Méthodes de Monte-Carlo par chaînes de Markov et algorithmes particulaires

Examen du mardi 23 avril 2024 9h00-12h00

Soit P un noyau markovien sur $(E, \mathcal{E})$.

1. Soient $\mu, \sigma \in \mathcal{P}(E)$. On note $\xi_+ - \xi_-$ la décomposition de Jordan-Hahn de $\mu - \sigma$.
   (a) Montrer que si μ est invariante par P, alors $d_{\mathrm{TV}}(\sigma, \sigma P) \leq 2 d_{\mathrm{TV}}(\sigma, \mu)$ et en déduire que l'ensemble des probabilités invariantes par P est un fermé de $\mathcal{P}(E)$ muni de la distance en variation totale.
   (b) Vérifier que $\xi_+ = \mu - \mu \wedge \sigma$ et $\xi_- = \sigma - \mu \wedge \sigma$.
   (c) Si μ et σ sont toutes deux invariantes par P, montrer que les mesures $\mu \wedge \sigma, \xi_+$ et ξ_- sont également invariantes par P.
   (d) Conclure que lorsque P admet deux probabilités invariantes distinctes, alors ce noyau admet deux probabilités invariantes $\tilde{\mu}$ et $\tilde{\sigma}$ mutuellement singulières au sens où il existe $B \in \mathcal{E}$ tel que $\tilde{\mu}(B^c) = 0 = \tilde{\sigma}(B)$.

Soit π une probabilité sur E invariante par P. Soit $(X_n)_{n \in \mathbb{N}}$ une chaîne de Markov de noyau P. Pour $n \in \mathbb{N}$, on pose $\mathcal{F}_n = \sigma(X_0, \cdots, X_n)$. Pour mettre en valeur le rôle joué par la loi μ de X_0, on note $\mathbb{P}_\mu$ et $\mathbb{E}_\mu$ la probabilité et l'espérance sur l'espace sous-jacent. Dans le cas particulier $\mu = \delta_x$ où $x \in E$, on simplifie ces notations en $(\mathbb{P}_x, \mathbb{E}_x)$. Pour $C \in \mathcal{E}$, on note $\tau_C = \inf\{k \geq 1 : X_k \in C\}$ avec la convention $\inf \emptyset = +\infty$. Un ensemble $C \in \mathcal{E}$ est dit π-accessible si $\pi(dx)$ p.p., $\mathbb{P}_x(\tau_C < \infty) > 0$.

2. Soit $C \in \mathcal{E}$ et $A = \{(x_k)_{k \in \mathbb{N}} \in E^{\mathbb{N}} : \exists k \geq 1 \text{ t.q. } x_k \in C\} \in \mathcal{E}^{\otimes \mathbb{N}}$. On pose $\varphi(x) = \mathbb{P}_x((X_k)_{k \in \mathbb{N}} \in A)$ pour $x \in E$.
   (a) Montrer que pour $x \in E$, $\varphi(x) = \mathbb{P}_x(\tau_C < \infty)$.
   (b) Vérifier que pour $n \in \mathbb{N}^*$ et $(x_k)_{k \in \mathbb{N}} \in E^{\mathbb{N}}$, $1_A((x_k)_{k \geq n}) \leq 1_A((x_k)_{k \in \mathbb{N}})$ et en déduire que $\mathbb{P}_\pi$ p.s., $1_A((X_k)_{k \geq n}) = 1_A((X_k)_{k \in \mathbb{N}})$.
   (c) Montrer $(\varphi(X_n))_{n \in \mathbb{N}}$ est une $\mathcal{F}_n$-martingale sous $\mathbb{P}_\pi$. Préciser sa limite pour $n \to \infty$ et vérifier que pour tout $n \in \mathbb{N}$, $\varphi(X_n)$ lui est $\mathbb{P}_\pi$ p.s. égale.
   (d) En déduire que $\pi(dx)$ p.p., $\varphi(x) \in \{0, 1\}$.
   (e) Conclure que si C est π-accessible, alors $\pi(dx)$ p.p., $\mathbb{P}_x(\tau_C < \infty) = 1$.

