# Examen du cours de M2 :
# Méthodes numériques probabilistes

16 janvier 2008, 09h00 - 12h00.

*Les notes de cours sont autorisées.*

## Algorithme "Coupling From The Past"

On considère un espace d'états $M$ discret, muni d'une probabilité $\pi$. Soit $(\theta_n)_{n\geq 1}$ une suite de variables aléatoires i.i.d., à valeurs dans un espace mesurable $(\Theta, \mathcal{A})$. Soit une fonction $f : \Theta \times M \to M$ que l'on suppose mesurable. Dans la suite, pour $\theta \in \Theta$, on notera $f_\theta : M \to M$ la fonction $f(\theta, \cdot)$.

On introduit la suite de variables aléatoires $(X_n)_{n\geq 0}$ avec $X_0$ indépendant des $(\theta_n)_{n\geq 1}$ et

$$X_{n+1} = f(\theta_{n+1}, X_n). \tag{1}$$

**1** Montrer que la suite $(X_n)_{n\geq 0}$ est une chaîne de Markov, de matrice de transition $P$ que l'on précisera.

**2** Comme exemple d'une telle dynamique, on considère le modèle d'Ising sur le réseau de dimension 1 : $\{1, \ldots, N\}$ où $N$ désigne le nombre de sites. Dans ce cas, $M = \{-1, +1\}^N$, et $\pi(x) = Z^{-1} \exp(-\beta V(x))$, où $\beta$ est une constante positive, $Z = \sum_{x \in M} \exp(-\beta V(x))$, et $V(x) = -H \sum_{i=1}^N x_i - \sum_{i,j=1}^N J_{i,j} x_i x_j$, où $J$ est une matrice symétrique. On considère la dynamique de Metropolis Hasting suivante : pour passer de la configuration $X_n$ à la configuration $X_{n+1}$,

- *Proposition* : choisir (indépendemment) un site $k_{n+1}$ de manière uniforme sur $\{1, \ldots, N\}$ et un spin $S_{n+1}$ de manière uniforme dans $\{-1, +1\}$, et proposer la configuration $\tilde{X}_{n+1}$ obtenue à partir de $X_n$ en changeant le spin du site $k_{n+1}$ en $S_{n+1}$ : $\tilde{X}_{n+1,i} = X_{n,i}$ pour tout $i \in \{1, \ldots, k_{n+1} - 1, k_{n+1} + 1, \ldots, N\}$ et $\tilde{X}_{n+1,k_{n+1}} = S_{n+1}$. (On a noté $X_n = (X_{n,1}, \ldots, X_{n,N})$ les composantes de $X_n$).
- *Acceptation / Rejet* : accepter ou rejeter la configuration $\tilde{X}_{n+1}$ suivant la règle de Metropolis Hasting.

**2.a** Ecrire précisément l'algorithme qui permet d'obtenir $X_{n+1}$ en fonction de $X_n$.

**2.b** Justifier le fait que la chaîne de Markov $(X_n)_{n\geq 0}$ est irréductible et apériodique de mesure invariante $\pi$.

**2.c** Ecrire la dynamique sur $X_n$ sous la forme (1) en précisant les $(\theta_n)_{n\geq 1}$.

On revient maintenant au cadre général du début du problème. L'objectif est de construire une méthode de simulation exacte de la mesure de probabilité $\pi$. Pour cela, on introduit l'application aléatoire $D_n : M \to M$ définie par : $D_0 = \mathrm{Id}$ et

$$D_n = f_{\theta_1} \circ f_{\theta_2} \circ \ldots \circ f_{\theta_n}.$$

On introduit également l'application aléatoire $G_n : M \to M$ définie par : $G_0 = \mathrm{Id}$ et

$$G_n = f_{\theta_n} \circ f_{\theta_{n-1}} \circ \ldots \circ f_{\theta_1}.$$

**3** On suppose que $P$ est une matrice de transition irréductible et apériodique, de mesure invariante $\pi$. Que peut-on dire de $\mathbb{P}(G_n(x) = y)$ pour $x, y \in M$ et $n \to \infty$ ?

