Chaînes de Markov conditionnées

Emmanuel Lesigne,
Travaux en cours avec Marc Peigné,
Rencontres en l'honneur d'Émile Le Page, septembre 2011.

Cadre général

E est un espace dénombrable, et $B \subset E$.

$P = (p(i,j))_{i,j \in E}$ est une matrice de probabilité de transition sur E.

On note $(Z_m)_{m \geq 0}$ une chaîne de Markov associée aux données (E,P), et définie sur une espace probabilisé $(\Omega, \mathscr{T}, \mathbb{P})$.

On suppose que : $\forall i \in B,\ \exists j \in B,\ p(i,j) > 0$.

Fait

Soit $m \in \mathbb{N}$. Sous la probabilité $\mathbb{P}$ conditionnée par l'événement $(Z_\ell \in B,\ 0 \leq \ell \leq n)$, la suite de variables aléatoires $(Z_m)_{m \geq 0}$ est une chaîne de Markov inhomogène dont les probabilités de transition $q(\cdot,\cdot,n,m)$ sont données par :

$$q(i,j,m,n) = \mathbb{P}(Z_{m+1} = j \mid Z_m = i, Z_{m+1} \in B, Z_{m+2} \in B, \ldots, Z_n \in B)$$

si $0 \leq m < n$ et $i,j \in B$;

$$q(i,j,m,n) = p(i,j) \quad \text{si } (m = n, i \in B \text{ et } j \in E) \text{ ou } (m > n \text{ et } i,j \in E).$$

Notations.

$$P_B := (p(i,j))_{i,j \in B} \ , \qquad P_B^n = (p_B^{(n)}(i,j))_{i,j \in B} \ , \ (n \in \mathbb{N}).$$

La somme des entrées de la ligne i de la matrice P_B^n est notée $L(i,n)$:

$$L(i,n) := \sum_{j \in B} p_B^{(n)}(i,j).$$

Proposition.

Soient $i,j \in B$. On a, pour $m < n$,

$$q(i,j,m,n) = q(i,j,0,n-m)$$

et

$$q(i,j,0,n) = p(i,j)\frac{L(j,n-1)}{L(i,n)} = p(i,j)\frac{L(j,n-1)}{\sum_{k \in B} p(i,k)L(k,n-1)}.$$

Deux questions

(Q1) $q(i,j) := \lim_{n\to\infty} q(i,j,0,n)$ existe ?

(Q2) $(q(i,j))_{i,j\in B}$ est une matrice de probabilité de transition ?

Exemple.
La marche aléatoires simple sur $\mathbb{Z}$ conditionnée à rester dans $B = \{1,2,3,4\}$.

$$P_B = \frac{1}{2}\begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix} \quad \text{et} \quad (q(i,j)) = \begin{pmatrix} 0 & 1 & 0 & 0 \\ \frac{1}{\theta^2} & 0 & \frac{1}{\theta} & 0 \\ 0 & \frac{1}{\theta} & 0 & \frac{1}{\theta^2} \\ 0 & 0 & 1 & 0 \end{pmatrix}.$$

Théorème de Perron-Frobenius

Supposons B fini : $B = \{1, 2, \dots, d\}$.
On pose

$$L(n) := \begin{pmatrix} L(1,n) \\ L(2,n) \\ \vdots \\ L(d,n) \end{pmatrix} = P_B^n \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix}$$

$$\lim_{n \to \infty} p(i,j) \frac{L(j, n-1)}{L(i,n)} \qquad ??$$

Proposition.

Si la matrice P_B est primitive (= irréductible et apériodique), alors il existe $\rho > 0$ tel que $\lim \frac{1}{\rho^n} P_b^n$ existe et est > 0.

Proposition.

Si la matrice P_B est irréductible, alors il existe $\rho > 0$ tel que $\lim \frac{1}{\rho^n} L(n)$ existe et est > 0.

Quitte à conjuguer par une matrice de permutation, on a

$$P_B = \begin{pmatrix} A_{1,1} & 0 & 0 & 0 & \dots & 0 \\ A_{2,1} & A_{2,2} & 0 & 0 & \dots & 0 \\ A_{3,1} & A_{3,2} & A_{3,3} & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \dots & \vdots \\ A_{e,1} & A_{e,2} & A_{e,3} & A_{e,4} & \dots & A_{e,e} \end{pmatrix}$$

où les matrices $A_{k,k}$, $1 \leq k \leq e$, sont des matrices carrées irréductibles ou nulles.

Proposition.

