4} - \lim_{x \rightarrow +\infty} \int_0^{+\infty} \arctan\left(\frac{1}{tx}\right) \frac{1}{1+t^2} dt = \frac{\pi^2}{4}.$$

12. Application au calcul de $\sum_{n=0}^{+\infty} \frac{1}{(2n+1)^2}$.

12.a. Nous savons que pour $x > 0$, $\int_0^x \frac{\ln t}{t^2-1} dt$ converge et vaut $g(x)$. Or, à la question 11, il vient d'être démontré que

$$\lim_{x \rightarrow +\infty} g(x) = \frac{\pi^2}{4} \Leftrightarrow \lim_{x \rightarrow +\infty} \int_0^x \frac{\ln t}{t^2-1} dt = \frac{\pi^2}{4}.$$

Ceci nous garantit donc que $\int_0^{+\infty} \frac{\ln t}{t^2-1} dt$ converge et que

$$\int_0^{+\infty} \frac{\ln t}{t^2-1} dt = \frac{\pi^2}{4} \Leftrightarrow \boxed{\int_0^{+\infty} \frac{\ln t}{1-t^2} dt = -\frac{\pi^2}{4}.$$

12.b. La fonction $t \mapsto \frac{1}{t}$ réalise une bijection \mathcal{C}^1 de $]1, +\infty[$ sur $]0, 1[$. Procédons alors, dans $\int_1^{+\infty} \frac{\ln t}{1-t^2} dt$, au changement de variable $t = \frac{1}{u}$, de sorte que $dt = -\frac{1}{u^2} du$. Puisque l'intégrale $\int_1^{+\infty} \frac{\ln t}{1-t^2} dt$ converge, la convergence de l'intégrale obtenue après changement de variable est automatique et

$$\int_1^{+\infty} \frac{\ln t}{1-t^2} dt = \int_1^0 \frac{\ln \frac{1}{u}}{1-\frac{1}{u^2}} \left(-\frac{1}{u^2}\right) du = \int_0^1 \frac{-\ln u}{u^2-1} du = \int_0^1 \frac{\ln u}{1-u^2} du.$$

Et alors, par la relation de Chasles,

$$\begin{aligned} \int_0^{+\infty} \frac{\ln t}{1-t^2} dt &= \int_0^1 \frac{\ln t}{1-t^2} dt + \int_1^{+\infty} \frac{\ln t}{1-t^2} dt \\ &= \int_0^1 \frac{\ln t}{1-t^2} dt + \int_0^1 \frac{\ln u}{1-u^2} du \\ &= \boxed{2 \int_0^1 \frac{\ln t}{1-t^2} dt.} \end{aligned}$$

12.c. Pour $t \in]0, 1[$, on a⁷

$$\sum_{k=0}^n t^{2k} = \frac{1-t^{2n+2}}{1-t^2} \Leftrightarrow \frac{1}{1-t^2} = \sum_{k=0}^n t^{2k} + \frac{t^{2n+2}}{1-t^2}.$$

Chgt de variable

Une bijection \mathcal{C}^1 est automatiquement strictement monotone grâce au théorème des valeurs intermédiaires.

⁷ Rappelons que ce qui est important ici n'est pas le fait d'être dans $]0, 1[$, mais différent de 1, la formule donnant les sommes partielles de la série $\sum q^n$ étant valable pour tout $q \neq 1$, y compris si la série diverge.

Et donc, après multiplication par $\ln t$,

$$\frac{\ln t}{1-t^2} = \sum_{k=0}^n (\ln t)t^{2k} + \frac{\ln t}{1-t^2}t^{2n+2}.$$

De plus, pour $k \in \mathbf{N}$, pour tout $t \in]0, 1]$,

$$|(\ln t)t^{2k}| \leq |\ln t| = -\ln t.$$

En effet, $\ln t \leq 0$ et donc $|\ln t| = -\ln t$.

Mais $\int_0^1 \ln t \, dt$ converge car

$$\int_A^1 \ln t \, dt = [t \ln t]_A^1 = -1 + A - A \ln A \underset{A \rightarrow 0^+}{\rightarrow} 1.$$

Donc par comparaison, $\int_0^1 t^{2k} \ln t \, dt$ converge (absolument).

Alors, puisque nous avons déjà prouvé que $\int_0^1 \frac{\ln t}{1-t^2} \, dt$ converge, la convergence de

$\int_0^1 \frac{\ln t}{1-t^2} t^{2n+2} \, dt$ est automatique, par somme d'intégrales convergentes et

$$\int_0^1 \frac{\ln t}{1-t^2} \, dt = \sum_{k=0}^n \int_0^1 (\ln t)t^{2k} \, dt + \int_0^1 \frac{\ln t}{1-t^2} t^{2n+2} \, dt.$$

12.d. Nous venons de prouver que pour $n = 0$, on a $\int_0^1 \ln t \, dt = -1$.

De manière générale, procédons à une intégration par parties sur un segment de la forme $[A, 1-A] \subset]0, 1]$. En posant $u = \ln t$, $v' = t^{2n}$, fonctions de classe \mathcal{C}^1 sur $[A, 1-A]$, on a

$$\begin{aligned} \int_A^{1-A} (\ln t)t^{2n} \, dt &= \left[\ln t \frac{t^{2n+1}}{2n+1} \right]_A^{1-A} - \int_A^{1-A} \frac{t^{2n+1}}{t(2n+1)} \, dt \\ &= \frac{1}{2n+1} (\ln(1-A)(1-A)^{2n+1} - \ln(A)A^{2n+1}) - \int_A^{1-A} \frac{t^{2n}}{2n+1} \, dt \\ &= \frac{1}{2n+1} (\ln(1-A)(1-A)^{2n+1} - \ln(A)A^{2n+1}) - \left[\frac{t^{2n+1}}{(2n+1)^2} \right]_A^{1-A} \\ &= \frac{1}{2n+1} (\ln(1-A)(1-A)^{2n+1} - \ln(A)A^{2n+1}) - \frac{1}{(2n+1)^2} ((1-A)^{2n+1} - A^{2n+1}) \\ &\xrightarrow{A \rightarrow 0} \frac{1}{(2n+1)^2}. \end{aligned}$$

Astuce

On sait déjà que l'intégrale converge, si on souhaite gagner du temps, plutôt que de prendre un segment $[A, B]$ puis prendre deux limites, on peut se contenter d'une seule limite. Le gain de temps n'est pas énorme pour autant...

△ Ceci n'est valable que si l'on sait déjà que l'intégrale converge. Si vous n'êtes pas sûr de votre coup, mieux vaut prendre deux limites.

12.e. Coupons l'intégrale en deux à l'aide de la relation de Chasles :

$$\int_0^1 \frac{\ln t}{t^2-1} t^{2n+2} \, dt = \int_0^{1/2} \frac{\ln t}{t^2-1} t^{2n+2} \, dt + \int_{1/2}^1 \frac{\ln t}{t^2-1} t^{2n+2} \, dt.$$

La fonction $t \mapsto \frac{\ln t}{1-t^2}$ est continue et négative sur $[1/2, 1]$ et donc y admet un minimum $m \leq 0$.

Ainsi, pour tout $n \in \mathbf{N}$, on a, par croissance de l'intégrale,

$$0 \geq \int_{1/2}^1 \frac{\ln t}{t^2-1} t^{2n+2} \, dt \geq m \int_{1/2}^1 t^{2n+2} \, dt \geq \frac{m}{2n+3} \left(1 - \frac{1}{2^{2n+3}}\right) \geq \frac{m}{2n+3}.$$

Par le théorème des gendarmes, on a alors

$$\lim_{n \rightarrow +\infty} \int_{1/2}^1 \frac{\ln t}{1-t^2} t^{2n+2} \, dt = 0.$$

Méthode

Pour montrer que la limite est nulle, on a sûrement envie d'utiliser des majorations, et pourquoi pas d'utiliser

$$\lim_{n \rightarrow +\infty} \int_0^1 t^n \, dt = 0.$$

Le problème est que la fonction $t \mapsto \frac{\ln t}{1-t^2}$ tend à la fois vers $-\infty$ en 0 et vers $+\infty$ en 1, donc il est dur de majorer efficacement sur $]0, 1[$. C'est pour cette raison que l'on choisit de la couper en deux, pour traiter séparément le problème en 0 et le problème en 1.

Sur $]0, 1/2[$, on a $1 - t^2 \in [\frac{3}{4}, 1]$ et donc

$$\frac{4}{3}(\ln t)t^{2n+2} \leq \frac{\ln t}{1-t^2}t^{2n+2} \leq 0$$

Et donc

$$\frac{4}{3} \int_0^{1/2} (\ln t)t^{2n+2} dt \leq \int_0^{1/2} \frac{\ln t}{1-t^2}t^{2n+2} dt \leq 0.$$

Mais la fonction $t \mapsto (\ln t)t^{2n+2}$ est négative sur $]0, \frac{1}{2}]$, donc

$$\int_0^{1/2} (\ln t)t^{2n+2} dt \geq \int_0^1 (\ln t)t^{2n+2} dt = -\frac{1}{(2n+1)^2}.$$

Et enfin

$$-\frac{4}{3(2n+1)^2} \leq \int_0^{1/2} \frac{\ln t}{1-t^2}t^{2n+2} dt \leq 0.$$

Par le théorème des gendarmes, on a alors

$$\lim_{n \rightarrow +\infty} \int_0^{1/2} \frac{\ln t}{1-t^2}t^{2n+2} dt = 0.$$

Enfin, par somme de limites, il vient

$$\boxed{\lim_{n \rightarrow +\infty} \int_0^1 \frac{\ln t}{1-t^2}t^{2n+2} dt = 0.}$$

12.f. En combinant les résultats des questions 12.a à 12.d, il vient

$$-\frac{\pi^2}{8} = \frac{1}{2} \int_0^{+\infty} \frac{\ln t}{1-t^2} dt = \int_0^1 \frac{\ln t}{1-t^2} dt = -\sum_{k=0}^n \frac{1}{(2k+1)^2} + \int_0^1 \frac{\ln t}{1-t^2}t^{2n+2} dt.$$

En passant à la limite lorsque $n \rightarrow +\infty$, on a alors la convergence⁸ de la série de terme général $\frac{1}{(2k+1)^2}$ et

$$-\frac{\pi^2}{8} = -\sum_{k=0}^{+\infty} \frac{1}{(2k+1)^2} \Leftrightarrow \boxed{\sum_{n=0}^{+\infty} \frac{1}{(2n+1)^2} = \frac{\pi^2}{8}.}$$

Enfin, en séparant les termes pairs des termes impairs dans la série (convergente) de terme général $\frac{1}{n^2}$, on a

$$\sum_{n=1}^{+\infty} \frac{1}{n^2} = \sum_{n=0}^{+\infty} \frac{1}{(2n+1)^2} + \sum_{n=1}^{+\infty} \frac{1}{(2n)^2} = \sum_{n=0}^{+\infty} \frac{1}{(2n+1)^2} + \frac{1}{4} \sum_{n=1}^{+\infty} \frac{1}{n^2} \Leftrightarrow \frac{3}{4} \sum_{n=1}^{+\infty} \frac{1}{n^2} = \sum_{n=0}^{+\infty} \frac{1}{(2n+1)^2}.$$

Soit encore

$$\sum_{n=1}^{+\infty} \frac{1}{n^2} = \frac{4}{3} \sum_{n=0}^{+\infty} \frac{1}{(2n+1)^2} = \frac{4}{3} \frac{\pi^2}{8} = \boxed{\frac{\pi^2}{6}.}$$

Signes !

In t est négatif, donc on a bien pris soin de changer le sens des inégalités

Important

L'inégalité qui suit n'est valable que parce que la fonction est négative (c'est clair si on pense à l'interprétation de l'intégrale en terme d'aire), elle ne le serait pas nécessairement pour une fonction changeant de signe.

⁸ Cette convergence était aisée à obtenir par ailleurs par comparaison à une série de Riemann.

Classique

Une telle relation existe pour toutes les séries de Riemann : si l'on connaît la somme de la série des termes impairs, on peut en déduire la somme de la série entière et vice-versa.

DEVOIR MAISON 4 (À RENDRE LE 11.10.18)

Notations et objectifs du problème.

On désigne par I l'intervalle $[1, +\infty[$; on note E l'espace vectoriel des fonctions continues et bornées sur I à valeurs réelles et $\mathcal{C}^1(I, \mathbf{R})$ l'espace vectoriel des fonctions de classe \mathcal{C}^1 sur I à valeurs réelles.

On fixe enfin a un réel strictement positif.

Pour f un élément de E , on dit qu'une fonction y de $\mathcal{C}^1(I, \mathbf{R})$ est une solution du problème (E_f) si :

$$\forall x \in I, \quad y'(x) - ay(x) + f(x) = 0.$$

L'objectif de ce problème est de montrer qu'à tout élément f de E , on peut associer une unique solution g de (E_f) qui soit bornée sur I , puis d'étudier l'opérateur $U : f \mapsto g$.

Les trois parties du problème traitent, souvent à partir d'exemples, de propriétés de l'opérateur U .

I. Existence et propriétés élémentaires de l'opérateur U .

1. Étude de l'équation (E_f) .

- a. On considère $f \in E$ et $y \in \mathcal{C}^1(I, \mathbf{R})$. Écrire la dérivée de $x \mapsto e^{-ax}y(x)$. Montrer alors que y est solution du problème (E_f) si et seulement si il existe $K \in \mathbf{R}$ tel que : $\forall x \in I, \quad y(x) = e^{ax} \left(K - \int_1^x e^{-at} f(t) dt \right)$.
- b. Montrer que, s'il existe une solution de (E_f) qui soit bornée sur I , celle-ci est unique.
- c. Vérifier que l'intégrale $\int_1^{+\infty} e^{-at} f(t) dt$ est convergente.
- d. Démontrer que $g : x \mapsto e^{ax} \int_x^{+\infty} e^{-at} f(t) dt$ est l'unique solution de (E_f) qui soit bornée sur I .

Dans toute la suite du problème, si $f \in E$, on note $U(f)$ la fonction g obtenue à la question d.

2. Linéarité de U .

- a. Expliciter $U(f)$ dans le cas où $f = 1$.
- b. Montrer que U est un endomorphisme de E .
- c. U est-il injectif ?
- d. On définit les puissances successives de U par $U^0 = \text{id}_E$ et pour tout entier naturel n non nul, $U^n = U^{n-1} \circ U$.
Montrer que, pour tout entier naturel n , $U^{n+1}(f)$ est la fonction $x \mapsto e^{ax} \int_x^{+\infty} \frac{(t-x)^n}{n!} e^{-at} f(t) dt$.

3. Cas des fonctions exponentielles.

- a. Pour k un nombre réel positif et f_k la fonction $x \mapsto e^{-kx}$, expliciter $U(f_k)$.
- b. En déduire que, pour tout réel $\lambda \in \left] 0, \frac{1}{a} \right]$, $\text{Ker}(U - \lambda \text{id}_E) \neq \{0\}$.
- c. Pour tout entier naturel n , expliciter $U^n(f_k)$. Pour x élément de I , préciser $\lim_{n \rightarrow +\infty} [U^n(f)](x)$.

4. Cas des fonctions sinus et cosinus.

Dans cet exemple seulement (ensemble de la question 4), on prendra $a = 1$.

- a. Expliciter $U(\sin)$ et $U(\cos)$.
- b. Montrer que le sous-espace vectoriel P de E engendré par les fonctions \sin et \cos est stable par U et que (\sin, \cos) en est une base. Dans cette base, écrire la matrice M de l'endomorphisme $U_P : \begin{cases} P & \rightarrow P \\ f & \mapsto U(f) \end{cases}$
- c. Calculer M^2, M^3, M^4 . Expliciter M^n pour tout entier naturel n , puis préciser la limite des coefficients de M^n lorsque n tend vers l'infini.

5. Une autre famille de fonctions.

Pour tout entier naturel n , on considère la fonction de E $\varphi_n : x \mapsto e^{-x}x^n$ et on note ψ_n la fonction $U(\varphi_n)$.

- a. Pour n entier naturel non nul, établir une relation entre ψ_n, φ_n et ψ_{n-1} .
- b. Pour p entier naturel, montrer que le sous-espace F_p de E engendré par $(\varphi_0, \varphi_1, \dots, \varphi_p)$ est stable par U et admet pour base $(\varphi_0, \varphi_1, \dots, \varphi_p)$.

- c. On prend ici $p = 2$, écrire dans la base $(\varphi_0, \varphi_1, \varphi_2)$ de F_2 la matrice T_2 de l'endomorphisme $U_2 : \begin{cases} F_2 & \rightarrow F_2 \\ f & \mapsto U(f) \end{cases}$.
Calculer T_2^n pour tout entier naturel n , puis préciser la limite des coefficients de T_2^n lorsque n tend vers l'infini.

6. Une autre expression de $U(f)$.

Pour $f \in E$, montrer que : $\forall x \in I, U(f)(x) = \int_0^{+\infty} e^{-at} f(x+t) dt$.

7. Positivité de U .

- a. Pour $f \in E$, montrer que : $|U(f)| \leq U(|f|)$.

On considère maintenant φ un élément de E à valeurs positives et $\psi = U(\varphi)$.

- b. Montrer que ψ est à valeurs positives.

- c. On suppose que φ est décroissante. Montrer que $a\psi \leq \varphi$ puis que ψ est décroissante.

8. Commutation de U avec la dérivation.

On note $E_1 = \{f \in E \cap \mathcal{C}^1(I, \mathbf{R}) / f' \text{ bornée sur } I\}$ et D l'opérateur de dérivation qui, à tout élément de E_1 , associe sa dérivée.

- a. Pour f un élément de E_1 , montrer, en utilisant la question 6, que : $aU(f) = f + U(f')$.

- b. En déduire que, pour tout élément f de E_1 , $D(U(f)) = U(D(f))$.

- c. Pour f une fonction de E_1 à valeurs positives et décroissante, retrouver le résultat de la question 7.c : $U(f)$ est décroissante.

II. Comportement asymptotique de $U(f)$ au voisinage de $+\infty$.

9. Résultats préliminaires.

On considère ici α et β deux fonctions à valeurs réelles, continues sur I . On suppose que, $\forall x \in I, \beta(x) > 0$, et que $\int_1^{+\infty} \beta(t) dt$ converge.

- a. On suppose ici que $\alpha(x) = o_{x \rightarrow +\infty}(\beta(x))$ et on se propose de montrer que $\int_x^{+\infty} \alpha(t) dt = o_{x \rightarrow +\infty} \left(\int_x^{+\infty} \beta(t) dt \right)$.

Soit $\varepsilon > 0$, montrer que : $\exists A > 0 / \forall x \geq A, \left| \int_x^{+\infty} \alpha(t) dt \right| \leq \varepsilon \int_x^{+\infty} \beta(t) dt$. Conclure.

- b. On suppose maintenant que $\alpha(x) \sim_{x \rightarrow +\infty} \beta(x)$, montrer que $\int_x^{+\infty} \alpha(t) dt \sim_{x \rightarrow +\infty} \int_x^{+\infty} \beta(t) dt$.

10. Cas des fonctions admettant une limite en $+\infty$.

Si f est un élément de E admettant une limite finie b en $+\infty$ (b est un nombre réel), montrer que $g = U(f)$ admet une limite en $+\infty$ que l'on précisera (on pourra commencer par traiter le cas où $b = 0$ en écrivant, dans ce cas, $f(t) = o_{t \rightarrow +\infty}(1)$ et en utilisant la question 9).

11. Cas des fonctions puissances.

Dans cette question et la suivante, ω est un réel strictement positif, on note f_ω la fonction $t \mapsto \frac{1}{t^\omega}$ et $g_\omega = U(f_\omega)$.

- a. Montrer que $g_\omega(x) = \frac{f_\omega(x)}{a} - \frac{\omega}{a} g_{\omega+1}(x)$. En déduire que $g_\omega(x) \sim_{x \rightarrow +\infty} \frac{f_\omega(x)}{a}$.

- b. Montrer que pour tout x de I , $\int_1^x \frac{e^{-at}}{t} dt = \ln(x) + \sum_{k=1}^{+\infty} (-1)^k \frac{a^k}{k \cdot k!} (x^k - 1)$ (On pourra utiliser une inégalité de Taylor-Lagrange pour la fonction $t \mapsto e^{-at}$), et en déduire :

$$g_1(x) = e^{-ax} \left\{ -\ln x - \sum_{k=1}^{+\infty} (-1)^k \frac{a^k}{k \cdot k!} (x^k - 1) + \int_1^{+\infty} \frac{e^{-at}}{t} dt \right\}.$$

12. Cas des fonctions comparables aux fonctions puissances f_ω .

On note toujours f un élément de E et $g = U(f)$.

- a. Prouver que si f est négligeable devant f_ω au voisinage de $+\infty$, alors g est négligeable devant g_ω au voisinage de $+\infty$.

- b. Prouver que si f est équivalent à f_ω au voisinage de $+\infty$, alors g est équivalent à $\frac{f}{a}$ au voisinage de $+\infty$.

III. Convergence absolue de $\int_1^{+\infty} U(f)$.

On s'intéresse dans cette partie à la convergence de $\int_1^{+\infty} |[U(f)](t)| dt$ dans le cas où $\int_1^{+\infty} |f(t)| dt$ est elle-même convergente. On note toujours $g = U(f)$.

13. Étude d'exemples.

- Pour k un réel strictement positif et $f_k : t \mapsto e^{-kt}$, on note $g_k = U(f_k)$. Vérifier que $\int_1^{+\infty} g_k(t) dt$ est convergente.
- Pour ω un réel strictement positif, on note encore $f_\omega : t \mapsto \frac{1}{t^\omega}$ et $g_\omega = U(f_\omega)$. Pour quelles valeurs de ω l'intégrale $\int_1^{+\infty} g_\omega(t) dt$ est-elle convergente ?

14. Cas des fonctions positives.

Dans cette question, f est un élément de E , à valeurs positives tel que $\int_1^{+\infty} f(t) dt$ est convergente.

On note $F : x \in I \mapsto \int_1^x f(t) dt$, $g = U(f)$ et $G : x \in I \mapsto \int_1^x g(t) dt$.

- Vérifier que $G' - aG = -F + g(1)$.
- Justifier que F est dans E et montrer qu'il existe une constante réelle K telle que, pour tout $x \in I$,
$$G(x) = Ke^{ax} + [U(F)](x) - \frac{g(1)}{a}.$$
- Vérifier que la fonction $x \mapsto \frac{G(x)}{x}$ est bornée sur I .
- En déduire que $K = 0$ et que $G = U(F) - \frac{g(1)}{a}$.
- Montrer alors que $\int_1^{+\infty} g(t) dt$ est convergente.

15. Cas général.

Dans cette question, f est un élément de E tel que $\int_1^{+\infty} f(t) dt$ est absolument convergente.

Montrer que $\int_1^{+\infty} g(t) dt$ est absolument convergente.

CORRECTION DU DEVOIR MAISON 4

I. Existence et propriétés élémentaires de l'opérateur U .1. Étude de l'équation (E_f) .

1.a. La dérivée de $h : x \mapsto e^{-ax}y(x)$ est $h' : x \mapsto e^{-ax}(y'(x) - ay(x))$.

En particulier, y est solution de (E_f) si et seulement si pour tout $x \in I$,

$$e^{ax}h'(x) + f(x) = 0 \Leftrightarrow h'(x) = -e^{-ax}f(x).$$

C'est-à-dire si et seulement si h est une primitive de $x \mapsto e^{-ax}f(x)$.

Mais par le théorème fondamental de l'analyse, $x \mapsto -\int_1^x e^{-at}f(t)dt$ est une primitive de $x \mapsto -e^{-ax}f(x)$, et donc h est une primitive de $x \mapsto -e^{-ax}f(x)$ si et seulement si il existe $K \in \mathbf{R}$ tel que $\forall x \in I$, $h(x) = K - \int_1^x e^{-at}f(t)dt$.

Autrement dit, si et seulement si il existe $K \in \mathbf{R}$ tel que

$$\forall x \in I, y(x) = e^{ax} \left(K - \int_1^x e^{-at}f(t)dt \right).$$

1.b. Supposons que y_1, y_2 soient deux solutions bornées de (E_f) . Alors il existe deux constantes réelles K_1, K_2 telles que

$$\forall x \in I, y_1(x) = e^{ax} \left(K_1 - \int_1^x e^{-at}f(t)dt \right) \text{ et } y_2(x) = e^{ax} \left(K_2 - \int_1^x e^{-at}f(t)dt \right).$$

De plus, il existe deux réels positifs M_1, M_2 tels que

$$\forall x \in I, |y_1(x)| \leq M_1 \text{ et } |y_2(x)| \leq M_2.$$

Mais alors $y_1 - y_2$ est encore bornée car

$$\forall x \in I, |y_1(x) - y_2(x)| \leq |y_1(x)| + |y_2(x)| \leq M_1 + M_2.$$

Or, $|y_1(x) - y_2(x)| = e^{ax}|K_1 - K_2|$.

Si $K_1 \neq K_2$, alors $|K_1 - K_2| > 0$ et donc $\lim_{x \rightarrow +\infty} |y_1(x) - y_2(x)| = +\infty$, ce qui est impossible car $y_1 - y_2$ est bornée.

On en déduit que $K_1 = K_2$, et donc que $y_1 = y_2$.

Ainsi, (E_f) possède une solution bornée, celle-ci est unique.

1.c. Puisque $f \in E$, elle est bornée, et donc il existe $M > 0$ tel que $\forall t \in I, |f(t)| \leq M$.

Ainsi, pour tout $t \in I$, on a $0 \leq |e^{-at}f(t)| \leq Me^{-at}$.

Mais $a > 0$, et donc $\int_1^{+\infty} Me^{-at}dt$ converge.

On en déduit donc que $\int_1^{+\infty} |e^{-at}f(t)|dt$ converge, c'est-à-dire que $\int_1^{+\infty} e^{-at}f(t)dt$ converge absolument et donc converge.

1.d. Notons que par la question précédente, $\int_x^{+\infty} e^{-at}f(t)dt$ est bien convergente.

D'autre part, on a

$$\left| \int_x^{+\infty} e^{-at}f(t)dt \right| \leq \int_x^{+\infty} |e^{-at}f(t)|dt \leq M \int_x^{+\infty} e^{-at}dt.$$

Mais pour $A > x$, on a

$$\int_x^A e^{-at}dt = \left[-\frac{1}{a}e^{-at} \right]_x^A = \frac{1}{a}(e^{-ax} - e^{-aA}) \xrightarrow{A \rightarrow +\infty} \frac{1}{a}e^{-ax}.$$

Mieux

Nous savons même qu'il s'agit de l'unique primitive qui s'annule en 1.

Remarque

Notons que nous venons donc de prouver que (E_f) possède toujours une infinité de solutions.

Danger

Attention aux signes, il n'est pas question d'affirmer que

$$|y_1(x) - y_2(x)| \leq |y_1(x)| - |y_2(x)|,$$

l'inégalité triangulaire nous donne seulement un signe +.

Et donc il vient $\int_x^{+\infty} e^{-at} dt = \frac{e^{-ax}}{a}$ et par conséquent

$$|g(x)| = e^{ax} \left| \int_x^{+\infty} e^{-at} f(t) dt \right| \leq e^{ax} M \frac{e^{-ax}}{a} \leq \frac{M}{a}.$$

Ainsi, g est bien une fonction bornée sur I .

D'autre part, si on note $K = \int_1^{+\infty} e^{-at} f(t) dt$, alors pour tout $x \in I$, par la relation de Chasles,

$$g(x) = e^{ax} \int_x^{+\infty} e^{-at} f(t) dt = e^{ax} \left(K - \int_1^x e^{-at} f(t) dt \right)$$

et donc par la question 1.a, g est une solution de (E_f) .

Puisqu'elle est bornée, la question 1.c garantit qu'il s'agit de l'unique solution bornée de (E_f) .

2. Linéarité de U .

2.a. Si $f = 1$, on a pour tout $x \in I$,

$$[U(f)](x) = e^{ax} \int_x^{+\infty} e^{-at} dt = e^{ax} \frac{e^{-ax}}{a} = \frac{1}{a}.$$

2.b. Soient f et g deux éléments de E et $\lambda \in \mathbf{R}$. Alors, pour tout $x \in I$,

$$\begin{aligned} [U(\lambda f + g)](x) &= e^{ax} \int_x^{+\infty} e^{-at} (\lambda f(t) + g(t)) dt \\ &= \lambda e^{ax} \int_x^{+\infty} e^{-at} f(t) dt + e^{ax} \int_x^{+\infty} e^{-at} g(t) dt = \lambda [U(f)](x) + [U(g)](x). \end{aligned}$$

Ainsi, on a $U(\lambda f + g) = \lambda U(f) + U(g)$, donc U est linéaire.

De plus, si $f \in E$, alors $U(f)$ est continue¹ et bornée par définition de $U(f)$.

Et ainsi $U(f) \in E$: U est un endomorphisme de E .

2.c. Soit $f \in \text{Ker}(U)$. Alors pour tout $x \in E$,

$$e^{ax} \int_x^{+\infty} e^{-at} f(t) dt = 0.$$

En particulier, pour tout $x \in I$, $\int_x^{+\infty} e^{-at} f(t) dt = 0$.

Mais $\int_x^{+\infty} e^{-at} f(t) dt = \int_1^{+\infty} e^{-at} f(t) dt - \int_1^x e^{-at} f(t) dt$ a pour dérivée $x \mapsto -e^{-ax} f(x)$.

Et donc $x \mapsto -e^{-ax} f(x)$ est la fonction nulle, et donc pour tout $x \in I$, $f(x) = 0$.

Ainsi, $\text{Ker}(U) = \{0\}$ et donc U est injectif.

2.d. Procédons par récurrence sur n . Pour $n = 0$, le résultat est vrai d'après la question 1.d.

Supposons donc que pour tout $f \in E$ et pour tout $x \in I$, $[U^{n+1}(f)](x) = e^{ax} \int_x^{+\infty} \frac{(t-x)^n}{n!} e^{-at} f(t) dt$.

En particulier

$$[U^{n+2}(f)](x) = [U^{n+1}(U(f))](x) = e^{ax} \int_x^{+\infty} \frac{(t-x)^n}{n!} e^{-at} [U(f)](t) dt.$$

Mais nous savons² que $U(f)$ est dérivable sur I avec

$$[U(f)]'(x) = a[U(f)](x) - f(x).$$

Procédons alors à une intégration par parties sur $[x, A]$, $A \geq x$, en posant

$$u(t) = \frac{(t-x)^{n+1}}{(n+1)!} \text{ et } v(t) = e^{-at} [U(f)](t)$$

qui sont deux fonctions de classe \mathcal{C}^1 , avec

$$u'(t) = \frac{(t-x)^n}{n!} \text{ et } v'(t) = e^{-at} (-a[U(f)](t) + [U(f)]'(t)) = -e^{-at} f(t).$$

¹ Et même \mathcal{C}^1 car solution de (E_f) .

Rédaction

Rappelons qu'une fonction est nulle si et seulement si elle s'annule en tout point de son ensemble de définition. Il est donc important ici de préciser que ceci est vrai pour tout $x \in I$.

² C'est la définition de $U(f)$.

Il vient alors

$$\begin{aligned} \int_x^A \frac{(t-x)^n}{n!} e^{-at} [U(f)](t) dt &= \left[\frac{(t-x)^{n+1}}{(n+1)!} e^{-at} [U(f)](t) \right]_x^A + \int_x^A \frac{(t-x)^{n+1}}{(n+1)!} e^{-at} f(t) dt \\ &= -\frac{(A-x)^{n+1}}{(n+1)!} e^{-aA} [U(f)](A) + \int_x^A \frac{(t-x)^{n+1}}{(n+1)!} e^{-at} f(t) dt. \end{aligned}$$

Par croissances comparées, $\frac{(A-x)^{n+1}}{(n+1)!} e^{-aA} \xrightarrow{A \rightarrow +\infty} 0$, et puisque $U(f)$ est bornée,

$$\lim_{A \rightarrow +\infty} \frac{(A-x)^{n+1}}{(n+1)!} e^{-aA} [U(f)](A) = 0.$$

On en déduit donc que

$$\int_x^{+\infty} \frac{(t-x)^n}{n!} e^{-at} [U(f)](t) dt = \int_x^{+\infty} \frac{(t-x)^{n+1}}{(n+1)!} e^{-at} f(t) dt.$$

Ce qui après multiplication par e^{ax} nous donne $[U^{n+2}(f)](x) = e^{ax} \int_x^{+\infty} \frac{(t-x)^{n+1}}{(n+1)!} e^{-at} f(t) dt$.

Et donc par le principe de récurrence, pour tout $x \in I$, pour tout $n \in \mathbf{N}$ et pour tout $f \in E$,

$$\boxed{[U^{n+1}(f)](x) = e^{ax} \int_x^{+\infty} \frac{(t-x)^n}{n!} e^{-at} f(t) dt.}$$

3. Cas des fonctions exponentielles.

3.a. Pour $x \in I$, on a

$$[U(f_k)](x) = e^{ax} \int_x^{+\infty} e^{-(a+k)t} dt = e^{ax} \frac{e^{-(a+k)x}}{a+k} = \frac{e^{-kx}}{a+k}.$$

Et donc $\boxed{U(f_k) = \frac{1}{a+k} f_k.}$

3.b. Soit $\lambda \in \left] 0, \frac{1}{a} \right]$, et soit $k = \frac{1}{\lambda} - a > 0$, de sorte que $\lambda = \frac{1}{a+k}$.

Alors nous venons de prouver que $U(f_k) = \frac{1}{a+k} f_k = \lambda f_k$, et donc $f_k \in \text{Ker}(U - \lambda \text{id}_E)$.

On en déduit donc que $\boxed{\text{Ker}(U - \lambda \text{id}_E) \neq \{0\}.}$

3.c. Puisque $U(f_k) = \frac{1}{a+k} f_k$, on a donc $\boxed{U^n(f_k) = \frac{1}{(a+k)^n} f_k.}$

En particulier, pour $x \in I$, $U^n(f_k) = \frac{1}{(a+k)^n} e^{-kx}$.

Si $a+k > 1$, alors $\lim_{n \rightarrow +\infty} [U^n(f_k)](x) = 0$.

Si $a+k = 1$, alors $\lim_{n \rightarrow +\infty} [U^n(f_k)](x) = e^{-kx}$.

Enfin, si $a+k < 1$, alors $\lim_{n \rightarrow +\infty} [U^n(f_k)](x) = +\infty$.

4. Cas des fonctions sinus et cosinus.

4.a. On a $[U(\cos)](x) = e^x \int_x^{+\infty} e^{-t} \cos t dt$.

Or, sur un segment de la forme $[x, A]$, une intégration par parties nous donne

$$\int_x^A e^{-t} \cos t dt = [e^{-t} \sin t]_x^A + \int_x^A e^{-t} \sin t dt = -e^{-x} \sin x + e^{-A} \sin A + \int_x^A e^{-t} \sin t dt.$$

Or, lorsque $A \rightarrow +\infty$, $e^{-A} \rightarrow 0$ et la fonction \sin étant bornée, $e^{-A} \sin A \xrightarrow{A \rightarrow +\infty} 0$.

Et donc $\int_x^{+\infty} e^{-t} \cos t dt = -e^{-x} \sin x + \int_x^{+\infty} e^{-t} \sin t dt$.

Ainsi, $[U(\cos)](x) = -\sin x + [U(\sin)](x)$.

De même, on prouve à l'aide d'une intégration par parties que

$$[U(\sin)](x) = \cos(x) - [U(\cos)](x).$$

Valeurs propres

Ceci est un exemple de ce qui peut se produire pour des endomorphismes d'espaces vectoriels de dimension infinie (ce qui est le cas de E) : ils peuvent posséder une infinité de valeurs propres distinctes. Notons au passage que cela prouve que E est de dimension infinie : dans le cas contraire, il n'existerait qu'un nombre fini de $\lambda \in \mathbf{R}$ pour lesquels $\text{Ker}(U - \lambda \text{id}_E) \neq \{0\}$.

Et donc, il vient

$$\begin{cases} U(\cos) = -\sin + U(\sin) \\ U(\sin) = \cos - U(\cos) \end{cases} \Leftrightarrow \begin{cases} U(\cos) = \frac{\cos - \sin}{2} \\ U(\sin) = \frac{\cos + \sin}{2} \end{cases}$$

Soit $f \in P$. Alors il existe deux réels α et β tels que $f = \alpha \cos + \beta \sin$.

Et donc $U(f) = \alpha \underbrace{U(\cos)}_{\in P} + \beta \underbrace{U(\sin)}_{\in P} \in P$.

Ainsi, P est stable par U .

Par définition de P , la famille (\sin, \cos) en est génératrice. D'autre part, si λ_1, λ_2 sont deux réels tels que $\lambda_1 \cos + \lambda_2 \sin = 0$, alors en évaluant cette relation en $x = 0$, il vient $\lambda_1 = 0$.

Et donc $\lambda_2 \sin = 0$. En évaluant en $x = \frac{\pi}{2}$, il vient donc $\lambda_2 = 0$.

Ainsi, la famille (\cos, \sin) est libre : c'est une base de P .

Et d'après les relations montrées précédemment,

$$M = \text{Mat}_{(\sin, \cos)}(U) = \begin{pmatrix} \overset{U(\sin)}{\frac{1}{2}} & \overset{U(\cos)}{-\frac{1}{2}} \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix} \begin{matrix} \sin \\ \cos \end{matrix}.$$

4.b. On a

$$M^2 = \frac{1}{2} \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, M^3 = \frac{1}{4} \begin{pmatrix} -1 & -1 \\ 1 & -1 \end{pmatrix}, M^4 = -\frac{1}{4} I_2.$$

Il est alors évident que si $n = 4p$ est multiple de p , on a

$$M^n = M^{4p} = (M^4)^p = \frac{(-1)^p}{4^p} I_2.$$

Sinon, il faut distinguer suivant la valeur du reste lors de la division euclidienne de n par 4 :

- si $n = 4p + 1$, alors $M^n = M^{4p} M = \frac{(-1)^p}{2 \cdot 4^p} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}$
- si $n = 4p + 2$, alors $M^n = M^{4p} M^2 = \frac{(-1)^p}{2 \cdot 4^p} \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$
- si $n = 4p + 3$, alors $M^n = M^{4p} M^3 = \frac{(-1)^p}{4^{p+1}} \begin{pmatrix} -1 & -1 \\ 1 & -1 \end{pmatrix} = \frac{(-1)^{p+1}}{4^{p+1}} \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}.$

5. Une autre famille de fonctions.

5.a. On a, pour tout $x \in I$,

$$[U(\psi_n)](x) = e^{ax} \int_x^{+\infty} e^{-at} e^{-t} t^n dt = e^{ax} \int_x^{+\infty} e^{-(a+1)t} t^n dt.$$

Mais, une intégration par parties sur un segment de la forme $[x, A]$ nous donne

$$\int_x^A e^{-(a+1)t} t^n dt = \left[-\frac{1}{a+1} e^{-(a+1)t} t^n \right]_x^A + \frac{n}{a+1} \int_x^A e^{-(a+1)t} t^{n-1} dt = \frac{x^n}{a+1} e^{-(a+1)x} - \frac{A^n}{a+1} e^{-(a+1)A}.$$

En passant à la limite lorsque $A \rightarrow +\infty$, $A^n e^{-(a+1)A} \xrightarrow{A \rightarrow +\infty} 0$ par croissances comparées, et donc

$$\int_x^{+\infty} e^{-(a+1)t} t^n dt = \frac{x^n}{a+1} e^{-(a+1)x} + \frac{n}{a+1} \int_x^{+\infty} e^{-at} e^{-t} t^{n-1} dt.$$

Après multiplication par e^{ax} , il vient donc $\psi_n(x) = \frac{\varphi_n(x)}{a+1} + \frac{n}{a+1} \psi_{n-1}(x)$.

5.b. Notons que le résultat précédent n'est valable que pour $n \geq 1$, et que pour $n = 0$, on a, avec les notations de la question 3, $\varphi_0 = f_1$, et donc en particulier

$$\psi_0 = U(\varphi_0) = \frac{1}{a+1} f_0 = \frac{1}{a+1} \varphi_0.$$

Plus généralement

Un sous-espace vectoriel F est stable par f si et seulement si l'image par f d'une base de F est encore dans F .

Montrons par récurrence sur p que $\psi_p \in F_p$.
 Pour $p = 0$, nous venons de prouver que $\psi_0 \in \text{Vect}(\varphi_0) = F_0$.
 Supposons donc que $\psi_p \in F_p$. Alors

$$\psi_{p+1} = \frac{1}{a+1} \underbrace{\varphi_{p+1}}_{\in F_{p+1}} + \frac{p+1}{a+1} \underbrace{\psi_p}_{\in F_p \subset F_{p+1}} \in F_{p+1}.$$

Ainsi, pour tout $p \in \mathbf{N}$, $\psi_p \in F_p$.

Soit à présent $p \in \mathbf{N}$ fixé, et soit $f = \sum_{i=0}^p \lambda_i \varphi_i \in F_p$.

Alors, par linéarité de U , $U(f) = \sum_{i=0}^p \lambda_i \psi_i$.

Et pour $i \in \llbracket 0, p \rrbracket$, nous venons de prouver que $\psi_i \in F_i \subset F_p$.

Ainsi, $U(f) \in F_p$, et donc F_p est stable par U .

Il est évident que la famille $(\varphi_0, \dots, \varphi_p)$ est une famille génératrice de F_p .

Soient donc $\lambda_0, \dots, \lambda_p$ des réels tels que $\sum_{i=0}^p \lambda_i \varphi_i = 0$.

Alors pour tout $x \in I$, on a

$$\sum_{i=0}^p \lambda_i x^i e^{-x} = 0 \Leftrightarrow \left(\sum_{i=0}^p \lambda_i x^i \right) e^{-x} = 0 \Leftrightarrow \sum_{i=0}^p \lambda_i x^i = 0.$$

Une exponentielle n'est jamais nulle.

Et donc le polynôme $\sum_{i=0}^p \lambda_i X^i$ s'annule en tous les $x \in I$, qui sont en nombre infini : c'est donc le polynôme nul, de sorte que les λ_i sont tous nuls.

Ainsi, la famille $(\varphi_0, \varphi_1, \dots, \varphi_p)$ est libre : c'est une base de F_p .

5.c. Notons $\mathcal{B} = (\varphi_0, \varphi_1, \varphi_2)$. On a alors

$$U(\varphi_0) = \frac{1}{a+1} \varphi_0, \quad U(\varphi_1) = \frac{1}{a+1} \varphi_1 + \frac{1}{a+1} \psi_0 = \frac{1}{a+1} \varphi_1 + \frac{1}{(a+1)^2} \varphi_0$$

$$\text{et } U(\varphi_2) = \frac{1}{a+1} \varphi_2 + \frac{2}{a+1} \psi_1 = \frac{1}{a+1} \varphi_2 + \frac{2}{(a+1)^2} \varphi_1 + \frac{2}{(a+1)^3} \varphi_0.$$

Donc la matrice de U est

$$T_2 = \text{Mat}_{\mathcal{B}}(U_2) = \begin{pmatrix} \overset{U(\varphi_0)}{\frac{1}{a+1}} & \overset{U(\varphi_1)}{\frac{1}{(a+1)^2}} & \overset{U(\varphi_2)}{\frac{2}{(a+1)^3}} \\ 0 & \frac{1}{a+1} & \frac{2}{(a+1)^2} \\ 0 & 0 & \frac{1}{a+1} \end{pmatrix} \begin{matrix} \varphi_0 \\ \varphi_1 \\ \varphi_2 \end{matrix}.$$

Notons que $T_2 = \frac{1}{(a+1)} I_3 + N$, où $N = \begin{pmatrix} 0 & \frac{1}{(a+1)^2} & \frac{2}{(a+1)^3} \\ 0 & 0 & \frac{2}{(a+1)^3} \\ 0 & 0 & 0 \end{pmatrix}$.

On a alors $N^2 = \begin{pmatrix} 0 & 0 & \frac{2}{(a+1)^4} \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$, et pour $k \geq 3$, $N^k = 0$.

Puisque N et I_3 commutent³, par la formule du binôme de Newton, on a alors

$$T_2^n = \left(\frac{1}{a+1} I_3 + N \right)^n = \sum_{k=0}^n \binom{n}{k} N^k \frac{1}{(a+1)^{n-k}} I_3 = \frac{1}{(a+1)^n} I_3 + \frac{n}{(a+1)^{n-1}} N + \frac{n(n-1)}{2(a+1)^{n-2}} N^2.$$

³ L'identité commute avec toute matrice.

$N^k = 0$ pour $k \geq 3$.

Soit encore

$$T_2^n = \begin{pmatrix} \frac{1}{(a+1)^n} & \frac{n}{(a+1)^{n+1}} & \frac{n(n+1)}{(a+1)^{n+2}} \\ 0 & \frac{1}{(a+1)^n} & \frac{n}{(a+1)^{n+1}} \\ 0 & 0 & \frac{1}{(a+1)^n} \end{pmatrix}.$$

En particulier, tous les coefficients de T_2^n tendent vers 0 lorsque n tend vers l'infini.

6. Une autre expression de $U(f)$.

D'après ce qui a été dit précédemment, pour tout $x \in I$, on a

$$[U(f)](x) = e^{ax} \int_x^{+\infty} e^{-at} f(t) dt = \int_x^{+\infty} e^{-a(t-x)} f(t) dt.$$

À x fixé, procédons alors au changement de variable affine $u = t - x \Leftrightarrow t = u + x$, de sorte que lorsque $t \rightarrow x$, $u \rightarrow 0$ et lorsque $t \rightarrow +\infty$, $x \rightarrow +\infty$. Il vient alors

$$[U(f)](x) = \int_0^{+\infty} e^{-au} f(u+x) du = \int_0^{+\infty} e^{-at} f(t+x) dt.$$

Remarque

Cette dernière intégrale converge puisque l'intégrale avant changement de variable converge.

7. Positivité de U .

7.a. Commençons par remarquer que si $f \in E$, alors $|f|$ est également dans E , car continue⁴ sur I et bornée.

Et donc $U(|f|)$ existe, de sorte que $\int_x^{+\infty} e^{-at} |f(t)| dt$ converge pour tout $x \in I$.

Par l'inégalité triangulaire, on a donc, pour tout $x \in I$,

$$\left| \int_x^{+\infty} e^{-at} f(t) dt \right| \leq \int_x^{+\infty} |e^{-at} f(t)| dt = \int_x^{+\infty} e^{-at} |f(t)| dt.$$

Après multiplication par $e^{ax} \geq 0$, on a donc

$$|[U(f)](x)| = \left| e^{ax} \int_x^{+\infty} e^{-at} f(t) dt \right| \leq e^{ax} \int_x^{+\infty} e^{-at} |f(t)| dt = [U(|f|)](x).$$

Et ceci étant vrai pour tout $x \in I$, on a bien $|U(f)| \leq U(|f|)$.

7.b. Pour $x \in I$, on a $\psi(x) = \int_0^{+\infty} e^{-at} \varphi(t+x) dt$.

Mais la fonction $t \mapsto e^{-at} \varphi(t+x)$ est positive sur \mathbf{R}_+ , et donc, par positivité de l'intégrale,

$$\psi(x) \geq 0.$$

7.c. Soit $x \in I$. Alors $a\psi(x) = ae^{ax} \int_x^{+\infty} e^{-at} \varphi(t) dt$.

Mais par décroissance de φ , pour tout $t \in]x, +\infty[$, $\varphi(t) \leq \varphi(x)$ et donc par croissance de l'intégrale,

$$a\psi(x) = ae^{ax} \int_x^{+\infty} e^{-at} \varphi(t) dt \leq ae^{ax} \int_x^{+\infty} e^{-at} \varphi(x) dt \leq \varphi(x) a \underbrace{e^{ax} \int_x^{+\infty} e^{-at} dt}_{=[U(1)](x)} = \varphi(x).$$

Donc on a bien $a\psi \leq \varphi$.

Soient $x < y$ deux réels dans I . Alors $\psi(x) = \int_0^{+\infty} e^{-at} \varphi(t+x) dt$ et $\psi(y) = \int_0^{+\infty} e^{-at} \varphi(t+y) dt$.

Mais par décroissance de φ , pour tout $t \geq 0$, $\varphi(t+x) \geq \varphi(t+y)$ et donc par croissance de l'intégrale, $\psi(x) \geq \psi(y)$: la fonction ψ est décroissante.

8. Commutation de U avec la dérivation.

8.a. Pour $x \in I$, on a

$$a[U(f)](x) = \int_0^{+\infty} ae^{-at} f(x+t) dt.$$

Procédons à une intégration par parties sur un segment de la forme $[0, A]$ en posant $u(t) = f(x+t)$ et $v(t) = -e^{-at}$ qui sont deux fonctions de classe \mathcal{C}^1 sur $[0, A]$, avec $u'(t) = f'(x+t)$ et $v'(t) = ae^{-at}$. Alors

$$\int_0^A ae^{-at} f(x+t) dt = [-f(x+t)e^{-at}]_0^A + \int_0^A e^{-at} f'(x+t) dt$$

⁴ Par composition de fonctions continues.

$$= f(x) - f(x+A)e^{-aA} + \int_0^A e^{-at} f'(x+t) dt.$$

Et donc en passant à la limite lorsque $A \rightarrow +\infty$, il vient

$$\int_0^{+\infty} ae^{-at} f(x+t) dt = f(x) + \int_0^{+\infty} e^{-at} f'(x+t) dt$$

soit encore $a[U(f)](x) = f(x) + [U(f')](x)$.

Ceci étant vrai pour tout $x \in I$, on en déduit que $aU(f) = f + U(f')$.

- 8.b. Rappelons que par définition de U , $U(f)$ est une fonction de classe \mathcal{C}^1 . Et puisque f' est bornée, $f' \in E$ et donc $U(f')$ est également de classe \mathcal{C}^1 .

En dérivant la relation de la question précédente, on obtient

$$aD(U(f)) = f' + D(U(f')).$$

Or, par définition, $U(f')$ est une solution de l'équation $(E_{f'})$ de sorte que

$$D(U(f')) - aU(f') + f' = 0 \Leftrightarrow D(U(f')) = aU(f') - f'.$$

Et donc

$$aD(U(f)) = f' + aU(f') - f' = aU(f') = aU(D(f)).$$

Puisque $a \neq 0$, on en déduit que $D(U(f)) = U(D(f))$.

- 8.c. Si f est positive et décroissante, f' est négative, et donc $D(U(f)) = U(f') = -U(-f')$. Or $-f' \geq 0$ et donc par la question 7.b, $U(f') \geq 0$.

On en déduit que $D(U(f)) \leq 0$, et donc $U(f)$ est décroissante.

II. Comportement asymptotique de $U(f)$ au voisinage de $+\infty$.

9. Résultats préliminaires.

- 9.a. Par hypothèse, nous avons $\lim_{x \rightarrow +\infty} \frac{\alpha(x)}{\beta(x)} = 0$.

Et donc pour tout $\varepsilon > 0$, il existe $A > 0$ tel que pour $t \geq A$, $\left| \frac{\alpha(t)}{\beta(t)} \right| \leq \varepsilon \Leftrightarrow |\alpha(t)| \leq \varepsilon \beta(t)$.

Par croissance de l'intégrale, on en déduit que pour tout $x \geq A$,

$$\int_x^{+\infty} |\alpha(t)| dt \leq \varepsilon \int_x^{+\infty} \beta(t) dt.$$

Et alors, par l'inégalité triangulaire, pour $x \geq A$,

$$\left| \int_x^{+\infty} \alpha(t) dt \right| \leq \int_x^{+\infty} |\alpha(t)| dt \leq \varepsilon \int_x^{+\infty} \beta(t) dt.$$

On en déduit que pour $x \geq A$,

$$\left| \frac{\int_x^{+\infty} \alpha(t) dt}{\int_x^{+\infty} \beta(t) dt} \right| \leq \varepsilon.$$

Et donc, sen revenant à la définition d'une limite, nous venons de prouver que

$$\lim_{x \rightarrow +\infty} \frac{\int_x^{+\infty} \alpha(t) dt}{\int_x^{+\infty} \beta(t) dt} = 0 \iff \int_x^{+\infty} \alpha(t) dt = o_{x \rightarrow +\infty} \left(\int_x^{+\infty} \beta(t) dt \right).$$

- 9.b. Par définition d'un équivalent, on a $\alpha(x) = \beta(x) + \gamma(x)$, où $\gamma(x) = o_{x \rightarrow +\infty}(\beta(x))$.

Et donc, en appliquant le résultat de la question précédente, il vient

$$\int_x^{+\infty} \alpha(t) dt = \int_x^{+\infty} \beta(t) dt + \int_x^{+\infty} \gamma(t) dt$$

Remarque

◀ Cette dernière intégrale converge car $f' \in E$.

Notation

Nous notons D pour la dérivée, comme nous l'indique l'énoncé, mais remarquons que $D(U(f)) = [U(f)]'$ et

$$D(U(f')) = [U(f')]'$$

Convergence

◀ La convergence de ces intégrales est garantie par le fait que $\int_1^{+\infty} \beta(t) dt$ converge.

$$= \int_x^{+\infty} \beta(t) dt + \underset{x \rightarrow +\infty}{o} \left(\int_x^{+\infty} \beta(t) dt \right)$$

$$\boxed{\underset{x \rightarrow +\infty}{\sim} \int_x^{+\infty} \beta(t) dt.}$$

10. Cas des fonctions admettant une limite en $+\infty$.

Comme indiqué, commençons par supposer que $b = 0$, c'est-à-dire que $f(t) = \underset{t \rightarrow +\infty}{o} (1)$.

Alors $e^{-at} f(t) = \underset{t \rightarrow +\infty}{o} (e^{-at})$ et donc par la question précédente,

$$\int_x^{+\infty} e^{-at} f(t) dt = \underset{x \rightarrow +\infty}{o} \left(\int_x^{+\infty} e^{-at} dt \right).$$

Or nous avons déjà prouvé⁵ que $\int_x^{+\infty} e^{-at} dt = \frac{1}{a} e^{-ax}$.

⁵ Voir question 1.d.

Après multiplication par e^{ax} , on a donc

$$[U(f)](x) = \underset{x \rightarrow +\infty}{o} \left(e^{ax} \frac{1}{a} e^{-ax} \right) = \underset{x \rightarrow +\infty}{o} (1).$$

Et donc $\boxed{\lim_{x \rightarrow +\infty} [U(f)](x) = 0.}$

Passons à présent au cas général, en posant $f_1 = f - b$, de sorte que $f_1 \in E$ et

$$\lim_{x \rightarrow +\infty} f_1(x) = \lim_{x \rightarrow +\infty} f(x) - b = b - b = 0.$$

Par ce qui précède, on a alors $\lim_{x \rightarrow +\infty} [U(f_1)](x) = 0$.

Or, par linéarité de U , $U(f_1) = U(f) - bU(1) = U(f) - \frac{b}{a}$.

Et donc $\lim_{x \rightarrow +\infty} [U(f)](x) = \lim_{x \rightarrow +\infty} [U(f_1)](x) + \frac{b}{a} = \boxed{\frac{b}{a}}.$

Rappel

↳ Nous avons calculé $U(1)$ à la question 2.a.

11. Cas des fonctions puissances.

11.a. On a, pour tout $x \in I$, par intégration par parties,

$$\int_x^{+\infty} e^{-at} \frac{1}{t^\omega} dt = \lim_{A \rightarrow +\infty} \left(\left[-\frac{1}{a} e^{-at} \frac{1}{t^\omega} \right]_x^A - \int_x^A \frac{\omega}{a} e^{-at} \frac{1}{t^{\omega+1}} dt \right) = \frac{1}{a} e^{-ax} \frac{1}{x^\omega} - \frac{\omega}{a} \int_x^{+\infty} e^{-at} \frac{1}{t^\omega} dt.$$

Et donc après multiplication par e^{ax} , on a donc

$$U(f_\omega)(x) = \frac{1}{a} \frac{1}{x^\omega} - \frac{\omega}{a} e^{ax} \int_x^{+\infty} e^{-at} \frac{1}{t^{\omega+1}} dt = \frac{1}{a} f_\omega(x) - \frac{\omega}{a} g_{\omega+1}(x).$$

Et donc on a bien $\boxed{g_\omega = \frac{1}{a} f_\omega - \frac{\omega}{a} g_{\omega+1}.}$

Puisque $f_{\omega+1}(x) = \underset{x \rightarrow +\infty}{o} (f_\omega(x))$, on a $e^{-at} f_{\omega+1}(t) = \underset{t \rightarrow +\infty}{o} (e^{-at} f_\omega(t))$, et donc d'après 9.a,

$$g_{\omega+1}(x) = e^{ax} \int_x^{+\infty} e^{-at} f_{\omega+1}(t) dt = \underset{x \rightarrow +\infty}{o} \left(e^{ax} \int_x^{+\infty} e^{-at} f_\omega(t) dt \right) = \underset{x \rightarrow +\infty}{o} (g_\omega(x)).$$

On a donc $\frac{f_\omega(x)}{a} = g_\omega(x) + \underset{x \rightarrow +\infty}{o} (g_\omega(x))$ et donc $\boxed{\frac{f_\omega(x)}{a} \underset{x \rightarrow +\infty}{\sim} g_\omega(x).}$

11.b. Comme indiqué, appliquons l'inégalité de Taylor-Lagrange à la fonction $f_a : u \mapsto e^{-au}$ entre 0 et $t \in I$.

f_a est de classe \mathcal{C}^∞ sur $[0, t]$ et pour tout $k \in \mathbf{N}$, $f_a^{(k)}(u) = (-1)^k a^k e^{-au}$.

En particulier, pour tout $u \in [0, t]$, $|f_a^{(k)}(u)| = a^k e^{-au} \leq a^k$.

Et donc, pour tout $t \in I$

$$\left| f_a(t) - \sum_{k=0}^n \frac{f_a^{(k)}(0)}{k!} t^k \right| \leq \frac{t^{n+1} a^{n+1}}{(n+1)!}.$$

Soit encore

$$\left| e^{-at} - \sum_{k=0}^n \frac{(-1)^k a^k}{k!} t^k \right| \leq \frac{t^{n+1} a^{n+1}}{(n+1)!}.$$

En divisant par $t \neq 0$, il vient

$$\left| \frac{e^{-at}}{t} - \sum_{k=0}^n \frac{(-1)^k a^k}{k!} t^{k-1} \right| \leq \frac{t^n a^{n+1}}{(n+1)!}.$$

Et donc par croissance de l'intégrale, pour tout $x \in I$,

$$\int_1^x \left| \frac{e^{-at}}{t} - \sum_{k=0}^n \frac{(-1)^k a^k}{k!} t^{k-1} \right| dt \leq \int_1^x \frac{t^n a^{n+1}}{(n+1)!} dt = a^{n+1} \frac{x^{n+1} - 1}{(n+1)(n+1)!}.$$

D'autre part, on a

$$\begin{aligned} \int_1^x \frac{e^{-at}}{t} dt - \ln(x) - \sum_{k=1}^n \frac{(-1)^k a^k}{k \cdot k!} (x^k - 1) &= \int_1^x \frac{e^{-at}}{t} dt - \int_1^x \frac{dt}{t} - \sum_{k=1}^n \frac{(-1)^k a^k}{k!} \int_1^x t^{k-1} dt \\ &= \int_1^x \left(\frac{e^{-at}}{t} - \frac{1}{t} - \sum_{k=1}^n \frac{(-1)^k a^k}{k!} t^{k-1} \right) dt. \end{aligned}$$

Et donc par l'inégalité triangulaire,

$$\begin{aligned} \left| \int_1^x \frac{e^{-at}}{t} dt - \ln(x) - \sum_{k=1}^n \frac{(-1)^k a^k}{k \cdot k!} (x^k - 1) \right| &= \left| \int_1^x \left(\frac{e^{-at}}{t} - \sum_{k=0}^n \frac{(-1)^k a^k}{k!} t^{k-1} \right) dt \right| \\ &\leq \int_1^x \left| \frac{e^{-at}}{t} - \sum_{k=0}^n \frac{(-1)^k a^k}{k!} t^{k-1} \right| dt \\ &\leq a^{n+1} \frac{x^{n+1} - 1}{(n+1)(n+1)!} \\ &\xrightarrow{n \rightarrow +\infty} 0. \end{aligned}$$

On en déduit donc que la série de terme général $\frac{(-1)^k a^k}{k \cdot k!} (x^k - 1)$ converge⁶ et que

⁶ Car la suite de ses sommes partielles admet une limite.

$$\boxed{\sum_{k=0}^{+\infty} \frac{(-1)^k a^k}{k \cdot k!} (x^k - 1) = \int_1^x \frac{e^{-at}}{t} dt - \ln(x).}$$

On en déduit que pour tout $x \in I$,

$$\begin{aligned} g_1(x) &= e^{ax} \int_x^{+\infty} e^{-at} f_1(t) dt = e^{ax} \int_x^{+\infty} \frac{e^{-at}}{t} dt \\ &= e^{ax} \left(\int_1^{+\infty} \frac{e^{-at}}{t} dt - \int_1^x \frac{e^{-at}}{t} dt \right) \\ &= \boxed{e^{ax} \left(\int_1^{+\infty} \frac{e^{-at}}{t} dt - \ln(x) - \sum_{k=1}^{+\infty} \frac{(-1)^k a^k}{k \cdot k!} (x^k - 1) \right)}. \end{aligned}$$

12. Cas des fonctions comparables aux fonctions puissances f_ω .

12.a. Si $f = o_{x \rightarrow +\infty}(f_\omega)$, alors $e^{-at} f(t) = o_{t \rightarrow +\infty}(e^{-at} f_\omega(t))$ et alors le résultat de la question 9.a s'applique car $t \mapsto e^{-at} f_\omega$ est strictement positive et $\int_1^{+\infty} e^{-at} f_\omega(t) dt$ converge.

Et donc $\int_x^{+\infty} e^{-at} f(t) dt = o_{x \rightarrow +\infty} \left(\int_x^{+\infty} e^{-at} f_\omega(t) dt \right)$, de sorte qu'après multiplication par e^{ax} ,

$$g(x) = e^{ax} \int_x^{+\infty} e^{-at} f(t) dt = o_{x \rightarrow +\infty} \left(e^{ax} \int_x^{+\infty} e^{-at} f_\omega(t) dt \right) = \boxed{o_{x \rightarrow +\infty}(g_\omega(x))}.$$

Détails

Cette dernière intégrale converge car

$$e^{-at} f_\omega(t) = o_{t \rightarrow +\infty}(e^{-at}).$$

- 12.b. De même, si $f(t) \underset{t \rightarrow +\infty}{\sim} f_\omega(t)$, alors $f(t) = f_\omega(t) + \underset{t \rightarrow +\infty}{o}(f_\omega(t))$.
Et donc en multipliant par e^{-at} , $f(t)e^{-at} = f_\omega(t)e^{-at} + \underset{t \rightarrow +\infty}{o}(e^{-at}f_\omega(t))$. Par la question 9.b, on a alors

$$\int_x^{+\infty} e^{-at} f(t) dt = \int_x^{+\infty} e^{-at} f_\omega(t) dt + \underset{x \rightarrow +\infty}{o} \left(\int_x^{+\infty} f_\omega(t) e^{-at} dt \right).$$

Et donc après multiplication par e^{ax} ,

$$g(x) = g_\omega(x) + \underset{x \rightarrow +\infty}{o}(g_\omega(x)) \underset{x \rightarrow +\infty}{\sim} g_\omega(x).$$

Mais d'après la question 11.a, $g_\omega(x) \underset{x \rightarrow +\infty}{\sim} \frac{f_\omega(x)}{a}$.

On en déduit donc que

$$g(x) \underset{x \rightarrow +\infty}{\sim} \frac{f_\omega(x)}{a} \left[\underset{x \rightarrow +\infty}{\sim} \frac{f(x)}{a} \right].$$

III. Convergence absolue de $\int_1^{+\infty} U(f)$.

13. Étude d'exemples.

- 13.a. Nous avons déjà prouvé que $g_k = U(f_k) = \frac{1}{a+k} f_k$.

Et alors $\int_1^{+\infty} f_k(t) dt$ est une intégrale de référence convergente, de sorte que $\int_1^{+\infty} g_k(t) dt$ converge.

- 13.b. Nous avons prouvé à la question 11.a que $g_\omega(t) \underset{t \rightarrow +\infty}{\sim} \frac{f_\omega(t)}{a}$.

Or, $\int_1^{+\infty} f_\omega(t) dt = \int_1^{+\infty} \frac{1}{t^\omega} dt$ converge si et seulement si $\omega > 1$.

D'après le critère des équivalents pour les intégrales de fonctions positives⁷, $\int_1^{+\infty} g_\omega(t) dt$ converge si et seulement si $\omega > 1$.

⁷ D'après le résultat de la question 7.b que g_ω est positive.

14. Cas des fonctions positives.

- 14.a. Puisque g est solution de (E_f) , alors, pour tout $t \in E$, $g'(t) - ag(t) = -f(t)$.
En intégrant cette relation entre 1 et x , il vient

$$\int_1^x g'(t) dt - a \int_1^x g(t) dt = - \int_1^x f(t) dt \Leftrightarrow g(x) - g(1) - aG(x) = -F(x).$$

De plus, g étant continue sur I , G est \mathcal{C}^1 sur I et $G'(x) = g(x)$. Et donc

$$\boxed{G'(x) - aG(x) = -F(x) + g(1)}.$$

- 14.b. Puisque F est la primitive de f qui s'annule en 1, elle est continue.
Pour montrer que F est dans E , il s'agit donc de prouver qu'elle est bornée. Notons que par croissance de l'intégrale, F est à valeurs positives.

D'autre part, puisque $\int_1^{+\infty} f(t) dt$ est convergente, on a $\lim_{x \rightarrow +\infty} F(x) = \int_1^{+\infty} f(t) dt$.

Et donc il existe $A \geq 1$ tel que pour $x \geq A$,

$$\left| F(x) - \int_1^{+\infty} f(t) dt \right| \leq 1 \implies F(x) \leq 1 + \int_1^{+\infty} f(t) dt.$$

Comme de plus, F est continue sur le segment $[1, A]$ elle y est bornée : il existe $M > 0$ tel que pour tout $x \in [1, A]$, $|F(x)| \leq M$.

Ainsi, pour tout $x \in I$, on a

$$F(x) \leq \max \left(M, 1 + \int_1^{+\infty} f(t) dt \right).$$

Et donc F est bornée, et donc $\boxed{\text{est un élément de } E.}$

Le résultat de la question 14.a montre que G est une solution de l'équation $(E_{F-g(1)})$. Et donc par la question 1.a, il existe $K \in \mathbf{R}$ tel que

$$\forall x \in I, G(x) = Ke^{ax} - e^{ax} \int_1^x e^{-at} (F(t) - g(1)) dt = Ke^{-ax} + [U(F)](x) - [U(g(1))](x).$$

Mais $U(g(1)) = g(1)U(1) = \frac{g(1)}{a}$, de sorte que

$$\boxed{\forall x \in I, G(x) = Ke^{ax} + [U(F)](x) - \frac{g(1)}{a}.}$$

14.c. Puisque $g = U(f)$ est dans E , elle est bornée : il existe $M > 0$ tel que pour tout $x \in I$, $|g(x)| \leq M$. Et alors, par croissance de l'intégrale,

$$|G(x)| = \left| \int_1^x g(t) dt \right| \leq \int_1^x |g(t)| dt \leq \int_1^x M dt \leq M(x-1) \leq Mx.$$

On en déduit donc que pour tout $x \in I$,

$$\left| \frac{G(x)}{x} \right| \leq \frac{Mx}{x} \leq M.$$

Et donc $\boxed{g \text{ est bornée sur } I.}$

14.d. Par définition de U , $U(F)$ est bornée sur I .

Si $K \neq 0$, alors $\lim_{x \rightarrow +\infty} K \frac{e^{ax}}{x} = \pm\infty$, suivant le signe de K . Et donc

$$\lim_{x \rightarrow +\infty} \frac{G(x)}{x} = \lim_{x \rightarrow +\infty} \frac{Ke^{ax}}{x} + [U(F)](x) - \frac{g(1)}{a} = \pm\infty,$$

contredisant le fait que $x \mapsto \frac{G(x)}{x}$ est bornée.

On en déduit donc que $\boxed{K = 0.}$ Et donc $\boxed{G = U(F) - \frac{g(1)}{a}.}$

14.e. Pour tout $x \in I$, on a

$$\int_1^x g(t) dt = G(x) = U(F)(x) - \frac{g(1)}{a}.$$

Mais puisque $U(F)$ est bornée, G est également bornée, et en particulier majorée.

Or, g étant positive, $G : x \mapsto \int_1^x g(t) dt$ est croissante⁸. Puisqu'elle est bornée, elle admet une limite en $+\infty$, et donc $\lim_{x \rightarrow +\infty} \int_1^x g(t) dt$ existe.

Ceci prouve donc que $\boxed{\int_1^{+\infty} g(t) dt \text{ converge.}}$

15. Cas général

Nous avons prouvé à la question 7.a que $|U(f)| \leq U(|f|)$.

Or, $|f|$ est une fonction positive, et par hypothèse, $\int_1^{+\infty} |f(t)| dt$ converge.

D'après le résultat de la question 14, $\int_1^{+\infty} [U(|f|)](t) dt$ est donc convergente.

Et alors, par critère de comparaison pour les intégrales de fonctions positives, $\int_1^{+\infty} |[U(f)](t)| dt$

converge, de sorte que $\boxed{\int_1^{+\infty} g(t) dt \text{ converge absolument.}}$

⁸ Par le théorème fondamental de l'analyse, sa dérivée est g .

DEVOIR MAISON 5 (À RENDRE LE 18.10.18)

On considère la matrice $A = \begin{pmatrix} 1 & 6 \\ 1 & 0 \end{pmatrix}$, et on note f l'unique endomorphisme de \mathbf{R}^2 dont la matrice dans la base canonique \mathcal{B} est A .

1. On suppose dans cette question que A est semblable à une matrice diagonale $D = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}$, avec $\lambda_1 \geq \lambda_2$. Autrement dit, on suppose qu'il existe $P_1 \in GL_2(\mathbf{R})$ telle que $A = P_1^{-1}DP_1$.

a. Calculer A^2 . Vérifier que $A^2 = A + 6I_2$, et en déduire un polynôme $P \in \mathbf{R}_2[X]$, non nul, annulateur de A .

b. Montrer que P est annulateur de D , et en déduire que λ_1 et λ_2 sont racines de P .

c. Calculer la trace de A , et en déduire que $D = \begin{pmatrix} 3 & 0 \\ 0 & -2 \end{pmatrix}$.

2. a. Montrer qu'il existe une base $\mathcal{B}' = (e_1, e_2)$ de \mathbf{R}^2 telle que $\text{Mat}_{\mathcal{B}'}(f) = D = \begin{pmatrix} 3 & 0 \\ 0 & -2 \end{pmatrix}$.

b. En déduire que A et D sont semblables, et déterminer une matrice P_1 inversible telle que $A = P_1^{-1}DP_1$.

c. Pour $n \in \mathbf{N}$, calculer D^n , puis en déduire que :

$$A^n = \frac{1}{5} \begin{pmatrix} 3^{n+1} - (-2)^{n+1} & 2 \cdot 3^{n+1} + 3 \cdot (-2)^{n+1} \\ 3^n - (-2)^n & 2 \cdot 3^n + 3 \cdot (-2)^n \end{pmatrix}.$$

On considère la suite $(u_n)_{n \in \mathbf{N}}$ définie par ses deux premiers termes $u_0 = -\frac{1}{2}$ et $u_1 = 1$, et pour tout $n \in \mathbf{N}$:

$$u_{n+2} = u_{n+1} + 6u_n + 3.$$

3. Calculer u_2 et u_3 .

4. On considère le programme Scilab suivant

```

1 n = input('n ?')
2 v = -1/2, u=1
3 for i=2 :n
4     a=u
5     .....
6     v=a
7 end
8 disp(u)
```

Quelle instruction faut-il ajouter en ligne 5 pour que ce programme calcule et affiche la valeur de u_n , l'entier n non nul étant donné par l'utilisateur ?

5. On pose $B = \begin{pmatrix} 3 \\ 0 \end{pmatrix}$, $L = \begin{pmatrix} -\frac{1}{2} \\ -\frac{1}{2} \end{pmatrix}$ et pour tout entier naturel n , $X_n = \begin{pmatrix} u_{n+1} \\ u_n \end{pmatrix}$.

a. Montrer que pour tout entier naturel n , on a : $X_{n+1} = AX_n + B$.

b. Vérifier que $AL + B = L$.

c. Montrer que pour tout entier naturel n on a : $X_n = A^n(X_0 - L) + L$.

d. Déduire des questions précédentes que pour tout entier naturel n on a :

$$u_n = \frac{3}{10} (3^n - (-2)^n) - \frac{1}{2}.$$

CORRECTION DU DEVOIR MAISON 5

ADAPTÉ DE BSB ECT 2018

$$1.a. \text{ On a } A^2 = \begin{pmatrix} 1 & 6 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 6 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 7 & 6 \\ 1 & 6 \end{pmatrix}.$$

Et donc on a bien $A^2 = A + 6I_2 \Leftrightarrow A^2 - A - 6I_2 = 0$, de sorte que le polynôme

$$P(X) = X^2 - X - 6 \text{ est annulateur de } A.$$

1.b. Puisque $A = P_1^{-1}DP_1$, il vient donc

$$0 = P(A) = P_1^{-1}P(D)P_1.$$

Et donc $P(D)$ est semblable à la matrice nulle. Or, la seule matrice semblable à la matrice nulle est la matrice nulle, donc $P(D) = 0$: P est annulateur de D .

Mais d'autre part,

$$P(D) = D^2 - D - 6I_2 = \begin{pmatrix} \lambda_1^2 - \lambda_1 - 6 & 0 \\ 0 & \lambda_2^2 - \lambda_2 - 6 \end{pmatrix} = \begin{pmatrix} P(\lambda_1) & 0 \\ 0 & P(\lambda_2) \end{pmatrix}.$$

Donc par identification des coefficients¹, on doit avoir $P(\lambda_1) = P(\lambda_2) = 0$: λ_1 et λ_2 sont deux racines de P .1.c. Il est clair que $\text{tr}(A) = 1$.Mais D étant semblable à A , on a également $\text{tr}(D) = 1 \Leftrightarrow \lambda_1 + \lambda_2 = 1$.D'autre part, $P(X) = X^2 - X - 6 = (X - 3)(X + 2)$, et donc les racines de P sont 3 et -2 , de sorte que $\lambda_1, \lambda_2 \in \{3, -2\}$.Et donc $(\lambda_1, \lambda_2) \in \{(3, 3), (3, -2), (-2, 3), (-2, -2)\}$.Mais $\lambda_1 + \lambda_2 = 1$, donc les λ_i ne peuvent valoir tous les deux 3, ni tous les deux -2 .Enfin, nous savons que $\lambda_1 \geq \lambda_2$, il n'y a qu'une solution possible : $\lambda_1 = 3$ et $\lambda_2 = -2$.

$$\text{Et donc nécessairement, } D = \begin{pmatrix} 3 & 0 \\ 0 & -2 \end{pmatrix}.$$

$$2.a. \text{ On a } \text{Mat}_{\mathcal{B}'}(f) = \begin{pmatrix} 3 & 0 \\ 0 & -2 \end{pmatrix} \begin{matrix} f(e_1) \\ f(e_2) \\ e_1 \\ e_2 \end{matrix} \text{ si et seulement si } \begin{cases} f(e_1) = 3e_1 \\ f(e_2) = -2e_2 \end{cases}.$$

Montrons donc qu'il existe bien une base $\mathcal{B}' = (e_1, e_2)$ vérifiant ces deux conditions.Soit $e_1 = (x, y) \in \mathbf{R}^2$. Alors

$$f(e_1) = 3e_1 \Leftrightarrow (f(e_1))_{\mathcal{B}} = (3e_1)_{\mathcal{B}} \Leftrightarrow A \begin{pmatrix} x \\ y \end{pmatrix} = 3 \begin{pmatrix} x \\ y \end{pmatrix} \Leftrightarrow \begin{cases} x + 6y = 3x \\ x = 3y \end{cases} \Leftrightarrow x = 3y \Leftrightarrow (x, y) = y(3, 1) \in \text{Vect}(3, 1).$$

On peut donc prendre $e_1 = (3, 1)$.De même, pour $e_2 = (x, y) \in \mathbf{R}^2$, on a $f(e_2) = -2e_2$ si et seulement si

$$A \begin{pmatrix} x \\ y \end{pmatrix} = -2 \begin{pmatrix} x \\ y \end{pmatrix} \Leftrightarrow \begin{cases} x + 6y = -2x \\ x = -2y \end{cases} \Leftrightarrow x = -2y \Leftrightarrow (x, y) = y(-2, 1) \in \text{Vect}(-2, 1).$$

On peut donc prendre $e_2 = (-2, 1)$.La famille (e_1, e_2) est alors libre car formée de deux vecteurs non colinéaires : c'est une base \mathcal{B}' de \mathbf{R}^2 dans laquelle $\text{Mat}_{\mathcal{B}'}(f) = D$.

2.b. Par la formule de changement de base, on a

$$A = \text{Mat}_{\mathcal{B}}(f) = P_{\mathcal{B}', \mathcal{B}}^{-1} \text{Mat}_{\mathcal{B}'}(f) P_{\mathcal{B}', \mathcal{B}} = P_{\mathcal{B}', \mathcal{B}}^{-1} D P_{\mathcal{B}', \mathcal{B}}.$$

Donc A et D sont semblables, et on peut prendre $P_1 = P_{\mathcal{B}', \mathcal{B}}$.

$$\text{Mais nous savons que } P_{\mathcal{B}', \mathcal{B}} = \begin{pmatrix} 3 & -2 \\ 1 & 1 \end{pmatrix}.$$

$$\text{Et donc } P_1 = P_{\mathcal{B}', \mathcal{B}} = P_{\mathcal{B}, \mathcal{B}'}^{-1} = \frac{1}{5} \begin{pmatrix} 1 & 2 \\ -1 & 3 \end{pmatrix}.$$

Rappel

Si $A = Q^{-1}DQ$, alors pour tout polynôme P ,

$$P(A) = Q^{-1}P(D)Q.$$

¹ $P(D) = 0$ car P est annulateur de D .

⚠ Attention !

Nous savons que λ_1 et λ_2 sont des racines de P , mais à ce stade, rien ne garantit qu'elles soient distinctes, elles peuvent être toutes les deux égales à 3 ou toutes les deux égales à -2 .

Rappel

Si on ne souhaite pas calculer un inverse par un pivot, rappelons qu'il existe une formule pour l'inverse d'une matrice $\begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathcal{M}_2(\mathbf{R})$:

$$A^{-1} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

2.c. Puisque $A = P_1^{-1}DP_1$, alors pour tout entier $n \in \mathbf{N}$, $A^n = P_1^{-1}D^nP_1$.

$$\text{Or, } D_1^n = \begin{pmatrix} 3^n & 0 \\ 0 & (-2)^n \end{pmatrix}.$$

Et donc

$$\begin{aligned} A^n &= \begin{pmatrix} 3 & -2 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 3^n & 0 \\ 0 & (-2)^n \end{pmatrix} \frac{1}{5} \begin{pmatrix} 1 & 2 \\ -1 & 3 \end{pmatrix} \\ &= \frac{1}{5} \begin{pmatrix} 3^{n+1} & (-2)^{n+1} \\ 3^n & (-2)^n \end{pmatrix} \begin{pmatrix} 1 & 2 \\ -1 & 3 \end{pmatrix} \\ &= \boxed{\begin{pmatrix} 1 & 3^{n+1} - (-2)^{n+1} & 2 \cdot 3^{n+1} + 3 \cdot (-2)^{n+1} \\ 3^n - (-2)^n & 2 \cdot 3^n + 3 \cdot (-2)^n \end{pmatrix}} \end{aligned}$$

3. On a $u_2 = u_1 + 6u_0 + 3 = 1$ et $u_3 = u_2 + 6u_1 + 3 = 10$.

4. La difficulté réside ici dans le fait qu'à tout moment, le programme doit garder en mémoire deux termes consécutifs de la suite (u_n) , puisque pour calculer u_{n+1} , nous avons besoin de connaître à la fois u_n et u_{n-1} .

L'idée est donc qu'après n passages dans la boucle for, la variable v contienne u_n et u contienne u_{n+1} .

Le problème est qu'en remplaçant la valeur de u par l'expression de u_{n+2} , on «perdrait» la valeur de u_{n+1} , qui ne serait plus stockée dans la mémoire de Sci Lab.

Pour remédier à ceci, on utilise temporairement une variable nommée a , dont le seul but est de stocker temporairement la valeur de u_{n+1} .

Il suffit donc de remplacer la ligne 5 par

$$5 \quad u = u + 6 * v + 3$$

5.a. Pour $n \in \mathbf{N}$, on a

$$AX_n + B = \begin{pmatrix} 1 & 6 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} u_{n+1} \\ u_n \end{pmatrix} + \begin{pmatrix} 3 \\ 0 \end{pmatrix} = \begin{pmatrix} u_{n+1} + 6u_n + 3 \\ u_{n+1} \end{pmatrix} = \begin{pmatrix} u_{n+2} \\ u_{n+1} \end{pmatrix} = \boxed{X_{n+1}}.$$

5.b. On a

$$AL + B = \begin{pmatrix} 1 & 6 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} -\frac{1}{2} \\ -\frac{1}{2} \end{pmatrix} + \begin{pmatrix} 3 \\ 0 \end{pmatrix} = \begin{pmatrix} -\frac{1}{2} \\ -\frac{1}{2} \end{pmatrix} = \boxed{L}.$$

5.c. Montrons par récurrence sur $n \in \mathbf{N}$ que pour tout $n \in \mathbf{N}$, $X_n = A^n(X_0 - L) + L$.

Pour $n = 0$, on a $A^0 = I_2$ et donc $X_0 = A^0(X_0 - L) + L$.

Supposons que $X_n = A^n(X_0 - L) + L$. Alors

$$X_{n+1} = AX_n + B = A(A^n(X_0 - L) + L) + B = A^{n+1}(X_0 - L) + AL + B = A^{n+1}(X_0 - L) + L.$$

Donc la formule est vraie au rang $n + 1$, et par le principe de récurrence, pour tout $n \in \mathbf{N}$,

$$\boxed{X_n = A^n(X_0 - L) + L}.$$

5.d. Nous savons que $X_0 = \begin{pmatrix} u_1 \\ u_0 \end{pmatrix} = \begin{pmatrix} 1 \\ -\frac{1}{2} \end{pmatrix}$.

Et donc

$$\begin{aligned} X_n &= \begin{pmatrix} u_{n+1} \\ u_n \end{pmatrix} = A^n(X_0 - L) + L \\ &= \frac{1}{5} \begin{pmatrix} 3^{n+1} - (-2)^{n+1} & 2 \cdot 3^{n+1} + 3 \cdot (-2)^{n+1} \\ 3^n - (-2)^n & 2 \cdot 3^n + 3 \cdot (-2)^n \end{pmatrix} \left(\begin{pmatrix} 1 \\ -\frac{1}{2} \end{pmatrix} - \begin{pmatrix} -\frac{1}{2} \\ -\frac{1}{2} \end{pmatrix} \right) + \begin{pmatrix} -\frac{1}{2} \\ -\frac{1}{2} \end{pmatrix}. \end{aligned}$$

Et donc, en identifiant les coefficients de la seconde ligne

$$u_n = \frac{1}{5} \left((3 - (-2)^n) \frac{3}{2} \right) - \frac{1}{2} = \boxed{\frac{3}{10} (3 - (-2)^n) - \frac{1}{2}}.$$

terpt

Matrices diagonales

Élever une matrice diagonale à la puissance n , c'est élever chacun de ses coefficients à la puissance n . Bien entendu, ceci n'est plus valable pour une matrice qui ne serait pas diagonale.

Remarque

Notons qu'avant d'entrer dans la boucle for (c'est-à-dire après 0 passages dans la boucle), v contient bien u_0 et u contient bien u_1 .

Astuce

Ici, seule la valeur de u_n nous intéresse, il n'est donc pas nécessaire de calculer tout le produit $A^n(X_0 - L)$, seule sa deuxième ligne nous est utile.

DEVOIR MAISON 6 (À RENDRE LE 15.11.18)

I - Étude de la fonction $x \mapsto \int_0^{+\infty} \frac{\sin t}{t+x} dt$.

On note $F :]0, +\infty[\rightarrow \mathbf{R}$ et $G :]0, +\infty[\rightarrow \mathbf{R}$ les applications définies, pour tout réel $x \in]0, +\infty[$ par :

$$F(x) = \int_1^x \frac{\sin u}{u} du \text{ et } G(x) = \int_1^x \frac{\cos u}{u} du.$$

1. **a.** Montrer, pour tout réel $x \in]0, +\infty[$: $F(x) = -\frac{\cos x}{x} + \cos 1 - \int_1^x \frac{\cos u}{u^2} du$.
En déduire que F admet une limite finie en $+\infty$. On note α cette limite.
- b.** De manière analogue, montrer que G admet une limite finie en $+\infty$, on note β cette limite.
- c.** En déduire que, pour tout réel $x \in]0, +\infty[$, les intégrales $\int_x^{+\infty} \frac{\sin u}{u} du$ et $\int_x^{+\infty} \frac{\cos u}{u} du$ convergent, et que :
$$\int_x^{+\infty} \frac{\sin u}{u} du = \alpha - F(x) \text{ et } \int_x^{+\infty} \frac{\cos u}{u} du = \beta - G(x).$$
2. **a.** Montrer, pour tout réel $x \in]0, +\infty[$ et tout réel $T \in]0, +\infty[$:

$$\int_0^T \frac{\sin t}{t+x} dt = \cos x \int_x^{x+T} \frac{\sin u}{u} du - \sin x \int_x^{x+T} \frac{\cos u}{u} du.$$

- b.** En déduire que, pour tout réel $x \in]0, +\infty[$, l'intégrale $\int_0^{+\infty} \frac{\sin t}{t+x} dt$ converge et que

$$\int_0^{+\infty} \frac{\sin t}{t+x} dt = \cos x \int_x^{+\infty} \frac{\sin u}{u} du - \sin x \int_x^{+\infty} \frac{\cos u}{u} du.$$

On note $A :]0, +\infty[\rightarrow \mathbf{R}$ l'application définie, pour tout réel $x \in]0, +\infty[$, par :

$$A(x) = \int_0^{+\infty} \frac{\sin t}{t+x} dt.$$

3. Montrer que l'application A est de classe \mathcal{C}^2 sur $]0, +\infty[$ et que, pour tout réel $x \in]0, +\infty[$:

$$A''(x) + A(x) = \frac{1}{x}.$$

4. Établir que $A(x)$ et $A'(x)$ tendent vers 0 lorsque x tend vers $+\infty$.

5. **a.** Montrer : $\forall x \in]0, 1], 0 \leq \int_x^1 \frac{\cos u}{u} du \leq -\ln x$.

- b.** En déduire que $\sin x \int_x^{+\infty} \frac{\cos u}{u} du$ tend vers 0 lorsque x tend vers 0 par valeurs strictement positives.

- c.** Montrer que l'intégrale $\int_0^{+\infty} \frac{\sin u}{u} du$ converge, et établir que $A(x)$ tend vers $\int_0^{+\infty} \frac{\sin u}{u} du$ lorsque x tend vers 0 par valeurs strictement positives.

II - Étude de la fonction $x \mapsto \int_0^{+\infty} \frac{e^{-xt}}{1+t^2} dt$.

6. Montrer que, pour tout réel $x \in]0, +\infty[$ et tout entier naturel k , l'application $t \mapsto t^k e^{-xt}$ est bornée sur $[0, +\infty[$ et en déduire que l'intégrale $\int_0^{+\infty} \frac{t^k e^{-xt}}{1+t^2} dt$ converge.

On note, pour tout entier naturel k , $B_k :]0, +\infty[\rightarrow \mathbf{R}$ l'application définie, pour tout réel $x \in]0, +\infty[$ par :

$$B_k(x) = \int_0^{+\infty} \frac{t^k e^{-xt}}{1+t^2} dt.$$

7. **a.** Montrer, en utilisant par exemple l'inégalité de Taylor-Lagrange :

$$\forall u \in \mathbf{R}, |e^u - 1 - u| \leq \frac{u^2}{2} e^{|u|}.$$

- b. En déduire, pour tout réel $x \in]0; +\infty[$, pour tout entier naturel k et pour tout réel h tel que $0 < |h| \leq \frac{x}{2}$:

$$\left| \frac{B_k(x+h) - B_k(x)}{h} + B_{k+1}(x) \right| \leq \frac{|h|}{2} B_{k+2}\left(\frac{x}{2}\right).$$

- c. En déduire que B_0 est de classe \mathcal{C}^2 sur $]0; +\infty[$ et que, pour tout réel $x \in]0; +\infty[$:

$$B_0''(x) + B_0(x) = \frac{1}{x}.$$

8. Montrer, pour tout réel $x \in]0; +\infty[$:

$$0 \leq B_0(x) \leq \frac{1}{x} \text{ et } 0 \leq -B_0'(x) \leq \frac{1}{x^2},$$

et en déduire les limites de $B_0(x)$ et $B_0'(x)$ lorsque x tend vers $+\infty$.

9. a. Montrer : $\forall x \in]0; +\infty[$, $e^{-\sqrt{x}} \int_0^{\frac{1}{\sqrt{x}}} \frac{1}{1+t^2} dt \leq B_0(x) \leq \int_0^{+\infty} \frac{1}{1+t^2} dt$.

- b. En déduire la limite de $B_0(x)$ lorsque x tend vers 0 par valeurs strictement positives.

III - Calcul de l'intégrale $\int_0^{+\infty} \frac{\sin u}{u} du$.

On considère l'application $\varphi :]0; +\infty[\rightarrow \mathbf{R}$ définie, pour tout réel $x \in]0; +\infty[$, par :

$$\varphi(x) = A(x) - B_0(x)$$

où A a été définie dans la partie I et B_0 a été définie dans la partie II.

On note $U :]0; +\infty[\rightarrow \mathbf{R}$ l'application définie, pour tout réel $x \in]0; +\infty[$, par :

$$U(x) = (\varphi(x))^2 + (\varphi'(x))^2.$$

10. Montrer que U est constante sur $]0; +\infty[$.
11. Quelle est la limite de $U(x)$ lorsque x tend vers $+\infty$?
12. En déduire : $\forall x \in]0; +\infty[$, $A(x) = B_0(x)$?
13. Quelle est la valeur de $\int_0^{+\infty} \frac{\sin u}{u} du$?

CORRECTION DU DEVOIR MAISON 6

EML 2004

I - Étude de la fonction $x \mapsto \int_0^{+\infty} \frac{\sin t}{t+x} dt$.

- 1.a. Procédons à une intégration par parties, en posant $f(u) = -\cos u$ et $g(u) = \frac{1}{u}$, qui sont deux fonctions de classe \mathcal{C}^1 sur le segment $[1, x]$ avec $f'(u) = \sin u$ et $g'(u) = -\frac{1}{u^2}$. Alors

$$F(x) = \int_1^x \frac{\sin u}{u} du = \left[-\frac{\cos u}{u} \right]_1^x - \int_1^x \frac{\cos u}{u^2} du = \boxed{-\frac{\cos x}{x} + \cos 1 - \int_1^x \frac{\cos u}{u^2} du.}$$

La fonction \cos étant bornée, puisque $\lim_{x \rightarrow +\infty} \frac{1}{x} = 0$, alors $\lim_{x \rightarrow +\infty} \frac{\cos x}{x} = 0$.

De plus, pour tout $u \geq 1$, on a $0 \leq \left| \frac{\cos u}{u^2} \right| \leq \frac{1}{u^2}$.

Puisque $\int_1^{+\infty} \frac{1}{u^2} du$ converge, par critère de comparaison pour les intégrales de fonctions positives, il en est de même de $\int_1^{+\infty} \left| \frac{\cos u}{u^2} \right| du$ et donc $\int_1^{+\infty} \frac{\cos u}{u^2} du$ converge absolument, et en particulier, converge.

On en déduit que lorsque $x \rightarrow +\infty$, $\lim_{x \rightarrow +\infty} \int_1^x \frac{\cos u}{u^2} du = \int_1^{+\infty} \frac{\cos u}{u^2} du$.

Et donc,

$$\lim_{x \rightarrow +\infty} F(x) = \cos 1 - \int_1^{+\infty} \frac{\cos u}{u^2} du.$$

- 1.b. De même, une intégration par parties sur le segment $[1, x]$ nous donne

$$\int_1^x \frac{\cos u}{u} du = \left[\frac{\sin u}{u} \right]_1^x + \int_1^x \frac{\sin u}{u^2} du = \frac{\sin x}{x} - \sin 1 + \int_1^x \frac{\sin u}{u^2} du.$$

On a alors $\lim_{x \rightarrow +\infty} \frac{\sin x}{x} = 0$ et l'intégrale $\int_1^{+\infty} \frac{\sin u}{u^2} du$ converge absolument car $\left| \frac{\sin u}{u^2} \right| \leq \frac{1}{u^2}$.

Et donc

$$\lim_{x \rightarrow +\infty} G(x) = -\sin(1) + \int_1^{+\infty} \frac{\sin u}{u^2} du.$$

- 1.c. Soit $x > 0$ fixé et soit $A > 0$. Alors, par la relation de Chasles,

$$\int_x^A \frac{\sin u}{u} du = \int_1^A \frac{\sin u}{u} du - \int_1^x \frac{\sin u}{u} du = \int_x^A \frac{\sin u}{u} du - F(x) \xrightarrow{A \rightarrow +\infty} \int_1^{+\infty} \frac{\sin u}{u} du - F(x) = \alpha - F(x).$$

En particulier, puisque cette limite existe, on en déduit que $\int_x^{+\infty} \frac{\sin u}{u} du$ converge et vaut $\alpha - F(x)$.

De même, on prouve que $\int_1^{+\infty} \frac{\cos u}{u} du$ converge et vaut $\beta - G(x)$.

- 2.a. Procédons au changement de variable $u = x + t$, qui est légitime car affine :

$$\begin{aligned} \int_0^T \frac{\sin t}{x+t} dt &= \int_x^{x+T} \frac{\sin(u-x)}{u} du \\ &= \int_x^{x+T} \frac{\cos(x) \sin(u) - \sin(x) \cos(u)}{u} du \\ &= \boxed{\cos x \int_x^{x+T} \frac{\sin u}{u} du - \sin x \int_x^{x+T} \frac{\cos u}{u} du.} \end{aligned}$$

Rappel

Le produit d'une fonction bornée par une fonction qui tend vers 0 tend vers 0.

Méthode

Pour prouver la convergence de l'intégrale d'une fonction qui n'est pas de signe constant, on étudie en priorité la convergence absolue. On peut éventuellement utiliser un o (ici on aurait pu dire que $\frac{\cos u}{u^2} = o_{u \rightarrow +\infty} \left(\frac{1}{x^{3/2}} \right)$), mais surtout pas de majoration ou d'équivalents, car ces critères ne valent que pour des fonctions de signe constant.

Rappel

Pour tous $a, b \in \mathbf{R}$,
 $\sin(a-b) =$
 $\sin a \cos b - \cos a \sin b.$

2.b. Lorsque T tend vers $+\infty$, on a alors¹

$$\lim_{T \rightarrow +\infty} \int_0^T \frac{\sin t}{x+t} dt = \cos x \int_x^{+\infty} \frac{\sin u}{u} du - \sin x \int_x^{+\infty} \frac{\cos u}{u} du.$$

Et donc ceci prouve que $\int_0^{+\infty} \frac{\sin t}{x+t} dt$ converge et que

$$\int_0^{+\infty} \frac{\sin t}{t+x} dt = \cos x \int_x^{+\infty} \frac{\sin u}{u} du - \sin x \int_x^{+\infty} \frac{\cos u}{u} du.$$

3. D'après la question 1.c, pour tout $x > 0$, on a $\int_x^{+\infty} \frac{\sin u}{u} du = \alpha - F(x)$.

Mais par le théorème fondamental de l'analyse, F est \mathcal{C}^1 sur \mathbf{R}_+^* , de dérivée égale à $\frac{\sin x}{x}$.

Et donc la fonction $x \mapsto \int_x^{+\infty} \frac{\sin u}{u} du$ est \mathcal{C}^1 sur \mathbf{R}_+^* , de dérivée égale à $-\frac{\sin x}{x}$.

De même, $x \mapsto \int_x^{+\infty} \frac{\cos u}{u} du$ est \mathcal{C}^1 sur \mathbf{R}_+^* , de dérivée égale à $x \mapsto -\frac{\cos x}{x}$.

Et donc A est \mathcal{C}^1 sur \mathbf{R}_+^* car somme et produit de fonctions \mathcal{C}^1 . On a alors

$$\begin{aligned} A'(x) &= -\sin x \int_x^{+\infty} \frac{\sin u}{u} du - \frac{\cos x \sin x}{x} - \cos x \int_x^{+\infty} \frac{\cos u}{u} du + \frac{\sin x \cos x}{x} \\ &= -\sin x \int_x^{+\infty} \frac{\sin u}{u} du - \cos x \int_x^{+\infty} \frac{\cos u}{u} du. \end{aligned}$$

A' est alors encore \mathcal{C}^1 car somme de produits de fonctions \mathcal{C}^1 , donc A est \mathcal{C}^2 , et

$$\begin{aligned} A''(x) &= -\cos x \int_x^{+\infty} \frac{\sin u}{u} du + \frac{\sin^2 x}{x} + \sin x \int_x^{+\infty} \frac{\cos u}{u} du + \frac{\cos^2 x}{x} \\ &= -A(x) + \frac{\sin^2 x + \cos^2 x}{x} \\ &= -A(x) + \frac{1}{x}. \end{aligned}$$

Et alors on a bien $A''(x) + A(x) = \frac{1}{x}$.

4. Lorsque $x \rightarrow +\infty$, on a $\int_x^{+\infty} \frac{\sin u}{u} du \rightarrow 0$ car il s'agit du reste d'une intégrale convergente.

La fonction \cos étant bornée sur \mathbf{R} , on a donc $\lim_{x \rightarrow +\infty} \cos x \int_x^{+\infty} \frac{\sin u}{u} du = 0$.

De même, on a $\lim_{x \rightarrow +\infty} \sin x \int_x^{+\infty} \frac{\cos u}{u} du = 0$.

Et donc, par somme de limites, $\lim_{x \rightarrow +\infty} A(x) = 0$.

On montre de même que $\lim_{x \rightarrow +\infty} A'(x) = 0$.

5.a. Nous savons que la fonction cosinus est positive sur $\left[0, \frac{\pi}{2}\right]$ et que $\frac{\pi}{2} > 1$, et donc pour $u \in]0, 1]$, on a $0 \leq \cos u \leq 1$, et donc

$$0 \leq \frac{\cos u}{u} \leq \frac{1}{u}.$$

Par croissance de l'intégrale, il vient donc

$$0 \leq \int_x^1 \frac{\cos u}{u} du \leq \int_x^1 \frac{1}{u} du = [\ln u]_x^1 = -\ln(x).$$

5.b. On a $\sin x \int_x^{+\infty} \frac{\cos u}{u} du = \sin x \int_x^1 \frac{\cos u}{u} du + \sin x \int_1^{+\infty} \frac{\cos u}{u} du$.

Puisque $\int_1^{+\infty} \frac{\cos u}{u} du$ est une constante et que $\lim_{x \rightarrow 0^+} \sin x = 0$, alors $\lim_{x \rightarrow 0^+} \sin x \int_1^{+\infty} \frac{\cos u}{u} du = 0$.

¹ Car nous avons prouvé en 1.c que les intégrales en question convergent.

\mathcal{C}^k

Rappelons qu'une fonction \mathcal{C}^k est par définition une fonction dérivable k fois, de dérivée $k^{\text{ème}}$ continue. Et une fonction dérivable est \mathcal{C}^k si et seulement si sa dérivée est \mathcal{C}^{k-1} .

De plus, la fonction sinus étant positive sur $[0, 1]$, on a, pour tout $x \in]0, 1]$,

$$0 \leq \sin x \int_x^1 \frac{\cos u}{u} du \leq -\sin x \ln x.$$

Or, au voisinage de 0, $\sin x \sim x$ et donc $\sin x \ln x \sim x \ln x \xrightarrow{x \rightarrow 0^+} 0$.

Et par le théorème des gendarmes, il vient alors $\lim_{x \rightarrow 0^+} \sin x \int_x^{+\infty} \frac{\cos u}{u} du = 0$.

Par somme de limites, on a alors

$$\lim_{x \rightarrow 0^+} \sin x \int_x^{+\infty} \frac{\cos u}{u} du = 0.$$

5.c. Nous savons déjà que $\int_1^{+\infty} \frac{\sin u}{u} du$ converge, et donc il suffit d'établir la convergence de $\int_0^1 \frac{\sin u}{u} du$.

La fonction $u \mapsto \frac{\sin u}{u}$ étant continue sur $]0, 1]$, le seul problème de convergence est au voisinage de 0.

Mais lorsque $u \rightarrow 0$, on a $\sin u \sim u$ et donc $\frac{\sin u}{u} \sim \frac{u}{u} = 1 \xrightarrow{u \rightarrow 0^+} 1$.

Ainsi, $\int_0^1 \frac{\sin u}{u} du$ est une intégrale faussement impropre², donc convergente.

Et donc $\int_0^{+\infty} \frac{\sin u}{u} du$ converge.

On a donc

$$A(x) = \underbrace{\cos x}_{\xrightarrow{x \rightarrow 0^+} 1} \underbrace{\int_x^{+\infty} \frac{\sin u}{u} du}_{\xrightarrow{x \rightarrow 0^+} \int_0^{+\infty} \frac{\sin u}{u} du} - \underbrace{\sin(x)}_{\xrightarrow{x \rightarrow 0^+} 0} \underbrace{\int_x^{+\infty} \frac{\cos u}{u} du}_{\xrightarrow{x \rightarrow 0^+} 0} \xrightarrow{x \rightarrow 0^+} \int_0^{+\infty} \frac{\sin u}{u} du.$$

² Autrement dit, $x \mapsto \frac{\sin x}{x}$ est prolongeable par continuité en 0.

II - Étude de la fonction $x \mapsto \int_0^{+\infty} \frac{e^{-xt}}{1+t^2} dt$.

6. Pour $x > 0$ fixé, la fonction $f_k : t \mapsto t^k e^{-xt}$ est dérivable sur \mathbf{R}_+ , de dérivée égale à

$$f'_k(t) = kt^{k-1} e^{-xt} - xt^k e^{-xt} = t^{k-1} e^{-xt} (k - xt).$$

De plus, on a $\lim_{t \rightarrow +\infty} t^k e^{-xt} = 0$ par croissance comparée.

Son tableau de variations est donc le suivant

x	0	$\frac{k}{x}$	$+\infty$
$f'_k(x)$		+	0
$f_k(x)$	$f_k(0)$	$f_k\left(\frac{k}{x}\right)$	0

Et donc pour tout $t \geq 0$, on a $0 \leq f_k(0) \leq t^k e^{-xt} \leq f_k\left(\frac{k}{x}\right) \leq \left(\frac{k}{x}\right)^k e^{-k}$, et donc f_k est bornée.

En particulier, il vient, pour tout $t \geq 0$,

$$0 \leq \frac{t^k e^{-xt}}{1+t^2} \leq \left(\frac{k}{x}\right)^k e^{-k} \frac{1}{1+t^2}.$$

Fonction bornée
Rappelons qu'une fonction est bornée si elle est à la fois majorée et minorée, il ne suffit pas de la majorer.

Or, $\int_0^{+\infty} \frac{dt}{1+t^2}$ converge car le seul éventuel problème de convergence est au voisinage de $+\infty$ et que $\frac{1}{1+t^2} \underset{t \rightarrow +\infty}{\sim} \frac{1}{t^2}$.

Et donc³ $\int_0^{+\infty} \frac{t^k e^{-xt}}{1+t^2} dt$ converge également.

- 7.a. La fonction $f(t) = e^t$ est de classe \mathcal{C}^2 sur \mathbf{R} , et $f'(t) = f''(t) = e^t$.
Si $u \geq 0$, alors, par croissance de l'exponentielle, on a, pour tout $t \in [0, u]$,
 $0 \leq f''(t) \leq e^u = e^{|u|}$. Et donc $|f''(t)| \leq e^{|u|}$.
D'après l'inégalité de Taylor-Lagrange à l'ordre 1 en 0, on a donc

$$|f(u) - f(0) - f'(0)u| \leq \frac{|u|^2}{2!} e^{|u|} \Leftrightarrow |e^u - 1 - u| \leq \frac{u^2}{2} e^{|u|}.$$

Si $u < 0$, alors pour $t \in [u, 0]$, on a $0 \leq f''(t) \leq 1 \leq e^{|u|}$ et donc, toujours par l'inégalité de Taylor-Lagrange,

$$|f(u) - f(0) - f'(0)u| \leq \frac{|u|^2}{2!} e^{|u|} \Leftrightarrow |e^u - 1 - u| \leq \frac{u^2}{2} e^{|u|}.$$

- 7.b. D'après le résultat de la question précédente, on a, pour tout $t \geq 0$,

$$|e^{-th} - 1 + th| \leq \frac{t^2 h^2}{2} e^{th}.$$

En multipliant par $\frac{t^k e^{-xt}}{1+t^2}$, on a, pour tout $t > 0$,

$$\left| \frac{t^k e^{-t(x+h)}}{1+t^2} - \frac{t^k e^{-xt}}{1+t^2} + \frac{t^{k+1} h e^{-xt}}{1+t^2} \right| \leq \frac{t^{k+2} h^2}{2(1+t^2)} e^{-t(x+h)}.$$

Et donc, après division par $|h| > 0$, il vient

$$\left| \frac{1}{h} \left(\frac{t^k e^{-t(x+h)}}{1+t^2} - \frac{t^k e^{-xt}}{1+t^2} \right) + \frac{t^{k+1} e^{-xt}}{1+t^2} \right| \leq \frac{t^{k+2} |h|}{2(1+t^2)} e^{-t(x+h)}.$$

Notons que $|h| > \frac{x}{2}$ et donc $x+h > \frac{x}{2}$, de sorte que, pour $t > 0$, $e^{-t(x+h)} \geq e^{-t \frac{x}{2}}$.

Enfin, on a

$$\begin{aligned} \left| \frac{B_k(x+h) - B_k(x)}{h} + B_{k+1}(x) \right| &= \left| \frac{1}{h} \left(\int_0^{+\infty} \frac{t^k e^{-t(x+h)}}{1+t^2} dt - \int_0^{+\infty} \frac{t^k e^{-xt}}{1+t^2} dt \right) + \int_0^{+\infty} \frac{t^{k+1} e^{-xt}}{1+t^2} dt \right| \\ &= \left| \int_0^{+\infty} \left(\frac{1}{h} \left(\frac{t^k e^{-t(x+h)}}{1+t^2} - \frac{t^k e^{-xt}}{1+t^2} \right) + \frac{t^{k+1} e^{-xt}}{1+t^2} \right) dt \right| \\ &\leq \int_0^{+\infty} \left| \frac{1}{h} \left(\frac{t^k e^{-t(x+h)}}{1+t^2} - \frac{t^k e^{-xt}}{1+t^2} \right) + \frac{t^{k+1} e^{-xt}}{1+t^2} \right| dt \\ &\leq \frac{|h|}{2} \int_0^{+\infty} \frac{t^{k+2} e^{-tx/2}}{1+t^2} dt \\ &\leq \frac{|h|}{2} B_{k+2} \left(\frac{x}{2} \right). \end{aligned}$$

- 7.c. Soit $x > 0$ fixé. Alors, on a $\lim_{h \rightarrow 0} \frac{|h|}{2} B_{k+2} \left(\frac{x}{2} \right) = 0$, et donc, par le théorème des gendarmes,

$$\lim_{h \rightarrow 0} \left| \frac{B_k(x+h) - B_k(x)}{h} + B_{k+1}(x) \right| = 0 \Leftrightarrow \lim_{h \rightarrow 0} \frac{B_k(x+h) - B_k(x)}{h} = -B_{k+1}(x).$$

Ceci prouve donc que B_k est dérivable en x et que $B'_k(x) = -B_{k+1}(x)$.
En particulier, B_k est dérivable sur \mathbf{R}_+^* .

Alternative MAISON 6

On pourrait aussi calculer explicitement cette intégrale en notant qu'une primitive de $\frac{1}{1+t^2}$ est $\text{Arctan}(t)$.

³ $\left(\frac{k}{x}\right)^k e^{-k}$ est une constante qui ne dépend pas de t .

Méthode

Pour appliquer l'inégalité de Taylor-Lagrange entre a et x , il faut un majorant de $|f^{(n+1)}|$ sur le segment d'extrémités a et x .
Lorsque, comme c'est le cas ici, $f^{(n)}$ n'est pas bornée sur \mathbf{R} , il faut donc faire preuve d'un peu de finesse pour majorer $|f^{(n+1)}|$.
Cela nous conduit en particulier à distinguer deux cas, car si $u \geq 0$, le segment d'extrémités 0 et u est $[0, u]$, alors que si $u \leq 0$, il s'agit de $[u, 0]$.

Précaution

On ne connaît pas le signe de h , donc $\frac{h^2}{|h|} = |h|$.

Valeur absolue

Ne pas oublier que l'intégrale de la valeur absolue n'est pas égale à la valeur absolue de l'intégrale, on dispose juste d'une inégalité (l'inégalité triangulaire).

Et donc B_0 est dérivable, et sa dérivée est $-B_1$, qui est elle-même dérivable, de dérivée B_2 .

Puisque $B_2 = (B_0)''$ est dérivable, elle est continue et donc B_0 est de classe \mathcal{C}^2 .

Et donc, pour tout $x > 0$, on a

$$B_0(x) + B_0''(x) = B_0(x) + B_2(x) = \int_0^{+\infty} \frac{e^{-xt}}{1+t^2} dt + \int_0^{+\infty} \frac{t^2 e^{-xt}}{1+t^2} dt = \int_0^{+\infty} \frac{e^{-xt}(1+t^2)}{1+t^2} dt = \int_0^{+\infty} e^{-xt} dt = \boxed{\frac{1}{x}}.$$

8. Il est clair que pour tout $k \in \mathbf{N}$, $B_k(x) \geq 0$ car intégrale d'une fonction positive. En particulier, on a $B_0(x) \geq 0$ et $B_2(x) \geq 0$, de sorte que

$$B_0(x) = \frac{1}{x} - B_0''(x) = \frac{1}{x} - B_2(x) \leq \frac{1}{x}.$$

De même, en dérivant la relation de la question 7.c, ce qui est légitime car $B_0'' = B_2$ est dérivable, de dérivée $-B_3$, il vient $-B_3(x) + B_0'(x) = -\frac{1}{x^2}$.

Et donc $-B_0'(x) = \frac{1}{x^2} - B_3(x) \leq \frac{1}{x^2}$.

De plus, $-B_0'(x) = B_1(x) \geq 0$.

Par le théorème des gendarmes, on en déduit que $\lim_{x \rightarrow 0} B_0(x) = \lim_{x \rightarrow 0} B_0'(x) = 0$.

- 9.a. Par la relation de Chasles, on a

$$B_0(x) = \int_0^{+\infty} \frac{e^{-xt}}{1+t^2} dt = \int_0^{\frac{1}{\sqrt{x}}} \frac{e^{-xt}}{1+t^2} dt + \int_{\frac{1}{\sqrt{x}}}^{+\infty} \frac{e^{-xt}}{1+t^2} dt.$$

Par croissance de l'intégrale, la seconde intégrale est positive car intégrale d'une fonction positive.

Et donc $B_0(x) \geq \int_0^{\frac{1}{\sqrt{x}}} \frac{e^{-xt}}{1+t^2} dt$.

De plus, pour $t \in \left[0, \frac{1}{\sqrt{x}}\right]$, $-xt \geq -\sqrt{x}$.

Par croissance de l'exponentielle, on a donc $e^{-xt} \geq e^{-\sqrt{x}}$, puis $\frac{e^{-xt}}{1+t^2} \geq \frac{e^{-\sqrt{x}}}{1+t^2}$.

Par croissance de l'intégrale, on a donc

$$B_0(x) \geq \int_0^{\frac{1}{\sqrt{x}}} \frac{e^{-xt}}{1+t^2} dt \geq e^{-\sqrt{x}} \int_0^{\frac{1}{\sqrt{x}}} \frac{1}{1+t^2} dt.$$

Enfin, pour tout $t \geq 0$, on a $e^{-xt} \leq 1$ et donc $\frac{e^{-xt}}{1+t^2} \leq \frac{1}{1+t^2}$, de sorte que, par croissance de l'intégrale,

$$B_0(x) = \int_0^{+\infty} \frac{e^{-xt}}{1+t^2} dt \leq \int_0^{+\infty} \frac{1}{1+t^2} dt.$$

Et donc on a bien

$$e^{-\sqrt{x}} \int_0^{\frac{1}{\sqrt{x}}} \frac{1}{1+t^2} dt \leq B_0(x) \leq \int_0^{+\infty} \frac{1}{1+t^2} dt.$$

- 9.b. D'une part, on a

$$\int_0^{+\infty} \frac{1}{1+t^2} dt = \lim_{A \rightarrow +\infty} \int_0^A \frac{1}{1+t^2} dt = \lim_{A \rightarrow +\infty} [\text{Arctan}(t)]_0^A = \lim_{A \rightarrow +\infty} \text{Arctan}(A) = \frac{\pi}{2}.$$

D'autre part, on a $\lim_{x \rightarrow 0^+} \frac{1}{\sqrt{x}} = +\infty$ et donc⁴,

$$\lim_{A \rightarrow +\infty} \int_0^{\frac{1}{\sqrt{x}}} \frac{1}{1+t^2} dt = \int_0^{+\infty} \frac{1}{1+t^2} dt = \frac{\pi}{2}.$$

⁴ Par composition de limites.

Enfin, $\lim_{x \rightarrow 0^+} e^{-\sqrt{x}} = 1$ et donc

$$\lim_{x \rightarrow 0^+} e^{-\sqrt{x}} \int_0^{\frac{1}{\sqrt{x}}} \frac{1}{1+t^2} dt = \frac{\pi}{2}.$$

D'après le théorème des gendarmes, on a donc $\lim_{x \rightarrow 0^+} B_0(x) = \frac{\pi}{2}$.

Partie III - Calcul de l'intégrale $\int_0^{+\infty} \frac{\sin u}{u} du$.

10. Puisqu'il a déjà été prouvé que A et B_0 sont toutes les deux de classe \mathcal{C}^2 , il en est de même de φ . Et donc φ' est dérivable, de sorte que U l'est également. On a alors, pour tout $x > 0$,

$$U'(x) = 2\varphi(x)\varphi'(x) + 2\varphi'(x)\varphi''(x) = 2\varphi'(x)(\varphi(x) + \varphi''(x)).$$

Or, on a $\varphi''(x) = A''(x) - B_0''(x)$.

D'après les relations obtenues aux questions 3 et 7.c, on a

$$A''(x) - B_0''(x) = \frac{1}{x} - A(x) - \frac{1}{x} + B_0(x) = -\varphi(x)$$

de sorte que $U'(x) = 0$.

Et donc U est constante sur l'intervalle $]0, +\infty[$.

11. Lorsque $x \rightarrow +\infty$, il a déjà été prouvé que $\lim_{x \rightarrow +\infty} A(x) = 0$ et $\lim_{x \rightarrow +\infty} B_0(x) = 0$, et donc $\lim_{x \rightarrow +\infty} \varphi(x) = 0$.

De même, il a été prouvé à la question 4 que $\lim_{x \rightarrow +\infty} A'(x) = 0$ et à la question 8 que

$$\lim_{x \rightarrow +\infty} B_0'(x) = 0, \text{ et donc } \lim_{x \rightarrow +\infty} \varphi'(x) = 0.$$

On en déduit donc que $\lim_{x \rightarrow +\infty} U(x) = 0$.

12. D'après les questions 10 et 11, U est la fonction constante égale à 0. Mais une somme de nombres positifs est nulle si et seulement si chacun de ces nombres est nul, donc pour tout $x \in]0, +\infty[$, $U(x) = (\varphi(x))^2 + (\varphi'(x))^2 = 0 \Leftrightarrow \varphi(x) = \varphi'(x) = 0$. En particulier, on a

$$\forall x \in]0, +\infty[, \varphi(x) = 0 \Leftrightarrow A(x) = B_0(x).$$

13. Nous savons d'une part que $\lim_{x \rightarrow 0^+} B_0(x) = \frac{\pi}{2}$ et d'autre part que $\lim_{x \rightarrow 0^+} A(x) = \int_0^{+\infty} \frac{\sin u}{u} du$. Mais les fonctions A et B_0 étant égales, elles ont mêmes limites et donc

$$\int_0^{+\infty} \frac{\sin u}{u} du = \frac{\pi}{2}.$$

Détail

Pour une fonction constante égale à λ , sa limite en $+\infty$ est nécessairement égale à λ .

DEVOIR MAISON 7 (À RENDRE LE 22.11.18)

Les parties I et II sont obligatoires, la partie III est facultative

Définitions et notations

- p désigne un entier naturel supérieur ou égal à 3.
- On note $\mathcal{M}_p(\mathbf{C})$ l'ensemble des matrices carrées d'ordre p à coefficients complexes, $\mathcal{M}_{1,p}(\mathbf{R})$ l'ensemble des matrices-lignes à p colonnes à coefficients réels, $\mathcal{M}_p(\mathbf{R})$ l'ensemble des matrices carrées d'ordre p à coefficients réels, I_p la matrice diagonale de $\mathcal{M}_p(\mathbf{C})$ et de $\mathcal{M}_p(\mathbf{R})$ dont tous les coefficients diagonaux sont égaux à 1.
- On note, pour toute matrice carrée A d'ordre p et tout $(i, j) \in \llbracket 1, p \rrbracket^2$, $(A)_{i,j}$ le coefficient de A situé à la ligne i et à la colonne j .
- On note, pour toute matrice-ligne L de $\mathcal{M}_{1,p}(\mathbf{R})$ et tout $j \in \llbracket 1, p \rrbracket$, $(L)_j$ le coefficient de L situé à la colonne j .
- On dit qu'une suite de matrices $(A_n)_{n \geq 1}$ de $\mathcal{M}_p(\mathbf{R})$ converge vers une matrice A de $\mathcal{M}_p(\mathbf{R})$, et on note $A_n \xrightarrow[n \rightarrow +\infty]{} A$, si et seulement si : $\forall (i, j) \in \llbracket 1, p \rrbracket^2, (A_n)_{i,j} \xrightarrow[n \rightarrow +\infty]{} (A)_{i,j}$.
- On dit qu'une suite de matrices $(L_n)_{n \geq 1}$ de $\mathcal{M}_{1,p}(\mathbf{R})$ converge vers une matrice L de $\mathcal{M}_{1,p}(\mathbf{R})$, et on note $L_n \xrightarrow[n \rightarrow +\infty]{} L$, si et seulement si : $\forall j \in \llbracket 1, p \rrbracket, (L_n)_j \xrightarrow[n \rightarrow +\infty]{} (L)_j$.
- On admet que, si la suite $(A_n)_{n \geq 1}$ de matrices de $\mathcal{M}_p(\mathbf{R})$ converge vers la matrice A et si la suite de matrice $(B_n)_{n \geq 1}$ de $\mathcal{M}_p(\mathbf{R})$ converge vers la matrice B , alors la suite $(A_n B_n)_{n \geq 1}$ converge vers la matrice AB .
- On admet que si la suite $(A_n)_{n \geq 1}$ de matrices de $\mathcal{M}_p(\mathbf{R})$ converge vers la matrice A et si L est une matrice-ligne de $\mathcal{M}_{1,p}(\mathbf{R})$, alors la suite $(L A_n)_{n \geq 1}$ de matrices converge vers LA .

- On appelle matrice stochastique toute matrice $A = (a_{i,j})$ de $\mathcal{M}_p(\mathbf{R})$ telle que :
$$\begin{cases} \forall (i, j) \in \llbracket 1, p \rrbracket^2, a_{i,j} \geq 0 \\ \forall i \in \llbracket 1, p \rrbracket, \sum_{j=1}^p a_{i,j} = 1, \end{cases}$$

et on note \mathcal{ST}_p l'ensemble des matrices stochastiques de $\mathcal{M}_p(\mathbf{R})$.

Partie I : Résultats généraux sur les matrices stochastiques - Illustrations

1. a. On note V la matrice-colonne à p lignes dont tous les coefficients sont égaux à 1.

Montrer que, pour tout $A \in \mathcal{M}_p(\mathbf{R})$: $A \in \mathcal{ST}_p \iff \begin{cases} \forall (i, j) \in \llbracket 1, p \rrbracket^2, (A)_{i,j} \geq 0 \\ AV = V. \end{cases}$

- b. En déduire que toutes les matrices de \mathcal{ST}_p ont une valeur propre commune.

2. Démontrer : $\forall A, B \in \mathcal{ST}_p, AB \in \mathcal{ST}_p$.

3. On note : $A_1 = \begin{pmatrix} 1 & 0 & 0 \\ 1/2 & 1/2 & 0 \\ 1/3 & 1/3 & 1/3 \end{pmatrix}, A_2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1/2 & 1/2 \\ 0 & 1/2 & 1/2 \end{pmatrix}, A_3 = \begin{pmatrix} 1 & 0 & 0 \\ 1/2 & 1/2 & 0 \\ 0 & 1/2 & 1/2 \end{pmatrix}$.

- a. Avec le minimum de calculs possibles, déterminer les valeurs propres de A_1 , et la dimension de ses sous-espaces propres.

En déduire qu'il existe une base de $\mathcal{M}_{3,1}(\mathbf{R})$ formée de vecteurs propres de A_1 , et déterminer la matrice dans cette base de l'application $f_{A_1} \in \mathcal{L}(\mathcal{M}_{3,1}(\mathbf{R}))$ définie par $f_{A_1}(X) = A_1 X$.

Enfin, en déduire que A_1 est semblable à une matrice diagonale.

- b. En utilisant éventuellement les matrices A_2 et A_3 :

- i. Montrer qu'il existe dans \mathcal{ST}_3 au moins un élément non semblable à une matrice diagonale de $\mathcal{M}_3(\mathbf{C})$.

- ii. Justifier si l'affirmation suivante est vraie ou fausse : « Pour tout élément A de \mathcal{ST}_3 , le sous-espace propre pour A associé à la valeur propre 1 est de dimension 1 ».

4. Soit $A \in \mathcal{ST}_p$ et λ une valeur propre de A dans \mathbf{C} .

On note $X = \begin{pmatrix} x_1 \\ \vdots \\ x_p \end{pmatrix}$ un vecteur propre de A associé à la valeur propre λ .

On note i un élément de $\llbracket 1, p \rrbracket$ tel que : $\forall k \in \llbracket 1, p \rrbracket, |x_k| \leq |x_i|$.

- a. Montrer : $|\lambda x_i| \leq |x_i|$.

- b. En déduire : $|\lambda| \leq 1$.

Partie II : Suites de moyennes de puissances de matrices stochastiques

Soit $A \in \mathcal{ST}_p$. On note $A^0 = I_p$.

5. a. Établir : $\forall n \in \mathbf{N}, A^n \in \mathcal{ST}_p$.

b. Montrer : $\forall n \in \mathbf{N}^*, \frac{1}{n} \sum_{k=0}^{n-1} A^k \in \mathcal{ST}_p$.

Dans la suite de cette partie II, on suppose qu'il existe $r \in \llbracket 1, p-1 \rrbracket$, $P \in \mathcal{M}_p(\mathbf{R})$ inversible, $D \in \mathcal{M}_p(\mathbf{R})$ diagonale dont tous les coefficients diagonaux $(D)_{i,i}$ sont égaux à 1 si $i \leq r$ et distincts de 1 si $i \geq r+1$, tels que : $A = PDP^{-1}$.

On note, pour tout $n \in \mathbf{N}^*$: $M_n = \frac{1}{n} \sum_{k=0}^{n-1} D^k$ et $B_n = PM_nP^{-1}$.

On note Δ la matrice de $\mathcal{M}_p(\mathbf{R})$ diagonale dont les coefficients diagonaux $(\Delta)_{i,i}$ sont égaux à 1 si $i \leq r$ et nuls sinon, et on note $B = P\Delta P^{-1}$.

6. Démontrer, pour tout $x \in \mathbf{R}$ fixé tel que $|x| \leq 1$: $\frac{1}{n} \sum_{k=0}^{n-1} x^k \xrightarrow{n \rightarrow +\infty} \begin{cases} 1 & \text{si } x = 1 \\ 0 & \text{si } x \neq 1. \end{cases}$

7. Montrer : $M_n \xrightarrow{n \rightarrow +\infty} \Delta$, et en déduire : $B_n \xrightarrow{n \rightarrow +\infty} B$.

8. a. Montrer : $\forall n \in \mathbf{N}^*, B_n \in \mathcal{ST}_p$.

b. En déduire : $B \in \mathcal{ST}_p$.

Partie III : Aspect probabiliste

On dispose d'un objet noté T et de trois urnes numérotées 1, 2 et 3.

À chaque instant n ($n \in \mathbf{N}$), T est dans une des trois urnes et une seule.

On note, pour tout $n \in \mathbf{N}$, X_n la variable aléatoire égale au numéro de l'urne dans laquelle se trouve l'objet à l'instant n et L_n la matrice suivante de $\mathcal{M}_{1,3}(\mathbf{R})$: $L_n = (P(X_n = 1) \quad P(X_n = 2) \quad P(X_n = 3))$.

On suppose connues la loi de X_0 et la matrice A de $\mathcal{M}_3(\mathbf{R})$ définie par :

$$\forall (i, j) \in \{1, 2, 3\}^2, (A)_{i,j} = P_{(X_0=i)}(X_1 = j).$$

On suppose : $\forall (i, j) \in \{1, 2, 3\}^2, P_{(X_n=i)}(X_{n+1} = j) = P_{(X_0=i)}(X_1 = j)$.

9. Montrer : $A \in \mathcal{ST}_3$.

10. Montrer : $\forall n \in \mathbf{N}, L_{n+1} = L_n A$ puis : $\forall n \in \mathbf{N}, L_n = L_0 A^n$.

On suppose dorénavant $A = A_1$, définie dans la question 3, et on note $D_1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1/2 & 0 \\ 0 & 0 & 1/3 \end{pmatrix}$.

11. Déterminer une matrice $P_1 \in \mathcal{M}_3(\mathbf{R})$, inversible et à coefficients diagonaux tous égaux à 1, telle que $A_1 = P_1 D_1 P_1^{-1}$ et calculer P_1^{-1} .

12. Déterminer la limite de la suite $(D_1^n)_{n \geq 1}$, puis la limite de la suite $(A_1^n)_{n \geq 1}$.

13. Déterminer la limite de la suite $(L_n)_{n \geq 1}$. Expliquer ce résultat par des arguments probabilistes.

CORRECTION DU DEVOIR MAISON 7

EML 2010

Partie I : Résultats généraux sur les matrices stochastiques. Illustrations

1.a. Il s'agit de montrer que $AV = V$ si et seulement si

$$\forall i \in \llbracket 1, p \rrbracket, \sum_{j=1}^p a_{i,j} = 1.$$

Mais pour $A = (a_{i,j})_{1 \leq i, j \leq p} \in \mathcal{M}_p(\mathbf{C})$, on a

$$AV = \begin{pmatrix} \sum_{j=1}^p a_{1,j} \\ \sum_{j=1}^p a_{2,j} \\ \vdots \\ \sum_{j=1}^p a_{p,j} \end{pmatrix}$$

Donc par identification des coefficients, il est clair que

$$AV = V \Leftrightarrow \forall i \in \llbracket 1, p \rrbracket, \sum_{j=1}^p a_{i,j} = 1.$$

1.b. Puisque V est non nul, $AV = V$ implique que V est un vecteur propre de A , associé à la valeur propre 1. On en déduit que toutes les matrices de \mathcal{ST}_p ont 1 comme valeur propre.

2. Soient $A = (a_{i,j})_{1 \leq i, j \leq p}$ et $B = (b_{i,j})_{1 \leq i, j \leq p}$ deux matrices de \mathcal{ST}_p . Alors il est clair que

$$(AB)V = A(BV) = AV = V.$$

De plus, notons $AB = (c_{i,j})_{1 \leq i, j \leq p}$. Par définition du produit matriciel, on a

$$\forall (i, j) \in \llbracket 1, p \rrbracket^2, c_{i,j} = \sum_{k=1}^p a_{i,k} b_{k,j} \geq 0$$

car tous les $a_{i,k}$ et tous les $b_{k,j}$ sont positifs.

Ainsi, d'après la question 1.a, on a bien $AB \in \mathcal{ST}_p$.

3.a. A_1 est triangulaire inférieure, donc ses valeurs propres sont ses coefficients diagonaux. En

l'occurrence, $\text{Spec}(A_1) = \left\{ 1, \frac{1}{2}, \frac{1}{3} \right\}$.

Et alors, puisque A_1 possède trois valeurs propres distinctes, ses sous-espaces propres sont de dimension 1.

Soit alors X_1 une base de $E_1(A_1)$, X_2 une base de $E_{1/2}(A_1)$ et X_3 une base de $E_{1/3}(A_1)$.

Alors (X_1, X_2, X_3) est une famille de vecteurs propres associés à des valeurs propres distinctes : elle est libre. Et étant de cardinal 3, c'est une base de $\mathcal{M}_{n,1}(\mathbf{R})$.

On a alors

$$f_{A_1}(X_1) = A_1 X_1 = X_1, \quad f_{A_1}(X_2) = A_1 X_2 = \frac{1}{2} X_2, \quad f_{A_1}(X_3) = A_1 X_3 = \frac{1}{3} X_3.$$

Et donc la matrice de f_{A_1} dans cette base est

$$\text{Mat}_{(X_1, X_2, X_3)}(f_{A_1}) = \begin{pmatrix} f_{A_1}(X_1) & f_{A_1}(X_2) & f_{A_1}(X_3) \\ 1 & 0 & 0 \\ 0 & \frac{1}{2} & 0 \\ 0 & 0 & \frac{1}{3} \end{pmatrix} \begin{matrix} X_1 \\ X_2 \\ X_3 \end{matrix}.$$

Et puisque cette matrice représente f_{A_1} , dont la matrice dans la base canonique est A_1 , elle est semblable à A_1 .

Mieux

Nous avons même prouvé que toutes les matrices de \mathcal{ST}_p ont V comme vecteur propre.

Alternative

On pourrait aussi montrer, à l'aide de la formule du produit matriciel que la somme des coefficients d'une ligne de AB est égale à 1, mais c'est plus laborieux.

3.b.i. A_3 est triangulaire, donc $\text{Spec}(A_3) = \left\{1, \frac{1}{2}\right\}$.

Puisque 1 n'apparaît qu'une fois sur la diagonale de A_3 , $\dim E_1(A_3) = 1$ et

$$\dim E_{1/2}(A_3) = 3 - \text{rg}\left(A_3 - \frac{1}{2}I_3\right) = 3 - \text{rg}\begin{pmatrix} 1/2 & 0 & 0 \\ 1/2 & 0 & 0 \\ 0 & 1/2 & 0 \end{pmatrix} = 3 - 2 = 1.$$

On a donc

$$\sum_{\lambda \in \text{Spec}(A_3)} \dim E_\lambda(A_3) = \dim E_1(A_3) + \dim E_{1/2}(A_3) = 2 < 3.$$

Or, si A_3 était semblable à une matrice diagonale D , celle-ci aurait comme coefficients diagonaux A et $1/2$. L'un de ces deux coefficients devrait alors apparaître deux fois sur la diagonale de D .

Mais alors le sous-espace propre associé de D devrait être de dimension 2, et il en serait de même du sous-espace propre de A_3 (rappelons que deux matrices semblables ont mêmes valeurs propres, et des sous-espaces propres associés de même dimension). Or, les deux sous-espaces propres de A_3 sont de dimension 1, d'où une contradiction :

A_3 n'est pas semblable à une matrice diagonale.

3.b.ii. Nous pourrions utiliser la matrice A_2 pour réfuter cette affirmation, mais plus simplement, notons que $I_3 \in \mathcal{ST}_3$, et que $\dim E_1(I_3) = 3 \neq 1$.

4.a. Par définition, on a $AX = \lambda X$. En particulier, en identifiant la i -ème coordonnée, il vient

$$\lambda x_i = \sum_{k=1}^p a_{i,k} x_k$$

et donc

$$|\lambda x_i| = \left| \sum_{k=1}^p a_{i,k} x_k \right| \leq \sum_{k=1}^p |a_{i,k} x_k| \leq \sum_{k=1}^p a_{i,k} |x_k| \leq \sum_{k=1}^p a_{i,k} |x_i| \leq |x_i| \underbrace{\sum_{k=1}^p a_{i,k}}_{=1} \leq |x_i|.$$

4.b. Puisque X est un vecteur propre de A , il est non nul¹, et donc $|x_i| = \max_{k \in \llbracket 1, p \rrbracket} |x_k| \neq 0$.

¹ Autrement dit, l'un au moins des x_k est non nul.

En divisant alors la relation précédente par $|x_i|$, on obtient

$$|\lambda| \leq 1.$$

Partie II : Suite de moyennes de puissances de matrices stochastiques

5.a. Procédons par récurrence sur \mathbf{N} .

On a $A^0 = I_n$, qui est clairement² stochastique

Supposons que A^n est stochastique.

Alors $A^{n+1} = AA^n \in \mathcal{ST}_p$, d'après la question 2 la partie I.

Par le principe de récurrence, $\forall n \in \mathbf{N}, A^n \in \mathcal{ST}_p$.

² Tous ses coefficients sont positifs, et la somme des coefficients de chaque ligne vaut 1.

5.b. Par la question précédente, chacune des A^k a tous ses coefficients positifs, donc $\sum_{k=0}^{n-1} A^k$ a

également tous ses coefficients positifs. En divisant par n , on a alors $\frac{1}{n} \sum_{k=0}^{n-1} A^k$ qui a tous ses coefficients positifs.

De plus,

$$\frac{1}{n} \left(\sum_{k=0}^{n-1} A^k \right) V = \frac{1}{n} \sum_{k=0}^{n-1} A^k V = \frac{1}{n} \sum_{k=0}^{n-1} V = \frac{1}{n} nV = V.$$

Donc d'après la question 1, la matrice $\frac{1}{n} \sum_{k=0}^{n-1} A^k \in \mathcal{ST}_p$.

6. Commençons par traiter le cas $x = 1$. On a alors $\sum_{k=0}^{n-1} 1^k = n$, et donc

$$\frac{1}{n} \sum_{k=0}^{n-1} 1^k = \frac{n}{n} = 1 \xrightarrow{n \rightarrow \infty} 1.$$

Si $x \neq 1$, alors $\sum_{k=0}^{n-1} x^k = \frac{1-x^n}{1-x}$.

On obtient alors une suite bornée car

$$\left| \frac{1-x^n}{1-x} \right| = \frac{|1-x^n|}{|1-x|} \leq \frac{1+|x|^n}{|1-x|} \leq \frac{2}{|1-x|}.$$

Puisque $\frac{1}{n}$ tend vers 0, on a alors :

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=0}^{n-1} x^k = 0.$$

7. Puisque A et D sont semblables, elles ont les mêmes valeurs propres. Or, A étant stochastique, ses valeurs propres sont de module inférieur à 1. Il en est donc de même de D , et ses valeurs propres étant ses coefficients diagonaux, on en déduit que tous les coefficients diagonaux de D sont de module³ inférieur ou égal à 1.

Ainsi, en utilisant les données de l'énoncé, on a

$$D = \text{diag}(\underbrace{1, \dots, 1}_{r \text{ fois}}, \lambda_{r+1}, \dots, \lambda_p) \text{ avec } \forall i \in \llbracket r+1, p \rrbracket, |\lambda_i| \leq 1 \text{ et } \lambda_i \neq 1.$$

Il vient alors $\forall k \in \mathbf{N}$, $D^k = \text{diag}(1, \dots, 1, \lambda_{r+1}^k, \dots, \lambda_p^k)$. Et donc

$$M_n = \frac{1}{n} \sum_{k=0}^{n-1} \text{diag}(1, \dots, 1, \lambda_{r+1}^k, \dots, \lambda_p^k) = \frac{1}{n} \text{diag} \left(\sum_{k=0}^{n-1} 1, \dots, \sum_{k=0}^{n-1} 1, \sum_{k=0}^{n-1} \lambda_{r+1}^k, \dots, \sum_{k=0}^{n-1} \lambda_p^k \right).$$

En utilisant la question précédente, on a alors :

- si $i \neq j$, $(M_n)_{i,j} = 0 \xrightarrow{n \rightarrow \infty} 0$
- si $i \in \llbracket 1, r \rrbracket$, $(M_n)_{i,i} = 1 \xrightarrow{n \rightarrow \infty} 1$
- si $i \in \llbracket r+1, p \rrbracket$, $(M_n)_{i,i} = \frac{1}{n} \sum_{k=0}^{n-1} \lambda_i^k \xrightarrow{n \rightarrow \infty} 0$

Ainsi, on a bien prouvé que $M_n \xrightarrow{n \rightarrow \infty} \Delta$.

De plus, en utilisant les résultats rappelés au début de l'énoncé : puisque $M_n \xrightarrow{n \rightarrow \infty} \Delta$, alors

$$PM_n \xrightarrow{n \rightarrow \infty} P\Delta, \text{ puis } \boxed{PM_n P^{-1} \xrightarrow{n \rightarrow \infty} P\Delta P^{-1} = B.}$$

- 8.a. Par définition, on a

$$B_n = PM_n P^{-1} = P \left(\frac{1}{n} \sum_{k=0}^{n-1} D^k \right) P^{-1} = \frac{1}{n} \sum_{k=0}^{n-1} (PDP^{-1})^k = \boxed{\frac{1}{n} \sum_{k=0}^{n-1} A^k \in \mathcal{ST}_p.}$$

- 8.b. Il s'agit de prouver qu'une limite de matrices stochastiques est encore une matrice stochastique.

Soit $(i, j) \in \llbracket 1, p \rrbracket^2$. Alors $\forall n \in \mathbf{N}$, $(B_n)_{i,j} \geq 0$ car $B_n \in \mathcal{ST}_p$.

Donc par passage à la limite, $B_{i,j} = \lim_{n \rightarrow \infty} (B_n)_{i,j} \geq 0$.

De plus, $\forall i \in \llbracket 1, p \rrbracket$ et pour tout $n \in \mathbf{N}$, $\sum_{j=1}^p (B_n)_{i,j} = 1$ et donc

$$\sum_{j=1}^p (B_n)_{i,j} = 1 \text{ et donc } \sum_{j=1}^p B_{i,j} = \sum_{j=1}^p \lim_{n \rightarrow \infty} (B_n)_{i,j} = \lim_{n \rightarrow \infty} \sum_{j=1}^p = 1.$$

Donc $\boxed{B \in \mathcal{ST}_p.}$

Danger

Attention à ne pas aller trop vite, pour $x = -1$, il n'est pas question d'utiliser le fait que $x^n \rightarrow 0$: c'est faux !

Rappel

Le produit d'une suite bornée et d'une suite de limite nulle tend vers 0.

³ En fait on sait que ces coefficients sont réels, donc leur module n'est autre que leur valeur absolue.

Détails

Pour $i \geq r+1$, $|\lambda_i| \leq 1$ et $\lambda_i \neq 1$, donc le résultat de la question 6 s'applique.

Partie III : Aspect probabiliste

9. Les coefficients de A sont tous des probabilités, donc sont évidemment positifs. Soit $i \in \llbracket 1, 3 \rrbracket$. Alors, puisque $\{[X_1 = j], j \in \llbracket 1, 3 \rrbracket\}$ forme un système complet d'événements, on a, par la formule des probabilités totales,

$$\begin{aligned} \sum_{j=1}^3 P_{[X_0=i]}(X_1 = j) &= \sum_{j=1}^3 \frac{P([X_0 = i] \cap [X_1 = j])}{P(X_0 = i)} \\ &= \frac{1}{P(X_0 = i)} \sum_{j=1}^3 P([X_0 = i] \cap [X_1 = j]) \\ &= \frac{P(X_0 = i)}{P(X_0 = i)}. \end{aligned}$$

Ainsi, la somme des coefficients de chaque ligne de A est bien égale à 1 : $A \in \mathcal{ST}_p$.

10. Soit $k \in \llbracket 1, 3 \rrbracket$. Alors, en appliquant la formule des probabilités totales au système complet d'événements $\{[X_n = i], i \in \llbracket 1, 3 \rrbracket\}$, il vient

$$P(X_{n+1} = j) = \sum_{i=1}^3 P(X_n = i)P_{(X_n=i)}(X_{n+1} = j) = \sum_{i=1}^3 P(X_n = i)P_{(X_0=i)}(X_1 = j).$$

Or, nous reconnaissons là le coefficient à la j -ème colonne du produit matriciel $L_n A$. Donc $L_{n+1} = L_n A$.

Par une récurrence rapide, on prouve donc que $\forall n \in \mathbf{N}, L_n = L_0 A^n$.

11. Soit f l'endomorphisme de \mathbf{R}^3 dont la matrice dans la base canonique de \mathbf{R}^3 est A_1 . Alors il s'agit de trouver une base $\mathcal{B} = (e_1, e_2, e_3)$ de \mathbf{R}^3 telle que

$$\text{Mat}_{\mathcal{B}}(f) = \begin{pmatrix} f(e_1) & f(e_2) & f(e_3) \\ 1 & 0 & 0 \\ 0 & 1/2 & 0 \\ 0 & 0 & 1/3 \end{pmatrix} \begin{matrix} e_1 \\ e_2 \\ e_3 \end{matrix}.$$

De plus, nous souhaiterions que $P_{\mathcal{B}_{can}, \mathcal{B}}$ possède des coefficients diagonaux égaux à 1, c'est-à-dire que si (f_1, f_2, f_3) est la base canonique de \mathbf{R}^3 , alors on souhaite avoir

$$P_{\mathcal{B}_{can}, \mathcal{B}} = \begin{pmatrix} e_1 & e_2 & e_3 \\ 1 & \star & \star \\ \star & 1 & \star \\ \star & \star & 1 \end{pmatrix} \begin{matrix} f_1 \\ f_2 \\ f_3 \end{matrix}.$$

Cherchons e_1 vérifiant $f(e_1) = e_1$, sous la forme $e_1 = (1, y, z)$.

Ceci revient à résoudre le système $A_1 \begin{pmatrix} 1 \\ y \\ z \end{pmatrix} = \begin{pmatrix} 1 \\ y \\ z \end{pmatrix}$. Soit encore

$$\begin{cases} 1 = 1 \\ 1/2 + 1/2y = y \\ 1/3 + 1/3y + 1/3z = z \end{cases} \Leftrightarrow \begin{cases} y = 1 \\ z = 1 \end{cases}$$

Et donc la seule solution est $e_1 = (1, 1, 1)$.

De même, on cherche $e_2 \in E_{1/2}(f)$ sous la forme $e_2 = (x, 1, z)$, ce qui après calculs, possède pour unique solution $e_2 = (0, 1, 2)$.

Enfin, on obtient de même $e_3 = (0, 0, 1)$.

Nous savons, sans calculs, que la famille (e_1, e_2, e_3) est libre car formée de trois vecteurs propres associés à des valeurs propres distinctes, et donc $\mathcal{B} = (e_1, e_2, e_3)$ est une base de \mathbf{R}^3 car libre et de cardinal 3.

Par construction, la matrice de f dans la base \mathcal{B} est D_1 , et

$$P_{\mathcal{B}_{can}, \mathcal{B}} = \begin{pmatrix} e_1 & e_2 & e_3 \\ 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 2 & 1 \end{pmatrix} \begin{matrix} f_1 \\ f_2 \\ f_3 \end{matrix}.$$

Alternative

On peut aussi directement affirmer que cette somme vaut 1 en remarquant qu'il s'agit de la formule des probabilités totales appliquée à la probabilité $P_{[X_0=i]}$ au lieu de l'appliquer à P .

Explication

Les \star sont des coefficients qui peuvent prendre n'importe quelle valeur, les seules contraintes portant sur les coefficients diagonaux.

Par la formule de changement de base, on a donc

$$A_1 = \text{Mat}_{\mathcal{B}_{can}}(f) = P_{\mathcal{B}_{can}, \mathcal{B}} \text{Mat}_{\mathcal{B}}(f) P_{\mathcal{B}, \mathcal{B}_{can}} = P_1 D_1 P_1^{-1}.$$

De plus, un calcul d'inverse nous donne $P_1^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 1 & -2 & 1 \end{pmatrix}$.

12. On a $D_1^n = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{2^n} & 0 \\ 0 & 0 & \frac{1}{3^n} \end{pmatrix}$, de sorte que

$$\lim_{n \rightarrow \infty} D_1^n = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} = D.$$

On en déduit comme dans la question 7 que

$$\lim_{n \rightarrow \infty} A_1^n = P_1 D P_1^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}.$$

13. Notons $L_0 = (a \ b \ c)$. Puisque $\{[X_0 = i], i \in \llbracket 1, 3 \rrbracket\}$ est un système complet d'événements, on a $a + b + c = 1$.
De plus, comme expliqué dans l'énoncé, on a

$$\lim_{n \rightarrow \infty} L_n = \lim_{n \rightarrow \infty} L_0 A_1^n = L_0 \left(\lim_{n \rightarrow \infty} A_1^n \right) = (a \ b \ c) \begin{pmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix} = (a + b + c \ 0 \ 0) = \begin{pmatrix} 1 & 0 & 0 \end{pmatrix}.$$

En termes de probabilités, cela signifie que

$$\lim_{n \rightarrow +\infty} P(X_n = 1) = 1, \quad \lim_{n \rightarrow +\infty} P(X_n = 2) = 0, \quad \lim_{n \rightarrow +\infty} P(X_n = 3) = 0.$$

Ce résultat s'explique par le fait que si l'objet est dans l'urne 1, il n'en sort pas :

$$P_{X_n=1}(X_{n+1} = 1) = 1.$$

De plus, si l'objet se trouve dans la seconde urne à un instant donné, il a une probabilité 1/2 de se trouver dans la première à l'instant suivant (et donc de ne plus en sortir), et de même, depuis la troisième urne, il a une probabilité 1/3 de se trouver dans la première à l'instant suivant, puis de ne plus en sortir.

Ainsi, quel que soit l'endroit où se trouve l'objet au départ, au bout d'un temps suffisamment long, il finira presque sûrement par se retrouver dans la première urne, puis n'en sortira plus.

Ceci correspond au résultat du calcul : quel que soit $(a \ b \ c)$ l'état de départ, alors au bout d'un temps assez long, l'objet se trouvera dans l'urne 1 avec probabilité 1 (et donc dans les autres avec probabilité nulle).

Remarque

Ce type de situations sera longuement étudié en TP, c'est ce qu'on appelle des chaînes de Markov.

DEVOIR MAISON 8 (À RENDRE LE 6.12.18)

EXERCICE 1

On considère deux variables aléatoires X et Y , définies toutes les deux sur le même espace probabilisé (Ω, \mathcal{A}, P) , indépendantes, et suivant la loi uniforme sur $[0, 1]$. On pose alors $Z = X + Y$.

1.
 - a. Déterminer une densité de Z
 - b. Montrer que pour tout $x \in]0, 1[$, $[Z > 1]$ et $[1 - x < Z < 1 + x]$ sont deux événements indépendants.
2. On pose $T = \text{Max}(X, Y)$. On admet que T est une variable aléatoire définie elle aussi sur l'espace probabilisé (Ω, \mathcal{A}, P) .
 - a. Montrer que T est une variable à densité, puis donner une densité de T .
 - b. En déduire que T possède une espérance $E(T)$ et la déterminer.
 - c. On pose $U = |X - Y|$, et on admet que U est une variable aléatoire définie elle aussi sur l'espace probabilisé (Ω, \mathcal{A}, P) . Montrer que U est combinaison linéaire de Z et de T , puis en déduire l'espérance de U .
 - d. Proposer un code Sci Lab permettant de vérifier numériquement le résultat de la question précédente, par exemple à l'aide de 10000 simulations de la variable U .

EXERCICE 2

Soit E un espace vectoriel euclidien de dimension n , dont le produit scalaire est noté $\langle \cdot, \cdot \rangle$.

L'objectif du problème est d'étudier les endomorphismes u de E tels que :

$$\forall x \in E, \langle u(x), x \rangle = 0.$$

Les endomorphismes vérifiant cette propriété sont appelés endomorphismes antisymétriques.

Partie I. Étude d'un exemple.

Dans cette partie, $E = \mathbf{R}_2[X]$. On rappelle que $(1, X, X^2)$ est une base de E .

On considère l'application $\varphi : E^2 \rightarrow \mathbf{R}$ définie pour tout couple (P, Q) d'éléments de E par :

$$\varphi(P, Q) = P(0)Q(0) + P(1)Q(1) + P(-1)Q(-1).$$

1. Vérifier que φ est un produit scalaire.

Dans cette première partie, on considère que E est muni de ce produit scalaire.

2. On considère l'endomorphisme u de E défini pour tout P de E par :

$$u(P) = 2P'(0)X^2 - (P(1) + P(-1))X.$$

- a. Vérifier : $\forall P \in E, 2P'(0) - P(1) + P(-1) = 0$.

- b. En déduire que u est un endomorphisme antisymétrique de l'espace vectoriel euclidien E .

3. Soient $P_1 = \frac{1}{2}(X^2 + X)$ et $P_2 = \frac{1}{2}u(P_1)$.

- a. Vérifier que P_1 est un vecteur propre de u^2 et que la famille (P_1, P_2) est orthonormale.

- b. Déterminer une base de $\text{Ker}(u)$.

- c. Déterminer une base orthonormale \mathcal{B} de E et un nombre réel a tels que la matrice associée à u relativement à

cette base soit $\begin{pmatrix} 0 & -a & 0 \\ a & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$.

Partie II. Caractérisation des endomorphismes antisymétriques.

Soit u un endomorphisme de E .

4. Pour tout couple (x, y) de E^2 , développer $\langle u(x + y), x + y \rangle$.

En déduire que u est un endomorphisme antisymétrique si et seulement si :

$$\forall (x, y) \in E^2, \langle u(x), y \rangle = -\langle x, u(y) \rangle.$$

5. On suppose dans cette question que la dimension n de E est non nulle.

Soient $\mathcal{B} = (e_1, e_2, \dots, e_n)$ une base orthonormale de E , et $M = (m_{i,j})_{1 \leq i, j \leq n}$ la matrice associée à u relativement à la base \mathcal{B} .

- a. Montrer : $\forall (i, j) \in \{1, \dots, n\}^2, m_{i,j} = \langle e_i, u(e_j) \rangle$.

- b. En déduire que u est un endomorphisme antisymétrique si et seulement si la matrice M associée à u relativement à la base \mathcal{B} vérifie ${}^t M = -M$.

CORRECTION DU DEVOIR MAISON 8

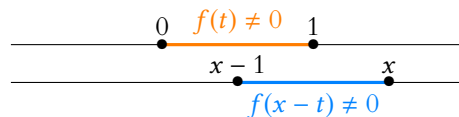
EXERCICE 1 (EDHEC 2004)

- 1.a. Notons $f : t \mapsto \begin{cases} 1 & \text{si } 0 \leq t \leq 1 \\ 0 & \text{sinon} \end{cases}$ une densité de la loi $\mathcal{U}([0, 1])$.

Puisque f est bornée, une densité de Z est donnée par

$$hx \mapsto \int_{-\infty}^{\infty} f(t)f(x-t) dt$$

On a $f(t) \neq 0 \Leftrightarrow 0 \leq t \leq 1$ et $f(x-t) \neq 0 \Leftrightarrow 0 \leq x-t \leq 1 \Leftrightarrow x-1 \leq t \leq x$.



Donc pour $x \leq 0$ ou $x-1 \geq x \Leftrightarrow x \geq 2$, on a $h(x) = 0$.

Si $x \in [0, 1]$, alors

$$h(x) = \int_0^x 1 \times 1 dt = x.$$

Et si $x \in]1, 2]$, alors

$$h(x) = \int_{x-1}^1 1 \times 1 dt = 1 - (x-1) = 2-x.$$

On en déduit que

$$h(x) = \begin{cases} x & \text{si } 0 \leq x \leq 1 \\ 2-x & \text{si } 1 < x \leq 2 \\ 0 & \text{sinon} \end{cases}$$

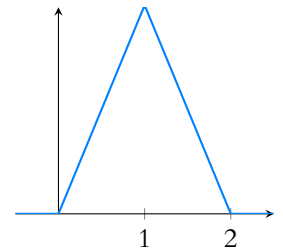


FIGURE 23.1– La densité h .

- 1.b. On a, pour $x \in]0, 1[$,

$$[Z > 1] \cap [1-x < Z < 1+x] = [1 < Z < 1+x]$$

Donc il vient

$$\begin{aligned} P([X > 1] \cap [1-x < Z < 1+x]) &= \int_1^{1+x} h(t) dt = \int_1^{1+x} (2-t) dt \\ &= \left[2t - \frac{t^2}{2} \right]_1^{1+x} = 2 + 2x - \frac{(1+x)^2}{2} - 2 + \frac{1}{2} \\ &= x - \frac{x^2}{2}. \end{aligned}$$

D'autre part, $P(Z > 1) = \int_1^{+\infty} h(t) dt = \int_1^2 (2-t) dt = \frac{1}{2}$ et

$$P(1-x < Z < 1+x) = \int_{1-x}^{1+x} h(t) dt = \int_{1-x}^1 t dt + \int_1^{1+x} (2-t) dt = \left[\frac{t^2}{2} \right]_{1-x}^1 + \left[2t - \frac{t^2}{2} \right]_1^{1+x} = 2x - x^2.$$

On a donc bien

$$P([Z > 1] \cap [1-x < Z < 1+x]) = P(Z > 1)P(1-x < Z < 1+x)$$

et donc les événements $[Z > 1]$ et $[1-x < Z < 1+x]$ sont indépendants.

- 2.a. Puisque X et Y sont à valeurs dans $[0, 1]$, T est également à valeurs dans $[0, 1]$. Ainsi, nous savons déjà que sa fonction de répartition F_T vérifie

$$F_T(x) = \begin{cases} 0 & \text{si } x < 0 \\ 1 & \text{si } x \geq 1 \end{cases}$$

Pour $x \in [0, 1[$, on a

$$[T \leq x] = [X \leq x] \cap [Y \leq x]$$

et donc, par indépendance de X et Y

$$F_T(x) = P(T \leq x) = P([X \leq x] \cap [Y \leq x]) = P(X \leq x)P(Y \leq x).$$

Mais X et Y suivent des lois uniformes sur $[0, 1]$, de sorte que $P(X \leq x) = P(Y \leq x) = x$, et on en déduit que $F_T(x) = x^2$. Donc

$$F_T(x) = \begin{cases} 0 & \text{si } x \leq 0 \\ x^2 & \text{si } 0 < x^2 \leq 1 \\ 1 & \text{si } x \geq 1 \end{cases}$$

Ainsi, la fonction F_T est continue sur \mathbf{R}_* , et sur $]1; +\infty[$ (car elle y est constante), ainsi que sur $]0, 1[$ car elle y est polynomiale. Enfin on a

$$\lim_{x \rightarrow 0^-} F_T(x) = 0 = F_T(0) = \lim_{x \rightarrow 0^+} F_T(x).$$

Donc F_T est continue en 0.

De même,

$$\lim_{x \rightarrow 1^-} F_T(x) = 1 = F_T(1) = \lim_{x \rightarrow 1^+} F_T(x)$$

et donc F_T est continue en 1. Pour conclure, F_T est continue sur \mathbf{R} .

De plus, F_T est dérivable sur \mathbf{R}_* et sur $]1; +\infty[$, de dérivée nulle. Elle est également dérivable sur $]0, 1[$, avec $F'_T(x) = 2x$.

Nous venons donc de prouver que F_T est continue sur \mathbf{R} , dérivable et de classe \mathcal{C}^1 sur \mathbf{R} privé de 0 et 1 : la variable T admet une densité.

De plus, une densité de T est donnée par toute fonction coïncidant avec F'_T là où celle-ci est définie, c'est-à-dire sur $\mathbf{R} \setminus \{0, 1\}$. Par exemple, on peut prendre

$$g : t \mapsto \begin{cases} 2t & \text{si } 0 \leq t \leq 1 \\ 0 & \text{sinon} \end{cases}$$

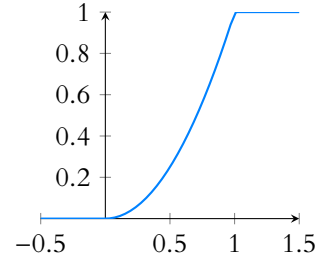


FIGURE 23.2– La fonction de répartition de T .

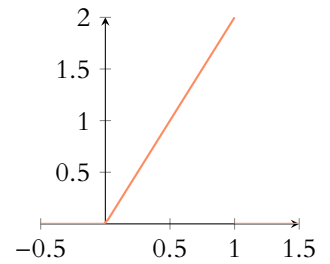


FIGURE 23.3– La densité de T .

2.b. T possède une espérance si et seulement si

$$\int_{-\infty}^{+\infty} tg(t) dt = \int_0^1 2t^2 dt \text{ converge (absolument, mais il s'agit d'une fonction positive).}$$

Or, il s'agit là d'une intégrale sur un segment, donc elle converge et vaut $\frac{2}{3}$. Ainsi

$$E(T) \text{ existe et vaut } \frac{2}{3}.$$

2.c. Notons que $Z - T = X + Y - \max(X, Y) = \min(X, Y)$.

Or, $U = |X - Y| = \max(X, Y) - \min(X, Y) = T - (Z - T) = 2T - Z$.

On en déduit que U admet une espérance, et que celle-ci vaut

$$E(U) = 2E(T) - E(Z) = 2E(T) - (E(X) + E(Y)) = 2 \cdot \frac{2}{3} - 1 = \frac{1}{3}.$$

2.d. Il s'agit de simuler 10000 réalisations de U et de calculer la moyenne de l'échantillon obtenu. Celle-ci devrait être proche de $1/3$.

```
1 X = grand(1, 10000, 'unf', 0, 1);
2 Y = grand(100, 10000, 'unf', 0, 1);
3 U = abs(X-Y);
4 disp(mean(U))
```

Logique

Si on a la somme des deux nombres ($X + Y$), et qu'on enlève le plus grand de ces deux nombres, il ne reste que le plus petit des deux.

EXERCICE 2 (EML 2002)

Partie I. Étude d'un exemple

1. Soient $P, Q \in \mathbf{R}_2[X]$. Alors

$$\varphi(P, Q) = \sum_{i=0}^2 P(i)Q(i) = \sum_{i=0}^2 Q(i)P(i) = \varphi(Q, P).$$

Donc φ est symétrique.

Soient $P, Q, R \in \mathbf{R}_2[X]$ et $\lambda \in \mathbf{R}$. Alors

$$\begin{aligned} \varphi(\lambda P + Q, R) &= \sum_{i=0}^2 (\lambda P + Q)(i)R(i) = \sum_{i=0}^2 (\lambda P(i) + Q(i))R(i) \\ &= \lambda \sum_{i=0}^2 P(i)R(i) + \sum_{i=0}^2 Q(i)R(i) \\ &= \lambda \varphi(P, R) + \varphi(Q, R). \end{aligned}$$

Donc φ est linéaire à gauche, et étant symétrique, est donc bilinéaire.

Pour $P \in \mathbf{R}_2[X]$, on a

$$\varphi(P, P) = \sum_{i=0}^2 P(i)^2 \geq 0.$$

Et alors $\varphi(P, P) = 0$ si et seulement si pour tout $i \in \llbracket 0, 2 \rrbracket$, $P(i)^2 = 0 \Leftrightarrow P(i) = 0$.

Le polynôme P possède alors trois racines¹ distinctes. Étant de degré au plus 2, il est donc nul.

Et donc on a $\varphi(P, P) = 0 \Leftrightarrow P = 0$.

Ceci achève bien de prouver que φ est un produit scalaire sur $\mathbf{R}_2[X]$.

2.a. Soit $P = aX^2 + bX + c \in E$. Alors $P'(X) = 2aX + b$ et donc

$$2P'(0) - P(1) + P(-1) = 2b - (2a - b + c) + (2a + b + c) = 0.$$

2.b. Soit $P \in E$. On a

$$\begin{aligned} \varphi(u(P), P) &= \underbrace{[u(P)](0)}_{=0} P(0) + [u(P)](1)P(1) + [u(P)](-1)P(-1) \\ &= (2P'(0) - P(1) - P(-1))P(1) + (2P'(0) + P(1) + P(-1))P(-1) \\ &= (2P'(0) - P(1) + P(-1))(P(1) + P(-1)) \\ &= 0. \end{aligned}$$

Et donc u est un endomorphisme antisymétrique.

3.a. On a $u(P_1) = 2P_1'(0)X^2 - (P_1(1) + P_1(-1))X = X^2 - X$.

Et donc $u^2(P_1) = -2X^2 - X = -4\left(\frac{1}{2}(X^2 + X)\right) = -4P_1$.

Ceci prouve donc que P_1 est un vecteur propre de u^2 , associé à la valeur propre -4 .

On a alors $\varphi(P_1, P_2) = \varphi\left(P_1, \frac{1}{2}u(P_1)\right) = \frac{1}{2}\varphi(P_1, u(P_1)) = 0$ car u est antisymétrique.

Donc P_1 et P_2 sont orthogonaux. D'autre part,

$$\varphi(P_1, P_1) = \underbrace{P_1(0)^2}_{=0} + P_1(1)^2 + \underbrace{P_1(-1)^2}_{=0} = P_1(1)^2 = 1.$$

Et donc $\|P_1\| = 1$.

Puisque $P_2 = \frac{1}{2}(X^2 - X)$, il vient

$$\varphi(P_2, P_2) = P_2(0)^2 + P_2(1)^2 + P_2(-1)^2 = P_2(-1)^2 = 1.$$

Et donc $\|P_2\| = 1$, ce qui achève de prouver que (P_1, P_2) est orthonormée.

¹ Qui sont 0, 1 et 2.

Notations

Attention ici aux notations : $u(P)$ est un polynôme de $\mathbf{R}_2[X]$.

Et donc $[u(P)](0)$ est le réel que l'on obtient en évaluant ce polynôme en $X = 0$.

Norme

Rappelons que par définition,

$$\|P_1\| = \sqrt{\varphi(P_1, P_1)}.$$

3.b. Soit $P = aX^2 + bX + c \in E$. Alors $P \in \text{Ker}(u)$ si et seulement si

$$\begin{aligned} 2P'(0)X^2 - (P(1) + P(-1))X = 0 &\Leftrightarrow \begin{cases} P'(0) = 0 \\ P(1) + P(-1) = 0 \end{cases} \\ &\Leftrightarrow \begin{cases} b = 0 \\ a + b + c = -(a - b + c) \end{cases} \\ &\Leftrightarrow \begin{cases} b = 0 \\ c = -a \end{cases} \\ &\Leftrightarrow P = aX^2 - a \Leftrightarrow P \in \text{Vect}(X^2 - 1). \end{aligned}$$

Deux polynômes sont égaux si et seulement si ils ont les mêmes coefficients.

Et donc $\text{Ker}(u)$ est de dimension 1 et possède $X^2 - 1$ comme base.

3.c. Si la base \mathcal{B} que nous cherchons est (Q_1, Q_2, Q_3) , alors nous voulons

$$\text{Mat}_{\mathcal{B}}(u) = \begin{pmatrix} u(Q_1) & u(Q_2) & u(Q_3) \\ 0 & -a & 0 \\ a & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{matrix} Q_1 \\ Q_2 \\ Q_3 \end{matrix}.$$

En particulier, on souhaite avoir $Q_3 \in \text{Ker}(u)$. Nous pouvons prendre $Q_3 = X^2 - 1$, en remarquant que $\|X^2 - 1\| = 1$.

D'autre part, nous savons que $u(P_1) = 2P_2$ et $u(P_2) = \frac{1}{2}u^2(P_1) = -2P_2$, donc nous pouvons prendre $Q_1 = P_1$ et $Q_2 = P_2$.

Mais il faut encore vérifier que $(P_1, P_2, X^2 - 1)$ est orthonormée. Comme nous savons déjà que (P_1, P_2) est orthonormée, et que $X^2 - 1$ est unitaire, il reste seulement à vérifier que P_1 et P_2 sont orthogonaux à $X^2 - 1$. Or

$$\varphi(P_1, X^2 - 1) = P_1(0) = 0 \text{ et } \varphi(P_2, X^2 - 1) = P_2(0) = 0.$$

Ainsi, $(P_1, P_2, X^2 - 1)$ est bien une famille orthonormée de E . Elle est donc libre, et étant de cardinal 3, c'est une base de E , dans laquelle la matrice de u est donnée par

$$\text{Mat}_{\mathcal{B}}(u) = \begin{pmatrix} u(P_1) & u(P_2) & u(X^2 - 1) \\ 0 & -2 & 0 \\ 2 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{matrix} P_1 \\ P_2 \\ X^2 - 1 \end{matrix}$$

Partie II. Caractérisation des endomorphismes antisymétriques.

4. Supposons u antisymétrique. Alors pour $(x, y) \in E^2$, il vient

$$\begin{aligned} \langle u(x + y), x + y \rangle &= \langle u(x) + u(y), x + y \rangle \\ &= \langle u(x), x + y \rangle + \langle u(y), x + y \rangle \\ &= \underbrace{\langle u(x), x \rangle}_{=0} + \langle u(x), y \rangle + \langle u(y), x \rangle + \underbrace{\langle u(y), y \rangle}_{=0} \\ &= \langle u(x), y \rangle + \langle u(y), x \rangle. \end{aligned}$$

Linéarité de u .
Linéarité à gauche de $\langle \cdot, \cdot \rangle$.
Linéarité à droite de $\langle \cdot, \cdot \rangle$.

D'autre part, u tant antisymétrique, on a $\langle u(x + y), x + y \rangle = 0$, de sorte que

$$\langle u(x), y \rangle = -\langle x, u(y) \rangle.$$

Inversement, si pour tout $(x, y) \in E^2$, on a $\langle u(x), y \rangle = -\langle x, u(y) \rangle$, alors en prenant $y = x$, il vient,

$$\forall x \in E, \langle u(x), x \rangle = -\langle x, u(x) \rangle = -\langle u(x), x \rangle.$$

Et donc $\forall x \in E, \langle u(x), x \rangle = 0$: u est antisymétrique.

² Le seul réel égal à son opposé est 0.

- 5.a. Par définition de la matrice M , la $j^{\text{ème}}$ colonne de M est le vecteur colonne des coordonnées de $u(e_j)$ dans la base \mathcal{B} . Soit encore

$$u(e_j) = \sum_{i=1}^n m_{i,j} e_i.$$

Et donc, par linéarité à droite du produit scalaire,

$$\langle e_i, u(e_j) \rangle = \left\langle e_i, \sum_{k=1}^n m_{k,j} e_k \right\rangle = \sum_{k=1}^n m_{k,j} \underbrace{\langle e_i, e_k \rangle}_{=0 \text{ si } k \neq i} = m_{i,j} \langle e_i, e_i \rangle = \boxed{m_{i,j}}.$$

- 5.b. Si u est antisymétrique, alors pour tout $(i, j) \in \llbracket 1, n \rrbracket^2$, on a

$$m_{i,j} = \langle e_i, u(e_j) \rangle = -\langle u(e_i), e_j \rangle = -m_{j,i}.$$

Et donc $\boxed{M \text{ est antisymétrique.}}$

Inversement, si M est antisymétrique, soient $(x, y) \in E^2$. Il existe alors des réels $\lambda_1, \dots, \lambda_n$ et μ_1, \dots, μ_n tels que

$$x = \sum_{i=1}^n \lambda_i e_i \text{ et } y = \sum_{j=1}^n \mu_j e_j.$$

Et donc $u(y) = \sum_{j=1}^n \mu_j u(e_j)$, de sorte que

$$\langle x, u(y) \rangle = \left\langle \sum_{i=1}^n \lambda_i e_i, \sum_{j=1}^n \mu_j u(e_j) \right\rangle = \sum_{i=1}^n \sum_{j=1}^n \lambda_i \mu_j \langle e_i, u(e_j) \rangle = \sum_{i=1}^n \sum_{j=1}^n \lambda_i \mu_j m_{i,j}.$$

Et de même³, on a

$$\langle u(x), y \rangle = \langle y, u(x) \rangle = \sum_{j=1}^n \mu_j \lambda_j m_{i,j}.$$

Si M est antisymétrique, on a alors pour tout $(i, j) \in \llbracket 1, n \rrbracket^2$, $m_{i,j} = -m_{j,i}$ et donc

$$\langle x, u(y) \rangle = \sum_{i=1}^n \sum_{j=1}^n \lambda_i \mu_j m_{i,j} = - \sum_{i=1}^n \sum_{j=1}^n \lambda_i \mu_j m_{j,i} = - \sum_{j=1}^n \sum_{i=1}^n \lambda_j \mu_i m_{i,j} = -\langle u(x), y \rangle.$$

Et donc $\boxed{u \text{ est antisymétrique.}}$

Rappel

Une matrice est antisymétrique si ${}^t M = -M$.
Par identification des coefficients, c'est le cas si et seulement si $\forall (i, j) \in \llbracket 1, n \rrbracket^2$,

$$m_{i,j} = -m_{j,i}.$$

³ Il suffit d'inverser les rôles de x et y .

Détails

Les variables i et j étant muettes, on peut sans problème les permuter, c'est-à-dire remplacer tous les i par des j et tous les j par des i .

DEVOIR MAISON 9 (À RENDRE LE 13.12.18)

Vous traiterez **au choix** l'un des deux exercices suivants, le second étant plus difficile que le premier.

EXERCICE 1

$\mathcal{M}_n(\mathbf{R})$ désigne l'espace vectoriel des matrices carrées d'ordre n à coefficients réels ($n \geq 1$) et E l'espace vectoriel des polynômes à coefficients réels de degré inférieur ou égal à $n - 1$.

On considère une matrice S de $\mathcal{M}_n(\mathbf{R})$ admettant n valeurs propres réelles $\lambda_1, \dots, \lambda_n$ distinctes deux à deux.

L'objet de cet exercice est de montrer que, si k est un entier naturel impair et si une matrice A de $\mathcal{M}_n(\mathbf{R})$ commute avec S^k , alors elle commute avec S .

1. Justifier l'existence d'une matrice P inversible telle que la matrice $P^{-1}SP$ soit une matrice D diagonale.

Dans la suite de l'exercice, un entier naturel impair k est fixé.

2. On considère l'application f de E dans \mathbf{R}^n qui à tout polynôme T fait correspondre le vecteur de \mathbf{R}^n défini par :

$$f(T) = (T(\lambda_1^k), T(\lambda_2^k), \dots, T(\lambda_n^k)).$$

- a. Montrer que f est un isomorphisme d'espaces vectoriels.
- b. En déduire l'existence d'un unique polynôme U de E tel que :

$$U(\lambda_1^k) = \lambda_1, U(\lambda_2^k) = \lambda_2, \dots, U(\lambda_n^k) = \lambda_n.$$

3. Prouver que le polynôme R , défini par $R(X) = U(X^k) - X$ est un polynôme annulateur de D puis de S .

4. Soit une matrice A de $\mathcal{M}_n(\mathbf{R})$ vérifiant $AS^k = S^kA$.

- a. Montrer que pour tout entier naturel p , $AS^{pk} = S^{pk}A$.
- b. En déduire que les matrices A et S commutent, c'est-à-dire que $AS = SA$.

5. On considère les deux matrices A et S de $\mathcal{M}_2(\mathbf{R})$ suivantes : $A = \begin{pmatrix} 1 & -1 \\ 2 & 2 \end{pmatrix}$, $S = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$.

- a. Vérifier que S possède deux valeurs propres distinctes.
- b. Montrer que A commute avec toute puissance paire de S , mais ne commute pas avec S .

EXERCICE 2

$\mathcal{M}_n(\mathbf{R})$ désigne l'ensemble des matrices carrées d'ordre n à coefficients réels.

Pour tout élément $A = (a_{i,j})_{1 \leq i,j \leq n}$ de $\mathcal{M}_n(\mathbf{R})$, on note $\text{tr}(A)$ la trace de A et tA la transposée de A .

Pour toutes matrices M, N de $\mathcal{M}_n(\mathbf{R})$, on pose :

$$\langle M|N \rangle = \text{tr}({}^tMN) = \sum_{i=1}^n \sum_{j=1}^n m_{i,j}n_{i,j}$$

où $m_{i,j}$ (resp. $n_{i,j}$) désigne le coefficient de M (resp. N) situé à l'intersection de la i -ième ligne et de la j -ième colonne. Soit A une matrice symétrique diagonalisable, on considère :

- l'application Φ_A de $\mathcal{M}_n(\mathbf{R})$ dans $\mathcal{M}_n(\mathbf{R})$ définie par : $\forall M \in \mathcal{M}_n(\mathbf{R}), \Phi_A(M) = AM - MA$.
- l'ensemble $\text{Sp}(A)$ formé des valeurs propres de A .
- l'ensemble $\text{Sp}(\Phi_A)$ formé des valeurs propres de Φ_A .
- l'ensemble $\Gamma = \{\lambda - \mu, (\lambda, \mu) \in \text{Sp}(A)^2\}$ formé des différences de deux valeurs propres quelconques de A .

Le but de cet exercice est d'établir que les deux propriétés suivantes sont valables pour toute matrice symétrique diagonalisable à coefficients réels A :

- ★ Φ_A est un endomorphisme diagonalisable,
- ★ les valeurs propres de Φ_A forment l'ensemble Γ , c'est-à-dire que $\text{Sp}(\Phi_A) = \Gamma$.

Partie I : Étude d'un cas particulier.

Dans cette partie uniquement, on suppose que $n = 2$, $A = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$ et on admet les deux propriétés suivantes :

- Φ_A est un endomorphisme de $\mathcal{M}_2(\mathbf{R})$,
- la famille (V_1, V_2, V_3, V_4) est une base de $\mathcal{M}_2(\mathbf{R})$ où l'on a posé :

$$V_1 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, V_2 = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, V_3 = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, V_4 = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}.$$

1. Justifier que la matrice T de l'endomorphisme Φ_A dans la base (V_1, V_2, V_3, V_4) s'écrit

$$T = \begin{pmatrix} 0 & -1 & 1 & 0 \\ -1 & 0 & 0 & 1 \\ 1 & 0 & 0 & -1 \\ 0 & 1 & -1 & 0 \end{pmatrix}.$$

2. Vérifier que $T^3 = 4T$. Qu'en déduit-on sur les valeurs propres de T ?
 3. Déterminer une base de l'espace propre associé à la valeur propre 0 de la matrice T .

4. Calculer TX_1 et TX_2 , où $X_1 = \begin{pmatrix} 1 \\ -1 \\ 1 \\ -1 \end{pmatrix}$ et $X_2 = \begin{pmatrix} 1 \\ 1 \\ -1 \\ -1 \end{pmatrix}$.

En déduire que Φ_A est diagonalisable.

5. Expliciter alors une matrice inversible P et une matrice diagonale D telles que $T = PDP^{-1}$ (on ne demande pas le calcul de P^{-1}).

Partie II : Réduction de Φ_A dans le cas général.

On revient désormais au cas général, A étant une matrice symétrique diagonalisable quelconque de $\mathcal{M}_n(\mathbf{R})$.

6. Montrer que Φ_A est un endomorphisme de $\mathcal{M}_n(\mathbf{R})$.
 7. Prouver que l'application $(M, N) \in (\mathcal{M}_n(\mathbf{R}))^2 \mapsto \langle M|N \rangle$ est un produit scalaire sur $\mathcal{M}_n(\mathbf{R})$.
 8. Établir que, pour toutes matrices M, N appartenant à $\mathcal{M}_n(\mathbf{R})$, on a :

$$\langle \Phi_A(M)|N \rangle = \langle M|\Phi_A(N) \rangle.$$

On admet (temporairement) que cette propriété suffit à justifier que Φ_A est diagonalisable.

9. Soient

- $X \in \mathcal{M}_{n,1}(\mathbf{R})$ un vecteur propre de A associé à la valeur propre λ ,
- $Y \in \mathcal{M}_{n,1}(\mathbf{R})$ un vecteur propre de A associé à la valeur propre μ .

On pose alors

$$M_{X,Y} = X^t Y \in \mathcal{M}_n(\mathbf{R}).$$

- a. Justifier que $M_{X,Y} \neq 0$, puis que ${}^t Y A = \mu {}^t Y$.
 b. Établir que $\Phi_A(M_{X,Y}) = (\lambda - \mu)M_{X,Y}$ puis que $\Gamma \subset \text{Sp}(\Phi_A)$.
 10. Soit $M \in \mathcal{M}_n(\mathbf{R})$ un vecteur propre de Φ_A associé à la valeur propre α .
 a. On suppose que pour tout vecteur propre Z de A , on a $MZ = 0$.
 Montrer alors que $M = 0$.
 En déduire qu'il existe au moins un vecteur propre Z_0 de A tel que $MZ_0 \neq 0$.
 On note μ la valeur propre associée à Z_0 .
 b. En revenant à l'expression de $\Phi_A(M)$, justifier que MZ_0 est un vecteur propre de A pour une valeur propre dont on précisera l'expression à l'aide de α et μ .
 c. Conclure.

CORRECTION DU DEVOIR MAISON 9

EXERCICE 1 (ÉCRICOME 2004)

1. Puisque $S \in \mathcal{M}_n(\mathbf{R})$ admet n valeurs propres distinctes, alors elle est diagonalisable. Il existe donc $P \in GL_n(\mathbf{R})$ telle que $P^{-1}SP$ soit diagonale.

2.a. Commençons par prouver que f est linéaire : soient $P, Q \in E$ et $\lambda \in \mathbf{R}$. Alors

$$\begin{aligned} f(\lambda P + Q) &= ((\lambda P + Q)(\lambda_1^k), \dots, (\lambda P + Q)(\lambda_n^k)) \\ &= (\lambda P(\lambda_1^k) + Q(\lambda_1^k), \dots, \lambda P(\lambda_n^k) + Q(\lambda_n^k)) = \lambda (P(\lambda_1^k), \dots, P(\lambda_n^k)) + (Q(\lambda_1^k), \dots, Q(\lambda_n^k)) \\ &= \lambda f(P) + f(Q). \end{aligned}$$

Donc f est linéaire.

De plus, la fonction $x \mapsto x^k$ réalise une bijection de \mathbf{R} dans \mathbf{R} car k est impair. Et donc, les λ_i étant deux à deux distincts, il en est de même des λ_i^k .

Ainsi, si $T \in \text{Ker}(f)$, alors $T(\lambda_1^k) = \dots = T(\lambda_n^k) = 0$. Et donc T possède n racines distinctes. Puisque $\deg T \leq n-1$, c'est donc que T est le polynôme nul et donc $\text{Ker } T = \{0\}$.

Ainsi, T est injective. Mais $\dim E = n = \dim \mathbf{R}^n$, et donc T est un isomorphisme.

2.b. Tout élément de \mathbf{R}^n possède un unique antécédent par f . En particulier, il existe un unique $U \in E$ tel que $f(U) = (\lambda_1, \dots, \lambda_n)$, c'est-à-dire tel que

$$U(\lambda_1^k) = \lambda_1, \dots, U(\lambda_n^k) = \lambda_n.$$

3. On a $D = \text{Diag}(\lambda_1, \dots, \lambda_n)$. Et donc

$$U(D^k) = U(\text{Diag}(\lambda_1^k, \dots, \lambda_n^k)) = \text{Diag}(U(\lambda_1^k), \dots, U(\lambda_n^k)) = \text{Diag}(\lambda_1, \dots, \lambda_n) = U.$$

Et donc $R(D) = U(D^k) - D = D - D = 0$.

Ainsi, R est un polynôme annulateur de D .

Et puisque $S = PDP^{-1}$, on a

$$R(S) = R(PDP^{-1}) = PR(D)P^{-1} = 0.$$

Donc R est aussi annulateur de S .

4.a. Montrons le résultat par récurrence sur p . Pour $p = 0$, c'est évident : $A = A$ et pour $p = 1$, c'est l'hypothèse faite sur A .

Supposons que $AS^{pk} = S^{pk}A$. Alors

$$AS^{(p+1)k} = AS^{pk}S^k = (AS^{pk})S^k = S^{pk}AS^k = S^{pk}S^kA = S^{(p+1)k}A.$$

Donc par le principe de récurrence, pour tout $p \in \mathbf{N}$, $AS^{pk} = S^{pk}A$.

4.b. Soit $U = \sum_{p=0}^{n-1} a_p X^p$ le polynôme de la question 2.b.

$$\text{Alors } U(X^k) = \sum_{p=0}^{n-1} a_p X^{kp}. \text{ Et donc } U(S^k) = \sum_{p=0}^{n-1} a_p S^{kp}.$$

On en déduit que

$$U(S^k)A = \sum_{p=0}^{n-1} a_p S^{kp}A = \sum_{p=0}^{n-1} a_p AS^{kp} = AU(S^k).$$

Mais, d'après la question 3, $R(S) = 0 \Leftrightarrow U(S^k) = S$, et donc $SA = AS$.

5.a. λ est valeur propre de S si et seulement si $S - \lambda I_2$ n'est pas inversible, soit si et seulement si $\det(S - \lambda I_2) = 0$. Or

$$\det(S - \lambda I_2) = \det \begin{pmatrix} -\lambda & 1 \\ 1 & -\lambda \end{pmatrix} = \lambda^2 - 1 = (\lambda - 1)(\lambda + 1).$$

Ainsi, S possède deux valeurs propres distinctes : 1 et -1.

Plus précisément

P est alors la matrice de passage de la base canonique de $\mathcal{M}_{n,1}(\mathbf{R})$ à une base de vecteurs propres de S .

Détails

Élever une matrice diagonale à la puissance k c'est élever chacun de ses coefficients diagonaux à la puissance k , et pour sommer deux matrices diagonales, il suffit de sommer leurs coefficients diagonaux.

Donc pour appliquer le polynôme U à D , il suffit de l'appliquer à chacun de ses coefficients diagonaux.

5.b. On a $S^2 = I_2$, et donc si $k = 2p$ est pair, $S^k = (S^2)^p = I_2^p = I_2$.

Donc A commute avec toute puissance paire de S .

On a $AS = \begin{pmatrix} 1 & -1 \\ 2 & -2 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} -1 & 1 \\ 2 & 2 \end{pmatrix}$ et $SA = \begin{pmatrix} 2 & 2 \\ 1 & -1 \end{pmatrix}$.

Donc S et A ne commutent pas.

Rappel

Toute matrice commute avec l'identité.

EXERCICE 2 (ÉCRICOME 2009)

Partie I : Étude d'un cas particulier.

1. Des calculs de produits matriciels nous donnent

$$\Phi_A(V_1) = AV_1 - V_1A = \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix} - \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} = -V_2 + V_3.$$

Et de même,

$$\Phi_A(V_2) = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} = -V_1 + V_4, \quad \Phi_A(V_3) = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} = V_1 - V_4, \quad \Phi_A(V_4) = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} = V_2 - V_3.$$

Et donc

$$\text{Mat}_{(V_1, V_2, V_3, V_4)}(\Phi_A) = \begin{pmatrix} \Phi_A(V_1) & \Phi_A(V_2) & \Phi_A(V_3) & \Phi_A(V_4) \\ 0 & -1 & 1 & 0 \\ -1 & 0 & 0 & 1 \\ 1 & 0 & 0 & -1 \\ 0 & 1 & -1 & 0 \end{pmatrix} \begin{matrix} V_1 \\ V_2 \\ V_3 \\ V_4 \end{matrix}.$$

Remarque

(V_1, V_2, V_3, V_4) est la base canonique de $\mathcal{M}_2(\mathbb{R})$.

2. Il s'agit encore d'un simple calcul :

$$T^2 = \begin{pmatrix} 2 & 0 & 0 & -2 \\ 0 & 2 & -2 & 0 \\ 0 & -2 & 2 & 0 \\ 0 & 1 & -1 & 0 \end{pmatrix}, \quad \text{puis } T^3 = \begin{pmatrix} 0 & -4 & 4 & 0 \\ -4 & 0 & 0 & 4 \\ 4 & 0 & 0 & -4 \\ 0 & 4 & -4 & 0 \end{pmatrix} = 4T.$$

Ainsi, $X^3 - 4X = X(X-2)(X+2)$ est un polynôme annulateur de T .

On en déduit que $\text{Sp}(T) \subset \{-2, 0, 2\}$.

3. On a $\begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix} \in E_0(T)$ si et seulement si

$$T \begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix} = 0 \Leftrightarrow \begin{cases} -y + z = 0 \\ -x + t = 0 \\ x - t = 0 \\ y - z = 0 \end{cases} \Leftrightarrow \begin{cases} x = t \\ y = z \end{cases} \Leftrightarrow \begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix} = \begin{pmatrix} x \\ y \\ y \\ x \end{pmatrix} \in \text{Vect} \left(\begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \end{pmatrix} \right).$$

Ainsi, $\begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \end{pmatrix}$ est une famille génératrice de $E_0(T)$. Puisqu'elle est formée de deux vecteurs

non colinéaires, elle est libre et donc est une base de $E_0(T)$.

4. Encore une fois, il s'agit de faire des calculs :

$$TX_1 = \begin{pmatrix} 0 & -1 & 1 & 0 \\ -1 & 0 & 0 & 1 \\ 1 & 0 & 0 & -1 \\ 0 & 1 & -1 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ -1 \\ 1 \\ -1 \end{pmatrix} = \begin{pmatrix} 2 \\ -2 \\ 2 \\ -2 \end{pmatrix}, \quad \text{et } TX_2 = \begin{pmatrix} -2 \\ -2 \\ 2 \\ 2 \end{pmatrix}.$$

On a donc $TX_1 = 2X_1$ et $TX_2 = -2X_2$.

Puisque X_1 et X_2 sont non nuls, ce sont des vecteurs propres de T , et donc 2 et -2 sont

Danger

À ce stade, nous ne pouvons en dire plus sur les valeurs propres de T , et sûrement pas que 0, 2 et -2 sont tous trois valeurs propres de T .

Nous savons juste que les valeurs propres sont parmi ses trois nombres, peut-être qu'aucun d'entre eux n'est valeur propre, peut-être le sont-ils les trois.

valeurs propres de T .

De plus, à la question précédente, nous avons prouvé que $\dim E_0(T) = 2$.
Et donc on a

$$4 \geq \dim E_0(T) + \dim E_2(T) + \dim E_{-2}(T) \geq 2 + 1 + 1 = 4$$

de sorte que $\sum_{\lambda \in \text{Sp}(T)} \dim E_\lambda(T) = 4$, et donc T est diagonalisable.

Et par conséquent, Φ_A est également diagonalisable.

5. Posons $D = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & -2 \end{pmatrix}$ et $P = \begin{pmatrix} 1 & 0 & 1 & 1 \\ 0 & 1 & -1 & 1 \\ 0 & 1 & 1 & -1 \\ 1 & 0 & -1 & -1 \end{pmatrix}$.

Alors P est la matrice de passage de la base canonique de $\mathcal{M}_{4,1}(\mathbf{R})$ à la base de vecteurs

propres $\begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \\ 1 \\ -1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ -1 \\ -1 \end{pmatrix}$.

Par la formule de changement de base, on a donc $T = PDP^{-1}$.

Partie II : Réduction de Φ_A dans le cas général.

6. Soient $M, N \in \mathcal{M}_n(\mathbf{R})$ et $\lambda \in \mathbf{R}$. Alors

$$\Phi_A(\lambda M + N) = A(\lambda M + N) - (\lambda M + N)A = \lambda(AM - MA) + (AN - NA) = \lambda\Phi_A(M) + \Phi_A(N).$$

Donc Φ_A est linéaire, et elle est bien¹ à valeurs dans $\mathcal{M}_n(\mathbf{R})$: c'est un endomorphisme de $\mathcal{M}_n(\mathbf{R})$.

¹ Un produit de matrices de $\mathcal{M}_n(\mathbf{R})$ est encore dans $\mathcal{M}_n(\mathbf{R})$.

