Examen du cours de M2 :
Méthodes numériques probabilistes

12 janvier 2009, 14h30 - 17h30.

Les notes de cours sont autorisées.

Exercice : chaîne de Markov sur $\{0,1\}^n$

On munit l'ensemble $E = \{0,1\}^n$ (où $n \geq 1$) de la distance : $\forall (x,y) \in E^2$,

$$d(x,y) = \sum_{i=1}^{n} (1 - 1_{\{x_i = y_i\}}),$$

où $x = (x_1, \dots, x_n)$ et $y = (y_1, \dots, y_n)$. On note $\mathcal{V}(x) = \{y \in E,\, d(x,y) \leq 1\}$. On considère la matrice de transition : $\forall (x,y) \in E^2$,

$$M(x,y) = \frac{1}{|\mathcal{V}(x)|}\; 1_{\{y \in \mathcal{V}(x)\}},$$

où $|\mathcal{V}(x)|$ désigne le cardinal de $\mathcal{V}(x)$. Vérifier que pour tout couple de points $(x,y) \in E^2$, on a $M^n(x,y) \geq 1/(n+1)^n$. Que peut-on en déduire sur la chaîne de Markov de matrice de transition M ?

Problème : discrétisation d'équations différentielles stochastiques

Soit $(\Omega, \mathcal{A}, \mathcal{F}_t, \mathbb{P})$ un espace de probabilité muni d'une filtration. On se donne des coefficients $b : [0,T] \times \mathbb{R} \to \mathbb{R}$ et $\sigma : [0,T] \times \mathbb{R} \to \mathbb{R}$ et on s'intéresse à l'équation différentielle stochastique

$$\begin{aligned} dX_t &= b(t, X_t)\, dt + \sigma(t, X_t)\, dW_t, \\ X_0 &= Y, \end{aligned} \tag{1}$$

où W_t est un $\mathcal{F}_t$-mouvement brownien unidimensionnel, et Y est une variable aléatoire de carré intégrable et $\mathcal{F}_0$-mesurable.

On suppose dans toute la suite que

$$\exists K > 0, \forall t \in [0,T], \begin{cases} \forall x \in \mathbb{R}, |b(t,x)| + |\sigma(t,x)| \leq K(1 + |x|), \\ \forall x, y \in \mathbb{R}, |b(t,x) - b(t,y)| + |\sigma(t,x) - \sigma(t,y)| \leq K|x - y|. \end{cases} \tag{2}$$

Nous avons vu en cours que sous ces hypothèses, l'équation différentielle stochastique (1) admet une unique solution forte.

1. L'objectif de cette question est de montrer qu'il existe $C > 0$ (indépendant de Y) tel que :

$$\mathbb{E}\left(\sup_{t \in [0,T]} |X_t|^2 \right) \leq C(1 + \mathbb{E}(|Y|^2)). \tag{3}$$

Nous utiliserons pour cela l'inégalité de Doob suivante (admise) : il existe $C > 0$ tel que pour tout processus $(H_t)_{t \geq 0}$ adapté et tel que, pour tout $t \geq 0$, $\int_0^t |H_r|^2\, dr < \infty$ p.s.,

$$\forall t \geq 0, \mathbb{E}\left(\sup_{s \leq t} \left| \int_0^s H_r dW_r \right|^2 \right) \leq C \int_0^t \mathbb{E}\left(|H_r|^2\right)\, dr.$$

1.a Pour $M > 0$, on introduit $\nu_M = \inf\{t \geq 0, |X_t| \geq M\}$ (avec la convention $\inf \emptyset = T$). Montrer qu'il existe une constante C indépendante de M telle que

$$\mathbb{E}\left(\sup_{0\leq s\leq T} |X_{s\wedge\nu_M}|^2\right) \leq C\left(1 + \mathbb{E}(|Y|^2)\right).$$

On rappelle que pour deux réels a et b, $a \wedge b = \min(a, b)$.

1.b En utilisant le Lemme de Fatou, en déduire (3).

