Examen de Chaînes de Markov

Randal Douc et Eric Moulines

Exercise 1

We first recall some usual notation, already used in the course.

Let $(\mathsf{X},\mathscr{X})$ be a measurable space. For a given Markov kernel P on $\mathsf{X}\times\mathscr{X}$, we denote by $\mathbb{P}_\xi$, the probability measure induced on $(\mathsf{X}^{\mathbb{N}},\mathscr{X}^{\otimes\mathbb{N}})$ by the Markov kernel P and initial distribution $\xi\in\mathbb{M}_1(\mathscr{X})$. The associated expectation operator is noted $\mathbb{E}_\xi$.

For a given measurable space $(\mathsf{G},\mathscr{G})$, $\mathsf{F}_+(\mathsf{G})$ denotes the set of measurable functions taking values in $\mathbb{R}^+\cup\{+\infty\}$.

Let $\mathsf{C}\in\mathscr{X}$ and let $\mathscr{C}$ be the σ-field $\mathscr{C}=\{\mathsf{A}\cap\mathsf{C}:\mathsf{A}\in\mathscr{X}\}$. Define by π_{C} the restriction of π on C, that is: $\pi_{\mathsf{C}}(\mathsf{A})=\frac{\pi(\mathsf{A}\cap\mathsf{C})}{\pi(\mathsf{C})}$ for all $\mathsf{A}\in\mathscr{X}$. Recall that the first hitting time of the set C is defined by

$$\sigma_{\mathsf{C}}=\inf\{k\geq 1: X_k\in\mathsf{C}\}$$

with the convention that $\sigma_{\mathsf{C}}=\infty$ if $\{k\geq 1: X_k\in\mathsf{C}\}=\emptyset$.

In what follows, we say that $\mathsf{C}\in\mathscr{X}$ is π-accessible if $\mathbb{P}_x(\sigma_{\mathsf{C}}<\infty)>0$ for π-almost all $x\in\mathsf{X}$. In this exercise, we assume that

(i) P is a Markov kernel on $\mathsf{X}\times\mathscr{X}$ with an invariant probability measure π such that C is π-accessible.

We recall that under (i), the Kac formula holds true: for all $h\in\mathsf{F}_+(\mathsf{X})$,

$$\pi(h)=\int_{\mathsf{C}}\pi(\mathrm{d}x)\mathbb{E}_x\left[\sum_{k=0}^{\sigma_{\mathsf{C}}-1}h(X_k)\right]=\int_{\mathsf{C}}\pi(\mathrm{d}x)\mathbb{E}_x\left[\sum_{k=1}^{\sigma_{\mathsf{C}}}h(X_k)\right]\tag{1.1}$$

1. Using the Kac formula (1.1) with a convenient choice of h, show that $\mathbb{P}_\pi(\sigma_{\mathsf{C}}<\infty)=1$.
2. Deduce that $\mathbb{P}_x(\sigma_{\mathsf{C}}<\infty)=1$ for π-almost all $x\in\mathsf{X}$

We assume in addition that

(ii) Q is a Markov kernel on $\mathsf{C}\times\mathscr{C}$ with an invariant probability measure π_{C}.

For any probability measure μ on $(\mathsf{X}\times\mathsf{C},\mathscr{X}\otimes\mathscr{C})$, consider a family of random elements $\{(Y_k,Z_k),k\in\mathbb{N}\}$ on the same probability space $(\Omega,\mathscr{F},\mathbb{P})$ where Y_k takes values on X and Z_k on C, obtained from Algorithm 1, where we have defined $\mathscr{F}_k=\sigma(Y_0,\ldots,Y_k,Z_0,\ldots,Z_k)$.

Algorithm 1 The teleportation process

1: **Initialization:** sample $(Y_0,Z_0)\sim\mu$, that is $\mathbb{P}((Y_0,Z_0)\in\mathsf{A})=\mu(\mathsf{A})$ for $\mathsf{A}\in\mathscr{X}\otimes\mathscr{C}$.
2: **for** $k=1$ to n **do**
3: sample $Y_k^\star\sim P(Y_{k-1},\cdot)$ that is $\mathbb{P}(Y_k^\star\in\mathsf{A}|\mathscr{F}_{k-1})=P(Y_{k-1},A)$ for $\mathsf{A}\in\mathscr{X}$.
4: **if** $Y_k^\star\notin\mathsf{C}$ **then**
5: set $(Y_k,Z_k)=(Y_k^\star,Z_{k-1})$.
6: **else**
7: sample $Z_k\sim Q(Z_{k-1},\cdot)$, that is $\mathbb{P}(Z_k\in\mathsf{B}|\mathscr{F}_{k-1}\cap\sigma(Y_k^\star))=Q(Z_{k-1},\mathsf{B})$ for all $\mathsf{B}\in\mathscr{C}$ on the event $\{Y_k^\star\in\mathsf{C}\}$. And set $Y_k=Z_k$.
8: **end if**
9: **end for**

3. Show that $\{(Y_k, Z_k), k \in \mathbb{N}\}$ is a Markov chain. Denoting by R the associated Markov kernel on $(\mathsf{X} \times \mathsf{C}) \times (\mathscr{X} \otimes \mathscr{C})$, write the expression of $Rh(y,z)$ for any $(y,z) \in \mathsf{X} \times \mathsf{C}$ and $h \in \mathsf{F}_+(\mathsf{X} \times \mathsf{C})$.

