Quadrature n° () 1-7

Sur la convexité du déterminant

par Erik THOMAS*

Résumé.

Dans cet article, nous montrons la convexité de la fonctionnelle $\mathscr{S}_n^{++} \ni M \mapsto \frac{1}{\kappa} \det^{-\kappa}(M)$ avec $\kappa \in \left[-\frac{1}{n}, +\infty\right[\setminus \{0\}$.
Puis, nous en déduisons une famille d'inégalités artihmético-géométriques qui généralisent l'inégalité arithmético-géométrique « usuelle ».

I Introduction

I.1 Notations utilisées

Dans toute la suite, n désigne un entier naturel supérieur à 1.

On note $\mathscr{M}_n(\mathbb{R})$ l'ensemble des matrices carrées à coefficients réelles d'ordre n et $\mathscr{G}_n(\mathbb{R})$ est l'ensemble des matrices carrées réelles d'ordre n inversibles.

$\mathscr{M}_n(\mathbb{R})$ sera muni de son produit scalaire usuel $\langle \cdot, \cdot \rangle$ défini par :

$$\forall (M,N) \in (\mathscr{M}_n(\mathbb{R}))^2, \quad \langle M, N \rangle = \operatorname{tr}\left({}^tMN\right).$$

On note $\|\cdot\|_{\mathrm{F}}$ la norme de Frobenius associée à ce produit scalaire.

$\mathscr{O}_n(\mathbb{R})$ est l'ensemble des matrices orthogonales réelles : $\mathscr{O}_n(\mathbb{R}) = \{A \in \mathscr{M}_n(\mathbb{R}),\, A^tA = I_n\}$.

$\mathscr{S}_n(\mathbb{R})$ est l'ensemble des matrices symétriques réelles d'ordre n.

$\mathscr{S}_n^+(\mathbb{R})$ (respectivement $\mathscr{S}_n^{++}(\mathbb{R})$) est l'ensemble des matrices symétrique réelles d'ordre n positives (respectivement définies positives), c'est-à-dire $M \in \mathscr{S}_n^+(\mathbb{R})$ (respectivement $M \in \mathscr{S}_n^{++}(\mathbb{R})$)

$$\forall X \in \mathscr{M}_{n,1}(\mathbb{R}), X \neq 0, \quad {}^tXMX \geq 0$$

$$\left(\text{respectivement } \forall X \in \mathscr{M}_{n,1}(\mathbb{R}),\, X \neq 0, {}^tXMX > 0\right).$$

Si $A \in \mathscr{M}_n(\mathbb{R})$, on note $\sigma(A)$ le spectre complexe de A.

On rappelle (conséquence du théorème spectral) qu'une matrice A symétrique réelle (respectivement positive, respectivement définie positive) vérifie :

$$\sigma(A) \subset \mathbb{R} \text{ (resp. } \sigma(A) \subset \mathbb{R}_+, \text{ resp. } \sigma(A) \subset \mathbb{R}_+^*).$$

I.2 But de l'article

Dans cet article, nous nous intéressons à la convexité de la fonctionnelle

$$\mathscr{S}_n^{++}(\mathbb{R}) \ni M \mapsto \frac{1}{\kappa} \det{}^{-\kappa}(M)$$

où $\kappa \in \left[-\frac{1}{n}, +\infty\right[\setminus \{0\}$.

La convexité de cette fonctionnelle a pour conséquence : pour tout $(M,N) \in (\mathscr{S}_n^+(\mathbb{R}))^2$,

$$\det{}^{1/n}(M+N) \geq \det{}^{1/n}(M) + \det{}^{1/n}(N).$$

L'inégalité précédente, inégalité, déjà intéressante en elle-même, est le point de départ de la preuve de l'inégalité de Brunn-Minkowski par méthode de transport optimal : si A et B sont des compacts de $\mathbb{R}^n$ et si $|\cdot|$ désigne la mesure de Lebesgue sur $\mathbb{R}^n$, alors

$$|A+B|^{1/n} \geq |A|^{1/n} + |B|^{1/n},$$

où $A+B = \{a+b,\, A \in A,\, b \in B\}$ est la somme de Minkowski de A et B.

Plus modestement, la concavité de la fonctionnelle introduite ci-dessus permettra de donner une famille d'inégalités de type arithmético-géométriques : pour tout $\kappa \in \left[-\frac{1}{n}, +\infty\right[\setminus \{0\}$,

$$\forall M \in \mathscr{S}_n^{++}(\mathbb{R}), \quad \frac{1}{\kappa} \det{}^{-\kappa}(M) \geq \frac{1}{\kappa} - \operatorname{tr}(M - I_n). \tag{1}$$

* erik.thomas@ens-rennes.fr

Cette inégalité, lorsque $\kappa = -\frac{1}{n}$, est l'inégalité arithmético-géométrique « classique ».

