Conditionnement d'une matrice

Florent Nacry (florent.nacry@univ-perp.fr)
Université Perpignan Via Domitia
Laboratoire LAMPS (Campus principal, Bâtiment B, Etage 1)

Le problème du conditionnement

Etant donnés $n \geq 1$ un entier, $A \in GL_n(\mathbb{K})$ et $b \in M_{n,1}(\mathbb{K})$, on s'intéresse aux effets de potentielles perturbations sur A et b (celles-ci pouvant provenir en pratique d'erreurs de mesures, de calculs, de troncatures,...) sur les solutions du système linéaire $Ax = b$ d'inconnue $x \in M_{n,1}(\mathbb{K})$.

Le problème du conditionnement

Etant donnés $n \geq 1$ un entier, $A \in GL_n(\mathbb{K})$ et $b \in M_{n,1}(\mathbb{K})$, on s'intéresse aux effets de potentielles perturbations sur A et b (celles-ci pouvant provenir en pratique d'erreurs de mesures, de calculs, de troncatures,...) sur les solutions du système linéaire $Ax = b$ d'inconnue $x \in M_{n,1}(\mathbb{K})$.

Commençons par considérer le cas où la matrice A est "perturbée" par une matrice P. Peut-on alors comparer la solution x_0 du système linéaire initial $Ax = b$ (connue théoriquement puisque $x_0 = A^{-1}b$) à la solution x_p du système linéaire perturbé $(A+P)x = b$? En particulier, si la perturbation est "petite" en un sens à préciser, l'erreur absolue $\|x_0 - x_p\|$ et/ou l'erreur relative $\|x_0 - x_p\|/\|x_p\|$ est-elle nécessairement "petite" ?

L'exemple classique

On pose

$$n = 4, \quad A = \begin{pmatrix} 10 & 7 & 8 & 7 \\ 7 & 5 & 6 & 5 \\ 8 & 6 & 10 & 9 \\ 7 & 5 & 9 & 10 \end{pmatrix} \quad \text{et} \quad b = \begin{pmatrix} 32 \\ 23 \\ 33 \\ 31 \end{pmatrix}.$$

On vérifie que

$$\{x \in M_{n,1}(\mathbb{K}) : Ax = b = (1,1,1,1)^T\},$$

i.e., le système linéaire $Ax = b$ d'inconnue $x \in M_{n,1}(\mathbb{K})$ a pour unique solution $(1,1,1,1)^T$. Avec

$$P = \begin{pmatrix} 0 & 0 & 0,1 & 0,2 \\ 0,08 & 0,04 & 0 & 0 \\ 0 & -0,02 & -0,11 & 0 \\ -0,01 & -0,01 & 0 & -0,02 \end{pmatrix}$$

on peut observer que

$$\left\{x \in M_{n,1}(\mathbb{K}) : (A+P)x = b\right\} = \left\{(-81,137,-34,22)^T\right\}.$$

Conclusion de l'exemple

De "petites perturbations" sur les coefficients de la matrice A peuvent donc induire de "grands effets" sur la solution. Remarquons que la matrice A de l'exemple ci-dessus a pourtant un "bon aspect" : elle est à coefficients entiers, symétrique, de déterminant 1 et d'inverse

$$A^{-1} = \begin{pmatrix} 25 & -41 & 10 & -6 \\ -41 & 68 & -17 & 10 \\ 10 & -17 & 5 & -3 \\ -6 & 10 & -3 & 2 \end{pmatrix}.$$

Conditionnement

Nous allons quantifier ce phénomène à travers la proposition suivante. Celle-ci fait appel à la notion de conditionnement (“condition number” en anglais) :

Conditionnement

Soient $n \geq 1$ un entier et $\|\cdot\|$ une norme sur $M_{n,1}(\mathbb{K})$. On appelle $\|\cdot\|$-conditionnement de $A \in GL_n(\mathbb{K})$ le réel

$$\mathrm{cond}_{\|\cdot\|}(A) = \mathscr{N}_{\|\cdot\|}(A)\mathscr{N}_{\|\cdot\|}(A^{-1}).$$

Erreur relative et conditionnement I

Le conditionnement permet une estimation de l'erreur relative mentionnée précédemment.

