Université Paris-Nord
Institut Galilée
Département de Mathématiques

Année 2014-2015
MACS - 1

C. Basdevant, B. Delourme

Corrigé de l'examen d'Analyse Numérique du 12 novembre 2014

Durée : 3h

Une feuille recto-verso de notes personnelles manuscrites est autorisée.

Exercice (8/20)

Soit $f(t, y) : [a, b] \times \mathbb{R}^n \to \mathbb{R}^n$ une fonction suffisamment continûment différentiable et Lipschitzienne par rapport à y :

$$\|f(t, y) - f(t, z)\| \leq L \|y - z\| \quad \forall t \in [a, b], \quad \forall y, z \in \mathbb{R}^n$$

Etant donnés les réels (a_1, a_2, b_1, b_2) et $c \in]0, 1]$, pour résoudre numériquement le problème :

$$y'(t) = f(t, y(t)) \quad \forall t \in [a, b], \quad y(a) = \eta$$

on utilise la méthode à un pas :

$$y_{k+1} = y_k + h\,\Phi(t_k, y_k, h) \tag{1}$$

avec $h = \frac{b-a}{N}$, $t_k = a + kh$, $k = 0 \dots N$, $y_0 = \eta$, et

$$\Phi(t, y, h) = b_1 f(t, y) + b_2 f(t + ch, y^*)$$
$$y^* = y + h\,[a_1 f(t, y) + a_2 f(t + ch, y^*)]$$

1. Pour quelles valeurs de (a_1, a_2, b_1, b_2) la méthode (1) est-elle implicite ?
   Corrigé : La méthode est implicite si $b_2 \neq 0$ et $a_2 \neq 0$.
2. Pour quelles valeurs de (a_1, a_2, b_1, b_2, c) retrouve-t-on la méthode d'Euler implicite ?
   Corrigé : La méthode d'Euler implicite est définie par $y_{k+1} = y_k + h\,f(t_{k+1}, y_{k+1})$, soit $y_{k+1} = y_k + h\,f(t_{k+1}, z)$ avec z l'unique solution de $z = y_k + h\,f(t_{k+1}, z)$, elle est obtenue pour $b_1 = 0$, $b_2 = 1$, $a_1 = 0$, $a_2 = 1$, $c = 1$.
3. Montrez que le schéma (1) est stable.
   Corrigé : Pour la stabilité du schéma il suffit de vérifier que Φ est Lipschitzienne par rapport à y, mais :

   $$\|\Phi(t, y, h) - \Phi(t, z, h)\| \leq |b_1| L\, \|y - z\| + |b_2| L\, \|y^* - z^*\|$$
   $$\|y^* - z^*\| \leq \|y - z\| + h|a_1|L\|y - z\| + h|a_2|L\|y^* - z^*\|$$

   De cette dernière inégalité on tire :

   $$\|y^* - z^*\| \leq \frac{1 + h|a_1|L}{1 - h|a_2|L}\|y - z\|$$

   d'où

   $$\|\Phi(t, y, h) - \Phi(t, z, h)\| \leq (|b_1|L + |b_2|L\frac{1 + h|a_1|L}{1 - h|a_2|L})\,\|y - z\|$$

4. Pour quelles valeurs de (a_1, a_2, b_1, b_2, c) le schéma (1) est-il convergent ?
   Corrigé : Le schéma étant stable, il sera convergent s'il est aussi consistant. Pour la consistance du schéma il faut vérifier que $\Phi(t, y, 0) = f(t, y)$, $\forall t, \forall y$, soit $b_1 f(t, y) + b_2 f(t, y) = f(t, y)$, il faut donc $b_1 + b_2 = 1$.
5. On suppose dans la suite $a_2 = 0$. Réécrivez le schéma ainsi obtenu.
   Corrigé : $\Phi(t, y, h) = b_1 f(t, y) + b_2 f(t + ch, y + ha_1 f(t, y))$
6. Pour quelles valeurs de (a_1, b_1, b_2, c) le schéma (1) est-il d'ordre 2 ? Ecrivez alors le schéma en fonction de c.
   Corrigé : Pour que le schéma soit d'ordre 2 il faut de plus que

$$\frac{\partial \Phi}{\partial h}(t, y, 0) = \frac{1}{2}\left(\frac{\partial f}{\partial t}(t, y) + \frac{\partial f}{\partial y}(t, y) f(t, y)\right), \forall t, \forall y$$

   soit

$$b_2 c \frac{\partial f}{\partial t}(t, y) + b_2 a_1 \frac{\partial f}{\partial y}(t, y) f(t, y) = \frac{1}{2}\left(\frac{\partial f}{\partial t}(t, y) + \frac{\partial f}{\partial y}(t, y) f(t, y)\right)$$

   il faut donc $b_1 + b_2 = 1$, $b_2 c = \frac{1}{2}$ et $b_2 a_1 = \frac{1}{2}$, soit $a_1 = c$, $b_2 = \frac{1}{2c}$ et $b_1 = 1 - \frac{1}{2c}$. Le schéma s'écrit alors :

$$\Phi(t, y, h) = (1 - \frac{1}{2c}) f(t, y) + \frac{1}{2c} f(t + ch, y + ch f(t, y))$$

