Corrigé : quelques inégalités de convexité autour du déterminant (Mines PC 2023)

Partie 1 – Questions préliminaires

1 ▷ Supposons $S \in \mathcal{S}_n^+(\mathbb{R})$, et considérons une valeur propre λ et un vecteur propre $X \in \mathcal{M}_{n,1}(\mathbb{R})$. On a $SX = \lambda X$ donc $\langle SX \mid X\rangle = \lambda \|X\|^2 \geqslant 0$, ce qui implique $\lambda \geqslant 0$ puisque $X \neq 0$.
Réciproquement, supposons $\mathrm{Sp}(S) \subset \mathbb{R}_+$. D'après le théorème spectral, il existe $P \in \mathcal{O}_n(\mathbb{R})$ et $D = \mathrm{diag}(\lambda_1, \dots, \lambda_n)$ tel que $S = PDP^T$. Alors $\langle SX \mid X\rangle = X^T PDP^T X = Y^T DY$ avec $Y = P^T X$, et en posant $Y^T = (y_1, \dots, y_n)$ alors $\langle SX \mid X\rangle = \sum_{k=1}^{n} \lambda_k y_k^2 \geqslant 0$ donc $S \in \mathcal{S}_n^+(\mathbb{R})$.
On a donc prouvé l'équivalence : $S \in \mathcal{S}_n^+(\mathbb{R}) \iff \mathrm{Sp}(S) \subset \mathbb{R}_+$.

2 ▷ Soit $A, B \in \mathcal{S}_n^+(\mathbb{R})$, et $t \in [0,1]$. Alors pour tout $X \in \mathcal{M}_{n,1}(\mathbb{R})$, $\langle AX \mid X\rangle \geqslant 0$ et $\langle BX \mid X\rangle \geqslant 0$ donc

$$\Big\langle \big((1-t)A + tB\big)X \;\Big|\; X \Big\rangle = (1-t)\langle AX \mid X\rangle + t\langle BX \mid X\rangle \geqslant 0$$

Ceci montre que $(1-t)A + tB \in \mathcal{S}_n^+(\mathbb{R})$ et donc que $\mathcal{S}_n^+(\mathbb{R})$ est convexe.
Supposons maintenant $A, B \in \mathcal{S}_n^{++}(\mathbb{R})$, et $t \in [0,1]$. Alors pour tout $X \in \mathcal{M}_{n,1}(\mathbb{R}) \setminus \{0\}$, $\langle AX \mid X\rangle > 0$ et $\langle BX \mid X\rangle > 0$ donc $(1-t)\langle AX \mid X\rangle + t\langle BX \mid X\rangle > 0$ car t et $(1-t)$ ne s'annulent pas simultanément. Ceci montre que $(1-t)A + tB \in \mathcal{S}_n^{++}(\mathbb{R})$ et donc que $\mathcal{S}_n^{++}(\mathbb{R})$ est convexe.
En revanche, $\mathcal{S}_n^+(\mathbb{R})$ et $\mathcal{S}_n^{++}(\mathbb{R})$ ne sont pas des sous-espaces vectoriels de $\mathcal{M}_n(\mathbb{R})$ puisque I_n appartient à chacun de ces deux ensembles, mais pas $-I_n$.

3 ▷ D'après le théorème spectral, il existe $P \in \mathcal{O}_n(\mathbb{R})$ et $D = \mathrm{diag}(\lambda_1, \dots, \lambda_n)$ tels que $A = PDP^T$. D'après la question 1 on a $\lambda_k > 0$ pour tout $k \in [\![1, n]\!]$ donc en posant $S = P\Delta P^T$ avec $\Delta = \mathrm{diag}(\sqrt{\lambda_1}, \dots, \sqrt{\lambda_n})$ on a $S \in \mathcal{S}_n^{++}(\mathbb{R})$ et $S^2 = A$.