Erreur relative et conditionnement I

Le conditionnement permet une estimation de l'erreur relative mentionnée précédemment.

Proposition

Soient $n \geq 1$ un entier, $A \in GL_n(\mathbb{K})$, $b \in M_{n,1}(\mathbb{K}) \setminus \{0\}$, $P \in M_n(\mathbb{K})$ et $x_0, x_p \in M_{n,1}(\mathbb{K})$ satisfaisant

$$Ax_0 = b \quad \text{et} \quad (A+P)x_p = b.$$

Alors, on a $x_p \neq 0$ et pour n'importe quelle norme $\|\cdot\|$ sur $M_{n,1}(\mathbb{K})$

$$\frac{\|x_0 - x_p\|}{\|x_p\|} \leq \operatorname{cond}_{\|\cdot\|}(A) \frac{\mathscr{N}_{\|\cdot\|}(P)}{\mathscr{N}_{\|\cdot\|}(A)}. \tag{1}$$

Démonstration

Démonstration. De l'égalité $(A+P)x_p = b = Ax_0$, on tire

$$A(x_p - x_0) + Px_p = 0.$$

Il vient alors $x_p - x_0 = -A^{-1}Px_p$ ce qui entraîne tout de suite

$$\|x_p - x_0\| \leq \mathscr{N}_{\|\cdot\|}(A^{-1})\mathscr{N}_{\|\cdot\|}(P)\|x_p\|.$$

D'autre part, puisque $b \neq 0$, on a évidemment $x_p \neq 0$ et ceci permet d'écrire

$$\frac{\|x_0 - x_p\|}{\|x_p\|} \leq \mathscr{N}_{\|\cdot\|}(A^{-1})\mathscr{N}_{\|\cdot\|}(P).$$

Ceci termine la preuve. ■

Optimalité

L'inégalité (1) est la meilleure possible ("sharp" en anglais) au sens suivant : étant donnée une norme $\|\cdot\|$ sur $M_{n,1}(\mathbb{K})$ (où $n \geq 1$ est un entier) et $A \in GL_n(\mathbb{K})$, on peut trouver $b \in M_{n,1}(\mathbb{K})$ avec $b \neq 0$ et $P \in GL_n(\mathbb{K})$ tels que

$$\frac{\|x_0 - x_p\|}{\|x_p\|} = \mathrm{cond}_{\|\cdot\|}(A) \frac{\mathscr{N}_{\|\cdot\|}(P)}{\mathscr{N}_{\|\cdot\|}(A)},$$

où $x_0 \in M_{n,1}(\mathbb{K})$ (resp. $x_p \in M_{n,1}(\mathbb{K})$ avec $x_p \neq 0$) est tel que $Ax_0 = b$ (resp. $(A+P)x_p = b$).

Démonstration de l'optimalité

Pour voir cela, fixons $\|\cdot\|$ une norme sur $M_{n,1}(\mathbb{K})$ avec $n \geq 1$ entier, $A \in GL_n(\mathbb{K})$ et choisissons $w \neq 0$ tel que

$$\left\|A^{-1}w\right\| = \mathscr{N}_{\|\cdot\|}(A^{-1})\,\|w\|\,.$$

Posons $P = \beta \mathrm{I}_n$ où β est un réel positif non nul tel que $-\beta \notin \mathrm{sp}(A)$ (de sorte que $A+P = A+\beta\mathrm{I}_n \in GL_n(\mathbb{K})$). Posons également $b = (A+P)w \neq 0$, $x_0 = A^{-1}b$ et $x_p = w \neq 0$ (l'unique solution du système linéaire $(A+P)x = b$ d'inconnue $x \in M_{n,1}(\mathbb{K})$). On vérifie alors que

$$\frac{\|x_0 - x_p\|}{\|x_p\|} = \frac{\left\|\beta A^{-1}w\right\|}{\|w\|} = \beta\,\mathscr{N}_{\|\cdot\|}(A^{-1}) = \mathrm{cond}_{\|\cdot\|}(A)\,\frac{\mathscr{N}_{\|\cdot\|}(P)}{\mathscr{N}_{\|\cdot\|}(A)}\,.$$

