DÉVELOPPEMENT 12

ENVELOPPE CONVEXE DU GROUPE ORTHOGONAL

On considère l'espace $\mathcal{M}_n(\mathbb{R})$ muni de la norme $\||\ \||_2$ induite par la norme euclidienne de $\mathbb{R}^n$.

Lemme. — *Les formes linéaires sur $\mathcal{M}_n(\mathbb{K})$ sont les applications*
$$\mathcal{M}_n(\mathbb{K}) \to \mathbb{K}, M \mapsto \text{Tr }(AM) \quad \textit{où} \quad A \in \mathcal{M}_n(\mathbb{K}).$$

Démonstration. — On considère le morphisme $f : \mathcal{M}_n(\mathbb{K}) \to \mathcal{M}_n(\mathbb{K})', A \mapsto f_A$ où $f_A(M) = \text{Tr }(AM)$. Il s'agit de montrer que f est un isomorphisme *i.e.* (puisque $\dim \mathcal{M}_n(\mathbb{K}) = \dim \mathcal{M}_n(\mathbb{K})'$) de montrer que f est injective. Si $A = (a_{i,j})$ est telle que $f_A = 0$ alors, pour tous $1 \leq i, j \leq n$, on a
$$0 = f_A(E_{i,j}) = \text{Tr }(AE_{i,j})$$
mais
$$AE_{i,j} = \sum_{1\leq k,l\leq n} a_{k,l}E_{k,l}E_{i,j} = \sum_{1\leq k,l\leq n} a_{k,l}\delta_{li}E_{k,l}E_{i,j} = \sum_{k=1}^{n} a_{k,i}E_{k,i}E_{i,j} = \sum_{k=1}^{n} a_{k,i}E_{k,j}$$
d'où
$$0 = \text{Tr }(AE_{i,j}) = \text{Tr } \sum_{k=1}^{n} a_{k,i}E_{k,j} = \sum_{k=1}^{n} a_{k,i}\text{Tr }(E_{k,j}) = \sum_{k=1}^{n} a_{k,i}\delta_{kj} = a_{j,i}$$
i.e. $A = 0$. $\square$

Théorème. — *L'enveloppe convexe de $\mathcal{O}(n)$ dans $\mathcal{M}_n(\mathbb{R})$ est la boule unité.*

Démonstration. — Il est clair que $\mathbb{B}_{\mathcal{M}_n(\mathbb{K})}$ contient l'enveloppe convexe de $\mathcal{O}(n)$, on considère donc une matrice $M \in \mathcal{M}_n(\mathbb{K})$ telle que $\|M\|_2 \leq 1$. D'après un corollaire du théorème de Hahn-Banach, pour montrer que M est dans l'enveloppe convexe de $\mathcal{O}(n)$, il suffit de montrer que
$$\varphi(M) \leq \sup_{O\in\mathcal{O}(n)} \varphi(O)$$
pour toute forme linéaire φ sur $\mathcal{M}_n(\mathbb{K})$. D'après le lemme, cela revient à montrer que
$$\text{Tr }(AM) \leq \sup_{O\in\mathcal{O}(n)} \text{Tr }(AO)\ ,\ \forall A \in \mathcal{M}_n(\mathbb{K}).$$
On considère une décomposition polaire $A = \Omega S$ de A (*i.e.* Ω est orthogonale et S est symétrique positive) et une base orthonormale $(e_1, \ldots, e_n)$ de $\mathbb{R}^n$ formée de vecteurs propres de S, alors
$$\sup_{O\in\mathcal{O}(n)} \text{Tr }(AO) \geq \text{Tr }(A\Omega^{-1}) = \text{Tr }(\Omega^{-1}A) = \text{Tr }(S) = \sum_{i=1}^{n} \|Se_i\|_2\,.$$
D'autre part, on a
$$\text{Tr }(AM) = \text{Tr }(MA) = \sum_{i=1}^{n}\langle MAe_i, e_i\rangle = \sum_{i=1}^{n}\langle Ae_i, M^*e_i\rangle$$

d'où d'après l'inégalité de Cauchy-Schwarz

$$\mathrm{Tr}\ (AM) \leq \sum_{i=1}^{n} \|Ae_i\|_2 \|M^* e_i\|_2 \leq \sum_{i=1}^{n} \|Ae_i\|_2 \,\|\!|M^*|\!\|_2 \,\|e_i\|_2 \,.$$