4 ▷ Raisonnons par récurrence sur $p \in \mathbb{N}^*$:

- si $p = 1$ alors $\lambda_1 = 1$ et $f(\lambda_1 x_1) = f(x_1) = \lambda_1 f(x_1)$;
- si $p \geqslant 2$, supposons le résultat acquis au rang $p-1$, et considérons $(\lambda_1, \dots, \lambda_p) \in (\mathbb{R}_+)^p$ tel que $\sum_{k=1}^{p} \lambda_k = 1$. Puisque $p \geqslant 2$ l'un au moins des λ_k est strictement inférieur à 1 ; supposons sans perte de généralité que $\lambda_p < 1$. Posons alors $x = \dfrac{1}{1-\lambda_p} \sum_{k=1}^{p-1} \lambda_k x_k$ et $y = x_p$. Alors $\sum_{k=1}^{p} \lambda_k x_k = (1-\lambda_p)x + \lambda_p y$.

  Puisque I est convexe et $\sum_{k=1}^{p-1} \dfrac{\lambda_k}{1-\lambda_p} = 1$, on a $x \in I$, et par hypothèse de récurrence, $f(x) \leqslant \sum_{k=1}^{p-1} \dfrac{\lambda_k}{1-\lambda_p} f(x_k)$.

  Ainsi, $f\Big(\sum_{k=1}^{p} \lambda_k x_k\Big) = f\big((1-\lambda_p)x + \lambda_p y\big) \leqslant (1-\lambda_p) f(x) + \lambda_p f(y) \leqslant \sum_{k=1}^{p} \lambda_k f(x_k)$: la récurrence se propage.

Partie 2 – Une première inégalité de convexité

5 ▷ On a $\mathrm{tr}(M) \geqslant 0$ (c'est la somme des valeurs propres, toutes positives) donc si $\det M = 0$ alors $\dfrac{1}{n} \mathrm{tr}\, M \geqslant (\det M)^{1/n}$.
Supposons maintenant $\det M > 0$, autrement dit $M \in \mathcal{S}_n^{++}(\mathbb{R})$. D'après le théorème spectral, $M = PDP^T$ avec $P \in \mathcal{O}_n(\mathbb{R})$ et $D = \mathrm{diag}(\lambda_1, \dots, \lambda_n)$. On a $\mathrm{tr}\, M = \sum_{k=1}^{n} \lambda_k$ et $\det M = \prod_{k=1}^{n} \lambda_k$.

La fonction $x \mapsto -\ln(x)$ est convexe sur $\mathbb{R}_+^*$ (sa dérivée seconde $x \mapsto \dfrac{1}{x^2}$ est positive) donc d'après la question 4, $-\ln\Big(\sum_{k=1}^{n} \dfrac{\lambda_k}{n}\Big) \leqslant -\sum_{k=1}^{n} \dfrac{1}{n} \ln(\lambda_k)$, soit en passant à l'exponentielle : $\dfrac{1}{n} \sum_{k=1}^{n} \lambda_k \geqslant \Big(\prod_{k=1}^{n} \lambda_k\Big)^{1/n}$, ou encore $\dfrac{1}{n} \mathrm{tr}\, M \geqslant (\det M)^{1/n}$.

6 ▷ Avec les notations de la question précédente, $\|M\|_2^2 = \operatorname{tr}(M^T M) = \operatorname{tr}(P^T D^2 P) = \operatorname{tr}(D^2) = \sum_{k=1}^{n} \lambda_k^2$ donc $\|M\|_2 = \left(\sum_{k=1}^{n} \lambda_k^2\right)^{1/2}$.

7 ▷ D'après l'inégalité admise et appliquée aux valeurs propres $\lambda_1, \dots, \lambda_n$ on a :

$$2\max(\lambda_1,\dots,\lambda_n)\left(\frac{1}{n}(\operatorname{tr} M) - (\det M)^{1/n}\right) \geqslant \frac{1}{n}\sum_{k=1}^{n}\left(\lambda_k - (\det M)^{1/n}\right)^2 = \frac{1}{n}\left\|M - (\det M)^{1/n} I_n\right\|^2$$

D'après la question précédente, $\|M\|_2 \geqslant \max(\lambda_1,\dots,\lambda_n)$ donc $\frac{1}{n}(\operatorname{tr} M) - (\det M)^{1/n} \geqslant \dfrac{\left\|M - (\det M)^{1/n} I_n\right\|^2}{2\|M\|_2}$.