Une nouvelle façon de perturber

En parallèle du système linéaire perturbé $(A+P)x = b$, nous pouvons également envisager le problème consistant à perturber le second membre b du système linéaire $Ax = b$ dont l'unique solution est toujours notée x_0. Ceci revient à étudier le système linéaire $Ax = b+q$ avec $q \in M_{n,1}(\mathbb{K})$ dont on suppose qu'il admet également une solution x_p. L'exemple de R.S. Wilson ci-dessous montre que ce nouveau problème souffre du même "effet papillon" que le précédent.

L'exemple classique

On pose

$$n = 4, \quad A = \begin{pmatrix} 10 & 7 & 8 & 7 \\ 7 & 5 & 6 & 5 \\ 8 & 6 & 10 & 9 \\ 7 & 5 & 9 & 10 \end{pmatrix} \quad \text{et} \quad b = \begin{pmatrix} 32 \\ 23 \\ 33 \\ 31 \end{pmatrix}.$$

Rappelons que

$$\{x \in M_{n,1}(\mathbb{K}) : Ax = b = (1,1,1,1)^T\},$$

Avec

$$q = \begin{pmatrix} 0,1 \\ -0,1 \\ 0,1 \\ -0,1 \end{pmatrix}.$$

on peut observer que

$$\left\{x \in M_{n,1}(\mathbb{K}) : Ax = b + q\right\} = \left\{(9,2; -12,6; 4,5; -1,1)^T\right\}.$$

Erreur relative et conditionnement II

Le résultat suivant offre la possibilité d'estimer l'erreur absolue $\|x_0 - x_p\|$ et l'erreur relative $\|x_0 - x_p\|/\|x_p\|$ en fonction du conditionnement de A.

Proposition

Soient $n \geq 1$ un entier, $A \in GL_n(\mathbb{K})$, $b \in M_{n,1}(\mathbb{K}) \setminus \{0\}$, $q \in M_{n,1}(\mathbb{K})$ et $x_0, x_p \in M_{n,1}(\mathbb{K})$ satisfaisant

$$Ax_0 = b \quad \text{et} \quad Ax_p = b + q.$$

Alors, $x_0 \neq 0$ et on a pour n'importe quelle norme $\|\cdot\|$ sur $M_{n,1}(\mathbb{K})$,

$$\frac{\|x_0 - x_p\|}{\|x_0\|} \leq \mathrm{cond}_{\|\cdot\|}(A) \frac{\|q\|}{\|b\|}. \tag{2}$$

Démonstration. Choisissons $\|\cdot\|$ une norme quelconque sur $M_{n,1}(\mathbb{K})$. Notons tout de suite que $x_0 \neq 0$ puisque $b \neq 0$. L'égalité $x_p - x_0 = A^{-1}q$ donne tout de suite

$$\|x_p - x_0\| \leq \mathscr{N}_{\|\cdot\|}(A^{-1})\,\|q\|\,.$$

De même, en utilisant $Ax_0 = b$, il vient

$$\|b\| \leq \mathscr{N}_{\|\cdot\|}(A)\,\|x_0\|\,,$$

ou encore

$$\frac{1}{\|x_0\|} \leq \frac{\mathscr{N}_{\|\cdot\|}(A)}{\|b\|}.$$

Il reste alors à écrire

$$\frac{\|x_0 - x_p\|}{\|x_0\|} \leq \mathscr{N}_{\|\cdot\|}(A^{-1})\,\|q\|\,\mathscr{N}_{\|\cdot\|}(A)\,\frac{1}{\|b\|}$$

pour terminer la démonstration.

