

162. Systèmes d'équations linéaires. Opérations élémentaires, aspects algorithmiques et conséquences théoriques.

Introduction : Les systèmes linéaires sont des problèmes coûteux à résoudre et intervenant pourtant largement en analyse numérique. Il est donc nécessaire de disposer de méthodes de résolution efficaces.

Soit $\mathbb{K} = \mathbb{R}$ ou \mathbb{C} .

1 Généralités sur les systèmes linéaires

1.1 Définitions

Définition 1. On appelle système linéaire une équation de la forme $Ax = b$, d'inconnue $x \in \mathbb{K}^n$, où $A \in \mathcal{M}_{p,n}(\mathbb{K})$ et $b \in \mathbb{K}^p$. On appelle rang du système le rang de A . Le système est dit compatible s'il possède au moins une solution.

Exemple 2. Exemples de situation nécessitant la résolution d'un système linéaire :

- Calcul d'une base antédurale.
- Résolution d'une EDP par un schéma implicite aux différences finies.
- Méthode de Newton dans le cas vectoriel.

Proposition 3. Le système est compatible si et seulement si $b \in \text{Im}(A)$.

Définition 4. Soit $A \in GL_n(\mathbb{K})$. Si $\|\cdot\|$ norme subordonnée, on définit le conditionnement de A relativement à cette norme par $\text{cond}(A) = \|A\| \cdot \|A^{-1}\|$.

Proposition 5. Si on note $x, x + \delta x$ les solutions de $Ax = b, A(x + \delta x) = b + \delta b$, on a $\frac{|\delta x|}{|x|} \leq \text{cond}(A) \frac{|\delta b|}{|b|}$, et l'inégalité est optimale.

Proposition 6. Si on note $x, x + \delta x$ les solutions de $Ax = b, (A + \delta a)(x + \delta x) = b$, on a $\frac{|\delta x|}{|x + \delta x|} \leq \text{cond}(A) \frac{\|\delta A\|}{\|A\|}$, et l'inégalité est optimale.

1.2 Systèmes de Cramer

Définition 7. Le système est dit de Cramer si $n = p$ et $A \in GL_n(\mathbb{K})$.

Théorème 8. Le système possède alors une unique solution $x \in \mathbb{K}^n$, donnée par $x_i = \frac{\det(A_1, \dots, A_{i-1}, b, A_{i+1}, \dots, A_n)}{\det(A)}$.

1.3 Cas général

Définition 9. Pour $s \leq \min(n, p)$, on appelle sous-matrice diagonale de taille s de A la matrice $A_s = (a_{i,j})_{1 \leq i, j \leq s}$.

Théorème 10 (Rouché-Fontené). On suppose que le système est de rang r , et quitte à renuméroter, que A_r est inversible. Le système est compatible, si et seulement si

$$\text{pour tout } t \in \llbracket r + 1, p \rrbracket, \begin{vmatrix} a_{1,1} & \cdots & a_{1,r} & b_1 \\ \vdots & & \vdots & \vdots \\ a_{r,1} & \cdots & a_{r,r} & b_r \\ a_{s,1} & \cdots & a_{s,r} & b_s \end{vmatrix} = 0. \text{ Le système est alors équivalent}$$

au système précédent dont on a ôté les $n - r$ dernières lignes.

Définition 11. Si le système est incompatible, le problème aux moindres carrés revient à chercher $\bar{x} = \arg \min_{x \in \mathbb{R}^n} \|Ax - b\|$.

Proposition 12. x est solution du problème aux moindres carrés, si et seulement si il vérifie $A^*Ax = b$.

Théorème 13. Le problème aux moindres carrés possède toujours une solution, et celle-ci est unique ssi $\ker(A) = \{0\}$.

