Mines MP 2013 – Épreuve de Mathématiques 2

A. Produit scalaire de matrices

1°) La i-ième composante d'un vecteur x de $\mathbb{R}^n$ dans une base orthonormée $(e_1, e_2, \dots, e_n)$ est donnée par $< x \,, e_i >$. En particulier, $< Ae_i \,, e_i >$ représente la i-ième composante du vecteur $f(e_i)$ où f désigne l'endomorphisme de $\mathbb{R}^n$ canoniquement associé à A. Par suite, $\sum\limits_{i=1}^{n} < Ae_i \,, e_i >$ est la somme des éléments diagonaux de la matrice de f dans la base $(e_1, e_2, \dots, e_n)$, c'est-à-dire la trace de f, encore égale à $\operatorname{tr}(A)$ par invariance de la trace par changement de base.

2°) – L'application $<\,,\,>$ est symétrique car $\forall A, B \in \mathcal{M}_n(\mathbb{R}),\ \operatorname{tr}({}^tAB) = \operatorname{tr}({}^t({}^tAB)) = \operatorname{tr}({}^tBA)$.

– Cette application étant clairement linéaire à droite par linéarité de la trace, elle est bilinéaire.

– Soit $(e_1, e_2, \dots, e_n)$ une base orthonormée de $\mathbb{R}^n$. Pour tout $A \in \mathcal{M}_n(\mathbb{R})$, on a $< A \,, A > = \sum\limits_{i=1}^{n} < {}^tAAe_i \,, e_i >$ soit $\sum\limits_{i=1}^{n} < Ae_i \,, Ae_i >$, d'où $< A \,, A > = \sum\limits_{i=1}^{n} \| Ae_i \|^2 \geqslant 0$ avec égalité si et seulement si $\forall i \in [\![1, n]\!],\ Ae_i = 0$, c'est-à-dire ssi A est nulle.

L'application $<\,,\,>$ définit donc un produit scalaire sur $\mathcal{M}_n(\mathbb{R})$.

3°) On rappelle (propriété du cours ou définition selon le point de vue adopté) qu'une matrice symétrique réelle A est positive si et seulement si $\forall x \in \mathbb{R}^n,\ < Ax \,, x > \geqslant 0$.

Toute matrice symétrique réelle étant diagonalisable dans $O_n(\mathbb{R})$ d'après le théorème spectral, il existe une base orthonormée $(e_1, e_2, \dots, e_n)$ de $\mathbb{R}^n$ formée de vecteurs propres de B. En appelant μ_i la valeur propre de B associée au vecteur e_i, il vient :

$$< A \,, B > = \operatorname{tr}({}^tAB) = \sum_{i=1}^{n} < {}^tABe_i \,, e_i > = \sum_{i=1}^{n} \mu_i < Ae_i \,, e_i >,$$

quantité qui est bien positive puisque $\forall x \in \mathbb{R}^n,\ < Ax \,, x > \geqslant 0$ et $\forall i \in [\![1, n]\!],\ \mu_i \geqslant 0$.

B. Décomposition polaire

4°) – La matrice ${}^tAA$ est symétrique réelle de manière immédiate. De plus, pour tout vecteur $x \in \mathbb{R}^n$, on a $< {}^tAAx \,, x > = \| Ax \|^2 \geqslant 0$ donc ${}^tAA$ est positive.

– Soit $\lambda_1, \lambda_2, \dots, \lambda_n$ les valeurs propres de ${}^tAA$, avec $0 \leqslant \lambda_1 \leqslant \lambda_2 \leqslant \cdots \leqslant \lambda_n$, et $(e_1, e_2, \dots, e_n)$ une base orthonormée de vecteurs propres associés. Alors, pour tout vecteur unitaire $X = \sum\limits_{i=1}^{n} x_i e_i$ de $\mathbb{R}^n$, on a

$$\| AX \|^2 = \sum_{i=1}^{n} \lambda_i x_i^2 \leqslant \lambda_n \| X \|^2 = \lambda_n,$$

d'où $\| A \|_2 \leqslant \sqrt{\lambda_n}$. Par ailleurs, le vecteur unitaire e_n réalise l'égalité $\| Ae_n \| = \sqrt{\lambda_n}$, d'où au final $\| A \|_2 = \sqrt{\lambda_n}$.

