Correction : Concours blanc MP/MP*

Jeudi 6 février 2025

Fonctions de matrices symétriques, continuité et convexité

Dans ce problème, on propose de définir la notion d'image d'une matrice réelle symétrique par une fonction d'une variable réelle, puis d'étudier quelques propriétés de cette notion (en particulier, relativement à la continuité et à la convexité). Ces notions présentent un intérêt en sciences physiques (statistique ou quantique).

Notations

Dans tout le problème :
- n désigne un entier naturel non nul ;
- si p et q sont des entiers naturels, l'ensemble des entiers k tels que $p \leqslant k \leq q$ est noté $[\![p, q]\!]$;
- si i et j sont des entiers naturels, alors $\delta_{i,j} = 1$ si $i = j$ et $\delta_{i,j} = 0$ sinon ;
- B_n désigne l'ensemble des bijections de $[\![1, n]\!]$ dans lui-même ;
- I est un intervalle de $\mathbb{R}$ qui n'est ni vide ni réduit à un singleton ;
- $\mathcal{C}^0(I, \mathbb{R})$ désigne l'ensemble des fonctions continues de I dans $\mathbb{R}$;
- une fonction φ de I dans $\mathbb{R}$ est dite polynomiale s'il existe P un polynôme réel tel que, pour tout $x \in I, \varphi(x) = P(x)$;
- $\mathcal{M}_n(\mathbb{R})$ (respectivement $D_n(\mathbb{R})$, resp. $S_n(\mathbb{R})$, resp. $O_n(\mathbb{R})$), désigne l'ensemble des matrices carrées (resp. diagonales, resp. symétriques, resp. orthogonales) d'ordre n à coefficients réels, et on confond un élément de $\mathcal{M}_1(\mathbb{R})$ avec son unique coefficient ;
- on note Tr l'application trace définie sur $\mathcal{M}_n(\mathbb{R})$;
- si $M \in \mathcal{M}_n(\mathbb{R})$, on note M^T sa transposée, on note $\mathrm{Sp}(M)$ son spectre réel, et si $(i,j) \in [\![1, n]\!]^2, [M]_{i,j}$ est le coefficient de M situé à la i-ème ligne et j-ème colonne ;
- on munit $\mathcal{M}_n(\mathbb{R})$ de sa norme infinie, notée $\|\cdot\|$ et définie par :

$$\forall M \in \mathcal{M}_n(\mathbb{R}), \quad \|M\| = \max\{|[M]_{i,j}|,\ 1 \leq i,j \leq n\}$$

- $S_n(I)$ désigne l'ensemble des matrices de $S_n(\mathbb{R})$ dont le spectre réel est inclus dans I ;
- si $u = (u_i)_{1\leq i\leq n} \in \mathbb{R}^n$, on dit que ce n-uplet est croissant si pour tout $(i,j) \in [\![1, n]\!]^2$,

$$(i \leq j) \Longrightarrow (u_i \leq u_j)$$

- si $i_0 \in [\![1, n]\!]$, on appelle nombre d'occurrences de u_{i_0} dans u le cardinal de l'ensemble $\{i \in [\![1, n]\!] \,; u_i = u_{i_0}\}$
- enfin $\mathrm{Diag}\left((u_i)_{1\leq i\leq n}\right)$ désigne l'élément D de $D_n(\mathbb{R})$ tel que :

$$\forall i \in [\![1, n]\!],\ [D]_{i,i} = u_i$$

on pourra noter cet élément en extension $D = \mathrm{Diag}(u_1, \ldots, u_n)$.

Matrices de permutations

Le but de cette partie est d'étudier l'action sur les matrices diagonales de la conjugaison par des matrices de permutations. On considère l'application ω de B_n dans $\mathcal{M}_n(\mathbb{R})$ définie par :

$$\forall \sigma \in B_n, \forall (i,j) \in [\![1, n]\!]^2, [\omega(\sigma)]_{i,j} = \delta_{i,\sigma(j)}$$

1) Démontrer que pour tout $(\sigma, \sigma') \in B_n^2, \omega(\sigma \circ \sigma') = \omega(\sigma)\omega(\sigma')$

Correction : Notons u_σ l'endomorphisme canoniquement associé à $\omega(\sigma)$. On a alors $u(e_j) = e_{\sigma(j)}$ (où les e_i sont les éléments de la base canonique). Ainsi, pour $\sigma, \sigma' \in B_n$,

$$\forall j,\ u_{\sigma\circ\sigma'}(e_j) = e_{\sigma\circ\sigma'(e_j)} = u_\sigma(u_{\sigma'}(e_j)) = u_\sigma \circ u_{\sigma'}(e_j)$$

En revenant aux matrices,

$$\boxed{\forall \sigma, \sigma' \in B_n,\ \omega(\sigma \circ \sigma') = \omega(\sigma)\omega(\sigma')}$$

2) Démontrer que $\omega(B_n) \subset O_n(\mathbb{R})$.