7. Soient $M, N \in \mathcal{M}_n(\mathbf{R})$. Alors

$$\langle M, N \rangle = \text{tr}({}^tMN) = \text{tr}({}^t({}^tMN)) = \text{tr}({}^tNM) = \langle N, M \rangle.$$

Ainsi, $\langle \cdot, \cdot \rangle$ est symétrique.

Soient $M, N, P \in \mathcal{M}_n(\mathbf{R})$ et $\lambda \in \mathbf{R}$. Alors

$$\langle \lambda M + N | P \rangle = \text{tr}({}^t(\lambda M + N)P) = \text{tr}(\lambda {}^tMP + {}^tNP) = \lambda \text{tr}({}^tMN) + \text{tr}({}^tNP) = \lambda \langle M | P \rangle + \langle N | P \rangle.$$

Donc $\langle \cdot, \cdot \rangle$ est linéaire à gauche, et donc, puisqu'elle est symétrique, c'est une forme bilinéaire symétrique.

Soit $M = (m_{i,j}) \in \mathcal{M}_n(\mathbf{R})$. Alors

$$\langle M, M \rangle = \text{tr}({}^tMM) = \sum_{i=1}^n ({}^tMM)_{i,i} = \sum_{i=1}^n \sum_{j=1}^n ({}^tM)_{i,j} M_{j,i} = \sum_{i=1}^n \sum_{j=1}^n m_{i,j}^2.$$

En particulier, $\langle M, M \rangle \geq 0$.

De plus, on a²

$$\langle M, M \rangle = 0 \Leftrightarrow \forall (i, j) \in \llbracket 1, n \rrbracket^2, m_{i,j}^2 = 0 \Leftrightarrow \forall (i, j) \in \llbracket 1, n \rrbracket^2, m_{i,j} = 0 \Leftrightarrow M = 0.$$

Ainsi, $\langle \cdot, \cdot \rangle$ est un produit scalaire sur $\mathcal{M}_n(\mathbf{R})$.

8. Soient M, N deux matrices de $\mathcal{M}_n(\mathbf{R})$. Alors

$$\langle \Phi_A(M) | N \rangle = \text{tr}({}^t(AM - MA)N) = \text{tr}({}^tMAN) - \text{tr}({}^tMNA).$$

D'autre part, on a

$$\langle M | \Phi_A(N) \rangle = \text{tr}({}^tM(AN - NA)) = \text{tr}({}^tMAN) - \text{tr}({}^tMNA).$$

Mais on a

$$\text{tr}({}^tMNA) = \text{tr}({}^t(MN)A) = \text{tr}({}^tMN).$$

Et donc

$$\langle M | \Phi_A(N) \rangle = \langle \Phi_A(M) | N \rangle.$$

Remarque

Ici le concepteur du sujet a été plutôt sympathique en donnant l'expression de $\langle M, N \rangle$ en fonction des $m_{i,j}$ et des $n_{i,j}$. Ce n'est pas toujours le cas et il faut savoir refaire ce calcul.

² Une somme de nombres positifs est nulle si et seulement si chacun de ces nombres est nul.

Trace d'un produit

Nous savons que

$$\text{tr}(AB) = \text{tr}(BA)$$

et comme nous venons de le faire pour trois matrices, on peut prouver que la trace d'un produit de matrices ne dépend pas de l'ordre dans lequel on fait ce produit.

9.a. Notons $X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ et $Y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$. Alors

$$M_{X,Y} = X^t Y = \begin{pmatrix} x_1 \\ \dots \\ x_n \end{pmatrix} \begin{pmatrix} y_1 & \dots & y_n \end{pmatrix} = \begin{pmatrix} x_1 y_1 & x_1 y_2 & \dots & x_1 y_n \\ x_2 y_1 & \dots & \dots & x_2 y_n \\ \vdots & & & \vdots \\ x_n y_1 & \dots & \dots & x_n y_n \end{pmatrix}.$$

Puisque X et Y sont des vecteurs propres, ils sont non nuls. Donc il existe i tel que $x_i \neq 0$ et j tel que $y_j \neq 0$.

Et alors, le coefficient (i, j) de $X^t Y$ est $x_i y_j \neq 0$, de sorte que $M_{X,Y} \neq 0$.

Puisque $AY = \mu Y$, en transposant cette relation, il vient ${}^t Y^t A = \mu^t Y \Leftrightarrow {}^t Y A = \mu^t Y$.

9.b. On a

$$\Phi_A(M_{X,Y}) = AM_{X,Y} - M_{X,Y}A = AX^t Y - X^t YA = \lambda X^t Y - X\mu^t Y = (\lambda - \mu)X^t Y = (\lambda - \mu)M_{X,Y}.$$

Ainsi, si $\alpha \in \Gamma$, il existe deux valeurs propres λ et μ de A telles que $\alpha = \lambda - \mu$.

Et alors si X et Y sont deux vecteurs propres respectivement associés à λ et μ , $M_{X,Y}$ est un vecteur propre de Φ_A , associé à la valeur propre $\alpha = \lambda - \mu$.

En particulier, $\alpha \in \text{Sp}(\Phi_A)$ et donc $\Gamma \subset \text{Sp}(\Phi_A)$.

10.a. Puisque A est diagonalisable, il existe une base de $\mathcal{M}_{n,1}(\mathbf{R})$ formée de vecteurs propres de A . Notons (Z_1, \dots, Z_n) une telle base.

Alors, pour tout $Z \in \mathcal{M}_{n,1}(\mathbf{R})$, il existe $\lambda_1, \dots, \lambda_n$ tels que $Z = \lambda_1 Z_1 + \dots + \lambda_n Z_n$.

$$\text{Et alors } MZ = \sum_{i=1}^n \lambda_i MZ_i = 0.$$

Ainsi, $MZ = 0$, pour tout $Z \in \mathcal{M}_{n,1}(\mathbf{R})$: M est la matrice nulle.

Ceci contredit³ le fait que M soit un vecteur propre de Φ_A , et donc, il existe au moins un vecteur propre Z_0 de A pour lequel $MZ_0 \neq 0$.

10.b. On a $\Phi_A(M) = \alpha M$, soit $AM - MA = \alpha M$.

$$\text{Et alors } AMZ_0 - MAZ_0 = \alpha MZ_0 \Leftrightarrow AMZ_0 - M\mu Z_0 = \alpha MZ_0 \Leftrightarrow A(MZ_0) = (\alpha + \mu)MZ_0.$$

Autrement dit, MZ_0 , qui est non nul, est un vecteur propre de A pour la valeur propre $\alpha + \mu$.

10.c. D'après ce qui précède, il existe une valeur propre λ de A telle que $\alpha + \mu = \lambda \Leftrightarrow \alpha = \lambda - \mu$. Et donc toute valeur propre α de Φ_A est de la forme $\lambda - \mu$, où λ et μ sont deux éléments de $\text{Sp}(A)$, donc $\alpha \in \Gamma$.

On a donc prouvé que $\text{Sp}(\Phi_A) \subset \Gamma$ et l'inclusion réciproque ayant été prouvée à la question 9.b, on a $\text{Sp}(\Phi_A) = \Gamma$.

Explication(s)

Si $MZ = 0$ pour tout $Z \in \mathcal{M}_{n,1}(\mathbf{R})$, alors l'endomorphisme de $\mathcal{M}_{n,1}(\mathbf{R})$ qui à Z associe MZ est l'endomorphisme nul. Et donc sa matrice dans la base canonique de $\mathcal{M}_{n,1}(\mathbf{R})$, qui est M , est la matrice nulle.

Il existe un moyen plus simple de se convaincre de ce résultat : si la $i^{\text{ème}}$ ligne de M , $(x_1 \dots x_n)$ est non nulle, alors le $i^{\text{ème}}$ coefficient de

$$M \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

est $x_1^2 + \dots + x_n^2 \neq 0$.

³ Un vecteur propre est non nul par définition.

DEVOIR MAISON 9 (HEC) (À RENDRE LE 20.12.18)

Dans tout le problème, pour tout couple $(p, q) \in (\mathbf{N}^*)^2$:

- on note $\mathcal{M}_{p,q}(\mathbf{R})$ l'espace vectoriel des matrices à p lignes et q colonnes à coefficients réels et $\mathcal{M}_p(\mathbf{R})$ est noté $\mathcal{M}_{p,p}(\mathbf{R})$;
- la matrice transposée d'une matrice A de $\mathcal{M}_{p,q}(\mathbf{R})$ est notée ${}^t A$;
- on note I_p la matrice identité de $\mathcal{M}_p(\mathbf{R})$ et pour toute matrice A , même nulle, de $\mathcal{M}_p(\mathbf{R})$, on pose par convention : $A^0 = I_p$;
- la matrice inverse d'une matrice inversible A de $\mathcal{M}_p(\mathbf{R})$ est notée A^{-1} .

Soit $(M_n)_{n \in \mathbf{N}}$ une suite de matrices de $\mathcal{M}_{p,q}(\mathbf{R})$. On pose pour tout $n \in \mathbf{N}$: $M_n = (m_{i,j}(n))_{\substack{1 \leq i \leq p \\ 1 \leq j \leq q}}$.

On dit que la suite $(M_n)_{n \in \mathbf{N}}$ converge vers la matrice $M = (m_{i,j})_{\substack{1 \leq i \leq p \\ 1 \leq j \leq q}}$ de $\mathcal{M}_{p,q}(\mathbf{R})$, si pour tout couple $(i, j) \in \llbracket 1, p \rrbracket \times \llbracket 1, q \rrbracket$, on a : $\lim_{n \rightarrow +\infty} m_{i,j}(n) = m_{i,j}$. On note alors : $M = \lim_{n \rightarrow +\infty} M_n$.

On admet sans démonstration que si $(A_n)_{n \in \mathbf{N}}$ et $(B_n)_{n \in \mathbf{N}}$ sont deux suites de matrices de $\mathcal{M}_{p,q}(\mathbf{R})$ qui convergent respectivement vers les matrices A et B , et si $(C_n)_{n \in \mathbf{N}}$ est une suite de matrices de $\mathcal{M}_{q,s}(\mathbf{R})$ ($s \geq 1$) qui converge vers $C \in \mathcal{M}_{q,s}(\mathbf{R})$, alors la suite $(A_n + B_n)_{n \in \mathbf{N}}$ converge vers $A+B$, la suite $(A_n C_n)_{n \in \mathbf{N}}$ converge vers AC , et pour tout réel α , la suite $(\alpha A_n)_{n \in \mathbf{N}}$ converge vers αA .

Le problème étudie quelques aspects mathématiques du contrôle de systèmes linéaires.

Partie I. Quelques propriétés de suites matricielles.

Pour toute matrice A de $\mathcal{M}_p(\mathbf{R})$, on pose pour tout x réel et pour tout $n \in \mathbf{N}$: $T_{A,n}(x) = \sum_{k=0}^n \frac{1}{k!} (xA)^k$.

1. *Exemple.* Dans cette question, $A = (a_{i,j})_{1 \leq i,j \leq p}$ est la matrice de $\mathcal{M}_p(\mathbf{R})$ définie par

$$\forall (i, j) \in \llbracket 1, p \rrbracket^2, a_{i,j} = 1.$$

- a. Justifier que la matrice A est diagonalisable.
- b. Soit V la matrice-colonne de $\mathcal{M}_{p,1}(\mathbf{R})$ dont tous les coefficients sont égaux à 1. Calculer le produit AV et en déduire une valeur propre de A .
- c. Montrer que 0 est une valeur propre de A et trouver la dimension du sous-espace propre associé.
- d. Exprimer A^2 en fonction de A .
Montrer que pour tout x réel et pour tout $n \in \mathbf{N}$, $T_{A,n}(x)$ appartient à $\text{Vect}(I_p, A)$.
- e. En déduire que pour tout x réel, la suite de matrice $(T_{A,n}(x))_{n \in \mathbf{N}}$ de $\mathcal{M}_p(\mathbf{R})$ converge vers la matrice $T_A(x)$ de $\mathcal{M}_p(\mathbf{R})$ définie par : $T_A(x) = I_p + \frac{e^{px} - 1}{p} A$.
- f. Calculer $T_A(0)$. Exprimer pour tout couple $(x, y) \in \mathbf{R}^2$, le produit $T_A(x)T_A(y)$ en fonction de $T_A(x+y)$.
En déduire que pour tout x réel, la matrice $T_A(x)$ est inversible et déterminer son inverse.

2. Soit $A = (a_{i,j})_{1 \leq i,j \leq p}$ une matrice de $\mathcal{M}_p(\mathbf{R})$.

On pose pour tout $k \in \mathbf{N}$: $A^k = (a_{i,j}^{(k)})_{1 \leq i,j \leq p}$ et $\mu_k = \max_{(i,j) \in \llbracket 1, p \rrbracket^2} |a_{i,j}^{(k)}|$.

- a. À l'aide de l'identité $A^{k+1} = AA^k$, montrer que pour tout $k \in \mathbf{N}$, on a : $\mu_{k+1} \leq p^k \mu_1^{k+1}$.

- b. En déduire que pour tout x réel, la série $\sum_{k \geq 0} \frac{\mu_k}{k!} x^k$ est convergente.

- c. Montrer que pour tout réel x et pour tout $(i, j) \in \llbracket 1, p \rrbracket^2$, la série $\sum_{k \geq 0} \frac{a_{i,j}^{(k)}}{k!} x^k$ est convergente.

- d. Montrer que pour tout x réel, la suite $(T_{A,n}(x))_{n \in \mathbf{N}}$ de $\mathcal{M}_p(\mathbf{R})$ converge vers une matrice $T_A(x)$ de $\mathcal{M}_p(\mathbf{R})$.
Que vaut $T_A(x)$ lorsque $p = 1$ et que l'unique coefficient de A est un réel a ?

3. Soit D une matrice diagonale de $\mathcal{M}_p(\mathbf{R})$.

- a. Vérifier que pour tout x réel et pour tout $n \in \mathbf{N}$, la matrice $T_{D,n}(x)$ est diagonale.
- b. En déduire que pour tout x réel, la matrice $T_D(x)$ est diagonale et donner l'expression de ses coefficients diagonaux en fonction de ceux de D .

- c. On pose pour tout $r \in \mathbf{N}^*$: $D_r = r \left(T_D \left(\frac{1}{r} \right) - I_p \right)$. Montrer que la suite $(D_r)_{r \in \mathbf{N}^*}$ converge vers D .
4. Soit A une matrice de $\mathcal{M}_p(\mathbf{R})$, P une matrice inversible de $\mathcal{M}_p(\mathbf{R})$ et $A' = P^{-1}AP$.
- a. Établir pour tout x réel, l'égalité : $T_{A'}(x) = P^{-1}T_A(x)P$.
- b. On suppose que A est diagonalisable. Montrer que :

$$\forall n \in \mathbf{N}, \lim_{r \rightarrow +\infty} r^{n+1} \left(T_A \left(\frac{1}{r} \right) - T_{A,n} \left(\frac{1}{r} \right) \right) = \frac{1}{(n+1)!} A^{n+1}. \quad (\star)$$

(On pourra traiter dans un premier temps le cas où A est diagonale).

On admet dans la suite du problème que la relation (\star) reste valable pour toute matrice A de $\mathcal{M}_p(\mathbf{R})$.

Partie II. Polynômes annulateurs et matrices de Kalman.

Soit E un espace vectoriel de dimension p sur \mathbf{C} , $\mathcal{L}(E)$ l'espace vectoriel des endomorphismes de E et φ un élément de $\mathcal{L}(E)$. On note id_E l'endomorphisme identité de E .

5. a. Rappeler la dimension de $\mathcal{L}(E)$ et justifier l'existence d'une suite finie $(z_k)_{1 \leq k \leq p^2}$ de nombres complexes tels que le polynôme $\prod_{k=1}^{p^2} (X - z_k)$ de $\mathbf{C}[X]$ soit un polynôme annulateur de φ .
- b. En considérant, pour tout $k \in \llbracket 1, p \rrbracket$, les endomorphismes $(\varphi - z_k \text{id}_E)$, montrer que φ possède au moins une valeur propre.
6. On suppose l'existence d'un entier k vérifiant $1 \leq k < p$ et d'un sous-espace vectoriel F de dimension k stable par φ . Soit H un supplémentaire de F dans E et π le projecteur de E sur H parallèlement à F .
- a. Montrer qu'il existe un vecteur non nul $v \in H$ et un nombre complexe λ vérifiant la relation : $\pi \circ \varphi(v) = \lambda v$.
- b. Montrer que la somme des deux sous-espaces vectoriels F et $\text{Vect}(v)$ est directe et stable par φ .
7. À l'aide des questions précédentes, établir par récurrence sur p l'existence d'une base $\mathcal{B} = (v_1, \dots, v_p)$ de E telle que pour tout $k \in \llbracket 1, p \rrbracket$, $\varphi(v_k) \in \text{Vect}(v_1, \dots, v_k)$.
8. Soit $M = (m_{i,j})_{1 \leq i, j \leq p}$ la matrice de φ dans la base \mathcal{B} . On pose, pour tout $k \in \llbracket 1, p \rrbracket$: $F_k = \text{Vect}(v_1, \dots, v_k)$.
- a. Montrer que pour tout $k \in \llbracket 2, p \rrbracket$, on a : $(\varphi - m_{k,k} \text{id}_E)(F_k) \subset F_{k-1}$.
- b. En déduire que le polynôme $\prod_{k=1}^p (X - m_{k,k})$ de $\mathbf{C}[X]$ est un polynôme annulateur de la matrice M .
9. Soit A une matrice de $\mathcal{M}_p(\mathbf{R})$. En utilisant la question 8.b, montrer que A admet un polynôme annulateur appartenant à $\mathbf{R}[X]$ et de degré p .
10. Soit A une matrice de $\mathcal{M}_p(\mathbf{R})$ et B une matrice-colonne de $\mathcal{M}_{p,1}(\mathbf{R})$. Pour tout $q \in \mathbf{N}^*$, on note \mathcal{G}_q le sous-espace vectoriel de $\mathcal{M}_{p,1}(\mathbf{R})$ engendré par $B, AB, A^2B, \dots, A^{q-1}B$ et K_q la matrice de $\mathcal{M}_{p,q}(\mathbf{R})$ dont les colonnes successives sont $B, AB, A^2B, \dots, A^{q-1}B$. La matrice K_q est appelée matrice de Kalman d'ordre q associée au couple (A, B) .
- a. Montrer que pour tout entier $q > p$, on a : $\mathcal{G}_q = \mathcal{G}_p$.
- b. Justifier l'existence d'un sous-espace vectoriel \mathcal{S} de $\mathcal{M}_{p,1}(\mathbf{R})$ vérifiant la propriété suivante : pour qu'une matrice G de $\mathcal{M}_{p,1}(\mathbf{R})$ appartienne à \mathcal{G}_p , il faut et il suffit que pour tout élément S de \mathcal{S} , on ait : ${}^tSG = 0$.
- c. En déduire que si une suite $(G_n)_{n \in \mathbf{N}}$ de matrices de \mathcal{G}_p est convergente, sa limite G appartient à \mathcal{G}_p .
- d. À l'aide des résultats précédents, montrer que pour tout x réel, la matrice-colonne $T_A(x)B$ appartient à \mathcal{G}_p , où $T_A(x)$ a été défini dans la question 2.d.

Partie III. Contrôle de système linéaires.

On conserve dans cette partie les définitions et notations de la question 10. Dans les questions 12, 13 et 14, on note p un entier supérieur ou égal à 2. Les questions 13 et 14 sont indépendantes des questions 11 et 12.

On note \mathcal{C}^0 l'espace vectoriel des fonctions continues sur $[0, 1]$ à valeurs dans \mathbf{R} .

11. Exemple : $p = 1$. Soit (a, b) un couple de réels.

- a. Soit $u \in \mathcal{C}^0$. On cherche une fonction f définie et dérivable sur $[0, 1]$, de dérivée f' , vérifiant $f(0) = 0$ et telle que pour tout $t \in [0, 1]$, $f'(t) = af(t) + bu(t)$.
Calculer la dérivée de la fonction $h : t \mapsto h(t) = f(t)e^{-at}$, et en déduire que f est donnée par :

$$\forall t \in [0, 1], f(t) = b \int_0^t u(x)e^{a(t-x)} dx. \quad (\star\star)$$

- b. On dit que le couple (a, b) est *contrôlable*, si pour tout réel y (appelé *cible*), il existe une fonction $u \in \mathcal{C}^0$ (appelée *contrôle*) telle que toute fonction f définie et dérivable sur $[0, 1]$ vérifiant $f(0) = 0$ et $f'(t) = af(t) + bu(t)$ pour tout $t \in [0, 1]$ atteint la cible en 1, c'est-à-dire vérifie $f(1) = y$.
Donner l'expression de la fonction f définie par $(\star\star)$ lorsque la fonction u est constante sur $[0, 1]$.
En déduire que le couple (a, b) est contrôlable si et seulement si $b \neq 0$.

12. Pour tout $x \in [0, 1]$, on pose : $W(x) = T_A(1-x)B \in \mathcal{M}_{p,1}(\mathbf{R})$ et $W(x) = (W_k(x))_{1 \leq k \leq p}$, où pour tout $k \in \llbracket 1, p \rrbracket$, $W_k(x)$ est le coefficient de la k -ième ligne de $W(x)$.
On admet que pour tout $k \in \llbracket 1, p \rrbracket$, la fonction $x \mapsto W_k(x)$ appartient à \mathcal{C}^0 et on définit alors, pour toute fonction $u \in \mathcal{C}^0$, la matrice-colonne $\int_0^1 u(x)W(x) dx$ de $\mathcal{M}_{p,1}(\mathbf{R})$ par :

$$\int_0^1 u(x)W(x) dx = \left(\int_0^1 u(x)W_k(x) dx \right)_{1 \leq k \leq p}.$$

Par analogie avec la question 11.b, on dit que le couple (A, B) est *contrôlable*, si pour tout matrice-colonne $Y \in \mathcal{M}_{p,1}(\mathbf{R})$ (*cible*), il existe une fonction $u \in \mathcal{C}^0$ (*contrôle*) vérifiant l'égalité : $\int_0^1 u(x)W(x) dx = Y$.

- a. Soit $u \in \mathcal{C}^0$. Justifier que pour tout $x \in [0, 1]$, $u(x)W(x)$ appartient à \mathcal{G}_p .
En déduire que $\int_0^1 u(x)W(x) dx$ appartient à \mathcal{G}_p .
- b. Soit Z un élément non nul de $\mathcal{M}_{p,1}(\mathbf{R})$ tel que pour toute fonction $u \in \mathcal{C}^0$, on ait : $\int_0^1 u(x)^t ZW(x) dx = 0$.
Montrer que pour tout $x \in [0, 1]$, on a : ${}^t ZW(x) = 0$.
- c. En déduire, à l'aide des relations (\star) (question 4.b), que pour tout $k \in \llbracket 1, p \rrbracket$, on a : ${}^t ZA^{k-1}B = 0$.
- d. Déduire des résultats précédents que le couple (A, B) est contrôlable, si et seulement si la matrice de Kalman K_p est inversible.

CORRECTION DU DEVOIR MAISON 9

Partie I. Quelques propriétés de suites matricielles.

1.a. La matrice A est symétrique, à coefficients réels, donc est diagonalisable.

1.b. On a

$$AV = \begin{pmatrix} 1 & \dots & 1 \\ \vdots & & \vdots \\ 1 & \dots & 1 \end{pmatrix} \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} = \begin{pmatrix} p \\ \vdots \\ p \end{pmatrix} = pV.$$

Donc¹ p est valeur propre de A .

¹ V est non nul, donc est un vecteur propre.

1.c. Toutes les colonnes de A sont égales et non nulles, donc A est de rang 1. En particulier, 0 est valeur propre de A et $\dim E_0(A) = p - \text{rg}(A) = p - 1$.

1.d. On a

$$A^2 = \begin{pmatrix} 1 & \dots & 1 \\ \vdots & & \vdots \\ 1 & \dots & 1 \end{pmatrix} \begin{pmatrix} 1 & \dots & 1 \\ \vdots & & \vdots \\ 1 & \dots & 1 \end{pmatrix} = \begin{pmatrix} p & \dots & p \\ \vdots & & \vdots \\ p & \dots & p \end{pmatrix} = pA.$$

Il vient alors $A^3 = AA^2 = ApA = pA^2 = p^2A$, puis $A^4 = p^3A$, etc.

Une récurrence rapide prouve alors que pour tout $k \in \mathbf{N}$, $A^k = p^{k-1}A$.

Et donc pour $n \in \mathbf{N}$ et $x \in \mathbf{R}$,

$$T_{A,n}(x) = \sum_{k=0}^n \frac{1}{k!} x^k A^k = \sum_{k=0}^n \frac{1}{k!} x^k p^{k-1} A = I_p + \left(\sum_{k=1}^n \frac{1}{k!} x^k p^{k-1} \right) A \in \text{Vect}(I_p, A).$$

1.e. Lorsque $n \rightarrow +\infty$, on a $\sum_{k=1}^n \frac{x^k p^{k-1}}{k!} = \frac{1}{p} \sum_{k=0}^n \frac{p^k x^k}{k!} - \frac{1}{p} \xrightarrow{n \rightarrow +\infty} \frac{e^{px}}{p} - \frac{1}{p} = \frac{e^{px} - 1}{p}$.

On en déduit que $\sum_{k=1}^n \frac{1}{k!} p^{k-1} x^k A \xrightarrow{n \rightarrow +\infty} \frac{e^{px} - 1}{p} A$, puis $T_{A,n}(x) \xrightarrow{n \rightarrow +\infty} I_p + \frac{e^{px} - 1}{p} A$.

1.f. En appliquant la formule obtenue à la question précédente avec $x = 0$, il vient

$$T_A(0) = I_p + \frac{e^0 - 1}{p} A = I_p.$$

Pour $(x, y) \in \mathbf{R}^2$, on a

$$\begin{aligned} T_A(x)T_A(y) &= \left(I_p + \frac{e^{px} - 1}{p} A \right) \left(I_p + \frac{e^{py} - 1}{p} A \right) \\ &= I_p + \frac{e^{px} - 1}{p} A + \frac{e^{py} - 1}{p} A + \frac{(e^{px} - 1)(e^{py} - 1)}{p^2} A^2 \\ &= I_p + \frac{e^{px} - 1}{p} A + \frac{e^{py} - 1}{p} A + \frac{e^{p(x+y)} - e^{px} - e^{py} + 1}{p^2} pA \\ &= I_p + \frac{e^{p(x+y)} - 1}{p} A = T_A(x+y). \end{aligned}$$

On en déduit que $T_A(x)T_A(-x) = T_A(x-x) = T_A(0) = I_p$, et donc $T_A(x)$ est inversible, d'inverse égale à $T_A(-x)$.

2.a. Montrons par récurrence sur $k \in \mathbf{N}$ que $\mu_{k+1} \leq p^k \mu_1^{k+1}$.

Pour $k = 0$, on a $\mu_1 = p^0 \mu_1^1$, donc la formule est évidemment vérifiée.

Supposons donc que $\mu_{k+1} \leq p^k \mu_1^{k+1}$. Alors, pour $(i, j) \in \llbracket 1, p \rrbracket^2$, on a

$$\begin{aligned} |a_{i,j}^{(k+2)}| &= \left| \sum_{\ell=1}^p a_{i,\ell} a_{\ell,j}^{(k+1)} \right| \\ &\leq \sum_{\ell=1}^p |a_{i,\ell}| |a_{\ell,j}^{(k+1)}| \end{aligned}$$

C'est la formule du produit matriciel appliquée à $A^{k+2} = AA^{k+1}$.

Inégalité triangulaire.

$$\begin{aligned} &\leq \sum_{\ell=1}^p \mu_1 \mu_{k+1} \\ &\leq p \mu_1 \mu_{k+1} \\ &\leq \mu_1 p p^k \mu_1^{k+1} \\ &\leq p^{k+1} \mu_1^{k+2}. \end{aligned}$$

Hypothèse de récurrence.

Cette inégalité étant vraie pour tout (i, j) , c'est donc que $\mu_{k+2} \leq p^{k+1} \mu_1^{k+2}$.

D'après le principe de récurrence, on a alors, pour tout $k \in \mathbf{N}$, $\mu_{k+1} \leq p^k \mu_1^{k+1}$.

2.b. On a alors, pour tout $k \in \mathbf{N}$, $0 \leq \left| \frac{\mu_k x^k}{k!} \right| \leq \frac{p^k \mu_1^{k+1} |x|^k}{k!} = \mu_1 \frac{(p \mu_1 |x|)^k}{k!}$.

On reconnaît alors le terme général d'une série exponentielle convergente, et donc par critère de domination pour les séries à termes positifs, $\sum_{k \geq 0} \frac{\mu_k x^k}{k!}$ est absolument convergente,

et donc convergente.

2.c. Par définition de μ_k , pour tout $(i, j) \in \llbracket 1, p \rrbracket$, $|a_{i,j}^{(k)}| \leq \mu_k$, et donc $0 \leq \frac{|a_{i,j}^{(k)} x^k|}{k!} \leq \frac{\mu_k |x|^k}{k!}$.

On reconnaît alors le terme général de la série étudiée à la question précédente, qui est convergente.

Et donc $\sum_{k \geq 0} \frac{a_{i,j}^{(k)} x^k}{k!}$ est absolument convergente, donc convergente.

2.d. Nous venons de prouver que pour tout $(i, j) \in \llbracket 1, p \rrbracket^2$, la suite de terme général $\sum_{k=0}^n \frac{a_{i,j}^{(k)} x^k}{k!}$

est convergente.

Mais il s'agit du coefficient (i, j) de la matrice $T_{A,n}(x)$.

Et donc, par définition de la convergence d'une suite de matrices, cela signifie que la suite $(T_{A,n}(x))_{n \in \mathbf{N}}$ converge.

Comme indiqué dans l'énoncé, on notera donc $T_A(x)$ sa limite.

Lorsque $p = 1$, l'unique coefficient de $T_{A,n}(x)$ est $\sum_{k=0}^n \frac{1}{k!} x^k a^k \xrightarrow{n \rightarrow +\infty} e^{ax}$.

Et donc $T_A(x)$ est la matrice 1×1 dont l'unique coefficient est e^{ax} .

3.a. Nous savons qu'un produit de deux matrices diagonales est une matrice diagonale et donc $\forall k \in \mathbf{N}$, D^k est diagonale.

De plus, les matrices diagonales forment un sous-espace vectoriel de $\mathcal{M}_p(\mathbf{R})$ et donc

$$\forall n \in \mathbf{N}, T_{D,n}(x) = \sum_{k=0}^n \frac{x^k}{k!} D^k \text{ est diagonale.}$$

3.b. Puisque $T_{D,n}(x)$ est diagonale, pour $i \neq j$, le coefficient (i, j) de $T_{D,n}(x)$ est nul.

Et donc en passant à la limite, le coefficient (i, j) de $T_D(x)$ est également nul, de sorte que

$T_D(x)$ est diagonale.

De plus, si on note $d_{i,j}$ le coefficient (i, j) de D , on a alors

$$T_{D,n}(x) = \text{Diag} \left(\sum_{k=0}^n \frac{x^k d_{1,1}^k}{k!}, \dots, \sum_{k=0}^n \frac{x^k d_{p,p}^k}{k!} \right) \xrightarrow{n \rightarrow +\infty} \text{Diag}(e^{d_{1,1}x}, \dots, e^{d_{p,p}x}).$$

Et donc le i -ème coefficient diagonal de $T_D(x)$ est $e^{d_{i,i}x}$.

3.c. D'après ce qui précède, $T_D \left(\frac{1}{r} \right)$ est une matrice diagonale, et si on note $D = \text{Diag}(\lambda_1, \dots, \lambda_p)$,

alors $T_D \left(\frac{1}{r} \right) = \text{Diag} \left(e^{\frac{\lambda_1}{r}}, \dots, e^{\frac{\lambda_p}{r}} \right)$.

Et donc

$$D_r = \text{Diag} \left(r \left(e^{\frac{\lambda_1}{r}} - 1 \right), \dots, r \left(e^{\frac{\lambda_p}{r}} - 1 \right) \right).$$

Pour la culture
Lorsque $x = 1$, on appelle $T_A(1)$ l'exponentielle de la matrice A .
Le fait que pour $p = 1$ on retrouve l'exponentielle légitime cette terminologie.

De plus, pour $i \in \llbracket 1, p \rrbracket$, on a

$$e^{\frac{\lambda_i}{r}} - 1 \underset{r \rightarrow +\infty}{\sim} \frac{\lambda_i}{r}$$

et donc $r \left(e^{\frac{\lambda_i}{r}} - 1 \right) \underset{r \rightarrow +\infty}{\sim} \lambda_i \xrightarrow{r \rightarrow +\infty} \lambda_i$.

Ceci étant vrai pour tout i , on en déduit que $D_r \xrightarrow{r \rightarrow +\infty} \text{Diag}(\lambda_1, \dots, \lambda_p) = D$.

4.a. Nous savons que pour tout polynôme $Q \in \mathbf{R}[X]$, on a $Q(A') = P^{-1}Q(A)P$.

En particulier, pour x fixé, si on note $Q_{x,n}(X) = \sum_{k=0}^n \frac{1}{k!} x^k X^k \in \mathbf{R}[X]$, alors

$$T_{A',n}(x) = Q_{x,n}(A') = P^{-1}Q_{x,n}(A)P = P^{-1}T_{A,n}(x)P.$$

Et donc lorsque $n \rightarrow +\infty$, il vient

$$T_{A'}(x) = \lim_{n \rightarrow +\infty} T_{A,n}(x) = \lim_{n \rightarrow +\infty} P^{-1}T_{A,n}(x)P = P^{-1} \left(\lim_{n \rightarrow +\infty} T_{A,n}(x) \right) P = \boxed{P^{-1}T_A(x)P}.$$

4.b. Commençons dans un premier temps par supposer que A est une matrice diagonale : $A = \text{Diag}(\lambda_1, \dots, \lambda_p)$.

Alors $T_A\left(\frac{1}{r}\right) = \text{Diag}\left(e^{\frac{\lambda_1}{r}}, \dots, e^{\frac{\lambda_p}{r}}\right)$ de sorte que pour $r \in \mathbf{N}^*$, $r^{n+1} \left(T_A\left(\frac{1}{r}\right) - T_{A,n}\left(\frac{1}{r}\right) \right)$ est une matrice diagonale, donc le i -ème coefficient diagonal est

$$\begin{aligned} r^{n+1} \left(e^{\frac{\lambda_i}{r}} - \sum_{k=0}^n \left(\frac{\lambda_i}{r} \right)^k \frac{1}{k!} \right) &= r^{n+1} \left(\sum_{k=0}^{+\infty} \left(\frac{\lambda_i}{r} \right)^k \frac{1}{k!} - \sum_{k=0}^n \left(\frac{\lambda_i}{r} \right)^k \frac{1}{k!} \right) \\ &= r^{n+1} \sum_{k=n+1}^{+\infty} \left(\frac{\lambda_i}{r} \right)^k \frac{1}{k!} \\ &= \frac{\lambda_i^{n+1}}{(n+1)!} + r^{n+1} \sum_{k=n+2}^{+\infty} \frac{\lambda_i^k}{r^k} \frac{1}{k!}. \end{aligned}$$

Montrons que $r^{n+1} \sum_{k=n+2}^{+\infty} \frac{\lambda_i^k}{r^k} \frac{1}{k!} \xrightarrow{r \rightarrow +\infty} 0$.

Pour $r \geq 1$, on a

$$\begin{aligned} r^{n+1} \left| \sum_{k=n+2}^{+\infty} \frac{\lambda_i^k}{r^k} \frac{1}{k!} \right| &\leq r^{n+1} \sum_{k=n+2}^{+\infty} \frac{|\lambda_i|^k}{r^k} \frac{1}{k!} \\ &\leq r^{n+1} \sum_{k=n+2}^{+\infty} \frac{|\lambda_i|^k}{r^{n+2}} \frac{1}{k!} \\ &\leq \frac{1}{r} \sum_{k=n+2}^{+\infty} \frac{|\lambda_i|^k}{k!} \\ &\leq \frac{1}{r} \sum_{k=0}^{+\infty} \frac{|\lambda_i|^k}{k!} \\ &\leq \frac{e^{|\lambda_i|}}{r} \xrightarrow{r \rightarrow +\infty} 0. \end{aligned}$$

Inégalité triangulaire.

Pour $r \geq 1$ et $k \geq n+2$,
 $r^k \geq r^{n+2}$.

Ainsi, le i -ème coefficient diagonal de $r^{n+1} \left(T_A\left(\frac{1}{r}\right) - T_{A,n}\left(\frac{1}{r}\right) \right)$ tend vers $\frac{\lambda_i^{n+1}}{(n+1)!}$ lorsque $r \rightarrow +\infty$ et donc

$$\boxed{\lim_{r \rightarrow +\infty} r^{n+1} \left(T_A\left(\frac{1}{r}\right) - T_{A,n}\left(\frac{1}{r}\right) \right) = \frac{1}{(n+1)!} \text{Diag}(\lambda_1^{n+1}, \dots, \lambda_p^{n+1}) = \frac{A^{n+1}}{(n+1)!}}.$$

Dans le cas général, si A est diagonalisable, alors il existe P inversible et D diagonale telles que $A = P^{-1}DP$ et alors, d'après la question précédente,

$$r^{n+1} \left(T_A\left(\frac{1}{r}\right) - T_{A,n}\left(\frac{1}{r}\right) \right) = P^{-1} r^{n+1} \left(T_D\left(\frac{1}{r}\right) - T_{D,n}\left(\frac{1}{r}\right) \right) P \xrightarrow{r \rightarrow +\infty} P^{-1} \frac{D^{n+1}}{(n+1)!} P = \boxed{\frac{A^{n+1}}{(n+1)!}}.$$

Partie II. Polynômes annulateurs et matrices de Kalman.

5.a. Nous savons que $\mathcal{L}(E)$ est de dimension p^2 .

Puisque tout polynôme à coefficients complexes est produit de facteurs de degré 1, la question peut se reformuler sous la forme suivante : «montrer que φ admet un polynôme annulateur unitaire de degré p^2 ».

La famille d'endomorphismes $(\text{id}_E, \varphi, \varphi^2, \dots, \varphi^{p^2})$, qui est de cardinal $p^2 + 1$ est nécessairement liée.

Il existe donc des complexes non tous nuls² a_0, \dots, a_{p^2} tels que $\sum_{k=0}^{p^2} a_k \varphi^k = 0$.

Par conséquent le polynôme $P = \sum_{k=0}^{p^2} a_k X^k$ est donc annulateur de φ , et n'est pas le polynôme nul.

Si $a_{p^2} \neq 0$, alors P est de degré p^2 , et donc $\frac{1}{a_{p^2}}P$ est unitaire, de degré p^2 , et est encore annulateur de φ , donc satisfait aux conditions de l'énoncé.

Si $a_{p^2} = 0$, alors $d = \deg P < p^2$.

Mais alors $Q = \frac{1}{a_d} X^{p^2-d} P$ est encore³ annulateur de φ , unitaire et de degré p^2 , et donc satisfait aux conditions de l'énoncé.

Dans les deux cas, si Q est un polynôme annulateur de φ , unitaire et de degré p^2 , alors en notant z_1, \dots, z_{p^2} les racines complexes, comptées avec multiplicité de Q , alors $Q = \prod_{k=1}^{p^2} (X - z_k)$ est annulateur de φ .

5.b. Rappelons que $\lambda \in \mathbb{C}$ est valeur propre de φ si et seulement si $\varphi - \lambda \text{id}_E$ est bijectif. Or, si Q est le polynôme obtenu à la question précédente, on a

$$Q(\varphi) = 0 \iff (\varphi - z_1 \text{id}_E) \circ (\varphi - z_2 \text{id}_E) \circ \dots \circ (\varphi - z_{p^2} \text{id}_E) = 0.$$

Si aucun des z_i n'était valeur propre de φ , alors $Q(\varphi)$ serait bijectif car composée d'endomorphismes bijectifs.

Mais l'endomorphisme nul n'est pas bijectif, et donc l'un⁴ des z_i est une valeur propre de φ .

6.a. Soit g l'endomorphisme de H défini pour tout $x \in H$ par $g(x) = \pi(\varphi(x))$. Alors, d'après ce qui précède⁵, g possède une valeur propre λ , et donc il existe un vecteur propre $v \in H$ tel que $g(v) = \lambda v \iff (\pi \circ \varphi)(v) = \lambda v$.

6.b. La somme est directe car $F \cap \text{Vect}(v) \subset F \cap H = \{0_E\}$, et donc $F \cap \text{Vect}(v) = \{0_E\}$. Il s'agit donc de prouver que $F \oplus \text{Vect}(v)$ est stable par φ .

Soit donc $x \in F \oplus \text{Vect}(v)$. Alors $x = u + \alpha v$, avec $u \in F$ et $\alpha \in \mathbb{C}$.

Et alors $\varphi(x) = \varphi(u) + \alpha \varphi(v)$.

Mais $\varphi(v)$ s'écrit de manière unique $x_1 + x_2$, avec $x_1 \in F$ et $x_2 \in H$, et d'après la question précédente, on a $x_2 = \pi(\varphi(v)) = \lambda v$.

Donc $\varphi(x) = \varphi(u) + \alpha(x_1 + \lambda v) = (\varphi(u) + \alpha x_1) + \alpha \lambda v$.

Il est clair que $\alpha \lambda v \in \text{Vect}(v)$, et puisque F est stable par φ par hypothèse, alors $\varphi(u) \in F$ et donc $\varphi(u) + \alpha x_1 \in F$.

Et donc on a bien $\varphi(v) \in F \oplus \text{Vect}(v)$: la somme est stable par φ .

7. Montrons par récurrence sur $n \leq p$ qu'il existe une famille libre (v_1, \dots, v_n) telle que $\forall k \in \llbracket 1, n \rrbracket, \varphi(v_k) \in \text{Vect}(v_1, \dots, v_k)$.

Puisque φ possède une valeur propre λ , il existe donc un vecteur propre v_1 associé à la valeur propre λ , pour lequel on a $f(v_1) = \lambda v_1 \in \text{Vect}(v_1)$. De plus, la famille v_1 étant formée d'un seul vecteur non nul, elle est libre.

Supposons donc la propriété vraie au rang n , et supposons que $n + 1 \leq p$.

Alors $F = \text{Vect}(v_1, \dots, v_n)$ est un sous-espace de E stable par φ car si $x = \sum_{j=1}^n \alpha_j v_j \in F$,

$$f(x) = \sum_{j=1}^n \alpha_j \underbrace{f(v_j)}_{\in \text{Vect}(v_1, \dots, v_j) \subset F} \in F.$$

Pol. unitaire

Notons bien que le polynôme donné par l'énoncé

$$\prod_{k=1}^{p^2} (X - z_k)$$

n'est pas n'importe quel polynôme de $\mathbb{C}_{p^2}[X]$, mais bien un polynôme de degré égal à p^2 , et dont le coefficient dominant vaut 1.

² À ne pas confondre avec tous non nuls : non tous nuls signifie que l'un **au moins** des a_k est non nul.

³ Un multiple d'un polynôme annulateur est annulateur car

$$(PQ)(\varphi) = P(\varphi) \circ Q(\varphi).$$

⁴ Au moins.

⁵ Le raisonnement tenu pour φ est encore valable pour g .

Récurrence finie

Notons qu'il s'agit là d'une récurrence finie, c'est-à-dire qu'elle ne porte que sur les $n \leq p$, et donc s'arrête à $n = p$. Elle ne prouve donc pas un résultat pour tout $n \in \mathbb{N}^*$, mais uniquement pour tout $n \in \llbracket 1, p \rrbracket$.

D'après la question 6, il existe donc un vecteur v_{n+1} tel que $F \oplus \text{Vect}(v_{n+1})$ soit stable par φ . Mais alors (v_1, \dots, v_{n+1}) est une base de $F \oplus \text{Vect}(v_{n+1})$ car obtenue par concaténation d'une base de F et d'une base de $\text{Vect}(v_{n+1})$, et on a

$$f(v_{n+1}) \in F \oplus \text{Vect}(v_{n+1}) = \text{Vect}(v_1, \dots, v_{n+1}).$$

Ainsi, l'hypothèse de récurrence est encore vraie au rang $n + 1$, et d'après le principe de récurrence, elle est vraie en particulier pour $n = p$: il existe une famille libre (v_1, \dots, v_p) de E telle que $\forall k \in \llbracket 1, p \rrbracket, f(v_k) \in \text{Vect}(v_1, \dots, v_k)$.

Et une telle famille étant libre et de cardinal $p = \dim E$, ce sera nécessairement une base de E .

8.a. Puisque v_1, \dots, v_k est génératrice de F_k , il suffit de vérifier que

$$\forall i \in \llbracket 1, k \rrbracket, (\varphi - m_{k,k} \text{id}_E)(v_i) \in F_{k-1}.$$

Si $i < k$, alors c'est évident car $\varphi(v_i) \in F_i \subset F_{k-1}$ et $v_i \in F_i \subset F_{k-1}$ de sorte que $\varphi(v_i) - m_{k,k}v_i \in F_k$.

Pour $i = k$, on a, par définition de M , $\varphi(v_k) = \sum_{\ell=1}^p m_{\ell,k}v_\ell$.

Mais d'autre part, $\varphi(v_k) \in \text{Vect}(v_1, \dots, v_k)$, donc nécessairement

$$m_{k+1,k} = m_{k+2,k} = \dots = m_{p,k} = 0.$$

Et donc $\varphi(v_k) - m_{k,k}v_k = \sum_{\ell=1}^k m_{\ell,k}v_\ell - m_{k,k}v_k = \sum_{\ell=1}^{k-1} m_{\ell,k}v_\ell \in \text{Vect}(v_1, \dots, v_{k-1}) = F_{k-1}$.

On en déduit donc que $(\varphi - m_{k,k} \text{id}_E)(F_k) \subset F_{k-1}$.

8.b. Prouvons d'abord qu'il s'agit d'un polynôme annulateur de φ .

Pour $x \in E = F_p$, on a, d'après la question 8.a, $(\varphi - m_{p,p} \text{id}_E)(x) \in F_{p-1}$.

Et donc $(\varphi - m_{p-1,p-1} \text{id}_E)((\varphi - m_{p,p} \text{id}_E)(x)) \in F_{p-2}$.

De proche en proche, on prouve ainsi que

$$\left(\prod_{k=2}^p (X - m_{k,k}) \right) (x) \in F_1 = \text{Vect}(v_1).$$

Or, $\varphi(v_1) \in F_1$, et donc $\varphi(v_1) = m_{1,1}v_1 \Leftrightarrow (\varphi - m_{1,1} \text{id}_E)(v_1) = 0$.

Et donc $F_1 \subset \text{Ker}(\varphi - m_{1,1} \text{id}_E)$.

On en déduit donc que

$$(\varphi - m_{1,1} \text{id}_E) \left(\left(\prod_{k=2}^p (X - m_{k,k}) \right) (x) \right) = 0.$$

Ceci étant vrai pour tout $x \in E$, $\prod_{k=1}^p (X - m_{k,k})$ est annulateur de φ , et donc de $M = \text{Mat}_{\mathcal{B}}(\varphi)$.

9. Soit $A \in \mathcal{M}_p(\mathbf{R})$, et soit φ l'unique endomorphisme de \mathbf{C}^n dont la matrice dans la base canonique est A .

Alors, nous venons de prouver qu'il existe un polynôme $P \in \mathbf{C}[X]$, unitaire et de degré p annulant φ , et donc A .

Notons $P = \sum_{k=0}^p \alpha_k X^k$, de sorte que $\sum_{k=0}^p \alpha_k A^k = 0$.

En conjuguant cette relation, on a alors $\sum_{k=0}^p \overline{\alpha_k} \overline{A^k} = 0$.

Mais A étant à coefficients réels, il en est de même des A^k et donc $\overline{A^k} = A^k$.

Ainsi, on a

$$0 = P(A) + \overline{P(A)} = \sum_{k=0}^p \alpha_k A^k + \sum_{k=0}^p \overline{\alpha_k} A^k = \sum_{k=0}^p (\alpha_k + \overline{\alpha_k}) A^k = 2 \sum_{k=0}^p \text{Re}(\alpha_k) A^k.$$

Matrice

Le fait que F_1, F_2, \dots, F_p soient stables par φ est en fait équivalent au fait que $\text{Mat}_{\mathcal{B}}(\varphi)$ soit triangulaire supérieure.

Remarque

Nous avons dit précédemment que la matrice de φ dans la base \mathcal{B} est triangulaire supérieure. Et nous venons donc de prouver que pour une matrice triangulaire $M \in \mathcal{M}_p(\mathbf{R})$,

$$\prod_{i=1}^p (X - m_{i,i})$$

est un polynôme annulateur de M .

Notons que ce résultat est évident dans le cas des matrices diagonales.

Remarque

Notons que nous avons presque répondu à la question, sauf que P possède des coefficients qui sont a priori complexes et non réels.

Partie réelle

Rappelons que pour tout complexe z ,

$$z + \bar{z} = 2\text{Re}(z).$$

Soit alors $Q(X) = 2 \sum_{k=0}^p \operatorname{Re}(\alpha_k) X^k \in \mathbf{R}_p[X]$. Nous venons de prouver que Q est un polynôme annulateur de A , et il est à coefficients réels. Enfin, il est bien de degré p car P étant unitaire, on a $\alpha_p = 1$ et donc $2\operatorname{Re}(\alpha_p) = 2 \neq 0$. Et donc, le coefficient de degré p de Q étant non nul, on a bien $\deg Q = p$.

Degré

Pour l'instant, Q est de degré inférieur ou égal à p , et l'énoncé demande un polynôme de degré p , donc il faut justifier ce dernier point.

Ainsi, nous avons prouvé qu'il existe un polynôme à coefficients réels, de degré p et annulateur de A .

10.a. Montrons que pour tout $r > p$, $A^{r-1} \in \operatorname{Vect}(I_p, A, A^2, \dots, A^{p-1})$.

Première méthode : par récurrence sur $r > p$.

Soit $P \in \mathbf{R}_p[X]$ un polynôme annulateur de A , de degré p . Quitte à diviser P par son coefficient dominant, on peut le supposer unitaire. Notons alors $Q = X^p + \sum_{k=0}^{p-1} \alpha_k X^k$.

On a alors $A^p = -\sum_{k=0}^{p-1} \alpha_k A^k \in \operatorname{Vect}(I_p, A, A^2, \dots, A^{p-1})$.

Supposons donc que $A^{r-1} = \sum_{k=0}^{p-1} \lambda_k A^k \in \operatorname{Vect}(I_p, A, A^2, \dots, A^{p-1})$. Alors

$$A^r = AA^{r-1} = \sum_{k=0}^{p-1} \lambda_k A^{k+1} = \sum_{k=0}^{p-2} \lambda_k A^{k+1} + \lambda_{p-1} A^p = \sum_{i=1}^{p-1} \lambda_{i-1} A^i - \sum_{k=0}^{p-1} \lambda_p \alpha_k A^k \in \operatorname{Vect}(I_p, A, A^2, \dots, A^{p-1}).$$

Par le principe de récurrence, on en déduit que $A^{r-1} \in \operatorname{Vect}(I_p, A, A^2, \dots, A^{p-1})$, et ce pour tout $r > p$.

Deuxième méthode : à l'aide d'une division euclidienne.

Notons $P \in \mathbf{R}_p[X]$ un polynôme annulateur de A de degré p .

Pour $r \in \mathbf{N}$, soit alors $X^{r-1} = Q_r(X)P(X) + R_r(X)$ la division euclidienne de X^{r-1} par P , avec $\deg R_r < \deg P$.

Alors

$$A^{r-1} = Q_r(A) \underbrace{P(A)}_{=0} + R_r(A) = R_r(A).$$

Mais puisque $R_r(X) \in \mathbf{R}_{p-1}[X]$, on a $A^{r-1} = R_r(A) \in \operatorname{Vect}(I_p, A, A^2, \dots, A^{p-1})$.

Dans les deux cas, pour tout $r \in \mathbf{N}$ il existe des réels $\lambda_0, \dots, \lambda_{p-1}$ tels que $A^{r-1} = \sum_{k=0}^{p-1} \lambda_k A^k$.

En particulier, si $q > p$, alors pour tout $r \in \llbracket 0, q-1 \rrbracket$,

$$A^r B = \sum_{k=0}^{p-1} \lambda_k A^k B \in \operatorname{Vect}(B, AB, A^2 B, \dots, A^{p-1} B) = \mathcal{G}_p.$$

Et donc $\mathcal{G}_q \subset \mathcal{G}_p$.

Puisque l'inclusion réciproque est évidente, on a donc $\mathcal{G}_p = \mathcal{G}_q$.

10.b. Notons $\mathcal{S} = \mathcal{G}_p^\perp$, où l'orthogonal est pris par rapport au produit scalaire canonique de $\mathcal{M}_{p,1}(\mathbf{R})$.

On a alors $\mathcal{G}_p = (\mathcal{G}_p^\perp)^\perp = \mathcal{S}^\perp$.

Et donc un vecteur $G \in \mathcal{M}_{p,1}(\mathbf{R})$ est dans \mathcal{G}_p si et seulement si pour tout $S \in \mathcal{S}$, $\langle G, S \rangle = 0 \Leftrightarrow {}^t S G = 0$.

10.c. Soit $(G_n)_{n \in \mathbf{N}}$ une suite convergente de matrices de \mathcal{G}_p , et soit $G = \lim_{n \rightarrow +\infty} G_n$.

Alors $\lim_{n \rightarrow +\infty} {}^t S G_n = \lim_{n \rightarrow +\infty} 0 = 0$.

D'autre part, d'après les résultats admis dans l'énoncé sur le produit de limites de matrices,

$${}^t S G = {}^t S \left(\lim_{n \rightarrow +\infty} G_n \right) = \lim_{n \rightarrow +\infty} {}^t S G_n = 0.$$

Et donc ${}^t S G = 0 \Leftrightarrow G \in \mathcal{G}_p$.

10.d. Soit $x \in \mathbf{R}$ et $n \in \mathbf{N}$. Alors $T_{A,n}(x)B = \sum_{k=0}^n \frac{1}{k!} (xA)^k B = \sum_{k=0}^n \frac{1}{k!} x^k A^k B \in \mathcal{G}_{n+1}$.

Mais d'après la question 10.a, $\mathcal{G}_{n+1} \subset \mathcal{G}_p$.

Rappel

Pour tout sous-espace vectoriel E d'un espace euclidien,

$$F = (F^\perp)^\perp.$$

D'autre part, nous avons prouvé à la question 2.d que la suite de matrices $(T_{A,n}(x))_{n \in \mathbf{N}}$ converge, et donc il en est de même de $(T_{A,n}(x)B)_{n \in \mathbf{N}}$, qui converge vers $T_A(x)B$. Et alors, d'après la question 10.c, $T_A(x)B \in \mathcal{G}_p$.

Partie III. Contrôle de systèmes linéaires.

11.a. Il est évident que h est dérivable sur $[0, 1]$ par produit de fonctions dérivables, et on a

$$\forall t \in [0, 1], h'(t) = f'(t)e^{-at} - af(t)e^{-at} = e^{-at}(f'(t) - af(t)) = \boxed{e^{-at}bu(t)}.$$

Si on a $f(0) = 0$, alors $h(0) = 0$ et donc il vient, pour tout $t \in [0, 1]$,

$$h(t) = h(t) - h(0) = \int_0^t h'(x) dx = b \int_0^t u(x)e^{-ax} dx.$$

Et donc $f(t) = e^{at}h(t) = b \int_0^t u(x)e^{a(t-x)} dx$.

11.b. Si u est constante égale à λ , alors si $a \neq 0$, on a

$$f(t) = b\lambda e^{at} \int_0^t e^{-ax} dx = b\lambda e^{at} \left[-\frac{1}{a}e^{-ax} \right]_0^t = \frac{b}{a}\lambda e^{at} (1 - e^{-at}) = \frac{b}{a}\lambda(e^{at} - 1).$$

En revanche, si $a = 0$, alors

$$f(t) = be^{at}\lambda \int_0^t 1 dx = b\lambda te^{at}.$$

Ainsi, si $b \neq 0$, alors pour tout $y \in \mathbf{R}$, si on note u_y la fonction constante égale à

En revanche, si $b = 0$, alors d'après 11.a, toute fonction vérifiant $f(0) = 0$ et $f'(t) = af(t)$ est nulle sur $[0, 1]$.

En particulier, pour toute cible $y \neq 0$, et quelle que soit la fonction $u \in \mathcal{C}^0$, on a $f(1) = 0 \neq y$.

Et donc le couple (a, b) n'est pas contrôlable.

12.a. Il a été prouvé à la question 10.d que $T_A(x)B \in \mathcal{G}_p$, quel que soit $x \in \mathbf{R}$. Et donc en particulier, pour tout $x \in [0, 1]$, $W(x) = T_A(1-x)B \in \mathcal{G}_p$.

Soit \mathcal{S} le sous-espace vectoriel de $\mathcal{M}_{p,1}(\mathbf{R})$ défini à la question 10.b, et soit $S = \begin{pmatrix} s_1 \\ \vdots \\ s_p \end{pmatrix} \in \mathcal{S}$.

Alors, pour tout $x \in [0, 1]$, $W(x) \in \mathcal{G}_p$ de sorte que ${}^tSW(x) = 0 \Leftrightarrow \sum_{k=1}^p s_k W_k(x) = 0$.

Alors

$${}^tS \left(\int_0^1 u(x)W(x) dx \right) = \sum_{k=1}^p s_k \int_0^1 u(x)W_k(x) dx = \int_0^1 u(x) \left(\sum_{k=1}^p s_k W_k(x) \right) dx = \int_0^1 u(x) \times 0 dx = 0.$$

Ceci étant vrai pour tout $S \in \mathcal{S}$, on en déduit que $\boxed{\int_0^1 u(x)W(x) dx \in \mathcal{G}_p}$.

12.b. Notons $Z = \begin{pmatrix} z_1 \\ \vdots \\ z_p \end{pmatrix}$.

Alors l'hypothèse est que pour toute fonction continue $u \in \mathcal{C}^0$, $\int_0^1 u(x) \left(\sum_{k=1}^p z_k W_k(x) \right) dx = 0$.

En particulier, c'est le cas si $u(x) = \sum_{k=1}^p s_k W_k(x) \in \mathcal{C}^0$. Et donc

$$\int_0^1 \left(\sum_{k=1}^p s_k W_k(x) \right)^2 dx = 0.$$

Mais l'intégrale d'une fonction positive et continue est nulle si et seulement si cette fonction est nulle, donc pour tout $x \in [0, 1]$, $\sum_{k=1}^p s_k W_k(x) = 0 \Leftrightarrow {}^t S W_k(x) = 0$.

12.c. Par définition de W , pour tout $x \in [0, 1]$, on a ${}^t ZW(x) = {}^t ZT_A(1-x)B = 0$.
 Quitte à changer x en $-x$, on a donc, pour tout $x \in [0, 1]$, ${}^t ZT_A(x)B = 0$.
 Reprenons comme indiqué les relations de la question 4.b : pour $n \in \mathbf{N}$,

$$\lim_{r \rightarrow +\infty} r^{n+1} \left(T_A \left(\frac{1}{r} \right) - T_{A,n} \left(\frac{1}{r} \right) \right) = \frac{1}{(n+1)!} A^{n+1}.$$

Après multiplication à gauche par ${}^t Z$ et à droite par B , on a donc

$$\lim_{r \rightarrow +\infty} r^{n+1} {}^t Z \left(T_A \left(\frac{1}{r} \right) - T_{A,n} \left(\frac{1}{r} \right) \right) B = \frac{1}{(n+1)!} {}^t Z A^{n+1} B.$$

Notons qu'il s'agit là d'une suite de réels : $v_r = r^{n+1} {}^t Z \left(T_A \left(\frac{1}{r} \right) - T_{A,n} \left(\frac{1}{r} \right) \right) B$ est une matrice 1×1 , donc un réel.

De plus, pour $r \geq 1$, ${}^t Z T_A \left(\frac{1}{r} \right) B = {}^t ZW \left(\frac{1}{r} \right)$, qui vaut donc 0 d'après la question précédente⁶.

Donc $v_r = -r^{n+1} {}^t Z T_{A,n} \left(\frac{1}{r} \right) B$.

Raisonnons par l'absurde, et supposons que l'un des ${}^t Z A^{k-1} B$, $k \in \llbracket 0, n \rrbracket$ soit non nul, et notons k_0 le plus petit des tels k . Alors

$$v_r = -r^{n+1} {}^t Z \left(\sum_{i=0}^n \frac{1}{i!} \left(\frac{1}{r} A \right)^i \right) B = -r^{n+1} \sum_{i=k_0}^n \frac{1}{i!} {}^t Z A^i B r^{n+1-i}.$$

Pour $i > k_0$, $r^{n+1-i} = \underset{r \rightarrow +\infty}{o} (r^{n+1-k_0})$.

Et donc⁷ $v_r \underset{r \rightarrow +\infty}{\sim} -\frac{1}{r^{n+1-k_0}} {}^t Z A^{k_0} B$, qui tend vers $\pm\infty$ suivant le signe du réel non nul ${}^t Z A^{k_0} B$.

Ceci contredit le fait que la suite (v_r) converge.

Et donc on en déduit que pour tout $k \in \llbracket 0, n \rrbracket$, ${}^t Z A^k B = 0$.

Ceci étant vrai pour tout $n \in \mathbf{N}$, nous avons en fait prouvé que $\forall k \in \mathbf{N}$, ${}^t Z A^k B = 0$.

Et donc bien entendu, ceci implique que pour tout $k \in \llbracket 1, p \rrbracket$, ${}^t Z A^{k-1} B = 0$.

12.d. Supposons que (A, B) est contrôlable, et montrons que K_p est inversible.

Alors, pour tout $Y \in \mathcal{M}_{p,1}(\mathbf{R})$, il existe une fonction $u \in \mathcal{C}^0$ telle que $\int_0^1 u(x)W(x) dx = Y$.

Et donc en particulier, $Y \in \mathcal{G}_p$.

On en déduit donc que $\mathcal{G}_p = \mathcal{M}_{p,1}(\mathbf{R})$. Mais $B, AB, \dots, A^{p-1}B$ est par définition une famille génératrice de \mathcal{G}_p , c'est donc une famille génératrice de $\mathcal{M}_{p,1}(\mathbf{R})$, et donc une famille de rang p .

Mais le rang de K_p étant égal au rang de la famille de ses vecteurs colonnes, on en déduit que $\text{rg } K_p = p \Leftrightarrow K_p$ est inversible.

Inversement, supposons K_p inversible, de sorte que $(B, AB, \dots, A^{p-1}B)$ est de rang p , et donc une famille libre de $\mathcal{M}_{p,1}(\mathbf{R})$. Alors, étant de cardinal $p = \dim \mathcal{M}_{p,1}(\mathbf{R})$, c'est une base de $\mathcal{M}_{p,1}(\mathbf{R})$. En particulier, si $Y \in \mathcal{M}_{p,1}(\mathbf{R})$,

$$\text{Soit } F = \left\{ Y \in \mathcal{M}_{p,1}(\mathbf{R}) : \exists u \in \mathcal{C}^0 \text{ tel que } \int_0^1 u(x)W(x) dx = Y \right\} = \left\{ \int_0^1 u(x)W(x) dx, u \in \mathcal{C}^0 \right\}$$

l'ensemble des cibles qui peuvent être atteintes.

Il est facile de voir que F est un sous-espace vectoriel de $\mathcal{M}_{p,1}(\mathbf{R})$ car $0 = \int_0^1 0 \times W(x) dx$

et que si $Y_1 = \int_0^1 u_1(x)W(x) dx$ et $Y_2 = \int_0^1 u_2(x)W(x) dx$, alors pour $\lambda \in \mathbf{R}$, on a

$$\lambda Y_1 + Y_2 = \int_0^1 (\lambda u_1(x) + u_2(x))W(x) dx \in F.$$

⁶ $\frac{1}{r}$ est dans $[0, 1]$.

⁷ Notons que n étant fixé, la somme comporte un nombre fixé de termes.

Remarque

Prouver le résultat pour tout $k \in \mathbf{N}$ ou pour tout $k \in \llbracket 0, p-1 \rrbracket$ revient au même, puisque nous savons que les A^k , $k \geq p$ sont combinaisons linéaires de I_p, A, \dots, A^{p-1} .

Remarque

Un argument de cardinalité prouve même qu'il s'agit d'une base de $\mathcal{M}_{p,1}(\mathbf{R})$.

Soit donc $Z \in F^\perp$. Alors pour tout $u \in \mathcal{C}^0$,

$${}^t Z \left(\int_0^1 u(x)W(x) dx \right) = 0 \Leftrightarrow \int_0^1 u(x) {}^t Z W(x) dx = 0.$$

D'après les question 12.b et 12.c, on a donc pour tout $k \in \llbracket 1, p \rrbracket$, ${}^t Z A^{k-1} B = 0$.

Mais $(B, AB, \dots, A^{p-1}B)$ étant une base de $\mathcal{M}_{p,1}(\mathbf{R})$, on en déduit que $Z \in (\mathcal{M}_{p,1}(\mathbf{R}))^\perp = \{0\}$, et donc $Z = 0$.

Ainsi, $F^\perp = \{0\}$ et donc $F = (F^\perp)^\perp = \mathcal{M}_{p,1}(\mathbf{R})$: pour tout $Y \in \mathcal{M}_{p,1}(\mathbf{R})$, il existe $u \in \mathcal{C}^0$

telle que $\int_0^1 u(x)W(x) dx = Y$. Et donc le système (A, B) est contrôlable.

DEVOIR MAISON 10 (À RENDRE LE 20.12.18)

Toutes les variables aléatoires considérées ici sont définies sur un même espace probabilisé (Ω, \mathcal{A}, P) .

Soit X une variable aléatoire, appelons que :

- F_X désigne sa fonction de répartition définie par : $\forall t \in \mathbf{R}, F_X(t) = P(X \leq t)$.
- X suit la loi exponentielle de paramètre $a \in]0; +\infty[$ si et seulement si sa fonction de répartition est définie par :

$$F_X(t) = \begin{cases} 0 & \text{si } t < 0 \\ 1 - \exp(-at) & \text{si } t \geq 0 \end{cases}$$

où $\exp(y)$ désigne l'exponentielle du réel y c'est-à-dire que : $\exp(y) = e^y$.

On considère une suite $(X_n)_{n \in \mathbf{N}^*}$ de variables aléatoires mutuellement indépendantes suivant la même loi exponentielle de paramètre 1. Pour tout entier naturel n non nul, on note Y_n et Z_n les deux variables aléatoires définies respectivement par :

$$Y_n = \max(X_1, \dots, X_n) = \max_{1 \leq i \leq n} X_i.$$

$$Z_n = \frac{X_1}{1} + \frac{X_2}{2} + \dots + \frac{X_n}{n} = \sum_{i=1}^n \frac{X_i}{i},$$

où $\max(X_1, \dots, X_n)$ désigne le maximum des valeurs de X_1, \dots, X_n .

Pour finir, on désigne par f_n la fonction définie sur \mathbf{R} par

$$f_n(t) = \begin{cases} 0 & \text{si } t < 0 \\ n \exp(-t) (1 - \exp(-t))^{n-1} & \text{si } t \geq 0 \end{cases}$$

1.
 - a. Pour tout réel t , exprimer le réel $F_{Y_n}(t)$ à l'aide des réels $F_{X_1}(t), \dots, F_{X_n}(t)$.
 - b. Pour tout réel t , donner alors l'expression de $F_{Y_n}(t)$ en fonction de n et de t en distinguant le cas $t < 0$ et le cas $t \geq 0$.
 - c. Vérifier alors que la fonction f_n est une densité de probabilité de la variable aléatoire Y_n .
2.
 - a. Préciser la fonction de répartition de la variable aléatoire $\frac{X_{n+1}}{n+1}$.
 - b. Démontrer que $\frac{X_{n+1}}{n+1}$ est une variable aléatoire à densité et proposer une densité d_{n+1} .
3. Pour tout réel x , vérifier que : $\int_0^x n \exp(nt) (1 - \exp(-t))^{n-1} dt = (\exp(x) - 1)^n$.
4. Montrer par récurrence que, pour tout entier naturel n non nul, Z_n est une variable aléatoire à densité dont f_n est une densité.
Indication : pour l'hérédité, on remarquera que $Z_{n+1} = Z_n + \frac{X_{n+1}}{n+1}$.
5. Sans calculer d'intégrale, montrer que Y_n admet une espérance, et calculer $E(Y_n)$.

CORRECTION DU DEVOIR MAISON 10

ECRICOME 2011

1.a. On a

$$F_{Y_n}(t) = P(Y_n \leq t) = P(\max(X_1, \dots, X_n) \leq t) = P\left(\bigcap_{i=1}^n [X_i \leq t]\right).$$

Mais les X_i étant indépendantes, il vient alors

$$F_{Y_n}(t) = \prod_{i=1}^n P(X_i \leq t) = \prod_{i=1}^n F_{X_i}(t).$$

1.b. Puisque les X_i sont identiquement distribuées, pour tout $i \in \llbracket 1, n \rrbracket$, $F_{X_i}(t) = F_{X_1}(t)$ et donc

$$F_{Y_n}(t) = (F_{X_1}(t))^n = \begin{cases} 0 & \text{si } t < 0 \\ (1 - e^{-t})^n & \text{si } t \geq 0 \end{cases}$$

1.c. Puisque X_1 est une variable à densité, F_{X_1} est continue sur \mathbf{R} et \mathcal{C}^1 sauf éventuellement en un nombre fini de points, et donc $F_{Y_n} = (F_{X_1})^n$ est également continue sur \mathbf{R} et \mathcal{C}^1 sauf éventuellement en un nombre fini de points.Donc Y_n est une variable à densité.D'autre part, toute fonction coïncidant avec F'_{Y_n} sauf éventuellement en un nombre fini de points est une densité de Y_n .Mais F_{Y_n} est dérivable sur \mathbf{R}^* et pour tout $t \in \mathbf{R}^*$,

$$F'_{Y_n}(t) = \begin{cases} 0 & \text{si } t < 0 \\ ne^{-t}(1 - e^{-t})^{n-1} & \text{si } t > 0 \end{cases} = f_n(t).$$

Et donc f_n est une densité de Y_n .2.a. Pour $t \in \mathbf{R}$, on a

$$P\left(\frac{X_{n+1}}{n+1} \leq t\right) = P(X_{n+1} \leq t(n+1)) = \begin{cases} 0 & \text{si } t < 0 \\ 1 - e^{-(n+1)t} & \text{si } t \geq 0 \end{cases}$$

2.b. La fonction de répartition de $\frac{X_{n+1}}{n+1}$ est clairement \mathcal{C}^1 sur $] -\infty, 0[$ et sur $]0, +\infty[$.

De plus, on a

$$\lim_{t \rightarrow 0^-} P\left(\frac{X_{n+1}}{n+1} \leq t\right) = 0 = \lim_{t \rightarrow 0^+} P\left(\frac{X_{n+1}}{n+1} \leq t\right) = P\left(\frac{X_{n+1}}{n+1} \leq 0\right)$$

et donc la fonction de répartition de $\frac{X_{n+1}}{n+1}$ est continue en 0. Ainsi, elle est continue sur \mathbf{R} et \mathcal{C}^1 sauf peut-être en 0 : $\frac{X_{n+1}}{n+1}$ est une variable à densité, et une densité en est toute fonction qui coïncide avec $F'_{\frac{X_{n+1}}{n+1}}$ sauf éventuellement en un nombre fini de points.Par exemple, on peut prendre $d_{n+1} : t \mapsto \begin{cases} 0 & \text{si } t < 0 \\ (n+1)e^{-(n+1)t} & \text{si } t \geq 0 \end{cases}$ 3. Soit $x \in \mathbf{R}$. Alors on a

$$\begin{aligned} \int_0^x ne^{nt}(1 - e^{-t})^{n-1} dt &= \int_0^1 ne^{nt}(e^{-t}(e^t - 1))^{n-1} dt \\ &= \int_0^1 ne^{nt}e^{-(n-1)t}(e^t - 1)^{n-1} dt = \int_0^x ne^t(e^t - 1)^{n-1} dt \\ &= [(e^t - 1)^n]_0^x = (e^x - 1)^n. \end{aligned}$$

Détails

Dans le cas d'une loi exponentielle, nous savons même que F_{X_1} est \mathcal{C}^1 sauf en 0.

Densité

Rappelons que si on change la valeur d'une densité en un nombre fini de points, on obtient encore une densité. Pour cette raison, nous ne nous préoccupons pas de la valeur de f_n en 0, où F_{Y_n} n'est pas dérivable.

Remarque

L'énoncé nous demandait de repasser par la fonction de répartition, mais c'est un résultat classique sur les lois exponentielles : si $X \leftrightarrow \mathcal{E}(1)$, alors $\lambda X \leftrightarrow \mathcal{E}\left(\frac{1}{\lambda}\right)$ et donc ici $\frac{X_{n+1}}{n+1} \leftrightarrow \mathcal{E}(n+1)$.

4. Montrons le résultat par récurrence sur $n \in \mathbf{N}^*$.
 Pour $n = 1$, $Z_1 = X_1 = Y_1$. Donc f_1 , qui est une densité de Y_1 , est également une densité de Z_1 .

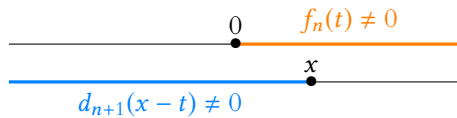
Supposons que f_n soit une densité de Z_n .

Alors $Z_{n+1} = Z_n + \frac{X_{n+1}}{n+1}$. Par le lemme des coalitions, Z_n , qui est une fonction de X_1, \dots, X_n est indépendante de $\frac{X_{n+1}}{n+1}$.

Puisque d_{n+1} est bornée¹ sur \mathbf{R} , on peut donc utiliser le produit de convolution : Z_{n+1} est une variable à densité, dont une densité est donnée par

$$g_{n+1}(x) = \int_{-\infty}^{+\infty} f_n(t)d_{n+1}(x-t) dt.$$

Or, on a $f_n(t) \neq 0 \Leftrightarrow t \geq 0$ et $d_{n+1}(x-t) \neq 0 \Leftrightarrow x-t \geq 0 \Leftrightarrow t \leq x$.



Ainsi, pour $x \leq 0$, on a $g_{n+1}(x) = 0 = f_{n+1}(x)$ et pour $x > 0$,

$$\begin{aligned} g_{n+1}(x) &= \int_0^x ne^{-t}(1-e^{-t})^{n-1}(n+1)e^{-(n+1)(x-t)} dt \\ &= (n+1)e^{-(n+1)x} \int_0^x e^{nt}(1-e^{-t})^{n-1} dt \\ &= (n+1)e^{-(n+1)x}(e^x-1)^n \\ &= (n+1)e^{-x}(1-e^{-x})^n = f_{n+1}(x). \end{aligned}$$

Et donc f_{n+1} est une densité de Z_{n+1} .

D'après le principe de récurrence, pour tout $n \in \mathbf{N}^*$, f_n est une densité de Z_n .

5. Puisque Z_n est Y_n ont même densité, elles suivent la même loi, et par conséquent, ont la même espérance.
 Or l'espérance de Z_n est facile à calculer grâce à la linéarité de l'espérance :

$$E(Z_n) = E\left(\sum_{k=1}^n \frac{X_k}{k}\right) = \sum_{k=1}^n \frac{1}{k}E(X_k) = \sum_{k=1}^n \frac{1}{k}.$$

Et donc $E(Y_n) = \sum_{k=1}^n \frac{1}{k}$.

bispt

¹ Par $n+1$.

Densité

La somme de deux variables à densité n'est pas toujours une variable à densité (penser à $X - X = 0$...).
 En revanche, c'est le cas si les deux variables sont indépendantes, et c'est le théorème de convolution qui l'affirme : « $X + Y$ est une variable à densité, dont une densité est donnée par ...»

DEVOIR MAISON 11 (À RENDRE LE 10.01.19)

Les questions marquées d'une étoile (★) sont réservées aux cubes. Le résultat de la question 14.c peut être admis pour aborder les deux questions suivantes.

Dans tout le problème :

- toutes les variables aléatoires sont supposées définies sur un espace probabilisé (Ω, \mathcal{A}, P) ;
- pour tout réel $t > 0$, X_t désigne une variable aléatoire à valeurs strictement positives qui suit la loi gamma de paramètre t , notée $\gamma(t)$;
- U désigne une variable aléatoire qui suit la loi uniforme sur $]0, 1]$;
- l'espérance et la variance d'une variable aléatoire A sont notées respectivement $E(A)$ et $V(A)$;
- la notation \exp désigne la fonction exponentielle de base e .

On rappelle ou on admet sans démonstration les résultats suivants :

- la fonction Γ définie pour tout réel $t > 0$ par $\Gamma(t) = \int_0^{+\infty} e^{-u} u^{t-1} du$, est de classe \mathcal{C}^∞ sur \mathbf{R}^{**} ; on note Γ' et Γ'' les dérivées premières et secondes de la fonction Γ . Pour tout réel t les intégrales $\int_0^{+\infty} (\ln u) e^{-u} u^{t-1} du$ et $\int_0^{+\infty} (\ln u)^2 e^{-u} u^{t-1} du$ sont convergentes et valent respectivement $\Gamma'(t)$ et $\Gamma''(t)$;
- on a pour tout réel $t > 0$: $\Gamma(t+1) = t\Gamma(t)$;
- pour tout réel $t > 0$, une densité f_{X_t} de X_t est donnée par : $f_{X_t}(x) = \begin{cases} \frac{1}{\Gamma(t)} e^{-x} x^{t-1} & \text{si } x > 0 \\ 0 & \text{si } x \leq 0 \end{cases}$;
- $\Gamma(1/2) = \sqrt{\pi}$.

L'objet du problème est de démontrer quelques propriétés de la fonction Γ en utilisant des méthodes essentiellement probabilistes.

Partie I. Quelques résultats préliminaires

1. On pose pour tout n de \mathbf{N}^* : $h_n = \sum_{k=1}^n \frac{1}{k}$. On considère les deux suites $(\gamma_n)_{n \geq 1}$ et $(v_n)_{n \geq 1}$ définies par : pour tout n de \mathbf{N}^* , $\gamma_n = h_n - \ln(n)$ et $v_n = \gamma_{n+1} - \gamma_n$.
 - a. Montrer que la série de terme général v_n est convergente.
 - b. En déduire la convergence de la suite $(\gamma_n)_{n \geq 1}$; on note γ sa limite.
 - c. On pose pour tout réel $t > 0$: $d_{n,t} = \gamma + \ln(t+n) - h_n$. Déterminer $\lim_{n \rightarrow +\infty} d_{n,t}$.
2.
 - a. Rappeler sans démonstration les valeurs respectives de $E(X_t)$ et $V(X_t)$.
 - b. On note pour tout réel $t > 0$, $\psi(t) = \Gamma'(t)/\Gamma(t)$, et ψ' la dérivée de ψ . Montrer que $E(\ln(X_t)) = \psi(t)$ et $V(\ln(X_t)) = \psi'(t)$.
3.
 - a. Montrer que pour tout réel $t > 1$, $E(1/X_t)$ existe et calculer sa valeur.
 - b. Établir pour tout réel $x > 0$, l'encadrement : $1 - \frac{1}{x} \leq \ln x \leq x - 1$.
En déduire que l'on a : $(\ln x)^2 \leq \left(1 - \frac{1}{x}\right)^2 + (x-1)^2$.
 - c. À l'aide des questions précédentes, établir les inégalités suivantes : pour tout réel $t > 0$, $E\left(\ln\left(\frac{X_t}{t}\right)\right) \leq 0$; pour tout réel $t > 1$, $E\left(\ln\left(\frac{X_t}{t}\right)\right) \geq -\frac{1}{t-1}$; pour tout réel $t > 2$, $E\left(\ln^2\left(\frac{X_t}{t}\right)\right) \leq \frac{2t}{(t-2)^2}$.
 - d. Soit t un réel fixé strictement positif. Montrer que la suite de variables aléatoires $\left(\ln\left(\frac{X_{t+n}}{t+n}\right)\right)_{n \geq 1}$ converge en probabilité vers 0.
4. (★) Soit $(A_n)_{n \geq 1}$, $(B_n)_{n \geq 1}$ et $(C_n)_{n \geq 1}$ trois suites de variables aléatoires à densité qui convergent en probabilité vers 0. On pose pour tout n de \mathbf{N}^* : $D_n = A_n + B_n + C_n$. Soit $(u_n)_{n \geq 1}$ une suite réelle qui converge vers u . On considère deux variables aléatoires réelles à densité M et N telles que pour tout n de \mathbf{N}^* , M est de même loi que $N + D_n + u_n$.

- a. Montrer pour tout réel $\varepsilon > 0$, l'inclusion : $\{|D_n| > \varepsilon\} \subset \{|A_n| > \varepsilon/3\} \cup \{|B_n| > \varepsilon/3\} \cup \{|C_n| > \varepsilon/3\}$.
En déduire que la suite $(D_n)_{n \geq 1}$ converge en probabilité vers 0.
 - b. On pose pour tout n de \mathbf{N}^* : $V_n = D_n + u_n - u$. Montrer que la suite de variables aléatoires $(V_n)_{n \geq 1}$ converge en probabilité vers 0. En déduire la limite en probabilité de la suite $((N + u) + V_n)_{n \geq 1}$.
 - c. On admet sans démonstration que la convergence en probabilité entraîne la convergence en loi. Montrer que les variables aléatoires M et $N + u$ sont de même loi.
5. Soit (α, β) un couple de réels strictement positifs, et X_α et X_β deux variables aléatoires indépendantes de lois respectives $\gamma(\alpha)$ et $\gamma(\beta)$. On pose : $T_{\alpha, \beta} = \frac{X_\alpha}{X_\beta}$, $Q_{\alpha, \beta} = \ln(T_{\alpha, \beta})$ et $B(\alpha, \beta) = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha + \beta)}$.
- a. Préciser $Q_{\alpha, \beta}(\Omega)$. Déterminer une densité de $\ln(X_\alpha)$ et de $-\ln(X_\beta)$ respectivement.
 - b. En déduire qu'une densité $f_{Q_{\alpha, \beta}}$ de $Q_{\alpha, \beta}$ est donnée par : pour tout réel x ,

$$f_{Q_{\alpha, \beta}}(x) = \frac{e^{-\beta x}}{\Gamma(\alpha)\Gamma(\beta)} \int_{-\infty}^{+\infty} e^{(\alpha+\beta)y} \exp(-e^y(1 + e^{-x})) dy$$

- c. À l'aide du changement de variable $u = e^y(1 + e^{-x})$, dont on justifiera la validité, établir la formule suivante : pour tout réel x , $f_{Q_{\alpha, \beta}}(x) = \frac{1}{B(\alpha, \beta)} \times \frac{e^{\alpha x}}{(1 + e^x)^{\alpha+\beta}}$.
- d. En déduire une densité $f_{T_{\alpha, \beta}}$ de $T_{\alpha, \beta}$.
- e. On pose : $J_{\alpha, \beta} = \frac{X_\alpha}{X_\alpha + X_\beta}$. Montrer qu'une densité $f_{J_{\alpha, \beta}}$ de $J_{\alpha, \beta}$ est donnée par :

$$f_{J_{\alpha, \beta}}(z) = \begin{cases} 0 & \text{si } z \notin]0, 1[\\ \frac{1}{B(\alpha, \beta)} z^{\alpha-1} (1-z)^{\beta-1} & \text{si } z \in]0, 1[\end{cases}$$

Partie II. Étude de la variable aléatoire $\ln(X_t)$

Soit $(Y_k)_{k \geq 1}$ une suite de variables aléatoires indépendantes de même loi exponentielle de paramètre 1. On suppose que pour tout réel $\alpha > 0$, X_α est indépendante de chacune des variables aléatoires de la suite $(Y_k)_{k \geq 1}$.

On pose : $S_0 = 0$ et pour tout k de \mathbf{N}^* , $S_k = \sum_{i=1}^k Y_i$.

- 6. Rappeler sans démonstration la loi de S_k ainsi que les valeurs respectives de $E(S_k)$ et $V(S_k)$.
- 7. a. Justifier pour tout n de \mathbf{N}^* , l'égalité suivante : $\ln(X_t) = \ln\left(\prod_{k=1}^n \frac{X_t + S_{k-1}}{X_t + S_k}\right) + \ln(X_t + S_n)$.
- b. Montrer que pour tout entier m de \mathbf{N}^* , la loi de $X_t + S_m$ est celle de X_{t+m} .

On admet jusqu'à la fin du problème les résultats suivants :

- soit n un entier de \mathbf{N}^* et A_1, \dots, A_n des variables aléatoires réelles mutuellement indépendantes. Alors, pour tout p de $\llbracket 1, n-1 \rrbracket$, pour toutes fonctions réelles φ_1 et φ_2 continues, les variables aléatoires $\varphi_1(A_1, \dots, A_p)$ et $\varphi_2(A_{p+1}, \dots, A_n)$ sont indépendantes ;
- si A et B sont deux variables aléatoires indépendantes admettant une espérance, alors AB admet une espérance et $E(AB) = E(A)E(B)$;
- pour tout couple (α, β) de réels strictement positifs, si les variables aléatoires X_α et X_β sont indépendantes, de lois respectives $\gamma(\alpha)$ et $\gamma(\beta)$, alors les variables aléatoires $\frac{X_\alpha}{X_\alpha + X_\beta}$ et $X_\alpha + X_\beta$ sont indépendantes.

- 8. On pose pour tout k de \mathbf{N}^* : $R_{t, k} = \frac{X_t + S_{k-1}}{X_t + S_k}$.
 - a. Montrer que $R_{t, 1}$ et $R_{t, 2}$ sont indépendantes. On admet dans la suite que les variables aléatoires $R_{t, k}$ ($k \geq 1$) sont indépendantes.
 - b. En déduire, à l'aide des questions précédentes que pour tout n de \mathbf{N}^* , les variables aléatoires $\ln(X_t)$ et $\sum_{k=1}^n \ln(R_{t, k}) + \ln(X_{t+n})$ sont de même loi.
- 9. a. Déterminer la loi de la variable aléatoire $-\ln(U)$.
 - b. À l'aide de la question 5.e, calculer une densité $f_{R_{t, k}}$ de la variable aléatoire $R_{t, k}$.

- c. Soit $(U_k)_{k \geq 1}$ une suite de variables aléatoires indépendantes de même loi que U . Montrer que pour tout k de \mathbf{N}^* , les variables aléatoires $R_{t,k}$ et $U_k^{\frac{1}{t+k-1}}$ sont de même loi.
- d. Dédurre des questions précédentes que pour tout n de \mathbf{N}^* , la variable aléatoire $\ln(X_t)$ est de même loi que la variable aléatoire $\sum_{k=0}^{n-1} \left(\frac{1}{k+1} - \frac{Y_{k+1}}{t+k} \right) + \ln \left(\frac{X_{t+n}}{t+n} \right) + d_{n,t} - \gamma$.

10. En utilisant les résultats des questions 1.c, 2.b, 3.c et 9.d, montrer que pour tout réel $t > 0$, on a :

$$E(\ln(X_t)) = \psi(t) = -\gamma + \sum_{k=0}^{+\infty} \left(\frac{1}{k+1} - \frac{1}{t+k} \right) \text{ et } V(\ln(X_t)) = \psi'(t) = \sum_{k=0}^{+\infty} \frac{1}{(t+k)^2}.$$

11. En utilisant la question 3.c, calculer $\lim_{t \rightarrow +\infty} (\psi(t) - \ln(t))$.
12. (★) On pose : $W = U^{1/\mu}$, où μ désigne un paramètre réel strictement positif inconnu. Afin d'estimer μ , on considère pour p supérieur ou égal à 3, un p -échantillon (W_1, W_2, \dots, W_p) i.i.d. de la loi de W .

On pose $G_p = -p \left(\sum_{i=1}^p \ln W_i \right)^{-1}$. Justifier que la variable aléatoire G_p est un estimateur du paramètre μ . Est-il sans biais ? Est-il convergent ?

13. On rappelle que l'appel à la fonction Scilab `rand()` a pour résultat un nombre réel pris au hasard dans l'intervalle $[0, 1]$, suivant une loi uniforme.

- a. Soit X la fonction Scilab suivante :

```
1 function y=X(lambda)
2     y = -log(1-rand())/lambda
3 endfunction
```

Cette fonction simule une variable aléatoire réelle. Donner sa loi. Justifier votre réponse.

- b. Écrire une fonction Scilab d'entête `function y = g(n)` simulant une variable aléatoire de loi $\gamma(n)$.

- c. (★) Soit m la fonction Scilab suivante :

```
1 function y = m(p)
2     m = p/g(p)
3 endfunction
```

On appelle la fonction `m` pour différentes valeurs de p , de plus en plus grandes. Que devrait-on constater ?

Partie III. Quelques propriétés de la fonction Γ

Les notations sont celles des parties I et II.

14. Premières applications : les formules de Wilks et Legendre.

- a. Soit $(a_k)_{k \in \mathbf{N}}$ une suite réelle. Pour tout n de \mathbf{N}^* et pour tout réel $t > 0$, établir l'égalité :

$$2 \sum_{k=0}^{2n-1} \left(\frac{1}{k+1} - \frac{a_{k+1}}{t+k} \right) = \sum_{k=0}^{n-1} \left(\frac{1}{k+1} - \frac{a_{2k+1}}{\frac{t}{2} + k} \right) + \sum_{k=0}^{n-1} \left(\frac{1}{k+1} - \frac{a_{2k+2}}{\frac{t+1}{2} + k} \right) + 2 \sum_{k=0}^{n-1} \left(\frac{1}{2k+1} - \frac{1}{2k+2} \right)$$

- b. Exprimer $w_n = \sum_{k=0}^{n-1} \left(\frac{1}{2k+1} - \frac{1}{2k+2} \right)$ en fonction de deux termes de la suite $(h_n)_{n \geq 1}$. En déduire que $\lim_{n \rightarrow +\infty} w_n = \ln 2$.

- c. (★) Pour $t > 0$, soit X_t et $X_{t+\frac{1}{2}}$ deux variables aléatoires indépendantes de lois respectives $\gamma(t)$ et $\gamma(t+\frac{1}{2})$. En utilisant les questions 4 et 9.d, montrer que la variable aléatoire $2 \ln(X_t)$ est de même loi que la variable aléatoire $\ln(X_{\frac{t}{2}}) + \ln(X_{\frac{t+1}{2}}) + 2 \ln 2$.

- d. On pose : $t = 2s$. Dédurre de la question précédente que pour tout réel $r > 0$, $(X_{2s})^{2r}$ et $2^{2r} (X_s)^r (X_{s+\frac{1}{2}})^r$ sont de même loi.

- e. En choisissant une valeur particulière de s , établir pour tout $r > 0$, la formule :

$$2^{2r-1} \Gamma(r) \Gamma \left(r + \frac{1}{2} \right) = \Gamma(2r) \sqrt{\pi}.$$

15. Deuxième application : la formule de Stirling

- a. Déterminer quatre réels a, b, c, d tels que pour tout réel $u > 0$, on a :

$$\frac{1}{u^2(u+1)^2} = \frac{a}{u^2} + \frac{b}{(u+1)^2} + \frac{c}{u} + \frac{d}{u+1}.$$

En déduire pour tout $t > 0$, la relation : $\psi'(t) = \frac{1}{t} + \frac{1}{2t^2} + \frac{1}{2} \sum_{k=0}^{+\infty} \frac{1}{(t+k)^2(t+k+1)^2}$.

On admet sans démonstration que pour tout $u > 0$, on a :

$$\frac{1}{3} \left(\frac{1}{\left(u + \frac{1}{14}\right)^3} - \frac{1}{\left(u + \frac{15}{14}\right)^3} \right) \leq \frac{1}{u^2(u+1)^2} \leq \frac{1}{3} \left(\frac{1}{u^3} - \frac{1}{(u+1)^3} \right)$$

- b. Déduire des deux résultats précédents, pour tout $t > 0$, les deux encadrements :

$$\frac{1}{2t^2} + \frac{1}{6\left(t + \frac{1}{14}\right)^3} \leq \psi'(t) - \frac{1}{t} \leq \frac{1}{2t^2} + \frac{1}{6t^3} \text{ et } \ln t - \frac{1}{2t} - \frac{1}{12t^2} \leq \psi(t) \leq \ln t - \frac{1}{2t} - \frac{1}{12\left(t + \frac{1}{14}\right)^2}$$

En déduire un équivalent de $E(\ln(X_t))$ et de $V(\ln(X_t))$ respectivement, lorsque t tend vers $+\infty$.

- c. Calculer pour tout y vérifiant $y > t > 0$, l'intégrale : $\int_t^y \left(\psi(x) - \ln(x) + \frac{1}{2x} \right) dx$.

Montrer pour t fixé, l'existence de $\lim_{y \rightarrow +\infty} \ln \left(\frac{\Gamma(y)}{y^{y-\frac{1}{2}} e^{-y}} \right)$; on note θ cette limite.

- d. En utilisant la question 14.e et l'identité $x^x = \left(x + \frac{1}{2}\right)^x \left(1 + \frac{1}{2x}\right)^{-x}$, valable pour $x > 0$, calculer e^θ .

En déduire que $\Gamma(x)$ est équivalent à $\sqrt{2\pi} x^{x-\frac{1}{2}} e^{-x}$, lorsque x tend vers $+\infty$.

CORRECTION DU DEVOIR MAISON 11

MATHS II 2010

Partie I. Quelques résultats préliminaires.

1.a. On a

$$v_n = \gamma_{n+1} - \gamma_n = \sum_{k=1}^{n+1} \frac{1}{k} - \ln(n) - \sum_{k=1}^n \frac{1}{k} - \ln(n+1) = \frac{1}{n+1} - \ln\left(1 + \frac{1}{n}\right).$$

Mais nous savons que $\ln\left(1 + \frac{1}{n}\right) = \frac{1}{n} - \frac{1}{2n^2} + o\left(\frac{1}{n^2}\right)$.

Et de même, $\frac{1}{1+n} = \frac{1}{n} \frac{1}{1+\frac{1}{n}} = \frac{1}{n} \left(1 - \frac{1}{n} + o\left(\frac{1}{n}\right)\right) = \frac{1}{n} - \frac{1}{n^2} + o\left(\frac{1}{n^2}\right)$.

Et donc $v_n = -\frac{1}{n^2} + o\left(\frac{1}{n^2}\right) + \frac{1}{2n^2} + o\left(\frac{1}{n^2}\right) = -\frac{1}{2n^2} + o\left(\frac{1}{n^2}\right) \underset{n \rightarrow +\infty}{\sim} -\frac{1}{2n^2}$.

Et donc, d'après le critère des équivalents pour les séries de signe constant¹, on en déduit que $\sum v_n$ converge.

1.b. Pour $N \geq 2$, on a

$$\sum_{k=1}^{N-1} v_k = \sum_{k=1}^{N-1} \gamma_{k+1} - \gamma_k = \sum_{k=1}^{N-1} \gamma_{k+1} - \sum_{k=1}^{N-1} \gamma_k = \sum_{i=2}^N \gamma_i - \sum_{k=1}^{N-1} \gamma_k = \gamma_N - \gamma_1.$$

Or, par définition d'une série convergente² $\sum_{k=1}^{N-1} v_k \xrightarrow{N \rightarrow +\infty} \sum_{k=1}^{+\infty} v_k$, et donc $\gamma_N \xrightarrow{N \rightarrow +\infty} \sum_{k=1}^{+\infty} v_k + \gamma_1$.

Ainsi $(\gamma_n)_{n \geq 1}$ converge.

1.c. On a

$$d_{n,t} = \gamma + \ln(n) + \ln\left(1 + \frac{t}{n}\right) - h_n = \gamma - \gamma_n + \ln\left(1 + \frac{t}{n}\right) \xrightarrow{n \rightarrow +\infty} \gamma - \gamma + 0 = 0.$$

2.a. On a $E(X_t) = V(X_t) = t$.

2.b. D'après le théorème de transfert, on a, sous réserve de convergence,

$$E(\ln(X_t)) = \int_0^{+\infty} \ln(x) \frac{1}{\Gamma(t)} x^{t-1} e^{-x} dx = \frac{1}{\Gamma(t)} \int_0^{+\infty} \ln(x) x^{t-1} e^{-x} dx = \frac{\Gamma'(t)}{\Gamma(t)} = \psi(t).$$

De même, on a

$$E(\ln(X_t)^2) = \int_0^{+\infty} (\ln x)^2 \frac{1}{\Gamma(t)} x^{t-1} e^{-x} dx = \frac{\Gamma''(t)}{\Gamma(t)}.$$

Et donc, par la formule de Huygens,

$$V(\ln(X_t)) = E(\ln(X_t)^2) - E(\ln(X_t))^2 = \frac{\Gamma''(t)}{\Gamma(t)} - \frac{\Gamma'(t)^2}{\Gamma(t)^2} = \frac{\Gamma''(t)\Gamma(t) - \Gamma'(t)^2}{\Gamma(t)^2}.$$

D'autre part, un calcul de dérivée prouve que

$$\psi'(t) = \frac{\Gamma''(t)\Gamma(t) - \Gamma'(t)^2}{\Gamma(t)^2} = V(\ln(X_t)).$$

3.a. Toujours d'après le théorème de transfert, on a, sous réserve de convergence absolue

$$\begin{aligned} E\left(\frac{1}{X_t}\right) &= \int_0^{+\infty} \frac{1}{x} \frac{1}{\Gamma(t)} x^{t-1} e^{-x} dx \\ &= \frac{1}{\Gamma(t)} \int_0^{+\infty} x^{t-2} e^{-x} dx = \frac{\Gamma(t-1)}{\Gamma(t)} \end{aligned}$$

Rappel

Au voisinage de 0,

$$\ln(1+x) = x - \frac{x^2}{2} + o(x^2)$$

et

$$\frac{1}{1+x} = 1 - x + o(x).$$

De plus, $\frac{1}{n} \xrightarrow{n \rightarrow +\infty} 0$.

$$^1 -\frac{1}{2n^2} \geq 0.$$

² La somme de la série est la limite de la suite des sommes partielles.

Convergence

La convergence de l'intégrale a en fait été donnée par l'énoncé.

On a reconnu des intégrales convergentes, donc l'espérance existe.

$$= \frac{\Gamma(t-1)}{(t-1)\Gamma(t-1)} = \frac{1}{t-1}.$$

Et donc $E\left(\frac{1}{X_t}\right)$ existe et vaut $\frac{1}{t-1}$.

3.b. La fonction \ln est concave et donc située en dessous de ses tangentes. En particulier, la tangente en $t = 1$ est la droite d'équation $y = x - 1$.

Et donc pour tout $x > 0$, $\ln(x) \leq x - 1$.

Pour prouver la seconde inégalité, posons $f(x) = \ln(x) - 1 + \frac{1}{x}$.

Alors f est dérivable sur \mathbf{R}_+ , de dérivée égale à

$$f'(x) = \frac{1}{x} - \frac{1}{x^2} = \frac{x-1}{x^2}.$$

f est donc décroissante sur $]0, 1]$ et croissante sur $[1; +\infty[$. Elle admet donc un minimum en $x = 1$, et $f(1) = 0$.

Par conséquent, pour tout $x > 0$, $f(x) \geq 0 \Leftrightarrow 1 - \frac{1}{x} \leq \ln(x)$.

Si $x \geq 1$, on a $\ln(x) \geq 0$ et donc $(\ln(x))^2 \leq (x-1)^2 \leq \left(1 - \frac{1}{x}\right)^2 + (x-1)^2$.

En revanche, si $x \leq 1$, alors $\ln(x) \leq 0$ et donc, la fonction carré étant décroissante sur \mathbf{R}_- ,

$$\ln(x)^2 \leq \left(1 - \frac{1}{x}\right)^2 \leq \left(1 - \frac{1}{x}\right)^2 + (x-1)^2.$$

Ainsi, pour tout $x > 0$, $\ln(x)^2 \leq \left(1 - \frac{1}{x}\right)^2 + (x-1)^2$.

3.c. Soit $t > 0$. Alors par le théorème de transfert³,

$$\begin{aligned} E\left(\ln\left(\frac{X_t}{t}\right)\right) &= \int_0^{+\infty} \frac{1}{\Gamma(t)} \ln\left(\frac{x}{t}\right) x^{t-1} e^{-x} dx \\ &\leq \int_0^{+\infty} \frac{1}{\Gamma(t)} \left(\frac{x}{t} - 1\right) x^{t-1} e^{-x} dx \\ &\leq \frac{1}{t\Gamma(t)} \int_0^{+\infty} x^t e^{-x} dx - \frac{1}{\Gamma(t)} \int_0^{+\infty} x^{t-1} e^{-x} dx \\ &\leq \frac{\Gamma(t+1)}{t\Gamma(t)} - \frac{\Gamma(t)}{\Gamma(t)} \\ &\leq 1 - 1 = \boxed{0}. \end{aligned}$$

D'autre part, d'après l'inégalité de la question 3.b, $\ln\left(\frac{X_t}{t}\right) \geq 1 - \frac{t}{X_t}$, de sorte que par croissance de l'espérance⁴,

$$E\left(\ln\left(\frac{X_t}{t}\right)\right) \geq 1 - E\left(\frac{t}{X_t}\right) \geq 1 - tE\left(\frac{1}{X_t}\right) \geq 1 - \frac{t}{t-1} \geq \boxed{-\frac{1}{t-1}}.$$

Enfin, d'après la question 3.b,

$$\ln^2\left(\frac{X_t}{t}\right) \leq \left(1 - \frac{t}{X_t}\right)^2 + \left(\frac{X_t}{t} - 1\right)^2 = 2 - 2\frac{t}{X_t} + \frac{t^2}{X_t^2} - 2\frac{X_t}{t} + \frac{X_t^2}{t^2}.$$

Comme à la question 3.a, nous pouvons appliquer le théorème de transfert pour prouver que pour $t > 2$,

$$\begin{aligned} E\left(\frac{1}{X_t^2}\right) &= \int_0^{+\infty} \frac{1}{x^2} \frac{1}{\Gamma(t)} x^{t-1} e^{-x} dx = \frac{1}{\Gamma(t)} \int_0^{+\infty} x^{t-3} e^{-x} dx \\ &= \frac{\Gamma(t-2)}{\Gamma(t)} = \frac{\Gamma(t-2)}{(t-1)\Gamma(t-1)} = \frac{\Gamma(t)}{(t-1)(t-2)\Gamma(t)} = \frac{1}{(t-1)(t-2)}. \end{aligned}$$

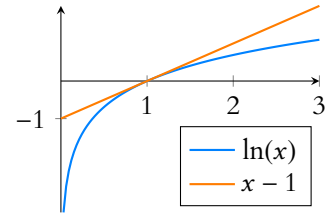


FIGURE 23.1- \ln et sa tangente en $x = 1$.

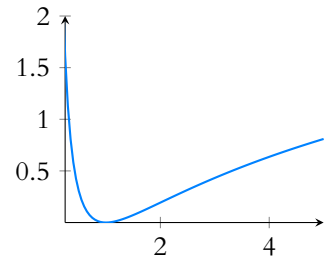


FIGURE 23.2- La fonction f .

³ Nous savons déjà que cette intégrale converge, car $\ln(X_t)$ admet une espérance d'après la question 2.b, et donc il en est de même de $\ln\left(\frac{X_t}{t}\right) = \ln(X_t) - \ln(t)$.

⁴ Qui s'applique puisque $\frac{t}{X_t} = t \times \frac{1}{X_t}$ admet une espérance.

Et donc toutes les variables aléatoires de l'inégalité précédemment prouvée possèdent une espérance, de sorte que, par croissance de l'espérance,

$$\begin{aligned}
 E\left(\ln^2\left(\frac{X_t}{t}\right)\right) &\leq 2 - 2tE\left(\frac{1}{X_t}\right) + t^2E\left(\frac{1}{X_t^2}\right) - \frac{2}{t}E(X_t) + \frac{1}{t^2}E(X_t^2) \\
 &\leq 2 - 2\frac{t}{t-1} + \frac{t^2}{(t-1)(t-2)} - 2 + \frac{V(X_t) + E(X_t)^2}{t^2} \\
 &\leq \frac{t^2 - 2t(t-2)}{(t-1)(t-2)} + \frac{t+t^2}{t^2} \\
 &\leq \frac{-t^2 + 4t}{(t-1)(t-2)} + \frac{1+t}{t} \\
 &\leq \frac{-t^3 + 4t^2 + (1+t)(t-1)(t-2)}{t(t-1)(t-2)} \\
 &\leq \frac{-t^3 + 4t^2 + t^3 - 2t^2 - t + 2}{t(t-1)(t-2)} \\
 &\leq \frac{2t^2 - t + 2}{t(t-1)(t-2)} \\
 &\leq \frac{2t^2}{t(t-1)(t-2)} \\
 &\leq \frac{2t}{(t-2)^2}.
 \end{aligned}$$

Formule de Huygens.

$t > 2$ donc $2 - t \leq 0$.

$t - 2 < t - 1$.

3.d. Par l'inégalité de Markov, appliquée à la variable aléatoire $\ln^2\left(\frac{X_{t+n}}{t+n}\right)$, qui est positive⁵ et possède une espérance, pour tout $\varepsilon > 0$, on a

$$P\left(\left|\ln\left(\frac{X_{t+n}}{t+n}\right)\right| > \varepsilon\right) = P\left(\ln^2\left(\frac{X_{t+n}}{t+n}\right) > \varepsilon^2\right) \leq \frac{1}{\varepsilon^2}E\left(\ln^2\left(\frac{X_{t+n}}{t+n}\right)\right) \leq \frac{2(t+n)}{\varepsilon^2(t+n-2)^2}.$$

⁵ Hypothèse indispensable pour appliquer Markov !

Or, $\frac{t+n}{(t+n-2)^2} \underset{n \rightarrow +\infty}{\sim} \frac{1}{n} \underset{n \rightarrow +\infty}{\rightarrow} 0$ et donc par le théorème des gendarmes,

$$\lim_{n \rightarrow +\infty} P\left(\left|\ln\left(\frac{X_{t+n}}{t+n}\right)\right| > \varepsilon\right) = 0$$

ce qui prouve que $\ln\left(\frac{X_{t+n}}{t+n}\right) \xrightarrow{P} 0$.

4.a. Travaillons avec les événements contraires : si l'événement $[|A_n| \leq \varepsilon/3] \cap [|B_n| \leq \varepsilon/3] \cap [|C_n| \leq \varepsilon/3]$ est réalisé, alors, par l'inégalité triangulaire,

$$|D_n| = |A_n + B_n + C_n| \leq |A_n| + |B_n| + |C_n| \leq \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} \leq \varepsilon.$$

Autrement dit, on a prouvé que

$$[|A_n| \leq \varepsilon/3] \cap [|B_n| \leq \varepsilon/3] \cap [|C_n| \leq \varepsilon/3] \subset [|D_n| \leq \varepsilon].$$

En passant aux événements contraires on a donc

$$\overline{[|D_n| \leq \varepsilon]} \subset \overline{[|A_n| \leq \varepsilon/3] \cap [|B_n| \leq \varepsilon/3] \cap [|C_n| \leq \varepsilon/3]}.$$

Soit encore

$$[|D_n| > \varepsilon] \subset [|A_n| > \varepsilon] \cup [|B_n| > \varepsilon] \cup [|C_n| > \varepsilon].$$

On en déduit que

$$P(|D_n| > \varepsilon) \leq P(|A_n| > \varepsilon) + P(|B_n| > \varepsilon) + P(|C_n| > \varepsilon).$$

Mais puisque (A_n) , (B_n) et (C_n) convergent les trois en probabilités vers 0, il vient donc,

par le théorème des gendarmes, $\lim_{n \rightarrow +\infty} P(|D_n| > \varepsilon) = 0 \Leftrightarrow D_n \xrightarrow{P} 0$.

Rappel
Si $A \subset B$, alors $\bar{B} \subset \bar{A}$.

Rappel
 $A \cap \bar{B} = \bar{A} \cup \bar{B}$.

4.b. Soit $\varepsilon > 0$. Alors, comme à la question précédente, on a

$$0 \leq P(|D_n| > \varepsilon) \leq P\left(|D_n| > \frac{\varepsilon}{2}\right) + P\left(|u_n - u| > \frac{\varepsilon}{2}\right).$$

Mais $u_n \xrightarrow[n \rightarrow +\infty]{} u$, et donc il existe $N \in \mathbf{N}$ tel que pour $n \geq N$, $|u_n - u| \leq \frac{\varepsilon}{2}$.

Et donc en particulier, pour $n \geq N$, l'événement $\left[|u_n - u| > \frac{\varepsilon}{2}\right]$ est l'événement vide (\emptyset), et donc est de probabilité nulle. Ainsi, pour $n \geq N$, on a

$$0 \leq P(|D_n + u_n - u| > \varepsilon) \leq P\left(|D_n| > \frac{\varepsilon}{2}\right).$$

Et donc par le théorème des gendarmes, $\lim_{n \rightarrow +\infty} P(|D_n + u_n - u| > \varepsilon) \xrightarrow[n \rightarrow +\infty]{} 0$.

Et donc on a bien $V_n \xrightarrow{P} 0$.

On en déduit que pour tout $\varepsilon > 0$,

$$P(|(N + u + V_n) - (N + u)| > \varepsilon) = P(|V_n| > \varepsilon) \xrightarrow[n \rightarrow +\infty]{} 0.$$

Et donc $N + u + V_n \xrightarrow{P} N + u$.

4.c. Notons que $N + D_n + u = N + u + V_n$. Et donc $N + D_n + u \xrightarrow{P} N + u$.

Et donc, d'après le résultat admis par l'énoncé, $N + D_n + u \xrightarrow{\mathcal{L}} N + u$.

Puisque la convergence en loi ne dépend que de la loi, et que $N + D_n + u$ et M ont même loi, on en déduit que $M \xrightarrow{\mathcal{L}} N + u$.

Or, M ne dépend pas de n ! Plus précisément : la convergence en loi signifie qu'en tout point de continuité de F_{N+u} , on a $F_M(x) = F_{N+u}(x)$.

Mais $N + u$ est une variable à densité, donc sa fonction de répartition est continue sur \mathbf{R} tout entier, et donc pour tout $x \in \mathbf{R}$, $F_M(x) = F_{N+u}(x)$: M et $N + u$ ont même loi.

5.a. Puisque X_α et X_β ne prennent que des valeurs strictement positives, il en est de même de $T_{\alpha,\beta}$. Et donc $Q_{\alpha,\beta}$ est bien définie, à valeurs dans \mathbf{R} .

Pour $x \in \mathbf{R}$, on a, par croissance de la fonction exponentielle :

$$P(\ln(X_\alpha) \leq x) = P(X_\alpha \leq e^x) = F_{X_\alpha}(e^x).$$

Mais puisque X_α est une variable à densité, sa fonction de répartition F_{X_α} est continue sur \mathbf{R} et \mathcal{C}^1 sauf en un nombre fini de points. Par composition par la fonction exponentielle, qui est \mathcal{C}^1 sur \mathbf{R} , il en est donc de même de la fonction de répartition de $\ln(X_\alpha)$.

Ainsi, X_α est une variable à densité, et une densité en est donnée par

$$f_\alpha : x \mapsto e^x F'_{X_\alpha}(e^x) = \frac{1}{\Gamma(\alpha)} e^x (e^x)^{\alpha-1} e^{-e^x} = \frac{1}{\Gamma(\alpha)} e^{\alpha x - e^x}.$$

De même, une densité de $\ln(X_\beta)$ est $x \mapsto \frac{1}{\Gamma(\beta)} e^{\beta x - e^x}$.

Par transformation affine, une densité de $-\ln(X_\beta)$ est donc

$$f_\beta : x \mapsto \frac{1}{\Gamma(\beta)} e^{-\beta x - e^{-x}}.$$

5.b. On a $Q_{\alpha,\beta} = \ln(X_\alpha) - \ln(X_\beta)$.

Or, $\ln(X_\alpha)$ et $-\ln(X_\beta)$ sont indépendantes car X_α et X_β le sont, et f_β est bornée par $\frac{1}{\Gamma(\beta)}$

de sorte que, par produit de convolution, $Q_{\alpha,\beta}$ est une variable à densité dont une densité est donnée par

$$f_{Q_{\alpha,\beta}}(x) = \int_{-\infty}^{+\infty} f_\alpha(t) f_\beta(x-t) dt = \int_{-\infty}^{+\infty} \frac{1}{\Gamma(\alpha)} e^{\alpha t - e^t} \frac{1}{\Gamma(\beta)} e^{-\beta(x-t) - e^{-(x-t)}} dt$$

Remarque

Nous parlons ici d'un événement qui n'a rien d'aléatoire : il ne contient aucune variable aléatoire. Et donc soit il vaut Ω (si l'événement est réalisé), soit il vaut \emptyset .

Remarque

La même preuve montre que si $X_n \xrightarrow{P} 0$, alors pour toute variable aléatoire fixée N , $X_n + N \xrightarrow{P} N$.

CV en loi

La convergence en loi est définie uniquement en termes de fonctions de répartitions, et donc ne dépend que des lois et non des variables.

Remarque

Le fait que X_α possède une densité continue sauf peut-être (c'est le cas si et seulement si $\alpha > 1$) en 0 indique que F_{X_α} est \mathcal{C}^1 , sauf peut-être en 0. Et donc la fonction de répartition de $\ln(X_\alpha)$ est \mathcal{C}^1 , sauf peut-être en $1 = e^0$.

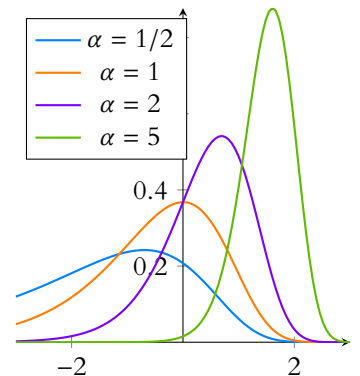


FIGURE 23.3— La densité f_α .

$$\begin{aligned}
&= \frac{e^{-\beta x}}{\Gamma(\alpha)\Gamma(\beta)} \int_{-\infty}^{+\infty} e^{(\alpha+\beta)t} e^{-e^t - e^{-t} e^{-x}} dt \\
&= \frac{e^{-\beta x}}{\Gamma(\alpha)\Gamma(\beta)} \int_{-\infty}^{+\infty} e^{(\alpha+\beta)t} e^{-e^t(1+e^{-x})} dt \\
&= \frac{e^{-\beta x}}{\Gamma(\alpha)\Gamma(\beta)} \int_{-\infty}^{+\infty} e^{(\alpha+\beta)y} e^{-e^y(1+e^{-x})} dy.
\end{aligned}$$

Remarque

Nous avons seulement changé la variable muette t en y pour obtenir le même résultat que dans l'énoncé.

- 5.c. Le changement de variable proposé est \mathcal{C}^1 et strictement croissant, donc légitime. Lorsque $y \rightarrow -\infty$, $u \rightarrow 0$ et lorsque $y \rightarrow +\infty$, $u \rightarrow +\infty$. De plus, on a

$$u = e^y(1 + e^{-x}) \Leftrightarrow e^y = \frac{u}{1 + e^{-x}} \Leftrightarrow y = \ln\left(\frac{u}{1 + e^{-x}}\right).$$

Et alors $du = e^y(1 + e^{-x})dy = u dy \Leftrightarrow dy = \frac{du}{u}$.

Ainsi, il vient

$$\begin{aligned}
f_{Q_{\alpha,\beta}}(x) &= \frac{e^{-\beta x}}{\Gamma(\alpha)\Gamma(\beta)} \int_0^{+\infty} \left(\frac{u}{1 + e^{-x}}\right)^{\alpha+\beta} e^{-u} \frac{du}{u} \\
&= \frac{e^{-\beta x}}{\Gamma(\alpha)\Gamma(\beta)(1 + e^{-x})^{\alpha+\beta}} \int_0^{+\infty} u^{\alpha+\beta-1} e^{-u} du \\
&= \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} \frac{e^{-\beta x}}{(1 + e^{-x})^{\alpha+\beta}} \\
&= \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} \frac{e^{\alpha x}}{(1 + e^x)^{\alpha+\beta}} \\
&= \boxed{\frac{1}{B(\alpha, \beta)} \frac{e^{\alpha x}}{(1 + e^x)^{\alpha+\beta}}}.
\end{aligned}$$

On a multiplié numérateur et dénominateur par $e^{(\alpha+\beta)x}$.

- 5.d. Notons que $T_{\alpha,\beta}$ ne prend que des valeurs positives, et donc pour $x < 0$, $P(T_{\alpha,\beta} \leq x) = 0$. De plus, $T_{\alpha,\beta} = 0$ si et seulement si $X_\alpha = 0$ et donc

$$P(T_{\alpha,\beta} \leq 0) = P(T_{\alpha,\beta} = 0) = P(X_\alpha = 0) = 0.$$

Soit à présent $x > 0$. Alors, par croissance de la fonction logarithme,

$$F_{T_{\alpha,\beta}}(x) = P(T_{\alpha,\beta} \leq x) = P(Q_{\alpha,\beta} \leq \ln(x)).$$

Mais $Q_{\alpha,\beta}$ est une variable aléatoire à densité, et l'une de ses densités, $f_{Q_{\alpha,\beta}}$ est continue sur \mathbf{R} , donc sa fonction de répartition est \mathcal{C}^1 sur \mathbf{R} .

Par composition avec la fonction logarithme, $F_{T_{\alpha,\beta}}$ est \mathcal{C}^1 sur \mathbf{R}_+^* .

De plus, on a

$$\lim_{x \rightarrow 0^+} F_{T_{\alpha,\beta}}(x) = \lim_{x \rightarrow 0^+} F_{Q_{\alpha,\beta}}(\ln(x)) = \lim_{y \rightarrow -\infty} F_{Q_{\alpha,\beta}}(y) = 0 = F_{T_{\alpha,\beta}}(0).$$

Et $F_{T_{\alpha,\beta}}$ étant constante sur \mathbf{R}_- , elle y est continue⁶ et même \mathcal{C}^1 .

Ainsi, $F_{T_{\alpha,\beta}}$ est continue sur \mathbf{R} et \mathcal{C}^1 sauf peut-être en 0.

Donc $T_{\alpha,\beta}$ est une variable à densité.

Une densité en est alors toute fonction qui coïncide avec $F'_{T_{\alpha,\beta}}$ sur \mathbf{R}^* . Par exemple, on peut prendre

$$f_{T_{\alpha,\beta}}(x) = \begin{cases} 0 & \text{si } x \leq 0 \\ \frac{1}{x} f_{Q_{\alpha,\beta}}(\ln x) & \text{si } x > 0 \end{cases} = \boxed{\begin{cases} 0 & \text{si } x \leq 0 \\ \frac{1}{B(\alpha, \beta)} \frac{x^{\alpha-1}}{(1+x)^{\alpha+\beta}} & \text{si } x > 0 \end{cases}}$$

- 5.e. Commençons par noter que $J_{\alpha,\beta}$ ne prend que des valeurs positives, et donc pour $x < 0$, $P(J_{\alpha,\beta} \leq x) = 0$. De plus, $J_{\alpha,\beta} = 0$ si et seulement si $X_\alpha = 0$, et donc

$$P(J_{\alpha,\beta} \leq 0) = P(J_{\alpha,\beta} = 0) = P(X_\alpha = 0) = 0.$$

 \mathcal{C}^1

Une variable de densité f_X possède

$$F_X : x \mapsto \int_{-\infty}^x f_X(t) dt$$

pour fonction de répartition. Si f_X est continue, alors par le théorème fondamental de l'analyse, F_X est \mathcal{C}^1 , de dérivée f_X .

⁶ En particulier, elle est continue à gauche en 0.

Détails

Rappelons que $F_{Q_{\alpha,\beta}}$ possède $f_{Q_{\alpha,\beta}}$ pour dérivée, comme indiqué plus haut.

D'autre part, $X_\alpha \leq X_\alpha + X_\beta$ et donc $J_{\alpha,\beta} \leq 1$. Donc pour $x \geq 1$, $P(J_{\alpha,\beta} \leq x) = 1$.
Soit à présent $x \in]0, 1[$. On a

$$\begin{aligned} F_{J_{\alpha,\beta}}(x) &= P(J_{\alpha,\beta} \leq x) = P\left(\frac{X_\alpha}{X_\alpha + X_\beta} \leq x\right) \\ &= P\left(\frac{X_\alpha + X_\beta}{X_\alpha} \geq \frac{1}{x}\right) \\ &= P\left(1 + \frac{X_\beta}{X_\alpha} \geq \frac{1}{x}\right) \\ &= P\left(\frac{X_\beta}{X_\alpha} \geq \frac{1-x}{x}\right) = P\left(\frac{X_\alpha}{X_\beta} \leq \frac{x}{1-x}\right) \\ &= P\left(T_{\alpha,\beta} \leq \frac{x}{1-x}\right) = F_{T_{\alpha,\beta}}\left(\frac{x}{1-x}\right). \end{aligned}$$

La fonction inverse est décroissante sur \mathbf{R}_+^* .

Il est donc évident⁷ que la fonction de répartition de $J_{\alpha,\beta}$ est \mathcal{C}^1 sauf éventuellement en 0 et en 1.

De plus, lorsque $x \rightarrow 0^+$, $\frac{x}{1-x} \xrightarrow{x \rightarrow 0^+} 0$ et donc $F_{J_{\alpha,\beta}}(x) \xrightarrow{x \rightarrow 0^+} 0$.

Et de même, lorsque $x \rightarrow 1^-$, $\frac{x}{1-x} \xrightarrow{x \rightarrow 1^-} +\infty$ et donc $F_{J_{\alpha,\beta}}\left(\frac{x}{1-x}\right) \xrightarrow{x \rightarrow 1^-} 1$.

On en déduit donc que $F_{J_{\alpha,\beta}}$ est continue sur \mathbf{R} , \mathcal{C}^1 sauf en 0 et en 1, et donc $J_{\alpha,\beta}$ est une variable à densité.

Pour $x \in]0, 1[$, par dérivation d'une composée, on a alors

$$F'_{J_{\alpha,\beta}}(x) = \frac{1}{(1-x)^2} \frac{1}{B(\alpha,\beta)} \left(\frac{x}{1-x}\right)^{\alpha-1} (1-x)^{\alpha+\beta} = \frac{1}{B(\alpha,\beta)} x^{\alpha-1} (1-x)^{\beta-1}.$$

Et donc une densité de $J_{\alpha,\beta}$ est :

$$f_{J_{\alpha,\beta}}(x) = \begin{cases} \frac{1}{B(\alpha,\beta)} x^{\alpha-1} (1-x)^{\beta-1} & \text{si } x \in]0, 1[\\ 0 & \text{sinon} \end{cases}$$

⁷ Car la fonction de répartition de $F_{T_{\alpha,\beta}}$ est \mathcal{C}^1 sauf en 0.

Partie II. Étude de la variable aléatoire $\ln(X_t)$.

6. Puisque la loi $\mathcal{E}(1)$ n'est autre que la loi $\gamma(1)$, alors, par stabilité⁸ des lois γ , $S_k \hookrightarrow \gamma(k)$.
Et donc en particulier $E(S_k) = k$ et $V(S_k) = k$.

⁸ Qui s'applique car les Y_i sont indépendantes.

7.a. Il s'agit de remarquer que le produit dans le logarithme est un produit télescopique :

$$\prod_{k=1}^n \frac{X_t + S_{k-1}}{X_t + S_k} = \frac{X_t}{X_t + S_1} \frac{X_t + S_1}{X_t + S_2} \dots \frac{X_t + S_{n-1}}{X_t + S_n} = \frac{X_t}{X_t + S_n}.$$

Et donc

$$\ln\left(\prod_{k=1}^n \frac{X_t + S_{k-1}}{X_t + S_k}\right) + \ln(X_t + S_n) = \ln\left(\frac{X_t}{X_t + S_n}\right) + \ln(X_t + S_n) = \ln(X_t).$$

7.b. Puisque X_t est indépendante de Y_1, \dots, Y_m , par le lemme des coalitions, elle est indépendante de S_m .

Et donc, par stabilité des lois gamma, $X_t + S_m$ suit la loi $\gamma(t+m)$, et donc $\boxed{\text{a même loi que } X_{t+m}}$.

8.a. On a $R_{t,1} = \frac{X_t}{X_t + S_1}$ et $R_{t,2} = \frac{X_t + S_1}{X_t + S_2}$.

D'après le résultat admis fourni par l'énoncé⁹, nous pouvons déjà affirmer que $\frac{X_t}{X_t + S_1}$ et $X_t + S_1$ sont indépendantes.

⁹ Qui s'applique bien puisque X_t et S_1 suivent toutes deux des lois gamma.

Puisque de plus, X_t, Y_1 et Y_2 sont mutuellement indépendantes, alors Y_2 et $\frac{X_t}{X_t + Y_1}$ sont indépendantes.

Ainsi, $\frac{X_t}{X_t + Y_1}$ et $X_t + S_2 = (X_t + Y_1) + Y_2$ sont indépendantes.

Et donc, toujours par le lemme des coalitions $\frac{X_t}{X_t + Y_1} = R_{t,1}$ et $\frac{X_t + Y_1}{X_t + S_2} = R_{t,2}$ sont indépendantes.

8.b. On a

$$\sum_{k=1}^n \ln(R_{t,k}) = \ln\left(\prod_{k=1}^n R_{t,k}\right) = \ln\left(\prod_{k=1}^n \frac{X_t + S_{k-1}}{X_t + S_k}\right) = \ln\left(\frac{X_t}{X_t + S_n}\right).$$

Mais $\prod_{k=1}^n R_{t,k} = \frac{X_t}{X_t + S_n}$, qui d'après le résultat admis, est indépendante de $X_t + S_n$.

De plus, $X_t + S_n$ a même loi que X_{t+n} , qui est elle-même indépendante de $\frac{X_t}{X_t + S_n}$ car indépendante de X_t et des Y_i .

Ainsi, les couples $\left(\frac{X_t}{X_t + S_n}, X_{t+n}\right)$ et $\left(\frac{X_t}{X_t + S_n}, X_t + S_n\right)$ ont même loi.

Ainsi, en appliquant la fonction $(x, y) \mapsto \ln(x) + \ln(y)$, continue sur $\mathbf{R}_+^* \times \mathbf{R}_+^*$, les variables $\ln\left(\frac{X_t}{X_t + S_n}\right) + \ln(X_t + S_n)$ et $\ln\left(\frac{X_t}{X_t + S_n}\right) + \ln(X_{t+n})$ ont même loi.

Et puisque $\ln\left(\frac{X_t}{X_t + S_n}\right) = \sum_{k=1}^n \ln(R_{t,k})$, $\sum_{k=1}^n \ln(R_{t,k}) + \ln(X_t + S_n)$ et $\sum_{k=1}^n \ln(R_{t,k}) + \ln(X_{t+n})$ ont même loi.

D'autre part, nous savons, d'après la question 7.a que

$$\ln(X_t) = \ln\left(\prod_{k=1}^n \frac{X_t + S_{k-1}}{X_t + S_k}\right) + \ln(X_t + S_n) = \ln\left(\prod_{k=1}^n R_{t,k}\right) + \ln(X_t + S_n) = \sum_{k=1}^n \ln(R_{t,k}) + \ln(X_t + S_n).$$

Donc $\ln(X_t)$ et $\sum_{k=1}^n \ln(X_{t+n})$ ont bien même loi.

9.a. Puisque U est à valeurs dans $]0, 1]$, $-\ln(U)$ est à valeurs dans \mathbf{R}_+ .

Donc déjà, pour $x < 0$, $P(-\ln(U) \leq x) = 0$.

Pour $x \geq 0$, on a, par croissance de la fonction exponentielle,

$$P(-\ln(U) \leq x) = P(U \geq e^{-x}) = 1 - P(U \leq e^{-x}) = 1 - e^{-x}.$$

Et donc $-\ln(U)$ suit la loi exponentielle de paramètre 1.

9.b. On a $R_{t,k} = \frac{X_t + S_{k-1}}{X_t + S_k} = \frac{X_t + S_{k-1}}{X_t + S_{k-1} + Y_k}$, et par le lemme des coalitions, $X_t + S_{k-1}$ et Y_k sont indépendantes.

Par la question 7.b, $X_t + S_{k-1}$ suit la loi $\gamma(t+k-1)$, et par définition, Y_k suit la loi $\mathcal{E}(1) = \gamma(1)$.

Et donc d'après le résultat de la question 5.e, une densité de $R_{t,k}$ est

$$f_{R_{t,k}} : z \mapsto \begin{cases} 0 & \text{si } z \notin [0, 1] \\ \frac{1}{B(t+k-1, 1)} z^{t+k-2} & \text{si } z \in [0, 1] \end{cases}$$

$$\text{Mais } B(t+k-1, 1) = \frac{\Gamma(t+k-1)\Gamma(1)}{\Gamma(t+k)} = \frac{\Gamma(t+k-1)}{\Gamma(t+k)} = \frac{\Gamma(t+k-1)}{(t+k-1)\Gamma(t+k-1)} = \frac{1}{t+k-1}.$$

On a donc

$$f_{R_{t,k}}(z) = \begin{cases} (t+k-1)z^{t+k-2} & \text{si } z \in [0, 1] \\ 0 & \text{sinon} \end{cases}$$

9.c. Puisque U_k est à valeurs dans $[0, 1]$, il en est de même de $U_k^{\frac{1}{t+k-1}}$.

Donc pour $z < 0$, $P\left(U_k^{\frac{1}{t+k-1}} \leq z\right) = 0$ et pour $z \geq 1$, $P\left(U_k^{\frac{1}{t+k-1}} \leq z\right) = 1$.

Soit donc à présent $z \in [0, 1]$. Alors

$$P\left(U_k^{\frac{1}{t+k-1}} \leq z\right) = P\left(U_k \leq z^{t+k-1}\right) = z^{t+k-1}.$$

Il est aisé de constater que la fonction de répartition de $U_k^{\frac{1}{t+k-1}}$ est alors continue en 0 et en 1, et donc sur \mathbf{R} , et qu'elle est de classe \mathcal{C}^1 sauf peut-être en 0 et en 1.

Détails

Il ne suffit pas de remarquer que ces deux couples ont les mêmes lois marginales : l'indépendance des deux variables est importante, c'est grâce à elle que la loi du couple est caractérisée par les lois marginales.

Donc $U_k^{\frac{1}{t+k-1}}$ est une variable à densité, et l'une de ses densités est donnée par

$$z \mapsto \begin{cases} 0 & \text{si } z \notin [0, 1] \\ (t+k-1)z^{t+k-2} & \text{si } z \in [0, 1] \end{cases}$$

Et donc $U_k^{\frac{1}{t+k-1}}$ et $R_{t,k}$ sont de même loi.

9.d. On a

$$\begin{aligned} \sum_{k=0}^{n-1} \left(\frac{1}{k+1} - \frac{Y_{k+1}}{t+k} \right) + \ln \left(\frac{X_{t+n}}{t+n} \right) + d_{n,t} - \gamma &= \underbrace{\sum_{k=1}^n \frac{1}{k}}_{=h_n} - \sum_{k=1}^n \frac{Y_k}{t+k-1} + \ln \left(\frac{X_{t+n}}{t+n} \right) + \ln(t+n) - h_n \\ &= - \sum_{k=1}^n \frac{Y_k}{t+k-1} + \ln(X_{t+n}). \end{aligned}$$

D'après la question 9.c, $\ln(R_{t,k})$ a même loi que $\ln \left(U_k^{\frac{1}{t+k-1}} \right) = -\frac{1}{t+k-1} (-\ln U_k)$.

Mais d'après 9.a, $-\ln(U_k)$ suit la loi exponentielle $\mathcal{E}(1)$, c'est-à-dire la même loi que Y_k .

Donc $-\frac{Y_k}{t+k-1}$ a même loi que $\ln(R_{t,k})$.

Et puisque les Y_k sont mutuellement indépendantes, et indépendantes de X_{t+n} , de même que les $R_{t,k}$ sont indépendantes et indépendantes de X_{t+n} , on en déduit que

$\sum_{k=1}^n -\frac{Y_k}{t+k-1} + \ln(X_{t+n})$ et $\sum_{k=1}^n \ln(R_{t,k}) + \ln(X_{t+n})$ ont même loi.

Et donc $\ln(X_{t+n})$ a même loi que $\sum_{k=0}^{n-1} \left(\frac{1}{k+1} - \frac{Y_{k+1}}{t+k} \right) + \ln \left(\frac{X_{t+n}}{t+n} \right) + d_{n,t} - \gamma$.

10. D'après 9.d, et par linéarité de l'espérance,

$$\begin{aligned} E(\ln(X_t)) &= \sum_{k=0}^{n-1} \left(\frac{1}{k+1} - \frac{1}{t+k} E(Y_{k+1}) \right) + E \left(\ln \left(\frac{X_{t+n}}{t+n} \right) \right) + d_{n,t} - \gamma \\ &= -\gamma + \sum_{k=0}^{n-1} \left(\frac{1}{k+1} - \frac{1}{t+k} \right) + E \left(\ln \left(\frac{X_{t+n}}{t+n} \right) \right) + d_{n,t}. \quad (\star) \end{aligned}$$

Mais nous savons d'après la question 1.c que $d_{n,t} \xrightarrow{n \rightarrow +\infty} 0$ et les encadrements de la question

3.c prouvent que $E \left(\ln \left(\frac{X_{t+n}}{t+n} \right) \right) \xrightarrow{n \rightarrow +\infty} 0$.

Et donc, en passant à la limite¹⁰ lorsque $n \rightarrow +\infty$, il vient donc

$$\psi(t) = E(\ln(X_t)) = -\gamma + \sum_{k=0}^{+\infty} \left(\frac{1}{k+1} - \frac{1}{t+k} \right).$$

De plus, les Y_k étant mutuellement indépendantes, et indépendantes de X_{t+n} , il vient

$$\begin{aligned} V(\ln(X_t)) &= \sum_{k=0}^{n-1} V \left(\frac{1}{k+1} - \frac{Y_{k+1}}{t+k} \right) + V \left(\ln \left(\frac{X_{t+n}}{t+n} \right) \right) \\ &= \sum_{k=0}^{n-1} \frac{1}{(t+k)^2} + V \left(\ln \left(\frac{X_{t+n}}{t+n} \right) \right). \end{aligned}$$

De plus, nous avons prouvé à la question 3.c que $V \left(\ln \left(\frac{X_{t+n}}{t+n} \right) \right) \xrightarrow{n \rightarrow +\infty} 0$.

Et donc en passant à la limite, $\psi'(t) = V(\ln(X_t)) = \sum_{k=0}^{+\infty} \frac{1}{(k+1)^2}$.

11. La question 3.c nous a prouvé que $E \left(\ln \left(\frac{X_t}{t} \right) \right) \xrightarrow{t \rightarrow +\infty} 0$.

Mais $E \left(\ln \left(\frac{X_t}{t} \right) \right) = E(\ln(X_t)) - \ln(t) = \psi(t) - \ln(t)$.

Et donc $\lim_{t \rightarrow +\infty} \psi(t) - \ln(t) = 0$.

Convergence

Au premier abord, il faudrait peut-être prouver la convergence de cette série (ce qui n'est pas très difficile). Mais l'égalité (\star) permet d'écrire les sommes partielles de cette série comme combinaison linéaire de termes admettant tous une limite lorsque $n \rightarrow +\infty$, et donc, automatiquement, ces sommes partielles admettent une limite, et donc la série converge.

12. On a $\ln(W) = \frac{1}{\mu} \ln(U)$. Or, nous avons prouvé à la question 9.a que $-\ln(U) \leftrightarrow \mathcal{E}(1)$.

Donc par stabilité¹¹ des lois γ , $-\mu \sum_{i=1}^p \ln W_i \leftrightarrow \gamma(p)$.

¹¹ Qui s'applique car les W_i sont indépendantes.

Autrement dit, $-\mu \sum_{i=1}^p \ln W_i$ a même loi que X_p . Et donc $G_p = \mu p \frac{1}{-\mu \sum_{i=1}^p \ln W_i}$ a même loi

que $\mu p \frac{1}{X_p}$.

Et donc en réutilisant le résultat de la question 3.a :

$$E(G_p) = \mu p E\left(\frac{1}{X_p}\right) = \frac{\mu p}{p-1}.$$

Et donc $E(G_p) \neq \mu$: G_p n'est pas un estimateur sans biais de μ .

Biais
Notons en revanche que G_p est asymptotiquement sans biais.

De même, en réutilisant un calcul de la question 3.c,

$$E(G_p^2) = (\mu p)^2 E\left(\frac{1}{X_p^2}\right) = \frac{\mu^2 p^2}{(p-1)(p-2)}.$$

Par la formule de Huygens, il vient alors

$$V(G_p) = \frac{\mu^2 p^2}{(p-1)(p-2)} - \frac{\mu^2 p^2}{(p-1)^2} = \frac{\mu^2 p^2}{(p-1)^2(p-2)}.$$

Ainsi, le risque quadratique de G_p , en tant qu'estimateur de μ vaut

$$\begin{aligned} r(G_p) &= b(G_p)^2 + V(G_p) \\ &= \left(\mu - \mu \frac{p}{p-1}\right)^2 + \frac{\mu^2 p^2}{(p-1)^2(p-2)} \\ &= \frac{\mu^2}{(p-1)^2} + \frac{\mu^2 p^2}{(p-1)^2(p-2)} \\ &= \frac{\mu^2(p^2 + p - 2)}{(p-1)(p-2)} \xrightarrow{p \rightarrow +\infty} 0. \end{aligned}$$

Et donc G_p est un estimateur convergent de μ .

Rappel
Un estimateur dont le risque tend vers 0 est convergent.

- 13.a. C'est un grand classique : cette fonction simule une loi exponentielle de paramètre λ . Pour le prouver, le plus simple serait probablement de déterminer la fonction de répartition de $-\frac{1}{\lambda} \ln(1-U)$, où $U \leftrightarrow \mathcal{U}([0, 1])$.

Remarquons plutôt que $1-U$ suit également la loi $\mathcal{U}([0, 1])$.

Et donc par 9.a, $-\ln(1-U) \leftrightarrow \mathcal{E}(1)$, de sorte que $-\frac{1}{\lambda} \ln(1-U) \leftrightarrow \mathcal{E}(\lambda)$.

- 13.b. Il s'agit de se rappeler que la somme de n lois $\gamma(1)$ indépendantes est une loi $\gamma(n)$, et que la loi $\gamma(1)$ n'est autre que la loi $\mathcal{E}(1)$. Et donc il suffit de simuler la somme de n lois $\mathcal{E}(1)$.

Rappel
Si $X \leftrightarrow \mathcal{E}(1)$, alors $\frac{X}{\lambda} \leftrightarrow \mathcal{E}(\lambda)$.

```

1  function y = g(n)
2      y = 0
3      for i=1 :n
4          y = y - log(rand())
5      end
6  endfunction

```

- 13.c. Le résultat de la question 12 est valable pour tout $\mu > 0$, et donc en particulier pour $\mu = 1$:

$$\frac{p}{-\sum_{i=1}^p \ln(W_i)} \xrightarrow{p \rightarrow +\infty} 1.$$

Mais les $-\ln(W_i)$ sont indépendantes, suivent la loi $\mathcal{E}(1)$, et donc leur somme suit la loi $\gamma(p)$.
Comme le résultat de la commande $g(p)$.
Et donc pour p grand, le résultat de la commande $g(p)$ doit être proche de 1.

Partie III. Quelques propriétés de la fonction Γ .

14. Premières applications : les formules de Wilks et Legendre.

14.a. On a

$$\begin{aligned} \sum_{k=0}^{n-1} \frac{1}{k+1} + \sum_{k=0}^{n-1} \left(\frac{1}{2k+1} - \frac{1}{2k+2} \right) &= \sum_{k=0}^{n-1} \frac{2}{2k+2} + \sum_{k=0}^{n-1} \frac{1}{2k+1} - \sum_{k=0}^{n-1} \frac{1}{2k+2} \\ &= \sum_{k=0}^{n-1} \frac{1}{2k+1} + \sum_{k=0}^{n-1} \frac{1}{2k+2} \\ &= \sum_{k=0}^{2n-1} \frac{1}{k+1}. \end{aligned}$$

D'autre part, en séparant les termes pairs des termes impairs, on a

$$\begin{aligned} 2 \sum_{k=0}^{2n-1} \frac{a_{k+1}}{t+k} &= 2 \sum_{k=0}^{n-1} \frac{a_{2k+2}}{2k+1+t} + 2 \sum_{k=0}^{n-1} \frac{a_{2k+1}}{2k+t} \\ &= \sum_{k=0}^{n-1} \frac{a_{2k+2}}{\frac{t+1}{2} + k} + \sum_{k=0}^{n-1} \frac{a_{2k+1}}{\frac{t}{2} + k}. \end{aligned}$$

Et donc en sommant les deux relations obtenues, il vient bien

$$2 \sum_{k=0}^{2n-1} \left(\frac{1}{k+1} - \frac{a_{k+1}}{t+k} \right) = \sum_{k=0}^{n-1} \left(\frac{1}{k+1} - \frac{a_{2k+1}}{\frac{t}{2} + k} \right) + \sum_{k=0}^{n-1} \left(\frac{1}{k+1} - \frac{a_{2k+2}}{\frac{t+1}{2} + k} \right) + 2 \sum_{k=0}^{n-1} \left(\frac{1}{2k+1} - \frac{1}{2k+2} \right).$$

14.b. On a

$$\begin{aligned} w_n &= \sum_{k=0}^{n-1} \left(\frac{1}{2k+1} - \frac{1}{2k+2} \right) \\ &= \sum_{k=0}^{n-1} \frac{1}{2k+1} + \sum_{k=0}^{n-1} \frac{1}{2k+2} - 2 \sum_{k=0}^{n-1} \frac{1}{2k+2} \\ &= \sum_{k=0}^{2n-1} \frac{1}{k+1} - \sum_{k=0}^{n-1} \frac{1}{k+1} \\ &= \sum_{k=1}^{2n} \frac{1}{k} - \sum_{k=1}^n \frac{1}{k} \\ &= \boxed{h_{2n} - h_n}. \end{aligned}$$

Et donc

$$w_n = \gamma_{2n} + \ln(2n) - \gamma_n - \ln(n) = \gamma_{2n} - \gamma_n + \ln(2) \xrightarrow{n \rightarrow +\infty} \gamma - \gamma + \ln(2) = \boxed{\ln(2)}.$$

14.c. D'après la question 9.d, la variable aléatoire $2 \ln(X_t)$ est de même loi que

$$\begin{aligned} 2 \sum_{k=0}^{2n-1} \left(\frac{1}{k+1} - \frac{Y_{k+1}}{t+k} \right) + 2 \ln \left(\frac{X_{t+2n}}{t+2n} \right) + 2d_{2n,t} - 2\gamma \\ = \sum_{k=0}^{n-1} \left(\frac{1}{k+1} - \frac{Y_{2k+1}}{\frac{t}{2} + k} \right) + \sum_{k=0}^{n-1} \left(\frac{1}{k+1} - \frac{Y_{2k+2}}{\frac{t+1}{2} + k} \right) + 2w_n + 2 \ln \left(\frac{X_{t+2n}}{t+2n} \right) + 2d_{2n,t} - 2\gamma. \end{aligned}$$

Mais la question 9.d, appliquée avec $\frac{t}{2}$ au lieu de t montre également que

$$\sum_{k=0}^{n-1} \left(\frac{1}{k+1} - \frac{Y_{k+1}}{\frac{t}{2} + k} \right) + \ln \left(\frac{X_{t/2+n}}{t+n} \right) + d_{n,t/2} - \gamma \text{ a même loi que } \ln(X_{t/2}).$$

De la même manière, $\sum_{k=0}^{n-1} \left(\frac{1}{k+1} - \frac{Y_{k+1}}{\frac{t+1}{2} + k} \right) + \ln \left(\frac{X_{(t+1)/2+n}}{t+n} \right) + d_{n,(t+1)/2} - \gamma$ a même loi que $\ln(X_{(t+1)/2})$.

Ainsi, $2 \ln(X_t)$ a même loi¹² que

¹² Toutes les variables considérées sont indépendantes.

$$\ln(X_{t/2}) + \ln(X_{(t+1)/2}) - \ln \left(\frac{X_{(t+1)/2+n}}{\frac{t+1}{2} + n} \right) - \ln \left(\frac{X_{t/2+n}}{\frac{t}{2} + n} \right) + 2 \ln \left(\frac{X_{t+n}}{t+n} \right) + 2w_n + 2d_{n,t} - 2d_{n,t/2} - 2d_{n,(t+1)/2}.$$

Or, à la question 3.c, nous avons prouvé que $\ln \left(\frac{X_{t+n}}{t+n} \right) \xrightarrow{P} 0$.

Et de même, $\ln \left(\frac{X_{t/2+n}}{t/2+n} \right) \xrightarrow{P} 0$ et $\ln \left(\frac{X_{(t+1)/2+n}}{(t+1)/2+n} \right) \xrightarrow{P} 0$.

Enfin, $2w_n + 2d_{n,t} - 2d_{n,t/2} - 2d_{n,(t+1)/2} \xrightarrow{n \rightarrow +\infty} 2 \ln(2)$.

Nous pouvons donc appliquer le résultat de la question 4, où $N = \ln(X_{t/2}) + \ln(X_{(t+1)/2})$, $u_n = 2w_n + 2d_{n,t} - 2d_{n,t/2} - 2d_{n,(t+1)/2}$ et

$$D_n = \underbrace{-\ln \left(\frac{X_{(t+1)/2+n}}{\frac{t+1}{2} + n} \right)}_{=A_n} - \underbrace{\ln \left(\frac{X_{t/2+n}}{\frac{t}{2} + n} \right)}_{=B_n} + \underbrace{2 \ln \left(\frac{X_{t+n}}{t+n} \right)}_{=C_n}.$$

Et donc $M = 2 \ln(X_t)$ a même loi que $N + u = \ln(X_{t/2}) + \ln(X_{(t+1)/2}) + 2 \ln(2)$.

14.d. D'après la question précédente, $2 \ln(X_{2s})$ et $\ln(X_s) + \ln \left(X_{s+\frac{1}{2}} \right) + 2 \ln(2)$ sont de même loi.

Et donc $2r \ln(X_{2s})$ et $r \ln(X_s) + r \ln \left(X_{s+\frac{1}{2}} \right) + 2r \ln(2)$ sont de même loi.

En passant à l'exponentielle, il vient $(X_{2s})^{2r}$ et $2^{2r} (X_s)^r \left(X_{s+\frac{1}{2}} \right)^r$ sont de même loi.

14.e. Pour tout $r > 0$ et tout $s > 0$, on a donc

$$E \left(X_s^{2r} \right) = 2^{2r} E \left(X_s^r X_{s+\frac{1}{2}}^r \right) = 2^{2r} E \left(X_s^r \right) E \left(X_{s+\frac{1}{2}}^r \right). \quad (\diamond)$$

Mais par le théorème de transfert, pour tout $r > 0$ et tout $t > 0$, on a

$$E(X_t^r) = \int_0^{+\infty} \frac{1}{\Gamma(t)} x^{t+r-1} e^{-x} dx = \frac{\Gamma(t+r)}{\Gamma(t)}.$$

Prenons $s = \frac{1}{2}$, de sorte que pour tout $r > 0$, la relation (\diamond) s'écrit

$$\frac{\Gamma(2r+1)}{\Gamma(1)} = 2^{2r} \frac{\Gamma(r+1)}{\Gamma(1)} \frac{\Gamma(r+\frac{1}{2})}{\Gamma(\frac{1}{2})}.$$

En se rappelant que $\Gamma(1) = 1$, et en multipliant par $\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}$, il vient

$$2^{2r} \Gamma(r+1) \Gamma\left(r+\frac{1}{2}\right) = \Gamma(2r+1) \sqrt{\pi}.$$

Enfin, notons que $\Gamma(r+1) = r\Gamma(r)$ et $\Gamma(2r+1) = 2r\Gamma(2r)$ de sorte qu'après division par $2r$,

$$2^{2r-1} \Gamma(r) \Gamma\left(r+\frac{1}{2}\right) = \Gamma(2r) \sqrt{\pi}.$$

15. Deuxième application : la formule de Stirling

15.a. Pour tout $u > 0$, on a

$$\begin{aligned} \frac{a}{u^2} + \frac{b}{(u+1)^2} + \frac{c}{u} + \frac{d}{u+1} &= \frac{a(u+1)^2 + bu^2 + cu(u+1)^2 + du^2(u+1)}{u^2(u+1)^2} \\ &= \frac{a(u^2 + 2u + 1) + bu^2 + c(u^3 + 2u^2 + u) + d(u^3 + u^2)}{u^2(u+1)^2} \end{aligned}$$

Détails

Les variables X_s et $X_{s+1/2}$ sont indépendantes, donc il en est de même de X_s^r et $X_{s+1/2}^r$.
Et donc l'espérance du produit est égale au produit des espérances.

Méthode

Nous choisissons $s = \frac{1}{2}$, principalement pour réussir à faire apparaître $\Gamma\left(r+\frac{1}{2}\right)$, qui viendra donc de $E\left(X_{1/2}^r\right)$.

$$= \frac{(c+d)u^3 + (a+b+2c+d)u^2 + (2a+c)u + a}{u^2(u+1)^2}.$$

Et donc $\frac{a}{u^2} + \frac{b}{(u+1)^2} + \frac{c}{u} + \frac{d}{u+1} = \frac{1}{u^2(u+1)^2}$ pour tout $u > 0$ si et seulement si $(c+d)u^3 + (a+b+2c+d)u^2 + (2a+c)u + a = 1$ pour tout $u > 0$.

Or deux polynôme coïncident en une infinité de points¹³ si et seulement si ils ont les mêmes coefficients, donc si et seulement si

¹³ Il y a une infinité de réels $u > 0$.

$$\begin{cases} a = 1 \\ 2a + c = 0 \\ a + b + 2c + d = 0 \\ c + d = 0 \end{cases} \Leftrightarrow \begin{cases} a = 1 \\ c = -2 \\ d = 2 \\ b = 1 \end{cases}$$

Et donc il vient, pour tout $t > 0$,

$$\begin{aligned} \sum_{k=0}^{+\infty} \frac{1}{(t+k)^2(t+k+1)^2} &= \sum_{k=0}^{+\infty} \left(\frac{1}{(t+k)^2} + \frac{1}{(t+k+1)^2} - \frac{2}{t+k} + \frac{2}{(t+k+1)} \right) \\ &= \sum_{k=0}^{+\infty} \frac{1}{(t+k)^2} + \sum_{i=1}^{+\infty} \frac{1}{(t+i)^2} - 2 \sum_{k=0}^{+\infty} \frac{1}{t+k} + 2 \sum_{i=1}^{+\infty} \frac{1}{t+i} \\ &= -\frac{1}{t^2} + 2 \sum_{k=0}^{+\infty} \frac{1}{(t+k)^2} - \frac{2}{t} \\ &= -\frac{1}{t^2} + 2\psi'(t) - \frac{2}{t}. \end{aligned}$$

On a utilisé l'expression de $\psi'(t)$ obtenue à la question 10.

Et donc

$$\psi'(t) = \frac{1}{t} + \frac{1}{2t^2} + \frac{1}{2} \sum_{k=0}^{+\infty} \frac{1}{(t+k)^2(t+k+1)^2}.$$

15.b. Puisque $\psi'(t) - \frac{1}{t} - \frac{1}{2t^2} = \frac{1}{2} \sum_{k=0}^{+\infty} \frac{1}{(t+k)^2(t+k+1)^2}$, en utilisant l'encadrement donné par l'énoncé, il vient donc

$$\frac{1}{6} \sum_{k=0}^{+\infty} \left(\frac{1}{(t+k+\frac{1}{14})^2} - \frac{1}{(t+k+\frac{15}{14})^2} \right) \leq \psi'(t) - \frac{1}{t} - \frac{1}{2t^2} \leq \frac{1}{6} \sum_{k=0}^{+\infty} \left(\frac{1}{(t+k)^3} - \frac{1}{(t+k+1)^3} \right).$$

Or, il est immédiat que la dernière série est une série télescopique :

$$\sum_{k=0}^{+\infty} \left(\frac{1}{(t+k)^3} - \frac{1}{(t+k+1)^3} \right) = \lim_{N \rightarrow +\infty} \sum_{k=0}^N \left(\frac{1}{(t+k)^3} - \frac{1}{(t+k+1)^3} \right) = \lim_{N \rightarrow +\infty} \left(\frac{1}{t^3} - \frac{1}{(t+N+1)^3} \right) = \frac{1}{t^3}.$$

De plus, en notant que $t+k+\frac{15}{14} = t+(k+1)+\frac{1}{14}$, la première série est également télescopique, et donc

$$\sum_{k=0}^{+\infty} \left(\frac{1}{(t+k+\frac{1}{14})^2} - \frac{1}{(t+k+\frac{15}{14})^2} \right) = \frac{1}{(t+\frac{1}{14})^2}.$$

Et donc nous avons prouvé que

$$\frac{1}{6} \frac{1}{(t+\frac{1}{14})^2} \leq \psi'(t) - \frac{1}{t} - \frac{1}{2t^2} \leq \frac{1}{6t^3}.$$

Intégrons cette inégalité entre t et A , avec $A > t$. Alors

$$\int_t^A \left(\frac{1}{2x^2} + \frac{1}{6(x+\frac{1}{14})^3} \right) dx \leq \int_t^A \left(\psi'(x) - \frac{1}{x} \right) dx \leq \int_t^A \left(\frac{1}{2x^2} - \frac{1}{6x^2} \right) dx.$$

Remarque

Toutes les séries en jeu sont évidemment convergentes par comparaison à des séries de Riemann.

Remarque

Si $A > t$, alors les bornes sont «dans le bon sens», et donc la croissance de l'intégrale s'applique bien.

On a alors

$$\int_t^A \left(\psi'(x) - \frac{1}{x} \right) dx = [\psi(x) - \ln(x)]_t^A = \psi(A) - \ln(A) - (\psi(t) - \ln(t)) \xrightarrow{A \rightarrow +\infty} -(\psi(t) - \ln(t)).$$

D'autre part,

$$\int_t^A \left(\frac{1}{2x^2} + \frac{1}{6(x + \frac{1}{14})^3} \right) dx = \left[-\frac{1}{2x} - \frac{1}{12(x + \frac{1}{14})^2} \right]_t^A = -\frac{1}{2A} - \frac{1}{12(A + \frac{1}{14})^2} + \frac{1}{2t} + \frac{1}{12(t + \frac{1}{14})^2} \xrightarrow{A \rightarrow +\infty} \frac{1}{2t} + \frac{1}{12(t + \frac{1}{14})^2}.$$

Enfin,

$$\int_t^A \left(\frac{1}{2x^2} + \frac{1}{6x^3} \right) dx = \left[-\frac{1}{2x} - \frac{1}{12x^2} \right]_t^A \xrightarrow{A \rightarrow +\infty} \frac{1}{2t} + \frac{1}{12t^2}.$$

Autrement dit, en faisant tendre A vers $+\infty$ dans l'inégalité ci-dessus, on a

$$\frac{1}{2t} + \frac{1}{12(t + \frac{1}{14})^2} \leq -(\psi(t) - \ln(t)) \leq \frac{1}{2t} + \frac{1}{12t^2}.$$

Et donc en multipliant le tout par -1 , on a enfin

$$\boxed{-\frac{1}{2t} - \frac{1}{12t^2} \leq \psi(t) - \ln(t) \leq -\frac{1}{2t} - \frac{1}{12(t + \frac{1}{14})^2}.$$

En divisant cette inégalité par $\ln t$, on obtient, pour $t > 1$,

$$1 - \frac{1}{2t \ln t} - \frac{1}{12t^2 \ln t} \leq \frac{\psi(t)}{\ln t} \leq 1 - \frac{1}{2t \ln t} - \frac{1}{12(t + \frac{1}{14})^2 \ln t}.$$

Et donc par le théorème des gendarmes, $\lim_{t \rightarrow +\infty} \frac{\psi(t)}{\ln t} = 1 \Leftrightarrow \psi(t) \underset{t \rightarrow +\infty}{\sim} \ln(t)$.

De même, on a

$$\frac{1}{2t} + \frac{t}{6(t + \frac{1}{14})^3} \leq t\psi'(t) - 1 \leq \frac{1}{2t} + \frac{1}{6t^2}$$

de sorte que par le théorème des gendarmes, $\lim_{t \rightarrow +\infty} t\psi'(t) - 1 = 0 \Leftrightarrow \psi'(t) \underset{t \rightarrow +\infty}{\sim} \frac{1}{t}$.

Et donc on en déduit que

$$\boxed{E(\ln(X_t)) \underset{t \rightarrow +\infty}{\sim} \ln(t) \text{ et } V(\ln(X_t)) \underset{t \rightarrow +\infty}{\sim} \frac{1}{t}.$$

15.c. Notons que $\psi : t \mapsto \frac{\Gamma'(t)}{\Gamma(t)}$ est la dérivée de $t \mapsto \ln(\Gamma(t))$.

Et donc

$$\int_t^y \left(\psi(x) - \ln(x) + \frac{1}{2x} \right) dx = \left[\ln(\Gamma(x)) - x \ln(x) + x + \frac{\ln(x)}{2} \right]_t^y.$$

Mais l'encadrement de la question 15.b montre que

$$-\frac{1}{12x^2} \leq \psi(x) - \ln(x) + \frac{1}{2x} \leq -\frac{1}{12(t + \frac{1}{14})^2}$$

et donc $\psi(x) - \ln(x) + \frac{1}{2x} \underset{x \rightarrow +\infty}{\sim} -\frac{1}{12x^2}$.

On en déduit, par critère de comparaison pour les intégrales de signe constant, que

$$\int_1^{+\infty} \left(\psi(x) - \ln(x) + \frac{1}{2x} \right) dx \text{ converge.}$$

 **Danger !**

Attention au sens des inégalités : on prend $t > 1$ pour diviser par un nombre positif, et donc préserver le sens des inégalités.

Autrement dit, $\lim_{y \rightarrow +\infty} \int_1^y \left(\psi(x) - \ln(x) + \frac{1}{2x} \right) dx$ existe.

Et donc $\lim_{y \rightarrow +\infty} \ln(\Gamma(y)) - y \ln(y) + y + \frac{\ln(y)}{2}$ existe.

Or $\ln(\Gamma(y)) - y \ln(y) + y + \frac{\ln(y)}{2} = \ln \left(\frac{\Gamma(y)}{y^{y-\frac{1}{2}} e^{-y}} \right)$.

Ceci prouve donc que $\lim_{y \rightarrow +\infty} \ln \left(\frac{\Gamma(y)}{y^{y-\frac{1}{2}} e^{-y}} \right)$ existe.

15.d. Commençons par nous assurer que l'identité donnée par l'énoncé est correcte :

$$\left(x + \frac{1}{2}\right)^x \left(1 + \frac{1}{2x}\right)^{-x} = \left(\frac{2x+1}{2}\right)^x \left(\frac{2x+1}{2x}\right)^{-x} = \left(\frac{2x+1}{2}\right)^x \left(\frac{2x}{2x+1}\right)^x = x^x.$$

Par définition, $e^\theta = \lim_{y \rightarrow +\infty} \frac{\Gamma(y)}{y^{y-\frac{1}{2}} e^{-y}}$.

Et donc en particulier,

$$e^\theta = \lim_{r \rightarrow +\infty} \frac{\Gamma(2r)}{(2r)^{2r-\frac{1}{2}} e^{-2r}}.$$

Or, d'après la question 14.e,

$$\begin{aligned} \frac{\Gamma(2r)}{(2r)^{2r-\frac{1}{2}} e^{-2r}} &= \frac{2^{2r-1} \Gamma(r) \Gamma\left(r + \frac{1}{2}\right)}{\sqrt{\pi} (2r)^{2r-\frac{1}{2}} e^{-2r}} \\ &= \frac{1}{\sqrt{2\pi}} \frac{\Gamma(r) \Gamma\left(r + \frac{1}{2}\right)}{r^{2r-\frac{1}{2}} e^{-2r}} \\ &= \frac{1}{\sqrt{2\pi}} \frac{\Gamma(r)}{r^{r-\frac{1}{2}} e^{-r}} \frac{\Gamma\left(r + \frac{1}{2}\right)}{r^r e^{-r}} \\ &= \frac{1}{\sqrt{2\pi}} \frac{\Gamma(r)}{r^{r-\frac{1}{2}} e^{-r}} \frac{\Gamma\left(r + \frac{1}{2}\right)}{\left(r + \frac{1}{2}\right)^r \left(1 + \frac{1}{2r}\right)^{-r} e^{-r}} \\ &= \frac{1}{\sqrt{2\pi}} \frac{\Gamma(r)}{r^{r-\frac{1}{2}} e^{-r}} \frac{\Gamma\left(r + \frac{1}{2}\right)}{\left(r + \frac{1}{2}\right)^{r+\frac{1}{2}-\frac{1}{2}} e^{-r-\frac{1}{2}}} e^{-\frac{1}{2}} \left(1 + \frac{1}{2r}\right)^r. \end{aligned}$$

C'est ici qu'on utilise l'identité donnée par l'énoncé.

Mais par définition de e^θ , on a $\lim_{r \rightarrow +\infty} \frac{\Gamma(r)}{r^{r-\frac{1}{2}} e^{-r}} = \lim_{r \rightarrow +\infty} \frac{\Gamma\left(r + \frac{1}{2}\right)}{\left(r + \frac{1}{2}\right)^{r+\frac{1}{2}-\frac{1}{2}} e^{-r-\frac{1}{2}}} = e^\theta$.

D'autre part, on a

$$\left(1 + \frac{1}{2r}\right)^r = e^{r \ln\left(1 + \frac{1}{2r}\right)}.$$

Or, $\ln\left(1 + \frac{1}{2r}\right) \underset{r \rightarrow +\infty}{\sim} \frac{1}{2r}$ et donc $r \ln\left(1 + \frac{1}{2r}\right) \underset{r \rightarrow +\infty}{\sim} \frac{1}{2}$.

Par continuité de l'exponentielle, on a donc $\lim_{r \rightarrow +\infty} \left(1 + \frac{1}{2r}\right)^r = e^{1/2}$.

Et donc il vient, par unicité de la limite,

$$e^\theta = \lim_{r \rightarrow +\infty} \frac{\Gamma(2r)}{(2r)^{2r-\frac{1}{2}} e^{-2r}} = \lim_{r \rightarrow +\infty} \frac{1}{\sqrt{2\pi}} \frac{\Gamma(r)}{r^{r-\frac{1}{2}} e^{-r}} \frac{\Gamma\left(r + \frac{1}{2}\right)}{\left(r + \frac{1}{2}\right)^{r+\frac{1}{2}-\frac{1}{2}} e^{-r-\frac{1}{2}}} e^{-\frac{1}{2}} \left(1 + \frac{1}{2r}\right)^r = \frac{1}{\sqrt{2\pi}} e^{2\theta}.$$

Et donc $e^\theta = \sqrt{2\pi}$.

On en déduit donc que $\frac{\Gamma(x)}{x^{x-\frac{1}{2}} e^{-x}} \underset{x \rightarrow +\infty}{\sim} \sqrt{2\pi} \Leftrightarrow \Gamma(x) \underset{x \rightarrow +\infty}{\sim} \sqrt{2\pi} x^{x-\frac{1}{2}} e^{-x}$.

Remarque : en particulier, pour $n \in \mathbf{N}$, on trouve $(n-1)! = \Gamma(n) \underset{n \rightarrow +\infty}{\sim} \sqrt{2\pi} n^{n-1/2} e^{-n}$, ce qui après multiplication par n nous donne l'équivalent «classique»

$$n! \underset{n \rightarrow +\infty}{\sim} \sqrt{2\pi} n^n e^{-n}.$$

C'est généralement cette formule que l'on appelle formule de Stirling. On en trouvera une preuve «simple» dans le sujet d'EML 2012.

EXTRAITS COMMENTÉS DU RAPPORT DU JURY

3.a — Dans cette question, on demandait explicitement de justifier la convergence absolue d'une intégrale. Le plus souvent, les candidats se contentent de faire le calcul pour conclure.

☞ De manière générale, lorsqu'on calcule une espérance, à part pour les variables aléatoires finies, il y a une convergence à justifier (d'une série pour les variables discrètes, d'une intégrale pour les variables à densité).

S'il n'y a généralement pas besoin d'écrire un roman pour justifier cette convergence (le calcul fait souvent apparaître une série/intégrale de référence), il faut tout de même indiquer à quel moment on a prouvé la convergence. Ceci permet notamment de montrer qu'on sait bien que le théorème de transfert est valable **sous réserve de convergence** de l'intégrale/la série qui apparaît.

4.a — Cette question est très mal traitée. Ainsi, l'inégalité $P(A \cup B \cup C) \leq P(A) + P(B) + P(C)$ a donné lieu à des hypothèses de validité assez curieuses telles que «les événements A , B et C doivent être indépendants», ou encore «ils doivent être incompatibles». On lit souvent des implications et des inégalités (\leq) entre événements, des inclusions entre variables aléatoires ou entre probabilité.

☞ Rappelons que la formule $P(A \cup B) \leq P(A) + P(B)$ est **toujours** valable. En revanche ce n'est une égalité que si les événements sont incompatibles.

Et l'indépendance n'a rien à faire ici, elle n'apparaît que si on s'intéresse à la probabilité de $A \cap B$, et non de $A \cup B$.

Enfin, les autres erreurs ne sont malheureusement pas surprenantes, et relèvent à la fois d'une mauvaise compréhension des concepts de probabilités, ainsi que de problèmes de logique.

Un événement est un ensemble (plus précisément : un sous-ensemble de l'univers Ω), et qu'à ce titre on peut écrire des inclusions et/ou des égalités entre événements. Mais pas d'implications (on ne peut avoir une implication qu'entre propositions logiques, par exemple des égalités), ni d'inégalités (on ne peut avoir une inégalité qu'entre deux nombres réels) ou de sommes.

Enfin, une probabilité est un nombre, et on ne peut donc pas avoir d'inclusion entre probabilités.

S'il est parfois possible de cacher son manque de compréhension de toutes ces notions dans des sujets plus classiques, il est plus dur d'aborder sérieusement un sujet de Maths II sans être à l'aise avec ces concepts.

4.b — Cette question et la suivante révèlent le faible niveau de compréhension des concepts de limite et de variable aléatoire : la plupart des candidats confondent l'égalité (ponctuelle) des variables aléatoires et leur égalité en loi.

☞ Il est évident que deux variables aléatoires égales ont la même loi, la réciproque étant clairement fausse.

Par exemple : si on lance deux dés, et qu'on note X_1 (respectivement X_2) le résultat du premier (resp. du second) lancer. Alors il est clair que X_1 et X_2 suivent la même loi (une loi uniforme sur $\llbracket 1, 6 \rrbracket$). En revanche, X_1 et X_2 ne sont pas égales : le résultat du premier dé n'est pas toujours le même que celui du premier dé (l'un des deux dés peut par exemple donner un 1 alors que l'autre donne un 6).

5.b — L'indépendance des variables aléatoires n'est pas assez souvent mentionnée pour justifier l'application de la formule de convolution.

☞ Le produit de convolution pour les variables à densité n'est valable que pour la somme de deux variables **indépendantes**. C'est toujours le cas lorsqu'on en a besoin, mais il est indispensable de vérifier cette indépendance (généralement, il suffit de la mentionner). En l'absence de cette hypothèse d'indépendance, nous ne disposons d'aucun résultat général sur la somme de deux variables à densité.

13.b — La fonction qui devait simuler les valeurs prises par une variable aléatoire ne fait souvent que calculer des valeurs de la densité de cette variable.

☞ Simuler une variable aléatoire devrait retourner un résultat aléatoire (et donc en SciLab devrait faire appel à `rand` ou à `grand`), c'est-à-dire que plusieurs appels successifs à la fonction devraient retourner des résultats différents. Alors que le calcul de la densité (qui doit tout de même dépendre d'un paramètre x : pour quelle valeur de x veut-on calculer $f(x)$?) n'a rien d'aléatoire, et doit toujours retourner la même valeur.

15.a — Une proportion non négligeable de candidats ont eu le courage de calculer la décomposition en éléments simples (en résolvant un système de quatre équations) et le mérite de trouver les bons coefficients.

☞ La première partie de cette question est somme toute assez classique, bien que particulièrement calculatoire.

S'il est assez clair dès le départ que les calculs vont être pénibles, il ne faut pas hésiter à se lancer : on est dans une épreuve difficile, et il n'y aura pas ou peu de questions «faciles», il faut se raccrocher aux questions que l'on pense savoir faire. En tous cas, il semble que certains candidats y arrivent, ce qui leur permet sûrement de faire la différence au final.

DEVOIR MAISON 12 (À RENDRE LE 17.01.19)

Le but de cet exercice est de prouver l'existence et de donner la valeur (par deux méthodes différentes) de

$$\Delta = \inf_{(x,y) \in \mathbf{R}^2} \int_0^{+\infty} (t^3 - xt - y)^2 e^{-t} dt.$$

Partie I : méthode utilisant un produit scalaire

Soit E l'espace vectoriel des polynômes à coefficients réels, de degré inférieur ou égal à 3 et F le sous-espace vectoriel de E engendré par les deux polynômes 1 et X .

1.
 - a. Rappeler pourquoi, pour tout entier naturel k , l'intégrale $\int_0^{+\infty} t^k e^{-t} dt$ converge.
 - b. Montrer que l'application de $E \times E$ dans \mathbf{R} qui, à tout couple (P, Q) d'éléments de E , associe $\langle P, Q \rangle = \int_0^{+\infty} P(t)Q(t)e^{-t} dt$ est un produit scalaire sur E , dont la norme associée sera notée $\| \cdot \|$.
2. Soit Q un polynôme de F défini par $Q = xX + y$, où x et y sont deux réels. Donner, sous forme d'intégrale, l'expression de $\|X^3 - Q\|^2$.
3.
 - a. Énoncer le théorème qui assure l'existence et l'unicité du polynôme Q_0 de F qui rend $\|X^3 - Q\|^2$ minimale.
 - b. En déduire sans calcul les valeurs de $\langle X^3 - Q_0, 1 \rangle$ et $\langle X^3 - Q_0, X \rangle$.
 - c. En notant $Q_0 = x_0X + y_0$, écrire le système que doit vérifier le couple (x_0, y_0) pour que $\int_0^{+\infty} (t^3 - xt - y)^2 e^{-t} dt$ soit minimale.
 - d. Déterminer la valeur de Δ .

Partie 2 : méthode utilisant une fonction de deux variables

On note f la fonction définie sur $\mathbf{R} \times \mathbf{R}$ par

$$\forall (x, y) \in \mathbf{R} \times \mathbf{R}, f(x, y) = \int_0^{+\infty} (t^3 - tx - y)^2 e^{-t} dt.$$

4. Écrire $f(x, y)$ comme une fonction polynomiale des deux variables x et y .
5. Déterminer le seul point critique (x_0, y_0) de f sur $\mathbf{R} \times \mathbf{R}$.
6. Établir que f possède un minimum global.

CORRECTION DU DEVOIR MAISON 12

EDHEC 2014

Partie I : méthode utilisant un produit scalaire

- 1.a. La fonction $t \mapsto t^k e^{-t}$ est continue sur $[0, +\infty[$, et on a $t^2 t^k e^{-t} \xrightarrow[t \rightarrow \infty]{} 0$, de sorte que

$$t^k e^{-t} = o_{t \rightarrow +\infty} \left(\frac{1}{t^2} \right).$$

Puisque $\int_1^{+\infty} \frac{1}{t^2} dt$ converge, il en est de même de $\int_0^{+\infty} t^k e^{-t} dt$.

- 1.b. Pour tout $(P, Q) \in E^2$, $t \mapsto P(t)Q(t)e^{-t}$ est continue sur $[0, +\infty[$.
De plus, si $a_i X^i$ est le terme de plus haut degré de P et $b_j X^j$ est le terme de plus haut degré de Q , alors

$$|P(t)Q(t)e^{-t}| \underset{t \rightarrow +\infty}{\sim} |a_i b_j| t^{i+j} e^{-t}$$

et l'intégrale $\int_0^{+\infty} |a_i b_j| t^{i+j} e^{-t} dt$ converge d'après la question précédente.

Ainsi, $\int_0^{+\infty} P(t)Q(t)e^{-t} dt$ converge absolument, et donc converge.

Si $(P, Q, R) \in E^3$ et $\lambda \in \mathbf{R}$, alors

$$\begin{aligned} \langle \lambda P + Q, R \rangle &= \int_0^{+\infty} (\lambda P(t) + Q(t))R(t)e^{-t} dt = \lambda \int_0^{+\infty} P(t)R(t)e^{-t} dt + \int_0^{+\infty} Q(t)R(t)e^{-t} dt \\ &= \lambda \langle P, R \rangle + \langle Q, R \rangle. \end{aligned}$$

Donc $\langle \cdot, \cdot \rangle$ est linéaire à gauche.

Pour tous $(P, Q) \in E \times E$, on a

$$\langle P, Q \rangle = \int_0^{+\infty} P(t)Q(t)e^{-t} dt = \int_0^{+\infty} Q(t)P(t)e^{-t} dt = \langle Q, P \rangle.$$

Donc $\langle \cdot, \cdot \rangle$ est symétrique, et donc¹ est bilinéaire symétrique.

Si $P \in E$, alors la fonction $t \mapsto P(t)^2 e^{-t}$ est positive sur $[0, +\infty[$, donc par positivité de l'intégrale,

$$\langle P, P \rangle = \int_0^{+\infty} P(t)^2 e^{-t} dt \geq 0.$$

Enfin, si $\langle P, P \rangle = 0$, la fonction $t \mapsto P(t)^2 e^{-t}$ étant continue, positive sur $[0, +\infty[$ et d'intégrale nulle, elle est nécessairement nulle.

Et donc $\forall t \in [0, +\infty[, P(t)^2 e^{-t} = 0 \Rightarrow P(t) = 0$.

Ainsi, P possède une infinité de racines, et donc est le polynôme nul.

Ainsi, $\langle \cdot, \cdot \rangle$ est un produit scalaire sur E .

2. On a, par définition de la norme

$$\|X^3 - Q\|^2 = \langle X^3 - xX - y, X^3 - xX - y \rangle = \int_0^{+\infty} (t^3 - xt - y)^2 e^{-t} dt.$$

- 3.a. L'énoncé exact du théorème en question (qui ne possède pas de nom officiel dans le programme d'ECS2, certains l'appellent parfois «théorème de la projection orthogonale») est le suivant : soit E un espace euclidien, F un sous-espace vectoriel de E et $x \in E$. Alors il existe un unique $y \in F$ tel que $\|x - y\| = \min \{\|x - z\|, z \in F\}$ et ce y est le projeté orthogonal de x sur F .

Ici, F est un sous-espace vectoriel de E , et $X^3 \in E$, donc il existe un unique $Q_0 \in F$ tel que

$$\|X^3 - Q_0\| = \min \{\|X^3 - P\|, P \in F\},$$

et Q_0 est alors le projeté orthogonal de X^3 sur F .

Remarque

L'énoncé n'est pas très clair à ce sujet, mais on peut aussi remarquer qu'il s'agit de l'intégrale définissant $\Gamma(k+1)$, dont on sait qu'elle converge (et dont nous connaissons la valeur !).

Convergence

L'énoncé ne demande pas clairement de montrer que cette intégrale converge, mais il faut bien le prouver pour être sûr que l'application $\langle \cdot, \cdot \rangle$ est correctement définie.

¹ Linéaire à gauche + symétrique implique automatiquement linéaire à droite.

Racines

Ne pas oublier l'argument de racines. Il semble évident que P étant nul sur $[0, +\infty[$, il doit être nul, mais ceci tient au fait que ce soit un polynôme.

En effet, pour une fonction non polynomiale, ce résultat n'est plus valable. Par exemple la fonction $t \mapsto \min(t, 0)$ est nulle sur $[0, +\infty[$, mais n'est pas la fonction nulle pour autant.

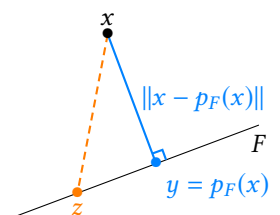


FIGURE 23.1– Illustration du théorème.

- 3.b. Si $Q_0 = p_F(X^3) \in F$, alors nécessairement $X^3 - Q_0 \in F^\perp$.
Et puisque $1 \in F$ et $X \in F$, alors

$$\langle X^3 - Q_0, 1 \rangle = \langle X^3 - Q_0, X \rangle = 0.$$

- 3.c. Le couple (x_0, y_0) doit vérifier les deux équations obtenues à la question précédente, à savoir

$$\begin{cases} \langle X^3 - Q_0, 1 \rangle = 0 \\ \langle X^3 - Q_0, X \rangle = 0 \end{cases} \Leftrightarrow \begin{cases} \int_0^{+\infty} (t^3 - x_0 t - y) e^{-t} dt = 0 \\ \int_0^{+\infty} (t^3 - x_0 t - y_0) t e^{-t} dt = 0 \end{cases}$$

Pour calculer les intégrales, remarquons tout de suite que

$$\int_0^{+\infty} t^k e^{-t} dt = \Gamma(k+1) = k!$$

Ainsi, le système se réécrit

$$\begin{cases} 6 - x_0 - y_0 = 0 \\ 24 - 2x_0 - y_0 = 0 \end{cases} \Leftrightarrow \begin{cases} 6 = x_0 + y_0 \\ 24 = 2x_0 + y_0 \end{cases} \Leftrightarrow \begin{cases} x_0 = 18 \\ y_0 = -12 \end{cases}$$

- 3.d. Notons que

$$X^3 = \underbrace{X^3 - Q_0}_{\in F^\perp} + \underbrace{Q_0}_{\in F}$$

et donc par le théorème de Pythagore,

$$\|X^3\|^2 = \|X^3 - Q_0\|^2 + \|Q_0\|^2.$$

Ainsi, on a

$$\begin{aligned} \Delta &= \|X^3 - Q_0\|^2 \\ &= \|X^3\|^2 - \|Q_0\|^2 \\ &= \int_0^{+\infty} t^6 e^{-t} dt - \int_0^{+\infty} (-18t + 12)^2 e^{-t} dt \\ &= 6! - 36 \int_0^{+\infty} (3t - 2)^2 e^{-t} dt \\ &= 720 - 36 \left(9 \int_0^{+\infty} t^2 e^{-t} dt - 12 \int_0^{+\infty} t e^{-t} dt + 4 \int_0^{+\infty} e^{-t} dt \right) \\ &= 720 - 36(18 - 12 + 4) = \boxed{360}. \end{aligned}$$

Partie II : méthode utilisant une fonction de deux variables

4. En développant le carré, on a, pour tout $t \geq 0$,

$$(t^3 - tx - y)^2 = t^6 + t^2 x^2 + y^2 - 2t^4 x - 2t^3 y + 2txy$$

et donc

$$\begin{aligned} f(x, y) &= \int_0^{+\infty} t^6 e^{-t} dt + x^2 \int_0^{+\infty} t^2 e^{-t} dt + y^2 \int_0^{+\infty} e^{-t} dt - 2x \int_0^{+\infty} t^4 e^{-t} dt - 2y \int_0^{+\infty} t^3 e^{-t} dt + 2xy \int_0^{+\infty} t e^{-t} dt \\ &= 6! + x^2 2! + y^2 1! - 2x 4! - 2y 3! + 2xy 1! \\ &= \boxed{720 + 2x^2 + y^2 - 48x - 12y + 2xy} \end{aligned}$$

5. La fonction f est de classe \mathcal{C}^1 sur \mathbf{R}^2 car polynomiale, et on a

$$\partial_1 f(x, y) = 4x - 48 + 2y, \quad \partial_2 f(x, y) = 2y - 12 + 2x.$$

Donc (x, y) est un point critique de f si et seulement si

$$\begin{cases} 4x - 48 + 2y = 0 \\ 2y - 12 + 2x = 0 \end{cases} \Leftrightarrow \begin{cases} 4x + 2y = 48 \\ -24 = 2y \end{cases} \Leftrightarrow \begin{cases} x = 18 \\ y = -12 \end{cases}$$

Donc f admet pour unique point critique $(18, -12)$.

6. Notons que pour tout $(x, y) \in \mathbf{R}^2$, on a

$$\begin{aligned} f(x, y) - f(x_0, y_0) &= 2x^2 - 48x + 2xy + y^2 - 12y + 720 - 360 \\ &= y^2 + 2y(x - 6) + 2x^2 - 48x + 360 \\ &= (y + (x - 6))^2 - (x - 6)^2 + 2x^2 - 48x + 360 \\ &= (y - x + 6)^2 + x^2 - 36x + 324 \\ &= (y - x + 6)^2 + (x - 18)^2 \geq 0 \end{aligned}$$

Et donc pour tout $(x, y) \in \mathbf{R}^2$, on a $f(x, y) \geq f(x_0, y_0)$: f atteint un minimum global en (x_0, y_0) .

Méthode alternative : à x fixé, $y \mapsto f(x, y) - f(x_0, y_0) = y^2 + (2x - 12)y + 360 - 48x + 2x^2$ est un polynôme de degré 2, dont le discriminant² est

$$\Delta(x) = (2x - 12)^2 - 4(360 - 48x + 2x^2) = 4x^2 - 48x + 144 - 1440 + 192x - 8x^2 = -4x^2 - 144x - 144 \times 9.$$

Pour étudier le signe de ce polynôme en x , nous pouvons à son tour calculer son discriminant, qui vaut

$$\Delta = 144^2 - 16 \times 144 \times 9 = 144 \times (144 - 16) = 0.$$

Puisque $\Delta \leq 0$, $\Delta(x)$ ne change pas de signe sur \mathbf{R} .

Or, $\Delta(0) < 0$, donc pour tout $x \in \mathbf{R}$, $\Delta(x) \leq 0$.

Et donc la fonction $y \mapsto f(x, y) - f(x_0, y_0)$ est elle aussi de signe constant, de même signe que $f(x, 0) - f(x_0, y_0)$.

Or $f(x, 0) - f(x_0, y_0) = 360 - 48x + 2x^2$. À son tour, cette expression est polynomiale de degré 2 et possède $48^2 - 8 \times 360 = 196 \times 12 - 240 \times 20 < 0$ comme discriminant, elle ne change donc pas de signe sur \mathbf{R} .

Et $f(0, 0) - f(x_0, y_0) = 360 \geq 0$, donc pour tout $x \in \mathbf{R}$, $f(x, 0) - f(x_0, y_0) \geq 0$.

Par conséquent, pour tout $(x, y) \in \mathbf{R}$, $f(x, y) - f(x_0, y_0) \geq 0$, et donc f possède un minimum global en (x_0, y_0) .

Méthode

On a mis sous forme canonique un polynôme de degré 2 en y . Rappelons qu'il s'agit de faire apparaître le début d'une identité remarquable, quitte à ajuster les constantes par la suite :

$$y^2 + 2ay = (y + a)^2 - a^2.$$

² Qui dépend de x .

Astuce

S'il faut calculer des produits sans calculatrice, autant factoriser au maximum les expressions afin d'éviter des calculs trop fastidieux.

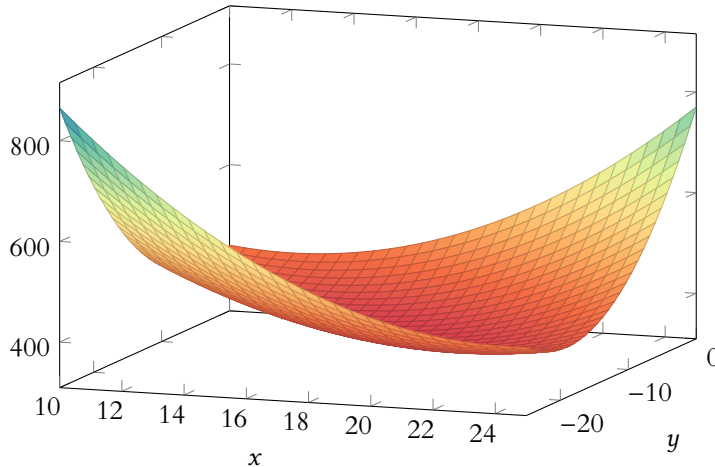


FIGURE 23.2 – Le graphe de la fonction f .

DEVOIR MAISON 13 (À RENDRE LE 25.01.19)

EXERCICE 1

Toutes les variables aléatoires intervenant dans cet exercice sont définies sur le même espace probabilisé (Ω, \mathcal{A}, P) .

On considère une suite $(X_n)_{n \geq 1}$ de variables aléatoires indépendantes et identiquement distribuées, suivant la loi exponentielle de paramètre $\lambda > 0$. On considère aussi, pour tout entier naturel n non nul la variable aléatoire M_n définie par $M_n = \max(X_1, X_2, \dots, X_n)$, c'est-à-dire que, pour tout ω de Ω , on a $M_n(\omega) = \max(X_1(\omega), X_2(\omega), \dots, X_n(\omega))$.

1. Soit g la fonction définie sur \mathbf{R} par : $\forall x \in \mathbf{R}, g(x) = e^{-x} \exp(-e^{-x})$.
 - a. Montrer que g est une densité de probabilité. On note G une variable aléatoire admettant g comme densité.
 - b. Déterminer la fonction de répartition, notée F_G , de la variable G .
2.
 - a. Donner, pour tout entier naturel n non nul, la fonction de répartition de la variable M_n .
 - b. Pour tout entier naturel n non nul, on pose $U_n = \lambda M_n - \ln(n)$. Montrer que la suite $(U_n)_{n \in \mathbf{N}^*}$ converge en loi vers une variable aléatoire dont on précisera la loi.

EXERCICE 2

On admet que si une suite $(T_n)_n$ de variables aléatoires converge en probabilité, alors la limite de cette suite est une variable aléatoire presque sûrement unique.

Plus précisément, si l'on a $T_n \xrightarrow{P} T$ et $T_n \xrightarrow{P} T'$, alors $P(T = T') = 1$.

On dit qu'une suite $(U_n)_{n \geq 1}$ de variables aléatoires converge en moyenne vers une variable aléatoire U si et seulement si : pour tout entier naturel n , la variable aléatoire $|U - U_n|$ possède une espérance et

$$\lim_{n \rightarrow +\infty} E(|U_n - U|) = 0.$$

On rappelle l'inégalité de Markov, valable pour une variable aléatoire V à valeurs positives et possédant une espérance mathématique :

$$\forall \varepsilon > 0, P(V > \varepsilon) \leq \frac{E(V)}{\varepsilon}.$$

1. Dans cette question, on considère une suite $(X_n)_{n \geq 1}$ de variables aléatoires définies sur un espace probabilisé (Ω, \mathcal{A}, P) et une variable aléatoire X , elle aussi définie sur cet espace probabilisé.
Montrer que, si la suite (X_n) converge en moyenne vers X , alors elle converge en probabilité vers X .

On se propose, dans la suite, d'étudier un exemple montrant que la réciproque de cette propriété est fautive.

Pour ce faire, on considère une suite $(Z_n)_{n \geq 1}$ de variables aléatoires définies sur un espace probabilisé (Ω, \mathcal{A}, P) , indépendantes, et suivant toutes la loi de Poisson de paramètre λ , avec $\lambda > 1$.

Pour tout entier naturel n non nul, on pose $Y_n = \prod_{k=1}^n Z_k$.

2.
 - a. Pour tout entier naturel n non nul, déterminer $P(Y_n \neq 0)$.
 - b. Pour tout ε strictement positif, comparer les événements $[Y_n > \varepsilon]$ et $[Y_n \neq 0]$.
 - c. En déduire que la suite (Y_n) converge en probabilité vers la variable certaine égale à 0.
3. Dans cette question, on suppose que la suite (Y_n) converge en moyenne vers une variable aléatoire Y .
 - a. Montrer que $P(Y = 0) = 1$.
 - b. Calculer l'espérance de Y_n .
 - c. Établir que $E(|Y_n - Y|) \geq E(Y_n) - E(Y)$, puis en déduire $\lim_{n \rightarrow +\infty} E(|Y_n - Y|)$.
4. Conclure.

CORRECTION DU DEVOIR MAISON 13

EXERCICE 1 (EXTRAIT DE EDHEC 2011)

- 1.a. La fonction g est évidemment positive car produit de fonctions positives, et elle est continue sur \mathbf{R} . De plus, pour $A < B$, on a

$$\int_A^B g(t) dt = [e^{-e^{-t}}]_A^B = e^{-e^{-B}} - e^{-e^{-A}}.$$

Or, lorsque $A \rightarrow -\infty$, $e^{-A} \rightarrow +\infty$ et donc $e^{-e^{-A}} \rightarrow 0$.