7. Ecrivez ce dernier schéma pour l'équation de Lotka-Volterra ci-dessous. Vous vous attacherez à définir une succession d'étapes permettant de passer de y_k à y_{k+1} de la façon la plus simple et la moins coûteuse.

$$\begin{aligned} p'(t) &= \alpha\, p(t) + \beta\, p(t) q(t) \\ q'(t) &= \gamma\, q(t) + \delta\, p(t) q(t) \end{aligned}$$

   Corrigé : Note : remarquez qu'ici on est en dimension 2 ($y_k = (p_k, q_k)$) et que f ne dépend pas explicitement du temps.

$$\begin{aligned} f_1 &= \alpha\, p_k + \beta\, p_k q_k, \quad p^* = p_k + ch f_1, \quad g_1 = \alpha\, p^* + \beta\, p^* q^* \\ f_2 &= \gamma\, q_k + \delta\, p_k q_k, \quad q^* = q_k + ch f_2, \quad g_2 = \gamma\, q^* + \delta\, p^* q^* \end{aligned}$$

$$p_{k+1} = p_k + h(1 - \frac{1}{2c}) f_1 + h \frac{1}{2c} g_1, \quad q_{k+1} = p_k + h(1 - \frac{1}{2c}) f_2 + h \frac{1}{2c} g_2$$

Problème (12/20)

On considère le système linéaire de $n > 2$ équations à n inconnues $(x_i\ ;\ i = 1 \ldots n)$:

$$(\alpha + 2\beta) x_i - \beta(x_{i+1} + x_{i-1}) = f_i \quad 1 \leq i \leq n \tag{2}$$

avec $\alpha \geq 0$ et $\beta > 0$ et pour lequel $\{f_i\ ;\ i = 1 \ldots n\}$, x_0 et x_{n+1} sont des données.

1. Mettre le système sous la forme $Ax = b$ avec $x \in \mathbb{R}^n$, $x = \{x_i \; ; \; i = 1 \dots n\}$, où A, une matrice $n \times n$ symétrique, et $b \in \mathbb{R}^n$ sont à déterminer.

   Corrigé : On trouve

$$A = \begin{pmatrix} \alpha+2\beta & -\beta & 0 & \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\ -\beta & \alpha+2\beta & -\beta & 0 & \dots\dots\dots\dots\dots\dots\dots\dots \\ 0 & -\beta & \alpha+2\beta & -\beta & 0 & \dots\dots\dots\dots\dots\dots \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\ \dots\dots\dots\dots & 0 & -\beta & \alpha+2\beta & -\beta & 0 & \dots\dots\dots\dots \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots & 0 & -\beta & \alpha+2\beta \end{pmatrix}$$

   et $b_1 = f_1 + \beta x_0$, $\{b_i = f_i \; ; \; i = 2 \dots n-1\}$, $b_n = f_n + \beta x_{n+1}$

2. En calculant $(Ax, x) = \sum_{i=1}^n (Ax)_i \, x_i$ Montrez que

$$(Ax, x) = \alpha \sum_{i=1}^{n} x_i^2 + \beta \left(x_1^2 + \sum_{i=1}^{n-1} (x_i - x_{i+1})^2 + x_n^2 \right)$$

   (indication on pourra poser pour faciliter les calculs $x_0 = x_{n+1} = 0$)

   Corrigé :

$$\begin{aligned} (Ax, x) =& \alpha \sum_{i=1}^{n} x_i^2 + \beta \sum_{i=1}^{n} [(x_i - x_{i+1})x_i + (x_i - x_{i-1})x_i] \\ =& \alpha \sum_{i=1}^{n} x_i^2 + \beta \left[\sum_{i=1}^{n} (x_i - x_{i+1})x_i + \sum_{j=0}^{n-1} (x_{j+1} - x_j)x_{j+1} \right] \\ =& \alpha \sum_{i=1}^{n} x_i^2 + \beta \left[x_1^2 + \sum_{i=1}^{n-1} (x_i - x_{i+1})^2 + x_n^2 \right] \end{aligned}$$

3. En déduire que la matrice A du système (2) est symétrique, définie positive.

   Corrigé : On a déjà vu que la matrice est symétrique, pour qu'elle soit définie positive il faut vérifier que $(Ax, x) \geq 0$ pour tout x et que $(Ax, x) = 0$ implique $x = 0$. Ceci se déduit immédiatement de la formule précédente puisque $\alpha \geq 0$ et $\beta > 0$, en effet si $(Ax, x) = 0$, on a $x_1 = 0$ et comme $x_i - x_{i+1} = 0$, de proche en proche tous les x_i sont nuls.