Partie 3 – On continue avec la convexité

8 ▷ D'après la question 3, il existe $S \in \mathcal{S}_n^{++}(\mathbb{R})$ telle que $A = S^2$. La matrice $S^{-1}BS^{-1}$ est symétrique réelle donc d'après le théorème spectral, il existe $P \in \mathcal{O}_n(\mathbb{R})$ et D diagonale telle que $S^{-1}BS^{-1} = PDP^T$, soit $B = SPDP^TS = QDQ^T$ en posant $Q = SP$, et $QQ^T = SPP^TS = S^2 = A$. Notons que Q est bien inversible car S et P le sont.

Posons $D = \operatorname{diag}(d_1,\dots,d_n)$. Pour tout $X \in \mathcal{M}_{n,1}(\mathbb{R})$ on a $\langle BX \mid X\rangle = X^T QDQ^T X = Y^T DY = \sum_{k=1}^{n} d_k y_k^2$ en posant $Y = Q^T X$.
Ainsi, si $B \in \mathcal{S}_n^{++}(\mathbb{R})$ alors pour tout $Y \in \mathcal{M}_{n,1}(\mathbb{R}) \setminus \{0\}$, $\sum_{k=1}^{n} d_k y_k^2 > 0$, ce qui impose $d_k > 0$ pour tout $k \in [\![1, n]\!]$.

9 ▷ Notons $f : t \mapsto \ln(1 + e^t)$. f est de classe $\mathscr{C}^\infty$ sur $\mathbb{R}$, et pour tout $t \in \mathbb{R}$, $f'(t) = \dfrac{e^t}{1+e^t}$, $f''(t) = \dfrac{e^t}{(1+e^t)^2} > 0$ donc f est strictement convexe.

10 ▷ D'après la question 8, il existe $Q \in \mathcal{GL}_n(\mathbb{R})$ et D diagonale à coefficients strictement positifs tels que $A = QQ^T$ et $B = QDQ^T$. Ainsi, $\det A = (\det Q)^2$, $\det B = \det D \times (\det Q)^2 = (d_1 \cdots d_n)(\det Q)^2$ et $\det(A + B) = \det(I + D) \times (\det Q)^2 = (1 + d_1)\cdots(1 + d_n)(\det Q)^2$.

La fonction f de la question 9 est convexe donc $f\left(\frac{1}{n}\sum_{k=1}^{n}\ln(d_k)\right) \leqslant \frac{1}{n}\sum_{k=1}^{n} f\big(\ln(d_k)\big)$, ce qui donne en passant à l'exponentielle : $1 + (d_1 \cdots d_n)^{1/n} \leqslant \big((1+d_1)\cdots(1+d_n)\big)^{1/n}$. Il reste à multiplier par $(\det Q)^{2/n}$ pour obtenir $(\det A)^{1/n} + (\det B)^{1/n} \leqslant \det(A+B)^{1/n}$.

11 ▷ Si $t = 0$ ou $t = 1$ le résultat est évident.
Si $t \in]0,1[$ les matrices $(1-t)A$ et tB sont définie positives donc d'après la question précédente,

$$\det\big((1-t)A + tB\big) \geqslant \left(\det((1-t)A)^{1/n} + \det(tB)^{1/n}\right)^n = \left((1-t)(\det A)^{1/n} + t(\det B)^{1/n}\right)^n$$

La fonction ln est concave donc $\ln\left((1-t)(\det A)^{1/n} + t(\det B)^{1/n}\right) \geqslant \frac{1}{n}\big((1-t)\ln(\det A) + t\ln(\det B)\big)$, et en passant à l'exponentielle on obtient $\left((1-t)(\det A)^{1/n} + t(\det B)^{1/n}\right)^n \geqslant (\det A)^{1-t}(\det B)^t$, ce qui permet de conclure.
Si A et B sont dans $\mathcal{S}_n^+(\mathbb{R})$, $(1-t)A + tB$ aussi puisque cet ensemble est convexe (question 2) et donc $\det\big((1-t)A + tB\big) \geqslant 0$. Or si l'une de ces deux matrices n'est pas dans $\mathcal{S}_n^{++}(\mathbb{R})$ alors son déterminant est nul et donc $(\det A)^{1-t}(\det B)^t = 0$. L'inégalité reste valable pour A et B dans $\mathcal{S}_n^+(\mathbb{R})$.

