Questions & réponses

Réponses

R987. Posé dans RMS 130 1.
Déterminer l'enveloppe convexe de $\mathrm{SO}(n,\mathbb{R})$.

(Clément de Seguins Pazzis)

Réponse d'après les indications de Geoffrey Delpluque et Mohamed Nassiri

On distingue trois cas, selon que $n=1$, $n=2$ ou $n \geqslant 3$.
Si $n=1$, $\mathrm{SO}_n(\mathbb{R})$ est réduit à I_1, donc son enveloppe convexe est $\{I_1\}$.
Supposons $n=2$. Alors $\mathrm{SO}_2(\mathbb{R})$ est le cercle unité de l'espace vectoriel de matrices

$$\mathcal{P} := \left\{ M_{a,b} := \begin{pmatrix} a & -b \\ b & a \end{pmatrix} \mid (a,b) \in \mathbb{R}^2 \right\}$$

pour la norme euclidienne associant à $M_{a,b}$ le scalaire $\sqrt{a^2+b^2}$, et il s'agit donc du disque unité fermé de ce plan euclidien.
Dans toute la suite, on supposera $n \geqslant 3$. Il est classique que l'enveloppe convexe de $\mathrm{O}_n(\mathbb{R})$ est constituée de la boule unité fermée de $\mathcal{M}_n(\mathbb{R})$ pour la norme d'opérateur associée à la norme euclidienne standard sur $\mathbb{R}^n$. Le cas de $\mathrm{SO}_n(\mathbb{R})$ est nettement plus compliqué.
Pour simplifier les choses, on notera $G := \mathrm{SO}_n(\mathbb{R})$ tout au long de notre réponse. Classiquement, G est compact, et comme $\mathcal{M}_n(\mathbb{R})$ est de dimension finie l'enveloppe convexe $\mathrm{Conv}(G)$ est également compacte. La multiplication à gauche ou à droite par une matrice fixée étant une fonction affine, elle conserve le barycentrage, si bien que $\mathrm{Conv}(G)$ est stable par multiplication à gauche ou à droite par les élements de G. Ainsi, $\mathrm{Conv}(G)$ est une réunion d'orbites pour l'action à gauche de G^2 sur $\mathcal{M}_n(\mathbb{R})$, dite action naturelle, définie comme suit :

$$(O_1,O_2).M := O_1 M O_2^{-1}.$$

Il est donc naturel de chercher, dans un premier temps, à analyser les orbites pour cette action. Dans un deuxième temps, nous résoudrons le problème posé grâce à une variante d'une inégalité classique sur les traces (Proposition 3).

Orbites sous l'action naturelle de G^2

Les orbites de $\mathcal{M}_n(\mathbb{R})$ sous l'action naturelle de G^2 sont des sous-ensembles des orbites pour l'action à gauche de $\mathrm{O}_n(\mathbb{R})^2$ sur $\mathcal{M}_n(\mathbb{R})$ définie par $(O_1,O_2).M := O_1 M O_2^{-1}$. Ces dernières sont bien connues, elles sont repérées par les *valeurs singulières*. Étant donné $M \in \mathcal{M}_n(\mathbb{R})$, il est en effet classique qu'il existe une unique suite décroissante $s_1(M) \geqslant \cdots \geqslant s_n(M) \geqslant 0$ de réels positifs telle qu'il existe $(O_1, O_2) \in \mathrm{O}_n(\mathbb{R})^2$ vérifiant

$$M = O_1 \,\mathrm{Diag}(s_1(M), \ldots, s_n(M))\, O_2.$$

Les $s_k(M)^2$ sont les valeurs propres, comptées avec ordre de multiplicité, de la matrice $M^T M$. En particulier, $s_1(M)$ est la norme d'opérateur de M, et $s_n(M)$ est nul si et seulement si M est singulière. Il sera utile, en vue de la suite, de remarquer que si l'on dispose d'une réduction $M = O_1 \,\mathrm{Diag}(\alpha_1, \ldots, \alpha_n)\, O_2$ où $\alpha_1, \ldots, \alpha_n$ sont seulement supposés réels, et O_1 et O_2 sont dans $\mathrm{O}_n(\mathbb{R})$, alors $|\alpha_1|, \ldots, |\alpha_n|$ sont, à l'ordre près, les valeurs singulières de M (comptées avec multiplicité).
Notons enfin que, dans une décomposition précédente, $D := \mathrm{Diag}(1, \ldots, 1, -1) \in \mathrm{GL}_n(\mathbb{R})$ vérifie $M = (O_1 D)\mathrm{Diag}(s_1(M), \ldots, s_n(M))(D O_2)$, et $O_1 D$ et $D O_2$ sont orthogonales. On peut donc toujours choisir pour O_1 une matrice de $\mathrm{SO}_n(\mathbb{R})$ mais il est alors possible que $\det O_2 = -1$. Dans ce cas

