D.M. 11 : convexité de fonctions sur $S_n^+(\mathbb{R})$

Pour le lundi 26 février 2024

Rappel de notations :

- On note $S_n(\mathbf{R})$ l'espace vectoriel des matrices symétriques à coefficients réels de taille $n \times n$.
- Sur $M_{n,1}(\mathbf{R})^2$, on définit l'application $\langle \cdot, \cdot \rangle$ par

$$\forall (X,Y) \in M_{n,1}(\mathbf{R})^2, \quad \langle X, Y \rangle = X^\top Y$$

où $X^\top$ est la transposée de X. On admet que l'on définit ainsi un produit scalaire sur $M_{n,1}(\mathbf{R})$. On note $\| \cdot \|$ la norme associée.
- On admet que l'application $A \in M_n(\mathbf{R}) \mapsto \|A\|_2 = \sqrt{\mathrm{Tr}\,(A^\top A)}$ est une norme sur $M_n(\mathbf{R})$.
- On note $S_n^+(\mathbf{R})$ (resp. $S_n^{++}(\mathbf{R})$) l'ensemble des matrices symétriques $S \in S_n(\mathbf{R})$ telles que

$$\forall X \in M_{n,1}(\mathbf{R}) \backslash \{0\}, \quad \langle SX, X \rangle \geq 0 \text{ (resp. } > 0) \, .$$

- Une application $f : C \to \mathbf{R}$ définie sur une partie convexe C d'un $\mathbf{R}$-espace vectoriel E est dite convexe si

$$\forall (x,y) \in C^2, \forall t \in [0,1], \quad f((1-t)x + ty) \leq (1-t)f(x) + tf(y)$$

Préliminaires sur les matrices symétriques

1) Montrer qu'une matrice $S \in S_n(\mathbf{R})$ appartient à $S_n^+(\mathbf{R})$ si, et seulement si, $\mathrm{Sp}(S) \subset \mathbf{R}_+$. De même, montrer que $S \in S_n^{++}(\mathbf{R})$ si, et seulement si, $\mathrm{Sp}(S) \subset \mathbf{R}_+^\star$.
2) Montrer que $S_n^+(\mathbf{R})$ et $S_n^{++}(\mathbf{R})$ sont des parties convexes de $M_n(\mathbf{R})$. Sont-elles des sous-espaces vectoriels de $M_n(\mathbf{R})$?

Préliminaires sur la convexité

3) Soit I un intervalle de $\mathbb{R}$ et $f : I \to \mathbb{R}$ une fonction convexe. Montrer que l'épigraphe $\mathcal{E}_f$ qui est le sous-ensemble défini par :

$$\mathcal{E}_f = \{(x,y) \in I \times \mathbb{R}, \ y \geq f(x)\}$$

est un sous-ensemble convexe de $\mathbb{R}^2$.

4) a) Soit C est une partie convexe d'un $\mathbf{R}$-espace vectoriel E. Montrer que pour tout $p \in \mathbf{N}^\star$, pour tout $(x_1, \dots, x_p) \in C^p$ et pour tout $(\lambda_1, \dots, \lambda_p) \in (\mathbf{R}_+)^p$ tel que $\sum_{i=1}^p \lambda_i = 1$, alors $\sum_{i=1}^p \lambda_i x_i \in C$.

Terminologie : une somme $\sum_{i=1}^p \lambda_i x_i$ avec $(\lambda_1, \dots, \lambda_p) \in (\mathbf{R}_+)^p$ tel que $\sum_{i=1}^p \lambda_i = 1$, s'appelle *une combinaison convexe* de $x_1, \dots, x_p$. La question précédente dit qu'un ensemble convexe est *stable par combinaison convexe.* On pourra raisonner par récurrence.

b) Justifier que pour un $x_1, \dots, x_p \in E$, l'ensemble $Conv(x_1, \dots, x_p)$ formé de toutes les combinaisons convexes de $x_1, \dots, x_p$ est le plus petit ensemble convexe de E contenant les points $x_1, x_2, \dots, x_p$, appelé *enveloppe convexe* de $x_1, \dots, x_p$.

c) Dessinez l'enveloppe convexe de trois points non alignés x_1, x_2, x_3 dans $E = \mathbb{R}^2$, en justifiant votre dessin.