Nous énoncerons une version quantitative de (1) lorsque $\kappa = -\frac{1}{n}$: pour tout $M \in \mathscr{S}_n^{++}(\mathbb{R})$,

$$\frac{\operatorname{tr}(M)}{n} - \det{}^{1/n}(M) \geq \frac{\|M - \det{}^{1/n}(M) I_n\|_{\mathrm{F}}^2}{2n\|M\|_{\mathrm{F}}},$$

où, on le rappelle, $\|M\|_{\mathrm{F}} = \sqrt{\operatorname{tr}({}^t MM)}$ est la norme de Frobenius de la matrice M

II Rappels sur la différentiabilité et sur les fonctions convexes

Dans cette partie, nous faisons des rappels de calcul différentiel et de convexité. Pour un exposé plus complet, nous renvoyons à [2] et [4].

II.1 Différentiabilité

Définition 1. *Différentiabilité.*

Soient $(E, \|\cdot\|_E)$ et $(F, \|\cdot\|_F)$ deux espaces vectoriels normés. Soit $\Omega \subset E$ un ouvert. Soit $f : \Omega \to F$ une application.

On dit que f est différentiable en $a \in \Omega$ s'il existe une application linéaire $df_a : E \to F$ continue telle que

$$\forall h \in E, \quad f(a+h) = f(a) + df_a(h) + o(\|h\|_E).$$

Remarque 2. *La continuité de df_a est automatique lorsque E est de dimension finie.*

Remarque 3. *La continuité de df_a implique la continuité de f en a.*

Proposition 4. *Soient $(E, \|\cdot\|_E)$ et $(F, \|\cdot\|_F)$ deux espaces vectoriels normés. Soit Ω un ouvert de E et soit $f : \Omega \to F$ une application différentiable en $a \in \Omega$.*

Alors, f est Gâteaux différentiable : pour tout $h \in E$,

$$df_a(h) = \lim_{t \to 0} \frac{f(a+th) - f(a)}{t}.$$

Remarque 5. *La réciproque de cette proposition est fausse : la Gâteaux différentiabilité n'entraîne pas la différentiabilité. L'existence de la limite, même pour tout $h \in E$, n'entraîne, à priori, pas la différentiabilité de f en a.*

Nous allons étudier deux exemples classiques qui seront utiles par la suite.

Exemple 6. *Dans cet exemple, nous allons calculer la différentielle du déterminant.*

On rappelle que $\mathscr{G}_n(\mathbb{R})$ est un ouvert dense de $\mathscr{M}_n(\mathbb{R})$.

L'application det *est différentiable en toute matrice $A \in \mathscr{M}_n(\mathbb{R})$ et*

$$\forall H \in \mathscr{M}_n(\mathbb{R}), \quad d(\det)_A(H) = \operatorname{tr}\left({}^t\operatorname{com}(A) H\right).$$

Démonstration. Avant de commencer la preuve, nous allons énoncer le lemme suivant.

Lemme 7. *Pour tout $A \in \mathscr{M}_n(\mathbb{R})$, on a :*

$$\det(M + I_n) = 1 + \operatorname{tr}(M) + o(\|M\|_{\mathrm{F}}).$$

Nous prouvons le lemme 7.

Démonstration. Nous donnons la preuve pour $n = 2$.

Si $M = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, on a

$$\begin{aligned}\det(M + I_2) &= 1 + (a+d) + ad - bc \\ &= 1 + \operatorname{tr}(M) + (ad - bc).\end{aligned}$$

On remarque que $ad - bc$ est un $o(\|M\|)$ car, en prenant la norme du plus grand coefficient, majoré par $o\left(\|M\|^2\right)$.

□

Reprenons la preuve de l'exemple 6.

• On commence par calculer la différentielle du déterminant sur $\mathscr{G}_n(\mathbb{R})$ Soit $A \in \mathscr{G}_n(\mathbb{R})$. Pour tout $H \in \mathscr{M}_n(\mathbb{R})$, on a :

$$\begin{aligned}\det(A+H) &= \det(A)\det\left(I_n + A^{-1}H\right) \\ &= \det(A)\left(1 + \operatorname{tr}\left(A^{-1}H\right) + o(\|H\|_{\mathrm{F}})\right) \\ &= \det(A) + \operatorname{tr}\left(\det(A)A^{-1}H\right) + o(\|H\|_{\mathrm{F}}) \\ &= \det(A) + \operatorname{tr}\left({}^t\operatorname{com}(A)H\right) + o(\|H\|_{\mathrm{F}}).\end{aligned}$$

Ainsi, pour tout $H \in \mathscr{M}_n(\mathbb{R})$, $d(\det)_A(H) = \operatorname{tr}({}^t\operatorname{com}(A)H)$.

• Les applications $X \mapsto d(\det)_X(\cdot)$ et $X \mapsto \operatorname{tr}({}^t\operatorname{com}(X)\cdot)$ sont continues et coincident sur $\mathscr{G}_n(\mathbb{R})$ qui est dense dans $\mathscr{M}_n(\mathbb{R})$: elles sont donc égales sur $\mathscr{M}_n(\mathbb{R})$.

