Examen du cours MOPSI

22 février 2008, 08h30-12h00.

Les notes de cours (notes manuscrites et polycopiés) sont autorisées.

L'examen comporte deux exercices et un problème indépendants. Nous vous demandons de rédiger les réponses sur deux copies distinctes :
- *sur la copie "Analyse numérique", rédiger les réponses à l'exercice 1, et aux parties I et II du problème,*
- *sur la copie "Probabilités", rédiger les réponses à l'exercice 2, et à la partie III du problème.*

Exercice 1 (copie "Analyse numérique").

On considère le problème unidimensionnel : pour $\varepsilon > 0$, trouver $u^\varepsilon : [0,1] \to \mathbb{R}$ tel que

$$\begin{cases} -\frac{d}{dx}\left(a\left(\frac{x}{\varepsilon}\right)\frac{d}{dx}u^\varepsilon\right) = f, \\ u^\varepsilon(0) = u^\varepsilon(1) = 0. \end{cases} \tag{1}$$

On suppose que $f : [0,1] \to \mathbb{R}$ est une fonction bornée et que $a : \mathbb{R} \to \mathbb{R}$ est une fonction 1-périodique telle que, $\forall x \in [0,1]$

$$m \leq a(x) \leq M,$$

pour deux constantes $m > 0$ et $M < +\infty$.

1 Donner un cadre fonctionnel et une formulation variationnelle permettant de définir précisément la solution u^ε du problème (1).

On veut calculer une approximation de la solution u^ε par la méthode des éléments finis. On introduit un maillage uniforme du segment $[0,1]$ en I intervalles de taille $h = 1/I$. Les noeuds du maillage sont les $(x_i = ih)_{0 \leq i \leq I}$. On définit sur ce maillage l'espace d'éléments finis P1 (fonctions continues, affines par morceaux et nulles en 0 et en 1) engendré par les fonctions chapeaux $(\phi_i)_{1 \leq i \leq I-1}$ qui satisfont : $\forall i \in \{1, \ldots, I-1\}$, $\forall j \in \{0, \ldots, I\}$,

$$\phi_i(x_j) = \delta_{i,j},$$

où $\delta_{i,j}$ est le symbole de Kronecker ($\delta_{i,j} = 1$ si $i = j$ et $\delta_{i,j} = 0$ si $i \neq j$).

2 On note u_h^ε la solution du problème discrétisé par la méthode des éléments finis. On a :

$$u_h^\varepsilon(x) = \sum_{i=1}^{I-1} U_i^\varepsilon \phi_i(x)$$

où $U_i^\varepsilon = u_h^\varepsilon(x_i)$. Montrer que le problème discret s'écrit sous la forme : trouver $U^\varepsilon = (U_1^\varepsilon, \ldots, U_{I-1}^\varepsilon)$ tel que

$$A^\varepsilon U^\varepsilon = F,$$

pour une matrice A^ε et un vecteur F à préciser en fonction de ε, a, f et des fonctions ϕ_i.

3 On considère le paramètre h fixé et on fait tendre ε vers 0. Calculer la limite A^0 de la matrice A^ε.

4 Déterminer l'équation continue dont le problème discret

$$A^0 U^0 = F$$

est une approximation.

5 Que dire de la validité de cette approche ?

Exercice 2 (copie "Probabilités").

On se donne $(W_t, t \geq 0)$ un mouvement brownien (standard) et on note $(\mathcal{F}_t, t \geq 0)$ la filtration usuelle qu'il engendre. On pose pour $t \geq 0$, $X_t = \int_0^t s dW_s$ et $Y_t = -\int_0^t W_s ds$.

1 Pourquoi le processus $(X_t, t \geq 0)$ est-il bien défini ? Montrer que $(X_t, t \geq 0)$ est une $\mathcal{F}_t$-martingale continue, donner la valeur de $\mathbb{E}[X_t^2]$, puis écrire X_t comme la limite dans L^2 d'une somme discrète.

2 Montrer que pour tout $t > 0$, X_t et Y_t suivent des lois normales de paramètres que l'on déterminera. En déduire que pour tout $t > 0$ fixé, X_t et Y_t ont la même loi.