Inverse et conditionnement

Le résultat ci-dessous examine le comportement de l'inverse d'une matrice dans le voisinage d'une matrice donnée.

Proposition

Soient $n \geq 1$ un entier, $\|\cdot\|$ une norme sur $M_{n,1}(\mathbb{K})$, $A \in GL_n(\mathbb{K})$ et $Q \in M_n(\mathbb{K})$. Si $A+Q \in GL_n(\mathbb{K})$, alors

$$\frac{\mathscr{N}_{\|\cdot\|}[(A+Q)^{-1} - A^{-1}]}{\mathscr{N}_{\|\cdot\|}[(A+Q)^{-1}]} \leq \mathrm{cond}_{\|\cdot\|}(A) \frac{\mathscr{N}_{\|\cdot\|}(Q)}{\mathscr{N}_{\|\cdot\|}(A)}.$$

Inverse et conditionnement

Le résultat ci-dessous examine le comportement de l'inverse d'une matrice dans le voisinage d'une matrice donnée.

Proposition

Soient $n \geq 1$ un entier, $\|\cdot\|$ une norme sur $M_{n,1}(\mathbb{K})$, $A \in GL_n(\mathbb{K})$ et $Q \in M_n(\mathbb{K})$. Si $A+Q \in GL_n(\mathbb{K})$, alors

$$\frac{\mathscr{N}_{\|\cdot\|}[(A+Q)^{-1}-A^{-1}]}{\mathscr{N}_{\|\cdot\|}[(A+Q)^{-1}]} \leq \mathrm{cond}_{\|\cdot\|}(A)\frac{\mathscr{N}_{\|\cdot\|}(Q)}{\mathscr{N}_{\|\cdot\|}(A)}.$$

Démonstration. Supposons que $A+Q \in GL_n(\mathbb{K})$. Commençons par écrire que

$$(A+Q)^{-1}(A+Q) = \mathrm{I}_n.$$

Cette égalité est évidemment équivalente à

$$(A+Q)^{-1}(\mathrm{I}_n + QA^{-1}) = A^{-1}.$$

Il découle alors de ceci

$$(A+Q)^{-1} - A^{-1} = -(A+Q)^{-1}QA^{-1}$$

et cette dernière relation entraîne l'égalité désirée.

Premières propriétés

Les premières propriétés du conditionnement sont données par la proposition suivante.

Proposition

Soient $n \geq 1$ un entier, $\|\cdot\|$ une norme sur $M_{n,1}(\mathbb{K})$ et $A \in GL_n(\mathbb{K})$. On a :

$$\mathrm{cond}_{\|\cdot\|}(A) \geq 1 \quad \text{et} \quad \mathrm{cond}_{\|\cdot\|}(A) = \mathrm{cond}_{\|\cdot\|}(A^{-1})$$

et

$$\mathrm{cond}_{\|\cdot\|}(\alpha A) = \mathrm{cond}_{\|\cdot\|}(A) \quad \text{pour tout } \alpha \in \mathbb{R} \setminus \{0\}.$$

Si $\|\cdot\|'$ désigne une autre norme sur $M_{n,1}(\mathbb{K})$, alors il existe des réels $\alpha, \beta \geq 0$ tels que

$$\alpha \mathrm{cond}_{\|\cdot\|}(B) \leq \mathrm{cond}_{\|\cdot\|'}(B) \leq \beta \mathrm{cond}_{\|\cdot\|}(B) \quad \text{pour tout } B \in GL_n(\mathbb{K}).$$

La démonstration

La première inégalité découle du caractère matriciel d'une norme subordonnée.

L'égalité $\mathrm{cond}_{\|.\|}(A) = \mathrm{cond}_{\|.\|}(A^{-1})$ est une conséquence directe de la définition du conditionnement.

La seconde égalité proposée découle de la propriété d'homogéneité d'une norme.

