Concours MINES - PONTS, épreuve 1 commune PC - PSI, 2007

corrigé proposé par François DEHAME, professeur en PSI, Lycée Raspail, Paris

I. Préliminaires

1. Posons $A = MN$; alors $A(i,k) = \sum_{j=1}^{r} M(i,j)\, N(j,k)$ avec $1 \le i \le n$, $1 \le k \le m$. Pour tout $i \in [\![1,n]\!]$, on a donc

$$\begin{aligned} \sum_{k=1}^{m} |A(i,k)| &= \sum_{k=1}^{m} \Big| \sum_{j=1}^{r} M(i,j)\, N(j,k) \Big| \\ &\le \sum_{k=1}^{m} \Big(\sum_{j=1}^{r} |M(i,j)|\, |N(j,k)| \Big) = \sum_{j=1}^{r} \Big(|M(i,j)| \sum_{k=1}^{m} |N(j,k)| \Big) \\ &\le \Big(\max_{1\le j\le r} \sum_{k=1}^{m} |N(j,k)| \Big) \times \Big(\sum_{j=1}^{r} |M(i,j)| \Big) = \|N\| \Big(\sum_{j=1}^{r} |M(i,j)| \Big) . \end{aligned}$$

En prenant le max sur tous les $i \in [\![1,n]\!]$, on obtient $\|A\| = \|MN\| \le \|M\|\, \|N\|$.

Remarque. Il est immédiat que $\|\cdot\|$ est bien une norme sur l'espace vectoriel $\mathcal{M}_{n,r}(\mathbb{K})$. Si l'on identifie cet espace à $\mathcal{L}(\mathbb{K}^r, \mathbb{K}^n)$, c'est en fait la norme subordonnée aux normes $\|\cdot\|_\infty$ sur $\mathbb{K}^r$ et $\mathbb{K}^n$ (mais cela prend un peu de temps de le démontrer, et ce n'est pas au programme PC). L'inégalité $\|MN\| \le \|M\|\, \|N\|$ découle alors d'un résultat du cours de PSI.

2. Si $P \in \mathcal{M}_n(\mathbb{R})$ est stochastique, alors on a $P(i,j) \ge 0$ pour tout couple (i,j) et l'égalité $PJ_n = J_n$ se traduit par la relation

$$\forall i \in [\![1,n]\!] \qquad \sum_{j=1}^{n} P(i,j) = \sum_{j=1}^{n} |P(i,j)| = 1 .$$

Donc $\|P\| = 1$ de façon évidente.

3. Par une récurrence immédiate, pour tout $k \in \mathbb{N}^*$, la matrice P^k est (strictement) positive et $P^k\, J_n = J_n$, donc P^k est stochastique.

II. Pseudo-inverse

4. Soit A' une matrice pseudo-inverse de A, notons a' l'endomorphisme de $\mathbb{R}^n$ canoniquement associé. On a l'inclusion immédiate $\mathrm{Im}(a^2) \subset \mathrm{Im}\, a$. Par ailleurs, des propriétés (1) et (2), on tire $a = a(a'a) = a(aa') = a^2\, a'$, d'où il découle que $\mathrm{Im}(a) \subset \mathrm{Im}(a^2)$. Finalement $\mathrm{Im}(a) = \mathrm{Im}(a^2)$, donc $\mathrm{rg}(a) = \mathrm{rg}(a^2)$.

5. Soit $x \in \mathrm{Im}(a) \cap \mathrm{Ker}(a)$; alors $a(x) = 0$ et il existe $y \in \mathbb{R}^n$ tel que $x = a(y)$. Donc $a^2(y) = 0$, autrement dit $y \in \mathrm{Ker}(a^2)$. Mais, de $\mathrm{rg}(a) = \mathrm{rg}(a^2)$ et du théorème du rang, on tire $\dim \mathrm{Ker}(a) = \dim \mathrm{Ker}(a^2)$ et, puisqu'on a l'inclusion évidente $\mathrm{Ker}(a) \subset \mathrm{Ker}(a^2)$, on a donc $\mathrm{Ker}(a) = \mathrm{Ker}(a^2)$. Finalement, $y \in \mathrm{Ker}(a)$ et $x = a(y) = 0$. On a prouvé que $\mathrm{Im}(a)$ et $\mathrm{Ker}(a)$ sont en somme directe, c'est-à-dire que $\mathrm{Im}(a) \cap \mathrm{Ker}(a) = \{0\}$. En considérant les dimensions (théorème du rang toujours), on conclut qu'ils sont supplémentaires : $\mathbb{R}^n = \mathrm{Im}(a) \oplus \mathrm{Ker}(a)$.

