Épreuve de mathématiques II
Correction

Exercice

Probabilité qu'une matrice aléatoire soit diagonalisable

1. Étude de la diagonalisabilité d'une matrice de $\mathscr{M}_2(\mathbb{R})$.

(a) α et β sont exactement les valeurs propres de A, donc si $\alpha \neq \beta$ le polynôme caractéristique de A sera scindé à racines simples, et dans ce cas A sera diagonalisable.

(b) Si $\alpha = \beta$ alors A ne sera pas diagonalisable car elle serait de la forme $A = \alpha I_2$ ce qui est absurde.

2. Calcul de la probabilité qu'une matrice aléatoire soit diagonalisable dans $\mathscr{M}_2(\mathbb{R})$.

(a) $\forall k \in \mathbb{N}^*$, $p(X=k) = p_1(1-p_1)^{k-1}$.

(b) On a $U(\Omega) = (X+Y)(\Omega) = \mathbb{N} \setminus \{0,1\}$ et, pour $k \in \mathbb{N} \setminus \{0,1\}$, on obtient par indépendance :

$$\begin{aligned} p(X+Y=k) &= p\left(\bigcup_{n=1}^{k-1} (X=n, Y=k-n)\right) \\ &= \sum_{n=1}^{k-1} p(X=n)p(Y=k-n) \\ &= \sum_{n=1}^{k-1} (1-p_1)^{n-1} p_1 (1-p_2)^{k-n-1} p_2 \\ &= \begin{cases} p_1 p_2 \dfrac{(1-p_2)^{k-1} - (1-p_1)^{k-1}}{p_1 - p_2} & \text{si } p_1 \neq p_2 \\ (k-1)p^2(1-p)^{k-2} & \text{si } p_1 = p_2 = p. \end{cases} \end{aligned}$$

Cela détermine la loi de la variable aléatoire $U = X+Y$.

(c) D'abord V prend ses valeurs dans $\mathbb{N}^*$. D'autre part, l'événement $(X \geq n)$ s'écrit comme réunion disjointe des événements élémentaires $(X=k)$, $k \geq n$. On a donc

$$p(X \geq n) = \sum_{k=n}^{\infty} p(X=k) = \sum_{k=n}^{\infty} (1-p_1)^{k-1} p_1 = \frac{(1-p_1)^{n-1} p_1}{1-(1-p_1)} = (1-p_1)^{n-1}.$$

Soit $w \in \Omega$. On a $V(w) \geq k$ si et seulement si $X(w) \geq k$ et $Y(w) \geq k$. Ces deux événements sont indépendants, et donc

$$p(V \geq k) = p(X \geq k)p(Y \geq k) = (1-p_1)^{k-1}(1-p_2)^{k-1} = [(1-p_1)(1-p_2)]^{k-1}.$$

Or,

$$\begin{aligned} p(V=k) &= p(V \geq k) - p(V \geq k+1) \\ &= [(1-p_1)(1-p_2)]^{k-1} - [(1-p_1)(1-p_2)]^{k} \\ &= (1-(1-p_1)(1-p_2)\,[(1-p_1)(1-p_2)]^{k-1}. \end{aligned}$$

Ainsi, V suit une loi géométrique de paramètre $(1-p_1)(1-p_2)$.

(d) L'événement $(X < Y)$ peut être décomposé en la réunion disjointes des événements $(X = k, Y > k)$ avec $k \in \mathbb{N}^*$. On a donc $p(X < Y) = \displaystyle\sum_{p=1}^{\infty} p(X = k, Y > k)$.

Par indépendance des variables X et Y, on a $p(X = k, Y > k) = p(X = k)p(Y > k)$ avec $p(X = k) = p_1(1 - p_1)^{k-1}$ et $p(Y > k) = p(Y \geq k + 1) = (1 - p_2)^k$. On en déduit

$$p(X < Y) = \sum_{k=1}^{\infty} p_1(1 - p_2)\left[(1 - p_1)(1 - p_2)\right]^{k-1} = \frac{p_1(1 - p_2)}{1 - (1 - p_1)(1 - p_2)} = \frac{p_1(1 - p_2)}{p_1 + p_2 - p_1p_2}.$$

(e) La probabilité recherchée n'est donc autre que $p(X \neq Y)$.