Define the measure $\bar{\pi}$ on $(\mathsf{X} \times \mathsf{C}, \mathscr{X} \otimes \mathscr{C})$ by: for all $h \in \mathsf{F}_+(\mathsf{X} \times \mathsf{C})$,

$$\bar{\pi}(h) = \int_{\mathsf{C}} \pi(\mathrm{d}x) \mathbb{E}_x \left[\sum_{k=0}^{\sigma_{\mathsf{C}}-1} h(X_k, x) \right], \tag{1.2}$$

where $\mathbb{E}_\xi$ is the expectation associated to the probability measure $\mathbb{P}_\xi$ induced on $(\mathsf{X}^{\mathbb{N}}, \mathscr{X}^{\otimes \mathbb{N}})$ by the Markov kernel P and initial distribution $\xi \in \mathbb{M}_1(\mathscr{X})$, and we set by convention $\mathbb{E}_x = \mathbb{E}_{\delta_x}$ to alleviate notation.

4. Express the first and second marginal distributions of $\bar{\pi}$.
5. Show that $\bar{\pi}(Rh) = A + B$ where

$$A = \int_{\mathsf{C}} \pi(\mathrm{d}x) \sum_{k=0}^{\infty} \mathbb{E}_x \left[\mathbb{1}_{\{k \le \sigma_{\mathsf{C}}-1\}} \int_{\mathsf{C}^c} P(X_k, \mathrm{d}y') h(y', x) \right],$$
$$B = \int_{\mathsf{C}} \pi(\mathrm{d}x) \sum_{k=0}^{\infty} \mathbb{E}_x \left[\mathbb{1}_{\{k \le \sigma_{\mathsf{C}}-1\}} P(X_k, \mathsf{C}) \right] \int_{\mathsf{C}} Q(x, \mathrm{d}z') h(z', z').$$

6. Show that $B = \int_{\mathsf{C}} \pi(\mathrm{d}z') h(z', z')$.
7. Deduce that $\bar{\pi}$ is an invariant probability measure for R.

Exercise 2

Nous considérons l'algorithme de Metropolis-Hastings sur $\mathsf{X} = \mathbb{R}^k$ muni de sa tribu Borélienne $\mathscr{X} = \mathscr{B}(\mathbb{R}^k)$. Nous notons h_π la densité cible, $q(x, \cdot)$ la densité de proposition (ces densités sont prises par rapport à la mesure de Lebesgue, notée Leb). On peut toujours choisir ces densités pour que $h_\pi(y) < \infty$ et $q(x,y) < \infty$ pour tout $x, y \in \mathbb{R}^k$. Pour $x \in \mathbb{R}^k$ et $A \in \mathscr{B}(\mathbb{R}^k)$, le noyau de transition est donné par

$$P(x, A) = \int \alpha(x,y) q(x,y) \mathbb{1}_A(y) \mathrm{Leb}(\mathrm{d}y) + (1 - \bar{\alpha}(x) \delta_x(A)$$

où nous avons posé

$$\alpha(x,y) = \begin{cases} 1 \wedge \frac{h_\pi(y) q(y,x)}{h_\pi(x) q(x,y)} & h_\pi(x) q(x,y) > 0 \\ 1 & h_\pi(x) q(x,y) = 0 \end{cases} \tag{1.3}$$

et $\bar{\alpha}(x) = \int \alpha(x,y) q(x,y) \mathrm{Leb}(\mathrm{d}y)$. Nous rappelons que $A \mapsto \pi(A) = \int_A h_\pi(y) \mathbb{1}_A(y) \mathrm{Leb}(\mathrm{d}y)$ est une loi stationnaire pour P. Nous supposons que

$$h_\pi(y) > 0 \Rightarrow (q(x,y) > 0, \quad \forall x \in \mathbb{R}^k)$$

1. Montrer que π est une mesure d'irréductibilité. Est elle maximale ?
2. Le noyau P est il récurrent ?

Nous supposons maintenant et pour le reste de l'exercice que les fonctions $h_\pi : \mathbb{R}^k \to \mathbb{R}^+$ et $q : \mathbb{R}^k \times \mathbb{R}^k \to \mathbb{R}^+$ sont **continues et strictement positives**.

