Cadre : on se place sur un IK-ev de dimension finie

(J) $Ax = b$ avec $A \in \mathcal{M}_{p,n}(K)$, $b \in \mathbb{R}^p$ avec $x \in \mathbb{R}^n$ l'inconnue de l'équation

I - Résultats d'existence et solutions

0 - Définitions générales

Considérons un système linéaire à p équations à n inconnues :

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n = b_1 \\ \vdots \\ a_{p1}x_1 + a_{p2}x_2 + \dots + a_{pn}x_n = b_p \end{cases} \quad (1) \quad \text{avec } a_{ij}, b_{ij} \in \mathbb{K}.$$

déf 1 : on appelle solution tout vecteur $x = (x_1 .. x_n) \in \mathbb{K}^n$ dont les composantes x_i satisfont toutes les équations

déf 2 : le système est compatible s'il admet au moins une solution

déf 3 : Soient E et F des IK-ev et $f \in \mathcal{L}(E,F)$.

on appelle rang de f l'entier $\dim(\mathrm{Im}\, f)$ noté $rg\, f$

Théorème 4 (du rang) : Soit E un IK-ev, F un IK-ev et $f \in \mathcal{L}(E,F)$

on a $\dim(E) = \dim(\mathrm{Ker}\, f) + \dim(\mathrm{Im}\, f) = \dim(\mathrm{Ker}\, f) + rg(f)$

Expression Matricielle du système

Soient $A = \| a_{ij} \| = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{p1} & \cdots & a_{pn} \end{pmatrix} \in \mathcal{M}_{p,n}(\mathbb{K})$ $B = \begin{pmatrix} b_1 \\ \vdots \\ b_p \end{pmatrix} \in \mathcal{M}_{p,1}(\mathbb{K})$

et $X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathcal{M}_{n,1}(\mathbb{K})$

Le système (1) peut s'écrire sous forme matricielle $AX = B$

Expression vectorielle du système

Notons $\vec{c_1} .. \vec{c_n}$ les vecteurs colonnes de A

$\vec{c_1} = (a_{11} \cdots a_{p1}) \in \mathbb{K}^p$ ..... $\vec{c_n} = (a_{1n} \cdots a_{pn}) \in \mathbb{K}^p$

on a $x_1\vec{c_1} = (a_{11}x_1 + \dots + a_{p1}x_1)$

$\vdots$

$x_n\vec{c_n} = (a_{1n}x_n + \dots + a_{pn}x_n)$

si $\vec{b} = (b_1 .. b_p) \in \mathbb{K}^p$ le système peut s'écrire $x_1\vec{c_1} + .. + x_n\vec{c_n} = \vec{b}$

1 - Systèmes de Cramer

déf 5 : On appelle système de Cramer un système linéaire dont la matrice A est carrée et inversible

Théorème 6 (Cramer)

Soit $A = \| a_{ij} \| = \| \vec{c_1} \dots \vec{c_n} \|$ et $\det(A) \neq 0$

Un système de Cramer $\begin{cases} a_{11}x_1 + \dots + a_{1n}x_n = b_1 \\ \vdots \\ a_{n1}x_1 + \dots + a_{nn}x_n = b_n \end{cases}$ admet toujours une et une seule solution quel que soit le vecteur $\vec{b} = (b_1 \dots b_n)$ solution donnée par les formules de Cramer : $x_i = \dfrac{\det \| \vec{c_1} \dots \vec{c_{i-1}}, b, \vec{c_{i+1}}, \dots \vec{c_n} \|}{\det(A)}$

Complexité : $O(n(n+1)! - 1)$

2 - Généralisation : le système de Rouché-Fontené

Considérons un système de p équations à n inconnues de rang r

$$\begin{cases} a_{11}x_1 + \dots + a_{1r}x_r + \dots + a_{1n}x_n = b_1 \\ \vdots \\ a_{r1}x_1 + \dots + a_{rr}x_r + \dots + a_{rn}x_n = b_r \\ \vdots \\ a_{p1}x_1 + \dots + a_{pr}x_r + \dots + a_{pn}x_n = b_p \end{cases} \quad (2)$$