On en déduit que $\int_{-\infty}^A g(t) dt$ converge et vaut $e^{-e^{-B}}$.

Mais lorsque $B \rightarrow +\infty$, alors $e^{-B} \rightarrow 0$ et donc $e^{-e^{-B}} \rightarrow e^0 = 1$.

Et donc $\int_{-\infty}^{+\infty} g(t) dt$ converge et vaut 1. Ainsi, g est une densité de probabilités.

- 1.b. Pour $x \in \mathbf{R}$, on a $F_G(x) = \int_{-\infty}^x e^{-t} e^{-e^{-t}} dt$.

Or, pour $A < x$, on a

$$\int_A^x e^{-t} e^{-e^{-t}} dt = [e^{-e^{-t}}]_A^x = e^{-e^{-x}} - e^{-e^{-A}} \xrightarrow{A \rightarrow -\infty} e^{-e^{-x}}.$$

Et donc $F_G(x) = e^{-e^{-x}}$.

- 2.a. Pour $x \in \mathbf{R}$, on a $[M_n \leq x] = \bigcap_{i=1}^n [X_i \leq x]$, et par indépendance des X_i , il vient donc

$$P(M_n \leq x) = \prod_{i=1}^n P(X_i \leq x) = P(X_1 \leq x)^n.$$

Et puisque $X_1 \hookrightarrow \mathcal{E}(\lambda)$, on a $P(X_1 \leq x) = \begin{cases} 1 - e^{-\lambda x} & \text{si } x \geq 0 \\ 0 & \text{sinon} \end{cases}$

Ainsi, il vient

$$P(M_n \leq x) = \begin{cases} (1 - e^{-\lambda x})^n & \text{si } x \geq 0 \\ 0 & \text{sinon} \end{cases}$$

- 2.b. Commençons par déterminer la fonction de répartition de U_n : pour $x \in \mathbf{R}$, on a

$$\begin{aligned} P(U_n \leq x) &= P(\lambda M_n - \ln(n) \leq x) = P(\lambda M_n \leq x + \ln(n)) = P\left(M_n \leq \frac{x + \ln(n)}{\lambda}\right) \\ &= \begin{cases} (1 - e^{-x - \ln(n)})^n & \text{si } x + \ln(n) \geq 0 \\ 0 & \text{sinon} \end{cases} = \begin{cases} \left(1 - \frac{1}{n} e^{-x}\right)^n & \text{si } x \geq -\ln(n) \\ 0 & \text{sinon} \end{cases} \end{aligned}$$

Soit $x \in \mathbf{R}$ fixé. Alors pour n suffisamment grand, on aura $x \geq -\ln(n)$, car $\lim_{n \rightarrow +\infty} -\ln(n) = -\infty$.

Il s'agit donc de déterminer $\lim_{n \rightarrow +\infty} \left(1 - \frac{e^{-x}}{n}\right)^n$. Mais

$$\left(1 - \frac{e^{-x}}{n}\right)^n = \exp\left(n \ln\left(1 - \frac{e^{-x}}{n}\right)\right).$$

Or, lorsque $n \rightarrow +\infty$, $\frac{e^{-x}}{n} \rightarrow 0$ et donc $\ln\left(1 - \frac{e^{-x}}{n}\right) \underset{n \rightarrow +\infty}{\sim} -\frac{e^{-x}}{n}$.

On en déduit que $n \ln\left(1 - \frac{e^{-x}}{n}\right) \underset{n \rightarrow +\infty}{\sim} -e^{-x} \xrightarrow{n \rightarrow +\infty} -e^{-x}$.

Et donc, par continuité de l'exponentielle, $P(M_n \leq x) \xrightarrow{n \rightarrow +\infty} e^{-e^{-x}} = F_G(x)$.

Et donc $(M_n)_{n \geq 1}$ converge en loi vers une variable aléatoire X de densité g .

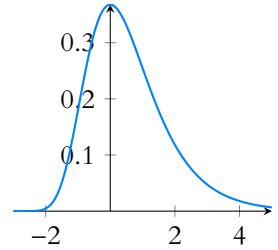


FIGURE 23.1– La densité g

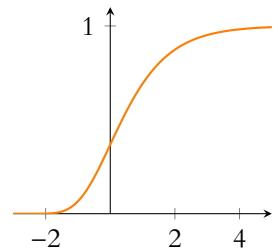


FIGURE 23.2– La fonction F_G .

☠ Forme indé.

On a bien $1 - \frac{e^{-x}}{n} \xrightarrow{n \rightarrow +\infty} 1$, mais il ne faut pas en déduire que

$$\left(1 - \frac{e^{-x}}{n}\right)^n \rightarrow 1,$$

car 1^∞ est une forme indéterminée. Pour lever l'indétermination, on repassera systématiquement par la forme exponentielle.

EXERCICE 2 (EDHEC 2009)

1. Si (X_n) converge en moyenne vers X , alors on a $\lim_{n \rightarrow +\infty} E(|X_n - X|) = 0$.

Mais, pour tout $\varepsilon > 0$, en appliquant l'inégalité de Markov à la variable $|X_n - X|$, qui est positive, et qui admet bien une espérance par hypothèse, on a

$$0 \leq P(|X_n - X| \geq \varepsilon) \leq \frac{E(|X_n - X|)}{\varepsilon} \xrightarrow{n \rightarrow +\infty} 0.$$

Par le théorème des gendarmes, on en déduit que

$$\forall \varepsilon > 0, P(|X_n - X| \geq \varepsilon) \xrightarrow{n \rightarrow +\infty} 0$$

et donc $X_n \xrightarrow{P} X$.

- 2.a. Un produit est nul si et seulement si l'un des facteurs est nul, donc $[Y_n \neq 0] = \bigcap_{k=1}^n [Z_k \neq 0]$.

Par indépendance des Z_k , il vient alors

$$P(Y_n \neq 0) = \prod_{k=1}^n P(Z_k \neq 0) = P(Z_1 \neq 0)^n.$$

Mais on a $P(Z_1 \neq 0) = 1 - P(Z_1 = 0) = 1 - e^{-\lambda}$. Et donc

$$P(Y_n \neq 0) = (1 - e^{-\lambda})^n.$$

- 2.b. Si $Y_n > \varepsilon$, alors $Y_n \neq 0$, de sorte que $[Y_n > \varepsilon] \subset [Y_n \neq 0]$.
 2.c. Notons que Y_n est une variable aléatoire à valeurs positives, et donc $|Y_n| = Y_n$. Pour $\varepsilon > 0$ fixé, on a

$$0 \leq P(|Y_n - 0| \geq \varepsilon) = P(Y_n \geq \varepsilon) \leq P(Y_n \neq 0) \leq (1 - e^{-\lambda})^n \xrightarrow{n \rightarrow +\infty} 0.$$

Par le théorème des gendarmes, on en déduit que

$$\forall \varepsilon > 0, \lim_{n \rightarrow +\infty} P(|Y_n - 0| \geq \varepsilon) = 0$$

et donc $Y_n \xrightarrow{P} 0$.

- 3.a. Si (Y_n) convergeait en moyenne vers une variable aléatoire Y , alors par la question 1 (Y_n) devrait converger en probabilité vers Y . Comme l'énoncé nous explique que la limite est presque sûrement unique, on doit avoir $P(Y = 0) = 1$.
 3.b. Les variables Z_k étant indépendantes, Y_n admet une espérance et

$$E(Y_n) = E\left(\prod_{k=1}^n Z_k\right) = \prod_{k=1}^n E(Z_k) = \lambda^n.$$

- 3.c. On a $Y_n - Y \leq |Y_n - Y|$ et donc, par croissance de l'espérance,

$$E(Y_n) - E(Y) \leq E(|Y_n - Y|).$$

Puisque $E(Y) = 0$, on en déduit que $E(|Y_n - Y|) \geq E(Y_n) - 0 = \lambda^n \xrightarrow{n \rightarrow +\infty} +\infty$.

Et alors nécessairement, $\lim_{n \rightarrow +\infty} E(|Y_n - Y|) = +\infty$.

4. Nous venons de prouver que $E(|Y_n - Y|)$ ne tend pas vers 0, et donc (Y_n) ne converge pas en moyenne.

Hypothèses

L'inégalité de Markov nécessite d'avoir une variable positive, donc avant de l'appliquer, on n'oubliera pas de vérifier que c'est le cas, et on le mentionnera clairement.

Inclusion

Pour deux événements A et B , $A \subset B$ signifie «si A est réalisé, alors B est réalisé».

Rappel

Si X et Y sont indépendantes et admettent une espérance, alors XY admet aussi une espérance et

$$E(XY) = E(X)E(Y).$$

Valeur absolue

Pour tout nombre x , on a $x \leq |x|$.
 S'il est positif, c'est trivial car $|x| = x$. Et sinon, $|x| \geq 0$, donc le résultat est immédiat.

DEVOIR MAISON 14 (À RENDRE LE 05.02.19)

Vous traiterez **au choix** l'un des deux sujets suivants, le second étant plus difficile que le premier.

EXERCICE 1

Soit n un entier naturel et soit $E = \mathbf{R}_n[X]$ l'ensemble des polynômes de degré inférieur ou égal à n .
Pour tout polynôme P de E , on pose : $\varphi(P) = P'' - 2XP'$.

1.
 - a. Montrer que φ est un endomorphisme de E .
 - b. Déterminer la matrice associée à φ dans la base canonique de E .
 - c. Montrer que φ est diagonalisable et donner ses valeurs propres.
2. Pour tout (P, Q) de E^2 , on note :

$$\langle P, Q \rangle = \int_{-\infty}^{+\infty} P(t)Q(t)e^{-t^2} dt.$$

- a. Montrer que pour tout $(P, Q) \in E^2$, $\langle P, Q \rangle$ est bien défini.
 - b. Montrer que $(P, Q) \mapsto \langle P, Q \rangle$ définit un produit scalaire sur E .
3. Montrer que φ est un endomorphisme symétrique pour le produit scalaire $\langle \cdot, \cdot \rangle$.
4. On définit une famille (P_0, P_1, \dots, P_n) de polynômes de E par :

$$P_0 = 1 \text{ et } \forall k \in \{1, 2, \dots, n\}, P_k = X^k - \sum_{i=0}^{k-1} \frac{\langle P_i, X^k \rangle}{\langle P_i, P_i \rangle} P_i.$$

- a. Montrer que (P_0, P_1, \dots, P_n) est une base orthogonale de $\mathbf{R}_n[X]$.
Indication : montrer par récurrence sur $k \leq n$ que (P_0, P_1, \dots, P_k) est orthogonale.
 - b. Montrer que pour tout $k \in \llbracket 0, n \rrbracket$, il existe un unique $(k+1)$ -uplet de réels $(\alpha_0, \dots, \alpha_k)$ tels que $\varphi(P_k) = \sum_{i=0}^k \alpha_i P_i$.
 - c. En calculant de deux manières différentes $\langle \varphi(P_k), P_i \rangle$, pour $0 \leq i \leq k \leq n$, montrer que les P_k , $k \in \llbracket 0, n \rrbracket$ sont des vecteurs propres de φ .
 - d. En déduire une base orthonormée de E formée de vecteurs propres de φ .

EXERCICE 2

Soit E un espace euclidien de dimension n , dont le produit scalaire est noté $\langle \cdot, \cdot \rangle$.
Soit u un endomorphisme non nul de E , antisymétrique c'est-à-dire vérifiant :

$$\forall (x, y) \in E^2, \langle u(x), y \rangle = -\langle x, u(y) \rangle.$$

On admet (voir DM 8) qu'un endomorphisme est antisymétrique si et seulement si sa matrice dans une base orthonormée est une matrice antisymétrique.

Partie I. Propriétés générales des endomorphismes antisymétriques.

1. Soit λ un nombre réel. Montrer que si λ est valeur propre de u , alors $\lambda = 0$.
2. Montrer que $\text{Im } u$ et $\text{Ker } u$ sont orthogonaux et supplémentaires dans E .
En déduire que $\text{Ker}(u) = \text{Ker}(u^2)$.
3. Montrer que u^2 est un endomorphisme symétrique de E et que toute valeur propre de u^2 est négative ou nulle.
4.
 - a. Montrer que u^2 admet au moins une valeur propre non nulle.
Soient x un vecteur propre de u^2 associé à une valeur propre non nulle, et F le sous-espace vectoriel de E engendré par $(x, u(x))$.
 - b. Montrer que $\dim F = 2$ et que F est stable par u .
 - c. Montrer que F^\perp , le supplémentaire orthogonal de F est stable par u .
 - d. On définit l'endomorphisme u_1 de F^\perp par : $\forall x \in F^\perp, u_1(x) = u(x)$.
Montrer que u_1 est un endomorphisme antisymétrique de F^\perp , et que $\text{Im } u = F \oplus \text{Im } u_1$.
5. Montrer que le rang d'un endomorphisme antisymétrique est pair. On pourra faire une récurrence sur la dimension de E .

Partie II : Application.

Dans cette partie, E est un espace euclidien de dimension 4 et $\mathcal{B} = (e_1, e_2, e_3, e_4)$ est une base orthonormale de E . Soit u l'endomorphisme de E dont la matrice dans la base \mathcal{B} est

$$A = \begin{pmatrix} 0 & 4 & 1 & -1 \\ -4 & 0 & -1 & -1 \\ -1 & 1 & 0 & -5 \\ 1 & 1 & 5 & 0 \end{pmatrix}.$$

6. Montrer que u est un endomorphisme antisymétrique de E .
Vérifier que le vecteur $f_1 = e_1 + e_2 - e_3$ est vecteur propre de u^2 .
7. Soit F le sous-espace vectoriel de E engendré par la famille $(f_1, u(f_1))$. Déterminer une base orthonormale de F et une base orthonormale de F^\perp .
8. En déduire une base orthonormale \mathcal{B}_0 de E et deux nombres réels a et b tels que la matrice de u dans la base \mathcal{B}_0 soit

$$\begin{pmatrix} 0 & -a & 0 & 0 \\ a & 0 & 0 & 0 \\ 0 & 0 & 0 & -b \\ 0 & 0 & b & 0 \end{pmatrix}.$$

CORRECTION DU DEVOIR MAISON 14

EXERCICE 1 (ÉCRICOME 2015)

- 1.a. Soit $P \in \mathbf{R}_n[X]$. Alors $\deg P \leq n$, de sorte que $\deg P' \leq n-1$ et $\deg P'' \leq n-2$.
Et alors $\deg(XP') \leq n$ de sorte que

$$\deg \varphi(P) = \deg(P'' - 2XP') \leq n.$$

Ainsi, φ est bien une application de $\mathbf{R}_n[X]$ dans lui-même.

Soient $P, Q \in \mathbf{R}_n[X]$ et $\lambda \in \mathbf{R}$. Alors

$$\varphi(\lambda P + Q) = (\lambda P + Q)'' - 2X(\lambda P + Q)' = \lambda(P'' - 2XP') + (Q'' - 2XQ') = \lambda\varphi(P) + \varphi(Q).$$

Donc φ est une application linéaire de $\mathbf{R}_n[X]$ dans lui-même : c'est un endomorphisme de $\mathbf{R}_n[X]$.

- 1.b. On a $\varphi(1) = 0$, $\varphi(X) = -2X$, et pour $k \in \llbracket 2, n \rrbracket$, $\varphi(X^k) = k(k-1)X^{k-2} - 2kX^k$, de sorte que la matrice de φ dans la base canonique de $\mathbf{R}_n[X]$ est

$$M = \begin{pmatrix} \varphi(1) & \varphi(X) & \varphi(X^2) & \varphi(X^3) & \dots & \varphi(X^{n-1}) & \varphi(X^n) \\ 0 & 0 & 2 & 0 & \dots & \dots & 0 \\ 0 & -2 & 0 & 6 & \dots & \dots & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & & \vdots \\ \vdots & & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & & \ddots & \ddots & \ddots & n(n-1) \\ \vdots & & & & \ddots & \ddots & 0 \\ 0 & \dots & \dots & \dots & \dots & 0 & -2n \end{pmatrix} \begin{matrix} 1 \\ X \\ X^2 \\ \vdots \\ X^{n-2} \\ X^{n-1} \\ X^n \end{matrix} \in \mathcal{M}_{n+1}(\mathbf{R}).$$

- 1.c. La matrice M est triangulaire supérieure, et ses coefficients diagonaux sont $0, -2, -4, \dots, -2n$.
Ce sont les valeurs propres de M qui sont alors au nombre de $n+1$.

Ainsi, M est une matrice carrée de taille $n+1$ possédant $n+1$ valeurs propres distinctes : elle est diagonalisable.

Et puisque M est la matrice de φ dans la base canonique de E , φ est diagonalisable.

- 2.a. La fonction $t \mapsto P(t)Q(t)e^{-t^2}$ est continue sur \mathbf{R} , donc les éventuels problèmes de convergence de l'intégrale sont en $\pm\infty$.

Notons $a_k X^k$ le terme dominant de $P(X)$ et $b_i X^i$ le coefficient dominant de $Q(X)$

Alors au voisinage de $\pm\infty$, on a

$$t^2 P(t)Q(t)e^{-t^2} \sim t^2 a_k t^k b_i t^i e^{-t^2} \sim a_k b_i t^{k+i+2} e^{-t^2} \xrightarrow[t \rightarrow \pm\infty]{} 0.$$

Donc $P(t)Q(t)e^{-t^2} = o\left(\frac{1}{t^2}\right)$, et puisque $\int_1^{+\infty} \frac{dt}{t^2}$ converge, il en est de même de $\int_1^{+\infty} P(t)Q(t)e^{-t^2} dt$.

De même, $\int_{-\infty}^{-1} P(t)Q(t)e^{-t^2} dt$ converge car $\int_{-\infty}^{-1} \frac{dt}{t^2}$ converge, et donc $\int_{-\infty}^{+\infty} P(t)Q(t)e^{-t^2} dt$ converge.

Ainsi, pour tout $(P, Q) \in E^2$, $\langle P, Q \rangle$ est bien défini.

- 2.b. Commençons par remarquer que $\langle \cdot, \cdot \rangle$ est bien une application de $\mathbf{R}_n[X] \times \mathbf{R}_n[X]$ dans \mathbf{R} .

Soient $P, Q, R \in \mathbf{R}_n[X]$ et $\lambda \in \mathbf{R}$. Alors

$$\begin{aligned} \langle \lambda P + Q, R \rangle &= \int_{-\infty}^{+\infty} (\lambda P(t) + Q(t))R(t)e^{-t^2} dt \\ &= \lambda \int_{-\infty}^{+\infty} P(t)R(t)e^{-t^2} dt + \int_{-\infty}^{+\infty} Q(t)R(t)e^{-t^2} dt \\ &= \lambda \langle P, R \rangle + \langle Q, R \rangle \end{aligned}$$

Remarque

On peut supposer a_k et b_i non nuls car si $P = 0$ ou $Q = 0$, alors le résultat est évident.

Ainsi, $\langle \cdot, \cdot \rangle$ est linéaire par rapport à la première variable.

Pour $P, Q \in \mathbf{R}_n[X]$,

$$\langle P, Q \rangle = \int_{-\infty}^{+\infty} P(t)Q(t)e^{-t^2} dt = \int_{-\infty}^{+\infty} Q(t)P(t)e^{-t^2} dt = \langle Q, P \rangle.$$

Donc $\langle \cdot, \cdot \rangle$ est symétrique, et par conséquent bilinéaire symétrique.

Si $P \in \mathbf{R}_n[X]$, la fonction $t \mapsto P^2(t)e^{-t^2}$ est continue et positive sur \mathbf{R} , donc par positivité de l'intégrale,

$$\langle P, P \rangle = \int_{-\infty}^{+\infty} P^2(t)e^{-t^2} dt \geq 0.$$

Enfin, si $\langle P, P \rangle = 0 \Leftrightarrow \int_{-\infty}^{+\infty} P^2(t)e^{-t^2} dt = 0$, puisque la fonction $t \mapsto P^2(t)e^{-t^2}$ est continue et positive sur \mathbf{R} , on a

$$\forall t \in \mathbf{R}, P^2(t)e^{-t^2} = 0 \Leftrightarrow \forall t \in \mathbf{R}, P(t) = 0.$$

Et donc $P = 0$.

Ainsi, $\langle \cdot, \cdot \rangle$ est un produit scalaire sur $\mathbf{R}_n[X]$.

3. Soient $P, Q \in E^2$. Alors

$$\langle \varphi(P), Q \rangle = \int_{-\infty}^{+\infty} (\varphi(P))(t)Q(t)e^{-t^2} dt = \int_{-\infty}^{+\infty} (P''(t) - 2tP'(t))Q(t)e^{-t^2} dt$$

Soient $A \leq B$. Alors

$$\int_A^B (P''(t) - 2tP'(t))Q(t)e^{-t^2} dt = \int_A^B P''(t)Q(t)e^{-t^2} dt - \int_A^B 2tP'(t)Q(t)e^{-t^2} dt$$

Procédons à une intégration par parties dans la première intégrale en posant $u(t) = P'(t)$ et $v(t) = Q(t)e^{-t^2}$, qui sont deux fonctions de classe \mathcal{C}^1 sur $[A, B]$, avec $u'(t) = P''(t)$ et $v'(t) = (Q'(t) - 2tQ(t))e^{-t^2}$. Alors

$$\begin{aligned} \int_A^B (P''(t) - 2tP'(t))Q(t)e^{-t^2} dt &= [P'(t)Q(t)e^{-t^2}]_A^B - \int_A^B P'(t)Q'(t)e^{-t^2} dt \\ &\quad + \int_A^B 2tP'(t)Q(t)e^{-t^2} dt - \int_A^B 2tP'(t)Q(t)e^{-t^2} dt. \end{aligned}$$

Lorsque $A \rightarrow -\infty$, alors $P'(A)Q(A)e^{-A^2} \rightarrow 0$ de sorte que

$$\int_{-\infty}^B (P''(t) - 2tP'(t))Q(t)e^{-t^2} dt = P'(B)Q(B)e^{-B^2} - \int_{-\infty}^B P'(t)Q'(t)e^{-t^2} dt.$$

Puis en faisant tendre B vers $+\infty$, il vient

$$\langle \varphi(P), Q \rangle = \int_{-\infty}^{+\infty} (P''(t) - 2tP'(t))Q(t)e^{-t^2} dt = - \int_{-\infty}^{+\infty} P'(t)Q'(t)e^{-t^2} dt.$$

De même, en inversant les rôles de P et Q , on a

$$\langle P, \varphi(Q) \rangle = \langle \varphi(Q), P \rangle = - \int_{-\infty}^{+\infty} Q'(t)P'(t)e^{-t^2} dt = \langle \varphi(P), Q \rangle.$$

Ceci prouve donc que φ est un endomorphisme symétrique de $\mathbf{R}_n[X]$.

4.a. Montrons par récurrence sur $i \leq n$ que (P_0, \dots, P_i) est une famille orthogonale de E . Pour $i = 0$, le résultat est évident : toute famille formée d'un seul vecteur non nul est orthogonale.

Rédaction

Il est important de mentionner à la fois la positivité et la continuité de l'intégrande. En effet, le résultat

$$\int_{-\infty}^{+\infty} f(t) dt = 0 \Leftrightarrow \forall t \in \mathbf{R}, f(t) = 0$$

nécessite ces deux hypothèses, et n'est plus vrai dès que l'on enlève une des deux hypothèses.

Détails

Si P est un polynôme, alors au voisinage de $+\infty$, $P(t)$ est équivalent à son terme de plus haut degré, qui est de la forme at^n .

Et alors, en posant, $x = t^2$, on a $P(t)e^{-t^2} \sim at^n e^{-t^2}$ et $at^n e^{-t^2} = ax^{n/2} e^{-x} \xrightarrow{x \rightarrow +\infty} 0$

Supposons donc que (P_0, \dots, P_i) est orthogonale.

Pour tout $j \in \llbracket 0, i \rrbracket$, on a

$$\langle P_{i+1}, P_j \rangle = \left\langle X^{i+1} - \sum_{\ell=0}^i \frac{\langle P_\ell, X^{i+1} \rangle}{\langle P_\ell, P_\ell \rangle} P_\ell, P_j \right\rangle = \langle X^{i+1}, P_j \rangle - \sum_{\ell=0}^i \frac{\langle P_\ell, X^{i+1} \rangle}{\langle P_\ell, P_\ell \rangle} \langle P_\ell, P_j \rangle.$$

Pour $\ell \neq j$, alors $\langle P_\ell, P_j \rangle = 0$ par hypothèse de récurrence.

Ainsi, il ne reste qu'un terme non nul dans la somme, correspondant à $\ell = j$:

$$\langle P_{i+1}, P_j \rangle = \langle X^{i+1}, P_j \rangle - \frac{\langle P_j, X^{i+1} \rangle}{\langle P_j, P_j \rangle} \langle P_j, P_j \rangle = \langle X^{i+1}, P_j \rangle - \langle X^{i+1}, P_j \rangle = 0.$$

Ainsi, P_{i+1} est orthogonal à P_0, \dots, P_i : la famille (P_0, \dots, P_{i+1}) est orthogonale.

Par le principe de récurrence, (P_0, \dots, P_n) est une famille orthogonale de E .

Une famille orthogonale ne contenant pas le vecteur nul est nécessairement libre, donc (P_0, \dots, P_n) est une famille libre de E , de cardinal $n + 1 = \dim E$: c'est une base de E .

Nous avons donc bien prouvé que (P_0, \dots, P_n) est une base orthogonale de E .

Remarques : • Le procédé de construction des P_i ressemble à un procédé de Gram-Schmidt, appliqué à la famille $(1, X, \dots, X^n)$.

Il en diffère toutefois légèrement en ceci que, à chaque étape, on ne normalise pas le vecteur obtenu en le divisant par sa norme. Nous venons de montrer qu'on obtenait alors une famille orthogonale, mais non nécessairement orthonormée, puisque les vecteurs n'ont pas de raison d'être de norme 1.

• En termes de projecteurs orthogonaux, on peut aussi reconnaître que $\sum_{i=0}^k \frac{\langle X^{k+1}, P_i \rangle}{\langle P_i, P_i \rangle} P_i$

est le projeté orthogonal de X^{k+1} sur $\text{Vect}(P_0, P_1, \dots, P_k)$, ce qui est un autre moyen de

prouver que $P_k = X^{k+1} - \sum_{i=0}^k \frac{\langle X^{k+1}, P_i \rangle}{\langle P_i, P_i \rangle} P_i$ est orthogonal à P_0, P_1, \dots, P_k .

4.b. Il est aisé de prouver par récurrence que $\deg P_k = k$.

En effet, $\deg(P_0) = 0$, puis $P_1 = X - \underbrace{\frac{\langle P_0, X \rangle}{\langle P_0, P_0 \rangle}}_{\in \mathbf{R}_0[X]} P_0$ est de degré 1, et plus généralement,

$$P_k = X^k - \underbrace{\sum_{i=0}^{k-1} \frac{\langle X^k, P_i \rangle}{\langle P_i, P_i \rangle} P_i}_{\in \mathbf{R}_{k-1}[X]}$$

Et donc (P_0, P_1, \dots, P_k) est une famille libre de $\mathbf{R}_k[X]$ car formée de polynômes de degrés deux à deux distincts.

Étant de cardinal $k + 1 = \dim \mathbf{R}_k[X]$, c'est une base de $\mathbf{R}_k[X]$.

De plus, il est clair que pour tout $P \in \mathbf{R}_n[X]$, $\deg \varphi(P) \leq \deg(P)$, et donc $\mathbf{R}_k[X]$ est stable par φ .

Ainsi, $\varphi(P_k) \in \mathbf{R}_k[X]$, et donc il existe un unique $(k + 1)$ -uplet de réels $(\alpha_0, \dots, \alpha_k)$ tel que

$$\varphi(P_k) = \sum_{i=0}^k \alpha_i P_i.$$

4.c. Fixons $k \in \llbracket 0, n \rrbracket$, et soit $i \in \llbracket 0, k \rrbracket$. Alors

$$\langle \varphi(P_k), P_i \rangle = \left\langle \sum_{j=0}^k \alpha_j P_j, P_i \right\rangle = \sum_{j=0}^k \alpha_j \underbrace{\langle P_j, P_i \rangle}_{=0 \text{ si } i \neq j} = \alpha_i.$$

D'autre part, φ étant symétrique, $\langle \varphi(P_k), P_i \rangle = \langle P_k, \varphi(P_i) \rangle$.

Mais puisque $\mathbf{R}_i[X]$ est stable par φ , $\varphi(P_i) \in \mathbf{R}_i[X]$. Et donc $\varphi(P_i)$ est combinaison linéaire de P_0, P_1, \dots, P_i .

La famille (P_0, P_1, \dots, P_k) étant orthogonale, on a donc $\langle P_k, \varphi(P_i) \rangle = 0$.

Ainsi, nous avons prouvé que pour $i < k$, $\alpha_i = 0$ et donc $\varphi(P_k) = \alpha_k P_k$.

Et donc¹ P_k est un vecteur propre de φ .

Détails

Les $\frac{\langle P_\ell, X^{i+1} \rangle}{\langle P_\ell, P_\ell \rangle}$ sont des réels.
On peut donc les sortir du produit scalaire par linéarité à gauche.

Détails

Les P_i sont non nuls car $\deg P_i = i$, ce qui se prouve aisément par récurrence sur i .

Rappel

Si (e_1, \dots, e_p) est une base orthonormée de F , alors le projecteur orthogonal sur F est donné par

$$p_F(x) = \sum_{i=1}^p \langle x, e_i \rangle e_i.$$

¹ P_k est non nul.

- 4.d. Nous venons de prouver que (P_0, P_1, \dots, P_n) est une base orthogonale de $\mathbf{R}_n[X]$, formée de vecteurs propres de φ .
 Donc $\left(\frac{P_0}{\|P_0\|}, \dots, \frac{P_n}{\|P_n\|}\right)$ est une base orthonormée de $\mathbf{R}_n[X]$. Elle est formée de vecteurs propres car $\frac{P_k}{\|P_k\|}$ est proportionnel au vecteur propre P_k .

Remarque : une étude un peu plus poussée prouve qu'en fait $\lambda_k = -2k$.
 Moyen simple de le constater est que la matrice de la restriction de φ à $\mathbf{R}_{k-1}[X]$ est triangulaire avec pour valeurs propres $0, -2, \dots, -2(k-1)$, alors que la matrice de la restriction de φ à $\mathbf{R}_k[X]$ est triangulaire avec pour valeurs propres $0, -2, \dots, -2k$.
 Donc φ possède un vecteur propre associé à la valeur propre $-2k$ dans $\mathbf{R}_k[X]$ et pas dans $\mathbf{R}_{k-1}[X]$. Et donc un tel vecteur propre est nécessairement de degré k . Puisque P_k est un vecteur propre de φ et qu'il est de degré k , c'est donc un vecteur propre associé à la valeur propre $-2k$.

EXERCICE 2 (EML 2002)

Partie I. Propriétés générales des endomorphismes antisymétriques.

1. Soit λ une valeur propre de u et soit x un vecteur propre associé.

Alors $\langle x, f(x) \rangle = 0$ car u est antisymétrique.
 Mais d'autre part, $\langle x, u(x) \rangle = \langle x, \lambda x \rangle = \lambda \langle x, x \rangle = \lambda \|x\|^2$.
 Puisque x est non nul², on a donc $\|x\|^2 \neq 0$, et par conséquent $\lambda = 0$.

2. Soit $x \in \text{Ker } u$ et $y \in \text{Im } u$. Alors il existe $z \in E$ tel que $y = u(z)$. Et alors

$$\langle x, y \rangle = \langle x, u(z) \rangle = \langle u(x), z \rangle = \langle 0_E, z \rangle = 0.$$

Ceci prouve donc que $\text{Ker } u \subset (\text{Im } u)^\perp$.
 Mais nous savons également que

$$\dim \text{Ker } u = \dim E - \dim \text{Im}(u) = \dim(\text{Im } u)^\perp.$$

Et donc puisque $\text{Ker } u$ possède la même dimension que $(\text{Im } u)^\perp$, tout en étant inclus dans $(\text{Im } u)^\perp$, on a $\text{Ker } u = (\text{Im } u)^\perp$.

Et donc $\text{Ker } u$ est le supplémentaire orthogonal de $\text{Im } u$.

On a toujours $\text{Ker } u \subset \text{Ker } u^2$, que l'endomorphisme u soit antisymétrique ou non.
 En effet, si $x \in \text{Ker } u$, alors $u(x) = 0_E$ et donc $u^2(x) = u(u(x)) = u(0_E) = 0_E$.
 Inversement, supposons que $x \in \text{Ker } (u^2)$. Alors $u(u(x)) = u^2(x) = 0_E$, de sorte que $u(x) \in \text{Ker } u$.

Et d'autre part, $u(x) \in \text{Im}(u)$. Autrement dit, $u(x) \in \text{Ker}(u) \cap \text{Im}(u)$.
 Or ces deux espaces étant supplémentaires, leur intersection est réduite au vecteur nul : $\text{Ker}(u) \cap \text{Im}(u) = \{0_E\}$. Et donc $u(x) = 0_E : x \in \text{Ker}(u)$.

Ainsi, on a bien $\text{Ker}(u) = \text{Ker}(u^2)$.

3. Soient $(x, y) \in E^2$. Alors

$$\langle u^2(x), y \rangle = -\langle u(x), u(y) \rangle = -(-\langle x, u^2(y) \rangle) = \langle x, u^2(y) \rangle.$$

Ainsi, u^2 est un endomorphisme symétrique.

Soit λ une valeur propre de u^2 , et soit x un vecteur propre associé. Alors

$$\|u(x)\|^2 = \langle u(x), u(x) \rangle = -\langle u^2(x), x \rangle = -\langle \lambda x, x \rangle = -\lambda \|x\|^2.$$

Puisque $\|u(x)\|^2 \geq 0$ et que $\|x\|^2 > 0$, on a donc $\lambda = \frac{-\|u(x)\|^2}{\|x\|^2} \leq 0$.

Et donc toute valeur propre de u^2 est négative ou nulle.

- 4.a. Raisonnons par l'absurde et supposons que la seule valeur propre de u^2 soit 0.
 Puisque u^2 est diagonalisable, car symétrique, on a $\sum_{\lambda \in \text{Spec}(u^2)} \dim E_\lambda(u^2) = \dim E = n$.

Soit encore $\dim E_0(u^2) = n \Leftrightarrow \dim \text{Ker}(u^2) = n$.
 Et donc $\text{Ker}(u^2) = E$, ce qui signifie que u^2 est l'endomorphisme nul.

Inclusion

Cela prouve également que $\text{Im } u \subset (\text{Ker } u)^\perp$, mais ces deux inclusions (qui sont en fait équivalentes) ne suffisent pas à prouver que $\text{Ker } u = (\text{Im } u)^\perp$.
 En effet, nous savons que tous les vecteurs de $\text{Ker } u$ sont orthogonal à tous les vecteurs de $\text{Im } u$, mais pour l'instant rien n'empêche qu'il existe des vecteurs orthogonaux à tous les vecteurs de $\text{Im } u$ mais qui ne sont pas dans $\text{Ker } u$.
 Nous n'en savons rien, tout bonnement car nous ne nous sommes pas posé la question, et ne nous sommes intéressés qu'aux vecteurs de $\text{Ker } u$.

Remarque

Cette égalité est vraie dès que $\text{Ker}(u)$ et $\text{Im}(u)$ sont supplémentaires, l'orthogonalité n'est ici pas utile.

Alternative

Si vous préférez travailler avec des matrices : la matrice de u^2 dans une base de vecteurs propres est diagonale, et ne possède que des 0 sur la diagonale. Donc c'est la matrice nulle, ce qui signifie que $u^2 = 0$.

Or nous avons prouvé précédemment que $\text{Ker}(u) = \text{Ker}(u^2)$, et donc $\text{Ker}(u) = E$, de sorte que u est l'endomorphisme nul.

Mais par hypothèse, $u \neq 0$, de sorte que notre hypothèse est absurde : u^2 possède au moins une valeur propre non nulle.

4.b. Notons λ la valeur propre de u^2 associée à x .

Puisque $(x, u(x))$ est une famille génératrice de F , celui-ci est de dimension au plus 2.

Supposons par l'absurde que F est de dimension 1. Alors la famille formée du seul vecteur x est libre³, et étant de cardinal $1 = \dim F$, c'est une base de F .

³ Car $x \neq 0_E$.

Mais puisque $u(x) \in F$, il existe $\alpha \in \mathbf{R}$ tel que $u(x) = \alpha x$. Et donc x est un vecteur propre de u , associé à la valeur propre α .

Mais nous avons prouvé à la question 1 que la seule valeur propre de u est 0, donc $\alpha = 0$ et donc $u(x) = 0_E$, de sorte que $u^2(x) = 0_E$.

Or, $u^2(x) = \lambda x$, avec $\lambda \neq 0$ et $x \neq 0_E$, de sorte que $u^2(x) \neq 0_E$.

Ceci est absurde, et donc F est de dimension 2.

Pour $y \in F$, il existe deux réels α_1 et α_2 tels que $y = \alpha_1 x + \alpha_2 u(x)$. Et donc

$$u(x) = u(\alpha_1 x + \alpha_2 u(x)) = \alpha_1 u(x) + \alpha_2 u^2(x) = \alpha_1 u(x) + \lambda \alpha_2 x \in F.$$

Et donc F est stable par u .

4.c. Soit $x \in F^\perp$. Alors, pour tout $y \in F$, on a

$$\langle u(x), y \rangle = -\langle x, u(y) \rangle.$$

Mais $u(y) \in F$ car F est stable par u , et donc $\langle x, u(y) \rangle = 0$ car $x \in F^\perp$.

Ainsi, $\langle u(x), y \rangle = 0$, de sorte que $u(x) \in F^\perp$. Et donc F^\perp est stable par u .

4.d. Notons que u_1 est la restriction de u à F^\perp : ce n'est rien d'autre que u , mais vu comme un endomorphisme de F^\perp et non plus de E tout entier.

Alors en particulier, pour $x, y \in F^\perp$, on a

$$\langle u_1(x), y \rangle = \langle u(x), y \rangle = -\langle x, u(y) \rangle = -\langle x, u_1(y) \rangle.$$

Donc u_1 est antisymétrique.

Notons que $F \subset \text{Im } u$. En effet, puisque $u(x) = \lambda x$, avec $\lambda \neq 0$, alors $x = \frac{1}{\lambda} u(x)$ et donc

$$x = u\left(\frac{1}{\lambda} x\right) \in \text{Im}(u).$$

Puisque d'autre part $u(x) \in \text{Im}(u)$, on a $F = \text{Vect}(x, u(x)) \subset \text{Im}(u)$.

Comme $\text{Im } u_1$ est évidemment⁴ inclus dans $\text{Im } u$, on a donc $F + \text{Im}(u_1) \subset \text{Im}(u)$.

⁴ Si $x = u_1(y) \in \text{Im}(u_1)$, alors

$$x = u(y) \in \text{Im}(u).$$

D'autre part, $\text{Im}(u_1) \subset F^\perp$ et $F \cap F^\perp = \{0_E\}$, de sorte que $\text{Im}(u_1) \cap F = \{0_E\}$.

Donc la somme $F + \text{Im}(u_1)$ est directe.

Enfin, si $y \in \text{Im}(u)$, alors il existe $z \in E$ tel que $y = u(z)$.

Or, z s'écrit de manière unique $z = z_F + z_{F^\perp}$ avec $z \in F$ et $z_{F^\perp} \in F^\perp$. Et donc $y = u(z) = u(z_F) + u(z_{F^\perp}) = u(z_F) + u_1(z_{F^\perp})$.

Puisque F est stable par u , $u(z_F) \in F$, et donc

$$y = \underbrace{u(z_F)}_{\in F} + \underbrace{u_1(z_{F^\perp})}_{\in \text{Im}(u_1)} \in F \oplus \text{Im}(u_1).$$

Donc $\text{Im}(u) \subset F \oplus \text{Im}(u_1)$, et l'inclusion réciproque ayant déjà été prouvée,

$$\text{Im}(u) = F \oplus \text{Im}(u_1).$$

5. Montrons par récurrence double sur $\dim E$ que si $u \in \mathcal{L}(E)$ est antisymétrique, alors $\text{rg}(u)$ est pair.

Plus précisément, soit $\mathcal{P}(n)$ la proposition : «pour tout espace euclidien E de dimension n et pour tout $u \in \mathcal{L}(E)$ antisymétrique, alors $\text{rg}(u)$ est pair».

Si $\dim E = 1$, et si u est un endomorphisme antisymétrique de E , alors la matrice de u dans une base orthonormée de E est une matrice antisymétrique. Mais la seule matrice

antisymétrique de $\mathcal{M}_1(\mathbf{R})$ est la matrice nulle⁵, et donc $u = 0$, qui est évidemment de rang pair.

Donc $\mathcal{P}(1)$ est vraie.

Si $\dim E = 2$, et si u est un endomorphisme antisymétrique de E , alors soit $u = 0$, auquel cas $\text{rg}(u) = 0$ est pair.

Soit $u \neq 0$, et alors il existe une valeur propre $\lambda \neq 0$ de u^2 . Notons x un vecteur propre associé, et soit $F = \text{Vect}(x, u(x))$. Alors nous avons prouvé à la question 4 que $\text{Im } u = F$ et que $\dim F = 2$, donc $\text{rg}(u) = 2$ est pair.

Ainsi, $\mathcal{P}(2)$ est vraie.

Supposons à présent que $\mathcal{P}(n - 1)$ et $\mathcal{P}(n)$ soient vraies, et soit E un espace euclidien de dimension $n + 1$, et u un endomorphisme antisymétrique de E .

Si $u = 0$, alors $\text{rg}(u) = 0$ est pair.

Supposons donc que $u \neq 0$.

Par la question 4.a, u^2 admet au moins une valeur propre non nulle λ . Et si x est un vecteur propre associé, alors en posant $F = \text{Vect}(x, u(x))$, F est de dimension 2, et $\text{Im } u = F \oplus \text{Im}(u_1)$, où u_1 est la restriction de u à F^\perp .

Mais u_1 est un endomorphisme antisymétrique de F^\perp , qui est de dimension $n - 1$. Et donc par hypothèse de récurrence, $\dim \text{Im}(u_1) = \text{rg}(u_1)$ est pair.

Et donc $\text{rg}(u) = \dim \text{Im}(u) = \dim F + \dim(\text{Im } u_1)$ est pair car somme de deux entiers pairs.

Ainsi, $\mathcal{P}(n + 1)$ est vraie.

Par le principe de récurrence, pour tout $n \in \mathbf{N}^*$, $\mathcal{P}(n)$ est vraie, et donc le rang d'un endomorphisme antisymétrique est toujours pair.

Remarque : cette récurrence double est un peu surprenante, car pour prouver $\mathcal{P}(n + 1)$, on n'utilise que $\mathcal{P}(n - 1)$ et pas $\mathcal{P}(n)$.

On pourrait en réalité faire deux récurrences simples : une si $\dim E$ est impair : $\mathcal{P}(1)$ est vraie, donc $\mathcal{P}(3)$ est vraie, donc $\mathcal{P}(5)$ est vraie, etc.

Et une pour $\dim E$ pair : $\mathcal{P}(2)$ est vraie, donc $\mathcal{P}(4)$ est vraie, donc $\mathcal{P}(6)$ est vraie, etc.

Partie II. Application

- 6. La base \mathcal{B} est orthonormée, et ${}^t A = -A$, de sorte que A est antisymétrique. On en déduit donc que u est un endomorphisme antisymétrique de E .

On a alors

$$A^2 \begin{pmatrix} 1 \\ 1 \\ -1 \\ 0 \end{pmatrix} = \begin{pmatrix} -18 & 0 & -9 & -9 \\ 0 & -18 & -9 & 9 \\ -9 & -9 & -27 & 0 \\ -9 & 9 & 0 & -27 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ -1 \\ 0 \end{pmatrix} = \begin{pmatrix} -9 \\ -9 \\ 9 \\ 0 \end{pmatrix} = -9 \begin{pmatrix} 1 \\ 1 \\ -1 \\ 0 \end{pmatrix}.$$

Donc f_1 est un vecteur propre de u^2 , associé à la valeur propre -9 .

- 7. Posons $f_2 = u(f_1) = 3e_1 - 3e_2 - 3e_4$, de sorte que (f_1, f_2) est une base de F .

Alors $x = \alpha_1 e_1 + \alpha_2 e_2 + \alpha_3 e_3 + \alpha_4 e_4 \in F^\perp$ si et seulement si

$$\begin{cases} \langle x, f_1 \rangle = 0 \\ \langle x, f_2 \rangle = 0 \end{cases} \Leftrightarrow \begin{cases} \alpha_1 + \alpha_2 - \alpha_3 = 0 \\ 3\alpha_1 - 3\alpha_2 - 3\alpha_4 = 0 \end{cases} \Leftrightarrow \begin{cases} \alpha_3 = \alpha_1 + \alpha_2 \\ \alpha_4 = \alpha_1 - \alpha_2 \end{cases}$$

$$\Leftrightarrow x = \alpha_1 e_1 + \alpha_2 e_2 + (\alpha_1 + \alpha_2)e_3 + (\alpha_1 - \alpha_2)e_4 = \alpha_1(e_1 + e_3 + e_4) + \alpha_2(e_2 + e_3 - e_4)$$

$$\Leftrightarrow x \in \text{Vect}(e_1 + e_3 + e_4, e_2 + e_3 - e_4).$$

Posons donc $f_3 = e_1 + e_3 + e_4$ et $f_4 = e_2 + e_3 - e_4$, qui est une base⁶ de F^\perp .

Pour obtenir des bases orthonormées de F et de F^\perp , il faut appliquer le procédé d'orthonormalisation de Gram-Schmidt aux deux familles (f_1, f_2) et (f_3, f_4) .

Mais notons que f_1 et f_2 sont déjà orthogonaux, de sorte que $\left(\frac{f_1}{\|f_1\|}, \frac{f_2}{\|f_2\|}\right)$ est une base orthonormée de F .

Donc $\left(\frac{f_1}{\sqrt{3}}, \frac{f_2}{3\sqrt{3}}\right)$ est une base orthonormée de F .

De même, (f_3, f_4) est orthogonale, donc $\left(\frac{f_3}{\|f_3\|}, \frac{f_4}{\|f_4\|}\right) = \left(\frac{f_3}{\sqrt{3}}, \frac{f_4}{\sqrt{3}}\right)$ est une base orthonormée de F^\perp .

Dans la suite, notons cette base $(\varepsilon_1, \varepsilon_2, \varepsilon_3, \varepsilon_4)$.

⁵ Par exemple car les coefficients diagonaux d'une matrice antisymétrique sont nuls.

Rappel
La dimension d'une somme directe est la somme des dimensions.

Astuce
Puisque nous connaissons u via sa matrice, il est bien plus simple de travailler avec des vecteurs colonnes et des matrices.
Et donc plutôt que de calculer $u(f_1)$, on préfère calculer $A \times (f_1)_{\mathcal{B}}$ qui est le vecteur colonne des coordonnées de $u(f_1)$ dans la base \mathcal{B} .

⁶ Les deux vecteurs n'étant pas colinéaires, ils forment une famille libre.

Rappel
Pour une famille orthogonale, le procédé de Gram-Schmidt revient à diviser chaque vecteur par sa norme.

8. La concaténation des deux bases orthonormées obtenues précédemment est une base orthonormée \mathcal{B}_0 de E .

On avait $u(f_1) = f_2$. Et donc $u(\varepsilon_1) = u\left(\frac{f_1}{\sqrt{3}}\right) = \frac{1}{\sqrt{3}}u(f_1) = \frac{1}{\sqrt{3}}(f_2) = 3\varepsilon_2$.

De plus, $u(f_2) = u^2(f_1) = -9f_1$. Et donc

$$u(\varepsilon_2) = \frac{1}{3\sqrt{3}}u(f_2) = -\frac{9}{3\sqrt{3}}f_1 = -\frac{3}{\sqrt{3}}f_1 = -3\varepsilon_1.$$

D'autre part, on a

$$A \begin{pmatrix} 1 \\ 0 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ -6 \\ -6 \\ 6 \end{pmatrix} \text{ et } A \begin{pmatrix} 0 \\ 1 \\ 1 \\ -1 \end{pmatrix} = \begin{pmatrix} 6 \\ 0 \\ 6 \\ 6 \end{pmatrix}$$

de sorte que $u(f_3) = -6f_4$ et $u(f_4) = 6f_3$.

Et donc on a

$$u(\varepsilon_3) = u\left(\frac{f_3}{\sqrt{3}}\right) = \frac{1}{\sqrt{3}}u(f_3) = -\frac{6}{\sqrt{3}}f_4 = -6\varepsilon_4.$$

Et de même,

$$u(\varepsilon_4) = \frac{1}{\sqrt{3}}u(f_4) = \frac{6}{\sqrt{3}}f_3 = 6\varepsilon_3.$$

Donc la matrice de u dans la base \mathcal{B}_0 est

$$\begin{pmatrix} u(\varepsilon_1) & u(\varepsilon_2) & u(\varepsilon_3) & u(\varepsilon_4) \\ 0 & -3 & 0 & 0 \\ 3 & 0 & 0 & 0 \\ 0 & 0 & 0 & 6 \\ 0 & 0 & -6 & 0 \end{pmatrix} \begin{matrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \\ \varepsilon_4 \end{matrix}.$$

Rappel

La concaténation d'une base orthonormée de F et d'une base orthonormée de F^\perp est une base orthonormée de E .

DEVOIR MAISON 15 (SciLab) (À RENDRE LE 14.03.19)

Vous me rendrez ce DM sous forme d'un fichier au format .sci ou .sce que vous m'enverrez par e-mail **au plus tard le 14 mars** à vienny@ecs2-fauriel.fr.

Vous commencerez pour cela par télécharger sur l'ENT le fichier nommé DM15.sci, et le complétez au fur et à mesure. **Ne changez pas les noms des fonctions !**

En revanche, pour les codes à compléter, vous pouvez au besoin changer tout ou partie du code compris entre fonction et endfunction.

Si votre programme ne retourne pas les résultats attendus, indiquez le en commentaire (précédé de //) dans votre fichier, en expliquant les problèmes constatés.

EXERCICE 1

Soit (u_n) la suite définie par récurrence par $u_0 = \frac{1}{2}$ et pour tout $n \in \mathbf{N}$, $u_{n+1} = u_n^2 - u_n \ln(u_n)$.

- Écrire une fonction suite qui prend comme paramètre un entier n et renvoie la valeur de u_n correspondante.
Vérification : $u_{10} \approx 0.8581$.
- On se propose de vérifier graphiquement que la suite (u_n) est croissante et tend vers 1.
Écrire à cet effet un programme qui trace un graphique sur lequel se trouvent les points (n, u_n) , pour $n \in \llbracket 0, 100 \rrbracket$.
- Écrire une fonction nommée plus_petit_n qui prend comme paramètre un entier p et retourne la plus petite valeur de n pour laquelle $|1 - u_n| < 10^{-p}$. Vérification : pour $p = 2$, le programme doit retourner 196.

EXERCICE 2

Soit (u_n) la suite définie par $u_0 = u_1 = u_2 = 1$ et $\forall n \in \mathbf{N}$, $u_{n+3} = u_{n+2} - 2u_{n+1} + u_n$.

Écrire une fonction d'entête fonction $y = \text{suite_rec}(n)$ qui prend comme paramètre $n \geq 3$ et retourne la valeur de u_n .
Vérification : $u_{20} = 82$.

EXERCICE 3

On définit une suite par récurrence en posant $u_0 = 3$ et pour tout $n \in \mathbf{N}$, $u_{n+1} = u_n - \sin(u_n)^2$.

On admet que la série de terme général $(-1)^k u_k$ est convergente, et on note $S = \sum_{k=0}^{+\infty} (-1)^k u_k$.

Pour tout $n \in \mathbf{N}$, on note $S_n = \sum_{k=0}^n (-1)^k u_k$, et on admet que $|S_n - S| \leq u_{n+1}$.

Soit $\varepsilon > 0$. En remarquant que si n est tel que $u_{n+1} \leq \varepsilon$, alors S_n est une valeur approchée de S à ε près, compléter le programme suivant pour qu'il prenne comme paramètre un réel ε et retourne une valeur approchée de S à ε près.

```

1  function --- = approx(epsilon)           6                               u = ---- ;
2      u = 3 ;                             7                               plus_moins = -plus_moins ;
3      plus_moins = 1 ;                    8                               S = ---- ;
4      S = plus_moins * u ;                9                               end
5      while -----                       10  endfunction

```

Vérification : $S \approx 1.3239$.

EXERCICE 4

- Compléter la fonction suivante pour qu'elle simule une réalisation d'une loi géométrique de paramètre p .

```

1  function y = geom(p)
2      y = 1 ;
3      while rand() -----
4          y = ----
5      end
6  endfunction

```

Vérification : pour $p = \frac{1}{2}$, à peu près un appel sur deux à la fonction `geom` doit retourner 1.

- On dit qu'une variable suit la loi binomiale négative de paramètres n et p si elle a la même loi que $\sum_{i=1}^n X_i$ où les X_i sont des variables i.i.d. suivant la loi $\mathcal{G}(p)$.
Écrire un programme d'entête fonction `y = bin_neg(n,p)` qui simule une réalisation d'une loi binomiale négative de paramètres n et p .

EXERCICE 5 : SIMULATION DE LA LOI LOGISTIQUE

Soit F la fonction définie sur \mathbf{R} par $F(x) = \frac{1}{1 + e^{-2x}}$.

- Montrer que F est la fonction de répartition d'une variable aléatoire à densité X , dont on déterminera une densité.
- Déterminer la bijection réciproque de F . Montrer que si $U \leftrightarrow \mathcal{U}([0, 1])$, alors $F^{-1}(U)$ a même loi que X .
- Écrire une fonction d'entête fonction `y = logistique()` qui simule une réalisation de la variable aléatoire X .

EXERCICE 6

Une urne contient 18 boules rouges et 2 boules vertes. On effectue des tirages sans remise dans cette urne, jusqu'à vider l'urne, et on note alors X le rang d'apparition de la première boule verte, et Y le rang d'apparition de la seconde boule verte.

- Compléter la fonction suivante de sorte que la commande `un_tirage(r,v)` simule un tirage dans une urne contenant r boules rouges et v boules vertes, et retourne 0 si la boule tirée est rouge et 1 si la boule tirée est verte.

```

1  fonction y = un_tirage(r,v)           5                               y = ---- ;
2      if rand() < --- then             6                               end
3          y = --- ;                     7  endfunction
4      else

```

- Compléter le programme suivant pour qu'il simule la variable aléatoire X .

```

1  fonction y = tirage_verte()           6                               nb_rouges = ....
2      y=1 ;                             7                               y = ----
3      nb_rouges = 18 ;                   8                               end
4      nb_vertes = 2 ;                    9  endfunction
5      while ---

```

- Adapter le programme de la question précédente pour qu'il simule une réalisation de la variable Y .
- Écrire un programme qui à l'aide de 10 000 simulations de Y , donne une valeur approchée de son espérance et de sa variance.

CORRECTION DU DEVOIR MAISON 15

EXERCICE 1

1. Pas de surprise, il s'agit de calculer les termes successifs de la suite à l'aide d'une boucle.

```

1 function u = suite(n)
2     u = 1/2 ;
3     for i=1 :n
4         u = u^2-u*log(u)
5     end
6 endfunction

```

2. Il s'agit donc de créer deux vecteurs : l'un nommé x contenant les nombres de 0 à 100, et l'autre nommé y contenant u_0, u_1, \dots, u_{100} .

```

1 x = 0 :100 ;
2 y= zeros(1,101) ;
3 for i=1 :101
4     y(i) = suite(i-1)
5 end
6 plot2d(x,y,-1)

```

Danger

Un piège ici : en SciLab les coefficients d'un vecteur sont forcément numérotés à partir de 1. Et donc u_0 se trouvera dans $y(1)$, u_1 dans $y(2)$, etc.

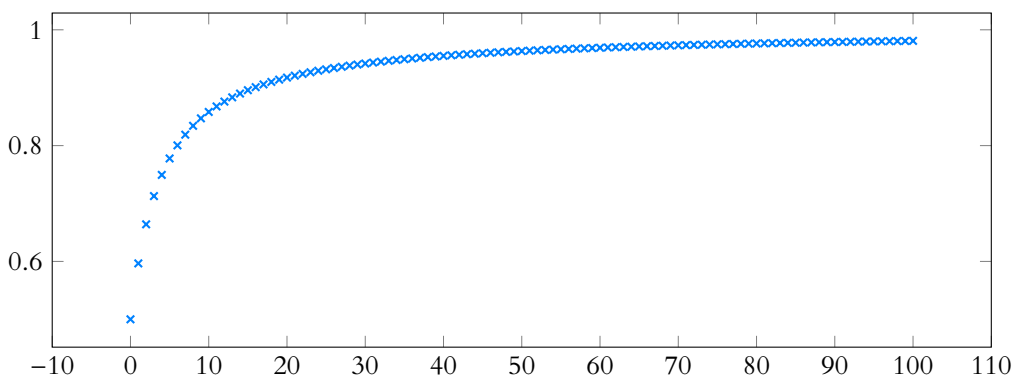


FIGURE 23.1 – Les premiers termes de la suite (u_n). La suite semble assez clairement converger vers un réel proche de 1.

3. Là non plus pas de surprise : on calcule successivement tous les termes de la suite, jusqu'à en trouver un vérifiant la condition requise.

```

1 function y = plus_petit(p)
2     y = 0 ;
3     u=1/2 ;
4     while abs(1-u)>=10^(-p)
5         u = u^2-u*log(u) ;
6         y = y+1 ;
7     end
8 endfunction

```

EXERCICE 2

Donnons deux solutions : une qui garde en mémoire toutes les valeurs de u_n , dans un vecteur u , et une qui à chaque étape ne garde en mémoire que les trois derniers termes¹ de la suite.

```

1 function y = suite_rec(n)
2     u = zeros(1,n+1) ;
3     u(1) = 1 ;
4     u(2) = 1 ;

```

¹ Les seuls nécessaires au calcul du terme suivant.


```

5     u(3) = 1 ;
6     for i=4 :n+1
7         u(i) = u(i-1) - 2*u(i-2) + u(i-3) ;
8     end
9     y = u(n+1) ;
10  endfunction

```

Une autre solution est bien plus similaire à l'exercice 6 du TP8, et nécessite de faire appel à une variable temporaire² pour garder en mémoire la valeur d'une variable avant de modifier cette variable.

L'idée du programme suivant est que si c contient la valeur de u_n , alors b contient u_{n-1} et a contient u_{n-2} .

Lorsqu'on calcule un «nouveau» terme de la suite, c'est-à-dire u_{n+1} , alors on affecte à c l'ancienne valeur contenue dans b (c'est-à-dire u_{n-1}) et on affecte à b l'ancienne valeur contenue dans a (c'est-à-dire u_n).

```

1  function y =suite_rec(n)
2      a = 1 ;
3      b = 1 ;
4      c = 1 ;
5      for i=3 :n
6          v = c ;
7          c = a-2*b+c ;
8          a = b ;
9          b = v ;
10     end
11     y = c
12 endfunction

```

EXERCICE 3

Il s'agit de calculer successivement les termes de la suite (u_n) jusqu'à en obtenir un qui soit inférieur ou égal à ε .

Dans le même temps, on calculera au fur et à mesure les valeurs correspondantes de S_n , à l'aide d'une variable S qui vaut initialement u_1 et à laquelle on ajoute à chaque fois le terme u_n que l'on vient de calculer.

```

1  function S = approx(epsilon)
2      u = 3 ;
3      plus_moins = 1 ;
4      S = plus_moins * u ;
5      while u>=epsilon
6          u = u - sin(u)^2 ;
7          plus_moins = -plus_moins ;
8          S = S + plus_moins*u ;
9      end
10 endfunction

```

EXERCICE 4

- C'est extrêmement classique : si on se souvient qu'une loi géométrique compte le nombre d'essais avant obtention du premier succès, il s'agit de répéter une expérience dont la probabilité de succès vaut p , jusqu'à obtention d'un premier succès. Ici, l'«expérience» consiste à tirer un nombre au hasard à l'aide de `rand()` et on déclarera que cette expérience est un succès si ce nombre est inférieur³ à p .

```

1  function y = geom(p)
2      y = 1
3      while rand()<p
4          y = y+1
5      end
6  endfunction

```

Danger !
Les éléments d'un vecteur étant numérotés à partir de 1, et non 0, il faut bien garder à l'esprit que $u(i)$ contient u_{i-1} et non u_i (puisque $u(1)$ contient u_0).

² C'est celle que nous avons appelée v dans le programme qui suit.

³ Ce qui se produit bien avec probabilité p .

Notons que la boucle `while` sera parcourue autant de fois qu'il y aura eu d'échecs. Et c'est donc pour cette raison qu'il est important d'initialiser `y` à 1, afin de compter également le succès final⁴.

⁴ Pour lequel la ligne 4 ne sera pas exécutée.

2. Rien à signaler ici...

```
1 function y = bin_neg(n,p)
2     y = 0
3     for i=1 :n
4         y = y+ geom(p)
5     end
6 endfunction
```

EXERCICE 5

1. La fonction F est évidemment \mathcal{C}^1 sur \mathbf{R} par opération sur les fonctions usuelles. Elle est croissante car la fonction $x \mapsto 1 + e^{-2x}$ est positive et décroissante, donc son inverse est croissante. Enfin, $\lim_{x \rightarrow -\infty} F(x) = 0$ et $\lim_{x \rightarrow +\infty} F(x) = 1$. Donc F est une fonction de répartition d'une variable à densité, dont une densité est donnée par exemple par

$$F' : x \mapsto \frac{2e^{-2x}}{(1 + e^{-2x})^2}.$$

2. Soit $y \in]0, 1[$. On a alors $y = F(x) \Leftrightarrow y = \frac{1}{1 + e^{-2x}} \Leftrightarrow e^{-2x} = \frac{1}{y} - 1 \Leftrightarrow x = -\frac{1}{2} \ln\left(\frac{1}{y} - 1\right)$.

Et donc la bijection réciproque de F est donnée par $F^{-1} : y \mapsto -\frac{1}{2} \ln\left(\frac{1}{y} - 1\right)$.

On a alors, pour tout $x \in \mathbf{R}$, par croissance⁵ de F^{-1} , $P(F^{-1}(U) \leq x) = P(U \leq F(x))$.

Mais $F(x) \in]0, 1[$, donc $P(U \leq F(x)) = F_U(F(x)) = F(x)$.

Et donc la fonction de répartition de $F^{-1}(U)$ est $F : F^{-1}(U)$ a même loi que X .

3. Il s'agit d'utiliser ce qui a été dit précédemment pour simuler $F^{-1}(U)$, en gardant à l'esprit que U se simule à l'aide de `rand()` :

```
1 function y = simul()
2     y = -1/2*log(1/rand() - 1);
3 endfunction
```

EXERCICE 6

1. La probabilité de tirer une boule verte est $\frac{v}{v+r}$. Donc le programme suivant fait l'affaire puisqu'il retourne 1 avec probabilité $\frac{v}{v+r}$.

```
1 function y = un_tirage(r,v)
2     if rand() < v/(v+r) then
3         y = 1
4     else
5         y = 0
6     end
7 endfunction
```

2. Il s'agit de tirer des boules dans une urne dont la composition initiale est 18 boules rouges et 2 vertes, jusqu'à obtenir une verte, et en gardant à l'esprit que si on obtient une boule rouge, on ne la remet pas dans l'urne.

```
1     function y = tirage_vert()
2         y = 1;
3         nb_rouges=18;
4         nb_vertes = 2;
5         while un_tirage(nb_rouges,nb_vertes)<1
6             nb_rouges = nb_rouges -1
```

Bijection

Nous n'avons pas prouvé que F réalise une bijection de \mathbf{R} sur $]0, 1[$, mais le fait que l'équation $y = F(x)$, possède une unique solution x le prouve ! En effet, cela signifie que tout élément de $]0, 1[$ possède un unique antécédent par F .

⁵ Ce qui garantit que l'on préserve le sens des inégalités.

```

7             y = y+1
8         end
9     endfunction

```

Notons que lors du dernier tirage⁶, la boucle `while` n'est pas parcourue, donc la commande `y=y+1` n'est pas exécutée. Donc si l'on souhaite que ce tirage soit bien pris en compte, il faut initialiser `y` à 1.

⁶ Celui qui donnera une boule verte.

3. Il faut cette fois tirer des boules jusqu'à ce que l'urne ne contienne plus de boules vertes.

```

1 function y = deux_vertes()
2     y = 0 ;
3     nb_rouges = 18 ;
4     nb_vertes = 2 ;
5     while nb_vertes > 0
6         y = y+1 ;
7         tirage = un_tirage(nb_rouges, nb_vertes) ;
8         if tirage == 1 then
9             nb_vertes = nb_vertes - 1 ;
10        else
11            nb_rouges = nb_rouges - 1 ;
12        end
13    end
14 endfunction

```

Notons que cette fois nous avons initialisé `y` à 0, puisque le tirage donnant la seconde boule verte sera bien compté dans `y`, étant donné qu'on augmente `y` à chaque tirage, indépendamment de la couleur de la boule obtenue.

4. Commençons par créer un grand vecteur rempli de 0, puis à l'aide d'une boucle `for`, stockons-y les simulations de Y .

```

1 A = zeros(1, 10000) ;
2 for i=1 : 10000
3     A(i) = deux_vertes() ;
4 end

```

Notons qu'un bon moyen de détecter des erreurs est de s'intéresser à la plus petite et à la plus grande valeur de A , qui s'obtiennent respectivement avec `min(A)` et `max(A)`.

En effet, celles-ci doivent valoir respectivement 2 et 20. Et donc si on obtient des valeurs inférieures à 2 ou supérieures à 20, il y a un problème dans le programme.

Une valeur approchée de l'espérance de Y s'obtient alors avec `mean(A)`, qui retourne environ 14, et une valeur approchée de la variance s'obtient alors à l'aide de `stdev(A)^2`, qui vaut alors environ 20.7.

Gardons en tête qu'il s'agit là de valeurs approchées, et que si nos simulations ne sont pas représentatives de la loi de Y , c'est-à-dire si on a eu un «mauvais» échantillon, on peut en théorie obtenir des valeurs assez éloignées de $E(Y)$ et $V(Y)$.

Faisons le pari que la taille de notre échantillon est assez grande pour que ceci ne se produise pas !

Variance

Rappelons que `stdev` nous donne l'écart-type et non la variance, raison pour laquelle on l'élève au carré.

DEVOIR MAISON 16 (MATHS II) (À RENDRE LE 14.03.19)

Dans tout le problème, N désigne un entier supérieur ou égal à 1.

On note $E(X)$ et $V(X)$ respectivement, l'espérance et la variance, lorsqu'elles existent, de toute variable aléatoire réelle X définie sur un espace probabilisé.

Soit $(U_n)_{n \geq 1}$ une suite de variables aléatoires définies sur un espace probabilisé (Ω, \mathcal{A}, P) , mutuellement indépendantes et de même loi uniforme discrète sur $\llbracket 1, N \rrbracket$.

On pose, pour tout n de \mathbf{N}^* : $T_n = \sup(U_1, U_2, \dots, U_n)$ et $Z_n = \inf(U_1, U_2, \dots, U_n)$. On admet que T_n et Z_n sont des variables aléatoires définies sur (Ω, \mathcal{A}, P) . Ainsi, pour tout ω de Ω , on a

$$T_n(\omega) = \max(U_1(\omega), U_2(\omega), \dots, U_n(\omega)) \text{ et } Z_n(\omega) = \min(U_1(\omega), U_2(\omega), \dots, U_n(\omega)).$$

On rappelle que si C désigne un élément de \mathcal{A} , on note $\mathbb{1}_C$ la variable aléatoire indicatrice de l'événement C , définie sur (Ω, \mathcal{A}, P) par

$$\mathbb{1}_C(\omega) = \begin{cases} 1 & \text{si } \omega \in C \\ 0 & \text{si } \omega \notin C \end{cases}$$

On pose, pour tout n de \mathbf{N}^* , $d_n(N) = \begin{cases} \sum_{k=1}^{N-1} \left(\frac{k}{N}\right)^n & \text{si } N \geq 2 \\ 0 & \text{si } N = 1 \end{cases}$

Préliminaire

- Soit Y une variable aléatoire définie sur (Ω, \mathcal{A}, P) , à valeurs dans $\llbracket 1, N \rrbracket$. Établir les deux relations suivantes :

$$E(Y) = \sum_{k=0}^{N-1} P(Y > k) \text{ et } E(Y^2) = \sum_{k=0}^{N-1} (2k+1)P(Y > k).$$

Partie I. Inf et Sup

- Rappeler, sans démonstration, les valeurs respectives de $E(U_1)$ et de $V(U_1)$.
- Calculer, pour tout k de $\llbracket 1, N \rrbracket$, $P(T_n \leq k)$.
 - En déduire la loi de probabilité de T_n .
- Montrer que la suite $(d_n(N))_{n \geq 1}$ est convergente, et calculer sa limite.
 - Exprimer $E(T_n)$ en fonction de N et de $d_n(N)$. En déduire la valeur de $\lim_{n \rightarrow \infty} E(T_n)$.
 - Établir la formule suivante : $V(T_n) = (2N-1)d_n(N) - 2Nd_{n+1}(N) - d_n^2(N)$.
En déduire la valeur de $\lim_{n \rightarrow \infty} V(T_n)$.
 - Montrer que si $N \geq 2$, on a $\lim_{n \rightarrow \infty} \frac{d_{n+1}(N)}{d_n(N)} = 1 - \frac{1}{N}$.
En déduire que, lorsque n tend vers $+\infty$, on a : $V(T_n) \sim d_n(N)$.
- Déterminer la loi de Z_n . Calculer $E(Z_n)$ et $V(Z_n)$.
- Écrire une fonction Scilab `simulmax(N, n)` qui simule la variable aléatoire T_n .

Partie II : Couple (Inf, Sup)

- On pose, pour tout n de \mathbf{N}^* et pour tout couple (k, ℓ) de \mathbf{N}^2 , $\phi_n(k, \ell) = P([T_n \leq k] \cap [Z_n \leq \ell])$.
 - Montrer, pour tout (k, ℓ) de $\llbracket 1, N \rrbracket^2$, la relation suivante :

$$\phi_n(k, \ell) = \begin{cases} \left(\frac{k}{N}\right)^n & \text{si } k \leq \ell \\ \left(\frac{k}{N}\right)^n - \left(\frac{k-\ell}{N}\right)^n & \text{si } k > \ell \end{cases}$$

- Établir, pour tout (k, ℓ) de $\llbracket 1, N \rrbracket^2$, la formule suivante :

$$P([T_n = k] \cap [Z_n = \ell]) = \phi_n(k, \ell) + \phi_n(k-1, \ell-1) - \phi_n(k-1, \ell) - \phi_n(k, \ell-1).$$

- En déduire, en distinguant les trois cas $k < \ell$, $k = \ell$ et $k > \ell$, l'expression de $P([T_n = k] \cap [Z_n = \ell])$ en fonction de k et ℓ .

8. On donne, pour tout couple (m, n) de $(\mathbf{N}^*)^2$, les deux relations suivantes :

$$\sum_{j=1}^m [(j+1)^n - 2j^n + (j-1)^n] = (m+1)^n - m^n - 1 \quad (21.1)$$

$$\sum_{j=1}^m j[(j+1)^n - 2j^n + (j-1)^n] = m(m+1)^n - (m+1)m^n \quad (21.2)$$

- a. En déduire, pour tout n de \mathbf{N}^* , la formule suivante : $E(T_n Z_n) = N(1 + d_{n+1}(N))$.
 b. On note, pour tout $n \in \mathbf{N}^*$, ρ_n le coefficient de corrélation linéaire entre T_n et Z_n .
 Calculer $\lim_{n \rightarrow \infty} \rho_n$ lorsque $N \geq 2$.
9. a. Pour tout n de \mathbf{N}^* , et pour tout couple (k, ℓ) de $\llbracket 1, N \rrbracket^2$, calculer la probabilité conditionnelle $P_{[T_n=k]}(Z_n = \ell)$.
 b. En déduire, pour tout n de \mathbf{N}^* et pour tout k de $\llbracket 1, N \rrbracket$, l'espérance conditionnelle $E(Z_n | T_n = k)$ de Z_n sachant $[T_n = k]$.

Partie III : Préviation

Pour n entier de \mathbf{N}^* , on dispose d'un $(n+1)$ -échantillon indépendant identiquement distribué (i.i.d.) $(U_1, U_2, \dots, U_{n+1})$ de la loi uniforme sur $\llbracket 1, N \rrbracket$.

On pose $T_n = \sup(U_1, U_2, \dots, U_n)$ et $T_{n+1} = \sup(U_1, \dots, U_{n+1}) = \sup(T_n, U_{n+1})$.

Pour tout $t = (t_1, \dots, t_N)$ de \mathbf{R}^N , on pose : $W_t(T_n) = \sum_{k=1}^N t_k \times \mathbb{1}_{[T_n=k]}$.

Dans cette partie, on se propose de déterminer la valeur de t pour laquelle les deux conditions suivantes sont vérifiées :

- i) $E(W_t(T_n)) = E(T_{n+1})$;
 ii) $E[(T_{n+1} - W_t(T_n))^2]$ est minimal.

10. Montrer, pour tout k de $\llbracket 1, N \rrbracket$, la relation : $P(W_t(T_n) = t_k) = P(T_n = k)$.

11. Établir, pour tout k de $\llbracket 1, N \rrbracket$, la formule suivante :

$$E(T_{n+1} \times \mathbb{1}_{[T_n=k]}) = E(T_{n+1} | T_n = k) \times P(T_n = k).$$

12. a. Calculer, pour tout couple (k, j) de $\llbracket 1, N \rrbracket^2$, $P([T_n = k] \cap [T_{n+1} = j])$.
 b. En déduire, pour tout couple (k, j) de $\llbracket 1, N \rrbracket^2$, la probabilité conditionnelle $P_{[T_n=k]}(T_{n+1} = j)$.
 c. Déterminer, pour tout k de $\llbracket 1, N \rrbracket$, l'expression de l'espérance conditionnelle $E(T_{n+1} | T_n = k)$ de T_{n+1} sachant $[T_n = k]$.
 d. En appliquant la formule de l'espérance totale, déduire de la question précédente la relation suivante :

$$E(T_{n+1}) = \frac{N+1}{2} + \frac{1}{2N} (E(T_n^2) - E(T_n)).$$

13. Établir l'égalité suivante : $(W_t(T_n))^2 = \sum_{k=1}^N t_k^2 \times \mathbb{1}_{[T_n=k]}$.

14. Soit g la fonction définie sur \mathbf{R}^N à valeurs réelles par :

$$g(t_1, t_2, \dots, t_N) = E[(T_{n+1} - W_t(T_n))^2].$$

- a. À l'aide des résultats des questions 11, 12 et 13, expliciter g en fonction des variables t_1, t_2, \dots, t_N .
 b. Montrer que g admet un minimum global sur \mathbf{R}^N , atteint en un point $\theta = (\theta_1, \theta_2, \dots, \theta_N)$ que l'on déterminera en fonction de $E(T_{n+1} | T_n = 1), E(T_{n+1} | T_n = 2), \dots, E(T_{n+1} | T_n = N)$.

15. Établir les deux relations suivantes :

$$E(W_\theta(T_n)) = E(T_{n+1}) \text{ et } V(W_\theta(T_n)) \leq V(T_{n+1}).$$

16. a. Établir, pour tout i de \mathbf{N}^* , l'égalité suivante :

$$\sum_{k=1}^N k^i \times \mathbb{1}_{[T_n=k]} = (T_n)^i.$$

b. En déduire la relation suivante :

$$W_\theta(T_n) = \frac{N+1}{2} + \frac{1}{2N} (T_n^2 - T_n).$$

Partie IV : Estimation

Soit U une variable aléatoire définie sur un espace probabilisé (Ω, \mathcal{A}, P) , de loi uniforme discrète sur $\llbracket 1, N \rrbracket$. On suppose que le paramètre N est inconnu.

Cette partie a pour objet la détermination d'un estimateur ponctuel de N , sans biais et de variance minimale.

Pour n entier supérieur ou égal à 1, soit (U_1, U_2, \dots, U_n) un n -échantillon i.i.d. de la loi de U .

17. Soit ε un réel strictement positif. On pose :

$$A_n(\varepsilon) = \llbracket T_n - N \geq \varepsilon \rrbracket \text{ et } B_n(\varepsilon) = \llbracket |T_n - E(T_n)| + |d_n(N)| \geq \varepsilon \rrbracket.$$

- Peut-on dire que $T_n + d_n(N)$ est un estimateur sans biais de N ?
- Montrer que la suite $(T_n)_{n \geq 1}$ est une suite d'estimateurs asymptotiquement sans biais du paramètre N .
- Montrer que $A_n(\varepsilon) \subset B_n(\varepsilon)$ et qu'il existe un entier naturel n_0 tel que, pour tout $n > n_0$, on a

$$B_n(\varepsilon) \subset \llbracket |T_n - E(T_n)| \geq \varepsilon/2 \rrbracket.$$

d. En déduire que la suite d'estimateurs $(T_n)_{n \geq 1}$ est convergente.

18. a. Calculer, pour tout n -uplet (u_1, u_2, \dots, u_n) de $\llbracket 1, N \rrbracket^2$, $P\left(\bigcap_{i=1}^n [U_i = u_i]\right)$.

b. En déduire que, pour tout k de $\llbracket 1, N \rrbracket$, la loi conditionnelle du vecteur aléatoire (U_1, \dots, U_n) sachant $[T_n = k]$ est donnée par

$$P_{[T_n=k]} \left(\bigcap_{i=1}^n [U_i = u_i] \right) = \begin{cases} \frac{1}{k^n - (k-1)^n} & \text{si pour tout } i \in \llbracket 1, n \rrbracket, 1 \leq u_i \leq N \text{ et } \max_{1 \leq i \leq n} (u_i) = k \\ 0 & \text{sinon} \end{cases}$$

On remarquera que cette loi conditionnelle ne dépend pas du paramètre N .

19. On pose, pour n entier de \mathbf{N}^* , $S_n = T_n + Z_n - 1$ et, pour tout k de $\llbracket 1, N \rrbracket$,

$$\psi_n(k) = \frac{k^{n+1} - (k-1)^{n+1}}{k^n - (k-1)^n}.$$

- Montrer que S_n est un estimateur sans biais de N .
- Établir, pour tout k de $\llbracket 1, N \rrbracket$, l'égalité : $\psi_n(k) = E(S_n | T_n = k)$.
- En déduire que $\psi_n(T_n)$ est un estimateur sans biais de N .
- On pose, pour tout k de $\llbracket 1, N \rrbracket$, $\varphi_n(k) = E(S_n^2 | T_n = k)$.
Établir, pour tout k de $\llbracket 1, N \rrbracket$, l'inégalité : $\psi_n^2(k) \leq \varphi_n(k)$ (on pourra utiliser la fonction définie sur \mathbf{R} par $\lambda \mapsto E((S_n - \lambda)^2 | T_n = k)$). En déduire que $V(\psi_n(T_n)) \leq V(S_n)$.
- Calculer $V(S_n)$. En déduire que $\psi_n(T_n)$ est un estimateur convergent de N .

20. Soit, pour n entier de \mathbf{N}^* , un estimateur sans biais R_n du paramètre N .

On pose, pour tout k de $\llbracket 1, N \rrbracket$, $f_n(k) = E(R_n | T_n = k)$.

- En utilisant une méthode analogue à celle de la question 19.d, montrer que $V(f_n(T_n)) \leq V(R_n)$.
- Soit F une fonction réelle. Montrer que pour n fixé dans \mathbf{N}^* , la condition «pour tout $N \in \mathbf{N}^*$, $E(F(T_n)) = N$ » est vérifiée si et seulement si, pour tout k de $\llbracket 1, N \rrbracket$, on a : $F(k) = \psi_n(k)$.
- En déduire que dans l'ensemble des estimateurs sans biais de N , l'estimateur $\psi_n(T_n)$ est optimal, dans le sens où $V(\psi_n(T_n))$ est minimale.

La partie IV constitue une démonstration du théorème de Lehmann-Scheffé dans le cas particulier d'une loi uniforme sur $\llbracket 1, N \rrbracket$, avec N inconnu.

CORRECTION DU DEVOIR MAISON 16

MATHS II 2009

Préliminaires

1. Soit Y une variable aléatoire à valeurs dans $\llbracket 1, N \rrbracket$. Alors

$$\begin{aligned} \sum_{k=0}^{N-1} P(Y > k) &= \sum_{k=0}^{N-1} \sum_{i=k+1}^N P(Y = i) = \sum_{i=1}^N \sum_{k=0}^{i-1} P(Y = i) \\ &= \sum_{i=1}^N P(Y = i) \sum_{k=0}^{i-1} 1 = \sum_{i=1}^N P(Y = i) i \\ &= \boxed{E(Y)}. \end{aligned}$$

De même, on a

$$\begin{aligned} \sum_{k=0}^{N-1} (2k+1)P(Y > k) &= \sum_{k=0}^{N-1} (2k+1) \sum_{i=k+1}^N P(Y = i) = \sum_{i=1}^N \sum_{k=0}^{i-1} (2k+1)P(Y = i) \\ &= \sum_{i=1}^N P(Y = i) \sum_{k=0}^{i-1} (2k+1) = \sum_{i=1}^N P(Y = i) \left(2 \sum_{k=0}^{i-1} k + \sum_{k=0}^{i-1} 1 \right) \\ &= \sum_{i=1}^N P(Y = i) \left(2 \frac{(i-1)i}{2} + i \right) = \sum_{i=1}^N P(Y = i) (i^2 - i + i) \\ &= \sum_{i=1}^N i^2 P(Y = i) = \boxed{E(Y^2)}. \end{aligned}$$

Remarque

Ce résultat reste en fait valable pour une variable aléatoire à valeurs dans \mathbf{N} admettant une espérance, en remplaçant la somme finie par la somme d'une série :

$$E(X) = \sum_{k=0}^{+\infty} P(X > k).$$

Partie I : Inf et Sup

2. On a $\boxed{E(U_1) = \frac{N+1}{2}} \text{ et } \boxed{V(U_1) = \frac{N^2-1}{12}}.$

- 3.a. Pour tout $k \in \llbracket 1, N \rrbracket$ on a, par indépendance des U_i ,

$$P(T_n \leq k) = P\left(\bigcap_{i=1}^n [U_i \leq k]\right) = \prod_{i=1}^n P(U_i \leq k) = (P(U_1 \leq k))^n = \boxed{\left(\frac{k}{N}\right)^n}.$$

- 3.b. Puisque T_n est à valeurs dans $\llbracket 1, N \rrbracket$, on en déduit, que

$$P(T_n = 1) = P(T_n \leq 1) = \frac{1}{N^n}$$

et que pour $k \in \llbracket 2, N \rrbracket$,

$$\boxed{P(T_n = k) = P(T_n \leq k) - P(T_n \leq k-1) = \left(\frac{k}{N}\right)^n - \left(\frac{k-1}{N}\right)^n}.$$

Notons que cette seconde formule reste valable pour $k = 1$.

- 4.a. Si $N = 1$, $d_n(N) = 0$ pour tout n , et donc

$$\lim_{n \rightarrow +\infty} d_n(N) = 0.$$

Si $N \geq 2$, pour tout $k \in \llbracket 1, N-1 \rrbracket$, on a $0 \leq \frac{k}{N} < 1$ et donc $\left(\frac{k}{N}\right)^n \xrightarrow[n \rightarrow +\infty]{} 0$.

Ainsi, par somme d'un nombre fini¹ de suites de limites nulles, $(d_n(N))_{n \geq 1}$ est convergente et

$$\lim_{n \rightarrow +\infty} d_n(N) = \sum_{k=1}^{N-1} \lim_{n \rightarrow +\infty} \left(\frac{k}{N}\right)^n = \boxed{0}.$$

¹ et fixé !

4.b. D'après la question préliminaire, on a $E(T_n) = \sum_{k=0}^{N-1} P(T_n > k)$. Or,

$$P(T_n > k) = 1 - P(T_n \leq k) = 1 - \left(\frac{k}{N}\right)^n$$

et donc

$$E(T_n) = \sum_{k=0}^{N-1} \left(1 - \left(\frac{k}{N}\right)^n\right) = N - \sum_{k=0}^{N-1} \left(\frac{k}{N}\right)^n = N - \sum_{k=1}^{N-1} \left(\frac{k}{N}\right)^n = \boxed{N - d_n(N)}.$$

On en déduit que $\lim_{n \rightarrow +\infty} E(T_n) = N$.

4.c. Toujours à l'aide du préliminaire, on a

$$\begin{aligned} E(T_n^2) &= \sum_{k=0}^{N-1} (2k+1)P(Y > k) = \sum_{k=0}^{N-1} (2k+1) \left(1 - \left(\frac{k}{N}\right)^n\right) \\ &= \sum_{k=0}^{N-1} (2k+1) - \sum_{k=0}^{N-1} (2k+1) \left(\frac{k}{N}\right)^n = 2 \sum_{k=0}^{N-1} k + N - \sum_{k=0}^{N-1} (2k+1) \left(\frac{k}{N}\right)^n \\ &= 2 \frac{N(N-1)}{2} + N - d_n(N) - 2 \sum_{k=1}^{N-1} k \left(\frac{k}{N}\right)^n \\ &= N^2 - d_n(N) - 2N \sum_{k=1}^{N-1} \left(\frac{k}{N}\right)^{n+1} \\ &= N^2 - d_n(N) - 2Nd_{n+1}(N). \end{aligned}$$

On a alors, par la formule de Huygens,

$$\begin{aligned} V(T_n) &= E(T_n^2) - E(T_n)^2 = N^2 - d_n(N) - 2Nd_{n+1}(N) - (N - d_n(N))^2 \\ &= N^2 - d_n(N) - 2Nd_{n+1}(N) - N^2 + 2Nd_n(N) - d_n(N)^2 \\ &= (2N-1)d_n(N) - 2Nd_{n+1}(N) - d_n(N)^2. \end{aligned}$$

Puisque $\lim_{n \rightarrow +\infty} d_n(N) = 0$, on a également $\lim_{n \rightarrow +\infty} d_n(N)^2 = 0$ et $\lim_{n \rightarrow +\infty} d_{n+1}(N) = 0$. Et donc

$$\lim_{n \rightarrow +\infty} V(T_n) = 0.$$

4.d. Commençons par remarquer que pour $k \in \llbracket 1, N-2 \rrbracket$, on a

$$\left(\frac{k}{N}\right)^n = o_{n \rightarrow +\infty} \left(\left(\frac{N-1}{N}\right)^n\right)$$

et donc

$$\sum_{k=1}^{N-2} \left(\frac{k}{N}\right)^n = o_{n \rightarrow +\infty} \left(\left(\frac{N-1}{N}\right)^n\right)$$

de sorte que

$$d_n(N) = \left(\frac{N-1}{N}\right)^n + o_{n \rightarrow +\infty} \left(\left(\frac{N-1}{N}\right)^n\right) \underset{n \rightarrow +\infty}{\sim} \left(\frac{N-1}{N}\right)^n.$$

On en déduit alors que

$$\frac{d_{n+1}(N)}{d_n(N)} \underset{n \rightarrow +\infty}{\sim} \frac{N-1}{N} = 1 - \frac{1}{N}$$

et donc

$$\lim_{n \rightarrow +\infty} \frac{d_{n+1}(N)}{d_n(N)} = 1 - \frac{1}{N}.$$

Intuition

Lorsque deux suites tendent vers 0, c'est celle qui tend le «plus vite» vers 0 qui est négligeable devant l'autre. Pour s'en convaincre, le plus simple est de revenir au quotient.

En particulier, pour deux suites géométriques de raison dans $] -1, 1[$, c'est celle de plus petite raison qui est négligeable devant l'autre.

Par conséquent, on a

$$\frac{V(T_n)}{d_n(N)} = (2N-1) - 2N \frac{d_{n+1}(N)}{d_n(N)} - d_n(N) \xrightarrow{n \rightarrow +\infty} 2N-1 - 2N \left(1 - \frac{1}{N}\right) = 2N-1 - 2N+2 = 1.$$

Et donc $V(T_n) \underset{n \rightarrow +\infty}{\sim} d_n(N)$.

5. Pour $k \in \llbracket 1, N \rrbracket$, on a, toujours par indépendance des U_i ,

$$P(Z_n \geq k) = P\left(\bigcap_{i=1}^n [U_i \geq k]\right) = \prod_{i=1}^n P(U_i \geq k) = (P(U_1 \geq k))^n.$$

Mais $P(U_1 \geq k) = \sum_{i=k}^N \frac{1}{N} = \frac{N-k+1}{N}$ et donc

$$\forall k \in \llbracket 1, N \rrbracket, P(Z_n \geq k) = \left(\frac{N-k+1}{N}\right)^n.$$

On a alors, $P(Z_n = N) = P(Z_n \geq N) = \left(\frac{1}{N}\right)^n$ et pour $k \in \llbracket 1, N-1 \rrbracket$,

$$P(Z_n = k) = P(Z_n \geq k) - P(Z_n \geq k+1) = \left(\frac{N-k+1}{N}\right)^n - \left(\frac{N-k}{N}\right)^n.$$

Afin d'éviter des calculs fastidieux, notons que

$$\forall k \in \llbracket 1, N \rrbracket, P(Z_n = k) = P(T_n = N+1-k) = P(N+1-T_n = k).$$

Donc Z_n et $N+1-T_n$ ont même loi.

On a alors

$$E(Z_n) = E(N+1-T_n) = N+1 - E(T_n) = d_n(N) + 1 \text{ et } V(Z_n) = V(T_n).$$

6. Il suffit de simuler n lois uniformes et de garder la plus grande d'entre elles.

```
1 function y = simulmax(N,n)
2   A = grand(1,n,'uin',1,N)
3   y = max(A)
4 endfunction
```

Partie II. Couple (Inf, Sup)

- 7.a. Notons qu'on a toujours $Z_n \leq T_n$, par définition.

Donc pour $k \leq \ell$, on a $[T_n \leq k] \cap [Z_n \leq \ell] = [T_n \leq k]$, de sorte que

$$\phi_n(k, \ell) = P(T_n \leq k) = \left(\frac{k}{N}\right)^n.$$

En revanche, si $k > \ell$, on a

$$P(T_n \leq k) = P([T_n \leq k] \cap [Z_n \leq \ell]) + P([T_n \leq k] \cap [Z_n \geq \ell + 1]).$$

Or,

$$P([T_n \leq k] \cap [Z_n \geq \ell + 1]) = P\left(\bigcap_{i=1}^n [\ell + 1 \leq U_i \leq k]\right) = \prod_{i=1}^n P(\ell + 1 \leq U_i \leq k) = \left(\frac{k-\ell}{N}\right)^n.$$

On en déduit que

$$\phi_n(k, \ell) = P(T_n \leq \ell) - P([T_n \leq k] \cap [Z_n \geq \ell + 1]) = \left(\frac{k}{N}\right)^n - \left(\frac{k-\ell}{N}\right)^n.$$

Même loi

Ces deux variables ont même loi, mais ne sont pas égales pour autant !

Détails

On a à la fois $T_n \leq k$ et $Z_n \geq \ell + 1$ si et seulement si les U_i sont tous plus petits que k (et donc le plus grand d'entre eux l'est aussi) et tous plus grands que $\ell + 1$ (et donc le plus petit d'entre eux l'est aussi).

7.b. On a

$$[T_n \leq k] \cap [Z_n \leq \ell] = ([T_n \leq k] \cap [Z_n \leq \ell - 1]) \cup ([T_n \leq k] \cap [Z_n = \ell])$$

et ces deux événements sont disjoints. Donc

$$\phi_n(k, \ell) = \phi_n(k, \ell - 1) + P([T_n \leq k] \cap [Z_n = \ell]).$$

Mais

$$P([T_n \leq k] \cap [Z_n = \ell]) = P([T_n \leq k - 1] \cap [Z_n = \ell]) + P([T_n = k] \cap [Z_n = \ell]).$$

et

$$P([T_n \leq k - 1] \cap [Z_n = \ell]) = P([T_n \leq k - 1] \cap [Z_n \leq \ell]) - P([T_n \leq k - 1] \cap [Z_n \leq \ell - 1]).$$

Au final, on a donc

$$\phi_n(k, \ell) = \phi_n(k, \ell - 1) + \phi_n(k - 1, \ell) - \phi_n(k - 1, \ell - 1) + P([Z_n = k] \cap [Z_n = \ell])$$

soit encore

$$P([T_n = k] \cap [Z_n = \ell]) = \phi_n(k, \ell) + \phi_n(k - 1, \ell - 1) - \phi_n(k - 1, \ell) - \phi_n(k, \ell - 1).$$

- 7.c. • Si $k < \ell$, alors nous pourrions utiliser le résultat précédent, mais il est plus simple de remarquer que $[T_n = k] \cap [Z_n = \ell] = \emptyset$, et donc $P([T_n = k] \cap [Z_n = \ell]) = 0$.
- Si $k = \ell$, alors

$$P([T_n = k] \cap [Z_n = \ell]) = P\left(\bigcap_{i=1}^n [U_i = k]\right) = \prod_{i=1}^n P(U_i = k) = \frac{1}{N^n}.$$

- Enfin, si $k > \ell$, alors on a $k \geq \ell - 1$, $k - 1 \geq \ell$ et $k - 1 \geq \ell - 1$, de sorte que

$$\begin{aligned} P([T_n = k] \cap [Z_n = \ell]) &= \Phi_n(k, \ell) + \Phi_n(k - 1, \ell - 1) - \Phi_n(k - 1, \ell) - \Phi_n(k, \ell - 1) \\ &= \frac{1}{N^n} (k^n - (k - \ell)^n + (k - 1)^n - (k - \ell)^n - (k - 1)^n + (k - 1 - \ell)^n - k^n + (k - \ell + 1)^n) \\ &= \frac{1}{N^n} ((k - 1 - \ell)^n + (k - \ell + 1)^n - 2(k - \ell)^n). \end{aligned}$$

Explication

Si à la fois le plus grand et le plus petit des U_i valent k , alors tous les U_i sont égaux à k .

8.a. Par le théorème de transfert, on a

$$E(T_n Z_n) = \sum_{k=1}^N \sum_{\ell=1}^N k \ell P([T_n = k] \cap [Z_n = \ell]) = \sum_{k=1}^N \sum_{\ell=1}^k k \ell P([T_n = k] \cap [Z_n = \ell]).$$

Nous avons mentionné que pour $\ell > k$,

$$P([T_n = k] \cap [Z_n = \ell]) = 0.$$

En remplaçant $P([T_n = k] \cap [Z_n = \ell])$ par l'expression obtenue à la question précédente, il vient, en séparant les cas $k = \ell$ et $k > \ell$,

$$\begin{aligned} E(T_n Z_n) &= \sum_{k=1}^N k^2 P([T_n = k] \cap [Z_n = k]) + \sum_{k=2}^N \sum_{\ell=1}^{k-1} k \ell P([T_n = k] \cap [Z_n = \ell]) \\ &= \frac{1}{N^n} \sum_{k=1}^N k^2 + \frac{1}{N^n} \sum_{k=2}^N k \sum_{\ell=1}^{k-1} \ell ((k - 1 - \ell)^n + (k - \ell + 1)^n - 2(k - \ell)^n). \end{aligned}$$

Dans la seconde somme, effectuons le changement d'indice $j = k - \ell$, de sorte que lorsque ℓ varie de 1 à $k - 1$, j varie de 1 à $k - 1$, et donc

$$\begin{aligned} \sum_{\ell=1}^{k-1} \ell ((k - 1 - \ell)^n + (k - \ell + 1)^n - 2(k - \ell)^n) &= \sum_{j=1}^{k-1} (k - j) ((j + 1)^n + (j - 1)^n - 2j^n) \\ &= k \sum_{j=1}^{k-1} ((j + 1)^n + (j - 1)^n - 2j^n) - \sum_{j=1}^{k-1} j ((j + 1)^n + (j - 1)^n - 2j^n) \end{aligned}$$

$$\begin{aligned}
 &= k(k^n - (k-1)^n - 1) - (k-1)k^n + k(k-1)^n \\
 &= k^{n+1} - (k-1)k^n - k = k^n - k
 \end{aligned}$$

On a utilisé les résultats fournis par l'énoncé, avec $m = k - 1$.

Et alors, il vient

$$\begin{aligned}
 \sum_{k=2}^N k \sum_{\ell=1}^{k-1} \ell ((j+1)^n + (j-1)^n - 2j^n) &= \sum_{k=2}^N k(k^n - k) \\
 &= \sum_{k=2}^N k^{n+1} - \sum_{k=2}^N k^2 \\
 &= \sum_{k=1}^N k^{n+1} - \sum_{k=1}^N k^2.
 \end{aligned}$$

Au final, on obtient donc

$$E(T_n Z_n) = \frac{1}{N^n} \sum_{k=1}^N k^{n+1}.$$

Si $N = 1$, on a alors² $E(T_n Z_n) = 1 = 1(1 + d_{n+1}(1))$, et si $N \geq 2$, on a

² $T_n Z_n$ est alors une variable certaine égale à 1.

$$E(T_n Z_n) = N \sum_{k=1}^N \left(\frac{k}{N}\right)^{n+1} = N \left(\sum_{k=1}^{N-1} \left(\frac{k}{N}\right)^n + 1 \right) = \boxed{N(1 + d_{n+1}(N))}.$$

8.b. Par définition, on a $\rho_n = \frac{\text{Cov}(T_n, Z_n)}{\sqrt{V(T_n)}\sqrt{V(Z_n)}} = \frac{\text{Cov}(T_n, Z_n)}{V(T_n)}$.

De plus, par la formule de Huygens,

$$\begin{aligned}
 \text{Cov}(T_n, Z_n) &= E(T_n Z_n) - E(T_n)E(Z_n) = N(1 + d_{n+1}(N)) - (N - d_n(N))(d_n(N) + 1) \\
 &= Nd_{n+1}(N) - Nd_n(N) + d_n(N)^2 + d_n(N).
 \end{aligned}$$

On en déduit³ que

³ Grâce à l'équivalent $V(T_n) \sim d_n(N)$ de la question 4.d.

$$\begin{aligned}
 \rho_n &= \frac{1}{V(T_n)} \text{Cov}(T_n, Z_n) \underset{n \rightarrow +\infty}{\sim} \frac{1}{d_n(N)} (Nd_{n+1}(N) - Nd_n(N) + d_n(N)^2 + d_n(N)) \\
 &\underset{n \rightarrow +\infty}{\sim} N \frac{d_{n+1}(N)}{d_n(N)} - N + d_n(N) + 1 \\
 &\xrightarrow{n \rightarrow +\infty} N \left(1 - \frac{1}{N}\right) - N + 1 = 0.
 \end{aligned}$$

On a utilisé la limite de $\frac{d_{n+1}(N)}{d_n(N)}$ calculée à la question 4.d.

9.a. Par définition d'une probabilité conditionnelle, on a $P_{[T_n=k]}(Z_n = \ell) = \frac{P([T_n = k] \cap [Z_n = \ell])}{P(T_n = k)}$.

En utilisant le résultat de la question 7, on en déduit que

$$P_{[T_n=k]}(Z_n = \ell) = \begin{cases} 0 & \text{si } k < \ell \\ \frac{1}{N^n} \frac{1}{\left(\frac{k}{N}\right)^n - \left(\frac{k-1}{N}\right)^n} = \frac{1}{k^n - (k-1)^n} & \text{si } k = \ell \\ \frac{\frac{1}{N^n} ((k-1-\ell)^n + (k-\ell+1)^n - 2(k-\ell)^n)}{\left(\frac{k}{N}\right)^n - \left(\frac{k-1}{N}\right)^n} = \frac{(k-1-\ell)^n + (k-\ell+1)^n - 2(k-\ell)^n}{k^n - (k-1)^n} & \text{si } k > \ell \end{cases}$$

9.b. Par définition de l'espérance conditionnelle, on a

$$E(Z_n | T_n = k) = \sum_{\ell=1}^N \ell P_{[T_n=k]}(Z_n = \ell).$$

À l'aide de l'expression de $P_{[T_n=k]}(Z_n = \ell)$ obtenue à la question 9.a, il vient

$$E(Z_n | T_n = k) = \frac{k}{k^n - (k-1)^n} + \frac{1}{k^n - (k-1)^n} \sum_{\ell=1}^{k-1} \ell ((k-1-\ell)^n + (k-\ell+1)^n - 2(k-\ell)^n).$$

Remarquons alors que la somme qui apparaît dans la formule ci-dessus a déjà été calculée à la question 8.a, et vaut $k^n - k$, de sorte que

$$E(Z_n | T_n = k) = \frac{k}{k^n - (k-1)^n} + \frac{k^n - k}{k^n - (k-1)^n} = \boxed{\frac{k^n}{k^n - (k-1)^n}}.$$

Partie III : Prévision

10. Notons que T_n ne peut prendre qu'une valeur à la fois, et donc les $\mathbb{1}_{[T_n=k]}$ sont toujours toutes nulles, sauf l'une d'entre elles qui vaut 1. Autrement dit, $W_t(T_n)$ prend la valeur t_k lorsque $T_n = k$.

$$\text{Plus précisément, pour } \omega \in \Omega, W_t(T_n)(\omega) = \sum_{k=1}^N t_k \mathbb{1}_{[T_n=k]}(\omega) = \begin{cases} t_1 & \text{si } T_n(\omega) = 1 \\ t_2 & \text{si } T_n(\omega) = 2 \\ \vdots & \\ t_N & \text{si } T_n(\omega) = N \end{cases}$$

Et donc $[W_t(T_n) = t_k] = [T_n = k]$.

Et donc $\boxed{P(W_t(T_n) = t_k) = P(T_n = k)}$.

11. Notons que $\{[T_n = i], i \in \llbracket 1, N \rrbracket\}$ est un système complet d'événements, et donc par la formule de l'espérance totale appliquée à $T_{n+1} \times \mathbb{1}_{[T_n=k]}$, on a

$$E(T_{n+1} \times \mathbb{1}_{[T_n=k]}) = \sum_{i=1}^N E(T_{n+1} \times \mathbb{1}_{[T_n=k]} | T_n = i) P(T_n = i).$$

Mais si $i \neq k$, alors pour tout $j \in \llbracket 1, N \rrbracket$, on a

$$P_{[T_n=i]}((T_{n+1} \times \mathbb{1}_{[T_n=k]}) = j) = \frac{P([T_{n+1} = j] \cap [T_n = k] \cap [T_n = i])}{P(T_n = i)} = 0$$

et donc pour $i \neq k$, $E(T_{n+1} \times \mathbb{1}_{[T_n=k]} | T_n = i) = 0$.

Au contraire, si $k = i$, alors pour tout $j \in \llbracket 1, N \rrbracket$, on a

$$\begin{aligned} P_{[T_n=k]}(T_{n+1} \times \mathbb{1}_{[T_n=k]} = j) &= \frac{P([T_{n+1} = j] \cap [T_n = k] \cap [T_n = k])}{P(T_n = k)} \\ &= \frac{P([T_{n+1} = j] \cap [T_n = k])}{P(T_n = k)} \\ &= P_{[T_n=k]}(T_{n+1} = j). \end{aligned}$$

Ainsi, la loi conditionnelle de $T_{n+1} \times \mathbb{1}_{[T_n=k]}$ sachant $[T_n = k]$ est la loi de T_{n+1} sachant $[T_n = k]$ et donc

$$E(T_{n+1} \times \mathbb{1}_{[T_n=k]} | T_n = k) = E(T_{n+1} | T_n = k).$$

Il ne reste donc qu'un seul terme dans la formule de l'espérance totale, et

$$\boxed{E(T_{n+1} \times \mathbb{1}_{[T_n=k]}) = E(T_{n+1} | T_n = k) \times P(T_n = k)}.$$

- 12.a. Commençons par remarquer que par le lemme des coalitions, T_n et U_{n+1} sont indépendantes, car T_n est fonction de U_1, \dots, U_n et que les U_i sont mutuellement indépendantes. Si $j < k$, alors $P([T_n = k] \cap [T_{n+1} = j]) = 0$ car on ne peut avoir $\max(U_1, \dots, U_{n+1}) < \max(U_1, \dots, U_n)$. Si $j = k$, alors

$$[T_n = k] \cap [T_{n+1} = k] = [T_n = k] \cap [U_{n+1} \leq k]$$

et donc

$$P([T_n = k] \cap [T_{n+1} = k]) = P(T_n = k) \times P(U_{n+1} \leq k) = \boxed{\left(\left(\frac{k}{N} \right)^n - \left(\frac{k-1}{N} \right)^n \right) \frac{k}{N}}.$$

Enfin, si $j > k$, alors

$$[T_n = k] \cap [T_{n+1} = j] = [T_n = k] \cap [U_{n+1} = j]$$

de sorte que

$$P([T_n = k] \cap [T_{n+1} = j]) = P(T_n = k) P(U_{n+1} = j) = \boxed{\left(\left(\frac{k}{N} \right)^n - \left(\frac{k-1}{N} \right)^n \right) \frac{1}{N}}.$$

Erreur

L'énoncé était faux ! Pour que le résultat de cette question soit valable, il faut supposer que les t_i sont deux à deux distincts, ce que nous faisons dans la suite. En effet, si on avait par exemple $t_1 = t_2$, alors on aurait

$$\begin{aligned} [W_t(T_n) = t_1] &= \\ &= [T_n = 1] \cup [T_n = 2]. \end{aligned}$$

Détails

On ne peut avoir $T_{n+1} \times \mathbb{1}_{[T_n=k]} = j$ que si $T_{n+1} = j$ et que $[T_n = k]$ est réalisé, car sinon $\mathbb{1}_{[T_n=k]} = 0$.

12.b. Par définition, on a

$$P_{[T_n=k]}(T_{n+1} = j) = \frac{P([T_n = k] \cap [T_{n+1} = j])}{P(T_n = k)} = \begin{cases} 0 & \text{si } j < k \\ \frac{k}{N} & \text{si } j = k \\ \frac{1}{N} & \text{si } j > k \end{cases}$$

12.c. Par définition de l'espérance conditionnelle, on a

$$E(T_{n+1}|T_n = k) = \sum_{j=1}^N jP_{[T_n=k]}(T_{n+1} = j).$$

Et donc à l'aide du résultat de la question précédente, on obtient $E(T_{n+1}|T_n = k) = N$ si $k = N$, et pour $k < N$,

$$\begin{aligned} E(T_{n+1}|T_n = k) &= \frac{k^2}{N} + \sum_{j=k+1}^N \frac{j}{N} \\ &= \frac{k^2}{N} + \sum_{j=1}^N \frac{j}{N} - \sum_{j=1}^k \frac{j}{N} \\ &= \frac{k^2}{N} + \frac{N(N+1)}{2N} - \frac{k(k+1)}{2N} \\ &= \frac{1}{2N} (N^2 + N + 2k^2 - k^2 - k) \\ &= \frac{N+1}{2} + \frac{k^2 - k}{2N}. \end{aligned}$$

Notons que cette dernière formule reste valable pour $k = N$.

12.d. D'après la formule de l'espérance totale appliquée au système complet d'événements $\{[T_n = k], k \in \llbracket 1, N \rrbracket\}$, il vient

$$\begin{aligned} E(T_{n+1}) &= \sum_{k=1}^N E(T_{n+1}|T_n = k)P(T_n = k) \\ &= \sum_{k=1}^N E(T_{n+1}|T_n = k)P(T_n = k) \\ &= \frac{N+1}{2} \underbrace{\sum_{k=1}^N P(T_n = k)}_{=1} + \frac{1}{2N} \sum_{k=1}^N k^2 P(T_n = k) - \frac{1}{2N} \sum_{k=1}^N k P(T_n = k) \\ &= \frac{N+1}{2} + \frac{1}{2N} E(T_n^2) - \frac{1}{2N} E(T_n) \\ &= \frac{N+1}{2} + \frac{1}{2N} (E(T_n^2) - E(T_n)). \end{aligned}$$

13. Par définition, on a $W_t(T_n) = \sum_{k=1}^N t_k \times \mathbb{1}_{[T_n=k]}$, et donc

$$W_t(T_n)^2 = \left(\sum_{k=1}^N t_k \times \mathbb{1}_{[T_n=k]} \right) \times \left(\sum_{i=1}^N t_i \times \mathbb{1}_{[T_n=i]} \right) = \sum_{k=1}^N \sum_{i=1}^N t_i t_k \mathbb{1}_{[T_n=i]} \mathbb{1}_{[T_n=k]}$$

Mais pour $i \neq k$, les événements $[T_n = i]$ et $[T_n = k]$ sont incompatibles, et donc

$$\mathbb{1}_{[T_n=i]} \times \mathbb{1}_{[T_n=k]} = 0.$$

Ainsi, dans la somme, il ne reste que

$$W_t(T_n)^2 = \sum_{k=1}^N t_k^2 \mathbb{1}_{[T_n=k]} \mathbb{1}_{[T_n=k]} = \sum_{k=1}^N t_k^2 \mathbb{1}_{[T_n=k]}.$$

Détails

On a toujours au moins l'un des deux événements $[T_n = k]$ et $[T_n = i]$ qui n'est pas réalisé, et donc l'indicatrice correspondante est nulle.

14.a. Par linéarité de l'espérance, on a

$$g(t_1, \dots, t_N) = E(T_{n+1}^2) - 2E(T_{n+1}W_t(T_n)) + E(W_t(T_n)^2).$$

Or, en utilisant la question 11, on a

$$E(T_{n+1}W_t(T_n)) = E\left(T_{n+1} \sum_{k=1}^N t_k \mathbb{1}_{[T_n=k]}\right) = \sum_{k=1}^N t_k E(T_{n+1} \mathbb{1}_{[T_n=k]}) = \sum_{k=1}^N t_k E(T_{n+1}|T_n = k)P(T_n = k).$$

Et par la question 13, on a

$$E(W_t(T_n)^2) = E\left(\sum_{k=1}^N t_k^2 \mathbb{1}_{[T_n=k]}\right) = \sum_{k=1}^N t_k^2 E(\mathbb{1}_{[T_n=k]}) = \sum_{k=1}^N t_k^2 P(T_n = k).$$

On en déduit que

$$g(t_1, \dots, t_N) = \sum_{k=1}^N (t_k^2 - 2t_k E(T_{n+1}|T_n = k) + E(T_{n+1}^2)) P(T_n = k).$$

14.b. Nous venons d'écrire $g(t_1, \dots, t_N)$ comme une somme de polynômes du second degré en les variables t_k .

Pour k fixé, le polynôme $(t_k^2 - 2t_k E(T_{n+1}|T_n = k) + E(T_{n+1}^2)) P(T_n = k)$ atteint un unique minimum global (sur \mathbf{R}) en $t_k = E(T_{n+1}|T_n = k)$, de sorte que g atteint un unique minimum global en

$$\theta = (E(T_{n+1}|T_n = 1), \dots, E(T_{n+1}|T_n = N))$$

et ce minimum vaut

$$\begin{aligned} g(\theta) &= \sum_{k=1}^N (E(T_{n+1}|T_n = k)^2 - 2E(T_{n+1}|T_n = k)^2 + E(T_{n+1}^2)) P(T_n = k) \\ &= E(T_{n+1}^2) - \sum_{k=1}^N E(T_{n+1}|T_n = k)^2 P(T_n = k). \end{aligned}$$

15. En remplaçant θ par son expression, on obtient

$$E(W_\theta(T_n)) = E\left(\sum_{k=1}^n E(T_{n+1}|T_n = k) \mathbb{1}_{[T_n=k]}\right) = \sum_{k=1}^N E(T_{n+1}|T_n = k) E(\mathbb{1}_{[T_n=k]}) = \sum_{k=1}^N E(T_{n+1}|T_n = k) P(T_n = k).$$

Mais d'après la formule de l'espérance totale, on reconnaît là l'espérance de T_{n+1} , et donc

$$E(W_\theta(T_n)) = E(T_{n+1}).$$

Par la formule de Huygens, on a $V(T_{n+1}) = E(T_{n+1}^2) - E(T_{n+1})^2$ et de même

$$V(W_\theta(T_n)) = E(W_\theta(T_n)^2) - E(W_\theta(T_n))^2 = E(W_\theta(T_n)^2) - E(T_{n+1})^2.$$

Il s'agit donc de prouver que $E(W_\theta(T_n)^2) \leq E(T_{n+1}^2)$.

Mais nous savons d'après la question 13 que $W_\theta(T_n)^2 = \sum_{k=1}^N \theta_k^2 \mathbb{1}_{[T_n=k]}$, de sorte que

$$E(W_\theta(T_n)^2) = \sum_{k=1}^N \theta_k^2 P(T_n = k) = \sum_{k=1}^N E(T_{n+1}|T_n = k)^2 P(T_n = k).$$

Ainsi,

$$E(T_{n+1}^2) - E(W_\theta(T_n)^2) = E(T_{n+1}^2) - \sum_{k=1}^n E(T_{n+1}|T_n = k)^2 P(T_n = k) = g(\theta).$$

Mais la fonction g est positive sur \mathbf{R}^N , car il s'agit de l'espérance d'une variable aléatoire positive.

Donc $g(\theta) \geq 0$, de sorte que $E(W_\theta(T_n)^2) \leq E(T_{n+1}^2)$, et ainsi

$$V(W_\theta(T_n)) \leq V(T_{n+1}).$$

Rappel

Pour tout événement A , $\mathbb{1}_A$ suit la loi de Bernoulli de paramètre $P(A)$, et donc son espérance vaut $P(A)$.

Détails

Il s'agit d'un résultat archi-classique : si $a > 0$, alors $ax^2 + bx + c$ est minimal pour $x = \frac{-b}{2a}$.

16.a. Pour $\omega \in \Omega$, on a

$$k = 1^N k^i \mathbb{1}_{[T_n=k]}(\omega) = \begin{cases} 1 & \text{si } T_n(\omega) = 1 \\ 2^i & \text{si } T_n(\omega) = 2 \\ \vdots & \\ N^i & \text{si } T_n(\omega) = N \end{cases} = T_n^i(\omega).$$

Et donc $T_n^i = \sum_{k=1}^N k^i \mathbb{1}_{[T_n=k]}.$

16.b. En remplaçant θ par son expression, on obtient

$$W_\theta(T_n) = \sum_{k=1}^N E(T_{n+1}|T_n = k) \mathbb{1}_{[T_n=k]}.$$

Mais nous avons calculé $E(T_{n+1}|T_n = k)$ à la question 12.c :

$$E(T_{n+1}|T_n = k) = \frac{N+1}{2} + \frac{k^2 - k}{2N}$$

de sorte que

$$\begin{aligned} W_\theta(T_n) &= \sum_{k=1}^N \left(\frac{N+1}{2} + \frac{k^2 - k}{2N} \right) \mathbb{1}_{[T_n=k]} \\ &= \frac{N+1}{2} \sum_{k=1}^N \mathbb{1}_{[T_n=k]} + \frac{1}{2N} \sum_{k=1}^N (k^2 - k) \mathbb{1}_{[T_n=k]} \\ &= \frac{N+1}{2} + \frac{1}{2N} \left(\sum_{k=1}^N k^2 \mathbb{1}_{[T_n=k]} - \sum_{k=1}^N k \mathbb{1}_{[T_n=k]} \right) \\ &= \frac{N+1}{2} + \frac{1}{2N} (T_n^2 - T_n). \end{aligned}$$

Partie IV : Estimation

17.a. A la question 4.b, nous avons montré que $E(T_n) = N - d_n(N)$, de sorte que

$$E(T_n + d_n(N)) = N - d_n(N) + d_n(N) = N.$$

On serait donc tenté de dire que $T_n + d_n(N)$ est un estimateur sans biais de N . Malheureusement, cette quantité dépend de N , qui est inconnu ! En fait, $T_n + d_n(N)$ n'est pas un estimateur : on ne peut l'écrire sous la forme $T_n + d_n(N) = \varphi(U_1, \dots, U_n)$, où φ est une fonction fixée (ne dépendant pas de N).

17.b. Toujours dans la question 4.b, nous avons prouvé que $\lim_{n \rightarrow +\infty} E(T_n) = N$, et donc

(T_n) est une suite d'estimateurs asymptotiquement sans biais de N .

17.c. Supposons $A_n(\varepsilon)$ réalisé, c'est-à-dire que $|T_n - N| \geq \varepsilon$. Alors par l'inégalité triangulaire, on a

$$\varepsilon \leq |T_n - N| = |T_n - E(T_n) + d_n(N)| \leq |T_n - E(T_n)| + |d_n(N)|$$

et donc $|T_n - E(T_n)| + |d_n(N)| \geq \varepsilon$.

Ainsi, $A_n(\varepsilon) \subset B_n(\varepsilon)$.

Nous avons prouvé précédemment que $d_n(N) \xrightarrow{n \rightarrow +\infty} 0$, donc il existe $n_0 \in \mathbf{N}$ tel que pour $n \geq n_0$, $|d_n(N)| \leq \frac{\varepsilon}{2}$.

Et alors si $|T_n - E(T_n)| + |d_n(N)| \geq \varepsilon$, il vient

$$|T_n - E(T_n)| \geq \varepsilon - |d_n(N)| \geq \varepsilon - \frac{\varepsilon}{2} = \frac{\varepsilon}{2}.$$

Et donc nous venons de prouver que pour $n \geq n_0$,

$$B_n(\varepsilon) \subset \left[|T_n - E(T_n)| \geq \frac{\varepsilon}{2} \right].$$

Remarque

Notons que T_n est bien un estimateur de N .

- 17.d. Soit $\varepsilon > 0$ fixé. Alors T_n possède une variance car il s'agit d'une variable aléatoire finie. Donc par l'inégalité de Bienaymé-Tchebychev, on a

$$P(|T_n - E(T_n)| \geq \varepsilon/2) \leq \frac{2V(T_n)}{\varepsilon}.$$

Or, il a été prouvé à la question 3.c. que $V(T_n) \xrightarrow{n \rightarrow +\infty} 0$, de sorte que

$$\lim_{n \rightarrow +\infty} P(|T_n - E(T_n)| \geq \varepsilon/2) = 0.$$

De plus, on a $P(A_n(\varepsilon)) \leq P(B_n(\varepsilon))$, et pour n suffisamment grand,

$$P(B_n(\varepsilon)) \leq P(|T_n - E(T_n)| \geq \varepsilon/2).$$

On en déduit que pour n assez grand, $P(A_n(\varepsilon)) \leq P(|T_n - E(T_n)| \geq \varepsilon/2)$ et donc par le théorème des gendarmes⁴,

⁴ Une probabilité est toujours positive.

$$\lim_{n \rightarrow +\infty} P(|T_n - N| \geq \varepsilon) = \lim_{n \rightarrow +\infty} P(A_n(\varepsilon)) = 0.$$

Ainsi, (T_n) converge en probabilité vers N , et donc (T_n) est une suite d'estimateurs convergente de N .

- 18.a. Il est évident que par indépendance des U_i ,

$$P\left(\bigcap_{i=1}^n [U_i = u_i]\right) = \prod_{i=1}^n P(U_i = u_i) = \frac{1}{N^n}.$$

- 18.b. Si $\max(u_i) \neq k$, alors $\bigcap_{i=1}^n [U_i = u_i] \cap [T_n = k] = \emptyset$ donc la probabilité conditionnelle cherchée est nulle.

En revanche, si le plus grand des u_i vaut k , on a $[T_n = k] \subset \bigcap_{i=1}^n [U_i = u_i]$, de sorte que

$$\bigcap_{i=1}^n [U_i = u_i] \cap [T_n = k] = \bigcap_{i=1}^n [U_i = u_i].$$

On en déduit que

$$P_{[T_n=k]}\left(\bigcap_{i=1}^n [U_i = u_i]\right) = \frac{P\left(\bigcap_{i=1}^n [U_i = u_i]\right)}{P(T_n = k)} = \frac{1}{N^n} \frac{1}{\left(\frac{k}{N}\right)^n - \left(\frac{k-1}{N}\right)^n} = \frac{1}{k^n - (k-1)^n}.$$

- 19.a. Commençons par remarquer que cette fois, S_n est bien un estimateur de N , car T_n et Z_n le sont. De plus, on a

$$E(S_n) = E(T_n) + E(Z_n) - 1 = N - d_n(N) + d_n(N) + 1 - 1 = N.$$

Et donc S_n est un estimateur sans biais de N .

- 19.b. Par linéarité de l'espérance (rappelons que l'espérance conditionnelle est l'espérance pour la probabilité $P_{[T_n=k]}$), on a

$$E(S_n|T_n = k) = E(T_n|T_n = k) + E(Z_n|T_n = k) - 1.$$

Il est évident que $E(T_n|T_n = k) = k$.

Et nous avons montré à la question 9.b que $E(Z_n|T_n = k) = \frac{k^n}{k^n - (k-1)^n}$. On a donc

$$E(S_n|T_n = k) = k + \frac{k^n}{k^n - (k-1)^n} - 1 = \frac{k^{n+1} - k(k-1)^n + k^n + (k-1)^n - k^n}{k^n - (k-1)^n} = \frac{k^{n+1} - (k-1)^{n+1}}{k^n - (k-1)^n} = \psi_n(k).$$

19.c. D'après le théorème de transfert, on a

$$E(\psi_n(T_n)) = \sum_{k=1}^n \psi_n(k)P(T_n = k) = \sum_{k=1}^N E(S_n|T_n = k)P(T_n = k).$$

Par la formule de l'espérance totale, nous reconnaissons là $E(S_n)$, et donc

$$E(\psi_n(T_n)) = E(S_n) = N.$$

Ainsi, $\psi_n(T_n)$ est un estimateur sans biais de N .

19.d. La fonction $\lambda \mapsto E((S_n - \lambda)^2|T_n = k)$ est positive, par croissance de l'espérance (l'espérance conditionnelle n'est autre que l'espérance pour la probabilité $P_{\{T_n=k\}}$) et car la variable aléatoire $(S_n - \lambda)^2$ est positive.

Or,

$$E((S_n - \lambda)^2|T_n = k) = E(S_n^2 - 2\lambda S_n + \lambda^2|T_n = k) = E(S_n^2|T_n = k) - 2\lambda E(S_n|T_n = k) + \lambda^2 = \varphi_n(k) - 2\lambda\psi_n(k) + \lambda^2$$

Ainsi, on a une fonction polynomiale de degré 2 en λ , qui est de signe constant, donc son discriminant est négatif ou nul.

Or, ce discriminant vaut $4\psi_n(k)^2 - 4\varphi_n(k)$.

On en déduit que $\psi_n(k)^2 \leq \varphi_n(k)$.

Par la formule de Huygens, on a $V(\psi_n(T_n)) = E(\psi_n(T_n)^2) - E(\psi_n(T_n))^2 = E(\psi_n(T_n)^2) - N^2$.
Mais par le théorème de transfert,

$$E(\psi_n(T_n)^2) = \sum_{k=1}^N \psi_n(k)^2 P(T_n = k) \leq \sum_{k=1}^N \varphi_n(k) P(T_n = k) = E(S_n^2).$$

Notons que le dernier point vient de la formule de l'espérance totale appliquée à la variable aléatoire S_n^2 :

$$E(S_n^2) = \sum_{k=1}^N E(S_n^2|T_n = k)P(T_n = k).$$

Enfin, par la formule de Huygens, on a $V(S_n) = E(S_n)^2 - N^2$, et donc

$$V(\psi_n(T_n)) \leq V(S_n).$$

19.e. On a

$$V(S_n) = V(T_n + Z_n - 1) = V(T_n + Z_n) = V(T_n) + V(Z_n) + 2\text{Cov}(Z_n, T_n) = 2V(T_n) + 2\text{Cov}(T_n, Z_n) = 2V(T_n)(1 + \rho_n).$$

Mais dans les parties I et II, nous avons déjà calculé $V(T_n)$ et ρ_n , de sorte que

$$\begin{aligned} V(S_n) &= 2 \left((2N - 1)d_n(N) - 2Nd_{n+1}(N) - d_n(N)^2 \right) + 2 \left(Nd_{n+1}(N) + d_n(N) - Nd_n(N) + d_n(N)^2 \right) \\ &= 2N(d_n(N) - d_{n+1}(N)) \end{aligned}$$

Or, nous savons que

$$\lim_{n \rightarrow +\infty} d_n(N) = \lim_{n \rightarrow +\infty} d_{n+1}(N) = 0.$$

On en déduit que $V(S_n) \xrightarrow{n \rightarrow +\infty} 0$. Et donc $V(\psi_n(T_n)) \xrightarrow{n \rightarrow +\infty} 0$.

Puisque $\psi_n(T_n)$ est un estimateur sans biais de N , son risque quadratique en N est égal à sa variance, et tend donc vers 0 lorsque $n \rightarrow +\infty$, de sorte que

$\psi_n(T_n)$ est un estimateur convergent de N .

20.a. Commençons par remarquer que $f_n(T_n)$ est un estimateur sans biais de N car

$$E(f_n(T_n)) = \sum_{k=1}^N f_n(k)P(T_n = k) = \sum_{k=1}^N E(R_n|T_n = k)P(T_n = k) = E(R_n) = N.$$

De la même manière que dans la question 19.d, on peut prouver que

$$\forall k \in \llbracket 1, N \rrbracket, E(R_n|T_n = k)^2 \leq E(R_n^2|T_n = k).$$

Et alors comme en 19.d, on a

$$E(f_n(T_n)^2) \leq E(R_n^2).$$

Et donc par la formule de Huygens,

$$V(f_n(T_n)) = E(f_n(T_n)^2) - N^2 \leq E(R_n^2) - N^2 = V(R_n).$$

20.b. D'après le théorème de transfert, on a

$$E(F(T_n)) = \sum_{k=1}^N F(k)P(T_n = k).$$

Donc il est clair que si $\forall k \in \llbracket 1, N \rrbracket$, $F(k) = \psi_n(k)$, alors

$$E(F(T_n)) = \sum_{k=1}^N \psi_n(k)P(T_n = k) = N.$$

Inversement, supposons que pour tout $N \in \mathbf{N}^*$, $E(F(T_n)) = N$.

Alors on a, pour tout $N \in \mathbf{N}^*$,

$$\sum_{k=1}^N F(k) \left(\left(\frac{k}{N} \right)^n - \left(\frac{k-1}{N} \right)^n \right) = N = \sum_{k=1}^N \psi_n(k) \left(\left(\frac{k}{N} \right)^n - \left(\frac{k-1}{N} \right)^n \right)$$

soit encore

$$\sum_{k=1}^N (F(k) - \psi_n(k)) (k^n - (k-1)^n) = 0.$$

Pour $N = 1$, on a directement $F(1) = 1 = \psi_n(1)$.

Pour $N \geq 2$, en soustrayant la relation correspondant à N et celle correspondant à $N - 1$,

on a

$$0 = \sum_{k=1}^N ((F(k) - \psi_n(k)) \left(\left(\frac{k}{N} \right)^n - \left(\frac{k-1}{N} \right)^n \right)) - \sum_{k=1}^{N-1} (F(k) - \psi_n(k)) \left(\left(\frac{k}{N-1} \right)^n - \left(\frac{k-1}{N-1} \right)^n \right) = F(N) - \psi_n(N) \left(1 - \left(\frac{N-1}{N} \right)^n \right).$$

Ainsi, $\boxed{\forall N \in \mathbf{N}^*, F(N) = \psi_n(N)}$.

20.c. Soit R_n un estimateur sans biais de N , et soit comme précédemment $f_n(k) = E(R_n | T_n = k)$.

Alors nous avons prouvé à la question 20.a que $f_n(T_n)$ est un estimateur sans biais de N .

Mais d'après la question 20.b, cela impose que $f_n = \psi_n$.

Et à la question 20.a, nous avons prouvé que

$$V(\psi_n(T_n)) = V(f_n(T_n)) \leq V(R_n).$$

Donc $\boxed{\psi_n(T_n)}$ est bien de variance minimale parmi tous les estimateurs sans biais de N .

DEVOIR MAISON 17 (À RENDRE LE 21.03.19)

Dans cet exercice, n désigne un entier naturel supérieur ou égal à 2.

On note I_n la matrice identité de $\mathcal{M}_n(\mathbf{R})$ et J_n la matrice de $\mathcal{M}_n(\mathbf{R})$ dont tous les éléments valent 1.

1. a. Déterminer le rang de J_n . En déduire que 0 est valeur propre de J_n et déterminer la dimension du sous-espace propre associé.
- b. Vérifier que le vecteur V_n élément de $\mathcal{M}_{n,1}(\mathbf{R})$, dont toutes les composantes sont égales à 1, est vecteur propre de J_n .
- c. À l'aide des questions précédentes, donner les valeurs propres de J_n .

Dans toute la suite, on considère la fonction f_n définie sur \mathbf{R}^n par :

$$\forall x = (x_1, x_2, \dots, x_n) \in \mathbf{R}^n, f_n(x) = \left(\sum_{k=1}^n x_k \right) \exp \left(- \sum_{k=1}^n x_k^2 \right).$$

2. Montrer que f_n est de classe \mathcal{C}^2 sur \mathbf{R}^n .

3. a. Montrer que, pour tout i de $\llbracket 1, n \rrbracket$, on a : $\partial_i(f_n)(x) = \left(1 - 2x_i \sum_{k=1}^n x_k \right) \exp \left(- \sum_{k=1}^n x_k^2 \right)$.

- b. En déduire que f_n possède deux points critiques $a = \frac{1}{\sqrt{2n}}(1, 1, \dots, 1)$ et $b = -a$.

4. a. Déterminer les dérivées partielles d'ordre 2 de f_n .

- b. Vérifier que la hessienne de f_n en a est $H_n(a) = \frac{-2}{\sqrt{2ne}}(nI_n + J_n)$.

- c. À l'aide de la première question, donner les valeurs propres de $H_n(a)$.

- d. En déduire que f_n possède un extremum local en a .

- e. Sans refaire tous les calculs, donner une conclusion concernant le point critique b .

5. a. Étudier la fonction h qui, à tout t de \mathbf{R}_+ , associe $h(t) = te^{-t^2}$.

- b. En appliquant l'inégalité de Cauchy-Schwarz à deux vecteurs bien choisis de \mathbf{R}^n , muni de son produit scalaire canonique, montrer que :

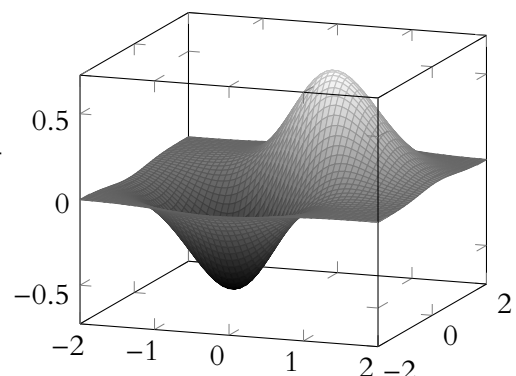
$$\forall (x_1, x_2, \dots, x_n) \in \mathbf{R}^n, \left(\sum_{k=1}^n x_k \right)^2 \leq n \sum_{k=1}^n x_k^2.$$

- c. Déduire des deux questions précédentes que f_n admet en a et en b des extrema globaux.

6. Question d'informatique.

- a. Écrire des commandes Sci Lab permettant de calculer et d'afficher $H_n(a)$ pour une valeur de n entrée par l'utilisateur.

- b. Dans le cas $n = 2$, la nappe suivante est-elle acceptable en tant que représentation graphique de la fonction f_2 ? Justifier.



CORRECTION DU DEVOIR MAISON 17

EDHEC 2018

- 1.a. Puisque toutes les colonnes de J_n sont égales, et que J_n n'est pas la matrice nulle, $\boxed{\text{rg}(J_n) = 1}$.

En particulier, elle n'est pas inversible, donc 0 est valeur propre de J_n et

$$\dim E_0(J_n) = n - \text{rg}(J_n) = \boxed{n - 1}.$$

- 1.b. On a

$$J_n V_n = \begin{pmatrix} 1 & \dots & 1 \\ \vdots & & \vdots \\ 1 & \dots & 1 \end{pmatrix} \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} = \begin{pmatrix} n \\ \vdots \\ n \end{pmatrix} = \boxed{nV_n}.$$

Puisque $V_n \neq 0$, c'est donc un $\boxed{\text{vecteur propre de } J_n, \text{ associé à la valeur propre } n}$.

- 1.c. On a $\dim E_n(J_n) \geq 1$ et $\dim E_0(J_n) = n - 1$.
Puisque d'autre part $n = \sum_{\lambda \in \text{Spec}(J_n)} \dim E_\lambda(J_n)$

Donc il n'y a pas de «place» pour d'autres valeurs propres : $\dim E_n(J_n) = 1$, et J_n ne possède pas d'autres valeurs propres, donc $\boxed{\text{Spec}(J_n) = \{0, n\}}$.

2. La fonction $(x_1, \dots, x_n) \mapsto -\sum_{k=1}^n x_k^2$ est polynomiale sur \mathbf{R}^n , donc elle y est de classe \mathcal{C}^2 .

Par composition avec la fonction exponentielle, qui est \mathcal{C}^2 sur \mathbf{R} , $(x_1, \dots, x_n) \mapsto \exp\left(-\sum_{k=1}^n x_k^2\right)$ est de classe \mathcal{C}^2 sur \mathbf{R}^n .

De plus, $(x_1, \dots, x_n) \mapsto \sum_{k=1}^n x_k$ est polynomiale sur \mathbf{R}^n , donc \mathcal{C}^2 .

Et donc, par produit de fonctions \mathcal{C}^2 , $\boxed{f_n \text{ est une fonction de classe } \mathcal{C}^2 \text{ sur } \mathbf{R}^n}$.

- 3.a. Soit $i \in \llbracket 1, n \rrbracket$. Alors par dérivation d'un produit,

$$\partial_i f_n(x) = -2x_i \left(\sum_{k=1}^n x_k \right) \exp\left(-\sum_{k=1}^n x_k^2\right) + \exp\left(-\sum_{k=1}^n x_k^2\right) = \left(1 - 2x_i \sum_{k=1}^n x_k\right) \exp\left(-\sum_{k=1}^n x_k^2\right).$$

- 3.b. Un point $(x_1, \dots, x_n) \in \mathbf{R}^n$ est un point critique de f_n si et seulement si

$$\forall i \in \llbracket 1, n \rrbracket, \left(1 - 2x_i \sum_{k=1}^n x_k\right) \exp\left(-\sum_{k=1}^n x_k^2\right) = 0 \Leftrightarrow 2x_i \sum_{k=1}^n x_k = 1.$$

Notons qu'alors, automatiquement, $\sum_{k=1}^n x_k \neq 0$, et donc pour tout $i \in \llbracket 1, n \rrbracket$, $x_i = \frac{1}{2 \sum_{k=1}^n x_k}$.

En particulier, on doit avoir $x_1 = x_2 = \dots = x_n$, de sorte que $\sum_{k=1}^n x_k = nx_1$.

Et alors $x_1 = \frac{1}{2nx_1} \Leftrightarrow x_1^2 = \frac{1}{2n} \Leftrightarrow x_1 = \pm \frac{1}{\sqrt{2n}}$.

Ainsi, f_n possède deux points critiques qui sont

$$a = \frac{1}{\sqrt{2n}}(1, 1, \dots, 1) \text{ et } b = -\frac{1}{\sqrt{2n}}(1, 1, \dots, 1) = -a.$$

- 4.a. Soient $(i, j) \in \llbracket 1, n \rrbracket^2$, avec $i \neq j$. Alors en dérivant $\partial_i f_n$ par rapport à x_j , on a

$$\partial_{i,j}^2 f_n(x) = -2x_j \left(1 - 2x_i \sum_{k=1}^n x_k\right) \exp\left(-\sum_{k=1}^n x_k^2\right) - 2x_i \exp\left(-\sum_{k=1}^n x_k^2\right) = -2 \left(x_i + x_j - 2x_i x_j \sum_{k=1}^n x_k\right) \exp\left(-\sum_{k=1}^n x_k^2\right).$$

⚠ Attention !

On ne peut pas se contenter de mentionner que les colonnes sont proportionnelles pour justifier qu'une matrice est de rang 1.
La matrice nulle a aussi ses colonnes proportionnelles, pourtant son rang est nul !

Égalité

Notons qu'en général on a juste une inégalité

$$\sum_{\lambda \in \text{Spec}(J_n)} \dim E_\lambda(J_n) \leq n,$$

mais J_n est symétrique, donc diagonalisable, et nous pouvons donc affirmer qu'il s'agit d'une égalité.

Et pour $i \in \llbracket 1, n \rrbracket$,

$$\begin{aligned} \partial_{i,i}^2 f_n(x) &= \left(-2 \sum_{k=1}^n x_k - 2x_i \right) \exp \left(- \sum_{k=1}^n x_k^2 \right) - 2x_i \left(1 - 2x_i \sum_{k=1}^n x_k \right) \exp \left(- \sum_{k=1}^n x_k^2 \right) \\ &= -2 \left(2x_i + \sum_{k=1}^n x_k - 2x_i^2 \sum_{k=1}^n x_k \right) \exp \left(- \sum_{k=1}^n x_k^2 \right). \end{aligned}$$

4.b. En a , $-\sum_{k=1}^n x_k^2 = -\sum_{k=1}^n \frac{1}{2n} = -\frac{1}{2}$. Il vient alors, pour $i \neq j$,

$$\partial_{i,j}^2 f_n(a) = -2 \left(\frac{1}{\sqrt{2n}} + \frac{1}{\sqrt{2n}} - \frac{2}{2n} \frac{n}{\sqrt{2n}} \right) e^{-1/2} = \frac{-2}{\sqrt{2ne}}.$$

Et pour $i = j$,

$$\partial_{i,i}^2 f_n(a) = -2 \left(\frac{2}{\sqrt{2n}} + \frac{n}{\sqrt{2n}} - \frac{2}{2n} \frac{n}{\sqrt{2n}} \right) e^{-1/2} = -2 \frac{n+1}{\sqrt{2ne}}.$$

Ainsi, les coefficients de $H_n(a)$ valent tous $-\frac{2}{\sqrt{2ne}}$, sauf les coefficients diagonaux, qui

valent $-\frac{2(n+1)}{\sqrt{2ne}} = -\frac{2}{\sqrt{2ne}}(n+1)$.

On reconnaît là les coefficients de $-\frac{2}{\sqrt{2ne}}(nI_n + J_n)$, et donc

$$H_n(a) = -\frac{2}{\sqrt{2ne}}(nI_n + J_n).$$

4.c. Un réel λ est valeur propre de $nI_n + J_n$ si et seulement si $nI_n + J_n - \lambda I_n$ n'est pas inversible, soit si et seulement si $J_n - (\lambda - n)I_n$ n'est pas inversible, si et seulement si $\lambda - n$ est valeur propre de J_n .

Puisque les valeurs propres de J_n sont 0 et n , c'est le cas si et seulement si $\lambda - n = 0 \Leftrightarrow \lambda = n$ ou $\lambda - n = n \Leftrightarrow \lambda = 2n$.

Donc les valeurs propres de $nI_n + J_n$ sont n et $2n$.

Par conséquent, les valeurs propres de $H(a) = -\frac{2}{\sqrt{2ne}}(nI_n + J_n)$ sont $\frac{-2n}{\sqrt{2ne}}$ et $\frac{-4n}{\sqrt{2ne}}$.

Une autre méthode, «à la main» : puisque J_n est diagonalisable, et que nous connaissons ses valeurs propres, il existe une matrice P inversible telle que $J_n = P^{-1} \text{Diag}(0, \dots, 0, n)P$. Mais $I_n = P^{-1}I_nP$, et donc

$$nI_n + J_n = P^{-1}nI_nP + P^{-1} \text{Diag}(0, \dots, 0, n)P = P^{-1} (nI_n + \text{Diag}(0, \dots, 0, n))P = P^{-1} \text{Diag}(n, \dots, n, 2n)P.$$

Et donc $nI_n + J_n$ est semblable à une matrice diagonale dont les coefficients diagonaux sont n et $2n$: $\text{Spec}(nI_n + J_n) = \{n, 2n\}$.

On conclut comme précédemment pour les valeurs propres de $H_n(a) = -\frac{2}{\sqrt{2ne}}(nI_n + J_n)$.

4.d. Les valeurs propres de $H_n(a)$ sont toutes deux strictement négatives, donc f_n possède un maximum local en a .

4.e. Il est facile de vérifier que si l'on change les x_i en leurs opposés, alors on change le signe des dérivées secondes de f_n :

$$\partial_{i,j}^2 f_n(-x_1, \dots, -x_n) = -\partial_{i,j}^2 f_n(x_1, \dots, x_n).$$

Et donc $H_n(b) = -H_n(a)$. Par conséquent, les valeurs propres de $H_n(b)$ sont toutes deux strictement positives, et donc f_n possède un minimum local en b .

5.a. La fonction h est dérivable sur \mathbf{R}_+ , et on a $h'(t) = -2t^2 e^{-t^2} + e^{-t^2} = e^{-t^2} (1 - 2t^2)$.

Par conséquent, on a $h'(t) \geq 0 \Leftrightarrow 1 - 2t^2 \geq 0 \Leftrightarrow t \leq \frac{1}{\sqrt{2}}$.

Et donc le tableau de variation de h est :

Astuce

Pour toute matrice inversible P , $I_n = P^{-1}I_nP$. Cela revient à dire que dans toute base, la matrice de l'endomorphisme identité de \mathbf{R}^n est la matrice identité.

t	0	$1/\sqrt{2}$	$+\infty$
$h'(t)$		+	-
$h(t)$	0		0

On en déduit que h possède un maximum en $t = \frac{1}{\sqrt{2}}$, et ce maximum vaut $h\left(\frac{1}{\sqrt{2}}\right) = \frac{1}{\sqrt{2e}}$.

5.b. Appliquons l'inégalité de Cauchy-Schwarz aux vecteurs (x_1, \dots, x_n) et $(1, \dots, 1)$. Alors

$$\left(\sum_{k=1}^n x_k \times 1\right)^2 \leq \left(\sum_{k=1}^n 1^2\right)\left(\sum_{k=1}^n x_k^2\right) \Leftrightarrow \left(\sum_{k=1}^n x_k\right)^2 \leq n\left(\sum_{k=1}^n x_k^2\right).$$

5.c. Soit $(x_1, \dots, x_n) \in \mathbf{R}^n$. Alors en utilisant les deux questions précédentes,

$$|f_n(x)| = \sqrt{\left(\sum_{k=1}^n x_k\right)^2} \exp\left(-\sum_{k=1}^n x_k^2\right) \leq \sqrt{n}\left(\sum_{k=1}^n x_k^2\right) \exp\left(-\sum_{k=1}^n x_k^2\right) = \sqrt{n}h\left(\sum_{k=1}^n x_k^2\right) \leq \sqrt{n} \frac{1}{\sqrt{2e}}.$$

Ainsi, on a $-\sqrt{\frac{n}{2e}} \leq f_n(x) \leq \sqrt{\frac{n}{2e}}$. Mais $\sqrt{\frac{n}{2e}} = f_n(a)$ et $-\sqrt{\frac{n}{2e}} = f_n(b)$.

Donc pour tout $x \in \mathbf{R}^n$, $f_n(b) \leq f_n(x) \leq f_n(a)$: f_n possède un maximum global en a et un minimum global en b .

6. Question d'informatique.

6.a. Le script suivant fonctionne :

```

1 function H(n)
2     M = -2/sqrt(2*n*e) * (n*eye(n,n) + ones(n,n))
3     disp(M)
4 endfunction
    
```

6.b. La nappe tracée correspond à une fonction qui possède un maximum local (et probablement¹ un maximum global) et un minimum local.

De plus, le maximum local est atteint en un point dont les deux coordonnées sont positives, qui pourrait donc être $a = \left(\frac{1}{2}, \frac{1}{2}\right)$, et de même le minimum local semble atteint en un point qui pourrait être b .

Donc avec les informations dont on dispose, il est plausible que la nappe soit une représentation de la fonction f_2 .

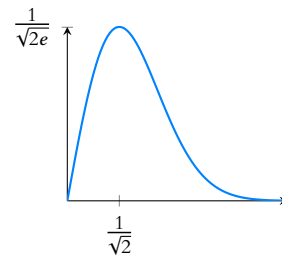


FIGURE 23.1– La fonction h .

Méthode

Comment choisir les bons vecteurs ? On sait que l'inégalité de Cauchy-Schwarz affirme que

$$\left(\sum_{k=1}^n a_k b_k\right)^2 \leq \left(\sum_{k=1}^n a_k^2\right)\left(\sum_{k=1}^n b_k^2\right)$$

Au vu du résultat demandé, on veut donc $a_k b_k = x_k$ et $b_k^2 = x_k^2$.

¹ Vu que nous ne voyons qu'une partie de la fonction, il est impossible d'être certain qu'elle ne prend pas ailleurs des valeurs plus grandes.

DEVOIRS EN CLASSE

Devoir surveillé 1 (08.09.18) 888
Correction 891
Devoir surveillé 2 (29.09.18) 899
Correction 902
Concours blanc 1 (07.11.18) 912
Correction 916
Devoir surveillé 4 (01.12.18) 928
Correction 931
Devoir surveillé 5 (12.01.19) 939
Correction 942
Devoir surveillé 6 (09.02.19) 951
Correction 955
Concours blanc 2 (05.03.19) 965
Correction 969
Concours blanc 2 : Maths II (08.03.19) 981
Correction 985
Devoir surveillé 8 (23.03.19) 998
Correction 1001
Devoir surveillé 8 (HEC) (23.03.19) 1010
Correction 1014
Devoir surveillé 8 : Maths II (23.03.19) 1027
Correction 1031

DEVOIR SURVEILLÉ 1 (08.09.18)

EXERCICE 1

On considère la matrice carrée $A = \begin{pmatrix} 1 & 2 \\ 3 & 6 \end{pmatrix}$ et les deux vecteurs colonnes $U = \begin{pmatrix} -2 \\ 1 \end{pmatrix}$ et $V = \begin{pmatrix} 1 \\ 3 \end{pmatrix}$.

1. Montrer que A n'est pas inversible.
2. Calculer AU et AV . Retrouver alors le résultat de la question précédente.

Dans la suite de cet exercice, on considère l'application f qui, à toute matrice M de $\mathcal{M}_2(\mathbf{R})$, associe $f(M) = AM$.

3. Montrer que f est un endomorphisme de $\mathcal{M}_2(\mathbf{R})$.
4.
 - a. Déterminer une base \mathcal{B} de $\text{Ker}(f)$ et vérifier que $\text{Ker}(f)$ est de dimension 2.
 - b. En déduire la dimension de $\text{Im}(f)$.
 - c. Montrer que $f(V^tU) = 7V^tU$ en déduire que $V^tU \in \text{Im}(f)$.
Montrer de même que $V^tV \in \text{Im}(f)$, et en déduire que $\mathcal{B}' = (V^tU, V^tV)$ est une base de $\text{Im}(f)$.
5. Montrer que la concaténation de \mathcal{B} et \mathcal{B}' forme une base de $\mathcal{M}_2(\mathbf{R})$. On note \mathcal{C} cette base.
6. Vérifier que la matrice de f dans la base \mathcal{C} est une matrice diagonale.
7. Les sous-espaces $\text{Im}(f)$ et $\text{Ker}(f)$ sont-ils supplémentaires dans $\mathcal{M}_2(\mathbf{R})$? Justifier votre réponse.

EXERCICE 2

On considère la fonction $f :]0, +\infty[\rightarrow \mathbf{R}$ définie, pour tout $x \in]0, +\infty[$ par :

$$f(x) = e^x - e \ln(x).$$

On admet les encadrements numériques suivants : $2,7 < e < 2,8$ $7,3 < e^2 < 7,4$ $0,6 < \ln(2) < 0,7$.

Partie I : Étude de la fonction f .

1.
 - a. Montrer que f est deux fois dérivable sur $]0, +\infty[$ et calculer, pour tout x de $]0, +\infty[$, $f'(x)$ et $f''(x)$.
 - b. Dresser le tableau de variations de f' avec les limites en 0 et en $+\infty$, et préciser $f'(1)$.
2. Dresser le tableau de variations de f avec les limites en 0 et en $+\infty$, ainsi que la valeur de $f(1)$.
3. Tracer l'allure de la courbe représentative de f .
4.
 - a. Étudier les variations de la fonction $u :]0, +\infty[\rightarrow \mathbf{R}$, définie par $u(x) = f'(x) - x$.
 - b. En déduire que l'équation $f'(x) = x$, d'inconnue $x \in]0, +\infty[$, admet une solution et une seule, notée α , et montrer : $1 < \alpha < 2$.

Partie II : Étude d'une suite et d'une série.

On considère la suite réelle $(u_n)_{n \in \mathbf{N}}$ définie par :

$$u_0 = 2 \text{ et pour tout } n \in \mathbf{N}, u_{n+1} = f(u_n).$$

5. Montrer que, pour tout n de \mathbf{N} , u_n existe et $u_n \geq 2$.
6.
 - a. Étudier les variations, puis le signe, de la fonction $g : [2; +\infty[\rightarrow \mathbf{R}, x \mapsto f(x) - x$.
 - b. En déduire que la suite $(u_n)_{n \in \mathbf{N}}$ est croissante.
7. Démontrer que la suite $(u_n)_{n \in \mathbf{N}}$ admet $+\infty$ pour limite.
8. Écrire un programme Scilab qui, étant donné un réel A , renvoie un entier naturel N tel que $u_N \geq A$.
9.
 - a. Démontrer : $\forall x \in [2, +\infty[, 2 \ln(x) \leq x \leq \frac{e^x}{3}$.
 - b. En déduire : $\forall n \in \mathbf{N}, u_{n+1} \geq \frac{6-e}{2} u_n$.
 - c. Déterminer la nature de la série de terme général $\frac{1}{u_n}$.

Partie III : Étude d'intégrales généralisées.

10. Montrer que l'intégrale $\int_0^1 f(x) dx$ converge et calculer cette intégrale.
11. L'intégrale $\int_0^{+\infty} f(x) dx$ converge-t-elle ?
12. Montrer que l'intégrale $\int_2^{+\infty} \frac{dx}{f(x)}$ converge. On pourra à cet effet utiliser le résultat de la question 9.a.

EXERCICE 3

La lettre n désigne un entier naturel non nul.

Soit f_n la fonction définie sur \mathbf{R}_+ par : $\forall x \in \mathbf{R}_+, f_n(x) = 1 - x - x^n$.

1. Montrer que l'équation $f_n(x) = 0$, d'inconnue x possède une seule solution notée u_n .
2.
 - a. Vérifier que u_n appartient à $]0, 1[$.
 - b. En déduire le signe de $f_{n+1}(u_n)$ puis établir que la suite (u_n) est croissante.
 - c. Conclure que la suite (u_n) converge et que sa limite appartient à $[0, 1]$.
 - d. Montrer par l'absurde que $\lim_{n \rightarrow +\infty} u_n = 1$.
3. Pour tout entier naturel n non nul, on pose $v_n = 1 - u_n$.
 - a. Justifier que v_n est strictement positif, puis montrer que $\ln v_n \underset{n \rightarrow +\infty}{\sim} -nv_n$.
 - b. Établir que $\lim_{n \rightarrow +\infty} \frac{\ln\left(\frac{-\ln v_n}{nv_n}\right)}{\ln v_n} = 0$ et en déduire que : $\ln v_n \underset{n \rightarrow +\infty}{\sim} -\ln n$.
 - c. Montrer enfin que : $v_n \underset{n \rightarrow +\infty}{\sim} \frac{\ln n}{n}$.
4. Donner la nature des séries de termes généraux v_n et v_n^2 .

PROBLÈME

Soit n un entier naturel non nul.

Dans une fête foraine, un stand propose le jeu suivant : le joueur lance n fois une pièce et compte le nombre de Pile obtenus. Si ce nombre est pair, le joueur est déclaré vainqueur, et s'il est impair, il est déclaré perdant.

Si le joueur est déclaré vainqueur, il gagne 10 euros pour chaque Pile obtenu, mais s'il a perdu, il doit payer 10 euros pour chaque Pile obtenu.

En particulier, s'il n'obtient aucun Pile, il est déclaré vainqueur, mais ne remporte rien. La pièce est truquée, et à chaque lancer, la probabilité d'obtenir Pile est égale à p ($p \in]0, 1[$), et celle d'obtenir Face est de $1 - p$.

On notera X la variable aléatoire égale au nombre de Pile obtenus, et G la variable aléatoire égale au gain algébrique du joueur.

Enfin, on notera A l'événement : «le joueur est déclaré vainqueur» et on dira que le jeu est favorable au joueur si l'espérance mathématique de la variable aléatoire G est positive.

Partie I

Dans cette partie, on suppose que $n = 3$ et $p = \frac{2}{3}$.

1. Reconnaître la loi de X , puis vérifier que : $P(A) = \frac{13}{27}$.
2. Montrer que $G(\Omega) = \{-30, -10, 0, 20\}$, puis expliciter la loi de G .
3. Calculer l'espérance de la variable aléatoire G . Le jeu est-il favorable au joueur ?

Partie II

Dans cette partie, on revient au cas général, où n est un entier naturel non nul et $p \in]0, 1[$.

Celui qui tient le stand souhaite rendre le jeu plus attractif en affichant le slogan «À ce jeu, il y a plus de gagnants que de perdants !», et cherche donc les conditions nécessaires sur p et n pour que son affichage ne soit pas mensonger.

Soit Y la variable aléatoire définie par : $Y = (-1)^X$.

Autrement dit, Y prend la valeur 1 lorsque X prend une valeur paire, et Y prend la valeur -1 lorsque X prend une valeur impaire.

4. a. On note $Z = \frac{Y+1}{2}$. Déterminer $Y(\Omega)$, puis montrer que Z suit une loi de Bernoulli de paramètre $P(A)$.
 b. Démontrer que : $E(Y) = 2P(A) - 1$.
 5. a. Donner la loi de X .
 b. En déduire qu'on a également :

$$E(Y) = \sum_{k=0}^n (-1)^k \binom{n}{k} p^k (1-p)^{n-k}$$

puis que : $E(Y) = (1 - 2p)^n$.

- c. Exprimer alors la valeur de $P(A)$ en fonction de n et p .
 d. Démontrer que :

$$P(A) \geq \frac{1}{2} \Leftrightarrow \left[p \leq \frac{1}{2} \text{ ou «n est pair»} \right].$$

Partie III

Le concepteur du jeu souhaite cependant vérifier que, tout en laissant son jeu attractif (c'est-à-dire en faisant en sorte que $P(A) \geq \frac{1}{2}$), son activité soit rentable pour lui, autrement dit que le jeu soit défavorable au joueur (c'est-à-dire que $E(G) \leq 0$).

6. Exprimer G en fonction de X et Y . En déduire que :

$$E(G) = 10 \sum_{k=0}^n (-1)^k k P(X = k).$$

7. Montrer que : $E(G) = -10np(1 - 2p)^{n-1}$.
 8. Démontrer alors que :

$$\begin{cases} P(A) \geq \frac{1}{2} \\ E(G) \leq 0 \end{cases} \Leftrightarrow p \leq \frac{1}{2}.$$

9. a. Étudier les variations de la fonction f définie sur $\left[0, \frac{1}{2}\right]$ par :

$$\forall x \in \left[0, \frac{1}{2}\right], f(x) = x(1 - 2x)^{n-1}.$$

- b. Pour une valeur de n fixée, comment le concepteur du jeu doit-il truquer sa pièce (c'est-à-dire quelle valeur doit-il donner à $p \in \left[0, \frac{1}{2}\right]$) pour optimiser la rentabilité de son activité ?

CORRECTION DU DEVOIR SURVEILLÉ 1

EXERCICE 1 (ADAPTÉ DE EDHEC ECE 2018)

1. On a $\det(A) = 6 - 3 \times 2 = 0$, et donc A n'est pas inversible.

2. On a $AU = \begin{pmatrix} 1 & 2 \\ 3 & 6 \end{pmatrix} \begin{pmatrix} -2 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ et $AV = \begin{pmatrix} 1 & 2 \\ 3 & 6 \end{pmatrix} \begin{pmatrix} 1 \\ 3 \end{pmatrix} = \begin{pmatrix} 7 \\ 21 \end{pmatrix}$.

Puisque $AU = 0$, alors que $U \neq 0$, on retrouve bien le fait que A n'est pas inversible.

En effet, si A était inversible, alors par multiplication à gauche par A^{-1} de la relation $AU = 0$, il viendrait $U = A^{-1}AU = A^{-1} \begin{pmatrix} 0 \\ 0 \end{pmatrix} = 0_{\mathcal{M}_{2,1}(\mathbf{R})}$.

3. Puisque A est une matrice de $\mathcal{M}_2(\mathbf{R})$, si $M \in \mathcal{M}_2(\mathbf{R})$, alors on a encore $f(M) = AM \in \mathcal{M}_2(\mathbf{R})$. Soient $M, N \in \mathcal{M}_2(\mathbf{R})$, et soit $\lambda \in \mathbf{R}$. Alors

$$f(\lambda M + N) = A(\lambda M + N) = \lambda AM + AN = \lambda f(M) + f(N).$$

Donc f est linéaire, et étant à valeurs dans $\mathcal{M}_2(\mathbf{R})$, c'est bien un endomorphisme de $\mathcal{M}_2(\mathbf{R})$.

4.a. Soit $M = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathcal{M}_2(\mathbf{R})$. Alors

$$\begin{aligned} M \in \text{Ker}(f) &\Leftrightarrow AM = 0 \Leftrightarrow \begin{pmatrix} 1 & 2 \\ 3 & 6 \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix} = 0 \Leftrightarrow \begin{pmatrix} a + 2c & b + 2d \\ 3a + 6c & 3b + 6d \end{pmatrix} = 0 \\ &\Leftrightarrow \begin{cases} a + 2c = 0 \\ b + 2d = 0 \\ 3a + 6c = 0 \\ 3b + 6d = 0 \end{cases} \Leftrightarrow \begin{cases} a + 2c = 0 \\ b + 2d = 0 \end{cases} \Leftrightarrow \begin{cases} a = -2c \\ b = -2d \end{cases} \end{aligned}$$

Et donc

$$M \in \text{Ker}(f) \Leftrightarrow M = \begin{pmatrix} -2c & -2d \\ c & d \end{pmatrix} \Leftrightarrow M = c \begin{pmatrix} -2 & 0 \\ 1 & 0 \end{pmatrix} + d \begin{pmatrix} 0 & -2 \\ 0 & 1 \end{pmatrix} \Leftrightarrow M \in \text{Vect} \left(\begin{pmatrix} -2 & 0 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & -2 \\ 0 & 1 \end{pmatrix} \right).$$

Par conséquent, la famille $\left(\begin{pmatrix} -2 & 0 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & -2 \\ 0 & 1 \end{pmatrix} \right)$ est génératrice de $\text{Ker}(f)$.

Puisque d'autre part ces deux matrices ne sont pas colinéaires, elles forment une famille libre, et donc une base de $\text{Ker}(f)$.

Cette base possédant deux éléments, on en déduit que $\dim \text{Ker}(f) = 2$.

4.b. Par le théorème du rang appliqué à f , on a

$$\dim \text{Ker}(f) + \dim \text{Im}(f) = \dim \mathcal{M}_2(\mathbf{R}) = 4.$$

Et donc $\dim \text{Im}(f) = 2$.

4.c. Puisque $U \in \mathcal{M}_{2,1}(\mathbf{R})$, tU est dans $\mathcal{M}_{1,2}(\mathbf{R})$, et donc le produit V^tU est dans $\mathcal{M}_{2,2}(\mathbf{R})$. Plus généralement, le produit d'un vecteur colonne par un vecteur ligne est une matrice carrée.

On a $V^tU = \begin{pmatrix} 1 \\ 3 \end{pmatrix} \begin{pmatrix} -2 & 1 \end{pmatrix} = \begin{pmatrix} -2 & 1 \\ -6 & 3 \end{pmatrix}$ et donc

$$f(V^tU) = \begin{pmatrix} 1 & 2 \\ 3 & 6 \end{pmatrix} \begin{pmatrix} -2 & 1 \\ -6 & 3 \end{pmatrix} = \begin{pmatrix} -14 & 7 \\ -42 & 21 \end{pmatrix} = 7V^tU.$$

On a donc $V^tU = \frac{1}{7}f(V^tU) = f\left(\frac{1}{7}V^tU\right) \in \text{Im}(f)$.

De même, on a $V^tV = \begin{pmatrix} 1 & 3 \\ 3 & 9 \end{pmatrix}$ et $f(V^tV) = \begin{pmatrix} 7 & 21 \\ 21 & 63 \end{pmatrix} = 7V^tV$.

On en déduit que $V^tV = f\left(\frac{1}{7}V^tV\right) \in \text{Im}(f)$.

Puisque les deux matrices V^tV et V^tU ne sont pas proportionnelles, elles forment une famille libre de $\text{Im}(f)$.

Puisque cette famille est de cardinal $2 = \dim \text{Im}(f)$, c'est une base de $\text{Im}(f)$.

Rappel

Une matrice 2×2 est inversible si et seulement si son déterminant est non nul.

Détails

La troisième ligne est proportionnelle à la première, donc on peut ne garder qu'une des deux.

De même, la quatrième est proportionnelle à la seconde.

Détails

Une famille de deux vecteurs est libre si et seulement si ils ne sont pas colinéaires, c'est-à-dire si l'un n'est pas un multiple de l'autre.

Image

Rappelons que $\text{Im}(f)$ est l'ensemble des éléments qui ont au moins un antécédent par f , c'est-à-dire qui s'écrivent sous la forme $f(\star)$, où \star désigne n'importe quel vecteur de E .

Il n'y a donc rien à justifier pour affirmer que $f\left(\frac{1}{7}V^tU\right)$ est dans $\text{Im} f$: c'est l'image d'un vecteur de E (en l'occurrence, le vecteur $\frac{1}{7}V^tU$), et donc il est automatiquement dans $\text{Im} f$.

5. Puisque \mathcal{B} et \mathcal{B}' sont de cardinal 2, leur concaténation est une famille de $4 = \dim \mathcal{M}_2(\mathbf{R})$ éléments de $\mathcal{M}_2(\mathbf{R})$. C'est donc une base de $\mathcal{M}_2(\mathbf{R})$ si et seulement si elle est libre. Soient donc $\alpha, \beta, \gamma, \delta$ des réels tels que

$$\alpha \begin{pmatrix} -2 & 0 \\ 1 & 0 \end{pmatrix} + \beta \begin{pmatrix} 0 & -2 \\ 0 & 1 \end{pmatrix} + \gamma \begin{pmatrix} -2 & 1 \\ -6 & 3 \end{pmatrix} + \delta \begin{pmatrix} 1 & 3 \\ 3 & 9 \end{pmatrix} = 0.$$

Il vient alors

$$\begin{pmatrix} -2\alpha - 2\gamma + \delta & -2\beta + \gamma + 3\delta \\ \alpha - 6\gamma + 3\delta & \beta + 3\gamma + 9\delta \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \Leftrightarrow \begin{cases} -2\alpha - 2\gamma + \delta = 0 \\ -2\beta + \gamma + 3\delta = 0 \\ \alpha - 6\gamma + 3\delta = 0 \\ \beta + 3\gamma + 9\delta = 0 \end{cases} \begin{matrix} \Leftrightarrow \\ \begin{matrix} L_3 \leftarrow 2L_3 + L_1 \\ L_4 \leftarrow 2L_4 + L_2 \end{matrix} \end{matrix} \Leftrightarrow \begin{cases} -2\alpha - 2\gamma + \delta = 0 \\ -2\beta + \gamma + 3\delta = 0 \\ -14\gamma + 7\delta = 0 \\ 7\gamma + 21\delta = 0 \end{cases}$$

$$\begin{matrix} \Leftrightarrow \\ L_4 \leftarrow 2L_4 + L_3 \end{matrix} \begin{cases} -2\alpha - 2\gamma + \delta = 0 \\ -2\beta + \gamma + 3\delta = 0 \\ -14\gamma + 7\delta = 0 \\ 49\delta = 0 \end{cases} \Leftrightarrow \begin{cases} 2\alpha = -2\gamma + \delta \\ 2\beta = \gamma + 3\delta \\ 14\gamma = 7\delta \\ \delta = 0 \end{cases} \Leftrightarrow \alpha = \beta = \gamma = \delta = 0.$$

Ainsi, \mathcal{C} est une famille libre de $\mathcal{M}_2(\mathbf{R})$, et donc en est une base.

6. Notons $E_1 = \begin{pmatrix} -2 & 0 \\ 1 & 0 \end{pmatrix}, E_2 = \begin{pmatrix} 0 & -2 \\ 0 & 1 \end{pmatrix}, E_3 = V^t U$ et $E_4 = V^t V$, de sorte que $\mathcal{C} = (E_1, E_2, E_3, E_4)$.
Puisque E_1 et E_2 sont dans $\text{Ker}(f)$, on a $f(E_1) = f(E_2) = 0$, et nous savons d'après la question 4.c que $f(E_3) = 7E_3$ et $f(E_4) = 7E_4$.
Donc la matrice de f dans la base \mathcal{C} est

$$\text{Mat}_{\mathcal{C}}(f) = \begin{pmatrix} f(E_1) & f(E_2) & f(E_3) & f(E_4) \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 7 & 0 \\ 0 & 0 & 0 & 7 \end{pmatrix} \begin{matrix} E_1 \\ E_2 \\ E_3 \\ E_4 \end{matrix}$$

qui est bien une matrice diagonale.

7. Puisque nous savons déjà que $\dim \text{Im}(f) + \dim \text{Ker}(f) = \dim \mathcal{M}_2(\mathbf{R})$, une option serait de déterminer $\text{Im}(f) \cap \text{Ker}(f)$: si cette intersection est égale à $\{0_{\mathcal{M}_2(\mathbf{R})}\}$, alors $\text{Im}(f)$ et $\text{Ker}(f)$ sont supplémentaires, sinon ils ne le sont pas.
Mais nous disposons d'une autre caractérisation des sous-espaces supplémentaires : deux sous-espaces F et G d'un espace vectoriel E sont supplémentaires si et seulement si la concaténation d'une base de F et d'une base de G est une base de E .
Or, nous avons prouvé à la question 5 que la concaténation de la base \mathcal{B} de $\text{Ker}(f)$ et de la base \mathcal{B}' de $\text{Im}(f)$ est une base de $\mathcal{M}_2(\mathbf{R})$.
Et donc nous pouvons directement conclure que $\mathcal{M}_2(\mathbf{R}) = \text{Ker}(f) \oplus \text{Im}(f)$: $\text{Ker}(f)$ et $\text{Im}(f)$ sont supplémentaires dans $\mathcal{M}_2(\mathbf{R})$.

EXERCICE 2 (EML ECE 2017)

Partie I : Étude de la fonction f .

- 1.a. Les fonctions exponentielle et logarithme sont deux fois dérivables sur \mathbf{R}_+^* , donc f est deux fois dérivable car différence de fonctions deux fois dérivables.

On a alors $f'(x) = e^x - \frac{e}{x}$ et $f''(x) = e^x + \frac{e}{x^2}$.

- 1.b. La fonction f'' est strictement positive sur \mathbf{R}_+^* et donc f' y est strictement croissante. De plus, on a $\lim_{x \rightarrow 0^+} f'(x) = -\infty$ et $\lim_{x \rightarrow +\infty} f'(x) = +\infty$.

Enfin, on a $f'(1) = e - e = \span style="border: 1px solid black; padding: 2px;">0.$

2. On a $\lim_{x \rightarrow 0^+} f(x) = +\infty$.

Et puisque $\ln(x) = o_{x \rightarrow +\infty}(e^x)$, on a $f(x) \sim_{x \rightarrow +\infty} e^x$ de sorte que $\lim_{x \rightarrow +\infty} f(x) = +\infty$. Enfin,

on a $f(1) = e$.

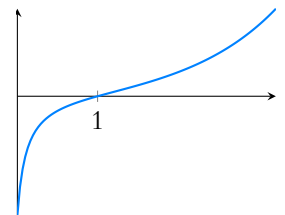
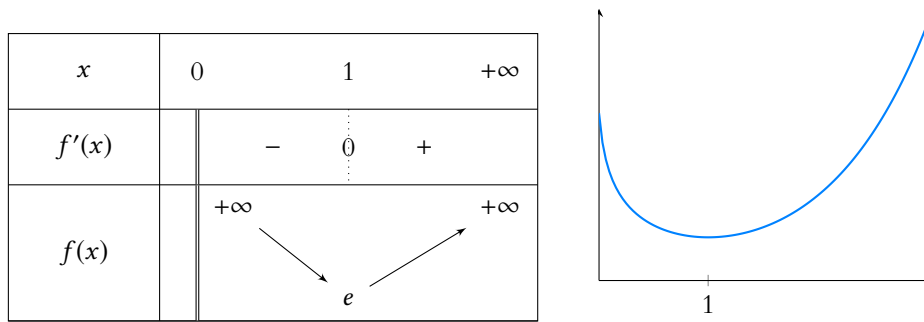


FIGURE 24.1– La fonction f' .

FIGURE 24.2 – La fonction f et son tableau de variations.

4.a. On a donc $u(x) = e^x - \frac{e}{x} - x$.

Cette fonction est dérivable sur \mathbf{R}_+^* et $u'(x) = e^x + \frac{e}{x^2} - 1 = \frac{x^2(e^x - 1) + e}{x^2}$.

Pour $x > 0$, $e^x - 1 > 0$, et donc $u'(x) > 0$.

On en déduit que la fonction u est strictement croissante sur \mathbf{R}_+^* .

4.b. La fonction u est continue sur \mathbf{R}_+^* car dérivable, elle y est strictement croissante, et on a $\lim_{x \rightarrow 0^+} u(x) = -\infty$ et $\lim_{x \rightarrow +\infty} u(x) = +\infty$.

Par le théorème de la bijection, il existe donc un unique $x \in]0, +\infty[$ tel que $u(x) = 0 \Leftrightarrow f'(x) = x$.

Ainsi, il existe un unique $\alpha \in \mathbf{R}_+^*$ tel que $f'(\alpha) = \alpha$.

D'autre part $u(1) = f'(1) - 1 = -1 < 0 = u(\alpha)$. Par croissance de u , on a donc $1 < \alpha$. De

même, $u(2) = e^2 - \frac{e}{2} - 2 > 7,3 - \frac{2,7}{2} - 2 > 5,3 - 1,35 > 0 = u(\alpha)$, et donc $\alpha < 2$.

Partie II : Étude d'une suite et d'une série.

5. Montrons par récurrence sur n que u_n est bien défini et supérieur à 2.

Il est évident que u_0 existe, et $u_0 \geq 2$.

Supposons donc que u_n existe et est supérieur ou égal à 2.

Alors $u_{n+1} = f(u_n)$ existe bien, puisque $u_n \geq 2 > 0$, et puisque f est croissante sur $[1, +\infty[$, on a alors

$$u_{n+1} = f(u_n) \geq f(2) = e^2 - e \ln(2) > 7,3 - 7,4 \times 0,7 = 7,3 - 5,18 \geq 2.$$

Et donc par le principe de récurrence, pour tout $n \in \mathbf{N}$, u_n est bien défini et $u_n \geq 2$.

6.a. La fonction g est dérivable sur $[2, +\infty[$ comme différence de fonction dérivables, et

$$g'(x) = f'(x) - 1 = e^x - \frac{e}{x} - 1 = \frac{x(e^x - 1) - e}{x}.$$

Mais pour $x \geq 2$, $x(e^x - 1) - e \geq 2(e^2 - 1) + e \geq 2(7,3 - 1) - 2,4 > 0$. On en déduit que $g'(x) > 0$, et donc que g est strictement croissante sur $[2, +\infty[$.

Comme de plus $g(2) = f(2) - 2 > 0$, on en déduit que pour tout $x \geq 2$, $g(x) \geq g(2) > 0$.

6.b. Pour tout $n \in \mathbf{N}$, on a $u_{n+1} - u_n = f(u_n) - u_n = g(u_n) > 0$.

Et donc la suite $(u_n)_{n \in \mathbf{N}}$ est croissante¹.

7. Par le théorème de la limite monotone, $(u_n)_{n \in \mathbf{N}}$ étant croissante, soit elle converge vers une limite finie ℓ , soit elle tend vers $+\infty$.

Supposons par l'absurde que $(u_n)_{n \in \mathbf{N}}$ converge vers une limite ℓ .

Alors nécessairement², $\ell \geq u_0 = 2$. Et alors on a aussi $\lim_{n \rightarrow +\infty} u_{n+1} = \ell$.

Mais par continuité de f , on a également $\lim_{n \rightarrow +\infty} u_{n+1} = \lim_{n \rightarrow +\infty} f(u_n) = f(\ell)$.

Et donc par unicité de la limite de $(u_{n+1})_{n \in \mathbf{N}}$, il vient donc $f(\ell) = \ell \Leftrightarrow g(\ell) = 0$.

Or, il a été montré à la question 6.a que g ne prend que des valeurs strictement positives sur $[2, +\infty[$, contredisant $f(\ell) = \ell$.

Ceci signifie donc que notre hypothèse est fautive, et donc (u_n) ne converge pas vers une limite finie, de sorte que $\lim_{n \rightarrow +\infty} u_n = +\infty$.

⚠ Danger !

Notons qu'il n'est pas question d'appliquer le théorème de la bijection directement à f' .

En effet, celui-ci permet de garantir l'existence et l'unicité de solutions à des équations du type $g(x) = \lambda$, où λ est une **constante**.

Mais en aucun cas on ne peut l'appliquer directement pour résoudre $g(x) = x$ ou $g(x) = x^2$.

📖 Défini ?

Puisque f n'est définie que sur \mathbf{R}_+^* , il n'est pas totalement évident que $u_{n+1} = f(u_n)$ soit toujours bien défini, puisque si pour une certaine valeur de n , $u_n \leq 0$, alors le terme suivant u_{n+1} n'est pas défini.

Il s'agit donc ici de prouver que ceci n'arrive jamais.

🔍 Détails

On a prouvé à la question précédente que $f(2) > 2$.

¹ Et même strictement croissante.

² La suite étant croissante, sa limite est plus grande que chacun de ses termes.

8. Il s'agit de calculer successivement les valeurs de u_n jusqu'à obtenir une valeur supérieure ou égale à A . Le programme suivant devrait convenir :

```

1  function n = DS(A)
2      n = 0
3      u = 2
4      while u < A
5          u = exp(u) - e*log(u) ;
6          n = n + 1 ;
7      end
8  endfunction
    
```

En effet, la boucle `while` va calculer les valeurs successives de u_n jusqu'à obtenir une valeur supérieure ou égale à A .

Lorsque c'est le cas, le programme s'arrête et renvoie la valeur de n correspondante.

- 9.a. Posons $h(x) = x - 2 \ln(x)$. Alors h est dérivable sur $[2, +\infty[$, avec $h'(x) = 1 - \frac{2}{x} = \frac{x-2}{x} \geq 0$, de sorte que h est croissante sur $[2, +\infty[$.
Et puisque $h(2) = 2 - 2 \ln(2) > 2 - 2 \times 0,7 > 0$, on en déduit que pour tout $x \geq 2$, $h(x) \geq 0 \Leftrightarrow x \geq 2 \ln(x)$.

Sur le même principe, posons $v(x) = \frac{e^x}{3} - x$, qui est dérivable avec $v'(x) = \frac{e^x}{3} - 1 = \frac{e^x - 3}{3}$.
Par croissance de l'exponentielle, pour tout $x \geq 2$, $e^x - 3 \geq e^2 - 3 \geq 7,3 - 3 \geq 0$.
Et donc pour tout $x \in [2, +\infty[$, $v'(x) \geq 0$, de sorte que v est croissante sur $[2, +\infty[$.

Puisque $v(2) = \frac{e^2}{3} - 2 = \frac{e^2 - 6}{3} > \frac{7,3 - 6}{2} \geq 0$, on en déduit que pour tout $x \geq 2$, $v(x) \geq 0 \Leftrightarrow \frac{e^x}{3} \geq x$.

- 9.b. Pour $n \in \mathbf{N}$, en utilisant les inégalités de la question précédente,

$$u_{n+1} = f(u_n) = e^{u_n} - e \ln(u_n) \geq 3u_n - e \frac{u_n}{2} = \frac{6-e}{2} u_n.$$

- 9.c. En passant à l'inverse dans l'inégalité précédente, il vient

$$\frac{1}{u_{n+1}} \leq \frac{2}{6-e} \frac{1}{u_n}.$$

Et alors, en utilisant plusieurs fois cette inégalité, il vient :

$$0 \leq \frac{1}{u_n} \leq \frac{2}{6-e} \frac{1}{u_{n-1}} \leq \left(\frac{2}{6-e}\right)^2 \frac{1}{u_{n-2}} \leq \dots$$

Plus précisément, une récurrence rapide prouve que pour tout $n \in \mathbf{N}$,

$$0 \leq \frac{1}{u_n} \leq \left(\frac{2}{6-e}\right)^n \frac{1}{u_0} = \left(\frac{2}{6-e}\right)^n \frac{1}{2}.$$

Mais nous avons $3,2 < 6 - e$ et donc $0 \leq \frac{2}{6-e} < \frac{2}{3,2} < 1$.

On en déduit que la série de terme général $\left(\frac{2}{6-e}\right)^n \frac{1}{2}$ est une série géométrique convergente³

Et donc par critère de comparaison pour les séries à termes positifs, la série de terme général $\frac{1}{u_n}$ converge.

Partie III : Étude d'intégrales généralisées.

10. La fonction f est continue sur $]0, 1]$. Mais elle n'est pas continue en 0, donc il se pourrait que l'intégrale ne converge pas en 0.

Pour $A \in]0, 1]$, on a

$$\int_A^1 f(x) dx = \int_A^1 (e^x - e \ln(x)) dx = [e^x - e(x \ln(x) - x)]_A^1$$

Méthode

Pour montrer une inégalité, il n'est pas rare d'avoir à introduire une fonction pour en étudier le signe via un tableau de variation.
Dans ce cas, on commencera par « passer du même côté » tous les termes de l'inégalité souhaitée, de manière à se ramener à une inéquation de la forme $h(x) \geq 0$ ou $h(x) \leq 0$.

Attention !

Le passage à l'inverse change le sens des inégalités car la fonction $x \mapsto \frac{1}{x}$ est décroissante sur \mathbf{R}_+ , et que toutes les quantités manipulées ici sont positives.

³ Car sa raison est dans $[0, 1]$.

Rédaction

Le fait que pour deux séries, si $v_n \leq w_n$ et si $\sum w_n$ converge alors $\sum v_n$ converge n'est valable que si $v_n \geq 0$, et donc il est important de mentionner ici la positivité de $\frac{1}{u_n}$, pour montrer que l'on connaît bien nos hypothèses, et qu'on a vérifié qu'elles s'appliquaient ici.

$$= e + e - e^A + e(A \ln(A) - A)$$

$$\xrightarrow{A \rightarrow 0^+} 2e - 1.$$

On en déduit donc que $\int_0^1 f(x) dx$ converge et que $\int_0^1 f(x) dx = 2e - 1$.

11. Nous savons déjà que $\int_0^1 f(x) dx$ converge, donc il s'agit d'étudier la convergence de $\int_1^{+\infty} f(x) dx$.

Or, au voisinage de $+\infty$, nous avons $f(x) \underset{x \rightarrow +\infty}{\sim} e^x$. Puisque $\int_1^{+\infty} e^x dx$ diverge, par critère de comparaison pour les intégrales de fonctions positives, il en est de même de $\int_1^{+\infty} f(x) dx$.

Et donc $\int_0^{+\infty} f(x) dx$ diverge.

12. La fonction $x \mapsto \frac{1}{f(x)}$ est continue sur $[2, +\infty[$, et donc le seul éventuel problème de convergence est au voisinage de $+\infty$. Or $\frac{1}{f(x)} \underset{x \rightarrow +\infty}{\sim} e^{-x}$.

Puisque $\int_2^{+\infty} e^{-x} dx$ converge, par critère de comparaison pour les fonctions positives,

$\int_2^{+\infty} \frac{dx}{f(x)}$ converge.

EXERCICE 3 (EDHEC 2018)

1. La fonction f_n est continue sur \mathbf{R}_+ , et elle y est même dérivable, avec $f'_n(x) = -1 - nx^{n-1} < 0$. Donc f_n est strictement décroissante sur \mathbf{R}_+ . Puisque d'autre part, $f_n(0) = 1$ et $\lim_{x \rightarrow +\infty} f_n(x) = -\infty$, par le théorème de la bijection, elle réalise une bijection strictement décroissante de \mathbf{R}_+ sur $]-\infty, 1]$.

Par conséquent, l'équation $f_n(x) = 0$ admet une unique solution sur \mathbf{R}_+ .

- 2.a. On a $f_n(0) = 1 > 0 = f(u_n)$, donc par décroissance de f_n , $u_n > 0$.
Et d'autre part, $f_n(1) = -1 < 0 = f_n(u_n)$, donc par décroissance de f_n , $u_n < 1$.
Ainsi, on a bien $u_n \in]0, 1[$.
- 2.b. Puisque $u_n \in]0, 1[$, on a $u_n^{n+1} < u_n^n$. Et donc

$$f_{n+1}(u_n) = 1 - u_n - u_n^{n+1} > 1 - u_n - u_n^n = f_n(u_n) = 0.$$

On a donc $f_{n+1}(u_n) > f_{n+1}(u_{n+1})$ et donc par décroissance de f_{n+1} , cela implique $u_n < u_{n+1}$.
Ainsi, la suite (u_n) est croissante⁴.

- 2.c. La suite (u_n) est croissante, et majorée⁵, et donc par le théorème de la limite monotone, elle converge.

De plus, puisque pour tout $n \in \mathbf{N}$, $0 < u_n < 1$, il vient alors $0 \leq \lim_{n \rightarrow +\infty} u_n \leq 1$.

- 2.d. Supposons par l'absurde que $\ell = \lim_{n \rightarrow +\infty} u_n < 1$.

On aurait alors, par continuité du logarithme, $\lim_{n \rightarrow +\infty} \ln(u_n) = \ln(\ell) < 0$ de sorte que $n \ln(u_n) \xrightarrow{n \rightarrow +\infty} -\infty$.

Mais $u_n = 1 - u_n^n = 1 - e^{n \ln u_n}$, et donc par passage à la limite, $\ell = 1 - 0 = 1$, ce qui est absurde.

Et donc on a bien $\lim_{n \rightarrow +\infty} u_n = 1$.

- 3.a. Puisque $u_n \in]0, 1[$, alors on a $v_n \in]0, 1[$.
D'autre part, $f_n(u_n) = 0$ s'écrit encore $v_n = u_n^n$.

Détails

Par croissances comparées

$$\lim_{A \rightarrow 0^+} A \ln(A) = 0.$$

Rédaction

Lorsqu'on utilise un équivalent pour déterminer la nature d'une intégrale, il est important de vérifier (et de mentionner) que toutes les fonctions en jeu sont positives, le critère ne fonctionne plus pour des fonctions qui ne seraient pas positives (ou au moins de signe constant).

Bornes

Le cours nous dit que $\int_0^{+\infty} e^{-x} dx$ converge, donc en réduisant le domaine d'intégration, ceci implique que $\int_2^{+\infty} e^{-x} dx$ converge également.

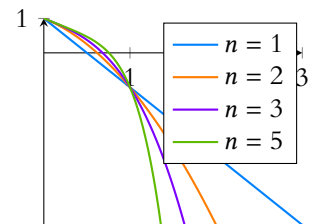


FIGURE 24.3— La fonction f_n pour quelques valeurs de n .

⁴ Et même strictement croissante.

⁵ Par 1 d'après la question 2.a.

⚠ Attention !

Le passage à la limite transforme les inégalités strictes en inégalités larges : on ne peut pas affirmer que

$$0 < \lim_{n \rightarrow +\infty} u_n < 1.$$

Soit $\ln(v_n) = n \ln(u_n) = n \ln(1 + (u_n - 1)) = n \ln(1 - v_n)$.

Puisque $v_n \xrightarrow{n \rightarrow +\infty} 0$, l'équivalent classique $\ln(1+x) \underset{x \rightarrow 0}{\sim} x$ nous donne alors $\ln(1 - v_n) \underset{n \rightarrow +\infty}{\sim} -v_n$.

Et donc $\boxed{\ln(v_n) \underset{n \rightarrow +\infty}{\sim} -nv_n}$.

3.b. L'équivalent de la question précédente nous indique que $\frac{\ln v_n}{-nv_n} \xrightarrow{n \rightarrow +\infty} 1$.

Et donc $\lim_{n \rightarrow +\infty} \ln\left(-\frac{\ln v_n}{nv_n}\right) = 0$.

D'autre part, comme $u_n \xrightarrow{n \rightarrow +\infty} 1$, on a $v_n \xrightarrow{n \rightarrow +\infty} 0$, et donc $\lim_{n \rightarrow +\infty} \ln v_n = -\infty$.

Par quotient de limites, on a donc

$$\boxed{\lim_{n \rightarrow +\infty} \frac{\ln\left(-\frac{\ln v_n}{nv_n}\right)}{\ln v_n} = 0.}$$

Ce qui s'écrit aussi sous la forme

$$\lim_{n \rightarrow +\infty} \frac{\ln(-\ln v_n) - \ln(n) - \ln(v_n)}{\ln v_n} = 0 \Leftrightarrow \lim_{n \rightarrow +\infty} \frac{\ln(-\ln v_n) - \ln n}{\ln v_n} = 1.$$

Mais puisque $-\ln(v_n) \xrightarrow{n \rightarrow +\infty} +\infty$, et que $\lim_{x \rightarrow +\infty} \frac{\ln(x)}{x} = 0$, il vient donc par composition de limites

$$\lim_{n \rightarrow +\infty} \frac{\ln(-\ln v_n)}{-\ln(v_n)} = 0.$$

Et donc

$$\lim_{n \rightarrow +\infty} \frac{-\ln n}{\ln v_n} = \lim_{n \rightarrow +\infty} \frac{\ln(-\ln v_n) - \ln(n)}{\ln v_n} - \lim_{n \rightarrow +\infty} \frac{\ln(-\ln v_n)}{\ln v_n} = 1.$$

Et donc $\boxed{\ln v_n \underset{n \rightarrow +\infty}{\sim} -\ln n}$.

3.c. Grâce aux équivalents des deux questions précédentes, on a $-nv_n \underset{n \rightarrow +\infty}{\sim} -\ln n$.

Et par conséquent, $\boxed{v_n \underset{n \rightarrow +\infty}{\sim} \frac{\ln n}{n}}$.

4. Pour $n \geq 3$, on a $0 \leq \frac{1}{n} \leq \frac{\ln n}{n}$. Puisque la série de terme général $\frac{1}{n}$ diverge, il en est donc de même de la série de terme général $\frac{\ln n}{n}$.

Et donc par le critère des équivalents pour les séries à termes **positifs**, on en déduit que la série de terme général v_n diverge.

D'autre part, on a $n^{3/2}v_n^2 \underset{n \rightarrow +\infty}{\sim} \frac{(\ln n)^2}{\sqrt{n}} \xrightarrow{n \rightarrow +\infty} 0$ par croissances comparées.

Et donc $v_n^2 = \underset{n \rightarrow +\infty}{o}\left(\frac{1}{n^{3/2}}\right)$.

Puisque la série⁶ de terme général $\frac{1}{n^{3/2}}$ converge, il en est de même de la série de terme général v_n .

⁶ De Riemann.

PROBLÈME (ÉCRICOME ECE 2018)

Partie I

1. Puisque les lancers de la pièce sont indépendants, nous reconnaissons bien évidemment une loi binomiale $\mathcal{B}\left(3, \frac{2}{3}\right)$.

Et on a alors

$$P(A) = P(X = 0) + P(X = 2) = \binom{3}{0} \left(\frac{2}{3}\right)^0 \left(\frac{1}{3}\right)^3 + \binom{3}{2} \left(\frac{2}{3}\right)^2 \left(\frac{1}{3}\right)^1 = \frac{1}{27} + 3 \frac{4}{27} = \boxed{\frac{13}{27}}.$$

Remarque

Cette majoration n'est pas valable pour $n = 1$ ou $n = 2$, mais comme les premiers termes ne changent rien à la nature de la série, le raisonnement qui suit reste valable.

Détails

X prend des valeurs entre 0 et 3 et donc ne prend une valeur paire que si elle vaut 0 ou 2.

