

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

162

Soient $n, p \in \mathbb{N}^*$, K un corps commutatif.

I) Systèmes linéaires: existence, unicité d'une solution

A) Définitions et premières propriétés

Def 1: On appelle système linéaire de p -équations à n inconnues tout système sur K de la forme:

$$(S) \begin{cases} \sum_{j=1}^n a_{1j} x_j = b_1 \\ \vdots \\ \sum_{j=1}^n a_{pj} x_j = b_p \end{cases} \text{ d'inconnues } (x_1, \dots, x_n) \in K^n.$$

(S) est dit compatible si l'ensemble des solutions \mathcal{S} est non vide.

Ex 2: $\begin{cases} x-y=1 \\ x+y=1 \\ 2x-y=1 \end{cases}$ n'est pas compatible. $\begin{cases} x-y=1 \\ x+y=1 \end{cases}$ l'est.

Rem 3: Soient $A = (a_{ij})_{\substack{1 \leq i \leq p \\ 1 \leq j \leq n}}$, $X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ et $b = \begin{pmatrix} b_1 \\ \vdots \\ b_p \end{pmatrix}$. (S) équivaut alors à $AX = b$. Notons $A = (A_1, \dots, A_n)$ en colonnes.

Prop 4: (S) est compatible si et seulement si $b \in \text{Vect}(A_1, \dots, A_n)$

Ex 5: Un système homogène ($b=0$) est toujours compatible.

Prop 6: Si $\mathcal{S} \neq \emptyset$, \mathcal{S} est un sous-espace affine de direction $\text{Ker } A$.

Cor 7: Si (S) est compatible, $\dim \mathcal{S} \geq n-p$

On suppose que $n=p$ jusqu'à la fin de cette partie.

Lem 8: ($p \geq 1$) La dimension du commutant de A est supérieure à n .

Thm 9: Soit $C(A)$ le commutant de A . $C(A) = K[A] \iff \lambda_A = \pi_A$

B) Système de Cramer

On suppose ici $n=p$. Alors $A \in \text{Mat}(K)$.

Def 10: (S) est un système de Cramer si $A \in \text{GL}_n(K)$.

Ex 11: Le deuxième système de Ex 2 est de Cramer.

Thm 12: Si (S) est de Cramer, il admet une unique solution donnée par: $x_i = \frac{\det(A_1, \dots, A_{i-1}, b, A_{i+1}, \dots, A_n)}{\det(A)} \quad \forall i \in [1; n]$.

Ex 13: Si $A = \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}$, $b = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$, $\det A = 2$, $\det(A, b) = 2$, $\det(A, b) = 0$ soit $x=1$ et $y=0$

Rem 14: En pratique, les formules de Cramer demandent environ $n^2 n!$ opérations ce qui rends son intérêt caduque pour de grandes valeurs de n .

C) Cas général

On ne suppose plus $n=p$.

Prop 15: Soit $r \in \mathbb{N}^*$. $\text{rg}(A) = r$ si et seulement si il existe une matrice carrée de taille r de A qui soit inversible, avec r maximal pour cette propriété.

Soit $r = \text{rg}(A)$ et A_{rj} une matrice extraite de A de taille r inversible correspondant aux lignes I et aux colonnes J , classés dans l'ordre croissant.

Def 16: Le système $\sum_{j=1}^n a_{ij} x_j = b_i \quad \forall i \in I$ est appelé système d'équations principales de A . On notera \mathcal{S} l'ensemble des solutions.

Lem 17: Si (S) est compatible, $\mathcal{S} = \mathcal{S}$.

Thm 18: (de Rouché-Forté) (S) est compatible si et seulement si $n=r$ ou $r \leq n-1$ et $\det \begin{pmatrix} a_{ij} & (b_i)_{i \in I} \\ (a_{kj})_{k \in [1; n] \setminus I} & b_k \end{pmatrix} = 0 \quad \forall k \in [1; n] \setminus I$. Dans ce cas, il est équivalent à $A_{r,r} x_r = b_r$ où $x_r = (x_j)_{j \in J}$ et $b_r = (b_i - \sum_{j \in J} a_{ij} x_j)_{i \in I}$

Rem 19: On est alors ramené à l'inversion d'un système carré inversible.

I) Systèmes échelonnés et méthode du pivot de Gauss

On suppose ici $n = p$ et $A \in GL_n(K)$.