4. Indiquez une méthode directe de résolution du système (2) (on citera un théorème assurant que la méthode est efficace) et évaluez l'ordre de grandeur du nombre d'opérations nécessaires.

   Corrigé : On a vu dans le cours que les matrices symétriques définies positives admettent une décomposition de Choleski BB^T. De plus la matrice A étant tridiagonale, la matrice B de Choleski ne comporte qu'une diagonale et une sous-diagonale. Le système (2) peut alors être résolu par une descente et une remontée en $O(n)$ opérations.

5. On considère, pour $k = 1 \dots n$, les vecteurs x^k de $\mathbb{R}^n$ définis par :

$$x^k = \left\{ x_i^k = \sin(\frac{i\,k\pi}{n+1}) \; ; \; i = 1 \dots n \right\}$$

Montrez que le vecteur x^k est vecteur propre de la matrice A pour une valeur propre λ^k que l'on déterminera.

Corrigé : Calculons les composantes de Ax^k (on utilise le fait que $x_0^k = 0$ et $x_{n+1}^k = 0$) :

$$\begin{aligned}(Ax^k)_i =&(\alpha + 2\beta)\sin(\frac{i\,k\pi}{n+1}) - \beta\left(\sin(\frac{(i+1)\,k\pi}{n+1}) + \sin(\frac{(i-1)\,k\pi}{n+1})\right)\\ =&(\alpha + 2\beta)\sin(\frac{i\,k\pi}{n+1}) - \beta\left(2\sin(\frac{i\,k\pi}{n+1})\cos(\frac{k\pi}{n+1})\right)\\ =&\left[\alpha + 2\beta(1-\cos(\frac{k\pi}{n+1}))\right]x_i^k = \left[\alpha + 4\beta\sin^2(\frac{k\pi}{2(n+1)})\right]x_i^k\end{aligned}$$

Ce qui prouve que x^k est vecteur propre de A pour la valeur propre $\lambda^k = \alpha + 4\beta\sin^2(\frac{k\pi}{2(n+1)})$

6. En déduire toutes les valeurs propres de A et que les vecteurs x^k sont orthogonaux deux à deux pour le produit scalaire euclidien. Retrouvez que A est définie positive.
Corrigé : Quand k varie de 1 à n les valeurs propres λ^k varient de $\alpha+4\beta\sin^2(\frac{\pi}{2(n+1)})$ à $\alpha + 4\beta\sin^2(\frac{n\pi}{2(n+1)})$ et sont distinctes deux à deux, on a donc trouvé toutes les valeurs propres de la matrice. Les sous espaces propres d'une matrice symétrique étant orthogonaux deux à deux on en déduit que les x^k sont orthogonaux deux à deux. Les valeurs propres étant strictement positives, cela confirme que A est définie positive.

7. Calculez le rayon spectral $\rho(J)$ de la matrice d'itération J de la méthode de Jacobi appliquée à la résolution du système (2) (indication la matrice J a la même forme que la matrice A avec des coefficients α' et β' à déterminer). En déduire que la méthode de Jacobi est convergente.
Corrigé : La matrice de Jacobi $J = D^{-1}(E+F)$ a la même forme que A avec $\beta' = \frac{\beta}{\alpha+2\beta}$ et $\alpha' = -2\beta' = \frac{-2\beta}{\alpha+2\beta}$, on en déduit ses valeurs propres :

$$\lambda'^k = \frac{-2\beta}{\alpha+2\beta}\left[1 - 2\sin^2\frac{k\pi}{2(n+1)}\right] = \frac{-2\beta}{\alpha+2\beta}\cos\frac{k\pi}{(n+1)}, \quad k = 1\ldots n$$

et donc $\rho(J) = \frac{2\beta}{\alpha+2\beta}\cos\frac{\pi}{(n+1)} < 1$, la méthode de Jacobi est donc convergente pour le système (2).