12 ▷ En composant par la fonction ln on obtient : $\forall A, B \in \mathcal{S}_n^{++}(\mathbb{R})$, $\ln\circ\det\big((1-t)A + tB\big) \geqslant (1-t)\ln\circ\det A + t\ln\circ\det B$, ce qui prouve que la fonction $\ln\circ\det$ est concave sur $\mathcal{S}_n^{++}(\mathbb{R})$.

Partie 4 – Encore de la convexité !

13 ▷ Pour tout $t \in \mathbb{R}^*$, $g(t) = t^n \det\left(\frac{1}{t} I_n + A\right) = t^n \chi_{-A}\left(\frac{1}{t}\right) = t^n \prod_{k=1}^{n}\left(\frac{1}{t} + \lambda_k\right) = \prod_{k=1}^{n}\left(1 + \lambda_k t\right)$ où $\operatorname{Sp}(A) = (\lambda_1, \dots, \lambda_n)$ (notées avec multiplicité). Cette égalité se prolonge par continuité en $t = 0$ car un déterminant est continu, et cette expression polynomiale montre que g est de classe $\mathscr{C}^\infty$ sur $\mathbb{R}$.

14 ▷ À l'aide de l'inégalité de convexité : $\forall x > -1, \ln(1+x) \leqslant x$ on obtient :

$$\forall t \geqslant 0, \quad \ln\Big(\det(\mathrm{I}_n + t\mathrm{A})\Big) = \sum_{k=1}^{n} \ln(1+\lambda_k t) \leqslant \Big(\sum_{k=1}^{n} \lambda_k\Big) t = (\operatorname{tr} \mathrm{A}) t$$

Partie 5 – Et pour finir… de la convexité !

15 ▷ Montrons par récurrence sur n que pour toutes matrices A et M de $\mathcal{M}_n(\mathbb{R})$, f_{A} est une fonction polynomiale (et donc de classe $\mathscr{C}^\infty$) :

– c'est clair si $n = 1$;

– si $n > 1$, supposons le résultat acquis jusqu'au rang $n-1$. Alors en développant le déterminant $f_{\mathrm{A}}(t)$ suivant sa dernière ligne on obtient une expression de la forme :

$$f_{\mathrm{A}}(t) = \sum_{j=1}^{n} (-1)^{n+j}(a_{n,j} + t m_{n,j})\Delta_{n,j}(t)$$

où $\Delta_{n,j}(t)$ est le mineur de rang (n,j) de la matrice $\mathrm{A} + t\mathrm{M}$. Mais on peut appliquer l'hypothèse de récurrence à chacun de ces mineurs, qui sont donc des fonctions polynomiales, et de ce fait il en est de même de f_{A} : la récurrence se propage.

16 ▷ D'après la question 8 il existe $\mathrm{Q} \in \mathcal{GL}_n(\mathbb{R})$ et D diagonale telle que $\mathrm{A} = \mathrm{Q}\mathrm{Q}^{\mathrm{T}}$ et $\mathrm{M} = \mathrm{QDQ}^{\mathrm{T}}$. Alors $\mathrm{A} + t\mathrm{M} = \mathrm{Q}(\mathrm{I} + t\mathrm{D})\mathrm{Q}^{\mathrm{T}}$. En posant $\mathrm{D} = \operatorname{diag}(d_1, \dots, d_n)$ on a déjà vu à la question 8 que pour tout $\mathrm{X} \in \mathcal{M}_{n,1}(\mathbb{R}) \setminus \{0\}$, $\mathrm{X}^{\mathrm{T}}(\mathrm{A} + t\mathrm{M})\mathrm{X} = \mathrm{Y}^{\mathrm{T}}(\mathrm{I} + t\mathrm{D})\mathrm{Y} = \sum_{k=1}^{n} (1 + t d_k) y_k^2$ avec $\mathrm{Y} = \mathrm{Q}^{\mathrm{T}}\mathrm{X}$. Pour t assez petit on a $1 + t d_k > 0$ pour tout $k \in [\![1, n]\!]$ et alors $\mathrm{X}^{\mathrm{T}}(\mathrm{A} + t\mathrm{M})\mathrm{X} > 0$, ce qui prouve que $\mathrm{A} + t\mathrm{M} \in \mathcal{S}_n^{++}(\mathbb{R})$.