Prop 7 : Il existe un déterminant Δ d'ordre r non nul extrait de A

déf 8 : Δ s'appelle le déterminant principal de (2)

- les équations dont les indices sont ceux des lignes de Δ s'appellent les équations principales

Les inconnues dont les indices sont ceux des colonnes de Δ s'appellent les inconnues principales

déf 9 : On appelle déterminant caractéristique les déterminants d'ordre $r+1$ de la forme

$$\left|\begin{array}{c|c} (a_{ij})_{i\in I, j\in J} & (b_i)_{i\in I} \\ \hline (a_{kj})_{j\in J} & b_k \end{array}\right| \quad \text{avec } k \notin J$$

Rq : Les déterminants caractéristiques n'existent que si $r<p$ et il y en a alors $p-r$

Théorème 10 (Rouché _ Fontené)

Le système (2) admet des solutions ssi $p=r$ ou les $p-r$ déterminants caractéristiques sont nuls

Le système est alors équivalent au système des équations principales, les inconnues principales étant déterminées par un système de Cramer à l'aide des inconnues non principales

④ - Utilisation pour des problèmes de géométrie plane complexe

déf 11 : Une courbe plane est une courbe qui est entièrement contenue dans un (unique) plan et qui est identifiable à une fonction continue $\alpha : I \to \mathbb{R}^2$ avec $I \subset \mathbb{R}$

déf 12 : L'image d'une courbe est appelée support de la courbe

Ex 13 : • Les courbes quadriques ou coniques de degré 2

• Les courbes cubiques de degré 3

Théorème 14 (Bezout faible)

Soient ϕ une courbe plane de degré m et ψ une courbe plane degré n

Si ϕ et ψ possèdent plus de $mn+1$ points alors les courbes ont une composante commune

Corollaire 15 En reprenant les notation du théorème on a :

• si ϕ ~~est~~ une droite et n'il y a $n+1$ points en commun alors ϕ est une composante de ψ.

• si ϕ est une conique et n'il y a plus de $2n+1$ points en commun alors :

• ϕ est une composante de ψ

ou • ϕ est produit de 2 droites dont l'une est aussi une composante de ψ.

Prop 16 : 5 points du plan affine déterminent une unique conique ssi il n'existe pas de droite passant par quatre de ces points

Théorème 17 Soient ϕ et ψ deux cubiques projectives, sans composante commune qui se coupent en 9 points distincts. alors toute cubique qui passe par 8 de ces 9 points passe par le neuvième. — Dév ①

II - Systèmes échelonnés

① - Action de groupe aux services des matrices

déf 18 : On considère l'action par translation à gauche de $Gl_p(K)$ sur $\mathcal{M}_{pn}(K)$ par

$$Gl_p(K) \times \mathcal{M}_{pn}(K) \to \mathcal{M}_{pn}(K)$$
$$P, A \to PA$$

a) Systèmes échelonnés

déf 19 : On appelle pivot d'une ligne non nulle le coefficient non nul le plus à gauche.

déf 20 : Une matrice est dite échelonnée en ligne lorsque :

• si une ligne est nulle toutes les lignes suivantes sont nulles

• Le pivot d'une ligne est strictement plus à droite que les pivots des lignes précédentes.

déf 21 : Une matrice est dite échelonnée réduite si de plus tous les pivots sont égaux à 1 et les pivots sont les seuls coefficients non nuls de leur colonne

b) Algorithme de Gauss

Théorème 22 : L'ensemble des matrices échelonnées réduites forment un système d'invariants pour l'action décrite précédemment.