Correction : Avec les notation précédentes, u_σ permute les éléments de la base canonique et envoie donc une base orthonormée sur une base orthonormée. $\omega(\sigma)$ représente donc une isométrie en b.o.n et est donc orthogonale.

$$\boxed{\omega(B_n) \subset O_n(\mathbb{R})}$$

3) Soit $\sigma \in B_n$ et $(d_i)_{1\leq i\leq n} \in \mathbb{R}^n$. Vérifier que :

$$\mathrm{Diag}\left((d_i)_{1\leq i\leq n}\right)\omega(\sigma) = \omega(\sigma)\,\mathrm{Diag}\left((d_{\sigma(i)})_{1\leq i\leq n}\right)$$

Correction : Multiplier à droite (resp à gauche) $\mathrm{diag}(d_1, \dots, d_n)$ revient à multiplier chaque colonne (resp ligne) par d_j. Ainsi

$$[\mathrm{diag}(d_1, \dots, d_n)\omega(\sigma)]_{i,j} = d_i\delta_{i,\sigma(j)}$$

$$[\omega(\sigma)\mathrm{diag}(d_{\sigma(1)}, \dots, d_{\sigma(n)})]_{i,j} = d_{\sigma(j)}\delta_{i,\sigma(j)} = d_i\delta_{i,\sigma(j)}$$

(pour la dernière égalité : les termes sont nuls si $i \neq \sigma(j)$ et donc égaux ; ils sont aussi égaux si $i = \sigma(j)$)
On a donc montré que

$$\boxed{\mathrm{diag}(d_1, \dots, d_n)\omega(\sigma) = \omega(\sigma)\mathrm{diag}(d_{\sigma(1)}, \dots, d_{\sigma(n)})}$$

4) En déduire l'équivalence suivante concernant deux éléments D et D' de $D_n(\mathbb{R})$,
i) D et D' ont le même ensemble de coefficients diagonaux, chacun ayant le même nombre d'occurrences dans D et D'.
ii) il existe $M \in \omega(B_n)$ telle que $D' = M^T DM$.

Correction : La propriété (i) signifie qu'il existe $\sigma \in B_n$ telle que $\forall i,\ d_i' = d_{\sigma(i)}$.
La propriété (ii) s'écrit, puisque les éléments de $\omega(B_n)$ sont des matrices orthogonales, $\omega(\alpha)D' = D\omega(\alpha)$.

Si (i) a lieu, alors (ii) aussi avec $\alpha = \sigma$ (question précédente). Lé réciproque est similaire toujours avec la question précédente (et avec $\sigma = \alpha$).

$$\boxed{\text{(i) et (ii) sont équivalentes}}$$

Fonctions de matrices symétriques

Cette partie a pour objectif de définir une correspondance entre l'espace des fonctions de I dans $\mathbb{R}$ et l'espace des fonctions de $S_n(I)$ dans $S_n(\mathbb{R})$, puis d'en démontrer quelques propriétés. Dans cette partie, f est une fonction de I dans $\mathbb{R}$.

5) Soit $S \in S_n(I)$. Justifier l'existence de $\Omega \in O_n(\mathbb{R})$ et de $(s_i)_{1\leq i\leq n} \in I^n$ tels que :

$$S = \Omega^T\,\mathrm{Diag}\left((s_i)_{1\leq i\leq n}\right)\Omega$$

Correction : Le théorème spectral indique que S est diagonalisable en base orthonormée (i.e. via une matrice de passage orthogonale). Comme l'inverse d'une matrice orthogonale est sa transposée et comme les valeurs propres de S sont supposées dans I,

$$\boxed{\forall S \in S_n(I),\ \exists \Omega \in O_n(\mathbb{R}),\ \exists (s_i) \in I^n,\ S = \Omega^\top \mathrm{diag}(s_1, \dots, s_n)\Omega}$$

6) Pour tout $(s_i)_{1\leq i\leq n} \in I^n$, justifier l'existence d'un élément P de $\mathbb{R}[X]$ tel que :

$$\forall i \in [\![1, n]\!], P(s_i) = f(s_i)$$

Correction : Notons $J \subset I$ un ensemble tel que $\{s_i,\ i \in I\} = \{s_j,\ j \in J\}$ et $\forall i, j \in J,\ i \neq j \implies s_i \neq s_j$. J est ainsi un ensemble d'indice donnant les éléments distincts parmi les s_i.
En notant $d = \mathrm{card}(J)$, l'application $P \in \mathbb{R}_{d-1}[X] \mapsto (P(s_j))_{j\in J} \in \mathbb{R}^d$ est linéaire et injective (si P de degré $\leq d-1$ admet d racines différentes, il est nul). Par dimension, c'est un isomorphisme. $(f(s_j))_{j\in J}$ admet donc un (unique) antécédent.
On a alors $P(s_i) = f(s_i)$ pour tout $i \in J$ et cela reste vrai pour les $i \in I$ par choix de J.

$$\boxed{\forall (s_i) \in I^n,\ \exists P \in\ \mathbb{R}[X],\ \forall i \in\ [\![1, n]\!],\ P(s_i) = f(s_i)}$$

On pourrait même, ce sont les formules d'interpolation de Lagrange, donner une expression explicite d'un P convenable :

$$P = \sum_{j \in J} f(s_j) L_j \quad \text{avec} \quad L_j = \prod_{i \in J \setminus \{j\}} \frac{X - s_i}{s_j - s_i}$$

Soit $S \in S_n(I)$. On suppose que l'on dispose des deux écritures :

$$S = \Omega^T \operatorname{Diag}\left((s_i)_{1 \le i \le n}\right) \Omega \text{ et } S = \Omega'^T \operatorname{Diag}\left((s'_i)_{1 \le i \le n}\right) \Omega'$$

avec $\Omega, \Omega' \in O_n(\mathbb{R})$ et $(s_i)_{1 \le i \le n}, (s'_i)_{1 \le i \le n} \in I^n$.