2 Opérations élémentaires, résolution directe

2.1 Opérations élémentaires sur les lignes et colonnes

Définition 14. Pour $1 \leq i \neq j \leq n, \lambda \neq 0$, on pose $T_{i,j}(\lambda) = I_n + \lambda E_{i,j}$. Pour $1 \leq i \leq n, \alpha \notin \{0, 1\}$, on pose $D_i(\alpha) = I_n + (\alpha - 1)E_{i,i}$. Pour $\sigma \in \mathcal{S}_n$, on note $P_\sigma = (\delta_{i,\sigma(j)})_{i,j}$. De telles matrices sont appelées respectivement transvections, dilatations et matrices de permutation.

Proposition 15. Ce sont des matrices inversibles, et multiplier à gauche ou à droite par une telle matrice revient à effectuer une opération élémentaire (voire annexe).

Théorème 16. Les transvections engendrent $SL_n(\mathbb{K})$, les transvections et les dilatations engendrent $GL_n(\mathbb{K})$.

2.2 Systèmes échelonnés

Définition 17. On appelle pivot d'une ligne non nulle le coefficient non nul le plus à gauche. Une matrice est dite échelonnée en lignes si lorsqu'une ligne est nulle, toutes les suivantes sont nulles, et si le pivot de chaque ligne est strictement plus à droite que ceux des lignes précédentes. Elle est dite échelonnée en colonnes si sa transposée est échelonnée en lignes.

Proposition 18. Un système échelonné se résout par un algorithme de remontée en $O(np)$ opérations.

2.3 Méthodes d'élimination et de décomposition

On suppose le système de Cramer.

Théorème 19. Il existe $M \in GL_n(\mathbb{K})$ telle que MA soit triangulaire supérieure.

On se ramène ainsi à un système échelonné avec l'algorithme du pivot de Gauss :

- On permute lignes et colonnes pour que le pivot $a_{1,1}$ soit non nul.
- On échelonne : $\forall i \geq 2, L_i \leftarrow L_i - \frac{a_{i,1}}{a_{1,1}}L_1$
- On recommence à partir de la deuxième ligne.

Théorème 20 (Bruhat). On note T_s le sous-groupe de $GL_n(\mathbb{K})$ formé des matrices triangulaires supérieures. Alors $G = \bigsqcup_{\sigma \in \mathcal{S}_n} T_s P_\sigma T_s$.

Application 21. L'ensemble \mathcal{D} des drapeaux de \mathbb{K}^n est en bijection avec $GL_n(\mathbb{K})/T_s$, et l'ensemble des orbites pour l'action de translation de $GL_n(\mathbb{K})$ sur $GL_n(\mathbb{K})/T_s \times GL_n(\mathbb{K})/T_s$ est en bijection avec \mathcal{S}_n .

Théorème 22. Si toutes les sous-matrices diagonales de A sont inversibles, il existe un unique couple (L, U) où L triangulaire inférieure avec des 1 sur la diagonale, et U triangulaire supérieure inversible, tel que $A = LU$.

Remarque 23. On résout alors le système en résolvant $Ly = b$ puis $Ux = y$ par des algorithmes de descente et remontée. Lorsque A est seulement supposée inversible, on a une décomposition PLU avec P matrice de permutation.

Théorème 24 (Cholesky). Si $A \in S_n^{++}(\mathbb{R})$, il existe une unique matrice B triangulaire inférieure à coefficients diagonaux positifs tels que $A = BB^*$.

Le calcul de ces décompositions se fait en respectivement $\simeq \frac{n^3}{3}$ et $\simeq \frac{n^3}{6}$ opérations. Ils permettent notamment le calcul du déterminant et de l'inverse d'une matrice.

Théorème 25. Il existe un unique couple (Q, R) avec Q orthogonale, R triangulaire supérieure inversible, tel que $A = QR$.

3 Méthodes itératives de résolution approchée

3.1 Décomposition régulière

Définition 26. On écrit $A = M - N$ avec M inversible. On dit que la méthode itérative associée à cette décomposition est convergente si, pour tout $b \in \mathbb{C}^n$, pour tout $x_0 \in \mathbb{C}^n$, la suite définie par $x_{k+1} = M^{-1}Nx_k + M^{-1}b$ converge.

