Université de Rennes 1 – Préparation
à l'agrégation de mathématiques
http://agreg-maths.univ-rennes1.fr

Auteur : Françoise Guimier
Juin 2004

Sur les matrices stochastiques.

Les matrices stochastiques (ou de transition) apparaissent dans l'étude des chaînes de Markov à espace d'états fini. Ce sont des matrices positives particulières. On peut leur attacher un graphe pondéré et donc on les retrouve aussi en théorie des graphes. On étudie ici leurs propriétés spectrales, qui permettent de démontrer un résultat de convergence de leurs itérées.

1. Rappels sur les matrices positives.

Notations. Soit $E = \{1, \ldots, d\}$. Un vecteur réel $v = (v_i)_{i \in E}$ sera dit positif (respectivement strictement positif) si $\forall i \in E$, $v_i \geq 0$ (resp. si $\forall i \in E$, $v_i > 0$). On notera $v \geq 0$ (resp. $v > 0$). De même pour une matrice réelle $A = (a_{i,j})_{(i,j) \in E^2}$. On rappelle que le spectre S de A est l'ensemble de ses valeurs propres et que son rayon spectral $\rho(A)$ est donné par $\rho(A) = \sup\{|\lambda| \mid \lambda \in S\}$. Pour $n \in \mathbb{N}$, on notera $A^n = (a_{i,j}^{(n)})$.

Soit $A = (a_{i,j})_{(i,j) \in E^2}$ une matrice réelle positive. On peut attacher à la matrice A un graphe orienté G ayant pour sommets les points i de l'ensemble $E = \{1, \ldots, d\}$ et pour arêtes les couples $(i, j) \in E^2$ tels que $a_{i,j} > 0$. On note alors $i \to j$ si $\exists n \geq 0$, $a_{i,j}^{(n)} > 0$, c'est-à-dire s'il existe un chemin $i = j_0, j_1, \ldots, j_n = j$, $a_{j_0,j_1} > 0, \ldots, a_{j_{n-1},j_n} > 0$, allant de i vers j sur le graphe. On vérifie facilement que la relation "$i \to j$ et $j \to i$" est une relation d'équivalence.

Définition 1. *On dit que A est irréductible s'il n'y a qu'une seule classe d'équivalence pour cette relation, c'est-à-dire si $\forall (i, j) \in E^2$, $\exists n \geq 0$, $a_{i,j}^{(n)} > 0$ et $\exists m \geq 0$, $a_{j,i}^{(m)} > 0$. Elle est réductible sinon.*

Exemples. Une matrice $A > 0$ est irréductible. De même, une matrice A ayant une puissance $A^n > 0$.
La matrice $\begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$ est réductible.
Il est facile de voir que A est réductible si et seulement si elle est, à une permutation des indices près, de la forme $\begin{pmatrix} B & 0 \\ C & D \end{pmatrix}$ où B et D sont des matrices carrées.

Théorème de Perron. (1907) *Soit A une matrice réelle strictement positive. Son rayon spectral $\rho(A)$ est une valeur propre simple et dominante (i.e. de module strictement supérieur à celui des autres valeurs propres). Elle admet un vecteur propre strictement positif.* [Gan]

Théorème de Frobenius. *Soit A une matrice réelle positive irréductible. Son rayon spectral $\rho(A)$ est une valeur propre simple et admet un vecteur propre strictement positif.* [Gan]

Exemple. Soit $A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$. La matrice A est positive et irréductible, les valeurs propres sont 1 et -1. C'est un exemple où $\rho(A)$ n'est pas dominante.

2. Matrices stochastiques.

Définition 2. *Une matrice stochastique (ou probabilité de transition) est une matrice positive P telle que $\forall i \in E$, $\sum\limits_{j \in E} p_{i,j} = 1$.*
Autrement dit, chaque ligne de P est un vecteur de probabilité.

Remarques.

