Un peu d'algèbre sur les matrices stochastiques

Éléments de correction

Exercice 1 *Carré de matrices orthogonales et matrices bistochastiques*
Soit n dans $\mathbb{N}^*$; on considère l'ensemble $D(n)$ des matrices bistochastiques, c'est-à-dire les matrices carrées réelles de taille n, à coefficients positifs (ou nuls) et dont la somme des coefficients sur chaque ligne comme sur chaque colonne vaut 1 :

$$D(n) := \left\{ M = (m_{ij})_{1\leq i,j\leq n} \mid m_{ij} \geq 0, \forall (i,j) \in [\![1;n]\!],\ \sum_{k=1}^{n} m_{kj} = 1,\ \sum_{k=1}^{n} m_{ik} = 1 \right\}.$$

On considère l'application de $\mathrm{M}_n(\mathbb{R})$ dans lui-même d'élévation au carré coefficient par coefficient :

$$\begin{array}{rccc} f: & \mathrm{M}_n(\mathbb{R}) & \to & \mathrm{M}_n(\mathbb{R}) \\ & (m_{ij})_{1\leq i,j\leq n} & \mapsto & (m_{ij}^2)_{1\leq i,j\leq n}. \end{array}$$

Une matrice dans l'image directe de $\mathrm{O}_n(\mathbb{R})$ par f est dite orthostochastique ; on note $\mathcal{O}(n) = f\left(\mathrm{O}_n(\mathbb{R})\right)$ l'ensemble de ces matrices.

1. Montrer que $\mathcal{O}(n)$ est un ensemble compact, connexe par arcs.
2. Montrer que $\mathcal{O}(n)$ et $D(n)$ coïncident pour n égal 1 ou 2. Qu'en est-il pour n plus grand ?
   Indice : on pourra par exemple considérer la matrice $D_3 := \frac{1}{2}\begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}$.
3. En déduire que $\mathcal{O}(n)$ n'est pas dense dans $D(n)$ pour n supérieur ou égal à 3.

Corrigé de l'exercice 1

1. L'application f est continue.

   Le groupe $\mathrm{O}_n(\mathbb{R})$ est fermé, comme image réciproque de la matrice nulle par l'application continue $M \mapsto {}^tMM - I_n$ de $\mathrm{M}_n(\mathbb{R})$ dans lui-même. Il est borné par 1 pour la norme infinie des coefficients (et aussi pour la norme subordonnée à la norme euclidienne canonique de $\mathbb{R}^n$, mais celle-ci est hors programme).

   Le piège est que $\mathrm{O}_n(\mathbb{R})$ n'est pas connexe (et que ce n'est pas grave). En effet, son image par le déterminant, continu, est $\{-1, 1\}$, non connexe.

   Le sous-groupe $\mathrm{SO}_n(\mathbb{R})$ est d'indice 2 dans $\mathbb{R}$. Et pour tout S dans $\mathrm{O}_n(\mathbb{R}) \setminus \mathrm{SO}_n(\mathbb{R})$, on a $\mathrm{O}_n(\mathbb{R}) = \mathrm{SO}_n(\mathbb{R}) \sqcup \mathrm{SO}_n(\mathbb{R})S$.

   Si on considère en particulier S la réflexion orthogonale dont la droite est engendrée par le premier vecteur de la base canonique, pour tout M dans $\mathrm{M}_n(\mathbb{R})$, M et MS ont même image par f (on a juste multiplié la première colonne par -1). Ainsi, $\mathcal{O}(n)$ est aussi l'image de $\mathrm{SO}_n(\mathbb{R})$ par f. Il suffit donc de démontrer que $\mathrm{SO}_n(\mathbb{R})$ est connexe par arcs.

   Soit R dans $\mathrm{O}_n(\mathbb{R})$. Par réduction des isométries en base orthonormée (au programme), il existe P dans $\mathrm{O}_n(\mathbb{R})$ telle que $PRP^{-1} = PR{}^{\mathrm{t}}P$ est diagonale par blocs, avec des blocs de la forme suivante :

   — diagonal pour la valeur propre 1 ;

— diagonal pour la valeur propre -1 ;

— de taille 2, matrice de rotation d'angle non multiple entier de π.

