Théorème de Perron-Frobenius et application aux mesures invariantes de chaines de Markov

Achille Méthivier

Théorème 1. *Soit $A \in \mathcal{M}_n(\mathbb{R})$ avec $n \geq 2$. On note $\rho(A) = \sup\{|\lambda| : \lambda \in \mathrm{Sp}(A)\}$. On suppose que A est à coefficients strictement positifs, alors $\rho(A)$ est l'unique valeur propre de A de module maximal et l'espace propre associé est une droite vectorielle engendrée par un vecteur à composantes réelles strictement positives.*

Lemme 2. *Si A est comme dans le théorème, alors $\rho(A) \neq 0$.*

Démonstration. Supposons $\rho(A) = 0$, alors toutes les valeurs propre de A sont nulles donc le polynôme caractéristique de A est X^n. Donc $A^n = 0$, ce qui contredit l'hypothèse que A est à coefficients strictement positifs. $\square$

Démonstration du théorème. Montrons tout d'abord que pour $y \in \mathbb{C}^{n \setminus \{0\}}$ vecteur propre de A associé à $\lambda \in \mathbb{C}$ telle que $|\lambda| = \rho(A)$, alors les composante de y sont non nulles et que $x = |y|$ est aussi vecteur propre associé à λ.
Soit $y \in \ker(A - \lambda I_n)$, on a, pour $1 \leq i \leq n$,

$$\sum_{j=1}^{n} a_{ij} y_j = \lambda y_i,$$

et

$$|\lambda||y_i| \leq \sum_{j=1}^{n} a_{ij}|y_j|.$$

Supposons que

$$\forall i \{1, \ldots, n\} \quad |\lambda||y_i| < \sum_{j=1}^{n} a_{ij}|y_j|.$$

Il existe donc $\varepsilon > 0$ tel que

$$\forall i \{1, \ldots, n\} \quad (|\lambda + \varepsilon)|y_i| < \sum_{j=1}^{n} a_{ij}|y_j|.$$

Si $x = |y|$, le vecteur défini par $x_i = |y_i|$, en notant $[Ax]_i$ la i-ème composante de Ax, on a

$$\forall i \{1, \ldots, n\} \quad (|\lambda + \varepsilon)x_i < [Ax]_i.$$

On obtient par récurrence, pour $k \geq 1$, que

$$\forall i \{1, \ldots, n\} \quad (|\lambda + \varepsilon)^k x_i < [A^k x]_i,$$

En effet, le résultat est vrai pour $k = 1$ et

$$[A^{k+1}x]_i > \sum_{j=1}^{n} [A^k x]_j > (|\lambda + \varepsilon)^k \sum_{j=1}^{n} a_{ij} x_j > (|\lambda + \varepsilon)^k [Ax]_i > (|\lambda + \varepsilon)^{k+1}.$$

On déduit alors

$$\forall k \geq 1 \quad (|\lambda + \varepsilon)^k \sum_{i=1}^{n} x_i < \sum_{i=1}^{n} [Ax]_i,$$

ce qui s'écrit aussi

$$\forall k \geq 1 \quad (|\lambda + \varepsilon)^k ||x||_1 < ||A^k x||_1 < ||A^k||_1 ||x||_1.$$

Le vecteur x étant non nul, on obtient

$$\forall k \geq 1 \quad (|\lambda + \varepsilon) < ||A^k||_1^{1/k},$$

ce qui contredit $\lim_{k \to +\infty} ||A^k||_1^{1/k} = \rho(A) = |\lambda|$.
On a montré qu'il existe $i \in \{1, \dot{,} n\}$ tel que

$$\sum_{j=1}^{n} a_{ij} x_j = |\lambda| x_i,$$

donc

$$|\lambda y_i| = \left| \sum_{j=1}^{n} a_{ij} y_j \right| = \sum_{j=1}^{n} a_{ij} |y_j|.$$