L'événement $(X \neq Y)$ est le complémentaire de l'événement $(X = Y)$ qui est la réunion d'événements deux à deux disjoints

$$(X = Y) = \bigcup_{n \in \mathbb{N}^*} (X = n, Y = n).$$

Par indépendance

$$p(X = n, Y = n) = p(X = n)p(Y = n) = p_1p_2(1 - p_1)^{n-1}(1 - p_2)^{n-1}.$$

Ainsi

$$p(X = Y) = \sum_{n=1}^{\infty} p(X = n)p(Y = n) = p_1p_2 \sum_{n=1}^{\infty} \left[(1 - p_1)(1 - p_2)\right]^{n-1} = \frac{p_1p_2}{p_1 + p_2 - p_1p_2}.$$

Finalement, la probabilité que la matrice soit diagonalisable vaut

$$p(X \neq Y) = 1 - p(X = Y) = \frac{p_1 + p_2 - 2p_1p_2}{p_1 + p_2 - p_1p_2}.$$

Problème

Première partie
Quelques résultats préliminaires

1. Il clair que ${}^t\left({}^tMM\right) = {}^tMM$. De plus, pour tout $X \in \mathscr{M}_{n,1}(\mathbb{R})$, on a :

$${}^tX{}^tMMX = {}^t(MX)\,MX = (MX|MX) = \|MX\|^2 \geq 0$$

où $\|.\|$ est la norme associée au produit scalaire $(.|.)$. Donc ${}^tMM$ est symétrique et positive.

2. (a) Les composantes de DX sont $d_1x_1, d_2x_2, ..., d_nx_n$, donc ${}^tXDX = (X|DX) = \displaystyle\sum_{i=1}^{n} d_ix_i^2$.

(b) Si les réels d_i sont positifs alors ${}^tXDX = \displaystyle\sum_{i=1}^{n} d_ix_i^2 \geq 0$ pour tout $X \in \mathscr{M}_{n,1}(\mathbb{R})$. Donc la matrice D est positive.

3. Caractérisation de positivité d'une matrice symétrique par le signe de ses valeurs propres

(a) Soient $A \in \mathscr{M}_n(\mathbb{R})$ une matrice symétrique et λ une valeur propre de A, alors il existe $X \in \mathscr{M}_{n,1}(\mathbb{R})$ non nul tel que $AX = \lambda X$ et donc $0 \leq {}^tXAX = \lambda{}^tXX = \lambda\|X\|^2$. Comme X est non nul, alors $\lambda \geq 0$.

(b) Soit A une matrice symétrique réelle.

i. C'est la traduction matricielle du théorème spectral : il existe $D = \mathrm{diag}(\lambda_1, \lambda_2, ..., \lambda_n)$ diagonale à éléments diagonaux réels et P une matrice orthogonale telles que $A = {}^tPDP$.

ii. Soit $X \in \mathscr{M}_{n,1}(\mathbb{R})$, on a ${}^tXAX = {}^tX{}^tPDPX = {}^t(PX)D(PX) = \sum_{i=1}^{n} \lambda_i y_i$ où les réels y_i sont les composantes de DX. Donc si les valeurs propres de A sont positives, alors A est positive.

4. La matrice ${}^tBB$ étant symétrique et positive(d'après 1.), donc d'après le théorème spectral il existe une matrice diagonale $\triangle$ à éléments diagonaux positifs (d'après 3.(b)) et P orthogonale telles que ${}^tP({}^tBB)P = \triangle$.

5. Il est très facile de vérifie que φ définit une forme bilinéaire symétrique. Reste à démontrer qu'elle est définie positive. Soit $A \in \mathscr{M}_n(\mathbb{R})$ et notons $(b_{ij}) = {}^tAA$. Alors

$$b_{ii} = \sum_{k=1}^{n} a_{ki}^2 \geq 0.$$

Ainsi,

$$\varphi(A, A) = \text{tr}({}^tAA) = \sum_{i=1}^{n} \sum_{k=1}^{n} a_{ki}^2 \geq 0.$$

On a bien affaire à une forme positive. De plus, si $\varphi(A) = 0$, alors pour tout $i = 1, 2, \ldots, n$ et tout $k = 1, 2, \ldots, n$, on a $a_{ki} = 0$, et donc $A = 0$. La forme est donc définie.

