Recasages possibles : 159, 161, 181.
Référence : Objectif agrégation, Beck, Malick, Peyré (p. 97) - Analyse pour l'agrégation, Queffélec, Zuily (p. 205)

Développement

Lemme 1 Soient E un espace euclidien, A une partie non vide de E et $x \in E$. Alors $x \in \overline{\mathrm{Conv}(A)}$ si et seulement si $\forall \varphi \in E^*$, $\varphi(x) \leqslant \sup_{y \in A} \varphi(y)$.

Théorème 2 On munit $\mathcal{M}_n(\mathbb{R})$ de la norme $\|\cdot\|$ subordonnée à la norme euclidienne $\|\cdot\|_2$ de $\mathbb{R}^n$. Alors, $\mathrm{Conv}\big(\mathcal{O}_n(\mathbb{R})\big)$ est la boule unité de $\mathcal{M}_n(\mathbb{R})$.

- *Preuve du Lemme 1* : On note $C = \overline{\mathrm{Conv}(A)}$; c'est un convexe fermé non vide de E. On commence par montrer que si $\varphi \in E^*$, alors $\sup_A \varphi = \sup_C \varphi$. Fixons donc $\varphi \in E^*$. Comme $A \subset C$, on a l'inégalité évidente $\sup_A \varphi \leqslant \sup_C \varphi$. Montrons l'inégalité réciproque. Soit $y \in \mathrm{Conv}(A)$. Par définition, il existe $\lambda_0, \ldots, \lambda_r \geqslant 0$, $a_1, \ldots, a_r \in A$ tels que
$$\sum_{i=1}^r \lambda_i = 1 \quad \text{et} \quad y = \sum_{i=1}^r \lambda_i a_i.$$
Alors,
$$\begin{aligned} \varphi(y) &= \varphi\left(\sum_{i=1}^r \lambda_i a_i\right) = \sum_{i=1}^r \lambda_i \varphi(a_i) \quad \text{par linéarité de } \varphi \\ &\leqslant \sup_A \varphi \sum_{i=1}^r \lambda_i \quad \text{car } \forall i \in [\![1, r]\!],\ a_i \in A \\ &\leqslant \sup_A \varphi \end{aligned}$$
Ainsi, en passant au sup sur $y \in \mathrm{Conv}(A)$, on obtient $\sup_{\mathrm{Conv}(A)} \varphi \leqslant \sup_A \varphi$. Enfin, par continuité de φ (une application linéaire est toujours continue en dimension finie), on a
$$\sup_C \varphi = \sup_{\mathrm{Conv}(A)} \varphi \leqslant \sup_A \varphi.$$
Il suffit donc de montrer l'équivalence $x \in C \iff \forall \varphi \in E^*,\ \varphi(x) \leqslant \sup_C \varphi(y)$. L'implication directe est évidente, donc il nous reste à montrer la réciproque. Raisonnons par contraposition : on considère $x \in E \setminus C$ et on va exhiber une forme linéaire φ sur E telle que $\varphi(x) > \sup_C \varphi$. Comme C est un convexe fermé non vide, on dispose du projeté de x sur C, que l'on note a : c'est l'unique vecteur tel que $\|x - a\| = d(x, C)$. Rappelons qu'il est caractérisé parmi les éléments de C par la propriété $\langle x - a \mid y - a\rangle \leqslant 0$ pour tout $y \in C$. On considère alors la forme linéaire $\varphi : y \mapsto \langle x - a \mid y\rangle$. On remarque alors que
$$\varphi(x-a) = \langle x - a \mid x - a\rangle = \|x-a\|^2 > 0$$
car $x \notin C$ donc $x \neq a$. Ainsi, par linéarité de φ, on a $\varphi(x) > \varphi(a)$. Or, si $y \in C$, la caractérisation du projeté sur un convexe fermé nous donne
$$\varphi(y-a) = \langle x - a \mid y - a\rangle \leqslant 0.$$
À nouveau par linéarité de φ, on a directement $\varphi(y) \leqslant \varphi(a)$. Ainsi, $\varphi(a) = \sup_C \varphi$, ce qui conclut la preuve de l'équivalence souhaitée, ainsi que la preuve du **Lemme 1**.