Donc y et x sont colinéaires, ce qui donne un vecteur propre réels, de coordonnées positives, associé à λ. Supposons que pour $i \in \{1, \ldots, n\}$, $x_i = 0$. On aurait $\sum a_{ij} x_j = \lambda x_i = 0$ et comme $a_{ij} > 0$, $x_j \geq 0$ on aurait $x = 0$, ce qui ne se peut. De plus, comme $\sum a_{ij} x_j = \lambda x_i$, $|\lambda|$ est un réel strictement positif.
Notons E_λ l'espace propre associé à λ et montrons qu'il est de dimension 1. Soient x, $y \in E_\lambda$, alors toutes leurs composantes sont non nulles d'après ce qui précède. Donc il existe $\alpha \in \mathbb{C}^{\setminus \{0\}}$ tel que $x_1 - \alpha y_1 = 0$ mais comme $x - \alpha y \in E_\lambda$ et admet une composante nulle, il est nul. Donc $\dim(E_\lambda) = 0$.
Il ne reste plus qu'à établir que $\lambda = \rho(A)$ est valeur propre simple de A. Supposons que la multiplicité de $|\lambda|$ soit supérieure ou égale à 2. Soit $x \in (\mathbb{R}_+^*)^n \cap E_\lambda$, alors il existe $y \in \mathbb{C}^n$, non colinéaire à x, tel que $Ay = x + \lambda y$. En effet, complétons la famille (x)

en une base B de $\mathbb{C}^n$ telle que A soit triangulaire supérieure. Dans la base B, x a pour coordonnées $(1, 0, \ldots, 0)$ et on cherche à résoudre le système suivant

$$\begin{pmatrix} \lambda & a_{12} & *\cdots & \cdots & \cdots * \\ 0 & \lambda & *\cdots & \cdots & \cdots * \\ 0 & 0 & & & \\ \vdots & \vdots & & T & \\ 0 & 0 & & & \end{pmatrix} \begin{pmatrix} y_1 \\ \vdots \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$$

Comme $\dim(E_\lambda) = 1$, $a_{12} \neq 0$ et on peut prendre le vecteur colonne $(0, a_{12}^{-1}, 0, \ldots, 0)$ qui vérifie bien les propriétés souhaitées. Comme A, x et λ sont réels, on a $A\bar{y} = x + \lambda\bar{y}$ et $A\frac{y+\bar{y}}{2} = x + \frac{y+\bar{y}}{2}$ avec $y + \bar{y} \in \mathbb{R}^n$. Il existe donc $z \in \mathbb{R}^n$ tel que $Az = x + \lambda z$. Pour α assez grand, on a $u = z + \alpha x \in (\mathbb{R}_+^*)^n$ et $Au = x + \alpha u$. Les composantes des vecteurs étant toutes positives, on a

$$||Au||_1 = ||x||_1 + \lambda||u||_1.$$

Pour $\varepsilon > 0$ tel que $||x||_1 > \varepsilon||u||_1$, on a

$$||Au||_1 > (\varepsilon + |\lambda|)||u||_1.$$

On en conclut, comme précédemment, que $\rho(A) < |\lambda|$, donc λ est de multiplicité 1. $\square$

Application 3. *Soit $(X_n)_{n\in\mathbb{N}}$ une chaîne de Markov à valeurs dans E ensemble fini. Soit P sa matrice de transition qu'on suppose ergodique, i.e. qu'il existe un entier $n \geq 1$ tel que $P^n > 0$. Alors P admet une unique mesure invariante $\mu = (\mu_1, \ldots, \mu_d)$. De plus,*

1. *les composantes de μ sont strictement positives,*
2. $$P^n \xrightarrow[n\to\infty]{} \begin{pmatrix} \mu \\ \vdots \\ \mu \end{pmatrix},$$
3. *Soit μ_0 une loi sur E, alors $\mu_0 P^n \xrightarrow[n\to\infty]{} \mu$.*

Démonstration. Dans un premier temps, supposons P strictement positive. Comme la somme des coefficients sur chaque ligne de P vaut 1 le vecteur de composantes constantes égales 1 est vecteur propre de P associé à la valeur propre 1, donc $\rho(P) \geq 1$. Montrons que $\rho(P) = 1$. Soit $\lambda \in \mathrm{Sp}(P)$, et $X = {}^t(x_1, \ldots, x_n)$ un vecteur propre de P associé à λ. Comme $X \neq 0$, on a x_i composante de X de module maximal. Quitte à diviser X par $|x_i|$, on peut supposer $|x_i| = 1$ et $|x_j| \leq 1$ pour tout $j \in \{1, \ldots, n\}$. La i-ème ligne du système $PX = \lambda X$, en prenant le module donne

$$|\lambda| = \left|\sum_{j=1}^{d} P(i,j)x_j\right| \leq \sum_{j=1}^{d} P(i,j)|x_j| \leq \sum_{j=1}^{d} P(i,j) = 1.$$