$$M = O_1 \mathrm{Diag}(s_1(M), \ldots, s_n(M)) O_2 = O_1 \mathrm{Diag}(s_1(M), \ldots, s_{n-1}(M), -s_n(M))(D O_2),$$

donc M est dans la G^2-orbite de $\mathrm{Diag}(s_1(M), \ldots, s_n(M))$ ou dans celle de la matrice $\mathrm{Diag}(s_1(M), \ldots, s_{n-1}(M), -s_n(M))$. Par comparaison des déterminants, seule la matrice $\mathrm{Diag}(s_1(M), \ldots, s_{n-1}(M), \mathrm{sgn}(\det M) s_n(M))$ est possible si M est singulière, sinon les deux matrices diagonales indiquées sont égales, et M est dans la G^2-orbite de la première.
Concluons :

Proposition 1. *Soit $M \in \mathcal{M}_n(\mathbb{R})$, dont on note ε le signe du déterminant, défini comme 0 si M est singulière. On pose $\widetilde{s_n}(M) := \varepsilon\, s_n(M)$.*
*Alors M est conjuguée à $\mathrm{Diag}(s_1(M), \ldots, s_{n-1}(M), \widetilde{s_n}(M))$, dite **forme réduite** de M, sous l'action naturelle du groupe G^2.*
En conséquence, deux éléments de $\mathcal{M}_n(\mathbb{R})$ sont conjugués sous l'action naturelle de G^2 si et seulement s'ils ont mêmes valeurs singulières et des déterminants de même signe.

Un résultat d'optimisation

Le résultat suivant (proposition 2) est classique, nous allons en examiner la version restreinte au groupe spécial orthogonal (proposition 3).

Proposition 2. *Soit A et B dans $\mathcal{M}_n(\mathbb{R})$. Alors*

$$\mathrm{tr}(AB) \leqslant \sum_{k=1}^{n} s_k(A)\, s_k(B).$$

Proposition 3. *Soit A et B dans $\mathcal{M}_n(\mathbb{R})$. Alors*

$$\max_{g \in G^2} \mathrm{tr}((g.A)B) = \widetilde{s_n}(A)\, \widetilde{s_n}(B) + \sum_{k=1}^{n-1} s_k(A)\, s_k(B).$$

Démonstration de la proposition 3. Commençons par constater que le membre de droite est bien atteint par la fonction $g \in G^2 \mapsto \mathrm{tr}((g.A)B)$. Pour cela, décomposons

$$A = O_1 \,\mathrm{Diag}\big(s_1(A), \dots, s_{n-1}(A), \widetilde{s_n}(A)\big)\, O_2$$

et

$$B = U_1 \,\mathrm{Diag}\big(s_1(B), \dots, s_{n-1}(B), \widetilde{s_n}(B)\big)\, U_2,$$

où O_1, O_2, U_1, U_2 sont dans G. On voit alors que

$$\mathrm{tr}\Big(\big(\underbrace{U_2^{-1}O_1^{-1}}_{\in G}\big)A\big(\underbrace{U_1O_2}_{\in G}\big)^{-1}B\Big) = \widetilde{s_n}(A)\,\widetilde{s_n}(B) + \sum_{k=1}^{n-1} s_k(A)\, s_k(B).$$

