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

Cadre : Soient $\mathbb{K}$ un corps et $n, p \in \mathbb{N}^\star$. Pour tout $i \in [\![1, n]\!]$ et tout $j \in [\![1, p]\!]$, soient $a_{i,j} \in \mathbb{K}$ et $b_i \in \mathbb{K}$. On considère le système de n équations à p inconnues à coefficients dans $\mathbb{K}$ défini par :

$$\left\{ \begin{array}{ccccccc} a_{1,1}x_1 & + & \dots & + & a_{1,p}x_p & = & b_1 \\ \vdots & & & & \vdots & & \vdots \\ a_{n,1}x_1 & + & \dots & + & a_{n,p}x_p & = & b_n \end{array} \right. \tag{S}$$

I Existence et unicité des solutions

1) Définitions et premières propriétés

Définition 1. On appelle solution de (S) tout vecteur $x \in \mathbb{K}^n$ dont les coordonnées vérifient chacune des équations de (S).

Définition 2. Le système est dit compatible si (S) admet une solution.

Exemple 3. $\left\{ \begin{array}{l} x+y=1 \\ x=0 \end{array} \right.$ *est compatible, mais pas* $\left\{ \begin{array}{l} x+y=1 \\ x+y=0 \end{array} \right.$.

Définition 4. Dans (S), on pose $A = (a_{i,j})_{1 \leqslant i \leqslant n, 1 \leqslant j \leqslant p} \in \mathcal{M}_{n,p}(\mathbb{K})$ et $b = (b_i)_{1 \leqslant i \leqslant n} \in \mathbb{K}^n$. Alors $x \in \mathbb{K}^n$ est solution de (S) si, et seulement si, $Ax = b$ dans $\mathbb{K}^n$.

Définition 5. On appelle rang du système (S) le rang de A.

Proposition 6. *Notons $A_1, \dots, A_n$ les colonnes de A. Alors (S) est compatible si, et seulement si, $b \in \operatorname{Vect}(A_1, \dots, A_n)$.*

Exemple 7. $\left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right)\left(\begin{smallmatrix} x \\ y \end{smallmatrix}\right) = \left(\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}\right)$ *est compatible, mais pas* $\left(\begin{smallmatrix} 1 & 1 \\ 1 & 1 \end{smallmatrix}\right)\left(\begin{smallmatrix} x \\ y \end{smallmatrix}\right) = \left(\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}\right)$.

Définition 8. Le système est dit compatible si $b = 0$.

Proposition 9. *Un système homogène est toujours compatible.*

2) Système de Cramer

Définition 10. On dit que (S) est de Cramer si A est inversible.

Proposition 11. *Un système de Cramer a une unique solution $x = A^{-1}b$.*

Théorème 12. *Supposons que (S) est de Cramer, et notons $A_1, \dots, A_n$ les colonnes de A. Alors :*

$$x = \frac{\det(A_1, \dots, A_{i-1}, b, A_{i+1}, \dots, A_n)}{\det A}$$

Remarque 13. *Numériquement, cette méthode utilise $(n+2)!$ opérations. C'est impossible à mettre en œuvre pour de grandes valeurs de n.*

Exemple 14. *Pour* $\left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right)\left(\begin{smallmatrix} x \\ y \end{smallmatrix}\right) = \left(\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}\right)$ *:* $x = \left|\begin{smallmatrix} 2 & 1 \\ 0 & 1 \end{smallmatrix}\right| = 2$ *et* $y = \left|\begin{smallmatrix} 1 & 2 \\ 0 & 0 \end{smallmatrix}\right| = 0$.

3) Cas général

Soient $A \in \mathcal{M}_{n,p}(\mathbb{K})$ de rang r, et $b \in \mathbb{K}^n$. On peut supposer $(a_{i,j})_{1 \leqslant i,j \leqslant r}$ inversible de déterminant δ.

Théorème 15. *(i) Le système est compatible si, et seulement si :*

$$\forall s \in [\![r+1, p]\!],\ \Delta_s = \begin{vmatrix} a_{1,1} & \cdots & a_{1,r} & b_1 \\ \vdots & & \vdots & \vdots \\ a_{r,1} & \cdots & a_{r,r} & b_r \\ a_{s,1} & \cdots & a_{s,r} & b_s \end{vmatrix} = 0$$

(ii) Si le premier point est réalisé, le système est équivalent à :

$$\left\{ \begin{array}{l} a_{1,1}x_1 + \dots + a_{1,r}x_r = b_1 - a_{1,r+1} - \dots - a_{1,n}x_n \\ \vdots \\ a_{r,1}x_1 + \dots + a_{r,r}x_r = b_r - a_{r,r+1} - \dots - a_{r,n}x_n \end{array} \right.$$

Il admet une infinité de solutions dépendant de $(n-r)$ paramètres (on donne à $x_{r+1}, \dots, x_n$ des valeurs arbitraires).

