DM PRB201 2019

Exercice 1. [Calculs de distance pour une chaîne de Markov]

Soit $a \in \mathbb{R}$ avec $0 < a < 1$, et soit une chaîne de Markov $(X_t)_{t\geq 0}$ sur l'espace d'état $\Omega = \{0, 1, \ldots, n-1\}$ de matrice de transition P telle que :

$$P(i,j) = a\,1_{\{j=i+1 \text{ mod } n\}} + (1-a)\,\frac{1}{n} \qquad\qquad i,j \in \Omega$$

1. Montrer que la chaîne de Markov $(X_t)_{t\geq 0}$ possède une unique loi stationnaire π et la calculer. La chaîne est-elle réversible ?
2. Montrer par une récurrence que, pour tout $(i,j) \in \Omega^2$, pour tout $t \in \mathbb{N}$
$$P^t(i,j) = a^t 1_{\{j=i+t \text{ mod } n\}} + (1-a^t)\frac{1}{n}.$$
Reconnaître la mesure de probabilité limite $\lim_{t\to\infty} P^t(i,\cdot)$, et retrouver le résultat de la question 1.
3. Montrer que, pour tout $i \in \Omega$, pour tout $t \in \mathbb{N}$, la distance en variation totale satisfait :
$$\frac{1}{2}\sum_k |P^t(i,k) - \pi(k)| = \left(1 - \frac{1}{n}\right) a^t$$
4. Montrer que, pour tout $i, j \in \Omega$ tel que $i \neq j$, pour tout $t \in \mathbb{N}$
$$\frac{1}{2}\sum_k |P^t(i,k) - P^t(j,k)| = a^t$$
5. Montrer que, pour tout $i \in \Omega$, pour tout $t \in \mathbb{N}$
$$\sum_{j\in\Omega} \frac{(P^t(i,j) - \pi(j))^2}{\pi(j)} = (n-1)\,a^{2t}$$

Correction. *La mesure stationnaire dans la première question est la mesure uniforme (on pouvait regarder un peu plus loin dans l'énoncé pour le deviner sans faire de calcul). Puisque la chaine est irréductible, la mesure stationnaire est unique et il suffit donc de vérifier que la mesure uniforme fonctionne, ce qui est évident :*

$$(\pi P)(j) = \sum_{i=0}^{n-1} \pi(i)P(i,j) = \sum_{i=0}^{n-1} \pi(i)(1-a)\frac{1}{n} + \pi(i-1 \ mod \ n)a = \frac{1-a}{n} + \frac{a}{n} = \frac{1}{n} = \pi(j)$$

La chaîne en revanche n'est pas réversible, puisque pour tout i,

$$\pi(i)P(i,i+1) \neq \pi(i+1)P(i+1,i)$$

(avec l'addition modulo n). La formule sur P^t se montre par récurrence : elle vaut pour $t = 0$ et $t = 1$ et si on la suppose vraie en t, on voit que

$$P^{t+1}(i,j) = \sum_k P^t(i,k)P(k,j) = \sum_k \left(a^t 1_{\{k=i+t \ mod \ n\}} + (1-a^t)\frac{1}{n}\right)\left(a 1_{\{j=k+1 \ mod \ n\}} + (1-a)\frac{1}{n}\right)$$

$$= a^{t+1} 1_{\{j=i+t+1 \ mod \ n\}} + \frac{1}{n}\left((1-a^t)(1-a) + a(1-a^t) + (1-a)a^t\right)$$

$$= a^{t+1} 1_{\{j=i+t+1 \ mod \ n\}} + \frac{1}{n}(1-a^{t+1})$$

La mesure de probabilité limite est la mesure uniforme, et c'est donc, par le théorème fondamental de conergence, l'unique mesure stationnaire. Ceci redémontre le résultat de la première question. On peut commencer par noter que si π et ν sont deux mesures de probabilité :

$$\sum_{x\in\Omega} \pi(x) - \nu(x) = 0$$

implique

$$\sum_{x\in\Omega:\pi(x)>\nu(x)} \pi(x) - \nu(x) = \sum_{x\in\Omega:\nu(x)>\pi(x)} \nu(x) - \pi(x)$$

et partant :

$$\frac{1}{2}\sum_{x\in\Omega} |\pi(x) - \nu(x)| = \sum_{x\in\Omega:\pi(x)>\nu(x)} \pi(x) - \nu(x).$$