3. Soit $f : E \to \mathbb{R}_+$ mesurable bornée.
   (a) Pour $n \in \mathbb{N}^*$, vérifier que $\{X_k \in C, X_{k+1} \notin C, X_{k+2} \notin C, \cdots, X_n \notin C\}_{1 \leq k \leq n-1}$ et $\{X_n \in C\}$ constituent une partition de l'événement $\{\tau_C \leq n\}$ et en déduire que

$$\pi(f) = \sum_{k=1}^{n} \mathbb{E}_\pi \left[1_C(X_k) \left(\prod_{\ell=k+1}^{n} 1_{C^c}(X_\ell) \right) f(X_n) \right] + \mathbb{E}_\pi[f(X_n) 1_{\{\tau_C > n\}}]. \tag{1}$$

(b) Vérifier que pour $k \in \{1, \cdots, n\}$, $\mathbb{E}_\pi \left[1_C(X_k) f(X_n) \prod_{\ell=k+1}^n 1_{C^c}(X_\ell)\right] = \int_C \mathbb{E}_x[f(X_{n-k}) 1_{\{\tau_C > n-k\}}] \pi(dx)$ et en déduire que le premier terme du second membre de (1) est égal à $\int_C \mathbb{E}_x \left[\sum_{j=0}^{(n-1)\wedge(\tau_C - 1)} f(X_j)\right] \pi(dx)$.

(c) En déduire que $\pi \geq \pi_C^0$ où π_C^0 est la mesure définie par $\pi_C^0(g) = \int_C \mathbb{E}_x \left[\sum_{j=0}^{\tau_C - 1} g(X_j)\right] \pi(dx)$ pour $g : E \to \mathbb{R}$ mesurable bornée.

On suppose maintenant que C est π-accessible.

(d) Quel est le comportement asymptotique pour $n \to \infty$ du second terme du second membre de (1)?

(e) En déduire que $\pi = \pi_C^0$ puis que $\pi(C) > 0$.

(f) En calculant $\pi_C^0(Pf)$, vérifier que π_C^1 définie par $\pi_C^1(g) = \int_C \mathbb{E}_x \left[\sum_{j=1}^{\tau_C} g(X_j)\right] \pi(dx)$ satisfait également $\pi_C^1 = \pi$.

On suppose désormais que π est l'unique probabilité invariante par P.

4. (a) Quel est le comportement asymptotique pour $n \to \infty$ de $\frac{1}{n} \sum_{k=0}^{n-1} 1_C(X_k)$ sous $\mathbb{P}_\pi$?

(b) En déduire que si $\pi(C) > 0$ alors $\pi(dx)$ p.p., $\mathbb{P}_x(\tau_C < \infty) = 1$ et C est π-accessible.

(c) Conclure que si $\pi(C) > 0$ alors $\pi = \pi_C^0 = \pi_C^1$.

(d) Lorsque P satisfait les conditions de dérive $(D1)$ et $(D2)$, vérifier que si $\pi(C) > 0$, alors $\mathbb{P}_x(\tau_C < \infty) = 1$ **pour tout** $x \in E$.

On suppose que C est π-accessible et on note π_C la probabilité sur $(E, \mathcal{E})$ définie par $\pi_C(A) = \frac{\pi(A \cap C)}{\pi(C)}$ pour tout $A \in \mathcal{E}$. On se donne également un noyau markovien Q qui laisse π_C invariante. Partant d'un couple initial (Y_0, Z_0) à valeurs dans $E \times E$, on construit une chaîne de Markov $((Y_n, Z_n))_{n \in \mathbb{N}}$ en itérant le passage suivant de l'instant n à l'instant $n+1$:

$$(Y_{n+1}, Z_{n+1}) = 1_{C^c}(W_{n+1})(W_{n+1}, Z_n) + 1_C(W_{n+1})(\zeta_{n+1}, \zeta_{n+1})$$

où W_{n+1} et ζ_{n+1} sont choisis conditionnellement indépendants et respectivement distribués suivant $P(Y_n, \cdot)$ et $Q(Z_n, \cdot)$. On note également $\hat{\pi}$ la mesure sur $E \times E$ définie par $\hat{\pi}(g) = \int_C \mathbb{E}_x \left[\sum_{j=0}^{\tau_C - 1} g(X_j, x)\right] \pi(dx)$ pour toute fonction $g : E \times E \to \mathbb{R}$ mesurable bornée.

5. Montrer que la première marginale de $\hat{\pi}$ est π (on pourra choisir $g(y, z) = f(y)$ avec $f : E \to \mathbb{R}$ mesurable bornée) et en déduire que $\hat{\pi}$ est une probabilité.

