

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 $\langle x, e_i \rangle$. En particulier, $\langle Ae_i, e_i \rangle$ 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, $\sum_{i=1}^n \langle Ae_i, e_i \rangle$ 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 $\text{tr}(A)$ par invariance de la trace par changement de base.
- 2°) – L'application $\langle, \rangle$ est symétrique car $\forall A, B \in \mathcal{M}_n(\mathbb{R}), \text{tr}({}^tAB) = \text{tr}({}^t({}^tAB)) = \text{tr}({}^tBA)$.
 – 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 $\langle A, A \rangle = \sum_{i=1}^n \langle {}^tAAe_i, e_i \rangle$ soit $\sum_{i=1}^n \langle Ae_i, Ae_i \rangle$, d'où $\langle A, A \rangle = \sum_{i=1}^n \|Ae_i\|^2 \geq 0$ avec égalité si et seulement si $\forall i \in \llbracket 1, n \rrbracket, Ae_i = 0$, c'est-à-dire ssi A est nulle.
 L'application $\langle, \rangle$ 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 $\forall x \in \mathbb{R}^n, \langle Ax, x \rangle \geq 0$.
 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 :

$$\langle A, B \rangle = \text{tr}({}^tAB) = \sum_{i=1}^n \langle {}^tABe_i, e_i \rangle = \sum_{i=1}^n \mu_i \langle Ae_i, e_i \rangle,$$

quantité qui est bien positive puisque $\forall x \in \mathbb{R}^n, \langle Ax, x \rangle \geq 0$ et $\forall i \in \llbracket 1, n \rrbracket, \mu_i \geq 0$.

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 $\langle {}^tAAx, x \rangle = \|Ax\|^2 \geq 0$ donc tAA est positive.
 – Soit $\lambda_1, \lambda_2, \dots, \lambda_n$ les valeurs propres de tAA , avec $0 \leq \lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_n$, et $(e_1, e_2, \dots, e_n)$ une base orthonormée de vecteurs propres associés. Alors, pour tout vecteur unitaire $X = \sum_{i=1}^n x_i e_i$ de $\mathbb{R}^n$, on a

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

d'où $\|A\|_2 \leq \sqrt{\lambda_n}$. 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 $\forall i \in \llbracket 1, n \rrbracket, h(e_i) = \sqrt{\lambda_i} e_i$ est alors autoadjoint positif et, par construction, $f^* \circ f = h^2$.
- 6°) Si $x \in \text{Ker } \tilde{h}$, alors il existe $y \in E$ tel que $x = h(y)$ et $h(x) = 0$.
 Or, par construction de h , on a clairement $\text{rg}(h) = \text{rg}(h^2)$ et, comme $\text{Ker } h \subset \text{Ker } h^2$, il vient $\text{Ker } h = \text{Ker } h^2$ 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 $\text{Im } h$ est un espace vectoriel de dimension finie, $\tilde{h}$ définit bien un automorphisme de $\text{Im } h$.

- 7°) – Pour tout $x \in \mathbb{R}^n$,

$$\|f(x)\|^2 = \langle f^* \circ f(x), x \rangle = \langle h^2(x), x \rangle = \langle h^* \circ h(x), x \rangle = \|h(x)\|^2,$$

d'où $\|h(x)\| = \|f(x)\|$.

- Il en ressort en particulier que $\text{Ker } h = \text{Ker } f$, d'où $\dim \text{Ker } h = \dim \text{Ker } f = \dim(\text{Im } f)^\perp$.
 – N'importe quelle application linéaire v envoyant une base orthonormée de $\text{Ker } h$ sur une base orthonormée de $(\text{Im } f)^\perp$ conserve alors la norme donc réalise un isomorphisme de $\text{Ker } h$ sur $(\text{Im } f)^\perp$.

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

– Les sous-espaces $\text{Ker } h$ et $\text{Im } h$ d'une part, et $\text{Im } f$ et $(\text{Im } f)^\perp$ 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 $\text{Im } h$ et v sur $\text{Ker } h$.

– De plus, l'application $f \circ \tilde{h}^{-1}$ conserve la norme : en effet, $\forall x \in \text{Im } h$, $\|f \circ \tilde{h}^{-1}(x)\| = \|h \circ \tilde{h}^{-1}(x)\| = \|x\|$. Comme v conserve également la norme des vecteurs de $\text{Ker } h$, 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 $\text{Ker } h$ et $\text{Im } h$, 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 $d_x : h \mapsto \|x - h\|$ 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,

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

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 $q(t) \geq \|x - h_0\|^2$. 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 :

$$\forall t \in [0, 1], \quad (1-t) \cdot \langle x - h_0, h_0 - h_1 \rangle + (1-t)^2 \|h_0 - h_1\|^2 \geq 0.$$

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

– Réciproquement, si on a $\forall h_1 \in H$, $\langle x - h_0, h_0 - h_1 \rangle \geq 0$, alors $\forall t \in [0, 1]$, $q(t) \geq \|x - h_0\|^2$ en remontant les calculs précédents. En particulier, $q(0) \geq \|x - h_0\|^2$, ce qui signifie que $\forall h_1 \in H$, $\|x - h_1\| \geq \|x - h_0\|$ 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 $CC(H) \subset \text{conv}(H)$.

