Séance 3 : normes de matrices et conditionnement

1) Un exemple important d'application : la matrice du laplacien discrétisé

Pour rechercher la déformation d'une structure (poutre, corde) fixée à ses deux extrémités, on est ramené à la résolution de l'EDO :

$$-u''(x) + c\,u(x) = f(x),\ x \in [0,1]$$

avec les conditions aux limites :

$$u(0) = u(1) = 0$$

(où f : force extérieure et $c > 0$)

L'existence et l'unicité de la solution peuvent être démontrées (voir séance à venir)

Par contre, il n'existe pas forcément de solution explicite.

On doit recourir au calcul d'une solution approchée, déterminée aux points $x_i = i/n$ $(0 \leq i \leq n)$;

celle-ci est donnée par la résolution du système suivant :

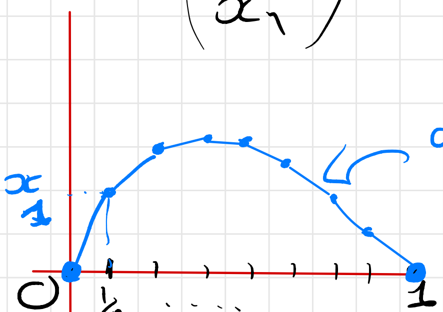
$$\boxed{(n^2 A + cI)\,X = F}$$ où

$$A = \begin{pmatrix} 2 & -1 & & & 0 \\ -1 & \ddots & \ddots & & \\ & \ddots & \ddots & \ddots & \\ & & \ddots & \ddots & -1 \\ 0 & & & -1 & 2 \end{pmatrix} \in \mathcal{S}_{n-1}(\mathbb{R})$$

matrice du laplacien discrétisé

$$F = \begin{pmatrix} f(\frac{1}{n}) \\ \vdots \\ f(\frac{n-1}{n}) \end{pmatrix}$$

$$\text{et } X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \approx \begin{pmatrix} u(\frac{1}{n}) \\ \vdots \\ u(\frac{n-1}{n}) \end{pmatrix}$$



approximation de u

2) Normes de matrices

Ici $\mathcal{M}_n(\mathbb{R})$ est considéré comme un espace vectoriel de dimension finie (égale à n^2).

Déf : soit E un e.v. On appelle norme sur E, toute application $N : E \to \mathbb{R}_+$

rq :

(i) $\forall x \in E$, $N(x) \geq 0$ et $N(x) = 0 \iff x = 0$ (séparation)

(ii) $\forall \lambda \in \mathbb{R}$, $\forall x \in E$, $N(\lambda x) = |\lambda| N(x)$ (homogénéité)

(iii) $\forall (x,y) \in E^2$, $N(x+y) \leq N(x) + N(y)$ (inégalité triangulaire)

Exemples :

* Sur $\mathbb{R}^n$, on dispose des normes $1, 2, p, +\infty$:

$$\|x\|_1 = \sum_{i=1}^{n} |x_i|$$

$$\|x\|_2 = \sqrt{\sum_{i=1}^{n} |x_i|^2}$$

$$\|x\|_p = \sqrt[p]{\sum_{i=1}^{n} |x_i|^p}$$

$$\|x\|_\infty = \underset{1 \leq i \leq n}{\mathrm{Max}} |x_i|$$

* Sur $\mathcal{M}_n(\mathbb{R})$, on peut disposer du même type de normes :

$$\|A\|_1 = \sum_{i,j=1}^{n^2} |A_{ij}|$$

On le résultat classique suivant :

Théorème : (en dimension finie) toutes les normes sont équivalentes :

$$c_1 N_1(x) \leq N_2(x) \leq c_2 N_1(x)$$

(c_1 et $c_2 > 0$ indépendants de x)

* Dans le cas présent, $\mathcal{M}_n(\mathbb{R})$ peut être muni d'une structure d'algèbre.

Def : si N est une norme sur $\mathcal{M}_n(\mathbb{R})$, on dit que N est une norme d'algèbre si $\forall (A, B) \in \mathcal{M}_n(\mathbb{R})$, $N(AB) \leq N(A)N(B)$

Remarque : voir TD 3 pour vérifier si les normes $\| \ \|_1$, $\| \ \|_2$ et $\| \ \|_\infty$ sont des normes d'algèbre sur $\mathcal{M}_n(\mathbb{R})$.