Mais $|\!|\!|M|\!|\!|_2 \leq 1$ implique que $|\!|\!|M^*|\!|\!|_2 \leq 1$ et la base $(e_1, \ldots, e_n)$ est orthonormale donc

$$\mathrm{Tr}\ (AM) \leq \sum_{i=1}^{n} \|Ae_i\|_2 \leq \sum_{i=1}^{n} \|\Omega S e_i\|_2 = \sum_{i=1}^{n} \|S e_i\|_2$$

et on a finalement bien $\mathrm{Tr}\ (AM) \leq \sup_{O \in \mathcal{O}(n)} \mathrm{Tr}\ (AO)$. □

On rappelle qu'un élément U de $\mathbb{B}$ est dit *extremal* si toute écriture du type $U = \frac{1}{2}(V + W)$ avec $V, W \in \mathbb{B}$ implique $U = V = W$.

Théorème. — *$\mathcal{O}(n)$ est l'ensemble des points extrémaux de la boule unité.*

Démonstration. — Notons tout d'abord que si $|\!|\!|U|\!|\!| < 1$ alors U n'est pas extrémal ; en effet, si $U = 0$ alors $U = \frac{1}{2}(\mathrm{I} + (-\mathrm{I}))$ et si $U \neq 0$ alors $U = \frac{1}{2}\left(\frac{1}{|\!|\!|U|\!|\!|}U + (2 - \frac{1}{|\!|\!|U|\!|\!|})U\right)$.

D'autre part, tout élément $U \in \mathcal{O}(n)$ est extrémal ; en effet, écrivons $U = \frac{1}{2}(V + W)$ alors, pour tout $x \in \mathbb{R}^n$, on a $2Ux = Vx + Wx$ d'où

$$4\|x\|^2 = \|2Ux\|^2 = \|Vx\|^2 + \|Wx\|^2 + 2\langle Vx, Wx\rangle \leq |\!|\!|V|\!|\!|^2 \|x\|^2 + |\!|\!|W|\!|\!|^2 \|x\|^2 + |\!|\!|V|\!|\!| |\!|\!|W|\!|\!| \|x\|^2 \leq 4\|x\|^2$$

ce qui implique que les inégalités ci-dessus sont en fait des égalités *i.e.* on a

$$\|Vx\| = \|x\| \ , \ \|Wx\| = \|x\| \quad \text{et} \quad \langle Vx, Wx\rangle = \|Vx\| \, \|Wx\| \, ;$$

la dernière égalité implique que Vx et Wx sont positivement liés et le deux premières montrent donc qu'on a en fait $Vx = Wx$, d'où $U = V = W$.

Soit A un élément extrémal de la boule unité, on en considère une décomposition polaire $A = SO$, ce qui peut aussi s'écrire

$$A = {}^t\Omega D \Omega O \quad \text{où} \quad D = \begin{bmatrix} d_1 & & \\ & \ddots & \\ & & d_n \end{bmatrix}$$

et $\Omega, O \in \mathcal{O}(n)$ et $\lambda_1, \ldots, \lambda_n \geq 0$. D'autre part on a $|\!|\!|A|\!|\!| = |\!|\!|D|\!|\!| = \sup_{1 \leq i \leq n} d_i$ donc $0 \leq d_i \leq 1$ pour tout i. Supposons que l'un des d_i soit non nul, par exemple $d_1 \neq 0$, et posons

$$D_1 = \begin{bmatrix} 1 & & & \\ & d_2 & & \\ & & \ddots & \\ & & & d_n \end{bmatrix} \quad \text{et} \quad D_2 = \begin{bmatrix} 2d_1 - 1 & & & \\ & d_2 & & \\ & & \ddots & \\ & & & d_n \end{bmatrix}$$