5°) $f^* \circ f$ étant autoadjoint positif (puisque sa matrice ${}^tAA$ dans une base orthonormée est symétrique positive), il se diagonalise dans une base orthonormée que l'on notera encore $(e_1, e_2, \dots, e_n)$. L'endomorphisme défini par $\forall i \in [\![1, n]\!],\ h(e_i) = \sqrt{\lambda_i} e_i$ est alors autoadjoint positif et, par construction, $f^* \circ f = h^2$.

6°) Si $x \in \operatorname{Ker} \tilde{h}$, alors il existe $y \in E$ tel que $x = h(y)$ et $h(x) = 0$.
Or, par construction de h, on a clairement $\operatorname{rg}(h) = \operatorname{rg}(h^2)$ et, comme $\operatorname{Ker} h \subset \operatorname{Ker} h^2$, il vient $\operatorname{Ker} h = \operatorname{Ker} h^2$ d'après le théorème du rang. L'égalité $h^2(y) = 0$ entraîne donc $h(y) = 0$, d'où $x = 0$.
L'endomorphisme $\tilde{h}$ est donc injectif et comme $\operatorname{Im} h$ est un espace vectoriel de dimension finie, $\tilde{h}$ définit bien un automorphisme de $\operatorname{Im} h$.

7°) – Pour tout $x \in \mathbb{R}^n$,

$$\| f(x) \|^2 = < f^* \circ f(x) \,, x > = < h^2(x) \,, x > = < h^* \circ h(x) \,, x > = \| h(x) \|^2,$$

d'où $\| h(x) \| = \| f(x) \|$.

– Il en ressort en particulier que $\operatorname{Ker} h = \operatorname{Ker} f$, d'où $\dim \operatorname{Ker} h = \dim \operatorname{Ker} f = \dim (\operatorname{Im} f)^\perp$.

– N'importe quelle application linéaire v envoyant une base orthonormée de $\operatorname{Ker} h$ sur une base orthonormée de $(\operatorname{Im} f)^\perp$ conserve alors la norme donc réalise un isomorphisme de $\operatorname{Ker} h$ sur $(\operatorname{Im} f)^\perp$.

8°) – On observe pour commencer que $\operatorname{Ker} h = (\operatorname{Im} h^*)^\perp = (\operatorname{Im} h)^\perp$.

– Les sous-espaces $\operatorname{Ker} h$ et $\operatorname{Im} h$ d'une part, et $\operatorname{Im} f$ et $(\operatorname{Im} f)^\perp$ d'autre part, étant supplémentaires orthogonaux dans E, il existe un unique endomorphisme u de E qui coïncide avec $f \circ \tilde{h}^{-1}$ sur $\operatorname{Im} h$ et v sur $\operatorname{Ker} h$.

– De plus, l'application $f \circ \tilde{h}^{-1}$ conserve la norme : en effet, $\forall x \in \operatorname{Im} h$, $\| f \circ \tilde{h}^{-1}(x) \| = \| h \circ \tilde{h}^{-1}(x) \| = \| x \|$. Comme v conserve également la norme des vecteurs de $\operatorname{Ker} h$, le théorème de Pythagore entraîne que $\forall x \in E$, $\| u(x) \| = \| x \|$, ce qui montre que u appartient à $O(E)$.

– Enfin, les endomorphismes f et $u \circ h$ coïncident par construction sur les sous-espaces supplémentaires $\operatorname{Ker} h$ et $\operatorname{Im} h$, donc sont égaux.

9°) Il s'agit de l'interprétation matricielle du résultat de la question **8°)** : si f est l'endomorphisme canoniquement associé à A, la relation $f = u \circ h$ se traduit matriciellement par $A = US$, avec $U \in O_n(\mathbb{R})$ et S symétrique positive (puisque la base canonique est orthonormée pour le produit scalaire usuel).