Pour conclure, il suffit donc de montrer que $\mathrm{tr}(OAUB) \leqslant \widetilde{s_n}(A)\,\widetilde{s_n}(B) + \sum\limits_{k=1}^{n-1} s_k(A)\, s_k(B)$ pour tout $(O, U) \in G^2$, ce qui est bien sûr la partie la plus délicate de la démonstration.
Nous commençons par le cas particulier où B est à valeurs singulières distinctes. On peut en outre supposer B sous forme réduite : écrivons en effet $B = U_1 B' U_2$ pour des matrices U_1 et U_2 dans G, où B' désigne la forme réduite de B. Si l'inégalité voulue est établie pour A et B' (qui conserve les propriétés supposées sur B et est en plus sous forme réduite), elle l'est pour A et B car

$$\forall (O, U) \in G^2,\ \mathrm{tr}(OAUB) = \mathrm{tr}(OAUU_1B'U_2) = \mathrm{tr}((\underbrace{U_2O}_{\in G})A(\underbrace{UU_1}_{\in G})B').$$

Nous supposerons désormais B sous forme réduite. Par compacité de G^2 et continuité de la trace, la fonction $\varphi : (O_1, O_2) \in G^2 \mapsto \mathrm{tr}(O_1AO_2B)$ possède un maximum. Quitte à remplacer A par un élément de sa G^2-orbite, on peut supposer que φ est maximale en (I_n, I_n). Comme l'espace vectoriel tangent à $\mathrm{SO}_n(\mathbb{R})$ en I_n est $\mathcal{A}_n(\mathbb{R})$, il vient $\mathrm{tr}(KAB) = 0$ et $\mathrm{tr}(AKB) = 0$, dernière égalité qui se récrit $\mathrm{tr}(BAK) = 0$, pour tout $K \in \mathcal{A}_n(\mathbb{R})$. Il s'ensuit que AB et BA sont symétriques. Ainsi, puisque B est elle-même symétrique, $AB = (AB)^T = BA^T$ et $BA = (BA)^T = A^TB$, puis $AB^2 = BA^TB = B^2A$. Or B^2 est diagonale à coefficients diagonaux deux à deux distincts (puisque B est sous forme réduite et possède n valeurs singulières distinctes). Ainsi A est diagonale.
Écrivons maintenant $A = \mathrm{Diag}(\alpha_1, \dots, \alpha_n)$. Ainsi,

$$\max \varphi = \varphi(I_n, I_n) = \mathrm{tr}(AB) = \alpha_n\, \widetilde{s_n}(B) + \sum_{k=1}^{n-1} \alpha_k\, s_k(B).$$

Notons que $s_1(B) \geqslant \cdots \geqslant s_{n-1}(B) \geqslant \widetilde{s_n}(B)$. Ainsi, l'inégalité de réordonnement (page 262 de [1]) donne $\mathrm{tr}(AB) \leqslant \alpha_{\sigma(n)}\, \widetilde{s_n}(B) + \sum\limits_{k=1}^{n-1} \alpha_{\sigma(k)}\, s_k(B)$ pour une permutation σ telle que $(\alpha_{\sigma(k)})_{1 \leqslant k \leqslant n}$ soit décroissante. La matrice $A' := \mathrm{Diag}(\alpha_{\sigma(1)}, \dots, \alpha_{\sigma(n)})$ a mêmes valeurs singulières et même déterminant que A, donc elle est dans la G^2-orbite de A. Quitte à remplacer A par A', on peut donc supposer que $\alpha_1 \geqslant \cdots \geqslant \alpha_n$.

Si $\alpha_n \geqslant 0$ alors $s_k(A) = \alpha_k$ pour tout $k \in \{1, \ldots, n\}$; par ailleurs $\det A \geqslant 0$ donc $\widetilde{s_n}(A) = \alpha_n$, et ainsi $\operatorname{tr}(AB) = \widetilde{s_n}(A)\,\widetilde{s_n}(B) + \sum_{k=1}^{n-1} s_k(A)\,s_k(B)$.