puis $V = {}^t\Omega D_1 \Omega O$ et $W = {}^t\Omega D_2 \Omega O$ alors $V \neq W$, $|\!|\!|V|\!|\!| = |\!|\!|D_1|\!|\!| \leq 1$, $|\!|\!|W|\!|\!| = |\!|\!|D_2|\!|\!| \leq 1$ et $A = \frac{1}{2}(U + W)$ ce qui contredit le caractère extrémal de A. Par conséquent, tous les d_i sont nuls *i.e.* $A = {}^t\Omega\Omega O = O \in \mathcal{O}(n)$. □

Leçons concernées

07 Groupe linéaire d'un espace vectoriel de dimension finie E, sous-groupes de $GL(E)$. Applications

30 Barycentres dans un espace affine réel de dimension finie ; convexité. Applications

40 Exemples de décompositions remarquables dans le groupe linéaire. Applications

45 Formes linéaires et hyperplans en dimension finie. Exemples et applications

Compléments

Applications de la caractérisation du dual de $\mathcal{M}_n(\mathbb{K})$. —
On a vu que les formes linéaires sur $\mathcal{M}_n(\mathbb{K})$ sont les applications

$$\mathcal{M}_n(\mathbb{K}) \to \mathbb{K}, M \mapsto \mathrm{Tr}\,(AM) \quad \text{où} \quad A \in \mathcal{M}_n(\mathbb{K}).$$

Proposition. — *Si f est une forme linéaire sur $\mathcal{M}_n(\mathbb{K})$ telle que $f(MN) = f(NM)$ pour tous M, N de $\mathcal{M}_n(\mathbb{K})$ alors il existe $\lambda \in \mathbb{K}$ tel que $f = \lambda \mathrm{Tr}$.*

Démonstration. — Soit $A \in \mathcal{M}_n(\mathbb{K})$ telle que $f(M) = \mathrm{Tr}\,(AM)$ pour tout $M \in \mathcal{M}_n(\mathbb{K})$; l'hypothèse s'écrit donc $\mathrm{Tr}\,(AMN) = \mathrm{Tr}\,(ANM)$, pour tous $M, N \in \mathcal{M}_n(\mathbb{K})$, *i.e.* $\mathrm{Tr}\,((AM - MA)N) = 0$. Puisque, pour M fixée, la forme linéaire $N \mapsto \mathrm{Tr}\,((AM - MA)N)$ est nulle, c'est donc que $AM = MA$. Ainsi, A commute avec toute matrice $M \in \mathcal{M}_n(\mathbb{K})$ or le centre de $\mathcal{M}_n(\mathbb{K})$ est composé des homothéties donc $A = \lambda \mathrm{I}$ et il s'ensuit que $f = \lambda \mathrm{Tr}$. □

Remarque. — Le fait que le centre de $\mathcal{M}_n(\mathbb{K})$ soit composé des homothéties peut se montrer de la façon suivante : $A = (a_{i,j})_{i,j}$ commute avec la matrice $E_{i,j}$ de la base canonique alors

$$\sum_{k=1}^{n} a_{k,i} E_{k,j} = \sum_{1 \leq k,l \leq n} a_{k,l} E_{k,l} E_{i,j} = AE_{i,j} = E_{i,j} A = \sum_{1 \leq k,l \leq n} a_{k,l} E_{i,j} E_{k,l} = \sum_{l=1}^{n} a_{j,l} E_{i,l}$$

d'où $a_{k,i} = 0$ pour $k \neq i$ et $a_{i,i} = a_{j,j}$.

Remarque. — On peut donner une preuve directe de la proposition. Soit $1 \leq i, j \leq n$ avec $i \neq j$, on a

$$f(E_{i,j}) = f(E_{i,i}E_{i,j}) = f(E_{i,j}E_{i,i}) = f(0) = 0 \quad \text{et} \quad f(E_{i,i}) = f(E_{i,j}E_{j,i}) = f(E_{j,i}E_{i,j}) = f(E_{j,j}).$$

Si on note λ la valeur commune des $f(E_{i,i})$, on vérifie que les formes linéaires f et $\lambda \mathrm{Tr}$ coïncident sur une base de $\mathcal{M}_n(\mathbb{K})$ donc sont égales.

