L'oral de maths à HEC et à l'ESCP

187 exercices corrigés.

M. Vienney
http ://www.ecs2-fauriel.fr
vienney@ecs2-fauriel.fr

Introduction

Avertissement

Ce document est en perpétuelle évolution, et bon nombre de corrigés n'ont pas été relus. Il est donc fort probable qu'il s'y trouve un certain nombre de coquilles, voire de questions non corrigées. La mise en page laisse également souvent à désirer...

Vous trouverez ici un recueil de quelques exercices posés ces dernières années à l'oral de l'ESCP et d'HEC.
Bien qu'il y ait parfois des ressemblances entre certains exercices, et des thèmes qui reviennent régulièrement, il n'y a pas ou peu d'exercices très classiques qu'il faut absolument savoir refaire, le but de ces oraux (comme de l'écrit des parisiennes) étant justement de vous sortir de votre «zone de confort» afin de voir comment vous réagissez à un contexte qui peut semble déroutant.
Il ne s'agit donc en aucun cas de faire tous ces exercices, mais plutôt d'essayer de bien en comprendre quelques uns.

Les exercices sont classés par thèmes, avec une indication (nécessairement subjective) de leur difficulté (Facile, Moyen ou Difficile), cette difficulté étant toute relative : un exercice «Facile» d'oral reste plutôt difficile par rapport à ce que nous avons fait dans l'année.

La quasi totalité des exercices avec préparation sont trop longs pour être finis dans le temps imparti, et ce n'est pas ce que l'on attend de vous. De même, il est rare de réussir à terminer les questions sans préparation dans les 10 à 15 minutes qui leur sont dévolues à l'oral. Mais peu importe, ce que l'examinateur va chercher à déceler, ce sont vos aptitudes à explorer des pistes (mêmes infructueuses) et votre capacité à réagir aux indications qui sont fournies en cours d'exercice.

Table des matières

1 Algèbre

1.1 Polynômes

Il existe des relations entres les racines d'un polynôme et ses coefficients.
La plus simple d'entre elles est pour les polynômes de degré deux : si $P(X) = X^2 + aX + b$ possède pour racines (réelles ou complexes) x_1 et x_2, alors

$$P(X) = (X - x_1)(X - x_2) = X^2 - (x_1 + x_2)X + x_1x_2.$$

Et donc par identification des coefficients, $a = -(x_1 + x_2)$ et $b = x_1x_2$.

Une application amusante de ces relations est la suivante : pour trouver deux nombres x_1 et x_2 tels que $\begin{cases} x_1 + x_2 = s \\ x_1x_2 = p \end{cases}$, il suffit de déterminer les racines de $X^2 - sX + p$ (ce qui pour des nombres réels n'est possible que si $s^2 - 4p > 0$).

Bien que l'exercice suivant relève plutôt de l'analyse, il fait apparaître des relations très importantes sur les polynômes, qui lient les racines et les coefficients.

Exercice 1.1 (ESCP 2009) [ESCP09.1.05] **Moyen**

Sommes de Newton
Abordable en première année : ✓
Thèmes du programme abordés : séries numériques, polynômes

1. Exemple introductif.
   (a) Montrer que pour tout n de $\mathbf{N}$, $\left(2+\sqrt{3}\right)^n + \left(2-\sqrt{3}\right)^n$ est un entier naturel.
   (b) En déduire que la série de terme général $\sin\left(\pi\left(2+\sqrt{3}\right)^n\right)$ est convergente.

Soit $P = X^3 + a_1X^2 + a_2X + a_3$ un polynôme unitaire de degré 3 à coefficients dans $\mathbf{N}$.
On suppose que P possède trois racines réelles x_1, x_2, x_3 telles que $|x_1| > 1, |x_2| < 1$ et $|x_3| < 1$.
On pose :

$$\sigma_1 = x_1 + x_2 + x_3,\ \sigma_2 = x_1x_2 + x_1x_3 + x_2x_3,\ \sigma_3 = x_1x_2x_3.$$

On pose enfin, pour tout $n \in \mathbf{N}^*$: $S_n = x_1^n + x_2^n + x_3^n$.

2. (a) Exprimer a_1, a_2, a_3 en fonction de $\sigma_1, \sigma_2, \sigma_3$.
   (b) Calculer $\sigma_1\sigma_2 - 3\sigma_3$. En déduire l'expression de S_1, S_2 et S_3 en fonction de $\sigma_1, \sigma_2, \sigma_3$.
   (c) Montrer que pour tout p de $\mathbf{N}^*$: $S_{p+3} - \sigma_1S_{p+2} + \sigma_2S_{p+1} - \sigma_3S_p = 0$.
   (d) En déduire que pour tout n de $\mathbf{N}^*$, S_n est un entier relatif.
3. Déterminer la nature de la série de terme général $\sin(\pi x_1^n)$.

Solution de l'exercice 1.1

1.a. D'après la formule du binôme, on a

$$\left(2+\sqrt{3}\right)^n+\left(2-\sqrt{3}\right)^n=\sum_{k=0}^{n}\binom{n}{k}2^{n-k}\left(\sqrt{3}\right)^k+\sum_{k=0}^{n}\binom{n}{k}2^{n-k}\left(-\sqrt{3}\right)^k=2\sum_{\substack{k=0\\ k\text{pair}}}^{n}\binom{n}{k}2^{n-k}\left(\sqrt{3}\right)^k.$$

Mais si $k=2p$ est pair, $\left(\sqrt{3}\right)^k=3^p$ est un entier, et donc $\left(2+\sqrt{3}\right)^n+\left(2-\sqrt{3}\right)^n$ est une somme d'entiers positifs, et donc est dans $\mathbf{N}$.
Mieux : c'est même un entier pair.

1.b. Notons que $\left(2+\sqrt{3}\right)^n$ tend vers $+\infty$ lorsque $n\to+\infty$, puisque $2+\sqrt{3}>1$.
En revanche, $0<2-\sqrt{3}<1$ et donc $\left(2-\sqrt{3}\right)^n\underset{n\to+\infty}{\longrightarrow}0$.
Mais d'après la question précédente,

$$\sin\left(X\pi\left(2+\sqrt{3}\right)^n\right)=\sin\left(\underbrace{\pi\left(\left(2+\sqrt{3}\right)^n+\left(2-\sqrt{3}\right)^n\right)}_{\in 2\pi\mathbf{Z}}-\pi\left(2-\sqrt{3}\right)^n\right)=-\sin\left(\pi\left(2-\sqrt{3}\right)^n\right).$$

Et puisque $\left(2-\sqrt{3}\right)^n\underset{n\to+\infty}{\longrightarrow}0$, on en déduit que

$$-\sin\left(\pi\left(2-\sqrt{3}\right)^n\right)\underset{n\to+\infty}{\sim}-\pi\left(2-\sqrt{3}\right)^n.$$

Par critère de comparaison pour les séries de signe constant, puisque la série[1] de terme général $-\pi\left(2-\sqrt{3}\right)^n$ converge, il en est de même de la série de terme général $\sin\left(\pi\left(2+\sqrt{3}\right)^n\right)$.

[1] Géométrique.

2.a. On a $X^3+a_1X^2+a_2X+a_3=(X-x_1)(X-x_2)(X-x_3)$.
Or en développant ce dernier produit, il vient

$$P(X)=X^3-(x_1+x_2+x_3)X+(x_1x_2+x_1x_3+x_2x_3)X-x_1x_2x_3=X-\sigma_1X^2+\sigma_2X-\sigma_3.$$

Et donc par identification des coefficients de P,

$$a_1=-\sigma_1,\ a_2=\sigma_2,\ a_3=-\sigma_3.$$

2.b. On a

$$\begin{aligned}\sigma_1\sigma_2-3\sigma_3&=(x_1+x_2+x_3)(x_1x_2+x_1x_3+x_2x_3)-3x_1x_2x_3\\&=x_1^2x_2+x_1^2x_3+x_1x_2x_3+x_1x_2^2+x_1x_2x_3+x_2^2x_3+x_1x_2x_3+x_1x_3^2+x_2x_3^2-3x_1x_2x_3\\&=x_1^2x_2+x_1x_2^2+x_1x_3^2+x_2^2x_3+x_2x_3^2\\&=(x_1^2+x_2^2+x_3^2)(x_1+x_2+x_3)-x_1^3-x_2^3-x_3^3\\&=S_1S_2-S_3.\end{aligned}$$

On a déjà $S_1=\sigma_1$.
De plus, $\sigma_1^2=x_1^2+x_2^2+x_3^2+2(x_1x_2+x_1x_3+x_2x_3)=S_2+2\sigma_2$.
Donc $S_2=\sigma_1^2-2\sigma_2$.
Et enfin, grâce au calcul réalisé au début de la question,

$$S_3=S_1S_2-\sigma_1\sigma_2+3\sigma_3=\sigma_1(\sigma_1^2-2\sigma_2)-\sigma_1\sigma_2+3\sigma_3=\sigma_1^3-3\sigma_1\sigma_2+3\sigma_3.$$

2.c. On pourrait faire un calcul direct et constater que tous les termes s'annulent.
Mais notons plutôt que pour tout $i\in[\![1,3]\!]$, $P(x_i)=0$.
Soit encore

$$x_i^p(x_i^3+a_1x_i^2+a_2x_i+a_3)=0\Leftrightarrow x_i^p(x_i^3-\sigma_1x_i^2+\sigma_2x_i-\sigma_3)=0.$$

En sommant ces trois relations, on obtient donc

$$S_{p+3}-\sigma_1S_{p+2}+\sigma_2S_{p+1}-\sigma_3S_p=0.$$

2.d. Puisque P est à coefficients dans $\mathbf{N}$, les a_i sont des entiers. Et par les relations de 2.a, $\sigma_1, \sigma_2, \sigma_3$ sont également des entiers.
Mais alors, les expressions de S_1, S_2, S_3 obtenues en 2.b prouvent que S_1, S_2, S_3 sont également des entiers.
Prouvons alors par une récurrence triple que S_p est un entier.
Pour $p \in \{0, 1, 2\}$, nous venons de le prouver.
Supposons que S_p, S_{p+1} et S_{p+2} sont des entiers.
Alors

$$S_{p+3} = \sigma_1 S_{p+2} - \sigma_2 S_{p+1} + \sigma_3 S_p \in \mathbf{Z}.$$

Et donc par le principe de récurrence, pour tout $p \in \mathbf{N}$, S_p est un entier relatif.

3. Pour tout $n \in \mathbf{N}$, on a

$$|\sin(\pi x_1^n)| = \left|\sin\left(\pi(\underbrace{S_n}_{\in\mathbf{Z}} - x_2^n - x_3^n)\right)\right| = \left|\sin\left(\pi(-x_2^n - x_3^n)\right)\right|.$$

Mais $x_2^n + x_3^n \underset{n\to+\infty}{\longrightarrow} 0$ car $|x_2| < 1$ et $|x_3| < 1$.
Et donc $|\sin(\pi x_1^n)| \underset{n\to+\infty}{\sim} \pi |x_2^n + x_3^n|$.
Or, par l'inégalité triangulaire, $0 \leqslant |x_2^n + x_3^n| \leqslant |x_2|^n + |x_3|^n$.
Mais la série de terme général $|x_2|^n + |x_3|^n$ est convergente car somme de deux séries géométriques convergentes, et donc par comparaison de séries à termes positifs, la série de terme général $|\sin(\pi x_1^n)|$ est convergente.
Et donc $\sum \sin(\pi x_1^n)$ est absolument convergente, donc convergente.

□

Valeur absolue
Si k est un entier,

$$\sin(x + k\pi) = \pm\sin(x)$$

suivant la parité de k.
C'est la raison pour laquelle nous utilisons des valeurs absolues, afin de ne pas avoir à distinguer suivant la partié de l'entier S_n.

Exercice 1.2 (QSP HEC 2012) [HEC12.QSP16] — Facile

Divisibilité de polynômes complexes
Abordable en première année : ✓

Soit $n \in \mathbf{N}$ et P_n le polynôme de $\mathbf{C}_n[X]$ défini par $P_n(X) = X^n + 1$.
Pour quelles valeurs de n, P_n est-il divisible par $X^2 + 1$?

Solution de l'exercice 1.2 Nous savons que $X^2 + 1 = (X - i)(X + i)$.
Et donc P_n est divisible par $X^2 + 1$ si et seulement si i et $-i$ en sont racines.
Or, $P_n(i) = i^n + 1$. Les puissances de i sont $i, -1, -i, 1, i, -1, -i, \ldots$. Autrement dit :

$$i^n = \begin{cases} 1 & \text{si } n = 4k \\ i & \text{si } n = 4k+1 \\ -1 & \text{si } n = 4k+2 \\ -i & \text{si } n = 4k+3 \end{cases}$$

Par conséquent, $P_n(i) = 0$ si et seulement si n est de la forme $4k + 2, k \in \mathbf{N}$.
De même, on a $P_n(-i) = (-i)^n + 1 = (-1)^n i^n + 1$. Donc $P_n(-i) = 0$ si et seulement si $(-1)^n i^n = -1$.
Or, i^n est réel si et seulement si n est pair, et alors $(-1)^n = 1$.
Donc $P_n(-i) = 0 \Leftrightarrow i^n = -1 \Leftrightarrow n = 4k + 2$.
Et donc P_n est divisible par $X^2 + 1$ si et seulement si n est de la forme $4k + 2$, $k \in \mathbf{N}$. □

Exercice 1.3 (QSP HEC 2008) [HEC08.QSP11] — Moyen

Autour de la division euclidienne
Abordable en première année : ✓

Soit $n \geqslant 2$ un entier naturel. Déterminer tous les polynômes de $\mathbf{R}_n[X]$ divisibles par $X+1$ et dont les restes de la division euclidienne par $X+2, X+3, \cdots, X+n+1$ sont égaux.

SOLUTION DE L'EXERCICE 1.3 Notons que pour $k \in \mathbf{R}$, le reste de la division euclidienne d'un polynôme P par $X+k$ est une constante α_k et qu'alors, si $P = (X+k)Q + \alpha_k$, nécessairement $P(-k) = \alpha_k$.
Ainsi, nous cherchons les polynômes divisibles par $X+1$ qui prennent la même valeur en $-2, -3, \cdots, -n-1$.
Soit donc P un tel polynôme, et soit α la valeur commune de $P(-2) = P(-3) = \cdots = P(-n-1)$.
Alors $P - \alpha$ est un polynôme de degré au plus n qui s'annule en $-2, -3, \cdots, -n-1$: il est donc de la forme $P - \alpha = \lambda(X+2)(X+3)\cdots(X+n+1)$, avec $\lambda \in \mathbf{R}$.
Et puisque P est divisible par $X+1$, on a $P(-1) = 0 \Rightarrow \lambda(-1+2)(-1+3)\cdots(-1+n+1)+\alpha = 0$.
Soit encore $\lambda n! + \alpha = 0$. Ainsi, P est de la forme $P = \lambda\left((X+2)(X+3)\cdots(X+n+1) - n!\right)$.

Remarque
Le fait que P soit de degré au plus n est important, sans cela on pourrait uniquement affirmer que $P-\alpha$ est divisible par $(X+2)\cdots(X+n+1)$, mais le quotient pourrait alors être un polynôme non constant.

Inversement, si $P = \lambda\left((X+2)(X+3)\cdots(X+n+1) - n!\right)$, avec $\lambda \in \mathbf{R}$, alors $P(-1) = \lambda(n! - n!) = 0$, de sorte que P est divisible par $X+1$.
D'autre part, $P(-2) = P(-3) = \cdots = P(-n-1) = -\lambda n!$. Et donc les restes de la division euclidienne de P par $X+2, X+3, \cdots, X+n+1$ sont tous égaux à $-\lambda n!$
Ainsi, les polynômes satisfaisant aux conditions de l'énoncé sont exactement les polynômes de la forme $\lambda\left((X+2)(X+3)\cdots(X+n+1) - n!\right)$. □

1.1.1 Polynômes de Lagrange

L'exercice suivant montre qu'en utilisant les polynômes de Lagrange, il est possible «d'approcher» toute fonction par une fonction polynomiale, en ce sens qu'elle prend les mêmes valeurs en n points fixés à l'avance.
Dans le cas où la fonction de départ est de classe $\mathscr{C}^n$, on obtient même une majoration de l'erreur commise en l'approchant par un tel polynôme.

EXERCICE 1.4 (ESCP 2012) [ESCP12.2.13] **Moyen**

Polynômes d'interpolation de Lagrange et majoration de l'erreur
Abordable en première année : ✓

Dans tout cet exercice, n désigne un entier naturel non nul et $(a_1, \ldots, a_n)$ 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 $i \in [\![1,n]\!]$. Montrer qu'il existe un unique polynôme $L_i \in \mathbf{R}_{n-1}[X]$ tel que pour tout $j \in [\![1,n]\!]$, $L_i(a_j) = \delta_{i,j}$.
   (b) Montrer que la famille $(L_i)_{i\in[\![1,n]\!]}$ est une base de $\mathbf{R}_{n-1}[X]$.
2. Soit $\pi : \mathbf{R}[X] \to \mathbf{R}[X]$ 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 $(L_i)_{i\in[\![1,n]\!]}$.
3. Soit $\varepsilon : \mathbf{R}_{n-1}[X] \to \mathbf{R}^n$, $P \mapsto (P(a_i))_{i\in[\![1,n]\!]}$.
   (a) Montrer que ε est un isomorphisme.
   (b) Soit $f \in \mathscr{F}(\mathbf{R},\mathbf{R})$. Montrer qu'il existe un unique polynôme $P \in \mathbf{R}_{n-1}[X]$ tel que $P(a_i) = f(a_i)$, pour tout $i \in [\![1,n]\!]$.
   Ce polynôme s'appelle le polynôme d'interpolation de Lagrange associé à la fonction f aux points $(a_1, \ldots, a_n)$.

4. Soient $a, b \in \mathbf{R}$ tels que $a < b$. Soient $f \in \mathscr{C}^n([a,b], \mathbf{R})$, $a_1, a_2, \dots, a_n$ tels que $a \leqslant a_1 < \cdots < a_n \leqslant b$ 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 :
   $$\varphi : [a,b] \to \mathbf{R},\ t \mapsto f(t) - P(t) - K\prod_{i=1}^{n}(t - a_i).$$
   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 :
   $$|f(x) - P(x)| \leqslant \frac{\prod_{i=1}^{n}|x - a_i|}{n!}\sup_{[a,b]}|f^{(n)}|.$$

Solution de l'exercice 1.4

1.a. Supposons qu'un tel polynôme L_i existe. Alors les $a_j, 1 \leqslant j \leqslant n, j \neq i$ 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 $L_i = Q\prod_{\substack{j=1\\j\neq i}}^{n}(X - a_j)$.
Mais alors $\deg L_i = \deg Q + (n-1)$, et puisque $\deg L_i \leqslant n-1$, $\deg Q \leqslant 0$.
Autrement dit, Q est une constante λ.
Et alors
$$1 = L_i(a_i) = \lambda\prod_{\substack{j=1\\j\neq i}}^{n}(a_i - a_j) \Leftrightarrow \lambda = \frac{1}{\prod_{\substack{j=1\\j\neq i}}^{n}(a_i - a_j)}.$$

> **Remarque**
> Le fait que les a_k soient deux à deux distincts garantit que
> $$\prod_{\substack{j=1\\j\neq i}}^{n}(a_i - a_j) \neq 0.$$

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 égal[1] à $n-1$, et vérifie $\forall j \in [\![1,n]\!], L_i(a_j) = \delta_{i,j}$.
Donc il existe bien un unique polynôme de $\mathbf{R}_{n-1}[X]$ vérifiant ces conditions.

[1] En réalité, il est de degré **égal** à $n-1$.

1.b. Soient $\lambda_1, \dots, \lambda_n$ des réels tels que $\sum_{i=1}^{n}\lambda_i L_i = 0$.
Alors pour $j \in [\![1,n]\!]$, en évaluant cette relation en a_j, il vient
$$\sum_{i=1}^{n}\lambda_i L_i(a_j) = 0 \Leftrightarrow \sum_{i=1}^{n}\lambda_i\delta_{i,j} = 0 \Leftrightarrow \lambda_j = 0.$$
Et donc $\lambda_1 = \cdots = \lambda_n = 0$: 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 $\mathbf{R}_{n-1}[X]$.

2.a. Notons que pour tout $j \in [\![1,n]\!]$, 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 évidente[2], π est un projecteur de $\mathbf{R}[X]$.

[2] L'évaluation des polynômes en a_i est linéaire.

2.b. Pour $P \in \mathbf{R}[X]$, on a $P \in \operatorname{Ker}(\pi) \Leftrightarrow \sum_{i=1}^{n}P(a_i)L_i = 0$, et par liberté de la famille $(L_1, \dots, L_n)$, c'est le cas si et seulement si $P(a_1) = \cdots = P(a_n) = 0$.

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 $P = \pi(P) = \sum_{P=1}^{n} P(a_i)L_i$.
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

$$P(a_1) = P(a_2) = \cdots = P(a_n) = 0.$$

Donc $a_1, \dots, a_n$ sont n racines distinctes[3] 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] Car les a_i le sont.

Rappel
Le seul polynôme qui possède strictement plus de racines que son degré est le polynôme nul.

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

$$\varepsilon(P) = (f(a_1, \dots, f(a_n)) \Leftrightarrow \forall i \in [\![1, n]\!],\ P(a_i) = f(a_i).$$

4.a. On a $\varphi(x) = f(x) - P(x) - K\prod_{i=1}^{n}(x - a_i)$.
Et donc $\varphi(x) = 0 \Leftrightarrow K = \dfrac{f(x) - P(x)}{\prod_{i=1}^{n}(x - a_i)}$.

4.b. Pour tout $i \in [\![1, n]\!]$, $\varphi(a_i) = f(a_i) - P(a_i) - \prod_{j=1}^{n}(a_i - a_j) = f(a_i) - P(a_i) = 0$.
Donc φ s'annule en $a_1, \dots, a_n$ et en x, qui sont bien $n+1$ réel 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, $K = \dfrac{f^{(n)}(\zeta)}{n!}$.

Dérivée $n^{\text{ème}}$
Si besoin, dérivez n fois

$$a_n X^n + \cdots + a_1 + a_0$$

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.

Et donc

$$\begin{aligned}|f(x)-P(x)| &= \left|\underbrace{\varphi(x)}_{=0} + K\prod_{i=1}^{n}(x-a_i)\right| \\ &= |K|\prod_{i=1}^{n}|x-a_i| \\ &= \frac{|f^{(n)}(\zeta)|}{n!}\prod_{i=1}^{n}|x-a_i| \\ &\leqslant \sup_{x\in[a,b]}|f^{(n)}|\frac{\prod_{i=1}^{n}|x-a_i|}{n!}.\end{aligned}$$

□

1.1.2 Racines des polynômes à coefficients réels : algèbre ou analyse ?

Exercice 1.5 (QSP HEC 2015) [HEC15.QSP106] — Moyen

Somme des dérivées d'un polynôme positif
Abordable en première année : ✓

1. Soit P un trinôme tel que : $\forall x\in\mathbf{R}$, $P(x)\geqslant 0$. Montrer que pour tout $x\in\mathbf{R}$, $P(x)+P'(x)+P''(x)\geqslant 0$.
2. Plus généralement, soit P un polynôme de degré n ($n\in\mathbf{N}$) tel que $\forall x\in\mathbf{R}$, $P(x)\geqslant 0$.
   (a) Que peut-on dire de la parité de n ?
   (b) Montrer que pour tout $x\in\mathbf{R}$, $P(x)+P''(x)+\cdots+P^{(n)}(x)\geqslant 0$.

Solution de l'exercice 1.5

1. Notons $P(x)=ax^2+bx+c$. Alors la limite en $+\infty$ de P est $\pm\infty$ suivant le signe de a.
Puisque P ne prend que des valeurs positives par hypothèse, on a donc $\lim_{x\to+\infty}P(x)=+\infty$ et donc $a\geqslant 0$.
D'autre part, on a $P'(x)=2ax+b$ et $P''(x)=2a$.
Notons alors $g(x)=P(x)+P'(x)+P''(x)$. Alors g est deux fois dérivable, et $g'(x)=P'(x)+P''(x)$ et $g''(x)=P''(x)=2a\geqslant 0$.
Donc g est convexe. En particulier, la courbe représentative de g est située au dessus de ses tangentes. Mais $g'(x)=2ax+b+2a$ s'annule en $x=-1+\dfrac{b}{2a}$, et $g\left(-1+\dfrac{b}{2a}\right)=P\left(-1+\dfrac{b}{2a}\right)+g'\left(-1+\dfrac{b}{2a}\right)\times 0$.
Ainsi, la tangente à g en $x=-1+\dfrac{b}{2a}$ est la droite[1] d'équation $y=g\left(-1+\dfrac{b}{2a}\right)$.
Et donc pour tout $x\in\mathbf{R}$, $g(x)\geqslant g\left(-1+\dfrac{b}{2a}\right)\geqslant 0$.

Minimum
Nous venons en fait de prouver que g possède un minimum en $-1+\frac{b}{2a}$.
Le même raisonnement prouve qu'une fonction convexe qui admet un point critique possède un minimum en ce point critique.

2. Notons $P(x)=a_nx^n+a_{n-1}x^{n-1}+\cdots+a_1x+a_0$.
2.a. Au voisinage de $\pm\infty$, $P(x)\underset{x\to\pm\infty}{\sim}a_nx^n\underset{x\to\pm\infty}{\longrightarrow}\pm\infty$.
Si n est impair, alors les limites de a_nx^n en $\pm\infty$ sont de signes opposés, l'une égale à $+\infty$ et l'autre à $-\infty$.
Puisque P est positif, ceci n'est pas possible, et donc n est pair.
2.b. Notons $g(x)=P(x)+P'(x)+\cdots+P^{(n)}(x)$, et montrons que $\forall x\in\mathbf{R}$, $g(x)\geqslant 0$.
Raisonnons par l'absurde, et supposons que g prend des valeurs strictement négatives.
Notons que g est encore un polynôme de degré n, équivalent en $+\infty$ à P, et donc $\lim_{x\to+\infty}g(x)=+\infty$.
Si $a\in\mathbf{R}$ est tel que $g(a)<0$, alors par le théorème des valeurs intermédiaires[2], il existe donc $b>a$ tel que $g(b)=0$.
Et de même, $g(x)\underset{x\to-\infty}{\sim}P(x)$, et donc $\lim_{x\to-\infty}g(x)=+\infty$, de sorte qu'il existe $c<a$ tel que

[2] Qui s'applique car un polynôme est continu.

$g(c) = 0$.
Ainsi, le polynôme g possède au moins deux racines.
Notons $x_1 < x_2 < \cdots < x_r$ les différentes racines de g. Par le théorème des valeurs intermédiaires, g est de signe constant entre deux racines consécutives.
Soient donc x_i et x_{i+1} deux racines consécutives de g telles que $\forall x \in]x_i, x_{i+1}[$, $g(x) < 0$.
Alors, pour tout $x \in]x_i, x_{i+1}[$, on a

$$g'(x) = P'(x) + P''(x) + P^{(n)}(x) + \underbrace{P^{(n+1)}(x)}_{=0} = g(x) - P(x) < 0.$$

Et donc g est strictement décroissante sur $[x_i, x_{i+1}]$.
Ceci contredit le fait que $g(x_i) = g(x_{i+1}) = 0$. Donc notre hypothèse de départ est fausse, et donc pour tout $x \in \mathbf{R}$, $g(x) \geqslant 0$.
Une autre solution : en notant toujours que $g' = g - P$, posons $f(x) = g(x)e^{-x}$, de sorte que

$$f'(x) = g'(x)e^{-x} - g(x)e^{-x} = -P(x)e^{-x} \leqslant 0.$$

Alors f est décroissante sur $\mathbf{R}$. Mais $\lim\limits_{x\to+\infty} f(x) = 0$, de sorte que $\forall x \in \mathbf{R}$, $f(x) \geqslant 0$.
Et donc pour tout $x \in \mathbf{R}$, $g(x) \geqslant 0$.

□

Remarque

Si r est une racine de g, il se peut tout de même que g ne change pas de signe en r. Penser par exemple au polynôme X^2, qui s'annule en 0, mais n'y change pas de signe : il est positif sur $\mathbf{R}$.

EXERCICE 1.6 (HEC 2015) [HEC15.75] — Moyen

Une inégalité sur les coefficients d'un polyôme scindé à racines simples.
Abordable en première année : ✓

1. Soient a, b, c des réels et T le trinôme $T(X) = aX^2 + bX + c$. On note T' et T'' respectivement, les dérivées premières et seconde de la fonction T.
   (a) Donner une condition nécessaire et suffisante portant sur les réels a, b et c pour que, pour tout $x \in \mathbf{R}$, on ait : $T(x) \geqslant 0$.
   (b) On suppose que T possède deux racines réelles distinctes. Déduire de la question précédente que :

$$\forall x \in \mathbf{R},\ T(x)T''(x) \leqslant (T')^2(x).$$

Dans la suite de l'exercice, on note n un entier supérieur ou égal à 2, et P un polynôme de $\mathbf{R}[X]$ de degré n, ayant n racines réelles distinctes.
On pose $P = \sum_{k=0}^{n} a_k X^k$. On note P' et P'' respectivement, les dérivées première et seconde de P.

2. Montrer que P' possède $(n-1)$ racines réelles.
3. (a) Montrer que la fonction $x \mapsto \dfrac{P'(x)}{P(x)}$ est décroissante sur chaque intervalle de son ensemble de définition.
   (b) En déduire que pour tout $x \in \mathbf{R}$, on a : $P(x)P''(x) \leqslant (P')^2(x)$.
4. À l'aide des questions précédentes, établir pour tout $k \in [\![0, n-2]\!]$, l'inégalité : $a_k a_{k+2} \leqslant a_{k+1}^2$.

SOLUTION DE L'EXERCICE 1.6

1.a. Pour que T ne change pas de signe, il faut et il suffit que son discriminant soit négatif ou nul.
Et pour que le signe de T soit positif, il faut alors que $a \geqslant 0$.
Ainsi, T est positif sur $\mathbf{R}$ tout entier si et seulement si $\begin{cases} a > 0 \\ b^2 - 4ac \leqslant 0 \end{cases}$.

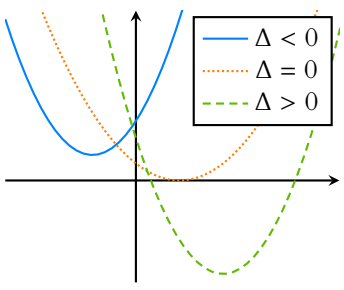


FIGURE 1.1– Un polynôme de degré 2 de signe constant est de discriminant négatif.

1.b. On a $T'(x) = 2ax + b$ et donc $(T')^2(x) = 4a^2x^2 + 4abx + b^2$.
D'autre part, $T''(x) = 2a$ et donc $T(x)T''(x) = 2a^2x^2 + 2abx + 2ac$.
Ainsi,

$$(T')^2(x) - T(x)T''(x) = 2a^2x^2 + 2abx + b^2 - 2ac.$$

Le discriminant de ce polynôme est

$$\Delta = 4a^2b^2 - 8a^2(b^2 - 2ac) = -4a^2b^2 + 16a^3c = -4a^2(b^2 - 4ac).$$

Mais puisque T possède deux racines réelles, $b^2 - 4ac > 0$ et donc $\Delta \leqslant 0$.
D'autre part, le coefficient dominant de $(T')^2 - TT''$ est $2a^2 > 0$, donc il vérifie les conditions de la première question :

$$\forall x \in \mathbf{R},\ (T')^2(x) - T(x)T''(x) \geqslant 0 \Leftrightarrow T(x)T''(x) \leqslant (T')^2(x).$$

2. Notons $x_1 < x_2 < \cdots < x_n$ les racines réelles de T, de sorte que $T(x_1) = T(x_2) = \cdots = T(x_n) = 0$.
Puisque T est dérivable sur $\mathbf{R}$, par le théorème de Rolle, T' s'annule au moins une fois sur chacun des intervalles $]x_i, x_{i+1}[$, $i \in [\![1, n-1]\!]$.
Et donc T' possède au moins $n-1$ racines réelles. Étant de degré $n-1$, il a au plus $n-1$ racines, et donc possède exactement $n-1$ racines.

3.a. Avec les notations précédentes, on a $P(X) = a\prod_{i=1}^{n}(X - x_i)$, où a est le coefficient dominant de P.
Or, $x \mapsto \dfrac{P'(x)}{P(x)}$, est la dérivée, sur chacun des intervalles $]-\infty, x_1[,]x_1, x_2[, \ldots,]x_{n-1}, x_n[,]x_n, +\infty[$, de la fonction $x \mapsto \ln(|P(x)|)$.

Or, pour $x \notin \{x_1, x_2, \ldots, x_n\}$, $\ln(|P(x)|) = \ln(|a|) + \sum_{k=1}^{n} \ln(|x - x_k|)$.

Et donc sa dérivée est $x \mapsto \sum_{k=1}^{n} \dfrac{1}{x - x_k}$.

Mais les fonctions $x \mapsto \dfrac{1}{x - x_k}$ sont toutes décroissantes sur leurs ensembles de définition, et donc $x \mapsto \dfrac{P'(x)}{P(x)}$ est décroissante car somme de fonctions décroissantes.

Danger !
Attention à ne pas oublier la valeur absolue, P n'est pas de signe constant puisqu'il change de signe en chacun des x_i.

Astuce
Le fait que pour un polynôme **scindé** de racines $x_1, \ldots, x_n$, on ait

$$\frac{P'(x)}{P(x)} = \sum_{k=1}^{n} \frac{1}{x - x_k}$$

est assez classique.
On parle alors de la dérivée logarithmique de P.

3.b. La fonction $x \mapsto \dfrac{P'(x)}{P(x)}$ est dérivable sur son ensemble de définition[1], de dérivée égale à $x \mapsto \dfrac{P(x)P''(x) - P'(x)^2}{P''(x)^2}$.
Or, par la question précédente, cette dérivée est négative, donc pour tout $x \notin \{x_1, x_2, \ldots, x_n\}$, on a

$$P(x)P''(x) - (P')^2(x) \leqslant 0 \Leftrightarrow P(x)P''(x) \leqslant (P')^2(x).$$

D'autre part, si $x \in \{x_1, x_2, \ldots, x_n\}$, alors l'inégalité est triviale puisque $\underbrace{P(x)}_{=0} P''(x) = 0$ et $(P')^2(x) \geqslant 0$.

[1] Qui est $\mathbf{R} \setminus \{x_1, x_2, \ldots, x_n\}$.

4. Évaluons l'égalité précédente en $x = 0$: $P(0)P''(0) \leqslant P'(0)^2$.
Or, $P(0) = a_0$, $P'(0) = a_1$ et $P''(0) = 2a_2$.
On a donc $2a_0a_2 \leqslant a_1^2$ et donc $a_0a_2 \leqslant a_1^2$.

De même, on a $P'(X) = \sum_{k=1}^{n} k a_k X^{k-1}$, et P' possède $n-1$ racines réelles d'après la question 2.
Par conséquent, le résultat de la question 3 s'applique : $\forall x \in \mathbf{R}$, $P'(x)P^{(3)}(x) \leqslant (P'')^2(x)$.
Mais $P^{(3)}(0) = 6a_3$, et donc $P'(0)P^{(3)}(0) \leqslant (P'')^2(0)$ soit encore

$$a_1 6a_3 \leqslant (2a_2)^2 \Rightarrow a_1a_3 \leqslant \frac{4}{6}a_2^2 \leqslant a_2^2.$$

Plus généralement, pour tout $k \in [\![0, n-2]\!]$ une application répétée de la question 2 montre que $P^{(k)}$ possède $n-k$ racines.
Et donc le résultat de la question 3 s'applique (en particulier avec $x = 0$) :

$$P^{(k)}(0)P^{(k+2)}(0) \leqslant (P^{(k+1)})^2(0).$$

Or, pour tout $i \leqslant n$, $P^{(i)}(X) = \sum_{k=i}^{n} a_k k(k-1)\cdots(k-i+1)X^{k-i}$, et donc $P^{(i)}(0) = a_i i!$.

Et donc la relation $P^{(k)}(0)P^{(k+2)}(0) \leqslant (P^{(k+1)})^2(0)$ s'écrit encore

$$k!a_k \times (k+2)!a_{k+2} \leqslant ((k+1)!a_{k+1})^2 \Leftrightarrow a_k a_{k+2} \leqslant a_{k+1}^2 \frac{((k+1)!)^2}{k!(k+2)!}$$

Mais $\dfrac{((k+1)!)^2}{k!(k+2)!} = \dfrac{k! \times (k+1) \times (k+1)!}{k! \times (k+1)! \times (k+2)} = \dfrac{k+1}{k+2} \leqslant 1$.

Et donc $a_k a_{k+2} \leqslant a_{k+1}^2$.

□

Taylor

Notons que le fait que $P^{(i)}(0) = i!\, a_i$ peut se retrouver plus simplement à l'aide de la formule de Taylor pour les polynômes :

$$P(X) = \sum_{i=0}^{n} \frac{P^{(i)}(0)}{i!} X^i$$

par identification des coefficients.

1.2 ESPACES VECTORIELS

EXERCICE 1.7 (ESCP 2016) [ESCP16.2.12] Facile

Base des polynômes factoriels et application au calcul de sommes de séries.
Abordable en première année : ✓

On note $\mathbf{R}[X]$ l'espace vectoriel des polynômes à coefficients réels et pour tout entier naturel n, $\mathbf{R}_n[X]$ désigne l'espace vectoriel des polynômes de degré inférieur ou égal à n.
On pose $S_0 = 1$ et pour tout entier k non nul, on désigne par S_k le polynôme défini par :

$$S_k(X) = \prod_{i=0}^{k-1}(X-i).$$

1. (a) Démontrer que, pour tout entier n, la famille $(S_k)_{0\leqslant k\leqslant n}$ est une famille libre de $\mathbf{R}[X]$.
   (b) Soit m un entier naturel. Prouver que pour tout polynôme P de $\mathbf{R}_m[X]$, il existe un unique $(m+1)$-uplet de réels $(a_k)_{0\leqslant k\leqslant m}$ tels que $P = \sum_{k=0}^{m} a_k S_k$.
2. (a) Pour tout entier naturel n et tout entier naturel k, calculer $S_k(n)$.
   (b) Soit P un polynôme de $\mathbf{R}_m[X]$ écrit dans la base $(S_k)_{0\leqslant k\leqslant m}$ sous la forme $P = \sum_{k=0}^{m} a_k S_k$.
      i. Démontrer que pour tout entier $N \geqslant m$, $\sum_{n=0}^{N} \frac{P(n)}{n!} = \sum_{k=0}^{m} a_k \sum_{n=k}^{N} \frac{1}{(n-k)!}$.
      ii. En déduire que la série $\sum_{n\geqslant 0} \frac{P(n)}{n!}$ converge et calculer sa somme en fonction des a_k, $k \in [\![0, m]\!]$.
   (c) Soit p un entier naturel non nul.
   Justifier la convergence de la série $\sum_{n\geqslant 0} \frac{n^2(n-1)(n-2)\cdots(n-p+1)}{n!}$ et calculer sa somme.

SOLUTION DE L'EXERCICE 1.7

1.a. Notons que S_k est le produit de k termes de degré 1, et donc est de degré k.
Ainsi, $(S_k)_{0\leqslant k\leqslant n}$ est une famille de polynômes de degrés deux à deux distincts : c'est donc une famille libre de $\mathbf{R}[X]$.

1.b. La famille $(S_k)_{0\leqslant k\leqslant n}$ est une famille libre de $n+1$ polynômes, tous dans $\mathbf{R}_n[X]$.
Puisque $\dim \mathbf{R}_n[X] = n+1$, c'est donc une base de $\mathbf{R}_n[X]$.
Et par conséquent, tout polynôme P de $\mathbf{R}_n[X]$ se décompose de manière unique dans cette base : il existe un unique $(n+1)$-uplet $(a_0, a_1, \ldots, a_n)$ de réels tels que $P = \sum_{k=0}^{n} a_k S_k$.

2.a. Si $k > n$, alors $X - n$ est l'un des facteurs de S_k, et donc n est racine de S_k, de sorte que $S_k(n) = 0$.

Rappel

Une famille de vecteurs d'un espace vectoriel E qui est de cardinal $\dim E$ est automatiquement une base, il n'y a pas besoin de vérifier qu'elle est génératrice. C'est d'ailleurs souvent le plus simple pour montrer qu'une famille est une base...si on connait la dimension de E.

Si $k \leqslant n$, alors

$$S_k(n) = \prod_{i=0}^{k-1}(n-i) = n(n-1)\cdots(n-k+1) = \frac{n!}{(n-k)!}.$$

2.b.i. D'après ce qui précède, on a

$$\begin{aligned}\sum_{n=0}^{N}\frac{P(n)}{n!} &= \sum_{n=0}^{N}\frac{1}{n!}\sum_{k=0}^{m}a_kS_k(n)\\ &= \sum_{k=0}^{m}a_k\sum_{n=0}^{N}\frac{S_k(n)}{n!}\\ &= \sum_{k=0}^{m}a_k\sum_{n=k}^{N}\frac{n!}{(n-k)!}\frac{1}{n!}\\ &= \sum_{k=0}^{m}a_k\sum_{n=k}^{N}\frac{1}{(n-k)!}.\end{aligned}$$

Si $n < k$, alors $S_k(n) = 0$.

2.b.ii. Pour $k \in [\![0, m]\!]$ fixé, on a

$$\sum_{n=k}^{N}\frac{1}{(n-k)!} = \sum_{i=0}^{N-k}\frac{1}{i!}.$$

Nous reconnaissons là une somme partielle de la série exponentielle de paramètre 1, et donc lorsque $N \to +\infty$,

$$\sum_{n=k}^{N}\frac{1}{(n-k)!} \underset{N\to+\infty}{\longrightarrow} \sum_{i=0}^{+\infty}\frac{1}{i!} = e^1 = e.$$

Et donc, par somme de limites, on a

$$\sum_{n=0}^{N}\frac{P(n)}{n!} \underset{N\to+\infty}{\longrightarrow} \sum_{k=0}^{m}a_k e.$$

Ceci signifie donc que la série $\displaystyle\sum_{n\geqslant 0}\frac{P(n)}{n!}$ converge et que

$$\sum_{n=0}^{+\infty}\frac{P(n)}{n!} = e\sum_{k=0}^{m}a_k.$$

Convergence

Rappelons que, **par définition**, une série converge si et seulement si la suite de ses sommes partielles admet une limite finie.
Cette limite est alors (toujours par définition) la somme de la série.

2.c. Soit P le polynôme défini par $P(X) = X^2(X-1)(X-2)\cdots(X-p+1) = XS_p$.

Alors, $P \in \mathbf{R}_{p+1}[X]$, et donc il existe des réels $a_0, \ldots, a_{p+1}$ tels que $P = \displaystyle\sum_{k=0}^{p+1}a_kS_k$.

En évaluant cette relation en $X = 0$, il vient donc

$$0 = P(0) = \sum_{k=0}^{p+1}a_kS_k(0) = a_0.$$

Et donc $P = \displaystyle\sum_{k=1}^{p+1}a_kS_k$. Puis en évaluant en $X = 1$, il vient

$$0 = P(1) = \sum_{k=1}^{p+1}a_kS_k(1) = a_1S_1(1) = a_1.$$

De proche en proche, en utilisant le fait que les racines de P sont $0, 1, \ldots, p-1$, on montre ainsi que $a_0, a_1, \ldots, a_{p-1}$ sont nuls.
Et donc $P = a_pS_p + a_{p+1}S_{p+1}$.

De plus, on a $P(p) = p \times p! = a_p S_p(p) + a_{p+1} \underbrace{S_{p+1}(p)}_{=0} = a_p p!$, de sorte que $a_p = p$.

Enfin, en évaluant en $X = p+1$, il vient,

$$P(p+1) = (p+1)^2 \times p \times \cdots \times 2 = (p+1) \times (p+1)! = p \underbrace{S_p(p+1)}_{=(p+1)!} + a_{p+1} \underbrace{S_{p+1}(p+1)}_{(p+1)!} = (p+a_{p+1})(p+1)!.$$

Et donc $a_{p+1} = 1$, de sorte que la décomposition de P dans la base $(S_0, \ldots, S_{p+1})$ est :

$$P = pS_p + S_{p+1}.$$

Et alors, en appliquant le résultat de la question 2.b.ii, la série $\sum_{n\geqslant 0} \frac{P(n)}{n!}$ converge et

$$\sum_{n=0}^{+\infty} \frac{P(n)}{n!} = \sum_{n=0}^{+\infty} \frac{n^2(n-1)\cdots(n-p+1)}{n!} = (p+1)e.$$

□

Alternative

Avec un peu d'astuce, ce résultat pouvait s'obtenir avec bien moins de calculs :

$$\begin{aligned} P &= XS_p \\ &= (X-p)S_p + pS_p \\ &= S_{p+1} + pS_p. \end{aligned}$$

EXERCICE 1.8 (QSP HEC 2013) [HEC13.QSP62]

Soit $E = \mathbf{R}_3[X]$ l'espace vectoriel des polynômes à coefficients réels de degré inférieur ou égal à 3. On pose :

$F = \{P \in E : P(0) = P(1) = P(2) = 0\}$, $G = \{P \in E : P(1) = P(2) = P(3) = 0\}$ et $H = \{P \in E : P(X) = P(-X)\}$.

Montrer que $E = F \oplus G \oplus H$.

SOLUTION DE L'EXERCICE 1.8 Soit $P \in E$. Alors $P \in F$ si et seulement si P est divisible par $X(X-1)(X-2)$.
Mais un tel polynôme s'écrit alors $P = QX(X-1)(X-2)$, et P étant de degré inférieur ou égal à 3, nécessairement, $\deg Q = 0$, c'est-à-dire que Q est une constante.
Ainsi, $F = \{\lambda X(X-1)(X-2),\ \lambda \in \mathbf{R}\} = \text{Vect}\,(X(X-1)(X-2))$.

On prouve de même que $G = \text{Vect}\,((X-1)(X-2)(X-3))$.

Enfin, soit $P = a_0 + a_1X + a_2X^2 + a_3X^3$. Alors

$$\begin{aligned} P \in H &\Leftrightarrow a_0 + a_1X + a_2X^2 + a_3X^3 = a_0 + a_1(-X) + a_2(-X)^2 + a_3(-X)^3 \\ &\Leftrightarrow a_0 + a_1X + a_2X^2 + a_3X^3 = a_0 - a_1X + a_2X^2 - a_3X^3 \\ &\Leftrightarrow a_1 = a_3 = 0 \\ &\Leftrightarrow P = a_0 + a_2X^2. \end{aligned}$$

Rappel

Deux polynômes sont égaux si et seulement si leurs coefficients sont égaux.

Autrement dit, $H = \{a_0 + a_2X^2, (a_0, a_2) \in \mathbf{R}^2\} = \text{Vect}(1, X^2)$.

Remarquons que $\dim F + \dim G + \dim H = 1 + 1 + 2 = 4 = \dim E$, ce qui est nécessaire pour avoir $E = F \oplus G \oplus H$.
En concaténant les bases obtenues précédemment de F, G et H, on obtient la famille $(X(X-1)(X-2), (X-1)(X-2)(X-3), 1, X^2)$ de E.
Montrons que cette famille est libre : soient $\lambda_1, \lambda_2, \lambda_3, \lambda_4$ des réels tels que

$$\lambda_1 X(X-1)(X-2) + \lambda_2(X-1)(X-2)(X-3) + \lambda_3 + \lambda_4 X^2 = 0.$$

En évaluant en $X = 1$ et $X = 2$, il vient $\lambda_3 + \lambda_4 = 0$ et $\lambda_3 + 4\lambda_4 = 0$.
Ainsi, $\lambda_3 = \lambda_4 = 0$.
Puis en évaluant en $X = 0$, il vient $-6\lambda_2 = 0$, et donc $\lambda_1 X(X-1)(X-2) = 0 \Leftrightarrow \lambda_1 = 0$.
Par conséquent, $(X(X-1)(X-2), (X-1)(X-2)(X-3), 1, X^2)$ est une famille libre de E, de cardinal $4 = \dim E$: c'est une base de E.
Et puisque la concaténation de bases de F, G et H est une base de E, ces trois sous-espaces

Danger !

Rappelons que ceci est nécessaire pour que la somme soit directe, mais n'est en aucun cas suffisant !

sont en somme directe, et $E = F \oplus G \oplus H$.

Remarque *: l'ensemble H est l'ensemble des polynômes pairs de $\mathbf{R}_2[X]$.*
On pourrait prouver comme nous l'avons fait ici que l'ensemble des polynômes pairs de $\mathbf{R}_n[X]$ a pour base $1, X^2, X^4$,etc, c'est-à-dire l'ensemble des X^k, avec k pair compris entre 0 et n.
Et de même, $\{P \in \mathbf{R}_n[X] : P(-X) = -P(X)\}$ l'ensemble des polynômes impairs de $\mathbf{R}_n[X]$ est engendré par X, X^3, X^5, etc, l'ensemble des X^k, avec k impair compris entre 1 et n. □

Exercice 1.9 (QSP HEC 2006) [HEC06.QSP01]

Soient A, B, C trois matrices de $\mathcal{M}_2(\mathbf{C})$. Montrer qu'il existe trois complexes α, β, γ non tous nuls tels que $\alpha A+\beta B+\gamma C$ possède une unique valeur propre.

Solution de l'exercice 1.9 Si la famille (A, B, C) est liée, alors il existe trois réels α, β, γ non tous nuls tels que $\alpha A + \beta B + \gamma C = 0$. Or la matrice nulle ne possède qu'une valeur propre : 0.

Supposons donc que (A, B, C) est libre. Alors $F = \text{Vect}(A, B, C)$ est de dimension 3.
Soit $G = \left\{ \begin{pmatrix} a & b \\ 0 & a \end{pmatrix} \in \mathcal{M}_2(\mathbf{C}),\ (a, b) \in \mathbf{C}^2 \right\}$. Alors G est formé de matrices ne possédant qu'une seule valeur propre, et est de dimension 2.
Si on avait $F \cap G = \{0\}$, alors F et G seraient en somme directe.
Mais alors $\dim(F \oplus G) = 5 \geqslant \dim \mathcal{M}_2(\mathbf{C})$, ce qui n'est pas possible.
On en déduit donc qu'il existe une matrice non nulle dans $F \cap G$, qui est donc une matrice possédant une seule valeur propre et combinaison linéaire[1] de A, B et C.
Commentaire *: le raisonnement que nous avons utilisé sur les dimensions se généralise très facilement : si F et G sont deux sous-espaces vectoriels d'un espace vectoriel E de dimension finie, et si $\dim F + \dim G > \dim E$, alors $F \cap G \neq \{0_E\}$.* □

Rappel
La dimension d'une somme directe est la somme des dimensions.

[1] Dont les coefficients sont nécessairement non tous nuls puisqu'il ne s'agit pas de la matrice nulle.

1.3 Applications linéaires, endomorphismes

L'exercice qui suit est un grand classique, et on est surpris de le voir figurer à l'oral d'HEC. Il nécessite néanmoins une bonne compréhension des projecteurs.

Exercice 1.10 (QSP HEC 2017) [HEC17.QSP213] Facile

Une condition nécessaire et suffisante pour commuter à un projecteur.

Soit E un espace vectoriel sur $\mathbf{R}$, p un projecteur de E et $u \in \mathcal{L}(E)$. Montrer que p et u commutent si et seulement si $\text{Ker}\, p$ et $\text{Im}\, p$ sont stables par u.

Solution de l'exercice 1.10 Il est classique que si u et p commutent, alors $\text{Ker}p$ et $\text{Im}p$ sont stables par u.
Inversement, supposons que ces deux espaces soient stables par u. 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 $u \circ p(x) = u(p(x_1) + p(x_2)) = u(x_1)$ et $p \circ u(x) = p(u(x_1)) + p(u(x_2))$.
Mais $\text{Ker}p$ est stable par u, de sorte que $u(x_2) \in \text{Ker}p$, et donc $p(u(x_2)) = 0$.
De même, $\text{Im}p$ est stable, donc $u(x_1) \in \text{Im}p$. Et nous savons que pour tout élément y dans $\text{Im}p$, $p(y) = y$. Donc $p \circ u(x) = p(u(x_1)) = u(x_1)$.
Ainsi, $\forall x \in E$, $p \circ u(x) = u \circ p(x)$, donc $p \circ u = u \circ p$: u et p commutent.

Rappel
Pour un projecteur p, si $x \in \text{Im}\, p$ alors $p(x) = x$. Ceci tient au fait que $\text{Im}\, p = E_1(p)$.

Remarque *: souvenons nous qu'un projecteur est diagonalisable, possède 0 et 1 comme valeurs propres et $E_0(p) = \text{Ker}(p)$ et $E_1(p) = \text{Im}(p)$.*
Nous venons donc de prouver que u commute avec p si et seulement si les sous-espaces propres de p sont stables par u.

On pourrait prouver plus généralement que pour f diagonalisable, g commute avec f si et seulement si les sous-espaces propres de f sont stables par g. □

EXERCICE 1.11 (QSP ESCP 2012) [ESCP12.QSP04] — Moyen

Étude d'une suite de polynôme.
Abordable en première année : ✓

Soient λ, μ deux réels avec $\lambda \neq 0$.
On considère la suite (P_n) de polynômes définie par $P_0 \in \mathbf{R}_2[X]$ et $\forall n \in \mathbf{N}$, $P_{n+1} = \lambda P_n + \mu P_n'$.

1. Montrer que $(P_n)_n$ est une suite d'éléments de $\mathbf{R}_2[X]$.
2. Soit $n \in \mathbf{N}$ fixé et $Q \in \mathbf{R}_2[X]$. Existe-t-il $P_0 \in \mathbf{R}_2[X]$ tel que $P_n = Q$?

SOLUTION DE L'EXERCICE 1.11

1. Montrons le par récurrence sur n. Par hypothèse, $P_0 \in \mathbf{R}_2[X]$.
Supposons que $P_n \in \mathbf{R}_2[X]$. Alors $\deg(\lambda P_n) \leqslant 2$ et $\deg(\mu P_n') \leqslant 1$, de sorte que $\deg(P_{n+1}) \leqslant 2$.
Ainsi, $P_{n+1} \in \mathbf{R}_2[X]$, et donc pour tout $n \in \mathbf{N}$, $P_n \in \mathbf{R}_2[X]$.

2. Il serait sûrement possible d'expliciter les coefficients de P_n en fonction de ceux de P_0, et de résoudre ensuite un système d'équations ayant pour inconnues les trois coefficients de P_0.
Mais ceci serait fastidieux, et il est plus simple de remarquer que $f : P \mapsto \lambda P + \mu P'$ est un endomorphisme de $\mathbf{R}_2[X]$, avec $P_n = f^n(P_0)$.
Il s'agit alors de déterminer si Q admet un antécédent par f^n.
Mais la matrice de f dans la base canonique de $\mathbf{R}_2[X]$ est

$$\mathrm{Mat}_{\mathcal{B}_{can}}(f) = \begin{array}{c} \begin{matrix} f(1) & f(X) & f(X^2) \end{matrix} \\ \begin{pmatrix} \lambda & \mu & 0 \\ 0 & \lambda & 2\mu \\ 0 & 0 & \lambda \end{pmatrix} \end{array} \begin{matrix} 1 \\ X \\ X^2 \end{matrix}.$$

Cette matrice est de rang 3 car $\lambda \neq 0$, et donc elle est inversible, de sorte que f est un isomorphisme.
Donc f^n est également un isomorphisme[1] et donc $Q = f^n(P_0)$ possède une unique solution : $P_0 = (f^n)^{-1}(Q)$.

[1] Une composée d'isomorphismes est un isomorphisme.

□

1.4 ENDOMORPHISMES ET MATRICES NILPOTENTS

TODO : escp 2006 2.08
Un endomorphisme f (respectivement une matrice carrée A) est dit nilpotent s'il existe un entier $p \in \mathbf{N}$ tel que $f^p = 0$ (resp. $A^p = 0$).
Aucune connaissance n'est exigible sur le sujet, mais il s'agit tout de même d'un thème récurrent de l'oral.
L'exercice suivant regroupe la plupart des questions classiques sur le sujet (la première étant très très classique, et il est difficile de deviner la méthode si on ne l'a jamais vue).
Si tout est formulé ici en termes d'endomorphismes, ces résultats se transposent aisément au cas des matrices nilpotentes.

EXERCICE 1.12 (ESCP 2017) [ESCP17.2.08] — Moyen

f est nilpotent si et seulement si $\mathrm{Spec}_{\mathbf{C}}(f) = \{0\}$.

Soit n un entier naturel $n \geqslant 2$ et E un espace vectoriel complexe de dimension n. Soit f un endomorphisme non nul de E nilpotent, ce qui signifie qu'il existe $p \geqslant 2$ tel que $f^p = 0$ et $f^{p-1} \neq 0$.

1. Montrer que $p \leqslant n$.
2. On suppose dans cette question uniquement que $p = n$. Résoudre l'équation $u^2 = f$, d'inconnue u endomorphisme de E.

3. Déterminer les valeurs propres de f. L'endomorphisme f est-il diagonalisable ?
4. Soit g un endomorphisme non nul de E.
   (a) Montrer qu'il existe un polynôme Q annulateur de g dont toutes les racines sont exclusivement les valeurs propres de g.
   (b) Montrer que dans l'ensemble des polynômes annulateurs non nuls de g, il existe un polynôme annulateur de degré minimal. Quel lien existe-t-il entre Q et ce polynôme. ?
   (c) On suppose dans cette question que g ne possède que 0 comme valeur propre. Montrer que g est nilpotent.

Solution de l'exercice 1.12

1. Soit $x \in E$ tel que $f^{p-1}(x) \neq 0_E$ (l'existence d'un tel x est garantie par le fait que f^{p-1} n'est pas l'endomorphisme nul).
Montrons alors que la famille $(x, f(x), f^2(x), \dots, f^{p-1}(x))$ est libre.
Soient donc $\lambda_0, \lambda_1, \dots, \lambda_{p-1}$ des réels tels que

$$\lambda_0 x + \lambda_1 f(x) + \cdots + \lambda_{p-1} f^{p-1}(x) = 0_E.$$

En appliquant f^{p-1} à cette relation, il vient

$$\lambda_0 f^{p-1}(x) + \lambda f^p(x) + \cdots + f^{2p-2}(x) = 0_E \Rightarrow \lambda_0 f^{p-1}(x) = 0_E.$$

Mais par hypothèse, $f^{p-1}(x) \neq 0_E$, et donc $\lambda_0 = 0$.
Ainsi, il reste $\lambda_1 f(x) + \lambda_2 f^2(x) + \cdots + \lambda_{p-1} f^{p-1}(x) = 0_E$.
Mais en appliquant f^{p-2}, on obtient alors $\lambda_1 f^{p-1}(x) = 0_E$, et donc $\lambda_1 = 0$.
Donc la relation de départ devient $\lambda_2 f^2(x) + \cdots + \lambda_{p-1} f^{p-1}(x) = 0_E$.
De proche en proche[1], on prouve ainsi que $\lambda_0 = \lambda_1 = \cdots = \lambda_{p-1} = 0$.
Et donc la famille $(x, f(x), \dots, f^{p-1}(x))$ est une famille libre de E, de cardinal p. Puisque $\dim E = n$, on en déduit que $p \leqslant n$.

Détails
Puisque f^p est nul, il est facile de voir que pour tout $k \geqslant p$, $f^k = 0$.

[1] Une récurrence serait plus rigoureuse, mais est-elle vraiment nécessaire ?

2. Puisque $f^n = 0$, il vient $u^{2n} = (u^2)^n = f^n = 0$.
Notons k le plus petit entier tel que $u^k = 0$ (un tel k existe puisque $\{m \in \mathbf{N} \text{ tel que } u^m = 0\}$ est non vide car il contient $2n$, et toute partie non vide de $\mathbf{N}$ possède un plus petit élément.
Puisque $u^{2n-2} = (u^2)^{n-1} = f^{n-1} \neq 0$, u^{2n-2} est non nul. Et donc $k = 2n-1$ ou $k = 2n$.
Mais le résultat de la question 1 s'applique également à l'endomorphisme u, et donc $k \leqslant 2$, ce qui est contradictoire[2] avec $k \geqslant 2n-1$.
Autrement dit, l'équation $u^2 = f$ ne possède pas de solution.

Rappel
Le cardinal d'une famille libre est toujours inférieur ou égal à la dimension de l'espace.

[2] Car $n \geqslant 2$ et donc $n < 2n-1$.

3. Puisque $f^p = 0$, le polynôme X^p est annulateur de f.
Et donc la seule valeur propre possible de f est 0, qui est la seule racine de X^p.
D'autre part, f n'est pas inversible, car sinon f^p le serait également car composée d'endomorphismes inversibles. Or $f^p = 0$, et l'endomorphisme nul n'est évidemment pas inversible.
Et donc 0 est bien valeur propre de f : $\mathrm{Spec}(f) = \{0\}$.
Remarque *: tout ce qui a été dit jusqu'à présent reste valable si E est un espace vectoriel sur* $\mathbf{R}$.

Valeurs possibles
À ce stade, rien ne permet de garantir que 0 est bien une valeur propre de f, nous savons uniquement que c'est le seul candidat.

4.a. Soit P un polynôme non nul, annulateur de g (nous savons qu'il en existe un).
Nous savons que toutes les valeurs propres de g sont racines de P, mais il se peut également qu'il existe des racines de P qui ne soient pas valeurs propres de g.
Par le théorème de d'Alembert-Gauss, nous savons que P se décompose en produit de facteurs de degré 1.
Notons $P = \prod_{i=1}^{k}(X-\lambda_i)\prod_{p=1}^{m}(X-\mu_p)$, où $\mathrm{Spec}(g) = \{\lambda_1, \dots, \lambda_k\}$, et où $\mu_1, \dots, \mu_m$ sont les racines de P qui ne sont pas valeurs propres de g.
Alors, pour tout $p \in [\![1, m]\!]$, $g - \mu_p \mathrm{id}_E$ est inversible.
Et donc en composant à gauche la relation $P(g) = 0$ par $(g-\mu_1\mathrm{id}_E)^{-1} \circ \cdots \circ (g-\mu_m\mathrm{id}_E)^{-1}$, il vient

$$(g-\mu_1\mathrm{id}_E)^{-1}\circ\cdots\circ(g-\mu_m\mathrm{id}_E)^{-1}\circ(g-\mu_1\mathrm{id}_E)\circ\cdots\circ(g-\mu_m\mathrm{id}_E)\circ(g-\lambda_1\mathrm{id}_E)\circ\cdots\circ(g-\lambda_k\mathrm{id}_E) = 0.$$

Or les $g - \mu_p\mathrm{id}_E$ sont des polynômes en g, et donc commutent deux à deux, donc il vient

$$(g-\mu_1\mathrm{id}_E)^{-1}\circ\cdots\circ(g-\mu_m\mathrm{id}_E)^{-1}\circ(g-\mu_m\mathrm{id}_E)\circ\cdots\circ(g-\mu_1\mathrm{id}_E)\circ(g-\lambda_1\mathrm{id}_E)\circ\cdots\circ(g-\lambda_k\mathrm{id}_E) = 0.$$

Remarque
Il se peut tout à fait que certains des μ_i soient égaux, si P possède des racines multiples.

Soit encore

$$(g-\mu_1 \mathrm{id}_E)\circ\cdots\circ(g-\mu_k \mathrm{id}_E)=0.$$

Ainsi, le polynôme $\prod_{i=1}^{k}(X-\lambda_i)$ est un polynôme annulateur de g, dont les seules racines réelles sont $\lambda_1,\dots,\lambda_k$, qui sont exactement les valeurs propres de g.

4.b. Notons $A=\{\deg P, P\in\mathbf{R}[X] \text{ non nul et annulateur de } g\}$.
Alors A est une partie non vide de $\mathbf{N}$ et donc possède un plus petit élément d.
Et donc un polynôme annulateur R de g et de degré d est donc de degré minimal parmi les polynômes annulateurs de g.

Notons $Q=RM+U$ la division euclidienne de Q par R, avec $\deg U<\deg R$. On obtient alors

$$0=Q(g)=(RM)(g)+U(g)=\underbrace{R(g)}_{=0}\circ M(g)+U(g).$$

Et donc U est annulateur de g. Puisqu'il est de degré strictement inférieur à Q, au vu de la définition de Q, on a donc $Q=0$.
Et ainsi, $Q=RM$: R divise Q.

4.c. Supposons que g ne possède que 0 comme valeur propre, et soit Q comme dans la question 4.a.
Alors Q ne possède que 0 comme racine. Et donc le polynôme R de la question 4.b ne possède également que 0 comme racine, et donc est de la forme $R=\lambda X^k$, avec $\lambda\neq 0$ et $k\in\mathbf{N}^*$.
Et donc $\lambda g^k=0$, de sorte que $g^k=0$.
De plus, $g^{k-1}\neq 0$, car le polynôme X^{k-1} ne peut être annulateur de g puisqu'il est de degré strictement inférieur à R.
Donc g es t nilpotent d'indice k.

Remarque *: le résultat de la dernière question n'est plus valable pour des espaces vectoriels réels. En effet, si f est l'endomorphisme de $\mathbf{R}^3$ dont la matrice dans la base canonique est $A=\begin{pmatrix}0&1&0\\-1&0&0\\0&0&0\end{pmatrix}$, alors f possède 0 comme unique valeur propre.*
Pourtant, f n'est pas nilpotent car $A^4=I_3$ et donc pour tout $p\in\mathbf{N}$, $A^{4p}=I_3$. On ne saurait donc avoir $A^k=0$ pour un certain $k\in\mathbf{N}$.
Ceci tient au fait que A possède des valeurs propres complexes non nulles, qui sont i et i. □

Danger !
Le fait que $R=\lambda X^k$ vient du fait qu'on travaille avec un polynôme complexe, qui possède nécessairement autant de racines (comptées avec multiplicités) que son degré.
Un polynôme réel ne possédant que 0 comme racine n'est pas forcément de cette forme, par exemple on peut prendre $X^3+X=X(X^2+1)$.

Dans une «bonne» base, la matrice d'un endomorphisme nilpotent est triangulaire supérieure, avec des 0 sur la diagonale. En petite dimension, il est même possible de donner explicitement une telle matrice.

EXERCICE 1.13 (QSP HEC 2014) [HEC14.QSP104] — Moyen

Réduction d'une matrice nilpotente d'indice 2.

Soit $A\in\mathcal{M}_3(\mathbf{R})$ une matrice non nulle telle que $A^2=0$. Montrer que A est semblable à $\begin{pmatrix}0&0&1\\0&0&0\\0&0&0\end{pmatrix}$.

SOLUTION DE L'EXERCICE 1.13 Soit f l'endomorphisme de $\mathbf{R}^3$ dont la matrice dans la base canonique est A.
Nous cherchons donc à prouver qu'il existe une base $\mathcal{B}=(e_1,e_2,e_3)$ de $\mathbf{R}^3$ telle que

$$\mathrm{Mat}_{\mathcal{B}}(f)=\begin{array}{c}\begin{matrix}f(e_1)&f(e_2)&f(e_3)\end{matrix}\\\begin{pmatrix}0&0&1\\0&0&0\\0&0&0\end{pmatrix}\end{array}\begin{matrix}e_1\\e_2\\e_3\end{matrix}.$$

Cela revient à demander que e_1,e_2 soient dans $\mathrm{Ker}\, f$ et que e_3 soit un antécédent de e_1 par f.

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

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 inversible non plus.
En particulier, $\dim \operatorname{Ker} f \geqslant 1$, et f étant non nulle, $\dim \operatorname{Ker} f \leqslant 2$.
Enfin, $f \circ f = 0$, donc $\operatorname{Im} f \subset \operatorname{Ker} f$. Ceci impose alors $\dim \operatorname{Im} f \leqslant \dim \operatorname{Ker} f$.
Comme le théorème du rang nous indique que $\dim \operatorname{Ker} f + \dim \operatorname{Im} f = 3$, il n'y a pas d'autre solution que $\dim \operatorname{Im} f = 1$ et $\dim \operatorname{Ker} f = 2$.
Soit alors $x \in \mathbf{R}^3$ tel que $f(x) \neq 0$.
On a $f(x) \in \operatorname{Ker} f$ car $f^2(x) = 0$. Autrement dit, la famille formée du seul vecteur $f(x)$ est une famille libre de $\operatorname{Ker} f$, qui peut être complétée en une base $(f(x), y)$ de $\operatorname{Ker} f$.
Enfin, $x \notin \operatorname{Ker} f$ par définition de x, et donc $\operatorname{Ker} f \cap \operatorname{Vect}(x) = \{0\}$.
On en déduit que $\mathbf{R}^3 = \operatorname{Ker} f \oplus \operatorname{Vect}(x)$, et donc $(f(x), y, x)$ est une base de $\mathbf{R}^3$ obtenue par concaténation de bases de $\operatorname{Ker} f$ et de $\operatorname{Vect}(x)$.
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. □

Si N est une matrice nilpotente, alors $I_n - N$ est inversible, et nous connaissons son inverse en fonction de N. Toutefois, la forme de cette inverse ne s'invente pas si on ne l'a jamais vu.

Exercice 1.14 (QSP ESCP 2016) [ESCP16.QSP01] **Moyen**

Inverse de $I_n - A$, A nilpotente

Soit N une matrice de $\mathcal{M}_n(\mathbf{R})$ $(n \geqslant 2)$ nilpotente, *i.e.* telle qu'il existe $p \geqslant 1$ tel que $N^p = 0$.

1. Montrer que la matrice $A = I_n - N$ est inversible et déterminer son inverse.
2. Montrer que $I - A^{-1}$ est nilpotente.

Solution de l'exercice 1.14

1. Montrer que A est inversible n'est pas trop difficile. En effet, X^p est un polynôme annulateur de N, et puisque sa seule racine est 0, la seule valeur propre possible de N est 0.
En particulier, 0 n'est pas valeur propre de N, et donc $I_n - N$ est inversible.

Toutefois, ceci ne permet pas de déterminer l'inverse de N. Pour cela, remarquons que $I_n = I_n - N^p$. Or

$$(I_n - N)(I_n + N + \cdots + N^{p-1}) = I_n + \cancel{N} + \cancel{N^2} + \cdots + \cancel{N^{p-1}} - \cancel{N} - \cancel{N^2} - \cdots - \cancel{N^{p-1}} - N^p = I_n - N^p = I_n.$$

Et donc $A = I_n - N$ est inversible, et $A^{-1} = I_n + N + \cdots + N^{p-1}$.

Remarque

Ceci est à mettre en parallèle avec l'identité classique

$$a^p - b^p = (a-b) \times (a^{p-1} + a^{p-2}b + \cdots + b^{p-1}).$$

Cette égalité reste en fait encore valable pour deux matrices A et B **si elles commutent** (développer le produit pour s'en convaincre).

Alternative : si l'on ne connaît pas l'astuce qui précède, il existe une autre manière de procéder à l'aide de polynômes annulateurs. Notons que $(A - I_n)^p = (-N)^p = 0$.
Et donc

$$\sum_{k=0}^{p} \binom{p}{k} (-1)^{p-k} A^k = 0 \Leftrightarrow A \sum_{k=1}^{p} (-1)^{p-k} A^{k-1} = (-1)^{p+1} I_n.$$

Donc $A^{-1} = \displaystyle\sum_{k=1}^{p} \binom{p}{k} (-1)^{p-k} A^{k-1}$.
Mais alors, une seconde application de la formule du binôme nous donne :

$$\begin{aligned} A^{-1} &= \sum_{k=1}^{p} \binom{p}{k} (-1)^{p-k} (I_n - N)^{k-1} \\ &= \sum_{k=1}^{p} \binom{p}{k} (-1)^{p-k} \sum_{i=0}^{k-1} \binom{k-1}{i} (-1)^i N^i \\ &= \sum_{i=0}^{p-1} (-1)^{p+i} N^i \sum_{k=1}^{i+1} \frac{p!}{k!(p-k)!} \frac{(k-1)!}{i!(k-1-i)!} (-1)^k \end{aligned}$$

$$= \sum_{i=0}^{p-1} (-1)^{p+i} N^i \sum_{k=1}^{i+1} \frac{p!}{k!(p-k)!} \frac{(k-1)!}{i!(k-1-i)!} (-1)^k$$

2. On a

$$I_n - A^{-1} = -\left(N + N^2 + \cdots + N^{p-1}\right) = -N\left(I_n + N + N^{p-2}\right).$$

Mais $I_n + N + \cdots + N^{p-2}$ est un polynôme en N, et donc commute avec N.
Par conséquent, il vient

$$(I_n - A^{-1})^p = (-1)^p \underbrace{N^p}_{=0} \left(I_n + N + \cdots + N^{p-2}\right)^p = 0.$$

Et donc $I_n - A^{-1}$ est une matrice nilpotente.

□

1.5 MATRICES

QSP ESCP 2018 sur le commutant d'un endo possédant des valeurs propres distinctes.

EXERCICE 1.15 (QSP HEC 2016) [HEC16.QSP185] — **Moyen**

Existence de vecteurs possédant les mêmes coordonnées dans deux bases.

Soit $\mathcal{B} = (e_1, e_2, e_3)$ une base de $\mathbf{R}^3$ et $\mathcal{B}' = (e_1 + e_2, 2e_2 + e_3, 3e_3)$.

1. Existe-t-il un vecteur non nul de $\mathbf{R}^3$ ayant les mêmes coordonnées dans ces deux bases ?
2. Plus généralement, si $\mathcal{B}$ et $\mathcal{B}'$ sont deux bases de $\mathbf{R}^n$, à quelle condition existe-t-il un vecteur non nul de $\mathbf{R}^n$ ayant les mêmes coordonnées dans ces deux bases ?

SOLUTION DE L'EXERCICE 1.15

1. Commençons par remarquer que $\mathcal{B}'$ est bien une base de $\mathbf{R}^3$ car $\mathrm{Mat}_{\mathcal{B}}(\mathcal{B}') = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 2 & 0 \\ 0 & 1 & 3 \end{pmatrix}$ est inversible[1].
Par la formule de changement de base, un vecteur $xe_1 + y_2 + ze_3$ a pour coordonnées dans la base $\mathcal{B}'$

$$\begin{pmatrix} x' \\ z' \\ y' \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 2 & 0 \\ 0 & 1 & 3 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x \\ x + 2y \\ y + 3z \end{pmatrix}.$$

Ces coordonnées sont égales à $\begin{pmatrix} x \\ y \\ z \end{pmatrix}$ si et seulement si

$$\begin{cases} x = x \\ x + 2y = y \\ y + 3z = z \end{cases} \Leftrightarrow \begin{cases} y = -x \\ x = 2z \end{cases} \Leftrightarrow \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathrm{Vect} \begin{pmatrix} 2 \\ -2 \\ 1 \end{pmatrix}.$$

Et donc, par exemple le vecteur $2e_1 - 2e_2 + e_3$ convient.

2. Notons $\mathcal{B} = (e_1, \ldots, e_n)$, et soit $x = \sum_{i=1}^{n} x_i e_i \in \mathbf{R}^n$ non nul.

Alors les coordonnées de x dans la base $\mathcal{B}'$ sont $P_{\mathcal{B}',\mathcal{B}} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$.

Elles sont égales aux coordonnées de x dans la base $\mathcal{B}$ si et seulement

$$\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = P_{\mathcal{B}',\mathcal{B}} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}.$$

[1] C'est une matrice triangulaire dont aucun des coefficients diagonaux n'est nul.

Remarque

L'énoncé précise bien qu'on cherche un vecteur non nul, puisque les coordonnées du vecteur nul dans toute base de $\mathbf{R}^3$ seront $(0, 0, 0)$.

Vérification

On vérifie facilement, même si c'est inutile, que

$$2e_1 - 2e_2 + e_3 = 2(e_1+e_2) - 2(2e_2+e_3) + 3e_3.$$

Autrement dit, si et seulement si $\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ est un vecteur propre[2] de $P_{\mathscr{B}',\mathscr{B}}$ associé à la valeur propre 1.
Et donc un tel vecteur existe si et seulement si 1 est valeur propre de $P_{\mathscr{B}',\mathscr{B}}$.
Notons que c'est le cas si et seulement si 1 est valeur propre de $P_{\mathscr{B},\mathscr{B}'} = P_{\mathscr{B}',\mathscr{B}}^{-1}$.

[2] Notons que ce vecteur est non nul car $x \neq 0$.

□

Les matrices magiques sont des matrices telles que la somme de chaque ligne et de chaque colonne soit la même (souvenez-vous des carrés magiques que vous avez peut-être rencontrés dans votre jeunesse).
C'est un thème qui revient régulièrement à l'oral. On peut l'aborder «à la main», comme dans l'exercice suivant, mais également en termes de valeurs propres.

Exercice 1.16 (ESCP 2009) [ESCP09.2.11] — Moyen

Matrices magiques.

Soit $n \geqslant 2$. On considère la matrice $J \in \mathscr{M}_n(\mathbf{R})$ dont tous les coefficients valent 1.
Soit $\mathscr{C}$ l'ensemble

$$\mathscr{C} = \left\{ A = (a_{i,j}) \in \mathscr{M}_n(\mathbf{R}), \exists \alpha \in \mathbf{R}, \forall i,j \in [\![1,n]\!]^2, \alpha = \sum_{k=1}^{n} a_{i,k} = \sum_{k=1}^{n} a_{k,j} \right\}.$$

1. Montrer que $\mathscr{C}$ est un $\mathbf{R}$-espace vectoriel.
2. Montrer que l'application d définie sur $\mathscr{C}$ et à valeurs réelles par $d(A) = \sum_{k=1}^{n} a_{1,k}$ est une application linéaire surjective, non injective.
3. (a) Soit $A \in \mathscr{M}_n(\mathbf{R})$. Montrer que A appartient à $\mathscr{C}$ si et seulement s'il existe un réel λ tel que $AJ = JA = \lambda J$.
   (b) Soient A, B deux matrices de $\mathscr{C}$. Montrer que AB appartient à $\mathscr{C}$, et calculer $d(AB)$.
   (c) Soit A une matrice inversible de $\mathscr{C}$. Montrer que A^{-1} appartient à $\mathscr{C}$, et trouver une relation entre $d(A)$ et $d(A^{-1})$.
4. Montrer que $\mathrm{Ker}(d)$ et $\mathrm{Vect}(J)$ sont deux sous-espaces vectoriels supplémentaires dans $\mathscr{C}$.
5. Soit $(r,s) \in [\![2,n]\!]^2$, et soit $A_{r,s} = (a_{i,j})$ la matrice dont tous les éléments sont nuls sauf $a_{1,1} = a_{r,s} = 1$ et $a_{1,s} = a_{r,1} = -1$.
   Montrer que la famille $(A_{r,s})_{(r,s)\in[\![2,n]\!]^2}$ forme une base de $\mathrm{Ker}(d)$, et en déduire la dimension de $\mathscr{C}$.
6. Soit p un entier naturel non nul et A une matrice de $\mathscr{C}$. Montrer que $B = \frac{d(A)}{n} J$ est solution de l'équation $A^p - B^p = (A-B)^p$.

Solution de l'exercice 1.16

1. Il faut faire les calculs pour le prouver proprement. Intuitivement, il faut comprendre que les matrices de $\mathscr{C}$ sont celles telles que la somme de n'importe quelle ligne soit égale à la somme de n'importe quelle colonne (la somme étant le α qui apparaît dans la définition de $\mathscr{C}$).
Il est clair que la matrice nulle vérifie les conditions définissant $\mathscr{C}$. Soient donc $A, B \in \mathscr{C}$, et $\lambda \in \mathbf{R}$.
Notons α, β les constantes telles que

$$\forall (i,j) \in [\![1,n]\!]^2, \alpha = \sum_{k=1}^{n} a_{i,k} = \sum_{k=1}^{n} a_{k,j} \text{ et } \beta = \sum_{k=1}^{n} b_{i,k} = \sum_{k=1}^{n} b_{k,j}.$$

Alors notons $C = \lambda A + B = (\lambda a_{i,j} + b_{i,j})$. On a, $\forall (i,j) \in [\![1,n]\!]^2$,

$$\sum_{k=1}^{n} c_{i,k} = \lambda \sum_{k=1}^{n} a_{i,k} + \sum_{k=1}^{n} b_{i,k} = \lambda\alpha + \beta$$

et de même

$$\sum_{k=1}^{n} c_{k,j} = \lambda \sum_{k=1}^{n} a_{k,j} + \sum_{k=1}^{n} b_{k,j} = \lambda\alpha + \beta.$$

On en déduit que $\lambda A + B \in \mathscr{C}$, et donc $\mathscr{C}$ est un sous-$\mathbf{R}$-espace vectoriel de $\mathscr{M}_n(\mathbf{R})$.

2. Notons que d est l'application qui à une matrice associe la somme de sa première ligne, et qui donc à une matrice de $\mathscr{C}$ associe la somme de n'importe laquelle de ses lignes, ou de n'importe laquelle de ses colonnes. Si $A, B \in \mathscr{C}$, et si $\lambda \in \mathbf{R}$, alors

$$d(\lambda A + B) = \sum_{k=1}^{n} (\lambda a_{1,k} + b_{1,k}) = \lambda \sum_{k=1}^{n} a_{1,k} + \sum_{k=1}^{n} b_{1,k} = \lambda d(A) + d(B).$$

Donc d est bien linéaire, et son image est un sous-espace vectoriel de $\mathbf{R}$, donc de dimension 0 ou 1.
Puisque $d(J) = n \neq 0$, on en déduit que $\dim \mathrm{Im} d = 1$, et donc d est surjective.
On a $J \in \mathscr{C}$, et $d(J) = n$. De même, $nI_n \in \mathscr{C}$ (car la somme de n'importe quelle ligne ou n'importe quelle colonne ne comporte qu'un seul coefficient non nul, égal à n) et donc $d(nI_n) = n$. Pourtant $nI_n \neq J$, donc d n'est pas injective.

> **Alternative**
> Si $f : E \to F$ est injective, alors $\dim E \leqslant \dim F$. Puisqu'ici $\dim \mathscr{M}_n(\mathbf{R}) > \dim \mathbf{R}$, d ne saurait être injective. D'ailleurs d est une forme linéaire non nulle, donc son noyau est un hyperplan de $\mathscr{M}_n(\mathbf{R})$, donc de dimension $n^2 - 1 \neq 0$.

3.a. Soit $A = (a_{i,j}) \in \mathscr{M}_n(\mathbf{R})$. Alors

$$AJ = \begin{pmatrix} \sum_{k=1}^n a_{1,k} & \sum_{k=1}^n a_{1,k} & \cdots & \sum_{k=1}^n a_{1,k} \\ \sum_{k=1}^n a_{2,k} & \sum_{k=1}^n a_{2,k} & \cdots & \sum_{k=1}^n a_{2,k} \\ \vdots & \vdots & & \vdots \\ \sum_{k=1}^n a_{n,k} & \sum_{k=1}^n a_{n,k} & \cdots & \sum_{k=1}^n a_{n,k} \end{pmatrix}$$

De même, on a

$$JA = \begin{pmatrix} \sum_{k=1}^n a_{k,1} & \sum_{k=1}^n a_{k,2} & \cdots & \sum_{k=1}^n a_{k,n} \\ \sum_{k=1}^n a_{k,1} & \sum_{k=1}^n a_{k,2} & \cdots & \sum_{k=1}^n a_{k,n} \\ \vdots & \vdots & & \vdots \\ \sum_{k=1}^n a_{k,1} & \sum_{k=1}^n a_{k,2} & \cdots & \sum_{k=1}^n a_{k,n} \end{pmatrix}.$$

Alors si $A \in \mathscr{C}$, soit α tel que

$$\forall (i,j) \in [\![1,n]\!]^2, \alpha = \sum_{k=1}^{n} a_{i,k} = \sum_{k=1}^{n} a_{k,j}.$$

D'après les calculs qui précèdent, $AJ = JA = \alpha J$.

Inversement, supposons que $AJ = JA = \lambda J$.
Si $i \in [\![1,n]\!]$, alors en identifiant n'importe que coefficient de la i-ème ligne de l'égalité $AJ = \lambda J$, il vient

$$\sum_{k=1}^{n} a_{i,k} = \lambda.$$

De même, si $j \in [\![1,n]\!]^2$, alors en identifiant n'importe que coefficient de la j-ième colonne de l'égalité $JA = \lambda J$, il vient

$$\sum_{k=1}^{n} a_{k,j} = \lambda.$$

On en déduit que A appartient donc à $\mathscr{C}$.
Notons que si $A \in \mathscr{C}$, l'unique $\lambda \in \mathbf{R}$ tel que $AJ = JA = \lambda J$ est en fait égal à $d(A)$.

3.b. Soient $A, B \in \mathscr{C}$ et soient $\lambda, \mu \in \mathbf{R}$ tels que

$$AJ = JA = \lambda J \text{ et } BJ = JB = \mu J.$$

> **Alternative**
> On peut aussi revenir à la définition de $\mathscr{C}$, mais alors les calculs sont interminables...

Alors $(AB)J = A(BJ) = A(JB) = (AJ)B = J(AB)$, et $(AB)J = A(BJ) = A(\mu J) = \mu(AJ) = \mu\lambda AJ$.
On en déduit que $AB \in \mathscr{C}$, et d'après ce que nous avons remarqué à la fin de la question précédente, nous avons nécessairement $d(AB) = \lambda\mu = d(A)d(B)$.

3.c. Soit $A \in \mathscr{C}$ une matrice inversible, et soit $\lambda \in \mathbf{R}$ tel que

$$AJ = JA = \lambda J.$$

Alors en multipliant à gauche cette égalité par A^{-1}, il vient

$$J = A^{-1}JA = \lambda A^{-1}J.$$

De même, en multipliant à droite par A^{-1}, on a

$$AJA^{-1} = J = \lambda JA^{-1}$$

Nécessairement, on a $\lambda \neq 0$, car sinon $J = \lambda JA^{-1}$ serait la matrice nulle. Alors

$$JA^{-1} = A^{-1}J = \frac{1}{\lambda}J.$$

On en déduit que $A^{-1} \in \mathscr{C}$ et $d(A^{-1}) = \frac{1}{\lambda} = \frac{1}{d(A)}$.

4. Par le théorème du rang appliqué à d, il est clair que

$$\dim \mathrm{Ker} d + 1 = \dim \mathscr{C}.$$

Ainsi, un supplémentaire de $\mathrm{Ker} d$ dans $\mathscr{C}$ est nécessairement de dimension 1.
Or, nous savons que $J \notin \mathrm{Ker} d$ car $d(J) = n$.
On en déduit que $\mathrm{Vect}(J)$ est un sous-espace vectoriel de dimension 1 de $\mathscr{C}$ tel que $\mathrm{Vect}(J) \cap \mathrm{Ker} d = \{0\}$.
Nous en déduisons que

$$\mathscr{C} = \mathrm{Ker} d \oplus \mathrm{Vect}(J).$$

Hyperplan

Nous l'avons déjà dit, mais $\mathrm{Ker}\, d$ est un hyperplan de $\mathscr{C}$, et donc tout supplémentaire en est de dimension 1.
Et même mieux, si $A \in \mathscr{C}$ vérifie $d(A) \neq 0$,alors $\mathrm{Vect}(A)$ est un supplémentaire de $\mathrm{Ker}\, d$.

5. Il est facile de vérifier que $A_{r,s} \in \mathscr{C}$, et que la somme de chaque ligne est nulle, autrement dit, que $d(A_{r,s}) = 0$, et donc que $A_{r,s} \in \mathrm{Ker} d$. Montrons que la famille $(A_{r,s})_{2 \leqslant r,s \leqslant n}$ est une famille libre de $\mathscr{C}$. Soient $(\lambda_{r,s})_{2 \leqslant r,s \leqslant n}$ des réels tels que

$$\sum_{2 \leqslant r,s \leqslant n} \lambda_{r,s} A_{r,s} = 0.$$

Alors, en identifiant le coefficient situé à la r-ème ligne et la s-ième colonne de l'égalité précédente, il vient $\lambda_{r,s} = 0$.
Ceci prouve donc qu'il s'agit d'une famille libre.
Montrons à présent qu'elle est génératrice de $\mathrm{Ker} d$, et soit $A \in \mathrm{Ker} d$. Alors, par définition,

$$\forall (i,j) \in [\![1,n]\!]^2, \sum_{k=1}^{n} a_{i,k} = \sum_{k=1}^{n} a_{k,j} = 0.$$

En particulier,

$$\forall s \in [\![2,n]\!], a_{1,s} = -\sum_{r=2}^{n} a_{r,s}$$

$$\forall r \in [\![2,n]\!], a_{r,1} = -\sum_{s=2}^{n} a_{r,s}$$

$$\text{et donc } a_{1,1} = -\sum_{s=2}^{n} a_{1,s} = \sum_{s=2}^{n}\sum_{r=2}^{n} a_{r,s}$$

On en déduit que

$$\begin{pmatrix} a_{1,1} & a_{1,2} & \dots & a_{1,n} \\ a_{2,1} & a_{2,2} & \dots & a_{2,n} \\ \vdots & \vdots & & \vdots \\ a_{n,1} & a_{n,2} & \dots & a_{n,n} \end{pmatrix} = \begin{pmatrix} -\sum_{r=2}^{n}\sum_{s=2}^{n} a_{r,s} - \sum_{s=2}^{n} a_{2,s} & -\sum_{r=2}^{n} a_{r,2} & \dots & -\sum_{r=2}^{n} a_{r,n} \\ a_{2,2} & \vdots & \dots & a_{2,n} \\ \vdots & \vdots & & \vdots \\ -\sum_{s=2}^{n} a_{n,s} & a_{n,2} & \dots & a_{n,n} \end{pmatrix}$$

$$= \sum_{r=2}^{n}\sum_{s=2}^{n} a_{r,s}A_{r,s}$$

Ainsi, la famille $(A_{r,s})_{2\leqslant r,s\leqslant n}$ est une famille génératrice de $\mathrm{Ker}d$: c'est une base de $\mathrm{Ker}d$. On en déduit que $\dim \mathrm{Ker}d = (n-1)^2$, et donc que

$$\dim \mathscr{C} = (n-1)^2 + 1.$$

> **Intuition**
> Une matrice est dans $\mathrm{Ker}\, d$ si et seulement si la somme de chaque ligne et chaque colonne est nulle.
> Sur chaque ligne, on peut choisir tous les coefficients des colonnes de 2 à n, mais alors on n'a plus le choix pour le premier, qui doit être égal à l'opposé de la somme des autres. Idem pour les colonnes. Ainsi, on a $(n-1)^2$ «degrés de liberté», et on doit donc avoir $\dim \mathrm{Ker}\,\mathscr{C} = (n-1)^2$.

6. Nous avons prouvé précédemment que si $A \in \mathscr{C}$, alors A et J commutent. Il en est donc de même de A et de $\frac{d(A)}{n}J$.
Nous pouvons alors utiliser la formule du binôme de Newton pour calculer

$$\left(A - \frac{d(A)}{n}J\right)^p = \sum_{k=1}^{p}\binom{p}{k}A^k\left(-\frac{d(A)}{J}\right)^{p-k}$$

Or, on a, pour $k < p$ (c'est-à-dire pour $p-k \neq 0$),

$$A^k J^{p-k} = A^{k-1}(AJ)J^{p-k-1} = A^{k-1}(d(A)J)J^{p-k-1} = d(A)A^{k-1}J^{p-k} = \cdots = d(A)^k J^{p-k}$$

De plus, nous avons $J^2 = nJ$, et donc une récurrence facile prouve que

$$\forall k \geqslant 2, J^k = n^{k-1}J.$$

Ainsi, il vient

$$\forall k \in [\![1, p-1]\!], \left(\frac{d(A)}{n}\right)^{p-k} A^k J^{p-k} = \frac{d(A)^{p-k}}{n^{p-k}}d(A)^k n^{p-k-1}J = \frac{d(A)^p}{n}J$$

On en déduit que

$$\left(A - \frac{d(A)}{n}J\right)^p = \left(-\frac{d(A)}{n}\right)^p J^p + \sum_{k=1}^{p-1}\binom{p}{k}(-1)^{p-k}\frac{d(A)^p}{n}J + A^p$$

Mais puisque $(1-1)^p = 0$, alors

$$\sum_{k=1}^{p-1}\binom{p}{k}(-1)^{p-k} = \underbrace{\sum_{k=0}^{p}\binom{p}{k}(-1)^{p-k}}_{=0} - 1 - (-1)^p = -(1+(-1)^p)$$

et donc

$$\begin{aligned}(A-J)^p &= (-1)^p\frac{d(A)^p}{n^p}J^p - (1+(-1)^p)\frac{d(A)^p}{n}J + A^p\\ &= (-1)^p\frac{d(A)^p}{n}J - (1+(-1)^p)\frac{d(A)^p}{n}J + A^p\\ &= A^p - \frac{d(A)^p}{n}J\\ &= A^p - \left(\frac{d(A)}{n}J\right)^p\end{aligned}$$

La dernière égalité vient du fait que

$$\left(\frac{d(A)}{n}J\right)^p = \frac{d(A)^p}{n^p}J^p = \frac{d(A)^p}{n^p}n^{p-1}J = \frac{d(A)^p}{n}J.$$

□

Exercice 1.17 (ESCP 2016) [ESCP16.2.15] — Moyen

Formule d'inversion de Pascal et asymptotique du nombre de permutations de $[\![1,n]\!]$ sans point fixe
Abordable en première année : ✓

1. Soient $(u_n)_{n\in\mathbf{N}}$ et $(v_n)_{n\in\mathbf{N}}$ deux suites réelles telles que : $\forall n\in\mathbf{N},\ u_n=\sum_{i=0}^{n}\binom{n}{i}v_i$.

   (a) Déterminer une matrice $P\in\mathcal{M}_{n+1}(\mathbf{R})$ telle qu'on ait l'égalité matricielle :
   $$(u_0\ \cdots\ u_n)=(v_0\ \cdots\ v_n)P.$$

   (b) On note $\mathbf{R}_n[X]$ l'espace vectoriel des polynômes à coefficients réels de degré inférieur ou égal à n et $\mathcal{B}_0=(1,X,X^2,\dots,X^n)$ sa base canonique.
   Montrer que $\mathcal{B}_1=(1,1+X,(1+X)^2,\dots,(1+X)^n)$ est une base de $\mathbf{R}_n[X]$.
   Montrer que P est inversible et calculer son inverse. *(On pourra montrer que P est la matrice de passage entre deux bases de $\mathbf{R}_n[X]$).*

   (c) En déduire l'expression de v_n en fonction des $u_i, 0\leqslant i\leqslant n$:
   $$v_n=\sum_{i=0}^{n}\binom{n}{i}(-1)^{n-i}u_i.$$

2. On rappelle qu'une permutation de $[\![1,n]\!]$ est une bijection de $[\![1,n]\!]$ dans $[\![1,n]\!]$.
   Si σ est une permutation de $[\![1,n]\!]$, on dit que $i\in[\![1,n]\!]$ est un point fixe de σ si $\sigma(i)=i$.
   On note d_n le nombre de permutations sans point fixe de $[\![1,n]\!]$ (i.e. de permutations σ de $[\![1,n]\!]$ telles que $\forall i,\sigma(i)\neq i$) et on pose : $d_0=1$.

   (a) Pour tout entier $n\geqslant 0$, montrer la relation : $n!=\sum_{i=0}^{n}\binom{n}{i}d_i$.

   (b) En déduire l'expression de d_n en fonction de n pour $n\geqslant 0$.

   (c) Soit $(a_n)_{n\geqslant 0}$ la suite définie par $a_0=1$ et $\forall n\in\mathbf{N},\ a_{n+1}=(n+1)a_n+(-1)^{n+1}$.
   Montrer que les deux suites (a_n) et (d_n) sont égales.

3. On note pour tout entier naturel n : $J_n=\int_0^1 x^n e^x\,dx$.

   (a) Montrer que : $\forall n\in\mathbf{N},\ \forall x\in\mathbf{R},\ e^x=\sum_{k=0}^{n}\frac{x^k}{k!}+\int_0^x\frac{(x-t)^n}{n!}e^t\,dt$.

   (b) En déduire que : $\forall n\in\mathbf{N},\ d_n=\frac{1}{e}(n!+(-1)^n J_n)$.

   (c) Montrer que lorsque n tend vers $+\infty$: $J_n\underset{n\to+\infty}{\sim}\frac{e}{n}$.
   En déduire un équivalent de d_n.

   (d) Déterminer la nature de la série de terme général $\frac{(n-1)!}{e}-\frac{d_n}{n}$.

Solution de l'exercice 1.17

1.a. Soit $P=\begin{pmatrix}\binom{0}{0}&\binom{1}{0}&\binom{2}{0}&\cdots&\binom{n}{0}\\0&\binom{1}{1}&\binom{2}{1}&\cdots&\binom{n}{1}\\0&0&\binom{2}{2}&\cdots&\binom{n}{2}\\\vdots&\ddots&\ddots&&\vdots\\0&\cdots&\cdots&0&\binom{n}{n}\end{pmatrix}\in\mathcal{M}_{n+1}(\mathbf{R})$.
Alors on a bien
$$(v_0\ \ v_1\ \ \dots\ \ v_n)P=\left(v_0\ \ v_0+v_1\ \ \dots\ \ \sum_{i=0}^{n}\binom{n}{i}v_i\right)=(v_0\ \ v_1\dots\ \ v_n).$$

1.b. La famille $(1,1+X),(1+X)^2,\dots,(1+X)^n)$ est formée de polynômes de degrés deux à deux distincts : c'est donc une famille libre de $\mathbf{R}_n[X]$.

Étant de cardinal $n+1 = \dim \mathbf{R}_n[X]$, c'est une base de $\mathbf{R}_n[X]$. Puisque $(1+X)^k = \sum_{i=0}^{k} \binom{k}{i} X^i$, la matrice de passage de $\mathscr{B}_1$ à $\mathscr{B}_0$ est :

$$P_{\mathscr{B},\mathscr{B}_1} = \begin{array}{c} \begin{matrix} 1 & 1+X & (1+X)^2 & \dots & (1+X)^n \end{matrix} \\ \begin{pmatrix} \binom{0}{0} & \binom{1}{0} & \binom{2}{0} & \dots & \binom{n}{0} \\ 0 & \binom{1}{1} & \binom{2}{1} & \dots & \binom{n}{1} \\ 0 & 0 & \binom{2}{2} & \dots & \binom{n}{2} \\ \vdots & \ddots & \ddots & & \vdots \\ 0 & \dots & \dots & 0 & \binom{n}{n} \end{pmatrix} \end{array} \begin{matrix} 1 \\ X \\ X^2 \\ \vdots \\ X^n \end{matrix} = P.$$

Puisqu'une matrice de passage est toujours inversible, on en déduit que P est inversible, et son inverse est $P^{-1} = P_{\mathscr{B}_1,\mathscr{B}_0}$.
Or, pour tout $k \in [\![0,n]\!]$, par la formule du binôme de Newton, on a

$$X^k = ((1+X)-1)^k = \sum_{i=0}^{k} \binom{k}{i} (1+X)^i (-1)^{i-k}.$$

Et donc

$$P^{-1} = P_{\mathscr{B}_1,\mathscr{B}_0} = \begin{array}{c} \begin{matrix} 1 & X & X^2 & \dots & X^n \end{matrix} \\ \begin{pmatrix} \binom{0}{0} & -\binom{1}{0} & \binom{2}{0} & \dots & (-1)^n\binom{n}{0} \\ 0 & \binom{1}{1} & -\binom{2}{1} & \dots & (-1)^{n-1}\binom{n}{1} \\ 0 & 0 & \binom{2}{2} & \dots & (-1)^{n-2}\binom{n}{2} \\ \vdots & \ddots & \ddots & & \vdots \\ 0 & \dots & \dots & 0 & \binom{n}{n} \end{pmatrix} \end{array} \begin{matrix} 1 \\ 1+X \\ (1+X)^2 \\ \vdots \\ (1+X)^n \end{matrix}.$$

1.c. On a donc

$$\begin{pmatrix} v_0 & v_1 & \dots & v_n \end{pmatrix} = \begin{pmatrix} u_0 & u_1 & \dots & u_n \end{pmatrix} P^{-1}.$$

En particulier, $u_n = \sum_{i=0}^{n} (-1)^{n-i} \binom{n}{i} u_i$.

2.a. Commençons par rappeler que le nombre de total de permutations de $[\![1,n]\!]$ est $n!$
En effet choisir une permutation σ, c'est choisir l'image de 1 (il y a n choix possibles), puis choisir l'image de 2 (il y a $n-1$ choix possibles, puisque cette image ne peut être égale[1] à $\sigma(1)$), etc, choisir l'image de n (et il n'y a alors plus qu'un choix possible pour être différent de $\sigma(1), \sigma(2), \dots, \sigma(n-1)$).
Soit $i \in [\![1,n]\!]$ fixé. Alors le nombre de permutations de $[\![1,n]\!]$ possédant exactement i points fixes est $\binom{n}{i} d_{n-i}$.
En effet, il y a $\binom{n}{i}$ manières de choisir les i points fixes, et d_{n-i} manières de permuter sans points fixes les $n-i$ autres éléments de $[\![1,n]\!]$.
Puisqu'une permutation de $[\![1,n]\!]$ a soit aucun point fixe, soit un point fixe, ... ,soit n points fixes, il vient

$$n! = \sum_{i=0}^{n} \binom{n}{i} d_{n-i} = \sum_{j=0}^{n} \binom{n}{n-j} d_j = \sum_{j=0}^{n} \binom{n}{j} d_j.$$

[1] Une bijection est injective, et en particulier ne peut prendre deux fois la même valeur.

Détails
Le nombre de permutations d'un ensemble ne dépend que de son cardinal : et donc si E possède d éléments, il a autant de permutations que $[\![1,d]\!]$, c'est-à-dire $d!$

2.b. Si l'on applique le résultat de la question 1.c, avec $u_n = n!$ et $v_n = d_n$, il vient donc

$$\forall n \in \mathbf{N},\ d_n = \sum_{i=0}^{n} \binom{n}{i} (-1)^{n-i} i!.$$

2.c. Montrons par récurrence sur n que $d_n = a_n$.
Par définition, on a $d_0 = a_0 = 1$, donc la récurrence est initialisée.
Supposons que $d_n = a_n$. Alors

$$d_{n+1} = \sum_{i=0}^{n+1} \binom{n+1}{i} (-1)^{n+1-i} i!$$

$$= (-1)^{n+1} + \sum_{i=1}^{n+1} \binom{n+1}{i} (-1)^{n+1-i} i!$$

$$= (-1)^{n+1} + \sum_{i=1}^{n+1} \frac{n+1}{i} \binom{n}{i-1} (-1)^{n-(i+1)} i!$$

$$= (-1)^{n+1} + (n+1) \sum_{i=1}^{n+1} \binom{n}{i-1} (-1)^{n-(i+1)} (i-1)!$$

$$= (-1)^{n+1} + (n+1) \sum_{j=0} \binom{n}{j} (-1)^{n-j} j!$$

$$= (-1)^{n+1} + (n+1) d_n$$

$$= (-1)^{n+1} + (n+1) a_n = a_{n+1}.$$

Rappel

On a

$$\binom{n}{k} = \frac{n}{k} \binom{n-1}{k-1}.$$

Par le principe de récurrence, pour tout $n \in \mathbf{N}$, $d_n = a_n$.

3.a. C'est la formule de Taylor avec reste intégral à l'ordre n appliquée à la fonction exponentielle[2] entre 0 et x.

[2] Ce qui est légitime puisque l'exponentielle est de classe $\mathscr{C}^\infty$.

3.b. Notons que $d_n = \sum_{i=0}^{n} \binom{n}{i} (-1)^{n-i} i! = n! \sum_{i=0}^{n} \frac{(-1)^{n-i}}{(n-i)!} = n! \sum_{i=0}^{n} \frac{(-1)^i}{i!}$.
Mais pour $x = -1$, la formule de la question précédente devient

$$e^{-1} = \sum_{k=0}^{n} \frac{(-1)^k}{k!} + \int_0^{-1} \frac{(-1-t)^n}{n!} e^t \, dt = \frac{d_n}{n!} + \int_0^{-1} \frac{(-1-t)^n}{n!} \, dt.$$

En procédant au changement de variable $u = 1 + t$, on a donc

$$\int_0^{-1} \frac{(-1-t)^n}{n!} \, dt = -\int_0^1 \frac{(-u)^n}{n!} e^{u-1} \, du = (-1)^{n+1} \frac{J_n}{e}.$$

Et par conséquent,

$$\frac{1}{e} = \frac{d_n}{n!} + (-1)^{n+1} \frac{J_n}{e} \Leftrightarrow d_n = \frac{1}{e} \left(n! + (-1)^n J_n \right).$$

Remarque

Notons que si l'équivalent donné dans l'énoncé est juste, on a forcément $J_n \to 0$.

3.c. Commençons par remarquer que $\lim_{n \to +\infty} J_n = 0$.
En effet, on a, pour $x \in [0,1]$, $e^x \leqslant e$ et donc

$$0 \leqslant J_n \leqslant \int_0^1 x^n \, dx = \frac{1}{n+1} \underset{n \to +\infty}{\longrightarrow} 0.$$

D'autre part, une intégration par parties, en posant $u(x) = x^{n+1}$ et $v(x) = e^x$ nous donne

$$nJ_n = \int_0^1 nx^n e^x \, dx = \left[x^{n+1} e^x \right]_0^1 - \int_0^1 x^{n+1} e^x \, dx = e - J_{n+1}.$$

Et donc $\frac{nJ_n}{e} = 1 - \frac{J_{n+1}}{e} \underset{n \to +\infty}{\longrightarrow} 1$, de sorte que $J_n \underset{n \to +\infty}{\sim} \frac{e}{n}$.
On a alors $(-1)^n J_n \underset{n \to +\infty}{=} o(n!)$, et donc

$$d_n = \frac{n!}{e} + \underset{n \to +\infty}{o}(n!) \underset{n \to +\infty}{\sim} \frac{n!}{e}.$$

3.d. On a

$$\frac{(n-1)!}{e} - \frac{d_n}{n} = \frac{(n-1)!}{e} - \frac{(n-1)!}{e} + \frac{(-1)^{n+1} J_n}{ne} = (-1)^{n+1} \frac{J_n}{ne}.$$

Or, $\left| (-1)^{n+1} \frac{J_n}{ne} \right| \underset{n \to +\infty}{\sim} \frac{1}{n^2}$, et donc la série de terme général $\frac{(n-1)!}{e} - \frac{d_n}{n}$ converge absolument, et donc converge.

□

EXERCICE 1.18 (QSP HEC 2016) [HEC16.QSP176] — Moyen

Inversibilité d'une matrice

Soit $n \geqslant 2$, et soit $M = (m_{i,j})_{1\leqslant i,j\leqslant n} \in \mathcal{M}_n(\mathbf{R})$ la matrice définie par $\forall (i,j) \in [\![1,n]\!]^2$, $m_{i,j} = i^{j-1}$.
Montrer que M est inversible

SOLUTION DE L'EXERCICE 1.18 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 $i \in [\![1,n]\!]$, $\displaystyle\sum_{j=1}^{n} x_j i^{j-1} = 0$.

Si on note $P = \displaystyle\sum_{j=1}^{n} x_j X^{j-1} \in \mathbf{R}_{n-1}[X]$, cela signifie donc que pour tout $i \in [\![1,n]\!]$, $P(i) = 0$.
Ainsi, P possède n racines distinctes, et étant de degré inférieur ou égal à $n-1$, il est donc nul : $\forall j \in [\![1,n]\!]$, $x_j = 0$.
Donc nécessairement $X = 0$, et donc M est inversible.

> **Rappel**
> Une matrice $M \in \mathcal{M}_n(\mathbf{K})$ est inversible si et seulement si pour tout $X \in \mathcal{M}_{n,1}(\mathbf{K})$,
> $MX = 0 \Rightarrow X = 0$.

Solution alternative : on peut remarquer que M est la matrice dans les bases canoniques de l'application linéaire $\varphi : \mathbf{R}_{n-1}[X] \to \mathbf{R}^n$ 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 $\mathbf{R}^n$:

$$\mathrm{Mat}(\varphi) = \begin{array}{c} \begin{matrix} \varphi(1) & \varphi(X) & \dots & \varphi(X^{n-1}) \end{matrix} \\ \begin{pmatrix} 1 & 1 & \dots & 1 \\ 1 & 2 & \dots & 2^{n-1} \\ \vdots & \vdots & & \vdots \\ 1 & n & \dots & n^{n-1} \end{pmatrix} \end{array} \begin{matrix} e_1 \\ e_2 \\ \vdots \\ e_n \end{matrix} = M.$$

Mais il est classique que φ est un isomorphisme, et donc M est inversible. □

Il est classique qu'une base de $\mathcal{M}_n(\mathbf{R})$ est formée des matrices $E_{i,j}$, où $E_{i,j}$ est la matrice dont tous les coefficients sont nuls, sauf celui de la $i^{\text{ème}}$ ligne et $j^{\text{ème}}$ colonne qui vaut 1.
Les exercices suivants nécessitent de calculer le produit de deux de ces matrices, ce qui demande un peu de rigueur.

EXERCICE 1.19 (QSP ESCP 2018) [ESCP18.QSP05] — Moyen

Une base de $\mathcal{M}_n(\mathbf{R})$ formée de matrices de projecteurs

On note $(E_{i,j})_{1\leqslant i,j\leqslant n}$ la base canonique de $\mathcal{M}_n(\mathbf{R})$. On rappelle que $E_{i,j}$ est la matrice dont tous les coefficient sont nuls, à l'exception de celui situé à la $i^{\text{ème}}$ ligne et $j^{\text{ème}}$ colonne, qui vaut 1.

1. Soit $(i,j) \in [\![1,n]\!]^2$, avec $i \neq j$. Calculer $(E_{i,i} + E_{i,j})^2$.
2. En déduire une base de $\mathcal{M}_n(\mathbf{R})$ formée de matrices de projecteurs.

SOLUTION DE L'EXERCICE 1.19

1. Un piège dans lequel il ne faudrait surtout pas tomber : utiliser le binôme de Newton. En effet, nous ne savons si $E_{i,i}$ et $E_{i,j}$ commutent, hypothèse indispensable pour appliquer le binôme.
On a donc

$$(E_{i,i}+E_{i,j})^2=(E_{i,i}+E_{i,j})(E_{i,i}+E_{i,j})=E_{i,i}^2+E_{i,i}E_{i,j}+E_{i,j}E_{i,i}+E_{i,j}^2.$$

Puisque $E_{i,i}$ est diagonale, il n'est pas bien dur de se convaincre que $E_{i,i}^2=E_{i,i}$.
Ensuite, puisque les colonnes autres que la $j^{\text{ème}}$ de $E_{i,j}$ sont nulles, toutes les colonnes du produit $E_{i,i}E_{i,j}$, sauf la $j^{\text{ème}}$ sont nulles.
De même : toutes les lignes de $E_{i,i}$, sauf la $i^{\text{ème}}$ sont nulles, donc toutes les lignes, sauf la $i^{\text{ème}}$ du produit $E_{i,i}E_{i,j}$ sont nulles.
Ainsi, tous les coefficients de $E_{i,i}E_{i,j}$ sont nuls, sauf peut-être celui situé à la $i^{\text{ème}}$ ligne et $j^{\text{ème}}$ colonne.
Et de fait, celui-ci vaut 1. Donc $E_{i,i}E_{i,j}=E_{i,j}$.

On prouve de manière similaire que $E_{i,j}E_{i,i}=0$, et que $E_{i,j}^2=0$.
Et donc il vient $(E_{i,i}+E_{i,j})^2=E_{i,i}+E_{i,j}$.

2. Nous venons de prouver que pour $i\neq j$, $E_{i,i}+E_{i,j}$ est une matrice de projecteur[1].
D'autre part, nous savons que les $E_{i,i}$ $1\leqslant i\leqslant n$ sont également des matrices de projecteurs.

Et donc la famille $\mathscr{B}$ formée des $E_{i,i}$ et des $E_{i,i}+E_{i,j}$, $i\neq j$ est une famille de $n^2=\dim\mathscr{M}_n(\mathbf{R})$ matrices de $\mathscr{M}_n(\mathbf{R})$, toutes étant des matrices de projecteurs.
Il suffit donc de prouver que la famille $\mathscr{B}$ est libre pour répondre à la question posée.
Soient donc $\alpha_1,\dots,\alpha_n,\beta_{1,2},\beta_{1,3},\dots,\beta_{n,n-1}$ des réels tels que

$$\sum_{i=1}^{n}\alpha_iE_{i,i}+\sum_{i=1}^{n}\sum_{\substack{j=1\\j\neq i}}^{n}\beta_{i,j}(E_{i,i}+E_{i,j})=0.$$

Alors

$$\sum_{i=1}^{n}\left(\sum_{\substack{j=1\\j\neq i}}^{n}\beta_{i,j}E_{i,j}+\left(\alpha_i+\sum_{\substack{j=1\\j\neq i}}^{n}\beta_{i,j}\right)E_{i,i}\right)=0.$$

Mais la famille des $E_{i,j}$, $1\leqslant i,j\leqslant n$ est libre[2], et donc pour $i\neq j$, $\beta_{i,j}=0$, et de plus pour tout i, $\alpha_i+\sum_{\substack{j=1\\j\neq i}}^{n}\beta_{i,j}=0\Rightarrow\alpha_i=0$.

Ainsi, la famille $\mathscr{B}$ est libre, et donc[3] est une base de $\mathscr{M}_n(\mathbf{R})$, formée de matrices de projecteurs.

□

Danger !
En réalité, les hypothèses du binôme matriciel sont bien souvent connues, l'erreur est plutôt de ne pas reconnaître un binôme : l'identité remarquable

$$(a+b)^2=\dots$$

n'est rien d'autre qu'un cas particulier du binôme de Newton !

Détails
Il s'agit ici de bien comprendre le produit matriciel : le coefficient (i,j) de $A\times B$ est obtenu à partir de la $i^{\text{ème}}$ ligne de A et de la $j^{\text{ème}}$ colonne de B.

[1] Autrement dit, dans n'importe quelle base, l'endomorphisme qu'elle représente est un projecteur.

Dénombrement
Il y a n matrices de la forme $E_{i,i}$ et $n\times(n-1)$ de la forme $E_{i,i}+E_{i,j}$. En effet, il y a n choix pour i, et une fois i choisi, il reste $n-1$ choix pour j.

[2] C'est une base de $\mathscr{M}_n(\mathbf{R})$.

[3] Par cardinalité.

Exercice 1.20 (ESCP 2008) [ESCP08.2.4] — Moyen

Hyperplans de $\mathscr{M}_n(\mathbf{R})$ stables pour le produit matriciel et ne contenant pas l'identité.
Abordable en première année : ✓

Soit n un entier supérieur ou égal à 2. Soit F un sous-espace vectoriel de $\mathscr{M}_n(\mathbf{R})$, 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 $\mathscr{M}_n(\mathbf{R})$.

1. Montrer que $\mathscr{M}_n(\mathbf{R})=F\oplus\text{Vect}(I_n)$.
2. (a) Soit p le projecteur de $\mathscr{M}_n(\mathbf{R})$ sur $\text{Vect}(I_n)$ parallèlement à F.
   Montrer que $\forall(M,M')\in(\mathscr{M}_n(\mathbf{R}))^2$, $p(MM')=p(M)P(M')$.
   (b) Montrer que pour tout matrice M de $\mathscr{M}_n(\mathbf{R})$ telle que $M^2\in F$, alors $M\in F$.
   (c) Soit $(E_{i,j})_{1\leqslant i,j\leqslant n}$ la base canonique de $\mathscr{M}_n(\mathbf{R})$.
   Calculer $E_{i,j}\times E_{k,\ell}$ pour $(i,j,k,\ell)\in[\![1,n]\!]^2$.
   (d) Montrer que F contient la base canonique de $\mathscr{M}_n(\mathbf{R})$.

(e) Conclure.

SOLUTION DE L'EXERCICE 1.20

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 $I_n = \dfrac{1}{\lambda} M \in F$, 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 $p(MM') = \lambda_1 \lambda_2 I_n = p(M)p(M')$.

> **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 $\text{Vect}(I_n)$.

2.b. Notons que $F = \text{Ker}(p)$.
Et donc $M^2 \in F \Leftrightarrow p\left(M^2\right) = 0 \Leftrightarrow p(M)p(M) = 0$.
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 $A = (a_{i,j})_{1 \leqslant i,j \leqslant n}$, on a

$$\left(AE_{k,\ell}\right)_{u,v} = \sum_{j=1}^{n} A_{u,j} (E_{k,\ell})_{j,v}.$$

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 $(AE_{k,\ell})_{u,\ell} = a_{u,k}$.
Ainsi, $(AE_{k,\ell})_{u,v} = \begin{cases} 0 & \text{si } v \neq \ell \\ a_{u,k} & \text{si } v = \ell \end{cases}$

> **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.

En particulier, $(E_{i,j}E_{k,\ell})_{u,v} = \begin{cases} 0 & \text{si } v \neq \ell \\ (E_{i,j})_{u,\ell} & \text{sinon} \end{cases}$
Mais puisque $(E_{i,j})_{u,\ell} = \begin{cases} 1 & \text{si } u = i \text{ et } j = \ell \\ 0 & \text{sinon} \end{cases}$ il vient

$$E_{i,j}E_{k,\ell} = \begin{cases} 0 & \text{si } j \neq k \\ E_{i,\ell} & \text{si } j = k \end{cases}$$

> **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.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 $E_{i,j}$, $1 \leqslant i, j \leqslant n$ sont dans F.

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.

□

1.6 VALEURS PROPRES, DIAGONALISATION

Pour un endomorphisme f d'un espace vectoriel de dimension finie, il est classique que f est bijectif si et seulement si $\text{Ker } f = \{0_E\}$.

Mais ceci se formule également en termes de valeurs propres : f est bijectif si et seulement si 0 n'est pas valeur propre de f.

Exercice 1.21 (QSP ESCP 2014) [ESCP14.QSP01] — Facile

Étude d'un endomorphisme de $\mathcal{M}_n(\mathbf{R})$.

Soit $n \in \mathbf{N}^*$, soit A une matrice donnée de $\mathcal{M}_n(\mathbf{R})$ et soit $\phi : \mathcal{M}_n(\mathbf{R}) \to \mathbf{R}$ une forme linéaire.
On définit l'application $f : \mathcal{M}_n(\mathbf{R}) \to \mathcal{M}_n(\mathbf{R})$ par

$$\forall M \in \mathcal{M}_n(\mathbf{R}),\ f(M) = M - \phi(M)A.$$

Donner une condition nécessaire et suffisante sur $\phi(A)$ pour que f soit bijective.

Solution de l'exercice 1.21 Nous savons que f est bijective si et seulement si $0 \notin \text{Spec}(f)$.
Or, il existe un vecteur propre M de f associé à la valeur propre 0 si et seulement si il existe $M \in \mathcal{M}_n(\mathbf{R})$, non nulle, telle que $f(M) = 0 \Leftrightarrow M = \phi(M)A$.
Une telle matrice est donc nécessairement colinéaire à A, donc de la forme λA, avec λ non nul[1]. Et alors $M = \phi(M)A \Leftrightarrow \lambda A = \phi(\lambda A)A \Leftrightarrow A = \phi(A)A \Leftrightarrow \phi(A) = 1$.
Donc 0 est valeur propre de f si et seulement si $\phi(A) = 1$, et donc f est bijective si et seulement si $\phi(A) \neq 1$. □

[1] Car $M \neq 0$ par hypothèse.

La première partie de l'exercice qui suit est assez classique, on y prouve (ici par des moyens matriciels, mais cela peut également se faire en termes d'endomorphisme, cf le problème 2 d'EML 2009) que si une matrice A est diagonalisable et possède des valeurs propres deux à deux distinctes, alors toute matrices B qui commute à A est également diagonalisable, et il existe une base de vecteurs propres commune à A et à B

Exercice 1.22 (ESCP 2012) [ESCP12.2.1] — Facile

Commutant d'une matrice de $\mathcal{M}_n(\mathbf{C})$ admettant n valeurs propres distinctes.

Soit $A \in \mathcal{M}_n(\mathbf{C})$ admettant n valeurs propres distinctes.

1. Montrer que la famille $(I_n, A, A^2, \ldots, A^{n-1})$ est libre.
2. On note $\mathcal{C} = \{M \in \mathcal{M}_n(\mathbf{C}) / AM = MA\}$. Montrer que $\mathcal{C}$ est un sous-espace vectoriel de $\mathcal{M}_n(\mathbf{C})$ de dimension supérieur ou égale à n.
3. Montrer l'existence d'une matrice P de $\mathcal{M}_n(\mathbf{C})$, inversible et d'une matrice Δ de $\mathcal{M}_n(\mathbf{C})$ diagonale, telles que
$$A = P\Delta P^{-1}.$$
4. Soit $M \in \mathcal{C}$. Montrer que tout vecteur colonne propre de A est un vecteur colonne propre de M. En déduire que la matrice $P^{-1}MP$ est diagonale.
En déduire que $\mathcal{C}$ est de dimension inférieure ou égale à n.
5. Montrer que $(I_n, A, \ldots, A^{n-1})$ est une base de $\mathcal{C}$.
6. Soit $A = \begin{pmatrix} 1 & 2 \\ 0 & -1 \end{pmatrix} \in \mathcal{M}_2(\mathbf{C})$. On note $\mathcal{R} = \{M \in \mathcal{M}_2(\mathbf{C}) / M^2 = A\}$.
   (a) Montrer que $\mathcal{R} \subset \text{Vect}(I, A)$.
   (b) Montrer que $\mathcal{R}$ est de cardinal 4. Déterminer les 4 matrices vérifiant $M^2 = A$.

Solution de l'exercice 1.22

1. Soient $\lambda_0, \lambda_1, \ldots, \lambda_{n-1}$ des complexes tels que $\lambda_0 I_n + \lambda_1 A + \cdots + \lambda_{n-1}A^{n-1} = 0$.
Alors le polynôme $P = \sum_{i=0}^{n-1} \lambda_i X^i$ est annulateur de A, et donc les valeurs propres de A sont parmi les racines de P.

Mais P est de degré au plus $n-1$, et toutes les valeurs propres de A, qui sont au nombre de n, en sont racines.
Donc P est le polynôme nul, et par identification des coefficients, $\lambda_0 = \lambda_1 = \cdots = \lambda_{n-1} = 0$.
Et donc $(I_n, A, \dots, A^{n-1})$ est une famille libre de $\mathcal{M}_n(\mathbf{C})$.

2. Soient M, N deux matrices de $\mathscr{C}$ et soit $\lambda \in \mathbf{C}$. Alors

$$A(\lambda M + N) = \lambda AM + AN = \lambda MA + NA = (\lambda M + N)A$$

et donc $\lambda M + N \in \mathscr{C}$.
D'autre part, pour tout $k \in \mathbf{N}$, A^k commute avec A, et donc $A^k \in \mathscr{C}$.
Ceci est en particulier vrai pour $k \in [\![0, n-1]\!]$, de sorte que $\text{Vect}(I_n, A, \dots, A^{n-1}) \subset \mathscr{C}$.
Or, la famille $I_n, A, \dots, A^{n-1}$ étant libre, $\dim \text{Vect}(I_n, A, \dots, A^{n-1}) = n$, et donc $\dim \mathscr{C} \geqslant n$.

3. Puisque A admet n valeurs propres distinctes, elle est diagonalisable, donc il existe P inversible et Δ diagonale telles que $A = P\Delta P^{-1}$.

4. Soit λ une valeur propre de A et soit $X \in \mathcal{M}_{n,1}(\mathbf{C})$ un vecteur propre associé, de sorte que $AX = \lambda X$.
On a alors $A(MX) = MAX = \lambda MX$, de sorte que $MX \in E_\lambda(A)$.
Mais puisque A possède n valeurs propres, ses sous-espaces propres sont tous de dimension 1.
Et X étant un vecteur non nul de $E_\lambda(A)$, c'est donc une base de $E_\lambda(A)$.
Par conséquent, il existe un complexe μ tel que $MX = \mu X$, et donc X est un vecteur propre de M.

S.e.v.
Par définition, $\text{Vect}(I_n, A, \dots, A^{n-1})$ est le plus petit sous-espace vectoriel de $\mathcal{M}_n(\mathbf{C})$ contenant $I_n, A, \dots, A^{n-1}$.
Et puisque nous avons déjà prouvé que $\mathscr{C}$ est un sous-espace vectoriel contenant ces matrices, on a donc bien

$$\text{Vect}(I_n, \dots, A^{n-1}) \subset \mathscr{C}.$$

Plus généralement
La même preuve montre en fait que si deux matrices (ou deux endomorphismes) commutent, tout sous-espace propre de l'une est stable par l'autre.

Ainsi, la matrice P, qui est la matrice de passage de la base canonique de $\mathcal{M}_{n,1}(\mathbf{R})$ à une base de vecteurs propres de A, est également la matrice de passage de la base canonique à une base de vecteurs propres de M.
Et donc $P^{-1}MP$ est une matrice diagonale.
Ainsi, $f : M \mapsto P^{-1}MP$ est une application linéaire de $\mathscr{C}$ dans l'ensemble $\mathscr{D}$ des matrices diagonales.
On constate facilement qu'elle est injective, et donc par le théorème du rang, $\dim \mathscr{C} = \dim \text{Ker } f + \dim \text{Im } f = \dim \text{Im } f \leqslant \dim \mathscr{D} = n$.
Remarque : *en fait, il ne serait pas beaucoup plus dur de prouver que f est surjective, et donc, puisqu'un isomorphisme préserve la dimension, on aurait directement que* $\dim \mathscr{C} = \dim \mathscr{D}$.

Détails
Pour vraiment prouver cela rigoureusement, il faudrait probablement repasser par des endomorphismes (par exemple introduire $X \mapsto AX$, endomorphisme de $\mathcal{M}_{n,1}(\mathbf{C})$) et des formules de changement de base.
Mais on peut raisonnablement supposer qu'un candidat admissible à une parisienne possède suffisamment de recul sur la diagonalisation pour savoir que $P^{-1}MP$ est diagonale si et seulement si les colonnes de P forment une base de vecteurs propres de M.

5. Nous avons prouvé d'une part que $\dim \mathscr{C} \geqslant n$ et d'autre part que $\dim \mathscr{C} \leqslant n$, et donc $\dim \mathscr{C} = n$.
Puisque $(I_n, A, \dots, A^{n-1})$ est une famille libre de $\mathscr{C}$, de cardinal $n = \dim \mathscr{C}$, c'est une base de $\mathscr{C}$.

6.a. Si $M \in \mathscr{R}$, alors $MA = MM^2 = M^2M = AM$.
Et donc, avec les notations précédentes, $\mathscr{R} \subset \mathscr{C}$.
Mais A possède deux valeurs propres distinctes, qui sont 1 et -1, de sorte que par la question 5, $\mathscr{C} = \text{Vect}(I_2, A)$ et donc $\mathscr{R} \subset \text{Vect}(I_2, A)$.

6.b. Soit $M = \lambda I + \mu A \in \text{Vect}(I_2, A)$. On a alors

$$M^2 = \lambda^2 I + 2\lambda\mu A + \mu^2 \underbrace{A^2}_{=I} = (\lambda^2 + \mu^2)I + 2\lambda\mu A.$$

Et donc $M^2 = A$ si et seulement si $\lambda^2 + \mu^2 = 0$ et $2\lambda\mu = 1$.
Soit encore $\mu = \dfrac{1}{2\lambda}$ et $\lambda^2 + \dfrac{1}{4\lambda^2} = 0 \Leftrightarrow \lambda^4 = -\dfrac{1}{4}$.
Et donc en particulier, $\lambda^2 = \pm i\dfrac{1}{2}$.
Pour $\lambda^2 = i\dfrac{1}{2} = \dfrac{1}{2}e^{i\frac{\pi}{2}}$, on obtient $\lambda = \dfrac{1}{\sqrt{2}}e^{i\frac{\pi}{4}} = \dfrac{1+i}{2}$ ou $\lambda = -\dfrac{1+i}{2}$.
Les valeurs de μ correspondantes sont respectivement $\dfrac{1-i}{2}$ et $\dfrac{-1+i}{2}$.
Et pour $\lambda^2 = -\dfrac{i}{2} = \dfrac{1}{2}e^{-i\frac{\pi}{2}}$, on obtient $\lambda = \dfrac{1}{\sqrt{2}}e^{-i\frac{\pi}{4}} = \dfrac{1-i}{2}$ et $\lambda = \dfrac{-1+i}{2}$.
Les valeurs de μ correspondantes sont alors $\dfrac{1+i}{2}$ et $\dfrac{-1-i}{2}$.
Donc au final, les 4 matrices M telles que $M^2 = A$ sont

$$\begin{pmatrix} 1 & 1-i \\ 0 & i \end{pmatrix}, \begin{pmatrix} -1 & i-1 \\ 0 & -i \end{pmatrix}, \begin{pmatrix} 1 & 1+i \\ 0 & -i \end{pmatrix}, \begin{pmatrix} -1 & -1-i \\ 0 & i \end{pmatrix}.$$

Méthode
Pour déterminer une racine carrée d'un nombre complexe z, le plus simple est de passer par la forme exponentielle : $z = re^{i\theta}$.
L'une des racines est alors $\sqrt{r}e^{i\theta/2}$, et la seconde est l'opposée de la première.

□

Exercice 1.23 (QSP ESCP 2018) [ESCP18.QSP02] — Facile

Un endomorphisme commutant à un endomorphisme possédant $\dim E$ valeurs propres est diagonalisable.

Soit $n \geqslant 2$ et soient f et g deux endomorphismes de $\mathbf{R}^n$ tels que $f \circ g = g \circ f$.
On suppose de plus que f admet n valeurs propres deux à deux distinctes.

1. Montrer que les sous-espaces propres de f sont stables par g.
2. En déduire que g est diagonalisable.

Solution de l'exercice 1.23

1. Soit λ une valeur propre de f, et soit $x \in E_\lambda(f)$.
Alors il s'agit de prouver que $g(x) \in E_\lambda(f) \Leftrightarrow f(g(x)) = \lambda g(x)$.
Mais $f(x) = \lambda x$, de sorte que $g(f(x)) = \lambda g(x)$ et donc[1] $f(g(x)) = \lambda g(x)$.
Ainsi, $E_\lambda(f)$ est stable par g.

[1] f et g commutent.

2. Puisque f possède $n = \dim \mathbf{R}^n$ valeurs propres, il est diagonalisable, et tous ses sous-espaces propres sont de dimension 1.
Soit alors $(e_1, \dots, e_n)$ une base de $\mathbf{R}^n$ formée de vecteurs propres de f, et notons λ_i la valeur propre associée à e_i.
Pour tout $i \in [\![1, n]\!]$, $E_{\lambda_i}(f)$ est de dimension 1, et contient e_i, de sorte que $E_{\lambda_i}(f) = \text{Vect}(e_i)$.
Mais nous venons de prouver que $g(e_i) \in E_{\lambda_i}(f)$, et par conséquent, il existe un réel α_i tel que $g(e_i) = \alpha_i e_i$.
Ceci prouve donc que $(e_1, \dots, e_n)$ est une base de $\mathbf{R}^n$ formée de vecteurs propres de g, ce qui garantit donc que g est diagonalisable.

> **Remarque**
> Nous avons prouvé au passage que tous les vecteurs propres de f sont également des vecteurs propres de g.

□

Si f et g sont deux endomorphismes de E, et s'il existe une base de E formée simultanément de vecteurs propres de f et de g, alors les matrices de f et g dans cette base sont toutes deux diagonales.
Mais deux matrices diagonales commutent toujours, de sorte que f et g commutent.
L'exercice suivant prouve que la réciproque est vraie : si f et g sont deux endomorphismes diagonalisables qui commutent, alors il existe une base formée de vecteurs propres de f et de g.

Exercice 1.24 (ESCP 2016) [ESCP16.2.10] — Moyen

Diagonalisation simultanée de deux endomorphismes

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 $(x_1, \dots, x_p) \in E_1 \times \cdots \times E_p$ tel que $x = x_1 + \cdots + x_p$.
2. On suppose désormais $x \neq 0$. Montrer que, quitte à modifier l'ordre, on peut supposer qu'il existe $r \in [\![1, p]\!]$ tel que $x_i = 0$ pour $i > r$ et $x_i \neq 0$ pour $i \leqslant r$. On a alors $x = x_1 + \cdots + x_r$.
   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 $j \geqslant 0$, $f^j(x) \in V_x$.
   (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 $P = \sum_{j=1}^{r} \alpha_j X^{j-1}$ est le polynôme nul. En déduire que A est inversible.

(d) Montrer que pour tout $i \in [\![1,p]\!]$, $x_i \in F$, 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.

SOLUTION DE L'EXERCICE 1.24

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 $x_{r+1} = \cdots = x_p = 0$.

> **Rappel**
> Le sous-espace propre est formé des vecteurs propres de f associés à la valeur propre λ_i et du vecteur nul.

3.a. Pour $i \in [\![1,r]\!]$, $x_i \in E_i$ 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, elle est génératrice de V_x, et donc c'est une base de V_x.

3.b. Pour $j \geqslant 0$, 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

$$\mathrm{Mat}_{(x_1,\dots,x_r)}(x, f(x), \dots, f^{r-1}(x)) = \begin{pmatrix} x & f(x) & \dots & f^{r-1}(x) \\ 1 & \lambda_1 & \dots & \lambda_1^{r-1} \\ 1 & \lambda_2 & \dots & \lambda_2^{r-1} \\ 1 & \vdots & & \vdots \\ 1 & \lambda_r & \dots & \lambda_r^{r-1} \end{pmatrix} \begin{matrix} \\ x_1 \\ x_2 \\ \vdots \\ x_r \end{matrix}.$$

Notons que $C_j = \begin{pmatrix} \lambda_1^{j-1} \\ \lambda_2^{j-1} \\ \vdots \\ \lambda_r^{j-1} \end{pmatrix}$ et donc

$$\sum_{j=1}^{r} \alpha_j C_j = 0 \iff \begin{pmatrix} \sum_{j=1}^{r} \alpha_j \lambda_1^{j-1} \\ \sum_{j=1}^{r} \alpha_1 \lambda_2^{j-1} \\ \vdots \\ \sum_{j=1}^{r} \alpha_1 \lambda_r^{j-1} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \iff \begin{pmatrix} P(\lambda_1) \\ P(\lambda_2) \\ \vdots \\ P(\lambda_r) \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$$

Et donc $P(\lambda_1) = \cdots = P(\lambda_r) = 0$.
Or P est de degré au plus $r-1$ et possède r racines, c'est donc qu'il est nul.
On en déduit[1] que $\alpha_1 = \cdots = \alpha_r = 0$.
Et donc la famille $(C_1, \dots, C_r)$ est libre. Elle est donc de rang r, de sorte que $\mathrm{rg}(A) = r$.
Et donc A est une matrice inversible.

[1] Par identification des coefficients.

> **Rang**
> Rappelons que par définition, le rang d'une matrice est le rang de la famille de ses vecteurs colonnes.

3.d. Puisque A est inversible, $(x, f(x), \dots, f^{r-1}(x))$ est une base de V_x.
En particulier, pour tout $i \in [\![1,r]\!]$, 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 : $x_i \in F$.

Pour le second point, commençons par prouver que les $F \cap E_i$ sont en somme directe.
Supposons donc que $0_E = y_1 + \cdots + y_p$, 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

$y_1 = \cdots = y_p = 0$.

De plus, nous venons de montrer que si $x \in F$, alors $x = x_1 + \cdots + x_p$, avec $x_i \in F \cap E_i$.

Donc $F = \sum_{i=1}^{p}(F \cap E_i)$ et puisque la somme est directe, $F = \bigoplus_{i=1}^{p}(F \cap E_i)$.

> **Remarque**
> Nous avons montré que $x_1, \ldots, x_r$ sont dans F, pour $x_{r+1}, \ldots, x_p$, c'est trivial car ils sont nuls et que le vecteur nul est dans F.

4. Soient $F_1, \ldots, F_q$ les sous-espaces propres de g.
Puisque f et g commutent, les F_j sont stables par f.

Et donc pour tout $j \in [\![1, q]\!]$, $F_j = \bigoplus_{i=1}^{p}(F_j \cap E_i)$.

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 $\mathrm{Mat}_{\mathcal{B}}(f)$ et $\mathrm{Mat}_{\mathcal{B}}(g)$ soient toutes les deux diagonales.

□

Le thème de l'exercice suivant est relativement classique, on le retrouve par exemple dans EML 2016.
Il s'agit de décomposer un endomorphisme diagonalisable comme combinaison linéaire de projecteurs sur ses sous-espaces propres.

Exercice 1.25 (HEC 2014) [HEC14.110] — Moyen

Projecteurs spectraux d'un endomorphisme diagonalisable

Soit E un espace vectoriel de dimension finie $n \geqslant 2$ sur $\mathbf{R}$. Soit f un endomorphisme de E pour lequel il existe un entier $m \geqslant 2$, des endomorphismes $p_1, p_2, \ldots, p_m$ non nuls de E et m réels distincts $\lambda_1, \lambda_2, \ldots, \lambda_m$ tels que pour tout entier $k \in \mathbf{N}$, on a : $f^k = \sum_{i=1}^{m} \lambda_i^k p_i$, où $f^k = \underbrace{f \circ f \circ \cdots \circ f}_{k \text{ fois}}$. On note id l'endomorphisme identité de E.

1. Soit P un polynôme de $\mathbf{R}[X]$. Exprimer $P(f)$ en fonction des $P(\lambda_i)$ et des p_i, pour $i \in [\![1, m]\!]$.
2. Soit Q le polynôme de $\mathbf{R}[X]$ définie par $Q(X) = \prod_{i=1}^{m}(X - \lambda_i)$. Calculer $Q(f)$.
   Qu'en déduit-on quant aux valeurs propres de f ?
3. Pour tout $k \in [\![1, m]\!]$, on pose : $L_k(X) = \prod_{\substack{i \in [\![1,m]\!] \\ i \neq k}} \frac{(X - \lambda_i)}{(\lambda_k - \lambda_i)}$.
   Calculer $L_k(f)$. En déduire que $\mathrm{Im}(p_k) \subset \mathrm{Ker}(f - \lambda_k \mathrm{id})$ ainsi que l'ensemble des valeurs propres de f.
4. Montrer que f est diagonalisable.
5. Vérifier que pour tout couple $(i, j) \in [\![1, m]\!]^2$, on a : $p_i \circ p_j = \begin{cases} 0 & \text{si } i \neq j \\ p_i & \text{si } i = j \end{cases}$
6. Soit F le sous-espace vectoriel de $\mathcal{L}(E)$ engendré par $(p_1, p_2, \ldots, p_m)$. Déterminer la dimension de F.

Solution de l'exercice 1.25

1. Soit $P = \sum_{k=0}^{d} a_k X^k \in \mathbf{R}[X]$. Alors

$$P(f) = \sum_{k=0}^{d} a_k f^k = \sum_{k=0}^{d} a_k \sum_{i=1}^{m} \lambda_i^k p_i = \sum_{i=1}^{m} \sum_{k=0}^{d} a_k \lambda_i^k p_i = \sum_{i=1}^{m} P(\lambda_i) p_i.$$

2. En utilisant le résultat de la question précédente, on a

$$Q(f) = \sum_{i=1}^{m} \underbrace{Q(\lambda_i)}_{=0} p_i = 0.$$

Donc Q est annulateur de f, et par conséquent, les valeurs propres de f sont parmi les racines de Q : $\mathrm{Spec}(f) \subset \{\lambda_1, \dots, \lambda_m\}$.

3. Puisque $L_k(\lambda_j) = \prod_{\substack{i \in [\![1,m]\!] \\ i \neq k}} \frac{\lambda_j - \lambda_i}{\lambda_k - \lambda_i}$, on a $L_k(\lambda_j) = \begin{cases} 0 & \text{si } j \neq k \\ 1 & \text{si } j = k \end{cases}$.

Et donc $L_k(f) = \sum_{i=1}^{m} L_k(\lambda_i) p_i = 1 \cdot p_k$.