- Algorithme de Gauss-Jordan

$r = 0$

Pour j de 1 jusqu'à m

Rechercher $\max(|A[i,j]|, r+1 \leq i \leq m)$

Si $A[k,j] \neq 0$ alors

$r = r+1$

diviser la ligne k par $A[k,j]$

Échanger les lignes k et r

Pour i de 1 jusqu'à p

Si $i \neq r$ alors

Soustraire à la ligne i la ligne r multipliée par $A[i,j]$

Fin Si

Fin Pour

Fin Si

Fin Pour

Fin Gauss-Jordan

Complexité : $O(n^3)$

- Application 23

1) Calcule du rang d'une matrice
2) Calcule d'inverse
3) Recherche d'un système d'équation d'un sev défini par une famille génératrice
4) Générateurs de $GL_n(\mathbb{R})$: dilatations et transvections
5) Générateurs de $SL_n(\mathbb{R})$: transvections

III - Analyse matricielle

① Stabilité et conditionnement

déf 24 : Le nombre $cond(A) = \|A^{-1}\| \|A\|$ est appelé conditionnement de A ; il dépend donc du choix des normes

Ex 25 : La matrice de Hilbert d'ordre n est la matrice H_n dont le coefficient d'indice i,j vaut $\frac{1}{i+j-1}$

$H_2 = \begin{pmatrix} 1 & 1/2 \\ 1/2 & 1/3 \end{pmatrix}$; $Cond_\infty(H_2) = 27$

Prop 26 : Pour une matrice symétrique inversible A, le conditionnement calculé relativement à la norme euclidienne est donné par la formule $Cond_2(A) = \rho(A)\rho(A^{-1})$ avec ρ rayon spectral de A.

②. Méthodes itératives

Méthode de Ricard : Pour résoudre $AX = B$ on utilise l'idée suivante :
on écrit $A = M - N$ avec M inversible il vient alors $AX = B \Leftrightarrow X = RX + C$
avec $R = M^{-1}N$ et $C = M^{-1}B$
on considère le système $X = RX + C$ comme un problème de point fixe que l'on résout en posant $\begin{cases} X_0 = 0 \\ X_{k+1} = RX_k + C \end{cases}$

Dans la suite nous considérons une matrice symétrique définie positive :

Théorème 27 : (Formulation Variationnelle)
$\bar{x}$ est solution du système $AX = B$ si il minimise la fonction $f : x \to \frac{1}{2}\langle Ax, x\rangle - \langle b, x\rangle$

Algorithme du gradient

Soit x_0 donné et $r_0 = b - Ax_0$. On se fixe un seuil de tolérance α

Tant que $\frac{\|r_n\|}{\|r_0\|} \geq \alpha$ faire

$$\alpha_{n+1} = \frac{\|r_n\|^2}{\langle Ar_n, r_n\rangle} \quad (1)$$

$$x_{n+1} = x_n + \alpha_{n+1} r_n \quad (2)$$

$$r_{n+1} = b - Ax_{n+1} \quad (3)$$

Théorème 28 L'algorithme du gradient converge vers la solution $\bar{x}$ du système $Ax = b$ et pour tout $n \in \mathbb{N}$ on a :

$$\|x_n - \bar{x}\| \leq \left(\frac{cond(A)-1}{cond(A)+1}\right)^n \sqrt{cond(A)} \|x_0 - \bar{x}\|$$

Dév ②

Références :

- H_2G_2 : Philippe Caldero, Germoni
- Bourdon Algèbre :
- Grifone : " Algèbre linéaire "
- Ramis, Warusfel : " Mathématiques : Tout-en-un pour la Licence 2 " p 931

Théorème de Cayley - Bacharach pour les cubiques

Références : • polycopié de M. Coste "système d'équations linéaires et courbes passant par des points fixés".
• P. Boyer, "algèbres et géométries"

Résultats préliminaires :

prop ① 5 points du plan affine determinent une unique conique si, et seulement s'il n'y en a pas 4 d'alignés.

Thm (Bézout Faible) : soient Φ une courbe de degré m et Ψ une courbe de degré n qui ont au moins $mn+1$ pts en commun. Alors Φ et Ψ ont une composante en commune.