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 $\forall i \in \llbracket 1, p \rrbracket$, $\lambda_i \geq 0$, $x_i \in H$ et $\sum_{i=1}^p \lambda_i = 1$, et de même $\forall j \in \llbracket 1, q \rrbracket$, $\mu_j \geq 0$, $y_j \in H$ 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 $CC(H) = \text{conv}(H)$.

13°) La famille $(x_2 - x_1, x_3 - x_1, \dots, x_p - x_1)$ est liée car de cardinal $p - 1 \geq n + 1$ dans un espace vectoriel de dimension n . Il existe donc des réels non tous nuls $\mu_2, \mu_3, \dots, \mu_p$ 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

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

- 14°) Pour un indice i fixé, l'ensemble des réels θ tels que $\lambda_i - \theta\mu_i \geq 0$ 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 $\{\theta \in \mathbb{R} / \forall i \in \llbracket 1, p \rrbracket, \lambda_i - \theta\mu_i \geq 0\}$ est un segment. Choisissons pour θ l'une de ses bornes : il existe alors un indice $j \in \llbracket 1, p \rrbracket$ tel que $\lambda_j - \theta\mu_j = 0$ et, pour tout $i \neq j$, $\lambda_i - \theta\mu_i \geq 0$. De plus,

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

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 $\forall i \in \llbracket 1, n+1 \rrbracket, 0 \leq t_i \leq 1$). D'après les questions 12°) et 14°), $\text{conv}(H)$ 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

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

(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, $\text{conv}(H) = \phi(\Lambda \times H^{n+1})$ est donc un compact de E .

E. Enveloppe convexe de $O_n(\mathbb{R})$

- 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 $M \mapsto {}^t M M$. 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 $\text{conv}(O_n(\mathbb{R}))$ 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 $(\mathcal{M}_n(\mathbb{R}), \|\cdot\|_2)$ et, comme la boule $\mathcal{B}$ est convexe, il vient $\text{conv}(O_n(\mathbb{R})) \subset \mathcal{B}$.
- 18°) D'après la question 11°), le projeté N de M sur $\text{conv}(O_n(\mathbb{R}))$ est caractérisé par

$$\forall V \in \text{conv}(O_n(\mathbb{R})), \quad \langle M - N, V - N \rangle \leq 0,$$

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

$$\text{tr}(AN) - \text{tr}(AM) = \langle M - N, N \rangle - \langle M - N, M \rangle = -\|M - N\|_2^2 < 0.$$

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

$$\text{tr}(US {}^t U) = \text{tr}({}^t UUS) = \text{tr}(S) < \text{tr}(USM).$$

- 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 $\lambda_i \geq 0$), l'inégalité de Cauchy-Schwarz donne alors :

$$\langle MUSE_i, e_i \rangle = \lambda_i \langle MUE_i, e_i \rangle \leq \lambda_i \|MUE_i\| \times \|e_i\| \leq \lambda_i \|Ue_i\| \times \|e_i\| = \lambda_i \|e_i\|^2 = \lambda_i.$$

En appliquant la question 1°),

$$\text{tr}(MUS) = \sum_{i=1}^n \langle MUSE_i, e_i \rangle \leq \sum_{i=1}^n \lambda_i = \text{tr}(S).$$

- 20°) Or $\text{tr}(MUS) = \text{tr}(USM)$: les inégalités des questions 18°) et 19°) conduisent alors à $\text{tr}(S) < \text{tr}(S)$.

L'hypothèse $M \notin \text{conv}(O_n(\mathbb{R}))$ amenant à une contradiction, on en déduit que $\mathcal{B} \subset \text{conv}(O_n(\mathbb{R}))$, d'où finalement $\text{conv}(O_n(\mathbb{R})) = \mathcal{B}$ d'après le 16°).

F. Points extrémaux

21°) D'après l'inégalité triangulaire,

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

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

Il en résulte que ${}^tV VX = X$ et ${}^tV WX$ 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 ${}^tV W$ 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 $S = Q^{-1} D Q$.

On obtient alors $A = (UQ^{-1}) D Q$ 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 $\|AX\| \leq 1$.

Or $AX = P D e_i = P(d_i e_i)$ donc, comme d_i est positif, $\|AX\| = d_i \|P e_i\| = d_i \|e_i\| = d_i$, ce qui conduit à $d_i \leq 1$.

– Si tous les coefficients d_i valaient 1, D serait égale à la matrice identité I_n , d'où $A = P Q \in O_n(\mathbb{R})$: impossible. Il existe donc un indice $j \in \llbracket 1, n \rrbracket$ 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 $A_\alpha = P D_\alpha Q$ et $A_{-\alpha} = P D_{-\alpha} Q$.

Pour tout vecteur unitaire X de $\mathbb{R}^n$,

$$\|A_\alpha X\| = \|P D_\alpha Q X\| = \|D_\alpha Q X\| \leq \|D_\alpha\|_2 \cdot \|QX\| = \|D_\alpha\|_2 \cdot \|X\| = \|D_\alpha\|_2.$$

Or, d'après le **4°)**, $\|D_\alpha\|_2 \leq 1$ 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 $\|D_{-\alpha}\|_2 \leq 1$.

On a ainsi construit deux matrices A_α et $A_{-\alpha}$ de $\mathcal{B}$ telles que $A = \frac{1}{2}(A_\alpha + A_{-\alpha})$ 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 $A \in O_n(\mathbb{R})$.