(Cette remarque n'est en rien obligatoire cependant). On obtient :

$$\frac{1}{2}\sum_{j=0}^{n-1} |P^t(i,j) - \pi(j)| = \sum_{j,P^t(i,j)>\pi(j)} P^t(i,j) - \pi(j) = a^t + (1-a^t)\frac{1}{n} - \frac{1}{n} = (1-\frac{1}{n})a^t$$

car seul l'entier $j = i+t$ vérifie la condition. De même,

$$\frac{1}{2}\sum_{k=0}^{n-1} |P^t(i,k) - P^t(j,k)| = \sum_{k,P^t(i,k)>P^t(j,k)} P^t(i,j) - P^t(j,k) = a^t + (1-a^t)\frac{1}{n} - (1-a^t)\frac{1}{n} = a^t$$

en retenant à nouveau le seul entier $k = i+t$. Enfin, pour le dernier calcul,

$$\sum_j \frac{(P^t(i,j) - \pi(j))^2}{\pi(j)} = \sum_j \frac{(a^t\, 1_{\{j=i+t \ mod \ n\}} + (1-a^t)\, 1/n - 1/n)^2}{1/n}$$

$$= \sum_j \frac{a^{2t}(1_{\{j=i+t \ mod \ n\}} - 1/n)^2}{1/n}$$

$$= na^{2t}\sum_j (1_{\{j=i+t \ mod \ n\}} - 1/n)^2$$

$$= na^{2t}[(n-1)1/n^2 + (1-1/n)^2]$$

$$= na^{2t}(1-1/n) = (n-1)a^{2t}$$

Exercice 2. [Critère de Kolmogorov].

Soit P matrice de transition irréductible sur Ω, et π son unique mesure de probabilité stationnaire. On note R la propriété : "Pour tout $t \geq 2$, pour tout $x_1, x_2, \ldots, x_t \in \Omega$ tels que $x_1 = x_n$,

$$\prod_{1 \leq i \leq t-1} P(x_i, x_{i+1}) = \prod_{1 \leq i \leq t-1} P(x_{i+1}, x_i)."$$

1. Supposons P réversible par rapport à π. Montrer que R est satisfaite.

Réciproquement, on suppose R satisfaite, et on souhaite montrer que P est réversible par rapport à π.

2. Montrer que, pour $t \geq 1$, et $x, y \in \Omega$:

$$P^t(x,y)P(y,x) = P(x,y)P^t(y,x)$$

3. Dans le cas où P est apériodique, en déduire que P est réversible par rapport à π.
4. On revient maintenant au cas où P est seulement supposée irréductible. En s'aidant de la conclusion de l'exercice 11 du TD1, trouver la limite de $\frac{\sum_{t=1}^{n} P^t(x,y)}{n}$ quand $n \to \infty$, puis conclure.

Correction. *P étant réversible par rapport à π, on vérifie tout d'abord par récurrence sur l'entier n que, quelque soit $x_1, x_2, \ldots, x_n$,*

$$\pi(x_1) \prod_{1 \leq i \leq n-1} P(x_i, x_{i+1}) = \pi(x_n) \prod_{1 \leq i \leq n-1} P(x_{i+1}, x_i).$$

cette récurrence facile ne pose pas de difficulté. Pour obtenir la propriété R il suffit ensuite de choisir $x_1 = x_n$ puis d'observer que $\pi(x_1) \neq 0$ (puisque la chaine est irréductible), ce qui permet de simplifier par $\pi(x_1)$.

Pour la réciproque, on procède comme suit. En sommant la relation de la propriété R sur $x_3, x_4, \ldots, x_{n-1}$ (de sorte que seuls $x_1 = x_n$ et x_2 sont fixés) on obtient par définition du produit matriciel que

$$P(x_1, x_2)P^n(x_2, x_1) = P^n(x_1, x_2)P(x_2, x_1).$$

Si P est de plus apériodique (et puisqu'elle est par hypothèse irréductible), $P^n(x,y) \to \pi(y)$ et donc en passant a la limite en n dans la relation précédente, on obtient que

$$P(x_1, x_2)\pi(x_1) = \pi(x_2)P(x_2, x_1),$$

soit la définition de la réversibilité. Dans le cas général où la chaîne n'est pas nécessairement apériodique, l'exercice 11 du TD1 assure que la somme de Césaro $\sum_{1 \leq m \leq n} P^m(x,y)/n$ converge vers (l'unique, P reste irréductible) mesure stationnaire $\pi(y)$. Il suffit alors de passer à la limite dans la relation

$$P(x_1, x_2)\frac{\sum_{1 \leq m \leq n} P^m(x_2, x_1)}{n} = \frac{\sum_{1 \leq m \leq n} P^m(x_1, x_2)}{n} P(x_2, x_1).$$

Cette formulation (appelée critère de Kolmogorov) présente l'avantage (tout théorique) de pouvoir vérifier la réversibilité sans connaître la mesure stationnaire π ; en revanche, cette condition implique les cycles de toute taille n ...

