Conditionnement. Théorème de Courant-Fisher

1. Conditionnement d'un système linéaire

Soit $A \in \mathcal{M}_{n,n}(\mathbb{R})$ une matrice inversible et $b \in \mathbb{R}^n$ un vecteur colonne. On cherche à étudier l'influence des erreurs d'arrondi de la matrice A et du vecteur b sur la solution $x \in \mathbb{R}^n$ du système $Ax = b$.

Un exemple classique. On cherche à résoudre le système $Ax = b$ avec

$$A = \begin{pmatrix} 10 & 7 & 8 & 7 \\ 7 & 5 & 6 & 5 \\ 8 & 6 & 10 & 9 \\ 7 & 5 & 9 & 10 \end{pmatrix} \text{ et } b = \begin{pmatrix} 32 \\ 23 \\ 33 \\ 31 \end{pmatrix}. \text{ La solution est } x = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}.$$

On considère les perturbations suivantes :

$$A+\Delta A = \begin{pmatrix} 10 & 7 & 8,1 & 7,2 \\ 7,08 & 5,04 & 6 & 5 \\ 8 & 5,98 & 9,89 & 9 \\ 6,99 & 4,99 & 9 & 9,98 \end{pmatrix}, \text{ c'est-à-dire } \Delta A = \begin{pmatrix} 0 & 0 & 0,1 & 0,2 \\ 0,08 & 0,04 & 0 & 0 \\ 0 & -0,02 & -0,11 & 0 \\ -0,01 & -0,01 & 0 & -0,02 \end{pmatrix}$$

$$\text{et } b + \delta b = \begin{pmatrix} 32,01 \\ 22,99 \\ 33,01 \\ 30,99 \end{pmatrix}, \text{ c'est-à-dire } \delta b = \begin{pmatrix} 0,01 \\ -0,01 \\ 0,01 \\ -0,01 \end{pmatrix}.$$

La solution de $Ay = b + \delta b$ est

$$y = \begin{pmatrix} 1,82 \\ -0,36 \\ 1,35 \\ 0,79 \end{pmatrix}, \text{ c'est-à-dire } \delta x = y - x = \begin{pmatrix} 0,82 \\ -1,36 \\ 0,35 \\ -0,21 \end{pmatrix}.$$

La solution de $(A + \Delta A)z = b$ est

$$z = \begin{pmatrix} -81 \\ 137 \\ -34 \\ 22 \end{pmatrix}, \text{ c'est-à-dire } \Delta x = z - x = \begin{pmatrix} -82 \\ 136 \\ -35 \\ 21 \end{pmatrix}.$$

On remarque que dans les deux cas, une petite erreur sur les coefficients de la matrice ou sur les coefficients du vecteur colonne entraîne une erreur importante sur les coefficients de la solution.

Exercice 1 : Définition du conditionnement. On note $||.||$ la norme vectorielle utilisée sur $\mathbb{R}^n$ et $|||.|||$ la norme matricielle sur $\mathcal{M}_{n,n}(\mathbb{R})$ subordonnée, c'est-à-dire la norme matricielle définie par $|||A||| = \sup\limits_{x \neq 0} \dfrac{||Ax||}{||x||}$.

1.1. Soit le système $Ax = b$ et le système perturbé $Ay = A(x + \delta x) = b + \delta b$. Montrer que $\dfrac{||\delta x||}{||x||} \leq |||A||| \times |||A^{-1}||| \dfrac{||\delta b||}{||b||}$.

1.2. Soit le système $Ax = b$ et le système perturbé $(A + \Delta A)z = (A + \Delta A)(x + \Delta x) = b$. Montrer que $\dfrac{||\Delta x||}{||x + \Delta x||} \leq |||A||| \times |||A^{-1}||| \dfrac{|||\Delta A|||}{|||A|||}$.

On remarque que, dans les deux cas, l'erreur relative sur les coefficients des données (A ou b) est amplifiée par la même constante qui ne dépend que de la matrice A. On appelle `conditionnement de la matrice` A inversible, le nombre $\mathrm{cond}(A) = |||A||| \times |||A^{-1}|||$.

Remarque 1 : Ces inégalités sont optimales au sens où il existe des matrices A, des vecteurs colonnes b et des perturbations ΔA et δb pour lesquels les inégalités deviennent des égalités (cf exercice 6).

Remarque 2 : Le déterminant et le conditionnement sont deux quantités indépendantes l'une de l'autre (cf exercice 7).

Remarque 3 : Un exemple classique de matrice mal conditionnée est la matrice de Hilbert dont les coefficients sont définis par $H_{i,j} = \dfrac{1}{i+j-1}$.