Proposition. — *$GL_n(\mathbb{K})$ coupe tout hyperplan de $\mathcal{M}_n(\mathbb{K})$.*

Démonstration. — Un hyperplan de $\mathcal{M}_n(\mathbb{K})$ est le noyau d'une forme linéaire $M \mapsto \mathrm{Tr}\,(AM)$ donc il s'agit de trouver M inversible telle que $\mathrm{Tr}\,(AM) = 0$. Notons r le rang de A alors il existe P, Q inversibles telles que

$$PAQ = \begin{bmatrix} \mathrm{I}_r & 0 \\ 0 & 0 \end{bmatrix} =: J_r$$

or $\mathrm{Tr}\,(AM) = \mathrm{Tr}\,(P^{-1}J_rQ^{-1}M) = \mathrm{Tr}\,(J_rQ^{-1}MP^{-1})$ *i.e.* il s'agit de trouver N inversible telle que $\mathrm{Tr}\,(J_rN) = 0$ (on pose alors $M = QNP$). On considère la matrice de permutation

$$N = \begin{bmatrix} 0 & 1 \\ I_{n-1} & 0 \end{bmatrix}$$

qui est bien inversible et telle que J_rN soit de diagonale nulle. □

Proposition. — *Soit $A, B \in \mathcal{M}_n(\mathbb{K})$, on a équivalence entre*

(*i*) *il existe $M \in \mathcal{M}_n(\mathbb{K})$ telle que $AM + MA = B$*

(ii) pour tout $C \in \mathcal{M}_n(\mathbb{K})$ telle que $AC + CA = 0$, on a $\mathrm{Tr}\,(BC) = 0$

Démonstration. — $(i) \Rightarrow (ii)$ Si $C \in \mathcal{M}_n(\mathbb{K})$ vérifie $AC + CA = 0$ alors

$$\mathrm{Tr}\,(BC) = \mathrm{Tr}\,((AM + MA)C) = \mathrm{Tr}\,(AMC) + \mathrm{Tr}\,(MAC) = \mathrm{Tr}\,(CAM) + \mathrm{Tr}\,(ACM)$$

i.e. $\mathrm{Tr}\,(BC) = \mathrm{Tr}\,((CA + AC)M) = 0$.

$(ii) \Rightarrow (i)$ On considère l'endomorphisme $f : \mathcal{M}_n(\mathbb{K}) \to \mathcal{M}_n(\mathbb{K}), M \mapsto AM + MA$, il s'agit de montrer que $B \in \mathrm{Im}\ f$. L'application $T : C \mapsto T_C$, où $T_C(M) = \mathrm{Tr}\,(CM)$, définit un isomorphisme de $\mathcal{M}_n(\mathbb{K})$ sur son dual et la condition (ii) signifie que $T_C(B) = 0$ dès que $C \in \ker f$ *i.e.*

$$B \in (T(\ker f))^\circ = \{N \in \mathcal{M}_n(\mathbb{K})\ ;\ \forall \varphi \in T(\ker f), \varphi(N) = 0\}.$$

Mais la première implication donne $\mathrm{Im}\ f \subset (T(\ker f))^\circ$ or (puisque T est un isomorphisme)

$$\dim(T(\ker f))^\circ = \dim \mathcal{M}_n(\mathbb{K}) - \dim T(\ker f) = \dim \mathcal{M}_n(\mathbb{K}) - \dim \ker f = \dim \mathrm{Im}\ f$$

d'où $B \in \mathrm{Im}\ f$. □

Preuve du théorème de Hahn-Banach et de ses corollaires. —

Théorème de Hahn-Banach. — *Soit E un espace vectoriel normé, M un sous-espace de E et A un ouvert convexe non vide de E tel que $M \cap A = \emptyset$. Alors il existe un hyperplan linéaire fermé H de E tel que $M \subset H$ et $H \cap A = \emptyset$.*