7) Montrer que l'on a alors :

$$\Omega'^T \operatorname{Diag}\left((f(s'_i))_{1 \le i \le n}\right) \Omega' = \Omega^T \operatorname{Diag}\left((f(s_i))_{1 \le i \le n}\right) \Omega$$

puis que $\Omega^T \operatorname{Diag}\left((f(s_i))_{1 \le i \le n}\right) \Omega \in S_n(\mathbb{R})$.

Correction : Montrons par récurrence sur k que pour toute matrice inversible Q et toute matrice M, on a

$$(Q^{-1}MQ)^k = Q^{-1}M^kQ$$

- C'est vrai pour $k = 0$.
- Si c'est vrai au rang k alors

$$(Q^{-1}MQ)^{k+1} = (Q^{-1}MQ)^k Q^{-1}MQ = Q^{-1}M^kQQ^{-1}MQ = Q^{-1}M^{k+1}Q$$

et le résultat est vrai au rang $k + 1$.

Par combinaisons linéaires, on en déduit que

$$\forall P \in \mathbb{R}[X],\ \forall Q \in GL_n(\mathbb{R}),\ \forall M \in \mathcal{M}_n(\mathbb{R}),\ P(Q^{-1}MQ) = Q^{-1}P(M)Q$$

Par ailleurs, on a aussi

$$\forall k \in \mathbb{R},\ \operatorname{diag}(d_1, \dots, d_n)^k = \operatorname{diag}(d_1^k, \dots, d_n^k)$$

et en combinant linéairement

$$\forall P \in \mathbb{R}[X],\ P(\operatorname{diag}(d_1, \dots, d_n)) = \operatorname{diag}(P(d_1), \dots, P(d_n))$$

Ici, on a donc (Ω est orthogonale et donc égale à son inverse), P étant le polynôme de la question 6,

$$\Omega^\top \operatorname{diag}(P(s_1), \dots, P(s_n))\Omega = P(S) = (\Omega')^\top \operatorname{diag}(P(s'_1), \dots, P(s'_n))\Omega'$$

De plus, $\operatorname{Sp}(S) = \{s_1, \dots, s_n\} = \{s'_1, \dots, s'_n\}$ et donc $P(s_i) = f(s_i)$ et $P(s'_i) = f(s'_i)$ pour tout i. Ainsi

$$\boxed{(\Omega')^\top \operatorname{diag}(f(s'_1), \dots, f(s'_n))\Omega' = \Omega^\top \operatorname{diag}(f(s_1), \dots, f(s_n))\Omega}$$

Comme $(AB)^\top = B^\top A^\top$, il est immédiat que

$$\boxed{\Omega^\top \operatorname{diag}(f(s_1), \dots, f(s_n))\Omega \in S_n(\mathbb{R})}$$

Dans la suite du problème, on note u l'application qui, à toute fonction φ de I dans $\mathbb{R}$, associe $u(\varphi)$ la fonction de $S_n(I)$ dans $S_n(\mathbb{R})$ définie par :

$$\forall S \in S_n(I), u(\varphi)(S) = \Omega^T \operatorname{Diag}\left((\varphi(s_i))_{1 \le i \le n}\right) \Omega$$

où $S = \Omega^T \operatorname{Diag}\left((s_i)_{1 \le i \le n}\right) \Omega$, avec $\Omega \in O_n(\mathbb{R})$ et $(s_i)_{1 \le i \le n} \in I^n$.

Cette fonction est bien définie puisque, d'après la question précédente, $u(\varphi)(S)$ ne dépend pas du choix des matrices $\Omega \in O_n(\mathbb{R})$ et $D = \operatorname{Diag}\left((s_i)_{1 \le i \le n}\right)$ avec $(s_i)_{1 \le i \le n} \in I^n$, tel que $S = \Omega^T D \Omega$.

Enfin, on désigne par v l'application $\operatorname{Tr} \circ u$.

8) Vérifier que u et v sont linéaires, puis calculer, pour toute fonction φ de I dans $\mathbb{R}$ et pour tout $x \in I, u(\varphi)(xI_n)$.

Correction : On a $\Omega^\top (D_1 + \lambda D_2)\Omega = \Omega^\top D_1 + \lambda \Omega^\top D_2 \Omega$ qui me semble donner aisément

$$u(\varphi_1 + \lambda\varphi_2)(S) = u(\varphi_1)(S) + \lambda u(\varphi_2)(S)$$

Ceci étant vrai pour tout S, $u(\varphi_1 + \lambda\varphi_2) = u(\varphi_1) + \lambda u(\varphi_2)$ et $\boxed{u \text{ est linéaire}}$.

La trace étant linéaire, par composition, $\boxed{v \text{ est linéaire}}$.

Soient $\varphi : I \to \mathbb{R}$ et $x \in I$. On a $\operatorname{diag}(x, \dots, x) = I_n^\top (xI_n) I_n$ et donc

$$u(\varphi)(xI_n) = I_n^\top \operatorname{diag}(\varphi(x), \dots, \varphi(x)) I_n = \varphi(x) I_n$$

$$\boxed{\forall x \in \mathbb{R},\ u(\varphi)(xI_n) = \varphi(x)I_n}$$

9) Étudier l'injectivité et la surjectivité de u.
Correction : Supposons que $u(\varphi) = 0$. Alors $\forall S \in S_n(\mathbb{R}),\ u(\varphi)(S) = 0$. Avec la question précédente, $\forall x \in I,\ \varphi(x) = 0$ et donc $\varphi = 0$. Ainsi $\boxed{\varphi \text{ est injective}}$ (noyau restreint au neutre).