Deuxième partie
Décomposition polaire d'une matrice de $\mathscr{M}_n(\mathbb{R})$

1. (a) Si $A_1, A_2, ..., A_n$ désignent les colonnes de la matrice A, alors ${}^tA_1, {}^tA_2, ..., {}^tA_n$ sont exactement les lignes de ${}^tA$, d'où :

$${}^tAA = \left({}^tA_iA_j\right)_{1 \leq i,j \leq n}.$$

Ainsi ${}^tAA = D^2 = \text{diag}(\lambda_1^2, \lambda_2^2, ..., \lambda_n^2)$ si, et seulement si, ${}^tA_iA_j = \lambda_i^2\delta_{ij}$ pour tout $(i, j) \in [\![1, n]\!]^2$. En particulier, si $d_i = 0$, $\|A_i\|^2 = {}^tA_iA_i = d_i^2 = 0$ et la colonne $A_i = 0$.

(b) Notons I l'ensemble des i tels que $d_i \neq 0$ et, pour chaque $i \in I$, posons $E_i = \dfrac{A_i}{||A_i||} = \dfrac{A_i}{d_i}$. La famille $(E_i)_{i \in I}$ est alors une famille orthonormale, que nous pouvons complétée en une base orthonormale $(E_i)_{1 \leq i \leq n}$. Comme $d_i = 0$ et $A_i = 0$ pour $i \notin I$, l'égalité $A_i = d_iE_i$ est vraie pour tout $i \in [\![1, n]\!]$.

(c) Soit E la matrice de passage de la base canonique de $\mathscr{M}_n(\mathbb{R})$ à la base $(E_i)_{1 \leq i \leq n}$. Cette base étant orthonormale, E est une matrice orthogonale et $A_i = d_iE_i$ pour tout i se traduit par $A = ED$.

2. (a) ${}^tBB$ est symétrique réelle, donc il existe P orthogonale telle que ${}^tP\left({}^tBB\right)P$ soit diagonale. D'autre part, ${}^tBB$ est positive donc ses valeurs propres sont positives. On en déduit que $\triangle = {}^tP{}^tBBP$ est une matrice diagonale à termes positifs. Donc il suffit de prendre $D = \text{diag}\left(\sqrt{\lambda_1}, \sqrt{d_2}, ..., \sqrt{d_n}\right)$.

(b) On a ${}^tP{}^tBBP = {}^t(BP)BP = D^2$, donc d'après l'étude précédente, il existe une matrice orthogonale E telle que $BP = ED$.
L'égalité $BP = ED$ entraîne $B = E{}^tPPD{}^tP = OS$ avec $O = E{}^tP \in \mathscr{O}_n(\mathbb{R})$ (produit de deux matrices orthogonales) et $S = PD{}^tP \in \mathscr{S}_n(\mathbb{R})$.

3. On obtient directement $M = {}^tCC$:

$$M = \begin{pmatrix} 6 & 5 & 5 \\ 5 & 6 & 5 \\ 5 & 5 & 6 \end{pmatrix} = I_3 + 5\,J \text{ avec } J = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}.$$

J est clairement de rang 1, donc 1 est valeur propre double de ${}^tCC$, la troisième valeur propre étant égale à 16 puisque $\text{Tr}({}^tCC) = 18$.

Comme $(1,1,1)$ est un vecteur propre évident associé à la valeur propre 16, posons

$$e_1 = \frac{\sqrt{3}}{3}\begin{pmatrix}1\\1\\1\end{pmatrix}.$$

Le noyau de $M - I_3$ est l'hyperplan d'équation $x+y+z=0$. Nous posons donc $e_2 = \frac{\sqrt{2}}{2}\begin{pmatrix}1\\-1\\0\end{pmatrix}$ et $e_3 = \frac{\sqrt{6}}{6}\begin{pmatrix}1\\1\\-2\end{pmatrix}$, pour obtenir ${}^tPMP = D^2$ avec $P = \begin{pmatrix}\frac{\sqrt{3}}{3} & \frac{\sqrt{2}}{2} & \frac{\sqrt{6}}{6}\\ \frac{\sqrt{3}}{3} & -\frac{\sqrt{2}}{2} & \frac{\sqrt{6}}{6}\\ \frac{\sqrt{3}}{3} & 0 & -\frac{\sqrt{6}}{3}\end{pmatrix}$ et $D = \begin{pmatrix}4&0&0\\0&1&0\\0&0&1\end{pmatrix}$.