Corollaire : ⓐ Dans le cas particulier où Ψ est une droite, si Φ et Ψ ont $m+1$ pts en commun, alors Ψ est une composante de Φ ;

ⓑ si Ψ est une conique et si il y a $2m+1$ pts en commun entre Φ et Ψ, alors
→ Ψ est le produit de 2 droites, alors l'une d'elles est une composante commune à Ψ et Φ ;
→ sinon Ψ est une composante de Φ.

Résultat principal :

Thm : soient Φ et Ψ deux cubiques projectives, sans composante commune, qui se coupent en neuf points distincts. Toute cubique qui passe par 8 points passe par le neuvième.

Demo : une cubique est la donnée de 10 coefficient et la condition de passer par 9 points fixés s'exprime par un système de neufs équations linéaires homogènes en ces 10 coefficients, on note S la matrice du système (S).

• La matrice formé par les 9 intersections forme un système d'équation de rang au plus 8 : l'espace des solutions du système homogène contient le plan issue de Φ et Ψ : $\{\lambda\Phi+\mu\Psi,\ \lambda,\mu\in\mathbb{C}^2\}$. Par le théorème du rang, on a alors :

$$rg(S) = \dim(\mathbb{C}^{10}) - \ker(S) \leq 10-2 = 8$$

• Prenons un sous-système issue de (S) formé par les équations de 8 points d'intersection. [But : montrer que le rang(S) est exactement 8]

lemme ① Parmis les 9 points d'intersection, il n'y en a pas 4 d'alignés.
② Il n'y en a pas 7 sur une conique.

démo (lemme)

① Supposons qu'il y en a 4 d'alignés. On peut considerer la droite qui les contient. Ayant $4 \geq 3\times1+1$ pts en commun, le Théorème de Bezout (Corollaire pour une cubique et une droite) en commun.

La seule composante de la droite étant elle-même, on en déduit que Φ et Ψ ont une composante en commun, contradiction.

② Supposons qu'il existe une conique Γ contenant 7 points. Deux cas :

→ si Γ est dégénérée (càd produit de 2 droites) alors une des 2 droites contient 4 pts parmi les 7, et on est ramené au cas précédent ;

→ sinon Γ est la composante commune à Φ et à Ψ

(on utilise le corollaire dans le 2^{nd} cas et $7 \geq 3 \times 2 + 1$)

Prenons A un des 8 points. Puisqu'il n'y a pas 4 pts alignés, □ lemme

on peut choisir B parmis les 7 pts restant, puis C hors de (AB), et enfin D hors des droites (AB) ; (AC) et (BC).

Ainsi les cotés du triangle BCD ne passent pas par A.

De plus, on peut considérer 4 coniques passant par 5 points :

- Γ_A qui passe par A, E, F, G, H ;
- Γ_B ———— B ———— ;
- Γ_C ———— C ———— ;
- Γ_D ———— D ———— ;

ceci se prouve grâce à la proposition ①.

Les coniques Γ_B, Γ_C et Γ_D ne sont pas toutes identiques, sinon elles passeraient toutes par 7 pts du système, ce qui contredit lemme ②. Ainsi il y a au moins 1 de ces 3 coniques distinctes de Γ_A, ~~et puisque~~ disons Γ_D, et puisque Γ_A et Γ_D ont déjà 4 points en commun, on sait que Γ_D ne passe pas par A.

Enfin, en considérant le produit Γ_D par (BC), on obtient une cubique passant par 7 pts mais pas par le $8^{ème}$ point. D'où l'indépendance des équations du sous-système considéré, et finalement, (S) est de rang 8.

En reformulant ceci, on aboutit bien au résultat suivant :

toute cubique dont le vecteur des coefficient satisfait 8 des 9 équations de (S), satisfait automatiquement la $9^{ème}$.

ie si la cubique passe par 8 des 9 points, elle passe par le $9^{ème}$.