On introduit le *temps de coalescence*

$$T = \inf\{n \in \mathbb{N},\, D_n \text{ est une application constante}\},$$

avec la convention $\inf \emptyset = +\infty$ et on note

$$\Phi = 1_{T<\infty} D_T.$$

**4** On suppose que $\mathbb{P}(T < \infty) = 1$. Montrer que $\Phi$ a pour loi $\pi$. *Indication : on pourra utiliser le fait que, pour $x \in M$ et $n \geq 1$ fixés, $G_n(x)$ et $D_n(x)$ ont même loi.*

**5** *Cette question est facultative et indépendante du reste du problème.* Dans cette question, on propose de donner une condition suffisante pour que $\mathbb{P}(T < \infty) = 1$. On suppose que $M$ est composé d'un nombre fini d'états et que $\exists \alpha \in (0, 1]$,

$$\forall A \subset M, |A| > 1 \implies \mathbb{P}(|f_{\theta_1}(A)| < |A|) \geq \alpha,$$

où $|\cdot|$ désigne ici le cardinal. On note $d = |M|$. Montrer que

$$\mathbb{P}(D_{d-1} \text{ est une application constante}) \geq \alpha^{d-1}.$$

En déduire que pour tout $k \in \mathbb{N}$, $\mathbb{P}(T > k(d-1)) \leq (1 - \alpha^{d-1})^k$ puis que $\mathbb{P}(T < \infty) = 1$.

**6** Quel(s) inconvénient(s) numérique(s) voyez-vous à l'algorithme de simulation de $\Phi$ ?

**7** On suppose que $M$ est muni d'une relation d'ordre notée $\leq$, et qu'il existe deux éléments $x_{\min}$ et $x_{\max}$ de $M$ tels que :

$$\forall x \in M,\, x_{\min} \leq x \leq x_{\max}.$$

On suppose de plus que l'application $f$ est monotone au sens suivant :

$$\forall \theta \in \Theta, \forall x, y \in M, x \leq y \implies f(\theta, x) \leq f(\theta, y).$$

Dans ce cas, on dit que *la dynamique* (1) *est monotone.*

**7.a** Comment calculer en pratique le temps de coalescence pour une dynamique monotone ? Décrire l'algorithme complet sous ces hypothèses. *Indication : on pourra considérer $D_n(x_{\min})$ et $D_n(x_{\max})$.*

**7.b** On suppose que la chaîne de Markov $(X_n)_{n\geq 0}$ est irréductible et de mesure invariante $\pi$. On note $Y_n = G_n(x_{\min})$. Justifier la suite d'inégalités :

$$\begin{aligned}
\mathbb{P}(T > n) &= \mathbb{P}(D_n \text{ n'est pas une application constante}),\\
&= \mathbb{P}(G_n \text{ n'est pas une application constante}),\\
&\leq \mathbb{P}(\forall k \in \{1, \dots, n\},\, Y_k \neq x_{\max}),\\
&\leq \mathbb{P}(S > n),
\end{aligned}$$

où $S = \inf\{n \geq 0,\, Y_n = x_{\max}\}$. En déduire que $\mathbb{P}(T = \infty) = 0$.

**8** Dans cette question, on considère à nouveau l'exemple du modèle d'Ising introduit dans la question 2. On introduit la relation d'ordre (partielle) suivante sur $M$ :

$$x \leq y \iff \forall i \in \{1, \dots, N\}, x_i \leq y_i.$$

On suppose $J_{i,j} \geq 0$. Montrer que la dynamique introduite dans la question 2 est monotone. En déduire une méthode de simulation exacte du modèle d'Ising.