3 Montrer que les processus $(X_t, t \geq 0)$ et $(Y_t, t \geq 0)$ sont des processus gaussiens.

4 Soient $0 < t < t'$. En écrivant $X_{t'} - X_t = \int_t^{t'} s dW_s$ comme la limite d'une somme discrète, montrer que $X_{t'} - X_t$ est indépendant de X_t. Calculer ensuite $\mathbb{E}[X_t X_{t'}]$ et $\mathbb{E}[Y_t Y_{t'}]$, et donner la loi des processus $(X_t, t \geq 0)$ et $(Y_t, t \geq 0)$? Ces processus ont-ils la même loi ? Est-ce contradictoire avec la question 2 ? Déduire des calculs précédents que $(X_t, t \geq 0)$ suit la même loi que le processus $(W_{t^3/3}, t \geq 0)$.

5 Montrer que pour $t \geq 0$, $tW_t = X_t - Y_t$. Est-ce contradictoire avec la question 2 ?

6 On pose $\tau_1 = \inf\{t \geq 0, X_t \geq 1\}$. Montrer que τ_1 est un $\mathcal{F}_t$-temps d'arrêt et que p.s., $\tau_1 = \inf\{t \geq 0, X_t = 1\}$. Montrer, à l'aide du théorème d'arrêt, que τ_1 ne peut être borné (i.e. montrer que $\forall T > 0$, $\mathbb{P}(\tau_1 \leq T) < 1$).

7 On pose $\tilde{\tau}_1 = \inf\{t \geq 0, W_t \geq 1\}$. Montrer que τ_1 a la même loi que $(3\tilde{\tau}_1)^{1/3}$. Calculer la densité de la loi de τ_1. A-t-on $\mathbb{E}[\tau_1] < \infty$?

Problème : Convergence de l'algorithme du recuit simulé.

Le problème comporte trois parties. Les résultats de la partie I peuvent être admis pour la suite. Les parties II et III sont indépendantes.

Partie I : Comportement en temps long d'une équation différentielle ordinaire (copie "Analyse numérique").

Soient $A, B : \mathbb{R}_+ \to \mathbb{R}$ deux fonctions telles que $B \geq 0$ et $\exists t > 0, \forall s \geq t,\ A(s) > 0$. On suppose $\int_0^\infty A(t)\, dt = \infty$ et $\lim_{t\to\infty} \frac{B(t)}{A(t)} = 0$. Montrer que si $f : \mathbb{R}_+ \to \mathbb{R}_+$ est de classe $\mathcal{C}^1$ et satisfait $f'(t) \leq -A(t)f(t) + B(t)$, alors

$$\lim_{t\to\infty} f(t) = 0.$$

Indication : on pourra considérer $\frac{d}{dt}\left(\exp\left(\int_0^t A(s)\,ds\right) f(t)\right)$ *et montrer que pour* t_0 *assez grand, et* $\forall t \geq t_0$,

$$f(t) \leq \left(\int_0^{t_0} \exp\left(\int_0^s A(r)\,dr\right) B(s)\,ds\right) \exp\left(-\int_0^t A(s)\,ds\right) + \sup\left(\frac{B(t)}{A(t)},\, t \geq t_0\right) \exp\left(-\int_{t_0}^t A(s)\,ds\right).$$

Nous admettrons le résultat analogue dans le cas discret. Soient $(A_n)_{n\in\mathbb{N}}, (B_n)_{n\in\mathbb{N}} \in (\mathbb{R}_+^*)^{\mathbb{N}}$ deux suites telles que $\sum_{n\in\mathbb{N}} A_n = +\infty$, $(A_n)_{n\geq 0}$ est une suite bornée et $\lim_{n\to\infty} \frac{B_n}{A_n} = 0$. Si $(f_n)_{n\in\mathbb{N}}$ est une suite de réels positifs telle que $f_{n+1} \leq (1 - A_n) f_n + B_n$, alors on peut montrer que

$$\lim_{n\to\infty} f_n = 0.$$

Partie II : Analyse de la convergence en temps et espace continus (copie “Analyse numérique”).

