Contrôle de Modèles et algorithmes stochastiques 1 et 2

Durée 2h00

Les notes de cours sont autorisées. Les calculatrices et les téléphones portables sont interdits.
Le barème est approximatif.

Exercice 1 (Réduction de variance) (12 points)

La mise en oeuvre de méthodes de réduction de variance est souvent nécessaire pour rendre les algorithmes de Monte-Carlo efficaces. Parmi les méthodes de réduction de variance, il existe la méthode d'"Echantillonnage préférentiel" qui consiste à simuler des variables selon une loi adaptée à la fonction que l'on cherche à estimer. La question posée dans cet exercice est la suivante : comment bien choisir l'échantillonnage ? On se donne une loi μ sur $\mathbb{R}$ absolument continue par rapport à la mesure de Lebesgue et on note $p : \mathbb{R} \to \mathbb{R}$ sa densité. On considère $f : \mathbb{R} \to \mathbb{R}$ une fonction mesurable bornée et on s'intéresse au calcul de $\mathbb{E}[f(X)] = \int_{\mathbb{R}} f(x)p(x)dx$. On suppose pour simplifier que $\mathbb{P}(f(X) = 0) = 0$.

1. (a) Montrer que pour tout $\theta \in \mathbb{R}$,

$$\mathbb{E}[f(X)] = \mathbb{E}[f(X+\theta)\frac{p(X+\theta)}{p(X)}].$$

On notera

$$Y_\theta = f(X+\theta)\frac{p(X+\theta)}{p(X)}.$$

On a donc pour tout $\theta \in \mathbb{R}$, $\mathbb{E}[Y_\theta] = \mathbb{E}[f(X)]$. On peut donc envisager de choisir Y_θ ayant une variance minimale.

(b) Supposons qu'il existe θ^* tel que $\mathbb{E}[(Y_{\theta^*})^2] = \min_{\theta\in\mathbb{R}} \mathbb{E}[Y_\theta^2]$. A l'aide de la question précédente, montrer qu'alors

$$\mathrm{Var}(Y_{\theta^*}) = \min_{\theta\in\mathbb{R}} \mathrm{Var}(Y_\theta).$$

(c) On note $Q(\theta) = \mathbb{E}[(Y_\theta)^2]$. Montrer que

$$Q(\theta) = \mathbb{E}[f^2(X)\frac{p(X)}{p(X-\theta)}].$$

2. On suppose maintenant que $X \sim \mathcal{N}(0,1)$. Le but est de construire un algorithme permettant d'approcher θ^* (s'il existe).

(a) Montrer qu'alors

$$Q(\theta) = \mathbb{E}[f^2(X)e^{\frac{\theta^2}{2}-\theta X}].$$

(b) On admet que Q est 2 fois dérivable (et satisfait les règles de dérivation des intégrales à paramètre). Calculer Q' et Q''.

(c) En déduire que Q' est strictement croissante. En admettant que $\lim_{\theta\to\pm\infty} Q(\theta) = +\infty$, en déduire que Q admet un unique minimum en un point que l'on notera θ^*.

3. Le but de la fin de l'exercice est de construire un algorithme permettant d'approcher θ^*. La fonction Q' ayant une croissance trop forte, on introduit une nouvelle fonction h définie par

$$h(\theta) = \mathbb{E}[f^2(X-\theta)(2\theta - X)].$$

(a) En faisant un nouveau changement de variable sur la fonction Q', montrer que $Q'(\theta) = h(\theta)e^{\theta^2}$. En déduire que θ^* est aussi l'unique point où la fonction h s'annule (et que h a le signe de Q').

(b) Montrer que

$$h(\theta) \leq C(1 + |\theta|)$$

et que

$$\mathbb{E}[\big(f^2(X-\theta)(2\theta - X)\big)^2] \leq C(1+\theta^2).$$

(c) Proposer un algorithme stochastique permettant d'approcher θ^* (où θ^* est vu comme l'unique zéro de la fonction h).

(d) Montrer à l'aide d'un théorème de Robbins-Monro (et de la question précédente) que cet algorithme converge vers θ^*.

Exercice 2 (Variante de l'algorithme de Metropolis-Hastings) (4 points)

Soit E un espace fini et $\pi > 0$ une mesure de probabilité sur E. Soit Q une matrice de transition sur E, irréductible et symétrique, appelée matrice de sélection. Soit $h :]0, \infty[\to]0, 1[$ une fonction vérifiant

$$h(u) = uh\left(\frac{1}{u}\right), \text{ pour tout } u > 0. \tag{1}$$

On construit une chaîne de Markov $(X_n)_{n\geq 0}$ sur E de la façon suivante. Pour tout $n \geq 0$, sachant $X_n = i$, on propose le prochain état de la chaîne suivant les probabilités données par la matrice de sélection Q. Si un état $j \neq i$ est proposé, on accepte cette transition avec probabilité

$$\delta_{ij} = h\left(\frac{\pi_j}{\pi_i}\right), \tag{2}$$

dans quel cas on pose $X_{n+1} = j$, sinon la chaîne reste dans l'état i, donc $X_{n+1} = i$.

On note P la matrice de transition de la chaîne de Markov construite ainsi. Pour tout $i \in E$ on pose $P_{ii} = 1 - \sum_{j\neq i} P_{ij}$.

1. Donner P_{ij} pour tout $i \neq j$ et montrer que P est réversible par rapport à π.
2. Montrer que, pour n'importe quelle loi initiale, la chaîne $(X_n)_{n\geq 0}$ converge en loi vers π quand $n \to \infty$.
3. Rappeler la formule de la probabilité d'acceptation δ_{ij} dans la version classique de l'algorithme de Metropolis-Hastings (pour le cas Q symétrique) et montrer qu'elle peut s'écrire sous la forme (2), avec une certaine fonction h qui vérifie la condition (1).

Exercice 3 (Chaîne de Markov réversible) (4 points)

Soit $(X_n)_{n\geq 0}$ une chaîne de Markov irréductible sur un espace E fini, de matrice de transition P et de loi initiale π, son unique probabilité invariante. Pour $N \in \mathbb{N}^*$ fixé, on définit $\hat{X}_n := X_{N-n}$, pour tout $n = 0, \ldots, N$.

1. Montrer que $(\hat{X}_n)_{0\leq n\leq N}$ est aussi une chaîne de Markov (appelée *chaîne retournée*), de même loi initiale π et de matrice de transition

$$\hat{P}_{ij} = \frac{\pi_j P_{ji}}{\pi_i}, \ \forall i, j \in E.$$

2. Montrer que $\frac{P+\hat{P}}{2}$ est une matrice de transition réversible par rapport à π.
3. Montrer que si P est réversible par rapport à une mesure de probabilité μ sur E, alors $\mu = \pi$ et la chaîne retournée $(\hat{X}_n)_{0\leq n\leq N}$ a la même loi que la chaîne $(X_n)_{0\leq n\leq N}$.
4. Montrer que si P est symétrique, alors elle est réversible par rapport à une mesure de probabilité qu'il faudra préciser.