□ Thm

Méthode du gradient à pas optimal

Ce développement est extrait du *Cours de mathématiques pures et appliquées, volume 1* de Ramis. Warusfel et Moulin. La preuve de l'inégalité de Kantorovitch se trouve dans les oraux X-ENS.
Soit A une matrice symétrique définie positive. On utilise sans démonstration le résultat suivant : $\overline{x}$ est la solution du système $Ax = b$ si et seulement s'il minimise la fonction $f : x \mapsto \frac{1}{2}\langle Ax, x\rangle - \langle b, x\rangle$. Il s'agit d'appliquer l'algorithme de minimisation du gradient à pas optimal à la fonction f : on construit la suite (x_n) en posant

$$x_{n+1} = x_n + \alpha_{n+1} r_n$$

où $r_n = -\nabla f(x_n) = b - Ax_n$ (direction de plus grande pente) et α_{n+1} est le pas optimal, c'est-à-dire que

$$f(x_n + \alpha_{n+1} r_n) = \min_{\alpha > 0} f(x_n + \alpha r_n).$$

On vérifie que $\alpha_{n+1} = \frac{\|r_n\|^2}{\langle Ar_n, r_n\rangle}$, défini tant que $x_n \neq \overline{x}$.
On rappelle l'algorithme :

Soit x_0 donné et $r_0 = b - Ax_0$. On se fixe un seuil de tolérance TOL.
Tant que $\frac{\|r_n\|}{\|r_0\|} \geqslant TOL$, faire

$$\alpha_{n+1} = \frac{\|r_n\|^2}{\langle Ar_n, r_n\rangle} \tag{1}$$

$$x_{n+1} = x_n + \alpha_{n+1} r_n \tag{2}$$

$$r_{n+1} = b - Ax_{n+1}. \tag{3}$$

Théorème. *L'algorithme du gradient converge vers la solution $\overline{x}$ du système $Ax = b$ et, pour tout $n \in \mathbb{N}$*

$$\|x_n - \overline{x}\| \leqslant \left(\frac{\mathrm{Cond}(A) - 1}{\mathrm{Cond}(A) + 1}\right)^n \sqrt{\mathrm{Cond}(A)}\|x_0 - \overline{x}\|.$$

Démonstration.

Lemme. *Inégalité de Kantorovitch Soit $A \in \mathcal{M}_N(\mathbb{R})$ symétique définie positive, de valeurs propres minimale et maximale λ_m et λ_M. Alors*

$$\forall x \in \mathbb{R}, \quad \frac{\|x\|^4}{\|x\|_{A^{-1}}^2 \|x\|_A^2} \geqslant 4\frac{\lambda_m \lambda_M}{(\lambda_M + \lambda_m)^2}$$

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

Démonstration. Il suffit de montrer cette inégalité pour x de norme 1, ce qu'on supposera dans la suite. A est symétrique donc il existe une base $(e_1, \cdots, e_n)$ orthonormale de vecteurs propres de A. On note $x = \sum x_i e_i$ et $0 < \lambda_1 \leqslant \cdots \leqslant \lambda_n$ les valeurs propres de A.

$$\begin{aligned}\langle Ax, x\rangle \langle A^{-1}x, x\rangle &= \sum_{i=1}^n \lambda_i x_i^2 \sum_{i=1}^n \frac{1}{\lambda_i} x_i^2 = \frac{\lambda_1}{\lambda_n} \sum_{i=1}^n \frac{\lambda_i}{\lambda_n} x_i^2 \sum_{i=1}^n \frac{\lambda_1}{\lambda_i} x_i^2 \\ &\leqslant \frac{\lambda_1}{4\lambda_n} \left(\sum_{i=1}^n \left(\frac{\lambda_i}{\lambda_n} + \frac{\lambda_1}{\lambda_i}\right) x_i^2\right)^2.\end{aligned}$$

Après étude de la fonction $x \mapsto x + \frac{1}{x}$ qui atteint son maximum $1 + \frac{\lambda_n}{\lambda_1}$ en λ_1 et en λ_n, on en déduit que

$$\langle Ax, x\rangle \langle A^{-1}x, x\rangle \leqslant \frac{\lambda_1}{4\lambda_n} \left(\left(1 + \frac{\lambda_n}{\lambda_1}\right) \sum_{i=1}^{n} x_i^2 \right)^2 = \leqslant \frac{\lambda_1}{4\lambda_n} \left(1 + \frac{\lambda_n}{\lambda_1}\right)^2 = \frac{(\lambda_1 + \lambda_n)^2}{4\lambda_1\lambda_n}.$$

Il suffit d'inverser cette formule pour obtenir l'inégalité escomptée.