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

$$dX_t = -\frac{1}{2}\frac{dH}{dx}(X_t)\,dt + \sqrt{\beta^{-1}(t)}\,dW_t \tag{2}$$

avec β une fonction du temps, scalaire, positive, croissante et qui tend vers $+\infty$ quand $t \to \infty$ et $(W_t, t \geq 0)$ un mouvement brownien de dimension 1. On note $\mathcal{M}$ l'ensemble des minima (globaux) de la fonction H :

$$\mathcal{M} = \{x \in \mathbb{T}, H(x) = \min_{\mathbb{T}} H\}.$$

L'objectif est de chercher une bonne loi de croissance pour $\beta(t)$ de manière à faire converger X_t vers $\mathcal{M}$.

II.1 On suppose dans cette question que $\beta^{-1}(t) = 0$, si bien que X_t est solution d'une équation différentielle ordinaire : $dX_t = -\frac{1}{2}\frac{dH}{dx}(X_t)\,dt$. Expliquer en le justifiant rapidement le comportement en temps long de X_t.

Dans la suite, on suppose $0 < \beta < \infty$. Pour β donné, on introduit la densité de probabilité $q_\beta = Z_\beta^{-1}\exp(-\beta H)$, où $Z_\beta = \int_{\mathbb{T}} \exp(-\beta H)$. Il faut comprendre $\int_{\mathbb{T}} \exp(-\beta H)$ commme l'intégrale de la fonction 1-périodique $\exp(-\beta H)$ sur une période :

$$\int_{\mathbb{T}} \exp(-\beta H) = \int_0^1 \exp(-\beta H(x))\,dx.$$

II.2 On admet que X_t admet une densité régulière $p(t,\cdot)$. On rappelle (cf. Equation (6.9) du cours de probabilités) que p satisfait l'équation de Fokker-Planck (ou Kolmogorov forward) :

$$\begin{cases} \dfrac{\partial p(t,x)}{\partial t} = \dfrac{1}{2}\dfrac{\partial}{\partial x}\left(\dfrac{dH}{dx} p(t,x)\right) + \dfrac{1}{2}\beta^{-1}(t)\dfrac{\partial^2 p}{\partial x^2}(t,x), \\ p(0,x) = p_0(x), \end{cases}$$

où p_0 est la densité de la condition initiale X_0. La fonction $x \to p(t,x)$ est 1-périodique, et on peut donc voir p comme une fonction définie pour $t \in \mathbb{R}_+$ et $x \in \mathbb{T}$. Vérifier que p satisfait l'équation aux dérivées partielles :

$$\frac{\partial p(t,x)}{\partial t} = \frac{\beta^{-1}(t)}{2}\frac{\partial}{\partial x}\left(q_{\beta(t)}(x)\frac{\partial}{\partial x}\left(\frac{p(t,x)}{q_{\beta(t)}(x)}\right)\right). \tag{3}$$

II.3 Pour $\beta > 0$ fixé, on introduit *le trou spectral* pour la mesure q_β défini par :

$$\lambda_1(\beta) = \frac{\beta^{-1}}{2}\inf\left\{\frac{\int_\mathbb{T}\phi'^2\,q_\beta}{\int_\mathbb{T}\phi^2\,q_\beta},\ \phi:\mathbb{T}\to\mathbb{R} \text{ tel que } \int_\mathbb{T}(\phi^2+\phi'^2)\,q_\beta < \infty \text{ et } \int_\mathbb{T}\phi\,q_\beta = 0\right\}.$$

II.3.a Démontrer que pour toute fonction $\phi : \mathbb{T} \to \mathbb{R}$, telle que $\int_\mathbb{T}(\phi^2+\phi'^2) < \infty$ et $\int_\mathbb{T}\phi\,q_\beta = 0$, $\int_\mathbb{T}\phi^2 \le \frac{1}{2}\int_\mathbb{T}(\phi')^2$. Comment s'appelle une inégalité de ce type ? *Indication : on pourra vérifier que* $\exists x_0$, $\phi(x_0) = 0$, *poser* $\psi(x) = \phi(x_0+x)$ *puis écrire* $\psi(x)^2 = \left(\int_0^x\psi'(y)\,dy\right)^2 \le x\int_0^x\psi'(y)^2\,dy$.

