Examen du cours de M2 :
Méthodes numériques probabilistes

13 janvier 2016, 9h30-12h30.

Les notes de cours sont autorisées.

Exercice : Equation de Langevin

On considère l'équation de Langevin

$$\begin{cases} dq_t = v_t\,dt \\ m\,dv_t = -\nabla\pi(q_t)\,dt - v_t\,dt + \sqrt{2\beta^{-1}}dW_t \end{cases}$$

avec m et β deux paramètres strictement positifs, $(q_t, v_t)_{t\geq 0}$ un processus à valeurs dans $\mathbb{R}^d \times \mathbb{R}^d$, $(W_t)_{t\geq 0}$ un brownien d-dimensionnel et $\pi : \mathbb{R}^d \to \mathbb{R}$ une fonction régulière. On suppose dans toute la suite que $\int_{\mathbb{R}^d} \exp(-\beta\pi(q))\,dq < \infty$ et que $\nabla\pi$ est une fonction lipschitzienne.

1 On s'intéresse tout d'abord à la mesure stationnaire pour ce processus.

1.1 Ecrire l'équation de Fokker-Planck qui donne l'évolution de la loi de (q_t, v_t).

1.2 Montrer que la mesure de probabilité sur $\mathbb{R}^d \times \mathbb{R}^d$ de densité

$$f(q,v) = Z^{-1}\exp\left(-\beta\left(\pi(q) + \frac{m|v|^2}{2}\right)\right)$$

où $Z = \int_{\mathbb{R}^d\times\mathbb{R}^d} \exp\left(-\beta\left(\pi(q) + \frac{m|v|^2}{2}\right)\right) dqdv < \infty$, est invariante pour la dynamique de Langevin.

2 Dans cette question, on s'intéresse à la limite $m \to 0$.

2.1 Montrer que pour tout $t \geq 0$,

$$v_t = v_0\,e^{-\frac{t}{m}} + \frac{1}{m}\int_0^t e^{-\frac{t-s}{m}}(-\nabla\pi(q_s)\,ds + \sqrt{2\beta^{-1}}dW_s).$$

2.2 En déduire que pour tout $t \geq 0$,

$$q_t = q_0 + v_0 m\left(1 - e^{-\frac{t}{m}}\right) + \int_0^t \left(1 - e^{-\frac{t-r}{m}}\right)(-\nabla\pi(q_r)\,dr + \sqrt{2\beta^{-1}}dW_r).$$

On introduit le processus X_t à valeurs dans $\mathbb{R}^d$ et solution de

$$dX_t = -\nabla\pi(X_t)\,dt + \sqrt{2\beta^{-1}}dW_t$$

avec comme condition initiale $X_0 = q_0$. Montrer que pour tout $t \geq 0$,

$$q_t - X_t = v_0 m\left(1 - e^{-\frac{t}{m}}\right) - \int_0^t \left(1 - e^{-\frac{t-r}{m}}\right)(\nabla\pi(q_r) - \nabla\pi(X_r))\,dr$$
$$+ \int_0^t e^{-\frac{t-r}{m}}\nabla\pi(X_r)\,dr - \sqrt{2\beta^{-1}}\int_0^t e^{-\frac{t-r}{m}}\,dW_r.$$

2.3 Montrer que

$$\int_0^t e^{-\frac{t-r}{m}} dW_r = W_t e^{-\frac{t}{m}} - \frac{1}{m}\int_0^t e^{-\frac{t-r}{m}} (W_r - W_t)\, dr.$$

2.4 *(Question plus difficile)* En déduire que pour tout $T > 0$, presque sûrement,

$$\lim_{m\to 0} \sup_{t\in[0,T]} \left| \int_0^t e^{-\frac{t-r}{m}} dW_r \right| = 0.$$

2.5 Montrer qu'il existe $C > 0$ tel que, pour $T > 0$ fixé : presque sûrement

$$\forall t \in [0,T],\ \sup_{s\le t} |q_s - X_s| \le C \int_0^t \sup_{r\le s} |q_r - X_r|\, ds + r_T(m)$$

avec $\lim_{m\to 0} |r_T(m)| = 0$. En déduire que presque sûrement

$$\lim_{m\to 0} \sup_{t\in[0,T]} |q_t - X_t| = 0.$$

Problème : Modèles de spins

On considère un modèle de spins sur le réseau $\{1,\dots,N\}$: à chaque point $i \in \{1,\dots,N\}$, on associe un spin $x_i \in \{-1,+1\}$. Une configuration est donnée par $x = (x_1,\dots x_N)$ et l'espace des configurations est