* Dans le cas présent, $\mathcal{M}_n(\mathbb{R})$ s'identifie aussi à $\mathcal{L}(\mathbb{R}^n)$.

Def soit N une norme sur $\mathbb{R}^n$,

On construit la norme $||| \ |||_N$ suivante sur $\mathcal{M}_n(\mathbb{R})$:

$$\boxed{|||A|||_N = \sup_{x \neq 0} \left(\frac{N(Ax)}{N(x)} \right)}$$

On parle de norme subordonnée par rapport à N.

Il s'agit bien d'une norme sur $\mathcal{M}_n(\mathbb{R})$:

$\rightarrow$ $|||A|||_N = \sup\limits_{\substack{x \in \mathbb{R}^n \\ N(x)=1}} (N(Ax))$ continue, compact

et $|||A|||_N < +\infty$.

$\rightarrow$ separation : si $|||A|||_N = 0$,

$\forall x \in \mathbb{R}^n$, $(x \neq 0)$, $N(Ax) = 0 \Rightarrow A = 0$

$\rightarrow$ homogénéité : OK

$\rightarrow$ IT :

$$N((A+B)x) \leq N(Ax) + N(Bx)$$
$$\leq \sup_{B_N(0,1)} (N(Ax)) + \sup_{B_N(0,1)} (N(Bx))$$

$\Rightarrow |||A+B|||_N \leq |||A|||_N + |||B|||_N$

Exemples : on dispose des normes subordonnées $|||\ |||_1$, $|||\ |||_2$ et $|||\ |||_\infty$.

On remarque que toute norme subordonnée est une norme d'algèbre :

$$|||AB|||_N = \max_{B_N(0,1)} \left(N(ABx) \right)$$

$$\leq |||A|||_N \max_{B_N(0,1)} \left(N(Bx) \right)$$

$$\leq |||A|||_N \, |||B|||_N .$$

Les trois normes ainsi construites sur $\mathcal{M}_n(\mathbb{R})$ possèdent les expressions suivantes :

Théorème : soit $A \in \mathcal{M}_n(\mathbb{R})$, on a

$$\boxed{|||A|||_1 = \max_{1 \leq j \leq n} \left(\sum_{i=1}^{n} |A_{ij}| \right)} \equiv 1$$

(max de la somme absolue des colonnes)

$$\boxed{|||A|||_\infty = \max_{1 \leq i \leq n} \left(\sum_{j=1}^{n} |A_{ij}| \right)} \equiv \infty$$

(max de la somme absolue des lignes)

$$\boxed{|||A|||_2 = \sqrt{\rho({}^tAA)}}$$ où

$\rho(B)$ désigne le rayon spectral de la matrice B :

$$\boxed{\rho(B) = \max \{ |\lambda| , \lambda \in \underbrace{S_{p_{\mathbb{C}}}(B)}_{\text{spectre complexe de } B} \}}$$

(ii) Soit $X \in \mathbb{R}^n$, $\|X\|_\infty = 1$

et

$$\|AX\|_\infty = \underset{1 \le i \le N}{\text{Max}} |(AX)_i|$$

$$= \underset{1 \le i \le n}{\text{Max}} \left(\left| \sum_{j=1}^n A_{ij} X_j \right| \right)$$

$$\le \underbrace{\underset{1 \le i \le n}{\text{Max}} \left(\sum_{j=1}^n |A_{ij}| \right)}$$

$$\Rightarrow |||A|||_\infty \le \quad ''$$

Réciproquement, on note i_0 une ligne telle que :

$\underset{1 \le i \le n}{\text{Max}} \left(\sum_j |A_{ij}| \right)$ est atteint

On note alors $X^* = \begin{pmatrix} x_1 \\ \vdots \\ x_j \end{pmatrix}$ le vecteur

tel que $X_j^* = \text{sgn}(A_{i_0,j})$ (± 1)

On a alors

$$\|AX^*\|_\infty = \underset{1 \le i \le n}{\text{Max}} \left(\left| \sum_{j=1}^n A_{ij} X_j^* \right| \right)$$

$$= \sum_{j=1}^n |A_{i_0,j}|$$

$$= \underset{1 \le i \le n}{\text{Max}} \left(\sum_{j=1}^n |A_{ij}| \right)$$

Comme $\|X^*\| = 1$, on a bien

$$|||A|||_\infty \ge \underset{1 \le i \le n}{\text{Max}} \left(\sum_{j=1}^n |A_{ij}| \right)$$

(pour $||| \ |||_1$: exercice)

Par l'expression de $|||\ |||_2$), on a

$$\|AX\|_2^2 = \langle AX, AX\rangle^2$$

produit scalaire usuel sur $\mathbb{R}^n$

$$= \langle \underbrace{{}^tAA}_{\text{symétrique, positive}} X, X\rangle^2$$

On note $0 \leq \lambda_1 \leq \dots \leq \lambda_n$ les valeurs propres (positives) de ${}^tAA$.

