162 SYSTÈMES D'ÉQUATIONS LINÉAIRES ; OPÉRATIONS ÉLÉMENTAIRES, ASPECTS ALGORITHMIQUES ET CONSÉQUENCES THÉORIQUES.

Soit $\mathbb{K}$ un corps et n, p des entiers non nuls.
On considère le système de n équations à p inconnues à coefficients dans $\mathbb{K}$ défini par :

$$(\mathscr{S}_B) \quad \left\{ \begin{array}{c} a_{1\,1}\,x_1 + \cdots + a_{1\,p}\,x_p = b_1 \\ \vdots \\ a_{n\,1}\,x_1 + \cdots + a_{n\,p}\,x_p = b_n \end{array} \right.$$

où les $(a_{i\,j})_{1 \leq i \leq n,\, 1 \leq j \leq p}$ et les $(b_i)_{1 \leq i \leq n}$ sont des éléments de $\mathbb{K}$.

I. Généralités [CG13, ChIV, p127]

I. A. Conditions d'existence et d'unicité de solutions

Définition 1. On appelle solution de $(\mathscr{S}_B)$ tout vecteur $x = (x_1, \ldots, x_p) \in \mathbb{K}^p$ satisfaisant les n équations. On dira que $(\mathscr{S}_B)$ est compatible s'il admet au moins une solution. On notera alors S_B l'ensemble des solutions.
Lorsque $b_1 = \cdots = b_n = 0$, on dit que $(\mathscr{S}_B)$ est homogène.

Exemple 2. Le système $\left\{ \begin{array}{l} x + y = 2 \\ x - y = 0 \\ 2x + y = 0 \end{array} \right.$ n'est pas compatible.

En notant $A = (a_{i\,j})_{1 \leq i \leq n,\, 1 \leq j \leq p} \in \mathcal{M}_{n\,p}(\mathbb{K})$, $B = (b_1 \ldots b_n)^\top \in \mathcal{M}_{n\,1}(\mathbb{K})$ et $X = (x_1, \ldots, x_p)^\top \in \mathcal{M}_{p\,1}(\mathbb{K})$, le système $(\mathscr{S}_B)$ se réécrit $AX = B$.

Théorème 3. *Le système est compatible si et seulement si $B \in \mathrm{Im}(A)$. Dans ce cas, S_B est un espace affine dirigé par $\ker(A)$ et de dimension $p - \mathrm{rg}(A)$.*

Définition 4. [Système de Cramer]
Lorsque $n = p$ et $A \in \mathcal{GL}_n(\mathbb{K})$, on dit que $(\mathscr{S}_B)$ est un système de Cramer.

Théorème 5. [Système de Cramer]
Si $(\mathscr{S}_B)$ est un système de Cramer, il admet une unique solution, égale à $X = A^{-1}B$. De plus, on a les formules de Cramer :

$$\forall i \in [\![1, p]\!], x_i = \frac{\det(C_1, \ldots, C_{i-1}, B, C_{i+1}, \ldots, C_n)}{\det(A)}$$

où les $(C_j)_{1 \leq j \leq n}$ sont les colonnes de A.

I. B. Cas de matrices triangulaires

Application 6. [Méthode des remontées]
Si $A \in \mathcal{GL}_n(\mathbb{K}) \cap \mathcal{TS}_n(\mathbb{K})$, le système se résout en résolvant chacune des équations de la dernière à la première : on calcule x_n solution de $x_n = \frac{1}{a_{n\,n}} b_n$, puis $x_{n-1} = \frac{1}{a_{n-1\,n-1}}(b_{n-1} - a_{n-1\,n} x_n)$, …, jusqu'à $x_1 = \frac{1}{a_{1\,1}}(b_1 - \sum_{j=2}^n a_{1\,j} x_j)$.

Remarque 7. Lorsque A n'est plus inversible, on peut toujours effectuer la méthode des remontées : les étapes i telles que $a_{i\,i} = 0$ constituent alors des équations de compatibilité.

Exemple 8. Résolution de $\left\{ \begin{array}{l} 2x - z = 0 \\ my + 3z = 12 \\ 2z = 8 \end{array} \right.$ où $m \in \mathbb{R}$ est un paramètre.