$$\mathcal{X}_N = \{-1,+1\}^N.$$

La magnétisation d'une configuration est définie par :

$$\forall x \in \mathcal{X}_N,\ m_N(x) = \frac{1}{N}\sum_{i=1}^N x_i$$

et $m_N \in \mathcal{M}_N$ avec

$$\mathcal{M}_N = \left\{-1, -1+\frac{2}{N}, \dots, 1-\frac{2}{N}, 1\right\}.$$

Les modèles de spins ont pour objectif d'expliquer le ferromagnétisme, c'est-à-dire le fait que certains matériaux sont magnétisés (leur magnétisation est non nulle), même en l'absence d'un champ magnétique extérieur.

1 Dans cette première partie, on suppose que les spins interagissent avec leurs plus proches voisins : c'est le modèle d'Ising. On se donne la fonction potentiel :

$$\forall x \in \mathcal{X}_N,\ V(x) = -\sum_{i=1}^N x_i x_{i+1}$$

avec comme convention $x_{N+1} = x_1$ (conditions aux limites périodiques). La mesure de Boltzmann associée donne à la configuration $x \in \mathcal{M}_N$ la probabilité

$$\nu_N(x) = (Z_N^{\text{Ising}})^{-1} \exp(-\beta V(x))$$

avec $\beta > 0$ un paramètre positif et $Z_N^{\text{Ising}} = \sum_{x\in\mathcal{X}_N} \exp(-\beta V(x))$. On considère l'algorithme suivant : $X^0 \in \mathcal{X}_N$ est donné, et, pour $k \ge 0$, X^{k+1} est construit à partir de X^k de la manière suivante :

— On tire un site I^k au hasard uniformément parmi $\{1, \dots N\}$ et on considère

$$Y^k = (X_1^k, \dots, X_{I^k-1}^k, -X_{I^k}^k, X_{I^k+1}^k, \dots X_N^k).$$

— On calcule $\alpha^k = \min\Big(1, \exp\big(-\beta(V(Y^k) - V(X^k))\big)\Big)$ et on tire une variable aléatoire U^k uniformément sur $[0,1]$. Si $U^k \le \alpha^k$ on pose $X^{k+1} = Y^k$, sinon, on pose $X^{k+1} = X^k$.

1.1 En utilisant des résultats du cours, rappeler pourquoi cet algorithme est ergodique par rapport à la mesure ν_N au sens suivant : pour toute fonction $\varphi : \mathcal{X}_N \to \mathbb{R}$, presque sûrement,

$$\lim_{K\to\infty} \frac{1}{K} \sum_{k=1}^{K} \varphi(X^k) = \sum_{x\in\mathcal{X}_N} \varphi(x)\nu_N(x).$$

On note $\tau_x = \inf\{n \ge 1, X^n = x$ pour la chaîne de condition initiale $X^0 = x\}$, pour tout $x \in \mathcal{X}_N$. On suppose N pair et on introduit les points de $\mathcal{X}_N$: $x_- = (-1, \dots, -1)$, $x_0 = (-1, +1, \dots, -1, +1)$, et $x_+ = (+1, \dots, +1)$. Comparer $\mathbb{E}(\tau_{x_-})$, $\mathbb{E}(\tau_{x_+})$ et $\mathbb{E}(\tau_{x_0})$.

1.2 Proposer une modification de l'algorithme pour échantillonner la mesure

$$\nu_N^m(x) = (Z_N^m)^{-1} \exp(-\beta V(x)) 1_{\{m_N(x)=m\}}$$

où la magnétisation moyenne $m \in \mathcal{M}_N$ est fixée, et $Z_N^m = \sum_{x\in\mathcal{X}_N} \exp(-\beta V(x)) 1_{\{m_N(x)=m\}}$.

2 Dans cette deuxième partie, on suppose que les spins interagissent tous entre eux selon la fonction potentiel (modèle de Curie-Weiss) :

$$\forall x \in \mathcal{X}_N,\ H(x) = -\frac{1}{2N} \sum_{1\le i,j\le N} x_i x_j.$$

La mesure de Boltzmann associée donne à la configuration $x \in \mathcal{M}_N$ la probabilité

$$\mu_N(x) = (Z_N)^{-1} \exp(-\beta H(x))$$

avec $\beta > 0$ un paramètre positif et $Z_N = \sum_{x\in\mathcal{X}_N} \exp(-\beta H(x))$. On s'intéresse à la loi de la magnétisation moyenne $m_N(X)$ quand X est distribué selon μ_N, noté $X \sim \mu_N$ dans la suite.