2. Si $X = 0$, alors $G = 0$. Si $X = 2$, alors le joueur gagne 20 euros, et donc $G = 20$.
Et dans le cas où le joueur perd, soit $X = 1$, et donc $G = -10$, soit $X = 3$ et donc $G = -30$.
Puisqu'on a épuisé tous les cas possibles, on a donc bien $G(\Omega) = \{-30, -10, 0, 20\}$.
Et alors la loi de G est donnée par

$$P(G = -30) = P(X = 3) = \frac{8}{27}, P(G = -10) = P(X = 1) = \frac{6}{27}, P(G = 0) = P(X = 0) = \frac{1}{27}, P(G = 20) = P(X = 2) = \frac{12}{27}.$$

3. On a alors

$$E(G) = -30P(G = -30) - 10P(G = -10) + 0 \times P(G = 0) + 20P(G = 20) = -\frac{60}{27}.$$

Puisque cette espérance est négative, le jeu n'est pas favorable au joueur.

Partie II

- 4.a. La variable aléatoire Y ne peut prendre que les valeurs -1 et 1 , donc $Y(\Omega) = \{-1, 1\}$.
Et donc Z ne peut prendre que les valeurs $\frac{-1+1}{2} = 0$ et $\frac{1+1}{2} = 1$, de sorte que Y suit une loi de Bernoulli.
Son paramètre est alors $P(Z = 1) = P(Y = 1) = P(\text{«}X \text{ est pair}\text{»}) = P(A)$, puisque Y vaut 1 si et seulement si X est pair.

- 4.b. On a alors $E(Z) = P(A)$. Mais d'autre part, par linéarité de l'espérance,

$$E(Z) = \frac{E(Y) + 1}{2} \Leftrightarrow E(Y) = 2E(Z) - 1 = 2P(A) - 1.$$

- 5.a. Sur le même principe qu'à la question 1, X compte le nombre de succès⁷ lors d'une répétition de n épreuves de Bernoulli indépendantes, et donc $X \hookrightarrow \mathcal{B}(n, p)$.
- 5.b. Par le théorème de transfert, puisque $Y = (-1)^X$, il vient

$$E(Y) = \sum_{k=0}^n (-1)^k P(X = k) = \sum_{k=0}^n (-1)^k \binom{n}{k} p^k (1-p)^{n-k}.$$

Ceci s'écrit encore

$$E(Y) = \sum_{k=0}^n \binom{n}{k} (-p)^k (1-p)^{n-k} = (-p + 1 - p)^n = (1 - 2p)^n.$$

- 5.c. En utilisant les résultats de 4.b et 5.b, on a donc

$$2P(A) - 1 = (1 - 2p)^n \Leftrightarrow P(A) = \frac{(1 - 2p)^n + 1}{2}.$$

- 5.d. Il vient donc $P(A) \geq \frac{1}{2} \Leftrightarrow (1 - 2p)^n \geq 0$.

Si $1 - 2p \geq 0 \Leftrightarrow p \leq \frac{1}{2}$, alors pour tout $n \in \mathbf{N}^*$, $(1 - 2p)^n \geq 0$.

En revanche, si $1 - 2p < 0 \Leftrightarrow p > \frac{1}{2}$, alors $(1 - 2p)^n$ est positif si et seulement si n est pair.

Et donc on a bien $P(A) \geq 0 \Leftrightarrow \left[p \leq \frac{1}{2} \text{ ou } n \text{ pair} \right]$.

Partie III

6. La somme qui change de main⁸ est égale à $10X$.
Or, c'est le joueur qui la reçoit si X est pair, c'est-à-dire si $Y = 1$, et c'est le joueur qui paye si X est impair, c'est-à-dire si $Y = -1$. Autrement dit, $G = 10 \times X \times Y$.
Soit encore $G = 10(-1)^X X$. Et alors par le théorème de transfert,

$$E(G) = \sum_{k=0}^n 10(-1)^k k P(X = k).$$

Astuce

Toute variable aléatoire qui ne prend que les valeurs 0 et 1 suit une loi de Bernoulli. Le paramètre en est alors la probabilité qu'elle prenne la valeur 1.

Convergence

Notons que la somme qui apparaît ici est finie, et donc automatiquement convergente. Par le théorème de transfert, cela garantit l'existence de $E(X)$.
Plus simplement : Y ne peut prendre que deux valeurs (-1 et 1), et donc, comme toute variable aléatoire finie, admet une espérance.

Parité

Si $q < 0$, alors $q^n > 0$ si n est pair et $q^n < 0$ si n est impair.

⁸ Qui n'est autre que $|G|$.

7. On a donc

$$\begin{aligned}
 E(G) &= 10 \sum_{k=0}^n (-1)^k k \binom{n}{k} p^k (1-p)^{n-k} = 10 \sum_{k=0}^n (-1)^k k \binom{n}{k} p^k (1-p)^{n-k} \\
 &= 10 \sum_{k=1}^n (-1)^k n \binom{n-1}{k-1} p^k (1-p)^{n-k} \\
 &= 10n \sum_{i=0}^{n-1} \binom{n-1}{i} (-p)^{i+1} (1-p)^{n-1-i} \quad (i = k-1) \\
 &= -10np \sum_{i=0}^{n-1} \binom{n-1}{i} (-p)^i (1-p)^{n-1-i} \\
 &= -10np(-p+1-p)^{n-1} = \boxed{-10np(1-2p)^{n-1}}.
 \end{aligned}$$

Le terme correspondant à $k = 0$ est nul.

Rappel

On a toujours

$$k \binom{n}{k} = n \binom{n-1}{k-1}.$$

Bien que très classique, cette formule n'est pas nécessairement à connaître par cœur, il suffit de savoir qu'une telle formule existe et d'être capable de la retrouver en revenant à la définition de coefficients binomiaux à l'aide de factorielles.

8. Nous savons déjà que $P(A) \geq 0$ si et seulement si $p \leq \frac{1}{2}$ ou n est pair.

Si $p > \frac{1}{2}$ et n est pair, alors $n-1$ est impair, et donc $(1-2p)^{n-1} < 0$, de sorte que $E(G) > 0$.

Donc si $P(A) \geq \frac{1}{2}$ et $E(G) \leq 0$, c'est nécessairement que $p \leq \frac{1}{2}$.

Inversement, si $p \leq \frac{1}{2}$, alors nous savons déjà que $P(A) \geq \frac{1}{2}$, et $E(G) = -10np(1-2p)^{n-1} \leq 0$.

Et donc nous avons bien prouvé que

$$\boxed{\begin{cases} P(A) \geq \frac{1}{2} \\ E(G) \leq 0 \end{cases} \Leftrightarrow p \leq \frac{1}{2}.$$

9.a. La fonction f est une fonction polynomiale, qui est donc dérivable, et on a

$$f'(x) = (1-2x)^{n-1} - 2(n-1)x(1-2x)^{n-2} = (1-2x)^{n-2} (1-2x - 2(n-1)x) = (1-2x)^{n-2}(1-2nx).$$

Puisque sur $\left[0, \frac{1}{2}\right]$, $(1-2x)^{n-2} \geq 0$, $f'(x)$ est donc du signe de $1-2nx$.

Et donc $f'(x) \geq 0 \Leftrightarrow 1-2nx \geq 0 \Leftrightarrow x \leq \frac{1}{2n}$.

Donc f est croissante sur $\left[0, \frac{1}{2n}\right]$ et décroissante sur $\left[\frac{1}{2n}, 1\right]$.

Elle admet donc un maximum en $\frac{1}{2n}$.

9.b. Optimiser la rentabilité du jeu pour le forain, c'est minimiser la valeur de $E(G)$.

Mais, $E(G) = -10np(1-2p)^{n-1} = -10nf(p)$.

Puisque la fonction f admet un maximum en $\frac{1}{2n}$, $E(G)$ admet un minimum pour $p = \frac{1}{2n}$.

Et donc le forain a intérêt à choisir une pièce qui tombe sur Pile avec une probabilité égale à $\frac{1}{2n}$.

Remarques : notons qu'on aura alors toujours $P(A) \geq \frac{1}{2}$, de sorte que le forain peut bien annoncer qu'il y a plus de gagnants que de perdants.

Toutefois, un joueur qui ne gagne rien (c'est-à-dire qui obtient 0 fois Pile) est considéré comme gagnant, ce qui est largement discutable.

Notons enfin que pour de grandes valeurs de n , $p = \frac{1}{2n}$ est vraiment petit, c'est-à-dire que la probabilité de faire Pile est faible, ce qui pourrait rebuter d'éventuels joueurs.

Détails

Plus $E(G)$ est petit, plus le joueur perd d'argent en moyenne, et donc plus le forain gagne d'argent en moyenne.

DEVOIR SURVEILLÉ 2 (29.09.18)

EXERCICE 1

Soit $n \in \mathbf{N}^*$, on note E l'ensemble des fonctions $f : \mathbf{R}_+^* \rightarrow \mathbf{R}$ telles qu'il existe deux polynômes P, Q appartenant à $\mathbf{R}_{n-1}[X]$ avec :

$$\forall x \in \mathbf{R}_+^*, f(x) = xP(x) + x \ln(x)Q(x).$$

Pour tout entier $k \in \{1, \dots, n\}$, on pose :

$$u_k : \begin{cases} \mathbf{R}_+^* & \longrightarrow \mathbf{R} \\ x & \longmapsto x^k \end{cases} \quad \text{et} \quad v_k : \begin{cases} \mathbf{R}_+^* & \longrightarrow \mathbf{R} \\ x & \longmapsto x^k \ln(x) \end{cases}$$

Pour toute fonction f appartenant à E , on note $\varphi(f)$ la fonction définie sur \mathbf{R}_+^* par :

$$\forall x \in \mathbf{R}_+^*, \varphi(f)(x) = \frac{1}{x} \int_0^x f(t) dt$$

et on note φ l'application qui à $f \in E$ associe $\varphi(f)$.

1. Prouver que E est un \mathbf{R} -espace vectoriel et que $E = \text{Vect}(u_1, v_1, \dots, u_n, v_n)$ (c'est-à-dire que E est l'espace vectoriel engendré par les fonctions $u_1, v_1, \dots, u_n, v_n$).

On admettra que la famille $\mathcal{B} = (u_1, v_1, \dots, u_n, v_n)$ est une base de E .

2. Justifier que chaque fonction f de E se prolonge en une fonction continue sur \mathbf{R}_+ et, pour tout $k \in \{1, \dots, n\}$, calculer $\varphi(u_k)$ et $\varphi(v_k)$.
Indication : pour calculer $\varphi(v_k)(x)$, on pourra commencer par procéder à une intégration par parties sur un segment de la forme $[A, x]$, avec $0 < A < x$.
3. Démontrer que φ est linéaire. En déduire que $\varphi(f) \in E$ lorsque $f \in E$.
4. Écrire la matrice de φ dans la base \mathcal{B} . L'endomorphisme φ est-il bijectif ?
5. Dans cette question, on considère un entier $k \in \llbracket 1, n \rrbracket$, et on note $\mathcal{E}_k = \left\{ f \in E : \varphi(f) = \frac{1}{k+1} f \right\}$.
 - a. Montrer que \mathcal{E}_k est un sous-espace vectoriel de E , qui contient u_k .
 - b. Soit f un élément non nul de \mathcal{E}_k , et soit g la fonction définie sur \mathbf{R}_+^* par

$$\forall x \in \mathbf{R}_+^*, g(x) = \frac{1}{x^{k+1}} \int_0^x f(t) dt.$$

Montrer que g est constante sur \mathbf{R}_+^* . En déduire l'expression de la fonction $x \mapsto \int_0^x f(t) dt$, puis celle de f .

- c. Montrer que $\mathcal{E}_k = \text{Vect}(u_k)$.

EXERCICE 2

Un ascenseur dessert n étages. Des personnes montent l'une après l'autre dans cet ascenseur pendant qu'il est au rez-de-chaussée.

On suppose :

- que personne ne monte dans l'ascenseur aux étages
- qu'au moment où elle rentre dans l'ascenseur, chaque personne choisit son étage au hasard parmi les n étages en pressant l'un des n boutons.
- que les choix des différentes personnes sont indépendants les uns des autres.

Pour tout $k \in \mathbf{N}$. On note S_k la variable aléatoire égale au nombre d'étages choisis par les k premières personnes, c'est-à-dire au nombre de boutons qui sont allumés avant que la $(k+1)^{\text{ème}}$ personne ne monte dans l'ascenseur. Remarquons que $S_0 = 0$.

1. Montrer que pour $k \in \mathbf{N}$ et $1 \leq j \leq n$, on a

$$P(S_{k+1} = j) = \frac{j}{n} P(S_k = j) + \frac{n-j+1}{n} P(S_k = j-1)$$

2. Pour $k \in \mathbf{N}$, justifier l'existence de $E(S_k)$, puis prouver que

$$\forall k \in \mathbf{N}, E(S_{k+1}) = 1 + \left(1 - \frac{1}{n}\right) E(S_k).$$

3. Déterminer la valeur de $E(S_0)$. Puis, en remarquant que la suite $(E(S_k))_{k \in \mathbf{N}}$ est arithmético-géométrique, déterminer la valeur de $E(S_k)$, pour tout $k \in \mathbf{N}$.
4. On suppose à présent que le nombre de personnes qui montent dans l'ascenseur au rez-de-chaussée est une variable aléatoire X qui suit une loi de Poisson de paramètre λ , et on note S la variable aléatoire égale au nombre d'étages où l'ascenseur devra s'arrêter.
- Montrer que S admet une espérance.
 - Justifier que pour tout $k \in \mathbf{N}$, pour tout $j \in \llbracket 0, n \rrbracket$, $P_{[X=k]}(S = j) = P(S_k = j)$.
En déduire l'expression de $E(S|X = k)$.
 - Montrer que $E(S) = n \left(1 - e^{-\frac{\lambda}{n}}\right)$.

EXERCICE 3

Soit p un entier naturel et f une fonction continue, strictement positive, décroissante sur $[p, +\infty[$ et telle que $\int_p^{+\infty} f(t) dt$ converge.

Pour tout entier naturel n supérieur ou égal à p , on pose $S_n = \sum_{k=p}^n f(k)$.

1. a. Utiliser la décroissance de f pour montrer que, pour tout entier naturel k supérieur ou égal à p , on a :

$$f(k+1) \leq \int_k^{k+1} f(t) dt.$$

$$\text{En déduire que pour } n \geq p : S_n - f(p) \leq \int_p^n f(t) dt.$$

- b. En déduire que la série de terme général $f(n)$ est convergente.

On pose désormais, pour tout entier naturel n supérieur ou égal à p , $R_n = \sum_{k=n+1}^{+\infty} f(k)$.

2. a. Montrer que, pour tout entier naturel n supérieur ou égal à p , on a :

$$\int_n^{+\infty} f(t) dt - f(n) \leq R_n \leq \int_n^{+\infty} f(t) dt.$$

- b. En déduire une condition suffisante portant sur $f(n)$ et $\int_n^{+\infty} f(t) dt$ pour que : $R_n \underset{n \rightarrow +\infty}{\sim} \int_n^{+\infty} f(t) dt$.

3. Dans cette question, pour tout réel x de $[2, +\infty[$, on pose $f(x) = \frac{1}{x(\ln x)^2}$.

- a. Montrer que cette fonction vérifie les quatre hypothèses de l'énoncé ainsi que la condition trouvée à la question 2.b.

- b. En déduire un équivalent, lorsque n est au voisinage de $+\infty$, de $R_n = \sum_{k=n+1}^{+\infty} \frac{1}{k(\ln k)^2}$.

- c. La série de terme général R_n est-elle convergente ?

PROBLÈME

On admettra dans la suite que $\sum_{n=1}^{+\infty} \frac{1}{n^2} = \frac{\pi^2}{6}$.

1. a. Montrer que, pour tout couple $(x, y) \in (\mathbf{R}_+^*)^2$, la série $\sum_{n \geq 1} \frac{1}{(n+x)(n+y)}$ et la série $\sum_{n \geq 1} \frac{1}{(n+x)^2(n+y)}$ convergent.

- b. Montrer que, pour tout $x \in [0, +\infty[$, la série $\sum_{n \geq 1} \left(\frac{1}{n} - \frac{1}{n+x}\right)$ converge.

On note S l'application définie, pour tout $x \in [0, +\infty[$ par $S(x) = \sum_{n=1}^{+\infty} \left(\frac{1}{n} - \frac{1}{n+x}\right)$.

2. Calculer $S(0)$ et $S(1)$.

3. a. Établir, pour tous $(x, y) \in (\mathbf{R}_+)^2$, $S(y) - S(x) = (y - x) \sum_{n=1}^{+\infty} \frac{1}{(n+x)(n+y)}$.

b. En déduire : $\forall (x, y) \in (\mathbf{R}_+)^2$, $|S(x) - S(y)| \leq \frac{\pi^2}{6} |y - x|$.

c. Montrer alors que la fonction S est continue sur $[0, +\infty[$.

4. a. Montrer, pour tout couple $(x, y) \in (\mathbf{R}_+)^2$ tel que $x \neq y$:

$$\left| \frac{S(y) - S(x)}{y - x} - \sum_{n=1}^{+\infty} \frac{1}{(n+x)^2} \right| \leq |y - x| \sum_{n=1}^{+\infty} \frac{1}{n^3}.$$

b. En déduire que la fonction S est dérivable sur $[0, +\infty[$ et que :

$$\forall x \in [0, +\infty[, \quad S'(x) = \sum_{n=1}^{+\infty} \frac{1}{(n+x)^2}.$$

c. Préciser les valeurs de $S'(0)$ et de $S'(1)$.

5. On admet que S est \mathcal{C}^1 . Montrer que S est concave.

6. Soit $x \in]0, +\infty[$ fixé. On note φ la fonction définie sur $[1, +\infty[$ par

$$\forall t \in [1, +\infty[, \varphi(t) = \frac{1}{t} - \frac{1}{t+x}.$$

a. Montrer que l'intégrale $\int_1^{+\infty} \varphi(t) dt$ converge et calculer sa valeur.

b. Montrer : $\forall n \in \mathbf{N}^*$, $\varphi(n+1) \leq \int_n^{n+1} \varphi(t) dt \leq \varphi(n)$, et en déduire :

$$\int_1^{+\infty} \varphi(t) dt \leq S(x) \leq 1 + \int_1^{+\infty} \varphi(t) dt.$$

c. Conclure : $S(x) \underset{x \rightarrow +\infty}{\sim} \ln(x+1)$.

En déduire que $S(x) \underset{x \rightarrow +\infty}{\sim} \ln(x)$.

7. a. Dresser le tableau de variations de S , en précisant la limite de S en $+\infty$.

b. Tracer l'allure de la courbe représentative de S .

CORRECTION DU DEVOIR SURVEILLÉ 2

EXERCICE 1 : EXTRAIT DE ECRICOME 2014

1. Nous allons prouver que E est un sous-espace vectoriel de l'ensemble des fonctions continues sur \mathbf{R}_+^* .

E est clairement formé de fonctions continues, donc inclus dans $\mathcal{C}^0(\mathbf{R}_+^*, \mathbf{R})$, et la fonction nulle est dans E : elle correspond à $P = Q = 0$.

Soient f, g deux éléments de E et $\lambda \in \mathbf{R}$. Il existe donc quatre polynômes P_1, P_2, Q_1, Q_2 de $\mathbf{R}_{n-1}[X]$ tels que pour tout $x \in \mathbf{R}_+^*$,

$$f(x) = xP_1(x) + xQ_1(x) \ln(x) \text{ et } g(x) = xP_2(x) + xQ_2(x) \ln(x).$$

Et donc, pour tout $x \in \mathbf{R}_+^*$, on a

$$(\lambda f + g)(x) = x(\lambda P_1(x) + P_2(x)) + x(\lambda Q_1(x) + Q_2(x)) \ln(x).$$

Puisque $\lambda P_1 + P_2$ et $\lambda Q_1 + Q_2$ sont encore dans $\mathbf{R}_{n-1}[X]$, on a bien $\lambda f + g \in E$.

Ainsi, E est un sous-espace vectoriel de $\mathcal{C}^0(\mathbf{R}_+^*, \mathbf{R})$, et donc est un espace vectoriel.

Si $f \in E$, soient $P = \sum_{k=0}^{n-1} a_k X^k$ et $Q = \sum_{k=0}^{n-1} b_k X^k$ deux polynômes de $\mathbf{R}_{n-1}[X]$ tels que

$$\forall x \in \mathbf{R}_+^*, f(x) = xP(x) + xQ(x) \ln(x).$$

Alors, pour $x \in \mathbf{R}_+^*$, on a

$$f(x) = x \left(\sum_{k=0}^{n-1} a_k x^k \right) + x \left(\sum_{k=0}^{n-1} b_k x^k \right) \ln(x) = \sum_{k=0}^{n-1} a_k u_{k+1}(x) + \sum_{k=0}^{n-1} b_k v_{k+1}(x).$$

Donc $f = \sum_{k=0}^{n-1} a_k u_{k+1} + \sum_{k=0}^{n-1} b_k v_{k+1} \in \text{Vect}(u_1, v_1, \dots, u_n, v_n)$.

Et donc $E \subset \text{Vect}(u_1, v_1, \dots, u_n, v_n)$.

Inversement, pour tout $k \in \llbracket 1, n \rrbracket$, il est clair que $u_k \in E$ et $v_k \in E$, de sorte que $\text{Vect}(u_1, v_1, \dots, u_n, v_n) \subset E$.

On a alors¹, $E = \text{Vect}(u_1, v_1, \dots, u_n, v_n)$.

2. Soit $f = \sum_{k=1}^n a_k u_k + \sum_{k=1}^n b_k v_k \in E$.

Pour tout $k \in \llbracket 1, n \rrbracket$, on a $\lim_{x \rightarrow 0^+} u_k(x) = 0$ et par croissances comparées, $\lim_{x \rightarrow 0^+} v_k(x) = 0$.

Ainsi, $\lim_{x \rightarrow 0^+} f(x) = 0$, et donc f se prolonge par continuité en 0.

Puisque f est continue sur \mathbf{R}_+^* en tant que combinaison linéaire de fonctions continues, f est donc prolongeable par continuité en une fonction continue sur \mathbf{R}_+ .

Pour $k \in \llbracket 1, n \rrbracket$ et $x \in \mathbf{R}_+^*$, on a

$$\varphi(u_k)(x) = \frac{1}{x} \int_0^x t^k dt = \frac{1}{x} \left[\frac{t^{k+1}}{k+1} \right]_0^x = \frac{x^k}{k+1} = \frac{1}{k+1} u_k(x).$$

De même, pour $A \in]0, x[$, par une intégration par parties, en posant $u(t) = \ln(t)$ et $v(t) = \frac{t^{k+1}}{k+1}$, qui sont de classe \mathcal{C}^1 sur $[A, x]$, avec $u'(t) = \frac{1}{t}$ et $v'(t) = t^k$, il vient

$$\begin{aligned} \int_A^x v_k(t) dt &= \int_A^x t^k \ln(t) dt = \left[\frac{t^{k+1}}{k+1} \ln(t) \right]_A^x - \int_A^x \frac{t^{k+1}}{k+1} \frac{1}{t} dt \\ &= \frac{x^{k+1}}{k+1} \ln(x) - \frac{A^{k+1}}{k+1} \ln(A) - \int_A^x \frac{t^k}{k+1} dt \\ &= \frac{x^{k+1}}{k+1} \ln(x) - \frac{A^{k+1}}{k+1} \ln(A) - \frac{x^{k+1}}{(k+1)^2} + \frac{A^{k+1}}{(k+1)^2}. \end{aligned}$$

Méthode

A priori, pour prouver qu'un ensemble E est un espace vectoriel, il faudrait vérifier les huit propriétés d'un espace vectoriel (les connaissez-vous les huit ?).

Ceci est très fastidieux, et en pratique, dans les sujets de concours, il est toujours possible de s'en passer en montrant que E est un sous-espace vectoriel d'un espace vectoriel de référence.

Ici E est inclus dans l'ensemble des fonctions continues sur \mathbf{R}_+^* , dont nous savons déjà que c'est un espace vectoriel. Nous nous contentons donc de prouver que E en est un sous-espace vectoriel, ce qui ne nécessite de vérifier que trois propriétés.

Vect

Rappelons que

$$\text{Vect}(u_1, v_1, \dots, u_n, v_n)$$

est l'ensemble des combinaisons linéaires de $u_1, v_1, \dots, u_n, v_n$.

¹ Par double inclusion.

Remarque

Notons que ceci justifie la convergence de l'intégrale définissant φ : il s'agit d'une intégrale faussement impropre, donc convergente.

Bornes

On ne sait procéder à une intégration par parties que sur un **segment**. Or la fonction $t \mapsto t^k \ln(t)$ n'est définie que sur $]0, x[$, qui n'est pas un segment (intervalle ouvert en 0). Et donc on ne peut faire une IPP directement sur l'intervalle $]0, x[$: on commence donc par la faire sur un segment $[A, x]$, puis on fait tendre A vers 0.

Lorsque $A \rightarrow 0^+$, il vient, par croissances comparées,

$$\int_0^x v_k(t) dt = \lim_{A \rightarrow 0^+} \int_A^x v_k(t) dt = \frac{x^{k+1}}{k+1} \ln(x) - \frac{x^{k+1}}{(k+1)^2}.$$

On en déduit que

$$\varphi(v_k)(x) = \frac{1}{x} \left(\frac{x^{k+1}}{k+1} \ln(x) - \frac{x^{k+1}}{(k+1)^2} \right) = \frac{1}{k+1} v_k(x) - \frac{1}{(k+1)^2} u_k(x).$$

Et donc $\varphi(u_k) = \frac{1}{k+1} u_k$ et $\varphi(v_k) = -\frac{1}{(k+1)^2} u_k + \frac{1}{k+1} v_k$.

3. Si $f, g \in E$, et $\lambda \in \mathbf{R}$, alors $\forall x \in \mathbf{R}_+^*$,

$$\varphi(\lambda f + g)(x) = \frac{1}{x} \int_0^x (\lambda f(t) + g(t)) dt = \lambda \frac{1}{x} \int_0^x f(t) dt + \frac{1}{x} \int_0^x g(t) dt = \lambda \varphi(f)(x) + \varphi(g)(x).$$

On a donc $\varphi(\lambda f + g) = \lambda \varphi(f) + \varphi(g)$: φ est linéaire.

Si $f \in E$, alors il existe $\lambda_1, \dots, \lambda_n$ et μ_1, \dots, μ_n des réels tels que $f = \sum_{i=1}^n (\lambda_i u_i + \mu_i v_i)$.

Par linéarité de φ , on a alors

$$\begin{aligned} \varphi(f) &= \sum_{i=1}^n (\lambda_i \varphi(u_i) + \mu_i \varphi(v_i)) \\ &= \sum_{i=1}^n \lambda_i \frac{1}{i+1} u_i + \sum_{i=1}^n \mu_i \left(\frac{1}{i+1} v_i - \frac{1}{(i+1)^2} u_i \right) \\ &= \sum_{i=1}^n \left(\frac{\lambda_i}{i+1} - \frac{\mu_i}{(i+1)^2} \right) u_i + \sum_{i=1}^n \frac{\mu_i}{i+1} v_i \in \text{Vect}(u_1, v_1, \dots, u_n, v_n) = E. \end{aligned}$$

4. D'après les calculs effectués à la question 2, on a

$$\text{Mat}_{\mathcal{B}}(\varphi) = \begin{pmatrix} \varphi(u_1) & \varphi(v_1) & \varphi(u_2) & \varphi(v_2) & \dots & \dots & \varphi(u_n) & \varphi(v_n) \\ \frac{1}{2} & -\frac{1}{4} & 0 & 0 & \dots & \dots & 0 & 0 \\ 0 & \frac{1}{2} & 0 & 0 & \dots & \dots & 0 & 0 \\ 0 & 0 & \frac{1}{3} & -\frac{1}{9} & & & \vdots & \vdots \\ 0 & 0 & 0 & \frac{1}{3} & \ddots & & \vdots & \vdots \\ \vdots & \vdots & & \ddots & \ddots & \ddots & \vdots & \vdots \\ \vdots & \vdots & & & \ddots & \ddots & 0 & 0 \\ \vdots & \vdots & & & & \ddots & \frac{1}{n+1} & -\frac{1}{(n+1)^2} \\ 0 & 0 & \dots & \dots & \dots & \dots & 0 & \frac{1}{n+1} \end{pmatrix} \begin{matrix} u_1 \\ v_1 \\ u_2 \\ v_2 \\ \dots \\ u_n \\ v_n \end{matrix}$$

cette matrice est triangulaire supérieure, et tous ses coefficients diagonaux sont non nuls², donc elle est inversible, et par conséquent, φ est bijectif.

5.a. Par définition, $\mathcal{E}_k \subset E$, et puisque $\varphi(0_E) = 0_E = \frac{1}{k+1} 0_E$, on a bien $0_E \in \mathcal{E}_k$. Soient $f, g \in \mathcal{E}_k$ et soit $\lambda \in \mathbf{R}$. Alors

$$\varphi(\lambda f + g) = \lambda \varphi(f) + \varphi(g) = \lambda \frac{1}{k+1} f + \frac{1}{k+1} g = \frac{1}{k+1} (\lambda f + g).$$

Et donc ceci prouve que $\lambda f + g \in \mathcal{E}_k$, de sorte que \mathcal{E}_k est un sous-espace vectoriel de E .

Nous avons prouvé à la question 4 que $\varphi(u_k) = \frac{1}{k+1} u_k$, et donc $u_k \in \mathcal{E}_k$.

Égalité

Deux fonctions sont égales sur \mathbf{R}_+^* si et seulement si elles prennent la même valeur en tout $x \in \mathbf{R}_+^*$. Et puisque pour tout $x \in \mathbf{R}_+^*$, on a

$$\begin{aligned} \varphi(\lambda f + g)(x) &= \lambda \varphi(f)(x) + \varphi(g)(x) \end{aligned}$$

on en déduit que

$$\varphi(\lambda f + g) = \lambda \varphi(f) + \varphi(g).$$

Endomorphisme

Ceci prouve que φ , qui est définie sur E prend également ses valeurs dans E (a priori, on savait uniquement que φ prenait ses valeurs dans l'ensemble des fonctions définies sur \mathbf{R}_+^*). Et comme de plus φ est linéaire, ceci prouve que c'est un endomorphisme de E .

² Autrement dit elle est échelonnée avec $2n$ pivots, donc de rang $2n$ et donc inversible.

- 5.b. Puisque f est continue, par le théorème fondamental de l'analyse, la fonction $x \mapsto \int_0^x f(t) dt$ est dérivable sur \mathbf{R}_+^* , de dérivée égale à f .
Donc g est dérivable sur \mathbf{R}_+^* car produit de fonctions dérivables, et on a

$$\begin{aligned} \forall x \in \mathbf{R}_+^*, g'(x) &= \frac{1}{x^{k+1}} f(x) - \frac{k+1}{x^{k+2}} \int_0^x f(t) dt \\ &= \frac{1}{x^{k+1}} f(x) - \frac{k+1}{x^{k+1}} [\varphi(f)](x) = \frac{1}{x^{k+1}} f(x) - \frac{k+1}{x^{k+1}} \frac{1}{k+1} f(x) = 0. \end{aligned}$$

Et donc g' est nulle, de sorte que g est une fonction constante.

Notons λ la valeur de cette constante. Alors, pour tout $x \in \mathbf{R}_+^*$, $\int_0^x f(t) dt = \lambda x^{k+1}$.

Et donc en dérivant cette relation, il vient $f(x) = (k+1)\lambda x^k$.

- 5.c. Nous venons de prouver qu'un élément non nul de \mathcal{E}_k est nécessairement de la forme $(k+1)\lambda u_k$, où λ est un réel. Et donc $\mathcal{E}_k \subset \text{Vect}(u_k)$.
Puisque nous savons déjà que $u_k \in \mathcal{E}$, alors $\text{Vect}(u_k) \subset \mathcal{E}_k$.
Et donc par double inclusion, $\mathcal{E}_k = \text{Vect}(u_k)$.

Pour les cubes
Notons que \mathcal{E} n'est rien d'autre que le sous-espace propre de φ associé à la valeur propre $\frac{1}{k+1}$, valeur propre qui se lisait sur la diagonale de la matrice de φ . Nous venons donc de prouver que tous les sous-espaces propres de φ sont de dimension 1.

EXERCICE 2 (ADAPTÉ DE EDHEC 1995)

1. Appliquons la formule des probabilités totales avec le système complet d'événements $\{[S_k = i], i \in \llbracket 0, n \rrbracket\}$.

$$P(S_{k+1} = j) = \sum_{i=0}^n P(S_k = i) P_{[S_k=i]}(S_{k+1} = j).$$

Puisque la $(k+1)$ ème personne a soit choisi un étage déjà choisi par les précédentes, soit a ajouté un nouvel étage au nombre d'étages déjà choisis, pour $i \notin \{j, j+1\}$,

$$P_{[S_k=i]}(S_{k+1} = j) = 0.$$

Et donc il ne reste que

$$P(S_{k+1} = j) = P(S_k = j-1) P_{[S_k=j-1]}(S_{k+1} = j) + P(S_k = j) P_{[S_k=j]}(S_{k+1} = j).$$

Mais $P_{[S_k=j-1]}(S_{k+1} = j)$ est la probabilité que la $(k+1)$ ème personne n'ait pas choisi un étage parmi les $j-1$ déjà choisis par les k première personnes : c'est $\frac{n-(j-1)}{n}$.

Et de même, $P_{[S_k=j]}(S_{k+1} = j)$ est la probabilité que la $(k+1)$ ème personne ait choisi un étage parmi les j déjà choisis : elle vaut $\frac{j}{n}$.

Ainsi, $P(S_{k+1} = j) = \frac{n-j+1}{n} P(S_k = j-1) + \frac{j}{n} P(S_k = j)$.

2. La variable aléatoire S_k prend ses valeurs dans $\llbracket 0, n \rrbracket$, et donc étant à support fini, elle admet une espérance. On a alors

$$\begin{aligned} E(S_k) &= \sum_{j=0}^n j P(S_{k+1} = j) = \sum_{j=1}^n j P(S_{k+1} = j) = \frac{1}{n} \sum_{j=1}^n j^2 P(S_k = j) + \frac{1}{n} \sum_{j=1}^n j(n-j+1) P(S_k = j-1) \\ &= \frac{1}{n} \sum_{j=1}^n j^2 P(S_k = j) + \frac{1}{n} \sum_{i=0}^{n-1} (i+1)(n-i) P(S_k = i) \\ &= \frac{1}{n} \sum_{j=1}^n j^2 P(S_k = j) + \frac{1}{n} \sum_{i=0}^{n-1} (in + n - i^2 - i) P(S_k = i) \\ &= \frac{1}{n} \sum_{j=1}^n j^2 P(S_k = j) + \sum_{i=0}^{n-1} i P(S_k = i) + \sum_{i=0}^{n-1} P(S_k = i) - \frac{1}{n} \sum_{i=0}^{n-1} i^2 P(S_k = i) - \frac{1}{n} \sum_{i=0}^{n-1} i P(S_k = i) \\ &= n P(S_k = n) + \frac{n-1}{n} \sum_{i=0}^{n-1} i P(S_k = i) + \underbrace{\sum_{i=0}^{n-1} P(S_k = i)}_{=1} - P(S_k = n) \end{aligned}$$

Chgt d'indice
On a posé $i = j-1$.

$$\begin{aligned}
 &= nP(S_k = n) + \frac{n-1}{n} (E(S_k) - nP(S_k = n)) + 1 - P(S_k = n) \\
 &= 1 + \frac{n-1}{n} E(S_k) = \boxed{1 + \left(1 - \frac{1}{n}\right) E(S_k)}.
 \end{aligned}$$

3. Puisque $S_0 = 0$, on a $\boxed{E(S_0) = 0}$.

La relation $E(S_{k+1}) = \left(1 - \frac{1}{n}\right) E(S_k) + 1$ prouve que la suite $((E(S_k))_{k \in \mathbb{N}}$ est arithmético-géométrique de raison $1 - \frac{1}{n}$.
 Nous savons qu'il existe alors une unique constante λ telle que la suite $u_k = E(S_k) + \lambda$ soit géométrique de raison $1 - \frac{1}{n}$.

Or, on a

$$u_{k+1} = E(S_{k+1}) + \lambda = 1 + \lambda + \left(1 - \frac{1}{n}\right) E(S_k) = 1 + \lambda + \left(1 - \frac{1}{n}\right) (u_k - \lambda) = \left(1 - \frac{1}{n}\right) u_k + 1 + \frac{\lambda}{n}.$$

Et donc $(u_k)_k$ est géométrique de raison $1 - \frac{1}{n}$ si et seulement si $1 + \frac{\lambda}{n} = 0 \Leftrightarrow \lambda = -n$.

$$\text{On a alors } u_k = u_0 \left(1 - \frac{1}{n}\right)^k = (E(S_0) - n) \left(1 - \frac{1}{n}\right)^k = -n \left(1 - \frac{1}{n}\right)^k.$$

Et donc au final,

$$E(S_k) = u_k - \lambda = n - n \left(1 - \frac{1}{n}\right)^k = \boxed{n \left(1 - \left(1 - \frac{1}{n}\right)^k\right)}.$$

Remarque : ce résultat peut s'obtenir avec des méthodes totalement différentes, en utilisant des variables indicatrices. Voir l'exercice 8 du TD2.

4.a. Comme pour les S_k , S est une variable à valeurs dans $\llbracket 0, n \rrbracket$, et donc $\boxed{\text{possède une espérance}}$.

4.b. La probabilité conditionnelle $P_{[X=k]}(S = j)$ est la probabilité, sachant que k personnes sont dans l'ascenseur, que j étages soient choisis. C'est donc $P(S_k = j)$.

On a alors

$$E(S|X = k) = \sum_{j=0}^n j P_{[X=k]}(S = j) = \sum_{j=0}^n j P(S_k = j) = E(S_k).$$

4.c. Par la formule de l'espérance totale³ appliquée au système complet d'événements $\{[X = k], k \in \mathbb{N}\}$, ³ Qui s'applique bien puisque S admet une espérance.

$$\begin{aligned}
 E(S) &= \sum_{k=0}^{+\infty} P(X = k) E(S|X = k) \\
 &= \sum_{k=0}^{+\infty} e^{-\lambda} \frac{\lambda^k}{k!} n \left(1 - \left(1 - \frac{1}{n}\right)^k\right) \\
 &= n e^{-\lambda} \left(\sum_{k=0}^{+\infty} \frac{\lambda^k}{k!} - \sum_{k=0}^{+\infty} \frac{(\lambda(1 - \frac{1}{n}))^k}{k!} \right) \\
 &= n e^{-\lambda} \left(e^{\lambda} - e^{\lambda(1 - \frac{1}{n})} \right) = \boxed{n \left(1 - e^{-\frac{\lambda}{n}}\right)}.
 \end{aligned}$$

On a reconnu deux séries exponentielles.

EXERCICE 3 (EDHEC 2003)

1.a. Puisque f est décroissante, pour tout entier $k \geq p$, f admet son minimum sur $[k, k+1]$ en $k+1$. Et donc pour tout $t \in [k, k+1]$, on a $f(k+1) \leq f(t)$.

Par croissance de l'intégrale, il vient alors

$$f(k+1) = \int_k^{k+1} f(k+1) dt \leq \int_k^{k+1} f(t) dt.$$

Rédaction

Il y a sûrement dans votre cours de première année une formule toute faite pour l'expression du terme général d'une suite vérifiant

$$u_{n+1} = a u_n + b.$$

Si vous la connaissez par cœur, vous pouvez l'utiliser, mais les risques de se tromper sont grands, et il est conseillé de plutôt connaître la méthode pour retrouver cette formule.

Danger !

Attention à ne pas mélanger les indices : ici n est fixé, et on parle bien de la suite $(u_k)_{k \in \mathbb{N}}$ et non $(u_n)_{n \in \mathbb{N}}$.

En sommant ces relations pour k variant de p à $n - 1$, il vient

$$\sum_{k=p}^{n-1} f(k+1) \leq \sum_{k=p}^{n-1} \int_k^{k+1} f(t) dt = \int_p^n f(t) dt.$$

Relation de Chasles.

Mais d'autre part, on a

$$\sum_{k=p}^{n-1} f(k+1) = \sum_{i=p+1}^n f(i) = \sum_{i=p}^n f(i) - f(p) = S_n - f(p).$$

Chgt d'indice
 $i = k + 1.$

Et donc on a bien

$$S_n - f(p) \leq \int_p^n f(t) dt.$$

1.b. Puisque la série de terme général $f(n)$ est à termes positifs, elle converge si et seulement si la suite de ses sommes partielles est majorée.

Or, d'après la question 1.a, pour tout $n \geq p$, on a $S_n \leq f(p) + \int_p^n f(t) dt$.

D'autre part, f étant positive sur $[p, +\infty[$, on a

$$\int_p^n f(t) dt = \int_p^{+\infty} f(t) dt - \underbrace{\int_n^{+\infty} f(t) dt}_{\geq 0} \leq \int_p^{+\infty} f(t) dt.$$

Majoration
Cette inégalité ne prouve pas que la suite S_n est majorée : le majorant dépend encore de n !

Et donc pour tout $n \geq p$, $S_n \leq \int_p^{+\infty} f(t) dt + f(p)$: la suite $(S_n)_{n \geq p}$ est donc majorée, et ainsi la série de terme général $f(n)$ converge.

2.a. Soit $n \geq p$, et $k \geq n$. Alors, comme précédemment, par décroissance de f , on a, pour tout $t \in [k, k + 1]$, $f(k + 1) \leq f(t) \leq f(k)$, de sorte que, par croissance de l'intégrale,

$$f(k+1) = \int_k^{k+1} f(k+1) dt \leq \int_k^{k+1} f(t) dt \leq \int_k^{k+1} f(k) dt = f(k).$$

En sommant ces relations pour k variant de n à $m \geq n$, il vient

$$\sum_{k=n}^m f(k+1) \leq \int_n^{m+1} f(t) dt \leq \sum_{k=n}^m f(k).$$

Lorsque m tend vers $+\infty$, on a alors $\sum_{k=n}^{+\infty} f(k+1) \leq \int_n^{+\infty} f(t) dt \leq \sum_{k=n}^{+\infty} f(k)$.

Mais $\sum_{k=n}^{+\infty} f(k+1) = \sum_{i=n+1}^{+\infty} f(i) = R_n$, de sorte que

$$R_n \leq \int_n^{+\infty} f(t) dt \leq R_n + f(n).$$

On en déduit que $\int_n^{+\infty} f(t) dt - f(n) \leq R_n \leq \int_n^{+\infty} f(t) dt$.

Limite
Le passage à la limite est bien justifié car toutes les quantités en jeu convergent : par hypothèse, c'est le cas de $\int_n^{+\infty} f(t) dt$ et pour la série, cela a été prouvé à la question 1.b.

2.b. On a⁴

$$1 - \frac{f(n)}{\int_n^{+\infty} f(t) dt} \leq \frac{R_n}{\int_n^{+\infty} f(t) dt} \leq 1.$$

D'après le théorème des gendarmes, si $\lim_{n \rightarrow +\infty} \frac{f(n)}{\int_n^{+\infty} f(t) dt} = 0$, alors $\lim_{n \rightarrow +\infty} \frac{R_n}{\int_n^{+\infty} f(t) dt} = 1$,

de sorte que $R_n \underset{n \rightarrow +\infty}{\sim} \int_n^{+\infty} f(t) dt$.

Notons que la condition $\lim_{n \rightarrow +\infty} \frac{f(n)}{\int_n^{+\infty} f(t) dt} = 0$ s'écrit encore $f(n) = o_{n \rightarrow +\infty} \left(\int_n^{+\infty} f(t) dt \right)$.

⁴ f étant strictement positive, son intégrale est non nulle et donc la division est légitime.

- 3.a. Sur $[2; +\infty[$, les fonctions $x \mapsto x$ et $x \mapsto \ln(x)$ sont positives strictement, continues et croissantes, donc il en est de même de $x \mapsto x(\ln x)^2$. Et alors, par passage à l'inverse⁵, f est strictement positive, continue et décroissante.
De plus on a, pour $A > 2$,

$$\int_2^A f(t) dt = \left[-\frac{1}{\ln t} \right]_2^A = \frac{1}{\ln 2} - \frac{1}{\ln(A)} \xrightarrow{A \rightarrow +\infty} \frac{1}{\ln 2}.$$

Et donc $\int_2^{+\infty} f(t) dt$ converge.

De même, on prouve que pour $n \geq 2$, $\int_n^{+\infty} f(t) dt = \frac{1}{\ln(n)}$.

Enfin, on a $f(n) = \frac{1}{n \ln(n)^2} = \underset{n \rightarrow +\infty}{o} \left(\frac{1}{\ln n} \right)$ et donc la condition trouvée à la question 2.b est vérifiée.

- 3.b. D'après ce qui a été dit à la question 2.b, on a donc

$$R_n \underset{n \rightarrow +\infty}{\sim} \int_n^{+\infty} f(t) dt = \frac{1}{\ln n}.$$

- 3.c. La fonction \ln est concave (car dérivable, et de dérivée égale à $x \mapsto \frac{1}{x}$, qui est décroissante), et donc sa courbe est située en dessous de ses tangentes.
En particulier, sa tangente au point d'abscisse 1 est la droite d'équation $y = x - 1$, de sorte que

$$\forall x > 0, \ln(x) \leq x - 1.$$

On a alors notamment, pour $n \geq 2$, $\ln(n) \leq n - 1$ et donc $\frac{1}{\ln n} \geq \frac{1}{n-1} \geq 0$.

Puisque la série de terme général $\frac{1}{n-1}$ diverge⁶, par critère de comparaison pour les séries à termes positifs, il en est de même de la série de terme général $\frac{1}{\ln n}$.

Et donc, d'après le critère des équivalents pour les séries à termes positifs⁷, la série de terme général R_n diverge également.

Autre méthode : on a $\frac{1}{n} = \underset{n \rightarrow +\infty}{o} \left(\frac{1}{\ln n} \right)$.

Et donc si $\sum \frac{1}{\ln n}$ convergerait, il en serait de même de $\sum \frac{1}{n}$, ce qui n'est pas le cas, donc $\sum \frac{1}{\ln n}$ diverge, et donc $\sum R_n$ diverge également.

PROBLÈME (EML 2005)

- 1.a. On a $\frac{1}{(n+x)(n+y)} \underset{n \rightarrow +\infty}{\sim} \frac{1}{n^2}$ et $\frac{1}{(n+x)^2(n+y)} \underset{n \rightarrow +\infty}{\sim} \frac{1}{n^3}$.
De plus, il est évident qu'il s'agit là de séries à termes positifs. Puisque les séries de Riemann $\sum \frac{1}{n^2}$ et $\sum \frac{1}{n^3}$ convergent, par critère de comparaison pour les séries à termes positifs,

$$\sum_{n \geq 1} \frac{1}{(n+x)(n+y)} \text{ et } \sum_{n \geq 1} \frac{1}{(n+x)^2(n+y)} \text{ convergent.}$$

- 1.b. Pour $x \geq 0$, on a

$$\frac{1}{n} - \frac{1}{n+x} = \frac{n+x-n}{n(n+x)} = \frac{x}{n(n+x)} \underset{n \rightarrow +\infty}{\sim} \frac{x}{n^2}$$

Si $x = 0$, la convergence est immédiate, et si $x \neq 0$, alors la série de Riemann $\sum \frac{x}{n^2}$ converge, donc par critère de comparaison pour les séries positives,

$$\sum_{n \geq 1} \left(\frac{1}{n} - \frac{1}{n+x} \right) \text{ converge.}$$

⁵ L'inverse d'une fonction croissante positive est une fonction décroissante.

Astuce

Ces primitives sont classiques : $\frac{1}{x(\ln x)^k}$ est de la forme $\frac{u'}{u^k}$ avec $u = \ln$ et donc nous savons l'intégrer.

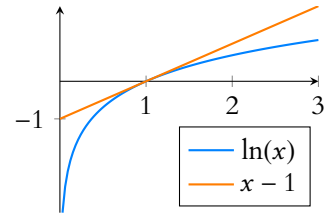


FIGURE 24.1— \ln et sa tangente en $x = 1$.

Alternative

Une telle inégalité se prouve également en étudiant la fonction $x \mapsto \ln(x) - (x-1)$, il n'est pas nécessaire d'utiliser un argument de concavité.

⁶ C'est la série harmonique.

⁷ R_n est évidemment positif car somme de termes positifs.

2. Nous avons

$$S(0) = \sum_{n=1}^{+\infty} \left(\frac{1}{n} - \frac{1}{n} \right) = \sum_{n=1}^{+\infty} 0 = \boxed{0}.$$

De même, on a $S(1) = \sum_{n=1}^{+\infty} \left(\frac{1}{n} - \frac{1}{n+1} \right).$

Nous reconnaissons là une série télescopique. Pour calculer sa somme, repassons par les sommes partielles. Soit donc $N \in \mathbf{N}^*$:

$$\sum_{n=1}^N \left(\frac{1}{n} - \frac{1}{n+1} \right) = \left(1 - \frac{1}{2} \right) + \left(\frac{1}{2} - \frac{1}{3} \right) + \dots + \left(\frac{1}{N-1} - \frac{1}{N} \right) + \left(\frac{1}{N} - \frac{1}{N+1} \right) = 1 - \frac{1}{N+1} \xrightarrow{N \rightarrow +\infty} 1.$$

Et donc $\boxed{S(1) = 1}.$

3.a. Soient $(x, y) \in \mathbf{R}^+$. Alors

$$\begin{aligned} S(y) - S(x) &= \sum_{n=1}^{+\infty} \left(\frac{1}{n} - \frac{1}{n+y} \right) - \sum_{n=1}^{+\infty} \left(\frac{1}{n} - \frac{1}{n+x} \right) \\ &= \sum_{n=1}^{+\infty} \left(\frac{1}{n} - \frac{1}{n+y} - \frac{1}{n} + \frac{1}{n+x} \right) \\ &= \sum_{n=1}^{+\infty} \left(\frac{1}{n+x} - \frac{1}{n+y} \right) \\ &= \sum_{n=1}^{+\infty} \frac{n+y - (n+x)}{(n+x)(n+y)} \\ &= \sum_{n=1}^{+\infty} \frac{y-x}{(n+x)(n+y)} \\ &= \boxed{(y-x) \sum_{n=1}^{+\infty} \frac{1}{(n+x)(n+y)}}. \end{aligned}$$

Notons que cette série converge car c'est la somme de deux séries convergentes.

3.b. En passant à la valeur absolue dans l'identité obtenue précédemment, il vient

$$|S(y) - S(x)| = |y-x| \left| \sum_{n=1}^{+\infty} \frac{1}{(n+x)(n+y)} \right| \leq |y-x| \sum_{n=1}^{+\infty} \left| \frac{1}{(n+x)(n+y)} \right|$$

Inégalité triangulaire.

Or, pour $x, y \in \mathbf{R}_+$, on a $0 \leq n^2 \leq (n+x)(n+y)$ et donc $0 \leq \frac{1}{(n+x)(n+y)} \leq \frac{1}{n^2}.$

On en déduit alors que $\sum_{n=1}^{+\infty} \frac{1}{(n+x)(n+y)} \leq \sum_{n=1}^{+\infty} \frac{1}{n^2}$ et donc

$$|S(y) - S(x)| \leq |y-x| \sum_{n=1}^{+\infty} \frac{1}{n^2} \leq \boxed{|y-x| \frac{\pi^2}{6}}.$$

3.c. Soit $x_0 \in \mathbf{R}_+$. Alors, grâce au résultat de la question précédente et au théorème des gendarmes, on a

$$\lim_{y \rightarrow x_0} |S(y) - S(x_0)| = 0.$$

Et donc $\lim_{y \rightarrow x_0} S(y) = S(x_0)$. Ainsi, S est continue en x_0 , et ceci étant valable pour tout

$x_0 \in \mathbf{R}_+$, $\boxed{S \text{ est continue sur } \mathbf{R}_+}.$

4.a. En réutilisant le résultat de la question 6.a, on a

$$\begin{aligned} \left| \frac{S(y) - S(x)}{y-x} - \sum_{n=1}^{+\infty} \frac{1}{(n+x)^2} \right| &= \left| \sum_{n=1}^{+\infty} \frac{1}{(n+x)(n+y)} - \sum_{n=1}^{+\infty} \frac{1}{(n+x)^2} \right| \\ &= \left| \sum_{n=1}^{+\infty} \left(\frac{1}{(n+x)(n+y)} - \frac{1}{(n+x)^2} \right) \right| \end{aligned}$$

$$\begin{aligned}
&= \left| \sum_{n=1}^{+\infty} \frac{n+x-(n+y)}{(n+x)^2(n+y)} \right| \\
&= \left| \sum_{n=1}^{+\infty} \frac{x-y}{(n+x)^2(n+y)} \right| \\
&= |y-x| \sum_{n=1}^{+\infty} \frac{1}{(n+x)^2(n+y)} \\
&\leq \boxed{|y-x| \sum_{n=1}^{+\infty} \frac{1}{n^3}}.
\end{aligned}$$

Détail

Pour tous $x, y \in \mathbf{R}_+^*$, on a

$$0 \leq \frac{1}{(n+x)^2} \leq \frac{1}{n^2}$$

et

$$0 \leq \frac{1}{n+y} \leq \frac{1}{n}.$$

4.b. \triangle Puisque la dérivation est linéaire, la dérivée d'une somme est la somme des dérivées.

Et comme la dérivée de $x \mapsto \frac{1}{n} - \frac{1}{n+x}$ est $x \mapsto \frac{1}{(n+x)^2}$, nous avons envie d'affirmer que

la dérivée de $x \mapsto \sum_{n=1}^{+\infty} \left(\frac{1}{n} - \frac{1}{n+x} \right)$ est $x \mapsto \sum_{n=1}^{+\infty} \frac{1}{(n+x)^2}$.

Nous allons prouver que ceci est vrai, mais ce n'est pas immédiat. En effet, le fait que la dérivée d'une somme soit la somme des dérivées est vrai pour une somme finie. Mais ce n'est pas vrai en toutes généralités pour une somme infinie. En effet, une somme infinie est en réalité en limite, et nous ne disposons d'aucun théorème affirmant que la dérivée d'une limite est la limite des dérivées.

C'est vrai sous certaines hypothèses, mais les énoncés en question ne sont pas au programme en ECS.

Donc lorsqu'on doit dériver une fonction définie par une série, nous n'avons qu'une option : revenir à la définition d'une dérivée en tant que limite d'un taux d'accroissement.

Puisque $\sum_{n=1}^{+\infty} \frac{1}{n^3}$ est une constante, lorsque $y \rightarrow x$, on a, par le théorème des gendarmes,

$$\lim_{y \rightarrow x} \left| \frac{S(y) - S(x)}{y - x} - \sum_{n=1}^{+\infty} \frac{1}{(n+x)^2} \right| = 0 \Leftrightarrow \lim_{y \rightarrow x} \frac{S(y) - S(x)}{y - x} = \sum_{n=1}^{+\infty} \frac{1}{(n+x)^2}.$$

Ainsi, S est dérivable en x , et de plus

$$S'(x) = \sum_{n=1}^{+\infty} \frac{1}{(n+x)^2}.$$

4.c. C'est un simple calcul : on a

$$S'(0) = \sum_{n=1}^{+\infty} \frac{1}{n^2} = \boxed{\frac{\pi^2}{6}} \text{ et } S'(1) = \sum_{n=1}^{+\infty} \frac{1}{(n+1)^2} = \sum_{n=2}^{+\infty} \frac{1}{n^2} = \sum_{n=1}^{+\infty} \frac{1}{n^2} - \frac{1}{1^2} = \boxed{\frac{\pi^2}{6} - 1}.$$

5. Puisque S est de classe \mathcal{C}^1 , elle est concave si et seulement si S' est décroissante.

Soient donc $(x, y) \in (\mathbf{R}_+)^2$, avec $x \leq y$. Alors pour tout $n \in \mathbf{N}$, on a $\frac{1}{(n+x)^2} \geq \frac{1}{(n+y)^2}$ et donc

$$S'(x) = \sum_{n=1}^{+\infty} \frac{1}{(n+x)^2} \geq \sum_{n=1}^{+\infty} \frac{1}{(n+y)^2} = S'(y).$$

Donc S' est décroissante et par conséquent, S est concave.

6.a. La fonction φ est continue sur $[1, +\infty[$, donc le seul éventuel problème de convergence de l'intégrale est en $+\infty$.

Soit $A \geq 1$. Alors

$$\int_1^A \varphi(t) dt = [\ln(t) - \ln(t+x)]_1^A = \ln(A) - \ln(A+x) + \ln(1+x) = \ln\left(\frac{A}{A+x}\right) + \ln(1+x).$$

Attention !

On a l'habitude de caractériser la convexité/concavité via le signe de la dérivée seconde. Ceci est valable pour une fonction de classe \mathcal{C}^2 . Mais il ne faut pas oublier que pour une fonction \mathcal{C}^1 , on a une caractérisation en fonction du sens de variation de la dérivée.

Lorsque $A \rightarrow +\infty$, $\frac{A}{A+x} \rightarrow 1$, et donc $\ln\left(\frac{A}{A+x}\right) \rightarrow 0$.

On en déduit que donc

$$\int_1^A \varphi(t) dt \xrightarrow{A \rightarrow +\infty} \ln(1+x)$$

et donc $\int_1^{+\infty} \varphi(t) dt$ converge et $\int_1^{+\infty} \varphi(t) dt = \ln(1+x)$.

6.b. La fonction φ est dérivable sur $[1, +\infty[$ et $\varphi'(t) = -\frac{1}{t^2} + \frac{1}{(t+x)^2} \leq 0$ car $\frac{1}{t^2} \geq \frac{1}{(t+x)^2}$.
Ainsi, φ est décroissante. Donc pour tout $t \in [n, n+1]$, on a

$$\varphi(n+1) \leq \varphi(t) \leq \varphi(n).$$

Par croissance de l'intégrale, on en déduit que

$$\varphi(n+1) = \int_n^{n+1} \varphi(n+1) dt \leq \int_n^{n+1} \varphi(t) dt \leq \int_n^{n+1} \varphi(n) dt = \varphi(n).$$

Soit $N \in \mathbf{N}^*$. Alors en sommant les relations précédemment obtenues pour n allant de 1 à N , il vient

$$\sum_{n=1}^{N+1} \varphi(n) - \varphi(1) = \sum_{n=2}^{N+1} \varphi(n) = \sum_{n=1}^N \varphi(n+1) \leq \int_1^{N+1} \varphi(t) dt \leq \sum_{n=1}^N \varphi(n).$$

En passant à la limite lorsque $N \rightarrow +\infty$, on a alors

$$S(x) - \varphi(1) = S(x) - 1 \leq \int_1^{+\infty} \varphi(t) dt \leq S(x).$$

Soit encore

$$\int_1^{+\infty} \varphi(t) dt \leq S(x) \leq \int_1^{+\infty} \varphi(t) dt + 1.$$

6.c. Puisque nous savons que $\int_1^{+\infty} \varphi(t) dt = \ln(1+x)$, l'inégalité précédente se réécrit

$$\ln(1+x) \leq S(x) \leq \ln(1+x) + 1.$$

En divisant cette inégalité par $\ln(1+x) \geq 0$ il vient

$$1 \leq \frac{S(x)}{\ln(1+x)} \leq 1 + \frac{1}{\ln(1+x)}.$$

D'après le théorème des gendarmes, lorsque $x \rightarrow +\infty$, on a alors

$$\lim_{x \rightarrow +\infty} \frac{S(x)}{\ln(1+x)} = 1 \text{ et donc } \boxed{S(x) \underset{x \rightarrow +\infty}{\sim} \ln(1+x)}.$$

Il reste donc à prouver que $\ln(1+x) \underset{x \rightarrow +\infty}{\sim} \ln x$.

Pour cela, calculons le quotient

$$\frac{\ln(x+1)}{\ln(x)} = \frac{\ln x + \ln\left(1 + \frac{1}{x}\right)}{\ln(x)} = 1 + \frac{\ln\left(1 + \frac{1}{x}\right)}{\ln x}.$$

Or, $\ln(x) \xrightarrow{x \rightarrow +\infty} +\infty$ et $\ln\left(1 + \frac{1}{x}\right) \xrightarrow{x \rightarrow +\infty} 0$, donc le quotient $\frac{\ln\left(1 + \frac{1}{x}\right)}{\ln x}$ tend vers 0 et donc

$$\frac{\ln(x+1)}{\ln(x)} \xrightarrow{x \rightarrow +\infty} 1.$$

On en déduit que $\ln(x+1) \underset{x \rightarrow +\infty}{\sim} \ln(x)$ et donc $\boxed{S(x) \underset{x \rightarrow +\infty}{\sim} \ln(x)}$.

Hypothèses

Pour utiliser la croissance de l'intégrale, il nous faut absolument avoir une inégalité valable **pour tout** t dans l'intervalle d'intégration.

Astuce

L'intégrale d'une constante sur un segment de longueur 1 est égale à cette constante.

Signes

Quand on multiplie ou divise les deux membres d'une inégalité par un réel λ , il est important de vérifier le signe de λ pour savoir si l'on change le sens de l'inégalité ou non.

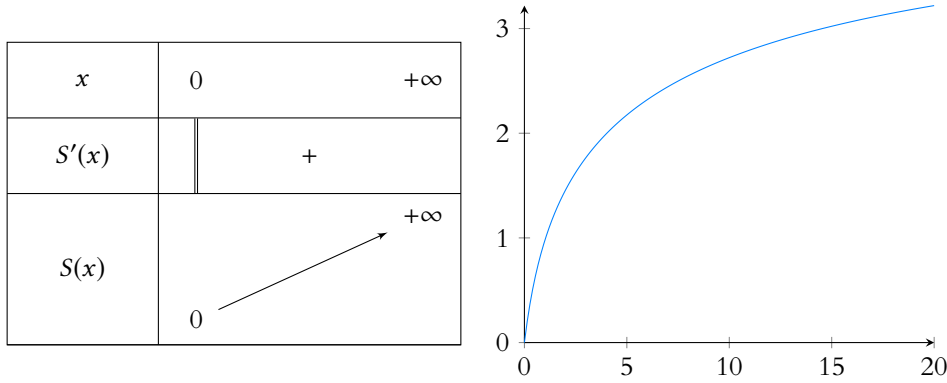
Danger !

Bien que $x+1 \sim x$, cela ne saurait suffire à conclure : on n'a pas le droit de composer les équivalents à gauche !

7. Pour tout $x \geq 0$ et pour tout $n \geq 1$, on a $\frac{1}{(n+x)^2} > 0$ et donc

$$S'(x) = \sum_{n=1}^{+\infty} \frac{1}{(n+x)^2} > 0.$$

On en déduit que S est strictement croissante sur \mathbf{R}_+ et donc on a



Notons qu'il est important de faire apparaître sur le graphique les propriétés de S mises en évidence dans les questions précédentes. Au delà des limites et de la croissance, on n'oubliera pas la concavité.

Asymptotiquement

L'équivalent obtenu précédemment nous dit qu'au voisinage de $+\infty$, $S(x)$ se comporte «comme» $\ln(x)$, et donc tend vers $+\infty$, mais lentement, comme le fait $\ln(x)$.

CONCOURS BLANC 1 (07.11.18)

EXERCICE 1

On considère la suite de nombres réels $(u_n)_{n \in \mathbf{N}}$ définie par la relation de récurrence :

$$\begin{cases} u_{n+1} = u_n + u_n^2 \\ u_0 = a, a \in \mathbf{R}_+^* \end{cases}$$

1. Convergence de $(u_n)_{n \in \mathbf{N}}$.

- a. Montrer que la suite $(u_n)_{n \in \mathbf{N}}$ est strictement positive et monotone.
- b. Montrer que cette suite diverge vers l'infini.

2. Comportement asymptotique de $(u_n)_{n \in \mathbf{N}}$.

On définit la suite $(v_n)_{n \in \mathbf{N}}$ par : $v_n = \frac{1}{2^n} \ln u_n$.

- a. Prouver que pour tout entier n de \mathbf{N} :

$$v_{n+1} - v_n = \frac{1}{2^{n+1}} \ln \left(1 + \frac{1}{u_n} \right).$$

En déduire que quels que soient les entiers naturels n et p :

$$0 < v_{n+p+1} - v_{n+p} \leq \frac{1}{2^{n+p+1}} \ln \left(1 + \frac{1}{u_n} \right).$$

- b. En déduire que quels que soient les entiers naturels k et n :

$$0 < v_{n+k+1} - v_n \leq \frac{1}{2^n} \ln \left(1 + \frac{1}{u_n} \right).$$

- c. Démontrer que la suite $(v_n)_{n \in \mathbf{N}}$ est majorée, puis qu'elle converge vers une limite notée α .
- d. Montrer que :

$$\forall n \in \mathbf{N}, \quad u_n \leq \exp(\alpha 2^n).$$

En passant à la limite pour n fixé dans l'encadrement de la question 2.b, montrer que :

$$\forall n \in \mathbf{N}, \quad \exp(\alpha 2^n) \leq u_n + 1.$$

En déduire, lorsque n tend vers l'infini, l'équivalent suivant :

$$u_n \underset{n \rightarrow +\infty}{\sim} \exp(\alpha 2^n).$$

- e. On pose :

$$\beta_n = \exp(\alpha 2^n) - u_n.$$

Montrer que la suite $(\beta_n)_{n \in \mathbf{N}}$ est bornée et qu'elle vérifie la relation suivante :

$$\forall n \in \mathbf{N}, \quad 2\beta_n - 1 = (\beta_{n+1} + \beta_n^2 - \beta_n) \exp(-\alpha 2^n).$$

- f. Prouver enfin que lorsque n tend vers l'infini :

$$u_n = -\frac{1}{2} + \exp(\alpha 2^n) + \underset{n \rightarrow +\infty}{o}(1).$$

EXERCICE 2

Définitions : soit E un espace vectoriel sur \mathbf{C} .

- On note id_E l'application identité de E .
- Pour tout endomorphisme f de E , on note $f^0 = \text{id}_E$, et pour tout entier naturel k , $f^{k+1} = f^k \circ f$.
- Soit $p \in \mathbf{N}^*$. On dit qu'un endomorphisme est **cyclique d'ordre p** s'il existe un élément x_0 de E vérifiant les conditions suivantes :

$$\star f^p(x_0) = x_0$$

• la famille $(x_0, f(x_0), \dots, f^{p-1}(x_0))$ est génératrice de E

• la famille $(x_0, f(x_0), \dots, f^{p-1}(x_0))$ est constituée d'éléments deux à deux distincts.

La famille $(x_0, f(x_0), \dots, f^{p-1}(x_0))$ est alors appelée un **cycle** de f .

Partie I : Étude d'un exemple

Dans cette partie, E est un espace vectoriel sur \mathbf{C} de dimension 3, et $\mathcal{B} = (e_1, e_2, e_3)$ est une base de E . On considère l'endomorphisme f de E dont la matrice dans la base \mathcal{B} est

$$A = \begin{pmatrix} 1 & 2 & 2 \\ 1 & 1 & 2 \\ -2 & -2 & -3 \end{pmatrix}.$$

1. Vérifier que $\mathcal{B}' = (e_1, f(e_1), f^2(e_1))$ est une base de E et déterminer la matrice de f dans cette base.
2. Montrer que f est cyclique d'ordre 4 et que $(e_1, f(e_1), f^2(e_1), f^3(e_1))$ est un cycle de f .
3. Montrer que $f^4 = \text{id}_E$. En déduire un polynôme annulateur de f .

Partie II : Cas général

Dans cette partie, E est un espace vectoriel sur \mathbf{C} de dimension n ($n \in \mathbf{N}^*$), et on considère un endomorphisme f de E , cyclique d'ordre p . Soit $(x_0, f(x_0), \dots, f^{p-1}(x_0))$ un cycle de f .

4. Montrer que $p \geq n$.
5. Montrer que $f^p = \text{id}_E$. En déduire que f est bijective.
6. On note m le plus grand des entiers naturels k tels que la famille $(x_0, f(x_0), \dots, f^{k-1}(x_0))$ est libre.
 - a. Montrer que $f^m(x_0)$ est combinaison linéaire des m vecteurs $x_0, f(x_0), \dots, f^{m-1}(x_0)$.
 - b. Montrer par récurrence que, pour tout entier naturel k supérieur ou égal à m , le vecteur $f^k(x_0)$ est combinaison linéaire des m vecteurs $x_0, f(x_0), \dots, f^{m-1}(x_0)$.
 - c. En déduire que $m = n$ et que la famille $(x_0, f(x_0), \dots, f^{n-1}(x_0))$ est une base de E .
7. On note a_0, a_1, \dots, a_{n-1} les n nombres complexes tels que

$$f^n(x_0) = a_0 x_0 + a_1 f(x_0) + \dots + a_{n-1} f^{n-1}(x_0).$$

- a. On considère l'endomorphisme g de E défini par $g = a_0 \text{id}_E + a_1 f + \dots + a_{n-1} f^{n-1}$.
Montrer que $\forall k \in \mathbf{N}, g(f^k(x_0)) = f^{n+k}(x_0)$.
En déduire que $f^n = a_0 \text{id}_E + a_1 f + \dots + a_{n-1} f^{n-1}$.
- b. Exprimer la matrice M de f dans la base $(x_0, f(x_0), \dots, f^{n-1}(x_0))$ en fonction des coefficients a_0, a_1, \dots, a_{n-1} .
- c. Montrer que

$$\forall \lambda \in \mathbf{C}, \text{rg}(f - \lambda \text{id}_E) \geq n - 1.$$

EXERCICE 3**Partie I : Étude d'une fonction f définie par une intégrale.**

1. Montrer que, pour tout $x \in]0; +\infty[$, l'intégrale $\int_0^{+\infty} \frac{e^{-t}}{x+t} dt$ converge.

On note $f :]0; +\infty[\rightarrow \mathbf{R}$ l'application définie, pour tout $x \in]0; +\infty[$ par : $f(x) = \int_0^{+\infty} \frac{e^{-t}}{x+t} dt$.

2. Montrer : $\forall x \in]0; +\infty[, f(x) \geq \int_0^1 \frac{e^{-t}}{x+t} dt$. En déduire : $f(x) \xrightarrow{x \rightarrow 0^+} +\infty$.
3. Montrer : $\forall x \in]0; +\infty[, 0 < f(x) \leq \frac{1}{x}$. En déduire : $f(x) \xrightarrow{x \rightarrow +\infty} 0$.
4. Montrer que l'intégrale $\int_0^{+\infty} t e^{-t} dt$ converge et que :

$$\forall x \in]0; +\infty[, \left| f(x) - \frac{1}{x} \right| \leq \frac{1}{x^2} \int_0^{+\infty} t e^{-t} dt.$$

En déduire : $f(x) \underset{x \rightarrow +\infty}{\sim} \frac{1}{x}$.

Partie II : Dérivabilité et expression de la dérivée de f sous forme d'une intégrale.

5. Soit $(x, h) \in]0; +\infty[\times \mathbb{R}^*$ tel que $h > -\frac{x}{2}$.

a. Montrer que l'intégrale $\int_0^{+\infty} \frac{e^{-t}}{(x+t)^2} dt$ converge.

b. Établir : $\forall t \in [0; +\infty[$, $\left| \frac{1}{h} \left(\frac{1}{x+h+t} - \frac{1}{x+t} \right) + \frac{1}{(x+t)^2} \right| \leq \frac{2|h|}{x^3}$.

c. En déduire : $\left| \frac{f(x+h) - f(x)}{h} + \int_0^{+\infty} \frac{e^{-t}}{(x+t)^2} dt \right| \leq \frac{2|h|}{x^3}$.

6. En déduire que f est dérivable sur $]0; +\infty[$ et que : $\forall x \in]0; +\infty[$, $f'(x) = -\int_0^{+\infty} \frac{e^{-t}}{(x+t)^2} dt$.

7. Montrer, pour tout $x \in]0; +\infty[$ et tout $(\varepsilon, A) \in]0, 1[\times [1; +\infty[$:

$$\int_{\varepsilon}^A \frac{e^{-t}}{(x+t)^2} dt = -\frac{e^{-A}}{x+A} + \frac{e^{-\varepsilon}}{x+\varepsilon} - \int_{\varepsilon}^A \frac{e^{-t}}{x+t} dt.$$

8. En déduire : $\forall x \in]0; +\infty[$, $f'(x) = -\frac{1}{x} + f(x)$.

9. Montrer que f est de classe \mathcal{C}^2 sur $]0; +\infty[$ et que : $\forall x \in]0; +\infty[$, $f''(x) = \frac{1}{x^2} + f'(x)$.

EXERCICE 4

1. **Question préliminaire** : montrer que si X et Y sont deux variables aléatoires sur le même espace probabilisé (Ω, \mathcal{A}, P) , suivant toutes deux des lois de Bernoulli, alors

$$\text{Cov}(X, Y) = P([X = 1] \cap [Y = 1]) - P(X = 1)P(Y = 1).$$

Une urne contient une boule noire et $(n-1)$ boules blanches, n désignant un entier supérieur ou égal à 2. On vide l'urne en effectuant des tirages d'une boule de la manière suivante : le premier tirage s'effectue sans remise, le deuxième s'effectue avec remise, le troisième s'effectue sans remise, le quatrième s'effectue avec remise, ... D'une manière générale, les tirages d'ordre impair s'effectuent sans remise et les tirages d'ordre pair s'effectuent avec remise de la boule tirée.

2. a. Quel est le nombre total N de tirages effectués lors de cette épreuve ?

b. Pour j élément de $\llbracket 1, n-1 \rrbracket$, combien reste-t-il de boules avant le $(2j)$ ^{ème} tirage ?
Combien en reste-t-il avant le $(2j+1)$ ^{ème} tirage ?

On désigne par X_k la variable aléatoire qui vaut 1 si la boule noire est obtenue au k ^{ème} tirage (que ce soit la première fois ou non) et 0 sinon. On désigne par X la variable aléatoire égale au nombre d'apparitions de la boule noire lors de cette épreuve.

3. a. Calculer $P(X_1 = 1)$, $P(X_2 = 1)$.

b. Pour tout entier naturel j de $\llbracket 1, n-1 \rrbracket$, calculer $P(X_{2j+1} = 1)$ et $P(X_{2j} = 1)$.

c. En déduire la loi suivie par toutes les variables X_k .

4. Pour tout j élément de $\llbracket 1, n \rrbracket$, on note U_j l'événement : «on obtient la boule noire pour la 1^{ère} fois au $(2j-1)$ ^{ème} tirage».

a. En considérant l'état de l'urne avant le $(2n-2)$ ^{ème} tirage, montrer que $P(U_n) = 0$.

Montrer que : $\forall j \in \llbracket 1, n \rrbracket$, $P(U_j) = \frac{n-j}{n(n-1)}$.

b. Exprimer l'événement $[X = 1]$ en fonction des U_j , puis en déduire la valeur de $P(X = 1)$.

c. Montrer que $P(X = n) = \frac{1}{n!}$.

5. Montrer que $X = \sum_{k=1}^{2n-1} X_k$, puis en déduire l'espérance de X .

6. Soit i un entier naturel compris entre 0 et $n-2$.

a. Pour tout j de $\llbracket 1, 2n-2i-2 \rrbracket$, donner la valeur de $P_{[X_{2i+1}=1]}(X_{2i+j+1} = 1)$.

b. En utilisant la question 1, en déduire que : $\forall j \in \llbracket 1, 2n-2i-2 \rrbracket$, $\text{Cov}(X_{2i+1}, X_{2i+j+1}) = -\frac{1}{n^2}$.

7. Soit i un entier naturel compris entre 1 et $n-1$.

- a. Montrer que, pour tout k de $\llbracket 1, n - i - 1 \rrbracket$, $P_{[X_{2i}=1]}(X_{2i+2k} = 1) = \frac{1}{n-i}$.
- b. Montrer que, pour tout k de $\llbracket 0, n - i - 1 \rrbracket$, $P_{[X_{2i=1}]}(X_{2i+2k+1} = 1) = \frac{1}{n-i}$.
- c. En déduire que $\forall j \in \llbracket 1, 2n - 2i - 1 \rrbracket$, $\text{Cov}(X_{2i}, X_{2i+j}) = \frac{i}{n^2(n-i)}$.
8. En notant que $V(X) = \text{Cov}(X, X)$, montrer que la variance de X est donnée par :

$$V(X) = \frac{(2n+1)(n-1)}{n^2} - \frac{2}{n} \sum_{j=1}^{n-1} \frac{1}{j}.$$

CORRECTION DU CONCOURS BLANC 1

EXERCICE 1 (ECRICOME 2003)

1. Convergence de $(u_n)_{n \in \mathbf{N}}$.

1.a. Pour tout $n \in \mathbf{N}$, on a $u_{n+1} - u_n = u_n + u_n^2 - u_n = u_n^2 \geq 0$.

Donc $(u_n)_{n \in \mathbf{N}}$ est croissante.

Et par conséquent, pour tout $n \in \mathbf{N}$, $u_n \geq u_0 > 0$.

1.b. Puisque $(u_n)_{n \in \mathbf{N}}$ est croissante, soit elle converge vers un réel ℓ , soit elle diverge vers $+\infty$.

Supposons par l'absurde que (u_n) converge vers ℓ .

Alors $\lim_{n \rightarrow +\infty} u_{n+1} = \ell$ et $u_n + u_n^2 \xrightarrow[n \rightarrow +\infty]{} \ell + \ell^2$.

Par unicité de la limite, on a donc $\ell = \ell + \ell^2 \Leftrightarrow \ell = 0$.

Or, pour tout $n \in \mathbf{N}$, $u_n \geq a$, donc $\ell \geq a > 0$, contredisant le fait que $\ell = 0$.

C'est donc que (u_n) ne converge pas, et par conséquent, $u_n \xrightarrow[n \rightarrow +\infty]{} +\infty$.

2. Comportement asymptotique de $(u_n)_{n \in \mathbf{N}}$.

2.a. Quel que soit $n \in \mathbf{N}$, on a

$$\begin{aligned} v_{n+1} - v_n &= \frac{1}{2^{n+1}} \ln(u_{n+1}) - \frac{1}{2^n} \ln(u_n) = \frac{1}{2^{n+1}} \ln(u_n + u_n^2) - \frac{1}{2^n} \ln(u_n) \\ &= \frac{1}{2^{n+1}} (\ln(u_n + u_n^2) - 2 \ln(u_n)) = \frac{1}{2^{n+1}} \ln(u_n + u_n^2) - \ln(u_n^2) = \frac{1}{2^{n+1}} \ln\left(\frac{u_n + u_n^2}{u_n^2}\right) \\ &= \frac{1}{2^{n+1}} \ln\left(1 + \frac{1}{u_n}\right). \end{aligned}$$

En particulier, pour $n, p \in \mathbf{N}$, on a

$$v_{n+p+1} - v_{n+p} = \frac{1}{2^{n+p+1}} \ln\left(1 + \frac{1}{u_{n+p}}\right).$$

Puisque $\frac{1}{u_{n+p}} > 0$, le logarithme est bien strictement positif, et donc $v_{n+p+1} - v_{n+p} > 0$.

D'autre part, puisque (u_n) est croissante et positive, $u_{n+p} \geq u_n \geq 0$ et donc $\frac{1}{u_{n+p}} \leq \frac{1}{u_n}$, de

sorte que, par croissance du logarithme, $\ln\left(1 + \frac{1}{u_{n+p}}\right) \leq \ln\left(1 + \frac{1}{u_n}\right)$.

On a donc bien $v_{n+p+1} - v_{n+p} \leq \frac{1}{2^{n+p}} \ln\left(1 + \frac{1}{u_n}\right)$.

2.b. Soient $k, n \in \mathbf{N}$. Alors

$$\begin{aligned} v_{n+k+1} - v_n &= (v_{n+k+1} - v_{n+k}) + (v_{n+k} - v_{n+k-1}) + \cdots + (v_{n+1} - v_n) \\ &= \sum_{i=0}^k (v_{n+i+1} - v_{n+i}) \\ &\leq \sum_{i=0}^k \frac{1}{2^{n+i+1}} \ln\left(1 + \frac{1}{u_n}\right) \\ &\leq \frac{1}{2^{n+1}} \ln\left(1 + \frac{1}{u_n}\right) \sum_{i=0}^k \frac{1}{2^i} \\ &\leq \frac{1}{2^{n+1}} \ln\left(1 + \frac{1}{u_n}\right) \sum_{i=0}^{+\infty} \frac{1}{2^i} \\ &\leq \frac{1}{2^{n+1}} \ln\left(1 + \frac{1}{u_n}\right) \times \frac{1}{1 - \frac{1}{2}} \\ &\leq \frac{1}{2^n} \ln\left(1 + \frac{1}{u_n}\right). \end{aligned}$$

Remarque

La croissance de (u_n) est ici importante, car sans cette hypothèse, on pourrait avoir une suite à termes strictement positifs, et de limite nulle. Penser par exemple à $u_n = \frac{1}{n}$.

Détails

Puisque la série de terme général $\frac{1}{2^i}$ est à termes positifs, ses sommes partielles sont toutes plus petites que la somme de la série.

Et d'autre part, puisque les $v_{n+i+1} - v_{n+i}$ sont tous strictement positifs, leur somme, égale à $v_{n+k+1} - v_n$ est également strictement positive.

En résumé, on a bien prouvé que $0 < v_{n+k+1} - v_n \leq \frac{1}{2^n} \ln\left(1 + \frac{1}{u_n}\right)$.

2.c. En prenant $n = 0$ dans l'inégalité précédente, on obtient, pour tout $k \in \mathbf{N}$,

$$v_{k+1} \leq v_0 + \ln\left(1 + \frac{1}{u_0}\right).$$

Autrement dit, pour tout $k \geq 1$, $v_k \leq v_0 + \ln\left(1 + \frac{1}{u_0}\right)$.

Cette inégalité est évidemment encore vraie pour $k = 0$, et donc $v_0 + \ln\left(1 + \frac{1}{u_0}\right)$ est un majorant de la suite (v_n) .

Comme d'autre part nous avons prouvé à la question 2.a que¹ $v_{n+1} - v_n > 0$, la suite (v_n) est croissante.

Et donc par le théorème de la limite monotone, elle converge.

2.d. Puisque (v_n) est croissante et converge vers α , pour tout $n \in \mathbf{N}$, $v_n \leq \alpha$.

Et donc $\ln(u_n) \leq 2^n \alpha$, de sorte que par croissance de l'exponentielle, $u_n \leq \exp(\alpha 2^n)$.

Reprenons à présent l'encadrement de la question 2.b, fixons n et faisons tendre k vers l'infini.

Il vient alors

$$\alpha - v_n \leq \frac{1}{2^n} \ln\left(1 + \frac{1}{u_0}\right) \Leftrightarrow \alpha \leq \frac{1}{2^n} \ln(u_n) + \frac{1}{2^n} \ln\left(1 + \frac{1}{u_n}\right).$$

Et donc après multiplication par 2^n , on a

$$\alpha 2^n \leq \ln(u_n) + \ln\left(1 + \frac{1}{u_n}\right) = \ln\left(u_n \left(1 + \frac{1}{u_n}\right)\right) = \ln(1 + u_n).$$

Par croissance de l'exponentielle, on a donc

$$\exp(\alpha 2^n) \leq u_n + 1.$$

On a donc $u_n \leq \exp(\alpha 2^n) \leq u_n + 1$, soit encore²

$$1 \leq \frac{\exp(\alpha 2^n)}{u_n} \leq 1 + \frac{1}{u_n}.$$

Mais puisque $u_n \xrightarrow{n \rightarrow +\infty} +\infty$, $1 + \frac{1}{u_n} \xrightarrow{n \rightarrow +\infty} 1$ et donc par le théorème des gendarmes,

$$\lim_{n \rightarrow +\infty} \frac{\exp(\alpha 2^n)}{u_n} = 1.$$

Soit encore $u_n \underset{n \rightarrow +\infty}{\sim} \exp(\alpha 2^n)$.

2.e. En reprenant les inégalités de la question précédente, on a

$$u_n \leq \exp(\alpha 2^n) \leq u_n + 1 \Leftrightarrow 0 \leq \underbrace{\exp(\alpha 2^n) - u_n}_{=\beta_n} \leq 1.$$

Et donc (β_n) est majorée et minorée, donc bornée.

Pour prouver la relation demandée, notons que $(\exp(\alpha 2^n))^2 = \exp(\alpha 2^{n+1})$. Et donc

$$\begin{aligned} \beta_{n+1} + \beta_n^2 - \beta_n &= \exp(\alpha 2^{n+1}) - u_{n+1} + (\exp(\alpha 2^n) - u_n)^2 - (\exp(\alpha 2^n) - u_n) \\ &= \exp(\alpha 2^{n+1}) - u_n - u_n^2 + \exp(\alpha 2^{n+1}) - 2u_n \exp(\alpha 2^n) + u_n^2 - \exp(\alpha 2^n) + u_n \\ &= 2 \exp(\alpha 2^{n+1}) - 2u_n \exp(\alpha 2^n) - \exp(\alpha 2^n). \end{aligned}$$

Et donc après multiplication par $\exp(-\alpha 2^n)$, il vient

$$(\beta_{n+1} + \beta_n^2 - \beta_n) \exp(-\alpha 2^n) = 2 \exp(\alpha 2^n) - 2u_n - 1 = 2\beta_n - 1.$$

¹ Il suffit de prendre $k = 0$ dans l'inégalité de 2.a

² $u_n > 0$, donc la division par u_n ne change pas le sens de l'inégalité.

- 2.f. Nous avons prouvé précédemment que pour tout $n \in \mathbf{N}$, $|\beta_n| \leq 1$.
Et donc

$$\begin{aligned} |2\beta_n - 1| &\leq |\beta_{n+1} + \beta_n^2 - \beta_n| \exp(-\alpha 2^n) \\ &\leq (|\beta_{n+1}| + |\beta_n^2| + |\beta_n|) \exp(-\alpha 2^n) \\ &\leq 3 \exp(-\alpha 2^n). \end{aligned}$$

Inégalité triangulaire.

Par le théorème des gendarmes, on en déduit que $\lim_{n \rightarrow +\infty} |2\beta_n - 1| = 0$.

Et donc $\lim_{n \rightarrow +\infty} 2\beta_n = 1 \Leftrightarrow \lim_{n \rightarrow +\infty} \beta_n = \frac{1}{2}$.

Donc $\beta_n = \frac{1}{2} + o_{n \rightarrow +\infty}(1)$.

Et donc

$$u_n = \exp(\alpha 2^n) - (\exp(\alpha 2^n) - u_n) = \exp(\alpha 2^n) - \beta_n = \boxed{\exp(\alpha 2^n) - \frac{1}{2} + o_{n \rightarrow +\infty}(1)}.$$

Remarque : ce résultat permet de retrouver l'équivalent $u_n \underset{n \rightarrow +\infty}{\sim} \exp(\alpha 2^n)$, mais il est bien plus précis, puisqu'il affirme que la différence entre u_n et $\exp(\alpha 2^n)$ tend vers $-\frac{1}{2}$, et donc que u_n ne s'éloigne jamais trop de $\exp(\alpha 2^n)$.

Ceci n'est pas vrai pour toute suite équivalente à $\exp(\alpha 2^n)$. Par exemple, la suite $v_n = \exp(\alpha 2^n) + n^2$ est également équivalente³ à $\exp(\alpha 2^n)$, mais pourtant la différence entre v_n et $\exp(\alpha 2^n)$ tend vers l'infini.

o(1)
Rappelons que la notation $o(1)$ désigne toute suite de limite nulle.

³ Car $n^2 = o_{n \rightarrow +\infty}(\exp(\alpha 2^n))$.

EXERCICE 2 (EXTRAIT DE EML 2001)

Partie I : Étude d'un exemple

1. En utilisant la matrice, on a $f(e_1) = e_1 + e_2 - 2e_3$ et

$$f^2(e_1) = f(e_1 + e_2 - 2e_3) = e_1 + e_2 - 2e_3 + 2e_1 + e_2 - 2e_3 - 4e_1 - 4e_2 + 6e_3 = -e_1 - 2e_2 + 2e_3.$$

Puisque \mathcal{B}' est une famille de cardinal 3 dans un espace vectoriel de dimension 3, il s'agit de prouver qu'elle est libre.

Soient donc $(\lambda_1, \lambda_2, \lambda_3) \in \mathbf{C}^3$ tels que $\lambda_1 e_1 + \lambda_2 f(e_1) + \lambda_3 f^2(e_1) = 0$. Alors puisque (e_1, e_2, e_3) est une base⁴ de E , il vient

$$\begin{cases} \lambda_1 + \lambda_2 - \lambda_3 = 0 \\ \lambda_2 - 2\lambda_3 = 0 \\ -2\lambda_2 + 2\lambda_3 = 0 \end{cases} \Leftrightarrow \begin{cases} \lambda_1 + \lambda_2 - \lambda_3 = 0 \\ \lambda_2 = 2\lambda_3 \\ \lambda_2 = \lambda_3 \end{cases} \Leftrightarrow \begin{cases} \lambda_1 = 0 \\ \lambda_2 = 0 \\ \lambda_3 = 0 \end{cases}$$

Ainsi, $(e_1, f(e_1), f^2(e_1))$ est une famille libre de E : c'est une base de E .

De plus, on a

$$\begin{aligned} f(f^2(e_1)) &= -f(e_1) - 2f(e_2) + 2f(e_3) = -(e_1 + e_2 - 2e_3) - 2(2e_1 + e_2 - 2e_3) + 2(2e_1 + 2e_2 - 3e_3) \\ &= -e_1 + e_2 = -e_1 - f(e_1) - f^2(e_1). \end{aligned}$$

On en déduit que la matrice de f dans la base \mathcal{B}' est

$$\begin{pmatrix} f(e_1) & f(f(e_1)) & f(f^2(e_1)) \\ 0 & 0 & -1 \\ 1 & 0 & -1 \\ 0 & 1 & -1 \end{pmatrix} \begin{matrix} e_1 \\ f(e_1) \\ f^2(e_1) \end{matrix}.$$

2. Les conditions 2 et 3 définissant un endomorphisme cyclique d'ordre 4 sont bien vérifiées.
De plus, on a

$$f^4(e_1) = f(f^3(e_1)) = f(-e_1 + e_2) = e_1.$$

Donc f est cyclique d'ordre 4, et $(e_1, f(e_1), f^2(e_1), f^3(e_1))$ est un cycle de f .

Détails
La famille $(x_0, \dots, f^3(x_0))$ contient la base de la question 1, donc est évidemment génératrice.

⁴ Et donc en particulier une famille libre.

3. Pour prouver que deux endomorphismes sont égaux, il suffit de vérifier qu'ils coïncident sur une base. Montrons que f^4 et id_E coïncident sur la base \mathcal{B}' .
On a $f^4(e_1) = e_1$, $f^4(f(e_1)) = f(f^4(e_1)) = f(e_1)$ et $f^4(f^2(e_1)) = f^2(f^4(e_1)) = f^2(e_1)$.
Ainsi, on a bien $f^4 = \text{id}_E$.

Un polynôme annulateur de f est alors $X^4 - 1$.

Partie II : Cas général

4. Puisque $(x_0, f(x_0), \dots, f^{p-1}(x_0))$ est génératrice, son cardinal est supérieur ou égal à n .
Donc $p \geq n$.
5. Par hypothèse, $f^p(x_0) = x_0$, et donc en appliquant f^k , $\forall k \in \llbracket 0, p-1 \rrbracket$, $f^{k+p}(x_0) = f^k(x_0)$.
Or, la famille $(x_0, f(x_0), \dots, f^{p-1}(x_0))$ est génératrice de E , donc si $x \in E$, il existe des scalaires $\lambda_0, \dots, \lambda_{p-1}$ tels que

$$x = \lambda_0 x_0 + \lambda_1 f(x_0) + \dots + \lambda_{p-1} f^{p-1}(x_0).$$

Et alors

$$f^p(x) = f^p(\lambda_0 x_0 + \lambda_1 f(x_0) + \dots + \lambda_{p-1} f^{p-1}(x_0)) = \sum_{i=0}^{p-1} \lambda_i f^{p+i}(x_0) = \sum_{i=0}^{p-1} \lambda_i f^i(x_0) = x.$$

Donc $\forall x \in E$, $f^p(x) = x$: on a bien $f^p = \text{id}_E$.

On en déduit que f est bijective, et que $f^{-1} = f^{p-1}$.

- 6.a. Par définition, $(x_0, f(x_0), \dots, f^m(x_0))$ est une famille liée. Donc il existe des scalaires $\lambda_0, \dots, \lambda_m \in \mathbb{C}^{m+1}$, non tous nuls, tels que

$$\sum_{i=0}^m \lambda_i f^i(x_0) = 0.$$

Supposons que $\lambda_m = 0$. Alors $\sum_{i=0}^{m-1} \lambda_i f^i(x_0) = 0$, et puisque la famille $(x_0, \dots, f^{m-1}(x_0))$

est libre, cela signifie que $\lambda_0 = \dots = \lambda_{m-1} = 0$. Ceci contredit le fait que les scalaires $\lambda_0, \dots, \lambda_m$ sont non tous nuls.

Donc $\lambda_m \neq 0$, et par conséquent

$$f^m(x_0) = - \sum_{i=0}^{m-1} \frac{\lambda_i}{\lambda_m} f^i(x_0).$$

Ceci prouve donc que $f^m(x_0)$ est combinaison linéaire de $x_0, \dots, f^{m-1}(x_0)$.

- 6.b. Montrons par récurrence que pour tout $k \geq m$, $f^k(x_0)$ est combinaison linéaire de $x_0, f(x_0), \dots, f^{m-1}(x_0)$, c'est-à-dire que $f^k(x_0) \in \text{Vect}(x_0, f(x_0), \dots, f^{m-1}(x_0))$.

La question précédente prouve que le résultat est vrai pour $k = m$.

Supposons que $f^k(x_0)$ est combinaison linéaire de $x_0, \dots, f^{m-1}(x_0)$, et soient $\lambda_0, \dots, \lambda_{m-1}$

tels que $f^k(x_0) = \sum_{i=0}^{m-1} \lambda_i f^i(x_0)$. Alors

$$f^{k+1}(x_0) = \sum_{i=0}^{m-1} \lambda_i f^{i+1}(x_0) = \sum_{i=0}^{m-2} \lambda_i f^{i+1}(x_0) + \lambda_{m-1} f^m(x_0).$$

Mais d'après la question précédente $f^m(x_0) \in \text{Vect}(x_0, \dots, f^{m-1}(x_0))$, et donc $f^{k+1}(x_0) \in \text{Vect}(x_0, \dots, f^{m-1}(x_0))$.

Par le principe de récurrence, on en déduit que

$$\forall k \geq m, f^k(x_0) \in \text{Vect}(x_0, f(x_0), \dots, f^{m-1}(x_0)).$$

Détails

Si deux endomorphismes f et g coïncident sur une base, alors ils ont la même matrice dans cette base. Et donc sont égaux.

Non tous nuls

Rappelons que non tous nuls n'est pas pareil que tous non nuls : cela signifie que l'un au moins de λ_i est non nul, mais les autres peuvent être nuls.

6.c. La famille $(x_0, \dots, f^{m-1}(x_0))$ est libre par hypothèse. On a alors $m \leq n$ (par cardinalité d'une famille libre), et donc par la question a, $m \leq p$. La question précédente prouve que tout élément de la famille $(x_0, \dots, f^{p-1}(x_0))$ est combinaison linéaire de $(x_0, \dots, f^{m-1}(x_0))$.

Mais $(x_0, \dots, f^{p-1}(x_0))$ est génératrice de E , ce qui signifie que tout vecteur de E peut s'écrire comme combinaison linéaire de vecteurs de $(x_0, \dots, f^{p-1}(x_0))$. Et donc tout vecteur de E peut s'écrire comme combinaison linéaire de $(x_0, \dots, f^{m-1}(x_0))$. Par conséquent, $(x_0, \dots, f^{m-1}(x_0))$ est génératrice de E : c'est une base de E .

Puisqu'il s'agit d'une base de E , de cardinal m , c'est donc que $m = n$.

7.a. Puisque g est un polynôme en f , f et g commutent. Et donc f^k et g commutent. Donc

$$g(f^k(x_0)) = f^k(g(x_0)) = f^k\left(\sum_{i=0}^{n-1} a_i f^i(x_0)\right) = f^k(f^n(x_0)) = \boxed{f^{n+k}(x_0)}.$$

En particulier, $\forall k \in \llbracket 0, n-1 \rrbracket$, on a

$$g(f^k(x_0)) = f^n(f^k(x_0)).$$

Autrement dit, les endomorphismes g et f^n coïncident sur la base $(x_0, f(x_0), \dots, f^{n-1}(x_0))$: ils sont donc égaux.

7.b. Puisque $f^n(e_1) = a_0 e_1 + a_1 f(e_1) + \dots + a_{n-1} f^{n-1}(e_1)$, la matrice M est donnée par

$$M = \begin{pmatrix} 0 & 0 & \dots & 0 & a_0 \\ 1 & 0 & & \vdots & a_1 \\ 0 & 1 & \ddots & \vdots & \vdots \\ \vdots & \ddots & \ddots & 0 & \vdots \\ 0 & \dots & 0 & 1 & a_{n-1} \end{pmatrix} \begin{matrix} e_1 \\ f(e_1) \\ \vdots \\ f^{n-2}(e_1) \\ f^{n-1}(e_1) \end{matrix}.$$

7.c. Soit $\lambda \in \mathbb{C}$. La matrice de $f - \lambda \text{id}_E$ dans la base $(x_0, f(x_0), \dots, f^{n-1}(x_0))$ est

$$M - \lambda I_n = \begin{pmatrix} -\lambda & 0 & \dots & 0 & a_0 \\ 1 & -\lambda & & \vdots & a_1 \\ 0 & 1 & \ddots & \vdots & \vdots \\ \vdots & \ddots & \ddots & -\lambda & \vdots \\ 0 & \dots & 0 & 1 & a_{n-1} - \lambda \end{pmatrix}.$$

Calculons son rang par la méthode du pivot :

$$M - \lambda I_n \xleftrightarrow[L_1 \leftrightarrow L_2]{} \begin{pmatrix} 1 & -\lambda & 0 & a_1 \\ -\lambda & 0 & \dots & \vdots & a_0 \\ 0 & 1 & \ddots & \vdots & \vdots \\ \vdots & \ddots & \ddots & -\lambda & \vdots \\ 0 & \dots & 0 & 1 & a_{n-1} - \lambda \end{pmatrix} \xleftrightarrow[L_2 \leftarrow L_2 + \lambda L_1]{} \begin{pmatrix} 1 & -\lambda & 0 & a_1 \\ 0 & -\lambda^2 & \vdots & a_0 + \lambda a_1 \\ 0 & 1 & \vdots & \vdots \\ \vdots & \ddots & \ddots & -\lambda & \vdots \\ 0 & \dots & 0 & 1 & a_{n-1} - \lambda \end{pmatrix}$$

$$\xleftrightarrow[L_3 \leftrightarrow L_2]{} \begin{pmatrix} 1 & -\lambda & 0 & a_1 \\ 0 & 1 & \vdots & a_2 \\ 0 & -\lambda^2 & \vdots & a_0 + \lambda a_1 \\ \vdots & \ddots & \ddots & -\lambda & \vdots \\ 0 & \dots & 0 & 1 & a_{n-1} - \lambda \end{pmatrix} \xleftrightarrow[L_3 + \lambda^2 L_2]{} \dots$$

De proche en proche, en répétant les opérations $L_i \leftrightarrow L_{i+1}$ et $L_{i+1} \leftarrow L_{i+1} + \lambda^i L_i$, on se ramène à une matrice triangulaire supérieure dont les $n-1$ premiers coefficients diagonaux sont non nuls (ils valent tous 1). Ceci signifie donc que le rang de $M - \lambda I_n$ est supérieur ou égal à $n-1$.

On a bien, comme annoncé, $\text{rg}(f - \lambda \text{id}_E) \geq n-1$.

Remarque

Cela signifie que si λ est une valeur propre de f , alors le sous-espace propre associé est de dimension 1.

EXERCICE 3 (EXTRAIT DE EML 2013)

Partie I : Étude d'une fonction f définie par une intégrale.

1. Pour $x > 0$ fixé, la fonction $t \mapsto \frac{e^{-t}}{x+t}$ est continue sur $[0; +\infty[$, donc le seul éventuel problème de convergence est au voisinage de $+\infty$.
Or, pour $t \geq 1$, on a

$$0 \leq \frac{e^{-t}}{x+t} \leq \frac{e^{-t}}{t} \leq e^{-t}.$$

Puisque $\int_1^{+\infty} e^{-t} dt$ converge⁵, il en est de même de $\int_1^{+\infty} \frac{e^{-t}}{x+t} dt$ et donc de $\int_0^{+\infty} \frac{e^{-t}}{x+t} dt$.

⁵ Par exemple car

$$\Gamma(1) = \int_0^{+\infty} e^{-t} dt \text{ converge.}$$

2. D'après la relation de Chasles, on a

$$f(x) = \int_0^1 \frac{e^{-t}}{x+t} dt + \int_1^{+\infty} \frac{e^{-t}}{x+t} dt.$$

Mais la fonction $t \mapsto \frac{e^{-t}}{x+t}$ est positive sur $[1; +\infty[$ et donc $\int_1^{+\infty} \frac{e^{-t}}{x+t} dt \geq 0$.

De plus, pour $t \in]0, 1]$, on a $\frac{e^{-t}}{x+t} \geq \frac{e^{-1}}{x+t}$ et donc, par croissance de l'intégrale, $\int_0^1 \frac{e^{-t}}{x+t} dt \geq \int_0^1 \frac{e^{-1}}{x+t} dt$.

On en déduit que $f(x) \geq \int_0^1 \frac{e^{-1}}{x+t} dt$.

Mais on a alors

$$\int_0^1 \frac{e^{-1}}{x+t} dt = e^{-1} \int_0^1 \frac{dt}{x+t} = e^{-1} [\ln(x+t)]_0^1 = e^{-1} \ln\left(\frac{x+1}{x}\right) = e^{-1} \ln\left(1 + \frac{1}{x}\right).$$

Lorsque $x \rightarrow 0^+$, $\frac{1}{x} \rightarrow +\infty$ et donc $\ln\left(1 + \frac{1}{x}\right) \xrightarrow{x \rightarrow 0^+} +\infty$.

On en déduit donc que $\lim_{x \rightarrow 0^+} f(x) = +\infty$.

3. Pour $x > 0$, on a $e^{-1} \ln\left(1 + \frac{1}{x}\right) > 0$ et donc, d'après l'inégalité de la question précédente, $f(x) > 0$.

Pour tout $t > 0$, on a $\frac{e^{-t}}{t+x} \leq \frac{e^{-t}}{x}$, de sorte que par croissance de l'intégrale,

$$f(x) = \int_0^{+\infty} \frac{e^{-t}}{t+x} dt \leq \int_0^{+\infty} \frac{e^{-t}}{x} dt = \frac{1}{x} \int_0^{+\infty} e^{-t} dt = \frac{1}{x} \Gamma(1) = \frac{1}{x}.$$

Puisque $\lim_{x \rightarrow +\infty} \frac{1}{x} = 0$, on en déduit, par le théorème des gendarmes, que $\lim_{x \rightarrow +\infty} f(x) = 0$.

4. L'intégrale en question n'est autre que $\Gamma(2)$ qui converge d'après le cours.
On a alors, pour $x > 0$,

$$\left| f(x) - \frac{1}{x} \right| = \left| \int_0^{+\infty} \frac{e^{-t}}{x+t} dt - \int_0^{+\infty} \frac{e^{-t}}{x} dt \right| = \left| \int_0^{+\infty} e^{-t} \left(\frac{1}{x+t} - \frac{1}{x} \right) dt \right| = \left| \int_0^{+\infty} \frac{-te^{-t}}{x(x+t)} dt \right|.$$

$$\text{Soit encore } \left| f(x) - \frac{1}{x} \right| = \int_0^{+\infty} \frac{te^{-t}}{x(x+t)} dt.$$

Mais pour $t > 0$, on a $\frac{1}{x(x+t)} \leq \frac{1}{x^2}$ et donc par croissance de l'intégrale⁶ :

$$\left| f(x) - \frac{1}{x} \right| \leq \int_0^{+\infty} \frac{te^{-t}}{x^2} dt = \frac{1}{x^2} \int_0^{+\infty} te^{-t} dt.$$

En particulier, $x \left| f(x) - \frac{1}{x} \right| \leq \frac{1}{x} \int_0^{+\infty} te^{-t} dt \xrightarrow{x \rightarrow +\infty} 0$, de sorte que

$$\lim_{x \rightarrow +\infty} x \left(f(x) - \frac{1}{x} \right) = 0 \iff f(x) - \frac{1}{x} = o_{x \rightarrow +\infty} \left(\frac{1}{x} \right).$$

Et donc

$$f(x) = \frac{1}{x} + \left(f(x) - \frac{1}{x} \right) = \frac{1}{x} + o_{x \rightarrow +\infty} \left(\frac{1}{x} \right) \sim_{x \rightarrow +\infty} \frac{1}{x}.$$

Valeur absolue

Notons qu'en général, la valeur absolue de l'intégrale n'est pas égale à l'intégrale de la valeur absolue !

Ici, l'intégrande est de signe constant (négatif), et donc $\int_0^{+\infty} \frac{-te^{-t}}{x(x+t)} dt$ est négatif, et donc sa valeur absolue est égale à son opposé, c'est-à-dire à $\int_0^{+\infty} \frac{te^{-t}}{x(x+t)} dt$.

⁶ Toutes les intégrales en jeu sont convergentes.

o / ~

Rappelons que $f \sim g$ si et seulement si $f = g + o(g)$.

Partie II : Dérivabilité et expression de la dérivée de f sous forme d'une intégrale.

5.a. La fonction $t \mapsto \frac{e^{-t}}{(x+t)^2}$ est continue sur $[0; +\infty[$, et donc le seul éventuel problème de convergence est au voisinage de $+\infty$.

Or, pour $t \geq 1$, on a $0 \leq \frac{e^{-t}}{(x+t)^2} \leq e^{-t}$, et puisque $\int_1^{+\infty} e^{-t} dt$ converge, il en est de même de $\int_1^{+\infty} \frac{e^{-t}}{(x+t)^2} dt$ et donc de $\int_0^{+\infty} \frac{e^{-t}}{(x+t)^2} dt$.

5.b. Soit $t \geq 0$. On a

$$\begin{aligned} \frac{1}{h} \left(\frac{1}{x+h+t} - \frac{1}{x+t} \right) + \frac{1}{(x+t)^2} &= -\frac{1}{h} \frac{h}{(x+h+t)(x+t)} + \frac{1}{(x+t)^2} \\ &= -\frac{x+t}{(x+h+t)(x+t)^2} + \frac{x+t+h}{(x+t+h)(x+t)^2} \\ &= \frac{h}{(x+h+t)(x+t)^2}. \end{aligned}$$

Puisque $h > -\frac{x}{2}$, on a $x+h > \frac{x}{2}$ et donc, t étant strictement positif,

$$\frac{1}{x+h+t} \leq \frac{1}{x+h} \leq \frac{2}{x}.$$

On en déduit que

$$\left| \frac{1}{h} \left(\frac{1}{x+h+t} - \frac{1}{x+t} \right) + \frac{1}{(x+t)^2} \right| \leq \frac{|h|}{(x+h+t)(x+t)^2} \leq \frac{2|h|}{x} \frac{1}{(x+t)^2} \leq \frac{2|h|}{x} \frac{1}{x^2} \leq \frac{2|h|}{x^3}.$$

5.c. En multipliant l'inégalité précédemment obtenue par e^{-t} on a, pour tout $t > 0$,

$$\left| \frac{1}{h} \left(\frac{e^{-t}}{x+t+h} - \frac{e^{-t}}{x+h} \right) + \frac{e^{-t}}{(x+t)^2} \right| \leq \frac{2e^{-t}|h|}{x^3}.$$

Puisque $\int_0^{+\infty} \frac{e^{-t}}{(x+t)^2} dt$ converge, il vient donc

$$\int_0^{+\infty} \left| \frac{1}{h} \left(\frac{e^{-t}}{x+t+h} - \frac{e^{-t}}{x+h} \right) + \frac{e^{-t}}{(x+t)^2} \right| dt \leq \int_0^{+\infty} \frac{2e^{-t}|h|}{x^3} dt.$$

D'après l'inégalité triangulaire généralisée, on a alors

$$\begin{aligned} \left| \frac{f(x+h) - f(x)}{h} + \int_0^{+\infty} \frac{e^{-t}}{(x+t)^2} dt \right| &= \left| \int_0^{+\infty} \frac{e^{-t}}{h} \left(\frac{1}{x+h+t} - \frac{1}{x+t} \right) dt + \int_0^{+\infty} \frac{e^{-t}}{(x+t)^2} dt \right| \\ &\leq \int_0^{+\infty} \left| \frac{1}{h} \left(\frac{e^{-t}}{x+t+h} - \frac{e^{-t}}{x+h} \right) + \frac{e^{-t}}{(x+t)^2} \right| dt \\ &\leq \frac{2|h|}{x^3} \int_0^{+\infty} e^{-t} dt \\ &\leq \frac{2|h|}{x^3}. \end{aligned}$$

6. Soit $x \in]0; +\infty[$. Lorsque h tend vers 0 dans l'inégalité précédente, on obtient

$$\lim_{h \rightarrow 0} \frac{f(x+h) - f(x)}{h} + \int_0^{+\infty} \frac{e^{-t}}{(x+t)^2} dt = 0 \iff \lim_{h \rightarrow 0} \frac{f(x+h) - f(x)}{h} = - \int_0^{+\infty} \frac{e^{-t}}{(x+t)^2} dt.$$

Ceci prouve donc que f est dérivable en x et que $f'(x) = - \int_0^{+\infty} \frac{e^{-t}}{(x+t)^2} dt$.

Ceci étant vrai pour tout $x > 0$, f est dérivable sur $]0; +\infty[$.

Majoration

N'oublions pas que si on souhaite utiliser une majoration pour prouver la convergence d'une intégrale, il faut absolument vérifier la positivité des fonctions en jeu, cette hypothèse étant indispensable. Par exemple, pour tout $t \geq 1$, on a $-\frac{1}{t} \leq -\frac{1}{t^2}$, et $\int_1^{+\infty} \frac{-1}{t} dt$ diverge alors que $\int_1^{+\infty} \frac{-1}{t^2} dt$ converge.

Inégalité triang.

Rappelons qu'on appelle inégalité triangulaire généralisée le fait suivant : si f est une fonction continue sur $]a, b[$ telle que $\int_a^b f(t) dt$ converge absolument, alors $\left| \int_a^b f(t) dt \right| \leq \int_a^b |f(t)| dt$.

$\Gamma(1)$

On aura reconnu que $\int_0^{+\infty} e^{-t} dt = \Gamma(1) = 0! = 1$.

Dérivabilité

f est dérivable en x car son taux d'accroissement entre $x+h$ et x admet une limite finie lorsque $h \rightarrow 0$. C'est la définition d'une fonction dérivable.

7. Procédons à une intégration par parties sur le segment $[\varepsilon, A]$ en posant $u(t) = e^{-t}$ et $v(t) = -\frac{1}{x+t}$, qui sont deux fonctions de classe \mathcal{C}^1 sur $[\varepsilon, A]$ avec $u'(t) = -e^{-t}$ et $v'(t) = \frac{1}{(x+t)^2}$.
Il vient alors

$$\int_{\varepsilon}^A \frac{e^{-t}}{(x+t)^2} dt = \left[-\frac{e^{-t}}{(x+t)} \right]_{\varepsilon}^A - \int_{\varepsilon}^A \frac{e^{-t}}{(x+t)^2} dt = \boxed{-\frac{e^{-A}}{x+A} + \frac{e^{-\varepsilon}}{x+\varepsilon} - \int_{\varepsilon}^A \frac{e^{-t}}{x+t} dt.}$$

8. Dans l'égalité précédente, faisons tendre ε vers 0. Il vient alors⁷

$$\int_0^A \frac{e^{-t}}{(x+t)^2} dt = -\frac{e^{-A}}{x+A} + \frac{1}{x} - \int_0^A \frac{e^{-t}}{x+t} dt.$$

Puis, en faisant tendre A vers $+\infty$, il vient

$$\int_0^{+\infty} \frac{e^{-t}}{(x+t)^2} dt = \frac{1}{x} - \int_0^{+\infty} \frac{e^{-t}}{x+t} dt \iff -f'(x) = \frac{1}{x} - f(x) \iff \boxed{f'(x) = -\frac{1}{x} + f(x).}$$

9. Puisque f est dérivable sur $]0, +\infty[$, elle y est continue. Et alors f' est également continue $]0, +\infty[$ car somme de la fonction $x \mapsto -\frac{1}{x}$ et de f , qui sont toutes deux continues.
De même, f' est dérivable car somme de fonctions dérivables. Et on a alors, en dérivant la relation de la question 8,

$$\boxed{\forall x \in]0, +\infty[, f''(x) = \frac{1}{x^2} + f'(x).}$$

Et puisque f' est continue, f'' est continue sur $]0, +\infty[$ car somme de fonctions continues. Ceci prouve donc que f est de classe \mathcal{C}^2 sur $]0, +\infty[$, car deux fois dérivable et à dérivée seconde continue.

EXERCICE 4 (EDHEC 1999)

1. **Question préliminaire** : par la formule de Huygens, on a $\text{Cov}(X, Y) = E(XY) - E(X)E(Y)$. Mais X et Y suivant des lois de Bernoulli, on a $E(X) = P(X = 1)$ et $E(Y) = P(Y = 1)$. D'autre part, XY ne prend que les valeurs 0 et 1 et est donc également une variable de Bernoulli, et donc

$$E(XY) = P(XY = 1) = P([X = 1] \cap [Y = 1]).$$

Et donc $\boxed{\text{Cov}(X, Y) = P([X = 1] \cap [Y = 1]) - P(X = 1)P(Y = 1).}$

- 2.a. Puisqu'on retire une boule tous les deux tirages, au bout de $2(n-1)$ tirages, il reste une seule boule dans l'urne, boule qui est retirée lors du $(2n-1)^{\text{ème}}$ tirage.

Ainsi, $\boxed{\text{l'urne est vidée au bout de } 2n-1 \text{ tirages.}}$

- 2.b. Avant le $(2j)^{\text{ème}}$ tirage, $2j-1$ tirages ont déjà eu lieu, dont j tirages d'ordre impair⁸ et $j-1$ tirages d'ordre pair. Ainsi, j boules ont été retirées de l'urne, qui contient donc $n-j$ boules. Puisqu'on ne retire aucune boule lors du $(2j)^{\text{ème}}$ tirage, il reste donc toujours $n-j$ boules avant le $(2j+1)^{\text{ème}}$ tirage.

- 3.a. Lors du premier tirage, il y a une chance sur n d'obtenir la boule noire, et donc $\boxed{P(X_1 = 1) = \frac{1}{n}.}$

Pour calculer $P(X_2 = 1)$, utilisons la formule des probabilités totales avec le système complet d'événements $\{[X_1 = 0], [X_1 = 1]\}$:

$$P(X_2 = 1) = P(X_1 = 0)P_{[X_1=0]}(X_2 = 1) + P(X_1 = 1)P_{[X_1=1]}(X_2 = 1).$$

Or, $P_{[X_1=1]}(X_2 = 1) = 0$, car si la boule noire a été obtenue lors du premier tirage, elle a été enlevée de l'urne, et donc il n'est pas possible de l'obtenir lors du second tirage.

D'autre part, si la boule noire n'a pas été obtenue lors du premier tirage, alors, avant le second tirage, l'urne contient $n-1$ boules, dont une seule noire : $P_{[X_1=0]}(X_2 = 1) = \frac{1}{n-1}$.

Et donc

$$P(X_2 = 1) = P(X_1 = 0)P_{[X_1=0]}(X_2 = 1) = \frac{n-1}{n} \frac{1}{n-1} = \boxed{\frac{1}{n}.}$$

⁷ Car toutes les intégrales qui suivent convergent, garantissant l'existence des limites.

Danger

Attention à ne pas s'arrêter à « f' est dérivable, donc f est deux fois dérivable. »

Une fonction \mathcal{C}^2 est une fonction deux fois dérivable dont la dérivée seconde est continue.

Remarque

Le même raisonnement pourrait se poursuivre pour prouver que f est \mathcal{C}^3 , puis \mathcal{C}^4 , etc. et donc que f est de classe \mathcal{C}^∞ .

⁸ Les tirages 1, 3, ..., $2j-1$.

- 3.b. Si la boule noire est obtenue lors du $(2j + 1)^{\text{ème}}$ tirage, c'est nécessairement qu'elle n'a pas été retirée de l'urne précédemment, et donc que tous les tirages d'ordre impair précédant le $(2j + 1)^{\text{ème}}$ ont donné une boule blanche :

$$[X_{2j+1} = 1] = [X_1 = 0] \cap [X_3 = 0] \cap \dots \cap [X_{2j-1} = 0] \cap [X_{2j+1} = 1].$$

Par la formule des probabilités composées, on a donc

$$\begin{aligned} P(X_{2j+1} = 1) &= P(X_1 = 0)P_{[X_1=0]}(X_3 = 0) \cdots P_{[X_1=0] \cap \dots \cap [X_{2j-3}=0]}(X_{2j-1} = 0)P_{[X_1=0] \cap \dots \cap [X_{2j-1}=0]}(X_{2j+1} = 1) \\ &= \frac{\cancel{n-1} \cancel{n-2}}{n \cancel{n-1}} \cdots \frac{\cancel{n-j}}{\cancel{n-j+1}} \frac{1}{\cancel{n-j}} \\ &= \boxed{\frac{1}{n}}. \end{aligned}$$

De même, on a $[X_{2j} = 1] = [X_1 = 0] \cap [X_3 = 0] \cap \dots \cap [X_{2j-1} = 0] \cap [X_{2j} = 1]$. Et donc

$$\begin{aligned} P(X_{2j} = 1) &= P(X_1 = 0)P_{[X_1=0]}(X_3 = 0) \cdots P_{[X_1=0] \cap \dots \cap [X_{2j-3}=0]}(X_{2j-1} = 0)P_{[X_1=0] \cap \dots \cap [X_{2j-1}=0]}(X_{2j} = 1) \\ &= \frac{\cancel{n-1} \cancel{n-2}}{n \cancel{n-1}} \cdots \frac{\cancel{n-j}}{\cancel{n-j+1}} \frac{1}{\cancel{n-j}} \\ &= \boxed{\frac{1}{n}}. \end{aligned}$$

- 3.c. Puisque les X_k sont des variables aléatoires ne prenant que les valeurs 0 et 1, elles suivent des lois de Bernoulli.

Et puisque pour tout $k \in \llbracket 1, 2n - 1 \rrbracket$, $P(X_k = 1) = \frac{1}{n}$, les X_k suivent donc toutes la loi de

$$\boxed{\text{Bernoulli de paramètre } \frac{1}{n}}.$$

- 4.a. Avant le $(2n - 2)^{\text{ème}}$ tirage, l'urne contient $n - (n - 1) = 1$ seule boule, qui sera remise dans l'urne à l'issue de ce tirage, et donc sera également la seule boule présente dans l'urne au moment du $(2n - 1)^{\text{ème}}$ tirage.

Mais si U_n est réalisé, cela signifie donc que cette boule est noire, et donc la boule noire sera tirée au $(2n - 2)^{\text{ème}}$ tirage, contredisant le fait qu'elle sera tirée pour la première fois au $(2n - 1)^{\text{ème}}$ tirage. Et donc $\boxed{P(U_n) = 0}$.

Soit $j \in \llbracket 1, n \rrbracket$. Alors $U_j = [X_1 = 0] \cap [X_2 = 0] \cap \dots \cap [X_{2j-2} = 0] \cap [X_{2j-1} = 1]$. Par la formule des probabilités composées, on a donc

$$\begin{aligned} P(U_j) &= P(X_1 = 0)P_{[X_1=0]}(X_2 = 0)P_{[X_1=0] \cap [X_2=0]}(X_3 = 0) \cdots P_{[X_1=0] \cap [X_2=0] \cap \dots \cap [X_{2j-2}=0]}(X_{2j-1} = 1) \\ &= \frac{\cancel{n-1} \cancel{n-2} \cancel{n-2} \cancel{n-3} \cancel{n-3}}{n \cancel{n-1} \cancel{n-1} \cancel{n-2} \cancel{n-2}} \cdots \frac{\cancel{n-j+1} \cancel{n-j+1}}{\cancel{n-j+2} \cancel{n-j+2}} \frac{n-j}{n-(j-1)} \frac{1}{n-(j-1)} \\ &= \boxed{\frac{n-j}{n(n-1)}}. \end{aligned}$$

- 4.b. Pour que l'événement $[X = 1]$ soit réalisé, il faut impérativement que la boule noire ne sorte à aucun des tirages d'ordre pair.

En effet, si tel est le cas, elle est remise dans l'urne, puis, étant donné que les boules sont tirées jusqu'à vider l'urne, elle sera nécessairement retirée au moins une fois lors d'un tirage d'ordre impair. Et alors $X \geq 2$.

Autrement dit, $[X = 1] = U_1 \cup U_2 \cup \dots \cup U_n$.

Les U_i étant deux à deux incompatibles par définition, on a donc

$$\begin{aligned} P(X = 1) &= \sum_{j=1}^n P(U_j) = \sum_{j=1}^n \frac{n-j}{n(n-1)} = \frac{1}{n(n-1)} \sum_{j=1}^n (n-j) \\ &= \frac{1}{n(n-1)} \sum_{i=0}^{n-1} i = \frac{1}{n(n-1)} \frac{(n-1)n}{2} = \boxed{\frac{1}{2}}. \end{aligned}$$

- 4.c. Notons que si la boule noire est obtenue lors d'un tirage d'ordre impair, elle est immédiatement retirée, et donc ne peut plus être obtenue lors des tirages suivants.

Remarque
Il se peut que la boule noire soit sortie lors d'un tirage d'ordre pair, mais dans ce cas elle a été remise dans l'urne, ce qui ne change donc rien à la composition de l'urne.

Rappel

Toute variable aléatoire ne prenant que les valeurs 0 et 1 suit une loi de Bernoulli.

Détails

À l'exception du premier et du dernier terme, tous apparaissent deux fois puisque la probabilité d'obtenir une boule blanche est la même au $(2k)^{\text{ème}}$ tirage et au $(2k + 1)^{\text{ème}}$.

Chgt d'indice

$i = n - j$.

Donc pour que $[X = n]$ soit réalisé, il faut que la boule noire soit obtenue lors de $n - 1$ tirages d'ordre pair, et lors d'un seul tirage d'ordre impair⁹.

Mais il n'y a que $n - 1$ tirages d'ordre pair, et donc nécessairement, la boule noire doit être tirée lors de tous les tirages d'ordre pair, ainsi que lors du dernier tirage, tous les autres tirages d'ordre impair donnant une boule blanche.

Autrement dit :

$$[X = n] = [X_1 = 0] \cap [X_2 = 1] \cap [X_3 = 0] \cap [X_4 = 1] \cap \dots \cap [X_{2n-2} = 1] \cap [X_{2n-1} = 1].$$

Et donc, toujours par la formule des probabilités composées,

$$\begin{aligned} P(X = n) &= P(X_1 = 0)P_{[X_1=0]}(X_2 = 1)P_{[X_1=0] \cap [X_2=1]}(X_3 = 0) \dots P_{[X_1=0] \cap [X_2=1] \cap \dots \cap [X_{2n-2}=1]}(X_{2n-1} = 1) \\ &= \frac{n-1}{n} \cdot \frac{1}{n-1} \cdot \frac{n-2}{n-1} \cdot \frac{1}{n-2} \cdot \frac{n-3}{n-2} \cdot \frac{1}{n-3} \dots \frac{1}{2} \cdot \frac{1}{2} \cdot \frac{1}{1} \\ &= \frac{1}{n(n-1) \dots 2 \cdot 1} = \boxed{\frac{1}{n!}}. \end{aligned}$$

5. Puisque X_k vaut 1 si et seulement si la boule noire est apparue lors du $k^{\text{ème}}$ tirage, le nombre de fois où la boule noire est apparue est $\sum_{k=1}^{2n-1} X_k$ et donc

$$X = \sum_{k=1}^{2n-1} X_k.$$

Par linéarité de l'espérance, on a alors

$$E(X) = E\left(\sum_{k=1}^{2n-1} X_k\right) = \sum_{k=1}^{2n-1} E(X_k) = \sum_{k=1}^{2n-1} \frac{1}{n} = \boxed{\frac{2n-1}{n}}.$$

- 6.a. Si la boule noire a été tirée lors du $(2i + 1)^{\text{ème}}$ tirage, puisqu'il s'agit d'un tirage d'ordre impair, elle a alors été retirée de l'urne, et ne peut donc pas être obtenue plus tard :

$$P_{[X_{2i+1}=1]}(X_{2i+j+1} = 1) = 0.$$

- 6.b. Puisque X_{2i+1} et X_{2i+j+1} suivent des lois de Bernoulli, on peut utiliser la question préliminaire :

$$\begin{aligned} \text{Cov}(X_{2i+1}, X_{2i+j+1}) &= P([X_{2i+1} = 1] \cap [X_{2i+j+1} = 1]) - P(X_{2i+1} = 1)P(X_{2i+j+1} = 1) \\ &= P(X_{2i+1} = 1)P_{[X_{2i+1}=1]}(X_{2i+j+1} = 1) - \frac{1}{n^2} \\ &= \boxed{-\frac{1}{n^2}}. \end{aligned}$$

- 7.a. Sachant que $[X_{2i} = 1]$ est réalisé, c'est qu'à l'issue du $(2i)^{\text{ème}}$ tirage, la boule noire est toujours dans l'urne.

Autrement dit, celle-ci contient une boule noire et $n - i - 1$ boules blanches.

Et donc, selon le même raisonnement qu'à la question 3.b, la probabilité qu'au $(2k)^{\text{ème}}$ tirage suivant le $(2i)^{\text{ème}}$, on obtienne une boule noire est de $\frac{1}{n-i}$. Et donc

$$P_{[X_{2i}=1]}(X_{2i+2k} = 1) = \frac{1}{n-i}.$$

De manière détaillée : par définition d'une probabilité conditionnelle

$$P_{[X_{2i}=1]}(X_{2i+2k} = 1) = \frac{P([X_{2i} = 1] \cap [X_{2i+2k} = 1])}{P(X_{2i} = 1)} = nP([X_{2i} = 1] \cap [X_{2i+2k} = 1]).$$

Mais

$$[X_{2i} = 1] \cap [X_{2i+2k} = 1] = [X_1 = 0] \cap [X_3 = 0] \cap \dots \cap [X_{2i-1} = 0] \cap [X_{2i=1}] \cap [X_{2i+1} = 1] \cap \dots \cap [X_{2i+2k-1} = 0] \cap [X_{2i+2k} = 1].$$

⁹ Puisqu'il faut bien qu'à un moment, la boule soit retirée de l'urne.

Détails

Puisqu'on vient de l'obtenir, c'est qu'elle était dans l'urne, et puisqu'il s'agit d'un tirage d'ordre pair, on la remet dans l'urne.

Et donc par la formule des probabilités composées,

$$\begin{aligned}
 P([X_{2i} = 1] \cap [X_{2i+2k} = 1]) &= P(X_1 = 0)P_{[X_1=0]}(X_3 = 0) \cdots P_{[X_1=0] \cap \cdots \cap [X_{2i-1}=0]}(X_{2i} = 1) \cdots P_{[X_1=0] \cap \cdots \cap [X_{2i+2k-1}=0]}(X_{2i+2k} = 1) \\
 &= \frac{n-1}{n} \frac{n-2}{n-1} \cdots \frac{n-(i-1)-1}{n-(i-1)} \frac{1}{n-i} \frac{n-i-1}{n-i} \cdots \frac{n-(i+k-2)-1}{n-(i+k-2)} \frac{1}{n-(i+k-1)} \\
 &= \frac{1}{n(n-i)}.
 \end{aligned}$$

Et donc
$$P_{[X_{2i}=1]}(X_{2i+2k} = 1) = \frac{1}{n-i}.$$

7.b. Le même raisonnement qu'à la question précédente peut être tenu, par analogie avec la question 3.b :

$$P_{[X_{2i}=1]}(X_{2i+2k+1} = 1) = \frac{1}{n-i}.$$

7.c. Nous venons de prouver¹⁰ que pour tout $j \in \llbracket 1, 2n - 2i - 1 \rrbracket$, $P_{[X_{2i}=1]}(X_{2i+j} = 1) = \frac{1}{n-i}$.
Utilisons là encore la question préliminaire :

¹⁰ En distinguant les cas j pair et j impair.

$$\begin{aligned}
 \text{Cov}(X_{2i}, X_{2i+k}) &= P([X_{2i} = 1] \cap [X_{2i+k} = 1]) - P(X_{2i} = 1)P(X_{2i+k} = 1) \\
 &= P(X_{2i} = 1)P_{[X_{2i}=1]}(X_{2i+k} = 1) - \frac{1}{n^2} \\
 &= \frac{1}{n(n-i)} - \frac{1}{n^2} = \frac{n - (n-i)}{n^2(n-i)} = \frac{i}{n(n-i)}.
 \end{aligned}$$

8. Par bilinéarité de la covariance, on a

$$\begin{aligned}
 V(X) &= V\left(\sum_{k=1}^{n-1} X_k\right) = \text{Cov}\left(\sum_{k=1}^{2n-1} X_k, \sum_{i=1}^{2n-1} X_i\right) \\
 &= \sum_{k=1}^{2n-1} \text{Cov}\left(X_k, \sum_{i=1}^{2n-1} X_i\right) && \text{Linéarité à gauche.} \\
 &= \sum_{k=1}^{2n-1} \sum_{i=1}^{2n-1} \text{Cov}(X_k, X_i) && \text{Linéarité à droite.} \\
 &= \sum_{k=1}^{2n-1} \text{Cov}(X_k, X_k) + \sum_{1 \leq k < i \leq 2n-1} \text{Cov}(X_k, X_i) + \sum_{1 \leq i < k \leq 2n-1} \text{Cov}(X_k, X_i) \\
 &= \sum_{k=1}^{2n-1} V(X_k) + \sum_{1 \leq k < i \leq 2n-1} \text{Cov}(X_k, X_i) + \sum_{1 \leq i < k \leq 2n-1} \text{Cov}(X_i, X_k) && \text{Symétrie de la covariance.} \\
 &= \sum_{k=1}^{2n-1} \frac{1}{n} \left(1 - \frac{1}{n}\right) + 2 \sum_{1 \leq i < k \leq 2n-1} \text{Cov}(X_i, X_k) && \text{Les deux sommes sont les mêmes.} \\
 &= \frac{(2n-1)(n-1)}{n^2} + 2 \sum_{1 \leq i < k \leq 2n-1} \text{Cov}(X_i, X_k).
 \end{aligned}$$

Mais, en séparant les cas où i est pair de ceux où i est impair, on obtient

$$\begin{aligned}
 \sum_{1 \leq i < k \leq 2n-1} \text{Cov}(X_i, X_k) &= \sum_{i=1}^{n-1} \sum_{k=2i+1}^{2n-1} \text{Cov}(X_{2i}, X_k) + \sum_{i=0}^{n-1} \sum_{k=2i+2}^{2n-1} \text{Cov}(X_{2i+1}, X_k) \\
 &= \sum_{i=1}^{n-1} \sum_{j=1}^{2n-2i-1} \text{Cov}(X_{2i}, X_{2i+j}) + \sum_{i=0}^{n-1} \sum_{j=1}^{2n-2i-2} \text{Cov}(X_{2i+1}, X_{2i+j+1}) \\
 &= \sum_{i=1}^{n-1} \sum_{j=1}^{2n-2i-1} \frac{i}{n^2(n-i)} + \sum_{i=1}^{n-1} \sum_{j=1}^{2n-2i-2} -\frac{1}{n^2} \\
 &= \sum_{i=1}^{n-1} \frac{(2n-2i-1)i}{n^2(n-i)} - \frac{2}{n^2} \sum_{i=1}^{n-1} (n-i-1)
 \end{aligned}$$

$$\begin{aligned}
&= \frac{1}{n^2} \sum_{i=1}^{n-1} \frac{(n-i)(2i+1) - n}{n-i} - \frac{2}{n^2} \sum_{j=0}^{n-2} j \\
&= \frac{1}{n^2} \sum_{i=1}^{n-1} (2i+1) - \frac{1}{n} \sum_{i=1}^{n-1} \frac{1}{n-i} - \frac{2}{n^2} \frac{(n-2)(n-1)}{2} \\
&= \frac{2}{n^2} \sum_{i=1}^{n-1} i + \frac{n-1}{n^2} - \frac{1}{n} \sum_{j=1}^{n-1} \frac{1}{j} - \frac{(n-1)(n-2)}{n^2} \\
&= \frac{n(n-1)}{n^2} + \frac{n-1}{n^2} - \frac{1}{n} \sum_{j=1}^{n-1} \frac{1}{j} - \frac{(n-1)(n-2)}{n^2} \\
&= \frac{2(n-1)}{n^2} - \frac{1}{n} \sum_{j=1}^{n-1} \frac{1}{j}
\end{aligned}$$

Chgt d'indice

Dans la seconde somme, on a posé

$$j = n - i - 1.$$

Nous pouvons donc enfin conclure :

$$\begin{aligned}
V(X) &= \frac{(2n-1)(n-1)}{n^2} + \frac{2(n-1)}{n^2} - \frac{2}{n} \sum_{j=1}^{n-1} \frac{1}{j} \\
&= \boxed{\frac{(2n+1)(n-1)}{n^2} - \frac{2}{n} \sum_{j=1}^{n-1} \frac{1}{j}}
\end{aligned}$$

DEVOIR SURVEILLÉ 4 (01.12.18)

EXERCICE 1

On admet que toutes les variables aléatoires considérées dans cet exercice sont définies sur le même espace probabilisé (Ω, \mathcal{A}, P) que l'on ne cherchera pas à déterminer.

On considère une variable aléatoire X suivant la loi exponentielle de paramètre $\frac{1}{2}$ et on pose $Y = \sqrt{X}$.

1. On rappelle qu'en SciLab, la commande `grand(1, 1, 'exp', 1/Lambda)` simule une variable aléatoire suivant la loi exponentielle de paramètre λ . Écrire une (ou des) commande(s) SciLab utilisant `grand` et permettant de simuler Y .
2.
 - a. Déterminer la fonction de répartition F_Y de Y .
 - b. En déduire que Y est une variable à densité, et donner une densité f_Y de Y .
3.
 - a. Rappeler la valeur du moment d'ordre 2 d'une variable aléatoire Z suivant la loi normale centrée réduite.
 - b. En déduire que Y a une espérance et donner sa valeur.
4. On pose $U = 1 - e^{-X/2}$.
 - a. Vérifier que $U(\Omega) = [0, 1[$.
 - b. Déterminer la fonction de répartition F_U de U et reconnaître la loi de U .
 - c. Exprimer X en fonction de U , puis en déduire une simulation SciLab de Y utilisant uniquement la fonction `rand`.

EXERCICE 2

Soit n un entier naturel supérieur ou égal à 2.

On note $E = \mathbf{R}_n[X]$ l'espace vectoriel des polynômes à coefficients réels de degré inférieur ou égal à n , et $\mathcal{B} = (1, X, \dots, X^n)$ la base canonique de E .

On note, pour tout polynôme P de E :

$$\varphi(P) = \frac{1}{n}X(1-X)P' + XP.$$

1.
 - a. Montrer que φ est une application linéaire.
 - b. Calculer $\varphi(X^n)$.
 - c. Montrer que φ est un endomorphisme de E .
2. Déterminer la matrice A de φ dans la base \mathcal{B} . Préciser le rang de cette matrice.
3. Donner les valeurs propres de φ , ainsi que la dimension de ses sous-espaces propres.
4.
 - a. L'endomorphisme φ est-il injectif ? Justifier votre réponse.
 - b. Soit P un polynôme non nul de $\text{Ker}(\varphi)$.
Montrer que P admet 1 comme unique racine (dans \mathbf{C}), et que P est de degré n .
 - c. En déduire une base de $\text{Ker}(\varphi)$.
5. On pose, pour tout k de $[[0, n]]$: $P_k = X^k(1-X)^{n-k}$.
 - a. Pour tout k de $[[0, n]]$, calculer $\varphi(P_k)$.
 - b. Montrer que la famille (P_0, P_1, \dots, P_n) est une base de E et expliciter la matrice de φ dans cette base.
 - c. Déterminer les sous-espaces propres de φ .

PROBLÈME

On confond polynôme et application polynomiale de \mathbf{R} dans \mathbf{R} .

On note E l'ensemble des applications $u : \mathbf{R} \rightarrow \mathbf{R}$ continues sur \mathbf{R} et telles que l'intégrale $\int_{-\infty}^{+\infty} (u(x))^2 e^{-x^2} dx$ converge.

On note F le \mathbf{R} -espace vectoriel des applications polynomiales de \mathbf{R} dans \mathbf{R} .

On note, pour tout $n \in \mathbf{N}$, F_n le \mathbf{R} -espace vectoriel des applications polynomiales de \mathbf{R} dans \mathbf{R} de degré inférieur ou égal à n .

Préliminaire : Valeur de l'intégrale de Gauss.

1. En considérant une variable aléatoire suivant une loi normale, justifier :

$$\forall m \in \mathbf{R}, \int_{-\infty}^{+\infty} e^{-(x-m)^2} dx = \sqrt{\pi}.$$

Partie I : Un produit scalaire sur E .

2. Établir, pour tout $(\alpha, \beta) \in [0; +\infty[{}^2$: $\alpha\beta \leq \frac{1}{2}(\alpha^2 + \beta^2)$.
3. En déduire que, pour tout $(u, v) \in E^2$, l'intégrale $\int_{-\infty}^{+\infty} u(x)v(x)e^{-x^2} dx$ converge.

On note $(\cdot|\cdot)$ l'application de E^2 dans \mathbf{R} qui à tout $(u, v) \in E^2$ associe $\frac{1}{\sqrt{\pi}} \int_{-\infty}^{+\infty} u(x)v(x)e^{-x^2} dx$.

On notera la présence du facteur $\frac{1}{\sqrt{\pi}}$.

4. a. Démontrer que E est un \mathbf{R} -espace vectoriel.
b. Montrer que l'application $(\cdot|\cdot)$ est un produit scalaire sur E .
5. Démontrer : $F \subset E$.

On note encore $(\cdot|\cdot)$ la restriction à F ou à F_n , pour $n \in \mathbf{N}$ du produit scalaire $(\cdot|\cdot)$ sur E . On admet que cette restriction est encore un produit scalaire sur F ou sur F_n .

On note $\|\cdot\|$ la norme sur E associée au produit scalaire $(\cdot|\cdot)$, définie, pour tout $u \in E$, par : $\|u\| = \sqrt{(u|u)}$.

Partie II : Polynômes d'Hermite

On note w l'application de \mathbf{R} dans \mathbf{R} , de classe \mathcal{C}^∞ , définie pour tout $x \in \mathbf{R}$ par $w(x) = e^{-x^2}$.
Pour tout $n \in \mathbf{N}$, on note H_n l'application de \mathbf{R} dans \mathbf{R} définie pour tout $x \in \mathbf{R}$ par

$$H_n(x) = (-1)^n e^{x^2} w^{(n)}(x), \text{ où } w^{(n)} \text{ désigne la dérivée } n\text{-ème de } w.$$

En particulier, $H_0(x) = 1$.

6. Calculer, pour tout $x \in \mathbf{R}$, $H_1(x), H_2(x), H_3(x)$.
Faire figurer les calculs sur la copie.
7. a. Montrer, pour tout $n \in \mathbf{N}$ et tout $x \in \mathbf{R}$:

$$H_{n+1}(x) = 2xH_n(x) - H'_n(x).$$

- b. En déduire que, pour tout $n \in \mathbf{N}$, H_n est un polynôme de degré n .
c. Contrôler alors les résultats obtenus question 6 et calculer H_4 .
Faire figurer les calculs sur la copie.
8. Déterminer, pour tout $n \in \mathbf{N}$, le coefficient du terme de plus haut degré de H_n .
9. Montrer, pour tout $n \in \mathbf{N}$ et tout $x \in \mathbf{R}$: $H_n(-x) = (-1)^n H_n(x)$.
Qu'en déduit-on, en terme de parité, pour l'application H_n ?

Partie III : Lien entre le produit scalaire et les polynômes d'Hermite

10. a. Montrer, pour tout $n \in \mathbf{N}^*$ et tout $P \in F$: $(P'|H_{n-1}) = (P|H_n)$, où $(\cdot|\cdot)$ est le produit scalaire défini à la question 5.
À cet effet, on pourra commencer par effectuer une intégration par parties sur un intervalle fermé borné.
b. En déduire, pour tout $n \in \mathbf{N}^*$ et tout $P \in F_{n-1}$: $(P|H_n) = 0$.
c. En déduire que, pour tout $n \in \mathbf{N}$, la famille (H_0, \dots, H_n) est orthogonale dans F .
11. Établir que, pour tout $n \in \mathbf{N}$, la famille (H_0, \dots, H_n) est une base de F_n .
12. Soit $n \in \mathbf{N}$.
a. Montrer : $\|H_n\|^2 = (H_n^{(n)}|H_0)$ où $\|\cdot\|$ est défini question 5.
b. En déduire la valeur de $\|H_n\|$.

Partie IV : Un endomorphisme symétrique.

On note f, g, h les application définies de F dans F , pour tout $P \in F$ par :

$$f(P) = -P'' + 2XP' + P, \quad g(P) = 2XP - P' \quad h(P) = P'.$$

Ainsi, par exemple, pour tout $P \in F$ et tout $x \in \mathbf{R}$: $(g(P))(x) = 2xP(x) - P'(x)$.

13. Montrer que f est un endomorphisme de F .

On admet que g et h sont aussi des endomorphismes de F , et on note Id_F l'application identique de F .

14. a. Établir : $g \circ h = f - \text{Id}_F$ et $h \circ g = f + \text{Id}_F$.
b. En déduire : $f \circ g - g \circ f = 2g$.
15. Montrer que, pour tout $\lambda \in \mathbf{R}$ et tout $P \in F$, si $f(P) = \lambda P$, alors $f(g(P)) = (\lambda + 2)g(P)$.
16. a. Calculer $f(H_0)$.
b. Calculer, pour tout $k \in \mathbf{N}$, $g(H_k)$ et en déduire, pour tout $k \in \mathbf{N}$: $f(H_k) = (2k + 1)H_k$.
17. Soit $n \in \mathbf{N}$.
a. Montrer que F_n est stable par f .
On note f_n l'endomorphisme de F_n défini par : $\forall P \in F_n, f_n(P) = f(P)$.
b. Montrer qu'il existe une base orthonormée de F_n constituée de vecteurs propres de f_n .

CORRECTION DU DEVOIR SURVEILLÉ 4

EXERCICE 1 (EDHEC 2018)

1. Le programme suivant fait l'affaire :

```
1 Y = sqrt(grand(1, 1, 'exp', 2))
```

- 2.a. Une racine carrée étant positive, il est évident que Y ne prend que des valeurs positives, et donc pour $x < 0$, $F_Y(x) = 0$. Soit à présent $x \geq 0$. Alors

$$\begin{aligned} F_Y(x) &= P(Y \leq x) = P(\sqrt{X} \leq x) \\ &= P(X \leq x^2) = 1 - e^{-x^2/2}. \end{aligned}$$

Et donc la fonction de répartition de Y est

$$F_Y : x \mapsto \begin{cases} 0 & \text{si } x < 0 \\ 1 - e^{-x^2/2} & \text{si } x \geq 0 \end{cases}$$

- 2.b. Commençons par prouver que Y est une variable à densité. Il est évident, par opérations sur les fonctions usuelles, que F_Y est de classe \mathcal{C}^1 sauf éventuellement en 0. De plus, on a $\lim_{x \rightarrow 0^-} F_Y(x) = 0 = \lim_{x \rightarrow 0^+} F_Y(x)$, de sorte que F_Y est continue en 0 et donc¹ sur \mathbf{R} .

Ainsi, Y est une variable à densité. Une densité en est alors toute fonction qui coïncide avec F'_Y là où celle-ci est définie, c'est-à-dire sur \mathbf{R}^* . On peut par exemple prendre

$$f_Y : x \mapsto \begin{cases} 0 & \text{si } x \leq 0 \\ xe^{-x^2/2} & \text{si } x > 0 \end{cases}$$

- 3.a. Soit N une variable aléatoire suivant la loi $\mathcal{N}(0, 1)$. Alors, par la formule de Huygens, on a

$$E(N^2) = V(N) + E(N)^2 = 1 + 0^2 = \boxed{1}.$$

- 3.b. Par définition de l'espérance, Y admet une espérance si et seulement si l'intégrale $\int_{-\infty}^{+\infty} t f_Y(t) dt = \int_0^{+\infty} t^2 e^{-t^2/2} dt$ converge.

Mais, une densité de N est $x \mapsto \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$, de sorte que

$$1 = E(N^2) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} t^2 e^{-t^2/2} dt \Leftrightarrow \int_{-\infty}^{+\infty} t^2 e^{-t^2/2} dt = \sqrt{2\pi}.$$

Puisque la fonction $t \mapsto t^2 e^{-t^2/2}$ est paire, on en déduit que

$$\int_0^{+\infty} t^2 e^{-t^2/2} dt = \frac{1}{2} \int_{-\infty}^{+\infty} t^2 e^{-t^2/2} dt = \frac{\sqrt{2\pi}}{2} = \sqrt{\frac{\pi}{2}}.$$

Et donc Y admet une espérance et $E(Y) = \sqrt{\frac{\pi}{2}}$.

- 4.a. Puisque X est à valeurs positives, $e^{-X/2}$ est à valeurs dans $]0, 1]$ et donc U est à valeurs dans $[0, 1[$.

- 4.b. D'après la question précédente, on a $F_U(x) = 0$ si $x < 0$ et $F_U(x) = 1$ si $x \geq 1$. Soit à présent $x \in [0, 1[$. On a alors

$$F_U(x) = P(U \leq x) = P(1 - e^{-X/2} \leq x)$$

Paramètre

Même si l'énoncé le rappelait, on prendra garde au fait que grand attend comme dernier paramètre la valeur de $1/\lambda$ (ici 2), et non de λ .

La fonction carré est croissante sur \mathbf{R}_+ .

¹ Puisque F_Y est \mathcal{C}^1 sur \mathbf{R}^* , elle y est en particulier continue.

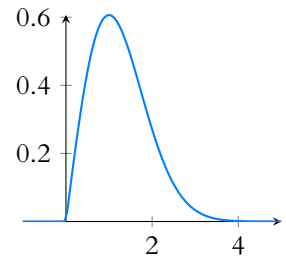


FIGURE 24.1– La densité f_Y .

$$\begin{aligned}
 &= P\left(e^{-X/2} \geq 1 - x\right) = P\left(-\frac{X}{2} \geq \ln(1 - x)\right) \\
 &= P(X \leq -2 \ln(1 - x)) = 1 - e^{\ln(1-x)} = x.
 \end{aligned}$$

La fonction \ln est croissante donc préserve le sens des inégalités.

Donc la fonction de répartition de U est

$$F_U : x \mapsto \begin{cases} 0 & \text{si } x < 0 \\ x & \text{si } 0 \leq x < 1 \\ 1 & \text{si } x \geq 1 \end{cases}$$

On reconnaît la fonction de répartition d'une loi uniforme sur $[0, 1]$, de sorte que $U \hookrightarrow \mathcal{U}([0, 1])$.

- 4.c. On a $e^{-X/2} = 1 - U \Leftrightarrow X = -2 \ln(1 - U)$.
Et donc on peut simuler X à l'aide de la commande suivante :

```
1 X = -2*log(1-rand())
```

EXERCICE 2 (EXTRAIT DE EML 2018)

- 1.a. Soient $P, Q \in \mathbf{R}_n[X]$ et $\lambda \in \mathbf{R}$. Alors, par linéarité de la dérivation,

$$\begin{aligned}
 \varphi(\lambda P + Q) &= \frac{1}{n} X(1 - X)(\lambda P + Q)' + X(\lambda P + Q) \\
 &= \frac{1}{n} X(1 - X)(\lambda P' + Q') + X(\lambda P + Q) \\
 &= \lambda \left(\frac{1}{n} X(1 - X)P' + XP \right) + \left(\frac{1}{n} X(1 - X)Q' + XQ \right) \\
 &= \lambda \varphi(P) + \varphi(Q).
 \end{aligned}$$

Donc φ est une application linéaire de $\mathbf{R}_n[X]$ dans $\mathbf{R}[X]$.

- 1.b. On a

$$\varphi(X^n) = \frac{1}{n} X(1 - X)nX^{n-1} + XX^n = (1 - X)X^n + X^{n+1} = X^n.$$

- 1.c. Notons qu'au premier abord, nous pouvons dire que $\deg P' \leq n - 1$ et donc $\deg X(1 - X)P' \leq n - 1 + 2 = n + 1$.
Et de même $\deg XP \leq n + 1$, de sorte que $\deg \varphi(P) \leq n + 1$.
Tout l'enjeu est donc de prouver que $\varphi(P)$ ne contient pas de terme en X^{n+1} .

Pour $Q \in \mathbf{R}_n[X]$, on a $\deg Q' \leq \deg Q - 1$ et donc

$$\deg(\varphi(Q)) \leq \max(\deg(Q') + \deg(X(1 - X)), \deg(Q) + 1) \leq \deg(Q) + 1.$$

Or, si $P \in \mathbf{R}_n[X]$, alors en séparant le terme de degré n des autres termes, on constate qu'il existe un réel a et un polynôme $Q \in \mathbf{R}_{n-1}[X]$ tels que $P = aX^n + Q$.
Et par linéarité de φ , on a alors

$$\varphi(P) = a\varphi(X^n) + \varphi(Q) = aX^n + \varphi(Q).$$

Or, $\deg \varphi(Q) \leq \deg(Q) + 1 \leq n$, de sorte que $\varphi(P) \in \mathbf{R}_n[X]$.

Ainsi, φ est à valeurs dans E , et étant linéaire, c'est un endomorphisme de E .

2. On a $\varphi(1) = X$, $\varphi(X^n) = X^n$ et pour $k \in \llbracket 1, n - 1 \rrbracket$,

$$\varphi(X^k) = \frac{k}{n} X(1 - X)X^{k-1} + X^{k+1} = \frac{k}{n} X^k - \frac{k}{n} X^{k+1} + X^{k+1} = \frac{k}{n} X^k + \frac{n - k}{n} X^{k+1}.$$

Degré

Notons que pour l'instant, nous ne pouvons rien dire du degré de $\varphi(P)$, même si on prouverait sans grandes difficultés que φ est à valeurs dans $\mathbf{R}_{n+1}[X]$.
Pour ne pas trop nous avancer, nous nous contentons pour l'instant de dire que φ est à valeurs dans $\mathbf{R}[X]$.
Notons que la question 1.c prouvera que φ est à valeurs dans $\mathbf{R}_n[X]$.

Danger !

Tout ce qu'on peut dire en toute généralité sur le degré d'une somme est que

$$\begin{aligned}
 \deg(P + Q) \\
 \leq \max(\deg P, \deg Q).
 \end{aligned}$$

Cette inégalité n'est pas nécessairement une égalité, il se peut que les termes de plus haut degré de P et Q se compensent.

Et donc la matrice A est donnée par

$$A = \text{Mat}_{\mathcal{B}}(\varphi) = \begin{pmatrix} \varphi(1) & \varphi(X) & \varphi(X^2) & \dots & \varphi(X^{n-1}) & \varphi(X^n) \\ 0 & 0 & \dots & \dots & \dots & 0 \\ 1 & \frac{1}{n} & 0 & \dots & \dots & 0 \\ 0 & \frac{n-1}{n} & \frac{2}{n} & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & \frac{n-1}{n} & 0 \\ 0 & \dots & \dots & 0 & 1 & 1 \end{pmatrix} \begin{matrix} 1 \\ X \\ X^2 \\ \vdots \\ X^{n-1} \\ X^n \end{matrix} \in \mathcal{M}_{n+1}(\mathbf{R}).$$

Puisque la première ligne de A est nulle, elle est de rang au plus n .

Puisque d'autre part elle est triangulaire² et possède n coefficients diagonaux non nuls, elle est de rang au moins n .

Et donc $\text{rg}(A) = 1$.

3. La matrice A est triangulaire inférieure, donc ses valeurs propres sont ses coefficients diagonaux : $\text{Spec}(\varphi) = \text{Spec}(A) = \left\{0, \frac{1}{n}, \frac{2}{n}, \dots, 1\right\} = \left\{\frac{k}{n}, k \in \llbracket 0, n \rrbracket\right\}$.

Notons qu'on obtient alors $n+1 = \dim E$ valeurs propres distinctes, donc tous les sous-espaces propres sont de dimension 1.

- 4.a. Puisque A est de rang $n < n+1$, elle n'est pas inversible.

Donc φ n'est pas un isomorphisme, et donc φ n'est pas injectif.

- 4.b. Soit $P \in \text{Ker } \varphi$. Alors $\frac{1}{n}X(1-X)P' = -XP \Leftrightarrow \frac{1}{n}(1-X)P' = -P$.

Soit a une racine complexe de P . Alors $P(a) = 0$ et donc $\frac{1}{n}(1-a)P'(a) = 0$, de sorte que $P'(a) = 0$ ou $a = 1$.

Supposons que $a \neq 1$, de sorte que $P'(a) = 0$. Notons alors $m \geq 1$ la multiplicité de a comme racine de P .

Alors a est racine de P' de multiplicité $m-1$.

Et donc a est racine de $\frac{1}{n}(1-X)P'$ de multiplicité $m-1$.

Mais $\frac{1}{n}(1-X)P' = P$, et donc a en est racine de multiplicité m , d'où une contradiction.

Ainsi, 1 est la seule racine de P .

Soit $P \in \text{Ker}(\varphi)$ non nul, et notons d le degré de P , de sorte que $P = \sum_{k=0}^d a_k X^k$, avec $a_d \neq 0$.

Alors le coefficient en X^{d+1} de $\varphi(P)$ est $-\frac{d}{n}a_d + a_d$. Or ce coefficient est nul puisque $\varphi(P) = 0$, et a_d étant non nul, on a donc $d = n$.

Ainsi, tout polynôme non nul de $\text{Ker}(\varphi)$ est de degré égal à n .

- 4.c. Nous savons que tout polynôme de $\mathbf{C}[X]$ se factorise en produit de termes de degré un. Or, un élément non nul de $\text{Ker}(\varphi)$ est nécessairement de degré n , donc produit de n termes de degré un.

Et puisque 1 en est la seule racine, nécessairement si $P \in \text{Ker}(\varphi)$, alors il existe $\lambda \in \mathbf{C}$ tel que $P = \lambda(X-1)^n$. Et donc $\text{Ker}(\varphi) = \text{Vect}((X-1)^n)$.

- 5.a. Pour $k \in \llbracket 1, n-1 \rrbracket$, on a $P'_k = kX^{k-1}(1-X)^{n-k} - (n-k)X^k(1-X)^{n-k-1}$. Et donc

$$\begin{aligned} \varphi(P_k) &= \frac{1}{n}X(1-X) \left(kX^{k-1}(1-X)^{n-k} - (n-k)X^k(1-X)^{n-k-1} \right) + X^{k+1}(1-X)^{n-k} \\ &= \frac{k}{n}X^k(1-X)^{n-k+1} - \frac{n-k}{n}X^{k+1}(1-X)^{n-k} + X^{k+1}(1-X)^{n-k} \\ &= \frac{k}{n}X^k(1-X)^{n-k} (X+1-X) = \frac{k}{n}P_k. \end{aligned}$$

² Inférieure.

Rappel

En dimension finie, un endomorphisme est injectif si et seulement si il est bijectif.

Multiplicité

Rappelons que la multiplicité d'une racine est le plus grand entier k tel que P soit divisible par $(X-a)^k$. C'est aussi le plus grand entier k tel que $P(a) = P'(a) = \dots = P^{(k-1)}(a) = 0$.

Degré

Lorsqu'on dit d'un polynôme qu'il est de degré d , cela signifie précisément qu'il possède un terme de degré d non nul, et aucun terme de degré plus élevé.

Réels/complexes

Nous avons ici considéré P comme un polynôme complexe. Mais puisque nous savons qu'il s'agit d'un polynôme à coefficients réels, nous pouvons en réalité affirmer que $\lambda \in \mathbf{R}$.

Remarque

Nous traitons à part les cas $k=0$ et $k=n$ pour éviter d'écrire des puissances négatives.

D'autre part, on a $P_0 = (1 - X)^n \in \text{Ker } \varphi$ et donc $\varphi(P_0) = 0$.

Et enfin $P_n = X^n$ et donc $\varphi(P_n) = X^n = P_n$.

5.b. Nous venons de prouver que pour tout $k \in \llbracket 0, n \rrbracket$, P_k est un vecteur propre de φ pour la valeur propre $\frac{k}{n}$.

Donc (P_0, P_1, \dots, P_n) est une famille de vecteurs propres associés à des valeurs propres distinctes : elle est libre. Étant de cardinal $n + 1 = \dim E$, c'est une base de E .

Et puisqu'il s'agit d'une base de vecteurs propres de φ , la matrice de φ dans cette base est diagonale :

$$\text{Mat}_{(P_0, \dots, P_n)}(\varphi) = \begin{pmatrix} \varphi(P_0) & \varphi(P_1) & \dots & \varphi(P_n) \\ 0 & \frac{1}{n} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & \frac{n}{n} \end{pmatrix} \begin{matrix} P_0 \\ P_1 \\ \vdots \\ P_n \end{matrix} = \text{Diag} \left(0, \frac{1}{n}, \frac{2}{n}, \dots, \frac{n-1}{n}, 1 \right).$$

5.c. Nous avons déjà dit que tous les $E_{\frac{k}{n}}(\varphi)$ sont de dimension 1.

Et puisque pour tout $k \in \llbracket 0, n \rrbracket$, P_k est un vecteur propre de φ associé à la valeur propre $\frac{k}{n}$, c'est une base de $E_{\frac{k}{n}}(\varphi)$.

Ainsi, pour tout $k \in \llbracket 0, n \rrbracket$, $E_{\frac{k}{n}}(\varphi) = \text{Vect}(P_k)$.

Dimension 1

Un sous-espace vectoriel de dimension 1 est engendré par n'importe lequel de ses vecteurs non nuls.

PROBLÈME (EML 2008)

Préliminaire : Valeur de l'intégrale de Gauss.

1. La densité d'une loi normale $\mathcal{N} \left(m, \frac{1}{2} \right)$ est $x \mapsto \frac{\sqrt{2}}{\sqrt{2\pi}} e^{-\frac{1}{2}(\sqrt{2})^2(x-m)^2} = \frac{1}{\sqrt{\pi}} e^{-(x-m)^2}$.

Et donc on a

$$\int_{-\infty}^{+\infty} \frac{1}{\sqrt{\pi}} e^{-(x-m)^2} dx = 1 \Leftrightarrow \int_{-\infty}^{+\infty} e^{-(x-m)^2} dx = \sqrt{\pi}.$$

Danger !

On a $\sigma^2 = \frac{1}{2}$, et donc $\sigma = \frac{1}{\sqrt{2}}$.

Partie I : Un produit scalaire sur E .

2. On a $(\alpha - \beta)^2 \geq 0 \Leftrightarrow \alpha^2 - 2\alpha\beta + \beta^2 \geq 0$. Et donc

$$2\alpha\beta \leq \alpha^2 + \beta^2 \Leftrightarrow \alpha\beta \leq \frac{1}{2}(\alpha^2 + \beta^2).$$

3. Soient u et v deux fonctions continues sur \mathbf{R} , telles que $\int_{-\infty}^{+\infty} u(x)^2 e^{-x^2} dx$ et $\int_{-\infty}^{+\infty} v(x)^2 e^{-x^2} dx$

convergent. Alors, pour tout $x \in \mathbf{R}$, on a, d'après l'inégalité de la question 2, $|u(x)v(x)| = |u(x)| \cdot |v(x)| \leq \frac{1}{2} (|u(x)|^2 + |v(x)|^2)$ et donc

$$0 \leq |u(x)v(x)|e^{-x^2} \leq \frac{1}{2}u(x)^2e^{-x^2} + \frac{1}{2}v(x)^2e^{-x^2}.$$

Puisque $\int_{-\infty}^{+\infty} \left(\frac{1}{2}u(x)^2e^{-x^2} + \frac{1}{2}v(x)^2e^{-x^2} \right) dx$ converge³, il en est de même de

$\int_{-\infty}^{+\infty} |u(x)v(x)e^{-x^2}| dx$. Et donc $\int_{-\infty}^{+\infty} u(x)v(x)e^{-x^2} dx$ converge absolument, donc converge.

³ Car somme d'intégrales convergentes.

Signes

Notons que le recours à l'absolue convergence était indispensable ici : u et v ne sont pas nécessairement de signe constant, et la positivité est une hypothèse indispensable pour prouver une convergence à l'aide d'une majoration.

4.a. Montrons que E est un sous-espace vectoriel de l'ensemble des fonctions continues de \mathbf{R} dans \mathbf{R} .

E est donc bien inclus dans l'ensemble des fonctions continues de \mathbf{R} dans \mathbf{R} , il contient bien la fonction nulle car $\int_{-\infty}^{+\infty} 0 dx$ converge.

Soient donc $u, v \in E$ et $\lambda \in \mathbf{R}$. Alors, pour tout $x \in \mathbf{R}$, on a

$$(\lambda u(x) + v(x))^2 e^{-x^2} = \lambda^2 u(x)^2 e^{-x^2} + 2\lambda u(x)v(x)e^{-x^2} + v(x)^2 e^{-x^2}.$$

Et alors $\int_{-\infty}^{+\infty} (\lambda u(x) + v(x))^2 e^{-x^2} dx$ converge car somme d'intégrales convergentes. Et donc $\lambda u + v \in E$.
Ainsi, E est un sous-espace vectoriel de l'ensemble des fonctions continues de \mathbf{R} dans \mathbf{R} .

4.b. Soient $u, v \in E$. Alors

$$(u|v) = \frac{1}{\sqrt{\pi}} \int_{-\infty}^{+\infty} u(x)v(x)e^{-x^2} dx = \frac{1}{\sqrt{\pi}} \int_{-\infty}^{+\infty} v(x)u(x)e^{-x^2} dx = (v|u).$$

Donc $(\cdot|\cdot)$ est symétrique.

Soient $u, v, w \in E$ et $\lambda \in \mathbf{R}$. Alors

$$\begin{aligned} (\lambda u + v|w) &= \frac{1}{\sqrt{\pi}} \int_{-\infty}^{+\infty} (\lambda u(x) + v(x))w(x)e^{-x^2} dx \\ &= \lambda \frac{1}{\sqrt{\pi}} \int_{-\infty}^{+\infty} u(x)w(x)e^{-x^2} dx + \frac{1}{\sqrt{\pi}} \int_{-\infty}^{+\infty} v(x)w(x)e^{-x^2} dx = \lambda(u|w) + (v|w). \end{aligned}$$

Donc $(\cdot|\cdot)$ est linéaire à gauche, et étant symétrique, est bilinéaire.

Si $u \in E$, alors $(u|u) = \frac{1}{\sqrt{\pi}} \int_{-\infty}^{+\infty} u(x)^2 e^{-x^2} dx$.

Or la fonction $x \mapsto u(x)^2 e^{-x^2}$ est positive sur \mathbf{R} , et donc, par positivité de l'intégrale, $(u|u) \geq 0$.

Enfin, puisque $x \mapsto u(x)^2 e^{-x^2}$ est continue et positive sur \mathbf{R} , son intégrale est nulle si et seulement si $\forall x \in \mathbf{R}, u(x)^2 e^{-x^2} = 0$, soit encore⁴ si et seulement si $\forall x \in \mathbf{R}, u(x) = 0$.

Et donc $(u|u) = 0$ si et seulement si $u = 0$.

Ainsi, $(\cdot|\cdot)$ est un produit scalaire sur E .

5. Les fonctions polynomiales sont bien évidemment continues sur \mathbf{R} , donc il s'agit de prouver que si $u \in F$, alors $\int_{-\infty}^{+\infty} u(x)e^{-x^2} dx$ converge.

Soit donc $u : x \mapsto \sum_{k=0}^n a_k x^k$ une fonction polynomiale de degré n .

Alors u est continue sur \mathbf{R} , et au voisinage de $+\infty$, on a $u(x) \sim a_n x^n$. Alors

$$x^2 u(x) e^{-x^2} \underset{x \rightarrow +\infty}{\sim} a_n x^{n+2} e^{-x^2}.$$

Mais puisque $e^{-x^2} = o_{x \rightarrow +\infty}(e^{-x})$, on a $x^{n+2} e^{-x^2} = o_{x \rightarrow +\infty}(x^{n+2} e^{-x})$.

Mais, par croissances comparées, $x^{n+2} e^{-x} \xrightarrow{x \rightarrow +\infty} 0$, de sorte que $x^{n+2} e^{-x^2} \xrightarrow{x \rightarrow +\infty} 0$. Et donc

$$u(x)e^{-x^2} = o_{x \rightarrow +\infty}\left(\frac{1}{x^2}\right).$$

Puisque $\int_1^{+\infty} \frac{dx}{x^2}$ est une intégrale absolument convergente, il en est de même de $\int_1^{+\infty} u(x)e^{-x^2} dx$.

De même, on prouve que $u(x)e^{-x^2} = o_{x \rightarrow -\infty}\left(\frac{1}{x^2}\right)$, et puisque $\int_{-\infty}^{-1} \frac{dx}{x^2}$ converge, il en est

de même de $\int_{-\infty}^{-1} u(x)e^{-x^2} dx$.

Et donc $\int_{-\infty}^{+\infty} u(x)e^{-x^2} dx$ converge. Ainsi, on a bien $F \subset E$.

Partie II : Polynômes d'Hermite.

6. On a, par dérivations successives,

$$w'(x) = -2xe^{-x^2}, w''(x) = e^{-x^2}(-2 + 4x^2), w^{(3)}(x) = e^{-x^2}(12x - 8x^3).$$

Et donc

$$H_1(x) = -e^{-x^2} e^{-x^2} (-2x) = 2x, H_2(x) = e^{-x^2} e^{-x^2} (4x^2 - 2) = 4x^2 - 2, H_3(x) = -e^{-x^2} e^{-x^2} (12x - 8x^3) = 8x^3 - 12x.$$

⁴ Une exponentielle n'est jamais nulle.

Degré

Rappelons que si un polynôme est de degré n , alors son coefficient de degré n est non nul. Ici on a donc $a_n \neq 0$.

Rédaction

La rédaction peut sembler fastidieuse, et on serait tentés de dire que «par croissances comparées, $x^{n+2} e^{-x^2} \xrightarrow{n \rightarrow +\infty} 0$ ». Mais il ne s'agit pas d'une des croissances comparées usuelles du cours, alors que c'est le cas de $x^{n+2} e^{-x}$.

Convergence

Notons qu'il n'y a pas besoin d'une étude de convergence de l'intégrale entre -1 et 1 , car il s'agit de l'intégrale d'une fonction continue sur un segment, elle converge donc.

7.a. On a $H_n(x) = (-1)^n e^{x^2} w^{(n)}(x)$. Et donc,

$$H'_n(x) = (-1)^n e^{x^2} w^{(n+1)}(x) + (-1)^n 2xe^{x^2} w^{(n)}(x) = -H_{n+1}(x) + 2xH_n(x).$$

Et donc $H_{n+1}(x) = 2xH_n(x) - H'_n(x)$.

7.b. Montrons par récurrence que H_n est un polynôme de degré n .

Pour $n = 0$, $H_0 = 1$ est bien un polynôme de degré 0.

Supposons donc que H_n soit un polynôme de degré n .

Alors $H_{n+1}(x) = 2xH_n(x) - H'_n(x)$ est un polynôme car H'_n est un polynôme, et $2xH_n(x)$ également.

De plus, $2xH_n(x)$ est de degré $n + 1$ et H'_n est de degré $n - 1$, donc H_{n+1} est de degré $n + 1$.

D'après le principe de récurrence, pour tout $n \in \mathbf{N}$, H_n est un polynôme de degré n .

7.c. Les fonctions H_1, H_2, H_3 obtenues à la question 6 sont bien des polynômes de degrés respectifs 1, 2 et 3. De plus, on a $2xH_0(x) - H'_0(x) = 2x = H_1(x)$ et

$$2xH_1(x) - H'_1(x) = 4x^2 - 2 = H_2(x) \text{ et } 2xH_2(x) - H'_2(x) = 8x^3 - 4x - 8x = 8x^3 - 12x = H_3(x).$$

Il vient alors

$$H_4(x) = 2xH_3(x) - H'_3(x) = 16x^4 - 24x^2 - 24x^2 + 8 = 16x^4 - 48x^2 + 8.$$

8. Les formules obtenues pour H_0, H_1, H_2, H_3 et H_4 laissent à penser que le coefficient de degré n de H_n est 2^n . Prouvons-le par récurrence.

Pour $n \leq 4$, ceci est vérifié, donc la récurrence est⁵ initialisée.

Supposons que le coefficient dominant de H_n est 2^n . Alors

$$H_{n+1}(x) = 2xH_n(x) - \underbrace{H'_n(x)}_{\in \mathbf{R}_{n-1}[X]},$$

de sorte que le coefficient de degré $n + 1$ de H_{n+1} est celui de $2xH_n(x)$, soit $2 \times 2^n = 2^{n+1}$.

Par le principe de récurrence, pour tout $n \in \mathbf{N}$, le coefficient dominant de H_n est 2^n .

9. Encore une fois, procédons par récurrence. Pour $n = 0$, on a bien $H_0(-x) = 1 = (-1)^0 H_0(x)$. Supposons que pour tout $x \in \mathbf{R}$, $H_n(-x) = (-1)^n H_n(x)$.

Alors

$$H_{n+1}(-x) = -2xH_n(-x) - H'_n(-x) = (-1)^{n+1} 2xH_n(x) - H'_n(-x).$$

Mais en dérivant $H_n(-x) = (-1)^n H_n(x)$, il vient $-H'_n(-x) = (-1)^n H'_n(x)$, soit encore $H'_n(-x) = (-1)^{n+1} H'_n(x)$. Et donc

$$H_{n+1}(-x) = (-1)^{n+1} 2xH_n(x) - (-1)^{n+1} H'_n(x) = (-1)^{n+1} (2xH_n(x) - H'_n(x)) = (-1)^{n+1} H_{n+1}(x).$$

D'après le principe de récurrence, on en déduit que pour tout $n \in \mathbf{N}$ et pour tout $x \in \mathbf{R}$,

$$H_n(-x) = (-1)^n H_n(x).$$

On en déduit que si n est un entier pair, H_n est une fonction paire, et si n est impair, H_n est impaire.

Partie III : Lien entre le produit scalaire et les polynômes d'Hermite.

10.a. On a, pour tout $n \in \mathbf{N}^*$, et pour $P \in F$,

$$(P'|H_{n-1}) = \frac{1}{\sqrt{\pi}} \int_{-\infty}^{+\infty} P'(x)H_{n-1}(x)e^{-x^2} dx.$$

Soit donc $[A, B]$ un segment inclus dans \mathbf{R} , et procédons à une intégration par parties sur ce segment en posant $u(x) = P(x)$ et $v(x) = H_{n-1}(x)e^{-x^2}$, qui sont deux fonctions \mathcal{C}^1 sur $[A, B]$, avec $u'(x) = P'(x)$ et $v'(x) = H'_{n-1}(x)e^{-x^2} - 2xH_{n-1}(x)e^{-x^2}$. Alors, il vient

$$\begin{aligned} \int_A^B P'(x)H_{n-1}(x)e^{-x^2} dx &= [P(x)H_{n-1}(x)e^{-x^2}]_A^B - \int_A^B P(x)(H'_{n-1}(x) - 2xH_{n-1}(x))e^{-x^2} dx \\ &= P(B)H_{n-1}(B)e^{-B^2} - P(A)H_{n-1}(A)e^{-A^2} + \int_A^B H_n(x)e^{-x^2} dx. \end{aligned}$$

Rappel

Si P et Q sont de degré distincts, alors $P + Q$ est de degré égal à $\max(\deg P, \deg Q)$.

Si P et Q ont même degré, alors on a juste une inégalité, les termes de plus haut degré de P et Q pouvant se compenser, faisant de $P + Q$ un polynôme de degré strictement inférieur à $\deg(P)$.

En revanche, si $\deg(P) \neq \deg(Q)$, c'est toujours une égalité, le polynôme de plus haut degré «imposant» son degré à la somme.

⁵ Largement !

On a utilisé la relation de la question 7.a :

$$H'_{n-1}(x) - 2xH_{n-1}(x) = -H_n(x).$$

Notons $a_k X^k$ le terme de plus haut degré de P , de sorte que, au voisinage de $\pm\infty$, $P(x) \underset{x \rightarrow \pm\infty}{\sim} a_k x^k$ et donc $P(x)H_{n-1}(x) \underset{x \rightarrow \pm\infty}{\sim} 2^{n-1} a_k x^{n-1+k}$.

Alors, en posant $t = x^2$, il vient

$$P(x)H_{n-1}(x)e^{-x^2} \underset{x \rightarrow \pm\infty}{\sim} 2^{n-1} a_k x^{n-1+k} e^{-x^2} = 2^{n-1} a_k t^{\frac{n-1+k}{2}} e^{-t} \xrightarrow[t \rightarrow +\infty]{} 0$$

par croissances comparées.

Et donc $\lim_{A \rightarrow -\infty} P(A)H_{n-1}(A)e^{-A^2} = \lim_{B \rightarrow +\infty} P(B)H_{n-1}(B)e^{-B^2} = 0$.

Ainsi, en prenant successivement la limite lorsque $A \rightarrow -\infty$, puis $B \rightarrow +\infty$, il vient

$$\int_{-\infty}^B P'(x)H_{n-1}(x) dx = -P(B)H_{n-1}(B)e^{-B^2} + \int_{-\infty}^B P(x)H_n(x) dx$$

puis

$$\int_{-\infty}^{+\infty} P'(x)H_{n-1}(x) dx = \int_{-\infty}^{+\infty} P(x)H_n(x) dx.$$

En multipliant par $\frac{1}{\sqrt{\pi}}$, il vient alors $(P'|H_{n-1}) = (P|H_n)$.

10.b. On a alors $(P|H_n) = (P'|H_{n-1}) = (P''|H_{n-2}) = \dots = (P^{(n)}|H_0)$.

En particulier, si $P \in F_{n-1}$, alors $P^{(n)} = 0$, et donc $(P^{(n)}|H_0) = 0$. Ainsi, $(P|H_n) = 0$.

10.c. Il s'agit de montrer que pour $i \neq j$, on a $(H_i|H_j) = 0$.

Supposons que $i > j$. Alors H_j est de degré $j < i$, et donc par la question 10.b, $(H_j|H_i) = 0$. De même si $j > i$, alors H_i est de degré strictement inférieur à i , et donc $(H_i|H_j) = 0$.

Ainsi, la famille (H_0, \dots, H_n) est une famille orthogonale.

11. (H_0, \dots, H_n) est une famille orthogonale ne contenant pas le vecteur nul⁶, elle est donc libre.

De plus, elle est de cardinal $n + 1 = \dim \mathbf{R}_n[X] = \dim F_n$. C'est donc une base de F_n .

12.a. On a $\|H_n\|^2 = (H_n|H_n)$ et donc, en utilisant n fois le résultat de la question 10.a, il vient

$$\|H_n\|^2 = (H_n|H_n) = (H_n'|H_{n-1}) = (H_n''|H_{n-2}) = \dots = (H_n^{(n)}|H_0).$$

12.b. Rappelons que $H_n(x) = 2^n x^n + Q_n(x)$, où Q_n est un polynôme de degré au plus $n - 1$.

En particulier, $Q_n^{(n)}(x) = 0$. Et on a donc

$$H_n'(x) = n2^n x^{n-1} + Q_n'(x), H_n''(x) = n(n-1)2^n + Q_n''(x)x^{n-2}, \dots, H_n^{(n)}(x) = n!2^n + \underbrace{Q_n^{(n)}(x)}_{=0}.$$

Ainsi,

$$\|H_n\|^2 = (n!2^n|1) = n!2^n(1|1) = \frac{n!2^n}{\sqrt{\pi}} \int_{-\infty}^{+\infty} e^{-x^2} dx = \frac{n!2^n}{\sqrt{\pi}} \sqrt{\pi} = n!2^n.$$

Et donc $\|H_n\| = \sqrt{2^n n!}$.

Partie IV : Un endomorphisme symétrique.

13. Si $P \in F$, alors P' et P'' sont encore des polynômes, de sorte que $f(P) = -P'' + 2XP' + P \in F$. Soient $P, Q \in F$ et $\lambda \in \mathbf{R}$. Alors

$$f(\lambda P + Q) = -(\lambda P + Q)'' + 2X(\lambda P + Q)' + (\lambda P + Q) = \lambda(-P'' + 2XP' + P) + (-Q'' + 2XQ' + Q) = \lambda f(P) + f(Q).$$

Ainsi, f est linéaire et donc est un endomorphisme de F .

14.a. Soit $P \in F$. Alors

$$(g \circ h)(P) = g(h(P)) = g(P') = 2XP' - P'' = (-P'' + 2XP' + P) - P = f(P) - \text{Id}_F(P).$$

Ceci étant vrai pour tout $P \in F$, on a donc $g \circ h = f - \text{Id}_F$.

De même, pour $P \in F$, on a

$$(h \circ g)(P) = h(g(P)) = (2XP - P')' = 2XP' + 2P - P'' = (-P'' + 2XP' + P) + P = f(P) + \text{Id}_F(P).$$

Et donc $h \circ g = f + \text{Id}_F$.

Coeff. dominant

Ne pas oublier le coefficient dominant lorsqu'on prend un équivalent en $\pm\infty$ d'un polynôme. Par exemple, $2x^2 + x + 1$ est équivalent à $2x^2$, mais pas à x^2 .

⁶ H_i possède $2^i \neq 0$ comme coefficient dominant, donc ne peut être égal au polynôme nul.

14.b. En utilisant les relations précédentes, on a

$$\begin{aligned} f \circ g - g \circ f &= (g \circ h + \text{Id}_F) \circ g - g \circ (h \circ g - \text{Id}_F) \\ &= g \circ h \circ g + g - g \circ h \circ g + g \\ &= \boxed{2g}. \end{aligned}$$

15. Puisque $f \circ g = g \circ f + 2g$, si $f(P) = \lambda P$, alors

$$f(g(P)) = g(f(P)) + 2g(P) = g(\lambda P) + 2g(P) = \lambda g(P) + 2g(P) = \boxed{(\lambda + 2)g(P)}.$$

16.a. Puisque $H_0 = 1$, $H'_0 = H''_0 = 0$. Et donc $f(H_0) = H_0 = 1$.

16.b. En utilisant le résultat de la question 7.a, on a $g(H_k) = 2XH_k - H'_k = H_{k+1}$.

Montrons alors par récurrence sur $k \in \mathbf{N}$ que $f(H_k) = (2k + 1)H_k$.

Pour $k = 0$, c'est le résultat de la question 16.a.

Supposons que $f(H_k) = (2k + 1)H_k$.

Alors $f(H_{k+1}) = f(g(H_k))$, et d'après la question 15, avec $\lambda = 2k + 1$, on a

$$f(H_{k+1}) = g(H_k) = ((2k + 1) + 2)g(H_k) = (2k + 3)H_{k+1} = (2(k + 1) + 1)H_{k+1}.$$

Ainsi, la relation est vérifiée au rang $k + 1$.

Par le principe de récurrence, $\boxed{\text{pour tout } k \in \mathbf{N}, f(H_k) = (2k + 1)H_k.}$

17.a. Soit $P \in F_n$. Alors $\deg P' \leq n - 1$ et $\deg P'' \leq n - 2$, de sorte que $\deg f(P) = \deg(-P'' + 2XP' + P) \leq n$.

Et donc $f(P) \in F_n$: $\boxed{F_n \text{ est stable par } f.}$

17.b. Nous avons prouvé à la question 16.b que H_0, H_1, \dots, H_n sont des vecteurs propres de f , associés à des valeurs propres distinctes.

Puisque de plus $\deg H_i = i$, H_0, \dots, H_n sont dans F_n , et donc sont des vecteurs propres de f_n .

D'autre part, nous avons prouvé dans les questions 10.c et 11 qu'ils forment une base orthogonale de F_n .

Enfin, puisque $\|H_n\| = \sqrt{2^n n!}$, la famille $\left(\frac{H_0}{\|H_0\|}, \frac{H_1}{\|H_1\|}, \dots, \frac{H_n}{\|H_n\|} \right) = \left(H_0, \frac{H_1}{\sqrt{2}}, \dots, \frac{H_n}{\sqrt{2^n n!}} \right)$

est une famille⁷ orthonormée de F_n .

C'est donc une base orthonormée de F_n formée de vecteurs propres.

Alternative

Si on n'a pas vu comment utiliser le résultat de la question précédente, il est tout de même possible de calculer directement

$$(f \circ g - g \circ f)(P)$$

et de constater qu'il s'agit de $2g(P)$, mais cette méthode est plus laborieuse.

⁷ Et même une base d'après la question 11.

Rappel

Puisque les sous-espaces propres sont des sous-espaces vectoriels, si on multiplie un vecteur propre par une constante non nulle, on obtient encore un vecteur propre.

DEVOIR SURVEILLÉ 5 (12.01.19)

EXERCICE

On se place dans un espace euclidien E de dimension n , où n désigne un entier naturel supérieur ou égal à 2.

Le produit scalaire des vecteurs x et y est noté $(x|y)$ et la norme de x est notée $\|x\|$.

On désigne par Id l'endomorphisme identique de E .

On considère un vecteur u de E dont la norme est égale à 1, un réel λ non nul et on note f_λ l'application qui, à tout vecteur x de E associe $f_\lambda(x) = \lambda(x|u)u + x$.

1. Donner la dimension de $(\text{Vect}(u))^\perp$.
2. Montrer que f_λ est un endomorphisme de E .
3. Montrer que le polynôme $X^2 - (\lambda + 2)X + (\lambda + 1)$ est un polynôme annulateur de f_λ .
4.
 - a. Déterminer $f_\lambda(u)$ et $f_\lambda(v)$, pour tout vecteur v de $(\text{Vect}(u))^\perp$.
 - b. Établir alors que f_λ possède deux valeurs propres distinctes et donner les sous-espaces propres associés à ces deux valeurs propres. L'endomorphisme f_λ est-il diagonalisable ?
5. Dans cette question on suppose que $\lambda = -1$.
 - a. Vérifier que f_{-1} est un projecteur.
 - b. Montrer plus précisément que f_{-1} est le projecteur orthogonal sur $(\text{Vect}(u))^\perp$.

PROBLÈME 1

Partie 1

Pour tout entier naturel n , on pose $u_n = \int_0^{\pi/2} (\cos t)^n dt$.

1.
 - a. Calculer u_0 et u_1 .
 - b. Montrer que la suite (u_n) est décroissante.
 - c. Établir que $\forall n \in \mathbf{N}, u_n > 0$.
2.
 - a. Montrer que la suite (u_n) converge. Dans la suite, on note $\ell = \lim_{n \rightarrow +\infty} u_n$.
 - b. Montrer que pour tout $a \in \left]0, \frac{\pi}{2}\right]$, on a

$$0 \leq u_n \leq \int_0^a (\cos t)^n dt + \int_a^{\pi/2} (\cos t)^n dt \leq a + \left(\frac{\pi}{2} - a\right) (\cos a)^n.$$

- c. En déduire que pour tout $a \in \left]0, \frac{\pi}{2}\right]$, $0 \leq \ell \leq a$.
 - d. Montrer que $\ell = 0$.
3.
 - a. Montrer, grâce à une intégration par parties, que $\forall n \in \mathbf{N}, (n+2)u_{n+2} = (n+1)u_n$.
 - b. Montrer que $\forall n \in \mathbf{N}, (n+1)u_{n+1}u_n = \frac{\pi}{2}$.
 - c. Utiliser la question précédente pour compléter les commandes Scilab suivantes afin qu'elles permettent de calculer u_n lorsque n est entré par l'utilisateur.

```

1 n = input('entrez la valeur de n : '),
2 u = %pi/2
3 for ----
4 -----
5 end
6 disp(u)

```

 - d. Écrire une fonction Scilab d'entête `function N = limite(epsilon)` telle que si on donne une valeur de ϵ , la fonction `limite(epsilon)` retourne le plus petit entier N tel que $u_N < \epsilon$.

Partie 2

On note f la fonction définie pour tout réel x par : $f(x) = \begin{cases} \sin(x) & \text{si } 0 \leq x \leq \pi/2 \\ 0 & \text{sinon} \end{cases}$

4. Vérifier que f est une densité de probabilité. Dans la suite, on considère une variable aléatoire réelle X définie sur un certain espace probabilisé (Ω, \mathcal{A}, P) , et ayant f pour densité.
5. Déterminer la fonction de répartition F de la variable aléatoire X .
6. a. Montrer que X possède une espérance et la calculer.
b. Montrer que X possède également une variance et la calculer.
7. On considère maintenant une suite $(X_n)_{n \in \mathbb{N}^*}$ de variables aléatoires toutes définies sur (Ω, \mathcal{A}, P) , mutuellement indépendantes et qui suivent toutes la même loi que X .
Pour tout entier naturel n supérieur ou égal à 2, on pose $I_n = \min(X_1, X_2, \dots, X_n)$ et on admet que I_n est une variable aléatoire, elle aussi définie sur (Ω, \mathcal{A}, P) .

- a. Déterminer la fonction de répartition, notée F_n de la variable aléatoire I_n .
En déduire que I_n est une variable à densité.
- b. Déterminer une densité de I_n .
- c. Montrer que I_n admet une espérance, et que $E(I_n) = u_n$.
- d. Montrer que I_n possède un moment d'ordre 2 :

$$E(I_n^2) = 2 \int_0^{\pi/2} x(\cos x)^n dx.$$

- e. Établir que : $E(I_n^2) \leq \pi u_n$. En déduire $\lim_{n \rightarrow +\infty} V(I_n)$.
8. Soit h la restriction de la fonction cosinus à $\left[0, \frac{\pi}{2}\right]$.
 - a. Montrer que h réalise une bijection de $\left[0, \frac{\pi}{2}\right]$ sur $[0, 1]$.
 - b. Justifier que l'on peut poser $Y = h(X)$. On admet alors que Y est une variable aléatoire, elle aussi définie sur (Ω, \mathcal{A}, P) . Déterminer la fonction de répartition G de Y , puis vérifier que Y suit une loi uniforme.
 - c. On rappelle que la commande `grand(1, 1, 'unf', a, b)` renvoie une simulation Sci Lab d'une variable aléatoire à densité suivant une loi uniforme sur $[a, b]$ et on admet que la fonction h^{-1} s'obtient par l'instruction `acos`. Compléter les commandes Sci Lab suivantes afin qu'elles permettent de simuler la variable aléatoire X .

```
1 Y = grand(1, 1, 'unf', ---, ---)
2 X = ---
```

- d. Compléter la fonction suivante pour qu'elle simule une réalisation de la variable I_n .

```
1 function y=I(n)
2     X = zeros(n, 1) ;
3     for i=...
4         X(i) = ...
5     end
6     y = ...
7 endfunction
```

En utilisant le résultat de la question 7.c, proposer un programme Sci Lab qui calcule une valeur approchée de u_{2019} .

PROBLÈME 2

Partie I : Étude d'un exemple

On considère, dans cette partie, les matrices de $\mathcal{M}_3(\mathbb{R})$: $A = \begin{pmatrix} 4 & 0 & 0 \\ -5 & 9 & 0 \\ -5 & 5 & 4 \end{pmatrix}$ et $I_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$.

1. Trouver, en fonction de I_3 et A , deux matrices P_1 et P_2 de $\mathcal{M}_3(\mathbb{R})$ telles que :

$$P_1 + P_2 = I_3 \text{ et } 4P_1 + 9P_2 = A.$$

Expliciter ensuite les coefficients de P_1 et ceux de P_2 .