Si chaque matrice $A_{k,k}$ est primitive ou nulle, alors pour tous $i,j \in B$

$$p(i,j)^{(n)} \sim r_{i,j}(n)\rho_{i,j}^n$$

avec $r_{i,j}$ polynôme et $\rho_{i,j} > 0$.

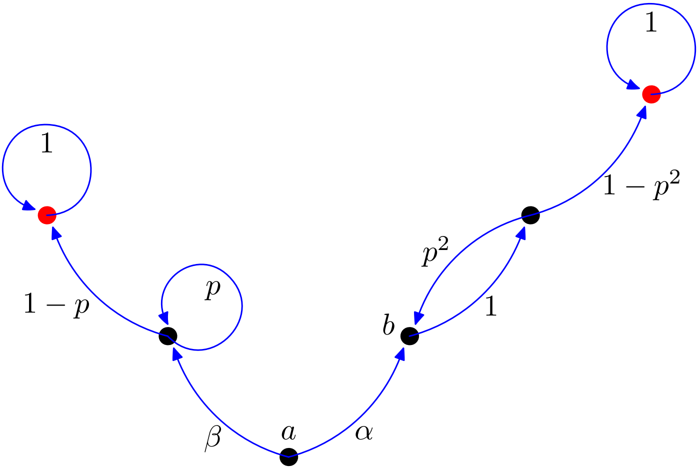
Conclusion provisoire : si la matrice P_B est irréductible ou si les blocs diagonaux de son écriture canonique sont primitifs ou nuls, alors pour tous i, j dans B,

$$\lim_{n\to\infty} \frac{L(j, n-1)}{L(i, n)} \qquad \text{existe.}$$

Par contre :

Si $P_B = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 0 & 2 \\ 0 & 1 & 0 \end{pmatrix}$, alors $L(2, 2n-1) = L(2, 2n) = 2^{n+1} - 1$.

Un contre-exemple

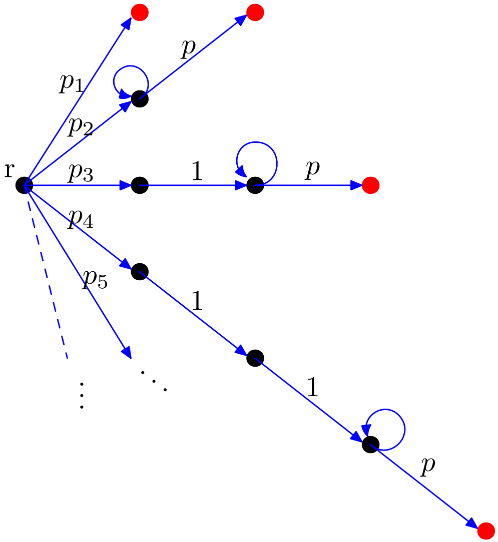


$B = \{\text{sites noirs}\}$,
$\mathbb{P}_a(Z_1 = b \mid Z_2, Z_3, \ldots, Z_{2n} \in B) = \alpha/(\alpha + p\beta)$,
$\mathbb{P}_a(Z_1 = b \mid Z_2, Z_3, \ldots, Z_{2n+1} \in B) = \alpha$.

Réponses aux deux questions

(Q1) $q(i,j) := \lim_{m\to\infty} q(i,j,0,m)$ existe-t-elle ?
Non, en général.

(Q2) Si ces limites existent, $(q(i,j))_{i,j\in B}$ est-elle une matrice de probabilité de transition ?
Oui évidemment si B est fini, mais non en général.



Si la suite (p_n) ne décroît pas trop vite, alors pour tout site i

$$\mathbb{P}_r(Z_1 = i \mid Z_1, Z_2, \ldots, Z_n \in B) \to 0 \; .$$

Marche aléatoire

μ est une loi de probabilité sur $\mathbb{Z}$, $(X_m)_{m\geq 0}$ est une suite de v.a. i.i.d. de loi μ, et $(S_m)_{mn\geq 0}$ est la marche aléatoire de loi μ :

$$S_m = S_0 + X_1 + X_2 + \ldots + X_m.$$

Objectif : Définition et description de la marche aléatoire conditionnée à rester négative.

Proposition.

Sous les hypothèses de moment appropriées on a, pour tous $i, j \in \mathbb{Z}^-$,

$$q(i,j) = \lim_{n\to\infty} \mathbb{P}\left(S_1 = j \mid S_0 = i, S_1 \leq 0, S_2 \leq 0, \ldots, S_n \leq 0\right)$$

existe et définit une probabilité de transition sur $\mathbb{Z}^-$.