C. Projeté sur un convexe compact

10°) – L'application $d_x : h \longmapsto \| x - h \|$ est 1-lipschitzienne donc continue de E dans $\mathbb{R}$. Comme H est compact, d_x est bornée et atteint sa borne inférieure sur H d'après le théorème des bornes, d'où l'existence de $h_0 \in H$ tel que $d(x, H) = \| x - h_0 \|$.

– Si $h_1 \neq h_0$ est un autre élément de H tel que $d(x, H) = \| x - h_1 \|$, alors d'après le théorème de la médiane,

$$\| x - \frac{1}{2}(h_0 + h_1) \|^2 = \frac{1}{2} \| x - h_0 \|^2 + \frac{1}{2} \| x - h_1 \|^2 - \frac{1}{4} \| h_0 - h_1 \|^2 < \big(d(x, H)\big)^2.$$

Or $\frac{1}{2}(h_0 + h_1) \in H$ vu que H est convexe, ce qui conduit à une contradiction avec la définition de la borne inférieure.

11°) – Utilisons cette fois l'indication de l'énoncé. Pour tout $h_1 \in H$ et tout $t \in [0, 1]$, on a $th_0 + (1 - t)h_1 \in H$ par convexité de H, donc $q(t) \geqslant \| x - h_0 \|^2$. En écrivant $q(t)$ sous la forme $\| x - h_0 + (1 - t)(h_0 - h_1) \|^2$ et en développant le carré scalaire, il vient alors :

$$\forall t \in [0, 1[, \qquad (1 - t). < x - h_0 \, , \, h_0 - h_1 > \; + (1 - t)^2 \, \| h_0 - h_1 \|^2 \geqslant 0.$$

En divisant par $(1 - t) > 0$ et en faisant tendre t vers 1 par valeurs inférieures, on obtient alors $\forall h_1 \in H$, $< x - h_0 \, , \, h_0 - h_1 > \geqslant 0$, ce qui équivaut à la condition demandée.

– Réciproquement, si on a $\forall h_1 \in H$, $< x - h_0 \, , \, h_0 - h_1 > \geqslant 0$, alors $\forall t \in [0, 1]$, $q(t) \geqslant \| x - h_0 \|^2$ en remontant les calculs précédents. En particulier, $q(0) \geqslant \| x - h_0 \|^2$, ce qui signifie que $\forall h_1 \in H$, $\| x - h_1 \| \geqslant \| x - h_0 \|$ et h_0 est bien (l'unique) point de H tel que $d(x, H) = \| x - h_0 \|$.

Remarque : cette condition signifie géométriquement que l'angle formé par les vecteurs $x - h_0$ et $h - h_0$ est obtus.

D. Théorème de Carathéodory et compacité

12°) Soit $CC(H)$ l'ensemble des combinaisons convexes des éléments de H. Toute partie convexe de E qui contient H contenant aussi les combinaisons convexes de leurs éléments, on a déjà l'inclusion immédiate $CC(H) \subset \operatorname{conv}(H)$.

De plus, H est clairement inclus dans $CC(H)$ (tout $x \in H$ s'écrit $x = 1.x$) et $CC(H)$ est convexe. En effet, si $x = \sum_{i=1}^{p} \lambda_i x_i \in CC(H)$ et $y = \sum_{j=1}^{q} \mu_j y_j \in CC(H)$ (avec $\forall i \in [\![1, p]\!]$, $\lambda_i \geqslant 0$, $x_i \in H$ et $\sum_{i=1}^{p} \lambda_i = 1$, et de même $\forall j \in [\![1, q]\!]$, $\mu_j \geqslant 0$, $y_j \in H$ et $\sum_{j=1}^{q} \mu_j = 1$) et si $\alpha \in [0, 1]$,

$$\alpha.x + (1 - \alpha).y = \sum_{i=1}^{p} \alpha \lambda_i x_i + \sum_{j=1}^{q} (1 - \alpha) \mu_j y_j$$

donc $\alpha.x + (1 - \alpha).y \in H$ vu que les scalaires $\alpha \lambda_i$ et $(1 - \alpha)\mu_j$ sont positifs et que $\sum_{i=1}^{p} \alpha \lambda_i + \sum_{j=1}^{q} (1 - \alpha) \mu_j = 1$.