6. Soit $\mathcal{B} = (e_1, \cdots, e_r, e_{r+1}, \cdots, e_n)$ une base de $\mathbb{R}^n$ adaptée à la décomposition $\mathbb{R}^n = \mathrm{Im}(a) \oplus \mathrm{Ker}(a)$. Puisque $\mathrm{Im}(a)$ est stable par a, la matrice de l'endomorphisme a relativement à une telle base est de la forme $\begin{pmatrix} B & 0 \\ 0 & 0 \end{pmatrix}$ avec $B \in \mathcal{M}_r(\mathbb{R})$. De plus, B est la matrice dans la base $(e_1, \cdots, e_r)$ de l'endomorphisme b de $\mathrm{Im}\, a$ induit par a, et cet endomorphisme b est un automorphisme de $\mathrm{Im}(a)$ puisque $\mathrm{Ker}(b) = \mathrm{Ker}(a) \cap \mathrm{Im}(a) = \{0\}$, donc la matrice B est inversible. En notant enfin W la matrice de passage de la base canonique de $\mathbb{R}^n$ vers la base $\mathcal{B}$, on a bien $A = W \begin{pmatrix} B & 0 \\ 0 & 0 \end{pmatrix} W^{-1}$, et $W \in \mathrm{GL}_n(\mathbb{R})$.

7. Posons $A' = W \begin{pmatrix} B^{-1} & 0 \\ 0 & 0 \end{pmatrix} W^{-1}$. Des calculs par blocs immédiats donnent $AA'A = A$, $A'AA' = A'$ et $AA' = A'A = W \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix} W^{-1}$. Donc A' est un pseudo-inverse de A.

8. L'endomorphisme a' commute avec a (propriété (1)), donc laisse stables les sous-espaces $\mathrm{Im}(a)$ et $\mathrm{Ker}(a)$. De plus, on a $a' = a'(aa') = a'(a'a) = a'^2\, a$, donc $\mathrm{Ker}(a) \subset \mathrm{Ker}(a')$: l'endomorphisme a' est nul sur $\mathrm{Ker}(a)$. La matrice de a' dans une base adaptée à la décomposition $\mathbb{R}^n = \mathrm{Im}(a) \oplus \mathrm{Ker}(a)$ est donc de la forme $\begin{pmatrix} D & 0 \\ 0 & 0 \end{pmatrix}$ avec $D \in \mathcal{M}_r(\mathbb{R})$.

9. On vérifie que $(aa')^2 = a(a'aa') = aa'$, donc aa' est un projecteur.

On a $\mathrm{Im}(aa') \subset \mathrm{Im}(a)$, mais de $a = (aa')a$, on tire l'inclusion inverse, donc $\mathrm{Im}(aa') = \mathrm{Im}(a)$.

On a $\mathrm{Ker}(a) \subset \mathrm{Ker}(a'a) = \mathrm{Ker}(aa')$ et, de $a = aa'a = a(aa')$, on tire l'inclusion inverse, donc $\mathrm{Ker}(aa') = \mathrm{Ker}(a)$.

$W^{-1}\,(AA')\,W$ est la matrice du projecteur aa' dans une base adaptée à la décomposition $\mathbb{R}^n = \mathrm{Im}(a) \oplus \mathrm{Ker}(a) = \mathrm{Im}(aa') \oplus \mathrm{Ker}(aa')$, donc est égale à $\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}$.

10. On a $W^{-1}\,(AA')\,W = \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix} = (W^{-1}AW)(W^{-1}A'W) = \begin{pmatrix} BD & 0 \\ 0 & 0 \end{pmatrix}$, d'où $D = B^{-1}$, ce qui détermine complètement A' : $A' = W \begin{pmatrix} B^{-1} & 0 \\ 0 & 0 \end{pmatrix} W^{-1}$ d'où l'unicité.

III. Calcul de $\mathbf{X}_\infty$

On commence par s'assurer que, si u est un endomorphisme de $\mathbb{R}^n$, alors u_c est bien un endomorphisme de $\mathbb{C}^n$: il faut pour cela vérifier la $\mathbb{C}$-linéarité de u_c, c'est-à-dire essentiellement le fait que, si λ est un nombre complexe et si $z = x + iy \in \mathbb{C}^n$, alors on a $u_c(\lambda z) = \lambda\, u_c(z)$. La vérification, facile, est laissée au lecteur.

On peut remarquer ensuite que u_c est l'endomorphisme de $\mathbb{C}^n$ qui, dans la base canonique de $\mathbb{C}^n$, est représenté par la matrice qui représente u dans la base canonique de $\mathbb{R}^n$.