Le théorème précédent assure que, 1 est valeur propre simple de P et que son sous-espace associé est de dimension 1.
Utilisons la réduction de Jordan de la matrice P,

$$Q \begin{pmatrix} 1 & & & \\ & J_1 & & \\ & & \ddots & \\ & & & J_r \end{pmatrix} Q^{-1} \quad J_k = \begin{pmatrix} \lambda_k & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ & & & \lambda_k \end{pmatrix} \quad |\lambda_k| < 1.$$

Regardons la convergence de la suite $(P^n)_{n\in\mathbb{N}}$. Chaque bloc de Jordan s'écrit $J_k = (\lambda_k I + N_k)$ avec N_k nilpotente. Comme ces deux matrices commutent, on peut appliquer la formule du binôme de Newton, et en notant p l'indice de nilpotence de N_k, pour $n > k$

$$(\lambda_k I + N_k)^n = \sum_{i=1}^{n} \binom{n}{i} \lambda_k^{n-i} N_k^i = \lambda_k^n \sum_{i=1}^{p} \frac{n(n-1)\cdots(n-i+1)}{i!} \lambda_k^{-i} N_k^i.$$

Donc

$$\begin{aligned} ||\lambda_k I + N_k||^n &\leq |\lambda_k|^n \sum_{i=1}^{p} \frac{n(n-1)\cdots(n-i+1)}{i!} |\lambda_k|^{-i} \cdot ||N_k||^i \\ &\leq |\lambda_k|^n n^p \underbrace{\sum_{i=1}^{p} \frac{|\lambda_k|^{-i}}{i!} ||N_k||^i}_{=M \text{ indépendant de } n}. \end{aligned}$$

Comme $|\lambda_k| < 1$, $|\lambda_k|^n n^p \xrightarrow[n\to\infty]{} 0$, donc $J_k^n \to 0$ et finalement

$$P^n \xrightarrow[n\to\infty]{} P^\infty = Q \begin{pmatrix} 1 & & & \\ & 0 & & \\ & & \ddots & \\ & & & 0 \end{pmatrix} Q^{-1}.$$

L'ensemble des matrices stochastiques est fermé, donc P^∞ est stochastique de rang 1. Il existe donc $\mu = (\mu_1, \cdots, \mu_n)$ tel que

$$P^\infty = \begin{pmatrix} \mu \\ \vdots \\ \mu \end{pmatrix},$$

comme $E_1({}^tP) \subset E_1({}^tP^\infty)$ et sont de dimension 1, on a $E_1({}^tP) = E_1({}^tP^\infty)$. Comme μ vérifie

$$\mu P^\infty = \mu$$

donc ${}^t\mu \in E_1({}^tP^\infty) = E_1({}^tP)$, autrement dit, μ est l'unique loi invariante de P. D'après le théorème, on a $\mu > 0$ et pour μ_0 une loi quelconque sur E,

$$\mu_n = \mu_0 P^n \to \mu_0 P^\infty = \mu_0 \begin{pmatrix} \mu \\ \vdots \\ \mu \end{pmatrix} = \mu.$$

Il reste à démontrer le cas où P est ergodique. On dispose de $n \geq 1$ tel que $P^n > 0$ et P^n est toujours stochastique. Par ce qui précède, 1 est sa seule valeur propre de module 1 et d'après le théorème, elle est simple et $\dim(E_1({}^tP^n)) = 1$. Or 1 est valeur propre de P, donc $E_1({}^tP) \subset E_1({}^tP^n)$, et on obtient $\dim(E_1({}^tP)) = 1$ et on peut recommencer le même raisonnement avec la décomposition de Jordan et conclure de même. $\square$

I Références

1. Probabilité pour les non-probabilistes, Walter Appel (page 566)
2. Probabilité, Philippe Barbe et Michel Ledoux, (chapitre VII)
3. Analyse matricielle, Jean-Étienne Rombaldi, (page 143)