

EXAMEN DE RATTRAPAGE D'ANALYSE NUMÉRIQUE I – CORRIGÉ

22 janvier 2009 – DURÉE 2h00

*La consultation et l'échange des documents, et l'utilisation des calculatrices sont interdits.
L'utilisation des 2 feuilles manuscrites recto-verso, format A4 est autorisée*

- Ne pas détacher les feuilles.
 - Utiliser l'espace blanc pour vos réponses et le verso pour brouillon.
 - Pensez à indiquer votre nom sur chaque feuille
-
-

Exercice 1

Soit b un entier positif qu'on prendra comme base de numérotation. Étant donné un entier non négatif x_0 , on aura la représentation en base b de ce nombre sous la forme

$$x_0 = d_0 + d_1b + d_2b^2 + \dots + d_nb^n \quad (0.1)$$

où d_i sont des entiers satisfaisant à la relation

$$0 \leq d_i < b; \quad i = 0, \dots, n \quad (0.2)$$

Dans cet exercice on va relaxer la condition que la base b soit un entier et nous allons considérer que b est un réel supérieur à 1. De même on va relaxer la condition que d_i soient des entiers et on considérera les nombres d_i comme pouvant être des réels. Le problème est que dans ce cas la représentation (0.1) n'est pas unique. Afin de rendre cette représentation unique, nous allons imposer des conditions supplémentaires aux conditions (0.2), à savoir que les quantités x_i , définies par

$$\begin{aligned} x_n &= d_n \\ x_i &= d_i + bx_{i+1}; \quad i = n-1, n-2, \dots, 0 \end{aligned} \quad (0.3)$$

soient toutes des quantités entières.

1 Montrer que, grâce aux conditions (0.2) et (0.3), la représentation (0.1) est unique.

Indication : Montrer d'abord que d_0 est l'unique entier non négatif et inférieur à b tel que $x_0 - d_0$ est un multiple entier de b . Généraliser, ensuite, en montrant que d_i est l'unique nombre non négatif et inférieur à b , tel que $x_i - d_i$ soit un multiple entier de b .

SOL.- En utilisant la recursion (0.3), nous avons $x_0 = d_0 + bx_1$, c'est-à-dire que d_0 est le reste de la division entière de x_0 par b . Étant donné que $0 \leq d_0 < b$, d_0 est unique par référence au théorème de l'arithmétique qui affirme que si x_0 est un entier non négatif et b un réel positif, alors il existe un unique d_0 réel et un unique x_1 entier, tels que

$$\begin{aligned} 0 &\leq d_0 < b \\ x_0 &= d_0 + bx_1 \end{aligned}$$

Pour la généralisation, nous observons que nous avons $x_i = d_i + bx_{i+1}$, avec x_i, x_{i+1} entiers et $0 \leq d_i < b$. Nous pouvons donc appliquer le raisonnement précédent.

La question précédente nous permet de nous assurer que (0.1) est unique quand la base b et les nombres d_i sont des réels non négatifs. Il convient maintenant de s'assurer que la précision des calculs ne souffre pas du fait que les d_i sont des réels. C'est l'objectif de la question suivante qui permet de montrer que dans la représentation (0.1) ce qui compte est la partie entière des nombres d_i .

2 Notons par d'_i la partie entière de d_i , c'est-à-dire

$$d'_i = \lfloor d_i \rfloor; \quad i = 0, \dots, n$$

où par $\lfloor x \rfloor$ on note l'entier immédiatement inférieur au réel x . Montrer que la suite x_0, x_1, \dots, x_n est déterminée de façon unique par la suite d'_0, d'_1, \dots, d'_n .

Indication. Montrer d'abord que d_n est un entier. Montrer, ensuite, qu'en construisant $x'_{n-1} = d'_{n-1} + bx_n$, nous avons que $x_{n-1} = \lceil x'_{n-1} \rceil$, où nous avons noté par $\lceil x \rceil$ l'entier immédiatement supérieur au réel x . Et, en général, montrer que le réel x_0 peut être évalué par la recursion

$$\begin{aligned} x_n &= d'_n \\ x_i &= \lceil d'_i + bx_{i+1} \rceil; \quad i = n-1, n-2, \dots, 0 \end{aligned}$$

SOL.- En utilisant la recursion (0.3), nous avons $x_n = d_n$ et comme x_n doit être un entier, d_n l'est aussi et donc $d'_n = d_n$.