Si $n = 1$ alors une application V de $S_n(I)$ dans $S_n(\mathbb{R})$ s'assimile à une application φ de I dans $\mathbb{R}$ ($V((x)) = (\varphi(x))$) et on a $V = u(\varphi)$. u est ainsi surjective.
Si $n \geq 2$, on peut trouver soit V l'application constante égale à $E_{1,2} + E_{2,1}$, définie sur $S_n(I)$. V est à valeurs dans $S_n(\mathbb{R})$. I étant non vide, il contient un élément x. On a $V(xI_n) = E_{1,2} + E_{2,1}$ qui n'est pas scalaire et n'est donc égale à $u(\varphi)(xI_n)$ pour aucune application φ. Ainsi, V n'a pas d'antécédent par u.

$$\boxed{u \text{ est non surjective sauf si } n = 1}$$

10) On suppose que f est polynomiale ; montrer qu'il existe $P \in \mathbb{R}[X]$ tel que pour tout $S \in S_n(I)$, $u(f)(S) = P(S)$.

Réciproquement, est-il vrai que, s'il existe $P \in \mathbb{R}[X]$ tel que pour tout $S \in S_n(I)$ $u(f)(S) = P(S)$, alors f est polynomiale ?
Correction : On suppose f polynomiale et il lui est donc associé un polynôme $P \in \mathbb{R}[X]$ tel que $\forall x \in I,\ P(x) = f(x)$.
Soit $S \in S_n(I)$. Il existe des éléments $s_i \in I$ et $\Omega \in O_n(\mathbb{R})$ telles que $S = \Omega^\top \mathrm{diag}(s_1, \ldots, s_n)\Omega$ et on a

$$u(f)(S) = \Omega^\top \mathrm{diag}(f(s_1), \ldots, f(s_n))\Omega = \Omega^\top \mathrm{diag}(P(s_1), \ldots, P(s_n))\Omega$$

Avec les remarques faites en question 7, ceci donne

$$u(f)(S) = P(\Omega^\top \mathrm{diag}(s_1, \ldots, s_n)\Omega) = P(S)$$

$$\boxed{\text{Si } f \text{ est polynomiale, il existe } P \in \mathbb{R}[X] \text{ tel que } \forall S \in S_n(I),\ u(f)(S) = P(S)}$$

On suppose, réciproquement, que f est telle qu'un tel polynôme P existe. On a en particulier

$$\forall x \in I,\ f(x)I_n = u(f)(xI_n) = P(xI_n) = P(x)I_n$$

et donc $\forall x \in I,\ f(x) = P(x)$. f est donc polynomiale et $\boxed{\text{la réciproque est vraie}}$.

11) Démontrer que, si $(\varphi_k)_{k\in\mathbb{N}}$ est une suite de fonctions de I dans $\mathbb{R}$ qui converge simplement sur I vers une fonction φ, alors les suites $(u(\varphi_k))_{k\in\mathbb{N}}$ et $(v(\varphi_k))_{k\in\mathbb{N}}$ convergent simplement sur $S_n(I)$.
Y a-t-il convergence uniforme sur $S_n(I)$ si l'on suppose que $(\varphi_k)_{k\in\mathbf{N}}$ converge uniformément sur I ?
Correction : On suppose que $\forall x \in I,\ \varphi_k(x) \to \varphi(x)$. On se donne alors $S \in S_n(I)$.Il existe des éléments $s_i \in I$ et $\Omega \in O_n(\mathbb{R})$ telles que $S = \Omega^\top \mathrm{diag}(s_1, \ldots, s_n)\Omega$ et on a

$$u(\varphi_k)(S) = \Omega^\top \mathrm{diag}(\varphi_k(s_1), \ldots, \varphi_k(s_n))\Omega = \Omega^\top D_k \Omega$$

(D_k) converge vers $D = \mathrm{diag}(\varphi(s_1), \ldots, \varphi(s_n))$ et $M \mapsto \Omega^\top M \Omega$ est continue (par exemple car elle est linéaire en dimension finie). Ainsi, $u(\varphi_k)(S)$ tend vers $u(\varphi)(S)$.

$$\boxed{\text{La convergence simple de } (\varphi_k) \text{ vers } \varphi \text{ sur } I \text{ entraîne celle de } (u(\varphi_k)) \text{ vers } u(\varphi) \text{ sur } S_n(I)}$$

La trace étant une application continue (linéaire en dimension finie)

$$\boxed{\text{La convergence simple de } (\varphi_k) \text{ vers } \varphi \text{ sur } I \text{ entraîne celle de } (v(\varphi_k)) \text{ vers } u(\varphi) \text{ sur } S_n(I)}$$

Avec les mêmes notations, on a

$$u(\varphi_k)(S) - u(\varphi)(S) = \Omega^\top \mathrm{diag}(\varphi_k(s_1) - \varphi(s_1), \ldots, \varphi_k(s_n) - \varphi(s_n))\Omega$$

