MÉTHODES ITÉRATIVES DE RÉSOLUTION D'UN SYSTÈME LINÉAIRE

Ciarlet p102

<u>Prérequis:</u> norme subordonnée, formule $\|\cdot\|_\infty$, rayon spectral.

Soit $A \in GL_n(\mathbb{R})$, $b \in \mathbb{R}^n$. On étudie le système $Ax = b$.

Def: si $(M,N) \in GL_n(\mathbb{R}) \times \mathcal{M}_n(\mathbb{R})$ est tel que $A = M - N$, on dit que la <u>méthode itérative</u> associée à (M,N) <u>converge</u> si pour tout $u_0 \in \mathbb{C}^n$, la suite de premier terme u_0 et définie par $\forall k \in \mathbb{N}\ u_{k+1} = M^{-1}(Nu_k + b)$ converge.

Thm: La méthode itérative associée à (M,N) converge ssi $\rho(M^{-1}N) < 1$.

On commence par montrer un lemme

Lemme: Soit $A \in \mathcal{M}_n(\mathbb{C})$. Soit $\varepsilon > 0$. Alors il existe une norme subordonnée $|||\cdot|||$ tq $|||A||| \leq \rho(A) + \varepsilon$.

Dem: Comme $A \in \mathcal{M}_n(\mathbb{C})$, A est trigonalisable : il existe $P \in GL_n(\mathbb{C})$ et $T = (t_{ij})$ triangulaire supérieure telles que $A = PTP^{-1}$.

- On note $(e_1, \dots, e_n)$ la base canonique de $\mathbb{C}^n$.
  Soit $\delta > 0$, on pose $e'_i = \delta^{i-1} e_i$ une <u>nouvelle base de $\mathbb{C}^n$</u>.
  On pose $D_\delta = \operatorname{diag}(1, \delta, \dots, \delta^{n-1})$ la matrice de changement de base de (e_i) vers (e'_i). $= \mathrm{Mat}_{B'B}(id)$

Soit $j \in [\![1, n]\!]$,

$$Te'_j = \delta^{j-1} Te_j = \delta^{j-1} \sum_{i=1}^{j} t_{ij} e_i = \delta^{j-1} \sum_{i=1}^{j} t_{ij}\, \delta^{1-i} e'_i$$

$$= \delta^{j-i} \sum_{i=1}^{j} t_{ij} e'_i =$$

Donc
$$T_\delta = D_\delta^{-1} T D_\delta = \begin{pmatrix} t_{11} & \delta t_{12} & \cdots & \delta^{n-1} t_{1n} \\ & \ddots & & \vdots \\ & & \ddots & \delta t_{n-1\,n} \\ 0 & & & t_{nn} \end{pmatrix}$$
matrice de l'endomorphisme canonique associé à T, dans la base (e'_i)

- Pour $x \in \mathbb{C}^n$, <u>on introduit</u> $\|\cdot\| : x \mapsto \|(PD_\delta)^{-1} x\|_\infty$. C'est bien <u>une norme</u>.
  On note $|||\cdot|||$ la norme subordonnée.

Soit $B \in \mathcal{M}_n(\mathbb{C})$, $|||B||| = \sup\limits_{x \in \mathbb{C}^n \setminus \{0\}} \dfrac{\|(PD_\delta)^{-1} Bx\|_\infty}{\|(PD_\delta)^{-1} x\|_\infty}$.

Or $|||(PD_\delta)^{-1} B PD_\delta|||_\infty = \sup\limits_{x \in \mathbb{C}^n \setminus \{0\}} \dfrac{\|(PD_\delta)^{-1} B PD_\delta x\|_\infty}{\|x\|_\infty}$ ← faire ds ce sens.