Soit maintenant $x'_{n-1} = d'_{n-1} + bx_n$. Puisque $x_{n-1} = d_{n-1} + bx_n$, on a $x_{n-1} - x'_{n-1} = d_{n-1} - d'_{n-1} < 1$. Nous avons aussi que $x_{n-1} \geq x'_{n-1}$. Donc $x_{n-1} = \lceil x'_{n-1} \rceil$. On peut généraliser facilement en suivant la même démarche.

Exercice 2

Soient deux matrices $\mathbf{A}, \mathbf{B} \in \mathbb{R}^{n \times n}$. On considère leur somme $\mathbf{C} = \mathbf{A} + \mathbf{B}$. À cause des erreurs d'arrondi, on récupère la matrice $\mathbf{m}(\mathbf{C})$ avec terme général $m(c_{ij}) = m(a_{ij} + b_{ij}); 1 \leq i \leq m, 1 \leq j \leq n$.

Calculer une borne supérieure pour la quantité

$$\frac{\|\mathbf{C} - \mathbf{m}(\mathbf{C})\|_\infty}{\|\mathbf{C}\|_\infty}$$

$$\begin{aligned} \text{SOL.-} \frac{\|\mathbf{C} - \mathbf{m}(\mathbf{C})\|_\infty}{\|\mathbf{C}\|_\infty} &= \frac{\max_i \sum_j |a_{ij} + b_{ij} - m(a_{ij} + b_{ij})|}{\max_i \sum_j |a_{ij} + b_{ij}|} = \\ \frac{\max_i \sum_j |a_{ij} + b_{ij} - (a_{ij} + b_{ij})(1 + \eta_{ij})|}{\max_i \sum_j |a_{ij} + b_{ij}|} &= \frac{\max_i \sum_j |(a_{ij} + b_{ij})\eta_{ij}|}{\max_i \sum_j |a_{ij} + b_{ij}|} \leq \frac{\max_i \sum_j \eta |a_{ij} + b_{ij}|}{\max_i \sum_j |a_{ij} + b_{ij}|} = \eta \end{aligned}$$

Exercice 3

Rappelons qu'un algorithme est une application $\phi : D \rightarrow \mathbb{R}^n$, avec $D \subset \mathbb{R}^m$, telle que pour un $\mathbf{x} \in D$, on a $\mathbf{y} = \phi(\mathbf{x})$.

Rappelons aussi qu'étant donné un algorithme ϕ , le facteur de conditionnement de cet algorithme est donné par la formule

$$\kappa = \mathbf{x} \cdot \frac{J(\phi(\mathbf{x}))}{\phi(\mathbf{x})}$$

où $J(\phi(\mathbf{x}))$ est le jacobien de ϕ .

Soit le système d'équations linéaires

$$\mathbf{Ax} = \mathbf{b}; \quad \mathbf{A} \in \mathbb{R}^{n \times n}; \mathbf{x}, \mathbf{b} \in \mathbb{R}^n \tag{0.4}$$

dont on cherche à calculer la solution \mathbf{x} .

1. Formuler la solution de ce problème en utilisant la notion de l'algorithme.
SOL.- $\mathbf{x} = \phi(\mathbf{b}) = \mathbf{A}^{-1}\mathbf{b}$.

2. Calculer la norme $\|\kappa\|_2$ du facteur de conditionnement pour ce problème.
SOL.- $J(\phi(\mathbf{x})) = \mathbf{A}^{-1}$. Donc

$$\|\kappa\|_2 = \frac{\|\mathbf{b}\|_2 \cdot \|\mathbf{A}^{-1}\|_2}{\|\mathbf{A}^{-1}\mathbf{b}\|_2} = \frac{\|\mathbf{Ax}\|_2 \cdot \|\mathbf{A}^{-1}\|_2}{\|\mathbf{x}\|_2} \leq \frac{\|\mathbf{A}\|_2 \cdot \|\mathbf{x}\|_2 \cdot \|\mathbf{A}^{-1}\|_2}{\|\mathbf{x}\|_2} = \|\mathbf{A}\|_2 \cdot \|\mathbf{A}^{-1}\|_2$$

3. La solution par de méthodes directes du système d'équations linéaires (0.4) conduit toujours à factoriser la matrice \mathbf{A} en deux matrices \mathbf{B} et \mathbf{C} . Nous avons ainsi

$$\mathbf{A} = \mathbf{B} \cdot \mathbf{C} \quad (0.5)$$

et la résolution de (0.4) se réalise par la résolution de deux systèmes

$$\mathbf{B}\mathbf{y} = \mathbf{b}; \quad \mathbf{C}\mathbf{x} = \mathbf{y} \quad (0.6)$$

reputés plus simples à résoudre.