11. Il est facile de vérifier que l'application $u \mapsto u_c$ est un morphisme de $\mathbb{R}$-algèbres de $\mathcal{L}(\mathbb{R}^n)$ vers $\mathcal{L}(\mathbb{C}^n)$, autrement dit que

$$\forall \lambda \in \mathbb{R} \quad \forall (a,b) \in \big(\mathcal{L}(\mathbb{R}^n)\big)^2 \qquad (\lambda a + b)_c = \lambda\, a_c + b_c \;;$$

$$\forall (a,b) \in \big(\mathcal{L}(\mathbb{R}^n)\big)^2 \qquad (ab)_c = a_c\, b_c \;;$$

$$(\mathrm{id}_{\mathbb{R}^n})_c = \mathrm{id}_{\mathbb{C}^n} \;.$$

Les vérifications sont des pures formalités laissées au lecteur. On en déduit en particulier $(a_c)^2 = (a^2)_c$.

12. Démontrons le lemme suivant :

Lemme : Soit u un endomorphisme d'un $\mathbb{C}$-espace vectoriel E de dimension finie n, soit λ une valeur propre de u de multiplicité m. On suppose que le sous-espace propre associé $F = \mathrm{Ker}(u - \lambda\,\mathrm{id}_E)$ est de dimension m. On a alors

$$\forall k \in \mathbb{N}^* \qquad \mathrm{Ker}(u - \lambda\,\mathrm{id}_E)^k = \mathrm{Ker}(u - \lambda\,\mathrm{id}_E) \;.$$

Démonstration du lemme : Soit $k \in \mathbb{N}^*$, posons $G = \mathrm{Ker}(u - \lambda\,\mathrm{id}_E)^k$; on a immédiatement l'inclusion $F \subset G$. Par ailleurs, G est stable par u puisque c'est le noyau d'un polynôme en u. Notons v l'endomorphisme de G induit par u. Comme $F \subset G$, il est clair que λ est valeur propre de v avec un sous-espace propre de dimension m, donc avec une multiplicité au moins égale à m. Mais on a aussi $\chi_v \mid \chi_u$, donc la multiplicité de λ en tant que valeur propre de v est m exactement. Enfin, λ est la seule valeur propre de v puisque v admet $(X-\lambda)^k$ comme polynôme annulateur. Le polynôme caractéristique de v, étant scindé, est donc nécessairement $(\lambda - X)^m$, ce qui entraîne que $\dim G = m = \dim F$, donc $G = F$, ce qu'il fallait démontrer.

Appliquons ce lemme à l'endomorphisme a_c de $E = \mathbb{C}^n$: cet endomorphisme est représenté canoniquement par la matrice $A = I_n - P$ et, comme le théorème 1 admis nous apprend que 1 est valeur propre simple de P, on a donc 0 qui est valeur propre simple de la matrice A, donc de l'endomorphisme a_c. On sait que, dans ce cas, le sous-espace propre associé $\mathrm{Ker}(a_c)$ est de dimension 1 donc, d'après le lemme ci-dessus, on a $\mathrm{Ker}(a_c^2) = \mathrm{Ker}(a_c)$. Du théorème du rang, on tire $\mathrm{rg}(a_c^2) = \mathrm{rg}(a_c) = n-1$ et, comme $\mathrm{Im}(a_c^2) \subset \mathrm{Im}(a_c)$, on a finalement $\mathrm{Im}(a_c^2) = \mathrm{Im}(a_c)$, c'est-à-dire $a_c^2(\mathbb{C}^n) = a_c(\mathbb{C}^n)$.

13. L'endomorphisme a de $\mathbb{R}^n$ étant représenté par la même matrice A, on a $\mathrm{rg}(A^2) = \mathrm{rg}(A) = n-1$, puis $\mathrm{rg}(a^2) = \mathrm{rg}(a) = n-1$. *Voir remarque à la fin du corrigé.*

14. On écrit

$$\begin{aligned}\Big(\sum_{j=0}^{k-1}(I_n - C)^j\Big)\, C &= \Big(\sum_{j=0}^{k-1}(I_n - C)^j\Big)\,\big(I_n - (I_n - C)\big)\\ &= \sum_{j=0}^{k-1}(I_n - C)^j - \sum_{j=0}^{k-1}(I_n - C)^{j+1}\\ &= (I_n - C)^0 - (I_n - C)^k = I_n - (I_n - C)^k\ ,\end{aligned}$$

ce qui donne la relation demandée.