- Le vecteur $v_0 = (1, \ldots, 1)$ est un vecteur propre de P correspondant à la valeur propre 1 et le rayon spectral de P est 1.
  En effet, soit λ une valeur propre de P de vecteur propre v et soit k tel que $|v_k| = \sup\limits_{1 \leq j \leq d} |v_j|$. On a $|\lambda v_k| = |\sum\limits_j p_{k,j} v_j| \leq \sum\limits_j p_{k,j} |v_j| \leq \sum\limits_j p_{k,j} |v_k| = |v_k|$. D'où $|\lambda| \leq 1$.

- D'après le théorème de Frobenius, si P est irréductible, 1 est une valeur propre simple de P.

Définition 3. *Une matrice stochastique est ergodique si elle admet une puissance $P^n > 0$.*

Proposition 1. *Une matrice stochastique ergodique est irréductible et admet 1 comme valeur propre simple dominante.*

En effet, soit n tel que $P^n > 0$. D'après le théorème de Perron appliqué à P^n, 1 est une valeur propre simple et dominante de P^n. Soit v un vecteur propre associé. L'espace caractéristique de P associé à la valeur propre 1 est contenu dans celui de P^n, qui est $\mathbb{R}v$, donc 1 est une valeur propre simple de P, de vecteur propre v. Si λ est une valeur propre de module 1 de P, on doit avoir $\lambda^n = 1$ et de même, l'espace caractéristique associé à λ est égal à $\mathbb{R}v$, donc $\lambda = 1$.

3. Probabilité invariante.

Définition 4. *Soit P une matrice stochastique. Une probabilité invariante pour P est un vecteur de probabilité π tel que $\pi P = \pi$.*
Autrement dit, pour la valeur propre 1, le vecteur de probabilité π est un vecteur propre gauche de P, c'est-à-dire un vecteur propre de la transposée P^t.

Théorème 1. *Soit P une matrice stochastique irréductible. Alors, P admet une unique probabilité invariante π et on a $\pi > 0$.*

Démonstration. Le théorème de Frobenius appliqué à P^t montre que l'espace propre associé à sa valeur propre 1 est engendré par un vecteur propre $v > 0$. En le normalisant, on obtient un vecteur de probabilité $\pi > 0$ et π est la seule probabilité invariante.

4. Itérées d'une matrice stochastique irréductible.

Dans la suite, P désignera une matrice stochastique irréductible.

Remarque. Supposons que (P^n) converge vers une matrice P^∞. Alors P^∞ est une matrice stochastique et puisque pour tout n, $PP^n = P^{n+1} = P^n P$, on a $PP^\infty = P^\infty = P^\infty P$. Comme 1 est une valeur propre simple de P de vecteur propre v_0, pour tout $j \in E$ il existe un réel π_j tel que $(P^\infty_{i,j})_i = \pi_j {v_0}^t$. Pour tout $(i,j) \in E^2$, $P^\infty_{i,j} = \pi_j$, donc $\pi = (\pi_j)_{j\in E}$ est un vecteur de probabilité. D'autre part, puisque $P^\infty = P^\infty P$, π est la probabilité invariante, qui est strictement positive.
Pour une matrice ergodique, les itérées convergent effectivement et on a :

Théorème 2. *Soit P une matrice stochastique ergodique et π sa probabilité invariante. Alors $\forall (i,j) \in E^2, P^{(n)}_{i,j}$ converge vers π_j et la vitesse de convergence est exponentielle, plus précisément, il existe $r \in]0,1[$ tel que $\lim_n r^{-n} \sup_{(i,j)\in E^2} |p^n_{i,j} - \pi_j| = 0$.*