□

Exemple 8. *Soit $(E, \|\cdot\|_E)$ un espace vectoriel normé. Soit $\mathscr{L}(E)$ l'ensemble des endomorphismes continus de E que l'on munit de la norme subordonnée $N_{\mathscr{L}(E)}$ définie par :*

$$\forall u \in \mathscr{L}(E), \quad N_{\mathscr{L}(E)}(u) = \sup_{\substack{x \in E \\ x \neq 0}} \frac{\|u(x)\|_E}{\|x\|_E}.$$

Rappelons que $N_{\mathscr{L}(E)}$ est une norme d'algèbre, c'est-à-dire que

$$\forall (u,v) \in \mathscr{L}(E)^2, \quad N_{\mathscr{L}(E)}(u \circ v) \leq N_{\mathscr{L}(E)}(u) \times N_{\mathscr{L}(E)}(v).$$

On note $\mathscr{G}(E)$ l'ensemble des automorphismes continus de E. Soit $\varphi : \mathscr{G}(E) \to \mathscr{G}(E)$ définie par : pour tout $u \in \mathscr{G}(E)$, $\varphi(u) = u^{-1}$. Alors, φ est différentiable sur $\mathscr{G}(E)$ et

$$\forall a \in \mathscr{G}(E), \forall h \in \mathscr{L}(E), \quad d\varphi_a(h) = -a^{-1} \circ h \circ a^{-1}.$$

Démonstration. On commence par calculer la différentielle en id_E.

Pour tout $h \in \mathscr{L}(E)$ tel que $N_{\mathscr{L}(E)}(h) < 1$, $id + h \in \mathscr{G}(E)$ et son inverse s'exprime comme une série absolument convergente dans $\mathscr{L}(E)$ (voir [5])

$$(id_E + h)^{-1} = \sum_{k=0}^{+\infty} (-1)^k h^k = id_E - h + o\left(N_{\mathscr{L}(E)}(h)\right).$$

Il s'ensuit que l'on a

$$\varphi(id_E + h) = \varphi(id_E) + d\varphi_{id_E}(h) + o\left(N_{\mathscr{L}(E)}(h)\right).$$

Ainsi φ est différentiable en id_E et $d\varphi_{id_E} = -id_{\mathscr{L}(E)}$.

On revient au cas général : soit $a \in \mathscr{G}(E)$, pour tout $h \in \mathscr{L}(E)$, on a $a + h = a \circ \left(id_E + a^{-1} \circ h\right)$.

Comme $N_{\mathscr{L}(E)}\left(a^{-1} \circ h\right) \leq N_{\mathscr{L}(E)}\left(a^{-1}\right) \times N_{\mathscr{L}(E)}(h)$, si l'on suppose $N_{\mathscr{L}(E)}\left(a^{-1}\right) \times N_{\mathscr{L}(E)}(h) < 1$, alors $id_E + a^{-1} \circ h$ est inversible et

$$\begin{aligned}
\varphi(a+h) &= (a+h)^{-1} \\
&= \left(id_E + a^{-1} \circ h\right) \circ a^{-1} \\
&= \left(id_E - a^{-1} \circ h + o\left(N_{\mathscr{L}(E)}(h)\right)\right) \circ a^{-1} \\
&= a^{-1} - a^{-1} \circ h \circ a^{-1} + o\left(N_{\mathscr{L}(E)}(h)\right) \\
&= \varphi(a) + d\varphi_a(h) + o\left(N_{\mathscr{L}(E)}(h)\right).
\end{aligned}$$

Il est clair que l'application $d\varphi_a$ est continue.

□

Définition 9. *Matrice hessienne.*

Soit $\Omega \subset \mathbb{R}^n$ un ouvert. Soit $f : \Omega \to \mathbb{R}$ une fonction de classe C^2. Soit $a \in \Omega$.

On définit la matrice hessienne de f en a par :

$$\mathrm{H}_f(a) = \left(\frac{\partial^2 f(a)}{\partial x_i \partial x_j}\right)_{(i,j) \in [\![1,n]\!]^2}.$$

Comme la matrice hessienne est symétrique (conséquence du lemme de Schwarz), la forme quadratique $q_a : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ définie par :

$$\forall (x,y) \in \mathbb{R}^n \times \mathbb{R}^n, \qquad q_a(x,y) = {}^t x \mathrm{H}_f(a) y$$

est symétrique.

Proposition 10. *Soit Ω un ouvert de $\mathbb{R}^n$ et soit $f : \Omega \to \mathbb{R}$ une application de classe C^2 sur Ω.*

Alors, pour tout $a \in \Omega$, pour tout $h \in \mathbb{R}^n$, on a :

$$q_a(h,h) = \lim_{t \to 0} \frac{df_{a+th}(h) - df_a(h)}{t}.$$

Proposition 11. *Taylor-Young.*

Soit Ω un ouvert de $\mathbb{R}^n$. Soit $f : \Omega \to \mathbb{R}$ une fonction de classe C^2 sur Ω. Soit $a \in \Omega$. Alors, pour tout $h \in \mathbb{R}^n$ tel que $a + h \in \Omega$

$$f(a+h) = f(a) + df_a(h) + \frac{1}{2} {}^t q_a(h,h) + o\left(\|h\|^2\right).$$

II.2 Fonctions convexes

Définition 12. *Soit E un espace vectoriel. Soit $C \subset E$.*

On dit que C est convexe si :