2. a. Calculer les matrices P_1^2, P_1P_2, P_2P_1 et P_2^2 .
b. En déduire : $\forall k \in \mathbf{N}, A^k = 4^k P_1 + 9^k P_2$.
3. Trouver au moins une matrice B de $\mathcal{M}_3(\mathbf{R})$ dont on explicitera les coefficients, telle que $B^2 = A$.
4. Quelles sont les valeurs propres de A ?

Dans toute la suite du problème, E désigne un \mathbf{R} -espace vectoriel de dimension finie supérieure ou égale à 1 et f un endomorphisme de E .

On note e l'endomorphisme identité de E qui, à chaque élément de E associe lui-même, et $\tilde{0}$ l'endomorphisme nul de E qui, à chaque élément de E , associe l'élément nul de E .

On suppose qu'il existe un entier m de \mathbf{N}^* , des réels $\lambda_1, \dots, \lambda_m$ deux à deux distincts et des endomorphismes p_1, \dots, p_m de E tous différents de $\tilde{0}$, tels que : $\forall k \in \llbracket 0, m \rrbracket, f^k = \sum_{i=1}^m \lambda_i^k p_i$.

Enfin, on considère les polynômes

$$N = \prod_{\ell=1}^m (X - \lambda_\ell), \text{ et pour tout } i \text{ de } \llbracket 1, m \rrbracket, M_i = \prod_{\substack{1 \leq \ell \leq m \\ \ell \neq i}} (X - \lambda_\ell) \text{ et } L_i = \frac{1}{M_i(\lambda_i)} M_i.$$

On admet que, pour tous polynômes P et Q de $\mathbf{R}[X]$: $(P \times Q)(f) = P(f) \circ Q(f)$.

Partie II : Étude des puissances de f .

5. Montrer, pour tout polynôme P de $\mathbf{R}_m[X]$: $P(f) = \sum_{i=1}^m P(\lambda_i) p_i$.
6. En déduire $N(f) = \tilde{0}$.
7. a. Montrer que, pour tout couple (i, j) de $\llbracket 1, m \rrbracket^2$, $L_i(\lambda_j)$ est égal à 1 si $i = j$ et à 0 si $i \neq j$.
b. En déduire, pour tout i de $\llbracket 1, m \rrbracket$: $L_i(f) = p_i$.
8. a. Montrer : $e = \sum_{i=1}^m p_i$.
b. En déduire que E est la somme des m sous-espaces vectoriels $\text{Im}(p_1), \dots, \text{Im}(p_m)$.
9. Soit i appartenant à $\llbracket 1, m \rrbracket$.
a. Vérifier : $N = M_i(\lambda_i)(X - \lambda_i)L_i$.
b. En déduire, en utilisant le résultat de la question 6 : $\text{Im}(p_i) \subset \text{Ker}(f - \lambda_i e)$.
10. Déduire des questions précédentes que f est diagonalisable, que les valeurs propres de f sont les réels $\lambda_1, \dots, \lambda_m$ et que, pour tout i de $\llbracket 1, m \rrbracket$, le sous-espace propre de f associé à la valeur propre λ_i est $\text{Im}(p_i)$.
11. a. Montrer, pour tout couple (i, j) de $\llbracket 1, m \rrbracket^2$ tel que $i \neq j$: $p_i \circ p_j = \tilde{0}$.
b. En déduire, en utilisant le résultat de la question 8.a, pour tout i de $\llbracket 1, m \rrbracket$: $p_i \circ p_i = p_i$.
c. Établir, pour tout i de $\llbracket 1, m \rrbracket$: $p_i \circ f = \lambda_i p_i$.
12. Montrer : $\forall k \in \mathbf{N}, f^k = \sum_{i=1}^m \lambda_i^k p_i$, puis, pour tout polynôme P de $\mathbf{R}[X]$: $P(f) = \sum_{i=1}^m P(\lambda_i) p_i$.

Partie III : Intervention de produits scalaires.

On munit le \mathbf{R} -espace vectoriel E d'un produit scalaire $\langle \cdot, \cdot \rangle$.

On considère l'application φ de $E \times E$ dans \mathbf{R} définie, pour tout $(x, y) \in E \times E$, par :

$$\varphi(x, y) = \sum_{i=1}^m \langle p_i(x), p_i(y) \rangle.$$

13. Montrer que φ est un produit scalaire sur E .

On remarquera qu'ainsi E est muni de deux produits scalaires $\langle \cdot, \cdot \rangle$ et φ .

14. Soit $i \in \llbracket 1, m \rrbracket$.
a. Montrer que pour tout $(x, y) \in E^2$, $\varphi(x, p_i(y)) = \varphi(p_i(x), y)$.
b. Montrer que $\text{Ker } p_i \subset (\text{Im } p_i)^\perp$, où $(\text{Im } p_i)^\perp$ désigne l'orthogonal de $\text{Im } p_i$ pour le produit scalaire φ .
c. En déduire que p_i est le projecteur orthogonal sur $E_{\lambda_i}(f)$, pour le produit scalaire φ .

CORRECTION DU DEVOIR SURVEILLÉ 5

EXERCICE (EDHEC 2010)

1. Vect(u) est de dimension 1 car u est non nul¹, et donc

¹ Car $\|u\| \neq 0$.

$$\dim(\text{Vect}(u))^\perp = \dim E - \dim \text{Vect}(u) = \boxed{n - 1.}$$

2. Il est clair que f_λ est à valeurs dans E , il s'agit donc de montrer que f_λ est linéaire. Soient donc $x, y, \in E$ et $\mu \in \mathbf{R}$. Alors

$$f_\lambda(\mu x + y) = \lambda(\mu x + y|u)u + \mu x + y = \lambda(\mu(x|u) + (y|u))u + \mu x + y = \mu(\lambda(x|u)u + x) + \lambda(y|u)u + y = \mu f_\lambda(x) + f_\lambda(y).$$

Ainsi, f_λ est bien un endomorphisme de E .

3. Soit $x \in E$. On a alors

$$\begin{aligned} f_\lambda^2(x) &= f_\lambda(f_\lambda(x)) = f_\lambda(\underbrace{\lambda(x|u)}_{\in \mathbf{R}} u + x) = \lambda(x|u)f_\lambda(u) + f_\lambda(x) \\ &= \lambda(x|u) \left(\lambda \underbrace{(u|u)}_{=\|u\|^2=1} u + u \right) + \lambda(x|u)u + x = \lambda^2(x|u)u + 2\lambda(x|u)u + x. \end{aligned}$$

D'autre part, on a

$$(\lambda + 2)f_\lambda(x) - (\lambda + 1)x = \lambda^2(x|u)u + \lambda x + 2\lambda(x|u)u + 2x - (\lambda + 1)x = \lambda^2(x|u)u + 2\lambda(x|u)u + x.$$

On a donc $f_\lambda^2(x) - (\lambda + 2)f_\lambda(x) + (\lambda + 1)\text{id}_E(x) = 0$.

Ceci étant vrai pour tout $x \in E$, il vient donc $f_\lambda^2 - (\lambda + 2)f_\lambda + (\lambda + 1)\text{id}_E = 0$.

Et alors $X^2 - (\lambda + 2)X + (\lambda + 1)$ est un polynôme annulateur de f_λ .

4.a. On a $f_\lambda(u) = \lambda(u|u)u + u = \lambda u + u = \boxed{(\lambda + 1)u}$.

Et si $v \in (\text{Vect}(u))^\perp$, alors

$$f_\lambda(v) = \lambda \underbrace{(v|u)}_{=0} u + v = \boxed{v}.$$

4.b. Notons que le polynôme annulateur obtenu précédemment a pour racines $\lambda + 1$ et 1 . Donc les seules valeurs propres possibles de f_λ sont $\lambda + 1$ et 1 .

Mais nous venons de prouver que $f_\lambda(u) = (\lambda + 1)u$, donc, puisque $u \neq 0$, $\lambda + 1$ est bien valeur propre, et $\text{Vect}(u) \subset E_{\lambda+1}(f_\lambda)$.

De même, si $v \in (\text{Vect}(u))^\perp$, alors $f_\lambda(v) = v$, donc 1 est bien² valeur propre de f_λ , et $(\text{Vect}(u))^\perp \subset E_1(f_\lambda)$.

Puisque $\lambda \neq 0$, on a donc deux valeurs propres distinctes, avec $\dim E_{\lambda+1}(f_\lambda) \geq 1$ et $\dim E_1(f_\lambda) \geq n - 1$.

Mais puisque $\dim E_{\lambda+1}(f_\lambda) + \dim E_1(f_\lambda) \leq n$, nécessairement

$$\dim E_{\lambda+1}(f_\lambda) = 1 \text{ et } \dim E_1(f_\lambda) = n - 1.$$

On en déduit donc que $E_{\lambda+1}(f_\lambda) = \text{Vect}(u)$ et $E_1(f_\lambda) = (\text{Vect}(u))^\perp$.

On a alors $\dim E_1(f_\lambda) + \dim E_{\lambda+1}(f_\lambda) = n = \dim E$, et donc f_λ est diagonalisable.

5.a. Le polynôme annulateur de f_{-1} obtenu précédemment est alors $X^2 - X + 0 = X^2 - X$.

On a donc $f_{-1}^2 = f_{-1}$: f_{-1} est un projecteur.

5.b. Nous savons que f_{-1} est la projection sur son image parallèlement à son noyau.

Or, $\text{Ker}(f_{-1}) = E_0(f_{-1}) = \text{Vect}(u)$ et $\text{Im}(f_{-1}) = (\text{Vect}(u))^\perp$.

Et donc f_{-1} est la projection sur $\text{Vect}(u)^\perp$, parallèlement à $\text{Vect}(u) = (\text{Vect}(u)^\perp)^\perp$: c'est le projecteur orthogonal sur $\text{Vect}(u)^\perp$.

Danger !

Rappelons (une fois encore) que les racines du polynôme annulateur ne sont pas nécessairement toutes des valeurs propres. Il faut vérifier une par une si elles sont valeurs propres ou non.

² $\text{Vect}(u)^\perp$ n'est pas réduit au vecteur nul.

Alternative

On peut aussi remarquer que f_{-1} est diagonalisable et ne possède que 0 et 1 comme valeurs propres, ce qui est une caractérisation des projecteurs.

Rappel

Si p est un projecteur, alors $\text{Im}(p) = E_1(p)$ et $\text{Ker}(p) = E_0(p)$.

EXERCICE 2 (ADAPTÉ DE EDHEC 2018)

1.a. On a $u_0 = \int_0^{\pi/2} (\cos t)^0 dt = \int_0^{\pi/2} 1 dt = \boxed{\frac{\pi}{2}}$.

Et $u_1 = \int_0^{\pi/2} \cos(t) dt = [\sin t]_0^{\pi/2} = \boxed{1}$.

1.b. Pour $t \in \left[0, \frac{\pi}{2}\right]$, on a $0 \leq \cos t \leq 1$. Et donc $(\cos t)^{n+1} \leq (\cos t)^n$.

Par croissance de l'intégrale, on a donc

$$u_{n+1} = \int_0^{\pi/2} (\cos t)^{n+1} dt \leq \int_0^{\pi/2} (\cos t)^n dt = u_n.$$

Et donc (u_n) est une suite décroissante.

1.c. La fonction $t \mapsto (\cos t)^n$ est positive sur $\left[0, \frac{\pi}{2}\right]$ et donc par croissance de l'intégrale,

$$\boxed{u_n \geq 0}.$$

D'autre part, $t \mapsto (\cos t)^n$ est continue, positive, et n'est pas la fonction nulle sur $\left[0, \frac{\pi}{2}\right]$. Et

donc³ $u_n \neq 0$, de sorte que $\boxed{u_n > 0}$.

2.a. Puisque (u_n) est décroissante et minorée⁴, par le théorème de la limite monotone, elle est convergente.

2.b. Par la relation de Chasles, on a $u_n = \int_0^a (\cos t)^n dt + \int_a^{\pi/2} (\cos t)^n dt$.

Or, pour $t \in [0, a]$, $(\cos t)^n \leq 1$ et donc $\int_0^a (\cos t)^n dt \leq \int_0^a 1 dt = a$.

D'autre part, la fonction \cos étant décroissante sur $\left[0, \frac{\pi}{2}\right]$, pour $t \in \left[a, \frac{\pi}{2}\right]$, on a $(\cos t)^n \leq (\cos a)^n$.

Et donc $\int_a^{\pi/2} (\cos t)^n dt \leq \int_a^{\pi/2} (\cos a)^n dt \leq \left(\frac{\pi}{2} - a\right) (\cos a)^n$.

On a donc bien

$$\boxed{u_n \leq a + \left(\frac{\pi}{2} - a\right) (\cos a)^n}.$$

2.c. Pour $a \in \left]0, \frac{\pi}{2}\right[$ fixé, on a $0 \leq \cos a < 1$, et donc $\lim_{n \rightarrow +\infty} (\cos a)^n = 0$.

Ainsi, en passant à la limite lorsque $n \rightarrow +\infty$ dans l'inégalité de la question précédente, il vient $\boxed{0 \leq \ell \leq a}$.

2.d. Puisque l'inégalité de la question 2.c est valable pour tout $a \in \left]0, \frac{\pi}{2}\right[$, en faisant tendre a vers 0, on obtient donc, par le théorème des gendarmes, $\boxed{\ell = 0}$.

3.a. Procédons à une intégration par parties en posant $u(t) = \sin(t)$ et $v(t) = (\cos t)^{n+1}$ qui sont deux fonctions de classe \mathcal{C}^1 sur $\left[0, \frac{\pi}{2}\right]$ avec $u'(t) = \cos(t)$ et $v'(t) = -(n+1)\sin(t)(\cos t)^n$.

On a alors

$$\begin{aligned} u_{n+2} &= \int_0^{\pi/2} \cos t (\cos t)^{n+1} dt \\ &= \underbrace{[\sin(t)(\cos t)^{n+1}]_0^{\pi/2}}_{=0} + \int_0^{\pi/2} (n+1) \sin^2(t) (\cos t)^n dt \\ &= (n+1) \int_0^{\pi/2} (1 - \cos^2 t) (\cos t)^n dt \\ &= (n+1) \int_0^{\pi/2} (\cos t)^n dt - (n+1) \int_0^{\pi/2} (\cos t)^{n+2} dt \\ &= (n+1)u_n - (n+1)u_{n+2}. \end{aligned}$$

Et donc en regroupant les termes en u_{n+2} , il vient

$$\boxed{(n+2)u_{n+2} = (n+1)u_n}.$$

⚠ Attention !

Si $x > 1$, alors la suite $(x^n)_n$ est croissante, mais pour $x \in [0, 1[$, la suite $(x^n)_n$ est décroissante.

³ Une fonction continue positive est d'intégrale nulle si et seulement si il s'agit de la fonction nulle.

⁴ Par 0.

Rappel

Pour tout $x \in \mathbf{R}$,
 $\cos^2(x) + \sin^2(x) = 1$.

3.b. Multiplions la relation de la question précédente par u_{n+1} :

$$\forall n \in \mathbf{N}, (n + 2)u_{n+2}u_{n+1} = (n + 1)u_{n+1}u_n.$$

Cela signifie donc que la suite $((n + 1)u_{n+1}u_n)_{n \in \mathbf{N}}$ est constante.

Or, nous avons déjà calculé u_0 et u_1 , et on a donc $1 \times u_1 u_0 = \frac{\pi}{2}$.

Et donc pour tout $n \in \mathbf{N}, (n + 1)u_{n+1}u_n = \frac{\pi}{2}$.

3.c. Il s'agit donc d'utiliser la relation de récurrence, couplée au fait que $u_0 = \frac{\pi}{2}$.

On prendra garde aux décalages d'indices : la relation précédemment obtenue s'écrit encore $u_n = \frac{\pi}{2nu_{n-1}}$.

```

1 n = input('entrez la valeur de n :');
2 u = %pi/2
3 for i=1 :n
4     u = %pi/(2*i*u)
5 end
6 disp(u)
```

3.d. Notons que nous sommes certains qu'un tel N existe, car $u_n \xrightarrow{n \rightarrow +\infty} 0$.

Il s'agit donc de calculer successivement les termes de la suite, jusqu'à en obtenir un qui soit plus petit que ε .

```

1 function N = limite(epsilon)
2     u = %pi/2 ;
3     N = 0 ;
4     while u >= epsilon
5         N = N+1 ;
6         u = %pi/(2*N*u)
7     end
8 endfunction
```

4. La fonction f est positive sur \mathbf{R} , et elle est continue sauf éventuellement⁵ en 0 et en $\frac{\pi}{2}$.

De plus,

$$\int_{-\infty}^{+\infty} f(t) dt = \int_0^{\pi/2} \sin t dt = [-\cos t]_0^{\pi/2} = 1.$$

Et donc f est une densité de probabilité.

5. Soit $x \in \mathbf{R}$. Alors $F(x) = \int_{-\infty}^x f(t) dt$.

Pour $x \leq 0$, on a $F(x) = \int_{-\infty}^x 0 dt = 0$.

Pour $0 \leq x \leq \frac{\pi}{2}$, $F(x) = \int_0^x \sin t dt = [-\cos t]_0^x = 1 - \cos x$.

Enfin, pour $x \geq \frac{\pi}{2}$,

$$F(x) = \int_{-\infty}^x f(t) dt = \int_{-\infty}^0 0 dt + \int_0^{\pi/2} \sin(t) dt + \int_{\pi/2}^x 0 dt = 1.$$

Donc

$$F(x) = \begin{cases} 0 & \text{si } x \leq 0 \\ 1 - \cos(x) & \text{si } 0 \leq x \leq \frac{\pi}{2} \\ 1 & \text{si } x \geq \frac{\pi}{2} \end{cases}$$

6.a. On a

$$E(X) = \int_{-\infty}^{+\infty} t f(t) dt = \int_0^{\pi/2} t \sin(t) dt.$$

⁵ En réalité, f est continue en 0, mais il n'est pas nécessaire de le vérifier.

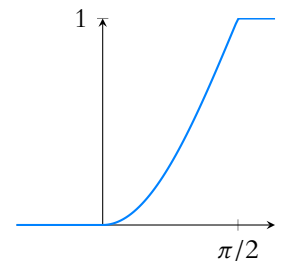


FIGURE 24.1– La fonction F .

Convergence

Notons que l'intégrale qui suit est une intégrale sur un segment, donc convergente, de sorte que X admet une espérance.

En procédant par une intégration par parties, on a

$$\int_0^{\pi/2} t \sin(t) dt = [-t \cos(t)]_0^{\pi/2} + \int_0^{\pi/2} \cos(t) dt = 1.$$

Donc $E(X) = 1$.

6.b. De même, par une double intégration par parties, on a⁶

$$\begin{aligned} E(X^2) &= \int_{-\infty}^{+\infty} t^2 f(t) dt = \int_0^{\pi/2} t^2 \sin(t) dt \\ &= [-t^2 \cos(t)]_0^{\pi/2} + 2 \int_0^{\pi/2} t \cos(t) dt = 2 \int_0^{\pi/2} t \cos(t) dt \\ &= 2 [t \sin t]_0^{\pi/2} - 2 \int_0^{\pi/2} \sin(t) dt \\ &= \pi - 2. \end{aligned}$$

Ainsi, X possède une variance, et par la formule de Huygens,

$$V(X) = E(X^2) - E(X)^2 = \pi - 2 - 1 = \pi - 3.$$

7.a. Pour tout $x \in \mathbf{R}$, $[I_n > x] = [\min(X_1, \dots, X_n) > x] = \bigcap_{i=1}^n [X_i > x]$.

Et donc par indépendance des X_i ,

$$\begin{aligned} F_n(x) &= P(I_n \leq x) = 1 - P(I_n > x) \\ &= 1 - \prod_{i=1}^n P(X_i > x) = 1 - \prod_{i=1}^n (1 - F(X_i \leq x)) \\ &= 1 - (1 - F(x))^n = \begin{cases} 0 & \text{si } x \leq 0 \\ 1 - (\cos x)^n & \text{si } 0 \leq x \leq \frac{\pi}{2} \\ 1 & \text{si } x \geq \frac{\pi}{2} \end{cases} \end{aligned}$$

Cette fonction est évidemment⁷ \mathcal{C}^1 sauf peut-être en 0 et en $\frac{\pi}{2}$.

D'autre part, on a $\lim_{x \rightarrow 0^+} 1 - (\cos x)^n = 0$ et $\lim_{x \rightarrow \frac{\pi}{2}} 1 - (\cos x)^n = 0$ et donc F est continue en 0 et en $\frac{\pi}{2}$.

Donc F est continue sur \mathbf{R} et \mathcal{C}^1 sauf peut-être en 0 et en $\frac{\pi}{2}$: I_n est une variable à densité.

7.b. Puisque I_n est une variable à densité, une densité de I_n est toute fonction coïncidant avec F'_n là où cette dernière est définie⁸.

Par exemple, on peut prendre $f_n : x \mapsto \begin{cases} n \sin(x)(\cos x)^{n-1} & \text{si } 0 \leq x \leq \frac{\pi}{2} \\ 0 & \text{sinon} \end{cases}$

7.c. On a, sous réserve de convergence,

$$E(I_n) = \int_{-\infty}^{+\infty} t f_n(t) dt = \int_0^{\pi/2} nt \sin(t)(\cos t)^{n-1} dt.$$

Notons qu'il s'agit là d'une intégrale sur un segment, et donc convergente, de sorte que I_n possède une espérance.

Une intégration par parties, en posant $u(t) = t$ et $v(t) = -(\cos t)^n$ nous donne alors

$$E(I_n) = [-t(\cos t)^n]_0^{\pi/2} + \int_0^{\pi/2} (\cos t)^n dt = \int_0^{\pi/2} (\cos t)^n dt = u_n.$$

7.d. De même, on a

$$E(I_n^2) = \int_{-\infty}^{+\infty} t^2 f_n(t) dt = \int_0^{\pi/2} nt^2 \sin t (\cos t)^{n-1} dt.$$

⁶ Là encore toutes les intégrales convergent car intégrales sur un segment.

⁷ Par opérations sur les fonctions usuelles.

⁸ Et nous venons de dire que F_n est dérivable sauf peut-être en 0 et en $\frac{\pi}{2}$.

Convergence

Notons qu'une fois encore, ces intégrales convergent puisqu'il s'agit d'intégrales sur un segment.

Procédons alors à une intégration par parties, en posant $u(t) = t^2$ et $v(t) = -(\cos t)^n$, qui sont deux fonctions de classe \mathcal{C}^1 avec $u'(t) = 2t$ et $v'(t) = n \sin t (\cos t)^{n-1}$.

$$E(I_n^2) = [t^2(\cos t)^n]_0^{\pi/2} + \int_0^{\pi/2} 2t(\cos t)^n dt = \boxed{2 \int_0^{\pi/2} t(\cos t)^n dt.}$$

7.e. Pour $t \in [0, \frac{\pi}{2}]$, on a $0 \leq t(\cos t)^n \leq \frac{\pi}{2}(\cos t)^n$.

Et donc par croissance de l'intégrale,

$$E(I_n^2) = 2 \int_0^{\pi/2} t(\cos t)^n dt \leq 2 \int_0^{\pi/2} \frac{\pi}{2}(\cos t)^n dt \leq \boxed{\pi u_n.}$$

On en déduit que $E(I_n^2) \xrightarrow{n \rightarrow +\infty} 0$.

Et donc par la formule de Huygens, $V(I_n) = E(I_n^2) - E(I_n)^2 = E(I_n^2) - u_n^2 \xrightarrow{n \rightarrow +\infty} 0$.

8.a. La fonction h est strictement décroissante sur $[0, \frac{\pi}{2}]$ puisque sa dérivée est $h'(t) = -\sin(t)$ qui est strictement négative sur $]0, \frac{\pi}{2}[$.

D'autre part, on a $h(0) = 1$ et $h(\frac{\pi}{2}) = 0$.

Puisque h est continue, par le théorème de la bijection, h réalise une bijection de $[0, \frac{\pi}{2}]$ sur $[0, 1]$.

8.b. L'expression de F prouve que X prend ses valeurs dans $[0, \frac{\pi}{2}]$, et donc il est légitime de poser $Y = h(X)$.

La variable aléatoire Y est alors à valeurs dans $[0, 1]$, de sorte que $P(Y \leq x) = 0$ si $x < 0$ et $P(Y \leq x) = 1$ si $x \geq 1$.

Pour $x \in \mathbf{R}$, on a

$$\begin{aligned} G(x) &= P(Y \leq x) = P(h(X) \leq x) = P(X \geq h^{-1}(x)) \\ &= 1 - P(X < h^{-1}(x)) = 1 - F_X(h^{-1}(x)) = 1 - (1 - h(h^{-1}(x))) = x. \end{aligned}$$

$$\text{Donc } G(x) = \begin{cases} 0 & \text{si } x \leq 0 \\ x & \text{si } 0 \leq x \leq 1 \\ 1 & \text{si } x \geq 1 \end{cases}$$

On reconnaît alors la fonction de répartition d'une loi uniforme sur $[0, 1]$, de sorte que

$$Y \leftrightarrow \mathcal{U}([0, 1]).$$

8.c. Puisqu'on a $X = h^{-1}(Y)$, le code suivant convient :

```
1 Y = grand(1, 1, 'unf', 0, 1)
2 X = acos(Y)
```

8.d. Il nous faut donc simuler n variables de même loi que X , et prendre leur minimum.

```
1 function y=I(n)
2     X = zeros(n, 1);
3     for i=1 : n
4         X(i) = acos(rand());
5     end
6     y = min(X)
7 endfunction
```

Enfin, puisque nous savons que $u_n = E(I_n)$, il nous faut estimer cette espérance. Pour cela, simulons un grand nombre de fois la variable I_n , et considérons la moyenne de ces simulations.

```
1 S = 0;
2 for i=1 : 10 000
3     S = S+I(2019);
4 end
5 disp(S/10000)
```

Méthode

Lors de l'étude d'une variable aléatoire X , se demander quelles valeurs elle peut prendre permet de restreindre les valeurs de x pour lesquelles il faudra vraiment faire un calcul pour déterminer $F_X(x)$.

Danger !

h étant décroissante, il en est de même de h^{-1} , et donc le sens des inégalité change !

PROBLÈME (EML 2016)

Partie I : Étude d'un exemple.

1. Il s'agit de résoudre un système de deux équations à deux inconnues matricielles :

$$\begin{cases} P_1 + P_2 = I_3 \\ 4P_1 + 9P_2 = A \end{cases} \xleftrightarrow{L_2 \rightarrow L_2 - 4L_1} \begin{cases} P_1 + P_2 = I_3 \\ 5P_2 = A - 4I_3 \end{cases} \Leftrightarrow \begin{cases} P_1 = \frac{9I_3 - A}{5} \\ P_2 = \frac{A - 4I_3}{5} \end{cases}$$

$$\text{Soit } P_1 = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & -1 & 1 \end{pmatrix}, P_2 = \begin{pmatrix} 0 & 0 & 0 \\ -1 & 1 & 0 \\ -1 & 1 & 0 \end{pmatrix}.$$

- 2.a. On a : $P_1^2 = P_1, P_1P_2 = P_2P_1 = 0$ et $P_2^2 = P_2$.

- 2.b. Prouvons par récurrence sur $k \in \mathbf{N}$ que $A^k = 4^k P_1 + 9^k P_2$.

Pour $k = 0$, on a $A^0 = I_3 = P_1 + P_2 = 4^0 P_1 + 9^0 P_2$.

Supposons que $A^k = 4^k P_1 + 9^k P_2$. Alors

$$\begin{aligned} A^{k+1} &= A^k A = (4^k P_1 + 9^k P_2)(4P_1 + 9P_2) \\ &= 4^{k+1} P_1^2 + 9 \times 4^k P_1^k P_2 + 4 \times 9^k P_2 P_1 + 9^{k+1} P_2^2 \\ &= 4^{k+1} P_1 + 9^{k+1} P_2. \end{aligned}$$

Alternative

On peut aussi utiliser le binôme de Newton matriciel avec

$$A^k = (4P_1 + 9P_2)^k$$

mais alors il ne faut pas oublier de signaler que P_1 et P_2 commutent (car nous avons prouvé que $P_1P_2 = P_2P_1 = 0$).

Ainsi, la propriété est héréditaire et donc, d'après le principe de récurrence, $\forall k \in \mathbf{N}, A^k = 4^k P_1 + 9^k P_2$.

3. Cherchons une telle matrice sous la forme $B = \lambda P_1 + \mu P_2$.

On a alors $B^2 = \lambda^2 P_1^2 + \lambda \mu P_1 P_2 + \lambda \mu P_2 P_1 + \mu^2 P_2^2 = \lambda^2 P_1 + \mu^2 P_2$.

Si l'on prend par exemple⁹ $\lambda = 2$ et $\mu = 3$, il vient

$$(2P_1 + 3P_2)^2 = 4P_1 + 9P_2 = A.$$

⁹ Il est clair qu'il y a d'autres solutions, par exemple $\lambda = 2$ et $\mu = -3$.

$$\text{Donc } B = 2P_1 + 3P_2 = \begin{pmatrix} 2 & 0 & 0 \\ -1 & 3 & 0 \\ -1 & 1 & -2 \end{pmatrix} \text{ convient.}$$

4. A est triangulaire¹⁰, donc ses valeurs propres sont ses coefficients diagonaux : $\text{Spec}(A) = \{4, 9\}$.

¹⁰ Inférieure.

Partie II : Étude des puissances de f .

5. Soit $P = \sum_{k=0}^m a_k X^k \in \mathbf{R}_m[X]$. Alors

$$P(f) = \sum_{k=0}^m a_k f^k = \sum_{k=0}^m a_k \sum_{i=1}^m \lambda_i^k p_i = \sum_{i=1}^m \left(\sum_{k=0}^m a_k \lambda_i^k \right) p_i = \sum_{i=1}^m P(\lambda_i) p_i.$$

6. Pour tout $i \in \llbracket 1, m \rrbracket$, λ_i est racine de N et donc $N(\lambda_i) = 0$.

On en déduit que

$$N(f) = \sum_{i=1}^m N(\lambda_i) p_i = \sum_{i=1}^m 0 \times p_i = \boxed{\tilde{0}}.$$

- 7.a. Les racines de M_i sont les $\lambda_j, 1 \leq j \leq m, j \neq i$.

En particulier, pour $i \neq j, M_i(\lambda_j) = 0$. Et donc $L_i(\lambda_j) = \frac{0}{M_i(\lambda_j)} = 0$.

Si $j = i$, alors $L_i(\lambda_i) = \frac{M_i(\lambda_i)}{M_i(\lambda_i)} = 1$.

- 7.b. Soit $i \in \llbracket 1, m \rrbracket$. Puisque $\deg(L_i) = \deg(M_i) = m - 1$, d'après la question 5 on a

$$L_i(f) = \sum_{j=1}^m L_i(\lambda_j) p_j = \sum_{\substack{j=1 \\ j \neq i}}^m \underbrace{L_i(\lambda_j)}_{=0} p_j + \underbrace{L_i(\lambda_i)}_{=1} p_i = \boxed{p_i}.$$

Remarque

Puisque λ_i n'est pas racine de M_i , on a bien $M_i(\lambda_i) \neq 0$ et donc la division par $M_i(\lambda_i)$ dans la définition de L_i est légitime.

Pour la culture

Les L_i sont appelés polynômes de Lagrange.

8.a. Par hypothèse, on a

$$e = f^0 = \sum_{i=1}^m \lambda_i^0 p_i = \boxed{\sum_{i=1}^m p_i.}$$

8.b. Soit $x \in E$. Alors

$$x = e(x) = \sum_{i=1}^m \underbrace{p_i(x)}_{\in \text{Im } p_i} \in \sum_{i=1}^m \text{Im}(p_i).$$

Et donc $E \subset \sum_{i=1}^m \text{Im}(p_i)$.

Les $\text{Im}(p_i)$ étant des sous-espaces vectoriels de E , on a évidemment $\sum_{i=1}^m \text{Im}(p_i) \subset E$ et donc

$$\boxed{E = \sum_{i=1}^m \text{Im}(p_i).}$$

9.a. On a

$$N = \prod_{\ell=1}^m (X - \lambda_\ell) = (X - \lambda_i) \prod_{\substack{\ell=1 \\ \ell \neq i}}^m (X - \lambda_\ell) = (X - \lambda_i) M_i = \boxed{(X - \lambda_i) M_i(\lambda_i) L_i.}$$

9.b. Soit $x \in \text{Im}(p_i)$. Alors il existe $z \in E$ tel que $x = p_i(z)$.

Et alors

$$(f - \lambda_i e)(x) = (f - \lambda_i e)(p_i(z)) = ((f - \lambda_i e) \circ L_i(f))(z).$$

Mais d'après la question 9.a, $(X - \lambda_i)L_i = \frac{1}{M_i(\lambda_i)}N$, de sorte que $(f - \lambda_i e) \circ L_i(f) = \frac{1}{M_i(\lambda_i)}N(f)$. On en déduit donc que

$$(f - \lambda_i e)(x) = \frac{1}{M_i(\lambda_i)}N(f)(z) = \frac{1}{\lambda_i}0 = 0.$$

Ainsi, $x \in \text{Ker}(f - \lambda_i e)$ et donc $\boxed{\text{Im}(p_i) \subset \text{Ker}(f - \lambda_i e)}$.

10. Soit $x \in E$. Alors il existe $x_1 \in \text{Im}(p_1), \dots, x_m \in \text{Im}(p_m)$ tels que $x = x_1 + x_2 + \dots + x_m$.

Mais puisque $x_i \in \text{Im}(p_i) \subset \text{Ker}(f - \lambda_i e)$, on en déduit que $x \in \sum_{i=1}^m \text{Ker}(f - \lambda_i e)$.

Et donc $E = \sum_{i=1}^m \text{Ker}(f - \lambda_i e)$.

De plus, les $\text{Ker}(f - \lambda_i e)$ ne sont pas réduits à $\{0\}$ car $\text{Im}(p_i) \subset \text{Ker}(f - \lambda_i e)$, et les p_i sont tous non nuls. Ce sont donc des¹¹ sous-espaces propres de f .

D'autre part, les sous-espaces propres de f sont en somme directe, donc

$$E = \bigoplus_{i=1}^m \text{Ker}(f - \lambda_i e), \text{ de sorte que } \dim E = \sum_{i=1}^m \dim \text{Ker}(f - \lambda_i e).$$

Comme de plus on a, comme toujours, $\dim E \geq \sum_{\lambda \in \text{Spec}(f)} E_\lambda(f)$, f ne possède donc pas

d'autres valeurs propres que les λ_i , $1 \leq i \leq m$ et donc E est somme des sous-espaces propres de f : $\boxed{f \text{ est diagonalisable.}}$

Enfin, puisque $E = \sum_{i=1}^m \text{Im}(p_i)$, alors $\dim E \leq \sum_{i=1}^m \dim \text{Im}(p_i)$.

Comme d'autre part, on a $\dim \text{Im}(p_i) \leq \dim \text{Ker}(f - \lambda_i e)$ et $\sum_{i=1}^m \dim \text{Ker}(f - \lambda_i e) = \dim E$,

on en déduit que

$$\forall i \in \llbracket 1, m \rrbracket, \dim \text{Im}(p_i) = \dim \text{Ker}(f - \lambda_i e).$$

Et puisque $\text{Im}(p_i) \subset \text{Ker}(f - \lambda_i e)$, ces deux sous-espaces vectoriels de E sont donc égaux :

$$\boxed{\forall i \in \llbracket 1, m \rrbracket, E_{\lambda_i}(f) = \text{Ker}(f - \lambda_i e) = \text{Im}(p_i).}$$

Somme

Rappelons que par définition,

$$\sum_{i=1}^m \text{Im}(p_i)$$

est l'ensemble des éléments de E qui peuvent s'écrire comme un élément de $\text{Im}(p_1)$ plus un élément de $\text{Im}(p_2)$, etc.

Détail

D'après 7.b, $p_i = L_i(f)$.

¹¹ Il est encore un peu tôt pour dire les sous-espaces propres de f , car f pourrait avoir d'autres valeurs propres que les λ_i .

Détails

La concaténation de bases des $\text{Im}(p_i)$ est de cardinal $\sum_{i=1}^m \dim \text{Im}(p_i)$, et elle est génératrice de E , donc

$$\dim E \leq \sum_{i=1}^m \dim \text{Im}(p_i).$$

- 11.a. Soit $x \in E$. Alors $p_j(x) \in \text{Im}(p_j) \subset \text{Ker}(f - \lambda_j e)$, de sorte que $f(p_j(x)) = \lambda_j p_j(x)$.
Donc pour tout polynôme $P \in \mathbf{R}[X]$, $P(f)(p_j(x)) = P(\lambda_j) \cdot p_j(x)$.
En particulier, si $P = L_i$, il vient

$$p_i(p_j(x)) = L_i(f)(p_j(x)) = \underbrace{L_i(\lambda_j)}_{=0 \text{ car } i \neq j} p_j(x) = 0.$$

Et donc $p_i \circ p_j = 0$.

- 11.b. Le même raisonnement que précédemment prouve que pour tout $x \in E$,

$$p_i \circ p_i(x) = L_i(\lambda_i) p_i(x) = p_i(x)$$

et donc $p_i \circ p_i = p_i$.

- 11.c. En combinant les résultats des deux questions précédentes, il vient

$$p_i \circ f = p_i \circ \left(\sum_{j=1}^m \lambda_j p_j \right) = \sum_{j=1}^m \lambda_j \underbrace{p_i \circ p_j}_{=0 \text{ si } i \neq j} = \lambda_i p_i.$$

12. Prouvons le résultat par récurrence sur k . La récurrence est déjà largement initialisée puisque par hypothèse, le résultat est vrai pour $k \in \llbracket 0, m \rrbracket$.

Supposons donc que $f^k = \sum_{i=1}^m \lambda_i^k p_i$. Alors

$$f^{k+1} = f \circ f^k = \left(\sum_{i=1}^m \lambda_i^k p_i \right) \circ f = \sum_{i=1}^m \lambda_i^k p_i \circ f = \sum_{i=1}^m \lambda_i^k \lambda_i p_i = \sum_{i=1}^m \lambda_i^{k+1} p_i.$$

Ainsi, la propriété est vraie au rang $k + 1$, et par le principe de récurrence :

$$\forall k \in \mathbf{N}, f^k = \sum_{i=1}^m \lambda_i^k p_i.$$

Enfin, si $P = \sum_{k=0}^d a_k X^k \in \mathbf{R}[X]$, alors, de même qu'à la question 5, il vient

$$P(f) = \sum_{k=0}^d a_k f^k = \sum_{k=0}^d a_k \sum_{i=1}^m \lambda_i^k p_i = \sum_{i=1}^m \left(\sum_{k=0}^d a_k \lambda_i^k \right) p_i = \sum_{i=1}^m P(\lambda_i) p_i.$$

Partie III : Intervention de produits scalaires.

13. Soient $x, y \in E$. Alors

$$\varphi(x, y) = \sum_{i=1}^m \langle p_i(x), p_i(y) \rangle = \sum_{i=1}^m \langle p_i(y), p_i(x) \rangle = \varphi(y, x).$$

Donc φ est symétrique.

Soient $x, y, z \in E$ et $\lambda \in \mathbf{R}$. Alors

$$\begin{aligned} \varphi(\lambda x + y, z) &= \sum_{i=1}^m \langle p_i(\lambda x + y), p_i(z) \rangle \\ &= \sum_{i=1}^m \langle \lambda p_i(x) + p_i(y), p_i(z) \rangle && \text{Linéarité des } p_i. \\ &= \sum_{i=1}^m \lambda \langle p_i(x), p_i(z) \rangle + \langle p_i(y), p_i(z) \rangle && \text{Linéarité à gauche de } \langle \cdot, \cdot \rangle. \\ &= \lambda \sum_{i=1}^m \langle p_i(x), p_i(z) \rangle + \sum_{i=1}^m \langle p_i(y), p_i(z) \rangle = \lambda \varphi(x, z) + \varphi(y, z). \end{aligned}$$

Complément

Nous venons de prouver que p_i est un projecteur, d'image égale à $E_{\lambda_i}(f)$.

La question précédente permettrait de prouver qu'on a également

$$\text{Ker } p_i = \bigoplus_{j \neq i} E_{\lambda_j}(f).$$

Et donc p_i est la projection sur $E_{\lambda_i}(f)$ parallèlement à la somme des autres sous-espaces propres. On parle de projecteur spectral.

Ainsi φ est linéaire à gauche, et étant symétrique, elle est bilinéaire symétrique.

Soit $x \in E$. Alors $\varphi(x, x) = \sum_{i=1}^m \langle p_i(x), p_i(x) \rangle = \sum_{i=1}^m \|p_i(x)\|^2 \geq 0$.

De plus, une somme de nombres positifs est nulle si et seulement si chacun de ses termes est nul, donc

$$\varphi(x, x) = 0 \Leftrightarrow \forall i \in \llbracket 1, m \rrbracket, \|p_i(x)\|^2 = 0 \Leftrightarrow \forall i \in \llbracket 1, m \rrbracket, p_i(x) = 0_E.$$

Mais alors, d'après le résultat de la question 8.a, $x = e(x) = \sum_{i=1}^m p_i(x) = 0_E$.

Donc $\varphi(x, x) = 0 \Rightarrow x = 0_E$ et la réciproque est toujours vérifiée.

Donc φ est bien un produit scalaire sur E .

14.a. Pour tous $x, y \in E$, on a

$$\varphi(p_i(x), y) = \sum_{j=1}^m \langle p_j(p_i(x)), p_j(y) \rangle = \langle p_i(x), p_i(y) \rangle.$$

De même, on a $\varphi(x, p_i(y)) = \langle p_i(x), p_i(y) \rangle$, et donc $\varphi(x, p_i(y)) = \varphi(p_i(x), y)$.

14.b. Soit $x \in \text{Ker } p_i$ et $y \in \text{Im } p_i$. Alors il existe $z \in E$ tel que $y = p_i(z)$. Et donc

$$\varphi(x, y) = \varphi(x, p_i(z)) = \varphi(\underbrace{p_i(x)}_{=0_E}, z) = 0.$$

Et donc x est orthogonal¹² à tout vecteur de $\text{Im } p_i$. On a donc prouvé que $\text{Ker } p_i \subset (\text{Im } p_i)^\perp$.

14.c. Nous savons déjà¹³ que $\text{Im } p_i = E_{\lambda_i}(f)$.

Il s'agit donc de prouver que $\text{Ker } p_i = (\text{Im } p_i)^\perp$.

Mais d'après le théorème du rang,

$$\dim \text{Ker } p_i = \dim E - \dim \text{Im } p_i = \dim(\text{Im } p_i)^\perp.$$

D'après la question précédente, $\text{Ker } p_i \subset (\text{Im } p_i)^\perp$, et ces deux espaces ayant même dimension, ils sont donc égaux : $\text{Ker } p_i = (\text{Im } p_i)^\perp = (E_{\lambda_i}(f))^\perp$.

Et donc p_i est le projecteur orthogonal sur $E_{\lambda_i}(f)$.

Rappel

Un vecteur est de norme nulle si et seulement si il est nul.

On a utilisé ici les résultats des questions 11.a et 11.b :

$$p_j \circ p_i = \begin{cases} 0 & \text{si } i \neq j \\ p_i & \text{si } i = j \end{cases}$$

¹² Pour le produit scalaire φ , et non pour le produit scalaire $\langle \cdot, \cdot \rangle$.

¹³ C'est le résultat de la question 10.

DEVOIR SURVEILLÉ 6 (09.02.19)

EXERCICE

On considère la fonction f définie sur \mathbf{R} par : $\forall t \in \mathbf{R}, f(t) = \begin{cases} 0 & \text{si } t < 0 \\ \frac{2}{\pi(1+t^2)} & \text{si } t \geq 0 \end{cases}$

1. Montrer que f est une densité.
2. On considère une variable aléatoire réelle X à densité, de densité f .
 - a. Déterminer la fonction de répartition F_X de X .
 - b. La variable X admet-elle une espérance ? Une variance ?
3. On considère une suite de variables aléatoires réelles $(X_n)_{n \in \mathbf{N}^*}$ à densité, à valeurs strictement positives, mutuellement indépendantes, dont chacun à pour densité f .

On définit, pour tout n de \mathbf{N}^* , les variables aléatoires $M_n = \max(X_1, \dots, X_n)$ et $Z_n = \frac{n}{M_n}$.

- a. Déterminer, pour tout n de \mathbf{N}^* , la fonction de répartition F_{M_n} de M_n .
- b. Justifier : $\forall u \in]0, +\infty[, \text{Arctan}(u) + \text{Arctan}\left(\frac{1}{u}\right) = \frac{\pi}{2}$ et $\text{Arctan}(u) \underset{u \rightarrow 0}{\sim} u$.
- c. Montrer alors, pour tout n de \mathbf{N}^* : $\forall x \in]0, +\infty[, P(Z_n \leq x) = 1 - \left(1 - \frac{2}{\pi} \text{Arctan}\left(\frac{x}{n}\right)\right)^n$.
- d. En déduire que la suite de variables aléatoires $(Z_n)_{n \in \mathbf{N}^*}$ converge en loi vers une variable aléatoire à densité dont on reconnaîtra la loi.

PROBLÈME 1

On effectue des lancers d'une pièce donnant «Pile» avec la probabilité p , élément de $]0, 1[$, et donnant «face» avec la probabilité $q = 1 - p$, les différents lancers étant supposés indépendants.

Pour tout entier naturel k non nul, on note P_k (resp. F_k) l'événement : «la pièce donne «pile» (resp. «face») au $k^{\text{ème}}$ lancer», on note également S_k le rang du $k^{\text{ème}}$ pile et T_k le rang d'apparition du dernier «pile» de la première série de k «piles» consécutifs. On suppose que S_k et T_k sont deux variables aléatoires toutes deux définies sur le même espace probabilisé (Ω, \mathcal{A}, P) .

Par exemple, si les lancers donnent $F_1, P_2, F_3, P_4, F_5, P_6, P_7, P_8$, alors S_1 et T_1 prennent la valeur 2, S_2 prend la valeur 4, T_2 prend la valeur 7, S_3 prend la valeur 6, T_3 prend la valeur 8, S_4 prend la valeur 7 et S_5 prend la valeur 8.

Partie 1 : simulations de S_k et T_k .

1. Compléter les lignes 7 et 10 du script SciLab suivant pour qu'il affiche la valeur prise par S_k lorsque k et p sont entrés par l'utilisateur :

```

1 k = input('donnez une valeur pour k :')
2 p = input('donnez une valeur pour p :')
3 n = 0
4 c = 0
5 while c < k
6     n = n + 1
7     if --- then c = c + 1
8     end
9 end
10 disp(---)

```

2. On souhaite que le script précédent affiche la valeur prise par T_k . Remplacer la ligne 7 par la suivante, dûment complétée :

```

7     if --- then c = c + 1, else ---

```

Partie 2 : calcul de l'espérance S_k .

3. Donner la loi de S_1 ainsi que son espérance.

4. Soit k un entier naturel supérieur ou égal à 2. Pour tout entier naturel n supérieur ou égal à k , on note X_{n-1} la variable aléatoire égale au nombre de « piles » obtenus lors des $n - 1$ premiers lancers.
- Donner la loi de X_{n-1} .
 - Donner $S_k(\Omega)$ puis écrire l'événement $[S_k = n]$ à l'aide de la variable X_{n-1} .
 - En déduire que la loi de S_k est donnée par :

$$\forall n \geq k, P(S_k = n) = \binom{n-1}{k-1} p^k q^{n-k}.$$

5. Soit k un entier naturel supérieur ou égal à 2. On pose $Z_1 = S_1$ et, pour tout entier naturel i supérieur ou égal à 2, on pose $Z_i = S_i - S_{i-1}$. On admet que $(Z_i)_{i \in \mathbf{N}^*}$ est une suite de variables aléatoires mutuellement indépendantes.
- Donner la loi des variables aléatoires Z_i .
 - Exprimer S_k à l'aide de certaines des variables Z_i .
 - En déduire que S_k possède une espérance et donner sa valeur.
6. **Estimation.**

On suppose le paramètre p inconnu et on souhaite trouver un estimateur de p . On admet que, si une suite de variables aléatoires $(Y_k)_{k \in \mathbf{N}^*}$ converge en probabilité vers une variable aléatoire Y , alors pour toute fonction f continue sur un intervalle I de \mathbf{R} tel que $P(Y \in I) = 1$, la suite $(f(Y_k))_{k \in \mathbf{N}^*}$ converge en probabilité vers $f(Y)$.

- Pour tout entier naturel k supérieur ou égal à 2, on pose $\overline{Z}_k = \frac{1}{k} \sum_{i=1}^k Z_i$.
Montrer que la suite $(\overline{Z}_k)_{k \in \mathbf{N}^*}$ converge en probabilité.
- En déduire que $\frac{k}{S_k}$ est un estimateur convergent de p .
- Donner sans calcul la valeur de $\sum_{j=k-1}^{+\infty} P(S_{k-1} = j)$. Montrer alors que la variable aléatoire $\frac{k-1}{S_k-1}$ possède une espérance et que l'on a : $E\left(\frac{k-1}{S_k-1}\right) = p$.
- En déduire que $\frac{k}{S_k}$ est un estimateur biaisé de p (on ne cherchera pas à calculer la valeur de ce biais).

Partie 3 : calcul de l'espérance de T_k .

7. Comparer les variables aléatoires S_1 et T_1 .
8. Soit k un entier naturel supérieur ou égal à 2. On admet que T_k possède une espérance que l'on se propose de déterminer.
- Justifier, en utilisant la variable aléatoire W égale au rang du premier face lors de l'expérience décrite au début de ce problème, que les événements $F_1, P_1 \cap F_2, P_1 \cap P_2 \cap F_3, P_1 \cap P_2 \cap P_3 \cap F_4, \dots, P_1 \cap \dots \cap P_{k-1} \cap F_k$ et $P_1 \cap \dots \cap P_{k-1} \cap P_k$, forment un système complet d'événements.
 - Montrer que, pour tout entier naturel n supérieur ou égal à k , on a $P_{F_1}(T_k = n) = P(T_k = n - 1)$, puis en déduire que l'espérance conditionnelle $E(T_k|F_1)$ est égale à $1 + E(T_k)$.
 - De la même façon, déterminer, pour tout i de $\llbracket 2, k \rrbracket$, la valeur de $E(T_k|P_1 \cap \dots \cap P_{i-1} \cap F_i)$.
 - Justifier que $E(T_k|P_1 \cap \dots \cap P_k) = k$.
9.
 - Déduire des questions précédentes, la relation : $E(T_k) = \sum_{j=0}^{k-1} (j+1)p^j q + E(T_k) \sum_{j=0}^{k-1} p^j q + kp^k$.
 - Établir finalement que :

$$E(T_k) = \frac{1-p^k}{qp^k}.$$

10. Justifier que $E(S_k) \leq E(T_k)$ puis utiliser certains résultats des parties 2 et 3 pour établir, sans étude de fonction, que l'on a :

$$\forall k \in \mathbf{N}^*, \forall p \in]0, 1[, (k-1)p^k \geq kp^{k-1} - 1.$$

PROBLÈME 2

Dans tout le problème, n est un entier tel que $n \geq 2$.

On confond polynôme de $\mathbf{R}_n[X]$ et fonction polynomiale sur \mathbf{R} ou sur $]0, +\infty[$ ou sur $]0; +\infty[$.

On note $\mathbf{R}_{n-1}[X]$ le sous-espace vectoriel de $\mathbf{R}[X]$ formé des polynômes de degré inférieur ou égal à $n - 1$.

Partie I : Interpolation polynomiale

Soient a_1, \dots, a_n des réels deux à deux distincts. On note

$$\varphi : \mathbf{R}_{n-1}[X] \rightarrow \mathbf{R}^n, P \mapsto (P(a_1), \dots, P(a_n)).$$

1. Montrer que φ est un isomorphisme.
2. En déduire que, pour tout $(b_1, \dots, b_n) \in \mathbf{R}^n$, il existe $P \in \mathbf{R}_{n-1}[X]$ unique tel que :

$$\forall i \in \{1, \dots, n\}, P(a_i) = b_i.$$

3. *Exemple* : déterminer le polynôme P_0 de $\mathbf{R}_3[X]$ tel que :

$$P_0(0) = 1, P_0(1) = 3, P_0(2) = 11, P_0(3) = 31.$$

Partie II : Polynômes spéciaux

On considère l'ensemble E des polynômes P de $\mathbf{R}[X]$ tels que

$$\forall x \in]0, +\infty[, (P(x) > 0 \text{ et } P'(x) > 0).$$

4. Donner un exemple d'élément de E .
5. Montrer que E est stable par multiplication par un réel strictement positif, par addition et par multiplication, c'est-à-dire que, pour tout $\alpha \in]0, +\infty[$ et tous $P, Q \in E$, on a :

$$\alpha P \in E, P + Q \in E, PQ \in E.$$

Est-ce que E est un sous-espace vectoriel de $\mathbf{R}[X]$?

6. Soit $P \in E$. On note $P_1 : \mathbf{R} \rightarrow \mathbf{R}, x \mapsto \int_0^x P(t) dt$.
Montrer que $P_1 \in E$.

7. Soit $P \in E$. Montrer : $\forall x \in [0, +\infty[, P(x) \geq P(0)$.

Pour tout $P \in E$, on note $\tilde{P} : [0, +\infty[\rightarrow [P(0), +\infty[, x \mapsto \tilde{P}(x) = P(x)$.

8. Montrer que l'application \tilde{P} est bijective.
9. Si, de plus, P est de degré au moins 2, est-ce que l'application réciproque \tilde{P}^{-1} de \tilde{P} est une application polynomiale ?

Partie III : Matrices symétriques positives

On note \mathbf{S}_n^+ l'ensemble des matrices symétriques A de $\mathcal{M}_n(\mathbf{R})$ telles que

$$\forall U \in \mathcal{M}_{n,1}(\mathbf{R}), {}^tUAU \geq 0.$$

Soit A une matrice symétrique de $\mathcal{M}_n(\mathbf{R})$.

10. La matrice A est-elle diagonalisable dans $\mathcal{M}_n(\mathbf{R})$? Justifier.
11. a. Montrer que si A est dans \mathbf{S}_n^+ , alors toutes les valeurs propres de A sont dans $[0, +\infty[$.
b. Réciproquement, montrer que si toutes les valeurs propres de A sont dans $[0, +\infty[$, alors A est dans \mathbf{S}_n^+ .

Partie IV : Matrice symétrique positive solution d'une équation polynomiale spéciale

Soit $P \in E$ de degré $n - 1$ (l'ensemble E a été défini dans la partie II), et soit $A \in \mathbf{S}_n^+$ admettant n valeurs propres deux à deux distinctes, notées $\lambda_1, \dots, \lambda_n$, appartenant toutes à $[P(0); +\infty[$.

On note D la matrice diagonale de $\mathcal{M}_n(\mathbf{R})$ dont les termes diagonaux sont successivement $\lambda_1, \dots, \lambda_n$, et Q une matrice inversible de $\mathcal{M}_n(\mathbf{R})$ telle que $A = QDQ^{-1}$.

On se propose de résoudre l'équation $P(S) = A$, d'inconnue $S \in \mathbf{S}_n^+$.

12. On suppose que l'équation $P(S) = A$ a une solution dans \mathbf{S}_n^+ .
Soit S appartenant à \mathbf{S}_n^+ telle que $P(S) = A$. On note $\Delta = Q^{-1}SQ$.
a. Montrer que $SA = AS$ et en déduire que $\Delta D = D\Delta$.

- b. Démontrer que Δ est diagonale et que les éléments diagonaux de Δ sont tous positifs ou nuls.
13. Établir que l'équation $P(S) = A$, d'inconnue $S \in \mathbf{S}_n^+$ admet une solution et une seule, et que celle-ci est $Q\Delta Q^{-1}$, où Δ est une matrice diagonale que l'on exprimera à l'aide de $\tilde{P}^{-1}(\lambda_1), \dots, \tilde{P}^{-1}(\lambda_n)$, où \tilde{P} a été définie dans la partie II.

14. *Exemple* : On prend ici $n = 4$, $P = X^3 + X + 1$ et $A = \begin{pmatrix} 2 & -1 & 0 & 0 \\ -1 & 2 & 0 & 0 \\ 0 & 0 & 21 & 10 \\ 0 & 0 & 10 & 21 \end{pmatrix}$.

- a. Vérifier que $P \in E$.
- b. Déterminer les valeurs propres de A et montrer : $A \in \mathbf{S}_4^+$.
- c. Déterminer une matrice diagonale D et une matrice orthogonale Q telles que $A = QDQ^{-1}$.
- d. Résoudre l'équation $P(S) = A$, d'inconnue $S \in \mathbf{S}_4^+$.

CORRECTION DU DEVOIR SURVEILLÉ 6

EXERCICE (EXTRAIT DE EML 2017)

1. Il est évident que f est à valeurs positives et continue sur \mathbf{R} , sauf en 0. De plus, pour $x \geq 0$, on a

$$\int_0^x f(t) dt = \int_0^x \frac{2}{\pi} \frac{1}{1+t^2} dt = \frac{2}{\pi} [\text{Arctan}(t)]_0^x = \frac{2}{\pi} \text{Arctan}(x) \xrightarrow{x \rightarrow +\infty} 1.$$

Donc $\int_0^{+\infty} f(t) dt$ converge et vaut 1. Puisque f est nulle sur \mathbf{R}^* , on en déduit que

$$\int_{-\infty}^{+\infty} f(t) dt \text{ converge et vaut 1. Donc } \boxed{f \text{ est bien une densité.}}$$

- 2.a. Soit $x \leq 0$. Alors $F_X(x) = \int_{-\infty}^x f(t) dt = \int_{-\infty}^x 0 dt = 0$.

Si $x > 0$, alors

$$F_X(x) = P(X \leq x) = \frac{2}{\pi} \int_0^x \frac{dt}{1+t^2} = \frac{2}{\pi} [\text{Arctan}(t)]_0^x = \frac{2}{\pi} \text{Arctan}(x).$$

Et donc
$$F_X(x) = \begin{cases} 0 & \text{si } x \leq 0 \\ \frac{2}{\pi} \text{Arctan}(x) & \text{si } x > 0 \end{cases}$$

- 2.b. X admet une espérance si et seulement si $\int_{-\infty}^{+\infty} t f(t) dt = \int_0^{+\infty} \frac{2}{\pi(1+t^2)} dt$ converge.

Mais au voisinage de $+\infty$, $t f(t) \underset{t \rightarrow +\infty}{\sim} \frac{2}{\pi} \frac{1}{t}$.

Puisque $\int_1^{+\infty} \frac{2}{\pi} \frac{dt}{t}$ diverge, par le critère des équivalents pour les fonctions positives,

$\int_1^{+\infty} t f(t) dt$ diverge également.

Et donc $\boxed{X \text{ n'admet pas d'espérance.}}$

Par conséquent, X n'admet pas non plus de moment d'ordre 2 et donc $\boxed{\text{pas de variance.}}$

- 3.a. Soit $x \in \mathbf{R}$. Alors

$$F_{M_n}(x) = P(M_n \leq x) = P\left(\bigcap_{i=1}^n [X_i \leq x]\right).$$

Mais par indépendance des X_i , il vient

$$P(M_n \leq x) = \prod_{i=1}^n P(X_i \leq x) = F_{X_1}(x)^n = \begin{cases} 0 & \text{si } x \leq 0 \\ \left(\frac{2}{\pi} \text{Arctan}(x)\right)^n & \text{si } x > 0 \end{cases}$$

- 3.b. Soit $g : u \mapsto \text{Arctan}(u) + \text{Arctan}\left(\frac{1}{u}\right)$. Alors g est dérivable sur $]0; +\infty[$, et

$$g'(u) = \frac{1}{1+u^2} - \frac{1}{u^2} \frac{1}{1+\frac{1}{u^2}} = \frac{1}{1+u^2} - \frac{1}{u^2+1} = 0.$$

Et donc g est constante sur l'intervalle¹ $]0; +\infty[$.

Puisque $g(1) = \text{Arctan}(1) + \text{Arctan}(1) = 2 \frac{\pi}{4} = \frac{\pi}{2}$, on en déduit que

$$\forall u \in]0; +\infty[, g(u) = \text{Arctan}(u) + \text{Arctan}\left(\frac{1}{u}\right) = \frac{\pi}{2}.$$

La fonction Arctan est de classe \mathcal{C}^1 , et donc par la formule de Taylor-Young à l'ordre 1,

$$\text{Arctan}(u) = \text{Arctan}(0) + \text{Arctan}'(0)u + \underset{u \rightarrow 0}{o}(u) = u + \underset{u \rightarrow 0}{o}(u) \underset{u \rightarrow 0}{\sim} u.$$

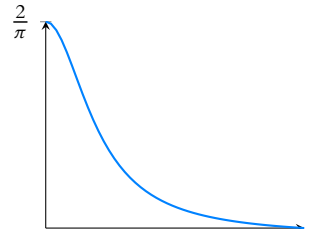


FIGURE 24.1– La densité f .

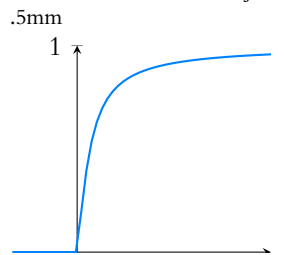


FIGURE 24.2– La fonction F_X .

Moments

Rappelons que si X admet un moment d'ordre p , alors elle admet un moment d'ordre q pour tout $q \leq p$.

Et donc par contraposée, si elle n'admet pas de moment d'ordre p , elle n'admet pas de moment d'ordre q , pour tout $q \geq p$.

¹ Notons que g est dérivable sur \mathbf{R}^* et que $g' = 0$ sur \mathbf{R}^* , mais \mathbf{R}^* n'est pas un intervalle, et donc ceci ne prouve pas que g y est constante.

Rappel

$$\tan \frac{\pi}{4} = 1 \Leftrightarrow \text{Arctan}(1) = \frac{\pi}{4}.$$

3.c. Soit $x \in]0, +\infty[$. Alors

$$\begin{aligned} P(Z_n \leq x) &= P\left(\frac{n}{M_n} \leq x\right) = P\left(M_n \geq \frac{n}{x}\right) \\ &= P\left(M_n > \frac{n}{x}\right) \\ &= 1 - P\left(M_n \leq \frac{n}{x}\right) \\ &= 1 - \left(\frac{2}{\pi} \operatorname{Arctan}\left(\frac{n}{x}\right)\right)^n. \end{aligned}$$

Mais en utilisant le résultat de la question 3.b,

$$\frac{2}{\pi} \operatorname{Arctan}\left(\frac{n}{x}\right) = 1 - \frac{2}{\pi} \operatorname{Arctan}\left(\frac{x}{n}\right)$$

de sorte que $P(M_n \leq x) = 1 - \left(1 - \frac{2}{\pi} \operatorname{Arctan}\left(\frac{x}{n}\right)\right)^n$.

3.d. Il est clair que Z_n est à valeurs positives, donc pour $x \leq 0$, $P(Z_n \leq x) = 0 \xrightarrow{n \rightarrow +\infty} 0$.
Pour $x > 0$, on a

$$\left(1 - \frac{2}{\pi} \operatorname{Arctan}\left(\frac{x}{n}\right)\right)^n = \exp\left(n \ln\left(1 - \frac{2}{\pi} \operatorname{Arctan}\left(\frac{x}{n}\right)\right)\right).$$

Mais $\frac{x}{n} \xrightarrow{n \rightarrow +\infty} 0$, et donc $\operatorname{Arctan}\left(\frac{x}{n}\right) \xrightarrow{n \rightarrow +\infty} 0$, de sorte que

$$\ln\left(1 - \frac{2}{\pi} \operatorname{Arctan}\left(\frac{x}{n}\right)\right) \underset{n \rightarrow +\infty}{\sim} -\frac{2}{\pi} \operatorname{Arctan}\left(\frac{x}{n}\right) \underset{n \rightarrow +\infty}{\sim} -\frac{2x}{\pi n}.$$

On en déduit que $n \ln\left(1 - \frac{2}{\pi} \operatorname{Arctan}\left(\frac{x}{n}\right)\right) \underset{n \rightarrow +\infty}{\sim} -\frac{2x}{\pi}$.

Et donc par continuité de l'exponentielle²,

$$\left(1 - \frac{2}{\pi} \operatorname{Arctan}\left(\frac{x}{n}\right)\right)^n = \exp\left(n \ln\left(1 - \frac{2}{\pi} \operatorname{Arctan}\left(\frac{x}{n}\right)\right)\right) \xrightarrow{n \rightarrow +\infty} e^{-\frac{2x}{\pi}}.$$

Ainsi, pour tout $x \in \mathbf{R}$, on a

$$\lim_{n \rightarrow +\infty} P(Z_n \leq x) = \begin{cases} 0 & \text{si } x \leq 0 \\ 1 - e^{-\frac{2x}{\pi}} & \text{si } x > 0 \end{cases}$$

Nous reconnaissons là la fonction de répartition d'une variable aléatoire suivant la loi exponentielle $\mathcal{E}\left(\frac{2}{\pi}\right)$.

Et donc si Z est une variable aléatoire suivant la loi exponentielle $\mathcal{E}\left(\frac{2}{\pi}\right)$, alors $Z_n \xrightarrow{\mathcal{L}} Z$.

PROBLÈME 1 (EDHEC 2018)

Partie 1 : simulations de S_k et T_k .

1. Puisque S_k est le rang d'apparition du $k^{\text{ème}}$ «Pile», il s'agit d'utiliser la variable c pour compter le nombre de «Pile».

Et pour simuler chacun des lancers, on peut simuler une loi uniforme sur $[0, 1]$ à l'aide de $\operatorname{rand}()$, qui prendra une valeur inférieure à p avec probabilité p .

Nous proposons donc le programme suivant :

Détails

M_n est une variable à densité car sa fonction de répartition, qui vaut $F_{X_1}^n$ est continue sur

\mathbf{R} et \mathcal{C}^1 sauf en un nombre fini de points.

Et donc il est possible de remplacer l'inégalité large par une inégalité stricte.

Danger !

Le piège classique : 1^∞ est une forme indéterminée, et il ne faudrait surtout pas conclure que $P(Z_n \leq x) \rightarrow 1$ au prétexte que

$$1 - \frac{2}{\pi} \operatorname{Arctan}\left(\frac{x}{n}\right) \rightarrow 1.$$

Il est indispensable de passer par la forme exponentielle pour lever l'indétermination.

² Ce qui nous permet de composer des limites, et non des équivalents !

Autrement dit

Chaque fois que $\operatorname{rand}()$ retourne un nombre inférieur à p , cela correspond à un «pile», et sinon à un «face».

```

1 k = input('donnez une valeur pour k :')
2 p = input('donnez une valeur pour p :')
3 n = 0
4 c = 0
5 while c < k
6     n = n + 1
7     if rand() < p then c = c + 1
8     end
9 end
10 disp(n)

```

2. Le programme proposé à la question précédente s'arrête dès que le $k^{\text{ème}}$ «pile» est obtenu, mais ces «piles» ne sont pas forcément consécutifs.

Pour que le programme s'arrête dès qu'on obtient k «piles» consécutifs, nous pouvons utiliser la variable c pour qu'elle compte le nombre de «piles» consécutifs, et donc la remettre à zéro à chaque «face».

7 `if rand() < p then c = c + 1, else c = 0 ;`

Le programme s'arrêtera donc lorsqu'on aura obtenu k «piles» consécutifs.

Partie 2 : calcul de l'espérance de S_k .

3. Puisque S_1 est le rang d'apparition du premier «pile», par indépendance des lancers, nous reconnaissons une loi géométrique de paramètre p . Et donc $E(S_1) = \frac{1}{p}$.

4.a. Les lancers étant indépendants, X_{n-1} suit une loi binomiale $\mathcal{B}(n-1, p)$.

4.b. Il faut au moins k lancers pour obtenir k «piles», et pour tout $n \geq k$, il est possible d'obtenir $n - k$ «faces» suivis de k «piles», de sorte que l'événement $[S_k = n]$ est possible.

Ainsi, $S_k(\Omega) = \llbracket k, +\infty \rrbracket = \{n \in \mathbf{N} : n \geq k\}$.

L'événement $[S_k = n]$ est réalisé si et seulement si le $n^{\text{ème}}$ lancer est un «pile» et qu'il y a eu exactement $k - 1$ «piles» lors des $n - 1$ premiers lancers.

Donc $[S_k = n] = [X_{n-1} = k - 1] \cap P_n$.

4.c. Puisque X_{n-1} ne dépend que des $n - 1$ premiers lancers, les événements $[S_{n-1} = k - 1]$ et P_n sont indépendants, de sorte que

$$\forall n \geq k, P(S_k = n) = P(X_{n-1} = k - 1)P(P_n) = \binom{n-1}{k-1} p^{k-1} q^{n-k} p = \boxed{\binom{n-1}{k-1} p^k q^{n-k}}.$$

5.a. La variable Z_i représente le nombre de lancers nécessaires à l'obtention du $i^{\text{ème}}$ pile après l'obtention du $(i - 1)^{\text{ème}}$.

Autrement dit, il s'agit du nombre de lancers nécessaires à l'obtention d'un pile. Et par indépendance des lancers, Z_i suit donc une loi géométrique de paramètre p .

En détails : l'explication précédente est largement suffisante au concours, mais peut donner l'impression qu'on s'en sort par une pirouette, sans vraiment justifier.

Appliquons la formule des probabilités totales avec le système complet d'événements $\{[S_{i-1} = n], n \geq i - 1\}$. Alors pour tout $k \geq 1$,

$$\begin{aligned}
 P(Z_i = k) &= \sum_{n=i-1}^{+\infty} P([Z_i = k] \cap [S_{i-1} = n]) \\
 &= \sum_{n=i-1}^{+\infty} P([S_i = n + k] \cap [S_{i-1} = n]) \\
 &= \sum_{n=i-1}^{+\infty} P([S_{i-1} = n] \cap F_{n+1} \cap F_{n+2} \cap \dots \cap F_{n+k-1} \cap P_{n+k}) \\
 &= \sum_{n=i-1}^{+\infty} P(S_{i-1} = n)P(F_{n+1})P(F_{n+2}) \cdots P(F_{n+k-1})P(P_{n+k})
 \end{aligned}$$

Les résultats des n premiers lancers sont indépendants de ceux des lancers suivants.

$$\begin{aligned}
 &= \sum_{n=i-1}^{+\infty} P(S_{i-1} = n) q^{n+k-1-(n+1)+1} p \\
 &= pq^{k-1} \sum_{n=i-1}^{+\infty} P(S_{i-1} = n) \\
 &= pq^{k-1}.
 \end{aligned}$$

Et donc on reconnaît bien une loi géométrique de paramètre p .

5.b. En faisant apparaître une somme télescopique, on a

$$S_k = (S_k - S_{k-1}) + (S_{k-1} - S_{k-2}) + \dots + (S_2 - S_1) + S_1 = \sum_{i=1}^k Z_i.$$

5.c. Puisque les Z_i ont une espérance, S_k possède aussi une espérance, et par linéarité de l'espérance,

$$E(S_k) = E\left(\sum_{i=1}^k Z_i\right) = \sum_{i=1}^k E(Z_i) = \frac{k}{p}.$$

Estimation

6.a. Les Z_i sont mutuellement indépendantes, suivent la même loi, et possèdent une espérance et une variance. Donc par la loi faible des grands nombres,

$$\overline{Z_k} \xrightarrow[k \rightarrow +\infty]{p} E(Z_1) = \frac{1}{p}.$$

6.b. Par continuité de la fonction $x \mapsto \frac{1}{x}, \frac{k}{S_k} = \frac{1}{\overline{Z_k}} \xrightarrow{p} \frac{1}{\frac{1}{p}} = p$.

6.c. Puisque S_{k-1} est à valeurs dans $\llbracket k-1, +\infty \rrbracket$, $\sum_{j=k-1}^{+\infty} P(S_{k-1} = j) = 1$.

D'après le théorème de transfert, sous réserve de convergence, on a

$$\begin{aligned}
 E\left(\frac{k-1}{S_k-1}\right) &= \sum_{i=k}^{+\infty} \frac{k-1}{i-1} P(S_k = i) \\
 &= \sum_{i=k}^{+\infty} \frac{k-1}{i-1} \binom{i-1}{k-1} p^k q^{i-k} \\
 &= \sum_{i=k}^{+\infty} \binom{i-2}{k-2} p^k q^{i-k} \\
 &= p \sum_{j=k-1}^{+\infty} \binom{j-1}{k-2} p^{k-1} q^{j-k} \\
 &= p \sum_{j=k-1}^{+\infty} P(S_{k-1} = j) = p.
 \end{aligned}$$

6.d. Puisque $S_{k+1} \geq S_k + 1$, il vient donc $\frac{k}{S_{k+1}-1} \leq \frac{k}{S_k}$.

Mais d'après la question précédente, $E\left(\frac{k}{S_k-1}\right) = p$ et donc par croissance de l'espérance,

$$E\left(\frac{k}{S_k}\right) \geq E\left(\frac{k}{S_k-1}\right) = p.$$

Il reste toutefois à constater que cette inégalité est stricte.

Mais S_{k+1} et $S_k + 1$ ne sont pas presque sûrement égales puisque $S_{k+1} - S_k = Z_{k+1}$ suit une loi géométrique, et donc n'est pas la variable certaine égale à 1.

La somme des probabilités d'un système complet d'événements vaut 1.

Remarque

En utilisant la loi de S_k et la définition de l'espérance, il aurait été bien difficile de calculer $E(S_k)$, sauf pour $k = 1$ et $k = 2$ où on aurait alors reconnu des séries géométriques dérivées.

Continuité

La fonction $x \mapsto \frac{1}{x}$ est continue sur \mathbf{R}_+^* , et $\overline{Z_k}$ prend ses valeurs dans \mathbf{R}_+^* (i.e. $P(\overline{Z_k} \in \mathbf{R}_+^*) = 1$), donc on peut appliquer le résultat rappelé dans l'énoncé.

Coeff. binomiaux

On a pour

$$\frac{k}{n} \binom{n}{k} = \binom{n-1}{k-1}.$$

Cette formule se retrouve sans grandes difficultés en revenant aux factorielles... à condition de savoir qu'une telle formule existe.

Convergence

Notons qu'on reconnaît ici une série convergente, ce qui garantit donc l'existence de $E\left(\frac{k-1}{S_k-1}\right)$.

Intuition

Si $S_k + 1$ et S_{k+1} étaient égales, cela signifierait qu'une fois que k « piles » sont obtenus, il suffit d'un lancer supplémentaire pour obtenir le « pile » suivant. Or ce lancer peut très bien donner un « face ».

Donc $\frac{k}{S_k} - \frac{k}{S_{k+1} - 1}$ est une variable positive, mais qui n'est pas presque sûrement nulle : son espérance est donc strictement positive³. Donc

$$E\left(\frac{k}{S_k} - \frac{k}{S_{k+1} - 1}\right) > 0 \Leftrightarrow E\left(\frac{k}{S_k}\right) > E\left(\frac{k}{S_{k+1} - 1}\right) = p.$$

Ceci prouve bien que $\frac{k}{S_k}$ est un estimateur biaisé de p , et même que son biais est positif.

³ Dans la somme qui définit l'espérance, au moins l'un des termes est strictement positif.

Partie 3 : calcul de l'espérance de T_k .

7. La variable T_1 désigne le rang d'apparition du dernier «pile» de la première série de **un** «pile». Autrement dit, c'est le rang du premier «pile», et donc $T_1 = S_1$.

8.a. Comme indiqué, considérons la variable W égale au rang d'apparition du premier «face». Alors $[W = 1], [W = 2], \dots, [W = k], [W > k]$ est un système complet d'événements.

Or, $[W = 1] = F_1, [W = 2] = P_1 \cap F_2, \dots, [W = k] = P_1 \cap \dots \cap P_{k-1} \cap F_k$ et enfin⁴ $[W > k] = P_1 \cap P_2 \cap \dots \cap P_k$.

Donc $F_1, P_1 \cap F_2, \dots, P_1 \cap \dots \cap P_{k-1} \cap F_k, P_1 \cap \dots \cap P_k$ est un système complet d'événements.

⁴ W est supérieur strictement à k si et seulement si les k premiers lancers donnent tous «pile».

8.b. Sachant que le premier lancer a donné «face», la première série de k «piles» consécutifs ne peut intervenir qu'à partir du second lancer.

Et donc la probabilité que T_k soit égal à n est la probabilité qu'il y ait besoin de $n - 1$ lancers⁵ pour obtenir la première série de k «piles» consécutifs. C'est donc $P(T_k = n - 1)$. Autrement dit, $P_{F_1}(T_k = n) = P(T_k = n - 1)$.

Notons que ceci est également vrai pour $n = k$: si le premier lancer donne «face», alors il n'est pas possible d'obtenir k «piles» en k lancers, et donc $P_{F_1}(T_k = k) = 0 = P(T_k = k - 1)$.

⁵ Entre le second et le $n^{\text{ème}}$, il y a $n - 1$ lancers.

Rédaction

Il serait en fait très désagréable de prouver ceci par le calcul. C'est toute la difficulté des probabilités : trouver le bon équilibre entre une intuition qu'on expliquera avec une phrase et une formalisation correcte à l'aide d'événements.

Si $n = k$, alors $P_{F_1}(T_k = n) = 0$.

Et alors, par définition de l'espérance conditionnelle,

$$\begin{aligned} E(T_k|F_1) &= \sum_{n=k}^{+\infty} nP_{F_1}(T_k = n) \\ &= \sum_{n=k+1}^{+\infty} nP_{F_1}(T_k = n) \\ &= \sum_{n=k+1}^{+\infty} nP(T_k = n - 1) \\ &= \sum_{i=k}^{+\infty} (i + 1)P(T_k = i) \\ &= \sum_{i=k}^{+\infty} iP(T_k = i) + \sum_{i=k}^{+\infty} P(T_k = i) = E(T_k) + 1. \end{aligned}$$

8.c. Sur le même principe, si $P_1 \cap \dots \cap P_{i-1} \cap F_i$ est réalisé, alors la première série de k «piles» consécutifs ne peut avoir lieu qu'à partir du $(i + 1)^{\text{ème}}$ lancer.

Et donc pour tout $n \geq k$, $P_{P_1 \cap \dots \cap P_{i-1} \cap F_i}(T_k = n) = P(T_k = n - i)$.

Il vient alors

$$\begin{aligned} E(T_k|P_1 \cap \dots \cap P_{i-1} \cap F_i) &= \sum_{n=k+i}^{+\infty} nP_{P_1 \cap \dots \cap P_{i-1} \cap F_i}(T_k = n) \\ &= \sum_{n=k+i}^{+\infty} nP(T_k = n - i) \\ &= \sum_{j=k}^{+\infty} (j + i)P(T_k = j) = i + E(T_k). \end{aligned}$$

Chgt d'indice

On a posé $j = n - i$ donc $n = i + j$.

8.d. Sachant que les k premiers lancers ont tous donné «pile», T_k est nécessairement égal à k . Autrement dit, la loi conditionnelle de T_k sachant $P_1 \cap \dots \cap P_k$ est la loi certaine égale à k . Et donc

$$E(T_k|P_1 \cap \dots \cap P_k) = k.$$

9.a. D'après la formule de l'espérance totale appliquée au système complet d'événements de la question 8.a, on a

$$\begin{aligned}
 E(T_k) &= P(F_1)E(T_k|F_1) + P(P_1 \cap F_2)E(T_k|P_1 \cap F_2) + \dots + P(P_1 \cap \dots \cap P_k)E(T_k|P_1 \cap \dots \cap P_k) \\
 &= \sum_{i=1}^k P(P_1 \cap \dots \cap P_{i-1} \cap F_i)(i + E(T_k)) + kP(P_1 \cap \dots \cap P_k) \\
 &= \sum_{i=1}^k p^{i-1}q(i + E(T_k)) + kp^k \\
 &= \boxed{\sum_{j=0}^{k-1} (j+1)p^j q + E(T_k) \sum_{j=0}^{k-1} p^j q + kp^k}.
 \end{aligned}$$

Chgt d'indice
 $j = i - 1.$

9.b. On a donc

$$E(T_k) \left(1 - \sum_{j=0}^{k-1} p^j q \right) = \sum_{j=0}^{k-1} (j+1)p^j q + kp^k.$$

Mais d'une part, nous savons que $\sum_{j=0}^{k-1} p^j q = q \frac{1-p^k}{1-p} = 1-p^k.$

D'autre part, pour tout $x \in]0, 1[$, on a $\sum_{i=0}^k x^i = \frac{1-x^{k+1}}{1-x}.$

En dérivant cette relation par rapport à x , il vient, pour tout $x \in]0, 1[$,

$$\sum_{i=1}^k ix^{i-1} = \frac{-(k+1)x^k(1-x) + 1 - x^{k+1}}{(1-x)^2} = \frac{1 - (k+1)x^k + kx^{k+1}}{(1-x)^2}.$$

Et donc en particulier, pour $x = p$,

$$\sum_{j=0}^{k-1} (j+1)p^j q = q \sum_{i=1}^k ip^{i-1} = \frac{1 - (k+1)p^k + kp^{k+1}}{q}.$$

On en déduit donc que

$$E(T_k)p^k = \frac{1 - (k+1)p^k + kp^{k+1}}{q} + kp^k = \frac{1 - (k+1)p^k + kp^{k+1} + (1-p)kp^k}{q} = \frac{1-p^k}{q}$$

et donc $E(T_k) = \frac{1-p^k}{qp^k}.$

Astuce
 Nous connaissons une formule pour les sommes partielles de séries géométriques. En dérivant une ou deux fois cette formule (ce qui est licite puisqu'on dérive alors une somme finie), on obtient des formules pour les sommes partielles de séries géométriques dérivées. C'est d'ailleurs ainsi qu'a été prouvée en cours la convergence des séries géométriques dérivées d'ordre 1 et 2.

10. Il est évident que lorsque la première série de k « piles » consécutifs s'achève, alors on a déjà obtenu au minimum k « piles ».

Et donc $S_k \leq T_k$. Par croissance de l'espérance, on a donc $E(S_k) \leq E(T_k)$.

Soit encore, pour tout $k \in \mathbf{N}^*$ et tout $p \in]0, 1[$,

$$\frac{k}{p} \leq \frac{1-p^k}{qp^k} \Leftrightarrow \underbrace{(1-p)kp^{k-1}}_{=q} \leq 1-p^k \Leftrightarrow kp^{k-1} - kp^k \leq 1-p^k \Leftrightarrow \boxed{kp^{k-1} - 1 \leq (k-1)p^k}.$$

PROBLÈME 2 (EML 2012)

Partie I : Interpolation polynomiale

1. Commençons par prouver que φ est linéaire : soient $P, Q \in \mathbf{R}_{n-1}[X]$, et soit $\lambda \in \mathbf{R}$. Alors

$$\begin{aligned}
 \varphi(\lambda P + Q) &= ((\lambda P + Q)(a_1), \dots, (\lambda P + Q)(a_n)) \\
 &= (\lambda P(a_1) + Q(a_1), \dots, \lambda P(a_n) + Q(a_n)) = \lambda(P(a_1), \dots, P(a_n)) + (Q(a_1), \dots, Q(a_n)) \\
 &= \lambda\varphi(P) + \varphi(Q).
 \end{aligned}$$

De plus, si $P \in \text{Ker}(\varphi)$, alors $P(a_1) = \dots = P(a_n) = 0$.

Donc P possède n racines distinctes. Or, P est de degré au plus $n-1$, et alors $P = 0$.

Nous venons donc de prouver que $\text{Ker} \varphi = \{0\}$, c'est-à-dire que φ est injectif.

Mais $\dim \mathbf{R}_{n-1}[X] = n = \dim \mathbf{R}^n$, et donc φ est un isomorphisme.

Rédaction
 Lorsque l'espace de départ et l'espace d'arrivée ne sont pas les mêmes, on oubliera pas de vérifier l'égalité des dimensions, sans quoi il n'est pas vrai qu'une application linéaire injective est un isomorphisme.

2. D'après ce qui précède, φ est surjectif, donc pour tout $(b_1, \dots, b_n) \in \mathbf{R}^n$, il existe un unique $P \in \mathbf{R}_{n-1}[X]$ tel que $\varphi(P) = (b_1, \dots, b_n)$, c'est-à-dire tel que

$$\forall i \in \llbracket 1, \dots, n \rrbracket, P(a_i) = b_i.$$

3. Soient $a, b, c, d \in \mathbf{R}$ tels que $P_0 = aX^3 + bX^2 + cX + d$. Alors

$$\begin{cases} P_0(0) = 1 \\ P_0(1) = 3 \\ P_0(2) = 11 \\ P_0(3) = 31 \end{cases} \Leftrightarrow \begin{cases} d = 1 \\ a + b + c + d = 3 \\ 8a + 4b + 2c + d = 11 \\ 27a + 9b + 3c + d = 31 \end{cases} \Leftrightarrow \begin{cases} d = 1 \\ a + b + c = 2 \\ 8a + 4b + 2c = 10 \\ 27a + 9b + 3c = 30 \end{cases} \Leftrightarrow \begin{cases} a = 1 \\ b = 0 \\ c = 1 \\ d = 1 \end{cases}$$

Ainsi, l'unique polynôme vérifiant les conditions demandées est $X^3 + X + 1$.

Partie II : polynômes spéciaux

4. $P = X^2$ convient, car $\forall x > 0$, on a $P(x) = x^2 > 0$ et $P'(x) = 2x > 0$.
 5. Soient $P, Q \in E$, et soit $\lambda \in \mathbf{R}_+^*$. Alors $\forall x \in]0, +\infty[$, on a

$$\alpha P(x) > 0 \text{ et } (\alpha P)'(x) = \alpha P'(x) > 0 \text{ donc } \alpha P \in E.$$

$$(P + Q)(x) = P(x) + Q(x) > 0 \text{ et } (P + Q)'(x) = P'(x) + Q'(x) > 0 \text{ donc } P + Q \in E.$$

$$(PQ)(x) = P(x)Q(x) > 0 \text{ et } (PQ)'(x) = P'(x)Q(x) + P(x)Q'(x) > 0 \text{ donc } PQ \in E.$$

Toutefois, E n'est pas un sous-espace vectoriel de $\mathbf{R}[X]$, par exemple car le polynôme nul n'est pas dans E .

6. Nous savons que P_1 est une primitive de P , et que $P_1(0) = 0$. En particulier, P_1 est un polynôme⁶.

En particulier, pour tout $x > 0$, $P_1'(x) = P(x) > 0$ car $P \in E$. Par conséquent, P_1 est strictement croissant sur $]0, +\infty[$, et $P_1(0) = 0$, donc pour $x > 0$, on a $P_1(x) > P_1(0) = 0$.

Ainsi, on a bien $P_1 \in E$.

7. Puisque $P \in E$, on a $P'(x) > 0$ pour $x \in]0, +\infty[$, de sorte que P est strictement croissant sur $]0, +\infty[$.

Et donc $\forall x \in]0, +\infty[, P(x) \geq P(0)$.

8. \tilde{P} est continue et strictement croissante car elle est dérivable, et sa dérivée est $\tilde{P}'(x) = P'(x) > 0$ sur $]0, +\infty[$.

De plus, on a $\tilde{P}(0) = P(0)$ et $\lim_{x \rightarrow +\infty} P(x) = +\infty$. En effet, \tilde{P} est polynomiale et non constante (car P' est non nul), et donc sa limite en $+\infty$ est soit $-\infty$, soit $+\infty$. Mais puisque $\forall x > 0, \tilde{P}(x) > 0$, on ne saurait avoir $\lim_{x \rightarrow +\infty} \tilde{P}(x) = -\infty$. Et donc $\lim_{x \rightarrow +\infty} \tilde{P}(x) = +\infty$.

Alors par le théorème de la bijection, \tilde{P} réalise une bijection de $]0, +\infty[$ sur $[P(0), +\infty[$.

9. Raisonnons par l'absurde en supposant que \tilde{P}^{-1} soit une application polynomiale, et notons

$$\tilde{P}^{-1} = \sum_{k=0}^d a_k X^k \text{ où } d = \deg \tilde{P}^{-1}, \text{ et donc } a_d \neq 0.$$

De même, notons $\tilde{P} = \sum_{i=0}^{\ell} b_i X^i$, avec $b_{\ell} \neq 0$.

Puisque \tilde{P}^{-1} est bijective, \tilde{P}^{-1} ne saurait être une constante, et donc $d \geq 1$.

Ainsi, on a

$$\forall x \in]P(0), +\infty[, \tilde{P}(\tilde{P}^{-1}(x)) = x.$$

Mais

$$\tilde{P}(\tilde{P}^{-1}(x)) = \sum_{i=0}^{\ell} b_i \left(\sum_{k=0}^d a_k x^k \right)^i.$$

En développant $\left(\sum_{k=0}^d a_k x^k \right)^{\ell}$, on obtient un terme de degré $\ell \times d$ qui est $a_d^{\ell} b_{\ell} X^{\ell d}$. Et tous les autres termes sont de degré inférieur strictement à ℓd .

Explication

E étant stable par addition et par multiplication par un réel positif, ce qu'il lui manque pour être un espace vectoriel est la stabilité par la multiplication par un réel négatif : si $\alpha \leq 0$ et si $P \in E$, alors $\alpha P \notin E$.

⁶ Car primitive d'un polynôme.

Et donc $\tilde{P}(\tilde{P}^{-1})$ est un polynôme de degré $\ell d \geq 2$ (car $\ell \geq 2$ et $d \geq 1$).
 Mais $\tilde{P}(\tilde{P}^{-1}(x)) = x$, donc $\tilde{P}(\tilde{P}^{-1})$ est de degré 1 : contradiction.
 Ainsi, \tilde{P}^{-1} n'est pas une application polynomiale.

Remarque
 Notons que si P est de degré 1, de la forme $aX + b$, alors $\tilde{P}^{-1} = \frac{X - b}{a}$ qui est de degré 1.

Partie III : Matrices symétriques positives

10. A est symétrique réelle, donc diagonalisable.

11.a. Soit λ une valeur propre de A , et soit X un vecteur propre associé.
 Alors ${}^tXAX \geq 0$ par hypothèse car $A \in \mathbf{S}_n^+$.
 Mais d'autre part, ${}^tXAX = {}^tX\lambda X = \lambda {}^tXX = \lambda \|X\|^2$, où $\|X\|$ est la norme de X pour le produit scalaire canonique de $\mathcal{M}_{n,1}(\mathbf{R})$.
 Et puisque $X \neq 0$ (c'est un vecteur propre), alors $\|X\|^2 > 0$. Et donc il vient $\lambda \geq 0$.
 Ainsi, toutes les valeurs propres de A sont positives.

Rédaction
 Il est important de préciser que $\|X\|^2 \neq 0$, car si $\|X\|^2 = 0$, alors $\lambda \times 0 \geq 0$ ne permet pas de conclure quoi que ce soit quant au signe de λ .

11.b. Supposons à présent que toutes les valeurs propres de A soient positives.
 Puisque A est symétrique réelle, il existe une base orthonormée de $\mathcal{M}_{n,1}(\mathbf{R})$ formée de vecteur propres de A .
 Notons (X_1, \dots, X_n) une telle base, et soit λ_i la valeur propre de A associée à X_i .

Si $X \in \mathcal{M}_{n,1}(\mathbf{R})$ alors X s'écrit de manière unique $X = \sum_{i=1}^n \alpha_i X_i$. Et alors

$$\begin{aligned} {}^tXAX &= \left(\sum_{j=1}^n \alpha_j {}^tX_j \right) A \left(\sum_{i=1}^n \alpha_i X_i \right) = \left(\sum_{j=1}^n \alpha_j {}^tX_j \right) \left(\sum_{i=1}^n \alpha_i AX_i \right) \\ &= \left(\sum_{j=1}^n \alpha_j {}^tX_j \right) \left(\sum_{i=1}^n \alpha_i \lambda_i X_i \right) = \sum_{i=1}^n \sum_{j=1}^n \alpha_i \alpha_j \lambda_i {}^tX_j X_i. \end{aligned}$$

Mais la famille (X_1, \dots, X_n) étant orthonormée⁷, on a ${}^tX_j X_i = \begin{cases} 0 & \text{si } i \neq j \\ 1 & \text{si } i = j \end{cases}$. Et donc

⁷ Pour le produit scalaire canonique de $\mathcal{M}_{n,1}(\mathbf{R})$.

$${}^tXAX = \sum_{i=1}^n \lambda_i \alpha_i^2 \geq 0.$$

On en déduit donc que $A \in \mathbf{S}_n^+$.

Partie IV : Matrice symétrique positive solution d'une équation polynomiale spéciale

12.a. Si $P(S) = A$, alors A est un polynôme en S . Et puisque S commute avec tout polynôme en S , A et S commutent, de sorte que $AS = SA$.
 On en déduit que

$$\Delta D = (Q^{-1}SQ)(Q^{-1}AQ) = Q^{-1}SAQ = Q^{-1}ASQ = (Q^{-1}AQ)(Q^{-1}SQ) = D\Delta.$$

12.b. Notons $d_{i,j}$ les coefficients de D et $\delta_{i,j}$ les coefficients de Δ .
 Puisque D et Δ commutent, pour tout $(i, j) \in \llbracket 1, n \rrbracket^2$, on a

$$\sum_{k=1}^n d_{i,k} \delta_{k,j} = (D\Delta)_{i,j} = (\Delta D)_{i,j} = \sum_{k=1}^n \delta_{i,k} d_{k,j}.$$

Mais puisque $D = \text{Diag}(\lambda_1, \dots, \lambda_n)$, $d_{i,j} = \begin{cases} \lambda_i & \text{si } i = j \\ 0 & \text{sinon} \end{cases}$ et donc $\sum_{k=1}^n d_{i,k} \delta_{k,j} = \lambda_i \delta_{i,j}$.

De même $\sum_{k=1}^n \delta_{i,k} d_{k,j} = \delta_{i,j} \lambda_j$.

Ainsi, $\delta_{i,j} \lambda_i = \delta_{i,j} \lambda_j \Leftrightarrow (\lambda_i - \lambda_j) \delta_{i,j} = 0$. Si $i \neq j$, alors $\lambda_i - \lambda_j \neq 0$, et donc $\delta_{i,j} = 0$.

Ainsi, Δ est diagonale.

De plus, puisque $\Delta = QSQ^{-1}$, alors $S = Q^{-1}\Delta Q$, et Δ étant diagonale, ses coefficients diagonaux sont ses valeurs propres, qui sont également⁸ les valeurs propres de S , qui sont toutes positives ou nulles car $S \in \mathbf{S}_n^+$.

Détail
 Les valeurs propres de A sont deux à deux distinctes par hypothèse.

⁸ Car S et Δ sont semblables.

13. Soit S une solution de l'équation $P(S) = A$. Alors nous venons de prouver qu'il existe une matrice diagonale Δ , à coefficients diagonaux positifs, telle que $S = Q^{-1}\Delta Q$. Notons d_1, \dots, d_n les coefficients diagonaux de Δ . Alors

$$Q^{-1}DQ = A = P(S) = Q^{-1}P(\Delta)Q = Q^{-1} \begin{pmatrix} P(d_1) & & \\ & \ddots & \\ & & P(d_n) \end{pmatrix} Q.$$

Donc en multipliant à gauche par Q et à droite par Q^{-1} , il vient

$$\begin{pmatrix} P(d_1) & & \\ & \ddots & \\ & & P(d_n) \end{pmatrix} = D = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}$$

Alors $P(d_1) = \lambda_1, \dots, P(d_n) = \lambda_n$.

Puisque P réalise une bijection de \mathbf{R}^+ sur $[P(0), +\infty[$ et que tous les λ_i sont dans $[P(0), +\infty[$ par hypothèse, alors nécessairement $\forall i \in \llbracket 1, n \rrbracket, d_i = \tilde{P}^{-1}(\lambda_i)$.

Ainsi, il existe au plus une solution de l'équation $P(S) = A$, qui est

$$S = Q^{-1} \begin{pmatrix} \tilde{P}^{-1}(\lambda_1) & & \\ & \ddots & \\ & & \tilde{P}^{-1}(\lambda_n) \end{pmatrix} Q.$$

Inversement, si on pose $S = Q^{-1} \begin{pmatrix} \tilde{P}^{-1}(\lambda_1) & & \\ & \ddots & \\ & & \tilde{P}^{-1}(\lambda_n) \end{pmatrix} Q$, alors

$$\begin{aligned} P(S) &= Q^{-1}P \left(\begin{pmatrix} \tilde{P}^{-1}(\lambda_1) & & \\ & \ddots & \\ & & \tilde{P}^{-1}(\lambda_n) \end{pmatrix} \right) Q = Q^{-1} \begin{pmatrix} P(\tilde{P}^{-1}(\lambda_1)) & & \\ & \ddots & \\ & & P(\tilde{P}^{-1}(\lambda_n)) \end{pmatrix} Q \\ &= Q^{-1} \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix} Q = Q^{-1}DQ = A. \end{aligned}$$

Donc l'équation $P(S) = A$ admet au moins une solution.

Ainsi, l'équation $P(S) = A$ admet une unique solution qui est $S = Q^{-1} \begin{pmatrix} \tilde{P}^{-1}(\lambda_1) & & \\ & \ddots & \\ & & \tilde{P}^{-1}(\lambda_n) \end{pmatrix} Q$.

- 14.a. Pour $x > 0$, on a $P(x) = x^3 + x + 1 > 1 > 0$, et $P'(x) = 3x^2 + 1 > 0$. Donc $\boxed{P \in E}$.

- 14.b. Soit $\lambda \in \mathbf{R}$. Alors λ est valeur propre de A si et seulement si $\text{rg}(A - \lambda I_4) < 4$. Or,

$$\begin{aligned} A - \lambda I_4 &= \begin{pmatrix} 2 - \lambda & -1 & 0 & 0 \\ -1 & 2 - \lambda & 0 & 0 \\ 0 & 0 & 21 - \lambda & 10 \\ 0 & 0 & 10 & 21 - \lambda \end{pmatrix} \begin{matrix} \xleftrightarrow{L_1 \leftrightarrow L_2} \\ \xleftrightarrow{L_3 \leftrightarrow L_4} \end{matrix} \begin{pmatrix} -1 & 2 - \lambda & 0 & 0 \\ 2 - \lambda & -1 & 0 & 0 \\ 0 & 0 & 21 - \lambda & 10 \\ 0 & 0 & 10 & 21 - \lambda \end{pmatrix} \\ &\begin{matrix} \xleftrightarrow{L_2 \leftarrow L_2 + (2-\lambda)L_1} \\ \xleftrightarrow{L_4 \leftarrow 10L_4 - (21-\lambda)L_3} \end{matrix} \begin{pmatrix} -1 & 2 - \lambda & 0 & 0 \\ 0 & \lambda^2 - 4\lambda + 3 & 0 & 0 \\ 0 & 0 & 21 - \lambda & 10 \\ 0 & 0 & 0 & -341 + 42\lambda - \lambda^2 \end{pmatrix} \end{aligned}$$

Ainsi, λ est valeur propre de A si et seulement si $\lambda^2 - 4\lambda + 3 = 0$ ou $\lambda^2 - 42\lambda + 341 = 0$. Et donc après résolution de ces deux équations (polynomiales de degré 2 en λ), on obtient $\lambda \in \{3, 1, 11, 31\}$.

Donc on a $\boxed{\text{Spec}(A) = \{1, 3, 11, 31\}}$.

Puisque toutes ces valeurs propres sont positives⁹, on a bien $\boxed{A \in \mathbf{S}_4^+}$.

Autrement dit

S est diagonalisable et une base de vecteurs propres de A est également une base de vecteurs propres de S .

Synthèse

Cette phase est indispensable : nous n'avons pas encore prouvé qu'une telle solution existe, mais avons juste dit que si une solution existe, c'est celle donnée précédemment. Il s'agit donc de prouver qu'il s'agit bien d'une solution de l'équation.

Calculs

Notons que le discriminant de la seconde équation est $42^2 - 4 \times 341$, ce qui n'est pas bien difficile à calculer... à condition de savoir encore faire une multiplication à la main !
Ce discriminant vaut $400 = 20^2$.

⁹ Et que A est symétrique.

14.c. Soit $X = \begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix} \in \mathcal{M}_{4,1}(\mathbf{R})$. Alors $X \in E_1(A) \Leftrightarrow AX = X$. Soit encore

$$\begin{pmatrix} 2 & -1 & 0 & 0 \\ -1 & 2 & 0 & 0 \\ 0 & 0 & 21 & 10 \\ 0 & 0 & 10 & 21 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix} = \begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix} \Leftrightarrow \begin{cases} 2x - y = x \\ -x + 2y = y \\ 21z + 10t = z \\ 10z + 21t = t \end{cases} \Leftrightarrow \begin{cases} x = y \\ t = -2z \\ z = -2t \end{cases} \Leftrightarrow \begin{cases} x = y \\ z = 0 \\ t = 0 \end{cases}.$$

Et donc $X \in E_1(A) \Leftrightarrow \begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix} = \begin{pmatrix} x \\ x \\ 0 \\ 0 \end{pmatrix} = x \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix} \in \text{Vect} \left(\begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix} \right)$. Ainsi, $E_1(A) = \text{Vect} \left(\begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix} \right)$.

De même, en résolvant trois autres systèmes, on obtient

$$E_3(A) = \text{Vect} \left(\begin{pmatrix} 1 \\ -1 \\ 0 \\ 0 \end{pmatrix} \right), E_{11}(A) = \text{Vect} \left(\begin{pmatrix} 0 \\ 0 \\ 1 \\ -1 \end{pmatrix} \right), E_{31}(A) = \text{Vect} \left(\begin{pmatrix} 0 \\ 0 \\ 1 \\ 1 \end{pmatrix} \right).$$

Notons que chacun des vecteurs propres obtenu est de norme $\sqrt{2}$, et donc, pour chaque sous-espace propre, une base orthonormée est obtenue en divisant le vecteur propre calculé par $\sqrt{2}$.

Et en concaténant ces bases orthonormées des sous-espaces propres, on obtient une base de $\mathcal{M}_{4,1}(\mathbf{R})$ formée de vecteurs propres de A .

$$Q = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 & 0 & 0 \\ 1 & -1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & -1 & 1 \end{pmatrix}, D = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 \\ 0 & 0 & 11 & 0 \\ 0 & 0 & 0 & 31 \end{pmatrix}.$$

Q est alors orthogonale car matrice de passage de la base canonique de $\mathcal{M}_{4,1}(\mathbf{R})$, qui est orthonormée, à la base orthonormée de vecteurs propres de A . Et on a

$$A = QDQ^{-1} \text{ (avec } Q^{-1} = {}^tQ\text{)}.$$

14.d. L'unique solution à l'équation $P(S) = A$ est

$$Q \begin{pmatrix} \tilde{P}^{-1}(1) & 0 & 0 & 0 \\ 0 & \tilde{P}^{-1}(3) & 0 & 0 \\ 0 & 0 & \tilde{P}^{-1}(11) & 0 \\ 0 & 0 & 0 & \tilde{P}^{-1}(31) \end{pmatrix} Q.$$

Mais ici, P est le polynôme obtenu à la question 3, de sorte que nous savons que

$$P(0) = 1, P(1) = 3, P(2) = 11 \text{ et } P(3) = 31.$$

Autrement dit, nous connaissons $\tilde{P}^{-1}(1)$ (qui vaut 0), $\tilde{P}^{-1}(3)$ (qui vaut 1), etc. Et donc l'unique solution à l'équation $P(S) = A$ est

$$S = Q \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \end{pmatrix} Q^{-1} = \boxed{\frac{1}{2} \begin{pmatrix} 1 & -1 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ 0 & 0 & 5 & 1 \\ 0 & 0 & 1 & 5 \end{pmatrix}}.$$

Q ou Q^{-1} ?

Attention à l'énoncé : si $A = QDQ^{-1}$, c'est que Q doit être la matrice de passage de la base canonique à la base de vecteurs propres. Et donc c'est bien celle que nous avons écrite facilement, pas besoin de calculer P^{-1} (qui ici vaut tP).

CONCOURS BLANC 2 (05.03.19)

EXERCICE 1

On note E l'espace vectoriel \mathbf{R}^n , muni du produit scalaire canonique.

Pour toutes matrices colonnes X et Y de $\mathcal{M}_{n,1}(\mathbf{R})$, on note $\langle X, Y \rangle = {}^tXY$.

Pour toute matrice colonne X de $\mathcal{M}_{n,1}(\mathbf{R})$, on note $\|X\| = \sqrt{{}^tXX}$.

On considère u et v deux endomorphismes de E , et on note A et B leurs matrices respectives dans la base canonique de E .

Partie I

Soit $a \in \mathbf{R}$. On suppose **dans cette partie uniquement** que $n = 2$ et que les matrices de u et v dans la base canonique sont respectivement :

$$A = \frac{1}{1+a^2} \begin{pmatrix} 1 & a \\ a & a^2 \end{pmatrix} \text{ et } B = \frac{1}{2} \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}.$$

1. Vérifier que u et v sont des projecteurs.
2.
 - a. Vérifier que les endomorphismes u , v et $u \circ v$ sont tous de rang 1.
 - b. Vérifier que le vecteur $x_0 = (1, a)$ est un vecteur propre de $u \circ v$.
 - c. Déterminer le spectre de $u \circ v$.
3.
 - a. Montrer que les valeurs propres de $u \circ v$ appartiennent à l'intervalle $[0; 1]$.
 - b. Pour quelle(s) valeur(s) de a , $u \circ v$ est-il un projecteur ?

Partie II

On revient dans cette partie au cas général, où n désigne un entier tel que $n \geq 2$.

On suppose que u et v sont des projecteurs symétriques de E et on pose : $C = BAB$.

4. Montrer que pour tout $X \in \mathcal{M}_{n,1}(\mathbf{R})$:

$$\|BX\|^2 = \langle BX, X \rangle.$$

En déduire que pour tout $X \in \mathcal{M}_{n,1}(\mathbf{R})$:

$$\|BX\| \leq \|X\|.$$

5. Montrer que C est diagonalisable dans $\mathcal{M}_n(\mathbf{R})$.
6. Soit λ une valeur propre de C et X un vecteur propre associé.
 - a. Exprimer $\|ABX\|^2$ en fonction de λ et $\|X\|$.
 - b. En déduire que les valeurs propres de C sont réelles positives.
7. Soit μ une (éventuelle) valeur propre de AB non nulle, et X un vecteur propre associé.
 - a. Montrer que BX est un vecteur propre de C .
 - b. Montrer : $ABX = \mu AX$. En déduire que : $AX = X$.
 - c. Montrer que : $\langle X, BX \rangle = \mu \|X\|^2$.
8. Déduire des questions précédentes que le spectre de AB est inclus dans $[0; 1]$.

EXERCICE 2

On considère la suite $(u_n)_{n \geq 0}$ définie par :

$$u_0 = 0, u_1 = 1 \text{ et } \forall n \in \mathbf{N}, u_{n+2} = u_{n+1} + u_n,$$

et on note f la fonction définie sur \mathbf{R}^2 par :

$$\forall (x, y) \in \mathbf{R}^2, f(x, y) = 2x^3 - 6xy + 3y^2 - 6y.$$

On pose enfin $\varphi = \frac{1 + \sqrt{5}}{2}$.

1. Vérifier que $\varphi > 1$ et que les réels φ et $\frac{-1}{\varphi}$ sont les solutions de l'équation : $x^2 - x - 1 = 0$.
2.
 - a. Montrer que f est de classe \mathcal{C}^2 sur \mathbf{R}^2 .

- b. Montrer que les seuls points critiques de f sont $(\varphi, \varphi + 1)$ et $\left(-\frac{1}{\varphi}, \frac{1}{\varphi + 1}\right)$.
- c. Étudier la nature locale des points critiques de f .
3. Montrer que, pour tout entier $n \in \mathbf{N}$: $u_n u_{n+2} - u_{n+1}^2 = (-1)^{n+1}$.
4. a. Recopier et compléter la fonction Scilab suivante afin que, prenant en argument un entier $n \geq 2$, elle calcule et renvoie la valeur du terme u_n de la suite $(u_n)_{n \geq 0}$.

```

1  function u=suite(n)
2      v=0
3      w=1
4      for k=2 :n
5          .....
6          .....
7          .....
8      end
9      u=.....
10 endfunction

```

- b. Justifier qu'il existe des réels λ et μ , que l'on déterminera, tels que :

$$\forall n \in \mathbf{N}, u_n = \lambda \varphi^n + \mu \left(\frac{-1}{\varphi}\right)^n.$$

- c. En déduire que la suite $\left(\frac{u_{n+1}}{u_n}\right)_{n \geq 1}$ converge et déterminer sa limite

5. On considère pour tout $n \in \mathbf{N}^*$: $S_n = \sum_{k=1}^n \frac{(-1)^k}{u_k u_{k+1}}$.

- a. Montrer, sans chercher à calculer de somme, que la série de terme général $\frac{1}{u_n u_{n+1}}$ converge.
- b. En déduire que la suite $(S_n)_{n \geq 1}$ converge.
- c. En utilisant le résultat de la question 3, montrer que pour tout $n \in \mathbf{N}^*$:

$$S_{n+1} - S_n = \frac{u_n}{u_{n+1}} - \frac{u_{n+1}}{u_{n+2}}.$$

- d. Montrer que : $\varphi = 1 - \sum_{k=1}^{+\infty} \frac{(-1)^k}{u_k u_{k+1}}$.

PROBLÈME

Partie I : Formule de Stirling

Pour tout $n \in \mathbf{N}$, on définit $W_n = \int_0^{\pi/2} (\cos t)^n dt$.

- Calculer W_0 et W_1 .
- Montrer, que la suite $(W_n)_{n \in \mathbf{N}}$ est décroissante.
 - Montrer, pour tout entier n tel que $n \geq 0$: $W_n > 0$.
- Montrer, pour tout entier n tel que $n \geq 0$: $(n+2)W_{n+2} = (n+1)W_n$.
 - En déduire, pour tout entier n tel que $n \geq 0$: $(n+1)W_{n+1}W_n = W_1W_0$.
- Montrer, pour tout entier n tel que $n \geq 0$: $W_n \geq W_{n+1} \geq \frac{n+1}{n+2}W_n$.
 - En déduire : $W_{n+1} \underset{n \rightarrow +\infty}{\sim} W_n$ puis $W_n \underset{n \rightarrow +\infty}{\sim} \sqrt{\frac{\pi}{2n}}$.
- Montrer, pour tout entier n tel que $n \geq 0$: $W_{2n} = \frac{(2n)!}{2^{2n}(n!)^2} \frac{\pi}{2}$.

On note, pour tout entier n tel que $n \geq 1$: $A_n = \frac{1}{n!} n^n e^{-n} \sqrt{n}$.

On note, pour tout entier n tel que $n \geq 2$: $a_n = -1 - \left(n - \frac{1}{2}\right) \ln \left(1 - \frac{1}{n}\right)$.

6. Montrer que la série $\sum_{n \geq 2} a_n$ converge.
7. Montrer, pour tout entier n tel que $n \geq 2$: $a_n = \ln(A_n) - \ln(A_{n-1})$.
8. En déduire que la suite $(A_n)_{n \in \mathbb{N}^*}$ converge et que sa limite ℓ est strictement positive.
9. a. Justifier : $n! \underset{n \rightarrow +\infty}{\sim} \frac{1}{\ell} n^n e^{-n} \sqrt{n}$.
- b. En utilisant l'expression de W_{2n} à l'aide de factorielles, en déduire la valeur de ℓ et l'équivalent suivant :

$$n! \underset{n \rightarrow +\infty}{\sim} n^n e^{-n} \sqrt{2\pi n}.$$

Partie II : Étude de variables aléatoires

Soit un réel a strictement positif et la fonction $f_a : \mathbf{R} \rightarrow \mathbf{R}$ définie, pour tout réel x par

$$f_a(x) = \begin{cases} 0 & \text{si } x \leq 0 \\ \frac{x}{a^2} e^{-\frac{x^2}{2a^2}} & \text{si } x > 0 \end{cases}$$

10. Montrer que f_a est une densité.

On considère une variable aléatoire X admettant f_a comme densité.

11. Déterminer la fonction de répartition de la variable aléatoire X .
12. Montrer que la variable aléatoire X admet une espérance $E(X)$ et calculer $E(X)$.
13. Montrer que la variable aléatoire X admet une variance $V(X)$ et calculer $V(X)$.
14. a. On considère une variable aléatoire V suivant la loi uniforme sur l'intervalle $]0, 1]$. Montrer que la variable aléatoire $Z = a\sqrt{-2 \ln(V)}$ suit la même loi que la variable aléatoire X .
- b. Écrire une fonction SciLab nommée `simul`, prenant en entrée le paramètre $a > 0$ et retournant une simulation de la variable aléatoire X .

Pour tout entier n strictement positif, on considère l'expérience suivante : on dispose de n urnes initialement vides, numérotées de 1 à n et on dispose d'un grand stock de boules que l'on dépose une à une dans ces urnes. Pour chaque boule, on choisit au hasard, de façon équiprobable, l'urne dans laquelle la boule est déposée.

On note T_n le rang du premier tirage pour lequel une des urnes contiendra deux boules.

15. Compléter la fonction SciLab suivante pour qu'elle simule une réalisation de la variable aléatoire T_n :

```

1  function T = tirage(n)
2      urnes = zeros(1,n)
3      T = 1
4      choix = floor((rand()*n))+1
5      while .....
6          urnes(choix) = urnes(choix)+1
7          choix = floor((rand()*n))+1
8          T = .....
9      end
10 endfunction

```

16. a. Justifier : $P(T_n > n + 1) = 0$.
- b. Déterminer, pour tout entier k tel que $k \leq n$: $P(T_n > k)$.
- c. En déduire, pour tout $k \in \llbracket 2, n + 1 \rrbracket$, que $P(T_n = k) = \frac{n!(k-1)}{n^k(n-k+1)!}$.
- d. On souhaite écrire une fonction SciLab qui calcule $E(T_n)$ en fonction de n . Compléter la fonction suivante à cet effet :

```

1  function E = esperance(n)
2      facto = prod([1 :n])
3      fac = facto
4      somme = 0
5      puissance = n
6      for k=2 : (n+1)
7          puissance = ....
8          fac = ....
9          somme = somme + k*(k-1)/(puissance*fac)
10     end
11     E = facto*somme
12 endfunction

```

On considère la variable aléatoire $Y_n = \frac{T_n}{\sqrt{n}}$. On se propose d'étudier la convergence en loi de la suite de variables aléatoires $(Y_n)_{n \geq 2}$.

Soit $y \in [0; +\infty[$. On note k_n l'entier naturel égal à la partie entière de $y\sqrt{n}$. On a donc : $k_n \leq y\sqrt{n} < 1 + k_n$.

17. Justifier $P(Y_n > y) = P(T_n > k_n)$.

18. En utilisant 9.b, montrer $P(Y_n > y) \underset{n \rightarrow +\infty}{\sim} e^{-k_n} \left(1 - \frac{k_n}{n}\right)^{k_n - n}$.

19. a. Déterminer le développement limité d'ordre 2 de $t \mapsto -t + (t-1)\ln(1-t)$ en 0.

b. En déduire : $\lim_{n \rightarrow +\infty} \left(-k_n + (k_n - n) \ln \left(1 - \frac{k_n}{n}\right)\right) = -\frac{y^2}{2}$.

20. Montrer que la suite de variables aléatoires $(Y_n)_{n \geq 2}$ converge en loi vers une variable aléatoire à densité dont on précisera une densité.

CORRECTION DU CONCOURS BLANC 2

EXERCICE 1 (ÉCRICOME 2018)

Partie I.

$$1. \text{ On a } A^2 = \frac{1}{(1+a^2)^2} \begin{pmatrix} 1+a^2 & a+a^3 \\ a+a^3 & a^2+a^4 \end{pmatrix} = \frac{1}{1+a^2} \begin{pmatrix} 1 & a \\ a & a^2 \end{pmatrix} = A.$$

Et donc $u^2 = u$: u est un projecteur de E .

$$\text{De même, } B^2 = \frac{1}{4} \begin{pmatrix} 2 & 2 \\ 2 & 2 \end{pmatrix} = \frac{1}{2} \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} = B.$$

Et donc $v^2 = v$: v est un projecteur.

2.a. Les deux colonnes de A sont colinéaires, la deuxième étant égale à a fois la première.

Puisque A est non nulle, on en déduit que $\text{rg}(A) = 1$, et donc $\text{rg}(u) = 1$.

De même, les colonnes de B sont non nulles, et égales, donc $\text{rg}(B) = 1$, et donc $\text{rg}(v) = 1$.

$$\text{Enfin, } AB = \frac{1}{2(1+a^2)} \begin{pmatrix} 1+a & 1+a \\ a+a^2 & a+a^2 \end{pmatrix}.$$

Les colonnes de AB sont donc égales, donc $\text{rg}(AB) \leq 1$.

Si $a \neq -1$, ces colonnes sont non nulles, et donc $\text{rg}(AB) = 1$, de sorte que $\text{rg}(u \circ v) = 1$.

En revanche, si $a = -1$, alors $AB = 0$, et donc $\text{rg}(u \circ v) = 0$.

$$2.b. \text{ On a } AB \begin{pmatrix} 1 \\ a \end{pmatrix} = \frac{1}{2(1+a^2)} \begin{pmatrix} 1+a+a+a^2 \\ a+a^2+a^2+a^3 \end{pmatrix} = \frac{1+2a+a^2}{2(1+a^2)} \begin{pmatrix} 1 \\ a \end{pmatrix} = \frac{(1+a)^2}{2(1+a^2)} \begin{pmatrix} 1 \\ a \end{pmatrix}.$$

$$\text{Et donc } (u \circ v)(x_0) = \frac{(1+a)^2}{2(1+a^2)} x_0.$$

Puisque x_0 est non nul¹, ceci prouve que x_0 est un vecteur propre de $u \circ v$, associé à la

$$\text{valeur propre } \frac{(1+a)^2}{2(1+a^2)}.$$

2.c. Si $a = -1$, alors $u \circ v = 0$, et donc $\text{Spec}(u \circ v) = \{0\}$.

Si $a \neq -1$, nous venons de déterminer une valeur propre de $u \circ v$, qui est $\frac{(1+a)^2}{2(1+a^2)} \neq 0$.

D'autre part, $u \circ v$ étant de rang 1 < $\dim \mathbf{R}^2$, 0 est valeur propre de $u \circ v$.

Nous avons donc deux valeurs propres distinctes et $u \circ v$, et puisque $\dim \mathbf{R}^2 = 2$, il ne peut y

avoir d'autres valeurs propres. Ainsi, $\text{Spec}(u \circ v) = \left\{ 0, \frac{(1+a)^2}{2(1+a^2)} \right\}$.

3.a. Il est évident que $0 \in [0; 1]$.

Soit f la fonction définie sur \mathbf{R} par $f(x) = \frac{(1+x)^2}{2(1+x^2)}$. Alors f est dérivable sur \mathbf{R} et

$$f'(x) = \frac{2(1+x)(1+x^2) - 2x(1+x)^2}{2(1+x^2)^2} = \frac{1+x+x^2+x^3-x-2x^2-x^3}{(1+x^2)^2} = \frac{1-x^2}{(1+x^2)^2}.$$

On a donc $f'(x) \leq 0 \Leftrightarrow 1-x^2 \leq 0 \Leftrightarrow -1 \leq x \leq 1$.

Le tableau de variation de f est donc donné par :

x	$-\infty$		-1		1		$+\infty$
$f'(x)$		+	0	-	0	+	
$f(x)$	$\frac{1}{2}$	↗		1	↘		$\frac{1}{2}$
					0		

⚠ Danger !

Une matrice dont toutes les colonnes sont colinéaires peut encore être la matrice nulle, et donc être de rang 0. Pour affirmer qu'une telle matrice est de rang 1, il ne faut pas oublier de vérifier qu'elle est non nulle (i.e. qu'au moins l'une de ses colonnes est non nulle).

⚠ Erreur d'énoncé ?

Le concepteur du sujet semble avoir oublié ce cas.

¹ Même si $a = 0$, sa première coordonnée est non nulle.

Rappel

0 est valeur propre de $f \in \mathcal{L}(E)$ si et seulement si $\text{rg}(f) < \dim E$.

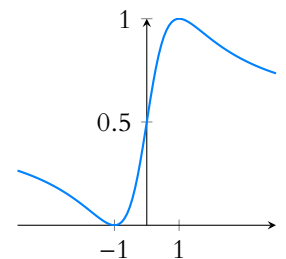


FIGURE 24.1– La fonction f .

Et alors, pour tout $a \in \mathbf{R}$, $f(a) = \frac{(1+a)^2}{2(1+a^2)} \in [0; 1]$.

On en déduit donc que $\text{Spec}(u \circ v) \subset [0; 1]$.

- 3.b. Si $a = -1$, alors $u \circ v = 0$, qui est bien un projecteur.
 Pour $a \neq -1$, $u \circ v$ possède deux valeurs propres distinctes, donc est diagonalisable.
 C'est un projecteur si et seulement si ses valeurs propres sont dans $\{0, 1\}$.
 C'est le cas si et seulement si $a = 1$.
 Donc $u \circ v$ est un projecteur si et seulement si $a = \pm 1$.

Partie II

Notons que u et v étant à la fois des projecteurs et des endomorphismes symétriques, ce sont des projecteurs orthogonaux. Ceci ne nous sera toutefois pas nécessaire dans la suite de l'exercice.

4. Puisque v est un endomorphisme symétrique, et que la base canonique de \mathbf{R}^n est orthonormée pour le produit scalaire canonique, B , qui est la matrice de v dans la base canonique est une matrice symétrique.
 Et donc pour $X \in \mathcal{M}_{n,1}(\mathbf{R})$, on a

$$\|BX\|^2 = {}^t(BX)(BX) = {}^tX{}^tBBX = {}^tXBBX = {}^tXBX = \langle X, BX \rangle = \langle BX, X \rangle.$$

On a alors, d'après l'inégalité de Cauchy-Schwarz,

$$\|BX\|^2 = \langle BX, X \rangle \leq \|BX\| \cdot \|X\|.$$

Si $\|BX\| \neq 0$, alors, en divisant par $\|BX\|$, il vient $\|BX\| \leq \|X\|$.

Et dans le cas où $\|BX\| = 0$, il est évident que $\|BX\| \leq \|X\|$.

5. Comme précédemment, puisque u est symétrique, A est une matrice symétrique.
 Et donc ${}^tC = {}^t(BAB) = {}^tB{}^tA{}^tB = BAB = C$.
 Donc C est une matrice symétrique réelle : elle est donc diagonalisable dans $\mathcal{M}_n(\mathbf{R})$.

- 6.a. On a

$$\begin{aligned} \|ABX\|^2 &= {}^t(ABX)(ABX) = {}^tX{}^tB{}^tAABX = {}^tXBA{}^2BX = {}^tXBABX = {}^tXCX \\ &= {}^tX\lambda X = \lambda {}^tXX = \lambda \|X\|^2. \end{aligned}$$

- 6.b. D'une part, $\|ABX\|^2 \geq 0$ car un carré est toujours positif.
 D'autre part, $\|X\|^2 \geq 0$, et puisque X est un vecteur propre, il est non nul et donc $\|X\|^2 > 0$.

On en déduit que $\lambda = \frac{\|ABX\|^2}{\|X\|^2} \geq 0$.

Et donc toutes les valeurs propres de C sont positives.

- 7.a. On a $CBX = BABBX = BABX = B\mu X = \mu BX$.
 Donc si $BX \neq 0$, c'est un vecteur propre de C , associé à la valeur propre μ .
 Or, si on avait $BX = 0$, alors il viendrait $ABX = 0$, et donc $\mu X = 0$.
 Puisque μ et X sont tous deux non nuls, ceci ne peut se produire.
 Donc on a bien $BX \neq 0$, donc BX est un vecteur propre de C , associé à la valeur propre μ .

- 7.b. On a $ABX = \mu X$. En multipliant à droite par A , il vient $A{}^2BX = \mu AX$, mais $A{}^2 = A$, donc $ABX = \mu AX$.
 Puisque d'autre part, $ABX = \mu X$, on en déduit que $\mu AX = \mu X$.
 Et μ étant non nul, il vient $AX = X$.

- 7.c. D'après ce qui précède,

$$\langle X, BX \rangle = \langle AX, BX \rangle = {}^tX{}^tABX = {}^tXABX = {}^tX\mu X = \mu {}^tXX = \mu \|X\|^2.$$

8. En combinant les résultats des questions 4 et 7.c, il vient d'une part

$$\mu \|X\|^2 = \langle X, BX \rangle = \|BX\|^2 \geq 0$$

Détails

Un endomorphisme diagonalisable est un projecteur si et seulement si ses valeurs propres sont dans $[0, 1]$.
 En effet, dans une «bonne base», sa matrice est diagonale, donc est égale à son carré si et seulement si ses coefficients diagonaux (ses valeurs propres) sont égaux à leurs propres carrés.
 Seuls 0 et 1 vérifient ceci (ce sont les seules solutions de $x^2 - x = 0$).

Détails

Puisque v est un projecteur, $v^2 = v$ et donc $B^2 = B$.

Remarque

Il est vraiment important ici de remarquer que $\|X\| > 0$, et de ne pas se contenter de $\|X\| \geq 0$.
 En effet, si on avait $\|X\| = 0$, on saurait seulement que $\lambda \times 0 \geq 0$, ce qui ne nous renseigne aucunement sur le signe de λ .

et donc $\mu = \frac{\|BX\|^2}{\|X\|^2} \geq 0$.

D'autre part, $\mu\|X\|^2 = \|BX\|^2 \leq \|X\|^2$, et donc $\mu \leq 1$.

Ainsi, les valeurs propres non nulles de AB sont bien dans $[0; 1]$.

Et si jamais 0 est valeur propre de AB , elle est bien évidemment incluse dans $[0; 1]$.

Ainsi, toutes les valeurs propres de AB sont dans $[0; 1]$.

EXERCICE 2 (ÉCRICOME 2018)

1. Puisque $5 > 1$, $\sqrt{5} > \sqrt{1} = 1$, et donc $\varphi > \frac{1+1}{2} = 1$.

D'autre part, le discriminant de $x^2 - x - 1$ est $\Delta = 1 + 4 = 5$, de sorte que ses solutions sont

$$x_1 = \frac{1 + \sqrt{5}}{2} = \varphi \text{ et } x_2 = \frac{1 - \sqrt{5}}{2}.$$

$$\text{Mais } -\frac{1}{\varphi} = -\frac{2}{1 + \sqrt{5}} = -\frac{2(1 - \sqrt{5})}{(1 + \sqrt{5})(1 - \sqrt{5})} = -\frac{2 - 2\sqrt{5}}{1 - 5} = \frac{1 - \sqrt{5}}{2} = x_2.$$

Donc φ et $-\frac{1}{\varphi}$ sont bien les deux solutions de $x^2 - x - 1$.

2.a. Il suffit de remarquer que f est polynomiale sur \mathbf{R}^2 , donc elle est de classe \mathcal{C}^2 .

2.b. On a $\partial_1 f(x, y) = 6x^2 - 6y$ et $\partial_2 f(x, y) = -6x + 6y - 6$.

Et donc $(x, y) \in \mathbf{R}^2$ est un point critique de f si et seulement si

$$\begin{cases} 6x^2 - 6y = 0 \\ -6x + 6y - 6 = 0 \end{cases} \Leftrightarrow \begin{cases} x^2 = y \\ x^2 - x - 1 = 0 \end{cases}$$

D'après ce qui précède, on a alors $x = \varphi$ ou $x = -\frac{1}{\varphi}$.

Dans le premier cas, $y = \varphi^2 = \varphi + 1$, et dans le second cas, $y = \frac{1}{\varphi^2} = \frac{1}{\varphi + 1}$.

Et donc f possède deux points critiques, qui sont $(\varphi, \varphi + 1)$ et $(-\frac{1}{\varphi}, \frac{1}{\varphi + 1})$.

2.c. Commençons par noter que f ne possède ni maximum global ni minimum global car $\lim_{x \rightarrow +\infty} f(x, 0) = \lim_{x \rightarrow +\infty} 2x^3 = +\infty$ et de même $\lim_{x \rightarrow -\infty} f(x, 0) = -\infty$.

Ceci signifie donc que f peut prendre des valeurs arbitrairement grandes et arbitrairement petites, et donc ne possède pas d'extremum global.

Pour étudier la nature locale des points critiques, calculons les dérivées partielles secondes de f .

On a $\partial_{1,1}^2 f(x, y) = 12x$, $\partial_{1,2}^2 f(x, y) = \partial_{2,1}^2 f(x, y) = -6$, $\partial_{2,2}^2 f(x, y) = 6$.

En particulier, la hessienne de f en $(\varphi, \varphi + 1)$ est

$$\nabla^2 f(\varphi, \varphi + 1) = \begin{pmatrix} 12\varphi & -6 \\ -6 & 6 \end{pmatrix} = \begin{pmatrix} 6(1 + \sqrt{5}) & -6 \\ -6 & 6 \end{pmatrix} = 6 \begin{pmatrix} 1 + \sqrt{5} & -1 \\ -1 & 1 \end{pmatrix}.$$

Mais un réel λ est valeur propre de $A = \begin{pmatrix} 1 + \sqrt{5} & -1 \\ -1 & 1 \end{pmatrix}$ si et seulement si $A - \lambda I_2$ n'est pas inversible, soit si et seulement si

$$\det \left(\begin{pmatrix} 1 + \sqrt{5} & -1 \\ -1 & 1 \end{pmatrix} - \lambda I_2 \right) = 0 \Leftrightarrow \lambda^2 - (2 + \sqrt{5})\lambda + \sqrt{5} = 0.$$

Le discriminant de cette dernière équation est

$$\Delta = (2 + \sqrt{5})^2 - 4\sqrt{5} = 4 + 4\sqrt{5} + 5 - 4\sqrt{5} = 9 > 0.$$

Il y a donc deux valeurs propres, qui sont $\lambda_1 = \frac{2 + \sqrt{5} - 3}{2} = \frac{\sqrt{5} - 1}{2}$ et $\frac{5 + \sqrt{5}}{2}$.

Ces deux valeurs propres étant positives strictement, on en déduit que celles de $\nabla^2 f(\varphi, \varphi + 1)$

Méthode

Lorsqu'au dénominateur d'une fraction, on a une expression de la forme $a + b\sqrt{c}$, on multiplie numérateur et dénominateur par $a - b\sqrt{c}$ (la « quantité conjuguée »), afin de ne garder des racines carrées qu'au numérateur.

² Et même les deux seules !

Astuce

Puisque φ est solution de $x^2 - x - 1 = 0$, on a $\varphi^2 = \varphi + 1$, il n'est pas nécessaire de recalculer la valeur de φ^2 .

le sont également , et donc f possède en $(\varphi, \varphi + 1)$ un minimum local.

De la même manière, on a

$$\nabla^2 f \left(-\frac{1}{\varphi}, \frac{1}{\varphi + 1} \right) = \begin{pmatrix} -\frac{12}{\varphi} & -6 \\ -6 & 6 \end{pmatrix} = 6 \begin{pmatrix} 1 - \sqrt{5} & -1 \\ -1 & 1 \end{pmatrix}.$$

Or, un réel λ est valeur propre de $B = \begin{pmatrix} 1 - \sqrt{5} & -1 \\ -1 & 1 \end{pmatrix}$ si et seulement si $B - \lambda I_2$ n'est pas inversible soit si et seulement si

$$\det(B - \lambda I_2) = 0 \Leftrightarrow \lambda^2 - (2 - \sqrt{5})\lambda - \sqrt{5} = 0.$$

Le discriminant de ce polynôme de degré 2 en λ est $\Delta = (2 - \sqrt{5})^2 + 4\sqrt{5} = 9$.

Ses racines, qui sont les valeurs propres de B , sont donc $\lambda_1 = \frac{2 - \sqrt{5} - 3}{2} = -\varphi < 0$ et $\lambda_2 = \frac{5 - \sqrt{5}}{2} > 0$.

Les valeurs propres de B étant de signes opposés, il en est de même de celles de $\nabla^2 f \left(-\frac{1}{\varphi}, \frac{1}{\varphi + 1} \right)$,

et donc $\left(-\frac{1}{\varphi}, \frac{1}{\varphi + 1} \right)$ est un point selle de f .

3. Montrons par récurrence sur n que $u_n u_{n+2} - u_n^2 = (-1)^{n+1}$.
 Pour $n = 0$, on a $u_0 = 0$ et $u_1 = u_2 = 1$, de sorte que

$$u_0 u_2 - u_1^2 = 0 - 1 = -1 = (-1)^{0+1}.$$

Supposons que $u_n u_{n+2} - u_{n+1}^2 = (-1)^{n+1}$. Alors

$$\begin{aligned} u_{n+1} u_{n+3} - u_{n+2}^2 &= u_{n+1} (u_{n+1} + u_{n+2}) - u_{n+2}^2 \\ &= u_{n+1}^2 + u_{n+1} u_{n+2} - u_{n+2}^2 \\ &= u_n u_{n+2} - (-1)^{n+1} + u_{n+1} u_{n+2} - u_{n+2}^2 \\ &= \underbrace{(u_n + u_{n+1})}_{=u_{n+2}} u_{n+2} - u_{n+2}^2 + (-1)^{n+2} = (-1)^{n+1+1}. \end{aligned}$$

Hypothèse de récurrence.

Ainsi, par le principe de récurrence, pour tout $n \in \mathbf{N}$, $u_n u_{n+2} - u_{n+1}^2 = (-1)^{n+1}$.

- 4.a. Nous allons calculer successivement les valeurs de u_k , en utilisant toujours la variable w pour stocker la dernière valeur de u_k calculée, et v pour stocker l'avant dernière valeur calculée.

```

1  function u=suite(n)
2      v=0
3      w=1
4      for k=2 :n
5          a = v+w
6          v =w
7          w = a
8      end
9      u = w
10 endfunction
    
```

Notons que nous avons utilisé une autre variable appelée a pour stocker temporairement la valeur de w , car si nous avions directement utilisé $w = v+w$, alors nous aurions perdu la valeur de w .

- 4.b. La suite (u_n) est une suite récurrente linéaire d'ordre 2, dont le polynôme caractéristique est $X^2 - X - 1$, dont les racines ont été calculées précédemment : il s'agit de φ et $-\frac{1}{\varphi}$.

Détails
 Si on multiplie une matrice par 6, ses valeurs propres sont également multipliées par 6.

Ces racines étant distinctes, il existe deux réels λ et μ tels que $\forall n \in \mathbf{N}, u_n = \lambda\varphi^n + \mu\left(-\frac{1}{\varphi}\right)^n$.

Afin de déterminer les valeurs de λ et μ , il suffit d'utiliser les valeurs de u_0 et u_1 :

$$\begin{cases} u_0 = 0 \\ u_1 = 1 \end{cases} \Leftrightarrow \begin{cases} \lambda + \mu = 0 \\ \lambda\varphi - \frac{\mu}{\varphi} = 1 \end{cases} \Leftrightarrow \begin{cases} \mu = -\lambda \\ \lambda\left(\varphi + \frac{1}{\varphi}\right) = 1 \end{cases}$$

Or, nous savons déjà que $-\frac{1}{\varphi} = \frac{1-\sqrt{5}}{2}$, de sorte que $\varphi + \frac{1}{\varphi} = \frac{1+\sqrt{5}}{2} - \frac{1-\sqrt{5}}{2} = \sqrt{5}$.

Et donc $\lambda = \frac{1}{\sqrt{5}}$, $\mu = -\lambda = -\frac{1}{\sqrt{5}}$. Ainsi, pour tout $n \in \mathbf{N}$,

$$u_n = \frac{1}{\sqrt{5}} \left(\varphi^n - \left(-\frac{1}{\varphi}\right)^n \right).$$

4.c. On a $\frac{\left(-\frac{1}{\varphi}\right)^n}{\varphi^n} = \left(\frac{-1}{\varphi^2}\right)^n \xrightarrow{n \rightarrow +\infty} 0$ car $\varphi > 1$ et donc $\left|\frac{-1}{\varphi^2}\right| < 1$.

Et donc $\left(-\frac{1}{\varphi}\right)^n = o_{n \rightarrow +\infty}(\varphi^n)$ de sorte que $u_n \underset{n \rightarrow +\infty}{\sim} \frac{1}{\sqrt{5}}\varphi^n$.

Et donc $u_{n+1} \underset{n \rightarrow +\infty}{\sim} \frac{1}{\sqrt{5}}\varphi^{n+1}$, de sorte que

$$\frac{u_{n+1}}{u_n} \underset{n \rightarrow +\infty}{\sim} \frac{\varphi^{n+1}}{\varphi^n} = \varphi \text{ et donc } \frac{u_{n+1}}{u_n} \xrightarrow{n \rightarrow +\infty} \varphi.$$

5.a. D'après ce qui précède, on a $\frac{1}{u_n u_{n+1}} \underset{n \rightarrow +\infty}{\sim} \frac{\sqrt{5}}{\varphi^n} \frac{\sqrt{5}}{\varphi^{n+1}} \underset{n \rightarrow +\infty}{\sim} \frac{5}{\varphi^{2n+1}} \underset{n \rightarrow +\infty}{\sim} \frac{5}{\varphi} \left(\frac{1}{\varphi^2}\right)^n$.

Mais $0 < \frac{1}{\varphi^2} < 1$, de sorte que la série de terme général $\left(\frac{1}{\varphi^2}\right)^n$ est une série géométrique convergente.

Et donc, par critère de comparaison pour les séries à termes positifs, la série de terme général $\frac{1}{u_n u_{n+1}}$ converge.

5.b. Puisque $\left|\frac{(-1)^k}{u_k u_{k+1}}\right| = \frac{1}{u_k u_{k+1}}$, nous venons de prouver que la série de terme général $\frac{(-1)^k}{u_k u_{k+1}}$ converge absolument.

Et donc en particulier elle converge. Par définition, cela signifie que la suite de ses sommes partielles est une suite convergente.

Or S_n est précisément la somme partielle d'ordre n de la série $\sum_{k \geq 1} \frac{(-1)^k}{u_k u_{k+1}}$, et donc la suite $(S_n)_{n \geq 1}$ converge.

5.c. D'une part, on a

$$S_{n+1} - S_n = \sum_{k=1}^{n+1} \frac{(-1)^k}{u_k u_{k+1}} - \sum_{k=1}^n \frac{(-1)^k}{u_k u_{k+1}} = \frac{(-1)^{n+1}}{u_n u_{n+1}}.$$

D'autre part, en utilisant le résultat de la question 3,

$$\frac{u_n}{u_{n+1}} - \frac{u_{n+1}}{u_{n+2}} = \frac{u_n u_{n+2} - u_{n+1}^2}{u_n u_{n+1}} = \frac{(-1)^{n+1}}{u_n u_{n+1}} = S_{n+1} - S_n.$$

5.d. Sommons les relations de la question précédente :

$$\sum_{k=1}^n (S_{k+1} - S_k) = \sum_{k=1}^n \left(\frac{u_k}{u_{k+1}} - \frac{u_{k+1}}{u_{k+2}} \right).$$

Nous reconnaissons alors deux sommes télescopiques :

$$\sum_{k=1}^n (S_{k+1} - S_k) = \sum_{k=1}^n S_{k+1} - \sum_{k=1}^n S_k = \sum_{i=2}^{n+1} S_i - \sum_{k=1}^n S_k = S_{n+1} - S_1 = S_{n+1} + 1.$$

Rappel

$u_n \approx v_n$ si et seulement si $u_n = v_n + o(v_n)$.

Rédaction

Ne pas oublier de mentionner la positivité (ou au moins le signe constant) lorsqu'on utilise un équivalent pour déterminer la nature des séries : le résultat n'est plus vrai pour des séries qui changent de signe !

Le mentionner permet de montrer au correcteur qu'on connaît bien les hypothèses.

Et de même,

$$\sum_{k=1}^n \left(\frac{u_k}{u_{k+1}} - \frac{u_{k+1}}{u_{k+2}} \right) = \frac{u_1}{u_2} - \frac{u_{n+1}}{u_{n+2}} = 1 - \frac{u_{n+1}}{u_{n+2}}.$$

En utilisant le résultat de 4.c, on a donc $S_{n+1} = -\frac{u_{n+1}}{u_{n+2}} \xrightarrow{n \rightarrow +\infty} -\frac{1}{\varphi}$.

Et donc

$$\sum_{k=1}^{+\infty} \frac{(-1)^k}{u_k u_{k+1}} = -\frac{1}{\varphi} = \frac{1 - \sqrt{5}}{2} = 1 - \frac{1 + \sqrt{5}}{2} = 1 - \varphi.$$

On en déduit enfin que

$$\varphi = 1 - \sum_{k=1}^{+\infty} \frac{(-1)^k}{u_k u_{k+1}}.$$

Quelques commentaires : la suite (u_n) est très classique, et est appelée suite de Fibonacci et doit son nom à Leonardo Fibonacci qui s'y est intéressé dès le XIII^{ème} siècle, via l'amusant problème suivant : «Un homme met un couple de lapins dans un lieu isolé de tous les côtés par un mur. Combien de couples obtient-on en un an si chaque couple engendre tous les mois un nouveau couple à compter du troisième mois de son existence ?»

Si on note x_n le nombre de couples de lapins au début du $n^{\text{ème}}$ mois, on a donc $x_0 = 0$ et $x_1 = 1$. De plus, au début du $n^{\text{ème}}$ mois, le nombre de couples est égal au nombre de couples déjà présents le mois précédent (x_{n-1}) , plus le nombre de couples nés depuis.

Mais puisqu'il faut deux mois à un couple pour atteindre la maturité sexuelle, le nombre de couples qui viennent de naître est égal au nombre de couples âgés de plus de deux mois, c'est-à-dire au nombre de couples qui étaient déjà nés deux mois avant : c'est x_{n-2} .

Et ainsi, on a la relation

$$x_n = x_{n-1} + x_{n-2} \Leftrightarrow x_{n+1} = x_n + x_{n-1}.$$

Cette suite a passionné des générations de mathématiciens, et il existe toujours une revue mathématique³ qui ne publie que des articles traitant des propriétés de (x_n) .

³ Le Fibonacci Quarterly.

PROBLÈME (EML 2012 + SciLab EXTRAIT D'ERICOME 2017)

Partie I : Formule de Stirling

1. On a $W_0 = \int_0^{\pi/2} 1 dt = \left[\frac{t}{1} \right]_0^{\pi/2} = \frac{\pi}{2}$ et $W_1 = \int_0^{\pi/2} \cos t dt = [\sin t]_0^{\pi/2} = 1$.

2.a. Pour $t \in [0, \pi/2]$, on a $0 \leq \cos t \leq 1$. Et donc $(\cos t)^{n+1} \leq (\cos t)^n$. Par croissance de l'intégrale, il vient alors

$$W_{n+1} = \int_0^{\pi/2} (\cos t)^{n+1} dt \leq \int_0^{\pi/2} (\cos t)^n dt = W_n.$$

Ainsi, la suite $(W_n)_n$ est décroissante.

2.b. La fonction $t \mapsto (\cos t)^n$ est continue et positive sur $[0, \pi/2]$. Comme elle est non nulle sur $\left] 0, \frac{\pi}{2} \right[$, son intégrale W_n est strictement positive.

3.a. Calculons $(n+2)W_{n+2}$ par intégration par parties en posant $u(t) = (\cos t)^{n+1}$ et $v(t) = \sin t$, qui sont \mathcal{C}^1 sur $[0, \pi/2]$, avec $u'(t) = -(n+1) \sin t (\cos t)^n$ et $v'(t) = \cos t$. Alors

$$W_{n+2} = [\sin t (\cos t)^{n+1}]_0^{\pi/2} + \int_0^{\pi/2} (n+1) \sin^2 t (\cos t)^n dt = (n+1) \int_0^{\pi/2} \sin^2 t (\cos t)^n dt.$$

Mais $\sin^2 t = 1 - \cos^2 t$, et donc

$$W_{n+2} = (n+1) \int_0^{\pi/2} (\cos t)^n dt - (n+1) \int_0^{\pi/2} (\cos t)^{n+2} dt = (n+1)W_n - (n+1)W_{n+2}.$$

Soit encore $(n+2)W_{n+2} = (n+1)W_n$.

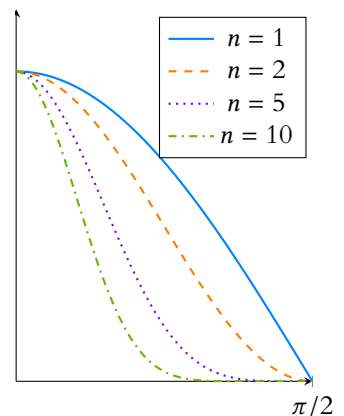


FIGURE 24.2– Les fonctions $t \mapsto (\cos t)^n$ pour quelques valeurs de n . On y voit bien que l'aire sous les courbes est de plus en plus petite, et donc que (W_n) est décroissante.

3.b. D'après le résultat de la question précédente, on a, pour tout $n \in \mathbf{N}$,

$$(n+2)W_{n+2}W_{n+1} = (n+1)W_nW_{n+1} = (n+1)W_{n+1}W_n.$$

Donc la suite $((n+1)W_{n+1}W_n)_{n \geq 0}$ est constante.

Puisque son premier terme vaut W_1W_0 , on a donc, pour tout $n \in \mathbf{N}$, $(n+1)W_{n+1}W_n = W_1W_0$.

4.a. Puisque (W_n) est décroissante, on a évidemment $W_n \geq W_{n+1}$.

De même on a $W_{n+1} \geq W_{n+2}$, et par la question 3.a, il vient alors $W_{n+1} \geq \frac{n+1}{n+2}W_n$.

On a donc prouvé que

$$W_n \geq W_{n+1} \geq \frac{n+1}{n+2}W_n.$$

4.b. On a $1 \geq \frac{W_{n+1}}{W_n} \geq \frac{n+1}{n+2}$, donc par le théorème des gendarmes, $\frac{W_{n+1}}{W_n} \xrightarrow[n \rightarrow +\infty]{} 1$.

Et donc $W_{n+1} \underset{n \rightarrow +\infty}{\sim} W_n$.

De la question 3.b, on déduit que

$$\frac{\pi}{2} = W_0W_1 = (n+1)W_{n+1}W_n \underset{n \rightarrow +\infty}{\sim} (n+1)W_n^2 \underset{n \rightarrow +\infty}{\sim} nW_n^2.$$

Et donc $nW_n^2 \underset{n \rightarrow +\infty}{\sim} \frac{\pi}{2}$ de sorte que $W_n \underset{n \rightarrow +\infty}{\sim} \sqrt{\frac{\pi}{2n}}$.

5. Montrons par récurrence sur n que $W_{2n} = \frac{(2n)!}{2^{2n}(n!)^2} \frac{\pi}{2}$.

Pour $n = 0$, on a déjà $W_0 = \frac{\pi}{2}$, donc la propriété est vraie.

Supposons donc que $W_{2n} = \frac{(2n)!}{2^{2n}(n!)^2} \frac{\pi}{2}$. Alors, par la question 3.a,

$$\begin{aligned} W_{2(n+1)} &= \frac{2n+1}{2n+2} W_{2n} = \frac{2n+1}{2n+2} \frac{(2n)!}{2^{2n}(n!)^2} \frac{\pi}{2} \\ &= \frac{(2n+1)(2n+2)}{(2n+2)^2} \frac{(2n)!}{2^{2n}(n!)^2} \frac{\pi}{2} \\ &= \frac{(2n+2)!}{(2(n+1))^2 2^{2n}(n!)^2} \frac{\pi}{2} \\ &= \frac{(2n+2)!}{2^{2n+2}((n+1)!)^2} \frac{\pi}{2}. \end{aligned}$$

Ainsi la propriété est vérifiée au rang $n+1$, et donc par le principe de récurrence,

$$\forall n \geq 0, W_{2n} = \frac{(2n)!}{2^{2n}(n!)^2} \frac{\pi}{2}.$$

6. Puisque $\frac{1}{n} \xrightarrow[n \rightarrow +\infty]{} 0$, on a

$$\ln\left(1 - \frac{1}{n}\right) = -\frac{1}{n} - \frac{1}{2n^2} - \frac{1}{3n^3} + o\left(\frac{1}{n^3}\right).$$

Et donc

$$a_n = -1 - \left(n - \frac{1}{2}\right) \left(-\frac{1}{n} - \frac{1}{2n^2} - \frac{1}{3n^3} + o\left(\frac{1}{n^3}\right)\right) = -1 + 1 + \frac{1}{2n} - \frac{1}{2n} - \frac{1}{4n^2} + \frac{1}{3n^2} + o\left(\frac{1}{n^2}\right).$$

$$\text{Ainsi, } a_n = \frac{1}{12n^2} + o\left(\frac{1}{n^2}\right) \underset{n \rightarrow +\infty}{\sim} \frac{1}{12n^2}.$$

Et alors, par critère de comparaison pour les séries à termes positifs⁴, la série de terme général a_n converge car la série de terme général $\frac{1}{12n^2}$ est une série de Riemann convergente.

Rappel

On a le droit d'élever des équivalents à une puissance fixée. En particulier à la puissance $1/2$, ce qui revient à prendre la racine des équivalents.

⚠ Attention !

Prouver la propriété au rang $n+1$, c'est vérifier l'égalité pour $W_{2(n+1)} = W_{2n+2}$, et non pour W_{2n+1} , qui n'est pas un terme de la suite $(W_{2n})_{n \in \mathbf{N}}$.

DL

Au voisinage de 0,

$$\ln(1-x) = -x - \frac{x^2}{2} - \frac{x^3}{3} + o(x^3)$$

Négligeabilité

Il manque dans ce développement des termes en $o\left(\frac{1}{n^3}\right)$ et en $\frac{1}{n^3}$. Tous ces termes sont négligeables devant $\frac{1}{n^2}$, et ont donc été inclus dans le $o\left(\frac{1}{n^2}\right)$.

⁴ Il n'est pas évident que a_n soit positif, mais son équivalent l'est, donc le critère de comparaison s'applique.

7. Pour $n \geq 2$, on a

$$\begin{aligned} \ln(A_n) - \ln(A_{n-1}) &= n \ln n - n + \frac{1}{2} \ln(2\pi) + \frac{1}{2} \ln(n) - \ln(n!) - (n-1) \ln(n-1) + (n-1) - \frac{1}{2} \ln(2\pi) - \frac{1}{2} \ln(n-1) + \ln((n-1)!) \\ &= -1 + n \ln(n) - (n-1) \ln(n-1) - \ln\left(\frac{n!}{(n-1)!}\right) + \frac{1}{2} \ln\left(\frac{n}{n-1}\right) \\ &= -1 + (n-1) \ln(n) - (n-1) \ln(n-1) + \frac{1}{2} \ln\left(\frac{n}{n-1}\right) \\ &= -1 + \left(n - \frac{1}{2}\right) \ln\left(\frac{n}{n-1}\right) \\ &= -1 - \left(n - \frac{1}{2}\right) \ln\left(\frac{n-1}{n}\right) \\ &= -1 - \left(n - \frac{1}{2}\right) \ln\left(1 - \frac{1}{n}\right) = \boxed{a_n} \end{aligned}$$

8. Pour $N \geq 2$, on a

$$\sum_{n=2}^N a_n = \sum_{n=2}^N (\ln(A_n) - \ln(A_{n-1})) = \ln(A_N) - \ln(A_1).$$

Somme télescopique.

Et donc en passant à l'exponentielle, $A_n = A_1 e^{\sum_{k=2}^n a_k}$. Or la série de terme général a_n converge, donc

$$\sum_{n=2}^{+\infty} a_n \xrightarrow{N \rightarrow +\infty} \sum_{n=2}^{+\infty} a_n = M.$$

Ainsi, par continuité de l'exponentielle, $A_N \xrightarrow{N \rightarrow +\infty} A_1 e^M = \ell$. Cette limite est alors strictement positive car $A_1 = e^{-1} > 0$.

9.a. On a $A_n \underset{n \rightarrow +\infty}{\sim} \ell$ et ainsi

$$n! \underset{n \rightarrow +\infty}{\sim} n^n e^{-n} \sqrt{n} \frac{1}{A_n} \underset{n \rightarrow +\infty}{\sim} \boxed{\frac{1}{\ell} n^n e^{-n} \sqrt{n}}.$$

9.b. Reprenons l'expression de la question 5 : on a

$$(2n)! \underset{n \rightarrow +\infty}{\sim} \frac{1}{\ell} (2n)^{2n} e^{-2n} \sqrt{2n} \text{ et } (n!)^2 \underset{n \rightarrow +\infty}{\sim} \frac{1}{\ell^2} n^{2n} e^{-2n} n$$

de sorte que $W_{2n} \underset{n \rightarrow +\infty}{\sim} \ell \frac{(2n)^{2n} e^{-2n} \sqrt{2n} \pi}{2^{2n} n^{2n} e^{-2n} n} \frac{\pi}{2} \underset{n \rightarrow +\infty}{\sim} \ell \frac{\pi}{\sqrt{2n}}$.

Comme d'autre part, on sait par la question 4.b que $W_{2n} \underset{n \rightarrow +\infty}{\sim} \sqrt{\frac{\pi}{4n}}$, alors

$$\ell \frac{\pi}{\sqrt{2n}} \underset{n \rightarrow +\infty}{\sim} \sqrt{\frac{\pi}{4n}} \Leftrightarrow \ell \sim \sqrt{\frac{1}{2\pi}}.$$

Mais deux constantes sont équivalentes si et seulement si elles sont égales, donc $\ell = \sqrt{\frac{1}{2\pi}}$

et donc

$$\boxed{n! \underset{n \rightarrow +\infty}{\sim} n^n e^{-n} \sqrt{2\pi n}}.$$

Partie II : Étude de variables aléatoires

10. La fonction f_a est évidemment positive sur \mathbf{R} , et elle est continue sauf peut-être en 0. De plus, pour $A > 0$, on a

$$\int_{-\infty}^A f_a(t) dt = \int_0^A f_a(t) dt = \int_0^A \frac{t}{a^2} e^{-\frac{t^2}{2a^2}} dt = \left[-e^{-\frac{t^2}{2a^2}} \right]_0^A = 1 - e^{-\frac{A^2}{2a^2}} \xrightarrow{A \rightarrow +\infty} 1.$$

Donc $\int_{-\infty}^{+\infty} f_a(t) dt$ converge et vaut 1.

Ainsi, f_a est bien une densité de probabilité.

Remarque
 Une suite convergente est équivalente à sa limite si celle-ci est non nulle (seule la suite nulle est équivalente à 0), et donc c'est pour cette raison qu'il était indispensable de vérifier que $\ell \neq 0$.

Pour la culture
 Cette formule est appelée formule de Stirling, et c'est en quelques sorte «le meilleur» équivalent que l'on puisse donner de $n!$ à l'aide de suites plus classiques. Il en existe une généralisation pour la fonction Γ

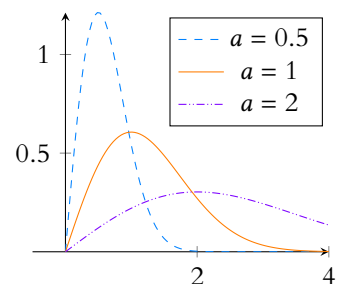
$$\Gamma(x) \underset{x \rightarrow +\infty}{\sim} x^x e^{-x} \sqrt{\frac{2\pi}{x}}$$


FIGURE 24.3— La densité f_a . La loi associée est appelée loi de Rayleigh.

11. Pour $x \leq 0$, on a

$$F_X(x) = \int_{-\infty}^x f_a(t) dt = \int_{-\infty}^x 0 dt = 0.$$

Et pour $x > 0$, le calcul a été fait à la question précédente :

$$F_X(x) = \int_{-\infty}^x f(t) dt = \int_0^x f(t) dt = 1 - e^{-\frac{x^2}{2a^2}}.$$

Ainsi,
$$F_X(x) = \begin{cases} 0 & \text{si } x \leq 0 \\ 1 - e^{-\frac{x^2}{2a^2}} & \text{si } x \geq 0 \end{cases}$$

12. Par définition, X admet une espérance si et seulement si $\int_0^{+\infty} t f_a(t) dt = \int_0^{+\infty} \frac{t^2}{a^2} e^{-\frac{t^2}{2a^2}} dt$ converge.

Notons que cette intégrale ressemble fortement à celle définissant le moment d'ordre 2 d'une loi normale.

Plus précisément, soit Y une variable aléatoire suivant la loi normale $\mathcal{N}(0, a^2)$, qui possède pour densité $t \mapsto \frac{1}{a\sqrt{2\pi}} e^{-\frac{t^2}{2a^2}}$.

Alors nous savons, par parité de l'intégrande, que

$$E(Y^2) = \int_{-\infty}^{+\infty} t^2 \frac{1}{a\sqrt{2\pi}} e^{-\frac{t^2}{2a^2}} dt = \frac{2}{a\sqrt{2\pi}} \int_0^{+\infty} t^2 e^{-\frac{t^2}{2a^2}} dt.$$

De plus, $E(Y^2) = V(Y) + E(Y)^2 = a^2 + 0 = a^2$.

On en déduit que

$$\int_0^{+\infty} t^2 e^{-\frac{t^2}{2a^2}} dt = a^3 \frac{\sqrt{2\pi}}{2}.$$

Et donc

$$E(X) = \frac{1}{a^2} \int_0^{+\infty} t^2 e^{-\frac{t^2}{2a^2}} dt = \frac{1}{a^2} \frac{a^3 \sqrt{2\pi}}{2} = a \sqrt{\frac{\pi}{2}}.$$

13. De même, sous réserve de convergence, on a

$$E(X^2) = \int_0^{+\infty} t^2 f_a(t) dt = \int_0^{+\infty} \frac{t^3}{a^2} e^{-\frac{t^2}{2a^2}} dt.$$

Soit $A > 0$. Procédons à une intégration par parties sur le segment $[0, A]$, en posant $u(t) = t^2$ et $v(t) = -e^{-\frac{t^2}{2a^2}}$. u et v sont alors bien de classe \mathcal{C}^1 sur $[0, A]$, avec $u'(t) = 2t$ et $v'(t) = \frac{t^2}{a^2} e^{-\frac{t^2}{2a^2}}$. Il vient alors

$$\begin{aligned} \int_0^A t^3 e^{-\frac{t^2}{2a^2}} dt &= \left[-t^2 e^{-\frac{t^2}{2a^2}} \right]_0^A + 2 \int_0^A t e^{-\frac{t^2}{2a^2}} dt \\ &= -A^2 e^{-\frac{A^2}{2a^2}} + 2 \left[-a^2 e^{-\frac{t^2}{2a^2}} \right]_0^A \\ &= -A^2 e^{-\frac{A^2}{2a^2}} + 2a^2 - 2a^2 e^{-\frac{A^2}{2a^2}} \\ &\xrightarrow{A \rightarrow +\infty} 2a^2 \end{aligned}$$

Et alors, par la formule de Huygens, il vient

$$V(X) = E(X^2) - E(X)^2 = 2a^2 - a^2 \frac{\pi}{2} = a^2 \frac{4 - \pi}{2}.$$

- 14.a. Notons que Z est à valeurs dans \mathbf{R}^+ . Donc pour $x < 0$, on a $F_Z(x) = 0$. Soit donc $x \geq 0$. Il vient

$$F_Z(x) = P(Z \leq x) = P\left(a\sqrt{-2 \ln V} \leq x\right) = P\left(\sqrt{-2 \ln V} \leq \frac{x}{a}\right) = P\left(-2 \ln V \leq \frac{x^2}{a^2}\right).$$

Convergence

Comme nous savons que $E(Y^2)$ existe (car $V(Y)$ existe), la convergence de l'intégrale est automatique, pas besoin de la prouver.

Astuce

Ce résultat pouvait également se retrouver par une intégration par parties. On avait alors besoin de la valeur de l'intégrale de Gauss $\int_0^{+\infty} e^{-t^2/2a^2}$ qui se retrouve en se rappelant que l'intégrale de la densité d'une loi normale vaut 1.

Inégalités

On a préservé le sens des inégalités par croissance de la fonction carré sur \mathbf{R}_+ .

Et alors, par croissance de l'exponentielle, il vient

$$F_Z(x) = P\left(\ln V \geq -\frac{x^2}{2a^2}\right) = P\left(V \geq e^{-\frac{x^2}{2a^2}}\right) = 1 - P\left(V \leq e^{-\frac{x^2}{2a^2}}\right).$$

Puisque V suit une loi uniforme sur $[0, 1]$, et que $0 \leq e^{-\frac{x^2}{2a^2}} \leq 1$, il vient donc

$$F_Z(x) = 1 - F_V\left(e^{-\frac{x^2}{2a^2}}\right) = 1 - e^{-\frac{x^2}{2a^2}}.$$

Ainsi, Z a la même fonction de répartition que X , et donc ces deux variables ont même loi.

Remarque

Il s'agit ici d'un cas particulier de la méthode d'inversion rencontrée en TP.

14.b. Le programme suivant convient :

```
1 function y = simul(a)
2   y = a*sqrt(-2*log(rand()))
3 endfunction
```

15. Rappelons que `floor((rand()*n))+1` simule une loi uniforme sur $\llbracket 1, n \rrbracket$. Et donc le programme choisit successivement les numéros des urnes dans lesquelles sont placées les boules, et doit donc s'arrêter dès qu'un numéro d'urne est tiré pour la seconde fois. Le contenu de chacune des urnes est représenté par le vecteur `urnes` qui au départ ne contient que des 0 (car toutes les urnes sont vides) et chaque fois qu'une boule est placée dans une urne, le 0 est remplacé par un 1.

```
1 function T = tirage(n)
2   urnes = zeros(1,n)
3   T = 1
4   choix = floor((rand()*n))+1
5   while urnes(choix) < 1
6       urnes(choix) = urnes(choix)+1
7       choix = floor((rand()*n))+1
8       T = T+1
9   end
10 endfunction
```

Avant la première entrée dans la boucle `while`, la variable `choix` contient le numéro de l'urne dans laquelle sera placée la première boule, et avant la $k^{\text{ème}}$ entrée dans la boucle, `choix` contient le numéro de l'urne dans laquelle doit être placée la $(k+1)^{\text{ème}}$ boule. Si cette urne contient déjà une boule (c'est-à-dire si `urnes(choix) = 1`), alors le programme s'arrête.

Si la boucle `while` a alors été parcourue k fois, c'est-à-dire si on a placé k boules en plus de la première dans des urnes vides, T vaut $k+1$, et donc le programme renvoie $k+1$ qui est bien la valeur attendue de T_n .

- 16.a. Puisqu'il n'y a que n urnes, si les n premières boules sont placées dans des urnes vides⁵, alors la $(n+1)^{\text{ème}}$ boule placée ira nécessairement dans une urne déjà occupée. Et donc T_n ne prend que des valeurs inférieures ou égales à $n+1$ de sorte que $P(T_n > n+1) = 0$.
- 16.b. Remarquons que pour $k=1$, on a $P(T_n > 1) = 1$, car T_n est à valeurs dans $\llbracket 2, n \rrbracket$. Notons A_i l'événement «la $i^{\text{ème}}$ boule est placée dans une urne vide». Alors

$$[T_n > k] = A_1 \cap A_2 \cap \dots \cap A_k.$$

Par la formule des probabilités composées, on a alors

$$P(T_n > k) = P(A_1)P_{A_1}(A_2)P_{A_1 \cap A_2}(A_3) \cdots P_{A_1 \cap \dots \cap A_{k-1}}(A_k).$$

Mais pour tout $i \in \llbracket 2, k \rrbracket$, $P_{A_1 \cap \dots \cap A_{i-1}}(A_i) = \frac{n-(i-1)}{n}$.

Et donc

$$P(T_n > k) = 1 \times \frac{n-1}{n} \times \dots \times \frac{n-(k-1)}{n} = \frac{(n-1) \cdots (n-k+1)}{n^{k-1}} = \frac{n!}{n^k(n-k)!}.$$

⁵ C'est-à-dire si $T_n \geq k$.

Explication

Cette probabilité conditionnelle est la probabilité de choisir une urne encore vide, sachant que $i-1$ urnes contiennent déjà une boule suite aux précédents tirages.

- 16.c. Puisque T_n est à valeurs entières, on a, pour tout $k \in \llbracket 2, n+1 \rrbracket$, $[T_n > k-1] = [T_n = k] \cup [T_n > k]$. Ces deux événements étant incompatibles, il vient donc

$$P(T_n > k-1) = P(T_n = k) + P(T_n > k) \Leftrightarrow P(T_n = k) = P(T_n > k-1) - P(T_n > k).$$

Soit encore

$$P(T_n = k) = \frac{n!}{n^{k-1}(n-k+1)!} - \frac{n!}{n^k(n-k)!} = n! \frac{n - (n-k+1)}{n^k(n-k+1)!} = \frac{n!(n-k)}{n^k(n-k+1)!}.$$

- 16.d. Par définition, $E(T_n) = \sum_{k=2}^{n+1} k \frac{n!(k-1)}{n^k(n-k+1)!}$.

Notons que $[1 : n]$ a pour effet de créer un vecteur contenant les nombres de 1 à n , et donc $\text{prod}([1 : n])$ renvoie $1 \times 2 \times \dots \times n = n!$.

Les valeurs successives de n^k sont calculées à la ligne 7 en remarquant que $n^k = n^{k-1} \times n$.

De même, les valeurs successives de $(n-k+1)!$ sont calculées à la ligne 8 en remarquant que

$$(n-k+1)! = \frac{(n-(k-1)+1)!}{(n-k+2)}.$$

Le programme calcule donc les valeurs successives de $k(k-1) \frac{1}{n^k(n-k+1)!}$, et les ajoute à

la variable somme, qui à la fin de la boucle vaut $\sum_{k=2}^{n+1} k(k-1) \frac{1}{n^k(n-k+1)!}$.

Puis la ligne 11 multiplie cette somme par facto qui vaut $n!$, permettant alors d'obtenir la valeur de $E(T_n)$.

```

1 fonction E = esperance(n)
2     facto = prod([1 : n])
3     fac = facto
4     somme = 0
5     puissance = n
6     for k=2 : (n+1)
7         puissance = puissance*n
8         fac = fac/(n-k+2)
9         somme = somme + k*(k-1)/(puissance*fac)
10    end
11    E = facto*somme
12 endfunction

```

17. Par définition de Y_n , on a

$$P(Y_n > y) = P\left(\frac{T_n}{\sqrt{n}} > y\right) = P(T_n > y\sqrt{n}).$$

Mais T_n étant à valeurs entières, on a

$$P(T_n > k_n) = \underbrace{P(k_n < T_n \leq y\sqrt{n})}_{=0} + P(T_n > y\sqrt{n}) \Leftrightarrow P(T_n > k_n) = P(T_n > y\sqrt{n}).$$

Et donc $P(Y_n > y) = P(T_n > k_n)$.

18. On a prouvé précédemment que $P(T_n > k_n) = \frac{n!}{n^{k_n}(n-k_n)!}$.

Notons que $y\sqrt{n} - 1 < k_n \leq y\sqrt{n}$ et donc $\frac{y\sqrt{n} - 1}{y\sqrt{n}} \leq \frac{k_n}{y\sqrt{n}} \leq 1$.

Par le théorème des gendarmes, on a alors $\frac{k_n}{y\sqrt{n}} \xrightarrow{n \rightarrow +\infty} 1$ soit $k_n \sim y\sqrt{n}$.

En utilisant le résultat de la question 9.b, il vient $n! \underset{n \rightarrow +\infty}{\sim} e^{-n} n^n \sqrt{2\pi n}$.

De même, on a $n - k_n \geq n - y\sqrt{n} \rightarrow +\infty$, et donc

$$(n - k_n)! \underset{n \rightarrow +\infty}{\sim} e^{-(n-k_n)} (n - k_n)^{n-k_n} \sqrt{2\pi(n - k_n)}.$$

Alternative

Si l'on ne remarque pas ceci, on peut remplacer la ligne 7 du programme par $\text{puissance} = n^k$

Alternative

Là aussi, on pouvait s'en tirer différemment, par exemple à l'aide de $\text{facto} = \text{prod}([1 : n-k+1])$

Détails

Il n'y a pas d'entier dans l'intervalle $]k_n, y\sqrt{n}]$, donc T_n ne peut y prendre de valeur.

Plus généralement

Si $u_n \rightarrow +\infty$, alors le même raisonnement prouve que

$$[u_n] \sim u_n.$$

On en déduit, que

$$P(Y_n > y) = P(T_n > k_n) \underset{n \rightarrow +\infty}{\sim} \frac{1}{n^{k_n}} \frac{e^{-n} n^n \sqrt{2\pi n}}{e^{-(n-k_n)} (n-k_n)^{n-k_n} \sqrt{2\pi(n-k_n)}} \\ \underset{n \rightarrow +\infty}{\sim} e^{-k_n} \left(\frac{n}{n-k_n}\right)^{n-k_n} \sqrt{\frac{n}{n-k_n}}.$$

Mais on a $\frac{n}{n-k_n} = \frac{n}{n} \frac{1}{1 - \frac{k_n}{n}} \underset{n \rightarrow +\infty}{\rightarrow} 1$ et donc $\sqrt{\frac{n}{n-k_n}} \underset{n \rightarrow +\infty}{\rightarrow} 1$. Il vient alors

$$P(Y_n > y) \underset{n \rightarrow +\infty}{\sim} \boxed{e^{-k_n} \left(\frac{1}{1 - \frac{k_n}{n}}\right)^{n-k_n} \underset{n \rightarrow +\infty}{\sim} e^{-k_n} \left(1 - \frac{k_n}{n}\right)^{k_n-n}}.$$

19.a. Nous savons qu'au voisinage de 0, on a $\ln(1-t) = -t - \frac{t^2}{2} + o(t^2)$ et donc

$$-t + (1-t)\ln(1-t) = -t + (t-1)\left(-t - \frac{t^2}{2} + o(t^2)\right) \\ = -t + t + \frac{t^2}{2} - t^2 + \underbrace{o(t^2) - \frac{t^3}{2}}_{=o(t^2)} + o(t^2) = \boxed{-\frac{t^2}{2} + o(t^2)}.$$

Remarque

Ce résultat peut bien entendu s'obtenir également à l'aide de Taylor-Young. Mais il faut alors calculer une dérivée seconde peu alléchante...

19.b. En factorisant par n , il vient

$$-k_n + (k_n - n) \ln\left(\frac{k_n}{n} - 1\right) = n \left(-\frac{k_n}{n} + \left(\frac{k_n}{n} - 1\right) \ln\left(1 - \frac{k_n}{n}\right)\right) \\ = n \left(-\frac{k_n^2}{2n^2} + o\left(\frac{k_n^2}{n^2}\right)\right) \\ \underset{n \rightarrow +\infty}{\sim} -\frac{k_n^2}{2n} \underset{n \rightarrow +\infty}{\sim} -\frac{y^2}{2}$$

C'est le résultat de la question précédente, qui s'applique car $\frac{k_n}{n} \rightarrow 0$.

Et donc

$$\boxed{\lim_{n \rightarrow +\infty} -k_n + (k_n - n) \ln\left(1 - \frac{k_n}{n}\right) = -\frac{y^2}{2}}.$$

20. Nous savons que pour $x \leq 0$, on a $P(Y_n \leq x) = 0$ car Y_n est à valeurs strictement positives. Pour $x > 0$, on a, grâce au résultat de la question précédente,

$$P(Y_n \leq y) = 1 - P(Y_n > y) = 1 - e^{-k_n + (k_n - n) \ln\left(1 - \frac{k_n}{n}\right)} \underset{n \rightarrow +\infty}{\rightarrow} 1 - e^{-y^2/2}.$$

On reconnaît alors la fonction de répartition d'une variable aléatoire de densité f_1 . Soit X une telle variable aléatoire.

On a alors, en tout $x \in \mathbf{R}$ (et donc en particulier en tout point de continuité de F_X), $\lim_{n \rightarrow +\infty} F_{Y_n}(x) = F_X(x)$.

Et donc $\boxed{Y_n \xrightarrow{\mathcal{L}} X}$.

Limite

On a ici utilisé la continuité de l'exponentielle pour passer à la limite dans l'exponentielle.

Remarque

On a ici un exemple d'une suite de variables discrètes, qui convergent en loi vers une variable à densité.

Quelques commentaires : la situation rencontrée ici est en fait celle que l'on rencontre dans le classique paradoxe des anniversaires : combien d'élèves doit contenir une classe pour que la probabilité que deux d'entre eux soient nés le même jour soit supérieure ou égale à 1/2 ?

Cela revient à considérer 365 urnes (correspondant aux jours de l'année) et à placer au hasard les élèves dans chacune de ces urnes. Si l'on considère que 365 est grand, on peut donc approcher la loi de $\frac{T_{365}}{\sqrt{365}}$ par celle de X .

Et donc on a

$$P(T_{365} \leq n) = P\left(\frac{T_{365}}{\sqrt{365}} \leq \frac{n}{\sqrt{365}}\right) \approx P\left(X \leq \frac{n}{\sqrt{365}}\right).$$

Or, $P\left(X \leq \frac{n}{\sqrt{365}}\right) \geq \frac{1}{2} \Leftrightarrow 1 - e^{-\frac{x^2}{2 \times 365}} \geq \frac{1}{2} \Leftrightarrow n \geq 23$.

Notons qu'un calcul exact utilisant la loi de T_{365} conduit au même résultat, et donc l'approximation faite plus haut semble valable.

CONCOURS BLANC 2 : MATHS II (08.03.19)

- On rappelle que la série $\sum_{n \geq 1} \frac{1}{n^x}$ converge si et seulement si le réel x est strictement supérieur à 1.
- On note ζ la fonction définie sur $]1, +\infty[$ par : $\forall x > 1, \zeta(x) = \sum_{n=1}^{+\infty} \frac{1}{n^x}$; on admet que $\zeta(2) = \frac{\pi^2}{6}$.
- Toutes les variables aléatoires introduites dans le problème sont supposées définies sur un même espace probabilisé (Ω, \mathcal{A}, P) .
- Si R est un élément de la tribu \mathcal{A} , on note \bar{R} l'événement contraire de R .

L'objet du problème est l'étude de la convergence de séries dont les termes sont des variables aléatoires.

La convergence de telles séries, en loi ou en probabilité, est celle de la suite des sommes partielles associées.

Autrement dit, pour toute suite de variables aléatoires $(U_n)_{n \in \mathbf{N}^*}$, on dit que la série $\sum_{n \geq 1} U_n$ converge (en loi ou en probabilité) lorsque

la suite de variables aléatoires $\left(\sum_{k=1}^n U_k \right)_{n \in \mathbf{N}^*}$ converge (en loi ou en probabilité).

Partie I. Séries télescopiques.

Dans cette partie, on considère une suite $(X_n)_{n \in \mathbf{N}^*}$ de variables aléatoires indépendantes, de même loi qu'une variable aléatoire X de référence, et on étudie la convergence de la série aléatoire $\sum_{n \geq 1} \left(\frac{X_n}{n} - \frac{X_{n+1}}{n+1} \right)$.

1. **a.** Justifier la convergence de la série $\sum_{n \geq 1} \frac{1}{n(n+1)}$. Calculer la somme de cette série.
- b.** Dans cet exemple, quelle est la loi de la variable aléatoire de référence X ?
2. Pour tout $n \in \mathbf{N}^*$, soit Y_n une variable aléatoire admettant pour densité la fonction f_n définie par :

$$f_n(x) = \begin{cases} 0 & \text{si } x < -\frac{1}{n+1} \text{ ou si } x > 1 \\ 1 + (n+1)x & \text{si } -\frac{1}{n+1} \leq x \leq 0 \\ c_n & \text{si } 0 < x < \frac{n}{n+1} \\ (n+1)(1-x) & \text{si } \frac{n}{n+1} \leq x \leq 1 \end{cases}$$

où c_n est une constante strictement positive.

- a.** Calculer la valeur de c_n et représenter graphiquement f_3 dans le plan rapporté à un repère orthonormé.
- b.** Déterminer la fonction de répartition F_n de Y_n . La fonction F_n est-elle de classe \mathcal{C}^1 sur \mathbf{R} ?
- c.** Montrer que la suite $(Y_n)_{n \in \mathbf{N}^*}$ converge en loi vers une variable aléatoire Y dont on précisera la loi.
3. Dans cette question, on suppose que la variable aléatoire de référence X possède une densité f bornée.

Pour tout $n \in \mathbf{N}^*$, on pose : $D_n = X_1 - \frac{X_{n+1}}{n+1}$.

- a.** Montrer que la suite de variables aléatoires $\left(\frac{X_{n+1}}{n+1} \right)_{n \in \mathbf{N}^*}$ converge en probabilité vers 0.
- b.** En déduire que la suite de variables aléatoires $(D_n)_{n \in \mathbf{N}^*}$ converge en loi vers X .
- c.** Justifier que la variable aléatoire D_n admet pour densité la fonction f_{D_n} telle que :

$$\forall x \in \mathbf{R}, f_{D_n}(x) = (n+1) \int_{-\infty}^{+\infty} f(t) f((n+1)(t-x)) dt.$$

- d.** En déduire une nouvelle démonstration du résultat obtenu dans la question 2.c.
 4. Dans cette question, on suppose que la variable aléatoire de référence X suit une loi normale centrée réduite.
- Pour tout $n \in \mathbf{N}^*$, on pose : $U_n = \frac{X_n}{n} - \frac{X_{n+1}}{n+1}$ et $T_n = \sum_{k=1}^n U_k$.

- a.** Pour tout $n \in \mathbf{N}^*$, déterminer la loi de la variable aléatoire U_n .

- b. Justifier la convergence en loi de la série $\sum_{n \geq 1} U_n$.
- c. Soit $(U'_n)_{n \in \mathbf{N}^*}$ une suite de variables aléatoires indépendantes vérifiant : $\forall n \in \mathbf{N}^*$, U_n et U'_n ont même loi.
 Pour tout $n \in \mathbf{N}^*$, on pose : $T'_n = \sum_{k=1}^n U'_k$.
- i. Justifier que la suite de variables aléatoires $(T'_n)_{n \in \mathbf{N}^*}$ converge en loi vers une variable aléatoire qui suit la loi normale centrée de variance $\frac{\pi^2}{3} - 1$.
- ii. Pourquoi ce résultat ne contredit-il pas ceux obtenus dans les questions 3.b et 4.b.

Partie II. Séries harmoniques «lacunaires».

Dans cette partie, on étudie les séries numériques obtenues à partir de la série harmonique divergente $\sum_{n \geq 1} \frac{1}{n}$ par effacement de certaines de ses termes.

Pour tout partie \mathcal{G} de \mathbf{N}^* , on note $\mathbb{1}_{\mathcal{G}}$ la fonction indicatrice de \mathcal{G} , c'est-à-dire la fonction définie sur \mathbf{N}^* à valeurs dans $\{0, 1\}$ telle que : $\forall k \in \mathbf{N}^*$, $\mathbb{1}_{\mathcal{G}}(k) = \begin{cases} 1 & \text{si } k \in \mathcal{G} \\ 0 & \text{sinon} \end{cases}$

Pour tout entier n supérieur ou égal à 1, on pose : $\forall \mathcal{G} \subset \mathbf{N}^*$, $h_n(\mathcal{G}) = \sum_{k=1}^n \frac{\mathbb{1}_{\mathcal{G}}(k)}{k}$.

Dans la question 5, on étudie deux cas de convergence et la question 6 est consacrée à un cas de divergence.

5. On pose : $\mathcal{D} = \{n^2, n \in \mathbf{N}^*\}$ et $\mathcal{T} = \{n^3, n \in \mathbf{N}^*\}$.

- a. Exprimer $h_n(\mathcal{D})$ à l'aide d'une somme partielle de la série de Riemann $\sum_{n \geq 1} \frac{1}{n^2}$.
- b. En déduire la convergence de la suite $(h_n(\mathcal{D}))_{n \in \mathbf{N}^*}$ et calculer $\sum_{n=1}^{+\infty} \frac{\mathbb{1}_{\mathcal{D}}(n)}{n}$.
- c. Justifier que $\mathcal{D} \cap \mathcal{T}$ est l'ensemble des entiers m pour lesquels $m^{1/6} \in \mathbf{N}^*$.
 Pour traiter cette question, on admet que la racine carrée d'un entier naturel qui n'appartient pas à \mathcal{D} est un nombre irrationnel, c'est-à-dire, un nombre qui ne peut pas s'écrire comme le quotient de deux entiers.
- d. Montrer que la suite $(h_n(\mathcal{D} \cup \mathcal{T}))_{n \in \mathbf{N}^*}$ est convergente et exprimer la somme $\sum_{n=1}^{+\infty} \frac{\mathbb{1}_{\mathcal{D} \cup \mathcal{T}}(n)}{n}$ à l'aide de certaines valeurs de la fonction ζ .

6. On note \mathcal{I} l'ensemble des entiers naturel impairs : $\mathcal{I} = \{2n - 1, n \in \mathbf{N}^*\}$.

Pour tout $n \in \mathbf{N}^*$, on pose : $u_n = \int_n^{n+1} \left(\frac{1}{2n-1} - \frac{1}{2t-1} \right) dt$.

- a. Pour tout $n \in \mathbf{N}^*$, établir l'encadrement : $0 \leq u_n \leq \frac{2}{(2n-1)^2}$.
- b. Montrer que pour tout $n \in \mathbf{N}^*$, on a : $h_n(\mathcal{I}) = \sum_{k=1}^{\lfloor \frac{n+1}{2} \rfloor} u_k + \int_1^{\lfloor \frac{n+3}{2} \rfloor} \frac{1}{2t-1} dt$.
- c. Pour tout $n \in \mathbf{N}^*$, justifier l'encadrement : $0 \leq \ln \left(\frac{1}{n} \left(2 \left\lfloor \frac{n+3}{2} \right\rfloor - 1 \right) \right) \leq \frac{2}{n}$.

- d. Montrer que la série de terme général u_n est convergente.

On pose $\delta = \sum_{n=1}^{+\infty} u_n$. Établir l'égalité : $\lim_{n \rightarrow +\infty} (h_n(\mathcal{I}) - \ln(\sqrt{n})) = \delta$.

- e. Pour tout $n \in \mathbf{N}^*$, montrer que l'on a : $\sum_{k=n+1}^{+\infty} u_k \leq \frac{1}{2n-1}$.

- f. Justifier pour tout entier $n \geq 3$, l'encadrement : $-\frac{1}{n} \leq \delta - (h_n(\mathcal{I}) - \ln(\sqrt{n})) \leq \frac{1}{n-2}$.

- g. La fonction SciLab suivante, dont le script est incomplet (ligne 6), permet de donner une valeur approchée de δ en calculant successivement des valeurs de $h_n(\mathcal{I}) - \ln(\sqrt{n})$ jusqu'à atteindre une précision donnée.

```

1 function s = delta(eps)
2     n=3 ;
3     s=1+1/3-(log(3)/2) ;
4     while 1/(n-2)>eps
5         n=n+2
6         s = s+1/n+..... ;
7     end
8 endfunction

```

- i. Quelles sont les valeurs de $h_n(\mathcal{I}) - \ln(\sqrt{n})$ affectées successivement à la variable s lorsqu'on applique cette fonction à $\text{eps}=0.2$?
- ii. Compléter la ligne 6.
- iii. Pour quelles raisons l'algorithme utilisé peut-il assurer une précision arbitraire au calcul de la valeur approchée de δ ?

Partie III. Séries de Riemann alternées.

Dans cette partie, on note $(X_n)_{n \in \mathbf{N}^*}$ une suite de variables aléatoires indépendantes, de même loi uniforme sur la paire $\{-1, +1\}$ c'est-à-dire : pour tout $n \in \mathbf{N}^*$, $P(X_n = 1) = P(X_n = -1) = \frac{1}{2}$.

Pour tout $n \in \mathbf{N}^*$, on pose : $S_n = \sum_{k=1}^n X_k$.

7. Soit $(x_n)_{n \in \mathbf{N}^*}$ une suite réelle. Pour tout $n \in \mathbf{N}^*$, on pose : $s_n = \sum_{k=1}^n x_k$.

On suppose l'existence d'un réel $\alpha \geq 0$ et d'un réel $M > 0$ tels que : $\forall n \in \mathbf{N}^*, |s_n| \leq Mn^\alpha$.

- a. Soit β un réel tel que $\beta > \alpha$.

i. Montrer pour tout entier $n \geq 2$, l'égalité :
$$\sum_{k=1}^n \frac{x_k}{k^\beta} = \frac{s_n}{n^\beta} + \sum_{k=1}^{n-1} s_k \left(\frac{1}{k^\beta} - \frac{1}{(k+1)^\beta} \right).$$

- ii. En déduire que la série $\sum_{n \geq 1} \frac{x_n}{n^\beta}$ est convergente.

- b. Justifier pour tout réel $x > 0$, la convergence de la série $\sum_{n \geq 1} \frac{(-1)^n}{n^x}$.

8. Soit s et t des réels strictement positifs et n un entier supérieur ou égal à 1.

- a. Calculer l'espérance de la variable aléatoire e^{tS_n} .

- b. En utilisant l'écriture de e^u ($u \in \mathbf{R}$) sous forme de somme d'une série, établir l'inégalité :

$$\frac{1}{2} (e^t + e^{-t}) \leq e^{\frac{t^2}{2}}.$$

- c. À l'aide de l'inégalité de Markov, montrer que : $P(S_n > s) \leq \exp\left(\frac{nt^2}{2} - ts\right)$.

- d. Justifier l'inégalité : $P(|S_n| > s) \leq 2 \exp\left(-\frac{s^2}{2n}\right)$.

9. Pour tout réel $\alpha \geq 0$, on pose : $\mathcal{C}_\alpha = \bigcap_{n=1}^{+\infty} \left(\bigcup_{k=n}^{+\infty} [|S_k| > k^\alpha] \right)$.

- a. Justifier que \mathcal{C}_α est un élément de la tribu \mathcal{A} .

- b. Montrer que si $\alpha > \frac{1}{2}$, alors la série $\sum_{n \geq 1} P(|S_n| > n^\alpha)$ est convergente.

- c. En déduire que pour tout réel $\alpha > \frac{1}{2}$, on a $P(\mathcal{C}_\alpha) = 0$.

10. Dans cette question, on s'intéresse à la série aléatoire $\sum_{n \geq 1} \frac{X_n}{n}$.

On note \mathcal{C} l'ensemble des $\omega \in \Omega$ pour lesquels la série numérique $\sum_{n \geq 1} \frac{X_n(\omega)}{n}$ converge et pour tout $n \in \mathbf{N}^*$, on

pose : $K_n = \sum_{k=1}^n \frac{X_k}{k}$.

Soit K l'application définie sur Ω par : $K(\omega) = \begin{cases} \lim_{n \rightarrow +\infty} K_n(\omega) & \text{si } \omega \in \mathcal{C} \\ 0 & \text{sinon} \end{cases}$

On admet sans démonstration que \mathcal{C} est un élément de la tribu \mathcal{A} et que K est une variable aléatoire.

a. En utilisant le résultat de la question 7.a, montrer que si α vérifie $0 \leq \alpha < 1$, alors on a : $\overline{\mathcal{C}_\alpha} \subset \mathcal{C}$.

b. À l'aide des résultats de la question 9, montrer que $P(\mathcal{C}) = 1$.

c. Pour tout réel $\varepsilon > 0$, on considère l'événement $E(\varepsilon)$ défini par : $E(\varepsilon) = \bigcap_{N=1}^{+\infty} \left(\bigcup_{n=N}^{+\infty} [|K - K_n| > \varepsilon] \right)$.

Montrer que $P(E(\varepsilon)) = 0$ et en déduire que la suite de variables aléatoires $(K_n)_{n \in \mathbf{N}^*}$ converge en probabilité vers la variable aléatoire K .

On admet sans démonstration que la suite $(K_n)_{n \in \mathbf{N}^*}$ converge alors en loi vers K .

Dans les questions 11 et 12, on considère une suite $(B_n)_{n \in \mathbf{N}^*}$ de variables aléatoires indépendantes suivant chacune la loi de Bernoulli de paramètre $\frac{1}{2}$. Pour tout $n \in \mathbf{N}^*$, on pose : $H_n = \sum_{k=1}^n \frac{B_k}{k}$.