II. Résolution pratique d'un système linéaire [CG13, ChI, p2]

II. A. Algorithme du pivot de Gauss

Proposition 9. *L'ensemble des solutions de $(\mathscr{S}_B)$ ne change pas si l'on effectue les opérations élémentaires suivantes :*
- *si l'on ajoute à une équation une combinaison linéaire des autres,*
- *si l'on multiplie une équation par un scalaire non nul,*
- *si l'on change l'ordre des équations.*

Définition 10. [Matrices de dilatation, transvection, permutation élémentaire]
Une matrice de dilatation est une matrice $D_{i,\alpha} \in \mathcal{GL}_n(\mathbb{K})$ pour un $\alpha \in \mathbb{K}^*$. Une matrice de transvection est une matrice $T_{i,j,\beta} = I_n + \beta E_{i,j} \in \mathcal{GL}_n(\mathbb{K})$ pour un $\beta \in \mathbb{K}^*$ et $i \neq j$. Une matrice de permutation élémentaire est une matrice $P_{i,j} \in \mathcal{GL}_n(\mathbb{K})$ pour un $i < j$.

$$D_{i,\alpha} = \begin{pmatrix} I_{i-1} & & \\ & \alpha & \\ & & I_{n-i} \end{pmatrix} \qquad P_{i,j} = \begin{pmatrix} I_{i-1} & & & & \\ & 0 & & 1 & \\ & & I_{j-i-1} & & \\ & 1 & & 0 & \\ & & & & I_{n-j} \end{pmatrix}$$

Proposition 11. *Soit $M \in \mathcal{M}_{n,p}(\mathbb{K})$. Les opérations élémentaires sur les lignes (resp. colonnes) de M sont obtenues par multiplication à gauche (resp. à droite) par les matrices précédemment introduites :*

Matrice	$D_{i,\alpha}M$	$T_{i,j\beta}M$	$P_{i,j}M$	$MD_{i,\alpha}$	$\dots$
Opération	$L_i \leftarrow \alpha L_i$	$L_i \leftarrow L_i + \beta L_j$	$L_i \leftrightarrow \alpha L_j$	$C_i \leftarrow \alpha C_i$	$\dots$

Définition 12. On appelle :
- pivot d'une ligne son coefficient non nul le plus à gauche.
- matrice échelonnée en lignes une matrice telle que dès qu'une ligne est nulle, les suivantes sont nulles, et pour les lignes non nulles le pivot d'une ligne est strictement à droite du pivot de la ligne précédente. On dit qu'une matrice échelonnée est réduite si ses pivots valent 1.

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

Exemple 13. $\begin{pmatrix} 1 & 2 & 5 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$ est réduite en lignes, $\begin{pmatrix} 2 & 0 & 0 \\ 1 & 3 & 0 \\ 2 & 4 & 0 \end{pmatrix}$ est échelonnée en colonnes.

Définition 14. [Action par translation] On définit les actions par translation/multiplication à gauche et à droite :

$$\begin{array}{ccc} \mathcal{GL}_n(\mathbb{K}) \times \mathcal{M}_{n,p}(\mathbb{K}) & \longrightarrow & \mathcal{M}_{n,p}(\mathbb{K}) \\ (P, M) & \longmapsto & PM \end{array} \quad \text{et} \quad \begin{array}{ccc} \mathcal{GL}_p(\mathbb{K}) \times \mathcal{M}_{n,p}(\mathbb{K}) & \longrightarrow & \mathcal{M}_{n,p}(\mathbb{K}) \\ (P, M) & \longmapsto & MP^{-1} \end{array}$$

Théorème 15. [Orbites de l'action par multiplication à gauche] *Pour l'action par multiplication à gauche :*
- *deux matrices A et A' sont dans la même orbite si et seulement si $\ker(A) = \ker(A')$,*
- *toute matrice est dans l'orbite d'une unique matrice réduite en lignes.*

Théorème 16. [Orbites de l'action par multiplication à droite] *Pour l'action par multiplication à droite :*
- *deux matrices A et A' sont dans la même orbite si et seulement si $\mathrm{Im}(A) = \mathrm{Im}(A')$,*
- *toute matrice est dans l'orbite d'une unique matrice réduite en colonnes.*

Application 17. [Algorithme du pivot de Gauss] 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.

Exemple 18. $\begin{pmatrix} 2 & 4 & 1 \\ 1 & 3 & 0 \\ 0 & 5 & 2 \end{pmatrix}$ a pour matrice réduite associée $\begin{pmatrix} 1 & 2 & 1/2 \\ 0 & 1 & -1/2 \\ 0 & 0 & 1 \end{pmatrix}$.

Remarque 19.
- Il y a $\mathrm{rg}(A)$ pivots. Les $n - r$ dernières lignes de la forme réduite donnent alors des équations de compatibilité du système sur le second membre.
- Lorsque $n = p$ et A est inversible, la forme réduite est triangulaire supérieure avec coefficients non nuls et on peut via des opérations élémentaires sur les colonnes se ramener à la matrice identité, ce qui permet un calcul pratique de l'inverse de A.
- Lorsque $p \geq n$ et la matrice est de rang n, on peut définir des inconnues principales associées aux colonnes des pivots, les autres inconnues devenant alors des paramètres.