Démonstration. — Notons $\mathcal{F}$ l'ensemble des sous-espaces N de E qui contiennent M et n'intersectent pas A, on ordonne $\mathcal{F}$ par inclusion de sorte que ce soit un ensemble inductif, alors le lemme de Zorn donne un élément maximal H ; on pose alors

$$\Omega = H + \bigcup_{\lambda > 0} \lambda A = \bigcup_{h \in H} \left(h + \bigcup_{\lambda > 0} \lambda A \right)$$

qui est un ouvert de E.

• On a $\Omega \cap (-\Omega) = \emptyset$. En effet, sinon il existe $x = h_1 + \lambda_1 a_1 = h_2 - \lambda_2 a_2$ avec $h_1, h_2 \in H$, $\lambda_1, \lambda_2 > 0$ et $a_1, a_2 \in A$; on peut alors écrire

$$\frac{\lambda_1}{\lambda_1 + \lambda_2} a_1 + \frac{\lambda_2}{\lambda_1 + \lambda_2} a_2 = \frac{1}{\lambda_1 + \lambda_2}(h_1 - h_2) \in H$$

alors que cet élément appartient à A par convexité, c'eest impossible puisque $H \cap A = \emptyset$.

• On a $E = H \cup \Omega \cup (-\Omega)$. En effet, sinon on considère $x \in E \setminus (H \cup \Omega \cup (-\Omega))$ puis on pose $\widetilde{H} = H \oplus \mathbb{R}x$ alors $H \subsetneq \widetilde{H}$ donc, par maximalité de H, on doit avoir $\widetilde{H} \cap A \neq \emptyset$ *i.e.* il existe $x \in H$ et $\lambda \neq 0$ tels que $y = h + \lambda x \in \widetilde{H} \cap A$. Mais comme $y \in A$, on a

$$x = -\frac{1}{\lambda} h + \frac{1}{\lambda} y \in \Omega \cup (-\Omega)$$

ce qui contredit le choix de x.

• On a $H \cap (\Omega \cup (-\Omega)) = \emptyset$. En effet, puisque H coupe Ω si et seulement si H coupe $(-\Omega)$, il suffit de montrer que $H \cap \Omega = \emptyset$, on suppose donc qu'il existe $x = h + \lambda a$ dans H où $h \in H$, $\lambda > 0$ et $a \in A$, alors $a = \frac{1}{\lambda}(x - h) \in A \cap H$ ce qui est impossible.

• Puisque Ω est ouvert et $H = E \setminus (\Omega \cup (-\Omega))$, H est fermé dans E.

• Enfin, H est un hyperplan linéaire. En effet, considérons un élément x non nul dans $\Omega \setminus H$ et posons $\widetilde{H} = H \oplus \mathbb{R}x$. Si $\widetilde{H} \neq E$ alors il existe $y \in (-\Omega)$ tel que $y \notin \widetilde{H}$ (on peut prendre y dans $(-\Omega)$ puisque $(-\Omega) \subset \widetilde{H}$ implique $\Omega \subset \widetilde{H}$) et on considère alors l'application

$$f : [0, 1] \to E, t \mapsto tx + (1 - t)y.$$

On a $0 \in f^{-1}(-\Omega)$ et $1 \in f^{-1}(\Omega)$ or $f^{-1}(-\Omega)$ et $f^{-1}(\Omega)$ sont deux ouverts non vides du connexe $[0,1]$ qui sont disjoints puisque $\Omega \cap (-\Omega) = \emptyset$, il s'ensuit que $f^{-1}(-\Omega) \cup f^{-1}(-\Omega) \subsetneq [0,1]$. Ainsi, il existe $t \in]0,1[$ tel que $f(t) \in H$ *i.e.*

$$y = \frac{1}{1-t}(f(t) - tx) \in H \oplus \mathbb{R}x = \widetilde{H}$$

ce qui est impossible par choix de y. On a donc $H \oplus \mathbb{R}x = \widetilde{H} = E$ *i.e.* H est un hyperplan. □