1.c Montrer qu'il existe $C > 0$ (indépendante de la condition initiale, mais dépendant de T) tel que, pour $0 \leq s < t \leq T$,

$$\mathbb{E}(|X_t - X_s|^2) \leq C(1 + \mathbb{E}(|Y|^2))(t - s).$$

On admettra dans la suite que (3) se généralise à des normes L^p sous la forme : pour tout $p \geq 1$, il existe $C > 0$ (C indépendant de Y) tel que

$$\mathbb{E}\left(\sup_{t\in[0,T]} |X_t|^{2p}\right) \leq C(1 + \mathbb{E}(|Y|^{2p})). \tag{4}$$

2. On introduit maintenant le schéma d'Euler pour discrétiser l'équation différentielle stochastique (1). L'intervalle $[0, T]$ est subdivisé en N sous-intervalles de longueur $\Delta t = T/N$. On note $t_k = k\Delta t$. On introduit $\overline{X}_{t_k}$ l'approximation de X_{t_k} obtenue par le schéma d'Euler : $\forall k \in \{0, \dots, N-1\}$,

$$\begin{aligned} \overline{X}_{t_{k+1}} &= \overline{X}_{t_k} + b(t_k, \overline{X}_{t_k})\Delta t + \sigma(t_k, \overline{X}_{t_k})\,(W_{t_{k+1}} - W_{t_k}), \\ \overline{X}_0 &= Y. \end{aligned} \tag{5}$$

Pour la preuve de convergence, il sera utile d'utiliser le processus interpolé $(\overline{X}_t)_{0\leq t\leq T}$ défini par :

$$\forall k \in \{0, \dots, N-1\}, \forall t \in [t_k, t_{k+1}], \overline{X}_t = \overline{X}_{t_k} + b(t_k, \overline{X}_{t_k})(t - t_k) + \sigma(t_k, \overline{X}_{t_k})\,(W_t - W_{t_k}).$$

Si nécessaire, nous utilisons la notation $\overline{X}_t^N$ pour indiquer explicitement la dépendance de $\overline{X}_t$ par rapport à N.

2.a Expliquer comment construire en pratique sur ordinateur une réalisation du processus discrétisé $(\overline{X}_{t_k})_{0\leq k\leq N}$.

2.b Vérifier que $\overline{X}_t$ est solution de l'équation différentielle stochastique

$$d\overline{X}_t = b(\tau_t, \overline{X}_{\tau_t})\,dt + \sigma(\tau_t, \overline{X}_{\tau_t})dW_t$$

où $\tau_t = \left\lfloor \frac{t}{\Delta t} \right\rfloor \Delta t$ (avec $\lfloor\cdot\rfloor$ qui désigne la partie entière).

Dans la suite, on suppose que les coefficients b et σ vérifient (en plus de (2)) la condition suivante :

$$\begin{aligned} &\exists \alpha > 1/2, \exists K > 0, \forall x \in \mathbb{R}, \forall (s,t) \in [0,T]^2, \\ &|\sigma(t,x) - \sigma(s,x)| + |b(t,x) - b(s,x)| \leq K(1 + |x|)(t - s)^\alpha. \end{aligned} \tag{6}$$

L'objectif de cette section est de montrer qu'il existe $C > 0$ tel que pour tout N

$$\mathbb{E}\left(\sup_{t\in[0,T]}\left|X_t - \overline{X}_t^N\right|^2\right) \leq \frac{C}{N}. \tag{7}$$

Cette notion de convergence est appelée *convergence forte* (ou convergence trajectorielle).

2.c Rappeler ce que devient ce résultat si $\sigma = 0$.

2.d Montrer que

$$\mathbb{E}\left(\sup_{u\leq t}|X_u - \overline{X}_u|^2\right) \leq C\left(t\int_0^t \mathbb{E}\left(|b(s,X_s) - b(\tau_s,\overline{X}_{\tau_s})|^2\right) ds + \int_0^t \mathbb{E}\left(|\sigma(s,X_s) - \sigma(\tau_s,\overline{X}_{\tau_s})|^2\right) ds\right).$$

En utilisant (2)–(6), en déduire que

$$\mathbb{E}\left(\sup_{u\leq t}|X_u - \overline{X}_u|^2\right) \leq C\left(\frac{1}{N} + \int_0^t \mathbb{E}\left(\sup_{u\leq s}|X_u - \overline{X}_u|^2\right) ds\right)$$

et conclure.