Exercice 3. [Décomposition de première entrée]

Soit P matrice stochastique sur Ω, et $(X_t)_{t\geq 0}$ la chaîne de Markov sur Ω de matrice de matrice de transition P.

1. Soit $y \in \Omega$, $n \in \mathbb{N}$. Posons $\tau_y^{(n)} = \min\{s : s \geq n, X_s = y\}$ le premier temps de passage en y après l'instant n. Observer que $\tau_y^{(n)}$ est un temps d'arrêt.
2. Montrer que, pour tout $t \geq n$,
$$P^t(x,y) = \sum_{n\leq s\leq t} \mathbb{P}_x(\tau_y^{(n)} = s)P^{t-s}(y,y), \quad x, y \in \Omega$$
3. En déduire :
$$\sum_{t=0}^{m} P^t(x,x) \geq \sum_{t=n}^{m+n} P^t(x,x).$$
4. Par un raisonnement probabiliste similaire (c'est-à-dire en introduisant un temps d'arrêt bien choisi), montrer que :
$$\sum_{t=0}^{m} P^t(x,x) \geq \sum_{t=0}^{m} P^t(y,x)$$

Commentaire : on pourra considérer la réécriture purement matricielle, c'est-à-dire sans introduire la chaîne de Markov associée, des preuves des deux inégalités obtenues en 3 et 4, afin de saisir l'intérêt du formalisme probabiliste.

Correction. *Notons* $\tau_y^{(\geq t)} = \min\{s \geq t, X_s = y\}$ *le premier temps d'atteinte de* y *après l'instant* t. *Il s'agit encore d'un temps d'arrêt :*

$$\{\tau_y^{(\geq t)} \leq s\} = \begin{cases} \emptyset & \text{si } s < t \\ \bigcup_{r=t}^{s}\{X_r = y\} & \text{si } s \geq t \end{cases}$$

Pour $t \geq n$, *on a, en notant que :*

$$\{X_t = y\} = \{X_t = y, n \leq \tau_y^{(n)} \leq t\} = \bigcup_{s=n}^{t}\{X_t = y, \tau_y^{(n)} = s\}$$

où la dernière réunion est disjointe :

$$\begin{aligned}
P^t(x,y) &= \mathbb{P}_x(X_t = y) \\
&= \sum_{n\leq s\leq t} \mathbb{P}_x(\tau_y^{(n)} = s, X_t = y) \\
&= \sum_{n\leq s\leq t} \mathbb{P}_x(\tau_y^{(n)} = s, X_{s+(t-s)} = y) \\
&= \sum_{n\leq s\leq t} \mathbb{P}_x(\tau_y^{(n)} = s, \theta_s(X)_{t-s} = y) \\
&= \sum_{n\leq s\leq t} \mathbb{P}_x(\tau_y^{(n)} = s)\mathbb{P}_x(X_{t-s} = y) \text{ de la propriété de Markov} \\
&= \sum_{n\leq s\leq t} \mathbb{P}_x(\tau_y^{(n)} = s)P^{t-s}(y,y)
\end{aligned}$$

Ceci permet d'écrire en appliquan cette identité en $x = y$, et en sommant de $t = n$ à $n + m$:

$$\begin{aligned}
\sum_{n \leq t \leq m+n} P^t(x,x) &= \sum_{n \leq s \leq t \leq m+n} \mathbb{P}_x(\tau_x^{(n)} = s) P^{t-s}(x,x) \\
&= \sum_{n \leq s \leq m+n} \mathbb{P}_x(\tau_x^{(n)} = s) \sum_{s \leq t \leq m+n} P^{t-s}(x,x) \\
&= \sum_{n \leq s \leq m+n} \mathbb{P}_x(\tau_x^{(n)} = s) \sum_{0 \leq r \leq m+n-s} P^r(x,x) \ \textit{posant } r = t - s \\
&\leq \sum_{n \leq s \leq m+n} \mathbb{P}_x(\tau_x^{(n)} = s) \sum_{0 \leq r \leq m} P^r(x,x) \ \textit{car } m + n - s \leq m \\
&= \sum_{0 \leq r \leq m} P^r(x,x) \mathbb{P}_x(n \leq \tau_x^{(n)} \leq n + m) \\
&\leq \sum_{0 \leq r \leq m} P^r(x,x)
\end{aligned}$$