Exemple 20. On cherche à résoudre le système $(\mathscr{S}_B)$, de paramètre $a \in \mathbb{R}$, $\begin{pmatrix} 1 & 2 & 3 \\ 3 & 6 & 10 \\ -1 & -2 & -2 \end{pmatrix} X = \begin{pmatrix} 4 \\ 17 \\ 1+a \end{pmatrix}$, équivalent au système $\begin{pmatrix} 1 & 2 & 3 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} X = \begin{pmatrix} 4 \\ 5 \\ a \end{pmatrix}$. On a donc une condition de compatibilité pour avoir existence d'une solution : il faut que $a = 0$. Si c'est le cas, on peut par exemple paramétrer les solutions selon x_2 en résolvant $\begin{pmatrix} 1 & 3 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_3 \end{pmatrix} = \begin{pmatrix} 4 - 2x_2 \\ 5 \end{pmatrix}$ qui admet bien une unique solution (à x_2 fixé).

II. B. Applications

[AK02, PIII/Ch7, p105/135]

Théorème 21. *$\mathcal{SL}_n(\mathbb{K})$ est engendré par les transvections, $\mathcal{GL}_n(\mathbb{K})$ est engendré par les transvections et les dilatations.*

Théorème 22. [Factorisation LU] *Soit $A \in \mathcal{M}_n(\mathbb{K})$ une matrice dont tous les mineurs principaux sont non nuls. Alors il existe un unique couple (L, U) de matrices tel que L est triangulaire inférieure avec diagonale de 1, U est triangulaire supérieure et $A = LU$.*

Théorème 23. [Factorisation de Cholesky] *Soit A une matrice symétrique réelle définie positive. Alors il existe une unique matrice B triangulaire inférieure telle que tous ses éléments diagonaux soient positifs et $A = BB^*$.*

Exemple 24. $\begin{pmatrix} 1 & 2 & 1 \\ 0 & -1 & 0 \\ 3 & 6 & 5 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 3 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 & 1 \\ 0 & -1 & 0 \\ 0 & 0 & 2 \end{pmatrix}$

$$\begin{pmatrix} 16 & 8 & 20 \\ 8 & 13 & 10 \\ 20 & 10 & 26 \end{pmatrix} = \begin{pmatrix} 4 & 0 & 0 \\ 2 & 3 & 0 \\ 5 & 0 & 1 \end{pmatrix} \begin{pmatrix} 4 & 2 & 5 \\ 0 & 3 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Application 25. [Résolution de $(\mathscr{S}_B)$]
Si $A \in \mathcal{GL}_n(\mathbb{K})$ admet une décomposition $A = LU$, on résout $(\mathscr{S}_B)$ en résolvant $LY = B$ puis $UX = Y$. Ces deux systèmes étant triangulaires, ils ont un coût moins important.
De même dans le cas de la factorisation de Cholesky.

Exemple 26. $\begin{pmatrix} 16 & 8 & 20 \\ 8 & 13 & 10 \\ 20 & 10 & 26 \end{pmatrix} X = \begin{pmatrix} 0 \\ 27 \\ -2 \end{pmatrix}$ donne $Y = \begin{pmatrix} 0 \\ 9 \\ -2 \end{pmatrix}$ puis $X = \begin{pmatrix} 1 \\ 3 \\ -2 \end{pmatrix}$.

III. Résolution numérique

III. A. Minimisation de fonctionnelle [FGN12, §1.21, p39–41]

On suppose dans cette section que $n = p$ et $A \in \mathcal{S}_n^{++}(\mathbb{R})$ est une matrice symétrique définie positive. On pose $\langle . \mid . \rangle$ le produit scalaire usuel sur $\mathbb{R}^n$.

Proposition 27. *La résolution de $(\mathscr{S}_B)$ est équivalente au problème de minimisation de $f : x \longmapsto \frac{1}{2}\langle AX \mid X\rangle - \langle B \mid X\rangle$: f admet en effet un unique minimum X_0, caractérisé par $AX_0 = B$.*

Théorème 28. [Méthode de gradient à pas optimal]
Soit $f : \mathbb{R}^p \longrightarrow \mathbb{R}$ elliptique, c'est-à-dire f est $\mathcal{C}^1$ et telle qu'il existe $\alpha > 0$ satisfaisant

$$\forall x, y \in \mathbb{R}^p, \langle \nabla f(x) - \nabla f(y) \mid x - y\rangle \geq \alpha \left\| x - y \right\|^2$$

Alors la suite $(x_n)_{n\in\mathbb{N}}$ définie par $x_0 \in \mathbb{R}^p$ et $x_{n+1} = x_n - \rho_n \nabla f(x_n)$ où $\rho_n = \operatorname{argmin}_{\rho>0} f(x_n - \rho \nabla f(x_n))$ converge vers l'unique minimum global de f.