L'étude de $\lim_{i\to-\infty} q(i, i+k)$ permettra de comparer les transitions de la marche conditionnée aux transitions de la marche initiale.

On pose $M_n = \max\{0, X_1, X_1 + X_2, \dots, X_1 + X_2 + \dots + X_n\}$.

$$\mathbb{P}\left(S_1 = j \mid S_0 = i, S_1 \leq 0, S_2 \leq 0, \dots, S_n \leq 0\right) = \frac{\mu(j-i)\ \mathbb{P}(M_{n-1} \leq -j)}{\mathbb{P}(M_n \leq -i)} \ .$$

Pour $k, \ell \geq 0$,

$$\mathbb{P}(M_n \leq k) = \sum_{\ell \geq 0} \mu(k-\ell)\ \mathbb{P}(M_{n-1} \leq \ell).$$

Hypothèse : $\mathbb{E}(|X|) < +\infty$.
Discussion suivant le signe de la dérive $\mathbb{E}(X)$.

dérive négative

Si $\mathbb{E}(X) < 0$ alors la suite (M_n) est presque sûrement majorée.
Pour tout $k \geq 0$,

$$\phi(k) := \lim_{n \to \infty} \mathbb{P}(M_n \leq k)$$

existe et ϕ est > 0 sur $\mathbb{N}$. La fonction $i \mapsto \phi(-i)$ est harmonique (pour $P_{\mathbb{Z}^-}$) :

$$\phi(k) = \sum_{\ell \geq 0} \phi(\ell) \mu(k - \ell) \ .$$

Les transitions de la marche conditionnée sont données par :

$$q(i,j) = \mu(j-i) \frac{\phi(-j)}{\phi(-i)} \ ,$$

et de plus

$$\lim_{i \to -\infty} q(i, i+k) = \mu(k).$$

(car $\lim_{+\infty} \phi = 1$).

dérive nulle, avec moment d'ordre 2

Si $\mathbb{E}(X) = 0$ et $\mathbb{E}(X^2) < +\infty$ alors pour tout $k \geq 0$,

$$\mathbb{P}(M_n \leq k) \sim_{n\to\infty} \frac{1}{\sqrt{n}}\phi(k) \qquad \text{(Kozlov (1976))}$$

avec $\phi > 0$ sur $\mathbb{N}$. Même relation d'harmonicité (pour $P_{\mathbb{Z}^-}$) :

$$\phi(k) = \sum_{\ell\geq 0} \phi(\ell)\,\mu(k-\ell)\ .$$

Les transitions de la marche conditionnée sont données par :

$$q(i,j) = \mu(j-i)\frac{\phi(-j)}{\phi(-i)}\ ,$$

et de plus

$$\lim_{i\to-\infty} q(i,i+k) = \mu(k).$$

(Avec une expression de ϕ et le théorème du renouvellement, on montre que $\phi(k) \sim_{k\to+\infty} ck$.)

dérive positive, avec moments exponentiels

On suppose ici que :
$\mathbb{E}(X) > 0$, $\mathbb{P}(X < 0) > 0$ et $\mathbb{E}\left(e^{tX}\right) < +\infty$ pour tout réel t.

Théorème.

Il existe $\rho \in]0,1[$ *et* $t_0 < 0$ *tels que*

$$\mathbb{P}(M_n \leq k) \sim_{n\to\infty} \frac{\rho^n}{n^{3/2}}\phi(k)$$

avec $\phi > 0$ *sur* $\mathbb{N}$, *et*

$$\phi(k) \sim_{k\to+\infty} ce^{-kt_0} \ .$$

La fonction $\phi(-\cdot)$ est propre pour $P_{\mathbb{Z}^-}$:

$$\rho\phi(k) = \sum_{\ell \geq 0} \phi(\ell)\mu(k-\ell) \ .$$

dérive positive, avec moments exponentiels

On suppose toujours que :
$\mathbb{E}(X) > 0$, $\mathbb{P}(X < 0) > 0$ et $\mathbb{E}\left(e^{tX}\right) < +\infty$ pour tout réel t.
Comme conséquence du théorème précédent, on a :

Les transitions de la marche conditionnée sont données par :

$$q(i,j) = \mu(j-i)\frac{\phi(-j)}{\rho\,\phi(-i)} \ ,$$

et de plus

$$\lim_{i\to-\infty} q(i,i+k) = \frac{e^{t_0 k}}{\rho}\mu(k).$$

où $t_0 < 0$.