3. Montrer que les ensembles compacts C tels que $\mathrm{Leb}(C) > 0$ sont "1-small" (vous préciserez la constante de minoration)

4. Montrer que le noyau P est apériodique.

Nous supposons maintenant (et pour le reste de l'exercice) que $q(x,y)$ ne dépend pas de x. On note alors $q(x,y) = q(y)$.

5. Montrer que si $q(y)/\pi(y) \geq \beta$ pour tout $y \in \mathbb{R}^k$, alors $\mathbb{R}^k$ est un "ensemble small".
6. Déterminer $\varepsilon > 0$ tel que pour tout $x \in \mathbb{R}^k$, $\|\delta_x P^n - \pi\|_{\mathrm{TV}} \leq 2(1-\varepsilon)^n$.

Nous supposons maintenant que pour tout $n > 0$, $\mathrm{Leb}(D_n) > 0$ où

$$D_n = \left\{ x \in \mathbb{R}^k : q(x)/h_\pi(x) \leq n^{-1} \right\}.$$

On sait que si le noyau Markovien est géométriquement ergodique alors il existe un "small set" C et $\kappa > 1$ tel que $\sup_{x \in C} \mathbb{E}_x[\kappa^{\sigma_C}] < \infty$.

7. Montrer que pour tout $x \in D_n$, pour tout $\ell \geq 1$ et tout $A \in \mathscr{B}(\mathbb{R}^k)$,
$$P^\ell(x,A) \leq n^{-1}\pi(A) + P^{\ell-1}(x,A),$$
avec la convention usuelle $P^0(x,A) = \mathbb{1}_A(x)$. En déduire que si C est un $(m, \varepsilon\nu)$-small set, alors pour tout $n > (2m)/\varepsilon$, on a $\mathrm{Card}(D_n \cap C) \leq 1$.

Pour $x \in \mathbb{R}^k$, nous notons

$$A_x = \left\{ y \in \mathbb{R}^k : [\pi(y)/q(y)][q(x)/\pi(x)] \geq 1 \right\}$$
$$R_x = \left\{ y \in \mathbb{R}^k : [\pi(y)/q(y)][q(x)/\pi(x)] < 1 \right\}$$

8. Soit $x \in D_n$. Montrer que pour tout $y \in A_x$, on a $1 \leq n^{-1}[\pi(y)/q(y)]$. En déduire que $P(x, A_x \setminus \{x\}) \leq n^{-1}$.
9. Montrer que pour tout $x \in D_n$, on a $P(x, R_x) \leq n^{-1}$.
10. Montrer que pour tout $x \in D_m$, on a $P(x,\{x\}) \geq (1-2/n)$. En déduire que pour tout $x \in D_n \cap C^c$, on a $\mathbb{P}_x(\sigma_C > k) \leq (1-2/n)^k$.
11. Montrer que le noyau Markovien P n'est pas géométriquement ergodique.

Exercise 3

Soit $(\mathsf{X}, \mathscr{X})$ un espace mesurable, P un noyau de Markov irréductible et récurrent, ψ une mesure d'irréductibilité maximale. Pour A un ensemble accessible, définissons pour $n \in \mathbb{N}^*$,

$$G_n = \{x \in A : \mathbb{P}_x(\sigma_A < \infty) < 1 - 1/n\},$$

où pour $B \in \mathscr{X}$, $\sigma_B = \{k \geq 1 : X_k \in B\}$ est le premier temps de retour en B.

1. Montrer que pour tout $x \in G_n$, $\mathbb{P}_x(\sigma_{G_n} < \infty) \leq 1 - 1/n$.
2. En déduire que $\sup_{x \in G_n} U(x, G_n) \leq n$ où U est le noyau potentiel de P.
3. Montrer que $\psi(A_0) = 0$ où $A_0 = \{x \in A : \mathbb{P}_x(\sigma_A < \infty) < 1\}) = 0$.

Nous posons $A_1 = A \setminus A_0 = \{x \in A : \mathbb{P}_x(\sigma_A < \infty) = 1\}$ et

$$A_{10} = \left\{ x \in A_1 : \mathbb{P}_x(\sigma_{A_0} < \infty) = 0 \right\}, \qquad A_{11} = \left\{ x \in A_1 : \mathbb{P}_x(\sigma_{A_0} < \infty) > 0 \right\},$$

4. Montrer que pour tout $x \in \mathsf{X}$,

$$\mathbb{P}_x(\sigma_{A_0} < \infty) \geq \mathbb{E}_x[\mathbb{1}_{\{\sigma_{A_{11}} < \infty\}} \mathbb{P}_{X_{\sigma_{A_{11}}}}(\sigma_{A_0} < \infty)]$$

   En déduire que $\psi(A_{11}) = 0$.
5. Montrer que pour $x \in A_{10}$, $\mathbb{P}_x(\sigma_{A_{11}} < \infty) = 0$.
6. Montrer que pour tout $x \in A_{10}$, $\mathbb{P}_x(N_{A_{10}} = \infty) = 1$ où N_B est le nombre de visites à B.