Munissons $\mathcal{M}_n(\mathbb{R})$ de la norme euclidienne canonique $\|.\|_2$ définie par

$$\|M\|_2 = \sqrt{\mathrm{Tr}(M^\top M)}$$

La trace étant invariante par similitude, deux matrices orthogonalement semblables ont même norme (calcul aisé). Ainsi

$$\begin{aligned}
\|u(\varphi_k)(S) - u(\varphi)(S)\|_2^2 &= \|\mathrm{diag}(\varphi_k(s_1) - \varphi(s_1), \ldots, \varphi_k(s_n) - \varphi(s_n))\|^2 \\
&= \sum_{i=1}^{n} |\varphi_k(s_i) - \varphi(s_i)|^2 \\
&\leq n\|\varphi_k - \varphi\|_{\infty,I}^2
\end{aligned}$$

Si (φ_k) converge uniformément vers φ sur I, le majorant, qui est indépendant de S, est de limite nulle. On a ainsi convergence uniforme de $(u(\varphi_k)(S))$ vers $u(\varphi)(S)$. Il est à noter que changer de norme sur $\mathcal{M}_n(\mathbb{R})$ est indifférent puisque toutes les normes sont équivalentes en dimension finie.

$$\boxed{\text{La convergence uniforme de } (\varphi_k) \text{ vers } \varphi \text{ sur } I \text{ entraîne celle de } (u(\varphi_k)) \text{ vers } u(\varphi) \text{ sur } S_n(I)}$$

Avec des notations, similaires, on a

$$|v(\varphi_k)(S) - v(\varphi)(S)| = \left| \sum_{i=1}^{n} (\varphi_k(s_i) - \varphi(s_i)) \right| \leq n\|\varphi_k - \varphi\|_{\infty,I}$$

et là encore

$$\boxed{\text{La convergence uniforme de } (\varphi_k) \text{ vers } \varphi \text{ sur } I \text{ entraîne celle de } (v(\varphi_k)) \text{ vers } v(\varphi) \text{ sur } S_n(I)}$$

Norme et convexité

L'objectif de cette partie est de munir $S_n(\mathbb{R})$ d'une nouvelle norme qui permettra de compléter l'étude des fonctions de matrices symétriques.

12) On note $\Sigma = \left\{X \in \mathcal{M}_{n,1}(\mathbb{R}); X^T X = 1\right\}$. Démontrer que si $S \in S_n(\mathbb{R})$ on a :

$$\min(\mathrm{Sp}(S)) = \min\left\{X^T SX; X \in \Sigma\right\} \text{ et } \max(\mathrm{Sp}(S)) = \max\left\{X^T SX; X \in \Sigma\right\}$$

Correction : Soit $S \in S_n(\mathbb{R})$. Par théorème spectral, il existe une base orthonormée $(X_1, \ldots, X_n)$ de $\mathbb{R}^n$ (que l'on assimile à $\mathcal{M}_{n,1}(\mathbb{R})$) formée de vecteurs propres pour S. Notons λ_i la valeur propre associée à X_i. Le spectre de S est ainsi constitué des λ_i.
Soit $X \in \mathbb{R}^n$ que l'on décompose en $X = x_1X_1 + \cdots + x_nX_n$. On a alors (la base étant orthonormée)

$$\min(\mathrm{Sp}(S)) \sum_{i=1}^{n} x_i^2 \leq X^\top SX = \sum_{i=1}^{n} \lambda_i x_i^2 \leq \max(\mathrm{Sp}(S)) \sum_{i=1}^{n} x_i^2$$

Comme $X^\top X = x_1^2 + \cdots + x_n^2$, on a donc

$$\forall X \in \Sigma,\ \min(\mathrm{Sp}(S)) \leq X^\top SX \leq \max(\mathrm{Sp}(S))$$

Le majorant (resp minorant) est atteint pour $X = X_j$ associé à une valeur propre maximale (resp minimal) et c'est donc un maximum (resp minimum).

$$\boxed{\min(\mathrm{Sp}(S)) = \min\{X^\top SX\ ;\ X \in \Sigma\} \text{ et } \max(\mathrm{Sp}(S)) = \max\{X^\top SX\ ;\ X \in \Sigma\}}$$

13) Montrer finalement que $S_n(I)$ est une partie convexe de $S_n(\mathbb{R})$ et que l'application ρ, de $S_n(\mathbb{R})$ dans $\mathbb{R}$, qui à toute matrice $M \in S_n(\mathbb{R})$ associe

$$\max\{|\lambda|; \lambda \in \mathrm{Sp}(M)\}$$

est une norme sur $S_n(\mathbb{R})$.
Correction : Soient $A, B \in S_n(\mathbb{R})$ et $\lambda \in [0, 1]$. On a

$$\forall X \in \Sigma,\ X^\top((1-\lambda)A + \lambda B)X = (1-\lambda)X^\top AX + \lambda X^\top BX$$

Comme $\lambda \geq 0$ et $1 - \lambda \geq 0$, la question précédente donne

$$\min((1-\lambda)\mathrm{Sp}(A)) + \lambda \min(\mathrm{Sp}(B)) \leq X^\top((1-\lambda)A + \lambda B)X \leq (1-\lambda)\max(\mathrm{Sp}(A)) + \lambda \max(\mathrm{Sp}(B))$$

puis

$$\mathrm{Sp}((1-\lambda)A + \lambda B) \subset [(1-\lambda)\min(\mathrm{Sp}(A)) + \lambda \min(\mathrm{Sp}(B)), (1-\lambda)\max(\mathrm{Sp}(A)) + \lambda \max(\mathrm{Sp}(B))]$$

Comme I est un intervalle, il est convexe. Comme A, B ont un spectre inclus dans I, les bornes de l'intervalle ci-dessus sont dans I et l'intervalle est donc inclus dans I. Ainsi $(1-\lambda)A + \lambda B \in S_n(I)$. On a montré que

$$\boxed{S_n(I) \text{ est une partie convexe de } S_n(\mathbb{R})}$$

Pour montrer que ρ est une norme sur $S_n(\mathbb{R})$, on a quatre propriétés à prouver.
- ρ est immédiatement positive.