## Convergence du recuit simulé

Soit une fonction $V : \mathbb{T} \to \mathbb{R}$ de classe $\mathcal{C}^2$, où $\mathbb{T} = \mathbb{R}/\mathbb{Z}$ désigne le tore de dimension 1 (autrement dit, $V \in \mathcal{C}^2(\mathbb{R}, \mathbb{R})$ et est périodique de période 1). On cherche les minima de la fonction $V$, et on considère pour cela la dynamique suivante :

$$dX_t = -\frac{1}{2}\frac{dV}{dx}(X_t)\, dt + \sigma(t) dB_t \tag{2}$$

avec $\sigma$ une fonction du temps, scalaire, positive, décroissante et qui tend vers 0 quand $t \to \infty$ et $B_t$ un mouvement brownien de dimension 1. On note $\mathcal{M}$ l'ensemble des minima (globaux) de la fonction $V$, et l'objectif est de chercher une bonne loi de décroissance pour $\sigma(t)$ de manière à faire converger $X_t$ vers $\mathcal{M}$.

Pour $\sigma > 0$ donné, on introduit la densité de probabilité $q_\sigma = Z_\sigma^{-1} \exp(-V/\sigma^2)$, où $Z_\sigma = \int_{\mathbb{T}} \exp(-V/\sigma^2)$.

**1** On admet que $X_t$ admet une densité $p(t, \cdot)$ par rapport à la mesure de Lebesgue sur $\mathbb{T}$. Montrer que $p$ satisfait l'équation aux dérivées partielles :

$$\frac{\partial p(t,x)}{\partial t} = \frac{\sigma(t)^2}{2} \frac{\partial}{\partial x} \left( q_{\sigma(t)}(x) \frac{\partial}{\partial x} \left( \frac{p(t,x)}{q_{\sigma(t)}(x)} \right) \right). \tag{3}$$

**2** Pour $\sigma > 0$ fixé, on introduit *le trou spectral* pour la mesure $q_\sigma$ défini par :

$$\lambda_1(\sigma) = \frac{\sigma^2}{2} \inf \left\{ \frac{\int_{\mathbb{T}} \phi'^2\, q_\sigma}{\int_{\mathbb{T}} \phi^2\, q_\sigma},\, \phi : \mathbb{T} \to \mathbb{R} \text{ tel que } \int_{\mathbb{T}} (\phi^2 + \phi'^2)\, q_\sigma < \infty \text{ et } \int_{\mathbb{T}} \phi\, q_\sigma = 0 \right\}. \tag{4}$$

**2.a** Démontrer que pour toute fonction $\phi : \mathbb{T} \to \mathbb{R}$, telle que $\int_{\mathbb{T}}(\phi^2 + \phi'^2) < \infty$ et $\int_{\mathbb{T}} \phi\, q_\sigma = 0$, $\int_{\mathbb{T}} \phi^2 \leq \frac{1}{2} \int_{\mathbb{T}} (\phi')^2$. *Indication : on pourra vérifier que* $\exists x_0,\ \phi(x_0) = 0$, *poser* $\psi(x) = \phi(x_0 + x)$ *puis écrire* $\psi(x)^2 = \left(\int_0^x \psi'(y)\, dy\right)^2 \leq x \int_0^x \psi'(y)^2\, dy$.

**2.b** En déduire que $\lambda_1(\sigma) \geq \sigma^2 \exp(-\mathrm{osc}(V)/\sigma^2)$, où $\mathrm{osc}(V) = \max_{\mathbb{T}}(V) - \min_{\mathbb{T}}(V)$.

**3** On suppose dans cette question que $\sigma(t)$ est une fonction constante : $\sigma(t) = \sigma_0$. On introduit $f(t) = \frac{1}{2} \int_{\mathbb{T}} \left( \frac{p(t,\cdot)}{q_{\sigma_0}} - 1 \right)^2 q_{\sigma_0}$. Montrer que $\frac{df}{dt} = -\frac{\sigma_0^2}{2} \int_{\mathbb{T}} \left( \frac{\partial}{\partial x} \left( \frac{p(t,\cdot)}{q_{\sigma_0}} \right) \right)^2 q_{\sigma_0}$. En déduire que $f(t) \leq f(0) \exp(-2\lambda_1(\sigma_0) t)$ et donc que $p(t, \cdot)$ converge vers $q_{\sigma_0}$ en norme $L^1(\mathbb{T})$ à vitesse exponentielle.