- *Preuve du Théorème 2* : Notons B la boule unité de $\mathcal{M}_n(\mathbb{R})$ pour la norme subordonnée à la norme euclidienne de $\mathbb{R}^n$. On a $\mathcal{O}_n(\mathbb{R}) \subset B$ car si $M \in \mathcal{O}_n(\mathbb{R})$, alors
$$\|M\| = \sup_{\|X\|_2 = 1} \|MX\|_2 = \sup_{\|X\|_2=1} \|X\|_2 = 1.$$
De plus, d'après l'inégalité triangulaire, B est convexe. Ainsi, $\mathrm{Conv}\big(\mathcal{O}_n(\mathbb{R})\big)$ étant l'intersection des convexes $\mathcal{M}_n(\mathbb{R})$ contenant $\mathcal{O}_n(\mathbb{R})$, on a $\mathrm{Conv}\big(\mathcal{O}_n(\mathbb{R})\big) \subset B$.

Montrons l'inclusion réciproque. La partie $\mathcal{O}_n(\mathbb{R})$ de $\mathcal{M}_n(\mathbb{R})$ est compacte, donc d'après un corollaire du théorème de Carathéodory, $\mathrm{Conv}\big(\mathcal{O}_n(\mathbb{R})\big)$ est compacte. En particulier,
$$\overline{\mathrm{Conv}\big(\mathcal{O}_n(\mathbb{R})\big)} = \mathrm{Conv}\big(\mathcal{O}_n(\mathbb{R})\big).$$
Soit $M \in B$, c'est-à-dire $M \in \mathcal{M}_n(\mathbb{R})$ telle que $\|M\| \leqslant 1$. D'après le **Lemme 1**, pour montrer que $M \in \mathrm{Conv}\big(\mathcal{O}_n(\mathbb{R})\big)$, il suffit de montrer que pour tout $\varphi \in \big(\mathcal{M}_n(\mathbb{R})\big)^*$, $\varphi(M) \leqslant \sup_{O \in \mathcal{O}_n(\mathbb{R})} \varphi(O)$. Or, la forme bilinéaire
$$\begin{array}{rcl} \mathcal{M}_n(\mathbb{R}) \times \mathcal{M}_n(\mathbb{R}) & \longrightarrow & \mathbb{R} \\ (A, B) & \longmapsto & \mathrm{Tr}(A^\top B) \end{array}$$
est non dégénérée (c'est le produit scalaire canonique de $\mathcal{M}_n(\mathbb{R})$), donc induit un isomorphisme entre $\mathcal{M}_n(\mathbb{R})$ et $\big(\mathcal{M}_n(\mathbb{R})\big)^*$ via $A \mapsto \big(B \mapsto \mathrm{Tr}(A^\top B)\big)$. Ainsi,

il suffit de montrer que pour tout $A \in \mathcal{M}_n(\mathbb{R})$, on a

$$\mathrm{Tr}(A^\top M) \leqslant \sup_{O \in \mathcal{O}_n(\mathbb{R})} \mathrm{Tr}(A^\top O).$$