Corollaire. — *Soit E un $\mathbb{R}$-espace vectoriel normé, F un convexe fermé de E et C un convexe compact de E tels que $F \cap C = \emptyset$. Alors il existe une forme linéaire continue φ telle que :*

$$\sup_{x\in C} \varphi(x) < \inf_{y\in F} \varphi(y).$$

Démonstration. — On pose $G = F - C$, alors G est fermé et ne contient pas 0. En effet, $0 \notin G$ puisque $F \cap C = \emptyset$ et considérons deux suites $(x_n)_n$ et $(y_n)_n$ respectivement dans F et C telles que la suite $(z_n)_n$, où $z_n = x_n - y_n$, converge vers $z \in E$. Puisque C est compact, il existe $\psi : \mathbb{N} \to \mathbb{N}$ strictement croissante telle que la suite $(y_{\psi(n)})_n$ converge vers $y \in C$. Notons $x = y + z$ alors

$$\lim_{n\to+\infty} x_{\psi(n)} = \lim_{n\to+\infty} \left(y_{\psi(n)} + z_{\psi(n)}\right) = y + z = x.$$

Or F est fermé donc $x \in F$ i.e. $z = x - y \in F - C = G$. En particulier, il existe $r > 0$ tel que $\mathbb{B}(0,r) \cap G = \emptyset$.

On pose $A = G + \mathbb{B}(0,r) = G - \mathbb{B}(0,r)$, alors $0 \notin A$ et il existe une forme linéaire continue $\varphi : E \to \mathbb{R}$ telle que $\varphi(z) > 0$ pour tout $z \in A$. En effet, on a $0 \notin A$ et puisque A est un ouvert convexe, on pose $M = \{0\}$ et on applique le théorème de Hahn-Banach géométrique donc il existe un hyperplan fermé H tel que $H \cap A = \emptyset$. En écrivant $H = \ker \varphi$ avec $\varphi : E \to \mathbb{R}$ linéaire, on voit que φ est continue. Comme $\ker \varphi \cap A = \emptyset$, on a $0 \notin \varphi(A)$ avec $\varphi(A)$ convexe dans $\mathbb{R}$ donc $\varphi(A)$ est un intervalle et on a donc soit $\varphi(A) \subset]0,+\infty[$, soit $\varphi(A) \subset]-\infty, 0[$; quitte à prendre $-\varphi$ au lieu de φ, on a bien $\varphi(z) > 0$ pour tout $z \in A$.

On a alors $m = \inf_{x\in G} \varphi(x) > 0$. En effet, supposons que $m = 0$, alors il existe une suite $(x_n)_n$ dans G telle que la suite $(\varphi(x_n))_n$ tende vers 0. Puisque φ est non nulle, il existe $u \in \mathbb{B}(0,r)$ tel que $\varphi(u) \neq 0$. On pose

$$v = -\frac{|\varphi(u)|}{\varphi(u)} \cdot u\ ,$$

alors on a $\|v\| = \|u\|$ donc $v \in \mathbb{B}(0,r)$. Comme $x_n + v \in G + \mathbb{B}(0,r) = A$, on a

$$0 < \varphi(x_n + v) = \varphi(x_n) + \varphi(v) = \varphi(x_n) - |\varphi(u)|$$

d'où

$$0 < |\varphi(u)| < \varphi(x_n) \xrightarrow[n\to+\infty]{} 0$$

ce qui est impossible. On a donc bien $m > 0$.