Exemple 16. *Soient* $A = \left(\begin{smallmatrix} 2 & 1 & -1 \\ 0 & 1 & 3 \\ 2 & k & 2 \end{smallmatrix}\right)$ *et* $b = \left(\begin{smallmatrix} \alpha \\ \beta \\ \gamma \end{smallmatrix}\right)$.

(i) Si $k \neq 2$, le système $Ax = b$ a une unique solution.

(ii) Si $k = 2$, et $\gamma \neq \alpha + \beta$, le système $Ax = b$ n'a pas de solution.

(iii) Si $k = 2$, et $\gamma = \alpha + \beta$, le système $Ax = b$ a une infinité de solutions.

Proposition 17. *Soit E un espace vectoriel de dimension finie n.*

(i) Soient $\varphi_1, \dots, \varphi_p \in E^\star$ telles que $\operatorname{rg}(\varphi_1, \dots, \varphi_p) = r$. Alors le sous-espace vectoriel $F = \{x \in E \,|\, \forall i, \varphi(x) = 0\}$ est de dimension $n - r$.

(ii) Réciproquement, si F est un sous-espace vectoriel de E de dimension q, il existe $n - q$ formes linéaires indépendantes $\varphi_1, \dots, \varphi_{n-q}$ telles que $F = \{x \in E \,|\, \forall i, \varphi(x) = 0\}$.

Remarque 18. *Les solutions de $Ax = 0$ forment un sous-espace vectoriel de dimension $n - \operatorname{rg}(A) = \dim \operatorname{Ker}(A)$. Les solutions de $Ax = b$ forment un sous-espace affine.*

Exemple 19. *Un hyperplan de E peut-être défini comme le noyau d'une forme linéaire sur E.*

II Méthode du pivot de Gauss

Définition 20. $\mathcal{GL}_n(\mathbb{K})$ agit sur $\mathcal{M}_n(\mathbb{K})$ par multiplication à gauche.

Définition 21. On appelle matrice de transvection toute matrice qui est de la forme $T_{i,j}(\lambda) = I_n + \lambda E_{i,j}$, où $\lambda \in \mathbb{K}$ et $i,j \in [\![1,n]\!]$ avec $i \neq j$.

Définition 22. On appelle matrice de dilatation toute matrice qui est de la forme $D_i(\lambda) = I_n + (\lambda - 1)E_{i,i}$, où $\lambda \in \mathbb{K}$ et $i \in [\![1,n]\!]$.

Définition 23. On appelle matrice de permutation toute matrice qui est de la forme $P_{i,j} = I_n - E_{i,i} - E_{j,j} + E_{i,j} + E_{j,i}$, où $i,j \in [\![1,n]\!]$ avec $i \neq j$.

Remarque 24. *Ces matrices représentent également des transformations élémentaires dans l'algorithme du pivot de Gauss. Soit $A \in \mathcal{M}_{m,n}(\mathbb{K})$:*

Opération	$T_{i,j}(\lambda)A$	$D_i(\lambda)A$	$P_{i,j}A$
Résultat	$L_i \leftarrow L_i + \lambda L_j$	$L_i \leftarrow \lambda L_i$	$L_i \leftrightarrow L_j$

On a les opérations analogues sur les colonnes en multipliant à droite.

Exemple 25. $T_{3,2}(1) \times T_{3,1}(-3) \times T_{2,1}(-1) \times \begin{pmatrix} 1 & 2 & 2 \\ 1 & 3 & -2 \\ 3 & 5 & 8 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 2 \\ 0 & 1 & -4 \\ 0 & 0 & -2 \end{pmatrix}$

Définition 26. On appelle :

(i) pivot d'une ligne son coefficient non nul le plus à gauche.

(ii) matrice échelonnée en lignes une matrice telle que dès qu'une ligne est nulle, les suivantes le sont, et pour les lignes non nulles le pivot d'une ligne est strictement à droite du pivot de la ligne précédente. Une matrice échelonnée est dite réduite si ses pivots valent 1.

On a la définition similaire de matrice échelonnée (réduite) en colonnes.

