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

↳ Refs : OA p 97 (lemme 1), Zuily Queffélec (lemme 2 et thm) p 205-20
Szpirglas (théorème et corollaire) p 344-34

lemme : (Hahn Banach géométrique, version faible dans les Hilbert)
Soit H un espace de Hilbert réel, C une partie (non vide) convexe et fermée de H, A une partie (non vide) quelconque de H, et on note $H' = \mathcal{L}c(H, \mathbb{R})$.
1) Si $x \in H \setminus C$, alors il existe $f \in H'$ telle que $f(x) > \sup_{y \in C} f(y)$
2) Si $x \in H$, on a l'équivalence : $x \in \overline{\mathrm{Conv}(A)} \Leftrightarrow \forall f \in H', f(x) \leq \sup_{y \in A} f(y)$.

Théorème : Soit $n \geq 1$. On note $\|\cdot\|_2$ la norme subordonnée de $M_n(\mathbb{R})$ associée à la norme euclidienne $\|\cdot\|_2$ sur $\mathbb{R}^n$ ie $\|M\|_2 = \sup_{\|x\|_2 = 1} \|Mx\|_2$ pour $M \in M_n(\mathbb{R})$.
On note $B = \{M \in M_n(\mathbb{R}), \|M\|_2 \leq 1\}$ la boule unité de $M_n(\mathbb{R})$ associée à cette norme. Alors $\mathrm{Conv}(O_n(\mathbb{R})) = B$.

Rappel : On dit que $M \in C$ un point d'un ensemble convexe est extrémal de C si M n'est pas le milieu de deux points distincts de C ie si $M = \frac{1}{2}(A+B)$ avec $A, B \in C \Rightarrow M = A = B$. On a alors les équivalences :
M est extrémal $\Leftrightarrow$ M n'est pas contenu dans un segment d'extrémités dans C et distinctes de M
$\Leftrightarrow$ M n'est pas combinaison convexe de points de C distincts de M
$\Leftrightarrow$ $C \setminus \{M\}$ est convexe

Corollaire : L'ensemble des points extrémaux de B est $O_n(\mathbb{R})$.

Preuve du lemme : 1) Soit $x \in H \setminus C$. Le théorème de projection sur un convexe fermé nous fournit $a = p_C(x)$ la projection de x sur C. On définit alors $f = \langle x-a, \cdot \rangle$. Tout d'abord, $f \in H'$ par Cauchy-Schwarz (et $\|f\| = \|x-a\|$). Ensuite, puisque $x \notin C$, x est différent de a, donc $f(x) - f(a) = f(x-a) = \|x-a\|^2 > 0$ ie $f(x) > f(a)$. De plus, la caractérisation de $p_C(x)$ par les angles obtus nous dit exactement que $\forall y \in C$, $f(y-a) \leq 0$ ie $f(y) \leq f(a)$. Ainsi on obtient $\sup_{y \in C} f(y) \leq f(a) < f(x)$, d'où $\sup_{y \in C} f(y) < f(x)$.

2) Tout d'abord, on applique la contraposée du point 1) à $C = \overline{\mathrm{Conv}(A)}$: soit $x \in H$, on a l'implication : $\forall f \in H', f(x) \leq \sup_{y \in \overline{\mathrm{Conv}(A)}} f(y) \Rightarrow x \in \overline{\mathrm{Conv}(A)}$.
L'implication réciproque est claire. Soit $f \in H'$, montrons $M_1 := \sup_{y \in \overline{\mathrm{Conv}(A)}} f(y) = \underbrace{\sup_{y \in A} f(y)}_{M_2}$
* $A \subset \overline{\mathrm{Conv}(A)}$ donc $M_2 \leq M_1$
* Soit $y \in \mathrm{Conv}(A)$. Il existe $n \in \mathbb{N}^*$, $\lambda_1, \dots, \lambda_n \in \mathbb{R}$ tels que $\sum_{i=1}^n \lambda_i = 1$ et $a_1, \dots, a_n \in A$ tels que $y = \sum_{i=1}^n \lambda_i a_i$. Alors $f(y) = \sum_{i=1}^n \lambda_i f(a_i) \leq M_2 \sum_{i=1}^n \lambda_i = M_2$, donc en passant au sup, on obtient $\sup_{y \in \mathrm{Conv}(A)} f(y) \leq M_2$, et donc par continuité de f, $M_1 \leq M_2$.

Ainsi $M_1 = M_2$, et on en déduit l'équivalence, $x \in \overline{\mathrm{Conv}(A)} \Leftrightarrow \forall f \in H', f(x) \leq \sup_{y \in A} f(y)$.