II.3.b En déduire que $\lambda_1(\beta) \ge \beta^{-1}\exp(-\beta\,\mathrm{osc}(H))$, où $\mathrm{osc}(H) = \max_\mathbb{T}(H) - \min_\mathbb{T}(H)$.

II.4 On suppose dans cette question que $\beta(t)$ est une fonction constante : $\beta(t) = \beta_0$. On introduit $f(t) = \frac{1}{2}\int_\mathbb{T}\left(\frac{p(t,\cdot)}{q_{\beta_0}} - 1\right)^2 q_{\beta_0}$. Montrer que $\frac{df}{dt} = -\frac{\beta_0^{-1}}{2}\int_\mathbb{T}\left(\frac{\partial}{\partial x}\left(\frac{p(t,\cdot)}{q_{\beta_0}}\right)\right)^2 q_{\beta_0}$. En déduire que $f(t) \le f(0)\exp(-2\lambda_1(\beta_0)t)$. En utilisant une inégalité de Cauchy Schwarz, montrer alors que $p(t,\cdot)$ converge vers q_{β_0} en norme $L^1(\mathbb{T})$ à vitesse exponentielle.

II.5 On suppose désormais que

$$\beta(t) = \frac{\ln(t+2)}{a},$$

où $a > \mathrm{osc}(H)$. On introduit la fonction $e(t) = \frac{1}{2}\int_\mathbb{T}\left(\frac{p(t,\cdot)}{q_{\beta(t)}} - 1\right)^2 q_{\beta(t)}$. Vérifier que $e(t) = \frac{1}{2}\left(\int_\mathbb{T}\frac{p(t,\cdot)^2}{q_{\beta(t)}} - 1\right)$. Montrer que

$$\frac{de}{dt} = \int_\mathbb{T}\frac{\partial p}{\partial t}\frac{p}{q_\beta} - \frac{1}{2}\int_\mathbb{T}\frac{\partial q_\beta}{\partial t}\left(\frac{p}{q_\beta}\right)^2. \tag{4}$$

En s'inspirant de la question II.4, vérifier que

$$\int_\mathbb{T}\frac{\partial p}{\partial t}\frac{p}{q_\beta} \le -2\lambda_1(\beta(t))e(t). \tag{5}$$

Montrer que

$$\frac{\partial q_\beta}{\partial t} = \frac{d\beta}{dt}\left(\int_\mathbb{T}Hq_\beta - H\right)q_\beta$$

et en déduire que

$$-\frac{1}{2}\int_\mathbb{T}\frac{\partial q_\beta}{\partial t}\left(\frac{p}{q_\beta}\right)^2 \le \frac{\mathrm{osc}(H)}{2a(t+2)}\,(2e+1). \tag{6}$$

En utilisant (4), (5) et (6), en déduire que

$$\frac{de}{dt} \leq -A(t)e(t) + B(t)$$

avec

$$A(t) = 2\lambda_1(\beta(t)) - \frac{\mathrm{osc}(H)}{a(t+2)}$$

et

$$B(t) = \frac{\mathrm{osc}(H)}{2a(t+2)}.$$

II.6 Montrer que $\int_0^\infty A(t)\,dt = \infty$ et que $\lim_{t\to\infty} \frac{B(t)}{A(t)} = 0$. En utilisant la partie I, obtenir que $\lim_{t\to\infty} e(t) = 0$.