On admettra dans la suite que le résultat de convergence forte (7) se généralise à des normes L^p sous la forme : pour tout $p \geq 1$, il existe $C > 0$ tel que pour tout N

$$\mathbb{E}\left(\sup_{t\in[0,T]}\left|X_t - \overline{X}_t^N\right|^{2p}\right) \leq \frac{C}{N^p}. \tag{8}$$

3 Soit f une fonction test. L'objectif de cette section est de donner une estimation de l'erreur $\left|\mathbb{E}(f(X_T)) - \mathbb{E}(f(\overline{X}_T^N))\right|$.

3.a En supposant f lipschitzienne et en utilisant le résultat de la section précédente, que peut-on dire de $\left|\mathbb{E}(f(X_T)) - \mathbb{E}(f(\overline{X}_T^N))\right|$?

3.b Expliquer comment construire numériquement une approximation de $\mathbb{E}(f(\overline{X}_T^N))$.

Nous allons montrer qu'en fait, si b, σ et f sont des fonctions suffisamment régulières

$$\left|\mathbb{E}(f(X_T)) - \mathbb{E}(f(\overline{X}_T^N))\right| \leq \frac{C}{N}.$$

Cette notion de convergence est appelée *convergence faible*. Pour cela, nous introduisons $u : [0,T]\times\mathbb{R} \to \mathbb{R}$ solution de l'équation aux dérivées partielles :

$$\begin{cases} \dfrac{\partial u}{\partial t} + \dfrac{\sigma^2}{2}\dfrac{\partial^2 u}{\partial x^2} + b\dfrac{\partial u}{\partial x} = 0, \\ u(T,x) = f(x). \end{cases} \tag{9}$$

On suppose dans la suite de cette section que b et σ sont de classe $\mathcal{C}^4$ avec des dérivées bornées, et que f est de classe $\mathcal{C}^4$, avec des dérivées à croissance polynomiale : $\forall n \in \{0,\ldots,4\}$, $\exists C, p > 0$, $\forall x \in \mathbb{R}$,

$$\left|\frac{d^n f}{dx^n}(x)\right| \leq C(1 + |x|^p).$$

On admet que sous ces hypothèses, la solution u de (9) est de classe $\mathcal{C}^4$, avec des dérivées à croissance polynomiale.

On suppose dans la suite de cette section que la condition initiale de l'équation différentielle stochastique (1) est déterministe : $Y = y$ où $y \in \mathbb{R}$.

3.c Montrer que $u(0, y) = \mathbb{E}(f(X_T))$.

3.d On introduit

$$\mathcal{E}_k = \mathbb{E}\left(u(t_{k+1}, \overline{X}_{t_{k+1}}) - u(t_k, \overline{X}_{t_k})\right).$$

Montrer que

$$\mathbb{E}(f(\overline{X}_T^N)) - \mathbb{E}(f(X_T)) = \sum_{k=0}^{N-1} \mathcal{E}_k.$$

3.e Montrer que

$$\mathcal{E}_k = \int_{t_k}^{t_{k+1}} \mathbb{E}\left(e(t, \overline{X}_t)\right) \, dt$$

où

$$e(t,x) = \frac{\partial u}{\partial t}(t,x) + \frac{\sigma^2}{2}(\tau_t, \overline{X}_{\tau_t}) \frac{\partial^2 u}{\partial x^2}(t,x) + b(\tau_t, \overline{X}_{\tau_t}) \frac{\partial u}{\partial x}(t,x).$$

3.f Vérifier que $\forall k \in \{0, \ldots, N-1\}$, $\forall t \in [t_k, t_{k+1}]$,

$$\mathbb{E}\left(e(t, \overline{X}_t)\right) = \int_{t_k}^{t} \mathbb{E}\left(\frac{\partial e}{\partial t}(s, \overline{X}_s) + \frac{\sigma^2}{2}(t_k, \overline{X}_{t_k}) \frac{\partial^2 e}{\partial x^2}(s, \overline{X}_s) + b(t_k, \overline{X}_{t_k}) \frac{\partial e}{\partial x}(s, \overline{X}_s)\right) ds$$

En déduire que

$$|\mathcal{E}_k| \leq \frac{C}{N^2}$$

et conclure.