Démonstration. Pour $x \in \mathbb{R}^d$ et $f \in \mathcal{L}(\mathbb{R}^d)$, on notera $||x|| = \sup_i |x_i|$ et $||f|| = \sup\{||f(x)|| \mid ||x|| = 1\}$. On rappelle que le rayon spectral $\rho(f)$ de f est alors égal à $\lim_n ||f^n||^{\frac{1}{n}}$.
On a vu que 1 est une valeur propre simple et dominante de P, de vecteur propre $v_0 = (1, \ldots, 1)$. Considérons l'application linéaire f de matrice P dans la base canonique $(e_1, \ldots, e_d)$ de $\mathbb{R}^d$. On note V un supplémentaire de $\mathbb{R}v_0$ stable par f et p la projection sur $\mathbb{R}v_0$ parallèlement à V (pour $x = av_0 + v, a \in \mathbb{R}, v \in V$, on a $p(x) = av_0$). Si g est la restriction de f à V, le rayon spectral $\rho(g)$ est < 1.
Il est facile de voir que si $x = av_0 + v$ où $a \in \mathbb{R}, v \in V$, on a $f^n(x) = p(x) + g^n(v)$ pour tout $n \in \mathbb{N}$, d'où $||f^n(x) - p(x)|| \leq ||g^n||\, ||v||$ où $||g^n|| \sim \rho(g)^n$. Fixons r tel que $\rho < r < 1$. On a alors, $\forall x \in \mathbb{R}^d$, $\lim_n r^{-n} ||f^n(x) - p(x)|| = 0$. Soit $\pi' \in \mathbb{R}^d$ tel que $\forall j$, $p(e_j) = \pi'_j v_0$. On obtient $r^{-n} \sup_{i,j} |p^{(n)}_{i,j} - \pi'_j| = r^{-n} \sup_j ||f^n(e_j) - p(e_j)|| \to 0$.
En particulier, pour tout i, j, $p^{(n)}_{i,j} \longrightarrow \pi'_j$; la remarque préalable au théorème permet d'en déduire que $\pi' = \pi$ et le résultat.

Remarques.

- On peut aussi démontrer ce résultat en utilisant une forme de Jordan de P, mais c'est plus calculatoire. Voir par exemple [Foa] ou [Ser]. (En fait, cette méthode aboutit aussi à la relation $\rho(f) = \lim_n ||f^n||^{\frac{1}{n}}$; les calculs sont cachés dans cette formule.)
- Ce résultat particularise la méthode classique des puissances pour le calcul numérique d'un vecteur propre dominant. ([Ser])

Si on supprime l'hypothèse d'ergodicité, les choses sont plus compliquées, car il peut y avoir plusieurs valeurs propres de module 1. On peut quand même voir que les itérées de P sont encore des matrices stochastiques et comme l'ensemble des matrices stochastiques est fermé et borné, il y a des sous-suites de (P^n) qui convergent et leurs limites sont stochastiques. Par exemple, si $P = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$, les sous-suites (P^{2n}) et (P^{2n+1}) convergent, mais (P^n) ne converge pas. Avec des raisonnements du même style que ci-dessus (voir [FoF]), on obtient :

Théorème 2'. *Soit P une matrice stochastique irréductible et π sa probabilité invariante. Alors $\forall (i,j) \in E^2, P_{i,j}^{(n)}$ converge en moyenne de Césaro vers π_j.*

5. Critères d'ergodicité.

Définition 5. *Soit P une matrice stochastique, $i \in E$ et $S_i = \{n \geq 0 \,|\, p_{i,i}^{(n)} > 0\}$. La période de i est $d_i = p.g.c.d\ (S_i)$. On dit que i est apériodique si $d_i = 1$.*

Proposition 2. *Si P est irréductible, tous les états $i \in E$ ont la même période.*

En effet, soit $(i,j) \in E^2$, $(n,m) \in \mathbb{N}$ tels que $p_{i,j}^{(m)} > 0$ et $p_{j,i}^{(n)} > 0$. Si $k \in S_j$, on a $p_{i,i}^{(m+k+n)} \geq p_{i,j}^{(m)} p_{j,j}^{(k)} p_{j,i}^{(n)} > 0$, donc $m+k+n \in S_i$ et d_i divise $m+k+n$. En particulier d_i divise $m+n$ et donc d_i est un diviseur commun à tous les éléments k de S_j. Par suite d_i divise d_j. Par symétrie, $d_i = d_j$.