Supposons maintenant que $\alpha_n < 0$ et montrons qu'alors $\alpha_{n-1} \geqslant |\alpha_n|$. Notons A'' la matrice obtenue à partir de A par échange des coefficients diagonaux d'indices respectifs $n-1$ et n, puis multiplication de ceux-ci par -1. On note que A'' est dans la G^2-orbite de A puisqu'elle a même déterminant et mêmes valeurs singulières que A, si bien que $\operatorname{tr}(A''B) \leqslant \operatorname{tr}(AB)$. Or

$$\begin{aligned} \operatorname{tr}(A''B) - \operatorname{tr}(AB) &= -\alpha_{n-1}\,\widetilde{s_n}(B) - \alpha_n\,s_{n-1}(B) - \alpha_{n-1}\,s_{n-1}(B) - \alpha_n\,\widetilde{s_n}(B) \\ &= -\big(s_{n-1}(B) + \widetilde{s_n}(B)\big)\,(\alpha_{n-1} + \alpha_n). \end{aligned}$$

Puisque $s_{n-1}(B) + \widetilde{s_n}(B) > 0$, il vient $\alpha_{n-1} \geqslant -\alpha_n = |\alpha_n|$. On en déduit que $\alpha_k = s_k(A)$ pour tout $k \in \{1, \ldots, n-1\}$ et $\alpha_n = \widetilde{s_n}(A)$, d'où $\operatorname{tr}(AB) = \widetilde{s_n}(A)\,\widetilde{s_n}(B) + \sum_{k=1}^{n-1} s_k(A)\,s_k(B)$, ce qui est le résultat recherché. Ceci achève la démonstration dans le cas particulier où B est à valeurs singulières distinctes.

Passons au cas général, que nous déduirons du cas particulier précédent grâce à un argument de densité. Nous réduisons $B = U_1 B' U_2$ avec U_1, U_2 dans G et B' la forme réduite de B. Posons $\Delta := \operatorname{Diag}(n-1, n-2, \ldots, 1, 0)$, et pour $t > 0$, $B_t := B + t\,U_1 \Delta U_2$, si bien que les valeurs singulières de B_t sont les $s_k(B) + (n-k)t$, en ordre strictement décroissant, et $\widetilde{s_n}(B_t) = \widetilde{s_n}(B)$. Fixons $(O, U) \in G^2$. Le cas particulier précédent donne donc, pour tout réel $t > 0$,

$$\operatorname{tr}(OAUB_t) \leqslant \widetilde{s_n}(A)\,\widetilde{s_n}(B_t) + \sum_{k=1}^{n-1} s_k(A)\,s_k(B_t)$$

autrement dit

$$\operatorname{tr}(OAUB) + t\,\operatorname{tr}(OAUU_1\Delta U_2) \leqslant \widetilde{s_n}(A)\,\widetilde{s_n}(B) + \sum_{k=1}^{n-1} s_k(A)\,(s_k(B) + (n-k)t).$$

L'inégalité $\operatorname{tr}(OAUB) \leqslant \widetilde{s_n}(A)\,\widetilde{s_n}(B) + \sum_{k=1}^{n-1} s_k(A)\,s_k(B)$ s'obtient donc en faisant tendre t vers 0. $\square$

Calcul de l'enveloppe convexe de $\mathrm{SO}_n(\mathbb{R})$

Nous allons maintenant utiliser une méthode classique de dualité pour étudier l'enveloppe convexe de $G = \mathrm{SO}_n(\mathbb{R})$. On pose d'abord

$$G^\star := \{(\varphi, \lambda) \in \mathcal{M}_n(\mathbb{R})^\star \times \mathbb{R} : \ \forall M \in G,\ \varphi(M) \leqslant \lambda\}.$$

Comme $\operatorname{Conv}(G)$ est un convexe fermé d'un espace vectoriel de dimension finie, le lemme de Hahn-Banach donne

$$\operatorname{Conv}(G) = \{M \in \mathcal{M}_n(\mathbb{R}) : \ \forall (\varphi, \lambda) \in G^\star,\ \varphi(M) \leqslant \lambda\}.$$

