Systèmes linéaires, opérations, aspects algorithmiques et conséquences théoriques

162

140

I) Définitions et compatibilité

1) Définitions. Interprétation matricielle.

A un anneau, $p, n \in \mathbb{N}^*$, $(a_{i,j})_{\substack{1\leq i\leq p\\ 1\leq j\leq n}}$, $(b_i)_{1\leq i\leq p} \in A$.

On considère le système suivant :

$$(S)\begin{cases} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n = b_1 \\ \vdots \\ a_{p1}x_1 + a_{p2}x_2 + \dots + a_{pn}x_n = b_n \end{cases}$$

On cherche les vecteurs $x=(x_1,\dots,x_n)$ satisfiant toutes les équations du système.

Si une telle solution existe, le système est dit compatible

En posant $A=(a_{ij}) \in \mathcal{M}_{p,n}(A)$, $B=\begin{pmatrix} b_1 \\ \vdots \\ b_p\end{pmatrix}$, $X=\begin{pmatrix} x_1 \\ \vdots \\ x_n\end{pmatrix}$,

le système (S) se ramène à la résolution de $\boxed{AX=B}$

Def 0 : rang(S) = rang(A)

Autre approche : On note $\vec{C_1},\dots,\vec{C_n}$ les colonnes de A. Alors,

$(S) \Leftrightarrow x_1\vec{C_1}+\dots+x_n\vec{C_n}=\vec{b}$ et $\boxed{(S) \text{ compatible} \Leftrightarrow \vec{b} \in \mathrm{Vect}(\vec{x_1},\dots,\vec{x_n})}$

2) Systèmes de Cramer

Jusqu'à mention contraire, $A=K$ est un corps

Def 1 : On appelle système de Cramer tout système linéaire dont la matrice A est carrée (c-à-d $p=n$) et inversible

Dans ce cas, on sait que $X=A^{-1}B$ est l'unique solution.

De plus, $\det A \neq 0$ et on peut utiliser les

Formules de Cramer : On note $\begin{pmatrix} x_1 \\ \vdots \\ x_n\end{pmatrix}=A^{-1}B$.

Alors, pour $1\leq i\leq n$, $x_i = \dfrac{\det(C_1,\dots,C_{i-1},b,C_{i+1},\dots,C_n)}{\det A}$

[DEV 1] Application : il existe une ~~suite~~ suite $(a_n)_{\mathbb{N}}$ de réels telle que, pour tout $n\in\mathbb{N}$, $\sum_{k=0}^{\infty} 2^{nk} a_k = (-1)^n$

Complexité : $O(n \cdot n!)$, ce qui est très mauvais.

3) Un théorème de compatibilité

Thm de Rouché-Fontené :

Soit $\begin{cases} \boxed{\begin{matrix} a_{11}x_1 + \dots + a_{1r}x_1 \\ \vdots \\ a_{r1}x_1+\dots+a_{rr}x_r\end{matrix}} + \dots + a_{1n}x_n = b_1 \\ \qquad\qquad + \dots + a_{rn}x_n = b_r \\ \vdots \\ a_{p1}x_1+\dots+a_{pr}x_r+\dots+a_{pn}x_n = b_p\end{cases}$

un système de rang $r\leq p$

On peut supposer que $\delta = \det\begin{pmatrix} a_{11} & - & a_{1r} \\ | & & | \\ a_{r1} & - & a_{rr}\end{pmatrix} \neq 0$

Alors :

1) (S) compatible $\Leftrightarrow \forall r+1\leq s\leq p$, $\Delta_s = \det\left(\begin{array}{cc|c} a_{11} - & a_{1r} & b_1 \\ a_{r1} - & a_{rr} & b_r \\ a_{s1} - & a_{sr} & b_s\end{array}\right)=0$

2) Dans ce cas, le système équivaut au système de Cramer à paramètres :

$$\begin{cases} a_{11}x_1+\dots+a_{1r}x_r = b_1 - a_{1,r+1}x_{r+1} - \dots - a_{1n}x_n \\ \vdots \\ a_{r1}x_1+\dots+a_{rr}x_r = b_r - a_{r,r+1}x_{r+1} - \dots - a_{rn}x_n\end{cases}$$

Et les solutions forment un espace affine de dimension $n-r$, et de direction $\mathrm{Ker}\, A$.