- Si $\rho(S) = 0$ alors 0 est la seule valeur propres de S et comme S est diagonalisable, elle est nulle. Ceci montre l'axiome de séparation.
- Soient $S \in S_n(\mathbb{R})$ et $\mu \in \mathbb{R}$. Les valeurs propres de μS sont les $\mu\lambda$ pour $\lambda \in \operatorname{Sp}(S)$. Comme $x \mapsto |\mu|x$ est croissante, le maximum des $|\mu||\lambda|$ est égal à $|\mu|$ fois le maximum des $|\lambda|$. Ainsi $\rho(\mu S) = |\mu|\rho(S)$ et on a l'homogénéité.
- Soient $A, B \in S_n(\mathbb{R})$. On a alors $\operatorname{Sp}(A) \subset [-\rho(A), \rho(A)]$ et idem pour B. On montre Comme plus haut que le spectre de $A+B$ est dans $[-\rho(A)-\rho(B), \rho(A)+\rho(B)]$ et donc $\rho(A+B) \leq \rho(A)+\rho(B)$. Ceci donne l'inégalité triangulaire.

$$\boxed{\rho \text{ est une norme sur } S_n(\mathbb{R})}$$

Continuité des fonctions de matrices symétriques

Dans cette partie, à l'aide de la norme précédemment introduite, on démontre quelques résultats relatifs à la continuité des fonctions de matrices symétriques. On suppose désormais $S_n(\mathbb{R})$ muni de la norme ρ et on appelle χ l'application de $S_n(\mathbb{R})$ dans $\mathbb{R}[X]$ qui, à tout élément de $S_n(\mathbb{R})$, associe son polynôme caractéristique.

On définit aussi l'application, notée $\operatorname{Sp}_\uparrow$, qui à toute matrice $S \in S_n(\mathbb{R})$, associe son spectre croissant (c'est-à-dire le n-uplet croissant des valeurs propres de S dans lequel le nombre d'occurrences de chaque valeur propre coïncide avec son ordre de multiplicité).

14) Démontrer que χ est continue.

Correction : χ va de $S_n(\mathbb{R})$ dans $\mathbb{R}_n[X]$. Les espaces étant de dimension finie, le choix de norme n'importe pas dans l'étude de continuité.

Les coefficients de $\chi(S)$ sont des fonctions polynomiales de ceux de S et donc continues (ce qui est immédiat quand on munit $S_n(\mathbb{R})$ de la norme infinie). Ainsi, les fonctions coordonnées de χ (dans la base canonique de $\mathbb{R}_n[X]$) sont continues :

$$\boxed{\chi \text{ est continue}}$$

On souhaite maintenant prouver que $\operatorname{Sp}_\uparrow$ est continue. À cet effet, on introduit un élément M de $S_n(\mathbb{R})$ et une suite $(M_k)_{k\in\mathbb{N}}$ à valeurs dans $S_n(\mathbb{R})$ qui converge vers M. Si $k \in \mathbb{N}$, on note $\Lambda_k = \operatorname{Sp}_\uparrow(M_k)$

15) Démontrer que la suite $(\Lambda_k)_{k\in\mathbb{N}}$ admet une valeur d'adhérence croissante.

Correction : On suppose que $\rho(M_k - M) \to 0$ et on a donc $\rho(M_k) \to \rho(M)$ (par seconde forme de l'inégalité triangulaire).

Ainsi, la suite $(\rho(M_k))$ est une suite bornée, disons majorée par un réel M. En notant $\Lambda_k = (\lambda_{1,k}, \ldots, \lambda_{n,k})$, on a

$$\forall k,\ \|\Lambda_k\| = \max_{1\leq i\leq n} |\lambda_{i,k}| \leq M$$

On a ainsi $(\Lambda_k)_{k\in\mathbb{N}}$ qui est bornée dans $\mathbb{R}^n$ et admet une valeur d'adhérence $\Lambda = (\lambda_1, \ldots, \lambda_n)$. En notant φ l'extractrice associé, $\lambda_{i,\varphi(k)} \to \lambda_i$ et le caractère croissant des Λ_k entraîne celui de Λ.

$$\boxed{(\Lambda_k)_{k\in\mathbb{N}} \text{ possède une valeur d'adhérence croissante}}$$

16) Montrer que, si α est une application strictement croissante de $\mathbb{N}$ dans $\mathbb{N}$ telle que la suite $\left(\Lambda_{\alpha(k)}\right)_{k\in\mathbb{N}}$ converge, alors : $\Lambda_{\alpha(k)} \underset{k\to+\infty}{\longrightarrow} \operatorname{Sp}_\uparrow(M)$.