Fixonx $A \in \mathcal{M}_n(\mathbb{R})$ et soit $\Omega S = A$ une décomposition polaire de A, c'est-à-dire $\Omega \in \mathcal{O}_n(\mathbb{R})$ et $S \in \mathcal{S}_n^+(\mathbb{R})$. D'après le théorème spectral, si $\lambda_1, \ldots, \lambda_n \geqslant 0$ sont les valeurs propres de S, il existe une base orthonormée $(e_1, \ldots, e_n)$ de $\mathbb{R}^n$ telle que $Se_i = \lambda_i e_i$ pour tout $i \in [\![1, n]\!]$. On a alors

$$\begin{aligned}
\mathrm{Tr}(A^\top M) &= \sum_{i=1}^n \left\langle A^\top M e_i \mid e_i \right\rangle = \sum_{i=1}^n \langle M e_i \mid A e_i \rangle \\
&\leqslant \sum_{i=1}^n \|Me_i\| \, \|Ae_i\| \quad \text{d'après Cauchy-Schwarz} \\
&\leqslant \sum_{i=1}^n \|\Omega S e_i\| = \sum_{i=1}^n \|Se_i\| \quad \text{car } M \in B \text{ et } \Omega \in \mathcal{O}_n(\mathbb{R}) \\
&\leqslant \sum_{i=1}^n \lambda_i = \mathrm{Tr}(S) = \mathrm{Tr}(\Omega^\top A) \\
&\leqslant \mathrm{Tr}(A^\top \Omega) \quad \text{car } \forall X \in \mathcal{M}_n(\mathbb{R}),\ \mathrm{Tr}(X^\top) = \mathrm{Tr}(X) \\
&\leqslant \sup_{O \in O_n(\mathbb{R})} \mathrm{Tr}(A^\top O)
\end{aligned}$$

Ceci montre bien que $M \in \mathrm{Conv}\big(\mathcal{O}_n(\mathbb{R})\big)$ d'après les raisonnements précédents, donc $B \subset \mathrm{Conv}\big(\mathcal{O}_n(\mathbb{R})\big)$ et on a finalement montré l'égalité $B = \mathrm{Conv}\big(\mathcal{O}_n(\mathbb{R})\big)$.

Commentaires et prolongements :

- On a utilisé le résultat classique (vrai et à connaître jusque dans les espaces de Hilbert) de projection sur un convexe fermé non vide. Remontrons le ici dans le cadre d'un espace euclidien pour coller au présent développement. Soit donc C un convexe fermé non vide d'un espace euclidien E, et $x \in E$. Notons

$$d = d(x, C) = \inf_{y \in C} \|x - y\|.$$

Notons que si $y \in C$ et $y \notin \overline{B} := \overline{B(x, 2d)}$, alors $\|x - y\| \geqslant 2d$. Ainsi, on a

$$\inf_{y \in C} \|x - y\| = \inf_{y \in C \cap \overline{B}} \|x - y\|.$$

Or, $C \cap \overline{B}$ est une partie compacte en tant que partie fermée et bornée dans un espace de dimension finie (ou plus simplement fermée dans le compact $\overline{B}$). Ainsi, l'application continue $y \mapsto \|x - y\|$ admet un minimum sur $C \cap \overline{B}$ atteint en un certain $a \in C$. Ceci justifie l'existence du projeté de x sur C.

Montrons alors l'unicité d'un tel point. On suppose que $a_1, a_2 \in C$ vérifient $\|x - a_1\| = \|x - a_2\| = d$. On va utiliser l'identité du parallélogramme, qui caractérise les normes euclidiennes :

$$\forall u, v \in E,\ \|u + v\|^2 + \|u - v\|^2 = 2\|u\|^2 + 2\|v\|^2.$$

On applique cette identité à $u = x - a_1$ et $v = x - a_2$, pour lesquels on a $u + v = 2\big(x - \frac{a_1 + a_2}{2}\big)$ et $u - v = a_2 - a_1$. Comme C est convexe, $\frac{a_1 + a_2}{2} \in C$ et donc on obtient

$$\begin{aligned}
\|a_2 - a_1\|^2 &= 2\|x - a_1\|^2 + 2\|x - a_2\|^2 - 4\left\|x - \frac{a_1 + a_2}{2}\right\|^2 \\
&= 4d^2 - 4\left\|x - \tfrac{a_1 + a_2}{2}\right\|^2 \leqslant 4d^2 - 4d^2 = 0
\end{aligned}$$

Ainsi, $a_1 = a_2$ et on a bien montré l'unicité d'un tel projeté.