Comme R est de déterminant 1, le nombre de -1 dans cette réduction est pair. On peut donc les regrouper 2 par 2 et considérer chaque paire comme la matrice de la rotation d'angle π. Or, pour tout réel θ, l'application

$$\begin{array}{ccc} [0;1] & \to & \mathrm{SO}_2(\mathbb{R}) \\ t & \mapsto & \begin{pmatrix} \cos(t\theta) & -\sin(t\theta) \\ \sin(t\theta) & \cos(t\theta) \end{pmatrix} \end{array}$$

est un chemin continu de I_2 à la rotation d'angle θ.

En concaténant les blocs et conjuguant par P^{-1} (opération continue), on reste dans $\mathrm{SO}_n(\mathbb{R})$ et on obtient un chemin continu reliant I_n et R.

2. On remarque que, pour tout n dans $\mathbb{N}^*$, on a $\mathcal{O}(n) \subseteq D(n)$. Il s'agit donc d'étudier l'inclusion réciproque. Le cas $n = 1$ est trivial. Pour $n = 2$, on a

$$D(2) = \left\{ \begin{pmatrix} \alpha & 1-\alpha \\ 1-\alpha & \alpha \end{pmatrix}, \alpha \in [0;1] \right\}$$

Or

$$\begin{pmatrix} \alpha & 1-\alpha \\ 1-\alpha & \alpha \end{pmatrix} = f(R_\alpha), \quad \text{avec} \quad R_\alpha := \begin{pmatrix} \sqrt{\alpha} & \sqrt{1-\alpha} \\ -\sqrt{1-\alpha} & \sqrt{\alpha} \end{pmatrix} \in \mathrm{SO}_2(\mathbb{R}).$$

Pour $n = 3$, on considère la matrice D_3 de l'énoncé. C'est une matrice bistochastique. Par l'absurde, supposons qu'il existe une matrice orthogonale $U = (U_{ij})_{1\le i,j\le 3}$ telle que $f(U) = D_3$. Le produit scalaire entre la première et la seconde colonne de U vaut alors

$$\langle U_{:,1}, U_{:,2}\rangle = U_{11}U_{12} + U_{21}U_{22} + U_{31}U_{32} = U_{31}U_{32}$$

car $U_{11}^2 = U_{22}^2 = 0$. Comme $U_{31}^2 = U_{32}^2 = 1/2$, on a $U_{31} \neq 0$ et $U_{32} \neq 0$ de sorte que $\langle U_{:,1}, U_{:,2}\rangle \neq 0$ et U n'est donc pas orthogonale. Contradiction.

3. Si $\mathcal{O}(n)$ était dense dans $D(n)$, on aurait $\overline{\mathcal{O}(n)} = D(n)$ mais comme $\mathcal{O}(n)$ est compact, cela reviendrait à dire que $\mathcal{O}(n) = D(n)$ et on vient de montrer que ce n'est pas le cas si $n \geq 3$. Question un peu gadget, mais visiblement des gens s'étaient posé la question dans le cas complexe.

Exercice 2 *Vers le théorème de Perron–Frobenius*

Soient $n \geq 2$ un entier et M une matrice carrée de taille $n \times n$ dont tous les coefficients sont strictement positifs et dont la somme des coefficients sur chaque ligne est constante égale à 1.

1. Montrer qu'il existe un vecteur ligne $v = (v_j)_{1\le j\le n}$ tel que $vM = v$ et que le vecteur $|v| := (|v_j|)_{1\le j\le n}$ vérifie aussi $|v|M = |v|$ et a toutes ses coordonnées strictement positives.
2. Montrer qu'il existe un unique vecteur ligne $m = (m_j)_{1\le j\le n}$ vérifiant $mM = m$, dont toutes les coordonnées sont strictement positives et dont la somme vaut 1.
3. On considère maintenant la matrice M^∞ dont toutes les lignes sont égales à m. Montrer que $MM^\infty = M^\infty M = M^\infty$.