Posons $S = PD{}^tP$ qui est inversible (les valeurs propres de S sont égales à celles de D), on peut poser $O = CS^{-1} = CPD^{-1}{}^tP$. Nous avons ensuite :

$${}^tOO = {}^tS^{-1}{}^tCCS^{-1} = S^{-1}MS^{-1} = PD^{-1}{}^tPPD^2{}^tPPD^{-1}{}^tP = I_3$$

et O est bien orthogonale.
Il reste à calculer S et O :

$$S = PD{}^tP = \begin{pmatrix}2&1&1\\1&2&1\\1&1&2\end{pmatrix},\quad O = CPD^{-1}{}^tP = \begin{pmatrix}0&-1&0\\1&0&0\\0&0&1\end{pmatrix}.$$

Troisième partie
Application à un calcul de distance

1. (a) L'application $M \mapsto {}^tM$ est linéaire en dimension finie, donc elle est continue sur $\mathscr{M}_n(\mathbb{R})$.

(b) L'application $\tau : M \mapsto {}^tMM$ est la composée de l'application linéaire $M \mapsto {}^tM$ et de l'application bilinéaire $(M,N) \mapsto MN$ et puisque on est en dimension finie, l'application τ est continue sur $\mathscr{M}_n(\mathbb{R})$.

(c) • On remarque que $\mathscr{O}_n(\mathbb{R}) = \tau^{-1}(I_n)$. Donc c'est un fermé comme image réciproque d'un fermé par une application continue.

• $\mathscr{O}_n(\mathbb{R})$ est une partie bornée de $\mathscr{M}_n(\mathbb{R})$. En effet, toutes les normes sur $\mathscr{M}_n(\mathbb{R})$ sont équivalentes. On note $N_\infty(M) = \sup\limits_{(i,j)\in[\![1,n]\!]^2} |m_{ij}|$ où $M = (m_{ij})_{1\le i,j\le n}$. N_∞ est une norme. On montre facilement que : $\forall M \in \mathscr{O}_n(\mathbb{R})$, $N_\infty(M) \le 1$. Donc $\mathscr{O}_n(\mathbb{R})$ est borné.

2. L'application $f : M \mapsto \|A-M\|_2$ est continue sur $\mathscr{M}_n(\mathbb{R})$. En effet, $\forall M, N \in \mathscr{M}_n(\mathbb{R})$, on a :

$$|f(M)-f(N)| = |\|A-M\|_2 - \|A-M\|_2| \le \|N-N\|_2.$$

Donc f est lipschitzienne et par conséquent elle est continue. Comme $\mathscr{O}_n(\mathbb{R})$ est un compact, alors d'après le théorème de Heine, f est bornée et atteint ses bornes sur le compact $\mathscr{O}_n(\mathbb{R})$ ($\mathscr{O}_n(\mathbb{R})$ est un fermé borné dans un espace de dimension finie). En particulier il existe $M_0 \in \mathscr{O}_n(\mathbb{R})$ tel que

$$\inf_{M\in\mathscr{O}_n(\mathbb{R})} \|A-M\|_2 = \|A-M_0\|_2.$$

3. Nous avons :

$$\|\Omega A\|_2^2 = \mathrm{Tr}\left({}^t(\Omega A)\Omega A\right) = \mathrm{Tr}\left({}^tA{}^t\Omega\Omega A\right) = \mathrm{Tr}\left({}^tAA\right) = \|A\|_2^2$$

et en utilisant la propriété classique $\mathrm{Tr}\,(MN) = \mathrm{Tr}\,(NM)$:

$$\|A\Omega\|_2^2 = \mathrm{Tr}\left({}^t\Omega{}^tAA\Omega\right) = \mathrm{Tr}\left({}^tAA\Omega{}^t\Omega\right) = \mathrm{Tr}\left({}^tAA\right) = \|A\|_2^2,$$

ce qui donne bien $\|\Omega A\|_2 = \|A\Omega\|_2 = \|A\|_2$.