15. De la question **6.**, on déduit que l'on peut écrire A sous la forme $A = W \begin{pmatrix} B & 0 \\ 0 & 0 \end{pmatrix} W^{-1}$, avec $W \in \mathrm{GL}_n(\mathbb{R})$ et $B \in \mathrm{GL}_{n-1}(\mathbb{R})$. On a alors $P = I_n - A = W \begin{pmatrix} I_{n-1} - B & 0 \\ 0 & 1 \end{pmatrix} W^{-1}$, donc $\displaystyle\sum_{j=0}^{k-1} P^j = W \begin{pmatrix} \displaystyle\sum_{j=0}^{k-1}(I_{n-1} - B)^j & 0 \\ 0 & k \end{pmatrix} W^{-1}$. Mais, B étant inversible, d'après **14.**, on a $\displaystyle\sum_{j=0}^{k-1}(I_{n-1} - B)^j = \big(I_{n-1} - (I_{n-1} - B)^k\big)\, B^{-1}$, donc

$$\sum_{j=0}^{k-1} P^j = W \begin{pmatrix} \big(I_{n-1} - (I_{n-1} - B)^k\big)\, B^{-1} & 0 \\ 0 & 0 \end{pmatrix} W^{-1} + k\, W \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} W^{-1}\ ,$$

et on vérifie par ailleurs que

$$I_n - AA' = W \begin{pmatrix} I_{n-1} & 0 \\ 0 & 1 \end{pmatrix} W^{-1} - W \begin{pmatrix} B & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} B^{-1} & 0 \\ 0 & 0 \end{pmatrix} W^{-1} = W \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} W^{-1}$$

et

$$(I_n - P^k)A' = W \begin{pmatrix} \left(I_{n-1} - (I_{n-1} - B)^k\right) B^{-1} & 0 \\ 0 & 0 \end{pmatrix} W^{-1} ,$$

d'où l'égalité demandée.

16. On a $\displaystyle\frac{1}{k}\sum_{j=0}^{k-1} P^j = \frac{1}{k}(I_n - P^k)\, A' + I_n - AA'$. Or, $\|I_n - P^k\| \leq \|I_n\| + \|P^k\| = 2$, donc $\displaystyle\left\|\frac{1}{k}(I_n - P^k)A'\right\| \leq \frac{2}{k}\|A'\|$ (en utilisant les questions **1.**, **2.** et **3.** du problème) et $\displaystyle\lim_{k\to+\infty} \frac{1}{k}(I_n - P^k)A' = 0$. Il reste donc $\displaystyle\lim_{k\to+\infty} \frac{1}{k}\sum_{j=0}^{k-1} P^j = I_n - AA'$.

17. Pour tout $k \in \mathbb{N}^*$, la matrice $U_k := \displaystyle\frac{1}{k}\sum_{j=0}^{k-1} P^j$ a tous ses coefficients strictement positifs donc, par passage à la limite, la matrice $I_n - AA'$ est positive (au sens large). De plus, chaque matrice P^j est stochastique donc $P^j\, J_n = J_n$ pour tout j, d'où il résulte facilement que $U_k J_n = J_n$ pour tout k, puis que $(I_n - AA')J_n = J_n$ par passage à la limite (et continuité du produit matriciel). Donc la matrice $I_n - AA'$ est stochastique. Enfin, l'égalité $(I_n - AA')A = 0$ n'est autre que la propriété (2) de la définition du pseudo-inverse.

18. Notons $L_1, \cdots, L_n$ les lignes de la matrice $I_n - AA'$. Chaque ligne L_i est stochastique (coefficients positifs ou nuls, de somme égale à 1). De plus, on a $(I_n - AA')A = 0$, c'est-à-dire $I_n - AA' = (I_n - AA')P$, ce qui signifie que chaque matrice-ligne L_i vérifie la relation $L_i = L_i P$. Du théorème 1 admis, on déduit que $L_i = X_\infty$ pour tout $i \in [\![1, n]\!]$. Donc $I_n - AA'$ est la matrice carrée dont chaque ligne coïncide avec X_∞, c'est-à-dire $I_n - AA' = J_n\, X_\infty$.

