Analyse Numérique

Matrices et systèmes linéaires

Laurence Lamoulie

EISTI

Ecole Internationale des Sciences du Traitement de l'Information

Problèmes envisagés

Deux problèmes types

- Résolution de système linéaire

$$AX = B$$

  où A est une matrice carrée de taille nxn

- Détermination de l'inverse d'une matrice A :

$$B = A^{-1}$$

Les méthodes applicables "à la main" ne sont pas utilisables à la machine.

Méthodes directes pour résoudre $AX = B$

Principe des méthodes directes

On remplace le problème initial $AX = B$ par un problème équivalent plus simple à résoudre. On s'appuie en général sur une transformation de A.
Exemples : Gauss, LU, Cholesky...

Problèmes des méthodes directes

- Stabilité
- Erreurs numériques
- Structures matricielles (remplissage)

Méthodes itératives pour résoudre $AX = B$

Principe des méthodes itératives

On approche la solution X du problème initial $AX = B$ par une suite de vecteurs X_k qui converge vers X selon un algorithme donné
Exemples : Jacobi, Gauss-Seidel, Relaxation, SSOR, Gradient conjugué...

Problèmes des méthodes itératives

- Convergence
- Vitesse de convergence
- Préconditionnement...

Besoin d'un arsenal théorique

Prérequis

A voir/revoir absolument

- Applications linéaires : injection, surjection, famille libre, famille génératrice, noyau, image, base, vecteurs
- Algèbre générale : norme, réduction des matrices (valeurs et vecteurs propres, déterminant)
- Théorie des matrices : singularité, rang, orthogonalité, diagonale dominante, définie positivité, symétrie

Pour vous aider : en ligne sur Arel, *Algèbre Matricielle et Numérique*, Claude Brezinski

Deux exemples de résultats classiques

Majoration du rayon spectral $\rho(A) = \max_{1 \leq i \leq n} |\lambda_i|$

Théorème de Gershgorin : Localisation des valeurs propres
Toutes les valeurs propres d'une matrice se trouvent dans l'union des cercles de Gershgorin

$$C_i = \left\{ z \in \mathbb{R}, |z - a_{ii}| \leq \sum_{k=1, k \neq i}^{n} |a_{ik}| \right\} \qquad i = 1...n$$

Exemple : La matrice A donnée par $A = \begin{pmatrix} 3 & -1 & 1 \\ -2 & 4 & 2 \\ -1 & 1 & 5 \end{pmatrix}$ a ses valeurs propres dans la réunion des cercles $C_1(3, 2)$, $C_2(4, 4)$, $C_3(5, 2)$. En fait les valeurs propres sont 2, 4, 6.

Deux exemples de résultats classiques

Majoration du quotient de Rayleigh

Définition : Quotient de Rayleigh

Si A est une matrice carrée, le quotient de Rayleigh est

$$R(v) = \frac{v^T A v}{v^T v}, \qquad v \in \mathbb{R}^n, v \neq 0$$

Résultat : Si A est une matrice carrée **symétrique** alors on a

$$|\lambda_{\min}| \leq |R(v)| \leq |\lambda_{\max}|$$

où $\lambda_{\min}$ et $\lambda_{\max}$ sont respectivement la plus petite et la plus grande valeur propre de A en module.

Normes vectorielles

Définitions

Pour tout vecteur $x \in \mathbb{R}^n$, on note :

- Norme euclidienne

$$\|x\|_2 = \sqrt{x^T x} = \sqrt{\sum_{i=1}^{n} x_i^2}$$

- Norme l_p

$$\|x\|_p = \left(\sum_{i=1}^{n} x_i^p \right)^{1/p}$$

- Norme du sup

$$\|x\|_\infty = \max_{i=1,\dots n} |x_i|$$

Normes vectorielles

Théorème 3.1.2 : Normes équivalentes

Toutes les normes $\|.\|$ de $\mathbb{R}^n$ sont **équivalentes**, i.e. pour tout couple de normes $\|.\|_\alpha$ et $\|.\|_\beta$, il existe deux constantes positives C_1 et C_2 telles que

$$C_1 \|x\|_\alpha \leq \|x\|_\beta \leq C_2 \|x\|_\alpha , \qquad \forall x \in \mathbb{R}^n$$