Exercice 2 : Propriétés du conditionnement. Montrer les propriétés suivantes :

2.1. $\mathrm{cond}(A) \geq 1$, $\mathrm{cond}(A) = \mathrm{cond}(A^{-1})$ et $\mathrm{cond}(A) = \mathrm{cond}(\alpha A)$ pour tout $\alpha \in \mathbb{R}^*$.

2.2. $\mathrm{cond}_2(A) = \dfrac{\mu_{\max}(A)}{\mu_{\min}(A)}$, où $\mu_{\max}(A)$ et $\mu_{\min}(A)$ sont la plus grande et la plus petite valeur singulière de A, c'est-à-dire la racine carrée de la plus grande et de la plus petite valeur propre de la matrice AA^*.

$\mathrm{cond}_2(A) = \dfrac{|\lambda_{\max}(A)|}{|\lambda_{\min}(A)|}$, où $\lambda_{\max}(A)$ et $\lambda_{\min}(A)$ sont la plus grande et la plus petite valeur propre de A, si A est une matrice normale (i.e. telle que $AA^* = A^*A$).

2.3. $\mathrm{cond}_2(A) = 1$ ssi $A = \alpha Q$, où $\alpha \in \mathbb{R}^*$ et Q est une matrice orthogonale.

2.4. $\mathrm{cond}_2(A) = \mathrm{cond}_2(AQ) = \mathrm{cond}_2(QA)$ pour toute matrice Q orthogonale.

Un système linéaire est `bien conditionné` si le conditionnement de sa matrice est proche de 1. Sinon, si son conditionnement est très grand, on dit que le système linéaire est `mal conditionné`.

2. Théorème de Courant-Fisher

Soit $A \in \mathcal{M}_{n,n}(\mathbb{R})$ une matrice symétrique. On appelle `quotient de Rayleigh` la quantité $r_A(x) = \dfrac{x^*Ax}{x^*x}$ définie pour tout $x \neq 0$.

Théorème de Courant-Fisher. *Soit $A \in \mathcal{M}_{n,n}(\mathbb{R})$ une matrice symétrique. On note ses valeurs propres $\lambda_1 \leq \cdots \leq \lambda_n$. Alors :*

(i) $\max\limits_{x\neq 0} r_A(x) = \lambda_n$ *et* $\min\limits_{x \neq 0} r_A(x) = \lambda_1$.

(ii) $\min\limits_{\dim V = k} \max\limits_{x \in V\setminus\{0\}} r_A(x) = \lambda_k$ *et* $\max\limits_{\dim V = k} \min\limits_{x \in V\setminus\{0\}} r_A(x) = \lambda_{n-k+1}$.

Exercice 3 : Démonstration du théorème.

3.1. En utilisant une base orthonormale $(e_i)_{1\leq i\leq n}$ de vecteurs propres de A, montrer le (i).

On note V_k le sous-espace vectoriel engendré par les vecteurs $(e_i)_{1\leq i \leq k}$ et W_k le sous-espace vectoriel engendré par les vecteurs $(e_i)_{k\leq i\leq n}$.

3.2. En utilisant l'espace V_k, montrer que $\min\limits_{\dim V = k} \max\limits_{x \in V\setminus\{0\}} r_A(x) \leq \lambda_k$.

3.3. Soit un sous-espace vectoriel V de dimension k. Montrer que $V \bigcap W_k \neq \{0\}$. En déduire que $\min\limits_{\dim V = k} \max\limits_{x \in V\setminus\{0\}} r_A(x) \geq \lambda_k$.

3.4. Vérifier que la deuxième égalité de (ii) découle de la première égalité. $\square$

3. Conditionnement d'un problème aux valeurs propres - Théorème de Bauer-Fike

Soit $A \in \mathcal{M}_{n,n}(\mathbb{R})$ une matrice. On cherche à étudier l'influence des erreurs d'arrondi de la matrice A sur le calcul des valeurs propres de A.

Un premier exemple. Soit $A = \begin{pmatrix} 0 & & & \varepsilon \\ 1 & \ddots & (0) & \\ & \ddots & \ddots & \\ (0) & & 1 & 0 \end{pmatrix}$. Le polynôme caractéristique de A vaut $X^n - \varepsilon$. Si $\varepsilon = 0$, les valeurs propres sont nulles. En revanche, si $n = 40$ et $\varepsilon = 10^{-40}$, les valeurs propres sont toutes de module 10^{-1}.