Remarque. Le résultat de la question **13.** peut facilement être obtenu sans utiliser le "complexifié" a_c introduit dans les questions **11.** et **12.** En effet, on sait (théorème 1 admis) que l'endomorphisme a de $\mathbb{R}^n$ admet 0 comme valeur propre simple, on a d'autre part l'inclusion triviale $\mathrm{Ker}(a) \subset \mathrm{Ker}(a^2)$. S'il existait un élément x dans $\mathrm{Ker}(a^2) \setminus \mathrm{Ker}(a)$, en posant $y = a(x)$, on vérifie immédiatement que la famille (x, y) est libre, que le sous-espace $F = \mathrm{Vect}(x, y)$ est stable par a, l'endomorphisme induit b ayant pour matrice $\begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$ dans la base (x, y). Cet endomorphisme b de F admet donc pour polynôme caractéristique $\chi_b(X) = X^2$, mais d'autre part $\chi_b \mid \chi_a$, ce qui est absurde puisque 0 est racine simple de χ_a. En conclusion, on a donc $\mathrm{Ker}(a) = \mathrm{Ker}(a^2)$ qui est de dimension 1 puisque 0 est valeur propre simple puis, par le théorème du rang, $\mathrm{rg}(a) = \mathrm{rg}(a^2) = n - 1$.

ANNEXE

Une démonstration du "théorème 1"

Je me propose de donner une démonstration du théorème 1 admis dans l'énoncé de ce problème. Le susdit énoncé renvoie à l'épreuve Mines-Ponts 2006, filière MP, mais en fait il ne s'agit pas tout à fait du même théorème (les hypothèses faites sur la matrice ne sont pas exactement les mêmes, on ne s'intéresse pas au caractère **strictement** positif du vecteur propre, en revanche on démontre que la valeur propre 1 est simple). La preuve que je propose est différente de celle du type Perron-Frobenius proposée dans l'épreuve MP 2006. Cette démonstration dépasse un tout petit peu le cadre du programme PSI (théorème de décomposition des noyaux), mais est lisible par un étudiant de la filière MP.

On réutilisera les questions **1.**, **2.** et **3.** du problème : si $A \in \mathcal{M}_n(\mathbb{R})$ est une matrice stochastique, on a $\|A^k\| = 1$ pour tout $k \in \mathbb{N}$. La suite $(A^k)_{k\in\mathbb{N}}$ est donc bornée.

Lemme 1 : *Soit $A \in \mathcal{M}_n(\mathbb{R}) \subset \mathcal{M}_n(\mathbb{C})$ une matrice stochastique. Alors*
- toutes ses valeurs propres complexes ont un module inférieur ou égal à 1 ;
- si $\lambda \in \mathbb{C}$ est une valeur propre de A de module 1, alors la dimension du sous-espace propre associé est égale à la multiplicité de cette valeur propre.

Démonstration du lemme 1 : notons a l'endomorphisme de $E = \mathbb{C}^n$ canoniquement associé à la matrice A, notons $\lambda_1, \cdots, \lambda_r$ les valeurs propres complexes (distinctes) de a, notons $m_1, \cdots, m_r$ leurs multiplicités respectives. Le polynôme caractéristique de a est alors $\chi_a(X) = \prod_{i=1}^{r} (\lambda_i - X)^{m_i}$. On a $\chi_a(a) = 0$ (Cayley-Hamilton), le théorème de décomposition des noyaux permet alors d'écrire $E = \bigoplus_{i=1}^{r} C_i(a)$ avec $C_i(a) = \mathrm{Ker}(a - \lambda_i\ \mathrm{id}_E)^{m_i}$ (sous-espaces caractéristiques). Chaque sous-espace $C_i(a)$ est stable par a, notons a_i l'endomorphisme induit. On a alors $(a_i - \lambda_i\ \mathrm{id}_{C_i(a)})^{m_i} = 0$, autrement dit $a_i = \lambda_i\ \mathrm{id}_{C_i(a)} + \nu_i$, où ν_i est un endomorphisme nilpotent de $C_i(a)$, d'indice de nilpotence β_i avec $\beta_i \leq m_i$.

De plus, on a $\dim C_i(a) = m_i$: en effet, l'endomorphisme a_i de $C_i(a)$ admet λ_i pour seule valeur propre puisqu'il admet $(X - \lambda_i)^{m_i}$ pour polynôme annulateur, son polynôme caractéristique (qui, sur $\mathbb{C}$, est scindé), est donc $\chi_{a_i}(X) = (\lambda_i - X)^{\dim C_i(a)}$. Comme χ_{a_i} divise χ_a, on déduit que $\dim C_i(a) \leq m_i$ et ceci pour tout $i \in [\![1, r]\!]$. Mais on a aussi $\sum_{i=1}^{r} \dim C_i(a) = \sum_{i=1}^{r} m_i = n$, ce qui entraîne $\dim C_i(a) = m_i$ pour tout i.