□

Pour la preuve générale, remarquons que

$$f(x) = \frac{1}{2} \langle Ax, x\rangle - \langle b, x\rangle = \frac{1}{2} \langle Ax, x\rangle - \langle A\overline{x}, x\rangle = \frac{1}{2} \langle A(x - \overline{x}), x - \overline{x}\rangle - \frac{1}{2} \langle A\overline{x}, \overline{x}\rangle ,$$

c'est-à-dire $f(x) = \frac{1}{2}\|x - \overline{x}\|_A - \frac{1}{2} \langle A\overline{x}, \overline{x}\rangle$. L'algorithme garantit la décroissance de f à chaque itération. Reste à voir qu'on a en fait une contractance de f :

$$\frac{\|x_{n+1} - \overline{x}\|_A^2}{\|x_n - \overline{x}\|_A^2} = \frac{\|x_n + \alpha_{n+1} r_n - \overline{x}\|_A^2}{\|x_n - \overline{x}\|_A^2} = \frac{\|x_n - \overline{x}\|_A^2 + \alpha_{n+1}^2 \|r_n\|_A^2 + 2 \langle A(x_n - \overline{x}), \alpha_{n+1} r_n\rangle}{\|x_n - \overline{x}\|_A^2} \tag{4}$$

$$= 1 + \frac{\|r_n\|^4/\|r_n\|_A^2 + 2\alpha_{n+1} \langle Ax_n - b, b - Ax_n\rangle}{\|x_n - \overline{x}\|_A^2} \tag{5}$$

$$= 1 - \frac{\|r_n\|^4/\|r_n\|_A^2}{\|x_n - \overline{x}\|_A^2} = 1 - \frac{\|r_n\|^4}{\|r_n\|_A^2 \|A^{-1} r_n\|_A^2} = 1 - \frac{\|r_n\|^4}{\langle Ar_n, r_n\rangle \langle A^{-1} r_n, r_n\rangle}. \tag{6}$$

On peut alors appliquer l'inégalité de Kantorovitch qui donne

$$\|x_{n+1} - \overline{x}\|_A^2 \leqslant \left(1 - 4\frac{\lambda_1\lambda_n}{(\lambda_1 + \lambda_n)^2}\right) \|x_n - \overline{x}\|_A^2 \tag{7}$$

$$\leqslant \left(\frac{\lambda_1 - \lambda_n}{\lambda_1 + \lambda_n}\right)^2 \|x_n - \overline{x}\|_A^2. \tag{8}$$

$$\leqslant \left(\frac{1 - \mathrm{Cond}(A)}{1 + \mathrm{Cond}(A)}\right)^2 \|x_n - \overline{x}\|_A^2. \tag{9}$$

Ceci montre la convergence linéaire de (x_n) vers $\overline{x}$ pour $\|\cdot\|_A$.
Or, pour tout x, $\leqslant \lambda_1\|x\|^2 \langle Ax, x\rangle = \|x\|_A^2 \leqslant \lambda_n\|x\|^2$, c'est-à-dire

$$\|x_n - x\| \leqslant \left(\frac{1 - \mathrm{Cond}(A)}{1 + \mathrm{Cond}(A)}\right)^n \sqrt{\lambda_n/\lambda_1}\|x_0 - \overline{x}\| = \left(\frac{1 - \mathrm{Cond}(A)}{1 + \mathrm{Cond}(A)}\right)^n \sqrt{\mathrm{Cond}(A)}\|x_0 - \overline{x}\|,$$

ce qu'il fallait montrer.

□