**4** On suppose désormais que
$$\sigma^{-2}(t) = \frac{\ln(t+2)}{a},$$
où $a > \operatorname{osc}(V)$. On introduit la fonction $e(t) = \frac{1}{2}\int_{\mathbb{T}} \left(\frac{p(t,\cdot)}{q_{\sigma(t)}} - 1\right)^2 q_{\sigma(t)}$. Vérifier que $e(t) = \frac{1}{2}\left(\int_{\mathbb{T}} \frac{p(t,\cdot)^2}{q_{\sigma(t)}} - 1\right)$. Montrer que
$$\frac{de}{dt} = \int_{\mathbb{T}} \frac{\partial p}{\partial t}\frac{p}{q_\sigma} - \frac{1}{2}\int_{\mathbb{T}} \frac{\partial q_\sigma}{\partial t}\left(\frac{p}{q_\sigma}\right)^2. \tag{5}$$
En s'inspirant de la question 3, vérifier que
$$\int_{\mathbb{T}} \frac{\partial p}{\partial t}\frac{p}{q_\sigma} \leq -2\lambda_1(\sigma(t))e(t). \tag{6}$$
Montrer que
$$\frac{\partial q_\sigma}{\partial t} = \frac{d\sigma^{-2}(t)}{dt}\left(\int_T V q_\sigma - V\right) q_\sigma$$
et en déduire que
$$-\frac{1}{2}\int_{\mathbb{T}} \frac{\partial q_\sigma}{\partial t}\left(\frac{p}{q_\sigma}\right)^2 \leq \frac{\operatorname{osc}(V)}{2a(t+2)}\left(2e+1\right). \tag{7}$$
En utilisant (5), (6) et (7), en déduire que
$$\frac{de}{dt} \leq -\alpha(t)e(t) + \beta(t)$$
avec
$$\alpha(t) = 2\lambda_1(\sigma(t)) - \frac{\operatorname{osc}(V)}{a(t+2)}$$
et
$$\beta(t) = \frac{\operatorname{osc}(V)}{2a(t+2)}.$$

**5** Montrer que $\int_0^\infty \alpha(t)\,dt = \infty$ et que $\lim_{t\to\infty} \frac{\beta(t)}{\alpha(t)} = 0$. En déduire que $\lim_{t\to\infty} e(t) = 0$. *Indication : on pourra considérer* $\frac{d}{dt}\left(\exp\left(\int_0^t \alpha(s)\,ds\right)e(t)\right)$ *et montrer que, pour* $t_0$ *assez grand et* $0 \leq t_0 \leq t$,
$$e(t) \leq \left(\int_0^{t_0} \exp\left(\int_0^s \alpha(r)\,dr\right)\beta(s)\,ds\right)\exp\left(-\int_0^t \alpha(s)\,ds\right) + \sup\left(\frac{\beta(t)}{\alpha(t)},\, t \geq t_0\right).$$

**6** Soit $\mathcal{M}^\delta$ un $\delta$-voisinage de $\mathcal{M}$. Vérifier que
$$\begin{aligned} e(t) &\geq \frac{1}{2}\int_{\mathcal{M}^\delta} \left(\frac{p(t,\cdot)}{q_{\sigma(t)}} - 1\right)^2 q_{\sigma(t)}, \\ &\geq \frac{1}{2}\left(\int_{\mathcal{M}^\delta} |p(t,\cdot) - q_{\sigma(t)}|\right)^2, \\ &\geq \frac{1}{2}\left|\mathbb{P}(X_t \in \mathcal{M}^\delta) - \int_{\mathcal{M}^\delta} q_{\sigma(t)}\right|^2. \end{aligned}$$
En déduire que $\lim_{t\to\infty} \operatorname{dist}(X_t, \mathcal{M}) = 0$ en probabilité.