$y = \underset{\text{inversible}}{\underline{PD_\delta}}\, x$ $\quad = \sup\limits_{y \in \mathbb{C}^n \setminus \{0\}} \dfrac{\|(PD_\delta)^{-1} By\|_\infty}{\|(PD_\delta)^{-1} y\|_\infty} = |||B|||$

Or si $C = (c_{ij}) \in \mathcal{M}_n(\mathbb{C})$, on a $|||C|||_\infty = \sup\limits_{1 \leq i \leq n} \sum\limits_{j=1}^{n} |c_{ij}|$.

- On a donc $|||A||| = |||(PD_\delta)^{-1} A PD_\delta|||_\infty = |||D_\delta^{-1} T D_\delta|||_\infty = |||T_\delta|||_\infty$ par définit° de T et T_δ

$$= \sup_{1 \leq i \leq n} \sum_{j=i}^{n} |\delta^{j-i} t_{ij}| \leq \underbrace{\sup_{1 \leq i \leq n} |t_{ii}|}_{= \rho(A)} + \sup_{1 \leq i \leq n-1} \sum_{j=i+1}^{n} \delta^{j-i} |t_{ij}|$$

[illegible] en particulier sans le sup

On choisit $\delta > 0$ tq $\forall\, 1 \le i \le n-1 \quad \sum_{j=i+1}^{n} \delta^{j-i} |t_{ij}| < \varepsilon$.

d'où $|||A||| \le \rho(A) + \varepsilon$.

Dém thm: Soit $u \in \mathbb{C}^n$ tq $Au = b$, c'est à dire, comme $A = M - N$, $Mu = Nu + b$.

On pose $e_k = u_k - u$ l'erreur. Alors

$e_{k+1} = u_{k+1} - u = M^{-1}(Nu_k + b) - M^{-1}(Nu + b) = M^{-1}N e_k$ ← convergence géométrique.

Par récurrence immédiate, on a donc $\forall k \in \mathbb{N} \quad e_k = (M^{-1}N)^k e_0$. On a alors deux cas.

→ Si $\rho(M^{-1}N) < 1$, on pose $\varepsilon = \dfrac{1 - \rho(M^{-1}N)}{2} > 0$ et d'après le lemme, il existe une norme subordonnée $|||\cdot|||$ tq $|||M^{-1}N||| \le \rho(M^{-1}N) + \varepsilon = \dfrac{1 + \rho(M^{-1}N)}{2} < 1$.

Donc, pour la norme $\|\cdot\|$ associée, on a pour $k \in \mathbb{N}$

$\|e_k\| \le |||M^{-1}N|||^k \|e_0\| \xrightarrow[k \to +\infty]{} 0$ car $|||M^{-1}N||| < 1$. donc $\lim_{k \to +\infty} e_k = 0$.

norme d'algèbre car subordonnée

Donc (u_k) converge vers u.

→ Si $\rho(M^{-1}N) \ge 1$. Soit λ une valeur propre, a priori complexe, de $M^{-1}N$ de module $\rho(M^{-1}N)$. Soit v un vecteur propre associé non nul.

Comme $\forall k \in \mathbb{N} \quad (M^{-1}N)^k v = \lambda^k v$, on prend $u_0 = u + v$ alors $e_0 = v$

et $\|e_k\| = \|(M^{-1}N)^k v\| = |\lambda|^k \|v\| \underset{k \to +\infty}{\not\to} 0$ car $|\lambda| \ge 1$.

Donc la méthode itérative ne converge pas pour $u_0 = u + v$.

L.

Quelques cas particuliers de méthodes itératives: On note $D = \mathrm{diag}(A)$, $E = -A_{inf}$, $F = -A_{sup}$. $A = D - E - F$

→ Méthode de Jacobi: $M = D$ et $N = D - A = E + F$. On note $J = \underbrace{D^{-1}(D-A)}_{M^{-1}N}$
si D est inversible !

→ Méthode de Gauss-Seidel: $M = D - E$ et $N = F$. On note $\mathcal{L}_1 = \underbrace{(D-E)^{-1}F}_{M^{-1}N}$.
si D inversible !