17 ▷ Dans le cas où $\mathrm{A} = \mathrm{I}_n$ on a $f_{\mathrm{I}_n}(t) = \det(\mathrm{I} + t\mathrm{M}) = \prod_{k=1}^{n} (1 + t\lambda_k)$ où $\lambda_1, \dots, \lambda_n$ sont les valeurs propres de M (question 13) et donc $f_{\mathrm{I}_n}(t) \underset{t\to 0}{=} 1 + t\sum_{k=1}^{n} \lambda_k + \mathrm{o}(t) \underset{t\to 0}{=} 1 + (\operatorname{tr} \mathrm{M})t + \mathrm{o}(t)$.

Dans le cas général, $f_{\mathrm{A}}(t) = \det \mathrm{A} \det(\mathrm{I}_n + t\mathrm{A}^{-1}\mathrm{M})$ et d'après le premier cas,

$$f_{\mathrm{A}}(t) \underset{t\to 0}{=} \det \mathrm{A}\Big(1 + t\operatorname{tr}(\mathrm{A}^{-1}\mathrm{M}) + \mathrm{o}(t)\Big) \underset{t\to 0}{=} \det \mathrm{A} + \det \mathrm{A} \operatorname{tr}(\mathrm{A}^{-1}\mathrm{M})t + \mathrm{o}(t)$$

18 ▷ Soit $t \in \,]-\epsilon_0, \epsilon_0[$, et $h \in \mathbb{R}^*$. Puisque $\mathrm{A} + t\mathrm{M}$ est inversible, on a

$$f_{\mathrm{A}}(t+h) = \det(\mathrm{A} + t\mathrm{M} + h\mathrm{M}) = \det(\mathrm{A} + t\mathrm{M})\det\Big(\mathrm{I}_n + h(\mathrm{A} + t\mathrm{M})^{-1}\mathrm{M}\Big)$$

et d'après la question précédente,
$$\begin{aligned} f_{\mathrm{A}}(t+h) &\underset{h\to 0}{=} \det(\mathrm{A} + t\mathrm{M})\Big(1 + h\operatorname{tr}\Big((\mathrm{A} + t\mathrm{M})^{-1}\mathrm{M}\Big) + \mathrm{o}(h)\Big) \\ &\underset{h\to 0}{=} f_{\mathrm{A}}(t) + h\det(\mathrm{A} + t\mathrm{M})\operatorname{tr}\Big((\mathrm{A} + t\mathrm{M})^{-1}\mathrm{M}\Big) + \mathrm{o}(h) \end{aligned}$$

donc $f_{\mathrm{A}}'(t) = \det(\mathrm{A} + t\mathrm{M})\operatorname{tr}\Big((\mathrm{A} + t\mathrm{M})^{-1}\mathrm{M}\Big)$.

19 ▷ En dérivant la relation $\phi(t) \times (\mathrm{A} + t\mathrm{M}) = \mathrm{I}_n$ on obtient $\phi'(t) \times (\mathrm{A} + t\mathrm{M}) + \phi(t)\mathrm{M} = 0$ et en particulier $\phi'(0) = -\phi(0)\mathrm{M}\mathrm{A}^{-1} = -\mathrm{A}^{-1}\mathrm{M}\mathrm{A}^{-1}$ donc d'après la formule de Taylor-Young, $\phi(t) \underset{t\to 0}{=} \mathrm{A}^{-1} - t\mathrm{A}^{-1}\mathrm{M}\mathrm{A}^{-1} + \mathrm{o}(t)$.

20 ▷ $\varphi_\alpha(t) = \dfrac{1}{\alpha f_{\mathrm{A}}(t)^\alpha}$ est dérivable comme composée de fonctions dérivables, et d'après la question 18,

$$\varphi_\alpha'(t) = -\frac{f_{\mathrm{A}}'(t)}{f_{\mathrm{A}}(t)^{\alpha+1}} = -\det(\mathrm{A} + t\mathrm{M})^{-\alpha}\operatorname{tr}\Big((\mathrm{A} + t\mathrm{M})^{-1}\mathrm{M}\Big)$$

21 ▷ On a $\varphi_\alpha'(t) = -\alpha\operatorname{tr}\Big(\phi(t)\mathrm{M}\Big)\varphi_\alpha(t)$. À l'aide de la question 19 et de la linéarité de la trace on peut écrire :

$$\operatorname{tr}\Big(\phi(t)\mathrm{M}\Big) \underset{t\to0}{=} \operatorname{tr}(\mathrm{A}^{-1}\mathrm{M}) - t\operatorname{tr}\Big((\mathrm{A}^{-1}\mathrm{M})^2\Big) + o(t)$$