Montrons enfin la caractérisation du projeté par le produit scalaire négatif, qui signifie que si $y \in C$, l'angle entre les vecteurs $\overrightarrow{ax}$ et $\overrightarrow{ay}$ est obtus. Supposons tout d'abord qu'un point $a \in C$ vérifie $\langle x - a \mid y - a \rangle \leqslant 0$ pour tout $y \in C$. Alors, si $y \in C$, on a

$$\begin{aligned}
\|x - y\|^2 &= \|x - a + a - y\|^2 \\
&= \|x - a\|^2 + 2\langle x - a \mid a - y \rangle + \|a - y\|^2 \\
&\geqslant \|x - a\|^2 - 2\langle x - a \mid y - a \rangle \geqslant \|x - a\|^2
\end{aligned}$$

Ainsi, a vérifie $\|x - a\| = d(x, C)$, c'est donc bien le projeté de x sur C. Réciproquement, supposons que a est le projeté de x sur C. Fixons $y \in C$. Par convexité de C, si $\lambda \in]0, 1[$, $(1 - \lambda)a + \lambda y = a + \lambda(y - a) \in C$. Ainsi,

$$\begin{aligned}
\|x - a\|^2 &\leqslant \|x - a - \lambda(y - a)\|^2 \\
&\leqslant \|x - a\|^2 - 2\lambda\langle x - a \mid y - a \rangle + \lambda^2\|y - a\|^2
\end{aligned}$$

Alors, en simplifiant les $\|x - a\|^2$, puis en divisant par λ, on obtient

$$2\langle x - a \mid y - a \rangle \leqslant \lambda\|y - a\|^2.$$

Enfin, en faisant tendre λ vers 0, on obtient $\langle x - a \mid y - a \rangle \leqslant 0$, ce qui conclut.

- On a également utilisé l'important théorème de Carathéodory pour justifier que $\mathrm{Conv}\big(\mathcal{O}_n(\mathbb{R})\big)$ est une partie compacte donc fermée dans E. On peut d'ailleurs se demander pourquoi montrer la compacité avec un théorème difficile alors qu'on ne se sert que du caractère fermé. Le problème est que l'enveloppe convexe d'un fermé n'est pas toujours fermée. Pour s'en convaincre, on peut considérer la réunion d'une droite et d'un point en dehors de la droite dans le plan euclidien. Énonçons et montrons ce théorème :

3 Théorème : (Carathéodory) *Soit $A \subset E$. Alors, tout élément de* $\mathrm{Conv}(A)$ *s'écrit comme combinaison convexe d'au plus* $\dim(E) + 1$ *vecteurs de E.*

Pour montrer ce théorème, considérons $x \in \mathrm{Conv}(A)$. Par définition, il existe $a_1, \dots, a_q \in A$ et $\lambda_1, \dots, \lambda_q \geqslant 0$ tels que $\lambda_1 + \cdots + \lambda_q = 1$ et $x = \lambda_1 a_1 + \cdots + \lambda_q a_q$. Si $q \leqslant \dim(E)+1$, il n'y a rien à montrer. Supposons donc $q > \dim(E)+1$. Alors, la famille $(a_1 - a_2, a_1 - a_3, \dots, a_1 - a_q)$ contient $k - 1 > \dim(E)$ éléments, donc est liée. Autrement dit il existe $\mu_2, \dots, \mu_q \in \mathbb{R}$ non tous nuls tels que

$$\mu_2(a_1 - a_2) + \mu_3(a_1 - a_3) + \cdots + \mu_q(a_1 - a_q) = 0,$$

et donc

$$\left(\sum_{i=2}^{q} \mu_i\right) a_1 - \sum_{i=2}^{n} \mu_i a_i = 0.$$

Posons alors $\nu_1 = \sum_{i=2}^{q} \mu_i$ et pour $i \in [\![2, q]\!]$, $\nu_i = -\mu_i$. On obtient alors