→ Méthode de relaxation: pour $\omega \in \mathbb{R}^{+*}$, $M = \dfrac{D}{\omega} - E$ et $N = \dfrac{1-\omega}{\omega} D + F$.
si D inversible

On pose $\mathcal{L}_\omega = M^{-1}N$

Prop: Si A est une matrice tridiagonale, $\rho(\mathcal{L}_1) = \rho(J)^2$. La méthode de Gauss Seidel a donc une vitesse de convergence double de celle de la méthode de Jacobi.

Dém: Pour $\mu \neq 0$, on note $B(\mu) = \begin{pmatrix} b_1 & \mu^{-1}c_2 & & 0 \\ \mu a_2 & \ddots & \ddots & \\ & \ddots & \ddots & \mu^{-1}c_n \\ 0 & & \mu a_n & b_n \end{pmatrix}$ où ~~B = ...~~.

Si on note $Q(\mu) = \mathrm{diag}(\mu, \mu^2, \dots, \mu^n)$ alors $B(\mu) = Q(\mu) B(1) Q(\mu)^{-1}$. matrice de chgt de base comme T et TS.

donc $\det B(\mu) = \det B(1) \quad \forall \mu \neq 0$.

Les valeurs propres de J sont les racines du polynôme caractéristique

$$P_J(X) = \det(J - XI_n) = \det(D^{-1}(\underbrace{D-A}_{E+F}) - XI_n) = \det(D^{-1}) \underbrace{\det(E+F-XD)}_{q_J(X)}$$

De même, les valeurs propres de $\mathcal{L}_1$ sont les racines du polynome

$$P_{\mathcal{L}_1}(X) = \det(\mathcal{L}_1 - XI_n) = \det((D-E)^{-1}F - XI_n) = \det((D-E)^{-1}) \underbrace{\det(F - XD + XE)}_{q_{\mathcal{L}_1}}$$

Or $\lambda \in \mathbb{C}^*$

$$q_{\mathcal{L}_1}(\lambda^2) = \det(F - \lambda^2 D + \lambda^2 E) = \lambda^n \det(\underbrace{\lambda^{-1}F - \lambda D + \lambda E}_{= -(\lambda D - E - F)(\lambda) \text{ à la diagonale près}})$$

$$= \lambda^n \det(\overbrace{-\lambda D + F + E}^{-(\lambda D - E - F)(1)})$$

$$= \lambda^n q_J(\lambda)$$

Donc les valeurs propres non nulles de $\mathcal{L}_1$ sont les carrés des valeurs propres non nulles de J. Ce qui permet de conclure.

Bonus:

Prop Le rayon spectral de $\mathcal{L}_\omega$ est ~~strictement~~ supérieur à $|\omega - 1|$. La méthode de relaxation ne peut donc converger que si $\omega \in]0, 2[$.

Dém:

On a $$\det(\mathcal{L}_\omega) = \frac{\det\left(\frac{1-\omega}{\omega} D + F\right)}{\det\left(\frac{D}{\omega} - E\right)} = \frac{\prod_{i=1}^n \frac{1-\omega}{\omega} a_{ii}}{\prod_{i=1}^n \frac{a_{ii}}{\omega}} = (1-\omega)^n.$$

On note $\lambda_1, \dots, \lambda_n$ les valeurs propres avec multiplicités de $\mathcal{L}_\omega$, alors.

$$\rho(\mathcal{L}_\omega)^n \geq \left|\prod_{i=1}^n \lambda_i\right| = |\det \mathcal{L}_\omega| = |1-\omega|^n.$$

donc $\rho(\mathcal{L}_\omega) \geqslant |\omega - 1|$.

Si $\omega \geq 2$, $\rho(\mathcal{L}_\omega) \geqslant 1$ donc la méthode itérative ne converge pas.