$CC(H)$ est donc le plus petit convexe de E contenant H, d'où l'égalité $CC(H) = \operatorname{conv}(H)$.

13°) La famille $(x_2 - x_1, x_3 - x_1, \ldots, x_p - x_1)$ est liée car de cardinal $p - 1 \geqslant n + 1$ dans un espace vectoriel de dimension n. Il existe donc des réels non tous nuls $\mu_2, \mu_3, \ldots, \mu_p$ tels que $\sum_{i=2}^{p} \mu_i (x_i - x_1) = 0$. En posant $\mu_1 = - \sum_{i=2}^{p} \mu_i$, on a bien trouvé p réels non tous nuls tels que

$$\sum_{i=1}^{p} \mu_i x_i = 0 \qquad \text{et} \qquad \sum_{i=1}^{p} \mu_i = 0.$$

14°) Pour un indice i fixé, l'ensemble des réels θ tels que $\lambda_i - \theta\mu_i \geqslant 0$ est un intervalle fermé contenant 0, égal à $\mathbb{R}$ si $\mu_i = 0$, borné supérieurement si $\mu_i > 0$ et borné inférieurement si $\mu_i < 0$. Comme l'un au moins des μ_i est strictement positif et l'un au moins strictement négatif, l'ensemble $\{\theta \in \mathbb{R} \,/\, \forall i \in [\![1,\, p]\!],\ \lambda_i - \theta\mu_i \geqslant 0\}$ est un segment. Choisissons pour θ l'une de ses bornes : il existe alors un indice $j \in [\![1,\, p]\!]$ tel que $\lambda_j - \theta\mu_j = 0$ et, pour tout $i \neq j$, $\lambda_i - \theta\mu_i \geqslant 0$. De plus,

$$x \;=\; \sum_{i=1}^{p} (\lambda_i - \theta\mu_i)x_i \qquad \text{et} \qquad \sum_{i=1}^{p} (\lambda_i - \theta\mu_i) \;=\; 1.$$

Il en ressort que x est combinaison convexe d'au plus $p-1$ éléments de H. Si ce nombre d'éléments est encore supérieur ou égal à $n+2$, on peut recommencer le raisonnement et, par une itération finie, on se ramène à une combinaison convexe d'au plus $n+1$ éléments de H.

15°) L'ensemble Λ est bien un compact de $\mathbb{R}^{n+1}$: il est en effet fermé (comme intersection de demi-espaces fermés et d'un hyperplan affine) et borné (si $(t_1, t_2, \dots, t_{n+1} \in \Lambda$, alors $\forall i \in [\![1,\, n+1]\!]$, $0 \leqslant t_i \leqslant 1$).
D'après les questions **12°)** et **14°)**, $\mathrm{conv}(H)$ est précisément l'ensemble des combinaisons convexes d'au plus $n+1$ points de H, donc l'image de $\Lambda \times H^{n+1}$ par l'application

$$\phi : \Big((t_1, t_2, \dots, t_{n+1}),\, (x_1, x_2, \dots, x_{n+1})\Big) \longmapsto \sum_{i=1}^{n+1} t_i x_i.$$

(À noter qu'une combinaison convexe de moins de $n+1$ points s'obtient aussi de la sorte, en prenant certains t_i nuls).
L'application ϕ étant continue sur $\mathbb{R}^{n+1} \times E^{n+1}$ (car par exemple bilinéaire en dimension finie) et $\Lambda \times H^{n+1}$ étant compact en tant que produit de compacts, $\mathrm{conv}(H) = \phi\big(\Lambda \times H^{n+1}\big)$ est donc un compact de E.