Correction : On sait que χ est continue et donc $\chi(M_k) \to \chi(M)$. A fortiori, $\chi(M_{\alpha(k)}) \to \chi(M)$. Or, avec les notation utilisées en question précédente,

$$\chi_{M_{\alpha(k)}} = \prod_{i=1}^{n}(X - \lambda_{i,\alpha(k)})$$

Notons μ_i la limite de $\lambda_{i,\alpha(k)}$, les théorèmes d'opération donnent

$$\chi_{M_{\alpha(k)}} = \prod_{i=1}^{n}(X - \mu_i)$$

Par unicité de la limite, on a donc

$$\chi_M = \prod_{i=1}^{n}(X - \mu_i)$$

La suite (μ_i) étant croissante (par croissance des λ_k), c'est la suite croissante des valeurs propres de M :

$$\boxed{\Lambda_{\alpha(k)} \underset{k\to+\infty}{\longrightarrow} \mathrm{Sp}_{\uparrow}(M)}$$

17) En déduire que $\mathrm{Sp}_{\uparrow}$ est continue.
Correction : La suite (Λ_k) possède une valeur d'adhérence (question 15) et celle-ci est forcément $\mathrm{Sp}_{\uparrow}(M)$. On a une suite à valeurs dans un compact (ses élément appartiennent à une boule fermée, on l'a noté en question 15) qui possède une unique valeur d'adhérence et cette suite est donc convergente. On a montré que $\mathrm{Sp}_{\uparrow}(M_k) \to \mathrm{Sp}_{\uparrow}(M)$. Par caractérisation séquentielle de la limite,

$$\boxed{\mathrm{Sp}_{\uparrow} \text{ est continue}}$$

Si on utilise le résultat de cours sur les suites à valeurs dans un compact, la question 15 ne sert pas. Plus précisément, elle sert juste à justifier, dans sa preuve, que l'on a une suite bornée en dimension finie et donc à valeurs dans un compact.

18) Démontrer que $O_n(\mathbb{R})$ est une partie compacte de $\mathcal{M}_n(\mathbb{R})$
Correction : $O_n(\mathbb{R})$ est fermé comme image réciproque du fermé $\{I_n\}$ par l'application continue $M \mapsto M^\top M$.
C'est aussi une partie bornée car $\forall M \in O_n(\mathbb{R}),\ \|M\| \leq 1$ (chaque colonne est de norme 1 dans $\mathbb{R}^n$ euclidien).
Comme $\mathcal{M}_n(\mathbb{R})$ est de dimension finie,

$$\boxed{O_n(\mathbb{R}) \text{ est une partie compacte de } \mathcal{M}_n(\mathbb{R})}$$

19) Démontrer que, si $\varphi \in \mathcal{C}^0(I, \mathbb{R})$, alors $u(\varphi)$ et $v(\varphi)$ sont continues.
Correction : Soit $\varphi \in C^0(I, \mathbb{R})$. Soit (M_k) une suite d'éléments de $S_n(I)$ qui converge dans $S_n(I)$ vers une matrice M. On note encore $\Lambda_k = (\lambda_{1,k}, \ldots, \lambda_n(k))$ son spectre ordonné. Pour chaque entier k, il existe $\Omega_k \in O_n(\mathbb{R})$ telle que

$$M_k = \Omega_k^\top \mathrm{diag}(\lambda_{i,k})\Omega_k$$

On a alors

$$u_\varphi(M_k) = \Omega_k^\top \mathrm{diag}(\varphi(\lambda_{i,k}))\Omega_k$$

On vient de voir que Λ_k converge vers le spectre ordonné $\Lambda = (\lambda_1, \ldots, \lambda_n)$ de M.
Par ailleurs, comme $O_n(\mathbb{R})$ est un compact, il existe une extraite $(\Omega_{\alpha(k)})$ qui converge vers $\Omega \in O_n(\mathbb{R})$. On a alors

$$M_{\alpha(k)} \to \Omega^\top \mathrm{diag}(\varphi(\lambda_i))\Omega \quad \text{et} \quad u_\varphi(M_{\alpha(k)}) \to \Omega^\top \mathrm{diag}(\varphi(\lambda_i))\Omega$$

Par unicité de la limite, on a $M = \Omega^\top \mathrm{diag}(\varphi(\lambda_i))\Omega$ et la seconde limite vaut ainsi $u(\varphi)(M)$. On a donc

$$u_\varphi(M_{\alpha(k)}) \to u(\varphi)(M)$$

Si on considère la suite $(u(\varphi)(M_k))$, on peut en fait, en travaillant sur une extraire quelconque comme ci-dessus, montrer que $u(\varphi)(M)$ est la seule valeur d'adhérence possible. Or, toutes les suites $(\varphi(\lambda_{i,k}))_{k\in\mathbb{N}}$ étant bornée, la suite $(u(\varphi)(M_k))$ est bornée (au sens de ρ, c'est quasi immédiat et le choix de norme n'importe pas). On est dans la même situation qu'en question 17 et on peut affirmer que $(u(\varphi)(M_k))$ converge vers $u(\varphi)(M)$. $u(\varphi)$ est donc continue. Comme la trace est continue, $v(\varphi)$ est aussi continues.

$$\boxed{\text{si } \varphi \in \mathcal{C}^0(I, \mathbf{R}), \text{ alors } u(\varphi) \text{ et } v(\varphi) \text{ sont continues}}$$

Convexité des fonctions de matrices symétriques

On démontre maintenant quelques résultats relatifs à la convexité des fonctions de matrices symétriques. Dans cette partie, f est une fonction de I dans $\mathbb{R}$.