En utilisant le conditionnement des matrices, établir un critère qui permet de savoir si la méthode de résolution donnée par (0.6) peut être utilisée sans perte notable de la précision du résultat à la place de la résolution par inversion de la matrice \mathbf{A} , c'est-à-dire $\mathbf{x} = \mathbf{A}^{-1}\mathbf{b}$.

SOL.- On forme la quantité

$$\kappa = \frac{\text{cond}(\mathbf{B}) \cdot \text{cond}(\mathbf{C})}{\text{cond}(\mathbf{A})}$$

Si $\kappa \gg 1$, on a une perte importante de la précision du résultat.

Exercice 4

Soit \mathbf{A} une matrice inversible d'ordre n , \mathbf{b} un élément de \mathbb{R}^n ; on note \mathbf{x} la solution du système $\mathbf{Ax} = \mathbf{b}$.

Soit $\Delta\mathbf{A}$ une matrice carrée d'ordre n . On suppose que la matrice $\mathbf{A} + \Delta\mathbf{A}$ est également inversible.

On considère le système linéaire perturbé $(\mathbf{A} + \Delta\mathbf{A})\mathbf{x}' = \mathbf{b}$.

On écrit sa solution \mathbf{x}' sous la forme $\mathbf{x}' = \mathbf{x} + \delta\mathbf{x}$ où \mathbf{x} est la solution du système initial $\mathbf{Ax} = \mathbf{b}$.

1. Établir une expression de $\delta\mathbf{x}$ en fonction de \mathbf{x} , \mathbf{A} et $\Delta\mathbf{A}$

SOL.- $(\mathbf{A} + \Delta\mathbf{A})\mathbf{x}' = \mathbf{b} \Rightarrow (\mathbf{A} + \Delta\mathbf{A})(\mathbf{x} + \delta\mathbf{x}) = \mathbf{b} \Leftrightarrow (\mathbf{A} + \Delta\mathbf{A})\delta\mathbf{x} + \Delta\mathbf{A}\mathbf{x} = \mathbf{0} \Leftrightarrow \delta\mathbf{x} = -(\mathbf{A} + \Delta\mathbf{A})^{-1}\Delta\mathbf{A}\mathbf{x}$

2. Que se passe-t-il si $\Delta\mathbf{A} = \mathbf{0}$?

SOL.- On trouve $\delta\mathbf{x} = -(\mathbf{A} + \mathbf{0})^{-1}\mathbf{0}\mathbf{x} = \mathbf{0}$: la perturbation sur la solution est nulle si celle que la matrice initiale du système l'est aussi.

3. Montrer que l'on peut obtenir une expression de $\delta\mathbf{x}$ en fonction de $\mathbf{x} + \delta\mathbf{x}$, \mathbf{A} et $\Delta\mathbf{A}$ sous la forme :

$$\delta\mathbf{x} = -\mathbf{A}^{-1}\Delta\mathbf{A}(\mathbf{x} + \delta\mathbf{x})$$

SOL.- $(\mathbf{A} + \Delta\mathbf{A})(\mathbf{x} + \delta\mathbf{x}) = \mathbf{b} \Leftrightarrow \mathbf{A}\delta\mathbf{x} = -\Delta\mathbf{A}(\mathbf{x} + \delta\mathbf{x}) \Leftrightarrow \delta\mathbf{x} = -\mathbf{A}^{-1}\Delta\mathbf{A}(\mathbf{x} + \delta\mathbf{x})$

4. En déduire une majoration de $\frac{\|\delta\mathbf{x}\|}{\|\mathbf{x}\|}$ sous la forme :

$$\frac{\|\delta\mathbf{x}\|}{\|\mathbf{x}\|} \leq \text{cond}(\mathbf{A}) \frac{\|\Delta\mathbf{A}\|}{\|\mathbf{A}\|} \frac{1}{(1 - \|\mathbf{A}^{-1}\| \|\Delta\mathbf{A}\|)}$$

SOL.- De $\delta\mathbf{x} = -\mathbf{A}^{-1}\Delta\mathbf{A}(\mathbf{x} + \delta\mathbf{x})$ on déduit

$$\|\delta\mathbf{x}\| = \|\mathbf{A}^{-1}\Delta\mathbf{A}(\mathbf{x} + \delta\mathbf{x})\| \leq \|\mathbf{A}^{-1}\| \|\Delta\mathbf{A}\| (\|\mathbf{x}\| + \|\delta\mathbf{x}\|)$$

donc

$$(1 - \|\mathbf{A}^{-1}\| \|\Delta\mathbf{A}\|) \|\delta\mathbf{x}\| \leq \|\mathbf{A}^{-1}\| \|\Delta\mathbf{A}\| \|\mathbf{x}\|$$