Théorème de Bauer-Fike. *Soit $A \in \mathcal{M}_{n,n}(\mathbb{R})$ une matrice diagonalisable dans $\mathbb{C}$ et P une matrice telle que $P^{-1}AP = D = \operatorname{diag}(\lambda_i,\, 1 \leq i \leq n)$. Soit $|||.|||$ une norme matricielle telle que, pour toute matrice diagonale, $|||\operatorname{diag}(d_i,\, 1 \leq i \leq n)||| = \max\limits_{1\leq i\leq n} |d_i|$. Alors, pour toute matrice δA, $\operatorname{Sp}(A + \delta A) \subset \bigcup\limits_{i=1}^{n} D_i$, où $D_i = \{z \in \mathbb{C},\, |z - \lambda_i| \leq \operatorname{cond}(P)|||\delta A|||\}$.*

Exercice 4 : Démonstration du théorème.
4.1. Montrer que si une matrice $I_n + M \in \mathcal{M}_{n,n}(\mathbb{R})$ est singulière, alors $|||M||| \geq 1$.
4.2. Soit λ une valeur propre de $A + \delta A$. Si $\lambda \neq \lambda_j$, $1 \leq j \leq n$, montrer l'égalité suivante :

$$P^{-1}\left(A + \delta A - \lambda I_n\right)P = \left(D - \lambda I_n\right)\left(I_n + (D - \lambda I_n)^{-1}P^{-1}(\delta A)P\right).$$

4.3. Conclure à partir des questions 4.1. et 4.2. $\square$

Remarque 4 : Les matrices normales sont très bien conditionnées pour le problème aux valeurs propres, puisqu'on peut choisir une matrice P orthogonale.

On considère maintenant le cas particulier d'une matrice symétrique et d'une perturbation symétrique.

Théorème. *Soit $A \in \mathcal{M}_{n,n}(\mathbb{R})$ et $B = A + \delta A \in \mathcal{M}_{n,n}(\mathbb{R})$ deux matrices symétriques de valeurs propres $\alpha_1 \leq \cdots \leq \alpha_n$ et $\beta_1 \leq \cdots \leq \beta_n$. Alors, on a $|\alpha_k - \beta_k| \leq |||\delta A|||_2$, $1 \leq k \leq n$.*

Exercice 5 : Démonstration du théorème. On utilise les mêmes notations que dans l'exercice 3.
5.1. En utilisant le théorème de Courant-Ficher, montrer que $\beta_k \leq \alpha_k + \max\limits_{x\in V_k} r_{\delta A}(x)$.
5.2. Montrer que $\max\limits_{x\in V_k} r_{\delta A}(x) \leq |||\delta A|||_2$.
5.3. Conclure. $\square$

Un second exemple. Soit la matrice $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$. Elle n'est pas diagonalisable et ses valeurs propres valent 1. La matrice $A_\varepsilon = \begin{pmatrix} 1 & 1 \\ \varepsilon & 1 \end{pmatrix}$ est, elle, diagonalisable et ses valeurs propres valent $1 \pm \sqrt{\varepsilon}$. On ne peut donc pas écrire d'estimation linéaire du type du théorème de Bauer-Fike pour des matrices non diagonalisables.

4. Exercices supplémentaires

Exercice 6 : Erreurs relatives et conditionnement
Schatzman, Analyse numérique
6.1. Soit $A = \begin{pmatrix} 1 & 100 \\ 0 & 1 \end{pmatrix}$. Calculer $\operatorname{cond}_1(A)$.

6.2. Considérons les systèmes suivants :

$$Ax = b = \begin{pmatrix} 100 \\ 1 \end{pmatrix} \text{ et } A(x + \delta x) = b + \delta b = \begin{pmatrix} 100 \\ 0 \end{pmatrix}.$$

Calculer la variation relative $||\delta x||_1/||x||_1$ de la solution et celle du second membre $||\delta b||_1/||b||_1$. Quel est le facteur d'amplification de l'erreur, c'est-à-dire $\dfrac{||\delta x||_1}{||x||_1} \times \dfrac{||b||_1}{||\delta b||_1}$?

Exercice 7 : Déterminant et conditionnement

Lascaux-Théodor, Analyse numérique matricielle appliquée à l'art de l'ingénieur

7.1. Calculer en fonction de n le déterminant et le conditionnement (en norme 1 ou ∞) de la matrice $A \in \mathcal{M}_{n,n}(\mathbb{R})$ définie par diag $(1, 10, \cdots, 10)$.

7.2. Faire de même avec la matrice $B \in \mathcal{M}_{n,n}(\mathbb{R})$ définie par

$$\begin{cases} B_{i,i} = 1, & 1 \le i \le n, \\ B_{i,i+1} = 2, & 1 \le i \le n-1, \\ B_{i,j} = 0, & \text{sinon.} \end{cases}$$

(Remarque : on peut montrer que $\mathrm{cond}_\infty(B) = \mathrm{cond}_1(B) = 3(2^n - 1)$.)

7.3. Qu'en concluez-vous ?

5. Références

Ciarlet, *Introduction à l'analyse numérique matricielle et à l'optimisation*, Dunod.

Lascaux-Théodor, *Analyse numérique matricielle appliquée à l'art de l'ingénieur*, Dunod.

Schatzman, *Analyse numérique*, Dunod.