Application 29. Prenant f la fonctionnelle quadratique définie précédemment, on obtient la convergence de toute suite $(x_n)_{n\in\mathbb{N}}$ définie par $x_0 \in \mathbb{R}^p$ et $x_{n+1} = x_n - \rho_n \nabla f(x)$

III. B. Méthode des moindres carrés [AK02, Ch7, p135]

Proposition 30. *Soit $A \in \mathcal{M}_{n,\,p}(\mathbb{R})$ avec $n \geq p$ telle que $\operatorname{rg}(A) = p$ (ou encore $\ker(A) = \{0\}$). Alors $A^\top A \in \mathcal{S}_p^{++}(\mathbb{R})$.*

Application 31. [Moindres carrés]
Soient $n, p \in \mathbb{N}$ et $A \in \mathcal{M}_{n,\,p}(\mathbb{R})$ et $b \in \mathbb{R}^n$. On cherche un vecteur $x \in \mathbb{R}^p$ tel que $\|b - Ax\|_2 = \inf_{y\in\mathbb{R}^p} \|b - Ay\|_2$.

(i) Il existe toujours une solution au problème,
(ii) x est solution si et seulement si $A^\top Ax = A^\top b$,
(iii) Si $n \geq p$ et $\operatorname{rg}(A) = p$, $A^\top A$ est inversible et il existe une unique matrice réelle B triangulaire inférieure, à diagonale positive, telle que $A^\top A = BB^\top$, et alors $x = (BB^\top)^{-1}A^\top b$ est une solution.

Application 32. [Régression linéaire en statistiques]
Etant donnés des points $(x_i, y_i)_{1\leq i\leq n}$, on souhaite tester si les points sont plus ou moins alignés. Pour cela, on cherche la droite $y = ax + b$ de meilleure approximation, c'est-à-dire on cherche a et b minimisant $\sum_{1\leq i\leq n} |y_i - ax_i - b|^2$. On se ramène donc au problème précédent en posant $A = \begin{pmatrix} 1 & x_1 \\ \vdots & \vdots \\ 1 & x_n \end{pmatrix}$ et $X = \begin{pmatrix} b \\ a \end{pmatrix}$.

ANNEXE

Moindres carrés

SPEECH

La formule de Cramer permet la résolution de systèmes linéaires mais est beaucoup trop complexe en pratique, on envisage d'autres méthodes de résolution. Dans le cas d'une matrice triangulaire, par exemple, on a une complexité en n^2 avec la méthode des remontées.

QUESTIONS

Q Soit $n \geq 2$ et $a_1, \ldots, a_n \in \mathbb{K}$. Soient $P_1, \ldots, P_n \in \mathbb{R}[X]$ tels que $\deg P_k = k-1$. Calculer $\det(P_i(a_j))_{i,j}$.

R Notons M la matrice en question. On suppose les polynômes unitaires pour simplifier. La famille est une base car est échelonnée en degrés.
On se ramène à la base canonique par transvections (ou par procédé d'orthogonalisation de Gram-Schmidt), ce qui donne le déterminant de Vandermonde, soit $\prod_{i<j}(a_j - a_i)$.

Q Calculer le déterminant de $\begin{pmatrix} \alpha_1 & a & \ldots & a \\ b & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & a \\ b & \ldots & b & \alpha_n \end{pmatrix}$.

R Regardons le polynôme $x \longmapsto \det((m_{ij} + x))$ de degré au plus n. Par des opérations sur les lignes, on enlève les x de chaque ligne sauf la première, ce qui montre en fait que le polynôme est de degré 1.
Appliquons ceci à la matrice. En évaluant en $-a$ et en $-b$, on obtient deux points d'annulation si $a \neq b$, et donc le déterminant est nul.
Si $a = b$, on utilise la continuité du déterminant par rapport aux coefficients. On pose $f(x) = u(a,b)x + v(a,b)$ où $u, v \in \mathcal{C}^\infty$.

Q Montrer que $\mathcal{SL}_n(\mathbb{R})$ est connexe par arcs.

R Utiliser le fait que $\mathcal{SL}_n(\mathbb{R})$ est engendré par les transvections.

Q Montrer que $\mathcal{GL}_n(\mathbb{R})$ est engendré par les inversibles diagonalisables.

BIBLIOGRAPHIE

[AK02] G. Allaire et S.-M. Kaber : *Algèbre linéaire numérique*. Ellipses, 2002.

[CG13] P. Caldero et J. Germoni : *Histoires hédonistes de groupes et de géométries - Tome 1*. Calvage et Mounet, 2013.

[FGN12] S. Francinou, H. Gianella et S. Nicolas : *Oraux X-ENS - Analyse 4*. Cassini, 2012.