8. On pose dans la suite $A = D - E - F$ avec les conventions habituelles (D diagonale de A, $-E$ triangle inférieur, $-F$ triangle supérieur).
Montrez que le polynôme caractéristique de la matrice d'itération $\mathcal{L}_1 = (D-E)^{-1}F$ de la méthode de Gauss-Seidel s'écrit pour $\mu \neq 0$:

$$P_{\mathcal{L}_1}(\mu) = \frac{\mu^{\frac{n}{2}}}{\det(D-E)}\det(\mu^{-\frac{1}{2}}F + \mu^{\frac{1}{2}}E - \mu^{\frac{1}{2}}D)$$

Corrigé :

$$\begin{aligned}P_{\mathcal{L}_1}(\mu) =&\det((D-E)^{-1}F - \mu I) = \det((D-E)^{-1})\det(F - \mu(D-E))\\ =&\det((D-E)^{-1})\det(\mu^{\frac{1}{2}}(\mu^{-\frac{1}{2}}F - \mu^{\frac{1}{2}}(D-E)))\\ =&\mu^{\frac{n}{2}}\det((D-E)^{-1})\det(\mu^{-\frac{1}{2}}F + \mu^{\frac{1}{2}}E - \mu^{\frac{1}{2}}D)\end{aligned}$$

9. Vérifiez que pour toute matrice tridiagonale $A' = D' - E' - F'$ et $\rho \neq 0$ on a l'égalité :

$$R(D' - F' - E')R^{-1} = D' - \rho^{-1}F' - \rho E'$$

avec R la matrice diagonale de terme général $R_{i,i} = \rho^{i-1}$ pour $i = 1 \dots n$.

Corrigé : Dans un produit matriciel $C = AB$, un élément du produit est donné par $c_{i,j} = \sum_k a_{i,k}b_{k,j}$. Si la matrice A est diagonale il vient $c_{i,j} = a_{i,i}b_{i,j}$, soit chaque ligne de B est multipliée par l'élément de même ligne de A, inversement si B est diagonale il vient $c_{i,j} = a_{i,j}b_{j,j}$ soit chaque colonne de A est multipliée par l'élément de même colonne de B. Ceci appliqué au produit $R(D' - F' - E')R^{-1}$ montre que l'élément $a'_{i,j}$ de A' devient $r_{i,i}a'_{i,j}r^{-1}_{j,j}$, soit $a'_{i,j}\rho^{i-j}$. La matrice A' étant tridiagonale, les éléments de E' non nuls correspondent à $i - j = 1$, ils sont donc multipliés par ρ, ceux de F' correspondent à $i - j = -1$, ils sont multipliés par ρ^{-1} et les éléments Diagonaux de D' sont inchangés, d'où le résultat.

10. En déduire que le déterminant de $\rho^{-1}F' + \rho E' - D'$ est indépendant de ρ pour $\rho \neq 0$.

Corrigé :

$$\det(\rho^{-1}F' + \rho E' - D') = \det(R(F' + E' - D')R^{-1}) = \det(F' + E' - D')$$

11. Déduire des questions précédentes que le rayon spectral de la matrice d'itération de Gauss-Seidel est donné par :

$$\rho(\mathcal{L}_1) = \rho(J)^2$$

pour toute système à matrice tridiagonale.

Corrigé : Des questions précédentes on déduit que pour tout $\rho \neq 0$:

$$P_{\mathcal{L}_1}(\mu) = \mu^{\frac{n}{2}} \det((D - E)^{-1}) \det(\rho^{-1}\mu^{-\frac{1}{2}}F + \rho\mu^{\frac{1}{2}}E - \mu^{\frac{1}{2}}D)$$

Pour $\rho = \mu^{-\frac{1}{2}}$ on obtient :

$$P_{\mathcal{L}_1}(\mu) = \mu^{\frac{n}{2}} \det((D - E)^{-1}) \det(F + E - \mu^{\frac{1}{2}}D) = \mu^{\frac{n}{2}} \det((D - E)^{-1}) \det(D) P_J(\mu^{\frac{1}{2}})$$

Ce qui prouve que les valeurs propres de la matrice d'itération de Gauss-Seidel sont les carrés de celles de la matrice de Jacobi et donc $\rho(\mathcal{L}_1) = \rho(J)^2$.

12. En déduire que la méthode de Gauss-Seidel appliquée à la résolution du système (2) est convergente. Que peut-on dire de la vitesse de convergence des méthodes de Jacobi et Gauss-Seidel appliquées au sytème (2) quand n augmente ? Quelle méthode vaut il mieux utiliser pour résoudre (2) quand n est grand ?

Corrigé : On a montré que $\rho(J) < 1$ pour le système (2), on en déduit $\rho(\mathcal{L}_1) < \rho(J) < 1$ et la méthode de Gauss-Seidel est convergente et elle converge plus vite que la méthode de Jacobi.