Mais $Q = \underbrace{\prod_{\substack{i \in [\![1,m]\!] \\ i \neq k}} (\lambda_k - \lambda_i)}_{\in \mathbf{R}^*} \times L_k \times (X - \lambda_k)$, et donc

$$0 = Q(f) = \prod_{\substack{i \in [\![1,m]\!] \\ i \neq k}} (\lambda_k - \lambda_i) \times (f - \lambda_k \mathrm{id}) \circ L_k(f) = \prod_{\substack{i \in [\![1,m]\!] \\ i \neq k}} (\lambda_k - \lambda_i) \times (f - \lambda_k \mathrm{id}) \circ p_k.$$

Et donc $(f - \lambda_k \mathrm{id}) \circ p_k = 0$. Par conséquent $\mathrm{Im}(p_k) \subset \mathrm{Ker}(f - \lambda_k \mathrm{id})$.
Mais par hypothèse, $p_k \neq 0$, de sorte que $\mathrm{Im}(p_k) \neq \{0_E\}$. Et donc $\mathrm{Ker}(f - \lambda_k \mathrm{id}) \neq \{0_E\}$, de sorte que λ_k est bien valeur propre de f.
On en déduit que $\mathrm{Spec}(f) = \{\lambda_1, \dots, \lambda_m\}$.

Classique
Pour deux endomorphismes f et g, on a $g \circ f = 0$ si et seulement si $\mathrm{Im}(f) \subset \mathrm{Ker}(g)$.

4. Puisque $\mathrm{id} = f^0 = \sum_{i=0}^{m} p_i$, si $x \in E$, alors

$$x = \sum_{i=1}^{m} p_i(x) \in \sum_{i=1}^{m} \mathrm{Im}(p_i) \subset \sum_{i=1}^{m} \mathrm{Ker}(f - \lambda_i \mathrm{id}).$$

Et donc $E = \sum_{i=1}^{m} E_{\lambda_i}(f)$.

Mais nous savons qu'une somme de sous-espaces propres est directe, et donc $E = \bigoplus_{i=1}^{m} E_{\lambda_i}(f)$, de sorte que f est diagonalisable.

5. L'inclusion prouvée à la question 3 montre que si $x \in E$, alors $p_j(x) \in E_{\lambda_j}(f)$ et donc $f(p_j(x)) = \lambda_j p_j(x)$.
Et alors, pour tout polynôme $P \in \mathbf{R}[X]$, $P(f)(p_j(x)) = P(\lambda_j) p_j(x)$.
En particulier, $L_i(f)(p_j(x)) = L_i(\lambda_j) p_j(x)$.
Mais $L_i(\lambda_j) = \begin{cases} 0 & \text{si } i \neq j \\ 1 & \text{si } i = j \end{cases}$

Puisque d'autre part, $L_i(f) = p_i$, il vient donc $p_i(p_j(x)) = L_i(f)(p_j(x)) = \begin{cases} 0 & \text{si } i \neq j \\ p_j(x) & \text{si } i = j \end{cases}$

Et donc on a bien $p_i \circ p_j = \begin{cases} 0 & \text{si } i \neq j \\ p_i & \text{si } i = j \end{cases}$
Ceci prouve en particulier que $p_i^2 = p_i$: les p_i sont des projecteurs.

Rappel
Si $x \in E_\lambda(f)$, alors pour tout $P \in \mathbf{R}[X]$, $P(f)(x) = P(\lambda) \cdot x$.

6. Montrons que les p_i forment une famille libre.
Soient donc $\mu_1, \dots, \mu_m$ des réels tels que $\mu_1 p_1 + \cdots + \mu_m p_m = 0$.
Soit alors x_i non nul[1] dans $\mathrm{Im}(p_i)$. Puisque p_i est un projecteur, $x_i = p_i(x_i)$.
Et donc

$$0 = \left(\sum_{j=1}^{m} \mu_j p_j \right)(x_i) = \sum_{j=1}^{m} \mu_j p_j(p_i(x_i)) = \mu_i p_i^2(x_i) = \mu_i x_i.$$

Puisque $x_i \neq 0_E$, on a donc $\mu_i = 0$.
Ainsi, $\mu_1 = \cdots = \mu_m = 0$, et donc la famille $(p_1, \dots, p_m)$ est libre.
Puisqu'elle est génératrice de F, c'est une base de F, qui est donc de dimension m.

[1] Puisque $\mathrm{Im}(p_i) \neq \{0_E\}$, un tel vecteur existe.

□

Le résultat de l'exercice qui suit est relativement classique.

Exercice 1.26 (QSP HEC 2009) [HEC09.QSP03] **Facile**

Un endomorphisme diagonalisable vérifie $E = \operatorname{Im} f \oplus \operatorname{Ker} f$.

Soit f un endomorphisme d'un espace vectoriel E de dimension n.
Montrer que si $\operatorname{Ker} f$ et $\operatorname{Im} f$ ne sont pas supplémentaires, alors f n'est pas diagonalisable.

Solution de l'exercice 1.26 Nous allons prouver la contraposée, à savoir : si f est diagonalisable, alors $E = \operatorname{Im} f \oplus \operatorname{Ker} f$.
Si f est diagonalisable, alors il existe une base $(e_1, \dots, e_n)$ de E formée de vecteurs propres de f. Notons λ_i la valeur propre associée à e_i.
Soit alors $x \in \operatorname{Im} f \cap \operatorname{Ker} f$.

Alors il existe $y \in E$ tel que $x = f(y)$. De manière unique, $y = \sum_{i=1}^{n} y_i e_i$, où les y_i sont des scalaires[1].
Et alors

$$f(x) = f^2(y) = \sum_{i=1}^{n} y_i f^2(e_i) = \sum_{i=1}^{n} y_i \lambda_i^2 e_i.$$

Mais puisque $x \in \operatorname{Ker} f$, $f(x) = 0_E$. La famille $(e_1, \dots, e_n)$ étant une base de E, on a donc pour tout $i \in [\![1, n]\!]$, $y_i \lambda_i^2 = 0$.
Autrement dit, soit $\lambda_i = 0$ (et donc e_i est un vecteur propre de f associé à la valeur propre 0, *i.e.* $e_i \in \operatorname{Ker} f$), soit $y_i = 0$.
Il vient alors $x = \sum_{i=1}^{n} y_i f(e_i) = \sum_{i=1}^{n} y_i \lambda_i$.
Mais par ce qui précède, pour tout $i \in [\![1, n]\!]$, $y_i \lambda_i = 0$, et donc $x = 0_E$.
On a donc prouvé que $\operatorname{Ker} f \cap \operatorname{Im} f = \{0_E\}$.
D'autre part, d'après le théorème du rang, $\dim E = \dim \operatorname{Im} f + \dim \operatorname{Ker} f$, ce qui prouve que $E = \operatorname{Im} f \oplus \operatorname{Ker} f$.

[1] Réels ou complexes suivant que E est un espace vectoriel sur **R** ou sur **C**.

Seconde méthode : nous pouvons aussi remarquer que f étant diagonalisable,

$$E = \bigoplus_{\lambda \in \operatorname{Spec}(f)} E_\lambda(f) = E_0(f) \oplus \bigoplus_{\substack{\lambda \in \operatorname{Spec}(f) \\ \lambda \neq 0}} E_\lambda(f) = \operatorname{Ker} f \oplus \bigoplus_{\substack{\lambda \in \operatorname{Spec}(f) \\ \lambda \neq 0}} E_\lambda(f).$$

Mais si $\lambda \neq 0$, et si $x \in E_\lambda(f)$, alors $f(x) = \lambda x$, de sorte que $x = \frac{1}{\lambda} f(x) = f\left(\frac{1}{\lambda} x\right) \in \operatorname{Im} f$.
Ceci prouve alors que pour tout $\lambda \in \operatorname{Spec}(f) \setminus \{0\}$, $E_\lambda(f) \subset \operatorname{Im} f$.

□

Les deux exercices qui suivent nécessitent un peu d'intuition pour «deviner» les valeurs propres/vecteurs propres.

Exercice 1.27 (QSP ESCP 2007) [ESCP07.QSP02] **Facile**

Valeurs propres d'un endomorphisme défini à partir d'une forme linéaire

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.

Solution de l'exercice 1.27 Pour tout $x \in \operatorname{Ker}(\varphi)$, on a $f(x) = x$. Donc 1 est valeur propre de f et $\operatorname{Ker}(\varphi) \subset E_1(f)$, qui est donc de dimension au moins égale[1] à

[1] $\operatorname{Ker}(\varphi)$ est un hyperplan de E.

$\dim E - 1$.
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 $\dim E_1(f) + \dim E_{1+\varphi(u)} \geqslant \dim E$, et donc cette inégalité est une égalité, de sorte que $\dim E_1(f) = \dim E - 1$ et $\dim E_{1+\varphi(u)} = 1$.
On en déduit que $E_1(f) = \text{Ker}(\varphi)$ 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 $x = \dfrac{\varphi(x)}{\lambda - 1}u$.
Mais puisque $\varphi(u) = 0$, on a alors $\varphi(x) = \dfrac{\varphi(x)}{\lambda - 1}\varphi(u) = 0$. D'où une contradiction : f ne possède qu'une seule valeur propre qui est 1, et $E_1(f) = \text{Ker}(\varphi)$. □

EXERCICE 1.28 (QSP HEC 2016) [HEC16.QSP01] — Moyen

Diagonalisation d'un endomorphisme de $\mathcal{M}_n(\mathbf{R})$.

Soit f l'endomorphisme de $\mathcal{M}_n(\mathbf{R})$ défini par $f(M) = M + 2{}^tM + 3\text{tr}(M)I_n$. f est-il diagonalisable ?

SOLUTION DE L'EXERCICE 1.28 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.
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 nulle[1], 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 $\varphi : \mathcal{S}_n(\mathbf{R}) \to \mathbf{R}$ 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 sait[2] 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

$$\dim E_{3(n+1)}(f) \geqslant 1, \dim E_{-1}(f) \geqslant \dim \mathcal{A}_n(\mathbf{R}) \text{ et } \dim E_3(f) \geqslant \dim \mathcal{S}_n(\mathbf{R}) - 1.$$

Et donc

$$\dim E_{-1}(f) + \dim E_3(f) + \dim E_{3(n+1)}(f) \geqslant \dim \mathcal{A}_n(\mathbf{R}) + \dim \mathcal{S}_n(\mathbf{R}) = \dim \mathcal{M}_n(\mathbf{R}).$$

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 des sous-espaces propres de f.
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 $C = \lambda I_n$.

[1] Prendre par exemple $D = \text{Diag}(1, \ldots, 1, 1 - n)$.

[2] 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 :

$$\dim \mathcal{A}_n(\mathbf{R}) = \frac{n(n-1)}{2}$$

$$\dim \mathcal{S}_n(\mathbf{R}) = \frac{n(n+1)}{2}.$$

Et alors $\text{tr}(M) = \text{tr}(A) + \text{tr}(B) + \lambda\text{tr}(I_n) = 0 + 0 + \lambda n$.
Et donc $\lambda = \dfrac{\text{tr}(M)}{n}$.
De plus, on a ${}^tM = {}^tA + {}^tB + \lambda I_n = -A + B + \lambda I_n$.
Et donc $A = \dfrac{M - {}^tM}{2}$ et $B = \dfrac{M + {}^tM}{2} - \lambda I_n = \dfrac{M + {}^tM}{2} - \dfrac{\text{tr}(M)}{n} I_n$.
Ainsi, si une telle décomposition $M = A + B + C$ existe, elle est unique.
Inversement, pour $M \in \mathcal{M}_n(\mathbf{R})$, posons

$$A = \frac{M - {}^tM}{2},\ B = \frac{M + {}^tM}{2} - \frac{\text{tr}(M)}{n} I_n \text{ et } C = \frac{\text{tr}(M)}{n} I_n.$$

Alors il est clair que $C \in \text{Vect}(I_n)$, A est antisymétrique car ${}^tA = \dfrac{{}^tM - M}{2} = -A$.
De plus, B est symétrique car ${}^tB = \dfrac{{}^tM + M}{2} - \dfrac{\text{tr}(M)}{n} I_n = B$, et elle est de trace nulle car

$$\text{tr}(B) = \frac{1}{2}\text{tr}({}^tM) + \frac{1}{2}\text{tr}(M) - \frac{\text{tr}(M)}{n}\text{tr}(I_n) = \frac{1}{2}\text{tr}(M) + \frac{1}{2}\text{tr}(M) - \text{tr}(M) = 0.$$

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. □

L'exercice suivant nécessite une bonne compréhension du produit matriciel : la $j^{\text{ème}}$ colonne d'un produit AB de matrices s'obtient en multipliant A par la $j^{\text{ème}}$ colonne de B.
Autrement dit, si $B_1, \dots, B_n$ sont les vecteurs colonnes de B, alors

$$AB = A\left(\begin{array}{c|c|c|c} B_1 & B_2 & \dots & B_n \end{array}\right) = \left(\begin{array}{c|c|c|c} AB_1 & AB_2 & \dots & AB_n \end{array}\right).$$

Exercice 1.29 (QSP HEC 2017) [HEC17.QSP216] — Moyen

Valeurs propres de $M \mapsto AM$

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.

Solution de l'exercice 1.29

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.

Soit alors $M = \begin{pmatrix} x_1 & 0 & \dots & 0 \\ x_2 & 0 & \dots & 0 \\ \vdots & \vdots & & \vdots \\ x_n & 0 & \dots & 0 \end{pmatrix} = \left(\begin{array}{c|c|c|c} x_1 & 0 & \dots & 0 \\ \vdots & \vdots & & \vdots \\ x_n & 0 & & 0 \end{array}\right) \in \mathcal{M}_n(\mathbf{R})$.

On a alors

$$\varphi_A(M) = AM = \left(\begin{array}{c|c|c|c} AX & 0 & \dots & 0 \end{array}\right) = \left(\begin{array}{c|c|c|c} \lambda X & 0 & \dots & 0 \end{array}\right) = \lambda M.$$

Remarque

On a choisi ici de mettre X dans la première colonne de M. Mais on aurait le même résultat en le mettant dans n'importe quelle autre colonne.

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 $(i,j) \in [\![1,n]\!]^2$, notons alors $M_{i,j} = \left(\begin{array}{c|c|c|c|c|c|c} 0 & \dots & 0 & X_i & 0 & \dots & 0 \end{array}\right) \in \mathcal{M}_n(\mathbf{R})$ 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 $(M_{i,j})_{1 \leqslant i,j \leqslant n}$ est une famille libre de $\mathcal{M}_n(\mathbf{R})$.
Soient donc $(\lambda_{i,j})_{1 \leqslant i,j \leqslant n}$ des réels tels que $\sum_{i=1}^{n} \sum_{j=1}^{n} \lambda_{i,j} M_{i,j} = 0$. Alors

$$0 = \left(\begin{array}{c|c|c|c} \displaystyle\sum_{i=1}^{n} \lambda_{i,1} X_i & \displaystyle\sum_{i=1}^{n} \lambda_{i,2} X_i & \dots & \displaystyle\sum_{i=1}^{n} \lambda_{i,n} X_i \end{array}\right).$$

Par identification des colonnes[1], on a donc

$$\forall j \in [\![1,n]\!] \quad \sum_{i=1}^{n} \lambda_{i,j} X_i = 0.$$

Mais la famille $(X_1, \dots, X_n)$ étant libre, il vient donc $\forall (i,j) \in [\![1,n]\!]^2$, $\lambda_{i,j} = 0$.
Ainsi, la famille $(M_{i,j})_{1 \leqslant i,j \leqslant n}$ 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.

□

[1] 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

$$\dim E_\lambda(\varphi_A) = n \times \dim E_\lambda(A).$$

EXERCICE 1.30 (ESCP 2014) [ESCP14.2.15] — Difficile

Racines d'ordre p d'une matrice unipotente

Dans tout cet exercice, n est un entier supérieur ou égal à 2. On note :

$$\mathcal{P}_n(\mathbf{R}) = \left\{A \in \mathcal{M}_n(\mathbf{R}) \,/\, \forall p \in \mathbf{N}^*, \exists B \in \mathcal{M}_n(\mathbf{R}) \text{ tel que } A = B^p\right\}.$$

1. Soit $A = \begin{pmatrix} 1 & 0 & 2 \\ 0 & 2 & 4 \\ 0 & 0 & 4 \end{pmatrix}$.
   (a) Montrer que A est diagonalisable.
   (b) Montrer que $A \in \mathcal{P}_3(\mathbf{R})$.

2. Soit $A = \begin{pmatrix} 1 & 0 \\ 0 & -2 \end{pmatrix}$. Montrer que $A \notin \mathscr{P}_2(\mathbf{R})$.
3. Soit $N \in \mathscr{M}_n(\mathbf{R})$ non nulle vérifiant $N^n = 0$. Montrer que $N \notin \mathscr{P}_n(\mathbf{R})$.
4. Dans cette question, N est une matrice non nulle, telle qu'il existe $k \in \mathbf{N}^*$ tel que $N^k = 0$. On pose $A = I_n + N$.
   (a) Déterminer les valeurs propres de A. La matrice A est-elle diagonalisable ?
   (b) Soit V un polynôme de $\mathbf{R}[X]$. On suppose qu'au voisinage de 0, on a : $V(x) = \underset{x\to 0}{o}(x^q)$, où $q \in \mathbf{N}$. Montrer qu'il existe un polynôme Q tel que $V(X) = X^q Q(X)$.
   (c) Soit $p \in \mathbf{N}^*$. Montrer que pour tout $q \in \mathbf{N}^*$, il existe un polynôme $U_q \in \mathbf{R}[X]$ tel qu'au voisinage de 0, on a : $1 + t = (U_q(x))^p + \underset{x\to 0}{o}(x^q)$.
   On pourra utiliser le développement limité de $(1+x)^\alpha$.
   (d) En déduire que $I_n + N \in \mathscr{P}_n(\mathbf{R})$.

Solution de l'exercice 1.30

1.a. Puisque A est triangulaire, ses valeurs propres sont ses coefficients diagonaux, et donc $\mathrm{Spec}(A) = \{1, 2, 4\}$. Et donc A possède trois valeurs propres distinctes, étant de taille 3, elle est diagonalisable.

1.b. Il existe donc une matrice P inversible telle que $A = P^{-1}\begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 4 \end{pmatrix}P$.

Et alors, pour $p \in \mathbf{N}^*$, si on pose $B_p = P^{-1}\begin{pmatrix} 1 & 0 & 0 \\ 0 & 2^{1/p} & 0 \\ 0 & 0 & 4^{1/p} \end{pmatrix}P$, on a

$$B_p^p = P^{-1}\begin{pmatrix} 1 & 0 & 0 \\ 0 & 2^{1/p} & 0 \\ 0 & 0 & 4^{1/p} \end{pmatrix}^p P = P^{-1}\begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 4 \end{pmatrix}P = A.$$

Et donc pour tout $p \in \mathbf{N}^*$, il existe $B_p \in \mathscr{M}_3(\mathbf{R})$ tel que $B_p^p = A$, de sorte que $A \in \mathscr{P}_3(\mathbf{R})$.

2. Supposons qu'il existe $B = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathscr{M}_2(\mathbf{R})$ telle que $B^2 = A$.

Alors $B^2 = \begin{pmatrix} a^2+bc & ab+bd \\ ac+cd & bc+d^2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & -2 \end{pmatrix}$. Soit encore $\begin{cases} a^2+bc = 1 \\ b(a+d) = 0 \\ c(a+d) = 0 \\ bc+d^2 = -2 \end{cases}$

De la seconde équation, il vient $a + d = 0$ ou $b = 0$.
Mais si $b = 0$, la dernière équation est $d^2 = -2$, qui n'est pas possible.
Donc $a = -d$, de sorte que $a^2 = d^2$, et donc $bc = 1 - a^2 = -2 - a^2$.
Ceci est impossible.
Donc il n'existe pas de matrice $B \in \mathscr{M}_2(\mathbf{R})$ telle que $B^2 = A$, et donc $A \notin \mathscr{P}_2(\mathbf{R})$.

> **Remarque**
> On prouverait de même que pour tout p pair, il n'existe pas de matrice $A \in \mathscr{M}_2(\mathbf{R})$ telle que $B^p = A$.
> En revanche, si p est impair, on peut prendre $B = \begin{pmatrix} 1 & 0 \\ 0 & -2^{1/p} \end{pmatrix}$.

3. Notons k le plus petit entier tel que $N^k = 0$, de sorte que $N^{k-1} \neq 0$.
Puisque $N \neq 0$, on a alors $k \geqslant 2$. Soit alors $B \in \mathscr{M}_n(\mathbf{R})$ telle que $B^n = N$.
On a alors $B^{nk} = N^n = 0$. Soit $p \in \mathbf{N}^*$ le plus petit entier tel que $B^p = 0$. Puisque $B^{n(k-1)} = N^{k-1} \neq 0$, on a donc $p > n(k-1) \geqslant n$.
Considérons donc $X \in \mathscr{M}_{n,1}(\mathbf{R})$ tel que $B^{p-1}X \neq 0$, et montrons que la famille $(X, BX, \dots, B^{p-1}X)$ est libre dans $\mathscr{M}_{n,1}(\mathbf{R})$.

Soient donc $\lambda_0, \lambda_1, \dots, \lambda_{p-1}$ des réels tels que $\sum_{i=0}^{p-1} B^i X = 0$.

Alors, en multipliant à gauche cette égalité par B^{p-1}, il vient $\lambda_0 B^{p-1}X = 0$. Et donc $\lambda_0 = 0$.

Il reste donc $\sum_{i=1}^{p-1} \lambda_i B^i X = 0$.

En multipliant par B^{p-2}, on obtient donc $\lambda_1 B^{p-1}X = 0$, et donc $\lambda_1 = 0$.
De proche en proche, on trouve donc $\lambda_0 = \lambda_1 = \cdots = \lambda_{p-1} = 0$, de sorte que la famille est libre.
Son cardinal, qui est p, est donc inférieur ou égal à la dimension de $\mathscr{M}_{n,1}(\mathbf{R})$, qui vaut n.

Donc $p \leqslant n$, contredisant ce qui a été dit plus tôt.
Par conséquent, $N \notin \mathscr{P}_n(\mathbf{R})$.

4.a. On a $(A - I_n)^k = N^k = 0$. Donc le polynôme $(X-1)^k$ est annulateur de A.
Puisqu'il ne possède que 1 comme racine, c'est donc la seule valeur propre possible de A.
De plus, $A - I_n = N$ n'est pas inversible car sinon N^k le serait, et donc 1 est bien valeur propre de A : $\mathrm{Spec}(A) = \{1\}$.
Si la matrice A était diagonalisable, elle serait semblable à I_n, et donc égale à I_n, ce qui n'est pas le cas puisque $N \neq 0$.

4.b. Procédons par récurrence sur q.
Si $q = 0$, c'est trivial : $Q = V$.
Supposons donc la propriété vraie pour q, et soit V un polynôme tel que $V(x) = \underset{x\to 0}{o}\left(x^{q+1}\right)$.
Puisque $x^{q+1} \underset{x\to 0}{\longrightarrow} 0$, on en déduit que $V(0) = \lim\limits_{x\to 0} V(x) = 0$.
Et donc 0 est racine de V : il existe un polynôme $P \in \mathbf{R}[X]$ tel que

□

EXERCICE 1.31 (ESCP 2007) [ESCP07.2.13] — Facile

Diagonalisation d'un endomorphisme de $\mathbf{R}_n[X]$.

Soit n un entier naturel non nul et Q un polynôme à coefficients réels de degré $d \leqslant n$.
On définit l'application suivante :

$$\varphi : \begin{cases} \mathbf{R}_n[X] & \longrightarrow \mathbf{R}_n[X] \\ P & \longmapsto (PQ)^{(n)} \end{cases}$$

où $(PQ)^{(n)}$ indique que l'on prend la dérivée $n^{\text{ème}}$ du produit de P par Q.

1. Justifier que φ est un endomorphisme de $\mathbf{R}_n[X]$.
2. Donner une condition nécessaire et suffisante sur le polynôme Q pour que φ soit un automorphisme de $\mathbf{R}_n[X]$.
3. Montrer que la matrice de φ dans la base canonique de $\mathbf{R}_n[X]$ est triangulaire supérieure. Que dire de plus dans le cas particulier où $d < n$.
4. Déterminer une condition nécessaire et suffisante sur le polynôme Q pour que φ soit diagonalisable.

SOLUTION DE L'EXERCICE 1.31

1. Soit $P \in \mathbf{R}_n[X]$. Alors $\deg(PQ) = \deg P + \deg Q \leqslant n + d \leqslant 2n$.
Et alors, $\deg(\varphi(P)) \leqslant 2n - n = n$. Ainsi, $\varphi(P) \in \mathbf{R}_n[X]$.
D'autre part, pour $P_1, P_2 \in \mathbf{R}_n[X]$ et $\lambda \in \mathbf{R}$, on a, par linéarité de la dérivation,

$$\varphi(\lambda P_1 + P_2) = ((\lambda P_1 + P_2)Q)^{(n)} = (\lambda P_1 Q + P_2 Q)^{(n)} = \lambda (P_1 Q)^{(n)} + (P_2 Q)^{(n)} = \lambda\varphi(P_1) + \varphi(P_2).$$

Donc φ est linéaire, et donc est un endomorphisme de $\mathbf{R}_n[X]$.

Détails
Il est classique que
$$\deg(P') \leqslant \deg(P) - 1.$$
Et donc une récurrence rapide prouve que pour tout k,
$$\deg(P^{(k)}) \leqslant \deg(P) - k.$$
C'est même une égalité si $\deg P \geqslant k$ (*i.e.* si $P^{(k)} \neq 0$).

2. Puisque $\mathbf{R}_n[X]$ est de dimension finie et que φ en est un endomorphisme, il est bijectif si et seulement si il est injectif, si et seulement si il est surjectif.
Si $d < n$, alors pour tout $P \in \mathbf{R}_n[X]$, $\deg(\varphi(P)) \leqslant n + d - n \leqslant d < n$.
Et donc $\mathrm{Im}\,\varphi \subset \mathbf{R}_d[X]$. En particulier, un polynôme de degré n n'est pas dans l'image de φ, qui n'est donc pas surjectif, et donc pas bijectif.
Si $d = n$, alors pour tout $P \in \mathbf{R}_n[X]$, on a $\deg(\varphi(P)) = \deg P + n - n = \deg P$.
Et en particulier, si $P \neq 0$, $\deg P \geqslant 0$ et donc $\deg\varphi(P) \geqslant 0$, de sorte que $\varphi(P) \neq 0$.
Et donc $\mathrm{Ker}\,\varphi = \{0\}$: φ est injectif, et donc bijectif.

Rappel
Par convention, le degré du polynôme nul est égal à $-\infty$.

On en déduit que φ est un automorphisme si et seulement si $d = n$.

3. On a $\deg(\varphi(X^k)) \leqslant k + d - n \leqslant k$.

Et donc il existe des réels $a_{0,k}, a_{1,k}, \ldots, a_{k,k}$ tels que $\varphi(X^k) = \sum\limits_{i=0}^{k} a_{i,k} X^i$.

Autrement dit, la matrice de φ dans la base canonique est

$$\mathrm{Mat}_{\mathcal{B}_{can}}(\varphi) = \begin{array}{c} \begin{matrix} \varphi(1) & \varphi(X) & \varphi(X^2) & \dots & \varphi(X^n) \end{matrix} \\ \begin{pmatrix} a_{0,0} & a_{0,1} & a_{0,2} & \dots & a_{0,n} \\ 0 & a_{1,1} & a_{1,2} & \dots & a_{1,n} \\ 0 & 0 & a_{2,2} & \dots & a_{2,n} \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & 0 & \dots & 0 & a_{n,n} \end{pmatrix} \end{array} \begin{matrix} 1 \\ X \\ X^2 \\ \vdots \\ X^n \end{matrix} \in \mathcal{M}_{n+1}(\mathbf{R}).$$

Ainsi, cette matrice est triangulaire supérieure.

Remarque

Le fait que cette matrice soit triangulaire supérieure est équivalent à dire que pour tout $k \in [\![0, n]\!]$, $\mathbf{R}_k[X]$ est stable par φ.

Dans le cas où $d < n$, on peut affirmer que pour tout $k \in [\![0, n]\!]$,

$$\deg \varphi(X^k) \leqslant k + d - n \leqslant k - 1.$$

Et donc en particulier, $a_{k,k} = 0$: la matrice possède une diagonale nulle.

Autrement dit

Pour tout $k \in [\![1, n]\!]$, l'image de $\mathbf{R}_k[X]$ par φ est inclus dans $\mathbf{R}_{k-1}[X]$.

4. Puisque la matrice de φ dans la base canonique est triangulaire, ses valeurs propres sont ses coefficients diagonaux.
En particulier, pour $d < n$, $\mathrm{Spec}(\varphi) = \{0\}$. Et donc φ ne peut être diagonalisable que si elle est nulle, c'est-à-dire si et seulement si $Q = 0$.

Rappel

Une matrice diagonalisable qui possède une unique valeur propre λ est nécessairement égale à λI_n.

Si $d = n$, notons $Q = \sum_{i=0}^{n} b_i X^i$, de sorte que $b_n \neq 0$. Ainsi,

$$\varphi(X^k) = \left(\sum_{i=0}^{n} b_i X^{i+k}\right)^{(n)} = \sum_{i=n-k}^{n} b_i \frac{(i+k)!}{(i+k-n)!} X^{i+k-n}.$$

Détails

Il est facile de prouver par récurrence sur $d \leqslant k$

$$\left(X^k\right)^{(d)} = k(k-1)\cdots(k-d+1)X^{k-d} = \frac{k!}{(k-d)!}X^{k-d}.$$

et que pour $d > k$,

$$(X^k)^{(d)} = 0.$$

En particulier, le coefficient de degré k de $\varphi(X^k)$ est $b_n(n+k)!$.
Ainsi, les coefficients diagonaux de $\mathrm{Mat}_{\mathcal{B}_{can}}(\varphi)$ sont les $b_n(n+k)!$, qui sont deux à deux distincts car $b_n \neq 0$.
Et donc φ possède $n+1 = \dim \mathbf{R}_n[X]$ valeurs propres distinctes, et donc est diagonalisable.

En résumé, φ est diagonalisable si et seulement si $Q = 0$ ou si $\deg Q = n$.

□

Exercice 1.32 (ESCP 2012) [ESCP12.2.17] — Moyen

Étude d'une équation matricielle.

Dans tout l'exercice, (A, B) désigne un couple de matrices de $\mathcal{M}_n(\mathbf{R})$ tel que :

$$AB - BA = A \text{ avec } A \text{ non nulle} \quad (\star)$$

1. Montrer que $\mathrm{tr}(A) = 0$ et que A n'est pas inversible.
2. Montrer que : $\forall k \in \mathbf{N}$, $A^k B - BA^k = kA^k$.
3. En considérant les valeurs propres de l'endomorphisme φ de $\mathcal{M}_n(\mathbf{R})$ défini par :

$$\varphi : M \mapsto MB - BM$$

montrer qu'il existe un entier $p \in [\![2, n]\!]$ tel que $A^p = 0$.

4. Pour $n = 2$, déterminer les couples (A, B) solutions du problème $(\star)$.

Solution de l'exercice 1.32

1. Les propriétés élémentaires de la trace indiquent que

$$\mathrm{tr}(A) = \mathrm{tr}(AB - BA) = \mathrm{tr}(AB) - \mathrm{tr}(BA) = \mathrm{tr}(AB) - \mathrm{tr}(AB) = 0.$$

D'autre part, si A était inversible, alors en multipliant à gauche par A^{-1}, il viendrait $B - A^{-1}BA = I_n$.
Puisque B et $A^{-1}BA$ sont semblables, elles ont donc même trace, de sorte que $\text{tr}(I_n) = \text{tr}(B) - \text{tr}(A^{-1}BA) = 0$.
Or $\text{tr}(I_n) = n \neq 0$, d'où une contradiction : A n'est donc pas inversible.

2. Montons le résultat par récurrence sur k.
Pour $k = 0$, on a évidemment $A^0B - BA^0 = B - B = 0 = 0A^0$.
Supposons que $A^kB - BA^k = kA^k$.
Alors, en multipliant à droite par A, il vient

$$A^kBA - BA^{k+1} = kA^{k+1}.$$

Mais d'après $(\star)$, on a $BA = AB - A$, de sorte que

$$A^kBA = A^kAB - A^kA = A^{k+1}B - A^{k+1}.$$

Et donc

$$A^{k+1}B - A^{k+1} - BA^{k+1} = kA^{k+1} \Leftrightarrow A^{k+1}B - BA^{k+1} = (k+1)A^{k+1}.$$

Par le principe de récurrence, pour tout $k \in \mathbf{N}$, $A^kB - BA^k = kA^k$.

3. La question précédente montre que si $A^k \neq 0$, alors A^k est vecteur propre de φ, associé à la valeur propre k.
Donc si pour tout $k \in \mathbf{N}$, $A^k \neq 0$, alors φ possède une infinité de valeurs propres, ce qui n'est pas possible puisqu'il doit posséder un nombre fini de valeurs propres.
Il existe donc un entier k tel que $A^k = 0$.
Il reste donc à prouver que le plus petit entier k tel que $A^k = 0$ vérifie $k \leqslant n$.
Mais c'est un résultat classique : une matrice de $\mathcal{M}_n(\mathbf{R})$ nilpotente possède forcément un indice de nilpotence inférieur ou égal à n.
Reprouvons ce résultat : puisque $A^{k-1} \neq 0$, il existe $X \in \mathcal{M}_{n,1}(\mathbf{R})$ tel que $A^{k-1}X \neq 0$.
Alors la famille $(X, AX, A^2X, \dots, A^{k-1}X)$ est une famille libre de $\mathcal{M}_{n,1}(\mathbf{R})$.
En effet, soient $\lambda_0, \dots, \lambda_{k-1}$ des réels tels que $\sum_{i=0}^{k-1} \lambda_i A^iX = 0$.
En multipliant à gauche par A^{k-1}, il vient $\lambda_0 A^{k-1}X + \sum_{i=1}^{k-1} \lambda_i \underbrace{A^{k-1+i}X}_{=0} = 0$.
Puisque $A^{k-1}X \neq 0$ c'est donc que $\lambda_0 = 0$.
Il reste donc $\lambda_1 AX + \lambda_2 A^2X + \cdots + \lambda_{k-1}A^{k-1}X = 0$.
En multipliant cette fois par A^{k-2}, on obtient $\lambda_1 A^{k-1}X = 0$, et donc $\lambda_1 = 0$.
De proche en proche, on prouve ainsi que $\lambda_0 = \lambda_1 = \cdots = \lambda_{k-1} = 0$.
Ainsi, $(X, AX, \dots, A^{k-1}X)$ est une famille libre de $\mathcal{M}_{n,1}(\mathbf{R})$, de cardinal k.
Et donc $k \leqslant \dim \mathcal{M}_{n,1}(\mathbf{R}) = n$.
Nous avons donc bien prouvé qu'il existe $k \in [\![2, n]\!]$ tel que $A^k = 0$.

Mieux
Le nombre de valeurs propres de φ est au plus égal à $\dim \mathcal{M}_n(\mathbf{R}) = n^2$.

Remarque
$k \geqslant 2$ puisque $A \neq 0$ par hypothèse.

4. D'après la question précédente, si (A, B) est solution, alors $A^2 = 0$.
Soit alors, comme précédemment, un vecteur $X \in \mathcal{M}_{2,1}(\mathbf{R})$ tel que $AX \neq 0$. Nous venons de prouver que (X, AX) est libre, et donc est une base de $\mathcal{M}_{2,1}(\mathbf{R})$.
La matrice de $f : U \mapsto AU$ dans cette base est alors $\begin{array}{c} \\ X \\ AX \end{array} \begin{array}{c} \begin{array}{cc} AX & A^2X \end{array} \\ \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \end{array} = T$.

Puisque d'autre part, la matrice de f dans la base canonique de $\mathcal{M}_{2,1}(\mathbf{R})$ est A, A est donc semblable à T : il existe une matrice P inversible telle que $A = P^{-1}TP$.
Posons alors $C = PBP^{-1}$, de sorte que $B = P^{-1}CP$.
L'équation $(\star)$ s'écrit encore

$$P^{-1}TCP - P^{-1}CTP = P^{-1}TP$$

ce qui, après multiplication à gauche par P et à droite par P^{-1} devient $TC - CT = T$.
Soit donc $C = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$. On a

$$TC - CT = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} a & b \\ c & d \end{pmatrix} - \begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} c & d-a \\ 0 & -c \end{pmatrix}.$$

On a alors $TC - CT = T$ si et seulement si $c = 0$ et $d = a + 1$.
Et donc C est de la forme $\begin{pmatrix} a & b \\ 0 & a+1 \end{pmatrix}$.

Inversement, s'il existe P inversible telle que $A = P^{-1}TP$ et deux a et b tels que $B = P^{-1}\begin{pmatrix} a & b \\ 0 & a+1 \end{pmatrix}P$, alors on a

$$AB - BA = P^{-1}\left(T\begin{pmatrix} a & b \\ 0 & a+1 \end{pmatrix} - \begin{pmatrix} a & b \\ 0 & a+1 \end{pmatrix}T\right)P = P^{-1}TP = A.$$

Ainsi, l'ensemble des solutions de $(\star)$ est

$$\left\{\left(P^{-1}TP, P^{-1}\begin{pmatrix} a & b \\ 0 & a+1 \end{pmatrix}P\right),\ (a,b) \in \mathbf{R}^2,\ P \in GL_2(\mathbf{R})\right\}.$$

□

Exercice 1.33 (HEC 2018) [HEC18.246] — Moyen

Dimension de l'ensemble des endomorphismes de $\mathcal{M}_n(\mathbf{R})$ commutant avec la transposition

Dans tout l'exercice, n désigne un nombre supérieur ou égal à 2.
On note $\mathcal{M}_n(\mathbf{R})$ l'ensemble des matrices carrées d'ordre n à coefficients réels, et T l'endomorphisme de $\mathcal{M}_n(\mathbf{R})$ qui associe à toute matrice M de $\mathcal{M}_n(\mathbf{R})$ sa transposée ${}^tM$.
Pour tout espace vectoriel E, on note $\mathcal{L}(E)$ l'ensemble des endomorphismes de E.
Le but de l'exercice est l'étude de l'ensemble

$$\mathcal{C} = \{f \in \mathcal{L}(\mathcal{M}_n(\mathbf{R})) : \forall M \in \mathcal{M}_n(\mathbf{R}),\ f(T(M)) = T(f(M))\}.$$

1. (a) **Question de cours** : donner la dimension de l'espace vectoriel $\mathcal{L}(E)$ lorsque E est un espace de dimension finie. En déduire la dimension de $\mathcal{L}(\mathcal{M}_n(\mathbf{R}))$.
   (b) Pourquoi la trace $\mathrm{tr}(M)$ de la matrice M de T dans une base de $\mathcal{M}_n(\mathbf{R})$ ne dépend pas de la base choisie ? Que vaut cette trace ?
2. Trouver une application linéaire dont $\mathcal{C}$ est le noyau et en déduire que $\mathcal{C}$ est un espace vectoriel.
3. On note $\mathcal{S}_n(\mathbf{R}) = \{S \in \mathcal{M}_n(\mathbf{R}) : S = {}^tS\}$ et $\mathcal{A}_n(\mathbf{R}) = \{A \in \mathcal{M}_n(\mathbf{R}) : A = -{}^tA\}$.
   (a) Établir que : $\mathcal{M}_n(\mathbf{R}) = \mathcal{S}_n(\mathbf{R}) \oplus \mathcal{A}_n(\mathbf{R})$.
   (b) En déduire qu'un endomorphisme f de $\mathcal{M}_n(\mathbf{R})$ appartient à $\mathcal{C}$ si et seulement si les deux sous-espaces $\mathcal{S}_n(\mathbf{R})$ et $\mathcal{A}_n(\mathbf{R})$ sont stables par f.
4. (a) Déterminer les dimensions respectives de $\mathcal{S}_n(\mathbf{R})$ et de $\mathcal{A}_n(\mathbf{R})$.
   (b) En déduire l'égalité : $\dim \mathcal{C} = \dfrac{n^2(n^2+1)}{2}$.
5. (a) Démontrer que l'endomorphisme T est diagonalisable
   (b) Justifier qu'il existe des éléments f de $\mathcal{C}$ qui ne sont pas diagonalisables, et en donner un exemple.

Solution de l'exercice 1.33

1.a. Nous savons que $\dim \mathcal{L}(E) = \dim E \times \dim E$.
En particulier, puisque $\dim \mathcal{M}_n(\mathbf{R}) = n^2$, il vient $\dim \mathcal{L}(\mathcal{M}_n(\mathbf{R})) = n^2 \times n^2 = n^4$.

1.b. Si $\mathcal{B}$ et $\mathcal{B}'$ sont deux bases de $\mathcal{M}_n(\mathbf{R})$, alors les matrices $\mathrm{Mat}_{\mathcal{B}}(T)$ et $\mathrm{Mat}_{\mathcal{B}'}(T)$ sont semblables[1] et par conséquent ont la même trace.
Donc il nous suffit de déterminer la trace de la matrice de T dans la base de notre choix.
Et puisque nous n'en connaissons pas 50, prenons la plus classique d'entre elles : la base canonique $\mathcal{B}_{can}$ et notons $M = \mathrm{Mat}_{\mathcal{B}_{can}}(T)$.
Rappelons que $\mathcal{B}_{can}$ est la famille des $(E_{i,j})_{\substack{1\leqslant i\leqslant n \\ 1\leqslant j\leqslant n}}$, où $E_{i,j}$ est la matrice dont tous les coefficients sont nuls, à l'exception du coefficient (i,j) qui vaut 1.
Pour déterminer la trace de la matrice de T dans cette base, il nous faut en déterminer les coefficients diagonaux.

[1] Puisqu'elles représentent toutes les deux l'endomorphisme T.

C'est-à-dire la composante de $T(E_{i,j})$ dans la base $\mathscr{B}_{can}$. Mais $T(E_{i,j}) = E_{j,i}$. Et donc si $i = j$, alors le coefficient diagonal de la colonne de M correspondant à $E_{i,j}$ vaut 1.
Et si $i \neq j$, alors ce coefficient diagonal est nul.
Ainsi, M possède n coefficients diagonaux non nuls, dans les colonnes correspondant à $E_{1,1}, E_{2,2}, \dots, E_{n,n}$, et ces n coefficients valent 1.
Et donc $\text{tr}(M) = n$.

Si A est la matrice d'un endomorphisme f dans une base $(e_1, \dots, e_n)$, alors le coefficient diagonal $a_{i,i}$ est la composante de $f(e_i)$ suivant e_i :

$$\begin{array}{c} \begin{array}{ccc} f(e_1) & \dots & f(e_n) \end{array} \\ \begin{pmatrix} a_{1,1} & \dots & \\ \vdots & \ddots & \\ & & a_{n,n} \end{pmatrix} \end{array} \begin{array}{c} e_1 \\ \vdots \\ e_n \end{array}$$

2. Notons qu'on a $\mathscr{C} = \{f \in \mathscr{L}(\mathscr{M}_n(\mathbf{R})) \ : T \circ f = f \circ T\} = \{f \in \mathscr{L}(\mathscr{M}_n(\mathbf{R})) : T \circ f - f \circ T = 0\}$.
Or, il est facile de prouver que $f \mapsto T \circ f - f \circ T$ est un endomorphisme de $\mathscr{L}(\mathscr{M}_n(\mathbf{R}))$, dont le noyau est donc $\mathscr{C}$, d'après l'égalité qui précède.
Et comme tout noyau, $\mathscr{C}$ est donc un sous-espace vectoriel de $\mathscr{L}(\mathscr{M}_n(\mathbf{R}))$.

3.a. C'est un grand classique. Puisque nous ne connaissons pas les dimensions de $\mathscr{S}_n(\mathbf{R})$ et de $\mathscr{A}_n(\mathbf{R})$, on ne peut utiliser tout critère faisant intervenir la dimension. Et donc il ne nous reste à peu près que la définition de supplémentaires : il va falloir prouver que toute matrice de $\mathscr{M}_n(\mathbf{R})$ s'écrit de manière unique comme somme d'un élément de $\mathscr{S}_n(\mathbf{R})$ et d'un élément de $\mathscr{A}_n(\mathbf{R})$.

Soit donc $M \in \mathscr{M}_n(\mathbf{R})$.
Procédons par analyse synthèse, et supposons qu'il existe $S \in \mathscr{S}_n(\mathbf{R})$ et $A \in \mathscr{A}_n(\mathbf{R})$ telles que $M = A + S$.
Alors ${}^tM = {}^tA + {}^tS = -A + S$.
Par conséquent, $M + {}^tM = 2S \Leftrightarrow S = \dfrac{M + {}^tM}{2}$.
Et de même, $M - {}^tM = 2A \Leftrightarrow A = \dfrac{M - {}^tM}{2}$.
Ainsi, nous venons de prouver que si une décomposition de M comme somme d'un élément de $\mathscr{A}_n(\mathbf{R})$ et d'un élément de $\mathscr{S}_n(\mathbf{R})$ existe, alors elle est unique, et donnée par les formules ci-dessus.

Il nous reste donc à faire la synthèse[2], c'est-à-dire à vérifier que $S = \dfrac{M + {}^tM}{2}$ est bien dans $\mathscr{S}_n(\mathbf{R})$, que $A = \dfrac{M - {}^tM}{2}$ est bien dans $\mathscr{A}_n(\mathbf{R})$, et que leur somme vaut M, ce qui est une formalité.
Et donc toute matrice de $\mathscr{M}_n(\mathbf{R})$ s'écrit de manière unique comme un élément de $\mathscr{S}_n(\mathbf{R})$ plus un élément de $\mathscr{A}_n(\mathbf{R})$.
Et donc $\mathscr{M}_n(\mathbf{R}) = \mathscr{S}_n(\mathbf{R}) \oplus \mathscr{A}_n(\mathbf{R})$.

[2] C'est-à-dire à prouver qu'une telle écriture existe.

3.b. Si $f \in \mathscr{L}(\mathscr{M}_n(\mathbf{R}))$ est tel que $\mathscr{S}_n(\mathbf{R})$ et $\mathscr{A}_n(\mathbf{R})$ soient stables par f.
Alors pour $M \in \mathscr{M}_n(\mathbf{R})$, il existe $S \in \mathscr{S}_n(\mathbf{R})$ et $A \in \mathscr{A}_n(\mathbf{R})$ telles que $M = S + A$.
Et alors $f(M) = f(S) + f(A)$, de sorte que

$$T(f(M)) = T(\underbrace{f(S)}_{\in \mathscr{S}_n(\mathbf{R})}) + T(\underbrace{f(A)}_{\in \mathscr{A}_n(\mathbf{R})}) = f(S) - f(A) = f({}^tS) + f({}^tA) = f(T(S + A)) = f(T(M)).$$

Et donc f et T commutent, de sorte que $f \in \mathscr{C}$.

Inversement, supposons que $f \in \mathscr{C}$. Alors pour $S \in \mathscr{S}_n(\mathbf{R})$, on a

$${}^tf(S) = T(f(S)) = f(T(S)) = f({}^tS) = f(S).$$

Ainsi, $f(S)$ est symétrique, et donc $\mathscr{S}_n(\mathbf{R})$ est stable par f.

Et de même, si $A \in \mathscr{A}_n(\mathbf{R})$, alors

$${}^tf(A) = T(f(A)) = f(T(A)) = f(-A) = -f(A)$$

de sorte que $f(A) \in \mathscr{A}_n(\mathbf{R})$ et donc $\mathscr{A}_n(\mathbf{R})$ est stable par f.

Méthode

La notation T est ici parfois un peu trompeuse, mais ne perdons pas de vue que nous cherchons à prouver que $f(S) \in \mathscr{S}_n(\mathbf{R})$. Autrement dit que $f(S)$ est égale à sa propre tansposée.

4.a. Le plus classique pour obtenir ces dimensions est d'exhiber une base de $\mathscr{S}_n(\mathbf{R})$.
Toutefois, notons qu'il y a bien plus simple avec le travail réalisé précédemment.
En effet, notons que $\mathscr{S}_n(\mathbf{R}) = \{M \in \mathscr{M}_n(\mathbf{R}) : M = {}^tM\} = \{M \in \mathscr{M}_n(\mathbf{R}) : T(M) = M\} = E_1(T)$ est un sous-espace propre de T.
Et de même, $\mathscr{A}_n(\mathbf{R}) = E_{-1}(T)$.

Comme nous avons prouvé à la question 3.a que $\mathcal{M}_n(\mathbf{R}) = E_1(T) \oplus E_{-1}(T)$, cela prouve que T est diagonalisable, et que $\mathrm{Spec}(T) = \{-1, 1\}$.
Et donc la matrice de T dans une base de $\mathcal{M}_n(\mathbf{R})$ formée de vecteurs propres de T est diagonale, avec $\dim \mathcal{S}_n(\mathbf{R})$ coefficients diagonaux égaux à 1 et $\dim \mathcal{A}_n(\mathbf{R})$ coefficients diagonaux égaux à -1.
Sa trace est donc $\dim \mathcal{S}_n(\mathbf{R}) - \dim \mathcal{A}_n(\mathbf{R})$.
Mais comme mentionné à la question 1.b, cette trace vaut n.
Nous avons donc le système suivant : $\begin{cases} \dim \mathcal{S}_n(\mathbf{R}) + \dim \mathcal{A}_n(\mathbf{R}) = \dim \mathcal{M}_n(\mathbf{R}) = n^2 \\ \dim \mathcal{S}_n(\mathbf{R}) - \dim \mathcal{A}_n(\mathbf{R}) = n \end{cases}$ qui conduit facilement à

$$\dim \mathcal{S}_n(\mathbf{R}) = \frac{n(n+1)}{2} \text{ et } \dim \mathcal{A}_n(\mathbf{R}) = \frac{n(n-1)}{2}.$$

Si vous préférez donner une base de $\mathcal{S}_n(\mathbf{R})$, il est bon de se souvenir qu'une telle base est donnée par la famille formée à la fois par les $E_{i,i}$, $1 \leqslant i \leqslant n$ et les $E_{i,j} - E_{j,i}$, $1 \leqslant i < j \leqslant n$, les détails sont laissés à titre d'exercice.

Base

Notons que cette base ne sort pas de nulle part : pour choisir une matrice de $\mathcal{S}_n(\mathbf{R})$, il suffit de choisir les n coefficients diagonaux (d'où les $E_{i,j}$) et les coefficients au dessus de la diagonale (d'où le $1 \leqslant i < j \leqslant n$). Ceci impose alors automatiquement la valeur des coefficients sous la diagonale.

4.b. D'après la question 3.b, choisir un endomorphisme de $\mathcal{C}$ revient[3] à choisir un endomorphisme de $\mathcal{A}_n(\mathbf{R})$ et un endomorphisme de $\mathcal{S}_n(\mathbf{R})$.
Or, $\dim \mathcal{L}(\mathcal{S}_n(\mathbf{R})) = \left(\dim \mathcal{S}_n(\mathbf{R})\right)^2 = \dfrac{n^2(n+1)^2}{2}$ et de même, $\dim \mathcal{L}(\mathcal{A}_n(\mathbf{R})) = \dfrac{n^2(n-1)^2}{4}$.
Et par conséquent, $\dim \mathcal{C} = \dim \mathcal{L}(\mathcal{S}_n(\mathbf{R})) + \dim \mathcal{L}(\mathcal{A}_n(\mathbf{R})) = \dfrac{n^2(n^2+1)}{2}$.

[3] Ce n'est pas vraiment dit ainsi dans 3.b, mais c'est ainsi qu'il faudrait le comprendre. Les détails sont ci-dessous si besoin.

Quelques détails : les détails qui suivent sont très probablement inutiles à l'oral, notamment car il serait trop long de tout écrire en détail, mais il faudrait au moins comprendre les grandes idées de ce qui suit.
Nous avons prouvé en 3.a que $f \in \mathcal{C}$ si et seulement si $\mathcal{S}_n(\mathbf{R})$ et $\mathcal{A}_n(\mathbf{R})$ sont stables par f.
Considérons alors l'application

$$\Phi : \left| \begin{array}{ccc} \mathcal{C} & \longrightarrow & \dim \mathcal{L}(\mathcal{S}_n(\mathbf{R})) \times \dim \mathcal{L}(\mathcal{A}_n(\mathbf{R})) \\ f & \longmapsto & (f_{|\mathcal{S}_n(\mathbf{R})}, f_{|\mathcal{A}_n(\mathbf{R})}) \end{array} \right. ,$$

où $f_{|\mathcal{S}_n(\mathbf{R})}$ désigne la restriction[4] de f à $\mathcal{S}_n(\mathbf{R})$.
Alors il est facile de voir que Φ est injective. En effet, puisque $\mathcal{S}_n(\mathbf{R})$ et $\mathcal{A}_n(\mathbf{R})$ sont supplémentaires, connaître f sur chacun de ces deux sous-espaces permet de connaître f sur $\mathcal{M}_n(\mathbf{R})$ tout entier.
D'autre part, Φ est bijective. En effet, soit $(u, v) \in \mathcal{L}(\mathcal{S}_n(\mathbf{R})) \times \mathcal{L}(\mathcal{A}_n(\mathbf{R}))$, et soit f l'application définie sur $\mathcal{M}_n(\mathbf{R})$, qui à toute matrice $M = \underbrace{S}_{\in \mathcal{S}_n(\mathbf{R})} + \underbrace{A}_{\in \mathcal{A}_n(\mathbf{R})} \in \mathcal{M}_n(\mathbf{R})$ associe $u(S) + v(A)$.
Alors f est linéaire, et laisse stable $\mathcal{S}_n(\mathbf{R})$ et $\mathcal{A}_n(\mathbf{R})$, donc est dans $\mathcal{C}$. Et alors $\Phi(f) = (u, v)$.
Autrement dit, Φ est un isomorphisme.
Et donc $\dim \mathcal{C} = \dim(\mathcal{L}(\mathcal{S}_n(\mathbf{R})) \times \mathcal{L}(\mathcal{A}_n(\mathbf{R}))) = \dim \mathcal{L}(\mathcal{S}_n(\mathbf{R})) + \dim \mathcal{L}(\mathcal{A}_n(\mathbf{R}))$.

[4] C'est toujours f, sauf qu'on considère cette fois que son espace de départ est seulement $\mathcal{S}_n(\mathbf{R})$ et non $\mathcal{M}_n(\mathbf{R})$ tout entier.

Enfin, mentionnons l'interprétation matricielle de tout cela : considérons une base $\mathcal{B} = (E_1, \dots, E_r, F_1, \dots, F_p)$ de $\mathcal{M}_n(\mathbf{R})$ obtenue par concaténation d'une base $(E_1, \dots, E_r)$ de $\mathcal{S}_n(\mathbf{R})$ et d'une base $(F_1, \dots, F_p)$ de $\mathcal{A}_n(\mathbf{R})$.
Si $\mathcal{S}_n(\mathbf{R})$ est stable par f, alors pour tout $i \in [\![1, r]\!]$, $f(E_i) \in \mathcal{S}_n(\mathbf{R}) = \mathrm{Vect}(E_1, \dots, E_r)$. Et donc la $i^{\text{ème}}$ colonne de $\mathrm{Mat}_{\mathcal{B}}(f)$ se termine par p zéros.
Et de même, si $\mathcal{A}_n(\mathbf{R})$ est stable, alors les p dernière colonnes de $\mathrm{Mat}_{\mathcal{B}}(f)$ commencent par r zéros.
Autrement dit, si $f \in \mathcal{C}$, alors la matrice de f dans la base $\mathcal{B}$ est de la forme

$$\mathrm{Mat}_{\mathcal{B}}(f) = \begin{array}{c} \begin{array}{cccccc} f(E_1) & \dots & f(E_r) & F_1 & \dots & F_p \end{array} \\ \begin{pmatrix} \star & \dots & \star & 0 & \dots & 0 \\ \vdots & & \vdots & \vdots & & \vdots \\ \star & \dots & \star & 0 & \dots & 0 \\ 0 & \dots & 0 & \star & \dots & \star \\ \vdots & & \vdots & \vdots & & \vdots \\ 0 & \dots & 0 & \star & \dots & \star \end{pmatrix} \end{array} \begin{array}{c} E_1 \\ \vdots \\ E_r \\ F_1 \\ \vdots \\ F_p \end{array}$$

où le bloc carré en haut à gauche est de taille $r^2 = \dim \mathcal{S}_n(\mathbf{R}) \times \dim \mathcal{S}_n(\mathbf{R})$ et le bloc carré en bas à droite est de dimension $p^2 = \dim \mathcal{A}_n(\mathbf{R}) \times \dim \mathcal{A}_n(\mathbf{R})$.

Inversement, si $f \in \mathcal{L}(\mathcal{M}_n(\mathbf{R}))$ possède dans la base $\mathcal{B}$ une matrice de cette forme, alors $\mathcal{S}_n(\mathbf{R})$ et $\mathcal{A}_n(\mathbf{R})$ sont stables par f et donc $f \in \mathcal{C}$.
Il s'agit donc de déterminer la dimension du sous-espace vectoriel de $\mathcal{M}_{n^2}(\mathbf{R})$ des matrices qui sont de cette forme, et il est facile de se convaincre que cette dimension vaut $(\dim \mathcal{S}_n(\mathbf{R}))^2 + (\dim \mathcal{A}_n(\mathbf{R}))^2 = \dfrac{n^2(n^2+1)}{4}$.

5.a. Cela a déjà été dit précédemment !

5.b. La dimension de $\mathcal{S}_n(\mathbf{R})$ est strictement supérieure à 2, et donc il existe des endomorphismes de $\mathcal{S}_n(\mathbf{R})$ qui ne sont pas diagonalisables. Par exemple, tous ceux dont la matrice dans une base est triangulaire supérieure, non nulle, et avec uniquement des 0 sur la diagonale.
De tels endomorphismes ne peuvent être diagonalisables puisqu'ils n'ont que 0 comme valeur propre, sans être nuls pour autant.
À partir d'un tel endomorphisme $u \in \mathcal{L}(\mathcal{S}_n(\mathbf{R}))$, on peut fabriquer l'endomorphisme $\Phi^{-1}(u, 0)$, qui ne sera pas diagonalisable pour des raisons évidentes[5]
Pour donner un exemple de tel endomorphisme, considérons celui à qui $E_{2,2}$ associe $E_{1,1}$, et qui à toute autre matrice de la base de $\mathcal{S}_n(\mathbf{R})$ évoquée ci-dessus associe 0.
Alors sa matrice dans la base $(E_{1,1}, E_{2,2}, \dots)$ de $\mathcal{S}_n(\mathbf{R})$ est nulle, à l'exception du coefficient situé sur la première ligne et seconde colonne, qui vaut 1, donc n'est pas diagonalisable.
Et alors, l'endomorphisme $f \in \mathcal{C}$ correspondant est

$$M = (m_{i,j})_{1 \leqslant i,j \leqslant n} \mapsto m_{2,2} E_{1,1}.$$

□

[5] Just kidding ! Ce n'est pas totalement évident, et il faudrait sûrement en écrire un peu plus pour se convaincre qu'il n'est pas diagonalisable.

EXERCICE 1.34 (QSP HEC 2018) [HEC18.QSP247] — Moyen

Diagonalisation d'une matrice complexe

On considère la matrice de $\mathcal{M}_3(\mathbf{C})$: $M = \begin{pmatrix} e^{2i\pi/3} & e^{4i\pi/3} & 1 \\ e^{4i\pi/3} & 1 & e^{2i\pi/3} \\ 1 & e^{2i\pi/3} & e^{4i\pi/3} \end{pmatrix}$.

1. Calculer la trace et le rang de M.
2. La matrice M est-elle diagonalisable ?
3. Justifier que M est semblable à la matrice $M' = \begin{pmatrix} e^{4i\pi/3} & e^{2i\pi/3} & 1 \\ e^{2i\pi/3} & 1 & e^{4i\pi/3} \\ 1 & e^{4i\pi/3} & e^{2i\pi/3} \end{pmatrix}$

SOLUTION DE L'EXERCICE 1.34

1. La trace de M est $\mathrm{tr}(M) = 1 + e^{2i\pi/3} + e^{4i\pi/3}$.
Un moyen simple de conclure est de calculer ces deux dernières exponentielles :

$$e^{2i\pi/3} = \cos\frac{2\pi}{3} + i\sin\frac{2\pi}{3} = -\frac{1}{2} + i\frac{\sqrt{3}}{2} \text{ et } e^{4i\pi/3} = -\frac{1}{2} - i\frac{\sqrt{3}}{2}$$

de sorte que $\mathrm{tr}(M) = 0$.

Pour la culture, donnons tout de même une autre méthode : $1 + e^{2i\pi/3} + e^{4i\pi/3}$ est la somme de termes en progression géométrique, de raison $e^{2i\pi/3}$. Et donc

$$1 + e^{2i\pi/3} + e^{4i\pi/3} = \sum_{k=0}^{2} \left(e^{2i\pi/3}\right)^k = \frac{1 - \left(e^{2i\pi/3}\right)^3}{1 - e^{2i\pi/3}} = \frac{1-1}{1 - e^{2i\pi/3}} = 0.$$

Cette méthode a l'avantage de se généraliser sans efforts à toutes les sommes du type $\sum_{k=0}^{p-1} e^{2ik\pi/p}$.

Pour le rang, notons que $e^{2i\pi/3}\times e^{4i\pi/3}=e^{6i\pi/3}=1$, et donc que toutes les colonnes de M sont colinéaires[1].
Et donc $\text{rg}(M)=1$.

[1] On a les relations $L_2=e^{2i\pi/3}L_1,\ L_3=e^{4i\pi/3}L_1$.

2. Puisque $\text{rg}(M)=1$, 0 est valeur propre de M avec un sous-espace propre de dimension 2.
Si M était diagonalisable, alors elle serait semblable à une matrice diagonale $D=\text{Diag}(0,0,\lambda)$ avec $\lambda\neq 0$.
Mais alors M et D ont même trace, de sorte que $\lambda=0$, contredisant ce qui vient d'être dit.
Donc M n'est pas diagonalisable.

3. Soit $\mathscr{B}=(e_1,e_2,e_3)$ une base de $\mathbf{C}^3$ et soit f l'unique endomorphisme de $\mathbf{C}^3$ dont la matrice dans la base $\mathscr{B}$ est M.
Alors $\mathscr{B}'=(e_3,e_2,e_1)$ est encore une base de $\mathbf{C}^3$ et on a

$$f(e_3)=e^{4i\pi/3}e_3+e^{2i\pi/3}e_2+e_1,\ f(e_2)=e^{2i\pi/3}e_3+e_2+e^{4i\pi/3}e_1,\ f(e_1)=e_3+e^{4i\pi/3}e_2+e^{2i\pi/3}e_1$$

de sorte que la matrice de f dans la base (e_3,e_2,e_1) est $\begin{array}{c}\begin{matrix}f(e_3)&f(e_2)&f(e_1)\end{matrix}\\ \begin{pmatrix}e^{4i\pi/3}&e^{2i\pi/3}&1\\ e^{2i\pi/3}&1&e^{4i\pi/3}\\ 1&e^{4i\pi/3}&e^{2i\pi/3}\end{pmatrix}\begin{matrix}e_3\\e_2\\e_1\end{matrix}\end{array}$.

Et donc nous avons deux matrices représentant le même endomorphisme de $\mathbf{C}^3$ dans deux bases différentes : ces matrices sont semblables.

□

1.6.1 Polynômes annulateurs

Bien que l'existence d'un polynôme annulateur non nul (pour un endomorphisme ou pour une matrice) soit directement du cours, il est fréquemment demandé de le redémontrer, et mieux vaut donc connaître la méthode, qui repose sur un argument de dimension.

Exercice 1.35 (QSP HEC 2012) [HEC12.QSP36] — Moyen

A^{-1} est un polynôme en A

Soit $n\in\mathbf{N}^*$ et $A\in\mathscr{M}_n(\mathbf{R})$.

1. Établir l'existence d'un polynôme non nul $P\in\mathbf{R}[X]$ tel que $P(A)=0$.
2. On suppose que la matrice A est inversible. Montrer que A^{-1} s'écrit comme un polynôme en A.

Solution de l'exercice 1.35

1. La famille $\left(I_n,A,A^2,\dots,A^{n^2}\right)$ est une famille de n^2+1 éléments de $\mathscr{M}_n(\mathbf{R})$, qui est de dimension n^2.
Par conséquent, elle est nécessairement liée : il existe des réels $a_0,a_1,\dots,a_{n^2}$, non tous nuls, tels que

$$a_0I_n+a_1A+a_2A^2+\cdots+a_{n^2}A^{n^2}=0.$$

Ainsi, si on pose $P=\sum_{k=0}^{n^2}a_kX^k$, P est un polynôme non nul de $\mathbf{R}[X]$ tel que $P(A)=0$.

2. Soit donc $P=\sum_{k=0}^{n^2}a_kX^k$ un polynôme annulateur non nul de A.

Notons d le plus entier entier k tel que $a_k\neq 0$, de sorte que $P=X^d\sum_{k=d}^{n^2}a_kX^{k-d}$.

Et donc, si on pose $Q=\sum_{k=d}^{n^2}a_kX^{k-d}$, on a $P(X)=X^dQ(X)$ et donc $0=P(A)=A^dQ(A)$.
Puisque A est inversible, en multipliant à gauche par A^{-d}, il vient donc $Q(A)=0$.

Remarque
Nous avons ainsi prouvé qu'on peut prendre P de degré au plus n^2.
En réalité, on peut faire mieux et trouver un tel P de degré n, mais la preuve en est bien plus délicate.

Autrement dit
d est la multiplicité de 0 en tant que racine de P. Elle peut éventuellement être nulle si 0 n'est pas racine de P.

Mais Q possède $a_d \neq 0$ comme coefficient constant.

Donc $Q(A) = 0 \Leftrightarrow \sum_{k=d+1}^{n^2} a_k A^{k-d} = -a_d I_n$, et donc

$$A\left(-\frac{1}{a_d}\sum_{k=d+1}^{n^2} a_k A^{k-d-1}\right) = I_n$$

de sorte que $A^{-1} = -\frac{1}{a_d}\sum_{k=d+1}^{n^2} a_k A^{k-d-1}$ est bien un polynôme en A.

□

Il est bien connu que les valeurs propres d'un endomorphisme sont racines de tout polynôme annulateur. Dans le cas d'un polynôme annulateur de degré minimal, l'exercice qui suit établit la réciproque.

EXERCICE 1.36 (QSP HEC 2015) [HEC15.QSP139] — Facile

Toute racine d'un polynôme minimal est valeur propre

Soit E un espace vectoriel réel de dimension finie et f un endomorphisme de E.

1. Établir l'existence d'un polynôme P non nul tel que $P(f) = 0$.
2. Soit Q un polynôme tel que $Q(f) = 0$ et de degré minimal parmi les polynômes non nuls tels que $P(f) = 0$. Montrer que toute racine de Q est valeur propre de f.

SOLUTION DE L'EXERCICE 1.36

1. Notons n la dimension de E. Alors $\mathscr{L}(E)$ est un espace vectoriel de dimension n^2. La famille d'endomorphismes $\mathrm{id}_E, f, \ldots, f^{n^2}$ est de cardinal n^2+1, et donc elle est nécessairement liée : il existe des scalaires $a_0, \ldots, a_{n^2}$, **non tous nuls** tels que $a_0\mathrm{id}_E + a_1 f + \cdots + a_{n^2} f^{n^2} = 0$.
Si on note alors $P = a_0 + a_1 X + \cdots + a^{n^2} X^{n^2}$, alors P est non nul[1] et vérifie $P(f) = 0$.

[1] Car ses coefficients ne sont pas tous nuls.

2. Soit $\lambda \in \mathbf{R}$ une racine de Q. Alors il existe un polynôme $R \in \mathbf{R}[X]$, non nul, tel que $Q = (X-\lambda)R$.
En particulier, on a $0 = Q(f) = (f - \lambda\mathrm{id}_E) \circ R(f)$ $(\star)$.
Supposons que λ ne soit pas valeur propre de f.
Alors $f - \lambda\mathrm{id}_E$ est inversible, et donc, en composant la relation $(\star)$ à gauche par $(f-\lambda\mathrm{id}_E)^{-1}$, il vient $R(f) = 0$.
Le degré de $R(f)$ est strictement inférieur à celui de Q, contredisant la minimalité du degré de Q parmi les polynômes annulateurs non nuls.
Donc λ est une valeur propre de f.

□

EXERCICE 1.37 (QSP ESCP 2012) [ESCP12.QSP06] — Facile

Pour tout $A \in \mathscr{M}_n(\mathbf{R})$, on pose $v(A) = A + {}^tA$.

1. Montrer que v est un endomorphisme de $\mathscr{M}_n(\mathbf{R})$ et calculer v^2.
2. Déterminer les valeurs propres de v, et montrer que v est diagonalisable.
3. Déterminer $\mathrm{Im}(v)$ et $\mathrm{Ker}(v)$.

SOLUTION DE L'EXERCICE 1.37

1. Il est évident que v est un endomorphisme de $\mathscr{M}_n(\mathbf{R})$. De plus, pour $A \in \mathscr{M}_n(\mathbf{R})$, on a

$$v^2(A) = v(A) + {}^tv(A) = A + {}^tA + {}^t(A + {}^tA) = 2(A + {}^tA) = 2v(A).$$

2. Nous venons de prouver que $X^2 - 2X$ est un polynôme annulateur de v car $v^2 - 2v = 0$.
Donc les valeurs propres potentielles de v sont 0 et 2.
Or, on a $v(A) = 0 \Leftrightarrow A = -{}^tA \Leftrightarrow A$ est antisymétrique.
Et de même, $v(A) = 2A \Leftrightarrow A = {}^tA \Leftrightarrow A$ est symétrique.
Or, $\mathcal{A}_n(\mathbf{R})$ et $\mathcal{S}_n(\mathbf{R})$ sont supplémentaires[1] dans $\mathcal{M}_n(\mathbf{R})$, donc la somme des sous-espaces propres de b est égale à $\mathcal{M}_n(\mathbf{R})$ tout entier : v est diagonalisable.

[1] Résultat qui n'est pas au programme mais qu'il est bon de connaître pour l'oral. Vous en avez forcément déjà rencontré la preuve : toute matrice carrée s'écrit de manière unique comme somme d'une matrice symétrique et somme d'une matrice antisymétrique.

3. Nous avons déjà prouvé que $\operatorname{Ker}(v) = E_0(v) = \mathcal{A}_n(\mathbf{R})$.
De plus, il est clair que $\operatorname{Im}(v) \subset \mathcal{S}_n(\mathbf{R})$.
Or, par le théorème du rang, $\dim \operatorname{Im}(v) = \dim \mathcal{M}_n(\mathbf{R}) - \dim \operatorname{Ker}(v) = \dim \mathcal{S}_n(\mathbf{R})$, donc $\operatorname{Im}(v) = \mathcal{S}_n(\mathbf{R})$.

Ce dernier résultat est trivial si l'on sait que pour u diagonalisable, on a $\operatorname{Im}(u) = \bigoplus_{\substack{\lambda \in \operatorname{Spec}(u) \\ \lambda \neq 0}} E_\lambda(u)$.

□

Les exercices qui suivent sont très classiques : déterminer, à l'aide d'un polynôme annulateur et d'informations supplémentaires, si un endomorphisme est diagonalisable ou non.

Exercice 1.38 (QSP ESCP 2015) [ESCP15.QSP02] — Facile

Étude de diagonalisabilité à partir d'un polynôme annulateur.

Soit f un endomorphisme d'un espace vectoriel E de dimension finie $n \geqslant 1$, et a et b deux réels distincts. On suppose que

$$(f - a\mathrm{id}_E)^3 \circ (f - b\mathrm{id}_E) = 0 \text{ et } (f - a\mathrm{id}_E)^2 \circ (f - b\mathrm{id}_E) \neq 0.$$

Étudier la diagonalisabilité de f.

Solution de l'exercice 1.38 Les valeurs propres de f sont parmi les racines du polynôme annulateur $(X - a)^3(X - b)$, et donc $\operatorname{Spec}(f) \subset \{a, b\}$.
Si f était diagonalisable, alors tout polynôme P dont toutes les valeurs propres de f sont racines serait 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. □

Explication
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.

Exercice 1.39 (QSP ESCP 2017) [ESCP17.QSP06] — Facile

Soit E un espace vectoriel de dimension finie sur $\mathbf{R}$ et u un endomorphisme de E.
On suppose que $u^3 = u^2$, $u \neq \mathrm{id}$, $u^2 \neq 0$ et $u^2 \neq u$.

1. Déterminer les valeurs propres de u.
2. L'endomorphisme u est-il diagonalisable ?

Solution de l'exercice 1.39

1. Puisque $u^3 = u^2$, le polynôme $P(X) = X^3 - X^2 = X^2(X - 1)$ est annulateur de u.
Les valeurs propres de u sont donc parmi les racines de P, qui sont 0 et 1.

▶ Si 0 n'était pas valeur propre de u, alors u serait bijectif.
Et donc en composant[1] $u^3 = u^2$ par u^{-2}, on obtiendrait $u = \mathrm{id}$, ce qui n'est pas le cas par hypothèse. Donc 0 est valeur propre de u.

[1] À droite ou à gauche.

Rappel
λ est valeur propre si et seulement si $u - \lambda\mathrm{id}$ n'est pas bijectif.

▶ Si 1 n'était pas valeur propre de u, alors $u - \mathrm{id}$ serait bijectif.
Et donc en composant $u^2 \circ (u - \mathrm{id})$ par $(u - \mathrm{id})^{-1}$, il viendrait $u^2 = 0$, contredisant là aussi l'énoncé.

Donc 1 est valeur propre de u.

Et donc $\text{Spec}(u) = \{0, 1\}$.

2. Si u était diagonalisable, alors $\prod_{\lambda \in \text{Spec}(u)} (X - \lambda) = X(X-1)$ serait un polynôme annulateur de u, de sorte que $u^2 = u$.
Mais l'énoncé nous informe que $u^2 \neq u$, et donc u n'est pas diagonalisable. □

> **Remarque**
> En d'autres termes : un endomorphisme diagonalisable qui ne possède que 1 et −1 comme valeurs propres est un projecteur.

EXERCICE 1.40 (QSP HEC 2006) [HEC06.QSP02] — Facile

Soit $E = \mathbf{R}^3$, et soit $f \in \mathcal{L}(E)$ tel que $f^4 = f^2$ et dont -1 et 1 sont valeurs propres.
Montrer que f est diagonalisable.

SOLUTION DE L'EXERCICE 1.40 Distinguons deux cas : $\text{Ker} f = \{0\}$ et $\text{Ker} f \neq \{0\}$.
► **Si** $\text{Ker}(f) \neq \{0\}$. Alors 0 est une valeur propre de f, qui possède donc trois valeurs propres distinctes, alors que E est de dimension 3 : f est diagonalisable.

►**Si** $\text{Ker}(f) = \{0\}$. Alors f est inversible, et donc $f^2 = \text{id}_E$.
Alors tout vecteur $x \in E$ s'écrit $x = \frac{x+f(x)}{2} + \frac{x-f(x)}{2}$, avec $\frac{x+f(x)}{2} \in \text{Ker}(f - \text{id}_E) = E_1(f)$ et $\frac{x-f(x)}{2} \in \text{Ker}(f + \text{id}_E) = E_{-1}(f)$.
Donc $E = E_1(f) + E_{-1}(f)$, et les sous-espaces propres de f étant en somme directe, alors $E = E_1(f) \oplus E_{-1}(f)$, et donc f est diagonalisable. □

> **Remarque**
> Ce qui suit se transpose très bien dans un cas plus général (et c'est un grand classique de l'oral) : un endomorphisme possédant un polynôme annulateur de degré deux avec deux racines distinctes est diagonalisable. Voir l'exercice suivant à ce sujet.

EXERCICE 1.41 (QSP HEC 2015) [HEC15.QSP136] — Facile

Diagonalisabilité d'un endomorphisme possédant un polynôme annulateur de degré deux, scindé et à racines simples.

Soit E un espace vectoriel sur $\mathbf{K}$ ($\mathbf{R}$ ou $\mathbf{C}$) de dimension finie $n \geqslant 2$. On considère des endomorphismes f, p et q de E, ainsi que des scalaires distincts λ et μ tels que pour tout $k \in \{0, 1, 2\}$, on a $f^k = \lambda^k p + \mu^k q$.

1. Montrer que $(f - \lambda \text{id}_E) \circ (f - \mu \text{id}_E) = 0_{\mathcal{L}(E)}$.
2. En déduire que l'ensemble des valeurs propres de f est inclus dans $\{\lambda, \mu\}$ et que f est diagonalisable.

SOLUTION DE L'EXERCICE 1.41

1. On a $(f - \lambda \text{id}_E) \circ (f - \mu \text{id}_E) = f^2 - (\lambda + \mu) f + \lambda\mu \text{id}_E$. Et en remplaçant f^2, f et $\text{id}_E = f^0$ par les expressions données dans l'énoncé, il vient :

$$(f - \lambda \text{id}_E) \circ (f - \mu \text{id}_E) = (\lambda^2 p + \mu^2 q) - (\lambda + \mu)(\lambda p + \mu q) + \lambda\mu(p + q) = 0.$$

2. Nous venons de prouver que $(X - \lambda)(X - \mu)$ est annulateur de f. Et donc les valeurs propres de f sont parmi les racines de ce polynôme : $\text{Spec}(f) \subset \{\lambda, \mu\}$.
Supposons que λ ne soit pas valeur propre de f. Alors $f - \lambda \text{id}_E$ est inversible, et donc

$$f - \mu \text{id}_E = (f - \lambda \text{id}_E)^{-1} \circ \underbrace{(f - \lambda \text{id}_E) \circ (f - \mu \text{id}_E)}_{=0} = 0.$$

Et donc $f = \mu \text{id}_E$, qui est diagonalisable.
De même si μ n'est pas valeur propre, alors $f = \lambda \text{id}_E$ est diagonalisable.
Si λ et μ sont tous deux valeurs propres de f, on a $E_\lambda(f) \neq \{0\}$ et $E_\mu(f) \neq \{0\}$.
De plus, puisque $(f - \lambda\lambda \text{id}_E) \circ (f - \mu \text{id}_E) = 0$, alors $\text{Im}(f - \mu \text{id}_E) \subset \text{Ker}(f - \lambda \text{id}_E) = E_\lambda(f)$.

Et de même, $(f - \mu \text{id}_E) \circ (f - \lambda \text{id}_E) = 0$, donc $\text{Im}(f - \lambda \text{id}_E) \subset E_\mu(f)$.

> **Classique**
> Si $f \circ g = 0$, alors $\text{Im}(g) \subset \text{Ker}(f)$.

Enfin, si $x \in E$, alors on a

$$x = \frac{1}{\lambda-\mu}\Big(\underbrace{(f(x)-\lambda x)}_{\in E_\mu(f)} - \underbrace{(f(x)-\mu x)}_{\in E_\lambda(f)}\Big) \in E_\lambda(f) + E_\mu(f).$$

Donc $E = E_\lambda(f) + E_\mu(f)$. Puisque deux sous-espaces propres associés à des valeurs propres distinctes sont toujours en somme directe, on en déduit que $E = E_\lambda(f) \oplus E_\mu(f)$.
Et donc f est diagonalisable.

□

Le résultat que nous venons de prouver à la question 2 de l'exercice précédent reste valable pour une matrice (ou un endomorphisme) qui aurait un polynôme annulateur de la forme $(X-\alpha)(X-\beta)$ avec $\alpha \neq \beta$.
Nous prouvons dans l'exercice qui suit une généralisation beaucoup plus vaste de ce résultat : un endomorphisme (ou une matrice carrée) qui possède un polynôme annulateur scindé à racines simples (c'est-à-dire de la forme $\prod_{i=1}^{d}(X-\lambda_i)$ avec les λ_i deux à deux distincts) est diagonalisable.

Exercice 1.42 (ESCP 2014) [ESCP14.2.12] — Moyen

Une condition nécessaire et suffisante de diagonalisabilité

Soit $n \in \mathbf{N}^*$. On note I_n la matrice identité d'ordre n. Soit A une matrice carrée d'ordre n à coefficients complexes.

1. (a) Montrer que si A est diagonalisable, alors A^2 l'est également.
   (b) En s'intéressant à la matrice d'ordre $n \geqslant 2$ qui a un 1 dans le coin supérieur droit et des 0 partout ailleurs, montrer que la réciproque est fausse.
2. Montrer que si A est diagonalisable, alors il existe $k \in \mathbf{N}^*$ et $\alpha_1, \dots, \alpha_k \in \mathbf{C}$ **deux à deux distincts** tels que :
$$(A-\alpha_1 I_n)\cdots(A-\alpha_k I_n) = 0.$$
3. On veut montrer la réciproque, à savoir : «pour tout endomorphisme f d'un espace vectoriel complexe E de dimension finie, s'il existe $k \in \mathbf{N}^*$ et $\alpha_1, \dots, \alpha_k \in \mathbf{C}$ deux à deux distincts tels que :
$$(f-\alpha_1 \mathrm{id}_E)\circ\cdots\circ(f-\alpha_k \mathrm{id}_E) = 0,$$
alors f est diagonalisable.
On procède par récurrence sur $k \geqslant 1$.
   (a) Pour tout $\alpha_1, \dots, \alpha_k, \alpha_{k+1} \in \mathbf{C}$ deux à deux distincts, montrer qu'il existe un polynôme P et un nombre complexe β tels qu'on ait l'égalité suivante entre polynômes :
$$1 = (X-\alpha_{k+1})P + \beta(X-\alpha_1)\cdots(X-\alpha_k).$$
   (b) On suppose que le résultat voulu est vrai pour l'entier k et que l'on a un endomorphisme f (d'un espace E de dimension finie) et des nombres $\alpha_1, \dots, \alpha_k, \alpha_{k+1} \in \mathbf{C}$ deux à deux distincts tels que :
$$(f-\alpha_1 \mathrm{id}_E)\circ\cdots\circ(f-\alpha_k \mathrm{id}_E)\circ(f-\alpha_{k+1}\mathrm{id}_E) = 0.$$
      i. Montrer que le sous-espace vectoriel $F = \mathrm{Im}(f-\alpha_{k+1}\mathrm{id}_E)$ est stable par f.
      On note alors g l'endomorphisme de F défini par : $\forall v \in F,\ g(v) = f(v)$.
      ii. Calculer $(g-\alpha_1 \mathrm{id}_F)\circ\cdots\circ(g-\alpha_k \mathrm{id}_F)$.
      iii. Montrer que tout vecteur $v \in F$ s'écrit sous la forme : $v = v_1 + \cdots + v_k$ avec $v_j \in \mathrm{Ker}(f-\alpha_j \mathrm{id}_E)$ pour tout j de $[\![1,k]\!]$.
   (c) Conclure.
4. Montrer que si A^2 est diagonalisable et A est inversible, alors A est diagonalisable.

Solution de l'exercice 1.42

1.a. Si A est diagonalisable, alors il existe P inversible et D diagonale telles que $A = P^{-1}DP$.
Et donc $A^2 = \left(P^{-1}DP\right)^2 = P^{-1}D^2P$.

Mais D^2 est encore une matrice diagonale, et donc A^2 est semblable à une matrice diagonale, donc diagonalisable.

1.b. Si A est la matrice de l'énoncé, il est facile de constater que $A^2 = 0$, et donc est diagonalisable[1].
En revanche, A est triangulaire, et ne possède que 0 comme coefficient diagonal, de sorte que $\text{Spec}(A) = \{0\}$.
Si elle était diagonalisable, elle serait semblable à la matrice diagonale ne possédant que des 0 sur sa diagonale, c'est-à-dire à la matrice nulle. Et donc serait égale à la matrice nulle, ce qui n'est pas le cas.

[1] La matrice nulle est diagonale !

Plus généralement
Une matrice diagonalisable qui ne possède qu'une valeur propre λ est nécessairement égale à λI_n.

2. Soit A une matrice diagonalisable, et soient $\alpha_1, \dots, \alpha_k$ les différentes valeurs propres de A.
Notons également, pour tout $i \in [\![1, k]\!]$, $r_i = \dim E_{\alpha_i}(A)$.

Alors A est semblable à la matrice $D = \text{Diag}\left(\underbrace{\alpha_1, \dots, \alpha_1}_{r_1 \text{ fois}}, \dots, \underbrace{\alpha_k, \dots, \alpha_k}_{r_k \text{ fois}}\right)$.

Notons alors $P = \prod_{i=1}^{k}(X - \alpha_i)$, de sorte que

$$P(D) = \text{Diag}\left(\underbrace{P(\alpha_1), \dots, P(\alpha_1)}_{r_1 \text{ fois}}, \dots, \underbrace{P(\alpha_k), \dots, P(\alpha_k)}_{r_k \text{ fois}}\right) = 0.$$

Détails
Appliquer un polynôme à une matrice diagonale, c'est l'appliquer à chacun de ses coefficients diagonaux.

Et donc $P(A)$, qui est semblable à $P(D) = 0$ est nulle, de sorte que

$$(A - \alpha_1 I_n) \cdots (A - \alpha_k I_n) = 0.$$

3.a. Réalisons la division euclidienne de $(X - \alpha_1) \cdots (X - \alpha_k)$ par $X - \alpha_{k+1}$:

$$(X - \alpha_1) \cdots (X - \alpha_k) = (X - \alpha_{k+1})Q + R \text{ avec } \deg R < \deg(X - \alpha_{k+1}) = 1 \qquad (\star).$$

Alors R est une constante λ. De plus, en évaluant la relation $(\star)$ en α_{k+1}, il vient $\lambda = \prod_{i=1}^{k}(\alpha_{k+1} - \alpha_k) \neq 0$.
Et donc en divisant les deux membres de $(\star)$ par $-\lambda$, il vient

$$-\frac{1}{\lambda}(X - \alpha_1) \cdots (X - \alpha_k) = (X - \alpha_{k+1})\left(-\frac{Q}{\lambda}\right) + 1.$$

Et donc en posant $\beta = -\frac{1}{\lambda}$ et $P = \frac{Q}{\lambda}$, on a bien le résultat désiré.

Alternative
$f - \alpha_{k+1}\text{id}_E$ est un polynôme en f, donc commute avec f. Or si deux endomorphismes commutent, il est classique que l'image de l'un est stable par l'autre.

4.a.i. Soit $y \in \text{Im}(f - \alpha_{k+1}\text{id}_E)$. Alors il existe $x \in E$ tel que $y = f(x) - \alpha_{k+1}x$.
On a alors $f(y) = f^2(x) - \alpha_{k+1}f(x) = (f - \alpha_{k+1}\text{id}_E)(f(x)) \in F$.

4.a.ii. Soit $y \in F$. Alors il existe $x \in E$ tel que $y = (f - \alpha_{k+1}\text{id}_E)(x)$.
Et donc, en notant que pour $v \in F$, $g(v) = f(v)$ et $\text{id}_F(v) = \text{id}_E(v) = v$, il vient

$$(g - \alpha_1\text{id}_F) \circ \cdots \circ (g - \alpha_k\text{id}_F)(y) = (f - \alpha_1\text{id}_E) \circ \cdots \circ (g - \alpha_k\text{id}_E) \circ (f - \alpha_{k+1}\text{id}_E)(x) = 0_E.$$

4.a.iii. La relation que nous venons de prouver, couplée à l'hypothèse de récurrence prouve que g est diagonalisable.

D'autre part, nous venons de prouver que $\prod_{i=1}^{k}(X - \alpha_i)$ est annulateur de g, et donc les valeurs propres de g sont parmi $\alpha_1, \dots, \alpha_k$.
Puisque g est diagonalisable, F est somme directe des sous-espaces propres de g et on a donc $F = \bigoplus_{i=1}^{k} \text{Ker}(g - \alpha_i\text{id}_F)$, où certains de ces noyaux sont éventuellement réduits à $\{0_E\}$ si α_i n'est pas valeur propre de g.
Et par conséquent $v = v_1 + \dots v_k$, avec $v_j \in \text{Ker}(g - \alpha_j\text{id}_F)$.
Mais il est évident que $\text{Ker}(g - \alpha_i\text{id}_F) \subset \text{Ker}(f - \alpha_i\text{id}_E)$.
Et donc on a bien $v = v_1 + \cdots + v_k$, avec $v_j \in \text{Ker}(f - \alpha_j\text{id}_E)$.

4.b. En utilisant la relation de la question 3.a, on a

$$\mathrm{id}_E = (f-\alpha_{k+1}\mathrm{id}_E)\circ P(f) + \beta(f-\alpha_1\mathrm{id}_E)\circ\cdots\circ(f-\alpha_k\mathrm{id}_E).$$

Et donc pour $x\in E$, on a

$$x = \mathrm{id}_E(x) = \underbrace{(f-\alpha_{k+1}\mathrm{id}_E)(P(f)(x))}_{\in F} + \beta(f-\alpha_1\mathrm{id}_E)\circ\cdots\circ(f-\alpha_k\mathrm{id}_E)(x).$$

Mais

$$\begin{aligned}(f-\alpha_{k+1}\mathrm{id}_E)\left(\beta(f-\alpha_1\mathrm{id}_E)\circ\cdots\circ(f-\alpha_k\mathrm{id}_E)(x)\right) &= \beta(f-\alpha_{k+1}\mathrm{id}_E)\circ(f-\alpha_1\mathrm{id}_E)\circ\cdots\circ(f-\alpha_k\mathrm{id}_E)(x)\\ &= \beta(f-\alpha_1\mathrm{id}_E)\circ\cdots\circ(f-\alpha_k\mathrm{id}_E)\circ(f-\alpha_{k+1}\mathrm{id}_E)(x)\end{aligned}$$

Détails
Les $f-\alpha_i\mathrm{id}_E$ sont des polynômes en f, et donc commutent deux à deux.

Et donc $\beta(f-\alpha_1\mathrm{id}_E)\circ\cdots\circ(f-\alpha_k\mathrm{id}_E)(x)\in\mathrm{Ker}(f-\alpha_{k+1}\mathrm{id}_E)$.
Puisque $(f-\alpha_{k+1}\mathrm{id}_E)(P(f)(x))\in F$, il s'écrit sous la forme $v_1+\cdots+v_k$, avec $v_i\in\mathrm{Ker}(f-\alpha_i\mathrm{id}_E)$.
Et donc $x\in\sum_{i=1}^{k+1}\mathrm{Ker}(f-\alpha_i\mathrm{id}_E)$.
Ainsi, $E=\sum_{i=1}^{k+1}\mathrm{Ker}(f-\alpha_i\mathrm{id}_E)$, et puisqu'on sait que des sous-espaces propres de f sont en somme directe, il vient

$$E=\bigoplus_{i=1}^{k+1}E_{\alpha_i}(f).$$

Remarque
Si certains des α_i ne sont pas valeurs propres de f, alors $\mathrm{Ker}(f-\alpha_i\mathrm{id}_E)=\{0_E\}$, et donc peut être enlevé de la somme.

Ainsi, E est somme direct des sous-espaces propres de f : f est diagonalisable.
Remarque *: tout ce que nous venons de faire peut être transposé au cas d'un espace vectoriel réel.*

5. Puisque A^2 est diagonalisable, il existe des complexes $\alpha_1,\dots,\alpha_k$ deux à deux distincts tels que $(A^2-\alpha_1I_n)\cdots(A^2-\alpha_kI_n)=0$.
De plus, nous avons dit à la question 2 que les α_i peuvent être choisis de telle sorte que ce soient exactement les valeurs propres de A^2.
Mais A étant inversible, il en est de même de A^2, et donc $0\notin\mathrm{Spec}(A^2)$, de sorte que les α_i sont tous non nuls.
Mais pour tout $i\in[\![1,k]\!]$, le polynôme $X^2-\alpha_i$ admet deux racines distinctes[2]
Et donc $X^2-\alpha_i=(X-\lambda_i)(X-\mu_i)$ avec $\lambda_i\neq\mu_i$.
D'autre part, pour $i\neq j$, $\lambda_i\neq\lambda_j$ et $\lambda_i\neq\mu_j$ puisque $\lambda_i^2=\alpha_i$ et $\lambda_j^2=\mu_j^2=\alpha_j$.
Donc le polynôme $\prod_{i=1}^{k}(X-\lambda_i)(X-\mu_i)$ est annulateur de A, et à racines simples.
D'après la question 3, la matrice A est donc diagonalisable.
Remarque *: ce résultat n'est plus valable pour des matrices réelles, comme le montre par exemple le cas de la matrice* $\begin{pmatrix}0&1\\-1&0\end{pmatrix}\in\mathcal{M}_2(\mathbf{R})$. *En effet,* $A^2=-I_2$ *est diagonalisable, mais* A *n'est pas diagonalisable sur* $\mathcal{M}_2(\mathbf{R})$ *car elle ne possède pas de valeur propre réelle.*

[2] Car son polynôme dérivé est $2X$ qui n'admet que 0 comme racine. Or 0 n'est pas racine de $X^2-\alpha_i$, et donc ce polynôme ne peut admettre de racine double.

Autrement dit
Le complexe α_i possède deux racines carrées distinctes.

□

1.6.2 Trigonalisation

Diagonaliser un endomorphisme (ou une matrice), c'est trouver une base dans laquelle la matrice de cet endomorphisme est «simple», à savoir diagonale.
Malheureusement, cela n'est pas toujours possible, et lorsqu'on ne peut trouver une telle base, on peut tout de même se poser la question de savoir s'il existe une base dans laquelle la matrice est triangulaire.

Exercice 1.43 (QSP HEC 2015) [HEC15.QSP131] **Facile**

Étude d'une matrice à paramètre

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 $B = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$.

SOLUTION DE L'EXERCICE 1.43

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,

$$M_x - \lambda I_3 = \begin{pmatrix} 1-\lambda & x & 0 \\ 0 & -\lambda & 1 \\ 0 & 1 & -\lambda \end{pmatrix} \underset{L_2 \leftrightarrow L_3}{\longleftrightarrow} \begin{pmatrix} 1-\lambda & x & 0 \\ 0 & 1 & -\lambda \\ 0 & -\lambda & 1 \end{pmatrix} \underset{L_3 \leftarrow L_3 + \lambda L_2}{\longleftrightarrow} \begin{pmatrix} 1-\lambda & x & 0 \\ 0 & 1 & -\lambda \\ 0 & 0 & 1-\lambda^2 \end{pmatrix}.$$

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.
Nous cherchons alors à prouver qu'il existe une base $\mathscr{B} = (e_1, e_2, e_3)$ de $\mathbf{R}^3$ dans la quelle la matrice de f est B.

On a $\text{Mat}_{\mathscr{B}}(f) = B = \begin{array}{c} \begin{matrix} f(e_1) & f(e_2) & f(e_3) \end{matrix} \\ \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix} \end{array} \begin{matrix} e_1 \\ e_2 \\ e_3 \end{matrix}$ si et seulement si $\begin{cases} f(e_1) = -e_1 \\ f(e_2) = e_2 \\ f(e_3) = e_2 + e_3 \end{cases}$

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 $e_2 = (1, 0, 0)$.
Il reste donc à trouver $e_3 = (a, b, c)$ 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.

□

Remarque

Sans calcul, en regardant la première colonne de M_x, il est déjà évident que 1 est valeur propre.
Mais cela ne nous aide pas à trouver les autres valeurs propres (si elles existent).

Détails

Si $x = 0$, toutes les colonnes de M_x sont proportionnelles.

$E_1(f)$

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).

Toute matrice de $\mathscr{M}_n(\mathbf{C})$ est semblable à une matrice triangulaire. Ce résultat est délicat à établir dans le cas général, mais pour les matrices 2×2, on peut utiliser le déterminant pour le prouver.

Exercice 1.44 (QSP ESCP 2017) [ESCP17.QSP01] — Difficile

Une équation matricielle

L'équation $A^2 = A - I_2$, d'inconnue $A \in \mathcal{M}_2(\mathbf{C})$ admet-elle des solutions non diagonalisables ?

Solution de l'exercice 1.44 Commençons par rappeler qu'une matrice de $\mathcal{M}_2(\mathbf{C})$ possède toujours au moins une valeur propre complexe.
Ce résultat est même vrai pour toute matrice de $\mathcal{M}_n(\mathbf{C})$, avec n quelconque, mais il est facile de le retrouver dans le cas d'une matrice 2×2, car un complexe λ est valeur propre de A si et seulement si $A - \lambda I_2$ n'est pas inversible.
Soit encore si et seulement si $\det(A-\lambda I_2) = 0 \Leftrightarrow \lambda^2 - \mathrm{tr}(A)\lambda + \det(A) = 0$.
Or, un polynôme à coefficients complexes possède toujours au moins une solution dans $\mathbf{C}$, et donc A possède au moins une valeur propre complexe.

Si $A \in \mathcal{M}_2(\mathbf{C})$ n'est pas diagonalisable, elle ne peut posséder deux valeurs propres distinctes, et donc possède une unique valeur propre λ, avec $\dim E_\lambda(A) = 1$.
Soit f_A l'endomorphisme de $\mathcal{M}_{2,1}(\mathbf{C})$ canoniquement associé à A. Alors f_A possède λ comme unique valeur propre.
Soit $e_1 \in \mathcal{M}_{2,1}(\mathbf{C})$ un vecteur propre de f_A, et soit $e_2 \in \mathcal{M}_{2,1}(\mathbf{C})$ tel que (e_1, e_2) forme une base de $\mathcal{M}_{2,1}(\mathbf{C})$.
Alors la matrice de f_A dans la base (e_1, e_2) est de la forme

> **Remarque**
> Un tel e_2 existe nécessairement d'après le théorème de la base incomplète.

$$\begin{array}{c} \begin{matrix} f_A(e_1) & f_A(e_2) \end{matrix} \\ \begin{pmatrix} \lambda & a \\ 0 & \mu \end{pmatrix} \begin{matrix} e_1 \\ e_2 \end{matrix} \end{array}.$$

Puisqu'elle est triangulaire, ses valeurs propres sont λ et μ, et puisque f_A ne possède que λ comme valeur propre, nécessairement $\lambda = \mu$.
Ainsi, A est semblable à $T = \begin{pmatrix} \lambda & a \\ 0 & \lambda \end{pmatrix}$, où $a \neq 0$ puisque A n'est pas diagonalisable.
Il existe donc $P \in GL_2(\mathbf{C})$ telle que $A = P^{-1}TP$.
On a alors $A^2 = P^{-1}T^2P$, et donc

> **Terminologie**
> On dit alors qu'on a trigonalisé A : elle est semblable à une matrice triangulaire.
> En réalité, toute matrice de $\mathcal{M}_n(\mathbf{C})$ est semblable à une matrice triangulaire, mais il serait fastidieux de le prouver avec les outils du programme d'ECS.

$$A^2 = A - I_2 \Leftrightarrow P^{-1}T^2P - P^{-1}TP + I_2 = 0 \Leftrightarrow P^{-1}\left(T^2 - T + I_2\right)P = 0 \Leftrightarrow T^2 - T + I_2 = 0.$$

Mais on a $T^2 = \begin{pmatrix} \lambda^2 & 2\lambda a \\ 0 & \lambda^2 \end{pmatrix}$, de sorte qu'on cherche à résoudre

$$\begin{pmatrix} \lambda^2 - \lambda + 1 & 2\lambda a - a \\ 0 & \lambda^2 - \lambda + 1 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}.$$

On doit alors avoir $a(2\lambda - 1) = 0$, et a étant non nul, ceci est équivalent à $2\lambda - 1 = 0 \Leftrightarrow \lambda = \dfrac{1}{2}$.
Or, pour $\lambda = \dfrac{1}{2}$, $\lambda^2 - \lambda + 1 \neq 0$. Et donc il n'existe pas de matrice $T = \begin{pmatrix} \lambda & a \\ 0 & \lambda \end{pmatrix}$ telle que $T^2 - T + I_2 = 0$, et donc pas de solution non diagonalisable de $A^2 = A - I_2$.

En revanche, il existe bien des solutions diagonalisables, car si λ, μ sont deux racines[1] de $X^2 - X + 1 = 0$, alors $A = \begin{pmatrix} \lambda & 0 \\ 0 & \mu \end{pmatrix}$ vérifie bien $A^2 = A - I_2$. □

[1] Éventuellement égales.

Exercice 1.45 (QSP HEC 2016) [HEC16.QSP170] — Moyen

Endomorphismes de noyau et d'image prescrits

On considère les deux sous-espaces vectoriels F et G de $\mathbf{R}^3$ définis par :

$$F = \text{Vect}((1,1,1)),\ G = \{(x,y,z) \in \mathbf{R}^3 : x+y-2z = 0\}.$$

Trouver un endomorphisme de $\mathbf{R}^3$ dont le noyau est F et l'image G.

SOLUTION DE L'EXERCICE 1.45 Une première idée serait de prendre la projection sur G parallèlement à F. Malheureusement, $(1,1,1) \in F \cap G$, de sorte que F et G ne sont pas supplémentaires dans $\mathbf{R}^3$, donc cette projection n'est pas définie.

Une base de G est $(1,1,1),(1,0,-2)$.
Complétons la famille libre $(1,1,1),(1,0,-2)$ en une base de $\mathbf{R}^3$, par exemple en posant $e_1 = (1,1,1), e_2 = (1,0,-2)$ et $e_3 = (0,1,0)$.
Soit alors f l'endomorphisme de $\mathbf{R}^3$ dont la matrice dans la base $\mathscr{B} = (e_1,e_2,e_3)$ est

$$\text{Mat}_{\mathscr{B}}(f) = \begin{array}{c} \begin{matrix} f(e_1) & f(e_2) & f(e_3) \end{matrix} \\ \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \end{array} \begin{matrix} e_1 \\ e_2 \\ e_3 \end{matrix}.$$

Détails
Le fait qu'il s'agisse bien d'une base de $\mathbf{R}^3$ est laissé en détails au lecteur.

Rappel
L'image d'une base est toujours une famille génératrice de $\text{Im}(f)$.

Alors il est évident que $\text{Im}(f) = \text{Vect}(f(e_1),f(e_2),f(e_3)) = \text{Vect}(e_1,e_2) = G$.
De plus, $x = \lambda_1 e_1 + \lambda_2 e_2 + \lambda_3 \in \mathbf{R}^3$ est dans $\text{Ker}\, f$ si et seulement si $\lambda_2 = \lambda_3 = 0$, c'est-à-dire si et seulement si $x \in \text{Vect}(e_1) = F$.
Donc $\text{Ker}\, f = F$.

L'endomorphisme que nous avons choisi n'est pas diagonalisable.
Supposons qu'il soit possible de choisir un tel f diagonalisable. Alors nécessairement $E_0(f) = \text{Ker}(f) = F$.
De plus, il existerait alors une base (u_1,u_2,u_3) de $\mathbf{R}^3$ formée de vecteurs propres de f.
Quitte à renuméroter les vecteurs, on peut supposer que $u_1 \in E_0(f)$.
Comme $\dim E_0(f) = 1$, on a alors $F = E_0(f) = \text{Vect}(u_1)$.
D'autre part, $E_0(f)$ étant de dimension 1, les deux autres vecteurs u_2 et u_3 sont nécessairement associés à des valeurs propres non nulles λ_2 et λ_3.
Mais alors $f(u_2) = \lambda_2 u_2 \Leftrightarrow u_2 = \dfrac{1}{\lambda_2} f(u_2) \in \text{Im}\, f$.
Et de même $u_3 \in \text{Im}(f)$.
Ainsi, (u_2,u_3) est une famille libre[1] de $\text{Im}(f)$, qui est de dimension 2 : c'est une base de $\text{Im}(f)$.
Donc $G = \text{Im}(f) = \text{Vect}(u_2,u_3)$. Or, (u_1,u_2,u_3) étant une base de $\mathbf{R}^3$, la concaténation de bases de F et de G est une base de $\mathbf{R}^3$, de sorte que $\mathbf{R}^3 = F \oplus G$, ce qui n'est pas le cas.
Donc il n'existe pas d'endomorphisme diagonalisable tel que $\text{Ker}(f) = F$ et $\text{Im}(f) = G$.

[1] Car sous-famille d'une base de $\mathbf{R}^3$.

Remarque *: nous venons en fait de prouver sur un cas particulier un résultat qui reste valable dans un cadre bien plus général : si f est diagonalisable, alors $E = \text{Im}(f) \oplus \text{Ker}(f)$.* □

1.6.3 A classer

EXERCICE 1.46 (HEC 2007) [HEC07.109] **Moyen**

Preuve d'une formule sur les coefficients binomiaux via l'étude d'un endomorphisme de $\mathbf{R}_n[X]$

Soit n un entier naturel non nul et $E = \mathbf{R}_n[X]$. Soit φ l'application définie sur E par :

$$\forall P \in E,\ [\varphi(P)](X) = P(X+1) - P(X).$$

1. (a) Vérifier que $\varphi \in \mathscr{L}(E)$ et expliciter sa matrice A dans la base canonique $\mathscr{B} = (1,X,X^2,\ldots,X^n)$ de E. On précisera l'élément $A_{i,j}$ situé à la $i^{\text{ème}}$ ligne et $j^{\text{ème}}$ colonne.
   (b) Déterminer le noyau et l'image de φ.

(c) Déterminer un polynôme annulateur de φ.

(d) Étudier la diagonalisabilité de φ.

2. (a) Soit $P \in E$. Montrer que :

$$\varphi^n(P)(X) = (-1)^n \sum_{k=0}^{n} \left[(-1)^k \binom{n}{k} P(X+k) \right],$$

où φ^n désigne la composée n fois : $\varphi \circ \varphi \circ \cdots \circ \varphi$.

(b) En déduire, pour $j \in \{0, 1, \ldots, n-1\}$ la valeur de :

$$S_j = \sum_{k=0}^{n} \left[(-1)^k \binom{n}{k} k^j \right].$$

3. Retrouver le résultat de la question précédente pour $j \in \{0, 1, 2\}$ en considérant la fonction f_n définie sur $\mathbf{R}$ par $f_n(x) = (1-x)^n$.

Solution de l'exercice 1.46

1.a. Il est évident que φ est linéaire.
De plus, si $P \in \mathbf{R}_n[X]$, alors $\deg(P(X+1)) = \deg(P)$ et donc $\deg(\varphi(P)) \leqslant \deg(P) \leqslant n$.
Et donc φ est à valeurs dans E : c'est un endomorphisme de E.

Danger !
Attention à ne pas affirmer que le degré de $\varphi(P)$ est **égal** à celui de P, le degré d'une somme n'est pas forcément égal au degré maximal des termes de la somme.
Par exemple,
$X^2 = (X^3) - (X^3 - X^2)$ n'est pas de degré 3.

On a alors, pour tout $j \in [\![0, n]\!]$,

$$\varphi\left(X^j\right) = (X+1)^j - X^j = \sum_{k=0}^{j} \binom{j}{k} X^k - X^j = \sum_{k=0}^{j-1} \binom{j}{k} X^j.$$

Et donc la matrice A est donnée par

Pour donner une formule pour $a_{i,j}$, il faut prendre garde aux décalages d'indices : le $j^{\text{ème}}$ élément de la base $\mathcal{B}$ est X^{j-1} et non X^j.
Et donc pour $(i,j) \in [\![1, n+1]\!]$, $a_{i,j}$ est la composante suivant X^{i-1} de $\varphi\left(X^{j-1}\right)$, c'est donc $a_{i,j} = \binom{j-1}{i-1}$.

1.b. Les calculs effectués précédemment prouvent que pour $k \in [\![1, n]\!]$, $\varphi(X^k)$ est de degré $k-1$, et $\varphi(1) = 0$.
Ceci prouve donc déjà que $1 \in \mathrm{Ker}(\varphi)$, et donc $\mathbf{R}_0[X] = \mathrm{Vect}(1)$, l'ensemble des polynômes constants est inclus dans $\mathrm{Ker}(\varphi)$.

D'autre part, si $P = \sum_{k=0}^{d} a_k X^k$ est un élément de E de degré égal à $d \geqslant 1$, on a alors

$$\varphi(P) = \sum_{k=0}^{d} a_k \varphi(X^k).$$

Mais $a_d \varphi(X^d)$ est degré égal[1] à $d-1$, et pour tout $k \in [\![0, d-1]\!]$, $a_k \varphi(X^k)$ est de degré inférieur ou égal à $d-2$.
Et donc $\varphi(P)$ est de degré égal à $d-1$. Et en particulier, $\varphi(P) \neq 0$.
Ceci prouve donc que $\mathrm{Ker}(\varphi) \subset \mathbf{R}_0[X]$, et l'inclusion réciproque ayant déjà été prouvée, on a $\mathrm{Ker}(\varphi) = \mathbf{R}_0[X]$.

[1] Il est important de remarquer que $a_d \neq 0$ puisque P est de degré **égal** à d.

Par le théorème du rang, on en déduit que $\dim \mathrm{Im}(\varphi) = \dim E - \dim \mathrm{Ker}(\varphi) = n+1-1 = n$.
D'autre part, puisque nous venons de prouver que $\deg(\varphi(P)) \leqslant \deg(P)-1$, $\mathrm{Im}(\varphi) \subset \mathbf{R}_{n-1}[X]$, qui est de dimension n.
Et donc on a $\mathrm{Im}(\varphi) = \mathbf{R}_{n-1}[X]$.

1.c. Puisque φ diminue le degré d'une unité, si $P \in \mathrm{Ker}(\varphi)$, on a $\deg(\varphi(P)) \leqslant \deg(P) - 1$.
Puis $\deg(\varphi^2(P)) \leqslant \deg(\varphi(P)) - 1 \leqslant \deg(P) - 2$, et de proche en proche, $\deg(\varphi^k(P)) \leqslant$

$\deg(P) - k$.
Et en particulier, $\deg(\varphi^{n+1}(P)) \leqslant \deg(P) - (n+1) \leqslant n - (n+1) \leqslant -1$.
Cela signifie donc que pour tout $P \in E$, $\varphi^{n+1}(P) = 0$ et donc $\varphi^{n+1} = 0$: X^{n+1} est un polynôme annulateur de φ.

1.d. Puisque A est triangulaire, ses valeurs propres sont ses coefficients diagonaux. Et donc φ ne possède que 0 comme valeur propre.
On a alors $\dim E_0(\varphi) = \dim \operatorname{Ker}(\varphi) = 1 < \dim E$, et donc φ n'est pas diagonalisable.

2.a. Prouvons par récurrence sur $i \in \mathbf{N}^*$ que pour tout $P \in E$,

$$\varphi^i(P)(X) = (-1)^i \sum_{k=0}^{i} \left[(-1)^k \binom{i}{k} P(X+i) \right].$$

Pour $i = 1$, le membre de droite est

$$(-1)^1 \left[(-1)^0 P(X) + (-1)^1 P(X+1) \right] = P(X+1) - P(X) = \varphi(P).$$

Donc la récurrence est initialisée.

Supposons que $\varphi^i(P)(X) = (-1)^i \sum_{k=0}^{i} \left[(-1)^k \binom{i}{k} P(X+i) \right]$. Alors

$$\begin{aligned}
\varphi^{i+1}(P)(X) = \varphi\left(\varphi^i(P)\right)(X) &= (-1)^i \sum_{k=0}^{i} \left[(-1)^k \binom{i}{k} \varphi(P(X+k)) \right] \\
&= (-1)^i \sum_{k=0}^{i} \left[(-1)^k \binom{i}{k} (P(X+k+1) - P(X+k)) \right] \\
&= (-1)^i \left[\sum_{k=0}^{i} (-1)^k \binom{i}{k} P(X+k+1) - \sum_{k=0}^{i} (-1)^k \binom{i}{k} P(X+k) \right] \\
&= (-1)^i \left[\sum_{k=1}^{i+1} (-1)^{k-1} \binom{i}{k-1} P(X+k) - \sum_{k=0}^{i} (-1)^k \binom{i}{k} P(X+k) \right] \\
&= (-1)^{i+1} \left[(-1)^0 \binom{i}{0} P(X) + \sum_{k=1}^{i} (-1)^k \left(\binom{i}{k-1} + \binom{i}{k} \right) P(X+k) + (-1)^{i+1} P(X+i+1) \right] \\
&= (-1)^{i+1} \left[(-1)^0 \binom{i+1}{0} P(X) + \sum_{k=1}^{i} (-1)^k \binom{i+1}{k} P(X+k) + (-1)^{i+1} \binom{i+1}{i+1} P(X+i+1) \right] \quad \binom{i}{k-1} + \binom{i}{k} = \binom{i+1}{k}. \\
&= (-1)^{i+1} \left[\sum_{k=0}^{i+1} (-1)^k \binom{i+1}{k} P(X+k) \right].
\end{aligned}$$

Et donc par le principe de récurrence, notre formule est vraie pour tout $i \geqslant 1$.
Et donc en particulier elle l'est pour $i = n$:

$$\varphi^n(P)(X) = \sum_{k=0}^{n} \left[(-1)^k \binom{n}{k} P(X+k) \right].$$

2.b. Puisque φ^n abaisse le degré de n unités, pour $j \in \{0, 1, \ldots, n-1\}$, $\varphi^n\left(X^j\right) = 0$.
Et donc par la formule de la question précédente,

$$\sum_{k=0}^{n} (-1)^k \binom{n}{k} (X+k)^j = 0.$$

En particulier, pour $X = 0$,

$$S_j = \sum_{k=0}^{n} (-1)^k \binom{n}{k} k^j = 0.$$

3. On a $f_n(1) = 1 = \sum_{k=0}^{n} \binom{n}{k}(-1)^{n-k} = (-1)^n S_0$, et donc $S_0 = 0$.

Comme de plus on a $f_n(x) = \sum_{k=0}^{n} \binom{n}{k}(-1)^k x^k$, alors

$$-n(1-x)^{n-1} = f_n'(x) = \sum_{k=1}^{n} \binom{n}{k}(-1)^k k x^{k-1}.$$

Et là encore, en évaluant cette relation en $x = 1$, il vient

$$\sum_{k=0}^{n} \binom{n}{k}(-1)^k k = 0 \Leftrightarrow S_1 = 0.$$

Enfin, on a

$$n(n-1)(1-x)^{n-2} = f_n''(x) = \sum_{k=2}^{n} k(k-1)\binom{n}{k}(-1)^k x^{k-2}.$$

□

1.7 Espaces euclidiens

Une conséquence classique de la bilinéarité du produit scalaire est l'identité $\|x+y\|^2 = \|x\|^2 + \|y\|^2 + 2\langle x, y\rangle$, qui permet de calculer la norme d'une somme si on connaît les normes des deux vecteurs x et y ainsi que leur produit scalaire.
Notons que cette identité permet de démontrer immédiatement le théorème de Pythagore.
Mais dès qu'on connaît trois des quatre termes de cette égalité, on peut en déduire la valeur du quatrième. Et en particulier si l'on connaît les normes de $x+y$, de x et de y alors on obtient la valeur du produit scalaire $\langle x, y\rangle$:

$$\langle x, y\rangle = \frac{1}{2}\left(\|x+y\|^2 - \|x\|^2 - \|y\|^2\right).$$

Cette identité est appelée identité de polarisation, et bien qu'elle ne figure pas explicitement au programme, il est bon de savoir qu'elle existe et d'être capable de la retrouver.

Il existe aussi une seconde identité de polarisation, qui se démontre sur le même principe, et qui est $\langle x, y\rangle = \frac{1}{4}\left(\|x+y\|^2 - \|x-y\|^2\right)$.

Exercice 1.47 (QSP ESCP 2011) [ESCP11.QSP01] **Moyen**

Une conditions suffisante pour qu'une famille de vecteurs unitaires soit une base.

Soit E un espace euclidien de dimension n, et soit $(e_1, \dots, e_n)$ une famille de n vecteurs unitaires.
On suppose que

$$\forall (i,j) \in [\![1,n]\!]^2, i \neq j \Rightarrow \|e_i - e_j\| = 1.$$

Montrer que $(e_1, \dots, e_n)$ est une base de E.

Solution de l'exercice 1.47 D'après l'identité de polarisation, pour $i \neq j$, on a

$$\langle e_i, e_j\rangle = -\langle e_i, -e_j\rangle = -\frac{1}{2}\left(\|e_i - e_j\|^2 - \|e_i\|^2 - \|-e_j\|^2\right) = \frac{1}{2}\left(\|e_i\|^2 + \|e_j\|^2 - \|e_i - e_j\|^2\right) = \frac{1}{2}.$$

Méthode
Nous connaissons les normes de e_i, e_j et de $e_i - e_j$, si on souhaite obtenir la valeur de $\langle e_i, e_j\rangle$, il faut donc penser à l'identité de polarisation.

Soient $\lambda_1, \dots, \lambda_n$ des réels tels que $\sum_{i=1}^{n} \lambda_i e_i = 0_E$.

Alors en prenant le produit scalaire avec e_j, il vient

$$\forall j \in [\![1,n]\!],\ \lambda_j \langle e_j, e_j\rangle + \sum_{\substack{i=1 \\ i\neq j}}^{n} \lambda_i \langle e_i, e_j\rangle = \langle 0_E, e_j\rangle = 0$$

soit encore

$$\forall j \in [\![1,n]\!],\ \lambda_j + \frac{1}{2}\sum_{\substack{i=1 \\ i\neq j}}^{n} \lambda_i = 0 \Leftrightarrow 2\lambda_j = -\sum_{\substack{i=1 \\ j\neq i}}^{n} \lambda_i \Leftrightarrow \lambda_j = -\sum_{i=1}^{n} \lambda_i.$$

En particulier, on a $\lambda_1 = \lambda_2 = \cdots = \lambda_n$, et donc

$$\forall j \in [\![1, n]\!], \lambda_j = -n\lambda_j$$

ce qui implique nécessairement $\lambda_1 = \cdots = \lambda_n = 0$.

Nous venons de prouver que la famille $(e_1, \ldots, e_n)$ est libre, et donc étant de cardinal $n = \dim E$, c'est une base de E. □

EXERCICE 1.48 (ESCP 2006) [ESCP06.2.1]

Endomorphismes orthogonaux et commutant de $\mathcal{O}_n(\mathbf{R})$.

Soit n un entier supérieur ou égal à 2.
On pose $E = \mathbf{R}^n$ muni de son produit scalaire canonique, noté $\langle \cdot, \cdot \rangle$. On note $\| \cdot \|$ la norme euclidienne associée.
Un endomorphisme g de E est dit *orthogonal* si pour tout $(x, y) \in E^2$, on a :

$$\langle g(x), g(y) \rangle = \langle x, y \rangle.$$

1. Soit g un endomorphisme de E. Montrer que les trois assertions suivantes sont équivalentes :
   i) g et orthogonal
   ii) Pour tout $x \in E$, $\|g(x)\| = \|x\|$.
   iii) L'image par g d'une base orthonormée de E est une base orthonormée de E.

   On note $\mathcal{O}_n(\mathbf{R})$ l'ensemble des endomorphismes orthogonaux de E.
2. Soit F un sous-espace vectoriel de E. Montrer que F est stable par tous les éléments de $\mathcal{O}_n(\mathbf{R})$ si et seulement si $F = \{0_E\}$ ou $F = E$ (on pourra montrer que si $F \neq \{0_E\}$ et $F \neq E$, il existe un élément de $\mathcal{O}_n(\mathbf{R})$, que l'on exhibera, qui ne laisse pas F stable).
3. Soit f un endomorphisme de E. On suppose, pour toute la suite de l'exercice, que f commute avec tous les éléments de $\mathcal{O}_n(\mathbf{R})$. Montrer que
   (a) $\operatorname{Ker} f$ est stable par tous les éléments de $\mathcal{O}_n(\mathbf{R})$.
   (b) Pour tout réel λ, $\operatorname{Ker}(f - \lambda \mathrm{id}_E)$ est stable par tous les éléments de $\mathcal{O}_n(\mathbf{R})$.
   (c) En déduire que pour tout réel λ, $\operatorname{Ker}(f - \lambda \mathrm{id}_E) = \{0_E\}$ ou $\operatorname{Ker}(f - \lambda \mathrm{id}_E) = E$.
4. On suppose que n est impair, et on admet qu'alors tout endomorphisme de $\mathbf{R}^n$ admet au moins une valeur propre réelle.
   Montrer qu'il existe $\lambda_0 \in \mathbf{R}$ tel que $f = \lambda_0 \mathrm{id}_E$.

SOLUTION DE L'EXERCICE 1.48

1. Commençons par prouver que *i)* ⇔ *ii)*.
i) ⇒ *ii)* Supposons que g soit orthogonal, et soit $x \in E$. Alors

$$\|g(x)\|^2 = \langle g(x), g(x) \rangle = \langle x, x \rangle = \|x\|^2.$$

Et donc $\|g(x)\|$.

ii) ⇒ *i)*. Inversement, supposons que pour tout $x \in E$, $\|g(x)\| = \|x\|$.
Soient alors $x, y \in E$. Nous savons que

$$\langle x, y \rangle = \frac{1}{2}\left(\|x + y\|^2 - \|x\|^2 - \|y\|^2\right)$$

et de même

$$\langle g(x), g(y) \rangle = \frac{1}{2}\left(\|g(x) + g(y)\|^2 - \|g(x)\|^2 - \|g(y)\|^2\right) = \frac{1}{2}\left(\|g(x + y)\|^2 - \|g(x)\|^2 - \|g(y)\|^2\right).$$

Et donc si g préserve la norme, il vient

$$\langle g(x), g(y) \rangle = \frac{1}{2}\left(\|x + y\|^2 - \|x\|^2 - \|y\|^2\right) = \langle x, y \rangle.$$

Et donc g est orthogonal.

i) $\Rightarrow$ *iii)*. Supposons g orthogonal, et soit $(e_1, \dots, e_n)$ une base orthonormée de E. Alors, pour tout $(i,j) \in [\![1,n]\!]^2$, on a

$$\langle g(e_i), g(e_j)\rangle = \langle e_i, e_j\rangle = \begin{cases} 1 & \text{si } i = j \\ 0 & \text{sinon} \end{cases}$$

Et donc $(g(e_1), \dots, g(e_n))$ est orthonormée. En particulier elle est libre, et étant de cardinal $n = \dim E$, c'est une base orthonormée de E.

Orthonormée
Toute famille orthonormée de cardinal $\dim E$ est une base de E (qui est bien entendu orthonormée).

iii) $\Rightarrow$ *ii)*. Supposons à présent que l'image d'une base orthonormée $(e_1, \dots, e_n)$ soit une base orthonormée de E.

Soit alors $x \in E$. De manière unique, $x = \sum_{i=1}^{n} \lambda_i e_i$. Et alors[1], $\|x\|^2 = \sum_{i=1}^{n} \lambda_i^2$.

[1] C'est ici qu'il est indispensable que la base soit orthonormée.

De plus, par linéarité de g, $g(x) = \sum_{i=1}^{n} \lambda_i g(e_i)$. Et puisque $(g(e_1), \dots, g(e_n))$ est orthonormée, alors on a

$$\|g(x)\|^2 = \sum_{i=1}^{n} \lambda_i^2 = \|x\|^2.$$

Et donc nous avons bien prouvé que *i)* $\Leftrightarrow$ *ii)* $\Leftrightarrow$ *iii)*.

Remarque
Un sev qui n'est ni réduit à $\{O_E\}$ ni égal à E tout entier n'est ni de dimension 0 ni de dimension n.

2. Soit F un sous-espace vectoriel de E, de dimension p, $1 \leqslant p < n$.
Soit $(e_1, \dots, e_p)$ une base orthonormée de F, et soit $(e_{p+1}, \dots, e_n)$ une base orthonormée de $F^\perp$, de sorte que $(e_1, \dots, e_n)$ soit une base orthonormée de E.
Soit alors g l'unique endomorphisme de E défini par

$$\forall i \in [\![1, n-1]\!],\ g(e_i) = e_{i+1} \text{ et } g(e_n) = e_1.$$

Unique ?
Pour définir une application linéaire, il suffit de définir l'image d'une base.
En effet, cela revient à donner sa matrice, ce qui caractérise de manière unique l'application linéaire.

Alors l'image de la base orthonormée $(e_1, \dots, e_n)$ est la base orthonormée $(e_2, \dots, e_n, e_1)$, et donc g est orthogonal d'après le *iii)* de la première question.
Et on a $g(e_p) = e_{p+1} \notin F$, de sorte que F n'est pas stable par g.
Ainsi, F n'est pas stable par tous les éléments de $\mathcal{O}_n(\mathbf{R})$.

À l'inverse, il est évident que $\{0_E\}$ et E tout entier sont stables par tous les endomorphismes de E, et en particulier par tous les endomorphismes orthogonaux.

3.a. C'est un grand classique : si g et g commutent, alors $\mathrm{Ker}(f)$ est stable par g.
En effet, si $x \in \mathrm{Ker}(f)$, alors

$$f(g(x)) = g(f(x)) = g(0_E) = 0_E$$

et donc $g(x) \in \mathrm{Ker}\, f$.

3.b. Il suffit de noter que si f et g commutent, alors $f - \lambda \mathrm{id}_E$ et g commutent car

$$(f - \lambda \mathrm{id}_E) \circ g = f \circ g - \lambda g = g \circ f - \lambda g = g \circ (f - \lambda \mathrm{id}_E).$$

Et il suffit alors d'appliquer la question précédente : pour tout $g \in \mathcal{O}_n(\mathbf{R})$, $\ker(f - \lambda \mathrm{id}_E)$ est stable par g.

3.c. Il suffit de combiner les questions 3.b et 2 : $\mathrm{Ker}(f - \lambda \mathrm{id}_E)$ est un sous-espace vectoriel de E, stable par tous les éléments de $\mathcal{O}_n(\mathbf{R})$: il est donc soit égal à E tout entier, soit réduit à $\{0_E\}$.

4. Le résultat admis dans l'énoncé nous dit que f admet une valeur propre réelle λ_0.
Et donc $\mathrm{Ker}(f - \lambda_0 \mathrm{id}_E)$ n'est pas réduit à $\{0_E\}$. D'après la question 3.c, on a donc $\mathrm{Ker}(f - \lambda_0 \mathrm{id}_E) = E$.
Et donc[2] $f - \lambda_0 \mathrm{id}_E = 0 \Leftrightarrow f = \lambda_0 \mathrm{id}_E$.

[2] Le seul endomorphisme de noyau égal à E est l'endomorphisme nul.

Commentaires *: disons quelques mots à propos du résultat admis. Il est fortement relié au fait que tout polynôme P à coefficients réels de degré impair possède au moins une racine. Ceci est une conséquence du théorème des valeurs intermédiaires : la limite en $+\infty$ est égale à $\pm\infty$ (le signe dépendant du signe du coefficient dominant), et par imparité du degré* $\lim_{x\to-\infty} P(x) = -\lim_{x\to+\infty} P(x)$.

*Or, si $M \in \mathcal{M}_n(\mathbf{R})$, il existe un polynôme P de degré n tel que les valeurs propres de M soient **exactement** les racines de P. Ce résultat ne figure pas au programme d'ECS, mais nous le connaissons pour les matrices 2×2 : un réel λ est valeur propre de $M = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ si et seulement si*

$$\det(M - \lambda I_2) = 0 \Leftrightarrow \lambda^2 - (a+d)\lambda + (ad - bc) = 0.$$

□

EXERCICE 1.49 (QSP HEC 17) [HEC17.QSP164] **Facile**

Stricte convexité de la sphère unité

Soit $(E, \langle\cdot,\cdot\rangle)$ un espace euclidien et soit $S = \{x \in E : \|x\| = 1\}$.
On note $\mathcal{P}$ la propriété suivant : $\forall (x,y) \in S^2$ avec $x \neq y$, $\forall t \in]0,1[$, $tx + (1-t)y \notin S$.

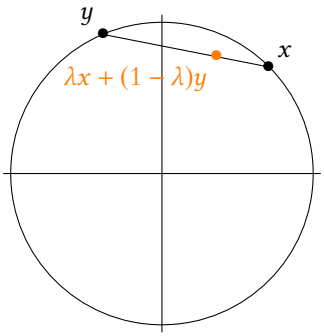
1. Illustrer graphiquement la propriété $\mathcal{P}$ lorsque $E = \mathbf{R}^2$ muni du produit scalaire canonique.
2. Établir la propriété $\mathcal{P}$ dans le cas général (on utilisera pour cela la fonction polynomiale P définie pour tout $t \in \mathbf{R}$ par $P(t) = \|tx + (1-t)y\|^2$).

SOLUTION DE L'EXERCICE 1.49

1. Lorsque $E = \mathbf{R}^2$ muni du produit scalaire canonique, on a

$$S = \{(x,y) \in R^2 : \|(x,y)\| = 1\} = \{(x,y) \in \mathbf{R}^2 : x^2 + y^2 = 1\}$$

qui est donc le cercle de centre $(0,0)$ et de rayon 1.
Soient donc x et y deux points de ce cercle. Les $\lambda x + (1-\lambda)y$, pour $\lambda \in [0,1]$ sont alors les points du segment qui joint x et y.
Et pour $\lambda \notin \{0,1\}$, ce sont les points du segment différents de x et de y.
La propriété $\mathcal{P}$ dit alors que les points de ce segment ne sont plus sur le cercle, mais bien à l'intérieur du cercle.



2. Soit x, y deux éléments distincts de S, et notons, comme indiqué dans l'énoncé

$$P(t) = \|tx + (1-t)y\|^2 = \langle tx + (1-t)y, tx + (1-t)y\rangle = t^2 \underbrace{\|x\|^2}_{=1} + 2t(1-t)\langle x,y\rangle + (1-t)^2 \underbrace{\|y\|^2}_{=1}$$

$$= t^2(2 - \langle x,y\rangle) + 2t(\langle x,y\rangle - 1) + 1.$$

D'après l'inégalité de Cauchy-Schwarz, $|\langle x,y\rangle| \leqslant \|x\|\|y\| \leqslant 1$. Et donc le coefficient de degré 2 de P est non nul, de sorte que P est un polynôme de degré 2.
Mais $P(0) = P(1) = 1$, et donc pour $t \notin \{0,1\}$, $P(t) \neq 1$.
Et donc pour $t \in]0,1[$, $\|tx + (1-t)y\|^2 \neq 1 \Leftrightarrow tx + (1-t)y \notin S$.

□

Détails
Un polynôme P de degré 2 ne peut prendre qu'au plus deux fois une valeur $a \in \mathbf{R}$. En effet, $P - a$ est un polynôme de degré 2 qui possède donc au plus deux racines distinctes.

Exercice 1.50 (QSP HEC 2011) [HEC11.QSP02] **Moyen**

Sur les matrices de passage entre deux bases orthonormées

Soit $A = (a_{i,j})_{1\leqslant i,j\leqslant n}$ la matrice de passage d'une base orthonormée de $\mathbf{R}^n$ à une autre base orthonormée de $\mathbf{R}^n$.

1. Justifier l'inversibilité de A et exprimer A^{-1} en fonction de A.
2. Montrer que $\left|\sum_{i=1}^{n}\sum_{j=1}^{n} a_{i,j}\right| \leqslant n$.

Solution de l'exercice 1.50

1. C'est du cours : la matrice de passage entre deux bases orthonormées est une matrice orthogonale, donc est inversible[1] et son inverse A^{-1} est égale à sa transposée ${}^tA$.

[1] Comme toute matrice de passage.

2. Notons $\langle\cdot,\cdot\rangle$ le produit scalaire canonique de $\mathcal{M}_{n,1}(\mathbf{R})$ défini par $\langle X, Y\rangle = {}^tXY$.

Notons alors $X = \begin{pmatrix}1\\ \vdots\\ 1\end{pmatrix} \in \mathcal{M}_{n,1}(\mathbf{R})$.

On a alors $AX = \begin{pmatrix}\sum_{j=1}^n a_{1,j}\\ \vdots\\ \sum_{j=1}^n a_{n,j}\end{pmatrix}$, et donc

$$\langle AX, X\rangle = \begin{pmatrix}1 & \dots & 1\end{pmatrix}\begin{pmatrix}\sum_{j=1}^n a_{1,j}\\ \vdots\\ \sum_{j=1}^n a_{n,j}\end{pmatrix} = \sum_{i=1}^{n}\sum_{j=1}^{n} a_{i,j}.$$

Par l'inégalité de Cauchy-Schwarz, on a alors

$$|\langle AX, X\rangle| \leqslant \|AX\|\cdot\|X\|.$$

Mais $\|X\| = \sqrt{\langle X, X\rangle} = \sqrt{n}$, et

$$\|AX\| = \sqrt{\langle AX, AX\rangle} = \sqrt{{}^t(AX)AX} = \sqrt{{}^tX{}^tAAX} = \sqrt{{}^tXX} = \sqrt{n}.$$

Et donc il vient bien

$$\left|\sum_{i=1}^{n}\sum_{j=1}^{n} a_{i,j}\right| = |\langle AX, X\rangle| \leqslant \sqrt{n}\sqrt{n} \leqslant n.$$

Remarque *: on serait tentés de raisonner de la même manière, mais en utilisant le produit scalaire canonique de $\mathcal{M}_n(\mathbf{R})$: $\langle A, B\rangle = \operatorname{tr}({}^tAB)$, et de considérer le produit scalaire de A avec la matrice J ne comportant que des 1.*

Mais l'inégalité de Cauchy-Schwarz nous donne alors $\left|\sum_{i=1}^{n}\sum_{j=1}^{n}\right| \leqslant n^{3/2}$, ce qui ne suffit pas.

□

Exercice 1.51 (QSP ESCP 2015) [ESCP15.QSP01] **Moyen**

Matrice de Gram

Soient $A_1, A_2, \dots, A_{n^2}$ n^2 matrice symétriques de $\mathcal{M}_n(\mathbf{R})$.
Montrer que la famille $(A_i)_{1\leqslant i\leqslant n^2}$ est une base de $\mathcal{M}_n(\mathbf{R})$ si et seulement si la matrice $G \in \mathcal{M}_{n^2}(\mathbf{R})$ dont le terme d'indice (i,j) est $\operatorname{tr}(A_iA_j)$ est inversible.

Solution de l'exercice 1.51 Rappelons que sur $\mathcal{M}_n(\mathbf{R})$ il existe un produit scalaire défini par $\langle M, N\rangle = \operatorname{tr}({}^tMN)$.
Puisque les A_i sont symétriques, on a alors, pour tout $(i,j) \in [\![1, n^2]\!]^2$, $\operatorname{tr}(A_iA_j) = \operatorname{tr}\left({}^tA_iA_j\right) =$

$\langle A_i, A_j\rangle$.

Supposons que $(A_i)_{1\leqslant i\leqslant n^2}$ est une base de $\mathcal{M}_n(\mathbf{R})$, et montrons que G est inversible. Notons $C_1,\dots,C_{n^2}$ les colonnes de G, de sorte que pour tout $j\in[\![1,n^2]\!]$,

$$C_j=\begin{pmatrix}\operatorname{tr}(A_1A_j)\\ \vdots\\ \operatorname{tr}(A_{n^2}A_j)\end{pmatrix}=\begin{pmatrix}\langle A_1,A_j\rangle\\ \vdots\\ \langle A_{n^2},A_j\rangle\end{pmatrix}.$$

Soient $\lambda_1,\dots,\lambda_{n^2}$ des réels tels que $\displaystyle\sum_{j=1}^{n^2}\lambda_jC_j=0$.
Alors, par identification des coefficients, il vient

$$\forall i\in[\![1,n]\!],\ \sum_{j=1}^{n^2}\lambda_j\langle A_i,A_j\rangle=0\Leftrightarrow\left\langle A_i,\sum_{j=1}^{n^2}\lambda_jA_j\right\rangle=0.$$

Et donc $\displaystyle\sum_{j=1}^{n^2}\lambda_jA_j\in\operatorname{Vect}(A_1,\dots,A_{n^2})^\perp=\mathcal{M}_n(\mathbf{R})^\perp=\{0\}$.

Par conséquent, $\displaystyle\sum_{j=1}^{n^2}\lambda_jA_j=0$. Puisque $(A_i)_{1\leqslant i\leqslant n^2}$ est une base de $\mathcal{M}_n(\mathbf{R})$, c'est en particulier une famille libre, et donc $\lambda_1=\cdots=\lambda_{n^2}=0$.
On en déduit que la famille $C_1,\dots,C_{n^2}$ est une famille libre de $\mathcal{M}_{n^2,1}(\mathbf{R})$, et étant de cardinal $n^2=\dim\mathcal{M}_{n^2,1}(\mathbf{R})$, c'en est une base.
Et donc $\operatorname{rg}(G)=\dim\operatorname{Vect}(C_1,\dots,C_{n^2})=n^2$, de sorte que G est une matrice inversible de $\mathcal{M}_{n^2}(\mathbf{R})$.

Inversement, supposons que G est inversible. Montrons alors que $(A_i)_{1\leqslant i\leqslant n^2}$ est une famille libre de $\mathcal{M}_n(\mathbf{R})$.

Soient $\lambda_1,\dots,\lambda_{n^2}$ une famille de réels tels que $\displaystyle\sum_{i=1}^{n^2}\lambda_iA_i=0$.

Alors pour tout $j\in[\![1,n^2]\!]$, $\displaystyle\left\langle\sum_{i=1}^{n^2}\lambda_iA_i,A_j\right\rangle=0$.
Et donc avec les notations précédentes,

$$\sum_{i=1}^{n^2}\lambda_iC_i=\begin{pmatrix}\displaystyle\left\langle\sum_{i=1}^{n^2}\lambda_iA_i,A_1\right\rangle\\ \vdots\\ \displaystyle\left\langle\sum_{i=1}^{n^2}\lambda_iA_i,A_{n^2}\right\rangle\end{pmatrix}=0.$$

Puisque G est inversible, la famille de ses vecteurs colonnes est de rang n^2, et donc est une base de $\mathcal{M}_{n^2,1}(\mathbf{R})$, et en particulier est libre.
Et donc $\lambda_1=\cdots=\lambda_{n^2}=0$, de sorte que $(A_i)_{1\leqslant i\leqslant n^2}$ est libre, et donc est une base de $\mathcal{M}_n(\mathbf{R})$.

Quelques commentaires *: le fait que les matrices A_i soient symétriques était en fait peu important ici. On serait notamment tentés de faire référence au théorème spectral et d'utiliser le fait qu'elles soient diagonalisables, ce qui est totalement inutile ici.*
Le même raisonnement montrerait que dans un espace euclidien E de dimension n, un famille de n vecteurs $e_1,\dots,e_n$ est une base si et seulement si la matrice de terme général $\langle e_i,e_j\rangle$ (appelée matrice de Gram[1] de la famille $(e_1,\dots,e_n)$) est inversible. □

[1] C'est le même Gram que celui de Gram-Schmidt : Jørgen Pedersen Gram, mathématicien danois du XIX$^{\text{ème}}$ siècle.

Exercice 1.52 (QSP HEC 2007) [HEC07.QSP104]

Endomorphismes préservant l'orthogonalité

Soit E un espace euclidien. On note $\langle\cdot,\cdot\rangle$ le produit scalaire, et $\|\|$ la norme associée.
Soit f un endomorphisme de E qui vérifie la propriété suivante :

$$\forall (x,y)\in E^2,\ \langle x,y\rangle = 0 \Rightarrow \langle f(x), f(y)\rangle = 0.$$

Montrer qu'il existe $k\in\mathbf{R}_+$ tel que pour tout $x\in E$, $\|f(x)\| = k\|x\|$.

Solution de l'exercice 1.52 Pour tout $x\in E$ non nul, notons $k_x = \dfrac{\|f(x)\|}{\|x\|}\geqslant 0$. Il s'agit alors de prouver que k_x est une constante qui ne dépend pas de x.

Soient donc x et y deux vecteurs non nuls de E. Alors par bilinéarité du produit scalaire,

$$\langle x+y, x-y\rangle = \langle x,x\rangle + \langle x,y\rangle - \langle x,y\rangle - \langle y,y\rangle = \|x\|^2 - \|y\|^2.$$

En particulier, si x et y sont de même norme, alors $x-y$ et $x+y$ sont orthogonaux.
Soient donc $u = \dfrac{x}{\|x\|}$ et $v = \dfrac{y}{\|y\|}$, qui sont deux vecteurs de norme 1.
Alors $u-v$ et $u+v$ sont orthogonaux, de sorte que $\langle f(u+v), f(u-v)\rangle = 0$. Mais

$$\langle f(u+v), f(u-v)\rangle = \langle f(u)+f(v), f(u)-f(v)\rangle = \|f(u)\|^2 - \|f(v)\|^2.$$

Et donc[1], $\|f(u)\|^2 = \|f(v)\|^2$.
Puisque $\|u\| = \|v\| = 1$, on a donc

$$k_u = \frac{\|f(u)\|}{\|u\|} = \|f(u)\| = \|f(v)\| = \frac{\|f(v)\|}{\|v\|} = k_v.$$

Mais puisque $x = \|x\|u$, alors $f(x) = \|x\|f(u)$ et donc

$$k_x = \frac{\|x\|\cdot\|f(u)\|}{\|x\|\|u\|} = k_u.$$

Et de même, $k_y = k_v$, de sorte que $k_x = k_y$.
Ainsi, k_x ne dépend pas de x : il existe $k\in\mathbf{R}_+$ tel que pour tout $x\in E$ non nul, $\dfrac{\|f(x)\|}{\|x\|} = k \Leftrightarrow \|f(x)\| = k\|x\|$.
Et bien évidemment, cette relation reste valable pour $x = 0_E$ car $\|f(0_e)\| = \|0_E\| = 0 = k\|0_E\|$.
Et donc nous avons bien prouvé que pour tout $x\in E$, $Vert f(x)\| = k\|x\|$. □

[1] Une norme est positive.

