Analyse Numérique - Correction de l'examen du Lundi 12 Janvier 2009

Aucun document autorisé - Durée : 3h

Exercice 1

Partie 1 : Résolution explicite dans un cas particulier

(a) On obtient les fonctions

$$u(x) = \frac{f}{c}x + K_1 + K_2 e^{cx},$$

où $K_1, K_2 \in \mathbb{R}$ sont des constantes d'intégration.

(b) On vérifie qu'il existe un unique jeu de valeurs pour K_1 et K_2 qui convient et on trouve

$$u(x) = \frac{1}{1-e^c}\left(\frac{f}{c}x(1-e^c) - \frac{f}{c}(1-e^{cx})\right).$$

(c) On calcule la dérivée seconde de u, on trouve

$$u''(x) = \frac{c}{1-e^c}e^{cx}f.$$

Or c et $1-e^c$ sont de signes opposés, ce qui montre que pour $f \geq 0$, u est bien concave. Comme par ailleurs $u(0) = u(1) = 0$, on obtient bien que u est positive (car la courbe est au-dessus de sa corde).

Partie 2 : Principe du maximum dans le cas général

(a) Comme u est continue sur le compact $[0,1]$ on sait qu'elle atteint son infimum, mais comme $u(0) = u(1) = 0$ et que $\inf u < 0$, cet infimum ne peut être atteint sur le bord, d'où l'existence d'un tel x_0.

(b) On peut par exemple utiliser l'inégalité de Young

$$\|c\|_\infty |y| = 2\frac{\|c\|_\infty |y|}{2} \leq \frac{1}{2}2^2 + \frac{1}{2}\frac{\|c\|_\infty^2 y^2}{4} = 2 + \frac{\|c\|_\infty^2}{8}y^2.$$

On obtient le résultat avec $R > \frac{\|c\|_\infty^2}{8}$. (N.B. : l'inégalité est bien stricte y compris pour $y = 0$! !).

(c) Comme $e^0 = 1$, le terme qui contient ε dans la définition de v, s'annulle en $x = 0$ et $x = 1$, on a donc bien $v(0) = v(1) = 0$. De plus, on a

$$v(x_0) = u(x_0) + \varepsilon(1 - e^{-Rx_0(1-x_0)}) \leq u(x_0) + \varepsilon,$$

et comme $u(x_0) < 0$ et $\varepsilon < |u(x_0)|$, on a bien que $v(x_0) < 0$.

(d) L'existence de $\tilde{x}$ s'obtient de la même manière qu'à la première question : v est continue sur $[0,1]$, elle y atteint donc son infimum qui est forcément négatif car $v(x_0) < 0$ donc elle ne peut pas atteindre son infimum sur le bord.

Comme $\tilde{x}$ est un point de minimum de v qui est une fonction de classe $\mathcal{C}^2$, on sait que sa dérivée première s'annulle en ce point et sa dérivée seconde est positive ou nulle en ce point.

(e) On calcule

$$\partial_x u = \partial_x v - \varepsilon R(1-2x)e^{-Rx(1-x)},$$
$$\partial_x^2 u = \partial_x^2 v + \varepsilon R(2+R(1-2x)^2)e^{-Rx(1-x)}.$$

Les résultats demandés s'obtiennent en évaluant ces formules en $\tilde{x}$ et en utilisant la question précédente.

(f) Comme u est solution du problème $-\partial_x^2 u + c(x)\partial_x u = f(x)$, on peut évaluer l'équation au point $\tilde{x}$ et on obtient

$$f(\tilde{x}) = -\partial_x^2 u(\tilde{x}) + c(\tilde{x})\partial_x u(\tilde{x}) \leq -\varepsilon R(2 + R(1-2\tilde{x})^2 + c(\tilde{x})(1-2\tilde{x}))e^{-R\tilde{x}(1-\tilde{x})}.$$

Par ailleurs $c(\tilde{x})(1-2\tilde{x}) \geq -|c(\tilde{x})||1-2\tilde{x}| \geq -\|c\|_\infty|1-2\tilde{x}|$, d'où le résultat demandé.