Il est relativement incommode d'utiliser ce point de vue à cause du paramètre λ. Par bonheur, on va observer (ce n'est pas immédiat !) que $\max \varphi(G) > 0$ pour tout $\varphi \in \mathcal{M}_n(\mathbb{R})^\star \setminus \{0\}$ (ce qui est traduit le fait que 0 est intérieur à $\mathrm{Conv}(G)$ dans $\mathcal{M}_n(\mathbb{R})$) : cela va nous permettre de prendre systématiquement $\lambda = 1$.
D'abord, les formes linéaires sur $\mathcal{M}_n(\mathbb{R})^\star$ sont les fonctions de la forme $\varphi_A : M \mapsto \mathrm{tr}(AM)$ pour un $A \in \mathcal{M}_n(\mathbb{R})$. Soit donc $A \in \mathcal{M}_n(\mathbb{R})$. On note que, vu les propriétés de la trace et le fait que G soit un sous-groupe de $\mathrm{GL}_n(\mathbb{R})$,

$$\varphi_A(G) = \{\mathrm{tr}(AO) \mid O \in G\} = \{\mathrm{tr}(AO_2^{-1}O_1) \mid (O_2, O_1) \in G^2\} = \{\mathrm{tr}(g.A) \mid g \in G^2\}$$

La proposition 3 appliquée à $B = I_n$ donne que φ_A a pour maximum $\widetilde{s_n}(A) + \sum_{k=1}^{n-1} s_k(A)$ sur G. On notera que ce maximum est supérieur ou égal à $s_1(A) + s_{n-1}(A) - s_n(A)$, donc à $s_1(A)$, car $n \geqslant 3$, et en particulier il est strictement positif si $A \neq 0$. Le point annoncé plus haut est donc établi, et par simple homogénéité on en déduit que, si l'on pose

$$G' := \{A \in \mathcal{M}_n(\mathbb{R}) : \ \forall M \in G, \ \varphi_A(M) \leqslant 1\},$$

alors

$$\mathrm{Conv}(G) = \{M \in \mathcal{M}_n(\mathbb{R}) : \forall A \in G', \ \mathrm{tr}(AM) \leqslant 1\}. \tag{1}$$

On vient en outre d'observer que

$$G' = \left\{A \in \mathcal{M}_n(\mathbb{R}) : \ \widetilde{s_n}(A) + \sum_{k=1}^{n-1} s_k(A) \leqslant 1\right\}.$$

De plus $|||A||| = s_1(A) \leqslant 1$ pour tout $A \in G'$, donc G' est borné.
Nous allons ensuite observer que les équations permettant de récupérer $\mathrm{Conv}(G)$ via la formule (1) peuvent être limitées à un sous-ensemble plus précis de G'. Pour cela, on va remarquer que tout élément de G' est combinaison convexe d'éléments appartenant à deux G^2-orbites très particulières. Remarquons que G', par sa définition-même, est convexe et fermé, et qu'il est en outre compact puisqu'on a démontré qu'il est borné. Le théorème de Krein-Milman montre donc que G' est l'enveloppe convexe de l'ensemble de ses points extrémaux. Nous allons maintenant obtenir des informations sur ces points extrémaux.

Définition 1. *Dans $\mathcal{M}_n(\mathbb{R})$, on considère les matrices diagonales*

$$E := \mathrm{Diag}(1, 0, \dots, 0) \quad \textit{et} \quad F := \frac{1}{n-2}\,\mathrm{Diag}(1, \dots, 1, -1).$$

Proposition 4. *Tout point extrémal de G' est dans la G^2-orbite de E ou de F.*

Réciproquement, on pourrait établir que E et F sont effectivement extrémaux dans G', mais ce ne sera pas utile.
Démonstration : Prenons un point extrémal A de G'. Clairement, tout élément de la G^2-orbite de A reste extrémal dans G'. On peut donc supposer que $A = \mathrm{Diag}(\alpha_1, \dots, \alpha_{n-1}, \varepsilon\alpha_n)$ où $\alpha_1 \geqslant \cdots \geqslant \alpha_n \geqslant 0$, $\varepsilon = \pm 1$, et $\sum_{k=1}^{n-1} \alpha_k \leqslant 1 - \varepsilon\alpha_n$.

