Contrôle final de Modèles et algorithmes stochastiques

Durée 2h00

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

Exercice 1 (Quantification de Voronoi et distorsion optimale) (14 points)

Soit Y une v.a. réelle de densité de probabilité f_Y, admettant une variance, et soit $N \in \mathbb{N}^*$.

On cherche la meilleure approximation quadratique de Y par une v.a. discrète de support à au plus N points.

Pour $x = (x_1, \ldots, x_N) \in \mathbb{R}^N$, on définit les sous-ensembles $(C_i(x))_{i=1,\ldots,N}$ de $\mathbb{R}$ par

$$C_i(x) := \{y \in \mathbb{R} : |y - x_i| < \min_{j \neq i} |y - x_j|\}$$

(appelés *mosaique de Voronoi*) et on souhaite approcher la v.a. Y par la v.a. discrète

$$\hat{Y}^x := \sum_{i=1}^{N} x_i 1_{C_i(x)}(Y),$$

appelée *quantifieur de Voronoi* (la projection, selon le plus proche voisin, de Y sur la grille x.).

On appelle *distorsion* l'erreur quadratique moyenne de cette approximation :

$$D(x) := \mathbb{E}[(Y - \hat{Y}^x)^2].$$

On fait l'hypothèse que la fonction D admet un unique point de minimum local $x^* = (x_1^*, \ldots, x_N^*)$. On suppose de plus que

$$\sum_{i=1}^{N} \int_{C_i(x)} (x_i - x_i^*)(x_i - y) f_Y(y) dy \geq 0, \text{ pour tout } x \in \mathbb{R}^N.$$

1. Montrer que

$$D(x_1, \ldots, x_N) = \sum_{i=1}^{N} \int_{C_i(x)} (y - x_i)^2 f_Y(y) dy = \mathbb{E}[\min_{i=1,\ldots,N} (Y - x_i)^2].$$

2. Montrer que pour toute v.a. discrète Z de support inclus dans $\{x_1, \ldots, x_N\}$ on a

$$D(x_1, \ldots, x_N) \leq \mathbb{E}[(Y - Z)^2].$$

En déduire que la meilleure approximation de Y par une v.a. discrète de support à au plus N points est donnée par $\hat{Y}^x$, avec $x = x^*$.

3. Montrer que si $x_1 < \ldots < x_N$, alors pour tout $i = 1, \ldots, N$ on a $C_i(x) = \left] \frac{x_{i-1}+x_i}{2}, \frac{x_i+x_{i+1}}{2} \right[$, avec par convention $x_0 := -\infty$ et $x_{N+1} := +\infty$. (Vous pourriez faire un dessin.)

4. Pour $i = 1, \ldots, N$, montrer que

$$\frac{\partial D}{\partial x_i}(x) = 2 \int_{C_i(x)} (x_i - y) f_Y(y) dy = 2\mathbb{E}\left[(x_i - Y) 1_{C_i(x)}(Y)\right].$$

Indication : Vous pourriez faire la démonstration dans le cas où $x_1 < \ldots < x_n$.

On rappelle la formule suivante de dérivation d'une intégrale à paramètre :

$$\frac{d}{dx} \int_{a(x)}^{b(x)} f(x, y) dy = \int_{a(x)}^{b(x)} \frac{\partial f}{\partial x}(x, y) dy + b'(x) f(x, b(x)) - a'(x) f(x, a(x)).$$

5. Montrer qu'il existe une constante $C > 0$ t.q.

$$||\nabla D(x)||^2 \leq C(||x||^2 + \mathbb{E}(Y^2)).$$

Indication : Vous pourriez utiliser les inégalités $(\mathbb{E}(X))^2 \leq \mathbb{E}(X^2)$ et $(a+b)^2 \leq 2(a^2+b^2)$.

6. Pour tenter de déterminer le point $x^* = (x_1^*, \ldots, x_N^*)$ qui minimise la distorsion D, on propose d'utiliser l'algorithme stochastique récursif suivant. On construit $X^k = (X_1^k, \ldots, X_N^k)$, avec $X^0 = x^0 \in \mathbb{R}^N$ et pour tout $k \geq 0$:

$$X_i^{k+1} = X_i^k - \frac{2}{k+1}(X_i^k - Y^{k+1})1_{C_i(X^k)}(Y^{k+1}), \quad \text{pour } i = 1, \ldots, N,$$

avec $(Y^k)_k$ une suite de v.a. i.i.d. de même loi que Y.

   (a) Montrer que l'algorithme stochastique ci-dessus peut s'écrire sous la forme

$$X^{k+1} = X^k - \frac{1}{k+1}\nabla D(X^k) + \frac{1}{k+1}\Delta M^{k+1},$$

   avec ΔM^{k+1} un accroissement de martingale à valeurs dans $\mathbb{R}^N$.

   (b) Soit $V : \mathbb{R}^N \to \mathbb{R}$ définie par $V(x) = ||x - x^*||^2$. Montrer que V est une fonction de Lyapounov sous-quadratique pour ∇D.

   (c) Montrer que X^k converge *p.s.* vers x^* quand $k \to \infty$.

Exercice 2 (Autour de l'algorithme de Metropolis-Hastings) (6 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.

1. Montrer que si π est la la loi uniforme sur E et Q est apériodique, alors la chaîne de Markov de matrice de transition Q converge en loi vers π quand $n \to \infty$, pour n'importe quelle loi initiale.

2. On suppose maintenant que π est différente de la loi uniforme sur E et Q n'est pas forcément apériodique. Soit $g : [0, \infty[\to [0, 1]$ définie par $g(u) = \min(1, u)$.
On construit une chaîne de Markov $(X_n)_{n\geq 0}$ sur E de matrice de transition P définie par

$$P_{ij} = Q_{ij} \; g\left(\frac{\pi_j}{\pi_i}\right), \; \text{pour } i \neq j. \tag{1}$$

   (a) Donner P_{ii} pour tout $i \in E$ et montrer que P est réversible par rapport à π.

   (b) Montrer (avec soin) que la matrice P est irréductible et apériodique.

   (c) Montrer que, pour n'importe quelle loi initiale, la chaîne $(X_n)_{n\geq 0}$ converge en loi vers π quand $n \to \infty$.

   (d) Dans la relation (1) définissant la matrice P, on prend maintenant la fonction $g(u) = \frac{u}{1+u}$. Montrer que P est encore réversible par rapport à π. Montrer ensuite que la chaîne $(X_n)_{n\geq 0}$ converge en loi vers π quand $n \to \infty$.