D'après le choix de R dans la question b (en prenant $y = 1-2\tilde{x}$), on obtient que le second membre de l'inégalité précédente est strictement négatif et donc que $f(\tilde{x}) < 0$. Ceci est en contradiction avec l'hypothèse $f \geq 0$ et prouve donc le principe du maximum.

Partie 3 : Schéma aux différences finies

(a) Le second membre F est donné par $F = (f(x_i))_{1\leq i\leq N}$ et la matrice A par

$$A = \begin{pmatrix} \frac{2}{\Delta x^2} + \frac{c(x_1)}{\Delta x} & -\frac{1}{\Delta x^2} & 0 & \cdots & 0 \\ -\frac{1}{\Delta x^2} - \frac{c(x_2)}{\Delta x} & \frac{2}{\Delta x^2} + \frac{c(x_2)}{\Delta x} & -\frac{1}{\Delta x^2} & 0 & \vdots \\ 0 & -\frac{1}{\Delta x^2} - \frac{c(x_3)}{\Delta x} & \frac{2}{\Delta x^2} + \frac{c(x_3)}{\Delta x} & -\frac{1}{\Delta x^2} & \vdots \\ \vdots & 0 & \ddots & \ddots & 0 \\ 0 & \cdots & \cdots & -\frac{1}{\Delta x^2} - \frac{c(x_N)}{\Delta x} & \frac{2}{\Delta x^2} + \frac{c(x_N)}{\Delta x} \end{pmatrix}.$$

(b) A n'est pas symétrique (sauf si $c = 0$) mais elle est tridiagonale.

(c) Le principe du maximum pour A dit que si v et b sont deux vecteurs de $\mathbb{R}^N$ tels que $Av = b$ et $b \geq 0$, alors $v \geq 0$. La démonstration est **exactement** celle vue en cours. On utilise ici la positivité de c. En effet, on raisonne par l'absurde en supposant qu'il existe un plus petit indice i_0 tel que

$$v_{i_0} = \min v_i < 0.$$

On a alors

$$b_{i_0} = \underbrace{-\frac{v_{i_0+1} - v_{i_0}}{\Delta x^2}}_{\geq 0} + \underbrace{\frac{v_{i_0} - v_{i_0-1}}{\Delta x^2}}_{<0} + \underbrace{c(x_{i_0})}_{\geq 0}\underbrace{\frac{v_{i_0} - v_{i_0-1}}{\Delta x}}_{<0},$$

ce qui prouve que $b_{i_0} < 0$. C'est une contradiction.

On en déduit que A est inversible car tout élément de son noyau doit être à la fois positif et négatif, il est donc nul.

(d) Soit u une solution de classe $\mathcal{C}^3$ su problème. L'erreur de consistance est définie classiquement par

$$R_i = -\frac{u(x_{i+1}) - 2u(x_i) + u(x_{i-1})}{\Delta x^2} + c(x_i)\frac{u(x_i) - u(x_{i-1})}{\Delta x} - f(x_i).$$

On peut aussi prendre la définition opposée, cela ne change essentiellement rien. On effectue ensuite des développements de Taylor standard pour obtenir l'estimation donnée dans l'énoncé.

(e) i. On sait que $\tilde{u}$ est de classe $\mathcal{C}^2$ et que $\partial_x^2 \tilde{u} = c\partial_x \tilde{u} - 1$ et donc, comme c est de classe $\mathcal{C}^1$, on voit bien que $\partial_x^2 \tilde{u}$ est aussi de classe $\mathcal{C}^1$ et donc que $\tilde{u}$ est de classe $\mathcal{C}^3$.

ii. A. Par définition de l'erreur de consistance donnée ci-dessus on a

$$\tilde{R} = A\tilde{U} - D,$$

d'où $A\tilde{U} = \tilde{R} + D$.