4. On définit $c := \min_{i,j} M^\infty_{ij}/M_{ij} > 0$. Montrer que pour tout $k \geq 1$, on a
$$\sup_{i,j} |M^k_{ij} - M^\infty_{ij}| \leq (1-c)^k.$$
Indice : lorsque $0 < c < 1$, on pourra considérer la matrice stochastique N définie par
$$N := \frac{1}{1-c}(M - cM^\infty), \text{ de sorte que } \quad M - M^\infty = (1-c)(N - M^\infty).$$
5. Si $M = \begin{pmatrix} 1-p & p \\ q & 1-q \end{pmatrix}$ où $0 < p < 1$ et $0 < q < 1$, déterminer le vecteur m.

Corrigé de l'exercice 2

1. Comme M est stochastique, 1 est valeur propre (associée au vecteur "vertical" $(1, \ldots, 1)^t$), donc 1 est aussi valeur propre de M^t i.e. il existe un vecteur "horizontal" v tel que $vM = v$. Considérons alors le vecteur $|v|$, par l'inégalité triangulaire, on a $\sum_i |v_i| M_{ij} \geq |v_j|$ pour tout $1 \leq j \leq n$. S'il existait une coordonnée j_0 avec $\sum_i |v_i| M_{ij_0} > |v_{j_0}|$, en faisant la somme sur les j, on aurait alors comme M est stochastique $\sum_i |v_i| = \sum_j \sum_i |v_i| M_{ij} > \sum_j |v_j|$, ce qui est absurde. On a donc bien $|v|M = |v|$ et cette dernière égalité implique que toutes les composantes de $|v|$ sont strictement positives.
2. On a un candidat tout désigné en la personne de $m := |v| / \sum_j |v_j|$. D'après ci-dessus, on a alors bien $mM = m$ et m a toutes ses coordonnées strictement positives. Supposons qu'un autre candidat m' vérifie les mêmes conditions.
   La question 1 montre que si $uM = u$ alors $u^+ = (u_i \vee 0)_{1 \leq i \leq n}$ est alors soit trivial, soit un vecteur propre (de même pour u^-. Ainsi si $u \neq 0$ vérifie $uM = u$, ses composantes sont soit toutes strictement positives soit toutes strictement négatives.
   Ainsi, le vecteur $m - m'$ est soit trivial soit un vecteur propre. Dans le dernier cas, comme ses composantes somment à $1 - 1 = 0$, il doit y avoir un changement de signe, ce qui est absurde, ainsi $m = m'$.
3. On écrit les deux produits...
4. Tout d'abord, comme M et M^∞ sont stochastiques, on a nécessairement $0 < c \leq 1$. Si $c = 1$, il n'y a plus rien à montrer. Si $0 < c < 1$, on considère la matrice stochastique N définie par
$$N := \frac{1}{1-c}(M - cM^\infty), \text{ de sorte que } \quad M - M^\infty = (1-c)(N - M^\infty).$$
   Toutes les entrées de $N - M^\infty$ sont bornées par 1 et on a aussi $NM^\infty = M^\infty N = M^\infty$ de sorte que
$$M^k - M^\infty = (M - M^\infty)^k = (1-c)^k (N - M^\infty)^k$$
   et uniformément en i et j, il vient
$$|M^k_{ij} - M^\infty_{ij}| \leq (1-c)^k.$$
5. Le spectre de M est $\mathrm{spec}(M) = \{1, \lambda := 1 - p - q\}$ et les vecteurs propres associés sont $(1,1)^t$ et $(p, -q)^t$. On a alors
$$M^k = PD^kP^{-1} = \frac{1}{p+q} \begin{pmatrix} q & p \\ q & p \end{pmatrix} + \frac{\lambda^k}{p+q} \begin{pmatrix} p & -p \\ -q & q \end{pmatrix}.$$
   Donc M^k converge à la vitesse λ^k vers $m = \left(\frac{q}{p+q}, \frac{p}{p+q}\right)$.