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

Définition 13. *Soit $(E, \|\cdot\|_E)$ un espace vectoriel normé. Soit $\Omega \subset E$ un convexe. Soit $f : \Omega \to \mathbb{R}$. On dit que f est convexe sur Ω si : pour tout $(x,y) \in \Omega^2$ et pour tout $t \in [0,1]$,*

$$f((1-t)x + ty) \leq (1-t) f(x) + t f(y).$$

Proposition 14. *Soit Ω un ouvert convexe de $\mathbb{R}^n$. Soit $f : \Omega \to \mathbb{R}$ une fonction de classe C^2 sur Ω. Si $x \in \Omega$, on note q_x la forme quadratique associée à la matrice hessienne de f en x. Les propositions suivantes sont équivalentes :*

1) f est convexe sur Ω

2) pour tout $(x,y) \in \Omega^2$,

$$f(y) \geq f(x) + df_x(y-x) \geq 0$$

3) pour tout $(x,y) \in \Omega^2$,

$$(df_y - df_x)(y-x) \geq 0$$

4) pour tout $x \in \Omega$, pour tout $h \in \mathbb{R}^n$,

$$q_a x(h,h) \geq 0.$$

Démonstration. • 1) ⇒ 2)

Soient $(x,y) \in \Omega^2$, soit $t \in [0,1]$ Par convexité de f, on a :

$$f(x + t(y-x)) = f((1-t)x + ty) \leq (1-t) f(x) + t f(y).$$

En supposant $t > 0$, on récupère

$$\frac{f(x + t(y-x)) - f(x)}{t} \leq f(y) - f(x).$$

Comme $\lim_{t \to 0} \frac{f(x+t(y-x)) - f(x)}{t} = df_x(y-x)$, on a bien

$$df_x(y-x) + f(x) \leq f(y).$$

• 2) ⇒ 1)

Soient $(x,y) \in \Omega^2$ et $t \in [0,1]$.

On écrit 2) avec les couples $(x, x + t(y-x))$ et $(y, x + t(y-x))$, on a :

$$f(y) \geq f(x + t(y-x)) + df_{x+t(y-x)}((1-t)(y-x))$$

et

$$f(y) \geq f(x + t(y-x)) + df_{x+t(y-x)}(-t(y-x)).$$

En multipliant la première ligne par t, la seconde par $1-t$ et en sommant, on obtient

$$tf(y)+(1-t)f(x)\geq f(x+t(y-x))=f((1-t)x+ty).$$

f est convexe sur Ω.

• 2) $\Rightarrow$ 3)

Soit $(x,y)\in\Omega^2$. On a

$$f(y)\geq f(x)+df_x(y-x)$$

et

$$f(x)\geq f(y)+df_y(x-y).$$

En sommant ces deux inégalités, on obtient 3).

• 3) $\Rightarrow$ 2)

Soit la fonction $\varphi:[0,1]\to\Omega$ définie par $\varphi(t)=f(x+t(y-x))$.

Comme f est de classe C^1 sur Ω, φ est dérivable sur $[0,1]$ et $\varphi'(t)=df_{x+t(y-x)}(y-x)$ et $\forall t\in[0,1]$, on a

$$\varphi'(t)-\varphi'(0)=\frac{1}{t}\left(df_{x+t(y-x)}-df_x\right)(t(y-x))\geq 0.$$

On remarque que cette inégalité reste valable pour $t=0$. Par l'inégalité des accroissements finis, on a donc :

$$\varphi(1)-\varphi(0)\geq\varphi'(0)\Leftrightarrow f(y)\geq f(x)+df_x(y-x).$$

• 3) $\Rightarrow$ 4)

Soit $h\in\mathbb{R}^n$ et $t\in\mathbb{R}^*$ assez petit. D'après 3), on a

$$(df_{x+th}-df_x)(th)\geq 0\Leftrightarrow\left(\frac{df_{x+th}-df_x}{t}\right)(h)\geq 0.$$

En faisant tendre t vers 0, on obtient

$$q_x(h,h)\geq 0.$$

• 4) $\Rightarrow$ 3)

Soit $(x,y)\in\Omega^2$ et soit la fonction $\psi:[0,1]\to\mathbb{R}$ définie par $\psi(t)=df_{x+t(y-x)}(y-x)$.

ψ est dérivable sur $[0,1]$ et $\psi'(t)={}^t(y-x)\mathrm{H}_f(x+t(y-x))(y-x)\geq 0$. D'où par l'inégalité des accroissements finis,

$$\psi(1)-\psi(0)\geq 0\Leftrightarrow(df_y-df_x)(y-x)\geq 0.$$

□

III Résultats principaux

III.1 Concavité de $\det^{1/n}$ sur $\mathscr{S}_n^+(\mathbb{R})$

Avant de s'intéresser aux questions de convexité du déterminant, on montre les propositions suivantes.