4. (a) On a $\|A - \Omega\|_2 = \|OS - \Omega\|_2 = \|O(S - O^{-1}\Omega)\|_2 = \|S - O^{-1}\Omega\|_2$ d'après la question précédente. Quand Ω décrit $\mathscr{O}_n(\mathbb{R})$, $O^{-1}\Omega$ décrit également $\mathscr{O}_n(\mathbb{R})$, donc $d(A, \mathscr{O}_n(\mathbb{R})) = d(S, \mathscr{O}_n(\mathbb{R}))$.

(b) Pour tout $\Omega \in \mathscr{O}_n(\mathbb{R})$, nous avons :

$$\|S - \Omega\|_2 = \|PDP^{-1} - \Omega\|_2 = \|P(D - P^{-1}\Omega P)P^{-1}\|_2 = \|D - P^{-1}\Omega P\|_2$$

car $P, P^{-1} \in \mathscr{O}_n(\mathbb{R})$. Une nouvelle fois, $P^{-1}\Omega P$ décrit $\mathscr{O}_n(\mathbb{R})$ quand Ω décrit $\mathscr{O}_n(\mathbb{R})$ donc

$$d(A, \mathscr{O}_n(\mathbb{R})) = d(S, \mathscr{O}_n(\mathbb{R})) = d(D, \mathscr{O}_n(\mathbb{R})).$$

5. (a) Les λ_i sont les valeurs propres de la matrice S qui est symétrique et positive, donc les λ_i sont positives.

(b) En notant $\Omega = (w_{ij})_{1 \le i,j \le n}$ et $D = (d_{ij})_{1 \le i,j \le n}$, nous obtenons :

$$\mathrm{Tr}\,(D\Omega) = \sum_{i=1}^{n}\sum_{k=1}^{n} d_{ik}w_{ki} = \sum_{i=1}^{n} \lambda_i w_{ii} \le \sum_{i=1}^{n} \lambda_i$$

car Ω étant orthogonale, les w_{ij} sont éléments de $[-1, 1]$.

(c) $\|D - \Omega\|_2^2 = (D - \Omega | D - \Omega) = \|D\|_2^2 - 2(D|\Omega) + \|\Omega\|_2^2 = \sum_{i=1}^{n} \lambda_i^2 - 2\,\mathrm{Tr}\,(D\Omega) + n.$

(d) On en déduit que pour tout $\Omega \in \mathscr{O}_n(\mathbb{R})$:

$$\|D - \Omega\|_2^2 \ge \sum_{i=1}^{n} \lambda_i^2 - 2\sum_{i=1}^{n} \lambda_i + n = \sum_{i=1}^{n} (\lambda_i - 1)^2 = \|D - I_n\|_2^2.$$

Comme $I_n \in \mathscr{O}_n(\mathbb{R})$, ceci prouve que la distance de D à $\Omega \in \mathscr{O}_n(\mathbb{R})$ est minimale pour $\Omega = I_n$.

Nous venons de démontrer que $d(A, \mathscr{O}_n(\mathbb{R})) = d(D, \mathscr{O}_n(\mathbb{R})) = d(D, I_n) = \sqrt{\sum_{i=1}^{n} (\lambda_i - 1)^2}$, où les λ_i sont les racines carrées des valeurs propres de ${}^tAA$.
D'après 4.(b), on a : $d(A, \mathscr{O}_n(\mathbb{R})) = d(D, \mathscr{O}_n(\mathbb{R})$, d'où :

$$\begin{aligned}
d(A, \mathscr{O}_n(\mathbb{R})) &= \|D - I_n\|_2 \\
&= \|P^{-1}SP - I_n\|_2 \\
&= \|P^{-1}(S - I_n)P\|_2 \\
&= \|S - I_n\|_2 \\
&= \|OS - O\|_2 = \|A - O\|_2
\end{aligned}$$

6. **Application :** Les valeurs propres de ${}^tCC = \begin{pmatrix} 6 & 5 & 5 \\ 5 & 6 & 5 \\ 5 & 5 & 6 \end{pmatrix}$ sont 16 et 1 double, donc $\lambda_1 = 4$ et $\lambda_2 = \lambda_3 = 1$ et par conséquent $d(C, \mathscr{O}_n(\mathbb{R})) = 3$.

• • • • • • • • • •