Considérons maintenant $x \in C$ et $y \in F$, alors $y - x \in G$ d'où $\varphi(y) - \varphi(x) \geq m$ donc

$$\forall y \in F,\ \sup_{x\in C} \varphi(x) < \sup_{x\in C} (m + \varphi(x)) \leq m + \varphi(y)$$

d'où

$$\sup_{x\in C} \varphi(x) \leq m \inf_{y\in F} \varphi(y) < \inf_{y\in F} \varphi(y).$$

□

Corollaire. — *Soit E un espace vectoriel normé et A une partie compacte de E. Alors $x \in E$ est adhérent à l'enveloppe convexe de A si et seulement si pour tout $\varphi \in E'$, on a*

$$\varphi(x) \leq \sup_{y\in A} \varphi(y).$$

Démonstration. — On pose $F = \{x\}$ et on note C l'adhérence de l'enveloppe convexe de A, alors F est un convexe fermé et C est un convexe compact (d'après le théorème de Caratheodory). Si $x \notin F$ alors $C \cap F = \emptyset$ avec C convexe compact et F convexe fermé donc le corollaire précédent donne une forme linéaire φ telle que

$$\sup_{y \in C} \varphi(y) < \inf_{y \in F} \varphi(y) \quad i.e. \quad \sup_{y \in C} \varphi(y) < \varphi(x).$$

Réciproquement, si $x \in C$ alors il existe $(x_n)_n$ dans l'enveloppe convexe de A tendant vers x, on a donc $\varphi(x_n) \leq \sup_{y \in C} \varphi(y)$ pour tout $\varphi \in E'$ et tout n, d'où $\varphi(x) \leq \sup_{y \in C} \varphi(y)$ par continuité de φ. □

Une caractérisation géométrique de $\mathcal{SO}(n)$ dans $SL_n(\mathbb{R})$. —

Proposition. — *On a $d_2(0, SL_n(\mathbb{R})) = \inf_{M \in SL_n(\mathbb{R})} \|M\|_2 = \sqrt{n}$ et le lieu de $SL_n(\mathbb{R})$ où cette distance est atteinte est exactement $\mathcal{SO}(n)$.*

Démonstration. — On considère les applications f et q de $\mathcal{M}_n(\mathbb{R})$ dans $\mathbb{R}$ définies pour $M = (m_{i,j})$ par

$$f(M) = \det M - 1 \quad \text{et} \quad q(M) = \|M\|_2^2 = \sum_{1 \leq i,j \leq n} m_{i,j}^2 = \text{Tr}\ ({}^tMM).$$

Il s'agit de deux fonctions de classe $\mathcal{C}^\infty$ puisque ce sont des fonctions polynômiales en les n^2 variables réelles $m_{i,j}$. De plus, on a $\frac{\partial q}{\partial m_{i,j}}(M) = 2m_{i,j}$ pour tous i, j, *i.e.* $\nabla q(M) = 2M$. Si $M_{i,j}$ désigne le cofacteur de $m_{i,j}$ alors $\det M = \sum_{j=1}^{n} m_{i,j} M_{i,j}$ mais $M_{i,j}$ ne dépend pas de la variable $m_{i,j}$ d'où $\frac{\partial f}{\partial m_{i,j}}(M) = M_{i,j}$ donc $\nabla f(M)$ est la comatrice $Com(M)$ de M. On souhaite minimiser l'expression $q(M)$ sous la contrainte $f(M) = 0$ (ce minimum existe bien puisque $SL_n(\mathbb{R})$ est un fermé de $\mathcal{M}_n(\mathbb{R})$) ; on rappelle le théorème des extrema liés :

Lemme. — *Soit $\mathcal{U}$ un ouvert de $\mathbb{R}^N$ et $u, v : \mathcal{U} \to \mathbb{R}^N$ de classe $\mathcal{C}^1$ telles que $\mathcal{V} = \{x \in \mathcal{U}; v(x) = 0\} \neq \emptyset$, $u_{|\mathcal{V}}$ a un extremum local en $a \in \mathcal{V}$ et $\nabla v(a) \neq 0$. Alors il existe $\lambda \in \mathbb{R}$ tel que $\nabla u(a) = \lambda \nabla v(a)$.*