2.1 Montrer que, pour $X \sim \mu_N$,

$$\forall m \in \mathcal{M}_N(x),\ \mathbb{P}(m_N(X) = m) = (Z_N)^{-1} \binom{N}{\frac{N(1-m)}{2}} \exp\left(\beta \frac{N}{2} m^2\right),$$

où pour $n \ge 1$ et $k \in \{0, \dots, n\}$, $\binom{n}{k} = \frac{n!}{k!(n-k)!}$ dénote le coefficient binomial.

2.2 *(Question optionnelle)* L'objectif de cette question est d'exhiber des bornes sur le coefficient binomial $\binom{n}{k}$. En utilisant la formule $\sum_{\ell=0}^{n} \binom{n}{\ell} p^\ell (1-p)^{n-\ell} = 1$ (valable pour tout p), montrer que

$$\forall k \in \{0, \dots n\},\ \frac{1}{n+1} \exp\left(nh\left(\frac{k}{n}\right)\right) \le \binom{n}{k} \le \exp\left(nh\left(\frac{k}{n}\right)\right)$$

où, pour tout $x \in [0,1]$, $h(x) = -x \ln x - (1-x)\ln(1-x)$. *Indication : pour* $p \in [0,1]$, *écrire* $\binom{n}{k} p^k (1-p)^{n-k} \leq \sum_{\ell=0}^{n} \binom{n}{\ell} p^\ell (1-p)^{n-\ell} \leq (n+1) \max_{\ell \in \{0,\dots,n\}} \binom{n}{\ell} p^\ell (1-p)^{n-\ell}$ *et considérer* $p = k/n$.

2.3 Montrer que pour $X \sim \mu_N$,

$$\forall m \in \mathcal{M}_N, \; \frac{1}{N+1} \exp(N\psi(m)) \leq Z_N \, \mathbb{P}(m_N(X) = m) \leq \exp(N\psi(m))$$

où, pour tout $m \in [-1,1]$, $\psi(m) = \frac{\beta}{2} m^2 + h\left(\frac{1-m}{2}\right)$.

2.4 On note $\psi^* = \sup_{m \in [-1,1]} \psi(m)$. Montrer que le suprémum est atteint en un point $m^* \in (-1,1)$ tel que

$$m^* = \tanh(\beta m^*)$$

où pour tout $x \in \mathbb{R}$, $\tanh(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}}$. Montrer que pour $\beta \leq 1$, le suprémum est atteint en un unique point $m^* = 0$, alors que pour $\beta > 1$, le suprémum est atteint en deux points $\{m^*_-, m^*_+\}$ avec $m^*_- = -m^*_+$ et $m^*_+ > 0$.

2.5 Montrer que pour tout $N \geq 1$,

$$\frac{1}{N} \ln Z_N \leq \psi^* + \frac{1}{N} \ln(1+N)$$

2.6 Montrer que qu'il existe $C_0 > 0$ et N_0 tel que pour tout $N \geq N_0$,

$$\frac{1}{N} \ln Z_N \geq \psi^* - \frac{1}{N} \ln(1+N) - \frac{C_0}{N^2}$$

2.7 On suppose $\beta \leq 1$ et $X \sim \mu_N$. Montrer que pour tout $\epsilon > 0$, il existe $C_1 > 0$ et N_1 tel que pour tout $N \geq N_1$

$$\mathbb{P}(|m_N(X)| \geq \epsilon) \leq \exp(-C_1 N).$$

2.8 On suppose $\beta > 1$ et $X \sim \mu_N$. Rappeler pourquoi pour tout $\epsilon > 0$, $\mathbb{P}(|m_N(X) - m^*_-| \geq \epsilon) = \mathbb{P}(|m_N(X) - m^*_+| \geq \epsilon)$. Montrer que pour tout $\epsilon > 0$, il existe $C_2 > 0$ et N_2 tel que pour tout $N \geq N_2$

$$\frac{1}{2} \leq \mathbb{P}(|m_N(X) - m^*_+| \geq \epsilon) \leq \frac{1}{2} + \exp(-C_2 N).$$

2.9 Déduire des questions précédentes que pour $X \sim \mu_N$, dans la limite $N \to \infty$, la loi de $m_N(X)$ converge vers δ_0 si $\beta \leq 1$ et vers $\frac{1}{2}\left(\delta_{m^*_-} + \delta_{m^*_+}\right)$ si $\beta > 1$. Commenter.