3.g Adapter la preuve précédente pour montrer qu'on peut obtenir un développement limité de l'erreur :

$$\mathbb{E}(f(\overline{X}_T^N)) - \mathbb{E}(f(X_T)) = \frac{C}{N} + O\left(\frac{1}{N^2}\right)$$

En itérant le procédé, on peut démontrer (sous des hypothèses de régularité suffisante sur f, b et σ) que

$$\mathbb{E}(f(\overline{X}_T^N)) - \mathbb{E}(f(X_T)) = \sum_{k=1}^{p} \frac{C_k}{N^k} + O\left(\frac{1}{N^{k+1}}\right).$$

3.h En déduire une méthode numérique pour construire une approximation de $\mathbb{E}(f(X_T))$ d'ordre $(1/N)^2$, et $(1/N)^p$ pour tout $p \geq 2$. Commenter l'intérêt pratique de ces méthodes.

4 Dans cette dernière partie, nous revenons sur la notion de convergence forte étudiée dans la partie 2. Nous supposons dans cette partie que b et σ ne dépendent pas du temps, pour simplifier ($b : \mathbb{R} \to \mathbb{R}$ et $\sigma : \mathbb{R} \to \mathbb{R}$). La fonction b est supposée de classe $\mathcal{C}^2$ avec des dérivées première et seconde bornées. la fonction σ est supposée de classe $\mathcal{C}^1$.

Dans un premier temps, nous supposons de plus que σ est une constante et, sans perte de généralité, que

$$\sigma = 1.$$

Nous allons montrer que dans ce cas (comparer avec (7))

$$\mathbb{E}\left(\sup_{t\in[0,T]} \left|X_t - \overline{X}_t^N\right|\right) \leq \frac{C}{N}.$$

4.a Montrer que $\forall t \in [0,T]$,

$$|X_t - \overline{X}_t| \leq \sup|b'| \int_0^t |X_{\tau_s} - \overline{X}_{\tau_s}|\, ds + \left|\int_0^t (b(X_s) - b(X_{\tau_s}))\, ds\right|.$$

4.b Vérifier que pour $k \in \{0,\dots,N-1\}$,

$$\int_{t_k}^{t_{k+1}} (b(X_s)-b(X_{t_k}))\, ds = \int_{t_k}^{t_{k+1}} (t_{k+1}-r)\left(b\,b' + \frac{b''}{2}\right)(X_r)\, dr + \int_{t_k}^{t_{k+1}} (t_{k+1}-r) b'(X_r)\, dW_r.$$

En déduire que

$$\mathbb{E}\left(\sup_{u\in[0,t]} \left|\int_0^u (b(X_s) - b(X_{\tau_s}))\, ds\right|\right) \leq \frac{C}{N}\left(\int_0^t \mathbb{E}\left|\left(b\,b' + \frac{b''}{2}\right)(X_r)\right| dr + \sqrt{\int_0^t \mathbb{E}\big((b'(X_r))^2\big)\, dr}\right),$$

et conclure.

4.c Dans cette dernière question, on suppose que $\sigma : \mathbb{R} \to \mathbb{R}$ n'est plus une fonction constante. Au vu des questions précédentes, on comprend que ce qui limite le taux de convergence forte (7), c'est la discrétisation du terme de diffusion et non pas celle du terme de drift. On introduit alors le schéma de Milshtein : $\forall k \in \{0,\dots,N-1\}$,

$$\begin{aligned}\tilde{X}_{t_{k+1}} &= \tilde{X}_{t_k} + b(\tilde{X}_{t_k})\Delta t + \sigma(\tilde{X}_{t_k})\,(W_{t_{k+1}} - W_{t_k}) + \frac{1}{2}\sigma\sigma'(\tilde{X}_{t_k})\,\big((W_{t_{k+1}} - W_{t_k})^2 - \Delta t\big),\\ \tilde{X}_0 &= Y.\end{aligned}$$

En raisonnant *formellement* sur l'erreur de consistance sur un pas de temps, justifier que l'ordre de convergence forte de ce schéma est 1 et non plus 1/2 :

$$\sup_{t\in[0,T]} \mathbb{E}\left(\left|X_t - \tilde{X}_t^N\right|^2\right) \leq \frac{C}{N^2}.$$