L'exercice qui suit, bien trop long pour une question sans préparation, nécessite à la fois de l'aisance avec le procédé d'orthonormalisation de Gram-Schmidt et avec la notion de matrice orthogonale.

Exercice 1.53 (QSP HEC 2014) [HEC14.QSP50] Difficile

Existence de bases orthonormées dans lesquelles les coordonnées de vecteurs sont prescrites.

Soit E un espace euclidien de dimension n.

1. Soit $x\in E$. Montrer que $\|x\| = \sqrt{n}$ si et seulement si il existe une base orthonormée $(e_1,\dots,e_n)$ telle que $x = \sum_{i=1}^{n} e_i$.
2. Soient x et y deux vecteurs de E. Donner une condition nécessaire et suffisante pour qu'il existe une base orthonormée $(e_1,\dots,e_n)$ telle que $x = \sum_{k=1}^{n} e_k$ et $y = \sum_{k=1}^{n} k e_k$.

SOLUTION DE L'EXERCICE 1.53

1. Supposons qu'il existe une base orthonormée $(e_1, \dots, e_n)$ telle que $x = \sum_{k=1}^{n} e_k$. Alors

$$\|x\|^2 = \sum_{k=1}^{n} 1^2 = n$$

et donc $\|x\| = \sqrt{n}$.

Inversement, soit $x \in E$ tel que $\|x\| = \sqrt{n}$.
Alors $\dfrac{x}{\sqrt{n}}$ est un vecteur de norme 1. Soit $\mathcal{B} = \left(\dfrac{x}{\sqrt{n}}, f_2, \dots, f_n\right)$ une base orthonormée de E obtenue par complétion de la famille libre[1] $\dfrac{x}{\sqrt{n}}$.

[1] Et même orthonormée.

Existence

Admettons pour l'instant qu'une telle matrice existe, nous prouverons son existence par la suite.

Soit $P \in \mathcal{M}_n(\mathbf{R})$ une matrice orthogonale dont la première colonne est $\dfrac{1}{\sqrt{n}}\begin{pmatrix}1\\ \vdots \\ 1\end{pmatrix}$.

Alors soit $\mathcal{B}' = (e_1, \dots, e_n)$ la base de E telle que $P = P_{\mathcal{B}',\mathcal{B}} = P$.
Puisque P est orthogonale, $\mathcal{B}'$ est encore une base orthonormée. Et alors, puisque nous avons fixé la première colonne de P, il vient

$$\frac{x}{\sqrt{n}} = \sum_{k=1}^{n} \frac{1}{\sqrt{n}} e_k \Leftrightarrow x = \sum_{k=1}^{n} e_k.$$

Rappel

Une matrice est orthogonale si et seulement si c'est la matrice de passage entre deux bases orthonormées.

Il nous reste donc à prouver qu'il existe bien une matrice orthogonale P dont la première colonne est $\dfrac{1}{\sqrt{n}}\begin{pmatrix}1\\ \vdots \\ 1\end{pmatrix}$.
Dans $\mathbf{R}^n$ muni de sa structure euclidienne canonique, le vecteur $(1, \dots, 1)$ est de norme $\sqrt{n}$, et donc $\dfrac{1}{\sqrt{n}}(1, \dots, 1)$ est de norme 1.
Soit $\mathcal{B}'$ une base orthonormée de $\mathbf{R}^n$ dont le premier vecteur est $\dfrac{1}{\sqrt{n}}(1, \dots, 1)$. Notons $\mathcal{B}$ la base canonique de $\mathbf{R}^n$ qui, rappelons-le, est une base orthonormée de $\mathbf{R}^n$.
Soit alors $P = P_{\mathcal{B},\mathcal{B}'}$. Alors P est orthogonale car matrice de passage entre deux bases orthonormées, et la première colonne de P est formée des coordonnées de $\dfrac{1}{\sqrt{n}}(1, \dots, 1)$ dans la base canonique, c'est donc $\dfrac{1}{\sqrt{n}}\begin{pmatrix}1\\ \vdots \\ 1\end{pmatrix}$.

2. Supposons qu'il existe une base orthonormée $(e_1, \dots, e_n)$ telle que $x = \sum_{k=1}^{n} e_k$ et $y = \sum_{k=1}^{n} k e_k$.
Alors $\|x\| = \sqrt{n}, \|y\| = \sqrt{\sum_{k=1}^{n} k^2} = \sqrt{\dfrac{n(n+1)(2n+1)}{6}}$ et $\langle x, y\rangle = \sum_{k=1}^{n} k = \dfrac{n(n+1)}{2}$.

Montrons à présent que ces trois conditions sont suffisantes, et supposons que

$$\|x\| = \sqrt{n}, \|y\| = \sqrt{\frac{n(n+1)(2n+1)}{6}} \text{ et } \langle x, y\rangle = \frac{n(n+1)}{2}.$$

Nous pouvons directement supposer que $n \geqslant 2$, car pour $n = 1$, le résultat est évident.
Alors x et y ne sont pas colinéaires, car on n'a pas l'égalité dans Cauchy-Schwarz :

$$\langle x, y\rangle = \frac{n(n+1)}{2} < \sqrt{n}\sqrt{\frac{n(n+1)(2n+1)}{6}} = \|x\| \cdot \|y\|.$$

Rappel

Deux vecteurs sont colinéaires si et seulement si l'inégalité de Cauchy-Schwarz est en fait une égalité.

Ainsi, la famille (x, y) est libre. Orthonormalisons-la à l'aide du procédé de Gram-Schmidt.
Pour cela, considérons $f_1 = \dfrac{x}{\|x\|} = \dfrac{x}{\sqrt{n}}$.

Soit alors $f_2^* = y - \langle f_1, y\rangle f_1 = y - \langle x, y\rangle \dfrac{x}{n} = y - \frac{n+1}{2}x$.
On a alors, par bilinéarité du produit scalaire,

$$\|f_2^*\|^2 = \langle f_2^*, f_2^*\rangle = \|y\|^2 - 2\frac{n+1}{2}\langle x, y\rangle + \frac{(n+1)^2}{4}\|x\|^2 = \frac{n(n+1)(2n+1)}{6} + \frac{n(n+1)^2}{4} - \frac{n(n+1)^2}{2} = \frac{(n-1)n(n+1)}{12}.$$

On pose alors

$$f_2 = \frac{f_2^*}{\|f_2^*\|} = \sqrt{\frac{12}{(n-1)n(n+1)}}\left(y - \frac{n+1}{2}x\right).$$

Ainsi, (f_1, f_2) est une famille orthonormée, qui peut être complétée en une base orthonormée $(f_1, \dots, f_n)$ de E.

Soit alors $P \in \mathcal{M}_n(\mathbf{R})$ une matrice orthogonale dont la première colonne est $\dfrac{1}{\sqrt{n}}\begin{pmatrix}1\\ \vdots \\ 1\end{pmatrix}$ et

dont la seconde colonne est

$$\sqrt{\frac{12}{(n-1)n(n+1)}}\left(\begin{pmatrix}1\\2\\ \vdots \\ n\end{pmatrix} - \frac{n+1}{2}\begin{pmatrix}1\\1\\ \vdots \\ 1\end{pmatrix}\right).$$

Comme précédemment, nous admettons pour le moment qu'une telle matrice existe.
Soit alors $\mathcal{B}' = (f_1, f_2, \dots, f_n)$, et soit $\mathcal{B} = (e_1, \dots, e_n)$ la base de E telle que $P = P_{\mathcal{B},\mathcal{B}'}$.
Alors $\mathcal{B}$ est orthonormée car P est une matrice orthogonale.
De plus, on a alors

$$f_1 = \sum_{k=1}^{n} \frac{1}{\sqrt{n}} e_k \Leftrightarrow x = \sum_{k=1}^{n} e_k.$$

De même, on a

$$f_2 = \sum_{k=1}^{n} \sqrt{\frac{12}{(n-1)n(n+1)}}\left(k - \frac{n+1}{2}\right)e_k.$$

En multipliant par $\sqrt{\dfrac{(n-1)n(n+1)}{12}}$, on a donc

$$y - \frac{n+1}{2}x = \sum_{k=1}^{n}\left(k - \frac{n+1}{2}\right)e_k$$

et donc en remplaçant x par $\displaystyle\sum_{k=1}^{n} e_k$, il vient $y = \displaystyle\sum_{k=1}^{n} k e_k$.

Prouvons à présent qu'il existe bien une matrice P orthogonale dont les deux premières colonnes sont celles souhaitées.
Dans $\mathbf{R}^n$ muni de sa structure euclidienne canonique, considérons la famille $(1, \dots, 1), (1, 2, \dots, n)$.
Il s'agit d'une famille libre, et donc nous pouvons lui appliquer le procédé d'orthonormalisation de Gram-Schmidt.
Soit $g_1 = \dfrac{1}{\sqrt{n}}(1, \dots, 1)$, et soit

$$g_2^* = (1, 2, \dots, n) - \langle \frac{1}{\sqrt{n}}(1, \dots, 1), (1, 2, \dots, n)\rangle \frac{1}{\sqrt{n}}(1, \dots, 1) = (1, 2, \dots, n) - \frac{n+1}{2}(1, \dots, 1).$$

Alors par bilinéarité du produit scalaire,

$$\|g_2^*\|^2 = \langle g_2^*, g_2^*\rangle = \|(1, 2, \dots, n)\|^2 + \frac{(n+1)^2}{4}\|(1, \dots, 1)\|^2 - 2\frac{n+1}{2}\langle (1, 2, \dots, n), (1, \dots, 1)\rangle$$
$$= \frac{n(n+1)(2n+1)}{6} + \frac{n(n+1)^2}{4} - \frac{n(n+1)^2}{2} = \frac{(n-1)n(n+1)}{12}.$$

On pose alors $g_2 = \sqrt{\dfrac{12}{(n-1)n(n+1)}}g_2^* = \sqrt{\dfrac{12}{(n-1)n(n+1)}}\left((1,2,\dots,n) - \dfrac{n+1}{2}(1,\dots,1)\right)$.
Il est alors possible de compléter (g_1, g_2) en une base orthonormée $(g_1, \dots, g_n)$. Et alors si P est la matrice de passage de la base canonique à la base $(g_1, \dots, g_n)$, P est orthogonale (car matrice de passage entre deux bases orthonormées de $\mathbf{R}^n$), et les deux premières colonnes de P sont bien celles que l'on souhaitait.

□

EXERCICE 1.54 (ESCP 2016) [ESCP16.2.20] — Moyen

Familles μ-presque orthogonales

Soit $\mu \geqslant 1$, 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 $i \in [\![1, n]\!]$, $\|u_i\| = 1$
- pour tout $(x_1, \dots, x_n) \in \mathbf{R}^n$, $\dfrac{1}{\mu}\sum_{i=1}^n x_i^2 \leqslant \left\|\sum_{i=1}^n x_i u_i\right\|^2 \leqslant \mu \sum_{i=1}^n x_i^2$.

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.
   (a) Montrer qu'il existe un réel k tel que $\forall x \in E$, $\|f(x)\| \leqslant k\|x\|$.
   (b) Montrer que si f est un automorphisme de E, alors il existe un réel $\lambda \geqslant 1$ tel que $\forall x \in E$, $\dfrac{1}{\lambda}\|x\| \leqslant \|f(x)\| \leqslant \lambda\|x\|$.
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 $\mu \geqslant 1$ tel que $(u_1, \dots, u_n)$ soit μ-presque orthogonale.

SOLUTION DE L'EXERCICE 1.54

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

$$0 \leqslant \frac{1}{\mu}\sum_{i=1}^n x_i^2 \leqslant \underbrace{\left\|\sum_{i=1}^n x_i u_i\right\|^2}_{=\|0_E\|^2=0}.$$

On en déduit que $\sum_{i=1}^n x_i^2 = 0$, et donc pour tout $i \in [\![1, n]\!]$, $x_i = 0$.

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

$$\left\|\sum_{i=1}^n x_i u_i\right\|^2 = \sum_{i=1}^n x_i^2.$$

Détails
C'est le corollaire du théorème de Pythagore.

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 $(i, j) \in [\![1, n]\!]^2$, avec $i \neq j$. On a alors

$$\|u_i + u_j\|^2 = \|u_i\|^2 + \|u_j\|^2 + 2\langle u_i, u_j\rangle = 2\left(1 + \langle u_i, u_j\rangle\right).$$

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

$$\frac{1}{1}\sum_{k=1}^{n} x_k^2 \leqslant \|u_i + u_j\|^2 \leqslant \sum_{k=1}^{n} x_k^2 \Leftrightarrow 2 \leqslant 2(1 + \langle u_i, u_j\rangle) \leqslant 2.$$

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

$$\|f(x)\|^2 = \left\|\sum_{i=1}^{n} x_i f(e_i)\right\| \leqslant \left(\sum_{i=1}^{n} |x_i| \cdot \|f(e_i)\|\right)^2.$$

Appliquons alors l'inégalité de Cauchy-Schwarz dans $\mathbf{R}^n$:

$$\left(\sum_{i=1}^{n} |x_i| \cdot \|f(e_i)\|^2\right)^2 \leqslant \left(\sum_{i=1}^{n} |x_i|^2\right)\left(\sum_{i=1}^{n} \|f(e_i)\|^2\right).$$

Et donc $\|f(x)\|^2 \leqslant \|x\| \times \left(\sum_{i=1}^{n} \|f(e_i)\|^2\right)$.

Ainsi, en posant $k = \sqrt{\sum_{i=1}^{n} \|f(e_i)\|^2}$, qui est bien une constante indépendante de x, on a $\|f(x)\| \leqslant k\|x\|$.

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 $x \in E$, $\|f^{-1}(x)\| \leqslant k\|x\|$.
En particulier, si l'on applique ceci à $f(x)$ au lieu de x, il vient

$$\|f^{-1}(f(x))\| \leqslant k\|f(x)\| \Leftrightarrow \|x\| \leqslant k\|f(x)\|.$$

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 $\frac{1}{k}\|x\| \leqslant \|f(x)\|$.
D'autre part, il existe une constante M telle que $\forall x \in E$, $\|f(x)\| \leqslant M\|x\|$.
Notons alors $\lambda = \max(M, k)$, de sorte que $M \leqslant \lambda$ et $\frac{1}{\lambda} \leqslant \frac{1}{k}$.
On a alors bien, pour tout $x \in E$,

$$\frac{1}{\lambda}\|x\| \leqslant \|f(x)\| \leqslant \lambda\|x\|.$$

Remarque

Puisqu'il vient $\frac{1}{\lambda} \leqslant \lambda$, nécessairement $\lambda \geqslant 1$.

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 $\lambda \geqslant 1$ telle que

$$\forall x \in F,\ \frac{1}{\lambda}\|x\| \leqslant \|f(x)\| \leqslant \lambda\|x\|.$$

En particulier, pour $(x_1,\dots,x_n)\in\mathbf{R}^n$, soit $x=\sum_{i=1}^n x_ie_i$.

Alors $\|x\|^2=\sum_{i=1}^n x_i^2$ et $\|f(x)\|^2=\left\|\sum_{i=1}^n x_iu_i\right\|^2$.

Et donc

$$\frac{1}{\lambda^2}\|x\|^2\leqslant\|f(x)\|^2\leqslant\lambda^2\|x\|^2\Leftrightarrow\frac{1}{\lambda^2}\sum_{i=1}^n x_i^2\leqslant\left\|\sum_{i=1}^n x_iu_i\right\|^2\leqslant\lambda^2\sum_{i=1}^n x_i^2.$$

□

EXERCICE 1.55 (QSP ESCP 2010) [ESCP10.QSP03] — Moyen

Familles obtusangles

Soit E un espace euclidien de dimension n et soient $e_1,\dots,e_{n+1}$ des vecteurs tels que pour tout $(i,j)\in[\![1,n+1]\!]^2$, $i\neq j$, $\langle e_i,e_j\rangle<0$.

1. Montrer, en utilisant la norme de u que si $u=\sum_{i=1}^n\lambda_ie_i=0_E$, alors $\sum_{i=1}^n|\lambda_i|e_i=0_E$.
2. Montrer que n quelconques de ces vecteurs forment une base de E.

SOLUTION DE L'EXERCICE 1.55

1. Si $u=0$, alors $\|u\|^2=0$. Or, par bilinéarité du produit scalaire,

$$\|u\|^2=\sum_{i=1}^n\sum_{j=1}^n\lambda_i\lambda_j\langle e_i,e_j\rangle=\sum_{i=1}^n\lambda_i^2\|e_i\|^2+\sum_{i=1}^n\sum_{\substack{j=1\\j\neq i}}^n\lambda_i\lambda_j\langle e_i,e_j\rangle.$$

Soit encore $\sum_{i=1}^n\sum_{\substack{j=1\\j\neq i}}^n\lambda_i\lambda_j\langle e_i,e_j\rangle=-\sum_{i=1}^n\lambda_i^2\|e_i\|^2\leqslant 0$.

On a donc[1]

$$\left|\sum_{i=1}^n\sum_{\substack{j=1\\i\neq j}}^n\lambda_i\lambda_j\langle e_i,e_j\rangle\right|=-\sum_{i=1}^n\sum_{\substack{j=1\\j\neq i}}^n\lambda_i\lambda_j\langle e_i,e_j\rangle.$$

[1] La valeur absolue d'un nombre négatif est son opposé.

Et donc, par l'inégalité triangulaire,

$$-\sum_{i=1}^n\sum_{\substack{j=1\\j\neq i}}^n\lambda_i\lambda_j\langle e_i,e_j\rangle\leqslant\sum_{i=1}^n\sum_{\substack{j=1\\j\neq i}}^n|\lambda_i|\cdot|\lambda_j|\cdot|\langle e_i,e_j\rangle|=-\sum_{i=1}^n\sum_{\substack{j=1\\j\neq i}}^n|\lambda_i|\cdot|\lambda_j|\langle e_i,e_j\rangle.$$

Détails

On a ici utilisé le fait que $\langle e_i,e_j\rangle<0$ et donc $|\langle e_i,e_j\rangle|=-\langle e_i,e_j\rangle$.

Et alors il vient

$$\begin{aligned}\left\|\sum_{i=1}^n|\lambda_i|e_i\right\|^2&=\sum_{i=1}^n\lambda_i^2\|e_i\|^2+\sum_{i=1}^n\sum_{\substack{j=1\\j\neq i}}^n|\lambda_i|\cdot|\lambda_j|\langle e_i,e_j\rangle\\&\leqslant\sum_{i=1}^n\lambda_i^2\|e_i\|^2+\sum_{i=1}^n\sum_{\substack{j=1\\j\neq i}}^n\lambda_i\lambda_j\langle e_i,e_j\rangle\\&\leqslant\sum_{i=1}^n\lambda_i^2\|e_i\|^2-\sum_{i=1}^n\lambda_i^2\|e_i\|^2\\&\leqslant 0.\end{aligned}$$

On en déduit[2] que $\left\|\sum_{i=1}^n|\lambda_i|e_i\right\|^2=0$ et donc que $\sum_{i=1}^n|\lambda_i|e_i=0_E$.

[2] Un carré est positif ou nul.

2. Montrons que $(e_1, \dots, e_n)$ est une base, la preuve serait la même pour n'importe quelle autre famille de n vecteurs de $\{e_1, \dots, e_n\}$.
Puisqu'il s'agit déjà d'une famille de cardinal $n = \dim E$, il suffit de prouver qu'elle est libre.

Soient donc $\lambda_1, \dots, \lambda_n$ des réels tels que $\sum_{i=1}^{n} \lambda_i e_i = 0_E$.

Alors, par la question précédente, nous savons que $\sum_{i=1}^{n} |\lambda_i| e_i = 0_E$.

En particulier, il vient

$$\langle 0_E, e_{n+1}\rangle = \left\langle \sum_{i=1}^{n} |\lambda_i| e_i, e_{n+1} \right\rangle = \sum_{i=1}^{n} |\lambda_i| \underbrace{\langle e_i, e_{n+1}\rangle}_{<0} \leqslant 0.$$

Et si l'un des λ_i est non nul, on a même $\langle 0_E, e_{n+1}\rangle < 0$, ce qui n'est pas possible puisque $\langle 0_E, e_{n+1}\rangle = 0$.
On en déduit donc que $\lambda_1 = \cdots = \lambda_n = 0$, et donc que $(e_1, \dots, e_n)$ est libre, donc est une base de E.

□

1.7.1 Polynômes orthogonaux

Exercice 1.56 (ESCP 2016) [ESCP16.2.01] **Moyen**

Polynômes de Tchebychev

1. Montrer que pour tous réels a et b, on a :
$$\cos(a)\cos(b) = \frac{1}{2}\left(\cos(a+b) + \cos(a-b)\right).$$

Soit la suite de polynômes $(T_n)_{n\in\mathbf{N}^*}$ définie par :
$$T_0 = 1,\ T_1 = X \text{ et } \forall n \in \mathbf{N},\ T_{n+2} = 2XT_{n+1} - T_n.$$

2. Montrer que pour tout $n \in \mathbf{N}^*$, T_n est un polynôme de degré n, de coefficient dominant que l'on explicitera.
Montrer que
$$\forall \theta \in \mathbf{R},\ T_n(\cos\theta) = \cos(n\theta).$$
3. Pour $(p, q) \in \mathbf{N}^2$, on pose $I_{p,q} = \int_0^{\pi} \cos(p\theta)\cos(q\theta)\, d\theta$.
Calculer $I_{p,q}$.
4. Pour tout couple (P, Q) d'éléments de $\mathbf{R}[X]$, montrer que l'intégrale $\int_{-1}^{1} \frac{P(t)Q(t)}{\sqrt{1-t^2}}\, dt$ est convergente. On la note $\langle P, Q\rangle$.
Montrer que $(P, Q) \mapsto \langle P, Q\rangle$ est un produit scalaire sur $\mathbf{R}[X]$. On note $\|\cdot\|$ la norme associée.
5. Montrer que pour tout $n \in \mathbf{N}$, la famille $(T_0, T_1, \dots, T_n)$ est une base orthogonale de $\mathbf{R}_n[X]$.
Cette base est-elle orthonormale ?
6. Pour tout $n \in \mathbf{N}^*$, calculer $\langle X^n, T_n\rangle$.
7. En déduire, pour tout entier $n \geqslant 1$, la valeur de :
$$d = \inf_{(a_0, a_1, \dots, a_{n-1}) \in \mathbf{R}^n} \sqrt{\int_{-1}^{1} \frac{\left(t^n - (a_{n-1}t^{n-1} + \cdots + a_0)\right)^2}{\sqrt{1-t^2}}\, dt}.$$

Solution de l'exercice 1.56

1. On a

$$\begin{aligned} cos(a+b) + \cos(a-b) &= \cos(a)\cos(b) - \sin(a)\sin(b) + \cos(a)\cos(b) + \sin(a)\sin(b) \\ &= 2\cos(a)\cos(b) \end{aligned}$$

d'où la formule voulue.

Remarque

Comme la plupart des formules de trigonométrie, ce résultat peut également se prouver à l'aide de la formule d'Euler :
$$\cos(a) = \frac{e^{ia} + e^{-ia}}{2}.$$
M. Vienney

2. On a $T_2 = 2X^2 - 1$, $T_3 = 4X^3 - 2X - X = 4X^3 - 3X$, et donc il est raisonnable de supposer que T_n a pour coefficient dominant 2^{n-1}.
Montrons donc, par récurrence double[1] sur $n \in \mathbf{N}^*$ que T_n est un polynôme de degré n de coefficient dominant égal à 2^{n-1}.
Pour $n = 1$ et $n = 2$, c'est vrai.
Supposons donc T_n de degré n et de coefficient en X^n égal à 2^{n-1} et T_{n+1} de degré $n+1$ et de coefficient dominant égal à 2^n.
Alors $T_{n+1} = 2^n X^{n+1} + Q_{n+1}$, avec $\deg Q_{n+1} \leqslant n$, de sorte que

$$T_{n+2} = 2XT_{n+1} - T_n = 2^{n+1}X^{n+2} + \underbrace{2XQ_{n+1} - T_n}_{\in \mathbf{R}_{n+1}[X]}$$

est donc bien de degré $n+2$, et de coefficient dominant égal à 2^{n+1}.
Par le principe de récurrence, pour tout $n \in \mathbf{N}^*$, T_n st de degré n et de coefficient dominant 2^{n-1}.

[1] Puisque T_{n+2} dépend de T_n et de T_{n+1}.

Remarque

Si nous n'utilisons pas le coefficient dominant de T_n, nous utilisons tout de même le fait que son degré soit n, et donc la récurrence double était inévitable (au moins pour le degré).

Prouvons également par récurrence[2] sur n que pour tout $n \in \mathbf{N}$ que pour tout $\theta \in \mathbf{R}$, $T_n(\cos\theta) = \cos(n\theta)$.
Pour $n = 0$, c'est évident car pour tout θ, $\cos(0 \times \theta) = 1 = T_0(\cos\theta)$.
Et pour $n = 1$, pour tout θ, $T_n(\cos\theta) = \cos\theta$.
Supposons donc que pour tout $\theta \in \mathbf{R}$, $T_n(\cos\theta) = \cos(n\theta)$ et $T_{n+1}(\cos\theta) = \cos((n+1)\theta)$.
Alors en utilisant la formule prouvée à la question 1,

$$T_{n+2}(\cos\theta) = 2\cos(\theta)\cos((n+1)\theta) - \cos(n\theta) = \cos((n+2\theta)) + \cos(n\theta) - \cos(n\theta) = \cos((n+2)\theta).$$

Par le principe de récurrence, pour tout $n \in \mathbf{N}$ et tout $\theta \in \mathbf{R}$, $T_n(\cos\theta) = \cos(n\theta)$.

[2] Double toujours.

3. D'après la formule de la question 1, on a

$$I_{p,q} = \frac{1}{2}\int_0^\pi \cos((p+q)\theta)\,d\theta + \frac{1}{2}\int_0^\pi \cos((p-q)\theta)\,d\theta.$$

Or, pour tout $n \in \mathbf{Z}^*$, on a

$$\int_0^\pi \cos(n\theta)\,d\theta = \left[\frac{\sin(n\theta)}{n}\right]_0^\pi = \frac{\sin(n\pi)}{n} - \frac{\sin(0)}{n} = 0 - 0 = 0.$$

Toutefois, ce résultat n'est pas valable pour $n = 0$, mais on a alors

$$\int_0^\pi \cos(0 \times \theta)\,d\theta = \int_0^\pi 1\,d\theta = \pi.$$

Et donc pour $(p,q) \in \mathbf{N}^2$, il vient

$$\int_0^\pi \cos((p+q)\theta)\,d\theta = \begin{cases} \pi & \text{si } p = q = 0 \\ 0 & \text{sinon} \end{cases}$$

et de même

$$\int_0^\pi \cos((p-q)\theta)\,d\theta = \begin{cases} \pi & \text{si } p = q \\ 0 & \text{sinon} \end{cases}$$

On en déduit que

$$I_{p,q} = \begin{cases} \pi & \text{si } p = q = 0 \\ \frac{\pi}{2} & \text{si } p = q \neq 0 \\ 0 & \text{si } p \neq q \end{cases}$$

4. La fonction $t \mapsto \dfrac{P(t)Q(t)}{\sqrt{1-t^2}}$ est continue sur $]-1,1[$, donc les problèmes de convergence sont au voisinage de ± 1.
Notons que PQ est continue sur le segment $[-1,1]$, et donc elle y est bornée : il existe $M > 0$ tel que $\forall t \in [-1,1]$, $|P(t)Q(t)| \leqslant M$.
Par conséquent, $\left|\dfrac{P(t)Q(t)}{\sqrt{1-t^2}}\right| \leqslant \dfrac{M}{\sqrt{1-t^2}}$, et il nous suffit donc de prouver que $\displaystyle\int_{-1}^1 \frac{M}{\sqrt{1-t^2}}\,dt$

converge.

Au voisinage de 1, on a $\sqrt{1-t^2} = \sqrt{(1-t)(1+t)} \underset{t\to 1^-}{\sim} \sqrt{2}\sqrt{1-t}$. Mais $\displaystyle\int_0^1 \frac{1}{\sqrt{1-t}}\,dt$ est une intégrale de Riemann convergente.

Par conséquent, $\displaystyle\int_0^1 \frac{P(t)Q(t)}{\sqrt{1-t^2}}\,dt$ converge absolument et donc converge.

On traite de même le cas de $\displaystyle\int_{-1}^0 \frac{P(t)Q(t)}{\sqrt{1-t^2}}\,dt$ en remarquant qu'au voisinage de -1, $\sqrt{1-t^2} \underset{t\to -1^+}{\sim} \sqrt{2}\sqrt{1+t}$ et que $\displaystyle\int_{-1}^0 \frac{dt}{\sqrt{1+t}}$ est une intégrale de Riemann convergente.

Et donc $\langle P, Q\rangle$ est bien défini.

La bilinéarité et la symétrie de $\langle\cdot,\cdot\rangle$ sont immédiates.

Si $P \in \mathbf{R}[X]$, alors $\langle P, P\rangle = \displaystyle\int_{-1}^1 \frac{P(t)^2}{\sqrt{1-t^2}}\,dt \geqslant 0$ car intégrale d'une fonction positive.

Enfin, la fonction $t \mapsto \dfrac{P(t)^2}{\sqrt{1-t^2}}$ est continue et positive sur $]-1,1[$.

Et donc on a $\langle P, P\rangle = 0$ si et seulement si pour tout $t \in]-1,1[$, $\dfrac{P(t)^2}{\sqrt{1-t^2}} = 0 \Leftrightarrow P(t) = 0$.

Mais dans ce cas, P possède alors une infinité de racines, et donc est le polynôme nul.

Ainsi, $\langle P, P\rangle = 0 \Rightarrow P = 0$, ce qui achève de prouver que $\langle\cdot,\cdot\rangle$ est un produit scalaire sur $\mathbf{R}[X]$.

5. Pour $(i,j) \in [\![0,n]\!]^2$, on a $\langle T_i, T_j\rangle = \displaystyle\int_{-1}^1 \frac{T_i(t)T_j(t)}{\sqrt{1-t^2}}\,dt$.

Procédons alors au changement de variable $t = \cos\theta$, qui est légitime car la fonction cos est de classe $\mathscr{C}^1$ sur $]0,\pi[$, strictement décroissante, avec $\cos(0) = 1$ et $\cos(\pi) = -1$.

On a alors $dt = \sin\theta\,d\theta = \sqrt{\sin^2\theta}\,d\theta = \sqrt{1-t^2}\,d\theta$, de sorte que $d\theta = \dfrac{dt}{\sqrt{1-t^2}}$.

Et donc par le théorème de changement de variable,

$$\langle T_i, T_j\rangle = \int_0^\pi T_i(\cos\theta)T_j(\cos\theta)\,d\theta = \int_0^\pi \cos(i\theta)\cos(j\theta)\,d\theta = I_{i,j}.$$

Et en particulier, si $i \neq j$, $\langle T_i, T_j\rangle = 0$, de sorte que $(T_0, T_1, \dots, T_n)$ est une famille orthogonale de $\mathbf{R}_n[X]$.

Puisque de plus, $\langle T_0, T_0\rangle = I_{0,0} = \pi \neq 1$, la famille n'est pas orthonormée.

Signe

On a bien $\sin\theta = \sqrt{\sin^2\theta}$ car sur $]0,\pi[$, $\sin\theta \geqslant 0$.
Ce ne serait pas forcément vrai sur un autre intervalle.

6. Puisque $T_n = 2^{n-1}X^n + Q_n$, avec $Q_n \in \mathbf{R}_{n-1}[X]$, on a donc $X^n = \dfrac{1}{2^{n-1}}T_n + \dfrac{Q_n}{2^{n-1}}$.

Mais $(T_0, T_1, \dots, T_{n-1})$ est une famille de polynômes de $\mathbf{R}_{n-1}[X]$, de degrés deux à deux distincts, et donc libre, de cardinal $n = \dim \mathbf{R}_{n-1}[X]$: c'est donc une base de $\mathbf{R}_{n-1}[X]$.

Et puisque la famille $(T_0, T_1, \dots, T_n)$ est orthogonale, $T_n \in (\mathbf{R}_{n-1}[X])^\perp$.

On en déduit que

$$\langle X^n, T_n\rangle = \frac{1}{2^{n-1}}\langle T_n, T_n\rangle + \frac{1}{2^{n-1}}\underbrace{\langle Q_n, T_n\rangle}_{=0} = \frac{1}{2^{n-1}}\|T_n\|^2.$$

Mais par ce qui a été dit précédemment, $\|T_n\|^2 = I_{n,n} = \dfrac{\pi}{2}$.

Et donc $\langle X^n, T_n\rangle = \dfrac{\pi}{2^n}$.

7. Lorsque $(a_0, a_1, \dots, a_{n-1})$ parcourt $\mathbf{R}^n$, $a_0 + a_1X + \cdots + a_{n-1}X^{n-1}$ parcourt $\mathbf{R}_{n-1}[X]$, et donc d n'est autre que la distance de X^n à $\mathbf{R}_{n-1}[X]$.

Un théorème du cours garantit alors qu'elle est atteinte lorsque $a_0 + a_1X + \cdots + a_{n-1}X^{n-1}$ est le projeté orthogonal de X^n sur $\mathbf{R}_{n-1}[X]$.

Puisque $(T_0, T_1, \dots, T_n)$ est une base orthogonale de $\mathbf{R}_n[X]$, $\left(\dfrac{T_0}{\|T_0\|}, \dots, \dfrac{T_n}{\|T_n\|}\right)$ en est une base orthonormée.

Et donc

$$X^n = \sum_{k=0}^{n} \left\langle X^n, \frac{T_k}{\|T_k\|} \right\rangle \frac{T_k}{\|T_k\|} = \underbrace{\sum_{k=0}^{n-1} \langle X^n, T_k\rangle \frac{T_k}{\|T_k\|^2}}_{\in \mathbf{R}_{n-1}[X]} + \underbrace{\langle X^n, T_n\rangle \frac{T_n}{\|T_n\|^2}}_{\in \mathbf{R}_{n-1}[X]^\perp}.$$

Sans surprise[3] on obtient alors $p_{\mathbf{R}_{n-1}[X]}(X^n) = \sum_{k=0}^{n-1} \langle X^n, T_k\rangle \frac{T_k}{\|T_k\|^2}$.

Mais alors

$$d = \|X^n - p_{\mathbf{R}_{n-1}[X]}(X^n)\| = |\langle X^n, T_n\rangle| \left\| \frac{T_n}{\|T_n\|^2} \right\| = \frac{\pi}{2^n} \frac{1}{\|T_n\|} = \frac{\pi}{2^n} \sqrt{\frac{2}{\pi}} = \frac{\sqrt{2\pi}}{2^n}.$$

Commentaire *: notons qu'il existe un moyen assez agréable, même si peut-être un peu moins naturel de dire tout ça : $X^n - p_{\mathbf{R}_{n-1}[X]}(X^n)$ n'est autre que le projeté orthogonal de X^n sur $(\mathbf{R}_{n-1}[X])^\perp$.*
Mais nous savons que $(\mathbf{R}_{n-1}[X])^\perp$ est de dimension 1, et qu'il contient T_n.
Une base orthonormée en est donc formée par le seul vecteur $\frac{T_n}{\|T_n\|}$, de sorte que le projeté orthogonal de X^n sur $(\mathbf{R}_{n-1}[X])^\perp$ est $\langle X^n, T_n\rangle \frac{T_n}{\|T_n\|^2}$.
Et donc comme annoncé ci-dessus, $d = |\langle X^n, T_n\rangle| \left\| \frac{T_n}{\|T_n\|^2} \right\|$.

□

[3] Car la famille $\left(\frac{T_0}{\|T_0\|}, \dots, \frac{T_{n-1}}{\|T_{n-1}\|}\right)$ est une base orthonormée de $\mathbf{R}_{n-1}[X]$.

$\|T_n\|$
La norme de T_n a été calculée à la question 5 et vaut $\sqrt{I_{n,n}}$.

EXERCICE 1.57 (ESCP 2014) [ESCP14.2.09] — Facile

Étude d'une famille de polynômes orthogonaux et d'un projecteur de $\mathbf{R}_n[X]$

On note $\mathbf{R}[X]$ l'ensemble des polynômes à coefficients réels et, pour tout $k \in \mathbf{N}$, on note $\mathbf{R}_k[X]$ l'ensemble de ces polynômes de degré au plus k. On munit $\mathbf{R}[X]$ du produit scalaire défini par :

$$\langle P, Q\rangle = \int_0^1 P(t)Q(t)\,dt.$$

L'orthogonal d'un sous-espace vectoriel F de $\mathbf{R}[X]$ pour ce produit scalaire est noté $F^\perp$.
Pour tout $j \in \mathbf{N}$, soit le polynôme $L_j = (X^j(1-X)^j)^{(j)}$ (polynôme dérivé d'ordre j du polynôme $X^j(1-X)^j$).

1. (a) Montrer que si $P(0) = P(1) = 0$, alors $\langle P', Q\rangle = -\langle P, Q'\rangle$.
   (b) Pour tout $k \in \mathbf{N}^*$, montrer que si 0 et 1 sont racines d'ordre k de P, alors $\langle P^{(k)}, Q\rangle = (-1)^k \langle P, Q^{(k)}\rangle$.
2. (a) Déterminer, pour tout $j \geqslant 1$, le degré de L_j et montrer que L_j est dans $\mathbf{R}_{j-1}[X]^\perp$.
   (b) Montrer que pour tout $n \in \mathbf{N}$, $(L_j)_{0\leqslant j\leqslant n}$ est une base orthogonale de $\mathbf{R}_n[X]$.
3. Soit $n \in \mathbf{N}^*$ et $E = \mathbf{R}_n[X]$. Soit un polynôme $A \in E$ non nul, et soit $f_A : E \to E$ l'application qui à tout polynôme $P \in E$ associe le reste de sa division euclidienne par A.
   (a) Montrer que f_A est un projecteur de E. Déterminer son noyau et son image.
   (b) Montrer que si A n'est pas de degré n et possède au moins une racine réelle α, alors f_A n'est pas un projecteur orthogonal.

SOLUTION DE L'EXERCICE 1.57

1.a. Les fonctions P et Q sont $\mathscr{C}^1$ sur $[0,1]$, donc une intégration par parties nous donne

$$\langle P', Q\rangle = \int_0^1 P'(t)Q(t)\,dt = \underbrace{[P(t)Q(t)]_0^1}_{=0} - \int_0^1 P(t)Q'(t)\,dt = -\langle P, Q'\rangle.$$

1.b. Procédons par récurrence sur $k \in \mathbf{N}^*$, où notre hypothèse de récurrence est «pour tout polynôme P dont 0 et 1 sont racines d'ordre k et pour tout polynôme Q, $\langle P^{(k)}, Q\rangle = (-1)^k\langle P, Q^{(k)}\rangle$».
Pour $k = 1$, c'est le résultat de la question 1.a.

Supposons donc l'hypothèse vraie au rang k, et soit P, Q deux polynômes tels que 0 et 1 soient racines d'ordre $k+1$ de P. Alors 0 et 1 sont des racines d'ordre k de P, de sorte que

$$\langle P, Q^{(k+1)}\rangle = (-1)^k \langle P^{(k)}, Q'\rangle.$$

Racines

Rappelons qu'une racine d'ordre $k+1$ de P est une racine de tous les polynômes $P, P', \cdots, P^{(k)}$.

Mais 0 et 1 sont racines de $P^{(k)}$ de sorte que par la question 1.a,

$$\langle P^{(k)}, Q'\rangle = -\langle P^{(k+1)}, Q\rangle$$

et donc $\langle P^{(k+1)}, Q\rangle = (-1)^{k+1}\langle P, Q^{(k+1)}\rangle$.
Ainsi, la propriété est vraie au rang $k+1$, et par le principe de récurrence, elle est vraie pour tout $k \in \mathbf{N}^*$.

2.a. On a $\deg(X^j(1-X)^j) = \deg(X^j) + \deg((1-X)^j) = j + j = 2j$, et donc en dérivant j fois, il vient $\deg L_j = 2j - j = j$.
Soit $P \in \mathbf{R}_{j-1}[X]$. Alors 0 et 1 sont racines d'ordre j de $X^j(1-X)^j$, et donc d'après la question 1.b,

$$\langle L_j, P\rangle = (-1)^j \left\langle X^j(1-X)^j, \underbrace{P^{(j)}}_{=0}\right\rangle = 0.$$

Donc L_j est orthogonal à tout polynôme de $\mathbf{R}_{j-1}[X]$ et donc dans $(\mathbf{R}_{j-1}[X])^\perp$.

2.b. Soient $(i,j) \in [\![1,n]\!]^2$, avec $i \neq j$.
Supposons que $i < j$. Alors $\langle L_i, L_j\rangle = 0$ car $L_i \in \mathbf{R}_{j-1}[X]$.
De même, si $i > j$, on a $\langle L_i, L_j\rangle = 0$, donc la famille $(L_i)_{0\leqslant i\leqslant n}$ est orthogonale. Elle est par conséquent libre[1], et de cardinal $n+1 = \dim \mathbf{R}_n[X]$: c'est donc une base orthogonale de $\mathbf{R}_n[X]$.

[1] Car orthogonale et formée de polynômes non nuls.

3.a. Soient $P_1, P_2 \in \mathbf{R}[X]$ et soit $\lambda \in \mathbf{R}$. Notons $P_1 = AQ_1 + R_1$ et $P_2 = AQ_2 + R_2$ les divisions euclidiennes de P_1 et P_2 par A, avec $\deg R_1 < \deg A$ et $\deg R_2 < \deg A$.
Alors on a $R_1 = f_A(P_1)$ et $R_2 = f_A(P_2)$.
De plus, $\lambda P_1 + P_2 = A(\lambda Q_1 + Q_2) + \lambda R_1 + R_2$, avec $\deg(\lambda R_1 + R_2) < \deg A$.
Par unicité de la division euclidienne, ceci est donc la division euclidienne de $\lambda P_1 + P_2$ par A, de sorte que $f_A(\lambda P_1 + P_2) = \lambda R_1 + R_2 = \lambda f_A(P_1) + f(P_2)$. Ainsi, f_A est linéaire.
De plus, si $\deg R < \deg A$, alors $R = A\times 0 + R$ de sorte que le reste de la division euclidienne de R par A est égal à R.
Soit donc $P = AQ + R$ la division euclidienne de $P \in \mathbf{R}_n[X]$. Alors

$$f_A^2(P) = f_A(R) = R = f_A(P).$$

Et donc $f_A^2 = f_A$: f_A est un projecteur.
Un élément P de $\mathbf{R}_n[X]$ est dans $\operatorname{Ker} f_A$ si et seulement si il s'écrit $P = AQ$, avec $Q \in \mathbf{R}[X]$.
Donc $\operatorname{Ker} f_A = \{P \in \mathbf{R}_n[X] : \exists Q \in \mathbf{R}[X] \text{ tel que } P = AQ\}$.
Si $R \in \operatorname{Im} f_A$, alors il existe $P \in E$ tel que R soit le reste de la division euclidienne de P par A.
En particulier, $\deg R < \deg(A)$ et donc $\operatorname{Im} f_A \subset \mathbf{R}_{\deg A-1}[X]$.
Inversement, si $R \in \mathbf{R}_{\deg A-1}[X]$, alors $R = A\times 0 + R$ et donc $R = f_A(R) \in \operatorname{Im} f_A$.
Ainsi, $\operatorname{Im} f_A = \mathbf{R}_{\deg A-1}[X]$.

Remarque

La même preuve montrerait que l'application qui à P associe son quotient dans la division euclidienne par A est également linéaire.

Projecteur

Notons qu'il n'était pas clair d'après l'énoncé que f_A soit linéaire, donc prouver $f_A^2 = f_A$ ne suffit pas, il faut également prouver la linéarité.

⚠ Subtilité

$\operatorname{Ker} f_A$ n'est pas l'ensemble des multiples de A, mais uniquement l'ensemble des multiples de A qui sont de degré inférieur ou égal à n. C'est donc l'ensemble des QA, avec $\deg Q \leqslant n - \deg A$.

3.b. Si A n'est pas de degré n et possède une racine α, alors $A = (X-\alpha)Q$, avec $\deg Q < n-1$.
Puisque nous savons déjà que f_A est un projecteur, c'est un projecteur orthogonal si et seulement si c'est un endomorphisme symétrique.
Or, on a $\langle f_A((X-\alpha)A), Q\rangle = \langle 0_E, Q\rangle = 0$.
D'autre part, puisque $\deg < \deg A$, $f_A(Q) = Q$ et donc

$$\langle (X-\alpha)A, f_A(Q)\rangle = \langle (X-\alpha)A, Q\rangle = \int_0^1 (t-\alpha)A(t)Q(t)\,dt = \int_0^1 A(t)^2\,dt.$$

Mais la fonction $t \mapsto A(t)^2$ est continue et positive sur $[0,1]$ est n'est pas nulle car il ne s'agit pas de la fonction nulle.
On en déduit que $\langle (X-\alpha)A, f_A(Q)\rangle > 0$ et donc $\langle (X-\alpha)A, f_A(Q)\rangle \neq \langle f_A((X-\alpha)A), Q\rangle$, de sorte que f_A n'est pas un projecteur orthogonal.

Remarque

Puisque $\deg(Q) < n-1$, on a $(X-\alpha)A \in E$. C'est ici qu'on utilise l'hypothèse $\deg A < n$.

□

1.8 Endomorphismes et matrices symétriques

EXERCICE 1.58 (QSP HEC 2005) [HEC05.QSP01]

Matrices symétriques orthogonales

Soit $M \in \mathcal{M}_3(\mathbf{R})$ une matrice à la fois symétrique et orthogonale. Que peut-on dire de M si sa première ligne est $(1 \quad 0 \quad 0)$?

SOLUTION DE L'EXERCICE 1.58 M étant symétrique, elle est donc de la forme $\begin{pmatrix} 1 & 0 & 0 \\ 0 & a & b \\ 0 & b & c \end{pmatrix}$.
Mais étant orthogonale, on a donc

$$M\,{}^tM = M^2 = I_3 \Leftrightarrow \begin{pmatrix} 1 & 0 & 0 \\ 0 & a^2+b^2 & b(a+c) \\ 0 & b(a+c) & b^2+c^2 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Et donc en particulier, $\begin{cases} a^2+b^2=1 \\ b(a+c)=0 \\ b^2+c^2=1 \end{cases}$.
Si $b=0$, alors $a=\pm 1$ et $c=\pm 1$.
Si $b \neq 0$, alors $a=-c$. De plus, $b^2 \in [0,1]$, de sorte que $b \in [-1,1]$.
Et donc M est de la forme $\begin{pmatrix} 1 & 0 & 0 \\ 0 & \pm\sqrt{1-b^2} & b \\ 0 & b & \pm\sqrt{1-b^2} \end{pmatrix}$.
Inversement, si M est de la forme $\begin{pmatrix} 1 & 0 & 0 \\ 0 & \pm\sqrt{1-b^2} & b \\ 0 & b & \pm\sqrt{1-b^2} \end{pmatrix}$, avec $b \in [-1,1]$, alors elle est bien symétrique, et orthogonale.
Et donc l'ensemble des matrices de $\mathcal{M}_3(\mathbf{R})$, symétriques, orthogonales, et dont la première ligne vaut $(1 \quad 0 \quad 0)$ est

$$\left\{ \begin{pmatrix} 1 & 0 & 0 \\ 0 & \pm\sqrt{1-b^2} & b \\ 0 & b & \pm\sqrt{1-b^2} \end{pmatrix},\ b \in [-1,1] \right\}.$$

□

Remarque
Notons que les matrices obtenues précédemment pour $b=0$ sont bien de cette forme.

Un résultat classique prouve que si $A \in \mathcal{M}_n(\mathbf{R})$, alors ${}^tAA$ et A ont même rang. Mieux : elles ont mêmes noyaux.
Afin de le prouver, il faut penser à faire apparaître un produit scalaire dans $\mathcal{M}_{n,1}(\mathbf{R})$: ${}^tX\,{}^tAAX = \|AX\|^2$.

EXERCICE 1.59 (QSP ESCP 2014) [ESCP14.QSP04]

Matrices telles que ${}^tAA = {}^tBB$.

Soient A et B deux matrices de $\mathcal{M}_n(\mathbf{R})$ telles que ${}^tA \times A = {}^tB \times B$.

1. Montrer que A et B ont même noyau.
2. On suppose B inversible. Montrer qu'il existe une matrice orthogonale U telle que $A = U \times B$.

SOLUTION DE L'EXERCICE 1.59

1. Commençons par montrer que A et ${}^tA \times A$ ont le même rang.
Pour cela, rappelons-nous que $\text{rg}(A) = n - \dim E_0(A)$.
Soit donc $X \in E_0(A)$, de sorte que $AX = 0$.
Alors évidemment, ${}^tAAX = 0$, et donc $X \in E_0({}^tAA)$, de sorte que $E_0(A) = E_0({}^tAA)$.
Inversement, si $X \in E_0({}^tAA)$, alors ${}^tAAX = 0$. En multipliant à gauche par ${}^tX$, il vient alors

$${}^tXAAX = 0 \Leftrightarrow {}^t(AX)AX = 0 \Leftrightarrow \|AX\|^2 = 0.$$

Remarque
Nous parlons ici de la norme associée au produit scalaire canonique de $\mathcal{M}_{n,1}(\mathbf{R})$ défini par $\langle X, Y\rangle = {}^tXY$.

Mais ceci implique donc[1] que $AX = 0$, et donc $X \in E_0(A)$. Ainsi, $E_0({}^tAA) \subset E_0(A)$ et donc $E_0(A) = E_0({}^tAA)$.
Le même raisonnement prouve que $E_0(B) = E_0({}^tBB)$.
On en déduit que

$$\text{rg}(A) = n - \dim E_0(A) = n - \dim E_0({}^tAA) = n - \dim E_0({}^tBB) = n - \dim E_0(B) = \text{rg}(B).$$

[1] Seul le vecteur nul est de norme nulle.

2. Si B est inversible, pour avoir $A = UB$, nécessairement on doit avoir $U = AB^{-1}$.
Il s'agit donc de prouver que $U = AB^{-1}$ est une matrice orthogonale.
Or

$${}^tUU = {}^t\left(AB^{-1}\right)\left(AB^{-1}\right) = {}^t\left(B^{-1}\right){}^tAAB^{-1}.$$

Mais la relation ${}^tAA = {}^tBB$ devient, après multiplication à droite par B^{-1} et à gauche par ${}^t\left(B^{-1}\right) = \left({}^tB\right)^{-1}$:

$${}^t\left(B^{-1}\right){}^tAAB^{-1} = I_n \text{ et donc } {}^tUU = I_n$$

de sorte que I_n est orthogonale.

□

Exercice 1.60 (QSP HEC 2013) [HEC13.QSP02] — Facile

Matrices symétrique dont une puissance vaut I_n.

Soit $A \in \mathcal{M}_n(\mathbf{R})$ une matrice symétrique telle qu'il existe un entier $k \in \mathbf{N}^*$ tel que $A^k = I_n$.
Que peut-on dire de A^2 ?

Solution de l'exercice 1.60 Puisque A est symétrique à coefficients réels, elle est diagonalisable : il existe des réels $\lambda_1, \dots, \lambda_n$ et une matrice inversible P tels que $A = P^{-1}\text{Diag}(\lambda_1, \dots, \lambda_n)P$.
Et donc $I_n = A^k = P^{-1}\text{Diag}(\lambda_1^k, \dots, \lambda_n^k)P$.
Autrement dit $\text{Diag}(\lambda_1^k, \dots, \lambda_n^k)$ est semblable à I_n. Mais la seule matrice semblable à I_n est I_n elle-même, et donc $\lambda_1^k = \cdots = \lambda_n^k = 1$.
Mais les seuls réels qui élevés à la puissance k valent 1 sont 1 et -1 (dans le cas où k est pair), de sorte que tous les λ_i valent ± 1.
Et donc $A^2 = P^{-1}\text{Diag}(\lambda_1^2, \dots, \lambda_n^2)P = P^{-1}\text{Diag}(1, 1, \dots, 1)P = P^{-1}I_nP = I_n$. □

L'exercice qui suit demande de diagonaliser explicitement des matrices de taille $2n$. Il faut alors un peu d'intuition pour être capable de voir les relations liant les colonnes et de les exploiter pour en déduire les valeurs propres et les vecteurs propres. Un pivot pour déterminer les valeurs propres serait pour le moins fastidieux.

Exercice 1.61 (QSP HEC 2015) [HEC15.QSP149] — Moyen

Étude d'une famille de matrices.

On note E_n l'espace vectoriel des matrices à $2n$ lignes et $2n$ colonnes, à coefficients réels, de la forme

$$\begin{pmatrix} a_1 & 0 & \dots & 0 & 0 & \dots & 0 & b_1 \\ 0 & a_2 & & 0 & 0 & & b_2 & 0 \\ \vdots & & \ddots & & & & & \vdots \\ 0 & 0 & & a_n & b_n & & 0 & 0 \\ 0 & 0 & & b_n & a_n & & 0 & 0 \\ \vdots & & & & & \ddots & & \vdots \\ 0 & b_2 & & 0 & 0 & & a_2 & 0 \\ b_1 & 0 & \dots & 0 & 0 & \dots & 0 & a_1 \end{pmatrix}.$$

1. Déterminer une base et la dimension de E_n.
2. Justifier la diagonalisabilité des matrices de E_n et trouver les vecteurs colonnes propres.

3. Quelles sont les matrices inversibles de E_n ?

SOLUTION DE L'EXERCICE 1.61

1. Si A est une matrice de E_n, alors

$$A = \begin{pmatrix} a_1 & 0 & \dots & 0 & 0 & \dots & 0 & b_1 \\ 0 & a_2 & & 0 & 0 & & b_2 & 0 \\ \vdots & & \ddots & & & & & \vdots \\ 0 & 0 & & a_n & b_n & & 0 & 0 \\ 0 & 0 & & b_n & a_n & & 0 & 0 \\ \vdots & & & & & \ddots & & \vdots \\ 0 & b_2 & & 0 & 0 & & a_2 & 0 \\ b_1 & 0 & \dots & 0 & 0 & \dots & 0 & a_1 \end{pmatrix}$$

$$= a_1 \begin{pmatrix} 1 & 0 & \dots & 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & & 0 & 0 & & 0 & 0 \\ \vdots & & \ddots & & & & & \vdots \\ 0 & 0 & & 0 & 0 & & 0 & 0 \\ 0 & 0 & & 0 & 0 & & 0 & 0 \\ \vdots & & & & & \ddots & & \vdots \\ 0 & 0 & & 0 & 0 & & 0 & 0 \\ 0 & 0 & \dots & 0 & 0 & \dots & 0 & 1 \end{pmatrix} + \dots + a_n \begin{pmatrix} 0 & 0 & \dots & 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & & 0 & 0 & & 0 & 0 \\ \vdots & & \ddots & & & & & \vdots \\ 0 & 0 & & 1 & 0 & & 0 & 0 \\ 0 & 0 & & 0 & 1 & & 0 & 0 \\ \vdots & & & & & \ddots & & \vdots \\ 0 & 0 & & 0 & 0 & & 0 & 0 \\ 0 & 0 & \dots & 0 & 0 & \dots & 0 & 0 \end{pmatrix}$$

$$+ b_1 \begin{pmatrix} 0 & 0 & \dots & 0 & 0 & \dots & 0 & 1 \\ 0 & 0 & & 0 & 0 & & 0 & 0 \\ \vdots & & \ddots & & & & & \vdots \\ 0 & 0 & & 0 & 0 & & 0 & 0 \\ 0 & 0 & & 0 & 0 & & 0 & 0 \\ \vdots & & & & & \ddots & & \vdots \\ 0 & 0 & & 0 & 0 & & 0 & 0 \\ 1 & 0 & \dots & 0 & 0 & \dots & 0 & 0 \end{pmatrix} + \dots + b_n \begin{pmatrix} 0 & 0 & \dots & 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & & 0 & 0 & & 0 & 0 \\ \vdots & & \ddots & & & & & \vdots \\ 0 & 0 & & 0 & 1 & & 0 & 0 \\ 0 & 0 & & 1 & 0 & & 0 & 0 \\ \vdots & & & & & \ddots & & \vdots \\ 0 & 0 & & 0 & 0 & & 0 & 0 \\ 0 & 0 & \dots & 0 & 0 & \dots & 0 & 0 \end{pmatrix}$$

$$= \sum_{i=1}^{n} a_i(E_{i,i} + E_{2n-i+1,2n-i+1}) + \sum_{j=1}^{n} bj(E_{i,2n+1-j} + E_{2n+1-i,i}).$$

Notation

Rappelons que $E_{i,j}$ désigne les matrices de la base canonique de $\mathcal{M}_{2n}(\mathbf{R})$, où $E_{i,j}$ a tous ses coefficients nuls, sauf celui situé à la $i^{\text{ème}}$ ligne et $j^{\text{ème}}$ colonne qui vaut 1.

Donc déjà, la famille des $2n$ matrices $(E_{i,i}+E_{2n-i+1,2n-i+1})_{1\leqslant i\leqslant n}$ et $(E_{i,2n+1-j}+E_{2n+1-i,i})_{1\leqslant i\leqslant n}$ est génératrice de E.
Elle est libre car si $a_1, \dots, a_n, b_1, \dots, b_n$ sont tels que

$$\sum_{i=1}^{n} a_i(E_{i,i}+E_{2n-i+1,2n-i+1})+\sum_{j=1}^{n} bj(E_{i,2n+1-j}+E_{2n+1-i,i}) = 0 \Leftrightarrow \begin{pmatrix} a_1 & 0 & \dots & 0 & 0 & \dots & 0 & b_1 \\ 0 & a_2 & & 0 & 0 & & b_2 & 0 \\ \vdots & & \ddots & & & & & \vdots \\ 0 & 0 & & a_n & b_n & & 0 & 0 \\ 0 & 0 & & b_n & a_n & & 0 & 0 \\ \vdots & & & & & \ddots & & \vdots \\ 0 & b_2 & & 0 & 0 & & a_2 & 0 \\ b_1 & 0 & \dots & 0 & 0 & \dots & 0 & a_1 \end{pmatrix} = 0$$

alors, par identification des coefficients, $a_1 = \dots = a_n = b_1 = \dots = b_n = 0$.
Et donc il s'agit d'une base de E_n, qui est donc de dimension $2n$.

2. Les matrices de E_n sont symétriques, à coefficients réels donc sont diagonalisables.

Si $A = \begin{pmatrix} a_1 & 0 & \dots & 0 & 0 & \dots & 0 & b_1 \\ 0 & a_2 & & 0 & 0 & & b_2 & 0 \\ \vdots & & \ddots & & & & & \vdots \\ 0 & 0 & & a_n & b_n & & 0 & 0 \\ 0 & 0 & & b_n & a_n & & 0 & 0 \\ \vdots & & & & & \ddots & & \vdots \\ 0 & b_2 & & 0 & 0 & & a_2 & 0 \\ b_1 & 0 & \dots & 0 & 0 & \dots & 0 & a_1 \end{pmatrix}$, on a alors

$$\begin{pmatrix} a_1 & 0 & \dots & 0 & 0 & \dots & 0 & b_1 \\ 0 & a_2 & & 0 & 0 & & b_2 & 0 \\ \vdots & & \ddots & & & & & \vdots \\ 0 & 0 & & a_n & b_n & & 0 & 0 \\ 0 & 0 & & b_n & a_n & & 0 & 0 \\ \vdots & & & & & \ddots & & \vdots \\ 0 & b_2 & & 0 & 0 & & a_2 & 0 \\ b_1 & 0 & \dots & 0 & 0 & \dots & 0 & a_1 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} a_1 + b_1 \\ 0 \\ \vdots \\ 0 \\ \vdots \\ a_1 + b_1 \end{pmatrix}$$

de sorte que $\begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$ est un vecteur propre de A, associé à la valeur propre $a_1 + b_1$.

Et de même,

$$\begin{pmatrix} a_1 & 0 & \dots & 0 & 0 & \dots & 0 & b_1 \\ 0 & a_2 & & 0 & 0 & & b_2 & 0 \\ \vdots & & \ddots & & & & & \vdots \\ 0 & 0 & & a_n & b_n & & 0 & 0 \\ 0 & 0 & & b_n & a_n & & 0 & 0 \\ \vdots & & & & & \ddots & & \vdots \\ 0 & b_2 & & 0 & 0 & & a_2 & 0 \\ b_1 & 0 & \dots & 0 & 0 & \dots & 0 & a_1 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \\ \vdots \\ 0 \\ -1 \end{pmatrix} = \begin{pmatrix} a_1 - b_1 \\ 0 \\ \vdots \\ 0 \\ \vdots \\ -a_1 + b_1 \end{pmatrix} = (a_1 - b_1) \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \\ -1 \end{pmatrix}$$

et donc nous avons là un vecteur propre associé à la valeur propre $a_1 - b_1$.

Plus généralement, si E_k désigne le $k^{\text{ème}}$ vecteur de la base canonique de $\mathcal{M}_{2n,1}(\mathbf{R})$, alors on a, pour $k \in [\![1, n]\!]$,

$$A(E_i + E_{2n+1-i}) = (a_i + b_i)(E_i + E_{2n+1-i}) \text{ et } A(E_i - E_{2n+1-i}) = (a_i - b_i)(E_i + E_{2n+1-i}).$$

Et donc $a_i + b_i$ et $a_i - b_i$ sont des valeurs propres de A et $E_i + E_{2n+1-i}$ et $E_i - E_{2n+1-i}$ en sont des vecteurs propres associés.

Détails

$E_i + E_{2n+1-i}$ est le vecteur colonne dont tous les coefficients sont nuls à l'exception du $i^{\text{ème}}$ et du $2n + 1 - i^{\text{ème}}$, c'est-à-dire le $i^{\text{ème}}$ en partant du bas.

De plus, la famille formée des $E_i + E_{2n+1-i}$ et $E_i - E_{2n+1-i}$, $1 \leqslant i \leqslant n$ est une famille libre[1]. Étant de cardinal $2n$, elle forme une base de $\mathcal{M}_{2n,1}(\mathbf{R})$, qui est donc formée de vecteurs propres de A.

[1] Il suffit de regarder les coefficients non nuls.

3. Une matrice $A \in E_n$ est inversible si et seulement si elle ne possède pas 0 comme valeur propre.
Or, nous venons d'obtenir toutes[2] les valeurs propres de A :

$$\text{Spec}(A) = \{a_i + b_i,\ i \in [\![1, n]\!]\} \cup \{a_i - b_i,\ i \in [\![1, n]\!]\}.$$

[2] Puisque nous avons une base de vecteurs propres.

Et donc A est inversible si et seulement si pour tout $i \in [\![1, n]\!]$, $a_i \neq \pm b_i$.

□

EXERCICE 1.62 (QSP HEC 2012) [HEC12.QSP12] Facile

Carré d'une similitude symétrique.

Soit E un espace euclidien et soit f un endomorphisme symétrique de E. On suppose qu'il existe une constante réelle $\alpha \geqslant 0$ tel que $\forall x \in E, \|f(x)\| = \alpha\|x\|$.
Montrer que $f^2 = \alpha^2 \mathrm{id}_E$.

SOLUTION DE L'EXERCICE 1.62 Puisque f est symétrique, il est diagonalisable : il existe donc une base $(e_1, \dots, e_n)$ de E formée de vecteurs propres de f.
Notons λ_i la valeur propre associée à e_i.
On a alors pour tout i de $[\![1, n]\!]$, $\|f(e_i)\| = \|\lambda_i e_i\| = |\lambda_i|\|e_i\|$.
Mais puisque $\|f(e_i)\| = \alpha\|e_i\|$, on en déduit que $\lambda_i = \pm\alpha$.

Soit à présent $x \in E$. De manière unique, x s'écrit $x = \sum_{i=1}^{n} x_i e_i$, où les x_i sont des réels. Et alors

$$f^2(x) = \sum_{i=1}^{n} x_i f^2(e_i) = \sum_{i=1}^{n} x_i \alpha^2 e_i = \alpha^2 \sum_{i=1}^{n} x_i e_i = \alpha^2 x.$$

Ceci étant vrai pour tout $x \in E$, on en déduit que $f^2 = \alpha^2 \mathrm{id}_E$.
Remarque *: notons que supposer f diagonalisable à la place de symétrique est suffisant.* □

EXERCICE 1.63 (HEC 2016) [HEC16.182] Moyen

Étude d'une famille d'endomorphisme symétriques

On considère un espace euclidien $(E, (\,|\,))$ de dimension $n \geqslant 2$.
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 $x \in E$, $(x|u(x)) \geqslant 0$.

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 $\mathrm{Im}(u)$.
   (a) Montrer que b est un vecteur propre de u associé à une valeur propre $\mu \geqslant 0$.
   (b) Montrer que pour tout vecteur x de E, $u(x) = \dfrac{\mu}{\|b\|^2}(x|b)b$.
   (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 $f \circ u_a$ appartienne à $T(E)$.
4. Donner une condition nécessaire et suffisante sur $(a, b) \in E^2$ pour que $u_a \circ u_b = u_b \circ u_a$.

SOLUTION DE L'EXERCICE 1.63

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 dans $T(E)$.
Supposons donc que $a \neq 0$. Si $x \in \mathrm{Vect}(a)^\perp$, alors $u_a(x) = (x|a)a = 0$.
Par conséquent, dans une base orthonormée de E obtenue par concaténation de $\dfrac{a}{\|a\|}$ et

d'une base de $\text{Vect}(a)^\perp$, la matrice de u_a est

$$\begin{pmatrix} \|a\|^2 & 0 & \dots & 0 \\ 0 & \vdots & & \vdots \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & 0 \end{pmatrix}.$$

Détails
a est une base de $\text{Vect}(a)$ et donc $\frac{a}{\|a\|}$ est une base orthonormée de $\text{Vect}(a)$.

Cette matrice est de rang 1 et symétrique, donc u_a est de rang 1 et est un endomorphisme symétrique[1].
Enfin, si $x \in E$, alors $(x|u_a(x)) = (x|(x|a)a) = (x|a)(x|a) = (x|a)^2 \geqslant 0$.
Et donc u_a est bien un élément de $T(E)$.

[1] Car nous nous sommes placés dans une base **orthonormée**.

1.b. Soit $x \in E$. Alors $x \in \text{Ker}(u_a) \Leftrightarrow (x|a) = 0 \Leftrightarrow x \in (\text{Vect}(a))^\perp$.
Et donc 0 est valeur propre de u_a, avec

$$\dim E_0(u_a) = \dim (\text{Vect}(a))^\perp = \dim E - \dim \text{Vect}(a) = n - 1.$$

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 $E_{\|a\|^2}(u_a) = \text{Vect}(a)$.

2.a. Si $b \in \text{Im}\, u$, alors il existe $x \in E$ tel que $b = u(x)$.
Puisque u est non nul, $\text{rg}(u) \geqslant 1$, 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 $(x|u(x)) \geqslant 0$ par hypothèse et que $\|x\|^2 > 0$, nécessairement $\mu \geqslant 0$.

Détails
u étant non nul, son rang est supérieur ou égal à 1. Et donc si $u \in T(E)$, nécessairement $\text{rg}(u) = 1$.

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

$$x = \left(x\middle|\frac{b}{\|b\|}\right)\frac{b}{\|b\|} + \sum_{i=2}^{n}(x|e_i)e_i.$$

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 $x \in E$,

$$u(x) = \frac{\mu}{\|b\|^2}(x|b)b = \left(x\middle|\frac{\sqrt{\mu}}{\|b\|}b\right)\frac{\sqrt{\mu}}{\|b\|}b.$$

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 $u = u_a$.

Danger !
Une erreure 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 !

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 $\dfrac{a}{\|a\|}$ forme une base orthonormée de F.
Posons alors $e_1 = \dfrac{a}{\|a\|}$, et soit $(e_2, \dots, e_n)$ est une base orthonormée de $F^\perp$.
Alors $\mathscr{B} = (e_1, e_2, \dots, e_n)$ est une base orthonormée[2] de E.
Or, pour $x \in F^\perp$, on a $u_a(x) = 0$ et donc $f \circ u_a(x) = 0$.
Et donc la matrice de $f \circ u_a$ dans la base orthonormée $\mathscr{B}$ est de la forme

[2] Car concaténation d'une base orthonormée de F et d'une base orthonormée de $F^\perp$.

$$\text{Mat}_{\mathscr{B}}(f \circ u_a) = \begin{array}{c} \begin{matrix} f\circ u_a(e_1) & f\circ u_a(e_2) & \dots & f\circ u_a(e_n) \end{matrix} \\ \begin{pmatrix} \star & 0 & \dots & 0 \\ \star & \vdots & & \vdots \\ \vdots & \vdots & & \vdots \\ \star & 0 & \dots & 0 \end{pmatrix} \end{array} \begin{matrix} e_1 \\ e_2 \\ \vdots \\ e_n \end{matrix}.$$

La base $\mathscr{B}$ étant symétrique, $f \circ u_a$ est symétrique si et seulement si $\text{Mat}_{\mathscr{B}}(f \circ u_a)$ est une matrice symétrique.
On constate que c'est le cas 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 alors 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

$$(f \circ u_a(x)|x) = (\mu\lambda a|\mu a + y) = \lambda\mu^2(a|a) + \lambda\mu \underbrace{(a|y)}_{=0} = \lambda\mu^2\|a\|^2.$$

Pour que ceci soit positif pour tout x, il faut et il suffit que $\lambda \geqslant 0$.

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 $u_a \circ u_b(x) = (u_b(x)|a)\, a = (x|b)(a|b)a$ 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 $\underbrace{\|a\|^2}_{>0} b = (a|b)a$.

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.

□

L'encadrement de la question 3.a de l'exercice qui suit est très très classique, la quantité encadrée s'appelant quotient de Rayleigh.

EXERCICE 1.64 (ESCP 18) [ESCP18.2.10] — Moyen

Encadrement des valeurs propres des solutions d'un équation matricielle

Soit n un entier naturel supérieur ou égal à 3. On identifie un endomorphisme f de $\mathbf{R}^n$ à sa matrice M dans la base canonique et on note donc, par abus de notation, $\text{Ker}(M)$ au lieu de $\text{Ker}(f)$.
Autrement dit, $\text{Ker}(M) = \{X \in \mathscr{M}_{n,1}(\mathbf{R}) : MX = 0\}$.
On note J_n la matrice de $\mathscr{M}_n(\mathbf{R})$ dont tous les coefficients valent 1, et on considère le sous-ensemble $\mathscr{F}$ de $\mathscr{M}_n(\mathbf{R})$ défini par

$$\mathscr{F} = \left\{A \in \mathscr{M}_n(\mathbf{R}) : A + {}^tA = J_n - I_n\right\}.$$

1. Déterminer le spectre de J_n.
2. Soit A une matrice appartenant à $\mathscr{F}$. On suppose que le rang de A est inférieur ou égal à $n-2$.
   (a) Montrer que $\mathrm{Ker}(J_n) \cap \mathrm{Ker}(A) \neq \{0\}$.
   (b) Soit X une matrice colonne non nulle de $\mathscr{M}_{n,1}(\mathbf{R})$ appartenant à $\mathrm{Ker}(A) \cap \mathrm{Ker}(J_n)$.
   Évaluer de deux manières différentes ${}^tX(A+{}^tA)X$ et en déduire une contradiction.
   (c) Quelles sont les valeurs possibles pour le rang de A ?
3. (a) Soit M une matrice symétrique de $\mathscr{M}_n(\mathbf{R})$ et $D = \mathrm{Diag}(\lambda_1, \lambda_2, \dots, \lambda_n)$ une matrice diagonale semblable à M, avec $\lambda_1 \leqslant \lambda_2 \leqslant \cdots \leqslant \lambda_n$.
   Établir pour tout matrice colonne X non nulle de $\mathscr{M}_{n,1}(\mathbf{R})$ l'encadrement suivant : $\lambda_1 \leqslant \dfrac{{}^tXMX}{{}^tXX} \leqslant \lambda_n$.
   (b) En déduire que pour toute matrice A appartenant à $\mathscr{F}$, les valeurs propres réelles de A sont dans l'intervalle $\left[-\dfrac{1}{2}, \dfrac{n-1}{2}\right]$.

Solution de l'exercice 1.64

1. C'est assez classique, et il existe de nombreuses manières d'arriver au résultat.
Commençons par noter que les colonnes de J_n sont toutes égales, et donc $\mathrm{rg}(J_n) = 1$, de sorte que 0 est valeur propre de J_n, avec un sous-espace propre de dimension $n-1$.
D'autre part, $J_n^2 = nJ_n$, de sorte que $X^2 - nX$ est un polynôme annulateur de J_n, dont les racines sont 0 et n.
Et donc si J_n possède une autre valeur propre, c'est nécessairement n, avec un sous-espace propre qui ne peut être que de dimension 1.
Mais J_n est symétrique, et donc est diagonalisable. Elle possède donc nécessairement[1] une autre valeur propre, qui est donc n.
Ainsi, $\mathrm{Spec}(J_n) = \{0, n\}$.

[1] De sorte que la somme des dimensions des sous-espaces propres soit égale à n.

2.a. Supposons que $\mathrm{Ker}(A) \cap \mathrm{Ker}(J_n) = \{0\}$. Alors ces deux sous-espaces propres sont en somme directe, et donc $\mathrm{Ker}(A) \oplus \mathrm{Ker}(J_n)$ est un sous-espace vectoriel de $\mathscr{M}_{n,1}(\mathbf{R})$ de dimension $\dim \mathrm{Ker}(A) + \dim \mathrm{Ker}(J_n) = n - \mathrm{rg}(A) + n - \mathrm{rg}(J_n) = 2n - 1 - \mathrm{rg}(A)$.
Par hypothèse, nous avons $\mathrm{rg}(A) \leqslant n-2$, de sorte que, par le théorème du rang, $\dim \mathrm{Ker}(A) + \dim \mathrm{Ker}(J_n) \geqslant 2n - 1 - (n-2) = n+1$.
C'est trop pour un sous-espace vectoriel de $\mathscr{M}_{n,1}(\mathbf{R})$, puisque $\dim \mathscr{M}_{n,1}(\mathbf{R}) = n$.
Et donc nécessairement, $\mathrm{Ker}(A) \cap \mathrm{Ker}(J_n) \neq \{0\}$.

Dimension
Rappelons que la dimension d'une somme directe est la somme des dimensions. Ceci n'est en revanche plus vrai pour une somme qui n'est pas directe.

2.b. Si $X \in \mathrm{Ker}(A)$, alors $AX = 0$ et donc ${}^tX{}^tA = {}^t(AX) = 0$.
Par conséquent, ${}^tX(A+{}^tA)X = {}^tX(AX) + ({}^tX{}^tA)X = 0$.
D'autre part, si $X \in \mathrm{Ker}(J_n)$, alors ${}^tX(J_n - I_n)X = -{}^tXI_nX = -{}^tXX = -\|X\|^2 < 0$.
Et donc, s'il existe X non nul dans $\mathrm{Ker}(A) \cap \mathrm{Ker}(J_n)$, alors on a simultanément ${}^tX(A+{}^tA)X = 0$ et ${}^tX(A+{}^tA)X = {}^tX(J_n - I_n)X = -\|X\|^2 < 0$, ce qui n'est pas possible.
Donc on ne peut pas avoir $\mathrm{Ker}(A) \cap \mathrm{Ker}(J_n) \neq \{0\}$.

2.c. De ce qui précède, on déduit que nécessairement $\mathrm{rg}(A) \geqslant n-1$, et donc $\mathrm{rg}(A)$ ne peut valoir que n ou $n-1$.

3.a. Puisque M est symétrique, il existe une base orthonormée[2] $(X_1, \dots, X_n)$ de $\mathscr{M}_{n,1}(\mathbf{R})$ formée de vecteurs propres de M. Et quitte à renuméroter ces vecteurs, on peut supposer que X_i est associé à la valeur propre λ_i.

[2] Pour le produit scalaire canonique de $\mathscr{M}_{n,1}(\mathbf{R})$.

Alors, tout vecteur X non nul de $\mathscr{M}_{n,1}(\mathbf{R})$ s'écrit de manière unique $X = \sum_{i=1}^{n} \alpha_i X_i$, où $\alpha_1, \dots, \alpha_n$ ne sont pas tous nuls.
On a alors

$$ {}^tXMX = \langle X, MX\rangle = \left\langle \sum_{i=1}^{n} \alpha_i X_i, \sum_{j=1}^{n} \alpha_j \underbrace{MX_j}_{\lambda_j X_j} \right\rangle = \sum_{i=1}^{n}\sum_{j=1}^{n} \alpha_i\alpha_j\lambda_j\langle X_i, X_j\rangle = \sum_{i=1}^{n} \alpha_i^2\lambda_i. $$

Détails
Puisque $(X_1, \dots, X_n)$ est orthonormée

$$ \langle X_i, X_j\rangle = \begin{cases} 0 & \text{si } i \neq j \\ 1 & \text{si } i = j \end{cases} $$

Et d'autre part, ${}^tXX = \left\langle \sum_{i=1}^{n} \alpha_i X_i, \sum_{j=1}^{n} \alpha_j X_j \right\rangle = \sum_{i=1}^{n}\sum_{j=1}^{n} \alpha_i\alpha_j\langle X_i, X_j\rangle = \sum_{i=1}^{n} \alpha_i^2$.

On a alors

$$ \lambda_1 \sum_{i=1}^{n} \alpha_i^2 \leqslant {}^tXMX \leqslant \lambda_n \sum_{i=1}^{n} \alpha_i^2. $$

Et donc il vient bien $\lambda_1 \leqslant \dfrac{{}^tXMX}{{}^tXX} \leqslant \lambda_n$.

3.b. Soit $A \in \mathcal{F}$, soit $\lambda \in \text{Spec}(A)$ et soit X un vecteur propre de A associé à la valeur propre λ.
Alors ${}^tX(A+{}^tA)X = {}^tX\lambda X + (\lambda {}^tX)X = 2\lambda\|X\|^2$.
D'autre part, ${}^tX(J_n - I_n)X = {}^tXJ_nX - {}^tXI_nX = {}^tXJ_nX - \|X\|^2$.
Et donc $2\lambda + 1 = \dfrac{{}^tXJ_nX}{\|X\|^2}$.
D'après la question précédente, en se rappelant que la plus petite valeur propre de J_n est 0 et que sa plus grande est n, on a donc $0 \leqslant 2\lambda + 1 \leqslant n$, et donc $\lambda \in \left[-\dfrac{1}{2}, \dfrac{n-1}{2}\right]$.

□

1.8.1 Projecteurs orthogonaux

EXERCICE 1.65 (HEC 2018) [HEC18.QSP261] **Facile**

Un calcul de projeté orthogonal

Soit $E = \mathbf{R}_3[X]$ l'espace vectoriel des polynômes de degré inférieur ou égal à 3.
On note H le sous-espace vectoriel de E engendré par les trois polynômes $(X-1)(X-2)(X-4)$, $(X-1)(X-3)(X-4)$ et $(X-2)(X-3)(X-4)$.

1. (a) Justifier que H est un hyperplan de E.
   (b) Trouver une forme linéaire dont le noyau est égal à H. Est-elle unique ?
2. On considère le produit scalaire sur E défini par

$$\forall (P,Q) \in E^2,\ \langle P,Q\rangle = P(1)Q(1) + P(2)Q(2) + P(3)Q(3) + P(4)Q(4).$$

Calculer, pour tout polynôme $P \in E$, la projection orthogonale de P sur $H^\perp$.

SOLUTION DE L'EXERCICE 1.65

1.a. Notons P_1, P_2, P_3 les trois polynômes engendrant H.
Alors la famille (P_1, P_2, P_3) est libre dans E. En effet, si $\lambda_1 P_1 + \lambda_2 P_2 + \lambda_3 P(3) = 0$, alors en évaluant cette relation en $X = 1$, il vient $\lambda_3 \underbrace{P_3(1)}_{\neq 0} = 0$, et donc $\lambda_3 = 0$.
On prouve de même, par évaluation en 2 et en 3 que $\lambda_1 = \lambda_2 = 0$.
Et donc (P_1, P_2, P_3) est une base de H, qui est donc de dimension 3.
Puisque $\dim E = 4 = 3 + 1$, on en déduit que H est un hyperplan de E.

1.b. Il s'agit de remarquer que P_1, P_2, P_3 sont tous les trois dans le noyau de $\varphi : P \mapsto P(4)$.
Et donc $H \subset \text{Ker}\,\varphi$.
Mais puisque $\text{Ker}\,\varphi$ est le noyau d'une forme linéaire non nulle[1], c'est un hyperplan de E.
Et donc $\dim E = \dim \text{Ker}\,\varphi$, de sorte que $E = \text{Ker}\,\varphi$.
Cette forme linéaire n'est pas unique, puisque 2φ possède le même noyau de φ.

> **Remarque**
> De manière générale, nous savons qu'un hyperplan est le noyau d'une forme linéaire. Mais il n'y a jamais unicité de cette forme linéaire, puisque la multiplier par un scalaire **non nul** ne change pas son noyau.

2. Il serait possible de déterminer $H^\perp$, et d'en déterminer une base. Mais remarquons plutôt que $Q = (X-1)(X-2)(X-3)$ est dans $H^\perp$.
En effet, si $P \in H$, alors

$$\langle P,Q\rangle = P(1)\underbrace{Q(1)}_{=0} + P(2)\underbrace{Q(2)}_{=0} + P(3)\underbrace{Q(3)}_{=0} + \underbrace{P(4)}_{=0}Q(4) = 0.$$

Et puisque $\dim H^\perp = \dim E - \dim H = 1$, Q forme donc une base de $H^\perp$.
Et donc $\dfrac{Q}{\|Q\|}$ est une base orthonormée de $H^\perp$, de sorte que

$$p_{H^\perp}(P) = \left\langle P, \frac{Q}{\|Q\|}\right\rangle \frac{Q}{\|Q\|} = \frac{\langle P,Q\rangle}{\|Q\|^2}Q.$$

□

Exercice 1.66 (QSP ESCP 2017) [ESCP17.QSP02] — Moyen

Projecteurs orthogonaux qui commutent

Soit $n \in \mathbf{N}$ tel que $n \geqslant 3$. $\mathbf{R}^n$ est muni de sa structure euclidienne canonique. Soient p, q deux projecteurs de $\mathbf{R}^n$ orthogonaux. On supposer que $p \circ q$ est un projecteur orthogonal.

1. Montrer que $p \circ q = q \circ p$.
2. Montrer que les valeurs propres possibles de $p + q$ sont $\{0, 1, 2\}$.
   Donner un exemple où ces trois nombres sont effectivement valeurs propres de $p + q$.

Solution de l'exercice 1.66

1. Notons A et B les matrices respectives de p et q dans la base canonique.
Puisque p et q sont des projecteurs orthogonaux, ce sont des endomorphismes symétriques, et la base canonique étant orthonormée[1], A et B sont symétriques.
De même, AB, qui est la matrice de $p \circ q$ est symétrique.
Et donc ${}^t(AB) = AB \Leftrightarrow {}^tB{}^tA = AB \Leftrightarrow BA = AB$.
On en déduit que A et B commutent, et donc que p et q commutent également.

2. On a

$$\begin{aligned}(p+q-2\mathrm{id}) \circ (p+q-\mathrm{id}) \circ (p+q) &= \left((p+q)^2 - 3(p+q) + 2\mathrm{id}\right) \circ (p+q) \\ &= \left((p+q)^3 - 3(p+q)^2 + 2(p+q)\right) \\ &= (p^3 + 3p^2 \circ q + 3p \circ q^2 + q^3) - 3(p^2 + 2p \circ q + q^2) + 2p + 2q \\ &= p + q + 6p \circ q - 3p - 6p \circ q - 3q + 2p + 2q = 0.\end{aligned}$$

Et donc $P = (X-2)(X-1)X$ est un polynôme annulateur de $p+q$.
Puisque les valeurs propres de $p+q$ sont parmi les racines de P, on en déduit que les valeurs propres possibles de $p+q$ sont $0, 1$ et 2.

Soient p et q les endomorphismes de $\mathbf{R}^3$ dont les matrices respectives dans la base canonique sont

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \text{ et } B = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Alors il est évident que p et q sont des projecteurs car ils sont diagonalisables et ne possèdent que 0 et 1 comme valeurs propres, et sont symétriques car leurs matrices dans la base canonique sont symétriques.
Ainsi, p et q sont des projecteurs orthogonaux.
De plus, on a $AB = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$, et donc par le même raisonnement, $p \circ q$ est encore un projecteur orthogonal.
Enfin, on a $A + B = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 0 \end{pmatrix}$, qui possède $0, 1$ et 2 comme valeurs propres, et donc $\mathrm{Spec}(p+q) = \{0, 1, 2\}$.

□

Remarque

Nous n'avons pas vraiment utilisé le fait que p et q soient des projecteurs, mais seulement qu'ils soient symétriques.
Plus généralement, la composée de deux endomorphismes symétriques est symétrique si et seulement si ces deux endomorphismes commutent.

Binôme

Puisque p et q commutent, on peut utiliser le binôme de Newton pour calculer $(p+q)^k$.

Exercice 1.67 (ESCP 2016) [ESCP16.2.17] — Moyen

La moyenne des puissances d'une contraction est un projecteur orthogonal

Soit $(E, \langle \cdot, \cdot \rangle)$ un espace euclidien muni de la norme $\| \cdot \|$ associée au produit scalaire, et $u \in \mathscr{L}(E)$ vérifiant :

$$\forall x \in E,\ \|u(x)\| \leqslant \|x\|.$$

1. (a) Soit x un vecteur appartenant à $\mathrm{Ker}(u - \mathrm{id}) \cap \mathrm{Im}(u - \mathrm{id})$. Justifier qu'il existe $y \in E$ tel que : $\forall n \in$

$\mathbf{N}$, $nx = u^n(y) - y$.

(b) En déduire que $E = \mathrm{Ker}(u - \mathrm{id}) \oplus \mathrm{Im}(u - \mathrm{id})$.

2. On pose : $\forall p \in \mathbf{N}$, $v_p = \dfrac{1}{p+1}\displaystyle\sum_{k=0}^{p} u^k$, et on note w le projecteur sur $\mathrm{Ker}(u - \mathrm{id})$ parallèlement à $\mathrm{Im}(u - \mathrm{id})$.
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
$$\forall x \in E,\ \lim_{p\to+\infty} \|v_p(x) - w(x)\| = 0.$$

3. Soit Q un projecteur de E, distinct de l'application nulle.
(a) Montrer que si $\mathrm{Ker}(Q)$ et $\mathrm{Im}(Q)$ sont orthogonaux, alors $\forall x \in E$, $\|Q(x)\| \leqslant \|x\|$.
(b) Réciproquement, on suppose que : $\forall x \in E$, $\|Q(x)\| \leqslant \|x\|$. Soit $x \in \mathrm{Im}(Q)$ et $y \in \mathrm{Ker}(Q)$. En considérant les vecteurs $z = x + \lambda y$ pour $\lambda \in \mathbf{R}$, montrer que $\langle x, y\rangle = 0$.
(c) En déduire que projecteur Q non nul est orthogonal si et seulement si $\forall x \in E$, $\|Q(x)\| \leqslant \|x\|$.

4. En déduire que w est un projecteur orthogonal.

SOLUTION DE L'EXERCICE 1.67

1.a. Puisque $x \in \mathrm{Ker}(u - \mathrm{id}) = E_1(u)$, on a $u(x) = x$.
D'autre part, $x \in \mathrm{Im}(u - \mathrm{id})$, 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 $n \in \mathbf{N}^*$,
$$u^n(y) - y = \sum_{k=1}^{n} \left(u^k(y) - u^{k-1}(y)\right) = \sum_{k=1}^{n} x = n \cdot x.$$
Ce résultat est évidemment encore valable pour $n = 0$ car alors $u^n(y) = y$.

1.b. Par le théorème du rang, il est évident que $\dim E = \dim \mathrm{Ker}(u - \mathrm{id}) + \dim \mathrm{Im}(u - \mathrm{id})$.
Il suffit donc de prouver que $\mathrm{Ker}(u - \mathrm{id}) \cap \mathrm{Im}(u - \mathrm{id}) = \{0_E\}$.
Soit donc $x \in \mathrm{Ker}(u - \mathrm{id}) \cap \mathrm{Im}(u - \mathrm{id})$, et soit y comme dans la question précédente.
On a alors pour tout $n \in \mathbf{N}$,
$$\|nx\| = n\|x\| = \|u^n(y) - y\| \leqslant \|u^n(y)\| + \|y\|.$$
Mais nous savons[1] que $\|u(y)\| \leqslant \|y\|$.
Et donc $\|u^2(y)\| = \|u(u(y))\| \leqslant \|u(y)\| \leqslant \|y\|$.
Une récurrence rapide prouve donc que pour tout $n \in \mathbf{N}$, $\|u^n(y)\| \leqslant \|y\|$.
Et donc $n\|x\| \leqslant \|y\| + \|y\| \leqslant 2\|y\|$.
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 $\mathrm{Ker}(u-\mathrm{id})\cap\mathrm{Im}(u-\mathrm{id}) = \{0_E\}$ et donc que $E = \mathrm{Ker}(u-\mathrm{id})\oplus\mathrm{Im}(u-\mathrm{id})$.

[1] C'est l'hypothèse faite sur u.

2. Soit $x \in E$. Alors, de manière unique, $x = x_1 + x_2$, avec $x_1 \in \mathrm{Ker}(u - \mathrm{id})$ et $x_2 \in \mathrm{Im}(u - \mathrm{id})$.
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 $v_p(x) = x_1 + \dfrac{1}{p+1}\displaystyle\sum_{k=0}^{p} u^k(x_2)$.
Et donc $v_p(x) - w(x) = \dfrac{1}{p+1}\displaystyle\sum_{k=0}^{p} u^k(x_2)$.
Puisque $x_2 \in \mathrm{Im}(u - \mathrm{id})$, 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
$$v_p(x_2) = \frac{1}{p+1}\sum_{k=0}^{p}\left(u^{k+1}(y) - u^k(y)\right) = \frac{1}{p+1}(u^{p+1}(y) - y).$$
Et donc
$$\|v_p(x) - w(x)\| = \|v_p(x_2)\|$$

$$= \frac{1}{p+1}\|u^p(y) - y\|$$
$$\leqslant \frac{1}{p+1}\left(\|u^p(y)\| + \|y\|\right) \quad \text{Inégalité triangulaire.}$$
$$\leqslant \frac{1}{p+1}\left(\|y\| + \|y\|\right)$$
$$\leqslant \frac{2\|y\|}{p+1} \underset{p\to+\infty}{\longrightarrow} 0.$$

Et donc on a bien $v_p(x)$ qui converge vers $w(x)$.

3.a. Si $\mathrm{Im}(Q)$ et $\mathrm{Ker}(Q)$ sont orthogonaux[2], alors pour tout $x \in E$, il existe $x_1 \in \mathrm{Im}(Q)$ et $x_2 \in \mathrm{Ker}(Q)$ tels que $x = x_1 + x_2$, et $Q(x) = x_1$.
Et donc, par le théorème de Pythagore,

$$\|x\|^2 = \|x_1 + x_2\|^2 = \|x_1\|^2 + \|x_2\|^2 \geqslant \|x_1\|^2 \geqslant \|Q(x)\|^2.$$

Et donc on a bien $\|Q(x)\| \leqslant \|x\|$.

[2] C'est-à-dire si Q est un projecteur orthogonal.

3.b. Par hypothèse, pour tout $\lambda \in \mathbf{R}$, on a $\|Q(x+\lambda y)\| \leqslant \|x+\lambda y\| \Leftrightarrow \|x+\lambda y\|^2 - \|x\|^2$.
Or, $\|x+\lambda y\|^2 = \|x\|^2 + 2\lambda\langle x, y\rangle + \lambda^2\|y\|^2$.
Et donc, la fonction $\lambda \mapsto \lambda^2\|x\|^2 + 2\lambda\langle x, y\rangle = \lambda(\lambda\|x\| + 2\langle x, y\rangle)$ 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 $x \in E$, $\|Q(x)\| \leqslant \|x\|$.

Inversement, si Q vérifie pour tout $x \in E$, $\|Q(x)\| \leqslant \|x\|$, alors nous venons de prouver que tout vecteur de $\mathrm{Ker}(Q)$ est orthogonal à tout vecteur de $\mathrm{Im}(Q)$. Et donc $\mathrm{Ker}(Q) \subset (\mathrm{Im}\, Q)^\perp$.
Comme de plus on a $\dim(\mathrm{Im}\, Q)^\perp = \dim E - \dim \mathrm{Im}\, Q = \dim \mathrm{Ker}\, Q$, il vient $\mathrm{Ker}\, Q = (\mathrm{Im}\, Q)^\perp$.
Et donc Q est le projecteur orthogonal sur $\mathrm{Im}\, Q$.

Danger !
Nous n'avons pas prouvé l'égalité entre $\mathrm{Ker}(Q)$ et $(\mathrm{Im}\, Q)^\perp$.
Il se pourrait encore qu'il y ait d'autres vecteurs dans $(\mathrm{Im}\, Q)^\perp$ que ceux de $\mathrm{Ker}\, Q$.

4. Par définition, w est un projecteur, et donc il s'agit de prouver que pour tout $x \in E$, $\|w(x)\| \leqslant \|x\|$.
Or, pour tout $p \in \mathbf{N}$, $\|w(x)\| = \|w(x) - v_p(x) + v_p(x)\| \leqslant \|w(x) - v_p(x)\| + \|v_p(x)\|$.
Et

$$\|v_p(x)\| \leqslant \frac{1}{p+1}\sum_{k=0}^{p}\|u^k(x)\| \leqslant \frac{1}{p+1}\sum_{k=0}^{p}\|x\| \leqslant \|x\|.$$

Puisque $\lim_{p\to+\infty}\|w(x) - v_p(x)\| = 0$, il vient, par passage à la limite,

$$\|w(x)\| \leqslant \|x\|.$$

Et donc w est un projecteur orthogonal.

□

Exercice 1.68 (ESCP 2008) [ESCP08.2.16] — Moyen

Étude d'un endomorphisme de $\mathbf{R}^n$.

Soit $n \in \mathbf{N}^*$. On pose $C = \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix} \in \mathcal{M}_{n,1}(\mathbf{R})$ et $A = I - C\,{}^tC \in \mathcal{M}_n(\mathbf{R})$, où C est un vecteur colonne non nul. On confond $\mathcal{M}_{n,1}(\mathbf{R})$ et $\mathbf{R}^n$.
On note $f \in \mathcal{L}(\mathbf{R}^n)$ l'endomorphisme de matrice A dans la base canonique de $\mathbf{R}^n$. On munit $\mathbf{R}^n$ de son produit scalaire usuel, et on note $\|\cdot\|$ la norme associée.

1. La matrice A est-elle diagonalisable ? Déterminer ses valeurs propres et les vecteurs propres associés.
2. À quelle condition sur C l'application f est-elle un projecteur ? Préciser alors de quel projecteur il s'agit.

3. Dans cette question, $n = 4$, et $C = \frac{1}{2}\begin{pmatrix}1\\-1\\1\\-1\end{pmatrix}$. On note H le sous-espace vectoriel de $\mathbf{R}^4$ d'équation $x-y+z-t = 0$.

   (a) Quelle est la dimension de H ?

   (b) Soit $U = \begin{pmatrix}1\\0\\1\\0\end{pmatrix}$. Déterminer le réel α défini par $\alpha = \inf\{\|U - X\|, X \in H\}$.

   La valeur α est-elle atteinte ? Si oui, préciser pour quel(s) vecteur(s) de H.

   (c) Déterminer, dans la base canonique de $\mathbf{R}^4$, la matrice B de la projection orthogonale sur H.

SOLUTION DE L'EXERCICE 1.68

1. Remarquons que $C^tC = \begin{pmatrix} c_1^2 & c_1c_2 & \dots & c_1c_n \\ c_2c_1 & c_2^2 & \dots & c_2c_n \\ \vdots & \vdots & & \vdots \\ c_nc_1 & c_nc_2 & \dots & c_n^2 \end{pmatrix}$

Notons que A est symétrique, et que nous verrons plus tard qu'une matrice symétrique réelle est nécessairement diagonalisable. Pour l'instant, nous ne pouvons conclure sans chercher les valeurs propres de A.
On a $AX = \lambda X \Leftrightarrow X - C^tCX = \lambda X \Leftrightarrow C^tCX = (1-\lambda)X$.
Ainsi, λ est valeur propre de A si et seulement si $1-\lambda$ est valeur propre de C^tC et $E_\lambda(A) = E_{1-\lambda}(C^tC)$.
On est donc ramenés à chercher les valeurs propres de C^tC.
Mais toutes les colonnes de C^tC sont proportionnelles, donc C est de rang 1, de sorte que 0 est valeur propre de C^tC et $\dim E_0(C^tC) = n-1$.
De plus, si X est orthogonal, alors ${}^tCX = 0$, et donc $C^tCX = 0$.
On en déduit que $\text{Vect}(C)^\perp \subset E_0(C^tC)$. Mais ces deux sous-espaces ont pour dimension $n-1$, donc ils sont égaux : $E_0(C^tC) = \text{Vect}(C)^\perp$.
De plus, $C^tCC = C\|C\|^2 = \|C\|^2C$. Donc $\|C\|^2$ est valeur propre de C^tC.
Comme le sous-espace propre associé est de dimension au plus un, et que C est dans ce sous-espace propre, on en déduit que $E_{\|C\|^2}(C^tC) = \text{Vect}(C)$.
Et donc C^tC est diagonalisable car la somme des dimensions de ses sous-espaces propres vaut n.
Et alors A est diagonalisable, avec pour valeurs propres 1 et $1-\|C\|^2$ et

$$E_1(A) = \text{Vect}(C)^\perp \text{ et } E_{1-\|C\|^2}(A) = \text{Vect}(C).$$

2. f est un projecteur si et seulement si ses valeurs propres sont 1 et 0. C'est-à-dire si et seulement si $\|C\| = 1$ ou $\|C\| = 0$. Ce second cas est exclus car par hypothèse, $C \neq 0$.
Donc f est un projecteur si et seulement si $\|C\| = 1$.
On a alors $E_1(f) = \text{Vect}(C)^\perp$ et $E_0(f) = \text{Vect}(C)$, donc f est le projecteur orthogonal sur $\text{Vect}(C)^\perp$.

Projecteur

On connaît déjà le sens direct. Inversement, si f est diagonalisable et possède 0 et 1 comme valeurs propres, alors il est annulé par $X(X-1) = X^2 - X$, et donc c'est un projecteur.

3.a. H est le noyau de la forme linéaire non nulle $(x,y,z,t) \mapsto x - y + z - t$. C'est donc un hyperplan de $\mathbf{R}^4$, et alors $\dim H = 3$.

3.b. On sait que α est égal à $\|U - p_H(U)\|$, et que ce minimum est atteint uniquement en $X = p_H(U)$.
Notons que $H = \text{Vect}(C)^\perp$, et que C est de norme 1.
On peut donc appliquer le résultat de la question 2 :

$$p_H(U) = (I_4 - {}^tCC)U = U - C^tCU = U - \langle C, U\rangle C = U - C = \frac{1}{1}\begin{pmatrix}1\\1\\1\\1\end{pmatrix}.$$

3.c. Toujours d'après la question 2, la matrice de p_H dans la base canonique est

$$A = I_4 - C^tC = \frac{1}{4}\begin{pmatrix}3&1&-1&1\\1&3&1&-1\\-1&1&3&1\\1&-1&1&3\end{pmatrix}.$$

□

1.8.2 Matrices et endomorphismes positifs

Un thème récurrent des sujets de concours, aussi bien à l'écrit qu'à l'oral, bien que rien ne figure explicitement au programme à ce sujet, est l'étude d'endomorphismes (ou des matrices) symétriques positifs.
Il en existe plusieurs définitions (équivalentes), mais il s'agit des endomorphismes symétriques dont toutes les valeurs propres sont positives, ou de manière équivalentes, ceux tels que pour tout x, $\langle f(x), x\rangle \geqslant 0$ (en termes de matrices : pour tout vecteur colonne X, ${}^tXAX \geqslant 0$).
L'exercice suivant résume bien les méthodes classiques sur les sujet.

Exercice 1.69 (ESCP 2014) [ESCP14.2.16] **Moyen**

Matrices symétriques définies positives

On dit qu'une matrice symétrique M de $\mathcal{M}_n(\mathbf{R})$ est définie positive si pour tout $X \in \mathcal{M}_{n,1}(\mathbf{R})$ non nul, ${}^tXMX > 0$.

1. Soit M une matrice symétrique réelle. Montrer l'équivalence des quatre propositions suivantes :
   i) M est définie positive.
   ii) Les valeurs propres de M sont strictement positives.
   iii) Il existe une matrice orthogonale P et une matrice diagonale D à coefficients diagonaux strictement positifs telles que $A = {}^tPDP$.
   iv) Il existe une matrice inversible L et symétrique telle que $M = L^2$.
2. Soit A et B deux matrices symétriques réelles avec B définie positive. Par la question précédente, on écrit $B = {}^tLL$. Trouver une matrice C symétrique réelle telle que pour tout réel λ, pour tout $X \in \mathcal{M}_{n,1}(\mathbf{R})$:
$$AX = \lambda BX \text{ si et seulement si } CZ = \lambda Z \text{ avec } Z = LX.$$
3. (a) Montrer que l'application Φ définie sur $(\mathcal{M}_{n,1}(\mathbf{R}))^2$ par $\Phi(X,Y) = {}^tXBY$ est un produit scalaire sur $\mathcal{M}_{n,1}(\mathbf{R})$.
   (b) Montrer qu'il existe une base orthonormée $(e_1,\dots,e_n)$ de $\mathcal{M}_{n,1}(\mathbf{R})$ pour ce produit scalaire et $\lambda_1,\dots,\lambda_n$ réels tels que pour tout $i \in [\![1,n]\!]$, $Ae_i = \lambda_i Be_i$.

Solution de l'exercice 1.69

1. Nous allons prouver que *i*) ⇒ *ii*) ⇒ *iii*) ⇒ *iv*).
Supposons que M est définie positive, et soit λ une valeur propre de M et X un vecteur propre associé. Alors
$${}^tXMX = {}^tX\lambda X = \lambda {}^tXX > 0.$$
Or ${}^tXX = \|X\|^2 > 0$, donc $\lambda > 0$.
Ainsi *ii*) est vérifié.

Supposons que toutes les valeurs propres de M sont strictement positives. Puisque M est symétrique réelle, elle est diagonalisable en base orthonormée : il existe une matrice diagonale D et une matrice orthogonale P telles que $M = {}^tPDP$. Mais les coefficients diagonaux de D sont alors les valeurs propres de M et donc sont tous strictement positifs. Ainsi iii) est vérifié et donc *i*) ⇒ *ii*).

Supposons iii) vérifié, avec $D = \mathrm{Diag}(\lambda_1,\dots,\lambda_n)$, et posons $D_1 = \mathrm{Diag}(\sqrt{\lambda_1},\dots,\sqrt{\lambda^n})$.
Soit alors $L = {}^tPD_1P$. L est inversible car ses valeurs propres sont $\sqrt{\lambda_1},\dots,\sqrt{\lambda_n}$ qui sont tous non nuls car les λ_i le sont.
D'autre part, on a ${}^tL = {}^t({}^tPD_1P) = {}^tP{}^tD_1{}^t({}^tP) = {}^tPD_1P = L$. Donc L est symétrique.
Enfin, $L^2 = {}^tPD_1P{}^tPD_1P = {}^tPD_1^2P = {}^tPDP = M$.
Donc iv) est vérifié.

Orthogonalité
La matrice P est orthogonale, donc ${}^tPP = I_n$.

Enfin, si on suppose que $M = L^2$ avec L symétrique inversible, on a, pour $X \in \mathcal{M}_{n,1}(\mathbf{R})$ non nul :
$${}^tXMX = {}^tXL^2X = {}^tX{}^tLLX = {}^t(LX)LX = \|LX\|^2.$$

Mais L étant inversible et X non nul, $LX \neq 0$ et donc $\|LX\|^2 > 0$. Et donc i) est vérifié.

Ainsi, nous avons bien prouvé que les quatre propriétés sont équivalentes.

2. Notons qu'on a

$$AX = \lambda BX \Leftrightarrow AX = \lambda {}^tLLX \Leftrightarrow {}^tL^{-1}AX = \lambda LX \Leftrightarrow {}^tL^{-1}AL^{-1}LX = \lambda LX.$$

Donc si on pose $C = {}^tL^{-1}AL^{-1}$, on a bien, quel que soit $X \in \mathcal{M}_{n,1}(\mathbf{R})$ et quel que soit $\lambda \in \mathbf{R}$,

$$AX = \lambda BX \Leftrightarrow CLX = \lambda LX.$$

Il ne reste qu'à vérifier que C est symétrique, mais on a

$${}^tC = {}^t({}^tL^{-1}AL^{-1}) = {}^tL^{-1}{}^tAL^{-1} = {}^tL^{-1}AL^{-1} = C.$$

3.a. Il est facile de voir que l'application Φ est bilinéaire.
Elle est symétrique car

$$\Phi(X,Y) = {}^tXBY = {}^tX{}^tLLY = \langle LX, LY\rangle = \langle LY, LX\rangle = \Phi(Y,X).$$

> **Notation**
> Ici, on a noté $\langle\cdot,\cdot\rangle$ le produit scalaire canonique de $\mathcal{M}_{n,1}(\mathbf{R})$.

Si $X \in \mathcal{M}_{n,1}(\mathbf{R})$, alors

$$\Phi(X,X) = {}^tXBX = {}^t(LX)LX = \langle LX, LX\rangle = \|LX\|^2 \geqslant 0.$$

Enfin, si $\Phi(X,X) = 0$, le calcul précédent prouve que $\|LX\|^2$, et donc $LX = 0$. Puisque L est inversible, c'est donc que $X = 0$.
Ainsi, Φ est bien un produit scalaire sur $\mathcal{M}_{n,1}(\mathbf{R})$.

3.b. C étant symétrique, il existe une base de $\mathcal{M}_{n,1}(\mathbf{R})$, orthonormée **pour le produit scalaire canonique** et formée de vecteurs propres de C. Notons $X_1, \dots, X_n$ une telle base, et notons $e_i = L^{-1}X_i$.
Alors $(e_1, \dots, e_n)$ est bien une base de $\mathcal{M}_{n,1}(\mathbf{R})$, car image de la base $X_1, \dots, X_n$ par l'automorphisme $X \mapsto L^{-1}X$.
Si λ_i est la valeur propre de C associée à X_i, alors, par la question 2,

> **Rappel**
> L'image d'une base par un isomorphisme est encore une base.

$$CX_i = \lambda_i X_i \Leftrightarrow C(Le_i) = \lambda_i(Le_i) \Leftrightarrow Ae_i = \lambda_i Be_i.$$

De plus, on a alors

$$\Phi(e_i,e_j) = {}^te_iBe_j = {}^t(L^{-1}X_i){}^tLLL^{-1}X_j = {}^tX_i({}^tL^{-1}{}^tL)(LL^{-1})X_j = {}^tX_iX_j = \langle X_i, X_j\rangle = \begin{cases} 0 & \text{si } i \neq j \\ 1 & \text{sinon} \end{cases}$$

Et donc ceci prouve que $(e_1, \dots, e_n)$ est une base orthonormée **pour le produit scalaire Φ.**

□

EXERCICE 1.70 (ESCP 2013) [ESCP13.2.19] — Difficile

Produit de matrices symétriques positives.

Soit $n \geqslant 2$. 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.

1. (a) Soit $S \in \mathcal{S}_n$. Montrer que $S \in \mathcal{S}_n^+ \Leftrightarrow \forall X \in \mathcal{M}_{n,1}(\mathbf{R})$, ${}^tXSX \geqslant 0$.
   (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 $\mathcal{S}^+$. 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.

Solution de l'exercice 1.70

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 orthonormée[1] 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 $X \in \mathcal{M}_{n,1}(R)$, il existe donc $\alpha_1, \dots, \alpha_n$ des réels tels que $X = \sum_{i=1}^{n} \alpha_i X_i$. Alors

$$\begin{aligned}
{}^tXSX &= \left(\sum_{j=1}^{n} \alpha_j {}^tX_j\right) S \left(\sum_{i=1}^{n} \alpha_i X_i\right) \\
&= \left(\sum_{j=1}^{n} \alpha_j {}^tX_j\right)\left(\sum_{i=1}^{n} \alpha_i SX_i\right) \\
&= \left(\sum_{j=1}^{n} \alpha_j {}^tX_j\right)\left(\sum_{i=1}^{n} \alpha_i \lambda_i X_i\right) \\
&= \sum_{i=1}^{n}\sum_{j=1}^{n} \alpha_j \alpha_i \lambda_i {}^tX_j X_i \\
&= \sum_{i=1}^{n} \alpha_i^2 \lambda_i \underbrace{\|X_i\|^2}_{=1} \\
&\geqslant 0.
\end{aligned}$$

[1] Pour le produit scalaire canonique de $\mathcal{M}_{n,1}(\mathbf{R})$, qui est défini par
$\langle X, Y\rangle = {}^tXY$.

Les X_i étant deux à deux orthogonaux, ${}^tX_jX_i = 0$ si $i \neq j$.

Inversement, supposons que pour tout $X \in \mathcal{M}_{n,1}(\mathbf{R})$, ${}^tXSX \geqslant 0$. Soit λ une valeur propre de S et soit $X \in \mathcal{M}_{n,1}(\mathbf{R})$ un vecteur propre associé.
Alors ${}^tXSX = {}^tX\lambda X = \lambda\|X\|^2$.
Et donc[2] $\lambda = \dfrac{{}^tXSX}{\|X\|^2} \geqslant 0$.
Ainsi, toutes les valeurs propres de S sont positives, et donc $S \in \mathcal{S}_n^+$.

[2] $\|X\|$ est non nul car il s'agit d'un vecteur propre.

1.b. Il est clair que S est symétrique car ${}^tS = {}^t({}^tAA) = {}^tAA = S$.
Soit $X \in \mathcal{M}_{n,1}(\mathbf{R})$. Alors ${}^tXSX = {}^tX{}^tAAX = {}^t(AX)AX = \|AX\|^2 \geqslant 0$.
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 $D = \mathrm{Diag}(\lambda_1, \dots, \lambda_n)$ et une matrice orthogonale P telles que $S = {}^tPDP$. Notons que les λ_i sont tous positifs, puisque $S \in \mathcal{S}_n^+$.
Soit alors $D_1 = \mathrm{Diag}(\sqrt{\lambda_1}, \dots, \sqrt{\lambda_n})$, de sorte que ${}^tD_1D_1 = D_1^2 = D$. Et donc en posant $A = D_1P$, il vient

$${}^tAA = {}^tP{}^tD_1D_1P = {}^tPDP = S.$$

2.a. Si 0 est valeur propre de UV, alors UV n'est pas inversible, et donc l'une[3] 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 canonique sont égales à U et V.
Il est classique que $\mathrm{Im}(v \circ u) \subset \mathrm{Im}(v)$ et $\mathrm{Ker}(u) \subset \mathrm{Ker}(v \circ u)$.
Donc si v n'est pas bijectif, il n'est pas surjectif, et donc $\mathrm{Im}(v) \neq \mathbf{R}^n$. On en déduit que $\mathrm{Im}(v \circ u) \neq \mathbf{R}^n$ et donc $v \circ u$ n'est pas bijectif.
Et si u n'est pas bijectif, $\mathrm{Ker}(u) \neq \{0\}$ et donc $\mathrm{Ker}(v \circ u) \neq \{0\}$, 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.

[3] Au moins.

Rappel
Pour un endomorphisme d'un espace de dimension finie, injectivité, bijectivité et surjectivité sont équivalentes.

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.

Réels/complexe
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.

2.c. Puisque $\mathcal{S}_n$ est un sous-espace vectoriel de $\mathcal{M}_n(R)$, $S+T$ est encore symétrique.
Et si $X \in \mathcal{M}_{n,1}(\mathbf{R})$, alors

$$^tX(T+S)X = {}^tXTX + {}^tXSX \geqslant 0.$$

Et donc $S+T \in \mathcal{S}_n^+$.

Sev de $\mathcal{M}_n(\mathbf{R})$
Un bon moyen de prouver que $\mathcal{S}_n$ est un sous-espace vectoriel de $\mathcal{M}_n(R)$ 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.

2.d. 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 $-1 < 0$.

3. 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. 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.

□

⚠ 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.

Exercice 1.71 (QSP HEC 2009) [HEC09.QSP04] — Moyen

Noyau de la somme de deux endomorphismes positifs

Soient f et g deux endomorphismes symétriques d'un espace euclidien E, dont toutes les valeurs propres sont positives ou nulles.

1. Montrer qu'il existe un endomorphisme symétrique ϕ tel que $f = \phi^2$.
2. Montrer que $\mathrm{Ker}(f+g) = \mathrm{Ker}(f) \cap \mathrm{Ker}(g)$.

Solution de l'exercice 1.71

1. Puisque f est symétrique, il existe une base orthonormée de E formée de vecteur propres de f. Notons $\mathcal{B}$ une telle base.
Alors la matrice de f dans cette base est $D = \mathrm{Diag}(\lambda_1, \dots, \lambda_n)$ où les λ_i sont les valeurs propres de f (et donc sont positifs).
Soit alors $A = \mathrm{Diag}(\sqrt{\lambda_1}, \dots, \sqrt{\lambda_n})$, 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.

Rappel
La matrice d'un endomorphisme dans une base formée de vecteurs propres est diagonale.

2. Il est évident que si $x \in \mathrm{Ker}\, f \cap \mathrm{Ker}\, g$, alors $(f+g)(x) = 0$ et donc $x \in \mathrm{Ker}(f+g)$.
Considérons à présent deux endomorphismes symétriques ϕ et ψ tels que $f = \phi^2$ et $g = \psi^2$.
Soit alors $x \in \mathrm{Ker}(f+g)$, 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, $x \in \mathrm{Ker}(f) \cap \mathrm{Ker}(g)$ et donc $\mathrm{Ker}(f+g) \subset \mathrm{Ker}(f) \cap \mathrm{Ker}(g)$.
Puisque l'inclusion réciproque a été prouvée précédemment, on a bien $\mathrm{Ker}(f+g) = \mathrm{Ker}(f) \cap \mathrm{Ker}(g)$.

□

1.8.3 Endomorphismes antisymétriques

Les questions 1 et 2 de l'exercice qui suit prouvent les principales propriétés des endomorphismes antisymétriques (encore une notion qui n'est pas au programme mais que l'on rencontre régulièrement). La dernière question est bien plus délicate, avec une récurrence que l'on rencontre parfois (à l'oral comme à l'écrit des parisiennes), qui porte sur la dimension d'un espace vectoriel.

Exercice 1.72 (HEC 2017) [HEC17.215] **Moyen**

Matrice d'un endomorphisme antisymétrique dont le carré ne possède qu'une valeur propre.

Soit E un espace euclidien de dimension $n \geqslant 1$ muni d'un produit scalaire noté $\langle \cdot, \cdot \rangle$.
Dans tout l'exercice, on considère un endomorphisme φ de E antisymétrique, c'est-à-dire tel que

$$\forall (x,y) \in E^2,\ \langle x, \varphi(y)\rangle = -\langle \varphi(x), y\rangle.$$

1. Établir les propriétés suivantes :
   (a) Pour tout $x \in E$, on a : $\langle x, \varphi(x)\rangle = 0$.
   (b) $\mathrm{Im}(\varphi) = (\mathrm{Ker}(\varphi))^\perp$.
   (c) Soit F un sous-espace vectoriel de E. Montrer que si F est stable par φ, alors $F^\perp$ est stable par φ.
   (d) $\mathrm{Ker}(\varphi) = \mathrm{Ker}\,(\varphi^2)$, où $\varphi^2 = \varphi \circ \varphi$.
   (e) Le spectre de φ est soit vide, soit réduit à $\{0\}$.
2. Montrer que toutes les valeurs propres de φ^2 sont négatives ou nulles.
3. Soit :
   - F un sous-espace vectoriel de E de dimension $p \geqslant 2$.
   - α un réel strictement positif.
   - u un endomorphisme antisymétrique de F tel que $u^2 = -\alpha^2 \mathrm{id}_F$, où id_F est l'endomorphisme identité de F.

   (a) On suppose que $p = 2$. Établir l'existence d'une base orthonormale de F dans laquelle la matrice A_α de u est donnée par : $A_\alpha = \begin{pmatrix} 0 & -\alpha \\ \alpha & 0 \end{pmatrix}$.
   (b) À l'aide d'un raisonnement par récurrence sur p, montrer qu'il existe une base orthonormale de F dans laquelle la matrice B_α de u est de la forme : $B_\alpha = \begin{pmatrix} A_\alpha & (0) & \dots & (0) \\ (0) & A_\alpha & \ddots & \vdots \\ \vdots & \ddots & \ddots & (0) \\ (0) & \dots & (0) & A_\alpha \end{pmatrix}$.

Solution de l'exercice 1.72

1.a. Soit $x \in E$. Alors on a

$$\langle x, \varphi(x)\rangle = -\langle \varphi(x), x\rangle = -\langle x, \varphi(x)\rangle.$$

Et donc nécessairement, $\langle \varphi(x), x\rangle = 0$.

1.b. Soit $x \in \mathrm{Ker}\,\varphi$ et $y \in \mathrm{Im}(\varphi)$. Alors il existe $z \in E$ tel que $y = \varphi(z)$, et donc

$$\langle x, y\rangle = \langle x, \varphi(z)\rangle = -\langle \varphi(x), z\rangle = -\langle 0_E, z\rangle = 0.$$

Donc $\mathrm{Im}\,\varphi \subset (\mathrm{Ker}(\varphi))^\perp$. D'autre part, d'après le théorème du rang,

$$\dim \mathrm{Im}\,\varphi = \dim E - \dim \mathrm{Ker}(\varphi) = \dim (\mathrm{Ker}(\varphi))^\perp.$$

Et donc $\mathrm{Im}(\varphi) = (\mathrm{Ker}(\varphi))^\perp$.

1.c. Soit $x \in F^\perp$. Alors, pour tout $y \in F$, $\langle x, y\rangle = 0$.
Et alors, $\langle \varphi(x), y\rangle = -\langle x, \varphi(y)\rangle$.
Mais $\varphi(y) \in F$ car F est stable par φ, et donc $\langle x, \varphi(y)\rangle = 0$.
On en déduit que $\langle \varphi(x), y\rangle = 0$, et donc $\varphi(x) \in F^\perp$.
Ainsi, $F^\perp$ est stable par φ.

1.d. On a déjà[1], $\mathrm{Ker}(\varphi) \subset \mathrm{Ker}\left(\varphi^2\right)$.
Soit $x \in \mathrm{Ker}\left(\varphi^2\right)$. Alors

$$\|\varphi(x)\|^2 = \langle \varphi(x), \varphi(x)\rangle = -\langle x, \varphi^2(x)\rangle = 0.$$

Et donc $\varphi(x) = 0$: $x \in \mathrm{Ker}(\varphi)$. Ainsi, $\mathrm{Ker}\left(\varphi^2\right) \subset \mathrm{Ker}(\varphi)$.
On a donc bien prouvé que $\mathrm{Ker}(\varphi) = \mathrm{Ker}\left(\varphi^2\right)$.

[1] Comme pour tout endomorphisme.

1.e. Il s'agit donc de prouver que la seule valeur propre possible de φ est 0.
Soit donc λ une valeur propre de φ et x un vecteur propre associé.
On a alors, d'après 1.a, $0 = \langle x, \varphi(x)\rangle = \langle x, \lambda x\rangle = \lambda\|x\|^2$.
Puisque $x \neq 0_E$, on a $\|x\|^2 \neq 0$, et donc $\lambda = 0$.
Ainsi, λ est la seule valeur propre possible de φ, et donc le spectre de φ est soit vide, soit réduit à $\{0\}$.

2. Soit λ une valeur propre de φ^2, et soit x un vecteur propre associé. Alors

$$\|\varphi(x)\|^2 = \langle \varphi(x), \varphi(x)\rangle = -\langle x, \varphi^2(x)\rangle = -\langle x, \lambda x\rangle = -\lambda\|x\|^2.$$

Et donc $\lambda = -\dfrac{\|\varphi(x)\|^2}{\|x\|^2} \leqslant 0$.

3.a. Soit e_1 un vecteur unitaire de F, et soit $e_2 = \dfrac{1}{\alpha}u(e_1)$.
Alors $u(e_1) = \alpha e_2$, et $u(e_2) = \dfrac{1}{\alpha}u^2(e_1) = -\alpha e_1$.
De plus, on a $\langle e_1, e_2\rangle = \dfrac{1}{\alpha}\langle e_1, u(e_1)\rangle = 0$.
Donc (e_1, e_2) est orthogonale, et en particulier est libre : c'est une base de F. Enfin,

$$\|e_2\|^2 = \frac{1}{\alpha^2}\|u(e_1)\|^2 = \frac{1}{\alpha^2}\langle u(e_1), u(e_1)\rangle = -\frac{1}{\alpha^2}\langle e_1, u^2(e_1)\rangle = -\frac{1}{\alpha^2}\langle e_1, -\alpha^2 e_1\rangle = \langle e_1, e_1\rangle = 1.$$

Donc (e_1, e_2) est une base orthonormée de F, dans laquelle la matrice de u est

$$\begin{array}{c} \begin{matrix} u(e_1) & u(e_2) \end{matrix} \\ \begin{pmatrix} 0 & -\alpha \\ \alpha & 0 \end{pmatrix} \begin{matrix} e_1 \\ e_2 \end{matrix} \end{array} = A_\alpha.$$

3.b. Notons que le résultat sous-entend que $p = \dim F$ est nécessairement pair.
Procédons donc à une récurrence sur la dimension p de F. Autrement dit, cherchons à prouver par récurrence sur $p \geqslant 2$ la propriété $\mathscr{P}(p)$: «pour tout sous-espace vectoriel F de E de dimension inférieure ou égale à p, et pour tout endomorphisme antisymétrique u de F vérifiant $u^2 = -\alpha^2\mathrm{id}_F$, il existe une base orthonormée de F dans laquelle la matrice de u est de la forme B_α».
La récurrence a été initialisée à la question précédente pour $p = 2$. Supposons donc que $\mathscr{P}(p)$ est vraie, et soit F un sous-espace vectoriel de E de dimension inférieure ou égale à $p+1$, et soit u un endomorphisme antisymétrique de F tel que $u^2 = -\alpha^2\mathrm{id}_F$.
Si $\dim F \leqslant p$, alors il n'y a rien à prouver : le résultat est directement dans l'hypothèse de récurrence.
Si $\dim F = p+1$, choisissons, comme dans la question précédente, un vecteur unitaire e_1 de F et posons $e_2 = \dfrac{1}{\alpha}u(e_1)$.
Posons alors $G = \mathrm{Vect}(e_1, e_2)$, de sorte que G est stable par u. Comme prouvé à la question 1.c, $G^\perp$ est également stable par u. Notons qu'alors $\dim G^\perp = \dim F - \dim G = \dim F - 2 \leqslant p$.
Soit alors $v : G^\perp \to G^\perp$, $x \mapsto u(x)$ la restriction de u à $G^\perp$.
Alors v est un endomorphisme de $G^\perp$, qui est encore antisymétrique et qui vérifie $g^2 = -\alpha^2\mathrm{id}_{G^\perp}$.

Détails

Par $G^\perp$, nous désignons ici le supplémentaire orthogonal de G dans F, c'est-à-dire l'ensemble de vecteurs de F (et non de E), qui sont orthogonaux à tous les vecteurs de G.

Par hypothèse de récurrence, il existe donc une base orthonormée $(e_3, \dots, e_n)$ de $G^\perp$ dans laquelle la matrice de g est $\begin{pmatrix} A_\alpha & (0) & \dots & (0) \\ (0) & A_\alpha & \ddots & \vdots \\ \vdots & \ddots & \ddots & (0) \\ (0) & \dots & (0) & A_\alpha \end{pmatrix}$.

Blocs
Cette matrice contient donc $\dim(G^\perp)/2$ blocs diagonaux égaux à A_α.

Alors la famille $(e_1, e_2, e_3, \dots, e_n)$ est une base orthonormée[2] de F, dans laquelle la matrice de u est $\begin{pmatrix} A_\alpha & (0) & \dots & (0) \\ (0) & A_\alpha & \ddots & \vdots \\ \vdots & \ddots & \ddots & (0) \\ (0) & \dots & (0) & A_\alpha \end{pmatrix}$, avec un bloc diagonal égal à A_α de plus que la matrice de v.

[2] Car concaténation d'une base orthonormée de G et d'une base orthonormée de $G^\perp$.

Ainsi, la propriété est vraie au rang $p+1$, et donc est vraie pour tout $p \geqslant 2$.

□

L'exercice suivant nécessite une bonne compréhension du procédé d'orthogonalisation de Gram-Schmidt : si $(e_1, \dots, e_n)$ est une famille libre d'un espace euclidien E, alors le procédé de Gram-Schmidt nous permet de construire une famille orthonormée $(f_1, \dots, f_n)$ à partir de $(e_1, \dots, e_n)$.
Si l'orthonormalité de $(f_1, \dots, f_n)$ est généralement connue, de même que les formules donnant les f_i en fonction de e_i, on oublie trop souvent que cette nouvelle famille possède une propriété supplémentaire : pour tout $i \in [\![1, n]\!]$, $\text{Vect}(f_1, \dots, f_i) = \text{Vect}(e_1, \dots, e_i)$.

Exercice 1.73 (QSP HEC 2011) [HEC11.QSP108] — Moyen

Endomorphismes antisymétriques trigonalisables

Soit $(E, \langle \cdot, \cdot \rangle)$ un espace euclidien et soit u un endomorphisme de E pour lequel il existe une base de E dans laquelle la matrice de u est triangulaire supérieure.

1. Montrer qu'il existe une base orthonormée de E dans laquelle la matrice de u est triangulaire supérieure (*on pourra penser au procédé d'orthonormalisation de Gram-Schmidt*).
2. On suppose de plus que pour tout x de E, $\langle u(x), x \rangle = 0$.
   Montrer que $\forall (x, y) \in E^2$, $\langle u(x), y \rangle = -\langle u(y), x \rangle$. Que peut-on dire de u ?

Solution de l'exercice 1.73

1. Soit $(e_1, \dots, e_n)$ une base de E dans laquelle la matrice $A = (a_{i,j})_{1 \leqslant i,j \leqslant n}$ de u soit triangulaire supérieure.
C'est le cas si et seulement si pour tout $j \in [\![1, n]\!]$, $u(e_j) \in \text{Vect}(e_1, \dots, e_j)$.
Soit alors $(f_1, \dots, f_n)$ la base orthonormée obtenue en appliquant le procédé d'orthonormalisation de Gram-Schmidt à la base $(e_1, \dots, e_n)$.
On sait alors[1] que pour tout $j \in [\![1, n]\!]$, $f_j \in \text{Vect}(e_1, \dots, e_j)$. Et donc par linéarité de u, $u(f_j)$ est combinaison linéaire de $u(e_1), \dots, u(e_j)$.
Mais comme il a été dit précédemment, pour tout $i \in [\![1, j]\!]$,

$$u(e_i) \in \text{Vect}(e_1, \dots, e_i) \subset \text{Vect}(e_1, \dots, e_j) = \text{Vect}(f_1, \dots, f_j).$$

Et donc $u(f_j) \in \text{Vect}(f_1, \dots, f_j)$, de sorte que la matrice de u dans la base orthonormée $(f_1, \dots, f_n)$ est triangulaire supérieure.

[1] Voir la remarque précédant l'exercice.

2. Soient $x, y \in E$. Alors, l'hypothèse faite sur u implique que

$$\langle u(x+y), x+y \rangle = 0 \Leftrightarrow \langle u(x) + u(y), x+y \rangle = 0.$$

Mais par bilinéarité du produit scalaire, on a donc

$$\underbrace{\langle u(x), x \rangle}_{=0} + \langle u(x), y \rangle + \langle u(y), x \rangle + \underbrace{\langle u(y), y \rangle}_{=0} = 0.$$

Soit $\langle u(x), y \rangle = -\langle x, u(y) \rangle$.

Mais si on note $B = (b_{i,j})_{1\leqslant i,j\leqslant n}$ la matrice de u dans la base orthonormée obtenue précédemment, on a $b_{i,j} = \langle u(f_j), f_i\rangle$.
En particulier, pour tout $i \in [\![1, n]\!]$, $b_{i,i} = \langle u(f_i), f_i\rangle = -\langle f_i, u(f_i)\rangle = -b_{i,i}$ et donc $b_{i,i} = 0$.
Pour $i \neq j$, alors soit $i > j$, et alors $b_{i,j} = 0$ car la matrice B est triangulaire supérieure, soit $i < j$, et alors

$$b_{i,j} = \langle u(f_j), f_i\rangle = -\langle f_j, u(e_i)\rangle = -b_{j,i} = 0.$$

Et donc tous les coefficients de B sont nuls, de sorte que $u = 0$.

□

Détails

Dans une base **orthonormée** $(f_1, \dots, f_n)$, tout vecteur x se décompose sous la forme

$$x = \sum_{i=1}^{n} \langle x, f_i\rangle f_i.$$

Or, $b_{i,j}$ est précisément la coordonnée de $u(f_j)$ en f_i.

2 Analyse

2.1 Fonctions d'une variable

Exercice 2.1 (QSP ESCP 2018) [ESCP18.QSP01]

Une fonction continue possède moins de points fixes que son carré

Soit $f : \mathbf{R} \to \mathbf{R}$ continue. On appelle point fixe de f tout réel x tel que $f(x) = x$.

1. Montrer que f admet un point fixe si et seulement si $f \circ f$ admet un point fixe.
2. Montrer que le nombre de points fixes de f est inférieur ou égal à celui de $f \circ f$.

Solution de l'exercice 2.1

1. Si x est un point fixe de f, alors $(f \circ f)(x) = f(f(x)) = f(x) = x$, donc x est un point fixe de $f \circ f$.
Donc déjà, si f admet un point fixe, alors $f \circ f$ admet un point fixe.

Inversement, supposons que f ne possède pas de point fixe. Alors, la fonction $x \mapsto f(x) - x$ est de signe constant. En effet, si elle changeait de signe, par le théorème des valeurs intermédiaires[1], il existerait $x \in \mathbf{R}$ tel que $f(x) - x = 0 \Leftrightarrow f(x) = x$.
Supposons donc qu'elle soit strictement positive, c'est-à-dire que pour tout $x \in \mathbf{R}$, $f(x) - x > 0 \Leftrightarrow f(x) > x$.
Alors pour tout $x \in \mathbf{R}$, $f(f(x)) > f(x) > x$. Et donc $f \circ f$ n'admet pas de point fixe.
De même, si $\forall x \in \mathbf{R}$, $f(x) < x$, alors pour tout x, $f(f(x)) < f(x) < x$, et donc $f \circ f$ n'admet pas non plus de point fixe.

Nous avons donc prouvé que f admet un point fixe si et seulement si $f \circ f$ admet un point fixe.

2. Nous avons prouvé précédemment que si x est un point fixe de f, alors c'est également un point fixe de $f \circ f$.
Et donc $f \circ f$ possède au moins autant de points fixes que f.

□

> **Remarque**
> Notons que la continuité de f ne nous a été d'aucun utilité ici.

[1] Et c'est ici que la continuité de f est indispensable !

Exercice 2.2 (QSP HEC 2007) [HEC07.QSP115] Facile

Généralisation de la fonction W de Lambert

Soit α un réel strictement positif. Montrer que pour tout réel x positif, il existe un unique réel positif noté $f(x)$ tel que $f(x)e^{f(x)} = x^{\alpha}$.
Étudier ensuite la dérivabilité de $f(x)$, et exprimer f' en fonction de f le cas échéant.

SOLUTION DE L'EXERCICE 2.2 La fonction $g : t \mapsto te^t$ est dérivable sur $\mathbf{R}_+$ et $g'(t) = e^t + te^t = e^t(1+t) > 0$.
Donc g est strictement croissante sur $\mathbf{R}_+$, avec $g(0) = 0$ et $\lim_{t\to+\infty} g(t) = +\infty$.
Par le théorème de la bijection[1], g réalise une bijection de $\mathbf{R}_+$ sur lui-même.
Et donc en particulier, pour tout $x \in \mathbf{R}_+$, il existe un unique $f(x)$ tel que $g(f(x)) = x^\alpha \Leftrightarrow f(x)e^{f(x)} = x^\alpha$.

[1] g étant dérivable, elle est continue.

Pour $\alpha = 1$, notons $W(x)$ l'unique solution de $te^t = x$. Autrement dit, soit W la bijection réciproque de g.
Alors pour tout $x \in \mathbf{R}_+$, $W(x)e^{W(x)} = x \Leftrightarrow e^{W(x)} = \dfrac{x}{W(x)}$.
Puisque g est dérivable sur $\mathbf{R}_+$, avec une dérivée qui ne s'annule pas, W est également dérivable, avec

$$W'(x) = \frac{1}{g'(W(x))} = \frac{1}{(1+W(x))e^{W(x)}} = \frac{W(x)}{x(1+W(x))}.$$

Dans le cas où α est quelconque, on a $f(x) = W(x^\alpha)$, qui est donc dérivable sur $\mathbf{R}_+^*$ car composée de fonctions qui le sont. Et on a alors

$$f'(x) = \alpha x^{\alpha-1}W'(x^\alpha) = \alpha x^{\alpha-1}\frac{W(x^\alpha)}{x^\alpha(1+W(x^\alpha))} = \frac{\alpha}{x}\frac{f(x)}{1+f(x)}.$$

Il reste donc juste à étudier la dérivabilité de f en 0.
Si $\alpha \geqslant 1$, alors $x \mapsto x^\alpha$ est dérivable en 0, de dérivée nulle, et donc $f'(0) = 0$.
Enfin, si $\alpha < 1$, alors $x \mapsto x^\alpha$ n'est pas dérivable en 0. Et donc f ne l'est pas non plus, car si elle l'était, $x^\alpha = f(x)e^{f(x)}$ le serait également par opération sur les fonctions dérivables. □

Astuce

Rappelons qu'il existe une formule pour la dérivée d'une bijection réciproque :

$$\left(f^{-1}\right)'(x) = \frac{1}{f'(f^{-1}(x))}$$

qui se retrouve facilement en dérivant la relation

$$f(f^{-1}(x)) = x.$$

EXERCICE 2.3 (QSP HEC 2013) [HEC13.QSP03] — Facile

Non annulation de fonctions vérifiant une équation fonctionnelle.

Soit f une application de $\mathbf{R}$ dans $\mathbf{R}$, continue sur $\mathbf{R}$ et telle que $\forall (x,y) \in \mathbf{R}^2$, $f(x+y)f(x-y) = f(x)^2$.
On suppose que f n'est pas la fonction nulle et on se propose de montrer que f ne s'annule pas sur $\mathbf{R}$.

1. Montrer que $f(0)$ n'est pas égal à 0.
2. Soit $a \in \mathbf{R}$ tel que $f(a) = 0$. Montrer que $f\left(\frac{a}{2}\right) = 0$.
3. Conclure. Donner un exemple de telle fonction

SOLUTION DE L'EXERCICE 2.3

1. Puisque f est non nulle, il existe $x \in \mathbf{R}$ telle que $f(x) \neq 0$. Et alors

$$f(x+x)f(x-x) = f(x)^2 \Leftrightarrow f(2x)f(0) = f(x)^2 \neq 0.$$

Et donc en particulier, $f(0) \neq 0$.

2. Supposons que $f(a) \neq 0$. Alors

$$f\left(\frac{a}{2}+\frac{a}{2}\right)f\left(\frac{a}{2}-\frac{a}{2}\right) = f\left(\frac{a}{2}\right)^2 \Leftrightarrow f(a)f(0) = f\left(\frac{a}{2}\right)^2 \Leftrightarrow 0 = f\left(\frac{a}{2}\right)^2.$$

Et donc $f\left(\frac{a}{2}\right) = 0$.

3. Supposons par l'absurde qu'il existe $a \in \mathbf{R}$ tel que $f(a) = 0$.
Alors par la question précédente, $f\left(\frac{a}{2}\right) = 0$. Et donc $f\left(\frac{a}{4}\right) = 0$. Et une récurrence rapide utilisant le résultat de la question précédente prouve que pour tout $n \in \mathbf{N}$, $f\left(\frac{a}{2^n}\right) = 0$.
Mais $\frac{a}{2^n} \underset{n\to+\infty}{\longrightarrow} 0$, et par continuité de la fonction f en 0

$$f(0) = \lim_{n\to+\infty} f\left(\frac{a}{2^n}\right) = \lim_{n\to+\infty} 0 = 0.$$

Remarque

C'est le seul moment où l'on utilise la continuité de f : il n'était donc pas nécessaire de la supposer continue sur $\mathbf{R}$, seule la continuité en 0 nous est utile.

Ceci contredit le résultat de la question 1, et donc notre hypothèse n'est pas valable : il n'existe pas de réel a tel que $f(a) = 0$, la fonction f ne s'annule donc pas sur $\mathbf{R}$.

Un exemple de telle fonction est par exemple donné par la fonction $x \mapsto e^x$ puisque

$$e^{x+y}e^{x-y} = e^{2x} = (e^x)^2.$$

□

2.1.1 Théorème de Rolle et conséquences

Exercice 2.4 (QSP ESCP 2012) [ESCP12.QSP05] — Moyen

Une suite définie par une équation implicite
Abordable en première année : ✓

Pour tout $n \in \mathbf{N}^*$, on pose $P_n(X) = \prod_{k=0}^{n}(X-k)$.

1. Montrer qu'il existe un unique réel $r_n \in]0,1[$ tel que $P_n'(r_n) = 0$.
2. Montrer que pour $x \notin [\![0,n]\!]$, $\dfrac{P_n'(x)}{P_n(x)} = \sum_{k=0}^{n}\dfrac{1}{x-k}$. En déduire $\lim\limits_{n\to+\infty} r_n$.

Solution de l'exercice 2.4

1. La fonction P_n est polynomiale, donc dérivable, et elle vérifie $P_n(0) = P_n(1) = 0$.
Donc par le théorème de Rolle, il existe un réel $r_n \in]0,1[$ tel que $P_n'(r_n) = 0$.
Toutefois, le théorème de Rolle ne saurait en rien garantir l'unicité d'un tel r_n.
Pour cela, notons que P_n' est un polynôme de degré n, puisque P_n est de degré $n+1$. Il possède donc au plus n racines.
Le même raisonnement que précédemment montre que P_n' possède[1] une racine dans $]0,1[$, une dans $]1,2[$, etc, une dans $]n-1,n[$.
Ce qui fait déjà au total n racines, soit le nombre maximal possible. Donc P_n' possède exactement une racine dans chacun de ces intervalles, et en particulier une unique racine dans $]0,1[$.

[1] Au moins.

⚠ Attention !
P_n n'est pas de signe constant, et change de signe en 0, en 1, en 2, etc.
On n'oubliera donc pas la valeur absolue.

2. Notons que $\dfrac{P_n'(x)}{P_n(x)}$ est la dérivée de $x \mapsto \ln(|P_n(x)|)$.
Mais $\ln(|P_n(x)|) = \ln\left(\prod_{k=0}^{n}|x-k|\right) = \sum_{k=0}^{n}\ln|x-k|$.
Et donc en dérivant, ce qui est possible sur $\mathbf{R}\setminus[\![0,n]\!]$, il vient :

$$\forall x \notin [\![0,n]\!],\ \frac{P_n'(x)}{P_n(x)} = \sum_{k=0}^{n}\frac{1}{x-k}.$$

En particulier, en évaluant cette relation en r_n, il vient

$$0 = \frac{P_n'(r_n)}{P_n(r_n)} = \sum_{k=0}^{n}\frac{1}{r_n-k}.$$

Et donc $\dfrac{1}{r_n} = \sum_{k=1}^{n}\dfrac{1}{k-r_n}$.
Mais pour tout $k \geqslant 1$, puisque $r_n \in]0,1[$, on a $k-1 < k-r_n < k$.
Et donc $\dfrac{1}{k-r_n} > \dfrac{1}{k}$.
Par conséquent, $\dfrac{1}{r_n} > \sum_{k=1}^{n}\dfrac{1}{k}$.
Mais puisque $\sum_{k=1}^{n}\dfrac{1}{k} \underset{n\to+\infty}{\longrightarrow} +\infty$., on en déduit que $\dfrac{1}{r_n} \underset{n\to+\infty}{\longrightarrow} +\infty$ et donc $r_n \underset{n\to+\infty}{\longrightarrow} 0$.