Les poids $\frac{e^{t_0 k}}{\rho}\mu(k)$ sont ceux de la marche recentrée (ou "relativisée") au sens de Cramer.

dérive positive, avec moments exponentiels

Résumé d'une démonstration du théorème :

1. Procédé de relativisation de Cramer : $L_\mu(t) = \mathbb{E}(e^{tX})$ atteint son minimum en un point $t_0 < 0$. Ce minimum est noté ρ. la mesure

$$\mu'(k) = \frac{e^{t_0 k}}{\rho}\mu(k)$$

est une probabilité centrée sur $\mathbb{Z}$.

2. Théorème local de Le Page - Peigné pour le couple $(M'_n, M'_n - S'_n)$.
Un petit calcul donne

$$\mathbb{P}(M_n \leq k) = \rho^n \mathbb{E}\left(e^{t_0(M'_n - S'_n)}\mathbf{1}_{M'_n - S'_n > 0} \times e^{-t_0 M'_n}\mathbf{1}_{M'_n \leq k}\right).$$

et le théorème d'Émile et Marc donne précisément un équivalent de

$$\mathbb{E}(\varphi(M'_n, M'_n - S'_n)).$$

Théorème local de Le Page - Peigné

(On suppose ici la loi μ apériodique dans le groupe que son support engendre.)

$\tau_+ := \inf\{n \geq 1 \mid S'_n > 0\}$ et $U_+ = \sum_{n=0}^{+\infty} (\text{loi de } S'_{\tau_+})^{*n}$,

$\tau_- := \inf\{n \geq 1 \mid S'_n \leq 0\}$ et $U_- = \sum_{n=0}^{+\infty} (\text{loi de } -S'_{\tau_-})^{*n}$.

$$\mathbb{E}(\varphi(M'_n, M'_n - S'_n)) \sim \frac{1}{n^{3/2}} \int_0^{+\infty} \int_0^{+\infty} \varphi(x,y) m(dx,dy) \ ,$$

où

$$m = \frac{1}{\sigma\sqrt{2\pi}} ((\lambda * U_+) \otimes U_- + U_+ \otimes (\lambda * U_-)) \ ,$$

et λ est la mesure de Lebesgue sur $\mathbb{R}^+$.

Conséquence du théorème local de Le Page - Peigné

$$\mathbb{P}(M_n \leq k) \sim \frac{\rho^n}{n^{3/2}} \phi(k)$$

avec

$$\phi(k) = \frac{1}{\sigma\sqrt{2\pi}} \Big(\int_0^{+\infty} e^{t_0 x} \, \lambda * U_+(dx) \cdot \int_0^k e^{-t_0 x} \, U_-(dx) \\ + \int_0^{+\infty} e^{t_0 x} \, U_+(dx) \cdot \int_0^k e^{-t_0 x} \, \lambda * U_-(dx) \Big) .$$

La marche aux plus proches voisins.

Si $\mu = q\delta_{-1} + r\delta_0 + p\delta_1$, on obtient des formules explicites.

Quand la dérive est négative $(q > p)$: on a $\mathbb{P}(M_n \leq k) \to 1 - \left(\frac{p}{q}\right)^{k+1}$,

$$q(i,i-1) = \frac{q^{2-i} - p^{2-i}}{q^{1-i} - p^{1-i}} \ , \quad q(i,i) = r \ , \quad q(i,i+1) = pq\frac{q^{-i} - p^{-i}}{q^{1-i} - p^{1-i}} \ .$$

Ces formules s'appliquent encore quand la dérive est nulle $(q = p)$:

$$q(i,i-1) = p\frac{2-i}{1-i} \ , \quad q(i,i) = r \ , \quad q(i,i+1) = p\frac{-i}{1-i} \ .$$

On retrouve les probabilités de transition de la chaîne de Markov $2\min_{0\leq t\leq n} S_t \ - S_n$. C'est le théorème de Pitman !

La marche aux plus proches voisins.

Quand la dérive est positive $(q < p)$:

$$q(i,i-1) = \frac{\sqrt{pq}}{2\sqrt{pq}+r}\,\frac{2-i}{1-i}\ ,$$

$$q(i,i) = \frac{r}{2\sqrt{pq}+r}\ ,$$

$$q(i,i+1) = \frac{\sqrt{pq}}{2\sqrt{pq}+r}\,\frac{-i}{1-i}\ .$$

Si $r=0$ ces transitions sont indépendantes de (p,q).
Il y a un centrage asymptotique.
La chaîne de Markov $2\min_{0\le t\le n} S_t\ - S_n$ est toujours bien définie mais c'est autre chose.