Proposition 27. *Les orbites sont en bijections avec les sous-espaces vectoriels de $\mathbb{K}^n$: $A \sim B \Leftrightarrow \operatorname{Ker} A = \operatorname{Ker} B$.*

Proposition 28. *Toute matrice est dans l'orbite d'une unique matrice échelonnée.*

Lemme 29. *On suppose E de dimension $n \geqslant 2$. Soient $x, y \in E \setminus \{0\}$. Il existe une transvection u ou un produit de deux transvections uv, tel que $u(x) = y$ ou $uv(x) = y$.*

Théorème 30. *Les transvections engendrent $\mathcal{SL}(E)$.*

Théorème 31. *Les transvections et les dilatations engendrent $\mathcal{GL}(E)$.*

Application 32. *L'algorithme du pivot de Gauss permet de se ramener à la matrice réduite associée à une matrice via des opérations élémentaires sur les lignes.*

III Autres algorithmes de résolution

1) Décomposition LU

Théorème 33 (Décomposition LU)**.** *Soit $A = (a_{i,j})_{1 \leqslant i,j \leqslant n} \in \mathcal{M}_n(\mathbb{K})$ telle que les n sous-matrices diagonales $\Delta_k = (a_{i,j})_{1 \leqslant i,j \leqslant k} \in \mathcal{M}_k(\mathbb{K})$ soient inversibles. Alors il existe une matrice triangulaire inférieure L dont les termes diagonaux valent 1 et une matrice triangulaire supérieure U telles que $A = LU$. Cette factorisation est unique.*

Exemple 34. $\begin{pmatrix} 5 & 2 & 1 \\ 5 & -6 & 2 \\ -4 & 2 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ -\frac{4}{5} & -\frac{9}{20} & 1 \end{pmatrix} \begin{pmatrix} 5 & 2 & 1 \\ 0 & -8 & 1 \\ 0 & 0 & \frac{9}{4} \end{pmatrix}$

Méthode 35. *On résout alors $Ly = b$ puis $Ux = y$ par remontée.*

Théorème 36 (Décomposition de Choleski)**.** *Soit $A \in \mathcal{GL}_n(\mathbb{K})$ symétrique définie positive. Alors il existe une matrice C réelle triangulaire inférieure telle que $A = C\,{}^{t}C$. Cette factorisation est unique si on impose aux termes diagonaux d'être strictement positifs.*

Exemple 37. $\begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 5 & 5 & 5 \\ 1 & 5 & 14 & 14 \\ 1 & 5 & 14 & 15 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 1 & 2 & 0 & 0 \\ 1 & 2 & 3 & 0 \\ 1 & 2 & 3 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 2 & 2 & 2 \\ 0 & 0 & 3 & 3 \\ 0 & 0 & 0 & 1 \end{pmatrix}$

Méthode 38. *On résout alors ${}^{t}Sy = b$ puis $Sx = y$ par remontée.*

2) Généralités et notions de convergence

Méthode 39. *On cherche à approximer la solution x du système $Ax = b$, où $A \in \mathcal{GL}_n(\mathbb{K})$ et $b \in \mathbb{K}^n$. On pose pour cela $A = M - N$, où $M \in \mathcal{GL}_n(\mathbb{K})$ est facile à inverser (diagonale, triangulaire, orthogonale...). On obtient la méthode itérative :*

$$x^{(k+1)} = M^{-1}Nx^{(k)} + M^{-1}b = F(x^{(k)}) \tag{M}$$

Si la suite $(x^{(k)})_{k \in \mathbb{N}}$ converge vers x^{∞}, alors x^{∞} est solution de $Ax = b$. De plus, x^{∞} est un point fixe de F.

Définition 40. Soients $A = M - N$ et $\|.\|$ une norme sur $\mathbb{K}^n$. La méthode itérative (M) est dite convergente lorsque :

$$\forall\, b \in \mathbb{K}^n,\ \forall\, x^{(0)} \in \mathbb{K}^n,\ \lim_{k \to \infty} \left\| x^{(k)} - x \right\| = 0$$

Remarque 41. *En posant $e^{(k)} = x^{(k)} - x$ et $B = M^{-1}N$, on a $e^{(k+1)} = Be^{(k)}$, et la méthode itérative (M) converge ainsi lorsque, par exemple, $\rho(B) < 1$.*

3) Méthodes itératives

Définition 42. Soit $A \in \mathcal{GL}_n(\mathbb{K})$. On considère les sous-matrices D, E, F de A définies comme suit :

$$\begin{cases} (D)_{i,j} = a_{i,j} \text{ si } i = j, \ 0 \text{ sinon} \\ (-E)_{i,j} = a_{i,j} \text{ si } i > j, \ 0 \text{ sinon} \\ (-F)_{i,j} = a_{i,j} \text{ si } i < j, \ 0 \text{ sinon} \end{cases} \qquad A = \begin{pmatrix} \boxed{D} & -F \\ -E & \end{pmatrix}$$