Proposition 15. *Les ensembles $\mathscr{S}_n^+(\mathbb{R})$ et $\mathscr{S}_n^{++}(\mathbb{R})$ sont convexes.*

Démonstration. Soient $(A,B)\in\mathscr{S}_n^+(\mathbb{R})$ (respectivement $\mathscr{S}_n^{++}(\mathbb{R})$). Soit $t\in[0,1]$.

Il est clair que la matrice $(1-t)A+tB$ est symétrique réelle et pour tout $X\in\mathscr{M}_{n,1}(\mathbb{R})$, $X\neq 0$, on a :

$${}^tX((1-t)A+tB)X=(1-t)\,{}^tXAX+t\,{}^tXBX\geq 0$$

$$(\text{respectivement }>0).$$

Il s'ensuit que la matrice $(1-t)A+tB\in\mathscr{S}_n^+(\mathbb{R})$ (respectivement $\mathscr{S}_n^{++}(\mathbb{R})$).

□

Proposition 16. *$\mathscr{S}_n^{++}(\mathbb{R})$ est dense dans $\mathscr{S}_n^+(\mathbb{R})$.*

Démonstration. Soit $M\in\mathscr{S}_n^+(\mathbb{R})$: il existe $P\in\mathscr{O}_n(\mathbb{R})$ et $D=\operatorname{diag}(\lambda_1,\cdots,\lambda_n)$, avec pour tout $k\in[\![1,n]\!]$ $\lambda_k\geq 0$ tels que $M={}^tPDP$.

Soit $(M_m)_{m\in\mathbb{N}}$ la suite de matrices définies par : pour tout $m\in\mathbb{N}$, $M_m=M+\frac{1}{m+1}I_n$.

Il est facile de voir que pour tout $m\in\mathbb{N}$, $M_m\in\mathscr{S}_n^{++}(\mathbb{R})$ que $\lim_{m\to+\infty}M_m=M$.

Ainsi toute matrice de $S_n^+(\mathbb{R})$ est limite d'une suite de matrices de $\mathscr{S}_n^{++}(\mathbb{R})$: on a bien montré que $\mathscr{S}_n^{++}(\mathbb{R})$ est dense dans $\mathscr{S}_n^+(\mathbb{R})$.

□

Le but de ce paragraphe est de montrer la proposition suivante.

Proposition 17.

$$\forall A,B\in\mathscr{S}_n^+(\mathbb{R}),\quad \det{}^{1/n}(A+B)\geq\det{}^{1/n}(A)+\det{}^{1/n}(B).$$

Démonstration. • On suppose que $(A,B)\in\left(\mathscr{S}_n^{++}(\mathbb{R})\right)^2$. Par le théorème de réduction simultanée (voir [3]), il existe $P\in GL_n(\mathbb{R})$ et D diagonale telles que : $A={}^tPP$ et $B={}^tPDP$. Ainsi

$$\begin{aligned}&\det{}^{1/n}(A+B)\geq\det{}^{1/n}(A)+\det{}^{1/n}(B)\\ \Leftrightarrow\ &\det{}^{1/n}(I_n+D)\geq\det{}^{1/n}(I_n)+\det{}^{1/n}(D).\end{aligned}$$

On écrit $D=\operatorname{diag}(\lambda_1,\cdots,\lambda_n)$. Comme $B\in S_n^{++}(\mathbb{R})$, pour tout $i\in[\![1,n]\!]$, $\lambda_i>0$, ainsi pour tout $i\in[\![1,n]\!]$, il existe $\mu_i\in\mathbb{R}$ tel que $\lambda_i=e^{\mu_i}$ de sorte que

$$\det{}^{1/n}(D+I_n)=\left(\prod_{i=1}^{n}(1+e^{\mu_i})\right)^{1/n}$$

$$\text{et }\ \det{}^{1/n}(I_n)+\det{}^{1/n}(D)=1+\prod_{i=1}^{n}e^{\frac{\mu_i}{n}}.$$

Soit la fonction f définie sur $\mathbb{R}$ par $f(t)=\ln(1+e^t)$. f est deux fois dérivable sur $\mathbb{R}$ et

$$\forall t\in\mathbb{R},\qquad f'(t)=\frac{e^t}{1+e^t}\text{ et }f''(t)=\frac{1}{(1+e^t)^2}\geq 0.$$

Il s'ensuit que f est convexe sur $\mathbb{R}$ ainsi

$$f\left(\frac{1}{n}\sum_{i=1}^{n}\mu_i\right) \leq \frac{1}{n}\sum_{i=1}^{n} f(\mu_i).$$

Or

$$f\left(\frac{1}{n}\sum_{i=1}^{n}\mu_i\right) = \ln\left(1+e^{\frac{1}{n}\sum_{i=1}^{n}\mu_i}\right) = \ln\left(1+\prod_{i=1}^{n} e^{\frac{\mu_i}{n}}\right)$$

et

$$\frac{1}{n}\sum_{i=1}^{n} f(\mu_i) = \frac{1}{n}\sum_{i=1}^{n}\ln(1+e^{\mu_i}) = \ln\left(\left(\prod_{i=1}^{n}(1+e^{\mu_i})\right)^{1/n}\right).$$

Ainsi

$$\ln\left(1+\prod_{i=1}^{n} e^{\frac{\mu_i}{n}}\right) \leq \ln\left(\left(\prod_{i=1}^{n}(1+e^{\mu_i})\right)^{1/n}\right).$$

On conclut en composant cette inégalité par la fonction exponentielle.