Détails
On a reconnu les sommes partielles de la série harmonique. Puisque cette série diverge et est à termes positifs, ses sommes partielles tendent vers $+\infty$.

□

EXERCICE 2.5 (QSP ESCP 2010) [ESCP10.QSP01] — Moyen

L'équation $P(x) = e^x$, $P \in \mathbf{R}[X]$ n'admet qu'un nombre fini de solutions.
Abordable en première année : ✓

Soit $P \in \mathbf{R}[X]$. Montrer que l'équation $P(x) = e^x$ possède un nombre fini de solutions.

SOLUTION DE L'EXERCICE 2.5 Notons qu'il se peut tout à fait que l'équation n'admette aucune solution, par exemple si $P(X) = -1$ ou $P(X) = -X^2$.

Si P est de degré 0 (c'est-à-dire un polynôme constant), alors par injectivité de l'exponentielle, l'équation admet au plus une seule solution. Plus précisément, elle admet une solution si P est strictement positif, et aucune solution sinon.

Montrons par récurrence sur $n = \deg P$ que l'équation $P(x) = e^x$ admet au plus $n+1$ solutions.
Pour $n = 0$, nous venons de prouver le résultat.
Supposons le résultat vrai pour tout polynôme de degré n, et soit $P \in \mathbf{R}[X]$ de degré $n+1$.
Soit alors $\varphi : x \mapsto P(x) - e^x$.
Supposons par l'absurde que φ s'annule strictement plus de $n+2$ fois.
Alors d'après le théorème de Rolle (φ est dérivable sur $\mathbf{R}$), φ' s'annule au moins $n+1$ fois.
Mais $\varphi'(x) = P'(x) - e^x$, et donc $\varphi'(x) = 0 \Leftrightarrow P'(x) = e^x$.
Or P' est un polynôme de degré n, donc l'équation $P'(x) = e^x$ possède au plus $n+1$ solutions.
Donc $P(x) = e^x$ possède au plus $n+2$ solutions.
Par le principe de récurrence, le résultat est vrai pour tout polynôme à coefficients réels.

Plus rapide *: nous avons ici donné une majoration du nombre de solutions, mais il était possible d'aller bien plus vite si on souhaite juste prouver qu'il existe un nombre fini de solutions.*
Si $P(x) = e^x$ possède une infinité de solutions, alors la fonction $\varphi : e^x - P(x)$ s'annule une infinité de fois. Donc sa dérivée également[1]*, donc sa dérivée seconde aussi, etc.*
Or en dérivant plus de $\deg(P)$ fois, on arrive à la conclusion que e^x s'annule une infinité de fois, ce qui est absurde. □

Remarque
Nous avons en fait prouvé un peu mieux que le résultat demandé : l'équation admet au maximum $\deg(P) + 1$ solutions.

[1] C'est encore le théorème de Rolle.

Rappelons qu'une fonction T-périodique est une fonction telle que pour tout $x \in \mathbf{R}$, $f(x+T) = f(x)$. De tels exemples de fonctions sont donnés par les fonctions trigonométriques : sin et cos sont 2π-périodiques.
Aucune connaissance n'est exigée sur le sujet

EXERCICE 2.6 (QSP ESCP 2012) [ESCP12.QSP03] — Facile

La dérivée d'une fonction périodique f s'annule au moins autant de fois que f.
Abordable en première année : ✓

Soit f une fonction T-périodique, $T > 0$, dérivable sur $\mathbf{R}$.
On supposer que f s'annule en p points distincts $x_1 < x_2 < \cdots < x_p$ de $[0, T[$.
Montrer que f' s'annule en au moins p points distincts de $[0, T[$ et distincts de $x_1, \ldots, x_p$.

SOLUTION DE L'EXERCICE 2.6 Par le théorème de Rolle[1], pour tout $i \in [\![1, p-1]\!]$, il existe $y_i \in]x_i, x_{i+1}[$ tel que f' s'annule en y_i.
D'autre part, $f(x_p) = f(x_1 + T)$, et donc f' s'annule en un point $y_p \in]x_p, x_1 + T[$.
Si $y_p \in [0, T[$, alors on a bien trouvé p points de $[0, T[$, distincts des x_i en lesquels f' s'annule.

[1] Qui s'applique puisque f est dérivable et que $f(x_i) = f(x_{i+1})$.

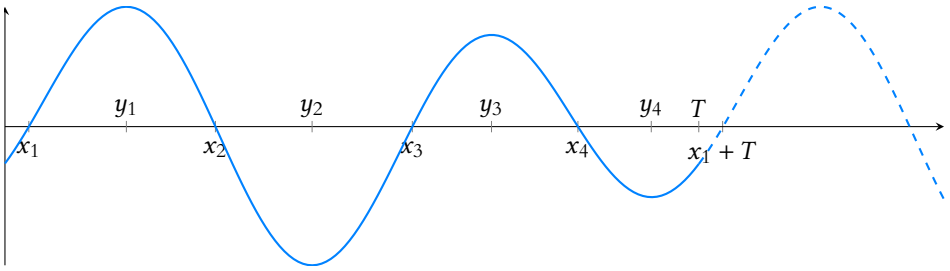


FIGURE 2.1 – Le cas le plus favorable : celui où y_p est dans $[0,T[$.

Sinon, il s'agit de noter que f' est également T-périodique. En effet,

$$f'(x+T)=\lim_{h\to 0}\frac{f(x+T+h)-f(x+T)}{h}=\lim_{h\to 0}\frac{f(x+h)-f(x)}{h}=f'(x).$$

Et donc on a également $f'(y_p-T)=0$, avec $y_p-T\in[0,T[$, car $y_p\in[T,x_1+T]$, de sorte que $y_p-T\in[0,x_1[$.
Et donc f' s'annule en $y_p-T<y_1<\cdots,y_{p-1}$, qui sont tous dans $[0,T[$, et tous distincts des x_i.

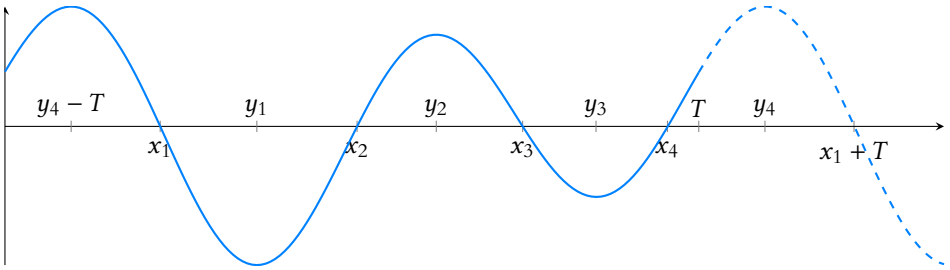


FIGURE 2.2 – Le cas où y_p n'est pas dans $[0,T[$.

Remarque *: il se peut toutefois que f' s'annule strictement plus de fois que f dans l'intervalle $[0,T[$: il n'y a pas forcément une seule solution de $f'(x)=0$ entre deux solutions de $f(x)=0$.*

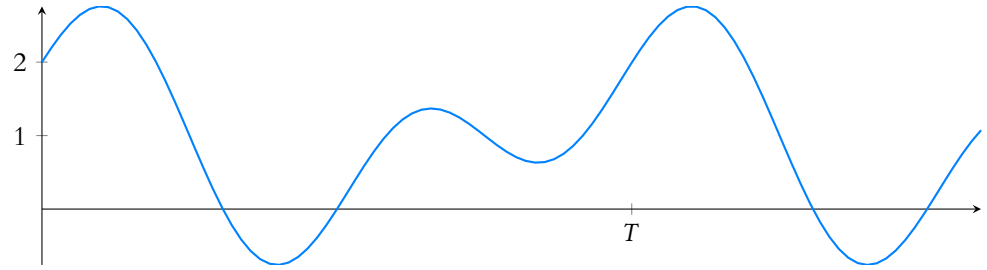


FIGURE 2.3 – Une fonction telle que f' s'annule 4 fois dans $[0,T[$ alors que f ne s'y annule que deux fois.

□

2.1.2 Fonctions convexes/concaves

EXERCICE 2.7 (QSP ESCP 17) [ESCP17.QSP03] **Moyen**

Majoration de la moyenne d'une fonction concave.
Abordable en première année : ✓

On considère une fonction f deux fois dérivable sur $[0,1]$ telle que pour tout $x\in[0,1]$, $f''(x)\leqslant 0$.
Montrer que $\displaystyle\int_0^1 f(t)\,dt\leqslant f\left(\frac{1}{2}\right)$. *On pourra commencer par le cas où* $f'\left(\frac{1}{2}\right)=0$.

SOLUTION DE L'EXERCICE 2.7

• Commençons par supposer que $f'\left(\frac{1}{2}\right) = 0$.

Puisque f est concave[1], sa courbe représentative est située sous ses tangentes. Et en particulier sous la tangente en $x = \frac{1}{2}$, qui est la droite d'équation $y = \underbrace{f'\left(\frac{1}{2}\right)}_{=0} x + f\left(\frac{1}{2}\right)$.

[1] Sa dérivée seconde est négative.

Et donc pour tout $t \in [0,1]$, $f(t) \leqslant f\left(\frac{1}{2}\right)$.

Et alors par croissance de l'intégrale, $\int_0^1 f(t)\,dt \leqslant \int_0^1 f\left(\frac{1}{2}\right) dt = f\left(\frac{1}{2}\right)$.

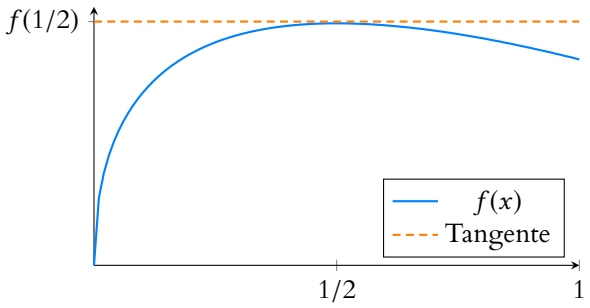


FIGURE 2.4 – La fonction f est située sous sa tangente en $x = 1/2$.

• Passons à présent au cas général, et posons $h(x) = f(x) - f'\left(\frac{1}{2}\right)x$.

La fonction h est alors deux fois dérivable et $h''(x) = f''(x) \leqslant 0$.

D'autre part, $h'\left(\frac{1}{2}\right) = f'\left(\frac{1}{2}\right) - f'\left(\frac{1}{2}\right) = 0$.

Il est donc possible d'appliquer ce qui a été fait plus haut à la fonction h : $\int_0^1 h(t)\,dt \leqslant h\left(\frac{1}{2}\right)$.

Soit encore

$$\int_0^1 f(t)\,dt - f'\left(\frac{1}{2}\right)\frac{1}{2} \leqslant f\left(\frac{1}{2}\right) - \frac{1}{2}f'\left(\frac{1}{2}\right)$$

et donc $\int_0^1 f(t)\,dt \leqslant f\left(\frac{1}{2}\right)$. □

Remarque

Ajouter une fonction affine à une fonction ne change pas sa convexité/concavité. En effet, les fonctions affines sont des fonctions qui sont à la fois convexe et concaves.

L'inégalité de la première question de l'exercice suivant est vraiment un grand classique.

EXERCICE 2.8 (QSP ESCP 2013) [ESCP13.QSP01] Facile

Autour de l'inégalité $\ln(x) \leqslant x - 1$.
Abordable en première année : ✓

1. Soit $u \geqslant 1$. Comparer $\ln(u)$ et $u - 1$.
2. Soit $f \in \mathscr{C}^1(\mathbf{R}_+, \mathbf{R}_+)$ telle que $f(0) = 1$ et pour tout $x > 0$, $f(x) > 1$.
   Montrer que pour tout $x > 0$,
$$f'(x) \geqslant \frac{1}{\ln(f(x))} \Rightarrow f(x) \geqslant 1 + \sqrt{2x}.$$

SOLUTION DE L'EXERCICE 2.8

1. La fonction ln est concave sur $\mathbf{R}_+^*$, et donc est située en dessous de ses tangentes.
Or l'équation de sa tangente en $x = 1$ est $y = x - 1$.
Et donc on en déduit que pour tout $u > 0$, $\ln(u) \leqslant u - 1$.

2. Puisque f est à valeurs supérieures à 1, $\ln(f(x))$ est positif, et donc pour tout $x > 0$, on a $f'(x)\ln(f(x)) \geqslant 1$.
Mais alors pour tout $x > 0$, par croissance de l'intégrale,

$$\int_0^x f'(t)\ln(f(t))\,dt \geqslant \int_0^x 1\,dt = x.$$

Nous reconnaissons que $x \mapsto f'(x)\ln(f(x))$ est la dérivée de $x \mapsto \frac{1}{2}(\ln(f(x)))^2$, et donc que

$$\int_0^x f'(t)\ln(f(t))\,dt = \left[\frac{1}{2}\ln^2(f(t))\right]_0^x = \frac{1}{2}\ln^2(f(x)).$$

On en déduit donc que $\ln^2(f(x)) \geqslant 2x$.
D'après l'inégalité de concavité de la première question, pour tout $x > 0$, $\ln(f(x)) \geqslant f(x)-1$ et donc[1], $\ln^2(f(x)) \geqslant (f(x)-1)^2$.
On a donc

$$(f(x)-1)^2 \geqslant 2x \Rightarrow f(x)-1 \geqslant \sqrt{2x} \Rightarrow f(x) \geqslant 1+\sqrt{2x}.$$

[1] Il s'agit de nombres positifs.

□

2.2 Suites numériques

Certains exercices nécessitent d'avoir les idées claires sur la définition de la convergence d'une suite.
Rappelons que $\lim\limits_{n\to+\infty} u_n = \ell$ si et seulement si

$$\forall \varepsilon > 0,\ \exists N \in \mathbf{N} \text{ tel que } n \geqslant N \Rightarrow |u_n - \ell| \leqslant \varepsilon.$$

Exercice 2.9 (ESCP 2017) [ESCP17.1.13] Moyen

Suites sous-additives, lemme de Fekete et chemins auto-évitants
Abordable en première année : ✓

Soit $(u_n)_{n\geqslant 1}$ une suite réelle vérifiant pour tous n, m entiers naturels non nuls,

$$u_{m+n} \leqslant u_n + u_m.$$

On pose, pour tout $n \in \mathbf{N}^*$, $v_n = \min\limits_{k\in[\![1,n]\!]} \frac{u_k}{k}$.

1. Montrer que la suite $(v_n)_{n\geqslant 1}$ admet une limite ℓ dans $\{-\infty\} \cup \mathbf{R}$.
2. Montrer que pour tous n, m entiers naturels non nuls, on a $u_{mn} \leqslant mu_n$.
3. On suppose dans cette question que ℓ ne vaut pas $-\infty$. Soit $\varepsilon > 0$.
   (a) Montrer qu'il existe $m \in \mathbf{N}$ tel que $\frac{u_m}{m} \leqslant \ell + \varepsilon$.
   (b) En utilisant la division euclidienne de n par m, montrer que la suite $\left(\frac{u_n}{n}\right)_{n\geqslant 1}$ converge vers ℓ lorsque n tend vers $+\infty$.
4. Dans la suite, on appelle *chemin sans croisement de longueur n* toute suite $M_0, M_1, \ldots, M_n$ de points du plan à coordonnées entières vérifiant :
   i) $M_0 = O$ (origine du plan)
   ii) pour tout $i \in [\![0, n-1]\!]$, la distance entre M_i et M_{i+1} est égale à 1.
   iii) pour tout $i \neq j$, on a $M_i \neq M_j$.
   On note N_n le nombre de chemins sans croisement de longueur n.
   (a) Montrer que $N_n \leqslant 4^n$.
   (b) Montrer que pour tous n, m entiers naturels non nuls, $N_{m+n} \leqslant N_n N_m$.
   (c) Quelle relation vérifie la suite $u_n = \ln N_n$?
   (d) En déduire que la suite $\left(N_n^{1/n}\right)_n$ converge.

SOLUTION DE L'EXERCICE 2.9

1. La suite (v_n) est décroissante puisque le minimum définissant v_{n+1} porte sur $n+1$ termes, dont ceux définissant v_n, donc $v_{n+1} \leqslant v_n$.
Par le théorème de la limite monotone, soit elle est minorée et admet alors une limite finie, soit elle tend vers $-\infty$.

> **Méthode**
> Lorsqu'on nous demande de prouver l'existence d'une limite, mais qu'on ne demande pas la valeur de cette limite, le théorème de la limite monotone doit être le premier réflexe, et donc il faut penser à étudier la monotonie de la suite.

2. Fixons $n \in \mathbf{N}^*$ et prouvons le résultat par récurrence sur m. Pour $m = 1$, c'est évident.
Supposons que $u_{mn} \leqslant mu_n$. Alors

$$u_{(m+1)n} = u_{mn+n} \leqslant u_{mn} + u_n \leqslant mu_n + u_n \leqslant (m+1)u_n.$$

Donc par le principe de récurrence, pour tout $m \in \mathbf{N}^*$, $u_{mn} \leqslant mu_n$.

3.a. Puisque $v_n \underset{n\to+\infty}{\longrightarrow} \ell$, il existe $N \in \mathbf{N}^*$ tel que pour $n \geqslant N$, $|v_n - \ell| \leqslant \varepsilon$, et donc en particulier, $v_n - \ell \leqslant \varepsilon$.
En particulier, pour $n = N$, on a $v_N \leqslant \ell + \varepsilon$.
Mais puisque $v_n = \min_{k\in[\![1,N]\!]} \dfrac{u_k}{k}$, il existe $m \in [\![1, N]\!]$ tel que $\dfrac{u_m}{m} \leqslant \ell + \varepsilon$.

3.b. Soit $n \geqslant m$. Notons alors $n = a_n m + b_n$, avec $b_n < m$.
On a alors

$$u_n = u_{a_n m + b_n} \leqslant u_{a_n m} + u_{b_n} \leqslant a_n u_m + u_{b_n}.$$

Et donc

$$\frac{u_n}{n} \leqslant \frac{a_n}{n} u_m + \frac{u_{b_n}}{n}.$$

Mais si on note $M = \max(|u_1|, |u_2|, \dots, |u_{m-1}|)$, alors on a $|u_{b_n}| \leqslant M$.
Et donc $\left|\dfrac{u_{b_n}}{n}\right| \leqslant \dfrac{M}{n}$, de sorte que $\lim\limits_{n\to+\infty} \dfrac{u_{b_n}}{n} = 0$.
Ainsi, il existe $N \in \mathbf{N}$ tel que pour $n \geqslant N$, $\dfrac{u_{b_n}}{n} \leqslant \varepsilon$.

D'autre part, on a $a_n m \leqslant n$ et donc $\dfrac{a_n}{n} \geqslant \dfrac{1}{m}$ et donc pour $n \geqslant N$,

$$u_n \leqslant \frac{u_m}{m} + \frac{u_{b_n}}{n} \leqslant \ell + \varepsilon + \varepsilon = \ell + 2\varepsilon.$$

> **Détails**
> La suite(v_n) est décroissante, donc tous ses termes sont supérieurs ou égaux à sa limite.

Enfin, pour tout $n \in \mathbf{N}$, on a $\dfrac{u_n}{n} \geqslant v_n \geqslant \ell$.
Nous venons donc de prouver que pour tout $\varepsilon > 0$, il existe $N \in \mathbf{N}$ tel que pour $n \geqslant N$, $\ell \leqslant \dfrac{u_n}{n} \leqslant \ell + \varepsilon \Rightarrow \left|\dfrac{u_n}{n} - \ell\right| \leqslant \varepsilon$.
Ceci prouve donc que $\dfrac{u_n}{n} \underset{n\to+\infty}{\longrightarrow} \ell$.

4. Avant de démarrer une question, il faut s'assurer qu'on a bien compris la définition qui est donnée. Un chemin sans croisement est une succession de points de $\mathbf{Z}^2$, telle que la distance entre deux points successifs soit égale à 1. Or, les seuls points de $\mathbf{Z}^2$ situés à distance 1 d'un point $A \in \mathbf{Z}^2$ sont les points qui sont à gauche, à droite, en haut et en bas de A.
Enfin, la dernière condition signifie qu'un chemin ne passe jamais deux fois par le même point[1].

[1] Ce qui explique le terme «sans croisement».

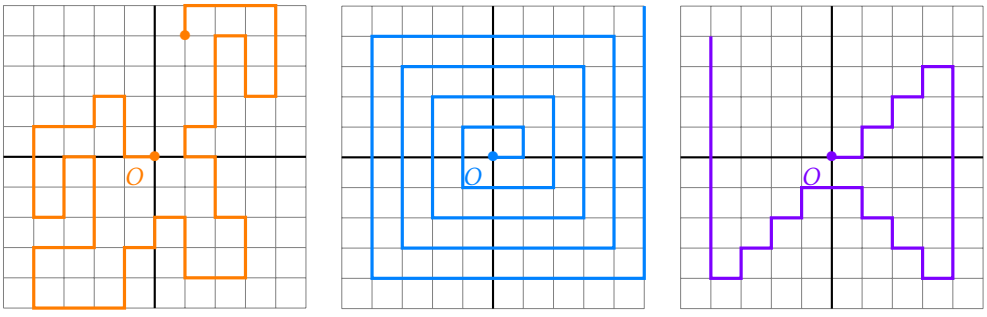


FIGURE 2.5 – Quelques exemples de chemins sans croisements.

4.a. Puisque les points d'un chemin doivent être à coordonnées entières, et que la distance entre deux points successifs doit être égale à 1, alors si $M_n = (a, b)$, il n'y a au maximum[2] que quatre choix possibles pour M_{n+1} : ce sont les points $(a+1, b), (a-1, b), (a, b+1)$ et $(a, b-1)$.
Et donc, on a $N_{n+1} \leqslant 4N_n$ et une récurrence rapide[3] prouve alors que $N_n \leqslant 4^n$.
Remarque : notons qu'il est en fait aisé de faire bien mieux : il y a 4 choix pour la position de M_1, mais pour chacun des points suivants, il n'y a au maximum que 3 choix possibles, puisqu'il n'est pas possible de revenir de la direction dont on vient.
Et donc $M_n \leqslant 4 \times 3^{n-1}$.

[2] L'un (ou plusieurs) de ces points peut déjà faire partie du chemin, et donc n'est pas possible.

[3] Pour l'initialisation, notons que $N_0 = 1$.

4.b. Si $M_0, M_1, \dots, M_{m+n}$ est un chemin sans croisement de longueur n, alors $M_0, \dots, M_n$ est un chemin sans croisement de longueur n.
Et $M_n, M_{n+1}, \dots, M_{n+m}$ est un chemin de longueur m, qui satisfait aux axiomes *ii*) et *iii*), mais pas à *i*).
Toutefois, si on impose l'origine d'un chemin, que ce soit O ou n'importe quel autre point ne change pas le nombre de chemins.
Il y a N_n choix pour le premier chemin de longueur n, et au plus N_m choix pour le second

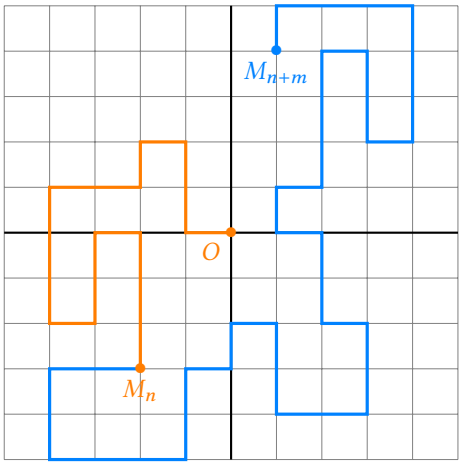


FIGURE 2.6 – Un chemin sans croisement de longueur $n+m$ peut être coupé en deux chemins sans croisement de longueurs respectives n et m.

de longueur m.
Et donc $N_{m+n} \leqslant N_n N_m$.

Inégalité
Notons qu'il y a strictement moins de N_m choix possibles pour le second car certains des chemins au départ de M_n vont croiser le premier chemin de longueur n, de sorte que la concaténation des deux n'est plus un chemin sans croisement.

4.c. En passant au logarithme, on a donc $\ln(N_{n+m}) \leqslant \ln(N_n) + \ln(N_m) \Leftrightarrow u_{n+m} \leqslant u_n + u_m$.

4.d. La suite (u_n) vérifie les hypothèses du début de l'exercice.
Puisqu'elle est positive, c'est également le cas de la suite (v_n), qui ne peut donc pas tendre vers $-\infty$.
D'après les questions 1 et 3, $\dfrac{u_n}{n}$ converge donc vers une limite ℓ.
Mais $\dfrac{u_n}{n} = \ln\left(M_n^{1/n}\right)$, de sorte que, par composition par l'exponentielle, $M_n^{1/n} \underset{n\to+\infty}{\longrightarrow} e^{\ell}$.
Si on a bien prouvé l'existence du nombre ℓ, sa valeur exacte reste inconnue. Des simulations informatiques permettent juste d'en donner une valeur approchée qui est 2.638.

□

EXERCICE 2.10 (ESCP 2008) [ESCP08.1.2] Moyen

Comparaison des vitesses de convergence de deux suites qui tendent vers π.
Abordable en première année : ✓

Soit $(u_n)_{n\in\mathbf{N}^*}$ la suite définie par $u_1 = 2$ et la relation de récurrence :

$$\forall n \geqslant 1,\ u_{n+1} = \frac{\sqrt{2}u_n}{\sqrt{1+\sqrt{1-\left(\frac{u_n}{2^n}\right)^2}}}.$$

1. Montrer que : $\forall n \in \mathbf{N}^*,\ u_n = 2^n \sin\left(\frac{\pi}{2^n}\right)$.
2. Déterminer la limite de la suite $(u_n)_{n\in\mathbf{N}^*}$.
3. Montrer que pour tout $n \in \mathbf{N}^*$, on a $|\pi - u_n| \leqslant \frac{\pi^3}{6\times 4^n}$.
4. Écrire une fonction `Scilab` , n'utilisant pas la constante `%pi`, qui prend comme paramètre un entier q, et retourne une valeur approchée de π à 10^{-q} près.
5. Soit $p \in \mathbf{N}$ fixé. Montrer que lorsque n tend vers $+\infty$, (u_n) admet le développement asymptotique suivant

$$u_n = \pi - \frac{\pi^3}{6\times 4^n} + \cdots + (-1)^p \frac{\pi^{2p+1}}{(2p+1)!\times 4^{pn}} + \underset{n\to+\infty}{o}\left(\frac{1}{4^{pn}}\right).$$

6. On définit une nouvelle suite (v_n) par :

$$\forall n \in \mathbf{N}^*,\ v_n = \frac{1}{3}\left(-u_n + 4u_{n+1}\right).$$

   Montrer que la suite (v_n) converge vers π et qu'au voisinage de $+\infty$, on a : $(v_n - \pi) = \underset{n\to+\infty}{o}(u_n - \pi)$.
7. Donner un équivalent de $(v_n - \pi)$ lorsque n tend vers $+\infty$.

SOLUTION DE L'EXERCICE 2.10

1. Montrons le résultat par récurrence sur n.
Pour $n = 1$, c'est évident car $2\sin(\pi/2) = 2 = u_1$.
Supposons donc que $u_n = 2^n \sin\left(\frac{\pi}{2^n}\right)$. Alors

$$u_{n+1} = \frac{\sqrt{2}2^n \sin\left(\frac{\pi}{2^n}\right)}{\sqrt{1+\sqrt{1-\sin^2\left(\frac{\pi}{2^n}\right)}}} = 2^n \frac{\sqrt{2}\sin\left(\frac{\pi}{2^n}\right)}{\sqrt{1+\cos\left(\frac{\pi}{2^n}\right)}}.$$

Signe

Puisque $\frac{\pi}{2^n}$ est dans $[0, \frac{\pi}{2}]$, son cosinus est positif et donc

$$\sqrt{1-\sin^2\left(\frac{\pi}{2^n}\right)} = \sqrt{\cos^2\left(\frac{\pi}{2^n}\right)} = \cos\left(\frac{\pi}{2^n}\right).$$

Une identité classique de trigonométrie est

$$\cos(2x) = 2\cos^2(x) - 1 \Leftrightarrow \frac{\cos(2x)+1}{2} = \cos^2(x).$$

En particulier, pour $x = \frac{\pi}{2^{n+1}}$, on obtient

$$\frac{1+\cos\left(\frac{\pi}{2^n}\right)}{2} = \cos^2\left(\frac{\pi}{2^n}\right).$$

Et donc[1], il vient

$$u_{n+1} = 2^n \frac{\sin\left(\frac{\pi}{2^n}\right)}{\cos\left(\frac{\pi}{2^{n+1}}\right)}.$$

[1] Toujours par positivité du cosinus sur $[0, \frac{\pi}{2}]$

Or, nous savons que pour tout x, $\sin(2x) = 2\sin(x)\cos(x)$ et donc pour $x = \frac{\pi}{2^{n+1}}$, il vient

$$\sin\left(\frac{\pi}{2^n}\right) = 2\sin\left(\frac{\pi}{2^{n+1}}\right)\cos\left(\frac{\pi}{2^{n+1}}\right).$$

Et donc on a bien $u_{n+1} = 2\times 2^n \sin\left(\frac{\pi}{2^{n+1}}\right) = 2^{n+1}\sin\left(\frac{\pi}{2^{n+1}}\right)$.
Par le principe de récurrence, pour tout $n \in \mathbf{N}^*$, $u_n = 2^n \sin\left(\frac{\pi}{2^n}\right)$.

2. Lorsque $n \to +\infty$, $\dfrac{\pi}{2^n} \underset{n\to+\infty}{\longrightarrow} 0$ et donc $\sin\left(\dfrac{\pi}{2^n}\right) \underset{n\to+\infty}{\sim} \dfrac{\pi}{2^n}$.
Après multiplication par 2^n, on a donc $u_n \underset{n\to+\infty}{\sim} \pi$, et donc $\lim\limits_{n\to+\infty} u_n = \pi$.

3. Appliquons la formule de Taylor-Lagrange à l'ordre 2 à la fonction sinus[2] entre 0 et $\dfrac{\pi}{2^n}$: la dérivée troisième de sin est la fonction − cos, qui est bornée par 1

$$\left|\sin\left(\frac{\pi}{2^n}\right) - \frac{\pi}{2^n}\right| \leqslant \frac{1\pi^3}{3!\,(2^n)^3}.$$

En multipliant par 2^n, il vient donc

$$|u_n - \pi| = \left|2^n \sin\left(\frac{\pi}{2^n}\right)\right| \leqslant \frac{\pi^3}{6\times 4^n}.$$

[2] Qui est bien de classe $\mathscr{C}^3$.

4. Notons que nous ne voulons pas utiliser la commande `%pi`, puisque notre but est de calculer une valeur approchée de π, sans utiliser la valeur que `Scilab` connaît déjà.
Pour cela, nous pouvons utiliser la majoration grossière $\pi \leqslant 4$.
Et alors, si n est tel que $\dfrac{4^3}{6\times 4^n} \leqslant 10^{-q}$, alors $|u_n - \pi| \leqslant 10^{-q}$.
Mais $\dfrac{4^3}{6\times 4^n} \leqslant 10^{-q} \Leftrightarrow 4^n \geqslant \dfrac{32\times 10^q}{3} \Leftrightarrow n \geqslant \dfrac{1}{\ln(4)}\left(q\ln(10) + \ln\left(\dfrac{32}{3}\right)\right)$.
Le programme suivant convient donc

```
function u = approx(q)
        u = 2 ;
        n=1 ;
        while n<(q*log(10)+log(32/3))/log(4)
                u = sqrt(2)*u/sqrt(1+sqrt(1-(u/2^n)^2)) ;
                n = n+1 ;
        end
endfunction
```

Remarque
Si on connaît un peu plus de décimales, on peut toujours dire que $\pi \leqslant 3.2$ ou encore $\pi \leqslant 3.15$.

5. Le développement limité de sinus à l'ordre $2p+2$ au voisinage de 0 de sin est

$$\sin(x) = x - \frac{x^3}{6} + \cdots + (-1)^p \frac{x^{2p+1}}{(2p+1)!} + \underset{x\to 0}{o}\left(x^{2p+2}\right) = \sum_{k=0}^{p} \frac{(-1)^k x^{2k+1}}{(2k+1)!} + \underset{x\to 0}{o}\left(x^{2p+2}\right).$$

Donc en particulier, en remplaçant x par $\dfrac{\pi}{2^n}$, qui tend bien vers 0 lorsque $n \to +\infty$, on a

$$\sin\left(\frac{\pi}{2^n}\right) = \sum_{k=0}^{p} (-1)^k \frac{\pi^{2k+1}}{(2k+1)!\times (2^n)^{2k+1}} + \underset{n\to+\infty}{o}\left(\frac{\pi^{2p+2}}{(2^n)^{2p+2}}\right).$$

Et donc après multiplication par 2^n,

$$u_n = \sum_{k=0}^{p} \frac{\pi^{2k+1}}{(2k+1)!\times 4^{nk}} + \underset{n\to+\infty}{o}\left(\frac{\pi^{2p+2}}{(2^n)^{2p+1}}\right).$$

On a donc bien le développement voulu,

$$u_n = \pi - \frac{\pi^3}{6!\times 4^n} + \cdots + (-1)^p \frac{\pi^{2p+1}}{(2p+1)!\times 4^{pn}} + \underset{n\to+\infty}{o}\left(\frac{\pi^{2p+2}}{2^{n(2p+1)}}\right)$$

à l'exception du reste.

Mais p étant fixé, la constante π^{2p+2} n'a pas d'importance, et donc $\underset{n\to+\infty}{o}\left(\dfrac{\pi^{2p+2}}{2^{n(2p+1)}}\right) = \underset{n\to+\infty}{o}\left(\dfrac{1}{2^{n(2p+1)}}\right)$.

Enfin, puisque $\dfrac{1}{2^{n(2p+1)}} = \underset{n\to+\infty}{o}\left(\dfrac{1}{4^n p}\right)$, une quantité négligeable devant $\dfrac{1}{2^{n(2p+1)}}$ est également négligeable devant $\dfrac{1}{4^{np}}$.
Et donc on a bien

$$u_n = \pi - \frac{\pi^3}{6\times 4^n} + \cdots + (-1)^p \frac{\pi^{2p+1}}{(2p+1)!\times 4^{pn}} + \underset{n\to+\infty}{o}\left(\frac{1}{4^{pn}}\right).$$

Remarque
Ce développement limité peut se retrouver par la formule de Taylor-Young, à condition d'être capable de retrouver les valeurs de $\sin^{(n)}(0)$, qui vaut

$$\begin{cases} (-1)^p & \text{si } n = 2p+1 \text{ est impair} \\ 0 & \text{si } n \text{ est pair} \end{cases}$$

6. D'après ce qui précède, on a

$$u_n = \pi - \frac{\pi^3}{6\times 4^n} + \underset{n\to+\infty}{o}\left(\frac{1}{4^n}\right)$$

et de même

$$4u_{n+1} = 4\pi - \frac{\pi^3}{6\times 4^n} + \underset{n\to+\infty}{o}\left(\frac{1}{4^n}\right).$$

Et donc en sommant ces deux expressions,

$$v_n = \pi + \underset{n\to+\infty}{o}\left(\frac{1}{4^n}\right).$$

Ceci prouve donc déjà que $\lim\limits_{n\to+\infty} v_n - \pi = 0$, et donc $\lim\limits_{n\to+\infty} v_n = \pi$.

Et comme d'autre part, $u_n - \pi = \frac{-\pi^3}{6\times 4^n} + \underset{n\to+\infty}{o}\left(\frac{1}{4^n}\right) \underset{n\to+\infty}{\sim} \frac{-\pi^3}{6\times 4^n}$, on en déduit donc que

$$v_n - \pi = \underset{n\to+\infty}{o}\left(\frac{1}{4^n}\right) = \underset{n\to+\infty}{o}\left(\frac{-\pi^3}{6\times 4^n}\right) = \underset{n\to+\infty}{o}(u_n - \pi).$$

Danger !

Les o ne se simplifient pas par addition :

$$o(u_n) - o(u_n) = o(u_n).$$

C'est là l'une des nombreuses subtilités de la notation o.

Rappel

Si deux suites u_n et v_n sont équivalentes, être négligeable devant u_n signifie la même chose qu'être négligeable devant v_n.

7. Il s'agit de pousser un ordre plus loin les développements limités utilisés à la question précédente :

$$u_n = \pi - \frac{\pi^3}{6\times 4^n} + \frac{\pi^5}{5!\times 4^{2n}} + \underset{n\to+\infty}{o}\left(\frac{1}{4^{2n}}\right)$$

et

$$4u_{n+1} = 4\pi - \frac{\pi^3}{6\times 4^n} + \frac{\pi^5}{5!\times 4^{2n+1}} + \underset{n\to+\infty}{o}\left(\frac{1}{4^{2n+1}}\right) = 4\pi - \frac{\pi^3}{6\times 4^n} + \underset{n\to+\infty}{o}\left(\frac{1}{4^{2n}}\right).$$

Et donc

$$v_n = \pi - \frac{\pi^5}{3\times 5!\times 4^{2n+1}} + \underset{n\to+\infty}{o}\left(\frac{1}{4^{2n}}\right), \text{ de sorte que } v_n - \pi \underset{n\to+\infty}{\sim} -\frac{\pi^5}{3\times 5!\times 4^{2n}}.$$

Quelques commentaires :

□

EXERCICE 2.11 (ESCP 2018) [ESCP18.1. ? ?] — Facile

Développement asymptotique d'une suite définie par récurrence
Abordable en première année : ✓

Soit un réel $x_0 > 0$. On définit la suite (x_n) par : $\forall n \in \mathbf{N}$, $x_{n+1} = x_n + \frac{1}{x_n}$.

On admet que $\sum\limits_{k=1}^{n} \frac{1}{k} \underset{n\to+\infty}{\sim} \ln(n)$.

1. Montrer que la suite (x_n) est monotone et préciser sa monotonie.
2. Déterminer $\lim\limits_{n\to+\infty} x_n$.
3. Déterminer la nature de la série de terme général $\frac{1}{x_n}$.
4. Montrer que la suite $(x_{n+1}^2 - x_n^2)$ converge.
5. Soit (a_n) et (b_n) des suites de réels strictement positifs tels que $a_n \underset{n\to+\infty}{\sim} b_n$ et $\sum\limits_n a_n$ diverge.

   On admet qu'alors $\sum\limits_{k=0}^{n} a_k \underset{n\to+\infty}{\sim} \sum\limits_{k=0}^{n} b_k$.

   (a) Montrer que la suite $\left(\frac{x_n^2}{n}\right)_{n\in\mathbf{N}^*}$ converge et en déduire un équivalent de (x_n).

(b) Déterminer un équivalent de $\sum_{k=0}^{n} \frac{1}{x_k^2}$.

(c) En déduire que $x_n = \sqrt{2n}\left(1 + \frac{1}{8}\frac{\ln n}{n} + \underset{n\to+\infty}{o}\left(\frac{\ln n}{n}\right)\right)$.

Solution de l'exercice 2.11

1. Une récurrence immédiate prouve que pour tout $n \in \mathbf{N}$, $x_n \geqslant 0$, et donc $x_{n+1} - x_n > 0$, de sorte que (x_n) est strictement croissante.

2. Par le théorème de la limite monotone[1], (x_n) converge vers une limite finie ℓ ou tend vers $+\infty$.
Supposons que $x_n \underset{n\to+\infty}{\longrightarrow} \ell$. Alors on a également $x_{n+1} \underset{n\to+\infty}{\longrightarrow} \ell$.
Et donc $\ell = \lim_{n\to+\infty} x_{n+1} = \lim_{n\to+\infty} x_n + \frac{1}{x_n} = \ell + \frac{1}{\ell}$.
Or l'équation $\ell = \ell + \frac{1}{\ell}$ ne possède pas de solution, et donc $x_n \underset{n\to+\infty}{\longrightarrow} +\infty$.

[1] (x_n) est minorée par 0 et croissante.

3. Notons que $\frac{1}{x_n} = x_{n+1} - x_n$. Et donc

$$\sum_{k=0}^{n} \frac{1}{x_k} = \sum_{k=0}^{n} (x_{k+1} - x_k) = x_{n+1} - x_0 \underset{n\to+\infty}{\longrightarrow} +\infty.$$

Et donc la série de terme général $\frac{1}{x_n}$ diverge.

4. On a $x_{n+1}^2 - x_n^2 = (x_{n+1} + x_n)(x_{n+1} - x_n) = \frac{x_{n+1} + x_n}{x_n} = 1 + \frac{x_{n+1}}{x_n} = 2 + \frac{1}{x_n^2} \underset{n\to+\infty}{\longrightarrow} 2$.

> **Remarque**
> On a toujours
> $$u_n \to \ell \iff u_n \sim \ell,$$
> sauf si $\ell = 0$!

5.a. Nous venons de prouver que $x_{n+1}^2 - x_n^2 \underset{n\to+\infty}{\sim} 2$.
Or la série de terme général 2 diverge grossièrement.
Donc par le résultat admis, les suites des sommes partielles des séries de terme général $x_{n+1}^2 - x_n^2$ et 2 sont équivalentes.
Autrement dit,

$$\sum_{k=0}^{n-1} (x_{k+1}^2 - x_k^2) \underset{n\to+\infty}{\sim} \sum_{k=0}^{n-1} 2 \Leftrightarrow x_n^2 - x_0^2 \underset{n\to+\infty}{\sim} 2n.$$

Puisque $x_n^2 \underset{n\to+\infty}{\longrightarrow} +\infty$, $x_0^2 = \underset{n\to+\infty}{o}(x_n^2)$ et donc $x_n^2 \underset{n\to+\infty}{\sim} 2n$.
Et donc $\frac{x_n^2}{n} \underset{n\to+\infty}{\longrightarrow} 2$.
On en déduit que $x_n \underset{n\to+\infty}{\sim} \sqrt{2n}$.

5.b. Puisque $\frac{1}{x_n^2} \underset{n\to+\infty}{\sim} \frac{1}{2n}$, et que la série de terme général $\frac{1}{2n}$ diverge, on a donc[2]

$$\sum_{k=0}^{n} \frac{1}{x_k^2} \underset{n\to+\infty}{\sim} \sum_{k=1}^{n} \frac{1}{2k} \underset{n\to+\infty}{\sim} \frac{1}{2}\ln(n).$$

[2] Toujours grâce au résultat admis.

> **Bornes**
> N'y aurait-il pas une petite arnaque sur les bornes de la somme ?
> Ceci afin d'éviter d'écrire $\sum_{k=0}^{n} \frac{1}{k} \cdots$
> En réalité, puisque la série diverge, et est à termes positifs, la suite de ses sommes partielles tend vers $+\infty$, et donc la somme commençant à $k = 0$ et la somme commençant à $k = 1$ sont équivalentes.

5.c. On a $x_n^2 = \sum_{k=0}^{n-1} (x_{k+1}^2 - x_k^2) + x_0^2 = x_0^2 + \sum_{k=0}^{n-1} \left(2 + \frac{1}{x_k^2}\right) = x_0^2 + 2n + \sum_{k=0}^{n-1} \frac{1}{x_k^2}$.
Et donc en utilisant le résultat de la question précédente,

$$x_n^2 = x_0^2 + 2n + \frac{\ln n}{2} + \underset{n\to+\infty}{o}(\ln n) = 2n + \frac{\ln n}{2} + \underset{n\to+\infty}{o}(\ln n).$$

En particulier, $\frac{x_n^2}{2n} = 1 + \frac{\ln n}{4n} + \underset{n\to+\infty}{o}\left(\frac{\ln n}{2n}\right)$.
Utilisons alors le développement limité de $\sqrt{1+x}$ en 0 : $\sqrt{1+x} = 1 + \frac{x}{2} + \underset{x\to0}{o}(x)$.
Il vient donc

$$\frac{x_n}{\sqrt{2n}} = \sqrt{\frac{x_n}{2n}} = \sqrt{1 + \frac{\ln n}{4n} + \underset{n\to+\infty}{o}\left(\frac{\ln n}{2n}\right)} = 1 + \frac{\ln n}{8n} + \underset{n\to+\infty}{o}\left(\frac{\ln n}{n}\right).$$

Et donc on a bien le résultat voulu.

□

EXERCICE 2.12 (ESCP 2014) [ESCP14.1.04] — Moyen

Sommation de Césaro

Soit $(u_n)_{n\geqslant 0}$ une suite réelle. On définit la suite $(s_n)_{n\geqslant 0}$ en posant :

$$\forall n \in \mathbf{N},\ s_n = \frac{u_0 + u_1 + \cdots + u_n}{n+1}.$$

1. On suppose dans cette question que la suite $(u_n)_{n\geqslant 0}$ est décroissante et tend vers $\ell \in \mathbf{R}$.
   (a) Si $x \in \mathbf{R}$, on note $\lfloor x \rfloor$ la partie entière de x. Vérifier que la suite $\left(\frac{n-\lfloor\sqrt{n}\rfloor}{n+1}\right)_{n\geqslant 0}$ converge vers 1 lorsque n tend vers $+\infty$.
   (b) Montrer que
   $$\ell \leqslant s_n \leqslant \frac{\lfloor\sqrt{n}\rfloor+1}{n+1}u_0 + \frac{n-\lfloor\sqrt{n}\rfloor}{n+1}u_{\lfloor\sqrt{n}\rfloor}.$$
   (c) Montrer que la suite $(s_n)_{n\geqslant 0}$ converge vers ℓ.
   (d) Réciproquement, on suppose que la suite $(s_n)_{n\geqslant 0}$ converge aussi vers $\ell \in \mathbf{R}$. Montrer que la suite $(u_n)_{n\geqslant 0}$ converge aussi vers ℓ (on pourra raisonner par l'absurde).
2. Dans cette question, on suppose que la suite $(u_n)_{n\geqslant 0}$ converge vers 0.
   (a) Montrer que $v_n = \sup\{|u_k|,\ k \geqslant n\}$ est bien défini.
   (b) Montrer que la suite $(v_n)_{n\geqslant 0}$ est décroissante et de limite nulle.
   (c) En déduire que la suite $(s_n)_{n\geqslant 0}$ converge vers 0.
3. On suppose maintenant que la suite $(u_n)_{n\geqslant 0}$ converge vers $\ell \in \mathbf{R}$. Montrer que la suite des moyennes $(s_n)_{n\geqslant 0}$ converge aussi vers ℓ. Donner un exemple simple prouvant que la réciproque est fausse.
4. On considère la suite $(w_n)_{n\geqslant 0}$ définie par récurrence en posant :
   $$w_0 = 1 \text{ et } w_{n+1} = w_n + e^{-w_n}.$$
   (a) Étudier la suite $(w_n)_{n\geqslant 0}$.
   (b) Prouver que : $\lim\limits_{n\to+\infty} (e^{w_{n+1}} - e^{w_n}) = 1$.
   (c) En déduire un équivalent de w_n lorsque n tend vers l'infini.

SOLUTION DE L'EXERCICE 2.12

1.a. Puisque $0 \leqslant \lfloor\sqrt{n}\rfloor \leqslant \sqrt{n}$, en divisant par n et en passant à la limite, on obtient, par le théorème des gendarmes, $\lim\limits_{n\to+\infty} \frac{\lfloor\sqrt{n}\rfloor}{n} = 0$, soit $\lfloor\sqrt{n}\rfloor = \underset{n\to+\infty}{o}(n)$.
Et donc $n - \lfloor\sqrt{n}\rfloor = n + \underset{n\to+\infty}{o}(n) \underset{n\to+\infty}{\sim} n$.
Ainsi, $\frac{n-\lfloor\sqrt{n}\rfloor}{n+1} \underset{n\to+\infty}{\sim} \frac{n}{n} \underset{n\to+\infty}{\longrightarrow} 1$.

1.b. Puisque (u_n) est décroissante de limite ℓ, pour tout $k \in \mathbf{N}$, $u_k \geqslant \ell$.
Et donc $s_n \geqslant \frac{\ell+\ell+\cdots+\ell}{n+1} = \frac{(n+1)\ell}{n+1} = \ell$.
De plus, pour $k \in [\![\lfloor\sqrt{n}\rfloor, n]\!]$, $u_k \geqslant u_{\lfloor\sqrt{n}\rfloor}$, et pour $k \in [\![0, \lfloor\sqrt{n}\rfloor - 1]\!]$, $u_k \geqslant u_0$. Et donc

$$s_n = \frac{\sum\limits_{k=0}^{\lfloor\sqrt{n}\rfloor-1} u_k + \sum\limits_{k=\lfloor\sqrt{n}\rfloor}^{n} u_k}{n+1} \leqslant \frac{\sum\limits_{k=0}^{\lfloor\sqrt{n}\rfloor-1} u_0 + \sum\limits_{k=\lfloor\sqrt{n}\rfloor}^{n} u_{\lfloor\sqrt{n}\rfloor}}{n+1} \leqslant \frac{\lfloor\sqrt{n}\rfloor}{n+1}u_0 + \frac{n-\lfloor\sqrt{n}\rfloor}{n+1}u_{\lfloor\sqrt{n}\rfloor}.$$

1.c. Lorsque $n \to +\infty$, $\lfloor\sqrt{n}\rfloor \to +\infty$ et donc $u_{\lfloor\sqrt{n}\rfloor} \underset{n\to+\infty}{\longrightarrow} \ell$.
On en déduit que $\frac{n-\lfloor\sqrt{n}\rfloor}{n+1}u_{\lfloor\sqrt{n}\rfloor} \underset{n\to+\infty}{\longrightarrow} \ell$.

Puisque d'autre part $\lfloor\sqrt{n}\rfloor = \underset{n\to+\infty}{o}(n)$, il vient $\dfrac{\lfloor\sqrt{n}\rfloor}{n+1} \underset{n\to+\infty}{\longrightarrow} 0$.
Et donc grâce à l'encadrement de la question précédente et au théorème des gendarmes, $s_n \underset{n\to+\infty}{\longrightarrow} \ell$.

1.d. Puisque $(u_n)_{n\geqslant 0}$ est décroissante, soit elle converge vers une limite a, soit elle tend vers $-\infty$.

Supposons qu'elle tende vers $-\infty$. Alors $\left(u_{\lfloor\sqrt{n}\rfloor}\right)$ tend aussi vers $-\infty$, de sorte que $\dfrac{n-\lfloor\sqrt{n}\rfloor}{n+1}u_{\lfloor\sqrt{n}\rfloor} \underset{n\to+\infty}{\longrightarrow} -\infty$.
Et donc la majoration de la question précédente prouve que $s_n \underset{n\to+\infty}{\longrightarrow} -\infty$ contredisant l'hypothèse qui est faite ici.
Donc (u_n) converge vers une limite a. Et alors la question précédente prouve que $s_n \underset{n\to+\infty}{\longrightarrow} a$, de sorte que par unicité de la limite, $a = \ell$, et donc $\lim\limits_{n\to+\infty} u_n = \ell$.

2.a. Nous savons que toute suite convergente est bornée. Et donc l'ensemble $\{|u_k|, k \geqslant n\}$ est borné. Toute partie bornée de $\mathbf{R}$ admet une borne supérieure, de sorte que v_n est bien défini.

2.b. Puisque $\{|u_k| n \geqslant k+1\} \subset \{|u_k|,\ k \geqslant n\}$, en passant aux bornes supérieures, $v_{n+1} \leqslant v_n$, et donc (v_n) est décroissante.

Soit $\varepsilon > 0$. Alors il existe $N \in \mathbf{N}$ tel que $n \geqslant N \Rightarrow |u_n| \leqslant \varepsilon$.
Et donc en particulier, pour $n \geqslant N$, et pour tout $k \geqslant n$, $k \geqslant N$ de sorte que $|u_k| \leqslant \varepsilon$.
Ceci étant vrai pour tout k, on en déduit que $0 \leqslant v_n \leqslant \varepsilon$.
Et donc on a bien prouvé que $v_n \underset{n\to+\infty}{\longrightarrow} 0$.

2.c. Notons (s'_n) la suite définie par $s'_n = \dfrac{v_0 + v_1 + \cdots + v_n}{n+1}$.
Puisque (v_n) est décroissante et de limite nulle, les résultats de la question 1 s'appliquent, et donc $s'_n \underset{n\to+\infty}{\longrightarrow} 0$.
Mais pour tout n, on a

$$0 \leqslant |s_n| \leqslant \frac{|u_0| + \cdots + |u_n|}{n+1} \leqslant \frac{v_0 + v_1 + \cdots + v_n}{n+1} \leqslant s'_n.$$

Et donc par le théorème des gendarmes, $\lim\limits_{n\to+\infty} |s_n| = 0 \Leftrightarrow \lim\limits_{n\to+\infty} s_n = 0$.

3. Posons pour tout $n \in \mathbf{N}$, $u'_n = u_n - \ell$, de sorte que $u'_n \underset{n\to+\infty}{\longrightarrow} 0$.
D'après la question 2, on a donc

$$\lim_{n\to+\infty} \frac{u'_0 + \cdots + u'_n}{n+1} = 0.$$

Or, $\dfrac{u'_0 + u'_1 + \cdots + u'_n}{n+1} = \dfrac{u_0-\ell+u_1-\ell+\cdots+u_n-\ell}{n+1} = \dfrac{u_0 + \cdots + u_n}{n+1} - \ell = s_n - \ell$.
Et donc $s_n - \ell \underset{n\to+\infty}{\longrightarrow} 0 \Leftrightarrow s_n \underset{n\to+\infty}{\longrightarrow} \ell$.

La réciproque est en revanche fausse, comme le montre le cas de la suite définie par $u_n = (-1)^n$.
En effet, on a alors $u_0 + u_1 + \cdots + u_n = \begin{cases} 1 & \text{si } n \text{ est pair} \\ 0 & \text{si } n \text{ est impair} \end{cases}$.
Et donc $|u_0 + \cdots + u_n| \leqslant 1$, de sorte que $|s_n| \leqslant \dfrac{1}{n+1}$ et donc $s_n \underset{n\to+\infty}{\longrightarrow} 0$ alors que (u_n) n'a pas de limite.

4.a. Puisque pour tout n, $e^{-w_n} \geqslant 0$, $w_{n+1} \geqslant w_n$, et donc (w_n) est croissante.
Par le théorème de la limite monotone, soit elle converge vers $\ell \geqslant 1$, soit elle tend vers $+\infty$.
Si elle tendait vers ℓ, puisque $w_{n+1} \underset{n\to+\infty}{\longrightarrow} \ell$, par unicité de la limite, il viendrait $\ell = \ell + e^{-\ell}$, ce qui est impossible.
Donc $w_n \underset{n\to+\infty}{\longrightarrow} +\infty$.

4.b. On a $e^{w_{n+1}} - e^{w_n} = e^{w_n + e^{-w_n}} - e^{w_n} = e^{w_n}\left(e^{e^{-w_n}} - 1\right)$.
Or, $e^{-w_n} \underset{n\to+\infty}{\longrightarrow} 0$, de sorte que

$$e^{e^{-w_n}} - 1 \underset{n\to+\infty}{\sim} e^{-w_n}$$

et donc $e^{w_{n+1}} - e^{w_n} \underset{n\to+\infty}{\sim} e^{w_n}e^{-w_n} = 1$.

4.c. Appliquons le résultat de la question 3 à la suite $e^{w_{n+1}} - e^{w_n}$:

$$\frac{e^{w_1} - e^{w_0} + e^{w_2} - e^{w_1} + \cdots + e^{w_n} - e^{w_{n-1}}}{n} \underset{n\to+\infty}{\longrightarrow} 1.$$

Et donc $\frac{e^{w_n}-e^{w_0}}{n} \underset{n\to+\infty}{\longrightarrow} 1$.
Soit encore $e^{w_n} \underset{n\to+\infty}{\sim} n \Leftrightarrow e^{w_n} = n + \underset{n\to+\infty}{o}(n)$.
Et donc

$$w_n = \ln(e^{w_n}) = \ln(n + \underset{n\to+\infty}{o}(n)) = \ln\left(n\left(1 + \underset{n\to+\infty}{o}(1)\right)\right) = \ln(n) + \ln\left(1 + \underset{n\to+\infty}{o}(1)\right).$$

Il vient donc enfin

$$\frac{w_n}{\ln n} = 1 + \underset{n\to+\infty}{o}\left(\frac{1}{\ln(n)}\right) \underset{n\to+\infty}{\longrightarrow} 1.$$

Ceci prouve donc que $w_n \underset{n\to+\infty}{\sim} \ln(n)$.

□

2.2.1 A classer

ESCP1.6 2010 : sommation de Cesaro et des suites récurrentes $u_{n+1} = f(u_n)$.

EXERCICE 2.13 (QSP ESCP 2006) [ESCP06.QSP01] **Moyen**

Développement asymptotique d'une suite définie par une équation.

Montrer que pour tout $n \in \mathbf{N}^*$, il existe un unique $y_n \in \mathbf{R}_+^*$ tel que $\ln(y_n) + y_n = \frac{1}{n}$.
Montrer que la suite $(y_n)_{n\in\mathbf{N}^*}$ converge et si on note ℓ sa limite, déterminer un équivalent de $y_n - \ell$.

SOLUTION DE L'EXERCICE 2.13 La fonction $f : x \mapsto \ln(x) + x$ est continue et strictement croissante sur $\mathbf{R}_+^*$.
Comme on a de plus $\lim_{x\to0^+} f(x) = -\infty$ et $\lim_{x\to+\infty} f(x) = +\infty$, 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 $\frac{1}{n} \leqslant 1 = f(1)$, 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

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

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

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

$$= \frac{1}{n} - \frac{1}{\ell}(y_n - \ell) + \underset{n\to+\infty}{o}(y_n - \ell).$$

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

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

□

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.

Exercice 2.14 (QSP HEC 2012) [HEC12.QSP42] — Moyen

Étude d'une suite définie par le produit d'un nombre variable de termes.

Soit $\alpha \in \mathbf{R}_+$. Pour tout $n \in \mathbf{N}^*$, on pose $u_n = \prod_{k=1}^{n}\left(1 + \frac{k^\alpha}{n^2}\right)$.

1. Montrer que si $\alpha = 2$, la suite $(u_n)_{n\geqslant 1}$ est divergente.
2. Montrer que si $0 \leqslant \alpha < 1$, la suite $(u_n)_{n\geqslant 1}$ converge vers 1.

Solution de l'exercice 2.14

1. Puisque nous préférons manipuler des sommes que des produits, commençons par passer au logarithme.

On a $\ln(u_n) = \sum_{k=1}^{n} \ln\left(1 + \frac{k^\alpha}{n^2}\right)$.

En particulier, pour $\alpha = 2$, on a

$$\ln(u_n) = \sum_{k=1}^{n} \ln\left(1 + \left(\frac{k}{n}\right)^2\right) = n \times \frac{1}{n}\sum_{k=1}^{n} \ln\left(1 + \left(\frac{k}{n}\right)^2\right).$$

Or, nous reconnaissons là des sommes de Riemann :

$$\frac{1}{n}\sum_{k=1}^{n} \ln\left(1 + \left(\frac{k}{n}\right)^2\right) \underset{n\to+\infty}{\longrightarrow} \int_0^1 \ln\left(1 + x^2\right) dx.$$

Nous n'avons pas besoin de calculer cette intégrale[1], juste de noter qu'elle est strictement positive, car intégrale d'une fonction continue et strictement positive.

[1] Ce qui pourrait se faire par une intégration par parties.

Et donc $\ln(u_n) \underset{n\to+\infty}{\sim} n\int_0^1 \ln(1+x^2)\,dx \underset{n\to+\infty}{\longrightarrow} +\infty$.

On en déduit par passage à l'exponentielle que $u_n \underset{n\to+\infty}{\longrightarrow} +\infty$.

2. Il s'agit donc ici de prouver que $\ln(u_n) \underset{n\to+\infty}{\longrightarrow} 0$.

Nous savons déjà que $\ln(u_n) \geqslant 0$, car somme de termes positifs.

Utilisons à présent l'inégalité de concavité classique : $\forall x \geqslant 0, \ln(1+x) \leqslant x$.

Il vient alors

$$\ln(u_n) = \sum_{k=1}^{n} \ln\left(1 + \frac{k^\alpha}{n^2}\right) \leqslant \sum_{k=1}^{n} \frac{k^\alpha}{n^2} \leqslant \sum_{k=1}^{n} \frac{n^\alpha}{n^2} \leqslant \frac{n^\alpha}{n}.$$

Puisque $0 \leqslant \alpha < 1$, $\frac{n^\alpha}{n} \underset{n\to+\infty}{\longrightarrow} 0$, de sorte que par le théorème des gendarmes, $\ln(u_n) \underset{n\to+\infty}{\longrightarrow} 0$.

Et donc par continuité de l'exponentielle, $u_n \underset{n\to+\infty}{\longrightarrow} e^0 = 1$.

□

2.3 Séries numériques

EXERCICE 2.15 (QSP ESCP 2007) [ESCP07.QSP01] — Facile

Série géométrique dérivée $p^{\text{ème}}$.
Abordable en première année : ✓

Soit $p \in \mathbf{N}$ fixé. Déterminer la nature de la série de terme général $\dfrac{n^p}{2^n}$.

SOLUTION DE L'EXERCICE 2.15 Commençons par noter que pour $p \in \{0,1,2\}$, nous avons affaire à une série géométrique ou géométrique dérivée de raison $\dfrac{1}{2}$, qui est donc convergente.
De manière générale, on a

$$n^2\frac{n^p}{2^n} = \frac{n^{p+2}}{2^n} \underset{n\to+\infty}{\longrightarrow} 0$$

par croissances comparées, de sorte que $\dfrac{n^p}{2^n} \underset{n\to+\infty}{=} o\left(\dfrac{1}{n^2}\right)$.

Et donc, par comparaison à une série de Riemann, la série de terme général $\dfrac{n^p}{2^n}$ converge. □

EXERCICE 2.16 (QSP ESCP 2008) [ESCP08.QSP01] — Moyen

Nature de la série $\sum \left(1+\dfrac{1}{n}\right)^{2n} - \left(1+\dfrac{2}{n}\right)^n$
Abordable en première année : ✓

Déterminer la nature de la série de terme général $u_n = \left(1+\dfrac{1}{n}\right)^{2n} - \left(1+\dfrac{2}{n}\right)^n$.

SOLUTION DE L'EXERCICE 2.16 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 $n\ln\left(1+\dfrac{2}{n}\right) \underset{n\to+\infty}{\sim} 2 \underset{n\to+\infty}{\longrightarrow} 2$ de sorte que, par continuité de l'exponentielle,

$$\left(1+\frac{2}{n}\right)^n \underset{n\to+\infty}{\longrightarrow} e^2.$$

D'autre part,

$$2n\ln\left(1+\frac{1}{n}\right) - n\ln\left(1+\frac{2}{n}\right) = 2n\left(\frac{1}{n} - \frac{1}{2n^2} + o\left(\frac{1}{n^2}\right)\right) - n\left(\frac{2}{n} - \frac{2}{n^2} + o\left(\frac{1}{n^2}\right)\right) = \frac{1}{n} + o\left(\frac{1}{n}\right) \underset{n\to+\infty}{\longrightarrow} 0.$$

Et donc, à l'aide de l'équivalent classique $e^x - 1 \underset{x\to0}{\sim} x$,

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

On en déduit que

$$u_n \underset{n\to+\infty}{\sim} \frac{e^2}{n}.$$

Et donc, d'après le critères 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 $u_n \geqslant 0$, mais puisque $\dfrac{e^{-2}}{n} \geqslant 0$, le critère des équivalents s'applique tout de même.

Exercice 2.17 (QSP HEC 2015) [HEC15.QSP151] **Moyen**

Nature de la série de terme général $\dfrac{u_{n+1}}{\sum_{k=1}^{n} u_k}$ en fonction de la nature de $\sum u_n$.

Soit $(u_n)_{n\in\mathbf{N}}$ une série convergente à termes positifs et de limite nulle. On pose pour tout $n \in \mathbf{N}$, $S_n = \sum_{k=0}^{n} u_k$ et $v_n = \dfrac{u_{n+1}}{S_n}$.

1. Étudier la nature de la série $\sum_{k\geqslant 0} v_k$ en fonction de la nature de la série $\sum_{k\geqslant 0} u_k$.
   Indication : dans le cas où $\sum_k u_k$ diverge, on pourra commencer par étudier $\sum_k \ln(1+v_k)$.
2. Quel résultat obtient-on dans le cas où $u_n = \dfrac{1}{n}$?

Solution de l'exercice 2.17

1. Commençons par le cas où $\sum_{k\geqslant 0} u_k$ converge.

Puisque $(u_n)_{n\geqslant 0}$ est positive, $(S_n)_{n\geqslant 0}$ est croissante et donc pour tout $n \in \mathbf{N}$, $S_n \geqslant S_0 = u_0$.
On en déduit que $0 \leqslant v_n \leqslant \dfrac{u_{n+1}}{u_0}$.
Puisque la série de terme général u_{n+1} est convergente, on en déduit que $\sum_{n\geqslant 0} v_n$ converge.

Supposons à présent que $\sum_{k\geqslant 0} u_k$ diverge.
Puisqu'il s'agit d'une série à termes positifs, la suite $(S_n)_{n\geqslant 0}$ de ses sommes partielles tend vers $+\infty$.
Et alors $v_n \underset{n\to+\infty}{\longrightarrow} 0$.
Pour $N \in \mathbf{N}$, on a

$$\begin{aligned}
\sum_{k=0}^{N} \ln(1+v_k) &= \sum_{k=0}^{N} \ln\left(1+\frac{u_{k+1}}{S_k}\right)\\
&= \sum_{k=0}^{N} \ln\left(\frac{u_{k+1}+S_k}{S_k}\right)\\
&= \sum_{k=0}^{N} \ln\left(\frac{S_{k+1}}{S_k}\right)\\
&= \sum_{k=0}^{N} \ln(S_{k+1}) - \ln(S_k)\\
&= \ln(S_{N+1}) - \ln(S_0) \qquad \text{Somme télescopique.}\\
&\underset{N\to+\infty}{\longrightarrow} +\infty.
\end{aligned}$$

On en déduit donc que la série de terme général $\ln(1+v_k)$ diverge.
Mais puisque $v_k \underset{k\to+\infty}{\longrightarrow} 0$, $\ln(1+v_k) \underset{k\to+\infty}{\sim} v_k$.
Et donc par critère de comparaison pour les séries à termes positifs, $\sum_{k\geqslant 0} v_k$ diverge.

En conclusion, les séries $\sum_{k\geqslant 0} u_k$ et $\sum_{k\geqslant 0} v_k$ sont de même nature.

2. Dans le cas où $u_n = \dfrac{1}{n}$, il est classique[1] $S_n = \sum_{k=1}^{n} \dfrac{1}{k} \underset{n\to+\infty}{\sim} \ln(n)$.

Et donc $v_n = \dfrac{1}{(n+1)S_n} \underset{n\to+\infty}{\sim} \dfrac{1}{n\ln n}$.

[1] Ce résultat s'obtient par exemple à l'aide d'une comparaison série/intégrale.

Puisque $\sum_{n\geqslant 1} u_n$ diverge, il en est de même de $\sum_{n\geqslant 2} \frac{1}{n\ln n}$.

Cette série est un cas particulier de ce qu'on appelle série de Bertrand.
Leur convergence peut également s'étudier à l'aide d'une comparaison série/intégrale (voir l'exercice 2.3.1).

□

Nous ne savons calculer que peu de sommes de séries : seules les séries géométriques, géométriques dérivées et exponentielles figurent au programme.
Si on nous demande de calculer la somme d'une série qui n'est pas de cette forme, cela passera quasi-systématiquement par un calcul de sommes partielles suivi d'un passage à la limite.
Sans indications supplémentaires pour le calcul des sommes partielles, c'est souvent que nous aurons affaire à une série télescopique (c'est-à-dire dont les termes successifs se «compensent»).

Exercice 2.18 (QSP HEC 2014) [HEC14.QSP61] — Moyen

Nature de la série de terme général $\ln(n) + a\ln(n+1) + b\ln(n+2)$.
Abordable en première année : ✓

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

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.

Solution de l'exercice 2.18

1. On a

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

Si $1+a+b \neq 0$, alors $u_n \underset{n\to+\infty}{\sim} (1+a+b)\ln(n)$ 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 $u_n \underset{n\to+\infty}{\sim} \frac{a+2b}{n}$ et donc[1] $\sum u_n$ diverge.
Il faut donc avoir de plus $a+2b=0$.
Or, $\begin{cases} a+b=-1 \\ a+2b=0 \end{cases} \Leftrightarrow \begin{cases} a=-2 \\ b=1 \end{cases}$
Et alors, on a $u_n \underset{n\to+\infty}{\sim} \frac{-1}{n^2}$, de sorte[2] que $\sum u_n$ converge.

[1] Par comparaison à une série de Riemann.

[2] Toujours par comparaison à une série de Riemann.

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

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

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

□

Exercice 2.19 (ESCP 2016) [ESCP16.1.10] Facile

Développement d'une intégrale sous forme d'une série.
Abordable en première année : ✓

Dans tout l'exercice, a est un réel fixé supérieur ou égal à 1. On note, sous réserve d'existence :

$$\forall n \in \mathbf{N},\ I_n = \int_0^{+\infty} e^{-at} t^n\, dt \text{ et } I = \int_0^{+\infty} e^{-at} \frac{\sin t}{t}\, dt.$$

1. (a) Prouver que pour tout entier n, l'intégrale définissant I_n est convergente.
   (b) Établir, pour tout $n \geqslant 1$, une relation de récurrence entre I_n et I_{n-1}. En déduire I_n en fonction de n et de a.
2. Démontrer que l'intégrale définissant I est absolument convergente.
3. (a) Soit x un réel positif. Montrer que pour tout entier naturel n,
$$\left|\sin x - \left(x - \frac{x^3}{6} + \cdots + (-1)^n \frac{x^{2n+1}}{(2n+1)!}\right)\right| \leqslant \frac{x^{2n+2}}{(2n+2)!}.$$
   (b) En déduire qu'il existe un réel K_n dépendant de n tel que :
$$\left|I - \sum_{k=0}^{n} (-1)^k \frac{I_{2k}}{(2k+1)!}\right| \leqslant K_n I_{2n+1}.$$
   (c) En déduire que la série $\displaystyle\sum_{k\geqslant 0} \frac{1}{(2k+1)a^{2k+1}}$ est convergente de somme I.
4. (a) Prouver que pour tout entier $n \in \mathbf{N}^*$ et tout réel $t \in [0,1]$:
$$\sum_{k=0}^{n} (-1)^k t^{2k} - \frac{1}{1+t^2} = (-1)^n \frac{t^{2n+2}}{1+t^2}.$$
   (b) En déduire que : $\forall n \in \mathbf{N}^*, \forall x \in [0,1], \left|\displaystyle\sum_{k=0}^{n} (-1)^k \frac{x^{2k+1}}{2k+1} - \operatorname{Arctan}(x)\right| \leqslant \frac{1}{2n+3}$.
   (c) En déduire une expression de I en fonction de a.

Solution de l'exercice 2.19

1.a. La fonction $t \mapsto e^{-at}t^n$ est continue sur $[0,+\infty[$, donc le seul éventuel problème de convergence se situe au voisinage de $+\infty$.
Or, au voisinage de $+\infty$, on a $t^2 e^{-at} t^n = t^{n+2}e^{-at} \underset{t\to+\infty}{\longrightarrow} 0$ par croissances comparées, de sorte que $t^n e^{-at} = \underset{t\to+\infty}{o}\left(\frac{1}{t^2}\right)$.
Puisque $\displaystyle\int_1^{+\infty} \frac{dt}{t^2}$ converge, il en est de même de $\displaystyle\int_1^{+\infty} t^n e^{-at}\, dt$ et donc de $\displaystyle\int_0^{+\infty} t^n e^{-at}\, dt$.

> **Détails**
> Puisque l'intégrande est continue en 0, l'intégrale entre 0 et 1 est l'intégrale d'une fonction continue sur un segment, donc est automatiquement convergente, il n'y a aucune étude à faire.

1.b. Soit $A > 0$. Posons alors $u(t) = t^{n+1}$ et $v(t) = -\frac{1}{a}e^{-at}$, qui sont deux fonctions de classe $\mathscr{C}^1$ sur le segment $[0,A]$, avec $u'(t) = (n+1)t^n$ et $v'(t) = e^{-at}$. Alors

$$\begin{aligned}\int_0^A t^{n+1}e^{-at}\, dt &= \left[-\frac{1}{a}e^{-at}t^{n+1}\right]_0^A + \frac{n+1}{a}\int_0^A t^n e^{-at}\\ &= -A^{n+1}\frac{e^{-aA}}{a} + \frac{n+1}{a}\int_0^A t^n e^{-at}\, dt\\ &\underset{A\to+\infty}{\longrightarrow} \frac{n+1}{a}\int_0^{+\infty} t^n e^{-at}\, dt.\end{aligned}$$

Et donc $I_{n+1} = \lim\limits_{A\to+\infty} \int_0^A t^n e^{-at}\,dt = \dfrac{n+1}{a} I_n$.
D'autre part, on a

$$I_0 = \lim_{A\to+\infty} \int_0^A e^{-at}\,dt = \lim_{A\to+\infty} \left[-\frac{1}{a}e^{-at}\right]_0^A = \lim_{A\to+\infty} \left(\frac{1}{a} - \frac{1}{a}e^{-aA}\right) = \frac{1}{a}.$$

Ainsi, $I_1 = \dfrac{1}{a}I_0 = \dfrac{1}{a^2}$, $I_2 = \dfrac{2}{a}I_1 = \dfrac{2}{a^3}$, $I_3 = \dfrac{3}{a}I_2 = \dfrac{3!}{a^4}$, et une récurrence facile prouve que pour tout $n \in \mathbf{N}$, $I_n = \dfrac{n!}{a^{n+1}}$.

Remarque

La valeur de I_n s'obtient bien plus facilement si on connait la fonction Γ, en procédant au changement de variable $x = at$ dans l'intégrale I_n.

2. La fonction $t \mapsto e^{-at}\dfrac{\sin t}{t}$ est continue sur $]0;+\infty[$.
Pour $t \in [1,+\infty[$, on a $\left|e^{-at}\dfrac{\sin t}{t}\right| \leqslant \dfrac{e^{-at}}{t}$.
Or, $t^2\dfrac{e^{-at}}{t} = te^{-at} \underset{t\to+\infty}{\longrightarrow} 0$, de sorte que $\dfrac{e^{-at}}{t} \underset{t\to+\infty}{=} o\left(\dfrac{1}{t^2}\right)$.
Puisque $\int_1^{+\infty} \dfrac{dt}{t^2}$ converge, il en est de même de $\int_1^{+\infty} \dfrac{e^{-at}}{t}\,dt$ et donc de $\int_1^{+\infty} \left|e^{-at}\dfrac{\sin t}{t}\right| dt$.
D'autre part, au voisinage de 0, $e^{-at}\dfrac{\sin t}{t} \underset{t\to0}{\sim} e^{-at}\dfrac{t}{t} = e^{-at}$.
Et donc $\lim\limits_{t\to0^+} e^{-at}\dfrac{\sin t}{t} = e^{-a\times0} = 1$.
Par conséquent, la fonction $t \mapsto \left|e^{-at}\dfrac{\sin t}{t}\right|$ est prolongeable par continuité en 0, de sorte que $\int_0^1 \left|e^{-at}\dfrac{\sin t}{t}\right| dt$ converge.
Et donc $\int_0^{+\infty} \left|e^{-at}\dfrac{\sin t}{t}\right| dt$ converge, de sorte que I converge absolument.

Astuce

Pour la fonction sinus, nous savons que les dérivées successives sont cos, $-$ sin, $-$ cos, sin, cos, $-$ sin, $-$ cos, ..., où les dérivées $4k$-ièmes sont toutes égales à sin, les dérivées $4k+1$-ièmes sont toutes égales à cos, etc. Un moyen pratique d'écrire ceci est donc

$$\sin^{(k)}(x) = \sin\left(x + k\frac{\pi}{2}\right).$$

Le même raisonnement vaut pour la fonction cosinus :

$$\cos^{(k)}(x) = \cos\left(x + k\frac{\pi}{2}\right).$$

3.a. Soit $n \in \mathbf{N}$. La fonction sin est de classe $\mathscr{C}^{n+2}$ sur $\mathbf{R}$.

De plus, nous savons que $\sin'(x) = \cos(x) = \sin\left(x + \frac{\pi}{2}\right)$,

$$\sin''(x) = -\sin(x) = \sin(x+\pi), \sin^{(3)}(x) = -\cos(x) = \sin\left(x + \frac{3\pi}{2}\right), \ldots$$

Et plus généralement, pour tout $k \in \mathbf{N}$, $\sin^{(k)}(x) = \sin\left(x + \dfrac{k\pi}{2}\right)$.
En particulier, $\sin^{(k)}(0) = \sin\left(k\dfrac{\pi}{2}\right) = \begin{cases} 0 & \text{si } k \text{ est pair} \\ (-1)^p & \text{si } k = 2p+1 \text{ est impair} \end{cases}$.
D'autre part, pour tout $t \in \mathbf{R}$, $\left|\sin^{(2n+2)}(t)\right| \leqslant 1$, de sorte que par l'inégalité de Taylor-Lagrange,

$$\left|\sin x - \left(\sum_{k=0}^{2n+1} \sin^{(k)}(0)\frac{x^k}{k!}\right)\right| \leqslant \frac{|x|^{2n+2}}{(2n+2)!}.$$

Et donc

$$\left|\sin x - \left(x - \frac{x^3}{3!} + \cdots + (-1)^n \frac{x^{2n+1}}{(2n+1)!}\right)\right| \leqslant \frac{x^{2n+2}}{(2n+2)!}.$$

Remarque

Notons que $\sum\limits_{k=0}^{2n+1} \sin^{(k)}(0)\dfrac{x^k}{k!}$ n'est autre que le développement limité d'ordre $2n+1$ de sin au voisinage de 0.
Ce développement limité est connu, c'est du cours, et donc il n'est pas indispensable de le calculer.

3.b. Pour $x > 0$, en multipliant l'inégalité précédente par $\dfrac{e^{-ax}}{x}$, il vient

$$\left|e^{-ax}\frac{\sin x}{x} - \sum_{k=0}^{n} (-1)^k e^{-ax}\frac{x^{2k}}{(2k+1)!}\right| \leqslant e^{-ax}\frac{x^{2n+1}}{(2n+2)!}.$$

En intégrant cette inégalité, par croissance de l'intégrale, on obtient donc[1]

$$\int_0^{+\infty} \left|e^{-ax}\frac{\sin x}{x} - \sum_{k=0}^{n} (-1)^k e^{-ax}\frac{x^{2k}}{(2k+1)!}\right| dx \leqslant \frac{1}{(2n+2)!}\int_0^{+\infty} e^{-ax}x^{2n+1}\,dx = \frac{I_{2n+1}}{(2n+2)!}.$$

[1] Ce qui est légitime car l'intégrale du terme de droite de l'inégalité est convergente.

Et alors, par l'inégalité triangulaire, il vient

$$\left|I-\sum_{k=0}^{n}(-1)^k\frac{I_{2k}}{(2k+1)!}\right| = \left|\int_0^{+\infty}\left(e^{-ax}\frac{\sin x}{x}-\sum_{k=0}^{n}(-1)^k e^{-ax}\frac{x^{2k}}{(2k+1)!}\right)dx\right|$$
$$\leqslant \int_0^{+\infty}\left|e^{-ax}\frac{\sin x}{x}-\sum_{k=0}^{n}(-1)^k e^{-ax}\frac{x^{2k}}{(2k+1)!}\right|dx$$
$$\leqslant \frac{1}{(2n+2)!}I_{2n+1}.$$

Ainsi, en posant $K_n=\dfrac{1}{(2n+2)!}$, on a bien l'inégalité souhaitée.

3.c. En combinant l'inégalité de la question précédente avec l'expression des I_k obtenue à la question 1.b, il vient, pour tout $n\in\mathbf{N}$,

$$\left|I-\sum_{k=0}^{n}\frac{1}{(2k+1)a^{2k+1}}\right| = \left|I-\sum_{k=0}^{n}(-1)^k\frac{I_{2k}}{(2k+1)!}\right| \leqslant \frac{1}{(2n+2)!}\frac{(2n+1)!}{a^{2n+2}} \leqslant \frac{1}{(2n+2)a^{2n+2}} \underset{n\to+\infty}{\longrightarrow} 0.$$

Et donc

$$\lim_{n\to+\infty} I-\sum_{k=0}^{n}(-1)^k\frac{1}{(2k+1)a^{2k+1}}=0 \Leftrightarrow \lim_{n\to+\infty}\sum_{k=0}^{n}(-1)^k\frac{1}{(2k+1)a^{2k+1}}=I.$$

Ceci prouve donc que la série $\displaystyle\sum_{k\geqslant 0}\frac{(-1)^k}{(2k+1)a^{2k+1}}$ converge, et que sa somme vaut I.

> **Rappel**
> Par **définition**, une série converge ei et seulement si la suite de ses sommes partielles converge, ce qui est bien le cas ici.
> Et dans ce cas, toujours par définition, la somme de la série est la limite de la suite des sommes partielles.

4.a. Pour tout $t\in[0,1]$ et pour tout $n\in\mathbf{N}^*$, on a

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

Et donc

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

4.b. Fixons $x\in[0,1]$, $n\in\mathbf{N}^*$, et intégrons l'égalité précédente entre 0 et x :

$$\sum_{k=0}^{n}(-1)^k\int_0^x t^{2k}\,dt-\int_0^x\frac{1}{1+t^2}\,dt=(-1)^n\int_0^x\frac{t^{2n+2}}{1+t^2}\,dt.$$

Soit encore

$$\sum_{k=0}^{n}(-1)^k\frac{x^{2k+1}}{2k+1}-\operatorname{Arctan}(x)=(-1)^n\int_0^x\frac{t^{2n+2}}{1+t^2}\,dt.$$

Mais, pour tout $t\in[0,x]$, on a $\dfrac{t^{2n+2}}{1+t^2}\leqslant t^{2n+2}$ de sorte que, par croissance de l'intégrale,

$$\int_0^x\frac{t^{2n+2}}{1+t^2}\,dt\leqslant\int_0^x t^{2n+2}\,dt=\frac{x^{2n+3}}{2n+3}\leqslant\frac{1}{2n+3}.$$

Et donc

$$\left|\sum_{k=0}^{n}(-1)^k t^{2k}-\operatorname{Arctan}(x)\right|=\int_0^x\frac{t^{2n+2}}{1+t^2}\,dt\leqslant\frac{1}{2n+3}.$$

4.c. Lorsque n tend vers $+\infty$ dans l'inégalité précédente, il vient, par le théorème des gendarmes,

$$\lim_{n\to+\infty}\sum_{k=0}^{n}(-1)^k\frac{x^{2k+1}}{2k+1}-\operatorname{Arctan}(x)=0\Leftrightarrow\lim_{n\to+\infty}\sum_{k=0}^{n}(-1)^k\frac{x^{2k+1}}{2k+1}=\operatorname{Arctan}(x).$$

Ainsi, la série de terme général $(-1)^k \dfrac{x^{2k+1}}{2k+1}$ converge et $\displaystyle\sum_{k=0}^{+\infty}(-1)^k \frac{x^{2k+1}}{2k+1} = \text{Arctan}(x)$.

En particulier, pour $x = \dfrac{1}{a} \in [0,1]$, il vient

$$\sum_{k=0}^{+\infty}(-1)^k \frac{1}{(2k+1)a^{2k+1}} = \text{Arctan}\left(\frac{1}{a}\right).$$

Et alors, en reprenant le résultat de la question 3.c, il vient

$$I = \sum_{k=0}^{+\infty}(-1)^k \frac{1}{(2k+1)a^{2k+1}} = \text{Arctan}\left(\frac{1}{a}\right).$$

□

EXERCICE 2.20 (ESCP 2012) [ESCP12.1.19] — Difficile

Théorème de convergence radiale d'Abel.

On considère une suite réelle bornée $(a_n)_{n\geqslant 0}$ et on pose, pour tout $x \in I = [0,1[$:

$$f(x) = \sum_{n=0}^{+\infty} a_n x^n.$$

On note $(s_n)_{n\geqslant 0}$ la suite des sommes partielles définie par :

$$\forall n \in \mathbf{N},\ s_n = a_0 + a_1 + \cdots + a_n.$$

On admet que si une suite $(u_n)_{n\geqslant 0}$ converge vers ℓ et si $x \mapsto K(x)$ est une fonction définie sur I à valeurs dans $\mathbf{N}$ telle que $\lim\limits_{x\to 1^-} K(x) = +\infty$, alors on a $\lim\limits_{x\to 1^-} u_{K(x)} = \ell$.

1. Vérifier que la série définissant la fonction f est absolument convergente pour tout x de I. La fonction f est donc bien définie sur I.
2. Soit N un entier strictement positif fixé. Montrer que

$$(1-x)\sum_{n=0}^{N} s_n x^n = \sum_{n=0}^{N} a_n x^n - s_N x^{N+1}.$$

En déduire que la série $\displaystyle\sum_{n=0}^{+\infty} s_n x^n$ est convergente pour tout $x \in I$ et que l'on a :

$$f(x) = (1-x)\sum_{n=0}^{+\infty} s_n x^n.$$

3. On suppose dans cette question que la série $\displaystyle\sum_{n=0}^{+\infty} a_n$ est convergente et de somme nulle.
   (a) Pour tout $x \in I$, on pose $K(x) = \left\lfloor \dfrac{1}{\sqrt{1-x}} \right\rfloor$. Vérifier que $K(x)$ tend vers $+\infty$ lorsque x tend vers 1 par valeurs inférieures.
   (b) On pose $u_n = \sup\{|s_k|, k \geqslant n\}$. Montrer que la suite (u_n) est bien définie et converge vers 0.
   (c) Montrer que l'on a :

$$|f(x)| \leqslant (1-x)\left((K(x)+1)u_0 + \sum_{k=K(x)+1}^{+\infty} |s_k| x^k\right) \leqslant (1-x)(K(x)+1)u_0 + u_{K(x)+1}.$$

   (d) En déduire que $\lim\limits_{x\to 1^-} f(x) = 0$.
4. On suppose dans cette question que la série $\displaystyle\sum_{n=0}^{+\infty} a_n$ est convergente et de somme s. Prouver que $f(x)$ converge vers s lorsque x tend vers 1 par valeurs inférieures.

Solution de l'exercice 2.20

1. La suite (a_n) étant bornée, notons M un réel tel que pour tout $n \in \mathbf{N}$, $|a_n| \leqslant M$.
Alors pour $x \in I$, on a $0 \leqslant |a_n x^n| \leqslant x^n$.
Mais la série de terme de terme général x^n est une série géométrique convergente, et donc $\sum_n |a_n x^n|$ converge, de sorte que la série définissant $f(x)$ converge absolument.

2. On a, en notant que pour $n \geqslant 1$, $a_n = s_n - s_{n-1}$,

$$\sum_{n=0}^{N} a_n x^n - s_N x^{N+1} = a_0 + \sum_{n=1}^{N} (s_n - s_{n-1}) x^n - s_N X^{N+1}$$

$$= a_0 + \sum_{n=1}^{N} s_n x^n - \sum_{n=1}^{N} s_{n-1} x^n - s_N x^{N+1}$$

$$= s_0 + \sum_{n=1}^{N} s_n x^n - \sum_{k=0}^{N-1} s_k x^{k+1} - s_N x^{N+1}$$

$$= \sum_{n=0}^{N} s_n x^n - \sum_{k=0}^{N} s_k x^{k+1}$$

$$= \sum_{n=0}^{N} s_n x^n - x \sum_{n=0}^{N} s_n x^n$$

$$= (1-x) \sum_{n=0}^{N} s_n x^n.$$

D'autre part, par l'inégalité triangulaire, on a $|s_n| \leqslant |a_0| + \cdots + |a_n| \leqslant (n+1)M$.
Et donc $0 \leqslant |s_N x^{N+1}| \leqslant (n+1) M x^{N+1} \underset{N \to +\infty}{\longrightarrow} 0$ par croissances comparées.
Et donc en faisant tendre N vers $+\infty$ dans l'égalité précédemment prouvée,

$$\lim_{N \to +\infty} (1-x) \sum_{n=0}^{N} s_n x^n = \sum_{n=0}^{+\infty} a_n x^n = f(x).$$

Ceci prouve que la série de terme général $s_n x^n$ converge et que $(1-x) \sum_{n=0}^{+\infty} s_n x^n = f(x)$.

3.a. Par le propriétés de la partie entière, on a $K(x) \geqslant \underbrace{\frac{1}{\sqrt{x-1}}}_{\underset{x \to 1^-}{\longrightarrow} +\infty} - 1$, et donc $\lim_{x \to 1^-} K(x) = +\infty$.

3.b. Puisque la suite de terme général a_n est convergente, la suite (s_n) de ses sommes partielles est convergente, de limite nulle.
Et donc $(|s_n|)$ est également convergente, et de limite nulle.
Or, toute suite convergente est bornée, donc $\{|s_k|, k \geqslant n\}$ est majoré, et donc admet une borne supérieure.
Soit $\varepsilon > 0$. Alors il existe $k_0 \in \mathbf{N}$ tel que pour $k \geqslant k_0$, $|s_k| \leqslant \varepsilon$.
Et donc pour $k \geqslant k_0$, $|s_k| \leqslant \varepsilon$.
Et en particulier, pour $n \geqslant k_0$, $0 \leqslant u_n \leqslant \varepsilon$.
Ainsi, pour tout $\varepsilon > 0$, il existe $k_0 \in \mathbf{N}$ tel que $n \geqslant k_0 \Rightarrow |u_n| \leqslant \varepsilon$, ce qui prouve que $\lim_{n \to +\infty} u_n = 0$.

> **Rappel**
> Toute partie majorée de $\mathbf{R}$ admet une borne supérieure (qui par définition en est le plus petit des majorants).

3.c. Par l'inégalité triangulaire,

$$|f(x)| = (1-x) \left| \sum_{n=0}^{+\infty} s_n x^n \right|$$

$$\leqslant (1-x) \sum_{n=0}^{+\infty} |s_n| x^n$$

$$\leqslant (1-x) \sum_{n=0}^{K(x)} s_n x^n + (1-x) \sum_{n=K(x)+1}^{+\infty} |s_n| x^n.$$

Or, $u_0 = \sup\{|s_k| k \in \mathbf{N}\}$ est un majorant de la suite $(|s_n|)$, de sorte que

$$\forall n \in [\![0, K(x)]\!],\ |s_n| \leqslant u_0.$$

Et donc

$$0 \leqslant \sum_{n=0}^{K(x)} |s_n| x^n \leqslant \sum_{n=0}^{K(x)} |s_n| \leqslant \sum_{n=0}^{K(x)} u_0 \leqslant (K(x)+1)u_0.$$

Et donc on a bien prouvé que

$$|f(x)| \leqslant (1-x)\left((K(x)+1)u_0 + \sum_{k=K(x)+1}^{+\infty} |s_k| x^k\right).$$

Pour la seconde inégalité, il s'agit de prouver que

$$(1-x)\sum_{k=K(x)+1}^{+\infty} |s_k| x^k \leqslant u_{K(x)+1}.$$

Mais pour $k \geqslant K(x)+1$, $|s_k| \leqslant u_{K(x)+1}$, par définition de la suite (u_n).
Et donc

$$\begin{aligned}
(1-x)\sum_{k=K(x)+1}^{+\infty} |s_k| x^k &\leqslant (1-x)\sum_{k=K(x)+1}^{+\infty} u_{K(x)+1} x^k \\
&\leqslant u_{K(x)+1}(1-x)\sum_{k=K(x)+1}^{+\infty} x^k \\
&\leqslant u_{K(x)+1}(1-x)\sum_{k=0}^{+\infty} x^k && \text{Les } x^k \text{ sont positifs.} \\
&\leqslant u_{K(x)+1}(1-x)\frac{1}{1-x} && \text{Somme d'une série géométrique de raison } x. \\
&\leqslant u_{K(x)+1}.
\end{aligned}$$

Et donc on a bien prouvé que

$$(1-x)\left((K(x)+1)u_0 + \sum_{k=K(x)+1}^{+\infty} |s_k| x^k\right) \leqslant (1-x)(K(x)+1)u_0 + u_{K(x)+1}.$$

3.d. On a d'une part

$$0 \leqslant (1-x)(K(x)+1)u_0 \leqslant (1-x)\left(\frac{1}{\sqrt{1-x}} + 1\right) \leqslant (1-x) + \sqrt{1-x}$$

de sorte que par le théorème des gendarmes, $\lim\limits_{x\to 1^-} (1-x)(K(x)+1) = 0$.
D'autre part, d'après le résultat admis dans l'énoncé, et puisque $u_{K(x)+1} \underset{x\to 1^-}{\longrightarrow} 0$..
Et donc $\lim\limits_{x\to 1^-} |f(x)| = 0$, et donc $\lim\limits_{x\to 1^-} f(x) = 0$.

4. Puisque nous savons que $\sum\limits_{n=0}^{+\infty} \frac{1}{2^n} = 2$, posons

$$\forall n \in \mathbf{N},\ b_n = a_n - \frac{s}{2^{n+1}}$$

de sorte que (b_n) est encore une suite bornée, avec

$$\sum_{n=0}^{+\infty} b_n = \sum_{n=0}^{+\infty} a_n - \frac{1}{2}\sum_{n=0}^{+\infty} \frac{1}{2^n} = s - \frac{1}{2}2s = 0.$$

Alors les résultats de la question précédente s'appliquent à (b_n) : $\sum\limits_{n=0}^{+\infty} b_n x^n \underset{x\to 1^-}{\longrightarrow} 0$.
Or

$$\sum_{n=0}^{+\infty} b_n x^n = \sum_{n=0}^{+\infty} a_n x^n - \frac{s}{2}\sum_{n=0}^{+\infty} \left(\frac{x}{2}\right)^n = f(x) - \frac{s}{2}\frac{1}{1-\frac{x}{2}}.$$

Lorsque $x \to 1^-$, $\dfrac{s}{2}\dfrac{1}{1-\frac{x}{2}} \underset{x\to 1^-}{\longrightarrow} \dfrac{s}{2}\dfrac{1}{1-\frac{1}{2}} = s$.

Et donc

$$\lim_{x\to 1^-} f(x) = \lim_{x\to 1^-} \sum_{n=0}^{+\infty} b_n x^n + \frac{s}{2}\frac{1}{1-\frac{x}{2}} = 0 + s = s.$$

□

Exercice 2.21 (ESCP 2017) [ESCP17.1.06] — Moyen

Quelques variations autour de la série harmonique.
Abordable en première année : ✓

1. Pour tout $n \in \mathbf{N}^*$, on pose $H_n = \sum_{k=1}^{n} \frac{1}{n}$ 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 $(\gamma_n)_{n\geqslant 1}$ 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 $\sum_{k\geqslant 1} \frac{(-1)^k}{k}$ 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 $\lim_{n\to+\infty} \sum_{k=1}^{3n+3} \frac{a_k}{k}$ 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 $\sum_{n=1}^{+\infty} \frac{a_n}{n} = \frac{\pi}{4} + \frac{1}{2}\ln(2)$.

Solution de l'exercice 2.21

1. On a

$$\begin{aligned}
\gamma_{n+1} - \gamma_n &= H_{n+1} - \ln(n+1) - H_n + \gamma_n \\
&= \frac{1}{n+1} + \ln\left(\frac{n}{n+1}\right) \\
&= \frac{1}{n+1} + \ln\left(1 - \frac{1}{n+1}\right) \\
&= \frac{1}{n+1} - \frac{1}{n+1} + \underset{n\to+\infty}{o}\left(\frac{1}{(n+1)^2}\right) \\
&= \underset{n\to+\infty}{o}\left(\frac{1}{(n+1)^2}\right).
\end{aligned}$$

Détails
On a utilisé le développement limité
$\ln(1+u) = u + \underset{u\to 0}{o}(u^2)$
qui est légitime puisque $\frac{1}{n} \underset{n\to+\infty}{\longrightarrow} 0$.

Mais puisque la série de terme général $\frac{1}{(n+1)^2}$ converge, il en est de même de $\sum_{n\geqslant 1} (\gamma_{n+1} - \gamma_n)$.

Pour $N \in \mathbf{N}^*$, on a

$$\sum_{n=1}^{N}(\gamma_{n+1}-\gamma_n) = \gamma_2 - \gamma_1 + \gamma_3 - \gamma_2 + \cdots + \gamma_N - \gamma_{N-1} + \gamma_{N+1} - \gamma_N = \gamma_{N+1} - \gamma_1.$$

Puisque la série de terme général $(\gamma_{n+1}-\gamma_n)$ converge, ses sommes partielles admettent une limite lorsque $N \to +\infty$. Et donc $\lim_{N\to+\infty} \gamma_{N+1} - \gamma_1$ existe, et donc la suite $(\gamma_N)_{N\geqslant 1}$ converge.

Remarque *: puisque* $\gamma = \underset{n\to+\infty}{o}(\ln n)$*, on a donc prouvé que* $H_n = \ln(n) + \underset{n\to+\infty}{o}(\ln n) \underset{n\to+\infty}{\sim} \ln n$.

> **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

$$\begin{aligned}\sum_{k=1}^{2n}\frac{(-1)^k}{k} + H_{2n} &= \sum_{k=1}^{n}\frac{1}{2k} - \sum_{k=1}^{n-1}\frac{1}{2k+1} + \sum_{k=1}^{2n}\frac{1}{k}\\ &= \sum_{k=1}^{n}\frac{1}{2k} - \sum_{k=1}^{n-1}\frac{1}{2k+1} + \sum_{k=1}^{n}\frac{1}{2k} + \sum_{k=1}^{n-1}\frac{1}{2k+1}\\ &= \sum_{k=1}^{n}\frac{1}{k} = H_n.\end{aligned}$$

Et donc $\displaystyle\sum_{k=1}^{2n}\frac{(-1)^k}{k} = H_n - H_{2n}$.

Lorsque $n \to +\infty$, on a $H_n = \ln(n) + \gamma + \underset{n\to+\infty}{o}(1)$, de sorte que

> **$o(1)$**
> Rappelons que la notation $o(1)$ désigne toute suite de limite nulle.

$$\sum_{k=1}^{2n}\frac{(-1)^k}{k} = \ln(n) + \gamma + \underset{n\to+\infty}{o}(1) - \ln(2n) - \gamma - \underset{n\to+\infty}{o}(1) = -\ln(2) + \underset{n\to+\infty}{o}(1) \underset{n\to+\infty}{\longrightarrow} -\ln(2).$$

D'autre part,

$$\sum_{k=1}^{2n+1}\frac{(-1)^k}{k} = \sum_{k=1}^{2n}\frac{(-1)^k}{k} - \frac{1}{2n+1} \underset{n\to+\infty}{\longrightarrow} -\ln(2).$$

Ainsi, si on note $S_n = \displaystyle\sum_{k=1}^{n}\frac{(-1)^k}{k}$ la somme partielle d'ordre n de la série de terme général $\dfrac{(-1)^k}{k}$, nous venons de montrer que les deux suites $(S_{2n})_{n\geqslant 1}$ et $(S_{2n+1})_{n\geqslant 0}$ convergent vers une même limite $-\ln(2)$, de sorte que $S_n \underset{n\to+\infty}{\longrightarrow} -\ln(2)$.

Et donc la série de terme général $\dfrac{(-1)^k}{k}$ converge et $\displaystyle\sum_{k=1}^{+\infty}\frac{(-1)^k}{k} = -\ln(2)$.

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

$$\sum_{k=1}^{3n+3}\frac{a_k}{k} = \ln(3n+3)+\gamma-\underset{n\to+\infty}{o}(1)-\frac{2}{3}\ln(n+1)-\frac{2}{3}\gamma-\underset{n\to+\infty}{o}(1) = \ln(3)+\frac{1}{3}\ln(n+1)+\frac{1}{3}\gamma+\underset{n\to+\infty}{o}(1) \underset{n\to+\infty}{\longrightarrow} +\infty.$$

Et donc la série de terme général $\dfrac{a_k}{k}$ diverge.

4.a. Notons que pour tout $k \in \mathbf{N}$, $\dfrac{1}{k+1} = \displaystyle\int_0^1 x^k\,dx$.

Et donc

$$\begin{aligned}\sum_{n=0}^{N}\left(\frac{1}{4n+1}-\frac{1}{4n+3}\right) &= \sum_{n=0}^{N}\int_0^1\left(x^{4n}-x^{4n+2}\right)dx\\ &= \int_0^1(1-x^2)\sum_{n=0}^{N}\left(x^4\right)^n dx\\ &= \int_0^1(1-x^2)\frac{1-\left(x^4\right)^{N+1}}{1-x^4}\,dx\\ &= \int_0^1\frac{1-x^{4N+4}}{1+x^2}\,dx\end{aligned}$$

> **Détails**
> On a une identité remarquable classique :
> $1-x^4=(1+x^2)(1-x^2)$.

4.b. Coupons en deux l'intégrale qui précède :

$$\int_0^1\frac{1-x^{4N+4}}{1+x^2}\,dx = \int_0^1\frac{dx}{1+x^2}-\frac{x^{4N+4}}{1+x^2}\,dx.$$

La première intégrale vaut $[\text{Arctan}(t)]_0^1=\dfrac{\pi}{4}$.

Pour la seconde, on peut noter que pour tout $x\in[0,1]$, $0\leqslant\dfrac{x^{4N+4}}{1+x^2}\leqslant x^{4N+4}$ et donc par croissance de l'intégrale,

$$0\leqslant\int_0^1\frac{x^{4N+4}}{1+x^2}\,dx\leqslant\int_0^1 x^{4N+4}\,dx=\frac{1}{4N+5}.$$

Et donc par le théorème des gendarmes, $\displaystyle\lim_{N\to+\infty}\frac{x^{4N+4}}{1+x^2}\,dx=0$.

Et donc $\displaystyle\lim_{N\to+\infty}\sum_{n=0}^{N}\left(\frac{1}{4n+1}-\frac{1}{4n+3}\right)dx=\frac{\pi}{4}$.

Ceci prouve donc que $\displaystyle\sum_{n\geqslant0}\left(\frac{1}{4n+1}-\frac{1}{4n+3}\right)$ converge et que $\displaystyle\sum_{n=0}^{+\infty}\left(\frac{1}{4n+1}-\frac{1}{4n+3}\right)=\frac{\pi}{4}$.

4.c. Sur le même principe, on prouve que

$$\begin{aligned}\sum_{n=0}^{N}\left(\frac{1}{4n+2}-\frac{1}{4n+4}\right) &= x(1-x^2)\int_0^1\sum_{n=0}^{n}\left(x^4\right)^n dx\\ &= \int_0^1\frac{x-x^{4N+5}}{1+x^2}\,dx\\ &= \int_0^1\frac{x}{1+x^2}\,dx-\int_0^1\frac{x^{4N+5}}{1+x^2}\,dx\\ &= \left[\frac{1}{2}\ln(1+x^2)\right]_0^1-\int_0^1\frac{x^{4N+5}}{1+x^2}\,dx\\ &= \frac{1}{2}\ln(2)-\int_0^1\frac{x^{4N+5}}{1+x^2}\,dx.\end{aligned}$$

Et donc on conclut comme à la question précédente :

$$\sum_{n=0}^{+\infty}\left(\frac{1}{4n+2}-\frac{1}{4n+4}\right)=\frac{1}{2}\ln(2).$$

5. Notons $S_n=\displaystyle\sum_{k=1}^{n}\frac{a_k}{k}$. Alors

$$S_{4N+4}=\sum_{n=1}^{N}\left(\frac{1}{4n+1}-\frac{1}{4n+3}\right)+\sum_{n=1}^{N}\left(\frac{1}{4n+2}-\frac{1}{4n+4}\right)\underset{N\to+\infty}{\longrightarrow}\frac{\pi}{4}+\frac{1}{2}\ln(2).$$

Toutefois, cela ne suffit pas à prouver que $\sum_{k\geqslant 1}\frac{a_k}{k}$ converge, nous avons prouvé que la suite $(S_{4N+4})_{N\geqslant 0}=(S_{4n})_{n\geqslant 0}$ converge, mais cela ne suffit pas à garantir que $(S_n)_{n\geqslant 1}$ converge. On pourra par exemple penser au cas de la suite $u_n=(-1)^n$.

En revanche, pour $n\in\mathbf{N}$, on a $4\left\lfloor\frac{n}{4}\right\rfloor$ est le multiple de 4 immédiatement inférieur à n, de sorte que $\left|n-4\left\lfloor\frac{n}{4}\right\rfloor\right|\leqslant 3$.
Et donc

$$\left|S_n-S_{4\left\lfloor\frac{n}{4}\right\rfloor}\right|\leqslant\frac{1}{4\left\lfloor\frac{n}{4}\right\rfloor+1}+\frac{1}{4\left\lfloor\frac{n}{4}\right\rfloor+2}+\frac{1}{4\left\lfloor\frac{n}{4}\right\rfloor+4}\underset{n\to+\infty}{\longrightarrow}0.$$

Et donc $S_n\underset{n\to+\infty}{\longrightarrow}\frac{\pi}{4}+\frac{1}{2}\ln(2)$, de sorte que

$$\sum_{n=0}^{+\infty}\frac{a_n}{n}=\frac{\pi}{4}+\frac{1}{2}\ln(2).$$

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 limite[1], de sorte que (S_n) converge également vers cette même limite.*

□

Autrement dit
La somme définissant S_n contient au maximum trois termes de plus que $S_{\left\lfloor\frac{n}{4}\right\rfloor}$. Nous allons utiliser dans la suite le fait que ces trois termes tendent vers 0 lorsque $n\to+\infty$.

[1] Qui est $\frac{\pi}{4}+\frac{1}{2}\ln(2)$.

EXERCICE 2.22 (ESCP 2018) [ESCP18.1. ? ?] — Facile

Calcul de la somme d'une série d'intégrales
Abordable en première année : ✓

On pose, pour tout entier naturel n : $a_n=\int_0^1 t^n\sqrt{1-t^2}\,dt$.

1. (a) Montrer que la suite (a_n) est décroissante. En déduire qu'elle converge.
   (b) À l'aide du changement de variable $t=\sin u$, calculer a_0.
2. (a) Montrer que pour tout $n\in\mathbf{N}$, on a : $(n+4)a_{n+2}=(n+1)a_n$.
   (b) Soit (w_n) la suite définie par : pour tout $n\in\mathbf{N}^*$, $w_n=n(n+1)(n+2)a_na_{n-1}$. Montrer que la suite (w_n) est constante, et calculer cette constante.
   (c) Montrer que $a_n\underset{n\to+\infty}{\sim}a_{n+1}$.
   (d) Établir l'existence d'un réel $K>0$ que l'on calculera, tel que $a_n\underset{n\to+\infty}{\sim}\frac{K}{n^{3/2}}$.
   (e) En déduire la nature de la série de terme général $b_n=(-1)^na_n$.
3. (a) Montrer que l'on a : $\lim_{n\to+\infty}\sum_{k=0}^{n}b_k=\int_0^1\sqrt{\frac{1-t}{1+t}}\,dt$.
   (b) En utilisant le changement de variable $u=\sqrt{\frac{1-t}{1+t}}$, calculer l'intégrale $\int_0^1\sqrt{\frac{1-t}{1+t}}\,dt$. Quelle est la somme de la série $\sum_{n\geqslant 0}b_n$?

SOLUTION DE L'EXERCICE 2.22

1.a. Pour $t\in[0,1]$, et $n\in\mathbf{N}$, on a $t^{n+1}\leqslant t^n$, de sorte que par croissance de l'intégrale, $\int_0^1 t^{n+1}\sqrt{1-t^2}\,dt\leqslant\int_0^1 t^n\sqrt{1-t^2}\,dt$.
Et donc (a_n) est décroissante. Étant clairement positive, par le théorème de la limite monotone, elle converge.

1.b. Posons donc $t=\sin u$. Notons que pour $u=0$, $t=0$, et pour $u=\frac{\pi}{2}$, $t=1$.
De plus, on a $dt=\cos u\,du$. Et donc

$$a_0=\int_0^1\sqrt{1-t^2}\,dt=\int_0^{\pi/2}\sqrt{1-\sin^2 u}\cos u\,du=\int_0^{\pi/2}\cos^2 u\,du.$$

Signe
Puisque $u\in[0,\frac{\pi}{2}]$, alors $\cos u\geqslant 0$, de sorte que $\sqrt{\cos^2 u}=\cos u$.

D'autre part, en utilisant la formule d'Euler, on a

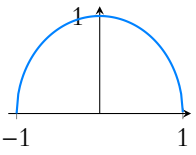
$$\cos^2 u = \left(\frac{e^{iu}+e^{-iu}}{2}\right)^2 = \frac{e^{2iu}+2+e^{-2iu}}{4} = \frac{1+\cos(2u)}{2}$$

de sorte que

$$a_0 = \int_0^{\pi/2} \frac{1+\cos(2u)}{2}\,du = \frac{1}{2}\left[u+\frac{\sin(2u)}{2}\right]_0^{\pi/2} = \frac{\pi}{4}.$$

Géométriquement

La courbe représentative de la fonction $x \mapsto \sqrt{1-x^2}$ est un demi-cercle centré en l'origine et de rayon 1. En effet, on a $y = \sqrt{1-x^2} \Rightarrow x^2+y^2 = 1$.



Et donc l'aire que nous calculons n'est autre que l'aire d'un quart de ce disque, qui vaut donc $\frac{\pi}{4}$.

2.a. Réalisons une intégration par parties dans l'intégrale a_{n+2}, en posant $u(t) = t^{n+1}$ et $v(t) = -\frac{1}{3}\left(1-t^2\right)^{3/2}$ qui sont deux fonctions de classe $\mathscr{C}^1$ sur $[0,1]$, avec $u'(t) = (n+1)t^n$ et $v'(t) = t\sqrt{1-t^2}$. On a alors

$$a_{n+2} = \int_0^1 t^{n+2}\sqrt{1-t^2}\,dt = \left[-\frac{t^{n+1}}{3}\left(1-t^2\right)^{3/2}\right]_0^1 + \frac{n+1}{3}\int_0^1 t^n(1-t^2)\sqrt{1-t^2}\,dt = \frac{n+1}{3}\left(a_n - a_{n+2}\right).$$

Et donc $\frac{n+4}{3}a_{n+2} = \frac{n+1}{3}a_n \Leftrightarrow (n+4)a_{n+2} = (n+1)a_n$.

2.b. En utilisant la question précédente, il vient pour tout $n \in \mathbf{N}^*$,

$$w_{n+1} = (n+1)(n+2)(n+3)a_{n+1}a_n = (n+1)(n+2)a_n n a_{n-1} = w_n.$$

Et donc la suite (w_n) est constante. On a alors, pour tout $n \in \mathbf{N}$, $w_n = w_1 = 6a_1a_0$.
Nous avons déjà calculé a_0, et

$$a_1 = \int_0^1 t\sqrt{1-t^2}\,dt = \left[-\frac{1}{3}\left(1-t^2\right)^{3/2}\right]_0^1 = \frac{1}{3}.$$

Et donc pour tout $n \in \mathbf{N}^*$, $w_n = \frac{\pi}{2}$.

2.c. On a, pour tout $n \in \mathbf{N}$, par décroissance de (a_n), $a_{n+2} \leqslant a_{n+1} \leqslant a_n$.
Mais $a_{n+2} = \frac{n+1}{n+4}a_n$ et donc $\frac{n+1}{n+4}a_n \leqslant a_{n+1} \leqslant a_n$.
On en déduit donc, après division par $a_n > 0$ et par application du théorème des gendarmes, que $\frac{a_{n+1}}{a_n} \underset{n\to+\infty}{\longrightarrow} 1$, et donc $a_{n+1} \underset{n\to+\infty}{\sim} a_n$.

2.d. En utilisant les deux questions précédentes, il vient

$$w_n \underset{n\to+\infty}{\sim} \frac{\pi}{2} \Leftrightarrow n^3 a_n \underset{n\to+\infty}{\sim} \frac{\pi}{2} \Leftrightarrow a_n \underset{n\to+\infty}{\sim} \sqrt{\frac{\pi}{2}}\frac{1}{n^{3/2}}.$$

On a $|b_n| = a_n \underset{n\to+\infty}{\sim} \frac{1}{n^{3/2}}\sqrt{\frac{\pi}{2}}$, et donc $\sum b_n$ converge absolument[1] par comparaison à une série de Riemann.

[1] Et donc converge.

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

$$\begin{aligned}\sum_{k=0}^n b_k &= \sum_{k=0}^n (-1)^k \int_0^1 t^k\sqrt{1-t^2}\,dt = \int_0^1 \sqrt{1-t^2}\sum_{k=0}^n (-t)^k\,dt \\ &= \int_0^1 \sqrt{(1-t)(1+t)}\frac{1-(-t)^{n+1}}{1+t}\,dt \\ &= \int_0^1 \sqrt{\frac{1-t}{1+t}}\,dt + (-1)^n\int_0^1 \sqrt{\frac{1-t}{1+t}}t^{n+1}\,dt.\end{aligned}$$

Somme des termes d'une suite géométrique de raison $-t \neq 1$.

Mais d'autre part,

$$\left|\int_0^1 \sqrt{\frac{1-t}{1+t}}t^{n+1}\,dt\right| \leqslant \int_0^1 \left|\sqrt{\frac{1-t}{1+t}}t^n\right|\,dt \leqslant \int_0^1 t^n\,dt \leqslant \frac{1}{n+1}.$$

Détails

Pour $t \in [0,1]$, $0 \leqslant 1-t \leqslant 1$ et $1 \leqslant 1+t$, de sorte que $\sqrt{\frac{1-t}{1+t}} \leqslant 1$.

Et donc par le théorème des gendarmes, $\int_0^1 \sqrt{\frac{1-t}{1+t}}t^{n+1}\,dt \underset{n\to+\infty}{\longrightarrow} 0$, de sorte que

$$\lim_{n\to+\infty}\sum_{k=0}^n (-1)^k a_k = \int_0^1 \sqrt{\frac{1-t}{1+t}}\,dt.$$

3.b. Procédons au changement de variable indiqué : on a alors $u^2 = \dfrac{1-t}{1+t} \Leftrightarrow t = \dfrac{1-u^2}{1+u^2}$ et donc $dt = \dfrac{-4u}{(1+u^2)^2}$, de sorte que

$$\int_0^1 \sqrt{\frac{1-t}{1+t}}\,dt = \int_0^1 \frac{4u^2}{(1+u^2)}\,du.$$

Procédons alors à une intégration par parties, en posant $f(u) = u$ et $g(u) = -\dfrac{2}{1+u^2}$, qui sont deux fonctions de classe $\mathscr{C}^1$ sur $[0,1]$, avec $f'(u) = 1$ et $g'(u) = \dfrac{4u}{(1+u^2)^2}$, de sorte que

$$\int_0^1 \frac{4u^2}{(1+u^2)}\,du = \left[\frac{-2u}{1+u^2}\right]_0^1 + \int_0^1 \frac{2}{1+u^2}\,du = -1 + [2\operatorname{Arctan}(u)]_0^1 = \frac{\pi}{2} - 1.$$

Et donc $\displaystyle\sum_{n=0}^{+\infty} b_n = \frac{\pi}{2} - 1$.

□

2.3.1 Comparaison série/intégrale

Bien que l'étude des séries et des intégrales soient très semblables, les intégrales sont souvent plus faciles à étudier que les séries, du fait que nous disposons d'un outil supplémentaire : le calcul de primitive.
Il existe toutefois un outil qui permet de faire un lien entre les deux, que l'on appelle *comparaison série/intégrale*.
Plus aucun résultat ne figure au programme à ce sujet, mais le thème n'en reste pas moins classique.
L'exercice suivant résume l'essentiel de cette méthode.

EXERCICE 2.23 (ESCP 2014) [ESCP14.1.05] **Facile**

Comparaison série/intégrale

On considère une fonction f continue, décroissante et strictement positive sur l'intervalle $[1;+\infty[$. On pose pour tout entier $n \geqslant 1$:

$$u_n = \int_1^n f(t)\,dt \text{ et } v_n = \sum_{k=1}^{n} f(k).$$

1. Montrer que la suite $(w_n)_{n\geqslant 1}$ définie par $w_n = v_n - u_n$ est décroissante et à termes positifs.
2. On suppose que l'intégrale $\displaystyle\int_1^{+\infty} f(t)\,dt$ est divergente. Justifier l'inégalité $u_n > 0$ pour tout entier $n \geqslant 2$.
   Déterminer la limite de la suite $\left(\frac{v_n}{u_n}\right)_{n\geqslant 2}$.
3. Dans cette question, on suppose que l'intégrale $\displaystyle\int_1^{+\infty} f(t)\,dt$ converge.
   (a) Que peut-on dire de la série de terme général $f(k)$?
   (b) Donner, à l'aide d'une intégrale, un équivalent du reste $R_n = \displaystyle\sum_{k=n+1}^{+\infty} f(k)$ lorsque $\displaystyle\lim_{n\to+\infty} \frac{f(n)}{\int_n^{+\infty} f(t)\,dt} = 0$.
4. **Applications** :
   (a) Donner un équivalent de $\displaystyle\sum_{k=1}^{n} \frac{1}{k}$ lorsque n tend vers $+\infty$.
   (b) Donner un équivalent de $\displaystyle\sum_{k=n+1}^{+\infty} \frac{1}{1+k^2}$ lorsque n tend vers $+\infty$.

SOLUTION DE L'EXERCICE 2.23

1. On a

$$w_{n+1} - w_n = v_{n+1} - v_n + u_{n+1} - u_n = f(n+1) - \int_n^{n+1} f(t)\,dt$$

Mais par décroissance de f, pour tout tout $t \in [n, n+1]$, $f(n+1) \leqslant f(t)$, et donc

$$f(n+1) = \int_n^{n+1} f(n+1)\,dt \leqslant \int_n^{n+1} f(t)\,dt.$$

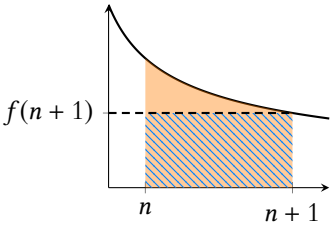


FIGURE 2.7– L'intégrale, qui est l'aire de la partie colorée est plus grande que l'aire de la partie hachurée, qui vaut $f(n+1)$.

On en déduit que $w_{n+1} - w_n \leqslant 0$ et donc (w_n) est décroissante.
De la même manière, on montre que pour tout $k \in \mathbf{N}^*$ et tout $t \in [k, k+1]$, $f(t) \leqslant f(k)$, et donc $\displaystyle\int_k^{k+1} f(t)\,dt \leqslant f(k)$.
En sommant ces relations pour k allant de 1 à $n-1$, il vient

$$\int_1^n f(t)\,dt \leqslant \sum_{k=1}^{n-1} f(k).$$

Et donc

$$w_n = f(n) + \sum_{k=1}^{n-1} f(k) - \int_1^n f(t)\,dt \geqslant f(n) > 0.$$

2. La fonction f est continue sur $[1, n]$ et y est strictement positive, donc pour tout $n \in \mathbf{N}$, $u_n = \displaystyle\int_1^n f(t)\,dt > 0$.
La fonction $F : x \mapsto \displaystyle\int_1^x f(t)\,dt$ est croissante car f est positive. Donc si l'intégrale diverge, c'est que F n'est pas majorée, et tend donc vers $+\infty$ lorsque $x \to +\infty$. En particulier,

$$\lim_{n\to+\infty} u_n = \lim_{n\to+\infty} F(n) = \lim_{n\to+\infty} \int_1^n f(t)\,dt = +\infty.$$

Positivité
La croissance de F se comprend bien sur un dessin en terme d'aire, mais si on souhaite le justifier par un calcul : f est positive, donc ses primitives sont croissantes.

Et comme $u_n \geqslant v_n$, on a de même $\lim\limits_{n\to+\infty} u_n = +\infty$.
Puisque (w_n) est décroissante et minorée, elle converge. Et par conséquent, $\lim\limits_{n\to+\infty} \dfrac{w_n}{u_n} = 0$.
Et donc

$$\frac{v_n}{u_n} = \frac{w_n + u_n}{u_n} = \underbrace{\frac{w_n}{u_n}}_{\to 0} + 1 \underset{n\to+\infty}{\longrightarrow} 1.$$

3.a. Si l'intégrale converge, alors F admet une limite finie α en $+\infty$, et donc est majorée (car elle est croissante). Et alors, pour tout $n \in \mathbf{N}^*$, $u_n = F(n) \leqslant \alpha$.
Or $w_n \leqslant w_1$, de sorte que

$$v_n = w_n + u_n \leqslant w_1 + \alpha.$$

Donc (v_n) est majorée, et puisqu'elle est croissante (par positivité de f), elle converge.
Et alors, puisque la suite des sommes partielles de $\sum_k f(k)$ converge[1], la série $\sum_k f(k)$ converge.

[1] La suite des sommes partielles de $\sum_k f(k)$ est précisément (v_n).

3.b. Pour tout $k \in \mathbf{N}$ et tout $t \in [k-1, k]$, on a $f(k) \leqslant f(t) \leqslant f(k-1)$ et donc

$$f(k) \leqslant \int_{k-1}^k f(t)\,dt \leqslant f(k-1).$$

Et donc, sommant ces relations pour k variant de $n+1$ à p $(p > n)$, il vient

$$\sum_{k=n+1}^{p} f(k) \leqslant \int_n^p f(t)\,dt \leqslant \sum_{k=n+1}^{p} f(k-1).$$

En passant à la limite lorsque $p \to +\infty$, il vient

$$\sum_{k=n+1}^{+\infty} f(k) \leqslant \int_n^{+\infty} f(t)\,dt \leqslant \sum_{k=n}^{+\infty} f(k).$$

Soit encore

$$R_n \leqslant \int_n^{+\infty} f(t)\,dt \leqslant R_{n-1} = R_n + f(n).$$

On a donc

$$\int_n^{+\infty} f(t)\,dt - f(n) \leqslant R_n \leqslant \int_n^{+\infty} f(t)\,dt.$$

Si l'on divise par $\int_1^n f(t)\,dt$, il reste alors

$$1 - \frac{f(n)}{\int_n^{+\infty} f(t)\,dt} \leqslant \frac{R_n}{\int_n^{+\infty} f(t)\,dt} \leqslant 1.$$

Grâce à l'hypothèse de l'énoncé et au théorème des gendarmes, on a alors

$$\lim_{n\to+\infty} \frac{R_n}{\int_n^{+\infty} f(t)\,dt} = 1 \Leftrightarrow R_n \underset{n\to+\infty}{\sim} \int_n^{+\infty} f(t)\,dt.$$

4. Applications

4.a. Nous avons ici $f(t) = \dfrac{1}{t}$, qui vérifie bien toutes les hypothèses : elle est continue, strictement positive et décroissante. De plus, $\int_1^{+\infty} f(t)\,dt$ diverge.
D'après la première question,

$$\sum_{k=1}^{n} \frac{1}{k} = \sum_{k=1}^{n} f(k) \underset{n\to+\infty}{\sim} \int_1^n f(t)\,dt = \ln(n).$$

> **Vitesse**
> Nous savions que
> $$\lim_{n\to+\infty} \sum_{k=1}^{n} \frac{1}{k} = +\infty$$
> mais nous savons désormais que cette divergence est «lente» car le ln ne tend pas très vite vers $+\infty$.

4.b. Nous avons ici $f(t) = \dfrac{1}{1+x^2}$, qui vérifie encore toutes les hypothèses.
Or, $\int_1^{+\infty} \dfrac{dt}{1+t^2}$ est convergente (par comparaison à une série de Riemann), et on a, pour tout $n \in \mathbf{N}^*$,

$$\int_n^{+\infty} f(t)\,dt = \lim_{A\to+\infty} \int_n^A \frac{dt}{1+t^2} = \lim_{A\to+\infty} \operatorname{Arctan}(A) - \operatorname{Arctan}(n) = \frac{\pi}{2} - \operatorname{Arctan}(n).$$

Alors

$$\lim_{n\to+\infty} \frac{f(n)}{\int_n^{+\infty} f(t)\,dt} = \lim_{n\to+\infty} \frac{1}{1+n^2} \frac{1}{\frac{\pi}{2} - \operatorname{Arctan}(n)}.$$

Considérons la fonction g définie sur $\mathbf{R}_+^*$ par $g(x) = \operatorname{Arctan}(x) + \operatorname{Arctan}\dfrac{1}{x}$. Alors g est de classe $\mathscr{C}^1$ sur $\mathbf{R}_+^*$, et on a

$$g'(x) = \frac{1}{1+x^2} - \frac{1}{x^2}\frac{1}{1+\frac{1}{x^2}} = \frac{1}{1+x^2} - \frac{1}{x^2+1} = 0.$$

> **Très dur**
> L'égalité que nous venons d'obtenir, $\forall x > 0$,
> $$\operatorname{Arctan} x + \operatorname{Arctan}\frac{1}{x} = \frac{\pi}{2}$$
> est «classique», mais n'est pas au programme, et on est d'habitude guidés pour la retrouver, et non sensés y penser seuls. Cela dit, si à l'oral vous arrivez jusqu'à cette question, elle servira surtout à décider si vous avez 19.5 ou 20 !

Donc g est constante sur $\mathbf{R}_+^*$. De plus,

$$\lim_{x\to+\infty} g(x) = \frac{\pi}{2} - \operatorname{Arctan}(0) = \frac{\pi}{2}$$

donc g est constante, égale à $\dfrac{\pi}{2}$.
En particulier, $\dfrac{\pi}{2} - \operatorname{Arctan}(n) = \operatorname{Arctan}\dfrac{1}{n} \underset{n\to+\infty}{\sim} \dfrac{1}{n}$.
Ainsi, $\dfrac{1}{1+n^2}\dfrac{1}{\frac{\pi}{2} - \operatorname{Arctan}(n)} \underset{n\to+\infty}{\sim} \dfrac{n}{1+n^2} \underset{n\to+\infty}{\longrightarrow} 0$.
Donc l'hypothèse technique de la question 3.b est vérifiée et par conséquent,

$$\sum_{k=n+1}^{+\infty} \frac{1}{1+k^2} \underset{n\to+\infty}{\sim} \int_n^{+\infty} \frac{dt}{1+t^2} = \frac{\pi}{2} - \operatorname{Arctan}(n) \underset{n\to+\infty}{\sim} \operatorname{Arctan}\frac{1}{n} \underset{n\to+\infty}{\sim} \frac{1}{n}.$$

□

Une comparaison série/intégrale peut nous aider à déterminer un équivalent de la suite des sommes partielles d'une série divergente. C'est le cas dans la question 2.c de l'exercice qui suit. Sans indication de la part de l'énoncé, toute la difficulté est d'y penser...

Exercice 2.24 (HEC 2012) [HEC12.40] Difficile

Une série dont le terme général est défini par une relation de récurrence

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

1. Écrire une fonction Scilab ayant pour argument un entier n et renvoyant $\displaystyle\sum_{k=0}^{n} u_k$.

2. Soit $\alpha \in \mathbf{R}$. On pose pour tout $n \in \mathbf{N}^*$, $v_n = \dfrac{(n+1)^\alpha u_{n+1}}{n^\alpha u_n}$.

   (a) Rappeler le développement limité à l'ordre deux au voisinage de 0 de $x \mapsto \ln(1+x)$.
   Montrer que $\ln v_n = (\alpha+1)\ln\left(1+\dfrac{1}{n}\right) - \ln\left(1+\dfrac{5}{2n}\right)$.
   Pour quelle valeur α_0 du réel α la série de terme général $\ln v_n$ est-elle convergente ?

   (b) Expliciter $\displaystyle\sum_{k=1}^{n} \ln v_k$ sans signe $\sum$, et en déduire qu'il existe un réel strictement positif C tel que $u_n \underset{n\to+\infty}{\sim} \dfrac{C}{n^{\alpha_0}}$.
   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 $\forall n \in \mathbf{N}$, $\displaystyle\sum_{k=0}^{n} k u_k \leqslant D\times\sqrt{n}$.

3. (a) Établir pour tout entier naturel n, la relation : $\displaystyle 2\sum_{k=1}^{n+1} k u_k + 3\sum_{k=1}^{n+1} u_k = 2\sum_{k=0}^{n} k u_k + 2\sum_{k=0}^{n} u_k$.

   (b) En déduire la valeur de $\displaystyle\sum_{n=0}^{+\infty} u_n$.