A) Méthode du pivot de Gauss et conséquences

Def 20: (S) est dit échelonné si $a_{ij} = 0 \forall i > j$, autrement dit que A est triangulaire supérieure dans notre cas.

Rem 21: On peut dans ce cas aisément résoudre le système par "ramontée". Le but de cette partie est d'abord transformer notre système en un système échelonné, plus facile à résoudre.

Prop 22: L ne change pas si:

- ⊗ On permute les lignes
 - ⊗ On multiplie une ligne par un scalaire non nul
 - ⊗ On ajoute à une ligne une combinaison linéaire des autres.
- Soit $(E_{ij})_{1 \leq i, j \leq n}$ la base canonique de $M_n(K)$.

Def 23: On appelle matrice de:

- ⊗ transvection $T_{ij}(\lambda) = I_n + \lambda E_{ij}$ avec $i \neq j, \lambda \neq 0$
- ⊗ dilatation: $D_i(\lambda) = \text{diag}(\lambda, \dots, \lambda, 1, \dots, 1)$ $\lambda \neq 1$

Rem 24: $T_{ij}(\lambda)A$ ajoute à la ligne i λ fois la j -ème ligne $D_i(\lambda)A$ multiplie par λ la ligne i et on a une description similaire à droite, avec les colonnes.

Alg 25: (méthode du pivot de Gauss)

- On te à permuter avec une autre ligne, $a_{ii} \neq 0$ (c'est le pivot)
- On annule les coefficients $a_{i+1, i}, i > 1$ en multipliant par les $T_{i+1, i}(-\frac{a_{i+1, i}}{a_{ii}})$ à gauche
- On itère avec $(a_{ij})_{2 \leq i, j \leq n}$
- On obtient à la fin un système échelonné.

EX 26:

$$\begin{cases} x+2y-z=1 \\ 2x+3y+z=2 \\ x+4y-6z=2 \end{cases} \Leftrightarrow \begin{cases} x+2y-z=1 \\ -y+3z=0 \\ 2y-5z=1 \end{cases} \Leftrightarrow \begin{cases} x+2y-z=1 \\ -y+3z=0 \\ z=1 \end{cases} \Leftrightarrow \begin{cases} x=-4 \\ y=3 \\ z=1 \end{cases}$$

Rem 27: En pratique, on a en $O(\frac{1}{3}n^3)$ opérations ce qui rends la méthode de Gauss plus efficace que celle de Cramer.

Rem 28: Cet algorithme se généralise de la même façon pour les matrices rectangulaires. A permutation des lignes près, on obtient alors une matrice échelonnée.

Thm 29: $SL_n(K)$ est engendré par les transvections.

Cor 30: $GL_n(K)$ est engendré par les transvections et les dilatations.

Thm 31: Soit $m \in \mathbb{N}^*$. On considère l'action à gauche de $GL_m(K)$ sur $M_{m, n}(K)$. Soient $B, \tilde{B} \in M_{m, n}(K)$, B et \tilde{B} sont dans la même orbite si et seulement si $\exists k \in GL_m(K)$ tel que $\tilde{B} = kB$. De plus, B est équivalente à une unique matrice échelonnée.

Rem 32: On a en résultat similaire à droite mais avec $\text{im}(B) = \text{im}(\tilde{B})$.

B) Décomposition de matrices (pour $K = \mathbb{R}$)

Thm 33: Soient $A^k = (a_{ij})_{1 \leq i, j \leq n}$, $k \in \mathbb{N}$, $S: \det A > 0 \forall k \in \mathbb{N}$, alors il existe $L \in GL_n(K)$ à diagonale unité et $U \in GL_n(K)$ triangulaire supérieure unique telle que $A^k = LU$.

EX 34: Ceci est en particulier vérifié pour les matrices définies positives.

Rem 35: Si A est juste inversible, $\exists P \in GL_n(K) / PA = LU$.

Cor 36: (décomposition de Cholesky) Si A est symétrique définie positive, il existe une unique matrice B triangulaire inférieure à diagonale positive telle que $A = B^t B$.

[7]

Rem 37: Dans ce cas, $AX=b \Leftrightarrow \begin{cases} BX=Y \\ BY=b \end{cases}$ qui sont des systèmes échelonnés.