Lemme 27. Si la méthode converge, alors sa limite x est l'unique solution du système linéaire $Ax = b$.

Proposition 28. La méthode converge, si et seulement si, $\rho(M^{-1}N) < 1$.

Théorème 29. Si $A \in H_n^{++}(\mathbb{C})$, alors $M^* + N$ est hermitienne. Si elle est de plus définie positive, alors la méthode itérative converge.

3.2 Méthodes de Jacobi, Gauss-Seidel et de gradient

Application 30. On note D la partie diagonale de A , E sa partie triangulaire inférieure, F sa partie triangulaire supérieure.

La méthode de Jacobi est associée à la décomposition $M = D, N = E + F$.

La méthode de Gauss-Seidel est associée à la décomposition $M = D - E, N = F$. Ainsi, si A est hermitienne définie positive, la méthode de Gauss-Seidel converge. Si de plus, $2D - A$ est définie positive, la méthode de Jacobi converge.

Exemple 31. La résolution de problèmes linéaires associés à la matrice du Laplacien $1D$, donnée par $A_n = n^2 \text{tridiag}(-1, 2, -1) \in \mathcal{M}_{n-1}(\mathbb{R})$, peuvent se faire par Jacobi et Gauss-Seidel.

Lemme 32. Soient $A \in S_n(\mathbb{R}), b \in \mathbb{R}^n$. On définit la fonction $f : \mathbb{R}^n \rightarrow \mathbb{R}$ par $f(x) = \frac{1}{2} \langle Ax, x \rangle + \langle b, x \rangle$. Alors :

- Si $A \in S_n^{++}(\mathbb{R})$, f admet un unique minimum en x_0 qui est la solution de $Ax = b$.
- Si $A \in S_n^+(\mathbb{R})$ et $b \in \text{Im}(A)$, les points où f atteint son minimum sont exactement les solutions de $Ax = b$.
- Sinon, f n'est pas minorée.

Théorème 33. On suppose $A \in S_n^{++}(\mathbb{R})$. Alors x_0 réalise le minimum de f ssi $\nabla f(x_0) = 0$. De plus, si $x \in \mathbb{R}^n$ tel que $\nabla f(x) \neq 0$, alors pour tout $\alpha \in]0, \frac{2}{\rho(A)}[$, $f(x - \alpha \nabla f(x)) < f(x)$.

Corollaire 34. Ainsi, pour $x_0 \in \mathbb{R}^n$, la suite définie par $x_{k+1} = x_k - \alpha \nabla f(x_k) = x_k + \alpha(b - Ax_k)$ converge vers la solution de $Ax = b$.

Remarque 35. Cette suite est celle de la décomposition régulière $M = \alpha^{-1}I_n, N = \alpha^{-1}I_n - A$.

Proposition 36 (Pas optimal). La suite définie par $x_{k+1} = x_k + \alpha_k(b - Ax_k)$, où α_k minimise la fonction $g_k(\alpha) = f(x_k - \alpha \nabla f(x_k))$, converge vers la solution de $Ax = b$.

Annexe

	Produit à gauche	Produit à droite
$T_{i,j}(\lambda)$	$L_i \leftarrow L_i + \lambda L_j$	$C_j \leftarrow C_j + \lambda C_i$
$D_i(\alpha)$	$L_i \leftarrow \alpha L_i$	$C_i \leftarrow \alpha C_i$
P_σ	Permutation des lignes selon σ^{-1}	Permutation des colonnes selon σ

Développements

- Générateurs du groupe linéaire.
- Méthodes itératives de décomposition régulière.
- Méthode du gradient à pas optimal.

Références

- [1] G. Allaire, S.-M. Kaber, ALGÈBRE LINÉAIRE NUMÉRIQUE, Ellipses.
- [2] P. Caldero, J. Germoni, HISTOIRES HÉDONISTES DE GROUPES ET DE GÉOMÉTRIES, Calvage & Mounet.
- [3] J. Grifone, ALGÈBRE LINÉAIRE, Cépaduès.