Exemple : $\begin{cases} 2x+y-z=\alpha \\ \quad y+3z=\beta \\ 2x+y+2z=\alpha+\beta\end{cases} \Leftrightarrow \begin{pmatrix} x\\y\\z\end{pmatrix} \in \begin{pmatrix} \frac{\alpha-\beta}{2} \\ \beta \\ 0\end{pmatrix} + \mathrm{Vect}\begin{pmatrix} 2 \\ -3 \\ 1\end{pmatrix}$.

II) Résolution dans un corps

1) Système échelonné, pivot de Gauss

Def 2 : Une matrice est dite échelonnée si ses lignes commencent par un nombre de zéros strictement croissant, jusqu'à ce que les lignes soient nulles. ex : $\begin{pmatrix} 2 & 3 & 1 & 0 \\ 0 & 0 & 7 & 1 \\ 0 & 0 & 0 & 0\end{pmatrix}$

Pour un tel système, le rang du système et ses solutions se "lisent" immédiatement par résolution "en cascade".

But: Se ramener à un système échelonné. Pour cela:

Prop 3: Les solutions de (S) sont invariantes par
1) changement de l'ordre de équations
2) Multiplication d'une équation par une constante
3) Ajout d'une combinaison linéaire des autres lignes à une équation.

On appelle ces opérations des opérations élémentaires

Thm 4: Toute matrice peut se ramener à une matrice échelonnée en ne faisant que des opérations élémentaires par le pivot de Gauss:

Pour $i = 1$ à $p-1$
  Pour $j = i+1$ à p
    Si $\exists a_{ki} \neq 0$, $k \geq i$ alors $a_{ii} \leftrightarrow a_{ki}$
      $l_j \leftarrow l_j - \frac{a_{ji}}{a_{ii}} l_i$
      Pour $k = i+1$ à n, $a_{jk} \leftarrow a_{jk} - \frac{a_{ji}}{a_{ii}} a_{ik}$
      $a_{ji} \leftarrow 0$

Complexité: Pivot en $O(n \times p^2)$ $(\sim 2np^2)$
puis: Résolution d'un système échelonné en $O(n \times p)$
→ Mieux que Cramer, mais toujours lent.

Autre problème: forte propagation des erreurs d'arrondi
⇒ Recherche de méthodes plus rapides et mieux conditionnés.

2) Décompositions LU, LUP

Thm 5: $A \in GL_n$. [DEV 2]
Si $\det(A[1..k, 1..k]) \neq 0$ pour tout $k = 1 \dots n$, alors il existe une unique U échelonnée et une unique L triang. inférieure à diagonale unité telle que $\boxed{A = LU}$

Cette décomposition est liée au pivot de Gauss, cependant, on peut l'améliorer en choisissant un pivot le plus grand de la ligne en cours et en l'amenant en position de pivot par permutation.

Prop 6: Pour toute $A \in GL_n$, il existe P matrice de permutation telle que $AP = LU$

Intérêt: Pas de gain de temps, mais le choix d'un meilleur pivot réduit fortement les erreurs de calcul

3) Décomposition de Cholesky

Thm 7: $A \in \mathbb{C}^{n \times n}$ hermitienne, définie positive.
Alors il existe une unique L triangulaire inférieure avec des réels ≥ 0 sur la diagonale telle que $A = LL^*$

De plus on dispose d'un algorithme de calcul en $O(n^3)$, mais avec une constante meilleure que pour LUP.

4) Décomposition QR, matrices de Householder

Def 8 [Matrice de Stiefel]: $M \in \mathbb{C}^{m \times n}_{(n \geq m)}$ est dite de Stiefel si ses n-vecteurs colonnes forment une base orthonormée de $\mathbb{C}^m$
→ Notation $M \in St_{m,n}$

Rem: C'est une généralisation de la notion de matrices unitaires

Prop 9: $A \in \mathbb{C}^{m \times n}$ de rang n possède une unique décomposition $A = QR$, avec $Q \in St_{m,n}$, et R triangulaire supérieure à diagonale > 0

Rem: L'existence est une conséquence du procédé d'orthonormalisation de Gramm-Schmidt.

Complexité : G-S effectue $\sim 2mn^2$ opérations.

$\to$ On peut faire mieux

Def 10 : $w \in \mathbb{C}^m$ unitaire.