Par ailleurs, $\varphi_\alpha(t) \underset{t\to0}{=} \varphi_\alpha(0) + t\varphi_\alpha'(0)t + o(t) \underset{t\to0}{=} \dfrac{1}{\alpha}(\det\mathrm{A})^{-\alpha} - t\operatorname{tr}(\mathrm{A}^{-1}\mathrm{M})(\det\mathrm{A})^{-\alpha} + o(t)$.
En faisant le produit de ces deux développements limités on obtient :

$$\varphi_\alpha'(t) \underset{t\to0}{=} -\operatorname{tr}(\mathrm{A}^{-1}\mathrm{M})(\det\mathrm{A})^{-\alpha} + t\Big(\alpha\operatorname{tr}(\mathrm{A}^{-1}\mathrm{M})^2 + \operatorname{tr}\Big((\mathrm{A}^{-1}\mathrm{M})^2\Big)\Big)(\det\mathrm{A})^{-\alpha} + o(t)$$

ce qui prouve que $\varphi_\alpha''(0) = \Big(\alpha\operatorname{tr}(\mathrm{A}^{-1}\mathrm{M})^2 + \operatorname{tr}\Big((\mathrm{A}^{-1}\mathrm{M})^2\Big)\Big)(\det\mathrm{A})^{-\alpha}$.

22 ▷ D'après la question 3 il existe S symétrique réelle telle que $\mathrm{A} = \mathrm{S}^2$. Alors $\mathrm{A}^{-1}\mathrm{M} = \mathrm{S}^{-1}\mathrm{S}^{-1}\mathrm{M} = \mathrm{S}^{-1}\Big(\mathrm{S}^{-1}\mathrm{M}\mathrm{S}^{-1}\Big)\mathrm{S}$ et la matrice $\mathrm{S}^{-1}\mathrm{M}\mathrm{S}^{-1}$ est bien symétrique donc $\mathrm{A}^{-1}\mathrm{M}$ est semblable à une matrice symétrique.

23 ▷ Notons $\mathrm{B} = \mathrm{S}^{-1}\mathrm{M}\mathrm{S}^{-1}$ la matrice symétrique introduite à la question précédente. La trace est un invariant de similitude donc $\varphi_\alpha''(0) = (\det\mathrm{A})^{-\alpha}\Big(\alpha(\operatorname{tr}\mathrm{B})^2 + \operatorname{tr}(\mathrm{B}^2)\Big)$.
On a $\det\mathrm{A} > 0$ et d'après l'inégalité de Cauchy-Schwarz, $\operatorname{tr}(\mathrm{B})^2 = \operatorname{tr}(\mathrm{B}^\mathrm{T}\mathrm{I})^2 = \langle\mathrm{B}\mid\mathrm{I}\rangle^2 \leqslant \|\mathrm{B}\|^2\|\mathrm{I}\|^2 = n\operatorname{tr}(\mathrm{B}^\mathrm{T}\mathrm{B}) = n\operatorname{tr}(\mathrm{B}^2)$ donc $\varphi_\alpha''(0) \geqslant (\det\mathrm{A})^{-\alpha}\Big(\alpha + \dfrac{1}{n}\Big)(\operatorname{tr}\mathrm{B})^2 \geqslant 0$.

24 ▷ Si $\varphi_\alpha''(0) > 0$ alors par continuité il existe $\eta > 0$ tel que pour tout $t \in]-\eta,\eta[$, $\phi_\alpha''(t) > 0$. Autrement dit, ϕ_α est convexe sur cet intervalle. Le graphe de cette fonction est donc situé au dessus de sa tangente en 0, ce qui s'écrit : $\varphi_\alpha(t) \geqslant \varphi_\alpha(0) + \varphi_\alpha'(0)t$, soit encore $\dfrac{1}{\alpha}\det(\mathrm{A}+t\mathrm{M})^{-\alpha} \geqslant \dfrac{1}{\alpha}(\det\mathrm{A})^{-\alpha} - \operatorname{tr}(\mathrm{A}^{-1}\mathrm{M})(\det\mathrm{A})^{-\alpha}t$.