- Lorsque l'on suppose $(A,B) \in (\mathscr{S}_n^{+}(\mathbb{R}))^2$, on conclut en utilisant la densité de $\mathscr{S}_n^{++}(\mathbb{R})$ et la continuité du déterminant.

$\square$

Corollaire 18. *log-concavité de* det *sur* $\mathscr{S}_n^{+}(\mathbb{R})$.

$M \mapsto \det^{1/n}(M)$ est log-concave sur $S_n^{+}(\mathbb{R})$: $\forall (A,B) \in (S_n^{+}(\mathbb{R}))^2$, $\forall t \in [0,1]$,

$$\det((1-t)A+tB) \geq \det^{1-t}(A)\det^{t}(B).$$

Démonstration. Nous aurons besoin du lemme suivant dont nous laissons la preuve au lecteur (l'argument crucial étant la concavité de la fonction ln).

Lemme 19. *Inégalité arithmético-géométrique.*

Soient a et b deux réels positifs. Soient p et q deux réels strictement positifs tels que $\frac{1}{p}+\frac{1}{q}=1$. Alors

$$ab \leq \frac{1}{p}a^p + \frac{1}{q}b^q.$$

L'argument de la preuve est l'homogénéité de $\det^{1/n}$: pour tout $A \in \mathscr{S}_n^{+}(\mathbb{R})$ et pour tout $\lambda \geq 0$,

$$\det^{1/n}(\lambda A) = \lambda \det^{1/n}(A).$$

Soient $(A,B) \in (\mathscr{S}_n^{+}(\mathbb{R}))^2$. Soit $t \in [0,1]$. En utilisant la proposition 17

$$\det^{1/n}((1-t)A+tB) \geq \det^{1/n}((1-t)A) + \det^{1/n}(tB).$$

L'homogénéité de det donne

$$\det^{1/n}((1-t)A) + \det^{1/n}(tB) = (1-t)\det^{1/n}(A) + t\det^{1/n}(B).$$

Le lemme 19 assure que

$$(1-t)\det^{1/n}(A) + t\det^{1/n}(B) \geq \left(\det^{1-t}(A)\det^{t}(B)\right)^{1/n}.$$

On a montré que

$$\det^{1/n}((1-t)A+tB) \geq \left(\det^{1-t}(A)\det^{t}(B)\right)^{1/n}.$$

On conclut la preuve en passant à la puissance n.

$\square$

III.2 Résultat principal

Voici le résultat principal :

Théorème 20. *Pour tout $\kappa \in \left[-\frac{1}{n},+\infty\right[\setminus \{0\}$, l'application $g : \mathscr{S}_n^{++}(\mathbb{R}) \to \mathbb{R}$ définie par :*

$$g(M) = \frac{1}{\kappa}\det^{-\kappa}(M)$$

est convexe sur $\mathscr{S}_n^{++}(\mathbb{R})$.

Démonstration. On rappelle que le déterminant d'une matrice dépend polynomialement des coefficients de cette matrice, donc $M \mapsto \det^{1/n}(M)$ est de classe C^∞ sur $\mathscr{S}_n^{++}(\mathbb{R})$ et d'après l'exemple 6, on a : $\forall S \in \mathscr{S}_n^{++}(\mathbb{R})$, $\forall T \in \mathscr{S}_n^{++}(\mathbb{R})$

$$\begin{aligned} dg_S(T) &= -\det^{-\kappa-1}(S)\det(S)\operatorname{tr}\left(S^{-1}T\right) \\ &= -\det^{-\kappa}(S)\operatorname{tr}\left(S^{-1}T\right). \end{aligned}$$

Soit $S \in \mathscr{S}_n^{++}(\mathbb{R})$. Pour tout $H \in \mathscr{S}_n(\mathbb{R})$ et pour tout $t \in \mathbb{R}$, t assez petit, on a :

$$\begin{aligned} \Delta_{S,H}(t) :=& dg_{S+tH}(H) - dg_S(H) \\ =& -\det^{-\kappa}(S+tH)\operatorname{tr}\left((S+tH)^{-1}H\right) + \\ & \det^{-\kappa}(S)\operatorname{tr}\left(S^{-1}H\right). \end{aligned}$$