6. Montrer que le noyau de la chaîne de Markov $((Y_n, Z_n))_{n \in \mathbb{N}}$ est

$$R((y, z), (ds, dt)) = 1_{C^c}(s) P(y, ds) \delta_z(dt) + P(y, C) Q(z, dt) \delta_t(ds).$$

7. Soit $g : E \times E \to \mathbb{R}_+$ mesurable.
On pose $T_1 = \int_{x\in C} \mathbb{E}_x \left[\sum_{k=0}^{\tau_C-1} \int_{y\in C^c} g(y,x)P(X_k,dy)\right] \pi(dx)$ et $T_2 = \int_{(x,z)\in C\times C} \mathbb{E}_x \left[\sum_{k=0}^{\tau_C-1} P(X_k,C)\right] g(z,z)\pi(dx)Q(x,dz)$.

   (a) Vérifier que $\hat{\pi}(Rg) = T_1 + T_2$.

   (b) Remarquer que $\mathbb{E}_x[1_{\{\tau_C>k\}}P(X_k,C)] = \mathbb{E}_x[1_{\{\tau_C>k\}}1_C(X_{k+1})]$ et en déduire que $\pi(dx)$ p.p., $\mathbb{E}_x \left[\sum_{k=0}^{\tau_C-1} P(X_k,C)\right] = 1$ puis que $T_2 = \int_{x\in C} g(x,x)\pi(dx)$.

   (c) Vérifier que
$$\mathbb{E}_x\left[1_{\{\tau_C>k\}} \int_{y\in C^c} g(y,x)P(X_k,dy)\right] = \mathbb{E}_x[1_{\{\tau_C>k\}}g(X_{k+1},x)1_{C^c}(X_{k+1})]$$
et en déduire que $T_1 = \int_{x\in C} \mathbb{E}_x \left[\sum_{j=1}^{\tau_C-1} g(X_j,x)\right] \pi(dx)$.

   (d) Conclure que $\hat{\pi}$ est invariante par R.

8. Supposons que π_C l'unique probabilité invariante par Q.

   (a) Pour η probabilité sur $E \times E$, vérifier que pour $B \in \mathcal{E}$,
$$\eta R(E\times B) = \int_{(y,z)\in E\times B} P(y,C^c)\eta(dy,dz) + \int_{(y,z)\in E\times E} P(y,C)Q(z,B)\eta(dy,dz).$$

   On suppose maintenant que η est invariante par R.

   (b) Montrer que pour tout $B \in \mathcal{E}$,
$$\int_{(y,z)\in E\times E} P(y,C)Q(z,B)\eta(dy,dz) = \int_{(y,z)\in E\times B} P(y,C)\eta(dy,dz)$$
et en déduire que la seconde marginale $\check{\eta}(dz)$ de la mesure $P(y,C)\eta(dy,dz)$ vérifie $\check{\eta} = \check{\eta}(E)\pi_C$.

   (c) Si $\check{\eta}(E) = 0$, vérifier que pour $B \in \mathcal{E}$, $\int_{(y,z)\in E\times E} P(y,C\cap B)\eta(dy,dz) = 0$ et en déduire que $\eta(B \times E) = \int_{(y,z)\in E\times E} P(y,B)\eta(dy,dz)$ puis que la première marginale η_1 de η est égale à π et enfin que $\pi(C) = \int_{(y,z)\in E\times E} P(y,C)\eta(dy,dz)$. Conclure que $\check{\eta}(E) > 0$.

   (d) Pour $g : E \times E \to \mathbb{R}_+$ mesurable, vérifier que
$$\eta(1_{C\times C}g) = \int_{z\in E} \int_{t\in E} 1_C(t)g(t,t)Q(z,dt)\check{\eta}(dz)$$
et en déduire que $\eta(1_{C\times C}g) = \frac{\check{\eta}(E)}{\pi(C)} \int_{z\in C} g(z,z)\pi(dz)$ puis que $\eta(g) \geq \frac{\check{\eta}(E)}{\pi(C)} \int_C \mathbb{E}_{(z,z)} \left[\sum_{k=0}^{\tau_{C\times C}-1} g(Y_k,Z_k)\right] \pi(dz)$ où $\tau_{C\times C} = \inf\{k \geq 1 : (Y_k,Z_k) \in C \times C\}$. Conclure que $\eta(g) \geq \frac{\check{\eta}(E)}{\pi(C)}\hat{\pi}(g)$.

   (e) Avec la question 1d, en déduire que R admet $\hat{\pi}$ comme unique probabilité invariante.

   (f) Pour $g : E \times E \to \mathbb{R}$ mesurable et telle que $\hat{\pi}(|g|) < \infty$, justifiez que
$$\hat{\pi}(dy,dz) \text{ } p.p., \text{ } \mathbb{P}_{(y,z)}\left(\lim_{n\to\infty} \frac{1}{n}\sum_{k=0}^{n-1} g(Y_k,Z_k) = \hat{\pi}(g)\right) = 1.$$