Dans l'espace vectoriel normé $\mathcal{L}(\mathbb{C}^n)$ (*peu importe la norme choisie, elles sont toutes équivalentes*), on sait que la suite $(a^k)_{k\in\mathbb{N}}$ doit rester bornée, ce qui entraîne facilement que, pour tout $i \in [\![1, r]\!]$, la suite $(a_i^k)_{k\in\mathbb{N}}$ doit rester bornée dans l'espace $\mathcal{L}\big(C_i(a)\big)$.

Fixons $i \in [\![1, r]\!]$, soit k un entier naturel supérieur à β_i. On a alors, par la formule du binôme,

$$a_i^k = \sum_{j=0}^{\beta_i - 1} \binom{k}{j}\ \lambda_i^{k-j}\ \nu_i^j\ .$$

Il est classique que la famille $\big(\mathrm{id}_{C_i(a)}, \nu_i, \nu_i^2, \cdots, \nu_i^{\beta_i - 1}\big)$ est libre. Les coordonnées de a_i^k

relativement à cette famille doivent alors rester bornées, autrement dit, pour tout $i \in [\![1, r]\!]$, pour tout $j \in [\![0, \beta_i - 1]\!]$, la suite numérique $\left(\binom{k}{j} \lambda_i^{k-j} \right)_{k \geq \beta_i}$ doit rester bornée. Cette condition est évidemment réalisée pour $\lambda_i = 0$, sinon on écrit

$$\binom{k}{j} \lambda_i^{k-j} = \frac{1}{j!\, \lambda_i^j}\, k(k-1)\cdots(k-j+1)\, \lambda_i^k\ .$$

• Par croissances comparées, il est clair que cette suite tend vers 0 lorsque $|\lambda_i| < 1$ et ceci pour tout j ; dans ce cas, on a $\lim\limits_{k \to +\infty} a_i^k = 0$ dans l'espace $\mathcal{L}\big(C_i(a)\big)$, ce qui servira dans la démonstration du théorème énoncé plus loin.

• Si $|\lambda_i| > 1$, il est clair que cette même suite n'est pas bornée, et ceci quel que soit j. Cela montre donc la première assertion du lemme : les valeurs propres d'une matrice stochastique sont de module au plus égal à 1.

• Si $|\lambda_i| = 1$, la suite $\left(\binom{k}{j} \lambda_i^{k-j} \right)_k$ est bornée seulement si $j = 0$. On voit donc que, si λ_i est une valeur propre de module 1 d'une matrice stochastique A, on doit avoir $\beta_i = 1$, ce qui signifie que $\nu_i = 0$ ou encore que $a_i = \lambda_i\ \mathrm{id}_{C_i(a)}$. Le sous-espace caractéristique $C_i(a)$ est alors confondu avec le sous-espace propre $\mathrm{Ker}(a - \lambda_i\ \mathrm{id}_E)$, et ce dernier a donc pour dimension la multiplicité m_i de la valeur propre λ_i.

Dans le cas des matrices stochastiques strictement positives, on peut préciser un peu plus .

Lemme 2 : *Soit $A \in \mathcal{M}_n(\mathbb{R}) \subset \mathcal{M}_n(\mathbb{C})$ une matrice stochastique strictement positive. Alors*
- le nombre 1 est valeur propre simple de A ;
- toutes les autres valeurs propres de A sont de module strictement inférieur à 1.

Démonstration du lemme 2 : On utilisera ici le théorème d'Hadamard qui dit que toute matrice à diagonale strictement dominante (i.e. toute matrice carrée $A = (a_{ij}) \in \mathcal{M}_n(\mathbb{C})$ vérifiant $\forall i \in [\![1, n]\!] \quad |a_{ii}| > \sum\limits_{j \neq i} |a_{ij}|$) est inversible, c'est un exercice très classique.

On sait déjà que toutes les valeurs propres de A ont un module inférieur ou égal à 1 (*lemme 1*). Soit λ un nombre complexe de module 1, mais différent de 1 ; pour tout $i \in [\![1, n]\!]$, puisque a_{ii} est un réel strictement positif et que $\lambda \notin \mathbb{R}_+$, l'inégalité triangulaire $|a_{ii} - \lambda| > \big||a_{ii}| - |\lambda|\big|$ est stricte, d'où

$$|a_{ii} - \lambda| > \big||a_{ii}| - |\lambda|\big| = |a_{ii} - 1| = 1 - a_{ii} = \sum_{j \neq i} a_{ij} = \sum_{j \neq i} |a_{ij}|\ .$$

La matrice $A - \lambda I_n$ est à diagonale strictement dominante, donc inversible, et λ n'est pas valeur propre de A, ce qui montre la deuxième assertion.