B. D'après la question d), on a

$$\|\tilde{R}\|_\infty \leq M\Delta x(\|\tilde{u}''\|_\infty + \|\tilde{u}'''\|_\infty) \leq \frac{1}{2},$$

dès que $\Delta x \leq \delta$, par définition de δ.
Pour tout i on a donc

$$(A\tilde{U})_i = \tilde{R}_i + 1 \geq 1 - \|\tilde{R}\|_\infty \geq 1/2,$$

soit encore

$$A\tilde{U} \geq \frac{1}{2}D.$$

(f) On a vu que A vérifie le principe du maximum discret, donc $A^{-1} \geq 0$. Si $\Delta x \leq \delta$, d'après la question précédente on a

$$0 \leq A^{-1}D \leq 2\tilde{U} \leq 2\|\tilde{u}\|_\infty D.$$

Ainsi, pour tout $U \in \mathbb{R}^N$, on a

$$-\|U\|_\infty D \leq U \leq \|U\|_\infty D,$$

$$-\|U\|_\infty A^{-1}D \leq A^{-1}U \leq \|U\|_\infty A^{-1}D,$$

$$\|A^{-1}U\|_\infty \leq \|U\|_\infty \|A^{-1}D\|_\infty \leq 2\|U\|\|\tilde{u}\|_\infty,$$

d'où le résultat.
On vient donc de démontrer la stabilité L^∞ du schéma sous la condition $\Delta x \leq \delta$.

(g) Par définition de l'erreur de consistance d'une part et du schéma d'autre part, on a

$$AU = F, \ \ A\bar{U} = R + F.$$

Donc, il vient

$$AE = R,$$

et ainsi, si $\Delta x \leq \delta$

$$\|E\|_\infty \leq \|A^{-1}\|_\infty \|R\|_\infty \leq 2\|\tilde{u}\|_\infty M\Delta x(\|u''\|_\infty + \|u'''\|_\infty).$$

(h) Il suffit de décentrer différemment le terme d'ordre 1.

Exercice 2

1. Les courbes caractéristiques associées à ce problème sont les solutions du problème suivant

$$\begin{cases} \dfrac{d}{dt}X(t,t_0,x_0) = c(X(t,t_0,x_0)), \\ X(t_0,t_0,x_0) = x_0. \end{cases}$$

   Elles sont bien définies globalement car la fonction c est bornée et donc il ne peut y avoir d'explosion en temps fini.

2. On pose $\varphi(t) = u(t, X(t,0,x_0))$. On a

$$\varphi'(t) = \partial_t u(t, X(t,0,x_0)) + c(X(t,0,x_0))\partial_x u(t, X(t,0,x_0)).$$

   De sorte que u est solution du problème étudié si et seulement si on a

$$\varphi'(t) = g(\varphi(t))$$

   et $\varphi(0) = u_0(x_0)$.

3. Quand $y_0 = 0$, l'unique solution du problème de Cauchy précédent est la fonction constante $y(t) = 0$, car $g(0) = 0$. De même si $y_0 = 1$, la solution est $y(t) = 1$.

4. Le problème de Cauchy est localement bien posé d'après le théorème de Cauchy-Lipschitz. En réalité, il ne peut y avoir d'explosion en temps fini car la solution pour $y_0 \in]0,1[$ ne peut atteindre les valeurs 0 et 1. En effet, par unicité de la solution du problème de Cauchy, s'il existait $t > 0$ tel que $y(t) \in \{0,1\}$, alors la solution y serait constante ce qui n'est pas possible car $y(0) = y_0 \in]0,1[$.

5. Il suffit de mettre bout à bout les ingrédients précédents pour conclure.

Exercice 3

Partie 1 : Un schéma éléments finis pour un problème elliptique

(a) Il suffit d'appliquer le théorème de Lax-Milgram (dont on vérifie évidemment les hypothèses).