III) Résolution approchée de systèmes linéaires

On suppose $n=p$ et $A \in GL_n(\mathbb{R})$

A) Méthodes itératives classiques

Le principe de ces méthodes est de décomposer $A=M-N$ avec M inversible. On définit alors:

$$\begin{cases} x_0 \in \mathbb{R}^n \\ Mx_{k+1} = Nx_k + b \quad \forall k \geq 1 \end{cases}$$

Ainsi, si (x_n) converge, c'est vers la solution $x^* = A^{-1}b$.
On dit que la méthode converge si quel que soit $x_0 \in \mathbb{R}^n$, (x_n) converge. Soit $e_n = x_n - x^*$ l'erreur.

[7]

Lem 38: (DEV2) Pour tout $\epsilon > 0$, et $B \in \mathcal{M}_n(\mathbb{R})$, il existe une norme $\|\cdot\|$ subordonnée telle que $\|B\| \leq \rho(B) + \epsilon$.

Thm 39: La méthode converge si et seulement si $\rho(M^{-1}N) < 1$

Thm 40: Si A est symétrique définie positive et que $M+N$ est aussi, alors la méthode converge.

Ex 41: Méthode de Jacobi; $M = \text{diag}(A)$, $N = M - A$; $T = M^{-1}N$

Ex 42: Méthode de Gauss-Seidel: Soient E la partie inférieure et F la partie supérieure de A ; $M = D + E$, $N = -F$.
Cette méthode converge dès que A est symétrique définie positive.

Thm 43: Si A est tridiagonale, les méthodes de Jacobi et Gauss-Seidel convergent ou divergent simultanément et $\rho(G) = \rho(T)^2$.

Rem 44: Dans ce cas, en cas de convergence, Gauss-Seidel converge plus vite que Jacobi.

B) Méthode du gradient

On suppose ici A symétrique définie positive et on définit la fonctionnelle quadratique: $J(x) = \frac{1}{2} \langle Ax, x \rangle - \langle b, x \rangle$.

Donc $\nabla J(x) = Ax - b$.

Prop 45: Le minimum de J sur \mathbb{R}^n existe et c'est x^* .

On souhaite donc minimiser J sur \mathbb{R}^n . On peut pour cela utiliser une méthode de descente:

- On choisit $x_0 \in \mathbb{R}^n$
- A l'étape k , si $d_k = \nabla J(x_k) = 0$, c'est fini. Sinon, on choisit $d \in \mathbb{R}^n$ tel que $\langle d, d_k \rangle < 0$ et on pose $x_{k+1} = x_k - \alpha_k d_k$ où $\alpha_k = \arg \min_{\alpha \geq 0} J(x_k - \alpha d_k)$

La méthode du gradient à pas optimal consiste à prendre $d = -d_k$.

Soit $\|x\|_A = \sqrt{\langle Ax, x \rangle}$

Lem 46: (DEV3) (inégalité de Kantorovitch)

$$\forall x \in \mathbb{R}^n, x \neq 0, \frac{\|x\|_A^4}{\|x\|_{A^{-1}}^2 \|x\|_A^2} \geq 4 \frac{\lambda_{\max} \lambda_{\min}}{(\lambda_{\max} + \lambda_{\min})^2}$$

$$\text{où } \lambda_{\max} = \max_{\lambda \in \text{Sp}(A)} \lambda, \quad \lambda_{\min} = \min_{\lambda \in \text{Sp}(A)} \lambda$$

Thm 47: La méthode du gradient à pas optimal converge et: $\forall k \in \mathbb{N}, \|x_k - x^*\| \leq \sqrt{\frac{\lambda_{\max}}{\lambda_{\min}}} \left(\frac{\lambda_{\max} - \lambda_{\min}}{\lambda_{\max} + \lambda_{\min}} \right)^k \|x_0 - x^*\|$

Rem 48: Plus $\text{cond}(A)$ est proche de 1, plus la convergence est rapide.

[7]

[7]

Références :

- ① Algèbre linéaire, Guifone [1]
- ② Algèbre et géométrie, Rombaldi [2]
- ③ Algèbre, Gourdon [3]
- ④ Cours d'algèbre, Perrin [4]
- ⑤ Développements d'analyse, Bernis [5]
- ⑥ Algèbre 2, FGN [6]
- ⑦ Algèbre linéaire numérique, Allaire [7]
- ⑧ M2G2, Germani-Caldero [8]