162

Systèmes d'équations linéaires; opérations, aspects algorithmiques et conséquences théoriques

Cadre : Soit $\mathbb{K}$ un corps commutatif

I] Systèmes linéaires

1) Définitions

Déf 1 : Le système $\begin{cases} a_{11}x_1 + \dots + a_{1n}x_n = b_1 \\ \vdots \\ a_{p1}x_1 + \dots + a_{pn}x_n = b_p \end{cases}$ est un système linéaire de p équations à n inconnues, où $a_{ij}, b_j \in \mathbb{K}$. On appelle solution tout vecteur $(x_1, \dots, x_n) \in \mathbb{K}^n$ dont les x_i vérifient toutes les équations. Le système est dit compatible s'il admet au moins une solution.

Exemple 2 : ① $\begin{cases} x_1 + x_2 = 2 \\ x_2 = 0 \end{cases}$ est compatible.

② $\begin{cases} x_1 + x_2 = 2 \\ x_1 + x_2 = 0 \end{cases}$ est non compatible

Expression matricielle : pour $A = (a_{ij}) \in \mathcal{M}_{p,n}(\mathbb{K})$, $b = \begin{pmatrix} b_1 \\ \vdots \\ b_p \end{pmatrix}$ et $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$, le système peut s'écrire sous la forme matricielle : $Ax = b$.

On appelle rang du système le rang de A.

Proposition 3 : En notant $(C_1, \dots, C_n)$ les colonnes de A, on a : le système $Ax = b$ est compatible ssi $b \in \mathrm{Vect}(C_1, \dots C_n)$

Exemple 4 : ① $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 2 \\ 0 \end{pmatrix}$ compatible et $\begin{pmatrix} 2 \\ 0 \end{pmatrix} = 2\begin{pmatrix} 1 \\ 0 \end{pmatrix} + 0\begin{pmatrix} 1 \\ 1 \end{pmatrix}$

② $\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 2 \\ 0 \end{pmatrix}$ non compatible et $\begin{pmatrix} 2 \\ 0 \end{pmatrix} \notin \mathbb{K}.\begin{pmatrix} 1 \\ 1 \end{pmatrix}$

Déf 5 : Notons $\|.\|$ une norme vectorielle quelconque et $|||.|||$ la norme matricielle subordonnée correspondante. Pour A inversible, on appelle conditionnement de A le nombre $\mathrm{Cond}(A) := |||A|||.|||A^{-1}|||$.

Thm 6 : Soient $A \in GL_n(\mathbb{K})$ et u, $u + \delta u$ les solutions de $Au = b$ et $A(u + \delta u) = b + \delta b$. Alors pour $b \neq 0$, on a : $\frac{\|\delta u\|}{\|u\|} \leqslant \mathrm{cond}(A) \cdot \frac{\|\delta b\|}{\|b\|}$ et pour A donnée, il existe $b \neq 0$ et $\delta b \neq 0$ tels qu'on ait égalité

2) Systèmes de Cramer

Déf 7 : On appelle système de Cramer un système linéaire avec A carrée et inversible.

Proposition 8 : Un système de Cramer admet une unique solution donnée par $x = A^{-1}b$

En pratique (théorème de Cramer) : La solution $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ d'un système de Cramer est donnée par les formules de Cramer suivantes :

$$x_i = \frac{\det(C_1, \dots, C_{i-1}, b, C_{i+1}, \dots, C_n)}{\det A}.$$

Exemple 9 : pour $\begin{cases} 2x - 5y + 2z = 7 \\ x + 2y - 4z = 3 \\ 3x - 4y - 6z = 5 \end{cases}$ on a $\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 5 \\ 1 \\ 1 \end{pmatrix}$

Complexité : recquiert de l'ordre de $(n+2)!$ opérations

3) Cas général

Théorème 10 : Théorème de Rouché-Fontené

À permutation près, on suppose que $\begin{vmatrix} a_{11} & & a_{1r} \\ a_{r1} & \dots & a_{rr} \end{vmatrix} \neq 0$, où $r = rg(A)$.