D'abord $A \neq 0$ sinon A serait le milieu du segment $[E, E']$ pour $E' := \mathrm{Diag}(-1, 0, \dots, 0)$, qui est dans la G^2-orbite de E et donc dans G'.
Supposons d'abord $\varepsilon = 1$. Alors directement

$$A = \left(1 - \sum_{k=1}^{n} \alpha_k\right).0 + \sum_{k=1}^{n} \alpha_k.D_k,$$

où D_k désigne la matrice diagonale dont le i-ème coefficient diagonal vaut $\delta_{i,k}$. Tous les points de cette combinaison convexe sont dans G', et ils sont distincts donc A ne peut être extrémal que si exactement l'un des coefficients vaut 1. Puisque $A \neq 0$ et $(\alpha_i)_{1 \leqslant i \leqslant n}$ est décroissante, la seule possibilité est que $\alpha_1 = 1$ et $\alpha_k = 0$ for tout $k \in \{2, \dots, n\}$. Ainsi $A = E$.
Supposons maintenant $\varepsilon = -1$ et $\alpha_n > 0$ (le cas $\varepsilon = -1$ et $\alpha_n = 0$ se réduisant au précédent). Supposons en outre que $(\alpha_i)_{1 \leqslant i \leqslant n}$ n'est pas constante et introduisons l'unique $k \in \{1, \dots, n-1\}$ tel que $\alpha_k > \alpha_{k+1} = \dots = \alpha_n$. Posons alors $\Delta := \mathrm{Diag}(0, \dots, 0, n-k-2, -1, \dots, -1, 1)$ où le coefficient $n-k-2$ est placé en position k. Pour t voisin de 0, on observe que les k premiers coefficients de $A_t := A + t\Delta$ sont strictement positifs, et tous supérieurs aux valeurs absolues des $n-k$ derniers, qui sont toutes égales et non nulles, et seul le dernier coefficient est négatif. On en déduit facilement que, pour t voisin de 0, la matrice A_t appartient à G' (et est de déterminant négatif). Ainsi, en écrivant $A = \frac{1}{2}A_t + \frac{1}{2}A_{-t}$ pour t voisin de 0^+, on contredirait l'extrémalité de A dans G'. Il s'ensuit que $\alpha_1 = \dots = \alpha_n = 1$, puis que $(n-2)\,\alpha_1 \leqslant 1$. Si cette dernière inégalité était stricte, on pourrait écrire $A = (n-2)\alpha_1.F + (1-(n-2)\alpha_1).0$, ce qui contredirait à nouveau l'extrémalité de A. Ainsi, $\alpha_1 = \dfrac{1}{n-2}$, et finalement $A = F$. □

Évidemment, E et F sont bien dans G', tout comme tout élément de leurs G^2-orbites respectives. Le théorème de Krein-Milman, combiné à la description (1), donne ainsi

$$\mathrm{Conv}(G) = \{M \in \mathcal{M}_n(\mathbb{R}) : \ \forall g \in G^2, \quad \mathrm{tr}((g.E)\,M) \leqslant 1 \quad \text{et} \quad \mathrm{tr}((g.F)\,M) \leqslant 1\}.$$

On peut alors conclure grâce à la Proposition 3.

Théorème 1. *Soit un entier naturel $n \geqslant 3$. Les éléments de l'enveloppe convexe de $\mathrm{SO}_n(\mathbb{R})$ sont les matrices $M \in \mathcal{M}_n(\mathbb{R})$ telles que*

$$s_1(M) \leqslant 1 \quad \textit{et} \quad \sum_{k=1}^{n-1} s_k(M) \leqslant n - 2 + \widetilde{s_n}(M).$$

On notera que la première inégalité peut se récrire $|||M||| \leqslant 1$, ce qui traduit simplement l'appartenance de M à l'enveloppe convexe de $\mathrm{O}_n(\mathbb{R})$, tandis que la seconde peut se récrire $\mathrm{tr}\, M \leqslant n-2$ sous réserve que $\det(M) \leqslant 0$. Il est clair enfin qu'aucune des deux inégalités du théorème 1 n'implique l'autre.
On notera enfin que le résultat du théorème 1 tombe en défaut lorsque $n = 2$. On a vu dans ce cas que $\mathrm{Conv}(\mathrm{SO}_2(\mathbb{R}))$ est le disque unité fermé du plan euclidien $\mathcal{P}$ et ne contient que des matrices de déterminant positif. Dans ce qui précède, la nécessité d'avoir $n > 2$ apparait clairement dans la démonstration du fait que toute forme linéaire non nulle sur $\mathcal{M}_n(\mathbb{R})$ a un maximum strictement positif sur G, qui conditionne tout le reste de l'étude.

Références

[1] G. Hardy, J. Littlewood et G. Pólya, Inequalities, Cambridge University Press, 1934.