11. a. Pour tout $n \in \mathbf{N}^*$, déterminer l'espérance et la variance de H_n et trouver leurs limites respectives lorsque n tend vers $+\infty$.

b. Montrer que, quel que soit le réel $r > 0$, on a : $\lim_{n \rightarrow +\infty} P(H_n \leq r) = 0$.

c. La fonction Scilab suivante, dont le script est incomplet (ligne 5) permet d'effectuer p simulations de la variable aléatoire $H_n - h_n(\mathcal{I})$, où $h_n(\mathcal{I})$ a été définie dans la partie II (préambule et question 6).

```

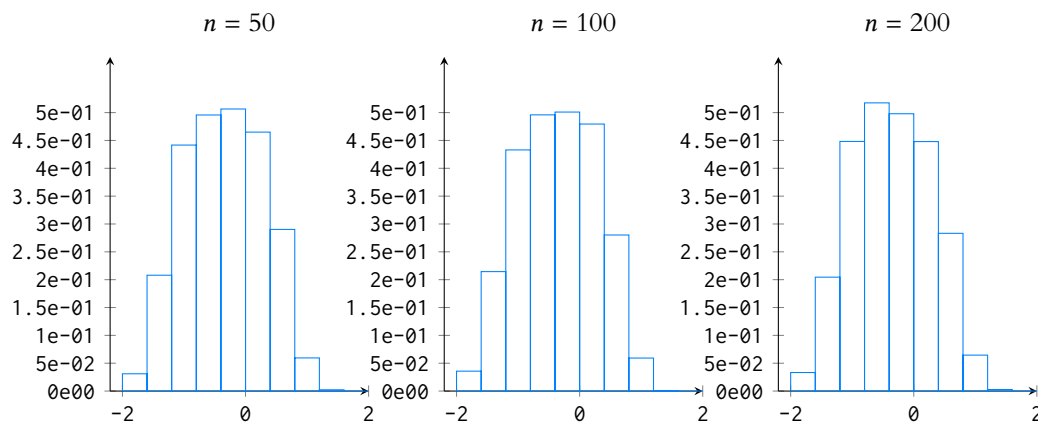
1 function y = simul(n,p)
2     y = zeros(p,1) ;
3     for i =1 :p
4         for k=1 :n
5             y(i,1) = y(i,1) + (grand(1,1,'bin',1,0.5)+.....)/k ;
6         end
7     end
8 endfunction

```

i. Compléter la ligne 5.

ii. Les trois histogrammes suivants représentent la distribution simulée de la variable aléatoire $H_n - h_n(\mathcal{I})$ pour $n = 50$, $n = 100$ et $n = 200$. Par quelles instructions ont-ils pu être obtenus ?

iii. Pourquoi ces histogrammes suggèrent-ils une convergence en loi de la suite $(H_n - h_n(\mathcal{I}))_{n \in \mathbf{N}^*}$?



12. Pour tout $n \in \mathbf{N}^*$, on pose : $B'_n = \frac{1 + X_n}{2}$.

a. Justifier pour tout $n \in \mathbf{N}^*$, la relation : $\sum_{k=1}^n \frac{B'_k}{k} - h_n(\mathcal{I}) = \frac{K_n}{2} + \sum_{k=1}^n \frac{(-1)^k}{2k}$.

b. En déduire que la suite de variables aléatoires $(H_n - h_n(\mathcal{I}))_{n \in \mathbf{N}^*}$ converge en loi vers une variable aléatoire de la forme $\lambda K + \mu$, où λ et μ sont des réels dont on précisera la valeur.

CORRECTION DU CONCOURS BLANC 2

Partie I. Séries télescopiques.

1.a. Pour $n \geq 1$, on a $\frac{1}{n(n+1)} = \frac{1}{n} - \frac{1}{n+1}$.

Et donc pour $N \in \mathbf{N}^*$,

$$\sum_{n=1}^N \frac{1}{n(n+1)} = \sum_{n=1}^N \left(\frac{1}{n} - \frac{1}{n+1} \right) = \sum_{n=1}^N \frac{1}{n} - \sum_{n=1}^N \frac{1}{n+1} = \sum_{n=1}^N \frac{1}{n} - \sum_{n=2}^{N+1} \frac{1}{n} = 1 - \frac{1}{N+1} \xrightarrow{N \rightarrow +\infty} 1.$$

Et donc la série $\sum_{n \geq 1} \frac{1}{n(n+1)}$ converge, et sa somme vaut 1.

1.b. Dans cet exemple, les X_n sont des variables certaines égales à 1.

2.a. Puisque f_n est une densité, son intégrale sur \mathbf{R} doit être égale à 1. Or on a

$$\begin{aligned} \int_{-1/(n+1)}^0 f_n(x) dx &= \int_{-1/(n+1)}^0 (1 + (n+1)x) dx = \left[x + \frac{(n+1)x^2}{2} \right]_{-1/(n+1)}^0 \\ &= \frac{1}{n+1} - \frac{1}{2(n+1)} = \frac{1}{2(n+1)}. \end{aligned}$$

De même,

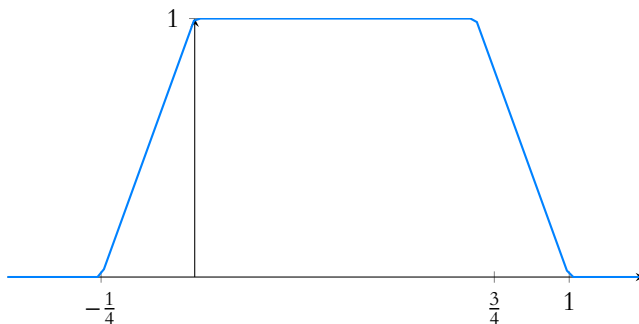
$$\int_{n/(n+1)}^1 f_n(x) dx = \int_{n/(n+1)}^1 (n+1)(1-x) dx = \left[-(n+1) \frac{(1-x)^2}{2} \right]_{n/(n+1)}^1 = \frac{1}{2(n+1)}.$$

Enfin, $\int_0^{n/(n+1)} c_n dx = \frac{n}{n+1} c_n$.

Et donc il vient, par la relation de Chasles,

$$\begin{aligned} \int_{-\infty}^{+\infty} f_n(x) dx = 1 &\Leftrightarrow \frac{1}{2(n+1)} + \frac{nc_n}{n(n+1)} + \frac{1}{2(n+1)} = 1 \\ &\Leftrightarrow c_n \frac{n}{n+1} = \frac{n}{n+1} \\ &\Leftrightarrow \boxed{c_n = 1}. \end{aligned}$$

Le graphique de la fonction f_3 est alors



2.b. Pour $x \in \mathbf{R}$, on a $F_n(x) = \int_{-\infty}^x f_n(t) dt$.

Donc déjà, pour $x \leq -\frac{1}{n+1}$, $F_n(x) = 0$, et pour $x \geq 1$, $F_n(x) = 1$.

Pour $x \in \left[-\frac{1}{n+1}, 0 \right]$,

$$F_n(x) = \int_{-\infty}^{-1/(n+1)} f_n(t) dt + \int_{-1/(n+1)}^x (1 + (n+1)t) dt = \frac{1}{2} \left(x + \frac{1}{n+1} \right) (1 + (n+1)x).$$

Convergence

Notons que s'il fallait juste établir la convergence, il serait plus rapide d'utiliser un équivalent. Mais puisqu'on nous demande de calculer la somme, il est indispensable de passer par les sommes partielles, qui nous permettent à la fois de prouver la convergence et de calculer la somme.

Astuce

Il n'est pas indispensable de passer par une primitive : l'intégrale que nous calculons n'est rien d'autre que l'aire d'un triangle de base $\frac{1}{n+1}$ et de hauteur 1 (voir figure ci-dessous), il n'est alors pas interdit d'utiliser une formule apprise au collège !

Pour $x \in \left[0, \frac{n}{n+1}\right]$, on a $F_n(x) = \underbrace{\int_{-\infty}^0 f(t) dt}_{=F_n(0)} + \int_0^x 1 dt = \frac{1}{2(n+1)} + x$.

Enfin, pour $x \in \left[\frac{n}{n+1}, 1\right]$, on a

$$\begin{aligned} F_n(x) &= \int_{-\infty}^{+\infty} f_n(t) dt - \int_x^{+\infty} f_n(t) dt \\ &= 1 - \int_x^1 (n+1)(1-t) dt \\ &= 1 - \frac{(n+1)(1-x)(1-x)}{2} \\ &= 1 - \frac{n+1}{2}(1-x)^2. \end{aligned}$$

Ainsi, la fonction de répartition F_n est donnée par

$$F_n(x) = \begin{cases} 0 & \text{si } x \leq -\frac{1}{n+1} \\ \frac{1}{2} \left(x + \frac{1}{n+1}\right) (1 + (n+1)x) & \text{si } -\frac{1}{n+1} \leq x \leq 0 \\ \frac{1}{2(n+1)} + x & \text{si } 0 \leq x \leq \frac{n}{n+1} \\ 1 - \frac{n+1}{2}(1-x)^2 & \text{si } \frac{n}{n+1} \leq x \leq 1 \\ 1 & \text{si } x \geq 1 \end{cases}$$

Il est évident que F_n est de classe \mathcal{C}^1 sauf peut-être en $-\frac{1}{n+1}$, en 0, en $\frac{n}{n+1}$ et en 1.

On pourrait vérifier qu'en tous ces points, F_n est dérivable à gauche et à droite, et que les deux dérivées latérales sont égales.

Il est plus simple de remarquer que f_n est continue en tous ces points, et donc sur \mathbf{R} , et alors par le théorème fondamental de l'analyse,

$$F_n : x \mapsto \int_{-\infty}^x f_n(t) dt = \int_{-1/(n+1)}^x f_n(t) dt$$

est une fonction de classe \mathcal{C}^1 sur \mathbf{R} .

2.c. Soit $x \in \mathbf{R}$.

• Si $x < 0$, alors pour n suffisamment grand, $x < -\frac{1}{n+1}$ de sorte que $F_n(x) = 0$. Et donc

$$\lim_{n \rightarrow +\infty} F_n(x) = 0.$$

• Si $x \in [0, 1]$, alors puisque $\frac{n}{n+1} \xrightarrow{n \rightarrow +\infty} 1$, pour n suffisamment grand, $x \leq \frac{n}{n+1}$, de sorte

$$\text{que } F_n(x) = \frac{1}{2(n+1)} + x \xrightarrow{n \rightarrow +\infty} x.$$

• Enfin, si $x \geq 1$, alors $F_n(x) = 1 \xrightarrow{n \rightarrow +\infty} 1$.

$$\text{Ainsi, } F_n(x) \xrightarrow{n \rightarrow +\infty} \begin{cases} 0 & \text{si } x < 0 \\ x & \text{si } 0 \leq x \leq 1 \\ 1 & \text{si } x > 1 \end{cases}$$

Nous reconnaissons là la fonction de répartition d'une loi uniforme sur $[0, 1]$, et donc

$$Y_n \xrightarrow{\mathcal{L}} Y, \text{ où } Y \text{ est une variable aléatoire qui suit la loi uniforme sur } [0, 1].$$

3.a. Soit $\varepsilon > 0$. Alors

$$P\left(\left|\frac{X_{n+1}}{n+1}\right| \geq \varepsilon\right) = P(|X_{n+1}| \geq (n+1)\varepsilon) = \int_{-\infty}^{-(n+1)\varepsilon} f(t) dt + \int_{(n+1)\varepsilon}^{+\infty} f(t) dt.$$

Mais $\lim_{n \rightarrow +\infty} \int_{-\infty}^{-(n+1)\varepsilon} f(t) dt = 0$, puisqu'il s'agit du reste d'une intégrale convergente¹.

Et pour les mêmes raisons, $\lim_{n \rightarrow +\infty} \int_{(n+1)\varepsilon}^{+\infty} f(t) dt = 0$.

Donc $\lim_{n \rightarrow +\infty} P\left(\left|\frac{X_{n+1}}{n+1}\right| \geq \varepsilon\right) = 0$, ce qui signifie bien que $\frac{X_{n+1}}{n+1} \xrightarrow{P} 0$.

Graphiquement

Là aussi il est permis de raisonner graphiquement, l'intégrale que nous calculons est l'aire d'un triangle, de hauteur $f_n(x) = (n+1)(1-x)$ et de base $(1-x)$.

Continuité

La continuité de f_n serait tout de même à vérifier par des calculs de limite, mais ceux-ci sont quand même très simples. À vous de les faire si vous n'êtes pas convaincu !

Rappel

La limite d'une suite ne dépend pas de ses premiers termes, et donc supposer que n est suffisamment grand ne change pas la limite.

¹ L'intégrale $\int_{-\infty}^0 f(t) dt$ est bien convergente puisque f est une densité.

3.b. Puisque X_1 et X ont même loi, $X_1 \xrightarrow{\mathcal{L}} X$, et donc par le théorème de Slutsky,

$$D_n = \underbrace{X_1}_{\xrightarrow{\mathcal{L}} X} - \underbrace{\frac{X_{n+1}}{n+1}}_{\xrightarrow{P} 0} \xrightarrow{\mathcal{L}} X - 0 = X.$$

3.c. Par transformation affine, une densité de $-\frac{X_{n+1}}{n+1}$ est $h_n : t \mapsto (n+1)f(-(n+1)t)$.

Or, les variables X_1 et $-\frac{X_{n+1}}{n+1}$ sont indépendantes, et X_1 possède une densité bornée², donc par convolution, une densité de D_n est

$$f_{D_n} : x \mapsto \int_{-\infty}^{+\infty} f(t)h_n(x-t) dt = \boxed{(n+1) \int_{-\infty}^{+\infty} f(t)f((n+1)(t-x)) dt.}$$

3.d. Supposons que la loi de référence soit une loi uniforme sur $[0, 1]$.

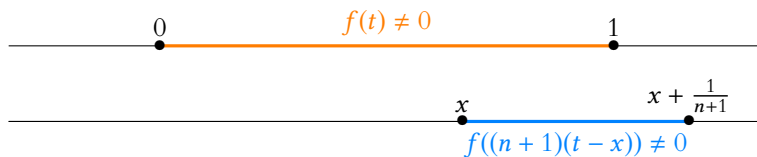
$$\text{Alors } f(x) = \begin{cases} x & \text{si } 0 \leq x \leq 1 \\ 0 & \text{sinon} \end{cases}.$$

Et par la question précédente, une densité de D_n est donc

$$f_{D_n} : x \mapsto (n+1) \int_{-\infty}^{+\infty} f(t)f((n+1)(t-x)) dt.$$

Or, on a $f(t) \neq 0 \Leftrightarrow 0 \leq t \leq 1$.

$$\text{Et } f((n+1)(t-x)) \neq 0 \Leftrightarrow 0 \leq (n+1)(t-x) \leq 1 \Leftrightarrow x \leq t \leq x + \frac{1}{n+1}.$$



Ainsi, pour $x \leq -\frac{1}{n+1}$, $f_{D_n}(x) = 0$.

$$\text{Pour } x \in \left[-\frac{1}{n+1}, 0\right], f_{D_n}(x) = (n+1) \int_0^{x+\frac{1}{n+1}} 1 dt = 1 + (n+1)x.$$

$$\text{Pour } x \in \left[0, \frac{n}{n+1}\right], f_{D_n}(x) = (n+1) \int_x^{x+\frac{1}{n+1}} 1 dt = 1.$$

$$\text{Et pour } x \in \left[\frac{n}{n+1}, 1\right], f_{D_n}(x) = (n+1) \int_x^1 1 dt = (n+1)(1-x).$$

Ainsi, $f_{D_n} = f_n$.

Autrement dit, D_n a même loi que la variable Y_n de la question 2.

Et donc en particulier, $Y_n \xrightarrow{\mathcal{L}} X$, qui suit une loi uniforme sur $[0, 1]$.

4.a. Puisque les X_i sont indépendantes, $\frac{X_n}{n}$ et $\frac{X_{n+1}}{n+1}$ sont également indépendantes. Par stabilité des lois normales, on a donc

$$U_n = \frac{X_n}{n} - \frac{X_{n+1}}{n+1} \Leftrightarrow \mathcal{N}\left(0 - 0, \frac{1}{n^2} + \left(\frac{-1}{n+1}\right)^2\right) = \boxed{\mathcal{N}\left(0, \frac{1}{n^2} + \frac{1}{(n+1)^2}\right)}.$$

4.b. On a, pour $N \in \mathbf{N}^*$,

$$\sum_{n=1}^N U_n = \sum_{n=1}^N \left(\frac{X_n}{n} - \frac{X_{n+1}}{n+1}\right) = X_1 - \frac{X_{N+1}}{N+1}.$$

Mais alors, le résultat de la question 3.b, qui s'applique car la densité d'une loi normale centrée réduite est bornée, montre que

$$\sum_{n=1}^N U_n \xrightarrow{\mathcal{L}} X_1.$$

Détails

$X_1 \xrightarrow{\mathcal{L}} X$ signifie que la suite constante égale à X_1 converge en loi vers X .

² Rappelons qu'il s'agit d'une des conditions auxquelles le théorème de convolution s'applique.

Intuition

À la question 3.b, nous avons prouvé que D_n converge en loi vers X .

Et donc si nous voulons retrouver le résultat de la question 2.c, avec une convergence en loi vers une loi $\mathcal{U}([0, 1])$, il faut que X suive cette loi uniforme.

Détails

Dans cette intégrale et les suivantes, l'intégrande vaut 1, car lorsque $f(t)f((n+1)(t-x))$ est non nul, il vaut $1 \times 1 = 1$.

⚠ Danger !

Attention au signe moins : si l'on soustrait bien les espérances, il faut tout de même ajouter les variances.

Et donc $\boxed{\text{la s\u00e9rie } \sum_{n \geq 1} U_n \text{ converge en loi.}}$

4.c.i. Puisque les U'_k sont ind\u00e9pendantes, et que $U'_k \hookrightarrow \mathcal{N}\left(0, \frac{1}{k^2} + \frac{1}{(k+1)^2}\right)$, nous pouvons de nouveau utiliser la stabilit\u00e9 des lois normales pour d\u00e9terminer la loi de T'_n :

$$T'_n \hookrightarrow \mathcal{N}\left(0, \sum_{k=1}^n \left(\frac{1}{k^2} + \frac{1}{(k+1)^2}\right)\right) = \mathcal{N}\left(0, \sum_{k=1}^n \frac{1}{k^2} + \sum_{k=2}^{n+1} \frac{1}{k^2}\right).$$

La variable al\u00e9atoire $\frac{T'_n}{\sqrt{V(T'_n)}}$ suit alors une loi normale centr\u00e9e r\u00e9duite, de sorte que la fonction de r\u00e9partition de T'_n est donn\u00e9e par :

$$\forall x \in \mathbf{R}, F_{T'_n}(x) = P(T'_n \leq x) = P\left(\frac{T'_n}{\sqrt{V(T'_n)}} \leq \frac{x}{\sqrt{V(T'_n)}}\right) = \Phi\left(\frac{x}{\sqrt{V(T'_n)}}\right).$$

Lorsque $n \rightarrow +\infty$,

$$V(T'_n) \xrightarrow{n \rightarrow +\infty} \sum_{k=1}^{+\infty} \frac{1}{k^2} + \sum_{k=2}^{+\infty} \frac{1}{k^2} = 2\zeta(2) - 1 = \frac{\pi^2}{3} - 1.$$

Et donc $F_{T'_n}(x) \xrightarrow{n \rightarrow +\infty} \Phi\left(\frac{x}{\sqrt{\frac{\pi^2}{3} - 1}}\right)$.

Nous reconnaissons alors la fonction de r\u00e9partition d'une loi normale centr\u00e9e de variance $\frac{\pi^2}{3} - 1$, et donc $\boxed{T'_n \xrightarrow{\mathcal{L}} T, \text{ o\u00f9 } T \hookrightarrow \mathcal{N}\left(0, \frac{\pi^2}{3} - 1\right)}$.

4.c.ii. La question est quelque peu \u00e9trange, nous venons d'\u00e9tudier deux suites diff\u00e9rentes, qui n'ont donc pas la m\u00eame limite en loi, mais en quoi ceci serait contradictoire ?

Un point pourrait tout de m\u00eame choquer : U_k et U'_k ont la m\u00eame loi, mais pas $\sum_{k=1}^n U_k$ et

$$\sum_{k=1}^n U'_k.$$

Ceci tient au fait que la loi d'une somme de variables al\u00e9atoires d\u00e9pend de la loi conjointe de ces variables al\u00e9atoires (et donc des \u00e9ventuels liens qui existent ou non entre ces variables) et pas seulement des lois de chacune des variables constituant la somme.

Or ici, les U'_k sont ind\u00e9pendantes, alors que les U_k ne le sont pas.

Partie II. S\u00e9ries harmoniques «lacunaires»

5.a. On a $h_n(\mathcal{D}) = \sum_{k=1}^n \frac{\mathbb{1}_{\mathcal{D}}(k)}{k}$.

Or, un entier k est dans \mathcal{D} si et seulement si il existe $p \in \mathbf{N}$ tel que $k = p^2$. Et si $k \leq n$, alors n\u00e9cessairement $p \leq \sqrt{n} \Leftrightarrow p \leq \lfloor \sqrt{n} \rfloor$.

Autrement dit, $\boxed{h_n(\mathcal{D}) = \sum_{p=1}^{\lfloor \sqrt{n} \rfloor} \frac{1}{p^2}}$.

5.b. Lorsque $n \rightarrow +\infty$, il vient donc $h_n(\mathcal{D}) \xrightarrow{n \rightarrow +\infty} \sum_{p=1}^{+\infty} \frac{1}{p^2} = \zeta(2) = \frac{\pi^2}{6}$.

Ceci signifie donc que la s\u00e9rie de terme g\u00e9n\u00e9ral $\frac{\mathbb{1}_{\mathcal{D}}(k)}{k}$ converge et que $\boxed{\sum_{k=1}^{+\infty} \frac{\mathbb{1}_{\mathcal{D}}(k)}{k} = \frac{\pi^2}{6}}$.

5.c. Montrons que $\mathcal{D} \cap \mathcal{T} = \{m \in \mathbf{N}^* \text{ tel que } m^{1/6} \in \mathbf{N}^*\} = \{k^6, k \in \mathbf{N}^*\}$.

D\u00e9j\u00e0, si $n = k^6$, alors $n = (k^3)^2 \in \mathcal{D}$ et $n = (k^2)^3 \in \mathcal{T}$.

Et donc $\{k^6, k \in \mathbf{N}^*\} \subset \mathcal{D} \cap \mathcal{T}$.

Inversement, soit $n \in \mathcal{D} \cap \mathcal{T}$. Il existe alors deux entiers p et q tels que $n = p^2 = q^3$.

D\u00e9tails

Le -1 vient du fait que la seconde s\u00e9rie commence \u00e0 $k = 2$ et non \u00e0 $k = 1$.

Non ind\u00e9p.

Il faudrait travailler un peu pour prouver que les U_k ne sont pas ind\u00e9pendantes (par exemple calculer une covariance), mais en tous cas il est clair qu'aucun argument facile ne prouve leur ind\u00e9pendance.

Montrons que $\sqrt{q} \in \mathbf{N}$ (c'est-à-dire que $q \in \mathcal{D}$), de sorte que $n = (\sqrt{q})^6$.

Raisonnons par l'absurde en supposant que $q \notin \mathcal{D}$.

Alors, le résultat admis donné dans l'énoncé nous dit \sqrt{q} est irrationnel.

Mais alors, $q\sqrt{q}$ est également un nombre irrationnel. En effet, s'il était rationnel, alors

$\sqrt{q} = \frac{q\sqrt{q}}{q}$ serait rationnel car quotient de deux rationnels.

Mais $n = p^2 = (q\sqrt{q})^2$, de sorte que $p = q\sqrt{q}$.

Et donc $\sqrt{q} = \frac{p}{q}$ est rationnel, ce qui contredit notre hypothèse de départ.

Ainsi, \sqrt{q} est un entier. Et donc $n = \sqrt{q}^6 \in \{k^6, k \in \mathbf{N}\}$.

Et donc nous venons de prouver que $\mathcal{D} \cap \mathcal{T} = \{k^6, k \in \mathbf{N}\}$.

5.d. On a $\mathbb{1}_{\mathcal{D} \cup \mathcal{T}} = \mathbb{1}_{\mathcal{D}} + \mathbb{1}_{\mathcal{T}} - \mathbb{1}_{\mathcal{D} \cap \mathcal{T}}$.

En effet, si $k \in \mathbf{N}$, on a quatre cas de figure :

- soit $k \in \mathcal{D} \cap \mathcal{T}$ et alors $\mathbb{1}_{\mathcal{D}}(k) + \mathbb{1}_{\mathcal{T}}(k) - \mathbb{1}_{\mathcal{D} \cap \mathcal{T}}(k) = 1 + 1 - 1 = 1 = \mathbb{1}_{\mathcal{D} \cup \mathcal{T}}(k)$.
- soit $k \in \mathcal{D}$ et $k \notin \mathcal{T}$ et alors $\mathbb{1}_{\mathcal{D}}(k) + \mathbb{1}_{\mathcal{T}}(k) - \mathbb{1}_{\mathcal{D} \cap \mathcal{T}}(k) = 1 + 0 - 0 = 1 = \mathbb{1}_{\mathcal{D} \cup \mathcal{T}}(k)$.
- soit $k \in \mathcal{T}$ et $k \notin \mathcal{D}$ et alors $\mathbb{1}_{\mathcal{D}}(k) + \mathbb{1}_{\mathcal{T}}(k) - \mathbb{1}_{\mathcal{D} \cap \mathcal{T}}(k) = 0 + 1 - 0 = 1 = \mathbb{1}_{\mathcal{D} \cup \mathcal{T}}(k)$.
- soit enfin $k \notin \mathcal{D} \cup \mathcal{T}$ et alors $\mathbb{1}_{\mathcal{D}}(k) + \mathbb{1}_{\mathcal{T}}(k) - \mathbb{1}_{\mathcal{D} \cap \mathcal{T}}(k) = 0 + 0 - 0 = 0 = \mathbb{1}_{\mathcal{D} \cup \mathcal{T}}(k)$.

Donc pour $n \in \mathbf{N}$, on a

$$h_n(\mathcal{D} \cup \mathcal{T}) = \sum_{k=1}^n \frac{\mathbb{1}_{\mathcal{D} \cup \mathcal{T}}(k)}{k} = \sum_{k=1}^n \frac{\mathbb{1}_{\mathcal{D}}(k)}{k} + \sum_{k=1}^n \frac{\mathbb{1}_{\mathcal{T}}(k)}{k} - \sum_{k=1}^n \frac{\mathbb{1}_{\mathcal{D} \cap \mathcal{T}}(k)}{k} = h_n(\mathcal{D}) + h_n(\mathcal{T}) - h_n(\mathcal{D} \cap \mathcal{T}).$$

Or, comme on a prouvé que $\lim_{n \rightarrow +\infty} h_n(\mathcal{D}) = \zeta(2)$, on peut prouver que $\lim_{n \rightarrow +\infty} h_n(\mathcal{T}) = \zeta(3)$.

Et en utilisant le fait que $\mathcal{D} \cap \mathcal{T} = \{k^6, k \in \mathbf{N}^*\}$, on prouve que $\lim_{n \rightarrow +\infty} h_n(\mathcal{D} \cap \mathcal{T}) = \zeta(6)$.

Il vient donc enfin

$$\lim_{n \rightarrow +\infty} h_n(\mathcal{D} \cup \mathcal{T}) = \sum_{k=1}^{+\infty} \frac{\mathbb{1}_{\mathcal{D} \cup \mathcal{T}}(k)}{k} = \zeta(2) + \zeta(3) - \zeta(6).$$

6.a. La fonction $t \mapsto -\frac{1}{2t-1}$ est croissante sur $[n, n+1]$ et donc pour tout $t \in [n, n+1]$, on a

$$\frac{1}{2n-1} - \frac{1}{2n-1} \leq \frac{1}{2n-1} - \frac{1}{2t-1} \leq \frac{1}{n-1} - \frac{1}{2(n+1)-1}.$$

Et donc par croissance de l'intégrale,

$$0 \leq u_n \leq \frac{1}{2n-1} - \frac{1}{2n+1} = \frac{2}{(2n-1)(2n+1)} \leq \frac{2}{(2n-1)^2}.$$

6.b. Pour $k \geq 1$, on a $u_k = \frac{1}{2k-1} - \int_k^{k+1} \frac{1}{2t-1} dt$.

Et donc

$$\sum_{k=1}^{\lfloor \frac{n+1}{2} \rfloor} u_k = \sum_{k=1}^{\lfloor \frac{n+1}{2} \rfloor} \frac{1}{2k-1} - \sum_{k=1}^{\lfloor \frac{n+1}{2} \rfloor} \int_k^{k+1} \frac{1}{2t-1} dt = \sum_{k=1}^{\lfloor \frac{n+1}{2} \rfloor} \frac{1}{2k-1} - \int_1^{\lfloor \frac{n+1}{2} \rfloor + 1} \frac{1}{2t-1} dt.$$

Mais $\lfloor \frac{n+1}{2} \rfloor + 1 = \lfloor \frac{n+3}{2} \rfloor$ et donc $\int_1^{\lfloor \frac{n+1}{2} \rfloor + 1} \frac{1}{2t-1} dt = \int_1^{\lfloor \frac{n+3}{2} \rfloor} \frac{1}{2t-1} dt$.

Donc il suffit de prouver que $\sum_{k=1}^{\lfloor \frac{n+1}{2} \rfloor} \frac{1}{2k-1} = h_n(\mathcal{I})$.

Mais par définition,

$$h_n(\mathcal{I}) = \sum_{k=1}^n \frac{\mathbb{1}_{\mathcal{I}}(k)}{k} = \sum_{1 \leq 2k-1 \leq n} \frac{1}{2k-1}.$$

Irrationalité

Vous avez probablement déjà rencontré une preuve de l'irrationalité de $\sqrt{2}$. On prouve de manière similaire que si \sqrt{n} n'est pas entier, alors il est irrationnel.

Remarque

Cette relation n'est pas sans rappeler les formules

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

et

$$\text{Card}(A \cup B) = \text{Card}(A) + \text{Card}(B) - \text{Card}(A \cap B)$$

Pour la culture

La valeur de $\zeta(n)$ est connue lorsque n est un entier pair, par exemple $\zeta(2) = \frac{\pi^2}{6}$ et $\zeta(6) = \frac{\pi^6}{945}$.

En revanche, on sait beaucoup moins de choses lorsque n est un entier impair. On sait que $\zeta(3)$ est irrationnel, et on ne sait à peu près rien de $\zeta(5), \zeta(7), \dots$

Astuce

On intègre des constantes, il n'est donc pas nécessaire de faire apparaître des primitives :

$$\int_a^b \lambda dt = \lambda(b-a).$$

Partie entière

Pour tout réel x , on a

$$\lfloor x \rfloor + 1 = \lfloor x + 1 \rfloor.$$

Or, $2k - 1 \leq n \Leftrightarrow k \leq \frac{n+1}{2}$.

Mais k est un entier et donc $k \leq \frac{n+1}{2} \Leftrightarrow k \leq \left\lfloor \frac{n+1}{2} \right\rfloor$.

Et ainsi, $h_n(\mathcal{I}) = \sum_{k=1}^{\lfloor \frac{n+1}{2} \rfloor} \frac{1}{2k-1}$. Au final, on a donc bien

$$h_n(\mathcal{I}) = \sum_{k=1}^{\lfloor \frac{n+1}{2} \rfloor} u_k + \int_1^{\lfloor \frac{n+3}{2} \rfloor} \frac{1}{2t-1} dt.$$

6.c. La fonction $t \mapsto \ln(t)$ est concave sur \mathbf{R}_+^* , et donc sa courbe représentative est située sous ses tangentes.

En particulier, la tangente au point d'abscisse 0 est la droite d'équation $y = x - 1$.

Et donc pour tout $t \in \mathbf{R}_+^*$, $\ln(t) \leq t - 1$.

Par conséquent, il vient

$$\ln\left(\frac{1}{n} \left(2 \left\lfloor \frac{n+3}{2} \right\rfloor - 1\right)\right) \leq \frac{1}{n} \left(2 \left\lfloor \frac{n+3}{2} \right\rfloor - 1\right) - 1 \leq \frac{1}{n} \left(2 \frac{n+3}{2} - 1\right) - 1 \leq 1 + \frac{2}{n} - 1 \leq \frac{2}{n}.$$

6.d. La série de terme général $\frac{2}{(2n-1)^2}$ converge car $\frac{2}{(2n-1)^2} \underset{n \rightarrow +\infty}{\sim} \frac{1}{2n^2}$.

Et donc, grâce à l'encadrement obtenu à la question 6.a, par comparaison de séries à termes positifs, la série de terme général u_n converge également.

D'après la question 6.b,

$$\begin{aligned} h_n(\mathcal{I}) - \ln(\sqrt{n}) &= \sum_{k=1}^{\lfloor \frac{n+1}{2} \rfloor} u_k + \int_1^{\lfloor \frac{n+3}{2} \rfloor} \frac{1}{2t-1} dt - \ln(\sqrt{n}) \\ &= \sum_{k=1}^{\lfloor \frac{n+1}{2} \rfloor} u_k + \left[\frac{1}{2} \ln(2t-1) \right]_1^{\lfloor \frac{n+3}{2} \rfloor} - \ln(\sqrt{n}) \\ &= \sum_{k=1}^{\lfloor \frac{n+1}{2} \rfloor} u_k + \frac{1}{2} \ln\left(2 \left\lfloor \frac{n+3}{2} \right\rfloor - 1\right) - \frac{1}{2} \ln(n) \\ &= \sum_{k=1}^{\lfloor \frac{n+1}{2} \rfloor} u_k + \frac{1}{2} \ln\left(\frac{1}{n} \left(2 \left\lfloor \frac{n+3}{2} \right\rfloor - 1\right)\right). \end{aligned}$$

Or, le théorème des gendarmes, appliqué à l'encadrement de la question 6.c prouve que

$$\ln\left(\frac{1}{n} \left(2 \left\lfloor \frac{n+3}{2} \right\rfloor - 1\right)\right) \underset{n \rightarrow +\infty}{\rightarrow} 0.$$

Et puisque $\left\lfloor \frac{n+1}{2} \right\rfloor \underset{n \rightarrow +\infty}{\rightarrow} +\infty$, en passant à la limite, on obtient

$$\lim_{n \rightarrow +\infty} (h_n(\mathcal{I}) - \ln(\sqrt{n})) = \sum_{k=1}^{+\infty} u_k = \delta.$$

6.e. Nous avons prouvé à la question 6.a que $u_k \leq \frac{1}{2k-1} - \frac{1}{2k+1}$.

Et donc pour $N \geq n+1$, il vient

$$\sum_{k=n+1}^N u_k \leq \sum_{k=n+1}^N \frac{1}{2k-1} - \sum_{k=n+1}^N \frac{1}{2k+1} = \frac{1}{2(n+1)-1} - \frac{1}{2N+1}.$$

Et donc en faisant tendre N vers $+\infty$, il vient alors

$$\sum_{k=n+1}^{+\infty} u_k \leq \frac{1}{2n+1} \leq \frac{1}{2n-1}.$$

Partie entière

Rappelons que par définition, $\lfloor x \rfloor$ est le plus grand entier inférieur ou égal à x .

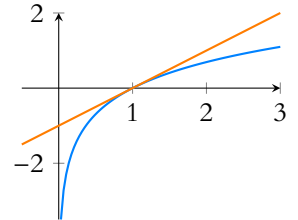


FIGURE 24.1- La fonction \ln et sa tangente en 1.

Partie entière

Rappelons que pour tout réel x , on a

$$\lfloor x \rfloor \leq x.$$

Somme télescopique.

Limite

Notons que le passage à la limite est justifié car nous savons que les deux membres de l'inégalité admettent une limite.

6.f. Puisque $\delta = \sum_{k=1}^{+\infty} u_k$, on a donc

$$\delta - (h_n(\mathcal{I}) - \ln(\sqrt{n})) = \sum_{k=\lfloor \frac{n+1}{2} \rfloor + 1}^{+\infty} u_k - \frac{1}{2} \ln \left(\frac{1}{n} \left(2 \left\lfloor \frac{n+3}{2} \right\rfloor - 1 \right) \right).$$

Et donc à l'aide des encadrements précédemment établis :

$$0 \leq \sum_{k=\lfloor \frac{n+1}{2} \rfloor + 1}^{+\infty} u_k \leq \frac{1}{2 \lfloor \frac{n+1}{2} \rfloor - 1} \text{ et } 0 \leq \ln \left(\frac{1}{n} \left(2 \left\lfloor \frac{n+3}{2} \right\rfloor - 1 \right) \right) \leq \frac{2}{n}$$

il vient

$$-\frac{1}{2n} \leq \delta - (h_n(\mathcal{I}) - \ln(\sqrt{n})) \leq \frac{1}{2 \lfloor \frac{n+1}{2} \rfloor - 1}.$$

Mais $\left\lfloor \frac{n+1}{2} \right\rfloor \geq \frac{n+1}{2} - 1 \geq \frac{n-1}{2}$ et donc par passage à l'inverse

$$\frac{1}{2 \lfloor \frac{n+1}{2} \rfloor - 1} \leq \frac{1}{n-1-1} = \frac{1}{n-2}.$$

On a donc bien prouvé que

$$\boxed{-\frac{1}{n} \leq \delta - (h_n(\mathcal{I}) - \ln(\sqrt{n})) \leq \frac{1}{n-2}}.$$

6.g.i. Pour $\text{eps} = 0.2$, la boucle `while` va s'exécuter tant que

$$\frac{1}{n-2} > 0.2 \Leftrightarrow n-2 < \underbrace{\frac{1}{0.2}}_{=5} \Leftrightarrow n < 7.$$

Mais puisqu'au départ $n = 3$, et que n augmente de 2 en 2, `s` ne va contenir que deux valeurs successives : sa valeur de départ qui est $h_3(\mathcal{I}) - \ln(\sqrt{3})$ et la valeur $h_5(\mathcal{I}) - \ln(\sqrt{5})$.

6.g.ii. Notons qu'à la ligne 3, le concepteur du programme a choisi d'écrire `log(3)/2` plutôt que `log(sqrt(3))`, ce qui calcule bien entendu la même valeur.

Restons dans le même esprit : pour passer de $h_{n-2}(\mathcal{I}) - \ln(\sqrt{n-2})$ à $h_n(\mathcal{I}) - \ln(\sqrt{n})$, il suffit d'ajouter $\frac{1}{n} + \ln(\sqrt{n-2})$ et de retirer $\ln(\sqrt{n})$. Soit encore d'ajouter $\frac{1}{n} - \frac{1}{2} \ln \left(\frac{n}{n-2} \right)$.

Nous pouvons donc compléter le code de la manière suivante

6 `s = s+1/n - log(n/(n-2))/2`

6.g.iii. Les encadrements de la question 6.f prouvent que

$$-\frac{1}{n-2} \leq \delta - (h_n(\mathcal{I}) - \ln(\sqrt{n})) \leq \frac{1}{n-2} \Leftrightarrow |\delta - (h_n(\mathcal{I}) - \ln(\sqrt{n}))| \leq \frac{1}{n-2}.$$

Et donc si n est tel que $\frac{1}{n-2} \leq \varepsilon$, alors $|\delta - (h_n(\mathcal{I}) - \ln(\sqrt{n}))| \leq \varepsilon$, de sorte que $h_n(\mathcal{I}) - \ln(\sqrt{n})$ est une valeur approchée de δ à ε près.

Partie III. Séries de Riemann alternées.

7.a.i. On a

$$\begin{aligned} \sum_{k=1}^{n-1} s_k \left(\frac{1}{k^\beta} - \frac{1}{(k+1)^\beta} \right) &= \sum_{k=1}^{n-1} \sum_{i=1}^k x_i \left(\frac{1}{k^\beta} - \frac{1}{(k+1)^\beta} \right) \\ &= \sum_{1 \leq i \leq k \leq n-1} x_i \left(\frac{1}{k^\beta} - \frac{1}{(k+1)^\beta} \right) \\ &= \sum_{i=1}^{n-1} x_i \sum_{k=i}^{n-1} \left(\frac{1}{k^\beta} - \frac{1}{(k+1)^\beta} \right) \end{aligned}$$

Rappel

Pour tout x , on a

$$x < \lfloor x \rfloor + 1 \Leftrightarrow \lfloor x \rfloor > x - 1.$$

$$\begin{aligned}
 &= \sum_{i=1}^{n-1} x_i \left(\frac{1}{i^\beta} - \frac{1}{n^\beta} \right) \\
 &= \sum_{i=1}^{n-1} \frac{x_i}{i^\beta} - \frac{1}{n^\beta} \sum_{i=1}^{n-1} x_i.
 \end{aligned}$$

Somme télescopique.

Et donc

$$\frac{1}{n^\beta} \sum_{i=1}^{n-1} x_i + \sum_{k=1}^{n-1} s_k \left(\frac{1}{k^\beta} - \frac{1}{(k+1)^\beta} \right) = \sum_{i=1}^{n-1} \frac{x_i}{i^\beta}.$$

En ajoutant $\frac{x_n}{n^\beta}$ de chaque côté, on obtient donc

$$\boxed{\frac{1}{n^\beta} \underbrace{\sum_{i=1}^n x_i}_{=s_n} + \sum_{k=1}^{n-1} s_k \left(\frac{1}{k^\beta} - \frac{1}{(k+1)^\beta} \right) = \sum_{i=1}^n \frac{x_i}{i^\beta}.}$$

7.a.ii. On a $\left| \frac{s_n}{n^\beta} \right| \leq M \frac{1}{n^{\beta-\alpha}}$ et puisque $\beta > \alpha$, $\frac{s_n}{n^\beta} \xrightarrow{n \rightarrow +\infty} 0$.

D'autre part, $|s_k| \leq Mk^\alpha$ et donc

$$\left| s_k \left(\frac{1}{k^\beta} - \frac{1}{(k+1)^\beta} \right) \right| \leq M \left(\frac{k^\alpha}{k^\beta} - \frac{k^\alpha}{(k+1)^\beta} \right) \leq M \frac{1}{k^{\beta-\alpha}} \left(1 - \frac{1}{\left(1 + \frac{1}{k}\right)^\beta} \right)$$

Mais à l'aide d'un développement limité, on obtient

$$1 - \frac{1}{\left(1 + \frac{1}{k}\right)^\beta} = 1 - \left(1 + \frac{1}{k}\right)^{-\beta} = 1 - \left(1 - \frac{\beta}{k} + o_{k \rightarrow +\infty} \left(\frac{1}{k}\right)\right) = \frac{\beta}{k} + o_{k \rightarrow +\infty} \left(\frac{1}{k}\right) \underset{k \rightarrow +\infty}{\sim} \frac{\beta}{k}.$$

Ainsi, $M \frac{1}{k^{\beta-\alpha}} \left(1 - \frac{1}{\left(1 + \frac{1}{k}\right)^\beta}\right) \underset{k \rightarrow +\infty}{\sim} \frac{M\beta}{k^{\beta-\alpha+1}}$.

Et puisque $\beta - \alpha + 1 > 1$, la série de terme général $M \frac{1}{k^{\beta-\alpha}} \left(1 - \frac{1}{\left(1 + \frac{1}{k}\right)^\beta}\right)$ converge.

Et donc, la série de terme général $s_k \left(\frac{1}{k^\beta} - \frac{1}{(k+1)^\beta} \right)$ converge absolument et donc converge.

Par conséquent, $\sum_{k=1}^{n-1} s_k \left(\frac{1}{k^\beta} - \frac{1}{(k+1)^\beta} \right)$ admet une limite finie lorsque $n \rightarrow +\infty$.

On en déduit donc que $\sum_{k=1}^n \frac{x_k}{k^\beta}$ admet également une limite finie lorsque $n \rightarrow +\infty$, c'est-à-dire que

la série de terme général $\frac{x_k}{k^\beta}$ converge.

7.b. Posons $x_n = (-1)^n$, de sorte que $s_n = \sum_{k=1}^n (-1)^k = \begin{cases} 0 & \text{si } n \text{ est pair} \\ -1 & \text{si } n \text{ est impair} \end{cases}$

Alors, puisque (s_n) est bornée, on a $|s_n| \leq 1 \leq 1n^0$.

Et donc en prenant $\alpha = 0$ et $\beta = x$ dans le résultat de la question précédente, on obtient la convergence de la série $\sum_{n \geq 1} \frac{(-1)^n}{n^x}$.

8.a. On a $e^{tS_n} = e^{t \sum_{i=1}^n X_i} = \prod_{i=1}^n e^{tX_i}$.

Or, les X_i étant indépendantes, les e^{tX_i} le sont aussi et donc³

$$E(e^{tS_n}) = \prod_{i=1}^n E(e^{tX_i}).$$

³ Les e^{tX_i} sont des variables à support fini, donc ont une espérance.

Mais, pour tout $i \in \llbracket 1, n \rrbracket$, par le théorème de transfert,

$$E(e^{tX_i}) = e^t P(X_i = 1) + e^{-t} P(X_i = -1) = \frac{e^t + e^{-t}}{2}.$$

Et donc
$$E(e^{tS_n}) = \left(\frac{e^t + e^{-t}}{2} \right)^n.$$

8.b. Rappelons que pour tout $u \in \mathbf{R}$, $e^u = \sum_{n=0}^{+\infty} \frac{u^n}{n!}$.

Et donc

$$e^t + e^{-t} = \sum_{n=0}^{+\infty} \frac{t^n}{n!} + \sum_{n=0}^{+\infty} \frac{(-t)^n}{n!} = \sum_{n=0}^{+\infty} \frac{1 + (-1)^n}{n!} t^n.$$

Mais $1 + (-1)^n = \begin{cases} 2 & \text{si } n \text{ est pair} \\ 0 & \text{si } n \text{ est impair} \end{cases}$ de sorte que $\frac{e^t + e^{-t}}{2} = \frac{1}{2} \sum_{n \text{ pair}} \frac{t^n}{n!} = \sum_{k=0}^{+\infty} \frac{t^{2k}}{(2k)!}$.

D'autre part, on a également

$$e^{t^2/2} = \sum_{n=0}^{+\infty} \frac{\left(\frac{t^2}{2}\right)^n}{n!} = \sum_{n=0}^{+\infty} \frac{t^{2n}}{2^n n!}.$$

Mais pour $k \in \mathbf{N}$, on a $(2k)! \geq 2^k k!$. Prouvons-le par une récurrence rapide : pour $k = 1$ et $k = 2$, c'est une égalité.

Supposons que $2^k k! \leq (2k)!$. Alors

$$2^{k+1}(k+1)! = 2 \times (k+1)2^k k! \leq 2 \times (k+1) \times (2k)! \leq (2k+1)(2k+2)(2k)! = (2(k+1))!.$$

Donc pour tout $k \in \mathbf{N}$, $2^k k! \leq (2k)!$ et donc $\frac{t^{2k}}{(2k)!} \leq \frac{t^{2k}}{2^k k!}$.

Ainsi, il vient

$$\frac{e^t + e^{-t}}{2} = \sum_{k=0}^{+\infty} \frac{t^{2k}}{(2k)!} \leq \sum_{k=0}^{+\infty} \frac{t^{2k}}{2^k k!} = e^{t^2/2}.$$

8.c. Notons que nous ne pouvons pas directement appliquer l'inégalité de Markov à la variable S_n , car elle n'est pas positive.

En revanche, la variable e^{tS_n} est positive, et $[S_n > s] = [e^{tS_n} > e^{ts}]$. Et donc

$$P(S_n > s) = P(e^{tS_n} > e^{ts}) \leq \frac{E(e^{tS_n})}{e^{ts}} \leq \left(\frac{e^t + e^{-t}}{2} \right)^n \frac{1}{e^{ts}} \leq \frac{e^{nt^2/2}}{e^{ts}} = \exp\left(n \frac{t^2}{2} - ts\right).$$

8.d. L'inégalité que nous venons d'établir est valable pour tout $t \in \mathbf{R}^+$.

Posons donc $f(t) = \exp\left(n \frac{t^2}{2} - ts\right)$. Par croissance de l'exponentielle, f possède les mêmes variations que $t \mapsto n \frac{t^2}{2} - ts$. Or, celle-ci possède un minimum⁴ en $t = \frac{s}{n}$.

$$f'(t) = (nt - s) \exp\left(n \frac{t^2}{2} - ts\right).$$

Donc f admet un minimum en $t = \frac{s}{n}$ et ce minimum vaut

$$f\left(\frac{s}{n}\right) = \exp\left(n \frac{s^2}{2n^2} - s \frac{s}{n}\right) = \exp\left(-\frac{s^2}{2n}\right).$$

Et donc en appliquant l'inégalité précédente pour $t = \frac{s}{n}$, il vient

$$P(S_n > s) \leq \exp\left(-\frac{s^2}{2n}\right).$$

D'autre part, pour tout $i \in \llbracket 1, n \rrbracket$, X_i et $-X_i$ ont même loi. Donc S_n et $-S_n$ ont également même loi.

Ainsi, $P(S_n < -s) = P(-S_n > s) \leq \exp\left(-\frac{s^2}{2n}\right)$. Et donc

$$P(|S_n| > s) = P(S_n > s) + P(S_n < -s) \leq 2 \exp\left(-\frac{s^2}{2n}\right).$$

⁴ Il s'agit d'une formule vue au lycée qui donne l'abscisse du sommet de la parabole d'équation $y = ax^2 + bx + c$.

Pour la culture

◀ Cette inégalité est appelée inégalité de Chernoff.

9.a. Puisque S_k est une variable aléatoire et que la fonction valeur absolue est continue, $|S_k|$ est également une variable aléatoire.
 Et donc, par définition d'une variable aléatoire, $[|S_k| \leq k^\alpha]$ est un élément de \mathcal{A} .
 Donc l'événement contraire, $[|S_k| > k^\alpha] \in \mathcal{A}$.

Donc pour tout $n \in \mathbf{N}^*$, $\bigcup_{k=n}^{+\infty} [|S_k| > k^\alpha] \in \mathcal{A}$.

Et enfin, $\mathcal{C}_\alpha = \bigcap_{n=1}^{+\infty} \bigcup_{k=n}^{+\infty} [|S_k| > k^\alpha] \in \mathcal{A}$.

9.b. D'après la question 8.d, on a

$$0 \leq P(|S_n| > n^\alpha) \leq 2 \exp\left(-\frac{n^{2\alpha}}{2n}\right) = 2 \exp\left(-\frac{n^{2\alpha-1}}{2}\right).$$

Mais $e^n \exp\left(-\frac{n^{2\alpha-1}}{2}\right) = \exp\left(n - \frac{n^{2\alpha-1}}{2}\right)$.

Et puisque $\alpha > \frac{1}{2}$, $n - \frac{n^{2\alpha-1}}{2} \underset{n \rightarrow +\infty}{\sim} -\frac{n^{2\alpha-1}}{2} \xrightarrow{n \rightarrow +\infty} -\infty$.

Et donc $e^n \exp\left(-\frac{n^{2\alpha-1}}{2}\right) \xrightarrow{n \rightarrow +\infty} 0$ de sorte que $\exp\left(-\frac{n^{2\alpha-1}}{2}\right) = o_{n \rightarrow +\infty}(e^{-n})$.

Or, la série de terme général e^{-n} est une série géométrique⁵ convergente, et donc la série de terme général $\exp\left(-\frac{n^{2\alpha-1}}{2}\right)$ converge également.

On en déduit, par majoration, que la série de terme général $P(|S_n| > n^\alpha)$ converge.

9.c. Soit $n \in \mathbf{N}^*$ et $N \geq n$. Alors

$$P\left(\bigcup_{k=n}^N [|S_k| > k^\alpha]\right) \leq \sum_{k=n}^N P(|S_k| > k^\alpha) \leq \sum_{k=n}^{+\infty} P(|S_k| > k^\alpha).$$

Par le théorème de la limite monotone on a donc

$$P\left(\bigcup_{k=n}^{+\infty} [|S_n| > n^\alpha]\right) \leq \sum_{k=n}^{+\infty} P(|S_k| > k^\alpha).$$

Mais ce majorant est le reste d'une série convergente, de sorte que $\lim_{n \rightarrow +\infty} \sum_{k=n}^{+\infty} P(|S_n| > k^\alpha) = 0$.

Donc par le théorème des gendarmes, $\lim_{n \rightarrow +\infty} P\left(\bigcup_{k=n}^{+\infty} [|S_k| > k^\alpha]\right) = 0$.

Enfin, par le théorème de la limite monotone,

$$P(\mathcal{C}_\alpha) = \lim_{n \rightarrow +\infty} P\left(\bigcup_{k=n}^{+\infty} [|S_k| > k^\alpha]\right) = 0.$$

10.a. Pour aborder cette question, il nous va falloir comprendre un peu plus finement ce qu'est l'événement \mathcal{C}_α .

L'événement $\bigcup_{k=n}^{+\infty} [|S_k| > k^\alpha]$ est l'événement «un des $[|S_k| > k^\alpha]$, $k \geq n$ est réalisé».

Et donc \mathcal{C}_α est l'événement «pour tout $n \in \mathbf{N}^*$, il existe un $k \geq n$ tel que $[|S_k| > k^\alpha]$ soit réalisé».

Donc en passant au contraire, si $\omega \in \overline{\mathcal{C}_\alpha}$, cela signifie qu'il existe un $n_0 \in \mathbf{N}^*$ tel que pour tout $k \geq n_0$, $[|S_k(\omega)| \leq k^\alpha]$.

Soit donc $\omega \in \overline{\mathcal{C}_\alpha}$. Alors, à partir d'un certain rang n_0 , $|S_k| \leq k^\alpha$.

Notons $M = \max_{1 \leq k \leq n_0-1} \frac{|S_k(\omega)|}{k^\alpha}$, qui existe bien puisqu'il s'agit du maximum d'un ensemble fini, qui admet donc un plus grand élément.

Alors pour tout $k \in \llbracket 1, n_0 - 1 \rrbracket$, $|S_k| \leq Mk^\alpha$.

Rappels

Une tribu est stable par passage au complémentaire, par union dénombrable et par intersection dénombrable.

⁵ De raison $e^{-1} \in]0, 1[$.

Rappel

Pour toute suite (A_k) d'événements,

$$P\left(\bigcup_{k=n}^{+\infty} A_k\right) = \lim_{N \rightarrow +\infty} P\left(\bigcup_{k=n}^N A_k\right).$$

Rappel

Pour toute suite décroissante (A_k) d'événements,

$$P\left(\bigcap_{k=1}^{+\infty} A_k\right) = \lim_{N \rightarrow +\infty} P(A_N).$$

Remarque

On pourrait encore reformuler ceci en disant que \mathcal{C}_α est l'événement «il existe une infinité de k tels que $|S_k| > k^\alpha$ ». Son contraire est alors «il n'existe qu'un nombre fini de k tels que $|S_k| > k^\alpha$ ».

Et donc pour tout $k \in \mathbf{N}^*$, $|S_k| \leq (M+1)k^\alpha$.

En effet, si $k \leq n_0 - 1$, alors $|S_k| \leq Mk^\alpha \leq (M+1)k^\alpha$. Et si $k \geq n_0$, alors $|S_k| \leq k^\alpha \leq (M+1)k^\alpha$.

Donc, en se rappelant que $S_k = \sum_{i=1}^k X_i$, le résultat de la question 7.a s'applique, avec

$\beta = 1 > \alpha$: la série $\sum_{k \geq 1} \frac{X_k(\omega)}{k}$ converge.

Et donc $\omega \in \mathcal{C}$.

Autrement dit, nous venons de prouver que $\overline{\mathcal{C}_\alpha} \subset \mathcal{C}$.

10.b. Ce qui précède est en particulier vrai pour $\alpha \in]\frac{1}{2}, 1[$, de sorte que

$$P(\overline{\mathcal{C}_\alpha}) \leq P(\mathcal{C}) \leq 1.$$

Et d'après la question 9.c, $P(\overline{\mathcal{C}_\alpha}) = 1 - P(\mathcal{C}_\alpha) = 1$.

Donc nécessairement $P(\mathcal{C}) = 1$.

10.c. Montrons que $E(\varepsilon) \subset \overline{\mathcal{C}}$, où plutôt⁶ que $\mathcal{C} \subset \overline{E(\varepsilon)}$.

Notons à cet effet que $\overline{E(\varepsilon)} = \bigcup_{N=1}^{+\infty} \left(\bigcap_{n=N}^{+\infty} [|K - K_n| \leq \varepsilon] \right)$.

Mais si $\omega \in \mathcal{C}$, alors la série de terme général $\frac{X_k(\omega)}{k}$ converge, de sorte que la suite de ses sommes partielles converge vers $K(\omega)$: $\lim_{n \rightarrow +\infty} K_n(\omega) = K(\omega)$.

Ainsi, il existe $n_0 \in \mathbf{N}^*$ tel que pour $n \geq n_0$, $|K_n(\omega) - K(\omega)| \leq \varepsilon$.

Et en particulier, l'événement $\bigcap_{n=n_0}^{+\infty} [|K_n - K| \leq \varepsilon]$ est réalisé.

Autrement dit, $\omega \in \bigcap_{n=n_0}^{+\infty} [|K_n - K| \leq \varepsilon]$.

Et donc $\omega \in \bigcup_{N=1}^{+\infty} \left(\bigcap_{n=N}^{+\infty} [|K_n - K| \leq \varepsilon] \right) = \overline{E(\varepsilon)}$.

Nous avons donc bien prouvé que $\mathcal{C} \subset \overline{E(\varepsilon)}$ et donc $E(\varepsilon) \subset \overline{\mathcal{C}}$.

Par croissance de la probabilité, on a alors $P(E(\varepsilon)) \leq P(\overline{\mathcal{C}})$, et puisque $P(\mathcal{C}) = 1$, $P(\overline{\mathcal{C}}) = 0$, et donc $P(E(\varepsilon)) = 0$.

Par le théorème de la limite monotone, on a

$$0 = P(E(\varepsilon)) = \lim_{N \rightarrow +\infty} P\left(\bigcup_{n=N}^{+\infty} |K - K_n| > \varepsilon\right).$$

Or, pour tout $N \in \mathbf{N}$,

$$0 \leq P(|K - K_N| > \varepsilon) \leq P\left(\bigcup_{n=N}^{+\infty} |K - K_n| > \varepsilon\right)$$

de sorte que par le théorème des gendarmes, $\lim_{N \rightarrow +\infty} P(|K - K_N| > \varepsilon) = 0$.

Ainsi, $K_n \xrightarrow{P} K$.

11.a. Puisque $E(B_k) = \frac{1}{2}$, par linéarité de l'espérance,

$$E(H_n) = \sum_{k=1}^n \frac{E(B_k)}{k} = \sum_{k=1}^n \frac{1}{2k} = \frac{1}{2} \sum_{k=1}^n \frac{1}{k}.$$

Mais la série de terme général $\frac{1}{k}$ est divergente, et puisqu'il s'agit d'une série à termes positifs, la suite de ses sommes partielles tend donc vers $+\infty$: $\lim_{n \rightarrow +\infty} E(H_n) = +\infty$.

Les B_k étant indépendantes, on a donc

$$V(H_n) = \sum_{k=1}^n V\left(\frac{B_k}{k}\right) = \sum_{k=1}^n \frac{V(B_k)}{k^2} = \sum_{k=1}^n \frac{1}{4k^2} = \frac{1}{4} \sum_{k=1}^n \frac{1}{k^2} \xrightarrow{n \rightarrow +\infty} \frac{1}{4} \zeta(2) = \frac{\pi^2}{24}.$$

⁶ Ce qui est équivalent.

CV en proba/loi

Le résultat admis dans l'énoncé n'est pas spécifique à ce cas particulier, on a toujours

$$X_n \xrightarrow{P} X \Rightarrow X_n \xrightarrow{\mathcal{L}} X.$$

En revanche la réciproque est fautive.

⚠ Attention !

Les sommes partielles d'une série divergente ne tendent pas toujours vers $+\infty$ (par exemple c'est faux pour la série de terme général $(-1)^n$), c'est vrai ici car il s'agit d'une série à termes positifs.

- 11.b. Soit $r > 0$ fixé. Alors, pour n suffisamment grand, on a $E(H_n) > r$.
 Et alors $[H_n \leq r] \subset [|H_n - E(H_n)| \geq E(H_n) - r]$.
 Par l'inégalité de Bienaymé-Tchebychev appliquée⁷ à H_n , on a alors

⁷ Ce qui est légitime puisque H_n admet une variance.

$$P(H_n \leq r) \leq P(|H_n - E(H_n)| \geq E(H_n) - r) \leq \frac{V(H_n)}{(E(H_n) - r)^2}.$$

Mais puisque $V(H_n)$ admet une limite finie, elle est bornée, et puisque $\frac{1}{(E(H_n) - r)^2} \xrightarrow{n \rightarrow +\infty} 0$,
 on en déduit que $\lim_{n \rightarrow +\infty} \frac{V(H_n)}{(E(H_n) - r)^2} = 0$.

Et donc par le théorème des gendarmes, $\boxed{\lim_{n \rightarrow +\infty} P(H_n \leq r) = 0}$.

Remarque : ceci prouve que (H_n) ne converge pas en loi car sinon $P(H_n \leq r)$ devrait converger vers la fonction de répartition d'une variable aléatoire. Or la fonction nulle n'est pas une fonction de répartition.

Notons que le raisonnement se généralise aisément à toute suite (X_n) de variables aléatoires telle que $E(X_n) \xrightarrow{n \rightarrow +\infty} +\infty$ et $V(X_n)$ est bornée.

- 11.c.i. Le code déjà présent permet de simuler H_n , et donc il faut ajouter $-h_n(\mathcal{I})$.
 Autrement dit, il faut être capable de calculer $\mathbb{1}_{\mathcal{I}(k)}$.
 Pour cela, nous proposons d'utiliser le fait suivant : $1 - (-1)^k = 2$ si k est impair et 0 sinon.
 Autrement dit, $\mathbb{1}_{\mathcal{I}(k)} = \frac{1 - (-1)^k}{2}$.
 Et donc on peut compléter le programme de la manière suivante :
- 5 `y(i, 1) = y(i, 1) + (grand(1, 1, 'bin', 1, 0.5) + ((-1)^k - 1)/2)/k ;`

- 11.c.ii. Les graphiques ont par exemple pu être obtenus avec `histplot(10, simul(n, 10000))`, qui trace un histogramme à 10 bâtons à partir de 10 000 simulations de $H_n - h_n(\mathcal{I})$.
- 11.c.iii. Le fait que ces trois histogrammes soient très similaires semble indiquer que la répartition des valeurs prises par $H_n - h_n(\mathcal{I})$ reste sensiblement la même pour de grandes valeurs de n . Et donc que les fonctions de répartition des $H_n - h_n(\mathcal{I})$ ne semblent que peu dépendre de la valeur de n , ce qui indique une convergence en loi.

- 12.a. Utilisons, comme précédemment le fait que $\mathbb{1}_{\mathcal{I}(k)} = \frac{1 - (-1)^k}{2}$. Ceci permet d'écrire

$$h_n(\mathcal{I}) = \sum_{k=1}^n \frac{\mathbb{1}_{\mathcal{I}(k)}}{k} = \sum_{k=1}^n \frac{1 - (-1)^k}{2k} = \frac{1}{2} \sum_{k=1}^n \frac{1}{k} - \frac{1}{2} \sum_{k=1}^n \frac{(-1)^k}{k}.$$

Et donc

$$\sum_{k=1}^n \frac{B'_k}{k} - h_n(\mathcal{I}) = \sum_{k=1}^n \frac{1}{2k} + \sum_{k=1}^n \frac{X_k}{k} - \frac{1}{2} \sum_{k=1}^n \frac{1}{k} + \sum_{k=1}^n \frac{(-1)^k}{k} = \boxed{\frac{K_n}{2} + \sum_{k=1}^n \frac{(-1)^k}{2k}}.$$

- 12.b. Nous savons⁸ que $K_n \xrightarrow{\mathcal{L}} K$, et $\sum_{k=1}^n \frac{(-1)^k}{2k} \xrightarrow{n \rightarrow +\infty} \frac{1}{2} \sum_{k=1}^{+\infty} \frac{(-1)^k}{k}$, qui est bien une série convergente d'après 7.b.

⁸ C'est le résultat de 10.c.

Mais en voyant cette suite de réels comme une suite de variables aléatoires certaines, on a

$$\sum_{k=1}^n \frac{(-1)^k}{2k} \xrightarrow{P} \frac{1}{2} \sum_{k=1}^{+\infty} \frac{(-1)^k}{k}.$$

Et donc par le lemme de Slutsky,

$$\sum_{k=1}^n \frac{B'_k}{k} - h_n(\mathcal{I}) \xrightarrow{\mathcal{L}} \frac{K}{2} + \frac{1}{2} \sum_{k=1}^{+\infty} \frac{(-1)^k}{k}.$$

Mais pour tout $i \in \mathbf{N}^*$, B'_i a même loi que X_i , et les B'_i sont indépendantes car les X_i le sont. Par conséquent, pour tout $n \in \mathbf{N}^*$, les vecteurs aléatoires (X_1, \dots, X_n) et (B'_1, \dots, B'_n) ont même loi.

Si l'on applique alors la fonction continue $(x_1, \dots, x_n) \mapsto \sum_{k=1}^n \frac{x_k}{k}$, on obtient alors que H_n

et $\sum_{k=1}^n \frac{B'_k}{k}$ ont même loi.

Et donc $H_n - h_n(\mathcal{I})$ et $\sum_{k=1}^n \frac{B'_k}{k} - h_n(\mathcal{I})$ ont même loi.

Par conséquent ,

$$H_n - h_n(\mathcal{I}) \xrightarrow{\mathcal{L}} \frac{K}{2} + \frac{1}{2} \sum_{k=1}^{+\infty} \frac{(-1)^k}{k}.$$

Remarque : il n'est en fait pas trop difficile de prouver que $\sum_{k=1}^{+\infty} \frac{(-1)^k}{k} = -\ln(2)$.

Ceci est à mettre en parallèle avec le développement limité de $\ln(1+x)$:

$$\ln(1+x) = \sum_{k=1}^n \frac{(-1)^{k-1}}{k} x^k + o(x^n).$$

Ce développement limité s'obtient à l'aide de la formule de Taylor-Young appliquée à $x \mapsto \ln(1+x)$. Si l'on applique la formule de Taylor avec reste intégral à l'ordre n à la même fonction entre 0 et 1, puis que l'on fait tendre n vers $+\infty$, on constate que le reste intégral tend vers 0 et donc que

$$\sum_{k=1}^{+\infty} \frac{(-1)^k}{k} = -\lim_{n \rightarrow +\infty} \sum_{k=1}^n \frac{(-1)^{k-1}}{k} = -\ln(1+1) = -\ln(2).$$

Détails

La convergence en loi ne regarde que les lois de variables en jeu, donc si pour tout n , X_n a même loi que Y_n , et que $X_n \xrightarrow{\mathcal{L}} X$, alors $Y_n \xrightarrow{\mathcal{L}} X$.

DEVOIR SURVEILLÉ 8 (23.03.19)

Si vous souhaitez que votre copie ne soit pas notée (mais bien entendu elle sera corrigée), inscrivez-le clairement sur la première page.

PROBLÈME 1

On considère l'application $\varphi : [0; +\infty[\rightarrow \mathbf{R}$ définie, pour tout réel $t \in [0; +\infty[$, par :

$$\varphi(t) = \begin{cases} \frac{\sin t}{t} & \text{si } t \neq 0 \\ 1 & \text{si } t = 0 \end{cases}$$

et on considère, pour tout entier $n \geq 1$, les intégrales :

$$I_n = \int_0^{+\infty} (\varphi(t))^n dt, \quad J_n = \int_0^1 (\varphi(t))^n dt, \quad K_n = \int_1^{+\infty} (\varphi(t))^n dt.$$

Partie I : Résultats généraux sur φ et J_n

1. Montrer que φ est continue sur $[0; +\infty[$ et que, pour tout entier $n \geq 1$, l'intégrale J_n existe.
2.
 - a. Montrer que φ est strictement positive sur $[0; 1]$ et que φ est strictement décroissante sur $[0; 1]$.
 - b. Établir, pour tout réel $t \in]0; +\infty[$: $|\varphi(t)| < 1$.
3.
 - a. Montrer, pour tout réel $t \in [0; +\infty[$: $\varphi(t) \geq 1 - t$.
(On pourra étudier les variations sur $[0; +\infty[$ de l'application $t \mapsto \sin t - t + t^2$).
 - b. En déduire, pour tout entier $n \geq 1$: $J_n \geq \frac{1}{n+1}$.

Partie II : Étude de I_1 .

4.
 - a. Montrer, pour tout réel $x \in [1; +\infty[$: $\int_1^x \frac{\sin t}{t} dt = \cos 1 - \frac{\cos x}{x} - \int_1^x \frac{\cos t}{t^2} dt$.
 - b. En déduire que les intégrales K_1 et I_1 sont convergentes.
5.
 - a. Montrer, pour tout réel $t \in [0; +\infty[$: $|\sin t| \geq \sin^2 t = \frac{1}{2}(1 - \cos(2t))$.
 - b. Montrer que l'intégrale $\int_1^{+\infty} \frac{\cos(2t)}{2t} dt$ converge.
 - c. Déduire des deux questions précédentes que l'intégrale I_1 n'est pas absolument convergente.

Partie III : Étude de I_n , pour $n \geq 2$.

6.
 - a. Montrer que, pour tout entier $n \geq 2$, l'intégrale K_n est convergente.
 - b. Établir, pour tout entier $n \geq 2$: $|K_n| \leq \frac{1}{n-1}$.
7.
 - a. Montrer que la suite $(J_n)_{n \geq 2}$ est décroissante.
 - b. Montrer que la suite $(J_n)_{n \geq 2}$ converge ; on note ℓ sa limite.
 - c. Établir, pour tout entier $n \geq 2$ et tout réel $a \in]0; 1[$:

$$\int_0^a (\varphi(t))^n dt \leq a \text{ et } \int_a^1 (\varphi(t))^n dt \leq (1-a)(\varphi(a))^n.$$

(On pourra utiliser la question 2.)

8.
 - d. En déduire, pour tout réel $a \in]0; 1[$: $0 \leq \ell \leq a$ et conclure : $\ell = 0$.
 - a. Montrer que, pour tout entier $n \geq 2$, l'intégrale I_n est convergente.
 - b. Établir : $\lim_{n \rightarrow +\infty} I_n = 0$.

Partie IV : Étude de la série de terme général $I_n, n \geq 2$.

9. Montrer, pour tout entier $p \geq 1$: $K_{2p} + K_{2p+1} \geq 0$.
10. En déduire, pour tout entier $N \geq 1$:

$$\sum_{p=1}^N (I_{2p} + I_{2p+1}) \geq \sum_{p=1}^N (J_{2p} + J_{2p+1}).$$

11. En déduire que la série $\sum_{n \geq 2} I_n$ diverge. (On pourra utiliser la question 3.b).

PROBLÈME 2

L'objectif du problème est d'étudier une suite de variables aléatoires $(Z_k)_{k \in \mathbf{N}^*}$.

Les deux premières parties sont indépendantes et la troisième utilise certains résultats obtenus dans les deux premières parties.

La partie I est consacrée à l'étude de deux endomorphismes sur $\mathbf{R}_n[X]$.

La partie II consiste à calculer l'espérance et la variance de Z_k ainsi qu'à calculer la somme $\sum_{k=0}^{+\infty} P(Z_k = r)$ sous réserve de convergence.

La partie III fournira la loi de Z_k ainsi que l'étude de la convergence de la série $\sum_{k \geq 0} P(Z_k = r)$.

Partie I : Étude de deux endomorphismes.

Soit n un entier naturel. On note $\mathbf{R}_n[X]$ l'ensemble des polynômes à coefficients réels de degré au plus n .

Pour tout entier $k \in \{0, 1, \dots, n\}$, on désigne par e_k le polynôme de $\mathbf{R}_n[X]$ défini par : $e_k = X^k$.

Rappelons que (e_0, e_1, \dots, e_n) est une base de $\mathbf{R}_n[X]$. Si $P \in \mathbf{R}_n[X]$, on définit les fonctions $f(P)$ et $g(P)$ par :

$$\forall x \in \mathbf{R} \setminus \{1\}, f(P)(x) = \frac{1}{x-1} \int_1^x P(t) dt \text{ et } f(P)(1) = P(1)$$

$$\forall x \in \mathbf{R}, g(P)(x) = [(X-1)P]'(x) = (x-1)P'(x) + P(x).$$

1. Prouver que g est un endomorphisme de $\mathbf{R}_n[X]$.
2. Soit $P \in \mathbf{R}_n[X]$. Calculer $f(g(P))$ puis justifier que $\text{Ker}(g) = \{0\}$.
3. Démontrer que g est un isomorphisme, que $g^{-1} = f$ et que f est un endomorphisme de $\mathbf{R}_n[X]$.
4. Écrire la matrice A de f dans la base (e_0, e_1, \dots, e_n) ainsi que la matrice B de g dans cette même base.
5. Montrer que f et g sont diagonalisables.

Partie II : Étude d'une suite de variables aléatoires

Soit n un entier naturel supérieur ou égal à 1. On dispose de $n+1$ urnes notées U_0, U_1, \dots, U_n et on suppose que $\forall i \in \{0, \dots, n\}$, l'urne U_i contient $i+1$ boules numérotées $0, 1, \dots, i$. On s'intéresse au jeu suivant :

- Au premier tirage, on pioche une boule dans l'urne U_n . Si la boule porte le numéro r alors on repose la boule dans l'urne U_n puis le tirage suivant s'effectue dans l'urne U_r .
- Plus généralement, pour tout entier naturel k non nul, si la boule numéro s a été piochée au k -ième tirage dans une certaine urne, on repose cette boule dans la même urne puis on effectue le $(k+1)$ -ième tirage dans l'urne U_s .

Pour tout entier naturel k , on note :

- Z_k est la variable aléatoire égale au numéro de la boule piochée au k -ième tirage.
On convient que $Z_0 = n$.

- F_k est le polynôme de $\mathbf{R}_n[X]$ défini par : $\forall x \in \mathbf{R}, F_k(x) = \sum_{r=0}^n P(Z_k = r)x^r$.

- $E(Z_k)$ désigne l'espérance de la variable aléatoire Z_k .

6. À l'aide de la formule des probabilités totales, prouver que :

$$\forall k \in \mathbf{N}, \quad \forall r \in \{0, 1, \dots, n\}, \quad P(Z_{k+1} = r) = \sum_{i=r}^n \frac{P(Z_k = i)}{i+1}.$$

7. Établir les deux formules suivantes valables pour tous entiers $k \in \mathbf{N}$ et $r \in \{0, 1, \dots, n-1\}$:

$$\begin{cases} (\mathcal{R}_1) : & (n+1)P(Z_{k+1} = n) = P(Z_k = n) \\ (\mathcal{R}_2) : & (r+1)P(Z_{k+1} = r) - (r+1)P(Z_{k+1} = r+1) = P(Z_k = r) \end{cases}$$

8. On admet dans cette question que la série $\sum_{k \geq 0} P(Z_k = r)$ converge pour tout $r \in \{1, \dots, n\}$ et on pose $S_r = \sum_{k=0}^{+\infty} P(Z_k = r)$.

En sommant les relations (\mathcal{R}_1) sur tous les entiers $k \in \mathbf{N}$, donner la valeur de S_n .

En sommant les relations (\mathcal{R}_2) sur tous les entiers $k \in \mathbf{N}$, donner la valeur de S_{n-1} et montrer que la suite $(rS_r)_{1 \leq r \leq n-1}$ est constante.

9. Soit $k \in \mathbf{N}$. Démontrer la relation :

$$(\mathcal{S}) : \forall x \in \mathbf{R}, \quad (x-1)F'_{k+1}(x) + F_{k+1}(x) = F_k(x).$$

- 10. a. Soit $k \in \mathbf{N}$. Établir que $F'_k(1) = E(Z_k)$ et $F''_k(1) = E(Z_k(Z_k - 1))$.
- b. En dérivant une fois puis deux fois la relation (\mathcal{S}) , donner la relation de récurrence vérifiée par la suite $(F'_k(1))_{k \in \mathbf{N}}$ ainsi que la relation de récurrence vérifiée par la suite $(F''_k(1))_{k \in \mathbf{N}}$.
- c. Donner la valeur de $F'_k(1)$ et de $F''_k(1)$ en fonction de k et n . Expliciter alors la variance $V(Z_k)$ de Z_k en fonction de k et n .

Partie III : Loi de chacune de ces variables aléatoires.

On reprend toutes les notations des parties I et II et on pourra admettre tous les résultats établis dans ces deux parties. Rappelons également qu'à la question 8 la relation (\mathcal{S}) est démontrée ce qui revient à écrire :

$$\forall k \in \mathbf{N}, \quad g(F_{k+1}) = F_k.$$

Pour finir, pour tout entier $k \in \{0, 1, \dots, n\}$ on désigne par u_k le polynôme de $\mathbf{R}_n[X]$ défini par : $u_k = (X-1)^k$.

- 11. Montrer que : $\forall k \in \mathbf{N}, \quad \sum_{r=0}^n P(Z_k = r)e_r = F_k = f^k(e_n)$.
- 12. Prouver que (u_0, u_1, \dots, u_n) est une base de $\mathbf{R}_n[X]$.
- 13. Calculer $f(u_r)$ pour $r \in \{0, 1, \dots, n\}$. Retrouver ainsi que f est diagonalisable.
- 14. Justifier que : $e_n = \sum_{r=0}^n \binom{n}{r} u_r$ et que : $\forall r \in \{0, 1, \dots, n\}, \quad u_r = \sum_{j=0}^r (-1)^{r-j} \binom{r}{j} e_j$.
- 15. Démontrer que : $\forall k \in \mathbf{N}, \quad f^k(e_n) = \sum_{r=0}^n \frac{\binom{n}{r}}{(r+1)^k} u_r$.
- 16. Soient $k \in \mathbf{N}$ et $j \in \{0, 1, \dots, n\}$. À l'aide des questions précédentes, établir que :

$$P(Z_k = j) = \sum_{r=j}^n (-1)^{r-j} \frac{\binom{n}{r} \binom{r}{j}}{(r+1)^k}.$$

17. Application.

a. Soit $j \in \{0, 1, \dots, n\}$. Déterminer un réel $M_{j,n}$ tel que :

$$\forall k \in \mathbf{N}, \quad |P(Z_k = j)| \leq \frac{M_{j,n}}{(j+1)^k}$$

puis justifier que la série $\sum_{k \geq 0} P(Z_k = j)$ converge lorsque $j \in \{1, 2, \dots, n\}$.

b. Déterminer un réel C_n tel que : $\forall k \in \mathbf{N}, \quad |P(Z_k = 0) - 1| \leq \frac{C_n}{2^k}$.

La série $\sum_{k \geq 0} P(Z_k = 0)$ est-elle convergente ?

CORRECTION DU DEVOIR SURVEILLÉ 8

PROBLÈME 1 (EML 2003)

Partie I : Résultats généraux sur φ et J_n .

1. La fonction φ est continue sur $]0, +\infty[$ car quotient de fonctions continues.

D'autre part, au voisinage de 0, $\varphi(t) \underset{t \rightarrow 0}{\sim} \frac{t}{t} = 1 \underset{t \rightarrow 0^+}{=} \varphi(0)$.

Et donc φ est continue en 0 et donc sur $[0; +\infty[$.

Par conséquent, pour tout entier $n \geq 1$, φ^n est continue et donc J_n est l'intégrale d'une fonction continue sur un segment, et donc converge.

- 2.a. Nous savons que la fonction sinus est strictement positive sur $]0, \pi[$ et que $1 < \pi$, donc pour tout $t \in]0, 1]$, $\sin(t) > 0$, de sorte que $\varphi(t) > 0$.

Et puisque $\varphi(0) > 0$, φ est strictement positive sur $[0, 1]$.

D'autre part, φ est dérivable sur $]0, 1]$ car quotient de fonctions dérivables, et

$$\forall t \in]0, 1], \varphi'(t) = \frac{t \cos t - \sin t}{t^2}.$$

Soit donc g la fonction définie sur $[0, 1]$ par $g(t) = t \cos t - \sin t$. Alors g est dérivable et

$$g'(t) = \cos t - t \sin t - \cos t = -t \sin t < 0.$$

Donc g est strictement décroissante sur $[0, 1]$, avec $g(0) = 0$.

On en déduit donc que pour tout $t \in]0, 1]$, $g(t) < 0$ et donc $\varphi'(t) = \frac{g(t)}{t^2} < 0$.

Par conséquent, φ est strictement décroissante sur $[0, 1]$.

- 2.b. Soit $t > 1$. Alors $|\sin t| \leq 1$ et donc $|\varphi(t)| = \left| \frac{\sin t}{t} \right| \leq \frac{1}{t} < 1$.

D'autre part, $\varphi(0) = 1$ et donc, par stricte décroissance de φ sur $[0, 1]$, pour tout $t \in]0, 1]$, $\varphi(t) < 1$.

Puisqu'enfin φ est positive sur $[0, 1]$, donc pour tout $t \in]0, 1]$, $|\varphi(t)| < 1$.

- 3.a. Pour $t = 0$, il est évident que $\varphi(0) = 1 \geq 1 - 1$.

Pour $t > 0$, on a $\varphi(t) - (1 - t) = \frac{\sin t - t + t^2}{t}$.

Soit donc $h : t \mapsto \sin t - t + t^2$.

Alors h est deux fois dérivable sur $[0; +\infty[$, et $h'(t) = \cos t - 1 + 2t$.

Donc $h''(t) = -\sin t + 2 > 0$, de sorte que h' est croissante.

Or, $h'(0) = 0$, donc pour tout $t \geq 0$, $h'(t) \geq 0$.

Ceci signifie donc que h est croissante, avec $h(0) = 0$, donc pour tout $t \geq 0$, $h(t) \geq 0$.

Et donc pour tout $t \in]0, +\infty[$, $\varphi(t) - 1 + t \geq 0 \Leftrightarrow$ $\varphi(t) \geq 1 - t$.

- 3.b. Pour tout $t \in [0, 1]$, et pour tout $n \geq 1$, on a $\varphi(t)^n \geq (1 - t)^n$.

Et donc, par croissance de l'intégrale,

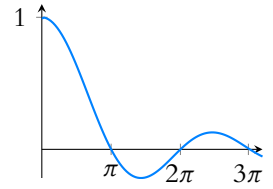
$$J_n = \int_0^1 (\varphi(t))^n dt \geq \int_0^1 (1 - t)^n dt = \left[-\frac{(1 - t)^{n+1}}{n + 1} \right]_0^1 = \frac{1}{n + 1}.$$

Partie II : Étude de I_1 .

- 4.a. Procédons à une intégration par parties sur le segment $[1, x]$, en posant $u(t) = \frac{1}{t}$, $v(t) = -\cos(t)$,

qui sont deux fonctions de classe \mathcal{C}^1 avec $u'(t) = -\frac{1}{t^2}$ et $v'(t) = \sin(t)$. Alors

$$\int_1^x \frac{\sin t}{t} dt = \left[-\frac{\cos t}{t} \right]_1^x - \int_1^x \frac{\cos t}{t^2} dt = \cos 1 - \frac{\cos x}{x} - \int_1^x \frac{\cos t}{t^2} dt.$$

FIGURE 24.1- La fonction φ .

Rappel

Si φ est continue sur $[0, 1]$ et que φ' est strictement négative sur $]0, 1]$, alors φ est strictement décroissante sur $[0, 1]$.

Il n'est pas utile de se préoccuper de la dérivabilité de φ en 0, ni du signe éventuel de $\varphi'(1)$.

Remarque

On pourrait aussi remarquer que h est convexe (car de dérivée seconde positive) et donc au dessus de toutes ses tangentes.

Mais en particulier, sa tangente en 0 est la droite horizontale d'équation $y = 0$, et donc pour tout t , $h(t) \geq 0$.

4.b. Puisque la fonction \cos est bornée sur \mathbf{R} , lorsque $x \rightarrow +\infty$, $\frac{\cos x}{x} \xrightarrow{x \rightarrow +\infty} 0$.

D'autre part, pour tout $t \in \mathbf{R}_+^*$, $\left| \frac{\cos t}{t^2} \right| \leq \frac{1}{t^2}$.

Donc par comparaison à une intégrale de Riemann convergente, $\int_1^{+\infty} \left| \frac{\cos t}{t} \right| dt$ converge,

et donc $\int_1^{+\infty} \frac{\cos t}{t^2} dt$ converge absolument, donc converge.

Par définition d'une intégrale convergente, cela signifie que $\lim_{x \rightarrow +\infty} \int_1^x \frac{\cos t}{t^2} dt$ existe.

Et donc $\lim_{x \rightarrow +\infty} \int_1^x \frac{\sin t}{t} dt$ existe, de sorte que $\int_1^{+\infty} \frac{\sin t}{t} dt = K_1$ converge.

Et puisque J_1 converge également, on en déduit que $I_1 = J_1 + K_1$ converge.

5.a. Soit $t \geq 0$. Puisque $|\sin t| \leq 1$, on a donc $|\sin t|^2 \leq |\sin t|$, soit encore $\sin^2 t \leq |\sin t|$.

D'autre part, nous savons que

$$\cos(2t) = \cos(t+t) = \cos^2 t - \sin^2 t = (1 - \sin^2 t) - \sin^2 t = 1 - 2\sin^2 t.$$

Et donc $\sin^2 t = \frac{1 - \cos(2t)}{2}$.

5.b. De même qu'à la question 4.a, une intégration par parties sur $[0, x]$ nous donne

$$\int_1^x \frac{\cos(2t)}{t} dt = \left[\frac{\sin(2t)}{2t} \right]_1^x + \int_1^x \frac{\sin(2t)}{2t^2} dt = \frac{\sin(2x)}{2x} - \frac{\sin(2)}{2} + \int_1^x \frac{\sin(2t)}{2t^2} dt.$$

On montre alors comme en 4.a que $\int_1^{+\infty} \frac{\sin(2t)}{2t^2} dt$ converge absolument, et donc que

$\int_1^x \frac{\cos(2t)}{t} dt$ admet une limite finie lorsque $x \rightarrow +\infty$ et donc que $\int_1^{+\infty} \frac{\cos(2t)}{t} dt$ converge.

5.c. En utilisant la question 5.a., il vient, pour tout $t \geq 1$,

$$\left| \frac{\sin t}{t} \right| \geq \frac{1}{2} \left(\frac{1}{t} - \frac{\cos(2t)}{t} \right).$$

Mais $\int_1^{+\infty} \frac{dt}{t}$ diverge et $\int_1^{+\infty} \frac{\cos(2t)}{t} dt$ converge, de sorte que $\int_1^{+\infty} \frac{1}{2} (1 - \cos(2t)) dt$ diverge.

Et donc par critère de comparaison pour les intégrales de fonctions positives, $\int_1^{+\infty} \left| \frac{\sin t}{t} \right| dt$ diverge.

Et donc I_1 ne converge pas absolument.

Partie III : Étude de I_n , pour $n \geq 2$.

6.a. Nous avons déjà prouvé que K_1 converge.

Et pour $n \geq 2$, on a pour tout $t > 0$, $|\varphi(t)^n| \leq \frac{1}{t^n}$.

Puisque $\int_1^{+\infty} \frac{dt}{t^n}$ converge, par critère de comparaison pour les intégrales de fonctions positives, il en est de même de $\int_1^{+\infty} |\varphi(t)^n| dt$, et donc K_n converge absolument. Et en particulier converge.

Ainsi, pour tout $n \in \mathbf{N}^*$, K_n converge.

6.b. D'après l'inégalité triangulaire, on a

$$\begin{aligned} |K_n| &= \left| \int_1^{+\infty} \varphi(t)^n dt \right| \\ &\leq \int_1^{+\infty} |\varphi(t)^n| dt \end{aligned}$$

Trigonométrie

Rappelons que

$$\begin{aligned} \cos(a+b) &= \\ \cos(a)\cos(b) - \sin(a)\sin(b). \end{aligned}$$

Rappel

La somme d'une intégrale convergente et d'une intégrale divergente diverge. En revanche, on ne peut rien dire de la somme de deux intégrales divergentes.

Remarque

Nous avons donc ici un exemple d'intégrale convergente, mais pas absolument convergente.

$$\leq \int_1^{+\infty} \frac{dt}{t^n}$$

$$= \frac{1}{n-1}.$$

- 7.a. Pour tout $t \in [0, 1]$, $0 \leq \varphi(t) \leq 1$, de sorte que $\varphi(t)^{n+1} \leq \varphi(t)^n$.
Et donc par croissance de l'intégrale,

$$J_{n+1} = \int_0^1 \varphi(t)^{n+1} dt \leq \int_0^1 \varphi(t)^n dt = J_n.$$

Ceci prouve donc que (J_n) est décroissante.

- 7.b. Puisque $\varphi(t)^n \geq 0$ sur $[0, 1]$, par positivité de l'intégrale, $J_n \geq 0$.
Et donc (J_n) est décroissante et minorée¹, donc par le théorème de la limite monotone, elle converge.
- 7.c. Il a été prouvé à la question 2.a. que φ est décroissante sur $[0, 1]$, avec $\varphi(0) = 1$ et $\varphi(1) \geq 0$.
Donc pour tout $t \in [0, 1]$, $0 \leq \varphi(t) \leq 1$ et donc $0 \leq \varphi(t)^n \leq 1$.
Par croissance de l'intégrale, on a donc $\int_0^a (\varphi(t))^n dt \leq \int_0^a 1 dt = a$.
De même, φ est décroissante sur $[a, 1]$ et donc pour tout $t \in [a, 1]$, $\varphi(t) \geq \varphi(a) \geq 0$ et donc $\varphi(t)^n \geq \varphi(a)^n$.
Donc par croissance de l'intégrale,

$$\int_a^1 \varphi(t)^n dt \geq \int_a^1 \varphi(a)^n dt = (1-a)\varphi(a)^n.$$

- 7.d. Pour $a \in]0, 1[$ fixé, et pour tout $n \in \mathbf{N}$, nous venons de prouver que

$$0 \leq J_n \leq \int_0^a \varphi(t)^n dt + \int_a^1 \varphi(t)^n dt \leq a^n + (1-a)\varphi(a)^n.$$

Mais $\varphi(a) < 1$ par stricte décroissance de φ sur $[0, 1]$, et donc $\varphi(a)^n \xrightarrow{n \rightarrow +\infty} 0$.

Ainsi, lorsque $n \rightarrow +\infty$, on a

$$0 \leq \ell \leq a.$$

Et ceci est vrai pour tout $a \in]0, 1[$, donc en faisant tendre a vers 0, il vient donc² $\ell = 0$.

- 8.a. Nous avons déjà prouvé que pour tout n , K_n est convergente, et J_n est l'intégrale d'une fonction continue sur un segment, donc convergente.
On en déduit que $I_n = K_n + J_n$ est convergente.
- 8.b. La majoration de la question 6.b prouve, à l'aide du théorème des gendarmes, que

$$\lim_{n \rightarrow +\infty} |K_n| = 0 \Leftrightarrow \lim_{n \rightarrow +\infty} K_n = 0.$$

Et d'autre part, nous venons de prouver que $\lim_{n \rightarrow +\infty} J_n = 0$.

Donc par somme de limites,

$$\lim_{n \rightarrow +\infty} I_n = \lim_{n \rightarrow +\infty} J_n + \lim_{n \rightarrow +\infty} K_n = 0.$$

Partie V : Étude de la série de terme général I_n , $n \geq 2$.

9. On a

$$K_{2p} + K_{2p+1} = \int_1^{+\infty} (\varphi(t)^{2p} + \varphi(t)^{2p+1}) dt = \int_1^{+\infty} \varphi(t)^{2p} (1 + \varphi(t)) dt.$$

Or il est évident que $\varphi(t)^{2p} \geq 0$ pour tout $t \geq 1$, et puisque nous avons prouvé à la question 2.b que $|\varphi(t)| \leq 1$, il vient

$$-1 \leq \varphi(t) \leq 1 \Leftrightarrow 0 \leq \varphi(t) + 1 \leq 2.$$

Et donc en particulier, $\varphi(t)^{2p} (1 + \varphi(t)) \geq 0$.

Par croissance de l'intégrale, on a donc

$$K_{2p} + K_{2p+1} = \int_1^{+\infty} \varphi(t)^{2p} (\varphi(t) + 1) dt \geq 0.$$

Intégrales de Riemann

Nous avons utilisé ici le fait que pour $\alpha > 1$,

$$\int_1^{+\infty} \frac{dt}{t^\alpha} = \frac{1}{\alpha-1}.$$

Si on ne se souvient pas de cette formule, il est aisé de la retrouver à l'aide d'une primitive de $\frac{1}{t^\alpha}$.

¹ Par 0.

² Par le théorème des gendarmes.

10. En sommant les relations de la question précédente pour p allant de 1 à N , il vient

$$\sum_{p=1}^N (K_{2p} + K_{2p+1}) \geq 0.$$

Mais d'autre part, $K_n = I_n - J_n$, donc

$$\sum_{p=1}^N (K_{2p} + K_{2p+1}) \geq 0 \Leftrightarrow \sum_{p=1}^N (I_{2p} + I_{2p+1} - J_{2p} - J_{2p+1}) \geq 0 \Leftrightarrow \sum_{p=1}^N (I_{2p} + I_{2p+1}) \geq \sum_{p=1}^N (J_{2p} + J_{2p+1}).$$

11. On a $\sum_{p=1}^N (J_{2p} + J_{2p+1}) = \sum_{k=2}^{2N+1} J_k$, et de même $\sum_{p=1}^N (I_{2p} + I_{2p+1}) = \sum_{p=1}^{2N+1} I_k$.

Or, à la question 3.b, nous avons prouvé que $J_k \geq \frac{1}{k+1}$.

Puisque la série de terme général $\frac{1}{k+1}$ diverge, et qu'il s'agit d'une série à termes positifs,

$$\lim_{N \rightarrow +\infty} \sum_{k=2}^{2N+1} \frac{1}{k+1} = +\infty.$$

Et donc $\lim_{N \rightarrow +\infty} \sum_{k=2}^{2N+1} I_k = +\infty$.

On en déduit que la série de terme général I_n diverge.

Positivité

La suite des sommes partielles de $\sum \frac{1}{k+1}$ tend vers $+\infty$ car il s'agit d'une série à termes positifs. En effet, la suite de ses sommes partielles est alors croissante, et ne convergeant pas, elle tend vers $+\infty$.
 Pour des séries divergentes qui ne sont pas de signe constant, il n'est pas obligatoire que la suite des sommes partielles tende vers $+\infty$. Par exemple, les sommes partielles de la série (grossièrement) divergente de terme général $(-1)^n$ sont bornées.

PROBLÈME 2 (ÉCRICOME 2012)

Partie I : Étude de deux endomorphismes.

1. Pour $P \in \mathbf{R}_n[X]$, il est clair que $g(P)$ est encore un polynôme de $\mathbf{R}[X]$, et $\deg P' \leq n - 1$, de sorte que $\deg(X - 1)P' \leq n$ et donc $\deg g(P) \leq n$.

Ainsi, $g(P) \in \mathbf{R}_n[X]$.

Il reste donc à montrer que g est linéaire : soient $P, Q \in \mathbf{R}_n[X]$ et $\lambda \in \mathbf{R}$. Alors

$$g(\lambda P + Q) = (X - 1)(\lambda P + Q)' + (\lambda P + Q) = \lambda(X - 1)P' + \lambda P + (X - 1)Q' + Q = \lambda g(P) + g(Q).$$

Ainsi, g est linéaire, et c'est donc un endomorphisme de $\mathbf{R}_n[X]$.

2. Si $x = 1$, alors $f(g(P))(1) = g(P)(1) = (1 - 1)P'(1) + P(1) = P(1)$.
 Si $x \neq 1$, alors

$$f(g(P))(x) = \frac{1}{x-1} \int_1^x ((t-1)P'(t) + P(t)) dt = \frac{1}{x-1} [(t-1)P(t)]_1^x = \frac{1}{x-1} (x-1)P(x) = P(x).$$

Ainsi, pour tout $x \in \mathbf{R}$, $f(g(P))(x) = P(x)$, et donc $(f \circ g)(P) = P$.

En particulier, si $g(P) = 0$, alors $P = f(g(P)) = 0$, de sorte que $\text{Ker } g = \{0_{\mathbf{R}_n[X]}\}$.

3. La question précédente prouve que g est injectif. Puisque c'est un endomorphisme de $\mathbf{R}_n[X]$, qui est de dimension finie, c'est un isomorphisme.

Nous venons de prouver que $f \circ g = \text{id}_{\mathbf{R}_n[X]}$. Donc pour $P \in \mathbf{R}_n[X]$,

$$f(P) = (f \circ g)(g^{-1}(P)) = g^{-1}(P) \in \mathbf{R}_n[X].$$

Ceci prouve donc que f est à valeurs dans $\mathbf{R}_n[X]$, et que $g^{-1} = f$.

Et en particulier, f est un endomorphisme³ de $\mathbf{R}_n[X]$.

4. Pour $k \in \llbracket 0, n \rrbracket$, et $x \neq 1$, on a

$$\begin{aligned} f(e_k)(x) &= \frac{1}{x-1} \int_1^x t^k dt = \frac{1}{x-1} \left[\frac{t^{k+1}}{k+1} \right]_1^x \\ &= \frac{1}{k+1} \frac{x^{k+1} - 1}{x-1} = \frac{1}{k+1} (1 + x + x^2 + \dots + x^k) \\ &= \frac{1}{k+1} (e_0(x) + \dots + e_n(x)). \end{aligned}$$

Remarque

Sans utiliser g et g^{-1} , il serait facile de prouver que f est linéaire, mais délicat de prouver que f est à valeurs dans $\mathbf{R}_n[X]$.

³ Et même un isomorphisme.

Factorisation

Rappelons que $a^n - b^n$ se factorise en

$$(a - b) \times \sum_{k=0}^{n-1} a^k b^{n-1-k}$$

Or, deux polynômes qui coïncident en une infinité de réels sont égaux, donc

$$f(e_k) = \frac{1}{k+1} (e_0 + e_1 + \dots + e_n).$$

Et donc la matrice de f dans la base (e_0, e_1, \dots, e_n) est

$$A = \begin{pmatrix} f(e_0) & f(e_1) & \dots & f(e_{n-1}) & f(e_n) \\ 1 & \frac{1}{2} & \dots & \frac{1}{n} & \frac{1}{n+1} \\ 0 & \frac{1}{2} & \dots & \frac{1}{n} & \frac{1}{n+1} \\ \vdots & \ddots & \ddots & \vdots & \vdots \\ 0 & 0 & \ddots & \frac{1}{n} & \frac{1}{n+1} \\ 0 & 0 & \dots & 0 & \frac{1}{n+1} \end{pmatrix} \begin{matrix} e_0 \\ e_1 \\ \vdots \\ e_{n-1} \\ e_n \end{matrix}.$$

De même, on a, pour $k \in \llbracket 0, n \rrbracket$,

$$g(e_k) = (X-1)kX^{k-1} + X^k = (k+1)X^k - kX^{k-1}$$

de sorte que la matrice de f dans la base (e_0, e_1, \dots, e_n) est

$$B = \begin{pmatrix} g(e_0) & g(e_1) & g(e_2) & \dots & g(e_{n-1}) & g(e_n) \\ 1 & -1 & 0 & \dots & 0 & 0 \\ 0 & 2 & -2 & \ddots & & 0 \\ 0 & 0 & 3 & \ddots & & \vdots \\ \vdots & & \ddots & \ddots & \ddots & \vdots \\ \vdots & & & 0 & n & -n \\ 0 & \dots & \dots & \dots & 0 & n+1 \end{pmatrix} \begin{matrix} e_0 \\ e_1 \\ e_2 \\ \vdots \\ e_{n-1} \\ e_n \end{matrix}.$$

5. Les matrices A et B sont triangulaires supérieures, donc leurs valeurs propres sont leurs coefficients diagonaux.

Or, dans les deux cas, ces coefficients diagonaux sont deux à deux distincts, donc A et B possèdent $n+1$ valeurs propres, et par conséquent sont diagonalisables.

On en déduit que f et g sont diagonalisables.

Attention !
 A et B sont bien des matrices de taille $n+1 = \dim \mathbf{R}_n[X]$.

Partie II : Étude d'une suite de variables aléatoires.

6. Puisqu'à chaque étape, on ne peut obtenir qu'une boule portant un numéro entre 0 et n , $\{[Z_k = i], i \in \llbracket 0, n \rrbracket\}$ est un système complet d'événements. Et donc, par la formule des probabilités totales,

$$P(Z_{k+1} = r) = \sum_{i=1}^n P(Z_k = i)P_{[Z_k=i]}(Z_{k+1} = r).$$

Or, $P_{[Z_k=i]}(Z_{k+1} = r) = 0$ si $i < r$, car alors le $(k+1)$ -ième tirage a lieu dans l'urne i qui ne contient que des boules portant un numéro inférieur ou égal à $r-1$.

En revanche, si $i \geq r$, puisque l'urne U_i contient $i+1$ boules numérotées de 0 à i , et que toutes ces boules sont équiprobables, $P_{[Z_k=i]}(Z_{k+1} = r) = \frac{1}{i+1}$.

Et donc on a bien

$$P(Z_{k+1} = r) = \sum_{i=r}^n \frac{P(Z_k = i)}{i+1}.$$

7. La formule précédente appliquée avec $r = n$ donne

$$P(Z_{k+1} = n) = \frac{P(Z_k = n)}{n+1} \Leftrightarrow (n+1)P(Z_{k+1} = n) = P(Z_k = n).$$

De même, on a, pour $r \in \llbracket, n - 1 \rrbracket$,

$$\begin{aligned} (r + 1)P(Z_{k+1} = r) - (r + 1)P(Z_{k+1} = r + 1) &= (r + 1) \sum_{i=r}^n \frac{P(Z_{k+1} = i)}{i + 1} - (r + 1) \sum_{i=r+1}^n \frac{P(Z_k = i)}{i + 1} \\ &= (r + 1) \frac{P(Z_k = r)}{r + 1} = \boxed{P(Z_k = r)}. \end{aligned}$$

8. Par définition, on a $S_n = \sum_{k=0}^{+\infty} P(Z_k = n)$. Et alors, en utilisant les relations (\mathcal{R}_1) , il vient

$$S_n = \sum_{k=0}^{+\infty} P(Z_k = n) = (n + 1) \sum_{k=0}^{+\infty} P(Z_{k+1} = n) = (n + 1) \sum_{i=1}^{+\infty} P(Z_i = n) = (n + 1)(S_n - P(Z_0 = n)).$$

Mais, comme indiqué dans l'énoncé, $P(Z_0 = n) = 1$, et donc

$$S_n = (n + 1)(S_n - 1) \Leftrightarrow \boxed{S_n = \frac{n + 1}{n}}.$$

Comme indiqué, sommons alors les relations (\mathcal{R}_2) sur n :

$$\begin{aligned} S_{n-1} &= \sum_{k=0}^{+\infty} P(Z_k = n - 1) = \sum_{k=0}^{+\infty} (nP(Z_{k+1} = n - 1) - nP(Z_{k+1} = n)) \\ &= n \sum_{k=0}^{+\infty} P(Z_{k+1} = n - 1) - n \sum_{k=0}^{+\infty} P(Z_{k+1} = n) \\ &= n \sum_{i=1}^{+\infty} P(Z_i = n - 1) - n \sum_{i=1}^{+\infty} P(Z_i = n) \\ &= n(S_{n-1} - P(Z_0 = n - 1)) - n(S_n - P(Z_0 = n)) \\ &= nS_{n-1} - n(S_n - 1). \end{aligned}$$

Puisque Z_0 est certaine égale à n , $P(Z_0 = n - 1) = 0$.

Donc $(n - 1)S_{n-1} = nS_n - n = 1 \Leftrightarrow S_{n-1} = \frac{1}{n - 1}$.

Sur le même principe, on a pour $r \leq n - 2$,

$$\begin{aligned} S_r &= \sum_{k=0}^{+\infty} P(Z_k = r) = (r + 1) \sum_{k=1}^{+\infty} P(Z_k = r) - (r + 1) \sum_{k=0}^{+\infty} P(Z_k = r + 1) \\ &= (r + 1)(S_r - P(Z_0 = r)) - (r + 1)(S_{r+1} - P(Z_0 = r + 1)) \\ &= (r + 1)S_r - (r + 1)S_{r+1}. \end{aligned}$$

Et donc $rS_r = (r + 1)S_{r+1}$: la suite $(rS_r)_{1 \leq r \leq n-1}$ est constante.

9. On a $F'_{k+1}(x) = \sum_{r=0}^n P(Z_{k+1} = r)rx^{r-1}$ et donc

$$\begin{aligned} (x - 1)F'_{k+1}(x) + F_{k+1}(x) &= (x - 1) \sum_{r=0}^n P(Z_{k+1} = r)rx^{r-1} + \sum_{r=0}^n P(Z_{k+1} = r)x^r \\ &= \sum_{r=0}^n rP(Z_{k+1} = r)x^r - F'_{k+1}(x) + \sum_{r=0}^n P(Z_{k+1} = r)x^r \\ &= \sum_{r=0}^n (r + 1)P(Z_{k+1} = r)x^r - F'_{k+1}(x) \\ &= (n + 1)P(Z_{k+1} = n)x^n + \sum_{r=0}^{n-1} (r + 1)P(Z_{k+1} = r)x^r - F'_{k+1}(x) \\ &= P(Z_k = n)x^n + \sum_{r=0}^{n-1} (r + 1)P(Z_{k+1} = r)x^r - F'_{k+1}(x) \end{aligned}$$

Remarque
Bien que ça ne soit pas demandé, notons que puisque $(n - 1)S_{n-1} = 1$, pour tout $r \in \llbracket 1, n - 1 \rrbracket$, $rS_r = 1$.

C'est la relation (\mathcal{R}_1) .

$$\begin{aligned}
&= P(Z_k = n)x^n + \sum_{r=0}^{n-1} (P(Z_k = r) + (r+1)P(Z_{k+1} = r+1))x^r - F'_{k+1}(x) \quad \text{C'est la relation } (\mathcal{R}_2). \\
&= P(Z_k = n)x^n + \sum_{r=0}^{n-1} P(Z_k = r)x^r + \sum_{i=1}^n iP(Z_{k+1} = i)x^{i-1} - F'_{k+1}(x) \\
&= P(Z_k = n)x^n + \sum_{r=0}^{n-1} P(Z_k = r)x^r = \boxed{F_k(x)}.
\end{aligned}$$

Chgt d'indice
 $i = k + 1.$

10.a. On a $F'_k(x) = \sum_{r=0}^n rP(Z_k = r)x^{r-1}$ et donc $F'_k(1) = \sum_{r=0}^n rP(Z_k = r) = \boxed{E(Z_k)}$.

De même, on a $F''_k(x) = \sum_{r=0}^n r(r-1)P(Z_k = r)x^{r-2}$ et donc $F''_k(1) = \sum_{r=0}^n r(r-1)P(Z_k = r)$.

Mais par le théorème de transfert, nous reconnaissons là l'espérance de $Z_k(Z_k - 1)$ et donc

$$\boxed{F''_k(1) = E(Z_k(Z_k - 1))}.$$

10.b. Comme indiqué, dérivons la relation (\mathcal{S}) : il vient alors

$$\forall x \in \mathbf{R}, (x-1)F''_{k+1}(x) + F'_{k+1}(x) + F'_{k+1}(x) = F'_k(x) \Leftrightarrow (x-1)F''_{k+1}(x) + 2F'_{k+1}(x) = F'_k(x).$$

En évaluant en $x = 1$, il vient donc $\boxed{2F'_{k+1}(1) = F'_k(1)}$.

Et en dérivant une seconde fois :

$$\forall x \in \mathbf{R}, (x-1)F'''_{k+1}(x) + F''_{k+1}(x) + 2F''_{k+1}(x) = F''_k(x).$$

En évaluant en $x = 1$, il vient donc $\boxed{3F''_{k+1}(1) = F''_k(1)}$.

10.c. Nous venons de prouver que la suite $(F'_k(1))_{k \in \mathbf{N}}$ est géométrique de raison $\frac{1}{2}$, et donc

$$\forall k \in \mathbf{N}, F'_k(1) = \frac{1}{2^k} F'_0(1) = \frac{1}{2^k} E(Z_0) = \boxed{\frac{n}{2^k}}.$$

Rappel
 La variable Z_0 est certaine égale à n , donc $E(Z_0) = n$.

De même, la suite $(F''_k(1))_{k \in \mathbf{N}}$ est géométrique de raison $\frac{1}{3}$, de sorte que

$$\forall k \in \mathbf{N}, F''_k(1) = \frac{1}{3^k} F''_0(1) = \frac{1}{3^k} E(Z_0(Z_0 - 1)) = \frac{n(n-1)}{3^k}.$$

On a donc, par la formule de Huygens,

$$V(Z_k) = E(Z_k^2) - E(Z_k)^2 = E(Z_k(Z_k - 1)) + E(Z_k) - E(Z_k)^2 = \boxed{\frac{n(n-1)}{3^k} + \frac{n}{2^k} - \frac{n^2}{4^k}}.$$

Partie III : Loi de chacune de ces variables aléatoires.

11. Notons que la première égalité $\sum_{r=0}^n P(Z_k = r)e_r = F_k$ n'est rien d'autre que la définition de F_k .

D'autre part, la question 9 nous donne $g(F_{k+1}) = F_k$ et puisque $g = f^{-1}$, il vient $F_{k+1} = f(F_k)$.

Et donc de proche en proche, $F_k = f^k(F_0)$.

Mais $F_0 = \sum_{r=0}^n P(Z_0 = r)e_r = e_n$ car Z_0 est une variable certaine égale à n .

Et donc, pour tout $k \in \mathbf{N}$, $F_k = f^k(F_0) = \boxed{f^k(e_n)}$.

12. Pour tout $k \in \llbracket 0, n \rrbracket$, u_k est de degré k .

Et donc (u_0, u_1, \dots, u_n) est une famille de polynômes à degrés deux à deux distincts, donc est libre.

Puisqu'elle est de cardinal $n+1 = \dim \mathbf{R}_n[X]$, $\boxed{\text{c'est une base de } \mathbf{R}_n[X]}$.

13. Revenons à la définition de f : pour $x \in \mathbf{R} \setminus \{1\}$, on a

$$f(u_r)(x) = \frac{1}{x-1} \int_1^x (t-1)^r dt = \frac{1}{x-1} \left[\frac{(t-1)^{r+1}}{r+1} \right]_1^x = \frac{(x-1)^r}{r+1} = \frac{u_r}{r+1}.$$

Et pour $x = 1$, on a $f(u_k)(1) = u_k(1) = 0 = \frac{u_k(1)}{k+1}$.

Ainsi, u_r est un vecteur propre de f associé à la valeur propre $\frac{1}{r+1}$, et donc (u_0, u_1, \dots, u_n) est une base de $\mathbf{R}_n[X]$ formée de vecteurs propres de f : f est donc diagonalisable.

14. On a $e_n = X^n = (X-1+1)^n$, et donc par la formule du binôme de Newton,

$$e_n = \sum_{r=0}^n \binom{n}{r} (X-1)^r 1^{n-r} = \sum_{r=0}^n \binom{n}{r} u_r.$$

De même, toujours par la formule du binôme, pour $r \in \llbracket 0, n \rrbracket$,

$$u_r = (X-1)^r = \sum_{j=0}^r \binom{r}{j} X^j (-1)^{r-j} = \sum_{j=0}^r \binom{r}{j} (-1)^{r-j} e_j.$$

15. Par linéarité de f^k , on a $f^k(e_n) = \sum_{r=0}^n \binom{n}{r} f^k(u_r)$. Mais u_r est un vecteur propre de f associé à la valeur propre $\frac{1}{r+1}$ et donc $f^k(u_r) = \frac{u_r}{(r+1)^k}$.

Ainsi,
$$f^k(e_n) = \sum_{r=0}^n \frac{\binom{n}{r}}{(r+1)^k} u_r.$$

16. En combinant les résultats des questions précédentes, nous avons

$$\begin{aligned} F_k &= f^k(e_n) = \sum_{r=0}^n \frac{\binom{n}{r}}{(r+1)^k} u_r = \sum_{r=0}^n \frac{\binom{n}{r}}{(r+1)^k} \sum_{j=0}^r (-1)^{r-j} \binom{r}{j} e_j \\ &= \sum_{0 \leq j \leq r \leq n} (-1)^{r-j} \frac{\binom{n}{r} \binom{r}{j}}{(r+1)^k} e_j \\ &= \sum_{j=0}^n \left(\sum_{r=j}^n (-1)^{r-j} \frac{\binom{n}{r} \binom{r}{j}}{(r+1)^k} \right) e_j. \end{aligned}$$

Mais d'autre part, par définition, $F_k = \sum_{j=0}^n P(Z_k = j) e_j$.

Et donc, par unicité des coordonnées⁴ de F_k dans la base (e_0, e_1, \dots, e_n) , on a

$$\forall j \in \llbracket 0, n \rrbracket, P(Z_k = j) = \sum_{r=j}^n (-1)^{r-j} \frac{\binom{n}{r} \binom{r}{j}}{(r+1)^k}.$$

17. Application

- 17.a. En appliquant l'inégalité triangulaire à la formule obtenue dans la question 16, il vient

$$|P(Z_k = j)| \leq \sum_{r=j}^n \frac{\binom{n}{r} \binom{r}{j}}{(r+1)^k}.$$

Rappel

C'est l'une des caractérisations de la diagonalisabilité : il existe une base formée de vecteurs propres.

Rappel

Si $f(x) = \lambda x$, alors pour tout $k \in \mathbf{N}$,

$$f^k(x) = \lambda^k x.$$

⁴ Ce qui n'est rien d'autre que l'unicité des coefficients d'un polynôme.

Remarque

Notons que la valeur absolue donnée dans l'énoncé est totalement inutile : une probabilité est toujours positive !

Or, pour $r \geq j$, on a $\frac{1}{(r+1)^k} \leq \frac{1}{(j+1)^k}$.

Et donc

$$|P(Z_k = j)| \leq \sum_{r=j}^n \frac{\binom{n}{r} \binom{r}{j}}{(j+1)^k}.$$

On peut donc poser $M_{j,n} = \sum_{r=j}^n \binom{n}{r} \binom{r}{j}$.

Pour $j \in \llbracket 1, n \rrbracket$ fixé, on a, pour $k \geq 0$, $\frac{1}{(j+1)^k} \leq \frac{1}{2^k}$.

Mais la série de terme général $\frac{1}{2^k}$ est une série géométrique convergente.

Donc par critère de domination pour les séries à termes positifs, la série de terme général

$\frac{M_{j,n}}{(j+1)^k}$ converge, et donc la série de terme général $P(Z_k = j)$ converge également.

Remarque

Il s'agit du résultat admis par l'énoncé à la question 8 : la série définissant S_n est bien convergente.

17.b. On a $P(Z_k = 0) = \sum_{r=0}^n (-1)^r \frac{\binom{n}{r}}{(r+1)^k}$.

Et donc $P(Z_k = 0) - 1 = \sum_{r=1}^n (-1)^r \frac{\binom{n}{r}}{(r+1)^k}$.

Par l'inégalité triangulaire, on a donc

$$|P(Z_k = 0) - 1| \leq \sum_{r=1}^n \frac{\binom{n}{r}}{(r+1)^k}.$$

Mais pour $r \geq 1$, $\frac{1}{(r+1)^k} \leq \frac{1}{2^k}$ et donc

$$|P(Z_k = 0) - 1| \leq \frac{1}{2^k} \underbrace{\sum_{r=1}^n \binom{n}{r}}_{=C_n}.$$

On a donc $\lim_{k \rightarrow +\infty} |P(Z_k = 0) - 1| = 0 \Leftrightarrow \lim_{k \rightarrow +\infty} P(Z_k = 0) = 1$.

Et donc la série de terme général $P(Z_k = 0)$ diverge grossièrement.

Rappel

Une série dont le terme général ne tend pas vers 0 n'est **jamais** convergente.

DEVOIR SURVEILLÉ 8 (HEC) (23.03.19)

On rappelle que :

- pour tout réel x strictement positif, l'intégrale $\int_0^{+\infty} t^{x-1} e^{-t} dt = \int_0^{+\infty} e^{(x-1)\ln t} e^{-t} dt$ est convergente ;
- la fonction Γ est définie sur \mathbf{R}_+^* , et associée à tout réel x strictement positif, le réel strictement positif $\Gamma(x) = \int_0^{+\infty} t^{x-1} e^{-t} dt$;
- pour tout réel x strictement positif, $\Gamma(x+1) = x\Gamma(x)$.

Pour tout entier naturel k non nul, et pour toute fonction f définie sur \mathbf{R}_+^* , k fois dérivable, on note $f^{(k)}$ la dérivée k -ième de la fonction f . Les dérivées première et seconde sont également notées f' et f'' .

Dans les parties II et III du problème, \exp désigne la fonction exponentielle. Les parties III et IV sont indépendantes. Le problème a pour objet la mise en évidence de certaines propriétés de la fonction Γ .

Partie I. Une expression de $\Gamma(x)$

1. Soit n un entier supérieur ou égal à 1.

a. Pour tout réel u tel que $0 \leq u < 1$, montrer que $\ln(1-u) \leq -u$.

En déduire, pour tout réel t de l'intervalle $[0, n]$, l'inégalité : $\left(1 - \frac{t}{n}\right)^n \leq e^{-t}$.

b. Étudier les variations de la fonction φ définie sur $[0, \sqrt{n}]$ qui, à tout réel t de $[0, \sqrt{n}]$ associe :

$$\varphi(t) = \ln\left(1 - \frac{t^2}{n}\right) - t - n \ln\left(1 - \frac{t}{n}\right).$$

c. Établir, pour tout réel t de $[0, \sqrt{n}]$, l'inégalité :

$$\left(1 - \frac{t^2}{n}\right) e^{-t} \leq \left(1 - \frac{t}{n}\right)^n.$$

d. Justifier, pour tout réel t de $[0, n]$, les inégalités :

$$e^{-t} - \frac{t^2}{n} e^{-t} \leq \left(1 - \frac{t}{n}\right)^n \leq e^{-t}.$$

En déduire que, pour tout réel x strictement positif :

$$\Gamma(x) = \lim_{n \rightarrow +\infty} \int_0^n \left(1 - \frac{t}{n}\right)^n t^{x-1} dt.$$

2. a. Pour tout réel x strictement positif et pour tout entier naturel n non nul, montrer que les intégrales $\int_0^1 y^{x-1} dy$ et $\int_0^1 y^{x-1} (1-y)^n dy$ sont convergentes.

On pose alors $B_0(x) = \int_0^1 y^{x-1} dy$ et pour tout n supérieur ou égal à 1, $B_n(x) = \int_0^1 y^{x-1} (1-y)^n dy$.

b. À l'aide d'une intégration par parties, montrer, pour tout n de \mathbf{N}^* , l'égalité :

$$B_n(x) = \frac{n!}{x(x+1)\dots(x+n)}.$$

En déduire, pour tout n de \mathbf{N} , la formule :

$$B_n(x) = \frac{\Gamma(x) \times \Gamma(n+1)}{\Gamma(x+n+1)}.$$

c. Montrer que, pour tout réel x strictement positif :

$$\Gamma(x) = \lim_{n \rightarrow +\infty} \frac{n^x n!}{x(x+1)\dots(x+n)}.$$

En déduire que, pour tout réel x strictement positif, $\Gamma(x+n) \underset{n \rightarrow +\infty}{\sim} n^x (n-1)!$

d. Pour tout n de \mathbf{N}^* , on pose $\lambda_n = \frac{\Gamma\left(\frac{n}{2}\right)}{\Gamma\left(\frac{n+1}{2}\right)}$. Montrer que $\lambda_n \underset{n \rightarrow +\infty}{\sim} \sqrt{\frac{2}{n}}$.

Partie II. Dérivabilité de la fonction Γ et conséquences

3. a. Montrer que, pour tout entier naturel k non nul, et pour tout réel x strictement positif, l'intégrale $\int_0^{+\infty} t^{x-1} (\ln t)^k e^{-t} dt$ est absolument convergente. On note $g_k(x)$ la valeur de cette intégrale.

b. Soit $[a, b]$ un segment de \mathbf{R}_+^* . Soit x_0 et x deux éléments distincts de $]a, b[$. Établir l'inégalité :

$$|\Gamma(x) - \Gamma(x_0) - (x - x_0)g_1(x_0)| \leq \frac{(x - x_0)^2}{2} \int_0^{+\infty} (\ln t)^2 \left(\max_{\alpha \in [a, b]} t^{\alpha-1} \right) e^{-t} dt.$$

c. Montrer l'inégalité suivante :

$$\int_0^{+\infty} (\ln t)^2 \left(\max_{\alpha \in [a, b]} t^{\alpha-1} \right) e^{-t} dt \leq \int_0^1 (\ln t)^2 t^{a-1} e^{-t} dt + \int_1^{+\infty} (\ln t)^2 t^{b-1} e^{-t} dt.$$

En déduire que la fonction Γ est dérivable en x_0 et que $\Gamma'(x_0) = g_1(x_0)$.

d. Établir que la fonction Γ est dérivable sur \mathbf{R}_+^* et que $\Gamma' = g_1$.

On montrerait de même que la fonction Γ est deux fois dérivable sur \mathbf{R}_+^* , et que $\Gamma'' = g_2$. Ce résultat est admis dans toute la suite du problème.

4. Pour tout n de \mathbf{N}^* , on pose $\gamma_n = -\ln n + \sum_{k=1}^n \frac{1}{k}$.

a. Établir, pour tout entier n supérieur ou égal à 2, la double inégalité suivante :

$$\sum_{k=2}^n \frac{1}{k} < \ln n < \sum_{k=1}^{n-1} \frac{1}{k}.$$

En déduire que, pour tout entier naturel n non nul, on a $0 < \gamma_n \leq 1$.

b. Montrer que la suite $(\gamma_n)_{n \in \mathbf{N}^*}$ est décroissante et convergente. On note γ sa limite.

5. a. Pour tout réel x strictement positif, et pour tout entier n strictement positif, montrer l'égalité :

$$\prod_{k=1}^n \left[\left(1 + \frac{x}{k} \right) \exp\left(-\frac{x}{k}\right) \right] = \exp(-x\gamma_n) \times \frac{(x+1)(x+2) \cdots (x+n)}{n^x n!}.$$

b. On pose $v_n(x) = \prod_{k=1}^n \left[\left(1 + \frac{x}{k} \right) \exp\left(-\frac{x}{k}\right) \right]$.

Montrer que la suite $(v_n(x))_{n \in \mathbf{N}^*}$ est convergente. On note $\ell(x)$ sa limite. Montrer la relation :

$$\ell(x) = \frac{\exp(-\gamma x)}{x\Gamma(x)}.$$

6. a. Soit x un réel strictement positif fixé.

Montrer que la série de terme général $\frac{x}{n} - \ln\left(1 + \frac{x}{n}\right)$, $n \geq 1$, est convergente.

b. Justifier, pour tout réel x strictement positif, l'égalité :

$$\ln(\ell(x)) = - \sum_{n=1}^{+\infty} \left[\frac{x}{n} - \ln\left(1 + \frac{x}{n}\right) \right].$$

En déduire, pour tout réel x strictement positif, la relation :

$$\ln(\Gamma(x)) = -\gamma x - \ln x + \sum_{n=1}^{+\infty} \left[\frac{x}{n} - \ln\left(1 + \frac{x}{n}\right) \right].$$

7. Soit ψ la dérivée de $x \mapsto \ln(\Gamma(x))$, qui est donc définie sur \mathbf{R}_+^* .

Établir, pour tout réel x strictement positif l'égalité : $\psi(x+1) = \psi(x) + \frac{1}{x}$.

Déterminer un équivalent simple de $\psi(x)$ lorsque x tend vers 0^+ .

Justifier, pour tout entier n supérieur ou égal à 2, la formule :

$$\psi(n) = \psi(1) + \sum_{k=1}^{n-1} \frac{1}{k}.$$

8. Pour tout entier n supérieur ou égal à 1, on considère la fonction U_n définie sur \mathbf{R}_+^* par :

$$U_n(x) = \ln\left(1 + \frac{x}{n}\right) - \frac{x}{n}.$$

On désigne par $A(x)$ la somme de la série de terme général $U_n(x)$.

- a. Montrer que A est deux fois dérivable sur \mathbf{R}_+^* . En particulier, exprimer, pour tout réel x strictement positif, $A'(x)$ et $A''(x)$ en fonction de $\Gamma(x)$, $\Gamma'(x)$ et $\Gamma''(x)$.
- b. Soit x un réel strictement positif fixé. Montrer que, pour tout entier k supérieur ou égal à 1, la série de terme général $U_n^{(k)}(x)$ est absolument convergente.

Dans toute la suite du problème, **on admet** les deux résultats suivants : pour tout réel x strictement positif on a

$$A'(x) = \sum_{n=1}^{+\infty} U_n'(x) \text{ et } A''(x) = \sum_{n=1}^{+\infty} U_n''(x).$$

- c. Calculer $\psi(1)$ en fonction de γ . En déduire la valeur de $\lim_{n \rightarrow +\infty} (\ln n - \psi(n))$.
9. On veut établir dans cette question que pour tout réel y strictement positif, on a $\psi'(y) > \frac{1}{y}$.
Soit x un réel strictement positif fixé. On considère la fonction G définie sur \mathbf{R}_+^* qui, à tout réel t strictement positif, associe $G(t) = \frac{1}{(t+x)^2}$.

- a. Montrer que sur \mathbf{R}_+^* , G est positive, strictement décroissante, et que l'intégrale $\int_1^{+\infty} G(t) dt$ est convergente.
- b. En déduire la double inégalité : $0 < \int_1^{+\infty} G(t) dt < \sum_{k=1}^{+\infty} G(k)$.
- c. Établir l'inégalité : $\psi'(x) > \frac{1}{x+1} + \frac{1}{x^2}$. Conclure.

Partie III. Estimation des paramètres d'une loi $\Gamma(\theta, r)$

Soit $(\theta, r) \in (\mathbf{R}_+^*)^2$, et soit f la fonction définie sur \mathbf{R} par

$$f(x) = \begin{cases} \frac{1}{\Gamma(r)\theta^r} \times x^{r-1} \exp\left(-\frac{x}{\theta}\right) & \text{si } x > 0 \\ 0 & \text{si } x \leq 0 \end{cases}$$

- 10. Montrer que f est une densité de probabilités.
Si X est une variable aléatoire de densité f , on dit que X suit la loi Γ de paramètres θ et r et on note $X \leftrightarrow \Gamma(\theta, r)$.

On considère une variable aléatoire X , qui suit une loi $\Gamma(\theta, r)$, les deux paramètres inconnus θ et r étant des réels strictement positifs.

Soit p un entier supérieur ou égal à 2. On considère un p -échantillon i.i.d. (X_1, X_2, \dots, X_p) de la loi de X : les variables aléatoires X_1, \dots, X_p sont mutuellement indépendantes et de même loi que X .

On désigne par x_1, \dots, x_p , un p -échantillon de réalisations des variables aléatoires X_1, \dots, X_p , respectivement ; les réels x_1, \dots, x_p sont fixés, strictement positifs et non tous égaux.

Soit L la fonction (appelée fonction de vraisemblance) définie sur $\mathbf{R}_+^* \times \mathbf{R}_+^*$ à valeurs dans \mathbf{R}_+^* qui, à tout couple (θ, r) de réels strictement positifs, associe :

$$L(\theta, r) = \prod_{i=1}^p f(x_i).$$

On pose $F(\theta, r) = \ln(L(\theta, r))$.

- 11. Montrer que la recherche du maximum de L sur $\mathbf{R}_+^* \times \mathbf{R}_+^*$ est équivalente à la recherche du maximum de F sur ce même ensemble.
- 12. a. Montrer que la fonction F est de classe \mathcal{C}^2 sur $\mathbf{R}_+^* \times \mathbf{R}_+^*$ et calculer ses dérivées partielles d'ordre 1 et 2 de la fonction F .

b. Montrer que les éventuels points critiques (θ^*, r^*) vérifient le système (S) d'équations suivant

$$(S) \quad \begin{cases} \theta^* r^* = \bar{x} & (1) \\ \ln r^* - \frac{\Gamma'(r^*)}{\Gamma(r^*)} = \ln \bar{x} - \frac{1}{p} \sum_{i=1}^p \ln x_i & (2) \end{cases}$$

dans lequel $\bar{x} = \frac{1}{p} \sum_{i=1}^p x_i$.

13. On pose $K_p = \ln \bar{x} - \frac{1}{p} \sum_{i=1}^p \ln x_i$.

- a. Justifier, pour tout réel $x > 0$ et différent de 1, l'inégalité : $\ln x < x - 1$. En déduire que $K_p > 0$.
 b. Soit h la fonction définie sur \mathbf{R}_+^* par :

$$h(y) = \ln y - \frac{\Gamma'(y)}{\Gamma(y)} - K_p.$$

Étudier les variations de h et dresser son tableau de variations.

- c. Montrer que l'équation (2) admet sur \mathbf{R}_+^* une unique solution r^* . En déduire que le système d'équations (S) admet une unique solution (θ^*, r^*) .
 14. Écrire la hessienne $\nabla^2 F$ de F au point (θ^*, r^*) . Déterminer le signe de son déterminant et de sa trace. En déduire qu'au point (θ^*, r^*) , la fonction L admet un maximum local.

On peut démontrer qu'en ce point, on obtient en fait un maximum global de L . On dit que le couple (θ^*, r^*) est une estimation du couple inconnu (θ, r) obtenue par la méthode du maximum de vraisemblance.

Partie IV. Estimateur sans biais de l'écart-type σ d'une loi normale centrée

Soit X une variable aléatoire qui suit une loi normale centrée et d'écart-type σ ; le paramètre réel inconnu σ , est strictement positif.

15. Montrer que la variable aléatoire $T = \frac{X^2}{2\sigma^2}$ suit une loi γ de paramètre $1/2$. En déduire la valeur de $\Gamma(1/2)$.
 16. Pour n entier naturel non nul, on considère un n -échantillon (X_1, X_2, \dots, X_n) i.i.d. (indépendant, identiquement distribué) de la loi de X .
 a. On désigne par S_n la variable aléatoire $S_n = \sum_{i=1}^n \frac{X_i^2}{2\sigma^2}$. Quelle est la loi de probabilité de S_n ?
 b. En déduire que la variable aléatoire Y_n définie par $Y_n = \frac{1}{n} \sum_{i=1}^n X_i^2$ est un estimateur sans biais de σ^2 .
 c. Montrer que l'espérance de $\sqrt{Y_n}$, notée $E(\sqrt{Y_n})$, vérifie : $E(\sqrt{Y_n}) < \sigma$.
 d. Donner l'expression de $E(\sqrt{Y_n})$ en fonction de n et σ .
 e. Montrer que la variable aléatoire $\widehat{\sigma}_n$ définie par

$$\widehat{\sigma}_n = \frac{\lambda_n}{\sqrt{2}} \cdot \left(\sum_{i=1}^n X_i^2 \right)^{1/2}$$

où λ_n a été défini dans la question 2.d, est un estimateur sans biais du paramètre σ .

17. a. Calculer la variance $V(\widehat{\sigma}_n)$ de l'estimateur $\widehat{\sigma}_n$ en fonction de n et σ .
 b. La suite $(\widehat{\sigma}_n)_{n \in \mathbf{N}^*}$ d'estimateurs de σ converge-t-elle en probabilité vers σ ?

CORRECTION DU DEVOIR SURVEILLÉ 8

HEC 2005

Partie I. Une expression de $\Gamma(x)$.

- 1.a. La fonction $x \mapsto \ln(1-x)$ est dérivable sur $[0, 1[$, et sa dérivée est $x \mapsto -\frac{1}{1-x}$ qui est décroissante, de sorte que $x \mapsto \ln(1-x)$ est concave. Sa courbe représentative se situe donc sous ses tangentes. En particulier, la tangente en $x=0$ est la droite d'équation $y=-x$, et donc $\forall u \in [0, 1[$, $\ln(1-u) \leq -u$.
En particulier, pour tout $t \in [0, n]$, on a $\ln\left(1-\frac{t}{n}\right) \leq -\frac{t}{n}$. Soit encore $n \ln\left(1-\frac{t}{n}\right) \leq -t$.
En passant à l'exponentielle¹, il vient donc

$$e^{n \ln\left(1-\frac{t}{n}\right)} = \left(1-\frac{t}{n}\right)^n \leq e^{-t}.$$

- 1.b. La fonction φ est dérivable² sur $[0, \sqrt{n}[$ et

$$\begin{aligned} \varphi'(t) &= -\frac{2t}{n\left(1-\frac{t^2}{n}\right)} - 1 + \frac{1}{1-\frac{t}{n}} \\ &= \frac{-\frac{2t}{n}\left(1-\frac{t}{n}\right) - \left(1-\frac{t^2}{n}\right)\left(1-\frac{t}{n}\right) + 1 - \frac{t^2}{n}}{1-\frac{t^2}{n}} \\ &= \frac{-\frac{2t}{n} + \frac{2t^2}{n^2} - 1 + \frac{t}{n} + \frac{t^2}{n} - \frac{t^3}{n^2} + 1 - \frac{t^2}{n}}{\left(1-\frac{t^2}{n}\right)\left(1-\frac{t}{n}\right)} \\ &= \frac{-\frac{t^3}{n^2} + \frac{2t^2}{n^2} - \frac{t}{n}}{\left(1-\frac{t^2}{n}\right)\left(1-\frac{t}{n}\right)} \\ &= \frac{t}{n^2} \frac{-t^2 + 2t - n}{\left(1-\frac{t^2}{n}\right)\left(1-\frac{t}{n}\right)}. \end{aligned}$$

Le polynôme $-t^2 + 2t - n$ a pour discriminant $\Delta = 4 - 4n \leq 0$, et donc il est de signe constant sur \mathbf{R} . Étant de coefficient dominant négatif, on a donc $\forall t \in \mathbf{R}$, $-t^2 + 2t - n \leq 0$. On en déduit que $\varphi'(t) \leq 0$, $\forall t \in [0, \sqrt{n}[$ et donc que φ est décroissante sur cet intervalle.

- 1.c. Il s'agit de prouver que pour $t \in [0, \sqrt{n}[$,

$$\left(1-\frac{t^2}{n}\right)e^{-t} \leq \left(1-\frac{t}{n}\right)^n \Leftrightarrow \ln\left(1-\frac{t^2}{n}\right) - t \leq n \ln\left(1-\frac{t}{n}\right) \Leftrightarrow \varphi(t) \leq 0.$$

Mais puisque φ est décroissante sur $[0, \sqrt{n}[$, avec $\varphi(0) = 0$, elle est négative sur $[0, \sqrt{n}[$ et donc

$$\forall t \in [0, \sqrt{n}[, \left(1-\frac{t^2}{n}\right)e^{-t} \leq \left(1-\frac{t}{n}\right)^n.$$

- 1.d. D'après la question précédente, on a pour $t \in [0, \sqrt{n}[$,

$$\left(1-\frac{t^2}{n}\right)e^{-t} \leq \left(1-\frac{t}{n}\right)^n \Leftrightarrow e^{-t} - \frac{t^2}{n}e^{-t} \leq \left(1-\frac{t}{n}\right)^n.$$

Pour $t \in [\sqrt{n}, n]$, cette inégalité reste valable puisque $\left(1-\frac{t^2}{n}\right) \leq 0$ et $\left(1-\frac{t}{n}\right)^n \geq 0$.

Et il a été prouvé à la question 1.a que pour tout $t \in [0, n]$, $\left(1-\frac{t}{n}\right)^n \leq e^{-t}$.

Après multiplication par $t^{x-1} \geq 0$, il vient, pour tout $t \in [0, n]$,

$$t^{x-1}e^{-t} - \frac{1}{n}t^{x+1}e^{-t} \leq \left(1-\frac{t}{n}\right)^n t^{x-1} \leq t^{x-1}e^{-t}.$$

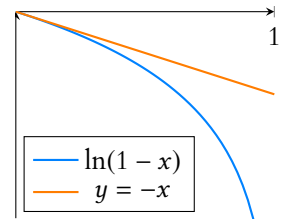


FIGURE 24.1– La fonction $\ln(1-x)$ et sa tangente en 0.

¹ Qui préserve le sens des inégalités car elle est croissante.

² Car somme de composées de fonctions usuelles dérivables.

Alternative

Notons que

$$\begin{aligned} -t^2 + 2t - n &= \\ -(t-1)^2 - (n-1) &\leq 0 \end{aligned}$$

Et donc par croissance de l'intégrale,

$$\int_0^n t^{x-1} e^{-t} dt - \frac{1}{n} \int_0^n t^{x+1} e^{-t} dt \leq \int_0^n \left(1 - \frac{t}{n}\right)^n t^{x-1} dt \leq \int_0^n t^{x-1} e^{-t} dt.$$

Lorsque $n \rightarrow +\infty$, $\int_0^n t^{x-1} e^{-t} dt \xrightarrow{n \rightarrow +\infty} \int_0^{+\infty} t^{x-1} e^{-t} dt = \Gamma(x)$.

De même $\int_0^{\sqrt{n}} t^{x+1} e^{-t} dt \xrightarrow{n \rightarrow +\infty} \Gamma(x+2)$ et donc $\lim_{n \rightarrow +\infty} \frac{1}{n} \int_0^n t^{x+1} e^{-t} dt = 0$.

Et donc, par le théorème des gendarmes, il vient

$$\lim_{n \rightarrow +\infty} \int_0^n \left(1 - \frac{t}{n}\right)^n t^{x-1} dt = \Gamma(x).$$

2.a. Si $x \geq 1$, il s'agit de deux intégrales sur un segment, donc le résultat est évident.

En revanche, pour $x \in]0, 1[$, les deux fonctions intégrées ne sont continues que sur $]0, 1[$.

Mais $\int_0^1 y^{x-1} dy = \int_0^1 \frac{1}{y^{1-x}} dy$, avec $1-x < 1$: il s'agit donc d'une intégrale de Riemann convergente.

Et au voisinage de 0, $y^{x-1}(1-y)^n \underset{y \rightarrow 0}{\sim} y^{x-1}$, de sorte que par critère de comparaison pour les

intégrales de fonctions positives, $\int_0^1 y^{x-1}(1-y)^n dy$ converge car $\int_0^1 y^{x-1} dy$ converge.

2.b. Notons que la valeur de $B_0(x)$ est facile à calculer³

$$B_0(x) = \lim_{A \rightarrow 0^+} \int_A^1 y^{x-1} dy = \lim_{A \rightarrow 0^+} \left[\frac{1}{x} y^x \right]_A^1 = \lim_{A \rightarrow 0^+} \frac{1}{x} - \frac{A^x}{x} = \frac{1}{x}.$$

Commençons par supposer que $x \geq 1$, de sorte qu'on peut directement procéder à une intégration par parties sur le segment $[0, 1]$.

Posons donc $u(y) = \frac{1}{x} y^x$ et $v(y) = (1-y)^n$, qui sont deux fonctions de classe \mathcal{C}^1 sur $[0, 1]$, avec $u'(y) = y^{x-1}$ et $v'(y) = -n(1-y)^{n-1}$.

Par intégration par parties, on a alors

$$B_n(x) = \left[\frac{1}{x} y^x (1-y)^n \right]_0^1 + \frac{n}{x} \int_0^1 y^x (1-y)^{n-1} dy = \frac{n}{x} \int_0^1 y^x (1-y)^{n-1} dy = \frac{n}{x} B_{n-1}(x+1).$$

De proche en proche, on a alors

$$\begin{aligned} B_n(x) &= \frac{n}{x} B_{n-1}(x+1) \\ &= \frac{n}{x} \frac{n-1}{x+1} B_{n-2}(x+2) \\ &= \dots = \frac{n}{x} \frac{n-1}{x+1} \dots \frac{1}{x+n-1} B_0(x+n) \\ &= \frac{n}{x} \frac{n-1}{x+1} \dots \frac{1}{x+n-1} \frac{1}{x+n} \\ &= \frac{n!}{x(x+1) \dots (x+n)}. \end{aligned}$$

Si $x \in]0, 1[$, le raisonnement est le même, sauf qu'il faut prendre un tout petit peu plus de précautions pour la première intégration par parties, puisque celle-ci doit être réalisée sur un segment de la forme $[A, 1]$, puis il faudra passer à la limite lorsque $A \rightarrow 0^+$.

On prouve alors de la même manière que $B_n(x) = \frac{x}{n} B_{n-1}(x+1)$, le reste du raisonnement est alors le même.

Nous savons déjà que $n! = \Gamma(n+1)$.

D'autre part, en utilisant plusieurs fois⁴ la relation $\Gamma(x+1) = x\Gamma(x)$, il vient

$$\Gamma(x+n+1) = (x+n)\Gamma(x+n) = (x+n)(x+n-1)\Gamma(x+n-1) = \dots = (x+n) \dots x\Gamma(x).$$

Convergence

Notons que ces intégrales convergent, bien qu'elles soient impropres en 0 si $x \in]0, 1[$.
Mais alors $t^{x-1} e^{-t} \underset{t \rightarrow 0}{\sim} t^{x-1}$, et nous reconnaissons alors une intégrale de Riemann convergente (car $1-x < 1$).

³ C'est presque du cours.

Limite

Notons que la limite de A^x est nulle car $x \geq 0$.

⁴ n fois pour être précis.

Et donc $\frac{\Gamma(x+n+1)}{\Gamma(x)} = (x+n)(x+n-1)\cdots x$.

On en déduit donc bien que $B_n(x) = \frac{\Gamma(x) \times \Gamma(n+1)}{\Gamma(x+n+1)}$.

2.c. Dans l'intégrale $\int_0^n \left(1 - \frac{t}{n}\right)^n t^{x-1} dt$, procédons au changement de variable⁵ $y = \frac{t}{n}$.
Alors il vient $dt = n dy$ et donc

⁵ Affine.

$$\begin{aligned} \int_0^n \left(1 - \frac{t}{n}\right)^n t^{x-1} dt &= \int_0^1 (1-y)^n (ny)^{x-1} n dy = n^x \int_0^1 (1-y)^n y^{x-1} dy \\ &= n^x B_n(x) = n^x \frac{n!}{x(x+1)\cdots(x+n)}. \end{aligned}$$

Et donc la limite de la question 1.d s'écrit encore

$$\Gamma(x) = \lim_{n \rightarrow +\infty} \frac{n^x n!}{x(x+1)\cdots(x+n)}.$$

On a donc

$$\begin{aligned} \frac{\Gamma(x+n)}{n^x(n-1)!} &= \frac{(x+n-1)(x+n-2)\cdots x \Gamma(x)}{n^x(n-1)!} \\ &\underset{n \rightarrow +\infty}{\sim} \frac{(x+n-1)(x+n-2)\cdots x}{n^x(n-1)!} \frac{n^x n!}{x(x+1)\cdots(x+n)} \\ &\underset{n \rightarrow +\infty}{\sim} \frac{x+n}{n} \xrightarrow[n \rightarrow +\infty]{} 1. \end{aligned}$$

Et donc on a bien $\Gamma(x+n) \underset{n \rightarrow +\infty}{\sim} n^x(n-1)!$

2.d. Commençons par traiter le cas des termes d'indice pair :

$$\lambda_{2n} = \frac{\Gamma(2n/2)}{\Gamma\left(\frac{2n+1}{2}\right)} = \frac{\Gamma(n)}{\Gamma\left(\frac{1}{2}+n\right)} = \frac{(n-1)!}{\Gamma\left(\frac{1}{2}+n\right)}.$$

Mais $\Gamma\left(\frac{1}{2}+n\right) \underset{n \rightarrow +\infty}{\sim} n^{1/2}(n-1)!$, de sorte que

$$\lambda_{2n} \underset{n \rightarrow +\infty}{\sim} \frac{(n-1)!}{\sqrt{n}(n-1)!} = \sqrt{\frac{1}{n}} = \sqrt{\frac{2}{2n}}.$$

De même, pour les termes impairs, on a $\lambda_{2n+1} = \frac{\Gamma\left(\frac{1}{2}+n\right)}{\Gamma(n+1)} = \frac{\Gamma\left(\frac{1}{2}+n\right)}{n!}$. Et donc

$$\lambda_{n+1} \underset{n \rightarrow +\infty}{\sim} \frac{n^{1/2}(n-1)!}{n!} \underset{n \rightarrow +\infty}{\sim} \frac{\sqrt{n}}{n} \underset{n \rightarrow +\infty}{\sim} \sqrt{\frac{1}{n}} \underset{n \rightarrow +\infty}{\sim} \sqrt{\frac{2}{2n+1}}.$$

Et donc on a bien $\lambda_n \underset{n \rightarrow +\infty}{\sim} \sqrt{\frac{2}{n}}$.

Détails

Pour que deux suites (u_n) et (v_n) soient équivalentes, il faut et il suffit que

$$\begin{cases} u_{2n} \sim v_{2n} \\ u_{2n+1} \sim v_{2n+1} \end{cases}$$

Cela découle du fait qu'une suite converge vers l si et seulement si ses sous-suites des termes d'indice pair et des termes d'indice impair convergent vers la même limite, que l'on applique ici à $\frac{u_n}{v_n}$.

Partie II. Dérivabilité de la fonction Γ et conséquences.

3.a. La fonction $t \mapsto t^{x-1} (\ln t)^k e^{-t}$ est continue sur $]0, +\infty[$.

Au voisinage de 0, $t^{x-1} (\ln t)^k e^{-t} \underset{t \rightarrow 0^+}{\sim} t^{x-1} (\ln t)^k$.

Et alors, si $\alpha \in]1-x, 1[$, on a

$$t^\alpha t^{x-1} (\ln t)^k e^{-t} \underset{t \rightarrow 0^+}{\sim} t^{x-1+\alpha} (\ln t)^k \xrightarrow[t \rightarrow 0^+]{} 0$$

Autrement dit, $t^{x-1} (\ln t)^k e^{-t} = o_{t \rightarrow 0^+} \left(\frac{1}{t^\alpha}\right)$.

Mais puisque $\alpha < 1$, $\int_0^1 \frac{dt}{t^\alpha}$ converge, et donc il en est de même de $\int_0^1 t^{x-1} (\ln t)^k e^{-t} dt$.

Détails

Il s'agit directement d'une croissance comparée, en notant que

$$x+1-\alpha > 0.$$

Au voisinage de $+\infty$, on a

$$e^{t/2} t^{x-1} (\ln t)^k e^{-t} = \underbrace{\frac{t^{x-1}}{e^{t/4}}}_{\xrightarrow{t \rightarrow +\infty} 0} \underbrace{\frac{(\ln t)^k}{e^{t/4}}}_{\xrightarrow{t \rightarrow +\infty} 0} \xrightarrow{t \rightarrow +\infty} 0.$$

Et donc $t^{x-1} (\ln t)^k e^{-t} = o_{t \rightarrow +\infty}(e^{-t/2})$.

On en déduit que $\int_1^{+\infty} t^{x-1} (\ln t)^k e^{-t} dt$ converge et donc que $\int_0^{+\infty} t^{x-1} (\ln t)^k e^{-t} dt$ converge.

3.b. Soit $t \in \mathbb{R}_+^*$ fixé.

Appliquons alors l'inégalité de Taylor-Lagrange à la fonction $f_t : u \mapsto t^{u-1} = e^{(u-1)\ln(t)}$.

Celle-ci est de classe \mathcal{C}^2 sur $[a, b]$, avec $f_t'(u) = (\ln t)t^{u-1}$ et $f_t''(u) = (\ln t)^2 t^{u-1}$.

En particulier, pour $u \in [a, b]$, on a $|f_t''(u)| \leq (\ln t)^2 \max_{\alpha \in [a, b]} t^{\alpha-1}$.

Et donc par l'inégalité de Taylor-Lagrange, pour x et x_0 dans $[a, b]$,

$$|f_t(x) - f_t(x_0) - (x - x_0)f_t'(x_0)| \leq \frac{(x - x_0)^2}{2} (\ln t)^2 \max_{\alpha \in [a, b]} t^{\alpha-1}.$$

Après multiplication par e^{-t} , il vient donc, pour tout $t > 0$,

$$|t^{x-1} e^{-t} - t^{x_0-1} e^{-t} - (x - x_0)t^{x_0-1} (\ln t) e^{-t}| \leq \frac{(x - x_0)^2}{2} (\ln t)^2 \max_{\alpha \in [a, b]} t^{\alpha-1} e^{-t}.$$

Et donc, par croissance de l'intégrale⁶ :

$$\int_0^{+\infty} |t^{x-1} e^{-t} - t^{x_0-1} e^{-t} - (x - x_0)t^{x_0-1} (\ln t) e^{-t}| dt \leq \frac{(x - x_0)^2}{2} \int_0^{+\infty} (\ln t)^2 \left(\max_{\alpha \in [a, b]} t^{\alpha-1} \right) e^{-t} dt.$$

D'autre part, d'après l'inégalité triangulaire,

$$\begin{aligned} |\Gamma(x) - \Gamma(x_0) - (x - x_0)g_1(x_0)| &= \left| \int_0^{+\infty} (t^{x-1} e^{-t} - t^{x_0-1} e^{-t} - (x - x_0)t^{x_0-1} (\ln t) e^{-t}) dt \right| \\ &\leq \int_0^{+\infty} |t^{x-1} e^{-t} - t^{x_0-1} e^{-t} - (x - x_0)t^{x_0-1} (\ln t) e^{-t}| dt \\ &\leq \frac{(x - x_0)^2}{2} \int_0^{+\infty} (\ln t)^2 \left(\max_{\alpha \in [a, b]} t^{\alpha-1} \right) e^{-t} dt. \end{aligned}$$

Il nous reste donc juste à justifier la convergence de cette dernière intégrale.

Pour $t \in]0, 1]$, et $\alpha \in [a, b]$, on a $t^{\alpha-1} = e^{(\alpha-1)\ln(t)}$, avec $\ln(t) \leq 0$, de sorte que

$$(\alpha - 1)\ln(t) \leq (a - 1)\ln(t) \Rightarrow |t^{\alpha-1}| \leq e^{(a-1)\ln t} = t^{a-1}.$$

Et donc $0 \leq (\ln t)^2 \left(\max_{\alpha \in [a, b]} t^{\alpha-1} \right) e^{-t} \leq (\ln t)^2 t^{a-1} e^{-t}$.

Puisqu'il a été prouvé à la question 3.a que $\int_0^1 (\ln t)^2 t^{a-1} e^{-t} dt$ converge, il en est de même

de $\int_0^1 (\ln t)^2 \left(\max_{\alpha \in [a, b]} t^{\alpha-1} \right) e^{-t} dt$.

Pour $t \geq 1$, le raisonnement est le même, sauf que $\ln t \geq 0$, de sorte que pour $\alpha \in [a, b]$, $t^{\alpha-1} = e^{(\alpha-1)\ln t} \leq t^{b-1}$.

Et puisque $\int_1^{+\infty} (\ln t)^2 t^{b-1} e^{-t} dt$ converge, il en est de même de $\int_1^{+\infty} (\ln t)^2 \left(\max_{\alpha \in [a, b]} t^{\alpha-1} \right) e^{-t} dt$.

En conclusion $\int_0^{+\infty} (\ln t)^2 \left(\max_{\alpha \in [a, b]} t^{\alpha-1} \right) e^{-t} dt$ converge et donc toutes les inégalités précédemment données sont bien légitimes.

3.c. Par la relation de Chasles,

$$\int_0^{+\infty} (\ln t)^2 \left(\max_{\alpha \in [a, b]} t^{\alpha-1} \right) e^{-t} dt = \int_0^1 (\ln t)^2 \left(\max_{\alpha \in [a, b]} t^{\alpha-1} \right) e^{-t} dt + \int_1^{+\infty} (\ln t)^2 \left(\max_{\alpha \in [a, b]} t^{\alpha-1} \right) e^{-t} dt.$$

⁶ Sous réserve que l'intégrale de gauche converge, ce qui n'est pas totalement évident et que nous prouverons plus tard.

Et alors les calculs ont été faits à la question précédente, lorsque nous avons prouvé la convergence de l'intégrale.

On en déduit que pour $x \in]a, b[\setminus \{x_0\}$, on a

$$\left| \frac{\Gamma(x) - \Gamma(x_0)}{x - x_0} - g_1(x_0) \right| \leq \frac{|x - x_0|}{2} \left(\int_0^1 (\ln t)^2 t^{a-1} e^{-t} dt + \int_1^{+\infty} (\ln t)^2 t^{b-1} e^{-t} dt \right).$$

Et donc par le théorème des gendarmes,

$$\lim_{x \rightarrow x_0} \left| \frac{\Gamma(x) - \Gamma(x_0)}{x - x_0} - g_1(x_0) \right| = 0 \Leftrightarrow \lim_{x \rightarrow x_0} \frac{\Gamma(x) - \Gamma(x_0)}{x - x_0} = g_1(x_0).$$

Ceci signifie donc que Γ est dérivable en x_0 et que $\Gamma'(x_0) = g_1(x_0)$.

3.d. Soit $x_0 \in \mathbf{R}_+^*$. Alors x_0 appartient à l'intérieur du segment $\left[\frac{x_0}{2}, 2x_0 \right]$, qui est bien inclus dans \mathbf{R}_+^* .

Et donc le raisonnement de la question précédente s'applique : Γ est dérivable en x_0 et $\Gamma'(x_0) = g_1(x_0)$.

Ceci étant vrai pour tout $x_0 \in \mathbf{R}_+^*$, Γ est dérivable sur \mathbf{R}_+^* et $\Gamma' = g_1$.

4.a. Soit $k \geq 1$. Alors la fonction $t \mapsto \frac{1}{t}$ est décroissante sur $[k, k+1]$. Et donc pour tout $t \in [k, k+1]$, on a

$$\frac{1}{k+1} \leq \frac{1}{t} \leq \frac{1}{k}.$$

Par croissance de l'intégrale, il vient donc

$$\frac{1}{k+1} = \int_k^{k+1} \frac{dt}{k+1} \leq \int_k^{k+1} \frac{dt}{t} \leq \int_k^{k+1} \frac{dt}{k} = \frac{1}{k}.$$

En sommant ces relations pour k variant de 1 à $n-1$, il vient

$$\sum_{k=1}^{n-1} \frac{1}{k+1} \leq \sum_{k=1}^{n-1} \int_k^{k+1} \frac{dt}{t} \leq \sum_{k=1}^{n-1} \frac{1}{k}.$$

La somme de gauche est évidemment égale à $\sum_{k=2}^n \frac{1}{k}$, et par la relation de Chasles,

$$\sum_{k=1}^{n-1} \int_k^{k+1} \frac{dt}{t} = \int_1^2 \frac{dt}{t} + \int_2^3 \frac{dt}{t} + \cdots + \int_{n-1}^n \frac{dt}{t} = \int_1^n \frac{dt}{t} = [\ln t]_1^n = \ln(n).$$

Donc déjà, $\sum_{k=2}^n \frac{1}{k} \leq \ln(n) \leq \sum_{k=1}^{n-1} \frac{1}{k}$.

Il nous reste donc à voir que ces inégalités sont strictes.

On peut par exemple constater que $t \mapsto \frac{1}{t} - \frac{1}{2}$ est continue, positive sur $[1, 2]$, et n'est pas la fonction nulle.

$$\text{Et donc } \int_1^2 \left(\frac{1}{t} - \frac{1}{2} \right) dt > 0 \Leftrightarrow \int_1^2 \frac{dt}{t} > \frac{1}{2}.$$

Ainsi, l'une des inégalités que nous avons sommées ci-dessus est stricte, on obtient donc une inégalité stricte : $\sum_{k=2}^n \frac{1}{k} < \ln(n)$.

On prouverait de la même manière que l'autre inégalité est stricte.

4.b. On a, pour tout $n \geq 1$,

$$\gamma_{n+1} - \gamma_n = -\ln(n+1) + \sum_{k=1}^{n+1} \frac{1}{k} + \ln(n) - \sum_{k=1}^n \frac{1}{k} = -\ln\left(\frac{n+1}{n}\right) + \frac{1}{n+1} = \frac{1}{n+1} - \ln\left(1 + \frac{1}{n}\right).$$

Détails

Les deux intégrales qui apparaissent ne dépendent pas de x .

Remarque

La dérivée Γ' que l'on obtient est exactement celle qu'on aurait obtenu en dérivant (par rapport à x) sans précaution la fonction dans l'intégrale définissant Γ .

Malheureusement, nous ne disposons d'aucun théorème nous assurant que la dérivée de l'intégrale est égale à l'intégrale de la dérivée, et donc il aurait été un peu cavalier d'affirmer directement que $\Gamma' = g_1$.

Or, la concavité de $x \mapsto \ln(1+x)$ prouve, comme à la question 1.a, que $\forall x \geq 0, \ln(1+x) \leq x$.

Et donc $\ln\left(1 + \frac{1}{n+1}\right) \leq \frac{1}{n+1}$, de sorte que $\gamma_{n+1} - \gamma_n \geq 0$.

Et donc la suite $(\gamma_n)_{n \in \mathbb{N}^*}$ est décroissante, et étant minorée⁷, par le théorème de la limite monotone, elle converge.

⁷ Par 0 d'après la question précédente.

5.a. On a

$$\begin{aligned} \prod_{k=1}^n \left[\left(1 + \frac{x}{k}\right) \exp\left(-\frac{x}{k}\right) \right] &= \prod_{k=1}^n \left(1 + \frac{x}{k}\right) \times \prod_{k=1}^n \exp\left(-\frac{x}{k}\right) \\ &= \prod_{k=1}^n \frac{k+x}{k} \times \exp\left(-\sum_{k=1}^n \frac{x}{k}\right) \\ &= \frac{(x+1)(x+2) \cdots (x+n)}{n!} \exp\left(-x \sum_{k=1}^n \frac{1}{k}\right) \\ &= \frac{(x+1)(x+2) \cdots (x+n)}{n!} \exp(-x(\gamma_n + \ln(n))) \\ &= \frac{(x+1)(x+2) \cdots (x+n)}{n!} \exp(-\gamma_n x) \frac{1}{e^{x \ln n}} \\ &= \boxed{\exp(-\gamma_n x) \frac{(x+1)(x+2) \cdots (x+n)}{n^x n!}}. \end{aligned}$$

5.b. Puisque $\gamma \xrightarrow{n \rightarrow +\infty} \gamma$, alors, par continuité de la fonction exponentielle, $\exp(-\gamma_n x) \xrightarrow{n \rightarrow +\infty} \exp(-\gamma x)$.

Et d'après la question 2.c, on a $\lim_{n \rightarrow +\infty} \frac{n^x n!}{x(x+1) \cdots (x+n)} = \Gamma(x)$, de sorte que⁸

$$\lim_{n \rightarrow +\infty} \frac{(x+1)(x+2) \cdots (x+n)}{n^x n!} = \frac{1}{x\Gamma(x)}.$$

⁸ Le passage à l'inverse est légitime puisque $\Gamma(x) \neq 0$.

Et donc la suite $(v_n(x))$ est bien convergente, et sa limite vérifie $\ell(x) = \frac{\exp(-\gamma x)}{x\Gamma(x)}$.

6.a. Puisque $\frac{x}{n} \xrightarrow{n \rightarrow +\infty} 0$, on a

$$\ln\left(1 + \frac{x}{n}\right) = \frac{x}{n} - \frac{x^2}{2n^2} + o_{n \rightarrow +\infty}\left(\frac{x^2}{n^2}\right)$$

et donc

$$\frac{x}{n} - \ln\left(1 + \frac{x}{n}\right) = \frac{x}{n} - \frac{x}{n} + \frac{x^2}{2n^2} + o_{n \rightarrow +\infty}\left(\frac{x^2}{n^2}\right) \underset{n \rightarrow +\infty}{\sim} \frac{x^2}{2n^2}.$$

Et donc, d'après le critère des équivalents pour les séries à termes positifs⁹, puisque la série¹⁰ de terme général $\frac{x^2}{2n^2}$ converge, il en est de même de la série de terme général

$$\frac{x}{n} - \ln\left(1 + \frac{x}{n}\right).$$

6.b. Par continuité de la fonction logarithme, on a $\ln(\ell(x)) = \lim_{n \rightarrow +\infty} \ln(v_n(x))$. Mais

$$\ln(v_n(x)) = \ln\left(\prod_{k=1}^n \left[\left(1 + \frac{x}{k}\right) \exp\left(-\frac{x}{k}\right)\right]\right) = \sum_{k=1}^n \left[\ln\left(1 + \frac{x}{k}\right) - \frac{x}{k}\right].$$

Et donc

$$\ln(\ell(x)) = \lim_{n \rightarrow +\infty} -\sum_{k=1}^n \left[\frac{x}{k} - \ln\left(1 + \frac{x}{k}\right)\right] = \boxed{-\sum_{k=1}^{+\infty} \left[\frac{x}{k} - \ln\left(1 + \frac{x}{k}\right)\right]}.$$

D'après 5.b, $\Gamma(x) = \frac{\exp(-\gamma x)}{x\ell(x)}$, et donc par continuité du logarithme,

$$\ln(\Gamma(x)) = -\gamma x - \ln x - \ln(\ell(x)) = \boxed{-\gamma x - \ln x + \sum_{n=1}^{+\infty} \left[\frac{x}{n} - \ln\left(1 + \frac{x}{n}\right)\right]}.$$

Équivalent

Rappelons que par **définition**, on a $u_n \sim v_n$ si et seulement si

$$u_n = v_n + o(v_n).$$

⁹ $\frac{x^2}{2n^2}$ est clairement positif.

¹⁰ De Riemann.

Continuité

Rappelons que lorsqu'on affirme que $u_n \rightarrow \ell$ donc $f(u_n) \rightarrow f(\ell)$, la continuité de f en ℓ est essentielle.

Par exemple, si $f(x) = \lfloor x \rfloor$, on a

$$1 - \frac{1}{n} \rightarrow 1$$

mais

$$f\left(1 - \frac{1}{n}\right) \neq \lfloor 1 \rfloor.$$

Ceci tient au fait que f n'est pas continue en 1.

7. Notons $u : x \mapsto \ln(\Gamma(x+1))$. Alors nous savons que $u(x) = \ln(x\Gamma(x)) = \ln(x) + \ln(\Gamma(x))$.

Et donc en dérivant, $u'(x) = \frac{1}{x} + \psi(x)$.

Mais d'autre part, par dérivation d'une composée, $u'(x) = \psi(x+1)$.

Et donc on a bien $\psi(x) = \psi(x+1) - \frac{1}{x}$.

Puisque la fonction Γ est de classe \mathcal{C}^1 , ψ est une fonction continue sur \mathbf{R}_+^* .

Et en particulier, $\psi(x+1) \xrightarrow{x \rightarrow 0} \psi(1)$.

Mais $\psi(1)$ est une constante, donc $\psi(1) = \lim_{x \rightarrow 0^+} \left(\frac{1}{x} \right)$.

Et donc $\psi(x) = -\frac{1}{x} + \lim_{x \rightarrow 0^+} \left(\frac{1}{x} \right) \underset{x \rightarrow 0^+}{\sim} -\frac{1}{x}$.

Enfin, pour $n \geq 2$, on a

$$\psi(n) = \psi(n-1) + \frac{1}{n-1} = \psi(n-2) + \frac{1}{n-2} + \frac{1}{n-1} = \dots = \psi(1) + \frac{1}{1} + \frac{1}{2} + \dots + \frac{1}{n-1} = \psi(1) + \sum_{k=1}^{n-1} \frac{1}{k}.$$

\mathcal{C}^1

Nous savons que Γ est deux fois dérivable. Donc Γ' est dérivable, et donc continue, de sorte que Γ est bien de classe \mathcal{C}^1 .

8.a. Grâce à la relation établie en 6.b, on a

$$A(x) = \sum_{n=1}^{+\infty} \left[\ln \left(1 + \frac{x}{n} \right) - \frac{x}{n} \right] = -\gamma x - \ln(x) - \ln(\Gamma(x)).$$

Mais Γ étant deux fois dérivable, il en est de même de $\ln \circ \Gamma$, et donc A est deux fois dérivable car somme de fonctions deux fois dérivables.

On a alors, pour tout $x \in \mathbf{R}_+^*$,

$$A'(x) = -\gamma - \frac{1}{x} - \frac{\Gamma'(x)}{\Gamma(x)}$$

$$A''(x) = \frac{1}{x^2} - \frac{\Gamma''(x)\Gamma(x) - \Gamma'(x)^2}{\Gamma(x)^2}.$$

8.b. Pour $k = 1$, on a

$$U_n^{(k)}(x) = U_n'(x) = \frac{1}{n} \left(\frac{1}{1 + \frac{x}{n}} - 1 \right) = \frac{1}{n} \left(\frac{n}{n+x} - 1 \right) = \frac{-x}{n(n+x)} \underset{n \rightarrow +\infty}{\sim} -\frac{x}{n^2}.$$

Et donc $|U_n'(x)| \underset{n \rightarrow +\infty}{\sim} \frac{x}{n^2}$, qui converge par comparaison à une série de Riemann convergente.

On a ensuite, $U_n''(x) = -\frac{1}{(n+x)^2}$, $U_n^{(3)}(x) = \frac{2}{(x+n)^3}$, $U_n^{(4)} = -\frac{2 \times 3}{(x+n)^4}$, et plus généralement, une récurrence rapide prouve que

$$\forall k \geq 2, U_n^{(k)}(x) = (-1)^{k-1} \frac{(k-1)!}{(n+x)^k}.$$

En particulier, $|U_n^{(k)}(x)| \underset{n \rightarrow +\infty}{\sim} \frac{(k-1)!}{n^k}$.

Et donc $\sum_{n \geq 1} U_n^{(k)}(x)$ converge absolument, par comparaison à une série de Riemann convergente.

8.c. La relation de la question 6.b s'écrit encore

$$A(x) = -\gamma x - \ln(x) - \ln(\Gamma(x)).$$

Ce qui en dérivant donne

$$A'(x) = -\gamma - \frac{1}{x} - \psi(x).$$

Et donc en particulier pour $x = 1$, $A'(1) = -\gamma - 1 - \psi(1)$.
Mais d'autre part, grâce au résultat admis,

$$A'(1) = \sum_{n=1}^{+\infty} U'_n(1) = - \sum_{n=1}^{+\infty} \frac{1}{n(n+1)}.$$

Or, il est classique que pour tout $n \in \mathbf{N}^*$, $\frac{1}{n(n+1)} = \frac{1}{n} - \frac{1}{n+1}$. Et alors, pour $N \in \mathbf{N}^*$

$$\sum_{n=1}^N \frac{1}{n(n+1)} = \sum_{n=1}^N \frac{1}{n} - \sum_{n=1}^N \frac{1}{n+1} = \sum_{n=1}^N \frac{1}{n} - \sum_{n=2}^{N+1} \frac{1}{n} = 1 - \frac{1}{N+1} \xrightarrow{N \rightarrow +\infty} 1.$$

Et donc $A'(1) = - \sum_{n=1}^{+\infty} \frac{1}{n(n+1)} = -1$.

Il vient donc

$$\psi(1) = -\gamma - 1 + 1 = \boxed{-\gamma}.$$

On a alors,

$$\ln(n) - \psi(n) = \sum_{k=1}^n \frac{1}{k} - \gamma_n - \left(\psi(1) - \sum_{k=1}^{n-1} \frac{1}{k} \right) = \gamma - \gamma_n + \frac{1}{n} \xrightarrow{n \rightarrow +\infty} \boxed{0}.$$

9.a. Il est évident que G est positive¹¹ sur \mathbf{R}_+ , et G est dérivable, avec $G'(t) = -\frac{2}{(x+t)^3} < 0$, donc G est strictement décroissante sur \mathbf{R}_+ .

De plus, G est continue sur $[1, +\infty[$, et au voisinage de $+\infty$, $G(t) \underset{t \rightarrow +\infty}{\sim} \frac{1}{t^2}$.

Puisque l'intégrale $\int_1^{+\infty} G(t) dt$ converge, il en est de même¹² de $\int_1^{+\infty} G(t) dt$.

9.b. On a, pour tout $k \in \mathbf{N}^*$ et tout $t \in [k, k+1]$, $G(k+1) \leq G(t) \leq G(k)$, et donc par croissance de l'intégrale,

$$G(k+1) = \int_k^{k+1} G(k+1) dt \leq \int_k^{k+1} G(t) dt \leq \int_k^{k+1} G(k) dt = G(k).$$

En sommant ces relations pour k allant de 1 à $n-1$, il vient, par la relation de Chasles,

$$\sum_{k=1}^{n-1} G(k+1) \leq \int_1^n G(t) dt \leq \sum_{k=1}^{n-1} G(k).$$

En particulier, en réalisant un changement d'indice $i = k+1$ dans la première somme, il vient, pour tout $n \in \mathbf{N}^*$,

$$\sum_{i=2}^n G(i) \leq \int_1^n G(t) dt.$$

Mais la fonction G étant positive, $\int_1^n G(t) dt \leq \int_1^n G(t) dt + \underbrace{\int_n^{+\infty} G(t) dt}_{\geq 0} \leq \int_1^{+\infty} G(t) dt$.

On en déduit que la suite des sommes partielles de la série de terme général $G(k)$ sont majorées par $\int_1^{+\infty} G(t) dt$.

Mais puisqu'il s'agit d'une série à termes positifs, cette suite des sommes partielles est croissante, et étant majorée, elle converge donc.

Autrement dit, la série de terme général $G(k)$ converge.

Et donc en faisant tendre n vers $+\infty$ dans l'inégalité précédente, il vient donc

$$\int_1^{+\infty} G(t) dt \leq \sum_{k=1}^{+\infty} G(k).$$

Remarque

Notons que les résultats admis nous expliquent qu'ici (et ici seulement !) on a le droit de dire que la dérivée de la somme infinie est égale à la somme infinie des dérivées. Ce qui est un problème difficile en général.

¹¹ Et même strictement positive !

¹² Il s'agit de fonctions de signe constant.

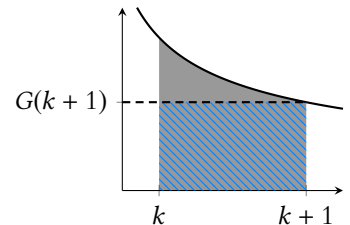


FIGURE 24.2— L'intégrale, qui est l'aire de la partie colorée est plus grande que l'aire de la partie hachurée, qui vaut $G(k+1)$ (c'est l'aire d'un rectangle de largeur 1 et de hauteur $G(k+1)$).

Majoration

Notons que nous avons bien majoré par une constante, ce que n'était pas $\int_1^n G(t) dt$.

Il s'agit presque du résultat voulu, sauf qu'il s'agit d'une inégalité large et non stricte. Nous pourrions refaire le même calcul en commençant les sommes¹³ à 2. On obtient alors

¹³ Et l'intégrale.

$$\int_2^{+\infty} G(t) dt \leq \sum_{k=2}^{+\infty} G(k).$$

Il reste donc à prouver que $\int_1^2 G(t) dt \leq G(1)$.

Or, par stricte décroissance de G , on a, pour tout $t \in]1, 2]$,

$$G(t) < G(1) \Leftrightarrow G(1) - G(t) > 0.$$

Et donc, la fonction $t \mapsto G(1) - G(t)$ étant continue, positive et non nulle sur $]1, 2]$,

$$\int_1^2 (G(1) - G(t)) dt > 0 \Leftrightarrow \int_1^2 G(t) dt > \int_1^2 G(1) dt = G(1).$$

Et donc on a bien

$$\int_1^{+\infty} G(t) dt = \int_1^2 G(t) dt + \int_2^{+\infty} G(t) dt < G(1) + \sum_{k=2}^{+\infty} G(k) = \sum_{k=1}^{+\infty} G(k).$$

Enfin, la fonction G est continue, positive et non nulle sur $[1, +\infty[$, donc $\int_1^{+\infty} G(t) dt > 0$.

Et donc on a bien prouvé que

$$0 < \int_1^{+\infty} G(t) dt < \sum_{k=1}^{+\infty} G(k).$$

9.c. Puisque $A'(x) = -\gamma - \frac{1}{x} - \psi(x)$, il vient

$$A''(x) = \frac{1}{x^2} - \psi'(x) \Leftrightarrow \psi'(x) = \frac{1}{x^2} - A''(x) = \frac{1}{x^2} - \sum_{k=1}^{+\infty} U_k''(x).$$

Mais $U_k''(x) = -\frac{1}{(k+x)^2} = -G(k)$.

Et donc $\psi'(x) = \frac{1}{x^2} + \sum_{k=1}^{+\infty} G(k) > \frac{1}{x^2} + \int_1^{+\infty} G(t) dt$.

Or, pour $A > 1$,

$$\int_1^A G(t) dt = \int_1^A \frac{dt}{(x+t)^2} = \left[-\frac{1}{x+t} \right]_1^A = \frac{1}{1+x} - \frac{1}{1+A} \xrightarrow{A \rightarrow +\infty} \frac{1}{1+x}.$$

Et donc on a bien $\psi'(x) > \frac{1}{x^2} + \frac{1}{x+1}$.

Enfin, on a, pour tout $x > 0$,

$$\frac{1}{x+1} + \frac{1}{x^2} = \frac{x^2 + x + 1}{x(x+1)} > \frac{x+1}{x(x+1)} = \frac{1}{x}$$

de sorte que $\forall x \in \mathbf{R}_+^*, \psi'(x) > \frac{1}{x}$.

Partie III. Estimation des paramètres d'une loi $\Gamma(\theta, r)$.

10. La fonction f est positive¹⁴ sur \mathbf{R} , et elle est continue sauf éventuellement en 0.

À l'aide du changement de variable affine $t = \frac{x}{\theta}$, il vient, sous réserve de convergence,

¹⁴ Car Γ ne prend que des valeurs positives par positivité de l'intégrale.

$$\int_{-\infty}^{+\infty} f(x) dx = \int_0^{+\infty} \frac{1}{\Gamma(r)\theta^r} x^{r-1} e^{-\frac{x}{\theta}} dx$$

$$\begin{aligned}
&= \int_0^{+\infty} \frac{1}{\Gamma(r)\theta^r} (\theta t)^{r-1} e^{-t} \theta dt \\
&= \frac{\theta^r}{\Gamma(r)\theta^r} \int_0^{+\infty} t^{r-1} e^{-t} dt \\
&= \frac{1}{\Gamma(r)} \Gamma(r) = 1.
\end{aligned}$$

Et donc f est bien une densité de probabilités.

11. L possède un maximum en (θ_0, r_0) si et seulement si

$$\forall (\theta, r) \in \mathbf{R}_+^* \times \mathbf{R}_+^*, L(\theta, r) \leq L(\theta_0, r_0).$$

Mais, par croissance de la fonction \ln , ceci est équivalent à

$$\forall (\theta, r) \in \mathbf{R}_+^* \times \mathbf{R}_+^*, \ln(L(\theta, r)) \leq \ln(L(\theta_0, r_0)).$$

Autrement dit, L présente un maximum en (θ_0, r_0) si et seulement si F possède un maximum en (θ_0, r_0) .

- 12.a. Notons que, les x_i étant strictement positifs,

$$F(\theta, r) = \sum_{i=1}^p \ln(f(x_i)) = \sum_{i=1}^p \ln\left(\frac{1}{\Gamma(r)\theta^r} x_i^{r-1} \exp\left(-\frac{x_i}{\theta}\right)\right) = \sum_{i=1}^p \left[-\ln(\Gamma(r)) - r \ln(\theta) + (r-1) \ln(x_i) - \frac{x_i}{\theta}\right].$$

Or, la fonction $(\theta, r) \mapsto r$ est classe \mathcal{C}^2 sur $\mathbf{R}_+^* \times \mathbf{R}_+^*$ car elle est polynomiale, donc par composition avec la fonction Γ qui est¹⁵ de classe \mathcal{C}^2 sur \mathbf{R}_+^* , à valeurs dans \mathbf{R}_+^* , puis avec la fonction logarithme, \mathcal{C}^2 sur \mathbf{R}_+^* , $(\theta, r) \mapsto -\ln(\Gamma(r))$ est \mathcal{C}^2 sur $\mathbf{R}_+^* \times \mathbf{R}_+^*$.

De la même manière, $(\theta, r) \mapsto \theta$ et $(\theta, r) \mapsto r$ sont \mathcal{C}^2 car polynomiales, donc par composition par la fonction logarithme, puis par produit, $(\theta, r) \mapsto r \ln(\theta)$ est \mathcal{C}^2 .

Et on montre de même que $(\theta, r) \mapsto (r-1) \ln(x_i) - \frac{x_i}{\theta}$ sont \mathcal{C}^2 , de sorte que F est \mathcal{C}^2 par somme de fonctions \mathcal{C}^2 .

On a alors, pour tout $(\theta, r) \in \mathbf{R}_+^* \times \mathbf{R}_+^*$,

$$\partial_1 F(\theta, r) = \sum_{i=1}^p \left[-\frac{r}{\theta} + \frac{x_i}{\theta^2}\right] = -\frac{pr}{\theta} + \frac{1}{\theta^2} \sum_{i=1}^p x_i.$$

$$\partial_2 F(\theta, r) = \sum_{i=1}^p [-\psi(r) - \ln(\theta) + \ln(x_i)] = \sum_{i=1}^p \ln(x_i) - p(\psi(r) + \ln(\theta)).$$

Et les dérivées d'ordre 2 sont données par

$$\partial_{1,1}^2 F(\theta, r) = \frac{pr}{\theta^2} - \frac{2}{\theta^3} \sum_{i=1}^p x_i \quad \partial_{2,2}^2 F(\theta, r) = -p\psi'(r) \quad \partial_{1,2}^2 F(\theta, r) = \partial_{2,1}^2 F(\theta, r) = -\frac{p}{\theta}.$$

- 12.b. Un couple $(\theta^*, r^*) \in \mathbf{R}_+^* \times \mathbf{R}_+^*$ est un point critique de F si et seulement si

$$\begin{aligned}
\begin{cases} -\frac{pr^*}{\theta^*} + \frac{1}{(\theta^*)^2} \sum_{i=1}^p x_i = 0 \\ \sum_{i=1}^p \ln(x_i) - p(\psi(r^*) + \ln(\theta^*)) = 0 \end{cases} &\Leftrightarrow \begin{cases} \frac{1}{(\theta^*)^2} (p\bar{x} - \theta^* pr^*) = 0 \\ \frac{1}{p} \sum_{i=1}^p \ln(x_i) = \psi(r^*) + \ln(\theta^*) \end{cases} &\Leftrightarrow \begin{cases} \theta^* r^* = \bar{x} \\ \frac{1}{p} \sum_{i=1}^p \ln(x_i) = \frac{\Gamma'(r^*)}{\Gamma(r^*)} + \ln\left(\frac{\bar{x}}{r^*}\right) \end{cases} \\
&\Leftrightarrow \begin{cases} \theta^* r^* = \bar{x} \\ \frac{1}{p} \sum_{i=1}^p \ln(x_i) - \frac{\Gamma'(r^*)}{\Gamma(r^*)} - \ln(r^*) = \ln(\bar{x}) \end{cases}
\end{aligned}$$

- 13.a. Soit $\varphi : x \mapsto x - 1 - \ln(x)$.

Alors φ est dérivable sur \mathbf{R}_+^* , et $\varphi'(x) = 1 - \frac{1}{x} = \frac{x-1}{x}$.

Convergence

On reconnaît ici une intégrale de référence convergente, donc par le théorème de changement de variable, l'intégrale de départ converge également.

¹⁵ On a admis que Γ est deux fois dérivable et que $\Gamma'' = g_1$. Nous admettrons également que g_1 est continue, de sorte que Γ est \mathcal{C}^2 .

Donc φ' est strictement négative sur $]0, 1[$, de sorte que φ est strictement décroissante sur $]0, 1[$, avec $\varphi(1) = 0$.

Et donc pour tout $x \in]0, 1[$, $\varphi(x) > 0 \Leftrightarrow \ln(x) < x - 1$.

Sur $]1, +\infty[$, φ' est strictement positive, donc φ est strictement croissante sur $]1, +\infty[$, et donc pour tout $x \in]1, +\infty[$, $\varphi(x) > \varphi(1) = 0 \Leftrightarrow \ln(x) < x - 1$.

Par conséquent, on a

$$\begin{aligned} K_p &= \ln \bar{x} - \frac{1}{p} \sum_{i=1}^p \ln x_i = \frac{1}{p} \sum_{i=1}^p [\ln \bar{x} - \ln x_i] = -\frac{1}{p} \sum_{i=1}^p \ln \frac{\bar{x}}{x_i} \\ &\geq -\frac{1}{p} \sum_{i=1}^p \left[\frac{x_i}{\bar{x}} - 1 \right] = \frac{1}{\bar{x}} \underbrace{\frac{1}{p} \sum_{i=1}^p x_i}_{=\bar{x}} - 1 = 0. \end{aligned}$$

De plus, les x_i ne sont pas tous égaux, et en particulier ne sont pas tous égaux à \bar{x} .

Et donc pour au moins l'un des $i \in \llbracket 1, p \rrbracket$, $\frac{x_i}{\bar{x}} \neq 1$, de sorte que $-\ln \frac{x_i}{\bar{x}} > \frac{x_i}{\bar{x}} - 1$.

Et donc on a

$$K_p > -\sum_{i=1}^p \left[\frac{x_i}{\bar{x}} - 1 \right] = 0.$$

13.b. On a donc $h(y) = \ln(y) - \psi(y) - K_p$.

Puisque ψ est dérivable, il en est de même de h , et $h'(y) = \frac{1}{y} - \psi'(y) < 0$.

On en déduit donc que h est strictement décroissante sur \mathbf{R}_+^* .

D'autre part, nous savons qu'au voisinage de 0, $\psi(y) \underset{y \rightarrow 0^+}{\sim} -\frac{1}{y}$.

Et donc $\ln(y) = \underset{y \rightarrow 0^+}{o}(\psi(y))$, de sorte que

$$\ln(y) - \psi(y) \underset{y \rightarrow 0^+}{\sim} -\psi(y) \underset{y \rightarrow 0^+}{\sim} \frac{1}{y}.$$

On en déduit en particulier que $\lim_{y \rightarrow 0^+} h(y) = +\infty$.

D'autre part, h étant décroissante, elle admet une limite¹⁶ en $+\infty$.

Et alors

$$\lim_{y \rightarrow +\infty} h(y) = \lim_{n \rightarrow +\infty} h(n) = \lim_{n \rightarrow +\infty} (\ln n - \psi(n)) - K_p = -k_p.$$

Le tableau de variation de h est donc donné par :

x	0	$+\infty$
$h'(x)$		-
$h(x)$	0	$-K_p$

13.c. Par le théorème de la bijection¹⁷, h réalise une bijection de \mathbf{R}_+^* sur $] -K_p, +\infty[$. Et donc en particulier il existe un unique $r^* \in \mathbf{R}_+^*$ tel que $h(r^*) = 0 \Leftrightarrow \ln r^* - \frac{\Gamma(r^*)}{\Gamma'(r^*)} = K_p$.

Et alors, $\left(\frac{\bar{x}}{r^*}, r^* \right)$ est l'unique solution du système (S).

14. D'après les calculs réalisés précédemment, on a

$$\partial_{1,1}^2 F \left(\frac{\bar{x}}{r^*}, r^* \right) = \frac{pr^*}{\left(\frac{\bar{x}}{r^*} \right)^2} - \frac{2p\bar{x}}{\left(\frac{\bar{x}}{r^*} \right)^3} = -\frac{p\bar{x}}{\left(\frac{\bar{x}}{r^*} \right)^3} = -\frac{p(r^*)^3}{\bar{x}^2}, \quad \partial_{2,2}^2 F \left(\frac{\bar{x}}{r^*}, r^* \right) = -p\psi'(r^*).$$

Remarque

Un argument de concavité de la fonction \ln nous permettrait de prouver l'inégalité large, mais pour obtenir une inégalité stricte, il faut passer par l'étude de fonction.

Détails

C'est le résultat de la question 9.

Détails

Ce résultat a été établi à la question 7.

¹⁶ Finie ou égale à $-\infty$.

Détails

La limite de $\psi(n) - \ln(n)$ a été calculée en 8.c.

¹⁷ Qui s'applique car h est continue et strictement décroissante.

Et de même, $\partial_{1,2}^2 F(\theta^*, r^*) = \partial_{2,1}^2 F(\theta^*, r^*) = -\frac{p\bar{x}}{r^*}$.

On a donc $\text{tr}(\nabla^2 F(\theta^*, r^*)) = \partial_{1,1}^2 F(\theta^*, r^*) + \partial_{2,2}^2 F(\theta^*, r^*) < 0$ et

$$\det(\nabla^2 F(\theta^*, r^*)) = \frac{p^2}{(\theta^*)^2} \left(\frac{\bar{x}}{\theta^*} \psi'(r^*) - 1 \right) = \frac{p^2}{(\theta^*)^2} (r^* \psi'(r^*) - 1).$$

Mais d'après la question 9, on a $\psi'(r^*) > \frac{1}{r^*}$ et donc $\det(\nabla^2 F(\theta^*, r^*)) > 0$.

Nous savons que les valeurs propres de $\det(\nabla^2 F(\theta^*, r^*))$ sont les racines de

$$P(X) = X^2 - \text{tr}(\nabla^2 F(\theta^*, r^*))X + \det(\nabla^2 F(\theta^*, r^*)).$$

Notons λ_1, λ_2 ces deux valeurs propres, de sorte que

$$P(X) = (X - \lambda_1)(X - \lambda_2) = X^2 - (\lambda_1 + \lambda_2)X + \lambda_1\lambda_2.$$

Par identification des coefficients de P , il vient donc

$$\det(\nabla^2 F(\theta^*, r^*)) = \lambda_1\lambda_2 \text{ et } \text{tr}(\nabla^2 F(\theta^*, r^*)) = \lambda_1 + \lambda_2.$$

On en déduit donc que $\lambda_1\lambda_2 > 0$, de sorte que λ_1 et λ_2 sont non nulles, et de même signe. Et puisque $\lambda_1 + \lambda_2 > 0$, toutes deux sont positives.

Et donc F possède un minimum local en (θ^*, r^*) .

Détails

Il y a bien deux valeurs propres distinctes puisque la matrice hessienne est diagonalisable, et ne peut avoir une seule valeur propre puisqu'elle n'est pas diagonale.

Partie IV. Estimateur sans biais de l'écart-type σ d'une loi normale.

15. Il est évident que T est à valeurs dans \mathbf{R}_+ .

Pour $x \geq 0$, on a

$$P(T \leq x) = P\left(\frac{X^2}{2\sigma^2} \leq x\right) = P(X^2 \leq 2\sigma^2 x) = P(-\sigma\sqrt{2x} \leq X \leq \sigma\sqrt{2x}) = P\left(-\sqrt{2x} \leq \frac{X}{\sigma} \leq \sqrt{2x}\right).$$

Mais $\frac{X}{\sigma} \leftrightarrow \mathcal{N}(0, 1)$, et donc

$$P(T \leq x) = \Phi(\sqrt{2x}) - \Phi(-\sqrt{2x}) = 2\Phi(\sqrt{2x}) - 1.$$

Donc la fonction de répartition de T est $F_T : x \mapsto \begin{cases} 2\Phi(\sqrt{2x}) - 1 & \text{si } x \geq 0 \\ 0 & \text{si } x < 0 \end{cases}$

Puisque Φ est \mathcal{C}^1 , F_T est \mathcal{C}^1 sur \mathbf{R}^* .

D'autre part, on a

$$\lim_{x \rightarrow 0^+} F_T(x) = 2\Phi(0) - 1 = 2 \cdot \frac{1}{2} - 1 = 0 = \lim_{x \rightarrow 0^-} F_T(x) = F_T(0).$$

Donc F_T est continue en 0 et donc sur \mathbf{R} .

Ainsi, T est une variable à densité, dont une densité est donnée par

$$f_T : x \mapsto \begin{cases} \frac{2}{\sqrt{2x}} \Phi'(\sqrt{2x}) & \text{si } x > 0 \\ 0 & \text{si } x \leq 0 \end{cases} = \begin{cases} \frac{2}{\sqrt{2x}} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}2x} & \text{si } x > 0 \\ 0 & \text{sinon} \end{cases} = \begin{cases} \frac{1}{\sqrt{\pi}} x^{1/2-1} e^{-x} & \text{si } x > 0 \\ 0 & \text{sinon} \end{cases}$$

Puisqu'il s'agit d'une densité, on a en particulier

$$\int_{-\infty}^{+\infty} f_T(x) dx = 1 \Leftrightarrow \frac{1}{\sqrt{\pi}} \int_0^{+\infty} x^{1/2-1} e^{-x} dx = 1 \Leftrightarrow \Gamma(1/2) = \sqrt{\pi}.$$

$$\text{Et donc } f_T(x) = \begin{cases} \frac{1}{\Gamma(1/2)} x^{1/2-1} e^{-x} & \text{si } x > 0 \\ 0 & \text{si } x \leq 0 \end{cases}$$

On reconnaît là une densité de la loi $\gamma(1/2)$, de sorte que $T \leftrightarrow \gamma(1/2)$.