Preuve du théorème : Tout d'abord, observons deux faits qui nous seront utiles pour la suite :

* Puisque $O_n(\mathbb{R})$ est compact, le théorème de Carathéodory permet de dire que $\mathrm{Conv}(O_n(\mathbb{R}))$ est compacte (cf rappel sur Carathéodory), donc en particulier fermée. Ainsi $\overline{\mathrm{Conv}(O_n(\mathbb{R}))} = \mathrm{Conv}(O_n(\mathbb{R}))$.
* L'espace $M_n(\mathbb{R})$ muni du produit scalaire $\langle A, B\rangle = \mathrm{Tr}({}^tAB)$ (il s'agit en fait du produit scalaire canonique de $\mathbb{R}^{n^2}$) est un espace de Hilbert. En particulier, le théorème de représentation de Riesz nous dit que : $\forall f \in (M_n(\mathbb{R}))', \exists! M \in M_n(\mathbb{R}), f = \mathrm{Tr}(\cdot\, {}^tM)$ (on peut aussi le faire à la main)

Et maintenant, allons y pour l'égalité, par double inclusion :

▶ Si $O \in O_n(\mathbb{R})$, alors $\|O\|_2 = 1$, donc $O \in B$. Ainsi $O_n(\mathbb{R}) \subset B$, et puisque la boule unité est convexe, $\mathrm{Conv}(O_n(\mathbb{R})) \subset B$.

▶ Soit $X \in B$. D'après le deuxième point du lemme 1, on a l'équivalence

$$X \in \mathrm{Conv}(O_n(\mathbb{R})) \iff \forall f \in E', f(X) \leq \sup_{U \in O_n(\mathbb{R})} f(U)$$

$$\iff \forall M \in M_n(\mathbb{R}), \mathrm{Tr}(XM) \leq \sup_{U \in O_n(\mathbb{R})} \mathrm{Tr}(UM).$$

Soit $M \in M_n(\mathbb{R})$. D'après le théorème de décomposition polaire, il existe un couple $(O, S) \in O_n(\mathbb{R}) \times S_n^+(\mathbb{R})$ tel que $M = OS$. D'après le théorème spectral, S est diagonalisable en base orthonormée : soit $\lambda_1, \dots \lambda_n \in \mathbb{R}_+$ les valeurs propres de S et $(e_1, \dots e_n)$ une base orthonormée de vecteurs propres associés. Puisque $O^{-1} \in O_n(\mathbb{R})$, on a donc d'une part :

$\sup_{U \in O_n(\mathbb{R})} \mathrm{Tr}(UM) \geq \mathrm{Tr}(O^{-1}M) = \mathrm{Tr}(S)$. Donc $\sup_{U \in O_n(\mathbb{R})} \mathrm{Tr}(UM) \geq \sum_{i=1}^n \lambda_i$.

D'autre part,

$$\mathrm{Tr}(XM) = \sum_{i=1}^n \langle XMe_i, e_i\rangle = \sum_{i=1}^n \langle O\underbrace{Se_i}_{\lambda_i e_i}, {}^tXe_i\rangle \overset{\text{Cauchy-Schwarz}}{\leq} \sum_{i=1}^n \underbrace{\|\lambda_i Oe_i\|_2}_{=\lambda_i \underbrace{\|e_i\|_2}_{=1}} \underbrace{\|{}^tXe_i\|_2}_{\leq \|{}^tX\|_2 \underbrace{\|e_i\|_2}_{=1}}$$

$$\leq \sum_{i=1}^n \lambda_i \underbrace{\|{}^tX\|_2}_{=\|X\|_2 \leq 1} \leq \sum_{i=1}^n \lambda_i$$

Donc $\mathrm{Tr}(XM) \leq \sum_{i=1}^n \lambda_i \leq \sup_{U \in O_n(\mathbb{R})} \mathrm{Tr}(UM)$, ainsi $X \in \mathrm{Conv}(O_n(\mathbb{R}))$

Donc $B \subset \mathrm{Conv}(O_n(\mathbb{R}))$. D'où l'égalité $B = \mathrm{Conv}(O_n(\mathbb{R}))$.

Preuve du corollaire : Tout d'abord, d'après ce qu'il précède, $B = \mathrm{Conv}(O_n(\mathbb{R}))$. Donc si l'on prend M dans B, alors il existe d de $\mathbb{N}^*$, $M_1, \dots M_d \in O_n(\mathbb{R})$ et des scalaires $\lambda_1, \dots \lambda_d$ dans $\mathbb{R}_+^*$ tels que $\sum_{i=1}^d \lambda_i = 1$ et $M = \sum_{i=1}^d \lambda_i M_i$. Ainsi, si M est extrémal dans $B = \mathrm{Conv}(O_n(\mathbb{R}))$ alors $d = 1$ et $M = M_1 \in O_n(\mathbb{R})$.