E. Enveloppe convexe de $\mathcal{O}_n(\mathbb{R})$

16°) $O_n(\mathbb{R})$ est un fermé de $\mathcal{M}_n(\mathbb{R})$ comme image réciproque du fermé $\{I_n\}$ par l'application continue $M \longmapsto {}^tMM$.
Il est de plus borné car tout vecteur colonne d'une matrice orthogonale est de norme euclidienne égale à 1.
$O_n(\mathbb{R})$ est ainsi un compact de $\mathcal{M}_n(\mathbb{R})$ et, d'après le résultat de la question **15°)**, son enveloppe convexe $\mathrm{conv}(O_n(\mathbb{R}))$ est donc également un compact de $\mathcal{M}_n(\mathbb{R})$.

17°) Soit M une matrice de $O_n(\mathbb{R})$. Alors, pour tout vecteur unitaire X de $\mathbb{R}^n$, on a $\| MX \| = 1$, d'où $\| M \|_2 = 1$.
L'ensemble $O_n(\mathbb{R})$ est donc contenu dans la boule unité $\mathcal{B}$ de $\big(\mathcal{M}_n(\mathbb{R}),\, \|\,.\,\|_2\ \big)$ et, comme la boule $\mathcal{B}$ est convexe, il vient $\mathrm{conv}(O_n(\mathbb{R})) \subset \mathcal{B}$.

18°) D'après la question **11°)**, le projeté N de M sur $\mathrm{conv}(O_n(\mathbb{R}))$ est caractérisé par

$$\forall V \in \mathrm{conv}(O_n(\mathbb{R})), \qquad < M - N\,,\, V - N > \;\leqslant\; 0,$$

ce qui se traduit par l'inégalité $\mathrm{tr}\,(AV) \leqslant \mathrm{tr}\,(AM)$. De plus, comme $M \notin \mathrm{conv}(O_n(\mathbb{R}))$,

$$\mathrm{tr}\,(AN) - \mathrm{tr}\,(AM) \;=\; < M - N\,,\, N > \;-\; < M - N\,,\, M > \;=\; -\, \| M - N \|_2^2 \;<\; 0.$$

En écrivant A sous la forme US avec $U \in O_n(\mathbb{R})$ et S symétrique positive, il vient $\forall V \in \mathrm{conv}(O_n(\mathbb{R}))$, $\mathrm{tr}\,(USV) < \mathrm{tr}\,(USM)$. En particulier, pour $V = {}^tU$, on obtient

$$\mathrm{tr}\,(US\,{}^tU) \;=\; \mathrm{tr}\,({}^tUUS) \;=\; \mathrm{tr}\,(S) \;<\; \mathrm{tr}\,(USM).$$

19°) S étant symétrique réelle, il existe une base orthonormée $(e_1, e_2, \dots, e_n)$ formée de vecteurs propres de S.
En posant $Se_i = \lambda_i e_i$ (avec $\lambda_i \geqslant 0$), l'inégalité de Cauchy-Schwarz donne alors :

$$< MUSe_i\,,\, e_i > \;=\; \lambda_i < MUe_i\,,\, e_i > \;\leqslant \lambda_i\, \| MUe_i \| \times \| e_i \| \;\leqslant\; \lambda_i\, \| Ue_i \| \times \| e_i \| \;=\; \lambda_i\, \| e_i \|^2 \;=\; \lambda_i.$$

En appliquant la question **1°)**,

$$\mathrm{tr}\,(MUS) \;=\; \sum_{i=1}^{n} < MUSe_i\,,\, e_i > \;\leqslant\; \sum_{i=1}^{n} \lambda_i \;=\; \mathrm{tr}\,(S).$$

20°) Or $\mathrm{tr}\,(MUS) = \mathrm{tr}\,(USM)$: les inégalités des questions **18°)** et **19°)** conduisent alors à $\mathrm{tr}\,(S) < \mathrm{tr}\,(S)$.
L'hypothèse $M \notin \mathrm{conv}(O_n(\mathbb{R}))$ amenant à une contradiction, on en déduit que $\mathcal{B} \subset \mathrm{conv}(O_n(\mathbb{R}))$, d'où finalement $\mathrm{conv}(O_n(\mathbb{R})) = \mathcal{B}$ d'après le **16°)**.