Si par exemple, $P = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$, la matrice est irréductible mais pas ergodique, les états 1 et 2 sont de période 2, les valeurs propres sont 1 et -1, 1 n'est pas dominante.
En général, il peut y avoir plusieurs valeurs propres de module 1, qui sont les h racines h-ème de l'unité pour un certain entier h et on peut montrer que la période est h. Le cas $h = 1$ est explicité par le théorème :

Théorème 3. *Soit P une matrice stochastique irréductible. Il y a équivalence entre :*

(i) La matrice P est ergodique.

(ii) La matrice P admet 1 comme valeur propre dominante.

(iii) La suite (P^n) converge vers une matrice strictement positive.

(iv) Pour tout $i \in E$, i est apériodique.

(v) Il existe $i \in E$ qui est apériodique.

Démonstration.
On a déjà vu que $(i) \implies (ii)$ (Proposition 1).
Pour montrer $(ii) \implies (iii)$, il suffit de reprendre la démonstration du théorème 2 qui n'utilise en fait que l'hypothèse (ii).
Supposons (iii) et soit $i \in E$. La suite $(p_{i,i}^{(n)})$ converge vers une limite strictement positive. Pour n assez grand, on a donc $p_{i,i}^{(n)} > 0$ et i est donc apériodique.
L'implication $(iv) \implies (v)$ est évidente.
Supposons (v). On peut d'abord remarquer que S_i est un sous-semi-groupe de $\mathbb{N}$. En effet, supposons que n et m soient dans S_i. On a alors $p_{i,i}^{(n+m)} \geq p_{i,i}^{(n)} p_{i,i}^{(m)} > 0$, donc $n+m \in S_i$.
On va utiliser un lemme d'arithmétique, dont la démonstration est donnée plus bas :

Lemme. *Soit S un sous-semi-groupe de $\mathbb{N}$ tel que $p.g.c.d\ S = 1$. Alors, il existe $N \in S$ tel que $[N, +\infty[\subset S$.*

Soit donc N tel que $[N, +\infty[\subset S_i$. Puisque la chaîne est irréductible, pour tout $(i,j) \in E^2$, il existe des entiers m_j et n_j tels que $p_{i,j}^{(m_j)} > 0$ et $p_{j,i}^{(n_j)} > 0$. Pour tout $(j,k) \in E^2$, si $n = n_j + n' + m_k$ où $n' \geq N$, on a $p_{j,k}^{(n)} \geq p_{j,i}^{(n_j)} p_{i,i}^{(n')} p_{i,k}^{(m_k)} > 0$. Soit $N' = \sup\{n_j + N + m_k | \, (j,k) \in E^2\}$. Comme **E est fini**, N' est fini et on obtient $P^{N'} > 0$ et (i).

Démonstration du lemme. On peut trouver $k \in \mathbb{N}$ et des éléments $s_1, \dots s_k \in S$ tels que p.g.c.d $(s_1, \dots, s_k) = 1$. Puis, en utilisant l'identité de Bezout et quitte à faire une permutation des indices, on peut trouver $p, 1 \leq p \leq k$ et des entiers positifs $n_1, \dots n_k$ tels que $\sum\limits_{1 \leq j \leq p} n_j s_j - \sum\limits_{p < j \leq k} n_j s_j = 1$. Les deux sommes sont dans S. On trouve donc $a, b \in S$, tels que $a - b = 1$. Si $b = 0$, on obtient $S = \mathbb{N}$. Sinon, pour $n \in \mathbb{N}$ tel que $n \geq b^2$ on peut écrire $n = bq + r, 0 \leq r < b \leq q$, soit $n = bq + r(a - b) = ra + (q - r)b$, donc $n \in S$. D'où $[b^2, +\infty[\subset S$.

Bibliographie.

[FoF] **D.Foata, A. Fuchs :** *Processus stochastiques. Processus de Poisson, chaînes de Markov et martingales. Cours et exercices corrigés.* Dunod, Paris, 2002.

[Gan]**F. R. Gantmacher :** *Théorie des matrices. Tome 2. Questions spéciales et applications.* Paris : Dunod , 1966. [Agusu et 65]

[Ser] **D. Serre.** *Les matrices. Théorie et pratique - Paris. Dunod. 2001*