Alors le système est compatible ssi :

$\forall s \in [\![r+1, n]\!]$ $\begin{vmatrix} a_{11} & a_{1r} & b_1 \\ a_{r1} & a_{rr} & b_r \\ a_{s1} & a_{sr} & b_s \end{vmatrix} = 0$ Dans ce cas, le système est équivalent à : $\begin{cases} a_{11}x_1 + \dots + a_{1r}x_r = b_1 - a_{1,r+1}x_{r+1} - \dots - a_{1n}x_n \\ \vdots \\ a_{r1}x_1 + \dots + a_{rr}x_r = b_r - a_{r,r+1}x_{r+1} - \dots - a_{rn}x_n \end{cases}$

où les variables $x_{r+1}, \dots, x_n$ prennent des valeurs arbitraires.

II] Systèmes échelonnés et résolution directe

1) Opérations élémentaires

Propriété 11 : L'ensemble des solutions d'un système linéaire ne change pas si on effectue sur les équations les opérations (dites élémentaires) suivantes :
- changement de l'ordre des équations
- multiplication d'une équation par un scalaire non nul
- ajout à une équation d'une combinaison linéaire des autres équations.

Déf 12 : matrice de dilatation $D_{i,\alpha}^{(p)} := \begin{pmatrix} I_{i-1} & & (0) \\ & \alpha & \\ (0) & & I_{p-i} \end{pmatrix} \in \mathcal{M}_p(\mathbb{K})$,

matrice de transvection : $T_{i,j,\beta}^{(p)} := I_p + \beta E_{i,j} \in \mathcal{M}_p(\mathbb{K})$,

matrice de permutation : $P_{i,j}^{(p)} := \begin{pmatrix} I_{i-1} & & & \\ & \begin{matrix} 0 & \cdots & 1 \\ \vdots & I_{j-i-1} & \vdots \\ 1 & \cdots & 0 \end{matrix} & \\ & & I_{p-j} \end{pmatrix} \in \mathcal{M}_p(\mathbb{K})$

Traduction matricielle :

Notons $L_1, \dots, L_p$ les lignes de A.

opération	$D_{i,\alpha}^{(p)} A$	$T_{i,j,\beta}^{(p)} A$	$P_{i,j}^{(p)} A$
résultat	$L_i \leftarrow \alpha L_i$	$L_i \leftarrow L_i + \beta L_j$	$L_i \leftrightarrow L_j$

idem avec les colonnes mais en multipliant à droite.

2) Systèmes échelonnés

Déf 13 : On appelle pivot d'une ligne non nulle le coefficient non nul situé le plus à gauche.

Déf 14 : Une matrice est dite échelonnée en lignes lorsque :
- si une ligne est nulle alors toutes les lignes suivantes sont nulles
- le pivot d'une ligne est strictement plus à droite que les pivots des lignes précédentes

Et elle est dite réduite si, de plus, tous les pivots sont égaux à 1 et les pivots sont les seuls coefficients non nuls de leur colonne.

Exemple 15 : $\begin{pmatrix} 0 & 1 & 0 & a & 0 & b & c \\ 0 & 0 & 1 & d & 0 & e & f \\ 0 & 0 & 0 & 0 & 1 & g & h \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$ est échelonnée.

Théorème 16 : Notons $\mathcal{E}_{p,n}$ l'ensemble des matrices échelonnées en lignes réduites de taille $p \times n$. Alors :

$$\mathcal{M}_{p,n}(\mathbb{K}) = \bigsqcup_{E \in \mathcal{E}_{p,n}} \{PE \mid P \in GL_p(\mathbb{K})\}.$$

3) Méthode du pivot de Gauss

Méthode :
- échange éventuel de lignes et colonnes afin que le pivot de la première ligne soit non nul.
- on échelonne ensuite le système : $\forall i \in [\![2,p]\!]\ \ L_i \leftarrow L_i - \frac{a_{i1}}{a_{11}} L_1$
- on réitère le procédé en partant de la deuxième ligne.

Exemple 17 :

$$\begin{cases} x+2y-2z+3w=2 \\ 2x+4y-3z+4w=5 \\ 5x+10y-8z+11w=12 \end{cases} \Leftrightarrow \begin{cases} x+2y-2z+3w=2 \\ z-2w=1 \end{cases} \Leftrightarrow \begin{cases} x=4-2y+w \\ z=1+2w \end{cases}$$

Applications 18 :
- calcul du rang d'une matrice
- calcul de l'inverse
- recherche d'un système d'équations d'un sev défini par une famille génératrice
- recherche d'une base d'un sev défini par un système d'équations
- engendrement de $SL_n(\mathbb{K})$ par les transvections
- décomposition de Bruhat

Complexité : $O(n^3)$ (pour $A \in \mathcal{M}_n(\mathbb{K})$)

4) Factorisation LU

Théorème 19 : Soit $A \in M_n(\mathbb{K})$ telle que $\forall k \in [\![1,n]\!]$ $\begin{pmatrix} a_{11} & a_{1k} \\ a_{k1} & a_{kk} \end{pmatrix}$ est inversible.