Solution de l'exercice 2.24

1. Le programme suivant convient :

```
1  function 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 semble 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 $u_n = \dfrac{2n}{2n+3}u_{n-1}$, et donc on pourra remplacer la ligne 5 par

```
5  S = (2*i)/(2*i+3)*u
```

2.a. On a $\ln(1+x) = x - \dfrac{x^2}{2} + \underset{x\to0}{o}(x^2)$.
De plus, pour tout $n \in \mathbf{N}^*$,

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

$$= \alpha \ln\left(\frac{n+1}{n}\right) + \ln\left(\frac{2(n+1)}{n\left(2+\frac{5}{n}\right)}\right)$$
$$= (\alpha+1)\ln\left(\frac{n+1}{n}\right) - \ln\left(\frac{2+\frac{5}{n}}{2}\right)$$
$$= (\alpha+1)\ln\left(1+\frac{1}{n}\right) - \ln\left(1+\frac{5}{2n}\right).$$

On en déduit donc que

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

En particulier, si $\alpha \neq \frac{3}{2}$, alors $\ln v_n \underset{n\to+\infty}{\sim} \left(\alpha - \frac{3}{2}\right)\frac{1}{n}$.
Et donc, par comparaison à une série de Riemann, $\sum \ln v_n$ diverge.
En revanche, si $\alpha = \frac{3}{2}$, alors $\ln v_n \underset{n\to+\infty}{\sim} \frac{15}{8n^2}$, et donc, toujours par comparaison à une série de Riemann, $\sum \ln v_n$ converge.
Ainsi, $\alpha_0 = \frac{3}{2}$.

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.

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

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

Pour $\alpha = \alpha_0$, la série $\sum \ln v_n$ converge et donc la suite de ses sommes partielles admet une limite finie ℓ.
Et donc $\ln\left((n+1)^{\alpha_0} u_{n+1}\right) - \ln\frac{2}{5} \underset{n\to+\infty}{\longrightarrow} \ell$.
On en déduit, par continuité de l'exponentielle que

$$(n+1)^{\alpha_0} u_{n+1} \underset{n\to+\infty}{\longrightarrow} \frac{2}{5}e^\ell.$$

Ainsi, si on pose $C = \frac{2}{5}e^\ell > 0$, il vient $u_n \underset{n\to+\infty}{\sim} \frac{C}{n^{\alpha_0}}$.
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 $k\sqrt{k}u_k \underset{k\to+\infty}{\sim} C \underset{k\to+\infty}{\longrightarrow} C$.
Et donc la suite $(k\sqrt{k}u_k)$ étant convergente, elle est bornée.
Soit donc $M \in \mathbf{R}$ tel que pour tout $k \in \mathbf{N}$, $k\sqrt{k}u_k \leqslant M \Leftrightarrow ku_k \leqslant \frac{M}{\sqrt{k}}$.
Alors, $\sum_{k=1}^{n} ku_k \leqslant M\sum_{k=1}^{n}\frac{1}{\sqrt{k}}$.
Mais, par décroissance de la fonction racine carrée sur $[k-1,k]$, on a, pour $k \geqslant 2$ et pour tout $t \in [k-1,k]$, $\frac{1}{\sqrt{k}} \leqslant \frac{1}{\sqrt{t}}$.
Par conséquent, par croissance de l'intégrale,

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

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

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

Relation de Chasles.

Mais cette dernière intégrale est facile à calculer :

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

Et donc $\sum_{k=1}^{n} \frac{1}{\sqrt{k}} \leqslant 2\sqrt{n} - 2 + 1 \leqslant 2\sqrt{n}$.
En en déduit donc que

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

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

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

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, $u_{n+1} \underset{n\to+\infty}{\longrightarrow} 0$, et d'après l'équivalent obtenu à la question 2.b, $(n+1)u_{n+1} \underset{n\to+\infty}{\sim} \frac{C}{\sqrt{n+1}} \underset{n\to+\infty}{\longrightarrow} 0$.
Ainsi, en passant à la limite, il vient $\sum_{k=1}^{+\infty} u_k = 2$.
En ajoutant $u_0 = 1$ on obtient donc enfin $\sum_{n=0}^{+\infty} u_n = 3$.

□

2.3.2 Séries alternées

Un résultat relativement classique (mais hors programme) sur la convergence des séries est le suivant (appelé critère de Leibniz) : si u_n est une suite décroissante de limite nulle, alors la série de terme général $(-1)^n u_n$ est convergente.
Il permet de prouver la convergence de certaines séries qui ne sont pas absolument convergentes, comme par exemple $\sum \frac{(-1)^n}{n}$.
Il n'est pas rare que des exercices demandent de le prouver soit en toute généralité, soit dans des cas particuliers. Dans les deux cas, la preuve utilise le fait que les deux suites des sommes partielles d'ordre pair et d'ordre impair sont adjacentes (voir les deux premières questions de l'exercice suivant).
On notera également (question 1.b de l'exercice suivant) qu'on dispose d'une majoration des restes de cette série.

Exercice 2.25 (ESCP 2017) [ESCP17.1.03] **Moyen**

Autour du reste de certaines séries alternées.
Abordable en première année : ✓

1. Soit $(b_p)_{p\in\mathbf{N}}$ une suite réelle décroissante et de limite nulle. Pour tout $n \in \mathbf{N}$, on pose $S_n = \sum_{p=0}^{n}(-1)^p b_p$.

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

   (b) En déduire que la suite $(S_n)_{n\in\mathbf{N}}$ converge vers une limite ℓ et que

   $$\forall k \in \mathbf{N},\ |\ell - S_k| \leqslant b_{k+1}.$$

   (c) Montrer que, pour tout $n \in \mathbf{N}$, le réel $\sum_{p=n+1}^{+\infty}(-1)^p b_p$ est bien défini.

Soit $f : \mathbf{R}_+ \to \mathbf{R}$ une fonction positive, dérivable, décroissante, convexe et telle que $\lim_{x\to+\infty} f(x) = 0$.

2. Pour tout $p \in \mathbf{N}$, on pose $a_p = f(p) - f(p+1)$.

   (a) Montrer que, pour tout $n \in \mathbf{N}$, le nombre $u_n = \sum_{p=n+1}^{+\infty}(-1)^p f(p)$ est bien défini.

   (b) Pour tout $n \in \mathbf{N}$, montrer que : $\sum_{p=n+1}^{+\infty}(-1)^p a_p = 2u_n - (-1)^{n+1} f(n+1)$.

   (c) Montrer que la suite $(a_p)_{p\in\mathbf{N}}$ est décroissante.

3. (a) Montrer que, pour tout $n \in \mathbf{N}$, $\left|\sum_{p=n+1}^{+\infty}(-1)^p a_p\right| \leqslant f(n+1) - f(n+2)$.

   (b) En déduire que la série de terme général u_n converge.

SOLUTION DE L'EXERCICE 2.25

1.a. Pour $n \in \mathbf{N}$, on a

$$S_{2n} - S_{2(n+1)} = \sum_{p=0}^{2n}(-1)^p b_p - \sum_{p=0}^{2n+2}(-1)^p b_p = b_{2n+1} - b_{2n+2}$$

qui est positif car $(b_p)_{p\in\mathbf{N}}$ est décroissante.
Donc $(S_{2n})_{n\in\mathbf{N}}$ est décroissante.
De même, $S_{2n+1} - S_{2n+3} = b_{2n+3} - b_{2n+2} \leqslant 0$, donc $(S_{2n+1})_{n\in\mathbf{N}}$ est croissante.

1.b. On a $S_{2n+1} - S_{2n} = -b_{2n+1} \underset{n\to+\infty}{\longrightarrow} 0$.
Donc par ce qui précède, les deux suites $(S_{2n})_{n\in\mathbf{N}}$ et $(S_{2n+1})_{n\in\mathbf{N}}$ sont adjacentes. Elles convergent donc vers une même limite ℓ.
Et alors, $(S_n)_{n\in\mathbf{N}}$ est également convergente, de limite ℓ.
On a alors, pour tout $n \in \mathbf{N}$, $S_{2n+1} \leqslant \ell \leqslant S_{2n+2} \leqslant S_{2n}$.
Et donc en particulier,

$$0 \leqslant \ell - S_{2n+1} \leqslant S_{2n+2} - S_{2n+1} = b_{2n+2}.$$

Donc $|S_{2n+1} - \ell| = \ell - S_{2n+1} \leqslant b_{2n+2}$.
De même, $S_{2n+1} - S_{2n} \leqslant \ell - S_{2n} \leqslant 0$, de sorte que

$$|S_{2n} - \ell| = S_{2n} - \ell \leqslant S_{2n} - S_{2n+1} = b_{2n+1}.$$

Ainsi, que k soit pair[1], ou que k soit impair[2], on a prouvé que

$$|S_k - \ell| \leqslant b_{k+1}.$$

Rappel
Une suite (u_n) converge si et seulement les deux suites (u_{2n}) et (u_{2n+1}) convergent vers **la même limite**.

[1] C'est-à-dire de la forme $k = 2n$.

[2] C'est-à-dire de la forme $k = 2n+1$.

1.c. Notons que la suite (S_n) n'est autre que la suite des sommes partielles de la série de terme général $(-1)^p b_p$.
Et donc puisque (S_n) converge, la série $\sum_p (-1)^p b_p$ converge.

Par conséquent, son reste d'ordre n, $\sum_{p=n+1}^{+\infty}(-1)^p b_p$ est bien défini.

2.a. La suite $(f(p))_{p\in\mathbf{N}}$ est positive, décroissante et de limite nulle, donc les résultats de la question 1 s'appliquent, et en particulier celui de la question 1.c, de sorte que u_n est bien défini.

2.b. Notons que $(-1)^p a_p = (-1)^p f(p) + (-1)^{p+1} f(p+1)$, et donc la série de terme général $(-1)^p a_p$ converge car somme des deux séries convergentes (voir la question précédente).
On a alors

$$\begin{aligned}\sum_{p=n+1}^{+\infty}(-1)^p a_p &= \sum_{p=n+1}^{+\infty}(-1)^p f(p) + \sum_{p=n+1}^{+\infty}(-1)^{p+1} f(p+1)\\ &= \sum_{p=n+1}^{+\infty}(-1)^p f(p) + \sum_{k=n+2}^{+\infty}(-1)^k f(k)\\ &= \sum_{p=n+1}^{+\infty}(-1)^p f(p) + \sum_{k=n+1}^{+\infty}(-1)^k f(k) - (-1)^{n+1} f(n+1)\\ &= 2u_n - (-1)^{n+1} f(n+1).\end{aligned}$$

2.c. D'après l'égalité de accroissements finis, pour tout $p \in \mathbf{N}$, il existe un réel $c_p \in]p, p+1[$ tel que

$$f(p+1) - f(p) = (p+1-p)f'(c_p) = f'(c_p).$$

Et donc $a_p = -f'(c_p)$.
Mais puisque f est convexe et dérivable, f' est croissante, de sorte que $f'(c_p) \leqslant f'(c_{p+1})$ et donc $a_p = -f'(c_p) \geqslant -f'(c_{p+1}) = a_{p+1}$.
Et donc (a_p) est décroissante.

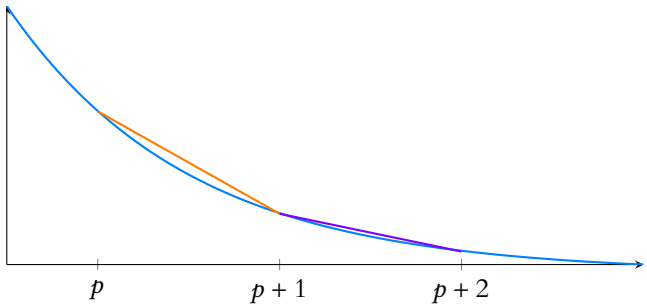


FIGURE 2.8 – La convexité de f implique que la pente de la corde passant par les points d'abscisses p à $p+1$, qui vaut $-a_p$ est plus faible que la pente de la corde passant par les points d'abscisses $p+1$ et $p+2$, qui est $-a_{p+1}$.

3.a. Puisque (a_p) est décroissante, positive, et de limite nulle[3], les résultats de la question 1 s'appliquent.
Et en particulier, d'après l'inégalité de 1.b, pour tout $n \in \mathbf{N}$,

$$\left|\sum_{p=n+1}^{+\infty}(-1)^p a_p\right| = \left|\sum_{p=0}^{n}(-1)^p b_p - \sum_{p=0}^{+\infty}(-1)^p b_p\right| \leqslant a_{n+1} = f(n+1) - f(n+2).$$

[3] Car différence de deux suites de limite nulle.

Détails
La somme $\sum_{p=0}^{+\infty}(-1)^p b_p$ n'est autre que le ℓ de la question 1.b.

3.b. D'après la question 2.b, on a donc

$$u_n = \frac{1}{2}\sum_{p=n+1}^{+\infty}(-1)^p a_p + \frac{1}{2}(-1)^{n+1} f(n+1).$$

Or, la série de terme général $(-1)^{n+1} f(n+1)$ converge, toujours par application de la question 1.
D'autre part, la série de terme général $f(n+1) - f(n+2)$ converge. En effet, il s'agit d'une série télescopique :

$$\sum_{n=0}^{N} f(n+1) - f(n+2) = f(1) - f(N+2) \underset{N\to+\infty}{\longrightarrow} f(1).$$

Danger !
Attention à ne pas conclure trop vite que toute série téelscopique est convergente, par exemple $\sum((n+1)-n)$ diverge.
Ici, on a tout de même utilisé le fait que f soit de limite nulle en $+\infty$.
Il est conseillé de systématiquement passer par les sommes partielles, comme nous l'avons fait ici.

Et donc si on pose $v_n = \sum_{p=n+1}^{+\infty} (-1)^p a_p$, alors le résultat de la question 3.a prouve que $|v_n| \leqslant f(n+1) - f(n+2)$, et donc la série de terme général v_n est absolument convergente, et donc convergente.
Ainsi, $u_n = \frac{1}{2}v_n + \frac{1}{2}(-1)^{n+1} f(n+1)$, et donc $\sum_n u_n$ est convergente car somme de deux séries convergentes.

□

2.4 INTÉGRALES

2.4.1 À classer

Attention, notamment sur les questions sans préparation, à ne pas tomber dans les pièges qui sont parfois tendus par l'énoncé.
Si un exercice vous semble trivial, c'est sûrement qu'il y a anguille sous roche !

EXERCICE 2.26 (QSP ESCP 2016) [ESCP16.QSP02] **Moyen**

Nature de $\int_1^{+\infty} \frac{f(t)}{t}\,dt$.

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.

SOLUTION DE L'EXERCICE 2.26 Un premier réflexe serait de dire que pour tout $t \geqslant 1$, $\frac{f(t)}{t} \leqslant f(t)$ et de penser que ceci suffit à conclure.
Malheureusement cette inégalité n'est vraie que si $f(t) \geqslant 0$, 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 $\lim_{x\to+\infty} f(x) = 0$, 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 $+\infty$.

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 $+\infty$.

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 $\mathscr{C}^1$.
Une intégration par parties nous donne alors, pour tout $x \geqslant 1$,

$$\int_1^x \frac{f(t)}{t}\,dt = \left[\frac{F(t)}{t}\right]_1^x + \int_1^x \frac{F(t)}{t^2}\,dt = \frac{F(x)}{x} + \int_1^x \frac{F(t)}{t^2}\,dt.$$

Mais puisque $\int_1^{+\infty} f(t)\,dt$ converge, $F(x)$ admet une limite ℓ lorsque $x \to +\infty$.
Et alors, par quotient de limites, $\frac{F(x)}{x} \underset{x\to+\infty}{\longrightarrow} 0$.
D'autre part, il existe[1] $A > 1$ tel que pour $x \geqslant A$, $\ell - 1 \leqslant F(x) \leqslant \ell + 1$.
Notons alors $M = \max(|\ell - 1|, |\ell + 1|)$, de sorte que pour tout $t \geqslant A$, $|F(t)| \leqslant M$.
Par conséquent, pour tout $t \geqslant A$, $\left|\frac{F(t)}{t^2}\right| \leqslant \frac{M}{t^2}$.
Et donc par comparaison à une intégrale de Riemann, $\int_A^{+\infty} \frac{F(t)}{t^2}\,dt$ converge absolument

[1] C'est la définition de limite, dans laquelle on a pris $\varepsilon = 1$.

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 $x \to +\infty$.

Et donc $\int_1^x \frac{f(t)}{t}\,dt$ admet une limite lorsque $x \to +\infty$ de sorte que $\int_1^{+\infty} \frac{f(t)}{t}\,dt$ converge. □

Exercice 2.27 (QSP HEC 2007) [HEC07.QSP01] **Facile**

Intégrale d'une fonction périodique
Abordable en première année : ✓

Donner un équivalent lorsque $x \to +\infty$ de $\int_0^x |\sin t|\,dt$.

Solution de l'exercice 2.27 Notons que la fonction $t \mapsto |\sin t|$ est π-périodique car $|\sin(x+\pi)| = |-\sin(x)| = |\sin(x)|$.
Ainsi, pour tout $k \in \mathbf{N}$,

$$\int_{k\pi}^{(k+1)\pi} |\sin(t)|\,dt = \int_0^{\pi} |\sin t|\,dt = \int_0^{\pi} \sin t\,dt = 2.$$

Détails
Pour prouver la première égalité, on peut procéder au changement de variable $t = x - k\pi$.

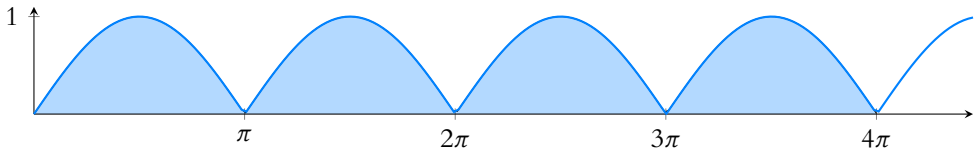


Figure 2.9 – Toutes les «arches» ont la même aire.

Dans ce qui suit, l'idée est que $\int_0^x |\sin t|\,dt$ doit être environ égal à l'aire d'une arche, fois le nombre d'arches qui se trouvent entièrement entre 0 et x.

Autrement dit
k est le nombre d'arches qui sont entièrement contenues entre les droites verticales d'abscisses 0 et x.

Soit donc $x \in \mathbf{R}$, et soit $k \in \mathbf{N}$ l'unique réel tel que $k\pi \leqslant x < (k+1)\pi$. Alors

$$\int_0^x |\sin t|\,dt = \sum_{i=0}^{k-1} \underbrace{\int_{i\pi}^{(i+1)\pi} |\sin t|\,dt}_{=2} + \int_{k\pi}^{x} |\sin t|\,dt = 2k + \int_{k\pi}^{x} |\sin t|\,dt.$$

Mais, par croissance de l'intégrale,

$$0 \leqslant \int_{k\pi}^{x} |\sin t|\,dt \leqslant \int_{k\pi}^{(k+1)\pi} |\sin t|\,dt \leqslant 2.$$

Et donc

$$2k \leqslant \int_0^x |\sin t|\,dt \leqslant 2k+2.$$

Remarque
Pour le dire autrement,
$k = \left\lfloor \frac{x}{\pi} \right\rfloor$.

Or, on a $\frac{x}{\pi} - 1 < k \leqslant \frac{x}{\pi}$, de sorte que

$$2\left(\frac{x}{\pi} - 1\right) \leqslant 2k \leqslant \int_0^x |\sin t|\,dt \leqslant 2k+2 \leqslant 2\frac{x}{\pi} + 2.$$

En divisant par $\frac{2x}{\pi}$, il vient

$$1 - \frac{\pi}{x} \leqslant \frac{\pi}{2x} \int_0^x |\sin t|\,dt \leqslant 1 + \frac{\pi}{x}.$$

Par le théorème des gendarmes, on a donc $\lim_{x\to+\infty} \frac{\pi}{2x}\int_0^x |\sin t|\,dt = 1$ et donc

$$\int_0^x |\sin t| \underset{x\to+\infty}{\sim} \frac{2x}{\pi}.$$

Généralisation *: ce résultat peut aisément se généraliser à toute fonction f qui soit continue et T-périodique, à condition que l'intégrale de f sur une période ne soit pas nulle. On a alors*

$$\int_0^x f(t)\,dt \underset{x\to+\infty}{\sim} \frac{x}{T}\int_0^T f(t)\,dt.$$

□

EXERCICE 2.28 (QSP ESCP 2009) [ESCP09.QSP01] — Moyen

Deux densités continues s'intersectent nécessairement.
Abordable en première année : ✓

Soit f une fonction continue et positive sur $\mathbf{R}_+$ telle que $\int_0^{+\infty} f(t)\,dt = 1$.
Montrer qu'il existe $\alpha \in \mathbf{R}_+$ tel que $f(a) = e^{-a}$.

SOLUTION DE L'EXERCICE 2.28

Soit φ la fonction définie sur $\mathbf{R}_+$ par $\varphi(x) = \int_0^x \left(f(t) - e^{-t}\right) dt$.
Alors φ est de classe $\mathcal{C}^1$ sur $\mathbf{R}_+$ et $\varphi'(x) = f(x) - e^{-x}$.
De plus, on a $\varphi(0) = \int_0^0 \left(f(t) - e^{-t}\right) dt = 0$ et

$$\lim_{x\to+\infty} \varphi(x) = \int_0^{+\infty} f(t)\,dt - \int_0^{+\infty} e^{-t}\,dt = 1 - 1 = 0.$$

Si φ' ne s'annulait pas, elle serait de signe constant.
Supposons par exemple que $\varphi'(x) > 0$, pour tout $x \geqslant 0$. Alors φ est strictement croissante sur $\mathbf{R}_+$. En particulier, pour tout $x \geqslant 1$, on a $\varphi(x) \geqslant \varphi(1) > 0$, et donc $\lim_{x\to+\infty} \varphi(x) \geqslant \varphi(1) > 0$.
Ceci est impossible.
On montre de même qu'il n'est pas possible que φ' soit strictement négative sur $\mathbf{R}_+$. Et donc φ' s'annule en au moins un réel $a \in \mathbf{R}_+$. Et donc $f(a) - e^{-a} = 0 \Leftrightarrow f(a) = e^{-a}$.
Remarque *: on pourrait tenir le même raisonnement en remplaçant $t \mapsto e^{-t}$ par n'importe quelle autre densité continue sur $\mathbf{R}_+$ et nulle sur $\mathbf{R}_-^*$.* □

Soyons rapides !
On pourrait calculer l'intégrale

$$\int_0^{+\infty} e^{-t}\,dt$$

mais si on reconnaît la densité d'une loi exponentielle $\mathcal{E}(1)$, ou encore si on reconnaît $\Gamma(1)$, il est immédiat que cette intégrale vaut 1.

EXERCICE 2.29 (ESCP 2012) [ESCP12.1.04] — Moyen

Fonction bêta d'Euler.

Pour tous réels strictement positifs x et y, on pose $B(x,y) = \int_0^1 t^{x-1}(1-t)^{y-1}\,dt$.

1. Prouver la convergence de l'intégrale définissant $B(x,y)$.
2. (a) Prouver que $\forall (x,y) \in (\mathbf{R}_+^*)^2, B(x,y) = B(y,x)$.
   (b) Montrer que $\forall (x,y) \in (\mathbf{R}_+^*)^2, B(x+1,y) = \frac{x}{y}B(x,y+1)$.
3. Montrer que $\forall (x,y) \in (\mathbf{R}_+^*)^2, B(x,y+1) = B(x,y) - B(x+1,y)$. En déduire que $B(x+1,y) = \frac{x}{x+y}B(x,y)$.
4. Soit n un entier naturel non nul, et soit $x > 0$.
   (a) Étudier le signe sur $[0,1]$ de la fonction $g : t \mapsto e^{-t} - 1 + t$. En déduire que pour tout $t \in [0,n]$, on a

$$\left(1 - \frac{t}{n}\right)^n \leqslant e^{-t}.$$

(b) Montrer que pour tout $t \in [0,n]$, on a $\left(1-\frac{t^2}{n}\right)e^{-t} \leqslant \left(1-\frac{t}{n}\right)^n$.

(c) Montrer que $\lim_{n\to\infty} \int_0^n \left(1-\frac{t}{n}\right)^n t^{x-1}\,dt = \int_0^\infty t^{x-1}e^{-t}\,dt = \Gamma(x)$.

5. Montrer que $\forall x > 0$, $\Gamma(x) = \lim_{n\to\infty} \frac{n^x n!}{x(x+1)\cdots(x+n)}$.

Solution de l'exercice 2.29

1. La fonction $t \mapsto t^{x-1}(1-t)^{y-1}$ est continue et positive sur $]0,1[$.

Au voisinage de 0, on a $t^{x-1}(1-t)^{y-1} \underset{t\to 0}{\sim} t^{x-1}$ et $\int_0^{1/2} t^{x-1}\,dt$ est une intégrale de Riemann convergente, donc, par le critère de comparaison pour les fonctions positives,

$$\int_0^{1/2} t^{x-1}(1-t)^{y-1}\,dt \text{ est convergente.}$$

De même, au voisinage de 1, $t^{x-1}(1-t)^{y-1} \underset{t\to 1}{\sim} (1-t)^{y-1}$.

Mais $\int_{1/2}^1 (1-t)^{y-1}\,dt$ est une intégrale de Riemann convergente et donc $\int_{1/2}^1 t^{x-1}(1-t)^{y-1}\,dt$ converge.

On en déduit que l'intégrale définissant $B(x,y)$ est convergente.

> **Remarque**
> Si $x \geqslant 1$ et $y \geqslant 1$, alors l'intégrande est continue sur le segment $[0,1]$, mais lorsque $x < 1$ ou $y < 1$ ce n'est plus vrai et donc il faut faire une vraie étude de convergence.

> **Astuce**
> En cas de doute pour les intégrales de Riemann de la forme $\int_a^b \frac{dt}{(t-a)^\alpha}$, faire le changement de variable $x = t-a$.
> Ici on pourrait poser $t = 1-x$.

2.a. Procédons au changement de variable $u = 1-t$ (qui est légitime car l'intégrale définissant $B(x,y)$ converge). On a alors $u = 1$ pour $t = 0$ et $u = 0$ pour $t = 1$. De plus, $du = -dt$, et alors

$$B(x,y) = \int_0^1 t^{x-1}(1-t)^{y-1}\,dt = -\int_1^0 (1-u)^{x-1}u^{y-1}\,du = \int_0^1 u^{y-1}(1-u)^{x-1}du = B(y,x).$$

2.b. On a, par intégration par parties sur un segment de la forme $[a;b] \subset]0,1[$,

$$\int_a^b t^x(1-t)^{y-1}\,dt = \left[-\frac{1}{y}t^x(1-t)^{y-1}\right]_a^b + \frac{x}{y}\int_a^b t^{x-1}(1-t)^y\,dt$$

En prenant la limite lorsque $a \to 0^+$, il vient

$$\int_0^b t^x(1-t)^{y-1}\,dt = -\frac{1}{y}b^x(1-b)^{y-1} + \frac{x}{y}\int_0^b t^{x-1}(1-t)^y\,dt$$

Puis en prenant la limite lorsque $b \to 0^+$

$$B(x+1,y) = \int_0^1 t^x(1-t)^{y-1}\,dt = \frac{x}{y}\int_0^1 t^{x-1}(1-t)^y\,dt = \frac{x}{y}B(x,y).$$

3. On a

$$B(x,y+1) = \int_0^1 t^{x-1}(1-t)^y\,dt = \int_0^1 t^{x-1}(1-t)(1-t)^{y-1}\,dt = \int_0^1 t^{x-1}(1-t)^{y-1}\,dt - \int_0^1 t^x(1-t)^{y-1}\,dt = B(x,y)-B(x+1,y).$$

On en déduit que

$$B(x+1,y) = \frac{x}{y}B(x,y) = \frac{x}{y}B(x,y) - \frac{x}{y}B(x+1,y)$$

soit

$$\frac{x+y}{y}B(x+1,y) = \frac{x}{y}B(x,y) \Leftrightarrow B(x+1,y) = \frac{x}{x+y}B(x,y).$$

4.a. Soit $g : t \mapsto e^{-t} - 1 + t$. Alors g est dérivable et $g'(t) = -e^{-t} + 1 \geqslant 0$, de sorte que g est croissante sur $[0,1]$.

De plus, $g(0) = 0$, et donc g est positive sur $[0,1]$.

On en déduit que pour $t \in [0,n]$, puisque $\frac{t}{n} \in [0,1]$, $1-\frac{t}{n} \leqslant e^{-t/n}$ et donc

$$\left(1-\frac{t}{n}\right)^n \leqslant \left(e^{-t/n}\right)^n = e^{-t}.$$

4.b. Si $t \in [\sqrt{n}, n]$, la relation est évidente car le membre de gauche est négatif lorsque celui de droite est positif.
Supposons donc que $t \in [0, \sqrt{n}]$. L'inégalité à prouver est équivalente (par croissance du logarithme) à

$$-t + \ln\left(1 - \frac{t^2}{n}\right) \leqslant n \ln\left(1 - \frac{t}{n}\right).$$

Étudions donc la fonction

$$h : t \mapsto -t + \ln\left(1 - \frac{t^2}{n}\right) - n \ln\left(1 - \frac{t}{n}\right)$$

Elle est dérivable sur $[0, \sqrt{n}]$, et on a

$$h'(t) = -1 - \frac{2t}{n - t^2} + \frac{n}{n - t} = -\frac{t((t-1)^2 + (n-1))}{(t^2 - n)(t - n)}.$$

Ainsi, elle est négative sur $[0, \sqrt{n}]$, de sorte que h y est décroissante. Comme $h(0) = 0$, on en déduit que

$$\forall t \in [0, \sqrt{n}], h(t) \leqslant 0 \Leftrightarrow -t + \ln\left(1 - \frac{t^2}{n}\right) \leqslant n \ln\left(1 - \frac{t}{n}\right) \Leftrightarrow \left(1 - \frac{t^2}{n}\right) e^{-t} \leqslant \left(1 - \frac{t}{n}\right)^n.$$

4.c. Des deux questions précédentes, il vient, pour $x > 0$

$$\int_0^n \left(1 - \frac{t^2}{n}\right) e^{-t} t^{x-1} dt \leqslant \int_0^n \left(1 - \frac{t}{n}\right)^n t^{x-1} dt \leqslant \int_0^n e^{-t} t^{x-1} dt$$

Mais par définition,

$$\lim_{n\to\infty} \int_0^n t^{x-1} e^{-t} dt = \int_0^\infty t^{x-1} e^{-t} dt = \Gamma(x).$$

De plus,

$$\int_0^n \left(1 - \frac{t^2}{n}\right) t^{x-1} e^{-t} dt = \int_0^n t^{x-1} e^{-t} - \frac{1}{n} \int_0^n t^{x+1} e^{-t} dt \underset{n\to\infty}{\longrightarrow} \Gamma(x) - 0\Gamma(x+2) = \Gamma(x).$$

Par le théorème des gendarmes, on en déduit que

$$\lim_{n\to\infty} \int_0^n \left(1 - \frac{t}{n}\right)^n = \Gamma(x).$$

5. En posant $u = \frac{t}{n}$, il vient

$$\int_0^n \left(1 - \frac{t}{n}\right)^n t^{x-1} dt = \int_0^1 (1-u)^n (un)^{x-1} n du = n^x \int_0^1 (1-u)^n u^{x-1} du = n^x B(x, n+1) = n^x B(n+1, x)$$

D'après la formule obtenue à la question 3, il vient

$$B(n+1, x) = \frac{n}{n+x} B(n, x) = \frac{n(n-1)}{(n+x)(n-1+x)} B(n-1, x) \cdots = \frac{n!}{(x+1)\cdots(x+n)} B(1, x)$$
$$= \frac{n!}{(x+1)\cdots(x+n)} \int_0^1 t^{x-1} dt = \frac{n!}{x(x+1)\cdots(x+n)}.$$

On en déduit par la question 4.*c* que

$$\Gamma(x) = \lim_{n\to\infty} \frac{n^x n!}{x(x+1)\cdots(x+n)}.$$

□

2.4.2 Fonctions définies par une intégrale

1.01 ESCP 2014 1.18 ESCP 2014 (intégrale dépendant de ses bornes)
Des exercices récurrents à l'oral (mais aussi à l'écrit) font apparaître des fonctions définies par une intégrale.

Exercice 2.30 (ESCP 2015) [ESCP15.1.03] — Moyen

Étude la fonction $x \mapsto \displaystyle\int_0^{+\infty} \frac{e^{-xt^2}}{1+t}\,dt$.

1. Soit x un réel. Démontrer que l'intégrale $\displaystyle\int_0^{+\infty} \frac{e^{-xt^2}}{1+t}\,dt$ est convergente si et seulement si x est strictement positif. On pose alors :
$$\forall x \in \mathbf{R}_+^*,\ F(x) = \int_0^{+\infty} \frac{e^{-xt^2}}{1+t}\,dt.$$
L'exercice a pour but l'étude de la fonction F.
2. Étudier le sens de variation de la fonction F.
3. (a) Démontrer que pour tout $x > 0$, on a $F(x) \geqslant \dfrac{1}{e}\displaystyle\int_0^{\frac{1}{\sqrt{x}}} \frac{1}{1+t}\,dt$. En déduire la limite de F en 0.
(b) Démontrer que pour tout $x > 0$, on a : $F(x) = \displaystyle\int_0^{+\infty} \frac{e^{-u^2}}{\sqrt{x}+u}\,du$. En déduire la limite de F en $+\infty$.
4. Prouver que pour tous x_0 et x réels strictement positifs, on a :
$$|F(x) - F(x_0)| \leqslant \frac{|\sqrt{x} - \sqrt{x_0}|}{\sqrt{x}\sqrt{x_0}} \times \frac{\sqrt{\pi}}{2}.$$
En déduire que F est continue sur $\mathbf{R}_+^*$.
5. Démontrer que $\displaystyle\lim_{x\to+\infty} \sqrt{x}\left(\sqrt{x}F(x) - \frac{\sqrt{\pi}}{2}\right) = -\frac{1}{2}$.
En déduire que $F(x) = \dfrac{\sqrt{\pi}}{2\sqrt{x}} - \dfrac{1}{2x} + o\left(\dfrac{1}{x}\right)$ au voisinage de $+\infty$.

Solution de l'exercice 2.30

1. La fonction $t \mapsto \dfrac{e^{-xt^2}}{1+t}$ est continue sur $[0,+\infty[$, de sorte que le seul problème de convergence éventuel est au voisinage de $+\infty$.

Si $x < 0$, alors $e^{-xt^2} \underset{t\to+\infty}{\longrightarrow} +\infty$, et donc pour t suffisamment grand, on a $\dfrac{e^{-xt^2}}{1+t} \geqslant \dfrac{1}{1+t}$.

Or, $\dfrac{1}{1+t} \underset{t\to+\infty}{\sim} \dfrac{1}{t}$ et $\displaystyle\int_1^{+\infty} \frac{1}{t}\,dt$ diverge.

On en déduit donc que $\displaystyle\int_0^{+\infty} \frac{e^{-xt^2}}{1+t}\,dt$ diverge.

De même, si $x = 0$, on a $\dfrac{e^{-xt^2}}{1+t} = \dfrac{1}{1+t}$ et donc $\displaystyle\int_0^{+\infty} \frac{e^{-xt^2}}{1+t}\,dt$ diverge pour les mêmes raisons.

Enfin, si $x > 0$, on a, par croissance comparée :
$$t^3 \frac{e^{-xt^2}}{1+t} \underset{t\to+\infty}{\sim} t^2 e^{-xt^2} \underset{t\to+\infty}{\longrightarrow} 0.$$
Et donc $\dfrac{e^{-xt^2}}{1+t} = \underset{t\to+\infty}{o}\left(\dfrac{1}{t^3}\right)$. Puisque $\displaystyle\int_1^{+\infty} \frac{dt}{t^3}$ est une intégrale[1] convergente, il en est de même de $\displaystyle\int_1^{+\infty} \frac{e^{-xt^2}}{1+t}\,dt$. Et $\displaystyle\int_0^1 \frac{e^{-xt^2}}{1+t}\,dt$ converge car intégrale d'une fonction continue sur un segment, de sorte que $\displaystyle\int_0^{+\infty} \frac{e^{-xt^2}}{1+t}\,dt$ converge.

Ainsi, nous avons bien prouvé que $\displaystyle\int_0^{+\infty} \frac{e^{-xt^2}}{1+t}\,dt$ converge si et seulement si $x > 0$.

2. Pour $0 < x \leqslant y$, et pour tout $t \geqslant 0$, on a $xt^2 \leqslant yt^2$ et donc, par décroissance de $u \mapsto e^{-u}$, $e^{-xt^2} \geqslant e^{-yt^2}$.

> **Remarque**
> On aurait pu prendre t^2 au lieu de t^3, mais la croissance comparée aurait alors nécessité quelques détails, par exemple un changement de variable en posant $T = t^2$.

[1] De Riemann.

En multipliant par $\frac{1}{1+t} \geqslant 0$, il vient $\frac{e^{-xt^2}}{1+t} \geqslant \frac{e^{-yt^2}}{1+t}$.
Par croissance de l'intégrale, il vient donc

$$F(x) = \int_0^{+\infty} \frac{e^{-xt^2}}{1+t}\,dt \geqslant \int_0^{+\infty} \frac{e^{-yt^2}}{1+t}\,dt = F(y).$$

La fonction F est donc décroissante.

3.a. D'après la relation de Chasles, on a

$$F(x) = \int_0^{1/\sqrt{x}} \frac{e^{-xt^2}}{1+t}\,dt + \int_{1/\sqrt{x}}^{+\infty} \frac{e^{-xt^2}}{1+t}\,dt.$$

La seconde intégrale est positive car intégrale d'une fonction positive.
Et donc $F(x) \geqslant \int_0^{1/\sqrt{x}} \frac{e^{-xt^2}}{1+t}\,dt$.
Mais pour $t \in \left[0, \frac{1}{\sqrt{x}}\right]$, on a $t^2 \leqslant \frac{1}{x}$ et donc $xt^2 \leqslant 1$.
On en déduit que $\frac{e^{-xt^2}}{1+t} \geqslant \frac{e^{-1}}{1+t}$.
Par croissance de l'intégrale, on a donc

$$\int_0^{1/\sqrt{x}} \frac{e^{-xt^2}}{1+t}\,dt \geqslant e^{-1} \int_0^{1/\sqrt{x}} \frac{1}{1+t}\,dt$$

et donc $F(x) \geqslant \frac{1}{e} \int_0^{1/\sqrt{x}} \frac{dt}{1+t}$.
De plus, on a, pour tout $x > 0$,

$$\int_0^{1/\sqrt{x}} \frac{dt}{1+t} = \left[\ln(1+t)\right]_0^{1/\sqrt{x}} = \ln\left(1+\frac{1}{\sqrt{x}}\right) \underset{x\to 0^+}{\longrightarrow} +\infty.$$

On en déduit donc que $F(x) \underset{x\to 0^+}{\longrightarrow} +\infty$.

3.b. Soit $A > 0$. Alors, en procédant au changement de variable $t = \sqrt{x}u \Leftrightarrow u = \frac{t}{\sqrt{x}}$, on a

$$\int_0^A \frac{e^{-xt^2}}{1+t}\,dt = \int_0^{A/\sqrt{x}} \frac{e^{-u^2}}{1+\frac{u}{\sqrt{x}}}\sqrt{x}\,du = \int_0^{A/\sqrt{x}} \frac{e^{-u^2}}{u+\sqrt{x}}\,du.$$

En passant à la limite lorsque $A \to +\infty$, il vient donc

$$F(x) = \lim_{A\to+\infty} \int_0^A \frac{e^{-xt^2}}{1+t}\,dt = \lim_{A\to+\infty} \int_0^{A/\sqrt{x}} \frac{e^{-u^2}}{\sqrt{x}+u}\,du = \int_0^{+\infty} \frac{e^{-u^2}}{\sqrt{x}+u}\,du.$$

Convergence

La convergence de cette seconde intégrale est automatique car nous savons que la limite lorsque $A \to +\infty$ existe car l'intégrale définissant F converge.

Pour $x > 0$, et pour tout $u \geqslant 0$, on a $\frac{e^{-u^2}}{\sqrt{x}+u} \leqslant \frac{e^{-u^2}}{\sqrt{x}}$. Et donc par croissance de l'intégrale,

$$F(x) = \int_0^{+\infty} \frac{e^{-u^2}}{1+u}\,du \leqslant \int_0^{+\infty} \frac{e^{-u^2}}{\sqrt{x}}\,du = \frac{1}{\sqrt{x}} \int_0^{+\infty} e^{-u^2}\,du.$$

Pour calculer cette dernière intégrale, considérons une variable aléatoire X suivant la loi normale $\mathcal{N}\left(0, \frac{1}{\sqrt{2}}\right)$.
Alors une densité de X est $t \mapsto \frac{1}{\sqrt{\pi}} e^{-t^2}$. Et donc

$$P(X \geqslant 0) = \int_0^{+\infty} \frac{1}{\sqrt{\pi}} e^{-t^2}\,dt.$$

D'autre part, on a $P(X \geqslant 0) = \dfrac{1}{2}$ et donc

$$\int_0^{+\infty} e^{-t^2}\,dt = \frac{\sqrt{\pi}}{2}.$$

On en déduit donc que

$$F(x) \leqslant \frac{1}{\sqrt{x}} \int_0^{+\infty} e^{-u^2}\,du \leqslant \frac{\sqrt{\pi}}{2\sqrt{x}} \underset{x\to+\infty}{\longrightarrow} 0.$$

> **Astuce**
> Pour toute loi normale d'espérance a, on a
> $$P(X \geqslant a) = P(X \leqslant a) = \frac{1}{2}.$$
> C'est une conséquence de la symétrie de la densité autour de a.

4. Pour $u \geqslant 0$, on a

$$\frac{e^{-u^2}}{\sqrt{x}+u} - \frac{e^{-u^2}}{\sqrt{x_0}+u} = e^{-u^2}\frac{\sqrt{x_0}-\sqrt{x}}{(\sqrt{x}+u)(\sqrt{x_0}+u)}$$

et donc

$$\left|\frac{e^{-u^2}}{\sqrt{x}+u} - \frac{e^{-u^2}}{\sqrt{x_0}+u}\right| = \left|\frac{e^{-u^2}(\sqrt{x_0}-\sqrt{x})}{(\sqrt{x_0}+u)(\sqrt{x}+u)}\right| \leqslant e^{-u^2}\left|\frac{\sqrt{x_0}-\sqrt{x}}{\sqrt{x}\sqrt{x_0}}\right|.$$

Par croissance de l'intégrale, on a donc

$$\begin{aligned}
|F(x)-F(x_0)| &= \left|\int_0^{+\infty}\left(\frac{e^{-u^2}}{\sqrt{x}+u} - \frac{e^{-u^2}}{\sqrt{x_0}+u}\right)du\right| \\
&\leqslant \int_0^{+\infty}\left|\frac{e^{-u^2}}{\sqrt{x}+u} - \frac{e^{-u^2}}{\sqrt{x_0}+u}\right|du \\
&\leqslant \int_0^{+\infty} e^{-u^2}\frac{|\sqrt{x}-\sqrt{x_0}|}{\sqrt{x}\sqrt{x_0}}\,du \\
&= \frac{|\sqrt{x}-\sqrt{x_0}|}{\sqrt{x}\sqrt{x_0}}\int_0^{+\infty} e^{-u^2}\,du = \frac{|\sqrt{x}-\sqrt{x_0}|}{\sqrt{x}\sqrt{x_0}}\frac{\sqrt{\pi}}{2}.
\end{aligned}$$

En particulier, lorsque $x \to x_0$, on a $\lim\limits_{x\to x_0} |F(x)-F(x_0)| = 0 \Leftrightarrow \lim\limits_{x\to x_0} F(x) = F(x_0)$.

Et donc F est continue en x_0. Ceci étant vrai pour tout $x_0 > 0$, on en déduit que F est continue sur $\mathbf{R}_+^*$.

5. En utilisant le résultat de la question 4.b, on a

$$\begin{aligned}
\sqrt{x}\left(\sqrt{x}F(x) - \frac{\sqrt{\pi}}{2}\right) &= x\int_0^{+\infty}\frac{e^{-u^2}}{\sqrt{x}+u}\,du - \sqrt{x}\int_0^{+\infty} e^{-u^2}\,du \\
&= \int_0^{+\infty} e^{-u^2}\left(\frac{x}{\sqrt{x}+u} - \sqrt{x}\right)du \\
&= \int_0^{+\infty} e^{-u^2}\frac{-\sqrt{x}u}{\sqrt{x}+u}\,du \\
&= \int_0^{+\infty} e^{-u^2}\left(\frac{u^2}{\sqrt{x}+u} - u\right)du \\
&= \int_0^{+\infty} e^{-u^2}\frac{u^2}{\sqrt{x}+u}\,du - \int_0^{+\infty} ue^{-u^2}\,du
\end{aligned}$$

> **Convergence**
> On a bien le droit de séparer l'intégrale en deux car les deux intégrales considérées convergent.

Mais on a

$$\int_0^{+\infty} ue^{-u^2}\,du = \lim_{A\to+\infty}\int_0^A ue^{-u^2}\,du = \lim_{A\to+\infty}\left[-\frac{1}{2}e^{-u^2}\right]_0^A = \lim_{A\to+\infty}\left(\frac{1}{2} - \frac{1}{2}e^{-A^2}\right) = \frac{1}{2}.$$

Et donc

$$\sqrt{x}\left(\sqrt{x}F(x) - \frac{\sqrt{\pi}}{2}\right) + \frac{1}{2} = \int_0^{+\infty} e^{-u^2}\frac{u^2}{\sqrt{x}+u}\,du.$$

Or on a

$$0 \leqslant \int_0^{+\infty} e^{-u^2}\frac{u^2}{\sqrt{x}+u}\,du \leqslant \frac{1}{\sqrt{x}}\int_0^{+\infty} u^2e^{-u^2}\,du \underset{x\to+\infty}{\longrightarrow} 0.$$

Et donc on a bien

$$\lim_{x\to+\infty} \sqrt{x}\left(\sqrt{x}F(x) - \frac{\sqrt{\pi}}{2}\right) = -\frac{1}{2}.$$

On en déduit que

$$xF(x) - \sqrt{x}\frac{\sqrt{\pi}}{2} = -\frac{1}{2} + o(1) \Leftrightarrow F(x) = \frac{\sqrt{\pi}}{2\sqrt{x}} - \frac{1}{2x} + o\left(\frac{1}{x}\right).$$

□

EXERCICE 2.31 (ESCP 2013) [ESCP13.1.06] — Moyen

Étude de $f(x) = \displaystyle\int_0^{+\infty} \frac{t}{1+xe^t}\,dt$.

1. Montrer que pour tout $x > 0$, l'intégrale $\displaystyle\int_0^{+\infty} \frac{t}{1+xe^t}\,dt$ converge.
   On note f l'application définie sur $\mathbf{R}_+^*$ par : pour tout $x > 0$, $f(x) = \displaystyle\int_0^{+\infty} \frac{t}{1+xe^t}\,dt$.
2. Montrer que f est strictement décroissante sur $\mathbf{R}_+^*$.
3. (a) Déterminer $\lim\limits_{x\to+\infty} f(x)$, puis un équivalent de $f(x)$ pour x au voisinage de $+\infty$.
   (b) Déterminer $\lim\limits_{x\to 0} f(x)$.
4. (a) En utilisant le changement de variable $u = x\cdot e^t$ que l'on justifiera, montrer que :
   $$f(x) = \ln(x)\big(\ln(x) - \ln(1+x)\big) + \int_x^{+\infty} \frac{\ln u}{u(1+u)}\,du.$$
   (b) En déduire que f est dérivable sur $\mathbf{R}_+^*$ et calculer $f'(x)$. Retrouver ainsi le sens de variation de f.

SOLUTION DE L'EXERCICE 2.31

1. La fonction $t \mapsto \dfrac{t}{1+xe^t}$ est continue sur $[0,+\infty[$, donc le seul éventuel problème de convergence est au voisinage de $+\infty$.
Or, au voisinage de $+\infty$, on a $\dfrac{t}{1+xe^t} \underset{t\to+\infty}{\sim} \dfrac{t}{xe^t} = \dfrac{1}{x}te^{-t}$.
Puisque $\displaystyle\int_0^{+\infty} te^{-t}\,dt$ converge[1], par critère de comparaison pour les fonctions positives, $\displaystyle\int_0^{+\infty} \frac{t}{1+xe^t}\,dt$ converge.

[1] C'est $\Gamma(2)$.

2. Soient $x < y$ deux réels strictement positifs.
Alors pour tout $t \geqslant 0$, on a $\dfrac{t}{1+xe^t} \geqslant \dfrac{1}{1+ye^t}$.
D'autre part, pour $t = 1$, on a $\dfrac{1}{1+xe} > \dfrac{1}{1+ye}$.
Autrement dit, la fonction $t \mapsto \dfrac{t}{1+xe^t} - \dfrac{t}{1+ye^t}$ 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

$$0 \leqslant \frac{t}{1+xe^t} \leqslant \frac{t}{xe^t} = \frac{1}{x}te^{-t}.$$

Et donc par croissance de l'intégrale,

$$0 \leqslant f(x) \leqslant \frac{1}{x}\int_0^{+\infty} te^{-t}\,dt \leqslant \frac{1}{x}\Gamma(2) = \frac{1}{x}.$$

Danger !
Une grosse erreur ici serait de dériver (par rapport à x) ce qui se trouve sous l'intégrale et donc d'affirmer que
$f'(x) = -\displaystyle\int_0^{+\infty} \frac{te^t}{(1+xe^t)^2}\,dt$.
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.

Et donc par le théorème des gendarmes, $\lim\limits_{x\to+\infty} f(x) = 0$.

On a alors

$$\begin{aligned}|xf(x)-1| &= \left|\int_0^{+\infty} \frac{xt}{1+xe^t}\,dt - \int_0^{+\infty} te^{-t}\,dt\right| \\ &= \left|\int_0^{+\infty} t\frac{x-e^{-t}-x}{1+xe^t}\,dt\right| \\ &= \left|\int_0^{+\infty} \frac{-te^{-t}}{1+xe^t}\,dt\right| \\ &\leqslant \int_0^{+\infty} \frac{te^{-t}}{1+xe^t}\,dt \\ &\leqslant \frac{1}{x}\int_0^{+\infty} te^{-2t}\,dt.\end{aligned}$$

Et donc en particulier[2], lorsque $x \to +\infty$, $xf(x) - 1 \underset{x\to+\infty}{\longrightarrow} 0$.

Soit encore $xf(x) \underset{x\to+\infty}{\longrightarrow} 1$, de sorte que $f(x) \underset{x\to+\infty}{\sim} \frac{1}{x}$.

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.

[2] 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.

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 $\lim\limits_{x\to0^+} f(x) = +\infty$.

Supposons donc que $\lim\limits_{x\to0^+} f(x) = \ell$, avec $\ell \in \mathbf{R}$.
Puisque la fonction intégrée est positive, pour tout $A \geqslant 0$, on a

$$f(x) = \int_0^{+\infty} \frac{t}{1+xe^t}\,dt \geqslant \int_0^{A} \frac{t}{1+xe^t}\,dt.$$

Mais pour $t \in [0, A]$, $e^t \leqslant e^A$ et donc

$$\frac{t}{1+xe^A} \leqslant \frac{t}{1+xe^t}.$$

Par croissance de l'intégrale, on en déduit que

$$\int_0^{A} \frac{t}{1+xe^t}\,dt \geqslant \int_0^{A} \frac{t}{1+xe^A}\,dt = \frac{A^2}{2}\frac{1}{1+xe^A}.$$

Nous venons donc de prouver que pour tout $x \in \mathbf{R}_+^*$ et tout $A \in \mathbf{R}_+$, $f(x) \geqslant \dfrac{A^2}{2(1+xe^A)}$.

Et donc en faisant tendre x vers 0, il vient $\ell \geqslant \dfrac{A^2}{2}$.

Or cette inégalité ne peut pas être vérifiée pour tout $A \geqslant 0$, puisque $\dfrac{A^2}{2} \underset{A\to+\infty}{\longrightarrow} +\infty$.
C'est donc que notre hypothèse de départ est fausse : f n'admet pas de limite finie en 0 et donc $\lim\limits_{x\to0^+} f(x) = +\infty$.

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 !

4.a. La fonction $t \mapsto xe^t$ est $\mathscr{C}^1$ et strictement croissante sur $\mathbf{R}_+$, donc le changement de variable est bien légitime.

Posons alors $u = xe^t$, de sorte que $t = \ln\left(\dfrac{u}{x}\right) = \ln(u) - \ln(x)$ et donc $dt = \dfrac{du}{u}$.
On a donc

$$\begin{aligned}f(x) &= \int_0^{+\infty} \frac{t}{1+xe^t}\,dt \\ &= \int_x^{+\infty} \frac{\ln(u)-\ln(x)}{1+u}\,\frac{du}{u} \\ &= \int_x^{+\infty} (\ln(u)-\ln(x))\,\frac{1}{u(1+u)}\,du \\ &= \int_x^{+\infty} \frac{\ln u}{u(1+u)}\,du - \ln(x)\int_x^{+\infty}\left(\frac{1}{u}-\frac{1}{1+u}\right)du.\end{aligned}$$

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

$$\frac{1}{u(1+u)} = \frac{1}{u} - \frac{1}{1+u}.$$

Or, pour $A \geqslant x$, on a

$$\int_x^A \left(\frac{1}{u} - \frac{1}{1+u}\right) du = \left[\ln(u) - \ln(1+u)\right]_x^A = \ln(A) - \ln(1+A) + \ln(x) - \ln(1+x)$$

$$= \ln\left(\underbrace{\frac{A}{1+A}}_{\underset{A\to+\infty}{\longrightarrow} 1}\right) + \ln(x) - \ln(1+x) \underset{A\to+\infty}{\longrightarrow} \ln(x) - \ln(1+x).$$

Et donc on a bien prouvé que

$$f(x) = \ln(x)\left(\ln(x) - \ln(1+x)\right) + \int_x^{+\infty} \frac{\ln(u)}{u(1+u)}\, du.$$

Convergence

Disons tout de même deux mots de la convergence de cette dernière intégrale : au voisinage de $+\infty$,

$$\frac{\ln u}{u(1+u)} \sim \frac{\ln u}{u^2} = o\left(\frac{1}{u^{3/2}}\right).$$

4.b. Notons que $\displaystyle\int_x^{+\infty} \frac{\ln u}{u(1+u)}\, du = \int_1^{+\infty} \frac{\ln u}{u(1+u)}\, du - \int_1^x \frac{\ln u}{u(1+u)}\, du.$

Or, par le théorème fondamental de l'analyse, $x \mapsto \displaystyle\int_1^x \frac{\ln u}{u(1+u)}\, du$ est $\mathscr{C}^1$ sur $\mathbf{R}_+^*$, de dérivée égale à $x \mapsto \dfrac{\ln(x)}{x(1+x)}$.

Et donc $x \mapsto \displaystyle\int_x^{+\infty} \frac{\ln u}{u(1+u)}\, du$ est également $\mathscr{C}^1$, de dérivée égale à $x \mapsto -\dfrac{\ln(x)}{x(1+x)}$.

On en déduit que f est dérivable sur $\mathbf{R}_+^*$ car somme et produit de fonctions dérivables sur $\mathbf{R}_+^*$, avec

$$f'(x) = \frac{1}{x}\left(\ln(x) - \ln(1+x)\right) + \ln(x)\left(\frac{1}{x} - \frac{1}{1+x}\right) - \frac{\ln(x)}{x(1+x)}$$

$$= \frac{1}{x}\ln\left(\frac{x}{1+x}\right).$$

Or, pour $x > 0$, $\dfrac{x}{1+x} < 1$, de sorte que $\ln\left(\dfrac{x}{1+x}\right) < 0$, et donc $f'(x) < 0$.

On retrouve alors que f est strictement décroissante sur $\mathbf{R}_+^*$.

□

2.5 FONCTIONS DE PLUSIEURS VARIABLES

EXERCICE 2.32 (HEC 2011) [HEC11.63] — Difficile

Étude du caractère $\mathscr{C}^1$ d'une fonction définie sur $\mathbf{R}^2$.

Soit f une fonction de $\mathbf{R}$ dans $\mathbf{R}$, de classe $\mathscr{C}^1$. On définit la fonction g de $\mathbf{R}^2$ dans $\mathbf{R}$ par :

$$\forall (x,y) \in \mathbf{R}^2,\ g(x,y) = \begin{cases} \dfrac{1}{y-x}\displaystyle\int_x^y f(t)\,dt & \text{si } x \neq y \\ f(x) & \text{si } x = y \end{cases}$$

1. **Exemple.** Dans cette question seulement, on suppose que f est définie par : $\forall x \in \mathbf{R},\ f(x) = x^2$. Déterminer la fonction g correspondante et montrer que g admet un minium global sur $\mathbf{R}^2$.
2. Soit $D = \{(x,y) \in \mathbf{R}^2 : x \neq y\}$. Montrer que g est de classe $\mathscr{C}^1$ sur D et calculer ses dérivées partielles sur D.
3. Soit $a \in \mathbf{R}$. Montrer que g admet des dérivées partielles du premier ordre en (a,a) et les exprimer en fonction de $f'(a)$, où f' désigne la dérivée de f.
4. Soit $a \in \mathbf{R}$ et $(x,y) \in D$.
   (a) Montrer que : $\partial_1 g(x,y) - \partial_1 g(a,a) = \dfrac{1}{(y-x)^2}\displaystyle\int_x^y (y-t)\left(f'(t) - f'(a)\right) dt$.
   (b) En déduire que : $|\partial_1 g(x,y) - \partial_1 g(a,a)| \leqslant \dfrac{1}{2}\sup\{|f'(t) - f'(a)|,\ t \in S\}$, où S désigne le segment d'extrémités x et y.
5. Déduire des questions précédentes que g est de classe $\mathscr{C}^1$ sur $\mathbf{R}^2$.

Solution de l'exercice 2.32

1. Pour $x \neq y$, on a alors

$$g(x,y) = \frac{1}{y-x}\int_x^y t^2\,dt = \frac{1}{y-x}\left[\frac{t^3}{3}\right]_x^y = \frac{1}{3}\frac{y^3-x^3}{y-x} = \frac{1}{3}\left(x^2+xy+y^2\right).$$

Pour $x = y$, on a $g(x,y) = x^2 = \frac{1}{3}\left(x^2+xy+y^2\right)$ et donc

$$\forall (x,y)\in\mathbf{R}^2,\ g(x,y) = \frac{1}{3}\left(x^2+xy+y^2\right) = \frac{1}{6}\left((x+y)^2+x^2+y^2\right).$$

Il est alors évident que $g(x,y)\geqslant 0$, car somme de carrés, et $g(x,y)=0$ si et seulement si

$$\begin{cases} x+y=0\\ x=0\\ y=0\end{cases} \Leftrightarrow (x,y)=(0,0).$$

Donc g admet un minimum global, égal à 0, et atteint uniquement en $(0,0)$.

2. Notons F une primitive de f. Alors F est de classe $\mathscr{C}^1$ sur $\mathbf{R}$ et

$$\forall (x,y)\in D,\ g(x,y) = \frac{1}{y-x}\left(F(y)-F(x)\right).$$

Or, les fonctions $(x,y)\mapsto x$ et $(x,y)\mapsto y$ sont $\mathscr{C}^1$ sur D, et donc par composition avec F, $(x,y)\mapsto F(x)$ et $(x,y)\mapsto F(y)$ sont deux fonctions $\mathscr{C}^1$ sur D.
Et alors, par somme de fonctions $\mathscr{C}^1$, $(x,y)\mapsto F(y)-F(x)$ est de classe $\mathscr{C}^1$ sur D.
Enfin, $(x,y)\mapsto y-x$ est polynomiale et donc $\mathscr{C}^1$ sur D, et ne s'y annule pas, de sorte que le quotient $(x,y)\mapsto \dfrac{F(y)-F(x)}{y-x}$ est de classe $\mathscr{C}^1$ sur D.

On a alors

$$\partial_1 g(x,y) = \frac{-f(x)(y-x)+(F(y)-F(x))}{(y-x)^2} = \frac{g(x,y)-f(x)}{y-x}.$$

En notant que $g(x,y) = \dfrac{F(y)-F(x)}{x-y}$, on constate que x et y jouent des rôles symétriques, et donc le même calcul nous donnerait $\partial_2 g(x,y) = \dfrac{g(x,y)-f(y)}{y-x}$.

3. Sous réserve d'existence de la limite, on a

$$\begin{aligned}\partial_1 g(a,a) &= \lim_{h\to 0}\frac{g(a+h,a)-g(a,a)}{h} = \lim_{h\to 0}\frac{1}{h}\left(\frac{1}{h}\int_a^{a+h} f(t)\,dt - f(a)\right)\\ &= \lim_{h\to 0}\frac{1}{h}\left(\frac{F(a+h)-F(a)}{h} - f(a)\right).\end{aligned}$$

Mais F étant $\mathscr{C}^1$, par la formule de Taylor-Young, on a

$$F(a+h) = F(a)+hF'(a)+\frac{h^2}{2}F''(a)+\underset{h\to 0}{o}(h^2).$$

Soit encore

$$\frac{F(a+h)-F(a)}{h} - f(a) = \frac{h}{2}f'(a)+\underset{h\to 0}{o}(h).$$

Et alors,

$$\frac{g(a+h,a)-g(a,a)}{h} = \frac{1}{h}\left(\frac{F(a+h)-F(a)}{h}-f(a)\right) = \frac{1}{2}f'(a)+\underset{h\to 0}{o}(1)\underset{h\to 0}{\longrightarrow}\frac{1}{2}f'(a).$$

On en déduit que $\partial_1 g(a,a)$ existe et vaut $\dfrac{1}{2}f'(a)$.
De la même manière, on a

$$\partial_2 g(a,a) = \lim_{h\to 0}\frac{g(a,a+h)-g(a,a)}{h} = \lim_{h\to 0}\frac{1}{h}\left(\frac{1}{h}\int_a^{a+h} f(t)\,dt - f(a)\right) = \frac{1}{2}f'(a).$$

Rappel

Pour tout n, on a

$$a^n - b^n = (a-b)\times\left(\sum_{k=0}^{n-1} a^{n-1-k}b^k\right).$$

Détails

C'est la définition de la dérivée partielle : c'est la dérivée (si elle existe) en a de la fonction $t\mapsto g(t,a)$.

4.a. On a

$$\partial_1 g(x,y) - \partial_1 g(a,a) = \frac{g(x,y)-f(x)}{y-x} - \frac{1}{2}f'(a) = \frac{1}{(y-x)^2}\left(\int_x^y f(t)\,dt - (y-x)f(x) - \frac{(y-x)^2}{2}f'(a)\right).$$

Or, une intégration par parties sur l'intégrale fournie par l'énoncé nous donne

$$\begin{aligned}\int_x^y (y-t)(f'(t)-f'(a))\,dt &= \left[(y-t)\left(f(t)-f'(a)t\right)\right]_x^y + \int_x^y (f(t)-f'(a)t)\,dt\\ &= -(y-x)(f(x)-f'(a)x) + \int_x^y f(t)\,dt - f'(a)\int_x^y t\,dt\\ &= -(y-x)(f(x)-f'(a)x) + \int_x^y f(t)\,dt - \frac{y^2-x^2}{2}f'(a)\\ &= \int_x^y f(t)\,dt - (y-x)f(x) - f'(a)(y-x)\left(\frac{y+x}{2}-x\right)\\ &= \int_x^y f(t)\,dt - (y-x)f(x) - f'(a)\frac{(y-x)^2}{2}.\end{aligned}$$

Factorisation

$y^2 - x^2 = (y-x)(y+x)$.

Et donc, après division par $(y-x)^2 \neq 0$, on obtient bien

$$\begin{aligned}\frac{1}{(y-x)^2}\int_x^y (y-t)\left(f'(t)-f'(a)\right)\,dt &= \frac{1}{(y-x)^2}\left(\int_x^y f(t)\,dt - (y-x)f(x) - \frac{(y-x)^2}{2}f'(a)\right)\\ &= \partial_1 g(x,y) - \partial_1 g(a,a).\end{aligned}$$

4.b. Notons que f' étant de classe $\mathcal{C}^1$ sur le segment S d'extrémités x et y, elle y est bornée et atteint ses bornes, donc $\max\{|f'(t)-f'(a)|,\ t\in S\}$ existe bien.
Notons M ce maximum. Alors pour $t \in S$, on a

$$|y-t| \leqslant |y-x| \text{ et } |f'(t)-f'(a)| \leqslant M.$$

Notation

L'énoncé note S le segment d'extrémités x et y, et non $[x,y]$, car si $y < x$, alors $S = [y,x]$.

Et donc pour $x < y$,

$$\begin{aligned}\left|\int_x^y (y-t)(f'(t)-f'(a))\,dt\right| &\leqslant \int_x^y |y-t|\cdot|f'(t)-f'(a)|\,dt\\ &\leqslant \int_x^y M\underbrace{|y-t|}_{=y-t}\,dt\\ &\leqslant M\left[-\frac{(y-t)^2}{2}\right]_x^y\\ &\leqslant M\frac{(y-x)^2}{2}.\end{aligned}$$

Sens des bornes

C'est ici que l'hypothèse $x < y$ est importante : pour appliquer l'inégalité triangulaire, il faut que les bornes soient «dans le bon sens».

Si $y < x$, alors il faut changer le sens des bornes de l'intégrale dans l'inégalité triangulaire

$$\begin{aligned}\left|\int_x^y (y-t)(f'(t)-f'(a))\,dt\right| &\leqslant \int_x^y |y-t|\cdot|f'(t)-f'(a)|\,dt\\ &\leqslant \int_y^x M\underbrace{|y-t|}_{=t-y}\,dt\\ &\leqslant M\left[\frac{(t-y)^2}{2}\right]_y^x\\ &\leqslant M\frac{(y-x)^2}{2}.\end{aligned}$$

5. Puisque nous savons déjà que $\partial_1 g$ et $\partial_2 g$ sont définies sur $\mathbf{R}^2$ tout entier, et continues sur D, il s'agit de prouver qu'elles sont continues en (a,a), pour tout $a \in \mathbf{R}$.
Il suffit donc de prouver que $\partial_1 g$ et $\partial_2 g$ sont continues en tout point de

$$\mathbf{R}^2 \setminus D = \{(a,a),\ a\in\mathbf{R}^2\}.$$

Nous allons le faire pour $\partial_1 g$, les mêmes types de calcul pourraient s'appliquer pour $\partial_2 g$ en raison de la symétrie jouée par les rôles de x et y.
Soit donc $a \in \mathbf{R}$, et soit $\varepsilon > 0$. Il s'agit de prouver que

$$\exists \eta > 0 : \forall (x,y) \in \mathbf{R}^2,\ \|(x,y)-(a,a)\| \leqslant \eta \Rightarrow |\partial_1 g(x,y) - \partial_1 g(a,a)| \leqslant \varepsilon.$$

Autrement dit
Si (x, y) est suffisamment proche de (a, a), alors $\partial_1 g(x, y)$ est suffisamment proche de $\partial_1 g(a, a)$.

Puisque f' est continue en a, il existe $\eta > 0$ tel que pour $x \in [a-\eta, a+\eta]$, $|f'(x)-f'(a)| \leqslant 2\varepsilon$.
Et alors, pour $(x,y) \in [a-\eta, a+\eta]^2 \cap D$, par la question précédente, on a

$$|\partial_1 g(x,y) - \partial_1 g(a,a)| \leqslant \frac{1}{2}\sup\{|f'(t)-f'(a)|,\ t \in [a-\eta, a+\eta]\} \leqslant \frac{1}{2} 2\varepsilon \leqslant \varepsilon.$$

D'autre part, si b est un réel tel que $|b-a| \leqslant \eta$, alors $|f'(b)-f'(a)| \leqslant 2\varepsilon$, et donc

$$|\partial_1 g(b,b) - \partial_1 g(a,a)| = \frac{1}{2}|f'(b)-f'(a)| \leqslant \varepsilon.$$

Soit donc $(x,y) \in \mathbf{R}^2$ tel que $\|(x,y)-(a,a)\| \leqslant \eta$.
Alors $\|(x,y)-(a,a)\|^2 \leqslant \eta^2 \Leftrightarrow (x-a)^2+(y-a)^2 \leqslant \eta^2$.
En particulier, $(x-a)^2 \leqslant \eta^2$ et donc$|x-a| \leqslant \eta$. De même, on a $|y-a| \leqslant \eta$.
Et donc d'après ce qui précède :

- soit $(x,y) \in D$, et alors $|\partial_1 g(x,y) - \partial_1 g(a,a)| \leqslant \varepsilon$
- soit (x,y) est de la forme (b,b), avec $|b-a| \leqslant \varepsilon$, et alors on a également

$$|\partial_1 g(x,y) - \partial_1 g(a,a)| \leqslant \varepsilon.$$

Autrement dit, nous avons bien prouvé que pour tout $(x,y) \in \mathbf{R}^2$,

$$\|(x,y)-(a,a)\| \leqslant \eta \Rightarrow |\partial_1 g(x,y) - \partial_1 g(a,a)| \leqslant \varepsilon.$$

Et donc $\partial_1 g$ est continue en (a,a). Ceci étant vrai pour tout $a \in \mathbf{R}$, $\partial_1 g$ est continue sur $\mathbf{R}^2$.
On prouve de même que $\partial_2 g$ est continue sur $\mathbf{R}^2$, de sorte que g est une fonction de classe $\mathscr{C}^1$ sur $\mathbf{R}^2$.

□

L'exercice suivant, bien que calculatoire, n'est pas si difficile... à condition d'être à l'aise avec les notions de dérivées directionnelles, notion délicate s'il en est !
Rappelons brièvement de quoi il s'agit : si f est une fonction disons $\mathscr{C}^2$, si $x, h \in \mathbf{R}^n$, alors nous pouvons considérer la fonction g **d'une seule variable** définie par $g : t \mapsto f(x+th)$.
Alors g est de classe $\mathscr{C}^2$ (ou $\mathscr{C}^1$ si f est seulement $\mathscr{C}^1$) et ses deux première dérivées sont données par

$$g'(t) = \langle \nabla f(x+th), h\rangle \text{ et } g''(t) = \frac{1}{2} q_{x+th}(h)$$

ou q_{x+th} désigne la forme quadratique associée à la hessienne $\nabla^2 f(x+th)$ de f en $x+th$.

Exercice 2.33 (HEC 2016) [HEC16.179] — Difficile

Équation des ondes en dimension 1

1. Soient a et b deux application de $\mathbf{R}$ dans $\mathbf{R}$, de classe $\mathscr{C}^\infty$ sur $\mathbf{R}$.
   Soit f l'application de $\mathbf{R}^2$ dans $\mathbf{R}$ telle que, pour tout $(x,y) \in \mathbf{R}^2$,
   $$f(x,y) = \frac{1}{2}\left[a(x+y) + a(x-y) + \int_{x-y}^{x+y} b(s)\,ds\right].$$
   Justifier que f est de classe $\mathscr{C}^2$ sur $\mathbf{R}^2$ et montrer que $\partial^2_{1,1}(f) - \partial^2_{2,2}(f)$ est l'application nulle.
   Pour $x \in \mathbf{R}$, préciser les valeurs de $f(x,0)$ et de $\partial_2(f)(x,0)$.
2. Dans cette question, f désigne une application de $\mathbf{R}^2$ dans $\mathbf{R}$ qui est de classe $\mathscr{C}^2$ sur $\mathbf{R}^2$.
   (a) Montrer qu'il existe une unique application g de $\mathbf{R}^2$ dans $\mathbf{R}$ telle que pour tout $(x,y) \in \mathbf{R}^2$,
   $$f(x,y) = g(x+y, x-y).$$

Dans la suite, on admettra que l'application g ainsi définie est de classe $\mathscr{C}^2$ sur $\mathbf{R}^2$.

(b) Si g désigne l'application définie au a), montrer que pour tout $(x,y)\in\mathbf{R}^2$,

$$\partial^2_{1,1}(f)(x,y)-\partial^2_{2,2}(f)(x,y)=4\partial^2_{1,2}(g)(x+y,x-y).$$

(c) En déduire que si $\partial^2_{1,1}(f)-\partial^2_{2,2}(f)$ est l'application nulle et que pour tout $x\in\mathbf{R}$, $f(x,0)=\partial_2(f)(x,0)=0$, alors f est l'application nulle.

3. (a) Montrer qu'il existe une unique application f de $\mathbf{R}^2$ dans $\mathbf{R}$ qui est de classe $\mathscr{C}^2$ sur $\mathbf{R}^2$ telle que $\partial^2_{1,1}(f)-\partial^2_{2,2}(f)$ soit l'application nulle et que pour tout $x\in\mathbf{R}$, $f(x,0)=x^2$ et $\partial_2(f)(x,0)=x$, et déterminer cette application.

(b) Étudier les extremums de f.

SOLUTION DE L'EXERCICE 2.33

1. Notons B une primitive de b, qui sera nécessairement de classe $\mathscr{C}^\infty$. Alors

$$f(x,y)=\frac{1}{2}\left[a(x+y)+a(x-y)+B(x+y)-B(x-y)\right].$$

Les fonctions $(x,y)\mapsto x+y$ et $(x,y)\mapsto x-y$ étant polynomiales sur $\mathbf{R}^2$, elles y sont de classe $\mathscr{C}^2$, et donc par somme de composées de fonctions $\mathscr{C}^2$, f est également $\mathscr{C}^2$ sur $\mathbf{R}^2$.
On a alors

$$\partial_1(f)(x,y)=\frac{1}{2}\left[a'(x+y)+a'(x-y)+b(x+y)-b(x-y)\right]$$
$$\partial_2(f)(x,y)=\frac{1}{2}\left[a'(x+y)-a'(x-y)+b(x+y)+b(x-y)\right].$$

Et donc

$$\partial^2_{1,1}(f)(x,y)=\frac{1}{2}\left[a''(x+y)+a''(x-y)+b'(x+y)-b'(x-y)\right],$$
$$\partial^2_{2,2}(f)(x,y)=\frac{1}{2}\left[a''(x+y)+a''(x-y)+b'(x+y)-b'(x-y)\right].$$

Et puisque $\partial^2_{1,1}(f)(x,y)=\partial^2_{2,2}(f)(x,y)$, alors leur différence est nulle.
Notons qu'on a $f(x,0)=a(x)$ et $\partial_2(f)(x,0)=b(x)$.

2.a. L'application $h:\mathbf{R}^2\to\mathbf{R}^2,(x,y)\mapsto(x+y,x-y)$ est un endomorphisme de $\mathbf{R}^2$.
Sa matrice dans la base canonique est $A=\begin{pmatrix}1&1\\1&-1\end{pmatrix}$, dont le déterminant est non nul, donc h est un isomorphisme[1] de $\mathbf{R}^2$.
Et alors, on a $f(x,y)=g(x+y,x-y)=(g\circ h)(x,y)$ pour tout $(x,y)\in\mathbf{R}^2$ si et seulement si

$$f=g\circ h\Leftrightarrow f\circ h^{-1}=g.$$

Ainsi, $g=f\circ h^{-1}$ est l'unique application de $\mathbf{R}^2$ dans $\mathbf{R}$ vérifiant la condition demandée.

[1] Et en particulier, h est une bijection.

Remarque

Notons au passage que $A^{-1}=\frac{1}{2}\begin{pmatrix}1&1\\1&-1\end{pmatrix}$, de sorte que $h^{-1}(x,y)=\left(\frac{x+y}{2},\frac{x-y}{2}\right)$.
Et donc on a $\forall(x,y)\in\mathbf{R}^2$,

$$g(x,y)=f\left(\frac{x+y}{2},\frac{x-y}{2}\right).$$

2.b. Soit $(x,y)\in\mathbf{R}^2$. Par définition, $\partial_1 f(x,y)$ est la dérivée en $t=x$ de $t\mapsto f(t,y)$.
Or, $f(t,y)=g((y,-y)+t(1,1))$.
Un résultat du cours nous assure alors que cette dérivée vaut

$$\partial_1(f)(x,y)=\langle\nabla g((y,-y)+x(1,1)),(1,1)\rangle=\partial_1(g)(x+y,x-y)+\partial_2(g)(x+y,x-y).$$

Le même raisonnement, appliqué aux fonctions $\partial_1(g)(x+y,x-y)$ et $\partial_2(g)(x+y,x-y)$ peut s'appliquer pour re-dériver par rapport à x :

$$\partial^2_{1,1}(f)(x,y)=\partial^2_{1,1}(g)(x+y,x-y)+\partial^2_{2,1}(g)(x+y,x-y)+\partial^2_{1,2}(g)(x+y,x-y)+\partial^2_{2,2}(g)(x+y,x-y).$$

De même, puisque $f(x,y)=g((x,x)+y(1,-1))$, on a

$$\partial_2(f)(x,y)=\langle\nabla g((x,x)+y(1,-1)),(1,1)\rangle=\partial_1(g)(x+y,x-y)-\partial_2(g)(x+y,x-y).$$

Et alors

$$\partial^2_{2,2}(f)(x,y)=\partial^2_{1,1}(g)(x+y,x-y)-\partial^2_{1,2}(g)(x+y,x-y)-\partial^2_{2,1}(g)(x+y,x-y)+\partial^2_{2,2}(g)(x+y,x-y).$$

Puisque g est $\mathscr{C}^2$, par le théorème de Schwarz, $\partial^2_{1,2}(g)=\partial^2_{2,1}(g)$ et donc

$$\partial^2_{1,1}(f)(x,y)-\partial^2_{2,2}(f)(x,y)=4\partial^2_{1,2}(g)(x+y,x-y).$$

2.c. Si $\partial^2_{1,1}(f) - \partial^2_{2,2}(f)$ est l'application nulle, alors, par la question précédente,

$$\forall (x,y) \in \mathbf{R}^2,\ \partial^2_{1,2}(g)(x+y, x-y) = 0.$$

Et puisque $(x,y) \mapsto (x+y, x-y)$ est une bijection de $\mathbf{R}^2$ sur lui-même, cela signifie que $\partial^2_{1,2}(g)$ est nulle.
Ainsi, pour tout $y \in \mathbf{R}$, $x \mapsto \partial_2(g)(x,y)$ est une application constante[2].
Notons $u(y)$ sa valeur, de sorte que pour tout $(x,y) \in \mathbf{R}^2$, $\partial_2(g)(x,y) = u(y)$.
La fonction u est alors de classe $\mathscr{C}^1$ sur $\mathbf{R}$, car $\partial_2(g)$ est de classe $\mathscr{C}^1$, et que

$$u(y) = \partial_2(g)(0,y) = \partial_2 g((0,0) + y(0,1)).$$

Si U est une primitive[3] de u, alors pour tout $x \in \mathbf{R}$, les fonctions $y \mapsto g(x,y)$ et U ont mêmes dérivées, donc elles diffèrent d'une constante[4] : il existe $v(x) \in \mathbf{R}$ tel que

$$\forall (x,y) \in \mathbf{R}^2,\ g(x,y) = U(y) + v(x).$$

La fonction v est alors de classe $\mathscr{C}^2$ car

$$v(t) = g(t,0) - U(0) = g((0,0) + t(1,0)) - U(0).$$

On a alors $f(x,0) = g(x,x) = U(x) + v(x)$ et $\partial_2 f(x,0) = \partial_1(g)(x,x) - \partial_2 g(x,x)$.
Mais $\partial_1(g)(x,y) = v'(x)$ et $\partial_2(g)(x,y) = U'(y) = u(y)$ et donc

$$\partial_2(f)(x,0) = v'(x) - u(x).$$

Ainsi, si $f(x,0) = 0$, il vient, pour tout $x \in \mathbf{R}$, $U(x) + v(x) = 0$ et donc $u(x) + v'(x) = 0$.
De même, si $\partial_2 f(x,0) = 0$, alors pour tout $x \in \mathbf{R}$, $u(x) - v'(x) = 0$.
Et donc $\begin{cases} u(x) + v'(x) = 0 \\ u(x) - v'(x) = 0 \end{cases} \Leftrightarrow u(x) = v'(x) = 0.$
On en déduit que U et v sont constantes : il existe $\lambda, \mu \in \mathbf{R}$ telles que $\forall x \in \mathbf{R}, U(x) = \lambda$ et $v(x) = \mu$.
Et puisque $f(x,0) = U(x) + v(x) = 0$, on en déduit que $\lambda + \mu = 0$ et donc

$$\forall (x,y) \in \mathbf{R},\ f(x,y) = U(y) + v(x) = \lambda + \mu = 0.$$

3.a. D'après la première question, en posant $a(t) = t^2$ et $b(t) = t$, qui sont bien de classe $\mathscr{C}^\infty$, alors la fonction

$$f : (x,y) \mapsto \frac{1}{2}\left[a(x+y) + a(x-y) + \int_{x-y}^{x+y} b(t)\,dt\right]$$

vérifie bien les conditions requises.
Notons qu'on a alors

$$\begin{aligned} f(x,y) &= \frac{1}{2}\left[(x+y)^2 + (x-y)^2 + \int_{x-y}^{x+y} t\,dt\right] \\ &= \frac{1}{2}\left[2x^2 + 2y^2 + \frac{1}{2}\left[(x+y)^2 - (x-y)^2\right]\right] \\ &= x^2 + xy + y^2. \end{aligned}$$

Il existe donc bien une fonction f vérifiant les conditions demandées.

Passons à l'unicité, et supposons qu'il existe une autre application f_1 vérifiant ces mêmes conditions, et soit alors $h = f - f_1$.
La fonction h est alors $\mathscr{C}^2$ sur $\mathbf{R}^2$, et vérifie

$$\begin{aligned} \partial^2_{1,1}(h) - \partial^2_{2,2}(h) &= \partial^2_{1,1}(f) - \partial^2_{1,1}(f_1) - \left(\partial^2_{2,2}(f) - \partial^2_{2,2}(f_1)\right) \\ &= \left(\partial^2_{1,1}(f) - \partial^2_{2,2}(f)\right) - \left(\partial^2_{1,1}(f_1) - \partial^2_{2,2}(f_1)\right) = 0. \end{aligned}$$

De plus, pour $x \in \mathbf{R}$, on a

$h(x,0) = f(x,0) - f_1(x,0) = x^2 - x^2 = 0$ et $\partial_2(h)(x,0) = \partial_2(f)(x,0) - \partial_2(f_1)(x,0) = x - x = 0$.

D'après la question 2.c, h est donc la fonction nulle, de sorte que $f = f_1$.
Ceci prouve bien l'unicité de f, et donc il existe une et une seule fonction f vérifiant les conditions requises, et il s'agit de $f(x,y) = x^2 + xy + y^2$.

[2] C'est-à-dire qui ne dépend pas de x, mais peut tout à fait dépendre de y.
[3] Qui est alors de classe $\mathscr{C}^2$.
[4] Qui peut dépendre de x.

3.b. Notons que $f(x,y) = \frac{1}{2}\left[(x+y)^2 + x^2 + y^2\right]$, et donc $f(x,y) \geqslant 0$, avec égalité si et seulement si $x = y = 0$.
Et donc f possède un minimum[5] en $(0,0)$, et ce minimum vaut 0.
D'autre part, $\lim\limits_{x\to+\infty} f(x,0) = \lim\limits_{x\to+\infty} x^2 = +\infty$, et donc f ne possède pas de maximum.

[5] Global.

□

2.5.1 Fonctions convexes

Il existe également une notion de convexité pour les fonctions de plusieurs variables, semblable à celle que l'on connaît pour les fonctions d'une variable.

EXERCICE 2.34 (ESCP 2001) [ESCP01.1.20] **Facile**

Fonctions convexes sur $\mathbf{R}^n$

Soit $n \in \mathbf{N}$ tel que $n \geqslant 2$. 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 $\mathscr{C}^1$ 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 $\lambda \in [0,1]$:

$$f((1-\lambda)x + \lambda y) \leqslant (1-\lambda)f(x) + \lambda f(y).$$

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 $\varphi'_{h,x}(0) \leqslant \varphi_{h,x}(1) - \varphi_{h,x}(0)$.
2. Montrer que pour tout $(x,y) \in (\mathbf{R}^n)^2$, on a

$$\langle \nabla f(x), y-x\rangle \leqslant f(y) - f(x).$$

   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 $\lambda \in]0,1[$:

$$f((1-\lambda)x + \lambda y) < (1-\lambda)f(x) + \lambda f(y).$$

   (a) Montrer que pour tout $x \in \mathbf{R}^n$, $f(0) \leqslant f(x)$, puis que si $x \neq 0$, alors $f(x) \neq 0$.
   (b) Montrer que $\inf\limits_{x\in\mathbf{R}^n, \|x\|=1} f(x)$ existe. On note α cette valeur. Montrer que $\alpha > 0$.
   (c) Montrer que pour tout $\|x\| > 1$, $\|f(x)\| \geqslant \alpha\|x\|$. En déduire la valeur de $\lim\limits_{\|x\|\to+\infty} f(x)$.

SOLUTION DE L'EXERCICE 2.34

1.a. Il s'agit de revenir à la définition : soient $t_1, t_2 \in \mathbf{R}$ et $\lambda \in [0,1]$. Alors

$$\begin{aligned}\varphi_{h,x}(\lambda t_1 + (1-\lambda)t_2) &= f(x + (\lambda t_1 + (1-\lambda)t_2)h)\\ &= f(\lambda(x+t_1h) + (1-\lambda)(x+t_2h))\\ &\leqslant \lambda f(x+t_1h) + (1-\lambda x)f(x+t_2h) = \lambda\varphi_{x,h}(t_1) + (1-\lambda)\varphi_{x,h}(t_2).\end{aligned}$$

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 $\varphi_{h,x}(1) \geqslant \varphi'_{h,x}(0) + \varphi_{h,x}(0)$ soit encore

$$\varphi'_{h,x}(0) \leqslant \varphi_{h,x}(1) - \varphi_{h,x}(0).$$

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

$$\langle \nabla f(x), y-x\rangle \leqslant \varphi_{h,x}(0) - \varphi_{h,x}(1) = f(x) - f(x+(y-x)) = f(x) - f(y).$$

En particulier, si x est un point critique de f, $\nabla f(x) = 0$ et donc pour tout $y \in \mathbf{R}^n$, $0 \leqslant f(y) - f(x) \Leftrightarrow f(x) \leqslant f(y)$.
f possède alors un minimum en x.

3.a. Pour $x = 0$, la relation précédente donne $\forall y \in \mathbf{R}^n$, $0 \leqslant f(y)$. S'il existait un $x \neq 0$ tel que $f(x) = 0$, alors il viendrait

$$f\left(\frac{1}{2} \times 0 + \frac{1}{2}x\right) < \frac{1}{2}f(0) + \frac{1}{2}f(x) \leqslant 0.$$

Ceci est contradictoire avec $f(y) \geqslant 0, \forall y \in \mathbf{R}^n$.

□

Plus généralement

Pour une fonction strictement convexe, le minimum, s'il existe, ne peut être atteint qu'en un unique point.

Pour les fonctions d'une variable deux fois dérivables, il est classique que les fonctions convexes sont celles dont la dérivée seconde est positive.
Ceci possède une généralisation simple pour les fonctions de plusieurs variables : les fonctions convexes sont celles qui, en tout point, possèdent une hessienne dont toutes les valeurs propres sont positives.
Il n'est pas trop dur (à l'aide de la formule de Taylor) de se convaincre que cela revient à demander à ce que f soit convexe dans toutes les directions, c'est-à-dire que l'intersection du graphe de f par n'importe quel plan vertical soit le graphe d'une fonction convexe (voir les dessins à la fin de l'exercice ci-dessous).

Exercice 2.35 (HEC 2018) [HEC18.268] — Moyen

Risque quadratique minimal d'une combinaison linéaire de variables i.i.d.

Soient $X_1, \ldots, X_n$ 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 $(x_1, \ldots, x_n) \in \mathbf{R}^n$, on note : $f(x_1, \ldots, x_n) = E\left(\left(\sum_{i=1}^n x_i X_i - \mu\right)^2\right)$.

1. (a) Justifier l'égalité : $f(x_1, \ldots, x_n) = \sigma^2 \sum_{i=1}^n x_i^2 + \mu^2 \left(\sum_{i=1}^n x_i - 1\right)^2$.

   (b) Justifier, pour tout $(x_1, \ldots, x_n) \in \mathbf{R}^n$, l'inégalité : $\left(\sum_{i=1}^n x_i\right)^2 \leqslant n \sum_{i=1}^n x_i^2$.

   (c) En déduire le minimum de f sur l'ensemble $\mathcal{H} = \{(x_1, \ldots, x_n) \in \mathbf{R}^n : x_1 + \cdots + x_n = 1\}$.

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 $h = (h_1, \ldots, h_n) \in \mathbf{R}^n$: $f(a + h) = f(a) + 2\int_0^1 (1-t)\left(\mu^2 \sum_{i=1}^n \sum_{j=1}^n h_i h_j + \sigma^2 \sum_{i=1}^n h_i^2\right) dt$.

   (c) En déduire que f admet un minimum global en a.

Solution de l'exercice 2.35

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,

$$f(x_1, \ldots, x_n) = E\left(\left(\sum_{i=1}^n x_i X_i\right)^2\right) - 2\mu \sum_{i=1}^n x_i E(X_i) + \mu^2 = E\left(\left(\sum_{i=1}^n x_i X_i\right)^2\right) - 2\mu^2 \sum_{i=1}^n x_i + \mu^2.$$

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 $y = (1, \dots, 1)$:

$$\langle x, y\rangle^2 \leqslant \|x\|^2 \cdot \|y\|^2 \Leftrightarrow \left(\sum_{i=1}^{n} x_i\right)^2 \leqslant \left(\sum_{i=1}^{n} x_i^2\right) \underbrace{\left(\sum_{i=1}^{n} 1^2\right)}_{=n}.$$

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, $f(x_1, \dots, x_n) \geqslant \dfrac{\sigma^2}{n} \left(\sum_{i=1}^{n} x_i\right)^2 = \dfrac{\sigma^2}{n}$.

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 $\left(\dfrac{1}{n}, \dots, \dfrac{1}{n}\right) \in \mathcal{H}$, $\dfrac{\sigma^2}{n}$ 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 en 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 $\left(\dfrac{1}{n}, \dots, \dfrac{1}{n}\right)$ et son opposé.

> ⚠ 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 $\mathcal{H}$.

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 $j \in [\![1, n]\!]$,

$$\partial_j f(x_1, \dots, x_n) = 2\sigma^2 x_j + 2\mu^2 \left(\sum_{i=1}^{n} x_i - 1\right).$$

Et donc en particulier, $(x_1, \dots, x_n) \in \mathbf{R}^n$ est un point critique de f si et seulement si

$$\forall j \in [\![1, n]\!],\ \sigma^2 x_j = -\mu^2 \left(\sum_{i=1}^{n} x_i - 1\right).$$

En particulier, on a doit avoir $x_1 = x_2 = \cdots = x_n = -\dfrac{\mu}{\sigma^2} \left(\sum_{i=1}^{n} x_i - 1\right)$.

Et donc $\sum_{i=1}^{n} x_i = n x_1$.

La première équation devient alors

$$\sigma^2 x_1 + \mu^2 (n x_1 - 1) = 0 \Leftrightarrow x_1 = \frac{\mu^2}{\sigma^2 + n\mu^2}.$$

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 fonction[1] $g : t \mapsto f(a+th)$.

[1] D'une variable.

En effet, nous savons que g est de classe $\mathscr{C}^2$, 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 $i \in [\![1, n]\!]$, on a

$$\partial^2_{i,i} f(x_1, \dots, x_n) = 2(\sigma^2+\mu^2).$$

Et pour $i \neq j$, $\partial^2_{i,j} f(x_1, \dots, x_n) = 2\mu^2$.
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 $t \in [0,1]$,

$$q_{a+th}(h) = \begin{pmatrix} h_1 & \dots & h_n \end{pmatrix} A \begin{pmatrix} h_1 \\ \vdots \\ h_n \end{pmatrix} = 2\begin{pmatrix} h_1 & \dots & h_n \end{pmatrix} \begin{pmatrix} (\sigma^2+\mu^2)h_1 + \mu^2 \sum_{j=2}^n h_j \\ \vdots \\ \mu^2 \sum_{j=1}^{n-1} h_j + (\mu^2+\sigma^2)h_n \end{pmatrix}$$

$$= 2\left(\mu^2 \sum_{i=1}^n \sum_{j=1}^n h_i h_j + \sigma^2 \sum_{i=1}^n h_i^2\right).$$

2.c. Quel que soit $h \in \mathbf{R}^n$, on a

$$\sum_{i=1}^n \sum_{j=1}^n h_i h_j = \sum_{i=1}^n h_i \left(\sum_{j=1}^n h_j\right) = \left(\sum_{i=1}^n h_i\right)\left(\sum_{j=1}^n h_j\right) = \left(\sum_{i=1}^n h_i\right)^2 \geqslant 0.$$

Et donc $\mu^2 \sum_{i=1}^n \sum_{j=1}^n h_i h_j + \sigma^2 \sum_{i=1}^n h_i^2 \geqslant 0$.

On en déduit, par positivité de l'intégrale que $\int_0^1 (1-t)\left(\mu^2 \sum_{i=1}^n \sum_{j=1}^n h_i h_j + \sigma^2 \sum_{i=1}^n h_i^2\right) dt \geqslant 0$.

Et donc que $f(a+h) \geqslant f(a)$, 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.*

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 $\sum_{i=1}^n x_i = 1$.

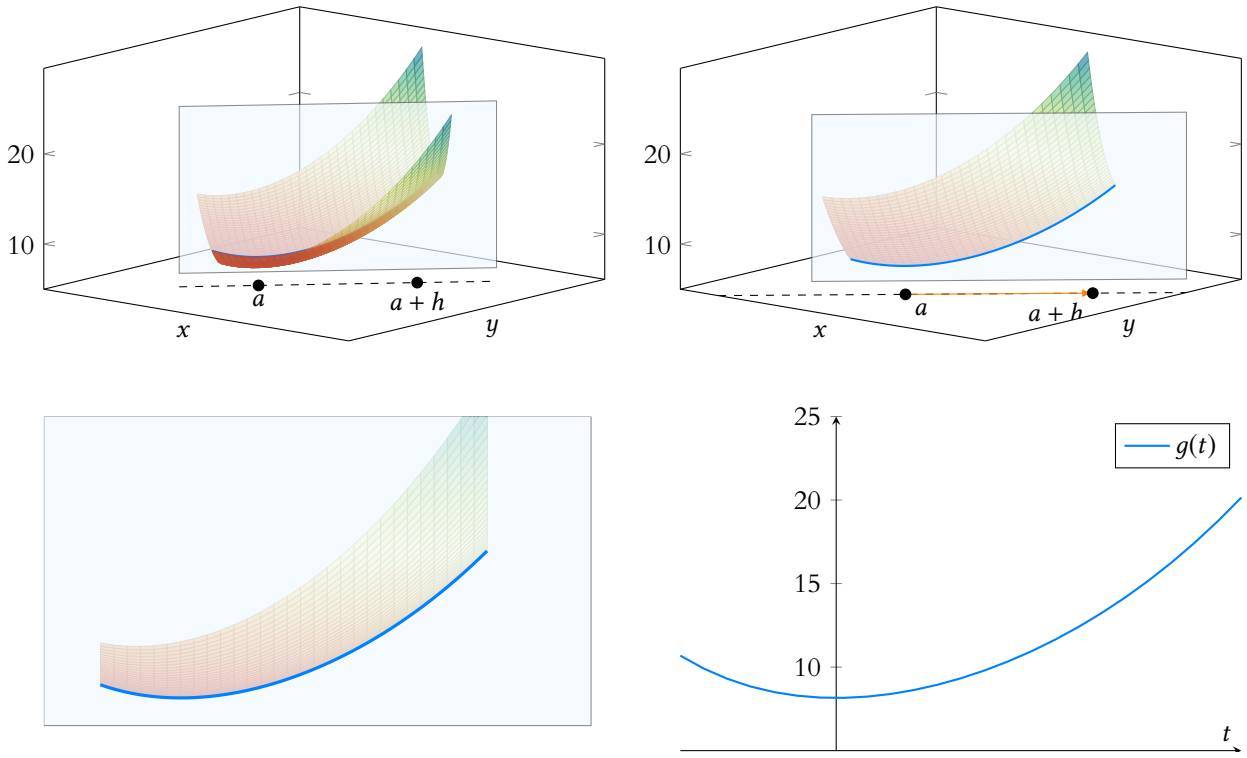


FIGURE 2.10 – Le fait que les valeurs propres de la hessienne en tout point soient positives se traduit par le fait que quel que soit le plan vertical par lequel on coupe le graphe de f, la fonction obtenue (ici g) est convexe.
On dit alors que f est une fonction convexe.

EXERCICE 2.36 (ESCP 2012) [ESCP12.1.13] — Difficile

Méthode du gradient pour une fonction quadratique convexe.

Soit n un entier, tel que $n \geqslant 2$. On considère $\mathbf{R}^n$ muni de son produit scalaire canonique noté $\langle ., . \rangle$, de la norme associée notée $\| \cdot \|$, et $A \in \mathcal{M}_n(\mathbf{R})$, symétrique réelle dont les valeurs propres sont toutes strictement positives. On confond vecteur de $\mathbf{R}^n$ et matrice colonne canoniquement associée et on pose, pour tout $X \in \mathbf{R}^n$, $\Phi(X) = {}^tXAX$.

1. Soit B un élément de $\mathbf{R}^n$. Montrer que l'équation $AX = B$ d'inconnue $X \in \mathbf{R}^n$ admet une unique solution qu'on notera R.
2. Montrer qu'il existe deux réels α et β strictement positifs tels que pour tout X de $\mathbf{R}^n$
$$\alpha\|X\|^2 \leqslant \Phi(X) \leqslant \beta\|X\|^2.$$
Dans la suite de l'exercice, on pose pour $X \in \mathbf{R}^n$: $F(X) = \Phi(X) - 2\,{}^tBX$.
3. (a) Déterminer le gradient ∇F_X de F en X.
   (b) Soient X et H deux éléments de $\mathbf{R}^n$. Montrer que
$$F(X + H) = F(X) + \langle \nabla F_X, H \rangle + \Phi(H).$$
   (c) En déduire que F possède un minimum sur $\mathbf{R}^n$. En quel point est-il atteint ?
4. Soit $X \in \mathbf{R}^n$ fixé, $X \neq 0$. Déterminer $\alpha \in \mathbf{R}$ de façon à ce que $F(X - \alpha\nabla F_X)$ soit minimal. Calculer ce minimum.
5. Soit $X_0 \in \mathbf{R}^n$. On définit une suite $(X_k)_{k\in\mathbf{N}}$ de vecteurs de $\mathbf{R}^n$ par, pour tout $k \in \mathbf{N}$: $X_{k+1} = X_k - \alpha_k \nabla F_{X_k}$, où $\alpha_k = \dfrac{\|\nabla F_{X_k}\|^2}{2\Phi(\nabla F_{X_k})}$ si $X_k \neq R$ et 0 sinon.
   (a) Montrer que la suite $(F(X_k))_{k\in\mathbf{N}}$ converge.
   (b) Exprimer $F(X_{k+1}) - F(X_k)$ en fonction de α_k et de ∇F_{X_k}.
6. Une suite $(Y_k)_{k\in\mathbf{N}}$ de vecteurs de $\mathbf{R}^n$ sera dite convergente vers un vecteur $Z \in \mathbf{R}^n$ si $\lim\limits_{k\to+\infty} \|Y_k - Z\| = 0$, ce qui revient à dire que les coordonnées de Y_k convergent vers les coordonnées correspondantes de Z.

(a) Montrer que la suite $(\nabla F_{X_k})_{k\in\mathbf{N}}$ converge vers 0.
(b) En déduire la limite de la suite $(X_k)_{k\in\mathbf{N}}$.

Solution de l'exercice 2.36

1. Les valeurs propres de A sont toutes non nulles, donc A est inversible. Et par conséquent, il existe un unique $X \in \mathbf{R}^n$, $X = A^{-1}B$ tel que $AX = B$.

2. Notons $\lambda_1, \dots, \lambda_n$ les valeurs propres de A, comptées autant de fois que la dimension du sous-espace propre associé.
Soit $\mathcal{B} = (e_1, \dots, e_n)$ une base orthonormée de vecteurs propres de A.

Alors tout vecteur X de $\mathbf{R}^n$ s'écrit de manière unique $X = \sum_{i=1}^{n} x_i e_i$.

On a alors $\Phi(X) = \sum_{i=1}^{n} \lambda_i x_i^2$.

Soit alors α (resp. β) la plus petite (resp. la plus grande) valeur propre de A, de sorte que pour tout $i \in [\![1, n]\!]$, $\alpha \leqslant \lambda_i \leqslant \beta$. Alors

$$\forall X \in \mathbf{R}^n, \alpha \sum_{i=1}^{n} x_i^2 \leqslant \Phi(X) \leqslant \beta \sum_{i=1}^{n} x_i^2 \Leftrightarrow \alpha\|X\|^2 \leqslant \Phi(X) \leqslant \beta\|X\|^2.$$

3.a. La fonction F est de classe $\mathcal{C}^1$ car polynomiale. En effet, on a

$$F(X) = \sum_{i=1}^{n}\sum_{j=1}^{n} a_{i,j}x_i x_j - 2\sum_{i=1}^{n} b_i x_i = \sum_{i=1}^{n} a_{i,i}x_i^2 + \sum_{i=1}^{n}\sum_{j\neq i} a_{i,j}x_i - \sum_{i=1}^{n} b_i x_i.$$

On a alors

$$\forall k \in [\![1, n]\!],\ \partial_k F(X) = 2a_{k,k}x_k + 2\sum_{j\neq k} a_{k,j}x_j - 2b_k = 2\left(\sum_{i=1}^{n} a_{k,i}x_i - b_k\right).$$

Autrement dit, en notant ∇F_X le gradient de F en X, on a[1]

$$\nabla F_X = 2(AX - B).$$

[1] On confond toujours vecteurs lignes et vecteurs colonnes, de sorte que ∇F_X est ici vu comme un vecteur colonne.

3.b. On a

$$\begin{aligned} F(X+H) - F(X) &= {}^t(X+H)A(X+H) - 2{}^tB(X+H) - {}^tXAX + 2{}^tBX \\ &= {}^tHAX + {}^tXAH + {}^tHAH - 2{}^tBH = 2{}^tHAX + \Phi(H) - 2{}^tBH. \end{aligned}$$

Mais on a

$$\langle \nabla F_X, H\rangle = {}^tH\nabla F_X = {}^tH2(AX - B) = -2{}^tHB + 2{}^tHAX.$$

Et donc $F(X + H) = F(X) + \langle \nabla F_X, H\rangle + \Phi(H)$.

3.c. Le gradient de F s'annule en X si et seulement si $AX = B$, c'est-à-dire uniquement en $X = R$.
De plus, le résultat obtenu à la question précédente prouve que si X est un point critique, alors

$$F(X + H) - F(X) = \Phi(H) \geqslant \alpha\|X\|^2 \geqslant 0.$$

Et donc F admet un minimum global en tout point critique.
Et puisque R est le seul point critique, F atteint un minimum global en R.

4. D'après la question 3.b,

$$F(X - \alpha\nabla F_X) = F(X) - \alpha\langle \nabla F_X, \nabla F_X\rangle + \alpha^2\Phi(\nabla F_X).$$

Si $X \neq R$, alors $\nabla F_X \neq 0$ et cette fonction est un polynôme du second degré en α, de coefficient dominant $\Phi(\nabla F_X) > 0$.

Alors ce trinôme admet un unique minimum atteint[2] en $\alpha = \dfrac{\|\nabla F_X\|^2}{2\Phi(\nabla F_X)}$.

Ce minimum vaut alors $F(X) = \dfrac{\|\nabla F_X\|^4}{4\Phi(\nabla F_X)^2}$.

Enfin, si $X = R$, alors $\alpha \mapsto F(X - \alpha\nabla F_X)$ est constante égale à $F(X)$.

Détails
Si $X = R$, alors $\nabla F_X = 0$, et donc le coefficient en α^2 est nul, nous avons donc un polynôme en α de degré 1. En revanche, dès que $X \neq R$, $\nabla F_X \neq 0$, et donc d'après la question 2, $\Phi(\nabla F_X) > 0$.

[2] C'est le résultat classique vu au lycée : une fonction polynomiale de degré 2 de la forme $ax^2 + bx + c$ admet un unique extremum (dont la nature dépend du signe de a), atteint en $-\frac{b}{2a}$.

5.a. Montrons que la suite $(F(X_k))_{k\in\mathbf{N}}$ est décroissante.
En effet, d'après la question précédente, $F(X_{k+1})$ est le minimum de $g_k : \alpha \mapsto F\left(X_k - \alpha\nabla F_{X_k}\right)$.
Ce minimum est alors inférieur ou égal à $g_k(0) = F(X_k)$. Donc $F(X_{k+1}) \leqslant F(X_k)$.
Comme de plus la suite $(F(X_k))_{k\in\mathbf{N}}$ est minorée par $F(R)$, elle est convergente.

5.b. Il vient donc

$$\begin{aligned} F(X_{k+1}) - F(X_k) &= F(X_k - \alpha_k\nabla F_{X_k}) - F(X_k) \\ &= \langle -\alpha_k\nabla F_{X_k}, \nabla F_{X_k}\rangle + \Phi(\alpha_k\nabla F_{X_k}) \\ &= -\alpha_k\langle \nabla F_{X_k}, \nabla F_{X_k}\rangle + \alpha_k^2\Phi(\nabla F_{X_k}) \\ &= -\frac{\|\nabla F_{X_k}\|^2}{2\Phi(\nabla F_{X_k})}\|\nabla F_{X_k}\|^2 + \frac{\|\nabla F_{X_k}\|^4}{4\Phi(\nabla F_{X_k})} \\ &= -\frac{\|\nabla F_{X_k}\|^4}{2\Phi(\nabla F_{X_k})}. \end{aligned}$$

6.a. La suite $(F(X_k))_{k\in\mathbf{N}}$ converge, donc $\lim\limits_{k\to+\infty} F(X_{k+1}) - F(X_k) = 0$, et donc

$$\lim_{k\to+\infty} -\frac{\|\nabla F_{X_k}\|^4}{2\Phi(\nabla F_{X_k})} = 0.$$

Or, en utilisant l'inégalité de la question 2, il vient

$$\frac{1}{\beta}\|\nabla F_{X_k}\|^2 \leqslant \frac{1}{\Phi(\nabla F_{X_k})} \leqslant \frac{1}{\alpha\|\nabla F_{X_k}\|^2}$$

et donc

$$\frac{\|\nabla F_{X_k}\|^4}{2\Phi(\nabla F_{X_k})} \geqslant \frac{1}{2\beta}\|\nabla F_{X_k}\|^2.$$

Par conséquent $\lim\limits_{k\to+\infty} \|\nabla F_{X_k}\| = 0$. Donc la suite (∇F_{X_k}) converge vers 0.

6.b. On en déduit que

$$\lim_{k\to+\infty} 2(AX - B) = 0 \Leftrightarrow \lim_{k\to+\infty} AX_k = B$$

et donc par multiplication par A^{-1},

$$\lim_{k\to+\infty} X_k = A^{-1}B = R.$$

> **Remarque**
> On admet ici un résultat relativement classique et assez intuitif, qui est que si $X_k \underset{k\to+\infty}{\longrightarrow} X$, alors on a également $A^{-1}X_k \underset{k\to+\infty}{\longrightarrow} AX$.

Quelques explications *: nous venons donc de prouver que nous avons un algorithme qui permet de déterminer le minimum de F : il suffit de partir de n'importe quel vecteur $X_0 \in \mathbf{R}^n$, et de construire la suite (X_k) comme décrit précédemment : elle converge alors automatiquement vers l'unique minimum de F.*
Notons que dans ce cas précis, nous avions déjà une expression de R, et donc l'utilisation de cet algorithme n'était pas nécessaire.
En revanche, cette méthode, nommée méthode du gradient, se généralise à d'autres fonctions convexes.
Rappelons-nous que le gradient indique la direction de plus forte pente.
Et donc en nous déplaçant dans la direction de $-\nabla F_{X_k}$, nous nous dirigeons vers le minimum de F.
Il faut toutefois veiller à ne pas prendre α_k trop grand, car alors on risque de «dépasser» le point critique. □

2.5.2 Optimisation sous contraintes

EXERCICE 2.37 (ESCP 2015) [ESCP15.108] — **Moyen**

Extremas d'une fonction sur la sphère unité.

Soit $n \in \mathbf{N}^*$, et $a_1, \ldots, a_n$ des réels non tous nuls. On considère les fonctions f, g et h définies sur $\mathbf{R}^n$ par

$$f(x_1,\ldots,x_n) = \prod_{k=1}^{n} x_k^2,\ g(x_1,\ldots,x_n) = \sum_{k=1}^{n} x_k^2 \text{ et } h(x_1,\ldots,x_n) = \sum_{k=1}^{n} a_k x_k.$$

1. (a) Justifier que f est de classe $\mathscr{C}^1$ sur $\mathbf{R}^n$. Déterminer son gradient en tout point.
   (b) Prouver que f admet un maximum sur la sphère unité S de $\mathbf{R}^n$ définie par $S = \left\{(x_1,\dots,x_n) \in \mathbf{R}^2 : \sum_{k=1}^{n} x_k^2 = 1\right\}$.
   (c) Déterminer le maximum de f sur S.
   (d) En déduire que pour tout point $u = (u_1,\dots,u_n)$ de $\mathbf{R}^n$, on a $\left|\prod_{k=1}^{n} u_k\right| \leqslant \left(\frac{\|u\|}{\sqrt{n}}\right)^n$ où $\|\cdot\|$ désigne la norme usuelle de $\mathbf{R}^n$.
2. (a) Soit $i \in [\![1,n]\!]$ tel que $a_i \neq 0$. On pose
$$H = \{(x_1,\dots,x_n) \in \mathbf{R}^n : h(x_1,\dots,x_n) = 1\} \text{ et } B = \left\{x \in \mathbf{R}^n : \|x\| \leqslant \frac{1}{|a_i|}\right\}$$
   Prouver que l'ensemble $B \cap H$ est non vide, puis qu'il est fermé borné.
   (b) 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 $h(x_1,\dots,x_n) = 1$.
   (c) Déterminer le minimum de g sur H.

Solution de l'exercice 2.37

1.a. f est de classe $\mathscr{C}^1$ sur $\mathbf{R}^n$ car elle est polynomiale. On a alors
$$\nabla f(x_1,\dots,x_n) = \left(2x_1x_2^2\cdots x_n^2, x_1^2 2x_2 \cdots x_n^2, x_1^2\cdots x_{n-1}^2 2x_n\right).$$
Autrement dit, pour tout $i \in [\![1,n]\!]$, 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 continue[1] 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 $i \in [\![1,n]\!]$, on a
$$x_i^2 = 1 - \sum_{\substack{j=1\\j\neq i}}^{n} x_j^2 \leqslant 1 \Rightarrow |x_i| \leqslant 1.$$
Puisque f est continue sur le fermé borné S, elle y admet un maximum.

[1] Car polynomiale.

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
$$\begin{cases} x_1^2+\cdots+x_n^2 = 1 \\ \nabla f(x_1,\dots,x_n) = \lambda \nabla g(x_1,\dots,x_n) \end{cases} \Leftrightarrow \begin{cases} x_1^2+\cdots+x_n^2 &= 1 \\ 2x_1x_2^2\dots x_n^2 &= 2\lambda x_1 \\ &\vdots \\ 2x_1^2x_2^2 &\dots x_n = 2\lambda x_n \end{cases}$$
Si l'un des x_i est nul, alors $f(x_1,\dots,x_n) = 0$, qui n'est clairement pas le minimum de f sur S.
Donc nous pouvons nous contenter de déterminer les points critiques (sous la contrainte S) dont toutes les coordonnées sont non nulles.
De la seconde équation, puisque $x_1 \neq 0$, il vient $x_2^2\cdots x_n^2 = \lambda$, et de la troisième équation, il vient $x_1^2x_3^2\cdots x_n^2 = \lambda$.
Et donc $x_1^2 = x_2^2$. On prouve de même que $x_1^2 = x_2^2 = \cdots = x_n^2$.
La première équation devient alors $nx_1^2 = 1 \Leftrightarrow x_1 = \pm\sqrt{\frac{1}{n}}$.
Dans ce cas, on a $f(x_1,\dots,x_n) = \frac{1}{n^n}$.

Alors que si l'un des x_i est nul, on a $f(x_1,\dots,x_n)=0$.

Et donc le maximum de f est $\dfrac{1}{n^n}$.

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 $\dfrac{1}{n}$ (et alors $f(x_1,,\dots,x_n)=\dfrac{1}{n^n}$. 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 $\dfrac{1}{n^n}$.

1.d. Si $u\in\mathbf{R}^n$ est non nul, alors $\dfrac{u}{\|u\|}\in S$. Et donc

$$f(u_1,\dots,u_n)=\prod_{i=1}^n\|u\|^2\frac{u_i^2}{\|u\|^2}=\|u\|^{2n}f\left(\frac{u}{\|u\|}\right)\leqslant\|u\|^{2n}\frac{1}{n^n}=\left(\frac{\|u\|}{\sqrt{n}}\right)^{2n}.$$

En prenant la racine carrée, il vient

$$\sqrt{f(u_1,\dots,u_n)}=\prod_{i=1}^n\sqrt{u_i^2}=\prod_{i=1}^n|u_i|=\left|\prod_{i=1}^n u_i\right|\leqslant\left(\frac{\|u\|}{\sqrt{n}}\right)^n.$$

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

$$h(x_1,\dots,x_n)=a_i\frac{1}{a_i}=1 \text{ et } \|(x_1,\dots,x_n)\|=\sqrt{\sum_{i=1}^n x_i^2}=\sqrt{\frac{1}{a_i^2}}=\frac{1}{|a_i|}\leqslant\frac{1}{|a_i|}.$$

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[2] sur $\mathbf{R}^n$ et donc $B=\left\{x\in\mathbf{R}^n : \|x\|\leqslant\frac{1}{|a_i|}\right\}$ 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 et de rayon $\dfrac{1}{|a_i|}$.

[2] En effet, on a $\|x\|=\sqrt{\sum_{i=1}^n x_i^2}$.

2.b. g est continue sur le fermé borné $B\cap H$, donc y admet un minimum.

Nous avons prouvé que $\left(0,\dots,0,\dfrac{1}{a_i},0,\dots,0\right)$ est dans $B\cap H$.

Or $g\left(0,\dots,0,\dfrac{1}{a_i},0,\dots,0\right)=\dfrac{1}{a_i^2}$.

On en déduit que le minimum de g sur $B\cap H$ est inférieur ou égal à $\dfrac{1}{a_i^2}$.

Si $x\in\mathbf{R}^n$ est dans H mais pas dans B, alors $\|x\|>\dfrac{1}{a_i}$ et donc $g(x)=\|x\|^2>\dfrac{1}{a_i^2}$.

Ainsi le minimum de g sur $B\cap H$ est également le minimum de g sur H tout entier, c'est-à-dire sous la contrainte $h(x_1,\dots,x_n)=1$.

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 $\forall i\in[\![1,n]\!],\ x_i=\dfrac{a_i}{\sum_{j=1}^n a_j^2}$.

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

$$g\left(\frac{a_1}{\sum_{i=1}^n a_i^2},\dots,\frac{a_n}{\sum_{i=1}^n a_i^2}\right)=\frac{1}{\sum_{i=1}^n a_i^2}.$$

□

2.6 Les inclassables

Les croissances comparées usuelles nous informent que $e^n = \underset{n\to+\infty}{o}(n!)$ et $n! = \underset{n\to+\infty}{o}(n^n)$.
La formule de Stirling, souvent admise dans les sujets d'écrit comme d'oral, vient préciser ceci en donnant un équivalent (assez surprenant) de $n!$ au voisinage de $+\infty$.
Il en existe plusieurs preuves, dont une entièrement analytique et ne faisant appel qu'à des outils élémentaires se trouve dans le problème 2 de Lyon 2012.
D'autres preuves à base de probabilités existent. La suivante, qui utilise le théorème central limite, est celle que l'on trouve dans le sujet de Maths II 2017.

Exercice 2.38 (ESCP 2016) [ESCP.16.1.07] **Moyen**

Formule de Stirling

Soit λ un réel strictement positif. Pour tout $n \in \mathbf{N}$, on pose

$$P_n(\lambda) = e^{-\lambda}\sum_{k=0}^{n}\frac{\lambda^k}{k!} \text{ et } I_n = \int_0^1 (1-x)^n e^{nx}\,dx.$$

1. (a) Montrer que $1 - P_n(\lambda) = \dfrac{1}{n!}\displaystyle\int_0^\lambda x^n e^{-x}\,dx$.
   (b) Déterminer une relation entre $P_n(n)$ et I_n.
2. Pour tout $x \in]0,1[$, on pose $\psi(x) = -\dfrac{x+\ln(1-x)}{x^2}$.
   Montrer que ψ admet un prolongement de classe $\mathscr{C}^1$ sur $[0,1[$, prolongement que l'on note encore ψ.
   Montrer que ψ réalise une bijection de $[0,1[$ sur un intervalle à préciser.
3. Pour $\sigma \in]0,1[$, on pose $\delta = \psi{-}1\left(\dfrac{1}{2\sigma^2}\right)$.
   (a) Montrer que pour tout $x \in]0,1[$, $((1-x)e^x)^n = e^{-nx^2\psi(x)}$.
   (b) Montrer que : $\dfrac{\sigma}{\sqrt{n}}\displaystyle\int_0^{\delta\sqrt{n}} e^{-x^2/2}\,dx \leqslant I_n \leqslant \sqrt{\dfrac{\pi}{2n}}$.
   (c) En déduire un équivalent de I_n lorsque n tend vers l'infini.
   (d) Montrer que $1 - P_n(n)$ est équivalent à $\dfrac{n^n}{n!}e^{-n}\sqrt{\dfrac{\pi}{2n}}$ lorsque n tend vers l'infini.
4. En utilisant le théorème central limite, déterminer un équivalent de $n!$ lorsque n tend vers l'infini.

3 Probabilités

Aussi bien à HEC qu'à l'ESCP, l'un des deux exercices (l'exercice avec préparation et la question sans préparation) portera forcément sur les probabilités. Le préparationnaire y accordera donc un soin tout particulier.

3.1 Généralités

Exercice 3.1 (ESCP 2012) [ESCP12.3.15]

Loi du zéro-un de Borel

Soit $(A_n)_{n\in\mathbf{N}^*}$ une suite d'événements d'un espace probabilisé $(\Omega, \mathcal{A}, P)$. On note $p_n = P(A_n)$.

On note B l'événement $\bigcap_{n\geqslant 1}\left(\bigcup_{k\geqslant n} A_k\right)$.

1. Expliquer pourquoi on a
$$B = \{\omega \in \Omega \,|\, \omega \text{ appartient à une infinité des } A_n\}.$$
2. On suppose que la série $\sum_k P(A_k)$ converge. Montrer que $P(B) = 0$.
3. On suppose que les événements (A_n) sont indépendants et que la série $\sum_n P(A_n)$ est divergente.
   (a) Montrer que l'événement $\overline{B}$ est égal à $\bigcup_{n\geqslant 1}\left(\bigcap_{k\geqslant n} \overline{A_k}\right)$.
   (b) Exprimer $P\left(\bigcap_{k=n}^{m} \overline{A_k}\right)$ en fonction des p_k.
   (c) Montrer que la série $\sum_k \ln(1-p_k)$ est divergente.
   (d) En déduire que $P(B) = 1$.
4. Soit α un réel strictement positif et (X_n) une suite de variables aléatoires indépendantes telles que pour tout n, X_n suit la loi de Bernoulli de paramètre $\frac{1}{n^\alpha}$.
   (a) Montrer que $\lim_{n\to+\infty} E(X_n) = 0$.
   (b) On suppose que $0 < \alpha < 1$. Montrer qu'avec une probabilité égale à 1, l'ensemble $\{n \in \mathbf{N}^* / X_n = 1\}$ contient une infinité d'éléments.
   (c) On suppose que $\alpha > 1$. Montrer qu'avec une probabilité égale à 1, l'ensemble $\{n \in \mathbf{N}^* / X_n = 1\}$ est fini.

Solution de l'exercice 3.1

1.

2. Pour tout $n \in \mathbf{N}$, on a $B \subset \bigcup_{k\geqslant n} A_k$. Et donc

$$P(B) \leqslant P\left(\bigcup_{k\geqslant n} A_k\right) \leqslant \sum_{k=n}^{+\infty} P(A_k).$$

Mais puisque la série de terme général $P(A_k)$ converge, son reste d'ordre n tend vers 0 lorsque n tend vers $+\infty$, et donc $0 \leqslant P(B) \leqslant 0$, de sorte que $P(B) = 0$.

3.a. Il s'agit juste d'utiliser les propriétés élémentaires du contraire : le contraire d'une union est l'intersection des contraires, et le contraire d'une intersection est l'union des contraires. Donc

$$\overline{B} = \bigcup_{n\geqslant 1} \overline{\bigcup_{k\geqslant n} A_k} = \bigcup_{n\geqslant 1} \bigcap_{k\geqslant n} \overline{A_k}.$$

3.b. Les A_k étant indépendants, leurs contraires le sont également, et donc

$$P\left(\bigcap_{k=n}^{m} \overline{A_k}\right) = \prod_{k=n}^{m} P\left(\overline{A_k}\right) = \prod_{k=n}^{m} (1-p_k).$$

3.c. La fonction $x \mapsto \ln(1-x)$ est concave sur $[0,1[$, et donc située sous ses tangentes.
En particulier, sa tangente en $x = 0$ est la droite d'équation $y = -x$, de sorte que pour tout $t \in [0,1[$, $\ln(1-t) \leqslant -t$.
Soit encore $0 \leqslant t \leqslant -\ln(1-t)$.

En particulier, pour tout $k \in \mathbf{N}^*$, $0 \leqslant p_k \leqslant -\ln(1-p_k)$.
Et puisque la série de terme général p_k diverge, il en est de même de la série de terme général $-\ln(1-p_k)$ et donc de $\sum_k \ln(1-p_k)$.

3.d. Puisque la série de terme général $\ln(1-p_k)$ est à termes négatifs, sa divergence implique que $\lim_{m\to+\infty} \sum_{n=1}^{m} \ln(1-p_k) = -\infty$.

Et même mieux : pour tout $n \geqslant 1$, $\lim_{m\to+\infty} \sum_{k=n}^{m} \ln(1-p_k) = +\infty$.

Et donc après passage à l'exponentielle, $\lim_{m\to+\infty} \prod_{k=n}^{m} (1-p_k) = 0$.

D'après le théorème de la limite monotone, on a donc

$$P\left(\bigcap_{k\geqslant n} \overline{A_k}\right) = \lim_{m\to+\infty} P\left(\bigcap_{k=n}^{m} \overline{A_k}\right) = \lim_{m\to+\infty} \prod_{k=n}^{m} (1-p_k) = 0.$$

On en déduit que

$$0 \leqslant P\left(\overline{B}\right) \leqslant \sum_{n=1}^{+\infty} P\left(\bigcap_{k\geqslant n} \overline{A_k}\right) = 0.$$

Et donc $P\left(\overline{B}\right) = 0$, de sorte que $P(B) = 1$.

4.a. On a $E(X_n) = \dfrac{1}{n^\alpha} \underset{n\to+\infty}{\longrightarrow} 0$.

4.b. Il s'agit d'appliquer le résultat de la question 3, en notant $A_k = P(X_k = 1)$.
Les X_k étant indépendantes, les A_k le sont, et on a alors $P(A_k) = \dfrac{1}{k^\alpha}$.
Puisque $0 < \alpha < 1$, la série de terme général $P(A_k)$ diverge, et donc $P(B) = 1$.
Autrement dit, presque sûrement[1], une infinité des A_k sont réalisés simultanément. Soit encore : il existe une infinité d'entiers n tels que $X_n = 1$.

[1] Ce qui veut dire «avec probabilité 1»

4.c. Appliquons cette fois le résultat de la question 2 : puisque $\alpha > 1$, $\sum P(A_k)$ diverge, et donc $P(B) = 0$.
Et donc $P\left(\overline{B}\right) = 1$. Mais $\overline{B} = \{\omega \in \Omega \mid \omega \text{ n'appartient qu'à un nombre fini des } A_n\}$.
Et donc, avec probabilité 1, l'ensemble $\{n \in \mathbf{N}^* / X_n = 1\}$ est un ensemble fini.

□

Exercice 3.2 (QSP ESCP 2014) [ESCP14.QSP03]

Indépendance d'événements

On considère une suite infinie de lancers d'une pièce, la probabilité de faire Pile étant égale à $p \in]0,1[$, et celle de faire face égale à $1-p$. Soit X_n le résultat du $n^{\text{ème}}$ lancer.
Les événements $A_n = [X_n \neq X_{n-1}]$ sont-ils indépendants ?

Solution de l'exercice 3.2 L'énoncé ne le précise pas, mais on peut supposer que X_n vaut 1 lorsqu'on obtient Pile au $n^{\text{ème}}$ lancer et 0 sinon.
Par la formule des probabilités totales appliquée au système complet d'événements $[X_n = 0], [X_n = 1]$, on a

$$P(A_n) = P(A_n \cap [X_n = 0]) + P(A_n \cap [X_n = 1]) = P([X_n = 0] \cap [X_{n-1} = 1]) P([X_n = 1] \cap [X_{n-1} = 0]).$$

Les lancers successifs étant indépendants[1], les variables X_n le sont et donc

$$P(A_n) = P(X_n = 0)P(X_{n-1} = 1) + P(X_n = 1)P(X_{n-1} = 0) = 2p(1-p).$$

[1] L'énoncé ne dit rien à ce sujet, mais cela semble être une hypothèse plus que raisonnable !

D'autre part, on a, toujours par les probabilités totales,

$$\begin{aligned} P(A_n \cap A_{n+1}) &= P(A_n \cap A_{n+1} \cap [X_n = 0]) + P(A_n \cap A_{n+1} \cap [X_n = 1]) = P([X_{n-1} = 1] \cap [X_n = 0] \cap [X_{n+1} = 1]) \\ &\quad + P([X_{n-1} = 0] \cap [X_n = 1] \cap [X_{n+1} = 0]) = p^2(1-p) + p(1-p)^2 = p(1-p). \end{aligned}$$

Et donc, A_n et A_{n+1} sont indépendants si et seulement si $p(1-p) = 4p^2(1-p)^2 \Leftrightarrow p(1-p) = \frac{1}{4} \Leftrightarrow p = \frac{1}{2}$.
Donc déjà, dans le cas où $p \neq \frac{1}{2}$, les événements A_n ne sont pas indépendants.

□

Exercice 3.3 (QSP ESCP 2018) [ESCP18.QSP03] Facile

Étude de l'indépendance de deux événements

On lance $n \geqslant 2$ fois une pièce équilibrée, et on considère les événements :

- A : «on obtient au plus une fois pile»
- B : «les résultats des différents lancers ne sont pas tous identiques»

Les événements A et B sont-ils indépendants ?

Solution de l'exercice 3.3 Notons que A et B sont indépendants si et seulement si A et $\overline{B}$ sont indépendants.
Or, $\overline{B}$ est l'événement «tous les lancers donnent le même résultat».
Considérons alors la variable aléatoire X égale au nombre de piles obtenus lors des n lancers.
Il est alors évident que $X \hookrightarrow \mathcal{B}\left(n, \frac{1}{2}\right)$ et donc

$$P\left(\overline{B}\right) = P(X = 0) + P(X = n) = \frac{1}{2^n} + \frac{1}{2^n} = \frac{1}{2^{n-1}}.$$

D'autre part,

$$P(A) = P(X = 0) + P(X = 1) = \frac{1}{2^n} + \binom{n}{1}\frac{1}{2^n} = \frac{n+1}{2^n}.$$

Et enfin,

$$P\left(A \cap \overline{B}\right) = P(X = 0) = \frac{1}{2^n}.$$

Détails
Si on a au plus un pile, et que tous les lancer donnent le même résultat, c'est qu'on n'a aucun pile !

Ainsi,

$$P\left(A \cap \overline{B}\right) = P(A)P\left(\overline{B}\right) \Leftrightarrow \frac{1}{2^n} = \frac{n+1}{2^{2n-1}} \Leftrightarrow 2^{n-1} = n+1,$$

ce qui n'est jamais le cas, sauf pour $n = 3$. En effet, une récurrence facile permet de prouver que pour $n \geqslant 4$, $2^{n-1} > n+1$.
Et donc A et $\overline{B}$ sont indépendants si et seulement si $n = 3$, et par conséquent il en est de même de A et B. □

3.2 DÉNOMBREMENTS

EXERCICE 3.4 (QSP ESCP 2017) [ESCP17.QSP04] Facile

Probabilité d'intersection de deux parties aléatoires de $[\![1,n]\!]$.

Soit n un entier supérieur ou égal à 1 et $E = [\![1,n]\!]$ l'ensemble des entiers compris entre 1 et n.
On choisit au hasard deux parties A et B de E, toutes les parties, y compris la partie vide ayant la même probabilité d'être choisie.
Calculer la probabilité de l'événement $[A \cap B = \emptyset]$.

SOLUTION DE L'EXERCICE 3.4 Commençons par dénombrer le nombre de parties de E.
Pour $k \in [\![0,n]\!]$, il y a $\binom{n}{k}$ parties de E formées de k éléments.
Et donc le nombre total de parties de E est $\sum_{k=0}^{n} \binom{n}{k} = 2^n$.
Et donc il y a $2^n \times 2^n = 4^n$ manières de choisir deux parties A et B de E.

Remarque
Ce résultat s'explique assez bien : pour chaque élément de E, on a deux choix : l'inclure dans la partie A ou non.
Et puisque E contient n éléments, cela fait $2\times2\times\cdots\times2 = 2^n$ choix possibles.

Dénombrons à présent le nombre de couples (A,B) de parties de E tels que $A \cap B = \emptyset$.
Pour k fixé, il y a $\binom{n}{k}$ manières de choisir une partie A formée de k éléments.
Et alors, pour choisir une partie B disjointe de E, il faut choisir une partie de l'ensemble des $n-k$ éléments qui ne sont pas dans A : il y a 2^{n-k} tels choix possibles.
Ainsi, le nombre de couples de parties de E cherché est

$$\sum_{k=0}^{n} \binom{n}{k} 2^{n-k} = \sum_{k=0}^{n} \binom{n}{k} 1^k 2^{n-k} = (1+2)^n = 3^n.$$

Et donc

$$P(A \cap B = \emptyset) = \frac{3^n}{4^n} = \left(\frac{3}{4}\right)^n.$$

□

EXERCICE 3.5 (QSP ESCP 2012) [ESCP12.QSP01] Moyen

Probabilité que des points soient alignés.
Abordable en première année : ✓

Soit $E = [\![1,4]\!] \times [\![1,4]\!] \subset \mathbf{R}^2$. On choisit une partie de E formée de 4 points.

1. Quelle est la probabilité que les quatre points soient alignés ?
2. Quelle est la probabilité d'avoir au moins trois points alignés ?

SOLUTION DE L'EXERCICE 3.5

1. Notons que E est de cardinal 16 et qu'il y a donc $\binom{16}{4}$ parties de E formées de 4 points.
Or, dans E, il y a en tout dix parties de E formées de 4 points alignés : les quatre droites verticales de E, les quatre droites horizontales, ainsi que les deux diagonales du carré E. Et donc la probabilité que les quatre points soient alignés vaut

$$\frac{10}{\binom{16}{4}} = \frac{10\times24}{16\times15\times14\times13} = \frac{1}{14\times13} = \frac{1}{182}.$$

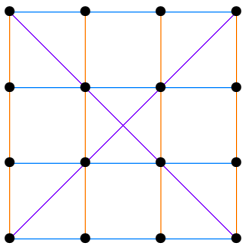


FIGURE 3.1 – L'ensemble E et ses dix droites.

2. Pour que trois points soient alignés, il faut soit :
- ►que trois des points soient situés sur l'une des droites précédemment évoquées
- ►que trois points soient sur l'une des droites situées au-dessus ou en-dessous d'une des diagonales :

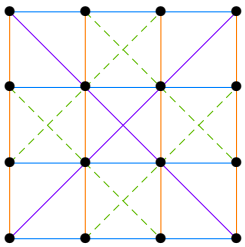


FIGURE 3.2 – Les dix droites et les quatre «sous-diagonales» de E.

Pour chacune des dix droites, il y a deux cas de figure à distinguer :
- ►soit on choisit les quatre points de la droite, et il n'y a qu'une manière de faire ceci.
- ►soit on choisit trois points de la droite : il y a $\binom{4}{3} = 4$ tels choix possibles, et le quatrième point non aligné avec les trois premiers : il reste donc 12 choix possibles.

En revanche, pour chacune des «sous-diagonales», il faut choisir les trois points de cette sous-diagonale, ainsi qu'un quatrième point parmi les 13 points restants.
Au total, la probabilité d'avoir au moins trois points alignés est

$$\frac{10\times(1+4\times 12)+4\times 13}{\binom{16}{4}} = \frac{542\times 24}{16\times 15\times 14\times 13} = \frac{271}{5\times 14\times 13} = \frac{271}{910}.$$

□

EXERCICE 3.6 (ESCP 2012) [ESCP12.3.2] Moyen

Nombre moyen de records lors de tirages sans remise.
Abordable en première année : ✓

Un urne contient initialement n boules numérotées depuis 1 jusqu'à n, avec $n \geqslant 2$. On vide l'urne en extrayant toutes les boules une à une, au hasard et sans remise.

1. Pour i compris entre 1 et n, on note X_i la variable aléatoire qui vaut 1 si la boule obtenue au $i^{\text{ème}}$ tirage porte le numéro i et 0 dans le cas contraire.
   Quelle est la loi de X_i ?
2. En déduire l'espérance du nombre de fois où il y a coïncidence entre le rang du tirage et le numéro de la boule obtenue, lorsque l'on vide l'urne.
3. Pour k compris entre 1 et n, on dit que le résultat du $k^{\text{ème}}$ tirage est un «record» si la boule obtenue à ce tirage porte un numéro strictement supérieur à tous les numéros obtenus jusqu'alors (par convention, le résultat du premier tirage sera toujours considéré comme un record).
   (a) Combien existe-t-il de façons de vider l'urne et pour lesquelles il n'y a qu'un seul record ? Pour lesquelles il y a n records ?

(b) Montrer que pour p et q entiers naturel, on a la relation suivante :

$$\binom{p}{p}+\binom{p+1}{p}+\cdots+\binom{p+q}{p}=\binom{p+q+1}{p+1}.$$

4. Soit k fixé entre 2 et n et j fixé entre k et n. Combien existe-t-il de façons de vider l'urne pour lesquelles la $k^{\text{ème}}$ boule obtenue porte le numéro j et $k^{\text{ème}}$ tirage constitue un record ?
5. Combien existe-t-il de façons de vider l'urne pour lesquelles le $k^{\text{ème}}$ tirage est un record ? En déduire la probabilité que le $k^{\text{ème}}$ tirage soit un record ? Pouvait-on avoir ce résultat directement ?
6. Pour k compris entre 1 et n, soit Y_k la variable aléatoire qui prend la valeur 1 si le résultat du $k^{\text{ème}}$ tirage est un record et 0 sinon.
Déterminer la loi de Y_k. Soit R le nombre aléatoire de records obtenus lorsqu'on vide l'urne. Déterminer l'espérance de R.

SOLUTION DE L'EXERCICE 3.6

1. Comme X_i ne prend que les valeurs 0 et 1, elle suit une loi de Bernoulli, et donc il s'agit de déterminer $P(X_i=1)$.
Puisqu'on effectue n tirages sans remise, le nombre total de tirages possibles est $n!$. En effet, il y a n choix pour la première boule, puis $n-1$ pour la deuxième, etc, et un seul choix pour la $n^{\text{ème}}$, soit $n\times(n-1)\times 1=n!$
Comptons alors le nombre de tirages tels que la $i^{\text{ème}}$ boule soit la boule numéro i.
Puisque le numéro de la $i^{\text{ème}}$ boule est fixé, il suffit de compter le nombre de manière des placer les $n-1$ boules ne portant pas le numéro i parmi les $n-1$ autres tirages.
Il y a alors $(n-1)!$ choix possibles, de sorte que $P(X_i=1)=\dfrac{(n-1)!}{n!}=\dfrac{1}{n}$.

2. Notons C le nombre de coïncidences entre rang du tirage et numéro de la boule tirée.
On a alors $C=\sum_{i=1}^{n}X_i$, et donc par linéarité de l'espérance,

$$E(C)=\sum_{i=1}^{n}E(X_i)=\sum_{i=1}^{n}\frac{1}{n}=1.$$

3.a. Au moment où la boule numéro n sort de l'urne, cela constitue nécessairement un record. Pour qu'il n'y ait qu'un seul record, il faut donc nécessairement que la boule numéro n sorte lors du premier tirage.
Il y a alors $(n-1)!$ tirages possibles vérifiant cette condition : il faut répartir les $n-1$ boules numérotées de 1 à $n-1$ sur les tirages de rang $2,\dots,n$.

Pour qu'il y ait n records, il faut que chaque tirage constitue un record. La boule n doit alors avoir été obtenue lors du $n^{\text{ème}}$ tirage, sinon les tirages suivant celui de la boule n ne pourraient être des records.
Et donc la boule $n-1$ doit être obtenue lors du $(n-1)^{\text{ème}}$ tirage, et plus généralement, la boule i doit être obtenue lors du $i^{\text{ème}}$ tirage.
Et donc il existe un seul tirage pour lequel il y a n records : celui où on tire les boules dans l'ordre croissant.

3.b. Pour tout $k\in\mathbf{N}$, on a $\binom{k}{p}+\binom{k}{p+1}=\binom{k+1}{p+1}$, et donc $\binom{k}{p}=\binom{k+1}{p+1}-\binom{k}{p+1}$.
On en déduit que

$$\begin{aligned}\binom{p}{p}+\binom{p+1}{p}+\binom{p+q}{p}&=\sum_{k=p}^{p+q}\binom{k}{p}\\&=\sum_{k=p}^{p+q}\binom{k+1}{p+1}-\binom{k}{p+1}\\&=\sum_{k=p}^{p+q}\binom{k+1}{p+1}-\sum_{k=p}^{p+q}\binom{k}{p+1}\end{aligned}$$

$$= \binom{p+q+1}{p+1} - \binom{p}{p+1} = \binom{p+q+1}{p+1}.$$

4. Si $k > j$, alors l'un[1] des tirages $1, \dots, k-1$ doit porter un numéro supérieur ou égal j, et donc si le $k^{\text{ème}}$ tirage donne la boule j, celui-ci ne peut constituer un record. Ainsi, pour $k > j$, il n'y a aucun tirage vérifiant les conditions requises.
Pour $k \leqslant j$, un tirage vérifie les conditions demandées si et seulement si les $k-1$ premières boules portent un numéro inférieur ou égal à $j-1$ et que la $k^{\text{ème}}$ porte le numéro j.
Or, il y a $\binom{j-1}{k-1}$ manières de choisir les numéros des boules des tirages 1 à $k-1$, et une fois ces numéros choisis, il y a $(k-1)!$ choix possibles pour l'ordre dans lequel sortent ces boules.
Les boules des tirages $k+1$ à n sont alors les boules qui ne sont pas sorties lors des k premiers tirages : il y en a $n-k$, et donc il y a $(n-k)!$ choix possibles pour les tirages $k+1, \dots, n$.
Donc le nombre de tirages cherché est $\binom{j-1}{k-1} \times (k-1)! \times (n-k)!$

[1] Au moins.

5. En distinguant différents cas suivant le numéro porté par la $k^{\text{ème}}$ boule, et en utilisant la question précédente, le nombre de façons de vider l'urne pour lesquelles le $k^{\text{ème}}$ tirage est un record est

$$\sum_{j=k}^{n} \binom{j-1}{k-1}(k-1)!(n-k)! = (k-1)!(n-k)! \sum_{j=k}^{n} \binom{j-1}{k-1} = (k-1)!(n-k)!\binom{n}{k} = \frac{n!}{k!(n-k)!}(k-1)!(n-k)! = \frac{n!}{k}.$$

Et donc, les $n!$ manières de vider l'urne étant équiprobables, la probabilité d'obtenir un record lors du $k^{\text{ème}}$ tirage est $\dfrac{n!}{k \times n!} = \dfrac{1}{k}$.

6. Grâce à ce que nous venons de dire, Y_k suit une loi de Bernoulli de paramètre $\dfrac{1}{k}$.
On a alors $R = \sum_{k=1}^{n} Y_k$, et donc par linéarité de l'espérance, $E(R) = \sum_{k=1}^{n} E(Y_k) = \sum_{k=1}^{n} \dfrac{1}{k}$.

□

3.3 Variables aléatoires discrètes

Exercice 3.7 (QSP ESCP 2017) [ESCP17.QSP05] — Difficile

Majoration de la variance d'une variable aléatoire à valeurs dans $[a, b]$

Soit X une variable aléatoire réelle prenant ses valeurs dans un segment $[a, b]$.

1. Montrer que X admet une variance et que $V(X) \leqslant \dfrac{(b-a)^2}{4}$.
2. Peut-on avoir égalité dans l'inégalité précédente ? Pour quelles variables aléatoires ?

Solution de l'exercice 3.7

1. Commençons par le cas particulier où $a = 0$ et $b = 1$.
Alors, X étant à valeurs dans $[0, 1]$, $0 \leqslant X^2 \leqslant X \leqslant 1$.
En particulier, la variable certaine égale à 1 admettant une espérance, il en est de même de X^2, et donc $E(X^2)$ existe, de sorte que X admet une variance.
D'autre part, par croissance de l'espérance, $E\left(X^2\right) \leqslant E(X)$ et donc

$$V(X) = E\left(X^2\right) - E(X)^2 \leqslant E(X) - E(X)^2 = E(X)\left(1 - E(X)\right).$$

Or, il est bien connu[1] que $\forall x \in [0, 1]$, $x(1-x) \leqslant \dfrac{1}{4}$.
Et donc $V(X) \leqslant \dfrac{1}{4}$.

[1] Une rapide étude de fonction permet de le retrouver.

Passons au cas général, et posons $Y = \dfrac{X-a}{b-a}$. Puisque $X - a$ est à valeurs dans $[0, b-a]$, Y est à valeurs dans $[0, 1]$.

Et donc par ce qui précède, Y admet une variance, et $V(Y) \leqslant \frac{1}{4}$.

Et donc $V(X) = V((b-a)Y + a) = (b-a)^2 V(Y) \leqslant \frac{(b-a)^2}{4}$.

> **Rappel**
> X admet une variance si et seulement si $\lambda X + \mu$ admet une variance et alors
> $$V(\lambda X + \mu) = \lambda^2 V(X).$$

2. Il s'agit de remonter les calculs faits précédemment et d'étudier à quelles conditions toutes les inégalités utilisées sont des égalités.

On a donc $V(X) = \frac{(b-a)^2}{4}$ si et seulement si $V(Y) = \frac{1}{4}$.

Or, pour $x \in [0,1]$, on a $x(1-x) = \frac{1}{4}$ si et seulement si $x = \frac{1}{2}$.

Donc pour que $V(Y) = \frac{1}{2}$, il faut déjà que $E(Y) = \frac{1}{2}$.

De plus, il faut également avoir $E\left(Y^2\right) = E(Y)$.

Nous avons déjà dit que $Y - Y^2 \leqslant 0$. Donc par l'inégalité de Markov, si $E(Y) = E\left(Y^2\right)$, alors pour tout $a > 0$,

$$P(Y - Y^2 \geqslant a) \leqslant \frac{E\left(Y - Y^2\right)}{a} = 0$$

de sorte que $Y - Y^2$ est (presque sûrement) égale à 0.

Et donc $Y^2 = Y$, ce qui n'est possible que si Y ne prend que les valeurs 0 et 1.

Et on a alors $E(Y) = P(Y = 1)$, de sorte que $P(Y = 1) = \frac{1}{2}$ et donc $P(Y = 0) = \frac{1}{2}$.

Autrement dit, Y est une variable de Bernoulli de paramètre $\frac{1}{2}$.

Et donc la loi de $X = (b-a)Y + a$ est donnée par

$$P(X = a) = P(Y = 0) = \frac{1}{2} \text{ et } P(X = b) = P(Y = 1) = \frac{1}{2}.$$

Inversement, si X ne prend que les valeurs a et b de manières équiprobables, il est facile de vérifier que

$$E(X) = \frac{b+a}{2},\ E\left(X^2\right) = \frac{b^2+a^2}{2} \text{ et donc } V(X) = \frac{a^2+b^2}{2} - \frac{(a+b)^2}{4} = \frac{(b-a)^2}{4}.$$

Et donc $V(X) = \frac{(b-a)^2}{4}$ si et seulement si X est une variable discrète, de support $\{a, b\}$, avec $P(X = a) = P(X = b) = \frac{1}{2}$.

□

EXERCICE 3.8 (QSP HEC 2013) [HEC13.QSP88] — Facile

La loi d'une variable finie est déterminée par ses moments.
Abordable en première année : ✓

1. Soit Y une variable aléatoire définie sur un espace probabilisé $(\Omega, \mathcal{A}, P)$, qui prend les valeurs 0, 1 et 2 avec les probabilités p_0, p_1 et p_2 respectivement. On suppose que $E(Y) = 1$ et $E(Y^2) = \frac{5}{3}$.
   Calculer p_0, p_1 et p_2.
2. Soit $x_0, x_1, \ldots, x_n$ $(n+1)$ réels distincts, et soit φ l'application de $\mathbf{R}_n[X]$ dans $\mathbf{R}^{n+1}$ qui, à tout polynôme Q de $\mathbf{R}_n[X]$, associe le $(n+1)$-uplet $(Q(x_0), Q(x_1), \ldots, Q(x_n))$.
   (a) Montrer que φ est une application linéaire bijective.
   (b) Déterminer la matrice Φ de φ dans les bases canoniques respectives de $\mathbf{R}_n[X]$ et $\mathbf{R}^{n+1}$.
3. Soit X une variable aléatoire discrète qui prend les valeurs $x_0, x_1, \ldots, x_n$.
   On suppose que $E(X), E(X^2), \ldots, E(X^n)$ sont connus. Peut-on déterminer la loi de X ?

SOLUTION DE L'EXERCICE 3.8

1. Puisque $Y(\Omega) = \{0, 1, 2\}$, nous savons que $p_0 + p_1 + p_2 = 1$.
D'autre part, en utilisant l'espérance et le moment d'ordre 2 de Y, il vient

$$0 \times p_0 + p_1 + 2p_2 = 1 \text{ et } 0^2 \times p_0 + 1^2 \times p_1 + 4p_2 = \frac{5}{3}.$$

Et donc une rapide résolution de système donne $p_0 = p_1 = p_2 = \frac{1}{3}$.
Autrement dit, Y suit la loi uniforme sur $[\![0, 2]\!]$.

2.a. La linéarité de φ est triviale.
Soit $Q \in \operatorname{Ker}\varphi$. Alors $Q(x_0) = Q(x_1) = \cdots = Q(x_n) = 0$, et donc Q possède $n+1$ racines distinctes[1], et est de degré au plus n : c'est donc le polynôme nul.
Ainsi, $\operatorname{Ker}\varphi = \{0\}$, et donc φ est injective.
Puisque $\dim \mathbf{R}_n[X] = \dim \mathbf{R}^{n+1} = n+1$, φ est donc bijective.

[1] Notons qu'ici il est très important que les x_i soient distincts.

2.b. Notons $e_0 = (1, 0, \ldots, 0), e_1 = (0, 1, 0, \ldots, 0), \ldots, e_n = (0, \ldots, 0, 1)$ la base canonique de $\mathbf{R}^{n+1}$.
On a alors $\varphi(1) = (1, 1, \ldots, 1) = e_0 + e_1 + \cdots + e_n$, et pour $k \in [\![1, n]\!]$,

$$\varphi(X^k) = (x_0^k, x_1^k, \ldots, x_n^k) = x_0^k e_0 + x_1^k e_1 + \cdots + x_n^k e_n.$$

Ainsi, la matrice Φ de φ dans les bases canoniques de $\mathbf{R}_n[X]$ et $\mathbf{R}^{n+1}$ est

$$\Phi = \begin{array}{c} \begin{matrix} \varphi(1) & \varphi(X) & \varphi(X^2) & \ldots & \varphi(X^n) \end{matrix} \\ \begin{pmatrix} 1 & x_0 & x_0^2 & \ldots & x_0^n \\ 1 & x_1 & x_1^2 & \ldots & x_1^n \\ 1 & x_2 & x_2^2 & \ldots & x_2^n \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & x_n & x_n^2 & \ldots & x_n^n \end{pmatrix} \end{array} \begin{matrix} e_0 \\ e_1 \\ e_2 \\ \vdots \\ e_n \end{matrix} \in \mathcal{M}_{n+1}(\mathbf{R}).$$

3. Notons $p_0 = P(X = x_0), p_1 = P(X = x_1), \ldots, p_n = P(X = x_n)$.
Il s'agit donc de déterminer les p_i.
Puisque $X(\Omega) = \{x_0, x_1, \ldots, x_n\}$, on a $p_0 + p_1 + \cdots + p_n = 1$, et par le théorème de transfert, pour tout $k \in [\![1, n]\!]$,

$$E\left(X^k\right) = x_0^k p_0 + x_1^k p_1 + \cdots + x_n^k p_n.$$

Il s'agit donc de déterminer si le système $\begin{cases} p_0 + p_1 + \cdots + p_n &= 1 \\ x_0 p_0 + x_1 p_1 + \cdots + x_n p_n &= E(X) \\ & \vdots \\ x_0^n p_0 + x_1^n p_1 + \cdots + x_n^n p_n &= E\left(X^n\right) \end{cases}$ possède une solution ou non.
Sous forme matricielle, ce système s'écrit

$$\begin{pmatrix} 1 & 1 & \ldots & 1 \\ x_0 & x_1 & \ldots & x_n \\ x_0^2 & x_1^2 & \ldots & x_n^2 \\ \vdots & \vdots & & \vdots \\ x_0^n & x_1^n & \ldots & x_n^n \end{pmatrix} \begin{pmatrix} p_0 \\ p_1 \\ p_2 \\ \vdots \\ p_n \end{pmatrix} = \begin{pmatrix} 1 \\ E(X) \\ E\left(X^2\right) \\ \vdots \\ E\left(X^n\right) \end{pmatrix} \Leftrightarrow {}^t\Phi \begin{pmatrix} p_0 \\ p_1 \\ p_2 \\ \vdots \\ p_n \end{pmatrix} = \begin{pmatrix} 1 \\ E(X) \\ E\left(X^2\right) \\ \vdots \\ E\left(X^n\right) \end{pmatrix}.$$

Mais φ étant bijective, sa matrice Φ est inversible, et donc il en est de même de ${}^t\Phi$.
Et donc le système possède une unique solution $(p_0, p_1, \ldots, p_n)$: la loi de X est uniquement déterminée par la donnée de $E(X), E\left(X^2\right), \ldots, E\left(X^n\right)$.

□

Précision
Ce qui nous intéresse est en fait de savoir si ce système possède **une unique** solution, c'est-à-dire de savoir si la loi d

Remarques ent déter-

Notons que l'énoncé faisait tout de même une hypothèse relativement forte : on suppose que $x_0, x_2, \ldots, x_n$ sont connus.
Si on ne connaît que les n premiers moments de X et pas son support, alors la méthode employée ici ne fonctionne plus.
Notons également que si X prend $n+1$ valeurs, il suffit de connaître les n premiers moments (soit un de moins que le cardinal du support).

L'exercice suivant pourrait tout à fait être posé à l'écrit à l'EDHEC, et il s'agit de probabilités on ne peut plus classiques.

Exercice 3.9 (ESCP 2015) [ESCP15.3.05] — Facile

Tirages dans une urne au contenu évolutif

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 $k \geqslant 1$, 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 $n \geqslant 2$, montrer que
   $$\forall j \in \mathbf{N}^*,\ \forall k \in [\![2, n]\!],\ P_{[I_n=k]}(X_n = j) = P(X_{k-1} = j - 1).$$
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 : $\forall n \geqslant 2,\ \forall j \geqslant 2,\ P(X_n = j) = \dfrac{1}{n}\displaystyle\sum_{k=1}^{n-1} P(X_k = j - 1)$.
   (d) Pour tout $n \geqslant 2$ et tout $j \geqslant 2$, calculer $nP(X_n = j) - (n-1)P(X_{n-1} = j)$.
   En déduire que pour tout $n \geqslant 2$ et tout $j \geqslant 1$:
   $$P(X_n = j) = \frac{n-1}{n}P(X_{n-1} = j) + \frac{1}{n}P(X_{n-1} = j - 1).$$
4. Montrer que : $\forall n \geqslant 2,\ E(X_n) = E(X_{n-1}) + \dfrac{1}{n}$. En déduire $E(X_n)$.

SOLUTION DE L'EXERCICE 3.9

1.a. Il est évident que I_n suit la loi uniforme sur $[\![1, n]\!]$.

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 tirages[1] est celle qu'on effectue $j - 1$ tirages après le premier, c'est-à-dire en partant d'une urne contenant $k - 1$ boules. Et donc
$$P_{[I_n=k]}(X_n = j) = P(X_{k-1} = j - 1).$$

[1] 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 $P(X_2 = 1) = \dfrac{1}{2}$.
Et donc $P(X_2 = 2) = 1 - P(X_2 = 1) = \dfrac{1}{2}$.
Autrement dit, X_2 suit la loi uniforme sur $[\![1, 2]\!]$, et donc $E(X_2) = \dfrac{3}{2}$.

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 $X_n(\Omega) = [\![1, n]\!]$.

3.b. On a $P(X_n = 1) = P(I_n = 1) = \dfrac{1}{n}$.
Et $[X_n = n] = [Y_1 = n] \cap [Y_2 = n - 2] \cap \cdots \cap [Y_n = 1]$.

Par la formule des probabilités composées, on a donc

$$\begin{aligned}P(X_n = n) &= P(Y_1 = n)P_{[Y_1=n]}(Y_2 = n-1) \times \cdots \times P_{[Y_1=1]\cap\cdots\cap[Y_{n-1}=2]}(Y_n = 1)\\ &= \frac{1}{n}\frac{1}{n-1}\cdots\frac{1}{1}\\ &= \frac{1}{n!}.\end{aligned}$$

3.c. Utilisons la formule des probabilités totales avec le système complet d'événements $\{[I_n = k],\ k \in [\![1, n]\!]\}$. Alors pour tout $j \geqslant 2$,

$$\begin{aligned}P(X_2 = j) &= \sum_{k=1}^{n} P(I_n = k)P_{[I_n=k]}(X_n = j)\\ &= \sum_{k=1}^{n} \frac{1}{n}P(X_{k-1} = j-1) && \text{C'est la question 1.c.}\\ &= \sum_{k=2}^{n} \frac{1}{n}P(X_{k-1} = j-1) && \text{Puisque } j \geqslant 2, \text{ on ne peut avoir tiré la boule 1 au premier tirage.}\\ &= \frac{1}{n}\sum_{k=1}^{n-1} P(X_k = j-1).\end{aligned}$$

3.d. En utilisant la relation de la question précédente, on a, pour tout $j \geqslant 2$,

$$nP(X_n = j) - (n-1)P(X_{n-1} = j) = \sum_{k=1}^{n-1} P(X_k = j-1) - \sum_{k=1}^{n-2} P(X_k = j-1) = P(X_{n-1} = j-1).$$

Et donc, toujours pour $j \geqslant 2$,

$$P(X_n = j) = \frac{1}{n}\left(nP(X_n = j) - (n-1)P(X_{n-1} = j)\right) + \frac{n-1}{n}P(X_{n-1} = j) = \frac{n-1}{n}P(X_{n-1} = j) + \frac{1}{n}P(X_{n-1} = j-1).$$

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,

> **Remarque**
> Puisque X_n est une variable finie, elle admet automatiquement une espérance.

$$\begin{aligned}E(X_n) &= \sum_{j=1}^{n} jP(X_n = j)\\ &= \sum_{j=1}^{n} \frac{n-1}{n}jP(X_{n-1} = j) + \sum_{j=1}^{n} \frac{1}{n}jP(X_{n-1} = j-1)\\ &= \frac{n-1}{n}E(X_{n-1}) + \frac{1}{n}\sum_{k=2}^{n} jP(X_{n-1} = j-1) && \text{Le terme en } j = 1 \text{ est nul.}\\ &= \frac{n-1}{n}E(X_{n-1}) + \frac{1}{n}\sum_{i=1}^{n-1}(i+1)P(X_{n-1} = i)\\ &= \frac{n-1}{n}E(X_{n-1}) + \frac{1}{n}\underbrace{\sum_{i=1}^{n-1} iP(X_{n-1} = i)}_{=E(X_{n-1})} + \frac{1}{n}\underbrace{\sum_{i=1}^{n-1} P(X_{n-1} = i)}_{=1}\\ &= E(X_{n-1}) + \frac{1}{n}.\end{aligned}$$

Il vient alors

$$E(X_n) = \sum_{k=2}^{n}\left(E(X_k) - E(X_{k-1})\right) + E(X_1) = \sum_{k=2}^{n}\frac{1}{k} + E(X_1).$$

Mais X_1 étant la variable certaine égale à 1, on a $E(X_1) = 1$ et donc

$$E(X_n) = \sum_{k=1}^{n} \frac{1}{k}.$$

□

EXERCICE 3.10 (QSP ESCP 2011) [ESCP11.QSP02] Facile

Nombre moyens de lancers d'une pièce avant l'obtention de deux résultats identiques consécutifs.

On lance une pièce équilibrée jusqu'à obtenir deux piles consécutifs ou deux faces consécutifs.
On note X la variable aléatoire égale au nombre de lancers effectués.
Donner la loi de X et son espérance.

SOLUTION DE L'EXERCICE 3.10 Remarquons que $X(\Omega) = [\![2, \infty[\![$.
Notons F_n (resp. P_n) l'événement «le $n^{\text{ème}}$ lancer a donné pile (resp. face)».
On a alors $[X = 2k] = (P_1 \cap F_2 \cap \cdots \cap F_{2k-2} \cap P_{2k-1} \cap P_{2k}) \cup (F_1 \cap P_2 \cap \cdots \cap P_{2k-2} \cap F_{2k-1} \cap F_{2k})$.
Par incompatibilité de ces deux événements, et par indépendance des lancers, on a donc

$$\begin{aligned} P(X = 2k) &= P(P_1)P(F_2)\cdots P(F_{2k-2})P(P_{2k-1})P(P_{2k}) + P(F_1)P(P_2)\cdots P(P_{2k-2})P(F_{2k-1})P(F_{2k}) \\ &= \frac{1}{2} \times \cdots \times \frac{1}{2} + \frac{1}{2} \times \cdots \times \frac{1}{2} \\ &= 2\frac{1}{2^{2k}} = \frac{1}{2^{2k-1}}. \end{aligned}$$

De la même manière,

$$[X = 2k+1] = (P_1 \cap F_2 \cap \cdots \cap P_{2k-1} \cap F_{2k} \cap F_{2k+1}) \cup (F_1 \cap P_2 \cap \cdots \cap F_{2k-1} \cap P_{2k} \cap P_{2k+1})$$

et donc $P(X = 2k+1) = 2\frac{1}{2^{2k+1}} = \frac{1}{2^{2k}}$.
Donc pour tout $n \geqslant 2$, $P(X = n) = \frac{1}{2^{n-1}}$.
Et alors, sous réserve de convergence,

$$\begin{aligned} E(X) &= \sum_{k=2}^{+\infty} kP(X = k) = \sum_{k=2}^{+\infty} k\frac{1}{2^{k-1}} \\ &= \sum_{k=1}^{+\infty} k\frac{1}{2^{k-1}} - 1 \\ &= \frac{1}{\left(1 - \frac{1}{2}\right)^2} - 1 \\ &= 3. \end{aligned}$$

On a reconnu une série géométrique dérivée de raison $\frac{1}{2}$, donc convergente, de sorte que $E(X)$ existe.

Pour aller plus loin *: l'exercice se généralise facilement au cas d'une pièce non équilibrée, même si le calcul de l'espérance est alors un peu plus subtil.* □

Remarque

Nous avons distingué le cas où n est pair ($n = 2k$) du cas où n est impair car il est plus facile d'écrire alors $[X = n]$ avec les événements P_i et F_i, mais il est sûrement possible de s'en tirer à l'oral sans réellement faire de distinction de cas.
Il faudra tout de même être capable d'écrire les choses proprement si l'examinateur le demande.

EXERCICE 3.11 (QSP HEC 2016) [HEC16.QSP167] Facile

Une variable dont les moments d'ordres 2 et 4 valent 1 ne prend que deux valeurs

Soit X une variable aléatoire discrète admettant des moments jusqu'à l'ordre 4 et telle que $\begin{cases} E(X) = \alpha \\ E\left(X^2\right) = E(X^4) = 1 \end{cases}$

1. Montrer que $\alpha \in [-1, 1]$.
2. Déterminer la loi de X.

Solution de l'exercice 3.11

1. Nous savons que $V(X) \geqslant 0$, et par la formule de Huygens, $V(X) = E\left(X^2\right) - E(X)^2 = 1 - \alpha^2$.
Et donc $1 - \alpha^2 \geqslant 0 \Leftrightarrow \alpha \in [-1, 1]$.
2. La variance de X^2 est donnée par $V(X^2) = E(X^4) - E\left(X^2\right)^2 = 1 - 1 = 0$.
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 $P(X = 1) = \dfrac{1+\alpha}{2}$ et donc $P(X = -1) = \dfrac{1-\alpha}{2}$.

□

3.3.1 Autour des lois usuelles

L'exercice qui suit ne nécessite en fait aucune connaissance de probabilités, si ce n'est celle de la loi de Poisson. Le reste est essentiellement un exercice d'analyse.

Exercice 3.12 (QSP HEC 2012) [HEC12.QSP34] — Facile

Étude asymptotique de la queue d'une loi de Poisson.

Soit X une variable aléatoire définie sur un espace probabilisé $(\Omega, \mathcal{A}, P)$, suivant une loi de Poisson de paramètre $\lambda > 0$.

1. Montrer que pour tout $n \geqslant \lambda - 1$, $P(X \geqslant n) \leqslant P(X = n)\dfrac{n+1}{n+1-\lambda}$.
2. En déduire que $P(X \geqslant n) \underset{n\to+\infty}{\sim} P(X = n)$.
3. Montrer que $P(X > n) = \underset{n\to+\infty}{o}(P(X = n))$.

Solution de l'exercice 3.12

1. On a
$$P(X \geqslant n) = \sum_{k=n}^{+\infty} P(X = k) = \sum_{k=n}^{+\infty} e^{-\lambda}\frac{\lambda^k}{k!} = P(X = n)\sum_{k=n}^{+\infty}\frac{\lambda^{k-n}n!}{k!}.$$
Mais $\dfrac{n!}{k!} = \dfrac{1}{(n+1)(n+2)\cdots k}$ de sorte que
$$\frac{\lambda^{n-k}n!}{k!} = \frac{\lambda}{n+1}\frac{\lambda}{n+2}\cdots\frac{\lambda}{k} \leqslant \left(\frac{\lambda}{n+1}\right)^{n-k}.$$
Puisque $\lambda < n + 1$, on a alors $\dfrac{\lambda}{n+1} < 1$, et donc
$$P(X \geqslant n) \leqslant P(X = n)\sum_{k=n}^{\infty}\left(\frac{\lambda}{n+1}\right)^{n-k} \leqslant P(X = n)\sum_{k=0}^{\infty}\left(\frac{\lambda}{n+1}\right)^{k} = P(X = n)\frac{n+1}{n+1-\lambda}.$$

Détails
La première inégalité vient du fait que
$$[X = n] \subset [X \geqslant n].$$

2. Nous savons que
$$P(X = n) \leqslant P(X \geqslant n) \leqslant P(X = n)\frac{n+1}{n+1-\lambda}$$
donc en divisant les deux membres par $P(X = n)$, et par le théorème des gendarmes, il vient
$$\lim_{n\to+\infty}\frac{P(X \geqslant n)}{P(X = n)} = 1 \Leftrightarrow P(X \geqslant n) \underset{n\to+\infty}{\sim} P(X = n).$$

Rappel
Par définition, on a $u_n \sim v_n$ si et seulement si $u_n = v_n + o(v_n)$.

3. Par l'équivalent précédent, on a $P(X \geqslant n) = P(X = n) + \underset{n\to+\infty}{o}(P(X = n))$.
Mais d'autre part $P(X \geqslant n) = P(X > n) + P(X = n)$ et donc
$$P(X > n) = P(X \geqslant n) - P(X = n) = P(X = n) + \underset{n\to+\infty}{o}(P(X = n)) - P(X = n) = \underset{n\to+\infty}{o}(P(X = n)).$$

□

EXERCICE 3.13 (QSP HEC 2012) [HEC12.QSP39] — Facile

Autour de la loi de Poisson

Soient X et Y deux variables aléatoires indépendantes définies sur le même espace probabilisé $(\Omega, \mathcal{A}, P)$ telles que X suive la loi de Poisson de paramètre $\lambda > 0$, et

$$Y(\Omega) = \{1, 2\},\ P(Y = 1) = P(Y = 2) = \frac{1}{2}.$$

On pose $Z = XY$.

1. Déterminer la loi de Z.
2. Quelle est la probabilité que Z prenne une valeur paire ?

SOLUTION DE L'EXERCICE 3.13

1. Il est évident que $Z(\Omega) = \mathbf{N}$.
Si $k \in \mathbf{N}$ est impair, alors on a

$$P(Z = k) = P([X = k] \cap [Y = 1]) = P(X = k)P(Y = 1) = \frac{1}{2}P(X = k) = \frac{\lambda^k}{2k!}e^{-\lambda}.$$

En revanche, si $k = 2n$ est pair, alors, par la formule des probabilités totales appliquée au système complet d'événements $\{[Y = 1], [Y = 2]\}$, il vient

$$\begin{aligned} P(Z = k) &= P([Z = k] \cap [Y = 1]) + P([Z = k] \cap [Y = 2]) \\ &= P([X = k] \cap [Y = 1]) + P([X = n] \cap [Y = 2]) \\ &= \frac{e^{-\lambda}}{2}\left(\frac{\lambda^k}{k!} + \frac{\lambda^n}{n!}\right) \\ &= \frac{e^{-\lambda}}{2}\lambda^{k/2}\left(\frac{1}{(k/2)!} + \frac{\lambda^{k/2}}{k!}\right). \end{aligned}$$

2. Notons $2\mathbf{N}$ l'ensemble des nombres pairs. Alors, par la formule des probabilités totales appliquée au système complet d'événements $\{[Y = 1], [Y = 2]\}$, on a

$$\begin{aligned} P(Z \in 2\mathbf{N}) &= P([Z \in 2\mathbf{N}] \cap [Y = 1]) + P([Z \in 2\mathbf{N}] \cap [Y = 2]) \\ &= P([X \in 2\mathbf{N}] \cap [Y = 1]) + P(Y = 2) \\ &= P(X \in 2\mathbf{N})P(Y = 1) + \frac{1}{2} \\ &= \frac{1}{2}\sum_{\substack{k=0 \\ k \text{ pair}}}^{+\infty} P(X = k) + \frac{1}{2} \\ &= \frac{e^{-\lambda}}{2}\sum_{\substack{k=0 \\ k \text{ pair}}}^{+\infty} \frac{\lambda^k}{k!} + \frac{1}{2}. \end{aligned}$$

Détails
Si $Y = 2$, alors $Z = 2X$ est automatiquement pair.

Cette dernière série peut se calculer en notant que

$$e^{\lambda} = \sum_{k=0}^{+\infty} \frac{\lambda^k}{k!} = \sum_{\substack{k=0 \\ k \text{ pair}}}^{+\infty} \frac{\lambda^k}{k!} + \sum_{\substack{k=0 \\ k \text{ impair}}}^{+\infty} \frac{\lambda^k}{k!}$$

alors que

$$e^{-\lambda} = \sum_{k=0}^{+\infty} \frac{(-\lambda)^k}{k!} = \sum_{\substack{k=0 \\ k \text{ pair}}}^{+\infty} (-1)^k \frac{\lambda^k}{k!} + \sum_{\substack{k=0 \\ k \text{ impair}}}^{+\infty} (-1)^k \frac{\lambda^k}{k!} = \sum_{\substack{k=0 \\ k \text{ pair}}}^{+\infty} \frac{\lambda^k}{k!} - \sum_{\substack{k=0 \\ k \text{ impair}}}^{+\infty} \frac{\lambda^k}{k!}.$$

Et donc

$$e^{\lambda} + e^{-\lambda} = 2\sum_{\substack{k=0 \\ k \text{ pair}}}^{+\infty} \frac{\lambda^k}{k!}.$$

On en déduit que

$$P(Z \in 2\mathbf{N}) = \frac{e^{-\lambda}}{2}\left(\frac{e^{\lambda}+e^{-\lambda}}{2}+1\right) = \frac{1}{2}\left(\frac{1+e^{-2\lambda}}{2}+1\right).$$

□

3.3.2 Lois conditionnelles, espérance totale

Exercice 3.14 (QSP ESCP 2016) [ESCP16.QSP04] **Facile**

Expérience dont les paramètres dépendent d'une autre expérience

On utilise une pièce de monnaie qui donne pile avec la probabilité $p \in]0,1[$.
On commence par lancer la pièce jusqu'à obtenir une première fois pile, et on note N le nombre de lancers nécessaires. Si le premier pile a été obtenu au $n^{\text{ème}}$ lancer, on lancer ensuite cette même pièce n^2 fois et on note X le nombre de pile obtenus au cours de ces n^2 lancers.

1. Quelle est la loi suivie par N ? Donner l'espérance et la variance de N.
2. Soit $n \in \mathbf{N}^*$. Déterminer la loi conditionnelle de X sachant l'événement $[N = n]$.
3. En déduire l'existence et la valeur de l'espérance de X.

Solution de l'exercice 3.14

1. Il est évident que N suit la loi géométrique de paramètre p.
Et donc $E(N) = \dfrac{1}{p}$, $V(N) = \dfrac{1-p}{p^2}$.

2. Si $[N = n]$, alors la seconde série de lancers de n pièces comporte n^2 lancers.
Et donc par indépendance des lancers successifs, la loi de X sachant $[N = n]$ est une loi $\mathcal{B}(n^2, p)$.

3. Grâce à ce qui vient d'être dit, nous pouvons affirmer que $E(X|N = n) = n^2 p$.
Et donc par la formule de l'espérance totale appliquée au système complet d'événements $\{[N = n],\ n \in \mathbf{N}^*\}$, on a, sous réserve de convergence,

$$\begin{aligned}
E(X) &= \sum_{n=1}^{+\infty} P(N=n)E(X|N=n) \\
&= \sum_{n=1}^{+\infty} p(1-p)^{n-1}n^2 p \\
&= p^2 \sum_{n=1}^{+\infty} n^2(1-p)^{n-1} \\
&= p^2\left((1-p)\sum_{n=1}^{+\infty} n(n-1)(1-p)^{n-2} + \sum_{n=1}^{+\infty} n(1-p)^{n-1}\right) \\
&= p^2\left((1-p)\frac{2}{(1-(1-p))^3} + \frac{1}{(1-(1-p))^2}\right) \\
&= p^2\left(\frac{2(1-p)}{p^3} + \frac{1}{p^2}\right) \\
&= \frac{2(1-p)}{p} + 1 = \frac{2-p}{p}.
\end{aligned}$$

On a reconnu deux séries qui convergent, donc $E(X)$ existe bien.

□

3.4 VECTEURS ALÉATOIRES

EXERCICE 3.15 (ESCP 2017) [ESCP17.3.23] — Moyen

Espérance d'une somme de lois de Rademacher

On admet que $\displaystyle\int_0^{+\infty} \frac{\sin t}{t}\,dt$ converge et que sa valeur est $\frac{\pi}{2}$.

1. Montrer que pour tout n entier naturel, $u_n = \displaystyle\int_0^{+\infty} \frac{1-\cos^n(t)}{t^2}\,dt$ existe et que $u_1 = \frac{\pi}{2}$.
2. (a) Montrer que pour tout s réel $\displaystyle\int_0^{+\infty} \frac{1-\cos(st)}{t^2}\,dt$ converge et vaut $\frac{\pi}{2}|s|$.
   (b) En déduire la valeur de u_2.

Dans la suite, on considère une suite $(X_n)_{n\in\mathbf{N}^*}$ de variables aléatoires mutuellement indépendantes, définies sur $(\Omega, \mathcal{A}, P)$, à valeurs dans $\{-1, 1\}$ et telles que, pour tout $k \in \mathbf{N}^*$,

$$P(X_k = 1) = P(X_k = -1) = \frac{1}{2}.$$

Pour tout $n \in \mathbf{N}^*$, on pose $S_n = X_1 + \cdots + X_n$.

3. Déterminer l'espérance et la variance de S_n.
4. (a) Calculer pour tout réel t, $E(\sin(X_1 t))$ et $E(\cos(X_1 t))$.
   (b) Montrer par récurrence sur n, que pour tout entier $n \in \mathbf{N}^*$ et tout réel t, on a : $E(\cos(S_n t)) = (\cos t)^n$.
5. Montrer que pour tout $n \in \mathbf{N}^*$, $E(|S_n|) = \frac{2}{\pi} u_n$.

SOLUTION DE L'EXERCICE 3.15

1. Notons que $u_0 = \displaystyle\int_0^{+\infty} 0\,dt$, et qu'il n'y donc rien à dire. Nous supposerons donc $n \geqslant 1$.

L'intégrale est impropre en 0 et en $+\infty$.

Au voisinage de $+\infty$, on a $-1 \leqslant \cos^n(t) \leqslant 1$ et donc $0 \leqslant \dfrac{1-\cos^n t}{t^2} \leqslant \dfrac{2}{t^2}$.

Et donc par comparaison à une intégrale de Riemann, $\displaystyle\int_1^{+\infty} \frac{1-\cos^n(t)}{t^2}\,dt$ converge.

Au voisinage de 0, on a

$$1-\cos^n(t) = (1-\cos t)(1+\cos t+\cos^2 t+\cdots+\cos^{n-1} t) \underset{t\to 0}{\sim} n(1-\cos t) \underset{t\to 0}{\sim} n\frac{t^2}{2}.$$

Et donc $\displaystyle\lim_{t\to 0^+} \frac{1-\cos^n t}{t^2} = \frac{n}{2}$, de sorte que $\displaystyle\int_0^1 \frac{1-\cos^n t}{t^2}\,dt$ est faussement impropre, donc convergente.

Pour calculer u_1, procédons à une intégration par parties sur un segment de la forme $[A, B] \subset \mathbf{R}_+^*$:

$$\int_A^B \frac{1-\cos t}{t^2}\,dt = \left[\frac{\cos t - 1}{t}\right]_A^B + \int_A^B \frac{\sin t}{t}\,dt = \frac{\cos(B)-1}{B} - \frac{\cos(A)-1}{A} + \int_A^B \frac{\sin t}{t}\,dt.$$

Or, lorsque $A \to 0$, $\dfrac{\cos A - 1}{A} \underset{A\to 0^+}{\sim} -\dfrac{A}{2} \underset{A\to 0^+}{\longrightarrow} 0$.

Et lorsque $B \to +\infty$, la fonction cos étant bornée, $\dfrac{\cos B - 1}{B} \underset{B\to+\infty}{\longrightarrow} 0$.

Et donc $u_1 = \displaystyle\int_0^{+\infty} \frac{\sin t}{t}\,dt = \frac{\pi}{2}$.

> **Rappel**
> Le produit d'une fonction bornée par une fonction de limite nulle tend vers 0.

2.a. Le cas $s = 0$ ne pose aucune difficulté puisque $\cos(0) = 1$.

Supposons donc $s > 0$, et procédons au changement de variable[1] $u = st$: alors $\displaystyle\int_0^{+\infty} \frac{1-\cos(st)}{t^2}\,dt$

[1] Affine, donc légitime.

converge si et seulement si $\displaystyle\int_0^{+\infty} \frac{1-\cos(u)}{u^2} s^2 \frac{du}{s} = s u_1$ converge, et en cas de convergence, ces deux intégrales sont égales.

Or nous venons de voir que u_1 converge et vaut $\dfrac{\pi}{2}$, de sorte que $\displaystyle\int_0^{+\infty} \frac{1-\cos(st)}{t^2}\,dt$ converge et vaut $\dfrac{\pi}{2}s = \dfrac{\pi}{2}|s|$.

Enfin, si $s<0$, notons que par parité du cosinus, remplacer s par $-s$ ne change rien[2], et donc $\displaystyle\int_0^{+\infty} \frac{1-\cos(st)}{t^2}\,dt = \int_0^{+\infty} \frac{1-\cos(-st)}{t^2}\,dt = \frac{\pi}{2}(-s) = \frac{\pi}{2}|s|$.

[2] Et $-s$ étant positif, on peut lui appliquer ce qui précède.

Alternative

Si vous n'aimez pas les formules de trigonométrie, on peut toujours utiliser les formules d'Euler pour linéariser

$$\cos^2 t = \left(\frac{e^{it}+e^{-it}}{2}\right)^2.$$

2.b. Nous savons que $\cos(2t) = 2\cos^2 t - 1$ et donc $\cos^2 t = \dfrac{1+\cos(2t)}{2}$ de sorte que $1-\cos^2 t = \dfrac{1-\cos(2t)}{2}$.

Et donc $\displaystyle u_2 = \int_0^{+\infty} \frac{1-\cos^2(t)}{t^2}\,dt = \frac{1}{2}\int_0^{+\infty} \frac{1-\cos(2t)}{t^2}\,dt = \frac{\pi}{2}$.

3. Il est évident que les X_k sont centrées, et que donc $V(X_k) = E\left(X_k^2\right) = 1$.
Et donc par linéarité de l'espérance, et en utilisant l'indépendance des X_k, il vient

$$E(S_n) = \sum_{k=1}^{n} E(X_k) = 0 \text{ et } V(S_n) = \sum_{k=1}^{n} V(X_k) = n.$$

4.a. Il s'agit d'un théorème de transfert[3] :

[3] Version discrète !

$$E(\sin(X_1 t)) = \frac{1}{2}\sin(t) + \frac{1}{2}\sin(-t) = 0 \text{ et } E(\cos(X_1 t)) = \frac{1}{2}(\cos t + \cos(-t)) = \cos t.$$

4.b. La récurrence est déjà initialisée pour $n=1$.
Supposons donc que $E(\cos(S_n t)) = (\cos t)^n$. Alors nous savons que

$$\cos(S_{n+1}t) = \cos\left((S_n + X_{n+1})\,t\right) = \cos(S_n t + X_{n+1} t) = \cos(S_n t)\cos(X_{n+1}t) - \sin(S_n t)\sin(X_{n+1}t).$$

Donc par linéarité de l'espérance,

$$E(\cos(S_{n+1}t)) = E\left(\cos(S_n t)\cos(X_{n+1}t)\right) - E\left(\sin(S_n t)\sin(X_{n+1}t)\right).$$

Et alors par indépendance[4] de $\cos(S_n t)$ et de $\cos(X_{n+1}t)$ (resp. de $\sin(S_n t)$ et de $\sin(X_{n+1}t)$), il vient

[4] Ce qui découle du lemme des coalitions.

$$E(\cos(S_{n+1}t)) = E(\cos(S_n t))E(\cos(X_{n+1}t)) - E(\sin(S_n t))\underbrace{E(\sin(X_{n+1}t))}_{=0} = (\cos t)^{n+1}.$$

Et donc par le principe de récurrence, pour tout $n \in \mathbf{N}^*$, $E(\cos(S_n t)) = (\cos t)^n$.

5. Par le théorème de transfert, on a

$$E(|S_n|) = \sum_{k \in S_n(\Omega)} |k| P(S_n = k).$$

Remarque

Puisque S_n est une variable finie (elle prend des valeurs dans $[\![-n, n]\!]$), la somme qui apparaît est une somme finie.

Or, nous savons d'après 2.a que $|k| = \dfrac{2}{\pi}\displaystyle\int_0^{+\infty} \frac{1-\cos(kt)}{t^2}\,dt$.

Et donc

$$\begin{aligned}
E(|S_n|) &= \sum_{k \in S_n(\Omega)} P(S_n = k)\frac{2}{\pi}\int_0^{+\infty} \frac{1-\cos(kt)}{t^2}\,dt \\
&= \frac{2}{\pi}\int_0^{+\infty} \sum_{k \in S_n(\Omega)} P(S_n = k)\frac{1-\cos(kt)}{t^2}\,dt \\
&= \frac{2}{\pi}\int_0^{+\infty} \frac{1}{t^2}\left(1 - \sum_{k \in S_n(\Omega)} P(S_n = k)\cos(kt)\right)dt \\
&= \frac{2}{\pi}\int_0^{+\infty} \frac{1}{t^2}\left(1 - E(\cos(S_n t))\right)\,dt \\
&= \frac{2}{\pi}\int_0^{+\infty} \frac{1-\cos^n(t)}{t^2}\,dt = \frac{2}{\pi}u_n.
\end{aligned}$$

Somme finie

La somme étant finie, on peut bien intervertir sans précaution la somme et l'intégrale.

Nous avons utilisé le fait que «la somme des probas vaut 1».

Théorème de transfert.

□

EXERCICE 3.16 (HEC 2007) [HEC07.103] Difficile

$X+Y$ a même loi que X (X et Y indépendantes) si et seulement si Y est certaine.

On considère deux variables aléatoires X et Y, discrètes, définies sur l'espace probabilisé $(\Omega, \mathscr{A}, P)$. On considère une partie D (respectivement Δ) de $\mathbf{R}$, en bijection avec $\mathbf{N}$, dans laquelle X (resp. Y) prend presque sûrement ses valeurs, et on indexe bijectivement les éléments de D (resp. Δ) de sorte que $D = \{x_n, n \in \mathbf{N}\}$ (resp. $\Delta = \{y_n, n \in \mathbf{N}\}$). On suppose que X et Y sont indépendantes et que X et $X+Y$ ont même loi.

1. On considère une série à termes positifs $\sum_{n\geqslant 0} a_n$, qui est convergente.
   Justifier l'existence du nombre $M = \max_n(a_n)$. En déduire que l'ensemble $\{P(X=x), x \in \mathbf{R}\}$ admet un plus grand élément.
   Soit alors a un réel tel que $P(X=a) = \max\{P(X=x), x \in \mathbf{R}\}$.
2. (a) Montrer que, pour tout $y \in \mathbf{R}$, $P(X=a-y) = P(X=a)$ ou $P(Y=y) = 0$.
   (b) En déduire que la variable aléatoire Y est finie.
3. Soit μ un réel appartenant à l'ensemble $\{y \in \mathbf{R}, P(Y=y) \neq 0\}$.
   Montrer que, pour tout $n \in \mathbf{N}$, $P(X=a-n\mu) = P(X=a)$.
4. Montrer que la variable Y est presque sûrement nulle.

SOLUTION DE L'EXERCICE 3.16

1. Si $\sum a_n$ est convergente, alors nécessairement (a_n) tend vers 0. Et donc, comme toute suite convergente, elle est bornée.
Si la suite (a_n) est nulle, elle admet évidemment un plus grand élément : 0.
Si elle est non nulle, soit $i \in \mathbf{N}$ tel que $a_i > 0$. Alors, en prenant $\varepsilon = \frac{a_i}{2}$, la définition de la convergence d'une suite nous indique que[1]

$$\exists N \in \mathbf{N} : n \geqslant N \Rightarrow a_n \leqslant \frac{a_i}{2}.$$

Notons alors $M = \max(N, i)$. Soit alors $j \in [\![0, M-1]\!]$ tel que a_j soit maximal (une famille **finie** de réels admet toujours un plus grand élément).
Alors pour $n \in \mathbf{N}$, on a :