Normes matricielles

Définition : Norme subordonnée

On appelle norme matricielle **subordonnée** à la norme vectorielle $||.||$ la norme définie par :

$$\forall A \in \mathbb{R}^{nxn}, \qquad |||A||| = \max_{v \in \mathbb{R}^n, v \neq 0} \frac{||Av||}{||v||}$$

Définition : Norme consistante

Une norme matricielle $||.||$ est dite **consistante** avec une norme vectorielle $||.||$ si :

$$\forall A \in \mathbb{R}^{nxn}, \forall v \in \mathbb{R}^n, \qquad ||Av|| \leq ||A|| \, ||v||$$

Normes matricielles

Conséquence des définitions

Toute norme subordonnée est consistante

Définition : Norme sous multiplicative

Une norme matricielle $\|.\|$ est dite **sous multiplicative** si elle vérifie :

$$\forall A, B \in \mathbb{R}^{nxn}, \qquad \|AB\| \leq \|A\| \, . \, \|B\|$$

Conséquence de cette définition

Les normes consistantes et subordonnées sont sous-multiplicatives

Normes matricielles

Définitions des normes subordonnées aux normes usuelles

Pour toute matrice $A \in \mathbb{R}^{nxn}$ on note

- Norme 1 : Somme des v.a. des termes en colonnes

$$\|A\|_1 = \max_{j=1,\dots n} \sum_{i=1}^{n} |a_{ij}|$$

- Norme 2

$$\|A\|_2 = \rho^{1/2}(A^* A) = \mu_1 \text{ plus grande valeur singulière de } A$$

- Norme infinie : Somme des v.a. des termes en lignes

$$\|A\|_\infty = \max_{i=1,\dots n} \sum_{j=1}^{n} |a_{ij}|$$

Normes matricielles

Exemple de norme non subordonnée

Pour toute matrice $A \in \mathbb{R}^{nxn}$ on appelle **norme de Frobenius ou de Schur**, la norme suivante

$$\|A\|_S = \left(\sum_{i,j=1}^{n} |a_{ij}|^2 \right)^{1/2} = \sqrt{tr(A^T A)}$$

Remarque importante

Pour toute matrice $A \in \mathbb{R}^{nxn}$ hermitienne $(A^* = A)$ on a

$$\|A\|_2 = \rho^{1/2}(A^2) = \lambda_1 \text{ plus grande valeur propre de } A$$

Cette norme ne doit pas être confondue avec la norme de Frobenius ou de Schur

Conditionnement

Définition du conditionnement d'une matrice

Pour toute matrice $A \in \mathbb{R}^{nxn}$ on appelle **conditionnement** de A relativement à la norme de Holder $\|.\|$ la quantité

$$cond(A) = \|A\| . \left\|A^{-1}\right\|$$

Résultat important

Pour toute matrice $A \in \mathbb{R}^{nxn}$ on a

$$cond(A) \geq 1$$

Conditionnement

Démonstration de $cond(A) \geq 1$

On a $\|A.A^{-1}\| = \|I\| = 1$
Comme les normes de Holder sont sous multiplicatives, on en déduit :

$$1 = \left\|A.A^{-1}\right\| \leq \|A\| . \left\|A^{-1}\right\| \qquad \text{soit} \qquad cond(A) \geq 1$$

Remarques

La valeur du conditionnement dépend de la norme utilisée pour l'évaluer, mais pour deux normes $\|.\|_\alpha$ et $\|.\|_\beta$ équivalentes, on a

$$C_1^2 cond_\alpha(A) \leq cond_\beta(A) \leq C_2^2 cond_\alpha(A)$$

Un **bon** conditionnement est un conditionnement proche de 1.

Principe

Résolution de système

On suppose que l'on résout $AX = B$ où

- A est une matrice de $\mathbb{R}^{nxn}$,
- X et B sont des vecteurs de $\mathbb{R}^{n}$.