On appelle matrice de Householder associée à w

$$H_w = I_m - 2\, w\, w^*$$

Prop 11 : H_w est la symétrie orthogonale par rapport à $\{w\}^\perp$

H_w est hermitienne unitaire

Prop. 12 : $u_1, u_2 \in \mathbb{K}^m$ indépendants. Alors $\exists w \in \mathbb{K}^m$ unitaire, et $k \in \mathbb{K}^*$, $H_w(u_1) = k u_2$

Thm 13 : $\exists Q \in U_n$ produit d'au plus n matrices de Householder, R triangulaire sup. $\in \mathbb{C}^{n \times n}$, t.q. $A = QR$

L'algorithme associé calcule R en $\sim 2mn^2 - \frac{2}{3}n^3$, ce qui est mieux que par G-S : Ici l'algorithme dans $\mathbb{R}$:

```
pour j=1 à m
    α² ← Σ_{i=j+1}^{m} |a_ij|²
    si α ≠ 0, alors
        β ← √(|a_jj|² + α²)
        w_j ← a_jj − β
        pour i = j+1 à m
            w_i ← a_ij
        H ← I_{m−j+1} − 2 w w^T / ||w||_2²
        A[j:m ; j:n] = H A[j:m ; j:n]
```

Conséquence : les opérations élémentaires engendrent $GL_n(\mathbb{K})$

III) Résolution dans un anneau principal

1) Facteurs invariants

Thm 14 : $A \in \mathcal{M}_{m,n}(\mathbb{A})$, avec $\mathbb{A}$ anneau principal.

Alors A est équivalente à $\mathrm{diag}(d_1, d_2, \dots, d_r, 0, \dots 0)$, avec $d_i \neq 0$ et $d_i \mid d_j$ si $i \leq j$

Algorithme : Dans le cas particulier où $\mathbb{A}$ est euclidien pour un stathme $|\cdot| : \mathbb{A} \to \mathbb{N}$ dispose d'une méthode de calcul des invariants :

1) Placer l'élément de stathme minimal en position $(1,1)$ par des opérations élémentaires

2) pour i de 2 à n : division euclidienne de a_{i1} par a_{11}
$a_{i1} = q_i a_{11} + r_i$, on effectue $L_i \leftarrow L_i - q_i L_1$

3) On obtient $A = \begin{pmatrix} a_{11} & * \cdots * \\ r_2 & \\ \vdots & A' \\ r_n & \end{pmatrix}$.

Si tous les r_i sont nuls, aller à 1) avec A'

Sinon, placer r_i de stathme minimal en $(1,1)$, et recommencer en 2).

Rem : C'est une généralisation du fait que toute matrice de $\mathcal{M}_{n,m}(\mathbb{K})$ est équivalente à $J_r = \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}$, $r = \mathrm{rg}\, A$.

2) Conséquences théoriques

Thm de la base adaptée :

Soit N un sous module de $M = \mathbb{A}^n$, de rang s.
Alors il existe $(e_1, \dots, e_n)$ base de $\mathbb{A}^n$, et $q_1 | \cdots | q_s \neq 0$ tels que $(q_1 e_1, \dots, q_s e_s)$ soit une $\mathbb{A}$-base de N

Théorème de structure des quotients

N sous-module de $\mathbb{A}^n$ de rang s. Alors, $X = \mathbb{A}^n / N$,
$X \cong \prod_{i=1}^{s} \mathbb{A}/q_i\mathbb{A} \times \mathbb{A}^r$, avec $q_1 | \cdots | q_s \neq 0$, $r = n - s$

Références :

Grifone – Algèbre linéaire (Partie I)).

Amodei, Dedieu – Analyse numérique matricielle (Partie II))

Gras, Gras – Algèbre fondamentale Arithmétique (Partie III)).

Oraux X-ENS, Algèbre 2, Analyse 2 (Développements)

Autres possibilités de parties :

* Méthodes de résolution itératives
* Méthodes de gradient
* Lien avec les équations différentielles (soit système d'EDO, soit linéarisé d'EDP)

Autres Devs possibles :

* Toutes les décompositions (LU, QR, Bruhat, ...)
* Rouché – Fontené
* Transformée de Fourier rapide
* Lemme de Siegel
* Convergence des méthodes itératives

⋮