5) Soit I intervalle de $\mathbf{R}$. Soit $f : I \to \mathbf{R}$ une fonction convexe. *A l'aide des deux questions précédentes*, montrer que pour tout $p \in \mathbf{N}^\star$, pour tout $(\lambda_1, \dots, \lambda_p) \in (\mathbf{R}_+)^p$ tel que $\sum_{i=1}^p \lambda_i = 1$ et pour tout $(x_1, \dots, x_p) \in I^p$, on a :

$$f\left(\sum_{i=1}^p \lambda_i x_i\right) \leq \sum_{i=1}^p \lambda_i f(x_i)$$

Premières inégalités sur les matrices symétriques

Soit $M \in S_n^+(\mathbf{R})$ une matrice non nulle.

6) A l'aide de ce qui précède, montrer l'inégalité : $\dfrac{\mathrm{Tr}(M)}{n} \geq \det^{1/n}(M)$.

Dans ce qui suit, on admet l'inégalité suivante :

$$\forall\,(x_1,\ldots,x_n) \in (\mathbf{R}_+)^n, \qquad 2\max\{x_1,\ldots,x_n\}\left(\frac{1}{n}\sum_{k=1}^n x_k - \prod_{k=1}^n x_k^{1/n}\right) \geq \frac{1}{n}\sum_{k=1}^n\left(x_k - \prod_{j=1}^n x_j^{1/n}\right)^2.$$

N.B. Un article contenant une démonstration de cette inégalité sera mise sur la page DM, pour les personnes les plus motivées !

7) a) Exprimer $\|M\|_2$ en fonction des valeurs propres de M.

b) En déduire que

$$\frac{\mathrm{Tr}(M)}{n} - \det^{1/n}(M) \geq \frac{\left\|M - \det^{1/n}(M)I_n\right\|_2^2}{2n\|M\|_2}.$$

Inégalités déduites d'un théorème de réduction simultanée

8) a) Soit $A \in M_n(\mathbb{R})$. Montrer que :

$$A \in S_n^+(\mathbb{R}) \Leftrightarrow \exists\, M \ \in M_n(\mathbb{R}) \quad A = M^\top.M.$$

Remarque : si $M = (C_1,\ldots,C_n)$ en colonne, alors pour tout i,j, $(M^\top.M)(i,j) = (C_i|C_j)$ cette matrice de p.s. s'appelle *matrice de Gram.* Donc on montre ici que les matrices symétriques positives sont exactement les matrices de Gram.

Indication pour le sens $\Rightarrow$ on pourra diagonaliser A en base orthonormée et utiliser la racine carrée de la matrice diagonale.

b) Soient $A \in S_n^{++}(\mathbf{R})$ et $B \in S_n(\mathbf{R})$. Montrer qu'il existe une matrice diagonale $D \in M_n(\mathbf{R})$ et une matrice $Q \in GL_n(\mathbf{R})$ telles que $B = Q^\top DQ$ et $A = Q^\top Q$.

Indication - Ecrire $A = M^\top.M$ comme au a) et considérer $(M^\top)^{-1}BM^{-1}$.

c) Avec les notations de la question précédente, montrer que si $B \in S_n^+(\mathbb{R})$ alors la matrice diagonale D de la question précédente est à coefficients positifs.

9) On se propose dans cette question de montrer que :

$$\forall (A,B) \in S_n^{++}(\mathbf{R})^2, \quad \det^{1/n}(A+B) \geq \det^{1/n}(A) + \det^{1/n}(B) \quad (\dagger)$$

a) Étudier la convexité de la fonction $t \mapsto \ln\left(1+\mathrm{e}^t\right)$.

b) En déduire que si $D \in D_n(\mathbf{R})$ est une matrice diagonale à coefficients positifs, alors :

$$(\det(I_n+D))^{\frac{1}{n}} \geq 1 + (\det(D))^{1/n}$$

c) Déduire de ce qui précède l'inégalité $(\dagger)$ annoncée en début de question.

10) a) En suivant la même stratégie qu'à la question précédente, montrer que, si A et B appartiennent $S_n^{++}(\mathbf{R})$, alors :

$$\forall t \in [0,1], \quad \det((1-t)A + tB) \geq \det(A)^{1-t}\det(B)^t.$$

b) Justifier que cette inégalité reste valable pour A et B seulement dans $S_n^+(\mathbf{R})$.

c) Que peut-on en déduire sur la fonction $\ln\circ\det$ sur $S_n^{++}(\mathbf{R})$?