II.7 Soit $\mathcal{M}_\delta$ un δ-voisinage de $\mathcal{M}$: $\mathcal{M}_\delta = \{x \in \mathbb{T}, \mathrm{dist}(x, \mathcal{M}) < \delta\}$. Vérifier que

$$\begin{aligned} e(t) &\geq \frac{1}{2} \int_{\mathcal{M}_\delta} \left(\frac{p(t,\cdot)}{q_{\beta(t)}} - 1 \right)^2 q_{\beta(t)}, \\ &\geq \frac{1}{2} \left(\int_{\mathcal{M}_\delta} |p(t,\cdot) - q_{\beta(t)}| \right)^2, \\ &\geq \frac{1}{2} \left| \mathbb{P}(X_t \in \mathcal{M}_\delta) - \int_{\mathcal{M}_\delta} q_{\beta(t)} \right|^2. \end{aligned}$$

Sans donner de démonstrations rigoureuses, justifier que $\lim_{t\to\infty} \int_{\mathcal{M}_\delta} q_{\beta(t)} = 1$. En déduire que $\lim_{t\to\infty} \mathrm{dist}(X_t, \mathcal{M}) = 0$ en probabilité.

Partie III : Analyse de la convergence en temps et espace discrets (copie "Probabilités").

Nous souhaitons minimiser une fonction $H : [0,1] \to \mathbb{R}_+$ que l'on suppose continue. On se donne $N \in \mathbb{N}^*$ et on considère l'espace d'états $E_N = \{k/N, k = 0, \ldots, N\}$.

III.1 Montrer que $\min_{x\in E_N} H(x) \underset{N\to+\infty}{\longrightarrow} \min_{x\in[0,1]} H(x)$.

Par la suite, l'entier N est fixé et on considère

$$\mathcal{M}^N = \{x \in E_N, H(x) = \min_{x\in E_N} H(x)\},$$

l'ensemble des minima globaux de H sur E_N.

III.2 L'algorithme d'Hastings-Metropolis.

On suppose que P est une matrice de transition sur E_N d'une chaîne de Markov irréductible, symétrique (i.e. $\forall x, y \in E_N, P(x,y) = P(y,x)$) et satisfaisant la condition de Doeblin :

il existe une probabilité π sur E_N , $l \in \mathbb{N}^*$, $c > 0$, t.q. $\forall x, y \in E_N$, $P^l(x,y) \geq c\,\pi(y)$.

III.2.a Montrer qu'il existe $y \in E_N$ t.q. $\forall x \in E_N, P^l(x,y) > 0$. En déduire que P est apériodique (cf. définition 4.6.1 du cours de probabilités).

III.2.b Pour $\beta > 0$, on définit pour $x \in E_N$, $\boxed{\mu_\beta(x) = e^{-\beta H(x)}/c_\beta}$ où $c_\beta = \sum_{x \in E_N} e^{-\beta H(x)}$, et

$$\forall x, y \in E_N, Q(x,y) = P(x,y)e^{-\beta(H(y)-H(x))^+}.$$

On rappelle que pour $x \in \mathbb{R}$, $x^+ = \max(x, 0)$ désigne la partie positive de x. A l'aide du problème résolu 4.6.4 du cours de probabilités, donner le comportement asymptotique de la chaîne de Markov $(X_n, n \in \mathbb{N})$ de matrice de transition Q. Que représente μ_β pour cette chaîne de Markov ? Rappeler le comportement de μ_β lorsque $\beta \to +\infty$.

III.2.c Donner un algorithme permettant de simuler la chaîne de Markov $(X_n,\ n \in \mathbb{N})$ de matrice de transition Q. Expliquer pourquoi, même si β est grand, la chaîne $(X_n, n \in \mathbb{N})$ peut rester longtemps piégée autour d'un minimum local de H.

III.2.d Donner un exemple concret de matrice de transition P sur E_N qui est symétrique, irréductible et satisfait la condition de Doeblin avec $l = 1$.

III.3 Convergence de l'algorithme du recuit simulé.

On suppose dans toute la suite que P est une matrice de transition sur E_N irréductible, symétrique et **satisfaisant la condition de Doeblin avec $\boxed{\mathbf{l{=}1}}$:**

$\forall x, y \in E_N,\ P(x,y) \geq c\ \pi(y)$ pour une certaine probabilité π sur E_N et un certain $c > 0$.