Le nombre 1 est valeur propre de A puisque $AJ_n = J_n$. Pour montrer que 1 est valeur propre simple, il suffit (*deuxième assertion du lemme 1*) de montrer que le sous-espace propre $\mathrm{Ker}(A - I_n)$ est de dimension 1, ou encore que la matrice $B = A - I_n$ est de rang $n - 1$. Considérons pour cela la matrice C, carrée d'ordre $n - 1$, obtenue à partir de B en supprimant la dernière ligne et la dernière colonne. Pour $i \in [\![1, n-1]\!]$, on a

$$|a_{ii} - 1| = 1 - a_{ii} = \sum_{\substack{1 \leq j \leq n \\ j \neq i}} a_{ij} = \sum_{\substack{1 \leq j \leq n-1 \\ j \neq i}} a_{ij} + a_{in} > \sum_{\substack{1 \leq j \leq n-1 \\ j \neq i}} a_{ij} = \sum_{\substack{1 \leq j \leq n-1 \\ j \neq i}} |a_{ij}|$$

puisque a_{in} est strictement positif. L'inégalité obtenue ci-dessus s'écrit $|c_{ii}| > \sum_{\substack{1 \leq j \leq n-1 \\ j \neq i}} |c_{ij}|$ et exprime donc que la matrice C est à diagonale strictement dominante, donc inversible. La matrice $B \in \mathcal{M}_n(\mathbb{R})$, admettant une matrice carrée extraite inversible d'ordre $n-1$, est de rang au moins $n-1$, donc de rang $n-1$ exactement, ce qu'il fallait démontrer.

Les matrices A et ${}^tA$ ayant même polynôme caractéristique, on en déduit que le nombre 1 est aussi valeur propre simple de ${}^tA$ et donc qu'il exsite un unique vecteur-colonne C, dont la somme des coordonnées vaut 1, tel que ${}^tAC = C$ et donc un unique vecteur-ligne $X_\infty = {}^tC$, dont la somme des coordonnées vaut 1, tel que $X_\infty\, A = X_\infty$, mais on ne sait pas encore que ses coordonnées sont positives.

On approche maintenant de la conclusion :

Théorème : *Soit $A \in \mathcal{M}_n(\mathbb{R})$ une matrice stochastique strictement positive. Alors la suite $(A^k)_{k\in\mathbb{N}}$ converge dans $\mathcal{M}_n(\mathbb{R})$, et la matrice $P = \lim_{k\to+\infty} A^k$ est stochastique et de rang 1. Ses lignes sont alors toutes égales et coïncident avec la matrice-ligne X_∞ mentionnée ci-dessus, qui a donc toutes ses coordonnées positives.*

Démonstration du théorème : Reprenons les notations de la démonstration du lemme 1. Toutefois indexons les valeurs propres $\lambda_1, \cdots, \lambda_r$ de façon que $\lambda_1 = 1$. Notons aussi u_i l'endomorphisme de $E = \mathbb{C}^n$ qui coïncide avec a (donc avec a_i) sur $C_i(a)$ et qui est nul sur $\bigoplus_{j\neq i} C_j(a)$. On a alors, dans $\mathcal{L}(E)$, $a = \sum_{i=1}^{r} u_i$ et, pour tout entier naturel non nul k, $a^k = \sum_{i=1}^{r} u_i^k$. Du lemme 2, il résulte que $m_1 = 1$ et que, pour $i \in [\![2, r]\!]$, on a $|\lambda_i| < 1$. De la démonstration du lemme 1, on déduit alors que, pour $i \geq 2$, on a $\lim_{k\to+\infty} u_i^k = 0$. Enfin, $C_1(a) = \mathrm{Ker}(a - \mathrm{id}_E)$ est une droite vectorielle et u_1 est clairement le projecteur sur cette droite suivant la direction de $\bigoplus_{i=2}^{r} C_i(a)$. Donc $u_1^k = u_1$ pour tout $k \in \mathbb{N}^*$ et $\lim_{k\to+\infty} a^k = u_1$.

En revenant à l'écriture matricielle, on a prouvé que la suite de matrices $(A^k)_{k\in\mathbb{N}}$ converge vers une matrice P de rang 1, qui de plus est une matrice de projecteur ($P^2 = P$). Enfin, pour tout k entier, la matrice A^k est stochastique, donc P est stochastique (il est immédiat que l'ensemble des matrices stochastiques est un fermé de $\mathcal{M}_n(\mathbb{R})$). Notons $L_1, \cdots, L_n$ les lignes de la matrice P ; elles sont toutes colinéaires entre elles et, sur chacune de ces lignes, la somme des coefficients vaut 1 : elles sont donc toutes égales. En passant à la limite dans la relation $A^k A = A^{k+1}$, on obtient $PA = P$, ce qui peut se lire “ligne par ligne” : pour tout $i \in [\![1, n]\!]$, $L_i A = L_i$. On a donc $L_i = X_\infty$ pour tout $i \in [\![1, n]\!]$, ce qui montre bien que les coordonnées du vecteur X_∞ sont toutes positives au sens large puisque P est stochastique.