Un ordre sur $S_n(\mathbb{R})$

11) a) Sur $S_n(\mathbb{R})$, on définit la relation $\geq$ par :

$$\forall\ (A,B) \in S_n(\mathbb{R})^2,\ A\ \geq B \Leftrightarrow A - B \in S_n^+(\mathbb{R}).$$

Montrer que $\geq$ est une relation d'ordre.

b) Un sous-ensemble C d'un $\mathbb{R}$-espace vectoriel E est appelé un *cône convexe* si, et seulement si, C est stable par combinaisons linéaires à coefficients strictement positifs, autrement dit, si et seulement si, pour tout $(x,y) \in C^2$ et tout $(\lambda,\mu) \in (\mathbb{R}^{+*})^2$, $\lambda x + \mu y \in C$. On dit qu'un cône convexe est *saillant* si $C \cap (-C) = \{0\}$.

Vérifier que $S_n^+(\mathbb{R})$ est un cône convexe saillant de $M_n(\mathbb{R})$.

c) Justifier que si E est un $\mathbb{R}$-e.v. quelconque et P est un cône convexe saillant de E, on peut associer à P un ordre défini par $\forall\,(u,v) \in E^2$, $u \geq v \Leftrightarrow u - v \in P$.

d) L'ordre défini au a) vérifie la propriété de la borne supérieure :

Soit $(A_k)_{k\in\mathbb{N}}$ une suite de matrices symétriques réelles, croissante pour la relation $\geq$ et majorée par une matrice B.

Montrer que cette suite converge dans $S_n(\mathbb{R})$.

Indication – On commencera par remarquer que pour tout $X \in M_{n,1}(\mathbb{R})$, $(X^\top A_k X)_{k\in\mathbb{N}}$ converge dans $\mathbb{R}$.

Croissance et concavité de la racine carrée sur $S_n^+(\mathbb{R})$ pour cet ordre

12) a) Soient S et T dans $S_n^+(\mathbb{R})$ tels que $S^2 \geq T^2$ (pour l'ordre dans $S_n^+(\mathbb{R})$ défini au 11a)). En écrivant $S^2 - T^2$ en fonction de S et $W = (S-T)$, montrer que toutes les valeurs propres de W sont positives, ce qui montre que $S \geq T$.

b) On admet ici qu'une matrice $S \in S_n^+(\mathbb{R})$ admet une unique racine carrée dans $S_n^+(\mathbb{R})$. Déduire du a) que la fonction $S \mapsto \sqrt{S}$ est concave sur $S_n^+(\mathbb{R})$ autrement dit que pour tout $(S,T) \in S_n^+(\mathbb{R})^2$,

$$\forall\ t \in [0,1],\ \sqrt{(1-t)S + tT} \geq (1-t)\sqrt{S} + t\sqrt{T}.$$

13) On se propose de retrouver le résultat de croissance du 12) a) et un autre à l'aide du théorème de réduction simultanée de la question 8b), mais dans le cadre de la *croissance stricte*.

Notation (attention) Pour tout $S,T \in S_n(\mathbb{R})$, on pose que $S > T$ ssi $(S-T) \in S_n^{++}(\mathbb{R})$ ce qui équivaut aussi à dire que pour toute colonne $X \in M_{n,1}(\mathbb{R}) \setminus \{0\}$, $(SX|X) > (TX|X)$.

N.B. Autrement dit par exemple $S > 0$ est une condition plus forte que ($S \geq 0$ et $S \neq 0$).

Soient A et B dans $S_n^{++}(\mathbb{R})$.

a) Avec 8b), montrer que $A^{-1}.B$ est diagonalisable dans $M_n(\mathbb{R})$ et que $\mathrm{Sp}(A^{-1}.B) \subset \mathbb{R}^{+*}$.

b) Montrer que $A - B > 0 \Leftrightarrow \mathrm{Sp}(A^{-1}.B) \subset]0,1[$.

c) On suppose que $A - B > 0$. Montrer que $B^{-1} - A^{-1} > 0$. Autrement dit $A \mapsto A^{-1}$ est décroissante strictement sur $S_n^{++}(\mathbb{R})$.

d) On suppose que $A - B > 0$. Montrer que $\sqrt{A} - \sqrt{B} > 0$ ce qui précise le résultat du 12 a).