On souhaite construire un algorithme dans l'esprit de l'algorithme d'Hastings-Metropolis, et qui converge effectivement vers un minimum global de H sur E_N en faisant évoluer le paramètre β vers des grandes valeurs au cours des itérations. Pour cela, nous allons considérer une suite $(\beta_n)_{n\in\mathbb{N}} \in (\mathbb{R}_+^*)^{\mathbb{N}}$, et une chaîne de Markov $(X_n, n \in \mathbb{N})$ **inhomogène en temps** telle que

$$\forall x, y \in E_N, \mathbb{P}(X_{n+1} = y | X_n = x) = Q_n(x,y) := P(x,y)e^{-\beta_n(H(y)-H(x))^+},$$

et de loi initiale ν_0. La loi de $(X_0, \ldots, X_n)$ est ainsi donnée par

$$\forall x_0, \ldots, x_n \in E_N, \mathbb{P}(X_0 = x_0, \ldots, X_n = x_n) = \nu_0(x_0)Q_0(x_0,x_1)Q_1(x_1,x_2)\ldots Q_{n-1}(x_{n-1},x_n).$$

On note ν_n la loi de X_n : $\forall x \in E_N, \nu_n(x) = \mathbb{P}(X_n = x)$. Pour une fonction $f : E_N \to \mathbb{R}$, $\mathbf{V}(f) = \sum_{x\in E^N} |f(x)|$ désigne la norme en variation de f. Enfin, on définit $\mathrm{osc}_N(H) = \max_{x\in E_N} H(x) - \min_{x\in E_N} H(x)$.

III.3.a Montrer que pour $n \in \mathbb{N}$, $\nu_{n+1} = \nu_n Q_n$ et $\mu_{\beta_n} Q_n = \mu_{\beta_n}$.

III.3.b Montrer que $Q_n(x,y) \geq ce^{-\beta_n \mathrm{osc}_N(H)}\pi(y)$. En déduire que

$$\mathbf{V}(\nu_{n+1} - \mu_{\beta_n}) \leq (1 - ce^{-\beta_n \mathrm{osc}_N(H)})\mathbf{V}(\nu_n - \mu_{\beta_n}).$$

Indication : on pourra s'inspirer de la question 2.a du Problème 4.6.5 du cours de probabilités.

III.3.c On pose pour $x \in E_N$, $\tilde{H}(x) = H(x) - \min_{x\in E_N} H(x)$. Montrer que

$$\mu_\beta(x) = \frac{e^{-\beta\tilde{H}(x)}}{\sum_{y\in E_N} e^{-\beta\tilde{H}(y)}}.$$

Montrer que pour $\beta' > \beta > 0$, on a

$$0 \leq e^{-\beta\tilde{H}(x)} - e^{-\beta'\tilde{H}(x)} \leq \mathrm{osc}_N(H)|\beta - \beta'|e^{-\beta\tilde{H}(x)}.$$

En déduire que

$$\mathbf{V}(\mu_\beta - \mu_{\beta'}) \leq 2\mathrm{osc}_N(H)|\beta - \beta'|.$$

Cette inégalité est-elle aussi vraie pour $\beta' \leq \beta$?

III.3.d On prend $\beta_n = \frac{\log(n+1)}{a}$. Montrer à l'aide des questions précédentes que

$$\mathbf{V}(\nu_{n+1} - \mu_{\beta_{n+1}}) \leq \left(1 - c(n+1)^{-\frac{\mathrm{osc}_N(H)}{a}}\right)\mathbf{V}(\nu_n - \mu_{\beta_n}) + 2\frac{\mathrm{osc}_N(H)}{a}\log\left(1 + \frac{1}{n+1}\right).$$

En déduire à l'aide de la partie I que $\mathbf{V}(\nu_n - \mu_{\beta_n}) \underset{n\to+\infty}{\to} 0$ si on prend $a > \mathrm{osc}_N(H)$.

III.3.e On suppose $a > \mathrm{osc}_N(H)$. Déduire des résultats précédents que

$$\mathbb{P}(X_n \in \mathcal{M}^N) \underset{n\to+\infty}{\to} 1.$$

Comparer ces résultats avec ceux obtenus dans la partie II.