Remarque. Le vecteur X_∞ est en fait à coordonnées strictement positives... mais je ne sais pas le démontrer simplement. Je renvoie donc au problème Mines-Ponts MP 2006, qui en donne la démonstration avec des hypothèses moins fortes sur la matrice A (A est stochastique et la matrice $(I_n + A)^{n-1}$ est strictement positive).

Ajoutons maintenant que, si $X \in \mathcal{K}_n$, alors $\lim\limits_{k\to+\infty} XA^k = XP \in \mathcal{K}_n$ (puisque $XPJ_n = XJ_n = 1 = J_1$ et que XP est positif) et $(XP)P = XP^2 = XP$, donc $XP = X_\infty$. Du théorème de Cesaro, on déduit alors la relation

$$\forall X \in \mathcal{K}_n \qquad X_\infty = \lim_{k\to+\infty} \frac{1}{k}\sum_{j=0}^{k-1} XA^j \ ,$$

qui est la dernière assertion du “théorème 1” admis en début d'énoncé... et qui n'est utilisée nulle part dans la suite du problème!

Remarque. À propos de cette convergence en moyenne (au sens de Cesaro), notons que si l'on suppose la matrice A stochastique mais non strictement positive, la suite (A^k) ne converge pas nécessairement - considérer par exemple la matrice $A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ - mais elle converge toujours en moyenne. En effet, reprenons les notations utilisées dans la démonstration du théorème ci-dessus, avec $\lambda_1 = 1$, notons $\lambda_2, \cdots, \lambda_s$ les valeurs propres de module 1 autres que 1, enfin $\lambda_{s+1}, \cdots, \lambda_r$ les valeurs propres de module strictement inférieur à 1. De la deuxième assertion du lemme 1, on déduit que u_1 est un projecteur sur $C_1(a) = \mathrm{Ker}(a - \mathrm{id}_E)$ (qui n'est plus nécessairement de rang 1) et que, pour $i \in [\![2, s]\!]$, $u_i = \lambda_i p_i$ où p_i est un projecteur sur $C_i(a) = \mathrm{Ker}(a - \lambda_i\ \mathrm{id}_E)$. On a donc $\dfrac{1}{k}\displaystyle\sum_{j=1}^{k} u_1^j = u_1$; si $i \in [\![2, s]\!]$, alors $\dfrac{1}{k}\displaystyle\sum_{j=1}^{k} u_i^j = \left(\dfrac{1}{k}\displaystyle\sum_{j=1}^{k} \lambda_i^j\right) p_i = \lambda_i\, \dfrac{1-\lambda_i^k}{k(1-\lambda_i)}\, p_i \underset{k\to+\infty}{\longrightarrow} 0$ puisque $\left|\dfrac{1-\lambda_i^k}{k(1-\lambda_i)}\right| \leq \dfrac{2}{k\,|1-\lambda_i|}$. Enfin, pour $i \in [\![s+1, r]\!]$, on a toujours $\lim\limits_{k\to+\infty} u_i^k = 0$, donc a fortiori la convergence en moyenne vers 0. Donc $\lim\limits_{k\to+\infty} \dfrac{1}{k}\displaystyle\sum_{j=1}^{k} a^j = u_1$.

En conclusion, si A est une matrice stochastique, alors

$$\lim_{k\to+\infty} \frac{1}{k}\sum_{j=1}^{k} A^j = \lim_{k\to+\infty} \frac{1}{k}\sum_{j=0}^{k-1} A^j = P\ ,$$

où la matrice P est stochastique et vérifie $P^2 = P$, mais cette matrice n'est pas nécessairement de rang 1.

Exemple. Avec $A = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$, on a immédiatement $\lim\limits_{k\to+\infty} \dfrac{1}{k}\displaystyle\sum_{j=0}^{k-1} A^j = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & 0 \\ \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & 0 & 1 \end{pmatrix}$.

Références.

• Une autre démonstration du théorème énoncé ci-dessus figure dans le problème de l'épreuve I-B de la Banque PT, session 2004.

• Toute la littérature sur les chaînes de Markov... (je citerai juste “Processus stochastiques”, de Foata et Fuchs, éditions Dunod, ISBN 2 10 006501 7, dont je me suis inspiré pour la démonstration proposée dans cet annexe).