- Soit $n \leqslant M-1$, et alors $a_n \leqslant a_j$ par définition d'un maximum
- Soit $n \geqslant M$, et alors $a_n \leqslant \frac{a_i}{2} \leqslant a_i \leqslant a_j$.

Donc pour tout $n \in \mathbf{N}$, $a_n \leqslant a_j$: la famille (a_n) admet bien un plus grand élément.
En particulier, puisque la série $\sum_n P(X=x_n)$ converge[2], elle admet un plus grand élément, correspondant à $a \in D$. Et alors pour $x \in \mathbf{R}$, on a soit $x \notin D$, et alors $P(X=x) = 0$. Soit $x \in D$, et alors il existe $n \in \mathbf{N}$ tel que $x = x_n$ et donc $P(X=x_n) \leqslant P(X=a)$.

2.a. Soit $y \in \mathbf{R}$. Si $y \notin \Delta$, alors par définition, $P(Y=y) = 0$.
Sinon, on a, par indépendance de X et Y,

$$P(X+Y=a) = \sum_{n=0}^{+\infty} P(Y=y_n)P(X=a-y_n).$$

Et puisque X et $X+Y$ ont la même loi,

$$P(X=a) = P(X+Y=a) = \sum_{n=0}^{+\infty} P(Y=y_n)P(X=a-y_n).$$

De plus, on a $1 = \sum_{n=0}^{+\infty} P(Y=y_n)$, donc

$$P(X=a)\sum_{n=0}^{+\infty} P(Y=y_n) = \sum_{n=0}^{+\infty} P(Y=y_n)P(X=a-y_n)$$

Toutefois
Être bornée ne suffit pas à admettre un plus grand élément. Par exemple, la suite $u_n = -\frac{1}{n}$ est bornée (car convergente), mais n'admet pas de plus grand élément.

[1] Il y a normalement une valeur absolue dans la définition de la convergence, mais on sait que a_n est à termes positifs.

[2] C'est la formule des probabilités totales :

$$\sum_{n=0}^{+\infty} P(X=x_n) = 1.$$

$$\Leftrightarrow \sum_{n=0}^{+\infty} P(Y=y_n)P(X=a) = \sum_{n=0}^{+\infty} P(Y=y_n)P(X=a-y_n)$$
$$\Leftrightarrow \sum_{n=0}^{+\infty} P(Y=y_n)\left(P(X=a)-P(X=a-x_n)\right)=0$$

Mais par définition de a, pour tout $n\in\mathbf{N}$, on a

$$P(X=a-y_n)\leqslant P(X=a)\Leftrightarrow P(X=a)-P(X=a-y_n)\geqslant 0.$$

Donc nous avons affaire à une somme de nombres positifs, dont la somme est nulle. Ceci signifie que chacun des nombres $P(Y=y_n)\left(P(X=a)-P(X=a-y_n)\right)$ est nul.
Et donc pour $n\in\mathbf{N}$, on a $P(Y=y_n=0)$ ou $P(X=a)=P(X=a-y_n)$.
Autrement dit, pour tout $y\in\Delta$, $P(Y=y)=0$ ou $P(X=a)=P(X=a-y)$.

2.b. Supposons que Y ne soit pas à support fini. Alors il existe une infinité de $y\in\mathbf{R}$ tels que $P(Y=y)\neq 0$. Et donc une infinité de $y\in\mathbf{R}$ tels que $P(X=a-y)=P(X=a)$.
Or $P(X=a)\neq 0$, car il existe au moins un $x\in\mathbf{R}$ tel que $P(X=x)\neq 0$, faute de quoi la somme de la série $\sum P(X=x_n)$ ne serait pas égale à 1.
Par conséquent, il existe une infinité de $n\in\mathbf{N}$ tels que $P(X=a-y_n)=P(X=a)$. Ceci est incompatible avec la convergence de la série $\sum_n P(X=x_n)$.
Par conséquent, il n'existe qu'un nombre fini de $y\in\mathbf{R}$ tels que $P(Y=y)\neq 0$: Y est une variable aléatoire finie.

> **Rappel**
> a est tel que $P(X=a)$ soit maximum.

3. Soit μ tel que $P(Y=\mu)\neq 0$. D'après la question 2.a, on a nécessairement

$$P(X=a-\mu)=P(X=a).$$

Autrement dit, $P(X=a-\mu)$ est un maximum de l'ensemble $\{P(X=x),x\in\mathbf{R}\}$. On peut donc refaire le raisonnement de la question 2.a : pour tout $y\in\mathbf{R}$, $P(X=a-\mu-y)=P(X=a-\mu)$ ou $P(Y=y)=0$.
Et puisque $P(Y=\mu)\neq 0$, alors $P(X=a-\mu-\mu)=P(X=a)$. Donc $a-2\mu$ est tel que $P(X=a-2\mu)$ est le plus grand élément de l'ensemble $\{P(X=x),x\in\mathbf{R}\}$. Etc...
De proche en proche : $\forall n\in\mathbf{N}, P(X=a-n\mu)=P(X=a)$.

[3] C'est ici qu'on a besoin de $\mu\neq 0$!

4. Supposons au contraire que Y ne soit pas presque sûrement nulle, c'est-à-dire qu'il existe $\mu\neq 0$ tel que $P(Y=\mu)\neq 0$.
Alors d'après la question précédente, pour tout $n\in\mathbf{N}$, $P(X=a-n\mu)=P(X=a)$. Mais les $a-n\mu$ sont deux à deux distincts[3], et alors la série de terme général $P(X=a-n\mu)$ diverge.
Or, $\bigcup_{n\in\mathbf{N}}[X=a-n\mu]$ est une union dénombrable d'événements deux à deux disjoints, donc cette série devrait converger, et avoir une somme inférieure ou égale à 1.
D'où une contradiction, et donc notre hypothèse de départ est fausse : pour tout $\mu\neq 0$, $P(Y=\mu)=0$.
Ceci ne laisse plus beaucoup de choix, car $P(\Omega)=1$: on a nécessairement $P(Y=0)=1$, et donc Y est presque sûrement nulle.

□

> **Convergence**
> Cette série devrait converger, car cette hypothèse fait partie de la définition d'une probabilité : pour une famille dénombrable d'événements (A_n) deux à deux incompatibles, $\sum P(A_n)$ converge, et sa somme vaut
> $$P\left(\bigcup_{n\in\mathbf{N}} A_n\right).$$

Exercice 3.17 (QSP HEC 2014) [HEC14.QSP101] — Moyen

Étude d'une loi conditionnelle

Soit $n_1\in\mathbf{N}^*$, $n_2\in\mathbf{N}^*$ et $p\in]0,1[$. On considère deux variables aléatoires X_1 et X_2, définies sur un même espace probabilisé $(\Omega,\mathcal{A},P)$, telles que X_1 suit la binomiale $\mathcal{B}(n_1,p)$ et X_2 suit la loi binomiale $\mathcal{B}(n_2,p)$.

1. Soit $n\in(X_1+X_2)(\Omega)$. Déterminer la loi conditionnelle de X_1 sachant $[X_1+X_2=n]$.
2. Calculer l'espérance conditionnelle $E(X_1|X_1+X_2=n)$.

Solution de l'exercice 3.17

1. Notons que nous savons, par stabilité des lois binomiales[1], que X_1+X_2 suit la loi binomiale de paramètres n_1+n_2 et p.

[1] Qui s'applique puisque X_1 et X_2 sont indépendantes.

Et donc $(X_1+X_2)(\Omega)=[\![0,n_1+n_2]\!]$.
On a alors, pour tout $n\in[\![0,n_1+n_2]\!]$, et tout $k\in[\![0,n_1]\!]$,

$$\begin{aligned}P_{[X_1+X_2=n]}(X_1=k)&=\frac{P([X_1+X_2=n]\cap[X_1=k])}{P(X_1+X_2=n)}\\&=\frac{P([X_2=n-k]\cap[X_1=k])}{P(X_1+X_2=n)}\\&=\frac{P(X_1=k)P(X_2=n-k)}{P(X_1+X_2=n)}\end{aligned}$$

X_1 et X_2 sont indépendantes.

Donc déjà, si $n-k>n_2\Leftrightarrow k<n-n_2$, $P_{[X_1+X_2=n]}(X_1=k)=0$.
De même, si $n-k<0\Leftrightarrow k>n$, $P_{[X_1+X_2=n]}(X_1=k)=0$.
Et sinon,

$$\begin{aligned}P_{[X_1+X_2=n]}(X_1=k)&=\frac{\binom{n_1}{k}p^k(1-p)^{n_1-k}\binom{n_2}{n-k}p^{n-k}(1-p)^{n_2-n+k}}{\binom{n_1+n_2}{n}p^n(1-p)^{n_1+n_2-n}}\\&=\frac{\binom{n_1}{k}\binom{n_2}{n-k}}{\binom{n_1+n_2}{n}}.\end{aligned}$$

2. On a donc, par définition de l'espérance conditionnelle,

$$E(X_1|X_1+X_2=n)=\sum_{k=n-n_2}^{n}k\binom{n_2}{n-k}p^{n-k}(1-p)^{n_2-n+k}.$$

□

EXERCICE 3.18 (QSP HEC 2014) [HEC14.QSP75] — Facile

Étude d'une loi conjointe

Soient X et Y deux variables aléatoires définies sur $(\Omega,\mathcal{A},P)$, à valeurs dans $\mathbf{N}$, dont la loi conjointe est donnée par

$$\forall(i,j)\in\mathbf{N}^2,\ P([X=i]\cap[Y=j])=\frac{a}{(i+j+1)!}.$$

1. Déterminer la valeur de a.
2. Les variables X et Y sont-elles indépendantes ?

SOLUTION DE L'EXERCICE 3.18

1. Puisqu'il s'agit d'une loi conjointe, la famille $(P([X=i]\cap[Y=j]))_{(i,j)\in\mathbf{N}^2}$ est sommable, et sa somme vaut 1.
Pour $k\in\mathbf{N}$, posons $I_k=\{(i,j)\in\mathbf{N}^2 : i+j=k\}$.
Alors I_k est un ensemble fini de cardinal $k+1$, et donc

$$\sum_{(i,j)\in I_k}P([X=i]\cap[Y=j])=\sum_{(i,j)\in I_k}\frac{a}{(k+1)!}=\text{Card}(I_k)\frac{a}{(k+1)!}=(k+1)\frac{a}{(k+1)!}=\frac{a}{k!}.$$

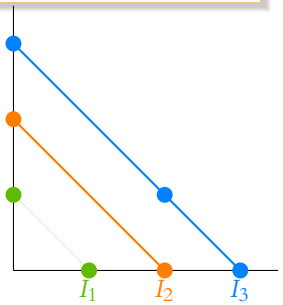


FIGURE 3.3– Les paquets I_k.

Et puisque $\mathbf{N}^2=\bigcup_{k=0}^{+\infty}I_k$, par le théorème de sommation par paquets,

$$\sum_{(i,j)\in\mathbf{N}^2}P([X=i]\cap[Y=j])=\sum_{k=0}^{+\infty}\sum_{(i,j)\in I_k}P([X=i]\cap[Y=j])=\sum_{k=0}^{+\infty}\frac{a}{k!}=ae.$$

Et donc $ae=1\Leftrightarrow a=e^{-1}$.

Convergence
Nous avons bien reconnu une série convergente, mais la convergence était de toutes façons garantie par le fait que la famille soit sommable.

2. Par la formule des probabilités totales appliquée au système complet d'événements $\{[X = i],\ i \in \mathbf{N}\}$, on a

$$P(Y=0) = \sum_{i=0}^{+\infty} P([X=i]\cap[Y=0]) = \sum_{i=0}^{+\infty} \frac{e^{-1}}{(i+1)!} = e^{-1}\sum_{i=1}^{+\infty}\frac{1}{i!} = e^{-1}\left(\sum_{i=0}^{+\infty}\frac{1}{i!}-1\right) = 1-e^{-1}.$$

Et de même, $P(X=0) = 1-e^{-1}$.
Or, on a $P([X=0]\cap[Y=0]) = e^{-1} \neq (1-e^{-1})^2 = 1-2e^{-1}+e^{-2}$.
En effet, si ces deux nombres étaient égaux, on aurait

$$(e^{-1})^2 - 3e^{-1} - 1 = 0.$$

et donc e^{-1} serait racine de X^2-3X-1. Mais ces racines sont $\dfrac{3+\sqrt{5}}{2}$ et $\dfrac{3-\sqrt{5}}{2}$.
On en déduit que X et Y ne sont pas indépendantes.

□

Exercice 3.19 (ESCP 2016) [ESCP16.3.11] Facile

Étude d'un couple de variables aléatoires

Toutes les variables aléatoires considérées dans cet exercice sont définies sur un espace probabilisé $(\Omega, \mathcal{A}, P)$.
Soient λ et p deux réels tels que $\lambda > 0$ et $p \in]0,1[$.
On considère le couple de variables aléatoires (X,Y) à valeurs dans $\mathbf{N}^2$, dont la loi conjointe est donnée par

$$\forall (k,n) \in \mathbf{N}^2,\ P([X=n]\cap[Y=k]) = \begin{cases} \dfrac{e^{-\lambda}\lambda^n p^k (1-p)^{n-k}}{k!(n-k)!} & \text{si } 0 \leqslant k \leqslant n \\ 0 & \text{sinon} \end{cases}$$

1. Vérifier que la relation ci-dessus définit bien une loi de probabilité sur $\mathbf{N}^2$.
2. Déterminer la loi marginale de la variable X, puis celle de la variable aléatoire Y.
   Les variables aléatoires X et Y sont-elles indépendantes ?
3. Déterminer la loi conditionnelle de la variable aléatoire Y sachant que $[X=n]$ est réalisé.
4. Soit Z la variable aléatoire définie par $Z = X - Y$. Déterminer la loi de la variable aléatoire Z.
5. Les variables aléatoires Y et Z sont-elles indépendantes ?

Solution de l'exercice 3.19

1. Pour n fixé, on a

$$\sum_{k=0}^{+\infty} P([X=n]\cap[Y=k]) = \sum_{k=0}^{n} \frac{e^{-\lambda}\lambda^n p^k (1-p)^{n-k}}{k!(n-k)!} = \frac{e^{-\lambda}\lambda^n}{n!}\sum_{k=0}^{n}\frac{n!}{k!(n-k)!}p^k(1-p)^{n-k}$$
$$= \frac{e^{-\lambda}\lambda^n}{n!}\sum_{k=0}^{n}\binom{n}{k}p^k(1-p)^{n-k} = \frac{e^{-\lambda}\lambda^n}{n!}.$$

Et alors, sous réserve de convergence,

$$\sum_{n=0}^{+\infty}\sum_{k=0}^{+\infty} P([X=n]\cap[Y=k]) = \sum_{n=0}^{+\infty} e^{-\lambda}\frac{\lambda^n}{n!} = e^{-\lambda}e^{\lambda} = 1.$$

On en déduit que la famille $P([X=n]\cap[Y=k])$ est bien sommable, de somme égale à 1.
Et donc définit une loi de probabilité sur $\mathbf{N}^2$.

2. D'après la formule des probabilités totales appliquée au système complet d'événements $\{[Y=k], k \in \mathbf{N}\}$, on a, pour tout $n \in \mathbf{N}$,

$$P(X=n) = \sum_{k=0}^{+\infty} P([X=n]\cap[Y=k]) = \frac{e^{-\lambda}\lambda^n}{n!}.$$

Cette somme a déjà été calculée dans la question précédente.

Et donc X suit la loi de Poisson de paramètre λ.
De même, la formule des probabilités totales appliquée au système complet d'événements $\{[X=n], n\in\mathbf{N}\}$ nous donne, pour tout $k\in\mathbf{N}$,

$$\begin{aligned}
P(Y=k) &= \sum_{n=0}^{+\infty} P([X=n]\cap[Y=k]) \\
&= \sum_{n=k}^{+\infty} \frac{e^{-\lambda}\lambda^n p^k (1-p)^{n-k}}{k!(n-k)!} \qquad \text{Les termes pour } n<k \text{ sont nuls.} \\
&= \frac{e^{-\lambda}p^k\lambda^k}{k!}\sum_{n=k}^{+\infty}\frac{\lambda^{n-k}(1-p)^{n-k}}{(n-k)!} \\
&= \frac{e^{-\lambda}p^k\lambda^k}{k!}\sum_{i=0}^{+\infty}\lambda^i\frac{(1-p)^i}{i!} \qquad \text{Chgt d'indice } i=n-k. \\
&= \frac{e^{-\lambda}p^k}{k!}e^{\lambda(1-p)} \\
&= \frac{e^{-\lambda p}(\lambda)^k}{k!}.
\end{aligned}$$

Et donc Y suit une loi de Poisson de paramètre λp.
Notons que X et Y ne sont pas indépendantes, car pour tout $k>n$, on a $P(X=n)\neq 0, P(Y=k)\neq 0$ et $P([X=n]\cap[Y=k])=0$.

3. On a, pour tout $k\in\mathbf{N}$,

$$P_{[X=n]}(Y=k) = \frac{P([X=n]\cap[Y=k])}{P(X=n)} = \begin{cases} \dfrac{n!}{k!(n-k)!}p^k(1-p)^{n-k} & \text{si } 0\leqslant k\leqslant n \\ 0 & \text{sinon} \end{cases}$$

On reconnaît alors une loi binomiale $\mathcal{B}(n,p)$, donc la loi de Y sachant $[X=n]$ est la loi $\mathcal{B}(n,p)$.

4. Notons qu'on a toujours[1] $X\geqslant Y$ car $P([X=n]\cap[Y=k])=0$ si $k>n$.
Et donc Z ne peut prendre que des valeurs entières positives ou nulles.
Pour $\ell\in\mathbf{N}$, on a

[1] Du moins presque sûrement.

$$\begin{aligned}
P(Z=\ell) &= \sum_{k=0}^{+\infty} P([Z=\ell]\cap[Y=k]) \\
&= \sum_{k=0}^{+\infty} P([X=\ell+k]\cap[Y=k]) \\
&= \sum_{k=0}^{+\infty} \frac{e^{-\lambda}\lambda^{k+\ell}p^k(1-p)^\ell}{k!\ell!} \\
&= \frac{e^{-\lambda}\lambda^\ell(1-p)^\ell}{\ell!}\sum_{k=0}^{+\infty}\frac{\lambda^k p^k}{k!} \\
&= \frac{e^{-\lambda}\lambda^\ell(1-p)^\ell}{\ell!}e^{\lambda p} \\
&= e^{-\lambda(1-p)}\frac{(\lambda(1-p))^k}{k!}.
\end{aligned}$$

Donc Z suit la loi de Poisson de paramètre $\lambda(1-p)$.

5. Pour $(k,\ell)\in\mathbf{N}^2$, on a

$$P([Y=k]\cap[Z=\ell]) = P([Y=k]\cap[X-Y=\ell]) = P([Y=k]\cap[X=k+\ell]) = \frac{e^{-\lambda}\lambda^{k+\ell}p^k(1-p)^\ell}{k!\ell!}.$$

D'autre part, on a

$$P(Y=k)P(Z=\ell) = e^{-\lambda p}\frac{(\lambda p)^k}{k!}e^{-\lambda(1-p)}\frac{(\lambda(1-p))^\ell}{\ell!} = \frac{e^{-\lambda}\lambda^{k+\ell}p^k(1-p)^\ell}{k!\ell!} = P([Y=k]\cap[Z=\ell]).$$

Et donc Y et Z sont indépendantes.

□

Exercice 3.20 (HEC 2010) [HEC10.2] Difficile

Étude d'un couple de variables aléatoires (somme et différence de deux géométriques).

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 $(\Omega, \mathcal{A}, P)$ dont la loi de probabilité est donnée par : pour tout entier $n \geqslant 2$, pour tout $t \in \mathbf{Z}$,

$$P([U=n]\cap[T=t]) = \begin{cases} p^2q^{n-2} & \text{si } |t| \leqslant n-2 \text{ et si } n, |t| \text{ de même parité} \\ 0 & \text{sinon} \end{cases}$$

1. Vérifier que $\displaystyle\sum_{n=2}^{+\infty} \sum_{\substack{|t|\leqslant n-2 \\ n,|t| \text{ de même parité}}} p^2q^{n-2} = 1$.
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 : $\forall t \in \mathbf{Z}$, $P(T=t) = \dfrac{pq^{|t|}}{1+q}$.
   (c) Calculer $E(T)$.
3. Soit n un entier supérieur ou égal à 2.
   (a) Déterminer la loi conditionnelle de T sachant $[U=n]$.
   (b) 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 $\mathrm{Cov}(U,T)$.
   (b) Les variables U et T sont-elles indépendantes ?

Solution de l'exercice 3.20

1. Si $n \geqslant 2$ est pair, $n = 2k$, alors les entiers pairs compris entre $-(n-2)$ et $n-2$ sont les $2i$, $-(k-1) \leqslant i \leqslant k-1$, de sorte que

$$\sum_{\substack{|t|\leqslant n-2 \\ |t| \text{pair}}} p^2q^{n-2} = \sum_{i=-(k-1)}^{k-1} p^2q^{n-2} = p^2q^{n-2}(2k-2+1) = (n-1)p^2q^{n-2}.$$

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 $2i+1$, $-k \leqslant i \leqslant k-1$, de sorte que

$$\sum_{\substack{|t|\leqslant n-2 \\ |t| \text{impair}}} p^2q^{n-2} = \sum_{i=-k}^{k-1} p^2q^{n-2} = p^2q^{n-2}(k-1+k+1) = (2k)p^2q^{n-2} = (n-1)p^2q^{n-2}.$$

Et donc

$$\sum_{n=2}^{+\infty} \sum_{\substack{|t|\leqslant n-2 \\ n,|t| \text{ de même parité}}} p^2q^{n-2} = \sum_{n=2}^{+\infty}(n-1)p^2q^{n-2} = p^2\sum_{m=1}^{+\infty} mq^{m-1} = p^2\frac{1}{(1-q)^2} = 1.$$

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

$$\begin{aligned} P([U=n]) &= \sum_{t\in\mathbf{Z}} P([U=n]\cap[T=t]) \\ &= \sum_{\substack{|t|\leqslant n-2 \\ |t|, n \text{ de même parité}}} p^2q^{n-2} \\ &= (n-1)p^2q^{n-2}. \end{aligned}$$

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 $\{[U=n], n \geqslant 2\}$.
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 $|t| \leqslant n-2 \Leftrightarrow n \geqslant t+2$. Et donc

$$P(T=t) = \sum_{\substack{n \geqslant t+2 \\ t,n \text{ de même parité}}}^{+\infty} p^2 q^{n-2}$$

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 $i \geqslant k+1$.

Et si t est impair, $t = 2k+1$, alors

$$P(T=t) = \sum_{i=k+1}^{+\infty} p^2 q^{2i+1-2} = p^2 q^{2k+1} \sum_{j=0}^{+\infty} (q^2)^j = p^2 q^{2k+1} \frac{1}{1-q^2} = \frac{pq^t}{1+q}.$$

Enfin, si t est négatif, alors

$$P(T=t) = \sum_{\substack{|t| \leqslant n-2 \\ |t|,n \text{ de même parité}}} p^2 q^{n-2} = P(T=|t|) = \frac{pq^{|t|}}{1+q}.$$

2.c. On a, sous réserve de convergence absolue,

$$E(T) = \sum_{t<0} tP(T=t) + 0 \cdot P(T=0) + \sum_{t>0} tP(T=t) = \sum_{t=1}^{+\infty} -t\frac{pq^t}{1+q} + \sum_{t=1}^{+\infty} t\frac{pq^{-t}}{1+q}.$$

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.

Les deux séries en jeu convergent absolument car $t\dfrac{pq^t}{1+q}$ est le terme général d'une série géométrique dérivée de raison $q \in]0,1[$, donc convergente.
Et alors on a

$$E(T) = \sum_{t=1}^{+\infty} (-t)\frac{pq^t}{1+q} + \sum_{t=1}^{+\infty} t\frac{pq^t}{1+t} = 0.$$

3.a. Notons que $P_{[U=n]}(T=t)$ est non nul si et seulement si $|t| \leqslant n-2$ et si n et t ont même parité.
Et donc, sachant que $[U=n]$, T prend ses valeurs dans $I_n = \{-(n-2), -(n-4), \ldots, n-4, n-2\}$.

Remarque
Si n est pair, alors $0 \in I_n$, et si n est impair, $0 \notin I_n$.

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 $I_n = \{-(n-2), -(n-4), \ldots, n-4, n-2\}$.
En particulier, la loi de T sachant $[U=n]$ est finie, donc $E(T|U=n)$ existe.

Danger !
Il ne s'agit pas d'une loi uniforme sur $[\![-(n-2), n-2]\!]$, car T ne prend que des valeurs de même parité que n.

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 $\{[U = n], n \geqslant 2\}$, on a

$$E(UT) = \sum_{n=2}^{+\infty} P(U = n)E(UT|U = n) = 0.$$

Et donc, par la formule de Huygens, $\mathrm{Cov}(U, T)$ existe et vaut

$$\mathrm{Cov}(U, T) = \underbrace{E(UT)}_{=0} - E(U)\underbrace{E(T)}_{=0} = 0.$$

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).

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.

□

Exercice 3.21 (QSP HEC 2009) [HEC09.QSP01] — Moyen

Espérance de $\dfrac{X}{X+Y}$.

Soient X et Y deux variables aléatoires indépendantes définies sur un même espace probabilisé $(\Omega, \mathcal{A}, P)$, de même loi, à valeurs dans $\mathbf{N}^*$. Montrer que $E\left(\dfrac{X}{X+Y}\right)$ existe et vaut $\dfrac{1}{2}$.

Solution de l'exercice 3.21 Montrons que $\dfrac{X}{X+Y}$ et $\dfrac{Y}{X+Y}$ ont même loi.
Il est clair que ces deux variables aléatoires ont le même support car X et Y ont même loi.
Soit donc r dans le support de $\dfrac{X}{X+Y}$ (qui est formé de nombres rationnels strictement positifs).
Alors, d'après la formule des probabilités totales, on a

Loi discrète
Notons que $\frac{X}{X+Y}$ est à valeurs dans $\mathbf{Q}$, qui est dénombrable, donc il s'agit d'une variable discrète.

$$P\left(\frac{X}{X+Y} = r\right) = \sum_{k=1}^{\infty} P\left([X = k] \cap \left[\frac{X}{X+Y} = r\right]\right) = \sum_{k=1}^{\infty} P\left([X = k] \cap \left[Y = \frac{k}{r} - k\right]\right) = \sum_{k=1}^{\infty} P(X = k)P\left(Y = \frac{k}{r} - k\right)$$

Le même calcul prouve que

$$P\left(\frac{Y}{X+Y} = r\right) = \sum_{k=1}^{\infty} P(Y = k)P\left(X = \frac{k}{r} - k\right)$$

Mais X et Y ayant même loi, on a donc $P\left(\dfrac{X}{X+Y} = r\right) = P\left(\dfrac{Y}{X+Y} = r\right)$.
Ainsi, les deux variables $\dfrac{X}{X+Y}$ et $\dfrac{Y}{X+Y}$ ont même loi.
Autre méthode : puisque X et Y sont indépendantes et de même loi, les vecteurs (X, Y) et (Y, X) sont de même loi[1]
La fonction $g : (x_1, x_2) \mapsto \dfrac{x_1}{x_1 + x_2}$ étant continue sur $(\mathbf{R}_+^*)^2$, les variables $g(X, Y) = \dfrac{X}{X+Y}$ et $g(Y, X) = \dfrac{Y}{X+Y}$ ont donc la même loi.

[1] I.e. possèdent la même fonction de répartition.

De plus, ces deux variables ont une espérance, car on a $0 \leqslant \dfrac{X}{X+Y} \leqslant 1$, et que $E(1)$ existe.
Enfin, on a

$$1 = E(1) = E\left(\frac{X}{X+Y} + \frac{Y}{X+Y}\right) = E\left(\frac{X}{X+Y}\right) + E\left(\frac{Y}{X+Y}\right) = 2E\left(\frac{X}{X+Y}\right).$$

On en déduit que $E\left(\dfrac{X}{X+Y}\right) = \dfrac{1}{2}$. □

Espérance
Si les variables ont même loi, et qu'elles admettent une espérance, elles ont nécessairement la même espérance.

EXERCICE 3.22 (QSP ESCP 2016) [ESCP16.QSP06] — Moyen

Une réciproque à la stabilité des lois binomiales

Soit n un entier supérieur ou égal à 1 et $p \in]0,1[$. Soient X et Y deux variables aléatoires indépendantes définies sur un même espace probabilisé et telles que X suit la loi binomiale $\mathcal{B}(n,p)$ et Y est à valeurs dans $[\![0,n]\!]$.
Que pensez vous de l'équivalence suivante : Y suit la loi $\mathcal{B}(n,p)$ si et seulement si $X+Y$ suit la loi $\mathcal{B}(2n,p)$?

SOLUTION DE L'EXERCICE 3.22 Notons que nous savons déjà que, X et Y étant indépendantes, si Y suit la loi binomiale $\mathcal{B}(n,p)$, alors $X+Y$ suit la loi $\mathcal{B}(2n,p)$, mais le cours ne dit absolument rien au sujet de la réciproque.

Inversement, supposons que $X+Y$ suit la loi $\mathcal{B}(2n,p)$.
Puisque X et Y sont indépendantes, on a alors, par produit de convolution, pour tout $k \in [\![0,2n]\!]$,

$$\sum_{i=0}^{k} P(X=i)P(Y=k-i) = P(X+Y=k) = \binom{2n}{k} p^k (1-p)^{2n-k}.$$

Remarque
On a ainsi un système de $2n+1$ équations vérifiées par les $P(X=k)$.

En particulier, pour $k=0$, on a $P(X=0)P(Y=0) = P(X+Y=0)$, soit encore

$$P(X=0) \times (1-p)^n = (1-p)^{2n} \Leftrightarrow P(X=0) = (1-p)^n.$$

Remarque
Pour $P(X=0)$ et $P(X=1)$, on retrouve les probabilités d'une loi binomiale, donc il semble que X suive la loi binomiale $\mathcal{B}(n,p)$.

Puis, pour $k=1$, on a $P(X=0)P(Y=1) + P(X=1)P(Y=0) = P(X+Y=1)$ soit

$$P(X=1)(1-p)^n + P(X=0)np(1-p)^{n-1} = np(1-p)^{2n-1} \Longleftrightarrow P(X=1) = np(1-p)^{n-1}.$$

Plus généralement, pour $k \in [\![0,n]\!]$, on a

$$P(Y=0)P(X=k) + P(Y=1)P(X=k-1) + \cdots + P(Y=k)P(X=0) = P(X+Y=k).$$

Et donc $P(X=k) = \dfrac{P(X+Y=k) - \sum_{i=0}^{k-1} P(X=i)P(Y=k-i)}{P(Y=0)}$ $(\star)$.

Soit Z une variable aléatoire indépendante de Y, suivant la loi binomiale $\mathcal{B}(n,p)$. Alors $Y+Z$ suit la loi binomiale $\mathcal{B}(2n,p)$, et donc pour tout $k \in [\![0,2n]\!]$, $P(Y+Z=k) = P(X+Y=k)$.

Réc. forte
Une récurrence forte est une récurrence où ne suppose pas seulement que la propriété est vraie au rang n pour la prouver au rang $n+1$, mais qu'elle est vraie aux rangs $0, 1, \ldots, n$ pour la prouver au rang $n+1$.

Prouvons, par récurrence forte sur $k \in [\![0,n]\!]$ que $P(X=k) = P(Z=k)$.
Nous venons de prouver que c'est bien le cas pour $k=0$ et $k=1$.
Soit donc $k \in [\![0,n-1]\!]$, et supposons que pour tout $i \in [\![0,k]\!]$, $P(X=i) = P(Z=i)$.
Alors, on a

$$P(X+Y=k+1) = \frac{P(X+Y=k+1) - \sum_{i=0}^{k} P(X=i)P(Y=k+1-i)}{P(Y=0)}$$

$$= \frac{P(Y+Z=k+1) - \sum_{i=0}^{k} P(Z=i)P(Y=k+1-i)}{P(Y=0)}$$

$$= \frac{\sum_{i=0}^{k+1} P(Z=i)P(Y=k+1-i) - \sum_{i=0}^{k} P(Z=i)P(Y=k+1-i)}{P(Y=0)}$$
$$= \frac{P(Z=k+1)P(Y=0)}{P(Y=0)}$$
$$= P(Z=k+1).$$

C'est le produit de convolution appliqué à Y et Z (qui sont indépendantes par hypothèse).

Par le principe de récurrence, pour tout $k \in [\![0, 2n]\!]$, $P(X=k) = P(Z=k)$, et donc X suit la loi binomiale $\mathscr{B}(n,p)$.

En conclusion, $X+Y \hookrightarrow \mathscr{B}(2n,p) \Leftrightarrow X \hookrightarrow \mathscr{B}(n,p)$.

Quelques explications : l'idée principale ici est de voir que la relation $(\star)$ détermine uniquement $P(X=0)$.
Puis, en utilisant la valeur de $P(X=0)$ ainsi obtenue, il n'y a donc qu'une valeur de $P(X=1)$ possible. Puis qu'une valeur de $P(X=2)$, etc.
D'autre part, nous savons que si $X \hookrightarrow \mathscr{B}(n,p)$, alors les $P(X=k)$ vérifient les $n+1$ relations $(\star)$: c'est donc l'unique solution du système d'équation mentionné plus tôt.
L'introduction de la variable Z n'a rien d'obligatoire, elle permet surtout d'écrire $P(Z=i)$ au lieu de $\binom{n}{i}p^i(1-p)^{n-i}$, et sert donc surtout à alléger les notations. □

Exercice 3.23 (ESCP 2018) [ESCP18.3.16] — Moyen

Étude des «sommets» d'une succession de tirages

1. Soit $\alpha > 0$. Donner un équivalent de $S_n = \sum_{j=1}^{n} j^\alpha$ lorsque $n \to +\infty$.
2. Soit $n \geqslant 2$, et soit $(X_j)_{j\geqslant 1}$ une suite de variables aléatoires définies sur un espace probabilisé $(\Omega, \mathscr{A}, P)$, mutuellement indépendantes, et suivant toutes la loi uniforme sur $[\![1,n]\!]$.
   Pour $k \geqslant 2$, on dit que X_k est un sommet si $X_{k-1} < X_k$ et $X_{k+1} < X_k$.
   Déterminer les probabilités des événements suivants :
   (a) «X_k est un sommet». Donner $\lim_{n\to+\infty} P(\text{«}X_k \text{ est un sommet »})$.
   (b) «X_k et X_{k+1} sont des sommets»
   (c) «X_k et X_{k+3} sont des sommets». Donner $\lim_{n\to+\infty} P(\text{«}X_k \text{ et } X_{k+3} \text{ sont des sommets»})$.
   (d) «X_2 et X_4 sont des sommets». Donner $\lim_{n\to+\infty} P(\text{«}X_2 \text{ et } X_4 \text{ sont des sommets»})$.

Solution de l'exercice 3.23

1. Le moyen le plus classique de trouver un tel équivalent est probablement de faire apparaître une intégrale : la fonction $t \mapsto t^\alpha$ est croissante sur $\mathbf{R}_+$ et donc pour tout $k \in \mathbf{N}$, et pour tout $t \in [k, k+1]$,

$$k^\alpha \leqslant t^\alpha \leqslant (k+1)^\alpha \Rightarrow k^\alpha \leqslant \int_k^{k+1} t^\alpha\,dt \leqslant (k+1)^\alpha.$$

Et donc on en déduit que

$$\int_{k-1}^{k} t^\alpha\,dt \leqslant k^\alpha \leqslant \int_k^{k+1} t^\alpha\,dt.$$

En sommant pour k allant de 1 à n, il vient donc, à l'aide de la relation de Chasles,

$$\int_0^n t^\alpha\,dt \leqslant S_n \leqslant \int_1^{n+1} t^\alpha\,dt.$$

Or, il est facile de calculer ces deux intégrales :

$$\int_0^n t^\alpha\,dt = \frac{n^{\alpha+1}}{\alpha+1} \text{ et } \int_1^{n+1} t^\alpha\,dt = \frac{1}{\alpha+1}\left((n+1)^{\alpha+1}-1\right).$$

Ces deux intégrales sont équivalentes, lorsque $n \to +\infty$, à $\dfrac{n^{\alpha+1}}{\alpha+1}$, et donc une application rapide du théorème des gendarmes donne $\displaystyle\sum_{k=1}^n k^\alpha \underset{n\to+\infty}{\sim} \frac{n^{\alpha+1}}{\alpha+1}$.

Notons qu'il existe une solution bien plus rapide si on essaie de faire apparaître des sommes de Riemann :

$$\sum_{k=1}^n k^\alpha = n^\alpha \sum_{k=1}^n \left(\frac{k}{n}\right)^\alpha = n^{\alpha+1}\frac{1}{n}\sum_{k=1}^n \left(\frac{k}{n}\right)^\alpha.$$

Mais $\dfrac{1}{n}\displaystyle\sum_{k=1}^n \left(\frac{k}{n}\right)^\alpha \underset{n\to+\infty}{\longrightarrow} \int_0^1 t^\alpha\,dt = \frac{1}{\alpha+1}$.

Et donc $\displaystyle\sum_{k=1}^n k^\alpha \underset{n\to+\infty}{\sim} \frac{n^{\alpha+1}}{\alpha+1}$.

2.a. Notons que X_k est un sommet si X_k est à la fois plus grand que X_{k-1} et que X_{k+1}.
Et donc si et seulement si $X_k > \max(X_{k-1}, X_{k+1})$.
Mais, pour $j \in [\![1,n]\!]$, on a, par indépendance des X_i,

$$P(\max(X_{k-1},X_{k+1}) \leqslant j) = P\left([X_{k-1}\leqslant j]\cap[X_{k+1}\leqslant j]\right) = P\left(X_{k-1}\leqslant j\right)P\left(X_{k+1}\leqslant j\right) = \frac{j^2}{n^2}.$$

Et donc, par la formule des probabilités totales appliquée au système complet d'événements $\{[X_k \leqslant i], i \in [\![1,n]\!]\}$,

$$\begin{aligned}
P(\max(X_{k-1},X_{k+1}) < X_k) &= \sum_{i=1}^n P\left([\max(X_{k-1},X_{k+1}) < X_k]\cap[X_k=i]\right) = \sum_{i=1}^n P\left([\max(X_{k-1},X_{k+1}) < i]\cap[X_k=i]\right)\\
&= \sum_{i=1}^n P\left([\max(X_{k-1},X_{k+1}) < i]\cap[X_k=i]\right)P(X_k=i)\\
&= \sum_{i=1}^n \frac{(i-1)^2}{n^2}\frac{1}{n} = \frac{1}{n^3}\sum_{i=1}^{n-1} i^2 = \frac{1}{n^3}\frac{(n-1)n(2n-1)}{6}\\
&= \frac{(n-1)(2n-1)}{6n^2}.
\end{aligned}$$

Et par conséquent, la limite de cette probabilité lorsque $n \to +\infty$ vaut $\dfrac{1}{3}$.

2.b. Il est évident que X_k et X_{k+1} ne peuvent être en même temps des sommets puisqu'on aurait alors $X_{k+1} < X_k$ et $X_k < X_{k+1}$...

2.c. L'événement «X_k est un sommet» ne dépend[1] que des variables aléatoires X_k, X_{k+1} et X_{k-1}.
De même l'événement «X_{k+3} est un sommet» ne dépend que des variables aléatoires X_{k+2}, X_{k+3} et X_{k+4}.
Par le lemme des coalitions, ces deux événements sont donc indépendants.
Et par conséquent,

[1] C'est ce que nous avons fait à la question 2.a.

$$\begin{aligned}
P(\text{«}X_k \text{ et } X_{k+3} \text{ sont des sommets»}) &= P(\text{«}X_k \text{ est un sommet»})P(\text{«}X_{k+3} \text{ est un sommet»})\\
&= \frac{(n-1)^2(2n-1)^2}{36n^4} \underset{n\to+\infty}{\longrightarrow} \frac{1}{9}.
\end{aligned}$$

2.d. Le raisonnement de la question précédente ne s'applique plus, puisque cette fois, les événements «X_2 est un sommet» et «X_4 est un sommet» dépendent tous deux de la variable X_3, et donc le lemme des coalitions ne s'applique plus.

On a donc «X_2 et X_4 sont des sommets» si et seulement si on a à la fois

$$X_2 > X_1,\ X_2 > X_3,\ X_4 > X_3 \text{ et } X_4 > X_5.$$

Soit si et seulement si $X_3 < \min(X_2, X_4)$, $X_4 > X_5$ et $X_2 > X_1$.
Utilisons donc une formule des probabilités totales, avec le système complet d'événements $\{[X_2 = i] \cap [X_4 = j],\ (i,j) \in [\![1,n]\!]^2\}$, en notant A l'événement «X_2 et X_4 sont des sommets» :

$$\begin{aligned}
P(A) &= \sum_{i=1}^{n}\sum_{j=1}^{n} P([X_1 < i] \cap [X_5 < j] \cap [X_3 < \min(i,j)] \cap [X_2 = i] \cap [X_4 = j]) \\
&= \sum_{i=1}^{n}\sum_{j=1}^{n} P(X_1 < i)P(X_5 < j)P(X_3 < \min(i,j))P(X_2 = i)P(X_4 = j) \\
&= \sum_{i=1}^{n}\sum_{j=1}^{n} \frac{(i-1)(j-1)}{n^4} P(X_3 < \min(i,j)) \\
&= \frac{1}{n^5}\sum_{i=2}^{n}\sum_{j=2}^{n} (i-1)(j-1)P(X_3 \leqslant \min(i-1, j-1)) \\
&= \frac{1}{n^5}\sum_{i=1}^{n-1}\sum_{j=1}^{n-1} ij\min(i,j). \qquad \text{Changement d'indice.}
\end{aligned}$$

Pour calculer cette dernière somme, il s'agit de séparer la seconde somme en deux, suivant que $i > j$ ou non :

$$\begin{aligned}
\sum_{i=1}^{n-1}\sum_{j=1}^{n-1} ij\min(i,j) &= \sum_{i=1}^{n-1} i\left(\sum_{j=1}^{i} j\min(i,j) + \sum_{j=i+1}^{n-1} j\min(i,j)\right) \\
&= \sum_{i=1}^{n-1} i\left(\sum_{j=1}^{i} j^2 + \sum_{j=i+1}^{n-1} ij\right) \\
&= \sum_{i=1}^{n-1} i\left(\frac{i(i+1)(2i+1)}{6} + i\left(\frac{n(n-1)}{2} - \frac{i(i+1)}{2}\right)\right) \\
&= \frac{1}{6}\sum_{i=1}^{n-1} i^2(i+1)(2i+1) + \frac{n(n-1)}{2}\sum_{i=1}^{n-1} i^2 - \sum_{i=1}^{n}\frac{i^3(i+1)}{2} \\
&= \frac{1}{6}\sum_{i=1}^{n-1} i^2(i+1)(2i+1) + \frac{n^2(n-1)^2(2n-1)}{12} - \sum_{i=1}^{n}\frac{i^3(i+1)}{2} \\
&= \frac{1}{3}\sum_{i=1}^{n-1} i^4 + \frac{1}{2}\sum_{i=1}^{n-1} i^3 + \frac{1}{6}\sum_{i=1}^{n-1} i^2 - \frac{1}{2}\sum_{i=1}^{n-1} i^3 - \frac{1}{2}\sum_{i=1}^{n-1} i^4 + \frac{n^2(n-1)^2(2n-1)}{12} \\
&= -\frac{1}{6}\sum_{i=1}^{n-1} i^4 + \frac{n(n-1)(2n-1)}{36} + \frac{n^2(n-1)^2(2n-1)}{12}.
\end{aligned}$$

Il est alors aisé d'obtenir des équivalents, même s'il va nous falloir prendre quelques précautions pour pouvoir les ajouter :

$$\frac{n^2(n-1)^2(2n-1)}{12} \underset{n\to+\infty}{\sim} \frac{n^5}{6} = \frac{n^5}{6} + \underset{n\to+\infty}{o}\left(n^5\right).$$

De même, $\dfrac{n^2(n-1)^2(2n-1)}{12} = \underset{n\to+\infty}{o}\left(n^5\right)$ et grâce à la question 1 :

$$-\frac{1}{6}\sum_{i=1}^{n-1} i^4 \underset{n\to+\infty}{\sim} -\frac{1}{6}\frac{n^5}{30} = -\frac{1}{6}\frac{n^5}{30} + \underset{n\to+\infty}{o}(n^5).$$

Et donc enfin :

$$n^5 P(A) = \frac{n^5}{6} - \frac{n^5}{30} + \underset{n\to+\infty}{o}(n^5) = \frac{2}{15}n^5 + \underset{n\to+\infty}{o}(n^5) \underset{n\to+\infty}{\sim} \frac{2}{15}n^5.$$

Et donc $P(A) \underset{n\to+\infty}{\longrightarrow} \frac{2}{15}$.

□

3.4.1 Covariance, corrélation linéaire

EXERCICE 3.24 (QSP ESCP 2014) [ESCP14.QSP02] **Facile**

Étude de l'indépendance de la somme et du produit de deux variables aléatoires.

Soient X et Y deux variables aléatoires discrètes sur $(\Omega, \mathcal{A}, P)$, 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 ?

SOLUTION DE L'EXERCICE 3.24

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 indépendantes[1], il vient

[1] Car X et Y le sont.

$$E((XY)^2) = E(X^2)E(Y^2) = (V+M^2)^2.$$

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 Huygens[2], on a

[2] Celle pour la covariance.

$$\begin{aligned}\mathrm{Cov}(X+Y, XY) &= E(X + Y(XY)) - E(X+Y)E(XY)\\ &= E(X^2Y) + E(XY^2) - (E(X)+E(Y))E(XY)\\ &= E(X^2)E(Y) + E(X)E(Y^2) - E(X)E(X)E(Y) - E(Y)E(X)E(Y)\\ &= (V+M^2)M + M(V+M^2) - M^3 - M^3 = 2MV.\end{aligned}$$

X^2 et Y sont indépendantes, de même que X et Y^2.

Et donc $\mathrm{Cov}(X+Y, XY) \neq 0$, de sorte que XY et $X+Y$ ne sont pas indépendantes.

□

EXERCICE 3.25 (QSP HEC 2014) [HEC14.QSP113] **Facile**

Soit $\mathcal{E}$ un ensemble de variables aléatoires discrètes centrées définies sur un même espace probabilisé, et admettant une variance.

1. Justifier l'existence de $V_0 = \inf\{V(X), X \in \mathcal{E}\}$.
2. On suppose que pour tout $X_1, X_2 \in \mathcal{E}$, $\frac{1}{2}(X_1+X_2) \in \mathcal{E}$.
   Soient $(X_1, X_2) \in \mathcal{E}^2$ telles que $V(X_1) = V(X_2) = V_0$. Montrer que $X_1 = X_2$ presque sûrement.

SOLUTION DE L'EXERCICE 3.25

1. Il s'agit d'une partie minorée (par 0 car une variance est toujours positive) de $\mathbf{R}$, donc elle admet une borne inférieure.

2. Notons que dans cette question, on suppose que la borne inférieure de la question précédente est en fait un minimum, atteint pour X_1 et pour X_2.
Commençons par traiter le cas où $V_0 = 0$. Puisque X_1 et X_2 sont de variance nulle, elle sont presque sûrement constantes, et étant centrées, elles sont nulles presque sûrement, et donc presque sûrement égales.

Supposons à présent que $V_0 \neq 0$ et que $V(X_1) = V(X_2) = V_0$.
Alors $\frac{1}{2}(X_1+X_2) \in \mathcal{E}$ et donc $V\left(\frac{1}{2}(X_1+X_2)\right) \geqslant V_0$. Mais

$$V\left(\frac{1}{2}(X_1+X_2)\right) = \frac{1}{4}V(X_1+X_2) = \frac{1}{4}(V_0 + V_0 + 2\mathrm{Cov}(X_1, X_2)).$$

On en déduit que $\text{Cov}(X_1, X_2) \geqslant V_0$.
Et ainsi $\rho(X_1, X_2) = \frac{\text{Cov}(X_1, X_2)}{\sqrt{V_0}\sqrt{V_0}} = \frac{V_0}{V_0} \geqslant 1$.
Puisqu'on sait que l'inégalité inverse est vérifiée, alors on a une égalité : $\rho(X_1, X_2) = 1$.
Alors il existe deux constantes a et b telles que $X_1 = aX_2 + b$ presque sûrement, et $a > 0$ (car le coefficient de corrélation linéaire vaut 1 et non -1).
Mais X_1 et X_2 étant centrées, $0 = E(X_1) = aE(X_2) + b = b$.
Et $V(X_1) = a^2 V(X_2)$, de sorte que $a^2 = 1$. On en déduit que $a = 1$, et donc que $X_1 = X_2$ presque sûrement.

Et même
Du coup on a
$\text{Cov}(X_1, X_2) = V_0$.

□

Si X et Y sont deux variables aléatoires indépendantes, alors pour tous $(h, k) \in (\mathbf{N}^*)^2$, X^h et Y^k sont indépendantes, de sorte que $\text{Cov}\left(X^h, Y^k\right) = 0$.
L'exercice suivant, aux notations un peu lourdes, s'intéresse à une réciproque dans le cas de variables à support fini.

Exercice 3.26 (HEC 2014) [HEC14.91] Moyen

Une condition nécessaire et suffisante d'indépendance

On confond vecteur de $\mathbf{R}^n$ et matrice colonne de $\mathcal{M}_{n,1}(\mathbf{R})$. Soit $x_1, x_2, \ldots, x_n$ n réels tous distincts.

1. On considère la matrice $A = \begin{pmatrix} 1 & x_1 & x_1^2 & \ldots & x_1^{n-1} \\ 1 & x_2 & x_2^2 & \ldots & x_2^{n-1} \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 1 & x_{n-1} & x_{n-1}^2 & \ddots & x_{n-1}^{n-1} \\ 1 & x_n & x_n^2 & \ldots & x_n^{n-1} \end{pmatrix}$.

   Soit $U = \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix}$ un vecteur de $\mathbf{R}^n$ tel que $AU = 0$.

   (a) Montrer que le polynôme $Q(T) = \sum_{j=1}^{n} \alpha_j T^{j-1}$ est nul.

   (b) En déduire que la matrice A est inversible.

2. Dans cette question, X et Y sont deux variables aléatoires définies sur un espace probabilisé $(\Omega, \mathcal{A}, P)$, suivant des lois de Bernoulli de paramètres respectifs p et p' ($0 < p < 1$ et $0 < p' < 1$) et telles que la covariance de X et Y est nulle.
   Montrer que X et Y sont indépendantes.

3. Soit (X, Y) un couple de variables aléatoires réelles discrètes finies définies sur $(\Omega, \mathcal{A}, P)$, et soit n et m deux entiers supérieurs ou égaux à 2.
   On suppose que $X(\Omega) = \{x_1, x_2, \ldots, x_n\}$ et $Y(\Omega) = \{y_1, y_2, \ldots, y_m\}$.
   On pose, pour tout $i \in [\![1, n]\!]$ et tout $j \in [\![1, m]\!]$:

   $$p_i = P(X = x_i),\ q_j = P(Y = y_j),\ \pi_{i,j} = P([X = x_i] \cap [Y = y_j]) \text{ et } \delta_{i,j} = \pi_{i,j} - p_i q_j.$$

   On suppose que pour tout $h \in [\![1, n-1]\!]$ et tout $k \in [\![1, m-1]\!]$, la covariance de X^h et Y^k est nulle.

   (a) Soit $k \in [\![1, m-1]\!]$. Montrer que pour tout $i \in [\![1, n]\!]$, on a : $\sum_{j=1}^{m} \delta_{i,j} y_j^k = 0$.

   (b) En déduire que X et Y sont indépendantes.

Solution de l'exercice 3.26

1.a. Notons que $Q(x_i) = \sum_{j=1}^{n} \alpha_j x_i^{j-1}$. Puisque $AU = 0$, on a

$$0 = AU = \begin{pmatrix} 1 & x_1 & x_1^2 & \dots & x_1^{n-1} \\ 1 & x_2 & x_2^2 & \dots & x_2^{n-1} \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 1 & x_{n-1} & x_{n-1}^2 & \ddots & x_{n-1}^{n-1} \\ 1 & x_n & x_n^2 & \dots & x_n^{n-1} \end{pmatrix} \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_{n-1} \\ \alpha_n \end{pmatrix} = \begin{pmatrix} \alpha_1 + \alpha_2 x_1 + \cdots + \alpha_n x_1^{n-1} \\ \alpha_1 + \alpha_2 x_2 + \cdots + \alpha_n x_2^{n-1} \\ \vdots \\ \alpha_1 + \alpha_2 x_{n-1} + \cdots + \alpha_n x_{n-1}^{n-1} \\ \alpha_1 + \alpha_2 x_n + \cdots + \alpha_n x_n^{n-1} \end{pmatrix} = \begin{pmatrix} Q(x_1) \\ Q(x_2) \\ \vdots \\ Q(x_{n-1}) \\ Q(x_n) \end{pmatrix}.$$

Et donc, pour tout $i \in [\![1, n]\!]$, $Q(x_i) = 0$.
Les x_i étant deux à deux distincts par hypothèses, Q est donc un polynôme de degré inférieur ou égal à $n-1$ qui possède au moins n racines : c'est nécessairement le polynôme nul.

> **Rappel**
> Une matrice $A \in \mathcal{M}_n(\mathbf{R})$ est inversible si et seulement si le seul vecteur $X \in \mathcal{M}_{n,1}(\mathbf{R})$ tel que $AX = 0$ est le vecteur nul.

1.b. Puisque Q est nul, tous ses coefficients sont nuls, et donc $\alpha_1 = \alpha_2 = \cdots = \alpha_n = 0$.
Ainsi, on a $AU = 0 \Rightarrow U = 0$, ce qui prouve que la matrice A est inversible.

2. Par la formule de Huygens, nous savons que

$$\mathrm{Cov}(X, Y) = E(XY) - E(X)E(Y) = 0 \Leftrightarrow E(XY) = E(X)E(Y).$$

Mais $E(X) = p, E(Y) = p'$, et par le théorème de transfert,

$$E(XY) = \sum_{k=0}^{1}\sum_{\ell=0}^{1} k\ell P([X = k] \cap [Y = \ell]) = P([X = 1] \cap [Y = 1]).$$

> **Détails**
> Si $k = 0$ ou $\ell = 0$, alors $k\ell = 0$, et donc le seul terme non nul de la somme est celui pour $k = \ell = 1$.

Et donc $P([X = 1] \cap [Y = 1]) = pp' = P(X = 1)P(Y = 1)$.
Or, deux variables de Bernoulli sont indépendantes si et seulement si les événements $[X = 1]$ et $[Y = 1]$ sont indépendants, donc X et Y sont indépendants.

> **Plus généralement**
> Si X est une variable aléatoire certaine, alors pour toute variable aléatoire Y,
> $$\mathrm{Cov}(X, Y) = 0.$$

3.a. Notons que pour $h = 0$, et pour tout $k \in [\![1, m-1]\!]$, on a encore $\mathrm{Cov}(X^h, Y^k) = 0$ puisque

$$\mathrm{Cov}\left(X^h, Y^k\right) = \mathrm{Cov}\left(1, Y^k\right) = E\left(Y^k\right) - E(1)E\left(Y^k\right) = 0.$$

Comme précédemment, grâce à la formule de Huygens, il vient

$$0 = \mathrm{Cov}\left(X^h, Y^k\right) = E\left(X^h Y^k\right) - E\left(X^h\right)E\left(Y^k\right).$$

Mais, par le théorème de transfert, $E\left(X^h Y^k\right) = \sum_{i=1}^{n}\sum_{j=1}^{m} x_i^h y_j^k \pi_{i,j}$.

Et d'autre part, toujours par le théorème de transfert, $E\left(X^h\right) = \sum_{i=1}^{n} x_i^h p_i$ et $E\left(Y^k\right) = \sum_{j=1}^{m} y_j^k q_j$.

Et donc, pour tout $h \in [\![0, n-1]\!]$ et tout $k \in [\![1, m-1]\!]$, on a

$$0 = \sum_{i=1}^{n}\sum_{j=1}^{m} x_i^h y_j^k \pi_{i,j} - \left(\sum_{i=1}^{n} x_i^h p_i\right)\left(\sum_{j=1}^{m} y_j^k q_j\right) = \sum_{i=1}^{n}\sum_{j=1}^{m} x_i^h y_j^k \pi_{i,j} - \sum_{i=1}^{n}\sum_{j=1}^{m} x_i^h y_j^k p_i q_j = \sum_{i=1}^{n} x_i^h \sum_{j=1}^{m} y_j^k \delta_{i,j}.$$

Fixons à présent $k \in [\![1, m-1]\!]$, et pour $i \in [\![1, n]\!]$, notons $\alpha_i = \sum_{j=1}^{m} y_j^k \delta_{i,j}$, et $U = \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix}$.

Alors, si A désigne la matrice de la première question, il vient[1]

$$\begin{cases} \alpha_1 + \alpha_2 + \cdots + \alpha_n & = 0 \\ x_1\alpha_1 + x_2\alpha_2 + \cdots + x_n\alpha_n & = 0 \\ x_1^2\alpha_1 + x_2^2\alpha_2 + \cdots + x_n^2\alpha_n & = 0 \\ & \vdots \\ x_1^{n-1}\alpha_1 + x_2^{n-1}\alpha_2 + \cdots + x_n^{n-1}\alpha_n & = 0 \end{cases} \Longleftrightarrow \begin{pmatrix} 1 & 1 & \dots & 1 \\ x_1 & x_2 & \dots & x_n \\ \vdots & \vdots & & \vdots \\ x_1^{n-1} & x_2^{n-1} & \dots & x_n^{n-1} \end{pmatrix} \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{pmatrix} = 0 \Leftrightarrow {}^tAU = 0.$$

[1] En prenant successivement $h = 0, h = 1, \dots, h = n-1$ dans la relation précédemment établie.

Mais A étant inversible[2], ${}^tA$ est également inversible, et donc ${}^tAU = 0 \Rightarrow U = 0$.
On en déduit que $\alpha_1 = \alpha_2 = \cdots = \alpha_n = 0$, et donc

$$\forall i \in [\![1, n]\!], \ \sum_{j=1}^{m} y_j^k \delta_{i,j} = 0.$$

[2] Les x_i sont deux à deux distincts.

3.b. Soit $i \in [\![1,n]\!]$ fixé.
Remarquons que le résultat de la question précédente est encore valable pour $k=0$ puisque

$$\sum_{j=1}^{m} \delta_{i,j} = \sum_{j=1}^{m} \pi_{i,j} - p_i \underbrace{\sum_{j=1}^{m} q_j}_{=1} = \sum_{j=1}^{m} P([X=x_i]\cap[Y=y_j]) - P(X=x_i) = P(X=x_i) - P(X=x_i) = 0.$$

Détails
Puisque $Y(\Omega)=\{y_1,\dots,y_m\}$, la formule des probabilités totales nous donne
$$\sum_{j=1}^{m} \pi_{i,j} = p_i.$$

On a donc[3]

$$\begin{cases} \delta_{i,1} + \delta_{i,2} + \cdots + \delta_{i,m} = 0 \\ y_1\delta_{i,1} + y_2\delta_{i,2} + \cdots + y_m\delta_{i,m} = 0 \\ \vdots \\ y_1^{m-1}\delta_{i,1} + y_2^{m-1}\delta_{i,2} + \cdots + y_m^{m-1}\delta_{i,m} = 0 \end{cases} \iff \begin{pmatrix} 1 & 1 & \dots & 1 \\ y_1 & y_2 & \dots & y_m \\ \vdots & \vdots & & \vdots \\ y_1^{m-1} & y_2^{m-1} & \dots & y_m^{m-1} \end{pmatrix} \begin{pmatrix} \delta_{i,1} \\ \delta_{i,2} \\ \vdots \\ \delta_{i,m} \end{pmatrix} = 0.$$

[3] En prenant successivement $k=0$, $k=1$, ..., $k=m-1$.

On conclut alors comme à la question précédente[4] : $\delta_{i,1} = \delta_{i,2} = \cdots = \delta_{i,m} = 0$.
Et donc pour tout $i \in [\![1,n]\!]$ et tout $j \in [\![1,m]\!]$,

$$\delta_{i,j} = 0 \Leftrightarrow \pi_{i,j} = p_i q_j \Leftrightarrow P([X=x_i]\cap[Y=y_j]) = P(X=x_i)P(Y=y_j).$$

[4] En utilisant cette fois le fait que les y_j sont deux à deux distincts.

C'est donc que X et Y sont deux variables aléatoires indépendantes.

□

Un résultat qui figure au programme est la formule $V(X+Y) = V(X) + 2\operatorname{Cov}(X,Y) + V(Y)$, qui permet de calculer la variance d'une somme de deux variables aléatoires si l'on en connait la covariance.
Cette formule se généralise sans grande difficultés au cas de n variables aléatoires, en se rappelant que $V(X) = \operatorname{Cov}(X,X)$.
En effet, la bilinéarité de la covariance nous donne alors

$$V(X_1+\cdots+X_n) = \operatorname{Cov}\left(\sum_{i=1}^{n} X_i, \sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n}\sum_{j=1}^{n} \operatorname{Cov}(X_i,X_j) = \sum_{i=1}^{n} V(X_i) + \sum_{\substack{1\leqslant i,j\leqslant n \\ i\neq j}} \operatorname{Cov}(X_i,X_j).$$

Cette dernière formule n'est pas au programme, mais il faut savoir la retrouver si besoin.

Exercice 3.27 (ESCP 2017) [ESCP17.3.20] — Facile

Variance d'une somme de lois de Bernoulli.

Soit n un entier naturel supérieur ou égal à 2. On dispose d'un paquet de n cartes $C_1, C_2, \dots, C_n$ que l'on distribue intégralement, les unes après les autres entre n joueur $J_1, \dots, J_n$ selon le protocole suivant :

- la première carte C_1 est donnée à J_1 ;
- la deuxième carte C_2 est distribuée de façon équiprobable entre J_1 et J_2 ;
- la troisième carte C_3 est distribuée de façon équiprobable entre J_1, J_2 et J_3 ;
- et ainsi de suite, jusqu'à la dernière carte C_n qui est donc distribuée de façon équiprobable entre $J_1, J_2, \dots, J_n$.

On suppose l'expérience modélisée par un espace probabilisé $(\Omega, \mathcal{A}, P)$.
On note X_n la variable aléatoire égale au nombre de joueurs qui n'ont reçu aucune carte à la fin de la distribution.

1. Déterminer $X_n(\Omega)$ et calculer $P(X_n=0)$ et $P(X_n=n-1)$.
2. Pour tout i de $[\![1,n]\!]$, on note B_i la variable aléatoire qui vaut 1 si J_i n'a reçu aucune carte et 0 sinon. Déterminer la loi de B_i. Exprimer la variable aléatoire X_n en fonction des variables aléatoires B_i, et en déduire l'espérance de X_n.
3. En faisant le moins de calculs possibles, donner la loi de X_4.
4. (a) Montrer que pour tout i et j dans $[\![1,n]\!]$ tels que $i<j$, on a
$$P([B_i=1]\cap[B_j=1]) = \frac{(i-1)(j-2)}{n(n-1)}.$$
En déduire la covariance des variables aléatoires B_i et B_j.
(b) Montrer que $V(X_n) = \dfrac{n+1}{12}$.

SOLUTION DE L'EXERCICE 3.27

1. Puisque J_1 a forcément au moins une carte[1], X_n prend au maximum la valeur $n-1$. Et dans le meilleur des cas, tous les joueurs ont eu une carte, de sorte que X_n peut prendre la valeur 0.
Et toutes les valeurs intermédiaires sont possibles, de sorte que $X_n(\Omega) = [\![0, n-1]\!]$.

[1] La première.

On a $[X_n = n-1]$ si et seulement si c'est le joueur 1 qui a eu toutes les cartes.
Notons donc A_i l'événement «J_1 a la carte C_i». Alors $[X_n = n-1] = A_1 \cap \cdots \cap A_n$ de sorte que par indépendance des différents choix,

$$P(X_n = n-1) = P(A_1)P(A_2)\cdots P(A_n) = 1 \times \frac{1}{2} \times \cdots \times \frac{1}{n} = \frac{1}{n!}.$$

De même, on a $[X_n = 0]$ si et seulement si pour tout i, c'est le joueur J_i qui a eu la carte C_i.

Détails
Si J_n a une carte, c'est nécessairement C_n. Et donc pour que J_{n-1} ait aussi une carte, ce ne peut être que la carte C_{n-1}, etc.

Et comme précédemment, ceci se réalise avec probabilité $1 \times \frac{1}{2} \times \cdots \times \frac{1}{n} = \frac{1}{n!}$.

2. Notons $A_{i,j}$ l'événement «J_i a la carte C_j». Alors

$$[B_i = 1] = \overline{A_{i,1}} \cap \overline{A_{i,2}} \cap \cdots \cap \overline{A_{i,n}} = \overline{A_{i,i}} \cap \overline{A_{i,i+1}} \cap \cdots \cap \overline{A_{n,i}}.$$

Et donc, par indépendance des différents choix

$$\begin{aligned} P(B_i = 1) &= P\left(\overline{A_{i,i}}\right) \times P\left(\overline{A_{i+1,i}}\right) \times \cdots \times P\left(\overline{A_{n,i}}\right) \\ &= \frac{i-1}{i}\frac{i}{i+1}\cdots\frac{n-1}{n} = \frac{i-1}{n}. \end{aligned}$$

Et donc B_i suit la loi de Bernoulli de paramètre $\frac{i-1}{n}$.
Puisque $X_n = \sum_{i=1}^{n} B_i$, par linéarité de l'espérance, il vient

$$E(X_n) = \sum_{i=1}^{n} E(B_i) = \sum_{i=1}^{n} \frac{i-1}{n} = \frac{1}{n}\sum_{j=0}^{n-1} j = \frac{n-1}{2}.$$

3. Nous savons que $X_4(\Omega) = [\![0, 3]\!]$, et que $P(X_4 = 0) = P(X_4 = 3) = \frac{1}{4!} = \frac{1}{24}$.

4.a. L'événement $[B_i = 1] \cap [B_j = 1]$ est réalisé si et seulement si les cartes $C_i, \ldots, C_{j-1}$ n'ont pas été données à J_i et que les cartes $C_j, \ldots, C_n$ n'ont été données ni à J_i, ni à J_j.
Avec les notations précédentes, cela donne

$$[B_i = 1] \cap [B_j = 1] = \left(\bigcap_{k=i}^{j-1} \overline{A_{k,i}}\right) \cap \left(\bigcap_{k=j}^{n} \left(\overline{A_{k,i}} \cap \overline{A_{k,j}}\right)\right).$$

Mais pour $k \geqslant j$, on a $P\left(\overline{A_{k,i}} \cap \overline{A_{k,j}}\right) = \frac{k-2}{k}$.
Et donc par indépendance des différents choix,

$$\begin{aligned} P([B_i = 1] \cap [B_j = 1]) &= \prod_{k=1}^{j-1} \frac{k-1}{k} \times \prod_{k=j}^{n} \frac{k-2}{k} \\ &= \prod_{k=i}^{n} \frac{1}{k} \times \prod_{k=i-1}^{j-2} k \times \prod_{k=j-2}^{n-2} k \\ &= \frac{(i-1)!}{n!}\frac{(j-2)!}{(i-2)!}\frac{(n-2)!}{(j-3)!} \\ &= \frac{(n-2)!}{n!}\frac{(i-1)!}{(i-2)!}\frac{(j-2)!}{(j-3)!} = \frac{(i-1)(j-2)}{n(n-1)}. \end{aligned}$$

Par le théorème de transfert, on en déduit que

$$E(B_iB_j) = \sum_{k=0}^{1}\sum_{\ell=0}^{1} k\ell P([B_i = k] \cap [B_j = \ell]) = P([B_i = 1] \cap [B_j = 1]) = \frac{(i-1)(j-2)}{n(n-1)}.$$

Explication

Le produit de deux variables de Bernoulli X et Y est encore une variable de Bernoulli, puisqu'il ne peut prendre que les valeurs 0 et 1. Et on a alors

$P(XY = 1) = P([X = 1] \cap [Y = 1])$.

Par la formule de Huygens, on a alors

$$\mathrm{Cov}(B_i, B_j) = E(B_iB_j) - E(B_i)E(B_j) = \frac{(i-1)(j-2)}{n(n-1)} - \frac{i-1}{n}\frac{j-1}{n} = \frac{(i-1)(n-j+1)}{n^2(n-1)}.$$

4.b. Par bilinéarité de la covariance, on a

$$\begin{aligned}
V(X_n) &= V(B_1 + \cdots + B_n)\\
&= \mathrm{Cov}(B_1 + \cdots + B_n, B_1 + \cdots + B_n)\\
&= \sum_{i=1}^{n}\sum_{j=1}^{n} \mathrm{Cov}(B_i, B_j)\\
&= \sum_{i=1}^{n} V(B_i) + \sum_{i=1}^{n}\sum_{\substack{j=1\\ j\neq i}}^{n} \mathrm{Cov}(B_i, B_j)\\
&= \sum_{i=1}^{n} V(B_i) + 2\sum_{1\leqslant i<j\leqslant n} \mathrm{Cov}(B_i, B_j)\\
&= \sum_{i=1}^{n} \frac{(i-1)(n-i+1)}{n^2} + 2\sum_{i=1}^{n}\sum_{j=i+1}^{n} \frac{(i-1)(n-j+1)}{n^2(n-1)}\\
&= \frac{n+1}{n^2}\sum_{i=1}^{n}(i-1) - \frac{1}{n^2}\sum_{i=1}^{n} i(i-1) + \frac{2}{n^2(n-1)}\sum_{i=1}^{n}(i-1)\sum_{k=1}^{n-i} k\\
&= \frac{n+1}{n^2}\frac{n(n-1)}{2} - \frac{1}{n^2}\sum_{i=1}^{n} i^2 + \frac{1}{n^2}\sum_{i=1}^{n} i + \frac{2}{n^2(n-1)}\sum_{i=1}^{n}(i-1)\frac{(n-i)(n-i+1)}{2}\\
&= \frac{(n+1)(n-1)}{2n} - \frac{(n+1)(2n+1)}{6n} + \frac{n+1}{2n} + \frac{2}{n^2(n-1)}
\end{aligned}$$

Chgt d'indice

Dans la dernière somme, on a posé $k = n - j + 1$

□

Exercice 3.28 (ESCP 2018) [ESCP18.3.22] Moyen

Soit $(X_n)_{n\geqslant 0}$ une suite de variables aléatoires réelles définies sur le même espace probabilisé $(\Omega, \mathcal{A}, P)$.
On suppose que les X_n sont indépendantes, et suivent toutes une même loi d'espérance m et d'écart-type σ. On définit alors la suite (Y_n) de variables aléatoires par

$$Y_0 = \frac{X_0}{2} \text{ et } \forall n \in \mathbf{N},\ Y_{n+1} = \frac{X_{n+1} + Y_n}{2}.$$

1. Montrer que pour tout $n \in \mathbf{N}$, on a $Y_n = \sum_{k=0}^{n} \frac{1}{2^{n+1-k}} X_k$. En déduire que Y_n admet une espérance et une variance.
2. Déterminer $E(Y_n)$. Quelle est sa limite lorsque $n \to +\infty$?
3. Calculer $V(Y_n)$ et préciser sa limite lorsque n tend vers $+\infty$.
4. Soient i et j deux entiers positifs tels que $i < j$. Prouver qu'il existe une constante c indépendante de i et j pour laquelle on a : $\mathrm{Cov}(Y_i, Y_j) \leqslant c\frac{1}{2^{j-i}}$.
5. Soit $n \in \mathbf{N}^*$. On pose $S_n = \frac{1}{n}\sum_{k=1}^{n} Y_k$.
   (a) Justifier l'existence de l'espérance et de la variance de S_n. Déterminer la limite ℓ de la suite $(E(S_n))_n$.

(b) Montrer que $V(S_n) = \dfrac{1}{n^2}\left(\sum_{k=1}^{n} V(Y_k) + 2\sum_{1\leqslant i<j\leqslant n} \mathrm{Cov}(Y_i, Y_j)\right)$.
En déduire que $\lim_{n\to+\infty} V(S_n) = 0$.

(c) Justifier que pour tout variable aléatoire X admettant un moment d'ordre 2, on a $E(|X|) \leqslant \sqrt{E(X^2)}$.
Prouver alors que pour tout $\varepsilon > 0$, on a :

$$P(|S_n - \ell| \geqslant \varepsilon) \leqslant \frac{1}{\varepsilon}\left(\sqrt{V(S_n)} + |E(S_n) - \ell|\right).$$

En déduire que la suite $(S_n)_n$ converge en probabilité vers la variable certaine égale à ℓ.

6. On suppose dans cette question que les variables aléatoires X_n sont indépendantes et suivent toutes la loi normale centrée réduite.
(a) Déterminer la loi de Y_n.
(b) Montrer que la suite (Y_n) converge en loi et préciser sa limite.

SOLUTION DE L'EXERCICE 3.28

1. La preuve se fait facilement par récurrence : pour $n = 0$, c'est trivial.

Supposons que $Y_n = \sum_{k=0}^{n} \dfrac{X_k}{2^{n+1-k}}$. Alors

$$Y_{n+1} = \frac{1}{2}\left(\sum_{k=0}^{n} \frac{X_k}{2^{n+1-k}} + X_{n+1}\right) = \sum_{k=0}^{n} \frac{X_k}{2^{n+2-k}} + \frac{X_{n+1}}{2} = \sum_{k=0}^{n+1} \frac{X_k}{2^{n+2-k}}.$$

Et donc par le principe de récurrence, pour tout $n \in \mathbf{N}$, $Y_n = \sum_{k=0}^{n} \dfrac{X_k}{2^{n+1-k}}$.

Puisque les X_k ont une variance, Y_n, qui est une combinaison linéaire des X_k admet une variance, et donc une espérance.

2. Par linéarité de l'espérance, on a

$$E(Y_n) = \sum_{k=0}^{n} E\left(\frac{X_k}{2^{n+1-k}}\right) = m\sum_{k=0}^{n} \frac{1}{2^{n+1-k}} = \frac{m}{2^{n+1}}\sum_{k=0}^{n} 2^k = \frac{m}{2^{n+1}}\left(2^{n+1}-1\right) = m\left(1 - \frac{1}{2^{n+1}}\right) \underset{n\to+\infty}{\longrightarrow} m.$$

3. Les X_i étant indépendantes, on a

$$V(Y_n) = \sum_{k=0}^{n} \frac{V(X_k)}{4^{n+1-k}} = \frac{\sigma^2}{4^{n+1}}\sum_{k=0}^{n} 4^k = \frac{\sigma^2}{4^{n+1}}\frac{4^{n+1}-1}{3} = \frac{\sigma^2}{3}\left(1 - \frac{1}{4^{n+1}}\right) \underset{n\to+\infty}{\longrightarrow} \frac{\sigma^2}{3}.$$

4. Par bilinéarité de la covariance, il vient

$$\begin{aligned}
\mathrm{Cov}(Y_i, Y_j) &= \mathrm{Cov}\left(\sum_{k=0}^{i} \frac{X_k}{2^{i+1-k}}, \sum_{\ell=0}^{j} \frac{X_\ell}{2^{j+1-\ell}}\right) \\
&= \sum_{k=0}^{i}\sum_{\ell=0}^{j} \frac{1}{2^{i+j+2-k-\ell}} \mathrm{Cov}(X_k, X_\ell) \\
&= \sum_{k=0}^{i} \frac{1}{2^{i+j+2-2k}} V(X_k) \\
&= \frac{\sigma^2}{2^{i+j+2}}\sum_{k=0}^{i} 4^k \\
&\leqslant \sigma^2 \frac{4^{i+1}}{3}\frac{1}{2^{i+j-2}} \\
&\leqslant \frac{\sigma^2}{3}\frac{1}{2^{j-i}}.
\end{aligned}$$

Détails
Les X_k sont indépendantes, donc $\mathrm{Cov}(X_k, X_\ell) = 0$ si $k \neq \ell$.

Donc $c = \dfrac{\sigma^2}{3}$ convient.

5.a. Les Y_k ont une variance, donc leur somme a également une variance, et donc une espérance.
On a alors, par linéarité de l'espérance,

$$E(S_n) = \frac{1}{n}\sum_{k=1}^{n} E(Y_k) = \frac{1}{n}\sum_{k=1}^{n} m\left(1-\frac{1}{2^{k+1}}\right) = m\left(1-\frac{1}{n}\sum_{k=1}^{n}\frac{1}{2^{k+1}}\right) = m\left(1-\frac{m}{4}\sum_{k=0}^{n-1}\frac{1}{2^k}\right)$$
$$= m\left(1-\frac{1}{2n}\left(1-\frac{1}{2^n}\right)\right) \underset{n\to+\infty}{\longrightarrow} m.$$

5.b. Par bilinéarité de la covariance, on a

$$V(S_n) = \mathrm{Cov}(S_n, S_n) = \frac{1}{n^2}\mathrm{Cov}\left(\sum_{k=1}^{n} Y_k, \sum_{\ell=1}^{n} Y_\ell\right)$$
$$= \frac{1}{n^2}\sum_{k=1}^{n}\sum_{\ell=1}^{n}\mathrm{Cov}(Y_k, Y_\ell)$$
$$= \frac{1}{n^2}\sum_{k=1}^{n} V(Y_k) + \frac{1}{n^2}\sum_{\substack{1\leqslant k,\ell\leqslant n\\ k\neq\ell}}\mathrm{Cov}(Y_k, Y_\ell)$$
$$= \frac{1}{n^2}\sum_{k=1}^{n} V(Y_k) + \frac{2}{n^2}\sum_{1\leqslant k<\ell\leqslant n}\mathrm{Cov}(Y_k, Y_\ell).$$

Détails
Chacun des termes de la somme de la ligne précédente apparaît deux fois.

Et par conséquent, il vient

$$V(S_n) = \frac{1}{n^2}\sum_{k=1}^{n}\frac{\sigma^2}{3}\left(1-\frac{1}{4^{k+1}}\right) + \frac{2}{n^2}\sum_{1\leqslant i<j\leqslant n}\mathrm{Cov}(Y_i, Y_j)$$
$$\leqslant \frac{\sigma^2}{3n^3}\left(n-\sum_{k=1}^{n}\frac{1}{4^{k+1}}\right) + \frac{1}{n^2}\sum_{i=1}^{n-1}\sum_{j=i+1}^{n} c\frac{1}{2^{j-i}}$$
$$\leqslant \frac{\sigma^2}{3n} + \frac{2}{n^2}\sum_{i=1}^{n-1} c\sum_{k=i}^{n-i}\frac{1}{2^k}$$
$$\leqslant\leqslant \frac{\sigma^2}{3n} + \frac{2c}{n^2}\sum_{i=1}^{n-1}\left(1-\frac{1}{2^{n-i}}\right)$$
$$\leqslant \frac{\sigma^2}{3n} + \frac{2c}{n^2}\sum_{i=1}^{n-1} 1$$
$$\leqslant \frac{\sigma^2}{3n} + \frac{2c}{n}.$$

Détails
C'est la formule classique pour la somme des termes consécutifs d'une suite géométrique : premier terme fois ...

Et donc on a bien $\lim_{n\to+\infty} V(S_n) = 0$.

5.c. Puisque X admet un moment d'ordre 2, $|X|$ aussi.
Et donc $V(|X|) = E(X^2) - E(|X|)^2 \geqslant 0$.
On en déduit donc que $E(|X|) \leqslant \sqrt{E\left(X^2\right)}$.
Alors, d'après l'inégalité de Markov appliquée à la variable $|S_n - \ell|$, qui est bien positive et admet une espérance, il vient

$$P(|S_n - \ell| \geqslant \varepsilon) \leqslant \frac{E(|S_n - \ell|)}{\varepsilon}.$$

Mais par l'inégalité triangulaire, $|S_n - \ell| \leqslant |S_n - E(S_n)| + |E(S_n) - \ell|$, de sorte que par croissance de l'espérance,

$$E(|S_n - \ell|) \leqslant E\left(|S_n - E(S_n)|\right) + E(|E(S_n) - \ell|) \leqslant \sqrt{\underbrace{E\left((S_n - E(S_n))^2\right)}_{=V(S_n)}} + |E(S_n) - \ell|$$

et donc on a bien

$$P(|S_n - \ell| \geqslant \varepsilon) \leqslant \frac{1}{\varepsilon}\left(\sqrt{V(S_n)} + |E(S_n) - \ell|\right).$$

Et donc en passant à la limite, il vient bien $P(|S_n - \ell| \geqslant \varepsilon) \underset{n\to+\infty}{\longrightarrow} 0$ de sorte que $S_n \xrightarrow{P} \ell$.

6.a. Par transformation affine de loi normale, $\dfrac{X_k}{2^{n+1-k}} \hookrightarrow \mathcal{N}\left(0, \dfrac{1}{4^{n+1-k}}\right)$.

Et donc, par stabilité des lois normales[1],

[1] Les X_k sont indépendantes.

$$Y_k \hookrightarrow \mathcal{N}\left(0, \sum_{k=0}^{n} \frac{1}{4^{n+1-k}}\right) = \mathcal{N}\left(0, \frac{1}{3}\left(1 - \frac{1}{4^{n+1}}\right)\right).$$

6.b. La fonction de répartition de Y_n est donc donnée par

$$P(Y_n \leqslant x) = P\left(\frac{Y_n}{\sqrt{V(Y_n)}} \leqslant \frac{x}{\sqrt{V(Y_n)}}\right) = \Phi\left(\frac{x}{\sqrt{V(Y_n)}}\right).$$

Mais puisque $\sqrt{V(Y_n)} \underset{n\to+\infty}{\longrightarrow} \sqrt{\dfrac{1}{3}}$, on en déduit que

$$P(Y_n \leqslant x) \underset{n\to+\infty}{\longrightarrow} \Phi\left(\sqrt{3}x\right).$$

Remarque
Notons qu'ici, $\sigma = 1$.

On reconnaît là la fonction de répartition d'une loi normale centrée d'écart-type égal à $\dfrac{1}{\sqrt{3}}$.

Et donc $Y_n \xrightarrow{\mathcal{L}} Y$, où $Y \hookrightarrow \mathcal{N}\left(0, \dfrac{1}{3}\right)$.

□

3.4.2 Probabilités et algèbre linéaire

Des exercices que l'on voit régulièrement définissent une matrice aléatoire, dépendant de deux (voire plus) variables aléatoires X et Y, et nous demandent de calculer la probabilité que cette matrice soit inversible/diagonalisable/etc.
De tels exercices s'abordent généralement de la manière suivante : considérer X et Y comme deux réels fixés, et déterminer, à l'aide d'outils d'algèbre linéaire, à quelle(s) condition(s) sur X et Y la matrice vérifie bien la condition demandée.
Une fois cette condition établie, il s'agit de déterminer la probabilité que cet événement soit réalisé.

EXERCICE 3.29 (QSP HEC 2007) Facile

Inversibilité d'une matrice aléatoire.

Soient X et Y deux variables aléatoires indépendantes sur $(\Omega, \mathcal{A}, P)$ 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.

SOLUTION DE L'EXERCICE 3.29 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 $\overline{A} = [X^2 = Y^2]$, et puisque X et Y sont à valeurs positives, on a $\overline{A} = [X = Y]$.
En appliquant la formule des probabilités totales au système complet d'événements $\{[X = k],\ k \in \mathbf{N}^*\}$, il vient

$$P\left(\overline{A}\right) = P(X = Y) = \sum_{k=1}^{+\infty} P([X = Y] \cap [X = k]) = \sum_{k=1}^{+\infty} P([X = k] \cap [Y = k]).$$

Par indépendance de X et Y, il vient donc

$$P\left(\overline{A}\right) = \sum_{k=1}^{+\infty} P(X = k)P(Y = k) = \sum_{k=1}^{+\infty} p^2(1-p)^{2(k-1)} = p^2 \sum_{i=0}^{+\infty} \left((1-p)^2\right)^i = p^2 \frac{1}{1-(1-p)^2} = \frac{p^2}{2p-p^2} = \frac{p}{2-p}.$$

Et donc

$$P(A) = 1 - P\left(\overline{A}\right) = 1 - \frac{p}{2-p} = \frac{2-2p}{2-p}.$$

□

Exercice 3.30 (QSP ESCP 2016) [ESCP16.QSP03] — Difficile

Endomorphismes aléatoires

Soit $p \in]0,1[$. Soient X et Y deux variables aléatoires réelles indépendantes, définies sur le même espace probabilisé $(\Omega, \mathcal{A}, P)$ et de même loi géométrique de paramètre p.
Soit $f : \mathbf{R}^2 \to \mathbf{R}^2$ définie par $f : (u,v) \mapsto (4Xu - 6Yv, 2Xu - 3Yv)$.
Déterminer la probabilité que f soit un projecteur.

Solution de l'exercice 3.30 Raisonnons sous forme matricielle : la matrice M de f dans la base canonique de $\mathbf{R}^2$ est $\begin{pmatrix} 4X & -6Y \\ 2X & -3Y \end{pmatrix}$.

Détails
On utilise ici la caractérisation suivante des projecteurs : un endomorphisme f est un projecteur si et seulement si il est diagonalisable avec $\mathrm{Spec}(f) \subset \{0,1\}$.

Ses deux colonnes sont proportionnelles, et sont non nulles, de sorte que $\mathrm{rg}(M) = 1$.
Ainsi, 0 est valeur propre de f, avec un sous-espace propre de dimension 1.
Donc f est un projecteur si et seulement si 1 est valeur propre de f.
Or, on a $M - I_2 = \begin{pmatrix} 4X-1 & -6Y \\ 2X & -3Y-1 \end{pmatrix}$, qui n'est pas inversible si et seulement si son déterminant est nul, c'est-à-dire si et seulement si

$$(4X-1)(-3Y-1) + 12XY = 0 \Leftrightarrow 4X = 1 + 3Y.$$

Donc la probabilité cherchée est $P(4X = 1+3Y)$.
Puisque X et Y sont à valeurs dans $\mathbf{N}^*$, cherchons tous les couples $(k,\ell) \in (\mathbf{N}^*)^2$ tels que $1 + 3k = 4\ell$.
Puisque 4ℓ est pair, $1+3k$ doit l'être également, ce qui est le cas si et seulement si $3k$ est impair, et donc si et seulement si k est impair.
Cherchons donc k sous la forme $2p+1$: on a $1 + 3(2p+1) = 4\ell$ si et seulement si $6p = 4(\ell - 1) \Leftrightarrow 3p = 2(\ell - 1)$.
Il faut alors que p soit pair, de la forme $p = 2n$, et alors $\ell = 3n + 1$.
Inversement, si $k = 4n+1$ et $\ell = 3n+1$, avec $n \in \mathbf{N}$ alors on a bien $3k + 1 = 4\ell$.
Ainsi, on a

$$\begin{aligned}
P(4X = 3Y+1) &= \sum_{n=0}^{+\infty} P([X = 3n+1] \cap [Y = 4n+1]) \\
&= \sum_{n=0}^{+\infty} P(X = 3n+1)P(Y = 4n+1) && X \text{ et } Y \text{ sont indépendantes.} \\
&= \sum_{n=0}^{+\infty} p(1-p)^{3n} p(1-p)^{4n} \\
&= p^2 \sum_{n=0}^{+\infty} (1-p)^{7n} \\
&= p^2 \sum_{n=0}^{+\infty} \left((1-p)^7\right)^n \\
&= p^2 \frac{1}{1-(1-p)^7}. && \text{Somme d'une série géométrique de raison } (1-p)^7 \in]0,1[.
\end{aligned}$$

□

Exercice 3.31 (QSP ESCP 2009) [ESCP09.QSP02] — Facile

Probabilité que deux matrices aléatoires soient semblables.

Soient X et Y deux variables aléatoires indépendantes qui suivent la loi géométrique de paramètre p.
Trouver la probabilité que les deux matrices $\begin{pmatrix} 1 & Y \\ 0 & 2 \end{pmatrix}$ et $\begin{pmatrix} X & X \\ 0 & Y \end{pmatrix}$ soient semblables.

SOLUTION DE L'EXERCICE 3.31 Notons que pour tout $y \in \mathbf{R}$, la matrice $\begin{pmatrix} 1 & y \\ 0 & 2 \end{pmatrix}$ possède deux valeurs propres 1 et 2, et est donc diagonalisable.
Par conséquent, si $\begin{pmatrix} 1 & y \\ 0 & 2 \end{pmatrix}$ et $\begin{pmatrix} x & x \\ 0 & y \end{pmatrix}$ sont semblables, elles doivent avoir les mêmes valeurs propres.
Or, les valeurs propres de la seconde matrice, qui est triangulaire, sont x et y.
Donc $\begin{pmatrix} 1 & y \\ 0 & 2 \end{pmatrix}$ et $\begin{pmatrix} x & x \\ 0 & y \end{pmatrix}$ possèdent les mêmes valeurs propres si et seulement si $\begin{cases} x = 1 \\ y = 2 \end{cases}$ ou $\begin{cases} x = 2 \\ y = 1 \end{cases}$.
Dans ce cas, les deux matrices sont diagonalisables[1] et possèdent les mêmes valeurs propres, avec des des sous-espaces propres tous deux de dimension 1.
Elles sont donc semblables.
Ainsi, $\begin{pmatrix} 1 & y \\ 0 & 2 \end{pmatrix}$ et $\begin{pmatrix} x & x \\ 0 & y \end{pmatrix}$ sont semblables si et seulement si $\begin{cases} x = 1 \\ y = 2 \end{cases}$ ou $\begin{cases} x = 2 \\ y = 1 \end{cases}$.
La probabilité cherchée est donc

$$P\left(([X = 1] \cap [Y = 2]) \cup ([X = 2] \cap [Y = 1])\right).$$

Par incompatibilité de ces deux événements, et par indépendance de X et Y, il s'agit donc de

$$\begin{aligned} P([X = 1] \cap [Y = 2]) + P([X = 2] \cap [Y = 1]) &= P(X = 1)P(Y = 2) + P(X = 2)P(Y = 1) \\ &= p^2(1-p) + p^2(1-p) = 2p^2(1-p). \end{aligned}$$

□

[1] Car possèdent deux valeurs propres distinctes.

Rappel

Deux matrices **diagonalisables** sont semblables si et seulement si elles ont les mêmes valeurs propres et des sous-espaces propres de mêmes dimensions.

3.4.3 Variables définies à partir d'un nombre aléatoire de variables aléatoires

À l'écrit comme à l'oral, il est très classique d'avoir à étudier une somme, un maximum ou un minimum d'un nombre fixé de variables aléatoires.
On peut parfois rencontrer des exercices plus délicats, où l'on étudie par exemple la somme d'un nombre aléatoire de variables aléatoires.
Autrement dit, si N est une variable aléatoire à valeurs dans $\mathbf{N}$ et $(X_i)_{i\in\mathbf{N}}$ sont des variables aléatoires, alors on peut définir une variable aléatoire S par :

$$\forall \omega \in \Omega,\ S(\omega) = \sum_{k=1}^{N(\omega)} X_i(\omega).$$

La clé pour étudier une telle variable aléatoire est généralement la formule des probabilités totales (ou son alter ego en termes d'espérance, qui est la formule de l'espérance totale), appliquée au système complet d'événements $\{[N = n],\ n \in \mathbf{N}\}$.

L'exercice suivant prouve des formules donnant l'espérance et la variance d'une somme d'un nombre aléatoire de variables aléatoires.

EXERCICE 3.32 (ESCP 2011) [ESCP11.3.14] **Facile**

Formules de Wald

On considère une suite de variables aléatoires réelles mutuellement indépendantes $N, X_1, \ldots, X_n, \ldots$ définies sur le même espace probabilisé $(\Omega, \mathcal{A}, P)$.
On suppose que N est une variable aléatoire réelle à valeurs dans $\mathbf{N}^*$, possédant un moment d'ordre 2 et que les variables aléatoires (X_i), $i \in \mathbf{N}^*$, suivent la même loi que X, où X est une variable aléatoire réelle à valeurs dans $\mathbf{N}$ et possédant un moment d'ordre 2.

On note Y la variable aléatoire définie par $Y = \sum_{i=1}^{N} X_i$, c'est-à-dire :

$$\forall \omega \in \Omega,\ Y(\omega) = \sum_{i=1}^{N(\omega)} X_i(\omega).$$

1. Déterminer l'espérance $E(Y)$ en fonction de $E(X)$ et de $E(N)$.
2. En utilisant la formule de l'espérance totale, déterminer $E\left(Y^2\right)$ en fonction de $E(X), V(X), E(N)$ et $E\left(N^2\right)$.
3. En déduire $V(Y)$ en fonction de $E(X), V(X), E(N)$ et $V(N)$.
4. Une urne contient n jetons numérotés de 1 à n et on dispose d'une pièce de monnaie qui donne le côté pile avec la probabilité p, où $0 < p < 1$. un joueur tire un jeton dans l'urne et lance ensuite la pièce de monnaie autant de fois que le numéro indiqué par le jeton.
   Calculer la moyenne et la variance de la variable aléatoire comptabilisant le nombre de piles obtenus.

Solution de l'exercice 3.32

1. Appliquons la formule de l'espérance totale au système complet d'événements $\{[N = n], n \in \mathbf{N}^*\}$.
On a alors, sous réserve de convergence,

$$\begin{aligned}
E(Y) &= \sum_{n=1}^{+\infty} E(Y|[N=n])P(N=n) \\
&= \sum_{n=1}^{+\infty} E(X_1 + X_2 + \cdots + X_n|[N=n])P(N=n) \\
&= \sum_{n=1}^{+\infty} E(X_1 + X_2 + \cdots + X_n)P(N=n) \\
&= \sum_{n=1}^{+\infty} nE(X)P(N=n) \\
&= E(X)\sum_{n=1}^{+\infty} nP(N=n) \\
&= E(X)E(N).
\end{aligned}$$

Détails
Par le lemme des coalitions, $X_1 + \cdots + X_n$ est indépendante de N, donc son espérance conditionnelle sachant $[N = n]$ est son espérance.

Puisque N admet une espérance, la série converge, et donc Y admet bien une espérance.

2. De la même manière, on a, toujours sous réserve de convergence,

$$E\left(Y^2\right) = \sum_{n=1}^{+\infty} E\left((X_1 + \cdots + X_n)^2 |N=n\right) P(N=n).$$

Or, les X_i étant mutuellement indépendantes, on a

$$V(X_1 + \cdots + X_n) = V(X_1) + \cdots + V(X_n) = nV(X)$$

et donc par la formule de Huygens,

$$E\left((X_1 + \cdots + X_n)^2\right) = V(X_1 + \cdots + X_n) + E(X_1 + \cdots + X_n)^2 = nV(X) + n^2E(X)^2.$$

Et donc

$$\begin{aligned}
E\left(Y^2\right) &= \sum_{n=1}^{+\infty} \left(nV(X) + n^2E(X)^2\right) P(N=n) \\
&\quad V(X)\sum_{n=1}^{+\infty} nP(N=n) + E(X)^2\sum_{n=1}^{+\infty} n^2P(N=n) \\
&= V(X)E(N) + E(X)^2E\left(N^2\right).
\end{aligned}$$

Toutes les séries en jeu convergent bien car N admet un moment d'ordre 2, donc $E\left(Y^2\right)$ existe.

3. Par la formule de Huygens appliquée à Y, on a donc

$$V(Y) = E\left(Y^2\right) - E(Y)^2 = V(X)E(N) + E(X)^2E\left(N^2\right) - E(X)^2E(N)^2 = V(X)E(N) + E(X)^2V(N).$$

4. Dans la situation décrite par l'énoncé, la variable N qui indique le numéro du jeton tiré suit une loi uniforme sur $[\![1,n]\!]$, et les X_i suivent des lois de Bernoulli de paramètre p.
Si on note Y le nombre total de piles obtenus, les formules précédentes s'appliquent :

$$E(Y) = E(N)E(X) = \frac{n+1}{2}p$$

et

$$V(Y) = V(X)E(N) + E(X)^2V(N) = p(1-p)\frac{n+1}{2} + p^2\frac{n^2-1}{12}$$
$$= p\frac{n+1}{12}(6(1-p)+p(n-1)) = p\frac{n+1}{12}(p(n-7)+6).$$

□

> **Remarque**
> Ce résultat s'obtiendrait directement en appliquant la formule de l'espérance totale comme à la question 1.

EXERCICE 3.33 (ESCP 2018) [ESCP18.3.01] — Moyen

Minimum d'un nombre aléatoire de lois de Pareto

Toutes les variables aléatoires de cet exercice sont définies sur un espace probabilisé $(\Omega, \mathcal{A}, P)$.

Soit X une variable aléatoire de densité $f : x \mapsto \begin{cases} \frac{2}{x^3} & \text{si } x > 1 \\ 0 & \text{sinon} \end{cases}$

Soit $(X_n)_{n\in\mathbf{N}^*}$ une suite de variables aléatoires indépendantes de même loi que X.

1. On définit, pour tout entier naturel n non nul, la variable aléatoire T_n par :

$$T_n = \frac{\max(X_1,\dots,X_n)}{\sqrt{n}}.$$

   (a) Montrer que la suite $(T_n)_{n\geqslant 1}$ converge en loi vers une variable aléatoire T.

   (b) Vérifier que T est une variable aléatoire à densité et donner une densité de T.

2. On considère une variable aléatoire N, indépendante des X_k, qui suit la loi géométrique de paramètre $1-q^2$, avec $0 < q < 1$.
On définit la variable aléatoire U par : pour tout $\omega \in \Omega$, $U(\omega) = \min(X_1(\omega),\dots,X_{N(\omega)}(\omega))$.

   (a) Montrer que pour tout réel x supérieur ou égal à 1, on a : $P(U > x) = \dfrac{1-q^2}{x^2-q^2}$.

   (b) Établir l'existence de l'espérance de la variable aléatoire U, notée $E(U)$.

   (c) Établir l'égalité suivante :

$$E(U) = 1 + \int_1^{+\infty} P(U > x)\,dx.$$

   (d) Calculer $E(U)$.

SOLUTION DE L'EXERCICE 3.33

1.a. La fonction de répartition de X est donnée par

$$F_X(x) = \int_{-\infty}^{x} f(t)\,dt = \begin{cases} 0 & \text{si } x < 1 \\ 1 - \dfrac{1}{x^2} & \text{si } x \geqslant 1 \end{cases}$$

Et donc par indépendance des X_i, il vient

$$P(T_n \leqslant x) = P(\max(X_1,\dots,X_n) \leqslant \sqrt{n}x) = P\left(\bigcap_{i=1}^{n}[X_i \leqslant x\sqrt{n}]\right) = F_X(x\sqrt{n})^n = \begin{cases} 0 & \text{si } x < \dfrac{1}{\sqrt{n}} \\ \left(1-\frac{1}{nx^2}\right)^n & \text{sinon} \end{cases}$$

Pour $x \leqslant 0$, on a donc $P(T_n \leqslant x) = 0 \underset{n\to+\infty}{\longrightarrow} 0$.

Et pour $x > 0$, alors à partir d'un certain rang, $x > \dfrac{1}{\sqrt{n}}$ et donc $P(T_n \leqslant x) = \left(1 - \dfrac{1}{nx^2}\right)^n$.

Nous reconnaissons là la très classique forme indéterminée 1^∞, et il nous faut passer par la forme exponentielle :

$$n\ln\left(1-\frac{1}{nx^2}\right) \underset{n\to+\infty}{\sim} -n\frac{1}{nx^2} \underset{n\to+\infty}{\longrightarrow} -\frac{1}{x^2}.$$

Et donc par continuité de l'exponentielle , il vient $\left(1-\frac{1}{nx^2}\right)^n \underset{n\to+\infty}{\longrightarrow} e^{-\frac{1}{x^2}}$.

Soit donc $F : x \mapsto \begin{cases} 0 & \text{si } x \leqslant 0 \\ e^{-\frac{1}{x^2}} & \text{si } x > 0 \end{cases}$.

Il est immédiat que F tend vers 0 en $-\infty$ et vers 1 en $+\infty$.
De plus, F est croissante car sur $\mathbf{R}_+$, elle est la composée des deux fonctions croissantes $t \mapsto -\frac{1}{t^2}$ et $t \mapsto e^t$.
Enfin, elle est clairement continue en tout point, sauf peut-être en 0.
Mais on a $\lim\limits_{x\to0^+} F(x) = \lim\limits_{x\to0^+} e^{-\frac{1}{x^2}} = 0 = \lim\limits_{x\to0^-} F(x)$.
Et donc F est continue en 0. Par conséquent, elle est continue à droite en tout point, et donc est la fonction de répartition d'une variable aléatoire T.
Puisqu'on a pour tout $x \in \mathbf{R}$, $\lim\limits_{n\to+\infty} F_{T_n}(x) = F(x) = F_T(x)$, c'est donc que $T_n \xrightarrow{\mathcal{L}} T$.

Rédaction
Rappelons que la continuité est indispensable pour composer les limites.

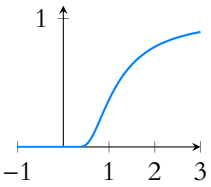


FIGURE 3.4– La fonction F.

1.b. Nous venons déjà de prouver que F est continue sur $\mathbf{R}$, il ne reste donc plus qu'à constater que F est $\mathscr{C}^1$ en tout point, sauf peut-être en 0.
Et donc T est une variable aléatoire à densité, dont une densité est donnée par

$$f_T : x \mapsto \begin{cases} 0 & \text{si } x \leqslant 0 \\ \frac{2}{x^3}e^{-\frac{1}{x^2}} & \text{si } x > 0 \end{cases}$$

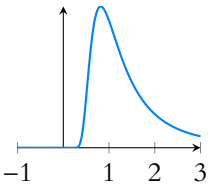


FIGURE 3.5– La densité f_T.

2.a. La difficulté vient ici du fait que la variable U n'est pas le minimum d'un nombre fixé à l'avance de variables i.i.d., mais que ce nombre dépend de la valeur de la variable aléatoire N. Le moyen de s'en sortir est alors de distinguer suivant la valeur de N, autrement dit, de faire appel à la formule des probabilités totales avec le système complet d'événements $\{[N = n],\ n \in \mathbf{N}^*\}$. On a alors, pour $x > 1$,

$$\begin{aligned}
P(U > x) &= \sum_{n=1}^{+\infty} P([U > x] \cap [N = n]) \\
&= \sum_{n=1}^{+\infty} P([\min(X_1, \dots, X_n) > x] \cap [N = n]) \\
&= \sum_{n=1}^{+\infty} P(\min(X_1, \dots, X_n) > x) P(N = n) \\
&= \sum_{n=1}^{+\infty} P(X > x)^n P(N = n) \\
&= \sum_{n=1}^{+\infty} \frac{1}{x^{2n}} \left(q^2\right)^{n-1} (1 - q^2) \\
&= \frac{1-q^2}{x^2} \sum_{n=1}^{+\infty} \left(\frac{q^2}{x^2}\right)^{n-1} \\
&= \frac{1-q^2}{x^2} \frac{1}{1-\frac{q^2}{x^2}} = \frac{1-q^2}{x^2-q^2}.
\end{aligned}$$

Indépendance
Par le lemme des coalitions, $\min(X_1, \dots, X_n)$ est indépendante de N.

Les X_i sont mutuellement indépendantes.

2.b. On a toujours $0 \leqslant U \leqslant X_1$. Or X_1 admet une espérance puisque $\int_1^{+\infty} t f(t)\,dt = \int_1^{+\infty} \frac{2}{t^2}\,dt$ est une intégrale de Riemann convergente.
Donc par domination, U admet également une espérance (et $E(U) \leqslant E(X_1)$).

2.c. Notons que U est à valeurs dans $[1, +\infty[$, et donc que $P(U \leqslant x) = 0$ pour $x < 1$.
D'autre part, pour $x \geqslant 1$, on a $P(U \leqslant x) = 1 - P(U > x) = 1 - \frac{1-q^2}{x^2-q^2} = \frac{x^2-1}{x^2-q^2}$.

Il est alors évident que F_U est $\mathscr{C}^1$ sauf peut-être en 1 et il est facile de constater qu'elle est continue en 1, donc sur $\mathbf{R}$, de sorte que U est une variable à densité.
Nous pourrions en donner une densité f_U, mais ce n'est pas demandé, et nous devrions pouvoir nous en passer dans la suite.
Puisque $E(U)$ existe, on a donc $E(U) = \int_1^{+\infty} t f_U(t)\,dt$.
Procédons alors à une intégration par parties, sur un segment de la forme $[1, A]$, en posant $u(t) = t$, $v(t) = F_U(t) - 1 = -P(U > t)$, qui sont deux fonctions de classe $\mathscr{C}^1$ sur $[1, A]$, avec $u'(t) = 1$ et $v'(t) = f_U(t)$, de sorte que

$$\int_1^A t f_U(t)\,dt = [t(F_U(t)-1)]_1^A + \int_1^A (1-F_U(t))\,dt = 1 - AP(U > A) + \int_1^A P(U > t)\,dt. \qquad (\star)$$

Or, grâce au calcul réalisé précédemment, on a

$$AP(U > A) = A\frac{1-q^2}{A^2-q^2} \underset{A\to+\infty}{\sim} \frac{1-q^2}{A} \underset{A\to+\infty}{\longrightarrow} 0.$$

Et donc en passant à la limite lorsque $A \to +\infty$ dans la relation $(\star)$, il vient

$$E(U) = \int_1^{+\infty} t f_U(t)\,dt = 1 + \int_1^{+\infty} P(U > t)\,dt.$$

Remarque
Puisque l'intégrale du membre de gauche converge, celle de droite converge automatiquement, il n'y a rien à vérifier.

2.d. Il s'agit donc de calculer $\displaystyle\int_1^{+\infty} \frac{1}{x^2-q^2}\,dx$.
Mais notons que $x^2 - q^2 = (x+q)(x-q)$ et donc il est classique[1] de chercher deux réels a et b tels que pour tout $x > q$,

$$\frac{1}{x^2-q^2} = \frac{a}{x+q} + \frac{b}{x-q} = \frac{(a+b)x + q(a-b)}{x^2-q^2}$$

ce qui nous conduit à $a + b = 0$ et $a - b = \dfrac{1}{q}$, de sorte que $a = \dfrac{1}{2q}$ et $b = -\dfrac{1}{2q}$.
Autrement dit, on a donc, pour $A > 1$,

$$\begin{aligned}\int_1^A \frac{1}{x^2-q^2}\,dx &= \frac{1}{2q}\int_1^A \left(\frac{1}{x-q} - \frac{1}{x+q}\right)dx \\ &= \frac{1}{2q}\left[\ln(x-q) - \ln(x+q)\right]_1^A \\ &= \frac{1}{2q}\left(\underbrace{\ln\left(\frac{A-q}{A+q}\right)}_{\underset{A\to+\infty}{\longrightarrow} 1} - \ln\left(\frac{1-q}{1+q}\right)\right) \underset{A\to+\infty}{\longrightarrow} \frac{1}{2q}\ln\left(\frac{1+q}{1-q}\right).\end{aligned}$$

Et par conséquent, on a donc

$$E(U) = 1 + \frac{1-q^2}{2q}\ln\left(\frac{1+q}{1-q}\right).$$

□

[1] Bien qu'aucune connaissance ne soit exigible à ce sujet...

EXERCICE 3.34 (ESCP 2014) [ESCP14.3.15] — Difficile

Somme et minimum d'un nombre aléatoire de lois exponentielles

Soit $(X_n)_{n\geqslant 1}$ une suite de variables aléatoires indépendantes, identiquement distribuées, définies sur un même espace probabilisé $(\Omega, \mathscr{A}, P)$ et suivant la loi exponentielle de paramètre $\lambda > 0$.

Pour tout entier $n \geqslant 1$, on pose $S_n = \sum_{k=1}^{n} X_k$ et $U_n = \min_{1\leqslant k\leqslant n} X_k$.

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 $(\Omega, \mathcal{A}, P)$, indépendante des (X_i) et qui suit la loi géométrique de paramètre p, avec $0 < p < 1$. On définit S et U par :

$$\forall \omega \in \Omega, S(\omega) = S_{N(\omega)}(\omega) \text{ et } U(\omega) = U_{N(\omega)}(\omega).$$

   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.

Solution de l'exercice 3.34

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

$$P(U_n > x) = P\left(\bigcap_{i=1}^{n} [X_i > x]\right) = \prod_{i=1}^{n} P(X_i > x) = (1 - F_{X_1}(x))^n = \begin{cases} 1 & \text{si } x \leqslant 0 \\ e^{-n\lambda x} & \text{si } x > 0 \end{cases}$$

On en déduit que

$$F_{U_n}(x) = 1 - P(U_n > x) = \begin{cases} 0 & \text{si } x \leqslant 0 \\ 1 - e^{-n\lambda x} & \text{si } x > 0 \end{cases}$$

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 $\lambda S_n = \lambda(X_1 + \cdots + X_n) \hookrightarrow \gamma(n)$.
Par transformation affine, une densité de $S_n = \frac{1}{\lambda}\lambda S_n$ est donc

$$x \mapsto \begin{cases} 0 & \text{si } x \leqslant 0 \\ \dfrac{\lambda^n}{(n-1)!} x^{n-1} e^{-\lambda x} & \text{si } x > 0 \end{cases}$$

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

$$\frac{\lambda^n}{(n-1)!} \int_x^A t^{n-1} e^{-\lambda t}\, dt = \frac{\lambda^{n-1}}{(n-1)!} \left[-t^{n-1} e^{-\lambda t}\right]_x^A + \frac{\lambda^{n-1}}{(n-2)!} \int_x^A t^{n-2} e^{-\lambda t}\, dt.$$

En passant à la limite lorsque $A \to +\infty$, il vient, par croissance comparée,

$$\int_x^{+\infty} \frac{\lambda^n}{(n-1)!} t^{n-1} e^{-\lambda t}\, dt = \frac{\lambda^{n-1}}{(n-1)!} x^{n-1} e^{-\lambda x} + \frac{\lambda^{n-1}}{(n-2)!} \int_x^{+\infty} t^{n-2} e^{-\lambda t}\, dt.$$

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.

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 facile[1] de calculer cette dernière intégrale, qui vaut $\frac{e^{-\lambda x}}{\lambda}$. Ainsi, on a bien

$$P(U_n > x) = e^{-\lambda x} \sum_{k=1}^{n-1} \frac{\lambda^k x^k}{k!} + e^{-\lambda x} = e^{-\lambda x} \sum_{k=0}^{n-1} \frac{\lambda^k x^k}{k!}.$$

[1] 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

$$P(U \leqslant x) = \sum_{k=1}^{+\infty} P([U \leqslant x] \cap [N = k]) = \sum_{k=1}^{+\infty} P([U_k \leqslant x] \cap [N = k]).$$

Détail
Si $N = k$, alors $U = U_k$.

Mais N étant indépendante des X_i, par le lemme des coalitions, elle est indépendante de chacune des U_k. Et donc

$$\begin{aligned}
P(U \leqslant x) &= \sum_{k=1}^{+\infty} P(U_k \leqslant x) P(N = k) \\
&= \sum_{k=1}^{+\infty} (1 - e^{-k\lambda x}) p(1-p)^{k-1} \\
&= \sum_{k=1}^{+\infty} p(1-p)^{k-1} - pe^{-\lambda x} \sum_{\ell=0}^{+\infty} e^{-\ell\lambda x} (1-p)^{\ell} \\
&= 1 - pe^{-\lambda x} \sum_{\ell=0}^{+\infty} \left(e^{-\lambda x}(1-p)\right)^{\ell} \\
&= 1 - pe^{-\lambda x} \frac{1}{1 - (1-p)e^{-\lambda x}} \\
&= 1 - \frac{p}{e^{\lambda x} - (1-p)}.
\end{aligned}$$

Cht d'indice
$\ell = k - 1$

Il est alors évident que F_U est $\mathscr{C}^1$ sur $\mathbf{R}_-^*$ ainsi que sur $\mathbf{R}_+^*$. De plus, on a

$$\lim_{x \to 0^+} F_U(x) = 1 - \frac{p}{1-(1-p)} = 0 = \lim_{x \to 0^-} F_U(x).$$

Par conséquent, F_U est continue en 0 et donc sur $\mathbf{R}$, et elle est $\mathscr{C}^1$ 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

$$f_U : x \mapsto \begin{cases} 0 & \text{si } x \leqslant 0 \\ \dfrac{p\lambda}{(e^{\lambda x} - (1-p))^2} & \text{si } x > 0 \end{cases}$$

2.b. De même, S est à valeurs dans $\mathbf{R}_+$, et pour $x \geqslant 0$, on a

$$P(S > x) = \sum_{k=1}^{+\infty} P(S_k > x) P(N = k) = \sum_{k=1}^{+\infty} e^{-\lambda x} \sum_{n=0}^{k-1} \frac{\lambda^n x^n}{n!} p(1-p)^{k-1}.$$

Méthode
Attention aux bornes lorsqu'on permute deux sommes (qu'elles soient finies ou infinies). On sommait pour

$$0 \leqslant n \leqslant k - 1$$

donc pour $k \geqslant n + 1$. Dans la nouvelle somme, une fois établi le fait que n peut prendre toutes les valeurs entières, on somme toujours pour $k \geqslant n + 1$, c'est-à-dire k allant de $n + 1$ à $+\infty$.

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 $F_S(x) = \begin{cases} 0 & \text{si } x \leqslant 0 \\ 1 - e^{-\lambda p x} & \text{si } x > 0 \end{cases}$ et donc S suit la loi exponentielle $\mathcal{E}(p\lambda)$.

□

Les calculs de la question 4 de l'exercice qui suit sont plutôt délirants, surtout en 30 minutes. Toutefois, les méthodes qui y sont développées, y compris dans les dernières questions sont très intéressantes.

Exercice 3.35 (HEC 2008) [HEC08.11] Moyen

Lois de probabilités infiniment divisibles

Les variables aléatoires considérées sont définies sur un espace probabilisé $(\Omega, \mathcal{A}, P)$.
Soit X une variable aléatoire réelle. On dit que X vérifie la propriété $(\mathcal{D})$ si, pour tout entier $n \in \mathbf{N}^*$, il existe n variables aléatoires réelles $X_{1,n}, X_{2,n}, \ldots, X_{n,n}$ mutuellement indépendantes, de même loi et dont la somme a même loi que X.

1. (a) Montrer que si X suit une loi de Poisson, alors X vérifie $(\mathcal{D})$.
   (b) Montrer qu'il en est de même si X suit une loi normale.
2. Une variable aléatoire prenant un nombre fini de valeurs vérifie-t-elle la propriété $(\mathcal{D})$?
3. Soit $(a, b) \in \mathbf{R}^2$ avec $a < b$. On considère dans cette question une variable aléatoire X à valeurs dans $[a, b]$ et vérifiant la propriété $(\mathcal{D})$.
   (a) Montrer que $V(X_{1,n}) \leqslant \dfrac{(b-a)^2}{n^2}$.
   (b) Que peut-on en déduire sur X ?
4. (a) Pour tout $\lambda \in \mathbf{R}^*$, déterminer le réel c_λ tel que la fonction $f_\lambda : x \mapsto \dfrac{\lambda c_\lambda}{\lambda^2 + x^2}$ soit une densité de probabilités.
   (b) Soit X_1 et X_2 deux variables aléatoires indépendantes de densité f_1. Montrer qu'une densité g de $X_1 + X_2$ vérifie :
   $$\forall x \in \mathbf{R},\ g(2x) = \int_{-\infty}^{+\infty} f_1(t+x) f_1(x-t)\, dt.$$
   (c) Soit $x \in \mathbf{R}^*$. Déterminer quatre réels (a, a', b, b'), dépendant de x, tels que
   $$\forall t \in \mathbf{R},\ \frac{1}{[1+(x-t)^2][1+(x+t)^2]} = \frac{a(x-t)+b}{1+(x-t)^2} + \frac{a'(x+t)+b'}{1+(x+t)^2}.$$
   (d) En déduire une expression simple de g sur $\mathbf{R}^*$.
   On admet que f est continue sur $\mathbf{R}$. Calculer $g(0)$.
   (e) On admet que si deux variables aléatoires indépendantes ont pour densités respectives f_λ et f_μ, alors leur somme admet pour densité $f_{\lambda+\mu}$.
   Une variable aléatoire X de densité f_λ vérifie-t-elle la propriété $(\mathcal{D})$?

Solution de l'exercice 3.35

1.a. Si X suit la loi de Poisson $\mathcal{P}(\lambda)$ et si $X_{1,n}, \ldots, X_{n,n}$ sont des variables aléatoires indépendantes de loi $\mathcal{P}\left(\dfrac{\lambda}{n}\right)$, alors par stabilité des lois de Poisson, $X_{1,n} + \cdots + X_{n,n}$ suit la même loi que X.
Et donc X vérifie la propriété $(\mathcal{D})$.

1.b. Si X suit la loi normale $\mathcal{N}(\mu,\sigma^2)$ et si $X_{1,n},\dots,X_{n,n}$ sont des variables indépendantes suivant toutes la loi $\mathcal{N}\left(\frac{\mu}{n},\frac{\sigma^2}{n}\right)$, alors par stabilité des lois normales, $X_{1,n}+\cdots+X_{n,n}$ suit la même loi que X.
Et donc X vérifie la propriété $(\mathscr{D})$.

2. Si X est une variable certaine[1], et si pour tout $i\in[\![1,n]\!]$, $X_{i,n}=\frac{X}{n}$, alors les $X_{i,n}$ sont indépendantes[2] et $X_{1,n}+\cdots+X_{n,n}=X$.
En revanche, si X prend $k\geqslant 2$ valeurs $x_1<x_2<\cdots<x_k$, alors X ne vérifie pas la propriété $(\mathscr{D})$.
Supposons par l'absurde que ce soit le cas, et soient $X_{1,n},\dots,X_{n,n}$ des variables i.i.d. dont la somme a même loi que X.
Alors les x_i doivent prendre des valeurs supérieures ou égales à $\frac{x_1}{n}$.
En effet, si on avait $P\left(X_{i,n}<\frac{x_1}{n}\right)>0$, puisque

$$\left[X_{1,n}<\frac{x_1}{n}\right]\cap\cdots\cap\left[X_{n,n}<\frac{x_1}{n}\right]\subset\left[X_{1,n}+\cdots+X_{n,n}<x_1\right]$$

il viendrait alors

$$0<P\left(X_{1,n}<\frac{x_1}{n}\right)^n\leqslant P(X_{1,n}+\cdots+X_{n,n}<x_1)=P(X<x_1)=0.$$

D'autre part, on doit avoir $P\left(X_{i,n}=\frac{x_1}{n}\right)\neq 0$.
En effet, dans le cas contraire, il viendrait alors $P\left(X_{i,n}>\frac{x_1}{n}\right)=1$, et puisque

$$\left[X_{1,n}>\frac{x_1}{n}\right]\cap\cdots\cap\left[X_{n,n}>\frac{x_1}{n}\right]\subset\left[X_{1,n}+\cdots+X_{n,n}>x_1\right],$$

on aurait donc $P(X>x_1)=1^n=1$, contredisant le fait que $P(X=x_1)\neq 0$.
Sur le même principe, on prouve que les $X_{i,n}$ prennent des valeurs inférieures ou égales à $\frac{x_k}{n}$ et que $P\left(X_{i,n}=\frac{x_k}{n}\right)\neq 0$.
Et donc pour tout $j\in[\![1,n]\!]$, $P\left(X=j\frac{x_1}{n}+(n-j)\frac{x_k}{n}\right)\neq 0$ puisque

$$\left[X_{1,n}=\frac{x_1}{n}\right]\cap\cdots\cap\left[X_{j,n}=\frac{x_1}{n}\right]\cap\left[X_{j+1,n}=\frac{x_k}{n}\right]\cap\cdots\cap\left[X_{n,n}=\frac{x_k}{n}\right]\subset\left[X_{1,n}+\cdots+X_{n,n}=j\frac{x_1}{n}+(n-j)\frac{x_k}{n}\right].$$

Et donc X peut prendre au moins n valeurs différentes.
Ceci est notamment vrai pour $n>k$ contredisant le fait que $X(\Omega)$ est de cardinal k.
Remarque *: on irait sûrement plus vite en disant dès le départ que les $X_{i,n}$ ne sont pas des variables certaines, et donc prennent au moins deux valeurs a_n et b_n, de sorte X peut prendre toutes les valeurs $ja_n+(n-j)b_n$ et conclure comme précédemment.*
Cela suppose tout de même que les $X_{i,n}$ sont des variables discrètes (ou au moins que $P(X_{i,n}=a)\neq 0$ et $P(X_{i,n}=b)\neq 0$).
Nous n'avons pas eu besoin de faire cette hypothèse dans notre raisonnement.

3.a. Puisque X est à valeurs dans $[a,b]$, les $X_{i,n}$ sont à valeurs dans $\left[\frac{a}{n},\frac{b}{n}\right]$, avec le même raisonnement qu'à la question précédente.
Mais alors $X_{i,n}-\frac{a}{n}$ est à valeurs dans $\left[0,\frac{b-a}{n}\right]$, de sorte que

$$0\leqslant\left(X_{i,n}-\frac{a}{n}\right)^2\leqslant\frac{(b-a)^2}{n^2}.$$

Et donc par domination, $X_{i,n}-\frac{a}{n}$ admet un moment d'ordre 2, et $E\left(\left(X_{i,n}-\frac{a}{n}\right)^2\right)\leqslant\frac{(b-a)^2}{n^2}$.
Et donc par la formule de Huygens[3],

$$V(X_{i,n})=V\left(X_{i,n}-\frac{a}{n}\right)=E\left(\left(X_{i,n}-\frac{a}{n}\right)^2\right)-E\left(X_{i,n}-\frac{a}{n}\right)^2\leqslant\frac{(b-a)^2}{n^2}.$$

[1] C'est-à-dire ne prenant qu'une valeur.

[2] Car certaines, et qu'une variable certaine est indépendante de toute autre variable aléatoire.

Remarque
Nous avons totalement détaillé ici, mais il y a fort à parier qu'à l'oral, un candidat a déjà l'intuition de ces résultats sans être forcément capable de les écrire ferait déjà forte impression.

[3] Puisque $X_{i,n}-\frac{a}{n}$ admet un moment d'ordre 2, elle a une variance, et donc il en est de même de $X_{i,n}$.

3.b. Les $X_{i,n}$ étant indépendantes, on a donc

$$V(X) = V\left(X_{1,n} + \cdots + X_{n,n}\right) \leqslant \sum_{i=1}^{n} V(X_{i,n}) \leqslant \frac{(b-a)^2}{n}.$$

Ceci étant vrai pour tout $n \in \mathbf{N}^*$, en faisant tendre n vers $+\infty$, par le théorème des gendarmes[4], on a donc $V(X) = 0$.
Et donc X est une variable certaine.

> **Remarque**
> Notons que cela généralise le résultat de la question précédente.

4.a. La fonction f_λ est évidemment continue sur $\mathbf{R}$, et est positive si et seulement si c_λ est du même signe que λ.
Si $\lambda > 0$ en procédant au changement de variable $t = \dfrac{x}{\lambda}$, on a[5]

$$\int_{-\infty}^{+\infty} f_\lambda(t)\,dx = c_\lambda \int_{-\infty}^{+\infty} \lambda^2 \frac{t}{\lambda^2 + \lambda^2 t^2} = \int_{-\infty}^{+\infty} \frac{1}{1+t^2}\,dt.$$

Or, il est classique que

$$\int_{-\infty}^{+\infty} \frac{dt}{1+t^2} = 2\int_{0}^{+\infty} \frac{dt}{1+t^2} = \lim_{A\to+\infty} 2\int_0^A \frac{dt}{1+t^2} = \lim_{A\to+\infty} 2\text{Arctan}(A) = \pi.$$

Et donc f_λ est une densité si et seulement si $c_\lambda = \dfrac{1}{\pi}$.
Dans le cas où $\lambda < 0$, le même changement de variable fonctionne, mais on prendra garde au fait que celui-ci change le sens des bornes :

$$\int_{-\infty}^{+\infty} f_\lambda(x)\,dx = c_\lambda \int_{+\infty}^{-\infty} \frac{dt}{1+t^2} = -c_\lambda \int_{-\infty}^{+\infty} \frac{dt}{1+t^2} = -\pi c_\lambda.$$

Et donc f_λ est une densité si et seulement si $c_\lambda = -\dfrac{1}{\pi}$.

4.b. La densité f_1 étant bornée (par $f_1(0)$), par le théorème de convolution, une densité de $X_1 + X_2$ est donnée par

$$\forall x \in \mathbf{R},\ g(2x) = \int_{-\infty}^{+\infty} f_1(u) f_1(2x-u)\,du.$$

Procédons alors au changement de variable affine $u = t + x$. Il vient alors

$$g(2x) = \int_{-\infty}^{+\infty} f_1(t+x) f_1(x-t)\,dt.$$

4.c. Armons-nous d'une bonne dose de courage, et mettons l'expression de gauche au même dénominateur :

$$\begin{aligned}
\frac{a(x-t)+b}{1+(x-t)^2} + \frac{a'(x+t)+b'}{1+(x+t)^2} &= \frac{(a(x-t)+b)\left(1+(x+t)^2\right) + (a'(x+t)+b')\left(1+(x-t)^2\right)}{[1+(x-t)^2][1+(x+t)^2]} \\
&= \frac{t^3(a-a') + t^2(-ax - a'x + b + b')}{[1+(x-t)^2][1+(x+t)^2]} \\
&+ \frac{t(ax^2 - a + 2bx - a'x^2 + a' - 2b'x)}{[1+(x-t)^2][1+(x+t)^2]} \\
&+ \frac{ax + ax^3 + b + bx^2 + a' + a'x^3 + b' + b'x^2}{[1+(x-t)^2][1+(x+t)^2]}.
\end{aligned}$$

Pour que cette expression soit égale à $\dfrac{1}{[1+(x-t)^2][1+(x+t)^2]}$ pour tout t, il faut et il suffit, par identification des coefficients du numérateur que

$$\begin{cases}
a - a' = 0 \\
-ax - a'x + b + b' = 0 \\
ax^2 - a + 2bx - a'x^2 + a' - 2b'x = 0 \\
ax + ax^3 + b + bx^2 + a' + a'x^3 + b' + b'x^2 = 1
\end{cases}$$

[4] Sous réserve de convergence, mais celle-ci sera établie dans quelques lignes.

La première équation nous donne $a = a'$, et donc la troisième devient $b = b'$.
La seconde équation devient alors $b = ax$, et donc après substitution dans la dernière, il vient

$$a = a' = \frac{1}{4x(x^2+1)} \text{ et } b = b' = \frac{1}{4(x^2+1)}.$$

4.d. On a donc

$$\begin{aligned} g(2x) &= \frac{1}{\pi^2}\int_{-\infty}^{+\infty} \frac{1}{[1+(x-t)^2][1+(x+t)^2]}\,dt \\ &= \frac{1}{\pi^2}\left(\int_{-\infty}^{+\infty}\left(\frac{a(x-t)}{1+(x-t)^2}+\frac{a(x+t)}{1+(x+t)^2}\right)dt + \int_{-\infty}^{+\infty}\frac{b}{1+(x-t)^2}\,dt + \int_{-\infty}^{+\infty}\frac{b}{1+(x+t)^2}\,dt\right) \end{aligned}$$

> **Danger !**
> Attention à ne pas séparer en deux la première intégrale, car
> $$\int_{-\infty}^{+\infty}\frac{x-t}{1+(x-t)^2}\,dt$$
> est une intégrale divergente, équivalente en $+\infty$ à $\frac{1}{x}$.

Les deux dernières intégrales se calculent à l'aide des changements de variable $u = x - t$ et $u = x + t$, et valent $b\pi$ toutes les deux.
D'autre part, une primitive de $t \mapsto \frac{a(x-t)}{1+(x-t)^2}+\frac{a(x+t)}{1+(x+t)^2}$ est $t \mapsto \frac{a}{2}(\text{Arctan}(x+t) - \text{Arctan}(x-t))$,
de sorte que

$$\int_{-\infty}^{+\infty}\left(\frac{a(x-t)}{1+(x-t)^2}+\frac{a(x+t)}{1+(x+t)^2}\right)dt = 0.$$

Donc au final, il vient

$$g(2x) = \frac{1}{\pi^2}2b\pi = \frac{1}{2\pi(x^2+1)}.$$

Soit enfin[6]

$$g(x) = \frac{1}{2\pi\left(\left(\frac{x}{2}\right)^2+1\right)} = \frac{2}{\pi(x^2+4)} = f_2(x).$$

[6] Pousser un gros «Ouf !» de soulagement.

Puisque f_2 est continue sur $\mathbf{R}$, on a donc $g(0) = \lim_{x\to 0} g(x) = \lim_{x\to 0} f_2(x) = f_2(0) = \frac{1}{2\pi}$.

4.e. Si X admet pour densité f_λ, soient alors $X_{1,n}, \dots, X_{n,n}$ des variables indépendantes de densité $f_{\lambda/n}$.
Alors $X_{1,n} + X_{2,n}$ admet pour densité $f_{2\lambda/n}$.
De plus, par le lemme des coalitions, elle est indépendante de $X_{3,n}$, de sorte que $X_{1,n} + X_{2,n} + X_{3,n}$ admet pour densité $f_{2\lambda/n+\lambda/n} = f_{3\lambda/n}$.
Et une récurrence rapide prouve que $X_{1,n} + \cdots + X_{n,n}$ admet pour densité f_λ.
Et donc X vérifie la propriété $(\mathscr{D})$.

□

3.5 VARIABLES À DENSITÉ

3.5.1 Généralités

Certains exercices, comme celui qui suit, sont vraiment trop durs pour qu'on puisse espérer tout écrire proprement dans le temps imparti (surtout pour les questions sans préparation).
L'examinateur en est forcément conscient, et va avant tout chercher à tester votre intuition, il ne faut alors pas hésiter à donner des réponses partielles, à traiter des cas particuliers, et quand cela peut faciliter la discussion, à faire des dessins !

EXERCICE 3.36 (QSP HEC 2013) [HEC13.QSP60] **Difficile**

Variance d'une variable à densité à valeurs dans $]-1,1[$.

Soit X une variable aléatoire possédant une densité de probabilité continue sur $\mathbf{R}$ et nulle en dehors de $]-1,1[$.

1. Montrer que X possède une variance, qui est strictement comprise en 0 et 1.
2. Montrer que toute valeur de l'intervalle ouvert $]0,1[$ est effectivement possible pour la variance de X.

SOLUTION DE L'EXERCICE 3.36 Notons f la densité de X mentionnée par l'énoncé.

1. Par le théorème de transfert, sous réserve de convergence,

$$E(X^2) = \int_{-\infty}^{+\infty} x^2 f(x)\,dx = \int_{-1}^{1} x^2 f(x)\,dx.$$

Puisque f est continue sur le segment $[-1,1]$, cette intégrale est en fait une intégrale sur un segment, et donc est convergente.
De plus, pour $x \in [-1,1]$, $0 \leqslant x^2 f(x) \leqslant f(x)$, de sorte que

$$0 \leqslant E(X^2) \leqslant \int_{-1}^{1} f(x)\,dx = 1.$$

Ceci prouve déjà que X admet une variance, qui est alors automatiquement strictement positive .
D'autre part, par la formule de Huygens, $V(X) = E(X^2) - E(X)^2 \leqslant E(X^2) \leqslant 1$.
Il reste donc à prouver que $V(X) < 1$, et pour cela, prouvons que $E(X^2) < 1$.
On a

$$1 - E(X^2) = \int_{-1}^{1} (1-x^2) f(x)\,dx.$$

Or, la fonction $x \mapsto (1-x^2)f(x)$ est continue et positive sur $]-1,1[$, et elle n'y est pas nulle.
En effet, f est positive, et $\int_{-1}^{1} f(x)\,dx = 1 > 0$.
Par conséquent, $1 - E(X^2) > 0$, et donc $E(X^2) < 1$.

> **Rappel**
> Une variance est toujours positive, et la variance d'une variable à densité est même strictement positive (les seules variables aléatoires de variance nulle sont les variables presque certaines, qui ne sont pas des variables à densité).

2. Soit $\alpha \in]0,1[$. Nous voulons prouver qu'il existe une densité f, continue sur $\mathbf{R}$ et nulle hors de $]-1,1[$ telle que si X possède f pour densité, alors $V(X) = \alpha$.
Notons tout de suite que la réponse n'est pas à chercher du côté des lois usuelles, car aucune des lois que nous connaissons ne possède une densité continue sur $\mathbf{R}$ et nulle en dehors de $]-1,1[$.

> **Méthode**
> Pour calculer $V(X)$, le plus simple est probablement de passer par la formule de Huygens, qui va nécessiter le calcul de l'espérance.
> Si f est paire, alors $E(X) = 0$, et on économise donc un calcul.

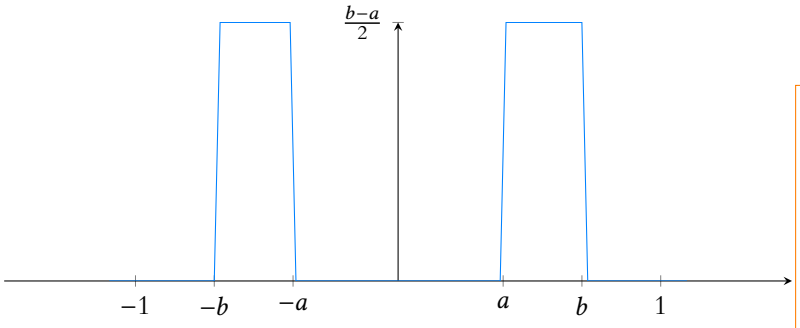
Cherchons de telles densités qui soient paires, de telle sorte que $E(X) = 0$.
Par parité, on doit donc avoir

$$\int_{-1}^{0} f(x)\,dx = \int_{0}^{1} f(x)\,dx = \frac{1}{2}\int_{-1}^{1} f(x)\,dx = \frac{1}{2}\int_{-\infty}^{+\infty} f(x)\,dx = \frac{1}{2}$$

et de même

$$\int_{-1}^{0} x^2 f(x)\,dx = \int_{0}^{1} x^2 f(x)\,dx = \frac{1}{2}\int_{-1}^{1} f(x)\,dx = \frac{E(X^2)}{2} = \frac{V(X)}{2} = \frac{\alpha}{2}.$$

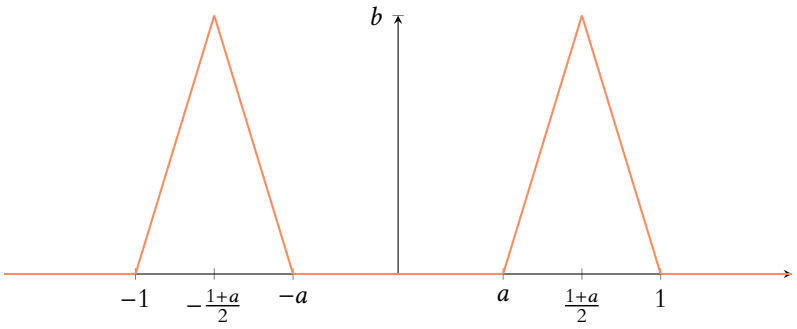
Sans l'hypothèse de continuité demandée par l'énoncé, il serait plus facile de répondre à la question, puisqu'on pourrait prendre une fonction «en escaliers» :



> **Intuition**
> La variance mesure la «dispersion» des valeurs prises par la variable aléatoire.
> Si a et b sont proches de 0, X ne prend que des valeurs proches de 0, et donc sa variance doit être faible.
> Alors qu'au contraire, si a et b sont proches de 1, X ne prend que des valeurs dont la distance à 0 $(= E(X))$ est proche de 1, donc la variance doit être proche de 1.

En prenant a et b proches de 0, on doit pouvoir obtenir une variance proche de 0, et en prenant a et b proches de 1, on doit pouvoir obtenir une variance proche de 1, tous les cas intermédiaires étant possibles.

Malheureusement, on nous demande ici une densité continue, ce qui n'est pas le cas d'une fonction en escaliers...



Pour simplifier les calculs, cherchons f sous la forme d'une fonction affine par morceaux, dont le graphique serait :
Puisque nous voulons une densité, l'intégrale doit être égale à 1. Plutôt que de calculer l'intégrale, remarquons que l'aire de chacun des triangles de la figure doit être égale à $\frac{1}{2}$, et donc on doit avoir

$$\frac{b(1-a)}{2} = \frac{1}{2} \Leftrightarrow b = \frac{1}{1-a}.$$

Détails
Ces expression ont été trouvée à partir du dessin : nous savons comment trouver l'équation de la droite passant par deux points donnés.

Cela revient à prendre :

$$f_a(x) = \begin{cases} 0 & \text{si } 0 \leqslant x \leqslant a \\ \frac{2}{(1-a)^2}(x-a) & \text{si } a < x \leqslant \frac{1+a}{2} \\ -\frac{2}{(1-a)^2}(x-1) & \text{si } \frac{1+a}{2} < x \leqslant 1 \end{cases}$$

avec $a \in [0,1[$.

Soit donc X_a une variable aléatoire de densité f_a, et notons alors

$$g(a) = V(X_a) = E\left(X_a^2\right) \int_{-1}^{1} x^2 f_a(x)\,dx = 2\int_a^1 x^2 f_a(x)\,dx.$$

On a alors

$$g(a) = 2\int_a^{\frac{1+a}{2}} \frac{2}{(1-a)^2} x^2(x-a)\,dx - 2\int_{\frac{1+a}{2}}^1 \frac{2}{(1-a)^2} x^2(x-1)\,dx.$$

On constate alors que g est une fonction continue sur $[0,1]$, par exemple car si on calcule l'intégrale[1], il s'agit d'un polynôme en a.
D'autre part, pour $x \in [a,1]$, on a $x^2 \geqslant a^2$ et donc

$$1 > g(a) = 2\int_a^1 x^2 f_a(x)\,dx \geqslant 2a^2 \underbrace{\int_a^1 f_a(x)\,dx}_{=\frac{1}{2}} \geqslant 2\frac{a^2}{2}.$$

[1] Ce que nous n'allons pas faire ici pour alléger la rédaction.

Par application du théorème des gendarmes, on a donc

$$\lim_{a\to+\infty} = 1.$$

D'autre part, on a

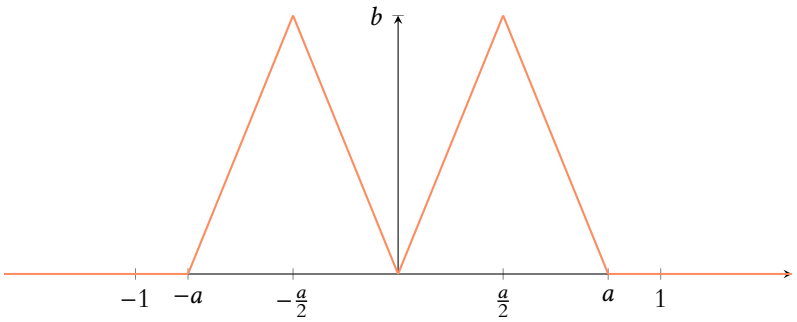
$$g(0) = 2\int_0^{\frac{1}{2}} 2x^3\,dx + 2\int_{\frac{1}{2}}^1 -2x^2(x-1)\,dx = \frac{7}{24}.$$

TVI
Nous n'avons rien dit de la monotonie de g (qui est très probablement croissante strictement), et donc nous faisons bien appel au théorème des valeurs intermédiaires, pas au théorème de la bijection.
Il se peut donc que g prenne des valeurs en dehors de cet intervalle, même si, d'après la question 1, g est à valeurs dans $]0,1[$. M. VIENNEY

Et donc, par le théorème des valeurs intermédiaires, lorsque a parcourt $[0,1]$, $g(a)$ prend toutes les valeurs de $\left[\frac{7}{24}, 1\right[$.

Il reste donc à traiter le cas des «petites[2]» variances. Pour cela, on peut modifier un peu la fonction ci-dessus, et chercher f_a sous la forme

[2] Entre 0 et $\frac{7}{24}$



On montrerait alors comme précédemment qu'on peut ainsi obtenir toutes les variances dans $\left]0, \frac{7}{24}\right]$.
Et donc, quel que soit $\alpha \in]0,1[$, il existe une variable aléatoire X_α possédant une densité continue sur $\mathbf{R}$, nulle en dehors de $]-1,1[$ telle que $V(X_\alpha) = \alpha$.

Et en temps limité ? *: il n'est bien évidemment pas question d'écrire autant de détails durant les 10 à 15 minutes consacrées à la question sans préparation lors d'un oral.*
Mais le jury n'attend pas ces détails, et la deuxième question avait ici pour objectif de tester l'intuition d'un candidat qui aurait déjà réussi la première question.
Un candidat capable de proposer une solution avec une fonction en escaliers (qui ne vérifie donc pas les conditions requises ici) fait déjà preuve d'une très bonne compréhension de la notion de variance.
Et un candidat capable de proposer un dessin d'une fonction affine par morceaux comme ci-dessus mérite une excellente note !

□

3.5.2 Autour des lois usuelles

Exercice 3.37 (HEC 2007) [HEC07.120] **Facile**

Loi de la plus grande valeur propre d'une matrice aléatoire

Soient U et V deux variables aléatoires indépendantes de loi uniforme sur $[0,1]$, définies sur un espace de probabilité noté $(\Omega, \mathcal{A}, P)$.

1. Quelle est la loi de $-\ln(U)$.
   Montrer que la densité de la variable aléatoire $Z = -\ln(U) - \ln(V)$ est donnée par

$$f_Z(x) = \begin{cases} xe^{-x} & \text{si } x \geqslant 0 \\ 0 & \text{sinon} \end{cases}$$

Soit a un réel supérieur ou égal à 1. On définit la matrice

$$M = \begin{pmatrix} 1 & -aU \\ aV & 3 \end{pmatrix}.$$

2. (a) Montrer que la probabilité p que la matrice M ait toutes ses valeurs propres réelles vaut

$$p = \frac{1 + 2\ln(a)}{a^2}.$$

   (b) Montrer que la probabilité que M soit diagonalisable dans $\mathbf{R}$ vaut également p.

3. Dans cette question, on prend $a = 1$. Pour tout $\omega \in \Omega$, on note $X(\omega)$ la plus grande valeur propre de $M(\omega)$.
   Déterminer une densité de X.

Solution de l'exercice 3.37

1. Il est évident que $-\ln(U)$ est à valeurs dans $\mathbf{R}_+$. Et pour $x \geqslant 0$, on a

$$P(-\ln(U) \leqslant x) = P(\ln(U) \geqslant -x) = P(U \geqslant e^{-x}) = 1 - P(U \leqslant e^{-x}) = 1 - e^{-x}.$$

Ainsi, la fonction de répartition de $-\ln(U)$ est

$$x \mapsto \begin{cases} 1-e^{-x} & \text{si } x \geqslant 0 \\ 0 & \text{sinon} \end{cases}$$

On reconnaît alors la fonction de répartition d'une loi exponentielle de paramètre 1 : $-\ln(U) \hookrightarrow \mathcal{E}(1)$.

Classique

Si $\lambda > 0$, alors

$$-\frac{1}{\lambda}\ln(U) \hookrightarrow \mathcal{E}(\lambda).$$

Puisque U et V sont indépendantes, il en est de même de $-\ln(U)$ et $-\ln(V)$, qui suivent toutes les deux la loi $\mathcal{E}(1)$, qui n'est autre que la loi $\gamma(1)$.
Par stabilité des lois gamma, $Z = -\ln(U) - \ln(V) \hookrightarrow \gamma(2)$. Donc une densité en est donnée par

$$f_Z(x) = \begin{cases} \dfrac{1}{\Gamma(2)}x^{2-1}e^{-x} & \text{si } x \geqslant 0 \\ 0 & \text{sinon} \end{cases} = \begin{cases} xe^{-x} & \text{si } x \geqslant 0 \\ 0 & \text{sinon} \end{cases}$$

2.a. Un complexe λ est valeur propre de M si et seulement si $M - \lambda I_2$ n'est pas inversible, soit si et seulement si

$$\det(M-\lambda I_2) = 0 \Leftrightarrow (1-\lambda)(3-\lambda) + a^2UV = 0 \Leftrightarrow \lambda^2 - 4\lambda + 3 + a^2UV = 0.$$

Complexes/réels

Une matrice réelle peut aussi être vue comme une matrice complexe, et donc nous pouvons chercher ses valeurs propres complexes.

Le discriminant de ce polynôme de degré 2 en λ vaut

$$\Delta = 16 - 12 - 4a^2UV = 4 - 4a^2UV = 4(1-a^2UV).$$

Ainsi, toutes les valeurs propres de M sont réelles si et seulement si $\Delta \geqslant 0$. Soit encore

$$\begin{aligned} p &= P(1-a^2UV \geqslant 0) = P\left(UV \leqslant \frac{1}{a^2}\right) \\ &= P\left(\ln(U)+\ln(V) \leqslant -2\ln(a)\right) = P(Z \geqslant 2\ln(a)) \\ &= 1 - \int_0^{2\ln(a)} f_Z(t)\,dt = 1 - \int_0^{2\ln(a)} te^{-t}\,dt \\ &= 1 - [-te^{-t}]_0^{2\ln(a)} - \int_0^{2\ln(a)} e^{-t}\,dt \\ &= 1 + \frac{2\ln(a)}{a^2} + [e^{-t}]_0^{2\ln(a)} \\ &= 1 + \frac{2\ln(a)}{a^2} + \frac{1}{a^2} - 1 = \frac{1+2\ln(a)}{a^2}. \end{aligned}$$

Intégration par parties.

2.b. Pour que M soit diagonalisable, il faut déjà que toutes ses valeurs propres soient réelles.
Si M ne possède qu'une valeur propre réelle λ, alors elle est diagonalisable si et seulement si $M = \lambda I_n$. Ceci n'est clairement pas le cas ici.
Et donc M est diagonalisable si et seulement si elle possède deux valeurs propres distinctes.
Soit si et seulement si $\Delta > 0$.
La probabilité que M soit diagonalisable est donc

$$P(1-a^2UV > 0) = P\left(UV < \frac{1}{a^2}\right) = P\left(-\ln U - \ln V > 2\ln(a)\right).$$

Mais puisque $-\ln(U) - \ln(V)$ est à densité,

$$P\left(-\ln(U)-\ln(V) < 2\ln(a)\right) = P\left(-\ln(U)-\ln(V) \leqslant 2\ln(a)\right) = p.$$

3. Notons que si $a = 1$, alors $p = 1$, et donc M ne possède que des valeurs propres réelles[1].
Nous avons prouvé que les valeurs propres de M sont les racines de $\lambda^2 - 4\lambda + 3 + 4UV$.
Et donc la plus grande valeur propre est

$$X = \frac{4+\sqrt{\Delta}}{2} = 2 + \sqrt{1-UV}.$$

[1] Et donc parler de sa plus grande valeur propre a bien un sens.

Notons que X ne prend que des valeurs entre 2 et 3, donc déjà, pour $x < 2$, $F_X(x) = 0$ et pour $x \geqslant 3$, $F_X(x) = 1$.
Pour $x \in [2,3[$, on a

$$\begin{aligned} F_X(x) &= P\left(2 + \sqrt{1-UV} \leqslant x\right) = P\left(\sqrt{1-UV} \leqslant x-2\right) \\ &= P\left(1 - UV \leqslant (x-2)^2\right) = P\left(UV \geqslant 1-(x-2)^2\right) \\ &= P\left(-\ln(U) - \ln(V) \leqslant -\ln\left(1-(x-2)^2\right)\right) \\ &= \int_0^{-\ln(1-(2-x)^2)} te^{-t}\,dt. \end{aligned}$$

Mais pour $u \in \mathbf{R}_+$, la même intégration par parties qu'à la question 2.a nous donne

$$\int_0^u te^{-t}\,dt = \left[-te^{-t}\right]_0^u + \int_0^u e^{-t}\,dt = 1 - (u+1)e^{-u}.$$

Et donc

$$F_X(x) = 1 - (1 - \ln(1-(2-x)^2))(1-(2-x)^2).$$

Il est alors évident que F_X est $\mathscr{C}^1$ sur $\mathbf{R}$, sauf éventuellement en 2 et en 3.
Et il est facile de vérifier que $\lim\limits_{x\to 2^+} F_X(x) = 0$ et $\lim\limits_{x\to 3^-} F_X(x) = 1$, de sorte que F_X est continue en 2 et en 3, et donc continue sur $\mathbf{R}$.
Ainsi, X est bien une variable à densité, et une densité en est donnée par

$$f_X : x \mapsto \begin{cases} (-2x+4)\ln\left(1-(2-x)^2\right) & \text{si } x \in [2,3[\\ 0 & \text{sinon} \end{cases}$$

□

Exercice 3.38 (QSP HEC 2009) [HEC09.QSP02] — Facile

Loi uniforme et loi géométrique.

Soit U une variable aléatoire définie sur un espace probabilisé $(\Omega, \mathscr{A}, P)$, 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.

Solution de l'exercice 3.38 Il est facile de remarquer que $X(\Omega) = \mathbf{N}^*$, et donc que X est une variable aléatoire discrète.
On a, pour $k \in \mathbf{N}^*$,

$$P(X=k) = P\left(\left\lfloor \frac{\ln U}{\ln q} \right\rfloor = k-1\right) = P\left(k-1 \leqslant \frac{\ln U}{\ln q} < k\right) = P(k\ln q < \ln U \leqslant (k-1)\ln q).$$

> **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 $\ln(q)$.

Par croissance de l'exponentielle, on en déduit que

$$P(X=k) = P(q^k < U \leqslant q^{k-1}) = q^{k-1} - q^k$$

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 $1-q$. □

Exercice 3.39 (QSP ESCP 2018) [ESCP18.QSP04] — Facile

Partie entière et mantisse d'une loi uniforme sur $[0,n]$.

Soit $n \in \mathbf{N}^*$ et soit X une variable aléatoire qui suit une loi uniforme sur $[0,n[$. On note Y la partie entière de X et $Z = X - Y$.
Déterminer la loi de Y, puis celle de Z.

SOLUTION DE L'EXERCICE 3.39 Il est clair que Y est à valeurs dans $[\![0, n-1]\!]$, car X est à valeurs dans $[0, n[$.
De plus, pour $k \in [\![0, n-1]\!]$, on a

$$P(Y = k) = P(\lfloor X \rfloor = k) = P(k \leqslant X < k+1) = \frac{k+1-k}{n-0} = \frac{1}{n}.$$

Et donc Y suit la loi uniforme sur $[\![0, n-1]\!]$.

Alors, Z est la partie décimale de X, et donc prend ses valeurs dans $[0, 1[$.
Et pour $x \in [0, 1[$, par la formule des probabilités totales avec le système complet d'événements $\{[Y = k],\ k \in [\![0, n-1]\!]\}$, il vient

$$P(Z \leqslant x) = \sum_{k=0}^{n-1} P([Z \leqslant x] \cap [Y = k]) = \sum_{k=0}^{n-1} P(k \leqslant X < k+x) = \sum_{k=0}^{n-1} \frac{(k+x)-k}{n} = n\frac{x}{n} = x.$$

Par conséquent, la fonction de répartition de Z est donnée par $F_Z(x) = \begin{cases} 0 & \text{si } x < 0 \\ x & \text{si } 0 \leqslant x < 1 \\ 1 & \text{si } x > 1 \end{cases}$,
de sorte que $Z \hookrightarrow \mathcal{U}([0, 1])$.
Remarque *: notons que ces résultats sont plutôt intuitifs : la partie entière et la partie décimale de X suivent toutes deux des lois uniformes, mais ce résultat ne serait plus valable si X suivait au départ une loi uniforme sur* $\left[0, n+\frac{1}{2}\right[$.
Un prolongement amusant de l'exercice serait le suivant : sachant que $X \hookrightarrow \mathcal{U}([a, b])$, à quelle condition Y suit-elle une loi uniforme ? À quelle condition Z suit-elle une loi uniforme ? □

Réponse
Y est uniforme ssi $\lfloor a \rfloor = \lfloor b \rfloor$ ou si $\lfloor b \rfloor = \lfloor a \rfloor + 1$ et $\frac{a+b}{2} \in \mathbf{Z}$, ou encore si $(a, b) \in \mathbf{Z}^2$.
Et Z est uniforme si et seulement si $b - a \in \mathbf{N}$.

EXERCICE 3.40 (QSP HEC 2013) [HEC13.QSP74]

Maximisation d'une loi normale sur un intervalle de longueur b.

Soit X une variable aléatoire définie sur un espace probabilisé $(\Omega, \mathcal{A}, P)$ suivant la loi normale d'espérance m et de variance égale à 1. Soit b un réel strictement positif fixé.

1. Montrer que l'application $a \mapsto P(a < X < a+b)$, de $\mathbf{R}$ dans $\mathbf{R}$, admet un maximum atteint en un point a_0 que l'on déterminera.
2. Exprimer la valeur de ce maximum à l'aide de la fonction de répartition d'une loi normale centrée réduite.
3. Interpréter géométriquement ce résultat.

SOLUTION DE L'EXERCICE 3.40

1. Notons f la fonction que l'on cherche à maximiser.
Puisque $X - m \hookrightarrow \mathcal{N}(0, 1)$, on a

$$f(a) = P(a < X < a+b) = P(a-m < X-m < a+b-m) = \Phi(a+b-m) - \Phi(a-m).$$

Mais nous savons que Φ est de classe $\mathcal{C}^1$ sur $\mathbf{R}$, et donc c'est également le cas de f, avec

$$f'(a) = \Phi'(a+b-m) - \Phi'(a-m) = \frac{1}{\sqrt{2\pi}}\left(e^{-\frac{1}{2}(a+b-m)^2} - e^{\frac{1}{2}(a-m)^2}\right).$$

On a donc $f'(a) \geqslant 0$ si et seulement si

$$e^{-\frac{1}{2}(a+b-m)^2} \geqslant e^{-\frac{1}{2}(a-m)^2} \Leftrightarrow (a+b-m)^2 \leqslant (a-m)^2 \Leftrightarrow 2(a-m)b+b^2 \leqslant 0 \Leftrightarrow a-m \leqslant -\frac{b}{2} \Leftrightarrow a \leqslant m-\frac{b}{2}.$$

Ainsi, le tableau de variation de f est le suivant.
Et donc f admet un maximum en $a = m - \frac{b}{2}$.

2. On a alors

$$f(a_0) = P\left(m-\frac{b}{2} < X < m+\frac{b}{2}\right) = P\left(-\frac{b}{2} < X-m < \frac{b}{2}\right) = \Phi\left(\frac{b}{2}\right) - \Phi\left(-\frac{b}{2}\right) = 2\Phi\left(\frac{b}{2}\right) - 1.$$

3. Puisque la densité de X est symétrique par rapport à la droite d'équation $x = m$, nous venons de prouver que l'aire sous la courbe, sur un segment de longueur b est maximale lorsque ce segment est centré en m.

□

Exercice 3.41 (QSP HEC 2013) [HEC13.QSP01] — Facile

Soient X et Y deux variables aléatoires indépendantes sur $(\Omega, \mathcal{A}, P)$ 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 $P(X \leqslant Y) \geqslant \frac{1}{2}$.

Solution de l'exercice 3.41 Soit $Z = X - Y$. Puisque $-Y \hookrightarrow \mathcal{N}(-\nu, \sigma^2)$ et que X et $-Y$ sont indépendantes, par stabilité des lois normales $Z \hookrightarrow \mathcal{N}(\mu - \nu, 2\sigma^2)$.
Mais $P(X \leqslant Y) = P(Z \leqslant 0)$.
Or, nous savons que $P(Z \leqslant \mu - \nu) = \frac{1}{2}$.
Donc si $\mu \leqslant \nu$, $P(Z \leqslant 0) \geqslant P(Z \leqslant \mu - \nu) = \frac{1}{2}$.
Inversement, si $\nu < \mu$, alors

$$P(Z \leqslant 0) < P(Z \leqslant \mu - \nu) = \frac{1}{2}.$$

Au final, on a

$$P(X \leqslant Y) \geqslant \frac{1}{2} \Leftrightarrow \mu \leqslant \nu.$$

□

Rédaction
Ne pas oublier de vérifier l'hypothèse d'indépendance pour utiliser la stabilité des lois normales (ou tout autre résultat de stabilité).

Exercice 3.42 (QSP HEC 2013) [HEC13.QSP51] — Facile

Diagonalisabilité d'une matrice 3×3 aléatoire

Soit X et Y deux variables aléatoires indépendantes, définies sur un espace probabilisé $(\Omega, \mathcal{A}, P)$ et de même loi $\mathcal{N}(0, 1)$.
On pose $M = \begin{pmatrix} 0 & X & 0 \\ Y & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$.

1. Calculer $P(X = Y)$ et $P(XY > 0)$.
2. Trouver la probabilité que la matrice M soit diagonalisable.

Solution de l'exercice 3.42

1. On a $P(X = Y) = P(X - Y) = 0$.
Or, par stabilité[1] des lois normales $X - Y \hookrightarrow \mathcal{N}(0, 2)$.
Et en particulier, $X - Y$ est à densité de sorte que $P(X = Y) = P(X - Y = 0) = 0$.
D'autre part, on a

$$\begin{aligned} P(XY > 0) &= P([X > 0] \cap [Y > 0]) + P([X < 0] \cap [Y < 0]) \\ &= P(X > 0)P(Y > 0) + P(X < 0)P(Y < 0) = (1 - \Phi(0))^2 + \Phi(0)^2 = \frac{1}{4} + \frac{1}{4} = \frac{1}{2}. \end{aligned}$$

[1] X et Y sont indépendantes.

Rappel
La probabilité qu'une loi normale centrée (pas nécessairement réduite) soit négative vaut $\Phi(0) = \frac{1}{2}$.

2. Soient x et y deux réels, et soit $M_{x,y} = \begin{pmatrix} 0 & x & 0 \\ y & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$.
Notons que 0 est toujours valeur propre de $M_{x,y}$ puisque sa dernière colonne est nulle.
• Si $x = y = 0$, alors $M_{x,y} = 0$ est diagonalisable.
• Si $x = 0$ et $y \neq 0$, alors $M_{x,y}$ est de rang 1, de sorte que $\dim E_0(M_{x,y}) = 2$.
De plus, si $M_{x,y}$ était diagonalisable, elle admettrait alors une seconde valeur propre λ, et on aurait $2 \times 0 + \lambda \times 1 = \mathrm{tr}(M_{x,y}) = 0$. Donc $\lambda = 0$, ce qui n'est pas possible : $M_{x,y}$ n'est pas diagonalisable.

• On raisonne de même dans le cas où $y = 0$ et $x \neq 0$.
• Si $x \neq 0$ et $y \neq 0$, alors $\text{rg}(M_{x,y}) = 2$, de sorte que $\dim E_0(M_{x,y}) = 1$.
De plus, un réel λ est valeur propre de $M_{x,y}$ si et seulement si $\text{rg}(M_{x,y} - \lambda I_3) < 3$. Or,

$$M_{x,y} - \lambda I_3 = \begin{pmatrix} -\lambda & x & 0 \\ y & -\lambda & 0 \\ 0 & 0 & -\lambda \end{pmatrix} \underset{L_1 \leftrightarrow L_2}{\longleftrightarrow} \begin{pmatrix} y & -\lambda & 0 \\ -\lambda & x & 0 \\ 0 & 0 & -\lambda \end{pmatrix} \underset{L_2 \leftarrow yL_2 + \lambda L_1}{\longleftrightarrow} \begin{pmatrix} y & -\lambda & 0 \\ 0 & yx - \lambda^2 & 0 \\ 0 & 0 & -\lambda \end{pmatrix}.$$

Et donc $\text{rg}(M_{x,y} - \lambda I_3) < 3$ si et seulement si $\lambda^2 - xy = 0$ ou $\lambda = 0$.
Donc si $xy \leqslant 0$, $M_{x,y}$ ne possède pas de valeurs propres non nulles, et si $xy \geqslant 0$, alors $M_{x,y}$ possède trois valeurs propres qui sont 0, $\sqrt{xy}$ et $-\sqrt{xy}$, et donc est diagonalisable.

Ainsi, la probabilité que M soit diagonalisable est

$$P(XY > 0) + P([X = 0] \cap [Y = 0]) = P(XY > 0) + 0 = \frac{1}{2}.$$

□

3.5.3 Autres lois

EXERCICE 3.43 (QSP HEC 2010) [HEC10.QSP01] **Moyen**

Autour de la loi de Cauchy

1. Montrer qu'il existe un réel c pour lequel la fonction $x \mapsto \dfrac{c}{1+x^2}$ est une densité de probabilité.
2. Une variable aléatoire réelle admettant une telle densité possède-t-elle une espérance ?
3. Montrer que si X est une variable aléatoire réelle de densité f, alors X et $\frac{1}{X}$ ont même loi.

SOLUTION DE L'EXERCICE 3.43

1. La fonction f est continue et positive sur $\mathbf{R}$, et il est classique que

$$\int_0^A \frac{dx}{1+x^2} = [\text{Arctan}(x)]_0^A \underset{A \to +\infty}{\longrightarrow} \frac{\pi}{2}$$

de sorte que par parité de l'intégrande, $\displaystyle\int_{-\infty}^{+\infty} \frac{dx}{1+x^2} = \pi$.
Et donc il faut choisir $c = \dfrac{1}{\pi}$ de manière à obtenir une densité de probabilité.

2. Supposons que X soit une variable aléatoire de densité f. Alors X admet une espérance si et seulement si $\dfrac{1}{\pi}\displaystyle\int_{-\infty}^{+\infty} \frac{x}{1+x^2}\,dx$ converge absolument.
Or, au voisinage de $+\infty$, $\dfrac{x}{1+x^2} \underset{x \to +\infty}{\sim} \dfrac{x}{x^2} = \dfrac{1}{x}$ et l'intégrale de Riemann $\displaystyle\int_1^{\infty} \frac{dx}{x}$ diverge.
Par conséquent[1], X n'admet pas d'espérance.

3. Essayons d'exprimer la fonction de répartition de $\frac{1}{X}$ en fonction de celle de X.
Notons que $P(X = 0) = 0$, de sorte que $\frac{1}{X}$ est bien définie.

Si $x = 0$, alors $P\left(\dfrac{1}{X} \leqslant 0\right) = P(X \leqslant 0) = F_X(0)$.
Si $x > 0$, alors

$$F_{1/X}(x) = P\left(\frac{1}{X} \leqslant x\right) = P(X \leqslant 0) + P\left(\frac{1}{x} \leqslant X\right) = F_X(0) + 1 - F_X\left(\frac{1}{x}\right).$$

Si $x < 0$, alors

$$F_{1/X}(x) = P(1/X \leqslant x) = P\left(\frac{1}{x} \leqslant X \leqslant 0\right) = F_X(0) - F_X\left(\frac{1}{x}\right).$$

Détails
En réalité, il existe peut-être des $\omega \in \Omega$ tels que $X(\omega) = 0$, mais ces ω forment un ensemble de mesure nulle. Sur cet ensemble, on peut par exemple décréter que $\frac{1}{X} = 0$, ce qui ne changera pas sa loi.

De plus, nous connaissons la fonction de répartition de X, puisque $\forall x \in \mathbf{R}$,

$$F_X(x) = \frac{1}{\pi}\int_{-\infty}^{x} \frac{dt}{1+t^2} = \frac{1}{\pi}\left(\text{Arctan}(x) + \frac{\pi}{2}\right).$$

Notons qu'en particulier, on a $F_X(0) = \dfrac{1}{2}$, et donc $F_{1/X}(0) = F_X(0)$.
Il s'agit à présent de prouver que

$$\begin{cases} 1 - \frac{1}{\pi}\text{Arctan}\left(\frac{1}{x}\right) = \frac{1}{\pi}\text{Arctan}(x) + \frac{1}{2} \text{ si } x > 0 \\ -\frac{1}{\pi}\text{Arctan}\left(\frac{1}{x}\right) = \frac{1}{\pi}\text{Arctan}(x) + \frac{1}{2} \text{ si } x < 0 \end{cases} \Leftrightarrow \begin{cases} \text{Arctan}(x) + \text{Arctan}\left(\frac{1}{x}\right) = \frac{\pi}{2} \text{ si } x > 0 \\ \text{Arctan}(x) + \text{Arctan}\left(\frac{1}{x}\right) = -\frac{\pi}{2} \text{ si } x < 0 \end{cases}$$

Or, la fonction $g : x \mapsto \text{Arctan}(x) + \text{Arctan}\left(\dfrac{1}{x}\right)$ est de classe $\mathscr{C}^1$ sur $\mathbf{R}^*$, et sa dérivée vaut

$$\frac{1}{1+x^2} - \frac{1}{x^2}\frac{1}{1+\left(\frac{1}{x}\right)^2} = \frac{1}{1+x^2} - \frac{1}{1+x^2} = 0.$$

Il serait dès lors tentant de conclure que g est constante, mais malheureusement, $\mathbf{R}^*$ n'est pas un intervalle, et nous pouvons juste conclure que g est constante sur chacun des intervalles $]-\infty, 0[$ et $]0, +\infty[$.
Or, on a

$$\lim_{x\to-\infty} g(x) = -\frac{\pi}{2} \text{ et } \lim_{x\to+\infty} g(x) = \frac{\pi}{2}$$

ce qui permet d'affirmer que

$$g(x) = \begin{cases} \frac{\pi}{2} & \text{si } x > 0 \\ -\frac{\pi}{2} & \text{si } x < 0 \end{cases}$$

Ceci suffit à prouver que pour tout $x \in \mathbf{R}$, on a $F_X(x) = F_{1/X}(x)$. Et donc X et $\dfrac{1}{X}$ suivent la même loi.

□

Exercice 3.44 (QSP ESCP 2012) [ESCP12.QSP02] — Moyen

Deux variables de même loi.

Soit $f : x \mapsto \begin{cases} \dfrac{1}{\ln 2}\dfrac{1}{1+x} & \text{si } 0 \leqslant x \leqslant 1 \\ 0 & \text{sinon} \end{cases}$.

1. Montrer que f est une densité.
2. Soit X une variable aléatoire admettant X pour densité. Montrer que $Y = \frac{1}{X} - \left\lfloor \frac{1}{X} \right\rfloor$ a même loi que X.

Solution de l'exercice 3.44

1. Il est facile de voir que f est positive et continue sauf en 0 et en 1.
De plus, $\displaystyle\int_0^1 \frac{dx}{1+x} = [\ln(1+x)]_0^1 = \ln 2$, donc $\displaystyle\int_0^1 f(x)\,dx = 1$.
Ainsi, f est bien une densité.

2. Puisque pour tout réel x, on a $0 \leqslant x - \lfloor x \rfloor \leqslant 1$, Y est une variable aléatoire à valeurs dans $[0,1]$.
Donc $\forall x < 0, F_Y(x) = 0$ et $\forall x \geqslant 1, F_Y(x) = 1$.
De plus, X étant à valeurs dans $]0,1[$, $\frac{1}{X}$ est à valeurs dans $[1, +\infty[$. Soit donc $x \in [0,1]$.
Alors

$$F_Y(x) = P(Y \leqslant x) = P\left(0 \leqslant \frac{1}{X} - \left\lfloor \frac{1}{X} \right\rfloor \leqslant x\right)$$

Mais $\frac{1}{X}$ étant à valeurs dans $[1, +\infty[$, on a

$$\left[0 \leqslant \frac{1}{X} - \left\lfloor \frac{1}{X} \right\rfloor \leqslant x\right] = \bigcup_{n=1}^{+\infty} \left[n \leqslant \frac{1}{X} \leqslant n + x\right].$$

Détails
Il s'agit en fait de dire que $\left\{\left[\left\lfloor \frac{1}{X} \right\rfloor = n\right],\ n \in \mathbf{N}^*\right\}$ est un système complet d'événements.

De plus, ces événements sont deux à deux incompatibles, de sorte que

$$F_Y(x) = \sum_{n=1}^{+\infty} P\left(n \leqslant \frac{1}{X} \leqslant n+x\right) = \sum_{n=1}^{+\infty} P\left(\frac{1}{n+x} \leqslant X \leqslant \frac{1}{n}\right).$$

Mais pour $n \geqslant 1$, on a

$$P\left(\frac{1}{n+x} \leqslant X \leqslant \frac{1}{n}\right) = \int_{\frac{1}{n+x}}^{\frac{1}{n}} \frac{1}{\ln 2}\frac{1}{1+t}\,dt = \frac{1}{\ln 2}\left[\ln(1+t)\right]_{1/(n+x)}^{1/n} = \frac{1}{\ln 2}\left(\ln\left(1+\frac{1}{n}\right) - \ln\left(1+\frac{1}{n+x}\right)\right).$$

Il nous reste donc à calculer

$$\sum_{n=1}^{+\infty}\left(\ln\left(1+\frac{1}{n}\right) - \ln\left(1+\frac{1}{n+x}\right)\right).$$

Soit donc $N \geqslant 1$. On a

$$\begin{aligned}\sum_{n=1}^{N}\left(\ln\left(1+\frac{1}{n}\right) - \ln\left(1+\frac{1}{n+x}\right)\right) &= \sum_{n=1}^{N}\left(\ln\left(\frac{n+1}{n}\right) - \ln\left(\frac{n+1+x}{n+x}\right)\right)\\ &= \sum_{n=1}^{N}\left(\ln(n+1) - \ln(n) - \ln(n+1+x) + \ln(n+x)\right)\\ &= \ln(N+1) - \ln(N+1+x) + \ln(1+x)\\ &= \ln\left(\frac{N+1}{N+1+x}\right) + \ln(1+x).\end{aligned}$$

Lorsque $N \to +\infty$, on a $\dfrac{N+1}{N+1+x} \underset{N\to+\infty}{\longrightarrow} 1$ et donc $\ln\left(\dfrac{N+1}{N+1+x}\right) \underset{N\to+\infty}{\longrightarrow} 0$. On en déduit que

$$\sum_{n=1}^{+\infty}\left(\ln\left(1+\frac{1}{n}\right) - \ln\left(1+\frac{1}{n+x}\right)\right) = \ln(1+x)$$

et donc

$$F_Y(x) = P(Y \leqslant x) = \frac{\ln(1+x)}{\ln 2}.$$

Mais pour $x \in [0,1]$,

$$P(X \leqslant x) = \int_0^x \frac{1}{\ln 2}\frac{dt}{1+t} = \frac{\ln(1+x)}{\ln 2}.$$

Ainsi, X et Y ont les mêmes fonctions de répartition : elles ont même loi.

□

EXERCICE 3.45 (ESCP 2014) [ESCP14.3.14] Facile

Loi de l'arcsinus.

1. (a) Montrer que la fonction $]0, \frac{\pi}{2}[\to]0,1[$, $t \mapsto \sin(t)$ réalise une bijection de classe $\mathscr{C}^\infty$. On note h sa bijection réciproque.
   (b) Montrer que h est dérivable sur $]0,1[$ et exprimer $h'(x)$ en fonction de x.

On définit la fonction $f : \mathbf{R} \to \mathbf{R}$ par

$$\forall x \in \mathbf{R},\ f(x) = \begin{cases} \dfrac{1}{\sqrt{x(1-x)}} & \text{si } x \in]0,1[\\ 0 & \text{sinon} \end{cases}$$

2. (a) Soient a et b deux réels tels que $0 < a \leqslant b < 1$. Montrer, à l'aide du changement de variable $u = \sqrt{t}$, que
$$\int_a^b f(t)\,dt = 2\left(h(\sqrt{b}) - h(\sqrt{a})\right).$$
(b) Montrer la convergence et calculer la valeur de $\displaystyle\int_0^1 f(t)dt$.

3. (a) Déterminer la constante k telle que la fonction $g = kf$ soit une densité de probabilité.
Désormais, on note X une variable aléatoire définie sur un espace probabilisé $(\Omega, \mathcal{A}, P)$, admettant g pour densité.
(b) Déterminer la fonction de répartition G de X.
(c) Pour $x \in \mathbf{R}$, exprimer $G(1-x)$ en fonction de x. En déduire $P\left(X < \frac{1}{2}\right)$.

4. Déterminer la fonction de répartition et reconnaître la loi de la variable aléatoire $Y = h(\sqrt{X})$.

Solution de l'exercice 3.45

1.a. La fonction $t \mapsto \sin t$ est de classe $\mathcal{C}^\infty$ sur $]0, \frac{\pi}{2}[$, et sa dérivée est $t \mapsto \cos(t)$ qui est strictement positive sur $]0, \frac{\pi}{2}[$. Donc $t \mapsto \sin t$ est strictement croissante sur $]0, \frac{\pi}{2}[$.
De plus, on a $\lim\limits_{x\to 0^+} \sin t = 0$ et $\lim\limits_{x\to \frac{\pi}{2}^-} \sin t = 1$, donc par le théorème de la bijection, $t \mapsto \sin t$ réalise une bijection de $]0, \frac{\pi}{2}[$ sur $]0; 1[$.

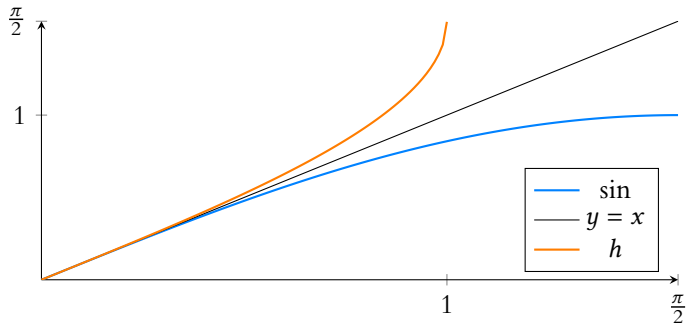


Figure 3.6 – Les courbes de sin et h sont symétriques par rapport à la première bissectrice $y = x$.

1.b. La fonction $t \mapsto \sin t$ est dérivable sur $]0, \frac{\pi}{2}[$, et nous avons déjà vu que sa dérivée ne s'annule pas sur $]0, \frac{\pi}{2}[$. Donc h est dérivable sur $]0, 1[$, et
$$\forall x \in]0,1[, h'(x) = \frac{1}{\cos(h(x))}.$$
Or, pour tout $x \in]0,1[$, $\cos^2(h(x)) + \underbrace{\sin^2(h(x))}_{=x^2} = 1 \Leftrightarrow \cos^2(h(x)) = 1 - x^2$.
Mais $h(x) \in]0, \frac{\pi}{2}[$, et donc $\cos(h(x)) \geqslant 0$. On en déduit donc que
$$\forall x \in]0,1[, h'(x) = \frac{1}{\sqrt{1-x^2}}.$$

Astuce
C'est un théorème du cours que f^{-1} est dérivable là où $f' \circ f$ ne s'annule pas et que
$$(f^{-1})' = \frac{1}{f' \circ f^{-1}}.$$
Cette formule peut se retrouver aisément en dérivant la relation
$$f \circ f^{-1}(x) = x.$$

2.a. La fonction $\varphi : u \mapsto u^2$ est de classe $\mathcal{C}^1$ sur $\mathbf{R}_+^*$, avec $\varphi'(u) = 2u$.
Ainsi, par le théorème de changement de variable[1] pour tous réels a et b tels que $0 < a < b < 1$, on a
$$\int_a^b f(t)\,dt = \int_{\sqrt{a}}^{\sqrt{b}} f(u^2)2u\,du = \int_{\sqrt{a}}^{\sqrt{b}} \frac{1}{u}\frac{1}{\sqrt{1-u^2}}2u\,du = 2\int_{\sqrt{a}}^{\sqrt{b}} h'(u)\,du = 2\left(h(\sqrt{b}) - h(\sqrt{a})\right).$$

[1] Sur un segment.

2.b. Soit $b \in]0,1[$ fixé. Alors
$$\lim_{a\to 0^+} \int_a^b f(t)\,dt = \lim_{a\to 0^+} 2\left(h(\sqrt{b}) - h(\sqrt{a})\right) = 2\left(h(\sqrt{b}) - \lim_{a\to 0^+} h(\sqrt{a})\right).$$

Mais $\lim\limits_{x\to 0^+} h(x) = 0$, et donc

$$\lim_{a\to 0^+} \int_a^b f(t)\,dt = 2h(\sqrt{b}).$$

Ainsi, $\int_0^b f(t)\,dt$ est convergente et $\int_0^b f(t)\,dt = 2h(\sqrt{b})$.

De même, on a $\lim\limits_{x\to 1^-} h(x) = \dfrac{\pi}{2}$, et donc

$$\lim_{b\to 1^-} \int_0^b f(t)\,dt = 2\frac{\pi}{2} = \pi.$$

On en déduit que $\int_0^1 f(t)\,dt$ converge et $\int_0^1 f(t)\,dt = \pi$.

Définition

Une fonction f est une densité de probabilité si :
- elle est positive sur $\mathbf{R}$
- elle est continue, **sauf éventuellement** en un nombre fini de points
- $\int_{-\infty}^{+\infty} f(t)\,dt$ converge et vaut 1.

3.a. La fonction f est positive sur $\mathbf{R}$, et elle y est continue sauf en 0 et en 1.

De plus, on a $\int_{-\infty}^{+\infty} f(t)\,dt = \pi$, donc si l'on pose $g = \frac{1}{\pi} f$, alors g est positive sur $\mathbf{R}$, continue sauf en 0 et en 1, et on a $\int_{-\infty}^{+\infty} g(t)\,dt = 1$.

Ainsi, pour $k = \frac{1}{\pi}$, $g = k \cdot f$ est une densité de probabilité.

Plus généralement

Pour toute fonction f positive, continue sauf en un nombre fini de points et telle que $\int_{-\infty}^{+\infty} f(t)\,dt$ converge, il existe une unique constante k telle que kf soit une densité de probabilités, et on a alors

$$\frac{1}{k} = \int_{-\infty}^{+\infty} f(t)\,dt.$$

3.b. Soit $x \leqslant 0$. Alors $G(x) = \int_{-\infty}^{x} g(t)\,dt = 0$ car g est nulle sur $]-\infty, x[$.

Si $x \in]0, 1]$, alors

$$G(x) = \int_{-\infty}^{x} g(t)\,dt = \int_0^x g(t)\,dt = \frac{2}{\pi} h(\sqrt{x}).$$

Enfin, si $x > 1$, alors $G(x) = \int_{-\infty}^{x} g(t)\,dt = \int_0^1 g(t)\,dt = 1$.

Par conséquent, la fonction de répartition de X est donnée par

$$G(x) = \begin{cases} 0 & \text{si } x \leqslant 0 \\ \frac{2}{\pi} h(\sqrt{x}) & \text{si } x \in]0,1[\\ 1 & \text{sinon} \end{cases}.$$

3.c. Il vient facilement $G(1-x) = 1$ si $x \leqslant 0$ et $G(1-x) = 0$ si $x \geqslant 1$. Pour $x \in]0,1[$, on a $G(1-x) = \int_0^{1-x} g(t)\,dt$.

Réalisons le changement de variables $u = 1 - t$ sur $]a, 1-x[$. Alors

$$\int_a^{1-x} g(t)\,dt = \int_x^{1-a} g(1-u)\,du = \int_x^{1-a} g(u)\,du = \frac{2}{\pi}\left(h(\sqrt{1-a}) - h(\sqrt{x})\right) \underset{a\to 0^+}{\longrightarrow} \frac{2}{\pi}\left(\frac{\pi}{2} - h(\sqrt{x})\right) = 1 - G(x).$$

Ainsi, pour tout $x \in \mathbf{R}$, on a $G(1-x) = 1 - G(x)$.

Puisque X est à densité, on a $P\left(X < \frac{1}{2}\right) = P\left(X \leqslant \frac{1}{2}\right) = G(1/2)$.

Mais $G(1/2) = 1 - G(1/2)$, et donc $G(1/2) = \frac{1}{2}$. On en déduit que

$$P\left(X < \frac{1}{2}\right) = \frac{1}{2}.$$

Rédaction

Commencer par déterminer le support de Y permet de déterminer directement où F_Y vaut 0 et où elle vaut 1.

4. X étant à support dans $]0,1[$, il en est de même de $\sqrt{X}$, et donc $Y = h(\sqrt{X})$ est à support dans $]0, \frac{\pi}{2}[$. Ainsi, pour $x \in]0, \frac{\pi}{2}[$, on a (par croissance de sin sur $]0, \frac{\pi}{2}[$),

$$P(Y \leqslant x) = P(h(\sqrt{X}) \leqslant x) = P(\sqrt{X} \leqslant \sin x) = P(X \leqslant \sin^2 x) = \frac{2}{\pi} h(\sin x) = \frac{2}{\pi} x.$$

Ainsi, la fonction de répartition de Y est donnée par $F_Y(x) = \begin{cases} 0 & \text{si } x \leqslant 0 \\ \frac{2}{\pi} x & \text{si } x \in]0, \frac{\pi}{2}[\\ 1 & \text{si } x \geqslant \frac{\pi}{2} \end{cases}$.

On reconnaît[2] alors la fonction de répartition d'une variable aléatoire suivant la loi uniforme sur $\left[0, \frac{\pi}{2}\right]$. Ainsi

$$Y \hookrightarrow \mathcal{U}\left(\left[0, \frac{\pi}{2}\right]\right).$$

□

[2] Remarque : une fonction de répartition qui est affine là où elle ne vaut pas 0 ou 1 est forcément la fonction de répartition d'une loi uniforme.

Exercice 3.46 (HEC 2018) [HEC18.247] Facile

Toutes les variables aléatoires introduites dans cet exercice sont définies sur un même espace probabilisé $(\Omega, \mathcal{A}, P)$.

1. Soit U une variable aléatoire à valeurs dans $]0,+\infty[$ et suivant la loi exponentielle de paramètre 1. On pose $X = U^{-1/\lambda}$.
   (a) Vérifier que la fonction de répartition de la variable aléatoire X est donnée par :
   $$\forall x \in \mathbf{R},\ F_X(x) = \begin{cases} e^{-x^{-\lambda}} & \text{si } x > 0 \\ 0 & \text{sinon} \end{cases}$$
   (b) En déduire que X est une variable à densité.
   (c) Proposer un script Scilab de simulation de la loi de X.
2. Soit $k \in \mathbf{N}^*$.
   (a) Démontrer que X possède un moment d'ordre k si et seulement k est strictement inférieur à λ et que ;
   $$\forall k < \lambda,\ E\left(X^k\right) = \Gamma\left(1 - \frac{k}{\lambda}\right).$$
   (b) En déduire que, pour tout réel $a > \dfrac{1}{2}$, on a l'inégalité stricte :
   $$(\Gamma(a))^2 < \Gamma(2a-1).$$
3. Soit Y une variable aléatoire positive, dont la fonction de répartition F_Y vérifie :
   $$1 - F_Y(x) \underset{x\to+\infty}{\sim} x^{-\lambda}.$$
   On considère une suite $(Y_n)_{n\in\mathbf{N}^*}$ de variables aléatoires indépendantes qui suivent toutes la loi de Y, et pour $n \in \mathbf{N}^*$, on pose $Z_n = n^{-1/\lambda}\max(Y_1,\dots,Y_n)$.
   Exprimer la fonction de répartition de Z_n à l'aide de celle de Y et en déduire que la suite (Z_n) converge en loi vers la variable X de la question 1.

Solution de l'exercice 3.46

1.a. Il est évident que $P(X \leqslant 0) = 0$ si $x \leqslant 0$, et pour $x > 0$, on a

$$P(X \leqslant x) = P\left(U^{-1/\lambda} \leqslant x\right) = P(U^{-1} \leqslant x^{\lambda}) = P\left(U \geqslant x^{-\lambda}\right) = 1 - P(\left(U \leqslant x^{-\lambda}\right) = e^{-x^{-\lambda}}.$$

1.b. Il est clair que F_X est $\mathscr{C}^1$, sauf peut-être en 0, et puisque

$$\lim_{x\to 0^+} F_X(x) = \lim_{x\to 0^+} e^{-x^{-\lambda}} = \lim_{u\to-\infty} e^u = 0 = \lim_{x\to 0^-} = F_X(x),$$

F_X est continue en 0 et donc sur $\mathbf{R}$.
Et par conséquent, X est une variable à densité.

1.c. Nous savons simuler aisément une loi exponentielle en utilisant la commande grand , et donc le script suivant convient :