En utilisant la différentielle du déterminant, on obtient :

$$\begin{aligned} \det^{-\kappa}(S+tH) =& \det^{-\kappa}(S) - \kappa\det^{-\kappa}(S)\operatorname{tr}\left(S^{-1}H\right)t \\ & + o(t), \end{aligned}$$

puis en utilisant la différentielle de l'inverse (voir exemple 8), on obtient :

$$\begin{aligned} \operatorname{tr}\left((S+tH)^{-1}H\right) &= \operatorname{tr}\left(\left(S^{-1} - S^{-1}HS^{-1}t + o(t)\right)H\right) \\ &= \operatorname{tr}\left(S^{-1}H\right) - \operatorname{tr}\left(\left(S^{-1}H\right)^2\right)t + o(t). \end{aligned}$$

On récupère :

$$\begin{aligned} \Delta_{S,H}(t) =& -\left(\det^{-\kappa}(S) - \kappa\det^{-\kappa}(S)\operatorname{tr}\left(S^{-1}H\right)t + o(t)\right) \\ & \times\left(\operatorname{tr}\left(S^{-1}H\right) - \operatorname{tr}\left(\left(S^{-1}H\right)^2\right)t + o(t)\right) + \\ & \det^{-\kappa}(S)\operatorname{tr}\left(S^{-1}H\right) \\ =& \, t\det^{-\kappa}(S)\left(\kappa\operatorname{tr}^2\left(S^{-1}H\right) + \operatorname{tr}\left(\left(S^{-1}H\right)^2\right)\right) + \\ & o(t). \end{aligned}$$

En supposant $t \neq 0$, on en déduit

$$\begin{aligned} \mathrm{H}_g(S)(H,H) &= \lim_{t\to 0} \frac{\Delta_{S,H}(t)}{t} \\ &= \det{}^{-\kappa}(S)\left(\kappa \mathrm{tr}^2\left(S^{-1}H\right) + \mathrm{tr}\left(\left(S^{-1}H\right)^2\right)\right) \end{aligned}$$

Lemme 21. *Soit* $(A,B) \in \mathscr{S}_n(\mathbb{R}) \times \mathscr{S}_n^+(\mathbb{R})$. *Alors,* $\sigma(AB) \subset \mathbb{R}$.

Nous prouvons le lemme.

Démonstration. On rappelle que si A et B sont deux matrices réelles, alors $\sigma(AB) = \sigma(BA)$.

Comme $B \in \mathscr{S}_n^+(\mathbb{R})$, il existe $P \in \mathscr{O}_n(R)$ et $D = \mathrm{diag}(\lambda_1, \cdots, \lambda_n)$ avec pour tout $i \in [\![1,n]\!]$, $\lambda_i \geq 0$ telles que $B = PDP^{-1} = PD^tP$.

On note $\widetilde{D} = \mathrm{diag}\left(\sqrt{\lambda_1}, \cdots, \sqrt{\lambda_n}\right)$ et $\widetilde{B} = P\widetilde{D}P^{-1} \in \mathscr{S}_n(\mathbb{R})$ de sorte que l'on ait $\widetilde{B}^2 = B$.

On a :

$$\sigma(AB) = \sigma\left(A\widetilde{B}\widetilde{B}\right) = \sigma\left(\widetilde{B}A\widetilde{B}\right) \subset \mathbb{R}$$

car la matrice $\widetilde{B}A\widetilde{B} \in \mathscr{S}_n(\mathbb{R})$.

□

On pose $T = S^{-1}H$ et on note $\lambda_1, \cdots, \lambda_n$ ses valeurs propres (réels par le lemme 21 ci-dessus). On a :

$$\kappa \mathrm{tr}^2\left(S^{-1}H\right) + \mathrm{tr}\left(\left(S^{-1}H\right)^2\right) = \kappa\left(\sum_{k=0}^{n} \lambda_k\right)^2 + \sum_{k=0}^{n} \lambda_k^2.$$

Comme $\kappa \geq -\frac{1}{n}$, on a :

$$\sum_{k=1}^{n} \lambda_k^2 - \frac{1}{n}\left(\sum_{k=1}^{n} \lambda_k\right)^2 \leq \sum_{k=1}^{n} \lambda_k^2 + \kappa\left(\sum_{k=1}^{n} \lambda_k\right)^2.$$

Par l'inégalité de Cauchy-Schwartz, on a bien $\sum_{k=1}^{n} \lambda_k^2 - \frac{1}{n}\left(\sum_{k=1}^{n} \lambda_k\right)^2 \geq 0$, puis

$$\mathrm{H}_g(S)(H,H) \geq 0.$$

Il s'ensuit que la fonction g est convexe sur $\mathscr{S}_n^{++}(\mathbb{R})$.

□

III.3 Conséquence

Le théorème 20 donne le corollaire suivant.

Corollaire 22. *Pour tout* $M \in \mathscr{S}_n^{++}(\mathbb{R})$*, on a :*

$$\frac{1}{\kappa}\det{}^{-\kappa}(M) \geq \frac{1}{\kappa} - \mathrm{tr}(M - I_n).$$

Démonstration. Il suffit d'appliquer le 2) de la proposition 14 à la fonction convexe g et avec $y = M$ et $x = I_n$.