\mathcal{C}^1

Φ est \mathcal{C}^1 sur \mathbf{R} tout entier car on sait qu'il existe une densité de la loi normale qui est continue sur \mathbf{R} tout entier.

16.a. Puisque les X_i sont mutuellement indépendantes, il en est de même des $\frac{X_i^2}{2\sigma^2}$.
 Mais nous venons de prouver que ces variables suivent toutes la loi $\gamma(1/2)$.
 Et donc par stabilité des lois γ , $S_n \leftrightarrow \gamma(n/2)$.

16.b. En particulier, on a $E(S_n) = \frac{n}{2}$, et donc par linéarité de l'espérance,

$$E(Y_n) = E\left(\frac{2\sigma^2}{n}S_n\right) = \frac{2\sigma^2}{n}E(S_n) = \sigma^2.$$

Puisque Y_n est une fonction¹⁸ des n variables aléatoires X_1, \dots, X_n , c'est bien un estimateur de σ^2 , qui est donc sans biais.

¹⁸ Ne dépendant pas de σ !

16.c. Puisque $(\sqrt{Y_n})^2 = Y_n$, et que Y_n possède une espérance, $\sqrt{Y_n}$ possède un moment d'ordre 2, et donc une variance.

De plus, cette variance est non nulle car Y_n est une variable à densité, et donc $\sqrt{Y_n}$ n'est pas une variable certaine.

On a alors $V(\sqrt{Y_n}) = E(Y_n) - E(\sqrt{Y_n})^2 = \sigma^2 - E(\sqrt{Y_n})^2$.

Et donc¹⁹ $V(\sqrt{Y_n}) > 0 \Leftrightarrow E(\sqrt{Y_n}) < \sigma$.

Variance nulle
 Rappelons que seules les variables (presque) certaines possèdent une variance nulle.

¹⁹ $E(\sqrt{Y_n}) \geq 0$ puisque $\sqrt{Y_n}$ est à valeurs positives.

16.d. Par linéarité de l'espérance, $E(\sqrt{Y_n}) = E\left(\sqrt{\frac{2\sigma^2}{n}}\sqrt{S_n}\right) = \sigma\sqrt{\frac{2}{n}}E(\sqrt{S_n})$.

Mais par le théorème de transfert,

$$E(\sqrt{S_n}) = \int_0^{+\infty} \sqrt{x} \frac{1}{\Gamma(n/2)} x^{n/2-1} e^{-x} dx = \frac{1}{\Gamma(n/2)} \int_0^{+\infty} x^{\frac{n+1}{2}-1} e^{-x} dx = \frac{\Gamma(\frac{n+1}{2})}{\Gamma(\frac{n}{2})} = \frac{1}{\lambda_n}.$$

Et donc $E(\sqrt{Y_n}) = \sqrt{\frac{2}{n}} \frac{\sigma}{\lambda_n}$.

16.e. D'après les calculs effectués précédemment, on a

$$E(\widehat{\sigma}_n) = \frac{\lambda_n}{\sqrt{2}} E\left(\left(\sum_{i=1}^n X_i^2\right)^{1/2}\right) = \frac{\lambda_n}{\sqrt{2}} \sigma \sqrt{2} E(\sqrt{S_n}) = \sigma \lambda_n \underbrace{\frac{\Gamma(\frac{n+1}{2})}{\Gamma(\frac{n}{2})}}_{=1/\lambda_n} = \sigma.$$

Et donc $\widehat{\sigma}_n$ est un estimateur sans biais de σ .

17.a. On a $\widehat{\sigma}_n^2 = \frac{\lambda_n^2}{2} \left(\sum_{i=1}^n X_i^2\right)^{1/2} = \frac{\lambda_n^2}{2} 2\sigma^2 S_n$.

Et donc par linéarité de l'espérance,

$$E(\widehat{\sigma}_n^2) = \lambda_n^2 \sigma^2 E(S_n) = \lambda_n^2 \sigma^2 \frac{n}{2}.$$

Par la formule de Huygens, on a alors

$$V(\widehat{\sigma}_n) = E(\widehat{\sigma}_n^2) - E(\widehat{\sigma}_n)^2 = \lambda_n^2 \sigma^2 \frac{n}{2} - \sigma^2 = \sigma^2 \left(\lambda_n^2 \frac{n}{2} - 1\right).$$

17.b. Soit $\varepsilon > 0$. Alors, par l'inégalité de Bienaymé-Tchebychev²⁰,

²⁰ Qui s'applique car $\widehat{\sigma}_n$ possède une variance.

$$P\left(\left|\widehat{\sigma}_n - \underbrace{E(\widehat{\sigma}_n)}_{=\sigma}\right| \geq \varepsilon\right) \leq \frac{V(\widehat{\sigma}_n)}{\varepsilon^2}.$$

Mais $\lambda_n \underset{n \rightarrow +\infty}{\sim} \sqrt{\frac{2}{n}}$, de sorte que $\lambda_n^2 \underset{n \rightarrow +\infty}{\sim} \frac{2}{n}$ et donc $\lambda_n^2 \frac{n}{2} \underset{n \rightarrow +\infty}{\sim} 1 \rightarrow 1$.

Et donc $V(\widehat{\sigma}_n) \underset{n \rightarrow +\infty}{\rightarrow} 0$.

Par le théorème des gendarmes²¹, on a donc $\lim_{n \rightarrow +\infty} P(|\widehat{\sigma}_n - \sigma| \geq \varepsilon) \underset{n \rightarrow +\infty}{\rightarrow} 0$, et donc

$$\widehat{\sigma}_n \xrightarrow{P} \sigma.$$

²¹ Une probabilité est positive.

DEVOIR SURVEILLÉ 8 : MATHS II (23.03.19)

Dans tout le problème :

- toutes les variables aléatoires introduites sont supposées définies sur le même espace probabilisé (Ω, \mathcal{A}, P) ;
- on note θ un paramètre réel.

Partie I. Une démonstration probabiliste de la formule de Stirling.

Pour tout $n \in \mathbf{N}^*$, soit h_n la fonction définie par : $\forall x \in [0, 1], h_n(x) = ((1-x)e^x)^n$.

Pour tout $n \in \mathbf{N}^*$, on pose : $I_n = \int_0^1 h_n(x) dx$.

1. a. À l'aide du changement de variable $u = n(1-x)$, montrer que : $\forall n \in \mathbf{N}^*, I_n = \frac{e^n}{n^{n+1}} \int_0^n u^n e^{-u} du$.
 b. Montrer que pour tout $x \in [0, 1]$, on a : $x + \ln(1-x) \leq -\frac{x^2}{2}$.
 c. En se référant à une densité de la loi normale centrée réduite, en déduire que : $\forall n \in \mathbf{N}^*, 0 \leq I_n \leq \sqrt{\frac{\pi}{2n}}$.
2. On note h_n^* la restriction à l'intervalle $]0, 1[$ de la fonction h_n .
 On pose pour tout $x \in]0, 1[$, $h_n^*(x) = \exp\left(-\frac{nx^2}{2}H(x)\right)$ et $g(x) = (1-x)\ln(1-x) + x - \frac{x^2}{2}$.
 a. Montrer que H est prolongeable par continuité en 0. On note encore H la fonction ainsi prolongée.
 b. Montrer que la fonction g est convexe et strictement positive sur $]0, 1[$.
 c. En déduire que la fonction H réalise une bijection strictement croissante de $[0, 1[$ sur $[1, +\infty[$.
3. Soit $(u_n)_{n \in \mathbf{N}^*}$ une suite convergente de limite nulle telle que : $\lim_{n \rightarrow +\infty} u_n \sqrt{n} = +\infty$ et $\forall n \in \mathbf{N}^*, 0 < u_n < 1$.
 a. Donner un exemple d'une telle suite $(u_n)_{n \in \mathbf{N}^*}$.
 b. Soit $(v_n)_{n \in \mathbf{N}^*}$ la suite définie par : $\forall n \in \mathbf{N}^*, v_n = H(u_n)$. Montrer que la suite $(v_n)_{n \in \mathbf{N}^*}$ est convergente et préciser sa limite.
 c. Établir pour tout $n \in \mathbf{N}^*$, l'encadrement : $I_n \geq \int_0^{u_n} h_n(x) dx \geq \frac{1}{\sqrt{nv_n}} \int_0^{u_n \sqrt{nv_n}} \exp\left(-\frac{y^2}{2}\right) dy$.
 d. Déduire des questions 1.c et 3.c un équivalent de I_n lorsque n tend vers $+\infty$.
4. Soit $(T_n)_{n \in \mathbf{N}^*}$ une suite de variables aléatoires définies sur (Ω, \mathcal{A}, P) , mutuellement indépendantes et de même loi exponentielle de paramètre 1. Pour tout $n \in \mathbf{N}^*$, on pose : $S_n = T_1 + T_2 + \dots + T_n$.
 a. Rappeler la loi suivie par la variable aléatoire S_n et montrer que $\lim_{n \rightarrow +\infty} P(S_n \leq n) = \frac{1}{2}$.
 b. Pour tout $n \in \mathbf{N}^*$, on pose : $U_n = \frac{T_{n+1}}{\sqrt{n}}$. Montrer que la suite $(U_n)_{n \in \mathbf{N}^*}$ converge en probabilités vers la constante 0.
 c. En déduire que $\lim_{n \rightarrow +\infty} P([S_{n+1} \leq n]) = \frac{1}{2}$.
5. Montrer que $n! \underset{n \rightarrow +\infty}{\sim} n^n e^{-n} \sqrt{2\pi n}$ (formule de Stirling).

Partie II. Quelques propriétés de la loi de Cauchy.

6. On rappelle que la fonction Arctan est la fonction réciproque de la restriction à l'intervalle ouvert $\left]-\frac{\pi}{2}, \frac{\pi}{2}\right[$ de la fonction tan, qu'elle est de classe \mathcal{C}^∞ sur \mathbf{R} et admet pour dérivée la fonction $x \mapsto \frac{1}{1+x^2}$ pour tout $x \in \mathbf{R}$ et qu'elle réalise une bijection de \mathbf{R} sur $\left]-\frac{\pi}{2}, \frac{\pi}{2}\right[$.
 a. Montrer que la fonction $\text{Arctan}(x)$ est impaire.
 b. Justifier l'existence d'un développement limité à l'ordre 3 de la fonction Arctan en 0 et le déterminer.
 c. Établir pour tout $x \in \mathbf{R}_+$, l'encadrement : $0 \leq \text{Arctan}(x) \leq x$.
 d. Montrer que pour tout $x \in \mathbf{R}_+^*$, on a : $\text{Arctan}(x) + \text{Arctan}\left(\frac{1}{x}\right) = \frac{\pi}{2}$.

7. a. Montrer que la fonction $x \mapsto \frac{1}{\pi} \times \frac{1}{1 + (x - \theta)^2}$ est une densité de probabilité sur \mathbf{R} .
 Dans toute la suite du problème, on note X une variable aléatoire à valeurs réelles, de densité f_X telle que :

$$\forall x \in \mathbf{R}, f_X(x) = \frac{1}{\pi} \times \frac{1}{1 + (x - \theta)^2}.$$

On dit que X suit une loi de Cauchy de paramètre θ et on note : $X \hookrightarrow \mathcal{C}_\theta$.

- b. La variable aléatoire X admet-elle une espérance ?
 c. Pour $\theta = 0$, tracer la courbe représentative de f_X dans le plan rapporté à un repère orthonormal.
 8. a. On note F_X la fonction de répartition de X . Pour tout $x \in \mathbf{R}$, calculer $F_X(x)$.
 b. Montrer que l'équation $F_X(x) = \frac{1}{2}$, d'inconnue x , admet une unique solution que l'on déterminera.
 Cette solution est la médiane théorique de X .

Partie III. La loi de la moyenne empirique.

9. Pour tout $n \in \mathbf{N}^*$, et pour tout $x \in \mathbf{R}$ soit $\varphi_{n,x}$ la fonction définie sur \mathbf{R} par :

$$\forall t \in \mathbf{R}, \varphi_{n,x}(t) = \frac{1}{(1 + t^2)(1 + (x - nt)^2)}.$$

On admet l'existence d'un unique quadruplet $(\alpha, \beta, \gamma, \delta)$ de réels, indépendants de t pour lesquels on a :

$$\forall t \in \mathbf{R}, \varphi_{n,x}(t) = \frac{\alpha t + \beta}{1 + t^2} + \frac{\gamma t + \delta}{1 + (x - nt)^2}.$$

Pour tout $n \in \mathbf{N}^*$ et pour tout $x \in \mathbf{R}$, on pose : $\sigma_{n,x} = (x^2 + (n + 1)^2)(x^2 + (n - 1)^2)$.

On admet sans démonstration que : $\alpha = \frac{2nx}{\sigma_{n,x}}$, $\beta = \frac{1 + x^2 - n^2}{\sigma_{n,x}}$, $\gamma = -\frac{2n^3x}{\sigma_{n,x}}$, $\delta = \frac{n^2(3x^2 + n^2 - 1)}{\sigma_{n,x}}$.

- a. Établir la convergence de l'intégrale $\int_{-\infty}^{+\infty} \varphi_{n,x}(t) dt$.
 b. À l'aide d'une primitive de la fonction $\psi_{n,x} : t \mapsto \frac{2t}{1 + t^2} - \frac{2n(nt - x)}{1 + (x - nt)^2}$, montrer que $\int_{-\infty}^{+\infty} \psi_{n,x}(t) dt = 0$.
 c. Établir la relation : $\forall n \in \mathbf{N}^*, \forall x \in \mathbf{R}, \int_{-\infty}^{+\infty} \varphi_{n,x}(t) dt = \frac{(n + 1)\pi}{x^2 + (n + 1)^2}$.
 10. On pose $Y = X - \theta$. Pour n entier de \mathbf{N}^* , soit (Y_1, Y_2, \dots, Y_n) un n -échantillon de variables aléatoires indépendantes et de même loi que Y .

Pour tout $n \in \mathbf{N}^*$, on pose : $S_n = \sum_{i=1}^n Y_i$ et $\bar{Y}_n = \frac{S_n}{n}$ (moyenne empirique de l'échantillon (Y_1, Y_2, \dots, Y_n)).

- a. Déterminer la fonction de répartition F_Y de Y . Quelle est la loi de Y ?
 b. Quelle est la fonction de répartition de S_2 ? En déduire la loi de \bar{Y}_2 .
 c. Déterminer, pour tout $n \in \mathbf{N}^*$, la loi de la variable aléatoire \bar{Y}_n .
 d. La loi faible des grands nombres s'applique-t-elle à la suite $(\bar{Y}_n)_{n \in \mathbf{N}^*}$? Pourquoi ?
 11. Soit $(N, n) \in (\mathbf{N}^*)^2$. On veut simuler N réalisations de la moyenne empirique \bar{Y}_n .
 On suppose que l'on connaît une fonction SciLab cauchy telle que la commande `A = cauchy(N, n)` retourne une matrice $A \in \mathcal{M}_{N,n}(\mathbf{R})$, réalisation d'une famille $(Y_{i,j})_{\substack{1 \leq i \leq N \\ 1 \leq j \leq n}}$ de variables aléatoires indépendantes de loi \mathcal{C}_0 .

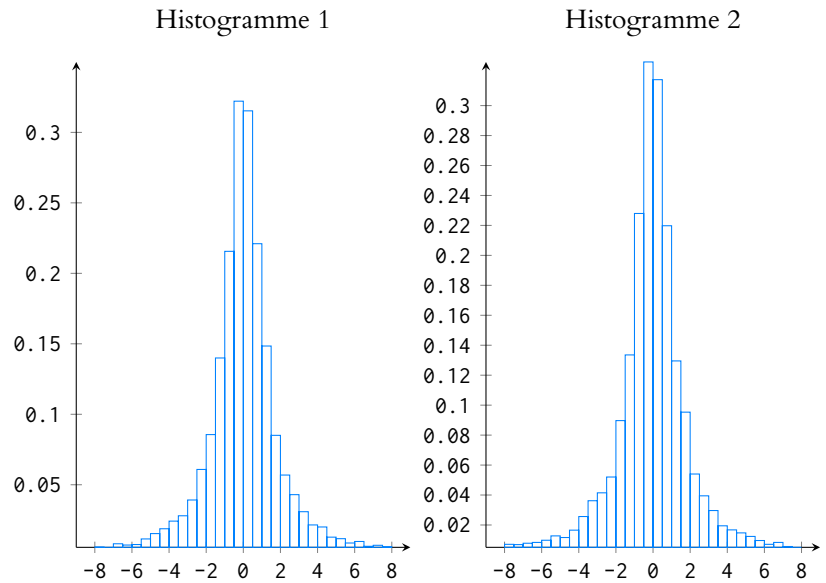
Soit M une matrice de $\mathcal{M}_{N,n}(\mathbf{R})$ avec $(N, n) \in (\mathbf{N}^*)^2$. On rappelle que dans le langage SciLab :

- la commande `sum(M)` retourne une matrice de $\mathcal{M}_{1,1}(\mathbf{R})$ contenant la somme de tous les éléments de M ;
- la commande `sum(M, 'r')` retourne un vecteur ligne de $\mathcal{M}_{1,n}(\mathbf{R})$ contenant les sommes des éléments de M , calculés colonne par colonne ;
- la commande `sum(M, 'c')` retourne un vecteur colonne de $\mathcal{M}_{N,1}(\mathbf{R})$ contenant les sommes des éléments de M , calculées ligne par ligne ;
- la commande `linspace(a, b, m)` retourne un vecteur ligne de m valeurs régulièrement espacées entre a et b , et l'on obtient le même vecteur avec la commande `(a : l : b)` en prenant $l = \frac{b - a}{m - 1}$;

- la commande `histplot(y, data)` permet de représenter les éléments du vecteur `data` sous la forme d'un histogramme : les classes de l'histogramme sont définies par le vecteur strictement croissant `y` : si ce vecteur contient m éléments $y(1), y(2), \dots, y(m)$ tels que $y(1) < y(2) < \dots < y(m)$, alors la première classe de l'histogramme est l'intervalle $]y(1), y(2)[$ et les autres classes sont les intervalles $]y(i), y(i+1)[$ pour $2 \leq i \leq m - 1$.
- a. Compléter le programme suivant afin que la matrice `MoyEmp` contienne 12000 réalisations de la moyenne empirique \overline{Y}_{200} .

```

1 N = 12000 ; n=200 ;
2 A = cauchy(N,n) ;
3 MoyEmp = .....
4 x= (-8 : 0.5 : 8)
5 histplot(x,MoyEmp) \ \ histogramme 1
6 histplot(x,A(:,1)) \ \ histogramme 2
    
```



- b. Les histogrammes 1 et 2 ont été obtenus à l'aide de ce programme. Expliquer en quoi ce couple d'histogrammes illustre le résultat de la question 10.c.

Partie IV. La loi de la médiane empirique.

Dans les questions 13, 14 et 15, on suppose que le paramètre θ est inconnu.

On rappelle que $X \leftrightarrow \mathcal{C}_\theta$. Pour n entier de \mathbf{N}^* , on note $(X_1, X_2, \dots, X_{2n+1})$ un $(2n + 1)$ -échantillon de variables aléatoires indépendantes et de même loi que X .

On admet l'existence de $(2n + 1)$ fonctions $g_1, g_2, \dots, g_{2n+1}$ continues sur \mathbf{R}^{2n+1} , à valeurs réelles, telles que les variables aléatoires réelles $\widehat{X}_1, \widehat{X}_2, \dots, \widehat{X}_{2n+1}$ définies par : $\forall k \in \llbracket 1, 2n + 1 \rrbracket, \widehat{X}_k = g_k(X_1, X_2, \dots, X_{2n+1})$ soient des variables aléatoires à densité et que pour tout $\omega \in \Omega$, les réels $\widehat{X}_1(\omega), \widehat{X}_2(\omega), \dots, \widehat{X}_{2n+1}(\omega)$ soient un réarrangement par ordre croissant de $X_1(\omega), X_2(\omega), \dots, X_{2n+1}(\omega)$: $\forall \omega \in \Omega, \widehat{X}_1(\omega) \leq \widehat{X}_2(\omega) \leq \dots \leq \widehat{X}_{2n+1}(\omega)$.

En particulier, la variable aléatoire $\widehat{X}_{n+1}(\omega)$ est la médiane empirique de l'échantillon $(X_1, X_2, \dots, X_{2n+1})$.

- 12. Pour tout $h \in \mathbf{R}_+^*$ et pour tout $x \in \mathbf{R}$, on note Z la variable aléatoire discrète définie par :

$$\forall \omega \in \Omega, Z(\omega) = \text{Card} \{i \in \llbracket 1, 2n + 1 \rrbracket : x < X_i(\omega) \leq x + h\}.$$

On note A et B les deux événements suivants : $A = [x < \widehat{X}_{n+1} \leq x + h]$ et $B = A \cap [Z = 1]$.

- a. Établir la relation : $P(B) = (n + 1) \binom{2n + 1}{n} (F_X(x))^n (F_X(x + h) - F_X(x)) (1 - F_X(x + h))^n$.
- b. On suppose que le réel x est fixé. Montrer qu'il existe un réel K indépendant de h pour lequel on a

$$0 \leq P(A) - P(B) \leq K (F_X(x + h) - F_X(x))^2.$$

- c. Montrer que \widehat{X}_{n+1} admet une densité $f_{\widehat{X}_{n+1}}$ donnée par

$$\forall x \in \mathbf{R}, f_{\widehat{X}_{n+1}}(x) = (n + 1) \binom{2n + 1}{n} (F_X(x))^n (1 - F_X(x))^n f_X(x).$$

- 13. a. Établir l'équivalence suivante : $x f_{\widehat{X}_{n+1}}(x) \underset{x \rightarrow +\infty}{\sim} (n + 1) \binom{2n + 1}{n} \frac{1}{\pi^{n+1}} \times \frac{1}{x^{n+1}}$.
- b. En déduire l'existence de l'espérance $E(\widehat{X}_{n+1})$ de la variable aléatoire \widehat{X}_{n+1} .
- c. Justifier que \widehat{X}_{n+1} est un estimateur du paramètre θ . Calculer $E(\widehat{X}_{n+1} - \theta)$. Conclure.

- d. À quelle condition nécessaire et suffisante portant sur n , la variable aléatoire \widehat{X}_{n+1} admet-elle une variance ?
14. On note $F_{\widehat{X}_{n+1}}$ la fonction de répartition de \widehat{X}_{n+1} . Soit ε un réel strictement positif.
- a. Établir la relation : $\forall t \in \mathbf{R}, f_{\widehat{X}_{n+1}}(2\theta - t) = f_{\widehat{X}_{n+1}}(t)$.
En déduire que $P(|\widehat{X}_{n+1} - \theta| \geq \varepsilon) = 2F_{\widehat{X}_{n+1}}(\theta - \varepsilon)$.
- b. Montrer que la suite d'estimateurs $(\widehat{X}_{n+1})_{n \in \mathbf{N}^*}$ converge en probabilité vers θ .
- c. La suite $(\widehat{X}_{n+1})_{n \in \mathbf{N}^*}$ converge-t-elle en loi vers la variable certaine θ ?

15. Pour tout entier $n \geq 2$, on pose : $W_{n+1} = \frac{2\sqrt{2n+1}}{\pi} (\widehat{X}_{n+1} - \theta)$.

- a. On note $f_{W_{n+1}}$ la densité continue sur \mathbf{R} de W_{n+1} . Montrer que :

$$\forall x \in \mathbf{R}, f_{W_{n+1}}(x) = \frac{n+1}{2\sqrt{2n+1}} \binom{2n+1}{n} \left[\frac{1}{4} - \frac{1}{\pi^2} \left(\operatorname{Arctan} \left(\frac{\pi x}{2\sqrt{2n+1}} \right) \right)^2 \right]^n \times \left(1 + \frac{\pi^2 x^2}{4(2n+1)} \right)^{-1}.$$

- b. Montrer que pour tout $x \in \mathbf{R}$, on a : $\lim_{n \rightarrow +\infty} f_{W_{n+1}}(x) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right)$.

On admet que ce résultat implique la convergence en loi de la suite de variables aléatoires $(W_{n+1})_{n \geq 2}$ vers une variable aléatoire T qui suit la loi normale centrée réduite.

- c. **Question réservée aux cubes** : on note Φ la fonction de répartition de T . Soit α un réel vérifiant $0 < \alpha < 1$; on pose : $t_\alpha = \Phi^{-1}\left(1 - \frac{\alpha}{2}\right)$.

Déterminer un intervalle de confiance asymptotique pour θ , centré sur \widehat{X}_{n+1} , au niveau de confiance $1 - \alpha$.

16. Dans le langage Sci Lab , la fonction `gsort` permet de trier les éléments d'une matrice réelle A :

- la commande `gsort(A, 'r')` renvoie une copie de A triée colonne par colonne, par ordre décroissant (chaque colonne est triée indépendamment des autres) ;
- la commande `gsort(A, 'c')` renvoie une copie de A triée ligne par ligne, par ordre décroissant (chaque ligne est triée indépendamment des autres).

On suppose que $\theta = 0$ et on considère p réalisations ($p \geq 10^4$) du $(2n+1)$ -échantillon $(X_1, X_2, \dots, X_{2n+1})$.

Recopier et compléter le code suivant afin que son exécution retourne un vecteur `MedianeEmp` de p réalisations de la médiane empirique \widehat{X}_{n+1} puis un vecteur `W` de p réalisations de W_{n+1} .

```

1 A = cauchy(p, 2*n+1)
2 S = gsort .....
3 MedianeEmp = .....
4 W = .....

```

CORRECTION DU DEVOIR SURVEILLÉ 8

Partie I. Une démonstration probabiliste de la formule de Stirling.

- 1.a. Procédons au changement de variable indiqué, qui est légitime car affine. Lorsque $x \rightarrow 0$, alors $u \rightarrow n$ et lorsque $x \rightarrow 1$, alors $u \rightarrow 0$. On a alors $x = 1 - \frac{u}{n}$ et donc $dx = -\frac{du}{n}$. Ainsi,

$$I_n = \int_n^0 \left(\frac{u}{n} e^{1-\frac{u}{n}}\right)^n \left(-\frac{1}{n}\right) du = \int_0^n \frac{u^n}{n^{n+1}} e^{n-u} du = \boxed{\frac{e^n}{n^{n+1}} \int_0^n u^n e^{-u} du.}$$

- 1.b. S'il serait possible de répondre à la question en étudiant le signe¹ de la fonction $x \mapsto x + \ln(1-x) + \frac{x^2}{2}$, notons qu'on a reconnu que $-x - \frac{x^2}{2}$ est le développement limité à l'ordre 2 de $\ln(1-x)$ et donc il s'agit de déterminer la «position» de $\ln(1-x)$ par rapport à son développement limité. C'est précisément ce à quoi sert la formule de Taylor avec reste intégral.
Soit donc $\varphi : [0, 1] \rightarrow \mathbb{R}$ la fonction définie par $\varphi(x) = \ln(1-x)$. Alors φ est de classe \mathcal{C}^3 sur $[0, 1]$, avec

$$\varphi'(x) = \frac{-1}{1-x}, \quad \varphi''(x) = \frac{-1}{(1-x)^2}, \quad \varphi^{(3)}(x) = \frac{-2}{(1-x)^3}.$$

Et donc, par la formule de Taylor, pour tout $x \in [0, 1]$,

$$\ln(1-x) + x + \frac{x^2}{2} = \varphi(x) - \varphi(0) - \varphi'(0)x - \frac{\varphi''(0)}{2}x^2 = \int_0^x \varphi^{(3)}(t) \frac{(x-t)^2}{2} dt.$$

Mais, pour tout $t \in [0, 1]$, $\varphi^{(3)}(t)(x-t)^2 \leq 0$ et donc, par croissance de l'intégrale,

$$\int_0^x \varphi^{(3)}(t) \frac{(x-t)^2}{2} dt \leq 0.$$

On en déduit que

$$\ln(1-x) + x + \frac{x^2}{2} \leq 0 \Leftrightarrow \boxed{\ln(1-x) + x \leq -\frac{x^2}{2}.}$$

- 1.c. Puisque h_n est positive sur $[0, 1]$, par croissance de l'intégrale, $I_n \geq 0$.
D'autre part, d'après la question précédente, pour tout $x \in [0, 1]$,

$$x + \ln(1-x) \leq -\frac{x^2}{2} \Rightarrow nx + n \ln(1-x) \leq -n \frac{x^2}{2}$$

et donc par passage à l'exponentielle²,

$$h_n(x) = e^{nx+n \ln(1-x)} \leq e^{-nx^2/2}.$$

On en déduit donc que $I_n \leq \int_0^1 e^{-nx^2/2} dx$.

Le changement de variable $t = \sqrt{nx}$ nous donne alors

$$\int_0^1 e^{nx^2/2} dx = \int_0^{\sqrt{n}} e^{-t^2/2} \frac{dt}{\sqrt{n}} \leq \frac{1}{\sqrt{n}} \int_0^{+\infty} e^{-t^2/2} dt.$$

Or, la densité d'une loi normale centrée réduite est $t \mapsto \frac{1}{\sqrt{2\pi}} e^{-t^2/2}$, et donc

$$\int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}} e^{-t^2/2} dt = 1 \Leftrightarrow \int_{-\infty}^{+\infty} e^{-t^2/2} dt = \sqrt{2\pi}.$$

Et alors, par parité de $t \mapsto e^{-t^2/2}$,

$$\int_0^{+\infty} e^{-t^2/2} dt = \frac{1}{2} \int_{-\infty}^{+\infty} e^{-t^2/2} dt = \sqrt{\frac{\pi}{2}}.$$

¹ Par exemple en dressant son tableau de variation.

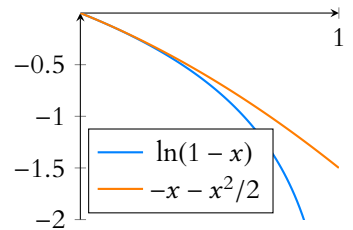


FIGURE 24.1– La fonction $\ln(1-x)$ et son développement limité en 0.

² Qui préserve le sens des inégalités car l'exponentielle est croissante.

Remarque

Le changement de variable était plus ou moins imposé par l'énoncé qui demandait de se référer à une loi normale **centrée réduite**, mais il aurait été possible de répondre à la question en faisant appel à une loi normale $\mathcal{N}(0, \frac{1}{n})$, car sa densité permet également de faire apparaître le terme en $e^{-nx^2/2}$.

Ainsi, on a bien

$$I_n \leq \frac{1}{\sqrt{n}} \int_0^{+\infty} e^{-t^2/2} dt \leq \sqrt{\frac{\pi}{2n}}.$$

2. Puisqu'on doit avoir $h_n^*(x) = \exp\left(-\frac{nx^2}{2}H(x)\right)$, alors $H(x) = -\frac{2}{nx^2} \ln(h_n^*(x)) = -\frac{2}{x^2}(x + \ln(1-x))$.

2.a. Lorsque $x \rightarrow 0^+$, $x + \ln(1-x) = x - x - \frac{x^2}{2} + o(x^2) \underset{x \rightarrow 0^+}{\sim} -\frac{x^2}{2}$.
Et donc $H(x) \underset{x \rightarrow 0^+}{\sim} 1 \xrightarrow{x \rightarrow 0^+} 1$.

Et donc H est prolongeable par continuité en 0 en posant $H(0) = 1$.

2.b. La fonction g est deux fois dérivable sur $]0, 1[$, et sa dérivée est

$$g'(x) = \frac{1-x}{x-1} - \ln(1-x) + 1-x = -\ln(1-x) - x.$$

Et donc sa dérivée seconde est

$$g''(x) = \frac{1}{1-x} - 1 = \frac{x}{1-x} \geq 0.$$

Ainsi, g est convexe sur $]0, 1[$.

D'autre part, en utilisant la question 1.b, il vient $g'(x) \geq \frac{x^2}{2} > 0$.

Et donc g est strictement croissante sur $]0, 1[$. Mais $\lim_{x \rightarrow 0^+} g(x) = 0$, et on en déduit donc que g est strictement positive sur $]0, 1[$.

2.c. L'expression donnée plus haut de H prouve que celle-ci est dérivable sur $]0, 1[$ car quotient de fonctions dérivables, et

$$\begin{aligned} H'(x) &= -2 \frac{\left(1 - \frac{1}{1-x}\right)x^2 - 2x(x + \ln(1-x))}{x^4} = -2 \frac{-x^2 - 2x \ln(1-x) - \frac{x^2}{1-x}}{x^4} \\ &= \frac{2}{(1-x)x^3} \left(2(1-x) \ln(1-x) - 2x + x^2\right) = \frac{4}{(1-x)x^3} g(x) > 0. \end{aligned}$$

Ainsi, H est strictement croissante sur $[0, 1[$.

Puisque H est continue³ sur $[0, 1[$, que $H(0) = 1$ et que $\lim_{x \rightarrow 1^-} H(x) = +\infty$, par le théorème de la bijection, H réalise une bijection strictement croissante de $[0, 1[$ sur $[1, +\infty[$.

3.a. Posons $u_n = \frac{1}{(n+1)^{1/3}}$.

Alors $\lim_{n \rightarrow +\infty} u_n = 0$, et pour tout $n \in \mathbf{N}^*$, $0 < u_n < 1$.

De plus, lorsque $n \rightarrow +\infty$, $\sqrt{n}u_n \underset{n \rightarrow +\infty}{\sim} \frac{n^{1/2}}{n^{1/3}} = n^{1/6} \xrightarrow{n \rightarrow +\infty} 0$.

Ainsi, la suite $(u_n)_{n \geq 0}$ définie par $u_n = \frac{1}{\sqrt[3]{n+1}}$ convient.

3.b. Puisque $u_n \xrightarrow{n \rightarrow +\infty} 0$, par continuité de la fonction H ,

$$\lim_{n \rightarrow +\infty} v_n = \lim_{n \rightarrow +\infty} H(u_n) = H(0) = 1.$$

3.c. La fonction h_n étant positive sur $[u_n, 1]$, on a

$$I_n = \int_0^{u_n} h_n(x) dx + \underbrace{\int_{u_n}^1 h_n(x) dx}_{\geq 0} \geq \int_0^{u_n} h_n(x) dx.$$

Dans cette dernière intégrale, procédons au changement de variable affine $y = x\sqrt{nv_n}$. Il vient alors

$$\int_0^{u_n} h_n(x) dx = \int_0^{u_n \sqrt{nv_n}} h_n\left(\frac{y}{\sqrt{nv_n}}\right) \frac{dy}{\sqrt{nv_n}}.$$

Remarque

Notons que H ne dépend pas de n , ce qui n'était pas évident sur la définition donnée dans l'énoncé.

Ouvert/fermé

Puisque H est continue sur $[0, 1[$, le fait que sa dérivée soit positive strictement sur $]0, 1[$ (ouvert en 0) garantit sa stricte croissance sur $[0, 1[$ (fermé en 0).

³ Elle trivialement continue sur $]0, 1[$, et a été prolongée par continuité en 0, donc est continue en 0.

Rédaction

N'oublions pas que pour composer des limites par une fonction, il est indispensable que cette fonction soit continue. Et donc **il faut mentionner cette continuité**.

Or, pour $y \in]0, u_n \sqrt{nv_n}]$, on a, par définition de H ,

$$h_n \left(\frac{y}{\sqrt{nv_n}} \right) = \exp \left(-\frac{ny^2}{2nv_n} H \left(\frac{y}{\sqrt{nv_n}} \right) \right) = \exp \left(-\frac{y^2}{2} \frac{H \left(\frac{y}{\sqrt{nv_n}} \right)}{H(u_n)} \right).$$

Mais, toujours pour $y \in]0, u_n \sqrt{nv_n}]$, $\frac{y}{\sqrt{nv_n}} \leq u_n$, et donc, par croissance de H , $\frac{H \left(\frac{y}{\sqrt{nv_n}} \right)}{H(u_n)} \leq 1$.

On en déduit donc que

$$\exp \left(-\frac{y^2}{2} \frac{H \left(\frac{y}{\sqrt{nv_n}} \right)}{H(u_n)} \right) \geq \exp \left(-\frac{y^2}{2} \right)$$

de sorte que, par croissance de l'intégrale,

$$I_n \geq \frac{1}{\sqrt{nv_n}} \int_0^{u_n \sqrt{nv_n}} \exp \left(-\frac{y^2}{2} \right) dy.$$

3.d. Lorsque $n \rightarrow +\infty$, $\sqrt{nv_n} \underset{n \rightarrow +\infty}{\sim} \sqrt{n}$.

D'autre part, on a alors

$$u_n \sqrt{nv_n} = \underbrace{u_n}_{\underset{n \rightarrow +\infty}{\rightarrow} +\infty} \underbrace{\sqrt{n}}_{\underset{n \rightarrow +\infty}{\rightarrow} 1} \underset{n \rightarrow +\infty}{\rightarrow} +\infty.$$

Et donc $\int_0^{u_n \sqrt{nv_n}} \exp \left(-\frac{y^2}{2} \right) dy \underset{n \rightarrow +\infty}{\rightarrow} \int_0^{+\infty} \exp \left(-\frac{y^2}{2} \right) dy = \sqrt{\frac{\pi}{2}}$.

Par conséquent, $\frac{1}{\sqrt{nv_n}} \int_0^{u_n \sqrt{nv_n}} \exp \left(-\frac{y^2}{2} \right) dy \underset{n \rightarrow +\infty}{\sim} \sqrt{\frac{\pi}{2n}}$.

En combinant les inégalités des questions 1.c et 3.c, on a donc

$$\frac{\frac{1}{\sqrt{nv_n}} \int_0^{u_n \sqrt{nv_n}} \exp \left(-\frac{y^2}{2} \right) dy}{\sqrt{\frac{\pi}{2n}}} \leq \frac{I_n}{\sqrt{\frac{\pi}{2n}}} \leq 1.$$

Et donc, par le théorème des gendarmes, $\lim_{n \rightarrow +\infty} \frac{I_n}{\sqrt{\frac{\pi}{2n}}} = 1$, de sorte que $I_n \underset{n \rightarrow +\infty}{\sim} \sqrt{\frac{\pi}{2n}}$.

4.a. Puisque la loi $\mathcal{E}(1)$ est également la loi $\gamma(1)$, par indépendance des T_i et par stabilité des lois gamma, $S_n \hookrightarrow \gamma(n)$.

D'autre part, les T_i sont indépendantes, identiquement distribuées, possèdent une espérance égale à 1 et une variance égale à 1.

Notons alors \overline{S}_n^* la variable centrée réduite associée à S_n , c'est-à-dire $\overline{S}_n^* = \frac{S_n - n}{\sqrt{n}}$.

Par le théorème central limite, $\overline{S}_n^* \xrightarrow{\mathcal{L}} X$ où X suit une loi normale centrée réduite.

En particulier, il vient

$$\lim_{n \rightarrow +\infty} P(S_n \leq n) = \lim_{n \rightarrow +\infty} P(S_n - n \leq 0) = \lim_{n \rightarrow +\infty} P(\overline{S}_n^* \leq 0) = P(X \leq 0) = \Phi(0) = \frac{1}{2}.$$

4.b. Soit $\varepsilon > 0$. Alors

$$P(|U_n| \geq \varepsilon) = P(U_n \geq \varepsilon) = P(T_{n+1} \geq \varepsilon \sqrt{n}).$$

Mais par l'inégalité de Markov, qui s'applique car T_{n+1} est une variable aléatoire positive possédant une espérance,

$$P(T_{n+1} \geq \varepsilon \sqrt{n}) \leq \frac{E(T_{n+1})}{\varepsilon \sqrt{n}} \leq \frac{1}{\varepsilon \sqrt{n}} \underset{n \rightarrow +\infty}{\rightarrow} 0.$$

Intégrale de Gauss

La valeur de $\int_0^{+\infty} e^{-y^2/2} dy$ a été calculée à la question 1.c

Markov

N'oublions pas que l'inégalité de Markov ne s'applique que pour des variables aléatoires positives, et donc qu'il faut vérifier (et mentionner) la positivité avant d'utiliser l'inégalité.

Ceci prouve bien que $U_n \xrightarrow{P} 0$.

4.c. Puisque $\overline{S}_n^* \xrightarrow{\mathcal{L}} X$ et que $\frac{T_{n+1}}{\sqrt{n}} \xrightarrow{P} 0$, par le théorème de Slutsky, $\overline{S}_n^* + \frac{T_{n+1}}{\sqrt{n}} \xrightarrow{\mathcal{L}} X + 0 = X$.

Et donc

$$P(S_{n+1} \leq n) = P(S_{n+1} - n \leq 0) = P\left(\frac{S_{n+1} - n}{\sqrt{n}} \leq 0\right) = P\left(\overline{S}_n^* + \frac{T_{n+1}}{\sqrt{n}} \leq 0\right) \xrightarrow{n \rightarrow +\infty} P(X \leq 0) = \boxed{\frac{1}{2}}.$$

5. En combinant les résultats des questions 1.a et 3.d, il vient

$$I_n = \frac{e^n}{n^{n+1}} \int_0^n u^n e^{-u} du \underset{n \rightarrow +\infty}{\sim} \sqrt{\frac{\pi}{2n}}.$$

$$\text{Et donc } \int_0^n u^n e^{-u} du \underset{n \rightarrow +\infty}{\sim} \sqrt{\frac{\pi}{2n}} \frac{n^{n+1}}{e^n}.$$

D'autre part, nous savons que S_{n+1} suit une loi $\gamma(n+1)$, donc une densité de S_{n+1} est la

fonction $x \mapsto \begin{cases} \frac{1}{n!} x^n e^{-x} & \text{si } x > 0 \\ 0 & \text{sinon} \end{cases}$ de sorte que

$$P(S_{n+1} \leq n) = \int_0^n \frac{1}{n!} x^n e^{-x} dx.$$

Mais d'après 4.c, cette probabilité tend vers $\frac{1}{2}$, et donc $\int_0^n \frac{1}{n!} x^n e^{-x} dx \underset{n \rightarrow +\infty}{\sim} \frac{1}{2}$.

On en déduit donc que

$$n! \underset{n \rightarrow +\infty}{\sim} 2 \int_0^n x^n e^{-x} dx \underset{n \rightarrow +\infty}{\sim} 2 \frac{n^{n+1}}{e^n} \sqrt{\frac{\pi}{2n}} \underset{n \rightarrow +\infty}{\sim} \boxed{n^n e^{-n} \sqrt{2\pi n}}.$$

Partie II. Quelques propriétés de la loi de Cauchy.

6.a. Dur de dire quelle réponse est attendue ici : on peut dire que l'imparité de l'arctangente est du cours, ou encore qu'il est connu que la bijection réciproque d'une fonction impaire est impaire.

Mais puisque l'énoncé semble attendre une preuve, redémontrons le résultat du cours.

Pour tout $x \in \mathbf{R}$, $\tan(\text{Arctan}(x)) = x$ et $\tan(\text{Arctan}(-x)) = -x$.

Or, la fonction \tan est impaire⁴, et donc $-x = \tan(-\text{Arctan}(x)) = \tan(\text{Arctan}(-x))$.

Puisque $-\text{Arctan}(x)$ et $\text{Arctan}(-x) \in \left]-\frac{\pi}{2}, \frac{\pi}{2}\right[$, par bijectivité⁵ de Arctan sur $\left]-\frac{\pi}{2}, \frac{\pi}{2}\right[$, on

a $\text{Arctan}(-x) = \text{Arctan}(x)$: $\boxed{\text{la fonction Arctan est impaire.}}$

6.b. La fonction Arctan est de classe \mathcal{C}^3 puisque sa dérivée est \mathcal{C}^2 . Donc par la formule de Taylor-Young, elle admet un développement limité à l'ordre 3 au voisinage de 0.

Or, on a

$$\text{Arctan}'(x) = \frac{1}{1+x^2}, \text{Arctan}''(x) = \frac{-2x}{(1+x^2)^2}, \text{Arctan}^{(3)}(x) = \frac{-2(1+x^2)^2 + 8x(1+x^2)}{(1+x^2)^3}.$$

Et donc⁶, le développement limité d'ordre 3 de Arctan au voisinage de 0 est

$$\begin{aligned} \text{Arctan}(x) &= \text{Arctan}(0) + \text{Arctan}'(0)x + \frac{\text{Arctan}''(0)}{2}x^2 + \frac{\text{Arctan}^{(3)}(0)}{3!}x^3 + o(x^3) \\ &= x - \frac{2}{6}x^3 + o(x^3) = \boxed{x - \frac{x^3}{3} + o(x^3)}. \end{aligned}$$

6.c. La dérivée seconde de $\text{Arctan}(x)$ est négative sur \mathbf{R}_+ , donc Arctan est concave sur \mathbf{R}_+ . Par conséquent, elle est située en dessous de ses tangentes.

Mais la tangente de $\text{Arctan}(x)$ en 0 est la droite d'équation $y = x$ et donc pour tout

$x \in \mathbf{R}_+$, $\boxed{\text{Arctan}(x) \leq x}$.

D'autre part, Arctan est croissante⁷ sur \mathbf{R} . Or, $\text{Arctan}(0) = 0$, et donc pour tout $x \in \mathbf{R}_+$,

$\boxed{0 \leq \text{Arctan}(x)}$.

Rappel

On a $\Gamma(n+1) = n!$

⁴ Car quotient de sin, qui est impaire, par cos qui est paire.

⁵ On n'utilise en fait ici que l'injectivité.

⁶ Toujours d'après la formule de Taylor-Young.

Tangente

Rappelons que l'équation de la tangente en 0 se lit sur le développement limité : c'est le développement limité d'ordre 1 au voisinage de 0.

⁷ Car sa dérivée est positive, mais aussi directement car c'est la bijection réciproque d'une fonction croissante.

- 6.d. Soit g la fonction définie sur \mathbf{R}_+^* par $g(x) = \text{Arctan}(x) + \text{Arctan}\left(\frac{1}{x}\right)$. Alors g est dérivable sur \mathbf{R}_+^* car somme de fonctions dérivables, et

$$\forall x \in \mathbf{R}_+^*, g'(x) = \frac{1}{1+x^2} - \frac{1}{x^2} \frac{1}{1+\frac{1}{x^2}} = \frac{1}{1+x^2} - \frac{1}{x^2+1} = 0.$$

Donc g est constante sur \mathbf{R}_+^* . De plus, $g(1) = 2\text{Arctan}(1) = 2\frac{\pi}{4} = \frac{\pi}{2}$ et donc pour tout $x \in \mathbf{R}_+^*$,

$$\boxed{\text{Arctan}(x) + \text{Arctan}\left(\frac{1}{x}\right) = \frac{\pi}{2}.$$

- 7.a. Il est évident que $x \mapsto \frac{1}{\pi} \frac{1}{1+(x-\theta)^2}$ est continue et positive sur \mathbf{R} .

De plus, pour $A < B$, on a

$$\int_A^B \frac{1}{\pi} \frac{1}{1+(x-\theta)^2} dx = \left[\frac{1}{\pi} \text{Arctan}(x-\theta) \right]_A^B = \frac{\text{Arctan}(B-\theta)}{\pi} - \frac{\text{Arctan}(A-\theta)}{\pi}$$

$$\xrightarrow[A \rightarrow -\infty]{B \rightarrow +\infty} \frac{\pi}{2\pi} + \frac{\pi}{2\pi} = 1.$$

Et donc $\int_{-\infty}^{+\infty} \frac{1}{\pi} \frac{1}{1+(x-\theta)^2} dx$ converge et vaut 1.

Ainsi, $x \mapsto \frac{1}{\pi} \frac{1}{1+(x-\theta)^2}$ est bien une densité de probabilités.

- 7.b. Par définition, X admet une espérance si et seulement $\int_{-\infty}^{+\infty} x f_X(x) dx$ converge.

Or, au voisinage de $+\infty$, $\frac{x}{\pi} \frac{1}{1+(x-\theta)^2} \underset{x \rightarrow +\infty}{\sim} \frac{x}{\pi} \frac{1}{x^2} = \frac{1}{\pi x}$.

Et puisque $\int_1^{+\infty} \frac{dx}{\pi x}$ diverge⁸, par critère de comparaison pour les intégrales de fonctions positives, il en est de même de $\int_1^{+\infty} x f_X(x) dx$.

Et donc X n'admet pas d'espérance.

- 7.c. Contrairement à ce que demande la question, nous n'utilisons pas un repère orthonormé, dans lequel la courbe serait toute «écrasée» et peu lisible.

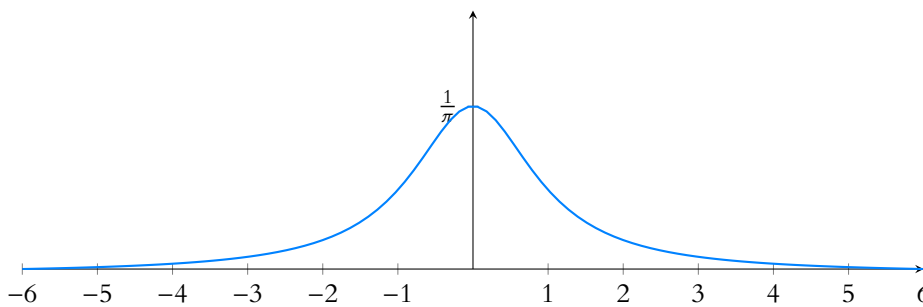


FIGURE 24.2 – La courbe représentative de f .

- 8.a. Soit $x \in \mathbf{R}$. Alors $F_X(x) = \int_{-\infty}^x f_X(t) dt$.

Mais pour $A < x$,

$$\int_A^x f_X(t) dt = \left[\frac{1}{\pi} \text{Arctan}(t-\theta) \right]_A^x = \frac{1}{\pi} \text{Arctan}(x-\theta) - \frac{1}{\pi} \text{Arctan}(A-\theta) \xrightarrow{A \rightarrow -\infty} \frac{\text{Arctan}(x-\theta)}{\pi} + \frac{1}{2}.$$

Ainsi,

$$\boxed{\forall x \in \mathbf{R}, F_X(x) = \frac{\text{Arctan}(x-\theta)}{\pi} + \frac{1}{2}.$$

Fonction constante

Le fait que g soit constante car de dérivée nulle est valable car \mathbf{R}_+^* est un **intervalle**. Par exemple, si on définit g de la même manière sur \mathbf{R}^* , alors sa dérivée est toujours nulle, mais elle n'est pas constante sur \mathbf{R}^* , qui n'est pas un intervalle. En revanche, elle est bien constante sur les deux intervalles \mathbf{R}_-^* et \mathbf{R}_+^* .

⁸ C'est une intégrale de Riemann avec $\alpha = 1$.

Remarque

On prouverait de même que l'intégrale de $x f_X(x)$ diverge aussi au voisinage de $-\infty$, mais ceci n'est pas nécessaire : puisqu'elle diverge au voisinage de $+\infty$, $\int_{-\infty}^{+\infty} x f_X(x) dx$ diverge, indépendamment de ce qu'il se passe en $-\infty$.

Méthode

Lorsqu'on demande de représenter une fonction f , il n'est pas besoin de faire un dessin parfait, ni d'essayer de calculer 10 valeurs de f pour placer correctement la courbe.

Le plus important est d'avoir une allure correcte, et de faire apparaître autant que possible les informations dont on dispose : les variations, les limites éventuelles, la parité/imparité, les valeurs qui auraient été calculées précédemment, la concavité, éventuellement une tangente si son équation est facile à obtenir.

Ici f est paire, de limite nulle en $+\infty$, décroissante sur \mathbf{R}_+ et vaut $1/\pi$ en 0. Toutes ces informations sont attendues sur la courbe.

8.b. Reprenons le résultat de la question précédente,

$$F_X(x) = \frac{1}{2} \Leftrightarrow \frac{\text{Arctan}(x - \theta)}{\pi} + \frac{1}{2} = \frac{1}{2} \Leftrightarrow \text{Arctan}(x - \theta) = 0 \Leftrightarrow x - \theta = 0 \Leftrightarrow x = \theta.$$

Ainsi, $F_X(x) = \frac{1}{2}$ possède une unique solution qui est θ .

Partie III. La loi de la moyenne empirique.

9.a. La fonction $\varphi_{n,x}$ est continue sur \mathbf{R} , donc les éventuels problèmes de convergence de l'intégrale sont en $\pm\infty$.

Or, aussi bien au voisinage de $+\infty$ qu'au voisinage de $-\infty$, on a

$$\varphi_{n,x}(t) \underset{t \rightarrow \pm\infty}{\sim} \frac{1}{n^2 t^4}.$$

Or, $\int_{-\infty}^{-1} \frac{dt}{n^2 t^4}$ et $\int_1^{+\infty} \frac{dt}{n^2 t^4}$ sont des intégrales convergentes, donc par critère de comparaison pour les intégrales de fonctions positives, $\int_{-\infty}^{-1} \varphi_{n,x}(t) dt$ et $\int_1^{+\infty} \varphi_{n,x}(t) dt$ convergent.

Puisque $\int_{-1}^1 \varphi_{n,x}(t) dt$ est l'intégrale d'une fonction continue sur un segment, elle converge,

et donc $\int_{-\infty}^{+\infty} \varphi_{n,x}(t) dt$ converge.

9.b. Une primitive de $t \mapsto \frac{2t}{1+t^2}$ est $t \mapsto \ln(1+t^2)$.

Une primitive de $t \mapsto \frac{2n(nt-x)}{1+(x-nt)^2}$ est $t \mapsto \ln(1+(x-nt)^2)$.

Par conséquent, pour $A < B$,

$$\begin{aligned} \int_A^B \psi_{n,x}(t) dt &= [\ln(1+t^2) - \ln(1+(x-nt)^2)]_A^B \\ &= \ln\left(\frac{1+B^2}{1+(x-nB)^2}\right) - \ln\left(\frac{1+A^2}{1+(x-nA)^2}\right). \end{aligned}$$

Or, lorsque $A \rightarrow -\infty$, $\frac{1+A^2}{1+(x-nA)^2} \underset{A \rightarrow -\infty}{\sim} \frac{1}{n^2}$, et donc

$$\lim_{A \rightarrow -\infty} \ln\left(\frac{1+A^2}{1+(x-nA)^2}\right) = \ln\left(\frac{1}{n^2}\right) = -2\ln(n).$$

Ainsi, $\int_{-\infty}^B \psi_{n,x}(t) dt = \ln\left(\frac{1+B^2}{1+(x-nB)^2}\right) - 2\ln(n)$.

De même, lorsque $B \rightarrow +\infty$, $\ln\left(\frac{1+B^2}{1+(x-nB)^2}\right) \xrightarrow{B \rightarrow +\infty} -2\ln(n)$ et donc $\int_{-\infty}^{+\infty} \psi_{n,x}(t) dt$ converge et

$$\int_{-\infty}^{+\infty} \psi_{n,x}(t) dt = -2\ln(n) + 2\ln(n) = \boxed{0}.$$

9.c. Utilisons les valeurs de $\alpha, \beta, \gamma, \delta$ données par l'énoncé :

$$\begin{aligned} \varphi_{n,x}(t) &= \frac{1}{\sigma_{n,x}} \left(\frac{2nxt}{1+t^2} + \frac{1+x^2-n^2}{1+t^2} - \frac{2n^3xt}{1+(x-nt)^2} + \frac{3n^2x^2+n^4-n^2}{1+(x-nt)^2} \right) \\ &= \frac{1}{\sigma_{n,x}} \left(nx \left(\frac{2t}{1+t^2} - \frac{2n^2t-2nx}{1+(x-nt)^2} \right) + \frac{1+x^2-n^2}{1+t^2} + \frac{n^2x^2+n^4-n^2}{1+(x-nt)^2} \right) \\ &= \frac{1}{\sigma_{n,x}} \left(nx\psi_{n,x}(t) + \frac{1+x^2-n^2}{1+t^2} + \frac{n^2x^2+n^4-n^2}{1+(x-nt)^2} \right). \end{aligned}$$

Par conséquent, il vient⁹

$$\int_{-\infty}^{+\infty} \varphi_{n,x}(t) dt = \frac{1}{\sigma_{n,x}} \left(\underbrace{nx \int_{-\infty}^{+\infty} \psi_{n,x}(t) dt}_{=0} + (1+x^2-n^2) \int_{-\infty}^{+\infty} \frac{dt}{1+t^2} + (n^2x^2+n^4-n^2) \int_{-\infty}^{+\infty} \frac{dt}{1+(x-nt)^2} \right).$$

Remarque

Nous venons de voir que $\varphi_{n,x}$ possède le même équivalent en $+\infty$ qu'en $-\infty$, mais ce n'est pas pour autant qu'il s'agit d'une fonction paire : il est impératif de donner des équivalents en $+\infty$ et en $-\infty$ pour étudier la nature de l'intégrale.

On a utilisé
 $3n^2x^2 = 2n^2x^2 + n^2x^2$.

⁹ Il est facile de vérifier que toutes ces intégrales convergent.

Mais

$$\int_A^B \frac{dt}{1+t^2} = [\text{Arctan}(t)]_A^B = \text{Arctan}(B) - \text{Arctan}(A) \xrightarrow[A \rightarrow -\infty]{B \rightarrow +\infty} \pi.$$

De même,

$$\int_A^B \frac{dt}{1+(x-nt)^2} = \left[\frac{1}{n} \text{Arctan}(nt-x) \right]_A^B \xrightarrow[A \rightarrow -\infty]{B \rightarrow +\infty} \frac{\pi}{n}.$$

Et donc

$$\begin{aligned} \int_{-\infty}^{+\infty} \varphi_{n,x}(t) dt &= \frac{1}{\sigma_{n,x}} \left((1+x^2-n^2)\pi + n(x^2+n^2-1)\pi \right) \\ &= \frac{1+x^2-n^2+nx^2+n^3-n}{(x^2+(n+1)^2)(x^2+(n-1)^2)} \pi. \end{aligned}$$

Or, on a

$$\begin{aligned} (n+1)(x^2+(n-1)^2) &= (n+1)(x^2+n^2-2n+1) \\ &= nx^2+n^3-2n^2+n+x^2+n^2-2n+1 \\ &= n^3+nx^2+x^2-n^2-n+1. \end{aligned}$$

Et donc

$$\int_{-\infty}^{+\infty} \varphi_{n,x}(t) dt = \frac{(n+1)(x^2+(n-1)^2)}{(x^2+(n+1)^2)(x^2+(n-1)^2)} \pi = \boxed{\frac{(n+1)\pi}{x^2+(n+1)^2}}.$$

Détails

Cette factorisation de $1+x^2-n^2+\dots$ ne sort pas de nulle part : au vu du résultat donné dans l'énoncé et de la valeur de l'intégrale obtenue précédemment, il fallait qu'elle soit vraie !

10.a. Soit $x \in \mathbf{R}$. Alors

$$F_Y(x) = P(Y \leq x) = P(X - \theta \leq x) = P(X \leq x + \theta) = \frac{\text{Arctan}(x)}{\pi} + \frac{1}{2}.$$

On reconnaît alors la loi de Cauchy de paramètre 0.

10.b. Puisque Y_1 et Y_2 sont indépendantes, et que leurs densités sont bornées¹⁰, une densité de S_2 est donnée par la formule du produit de convolution :

¹⁰ Par $\frac{1}{\pi}$.

$$f_{S_2}(x) = \int_{-\infty}^{+\infty} f_{Y_1}(t) f_{Y_2}(x-t) dt = \int_{-\infty}^{+\infty} \frac{1}{\pi} \frac{1}{1+t^2} \frac{1}{\pi} \frac{1}{1+(x-t)^2} dt = \frac{1}{\pi^2} \int_{-\infty}^{+\infty} \varphi_{1,x}(t) dt = \frac{2}{\pi} \frac{1}{x^2+4}.$$

Ainsi, pour $x \in \mathbf{R}$, on a

$$P(S_2 \leq x) = \int_{-\infty}^x f_{S_2}(t) dt = \frac{2}{\pi} \int_{-\infty}^x \frac{1}{t^2+4} dt.$$

Mais, pour $A \leq x$,

$$\begin{aligned} \int_A^x \frac{2}{\pi} \frac{1}{t^2+4} &= \frac{1}{2\pi} \int_A^x \frac{1}{1+(\frac{t}{2})^2} dt \\ &= \frac{1}{\pi} \left[\text{Arctan}\left(\frac{t}{2}\right) \right]_A^x = \frac{1}{\pi} \left(\text{Arctan}\left(\frac{x}{2}\right) - \text{Arctan}\left(\frac{A}{2}\right) \right) \\ &\xrightarrow[A \rightarrow -\infty]{} \frac{\text{Arctan}\left(\frac{x}{2}\right)}{\pi} + \frac{1}{2}. \end{aligned}$$

On a divisé numérateur et dénominateur par 4.

Et alors, $P(\overline{Y_2} \leq x) = P(S_2 \leq 2x) = \frac{\text{Arctan}(x)}{\pi} + \frac{1}{2}$.

On en déduit que $\overline{Y_2}$ suit la loi de Cauchy de paramètre 0.

10.c. Au vu des résultats précédents, il semble légitime de supposer que $\overline{Y_n}$ suit la loi \mathcal{C}_0 . Montrons-le par récurrence sur n .

La récurrence est largement initialisée puisque nous savons déjà que $\overline{Y_1}$ et $\overline{Y_2}$ suivent la loi \mathcal{C}_0 .

Remarque

Nous sommes ici passés par la fonction de répartition car l'énoncé le demandait, mais une fois une densité de S_2 connue, il était aisé d'obtenir une densité de $\overline{Y_2}$ par transformation affine. Et alors on aurait reconnu la densité de la loi \mathcal{C}_0 .

Supposons donc que \overline{Y}_n suit la loi \mathcal{C}_0 .

Alors, par transformation affine, une densité de $S_n = n\overline{Y}_n$ est donnée par

$$x \mapsto \frac{1}{n\pi} \frac{1}{1 + \left(\frac{x}{n}\right)^2}.$$

De plus, d'après le lemme des coalitions, S_n et Y_{n+1} sont indépendantes, et donc une densité de S_{n+1} est donnée par

$$f_{S_{n+1}}(x) = \int_{-\infty}^{+\infty} f_{S_n}(t) f_{Y_{n+1}}(x-t) dt = \int_{-\infty}^{+\infty} \frac{1}{n\pi} \frac{1}{1 + \left(\frac{t}{n}\right)^2} \frac{1}{\pi} \frac{1}{1 + (x-t)^2} dt.$$

Procédons au changement de variable¹¹ $u = \frac{t}{n}$, de sorte que

¹¹ Affine.

$$f_{S_{n+1}}(x) = \frac{1}{n\pi^2} \int_{-\infty}^{+\infty} \frac{1}{1+u^2} \frac{1}{1+(x-nu)^2} n du = \frac{1}{\pi^2} \int_{-\infty}^{+\infty} \varphi_{n,x}(u) du = \frac{1}{\pi} \frac{n+1}{x^2 + (n+1)^2}.$$

Alors, par transformation affine, une densité de $\overline{Y}_{n+1} = \frac{S_{n+1}}{n+1}$ est

$$x \mapsto (n+1) f_{S_{n+1}}((n+1)x) = \frac{1}{\pi} \frac{(n+1)^2}{((n+1)x)^2 + (n+1)^2} = \frac{1}{\pi} \frac{1}{1+x^2}.$$

Et donc $\overline{Y}_{n+1} \hookrightarrow \mathcal{C}_0$.

Par le principe de récurrence, pour tout $n \in \mathbf{N}^*$, $\overline{Y}_n \hookrightarrow \mathcal{C}_0$.

10.d. Les Y_i n'admettent pas d'espérance¹², donc la loi faible des grands nombres ne s'applique pas à $(\overline{Y}_n)_{n \geq 1}$.

¹² Et donc pas de variance.

11.a. La matrice A est une matrice à 12000 lignes et 200 colonnes. Si l'on calcule la moyenne de chacune des lignes, on obtient 12000 simulations de \overline{Y}_{200} .

```
1 N = 12000 ; n=200 ;
2 A = cauchy(N,n) ;
3 MoyEmp = sum(A,'c')/200
```

11.b. Le premier histogramme donne approximativement la répartition des valeurs prises par \overline{Y}_{200} .

Le deuxième histogramme donne la répartition des valeurs prises par Y_1 , puisque la première colonne de A contient 12 000 simulations de Y_1 .

Ces deux histogrammes sont très proches, et donc ceci semble bien confirmer le fait que Y_1 et \overline{Y}_{200} suivent toutes deux la même loi, qui est donc la loi \mathcal{C}_0 puisque $Y_1 \hookrightarrow \mathcal{C}_0$.

Remarque : notons que le sujet aurait également pu nous demander d'écrire la fonction cauchy en utilisant la méthode d'inversion vue en TP.

En effet, nous connaissons la fonction de répartition de X est $F_X(x) = \frac{1}{2} + \frac{\text{Arctan}(x)}{\pi}$.

Et donc sa bijection réciproque est $F_X^{-1} : y \mapsto \tan\left(\pi\left(y - \frac{1}{2}\right)\right)$.

Et donc si U suit la loi uniforme sur $]0, 1[$ (qui se simule en Scilab à l'aide de rand), $F_X^{-1}(U)$ suit la loi de Cauchy de paramètre 0.

Partie IV. La loi de la médiane empirique.

12.a. L'événement B est réalisé si et seulement si l'une des variables X_1, \dots, X_{2n+1} prend une valeur dans $]x, x+h]$, que n prennent une valeur inférieure ou égale à x et que les n dernières prennent une valeur supérieure strictement à $x+h$.

Or, il y a $\binom{2n+1}{n}$ manières de choisir les variables inférieures à x , et pour chacune d'entre elles, la probabilité qu'elle soit inférieure à x vaut $F_X(x)$.

Une fois choisies ces n variables, il y a $n+1$ manières de choisir la variable prenant ses valeurs dans $]x, x+h]$, ce qui se produit avec probabilité $F_X(x+h) - F_X(x)$.

Hasard

Pour obtenir la «vraie» répartition des valeurs prises par \overline{Y}_{200} , il faudrait tracer sa fonction de répartition. Ici, nous nous basons sur un échantillon (de 12 000 simulations), et donc il y a une part de hasard dans le résultat : il se pourrait que nous ayons simulé un «mauvais» échantillon et que l'histogramme obtenu ne reflète en rien les valeurs prises par \overline{Y}_{200} . La grande taille de l'échantillon rend ceci assez peu probable.

Enfin, la probabilité que chacune des n autres variables prenne une valeur supérieure à $x + h$ est $(1 - F_X(x + h))^n$.
Ainsi, par indépendance des X_i ,

$$P(B) = (n + 1) \binom{2n + 1}{n} (F_X(x))^n (F_X(x + h) - F_X(x)) (1 - F_X(x + h))^n .$$

Plus rigoureusement : la démonstration ci-dessus me satisfait totalement, et devrait satisfaire n'importe quel correcteur, même si elle n'est pas totalement rigoureuse. Essayons d'en donner une preuve un peu plus formalisée, que le lecteur déjà convaincu pourra passer.

Quant au lecteur en quête de rigueur, il pourra essayer de comprendre cette preuve... et se convaincre qu'il vaut finalement mieux adopter la rédaction ci-dessus !

Notons I l'ensemble défini par

$$I = \{(A, B, C) \in \mathcal{P}(\llbracket 1, 2n + 1 \rrbracket)^3 / A \cap B = B \cap C = A \cap C = \emptyset, \text{Card}(A) = \text{Card}(C) = n\} .$$

Autrement dit, un élément (A, B, C) de I est une partition de $\llbracket 1, 2n + 1 \rrbracket$ en trois sous-ensembles deux à deux disjoints, A et C contenant n éléments chacun et B contenant donc le dernier élément de $\llbracket 1, 2n + 1 \rrbracket$.

L'idée est que A doit correspondre au numéros des variables X_1, \dots, X_{2n+1} inférieure à x , le seul élément de B est le numéro de la variable comprise entre x et $x + h$ et les éléments de C sont les numéros des variables plus grandes que $x + h$.

Alors

$$B = \bigcup_{(A, B, C) \in I} \left(\bigcap_{i \in A} [X_i \leq x] \cap \bigcap_{j \in B} [x < X_j \leq x + h] \cap \bigcap_{\ell \in C} [X_\ell > x + h] \right),$$

où l'union est disjointe (i.e. les événements de l'union sont deux à deux incompatibles).

Or, pour $(A, B, C) \in I$ fixé, par indépendance de X_1, \dots, X_{2n+1} , on a

$$\begin{aligned} P\left(\bigcap_{i \in A} [X_i \leq x] \cap \bigcap_{j \in B} [x < X_j \leq x + h] \cap \bigcap_{\ell \in C} [X_\ell > x + h]\right) &= \prod_{i \in A} P(X_i \leq x) \times P(x < X_j < x + h) \times \prod_{\ell \in C} P(X_\ell > x + h) \\ &= (F_X(x))^n (F_X(x + h) - F_X(x)) (1 - F_X(x + h))^n . \end{aligned}$$

Et donc, d'après l'incompatibilité mentionnée plus haut,

$$P(B) = \sum_{(A, B, C) \in I} (F_X(x))^n (F_X(x + h) - F_X(x)) (1 - F_X(x + h))^n = \text{Card}(I) (F_X(x))^n (F_X(x + h) - F_X(x)) (1 - F_X(x + h))^n .$$

Il ne nous reste donc qu'à déterminer le cardinal de I .

Mais, choisir un élément de I , c'est choisir les n éléments de A : il y a $\binom{2n + 1}{n}$ manières de le faire.

Puis une fois les éléments de A choisis, il faut choisir, parmi les $n + 1$ éléments restants, celui qui ira dans B : il y a $n + 1$ choix possibles.

Et alors automatiquement, les n éléments restants forment C .

Et donc $\text{Card}(I) = (n + 1) \binom{2n + 1}{n}$, de sorte que

$$P(B) = (n + 1) \binom{2n + 1}{n} (F_X(x))^n (F_X(x + h) - F_X(x)) (1 - F_X(x + h))^n .$$

12.b. Puisque $B \subset A$, $P(B) \leq P(A)$ et donc $P(A) - P(B) \geq 0$.

D'autre part, par la formule des probabilités totales appliquée au système complet d'événements $\{[Z = 0], [Z = 1], [Z \geq 2]\}$, on a

$$P(A) = \underbrace{P(A \cap [Z = 0])}_{=0} + P(A \cap [Z = 1]) + P(A \cap [Z = 2]) = P(B) + P(A \cap [Z = 2]).$$

Dénombrement

Notons que, comme souvent en dénombrement, il y a plusieurs moyens d'arriver au résultat. Il était par exemple possible de choisir d'abord une variable parmi les $2n + 1$ (celle qui prend sa valeur entre x et $x + h$), puis d'en choisir n parmi les $2n$ restantes.

On obtenait alors $(2n + 1) \binom{2n}{n}$ au lieu de $(n + 1) \binom{2n + 1}{n}$. Fort heureusement, ces deux nombres sont égaux, et donc le résultat ne dépend pas de la méthode choisie !

Détails

Les événements A et $[Z = 0]$ sont incompatibles car on ne peut avoir à la fois la $\overline{X_{n+1}}$ qui prend une valeur dans $[x, x + h]$ et aucune des X_i qui prend ses valeurs dans le même intervalle.

Et donc $P(A) - P(B) = P(A \cap [Z \geq 2]) \leq P(Z \geq 2)$. Mais, il est aisé de calculer $P(Z \geq 2)$: il y a $\binom{2n+1}{2} = n(2n+1)$ manières de choisir 2 variables parmi $X_1, X_2, \dots, X_{2n+1}$, et pour chacune de ces variables, la probabilité qu'elle prenne une valeur comprise entre x et $x+h$ est $F_X(x+h) - F_X(x)$.

Ainsi, $P(Z \geq 2) = n(2n+1)(F_X(x+h) - F_X(x))^2$.

Et donc, en posant $K = n(2n+1)$, on a bien

$$0 \leq P(A) - P(B) \leq K(F_X(x+h) - F_X(x))^2.$$

12.c. Notons $F_{\widehat{X}_{n+1}}$ la fonction de répartition de la variable aléatoire \widehat{X}_{n+1} .

Alors $P(A) = P(x < \widehat{X}_{n+1} \leq x+h) = F_{\widehat{X}_{n+1}}(x+h) - F_{\widehat{X}_{n+1}}(x)$.

Et donc, en divisant l'inégalité de la question 12.b par $|h| > 0$, il vient

$$0 \leq \left| \frac{F_{\widehat{X}_{n+1}}(x+h) - F_{\widehat{X}_{n+1}}(h)}{h} - \frac{P(B)}{h} \right| \leq \left| K \frac{F_X(x+h) - F_X(x)}{h} (F_X(x+h) - F_X(h)) \right|.$$

Or, lorsque $h \rightarrow 0$, $\frac{F_X(x+h) - F_X(x)}{h} \xrightarrow{h \rightarrow 0} F'_X(x) = f_X(x)$.

Mais d'autre part, F_X étant continue, $\lim_{h \rightarrow 0} F_X(x+h) - F_X(x) = 0$ et donc

$$\lim_{h \rightarrow 0} K \frac{F_X(x+h) - F_X(x)}{h} (F_X(x+h) - F_X(h)) = 0.$$

D'après le théorème des gendarmes, on a donc

$$\lim_{h \rightarrow 0} \frac{F_{\widehat{X}_{n+1}}(x+h) - F_{\widehat{X}_{n+1}}(h)}{h} - \frac{P(B)}{h} = 0.$$

Mais d'après la question 12.a,

$$\frac{P(B)}{h} = (n+1) \binom{2n+1}{n} (F_X(x))^n \frac{F_X(x+h) - F_X(x)}{h} (1 - F_X(x))^n \xrightarrow{h \rightarrow 0} (n+1) \binom{2n+1}{n} (F_X(x))^n (1 - F_X(x))^n f_X(x).$$

On en déduit donc que

$$\lim_{h \rightarrow 0} \frac{F_{\widehat{X}_{n+1}}(x+h) - F_{\widehat{X}_{n+1}}(h)}{h} = (n+1) \binom{2n+1}{n} (F_X(x))^n (1 - F_X(x))^n f_X(x).$$

Ceci prouve que $F_{\widehat{X}_{n+1}}$ est dérivable sur \mathbf{R} , et que sa dérivée est $x \mapsto (n+1) \binom{2n+1}{n} (F_X(x))^n (1 - F_X(x))^n f_X(x)$.

Cette dérivée est continue car F_X et f_X sont continues, et donc $F_{\widehat{X}_{n+1}}$ est une fonction de

classe \mathcal{C}^1 sur \mathbf{R} , ce qui prouve que \widehat{X}_{n+1} est une variable à densité.

L'une de ses densités en est alors donnée par

$$f_{\widehat{X}_{n+1}} : x \mapsto (n+1) \binom{2n+1}{n} (F_X(x))^n (1 - F_X(x))^n f_X(x).$$

13.a. Au voisinage de $+\infty$, $f_X(x) \underset{x \rightarrow +\infty}{\sim} \frac{1}{\pi} \frac{1}{x^2}$ et $F_X(x) \underset{x \rightarrow +\infty}{\sim} 1$.

D'autre part, pour $x > \theta$,

$$\begin{aligned} 1 - F_X(x) &= \frac{1}{2} - \frac{\text{Arctan}(x - \theta)}{\pi} = \frac{1}{2} - \frac{1}{\pi} \left(\frac{\pi}{2} - \text{Arctan} \left(\frac{1}{x - \theta} \right) \right) \\ &= \frac{1}{\pi} \text{Arctan} \left(\frac{1}{x - \theta} \right) \\ &\underset{x \rightarrow +\infty}{\sim} \frac{1}{\pi} \frac{1}{x - \theta} \underset{x \rightarrow +\infty}{\sim} \frac{1}{\pi} \frac{1}{x}. \end{aligned}$$

Et donc $(1 - F_X(x))^n \underset{x \rightarrow +\infty}{\sim} \frac{1}{\pi^n} \frac{1}{x^n}$.

Par produit d'équivalents, on obtient donc

$$x f_{\widehat{X}_{n+1}}(x) \underset{x \rightarrow +\infty}{\sim} x(n+1) \binom{2n+1}{n} \frac{1}{\pi^n} \frac{1}{x^n} \frac{1}{\pi} \frac{1}{x^2} = (n+1) \binom{2n+1}{n} \frac{1}{\pi^{n+1}} \frac{1}{x^{n+1}}.$$

Rappel

F_X est \mathcal{C}^1 sur \mathbf{R} tout entier, et

$$F'_X = f_X.$$

C'est le résultat de la question 6.d.

Équivalent

On a $\text{Arctan}(u) \underset{u \rightarrow 0}{\sim} u$, ce qui s'obtient à l'aide du développement limité d'ordre 1 de l'arctangente en 0.

13.b. La fonction $x \mapsto x f_{\widehat{X}_{n+1}}(x)$ est continue sur \mathbf{R} .

Au voisinage de $+\infty$, l'équivalent de la question précédente prouve que $\int_1^{+\infty} x f_{\widehat{X}_{n+1}}(x) dx$ converge, par comparaison à une intégrale de Riemann convergente.

Afin d'étudier $f_{\widehat{X}_{n+1}}(x)$ au voisinage de $-\infty$, remarquons que $f_X(2\theta - x) = f_X(x)$.

Alors en procédant au changement de variable $u = 2\theta - t$, il vient pour $x \in \mathbf{R}$,

$$\begin{aligned} F_X(x) &= \int_{-\infty}^x f_X(t) dt = \int_{-\infty}^x f_X(2\theta - t) dt = \int_{2\theta-x}^{+\infty} f_X(u) du \\ &= P(X \geq 2\theta - x) = 1 - F_X(2\theta - x) \underset{x \rightarrow -\infty}{\sim} \frac{1}{\pi} \frac{1}{-x}. \end{aligned}$$

D'autre part, $1 - F_X(x) \xrightarrow{x \rightarrow -\infty} 1$ et donc

$$x f_{\widehat{X}_{n+1}}(x) \underset{x \rightarrow -\infty}{\sim} (n+1) \binom{2n+1}{n} \frac{1}{\pi^{n+1}} \frac{(-1)^n}{x^{n+1}}.$$

Puisque $\int_{-\infty}^{-1} \frac{dx}{x^{n+1}}$ converge, par critère de comparaison pour les fonctions de signe constant¹³, $\int_{-\infty}^{-1} x f_{\widehat{X}_{n+1}}(x) dx$ converge.

Et donc $\int_{-\infty}^{+\infty} x f_{\widehat{X}_{n+1}}(x) dx$ converge, de sorte que \widehat{X}_{n+1} admet une espérance.

13.c. La variable aléatoire \widehat{X}_{n+1} est égale à¹⁴ $g_{n+1}(X_1, X_2, \dots, X_{2n+1})$, où g_{n+1} ne dépend pas de θ : c'est donc un estimateur de θ .

Par transformation affine, une densité de $\widehat{X}_{n+1} - \theta$ est

$$\begin{aligned} h_{n+1} : x \mapsto f_{\widehat{X}_{n+1}}(x + \theta) &= (n+1) \binom{2n+1}{n} (F_X(x + \theta))^n (1 - F_X(x + \theta))^n f_X(x + \theta) \\ &= (n+1) \binom{2n+1}{n} \left(\frac{1}{2} + \frac{\text{Arctan}(x)}{\pi} \right)^n \left(\frac{1}{2} - \frac{\text{Arctan}(x)}{\pi} \right)^n \frac{1}{\pi} \frac{1}{1+x^2}. \end{aligned}$$

Cette densité est alors paire, et donc $x h_{n+1}(x)$ est impaire.

Par conséquent¹⁵,

$$E(\widehat{X}_{n+1} - \theta) = \int_{-\infty}^{+\infty} x h_{n+1}(x) dx = 0.$$

Ainsi, $E(\widehat{X}_{n+1}) = \theta$, et donc \widehat{X}_{n+1} est un estimateur sans biais de θ .

13.d. Nous savons que \widehat{X}_{n+1} admet une variance si et seulement si elle admet un moment d'ordre 2.

D'après le théorème de transfert, c'est le cas si et seulement si $\int_{-\infty}^{+\infty} x^2 f_{\widehat{X}_{n+1}}(x) dx$ converge.

Mais, au voisinage de $\pm\infty$, $x^2 f_{\widehat{X}_{n+1}}(x) \underset{x \rightarrow \pm\infty}{\sim} (n+1) \binom{2n+1}{n} \frac{1}{\pi^{n+1}} \frac{1}{x^n}$.

Et donc, par comparaison à une intégrale de Riemann, \widehat{X}_{n+1} admet une variance si et seulement si $n \geq 2$.

14.a. Nous avons déjà dit¹⁶ que $f_X(2\theta - t) = f_X(t)$ et que donc $F_X(x) = 1 - F_X(2\theta - x)$. Donc il vient

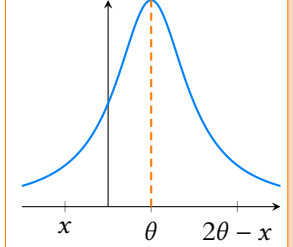
$$\begin{aligned} f_{\widehat{X}_{n+1}}(2\theta - x) &= (n+1) \binom{2n+1}{n} (F_X(2\theta - x))^n (1 - F_X(2\theta - x))^n f_X(2\theta - x) \\ &= (n+1) \binom{2n+1}{n} (1 - F_X(x))^n (1 - (1 - F_X(x)))^n f_X(x) \\ &= (n+1) \binom{2n+1}{n} (F_X(x))^n (1 - F_X(x))^n f_X(x) = f_{\widehat{X}_{n+1}}(x). \end{aligned}$$

Par conséquent, on a

$$P(|\widehat{X}_{n+1} - \theta| \geq \varepsilon) = P(\widehat{X}_{n+1} - \theta \geq \varepsilon) + P(\widehat{X}_{n+1} - \theta \leq -\varepsilon)$$

Graphiquement

Ceci signifie que la courbe de f_X est symétrique par rapport à la droite $x = \theta$.



¹³ $\frac{(-1)^n}{x^{n+1}}$ est négatif sur \mathbf{R}_- .

¹⁴ Avec les notations du début de la partie IV.

¹⁵ On sait déjà que l'intégrale converge, hypothèse indispensable pour utiliser l'imparité.

¹⁶ À la question 13.b.

$$\begin{aligned}
 &= P(\widehat{X}_{n+1} \geq \theta + \varepsilon) + P(\widehat{X}_{n+1} \leq \theta - \varepsilon) \\
 &= \int_{\theta+\varepsilon}^{+\infty} f_{\widehat{X}_{n+1}}(t) dt + F_{\widehat{X}_{n+1}}(\theta - \varepsilon).
 \end{aligned}$$

Mais, dans l'intégrale, procédons au changement de variable $u = 2\theta - t$. Il vient alors

$$\begin{aligned}
 \int_{\theta+\varepsilon}^{+\infty} f_{\widehat{X}_{n+1}}(t) dt &= \int_{\theta+\varepsilon}^{+\infty} f_{\widehat{X}_{n+1}}(2\theta - t) dt \\
 &= \int_{-\infty}^{\theta-\varepsilon} f_{\widehat{X}_{n+1}}(u) du \\
 &= F_{\widehat{X}_{n+1}}(\theta - \varepsilon).
 \end{aligned}$$

Et donc

$$P(|\widehat{X}_{n+1} - \theta| \geq \varepsilon) = 2F_{\widehat{X}_{n+1}}(\theta - \varepsilon).$$

14.b. Au vu du résultat de la question précédente, il va nous falloir prouver que $F_{\widehat{X}_{n+1}}(\theta - \varepsilon) \xrightarrow[n \rightarrow +\infty]{} 0$.

Rappelons donc que $F_{\widehat{X}_{n+1}}(\theta - \varepsilon) = \int_{-\infty}^{\theta-\varepsilon} (n+1) \binom{2n+1}{n} (F_X(x))^n (1 - F_X(x))^n f_X(x) dx$.

Commençons par déterminer un équivalent de $\binom{2n+1}{n}$.

Par définition, on a $\binom{2n+1}{n} = \frac{(2n+1)!}{n!(n+1)!}$.

En utilisant la formule de Stirling prouvée dans la partie I, il vient donc

$$\binom{2n+1}{n} \underset{n \rightarrow +\infty}{\sim} \frac{(2n+1)^{2n+1} e^{-2n-1} \sqrt{2\pi(2n+1)}}{n^n e^{-n} \sqrt{2\pi n} (n+1)^{n+1} e^{-n-1} \sqrt{2\pi(n+1)}} \underset{n \rightarrow +\infty}{\sim} \frac{(2n+1)^{2n+1}}{n^n (n+1)^{n+1}} \frac{1}{\sqrt{\pi n}}.$$

Or, $(2n+1)^{2n+1} = (2n)^{2n+1} \left(1 + \frac{1}{2n}\right)^{2n+1}$ avec

$$\left(1 + \frac{1}{2n}\right)^{2n+1} = e^{(2n+1)\ln(1+\frac{1}{2n})} = e^{(2n+1)(\frac{1}{2n} + o(\frac{1}{n}))} = e^{1 + \frac{1}{2n} + o(1)} \xrightarrow[n \rightarrow +\infty]{} e.$$

Et donc $(2n+1)^{2n+1} \underset{n \rightarrow +\infty}{\sim} 2^{2n+1} n^{2n+1} \times e$.

Le même type de calculs prouve que $(n+1)^{n+1} \underset{n \rightarrow +\infty}{\sim} e \times n^{n+1}$.

Et donc

$$\binom{2n+1}{n} \underset{n \rightarrow +\infty}{\sim} \frac{2^{2n+1} n^{2n+1} e}{n^n n^{n+1} e} \frac{1}{\sqrt{\pi n}} \underset{n \rightarrow +\infty}{\sim} \frac{2^{2n+1}}{\sqrt{\pi n}}.$$

Une rapide¹⁷ étude de la fonction $t \mapsto t(1-t)$ prouve que pour tout $t \in [0, 1]$, $t(1-t) \leq \frac{1}{4}$, avec égalité si et seulement si $t = \frac{1}{2}$.

En particulier pour tout $x \in \mathbf{R}$, $F_X(x)(1 - F_X(x)) \leq \frac{1}{4}$, avec égalité si et seulement si $F_X(x) = \frac{1}{2}$, soit si et seulement si $x = \theta$.

Comme la fonction $x \mapsto F_X(x)(1 - F_X(x))$ est croissante sur $]-\infty, \theta - \varepsilon[$, elle y est alors majorée par $A = F_X(\theta - \varepsilon)(1 - F_X(\theta - \varepsilon)) < \frac{1}{4}$.

Et donc il vient

$$0 \leq F_{\widehat{X}_{n+1}}(\theta - \varepsilon) \leq (n+1) \binom{2n+1}{n} A^n \int_{-\infty}^{\theta-\varepsilon} f_X(x) dx = A^n (n+1) \binom{2n+1}{n} F_X(\theta - \varepsilon).$$

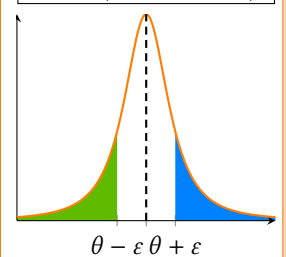
Mais en utilisant les équivalents précédemment trouvés, il vient

$$A^n (n+1) \binom{2n+1}{n} \underset{n \rightarrow +\infty}{\sim} \frac{2^{2n+1} n}{\sqrt{\pi n}} A^n = \frac{2}{\sqrt{\pi}} (4A)^n \sqrt{n}.$$

Graphiquement

Cette relation s'interprète graphiquement en termes d'aires. On pourra notamment comparer avec ce que l'on connaît pour une loi normale.

$$\begin{aligned}
 &\text{■ } P(\widehat{X}_{n+1} \leq \theta - \varepsilon) \\
 &\text{■ } P(\widehat{X}_{n+1} \geq \theta + \varepsilon)
 \end{aligned}$$



¹⁷ Et classique !

Or, $0 < A < \frac{1}{4}$ et donc $0 < 4A < 1$, de sorte que, par croissances comparées,

$$\sqrt{n}(4A)^n \xrightarrow{n \rightarrow +\infty} 0.$$

Et donc, par le théorème des gendarmes, $\lim_{n \rightarrow +\infty} F_{\widehat{X}_{n+1}}(\theta - \varepsilon) = 0$.

D'après la majoration de la question précédente, on a donc $\lim_{n \rightarrow +\infty} P(|\widehat{X}_{n+1} - \theta| \geq \varepsilon) = 0$:

$(\widehat{X}_{n+1})_{n \in \mathbb{N}^*}$ converge en probabilité vers θ .

14.c. Il est sûrement possible de prouver cette convergence en revenant à la définition de la convergence en loi et en utilisant le fait que $\lim_{n \rightarrow +\infty} P(|\widehat{X}_{n+1} - \theta| \geq \varepsilon) = 0$.

Soyons plus astucieux, et utilisons le théorème de Slutsky : on a $0 \xrightarrow{\mathcal{L}} 0$ et $\widehat{X}_{n+1} \xrightarrow{P} \theta$, où θ est une constante.

Alors $0 + \widehat{X}_{n+1} \xrightarrow{\mathcal{L}} 0 + \theta$: $(\widehat{X}_{n+1})_n$ converge en loi vers θ .

15.a. À la question 13.c, nous avons déjà donné une densité h_{n+1} de $\widehat{X}_{n+1} - \theta$.

Par transformation affine, une densité de $W_{n+1} = \frac{2\sqrt{2n+1}}{\pi}(\widehat{X}_{n+1} - \theta)$ est donc

$$\begin{aligned} f_{W_{n+1}}(x) &= \frac{\pi}{2\sqrt{2n+1}} h_{n+1}\left(\frac{\pi}{2\sqrt{2n+1}}x\right) \\ &= \frac{n+1}{2\sqrt{2n+1}} \binom{2n+1}{n} \left(\frac{1}{2} + \frac{\text{Arctan}\left(\frac{\pi}{2\sqrt{2n+1}}x\right)}{\pi}\right)^n \left(\frac{1}{2} - \frac{\text{Arctan}\left(\frac{\pi}{2\sqrt{2n+1}}x\right)}{\pi}\right)^n \frac{1}{1 + \left(\frac{\pi}{2\sqrt{2n+1}}x\right)^2} \\ &= \frac{n+1}{2\sqrt{2n+1}} \binom{2n+1}{n} \left[\frac{1}{4} - \frac{1}{\pi^2} \left(\text{Arctan}\left(\frac{\pi}{2\sqrt{2n+1}}x\right)\right)^2\right]^n \left(1 + \frac{\pi^2 x^2}{4(2n+1)}\right)^{-1}. \end{aligned}$$

Plus généralement

Le même raisonnement prouve qu'une suite de variables aléatoires qui converge en probabilités vers une constante converge en loi vers cette même constante.

Un résultat classique (mais hors programme) affirme qu'une suite de variables aléatoires converge en probabilité vers une constante si et seulement si elle converge en loi vers cette constante.

Détails

$$(a+b)(a-b) = a^2 - b^2.$$

15.b. Commençons par remarquer que

$$\begin{aligned} \left[\frac{1}{4} - \frac{1}{\pi^2} \left(\text{Arctan}\left(\frac{\pi}{2\sqrt{2n+1}}x\right)\right)^2\right]^{2^n} &= \frac{1}{4^n} \left[1 - \frac{4}{\pi^2} \left(\text{Arctan}\left(\frac{\pi}{2\sqrt{2n+1}}x\right)\right)^2\right]^{2^n} \\ &= \frac{1}{4^n} \exp\left(n \ln\left(1 - \frac{4}{\pi^2} \left(\text{Arctan}\left(\frac{\pi}{2\sqrt{2n+1}}x\right)\right)^2\right)\right). \end{aligned}$$

Mais $\frac{\pi}{2\sqrt{2n+1}}x \xrightarrow{n \rightarrow +\infty} 0$, de sorte que $\text{Arctan}\left(\frac{\pi}{2\sqrt{2n+1}}x\right) \xrightarrow{n \rightarrow +\infty} 0$ et donc

$$\begin{aligned} n \ln\left(1 - \frac{4}{\pi^2} \left(\text{Arctan}\left(\frac{\pi}{2\sqrt{2n+1}}x\right)\right)^2\right) &\underset{n \rightarrow +\infty}{\sim} -n \frac{4}{\pi^2} \left(\text{Arctan}\left(\frac{\pi}{2\sqrt{2n+1}}x\right)\right)^2 \\ &\underset{n \rightarrow +\infty}{\sim} -\frac{4n}{\pi^2} \frac{\pi^2}{4(2n+1)} x^2 \\ &\underset{n \rightarrow +\infty}{\sim} -\frac{x^2}{2} \xrightarrow{n \rightarrow +\infty} -\frac{x^2}{2}. \end{aligned}$$

Enfin, par continuité de l'exponentielle,

$$\left[1 - \frac{4}{\pi^2} \left(\text{Arctan}\left(\frac{\pi}{2\sqrt{2n+1}}x\right)\right)^2\right]^{2^n} \xrightarrow{n \rightarrow +\infty} \exp\left(-\frac{x^2}{2}\right).$$

D'autre part, nous connaissons déjà des équivalents des autres termes, donc

$$f_{W_{n+1}}(x) \underset{n \rightarrow +\infty}{\sim} \frac{n}{2\sqrt{2n}} \frac{2^{2n+1}}{\sqrt{\pi n}} \frac{1}{4^n} \exp\left(-\frac{x^2}{2}\right) \underset{n \rightarrow +\infty}{\sim} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right) \xrightarrow{n \rightarrow +\infty} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right).$$

Rappel

Il a été prouvé plus tôt que

$$\text{Arctan}(x) \underset{x \rightarrow 0}{\sim} x.$$

Remarque

Le résultat admis dans l'énoncé, qui s'appelle le lemme de Scheffé, nous dit que pour des variables à densité, il est possible de prouver la convergence en loi en étudiant la convergence de la suite des densités plutôt que celle de la suite des fonctions de répartition (qui est la définition de la convergence en loi).

15.c. Puisque W_{n+1} converge en loi vers T ,

$$\lim_{n \rightarrow +\infty} P(-t_\alpha \leq W_{n+1} \leq t_\alpha) = P(-t_\alpha \leq T \leq t_\alpha) = \Phi(t_\alpha) - \Phi(-t_\alpha) = \Phi(t_\alpha) - (1 - \Phi(t_\alpha)) = 1 - \alpha.$$

Mais d'autre part,

$$\begin{aligned} P(-t_\alpha \leq W_{n+1} \leq t_\alpha) &= P\left(-t_\alpha \leq \frac{2\sqrt{2n+1}}{\pi} (\widehat{X}_{n+1} - \theta) \leq t_\alpha\right) \\ &= P\left(-\frac{\pi}{2\sqrt{2n+1}} t_\alpha \leq \theta - \widehat{X}_{n+1} \leq \frac{\pi}{2\sqrt{2n+1}} t_\alpha\right) \\ &= P\left(\widehat{X}_{n+1} - \frac{\pi}{2\sqrt{2n+1}} t_\alpha \leq \theta \leq \widehat{X}_{n+1} + \frac{\pi}{2\sqrt{2n+1}} t_\alpha\right). \end{aligned}$$

Ainsi, nous venons de prouver que

$$\lim_{n \rightarrow +\infty} P\left(\widehat{X}_{n+1} - \frac{\pi}{2\sqrt{2n+1}} t_\alpha \leq \theta \leq \widehat{X}_{n+1} + \frac{\pi}{2\sqrt{2n+1}} t_\alpha\right) = 1 - \alpha$$

et donc $\left[\widehat{X}_{n+1} - \frac{\pi}{2\sqrt{2n+1}} t_\alpha, \widehat{X}_{n+1} + \frac{\pi}{2\sqrt{2n+1}} t_\alpha\right]$ est un intervalle de confiance asymptotique pour θ , au niveau de confiance $1 - \alpha$.

Remarquons que le milieu de cet intervalle est bien \widehat{X}_{n+1} .

16. Chaque ligne de A correspond à une réalisation du $(2n+1)$ -échantillon $(X_1, X_2, \dots, X_{2n+1})$. Si on trie cette matrice ligne par ligne, alors, à chaque fois, la $n+1$ -ème colonne contiendra une réalisation de \widehat{X}_{n+1} .

```
1 A = cauchy(p, 2*n+1) ;
2 S = gsort(A, 'c') ;
3 MedianeEmp = S(:, n+1) ;
4 W = MedianeEmp * 2*sqrt(2*n+1)/%pi ;
```

TRAVAUX PRATIQUES DE Scilab

Liste des commandes Scilab exigibles.	1046
TP 1 : Manipulation de matrices.	1048
Correction	1052
TP 2 : Simulation de lois discrètes	1056
Correction	1060
TP 3 : Statistiques à une variable.	1066
Correction	1069
TP 4 : Chaînes de Markov I	1073
Correction	1076
TP 5 : Chaînes de Markov II	1081
Correction	1085
TP 6 : Simulation de lois à densité	1087
Correction	1091
TP 7 : Statistiques bivariées.	1094
Correction	1098
TP 8 : Révisions	1103
Correction	1107
TP 9 : Fonctions de plusieurs variables	1115
Correction	1118
TP 10 : Estimation, méthodes de Monte-Carlo	1121
Correction	1124

25.1 RÉSUMÉ DES COMMANDES AU PROGRAMME

Voici la liste des commandes Scilab figurant au programme officiel.

Pour certaines d'entre elles, le fonctionnement doit être connu par cœur.

Pour d'autres, il est bon de les avoir déjà utilisées, mais en cas d'utilisation au concours, leur fonctionnement devrait être rappelé (au moins brièvement) par l'énoncé.

Enfin, le programme mentionne d'autres commandes sans qu'il soit clair qu'elles puissent être ou non utilisées dans un sujet de concours....

PROGRAMME DE PREMIÈRE ANNÉE

Les instructions conditionnelles **if ... then ...else ... end** et **if ... then ... end** pour exécuter une (ou des) instruction(s) seulement si une certaine condition est vérifiée.

Les boucles usuelles : **for ... end** et **while ... end**. Rappelons que la première permet de répéter une instruction un nombre prédéfini de fois, alors que la seconde permet de répéter une instruction tant qu'une certaine condition est vérifiée (et donc est particulièrement utile lorsqu'on ne sait pas à l'avance combien de fois on souhaite répéter l'instruction).

Les commandes **input** et **disp** pour demander à l'utilisateur d'entrer un nombre ou d'afficher le contenu d'une variable. Les commandes **%pi** et **%e** pour utiliser les nombres π et e .

- **log**, **exp**, **floor**, **abs**, **sqrt**, **sin** et **cos** : fonctions numériques usuelles : logarithme népérien, exponentielle, partie entière, valeur absolue, racine carrée, sinus et cosinus.
- **rand** : simule un nombre suivant une loi uniforme (à densité) sur $[0, 1]$.
- **rank** : calcule le rang d'une matrice
- **inv** : calcule l'inverse d'une matrice carrée inversible.
- **size** : retourne la taille d'une matrice (voir TP 1).
- **linspace** : `linspace(a,b,n)` retourne un vecteur de taille n , «découpant» l'intervalle $[a, b]$ en $n - 1$ segments de même longueur. Est particulièrement utile pour utiliser `plot2d` ou `fplot3d` (voir TP 9).
- **grand** : simule une (ou plusieurs) variables aléatoires suivant une loi usuelle. Les différents paramètres possibles sont 'uin' (uniforme discrète), 'poi' (poisson), 'geom' (géométrique), 'bin' (binomiale, et en particulier Bernoulli), 'unf' (uniforme à densité), 'exp' (exponentielle) et 'nor' (normale).
- **find** : permet de chercher les éléments vérifiant une certaine condition dans une matrice. Souvent couplée à la commande `length` lorsque le but est de déterminer combien de tels éléments se trouvent dans la matrice (voir TP 1).
- **bar** : trace un diagramme en bâtons.

PROGRAMME DE SECONDE ANNÉE

Commandes à connaître par cœur

- **sum** : calcule la somme de tous les coefficients d'une matrice (voir TP 1).
- **cumsum** : calcule les sommes partielles (ou sommes cumulées) d'un vecteur. Particulièrement utile pour l'étude de séries (voir TP 1).
- **max** : renvoie le plus grand coefficient d'une matrice (voir TP 1)
- **min** : renvoie le plus petit coefficient d'une matrice (voir TP 1)
- **zeros** : crée la matrice nulle. Plus précisément, `zeros(n,p)` renvoie la matrice nulle à n lignes et p colonnes (voir TP 1).
- **ones** : crée une matrice dont tous les coefficients valent 1. Plus précisément, `ones(n,p)` renvoie la matrice à n lignes et p colonnes dont tous les coefficients valent 1 (voir TP 1).
- **eye** : crée la matrice identité (voir TP 1).
- **mean** : calcule la moyenne des coefficients d'une matrice, est équivalent à `sum(A)/length(A)` (voir TP 3).
- **spec** : détermine les valeurs propres d'une matrice et (dans le cas d'une matrice diagonalisable) une base de vecteurs propres (voir TP5).

Commandes qui doivent avoir été déjà manipulées

Ces commandes ont une syntaxe plus ou moins compliquée, qui devrait vous être rappelée si besoin.

- **cdfnor** : permet de faire divers calculs sur les lois normales. En particulier de calculer Φ avec l'option 'PQ' (voir TP 6) et Φ^{-1} avec l'option 'X' (voir TP 10).

- **plot2d** : trace une fonction d'une variable réelle ou une suite. Prend en paramètres deux vecteurs x et y de même taille. Le premier contient les abscisses des points de la courbe, le second les ordonnées. Sans plus de précisions, Scilab place les points en question et les relie par des segments. Peut aussi s'utiliser avec l'option -1 , afin de ne pas relier les points (par exemple pour représenter graphiquement une suite). Voir première année.
- **fplot2d** : trace une fonction d'une variable réelle. Son fonctionnement est similaire à celui de `fplot3d`. Voir les TP d'ECS1.
- **plot3d** : n'a pas été vue en TP. Fonctionne de la même manière que `plot2d`, mais pour des graphiques en 3D.
- **fplot3d** : trace le graphique d'une fonction de deux variables (voir TP 9).

Autres commandes mentionnées par le programme

- **median** : calcule la médiane d'une série de nombres (voir TP 3).
- **tabul** : crée la matrice des modalités-effectifs d'une série de nombres (voir TP 3).
- **dsearch** : regroupe des données statistiques par classes (voir TP 3).
- **stdev** ou **st_deviation** : calcule l'écart-type d'une série de nombres (voir TP 3).
- **corr** : calcule le coefficient de corrélation linéaire de deux séries de nombres (voir TP 7).
- **grand** avec l'option `markov` : permet de simuler n étapes d'une chaîne de Markov si on lui fournit la matrice de transition en paramètre (voir TP 5).

LES AUTRES COMMANDES DÉJÀ RENCONTRÉES DANS DES SUJETS DE CONCOURS

La syntaxe `function ... endfunction` ne figure pas au programme, mais a été utilisée plusieurs fois à Ecricome. Et notamment dans le sujet de 2015 où elle permettait de créer une fonction ne retournant pas un seul nombre, mais un couple de nombres (`function [x,k] = h(e)`).

Le sujet d'HEC 2016 utilisait les variables booléennes, qui sont des variables pouvant prendre les valeurs `%T` (pour true) ou `%F` (pour false), qui sont en quelque sorte des analogues aux variables de Bernoulli (pouvant prendre les valeurs 0 ou 1). Voir la question 5.b d'HEC 2016 pour un exemple d'utilisation.

TP 1 : MANIPULATION DE MATRICES

MANIPULATION EFFICACE DE MATRICES

Rappelons que dans SciLab, toutes les variables sont des matrices, y compris les variables stockant des valeurs réelles, qui sont alors des matrices de taille 1×1 .

Pour créer une matrice, on entre ses valeurs entre crochets, ligne par ligne, en séparant les valeurs d'une même ligne par une virgule et les colonnes par des points-virgules. Par exemple $A = [1, 2, 3 ; 2, 3, -1]$ permet de créer la matrice

$$\begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & -1 \end{pmatrix}.$$

On accède (et on peut modifier) au coefficient (i, j) d'une matrice A grâce à la commande $A(i, j)$. Ainsi, si on entre $A(2, 1) = -4$, la matrice A devient

$$\begin{pmatrix} 1 & 2 & 3 \\ -4 & 3 & -1 \end{pmatrix}.$$

La commande $\text{size}(A)$ retourne une matrice 1×2 , dont le premier coefficient est le nombre de lignes de A et le second est le nombre de colonnes de A .

Toujours avec l'exemple précédent, si l'on entre $s = \text{size}(A)$, alors s est la matrice égale à $(2 \ 3)$, signifiant que A possède deux lignes et trois colonnes.

Les opérations usuelles sur les matrices (somme, produit, puissance) sont obtenues à l'aide des symboles $+$, $*$, $^$.

La transposée de A est obtenue avec A' , le rang de A avec $\text{rank}(A)$, et dans le cas où A est inversible, son inverse A^{-1} est obtenue avec $\text{inv}(A)$.

Si a et b sont deux entiers avec $a < b$, la commande $a : b$ crée le vecteur ligne $(a \ a + 1 \ \dots \ b)$.

Il est également possible d'aller de 2 en 2, ou de -1 en -1 . Par exemple, entrez les commandes $0 : 2 : 10$ et $10 : -1 : 1$.

On peut obtenir la somme des éléments d'une matrice A avec $\text{sum}(A)$ et le produit des éléments de A avec $\text{prod}(A)$. Ainsi, si l'on veut calculer la somme des entiers de 1 à 1000, il suffit d'entrer $\text{sum}(1 : 1000)$.

EXERCICE 1 Écrire deux lignes de commande qui demandent un entier non nul n et affichent $n!$.

PD

La commande $\text{ones}(n, p)$ crée une matrice de taille $n \times p$ dont tous les coefficients valent 1. De même pour $\text{zeros}(n, p)$ qui crée la matrice nulle de taille $n \times p$.

Enfin, $\text{eye}(n, p)$ a pour effet de créer une matrice $n \times p$ dont tous les coefficients sont nuls, à l'exception de ceux de la diagonale valant 1 (en particulier, si $n = p$, cela crée la matrice identité I_n .)

EXERCICE 2 Que vaut le produit d'un vecteur colonne $\begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}$ par le vecteur ligne $(1 \ 1 \ \dots \ 1)$?

PD

En déduire une ligne de commande créant la matrice

$$\begin{pmatrix} 1 & 1 & \dots & 1 \\ 2 & 2 & \dots & 2 \\ \vdots & \vdots & & \vdots \\ 10 & 10 & \dots & 10 \end{pmatrix}$$

EXERCICE 3 En une seule ligne de commande, et sans boucle, créer la matrice 10×10 dont tous les coefficients diagonaux valent 0 et dont les autres coefficients valent 1. *Indication : écrire cette matrice comme somme de deux matrices très simples.*

PD

Il est possible de récupérer uniquement la i -ème ligne d'une matrice A à l'aide de la commande $A(i, :)$, qui retourne un vecteur ligne contenant la i -ème ligne de A . De même, $A(:, j)$ retourne la j -ème colonne de A .

Les commandes $\text{min}(A)$ et $\text{max}(A)$ retournent respectivement le plus petit et le plus grand des coefficients de A .

RAPPEL : LA SYNTAXE fonction ... endfunction

La commande `function` permet de créer des fonctions dans `Scilab`, à l'image des fonctions que nous manipulons en mathématiques : elles attendent un ou plusieurs paramètres, et retournent un ou plusieurs nombres.

Par exemple, regardons le code suivant :

```
1 function y =carre(x)
2     y = x*x
3 endfunction
```

Il a pour effet de créer une fonction nommée `carre`, qui attend qu'on lui passe un paramètre `x` et retournera la valeur de `y`. Et si l'on en croit la ligne 2, `y` doit être le carré de `x`.

Autrement dit, nous avons une fonction qui calcule le carré du nombre qu'on lui rentre.

Pour l'utiliser pour calculer le carré de 13, il suffit d'entrer `carre(13)` dans la console, qui nous affichera alors 169. Notons qu'à aucun moment nous n'avons utilisé les instructions `input` et `disp`, qui ne sont pas nécessaires si on utilise `function`.

L'avantage de cette syntaxe est qu'à partir du moment où elle a été définie, on peut utiliser plusieurs fois une fonction, y compris au sein d'une autre fonction. Par exemple :

```
1 function S = somme_carres(n)
2     S = 0
3     for i=1 :n
4         S = S + carre(i)
5     end
6 endfunction
```

crée une fonction `somme_carres` qui prend comme paramètre un entier `n` et retourne $\sum_{i=1}^n i^2$.

Une fonction peut même prendre plusieurs paramètres, comme par exemple :

```
1 function S = somme_puissances(n,k)
2     S = 0
3     for i=1 :n
4         S = S+i^k
5     end
6 endfunction
```

qui crée une fonction qui attend deux paramètres `n` et `k` et retourne $\sum_{i=1}^n i^k$.

Par exemple, pour calculer $\sum_{i=1}^{100} i^5$, il suffit d'entrer `somme_puissances(100,5)` dans la console.

EXERCICE 4 Compléter la fonction `min_ligne` afin qu'elle prenne comme paramètres une matrice `A` et un entier `i` et retourne le plus petit coefficient de la `i`-ème ligne de la matrice `A`.

```
1 function y = min_ligne(A,i)
2     y =
3 endfunction
```

Écrire de même une fonction `max_colonne`.

Il est possible d'appliquer une fonction définie sous `Scilab` à tous les coefficients d'une matrice. Par exemple, si `A` est une matrice dont tous les coefficients sont positifs, alors `sqrt(A)` est la matrice dont les coefficients sont les racines de ceux de `A`.

⚠ La fonction `log` désigne le logarithme népérien, et non le logarithme décimal. Il n'y a pas de fonction nommée `ln` dans `Scilab`.

EXERCICE 5

- Créer une matrice contenant les racines carrées des entiers de 1 à 100 et une autre contenant `[ln(1), ln(2), ..., ln(100)]`.

2. Compléter la fonction suivante afin qu'elle retourne un couple $[a, b]$ où $a = \sum_{i=1}^n \sqrt{i}$ et $b = \sum_{i=1}^n [\ln(i)]$.

```

1 fonction [a,b] = sommes(n)
2     a = ...
3     b = ...
4 endfunction

```

RECHERCHE D'ÉLÉMENTS DANS UNE MATRICE

Dans Scilab, une matrice A de taille $n \times p$ est en fait stockée comme un vecteur ligne de taille np , dont les coefficients sont d'abord ceux de la première colonne de A , puis ceux de la deuxième colonne, etc. Par exemple, si $A \in \mathcal{M}_{2,3}(\mathbf{R})$, $A(1)$ renvoie le coefficient $(1, 1)$ de A , $A(2)$ renvoie le coefficient $(2, 1)$ de A et $A(5)$ renvoie le coefficient $(1, 3)$ de A .

EXERCICE 6 Reprendre la matrice de l'exercice 2, et deviner ce que vont renvoyer les commandes $A(8)$, $A(37)$ et $A(99)$. Le vérifier. F

La commande `find` permet de trouver les éléments d'une matrice vérifiant une certaine condition. Par exemple, `find(A==0)` va retourner une matrice contenant les positions des coefficients de A qui sont nuls et `find(A>3)` va chercher lesquels sont plus grands que 3.

Tester cette commande en prenant $A=\text{zeros}(3, 4)$, $A=\text{eye}(3, 3)$ et $A=\text{eye}(4, 4)$, ainsi que sur d'autres matrices de votre choix.

La commande `length(A)` retourne le nombre de coefficients d'une matrice A (soit $n \times p$).

Combinée à la fonction `find`, elle permet de compter le nombre d'éléments vérifiant une condition donnée. Par exemple, `length(find(A == 0))` retourne le nombre de coefficients nuls d'une matrice A .

Notons qu'il existe un autre moyen, plus simple mais plus astucieux pour compter le nombre d'éléments d'une matrice vérifiant une certaine condition : au lieu de `length(find(A==0))` et `length(find(A>3))`, on peut utiliser les commandes `sum(A==0)` et `sum(A>3)`. Pour en comprendre le fonctionnement, vous pouvez essayer de commencer par entrer $A>3$.

EXERCICE 7 Entrez une matrice A de votre choix. Écrire une ligne de commande, qui affiche le nombre de coefficients de A égaux à $m = \max_{i,j} a_{i,j}$ AD

EXERCICE 8 Le nombre de véhicules passant sur une petite route de campagne en une journée est une variable aléatoire X qui suit une loi de Poisson $\mathcal{P}(5)$. AD

On peut simuler 1000 journées en entrant $A = \text{grand}(1, 1000, 'poi', 5)$. On obtient ainsi une matrice A qui contient 1000 nombres tirés suivant la loi $\mathcal{P}(5)$.

1. En une seule ligne de commande, calculer le nombre moyen de véhicules par jour durant ces 1000 jours. Recommencez plusieurs fois en recréant la matrice A . Que constatez-vous ? Était-ce prévisible ?
2. En une seule ligne de commande, évaluez la probabilité qu'une journée voie passer plus de 8 véhicules. *Question subsidiaire : sauriez-vous calculer la valeur exacte de cette probabilité ?*
3. Toujours en une seule ligne, déterminer le nombre de jours où la route a connu son minimum de fréquentation.

LES COMMANDES `cumsum` ET `cumprod`

Si A est une matrice, alors `cumsum(A)` retourne une matrice de même taille que A , et dont les coefficients sont les sommes cumulées des coefficients de A , en commençant par la première colonne, puis la deuxième, etc. Tester `cumsum(1:n)`, puis appliquer `cumsum` à un vecteur ligne, un vecteur colonne, puis une matrice de votre choix.

De même, on dispose d'une commande `cumprod` qui fonctionne sur le même principe et calcule des produits au lieu des sommes.

EXERCICE 9 La série $\sum_{k \geq 2} \frac{(-1)^k}{k^2}$ converge absolument, donc converge. AD

1. Créer un vecteur ligne u , tel que pour $1 \leq i \leq 50$, $u(i)$ soit égal à $\frac{(-1)^i}{i^2}$. On pourra utiliser pour cela une boucle `for`.
2. Créer un vecteur v tel que pour $1 \leq i \leq 50$, $v(i)$ soit égal à $\sum_{k=1}^i \frac{(-1)^k}{k^2}$.
3. Illustrer la convergence de la série en représentant graphiquement la suite de ses sommes partielles et donner une valeur approchée à 10^{-1} près de $\sum_{k=1}^{+\infty} \frac{(-1)^k}{k^2}$.

EXERCICE 10 Il est évident que $\sum_{k=1}^n \sqrt{k} \xrightarrow{n \rightarrow +\infty} +\infty$.

En utilisant une boucle `while`, déterminer le plus petit entier n tel que $\sum_{k=1}^n \sqrt{k} > 10000$.

Même question, mais avec une seule ligne de commande, en remarquant que cet entier n est nécessairement plus petit que 10000.

OPÉRATIONS POINT PAR POINT

Il existe dans `Scilab` des opérations sur les matrices différentes de celles utilisées en maths.

Celles-ci utilisent les symboles `.*`, `./` et `.^`

Lorsque A et B sont deux matrices de **même taille** $n \times p$, alors $A.*B$ renvoie la matrice de taille $n \times p$ dont le coefficient (i, j) est $a_{i,j}b_{i,j}$.

De même, $A./B$ renvoie la matrice $n \times p$ dont le coefficient (i, j) est $\frac{a_{i,j}}{b_{i,j}}$. Notons que cette opération n'est licite que si tous les coefficients de B sont non nuls.

Enfin, $A.^B$ renvoie la matrice dont les coefficients sont les $a_{i,j}^{b_{i,j}}$.

EXERCICE 11 Observez le résultat des commandes suivantes : `a = 1 : 10`, `b = -1*ones(1,10)`, `b.^a` puis `a = 1 : 10 ; a.^2`. Comprenez-vous le fonctionnement de la commande `.^` ?

EXERCICE 12 Créer, sans utiliser de boucle, un vecteur u de taille 1000 dont le n -ième élément soit égal à $\frac{(-1)^n}{n^2}$.

EXERCICE 13 Observez, pour $n = 2$, $n = 3$ et $n = 4$ le résultat de la commande

```
cumprod((n :-1 :1) ./ (1 :n))
```

Que reconnaissez-vous ? Savez-vous justifier ceci mathématiquement ? Exploiter ce résultat pour créer une fonction `binomial(n,k)` qui prend en entrée deux entiers n et k et retourne $\binom{n}{k}$.

CORRECTION DU TP 1

SOLUTION DE L'EXERCICE 1

Puisque $n!$ est le produit des entiers de 1 à n , on peut créer un vecteur contenant tous ces nombres et calculer le produit de ses coefficients grâce à la commande `prod`.

```
1 n = input('Entrez un entier :')
2 prod(1:n)
```

Subtil...

Vous pouvez constater que ceci fonctionne aussi pour $n = 0$. Savez-vous expliquer pourquoi ?

□

SOLUTION DE L'EXERCICE 2

C'est une simple question de mathématiques : on a

$$\begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} \begin{pmatrix} 1 & 1 & \dots & 1 \end{pmatrix} = \begin{pmatrix} a_1 & \dots & a_n \\ \vdots & & \vdots \\ a_1 & \dots & a_n \end{pmatrix}.$$

En particulier, on a

$$\begin{pmatrix} 1 & 1 & \dots & 1 \\ 2 & 2 & \dots & 2 \\ \vdots & \vdots & & \vdots \\ 10 & 10 & \dots & 10 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \\ \vdots \\ 10 \end{pmatrix} \begin{pmatrix} 1 & 1 & \dots & 1 \end{pmatrix}.$$

Donc une solution rapide est

```
1 (1 :10)' * ones(1,10)
```

Précision

`1 :10` crée un vecteur ligne, et nous voulons un vecteur colonne. On peut passer de l'un à l'autre via la transposition.

□

SOLUTION DE L'EXERCICE 3

Notons que la matrice recherchée s'écrit

$$\begin{pmatrix} 0 & 1 & \dots & 1 \\ 1 & 0 & & \vdots \\ \vdots & & \ddots & \vdots \\ 1 & \dots & 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 1 & \dots & 1 \\ \vdots & \vdots & & \vdots \\ \vdots & & & \vdots \\ 1 & 1 & \dots & 1 \end{pmatrix} - \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & & \vdots \\ \vdots & & \ddots & \vdots \\ 0 & \dots & 0 & 1 \end{pmatrix}.$$

Mais ces deux matrices sont faciles à créer avec les commandes `ones` et `eye`.

```
1 ones(10,10) - eye(10,10)
```

□

SOLUTION DE L'EXERCICE 4

Rappelons rapidement la syntaxe de fonction : lorsqu'on écrit fonction $y = \text{min_ligne}(A, i)$, cela signifie qu'on crée une fonction nommée `min_ligne`, qui va prendre comme paramètres A et i et va retourner la valeur de la variable y ¹.

```
1 function y = min_ligne(A,i)
2     y = min(A(i, :))
3 endfunction
```

¹ Et donc ici il s'agit de modifier la valeur de la variable y afin qu'elle prenne la valeur du plus petit coefficient de la i -ème ligne de A .

□

SOLUTION DE L'EXERCICE 5

1. Il s'agit d'appliquer les fonctions $t \mapsto \sqrt{t}$ et $t \mapsto \lfloor \ln t \rfloor$ au vecteur `1 :100`.

```
1 a = sqrt(1 :100);
2 b = floor(log(1 :100));
```

Danger !

Sous Sci Lab, la fonction `ln` est en fait appelée `log`.

2. Une fonction n'est pas obligée de retourner un nombre, elle peut retourner un couple de nombres, une matrice, etc. Ici, on lui dit de retourner une matrice (a,b) à une ligne et deux colonnes. Il nous faut donc définir les valeurs de a et b à retourner.

```
1 function [a,b] = sommes(n)
2   a = sum(sqrt(1 :n));
3   b = sum(floor(log(1 :n)));
4   endfunction
```

□

SOLUTION DE L'EXERCICE 6

Le 8-ème élément de la matrice A est le huitième de la première colonne (donc vaut 8). Le 37-ème est le 7-ème de la quatrième colonne (donc 7) et le 99-ème élément de la matrice est le 9-ème de la dernière colonne, donc vaut 9.

□

SOLUTION DE L'EXERCICE 7

Il s'agit juste d'utiliser ce qui a été dit précédemment : `find(A == max(A))` permet de chercher les éléments de A égaux à m. Pour compter ces éléments, on ajoute la fonction `length` :

```
1 length(find(A == max(A)))
```

□

SOLUTION DE L'EXERCICE 8

1. Il existe une commande pour calculer la moyenne, mais nous ne l'avons pas encore rencontrée. Peu importe : il suffit de faire la somme des coefficients de A et de la diviser par 1000

```
1 sum(A)/1000
```

En faisant plusieurs essais, on constate que cette moyenne est proche de 5, ce qui est logique car l'espérance de X vaut 5

2. Nous n'aurons qu'une approximation de cette probabilité, mais on peut compter le nombre de jours où 8 voitures sont passées et diviser ce nombre par 1000

```
1 length(find(A>=8))/1000
```

3. Sur le même principe que dans l'exercice 7 :

```
1 length(find(A)==min(A))
```

□

SOLUTION DE L'EXERCICE 9

1. On utilise ici une boucle, même s'il existe une solution plus rapide que nous verrons plus loin.

```
1 u = zeros(1,50);
2 for i=1 :50 do
3   u(i) = (-1)^i/i^2;
4 end
```

2. Remarquons que `cumsum` fait précisément ce que l'on souhaite²

```
1 v = cumsum(u);
```

3. On peut utiliser pour cela la fonction `plot`

```
1 plot2d(1 :50,v);
```

Concours

Cette subtilité s'est déjà vue aux concours (ECRICOME 2015).

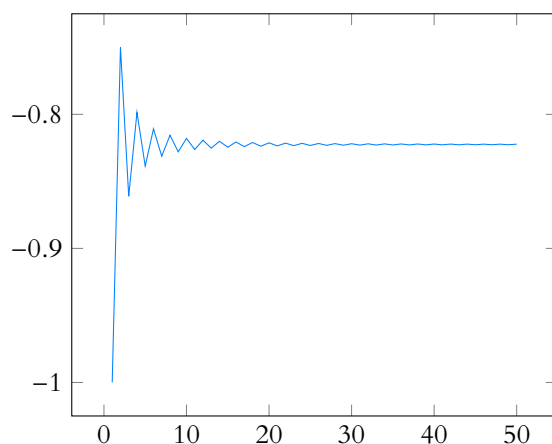
Rappel

Si $X \leftrightarrow \mathcal{P}(\lambda)$, alors $E(X) = \lambda$.

² Et c'est là le principal usage que l'on fera de `cumsum` : calculer les sommes partielles d'une série.

Rappel

`plot2d` prend deux arguments, qui sont deux vecteurs de même taille, qui contiennent respectivement les abscisses et les ordonnées des points à placer. Elle place ces points, puis les relie. Éventuellement, en ajoutant `style=-1` comme troisième paramètre, on empêche la fonction de relier les points, ce qui peut être pratique pour représenter une suite.



Le graphique semble indiquer que $\sum_{k=1}^{+\infty} \frac{(-1)^k}{k^2} \approx -0.8$. \square

SOLUTION DE L'EXERCICE 10

Nous allons calculer toutes les sommes partielles de la série de terme général \sqrt{k} , jusqu'à ce que l'une d'entre elles devienne supérieure à 10000.

```
1 S = sqrt(1);
2 n = 1;
3 while S<=10000
4     n = n+1;
5     S = S + sqrt(n);
6 end
7 disp(n)
```

On obtient alors $n = 608$. Une solution plus astucieuse (qui retourne elle aussi 608 !) en une seule ligne est la suivante :

```
1 min(find(cumsum(sqrt(1 :10000))>10000))
```

Explication : `sqrt(1 :10000)` contient les racines des entiers de 1 à 10000. Si on applique `cumsum` à ce vecteur, on a donc les sommes partielles de la série $\sum \sqrt{k}$.

`find` permet ensuite de trouver lesquelles parmi ces sommes partielles sont plus grandes que 10000 et `min` permet de trouver la plus petite valeur pour laquelle c'est le cas. \square

SOLUTION DE L'EXERCICE 12

Pour obtenir les $(-1)^n$, on peut créer un vecteur rempli de -1 , et l'«élever à la puissance» $1 : n$: le premier coefficient contiendra $(-1)^1$, le second $(-1)^2$, etc.

```
1 (-1*ones(1,1000)).^(1 :1000)./((1 :1000).^2);
```

\square

SOLUTION DE L'EXERCICE 13

Pour $n = 2$, on obtient $[2, 1]$, pour $n = 3$, on obtient $[3, 3, 1]$ et pour $n = 4$, on a $[4, 6, 4, 1]$.

On reconnaît là les lignes du triangle de Pascal, auxquelles on a retiré le premier coefficient (qui est toujours 1).

Justifions ceci : $[n \ -1 \ :1]$ est le vecteur $[n, n-1, \dots, 2, 1]$.

Et donc $[n \ -1 \ :1] ./ (1 : n)$ est le vecteur $[\frac{n}{1}, \frac{n-1}{2}, \dots, \frac{2}{n-1}, \frac{1}{n}]$.

Par conséquent, `cumprod((n : -1 : 1) ./ (1 : n))` est le vecteur dont le k -ème coefficient est

$$\frac{n}{1} \frac{n-1}{2} \times \dots \times \frac{n-k+1}{k} = \frac{n(n-1) \cdots (n-k+1)}{1 \times 2 \times \cdots \times k} = \frac{n(n-1) \cdots (n-k+1)}{k!} = \binom{n}{k}.$$

On peut en déduire une fonction binomial de la manière suivante

Boucles while

Si dans une boucle `while` la condition est toujours vérifiée, le programme ne s'arrête jamais. Ici, il est donc impératif que la valeur de `S` change au cours du programme. De plus, on n'oubliera pas d'augmenter la valeur de `n` au fur et à mesure, ce n'est pas automatique, contrairement aux boucles `for`.

```
1 function y = binomial(n,k)
2   if k == 0 then
3     y = 1 ;
4   else
5     A = cumprod((n :-1 :1)./(1 :n)) ;
6     y = A(k)
7   end
8 endfunction
```

Notons que si l'on veut éviter le recours à une disjonction de cas, il existe une solution encore plus astucieuse :

```
1 function y = binomial(n,k)
2   A = [1,cumprod((n :-1 :1)./(1 :n))];
3   y = A(k+1) ;
4 endfunction
```

On crée ainsi un vecteur A dont la première coordonnées vaut 1 et les suivantes sont les coefficients binomiaux que nous avons vu précédemment. \square

TP 2 : SIMULATION DE LOIS DISCRÈTES

On rappelle que la fonction `rand()` permet de simuler une variable aléatoire suivant une loi uniforme (à densité) sur $[0, 1]$. Le but de ce TP est d'apprendre à simuler des lois discrètes uniquement avec `rand`, et donc sans utiliser la commande `grand`, qui est capable de simuler la plupart des lois usuelles.

SIMULATION D'UNE LOI GÉOMÉTRIQUE

EXERCICE 1 On rappelle qu'une loi géométrique $\mathcal{G}(p)$ correspond au nombre d'essais nécessaires avant l'obtention d'un succès lors de la répétition d'épreuves de Bernoulli indépendantes de paramètre p .

F

1. Compléter la fonction suivante afin qu'elle simule une loi de Bernoulli de paramètre p :

```
1 function y = bernoulli(p)
2     u = rand();
3     y = 0;
4     if u ... then
5         y = ....
6     end
7 endfunction
```

2. En utilisant la fonction précédente, compléter la fonction suivante afin qu'elle simule une loi géométrique de paramètre p :

```
1 function y = geom(p)
2     y = ...
3     while bernoulli(p)....
4         ....
5     end
6 endfunction
```

3. En appelant 10000 fois (par exemple à l'aide d'une boucle `for`) la fonction `geom`, avec la valeur de p de votre choix, vérifier que la valeur moyenne retournée par `geom` est cohérente avec ce qu'on attend.

SIMULATION DU MOBILE D'EDHEC 2018

Nous reprenons dans cette partie l'énoncé de l'exercice de probabilités du DM1, extrait du sujet 2018 de l'EDHEC. Il s'agissait d'un mobile se déplaçant sur un axe de la manière suivante : lors du $n^{\text{ème}}$ déplacement, le mobile se rend de manière équiprobable sur l'un des points d'abscisse $0, 1, \dots, n$.

On note alors X le numéro du déplacement où, pour la première fois, le mobile revient à l'origine.

EXERCICE 2 Simulation du mobile, espérance de X .

AD

1. Compléter la fonction suivante pour qu'elle simule X .

```
1 function n = X()
2     n = 1
3     while grand(...) ....
4         n = n+1
5     end
6 endfunction
```

2. Créer un vecteur ligne A à 10 000 colonnes contenant 10 000 simulations de X .

3. Il a été prouvé que $P(X = n) = \frac{1}{n(n+1)}$.

Nous nous proposons ici de vérifier graphiquement cette formule à l'aide de la commande `bar`, qui permet de tracer des diagrammes «en bâtons». À cet effet, vous pourrez utiliser le code suivant

```
1 u = zeros(1,10) ;
2 for i=1 :10000
3     if A(i)<=10 then
4         u(A(i)) = u(A(i))+1 ;
5         end
6     end
7     u=u/10000
8     v = [1 :10] ;
```



```

9  for i=1 :10
10     v(i) = 1/(i*(i+1));
11  end
12  bar([1 :10],[u',v'])

```

La commande bar crée ainsi, pour chaque entier i entre 1 et 10, une barre (bleue) de hauteur $u(i)$ et une barre (verte) de hauteur $v(i)$.

- 3.a. Expliquer l'effet des lignes 2 à 6, puis celui de la ligne 7.
- 3.b. Quel effet ont les lignes 9 à 11 ?
- 3.c. Interpréter le graphique obtenu.
4. Il a été prouvé que X n'admet pas d'espérance. Pourtant, si on simule un grand nombre de fois X , ce que nous avons fait à la question 2, il est possible de calculer la moyenne de ces simulations. Calculer cette moyenne, puis exécuter de nouveau le code de la question 2 et calculer de nouveau la moyenne. Répéter l'opération plusieurs fois, et commenter les résultats obtenus.

EXERCICE 3 Simulation d'une loi $\mathcal{U}([0, n])$ à l'aide de rand()

Dans l'exercice précédent, nous avons utilisé grand pour simuler une loi uniforme à valeurs entières.

Imaginer un moyen de se passer de grand et de simuler une loi uniforme sur $[0, n]$ en n'utilisant que la fonction rand() et la fonction floor (qui calcule la partie entière).

Faire de même pour une uniforme sur $[1, n]$, puis pour une uniforme sur $[a, b]$, a et b étant deux entiers avec $a < b$.

EXERCICE 4 Modifier la fonction X de l'exercice 2 pour qu'il simule la variable X_k , qui correspond au nombre de déplacements du mobile avant le $k^{\text{ème}}$ retour à l'origine. On pourra commencer par fonction $n = X(k)$.

Peut-on le faire sans utiliser grand() ?

AD

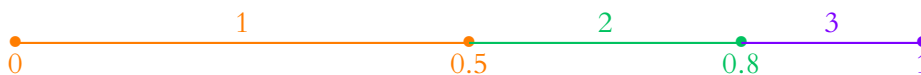
PD

LA MÉTHODE D'INVERSION

Nous expliquons dans ce paragraphe une méthode efficace, appelée méthode d'inversion, permettant de simuler n'importe quelle loi à valeurs dans \mathbf{N} .

Commençons par un exemple simple : pour simuler une variable prenant la valeur 1 avec probabilité 0.5, la valeur 2 avec probabilité 0.3 et la valeur 3 avec probabilité 0.2, on peut tirer un nombre au hasard dans $[0, 1]$ avec rand(), et retourner 1 si ce nombre est plus petit que 0.5 (ce qui se produit avec probabilité 0.5), retourner 2 s'il est dans $]0.5, 0.8]$ (ce qui se produit avec probabilité $0.8 - 0.5 = 0.3$) et retourner 3 s'il est plus grand que 0.8 (ce qui se produit avec probabilité $1 - 0.8 = 0.2$).

Autrement dit, on a coupé en trois le segment $[0, 1]$, et on renvoie 1, 2 ou 3 suivant dans quel petit intervalle se trouve le nombre produit par rand().



Plus rigoureusement : si X est une variable aléatoire discrète qui prend les valeurs $x_1 < x_2 < \dots < x_n < \dots$, alors sa fonction de répartition est donnée par :

$$F(x) = \begin{cases} 0 & \text{si } x < x_1 \\ \sum_{i=1}^n P(X = x_i) & \text{si } x_n \leq x < x_{n+1} \end{cases}$$

En particulier, si $U \hookrightarrow \mathcal{U}([0, 1])$, alors $P(U \leq F(x_1)) = F(x_1) = P(X = x_1)$ et

$$\forall k \geq 2, P(F(x_{k-1}) < U \leq F(x_k)) = F(x_k) - F(x_{k-1}) = P(X \leq x_k) - P(X \leq x_{k-1}) = P(X = x_k).$$

EXERCICE 5 Simulation de la loi binomiale par la méthode d'inversion

La méthode vue précédemment s'adapte facilement pour simuler toute variable aléatoire X ne prenant qu'un nombre fini de valeurs : il suffit de «couper» le segment $[0, 1]$ en autant de petits segments que le cardinal de $X(\Omega)$, les longueurs de ces segments étant les probabilités $P(X = x)$.

Nous allons essayer ici de simuler une loi binomiale $\mathcal{B}(10, 0.5)$.

Pour cela nous allons utiliser la commande binomial(0.5, 10) (qui ne figure pas au programme) et qui renvoie un vecteur de taille 11, contenant $P(X = 0), P(X = 1), \dots, P(X = 10)$.

Nous nous donnons à cet effet le code suivant :

```

1  function n = simul_binomiale()
2     u = binomial(0.5,10);
3     v = cumsum(u);
4     t = rand();
5     n = 0;
6     while t > v(n+1)

```

AD

```

7         n = n+1 ;
8     end
9 endfunction

```

1. Simuler 10 000 fois X à l'aide de la fonction `simule_binomiale()`, et calculer la moyenne de ces simulations. Le résultat était-il prévisible ?
2. Dans le code donné ci-dessus, que représente le vecteur v ?
3. Expliquer le fonctionnement de ce programme, et en particulier des lignes 6 et 7.

EXERCICE 6 Telle que nous l'avons utilisée précédemment, la méthode d'inversion ne permet de simuler que des lois à support fini. En effet, pour une variable discrète à support infini, il n'est pas possible de calculer dès le début toutes les $P(X = k)$, puisqu'il y en a une infinité.

Nous proposons ici de l'adapter pour simuler une loi de Poisson (dont le support est, rappelons-le, \mathbf{N}).

```

1 function k=poisson(lambda)
2     t = rand()
3     k = 0 ;
4     u = exp(-lambda) ;
5     s=u ;
6     while s<t
7         k = k+1 ;
8         u = u*lambda/k ;
9         s = s+u ;
10    end
11 endfunction
12
13 A = zeros(1,30) ;
14 for i =1 :100000
15     x = poisson(8)+1 ;
16     A(x) = A(x)+1 ;
17 end
18 A = A/100000 ;
19 u = zeros(1,30)
20 for i=1 :30
21     u(i) = exp(-8)*8^(i-1)/prod(1 :i-1) ;
22 end
23 bar(0 :29,[A',u']) ;

```

1. Exécuter ce code et interpréter le graphique ainsi obtenu. En particulier, à quoi correspond le vecteur u ?
2. Expliquer le fonctionnement de la fonction `poisson`.

EXERCICE 7 Deuxième simulation de la loi géométrique

1. Sur le même principe que la fonction `poisson`, écrire une fonction `geom2` qui prend en paramètre un nombre $p \in]0, 1[$, puis à l'aide de la méthode d'inversion, simule une loi géométrique de paramètre p .

2. *Question facultative* : montrer que si $X \hookrightarrow \mathcal{G}(p)$, alors $F_X(t) = \begin{cases} 0 & \text{si } t < 1 \\ 1 - (1 - p)^n & \text{si } t \in [n, n + 1[, n \in \mathbf{N}^* \end{cases}$

Essayer de simplifier la fonction `geom2` pour ne plus avoir besoin d'une boucle `while`. On pourra à cet effet utiliser la fonction `ceil` qui est la fonction partie entière supérieure, qui à x associe l'entier immédiatement supérieur à x .

3. *Facultatif* : il est possible de connaître le temps nécessaire à l'exécution d'une suite de commandes en utilisant les commandes `tic()` et `toc()`. Il faut pour cela utiliser la commande `tic()` juste avant le bloc de commandes qui nous intéresse, puis à la fin utiliser `temps = toc()`. La variable `temps` contient alors le temps (en secondes) qui s'est écoulé entre l'appel de `tic()` et celui de `toc()`.

Utiliser cette méthode pour comparer les temps d'exécution de `geom` et `geom2` lors de 100 000 simulations. On pourra également comparer avec l'utilisation de `grand`.

EXERCICE 8 Une autre simulation de la loi d'EDHEC 2018

1. En utilisant la méthode d'inversion, simuler la loi de la variable X de l'exercice 2.
2. En remarquant que $F_X(t) = \begin{cases} 0 & \text{si } t < 1 \\ 1 - \frac{1}{n+1} & \text{si } t \in [n, n + 1[\end{cases}$, simuler la loi de X à l'aide de la méthode d'inversion, mais sans utiliser de boucle.

UNE NOUVELLE LOI : LA LOI BINOMIALE NÉGATIVE.

Reprenons notre mobile, et changeons un peu les règles du jeu : on suppose à présent qu'à chaque déplacement, le mobile retourne en 0 avec probabilité p , et sinon, se rend de manière équiprobable en l'un des points d'abscisse $1, 2, \dots, n$.

Pour $i \in \mathbf{N}^*$, on note alors X_i la variable aléatoire correspondant au numéro du déplacement amenant le $i^{\text{ème}}$ retour à l'origine.

EXERCICE 9 Reconnaître la loi de X_1 .

2. En vous inspirant de ce qui a été fait dans les exercices précédents, écrire une fonction qui simule la loi de X_i .
3. Déterminer le support de X_i , et montrer que $\forall k \in X_i(\Omega), P(X_i = k) = \binom{k-1}{i-1} p^i (1-p)^{k-i}$.

4. Écrire un autre programme simulant la loi de X_i , mais en utilisant cette fois la méthode d'inversion. On pourra commencer par trouver une relation liant $\binom{k-1}{i-1}$ et $\binom{k}{i-1}$.

D'AUTRES SIMULATIONS

EXERCICE 10 Simuler une loi uniforme $\mathcal{U}(\llbracket 1, n \rrbracket)$ par la méthode d'inversion, sans utiliser de boucle.

EXERCICE 11 Un problème de supermarché : la loi de Borel

On considère une caisse de supermarché fonctionnant sur le principe suivant : chaque client passe une minute à la caisse, et pendant qu'un client est servi, le nombre de clients arrivant à la caisse est une variable aléatoire suivant une loi de Poisson $\mathcal{P}(\mu)$.

- Écrire une fonction caisse de paramètre μ estimant le nombre de minutes nécessaires après l'arrivée du premier client avant que la caisse soit vide. Essayer cette fonction à l'aide de plusieurs valeurs de μ .
À tout hasard, on rappelle que si un programme semble ne pas s'arrêter, on peut utiliser la combinaison de touches **Ctrl** + **C** suivie de abort.
- Constaté que pour $\mu < 1$, la moyenne est de l'ordre de $\frac{1}{1-\mu}$. Essayer d'expliquer ce résultat «avec les mains» en considérant le nombre moyen de clients se présentant à la caisse pendant que le premier est servi, puis le nombre moyen de clients arrivant à la caisse pendant que ces «clients de première génération» sont servis, etc.
- La probabilité que ce nombre de clients soit égal à n est égale à $\frac{e^{-\mu n}(\mu n)^{n-1}}{n!}$.
Vérifier ceci à l'aide d'un graphique semblable à celui utilisé dans l'exercice 2.

AD

D

CORRECTION DU TP 2

SOLUTION DE L'EXERCICE 1

1. On souhaite retourner 1 avec probabilité p et 0 avec probabilité $1 - p$. Or la probabilité que u prenne une valeur inférieure ou égale à p est égale à p , donc une solution est la suivante :

```

1 function y = bernoulli(p)
2     u = rand();
3     y = 0;
4     if u < p then
5         y = 1;
6     end
7 endfunction

```

2. La boucle sera parcourue autant de fois qu'il y a d'appels à la fonction `bernoulli` qui retournent 0. Il suffit donc de compter le nombre de passages dans la boucle, et d'ajouter un (car au moment de l'obtention du succès, le programme s'arrête, et le contenu de la boucle n'est pas exécuté.)

```

1 function y = geom(p)
2     y = 1;
3     while bernoulli(p)==0
4         y = y+1;
5     end
6 endfunction

```

3. On peut appeler 10000 fois la fonction `geom`, et faire la moyenne des résultats ainsi obtenus. On s'attend alors à avoir une valeur proche de $\frac{1}{p}$, qui est l'espérance d'une loi géométrique.

```

1 S = 0;
2 for i = 1 :10000
3     S = S + geom(0.2);
4 end;
5 S/10000

```

Exécutées plusieurs fois, ces lignes de code m'ont retourné les valeurs 4.9335, 4.9568, 4.9998 et 5.054, toute proches de la valeur attendue $\frac{1}{0.2} = 5$. □

SOLUTION DE L'EXERCICE 2

1. À chaque étape, il s'agit de simuler le déplacement du mobile, qui suit une loi uniforme sur $[[0, n]]$.

Si ce déplacement nous mène en 0, alors le programme s'arrête.

```

1 function n = X()
2     n = 1
3     while grand(1,1,'uin',0,n) <> 0
4         n = n+1
5     end
6 endfunction

```

2. L'idée est de commencer par créer un vecteur de la bonne taille ne contenant que des 0, puis de remplir ses coefficients un par un à l'aide d'une boucle, chacun des coefficients étant simulé à l'aide de `X`.

```

1 A = zeros(1,10000);
2 for i=1 :10000
3     A(i) = X();
4 end

```

Loi uniforme

Rappelons que la fonction de répartition d'une variable aléatoire suivant la loi uniforme $\mathcal{U}([0, 1])$ est

$$x \mapsto \begin{cases} 0 & \text{si } x \leq 0 \\ x & \text{si } 0 \leq x \leq 1 \\ 1 & \text{sinon} \end{cases}$$

Boucles while

Il est facile de se tromper avec une boucle `while`, et de retourner un de moins ou un de plus que ce qu'on avait prévu. Il est conseillé de toujours essayer «à la main», avec un petit nombre de passages que notre programme retourne bien le nombre attendu.

Espérance

Si l'intuition nous dit qu'on obtiendra bien ainsi l'espérance, le résultat mathématique le justifiant, appelé loi faible des grands nombres sera vu plus tard dans l'année, et se formule à l'aide d'une convergence en probabilité. Notons que si vraiment on n'a pas eu de chance sur nos 10000 tirages, rien n'exclut que le résultat obtenu soit très loin de $\frac{1}{p}$.

- 3.a. La boucle for sert à parcourir le vecteur A.
Le if de la ligne 3 nous indique qu'il ne se passe rien si on trouve des coefficients supérieurs à 10.
Dans le cas contraire, si A(i) vaut k, alors on ajoute 1 au k^{ème} coefficient de u.
Autrement dit, à l'issue de la boucle, la valeur de u(k) est égale au nombre de coefficients de A qui étaient égaux à k : on a compté le nombre de simulations de X qui étaient égales à 1, à 2, etc.
La ligne 7 a alors pour effet de

□

SOLUTION DE L'EXERCICE 3

Commençons par un exemple avant de passer au cas général : rand() retourne un nombre suivant une loi uniforme¹ sur [0, 1].

Et donc, par transformation affine de loi uniforme, rand()*10 retourne un nombre suivant une loi uniforme sur [0, 10].

Par conséquent, floor(10*rand()) est un entier entre 0 et 9. Ce nombre pourrait a priori prendre la valeur 10, mais pour que cela arrive, il faudrait que 10*rand() soit exactement égal à 10, et la probabilité que ceci arrive est nulle.

Et on se convainc facilement que les valeurs 0, 1, ..., 9 sont prises de manière équiprobable.

Donc floor(10*rand()) simule une loi $\mathcal{U}(\llbracket 0, 9 \rrbracket)$.

Plus généralement, floor((n+1)*rand()) simule une loi uniforme sur $\llbracket 0, n \rrbracket$.

Enfin, pour le cas d'une uniforme sur $\llbracket a, b \rrbracket$, notons que

$$X \leftrightarrow \mathcal{U}(\llbracket a, b \rrbracket) \Leftrightarrow X - a \leftrightarrow \mathcal{U}(\llbracket 0, b - a \rrbracket).$$

Et donc pour simuler une loi $\mathcal{U}(\llbracket a, b \rrbracket)$, nous pouvons utiliser a + floor((b-a+1)*rand()).

□

SOLUTION DE L'EXERCICE 5

- Lors de nos 10 000 simulations, on obtient une moyenne proche de 5, ce qui correspond bien au fait l'espérance d'une binomiale de paramètres 10 et 0.5 est égale à 5.
- Le vecteur v contient

$$\begin{aligned} & [P(X = 0), P(X = 0)+P(X = 1), P(X = 0)+P(X = 1)+P(X = 2), \dots, P(X = 0)+\dots+P(X = 10)] \\ & = [P(X \leq 0), P(X \leq 1), \dots, \underbrace{P(X \leq 10)}_{=1}]. \end{aligned}$$

Il correspond donc en quelque sorte à la fonction de répartition de X : celle-ci est en effet une fonction en escaliers, et les différents «paliers» sont de hauteur $P(X = 0), P(X = 0) + P(X = 1), \dots, 1$.

- La ligne 4 a pour effet de tirer un nombre au hasard t dans le segment [0, 1].
Puis, on parcourt le vecteur v à l'aide de la boucle while jusqu'à trouver un coefficient de v supérieur à t. Si ce coefficient vaut n+1, alors le programme retourne n.
Autrement dit, si $t \in]P(X \leq n - 1), P(X \leq n)]$, ce qui se produit avec probabilité $P(X \leq n) - P(X \leq n - 1) = P(X = n)$, alors le programme retourne n : il simule bien une loi $\mathcal{B}(10, 0.5)$.

□

SOLUTION DE L'EXERCICE 6

- u est un vecteur qui contient ce que l'on reconnaît être la loi de probabilité d'une variable aléatoire suivant la loi de Poisson.

Dans le même temps A est un vecteur qui compte le nombre d'occurrences de chaque nombre compris entre 0 et 29 pour 10000 appels à la fonction poisson.

Le graphique affiché par la seconde partie du programme est le suivant :

¹ À densité.

Détail

C'est en fait un exercice de probabilités assez classique, et pas inintéressant : si $X \leftrightarrow \mathcal{U}(\llbracket 0, n \rrbracket)$, alors

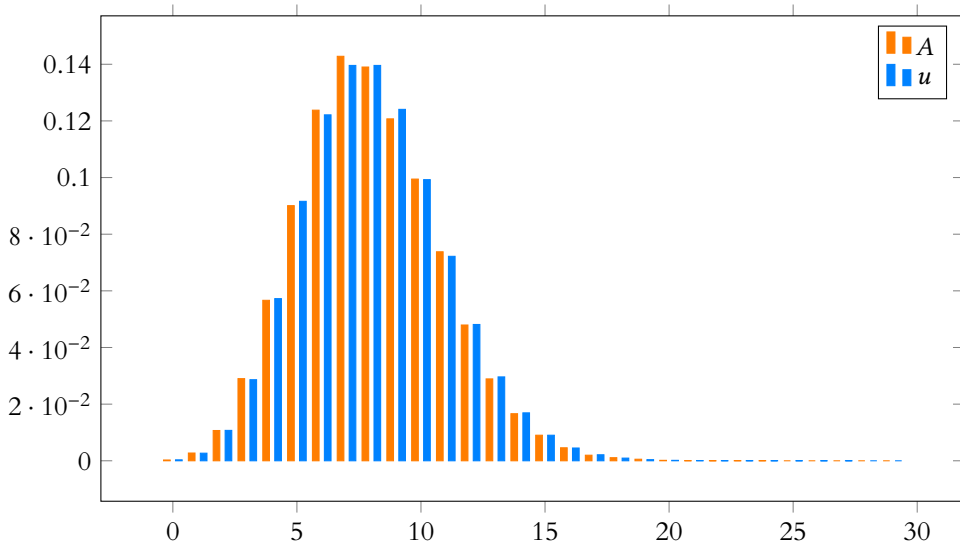
$$\lfloor U \rfloor \leftrightarrow \mathcal{U}(\llbracket 0, n - 1 \rrbracket).$$

⚠ Attention !

Le fait qu'on s'intéresse à t(n+1) et non à t(n) tient au fait que la loi binomiale prend ses valeurs dans $\llbracket 0, 10 \rrbracket$, alors que les éléments du vecteur t sont numérotés de 1 à 11.

Factorielle

La commande prod([1 : i-1]) sert à calculer (i-1)! puisqu'il s'agit du produit des nombres de 1 à i-1.



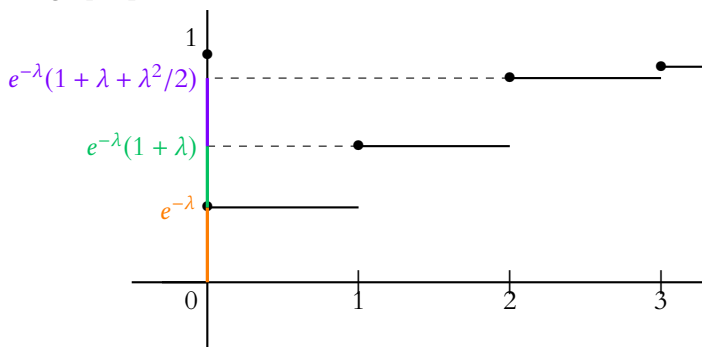
Les deux graphiques étant très proches, on peut donc penser que la fonction poisson simule une loi de Poisson de paramètre 8.

Probabilités
 Bien entendu, on n'est sûr de rien, et si les graphiques se ressemblent, peut-être est-ce un coup de chance ... Pourtant, si l'on répète plusieurs fois l'opération, les graphiques semblent toujours très proches (bien que celui de A puisse varier légèrement).

- L'idée de la méthode d'inversion est de découper le segment $[0, 1]$ en utilisant la fonction de répartition de la loi de Poisson. Rappelons que la fonction de répartition de la loi de Poisson est

$$F : x \mapsto P(X \leq x) = \begin{cases} 0 & \text{si } x < 0 \\ \sum_{k=0}^{\lfloor x \rfloor} P(X = k) & \text{sinon} \end{cases} = \begin{cases} 0 & \text{si } x < 0 \\ \sum_{k=0}^{\lfloor x \rfloor} e^{-\lambda} \frac{\lambda^k}{k!} & \text{sinon} \end{cases}$$

Son graphique est alors



En regardant le segment $[0, 1]$ sur l'axe des ordonnées, on obtient alors le découpage suivant :



La méthode d'inversion consiste alors à tirer un nombre au hasard² entre 0 et 1. Si ce nombre est dans le premier segment, alors on retourne 0, s'il est dans le second, on retourne 1, etc. Ainsi, la probabilité qu'on retourne 0 est la longueur du premier segment : c'est $e^{-\lambda}$. De même, la probabilité qu'on retourne 1 est la longueur du second segment : c'est $\lambda e^{-\lambda}$. Etc... Et la probabilité qu'on retourne k est $e^{-\lambda} \frac{\lambda^k}{k!}$.

² Par là, on entend un nombre suivant la loi uniforme sur $[0, 1]$. C'est exactement ce que fait la commande rand de Sci Lab .

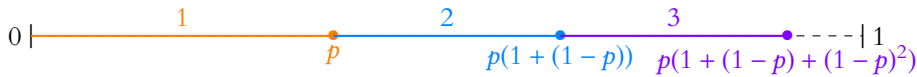
C'est exactement ce que fait la fonction poisson : elle commence par choisir un nombre t au hasard entre 0 et 1. Puis elle calcule successivement les $\sum_{i=0}^k e^{-\lambda} \frac{\lambda^i}{i!}$, jusqu'à ce qu'un de ces nombres soit strictement supérieur à t , et retourne alors k .

Détails
 Ce programme est en fait assez peu lisible car il utilise une astuce pour accélérer le calcul : plutôt que des calculer à chaque fois $\frac{\lambda^k}{k!}$ (ce qui demande $2k$ produits), il utilise le fait que

$$\frac{\lambda^{k+1}}{(k+1)!} = \frac{\lambda^k}{k!} \frac{\lambda}{k+1}.$$

SOLUTION DE L'EXERCICE 7

1. Sur le même principe, il s'agit de diviser le segment $[0, 1]$ en utilisant la loi de X . On obtient alors :



Il s'agit donc, comme pour la fonction poisson de choisir un nombre t au hasard, puis de calculer les sommes $\sum_{i=1}^k p(1-p)^{i-1}$ jusqu'à en obtenir une supérieure³ à t .

```

1  function k = geom2(p)
2      t = rand();
3      s = p;
4      k=1;
5      while s < t
6          s = s+p*(1-p)^k;
7          k=k+1;
8      end
9  endfunction

```

2. Soit X une variable aléatoire suivant la loi géométrique $\mathcal{G}(p)$. Alors pour $t \in [n, n+1[$, avec n entier non nul, on a

$$F_X(t) = P(X \leq t) = P(X \leq n) = \sum_{k=1}^n P(X = k) = \sum_{k=1}^n p(1-p)^{k-1} = p \sum_{i=1}^{n-1} (1-p)^i.$$

Mais on reconnaît alors les sommes partielles d'une série géométrique de raison $1-p$, de sorte que

$$F_X(t) = p \frac{(1-p)^n - 1}{(1-p) - 1} = 1 - (1-p)^n.$$

Ainsi, pour $u \in]0, 1[$, on peut se dispenser d'utiliser une boucle pour trouver le plus petit entier k tel que $F_X(k) \geq t$. En effet, on a

$$F_X(k) \geq t \Leftrightarrow 1 - (1-p)^k \geq t \Leftrightarrow (1-p)^k \leq 1-t \Leftrightarrow k \geq \frac{\ln(1-t)}{\ln(1-p)}.$$

On cherche donc le plus petit entier supérieur ou égal à $\frac{\ln(1-t)}{\ln(1-p)}$.

```

1  function k = geom2(p)
2      t = rand();
3      k = ceil(log(1-t)/log(1-p));
4  endfunction

```

3. Pour comparer évaluer le temps d'exécution de 100 000 appels à la fonction geom, on peut utiliser le code suivant :

```

1  tic();
2  for i=1 :100000
3      geom2(0.01);
4  end
5  temps=toc();
6  disp(temps)

```

Notons que geom2 s'exécute en des temps bien inférieurs à geom, ce qui est d'autant plus vrai que p est petit.

Notons également que grand(1, 1, 'geom', p) est légèrement plus rapide que geom2. En revanche, si l'on remplace la boucle par grand(1000, 100, 'geom', p) alors on obtient un temps bien plus court que toutes les autres méthodes⁴.

Précision

Comme une loi géométrique prend comme valeurs 1,2, etc, ici le premier segment correspond à 1 et non à 0 comme pour une loi de Poisson.

³ Supérieure strictement ou supérieure ou égale, peu importe : t est tiré suivant une loi à densité, et donc la probabilité qu'il prenne exactement n'importe quelle valeur est nulle.

Intuition

Si on pense à une loi géométrique comme un temps d'attente, alors $P(X > n)$ est la probabilité que les n premiers essais aient été des échecs.

Elle vaut donc $(1-p)^n$.
Et donc

$$P(X \leq n) = 1 - (1-p)^n.$$

Signes...

On a changé le sens de l'inégalité car $\ln(1-p) < 0$

Logique !

Plus p est petit, plus geom(p) fait de calculs car elle doit alors répéter un grand nombre d'épreuves de Bernoulli avant d'obtenir un succès. En revanche, le temps d'exécution de geom2 ne varie pas : il y a toujours le même nombre de calculs à faire.

⁴ SciLab possède ses propres méthodes pour générer encore plus rapidement des nombres aléatoires.

□

SOLUTION DE L'EXERCICE 9

1. La loi de MF_1 est une loi géométrique $\mathcal{G}\left(\frac{1}{576}\right)$.
2. L'événement $[MF_i = k]$ est réalisé si et seulement si la k -ième carte achetée est une carte Fourcade, et que sur les $k - 1$ premières cartes, $i - 1$ exactement soient des cartes Fourcade. Une fois que l'on a choisi les positions de ces cartes, la probabilité de tirer ainsi les cartes est $\left(\frac{1}{576}\right)^i \left(1 - \frac{1}{576}\right)^{k-i}$ (on a tiré i fois la carte Fourcade et $k - i$ fois une autre carte).
Mais il y a $\binom{k-1}{i-1}$ manières de choisir la position de $i - 1$ premières cartes Fourcade parmi les $k - 1$ premières cartes, et les événements ainsi obtenus sont deux à deux disjoints. Donc

$$\forall k \in \mathbf{N}, P(MF_i = k) = \binom{k-1}{i-1} \left(\frac{1}{576}\right)^i \left(1 - \frac{1}{576}\right)^{k-i}.$$

3. Commençons par utiliser une boucle `while` : on va répéter des épreuves de Bernoulli jusqu'à obtention du i -ème succès.

```

1  fonction y = MF(i,p)
2  y = 0 ;
3  s = 0 ;
4  while s < i
5      u=bernoulli(p) ;
6      y = y+1 ;
7      s = s + bernoulli(p) ;
8  end
9  endfunction

```

Si l'on préfère utiliser la méthode d'inversion (plus dure à mettre en œuvre, et pas véritablement plus rapide, sauf si l'on arrive à se passer de la boucle), notons que

$$\binom{k}{i-1} = \frac{k!}{(i-1)!(k-i+1)!} = \frac{k}{k-i+1} \frac{(k-1)!}{(i-1)!(k-i)!} = \frac{k}{k-i+1} \binom{k-1}{i-1}.$$

On peut alors procéder comme pour la fonction poisson

```

1  fonction k = MF(i,p)
2  t = rand() ;
3  k = i ;
4  s = p^i ;
5  u = p^i ;
6  while s < t
7      u = u*(1-p)*k/(k-i+1) ;
8      s = s+u ;
9      k=k+1 ;
10 end
11 endfunction

```

4. Le nombre de vignettes à acheter pour que les 6 petites sœurs aient chacune une vignette est MF_7 . On peut simuler ainsi un grand nombre de collections :

```

1  S = 0 ;
2  for i=1 :1000
3      S = S+MF(7,1/576) ;
4  end
5  S/1000

```

On constate alors qu'il faut environ 4000 vignettes pour obtenir 7 fois la vignette Fourcade (la moyenne théorique est $E(MF_7) = 7 \times 576 = 4032$).

□

SOLUTION DE L'EXERCICE 10

Il s'agit de remarquer que la fonction de répartition d'une loi uniforme $\mathcal{U}(\llbracket 1, n \rrbracket)$ est

$$F_X : x \mapsto \begin{cases} 0 & \text{si } x < 1 \\ \frac{k}{n} & \text{si } k \leq x < k+1, k \in \llbracket 1, n \rrbracket \\ 1 & \text{si } x \geq n \end{cases}$$

Autrement dit, pour $x \in [0, 1]$, l'unique entier k tel que $F(k) < x \leq F(k+1)$ vérifie

$$\frac{k}{n} < x \leq \frac{k+1}{n} \Leftrightarrow k < nx \leq (k+1) \Leftrightarrow k = \lceil nx \rceil$$

Donc

```
1 fonction y = uniforme(n)
2   y = ceil(n*rand());
3 endfunction
```

□

SOLUTION DE L'EXERCICE 11

- Utilisons une variable u pour simuler le nombre de clients dans la queue et une autre y pour simuler le temps.

```
1 fonction y= caisse(lambda)
2   y = 0
3   u = 1
4   while u <> 0
5       u = u+grand(1,1,'poi',lambda)-1;
6       y = y+1
7   end
8 endfunction
```

On constate que pour $\mu \geq 1$, le programme ne semble pas s'arrêter. En effet, puisque l'espérance d'une loi de Poisson $\mathcal{P}(\mu)$ est μ , pour chaque client servi, μ clients arrivent en moyenne. Si $\mu > 1$, la file a peu de chances de se vider.

- Pendant que le premier client est servi, μ clients arrivent en moyenne. Appelons ces clients des clients de «première génération». Pendant que chacun de ces clients est servi, μ clients arrivent, soit μ^2 «clients de seconde génération». Et ainsi de suite : il y a **en moyenne** μ^n clients de n -ème génération. Et donc le nombre moyen de clients est

$$1 + \mu + \mu^2 + \dots + \mu^n + \dots = \frac{1}{1 - \mu}.$$

- Testons par exemple pour $\mu = 1/2$. Commençons par créer un vecteur contenant les valeurs théoriques :

```
1 prob = zeros(1,30);
2 for i = 1 :30
3   prob(i) = exp(-i/2)*(i/2)^(i-1)/prod(1:i);
4 end
```

Puis, appelons 1000 fois la fonction `caisse` en comptant les résultats, et enfin traçons le tout. On remarque alors que les graphiques concordent...

```
1 u = zeros(1,30);
2 for i=1:1000
3   x = caisse(1/2);
4   u(x) = u(x)+1;
5 end
6 bar([1:30],[prob',u'/1000]);
```

□

Graphiquement

Si l'on préfère faire un dessin, la fonction de répartition d'une loi uniforme découpe le segment $[0, 1]$ en n morceaux de taille égale.

Notation

$\lceil x \rceil$ désigne la partie entière supérieure, c'est à dire le premier entier immédiatement supérieur à x .

Attention

On n'oublie pas de commencer avec un client dans la file ($u=1$) et d'en enlever un à chaque passage de la boucle, et pas seulement à en ajouter.

Notons qu'on a utilisé `grand`, mais on aurait aussi bien pu utiliser la fonction `poisson` des exercices précédents.

Pour la culture

Cette distribution de probabilités s'appelle loi de Borel-Tanner, et il est très délicat de prouver qu'il s'agit bien d'une loi de probabilités, c'est à dire que

$$\sum_{n=1}^{+\infty} e^{-\mu n} \frac{(\mu n)^{n-1}}{n!} = 1.$$

TP 3 : STATISTIQUES À UNE VARIABLE

Donner une définition exacte de ce qu'est la statistique n'est pas chose facile, et le statisticien britannique Maurice Kendall ironisait :

« Parmi les thèmes à propos desquels les statisticiens ne sont pas d'accord, se trouve la définition de leur science. »

Sur Wikipédia se trouve tout de même une définition qui nous conviendra bien :

La statistique est l'étude de la collecte de données, leur analyse, leur traitement, l'interprétation des résultats et leur présentation afin de rendre les données compréhensibles par tous.

Le but de ce TP est d'étudier, sur des exemples, quels indicateurs statistiques peuvent être pertinents dans l'étude d'une série de données, et comment utiliser SciLab afin de manipuler de grandes quantités de données.

VOCABULAIRE DES STATISTIQUES

Donnons ou redonnons quelques définitions de statistiques :

On étudie un ensemble $\Omega = \{\omega_1, \dots, \omega_n\}$ fini appelé **population**, et on appelle $n = \text{Card}(\Omega)$ l'**effectif** de cette population.

Un **caractère** est une application X définie sur Ω , et qui peut être à valeurs dans un ensemble quelconque.

Lorsque X est un caractère à valeurs réelles, on dit que X est un caractère **quantitatif** (par opposition aux caractères **qualitatifs**). Par exemple si la population est l'ensemble des étudiants d'ECS2, alors la taille (en cm) ou la note au dernier DS de maths sont des caractères quantitatifs, alors que la commune de naissance est un caractère qualitatif.

Nous ne traiterons dans la suite que du cas des caractères quantitatifs.

L'ensemble des valeurs prises par un caractère X est alors une **série statistique**, que l'on note généralement $X(\Omega) = \{x_1, \dots, x_n\}$, où $x_i = X(\omega_i)$.

Les valeurs effectivement prises par X sont aussi appelées **modalités** de la série statistique. Par exemple, les modalités de la série $\{5, 1, 4, 7, 7, 5, 2, 1, 1, 1\}$ sont 1, 2, 4, 5 et 7. L'**effectif d'une modalité** est le nombre de fois où cette modalité apparaît dans la série. Dans l'exemple précédent, l'effectif de 1 est 4.

Sous SciLab, nous représenterons les séries statistiques par des vecteurs (lignes ou colonnes).

La **moyenne** d'une série statistique est $\bar{x} = \frac{x_1 + \dots + x_n}{n}$. Elle est obtenue en SciLab par `mean(X)`, ce qui est équivalent à `sum(X)/length(X)`.

L'**écart-type** d'une série statistique $X = \{x_1, \dots, x_n\}$ est

$$\sigma_{n-1}(X) = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2}.$$

Il s'obtient à l'aide de la commande `stdev(X)` (le nom provient de l'anglais «standard deviation», l'écart-type).

L'**étendue** d'une série statistique X est la différence entre sa plus grande modalité et sa plus petite modalité. Il est possible de l'obtenir grâce à `max(X)-min(X)`.

La **médiane** est un nombre m qui sépare Ω en deux sous-populations $\Omega^- = \{\omega \in \Omega : X(\omega) < m\}$ et $\Omega^+ = \{\omega \in \Omega, X(\omega) > m\}$ de même effectif. Un tel nombre n'existe pas toujours, et n'est pas toujours unique.

Une convention couramment acceptée, et utilisée par SciLab est de dire que si les valeurs sont ordonnées :

$x_1 \leq x_2 \leq \dots \leq x_n$ alors

- si n est impair, la médiane est $x_{\frac{n+1}{2}}$ (la valeur du milieu)
- si n est pair égal à $2p$, la médiane est la demi-somme de x_p et x_{p+1} .

La médiane d'une série statistique s'obtient à l'aide la commande `median`.

La commande `tabul(X, 'i')` permet d'obtenir la **matrice des modalités-effectifs**. Il s'agit d'une matrice à deux colonnes $M = [m_1, m_2]$, où la première colonne contient l'ensemble des modalités triées par ordre croissant (on peut remplacer 'i' par 'd' pour les obtenir par ordre décroissant), et où la seconde colonne contient les effectifs correspondants. Par exemple,

si $X = [1, 4, 1, 6, 1, 1, 6, 1]$, alors $m_1 = \begin{pmatrix} 1 \\ 4 \\ 6 \end{pmatrix}$ et $m_2 = \begin{pmatrix} 5 \\ 1 \\ 2 \end{pmatrix}$.

La **fréquence d'une modalité** est son effectif divisé par l'effectif total de la population. Par exemple, dans la série précédente, la fréquence de la modalité 1 est $\frac{5}{8}$ et celle de la modalité 6 est $\frac{1}{4}$.

On appelle **fréquence cumulée** d'une modalité la somme de toutes les fréquences des modalités qui lui sont inférieures. Autrement dit, si les modalités sont $x_1 \leq \dots \leq x_n$, de fréquences associées f_1, \dots, f_n , alors la fréquence cumulée de la modalité x_i est $f_1 + \dots + f_i$.

Le polygone reliant les points d'abscisse x_i et d'ordonnée la fréquence cumulée associée à x_i est appelé **polygone des fréquences cumulées**.

Le **premier quartile** q_1 d'une série statistique X est la plus petite modalité de X ayant une fréquence cumulée supérieure ou égale à $\frac{1}{4}$. Autrement dit, c'est la plus petite valeur de la série telle que 25% des valeurs lui soient inférieures ou égales. De même, le **troisième quartile** q_3 est la plus petite modalité de la série telle que 75% des valeurs de la série lui soient inférieures ou égales.

Environ 50% des valeurs de la série statistique sont donc dans l'**intervalle interquartile** $[q_1, q_3]$.

La commande `quart(X)` permet d'obtenir les quartiles d'une série statistique X , voir l'exercice 2.

De même, on définit les **déciles** et les **centiles** d'une série statistique.

Remarque : toutes ces notions n'ont vraiment d'intérêt que si l'effectif de la population est beaucoup plus grand que 4 (pour les quartiles), 10 (pour les déciles) ou 100 (pour les centiles) : donner les centiles pour une population de 50 individus n'est pas pertinent.

EXERCICE 1 Créer un vecteur X de 1000 nombres choisis aléatoirement suivant une loi uniforme sur $[0, 10]$. Calculer la moyenne, l'écart-type, la médiane et les quartiles. Tous ces résultats étaient-ils prévisibles ?

F

EXERCICE 2 Tester l'instruction `quart` sur quelques séries statistiques, de petit effectif ($n \leq 10$). Que remarquez-vous ? En particulier concernant la deuxième ligne du vecteur retourné ?

F

EXERCICE 3 (Polygone des fréquences cumulées)

Une expérience consiste à lancer deux dés et à regarder la somme de ces deux dés.

AD

1. Créer un vecteur X simulant 1000 répétitions de cette expérience.
2. À l'aide de la commande `tabu`, créer la matrice des modalités-effectifs de X .
3. Créer un vecteur contenant les fréquences de chacune des modalités de cette série, puis un vecteur contenant les fréquences cumulées.
4. Tracer le polygone des fréquences cumulées. Les coordonnées du point le plus à droite de ce polygone étaient-elles prévisibles ?

REGROUPEMENT PAR CLASSES

EXERCICE 4 Simuler (à l'aide de la commande `grand`) 10000 nombres suivant une loi uniforme sur $[[1, 30]]$, et déterminer le nombre de modalités de cette série statistique.

PD

Même question avec une loi de Poisson de paramètre 30, puis une loi normale $\mathcal{N}(3, 4)$. Que constatez-vous ?

Lorsque $X = [x_1, \dots, x_n]$ est une série statistique dont le nombre de modalités est très grand, plutôt que de conserver toutes les valeurs, il est plus intéressant de les regrouper par classes.

Si l'on se donne une série statistique $X = [x_1, \dots, x_n]$, et $C = [c_1, \dots, c_k]$, une suite de réels rangés par ordre croissant, notons n_1 le nombre d'éléments de X qui sont dans l'intervalle $I_1 = [c_1, c_2]$, et pour $i \in \llbracket 2, k-1 \rrbracket$, notons n_i le nombre d'éléments de X qui sont dans l'intervalle $I_i =]c_i, c_{i+1}]$.

L'instruction `[ind, occ, non] = dsearch(X, C)` crée trois vecteurs :

- `ind` qui est de même taille que X , et qui à la position de x_i indique le numéro n de l'intervalle I_n dans lequel se trouve x_i .
- `occ` qui est de taille $k-1$ (où k est la taille de C), et qui contient en j -ème position le nombre d'éléments de X dans l'intervalle I_k .
- `non` est le nombre d'éléments qui ne sont dans aucun des intervalles I_n .

Par exemple, si $X = [4, 2, -1, 7, 10, 12, 15]$ et si $C = [0, 3, 6, 9, 12]$, alors :

- `ind = [2, 1, 0, 3, 4, 4, 0]` indiquant que 4 est dans l'intervalle $]3, 6]$ (le second intervalle), que 2 est dans l'intervalle $[0, 3]$ (le premier), que -1 n'est dans aucun intervalle, etc.
- `occ = [1, 1, 1, 2]` indiquant qu'il y a un élément dans le premier intervalle, un dans le second, etc.
- `non = 2` indiquant que deux des éléments de X ne sont dans aucun intervalle.

Remarque : les noms `ind`, `occ` et `non` peuvent être changés.

EXERCICE 5 Télécharger sur l'ENT le fichier nommé population_communes.sav

Dans SciLab, aller dans le menu Fichier/Charger un environnement, et ouvrir ce fichier.

Cela a pour effet de créer une matrice nommée population dont les 36 552 colonnes correspondent à la population de chacune des 36 552 communes de France métropolitaine (chiffres du recensement de 2012).

1. Quelle est la population totale de la France métropolitaine ? En moyenne, combien y a-t-il d'habitants par commune ?
2. Quel est le plus petit nombre d'habitants d'une commune ? Combien de communes possèdent ce nombre d'habitants ? Savez-vous quelles sont ces communes ?
3. Combien de communes comptent moins de 100 habitants ? Moins de 200 ? Combien d'habitants totalisent ces communes ? *On pourra au choix utiliser la commande tabul ou une boucle.*
4. Créer un vecteur contenant les nombres $-1, 999, 1999, \dots, 499999$. Puis, à l'aide de la commande dsearch, créer un vecteur P contenant le nombre de communes de moins de 1000 habitants, de 1000 à 1999 habitants, etc.
5. En utilisant le vecteur P, encadrer le nombre de français vivants dans des communes ayant entre 1000 et 9999 habitants. Peut-on retrouver ce nombre exact uniquement à l'aide de P ? Comment faire pour le retrouver ?

EXERCICE 6 Un résultat classique mentionne que si X suit une loi normale $\mathcal{N}(\mu, \sigma^2)$, alors

$P(\mu - \sigma \leq X \leq \mu) = P(\mu \leq X \leq \mu + \sigma) \approx 0.34$. De même, $P(\mu - 2\sigma \leq X \leq \mu - \sigma) = P(\mu + \sigma \leq X \leq \mu + 2\sigma) \approx 0.135$. En simulant un très grand nombre de réalisations d'une loi normale de votre choix, vérifier ces résultats, et déterminer la proportion de nombres dans chacun des intervalles $[\mu - 5\sigma, \mu - 4\sigma], \dots, [\mu + 4\sigma, \mu + 5\sigma]$.

En particulier, quelle proportion se trouvent dans l'intervalle $[\mu - 3\sigma, \mu + 3\sigma]$? Reconnaissez-vous là un résultat déjà rencontré ?

REPRÉSENTATIONS GRAPHIQUES

La fonction bar est particulièrement utile pour représenter graphiquement des données statistiques.

Si X et Y sont deux vecteurs de même taille, alors $\text{bar}(X, Y)$ représente des barres de hauteur y_i aux points d'abscisse x_i .

EXERCICE 7 Créer un vecteur X contenant 100 nombres tirés suivant une loi $\mathcal{B}(10, 0.5)$. À l'aide de la commande tabul, obtenir les fréquences de chacune des modalités $0, 1, \dots, 10$. Représenter alors les fréquences obtenues à l'aide d'un diagramme en bâtons.

Il est également possible de représenter plusieurs séries de données sur un même graphique afin de les comparer.

Si l'on se donne les effectifs de deux séries statistiques $Y1$ et $Y2$ sous forme de vecteurs colonnes (rappelons que pour passer d'un vecteur ligne à un vecteur colonne, il suffit de transposer), alors $\text{bar}(X, [Y1, Y2])$ trace sur un même graphique les diagrammes en barres correspondant aux deux séries.

EXERCICE 8 La commande binomial(p, n) crée un vecteur $X = (x_1, \dots, x_{n+1})$ tel que $x_k = P(Y = k - 1)$ où Y est une variable aléatoire de loi $\mathcal{B}(n, p)$ (autrement dit, X contient la loi de Y).

En utilisant cette commande, représenter graphiquement la loi d'une variable aléatoire de loi $\mathcal{B}(10, 0.3)$.

Simuler un échantillon de 1000 nombres distribués suivant une loi $\mathcal{B}(10, 0.3)$, représenter comme dans l'exercice précédent les fréquences associées à chacune des modalités, et **sur un même graphique**, représenter les probabilités théoriques, et celles observées sur l'échantillon.

Question subsidiaire : savez-vous obtenir la distribution théorique sans utiliser la commande binomial ?

EXERCICE 9 On reprend les données de l'exercice 5.

1. Créer un diagramme en bâtons représentant le nombre de communes comportant moins de 99 habitants, de 100 à 199 habitants, etc, de 2400 à 2499 habitants.
2. Créer un graphique similaire représentant cette fois, non plus le nombre de communes, mais le nombre de français habitant une commune de 0 à 99 habitants, de 100 à 199 habitants, etc.
3. Tracer une courbe (à l'aide de plot2d) sur laquelle, si on lit x en abscisse, on lit en ordonnée le nombre de français habitant une commune de moins de x habitants.

EXERCICE 10 Histogrammes

La fonction histplot permet de tracer des histogrammes. De manière précise, si $C = [c_1, \dots, c_k]$ est une suite de nombres rangés dans l'ordre croissant, et si $X = [x_1, \dots, x_n]$, alors en notant n_1 le nombre d'éléments de X compris dans $[c_1, c_2]$, n_i le nombre d'éléments de X dans l'intervalle $I_i =]c_i, c_{i+1}]$, la commande histplot(C, X) a pour effet de tracer une suite de rectangles de base I_i et d'aire proportionnelle à n_i (de manière précise, SciLab s'arrange pour que la somme des aires des rectangles soit égale à 1).

Créer un vecteur contenant 1000 réalisations d'une variable de loi $\mathcal{N}(0, 4)$.

Regrouper cette série statistique en 20 classes, et tracer l'histogramme associé.

Sur le même graphique, tracer, avec plot2d la fonction de densité de la loi normale $\mathcal{N}(0, 4)$. Que constatez-vous ?

CORRECTION DU TP 3

SOLUTION DE L'EXERCICE 1

```

1 X = grand(1,1000,'unf',0,10);
2 mean(X)
3 stdev(X)
4 median(X)
5 quart(X)

```

□

SOLUTION DE L'EXERCICE 2

On constate que les quartiles retournés par Sci Lab ne sont pas nécessairement des modalités de la série, mais sont plutôt calculés, à l'image de la médiane, comme la demi-somme de deux valeurs de la série.

Il s'agit essentiellement d'une convention, cela ne change en rien l'interprétation que l'on peut en faire. □

SOLUTION DE L'EXERCICE 3

```

1 X = grand(1,1000,'uin',1,6) + grand(1,1000,'uin',1,6);

```

1.

```

1 A = tabul(X,'i')

```

2. Il s'agit de remarquer que les effectifs sont dans la seconde colonne de la matrice obtenue avec `tabul`, et donc que les fréquences sont obtenues en divisant par 1000. De plus, les fréquences cumulées peuvent être obtenues sans douleur¹ en appliquant la commande `cumsum` au vecteur des fréquences.

```

1 frequences = (A(:,2))/1000;
2 frequences_cum = cumsum(frequences);

```

3. Comme les abscisses sont les nombres de 2 à 12, il s'agit de placer les points dont l'abscisse est un nombre compris entre 2 et 12 et l'ordonnée est la fréquence cumulée associée. Il s'agit ensuite de relier ces points².

```

1 plot2d(2:12, frequences_cum);

```

On constate que le point le plus à droite (correspondant à la plus grande modalité) a pour ordonnée 1, ce qui doit toujours être le cas, la somme des fréquences valant toujours 1. □

SOLUTION DE L'EXERCICE 4

```

1 A = grand(100,100,'uin',1,30);
2 B = tabul(A,'i')
3 disp(length(B)/2);

```

On constate que pour une loi uniforme, on obtient toujours 30 modalités, pour $\mathcal{P}(30)$, on obtient un nombre variable de modalités, mais toujours proche de 40. En revanche, pour une loi normale, on obtient 10 000 modalités. Cela est lié au fait qu'une variable à densité a une probabilité nulle de prendre une valeur donnée. □

SOLUTION DE L'EXERCICE 5

1. Il s'agit d'utiliser `sum(population)`, ce qui permet d'obtenir un nombre total d'habitants égal à 63375971.

Résultats

Pas de véritable surprise, on constate que la moyenne est proche de 5, tout comme la médiane, et que les deux quartiles sont proches de 2.5 et 7.5. Tout ceci est prévisible, car pour une loi uniforme $\mathcal{U}([0, 10])$, on a

$$P(U \leq 5) = 0.5$$

et

$$P(U \leq 2.5) = 0.25$$

Précision : le paramètre 'i' signifie *increasing* et permet de trier les valeurs dans l'ordre croissant. Si on omet ce paramètre, on obtient les valeurs dans l'ordre décroissant.

¹ En tous cas sans boucle !

² Et c'est exactement ce que fait la commande `plot2d` : placer les points dont on lui donne les coordonnées et les relier par des segments !

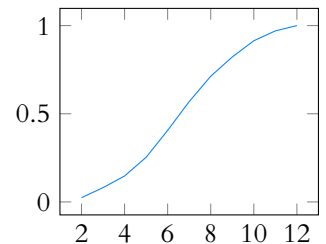


FIGURE 25.1– Le polygone des fréquences cumulées

Explication : on divise la taille de la matrice des modalités-effectifs car elle possède toujours deux colonnes, et donc s'il y a n modalités, elle comporte $2n$ coefficients.

2. Trouver le minimum n'est pas difficile, il suffit d'utiliser `min(population)`, qui nous apprend qu'il existe une (ou des) commune(s) possédant 0 habitants. Si l'on souhaite trouver ce nombre d'habitants, on peut utiliser, au choix `tabul(population, 'i')`, qui retourne une matrice dont la première ligne est `[0,6]`, et donc il y a six communes sans habitants. On peut également utiliser `length(find(population==0))`, qui retourne évidemment le même résultat.
3. On remarque que pour tout $k \in \llbracket 1, 200 \rrbracket$, il existe au moins une commune de k habitants. Donc il nous suffit de sommer les 100 (resp.) premières lignes de la matrice des modalités effectifs. On peut par exemple à cet effet utiliser astucieusement la commande `cumsum`.

```
1 A = tabul(population, 'i');
2 B = cumsum(A(:,2));
3 disp(B(100))
```

On peut aussi préférer faire une boucle parcourant les 200 premières lignes de la matrice des modalités-effectifs.

```
1 A = tabul(population, 'i');
2 s = 0;
3 for i=1 :100
4     s = s+A(i,2);
5 end
```

Enfin, on peut si on le souhaite parcourir toute la matrice M .

```
1 length(find(population<100))
```

Dans tous les cas, on trouve bien 3477 communes avec strictement moins de 100 habitants et 9381 avec strictement moins de 200 habitants.

En revanche, ceci ne suffit absolument pas à déterminer le nombre total d'habitants de ces communes. Au mieux, cela permet de donner un encadrement : il y a entre 0 et 99×3477 habitants dans les communes de moins de 100 habitants.

Le nombre total d'habitants est alors obtenu via

```
1 s = 0;
2 for i=1 :100
3     s = s+A(i,1)*A(i,2);
4 end
```

Les communes de moins de 100 habitants totalisent alors 225 930 habitants, et celles de moins de 200 habitants en représentent 1 096 574 habitants.

4. Il s'agit assurément d'un travail pour `dsearch` ! Et d'ailleurs, le P que l'on souhaite obtenir est précisément le `occ` décrit précédemment (nombre d'individus dans chaque catégorie).

```
1 u = [-1 :1000 :499999];
2 [ind,P,non] = dsearch(population,u);
```

Notons au passage que `non = 2`, ce qui signifie que seules Paris et Marseille ont plus de 500 000 habitants.

```
5.
1 mini = 0;
2 maxi = 0;
3 for i=2 :10
4     mini = mini + u(i)*P(i);
5     maxi = maxi + u(i+1)*P(i);
6 end
```

On obtient alors `mini = 19 688 228` et `maxi = 28 460 228`.

On ne peut se contenter de P pour retrouver le nombre exact, et il faudra au minimum la

Pour la culture

Ces six communes sont les communes «mortes pour la France», qui ont été intégralement détruites lors de la bataille de Verdun et jamais reconstruites.

Indices

Ce programme donne le nombre de villages de strictement moins de 100 habitants, car la 100^{ème} ligne de A correspond aux villages de 99 habitants, du fait de la présence de communes non habitées.

Méthode

C'est tout le défi des statistiques : essayer de résumer une grande série de nombres par quelques chiffres clés, tout en gardant à l'esprit ce qu'on peut faire dire à ces nombres et ce qu'on ne peut leur faire dire !

Ici, si on ne garde que les nombre d'habitants de communes de moins de 100, 200, ... habitants, alors on ne peut répondre à la question avec exactitude. En revanche, si on a la matrice des modalités-effectifs on peut y répondre. Toutefois, cette matrice possède 11264 coefficients, donc on peut légitimement se demander si on a vraiment résumé les données initiales...

matrice des modalités-effectifs et procéder comme précédemment. Le résultat exact est alors 23 386 596.

Notons qu'une solution très «Scilabique» est la suivante :

```
1 sum((A(:,1)<10000).*(A(:,1)>999).*A(:,2).*A(:,1))
```

□

SOLUTION DE L'EXERCICE 7

```
1 X = grand(1,100,'bin',10,0.5);
2 A = tabul(X,'i');
3 bar(A(:,1),A(:,2)/100);
```

□

SOLUTION DE L'EXERCICE 8

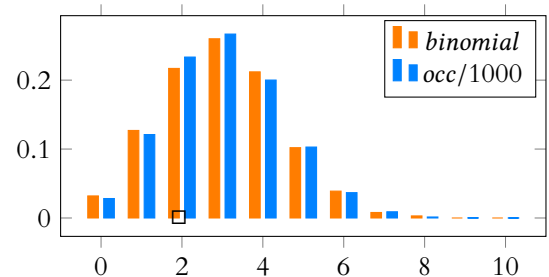
Pour représenter la loi $\mathcal{B}(10, 0.3)$, il suffit d'utiliser.

```
1 bar(0:10,binomial(0.3,10));
```

Simuler 1000 nombres suivant la loi binomiale ne pose pas de problème, et les regrouper par valeurs non plus si on utilise `tabul`. Toutefois, il est plus que probable qu'on n'obtienne pas tous les nombres de 0 à 10, le 9 et le 10 étant très peu probables. Et alors on aura un vecteur qui n'a pas la même taille que `binomial(0.3,10)`, de sorte que Scilab refuse de les tracer sur un même graphique.

Une option est par exemple d'utiliser `dsearch` plutôt que `tabul`, quitte à obtenir des 0.

```
1 u = -1 :10
2 A = grand(1,1000,'bin',10,0.3);
3 [ind,occ,non] = dsearch(A,u);
4 bar([0:10]',[binomial(0.3,10);occ/1000]')
```



SOLUTION DE L'EXERCICE 9

- On considère toujours que A est la matrice des modalités effectifs obtenue dans l'exercice 5.

```
1 u = [-1 :100 :2499];
2 [ind,occ,non] = dsearch(A,u);
3 bar([100 :100 :2500],occ)
```

- Il faut être un peu plus subtil que dans l'exercice 5 : si il y avait bien des villages avec 1, 2, 3, ..., 200, il est fort douteux qu'il y ait une commune avec 2342 habitants, une avec 2343 habitants, etc. Autrement dit, la i -ème ligne de la matrice des modalités-effectifs ne correspond pas nécessairement aux communes de $i + 1$ habitants.

```
1 u = [-1 :100 :2499];
2 v = zeros(1,25)
3 s = min(find(A(:,1)>=2500))-1; //le numero de la derniere ligne
4 // de A qui nous interesse
5 for i=1 :s
6     c = floor(A(i,1)/100)+1; // le numero de la categorie d'une commune
7     v(c) = v(c) + A(i,1)*A(i,2);
8 end
9 bar([100 :100 :2500],v)
```

- Il y a ici une petite astuce : $A(:,1). * A(:,2)$ représente le nombre de français vivant dans une commune de 0 habitants, de 1 habitant, etc... On peut donc tracer la courbe de la manière suivante :

```
1 B = cumsum(A(:,1).*A(:,2));
2 plot2d(A(:,1),B)
```

Remarque

Une telle courbe manque d'intérêt, tout simplement car elle n'est pas lisible : en abscisse se trouvent les nombres des 1 à 2000000, alors que seule une commune a plus d'un million d'habitants. Donc les informations que l'on pourrait en extraire sont en fait illisibles. Peut-être qu'une telle courbe a de l'intérêt si on se cantonne aux villes de moins de 10 000 ou 100 000 habitants.