```
1 lambda = input('Entrez lambda  :') ;
2 X = grand(1,1,'exp',1)^(-1/lambda)
```

Remarque

Nous saurions également simuler une loi exponentielle à partir de rand() en utilisant la méthode d'inversion !

2.a. La variable X possède un moment d'ordre k si et seulement si $E\left(X^{-k/\lambda}\right)$ existe.
D'après le théorème de transfert, c'est le cas si et seulement si $\displaystyle\int_0^{+\infty} t^{-k/\lambda}e^{-t}\,dt$ converge[1]
Nous reconnaissons là l'intégrale définissant $\Gamma\left(1-\dfrac{k}{\lambda}\right)$, qui converge si et seulement si $1 - \dfrac{k}{\lambda} > 0 \Leftrightarrow k < \lambda$.
Et lorsque c'est le cas, on a bien $E(X^k) = \Gamma\left(1-\dfrac{k}{\lambda}\right)$.

[1] Absolument, mais il s'agit de l'intégrale d'une fonction positive.

Strictement ?
A priori une variance est uniquement positive ou nulle, mais nous savons qu'une variable à densité ne peut avoir une variance nulle (seules les variables certaines ont une variance nulle).

2.b. Lorsque X possède une variance, c'est-à-dire pour $\lambda > 2$, la formule de Huygens nous donne

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

Et donc $\Gamma\left(1-\frac{1}{\lambda}\right)^2 < \Gamma\left(1-\frac{2}{\lambda}\right)$.
Soit donc $\alpha > \frac{1}{2}$, et soit $\lambda = \frac{1}{1-a} > 2$, de sorte que $a = 1 - \frac{1}{\lambda}$.
Alors $1 - \frac{2}{\lambda} = 1 - 2(1-a) = 2a - 1$ et donc l'inégalité précédente devient $\Gamma(a)^2 < \Gamma(2a-1)$.

3. Pour $x \in \mathbf{R}$, on a

$$P(Z_n \leqslant x) = P\left(\max(Y_1, \dots, Y_n) \leqslant n^{1/\lambda}x\right) = P\left(Y \leqslant n^{1/\lambda}x\right)^n = F_Y\left(n^{1/\lambda}x\right)^n.$$

Puisque Y est positive, lorsque $x < 0$, on a $P(Y \leqslant x) = 0$ et donc $P\left(Z_n \leqslant n^{1/\lambda}x\right) = 0$.
En revanche, pour $x \geqslant 0$, il vient donc $P(Z_n \leqslant x) = \left(1 - \left(1 - F_Y\left(n^{1/\lambda}x\right)\right)^n$, et puisque $n^{1/\lambda}x \underset{n\to+\infty}{\longrightarrow} +\infty$, alors $F_Y\left(n^{1/\lambda}x\right) \underset{n\to+\infty}{\longrightarrow} 1$, de sorte que $1 - F_Y\left(n^{1/\lambda}x\right) \underset{n\to+\infty}{\longrightarrow} 0$.
On a alors

$$F_{Z_n}(x) = \exp\left(n \ln\left(1 - \left(1 - F_Y\left(n^{1/\lambda}x\right)\right)\right)\right)$$

Mais $\ln\left(1 - \left(1 - F_Y\left(n^{1/\lambda}x\right)\right)\right) \underset{n\to+\infty}{\sim} F_Y\left(n^{1/\lambda}x\right)) - 1 \underset{n\to+\infty}{\sim} -\left(n^{1/\lambda}x\right)^{-\lambda} = -\frac{x^{-\lambda}}{n}$.
On en déduit que $n \ln\left(1 - \left(1 - F_Y\left(n^{1/\lambda}x\right)\right)\right) \underset{n\to+\infty}{\sim} -x^{-\lambda}$ et donc par continuité de l'exponentielle, $F_{Z_n}(x) \underset{n\to+\infty}{\longrightarrow} e^{-x^{-\lambda}}$.
Nous reconnaissons donc la fonction de répartition de la variable aléatoire X de la question 1, et donc $Z_n \xrightarrow{\mathcal{L}} X$.

□

3.5.4 Produit de convolution

EXERCICE 3.47 (ESCP 2008) [ESCP08.3.24] — Moyen

Somme des valeurs absolues de deux loi normales centrées réduites

Un joueur prend pour cible un mur muni d'un repère orthonormé (O, i, j). On note X et Y les variables aléatoires désignant respectivement l'abscisse et l'ordonnée du point d'impact du tir sur le mur, et on suppose que X et Y sont indépendantes, de même loi normale centrée réduite. On note Φ la fonction de répartition de X.

1. (a) Donner une densité de $|X|$.

   (b) Montrer que la variable aléatoire $Z = |X| + |Y|$ admet comme densité la fonction h définie par

$$h(x) = \begin{cases} \frac{2}{\pi} e^{-x^2/4} \int_{-x/2}^{x/2} e^{-v^2}\,dv & \text{si } x \geqslant 0 \\ 0 & \text{sinon} \end{cases}$$

2. (a) Déterminer la dérivée de la fonction φ définie sur $\mathbf{R}$ par : $\varphi(x) = \int_{-x/2}^{x/2} e^{-v^2}\,dv$.

   (b) Étudier les variations de φ et tracer l'allure de sa représentation graphique dans un repère orthonormé du plan.

3. On peint sur le mur un carré plein de sommets $(1,0), (0,1), (-1,0), (0,-1)$.
   On note p la probabilité pour que l'impact soit dans le carré. Exprimer p à l'aide de la variable aléatoire Z, et déterminer p en fonction de Φ.

SOLUTION DE L'EXERCICE 3.47

1.a. Il est évident que $|X|$ est à valeurs positives, et donc pour $x < 0$, $P(|X| \leqslant x) = 0$.
Soit à présent $x \geqslant 0$. On a alors

$$P(|X| \leqslant x) = P(-x \leqslant X \leqslant x) = \Phi(x) - \Phi(-x) = 2\Phi(x) - 1.$$

Et donc la fonction de répartition de $|X|$ est donnée par

$$F_{|X|}(x) = \begin{cases} 2\Phi(x) - 1 & \text{si } x \geqslant 0 \\ 0 & \text{si } x < 0 \end{cases}$$

Il est clair que $F_{|X|}$ est $\mathscr{C}^1$ sauf éventuellement en 0. De plus, $\lim\limits_{x\to 0^-} F_{|X|}(x) = 0 = \lim\limits_{x\to 0^+} F_{|X|}(x)$ et donc $F_{|X|}$ est continue en 0 et donc sur $\mathbf{R}$ tout entier.
On en déduit que $|X|$ est une variable à densité, et qu'une densité en est donnée par

$$f_{|X|}(x) = \begin{cases} 2\Phi'(x) & \text{si } x \geqslant 0 \\ 0 & \text{sinon} \end{cases} = \begin{cases} \sqrt{\dfrac{2}{\pi}} e^{-\frac{1}{2}x^2} & \text{si } x \geqslant 0 \\ 0 & \text{sinon} \end{cases}$$

1.b. Puisque $|X|$ et $|Y|$ sont indépendantes[1], et à densités bornées, une densité de Z est donnée par

[1] Car X et Y le sont.

$$f_Z(x) = \int_{-\infty}^{+\infty} f_{|X|}(t) f_{|Y|}(x-t)\, dt.$$

Mais $f_{|X|}(t) = 0$ si $t < 0$ et $f_{|Y|}(x-t) = 0$ si $x - t < 0 \Leftrightarrow t > x$. Ainsi, si $x < 0$, $f_Z(x) = 0$, et pour $x \geqslant 0$,

$$f_Z(x) = \int_0^x \sqrt{\frac{2}{\pi}} e^{-\frac{1}{2}t^2} \sqrt{\frac{2}{\pi}} e^{-\frac{1}{2}(x-t)^2}\, dt = \frac{2}{\pi} \int_0^x e^{-t^2+xt-x^2/2}\, dx = \frac{2}{\pi} e^{-\frac{x^2}{2}} \int_0^x e^{-(t^2-xt)}\, dt.$$

Notons que $t^2 - xt = \left(t - \dfrac{x}{2}\right)^2 - \dfrac{x^2}{4}$, de sorte que

$$f_Z(x) = \frac{2}{\pi} e^{-\frac{x^2}{4}} \int_0^x e^{-\left(t-\frac{x}{2}\right)^2}\, dt.$$

Procédons alors au changement de variable[2] $v = t - \dfrac{x}{2}$, de sorte que

[2] Affine.

$$f_Z(x) = \frac{2}{\pi} e^{-\frac{x^2}{4}} \int_{-x/2}^{x/2} e^{-v^2}\, dv.$$

2.a. D'après la relation de Chasles, on a $\varphi(x) = \displaystyle\int_0^{x/2} e^{-v^2}\, dv - \int_0^{-x/2} e^{-v^2}\, dv$.
Or, par le théorème fondamental de l'analyse, nous savons que $F : x \mapsto \displaystyle\int_0^x e^{-v^2}\, dv$ est une primitive de $x \mapsto e^{-x^2}$.
Et donc, $\varphi : x \mapsto F\left(\dfrac{x}{2}\right) - F\left(-\dfrac{x}{2}\right)$ est dérivable sur $\mathbf{R}$, et a pour dérivée $\varphi'(x) = \dfrac{1}{2} e^{-\frac{x^2}{4}} + \dfrac{1}{2} e^{-\frac{x^2}{4}} = e^{-\frac{x^2}{4}}$.

2.b. La fonction φ est donc croissante. De plus, elle est clairement impaire. Enfin, on a $\lim\limits_{x\to+\infty} \varphi(x) = \displaystyle\int_{-\infty}^{+\infty} e^{-v^2}\, dv$. Or, si N est une variable aléatoire suivant la loi normale $\mathcal{N}\left(0, \dfrac{1}{2}\right)$, une densité de N est donnée par $x \mapsto \dfrac{1}{\sqrt{\pi}} e^{-x^2}$ et donc

$$\int_{-\infty}^{\infty} \frac{1}{\sqrt{\pi}} e^{-x^2}\, dx = 1 \Leftrightarrow \int_{-\infty}^{+\infty} e^{-x^2}\, dx = \sqrt{\pi}.$$

Ainsi, $\lim\limits_{x\to+\infty} \varphi(x) = \sqrt{\pi}$ et donc[3], $\lim\limits_{x\to-\infty} \varphi(x) = -\sqrt{\pi}$.

[3] Par imparité.

3. Il s'agit de remarquer que le carré en question est $\{(x, y) \in \mathbf{R}^2 : |x| + |y| \leqslant 1\}$.
Et donc

$$p = P(|X| + |Y| \leqslant 1) = P(|Z| \leqslant 1) = \int_{-\infty}^{1} f_Z(t)\, dt = \int_0^1 \frac{2}{\pi} e^{-t^2/4} \varphi(t)\, dt.$$

Or, nous savons que $e^{-x^2/4}$ est la dérivée de φ, donc

$$p = \frac{2}{\pi}\int_0^1 \varphi'(t)\varphi(t)\,dt = \frac{2}{\pi}\left[\frac{1}{2}\varphi(t)^2\right]_0^1 = \frac{1}{\pi}\int_{-1/2}^{1/2} e^{-x^2}\,dx.$$

Procédons alors au changement de variable $x = \dfrac{t}{\sqrt{2}}$, de sorte que

$$\int_{-1/2}^{1/2} e^{-x^2}\,dx = \int_{-1/\sqrt{2}}^{1/\sqrt{2}} e^{-t^2/2}\frac{dt}{\sqrt{2}} = \frac{1}{\sqrt{2}}\left(\Phi\left(\frac{1}{\sqrt{2}}\right) - \Phi\left(-\frac{1}{\sqrt{2}}\right)\right) = 2\Phi\left(\frac{1}{\sqrt{2}}\right) - 1.$$

Il vient donc $p = \dfrac{1}{\pi}\left(2\Phi\left(\dfrac{1}{\sqrt{2}}\right) - 1\right)$.

□

EXERCICE 3.48 (HEC 2016) [HEC16.158] — Moyen

Différence des carrés de deux uniformes

1. Soit f la fonction définie sur $\mathbf{R}$ par : $\forall x \in \mathbf{R}$, $f(x) = \begin{cases} \frac{1}{2\sqrt{x}} & \text{si } 0 < x \leqslant 1 \\ 0 & \text{sinon} \end{cases}$
   (a) Montrer que f est une densité de probabilités.
   (b) Donner l'allure de la représentation graphique de la fonction de répartition d'une variable aléatoire réelle de densité f.
   (c) Trouver la loi de $Y = \sqrt{X}$ lorsque X est une variable aléatoire positive admettant f pour densité.
2. (a) Pour quelles valeurs réelles de s l'intégrale $\displaystyle\int_s^1 \frac{dx}{\sqrt{x(x-s)}}$ est-elle convergente ?
   (b) Calculer alors $\displaystyle\int_s^1 \frac{dx}{\sqrt{x(x-s)}}$ en utilisant le changement de variable $t = \sqrt{\dfrac{x-s}{x}}$.
3. On considère deux variables aléatoires X et Y, indépendantes, admettant chacune f pour densité.
   (a) Proposer une méthode de simulation en `Scilab` de la variable aléatoire $S = X - Y$.
   (b) Démontrer que S est une variable aléatoire à densité et en donner une densité continue sur $\mathbf{R}^*$.

SOLUTION DE L'EXERCICE 3.48

1.a. La fonction f est évidemment positive sur $\mathbf{R}$, et elle est continue sauf peut-être en 0 et en 1.

On a alors, sous réserve de convergence de cette intégrale, $\displaystyle\int_{-\infty}^{+\infty} f(t)\,dt = \int_0^1 \frac{dt}{2\sqrt{t}}$.

Mais on reconnaît alors une intégrale de Riemann convergente, et pour $A \in]0,1]$,

$$\int_A^1 \frac{dt}{2\sqrt{t}} = \left[\sqrt{t}\right]_A^1 = 1 - \sqrt{A} \underset{A\to 0^+}{\longrightarrow} 1.$$

On en déduit donc que $\displaystyle\int_{-\infty}^{+\infty} f(t)\,dt = 1$, et donc f est bien une densité de probabilités.

1.b. Notons F la fonction de répartition d'une variable aléatoire admettant f pour densité.

Alors $F(x) = 0$ si $x \leqslant 0$ et $F(x) = 1$ si $x \geqslant 1$.

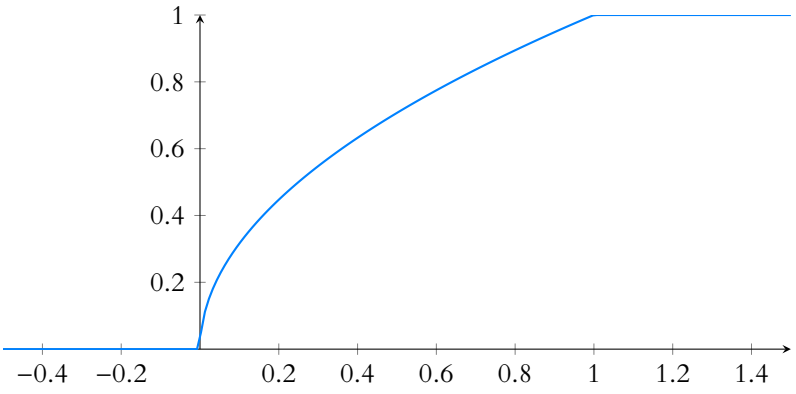
Pour $x \in]0,1]$, on a $F(x) = \displaystyle\int_{-\infty}^{x} f(t)\,dt = \int_0^x \frac{dt}{2\sqrt{t}}$.

Or, si $A \in]0,x]$,

$$\int_A^x \frac{dt}{2\sqrt{t}} = \left[\sqrt{t}\right]_A^x = \sqrt{x} - \sqrt{A} \underset{A\to 0^+}{\longrightarrow} \sqrt{x}.$$

Et donc $F(x) = \begin{cases} 0 & \text{si } x \leqslant 0 \\ \sqrt{x} & \text{si } 0 < x < 1 \\ 1 & \text{si } x \geqslant 1 \end{cases}$

Sa représentation graphique est alors donnée par



Figures

Un dessin à la main doit tout de même faire apparaître les caractéristiques essentielles de F.
Ici, il est légitime d'attendre du dessin qu'il fasse apparaître :

- la continuité de F
- sa croissance
- sa concavité sur $[0,1]$ (la dérivée y est décroissante)
- la tangente verticale en l'origine
- la non dérivabilité en 1

1.c. Puisque X est à valeurs dans $]0,1[$, il est évident que c'est également le cas de Y.
Et donc, on a $F_Y(x)=0$ si $x<0$ et $F_Y(x)=1$ si $x\geqslant 1$.
Pour $x\in]0,1[$, on a

$$F_Y(x)=P(Y\leqslant x)=P(\sqrt{X}\leqslant x)=P(X\leqslant x^2)=F(x^2)=\sqrt{x^2}=x.$$

Ainsi, $F_Y(x)=\begin{cases}0 & \text{si } x\leqslant 0\\ x & \text{si } 0<x\leqslant 1\\ 1 & \text{si } x>1\end{cases}$
On reconnaît là la fonction de répartition d'une loi uniforme sur $[0,1]$: $Y\hookrightarrow\mathcal{U}([0,1])$.

2.a. Si $s<0$, alors $x\mapsto\dfrac{1}{\sqrt{x(x-s)}}$ n'est pas définie sur $]s,0]$, et donc l'intégrale n'a pas de sens.
De même, si $s>1$, alors pour $x\in[1,s[$, on a $x(x-s)<0$, et donc l'intégrale n'a pas de sens.
Pour $s=0$, et $x\in]0,1]$, on a $\dfrac{1}{\sqrt{x(x-s)}}=\dfrac{1}{\sqrt{x^2}}=\dfrac{1}{x}$, dont l'intégrale entre 0 et 1 diverge.
Enfin, pour $s\in]0,1[$, la fonction $x\mapsto\dfrac{1}{\sqrt{x(x-s)}}$ est continue sur $]s,1]$, et au voisinage de s, on a $\dfrac{1}{\sqrt{x(x-s)}}\underset{x\to s}{\sim}\dfrac{1}{\sqrt{s}}\dfrac{1}{\sqrt{x-s}}$.
Or, $\displaystyle\int_s^1\frac{dx}{\sqrt{x-s}}$ converge, donc il en est de même de $\displaystyle\int_s^1\frac{dx}{\sqrt{x(x-s)}}$.

2.b. Posons $t=\sqrt{\dfrac{x-s}{x}}=\sqrt{1-\dfrac{s}{x}}$. Alors $x\mapsto\sqrt{1-\dfrac{s}{x}}$ est $\mathcal{C}^1$ et strictement croissante sur $]s,1]$, avec

$$\lim_{x\to s}\sqrt{1-\frac{s}{x}}=0 \text{ et } \lim_{x\to 1}\sqrt{1-\frac{s}{x}}=\sqrt{1-s}.$$

On a alors $t^2=1-\dfrac{x}{s}\Leftrightarrow x=\dfrac{s}{1-t^2}$, de sorte que $dx=\dfrac{2st}{(1-t^2)^2}$.
Et donc, par le théorème de changement de variable,

$$\int_s^1\frac{dx}{\sqrt{x(x-s)}}=\int_0^{\sqrt{1-s}}\frac{1}{\sqrt{\frac{s}{1-t^2}\left(\frac{s}{1-t^2}-s\right)}}\frac{2st}{(1-t^2)^2}=2\int_0^{\sqrt{1-s}}\frac{dt}{1-t^2}.$$

Or, $\dfrac{1}{1-t^2}=\dfrac{1}{(1-t)(1+t)}=\dfrac{1}{2}\left(\dfrac{1}{1-t}+\dfrac{1}{1+t}\right)$, de sorte que

$$\int_s^1\frac{dx}{\sqrt{x(x-s)}}=\int_0^{\sqrt{1-s}}\frac{dt}{1-t}+\int_0^{\sqrt{1-s}}\frac{dt}{1+t}=\left[-\ln(1-t)\right]_0^{\sqrt{1-s}}+\left[\ln(1+t)\right]_0^{\sqrt{1-s}}=\ln\left(\frac{1+\sqrt{1-s}}{1-\sqrt{1-s}}\right).$$

3.a. D'après la question 1.c, si X possède f pour densité, alors $\sqrt{X}$ suit une loi uniforme sur $[0,1]$.
Et donc si $U\hookrightarrow\mathcal{U}([0,1])$, alors U^2 possède f pour densité.
Donc on peut simuler S en utilisant