(b) Si u est de classe $\mathcal{C}^1$, la dérivée faible ∇u coïncide avec la dérivée usuelle u'. Si de plus u est de classe $\mathcal{C}^2$, on intègre alors par parties et on obtient

$$\int_0^1 (-\tau u'' + u - f)v \, dx = 0,$$

pour toute fonction $v \in H_0^1(I)$. Comme $H_0^1(\Omega)$ est dense dans L^2 par exemple, on en déduit que la fonction $-\tau u'' + u - f$ est identiquement nulle et donc que u est solution classique du problème.

(c) Il suffit bien entendu de vérifier que $(\varphi_i)_{1 \leq i \leq N}$ est une famille libre. Pour cela, on considère une combinaison linéaire nulle des φ_i

$$\sum_{i=1}^{N} \alpha_i \varphi_i = 0,$$

et l'on évalue en un point x_j avec $1 \leq j \leq N$. Par définition des φ_i, on trouve

$$\alpha_j = 0,$$

et le résultat est démontré.

(d) On applique à nouveau le théorème de Lax-Milgram à l'espace V_N qui est bien un Hilbert (car sous-espace de dimension finie, donc fermé, de $H_0^1(I)$).

(e) On note $u_N = \sum_{j=1}^N u_j \varphi_j$. On prend alors les fonctions tests v_N respectivement égales à tous les φ_i, ce qui donne les équations suivantes

$$\sum_{j=1}^{N} u_j \left(\tau \int_I \nabla\varphi_i \nabla\varphi_j \, dx + \int_I \varphi_i \varphi_j \, dx \right) = \int_I f \varphi_i \, dx.$$

On trouve donc $A = (\int_I \nabla\varphi_i \nabla\varphi_j \, dx)_{i,j}$ et $M = (\int_I \varphi_i \varphi_j \, dx)_{i,j}$ et $F = (\int_I f\varphi_i \, dx)$.

La formule $A = \Delta x \tilde{A}$ se vérifie immédiatement par le calcul des coefficients de A.

(f) On a vu en cours que $\tilde{A}$ est symétrique définie positive et vérifie le principe du maximum ce qui implique les mêmes propriétés sur A. De plus, on a montré la propriété de stabilité $\|\tilde{A}^{-1}\|_\infty \leq \frac{1}{8}$ ce qui implique le résultat attendu sur A.

(g) Tous les coefficients de M sont des intégrales de fonctions positives, donc $M \geq 0$.
On a donc $v = \sum_{i=1}^N v_i \varphi_i$ de sorte que

$$(MV, V) = \sum_{i=1}^N \sum_{j=1}^N v_i v_j \left(\int_I \varphi_i \varphi_j \, dx \right) = \int_I v^2 \, dx.$$

Donv (MV, V) est égale au carré de la norme L^2 de v. Ceci implique, en particulier, la définie positivité de M car si $(MV, V) = 0$, alors la fonction v est nulle et donc ces coefficients dans la base des φ_i sont tous nuls, i.e. $V = 0$.

(h) Supposons que $W \neq 0$. Il existe donc i tel que $w_i > 0$. Il existe alors $j \in \{i-1, i, i+1\} \cap \{1, \dots, N\}$ tel que $(MW)_j = 0$ puisqu'au plus un coefficient de MW est non nul. On voit qu'on obtient une contradiction puisque

$$0 < m_{j,i} w_i \leq (MW)_j = 0.$$

Soit C_i la i-ième colonne de M^{-1}. On a, par définition, $MC_i = e_i$ où e_i est le i-ième vecteur de la base canonique. Comme e_i a tous ses coefficients sauf 1 nuls, C_i ne peut pas être positif car sinon il serait nul d'après ce qui précède. Ainsi C_i contient au moins un coefficient négatif. En particulier la matrice M ne vérifie pas le principe du maximum discret.
De même, quand $\tau \to 0$, $(\tau A + M)^{-1}$ converge vers M^{-1} (car l'application *inverse* est continue sur l'ensemble des matrices inversibles) et donc, pour tout τ assez petit, $(\tau A + M)^{-1}$ a des coefficients négatifs dans chacune de ces colonnes et en particulier ne vérifie pas le principe du maximum discret.