soit

$$\begin{aligned} \frac{\|\delta\mathbf{x}\|}{\|\mathbf{x}\|} &\leq \|\mathbf{A}^{-1}\| \|\Delta\mathbf{A}\| \frac{1}{(1 - \|\mathbf{A}^{-1}\| \|\Delta\mathbf{A}\|)} = \|\mathbf{A}^{-1}\| \|\mathbf{A}\| \frac{\|\Delta\mathbf{A}\|}{\|\mathbf{A}\|} \frac{1}{(1 - \|\mathbf{A}^{-1}\| \|\Delta\mathbf{A}\|)} \\ &= \text{cond}(\mathbf{A}) \frac{\|\Delta\mathbf{A}\|}{\|\mathbf{A}\|} \frac{1}{(1 - \|\mathbf{A}^{-1}\| \|\Delta\mathbf{A}\|)} \end{aligned}$$

On considère le système linéaire $(S_\varepsilon) : \mathbf{A}_\varepsilon \mathbf{x} = \mathbf{b}$ sous la forme

$$(S_\varepsilon) \begin{cases} (100 + \varepsilon)x + y = 101 \\ (1 + \varepsilon)y = 1 \end{cases}$$

avec

$$\mathbf{A}_\varepsilon = \begin{bmatrix} 100 + \varepsilon & 1 \\ 0 & 1 + \varepsilon \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} 101 \\ 1 \end{bmatrix} \quad \text{et} \quad \mathbf{x} = \begin{bmatrix} x \\ y \end{bmatrix}$$

où ε représente une erreur sur les termes diagonaux de la matrice du système.

5. Résoudre le système non perturbé $(S_0) : \mathbf{A}_0 \mathbf{x} = \mathbf{b}$

SOL.- SOLUTION :

$$(S_0) \begin{cases} 100x + y = 101 \\ y = 1 \end{cases}$$

donc la solution est $(x, y) = (1, 1)$

6. Calculez le conditionnement pour la norme $\|\cdot\|_\infty$ de \mathbf{A}_0 , où \mathbf{A}_0 désigne la matrice du système non perturbé ($\varepsilon = 0$)

SOL.-

$$\text{cond}_\infty(\mathbf{A}_0) = \|\mathbf{A}_0\|_\infty \|\mathbf{A}_0^{-1}\|_\infty$$

avec

$$\mathbf{A}_0 = \begin{pmatrix} 100 & 1 \\ 0 & 1 \end{pmatrix}$$

$$\|\mathbf{A}_0\|_\infty = \max_{i=1,2} \sum_{j=1}^2 |\mathbf{A}_{ij}| = 101$$

et

$$\mathbf{A}_0^{-1} = \begin{pmatrix} \frac{1}{100} & -\frac{1}{100} \\ 0 & 1 \end{pmatrix}$$

$$\|\mathbf{A}_0^{-1}\|_\infty = \max_{i=1,2} \sum_{j=1}^2 |A_{ij}| = 1$$

donc

$$\text{cond}_\infty(\mathbf{A}_0) = 101$$

7. Donnez, en fonction de ε et sans calculer explicitement la solution du système perturbé une majoration de l'incertitude

$$\frac{\|\delta \mathbf{x}\|_\infty}{\|\mathbf{x}\|_\infty} = \frac{\left\| \begin{bmatrix} \delta x \\ \delta y \end{bmatrix} \right\|_\infty}{\left\| \begin{bmatrix} x \\ y \end{bmatrix} \right\|_\infty}$$

SOL.- D'après la question 4 on a

$$\frac{\|\delta \mathbf{x}\|_\infty}{\|\mathbf{x}\|_\infty} \leq \text{cond}_\infty(\mathbf{A}) \frac{\|\Delta \mathbf{A}\|_\infty}{\|\mathbf{A}\|_\infty} \frac{1}{(1 - \|\mathbf{A}^{-1}\|_\infty \|\Delta \mathbf{A}\|_\infty)}$$

Or on a

$$\Delta \mathbf{A} = \begin{pmatrix} \varepsilon & 0 \\ 0 & \varepsilon \end{pmatrix}$$

donc

$$\|\Delta \mathbf{A}\|_\infty = |\varepsilon|$$

En réutilisant les résultats précédents on obtient :

$$\frac{\|\delta \mathbf{x}\|_\infty}{\|\mathbf{x}\|_\infty} \leq 101 \frac{|\varepsilon|}{101} \frac{1}{(1 - |\varepsilon|)} = \frac{|\varepsilon|}{1 - |\varepsilon|}$$