```
S = rand()^2 - rand()^2
```

3.b. Une densité de $-Y$ est $f_Y : t \mapsto \begin{cases} \frac{1}{2\sqrt{-t}} & \text{si } -1 \leqslant t < 0 \\ 0 & \text{sinon} \end{cases}$

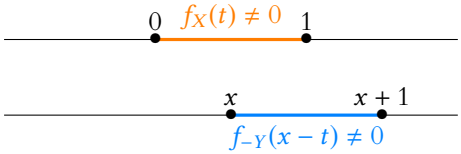
Puisque ni la densité de X ni celle de $-Y$ ne sont bornées, il faut prendre des précautions : le théorème sur le produit de convolution[1] nous dit que si la fonction g définie par

[1] Qui s'applique car X et $-Y$ sont indépendantes.

$$g(x) = \int_{-\infty}^{+\infty} f_X(t) f_{-Y}(x-t)\,dt$$

est bien définie et continue sauf éventuellement en un nombre fini de points, alors c'est une densité de S.
Notons qu'on a $f_X(t) \neq 0 \Leftrightarrow 0 < t \leqslant 1$ et de même

$$f_{-Y}(x-t) \neq 0 \Leftrightarrow -1 \leqslant x - t < 0 \Leftrightarrow x < t \leqslant x+1.$$



Donc déjà, pour $x \leqslant -1$ ou $x \geqslant 1$, on a $g(x) = 0$.

Pour $0 \leqslant x < 1$,

$$g(x) = \frac{1}{4}\int_x^1 \frac{dt}{\sqrt{t(t-x)}} = \frac{1}{4}\ln\left(\frac{1+\sqrt{1-x}}{1-\sqrt{1-x}}\right).$$

Pour $-1 < x < 0$, on a

$$g(x) = \int_0^{x+1} \frac{1}{2\sqrt{t}}\frac{1}{2\sqrt{t-x}}\,dt.$$

Et alors, en utilisant le changement de variable $u = t - x$, il vient alors

$$g(x) = \frac{1}{4}\int_{-x}^1 \frac{du}{\sqrt{u-(-x)}\sqrt{u}} = \frac{1}{4}\ln\left(\frac{1+\sqrt{1+x}}{1-\sqrt{1+x}}\right).$$

Remarque
$-x \in]0,1[$, et donc il est légitime d'utiliser le résultat de la question 2.b, qui n'était valable que pour $s \in]0,1[$.

Puisque l'intégrale définissant g est bien définie, et que la fonction g ainsi obtenue est continue sur $\mathbf{R}$, sauf éventuellement en $-1, 0$ et 1, nous pouvons affirmer que S est bien une variable à densité, et que g en est une densité.
Enfin, puisque l'énoncé nous demandait une densité continue sur $\mathbf{R}^*$, il nous reste juste à remarquer que g est continue en -1 et en 1 puisque

$$\lim_{x\to 1^-} \frac{1+\sqrt{1-x}}{1-\sqrt{1+x}} = 1$$

et donc $\lim\limits_{x\to 1^-} g(x) = 0 = \lim\limits_{x\to 1^+} g(x) = g(0)$.
Et de même en -1.
On a donc

$$g(x) = \begin{cases} 0 & \text{si } x < -1 \text{ ou } x > 1 \\ \frac{1}{4}\ln\left(\frac{1+\sqrt{1-x}}{1-\sqrt{1-x}}\right) & \text{si } 0 \leqslant x < 1 \\ \frac{1}{4}\ln\left(\frac{1+\sqrt{1+x}}{1-\sqrt{1+x}}\right) & \text{si } -1 < x < 0 \end{cases} = \begin{cases} \frac{1}{4}\ln\left(\frac{1+\sqrt{1-|x|}}{1-\sqrt{1-|x|}}\right) & \text{si } -1 < x < 1 \\ 0 & \text{sinon} \end{cases}$$

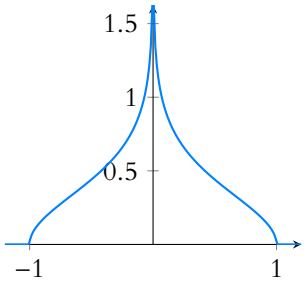


FIGURE 3.7– La densité g.

□

Exercice 3.49 (HEC 2015) [HEC15.139] Difficile

Distance entre deux points aléatoires

1. Soient a, b, α trois réels strictement positifs vérifiant $0 < \alpha < a^2 \leqslant b^2$.
   (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 $(\Omega, \mathcal{A}, P)$ 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 $P(D \leqslant \theta)$.

Solution de l'exercice 3.49

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, $f(t) \underset{t\to 0^+}{\sim} \frac{a}{\sqrt{t}}\left(\frac{b}{\sqrt{\alpha}} - 1\right)$.
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, $\left(\frac{a}{\sqrt{t}} - 1\right)\left(\frac{b}{\sqrt{\alpha - t}} - 1\right) \underset{t\to\alpha}{\sim} \left(\frac{a}{\sqrt{\alpha}} - 1\right)\frac{b}{\sqrt{\alpha - t}}$.
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.

> **Détails**
> -1 est négligeable devant $\frac{a}{\sqrt{t}}$ (qui tend vers $+\infty$), donc la somme des deux est équivalente au terme prépondérant : $\frac{a}{\sqrt{t}}$.

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

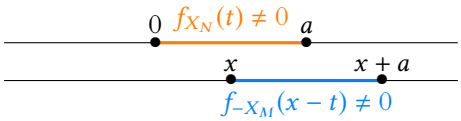
$$\begin{aligned}
\int_0^\alpha \left(\frac{a}{\sqrt{t}} - 1\right)\left(\frac{b}{\sqrt{\alpha - t}} - 1\right) dt &= -2\alpha \int_{\pi/2}^0 \left(\frac{a}{\sqrt{\alpha\cos^2 u}} - 1\right)\left(\frac{b}{\sqrt{\alpha(1 - \cos^2 u)}} - 1\right) \sin u \cos u\, du \\
&= 2\int_0^{\pi/2} (a - \sqrt{\alpha}\cos u)(b - \sqrt{\alpha}\sin u)\, du \\
&= 2\int_0^{\pi/2} ab\, du - 2a\sqrt{\alpha}\int_0^{\pi/2} \sin u\, du - 2b\sqrt{\alpha}\int_0^{\pi/2} \cos u\, du + 2\alpha\int_0^{\pi/2} \sin u \cos u\, du \\
&= \pi ab - 2\sqrt{\alpha}(a + b) + 2\alpha\left[\frac{\sin^2 u}{2}\right]_0^{\pi/2} \\
&= \pi ab - 2\sqrt{\alpha}(a + b) + \alpha.
\end{aligned}$$

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, $f_{X_N}(t) \neq 0 \Leftrightarrow 0 \leqslant t \leqslant a$ et $f_{-X_M}(x-t) \neq 0 \Leftrightarrow -a \leqslant x-t \leqslant 0 \Leftrightarrow x \leqslant t \leqslant x+a$.



Donc si $x \leqslant -a$, $f_{Z_a}(x) = 0$ et de même, si $x \geqslant a$, $f_{Z_a}(x) = 0$.
Si $-a \leqslant x \leqslant 0$, alors

$$f_{Z_a}(x) = \int_0^{x+a} \frac{1}{a}\frac{1}{a}\,dt = \frac{x+a}{a^2}.$$

Et si $0 \leqslant x \leqslant a$, alors

$$f_{Z_a}(x) = \int_x^{a} \frac{1}{a}\frac{1}{a}\,dt = \frac{a-x}{a^2}.$$

De même[1], on prouve qu'une densité de Z_b est donnée par

[1] En changeant les a en b.

$$f_{Z_b}(x) = \begin{cases} \frac{x+b}{b^2} & \text{si } -b \leqslant x \leqslant 0 \\ \frac{b-x}{b^2} & \text{si } 0 \leqslant x \leqslant b \\ 0 & \text{sinon} \end{cases} = \begin{cases} \frac{b-|x|}{b^2} & \text{si } -b \leqslant x \leqslant b \\ 0 & \text{sinon} \end{cases}$$

2.c. Puisque $Z_a(\Omega) = [-a, a]$, il vient $Z_a^2(\Omega) = [0, a^2]$.
Et donc $P(Z_a^2 \leqslant x) = 0$ si $x < 0$ et $P(Z_a^2 \leqslant x) = 1$ si $x \geqslant a^2$.
Si $x \in [0, a^2]$, alors

$$F_{Z_a^2}(x) = P(Z_a^2 \leqslant x) = P(-\sqrt{x} \leqslant Z_a \leqslant \sqrt{x}) = \int_{-\sqrt{x}}^{\sqrt{x}} f_{Z_a}(t)\,dt = 2\int_0^{\sqrt{x}} \frac{a-t}{a^2}\,dt = 2\left[\frac{t}{a} - \frac{t^2}{2a^2}\right]_0^{\sqrt{x}} = \frac{2\sqrt{x}}{a} - \frac{x}{a^2}.$$

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 $\mathscr{C}^1$ 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 coalitions[2], 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

[2] 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.

$$f_{D^2}(x) = \int_{-\infty}^{+\infty} f_{Z_a^2}(t) f_{Z_b^2}(x-t)\,dt = \int_0^{a^2} f_{Z_a^2}(t) f_{Z_b^2}(x-t)\,dt.$$

Notons que pour calculer $P(D \leqslant \theta) = P(0 \leqslant D^2 \leqslant \theta^2)$, nous n'avons besoin de connaître une densité de D^2 que sur l'intervalle $[0, \theta^2] \subset [0, a^2]$.
Or, pour $0 \leqslant x < a$, on a $f_{Z_a^2}(t) = \frac{1}{a^2}\left(\frac{a}{\sqrt{t}} - 1\right)$ et

$$f_{Z_b^2}(x-t) \neq 0 \Leftrightarrow 0 \leqslant x-t \leqslant b^2 \Leftrightarrow x - b^2 \leqslant t \leqslant x.$$

Donc, pour $x \in [0, a^2]$, on a

$$f_{D^2}(x) = \int_0^x \frac{1}{a^2}\left(\frac{a}{\sqrt{t}} - 1\right)\frac{1}{b^2}\left(\frac{b}{\sqrt{x-t}} - 1\right)dt = \frac{1}{a^2b^2} I_{a,b}(x).$$

Ainsi, il vient[3]

$$P(D \leqslant \theta) = P(D^2 \leqslant \theta^2) = \int_{-\infty}^{\theta^2} f_{D^2}(x)\,dx = \int_0^{\theta^2} \frac{1}{a^2b^2} I_{a,b}(x)\,dx.$$

En utilisant le résultat de la question 1, on a donc

$$\begin{aligned} P(D \leqslant \theta) &= \frac{1}{a^2b^2}\int_0^{\theta^2} \left(\pi ab - 2\sqrt{x}(a+b) + x\right)\,dx \\ &= \frac{\pi\theta^2}{ab} - \frac{4}{3}\frac{a+b}{a^2b^2}\theta^3 + \frac{\theta^4}{2a^2b^2}. \end{aligned}$$

□

[3] D ne prend que des valeurs positives, donc $[D \leqslant \theta] = [D^2 \leqslant \theta^2]$.

Exercice 3.50 (HEC 2014) [HEC14.116] Facile

Convergence en loi du produit de n lois uniformes sur $[0,1]$.

Soit $(X_n)_{n\in\mathbf{N}^*}$ une suite de variables aléatoires indépendantes, définies sur un espace probabilisé $(\Omega, \mathcal{A}, P)$ et de même loi uniforme sur l'intervalle $[0,1]$. On note F la fonction de répartition de X_1.

On pose pour tout $n \in \mathbf{N}^*$, $Z_n = \prod_{k=1}^{n} X_k$ et pour tout $k \in [\![1,n]\!]$, $Y_k = -\ln X_k$.

1. (a) Calculer pour tout $s \in \mathbf{N}$, $E\left(Z_n^s\right)$.
   (b) Quelle est la loi de Y_1 ?
   (c) En déduire la loi de $T_n = \sum_{k=1}^{n} Y_k$.
   (d) Déterminer une densité f_{Z_n} de la variable aléatoire Z_n.
2. Soit r un entier naturel et $z \in]0,1]$.
   (a) Établir la convergence de l'intégrale $\int_0^z (-\ln t)^r\,dt$.
   (b) À l'aide du changement de variable $y = -\ln t$ dont on justifiera la validité, montrer que l'on a :
   $$\frac{1}{r!}\int_{-\ln z}^{+\infty} y^r e^{-y}\,dy = z \times \sum_{k=0}^{r} \frac{(-\ln z)^k}{k!}.$$
   (c) En déduire la fonction de répartition F_{Z_n} de Z_n.
3. Étudier la convergence en loi de la suite de variables aléatoires $(Z_n)_{n\in\mathbf{N}^*}$.

Solution de l'exercice 3.50

1.a. Puisque les X_k sont indépendantes, il en est de même des X_k^s, et donc, sous réserve d'existence des $E\left(X_k^s\right)$, on a

$$E\left(Z_n^s\right) = E\left(\prod_{k=1}^{n} X_k^s\right) = \prod_{k=1}^{n} E\left(X_k^s\right) = \left(E(X_1^s)\right)^n.$$

Mais, par le théorème de transfert,

$$E\left(X_1^s\right) = \int_0^1 x^s\,dx = \frac{1}{s+1}.$$

Et donc $E\left(Z_n^s\right) = \dfrac{1}{(s+1)^k}$.

1.b. Puisque X_1 est à valeurs dans $]0,1]$, $Y_1 = -\ln X_1$ est à valeurs dans $\mathbf{R}_+$. Ainsi, pour $x < 0$, $P(Y_1 \leqslant x) = 0$. Et pour $x \geqslant 0$,

$$P(Y_1 \leqslant x) = P(-\ln(X_1) \leqslant x) = P(X_1 \geqslant e^{-x}) = 1 - P(X_1 < e^{-x}) = 1 - P(X_1 \leqslant e^{-x}) = 1 - e^{-x}.$$

Et donc la fonction de répartition de Y_1 est donnée par

$$F_{Y_1}(x) = \begin{cases} 0 & \text{si } x < 0 \\ 1 - e^{-x} & \text{si } x \geqslant 0 \end{cases}$$

Donc Y_1 suit la loi exponentielle $\mathcal{E}(1)$.

1.c. Puisque la loi $\mathcal{E}(1)$ est aussi la loi $\gamma(1)$, par stabilité des lois gamma[1], T_n suit la loi $\gamma(n)$.

[1] Les X_k sont indépendantes, donc les Y_k le sont aussi.

1.d. Il est évident que Z_n est à valeurs dans $]0,1]$, et donc $F_{Z_n}(x) = 0$ si $x \leqslant 0$ et $F_{Z_n}(x) = 1$ si $x \geqslant 1$.

D'autre part, on a $\ln(Z_n) = \sum_{k=1}^{n} \ln X_k = -\sum_{k=1}^{n} -\ln X_k = -T_n$, et donc, pour $x \in]0,1]$,

Croissance de l'exponentielle.

$$\begin{aligned} F_{Z_n}(x) &= P(Z_n \leqslant x) = P(\ln Z_n \leqslant \ln x) \\ &= P(T_n \geqslant -\ln x) \\ &= 1 - F_{T_n}(-\ln x) \end{aligned}$$

Détails

Nous connaissons une densité de T_n qui est continue sur $\mathbf{R}^*$, et donc F_{T_n} est $\mathcal{C}^1$, sauf éventuellement en 0.

Puisque T_n est à densité, F_{T_n} est continue sur $\mathbf{R}$ et $\mathcal{C}^1$ sauf éventuellement en un nombre fini de points.
Donc déjà, F_{Z_n} est continue sauf éventuellement en 0 et en 1.
Lorsque $x \to 0^+$, on a $\ln x \underset{x\to 0^+}{\longrightarrow} -\infty$ et donc $\lim\limits_{x\to 0^+} F_{Z_n}(x) = \lim\limits_{x\to 0^+} \left(1 - F_{T_n}(-\ln x)\right) = 0 = F_{Z_n}(0)$.
Puisque d'autre part, F_{Z_n} est continue à gauche[2] en 0, elle est continue en 0.
De même, lorsque $x \to 1^-$, $\lim\limits_{x\to 1^-} F_{Z_n}(x) = 1 - F_{T_n}(0)$.
Et puisque T_n est à valeurs positives, $F_{T_n}(0) = P(T_n \leqslant 0) = P(T_n = 0) = 0$.
Et donc F_{Z_n} est continue en 1, et donc continue sur $\mathbf{R}$.
De plus, par composition de fonctions $\mathcal{C}^1$, elle est $\mathcal{C}^1$ sauf éventuellement en 0 et en 1, de sorte que Z_n est une variable à densité.
Par dérivation d'une composée, une densité de Z_n est donc

[2] Elle est constante sur $\mathbf{R}_-$.

$$f_{Z_n} = \begin{cases} \dfrac{1}{x} f_{T_n}(-\ln x) & \text{si } x \in]0,1[\\ 0 & \text{sinon} \end{cases} = \begin{cases} \dfrac{1}{(n-1)!}(-\ln x)^{n-1} & \text{si } x \in]0,1[\\ 0 & \text{sinon} \end{cases}$$

2.a. Le problème de convergence de l'intégrale est au voisinage de 0.

Mais par croissances comparées, on a $(-\ln t)^r \underset{t\to 0^+}{=} o\left(\dfrac{1}{\sqrt{t}}\right)$.

Et donc, puisque l'intégrale[3] $\displaystyle\int_0^z \frac{1}{\sqrt{t}}\,dt$ converge, il en est de même de $\displaystyle\int_0^z (-\ln t)^r\,dt$.

[3] De Riemann

2.b. La fonction $t \mapsto -\ln t$ est $\mathcal{C}^1$ sur $]0,z]$, et elle y est strictement décroissante, donc le changement de variable est légitime.
On a alors, sous réserve de convergence,

$$\int_{-\ln z}^{+\infty} y^r e^{-y}\,dy = \int_0^z (-\ln t)^r t \frac{dt}{t} = \int_0^z (-\ln t)^r\,dt.$$

Or, nous avons prouvé à la question précédente que cette intégrale converge.
Montons donc par récurrence sur r que

$$\frac{1}{r!}\int_0^z (-\ln t)^r\,dt = z\sum_{k=0}^{r} \frac{(-\ln z)^k}{k!}.$$

Pour $k = 0$, c'est évident, puisque $\displaystyle\int_0^z (-\ln t)^0\,dt = z$.
Supposons donc que le résultat vrai au rang r. Une intégration par parties nous donne, pour tout $A \in]0,z]$,

$$\begin{aligned} \frac{1}{(r+1)!}\int_A^z (-\ln t)^{r+1}\,dt &= \frac{1}{(r+1)!}\left[t(-\ln t)^{r+1}\right]_A^z + \frac{1}{r!}\int_A^z (-\ln t)^r\,dt \\ &= \frac{z(-\ln z)^{r+1}}{(r+1)!} - \frac{A(-\ln A)^{r+1}}{(r+1)!} + \frac{1}{r!}\int_A^z (-\ln t)^r\,dt. \end{aligned}$$

Et donc après passage à la limite $A \to 0$, il vient

$$\frac{1}{(r+1)!}\int_0^z (-\ln t)^{r+1}\,dt = \frac{z(-\ln z)^{r+1}}{(r+1)!} + \frac{1}{r!}\int_0^z (-\ln t)^r\,dt = z\sum_{k=0}^{r+1}\frac{(-\ln z)^k}{k!}.$$

On en déduit donc que pour tout $r \in \mathbf{N}$,

$$\frac{1}{r!}\int_{-\ln z}^{+\infty} y^r e^{-y}\,dy = \frac{1}{r!}\int_0^z (-\ln t)^r\,dt = z\sum_{k=0}^{r}\frac{(-\ln z)^k}{k!}.$$

2.c. Nous savons déjà que $F_{Z_n}(x) = 0$ si $x \leqslant 0$ et $F_{Z_n}(x) = 1$ si $x \geqslant 1$.
Et pour $x \in]0,1[$, on a donc

$$F_{Z_n}(x) = P(Z_n \leqslant x) = 1 - P(T_n \geqslant -\ln x) = \frac{1}{(n-1)!}\int_{-\ln x}^{+\infty} y^{n-1}e^{-y}\,dy = x\times\sum_{k=0}^{n-1}\frac{(-\ln x)^k}{k!}.$$

3. Pour $x \in]0,1]$, on a

$$F_{Z_n}(x) = x\times\sum_{k=0}^{n-1}\frac{(-\ln x)^k}{k!} \underset{n\to+\infty}{\longrightarrow} x\times\sum_{k=0}^{+\infty}\frac{(-\ln x)^k}{k!} = xe^{-\ln x} = 1.$$

Et donc il vient

$$F_{Z_n}(x) \underset{n\to+\infty}{\longrightarrow} \begin{cases} 0 & \text{si } x \leqslant 0 \\ 1 & \text{si } x > 0 \end{cases}$$

Mais la fonction $F : x \mapsto \begin{cases} 0 & \text{si } x < 0 \\ 1 & \text{si } x \geqslant 0 \end{cases}$ est la fonction de répartition d'une variable aléatoire certaine égale à 0.
Elle est continue, sauf en 0, et donc en tout point x où F est continue, on a $\lim\limits_{n\to+\infty} F_{Z_n}(x) = F(x)$.
Et donc ceci prouve que $Z_n \xrightarrow{\mathcal{L}} 0$.
Remarque : *ce résultat n'est pas vraiment surprenant est était largement prévisible. En effet, Z_n est le produit de n nombres entre 0 et 1, et donc plus n est grand, plus la valeur de Z_n doit être petite. Notons que ceci reste une intuition, et qu'on ne peut se passer de calculs pour prouver ce résultat (un produit de n nombres entre 0 et 1 ne tend pas nécessairement vers 0 lorsque le nombre de termes tend vers $+\infty$, par exemple $\left(1-\frac{1}{n}\right)^n \underset{n\to+\infty}{\longrightarrow} e^{-1}$).*
Notons qu'il ne serait pas très difficile de prouver la convergence en probabilités de Z_n vers 0. Si on sait alors que la convergence en probabilités implique la convergence en loi , alors le résultat de cette question 3 s'obtient sans passer par ce qui précède.

□

Méthode
La limite ainsi obtenue n'est pas une fonction de répartition, car elle n'est pas continue à droite en 0.
Toutefois, en changeant sa valeur en 0, on obtient une fonction de répartition : celle d'une variable certaine.
Comme la convergence en loi ne regarde que les points de continuité, ce qui se passe en 0 n'a donc aucune importance.

3.6 Convergence des variables aléatoires

3.6.1 Inégalités de concentration : Markov, Bienaymé-Tchebychev et leurs variantes

Exercice 3.51 (QSP ESCP 2016) [ESCP16.QSP05] **Facile**

Majoration du reste d'une série exponentielle

On considère une variable aléatoire X définie sur un espace probabilisé $(\Omega, \mathcal{A}, P)$ 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 $P(X \geqslant n) \leqslant e^{-tn}E\left(e^{tX}\right)$.
3. En déduire que : $\sum_{k=n}^{+\infty}\frac{\lambda^k}{k!} \leqslant e^n\left(\frac{\lambda}{n}\right)^n$.

SOLUTION DE L'EXERCICE 3.51

1. Par le théorème de transfert, on a, sous réserve de convergence[1]

$$\begin{aligned} E\left(e^{tX}\right) &= \sum_{n=0}^{+\infty} e^{tn} P(X=n) \\ &= \sum_{n=0}^{+\infty} e^{tn} \frac{\lambda^n}{n!} e^{-\lambda} \\ &= e^{-\lambda} \sum_{n=0}^{+\infty} \left(e^t\right)^n \frac{\lambda^n}{n!} \\ &= e^{-\lambda} \exp\left(e^t \lambda\right) \\ &= \exp\left(\lambda(e^t - 1)\right). \end{aligned}$$

[1] Absolue, mais il s'agit d'une série à termes positifs.

On a reconnu une série exponentielle, donc convergente, donc $E\left(e^{tX}\right)$ existe.

2. Notons que $[X \geqslant n] = \left[e^{tX} \geqslant e^{tn}\right]$ et donc par l'inégalité de Markov appliquée à e^{tX}, qui est bien positive et admet une espérance,

$$P(X \geqslant n) = P\left(e^{tX} \geqslant e^{tn}\right) \leqslant \frac{E\left(e^{tX}\right)}{e^{tn}} \leqslant e^{-tn} E\left(e^{tX}\right).$$

3. En utilisant ce qui précède, il vient

$$P(X \geqslant n) \leqslant e^{-tn} \exp\left(\lambda(e^t - 1)\right).$$

Ceci étant vrai pour tout $t \geqslant 0$, cherchons la valeur de t qui minimise

$$e^{-tn} \exp\left(\lambda(e^t - 1)\right) = \exp\left(-tn + \lambda(e^t - 1)\right).$$

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

$$P(X \geqslant n) \leqslant \exp\left(-n\ln\left(\frac{n}{\lambda}\right) + n - \lambda\right) \leqslant \left(\frac{\lambda}{n}\right)^n e^n e^{-\lambda}.$$

Mais puisque d'autre part, X est à valeurs entières, on a

$$P(X \geqslant n) = \sum_{k=n}^{+\infty} P(X=k) = e^{-\lambda} \sum_{k=n}^{+\infty} \frac{\lambda^k}{k!}.$$

Et donc il vient, après division par $e^{-\lambda}$,

$$\sum_{k=n}^{+\infty} \frac{\lambda^k}{k!} \leqslant \left(\frac{\lambda}{n}\right)^n e^n.$$

□

3.6.2 Convergence en probabilités

EXERCICE 3.52 (QSP HEC 2007) [HEC07.QSP107] **Difficile**

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
- $\lim\limits_{n\to+\infty} E(X_n) = 1$
- $\lim\limits_{n\to+\infty} V(X_n) = +\infty$.

Solution de l'exercice 3.52 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 $P(X_n = 0) = 1 - \dfrac{1}{n}$.

On a alors, pour $\varepsilon > 0$,

$$P(|X_n - 0| \geqslant \varepsilon) \leqslant P(X_n \neq 0) \leqslant \frac{1}{n} \underset{n\to+\infty}{\longrightarrow} 0.$$

Et donc $X_n \xrightarrow{P} 0$.

Notons alors a_n l'autre valeur prise par X_n, de sorte que $P(X_n = a_n) = \dfrac{1}{n}$.

On a alors $E(X_n) = \dfrac{a_n}{n}$ et $E\left(X_n^2\right) = \dfrac{a_n^2}{n}$.

Par la formule de Huygens, on a $V(X_n) = E\left(X_n^2\right) - E(X_n)^2$, et donc si on arrive à avoir $E(X_n) \underset{n\to+\infty}{\longrightarrow} 1$ et $E\left(X_n\right)^2 \underset{n\to+\infty}{\longrightarrow} +\infty$, alors automatiquement, on aura $V(X_n) \underset{n\to+\infty}{\longrightarrow} +\infty$.

La condition $E(X_n) \underset{n\to+\infty}{\longrightarrow} 1$ est satisfaite si et seulement si $a_n \underset{n\to+\infty}{\sim} n$.
Prenons donc $a_n = n$. On a alors $E\left(X_n^2\right) = n \underset{n\to+\infty}{\longrightarrow} +\infty$.

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* $P(X_n = 0) = 1 - \dfrac{1}{n}$.

• *Il n'y a pas du tout unicité de la solution, par exemple, en gardant* $P(X_n = 0) = 1 - \dfrac{1}{n}$, *prendre n'importe quelle suite* (a_n) *équivalente à n pour la seconde valeur de* X_n *convenait. De même, nous aurions pu prendre* $P(X_n = 0) = 1 - \dfrac{1}{n^2}$ *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.

□

Exercice 3.53 (ESCP 2017) [ESCP17.3.15] Moyen

Convergence en probabilités de sommes et de produits.

Dans cet exercice, toutes les variables aléatoires sont définies sur le même espace probabilisé $(\Omega, \mathcal{A}, P)$. 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 $X_nY_n - XY = (X_n - X)Y + (Y_n - Y)X + (X_n - X)(Y_n - Y)$.

Soit $\varepsilon > 0$. On pose :

$$A_n = [|X_nY_n - XY| > \varepsilon]\,,\ B_n = [|X_n - X||Y| > \varepsilon/3]\,,\ C_n = [|Y_n - Y||X| > \varepsilon/3]\,,\ D_n = [|X_n - X||Y_n - Y| > \varepsilon/3]\,.$$

2. Montrer que $P(A_n) \leqslant P(B_n) + P(C_n) + P(D_n)$.
3. Soit $\delta > 0$.
   (a) Montrer qu'il existe un réel $A > 0$ tel que pour tout $t > A$, $P(|Y| > t) < \dfrac{\delta}{2}$.
   (b) Montrer que pour tout $t > 0$, $P(B_n) \leqslant P\left(|X_n - X| > \dfrac{\varepsilon}{3t}\right) + P(|Y| > t)$.
   (c) Soit $t > A$. Montrer qu'il existe un entier N_0 tel que pour tout $n \geqslant N_0$, $P\left(|X_n - X| > \dfrac{\varepsilon}{3t}\right) < \dfrac{\delta}{2}$.
   (d) En déduire que $\lim\limits_{n\to+\infty} P(B_n) = 0$.

On montrerait de même et on admet ici que $\lim\limits_{n\to+\infty} P(C_n) = 0$.

4. Montrer que $P(D_n) \leqslant P\left(|X_n - X| > \sqrt{\dfrac{\varepsilon}{3}}\right) + P\left(|Y_n - Y| > \sqrt{\dfrac{\varepsilon}{3}}\right)$.
   En déduire que $\lim\limits_{n\to+\infty} P(D_n) = 0$.
5. Montrer que la suite (X_nY_n) converge en probabilité vers XY.

SOLUTION DE L'EXERCICE 3.53

1. C'est un simple calcul...

2.a. Si on a à la fois $|X_n - X||Y| \leqslant \frac{\varepsilon}{3}$, $|Y_n - Y||X| \leqslant \frac{\varepsilon}{3}$ et $|X_n - X| \cdot |Y_n - Y| \leqslant \frac{\varepsilon}{3}$.
Alors, par l'inégalité triangulaire, on a

$$\begin{aligned}|X_nY_n - XY| &= |(X_n - X)Y + (Y_n - Y)X + (X_n - X)(Y_n - Y)| \\ &\leqslant |X_n - X||Y| + |Y_n - Y||X| + |X_n - X| \cdot |Y_n - Y| \\ &\leqslant \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} \leqslant \varepsilon.\end{aligned}$$

Autrement dit, nous venons de prouver que $\overline{B_n} \cap \overline{C_n} \cap \overline{D_n} \subset [|X_nY_n - XY| \leqslant \varepsilon]$.
En passant aux événements contraires, il vient

$$|X_nY_n - XY| > \varepsilon] \subset B_n \cup C_n \cup D_n.$$

Et donc

$$P\left(|X_nY_n - XY| > \varepsilon\right) \leqslant P(B_n \cup C_n \cup D_n) \leqslant P(B_n) + P(C_n) + P(D_n).$$

3.a.i. On a $P(|Y| > t) = 1 - P(|Y| \leqslant t)$.
Or, $|Y|$ est une variable aléatoire sur $(\Omega, \mathcal{A}, P)$, donc sa fonction de répartition tend vers 1 en $+\infty$ et donc $\lim\limits_{t\to+\infty} P(|Y| > t) = 0$. Et donc[1] il existe $A > 0$ tel que pour $t \geqslant A$, $g(x) < \dfrac{\delta}{2}$.
Et donc pour $t \geqslant A$, $0 \leqslant P(|Y| > t) \leqslant g(t) < \dfrac{\delta}{2}$.

[1] C'est la **définition** d'une limite.

3.a.ii. Comme à la question 2, on a

$$[|Y| \leqslant t] \cap \left[|X_n - X| \leqslant \frac{\varepsilon}{3t}\right] \subset \underbrace{\left[|X_n - X||Y| \leqslant \frac{\varepsilon}{3}\right]}_{=\overline{B_n}}.$$

Et donc

$$P(B_n) \leqslant P\left(|X_n - X| > \frac{\varepsilon}{3t}\right) + P(|Y| > t).$$

3.a.iii. C'est la définition de $X_n \xrightarrow{P} X$: puisque $\lim\limits_{n\to+\infty} P\left(|X_n - X| > \dfrac{\varepsilon}{3t}\right) = 0$, il existe donc $N_0 \in \mathbf{N}$ tel pour $n \geqslant N_0$, $P\left(|X_n - X| > \dfrac{\varepsilon}{3t}\right) < \dfrac{\delta}{2}$.

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 $n \in \mathbf{N}$, $P(B_n) \leqslant P\left(|X_n - X| > \dfrac{\varepsilon}{3t}\right) + \dfrac{\delta}{2}$.
Et en particulier, si N_0 est comme à la question 3.c, pour $n \geqslant N_0$, on a $0 \leqslant P(B_n) < \delta$.
Autrement dit, nous venons de prouver que pour tout $\delta > 0$, il existe $N_0 \in \mathbf{N}$ tel que pour $n \geqslant N_0$, $|P(B_n)| \leqslant \delta$.
Nous reconnaissons là la définition de $\lim\limits_{n\to+\infty} P(B_n) = 0$.

3.b. Toujours comme à la question 2, on a, par passage aux événements contraires,

$$\left[|X_n - X| \leqslant \sqrt{\frac{\varepsilon}{3}}\right] \cap \left[|Y_n - Y| \leqslant \sqrt{\frac{\varepsilon}{3}}\right] \subset \underbrace{\left[|X_n - X||Y_n - Y| \leqslant \frac{\varepsilon}{3}\right]}_{=\overline{D_n}}.$$

Et donc

$$P(D_n) \leqslant P\left(\left[|X_n - X| > \sqrt{\frac{\varepsilon}{3}}\right] \cup \left[|Y_n - Y| > \sqrt{\frac{\varepsilon}{3}}\right]\right) \leqslant P\left(\left[|X_n - X| > \sqrt{\frac{\varepsilon}{3}}\right]\right) + P\left(\left[|Y_n - Y| > \sqrt{\frac{\varepsilon}{3}}\right]\right).$$

Puisque $X_n \xrightarrow{P} X$, on a $\lim\limits_{n\to+\infty} P\left(\left[|X_n - X| > \sqrt{\frac{\varepsilon}{3}}\right]\right) = 0$.
Et de même, $Y_n \xrightarrow{P} Y$ et donc $\lim\limits_{n\to+\infty} P\left(\left[|Y_n - Y| > \sqrt{\frac{\varepsilon}{3}}\right]\right)$.
On en déduit que $\lim\limits_{n\to+\infty} P(D_n) = 0$.

3.c. D'après la question 2, et les résultats de 3.d et 4, on a

$$0 \leqslant P(A_n) \leqslant \underbrace{P(B_n) + P(C_n) + P(D_n)}_{\xrightarrow[n\to+\infty]{} 0}.$$

Et donc par le théorème des gendarmes, $\lim\limits_{n\to+\infty} P(A_n) = 0$.
Autrement dit, $X_n Y_n \xrightarrow{P} XY$.

□

3.6.3 Convergence en loi

Si la définition de la convergence en loi fait appel aux fonctions de répartitions, on n'oubliera pas que si les X_n et X sont des variables aléatoires à valeurs entières, alors $X_n \xrightarrow{\mathcal{L}} X$ si et seulement si $\forall k \in \mathbf{Z}, P(X_n = k) \underset{n\to+\infty}{\longrightarrow} P(X = k)$.

Exercice 3.54 (QSP HEC 2017) [HEC17.QSP102] **Facile**

Convergence en loi du minimum de lois géométriques.

Soit $(X_i)_{i\in\mathbf{N}^*}$ une suite de variables aléatoires indépendantes, définies sur un même espace probabilisé $(\Omega, \mathscr{A}, P)$ et soit $(p_i)_{i\in\mathbf{N}^*}$ une suite de réels de $]0,1[$ telle que pour tout $i \in \mathbf{N}^*$, X_i suit la loi géométrique de paramètre p_i. On pose, pour tout $i \in \mathbf{N}^*$, $q_i = 1 - p_i$.
On pose, pour tout $n \in \mathbf{N}^*$, $Z_n = \min(X_1, \ldots, X_n)$.

1. Calculer, pour tout $k \in \mathbf{N}^*$, la probabilité $P(Z_n \geqslant k)$. Quelle est la loi de Z_n ?
2. On suppose que pour tout $i \in \mathbf{N}^*$, on a $p_i = \dfrac{1}{(i+1)^2}$. Montrer que la suite de variables aléatoires $(Z_n)_{n\in\mathbf{N}^*}$ converge en loi vers une variable aléatoire dont on précisera la loi.

Solution de l'exercice 3.54

1. Pour $k \in \mathbf{N}^*$, on a

$$P(X_i \geqslant k) = \sum_{\ell=k}^{+\infty} P(X_i = \ell) = \sum_{\ell=k}^{+\infty} p_i q_i^{\ell-1} = p_i \sum_{j=0}^{+\infty} q_i^{j+k-1} = p_i q_i^{k-1} \sum_{j=0}^{+\infty} q_i^j = p_i q_i^{k-1} \frac{1}{1-q_i} = q_i^{k-1}.$$

Intuition
$P(X_i \geqslant k)$ est la probabilité d'avoir besoin d'au moins k essais avant l'obtention d'un premier succès.
C'est donc la probabilité d'avoir eu $k-1$ échecs lors des $k-1$ premiers essais, donc

Par indépendance des X_i, on a alors

$$P(Z_n \geqslant k) = P\left(\bigcap_{i=1}^{n}[X_i \geqslant k]\right) = \prod_{i=1}^{n} q_i^{k-1} = \left(\prod_{i=1}^{n} q_i\right)^{k-1}.$$

Et donc, Z_n étant à valeurs entières,

$$P(Z_n = k) = P(Z_n \geqslant k) - P(Z_n \geqslant k+1) = \left(\prod_{i=1}^{n} q_i\right)^{k-1} - \left(\prod_{i=1}^{n} q_i\right)^{k} = \left(\prod_{i=1}^{n} q_i\right)^{k-1}\left(1 - \prod_{i=1}^{n} q_i\right).$$

On reconnaît alors la loi géométrique de paramètre $1 - \prod_{i=1}^{n} q_i$.

2. Dans ce cas, on a alors, pour tout $n \in \mathbf{N}^*$,

$$\begin{aligned}\prod_{i=1}^{n} q_i &= \prod_{i=1}^{n}\left(1 - \frac{1}{(i+1)^2}\right) = \prod_{i=1}^{n}\frac{1-(i+1)^2}{(i+1)^2} = \prod_{i=1}^{n}\frac{i(i+2)}{(i+1)^2}\\ &= \frac{\prod_{i=1}^{n} i \prod_{i=1}^{n}(i+2)}{\left(\prod_{i=1}^{n}(i+1)\right)^2} = \frac{\prod_{i=1}^{n} i \prod_{k=3}^{n+2} k}{\prod_{j=2}^{n+1} j \prod_{j=2}^{n+1} j}\\ &= \frac{1}{n+1}\frac{n+2}{2} = \frac{n+2}{2(n+1)}.\end{aligned}$$

Et donc $\prod_{i=1}^{n} q_i \underset{n\to+\infty}{\longrightarrow} \frac{1}{2}$.
Ainsi, pour tout $k \in \mathbf{N}^*$, on a

$$P(Z_n = k) = \underbrace{\prod_{i=1}^{n} q_i}_{\underset{n\to+\infty}{\longrightarrow}\frac{1}{2}}\underbrace{\left(1 - \prod_{i=1}^{n} q_i\right)^{k-1}}_{\underset{n\to+\infty}{\longrightarrow}\frac{1}{2^{k-1}}} \underset{n\to+\infty}{\longrightarrow} \frac{1}{2^k}.$$

Mais si X suit une loi géométrique de paramètre $\frac{1}{2}$, alors pour tout $k \in \mathbf{N}^*$, $P(X = k) = \frac{1}{2^k}$.
Puisque les Z_n et X sont à valeurs dans $\mathbf{N}^*$, on en déduit que $Z_n \xrightarrow{\mathcal{L}} X$, où $X \hookrightarrow \mathcal{G}\left(\frac{1}{2}\right)$.

□

EXERCICE 3.55 (ESCP 2012) [ESCP12.3.3] — Facile

Loi d'Euler.

1. Soit g la fonction définie sur $\mathbf{R}$ par $g(x) = \dfrac{2}{\pi(e^x + e^{-x})}$.
   Montrer que g est une densité de probabilités (*on pourra utiliser le changement de variable* $t = e^x$.)

Dans la suite, on note X une variable aléatoire définie sur un espace probabilisé $(\Omega, \mathcal{A}, P)$ admettant g pour densité.

2. Déterminer la fonction de répartition F_X de X.
3. Montrer que X admet des moments de tous ordres et calculer son espérance.
4. On pose $Y = e^X$.
   (a) Montrer que Y est une variable aléatoire à densité et en déterminer une densité.
   (b) La variable aléatoire Y admet-elle une espérance ?
5. On considère une suite de variables aléatoires $(Y_n)_{n\in\mathbf{N}^*}$ définies sur le même espace probabilisé $(\Omega, \mathcal{A}, P)$, indépendantes, et suivant toutes la même loi que Y. Pour $n \in \mathbf{N}^*$, on pose $M_n = \max(Y_1, \ldots, Y_n)$.

(a) Déterminer la fonction de répartition de M_n.

(b) Pour $n \geqslant 1$, on pose $Z_n = \dfrac{n}{M_n}$. Montrer que la suite $(Z_n)_{n\in\mathbf{N}^*}$ converge en loi vers une variable aléatoire dont on donnera la loi.

Solution de l'exercice 3.55

1. Il est clair que g est positive et continue sur $\mathbf{R}$.
De plus, en utilisant le changement de variable $t = e^x$, qui est légitime car la fonction exponentielle est $\mathscr{C}^1$ et strictement croissante sur $\mathbf{R}$, on a $x = \ln t$ et donc $dx = \dfrac{dt}{t}$, de sorte que, sous réserve de convergence,

$$\int_{-\infty}^{+\infty} \frac{2}{\pi\left(e^x + e^{-x}\right)}\,dx = \int_0^{+\infty} \frac{2}{\pi\left(t+\frac{1}{t}\right)}\frac{dt}{t} = \frac{2}{\pi}\int_0^{+\infty}\frac{dt}{1+t^2}.$$

Chgt de variable

On n'oubliera pas de changer les bornes : lorsque $x \to -\infty$, $t = e^x \to 0$.

Or il est bien connu que cette dernière intégrale converge et vaut $\dfrac{\pi}{2}$, de sorte que $\displaystyle\int_{-\infty}^{+\infty} g(t)\,dt = 1$, et donc g est bien une densité de probabilités.

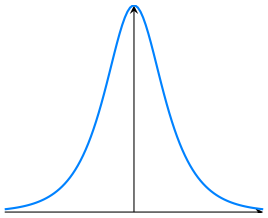


Figure 3.8– La densité g.

2. Soit $x \in \mathbf{R}$. Alors, à l'aide du même changement de variable que précédemment, on obtient

$$F_X(x) = P(X \leqslant x) = \int_{-\infty}^{x} g(t)\,dt = \frac{2}{\pi}\int_0^{e^x}\frac{1}{1+t^2}\,dt = \frac{2}{\pi}\mathrm{Arctan}\left(e^x\right).$$

3. Soit $k \in \mathbf{N}$. Par le théorème de transfert, on a, sous réserve de convergence,

$$E(X^k) = \int_{-\infty}^{+\infty}\frac{2t^k}{\pi\left(e^t+e^{-t}\right)}\,dt.$$

La fonction $t \mapsto \dfrac{t^k}{e^t+e^{-t}}$ est de même parité que k, et donc dans tous les cas, l'intégrale converge si et seulement si $\displaystyle\int_0^{+\infty}\frac{t^k}{e^t+e^{-t}}\,dt$ converge.
Mais au voisinage de $+\infty$, $\dfrac{t^k}{e^t+e^{-t}} \underset{t\to+\infty}{\sim} t^k e^{-t}$, et nous savons que $\displaystyle\int_0^{+\infty} t^k e^{-t}\,dt = \Gamma(k+1)$ converge.
Donc $\displaystyle\int_{-\infty}^{+\infty} t^k g(t)\,dt$ converge, et donc $E(X^k)$ existe.
En particulier, pour k impair, par imparité de la fonction $t \mapsto t^k g(t)$, on a $E(X^k) = 0$.
Et notamment, pour $k = 1$, $E(X) = 0$.

4.a. Il est évident que Y est à valeurs strictement positives, et donc pour $x \leqslant 0$, $P(Y \leqslant x) = 0$.
Pour $x > 0$, il vient, par croissance du logarithme,

$$P(Y \leqslant x) = P(e^X \leqslant x) = P(X \leqslant \ln x) = \frac{2}{\pi}\mathrm{Arctan}(x).$$

Ainsi, on a $F_Y(x) = \begin{cases} \frac{2}{\pi}\mathrm{Arctan}(x) & \text{si } x \geqslant 0 \\ 0 & \text{sinon} \end{cases}$
Cette fonction est clairement $\mathscr{C}^1$ sur $\mathbf{R}^*$.
Elle est continue en 0 car $\lim\limits_{x\to 0}\mathrm{Arctan}(x) = 0$.
Et donc Y est une variable à densité, dont une densité est

$$f_Y : t \mapsto \begin{cases} \frac{2}{\pi}\frac{1}{1+t^2} & \text{si } t \geqslant 0 \\ 0 & \text{sinon} \end{cases}$$

4.b. Au voisinage de $+\infty$, on a $t f_Y(t) \underset{t\to+\infty}{\sim} \dfrac{2}{\pi t}$, de sorte que $\displaystyle\int_{-\infty}^{+\infty} t f_Y(t)\,dt$ diverge, et donc Y n'admet pas d'espérance.

5.a. Par indépendance des Y_i, on a, pour $x \in \mathbf{R}$,

$$P(M_n \leqslant x) = P\left(\bigcap_{i=1}^{n}[Y_i \leqslant]\right) = \prod_{i=1}^{n} P(Y_i \leqslant x) = \begin{cases} \left(\frac{2}{\pi}\mathrm{Arctan}(x)\right)^n & \text{si } x \geqslant 0 \\ 0 & \text{sinon} \end{cases}$$

Remarque

Notons en particulier que M_n est une variable à densité, puisque cette fonction de répartition est continue sur $\mathbf{R}$ et $\mathscr{C}^1$ sur $\mathbf{R}^*$.

5.b. Il est évident que Z_n est à valeurs positives, et donc pour $x \leqslant 0$, $P(Z_n \leqslant x) = 0$.
Pour $x > 0$, on a

$$P(Z_n \leqslant x) = \left(\frac{n}{M_n} \leqslant x\right) = P\left(M_n \geqslant \frac{n}{x}\right) = 1 - P\left(M_n \leqslant \frac{n}{x}\right) = 1 - \left(\frac{2}{\pi}\text{Arctan}\left(\frac{n}{x}\right)\right)^n.$$

On a

$$\left(\frac{2}{\pi}\text{Arctan}\left(\frac{n}{x}\right)\right)^n = \exp\left(n\ln\left(\frac{2}{\pi}\text{Arctan}\left(\frac{n}{x}\right)\right)\right).$$

Or, il est classique que pour tout $x > 0$,

$$\text{Arctan}(x) + \text{Arctan}\left(\frac{1}{x}\right) = \frac{\pi}{2}.$$

Et donc en particulier,

$$\text{Arctan}\left(\frac{n}{x}\right) = \frac{\pi}{2} - \text{Arctan}\left(\frac{x}{n}\right).$$

Et donc $\ln\left(\frac{2}{\pi}\text{Arctan}\left(\frac{x}{n}\right)\right) = \ln\left(1 - \frac{2}{\pi}\text{Arctan}\left(\frac{x}{n}\right)\right)$.
Lorsque $n \to +\infty$, $\frac{x}{n} \underset{n\to+\infty}{\longrightarrow} 0$ et donc $\frac{2}{\pi}\text{Arctan}\left(\frac{x}{n}\right) \underset{n\to+\infty}{\longrightarrow} 0$. On a alors

$$\ln\left(1 - \frac{2}{\pi}\text{Arctan}\left(\frac{x}{n}\right)\right) \underset{n\to+\infty}{\sim} -\frac{2}{\pi}\text{Arctan}\left(\frac{x}{n}\right) \underset{n\to+\infty}{\sim} -\frac{2}{\pi}\frac{x}{n}.$$

Ainsi,

$$n\ln\left(\frac{2}{\pi}\text{Arctan}\left(\frac{x}{n}\right)\right) \underset{n\to+\infty}{\sim} -n\frac{2}{\pi}\frac{x}{n} \underset{n\to+\infty}{\longrightarrow} -\frac{2x}{\pi}.$$

Et donc par composition de limites,

$$P(Z_n \leqslant x) \underset{n\to+\infty}{\longrightarrow} 1 - e^{-\frac{2}{\pi}x}.$$

Et donc $Z_n \xrightarrow{\mathcal{L}} Z$, où $Z \hookrightarrow \mathcal{E}\left(\frac{2}{\pi}\right)$.

□

Classique ?

Disons que ce résultat est classique des sujets de parisiennes. Le candidat qui le connait aura le droit de l'utiliser tel quel.
Rappelons que pour le démontrer, il s'agit de dériver la fonction

$$x \mapsto \text{Arctan}(x) + \text{Arctan}\left(\frac{1}{x}\right)$$

sur $\mathbf{R}_+^*$ et de constater que la dérivée étant nulle, cette fonction est constante.

EXERCICE 3.56 (ESCP 2017) [ESCP17.3.11] — Difficile

Convergence en loi dans le problème des anniversaires.

Dans tout l'exercice, d désigne un entier naturel tel que $d \geqslant 2$.

1. Soit x un réel strictement positif. Déterminer $\lim\limits_{d\to+\infty} \frac{1}{d}\sum\limits_{k=1}^{\lfloor x\sqrt{d}\rfloor} k$.

2. (a) Montrer que la fonction f définie sur $]-1,1[\backslash\{0\}$ par $f(x) = \frac{\ln(1-x)+x}{x^2}$ admet un prolongement par continuité en 0.

   (b) Déterminer $\lim\limits_{d\to+\infty} \sum\limits_{k=1}^{\lfloor x\sqrt{d}\rfloor} \ln\left(1 - \frac{k}{d}\right)$.

Soit $(U_n)_{n\geqslant 1}$ une suite de variables aléatoires indépendantes définies sur le même espace probabilisé $(\Omega, \mathcal{A}, P)$ suivant toutes la loi loi uniforme discrète sur $[\![1, d]\!]$.
Soit N_d la variable aléatoire définie sur $(\Omega, \mathcal{A}, P)$ par

$$\forall \omega \in \Omega,\ N_d(\omega) = \min\{i \geqslant 2 \text{ tel que } U_i(\omega) \in \{U_1(\omega), \ldots, U_{i-1}(\omega)\}\}.$$

3. (a) Montrer que pour tout $n \in \mathbf{N}$ tel que $2 \leqslant n \leqslant d$, on a

$$P(N_d > n) = \left(1 - \frac{1}{d}\right)\left(1 - \frac{2}{d}\right)\cdots\left(1 - \frac{n-1}{d}\right).$$

(b) Déterminer la limite en loi de $\left(\frac{N_d}{\sqrt{d}}\right)$ lorsque d tend vers l'infini.

4. (a) Montrer que $E(N_d) = \sum_{k=0}^{d} P(N_d > k)$.

(b) En déduire que $E(N_d) = \int_0^{+\infty} \left(1 + \frac{t}{d}\right)^d e^{-t}\, dt$.

Solution de l'exercice 3.56

1. Une formule bien connue[1] nous dit que $\sum_{k=1}^{\lfloor x\sqrt{d}\rfloor} k = \frac{\left(\lfloor x\sqrt{d}\rfloor\right)\left(\lfloor x\sqrt{d}\rfloor + 1\right)}{2}$.

[1] $\lfloor x\sqrt{d}\rfloor$ est un entier comme un autre !

Mais, puisque $x\sqrt{d} - 1 < \lfloor x\sqrt{d}\rfloor \leqslant x\sqrt{d}$, on a $1 - \frac{1}{x\sqrt{d}} \leqslant \frac{\lfloor x\sqrt{d}\rfloor}{x\sqrt{d}} \leqslant 1$ et donc par le théorème des gendarmes,

$$\lim_{d\to+\infty} \frac{\lfloor x\sqrt{d}\rfloor}{x\sqrt{d}} = 1 \Leftrightarrow \lfloor x\sqrt{d}\rfloor \underset{d\to+\infty}{\sim} x\sqrt{d}.$$

On en déduit que

$$\frac{1}{d}\sum_{k=1}^{\lfloor x\sqrt{d}\rfloor} k \underset{d\to+\infty}{\sim} \frac{1}{d}\frac{\left(x\sqrt{d}\right)^2}{2} \underset{d\to+\infty}{\sim} \frac{x^2}{2} \underset{d\to+\infty}{\longrightarrow} \frac{x^2}{2}.$$

2.a. Au voisinage de 0, $\ln(1-x) + x = -x - \frac{x^2}{2} + \underset{x\to0}{o}\left(x^2\right) + x = -\frac{x^2}{2} + \underset{x\to0}{o}\left(x^2\right) \underset{x\to0}{\sim} -\frac{x^2}{2}$.

Et donc $\frac{\ln(1-x)+x}{x^2} \underset{x\to0}{\sim} -\frac{1}{2} \underset{x\to0}{\longrightarrow} -\frac{1}{2}$, de sorte que f est prolongeable par continuité en 0 en posant $f(0) = -\frac{1}{2}$.

2.b. Puisque f est continue en 0, il existe $\delta > 0$ tel que $|x| \leqslant \delta \Rightarrow |f(x) - f(0)| \leqslant 1$. Soit encore

$$\left|\frac{\ln(1-x)+x}{x^2} + \frac{1}{2}\right| \leqslant 1 \Leftrightarrow \left|\ln(1-x) + x + \frac{x^2}{2}\right| \leqslant x^2.$$

Mais pour $1 \leqslant k \leqslant \lfloor x\sqrt{d}\rfloor$, on a $0 \leqslant \frac{k}{d} \leqslant \frac{x\sqrt{d}}{d}$, et $\lim_{d\to+\infty} \frac{x\sqrt{d}}{d} = 0$, de sorte que pour d suffisamment grand, et pour tout $k \in [\![1, \lfloor x\sqrt{d}\rfloor]\!]$, $\left|\frac{k}{d}\right| \leqslant \delta$.

Il vient donc, pour tout $k \in [\![1, \lfloor x\sqrt{d}\rfloor]\!]$,

$$\left|\ln\left(1 - \frac{k}{d}\right) + \frac{k}{d} + \frac{k^2}{2d^2}\right| \leqslant \frac{k^2}{d^2}.$$

Et donc

$$\left|\sum_{k=1}^{\lfloor x\sqrt{d}\rfloor} \ln\left(1 - \frac{k}{d}\right) + \sum_{k=1}^{\lfloor x\sqrt{d}\rfloor} \frac{k}{d} + \frac{1}{2}\sum_{k=1}^{\lfloor x\sqrt{d}\rfloor} \frac{k^2}{d^2}\right| \leqslant \sum_{k=1}^{\lfloor x\sqrt{d}\rfloor} \left|\ln\left(1 - \frac{k}{d}\right) + \frac{k}{d} + \frac{1}{2}\frac{k^2}{d^2}\right| \leqslant \sum_{k=1}^{\lfloor x\sqrt{d}\rfloor} \frac{k^2}{d^2}.$$

Inégalité triangulaire.

Et par conséquent, pour d suffisamment grand,

$$-\frac{3}{2}\sum_{k=1}^{\lfloor x\sqrt{d}\rfloor} \frac{k^2}{d^2} \leqslant \sum_{k=1}^{\lfloor x\sqrt{d}\rfloor} \ln\left(1 - \frac{k}{d}\right) + \sum_{k=1}^{\lfloor x\sqrt{d}\rfloor} \frac{k}{d} \leqslant \frac{1}{2}\sum_{k=1}^{\lfloor x\sqrt{d}\rfloor} \frac{k^2}{d^2}.$$

Mais à l'aide des mêmes méthodes qu'à la question 1, on prouve que

$$\frac{1}{d^2}\sum_{k=1}^{\lfloor x\sqrt{d}\rfloor} k^2 = \frac{\left(\lfloor x\sqrt{d}\rfloor\right)\left(\lfloor x\sqrt{d}\rfloor + 1\right)\left(2\lfloor x\sqrt{d}\rfloor + 1\right)}{6d^2} \underset{d\to+\infty}{\sim} \frac{x^3}{3\sqrt{d}}.$$

Et donc en particulier, cette somme tend vers 0.

On en déduit donc que $\sum_{k=1}^{\lfloor x\sqrt{d}\rfloor} \ln\left(1-\frac{k}{d}\right) + \sum_{k=1}^{\lfloor x\sqrt{d}\rfloor} \frac{k}{d} \underset{d\to+\infty}{\longrightarrow} 0$.

Et donc en utilisant le résultat de la question 1, il vient donc

$$\lim_{d\to+\infty} \sum_{k=1}^{\lfloor x\sqrt{d}\rfloor} \ln\left(1-\frac{k}{d}\right) = -\frac{x^2}{2}.$$

3.a. Notons que la variable aléatoire prend comme valeur le numéro du premier tirage qui donne un numéro déjà sorti.

On a donc $[N_d > n]$ si et seulement si les n premiers tirages ont tous donné des numéros distincts.

Notons que $[N_d > n] = [N_d > 1] \cap [N_d > 2] \cap \cdots \cap [N_d > n]$. Ceci nous permet d'utiliser alors la formule des probabilités totales :

$$\begin{aligned} P(N_d > i) &= P\left([N_d > n] = [N_d > 1] \cap [N_d > 2] \cap \cdots \cap [N_d > n]\right) \\ &= P(N_d > 1) P_{[N_d>1]}(N_d > 2) P_{[N_d>1]\cap[N_d>2]}(N_d > 3) \cdots P_{[N_d>1]\cap\cdots\cap[N_d>n-1]}(N_d > n) \\ &= P(N_d > 1) P_{[N_d>1]}(N_d > 2) P_{[N_d>2]}(N_d > 3) \cdots P_{[N_d>n-1]}(N_d > n) \\ &= 1 \times \left(1-\frac{1}{d}\right)\left(1-\frac{2}{d}\right)\cdots\left(1-\frac{n-1}{d}\right). \end{aligned}$$

Astuce

L'intersection que nous venons d'écrire est absolument triviale puisque pour tout i, $[N_d > i] \subset [N_d > i-1]$. Toutefois, elle permet de faire apparaître les probabilités conditionnelles $P_{[N_d>i-1]}(N_d > i)$, qui sont faciles à calculer : il s'agit de la probabilité que U_i ne prenne pas l'une des $i-1$ valeurs différentes déjà prises par $U_1, \ldots, U_{i-1}$: c'est donc $1 - \frac{i-1}{d}$.

3.b. Notons que N_d prend ses valeurs dans $\mathbf{R}+$ et donc $\frac{N_d}{\sqrt{d}}$ prend également ses valeurs dans $\mathbf{R}_+$.

Pour $x \geqslant 0$, on a alors

$$P\left(\frac{N_d}{\sqrt{d}} \leqslant x\right) = P(N_d \leqslant x\sqrt{d}) = 1 - P(N_d > x\sqrt{d}).$$

Et N_d étant à valeurs entières, ceci est égal à $1 - P(N_d > \lfloor x\sqrt{d}\rfloor)$.

Mais puisque $x\sqrt{d} = \underset{d\to+\infty}{o}(d)$, pour x fixé, et pour d suffisamment grand, $2 \leqslant x\sqrt{d} \leqslant d$, de sorte qu'on peut utiliser le résultat de la question précédente :

$$P\left(\frac{N_d}{\sqrt{d}} \leqslant x\right) = 1 - \left(1-\frac{1}{d}\right)\cdots\left(1-\frac{1}{d}\right)\left(1-\frac{2}{d}\right)\cdots\left(1-\frac{\lfloor x\sqrt{d}\rfloor - 1}{d}\right).$$

Pour déterminer la limite de ce produit, passons plutôt aux logarithmes, pour faire apparaître des sommes :

$$\lim_{d\to+\infty} \sum_{k=1}^{\lfloor x\sqrt{d}\rfloor - 1} \ln\left(1-\frac{k}{d}\right) = \lim_{d\to+\infty} \left(\sum_{k=1}^{\lfloor x\sqrt{d}\rfloor} \ln\left(1-\frac{k}{d}\right) - \underbrace{\ln\left(1-\frac{\lfloor x\sqrt{d}\rfloor}{d}\right)}_{\underset{d\to+\infty}{\longrightarrow} 0} \right) = -\frac{x^2}{2}.$$

Et donc par continuité de l'exponentielle,

$$\lim_{d\to+\infty} \left(1-\frac{1}{d}\right)\left(1-\frac{2}{d}\right)\cdots\left(1-\frac{\lfloor x\sqrt{d}\rfloor - 1}{d}\right) = e^{-x^2/2}.$$

Et par conséquent,

$$\lim_{d\to+\infty} P\left(\frac{N_d}{\sqrt{d}} \leqslant x\right) = 1 - e^{-x^2/2}.$$

On a donc

$$\lim_{d\to+\infty} P\left(\frac{N_d}{\sqrt{d}} \leqslant x\right) = \begin{cases} 0 & \text{si } x < 0 \\ 1 - e^{-x^2/2} & \text{si } x \geqslant 0 \end{cases}$$

On vérifie aisément qu'il s'agit d'une fonction de répartition, d'une variable à densité admettant $g : x \mapsto \begin{cases} 0 & \text{si } x < 0 \\ xe^{-x^2/2} & \text{si } x \geqslant 0 \end{cases}$.

Et donc $\left(\frac{N_d}{\sqrt{d}}\right)$ converge en loi vers toute variable admettant g pour densité.

Remarque

On pourrait en fait faire beaucoup mieux en disant que N_d prend ses valeurs dans $[\![2, d+1]\!]$.

Pour la culture

Cette loi est appelée loi de Rayleigh. C'est par exemple la racine carrée d'une loi exponentielle $\mathcal{E}(1/2)$.

4. Puisque $N_d(\Omega) = [\![2, d]\!]$, on a

$$\sum_{k=0}^{d} P(N_d > k) = \sum_{k=0}^{d} \sum_{i=k+1}^{d} P(N_d = i) = \sum_{i=1}^{d} P(N_d = i) \sum_{k=0}^{i-1} = \sum_{i=1}^{d} P(N_d = i)i = E(N_d).$$

5. Notons que l'expression de $P(N_d > n)$ obtenue à la question 3.a peut encore s'écrire

$$P(N_d > n) = \frac{d-1}{d}\frac{d-2}{d} \cdots \frac{d-n+1}{d} = \frac{d!}{(d-n)!d^n}.$$

On a alors

$$\int_0^{+\infty} \left(1 + \frac{t}{d}\right)^d e^{-t}\, dt = \sum_{n=0}^{d} \binom{d}{n} \frac{1}{d^n} \int_0^{+\infty} t^n e^{-t}\, dt = \sum_{n=0}^{d} \binom{d}{n} \frac{n!}{d^n} = \sum_{n=0}^{d} P(N_d > n) = E(N_d).$$

Quelques commentaires *: un exercice relativement classique est que dans une classe de plus de 23 élèves, il y a plus d'une chance sur deux que deux élèves aient leur anniversaire le même jour. Pour faire le lien avec cet exercice : pour $d = 365$, pour quelle valeur de n a-t-on $P(N_d > n) \leqslant \frac{1}{2}$? La valeur 23 peut s'obtenir facilement à l'aide d'un programme qui testerait une à une toutes les valeurs de n jusqu'à obtenir $P(N_d > n) \leqslant \frac{1}{2}$.*

Mais si on approxime[2] $\frac{N_{365}}{\sqrt{365}}$ *par une loi de Rayleigh, on a alors*

$$P(N_{365} > n) \leqslant \frac{1}{2} \Leftrightarrow P\left(\frac{N_{365}}{\sqrt{365}} > \frac{n}{\sqrt{365}}\right) \leqslant \frac{1}{2} \Leftrightarrow 1 - e^{-x^2/(2\times 365)} \leqslant \frac{1}{2}$$

ce qui permet aisément de retrouver $n \geqslant 23$.
De même, combien d'élèves faut-il dans une classe pour qu'il y ait 9 chances sur 10 que deux aient la même date d'anniversaire ? □

[2] On peut considérer que 365 est suffisamment grand pour que cette approximation soit satisfaisante, mais aucun outil théorique ne nous le garantit.

Exercice 3.57 (HEC 2016) [HEC16.189] — Difficile

Convergence en loi et lois γ.

On considère un espace probabilisé $(\Omega, \mathcal{A}, P)$ et Z une variable aléatoire définie sur cet espace suivant la loi normale centrée réduite.
Pour tout réel $\nu > 0$, S_ν désigne une variable aléatoire sur $(\Omega, \mathcal{A}, P)$ suivant la loi $\gamma(\nu)$.

1. Soit $(\nu_n)_{n\in\mathbf{N}^*}$ une suite de nombres réels strictement positifs, strictement croissante et non majorée.
   Justifier la convergence et trouver la limite de la suite $\left(\frac{\lfloor \nu_n \rfloor}{\nu_n}\right)_{n\in\mathbf{N}^*}$.
2. On considère une suite $(X_n)_{n\in\mathbf{N}^*}$ de variables aléatoires définies sur $(\Omega, \mathcal{A}, P)$, mutuellement indépendantes et suivant chacune la loi exponentielle de paramètre 1, et pour tout $n \in \mathbf{N}^*$, on note $Z_n = \frac{1}{\sqrt{n}} \sum_{k=1}^{n} (X_k - 1)$.
   (a) Justifier la convergence en loi de la suite $(Z_n)_{n\in\mathbf{N}^*}$ vers Z.
   (b) En déduire la convergence en loi de la suite des variables aléatoires $\frac{1}{\sqrt{\nu_n}}(S_{\lfloor \nu_n \rfloor} - \lfloor \nu_n \rfloor)$.
3. Pour tout $n \in \mathbf{N}^*$, on note $\delta_n = \nu_n - \lfloor \nu_n \rfloor$.
   (a) Justifier, pour tout $x \geqslant 0$, l'inégalité $P(S_{\delta_n} \geqslant x) \leqslant P(S_1 \geqslant x)$.
   (b) Montrer que la suite des variables aléatoires $\frac{S_{\delta_n}}{\sqrt{\nu_n}}$ converge en probabilités vers 0.
4. À l'aide des résultats précédents, établir la convergence en loi vers Z de la suite des variables aléatoires $\frac{1}{\sqrt{\nu_n}}(S_{\nu_n} - \nu_n)$.

Solution de l'exercice 3.57

1. Pour tout $n \in \mathbf{N}^*$, on a $v_n - 1 < \lfloor v_n \rfloor \leqslant v_n$ et donc

$$1 - \frac{1}{v_n} < \frac{\lfloor v_n \rfloor}{v_n} \leqslant 1.$$

Puisque (v_n) est croissante et non majorée, on a alors $v_n \to +\infty$ et donc, par le théorème des gendarmes, la suite $\left(\dfrac{\lfloor v_n \rfloor}{v_n}\right)_n$ converge vers 1.

2.a. Puisque les X_i ont pour espérance 1 et pour variance 1, Z_n est la variable centrée réduite associée à $\sum_{k=1}^{n} X_k$. Et donc, les X_i étant indépendantes et identiquement distribuées, par le théorème central limite, $Z_n \xrightarrow{\mathcal{L}} Z$.

2.b. Notons que $S_{\lfloor v_n \rfloor}$ suit la loi $\gamma(\lfloor v_n \rfloor)$, tout comme[1] $\sum_{k=1}^{\lfloor v_n \rfloor} X_k$.

[1] C'est la stabilité des lois γ, puisqu'une loi $\mathcal{E}(1)$ est aussi une loi $\gamma(1)$.

D'après la question 1, $\lfloor v_n \rfloor \underset{n\to+\infty}{\sim} v_n \underset{n\to+\infty}{\longrightarrow} +\infty$ et donc $Z_{\lfloor v_n \rfloor} \xrightarrow{\mathcal{L}} Z$.
Et puisque $\dfrac{1}{\sqrt{n}}(S_{\lfloor v_n \rfloor} - \lfloor v_n \rfloor)$ a même loi que $Z_{\lfloor v_n \rfloor}$, on a donc

$$\frac{1}{\sqrt{\lfloor v_n \rfloor}}(S_{\lfloor v_n \rfloor} - \lfloor v_n \rfloor) \xrightarrow{\mathcal{L}} Z.$$

De plus, on a

$$\frac{1}{\sqrt{v_n}}(S_{\lfloor v_n \rfloor} - \lfloor v_n \rfloor) = \sqrt{\frac{\lfloor v_n \rfloor}{v_n}} \times \frac{1}{\sqrt{\lfloor v_n \rfloor}}(S_{\lfloor v_n \rfloor} - \lfloor v_n \rfloor).$$

Or, on a $\sqrt{\dfrac{\lfloor v_n \rfloor}{v_n}} \underset{n\to+\infty}{\longrightarrow} 1$ et donc, si on voit cette suite de réels comme une suite de variables aléatoires certaines, on a alors $\sqrt{\dfrac{\lfloor v_n \rfloor}{v_n}} \xrightarrow{P} 1$.
Par le lemme de Slutsky, on a alors

$$\frac{1}{\sqrt{v_n}}(S_{\lfloor v_n \rfloor} - \lfloor v_n \rfloor) \xrightarrow{\mathcal{L}} 1 \times Z = Z.$$

Astuce

L'utilisation faite ici de la stabilité des lois gamma est assez astucieuse.
Notons qu'un réflexe naturel serait d'essayer d'exprimer les probabilités de $[S_{\delta_n} \geqslant x]$ et $[S_1 \geqslant x]$ à l'aide de deux intégrales. Mais il est alors très difficile de prouver que la première est inférieure à la seconde (sauf si $x \geqslant 1$, auquel cas on peut utiliser la croissance de l'intégrale).

3.a. Commençons par remarquer que $\delta_n \in [0, 1[$.
Notons Y_n une variable aléatoire indépendante de S_{δ_n} suivant la loi $\gamma(1 - \delta_n)$.
Alors, par stabilité des lois γ, $S_{\delta_n} + Y_n$ suit la loi $\gamma(1)$, et donc a même loi que S_1.
Puisque Y_n est à valeurs positives, on a

$$[S_{\delta_n} \geqslant x] \subset [S_{\delta_n} + Y_n \geqslant x].$$

Et en particulier, on a

$$P(S_{\delta_n} \geqslant x) \leqslant P(S_{\delta_n} + Y_n \geqslant x) = P(S_1 \geqslant x).$$

3.b. Soit $\varepsilon > 0$. Alors, par positivité de S_{δ_n},

$$0 \leqslant P\left(\left|\frac{S_{\delta_n}}{\sqrt{v_n}}\right| \geqslant \varepsilon\right) = P(S_{\delta_n} \geqslant \sqrt{v_n}\varepsilon) \leqslant P(S_1 \geqslant \sqrt{v_n}\varepsilon).$$

Mais on a alors,

$$P(S_1 \geqslant \sqrt{v_n}\varepsilon) = 1 - P(S_1 \leqslant \sqrt{v_n}\varepsilon) \underset{n\to+\infty}{\longrightarrow} 1 - 1 = 0.$$

Et donc $\dfrac{S_{\delta_n}}{\sqrt{v_n}} \xrightarrow{P} 0$.

Danger !

On aimerait pouvoir dire que $X_n = S_{\delta_n}$, mais malheureusement, l'énoncé ne fait aucune hypothèse sur l'indépendance mutuelle des S_v.
Notons qu'on pourrait faire nous même cette hypothèse en expliquant qu'elle ne change rien au problème, la convergence en loi n'étant pas affectée par le fait que les variables soient indépen-

4. Pour tout n, soit X_n une variable aléatoire suivant la loi $\gamma(\delta_n)$, indépendante de $S_{\lfloor v_n \rfloor}$.
Alors $S_{\lfloor v_n \rfloor} + X_n$ a même loi que S_{v_n}.

Mais $\frac{1}{\sqrt{v_n}}\left(S_{\lfloor v_n\rfloor}+X_n-v_n\right)=\frac{1}{\sqrt{v_n}}\left(S_{\lfloor v_n\rfloor}-\lfloor v_n\rfloor\right)+\frac{1}{\sqrt{v_n}}\left(X_n-\delta_n\right)$.
Soit $\varepsilon>0$. Alors, par l'inégalité triangulaire, on a

$$\left[\left|\frac{X_n}{\sqrt{v_n}}\right|<\frac{\varepsilon}{2}\right]\cap\left[\left|\frac{\delta_n}{\sqrt{v_n}}\right|<\frac{\varepsilon}{2}\right]\subset\left[\left|\frac{1}{\sqrt{v_n}}(X_n-\delta_n)\right|<\varepsilon\right].$$

Et donc en passant à l'événement contraire[2],

$$\left[\left|\frac{1}{\sqrt{v_n}}(X_n-\delta_n)\right|\geqslant\varepsilon\right]\subset\left[\left|\frac{X_n}{\sqrt{v_n}}\right|\geqslant\frac{\varepsilon}{2}\right]\cup\left[\left|\frac{\delta_n}{\sqrt{v_n}}\right|\geqslant\frac{\varepsilon}{2}\right].$$

[2] **Rappel** : le contraire d'une intersection est l'union des contraires.

Et donc

$$0\leqslant P\left(\left|\frac{1}{\sqrt{v_n}}(X_n-\delta_n)\right|\geqslant\varepsilon\right)\leqslant P\left(\left|\frac{X_n}{\sqrt{v_n}}\right|\geqslant\frac{\varepsilon}{2}\right)+P\left(\left|\frac{\delta_n}{\sqrt{v_n}}\right|\geqslant\frac{\varepsilon}{2}\right).$$

Mais par la question 3.b,

$$P\left(\left|\frac{X_n}{\sqrt{v_n}}\right|\geqslant\frac{\varepsilon}{2}\right)\underset{n\to+\infty}{\longrightarrow}0.$$

Et puisque $0\leqslant\delta_n\leqslant1$, alors $\frac{\delta_n}{\sqrt{v_n}}\underset{n\to+\infty}{\longrightarrow}$ et donc pour n suffisamment grand, $\left|\frac{\delta_n}{\sqrt{v_n}}\right|<\frac{\varepsilon}{2}$ et donc $P\left(\left|\frac{\delta_n}{\sqrt{v_n}}\right|\geqslant\frac{\varepsilon}{2}\right)=0$.
On en déduit, par le théorème des gendarmes que

$$P\left(\left|\frac{1}{\sqrt{v_n}}(X_n-\delta_n)\right|\geqslant\varepsilon\right)\underset{n\to+\infty}{\longrightarrow}0$$

et donc $\frac{1}{\sqrt{v_n}}(X_n-\delta_n)\xrightarrow{P}0$.

Puisque de plus, $\frac{1}{\sqrt{v_n}}\left(S_{\lfloor v_n\rfloor}-\lfloor v_n\rfloor\right)\xrightarrow{\mathcal{L}}Z$ par la question 2.b, alors, par le lemme de Slutsky, $\frac{1}{\sqrt{v_n}}\left(S_{\lfloor v_n\rfloor}+X_n-v_n\right)\xrightarrow{\mathcal{L}}Z$.
Mais $\frac{1}{\sqrt{v_n}}\left(S_{\lfloor v_n\rfloor}+X_n-v_n\right)$ à même loi que $\frac{1}{\sqrt{v_n}}(S_{v_n}-v_n)$, et donc

$$\frac{1}{\sqrt{v_n}}\left(S_{v_n}-v_n\right)\xrightarrow{\mathcal{L}}Z.$$

□

Conv. en loi
La convergence en loi ne «regarde» que la fonction de répartition, donc si deux suites de variables (X_n) et (Y_n) sont telles que X_n et Y_n ont même loi, alors $X_n\xrightarrow{\mathcal{L}}Z$ si et seulement si $Y_n\xrightarrow{\mathcal{L}}Z$.

3.6.4 Théorème central limite

Exercice 3.58 (QSP HEC 2012) [HEC12.QSP8] **Facile**

Convergence en loi d'une suite de variables aléatoires

Soit $(U_n)_{n\in\mathbf{N}^*}$ une suite de variables aléatoires indépendantes, définies sur le même espace probabilisé $(\Omega,\mathcal{A},P)$, de même loi uniforme sur $]0,1]$.
On pose, pour tout $n\in\mathbf{N}^*$,

$$X_n=\prod_{i=1}^{n}U_i^{\frac{1}{n}}\text{ et }Y_n=(eX_n)^{\sqrt{n}}.$$

Montrer que la suite de variables aléatoires $(\ln Y_n)_{n\in\mathbf{N}^*}$ converge en loi vers une variable aléatoire dont on précisera la loi.

Solution de l'exercice 3.58 Soit $n\in\mathbf{N}^*$. Alors on a

$$\ln Y_n=\sqrt{n}\left(1+\ln X_n\right)=\sqrt{n}\left(1+\frac{1}{n}\sum_{i=1}^{n}\ln U_i\right)=\frac{1}{\sqrt{n}}\left(n+\sum_{i=1}^{n}\ln U_i\right)$$

Si l'on pose $S_n = \sum_{i=1}^n \ln U_i$, alors

$$\ln Y_n = \frac{S_n + n}{\sqrt{n}}.$$

Soit alors $T_i = \ln U_i$. Nous pourrions calculer une densité de T_i afin de déterminer son espérance et sa variance, mais il est plus simple d'utiliser le théorème de transfert : sous réserve de convergence absolue,

$$E(T_i) = \int_0^1 \ln t dt = \lim_{A\to 0^+} [t\ln t - t]_A^1 = -1$$

De même, on a

$$E(T_i^2) = \int_0^1 (\ln t)^2 dt = \lim_{A\to 0^+} \int_A^1 (\ln t)^2 dt$$

Mais par une intégration par parties, il vient

$$\int_A^1 (\ln t)^2 dt = \left[t(\ln t)^2 - t\ln t\right]_A^1 - \int_A^1 (\ln t - 1)dt = -A(\ln A)^2 + A\ln A + A\ln A + 1 + 1 - A \underset{A\to 0^+}{\longrightarrow} 2$$

Donc $E(T_1^2)$ existe et vaut 2. Et donc

$$V(T_1) = E(T_1^2) - E(T_1)^2 = 2 - 1^2 = 1.$$

De plus, les T_i sont indépendantes et identiquement distribuées, de sorte que

$$\frac{S_n - E(S_n)}{V(S_n)} = \frac{S_n + n}{\sqrt{n}} = \ln Y_n.$$

Mais les T_i étant i.i.d et possédant une variance, alors par le théorème central limite, $\ln Y_n$ converge en loi vers une variable aléatoire X suivant la loi normale centrée réduite. □

EXERCICE 3.59 (HEC 2016) [HEC2016.176] — Moyen

Preuve de non convergence en probabilité.

Soit $(X_n)_{n\in\mathbf{N}^*}$ une suite de variables aléatoires indépendantes, définies sur le même espace probabilisé $(\Omega, \mathcal{A}, P)$ et suivant toutes une même loi, qui admet un moment d'ordre deux.
On suppose que les variables aléatoires X_n sont centrées et réduites.

1. Pour tout $n \in \mathbf{N}^*$, on pose :

$$U_n = \frac{X_1 + X_2 + \cdots + X_n}{\sqrt{n}} \text{ et } V_n = \frac{X_{n+1} + X_{n+2} + \cdots + X_{2n}}{\sqrt{n}}.$$

Soit $\varepsilon > 0$.

(a) Montrer que $P(U_n \geqslant \varepsilon)$ tend vers $1 - \Phi(\varepsilon)$ quand n tend vers l'infini.

(b) Quelle est la limite de $P(V_n \leqslant -\varepsilon)$ quand n tend vers l'infini ?

(c) Pour tout $n \in \mathbf{N}^*$, on pose $Y_n = \left(\sqrt{2} - 1\right) U_n - V_n$.

i. Justifier l'inégalité :

$$P\left(Y_n \geqslant \varepsilon\sqrt{2}\right) \geqslant P([U_n \geqslant \varepsilon] \cap [V_n \leqslant -\varepsilon]).$$

ii. En déduire que la suite $(Y_n)_{n\in\mathbf{N}^*}$ ne converge pas en probabilité vers 0.

2. On conserve les notations de la question précédente.

(a) Exprimer $U_{2n} - U_n$ en fonction de Y_n.

(b) Démontrer qu'il n'existe pas de variable aléatoire Z telle que la suite $(U_n)_{n\in\mathbf{N}^*}$ converge en probabilité vers Z.

SOLUTION DE L'EXERCICE 3.59

1.a. Puisque les X_n sont indépendantes, centrées et réduites, on a $E(X_1+\cdots+X_n)=0$ et $V(X_1+\cdots+X_n)=V(X_1)+\cdots+V(X_n)=n$.
Ainsi, la variable centrée réduite associée à $X_1+\cdots+X_n$ est U_n, et donc par le théorème central limite[1], $U_n \xrightarrow{\mathcal{L}} X$, où X suit la loi normale $\mathcal{N}(0,1)$.
Et donc $P(U_n \geqslant \varepsilon) \underset{n\to+\infty}{\longrightarrow} P(X \geqslant \varepsilon) = 1-\Phi(\varepsilon)$.

[1] Qui s'applique puisque les X_i sont i.i.d. et possèdent un moment d'ordre 2.

1.b. De la même manière, on prouve que $V_n \xrightarrow{\mathcal{L}} X$ et donc

$$P(V_n \leqslant -\varepsilon) \underset{n\to+\infty}{\longrightarrow} P(X \leqslant -\varepsilon) = \Phi(-\varepsilon) = 1-\Phi(\varepsilon).$$

1.c.i. Si on a à la fois $[U_n \geqslant \varepsilon]$ et $V_n \leqslant -\varepsilon$, alors

$$Y_n = \left(\sqrt{2}-1\right)U_n - V_n \geqslant \left(\sqrt{2}-1\right)\varepsilon - (-\varepsilon) = \sqrt{2}\varepsilon.$$

Et donc $[U_n \geqslant \varepsilon] \cap [V_n \leqslant -\varepsilon] \subset \left[Y_n \geqslant \varepsilon\sqrt{2}\right]$, ce qui en passant aux probabilités nous donne l'inégalité désirée.

1.c.ii. Par le lemme des coalitions, les variables aléatoires U_n et V_n sont indépendantes, de sorte que

$$P([U_n \geqslant \varepsilon] \cap [V_n \leqslant -\varepsilon]) = P(U_n \geqslant \varepsilon)P(V_n \leqslant -\varepsilon) \underset{n\to+\infty}{\longrightarrow} (1-\Phi(\varepsilon))^2.$$

Et donc en particulier, $P\left(Y_n \geqslant \varepsilon\sqrt{2}\right) \underset{\not\to}{\longrightarrow}{}_{n+\infty} 0$.

Or, si on avait $Y_n \xrightarrow{P} 0$, alors on aurait

$$P\left(Y_n \geqslant \varepsilon\sqrt{2}\right) \leqslant P\left(|Y_n| \geqslant \varepsilon\sqrt{2}\right) \underset{n\to+\infty}{\longrightarrow} 0.$$

Et donc on en déduit que Y_n ne converge pas en probabilités vers 0.

2.a. On a

$$\begin{aligned} U_{2n}-U_n &= \frac{X_1+\cdots+X_{2n}}{\sqrt{2n}} - \frac{X_1+\cdots+X_n}{\sqrt{n}} \\ &= \frac{X_1+\cdots+X_{2n}-\sqrt{2}(X_1+\cdots+X_n)}{\sqrt{2n}} \\ &= \frac{1}{\sqrt{2}}\frac{X_{n+1}+\cdots+X_{2n}+(1-\sqrt{2})(X_1+\cdots+X_n)}{\sqrt{n}} = -\frac{Y_n}{\sqrt{2}}. \end{aligned}$$

2.b. Supposons par l'absurde qu'il existe Z telle que $U_n \xrightarrow{P} Z$.
Alors on aurait également $U_{2n} \xrightarrow{P} Z$, puisque pour tout $\varepsilon>0$, $\lim\limits_{n\to+\infty} P(|U_{2n}-Z| \geqslant \varepsilon) \underset{n\to+\infty}{\longrightarrow} 0$.
Mais alors, pour tout $\varepsilon>0$, on a, d'après l'inégalité triangulaire,

$$\left[|U_{2n}-Z| < \frac{\varepsilon}{2}\right] \cap \left[|U_n-Z| < \frac{\varepsilon}{2}\right] \subset \left[|U_{2n}-U_n| < \varepsilon\right].$$

Et donc par passage à l'événement contraire,

$$\left[|U_{2n}-U_n| \geqslant \varepsilon\right] \subset \left[|U_{2n}-Z| \geqslant \frac{\varepsilon}{2}\right] \cup \left[|U_n-Z| \geqslant \frac{\varepsilon}{2}\right].$$

On en déduit[2] que

$$P(|U_{2n}-U_n| \geqslant \varepsilon) \leqslant P\left(|U_{2n}-Z| \geqslant \frac{\varepsilon}{2}\right) + P\left(|U_n-Z| \geqslant \frac{\varepsilon}{2}\right) \underset{n\to+\infty}{\longrightarrow} 0.$$

[2] $P(A\cup B) \leqslant P(A)+P(B)$.

Et donc autrement dit, $U_{2n}-U_n \xrightarrow{P} 0$.
Mais alors, on devrait également avoir $Y_n = -\sqrt{2}(U_{2n}-U_n) \xrightarrow{P} 0$, ce qui n'est pas le cas d'après la question 1.c.
On en déduit donc qu'il n'existe pas de variable aléatoire Z telle que $U_n \xrightarrow{P} Z$.

□

L'exercice suivant, sans grande difficulté, est essentiellement calculatoire. Pas le plus intéressant que l'on puisse imaginer, mais un des rares qui applique de façon non triviale le théorème central limite.

EXERCICE 3.60 (ESCP 2012) [ESCP12.3.17] — Facile

Autour de la loi log-normale

On note φ la densité continue de $\mathbf{R}$ de la loi normale centrée réduite et Φ la fonction de répartition correspondante. On dira qu'une variable aléatoire réelle positive X suit la loi $LN(\mu,\theta)$, $\mu\in\mathbf{R}$, $\theta\in\mathbf{R}_+^*$ si sa fonction de répartition F_X est donnée par :

$$F_X(x)=\begin{cases}\Phi\left(\dfrac{1}{\theta}\left(\ln(x)-\mu\right)\right) & \text{si } x>0\\ 0 & \text{si } x\leqslant 0\end{cases}$$

1. Montrer que F_X est bien la fonction de répartition d'une variable aléatoire à densité, et en donner une densité.
2. On suppose que X suit la loi $LN(\mu,\theta)$ et on pose $Y=\ln X$. Déterminer la loi de Y.
3. Soient X_1 et X_2 deux variables aléatoires indépendantes, respectivement de lois $LN(\mu_1,\theta_1)$ et $LN(\mu_2,\theta_2)$. Déterminer la loi du produit X_1X_2.
4. Soit Z une variable aléatoire de loi $LN(\mu,\theta)$. Montrer que pour tout $c,\alpha\in\mathbf{R}_+^*$, la variable aléatoire cZ^α est de loi $LN(\mu',\theta')$, où on précisera les valeurs de μ' et θ'.
5. Soit $p\in]0,1[$ et $q=1-p$. On définit $X_1,\dots,X_n$ des variables aléatoires indépendantes par :
$$P(X_i=e)=p,\ p(X_i=1)=q.$$
   (a) Quelle est la loi de $S_n=\sum_{i=1}^{n}(\ln X_i)$?
   (b) Montrer que $T_n=e^{-p\sqrt{n}}\left(\prod_{i=1}^{n}X_i\right)^{\frac{1}{\sqrt{n}}}$ converge en loi vers une variable aléatoire do

Méthode

Puisqu'ici, on ne sait même pas que F_X est la fonction de répartition d'une variable aléatoire (même si la notation F_X est trompeuse), on n'oubliera pas de vérifier ls quatre points caractéristiques des fonctions de répartition : les limites en $\pm\infty$, la croissance de F_X et sa continuité à droite en tout point.
Prouver que F_X est continue sur $\mathbf{R}$ implique notamment ce dernier point !
Une fois ces quatre points prouvés, il faut encore vérifier que F_X est continue sur $\mathbf{R}$ et $\mathscr{C}^1$ sauf éventuellement en un nombre fini de points pour prouver que F_X est la fonction de répartition d'une variable **à densité.**

SOLUTION DE L'EXERCICE 3.60

1. Il est évident que F_X est de classe $\mathscr{C}^1$ sur $\mathbf{R}^*$, et qu'elle est croissante sur $\mathbf{R}_+$ par composition de fonctions croissantes[1].
Par conséquent, elle est croissante sur $\mathbf{R}$ tout entier.
De plus, on a $\lim\limits_{x\to0^+}F_X(x)=\lim\limits_{t\to-\infty}\Phi(t)=0=\lim\limits_{x\to0^-}F_X(x)=0$. Donc F_X est continue en 0 et donc sur $\mathbf{R}$.
Enfin, on a $\lim\limits_{x\to-\infty}F_X(x)=0$ et $\lim\limits_{x\to+\infty}F_X(x)=\lim\limits_{t\to+\infty}\Phi(t)=1$.
Et donc F_X est bien la fonction de répartition d'une variable à densité, dont une densité est toute fonction coïncidant avec F_X' sauf éventuellement en un nombre fini de points.
Par exemple, on peut prendre

$$f(x)=\begin{cases}\dfrac{1}{x}\varphi\left(\dfrac{1}{\theta}\left(\ln x-\mu\right)\right) & \text{si } x>0\\ 0 & \text{sinon}\end{cases}=\begin{cases}\dfrac{1}{x\sqrt{2\pi}}e^{-\frac{1}{2}\left(\frac{\ln(x)-\mu}{\theta}\right)^2} & \text{si } x>0\\ 0 & \text{sinon}\end{cases}$$

[1] La fonction Φ est croissante sur $\mathbf{R}$, puisque c'est une fonction de répartition.

2. Pour $x\in\mathbf{R}$, on a, par croissance de l'exponentielle,

$$P(Y\leqslant x)=P(\ln(X)\leqslant x)=P(X\leqslant e^x)=F_X(e^x)=\Phi\left(\frac{1}{\theta}(x-\mu)\right).$$

Nous reconnaissons là la fonction de répartition d'une loi $\mathcal{N}(\mu,\theta^2)$, et donc $Y\hookrightarrow\mathcal{N}(\mu,\theta^2)$.

Rappel

Pour obtenir la fonction de répartition d'une loi normale, on passe par la variable centrée réduite associée, qui suit la loi $\mathcal{N}(0,1)$ et donc possède Φ comme fonction de répartition.

3. Puisque X_1 et X_2 sont indépendantes, alors $Y_1=\ln(X_1)$ et $Y_2=\ln(X_2)$ le sont également, et suivent respectivement les lois $\mathcal{N}(\mu_1,\theta_1^2)$ et $\mathcal{N}(\mu_2,\theta_2^2)$.
Par stabilité des lois normales, $Y_1+Y_2=\ln(X_1X_2)$ suit alors la loi $\mathcal{N}(\mu_1+\mu_2,\theta_1^2+\theta_2^2)$.
Et donc X_1X_2 suit la loi $LN(\mu_1+\mu_2,\sqrt{\theta_1^2+\theta_2^2})$.

Détails

Le calcul de la question 2 est totalement réversible :

$$X\hookrightarrow LN(\mu,\theta)\iff \ln(X)\hookrightarrow\mathcal{N}(\mu,\theta^2).$$

4. On a $\ln(cZ^\alpha)=\ln c+\alpha\ln Z$.
Mais $\ln Z$ suit la loi $\mathcal{N}(\mu,\theta)$, si bien que, par transformation affine de loi normale,

$$\ln c+\alpha\ln Z\hookrightarrow\mathcal{N}\left(\alpha\mu+\ln c,\alpha^2\theta^2\right).$$

Et donc $cZ^\alpha\hookrightarrow LN\left(\alpha\mu+\ln c,\alpha\theta\right)$.

5.a. Les $\ln(X_i)$ prennent les valeurs 0 et 1, avec probabilités respectives q et p, donc suivent la loi $\mathcal{B}(p)$. Et puisqu'elles sont indépendantes, par stabilité des lois binomiales, $S_n \hookrightarrow \mathcal{B}(n,p)$.

5.b. On a $\ln(T_n) = -p\sqrt{n} + \dfrac{1}{\sqrt{n}}S_n = \dfrac{S_n - np}{\sqrt{n}}$.

Mais, par le théorème central limite, $\dfrac{S_n - np}{\sqrt{npq}} \xrightarrow{\mathcal{L}} Z$, où $Z \hookrightarrow \mathcal{N}(0,1)$.

Et alors, par composition par la fonction continue $x \mapsto x\sqrt{pq}$, $\ln(T_n) \xrightarrow{\mathcal{L}} \sqrt{pq}Z$, où $\sqrt{pq}Z \hookrightarrow \mathcal{N}(0,pq)$.

Et alors, par composition par la fonction exponentielle, $T_n \xrightarrow{\mathcal{L}} e^{\sqrt{pq}Z}$, qui suit la loi $LN(0,\sqrt{pq})$.

□

3.7 Estimation

3.7.1 Estimation ponctuelle

Exercice 3.61 (HEC 2009) [HEC09.8] **Moyen**

Estimation du paramètre d'une loi à densité

On note F la fonction définie par

$$F(x) = \begin{cases} 0 & \text{si } x \leqslant -1 \\ \dfrac{1}{2}(x+1)^2 & \text{si } -1 < x \leqslant 0 \\ 1 - \dfrac{1}{2}(x-1)^2 & \text{si } 0 < x < 1 \\ 1 & \text{si } x \geqslant 1 \end{cases}$$

1. (a) Montrer que F est de classe $\mathcal{C}^1$ sur $\mathbf{R}$.
   (b) On note f la fonction dérivée de F et, pour tout nombre réel θ, f_θ la fonction définie sur $\mathbf{R}$ par :
   $$\forall x \in \mathbf{R},\ f_\theta(x) = f(x-\theta).$$
   Vérifier que f_θ est une densité de probabilités.
   (c) Soit X une variable aléatoire définie sur un espace probabilisé $(\Omega, \mathcal{A}, P)$, admettant f_θ pour densité. Montrer que X possède une espérance et une variance, et les calculer.
2. Soit $(X_n)_{n\in\mathbf{N}^*}$ une suite de variables aléatoires, définies sur $(\Omega, \mathcal{A}, P)$, indépendantes, de même loi, admettant pour densité f_θ.
   Pour tout entier $n \geqslant 1$, on note $U_n = \min(X_1, \ldots, X_n)$ et $V_n = \max(X_1, \ldots, X_n)$.
   (a) Exprimer, à l'aide de F, θ et n les fonctions de répartition de U_n et V_n.
   (b) Justifier, pour tout $\varepsilon \in]0,1[$, l'inégalité :
   $$P\left([-1+\theta \leqslant U_n < -1+\theta+2\varepsilon] \cap [1+\theta-2\varepsilon < V_n \leqslant 1+\theta]\right) \leqslant P\left(\left|\frac{U_n+V_n}{2} - \theta\right| < \varepsilon\right).$$
   (c) En déduire que $\dfrac{U_n+V_n}{2}$ est un estimateur convergent de θ.
   (d) Est-il sans biais ?

Solution de l'exercice 3.61

1.a. Par opérations sur les fonctions usuelles, F est $\mathcal{C}^1$ sauf peut-être en -1, 0 et 1.
Mais on a $\lim\limits_{x\to-1^+} F(x) = 0 = \lim\limits_{x\to-1^-} F(x) = F(-1)$ et donc F est continue en -1.
De même, $\lim\limits_{x\to0^-} F(x) = \dfrac{1}{2} = F(0) = \lim\limits_{x\to0^+} F(x)$, donc F est continue en 0.
Enfin, $\lim\limits_{x\to1^-} F(x) = 1 = \lim\limits_{x\to1^+} F(X)$, donc F est également continue en 1, et donc sur $\mathbf{R}$.

De plus, on a, pour $x \notin \{-1, 0, 1\}$,

$$F'(x) = \begin{cases} 0 & \text{si } x < -1 \\ x+1 & \text{si } 1 < x < 0 \\ 1-x & \text{si } 0 < x < 1 \\ 0 & \text{si } x > 1 \end{cases}$$

Pour $x \in]-1, 0]$, on a $\dfrac{F(x)-F(-1)}{x+1} = \dfrac{(x+1)^2}{2(x+1)} \xrightarrow[x\to -1^+]{} 0$, et donc F est dérivable à droite en -1, avec une dérivée droite égale à 0.
D'autre part, il est évident qu'elle est dérivable à gauche, avec $F'_g(-1) = 0$. Donc F est dérivable en -1, et $F'(-1) = 0$.
On constate alors que $\lim\limits_{x\to -1^-} F'(x) = \lim\limits_{x\to -1^+} F'(x) = F'(-1)$, donc F' est continue en -1.

Détails
Sur $]-1, 0]$,
$$F'(x) = x+1$$
et donc en particulier, $F'_g(0) = 1$.

Puisque $F(x) = \dfrac{1}{2}(x+1)^2$ sur $]-1, 0]$, alors elle est en particulier dérivable à gauche en 0, avec $F'_g(0) = 1$. D'autre part, pour $x \in]0, 1]$,

$$\frac{F(x)-F(0)}{x} = \frac{1}{x}\left(1 - \frac{1}{2}(x-1)^2 - \frac{1}{2}\right) = \frac{1-(x-1)^2}{2x} = \frac{-x^2+2x}{2x} \xrightarrow[x\to 0^+]{} 1.$$

Et donc F est dérivable à droite en 0, avec $F'_d(0) = 1$.
Donc F est dérivable en 0 avec $F'(0) = 1$.
On constate alors que $\lim\limits_{x\to 0^-} F'(x) = \lim\limits_{x\to 0^+} F'(x) = F'(0)$ et donc F' est continue en 0.
On prouve de même que F est bien dérivable en 1, avec F' continue en 1.
Et donc que F est dérivable sur $\mathbf{R}$, à dérivée continue, et donc est de classe $\mathscr{C}^1$ sur $\mathbf{R}$.

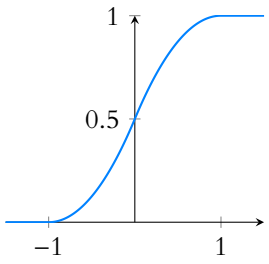


FIGURE 3.9– La fonction F.

1.b. Notons que F est croissante sur $\mathbf{R}$, et que $\lim\limits_{x\to -\infty} F(x) = 0$ et $\lim\limits_{x\to +\infty} F(x) = 1$.
Puisqu'elle est $\mathscr{C}^1$ sur $\mathbf{R}$, et en particulier continue sur $\mathbf{R}$, il s'agit de la fonction de répartition d'une variable à densité.
Sa dérivée f est donc une densité de probabilités. Notons que

$$f(x) = \begin{cases} 0 & \text{si } x \leqslant -1 \\ x+1 & \text{si } -1 < x < 0 \\ x-1 & \text{si } 0 \leqslant x < 1 \\ 0 & \text{si } x \geqslant 1 \end{cases}.$$

Remarque
Cette densité est aussi celle de la somme de deux lois uniformes indépendantes sur $\left[-\frac{1}{2}, \frac{1}{2}\right]$.

Et alors, par composition de fonctions continues[1], f_θ est encore continue.
De même, puisque f est positive sur $\mathbf{R}$, f_θ l'est également.
Enfin, le changement de variable $t = x - \theta$ prouve que

$$\int_{-\infty}^{+\infty} f_\theta(x)\,dx = \int_{-\infty}^{+\infty} f(x-\theta)\,dx = \int_{-\infty}^{+\infty} f(t)\,dt = 1.$$

Et donc f_θ est bien une densité de probabilités.

[1] f est continue car F est $\mathscr{C}^1$.

1.c. Notons que par transformation affine, X admet f_θ pour densité si et seulement si $X - \theta$ admet f pour densité.
En effet, si une densité de X est f_θ, alors une densité de $X - \theta$ est

$$x \mapsto f_\theta(x+\theta) = f(x+\theta-\theta) = f(x).$$

La réciproque se prouve de la même manière.

Remarque
Une transformation affine est toujours réversible, et donc si on a prouvé que si X a pour densité f_θ, alors $X - \theta$ a pour densité f, alors l'implication réciproque est automatique.

Soit alors Y une variable aléatoire de densité f. Puisque f est paire, $t \mapsto tf(t)$ est impaire, et on a alors[2],

$$E(Y) = \int_{-\infty}^{+\infty} tf(t)\,dt = 0.$$

[2] Il s'agit d'une intégrale sur un segment, donc la convergence est assurée.

En revanche, la fonction $t \mapsto t^2 f(t)$ est paire, et donc,

$$E\left(Y^2\right) = \int_{-\infty}^{+\infty} t^2 f(t)\,dt = 2\int_0^{+\infty} t^2 f(t)\,dt = 2\int_0^1 t^2(1-t)\,dt = \frac{1}{6}.$$

Et donc $V(Y) = E\left(Y^2\right) = \dfrac{1}{6}$.

On en déduit que $E(X-\theta) = E(Y) \Leftrightarrow E(X) = \theta$ et $V(X) = V(X-\theta) = V(Y) = \dfrac{1}{6}$.

2.a. Nous avons dit précédemment que X a pour densité f_θ si et seulement si $X-\theta$ a pour densité f (et donc F pour fonction de répartition).
Et donc, si X a pour densité f_θ, pour $x \in \mathbf{R}$,

$$P(X \leqslant x) = P(X - \theta \leqslant x - \theta) = F(x-\theta).$$

Alors, par indépendances des X_i, il vient

$$P(V_n \leqslant x) = P\left(\bigcap_{i=1}^{n}[X_i \leqslant x]\right) = \prod_{i=1}^{n} P(X_i \leqslant x) = F(x-\theta)^n.$$

Et de même, on a

$$\begin{aligned} P(U_n \leqslant x) &= 1 - P(U_n > x) = 1 - P\left(\bigcap_{i=1}^{n}[X_i > x]\right) \\ &= 1 - \prod_{i=1}^{n} P(X_i > x) = 1 - \prod_{i=1}^{n}(1 - P(X_i \leqslant x)) \\ &= 1 - (1 - F(x-\theta))^n. \end{aligned}$$

2.b. Supposons que les deux événements

$$[-1+\theta \leqslant U_n < -1+\theta+2\varepsilon] \text{ et } [1+\theta-2\varepsilon < V_n \leqslant 1+\theta]$$

soient réalisés simultanément. Alors, en sommant les inégalités, il vient

$$< 2\theta - 2\varepsilon < U_n + V_n < 2\theta + 2\varepsilon \Leftrightarrow -\varepsilon < \frac{U_n+V_n}{2} - \theta < \varepsilon \Leftrightarrow \left|\frac{U_n+V_n}{2} - \theta\right| < \varepsilon.$$

Autrement dit, nous venons de prouver l'inclusion d'événements

$$[-1+\theta \leqslant U_n < -1+\theta+2\varepsilon] \cap [1+\theta-2\varepsilon < V_n \leqslant 1+\theta] \subset \left[\left|\frac{U_n+V_n}{2} - \theta\right| < \varepsilon\right].$$

Et donc, par croissance de la probabilité[3]

$$P\left([-1+\theta \leqslant U_n < -1+\theta+2\varepsilon] \cap [1+\theta-2\varepsilon < V_n \leqslant 1+\theta]\right) \leqslant P\left(\left|\frac{U_n+V_n}{2} - \theta\right| < \varepsilon\right).$$

[3] Il s'agit là du terme savant pour une propriété bien connue :

$A \subset B \Rightarrow P(A) \leqslant P(B)$.

2.c. Passons aux événement contraires dans l'inégalité de la question précédente :

$$1 - P\left(\left[-1+\theta \leqslant U_n < -1+\theta+2\varepsilon\right] \cap \left[1+\theta-2\varepsilon < V_n \leqslant 1+\theta\right]\right) \leqslant 1 - P\left(\left|\frac{U_n+V_n}{2} - \theta\right| \geqslant \varepsilon\right).$$

Soit encore

$$P\left(\left|\frac{U_n+V_n}{2} - \theta\right| \geqslant \varepsilon\right) \leqslant P\left(\overline{[-1+\theta \leqslant U_n < -1+\theta+2\varepsilon]} \cup \overline{[1+\theta-2\varepsilon < V_n \leqslant 1+\theta]}\right).$$

Rappel

$\overline{A \cap B} = \overline{A} \cup \overline{B}$.

Mais on a alors

$$\begin{aligned} &P\left(\overline{[-1+\theta \leqslant U_n < -1+\theta+2\varepsilon]} \cup \overline{[1+\theta-2\varepsilon < V_n \leqslant 1+\theta]}\right) \\ &\quad \leqslant P\left(\overline{[-1+\theta \leqslant U_n < -1+\theta+2\varepsilon]}\right) + P\left(\overline{[1+\theta-2\varepsilon < V_n \leqslant 1+\theta]}\right) \\ &\quad \leqslant 1 - P(-1+\theta \leqslant U_n < -1+\theta+2\varepsilon) + 1 - P(1+\theta-2\varepsilon < V_n \leqslant 1+\theta) \end{aligned}$$

En utilisant les résultats de la question 2.a, il vient, pour $\varepsilon < \dfrac{1}{2}$,

$$1 - P(-1+\theta \leqslant U_n < -1+\theta+2\varepsilon) = 1 - (1 - (1 - F(-1+2\varepsilon))^n) + \left(1 - \left(1 - \underbrace{F(-1)}_{=0}\right)^n\right)$$

$$= 1 - \left(1 - \left(1 - 2\varepsilon^2\right)^n\right)$$
$$\xrightarrow[n\to+\infty]{} 0.$$

De même, on a

$$1 - P\left(1 + \theta - 2\varepsilon < V_n \leqslant 1 + \theta\right) = 1 - \left(\underbrace{F(1)^n}_{=1} - F(1-2\varepsilon)^n\right)$$
$$= \left(1 - 2\varepsilon^2\right)^n \xrightarrow[n\to+\infty]{} 0.$$

Et donc au final, par le théorème des gendarmes, on a bien, pour tout $\varepsilon \in \left]0, \frac{1}{2}\right[$,

$$\lim_{n\to+\infty} P\left(\left|\frac{U_n + V_n}{2} - \theta\right| \geqslant \varepsilon\right) = 0 \Leftrightarrow \frac{U_n + V_n}{2} \xrightarrow{P} \theta.$$

Et donc $\dfrac{U_n + V_n}{2}$ est bien un estimateur convergent de θ.

Remarque

Dans la définition de la convergence en probabilités, on peut remplacer le $\forall \varepsilon > 0$ par «pour tout ε suffisamment petit».
Et donc ici, il suffit de prouver le résultat pour $\varepsilon < \frac{1}{2}$.
En effet, pour $\varepsilon \geqslant \frac{1}{2}$, on a alors

$$P\left(\left|\frac{U_n + V_n}{2} - \theta\right| \geqslant \varepsilon\right) \leqslant P\left(\left|\frac{U_n + V_n}{2} - \theta\right| \geqslant \frac{1}{4}\right) \xrightarrow[n\to+\infty]{} 0.$$

2.d. Commençons par le cas où $\theta = 0$.
Puisque la densité f est paire, pour tout $i \in [\![1, n]\!]$, $-X_i$ et X_i ont même loi.
Et alors $\min(-X_1, \dots, -X_n) = -\max(X_1, \dots, X_n)$.
Les X_i ayant même loi que les $-X_i$, $\min(X_1, \dots, X_n)$ et $\min(-X_1, \dots, -X_n)$ suivent alors la même loi, et donc, sous réserve d'existence, la même espérance, qui est alors l'opposé de l'espérance de $\max(X_1, \dots, X_n)$.
Ainsi, si ces espérance existent, on a $E(U_n) = -E(V_n)$.
Et donc $E\left(\dfrac{U_n + V_n}{2}\right) = 0$.

Dans le cas général, notons que les $X_i - \theta$ ont pour densité f, et puisque $\min(X_1, \dots, X_n) - \theta = \min(X_1 - \theta, \dots, X_n - \theta)$, alors

$$E\left(U_n\right) = E\left(\min(X_1 - \theta, \dots, X_n - \theta) + \theta.\right.$$

De même, on a

$$E\left(V_n\right) = E\left(\max(X_1 - \theta, \dots, X_n - \theta)\right) + \theta.$$

Et comme les $X_i - \theta$ ont pour densité f, nous savons que

$$E\left(\min(X_1 - \theta, \dots, X_n - \theta) = -E\left(\max(X_1 - \theta, \dots, X_n - \theta)\right.\right.$$

et donc

$$E\left(\frac{U_n + V_n}{2}\right) = \frac{1}{2}(\theta + \theta) = \theta.$$

Ainsi, $\dfrac{U_n + V_n}{2}$ est un estimateur sans biais de θ.

□

EXERCICE 3.62 (QSP ESCP 2005) [ESCP05.QSP01 ; sujet] Facile

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 ?

SOLUTION DE L'EXERCICE 3.62

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.

2. T_1 et T_2 étant indépendants, on a

$$V(\Theta_a) = a^2V(T_1) + (1-a)^2V(T_2) = a^2(V(T_1)+V(T_2)) - 2aV(T_2) + V(T_2).$$

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 minimum[1] en $a = \dfrac{V(T_2)}{V(T_1)+V(T_2)}$.
Le biais de Θ_a étant nul, son risque quadratique est égal à sa variance, et donc est minimal pour $a = \dfrac{V(T_2)}{V(T_1)+V(T_2)}$.

□

> **Rappel**
> Une variable aléatoire est de variance nulle si et seulement si c'est une variable certaine.

[1] Une formule bien connue nous explique que le sommet d'une parabole d'équation $y = ax^2+bx+c$ a pour abscisse $x = -\dfrac{b}{2a}$.

Exercice 3.63 (ESCP 2016) [ESCP16.3.14] — Facile

Différents estimateurs du paramètre d'une loi (loi du maximum de deux uniformes sur $[0,\theta]$.)

Soit $(\Omega, \mathcal{A}, P)$ 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

$$f : t \mapsto \begin{cases} \dfrac{2t}{\theta^2} & \text{si } 0 \leqslant t \leqslant \theta \\ 0 & \text{sinon} \end{cases}$$

Pour tout $n \in \mathbf{N}^*$, on pose $\overline{X_n} = \dfrac{1}{n}\sum_{i=1}^{n} X_i$.

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 $T_n = \dfrac{3}{2}\overline{X_n}$ 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. Appliquer le théorème central limite à $\left(\overline{X_n}\right)$. En déduire un intervalle de confiance asymptotique de θ au niveau de confiance 95%, basé sur l'estimateur $\overline{X_n}$.
   Si Φ désigne la fonction de répartition d'une variable aléatoire qui suit une loi $\mathcal{N}(0,1)$, alors on donne $2\Phi(1.96)-1 \simeq 0.95$.
4. 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 $M'_n = \dfrac{2n+1}{2n}M_n$ en tant qu'estimateur du paramètre θ.

Solution de l'exercice 3.63

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.

> **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 $[0,\theta]$.

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 $E(\overline{X_n}) = \dfrac{2}{3}\theta$ 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 $E\left(X_1^2\right) = \dfrac{2}{\theta^2}\displaystyle\int_0^\theta t^3\,dt = \frac{\theta^2}{2}$, et donc par la formule de Huygens,

$$V(X_1) = E\left(X_1^2\right) - E(X_1)^2 = \frac{\theta^2}{2} - \frac{4}{9}\theta^2 = \theta^2\frac{1}{18}.$$

Et donc, par indépendance des X_i, $V(\overline{X_n}) = \dfrac{1}{n^2}\displaystyle\sum_{k=1}^{n} V(X_k) = \frac{V(X_1)}{n} = \frac{1}{18}\frac{\theta^2}{n}$.

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 $n \to +\infty$, on en déduit que T_n est un estimateur convergent de θ.

3. Puisque les X_i sont i.i.d. et admettent une variance, le théorème central limite s'applique, et par conséquent la suite des variables centrées réduites associée à $\left(\overline{X_n}\right)$ converge en loi vers une variable Y suivant la loi $\mathcal{N}(0,1)$.
Mais la variable centrée réduite associée à $\overline{X_n}$ n'est autre que

$$\frac{\overline{X_n} - E\left(\overline{X_n}\right)}{\sqrt{V\left(\overline{X_n}\right)}} = \frac{\overline{X_n} - \frac{2}{3}\theta}{\sqrt{\frac{\theta^2}{18n}}} = \frac{3\sqrt{2n}}{\theta}\overline{X_n} - 2\sqrt{2n}.$$

Et ainsi, par définition d'une convergence en loi :

$$\lim_{n\to+\infty} P\left(-1.96 \leqslant \frac{3\sqrt{2n}}{\theta}\overline{X_n} - 2\sqrt{2n} \leqslant 1.96\right) = P(-1.96 \leqslant Y \leqslant 1.96) = 0.95$$

$$\Leftrightarrow \lim_{n\to+\infty} P\left(\theta\left(-1.96 + 2\sqrt{2n}\right) \leqslant \overline{X_n} \leqslant \theta\left(1.96 + 2\sqrt{2n}\right)\right) = 0.95$$

$$\Leftrightarrow \lim_{n\to+\infty} P\left(\frac{3\sqrt{2n}}{2\sqrt{2n} + 1.96}\overline{X_n} \leqslant \theta \leqslant \frac{3\sqrt{2n}}{2\sqrt{2n} - 1.96}\overline{X_n}\right) = 0.95$$

Et donc l'intervalle de confiance asymptotique cherché est $\left[\dfrac{3\sqrt{2n}}{2\sqrt{2n} + 1.96}\overline{X_n}, \dfrac{3\sqrt{2n}}{2\sqrt{2n} - 1.96}\overline{X_n}\right]$.

4.a. Pour $x \in \mathbf{R}$, on a

$$P(X_1 \leqslant x) = \int_{-\infty}^{x} f(t)\,dt = \begin{cases} 0 & \text{si } x \leqslant 0 \\ \frac{x^2}{\theta^2} & \text{si } x \in [0,\theta] \\ 1 & \text{si } x > \theta \end{cases}$$

Et donc

$$G(x) = P(M_n \leqslant x) = P\left(\bigcap_{i=1}^{n}[X_i \leqslant x]\right) = \prod_{i=1}^{n} P(X_i \leqslant x) = \begin{cases} 0 & \text{si } x \leqslant 0 \\ \left(\frac{x}{\theta}\right)^{2n} & \text{si } x \in [0,\theta] \\ 1 & \text{si } x > \theta \end{cases}$$

4.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

$$g : x \mapsto \begin{cases} 0 & \text{si } x \leqslant 0 \\ 2n\dfrac{x^{2n-1}}{\theta^{2n}} & \text{si } 0 < x < \theta \\ 0 & \text{si } x \geqslant \theta \end{cases}$$

On a alors,

$$E(M_n) = \int_0^\theta tg(t)\,dt = \frac{2n}{\theta^{2n}}\int_0^\theta t^{2n}\,dt = \frac{2n}{2n+1}\theta.$$

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 $V(M_n) = E(M_n^2) - E(M_n)^2 = \dfrac{2n}{(2n+2)(2n+1)^2}\theta^2$.

On en déduit que $V(M'_n) = \dfrac{(2n+1)^2}{4n^2}V(M_n) = \dfrac{\theta^2}{2n(2n+2)}$.

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 $r(M'_n) \leqslant r(T_n)$ et donc que M'_n est un meilleur estimateur de θ que T_n.

□

Exercice 3.64 (QSP ESCP 2015) [ESCP15.QSP03] — Facile

Comparaison de deux estimateurs

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 ?

Solution de l'exercice 3.64 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 $r(V_n) = V(V_n) \geqslant r(T_n)$, de sorte T_n est un meilleur estimateur que V_n. □

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».

Exercice 3.65 (ESCP 2015) [ESCP15.3.18] — Moyen

Autour de l'estimation du paramètre d'une loi de Poisson

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

$$\overline{Y_n} = \frac{1}{n}\sum_{k=1}^{n} Y_k \text{ et } S_n = \sum_{k=1}^{n} X_k.$$

1. (a) Montrer que $\overline{Y_n}$ est un estimateur sans biais de $e^{-\lambda}$?
   (b) Calculer $V(\overline{Y_n})$ et en déduire que $\overline{Y_n}$ 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 $\overline{Y_n}$.

SOLUTION DE L'EXERCICE 3.65

1.a. On a $P(Y_k = 1) = P(X_k = 0) = e^{-\lambda}\dfrac{\lambda^0}{0!} = e^{-\lambda}$.
Donc les Y_k suivent la loi de Bernoulli de paramètre $e^{-\lambda}$.
Et donc $E(\overline{Y_n}) = e^{-\lambda}$, donc $\overline{Y_n}$ 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

$$V(\overline{Y_n}) = \frac{1}{n^2}\sum_{k=1}^{n} V(Y_k) = \frac{1}{n^2}\sum_{k=1}^{n} e^{-\lambda}(1-e^{-\lambda}) = \frac{e^{-\lambda}(1-e^{-\lambda})}{n} \xrightarrow[n\to+\infty]{} 0.$$

Et donc en particulier, $\overline{Y_n}$ est un estimateur convergent de $e^{-\lambda}$.

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.

2. Par définition, on a $\varphi(j) = \dfrac{P([X_1 = 0] \cap [S_n = j])}{P(S_n = j)}$.
Mais $[X_1 = 0] \cap [S_n = j] = [X_1 = 0] \cap \left[\sum_{k=2}^{n} X_k = j\right]$.
Par stabilité[1] des lois de Poisson, $\sum_{k=2}^{n} X_k$ suit une loi de Poisson de paramètre $(n-1)\lambda$ et par le lemme des coalitions, est indépendante de X_1.
Par conséquent,

[1] Les X_k sont indépendantes.

$$P([X_1 = 0] \cap [S_n = j]) = P\left([X = 1 = 0] \cap \left[\sum_{k=2}^{n} X_k = j\right]\right) = P(X_1 = 0)P\left(\sum_{k=2}^{n} X_k = j\right)$$
$$= e^{-\lambda}e^{-(n-1)\lambda}\frac{((n-1)\lambda)^j}{j!}.$$

D'autre part, S_n suit une loi de Poisson de paramètre $n\lambda$ et donc $P(S_n = j) = e^{-n\lambda}\dfrac{(n\lambda)^j}{j!}$.
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,

$$E(T_n) = \sum_{j=0}^{+\infty} \varphi(j)P(S_n = j) = \sum_{j=0}^{+\infty} \left(\frac{n-1}{n}\right)^j e^{-n\lambda}\frac{(n\lambda)^j}{j!}$$
$$= e^{-n\lambda}\sum_{j=0}^{+\infty}\frac{((n-1)\lambda)^j}{j!}$$
$$= e^{-n\lambda}e^{(n-1)\lambda} = e^{-\lambda}.$$

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 $V(T_n) = E\left(T_n^2\right) - e^{-2\lambda}$.
Et par le théorème de transfert,

$$\begin{aligned} E\left(T_n^2\right) &= \sum_{j=0}^{+\infty} \varphi(j)^2 P(S_n = j) \\ &= \sum_{j=0}^{+\infty} \left(\frac{n-1}{n}\right)^{2j} e^{-n\lambda} \frac{(n\lambda)^j}{j!} \\ &= e^{-n\lambda} \sum_{j=0}^{+\infty} \left(\frac{(n-1)^2\lambda}{n}\right)^j \frac{1}{j!} \\ &= e^{-n\lambda} e^{(n-1)^2\lambda/n} \\ &= e^{\frac{-2n+1}{n}\lambda}. \end{aligned}$$

On a reconnu une série exponentielle de paramètre $\frac{(n-1)^2\lambda}{n}$.

Et donc $V(T_n) = e^{\frac{-2n+1}{n}\lambda} - e^{-2\lambda} = e^{-2\lambda}\left(e^{\frac{\lambda}{n}} - 1\right) \underset{n\to+\infty}{\longrightarrow} 0$.

Et donc en particulier, T_n est un estimateur convergent de $e^{-\lambda}$.

4. Puisque $\overline{Y_n}$ et T_n sont des estimateurs sans biais de $e^{-\lambda}$, leurs risques quadratiques sont égaux à leurs variances.
On a alors

$$\frac{V(\overline{Y_n})}{V(T_n)} = \frac{e^{-\lambda}(1-e^{-\lambda})}{e^{-2\lambda}(e^{\lambda/}-1)} = \frac{e^{\lambda}-1}{e^{\lambda/n}-1}.$$

Mais $0 < e^{\lambda/n} - 1 \leqslant e^{\lambda-1}$ et donc $\dfrac{V(\overline{Y_n})}{V(T_n)} \geqslant 1 \Leftrightarrow V(\overline{Y_n}) \geqslant V(T_n)$.

On en déduit que T_n est un meilleur estimateur que $\overline{Y_n}$.

□

Exercice 3.66 (ESCP 2012) [ESCP12.3.1] — Moyen

Estimation par capture-recapture

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 relacher 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 $\dfrac{1}{nm}Y_n$ est un estimateur sans biais et convergent de $\dfrac{1}{N}$.
   (b) Pourquoi ne peut-on pas prendre $\dfrac{nm}{Y_n}$ comme estimateur de N ?
   (c) On pose $B_n = \dfrac{m(n+1)}{Y_n+1}$. 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 $i \in [\![2, n]\!], D_i = X_i - X_{i-1}$. On admet que les D_i sont mutuellement indépendantes.
   (a) Pour tout $i \in [\![2, n]\!]$, que représente concrètement D_i ?
   (b) Déterminer, pour tout $i \in [\![2, n]\!]$, la loi de D_i, son espérance et sa variance. En déduire l'espérance et la variance de X_n.
   (c) On pose $A_n = \dfrac{m}{n}X_n$. Montrer que A_n est un estimateur sans biais et convergent de N, et déterminer son risque quadratique.

SOLUTION DE L'EXERCICE 3.66

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 $\frac{m}{N}$, donc[1] $\frac{Y_n}{n}$ 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 $\frac{1}{N}$.

[1] 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 $\left(1 - \frac{m}{N}\right)^n \neq 0$.

1.c. D'après le théorème de transfert, on a[2]

$$\begin{aligned}
E(B_n) &= \sum_{k=0}^{n} \frac{m(n+1)}{k+1} P(Y_n = k) \\
&= \sum_{k=0}^{n} \frac{m(n+1)}{k+1} \binom{n}{k} \left(\frac{m}{N}\right)^k \left(1 - \frac{m}{N}\right)^{n-k} \\
&= m(n+1) \sum_{k=0}^{n} \frac{1}{n+1} \binom{n+1}{k+1} \left(\frac{m}{N}\right)^k \left(1 - \frac{m}{N}\right)^{n-k} \\
&= N \sum_{k=1}^{n+1} \binom{n+1}{k} \left(\frac{m}{N}\right)^{k-1} \left(\frac{m}{N}\right)^k \left(1 - \frac{m}{N}\right)^{n+1-k} \\
&= N\left(1^{n+1} - \left(1 - \frac{m}{N}\right)^{n+1}\right)
\end{aligned}$$

[2] 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

$$b(B_n) = E(B_n) - N = N\left(1 - \frac{m}{N}\right)^{n+1} \underset{n\to\infty}{\longrightarrow} 0.$$

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 $(i-1)^{\text{ème}}$ lion (exclu) et celle du $i^{\text{ème}}$ 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

$$r(A_n) = V(A_n) = \frac{m^2}{n^2}V(X_n) = \frac{N(N-m)}{n} \xrightarrow[n\to\infty]{} 0.$$

Puisque $r(A_n) \xrightarrow[n\to\infty]{} 0$, A_n est un estimateur convergent de N.

□

Exercice 3.67 (QSP HEC 2018) [HEC18.QSP266] **Moyen**

Divers estimateurs du paramètre d'une loi $\mathcal{U}([-\theta,\theta])$

Soit $(X_n)_{n\in\mathbf{N}^*}$ une suite de variables aléatoires indépendantes, suivant chacune la loi uniforme $\mathcal{U}([-\theta,\theta])$, où θ désigne un paramètre réel strictement positif inconnu.
Pour tout $n \in \mathbf{N}^*$, on note $U_n = \min(X_1,\dots,X_n)$ et $V_n = \max(X_1,\dots,X_n)$.

1. Démontrer que $(V_n)_{n\in\mathbf{N}^*}$ est une suite convergente d'estimateurs de θ.
2. Pour tout $n \in \mathbf{N}^*$, on note $T_n = \frac{1}{2}\sup\{X_j - X_i,\ (i,j) \in [\![1,n]\!]^2\}$.
   (a) Exprimer T_n en fonction de U_n et V_n.
   (b) En déduire que $(T_n)_{n\in\mathbf{N}^*}$ est une suite convergente d'estimateurs de θ.

Solution de l'exercice 3.67

1. Rappelons que la fonction de répartition des X_i est connue : il s'agit de

$$F_{X_i} : x \mapsto \begin{cases} 0 & \text{si } x \leqslant -\theta \\ \dfrac{x+\theta}{2\theta} & \text{si } -\theta \leqslant x \leqslant \theta \\ 1 & \text{si } x \geqslant \theta \end{cases}$$

Et donc, les X_i étant indépendantes, la fonction de répartition de V_n est

$$F_{V_n}(x) = \begin{cases} 0 & \text{si } x \leqslant -\theta \\ \left(\frac{x+\theta}{2\theta}\right)^n & \text{si } -\theta \leqslant x \leqslant \theta \\ 1 & \text{si } x \geqslant \theta \end{cases}$$

Pour x fixé, on a alors

$$\lim_{n\to+\infty} P(V_n \leqslant x) = \begin{cases} 0 & \text{si } x < \theta \\ 1 & \text{si } x \geqslant \theta \end{cases}$$

On reconnaît là la fonction de répartition d'une variable certaine égale à θ, vers laquelle (V_n) converge donc en loi.
La convergence en loi vers une constante est équivalente à la convergence en probabilité vers cette constante... mais ce résultat n'est malheureusement pas au programme, et notre définition de la convergence des estimateurs nécessite une convergence en probabilité.

Méthode

Nous savons qu'un bon moyen de prouver qu'un estimateur est convergent est de prouver que son risque quadratique tend vers 0.
C'est bien le cas ici, mais si on réfléchit 2 secondes avant de démarrer les calculs, on réalise rapidement que cela risque d'être bien long : il faudra une densité de V_n, puis son espérance, puis son moment d'ordre 2...
Ne peut-on pas plutôt utiliser directement la définition de la convergence en probabilité ?

Nous pouvons quand même nous en tirer sans grande difficulté : soit $\varepsilon > 0$. Alors

$$P(|V_n - \theta| \geqslant \varepsilon) = P(\theta - V_n \geqslant \varepsilon) = P(V_n \leqslant \theta - \varepsilon) = \begin{cases} 0 & \text{si } \varepsilon \geqslant \theta \\ \left(1 - \frac{\varepsilon}{2\theta}\right)^n & \text{sinon} \end{cases} \xrightarrow[n\to+\infty]{} 0.$$

Et donc $V_n \xrightarrow{P} \theta$, de sorte que V_n est bien un estimateur convergent de θ.

2.a. On a $T_n = \frac{1}{2}(V_n - U_n)$: l'écart entre X_j et X_i va être maximal lorsque X_j est le plus grand possible et X_i le plus petit possible.

2.b. De la même manière qu'on a prouvé que $V_n \xrightarrow{P} \theta$, on prouverait que $U_n \xrightarrow{P} -\theta$.
Pour la culture, notons tout de même que les $-X_i$ suivent la même loi que les X_i et que $U_n = -\max(-X_1, \dots, -X_n)$.
Ce maximum a la même loi que V_n et donc converge en probabilité vers θ, de sorte que U_n converge en probabilité vers $-\theta$.

On a alors $T_n = \frac{1}{2}(V_n - U_n) \xrightarrow{P} \frac{1}{2}(\theta - (-\theta)) = \theta$.
Bien que ceci[1] soit classique, il n'est pas au programme, et il faut savoir le redémontrer...
Soit donc $\varepsilon > 0$. Alors, d'après l'inégalité triangulaire,

$$\left[|V_n - \theta| \leqslant \frac{\varepsilon}{2}\right] \cap \left[|U_n + \theta| \leqslant \frac{\varepsilon}{2}\right] \subset \left[|V_n - U_n - 2\theta| \leqslant \varepsilon\right].$$

Et donc en passant à l'événement contraire,

$$\left[|V_n - U_n - 2\theta| > \varepsilon\right] \subset \left[|V_n - \theta| > \frac{\varepsilon}{2}\right] \cup \left[|U_n + \theta| > \frac{\varepsilon}{2}\right].$$

Et par conséquent

$$P\left(|V_n - U_n - 2\theta| > \varepsilon\right) \leqslant P\left(|V_n - \theta| > \frac{\varepsilon}{2}\right) + P\left(|U_n + \theta| > \frac{\varepsilon}{2}\right).$$

Et donc enfin, par le théorème des gendarmes, il vient $\lim_{n\to+\infty} P\left(|V_n - U_n - 2\theta| > \varepsilon\right) = 0$.

□

[1] La limite en proba d'une somme est la somme des limites en probas.

Rappel
On a toujours
$P(A \cup B) \leqslant P(A) + P(B)$.

3.7.2 Estimation par intervalle de confiance

EXERCICE 3.68 (QSP HEC 2011) [HEC11.QSP01] — Facile

Intervalle de confiance du paramètre d'une loi exponentielle décalée.

On désire estimer un paramètre réel θ. Soit n un entier naturel non nul, et soient $X_1, \dots, X_n$ des variables aléatoires définies sur un même espace probabilisé $(\Omega, \mathcal{A}, P)$, indépendantes, et suivant la même loi de fonction de répartition F_θ donnée par

$$F_\theta(x) = \begin{cases} 1 - e^{\theta - x} & \text{si } x \geqslant \theta \\ 0 & \text{si } x < \theta \end{cases}$$

1. Déterminer la fonction de répartition de la variable aléatoire T_n donnée par

$$T_n = n(\mu_n - \theta) \text{ avec } \mu_n = \min(X_1, \dots, X_n)$$

et reconnaître la loi de T_n.

2. En déduire un intervalle de confiance pour θ au niveau de risque $\alpha \in]0, 1[$.

SOLUTION DE L'EXERCICE 3.68

1. Commençons par remarquer que les X_i ont pour support $[\theta, +\infty[$, et donc il en est de même de μ_n. Alors T_n prend ses valeurs dans $[0, +\infty[$, de sorte que $F_{T_n}(x) = 0$ si $x < 0$.
Pour $x \geqslant 0$, on a

$$F_{T_n}(x) = P(T_n \leqslant x) = P\left(\mu_n \leqslant \frac{x}{n} + \theta\right).$$

Or, par indépendance des X_i, il vient

$$P\left(\mu_n \leqslant \frac{x}{n} + \theta\right) = 1 - P\left(\mu_n > \frac{x}{n} + \theta\right) = 1 - \prod_{i=1}^{n} P\left(X_i > \frac{x}{n} + \theta\right) = 1 - \left(e^{-\frac{x}{n}}\right)^n = 1 - e^{-x}.$$

Ainsi, T_n suit une loi exponentielle de paramètre 1.

2. Remarquons que T_n n'est pas un estimateur de θ car son expression dépend précisément de θ, que l'on cherche à exprimer.

En revanche, μ_n est un estimateur de θ, et $\mu_n = \dfrac{T_n}{n} + \theta$ (cette expression est trompeuse car elle laisse entendre que μ_n est fonction de θ, ce qui n'est pas le cas).

Puisqu'on a toujours $\theta \leqslant \mu_n$, nous pouvons[1] chercher un intervalle de confiance sous la forme $[\mu_n - a, \mu_n]$.

On a

$$P(\mu_n - a \leqslant \theta \leqslant \mu_n) = P(\mu_n - a \leqslant \theta) = P(\mu_n - \theta \leqslant a) = P(T_n \leqslant na) = 1 - e^{-na}.$$

Si l'on veut $P(\mu_n - a \leqslant \theta \leqslant \mu_n) = 1 - \alpha$, on peut donc prendre $a = -\dfrac{\ln \alpha}{n}$.

Ainsi, un intervalle de confiance de θ au niveau de confiance $1 - \alpha$ est $\left[\mu_n + \dfrac{\ln \alpha}{n}, \mu_n\right]$. □

[1] Mais il n'y a aucune obligation : on pourrait également chercher un intervalle de confiance centré en μ_n.

Remarque

a est alors positif car $\alpha \in]0, 1[$.

Exercice 3.69 (QSP HEC 2009) [HEC09.QSP06] — Difficile

Intervalle de confiance asymptotique pour le paramètre d'une loi de Poisson.

Soient $n \geqslant 1$ et $(X_1, \ldots, X_n)$ un échantillon identiquement distribué indépendant de loi de Poisson de paramètre $\lambda > 0$ inconnu.

On pose $\overline{X_n} = \dfrac{1}{n}\sum_{i=1}^{n} X_i$ et $T_n = \sqrt{n}\dfrac{\overline{X_n} - \lambda}{\sqrt{\lambda}}$.

En utilisant T_n, déterminer un intervalle de confiance asymptotique de λ au niveau de risque α.

Solution de l'exercice 3.69 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

$$\lim_{n\to\infty} P(a \leqslant T_n \leqslant b) = P(a \leqslant X \leqslant b) = \Phi(b) - \Phi(a).$$

En particulier, soit t_α l'unique réel tel que $\Phi(t_\alpha) = 1 - \dfrac{\alpha}{2}$.

Alors

$$\lim_{n\to\infty} P\left(-t_\alpha \leqslant \sqrt{n}\frac{\overline{X_n} - \lambda}{\sqrt{\lambda}} \leqslant t_\alpha\right) = 1 - \alpha$$

soit encore

$$\lim_{n\to\infty} P\left(\lambda - \frac{t_\alpha}{\sqrt{n}}\sqrt{\lambda} \leqslant X_n \leqslant \lambda + \frac{t_\alpha}{\sqrt{n}}\sqrt{\lambda}\right) = 1 - \alpha.$$

Mais $\lambda - \dfrac{t_\alpha}{\sqrt{n}}\sqrt{\lambda} = \left(\sqrt{\lambda} - \dfrac{t_\alpha}{2\sqrt{n}}\right)^2 - \dfrac{t_\alpha^2}{4n}$, et de même, $\lambda + \dfrac{t_\alpha}{\sqrt{n}}\sqrt{\lambda} = \left(\sqrt{\lambda} + \dfrac{t_\alpha}{2\sqrt{n}}\right)^2 - \dfrac{t_\alpha^2}{4n}$.

Donc on a

$$\lim_{n\to\infty} P\left(\left(\sqrt{\lambda} - \frac{t_\alpha}{2\sqrt{n}}\right)^2 - \frac{t_\alpha^2}{4n} \leqslant \overline{X_n} \leqslant \left(\sqrt{\lambda} + \frac{t_\alpha}{2\sqrt{n}}\right)^2 - \frac{t_\alpha^2}{4n}\right) = 1 - \alpha.$$

Détails

Il s'agit de la mise sous forme canonique de polynômes de degré 2 en $\sqrt{\lambda}$.

Mais on a, au moins pour n suffisamment grand pour que $\sqrt{\lambda} - \dfrac{t_\alpha}{2\sqrt{n}} \geqslant 0$,

$$\begin{aligned}
\left[\left(\sqrt{\lambda} - \frac{t_\alpha}{2\sqrt{n}}\right)^2 - \frac{t_\alpha^2}{4n} \leqslant \overline{X_n} \leqslant \left(\sqrt{\lambda} + \frac{t_\alpha}{2\sqrt{n}}\right)^2 - \frac{t_\alpha^2}{4n}\right] &= \left[\sqrt{\lambda} - \frac{t_\alpha}{2\sqrt{n}} \leqslant \sqrt{\overline{X_n} + \frac{t_\alpha^2}{4n}}\right] \cap \left[\sqrt{\lambda} + \frac{t_\alpha}{2\sqrt{n}} \geqslant \sqrt{\overline{X_n} + \frac{t_\alpha^2}{2\sqrt{n}}}\right] \\
&= \left[\sqrt{\overline{X_n} + \frac{t_\alpha^2}{4n}} - \frac{t_\alpha}{2\sqrt{n}} \leqslant \sqrt{\lambda} \leqslant \sqrt{\overline{X_n} + \frac{t_\alpha^2}{4n}} + \frac{t_\alpha}{2\sqrt{n}}\right] \\
&= \left[\left(\sqrt{\overline{X_n} + \frac{t_\alpha^2}{4n}} - \frac{t_\alpha}{2\sqrt{n}}\right)^2 \leqslant \lambda \leqslant \left(\sqrt{\overline{X_n} + \frac{t_\alpha^2}{4n}} + \frac{t_\alpha}{2\sqrt{n}}\right)^2\right].
\end{aligned}$$

Et donc on a

$$\lim_{n\to\infty} P\left(\left(\sqrt{\overline{X_n}+\frac{t_\alpha^2}{4n}}-\frac{t_\alpha}{2\sqrt{n}}\right)^2 \leqslant \lambda \leqslant \left(\sqrt{\overline{X_n}+\frac{t_\alpha^2}{4n}}+\frac{t_\alpha}{2\sqrt{n}}\right)^2\right) = 1-\alpha.$$

Ainsi, un intervalle de confiance asymptotique de λ au niveau de confiance $1-\alpha$ est

$$\left[\left(\sqrt{\overline{X_n}+\frac{t_\alpha^2}{4n}}-\frac{t_\alpha}{2\sqrt{n}}\right)^2, \left(\sqrt{\overline{X_n}+\frac{t_\alpha^2}{4n}}+\frac{t_\alpha}{2\sqrt{n}}\right)^2\right].$$

□

Méthode

Le but de ce calcul est d'isoler λ, afin d'obtenir un intervalle de confiance pour λ.

EXERCICE 3.70 (HEC 2012) [HEC12.28] — Facile

Intervalle de confiance pour le paramètre d'une loi de Bernoulli via l'inégalité de Chernoff

Soit p un paramètre réel inconnu vérifiant $0 < p < 1$. Pour $n \in \mathbf{N}^*$ soit $X_1, X_2, \ldots, X_n$ n variables aléatoires définies sur un espace probabilisé $(\Omega, \mathcal{A}, P)$, indépendantes et de même loi de Bernoulli de paramètre p.

On pose, pour tout $n \in \mathbf{N}^*$, $\overline{X_n} = \dfrac{1}{n}\sum_{k=1}^{n} X_k$.

1. Rappeler comment l'inégalité de Bienaymé-Tchebychev permet de définir à partir de $\overline{X_n}$ un intervalle de confiance de p au niveau de risque α.
2. Soit f la fonction de $\mathbf{R}$ dans $\mathbf{R}$ définie par : $f(t) = -pt + \ln(1-p+pe^t)$.
   (a) Montrer que f est de classe $\mathcal{C}^2$ sur $\mathbf{R}$ et que sa dérivée seconde vérifie : $\forall t \geqslant 0,\ f''(t) \leqslant \dfrac{1}{4}$.
   (b) Montrer à l'aide d'une formule de Taylor que pour tout $t \geqslant 0$, on a : $f(t) \leqslant \dfrac{t^2}{8}$.
   (c) En déduire que pour tout $t \geqslant 0$, et pour tout $k \in [\![1, n]\!]$, on a : $E\left(\exp\left(t\left(X_k - p\right)\right)\right) \leqslant \exp\left(\dfrac{t^2}{8}\right)$.
3. (a) Montrer que si S est une variable aléatoire discrète finie à valeurs positives et a un réel strictement positif, on a $P(S \geqslant a) \leqslant \dfrac{E(S)}{a}$.
   (b) À l'aide des questions 2.c et 3.a, établir pour tout couple $(t, \varepsilon) \in (\mathbf{R}_+)^2$, l'inégalité :
   $$P\left(\overline{X_n} - p \geqslant \varepsilon\right) \leqslant \exp\left(-t\varepsilon + \frac{t^2}{8n}\right).$$
   (c) Montrer que pour tout $\varepsilon > 0$, on a : $P\left(\left|\overline{X_n} - p\right| \geqslant \varepsilon\right) \leqslant 2\exp\left(-2n\varepsilon^2\right)$.
   (d) En déduire un intervalle de confiance de risque α pour le paramètre p et comparer sa longueur, lorsque α est proche de 0, à celle de l'intervalle de confiance demandé dans la question 1.

SOLUTION DE L'EXERCICE 3.70

1. Il est classique que, par indépendance des X_i, $E(\overline{X_n}) = p$ et $V(\overline{X_n}) = \dfrac{p(1-p)}{n} \leqslant \dfrac{1}{4n}$.
Et donc, par l'inégalité de Bienaymé-Tchebychev, pour tout $\varepsilon > 0$,

$$P\left(|\overline{X_n} - p| > \varepsilon\right) \leqslant \frac{1}{4n\varepsilon^2}.$$

Et donc en passant à l'événement contraire,

$$P\left(|\overline{X_n} - p| \leqslant \varepsilon\right) \geqslant 1 - \frac{1}{4n\varepsilon^2}.$$

Mais $\left[|\overline{X_n} - p| \leqslant \varepsilon\right] = \left[\overline{X_n} - \varepsilon \leqslant p \leqslant \overline{X_n} + \varepsilon\right]$.
Ainsi, $P\left(\overline{X_n} - \varepsilon \leqslant p \leqslant \overline{X_n} + \varepsilon\right) \geqslant 1 - \dfrac{1}{4n\varepsilon^2}$.

Or, nous avons là un intervalle de confiance de p au niveau de risque $\frac{1}{4n\varepsilon^2}$.

En particulier, pour $\frac{1}{4n\varepsilon^2} \leqslant \alpha \Leftrightarrow \varepsilon \geqslant \frac{1}{2\sqrt{n\alpha}}$, on a un intervalle de confiance de p au niveau de risque α.

Et donc $\left[\overline{X_n} - \frac{1}{2\sqrt{n\alpha}}, \overline{X_n} + \frac{1}{2\sqrt{n\alpha}}\right]$ est un intervalle de confiance de p au niveau de risque α.

2.a. f est de classe $\mathscr{C}^2$ par opérations sur les fonctions usuelles, et

$$f'(t) = -p + \frac{pe^t}{1-p+pe^t} = -p + \frac{p}{(1-p)e^{-t}+p},\ f''(t) = \frac{p(1-p)}{((1-p)e^{-t}+p)^2}.$$

Or, pour $t \geqslant 0$, $(1-p)e^{-t} + p \geqslant 1 - p + p = 1$ et donc $f''(t) \leqslant p(1-p)$.

Il est classique que sur $[0,1]$, la fonction $x \mapsto x(1-x)$ admet un maximum égal à $\frac{1}{4}$ et donc $f''(t) \leqslant \frac{1}{4}$.

2.b. Notons que $f(0) = \ln(1) = 0$ et $f'(0) = -p + \frac{p}{1-p+p} = 0$.
Donc par la formule de Taylor avec reste intégral, pour tout $t \geqslant 0$,

$$f(t) = f(0) + f'(0)t + \int_0^t \frac{f''(x)}{1!}(t-x)\,dx = \int_0^t \frac{f''(x)}{1!}(t-x)\,dx.$$

Mais grâce à l'inégalité précédemment obtenue, et par croissance de l'intégrale,

$$\int_0^t f''(x)(t-x)\,dt \leqslant \int_0^t \frac{1}{4}(t-x)\,dx = \left[-\frac{(t-x)^2}{8}\right]_0^t = \frac{t^2}{8}.$$

Et donc $f(t) \leqslant \frac{t^2}{8}$.

2.c. D'après le théorème de transfert, on a

$$E\left(\exp\left(t\left(X_k - p\right)\right)\right) = \exp(t(-p))P(X_k = 0) + \exp(t(1-p))P(X_k = 1) = e^{-pt}\left((1-p) + pe^t\right) = e^{f(t)} \leqslant e^{\frac{t^2}{8}}.$$

3.a. C'est simplement l'inégalité de Markov, qui s'applique puisque S est à valeurs positives, et possède une espérance puisqu'il s'agit d'une variable finie.

3.b. Par croissance de l'exponentielle, on a $P(\overline{X_n} - \varepsilon \geqslant p) = P\left(t(\overline{X_n} - p) \geqslant t\varepsilon\right) = P\left(e^{t(\overline{X_n}-p)} \geqslant e^{t\varepsilon}\right)$.
Mais $e^{t(\overline{X_n}-p)}$ est une variable aléatoire à valeurs finies (puisque $\overline{X_n}$ ne prend qu'un nombre fini de valeurs) et positive. Donc la question 3.a s'applique :

$$P(\overline{X_n} - p \geqslant \varepsilon) \leqslant \frac{E\left(\exp\left(t(\overline{X_n} - p)\right)\right)}{e^{t\varepsilon}}.$$

Par indépendance des X_i, on a

$$E\left(\exp(t(\overline{X_n} - p))\right) = E\left(\prod_{k=1}^n \exp\left(\frac{t}{n}(X_k - p)\right)\right) = \prod_{k=1}^n E\left(\exp\left(\frac{t}{n}(X_k - p)\right)\right) \leqslant \left(\exp\left(\frac{t^2}{8n^2}\right)\right)^n \leqslant \exp\left(\frac{t^2}{8n}\right).$$

Et donc

$$P(\overline{X_n} - p \geqslant \varepsilon) \leqslant \exp\left(-t\varepsilon + \frac{t^2}{8n}\right).$$

3.c. Notons que les $1 - X_i$ suivent la loi de Bernoulli de paramètre $1-p$, et sont encore indépendantes.
En appliquant le même raisonnement aux $1 - X_i$, on obtient alors

$$P\left(\frac{1}{n}\sum_{k=1}^n (1 - X_k) - (1-p) \geqslant \varepsilon\right) \leqslant \exp\left(-t\varepsilon + \frac{t^2}{8n}\right).$$

Mais

$$\left[\frac{1}{n}\sum_{k=1}^{n}(1-X_k)-(1-p)\geqslant\varepsilon\right]=\left[1-\overline{X_n}-1+p\geqslant\varepsilon\right]=\left[\overline{X_n}-p\leqslant-\varepsilon\right].$$

Et donc

$$P(|\overline{X_n}-p|\geqslant\varepsilon)=P(\overline{X_n}-p\geqslant\varepsilon)+P(\overline{X_n}-p\leqslant-\varepsilon)\leqslant 2\exp\left(-t\varepsilon+\frac{t^2}{8n}\right).$$

Cette inégalité est alors valable pour tout $t\geqslant 0$.

Mais sur $\mathbf{R}_+$, la fonction $t\mapsto -t\varepsilon+\dfrac{t^2}{8n}$ admet un minimum en $t=4n\varepsilon$, et ce maximum vaut $-4n\varepsilon^2+\dfrac{16n^2\varepsilon^2}{8n}=-2n\varepsilon^2$.

Et donc, par croissance de l'exponentielle,

$$P(|\overline{X_n}-p|\geqslant\varepsilon)\leqslant 2\exp(-2n\varepsilon^2).$$

Détails

C'est la formule vue en première pour l'extremum d'une fonction polynomiale de degré 2 :

$$t\mapsto at^2+bt+c$$

admet un extremum en $t=-\dfrac{b}{2a}$.

3.d. Comme à la question 1, $\left[\overline{X_n}-\varepsilon,\overline{X_n}+\varepsilon\right]$ est un intervalle de confiance de p au niveau de risque $2\exp(-2n\varepsilon^2)$.

Et donc en particulier, pour $2\exp(-2n\varepsilon^2)\leqslant\alpha\Leftrightarrow\varepsilon\geqslant\sqrt{-\dfrac{\ln\left(\frac{\alpha}{2}\right)}{2n}}$, on dispose d'un intervalle de confiance de p au niveau de risque α.

Ainsi, $\left[\overline{X_n}-\sqrt{-\dfrac{\ln\left(\frac{\alpha}{2}\right)}{2n}},\overline{X_n}+\sqrt{-\dfrac{\ln\left(\frac{\alpha}{2}\right)}{2n}}\right]$ est l'intervalle de confiance cherché.

La longueur de cet intervalle est $2\sqrt{-\dfrac{\ln\left(\frac{\alpha}{2}\right)}{2n}}$ alors que celle de l'intervalle de confiance obtenu dans la question 1 est $\dfrac{1}{\sqrt{n\alpha}}$.

Lorsque $\alpha\to 0$, on a $\ln(\alpha)=\underset{\alpha\to 0}{o}\left(\dfrac{1}{\alpha}\right)$ et donc l'intervalle que nous venons d'obtenir est beaucoup plus petit que celui obtenu via Bienaymé-Tchebychev.

□

Remarque

Ces deux longueurs tendent vers $+\infty$ lorsque α tend vers 0.
Si le fait qu'elles soient fonction croissante de α est logique (si on veut minimiser le risque, il faut augmenter la taille de l'intervalle), le fait qu'elles tendent vers $+\infty$ est le signe que ces intervalles ne sont pas optimaux. En effet, $p\in[0,1]$, alors que lorsque $\alpha\to 0$, nos deux intervalles deviennent bien plus grands que $[0,1]$!

3.8 A TRIER

EXERCICE 3.71 (HEC 2015) [HEC15.149]

Méthode de Monte-Carlo et réduction de la variance par des variables antithétiques.

Dans tout l'exercice, U désigne une variable aléatoire à densité suivant la loi uniforme sur $[0,1]$.

Soit g une fonction continue de $[0,1]$ dans $\mathbf{R}_+$. On pose $I=\displaystyle\int_0^1 g(t)\,dt$ et $J=\int_0^1 g^2(t)\,dt$.

1. Pour $n\in\mathbf{N}^*$, on définit la variable aléatoire T_n par : $T_n=\dfrac{1}{n}\lfloor 1+nU\rfloor$.
   (a) Déterminer la loi de T_n.
   (b) Calculer $\lim\limits_{n\to+\infty}E(g(T_n))$ en fonction de I.
   (c) Calculer $\lim\limits_{n\to+\infty}V(g(T_n))$ en fonction de I et de J.
2. On pose $X=g(U)$ et $Y=\dfrac{1}{2}(g(U)+g(1-U))$.
   (a) Calculer $E(X)$ et $E(Y)$ en fonction de I.
   (b) Soient f et h deux fonctions continues de $[0,1]$ dans $\mathbf{R}$. On pose, pour tout réel λ :

$$Q(\lambda)=\int_0^1(\lambda f(t)-h(t))^2\,dt.$$

Établir l'inégalité : $\left(\int_0^1 f(t)h(t)\,dt\right)^2 \leqslant \left(\int_0^1 f^2(t)\,dt\right)\times\left(\int_0^1 h^2(t)\,dt\right)$.

(c) En déduire que $V(Y) \leqslant V(X)$.

3. On suppose dans cette question que la fonction g est croissante sur $[0,1]$. Pour $n \in \mathbf{N}^*$, on considère deux variables aléatoires U_n et W_n indépendantes et de même loi que la variable aléatoire T_n de la question 1.

(a) Justifier l'inégalité :
$$E([g(U_n)-g(W_n)][g(1-U_n)-g(1-W_n)]) \leqslant 0.$$

(b) En déduire que $E(g(T_n)g(1-T_n)) \leqslant E(g(T_n))E(g(1-T_n))$.

4. Montrer que $V(Y) \leqslant \frac{1}{2}V(X)$.

Solution de l'exercice 3.71

1.a. Il est clair que $T_n(\Omega) = \left\{\frac{1}{n}, \frac{2}{n}, \dots, 1\right\}$, et pour tout $k \in [\![1,n]\!]$,

$$P\left(T_n = \frac{k}{n}\right) = P\left(\lfloor 1+nU\rfloor = k\right) = P(k \leqslant 1+nU < k+1) = P\left(\frac{k-1}{n} \leqslant U < \frac{k}{n}\right) = \frac{k}{n} - \frac{k-1}{n} = \frac{1}{n}.$$

Autrement dit, T_n suit la loi uniforme sur $\left\{\frac{1}{n}, \frac{2}{n}, \dots, 1\right\}$.

1.b. Par le théorème de transfert, on a

$$E(g(T_n)) = \sum_{k=1}^{n} \frac{1}{n} g\left(\frac{k}{n}\right) = \frac{1}{n}\sum_{k=1}^{n} g\left(\frac{k}{n}\right).$$

Puisque g est continue sur $[0,1]$, nous reconnaissons là une somme de Riemann :

$$E(g(T_n)) \underset{n\to+\infty}{\longrightarrow} \int_0^1 g(t)\,dt = I.$$

1.c. De même, le théorème de transfert prouve que $E\left(g(T_n)^2\right) \underset{n\to+\infty}{\longrightarrow} \int_0^1 g^2(t)\,dt = J$.

Et donc par la formule de Huygens,

$$V(g(T_n)) = E\left(g(T_n)^2\right) - E(g(T_n))^2 \underset{n\to+\infty}{\longrightarrow} J - I^2.$$

2.a. Par le théorème de transfert, on a

$$E(g(U)) = \int_{-\infty}^{+\infty} g(t)f_U(t)\,dt = \int_0^1 g(t)\,dt = I.$$

D'autre part, puisque $1-U$ suit la même loi que U, $g(1-U)$ suit la même loi que $g(U) = X$ et donc

$$E(Y) = \frac{1}{2}\left(E(g(U)) + E(g(1-U))\right) = E(X) = I.$$

2.b. Il s'agit de l'inégalité de Cauchy-Schwarz pour le produit scalaire $\langle f,g\rangle = \int_0^1 f(t)g(t)\,dt$.

Puisque l'énoncé semble en donner une preuve, donnons la : la fonction Q est positive sur $\mathbf{R}$, car intégrale d'une fonction positive.
Mais pour tout $\lambda \in \mathbf{R}$,

$$Q(\lambda) = \lambda^2\int_0^1 f^2(t)\,dt - 2\lambda\int_0^1 f(t)h(t)\,dt + \int_0^1 h^2(t)\,dt.$$

Il s'agit donc d'une fonction polynomiale de degré 2, qui ne change pas de signe sur $\mathbf{R}$.
Son discriminant est donc positif ou nul. Et donc

$$4\left(\int_0^1 f(t)h(t)\,dt\right)^2 - 4\left(\int_0^1 f^2(t)\,dt\right)\times\left(\int_0^1 h^2(t)\,dt\right) \leqslant 0 \Leftrightarrow \left(\int_0^1 f(t)h(t)\,dt\right)^2 \leqslant \left(\int_0^1 f^2(t)\,dt\right)\times\left(\int_0^1 h^2(t)\,dt\right).$$

2.c. On a

$$V(Y) = \frac{1}{4}\left(V(g(U)) + V(g(1-U)) + 2\operatorname{Cov}(g(U), g(1-U))\right) = \frac{1}{2}V(X) + \frac{1}{2}\operatorname{Cov}\left(g(U), g(1-U)\right).$$

Par la formule de Huygens,

$$\operatorname{Cov}(g(U), g(1-U)) = E(g(U)g(1-U)) - E(g(U))E(g(1-U)) = E(g(U)g(1-U)) - I^2.$$

Le théorème de transfert nous donne alors

$$\operatorname{Cov}(g(U), g(1-U)) = \int_0^1 g(t)g(1-t)\,dt - I^2.$$

Or, le résultat de la question précédente, appliqué aux fonction $t \mapsto g(t)$ et $t \mapsto g(1-t)$ affirme que

$$\left(\int_0^1 g(t)g(1-t)\,dt\right)^2 dt \leqslant \left(\int_0^1 g(t)^2\,dt\right) \times \left(\int_0^1 g(1-t)^2\,dt\right).$$

Un changement de variable $x = 1-t$ montre que

$$\int_0^1 g(1-t)^2\,dt = \int_0^1 g^2(x)\,dx = J.$$

Et donc, après passage à la racine carrée[1]

$$\int_0^1 g(t)g(1-t)\,dt \leqslant \sqrt{J^2} = J.$$

[1] Toutes les quantités en jeu sont positives.

Il vient donc

$$\operatorname{Cov}(g(U), g(1-U)) \leqslant J - I^2 = V(X).$$

On en déduit donc que

$$V(Y) = \frac{1}{2}V(X) + \frac{1}{2}\operatorname{Cov}(g(U), g(1-U)) \leqslant V(X).$$

3.a. $(1-U_n) - (1-W_n) = W_n - U_n$ est donc de signe opposé à $U_n - W_n$.
Et puisque g est croissante, $g(1-U_n) - g(1-W_n)$ est donc de signe opposé à $g(U_n) - g(W_n)$.
Et donc $[g(U_n) - g(W_n)][g(1-U_n) - g(1-W_n)] \leqslant 0$.
Par positivité de l'espérance, on a donc

$$E\left([g(U_n) - g(W_n)][g(1-U_n) - g(1-W_n)]\right) \leqslant 0.$$

Détails
Une fonction croissante est une fonction qui préserve le sens des inégalités.

3.b. Développons l'espérance précédente :

$$E\left([g(U_n) - g(W_n)][g(1-U_n) - g(1-W_n)]\right) = E(g(U_n)g(1-U_n)) - E(g(U_n)g(1-W_n)) - E(g(1-U_n)g(W_n)) + E(g(W_n)g(1-W_n)).$$

Mais par indépendance de U_n et W_n, on a

$$E(g(U_n)g(1-W_n)) = E(g(U_n))E(g(1-W_n)) = E(g(T_n))E(g(1-T_n)) \text{ et } E(g(1-U_n)g(W_n)) = E(g(1-U_n))E(g(W_n)) == E(g(T_n))E(g(1-T_n)).$$

D'autre part, $g(U_n)g(1-U_n)$ et $g(W_n)g(1-W_n)$ sont identiquement distribuées et ont donc même espérance, qui est également la même qe celle de $g(T_n)g(1-T_n)$. Il vient alors

$$2E(g(1-T_n)g(1-T_n)) - 2E(g(T_n))E(g(1-T_n)) \leqslant 0 \Leftrightarrow E(g(T_n)g(1-T_n)) \leqslant E(g(T_n))E(g(1-T_n)).$$

4. Posons $Y_n = \frac{1}{2}(g(T_n) + g(1-T_n))$.
Alors $V(Y_n) = \frac{1}{4}\left(E(g(T_n)) + E(g(1-T_n) + 2\operatorname{Cov}(g(T_n), g(1-T_n))\right)$.
Alors la formule de Huygens nous donne

$$\operatorname{Cov}(g(T_n)g(1-T_n)) = E(g(T_n)g(1-T_n)) - E(g(T_n))E(g(1-T_n)) \leqslant 0.$$

Et donc $V(Y_n) \leqslant \frac{1}{4}\left(V(g(T_n) + V(g(1-T_n))\right)$.
Comme à la question 1.c, on prouve que $\lim\limits_{n\to+\infty} V(Y_n) = V(Y)$ et que $\lim\limits_{n\to+\infty} V(g(1-T_n)) = V(X)$. Et donc

$$V(Y) = \lim_{n\to+\infty} V(T_n) \leqslant \lim_{n\to+\infty} \frac{1}{4}\left(V(g(T_n)) + V(g(1-T_n))\right) = \frac{1}{2}V(X).$$

□

Exercice 3.72 (QSP ESCP 2010) [ESCP10.QSP02] **Moyen**

Une urne contient $4n+2$ boules numérotées de 1 à $4n+2$. On tire sans remise $2n+1$ boules.
Quelle est la probabilité que la somme des numéros des boules tirées soit strictement supérieure à la somme des numéros des boules restées dans l'urne ?

Solution de l'exercice 3.72 Notons que la somme totale des numéros portés par les boules de l'urne vaut $\sum_{k=1}^{4n+2} k = \frac{(4n+2)(4n+3)}{2} = (2n+1)(4n+3)$.
Si on note X la somme des numéros portés par les boules tirées, on cherche donc $P(X > (2n+1)(4n+3)-X) = P\left(X > \frac{(2n+1)(4n+3)}{2}\right)$.
Nous pourrions être tentés d'essayer de déterminer la loi de X, mais on se rendrait vite compte que ceci est bien trop difficile[1].

[1] Vous pouvez essayer de le faire pour $n=1$, c'est-à-dire pour une urne contenant 6 boules, ce sera déjà fastidieux, et il semble difficile de mettre en place une stratégie qui se transposerait au cas général.

Notons plutôt qu'il y a trois cas possibles : soit la somme des boules tirées est strictement supérieure à celle des boules restantes, soit les deux sommes sont égales, soit la somme des boules restantes est strictement supérieure à celle des boules tirées. Et donc

$$1 = P(X > (2n+1)(4n+3)-X) + P(X = (2n+1)(4n+3)-X) + P(X < (2n+1)(4n+3)-X).$$

Mais $P(X = (2n+1)(4n+3)-X) = P\left(X = \frac{(2n+1)(4n+3)}{2}\right) = 0$ car $(2n+1)(4n+3)$ est impair[2] et donc $\frac{(2n+1)(4n+3)}{2}$ n'est pas entier, de sorte que $P(X = (2n+1)(4n+3)-X) = 0$.
D'autre part, X et $(2n+1)(4n+3)-X$ suit la même loi que X, puisque choisir les $2n+1$ boules que l'on tire revient à choisir les $2n+1$ boules que l'on laisse dans l'urne.
Et donc $P\left(X > \frac{(2n+1)(4n+3)}{2}\right) = P\left(X < \frac{(2n+1)(4n+3)}{2}\right)$.
On en déduit donc que

$$2P\left(X > \frac{(2n+1)(4n+3)}{2}\right) = 1 \Leftrightarrow P\left(X > \frac{(2n+1)(4n+3)}{2}\right) = \frac{1}{2}.$$

□

[2] Puisque produit de deux nombres impairs.

Détails
$(2n+1)(4n+3)-X$ est la somme des numéros portés par les boules restant dans l'urne.

4 Scilab

Nous regroupons dans cette partie quelques exercices portant spécifiquement sur `Scilab` , posés ces dernières années. La plupart d'entre eux relèvent du programme de probabilités.

Exercice 4.1 (QSP HEC 2016) [HEC16.QSP173] — Facile

Approximation de π par la méthode de Monte-Carlo

Donner la finalité du programme suivant

```
N = 100000 ;S = 0 ;
for i=1 :N
        u=rand() ;
        S = S+4/N*1/(1+u^2) ;
end
disp(S)
```

On pourra penser à la loi faible des grands nombres.

Solution de l'exercice 4.1 Ce programme simule à chaque passage dans la boucle une variable aléatoire X_i suivant la loi uniforme sur $[0,1]$, à l'aide de `rand`.
À la fin de la boucle, la variable `S` contient alors

$$\frac{1}{N}\sum_{i=1}^{N}\frac{4}{1+X_i^2}.$$

Or, lors de différents appels à `rand()`, on simule des variables X_i qui sont indépendantes.
Donc les $\dfrac{4}{1+X_i^2}$ sont indépendantes, et de même loi[1]
Par la loi faible des grands nombres[2], on a donc

$$\frac{1}{N}\sum_{i=1}^{N}\frac{4}{1+X_i^2}\xrightarrow[N\to+\infty]{P}E\left(\frac{4}{1+X_1^2}\right).$$

Soit donc U une variable aléatoire suivant la loi uniforme sur $[0,1]$.
Par le théorème de transfert, on a

$$E\left(\frac{4}{1+U^2}\right)=\int_0^1\frac{4}{1+x^2}\,dx=[4\operatorname{Arctan}(x)]_0^1=4\operatorname{Arctan}(1)=\pi.$$

De plus, $\dfrac{4}{1+U^2}$ admet bien un moment d'ordre 2 puisque $\displaystyle\int_0^1\left(\frac{4}{1+x^2}\right)^2dx$ converge en tant qu'intégrale sur un segment.
Donc la loi faible des grands nombres s'applique bien, et donc

$$\frac{1}{N}\sum_{i=1}^{N}\frac{4}{1+X_i^2}\xrightarrow{P}\pi.$$

Et donc pour N grand[3], le programme doit retourner une valeur approchée de π.
Quelques exécutions du programme affichent successivement 3.1413348, 3.1448606, 3.141795, ce qui semble bien confirmer ce qui vient d'être dit. □

[1] Car les X_i le sont.

[2] Sous réserve qu'elle s'applique.

Remarque

Notons que seules les deux premières décimales sont bonnes, ce qui ne nous apprend pas grand chose, puisque nous savons déjà depuis tout petits que $\pi\simeq 3.14$.

EXERCICE 4.2 (QSP HEC 2016) [HEC16.QSP179] — Moyen

Modélisation de tirages avec remise dans une urne

On tire avec remise, une boule dans une urne contenant n boules numérotées.

1. On note T la variable aléatoire égale au numéro du tirage où pour la première fois deux boules différentes ont été tirées.
   Déterminer l'espérance de T.
2. Quelle est la variable aléatoire V_n dont la fonction Scilab suivante calcule une simulation ?

```
1  function compt=V(n)
2          A = [] ; compt =0 ;
3          while length(A)<n
4                  u = floor(n*rand()+1) ;
5                  i = find(A==u) ; // renvoie les positions de u dans le vecteur A
6                  if length(i)==0 then
7                          A = [A,u] ;
8                  end
9                  compt=compt+1 ;
10         end
11 endfunction
```

SOLUTION DE L'EXERCICE 4.2

1. Bien que l'énoncé ne dise rien à ce sujet, nous pouvons considérer que les boules sont numérotées de 1 à n.
Pour $i \in [\![1,n]\!]$, notons A_i l'événement «la boule obtenue lors du premier tirage est la boule i».
Alors la loi conditionnelle de $T-1$ sachant A_i est une loi géométrique de paramètre $\dfrac{n-1}{n}$.

Détails

À chaque tirage, la probabilité de ne pas obtenir la boule i vaut $\dfrac{n-1}{n}$, et les tirages ayant lieu avec remise, ils sont indépendants.

En particulier, $E(T-1|A_i) = \dfrac{n}{n-1}$.
Par la formule de l'espérance totale appliquée au système complet d'événements $\{A_1,\dots,A_n\}$, on a

$$E(T-1) = \sum_{i=1}^{n} P(A_i)E(T-1|A_i) = \sum_{i=1}^{n} \frac{1}{n}\frac{n}{n-1} = \frac{n}{n-1}.$$

Et donc $E(T) = 1 + \dfrac{n}{n-1}$.

2. La ligne 4 permet de simuler le tirage d'une boule dans l'urne, puisqu'elle simule une loi uniforme sur $[\![1,n]\!]$.
Les lignes 5 à 7 indiquent que si on a obtenu un numéro qui n'était pas dans A, alors on l'ajoute à A.
Ainsi, à tout instant, A contient la liste des différents numéros obtenus lors des précédents tirages, chaque nombre ne pouvant y figurer au maximum qu'une fois[1].
Et la boucle while est parcourue jusqu'à ce que A soit de taille n, c'est-à-dire jusqu'à ce que tous les numéros entre 1 et n aient été obtenus au moins une fois chacun.
Et alors compt contient le nombre de passages dans la boucle, c'est-à-dire le nombre de tirages qui ont été nécessaires pour obtenir au moins une fois chaque boule.
Et donc V_n représente le nombre de tirages nécessaires pour obtenir au moins une fois chaque boule de l'urne.

[1] Si un nombre est déjà dans A et qu'on le tire de nouveau, on ne l'ajoute pas à A.

□

L'exercice qui suit est assez classique, mais le code fourni par l'énoncé est assez dense, et le comprendre «à la volée» demande une certaine aisance.
On notera qu'il y a là une commande (gsort) qui ne figure pas au programme officiel, mais dont le fonctionnement est expliqué (de manière succincte, mais largement compréhensible).

Exercice 4.3 (QSP HEC 2015) [HEC15.QSP143]

Problème des anniversaires
Abordable en première année : ✓

Dans une classe de 30 élèves, on considère une expérience consistant d'abord à demander à chaque élève sa date d'anniversaire. La suite de l'expérience (simulée 1000 fois sur ordinateur) est décrite par le programme Scilab suivant :

```
1  Nexp = 1000 ; Neleve=30 ;test=0 ;
2  for n=1 :Nexp
3          anniv=zeros(Neleve,1) ;
4          for i=1 :Neleve
5                  anniv(i) = floor(365*rand())+1 ;
6          end
7          anniv = gsort(anniv) ;ok=0 ; // gsort= tri par ordre croissant
8          for j=1 :Neleve-1
9                  if anniv(j) == anniv(j+1) then
10                         ok=1 ;
11                 end
12         end
13         test=test+ok ;
14 end
15 disp(test/Nexp) ;
```

1. Le code retourne une valeur (à chaque fois différente) autour de 0.71. Que représente cette valeur ?
2. Calculer la valeur exacte de la probabilité simulée par ce programme.
3. Écrire un programme Scilab permettant de déterminer le nombre d'élèves à partir duquel cette valeur dépasse 0.5.

Solution de l'exercice 4.3

1. Les lignes 3 à 5 ont pour effet de choisir au hasard la date de naissance de chacun des élèves (qui sont au nombre de Neleve, fixé ici à 30).
La liste anniv ainsi obtenue est ensuite triée à la ligne 7. Elle est ensuite parcourue grâce à la boucle de la ligne 8, et le code des lignes 9 à 11 a pour effet, de donner à ok la valeur 1 si la même date apparait deux fois dans anniv.
Enfin, la variable test compte le nombre de fois où, lors des Nexp répétitions de l'expérience, la variable ok valait 1.
Puisqu'enfin on divise test par Nexp, la valeur affichée par le programme correspond environ à la probabilité qu'il y ait deux élèves avec la même date d'anniversaire dans la classe. Cette probabilité est donc visiblement proche de 0.71.

365 ?
On suppose donc qu'aucun de ces élèves n'est né un 29 février.

Remarque
Le fait que anniv soit triée par ordre croissant implique que si la même date apparaît deux fois dans anniv, ce sera nécessairement à deux positions consécutives.

2. Il y a 365^{30} possibilités pour les choix des dates d'anniversaires des trente élèves.
Pour que tous ces 30 élèves aient des dates d'anniversaire différentes, il y a 365 choix possibles pour la date d'anniversaire du premier, puis 364 choix pour la date d'anniversaire du second, etc, 336 choix pour la date d'anniversaire du dernier.
Donc la probabilité que tous les élèves aient une date d'anniversaire différente vaut $\dfrac{365\times 364\times\cdots\times 336}{365^{30}} = \dfrac{365!}{365^{30}\times 335!}$.
Et donc la probabilité cherchée est, en passant à l'événement contraire,

$$1-\frac{365!}{365^{30}\times 335!} = 1-\prod_{i=1}^{30}\frac{365-i+1}{365}.$$

Proba/fréquence
Puisque nous réalisons ici des simulations, nous ne manipulons pas vraiment des probabilités, mais plutôt leur alter-ego statistique : des fréquences.
Nous répétons donc l'expérience jusqu'à ce que la fréquence des classes avec deux élèves nés le même jour soit supérieure à 1/2.

3. Donnons trois solutions à cette question. La première, peut-être la plus facile, nécessite juste une petite adaptation du code donné précédemment, pour répéter des simulations sur des classes à 2 élèves, à 3 élèves, etc, jusqu'à obtenir une probabilité supérieure ou égale à 0.5.

```
Nexp = 10000 ;
Neleve=1 ;proba=0 ;
while proba<0.5
        Neleve=Neleve+1 ;
        test=0 ;
        for n=1 :Nexp
                anniv=zeros(Neleve,1) ;
                for i=1 :Neleve
                        anniv(i) = floor(365*rand())+1 ;
                end
                anniv = gsort(anniv) ;ok=0 ;
                for j=1 :Neleve-1
                        if anniv(j) == anniv(j+1) then
                                ok=1 ;
                        end
                end
                test=test+ok ;
        end
        proba =test/Nexp ;
end
disp(Neleve) ;
```

Le principal inconvénient de ce code est qu'il repose sur des simulations, et donc qu'on n'aura aucune garantie quant à la valeur retournée.
Par exemple, en l'exécutant plusieurs fois, j'ai obtenu les valeurs 22, 23 et 24.
Il est sûrement possible d'augmenter Nexp, le nombre de répétitions, mais le temps de calcul risque d'augmenter d'autant.

Nous proposons plutôt d'utiliser le résultat de la question précédente : pour une classe de N élèves, la probabilité que deux élèves aient la même date d'anniversaire vaut

$$1-\prod_{i=1}^{N}\frac{365-i+1}{365}.$$

Factorielles
L'expression avec les factorielles n'est pas utilisable, puisque $\frac{365!}{365^{30}}$ est trop grand pour que Scilab parvienne à le calculer.

Voici une suggestion de code :

```
N=1 ;
p = 1 ;
while p>0.5
        p = p*(365-N)/365 ;
        N = N+1 ;
end
disp(N)
```

Ce programme retourne alors $N = 23$: dans une classe de 23 élèves, il y a plus d'une chance sur deux que deux étudiants soient nés le même jour.

Enfin, juste pour la beauté du geste, proposons une solution en une ligne, qui repose sur le même principe que la solution précédente, ainsi que sur le fait que nous savons déjà que $N < 30$:

```
disp(min(find(cumprod([365 :-1 :336]./(365*ones(1,30)))<0.5)))
```

Nous laissons les détails au lecteur, mais à part le fait qu'elle ne tienne qu'en une ligne, cette solution n'est pas plus intéressante que la précédente[1].

[1] Elle demande même plus de calculs !

□

Exercice 4.4 (QSP HEC 2018) [HEC18.QSP248] — Moyen

Autour de la loi binomiale négative

On considère la fonction Scilab suivante :

```
1  function f = bn(N,n,p)
2          X = grand(N,n,'geom',p) ;
3          k = 0 ;
4          for i=1 :N
5                  Y = sum(X(i, :)) ;
6                  if Y>=n/p then k = k+1
7                  end
8          end
9          f = k/N
10 endfunction
```

1. De quel nombre réel `bn(N,1,0.4)` fournit-il une valeur approchée lorsque N est grand ?
2. Donner une valeur approximative de `bn(1000,1000,0.4)`.

Solution de l'exercice 4.4

1. La ligne 2 sert à simuler N échantillons de n variables suivant la loi géométrique $\mathcal{G}(p)$.
Pour chacun de ces échantillons, on calcule la somme des n variables composant l'échantillon (ligne 5), et on compte le nombre de fois où cette somme est supérieure ou égale à $\frac{n}{p}$, qui n'est rien d'autre que l'espérance de cette somme.
Enfin, la ligne 9 nous indique qu'on s'intéresse non pas au nombre de fois où la somme dépasse $\frac{n}{p}$, mais plutôt à la fréquence du nombre de fois où cela se produit.
Autrement dit, si Y est une variable de même loi que la somme de n lois géométriques indépendantes de paramètre p, `bn(N,n,p)` retourne, pour N grand, une valeur approchée de $P(Y \geqslant E(Y))$.

En particulier, si $n = 1$, `bn(N,1,0.4)` retourne une valeur approchée de $P(Y \geqslant E(Y))$ où $Y \hookrightarrow \mathcal{G}(0.4)$.
Soit encore

$$P(Y \geqslant 2.5) = P(Y \geqslant 3) = 1 - P(Y = 1) - P(Y = 2) = 1 - 0.4 - 0.4 \times 0.6 = 0.36.$$

2. Lorsque n est grand, la somme de n variables géométriques i.i.d. a une loi proche de celle de la loi normale de même espérance et de même variance.
En effet, c'est une conséquence du théorème central limite : si $(X_n)_{n\in\mathbf{N}}$ est une suite de variables aléatoires i.i.d. suivant la loi géométrique $\mathcal{G}(p)$, et si $Y_n = \sum_{i=1}^{n} X_i$, alors $\frac{Y_n - E(Y_n)}{\sqrt{V(Y_n)}} \xrightarrow{\mathcal{L}} Y$ où $Y \hookrightarrow \mathcal{N}(0,1)$.
Autrement dit, pour n grand, on peut approcher la loi de $\frac{Y_n - E(Y_n)}{\sqrt{V(Y_n)}}$ par celle de Y.
Et donc celle de Y_n par celle de $\sqrt{V(Y_n)}(Y - E(Y_n))$, qui par transformation affine de loi normale, suit une loi $\mathcal{N}(E(Y_n), V(Y_n))$.
Ainsi, pour $n = 1000$, la loi de Y_n est proche d'une loi normale Y d'espérance $\frac{1000}{0.4} = 2500$, et de variance $\frac{1000 \times 0.6}{0.4^2}$.
Et alors, la probabilité que Y_n dépasse son espérance est approximativement égale à la probabilité que Y dépasse son espérance.
Mais pour une loi normale, cette probabilité vaut toujours[1] $\frac{1}{2}$!
Et donc `bn(1000,1000,0.4)` vaut environ 0.5.

□

Remarque

Ce que nous venons de faire n'est aucunement spécifique à la loi géométrique, et fonctionne de la même manière tant que le théorème central limite s'applique : la somme de N variables i.i.d. admettant une variance est, lorsque N est grand, «proche» de la loi normale de même variance et même espérance.

[1] Quelles que soient l'espérance et la variance.

Questions de cours posées à l'oral d'HEC

Les oraux d'HEC commencent systématiquement par une question de cours, plus ou moins difficile.
Cette question est généralement en lien avec la suite de l'exercice, et peut éventuellement être un indice pour la suite.
Voici une petite liste (non exhaustive) de questions posées ces dernières années.

1. Condition nécessaire et suffisante pour qu'une matrice soit diagonalisable.
2. Définition de la convergence en loi d'une suite de variables aléatoires.
3. Énoncer la formule du rang pour une application linéaire entre deux espaces vectoriels (sur $\mathbf{R}$ ou $\mathbf{C}$) de dimension finie.
4. Énoncer le théorème de stabilité de la loi de Poisson pour la somme.
5. Formule des probabilités totales.
6. Énoncer le théorème de Slutsky.
7. Polynômes annulateurs d'endomorphisme : définition et propriétés.
8. Stabilité de la loi γ pour la somme.
9. Définition et propriétés d'un produit scalaire.
10. Espérance d'un produit de variables aléatoires indépendantes.
11. Donner la définition de la convergence en loi et de la convergence en probabilités d'une suite de variables aléatoires.
12. Énoncer des conditions suffisantes de diagonalisabilité d'une matrice carrée réelle.
13. Indiquer pour quels nombres réels x les séries $\sum_{n\geqslant 1} x^n$ et $\sum_{n\geqslant 1} nx^{n-1}$ sont convergentes et préciser leurs sommes respectives.
14. Rappeler la définition d'un sous-espace stable par f.
15. Justifier que, si f est une fonction continue sur un intervalle I, alors la fonction $F : x \mapsto \int_a^x f(t)\,dt$ est définie et de classe $\mathscr{C}^1$ sur I.
16. Rappeler la formule donnant une densité de la somme de deux variables aléatoires à densité indépendantes et les conditions sous lesquelles cette formule est applicable.
17. i) Montrer que toute matrice de $\mathscr{M}_n(\mathbf{C})$ admet au moins un polynôme annulateur non nul.
    ii) En déduire que toute matrice de $\mathscr{M}_n(\mathbf{C})$ admet au moins une valeur propre complexe.
18. i) Rappeler l'énoncé du théorème central limite.
    ii) Comment en déduit-on un intervalle de confiance asymptotique pour l'espérance d'une loi admettant un moment d'ordre deux ?
19. Soit f une application de $\mathbf{R}^n$ dans $\mathbf{R}$ qui est de classe $\mathscr{C}^1$ sur $\mathbf{R}^n$. Donner la formule du développement limité à l'ordre un de f en un point et pour $(x, h) \in \mathbf{R}^n \times \mathbf{R}^n$, donner l'expression de la dérivée de l'application g définie sur $\mathbf{R}$ et qui à $t \in \mathbf{R}$ associe $g(t) = f(x + th)$.
20. Changement de variable dans une intégrale impropre.
21. Énoncer le théorème concernant la diagonalisation des endomorphismes symétriques d'un espace euclidien et démontrer la propriété concernant les sous-espaces propres d'un tel endomorphisme.
22. i) Donner, pour tout réel $\nu > 0$, l'expression d'une densité de la loi $\gamma(\nu)$.
    ii) Dans le cas où ν est strictement supérieur à 2, donner une représentation graphique de l'unique densité continue de la loi $\gamma(\nu)$.
23. Ordre de multiplicité d'une racine d'un polynôme.
24. i) Énoncer le théorème de la bijection.
    ii) Donner une condition suffisante pour qu'une fonction de répartition soit bijective.
25. Définition du moment d'ordre k d'une variable aléatoire réelle.
26. Propriétés de la fonction de répartition d'une variable aléatoire à densité.

27. Formule du rang pour une application linéaire. Application à la caractérisation des isomorphismes.
28. Formule de Taylor à l'ordre r avec reste intégral pour une fonction de classe $\mathscr{C}^{r+1}$.
29. Définition et propriétés d'un endomorphisme symétrique d'un espace vectoriel euclidien.
30. Soit f un endomorphisme de E, x un vecteur prope de f associé à la valeur propre θ et $P \in \mathbf{R}[X]$.
Exprimer $P(f)(x)$ en fonction de P, θ et x. Montrer que toute valeur propre de f est racine de n'importe quel polynôme annulateur de f.
31. Définition et propriétés de la loi exponentielle de paramètre $\lambda > 0$.
32. Rappeler la définition d'une application surjective et d'une application injective.
33. Théorème de transfert pour une variable à densité.
34. Énoncer le théorème de Pythagore.
35. Définition d'une valeur propre et d'un vecteur propre d'un endomorphisme.
36. Soit h une fonction numérique de classe $\mathscr{C}^1$ sur un ouvert Ω de $\mathbf{R}^n$.
    i) Qu'appelle-t-on point critique de h ?
    ii) Qu'appelle-t-on point critique pour l'optimisation de h sous contraintes d'égalités linéaires

$$\mathscr{C} : \begin{cases} g_1(X) &= b_1 \\ &\vdots \\ g_p(X) &= b_p \end{cases}$$

37. Théorème de d'Alembert-Gauss. Application à la factorisation de polynômes dans $\mathbf{R}[X]$.
38. Que peut-on dire d'un polynôme de $\mathbf{R}_n[X]$ qui admet plus de n racines ?
39. Formule de l'espérance totale pour une variable aléatoire discrète X et un système complet d'événements $(A_n)_{n\in\mathbf{N}^*}$.
40. Convergence des séries de Riemann.
41. Définition et propriétés de l'orthogonal d'un sous-espace vectoriel d'un espace euclidien.
42. Rappeler la définition du rang d'une matrice. Quelle est, selon les valeurs des réels a, b, c et d le rang de $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$?
43. Définition et propriétés d'un projecteur orthogonal.
44. Inégalité des accroissements finis pour une fonction réelle d'une variable réelle.
45. Rappeler la définition d'un endomorphisme symétrique d'un espace euclidien. Que peut-on dire de sa matrice dans une base orthonormale ?
46. Définition de la limite d'une suite de nombres réels ?
47. Donner deux conditions suffisantes et non nécessaires de diagonalisabilité d'une matrice.
48. Définition de la convergence d'une série réelle.
49. Définition et propriétés d'une fonction convexe sur un intervalle de $\mathbf{R}$.
50. Soit $X_1, \dots, X_n$ des variables indépendantes suivant la loi de Bernoulli $\mathscr{B}(p)$.
Rappeler comment l'inégalité de Bienaymé-Tchebychev permet de définir à partir de $\overline{X_n} = \dfrac{1}{n}\displaystyle\sum_{i=1}^{n} X_i$ un intervalle de confiance de p au niveau de confiance α ($\alpha \in]0,1[$).
51. Soient (u_n) et (v_n) deux suites de réels strictement positifs. Rappeler la signification de la relation $v_n = o(u_n)$.
52. Continuité d'une fonction réelle d'une variable réelle.
53. Sous-espaces vectoriels supplémentaires.
54. Comparaison de séries à termes positifs.
55. Définition de la covariance et du coefficient de corrélation linéaire $\rho_{X,Y}$ de deux variables aléatoires discrètes X et Y, prenant chacune au moins deux valeurs avec une probabilité strictement positive.
Indiquer dans quels cas $\rho_{X,Y}$ vaut 1 ou -1.
56. Rappeler la définition de la continuité en un point d'une fonction de $\mathbf{R}^n$ dans $\mathbf{R}$ ($n \geqslant 2$). Donner un exemple d'une fonction non continue sur $\mathbf{R}^2$.
57. Caractériser les isomorphismes d'espaces vectoriels de dimensions finies.
58. Inégalité de Bienaymé-Tchebychev. Estimateur sans biais, convergent.
59. Rappeler la définition du rang d'une matrice.
Une matrice et sa transposée ont-elles nécessairement le même rang ?

60. Donner une condition nécessaire et suffisante pour qu'une suite réelle décroissante soit convergente.
61. Loi d'un couple de variables aléatoires. Lois marginales.
62. Définition et propriétés d'une matrice inversible.
63. Rappeler la définition d'une série convergente. Montrer qu'une série à termes positifs est convergente si et seulement si la suite de ses sommes partielles est majorée.
Cette équivalence demeure-t-elle valable pour les séries à termes réels de signe quelconque ?