20) On suppose ici que f est convexe sur I et que $S \in S_n(I)$. On note

$$\mathcal{U}_S = \left\{\Omega^T S\Omega; \Omega \in O_n(\mathbb{R})\right\}$$

Justifier que pour tout $U \in \mathcal{U}_S$, pour tout $k \in [\![1, n]\!]$, $[U]_{k,k} \in I$.
Démontrer alors que :

$$\max\left\{\sum_{k=1}^{n} f\left([U]_{k,k}\right); U \in \mathcal{U}_S\right\} = v(f)(S)$$

Correction : Pour une matrice M quelconque, on a $[M]_{i,j} = E_j^\top M E_i$ où $E_1, \ldots, E_n$ sont les vecteurs colonnes associés à la base canonique.
Soit $\Omega \in O_n(\mathbb{R})$ et soit $U = \Omega^\top S\Omega$. On a alors

$$[U]_{j,j} = Y_j^\top S Y_j \quad \text{avec} \quad Y_j = \Omega E_j$$

Les vecteurs Y_j étant dans Σ (car $E_j \in \Sigma$ et O orthogonale), la question 12 montre que

$$\forall j,\ \min(\mathrm{Sp}(S)) \leq [U]_{j,j} \leq \max(\mathrm{Sp}(S))$$

Majorant et minorant sont dans I et I est un intervalle. Ainsi

$$\boxed{\forall j \in [\![1,n]\!],\ [U]_{j,j} \in I}$$

Notons $\lambda_1, \dots, \lambda_n$ les valeurs propres de S (comptées avec multiplicité).
Soit $\Omega \in \mathcal{U}_S$. U est orthogonalement semblable à S et il existe $\Omega \in O_n(\mathbb{R})$ telle que $U = \Omega^\top \mathrm{diag}((\lambda_i)_{1\leq i\leq n})\Omega$.
Avec les notations précédentes, on a

$$\forall j,\ f([U]_{j,j}) = f\left(Y_j^\top \mathrm{diag}((\lambda_i)_{1\leq i\leq n})Y_j\right) = f\left(\sum_{i=1}^n \lambda_i [Y_j]_i^2\right)$$

Comme $Y_j \in \Sigma$, les $[Y_j]_i^2$ sont positifs de somme 1. Par convexité de f, on a donc

$$\forall j,\ f([U]_{j,j}) \leq \sum_{i=1}^n [Y_j]_i^2 f(\lambda_i)$$

En sommant ces relations, on obtient

$$\sum_{j=1}^n f([U]_{j,j}) \leq \sum_{j=1}^n \sum_{i=1}^n [Y_j]_i^2 f(\lambda_i) = \sum_{i=1}^n \left(f(\lambda_i) \sum_{j=1}^n [Y_j]_i^2\right)$$

Les Y_j sont en fait les colonnes de Ω et $[Y_j]_i = [\Omega]_{i,j}$. Les lignes de Ω forment aussi une famille orthonormée et les sommes intérieures ci-dessus valent 1. On a donc

$$\sum_{j=1}^n f([U]_{j,j}) \leq \sum_{i=1}^n f(\lambda_i) = v(f)(S)$$

L'inégalité est une égalité quand $\Omega = I_n$ et on a donc

$$\boxed{\max\left\{\sum_{k=1}^n f([U]_{k,k});U \in \mathcal{U}_S\right\} = v(f)(S)}$$

21) En déduire que, si f est convexe sur I, pour tout $(A,B) \in S_n(I)^2$, pour tout $t \in [0,1]$, on a :

$$v(f)((1-t)A + tB) \leq (1-t)v(f)(A) + tv(f)(B)$$

Correction : Avec la question précédente, il existe une matrice $\Omega \in O_n(\mathbb{R})$ telle qu'en notant $U = \Omega^\top((1-t)A+tB)\Omega$ on ait

$$v(f)((1-t)A+tB) = \sum_{k=1}^n f([U]_{k,k})$$

Or, $[U]_{k,k} = (1-t)[\Omega^\top A\Omega]_{k,k} + t[\Omega^\top B\Omega]_{k,k}$ et ainsi

$$v(f)((1-t)A+tB) = (1-t)\sum_{k=1}^n f([\Omega^\top A\Omega]_{k,k}) + t\sum_{k=1}^n f([\Omega^\top B\Omega]_{k,k})$$

La question précédente permet de majorer les sommes par $v(f)(A)$ et $v(f)(B)$. En multipliant par t et $1-t$ on ne change pas le sens des inégalités. On trouve

$$\boxed{v(f)((1-t)A+tB) \leq (1-t)v(f)(A) + tv(f)(B)}$$

On dit qu'une fonction ψ de $S_n(I)$ dans $\mathbb{R}$ est convexe sur $S_n(I)$ si elle vérifie la relation :

$$\forall (A,B) \in S_n(I)^2, \forall t \in [0,1], \quad \psi(((1-t)A+tB) \leq (1-t)\psi(A) + t\psi(B)$$

22) Démontrer finalement que la fonction $v(f)$ est convexe sur $S_n(I)$ si, et seulement si, f est convexe sur I.
Correction : On vient de voir que la convexité de f entraîne celle de $v(f)$.
On suppose, réciproquement, que $v(f)$ est convexe. On applique cette propriété avec $A = xI_n$ et $B = yI_n$ pour $x, y \in I$. On obtient alors

$$nf((1-t)x+ty) = v(f)((1-t)xI_n + tyI_n) \leq (1-t)v(f)(xI_n) + tv(f)(yI_n) = n((1-t)f(x) + ty)$$

ce qui donne la convexité de f.

$$\boxed{f \text{ est convexe sur } I \text{ ssi } v(f) \text{ l'est sur } S_n(I)}$$