Montrons maintenant que tout élément de $O_n(\mathbb{R})$ est extrémal dans B. Soit $O \in O_n(\mathbb{R})$, supposons que $O = \frac{1}{2}(U+V)$ où $U, V \in B$. Alors pour $x \in \mathbb{R}^n$, on a $\|Ox\|_2^2 = \frac{1}{4}\|Ux + Vx\|_2^2$, donc d'après Cauchy-Schwarz,

$$4\|Ox\|_2^2 = \|Ux\|_2^2 + 2\langle Ux, Vx\rangle + \|Vx\|_2^2 \leq \|Ux\|_2^2 + 2\|Ux\|_2\|Vx\|_2 + \|Vx\|_2^2$$
$$\leq (\|U\|_2\|x\|_2)^2 + 2\|U\|_2\|x\|_2\|V\|_2\|x\|_2 + (\|V\|_2\|x\|_2)^2$$

Donc, puisque $U, V \in B$, on obtient $4\|Ox\|_2^2 \leq 4\|x\|_2^2$. Or $O \in O_n(\mathbb{R})$ donc $\|Ox\|_2^2 = \|x\|_2^2$, donc toutes les inégalités étaient des égalités, en particulier dans Cauchy Schwarz : $\langle Ux, Vx\rangle \overset{①}{=} \|Ux\|_2\|Vx\|_2$, et dans l'inégalité suivante : $\|Vx\|_2 = \|x\|_2 = \|Ux\|_2$ ②. Alors ① indique qu'il existe $\lambda \geq 0$ tel que $Vx = \lambda Ux$, et ② implique $|\lambda| = 1$, donc $Ux = Vx = Ox$, et ce pour tout $x \in \mathbb{R}^n$, donc $O = U = V$ et donc O est extrémal, d'où le résultat.

(cf Gourdon)

Avant de commencer les remarques, quelques rappels sur Carathéodory :

Théorème : Dans un espace affine de dimension n, l'enveloppe convexe d'un sous ensemble A est l'ensemble des barycentres à coefficients positifs ou nuls de familles de $n+1$ points de A. (Preuve dans le Gourdon)

Corollaire : Soit A un compact de E, alors $\mathrm{Conv}(A)$ est compacte.

Démo : On note $\Delta = \{(\lambda_0, \dots, \lambda_n) \in \mathbb{R}^{n+1} \mid \sum_{i=0}^{n} \lambda_i = 1$ et $\forall i, \lambda_i \geq 0\}$. Alors Δ est fermé dans le compact $[0,1]^{n+1}$ donc compact. On considère

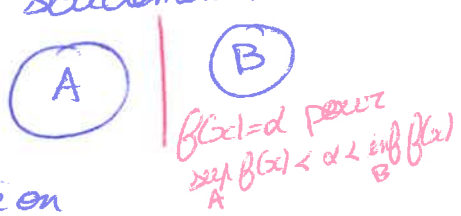
$$\varphi : \Delta \times A^{n+1} \longrightarrow E$$
$$((\lambda_0, \dots, \lambda_n), (x_0, \dots, x_n)) \longmapsto \sum_{i=0}^{n} \lambda_i x_i$$

D'après le théorème, $\mathrm{Conv}(A) = \varphi(\Delta \times A^{n+1})$ or $\Delta \times A^{n+1}$ est compact comme produit de compacts, et φ est continue, donc $\mathrm{Conv}(A)$ est compact.

⚠ Le théorème ne dit pas qu'il existe $x_0, \dots, x_{n+1}$ tels que $\mathrm{Conv}(A) = \mathrm{Conv}(x_0, \dots, x_{n+1})$!!!

Pour des rappels sur les théorèmes de projection sur un convexe fermé, ou de décomposition polaire, cf les devs concernés !

Remarques : ✱ L'idée du théorème de Hahn Banach, c'est de séparer des convexes disjoints (sous certaines hypothèses) par un hyperplan ; on peut généraliser le lemme 1 en prenant 2 convexes non vides disjoints A et B, on suppose A fermé et B compact, alors il existe un hyperplan qui sépare strictement A et B ie il existe $f \in H'$ telle que $\sup_{x \in A} f(x) < \inf_{x \in B} f(x)$



On peut le faire de façon générale sur un EVN E réel avec les mêmes hypothèses sur les convexes (et si on suppose seulement A ouvert et rien sur B on peut séparer au sens large : $\exists f \in E', \exists c \in \mathbb{R}$ $\sup f(x) \leq c \leq \inf f(x)$)