$$\sum_{i=1}^{q} \nu_i = 0 \quad \text{et} \quad \sum_{i=1}^{q} \nu_i a_i = 0.$$

Les ν_i étant non tous nuls, mais de somme nulle, il existe $i \in [\![1, q]\!]$ tel que $\nu_i > 0$. Considérons alors $\alpha = \min_{\nu_i > 0} \left\{\frac{\lambda_i}{\nu_i}\right\}$ de sorte que pour tout $i \in [\![1, q]\!]$, $\lambda_i - \alpha\nu_i \geqslant 0$. En effet, si $\nu_i \leqslant 0$, comme $\alpha \geqslant 0$, l'inégalité est triviale et si $\nu_i > 0$, par construction de α, on a $\frac{\lambda_i}{\nu_i} \geqslant \alpha$, d'où $\lambda_i - \alpha\nu_i \geqslant 0$. De plus,

$$\sum_{i=1}^{q} (\lambda_i - \alpha\nu_i) = \sum_{i=1}^{q} \lambda_i - \alpha \sum_{i=1}^{q} \nu_i = 1 - 0 = 1,$$

$$\text{et} \qquad \sum_{i=1}^{q} (\lambda_i - \alpha\nu_i) a_i = x - \alpha \sum_{i=1}^{q} \nu_i a_i = x.$$

Or α étant un minimum, il est atteint et donc il existe $i \in [\![1, q]\!]$ tel que $\lambda_i - \alpha\nu_i = 0$. Ainsi, on a écrit x comme combinaison convexe de $q - 1$ points, ce qui conclut en itérant le processus jusqu'à $q = \dim(E) + 1$.

Ce théorème permet de montrer que l'enveloppe convexe d'un compact est compacte (c'est ce dont on s'est servis dans ce développement). En effet si A est compact, et si on pose $n = \dim(E)$ et

$$\Delta = \left\{ \big(\lambda_1, \dots, \lambda_{n+1}\big) \in \mathbb{R}_+^{n+1} \mid \sum_{i=1}^{n+1} \lambda_i = 1 \right\},$$

alors $\Delta \times A^{n+1}$ est un compact en tant que produit de compacts. De plus, l'application

$$\Phi : \left| \begin{array}{ccc} \Delta \times A^{n+1} & \longrightarrow & \mathrm{Conv}(A) \\ \Big(\big(\lambda_1, \dots, \lambda_{n+1}\big), \big(a_1, \dots, a_{n+1}\big)\Big) & \longmapsto & \displaystyle\sum_{i=1}^{n+1} \lambda_i a_i \end{array} \right.$$

est continue et elle est surjective d'après le théorème de Carathéorody, donc $\mathrm{Conv}(A) = \Phi(\Delta \times A^{n+1})$ est compact.

Remarque : Pour le théorème de Carathéodory (dans le cadre vectoriel) et ce corollaire, on trouve ces preuves dans Analyse, Gourdon (p. 54,55).

- Enfin, on peut faire le lien entre ce calcul d'enveloppe convexe et le théorème de Krein-Milman en montrant que l'ensemble $\mathrm{Extr}(B)$ des points extrémaux de la boule unité B considérée est égal à $\mathcal{O}_n(\mathbb{R})$. Procédons pour cela par double inclusion :

$(\subseteq)$: Si $\Omega \in \mathcal{O}_n(\mathbb{R})$ s'écrit $\Omega = (1-t)U + tV$ avec $t \in]0, 1[$, $U, V \in B$, alors pour tout $X \in \mathbb{R}^n$ de norme 1, on a par inégalité triangulaire

$$1 = \|\Omega X\| \leqslant (1-t)\|UX\| + t\|VX\| \leqslant (1-t)\|X\| + t\|X\| = 1.$$