F. Points extrémaux

21°) D'après l'inégalité triangulaire,

$$\|X\| = \|UX\| = \frac{1}{2}\|VX + WX\| \leqslant \frac{1}{2}(\|VX\| + \|WX\|) = \frac{1}{2}(\|X\| + \|X\|) = \|X\| .$$

La norme $\|.\|$ étant euclidienne, les vecteurs VX et WX sont donc colinéaires (et de même sens).

Il en résulte que ${}^tVVX = X$ et ${}^tVWX$ sont colinéaires pour tout vecteur $X \in \mathbb{R}^n$ ce qui entraîne, par un résultat classique que nous admettrons à ce stade du problème, que ${}^tVW$ est une matrice scalaire de la forme αI_n. On obtient alors $W = \alpha V$ mais, comme V et W sont orthogonales, on a nécessairement $\alpha = \pm 1$.

Le cas $\alpha = -1$ est impossible (car on aurait $U = 0$). Il reste donc $\alpha = 1$, ce qui conduit à $U = V = W$, c'est-à-dire au fait que U soit extrémal dans $\mathcal{B}$.

22°) En utilisant à nouveau la décomposition polaire, A s'écrit sous la forme US avec $U \in O_n(\mathbb{R})$ et S symétrique positive.
Or, d'après le théorème spectral, il existe $Q \in O_n(\mathbb{R})$ et D diagonale à coefficients diagonaux positifs tels que $S = Q^{-1}DQ$.

On obtient alors $A = (UQ^{-1})DQ$ et il suffit de poser $P = UQ^{-1} \in O_n(\mathbb{R})$ pour conclure.

23°) – Soit $X = Q^{-1}e_i$, où e_i désigne le i-ième vecteur de la base canonique de $\mathbb{R}^n$. On a alors $\|QX\| = \|X\| = 1$ et comme A appartient à $\mathcal{B}$, il vient $\|AX\| \leqslant 1$.

Or $AX = PDe_i = P(d_ie_i)$ donc, comme d_i est positif, $\|AX\| = d_i\,\|Pe_i\| = d_i\,\|e_i\| = d_i$, ce qui conduit à $d_i \leqslant 1$.

– Si tous les coefficients d_i valaient 1, D serait égale à la matrice identité I_n, d'où $A = PQ \in O_n(\mathbb{R})$: impossible. Il existe donc un indice $j \in [\![1, n]\!]$ tel que $d_j < 1$.

24°) Soit $\alpha = 1 - \alpha_j$ avec les notations du **23°)**. Appelons D_α, resp. $D_{-\alpha}$, la matrice diagonale dont les coefficients diagonaux sont les mêmes que ceux de D, à l'exception du j-ième qui vaut $d_j + \alpha$, resp. $d_j - \alpha$. Posons enfin $A_\alpha = PD_\alpha Q$ et $A_{-\alpha} = PD_{-\alpha}Q$.
Pour tout vecteur unitaire X de $\mathbb{R}^n$,

$$\|A_\alpha X\| = \|PD_\alpha QX\| = \|D_\alpha QX\| \leqslant \|D_\alpha\|_2 \,.\, \|QX\| = \|D_\alpha\|_2 \,.\, \|X\| = \|D_\alpha\|_2 \,.$$

Or, d'après le **4°)**, $\|D_\alpha\|_2 \leqslant 1$ étant donné que D_α est symétrique positive et que ses valeurs propres sont majorées par 1.
En observant que $\|D_{-\alpha}\|_2 = \|\,|D_{-\alpha}|\,\|_2$, où $|D_{-\alpha}|$ est la matrice diagonale dont les coefficients sont les valeurs absolues de ceux de $D_{-\alpha}$, on obtient de même $\|D_{-\alpha}\|_2 \leqslant 1$.
On a ainsi construit deux matrices A_α et $A_{-\alpha}$ de $\mathcal{B}$ telles que $A = \dfrac{1}{2}\big(A_\alpha + A_{-\alpha}\big)$ et $A_\alpha \neq A$: la matrice A n'est donc pas extrémale.

Conclusion : les points extrémaux de $\mathcal{B}$ sont exactement les matrices orthogonales $A \in O_n(\mathbb{R})$.