On va supposer successivement :

1. que l'on commet une erreur sur le second membre B, i.e. on résout $A(X + \Delta X) = B + \Delta B$
2. que l'on commet une erreur sur la matrice A, i.e. on résout $(A + \Delta A)(X + \Delta X) = B$
3. que l'on commet une erreur sur les deux, i.e. on résout $(A + \Delta A)(X + \Delta X) = B + \Delta B$.

et on va estimer les répercutions sur la solution X du système.

Cas 1 : Perturbation de B

On résout

$$A(X + \Delta X) = B + \Delta B$$

or $AX = B$ donc

$$A\Delta X = \Delta B \qquad \text{soit} \qquad \Delta X = A^{-1}\Delta B$$

d'où

$$\|\Delta X\| \leq \left\|A^{-1}\right\| \|\Delta B\|$$

Or $AX = B$ donc $\|B\| \leq \|A\| \|X\|$ donc

$$\|\Delta X\| \|B\| \leq \|A\| \left\|A^{-1}\right\| \|\Delta B\| \|X\|$$

et si $\|B\| \neq 0$ on en déduit

$$\boxed{\frac{\|\Delta X\|}{\|X\|} \leq cond(A)\frac{\|\Delta B\|}{\|B\|}}$$

Cas 2 : Perturbation de A

On résout

$$(A+\Delta A)(X+\Delta X)=B$$

or $AX=B$ donc

$$A\Delta X=-\Delta A(X+\Delta X) \qquad \text{soit} \qquad \Delta X=-A^{-1}\Delta A(X+\Delta X)$$

d'où (en utilisant le fait que les normes matricielles sont sous-multiplicatives)

$$\|\Delta X\| \leq \left\|A^{-1}\right\| \|\Delta A\| \left(\|X\|+\|\Delta X\|\right)$$

Soit

$$\|\Delta X\| \left(1-\left\|A^{-1}\right\| \|\Delta A\|\right) \leq \left\|A^{-1}\right\| \|\Delta A\| \|X\|$$

et si $\|X\| \neq 0$ on en déduit

$$\frac{\|\Delta X\|}{\|X\|} \leq \frac{\|A^{-1}\| \|\Delta A\|}{1-\|A^{-1}\| \|\Delta A\|} = \frac{\|A^{-1}\| \|A\|}{1-\|A^{-1}\| \|\Delta A\|} \cdot \frac{\|\Delta A\|}{\|A\|}$$

Cas 2 : Perturbation de A

On suppose que la perturbation de A est "petite" au sens où

$$\left\| A^{-1} \right\| \|\Delta A\| \ll 1$$

et on déduit alors que

$$\boxed{\frac{\|\Delta X\|}{\|X\|} \leq cond(A).\frac{\|\Delta A\|}{\|A\|}}$$

Cas 3 : Perturbations de A et de B

On résout

$$(A + \Delta A)(X + \Delta X) = B + \Delta B$$

De façon analogue aux cas précédents, on montre que :

$$\boxed{\frac{\|\Delta X\|}{\|X\|} \leq cond(A). \left(\frac{\|\Delta A\|}{\|A\|} + \frac{\|\Delta B\|}{\|B\|} \right)}$$

Exemple (from M. Schatzman)

Importance du conditionnement matriciel

On considère la matrice $A = \begin{pmatrix} 1 & 1 \\ 1 & 1.0001 \end{pmatrix}$ et on résout successivement :

$$Ax = \begin{pmatrix} 2 \\ 2 \end{pmatrix} \qquad \text{et} \qquad A(x + \delta x) = \begin{pmatrix} 2 \\ 2.0001 \end{pmatrix}$$

On trouve :

$$x = \begin{pmatrix} 2.0 \\ 0 \end{pmatrix} \qquad \text{et} \qquad x + \delta x = \begin{pmatrix} 1.0 \\ 1.0 \end{pmatrix}$$

sachant que $cond_2(A) = 40004$.

Références

- *Analyse numérique*, M. Schatzman, InterEditions, 1991,
- *Analyse numérique matricielle appliquée à l'art de l'ingénieur*, P. Lascaux R. Theodor, Masson, 1993.
- *Algèbre matricielle numérique*, C. Brezinski, document pdf