Alors il existe une matrice triangulaire inférieure L et une matrice triangulaire supérieure U telles que $A = LU$ et cette décomposition est unique (avec $l_{ii} = 1$)

Théorème 20 : Soit $A = \begin{pmatrix} b_1 & c_1 & & (0) \\ a_2 & & & \\ & & & c_{n-1} \\ (0) & & a_n & b_n \end{pmatrix}$. On définit (δ_k) par $\delta_0 = 1$, $\delta_1 = b_1$ et $\delta_k = b_k \delta_{k-1} - a_k c_{k-1} \delta_{k-2}$, pour $k \in [\![2,n]\!]$

Si $\forall k \in [\![1,n]\!]$ $\delta_k \neq 0$ alors

$$A = LU = \begin{pmatrix} 1 & & & (0) \\ a_2 \frac{\delta_0}{\delta_1} & \ddots & & \\ & \ddots & \ddots & \\ (0) & & a_n \frac{\delta_{n-2}}{\delta_{n-1}} & 1 \end{pmatrix} \begin{pmatrix} \frac{\delta_1}{\delta_0} & c_1 & & (0) \\ & \ddots & \ddots & \\ & & \ddots & c_{n-1} \\ (0) & & & \frac{\delta_n}{\delta_{n-1}} \end{pmatrix}$$

Remarque : procédé simple de résolution d'un système linéaire avec A tridiagonale.

Complexité : $8n - 6$ opérations

III] Méthodes itératives de résolution de systèmes linéaires.

1) Principe des méthodes itératives

Déf 21 : Pour une méthode itérative, la solution x est la limite d'une suite $(x^{(k)})$ définie par $\begin{cases} x^{(0)} \text{ arbitraire} \\ x^{(k+1)} = Bx^{(k)} + c \end{cases}$, où B et c dépendent du système linéaire $Ax = b$.

Thm 22 Il y a équivalence entre
(i) la méthode itérative est convergente
et (ii) $\rho(B) < 1$.

2) Méthode de relaxation DVLPT 2

Déf 23 : Soit $\omega \neq 0$. La méthode de relaxation consiste à décomposer $A = \left(\frac{D}{\omega} - E\right) - \left(\frac{1-\omega}{\omega} D + F\right)$, où $D = (a_{ij}\delta_{ij})_{i,j}$, $E = (-a_{ij}\delta_{i>j})_{i,j}$ et $F = (-a_{ij}\delta_{j>i})_{i,j}$

Et on a : $x^{(k+1)} = \left(\frac{D}{\omega} - E\right)^{-1}\left(\frac{1-\omega}{\omega} D + F\right) x^{(k)} + b$.

Thm 24 : Soit A hermitienne définie positive.
Alors : (la méthode de relaxation converge) $\iff (\omega \in]0,2[)$

3) Méthode de gradient à pas optimal DVLPT3

Déf 25 : On note $J(v) := \frac{1}{2}\langle Av, v\rangle - \langle b, v\rangle$
(avec $A \in \mathcal{S}_n^{++}(\mathbb{R})$) $\quad = \frac{1}{2}\sum_{i,j=1}^{n} a_{ij} v_i v_j - \sum_{i=1}^{n} b_i v_i$

Déf 26 : La méthode de gradient à pas optimal est définie par les relations :

$$\begin{cases} J(x^{(k)} - \rho_k \nabla J(x^{(k)})) = \inf_{\rho \in \mathbb{R}} J(x^{(k)} - \rho \nabla J(x^{(k)})) \\ x^{(k+1)} = x^{(k)} - \rho_k \nabla J(x^{(k)}) \end{cases}$$

Thm 27 : La méthode de gradient à pas optimal converge.

Références

Caldero - Germoni, *Histoires hédonistes de groupes et de géométries*

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

Grifone, *Algèbre linéaire*

Oraux X-ENS algèbre 1 (pour BRUHAT)