Enfin, l'encadrement proposé est une conséquence directe de l'équivalence des normes en dimension finie. ■

Conditionnement relatif à la norme 2

On dispose de résultats particuliers pour le conditionnement relatif à la norme 2.

Proposition

Soient $n \geq 1$ et $A \in GL_n(\mathbb{C})$. Ont lieu :

(a) Si $\mu_1(A)$ et $\mu_n(A)$ désignent respectivement la plus petite et la plus grande valeur singulière de A (i.e., les racines carrées des valeurs propres de la matrice hermitienne positive $A^\star A$), alors on a

$$\mathrm{cond}_{\|\cdot\|_2}(A) = \frac{\mu_n(A)}{\mu_1(A)}.$$

(b) Si A est une matrice normale (i.e., $AA^\star = A^\star A$), alors on a

$$\mathrm{cond}_{\|\cdot\|_2}(A) = \frac{\max\limits_{\lambda \in \mathrm{sp}(A)} |\lambda|}{\min\limits_{\lambda \in \mathrm{sp}(A)} |\lambda|}.$$

(c) Si A est unitaire ou orthogonale, alors $\mathrm{cond}_{\|\cdot\|_2}(A) = 1$.

(d) Pour tout $U \in M_n(\mathbb{C})$ unitaire, on a

$$\mathrm{cond}_{\|\cdot\|_2}(A) = \mathrm{cond}_{\|\cdot\|_2}(AU) = \mathrm{cond}_{\|\cdot\|_2}(UA) = \mathrm{cond}_{\|\cdot\|_2}(U^\star AU).$$

(*a*) On a

$$\mathrm{cond}_{\|\cdot\|_2}(A) = \mathscr{N}_{\|\cdot\|_2}(A)\mathscr{N}_{\|\cdot\|_2}(A^{-1}) = \sqrt{\rho(B)}\sqrt{\rho(B^{-1})},$$

où $B = A^\star A$ (qui est hermitienne positive). Par définition, le réel $\sqrt{\rho(B)}$ est la plus grande valeur singulière de A, notée $\mu_n(A)$. D'autre part, pour une matrice inversible $M \in M_n(\mathbb{C})$, il n'est pas difficile de voir que $\lambda \in \mathbb{C}$ avec $\lambda \neq 0$ est valeur propre de M si et seulement si $1/\lambda$ est valeur propre de M^{-1}. Ceci entraîne que $\sqrt{\rho(B^{-1})}$ n'est nulle autre que $1/\mu_1(A)$ où $\mu_1(A)$ est la plus petite valeur singulière de A. On aboutit ainsi à l'égalité

$$\mathrm{cond}_{\|\cdot\|_2}(A) = \frac{\mu_n(A)}{\mu_1(A)}.$$

La démonstration de (b) et (c)

(b)Puisque A est normale, on a

$$\mathscr{N}_{\|\cdot\|_2}(A) = \rho(A) = \max_{\lambda \in \mathrm{sp}(A)} |\lambda|. \tag{3}$$

Le caractère normal de la matrice inversible A entraîne tout de suite le fait que A^{-1} est également normale. Il vient alors

$$\mathscr{N}_{\|\cdot\|_2}(A^{-1}) = \rho(A^{-1}) = \mathscr{N}_{\|\cdot\|_2}(A^{-1}) = \rho(A^{-1}) = \frac{1}{\min\limits_{\lambda \in \mathrm{sp}(A)} |\lambda|}. \tag{4}$$

Il reste à combiner (3) et (4) pour aboutir au résultat souhaité.

(c) Le fait que A soit unitaire donne

$$\mathrm{cond}_{\|\cdot\|_2}(A) = \sqrt{\rho(A)}\,\sqrt{\rho(A^\star)} = \sqrt{\rho(A^\star A)} = \sqrt{\rho(\mathrm{I}_n)} = 1.$$

(d) C'est une conséquence directe de l'invariance par transformation unitaire satisfaite par $\|\cdot\|_2$.