□

Remarque 23. *Lorsque* $\kappa = -\frac{1}{n}$*, on obtient : pour tout* $M \in \mathscr{S}_n^{++}(\mathbb{R})$*,*

$$\frac{\mathrm{tr}(M)}{n} \geq \det{}^{1/n}(M). \tag{2}$$

C'est l'inégalité arithmético-géométrique. On sait que l'inégalité est égalité si et seulement si M est une matrice scalaire.

IV Généralisation

Dans ce paragraphe, on s'intéresse à trouver une forme quantitative de l'inégalité (2) : on cherche une fonctionnelle positive $M \mapsto R(M)$ telle que :

$$\forall M \in \mathscr{S}_n^+(\mathbb{R}), \quad \frac{\mathrm{tr}(M)}{n} - \det{}^{1/n}(M) \geq R(M)$$

avec $R(M) = 0$ si et seulement si M est une matrice scalaire.

Nous allons montrer le théorème suivant :

Théorème 24. *Pour toute matrice* $M \in \mathscr{S}_n^{++}(\mathbb{R})$*, on a :*

$$\frac{\mathrm{tr}(M)}{n} - \det{}^{1/n}(M) \geq \frac{\|M - \det^{1/n}(M) I_n\|_{\mathrm{F}}^2}{2n\|M\|_{\mathrm{F}}}.$$

Remarque 25. *Cet énoncé fournit bien une forme quantitative de* (2).

Démonstration. Le point crucial de la preuve est le résultat suivant dû à Horst Alzer :

Théorème 26. *Inégalité d'Alzer.*

Soient $x_1, \cdots, x_n$ des réels strictement positifs. Soient $p_1, \cdots, p_n$ des réels positifs tels que $\sum_{k=1}^{n} p_k = 1$.

On note $A_n = \sum_{k=1}^{n} p_k x_k$ et $G_n = \prod_{k=1}^{n} x_k^{p_k}$. Alors, on a :

$$A_n - G_n \geq \frac{1}{2\sup_{1\leq k\leq n} x_k} \sum_{k=1}^{n} p_k (x_k - G_n)^2. \tag{3}$$

Pour la preuve du résultat de Alzer, nous renvoyons à [1].

Soit $M \in \mathscr{S}_n^{++}(\mathbb{R})$. Il existe $P \in \mathscr{O}_n(\mathbb{R})$ et $D = \mathrm{diag}(\lambda_1, \cdots, \lambda_n)$ avec pour tout $i \in [\![1,n]\!]$, $\lambda_i \geq 0$ telles que $M = PD^tP$

On remarque que :

$$\begin{aligned} \|M\|_{\mathrm{F}}^2 &= \mathrm{tr}\left(PD\underbrace{{}^tPP}_{=I_n}D^tP\right) \\ &= \mathrm{tr}\left(PD^{2\,t}P\right) \\ &= \mathrm{tr}\left({}^tPPD^2\right) \\ &= \mathrm{tr}\left(D^2\right) \\ &= \sum_{k=1}^{n} \lambda_k^2. \end{aligned}$$

En appliquant (3) avec $x_k = \lambda_k$ et $p_k = \frac{1}{n}$ pour tout $k \in [\![1,n]\!]$, en notant, $\lambda_{\max} = \sup_{1\leq k\leq n} \lambda_k$ et en remarquant que $G_n = \det^{1/n}(M)$, on obtient :

$$\frac{\mathrm{tr}(M)}{n} - \det{}^{1/n}(M) \geq \frac{1}{2n\lambda_{\max}} \sum_{k=1}^{n} (\lambda_k - G_n)^2.$$

Comme la matrice $M - \det^{1/n}(M)\, I_n$ est symétrique, on a

$$\sum_{k=1}^{n} (\lambda_k - G_n)^2 = \|M - \det^{1/n}(M)\, I_n\|_{\mathrm{F}}^2,$$

on obtient donc

$$\frac{\operatorname{tr}(M)}{n} - \det^{1/n}(M) \geq \frac{1}{2n\lambda_{\max}} \|M - \det^{1/n}(M)\, I_n\|_{\mathrm{F}}^2.$$

Comme

$$\lambda_{\max} \leq \sqrt{\sum_{k=1}^{n} \lambda_k^2} = \|M\|_{\mathrm{F}},$$

on a finalement

$$\frac{\operatorname{tr}(M)}{n} - \det^{1/n}(M) \geq \frac{\|M - \det^{1/n}(M)\, I_n\|_{\mathrm{F}}^2}{2n\|M\|_{\mathrm{F}}}.$$

□

Références

[1] H. Alzer, *A new refinement of the arithmetic mean-geometric mean inequality*. Rocky Mountain J. Math. 27, no. 3, pp. 663-667, 1997.

[2] H. Cartan, *Cours de calcul différentiel*. Hermann, 1977.

[3] X. Gourdon, *Les maths en tête : Algèbre*. Ellipses, 2009.

[4] J-H. Hiriart-Urruty, *Optimisation et analyse convexe*. Presses Universitaires de France, 1998.

[5] F. Rouvière *Petit guide de calcul différentiel : À l'usage de la licence et de l'agrégation*. Cassini, 2009.