On peut alors considérer plusieurs méthodes itératives :

(i) Jacobi : $M = D$, $N = E + F$

(ii) Gauss-Seidel : $M = D - E$, $N = F$

(iii) Relaxation : $M = \frac{1}{\omega} D - E$, $N = \frac{1-\omega}{\omega} D + F$

Remarque 43. *Avec $B = M^{-1}N$, la méthode itérative (M) devient :*

$$x^{(k+1)} = Bx^{(k)} + b \tag{M'}$$

Théorème 44. *Si la matrice A est à diagonale strictement dominante, alors la méthode de Jacobi converge.*

Exemple 45. *Si $A = \left(\begin{smallmatrix} 1 & a & a \\ a & 1 & a \\ a & a & 1 \end{smallmatrix}\right)$, la méthode de Jacobi converge si $|a| < \frac{1}{2}$.*

Théorème 46. *Si A est hermitienne définie positive, alors la méthode de relaxation converge pour $\omega \in]0, 2[$.*

Corollaire 47. *Si A est hermitienne définie positive, alors la méthode de Gauss-Seidel converge.*

4) Méthodes de gradient

Caractérisation de l'α-convexité

Définition 48. Pour $\alpha > 0$, on dit que la fonction $f : C \to \mathbb{R}$ est α-convexe si pour tous $a, b \in C$ distincts et tout $\lambda \in]0, 1[$, on a :

$$f((1-\lambda)a + \lambda b) \leqslant (1-\lambda) f(a) + \lambda f(b) - \frac{\alpha}{2} \|a - b\|^2 \lambda (1 - \lambda)$$

Théorème 49. *Soit $J : C \to \mathbb{R}$ différentiable. Il y a équivalence entre :*

(i) J est α-convexe sur C.

(ii) $\forall x, y \in C$, $\langle \nabla J(x) - \nabla J(y), x - y \rangle \geqslant \alpha \|x - y\|^2$.

(iii) $\forall x, y \in C$, $J(x) \geqslant J(y) + \langle \nabla J(y), x - y \rangle + \frac{\alpha}{2} \|x - y\|^2$.

Si J est deux fois différentiable, on a aussi : $\left\langle d^2 J(x) \cdot y, y \right\rangle \geqslant \alpha \|y\|^2$.

Méthode de gradient

Soit $J : \mathbb{R}^n \to \mathbb{R}$. On suppose J différentiable. On cherche, s'il existe, un élément $u \in \mathbb{R}^n$ tel que :

$$J(u) = \inf_{v \in \mathbb{R}^n} J(v)$$

Pour cela, on utilise les méthodes de gradient. On considère la suite :

$$u_0 \in \mathbb{R}^n \quad \text{et} \quad \forall k \in \mathbb{N}, \ u^{k+1} = u^k - \rho^k \nabla J(u^k)$$

Il existe plusieurs possibilités pour choisir les ρ^k, par exemple :

(i) Gradient à pas fixe : $\rho^k = \rho$ une constante positive fixée.

(ii) Gradient à pas optimal : ρ^k minimise $\rho \mapsto J\left(u^k - \rho \nabla J(u^k)\right)$.

Théorème 50. *Si J est α-convexe et différentiable, et que ∇J est L-lipschitzienne, alors la méthode de gradient à pas optimal converge vers l'unique minimum de J.*

Application 51. *Soient $A \in \mathcal{S}_n^{++}(\mathbb{R})$, $b \in \mathbb{R}^n$ et $c \in \mathbb{R}$. On considère la fonctionnelle quadratique $J : \mathbb{R}^n \to \mathbb{R}$ définie par :*

$$J(X) = \langle AX, X \rangle - \langle b, X \rangle + c$$

Cette fonctionnelle satisfait les conditions du théorème précédent. De plus, son minimum est atteint en $X_0 \in \mathbb{R}^n$ qui vérifie $\nabla J(X_0) = AX - b = 0$. On a donc une méthode itérative pour approcher la solution de $AX = b$.

Développements

— Générateurs de $\mathcal{GL}(E)$ et de $\mathcal{SL}(E)$ (30,31) [Per96]

— Algorithme de gradient à pas optimal (50) [Cia88]

Références

[Gou94] X. Gourdon. *Les Maths en Tête : Algèbre.* Ellipses, 2e édition

[Per96] D. Perrin. *Cours d'Algèbre.* Ellipses

[Gri11] J. Grifone. *Algèbre Linéaire.* Cépaduès, 4e édition

[Cia88] P. Ciarlet. *Introduction à l'analyse numérique et à l'optimisation.* Masson