□

SOLUTION DE L'EXERCICE 10

Simulons donc 1000 réalisations d'une loi $\mathcal{N}(0, 4)$:

```
A = grand(1000, 1, 'nor', 0, 2)
```

Comme nous ne savons pas a priori quels seront les plus grandes et les plus petites valeurs prises par A, nous pouvons utiliser

```
x = linspace(min(A), max(A), 20)
```

Et donc pour tracer l'histogramme, il suffit de

```
histplot(x, A)
```

Enfin, pour tracer la densité, nous pouvons créer une fonction f correspondant à cette densité, puis la tracer à l'aide de fplot2d :

```
function y = f(x)
    y = 1/(2*sqrt(2*pi))*exp(-1/2*(x/2)^2)
endfunction
```

```
fplot2d(x, f)
```

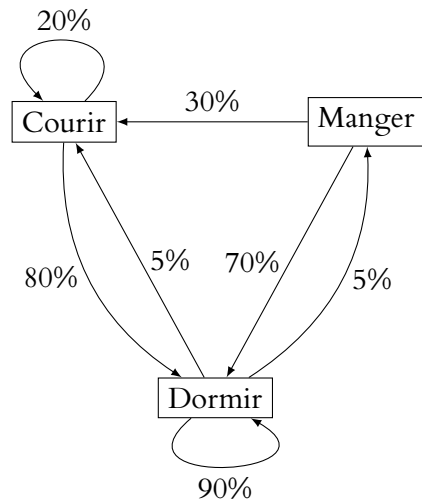
□

TP 4 : CHAÎNES DE MARKOV I

PRÉLUDE : DOUDOU LE HAMSTER

Doudou le hamster n'a que trois occupations dans sa cage : il dort, il court dans sa roue, et il mange. Chaque minute, il change d'activité ou continue celle qu'il était en train de faire. On suppose que :

- s'il dort à un instant donné, la minute suivante il y a 9 chances sur 10 qu'il dorme encore la minute suivante, mais s'il se réveille, il va soit manger, soit courir, les deux étant aussi probables l'un que l'autre.
- s'il mange, il y a 3 chances sur 10 qu'il parte dans sa roue, et 7 chances sur 10 qu'il retourne se coucher.
- s'il court, il y a 8 chances sur 10 qu'il aille se coucher à l'instant suivant, et sinon, il court encore un peu.



En moyenne, combien de temps Doudou consacre-t-il à chacune de ses activités ? S'il dort au moment où vous le regardez, quelle est la probabilité qu'il dorme 10 minutes plus tard ? Etc

LES CHAÎNES DE MARKOV

Considérons un ensemble E fini, dont les éléments sont notés x_1, \dots, x_N et appelés états.

Une **chaîne de Markov** sur E est une suite $(X_n)_{n \in \mathbb{N}}$ de variables aléatoires à valeurs dans E (pour simplifier, considérons dans la suite que $E = \llbracket 1, N \rrbracket$, de sorte qu'il s'agisse de variables aléatoires réelles, comme nous en avons l'habitude, même si ici cela n'a pas d'importance) telle que

$\forall n \in \mathbb{N}$, la loi de X_{n+1} sachant X_0, X_1, \dots, X_n est égale à la loi de X_{n+1} sachant X_n . De manière formelle, cela signifie que

$$\forall (i_0, \dots, i_n, i_{n+1}) \in \llbracket 1, N \rrbracket^{n+1}, P_{[X_0=i_0] \cap \dots \cap [X_n=i_n]}(X_{n+1} = i_{n+1}) = P_{[X_n=i_n]}(X_{n+1} = i_{n+1}).$$

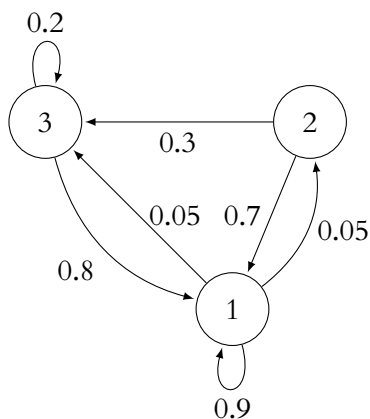
De manière bien plus informelle, cela signifie que la loi du futur (instant $n+1$) ne dépend que du présent (instant n) et non du passé (instants $0, 1, \dots, n-1$).

Nous supposons de plus que la propriété suivante est vérifiée :

$$\forall n \in \mathbb{N}, \text{ la loi de } X_{n+1} \text{ sachant } X_n \text{ est égale à la loi de } X_1 \text{ sachant } X_0.$$

Autrement dit, les lois régissant le passage d'un instant au suivant (de X_{n+1} à X_n) ne changent pas au cours du temps.

Exemple 1 : la chaîne de Markov de Doudou



Numérotons les trois activités de Doudou de 1 à 3 : par exemple, 1 pour «dormir», 2 pour «manger» et 3 pour «courir».

Supposons dans un premier temps que Doudou dort à l'instant 0, et notons X_0 la variable aléatoire certaine égale à 1, et pour tout n , notons X_n la variable aléatoire indiquant l'activité de Doudou à la n -ième minute.

Alors $(X_n)_{n \in \mathbb{N}}$ est une chaîne de Markov sur $\llbracket 1, 3 \rrbracket$.

En effet, les règles régissant son activité future ne dépendent que de son activité présente (ce qui est cohérent avec le peu de mémoire vive dont dispose un hamster), et ces règles ne changent pas au cours du temps.

De plus, les règles énoncées précédemment nous donnent la loi de X_{n+1} sachant X_n (on parle de **probabilités de transition**). Par exemple, $P_{[X_n=1]}(X_{n+1} = 1) = \frac{9}{10}$ et $P_{[X_n=3]}(X_{n+1} = 2) = 0$.

EXERCICE 1 En supposant que $X_0 = 1$ (c'est-à-dire que Doudou dort à l'instant 0), déterminer les lois de X_1 et X_2 , puis calculer leurs espérances. Que signifient ces espérances ?

SIMULATION NAÏVE

Si l'on souhaite simuler une chaîne de Markov, une première possibilité est d'écrire une fonction qui prend en entrée un état de la chaîne, et qui retourne la valeur (aléatoire) à l'instant suivant.

EXERCICE 2

- On souhaite écrire une fonction `doudou(i)` qui prend en entrée un paramètre $i \in \llbracket 1, 3 \rrbracket$ correspondant à l'activité de Doudou à l'instant n , et simule l'activité de Doudou à l'instant $n + 1$ (c'est-à-dire qui retourne un nombre de $\llbracket 1, 3 \rrbracket$ suivant les règles précédemment énoncées). Par exemple, si $i = 2$, `doudou(2)` devra retourner 3 avec probabilité 0.3 et 1 avec probabilité 0.7.
Télécharger le fichier `doudou.sci` sur l'ENT, l'ouvrir dans Scinotes et le compléter en remplaçant les `****`.
- On suppose que Doudou dort à l'instant 0. En appelant 10 000 fois `doudou`, et en étudiant les fréquences d'apparition de chacun des états, retrouver la loi de X_1 déterminée à l'exercice 1.

EXERCICE 3 Doudou mange à l'instant initial.

- Que va simuler la commande `doudou(doudou(2))` ? Et la commande `doudou(doudou(doudou(2)))` ?
- Écrire une fonction `plus_tard(n)` qui utilise la fonction `doudou` pour simuler l'activité de Doudou à l'instant n .
- À l'aide d'une (ou plusieurs) boucle et de la fonction `plus_tard`, créer une matrice 100×100 qui simule l'activité de 10000 Doudous quatre heures après l'instant initial.
- Quelle répartition des activités de Doudou observez-vous ?
- Modifier votre code en supposant que Doudou court à l'instant initial, puis que Doudou dort. Qu'observez-vous ?

MATRICE DE TRANSITION D'UNE CHAÎNE DE MARKOV

Exemple 2 : une compagnie d'assurances automobiles met en place un système de bonus-malus pour ses clients. Ceux-ci peuvent bénéficier de bonus de 10, 20, 30 ou 40%, et des mêmes malus.

Il sont donc répartis en 9 catégories, numérotées de 1 à 9, la première correspondant à un bonus de 40%, la suivante à un bonus de 30%, etc, jusqu'à la dernière qui correspond à un malus maximal.

On considère que chaque année, un conducteur a

- 0 accident avec probabilité $2/3$
- 1 accident avec probabilité $1/6$
- 2 accidents avec probabilité $1/12$
- 3 accidents avec probabilité $1/12$

Si un conducteur n'a pas d'accident, alors il gagne 10% de bonus (sauf si celui-ci est déjà maximal) et s'il a k accidents, il perd $k \times 10\%$ de bonus (sauf si son malus est déjà maximal).

En notant alors X_n le numéro de la catégorie du conducteur au bout de n années, nous avons une chaîne de Markov sur $\llbracket 1, 9 \rrbracket$.

Nous pourrions l'étudier comme nous l'avons fait pour Doudou en écrivant une fonction qui, étant donnée la catégorie du conducteur l'année n , retourne sa catégorie l'année $n + 1$.

Toutefois, l'écriture d'une telle fonction serait relativement laborieuse (il faudrait distinguer les conducteurs ayant déjà un bonus maximal, mais également ceux ayant un malus maximal) et nous proposons ici une manière commode de représenter les chaînes de Markov à l'aide de matrices.

Définition 4.1 – On appelle matrice de transition d'une chaîne de Markov sur $\llbracket 1, N \rrbracket$ la matrice $M = (m_{i,j}) \in \mathcal{M}_N(\mathbf{R})$ définie par

$$\forall (i, j) \in \llbracket 1, N \rrbracket^2, m_{i,j} = P_{[X_0=i]}(X_1 = j)$$

EXERCICE 4 Écrire la matrice de transition de la chaîne de Markov de Doudou. Calculer la somme de chacune des lignes. Que remarquez-vous ? Est-ce surprenant ?

Proposition 4.2 : La somme de chacune des lignes de la matrice de transition d'une chaîne de Markov vaut 1 (on dit alors que la matrice est *stochastique*).

EXERCICE 5 Prouver ce résultat.

Définition 4.3 – On appelle **état probabiliste** un vecteur ligne (p_1, \dots, p_N) tel que

$$\forall i \in \llbracket 1, N \rrbracket, p_i \geq 0 \text{ et } \sum_{i=1}^N p_i = 1.$$

Un état probabiliste permet de représenter facilement la loi d'une variable aléatoire à valeurs dans $\llbracket 1, N \rrbracket$. Par exemple, si Doudou dort à l'instant initial, alors la loi de X_0 est représentée par le vecteur $(1, 0, 0)$, et la loi de X_1 est représentée par $(0.9, 0.05, 0.05)$.

Proposition 4.4 : Si μ_i est l'état probabiliste représentant la loi de X_i , alors, par la formule des probabilités totales,

$$\forall i \in \mathbf{N}, \mu_{i+1} = \mu_i M.$$

Une récurrence permet alors de prouver que $\forall i \in \mathbf{N}, \mu_i = \mu_0 M^i$.

EXERCICE 6

1. Créer la matrice M qui est la matrice de transition de la chaîne de Markov de Doudou.
2. Déterminer sans boucle la loi de X_{1440} (1440 étant le nombre de minutes dans une journée) si Doudou dort à l'instant initial.
3. Faire de même si Doudou mange à l'instant initial, ou s'il court. Que remarquez-vous ?

EXERCICE 7 Revenons à notre compagnie d'assurance.

1. Écrire la matrice de transition de la chaîne de Markov associée. La rentrer dans Sci Lab (ou utiliser le fichier assurance.sci disponible sur l'ENT).
2. Un jeune conducteur commence son parcours d'assuré sans bonus ni malus. Quelle est la probabilité que 30 ans après l'obtention de son permis, il ait un malus maximal ? Un bonus maximal ? 10 % de bonus ?
3. Représenter sur un graphique la proportion de conducteurs dans chaque catégorie au bout de 40 ans, en supposant que tous commencent sans bonus ni malus.
4. La compagnie s'est spécialisée dans les mauvais conducteurs (qu'elle assure à prix d'or !), et on suppose que les nouveaux assurés ont tous un malus maximal. Au bout de 30 ans, quelle proportion d'assurés auront toujours un malus maximal ? *Question facultative : tracer sur un graphique l'évolution de la proportion d'assurés ayant un malus de 40 %.*

LA COMMANDE grand AVEC L'OPTION markov

La commande Sci Lab `grand(n, 'markov', M, x0)` permet de simuler les n états successifs d'une chaîne de Markov de matrice de transition M , en supposant que l'état initial est l'état x_0 . Le résultat est retourné sous forme d'un vecteur ligne. Par exemple, si M est la matrice de transition de Doudou, alors `grand(60, 'markov', M, 1)` simule une heure d'activité d'un Doudou initialement au repos.

EXERCICE 8 Doudou mange à l'instant 0. Simuler une journée complète de Doudou. Au cours de cette journée, combien de fois est-il passé de son lit à sa roue ? Combien de fois a-t-il fait de la roue deux minutes ou plus ?

EXERCICE 9 Le tarif de base de la compagnie d'assurance est de 500 euros, et une personne avec $x\%$ de bonus paiera donc $500(1 - \frac{x}{100})$ euros.

1. Montrer que si une année un assuré est dans la catégorie i ($i \in \llbracket 1, 9 \rrbracket$), alors il paiera cette année là $500 * (1 + \frac{i-5}{10})$ euros.
2. Simuler le parcours d'assuré d'un jeune conducteur qui resterait 30 ans chez le même assureur. Quelle est la somme totale versée à l'assureur durant ces 30 ans ?

3. Si x_0 n'est pas juste un état, mais un vecteur colonne $\begin{pmatrix} p_1 \\ \vdots \\ p_k \end{pmatrix}$ avec les $p_i \in \llbracket 1, N \rrbracket$, alors `grand(n, 'markov', M, x0)` retourne une matrice à k lignes et n colonnes, chacune des lignes représentant le parcours d'un assuré dont l'état initial serait p_i . Simuler ainsi le parcours d'assuré de 1000 jeunes conducteurs durant 30 ans. En moyenne, combien ont-ils versé à la compagnie ?

CORRECTION DU TP 4

SOLUTION DE L'EXERCICE 1

On a, par la formule des probabilités totales, appliquée au système complet d'événements $\{[X_0 = 1]\}$,

$$P(X_1 = 1) = P_{[X_0=1]}(X_1 = 1) = 0.9, P(X_1 = 2) = P_{[X_0=1]}(X_1 = 2) = 0.05, P(X_1 = 3) = P_{[X_0=1]}(X_1 = 3) = 0.05.$$

On en déduit que $E(X_1) = 0.9 + 2 \times 0.05 + 3 \times 0.05 = 1.15$.

De même, par la formule des probabilités totales appliquée au système complet d'événements $\{[X_1 = 1], [X_1 = 2], [X_1 = 3]\}$,

$$P(X_2 = 1) = 0.9P(X_1 = 1) + 0.8P(X_1 = 3) + 0.7P(X_1 = 2) = 0.885$$

et $P(X_2 = 2) = 0.045, P(X_2 = 3) = 0.07$.

On a alors $E(X_2) = 1.185$.

Ces espérances ne représentent **rien** ! En effet, nous avons arbitrairement numéroté les activités de Doudou de 1 à 3, et si on les numérote dans un autre ordre, les espérances de X_1 et X_2 vont changer. En revanche, les lois de X_1 et X_2 nous donneront toujours les probabilités que Doudou soit en train de faire telle ou telle activité à l'instant 1 ou à l'instant 2. \square

SOLUTION DE L'EXERCICE 2

- Il faut distinguer trois cas, et à chaque fois distinguer de nouveau les différents cas, ce qui se fait à l'aide de nombreux if

```

1  function y = doudou(n)
2      p = rand()
3      y=n
4      if n==1 then
5          if p <=0.05 then
6              y = 2
7          else if p<=0.1 then
8              y=3
9          end
10         end
11     end
12     if n==2 then
13         if p<=0.3 then
14             y=3
15         else
16             y=1
17         end
18     end
19     if n==3 then
20         if p<=0.8 then
21             y=1
22         else
23             y=3
24         end
25     end
26 endfunction

```

- Appelons 10000 fois la fonction doudou, et regardons combien de fois nous sont retournés chacun des résultat. Pour cela, on peut par exemple créer un vecteur A de taille 1×3 , qui stocke le nombre de fois où Doudou sera en train de dormir et le nombre de fois où Doudou sera en train de courir.

Remarque

On sait que si Doudou mange à l'instant initial, il ne mangera pas à l'instant 1.

```

1 A = zeros(1,3) ;
2 for i = 1 :10000
3     x = doudou(2) ;
4     A(x) = A(x)+1 ;
5 end
6 A/10000

```

On obtient alors quelque chose ressemblant à $A = [0.7047, 0, 0.2953]$, ce qui est très proche des valeurs théoriques (0.7 et 0.3).

3. `doudou(2)` représente l'activité de Doudou à l'instant 1, et donc `doudou(doudou(2))` représente son activité à l'instant suivant, soit à l'instant 2.

□

SOLUTION DE L'EXERCICE 3

1. `doudou(2)` simule l'activité de Doudou à l'instant 1 (puisque 2 est son activité, à savoir «manger» à l'instant 0).
Et donc `doudou(doudou(2))` simule l'activité de Doudou un instant plus tard que l'instant 1, c'est-à-dire à l'instant 2.
De même, si on ajoute un appel à `doudou`, nous obtiendrons l'activité de notre rongeur à l'instant 3
2. On utilise le principe de la question précédente : pour simuler l'activité de Doudou à l'instant n , il faut composer n fois la fonction `doudou` par elle-même.

```

1 function y = plus_tard(n)
2     y = 2 // l'instant initial
3     for i= 1 :n
4         y = doudou(y) ;
5     end
6 endfunction

```

3.

```

1 A = zeros(1,3) ;
2 for i = 1 :10000
3     x = plus_tard(240) ;
4     A(x)= A(x) +1 ;
5 end
6 A/10000

```

4. On observe que 4h après l'instant initial, Doudou dort avec une probabilité de 0.89, mange avec une probabilité 0.04 et court avec une probabilité 0.07.
5. Il suffit de modifier la seconde ligne de la fonction `plus_tard`, par exemple en remplaçant `y=2` par `y = 1` pour que Doudou soit en train de dormir au temps 0.
On remarque alors qu'au bout de 4h, la répartition des activités de Doudou reste globalement la même, quelle que soit son activité de départ.

□

SOLUTION DE L'EXERCICE 4

On a $M = \begin{pmatrix} 0.9 & 0.05 & 0.05 \\ 0.7 & 0 & 0.3 \\ 0.8 & 0 & 0.2 \end{pmatrix}$. On constate que la somme de chacune des lignes vaut 1,

ce qui n'est pas surprenant car $\{[X_1 = 1], [X_1 = 2], [X_1 = 3]\}$ forme un système complet d'événements, et donc

$$\forall i \in \llbracket 1, 3 \rrbracket, \sum_{j=1}^3 P_{[X_0=i]}(X_1 = j) = 1.$$

□

SOLUTION DE L'EXERCICE 5

Trivial

◀ Dans 4 heures, il y a 240 minutes.

Arrondis

◀ Ce sont des valeurs arrondies. Remarquons également qu'il s'agit de simulations, donc rien ne garantit que ce soit bien la loi de X_{240} , il reste une part de hasard derrière. Toutefois, si l'on appelle plusieurs fois le programme précédent, on remarque qu'on trouve à chaque fois des valeurs proches.

Le coefficient à la $i^{\text{ème}}$ ligne et $j^{\text{ème}}$ colonne de la matrice de transition est $P_{[X_0=i]}(X_1 = j)$. Mais X_1 est une variable aléatoire à valeurs dans $\llbracket 1, N \rrbracket$, de sorte que $\{[X_1 = j], j \in \llbracket 1, N \rrbracket\}$ est un système complet d'événements. Donc la somme des coefficients de la i -ème ligne est

$$\sum_{j=1}^n P_{[X_0=i]}(X_1 = j) = P_{[X_0=i]}(\Omega) = 1.$$

□

SOLUTION DE L'EXERCICE 6

- On a $M = \begin{pmatrix} 0.9 & 0.05 & 0.05 \\ 0.7 & 0 & 0.3 \\ 0.8 & 0 & 0.2 \end{pmatrix}$.
- D'après la proposition précédente, on a $\mu_{1440} = \mu_0 M^{1440}$ et donc on peut utiliser
`1 [1,0,0] * M^1440`
- Si Dou dou mange à l'état initial, il suffit de remplacer l'état initial, qui était (1, 0, 0) par (0, 1, 0). On observe alors que les probabilités obtenues sont les mêmes, quel que soit l'état initial qu'on a choisi.

□

SOLUTION DE L'EXERCICE 7

- La matrice de transition de cette chaîne de Markov est

$$A = \begin{pmatrix} 2/3 & 1/6 & 1/12 & 1/12 & 0 & 0 & 0 & 0 & 0 \\ 2/3 & 0 & 1/6 & 1/12 & 1/12 & 0 & 0 & 0 & 0 \\ 0 & 2/3 & 0 & 1/6 & 1/12 & 1/12 & 0 & 0 & 0 \\ 0 & 0 & 2/3 & 0 & 1/6 & 1/12 & 1/12 & 0 & 0 \\ 0 & 0 & 0 & 2/3 & 0 & 1/6 & 1/12 & 1/12 & 0 \\ 0 & 0 & 0 & 0 & 2/3 & 0 & 1/6 & 1/12 & 1/12 \\ 0 & 0 & 0 & 0 & 0 & 2/3 & 0 & 1/6 & 1/6 \\ 0 & 0 & 0 & 0 & 0 & 0 & 2/3 & 0 & 1/12 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 2/3 & 1/3 \end{pmatrix}$$

- Si un conducteur commence sans bonus ni malus, il commence son parcours d'assuré en catégorie 5, donc son état initial est (0 0 0 0 1 0 0 0 0).
- Notons X_i la catégorie de l'assuré à l'issue de la i -ème année. Alors la loi de X_i est donnée par $\mu_0 \times A^i$. En particulier, on a $\mu_{30} = \mu_0 \times A^{30}$. En Scilab, cela donne

```
1 u = zeros(1,9);
2 u(5) = 1;
3 v = u*A^30
```

On obtient alors le vecteur :

$$v = [0.2269 \quad 0.1136 \quad 0.1139 \quad 0.1141 \quad 0.1001 \quad 0.0931 \quad 0.0861 \quad 0.0790 \quad 0.0728]$$

On en déduit que la probabilité qu'un assuré possède au bout de 30 ans un malus maximal est de 0.073. La probabilité que son bonus soit maximal est 0.227 et celle qu'il n'ait que 10% de bonus est 0.114.

- Il s'agit de représenter le vecteur v que nous venons d'obtenir, ce qui est très facile en utilisant la commande `bar(v)`. On obtient alors le graphique :

Attention

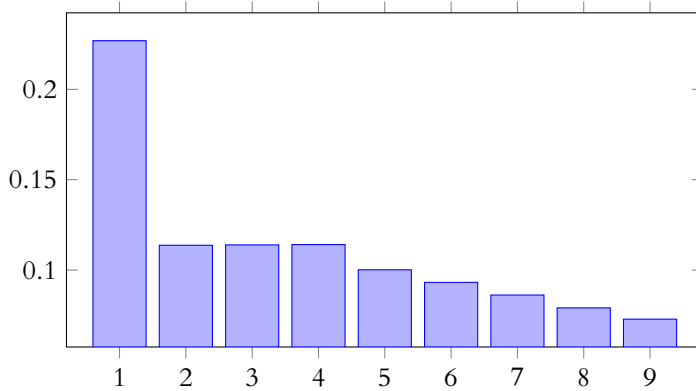
On se méfiera du fait que si un assuré est en première catégorie, alors il ne peut «gagner» encore une catégorie, donc reste dans cette catégorie s'il n'a pas d'accident dans l'année.
 De même pour les assurés qui auraient «trop» d'accidents : par exemple un assuré avec 40% de malus ne peut en avoir plus, et donc garde le malus maximal avec probabilité 1/3.

Autrement dit

Cet état représente une probabilité, mais si l'on est certain qu'il commence en catégorie 5, alors la variable X_0 désignant son état initial est la variable certaine égale à 5.

Remarque

Si l'on tape `sum(v)`, on remarque que la somme des composantes de ce vecteur vaut bien 1.



5. On suppose cette fois que l'état initial est $(0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 1)$. On a alors

```
1 u = zeros(1,9);
2 u(9) = 1;
3 v = u*A^30
```

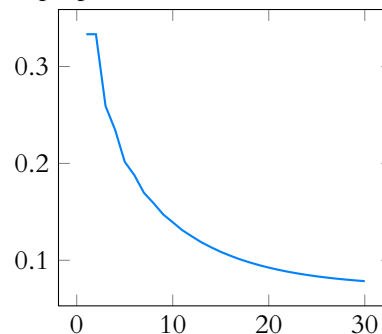
Ce code nous retourne

$[0.2125 \ 0.1088 \ 0.1111 \ 0.1133 \ 0.1022 \ 0.0971 \ 0.0913 \ 0.0848 \ 0.0784]$

On constate alors que les proportions d'assurés des différentes catégories sont très proches de celles obtenues précédemment.

Si l'on veut garder une trace du nombre d'assurés de malus maximal à chaque catégorie, il faut calculer un par un les produits $\mu_0 \times A^i$, par exemple en utilisant le fait que $\mu_{i+1} = \mu_0 \times A^{i+1} = \mu_i \times A$, et extraire à chaque étape la proportion d'assurés de malus maximal.

```
1 u = zeros(1,9);
2 u(9) = 1;
3 s = zeros(1,30);
4 for i=1 :30
5     u = u*A;
6     s(i) = u(9);
7 end
8 plot2d(s);
```



□

SOLUTION DE L'EXERCICE 8

Si Doudou mange à l'instant 0, alors l'état initial de Doudou est l'état 2. On va donc simuler une journée complète de Doudou en utilisant

```
1 B = grand(1440, 'markov', M, 2)
```

Le vecteur ligne ainsi obtenu représente les différentes activités de Doudou au cours de cette journée.

Si l'on souhaite connaître le nombre de fois où Doudou est passé de son lit à sa roue, il s'agit de trouver combien B contient de 1 suivis de 3.

```
1 s = 0
2 for i = 1 :1439
3     if B(i)==1 & B(i+1)==3 then
4         s = s+1
5     end
6 end
```

On obtient alors un nombre de l'ordre de 60 (mais qui peut-être très variable d'une simulation à l'autre).

Danger

Attention : on a simulé là l'activité de Doudou au cours d'une journée. Sur une autre journée, ses activités seront peut-être réparties différemment !

Astuce

Si l'on veut se dispenser d'utiliser un & dans la condition, on peut remarquer que cette condition est équivalente à $B(i)*B(i+1) == 3$.

Pour compter le nombre de fois où Doudou a fait de la roue deux minutes ou plus, on peut compter le nombre de fois où dans la matrice on a un 3 suivi d'un autre 3, lui-même suivi d'un 1 ou d'un 2.

```

1 s = 0
2 for i = 1 :1438
3     if B(i)==3 & B(i+1)==3 & B(i+2)<>3 then
4         s = s+1
5     end
6 end

```

□

SOLUTION DE L'EXERCICE 9

1. $5 - i$ représente le bonus¹ divisé par 10. Ainsi, si l'individu est dans la 2-ème catégorie, alors il a un bonus de 30% et $5 - i = 3$. Donc le pourcentage du tarif normal que va payé l'assuré est $100 - 10(5 - i)$, de sorte que la somme payée est $500 * \frac{100 - 10(5 - i)}{100} = 500 \left(1 + \frac{i - 5}{10}\right)$.
2. Pour simuler le parcours d'un assuré, on peut utiliser grand avec l'option markov. Il s'agit ensuite de calculer la somme totale en utilisant le vecteur ainsi obtenu et en calculant, pour chaque année² la somme payée.

```

1 p = grand(30, 'markov', A, 5) ;
2 s = 0 ;
3 for i=1 :30
4     s = s + 500*(1+(p(i)-5)/10) ;
5 end
6 disp(s)

```

Les résultats obtenus sont très variables³, et sont en gros compris entre 10000 et 20000.

3. Il s'agit de reprendre le même principe, en utilisant comme dernier paramètre de grand la matrice colonne à 1000 lignes dont tous les coefficients valent 5.

```

1 p = grand(30, 'markov', A, 5*ones(1000, 1)) ;
2 s = zeros(1, 1000) ;
3 for j = 1 :1000
4     for i=1 :30
5         s(j) = s(j) + 500*(1+(p(j, i)-5)/10) ;
6     end
7 end
8 disp(mean(s))

```

On obtient alors une somme moyenne proche de 14000 euros.

□

Subtil...

Si on ne demande pas à ce que le second 3 soit suivi d'un 1 ou d'un deux 2, on risque de compter plusieurs fois une longue séquence de roue. Par exemple si Doudou fait de la roue trois minutes d'affilée, B va contenir trois 3 consécutifs, que nous risquons de compter comme deux séances de roue alors qu'il s'agit d'une seule et même séance.

Signes

Un bonus est une remise : si on a 10% de bonus, on paye 10% moins cher !

¹ où un malus est considéré comme un bonus négatif

² c'est-à-dire pour chaque composante du vecteur obtenu

³ En effet, tout dépend du parcours de l'assuré ainsi simulé : s'il n'a jamais eu d'accident ou au contraire s'il en a eu tous les ans.

Moyenne

Cette fois, puisqu'on a fait la moyenne sur un grand nombre d'assurés, la somme obtenue ne devrait plus beaucoup dépendre des assurés qu'on a simulé : normalement (sauf accident), on a simulé à peu près tous les types d'assurés. Prendre encore plus d'assurés permettrait de réduire l'incertitude (même si l'on n'est jamais à l'abri d'une simulation exceptionnelle, où par exemple on n'aurait simulé que des conducteurs n'ayant jamais d'accidents, ce qui est extrêmement rare.)

TP 5 : CHAÎNES DE MARKOV II

Le contexte est le même que dans le TP 4.

LOI STATIONNAIRE D'UNE CHAÎNE DE MARKOV

Nous avons constaté sur l'exemple de Doudou le hamster que lorsque n devient grand, la loi de X_n ne dépend plus de l'état initial.

Plus précisément, (sous quelques hypothèses techniques dont nous reparlerons plus tard) si M est la matrice de transition de la chaîne de Markov, alors quel que soit le vecteur probabiliste μ_0 décrivant l'état initial, on a, pour n suffisamment grand,

$$\mu_n = \mu_0 M^n \simeq \mu$$

où μ ne dépend pas de n .

De manière plus formelle, on a le résultat suivant :

Proposition 5.5 : (Sous quelques hypothèses techniques) il existe un unique vecteur probabiliste μ tel que pour tout vecteur probabiliste μ_0 ,

$$\lim_{n \rightarrow +\infty} \mu_0 M^n = \mu.$$

μ est alors appelée **loi stationnaire** de la chaîne de Markov.

EXERCICE 1

- Rappeler pourquoi 1 est valeur propre de M .
- Montrer que μ vérifie $\mu M = \mu$.
- En déduire que ${}^t M {}^t \mu = {}^t \mu$, puis que 1 est valeur propre de ${}^t M$, et qu'un vecteur propre associé est ${}^t \mu$.

Sous Sci Lab la commande `spec(M)` permet d'obtenir les valeurs propres d'une matrice M .

Mieux, si on exécute la commande `[P, D] = spec(M)`, alors on obtient pour D une matrice diagonale dont les coefficients sont les valeurs propres de M , et pour P la matrice de passage de la base canonique à une base de vecteurs propres de M (au moins lorsque M est diagonalisable).

EXERCICE 2 Vérifier à l'aide de la commande `spec` que 1 est valeur propre de la matrice de transition M de Doudou, et que $\dim E_1(M) = 1$.

Utiliser cette même commande pour déterminer une base de $E_1({}^t M)$. Comment en déduire un vecteur probabiliste (c'est-à-dire un vecteur dont toutes les composantes sont positives, et dont la somme des composantes vaut 1) qui soit dans $E_1({}^t M)$?

Comparer ce vecteur propre à la loi stationnaire de la chaîne de Markov de Doudou (d'après ce qui précède, on pourra assimiler cette loi stationnaire à $\mu_0 M^i$ où μ_0 est un état initial quelconque et i est grand, de l'ordre de 10000).

Le résultat énoncé précédemment sur la loi stationnaire peut être précisé de la manière suivante :

Proposition 5.6 : (Sous quelques hypothèses) si M est la matrice de transition d'une chaîne de Markov, alors 1 est valeur propre de M . De plus, il existe alors un unique vecteur probabiliste μ tel que $\mu M = \mu$, c'est-à-dire un unique vecteur probabiliste dans $E_1({}^t M)$.

μ est alors le vecteur représentant la loi stationnaire de la chaîne de Markov.

Enfin, quel que soit l'état probabiliste de départ μ_0 , on a $\lim_{n \rightarrow \infty} \mu_0 M^n = \mu$.

EXERCICE 3 Pour l'exemple de la compagnie d'assurance, vérifier à l'aide de Sci Lab que 1 est valeur propre de la matrice de transition (le fichier créant la matrice se trouve sur l'ENT).

Proposer une méthode afin de déterminer approximativement la loi stationnaire de cette chaîne de Markov, et vérifier que l'on obtient ainsi un vecteur μ vérifiant $\mu M = \mu$.

CONVERGENCE EN LOI VERS LA LOI STATIONNAIRE

Précisons un peu comment la loi de X_n converge vers la loi stationnaire lorsque n devient grand.

Notons $\mu = (\mu_1, \dots, \mu_N)$ le vecteur probabiliste de la loi stationnaire. Alors,

$$\forall i \in \llbracket 1, N \rrbracket, \lim_{n \rightarrow +\infty} P(X_n = i) = \mu_i.$$

Cela signifie que la suite (X_n) converge en loi vers la loi stationnaire.

EXERCICE 4 On se place de nouveau dans le cas de la compagnie d'assurance, et on suppose que A est la matrice de transition de la chaîne de Markov associée (téléchargeable sur l'ENT).

On se donne le code suivant

```

1 C = 5*ones(100000,1) ;
2 B = grand(50,'markov',A,C) ;
3 loi_empirique = zeros(9,50)
4 for i=1 :9
5     for j=1 :50
6         loi_empirique(i,j) = length(find(B(:,j)==i))/100000 ;
7     end
8 end

```

- Déterminer la situation modélisée ici, et expliquer à quoi sert la seconde ligne.
- Exécuter ce code. À quoi correspond la matrice B ?
- Que représente la $j^{\text{ème}}$ colonne de la matrice loi_empirique ?
- Que constatez-vous sur les dernières colonnes de la matrice loi_empirique ? Comparer avec le résultat de l'exercice 3.
- On se propose de représenter graphiquement les lois de X_1, X_3, X_5, X_{10} et X_{50} . Pour cela on utilise le code suivant

```

1 l = loi_empirique ;
2 bar((1 :9)', [l(:,1), l(:,3), l(:,5), l(:,10), l(:,40), l(:,50)]))

```

Exécuter ce code, et interpréter le résultat. À quoi correspond chaque couleur ?

UNE CHAÎNE DE MARKOV CÉLÈBRE : LE Google PAGERANK

En 1997, deux étudiants de Stanford, Larry Page et Sergueï Brin ont l'idée d'utiliser les chaînes de Markov pour créer un moteur de recherche. Le problème des moteurs de recherche de l'époque est le suivant : une fois qu'on a associé une liste de mots-clés à chaque page du Web, comment les «ordonner» par ordre d'importance ? Comment choisir, pour un mot clé donné (par exemple «matrice») les pages les plus susceptibles d'intéresser l'internaute ?

À part demander à un humain de trier les pages par ordre d'intérêt (notion bien évidemment subjective), il semble difficile de proposer un classement pertinent de ces pages.

L'idée des fondateurs de Google est la suivante : numérotons l'ensemble des pages d'Internet de 1 à N (N est bien évidemment très très grand, de l'ordre de 10^{13} , mais il est fini).

On considère qu'il existe un nombre $c \in]0, 1[$ tel que, à chaque étape de sa navigation, un internaute :

- suit un des liens de la page sur laquelle il se trouve avec probabilité $1 - c$. Si la page ne contient pas de lien, alors on repart, de manière équiprobable vers n'importe laquelle des N pages.
- ou bien, avec probabilité c , repart vers n'importe laquelle des N pages d'Internet, toutes les pages étant équiprobables.

Donc si X_n est la variable aléatoire égale au numéro de la page sur laquelle se trouve l'internaute à la $n^{\text{ème}}$ étape de sa navigation, on a

$$P_{[X_n=i]}(X_{n+1} = j) = \begin{cases} \frac{1}{N} & \text{si la page } i \text{ ne contient aucun lien} \\ \frac{c}{N} & \text{si la page } i \text{ contient des liens, mais aucun vers la page } j \\ \frac{c}{N} + \frac{1-c}{\ell_i} & \text{sinon, où } \ell_i \text{ est le nombre total de liens que contient la page } i \end{cases}$$

La suite $(X_n)_{n \in \mathbb{N}}$ est alors une chaîne de Markov, et on appelle PageRank (du nom de Larry Page) de la page i le nombre $\lim_{n \rightarrow +\infty} P(X_n = i)$. Le PageRank d'une page est la probabilité, en surfant «au hasard» d'arriver sur cette page. En particulier, d'après ce que nous avons dit précédemment sur les chaînes de Markov, cette probabilité ne dépend pas de la page où commence la navigation.

On remarque notamment que plus une page possède de liens pointant vers elle-même, plus son PageRank est élevé, même si ce n'est pas le seul facteur à prendre en compte. Ainsi, une page avec un PageRank élevé est censée être plus pertinente étant donné qu'elle possède davantage de liens qui pointent vers elle.

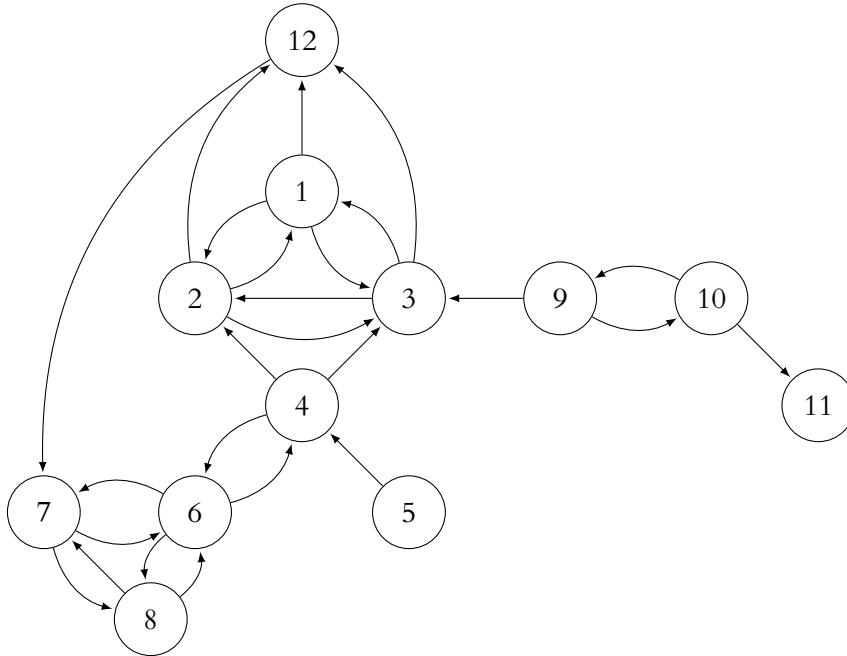
Afin de calculer le PageRank de chacune des pages référencées, les serveurs de Google doivent alors déterminer la loi stationnaire de cette chaîne de Markov, c'est-à-dire calculer un vecteur propre de la matrice de transition (qui, rappelons-le, est une très grande matrice). Nous ne nous attarderons pas sur cet aspect, mais il existe des algorithmes permettant de calculer rapidement un vecteur propre d'une matrice très grande.

EXERCICE 5 Afin d'augmenter le PageRank de mon site, je vais avoir besoin de davantage de liens pointant vers lui. Vaut-il mieux pour moi avoir un lien sur chacun des 36 blogs de mes étudiants, ou un seul lien sur la page d'accueil du site du Monde ?

PD

EXERCICE 6 On considère un modèle simple de Web comportant 12 pages, le schéma ci-dessous représentant les liens entre les pages.

AD



- Essayer de déterminer quelles pages sont susceptibles d'avoir un PageRank plus élevé que les autres, ou au contraire bien plus faible.
- Télécharger le fichier liens.sce qui va créer une matrice B comportant des 0 et des 1, un 1 à la $i^{\text{ème}}$ ligne et $j^{\text{ème}}$ colonne indiquant un lien de la page i vers la page j . Compléter le code suivant pour que, en prenant 0.15 comme valeur de c , il crée, à partir de B , la matrice de transition C de la chaîne de Markov. Puis déterminer le PageRank, c'est-à-dire la loi stationnaire de la chaîne de Markov.

```

1 C = zeros(12,12) ;
2 c = 0.15 ;
3 for i=1 :12
4     l = sum(B(i, :)) ; //nombre de liens sortant de la page i
5     for j = 1 :12
6         if l==0 then
7             C(i,j) = ****
8         else
9             C(i,j) = *****
10    end
11    end
12    end
13
14 [P,D] = spec(C') ;
15 PageRank = *****
16 disp(PageRank)

```

Que constatez-vous ?

- Modifier la matrice B pour supprimer le lien de la page 12 vers la page 7, et recalculer le PageRank. Même question, mais en ajoutant cette fois un lien de la page 12 vers la page 9.

EXERCICE 7 La constante c

Une étude démontre qu'en pratique, un internaute relance sa navigation au hasard au bout de 6 pages en moyenne. Expliquer pourquoi il est alors pertinent de choisir $c = 0.15$.

PD

POUR ALLER PLUS LOIN : CONDITIONS NÉCESSAIRES À L'EXISTENCE D'UNE LOI STATIONNAIRE

Modifions légèrement notre modèle d'assurance, et supposons que la compagnie change les règles pour deux catégories d'assurés : ceux au bonus maximal et ceux au malus maximal. Supposons à présent qu'une fois que l'assuré atteint l'une de ces deux catégories, il y reste à vie.

EXERCICE 8 Va-t-on, comme précédemment, avoir $\forall i \in \llbracket 1, 9 \rrbracket, P(X_n = i)$ qui tend vers une limite indépendante de l'état initial (c'est-à-dire de la loi de X_0) ?

Montrer que la matrice de transition de cette nouvelle chaîne de Markov possède au moins deux vecteurs propres pour la valeur propre 1 qui sont des vecteurs de probabilités. À quels états initiaux correspondent-ils ?

EXERCICE 9 Modifier la matrice de transition de manière à simuler ce nouveau système.

Reprendre le code de l'exercice 4, avec 1000 assurés initialement à l'état 5, puis avec 1000 assurés initialement à l'état 2, et enfin avec 1000 assurés dont les états initiaux sont choisis au hasard dans $\llbracket 1, 9 \rrbracket$.

Avant de tracer un nouveau graphique, on pourra utiliser la commande `scf()` qui permet d'ouvrir une nouvelle fenêtre graphique et donc de comparer les différents graphiques.

Que constatez-vous ?

Les conditions nécessaires à l'existence d'un unique vecteur de probabilité correspondant à une loi stationnaire demandent en gros que depuis chaque état, il soit possible en un temps fini d'atteindre n'importe quel autre état. Ceci assure qu'au bout d'un certain temps, l'influence de l'état initial se sera estompée, et que celui-ci n'aura plus aucune importance.

Notons que c'était bien le cas de la chaîne de Markov de Doudou le hamster : peu importe son activité, on bout de quelques étapes, il peut de nouveau être en train de pratiquer n'importe laquelle de ses trois activités favorites.

C'était également le cas pour la chaîne de Markov du PageRank, grâce à l'hypothèse qu'à tout moment, un internaute peut repartir vers n'importe quelle page.

De même pour le premier modèle d'assurance : peu importe la catégorie dans laquelle se trouve un assuré, au bout de quelques années, il peut se trouver dans n'importe laquelle des catégories (même si le temps nécessaire pour passer d'un malus maximal à un bonus maximal est de minimum 8 ans pour un assuré n'ayant pas d'accidents).

En revanche, notre second modèle d'assurance ne vérifiait pas cette condition : pour un assuré dans une des deux catégories extrêmes, il n'était plus possible de passer dans une autre catégorie. On parle alors d'états **absorbants**, et leur existence ne garantit plus l'existence d'une unique loi stationnaire vers laquelle la loi de X_n finirait par converger.

Une autre raison à l'absence de loi stationnaire est la présence d'états **périodiques**.

Imaginons l'exemple très simple d'une puce se déplaçant sur les six sommets d'un hexagone, numérotés de 1 à 6. À chaque instant, si elle se trouve sur le sommet n ($1 \leq n \leq 5$), elle va au sommet $n + 1$, et si elle se trouve sur le sommet 6, alors elle va au sommet 1.

Cette situation ne laisse pas de place au hasard, mais il s'agit bien d'une chaîne de Markov, dont la matrice de transition ne contient que des zéros et des 1.

EXERCICE 10 Écrire la matrice de transition de cette chaîne de Markov. Vérifier que 1 en est valeur propre, et déterminer un vecteur probabiliste qui soit vecteur propre de la transposée de la matrice de transition pour la valeur propre 1. Donner des états (probabiliste) initiaux pour lesquels la loi de X_n ne converge pas vers ce vecteur propre.

D

EXERCICE 11 Expliquer pourquoi la chaîne de Markov du PageRank ne possède ni états absorbants, ni états périodiques.

AD

CORRECTION DU TP 5

SOLUTION DE L'EXERCICE 1

- La matrice de transition M est une matrice stochastique¹. Et nous avons prouvé dans le DM7 qu'une telle matrice possède nécessairement 1 pour valeur propre, et qu'un vecteur propre associé est $\begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}$.
- On a $\mu M = \lim_{i \rightarrow +\infty} \mu_0 M^i = \lim_{i \rightarrow +\infty} \mu_0 M^{i+1} = \mu$.
- En transposant la relation $\mu M = \mu$, il vient ${}^t M {}^t \mu = {}^t \mu$.
Puisque μ est non nul², 1 est valeur propre de ${}^t M$ et ${}^t \mu$ est un vecteur propre associé. \square

¹ Rappelons que le terme n'est pas à savoir, mais nous avons étudié ces matrices lors du DM7 : ce sont les matrices à coefficients positifs telles que la somme des coefficients de chaque ligne vaille 1.

² Un vecteur probabiliste possède la somme de ses coordonnées égale à 1, donc au moins une d'entre elles est non nulle.

SOLUTION DE L'EXERCICE 2

La commande `[P,D] = spec(M)` retourne

$$D = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0.05 + 0.05i & 0 \\ 0 & 0 & 0.05 - 0.05i \end{pmatrix} \text{ et } P = \begin{pmatrix} 0.5773503 & 0.0724571 + 0.0154417i & 0.0724571 - 0.0154417i \\ 0.5773503 & -0.9241248 & -0.9241248 \\ 0.5773503 & -0.3230873 - 0.1900514i & -0.3230873 + 0.1900514i \end{pmatrix}.$$

En particulier, 1 est bien sur la diagonale de D , donc est une valeur propre de M .

De même, les valeurs propres de ${}^t M$ sont obtenus à l'aide de `[P,D] = spec(M')`, qui retourne alors

$$D = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0.05 + 0.05i & 0 \\ 0 & 0 & 0.05 - 0.05i \end{pmatrix} \text{ et } P = \begin{pmatrix} 0.9954800 & -0.4743416 - 0.1581139i & -0.4743416 + 0.1581139i \\ 0.0497740 & -0.3162278 + 0.1581139i & -0.3162278 - 0.1581139i \\ 0.0808828 & 0.7905694 & 0.7905694 \end{pmatrix}.$$

En particulier, la première colonne de P est un vecteur propre de ${}^t M$ associé à la valeur propre 1.

Ce n'est toutefois pas un vecteur probabiliste puisque la somme de ses coordonnées³ ne vaut pas 1.

Mais alors

```
1 P(:,1)/sum(P(:,1))
```

retourne alors $\begin{pmatrix} 0.8839779 \\ 0.0441989 \\ 0.0718232 \end{pmatrix}$ qui est bien un vecteur probabiliste, et qui est encore un vecteur

propre puisqu'il est colinéaire à celui obtenu précédemment.

Et on remarque alors qu'il s'agit de la loi stationnaire obtenue lors du TP4, et qu'on peut retrouver par exemple à l'aide de

```
1 [1,0,0]*M^1000
```

\square

SOLUTION DE L'EXERCICE 3

On peut procéder comme précédemment : déterminer un vecteur propre à l'aide de la commande `spec` et le diviser par la somme de ses coefficients de manière à ramener cette somme à 1.

```
1 [P,D] = spec(A) ;
2 P(:,1)/sum(P(:,1))
```

\square

SOLUTION DE L'EXERCICE 5

A priori⁴, vos blogs ont un PageRank bien plus faible que la page d'accueil du Monde. Et donc mieux vaut un lien depuis une page au PageRank élevé que de multiples liens depuis de nombreuses pages peu connues.

Une technique pour augmenter mon PageRank serait de créer de nombreuses pages pointant vers mon site, mais si aucune de ces pages n'a de lien depuis une page de plus forte audience, l'impact d'une telle manœuvre sera limitée. \square

³ Remarquons que ces coordonnées sont toutes positives.

Remarque

Notons qu'a priori, rien ne garantit que le vecteur ainsi obtenu est un vecteur probabiliste : il pourrait avoir une (ou plusieurs) coordonnées négatives.

Un résultat (compliqué) prouve qu'en fait il existe bien un vecteur propre de M , associé à la valeur propre 1 et qui est probabiliste. Mieux : tout vecteur de $E_1(A)$ dont la somme des coordonnées vaut 1 est probabiliste, et donc notre méthode fonctionne à tous les coups.

⁴ Sauf si l'un d'entre vous est un blogueur célèbre et que je l'ignore !

SOLUTION DE L'EXERCICE 6

1. La page numéro 5 doit être celle qui possède le PageRank le plus faible, puisqu'on ne peut arriver sur cette page que «par hasard» : elle ne possède aucun lien pointant vers elle. De même, les pages 9, 10 et 11 doivent avoir des PageRank peu élevés, car ne possèdent que des liens de l'une vers l'autre, mais aucun lien venant des autres pages. Donc un internaute ne peut arriver sur l'une de ces pages que «par hasard», puis risque de partir vers d'autres pages (via $9 \rightarrow 3$ ou $10 \rightarrow 11$) et ne peut alors revenir que par hasard. La page au PageRank le plus élevé est probablement la 7, non seulement car elle comporte de nombreux liens pointant vers elle, mais également car ces liens viennent de pages qui ont probablement un PageRank élevé.
- 2.

Détails

La variable l , qui est différente pour chaque ligne, sert à compter le nombre de liens que comporte une page.

```

1 C = zeros(12,12) ;
2 c = 0.15 ;
3 for i=1 :12
4     l = sum(B(i, :)) ;
5     for j = 1 :12
6         if l==0 then
7             C(i,j) = 1/12 ;
8         else
9             C(i,j) = c/12 + B(i,j)*(1-c)/l ;
10    end
11 end
12 end
13
14 [P,D] = spec(C') ;
15 PageRank = P(:,1)/sum(P(:,1))
    
```

□

SOLUTION DE L'EXERCICE 7

Si à chaque instant, l'internaute relance sa navigation avec une probabilité c , le nombre de pages visitées avant de repartir d'une page au hasard suit une loi géométrique de paramètre c .

Et donc en moyenne un internaute visite $\frac{1}{c}$ pages avant de recommencer d'une page au hasard.

Pour avoir $\frac{1}{c} = 6$, on prend donc $c = \frac{1}{6} \approx 0.15$.

□

SOLUTION DE L'EXERCICE 10

La matrice de transition est $M = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$.

Un vecteur propre de tM pour la valeur propre 1 est obtenu à l'aide de la commande `spec(M')`.

On constate alors que toutes ses coordonnées sont les mêmes, et donc $\frac{1}{6} \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}$ est un vecteur

probabiliste qui est un vecteur propre de tM .

Si on prend alors comme état initial $(1 \ 0 \ 0 \ 0 \ 0 \ 0)$, alors il n'y a pas convergence en loi vers la loi représentée par ce vecteur propre.

En effet, si $n = 6k$, alors la loi de X_n est la loi certaine égale à 1, si $n = 6k + 1$, la loi de X_n est la loi certaine égale à 2, etc.

Notons en revanche que si l'état initial est $(\frac{1}{6} \ \frac{1}{6} \ \frac{1}{6} \ \frac{1}{6} \ \frac{1}{6} \ \frac{1}{6})$, alors à tout instant n , X_n suit la loi uniforme sur $[[1, 6]]$, qui converge en loi vers la loi uniforme sur $[[1, 6]]$. □

Remarque

Ce vecteur est celui où la puce possède la même probabilité d'être sur chaque sommet.

Détails

Toutes les 6 itérations, la puce se trouve sur le sommet 1.

TP 6 : SIMULATION DE LOIS À DENSITÉ

Nous nous intéressons dans ce TP à différentes méthodes de simulation de variables à densité.

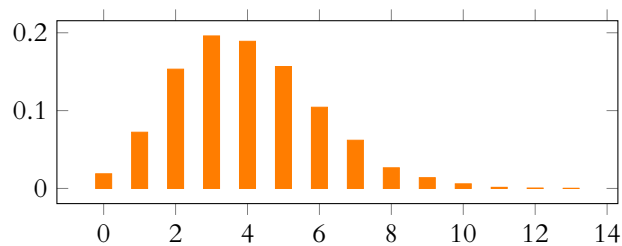
Bien entendu, nous connaissons déjà la commande `grand` de Scilab qui permet de simuler la plupart des lois usuelles. Mais, afin de comprendre au mieux le fonctionnement de Scilab, nous allons essayer de simuler d'autres lois uniquement en utilisant la commande `rand`.

LE POLYGONE DES FRÉQUENCES CUMULÉES, OU LA FONCTION DE RÉPARTITION EMPIRIQUE

Lors de la simulation de variables aléatoires discrètes, afin de comparer nos résultats à la théorie, nous avons pour habitude de regrouper les simulations par modalités, et de tracer un diagramme représentant les effectifs (ou leur fréquence) par modalité.

Par exemple, si Y est un vecteur contenant 10 000 simulations d'une loi de Poisson de paramètre 4, on peut utiliser

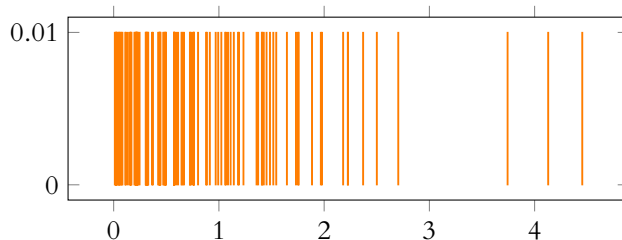
```
1 Y = grand(10000,1,'poi',4);
2 M = tabul(Y,'i');
3 bar(M(:,1),M(:,2)/10000);
```



Pour une variable à densité X , cette méthode n'a aucune chance de fonctionner, puisque pour tout $x \in \mathbf{R}$, $P(X = x) = 0$, et donc chaque modalité de notre échantillon risque d'avoir un effectif égal à 1.

Et donc par exemple pour une loi exponentielle, on obtient :

```
1 Y = grand(100,1,'exp',1);
2 M = tabul(Y,'i');
3 bar(M(:,1),M(:,2)/100);
```



Ce graphique est bien difficile à interpréter, et ne nous permet pas de reconnaître une loi exponentielle.

Nous proposons à la place de tracer le polygone des fréquences cumulées, qui a été défini dans le TP3.

Celui-ci doit être un analogue à la fonction de répartition, puisqu'un point d'abscisse x du polygone a pour ordonnée la fréquence d'apparition des simulations inférieures ou égales à x , qui doit donc être l'analogue de $P(X \leq x)$.

On utilisera pour cela la fonction suivante, qui prend en paramètre un vecteur Y (qui doit donc être un échantillon de la loi qui nous intéresse) et trace le polygone des fréquences cumulées associé.

```
1 fonction polygone(Y)
2     M = tabul(Y,'i');
3     N = length(Y);
4     plot2d(M(:,1),cumsum(M(:,2))/N)
5 endfunction
```

EXERCICE 1 Étudier rapidement le fonctionnement de cette fonction.

L'essayer avec différents échantillons de grande taille (supérieure à 10 000) suivant différentes lois à densité usuelles simulées avec `grand`, et notamment avec une loi exponentielle afin de comparer au graphique donné plus haut. Comment utiliser cette fonction `polygone` pour tracer la fonction Φ ?

Dans la suite, pour vérifier que les variables que nous simulons suivent bien la loi attendue, nous comparerons les graphiques obtenus à l'aide de `polygone` avec ceux des fonctions de répartition attendues.

RAPPEL : TRACER UNE FONCTION f AVEC Scilab

Le plus simple pour tracer une fonction avec Scilab est probablement d'utiliser `fplot2d`.

Rappelons que cela nécessite d'utiliser la syntaxe fonction `y = f(x) ... endfunction` afin de créer la fonction f .

Il faut ensuite choisir l'intervalle $[a, b]$ sur lequel on souhaite représenter cette fonction, et la précision souhaitée pour le dessin, c'est-à-dire le nombre n de «petits segments» en lequel on souhaite diviser $[a, b]$.

Ceci peut se réaliser à l'aide de `x = linspace(a, b, n)`.

Il suffit alors d'utiliser la commande `fplot2d(x, f)`. Par exemple, pour tracer la fonction de répartition d'une loi exponentielle de paramètre 1, on peut utiliser

```
function y = f(x)
    y = 1-exp(-x) ;
endfunction

x = linspace(0, 5, 100) ;
fplot2d(x, f) ;
```

LA MÉTHODE D'INVERSION POUR LES VARIABLES À DENSITÉ

Dans toute la suite, U désigne une variable aléatoire suivant la loi uniforme sur $]0, 1[$.

EXERCICE 2 Rappeler l'expression de la fonction de répartition de U .

EXERCICE 3 Soit X une variable à densité, de fonction de répartition F . On suppose qu'il existe un intervalle I de \mathbf{R} tel que F réalise une bijection de I sur $]0, 1[$.

Exprimer la fonction de répartition de $Y = F^{-1}(U)$ en fonction de F . En déduire que Y a même loi que X .

La méthode d'inversion consiste à exploiter le résultat de l'exercice précédent dans le cas où on connaît une expression de F^{-1} : alors si U est une variable suivant la loi $\mathcal{U}([0, 1])$ (que l'on sait simuler), $F^{-1}(U)$ simule une variable de même loi que X .

EXERCICE 4 Simulation d'une loi exponentielle

- Déterminer la fonction de répartition F_λ d'une loi exponentielle de paramètre λ , et montrer qu'elle réalise une bijection de \mathbf{R}_+^* sur $]0, 1[$. Déterminer sa bijection réciproque.
- Écrire un programme exploitant F_λ^{-1} pour simuler une loi exponentielle de paramètre λ (avec λ de votre choix) uniquement à partir de `rand()`. On pourra à cet effet créer un vecteur X de taille 10 000 (par exemple) contenant 10 000 simulations de cette loi exponentielle.
- En utilisant les commandes `mean` et `st_deviation`, vérifier que la moyenne et l'écart-type empiriques sont bien conformes à ceux qu'on attend.
- Utiliser la fonction `polygone` pour tracer la fonction de répartition empirique correspondant à cet échantillon.
- Toujours à l'aide de `polygone`, tracer la fonction de répartition empirique d'un échantillon de même taille simulé cette fois à l'aide de la commande `grand` (**attention** : `grand` prend en paramètre $1/\lambda$ et non λ !). Comparer les graphiques.

EXERCICE 5 Simulation d'une loi $\gamma(n)$

En utilisant l'exercice précédent, écrire une fonction fonction `y = simule_gamma(n)` qui prend comme paramètre un entier $n \in \mathbf{N}^*$ et retourne une simulation d'une loi $\gamma(n)$.

EXERCICE 6 La loi de Cauchy

Soit f la fonction définie par $f(x) = \frac{1}{\pi} \frac{1}{1+x^2}$.

- Montrer que f est la densité d'une variable aléatoire.
Dans la suite, on considère X une variable aléatoire admettant pour densité la fonction f , on dit que X suit la loi de Cauchy.
- Déterminer la fonction de répartition F_X de X .
- Montrer que $\pi U - \frac{\pi}{2}$ suit une loi uniforme sur $]-\frac{\pi}{2}, \frac{\pi}{2}[$.
- Montrer que $\tan(\pi U - \frac{\pi}{2})$ est une variable aléatoire suivant la loi de Cauchy.
- Montrer que F_X réalise une bijection de \mathbf{R} sur $]0, 1[$, déterminer F_X^{-1} et faire le lien avec la question précédente.
- Utiliser la question 4 pour créer un vecteur de taille 10 000 contenant 10 000 simulations d'une variable aléatoire suivant la loi de Cauchy.
Calculer sa moyenne à l'aide de `mean`, et recommencer sur plusieurs échantillons. Que constatez-vous ? Que dire de l'espérance d'une loi de Cauchy ?

UNE PREMIÈRE MÉTHODE DE SIMULATION D'UNE LOI NORMALE (MÉTHODE DE BOX-MULLER)

EXERCICE 7 On admet que si U et V sont deux variables aléatoires indépendantes suivant la loi $\mathcal{U}([0, 1])$, alors $X = \sqrt{-2 \ln(U)} \cos(2\pi V)$ suit la loi normale centrée réduite. PD

1. Utiliser ce qui a été dit précédemment pour créer un vecteur N contenant 10 000 simulations d'une loi $\mathcal{N}(0, 1)$.
2. En utilisant la fonction `grand`, créer un vecteur G contenant lui aussi 10 000 simulations de la même loi. Tracer sur un même graphique les deux polygones des fréquences cumulées correspondants, et les comparer.

SciLab dispose d'une fonction permettant de calculer la fonction de répartition d'une variable aléatoire suivant une loi normale.

Si $X \hookrightarrow \mathcal{N}(\mu, \sigma^2)$, alors on obtient $F_X(x) = P(X \leq x)$ à l'aide de la commande `cdfnor('PQ', x, mu, sigma)`.

Le fait que le premier paramètre soit 'PQ' vient du fait que si on entre `[p, q] = cdfnor('PQ', x, mu, sigma)`, alors `p` contient $P(X \leq x)$ et `q` contient $P(X > x) = 1 - p$ (pour une raison obscure, SciLab tient à nous retourner la valeur de p et celle de q , alors que nous serions tout à fait capables de nous débrouiller s'il ne nous donnait que l'une de ces deux valeurs...).

EXERCICE 8

1. Soit X une variable aléatoire suivant la loi $\mathcal{N}(3, 4)$. Calculer $P(X \leq 3)$, $P(X \leq 0)$ et $P(X > 10)$.
2. Calculer $\Phi(0)$, $\Phi(1)$ et $\Phi(1.96)$. F

EXERCICE 9

1. En utilisant le résultat de l'exercice 7 et les propriétés des lois normales, proposer une méthode de simulation d'une loi normale $\mathcal{N}(-1, 4)$. PD
2. Écrire une fonction d'entête `function y = rep_normale(mu, sigma, x)` qui, en utilisant `cdfnor`, calcule la valeur de $F_X(x)$, où $X \hookrightarrow \mathcal{N}(\mu, \sigma^2)$. Utiliser cette fonction pour vérifier que la loi simulée à la question précédente est bien une loi $\mathcal{N}(-1, 4)$.
3. Proposer une méthode pour vérifier la stabilité des lois normales.

POUR ALLER PLUS LOIN

EXERCICE 10 Histogrammes et densité

La fonction `histplot` permet de tracer des histogrammes. De manière précise, si $C = [c_1, \dots, c_k]$ est une suite de nombres rangés dans l'ordre croissant, et si $X = [x_1, \dots, x_n]$, alors en notant n_1 le nombre d'éléments de X compris dans $[c_1, c_2]$, n_i le nombre d'éléments de X dans l'intervalle $I_i =]c_i, c_{i+1}]$, la commande `histplot(C, X)` a pour effet de tracer une suite de rectangles de base I_i et d'aire proportionnelle à n_i (de manière précise, SciLab s'arrange pour que la somme des aires des rectangles soit égale à 1).

Créer un vecteur contenant 1000 réalisations d'une variable de loi $\mathcal{N}(0, 1)$ (à l'aide de la méthode de votre choix).

Regrouper cette série statistique en 50 classes, et tracer l'histogramme associé.

Sur le même graphique, tracer la fonction de densité de la loi normale $\mathcal{N}(0, 1)$. Commenter D

EXERCICE 11 La loi de Rayleigh

1. Soit $\sigma > 0$. Montrer que la fonction définie sur \mathbf{R}_+ par $x \mapsto \frac{1}{\sigma^2} x e^{-\frac{x^2}{2\sigma^2}}$ est une densité. D
Une variable aléatoire admettant une telle densité est dite suivre la loi de Rayleigh de paramètre σ .
2. Donner la fonction de répartition d'une variable aléatoire suivant la loi de Rayleigh de paramètre σ .
3. Montrer que cette fonction de répartition réalise une bijection de \mathbf{R}_+^* sur $]0, 1[$.
Utiliser alors la méthode d'inversion pour simuler une variable aléatoire suivant la loi de Rayleigh de paramètre σ .
4. Estimer numériquement l'espérance et la variance d'une variable aléatoire suivant une loi de Rayleigh de paramètre σ . Retrouver ces résultats par le calcul.
5. Si X suit la loi de Rayleigh de paramètre σ , quelle est la loi de $Y = X^2$?

EXERCICE 12 Différentes simulations de la loi de Pareto

Soit $k \in \mathbf{N}^*$ et $\lambda > 0$. D

1. Déterminer la valeur de r pour laquelle $f_\lambda : x \mapsto \begin{cases} 0 & \text{si } x \leq \lambda \\ \frac{r}{x^{k+1}} & \text{sinon} \end{cases}$ est une densité de probabilité.

Si X admet pour densité f_λ , on dit que X suit la loi de Pareto de paramètres λ et k .

2. Déterminer la fonction de répartition d'une variable suivant la loi de Pareto de paramètres λ et k .
3. En utilisant la méthode d'inversion, simuler une variable aléatoire simulant une loi de Pareto de paramètres λ et k .

4. Soient X_1, \dots, X_k k variables aléatoires i.i.d. suivant la loi uniforme sur $[0, 1]$. On pose alors $Y = \frac{\lambda}{\max(X_1, \dots, X_k)}$.
- 4.a. Montrer que Y suit une loi de Pareto de paramètres λ et k .
- 4.b. En déduire une autre méthode pour simuler la loi de Pareto.

CORRECTION DU TP 6

SOLUTION DE L'EXERCICE 2

Est-ce bien nécessaire de l'écrire ? Allez, au cas où...

$$F_X(x) = \begin{cases} 0 & \text{si } x \leq 0 \\ x & \text{si } 0 < x < 1 \\ 1 & \text{si } x \geq 1 \end{cases}$$

□

SOLUTION DE L'EXERCICE 3

Supposer que F réalise une bijection de I sur $]0, 1[$ revient à supposer que X est à valeurs dans I .

De plus, $Y = F^{-1}(U)$ est alors à valeurs dans I .

On a, pour $x \in I$,

$$P(Y \leq x) = P(F^{-1}(U) \leq x) = P(U \leq F(x)) = F(x).$$

Et pour $x \notin I$, on a soit $P(Y \leq x) = 0$ si x est plus petit que tout élément de I , soit $P(Y \leq x) = 1$ si x est plus grand que tout élément de I .

Et alors $P(X \leq x)$ prend les mêmes valeurs puisque X est à valeurs dans I .

Ainsi, X et Y ont même fonction de répartition : elles sont égales !

□

Remarque

Une fonction de répartition est toujours croissante !

SOLUTION DE L'EXERCICE 4

1. Nous savons que $F_\lambda(x) = \begin{cases} 0 & \text{si } x \leq 0 \\ 1 - e^{-\lambda x} & \text{si } x > 0 \end{cases}$.

Cette fonction est continue et strictement croissante sur \mathbf{R}_+^* , avec $\lim_{x \rightarrow 0} F_\lambda(x) = 0$ et

$$\lim_{x \rightarrow +\infty} F_\lambda(x) = 1.$$

Par le théorème de la bijection, F_λ réalise donc une bijection de \mathbf{R}_+^* sur $]0, 1[$.

Pour $y \in]0, 1[$, on a

$$F_\lambda(x) = y \Leftrightarrow 1 - e^{-\lambda x} = y \Leftrightarrow e^{-\lambda x} = 1 - y \Leftrightarrow x = -\frac{1}{\lambda} \ln(1 - y).$$

$$\text{Et donc } F_\lambda^{-1}(y) = -\frac{1}{\lambda} \ln(1 - y).$$

2. Si on utilise le résultat de la question 5, $F_\lambda^{-1}(U)$ suit une loi exponentielle $\mathcal{E}(\lambda)$.

```
1 X = zeros(10000,1);
2 for i=1 :10 000
3     X(i) = -1/lambda*log(1-rand());
4 end
```

3. En prenant par exemple $\lambda = 2$ dans la question précédente, on obtient bien moyenne de X proche de 0.5 et un écart-type également proche de 0.5.

Ceci est cohérent avec le fait que l'espérance d'une loi exponentielle $\mathcal{E}(2)$ vaut 1/2 et sa variance vaut 1/4 (et donc son écart-type vaut 1/2).

4 et 5. On simule ce deuxième échantillon à l'aide de

- 1 $Y = \text{grand}(10000, 1, 'exp', 1/\lambda)$;

On constate, sans surprise, que les deux graphiques sont proches, confirmant¹ que nous avons bien simulé une loi exponentielle.

□

¹ Mais avons-nous vraiment besoin de cette confirmation après les calculs de l'exercice 5 ?

SOLUTION DE L'EXERCICE 5

Il suffit de se rappeler que, par stabilité des lois γ , la somme de n variables aléatoires indépendantes suivant la loi $\mathcal{E}(1) = \gamma(1)$ est une variable suivant la loi $\gamma(n)$.

Et donc, en utilisant l'exercice précédent pour simuler des lois $\mathcal{E}(1)$:

```

1 fonction y = simule_gamma(n)
2     y = 0
3     for i = 1 : n
4         y = y - log(rand())
5     end
6 endfunction

```

□

SOLUTION DE L'EXERCICE 6

1. La fonction f est continue et positive sur \mathbf{R} . Pour $A > 0$, on a

$$\int_0^A f(x) dx = \frac{1}{\pi} \int_0^A \frac{1}{1+x^2} dx = \frac{1}{\pi} [\text{Arctan}(x)]_0^A = \frac{1}{\pi} \text{Arctan}(A) \xrightarrow{A \rightarrow +\infty} \frac{1}{\pi} \frac{\pi}{2} = \frac{1}{2}.$$

Puisque de plus f est paire, on en déduit que $\int_{-\infty}^{+\infty} f(x) dx$ est convergente et vaut $2 \frac{1}{2} = 1$.
Donc f est bien une densité de probabilités.

2. Pour $x \in \mathbf{R}$, on a

$$F_X(x) = \int_{-\infty}^x f(t) dt = \int_{-\infty}^0 f(t) dt + \int_0^x f(t) dt = \frac{1}{2} + \frac{1}{\pi} [\text{Arctan}(t)]_0^x = \frac{1}{2} + \frac{1}{\pi} \text{Arctan}(x).$$

3. C'est du cours : il s'agit d'une transformation affine de loi uniforme.
4. La fonction de répartition de $\pi U - \frac{\pi}{2}$ est

$$F : x \mapsto \begin{cases} 0 & \text{si } x \leq -\frac{\pi}{2} \\ \frac{x + \frac{\pi}{2}}{\pi} & \text{si } x - \frac{\pi}{2} < x < \frac{\pi}{2} \\ 1 & \text{sinon} \end{cases}$$

5. Pour $x \in \mathbf{R}$, on a

$$\begin{aligned} P\left(\tan\left(\pi U - \frac{\pi}{2}\right) \leq x\right) &= P\left(\pi U - \frac{\pi}{2} \leq \text{Arctan}(x)\right) \\ &= \frac{1}{\pi} \text{Arctan}(x) + \frac{1}{2} = F_X(x). \end{aligned}$$

Et donc, $\tan\left(U - \frac{\pi}{2}\right)$ suit bien une loi de Cauchy.

6. La fonction Arctan est continue et strictement croissante sur \mathbf{R} , donc il en est de même de F_X .

Puisqu'on a bien $\lim_{x \rightarrow -\infty} F_X(x) = 0$ et $\lim_{x \rightarrow +\infty} F_X(x) = 1$, par le théorème de la bijection, F_X réalise une bijection de \mathbf{R} dans $]0, 1[$.

On a alors, pour $y \in]0, 1[$,

$$F_X(x) = y \Leftrightarrow \frac{1}{\pi} \text{Arctan}(x) + \frac{1}{2} = y \Leftrightarrow \text{Arctan}(x) = \pi y - \frac{\pi}{2} \Leftrightarrow \tan\left(\pi y - \frac{\pi}{2}\right).$$

En particulier, en utilisant le résultat de l'exercice 5, on retrouve que $F_X^{-1} = \tan\left(\pi U - \frac{\pi}{2}\right)$ suit une loi de Cauchy.

□

SOLUTION DE L'EXERCICE 7

1. Puisque nous savons simuler des lois uniformes avec $\text{rand}()$, il s'agit d'utiliser la définition de X :

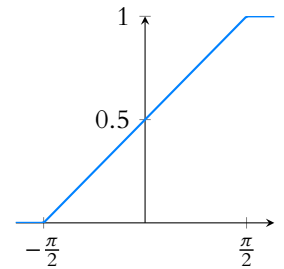
```

1 N = zeros(1, 10000);
2 for i = 1 : 10000
3     N(i) = sqrt(-2*log(rand()))*cos(2*%pi*rand());
4 end

```

Les moyennes et variances de l'échantillon ainsi obtenues s'obtiennent respectivement à l'aide des commandes $\text{mean}(N)$ et $\text{stdev}(N)$. On constate qu'elles valent² environ 0 et 1.

² Comme attendu.



La fonction Arctan est croissante.

Remarque

Il n'y a pas besoin de vérifier ces limites : ce sont celles de toute fonction de répartition.

2. Rappelons que pour tracer deux courbes sur le même graphique, il suffit de ne pas fermer le graphique entre les deux commandes traçant les courbes. Donc ici on peut se contenter de

```
1 G = grand(1,10000,'nor',0,1);
2 polygone(N)
3 polygone(G)
```

On constate alors que les courbes se superposent presque, ce qui laisse bien penser que nous avons simulé la bonne loi. □

SOLUTION DE L'EXERCICE 8

1. Il s'agit d'utiliser $[p, q] = \text{cdfnor}('PQ', x, 3, 2)$, où x vaut successivement 3, 0 et 10. Dans les deux premiers cas, c'est la valeur de p qui nous intéresse, et dans le dernier cas, c'est celle de q .
On obtient alors $P(X \leq 3) = 0.5$, $P(X \leq 0) \simeq 0.06$ et $P(X > 10) \simeq 0.00023$.
2. Pour calculer la valeur de $\Phi(x)$, il faut utiliser $[p, q] = \text{cdfnor}('PQ', x, 0, 1)$, et on obtient alors

$$\Phi(0) = 0.5, \Phi(1) = 0.84, \Phi(1.96) = 0.975.$$

□

□

SOLUTION DE L'EXERCICE 12

1. Il est clair que f est continue et positive.
De plus, on a

$$\int_{-\infty}^{+\infty} f_{\lambda}(x) dx = \int_{\lambda}^{+\infty} \frac{r}{x^{k+1}} dx = \frac{r}{k\lambda^k}.$$

Et donc cette intégrale vaut 1 si et seulement si $r = k\lambda^k$.

2. Pour $x \leq \lambda$, on a $P(X \leq x) = 0$, et si $x > \lambda$, alors

$$P(X \leq x) = \int_{-\infty}^x f(t) dt = \int_{\lambda}^x f(t) dt = 1 - \left(\frac{\lambda}{x}\right)^k.$$

3. F_X réalise une bijection croissante de $]\lambda, +\infty[$ sur $]0, 1[$. Pour trouver sa bijection réciproque, fixons $y \in]0, 1[$. On a alors

$$F_X(x) = y \Leftrightarrow 1 - \left(\frac{\lambda}{x}\right)^k = y \Leftrightarrow \left(\frac{\lambda}{x}\right)^k = 1 - y \Leftrightarrow x = \frac{\lambda}{(1-y)^{1/k}}.$$

Et donc $F_X^{-1}(y) = \frac{\lambda}{(1-y)^{1/k}}$.

Il s'agit donc d'appliquer la méthode d'inversion expliquée précédemment pour simuler une loi de Pareto :

```
1 lambda/((1- rand())^(1/k))
```

4.a.

□

Danger !

On pensera bien à entrer l'écart type (ici 2) et non la variance.

Remarque

L'une de ces valeurs était prévisible, savez-vous laquelle ?

TP 7 : STATISTIQUES BIVARIÉES

Les données statistiques ne vont pas toujours toutes seules, et pour un même individu, il est possible de s'intéresser à plusieurs caractères. Par exemple, pour une même personne on peut s'intéresser à son poids et sa taille, ou à ses revenus et à son niveau d'études.

Nous nous limiterons ici à l'étude simultanée de deux caractères, et nous intéresserons à la question suivante : *peut-on exprimer l'un en fonction de l'autre ? Plus précisément : l'un est-il une fonction affine de l'autre ?*

REPRÉSENTATION GRAPHIQUE

Nous considérons un échantillon de taille n , pour lequel nous étudions deux caractères, que l'on suppose contenus dans deux vecteurs X et Y de taille n , tels que $X(i)$ et $Y(i)$ correspondent au même individu.

On représente alors la série statistique à l'aide d'un **nuage de points**. C'est l'ensemble des n points de coordonnées $(X(i), Y(i))$. La question est alors de savoir si ces points sont distribués ou non sur une même droite.

EXERCICE 1 Étudions le lien entre la consommation de viande (en grammes par jour et par personne) et l'incidence du cancer du colon (pour 100 000 femmes et par an) dans les pays industrialisés. On se donne les données suivantes :

PD

Pays	Consommation de viande	Taux de cancer
Japon	26	7.5
Israël	124	16.4
États-Unis	284	34
Finlande	101	9.8
Grande-Bretagne	205	23.3

1. Entrer ces données sous la forme de deux vecteurs X et Y .
2. Représenter le nuage de points associé à l'aide de la commande `plot2d` (afin de ne pas relier les points entre eux, on utilisera `plot2d(X, Y, -1)`).
3. Les points sont-ils tous situés sur une même droite ? Si non, peut-on trouver une droite qui passe très près de tous les points ? Proposer l'équation d'une telle droite (approximativement).

EXERCICE 2 Créer deux vecteurs X et Y contenant chacun 1000 nombres tirés au hasard suivant une loi uniforme sur $[0, 1]$, puis représenter le nuage de points associé. Pouvez-vous proposer une droite qui s'approche «au mieux» du nuage de points ?

F

Dans toute la suite, on appelle **point moyen** du nuage le point de coordonnées (\bar{x}, \bar{y}) , avec

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i \text{ et } \bar{y} = \frac{1}{n} \sum_{i=1}^n y_i.$$

Ce point moyen ne signifie pas grand chose pris tout seul, essayez par exemple d'imaginer l'allure d'un nuage de points dont tous les points seraient très loin du point moyen.

LA MÉTHODE DES MOINDRES CARRÉS

Il y a très peu de chance que tous nos points soient sur une même droite, mais nous allons chercher «la meilleure droite» approximant cet ensemble de points.

Précisons un peu : si nos points sont $(x_1, y_1), \dots, (x_n, y_n)$, nous allons chercher une droite d'équation $y = ax + b$ qui minimise la «somme des carrés des écarts à la droite».

De manière plus précise, nous cherchons à savoir s'il existe une droite qui rend minimale

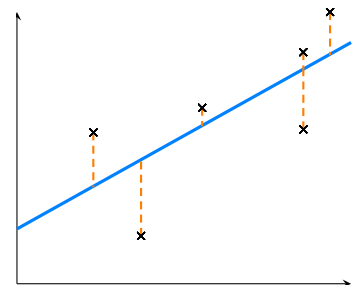
$$\sum_{i=1}^n (y_i - (ax_i + b))^2.$$

Une telle droite quand elle existe, est appelée **droite des moindres carrés**.

Il s'agit donc d'une droite minimisant la somme des carrés des longueurs des segments pointillés tracés sur la figure ci-contre.

Notons que plutôt que de chercher à minimiser la somme des carrés, on pourrait vouloir minimiser la somme des longueurs, ou encore minimiser la plus grande des longueurs.

Une des raisons pour lesquelles on s'intéresse aux carrés est qu'on dispose alors d'un résultat garantissant l'existence et l'unicité d'une telle droite.



AD

EXERCICE 3 Existence de la droite des moindres carrés

Considérons la matrice $A = \begin{pmatrix} x_1 & 1 \\ \vdots & \vdots \\ x_n & 1 \end{pmatrix} \in \mathcal{M}_{n,2}(\mathbf{R})$ et $B = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \in \mathcal{M}_{n,1}(\mathbf{R})$.

1. Pourquoi n'est-il pas trop restrictif de supposer que A est de rang 2 ?
2. Pour $X = \begin{pmatrix} a \\ b \end{pmatrix} \in \mathcal{M}_{2,1}(\mathbf{R})$, exprimer $\|AX - B\|^2$ en fonction des x_i et des y_i .
3. En utilisant le théorème des pseudo-solutions, montrer qu'il existe un unique $X_0 = \begin{pmatrix} a_0 \\ b_0 \end{pmatrix} \in \mathcal{M}_{2,1}(\mathbf{R})$ qui minimise l'ensemble $\{\|AX - B\|, X \in \mathcal{M}_{2,1}(\mathbf{R})\}$.
4. En déduire qu'il existe une unique droite des moindres carrés.

On admet qu'avec les notations précédentes, on a $\begin{pmatrix} a_0 \\ b_0 \end{pmatrix} = ({}^tAA)^{-1}{}^tAB$ (voir exercice 10).

EXERCICE 4 Reprendre les données de l'exercice 1 et appliquer la méthode qui précède afin d'obtenir les valeurs de a_0 et b_0 . Représenter cette droite sur le même graphique que le nuage de points. Le résultat est-il conforme à vos attentes ?

La commande `[a,b] = reglin(X,Y)` permet d'obtenir a_0 et b_0 sans passer par la matrice A . Cela nécessite absolument d'avoir X et Y des **vecteurs lignes**, si on a des vecteurs colonnes, on commencera par les transposer.

EXERCICE 5 Télécharger sur l'ENT le fichier PIB_naissances.sce, et ouvrez-le dans Scilab .

Cela a pour effet de créer une grande matrice M , à deux colonnes.

On lit dans ces deux colonnes, pour l'année 2014, le PIB par habitant en dollars et le nombre de naissances pour 1000 habitants (données issues du CIA World Factbook 2015).

1. Parcourir rapidement cette matrice, en chercher les valeurs extrêmes, etc. Pouvez-vous supposer qu'il existe un lien entre le PIB par habitant et le taux de natalité ?
2. Tracer le nuage de points associé, puis, sur le même graphique, la droite des moindres carrés. Est-il raisonnable de supposer que le taux de natalité est une fonction affine du PIB ?
3. Quel est le signe du coefficient directeur de cette droite ? Comment l'expliquez-vous ?

PD

COVARIANCE, COEFFICIENT DE CORRÉLATION

Si (x_1, \dots, x_n) est une série statistique, on définit la **variance** de X par $V(X) = \sigma_X^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2$.

Rappelons que σ_X s'obtient en Scilab par la commande `stdev(X)`.

On définit alors la covariance (empirique) de X et Y par $\text{Cov}(X, Y) = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})$.

$\text{Cov}(X, Y)$ se calcule en Scilab avec la commande `corr(X, Y, 1)`.

Tout comme la covariance de deux variables aléatoires, il n'est pas évident de donner une interprétation de la covariance seule. Son signe a toutefois une interprétation : s'il est positif, c'est que x_i « a tendance » à augmenter lorsque y_i augmente, alors que s'il est négatif, x_i diminue lorsque y_i augmente.

On définit alors le coefficient de corrélation linéaire de X et Y par $\rho(X, Y) = \frac{\text{Cov}(X, Y)}{\sigma_X \sigma_Y}$.

Il est possible de prouver qu'on obtient les mêmes résultats que pour le coefficient de corrélation linéaire de deux variables aléatoires :

- $|\rho(X, Y)| \leq 1$
- $\rho(X, Y) = \pm 1$ si et seulement si il existe a et b tels que $Y = aX + b$. Dans ce cas, le signe de a est le signe de ρ .

Ainsi, un coefficient de corrélation linéaire proche de ± 1 indique que la droite des moindres carrés approche bien le nuage de points, alors qu'une corrélation linéaire proche de 0 indique une absence de relation de dépendance linéaire entre X et Y .

En SciLab, on peut calculer le coefficient de corrélation linéaire «à la main», en utilisant $\text{corr}(X, Y, 1) / (\text{stdev}(X) * \text{stdev}(Y))$ ou directement à l'aide de la commande $\text{correl}(X, Y)$.

EXERCICE 6

1. Calculer les coefficients de corrélation linéaire pour les deux séries statistiques dont nous disposons. Êtes-vous surpris ?
2. Reprendre deux vecteurs X et Y comme dans l'exercice 2, et calculer le coefficient de corrélation linéaire. Comment expliquer le résultat obtenu ?

F

SENSIBILITÉ AUX VALEURS EXTRÊMES

EXERCICE 7 Revenons sur l'étude de la seconde série statistique (PIB/taux de natalité).

Commençons par trier ces données par PIB croissant à l'aide de la commande $M = \text{lex_sort}(M)$.

On constate qu'il existe quelques valeurs extrêmes, assez éloignées des autres. En particulier du côté des pays au PIB/habitant élevé (les cinq plus élevés sont le Qatar, le Liechtenstein, Macao, les Bermudes et Monaco : un producteur de pétrole et quatre paradis fiscaux, ça alors !)

On décide de limiter notre étude aux pays dont le PIB/habitant est inférieur à 60 000 dollars, ce qui revient à exclure 7 pays.

1. Créer une matrice reprenant les données de M mais uniquement avec les PIB/habitant inférieurs à 60 000 dollars. Si on souhaite ne garder d'une matrice M que les lignes numérotées entre 1 et 10, on peut utiliser la commande $N = M(1 : 10, :)$.
2. Calculer le coefficient de corrélation linéaire de cette nouvelle série.

PD

EXERCICE 8 Refaire le même exercice avec les données du fichier `PIB_religion.sce` qui donne, pour 40 pays, le PIB/habitant ainsi que le pourcentage d'habitants de ce pays déclarant que la religion a une importance.

Le fichier `PIB_religion_sans_extremes.sce` reprend les mêmes données, mais on a enlevé deux pays qui étaient très éloignés de la droite des moindres carrés (la Chine et les USA). Calculer le nouveau coefficient de corrélation linéaire.

Un des inconvénients de la méthode des moindres carrés est qu'elle est très sensible aux valeurs extrêmes : une seule valeur très éloignée de la droite des moindres carrés a une «grosse» influence sur le coefficient de corrélation linéaire (et également sur la position de la droite des moindres carrés).

F

CORRÉLATION NE VEUT PAS DIRE CAUSALITÉ !

Une bonne corrélation entre deux séries de données ne signifie pas pour autant qu'il existe un lien de cause à effet entre les deux. À titre d'exemple, considérons les séries statistiques suivantes (données authentiques !)

Ce tableau donne la consommation de margarine en livre/personne aux USA, et la taux de divorce dans l'état du Maine (en divorce pour 1000 personnes).

Année	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009
Consommation de margarine	8.2	7	6.5	5.3	5.2	4	4.6	4.5	4.2	3.7
Taux de divorce	5	4.7	4.6	4.4	4.3	4.1	4.2	4.2	4.2	4.1

Le tableau suivant donne le nombre de morts (pour un million d'habitants) sur les autoroutes américaines, ainsi que le nombre de tonnes de citrons mexicains importés aux États-Unis de 1996 à 2000.

Année	1996	1997	1998	1999	2000
Morts	15.85	15.7	15.39	15.32	14.85
Importations de citrons	230	280	360	410	525

EXERCICE 9 Calculer les coefficients de corrélation pour ces deux séries. En déduisez vous une information pertinente ?

F

POUR ALLER PLUS LOIN : UN PEU DE THÉORIE

EXERCICE 10 Pour faire le lien avec les probabilités

1. Montrer que $\sigma_X^2 = \left(\frac{1}{n} \sum_{i=1}^n x_i^2 \right) - (\bar{x})^2$.

D

2. Montrer que $\text{Cov}(X, Y) = \overline{xy} - \bar{x}\bar{y}$, où \overline{xy} désigne la moyenne de (x_1y_1, \dots, x_ny_n) .
3. Calculer tAA , puis son inverse, et enfin $({}^tAA)^{-1}{}^tAB = \begin{pmatrix} a_0 \\ b_0 \end{pmatrix}$.
4. En déduire que $a_0 = \frac{\text{Cov}(X, Y)}{\sigma_X^2}$. Déterminer de même une formule pour b_0 .
5. Montrer que la droite des moindres carrés passe par le point moyen (\bar{x}, \bar{y}) du nuage de points.

EXERCICE 11 Avec des produits scalaires

1. Montrer à l'aide de l'inégalité de Cauchy-Schwarz que $|\rho(X, Y)| \leq 1$.
2. Montrer que $\rho(X, Y) = \pm 1$ si et seulement si les (x_i, y_i) sont tous sur une même droite.
3. Expliquer en quoi la recherche de la droite des moindres carrés fait intervenir le projeté orthogonal de Y sur $F = \text{Vect}(X, 1_n)$, où 1_n désigne le vecteur $(1, \dots, 1)$.
4. En utilisant alors le fait que $Y - (a_0X + b_01_n) \in F^\perp$, retrouver le fait que le point moyen (\bar{x}, \bar{y}) du nuage de points appartient à la droite des moindres carrés.

EXERCICE 12 Retour sur le théorème des pseudo-solutions

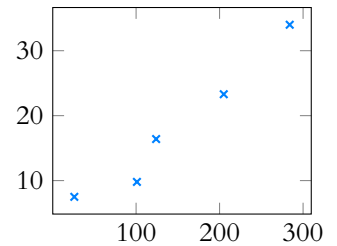
Le but de cet exercice est de montrer que l'unique X_0 minimisant $\|AX - B\|$ vérifie en fait ${}^tAAX_0 = {}^tAB$.
 On reprend les notations du cours, et on note $f : \mathcal{M}_{p,1}(\mathbf{R}) \rightarrow \mathcal{M}_{n,1}(\mathbf{R})$ l'application définie par $f(X) = AX$.
 Nous savons que $Y_0 = AX_0$ est le projeté orthogonal de B sur $\text{Im}(f)$.

1. Montrer que $Y_0 - B \in \text{Im}(f)^\perp$.
2. En déduire que $\forall X \in \mathcal{M}_{p,1}(\mathbf{R})$, on a ${}^tX({}^tAY_0 - {}^tAB) = 0$.
3. Conclure.

CORRECTION DU TP 7

SOLUTION DE L'EXERCICE 1

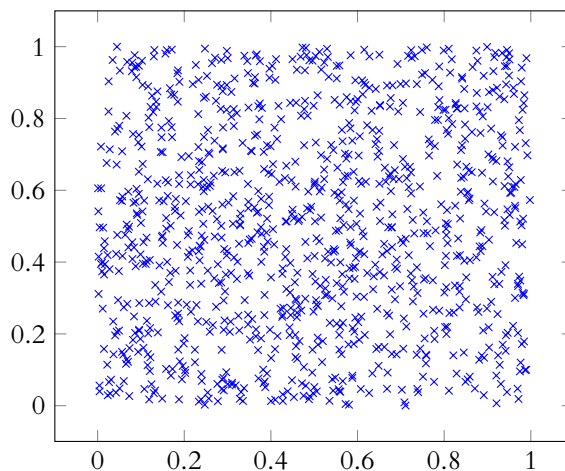
- À l'aide de la commande indiquée dans l'énoncé, on obtient la figure suivante :
- Les points ne sont pas totalement alignés, mais le sont presque, à l'exception peut-être du point correspondant à la Finlande.
On peut par exemple prendre comme droite la droite reliant le premier au dernier point, qui est alors de coefficient directeur $\Delta = \frac{34 - 7.5}{284 - 26} \approx 0.1$.
Puisqu'elle passe par le point de coordonnées (26, 7.5), il s'agit donc de la droite d'équation $y = 0.1x + 4.9$.



□

SOLUTION DE L'EXERCICE 2

À l'aide de la commande `plot2d(grand(1,1000,'unf',0,1),grand(1,1000,'unf',0,1),-1)`, on obtient un dessin semblable à



À première vue, aucune droite ne semble mieux approximer ces points que les autres. □

SOLUTION DE L'EXERCICE 3

- A possède deux colonnes, donc est de rang 1 ou 2 (elle ne peut être de rang 0 car la seconde colonne est non nulle).
Si elle était de rang 1, les deux colonnes seraient colinéaires, et donc cela signifierait que les x_i sont tous égaux.
Et donc les points (x_i, y_i) seraient tous situés sur la droite d'équation $x = x_1$.

- On a $AX = \begin{pmatrix} ax_1 + b \\ ax_2 + b \\ \vdots \\ ax_n + b \end{pmatrix}$ et donc $\|AX - B\|^2 = \sum_{i=1}^n (ax_i + b - y_i)^2$.

- Le théorème des pseudo-solutions s'applique car A est de rang 2, et donc il existe un unique X_0 tel que $\|AX_0 - B\|$ soit minimal.
- La droite d'équation $y = a_0x + b_0$ est donc l'unique droite des moindres carrés.

□

SOLUTION DE L'EXERCICE 4

Il s'agit de créer les matrices A et B , puis d'appliquer la formule obtenue précédemment.

```

1 A = [26, 1 ; 124, 1 ; 284, 1 ; 101, 1 ; 205, 1] ;
2 B = [7.5 ; 16.4 ; 34 ; 9.8 ; 23.3] ;
3 C = inv(A'*A)*A'*B ;
4 a = C(1, 1) ;
5 b = C(2, 1) ;
6 plot2d(A(:, 1), B, -1)
7 plot2d([0, 300], [b, 300*a+b])

```

En fait

Nous verrons qu'une telle droite existe nécessairement. Mais il est douteux que son usage soit pertinent...

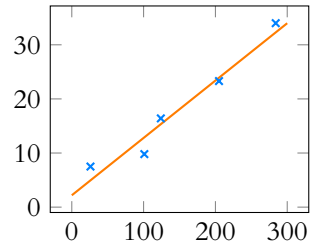
Pertinence

De plus, il serait assez peu pertinent de se demander si les y_i dépendent des x_i , puisque ceux-ci sont tous égaux. Imaginez que vous vous intéressiez à un éventuel lien entre l'âge et les revenus, et que vous décidiez de faire votre étude sur un échantillon de personnes ayant toutes 30 ans. Espérez-vous tirer une information d'un tel échantillon ?

Notons que pour tracer une droite, il suffit d'en placer deux points, SciLab se chargeant alors de les relier.

C'est ce que nous avons fait dans la dernière ligne en plaçant les deux points d'abscisses 0 et 300 sur la droite d'équation $y = ax + b$.

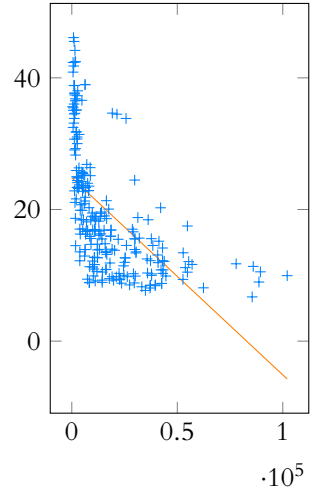
□



SOLUTION DE L'EXERCICE 5

- On peut par exemple utiliser $\min(M(:,1))$ et $\max(M(:,1))$ pour trouver les valeurs extrêmes du PIB par habitant, et procéder de même sur la seconde colonne pour trouver celles du taux de natalité.
On trouve alors des PIB/habitant compris entre 400 et 102100 et des taux de natalité compris entre 6.72 et 46.12.
- On procède comme à l'exercice précédent. Beaucoup de points sont éloignés de la droite des moindres carrés, et donc il semble difficile d'affirmer que le taux de naissance soit une fonction affine du PIB.
- On constate que le coefficient directeur est négatif, ce qui est cohérent avec l'intuition : même si le taux de natalité n'est pas une fonction affine du PIB, il a tendance à diminuer lorsque le PIB augmente.

□



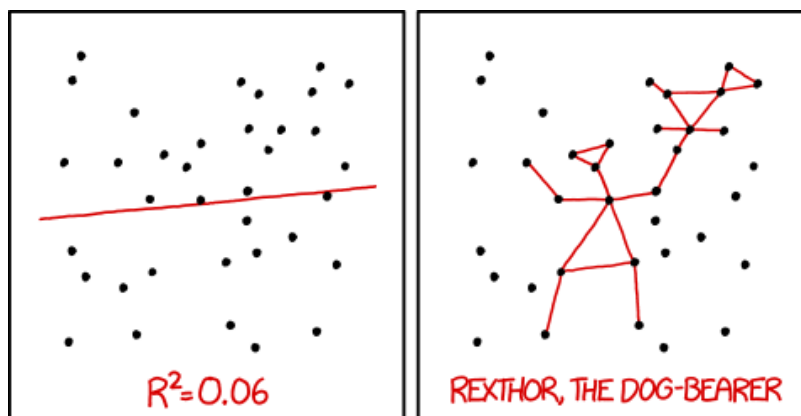
SOLUTION DE L'EXERCICE 6

- Le coefficient de corrélation de la première série (cancer du colon/consommation de viande) est de 0.97, ce qui correspond au fait que les points étaient presque alignés. En revanche, celui de la seconde est de -0.57 . Son signe correspond bien à ce qui a été constaté à l'exercice précédent, et le fait qu'il soit proche de 0.5, c'est-à-dire ni trop proche de 0 ni trop proche de 1 semble indiquer un lien entre PIB et taux de natalité, mais révèle qu'il y a d'autres facteurs à prendre en compte.
- Pour deux vecteurs pris au hasard¹ comme dans l'exercice 2, on obtient une corrélation proche de 0.
Ce résultat est à mettre en parallèle avec le fait que si X et Y sont deux variables aléatoires indépendantes, alors $\text{Cov}(X, Y) = 0$ et donc $\sigma_{X, Y} = 0$.

¹ Et donc ces deux vecteurs sont indépendants !

Notons enfin que même dans le cas du nuage de points de l'exercice 2, il existe une unique droite des moindres carrés.

Il est très dur, voire impossible de se figurer la direction de cette droite juste en regardant le nuage de points. Et le fait qu'elle « colle mal » au nuage de points est traduit par le fait que le coefficient de corrélation est très proche de 0.



I DON'T TRUST LINEAR REGRESSIONS WHEN IT'S HARDER TO GUESS THE DIRECTION OF THE CORRELATION FROM THE SCATTER PLOT THAN TO FIND NEW CONSTELLATIONS ON IT.

□

SOLUTION DE L'EXERCICE 7

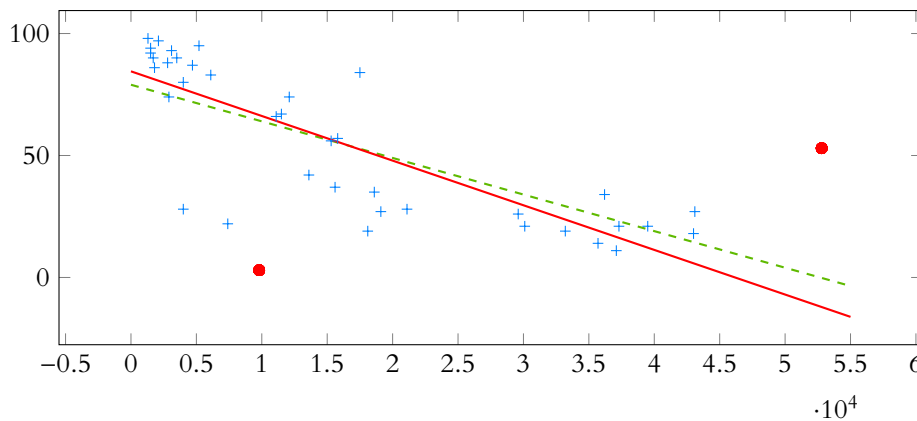
1. Une fois que M est triée par PIB, il suffit d'enlever les 7 dernières lignes, ce qui se fait aisément via $N = M(1 : 227, :)$
2. On trouve alors un coefficient de corrélation égal à -0.61 . L'écart n'est pas énorme avec l'ancien coefficient de corrélation linéaire, mais il faut garder à l'esprit qu'on n'a enlevé que 7 valeurs sur 234.

□

SOLUTION DE L'EXERCICE 8

On constate que le coefficient de corrélation de la première série vaut -0.7 , alors que celui de la seconde vaut -0.81 .

L'écart est significatif, surtout si on pense qu'on n'a enlevé que deux valeurs. Mais les deux points en question étaient les plus éloignés de la droite des moindres carrés, et donc avaient une influence notable sur la position de cette droite.



□

En rouge, les deux points qu'on a enlevé dans la deuxième série. La droite pointillée est la droite des moindres carrés avant suppression de ces deux pays. La droite pleine est la droite des moindres carrés pour les 38 pays restants. On constate que celle-ci «colle» davantage aux points que la droite pointillée. Ainsi, les deux points éloignés de cette droite avaient tendance à éloigner la droite des 38 autres points.

SOLUTION DE L'EXERCICE 9

La première corrélation vaut 0.99 et la seconde vaut -0.99 .

Pourtant il semble bien difficile² d'envisager un lien rationnel entre le nombre de morts sur les routes et les importations de citrons mexicains.

Le travail du mathématicien s'arrête donc au calcul de la corrélation, l'exploitation de ces chiffres nécessite d'autres compétences !

□

² Même si les adeptes de théorie du complot y arriveront probablement !

SOLUTION DE L'EXERCICE 10

1. On a

$$\begin{aligned}\sigma_X^2 &= \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2 = \frac{1}{n} \sum_{i=1}^n x_i^2 - \frac{2}{n} \sum_{i=1}^n x_i \bar{x} + \frac{1}{n} \sum_{i=1}^n \bar{x}^2 \\ &= \frac{1}{n} \sum_{i=1}^n x_i^2 - \frac{2}{n} \bar{x} \sum_{i=1}^n x_i + \bar{x}^2 = \frac{1}{n} \sum_{i=1}^n x_i^2 - 2\bar{x}^2 + \bar{x}^2 \\ &= \frac{1}{n} \sum_{i=1}^n x_i^2 - \bar{x}^2\end{aligned}$$

2. De même, il vient

$$\begin{aligned}\text{Cov}(X, Y) &= \frac{1}{n} \left(\sum_{i=1}^n x_i y_i - \frac{\bar{x}}{n} \sum_{i=1}^n y_i - \frac{\bar{y}}{n} \sum_{i=1}^n x_i + n\bar{x}\bar{y} \right) \\ &= \bar{x}\bar{y} - \bar{x}\bar{y} - \bar{x}\bar{y} + \bar{x}\bar{y} \\ &= \bar{x}\bar{y} - \bar{x}\bar{y}\end{aligned}$$

3. On a

$${}^t AA = \begin{pmatrix} x_1 & x_2 & \dots & x_n \\ 1 & 1 & \dots & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \\ 1 \end{pmatrix} = \begin{pmatrix} \sum_{i=1}^n x_i^2 & \sum_{i=1}^n x_i \\ \sum_{i=1}^n x_i & n \end{pmatrix} = n \begin{pmatrix} \sigma_X^2 + \bar{x}^2 & \bar{x} \\ \bar{x} & 1 \end{pmatrix}.$$

On constate que le déterminant de cette matrice est $\sigma_X^2 \neq 0$, et donc elle est inversible. De plus son inverse est

$$({}^tAA)^{-1} = \frac{1}{n\sigma_X^2} \begin{pmatrix} 1 & -\bar{x} \\ -\bar{x} & \sigma_X^2 + \bar{x}^2 \end{pmatrix}.$$

Puisque ${}^tAB = \begin{pmatrix} \sum_{i=1}^n x_i y_i \\ \sum_{i=1}^n y_i \end{pmatrix} = n \begin{pmatrix} \bar{x}\bar{y} \\ \bar{y} \end{pmatrix}$, il vient enfin

$$({}^tAA)^{-1}{}^tAB = \frac{1}{\sigma_X^2} \begin{pmatrix} 1 & -\bar{x} \\ -\bar{x} & \sigma_X^2 + \bar{x}^2 \end{pmatrix} \begin{pmatrix} \bar{x}\bar{y} \\ \bar{y} \end{pmatrix} = \frac{1}{\sigma_X^2} \begin{pmatrix} \bar{x}\bar{y} - \bar{x}\bar{y} \\ -\bar{x}\bar{x}\bar{y} + \sigma_X^2\bar{y} + \bar{x}^2\bar{y} \end{pmatrix}.$$

4. On en déduit notamment que $a_0 = \frac{\text{Cov}(X, Y)}{\sigma_X^2}$, et que

$$b_0 = \frac{-\bar{x}\bar{x}\bar{y} + \sigma_X^2\bar{y} + \bar{x}^2\bar{y}}{\sigma_X^2}.$$

5. Il s'agit de vérifier que $\bar{y} = a_0\bar{x} + b_0$. Mais on a

$$a_0\bar{x} + b_0 = \frac{\text{Cov}(X, Y)}{\sigma_X^2} \bar{x} + \frac{-\bar{x}\bar{x}\bar{y} + \sigma_X^2\bar{y} + \bar{x}^2\bar{y}}{\sigma_X^2} = \frac{\bar{x}(\bar{x}\bar{y} + \bar{x}\bar{y}) - \bar{x}\bar{x}\bar{y} + \sigma_X^2\bar{y} + \bar{x}^2\bar{y}}{\sigma_X^2} = \bar{y} \frac{\sigma_X^2}{\sigma_X^2} = \bar{y}.$$

□

SOLUTION DE L'EXERCICE 11

1. Appliquons l'inégalité de Cauchy-Schwarz dans \mathbf{R}^n aux vecteurs $(x_1 - \bar{x}, \dots, x_n - \bar{x})$ et $(y_1 - \bar{y}, \dots, y_n - \bar{y})$. Il vient alors

$$\left| \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y}) \right| \leq \sqrt{\sum_{i=1}^n (x_i - \bar{x})^2} \sqrt{\sum_{i=1}^n (y_i - \bar{y})^2}.$$

Après division par n , cette inégalité devient

$$|\text{Cov}(X, Y)| \leq \sigma_X \sigma_Y$$

et donc on a bien

$$|\rho(X, Y)| = \frac{|\text{Cov}(X, Y)|}{\sigma_X \sigma_Y} \leq 1.$$

2. Nous savons que dans Cauchy-Schwarz, il y a une égalité si et seulement si les deux vecteurs sont colinéaires. C'est-à-dire si et seulement si il existe un $\lambda \in \mathbf{R}$, tel que pour tout $i \in \llbracket 1, n \rrbracket$, $y_i - \bar{y} = \lambda(x_i - \bar{x})$.

Soit encore si et seulement si tous les points (x_i, y_i) sont sur la même droite d'équation $y = \lambda x + \bar{y} - \lambda \bar{x}$.

3. La preuve du théorème des pseudo-solutions indique que le minimum X_0 est tel que AX_0 est le projeté orthogonal de B sur $\text{Im } f_A = \{AU, U \in \mathcal{M}_{2,1}(\mathbf{R})\}$.

Mais $\text{Im } f_A$ est de dimension 2 car A est de rang 2, et engendré par $A \begin{pmatrix} 1 \\ 0 \end{pmatrix} = X$ et $A \begin{pmatrix} 0 \\ 1 \end{pmatrix} = 1_n$.

Autrement dit, $\text{Im } f_A = \text{Vect}(X, 1_n)$ et donc AX_0 est le projeté orthogonal de B sur $\text{Vect}(X, 1_n)$.

4. $Y_0 = AX_0 = A \begin{pmatrix} a_0 \\ b_0 \end{pmatrix} = a_0 A \begin{pmatrix} 1 \\ 0 \end{pmatrix} + b_0 A \begin{pmatrix} 0 \\ 1 \end{pmatrix} = a_0 X + b_0 1_n$.

Par propriété des projecteurs orthogonaux, on a alors $Y - (a_0 X + b_0 1_n) \in F^\perp$. En particulier,

$$\langle 1_n, Y - (a_0 X + b_0 1_n) \rangle = 0 \Leftrightarrow \sum_{i=1}^n y_i - (a_0 x_i + b_0) = 0.$$

En divisant cette relation par n , il vient

$$\frac{1}{n} \sum_{i=1}^n y_i - a_0 \frac{1}{n} \sum_{i=1}^n x_i - b_0 = 0 \Leftrightarrow \bar{y} = a_0 \bar{x} + b_0.$$

Et ainsi, le point moyen (\bar{x}, \bar{y}) est sur la droite des moindres carrés.

Pour une matrice 2×2 , on a une formule : l'inverse de $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ est $\frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$.

□

SOLUTION DE L'EXERCICE 12

1. C'est une propriété des projecteurs orthogonaux.
2. Pour tout $Y \in \text{Im } f$, on a $\langle Y, Y_0 - B \rangle = 0$. Mais pour $X \in \mathcal{M}_{p,1}(\mathbf{R})$, on a $AX \in \text{Im } f$, donc

$$\langle AX, Y_0 - B \rangle = 0 \Leftrightarrow t(AX)(Y_0 - B) = 0 \Leftrightarrow {}^tX({}^tAY_0 - {}^tAB) = 0.$$

3. Pour tout $X \in \mathcal{M}_{p,1}(\mathbf{R})$, on a

$${}^t({}^tAY_0 - {}^tAB)X = 0.$$

Notons que ${}^t({}^tAY_0 - {}^tAB)$ est un vecteur ligne à p colonnes.

Si son produit avec tout vecteur colonne est nul, cela signifie que l'application linéaire de \mathbf{R}^n dans \mathbf{R} canoniquement associée est nulle, et donc c'est la matrice nulle.

On en déduit que ${}^tAY_0 - {}^tAB = 0$, et donc que ${}^tAY_0 = {}^tAB$, soit encore

$${}^tAAX_0 = {}^tAB.$$

On peut alors montrer que tAA est inversible (nous l'avons fait précédemment dans le cas particulier où A est la matrice de l'exercice 3) et donc que

$$X_0 = ({}^tAA)^{-1}{}^tAB.$$

□

TP 8 : RÉVISIONS

ÉCHAUFFEMENT : fonction ... endfunction

L'indice de masse corporelle (IMC) d'une personne est défini par $I = \frac{P}{t^2}$ où P est le poids en kilogrammes et t la taille en mètres.

EXERCICE 1

1. Écrire une fonction d'entête `function y = IMC(P, t)` qui prend comme paramètres le poids et la taille, et retourne la valeur de l'IMC.
2. On considère que la taille d'une personne (en mètres) et son poids (en kg) sont deux variables aléatoires indépendantes, suivant respectivement des lois normales $\mathcal{N}(1.78, 0.0016)$ et $\mathcal{N}(72, 36)$.
Créer deux vecteurs `P` et `t` de taille 10 000 contenant respectivement les simulations du poids et de la taille de 10 000 personnes. *On prendra garde au fait que grand prend comme paramètres d'une loi normale son espérance et son écart-type.*
3. En utilisant la fonction `IMC` et les vecteurs `P` et `t`, donner la valeur moyenne de l'IMC pour notre échantillon de 10 000 personnes.
4. Une personne est considérée en surpoids lorsque son IMC dépasse 25. Déterminer la proportion de personnes en surpoids.

F

DES BOUCLES, DES BOUCLES ET ENCORE DES BOUCLES

EXERCICE 2 Soit (u_n) la suite définie par
$$\begin{cases} u_0 = 2 \\ u_{n+1} = \sin(u_n) \end{cases}$$

Écrire une fonction qui prend comme paramètre un entier n et calcule la valeur de u_n correspondante. *On pourra vérifier que pour $n = 100$, on obtient $u_n \approx 0.16926$.*

PD

EXERCICE 3 Même question que pour l'exercice précédent pour la suite définie par
$$\begin{cases} u_0 = 1 \\ u_{n+1} = u_n + \frac{(-1)^n}{n+1} \end{cases}$$

On vérifiera qu'on obtient bien $u_{10} \approx 1.645$.

PD

EXERCICE 4 Écrire un programme qui prend comme paramètre un entier n et retourne $u_n = \sum_{k=1}^n \frac{1}{k} - \ln(n)$.

Vérifier alors numériquement que la suite (u_n) est convergente, et donner une valeur approchée de sa limite.

F

EXERCICE 5 On considère à présent la suite définie par $u_1 = 2$ et pour tout $n \geq 1$, $u_{n+1} = \frac{\sqrt{2}u_n}{\sqrt{1 + \sqrt{1 - (\frac{u_n}{2^n})^2}}}$.

D

1. Comme à l'exercice précédent, écrire un programme qui calcule u_n . *On vérifiera notamment que pour $n = 1$, le programme retourne bien 2, et que pour $n = 2$, il retourne $2\sqrt{2}$.*
2. Essayer ce programme pour de grandes valeurs de n . Reconnaissez-vous la limite de cette suite ?
3. On admet que pour $n \in \mathbf{N}^*$, $|u_n - \pi| \leq \frac{\pi^3}{6 \times 4^n}$.

Sans utiliser la constante Scilab `% pi`, écrire un programme qui prend comme paramètre un entier p et retourne une valeur approchée de π à 10^{-p} près.

EXERCICE 6

1. Soit (u_n) la suite (appelée suite de Fibonacci) définie par $u_0 = u_1 = 1$ et pour tout $n \in \mathbf{N}$, $u_{n+2} = u_{n+1} + u_n$. Compléter la fonction suivante pour qu'elle prenne en paramètre un entier n et retourne la valeur de u_n .

```

1 function u = fibonacci(n)
2     v = 1 ;
3     u = 1 ;
4     for i=2 :n
5         a =
6         u =
7         v = a
8     end
9 endfunction

```

AD

On vérifiera notamment que $u_{10} = 89$.

2. Soient (a_n) et (b_n) les suites définies par

$$a_0 = 2, b_0 = 3, \text{ et } \forall n \in \mathbf{N}, a_{n+1} = \sqrt{a_n b_n}, b_{n+1} = \frac{a_n + b_n}{2}.$$

Écrire une fonction d'entête fonction `[a,b] = suites(n)` prenant comme paramètre un entier n et retournant les valeurs de a_n et b_n correspondantes.

Constater que ces suites convergent vers une même limite.

EXERCICE 7 Triangle de Pascal

On rappelle la formule de Pascal : $\binom{n-1}{k-1} + \binom{n-1}{k} = \binom{n}{k}$.

On souhaite exploiter cette formule pour créer une matrice nommée `triangle`, de taille 21×21 contenant les 21 premières lignes du triangle de Pascal, c'est-à-dire les coefficients binomiaux $\binom{n}{k}$, pour $n \in \llbracket 0, 20 \rrbracket$. Puisque les lignes et les colonnes d'une matrice sont numérotées à partir de 1 et non de 0, on souhaite que le coefficient situé à la $i^{\text{ème}}$ ligne et $j^{\text{ème}}$ colonne de la matrice `triangle` soit égal à $\binom{i-1}{j-1}$.

- Créer une matrice de taille 21×21 ne contenant que des 0, à l'exception des coefficients de la première colonne, tous égaux à 1.
- En utilisant la formule de Pascal, et sans calculer de factorielles, écrire un programme qui remplit les coefficients de la matrice `triangle`, et constater que les résultats sont bien ceux attendus.

Donner la valeur de $\binom{20}{16}$.

EXERCICE 8 Il est bien connu que la série de terme général $\frac{1}{n}$ diverge, et donc que $\sum_{k=1}^n \frac{1}{k} \xrightarrow{n \rightarrow +\infty} +\infty$.

Écrire une fonction qui prend comme paramètre un réel $A > 0$ et qui retourne la valeur du plus petit entier n tel que $\sum_{k=1}^n \frac{1}{k} > A$.

On vérifiera que pour $A = 10$, alors $n = 12\,367$.

Quel est le plus petit entier n tel que $\sum_{k=1}^n \frac{1}{k} > 15$?

EXERCICE 9 Soit (u_n) la suite définie par $\begin{cases} u_0 = 1 \\ u_{n+1} = e^{-u_n} \end{cases}$

Il n'est pas très difficile de prouver que (u_n) converge, que sa limite ℓ est l'unique solution de l'équation $\ell = e^{-\ell}$ et que pour tout $k \in \mathbf{N}$, $u_{2k+1} \leq \ell \leq u_{2k+2} \leq u_{2k}$.

- Montrer que pour tout n , ℓ est dans l'intervalle d'extrémités u_n et u_{n+1} .
- Écrire une fonction `SciLab` qui prend comme paramètre un réel $\epsilon > 0$, et qui retourne le plus petit entier n tel que $|u_n - u_{n+1}| \leq \epsilon$.
On vérifiera que pour $\epsilon = 0.001$, on obtient $n = 12$.
- En déduire une fonction qui, étant donné un entier $p \in \mathbf{N}^*$, retourne une valeur approchée de ℓ à 10^{-p} près.
On vérifiera que pour $p = 5$, on obtient $\ell \simeq 0.5671477$.

EXERCICE 10 Une conjecture célèbre

On appelle suite de Syracuse toute suite $(u_n)_{n \in \mathbf{N}}$ dont le premier terme est dans \mathbf{N}^* et vérifiant la relation de récurrence

$$\forall n \in \mathbf{N}, u_{n+1} = \begin{cases} \frac{u_n}{2} & \text{si } n \text{ est pair} \\ 3u_n + 1 & \text{sinon} \end{cases}$$

On conjecture que toute suite de Syracuse, indépendamment de la valeur de son premier terme, finit toujours par prendre la valeur 1, puis les valeurs 1, 4, 2, 1, 4, 2, 1, 4, 2, 1, ...

- Essayer à la main de vérifier cette conjecture pour de petites valeurs de u_0 .
- Soit $k \in \mathbf{N}^*$ et soit $(u_n)_{n \in \mathbf{N}}$ la suite de Syracuse telle que $u_0 = k$.
On appelle « temps de vol » d'un entier k le plus petit entier n tel que $u_n = 1$.
Écrire une fonction `vol` $n = \text{vol}(k)$ qui calcule le temps de vol d'un entier k . On vérifiera notamment que le temps de vol de 127 est 46.



- Utiliser la fonction précédente pour créer un graphique représentant en abscisse le premier terme k ($1 \leq k \leq 150$) et en ordonnée le temps de vol associé.
- On appelle «altitude» d'un entier k la valeur maximale atteinte par la suite de Syracuse de premier terme k .
Écrire une fonction `function [t,a] = syracuse(k)` qui prend en paramètre un entier k et retourne le temps de vol et l'altitude de k .
- Utiliser la fonction précédente pour tracer un graphique représentant l'altitude en fonction de k .

UN PEU DE PROBABILITÉS.

EXERCICE 11 Si (X_1, \dots, X_n) sont des variables aléatoires indépendantes suivant la loi $\mathcal{N}(0, 1)$, la loi suivie par la variable aléatoire $\sum_{i=1}^n X_i^2$ est appelée *loi du χ^2* (prononcer «chi-deux») de paramètre n .

- Écrire une fonction `function y = chi_deux(n)` qui prend en paramètre un entier $n \geq 1$, et qui simule une variable suivant la loi $\chi^2(n)$.
- Écrire une fonction qui simule la variable T_p , où $T_p = \max(Y_1, \dots, Y_p)$, où Y_1, \dots, Y_p sont p variables aléatoires indépendantes suivant la loi $\chi^2(n)$.
- Proposer une méthode pour obtenir une valeur approchée de $E(T_p)$.

EXERCICE 12 En utilisant `rand()`, mais sans `grand()`, écrire une fonction d'entête `function y = binomiale(n,p)` qui simule une réalisation d'une loi $\mathcal{B}(n, p)$. *Indication : on se rappellera qu'une loi binomiale correspond au nombre de succès lors d'une répétition d'épreuves de Bernoulli indépendantes.*

EXERCICE 13 Une urne contient initialement des boules numérotées de 2 à n .

On effectue un tirage dans cette urne, et on enlève de l'urne toutes les boules portant un numéro supérieur ou égal à celui de la boule tirée.

On ajoute alors la boule numéro 1 dans l'urne, et on effectue un nouveau tirage, et on note X le numéro de la boule obtenue.

Écrire un programme qui simule la variable aléatoire X .

Question subsidiaire : savez-vous le faire sans `grand()` ?

EXERCICE 14 On lance n fois une pièce équilibrée, et on note X le nombre de fois où l'on obtient deux résultats identiques consécutifs. Écrire un programme qui simule la variable aléatoire X .

Utiliser ensuite ce programme pour estimer la valeur de $E(X)$.

UN SUJET DE CONCOURS (ÉCRICOME 2016)

EXERCICE 15 Soient a et b deux entiers strictement positifs. Une urne contient initialement a boules rouges et b boules blanches. On effectue une succession d'épreuves, chaque épreuve étant constituée des trois étapes suivantes :

- on pioche une boule au hasard dans l'urne,
- on replace la boule tirée dans l'urne,
- on rajoute dans l'urne une boule de la même couleur que celle qui vient d'être piochée.

Après n épreuves, l'urne contient donc $a + b + n$ boules.

Pour tout $n \in \mathbf{N}^*$, on note X_n le nombre de boules rouges qui ont été ajoutées dans l'urne (par rapport à la composition initiale) à l'issue des n premières épreuves.

Pour tout $n \in \mathbf{N}^*$, on notera R_n l'événement «on pioche une boule rouge au n -ième tirage».

- Compléter la fonction suivante, qui simule le tirage d'une boule dans une urne contenant x boules rouges et y boules blanches et qui retourne la valeur 0 si la boule est rouge et 1 si elle est blanche.

```

1 function res = tirage(x,y)
2     r = rand()
3     if ..... then
4         res = 0
5     else
6         res=1
7     end
8 endfunction

```

- Compléter la fonction suivante, qui simule n tirages successifs dans une urne contenant initialement a boules rouges et b boules blanches (selon le protocole décrit ci-dessus) et qui retourne la valeur de X_n :

```
1 function Xn = experience(a,b,n)
2     x = a
3     y=b
4     for k=1 :n
5         r = tirage(x,y)
6         if r ==0 then
7             x = .....
8         else
9             .....
10        end
11    end
12    Xn = .....
13 endfunction
```

3. Écrire une fonction Scilab d'en tête :

```
1 function loi = simulation(a,b,n,m)
```

qui fait appel m fois à la fonction précédente pour estimer la loi de X_n . Le paramètre de sortie sera un vecteur contenant les approximations de $P(X_n = 0), P(X_n = 1), \dots, P(X_n = n)$.

CORRECTION DU TP 8

SOLUTION DE L'EXERCICE 1

1. Clarifions bien les choses : lorsqu'on n'utilise la syntaxe fonction ... endfunction, on n'utilise pas **input** ni **disp**.

En effet, si on a besoin de donner une valeur au programme, on en fera un paramètre de la fonction, comme c'est le cas avec la fonction IMC, qui attend deux paramètres P et t .

Si on souhaite calculer l'IMC d'une personne de 1m80 pour 72kg, on utilisera directement la commande `IMC(72, 1.8)`.

Et l'IMC en question sera alors la valeur retournée par la fonction (qui par défaut s'affiche si on a tapé la commande dans la console).

Ceci étant dit, la fonction est alors triviale à écrire :

```
1 function y = IMC(P,t)
2     y = P/t^2
3 endfunction
```

2. Les commandes `t = grand(1,10000, 'nor', 1.78, 0.04)` et `P = grand(1,10000, 'nor', 72, 6)` conviennent.

3. Il suffit, pour chacun des individus de notre échantillon, de calculer son IMC, puis de calculer la moyenne de ces IMC.

Notons que pour cela, nous n'avons pas besoin de conserver individuellement la valeur de chaque IMC, et qu'il nous suffit de produire une variable s qui contient la somme des IMC, puis de la diviser par 10 000.

```
1 s = 0 ;
2 for i=1 :10000
3     s = s+IMC(P(i),t(i)) ;
4 end
5 disp(s/10000)
```

4. On peut, entre autres solutions, procéder de manière similaire à la question précédente : calculer successivement les IMC, et lorsque celui-ci dépasse 25, ajouter 1 à une variable s .

```
1 s = 0 ;
2 for i=1 :10000
3     if IMC(P(i),t(i))>25 then
4         s = s+1 ;
5     end
6 end
7 disp(s/10000)
```

On obtient alors environ 14% de personnes en surpoids.

□

SOLUTION DE L'EXERCICE 2

C'est l'exemple typique de question où il faut utiliser une boucle `for` : on utilise une variable u , qui au départ vaut 2(= u_0), puis qui va prendre successivement les valeurs u_1, u_2, \dots, u_n .

Autrement dit, n fois on remplace u par $\sin(u)$.

```
1 function u = suite(n)
2     u=2
3     for i=1 :n
4         u = sin(u)
5     end
6 endfunction
```

□

Remarque

Mais a-t-on vraiment besoin d'afficher cette valeur ? Voir la question 4.

disp

Notons que si on avait utilisé à tort un `disp` dans notre fonction IMC, alors les 10000 appels à la fonction IMC afficheraient les valeurs correspondantes des IMC, ce qui ne nous intéresse pas du tout !

for

Notons que si i varie de 1 à n , le programme n'utilise pas du tout la valeur de i .

C'est bien normal, puisque nous répétons n fois la même opération (appliquer le sinus à u).

Autrement dit, la transformation réalisée sur u ne dépend pas de i , c'est la même à chaque étape, et elle n'a donc aucune raison de dépendre de la valeur de i .

SOLUTION DE L'EXERCICE 3

C'est globalement la même chose qu'à l'exercice précédent, sauf que la transformation réalisée dépend de la valeur de i .

Plus précisément : avant le $i^{\text{ème}}$ passage dans la boucle, u contient u_{i-1} .

Et donc les instructions situées dans la boucle `for` vont avoir pour effet de calculer u_i .

Or la relation de récurrence s'écrit encore $u_i = u_{i-1} + \frac{(-1)^{i-1}}{i}$.

```

1 fonction u = suite(n)
2     u = 1 ;
3     for i=1 :n
4         u = u + (-1)^(i-1)/i ;
5     end
6 endfunction

```

□

Danger !

Il faut faire très attention aux indices sur ce genre de question, et ne pas confondre n et $n + 1$ (ou i et $i + 1$).

SOLUTION DE L'EXERCICE 4

On peut par exemple commencer par calculer la somme à l'aide d'une boucle, puis seulement à la fin retirer $\ln(n)$.

```

1 fonction u = suite(n)
2     u = 0 ;
3     for i=1 :n
4         u = u+1/i ;
5     end
6     u = u - log(n) ;
7 endfunction

```

On constate que pour de grandes valeurs de n , on obtient toujours la même valeur, proche de 0.577.

□

SOLUTION DE L'EXERCICE 5

1. Au delà du fait que la formule donnée soit compliquée, il faut être vigilant, comme dans l'exercice 3 à un éventuel décalage d'indice.

```

1 fonction u = suite(n)
2     u = 2
3     for i=2 :n
4         u = sqrt(2)*u/sqrt(1+sqrt(1-(u/2^(i-1))^2))
5     end
6 endfunction

```

2. Pour $n = 100$, on obtient 3.1415927, qui ressemble curieusement à π . Il semble donc raisonnable de conjecturer que cette suite converge vers π .
3. Nous cherchons donc un entier n tel que $|u_n - \pi| \leq 10^{-p}$.

Et pour cela, le plus simple est de considérer n tel que $\frac{\pi^3}{6 \times 4^n} \leq 10^{-p}$.

Puisque nous n'avons pas le droit d'utiliser la commande `%pi` rusons un peu, et contentons nous de majorer π^3 par 4^3 .

Et donc

$$\frac{4^3}{6 \times 4^n} \leq 10^{-p} \Leftrightarrow 4^{n-3} \geq \frac{10^p}{6} \Leftrightarrow n \geq 3 + \frac{p \ln(10) - \ln(6)}{\ln(4)}.$$

```

1 fonction y = approx_pi(p)
2     n = floor(3+(p*log(10)-log(6))/log(4))+1 ;
3     y = suite(n)
4 endfunction

```

□

Interdiction

Pourquoi interdire cette commande ? Tout simplement car il s'agit d'un programme qui vise à calculer une valeur approchée de π . Si on s'autorise la commande `%pi`, alors ne nous embêtons pas : utilisons-la pour obtenir une valeur approchée de π !

SOLUTION DE L'EXERCICE 6

1. On ne peut pas tout à fait procéder comme dans les exercices 2 et 3, car à chaque étape, il faut garder en mémoire les deux dernières valeurs de u_n .
Autrement dit, on peut utiliser deux variables u et v telles qu'à tout instant, u contienne la valeur de u_n et v contienne la valeur de u_{n-1} .
Alors nous pourrions facilement calculer la valeur de u_{n+1} : c'est celle de $u+v$.
Et il faudra alors changer la valeur de v pour qu'elle contienne u_n , qui était dans u .
Une erreur à ne pas faire serait de se contenter de

1 $u = u+v$; $v=u$;

En effet, ce programme commence par remplacer u par $u+v$, et donc on «perd» la valeur de u_n qui était dans u .
Pour pallier à ce problème, nous proposons d'utiliser une troisième variable a qui stockera temporairement la valeur de u_n .

```

1  function u = fibonacci(n)
2      v = 1 ;
3      u = 1 ;
4      for i=2 :n
5          a = u
6          u = u+v
7          v = a
8      end
9  endfunction

```

2. Le principe est le même :

```

1  function [a,b] = suites(n)
2      a = 2 ;
3      b = 3 ;
4      for i=1 :n
5          c=a ;
6          a = sqrt(a*b) ;
7          b = (c+b)/2 ;
8      end
9  endfunction

```

□

SOLUTION DE L'EXERCICE 7

1. Nous savons facilement créer une matrice formée de 0 à l'aide de la commande `zeros`.
Pour y mettre des 1 sur la première colonne, on peut utiliser une boucle qui parcourt les coefficients de la première colonne :

```

1  triangle = zeros(21,21) ;
2  for i = 1 : 21
3      triangle(i,1) = 1 ;
4  end

```

2. Il s'agit de remplir successivement les lignes de la matrice `triangle` en utilisant la formule de Pascal.

On utilisera pour cela deux boucles : l'une parcourant les différentes lignes, et au sein cette première boucle, une seconde remplissant successivement les différentes colonnes de chaque ligne.

```

1  for i = 2 :21
2      for j =2 :i
3          triangle(i,j) = triangle(i-1,j-1) + triangle(i-1,j) ;
4      end
5  end

```

Le coefficient $\binom{20}{16}$ est alors obtenu à l'aide de `triangle(21,17)`, et il vaut 4845

□

SOLUTION DE L'EXERCICE 8

Notons qu'ici, un ordinateur est indispensable, et qu'on ne sait pas résoudre «à la main»

l'inéquation $\sum_{k=1}^n \frac{1}{k} > A$.

Il s'agit donc de calculer successivement les valeurs de $\sum_{k=1}^n \frac{1}{k}$ jusqu'à ce que l'un des termes dépasse A .

C'est l'exemple type de cas où l'on utilise une boucle `while`.

```

1 fonction n = serie_harmonique(A)
2     S = 1 ;
3     n=1 ;
4     while S<A
5         n = n+1 ;
6         S = S+1/n ;
7     end
8 endfunction

```

En particulier, pour $A = 15$, on obtient $n = 1\ 835\ 421$.

□

SOLUTION DE L'EXERCICE 9

1. L'inégalité $u_{2k+1} \leq \ell \leq u_{2k+2} \leq u_{2k}$ nous dit que pour $n = 2k$ pair, $u_{n+1} \leq \ell \leq u_n$, et pour $n = 2k + 1$ impair, $u_n = u_{2k+1} \leq \ell \leq u_{2k+2} = u_{n+1}$.
Donc dans les deux cas, ℓ est bien entre u_n et u_{n+1} .
2. Il s'agit de calculer successivement les termes de la suite, jusqu'à avoir $|u_n - u_{n+1}| < \varepsilon$.
On augmente alors à chaque fois de 1 une variable n qui stocke

```

1 fonction n = intervalle(epsilon)
2     u = 1
3     n = 0
4     while abs(u-exp(-u))>epsilon
5         u = exp(-u) ;
6         n = n+1 ;
7     end
8 endfunction

```

3. Nous allons adapter un peu le programme précédent : si la longueur de l'intervalle d'extrémités u_n et u_{n+1} est inférieure ou égale à 10^{-p} , alors la distance entre ℓ (qui se trouve dans cet intervalle) et u_n est nécessairement elle aussi inférieure ou égale à 10^{-p} .
Et donc u_n est une approximation de ℓ à 10^{-p} près.

```

1 fonction u = approximation(p)
2     u = 1 ;
3     while abs(u - exp(-u))>10^(-p)
4         u = exp(-u) ;
5     end
6 endfunction

```

□

SOLUTION DE L'EXERCICE 10

1. Essayons par exemple pour $u_0 = 14$. Alors on obtient 14, 7, 22, 11, 34, 17, 52, 26, 13, 40, 20, 10, 5, 16, 8, 4, 2, 1, 4, 2, 1, ...
Notons qu'une fois qu'on arrive à 1, les termes suivants de la suite seront toujours 4, 2, 1, 4, 2, 1, ... , se répétant une infinité de fois.
2. Le seul point qui nécessite un peu d'astuce est la détermination de la parité d'un entier en Scilab .
Nous proposons ici d'utiliser l'astuce suivante : un entier n est pair si et seulement si $n = 2 \left\lfloor \frac{n}{2} \right\rfloor$.

Remarque

Ceci semble indiquer que la divergence de la suite est très lente, et que pour avoir par exemple $\sum_{k=1}^n \frac{1}{k} > 100$, il doit falloir n très grand (ne demandez pas ceci à Sci Lab le temps d'exécution du programme serait bien trop long !).

```

1 function n = vol(k)
2     n = 0
3     u=k
4     while u>1
5         n = n+1 ;
6         if floor(u/2) == u/2 then
7             u = u/2 ;
8         else
9             u=3*u+1 ;
10        end
11    end
12 endfunction

```

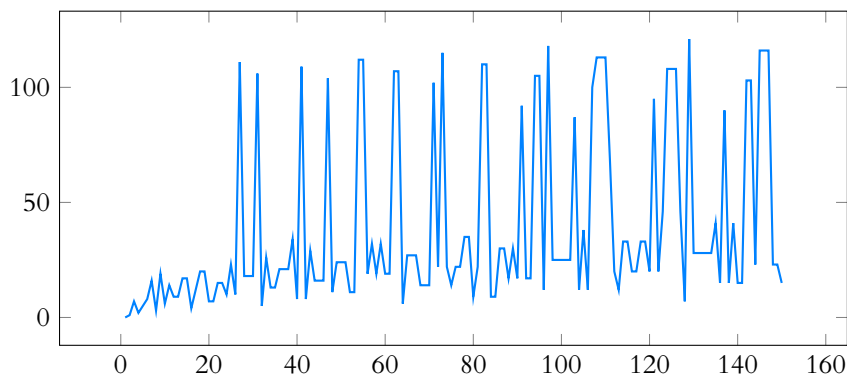
3. Utilisons une boucle pour calculer les différents temps de vol, puis la commande `plot2d` pour tracer le graphique.

```

1 A = zeros(1,150) ;
2 for i=1 :150
3     A(i) = vol(i) ;
4 end
5 plot2d([1 :150],A)

```

On obtient alors le graphique suivant :



4. Nous allons nous contenter de modifier très légèrement la fonction `vol` de la question 2. En effet, les étapes sont les mêmes, il s'agit juste d'intégrer le calcul de l'altitude à cette fonction.

Pour cela, notons qu'à chaque étape, l'altitude maximale atteinte est le maximum de l'altitude en cours, et de l'altitude maximale précédemment atteinte.

Autrement dit : $\max(u_1, \dots, u_n) = \max(u_n, \max(u_1, \dots, u_{n-1}))$.

```

1 function [t,a] = syracuse(k)
2     t = 0 ;
3     a = k ;
4     u=k ;
5     while u>1
6         t = t+1 ;
7         if floor(u/2) == u/2 then
8             u = u/2 ;
9         else
10            u=3*u+1 ;
11        end
12        a = max(a,u) ;
13    end
14 endfunction

```

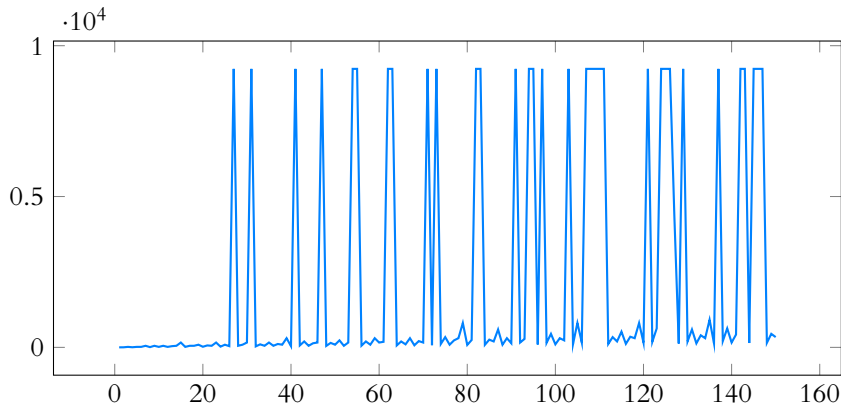
5. Il s'agit de procéder comme à la question 3 :

```

1 A = zeros(1,150) ;
2 for k=1 :150
3     [t,a] = syracuse(k)
4     A(k) = a
5 end
6 plot2d([1 :150],A)

```

On obtient alors le graphique suivant :



□

SOLUTION DE L'EXERCICE 11

- Il suffit de simuler n lois normales centrées réduites, et d'ajouter leurs carrés.

```

1 function y = chi_deux(n)
2     y=0
3     for i =1 :n
4         y = y + grand(1,1,'nor',0,1)^2
5     end
6 endfunction

```

Notons qu'il existe une solution un peu plus rapide si on se souvient (voir TP1) que $A.^2$ a pour effet d'élever tous les coefficients d'une matrice¹ au carré.

On peut alors se contenter de

```

1 function y =chi_deux(n)
2     y = sum(grand(1,n,'nor',0,1).^2)
3 endfunction

```

- Il nous suffit de faire appel p fois à la fonction `chi_deux`, et de ne garder que le plus grand résultat.

```

1 function y = T(p,n)
2     y = chi_deux(n) ;
3     for i = 1 :p-1
4         y = max(y,chi_deux(n)) ;
5     end
6 endfunction

```

- D'après la loi faible des grands nombres² si Z_1, \dots, Z_q sont q variables indépendantes de même loi que T_p , alors $\frac{1}{q} \sum_{k=1}^q Z_k \xrightarrow{q \rightarrow +\infty} E(T_p)$.

Bref, il s'agit de faire comme d'habitude : réaliser un grand nombre de simulations de T_p , et calculer la moyenne de ces simulations !

¹ Carrée ou non.

² Sous réserve qu'elle s'applique, c'est-à-dire si T_p admet une variance. C'est bien le cas, mais contentons nous de l'admettre sans démonstration.


```

1 n = input("Choisir la valeur de n")
2 p = input("Choisir la valeur de p") ;
3 N = 10 000 ;
4 s = 0 ;
5 for i=1 :n
6     s = s+T(p,n) ;
7     end
8 disp(s/N)

```

□

SOLUTION DE L'EXERCICE 12

Nous allons simuler n lois de Bernoulli $\mathcal{B}(p)$, en se rappelant³ qu'un moyen simple de procéder est de tirer un nombre avec `rand()` et de retourner 1 si ce nombre est inférieur ou égal à p et 0 sinon.

³ Voir TP2.

```

1 function y = binomiale(n,p)
2     y = 0 ;
3     for i = 1 :n
4         if rand()<p then
5             y = y+1
6         end
7     end
8 endfunction

```

□

SOLUTION DE L'EXERCICE 13

Il s'agit de tirer un premier nombre k suivant une loi uniforme sur $\llbracket 2, n \rrbracket$, puis d'utiliser ce nombre pour simuler une uniforme sur $\llbracket 1, k - 1 \rrbracket$.

```

1 function y = tirage(n)
2     k = grand(1,1,'uin',2,n) ;
3     y = grand(1,1,'uin',1,k-1) ;
4 endfunction

```

Il est possible de se passer de `grand()` en utilisant l'astuce suivante, expliquée en détail dans le TP2 (exercice 3) : $a + \text{floor}((b-a+1)*\text{rand}())$ simule une loi uniforme sur $\llbracket a, b \rrbracket$.

```

1 function y = tirage(n)
2     k = 2+floor((n-1)*rand()) ;
3     y = 1 + floor((k-1)*rand()) ;
4 endfunction

```

□

SOLUTION DE L'EXERCICE 14

Le plus simple est probablement de créer un grand vecteur contenant directement les résultats des n lancers, puis de chercher combien de fois ce vecteur contient deux coefficients consécutifs égaux.

```

1 function y = deux_fois(n)
2     X = grand(1,n,'bin',1,0.5) ;
3     y = 0 ;
4     for i=2 :n
5         if X(i) == X(i-1) then
6             y = y+1
7         end
8     end
9 endfunction

```

□

SOLUTION DE L'EXERCICE 15

1. Il s'agit de remarquer que la probabilité d'obtenir une boule rouge est $\frac{x}{x+y}$ (nombre de boules rouges divisé par le nombre total de boules), et donc on peut proposer le programme suivant :

```

1  function res = tirage(x,y)
2      r = rand()
3      if r<x/(x+y) then
4          res=0
5      else
6          res=1
7      end
8  endfunction

```

2. À chaque étape, si $r = 0$ (si la boule tirée est rouge), alors on rajoute une boule rouge dans l'urne, donc on augmente x de 1.
 Au contraire, si la boule tirée est blanche, on augmentera y de 1.
 À la fin, X_n est le nombre de boules rouges **ajoutées** dans l'urne : c'est la différence entre le nombre final de boules rouges (stocké dans la variable x) et le nombre initial de boules rouges (stocké dans la variable a).

```

1  function Xn = experience(a,b,n)
2      x = a
3      y = b
4      for k=1 :n
5          r = tirage(x,y)
6          if r==0 then
7              x = x+1
8          else
9              y = y+1
10         end
11     end
12     Xn = x-a
13 endfunction

```

3. Notons que X_n peut prendre $n + 1$ valeurs différentes (de 0 à n), et que donc le vecteur loi va devoir être de taille $n + 1$.
 Puisque dans Scilab les éléments d'un vecteur sont numérotés à partir de 1 et non de 0, on stockera $P(X_n = k)$ dans la $(k + 1)^{\text{ème}}$ colonne du vecteur loi, de sorte que $P(X_n = 0)$ corresponde à loi(1), $P(X_n = 1)$ corresponde à loi(2), etc.

```

1  function loi = simulation(a,b,n,m)
2      loi = zeros(1,n+1)
3      for i = 1 :m
4          u = experience(a,b,n)
5          loi(u+1) = loi(u+1)+1
6      end
7      loi=loi/m
8  endfunction

```

Remarque

Puisque l'énoncé nous indique qu'on souhaite des fréquences, on n'oublie pas de diviser le vecteur loi par le nombre total de répétitions.

□

TP 9 : FONCTIONS DE PLUSIEURS VARIABLES

Soit $f : \mathbf{R}^2 \rightarrow \mathbf{R}$. Son graphe est alors une surface dans \mathbf{R}^3 , que l'on peut tracer avec Scilab .

Le moyen le plus simple de procéder est le suivant :

- Commencer par définir la fonction dans Scilab à l'aide de la commande `function z = f(x,y) endfunction`. Par exemple, pour définir la fonction $f(x, y) = \sin(x) \sin(y)$, on utilisera


```
1 function z = f(x,y)
2   z = sin(x) * sin(y)
3 endfunction
```
- Comme on ne peut pas représenter la fonction sur \mathbf{R}^2 tout entier, on choisit un ensemble de la forme $[a, b] \times [c, d]$ sur lequel on souhaite représenter la fonction. Une fois ces valeurs choisies, on crée alors deux vecteurs X et Y , découpant ces intervalles en n petits intervalles de longueurs égales à l'aide des commandes `X = linspace(a,b,n)` et `Y = linspace(c,d,n)`.
- La commande `fplot3d(X,Y,f)` permet alors de représenter le graphe de la fonction. Nous utiliserons en fait `fplot3d1` au lieu de `fplot3d`, qui fournit des graphiques plus colorés. Toutefois, c'est la seconde qui est au programme (mais leurs fonctionnements sont identiques). **De toutes les commandes données dans ce TP, `fplot3d` est la seule qui soit au programme et dont vous devez connaître le fonctionnement.**
- Facultatif** : les couleurs n'étant pas des plus jolies, ni toujours des plus lisibles, on peut en choisir d'autres à l'aide des commandes suivantes, à exécuter alors que la fenêtre graphique est ouverte :

```
1 g =(gcf());
2 g.color_map = rainbowcolormap(128);
```

Si les couleurs ne sont toujours pas à votre goût, vous pouvez remplacer `rainbowcolormap` par `hotcolormap`, `coolcolormap`, `jetcolormap`, `summercolormap` ou `springcolormap`.

On peut alors faire tourner le graphe afin de l'orienter à sa guise à l'aide du clic droit (ou **Ctrl**) + Clic pour les adeptes de la marque à la pomme).

EXERCICE 1 Premier exemple

Considérons la fonction $f(x, y) = xy - xy^2 - x^2y$.

- Tracer le graphe de cette fonction sur $[-1, 1.5] \times [-1, 1.5]$, en prenant, à l'étape 2, $n = 30$.
- Fermer la fenêtre graphique (ou utiliser `clf()`) et faire de même avec $n = 3$, $n = 10$ et $n = 200$. Que constatez-vous ? Dans la suite, on pourra systématiquement prendre $n = 50$, qui est généralement un bon compromis.
- Déterminer graphiquement le(s) point(s) critique(s) de f . Quelle est la nature locale de ce(s) point(s) critique(s) ? (On pourra par exemple choisir de tracer f sur un domaine beaucoup plus petit que $[-1, 1.5] \times [-1, 1.5]$, afin de zoomer sur le(s) point(s) critique(s)). Retrouver ces points critiques par le calcul.
- La fonction f admet-elle un maximum global et/ou un minimum global ?

EXERCICE 2 Encore des points critiques

- La fonction $f(x, y) = x^2(1 + y)^3 + y^2$ possède $(0, 0)$ comme unique point critique. Le vérifier graphiquement et retrouver la nature de ce point critique.
- La fonction $f(x, y) = x(\ln(x)^2 + y^2)$ possède deux points critiques qui sont $(1, 0)$ et $(e^{-2}, 0)$. Les retrouver graphiquement, ainsi que la nature ces points critiques (on pourra choisir des fenêtres différentes pour chacun d'entre eux).

EXERCICE 3 Points selles, valeurs propres de hessiennes

- L'exemple typique de point selle est le point critique $(0, 0)$ de la fonction $f(x, y) = x^2 - y^2$. La tracer et observer le résultat. La hessienne de f en $(0, 0)$ possède -2 comme valeurs propres. Pouvez-vous, sans le moindre calcul, déterminer une base orthonormée formée de vecteurs propres de $\nabla^2 f(0, 0)$?
- La fonction définie sur $(\mathbf{R}_+^*)^2$ par $f(x, y) = x^y - y^x$ possède (e, e) comme point critique. Retrouver graphiquement sa nature.

- La fonction $f(x, y) = x^3 + xy + y^3$ possède deux points critiques. Les déterminer graphiquement, vérifier que l'un correspond à un maximum local et l'autre à un point selle.
On admet que la hessienne de f en le point selle a deux valeurs propres non nulles. Déterminer (graphiquement) un vecteur propre associé à chaque valeur propre.

EXERCICE 4 Soit $f(x, y) = e^{x^2+y^2} - \ln(x) - \ln(y)$.

- Calculer le gradient de f , et montrer que (x, y) est un point critique de f si et seulement si $\begin{cases} x = y \\ 2x^2 e^{2x^2} = 1 \end{cases}$.
- Montrer que l'équation $2x^2 e^{2x^2} = 1$ possède exactement une unique solution $\alpha \in \mathbf{R}_+^*$. Déterminer graphiquement la valeur de α à 0.01 près.
- Graphiquement, déterminer la nature locale de l'unique point critique de f .
- Calculer les dérivées partielles secondes de f , et à l'aide de la commande `spec`, déterminer les valeurs propres de la hessienne de f en (α, α) . Retrouver le résultat de la question précédente.

EXERCICE 5 Lignes de niveau

La commande `contour(X, Y, f, n)` permet de tracer n lignes de niveau de la fonction f .

Par exemple, essayer cette commande sur la fonction $f(x, y) = e^{-x^2} \sin(y)$ sur $[-2, 2] \times [-2, 2]$, et comparer au graphique de la fonction.

Pour tracer les lignes de niveau dans une autre fenêtre que le graphe de f , on pourra commencer par utiliser `fplot3d1`, puis ouvrir un autre graphique à l'aide de `scf()`, puis utiliser `contour`.

EXERCICE 6 Gradient et lignes de niveau

On reprend la fonction de l'exercice précédent.

- Créer deux fonctions nommées `D1f` et `D2f` qui correspondent à $\partial_1 f$ et $\partial_2 f$.
- SciLab permet de représenter graphiquement le gradient de f à l'aide de la commande

```
champ(X, Y, feval(X, Y, D1f), feval(X, Y, D2f))
```

Exécuter cette commande. Pouvez-vous interpréter le résultat ? On pourra pour cela tracer le graphique de f dans une nouvelle figure et comparer les résultats.

- Sur le même graphique que le gradient, représenter des lignes de niveau (cf. exercice précédent). Quel résultat évoqué en cours retrouvez-vous ?

EXERCICE 7 Un point selle dégénéré

La fonction $f(x, y) = (x^2 - y)(3x^2 - y)$ admet $(0, 0)$ comme unique point critique, et nous pourrions prouver qu'il s'agit d'un point selle.

Vérifier graphiquement qu'il s'agit d'un point critique.

Reconnaissez-vous un point selle ? Déterminer la hessienne en $(0, 0)$ et ses valeurs propres. Commenter.

EXERCICE 8 Lecture graphique

Soit $f(x, y) = \frac{1}{1+x^2} \sin(x^2 + y^2)$.

- Tracer le graphe de f , en choisissant une ou plusieurs fenêtres de votre choix. f semble-t-elle posséder des extrema locaux ? Globaux ? Pouvez-vous deviner en quel(s) point(s).
- Prouver les résultats observés graphiquement.

PLAN TANGENT

La formule de Taylor à l'ordre 1 nous dit qu'au voisinage d'un point $a = (x_0, y_0)$, on a

$$f(a + h) \simeq f(a) + \langle \nabla f(a), h \rangle.$$

Soit encore $f(x, y) \simeq f(x_0, y_0) + \partial_1 f(x_0, y_0)(x - x_0) + \partial_2 f(x_0, y_0)(y - y_0)$.

Cette dernière équation est alors l'équation au plan tangent au graphe de f en (x_0, y_0) .

EXERCICE 9 Quelques plans tangents

Pour tracer le plan tangent à une fonction f en un point (x_0, y_0) , on peut commencer par calculer le gradient de cette fonction en (x_0, y_0) , puis définir une fonction g par

$$g(x, y) = f(x_0, y_0) + \partial_1 f(x_0, y_0)(x - x_0) + \partial_2 f(x_0, y_0)(y - y_0).$$

En traçant cette fonction (sur le même graphique que le graphe de f), au moins au voisinage de (x_0, y_0) , on obtient le plan tangent.

Appliquer cette méthode pour tracer les plans tangents à $f(x, y) = (x - y)^2 + x^3 + y^3$ en $(0, -1)$, en $(0, 0)$ et en $(-1, -1)$.

EXERCICE 10 Position du plan tangent

D

1. Si $f : \mathbf{R} \rightarrow \mathbf{R}$ est une fonction \mathcal{C}^2 , rappeler en quoi la formule de Taylor-Young à l'ordre 2 permet, lorsque la dérivée seconde de f en a est non nulle, de déterminer la position de \mathcal{C}_f par rapport à sa tangente en $x = a$ (au moins au voisinage de a).
2. Soit à présent f une fonction \mathcal{C}^2 de \mathbf{R}^2 dans \mathbf{R} . En utilisant la formule de Taylor à l'ordre 2, formuler une hypothèse sur le lien entre les valeurs propres de la hessienne de f en a et la position du graphe de f par rapport à son plan tangent en a . En particulier, quel résultat retrouve-t-on pour un point critique ?
3. Vérifier cette hypothèse sur les plans tangents de l'exercice précédent en calculant les valeurs propres des hessiennes correspondantes.

EXERCICE 11 Calcul approché du gradient et de la hessienne

TD

Rappelons que les dérivées partielles de f en un point (x, y) sont définies par

$$\partial_1 f(x, y) = \lim_{h \rightarrow 0} \frac{f(x+h, y) - f(x, y)}{h} \text{ et } \partial_2 f(x, y) = \lim_{h \rightarrow 0} \frac{f(x, y+h) - f(x, y)}{h}.$$

1. Proposer une méthode pour calculer approximativement le gradient en un point.
2. En déduire une fonction calculant approximativement la hessienne de f en un point.
3. Appliquer cette méthode avec la fonction de l'exercice 4.

CORRECTION DU TP 9

SOLUTION DE L'EXERCICE 1

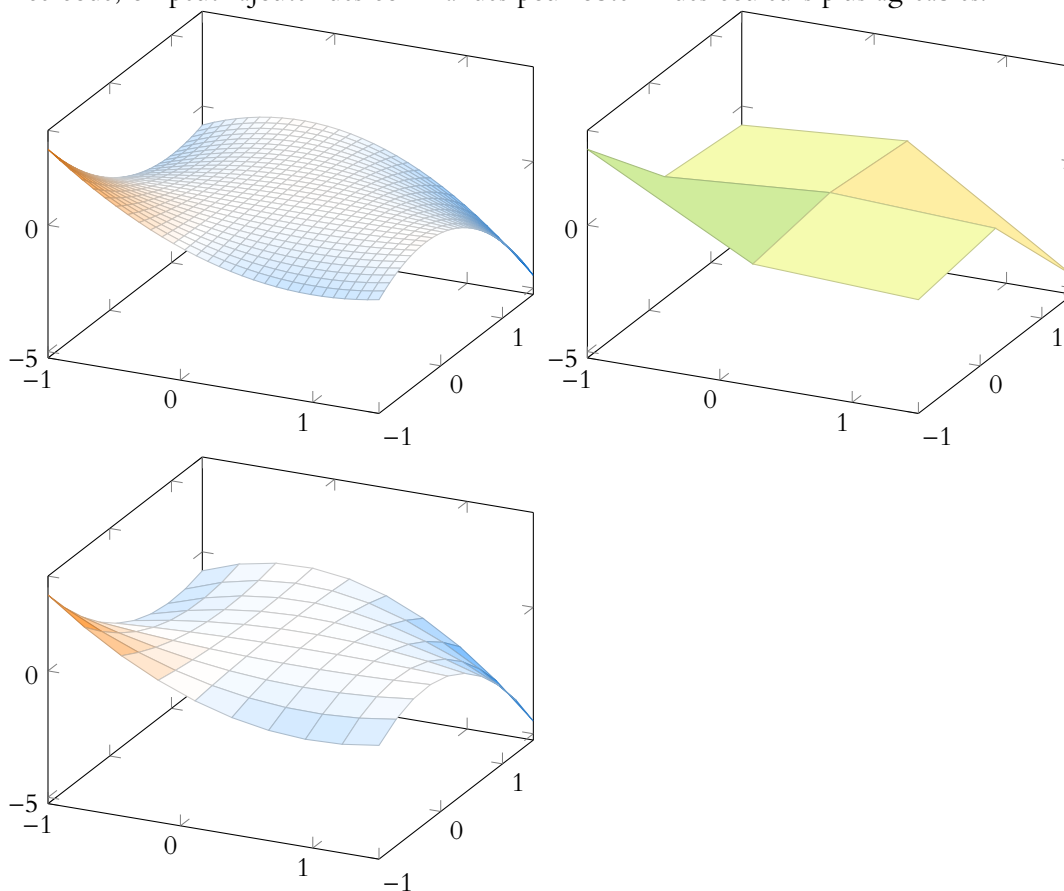
1. Le code indispensable est donc

```

1 function z = f(x,y)
2     z = x*y-x*y^2-y*x^2
3 endfunction
4 X = linspace(-1,1.5,30);
5 Y = linspace(-1,1.5,30);
6 fplot3d1(X,Y,f)

```

À ce code, on peut rajouter des commandes pour obtenir des couleurs plus agréables.



2. Un point critique est un point où le plan tangent est horizontal. Sur la figure, on semble en voir en $(0, 0)$, $(1, 0)$ et $(0, 1)$, le premier étant un point selle, les deux autres étant des maximum locaux.
3. Non, car on a $\lim_{x \rightarrow +\infty} f(x, x) = +\infty$ et $\lim_{x \rightarrow +\infty} f(x, 1) = -\infty$.

□

SOLUTION DE L'EXERCICE 4

1. Pour tracer la fonction $x \mapsto 2x^2e^{2x^2}$, nous pouvons utiliser le code suivant :

```

1 x = linspace(0,1,50)
2 y = zeros(1,50)
3 for i=1 :50
4     y(i) = 2*x(i)^2*exp(2*x(i)^2);
5 end
6 plot2d(x,y)

```

Remarque

Le dessin produit n'autorise pas forcément une excellente précision, mais donne comme première approximation $0.4 \leq \alpha \leq 0.6$. Il est possible de «réduire» la fenêtre en remplaçant la première ligne par $x = \text{linspace}(0.4, 0.6, 50)$

2. Graphiquement, il s'agit d'un minimum local.

3. On a

$$\partial_1 f(x, y) = 2xe^{x^2+y^2} - \frac{1}{x}, \quad \partial_2 f(x, y) = 2ye^{x^2+y^2} - \frac{1}{y}.$$

Et donc

□

SOLUTION DE L'EXERCICE 6

1. Puisque $\partial_1 f(x, y) = -2xe^{-x^2} \sin(y)$ et $\partial_2 f(x, y) = e^{-x^2} \cos(y)$, on peut utiliser

```
1 function z = D1f(x,y)
2     z = -2*x*exp(-x^2)*sin(y)
3 endfunction
4
5 function z = D2f(x,y)
6     z = exp(-x^2)*cos(y)
7 endfunction
```

□

SOLUTION DE L'EXERCICE 9

Commençons par calculer les dérivées partielles de f :

$$\partial_1 f(x, y) = 2(x - y) + 3x^2, \quad D_2 f(x, y) = 2(y - x) + 3y^2.$$

Puis créons deux fonctions : la fonction f et la fonction g définie par l'énoncé :

```
1 function z = f(x,y)
2     z = (x-y)^2 + x^3 + y^3
3 endfunction
4
5 function z = g(x,y)
6     z = 2*x+(y+1)
7 endfunction
```

Puis, il faut tracer ces deux fonctions sur le même graphique, par exemple avec

```
1 x = linspace(-0.5, 0.5, 30)
2 y = linspace(-1.5, -0.5, 30)
3 fplot3d1(x,y,f)
4 fplot3d1(x,y,g)
```

□

SOLUTION DE L'EXERCICE 10

1. La formule de Taylor-Young nous donne, dans le cas où $f''(a) \neq 0$ l'équivalent suivant

$$f(x) - (f(a) + f'(a)(x - a)) \underset{x \rightarrow a}{\sim} f''(a)(x - a)^2.$$

Mais $y = f(a) + f'(a)(x - a)$ est l'équation de la tangente à f en a , et donc $f(x) - (f(a) + f'(a)(x - a)) \geq 0$ si et seulement si la courbe de f est située au dessus de la tangente.

Mais, au voisinage de a , $f''(a)(x - a)^2$ est du signe de $f''(a)$, donc si $f''(a) > 0$, alors la courbe de f est située au dessus de la tangente. Et inversement, si $f''(a) < 0$, alors la courbe de f est située en dessous de la tangente.

2. La formule de Taylor à l'ordre 2 nous donne

$$f(a + h) - (f(a) + \langle \nabla f(a), h \rangle + \frac{1}{2} q_a(h) + \|h\|^2 \varepsilon(h))$$

de sorte que, pour h proche de 0 (c'est-à-dire $a + h$ au voisinage de a), $f(a + h) - (f(a) + \langle \nabla f(a), h \rangle)$ est du signe de $q_a(h)$.

Mais $z = f(a) + \langle \nabla f(a), h \rangle$ est l'équation du plan tangent à f en a , et donc $f(a + h) - (f(a) +$

Remarque

Dans le cas où $f''(a) = 0$, cet équivalent n'est plus valable car on écrit alors un équivalent à 0 !

$\langle \nabla f(a), h \rangle$ est positif si et seulement si la surface représentant f est située au dessus de la courbe.

Et donc, si $\nabla^2 f(a)$ ne possède que des valeurs propres positives strictement, le plan tangent est situé au dessus la surface.

En particulier, si a est un point critique, le plan tangent est horizontal, et dire que la courbe de f est située¹ au dessus de ce plan revient à dire que f possède un minimum local en a , résultat que nous connaissons déjà.

De plus, lorsque $\nabla^2 f(a)$ possède des valeurs propres de signes contraire, alors la position du plan tangent n'est pas constante : la surface «traverse» le plan tangent en a .

¹ Toujours au voisinage de a .

3. Avec la fonction de l'exercice 8, on a

$$\nabla^2 f(x, y) = \begin{pmatrix} 2 + 6x & -2 \\ -2 & 2 + 6y \end{pmatrix}.$$

En particulier, on trouve que les valeurs propres de $\nabla^2 f(0, 0)$ sont 0 et 4,

□

TP 10 : ESTIMATION, MÉTHODES DE MONTE-CARLO

RETOUR SUR LA SIMULATION D'UNE LOI NORMALE

On rappelle que SciLab dispose d'une fonction permettant de calculer la fonction de répartition d'une variable aléatoire suivant une loi normale.

Si $X \hookrightarrow \mathcal{N}(\mu, \sigma^2)$, alors on obtient $F_X(x) = P(X \leq x)$ à l'aide de la commande `cdfnor('PQ', x, mu, sigma)`.

EXERCICE 1 Soit $(U_k)_k$ une suite de variables aléatoires i.i.d. suivant la loi uniforme sur $[0, 1]$.

On pose, pour tout $n \in \mathbf{N}^*$, $X_n = \sqrt{\frac{12}{n}} \left(\sum_{k=1}^n U_k - \frac{n}{2} \right)$.

1. Montrer que (X_n) converge en loi vers une variable aléatoire X suivant la loi $\mathcal{N}(0, 1)$.
2. Pour $n = 12$, la définition de X_n est particulièrement simple. Compléter le code suivant afin qu'il crée un vecteur A contenant 10000 simulations de X_{12} .

```
1 A = -6*ones(1,10000)
2 for i = 1 :10000
3     for j = 1 :...
4         A(i) = A(i) + ...
5     end
6 end
```

3. On se redonne la fonction polygone définie au TD6, qui permet de tracer le polygone des fréquences cumulées d'une série de nombres.

```
1 function polygone(Y)
2     M = tabul(Y, 'i') ;
3     N = length(Y) ;
4     plot2d(M(:,1), cumsum(M(:,2))/N)
5 endfunction
```

En utilisant cette fonction, représenter la distribution des valeurs prises par X_{12} .

4. Créer une fonction `phi` qui calcule $\Phi(x)$ à l'aide `cdfnor`, et comparer alors le graphique de Φ à celui obtenu à la question précédente.
5. Faire de même avec X_{20} et X_{1000} , et commenter les résultats.
6. Comment exploiter ce qui précède pour simuler une loi normale $\mathcal{N}(2, 9)$? Comment vérifier alors le résultat ?

ESTIMATION PAR INTERVALLE DE CONFIANCE

Afin de déterminer des intervalles de confiance, nous aurons souvent besoin d'inverser la fonction Φ , notamment pour déterminer le réel t_α tel que $\Phi(t_\alpha) = 1 - \frac{\alpha}{2}$, soit encore $t_\alpha = \Phi^{-1} \left(1 - \frac{\alpha}{2} \right)$.

Nous utiliserons à cet effet la commande `cdfnor` de la manière suivante : pour déterminer l'unique réel x tel qu'une variable aléatoire $X \hookrightarrow \mathcal{N}(\mu, \sigma)$ vérifie $P(X \leq x) = p$, on saisit

```
x = cdfnor('X', mu, sigma, p, 1-p)
```

Par exemple, si l'on souhaite déterminer le nombre x tel que $\Phi(x) = 0.75$, on entre `x = cdfnor('X', 0, 1, 0.75, 0.25)`.

EXERCICE 2 Des valeurs classiques

Avec les notations habituelles, déterminer les valeurs $t_{0.05}$ et $t_{0.01}$.

EXERCICE 3 Un intervalle de confiance

Nous avons prouvé en TD que si X_1, \dots, X_n suivent une loi normale $\mathcal{N}(\theta, 1)$, alors $\left[\overline{X}_n - \frac{t_\alpha}{\sqrt{n}}, \overline{X}_n + \frac{t_\alpha}{\sqrt{n}} \right]$ est un intervalle de confiance de θ au niveau de confiance $1 - \alpha$.

Comme précédemment, choisissons θ au hasard, par exemple à l'aide de `theta = grand(1, 1, 'exp', 100)` ;

Dans la suite, on pourra prendre, par exemple, $\alpha = 0.05$. On crée ensuite une matrice contenant 10000 échantillons de X_1, \dots, X_{10} :

```
X = grand(10000, 10, 'nor', theta, 1) ;
```

1. Créer deux vecteurs A et B tels que, pour tout i , l'intervalle de confiance associé au i -ème échantillon soit $[A(i), B(i)]$.
2. Parmi les 10000 échantillons, quelle proportion vérifie vraiment $\theta \in [A(i), B(i)]$?
3. Écrire une fonction `y = etendue(n, alpha)` qui prend en paramètre un entier n et un réel $\alpha \in]0, 1[$ et renvoie l'étendue (=la longueur) de l'intervalle de confiance de θ correspondant.
Essayer ce programme pour $\alpha = 0.05$, en faisant varier n et constater l'effet que cela a sur l'étendue de l'intervalle.
De même, à n fixé, étudier l'étendue de l'intervalle pour $\alpha = 0.05, \alpha = 0.01, \alpha = 0.001$ et $\alpha = 0.0001$.

MÉTHODES DE MONTE-CARLO

L'appellation de «méthode(s) de Monte-Carlo» regroupe toute une famille d'algorithmes qui permettent de calculer des valeurs approchées de certaines quantités de manière probabiliste. C'est-à-dire qu'ils renvoient une valeur qui est probablement une «bonne» approximation de la quantité que l'on cherche à approcher, mais qu'il reste toujours un risque que la valeur retournée ne soit pas bonne.

Par exemple, nous avons déjà expliqué en cours que si f est une fonction continue sur $[0, 1]$, et si U_1, \dots, U_n sont des variables i.i.d. suivant la loi uniforme sur $[0, 1]$, alors une des conséquences de la loi faible des grands nombres est que

$$X_n = \frac{1}{n} \sum_{i=1}^n f(U_i) \text{ converge en probabilités vers } I = \int_0^1 f(t) dt.$$

Et donc pour n grand, une simulation de X_n est probablement une bonne approximation de $\int_0^1 f(t) dt$.

EXERCICE 4 On souhaite mettre en œuvre la méthode précédente pour déterminer une valeur approchée de $I = \int_0^1 \frac{e^{-t}}{1+t^4} dt$.

Compléter le programme suivant afin qu'il retourne une valeur approchée de I .

```

1  fonction y = f(x)
2      y = .....
3  endfunction
4
5  n = 10000 ;
6  S = 0 ;
7  for i=1 :n
8      S = S+....
9  end
10 disp(....)
    
```

Essayer ce programme avec différentes valeurs de n , en faisant plusieurs essais pour chaque valeur de n . Comparer les valeurs obtenues avec celle que donne `integrate('f(x)', 'x', 0, 1)`.

Dans l'exemple précédent, l'intérêt d'une méthode probabiliste n'est pas flagrant, puisque nous aurions également pu utiliser la méthode de rectangles (=sommes de Riemann) pour obtenir une valeur approchée de l'intégrale, sans risque d'erreur (il n'y a pas de hasard dans la méthode des rectangles).

En revanche, cette méthode peut aisément être adaptée pour des intégrales impropres, pour lesquelles la méthode des rectangles marche mal.

EXERCICE 5 (Extrait d'EDHEC 2017)

Pour $k \in \mathbf{N}^*$, on souhaite déterminer une valeur approchée de $I_k = \int_1^{+\infty} t^k e^{-t} dt$.

1. À l'aide d'un changement de variable bien choisi, montrer que $I_k = e^{-1} \int_0^{+\infty} (x+1)^k e^{-x} dx$.
2. Soient $(U_i)_{i \in \mathbf{N}^*}$ une suite de variables i.i.d. suivant la loi $\mathcal{E}(1)$. On pose alors, pour tout $i \in \mathbf{N}^*$, $Y_i = (U_i + 1)^k$.
 - 2.a. Montrer que les Y_i ont une espérance et la calculer. Ont-ils une variance ?
 - 2.b. Montrer que si on pose $X_n = \frac{e^{-1}}{n} \sum_{i=1}^n Y_i$, alors $X_n \xrightarrow{P} I_k$.
 - 2.c. Proposer un programme qui permette de calculer une valeur approchée de I_k . L'essayer pour différentes valeurs de n .

Il est également possible d'utiliser des raisonnements probabilistes pour approcher des aires.

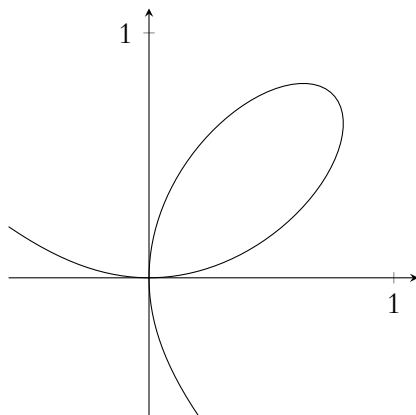
À titre d'exemple, considérons la courbe ci-dessous (appelée folium de Descartes), et cherchons à calculer une valeur approchée de l'aire formée par la boucle



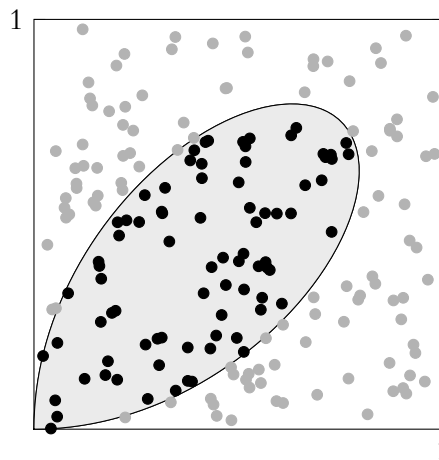
Pour cela, tirons n points au hasard dans le carré $[0, 1] \times [0, 1]$, uniformément répartis, c'est-à-dire des points dont l'abscisse et l'ordonnée sont tirées à l'aide d'une loi uniforme sur $[0, 1]$.

Comptons alors combien de points parmi les n sont dans la boucle de la courbe.

Le folium de Descartes.



200 points répartis aléatoirement dans $[0, 1]^2$.



La proportion de points qui sont dans la boucle de la courbe doit alors être proche du rapport $\frac{\text{aire de la boucle}}{\text{aire totale du carré}}$.

Puisqu'ici le carré est d'aire égale à 1, il suffit de compter le nombre de points « tombés » dans la boucle, et de diviser par n pour obtenir une approximation de l'aire de la boucle.

EXERCICE 6 La boucle dessinée ci-dessus est $\mathcal{B} = \left\{ (x, y) \in \mathbf{R}^2 : x \geq 0, y \geq 0, x^3 + y^3 \leq \frac{3xy}{2} \right\}$.

Compléter le programme suivant pour qu'il retourne une valeur approchée de l'aire de \mathcal{B} .

```

1 n = 10 000 ;
2 X = grand(1, n, 'unf', 0, 1) ;
3 Y = .....
4 S = 0 ;
5 for i=1 :n
6     if ..... then
7         S = S+1 ;
8     end
9 end
10 disp(S/n)

```

EXERCICE 7 Calcul d'une valeur approchée de π

Soit $\mathcal{D} = \{(x, y) \in \mathbf{R}^2 : x \geq 0, y \geq 0, x^2 + y^2 \leq 4\}$.

1. Représenter \mathcal{D} . Quelle est son aire ?
2. Soient $(X_i)_{1 \leq i \leq n}$ et $(Y_i)_{1 \leq i \leq n}$ des variables i.i.d. suivant la loi $\mathcal{U}([0, 2])$. On note T_i la variable qui vaut 1 si le point (X_i, Y_i) est dans \mathcal{D} et 0 sinon.
 - 2.a. Déterminer la loi des T_i .
 - 2.b. Montrer que si $\overline{T_n} = \frac{1}{n} \sum_{i=1}^n T_i$, alors $4\overline{T_n}$ converge en probabilités vers π .
 - 2.c. Écrire un programme qui calcule une valeur approchée de π .
3. Déterminer un intervalle de confiance asymptotique de π , au niveau de confiance $1 - \alpha$, dont les bornes dépendent de T_n .

CORRECTION DU TP 10

SOLUTION DE L'EXERCICE 1

1. Puisque les U_k sont indépendantes, on a $E\left(\sum_{k=1}^n U_k\right) = \sum_{k=1}^n E(U_k) = n\frac{1}{2}$ et

$$V\left(\sum_{k=1}^n U_k\right) = \sum_{k=1}^n V(u_k) = n\frac{1}{12}.$$

Par conséquent, la variable aléatoire centrée réduite associée à $\sum_{k=1}^n U_k$ est

$$\overline{S}_n^* = \frac{\sum_{k=1}^n U_k - \frac{n}{2}}{\sqrt{\frac{n}{12}}} = \sqrt{\frac{12}{n}} \left(\sum_{k=1}^n U_k - \frac{n}{2} \right).$$

Et donc, par le théorème central limite¹ $\overline{S}_n^* \xrightarrow{\mathcal{L}} X$, où $X \leftrightarrow \mathcal{N}(0, 1)$.

¹ Qui s'applique car les U_k sont indépendantes, de même loi, et possèdent une variance.

```
2.
1 A = -6*ones(1,10000)
2 for i = 1 :10000
3     for j = 1 :12
4         A(i) = A(i) + rand();
5     end
6 end
```

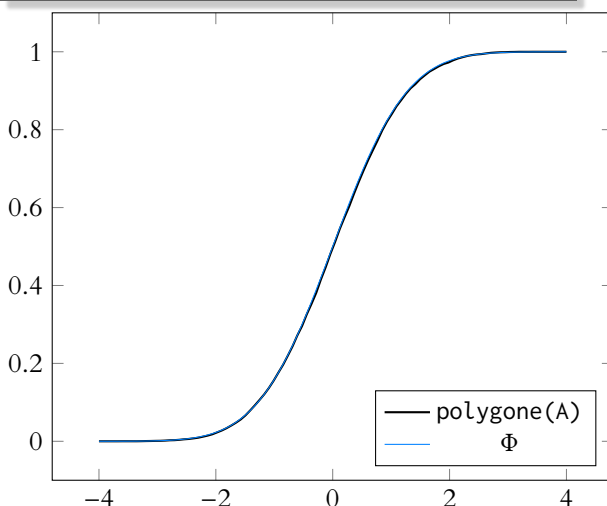
3. Pour tracer la fonction de répartition de la variable que nous venons de simuler, ou du moins une approximation de cette fonction de répartition, nous pouvons utiliser la fonction polygone.

```
1 polygone(A)
```

4. Si on souhaite comparer ce graphique à l'allure de Φ , il suffit de tracer (sur le même graphique²) la fonction Φ , à l'aide de cdfnor.

```
1 function y = phi(x)
2     [p,q] = cdfnor('PQ',x,0,1);
3     y = p
4 endfunction
5
6 x = linspace(-4,4,100);
7 fplot2d(x,phi);
```

² Et pour cela il suffit de ne pas fermer la fenêtre graphique !



CORRECTION

On constate alors que les deux courbes se superposent presque parfaitement. Et donc X_{12} semble être une bonne approximation d'une loi normale centrée réduite.

- L'approximation est encore meilleure avec 20 ou 1000 lois uniformes. Mais puisqu'elle était déjà très bonne pour 12, on ne décèle³ pas une grosse différence.
- Si l'on sait simuler une loi normale $\mathcal{N}(0, 1)$, alors en la multipliant par 3 et en ajoutant 2, on simule une loi $\mathcal{N}(2, 9)$.
Donc par exemple, il est possible d'utiliser le code suivant :

```
1 A = -6*ones(1,10000)
2 for i = 1 :10000
3     for j = 1 :12
4         A(i) = A(i) + rand();
5     end
6     A(i) = 3*A(i)+2;
7 end
```

Pour vérifier le résultat, il faut alors tracer la fonction de répartition d'une telle loi $\mathcal{N}(2, 9)$. Mais on sait que si $X \leftrightarrow \mathcal{N}(2, 9)$, alors

$$P(X \leq x) = P\left(\frac{X-2}{3} \leq \frac{x-2}{3}\right) = \Phi\left(\frac{x-2}{3}\right).$$

Et donc

□

SOLUTION DE L'EXERCICE 2

On a $1 - \frac{0.05}{2} = 0.975$ et $1 - \frac{0.01}{2} = 0.995$, et donc $t_{0.05}$ et $t_{0.01}$ sont obtenus respectivement à l'aide des commandes

```
1 cdfnorf('X', 0, 1, 0.975, 0.025)
2 cdfnorf('X', 0, 1, 0.995, 0.005)
```

On trouve alors $t_{0.05} \approx 1.96$ et $t_{0.01} \approx 2.57$.

□

SOLUTION DE L'EXERCICE 3

- Il faut utiliser une boucle afin de calculer les bornes de l'intervalle sur chacun des 10000 échantillons.

```
1 A = zeros(10000,1);
2 B = zeros(10000,1);
3 t = cdfnorf('X', 0, 1, 0.975, 0.025);
4 for i=1 :10000
5     A(i) = mean(X(i, :))-t/sqrt(10);
6     B(i) = mean(X(i, :))+t/sqrt(10);
7 end
```

- Il s'agit de compter parmi les 10000 échantillons pour combien d'entre eux $\theta \in [A(i), B(i)]$.

```
1 s = 0;
2 for i = 1 :10000
3     if (theta >=A(i)) & (theta <=B(i)) then
4         s = s+1;
5     end
6 end
7 disp(s)
```

On obtient environ 9500 échantillons pour lesquels θ se trouve dans l'intervalle de confiance. Ce qui est cohérent avec ce que l'on pouvait attendre, cela correspond bien à 95% des échantillons.

et avez obtenu un «mauvais tirage», pour lequel X_{12} approche très mal une loi normale. Toutefois, sur 10 000 variables aléatoires, cela est très peu probable.

³ À l'œil nu.

Rappel

Si $X \leftrightarrow \mathcal{N}(0, 1)$, alors $aX + b \leftrightarrow \mathcal{N}(b, a^2)$.

Φ

Rappelons que Φ est la fonction de répartition d'une loi normale centrée réduite.

3. Il suffit de constater que l'étendue de l'intervalle, indépendamment de l'échantillon choisi est $2\frac{t_\alpha}{\sqrt{n}}$.

```
1 function y = etendue(n,alpha)
2     y = 2*cdfnorf('X',0,1,1-alpha/2,alpha/2)/sqrt(n) ;
3 endfunction
```

Lorsqu'on essaie avec des valeurs de n de plus en plus grandes, on constate⁴ que l'intervalle est de plus en plus petit.
Au contraire, si α est de plus en plus petit, l'étendue est de plus en plus grande.

⁴ Sans surprise !

□

INDEX

- A**
 Application linéaire
 image, 82
 noyau, 82
 rang, 82
 Approximation
 d'une loi binomiale par une loi de Poisson, 437
 d'une loi binomiale par une loi normale, 440
 d'une loi de Poisson par une loi normale, 441
 Automorphisme, 83
- B**
 Base, 75
 adaptée à une somme directe, 78
 canonique
 de \mathbf{K}^n , 76
 de $\mathcal{M}_{n,p}(\mathbf{K})$, 76
 de $\mathbf{K}_n[X]$, 76
 orthonormée, 290
 Biais d'un estimateur, 483
 Borné de \mathbf{R}^n , 510
 Boule
 fermée, 508
 ouverte, 508
- C**
 Changement de variable pour les intégrales impropres, 114
 Coefficient de corrélation linéaire, 188
 Combinaison linéaire, 72
 Contrainte
 d'égalités linéaires, 564
 non critique, 561
 Convergence
 en loi, 432
 cas des variables discrètes, 435
 en probabilité, 429
 Coordonnées d'un vecteur dans une base, 76
 Covariance, 185
 Croissances comparées usuelles, 12
- D**
 Dénombrable, 19
 Densité, 245
 Dérivée
 directionnelle, 404
 partielle, 400
 Dérivée directionnelle
 seconde, 521
 Dérivée partielle
 d'ordre 2, 516
 Développements limités usuels, 12
 Droite (vectorielle), 76
- E**
 Échantillon, 482
 Endomorphisme, 82
 diagonalisable, 314
 induit, 154
 symétrique, 461
 Espace
 euclidien, 290
 probabilisable, 45
 probabilisé, 46
 Espace vectoriel
 de dimension finie, 76
 Espérance
 conditionnelle, 53
 d'une variable à densité, 246
 d'une variable aléatoire discrète, 49
 totale, 55
 Estimateur, 482
 asymptotiquement sans biais, 486
 convergent, 487
 condition nécessaire, 491
 sans biais, 483
 Extremum local, 514
- F**
 Famille
 génératrice, 74
 libre, 73
 liée, 73
 orthogonale, 287
 sommable, 20
 Fermé de \mathbf{R}^n , 509
 Fonction Γ , 117
 Fonction continue
 de \mathbf{R}^n dans \mathbf{R} , 396
 Fonction de classe \mathcal{C}^2 , 517
 Fonction de répartition, 47
 Fonction partielle, 398
 Forme bilinéaire, 281
 symétrique, 282
 Forme linéaire, 85
 Forme quadratique
 signe d'une, 468
 Forme quadratique associée à une matrice symétrique, 467
 Formule
 de Bayes, 52
 de changement de base, 151
 de Huygens, 50, 249
 pour la covariance, 186
 de l'espérance totale, 55
 des probabilités composées, 51
 des probabilités totales, 51
- G**
 Gradient, 400
 Gram-Schmidt
 procédé d'orthogonalisation de, 288

- H**
 Hessienne, 518
 Hyperplan, 86
- I**
 Image d'une application linéaire, 82
 Indépendance
 de deux événements, 175
 de deux variables aléatoires, 176
 Inégalité
 de Bienaymé–Tchebychev, 428
 de Cauchy–Schwarz, 284
 dans \mathbf{R}^n , 284
 pour les intégrales, 284
 de Markov, 427
 triangulaire
 généralisée, 16
 pour les intégrales, 111
 Intégrale
 convergente, 107
 de Riemann, 110
 d'une fonction paire/impair, 116
 faussement impropre, 107
 reste d'une intégrale convergente, 112
 Intégration par parties, 113
 Intervalle de confiance, 545
 asymptotique, 547
 Isomorphisme, 83
- L**
 Lemme des coalitions, 344
 Loi
 binomiale $\mathcal{B}(n, p)$, 50
 conditionnelle, 52
 conjointe, 177
 de Bernoulli $\mathcal{B}(p)$, 50
 de Cauchy, 246
 de Poisson $\mathcal{P}(\lambda)$, 50
 d'un couple de variables aléatoires, 177
 d'une variable aléatoire, 47
 exponentielle $\mathcal{E}(\lambda)$, 251
 gamma $\gamma(v)$, 252
 géométrique $\mathcal{G}(p)$, 50
 marginale, 177
 normale $\mathcal{N}(\mu, \sigma^2)$, 253
 triangulaire, 182
 uniforme
 à densité $\mathcal{U}([a, b])$, 251
 discrète $\mathcal{U}(\llbracket a, b \rrbracket)$, 50
 Loi faible des grands nombres, 430
- M**
 Matrice
 antisymétrique, 144
 de passage, 150
 de rang 1, 146
 diagonale, 144
 diagonalisable, 314
 hessienne, 518
 inversible, 144
 orthogonale, 293
 scalaire, 144
 symétrique, 144
 Matrices
 semblables, 152
 Maximum
 global, 406
 Maximum local, 514
 sous contrainte, 559
 Méthode de Monte-Carlo, 432
 Méthode d'inversion
 pour une variable discrète, 1062
 Minimum
 global, 406
 Minimum local, 514
 sous contrainte, 559
 Moments
 d'une variable à densité, 248
- N**
 Norme, 285
 Noyau d'une application linéaire, 82
- O**
 Orthogonalité
 de sous-espaces vectoriels, 286
 de vecteurs, 286
 Ouvert de \mathbf{R}^n , 508
- P**
 Plan (vectoriel), 76
 Point critique, 407
 sous contrainte d'égalité linéaire, 564
 sous une contrainte non critique, 562
 Point selle, 516
 Polynôme annulateur, 311
 Probabilité
 conditionnelle, 51
 Procédé d'orthogonalisation de Gram–Schmidt, 288
 Produit de convolution
 cas des variables à densité, 255
 cas des variables discrètes, 181
 Produit scalaire, 283
 canonique de \mathbf{R}^n , 283
 canonique de $\mathcal{M}_{n,1}(\mathbf{R})$, 292
 norme associée, 285
 Projecteur, 84
 orthogonal, 376
 Projection, 84
- R**
 Rang
 d'une application linéaire, 82
 d'une matrice, 146
 Règle $n^\alpha u_n$, 17
 Reste
 d'une intégrale convergente, 112
 d'une série convergente, 14
 Risque quadratique d'un estimateur, 489
 décomposition biais-variance, 489
- S**
 Série

- convergente, 13
 - absolument, 16
 - reste d'ordre n , 14
- de Riemann, 17
- exponentielle, 18
- géométrique, 18
 - dérivée, 18
- grossièrement divergente, 14
- harmonique, 17
- télescopique, 13
- Somme
 - de deux sous-espaces vectoriels, 77
 - de n sous-espace vectoriels, 79
 - directe
 - de deux sous-espaces vectoriels, 77
 - de n sous-espace vectoriels, 80
- Somme de Riemann, 105
- Sous-espace
 - propre
 - d'un endomorphisme, 216
 - d'une matrice, 211
 - vectorel, 72
- Sous-espace vectoriel
 - engendré par une partie, 72
 - stable par un endomorphisme, 86
- Sous-espaces supplémentaires, 78
- Spectre
 - d'un endomorphisme, 216
 - d'une matrice, 211
- Suite
 - bornée, 10
 - convergente, 9
 - divergente, 9
 - majorée, 10
 - minorée, 10
- Suite d'estimateurs, 486
 - asymptotiquement sans biais, 486
 - convergente, 487
- Suites adjacentes, 10
- Supplémentaire
 - orthogonal, 375
- Système d'événements
 - complet, 46
 - quasi complet, 46
- T**
- Théorème
 - central limite, 439
 - de Fubini, 20
 - de la base incomplète, 76
 - de la dimension, 76
 - de la limite monotone (suites), 10
 - de Pythagore, 287
 - de Schwarz, 518
 - de Slutsky, 438
 - de sommation par paquets, 20
 - de transfert
 - cas des variables à densité, 247
 - cas discret, 49
 - couples de variables discrètes, 180
 - des extrema liés, 562
 - du rang, 82
- Trace d'une matrice caréée, 155
- Tribu, 45
 - des boréliens, 47
 - engendrée par un couple de variables aléatoires, 179
 - engendrée par une variable aléatoire, 176
- V**
- Valeur propre
 - d'un endomorphisme, 216
 - d'une matrice, 210
- Variable
 - centrée, 250
 - réduite, 250
 - indicatrice, 48
- Variable aléatoire
 - à densité, 245
 - fonction de répartition, 47
 - loi, 47
 - réelle, 47
 - discrète, 48
- Variables aléatoires
 - i.i.d., 438
- Variance
 - d'une somme de variables discrètes, 187
 - d'une variable à densité, 249
 - d'une variable discrète, 49
- Vecteur aléatoire, 341
 - loi, 341
- Vecteur propre
 - d'un endomorphisme, 216
 - d'une matrice, 211