(i) On écrit

$$\tau A + M = (\tau - \Delta x^2/6) A + \Delta x \mathrm{Id},$$

d'où $\alpha = \tau - \Delta x^2/6$ et $\beta = \Delta x$. Ainsi, si $\tau \geq \Delta x^2/6$, on a $\alpha \geq 0$ et $\beta \geq 0$. On a alors vu en cours que la matrice $\alpha A + \beta \mathrm{Id}$ vérifie le principe du maximum discret.

Partie 2 : Résolution d'un problème parabolique du type "équation de la chaleur"

(a) On cherche la solution u sous la forme $u(t,x) = \varphi(t)u_0(x)$. Une telle fonction est solution du problème si et seulement si $\varphi(0) = 1$ et $\varphi'(t) + \pi^2\varphi(t) - \alpha\varphi(t) = 0$. On obtient donc

$$\varphi(t) = e^{(\alpha-\pi^2)t},$$

et donc

$$u(t,x) = e^{(\alpha-\pi^2)t}\sin(\pi x).$$

On a donc trois comportements possibles.
- Si $\alpha < \pi^2$, on a $u(t,x) \to 0$ quand $t \to \infty$.
- Si $\alpha = \pi^2$, on a $u(t,x) = u_0(x)$ pour tout $t > 0$.
- Si $\alpha > \pi^2$, on a $u(t,x) \to \infty$ quand $t \to \infty$.

(b) Le schéma s'écrit $BU^{n+1} = CU^n$, avec

$$B = \Delta t A + M, \text{ et } C = (1+\alpha\Delta t)M.$$

On a vu dans la partie précédente que B est inversible et donc cette suite est bien définie.

(c) On prend la formule qui définit le schéma au rang n et on prend le produit scalaire de cette formule avec U^{n+1}. On écrit alors

$$(M(U-V),U) = \frac{1}{2}(MU,U) - \frac{1}{2}(MV,V) + \frac{1}{2}(M(U-V),(U-V)),$$

et on utilise le fait que $(AU^{n+1},U^{n+1}) \geq 0$.

(d) Comme M est symétrique définie positive, l'application $(V,W) \mapsto (MV,W)$ est un produit scalaire. La première inégalité demandée est donc l'inégalité de Cauchy-Schwarz. Pour la deuxième inégalité, on applique juste l'inégalité de Young

$$ab \leq \frac{\beta}{2}a^2 + \frac{1}{2\beta}b^2.$$

(e) On utilise l'inégalité précédente avec $V = U^n$ et $W = U^{n+1} - U^n$ et $\beta = \alpha\Delta t$, on trouve

$$\alpha\Delta t(MU^n, U^{n+1} - U^n) \leq \frac{\alpha^2\Delta t^2}{2}(MU^n,U^n) + \frac{1}{2}(M(U^{n+1}-U^n), U^{n+1} - U^n).$$

Ainsi, en remettant cela dans l'inégalité de la question c) on trouve

$$\frac{1}{2}(MU^{n+1},U^{n+1}) \leq (\frac{1}{2} + \alpha\Delta t + \frac{1}{2}\alpha^2\Delta t^2)(MU^n,U^n),$$

ce qui fournit le résultat annoncé.

(f) L'inégalité précédente s'écrit

$$(MU^{n+1},U^{n+1}) \leq (1+\alpha\Delta t)^2(MU^n,U^n),$$

ce qui donne par récurrence

$$(MU^n,U^n) \leq (1+\alpha\Delta t)^{2n}(MU^0,U^0).$$

Or, si $n \leq T/\Delta t$ on a

$$(1+\alpha\Delta t)^{2n} = e^{2n\log(1+\alpha\Delta t)} \leq e^{2n\alpha\Delta t} \leq e^{2\alpha T},$$

ce qui donne l'inégalité souhaitée.

(g) D'après la question 1-g), la quantité (MU^n,U^n) représente la norme L^2 au carré de la solution approchée au temps $n\Delta t$, et on a donc établit la stabilité L^2 du schéma.