Si $\inf_{M \in SL_n(\mathbb{R})} \|M\|_2$ est atteint en $A \in SL_n(\mathbb{R})$ alors il existe un réel μ tel que $A = \mu Com(A)$ or $\det A = \det Com(A) = 1$ d'où $\mu = 1$. Or $A^{-1} = {}^tCom(A)$ donc ${}^tAA = \mathrm{I}$ *i.e.* $A \in \mathcal{O}(n)$ d'où $A \in \mathcal{SO}(n)$. Réciproquement, si $A \in \mathcal{SO}(n)$ alors on a $q(A) = n$. □

Distance au groupe orthogonal. —

Proposition. — *Pour tout $M \in \mathcal{M}_n(\mathbb{R})$, on a $d(M, \mathcal{O}(n)) = \|\!|\sqrt{{}^tMM} - \mathrm{I}|\!\|_2$.*

Démonstration. — Notons tout d'abord que si S et T sont symétriques positives et si $\langle\ ,\ \rangle$ désigne le produit scalaire euclidien sur $\mathcal{M}_n(\mathbb{R})$ alors $\langle S, T \rangle \geq 0$. Par densité, il suffit de vérifie cela pour T symétrique définie positive ; on note $\sqrt{T}$ l'unique racine carré de T alors $R = \sqrt{T} S \sqrt{T}$ est symétrique positive et on a $\langle S, T \rangle = \text{Tr}\ (ST) = \text{Tr}\ (\sqrt{T} ST \sqrt{T}^{-1}) = \text{Tr}\ (R) \geq 0$.

On considère l'action de $\mathcal{O}(n) \times \mathcal{O}(n)$ sur $\mathcal{M}_n(\mathbb{R})$ donnée par $(\Omega_1, \Omega_2) \star M = \Omega_1 M \Omega_2^{-1}$ alors on a $\|(\Omega_1, \Omega_2) \star M\|_2 = \|M\|_2$ donc tous les points d'une même orbite sont à la même distance de $\mathcal{O}(n)$. Soit $M = SO$ une décomposition polaire de M, alors il existe $\Omega \in \mathcal{O}(n)$ telle que $S = {}^t\Omega D \Omega$ avec D diagonale à coefficients positifs ; il s'ensuit que D est dans l'orbite de M or on a

$$\|D - \mathrm{I}\|_2 = \left\| {}^t\Omega(D - \mathrm{I})\Omega \right\|_2 = \left\| {}^t\Omega D \Omega - \mathrm{I} \right\|_2 = \|S - \mathrm{I}\|_2 = \left\| \sqrt{M {}^tM} - \mathrm{I} \right\|_2$$

donc il reste à montrer que $\|D - \mathrm{I}\|_2$ est la distance de D à $\mathcal{O}(n)$.

Soit $U \in \mathcal{O}(n)$ et $\delta := \|D-U\|_2^2 - \|D-\mathrm{I}\|_2^2$, montrons que $\delta \geq 0$. En développant, on obtient

$$\delta = 2\langle \mathrm{I}-U, D\rangle = 2\langle \mathrm{I}-E, D\rangle \quad \text{où} \quad E = \frac{1}{2}(U +^t U).$$

Si $\| \ \|_2$ est la norme induite sur $\mathcal{M}_n(\mathbb{R})$ par la norme $\| \ \|_2$ de $\mathbb{R}^n$ alors $\|U\|_2 = 1$ donc $\|E\|_2 \leq 1$ et il s'ensuit que la matrice symétrique $I-E$ est positive puisque

$$\langle (\mathrm{I}-E)X, X\rangle = \|X\|_2^2 - \langle EX, X\rangle \geq \|X\|_2^2 (1 - \|E\|_2) \geq 0.$$

D'après la remarque préliminaire, on a donc $\delta = 2\langle \mathrm{I}-E, D\rangle \geq 0$. $\square$

Références

M. Alessandri, *Thèmes de géométrie. Groupes en situation géométrique*, Dunod, 1999.

S. Francinou, H. Gianella et S. Nicolas, *Oraux X-ENS, algèbre 1*, Cassini, 2001.

H. Queffélec et C. Zuily, *Éléments d'analyse pour l'agrégation*, Dunod, 2002.