Ainsi, l'inégalité triangulaire utilisée est une égalité, ce qui signifie (cas d'égalité pour l'inégalité triangulaire d'une norme euclidienne) que UX et VX sont positivement liés. Autrement dit, quitte à échanger U et V, il existe $\lambda \geqslant 0$ tel que $UX = \lambda VX$. Or, en considérant la norme, on obtient $1 = |\lambda| = \lambda$, d'où $UX = VX$. Ceci étant valable pour tout $X \in \mathbb{R}^n$ de norme 1, puis pour tout $X \in \mathbb{R}^n$ par linéarité, on obtient $U = V$. En reportant à l'égalité avec Ω, on voit que $\Omega = U = V$, ce qui montre bien que $\mathcal{O}_n(\mathbb{R}) \subseteq \mathrm{Extr}(B)$.

($\supseteq$) : Réciproquement, soit $M \in B \setminus \mathcal{O}_n(\mathbb{R})$. On va montrer que $M \notin \mathrm{Extr}(B)$. Soit $M = \Omega S$ une décomposition polaire de M, avec donc $\Omega \in \mathcal{O}_n(\mathbb{R})$ et $S \in \mathcal{S}_n^+(\mathbb{R})$ différente de I_n. D'après le théorème spectral, il existe $P \in \mathcal{O}_n(\mathbb{R})$ telle que $S = PDP^{-1}$, où D est la matrice diagonale

$$D = \begin{pmatrix} \lambda_1 & & (0) \\ & \ddots & \\ (0) & & \lambda_n \end{pmatrix},$$

avec $0 \leqslant \lambda_1 \leqslant \cdots \leqslant \lambda_n$. Comme S est symétrique et Ω orthogonale, on a $\lambda_n = \|S\| = \|\Omega S\| = \|M\| \leqslant 1$. Par ailleurs, $S \neq I_n$ donc $\lambda_1 < \lambda_n$, d'où $\lambda_1 < 1$. Par conséquent, λ_1 n'est pas un point extrémal de $[-1, 1]$, et il existe donc $t \in]0, 1[$ tel que $\lambda_1 = (1-t) \times (-1) + t \times 1$. Posons alors

$$D_1 = \begin{pmatrix} 1 & & & (0) \\ & \lambda_2 & & \\ & & \ddots & \\ (0) & & & \lambda_n \end{pmatrix} \quad \text{et} \quad D_{-1} = \begin{pmatrix} -1 & & & (0) \\ & \lambda_2 & & \\ & & \ddots & \\ (0) & & & \lambda_n \end{pmatrix},$$

de sorte que $D = (1-t)D_{-1} + tD_1$, avec $D \neq D_1, D_{-1}$. Posons ensuite $S_1 = PD_1P^{-1}$ et $S_{-1} = PD_{-1}P^{-1}$. On a alors

$$M = \Omega PDP^{-1} = \Omega\big(P(1-t)D_{-1}P^{-1} + PtD_1P^{-1}\big) = (1-t)\Omega S_{-1} + t\Omega S_1.$$

Or, $S \neq S_1, S_{-1}$ et donc comme Ω est inversible, $M \neq \Omega S_1, \Omega S_{-1}$. De plus, $\|\Omega S_1\| = \|S_1\| = \rho(S_1) = \rho(D_1) = 1 = \rho(D_{-1}) = \|\Omega S_{-1}\|$. Ainsi, on a bien écrit M comme combinaison convexe non triviale de deux éléments de B distincts de M, donc $M \notin \mathrm{Extr}(B)$, ce qui conclut. Comme B est un convexe compact de $\mathcal{M}_n(\mathbb{R})$, le théorème de Krein-Milman nous dit que $B = \mathrm{Conv}\big(\mathrm{Extr}(B)\big) = \mathrm{Conv}\big(\mathcal{O}_n(\mathbb{R})\big)$ et on retrouve bien le résultat.