On a donc

$$\|AX\|_2^2 = \sum_{i=1}^{n} \lambda_i y_i^2$$ après passage dans une b.o.n adaptée à ${}^tAA$

Soit : $\|AX\|_2^2 \leq \lambda_n \|X\|_2^2$

avec cas d'égalité possible.

On en déduit que

$$|||A|||_2 = \sqrt{\lambda_n} = \sqrt{\rho({}^tAA)}$$

Remarque : il existe des normes d'algèbre qui ne sont pas subordonnées (cf TD 3).

Enfin, il existe un lien entre norme subordonnée et rayon spectral :

Théorème : Soit $A \in \mathcal{M}_n(\mathbb{R})$. On a pour toute norme subordonnée :

$$\boxed{\rho(A) \leq |||A|||}$$

Inversement, si $\varepsilon > 0$ est donné, il existe une norme subordonnée $||| \ |||_\varepsilon$ telle que

$$|||A|||_\varepsilon \leq \rho(A) + \varepsilon.$$

preuve : soit $\lambda \in Sp_{\mathbb{C}}(A)$ tq $|\lambda| = \rho(A)$ et X un vecteur propre ($\in \mathbb{C}^n$) associé. On a

$$\|AX\| = |\lambda| \, \|X\| = \rho(A) \|X\|$$

Comme $\|AX\| \leq |||A||| . \|X\|$, on a bien $\rho(A) \leq |||A|||$.

Inversement, on construit une norme subordonnée à l'aide de la norme $||| \ |||_\infty$ et d'un changement de base (voir Ciarlet, "Introduction à l'analyse numérique matricielle" ou G. Allaire "Algèbre linéaire numérique" ou poly etc...)

Une application de ce résultat en termes de convergence est le suivant :

Proposition : soit $A \in \mathcal{M}_n(\mathbb{R})$

(i) $\lim\limits_{i \to +\infty} A^i = 0$

(ii) $\lim\limits_{i \to +\infty} A^i x = 0$ pour tout $x \in \mathbb{R}^N$

(iii) $\rho(A) < 1$

(iv) il existe une norme subordonnée $|||\ |||_N$ telle que $|||A|||_N < 1$.

Ces 4 propositions sont équivalentes (preuve : exercice)

Remarque : si $\rho(A) < 1$, on peut montrer que la série $\sum\limits_{i=0}^{+\infty} A^i$ est convergente et que

$$\sum_{i=0}^{+\infty} A^i = (I - A)^{-1}$$

3) Conditionnement d'une matrice

Le conditionnement d'un problème (ici la résolution d'un système linéaire) désigne la sensibilité (forte ou non) de la solution par rapport aux erreurs d'arrondi commises durant sa solution.

* Ici, le problème à résoudre est le système : $\boxed{AX=b}$ où $A\in\mathcal{M}_n(\mathbb{R})$ inversible et $b\in\mathbb{R}^n$.
* On définit la notion de conditionnement d'une matrice afin de mesurer la sensibilité du calcul de X par rapport aux erreurs d'arrondis sur A et b.

Exemple : si $A=\begin{pmatrix} 1/3 & \pi \\ \sqrt{2} & 1/3 \end{pmatrix}$, on prendra numériquement

$$\tilde{A}=\begin{pmatrix} \overbrace{0.33\ldots3}^{16} & \overbrace{3,14\ldots}^{16} \\ 1,41\ldots & 0,33\ldots \end{pmatrix}$$

$\tilde{X}$ peut fortement différer de X...

Déf : soit $A\in GL_n(\mathbb{R})$. On note

$$\boxed{\underset{|||\ |||_N}{\mathrm{cond}}(A)=N(A).N(A^{-1})}$$

où $|||\ |||_N$ désigne une norme subordonnée quelconque.

On a la proposition :

Proposition : si $A\in GL_n(\mathbb{R})$, $b\in\mathbb{R}^n$ et $\delta b\in\mathbb{R}^n$. On résout :

$$Ax=b \text{ et } A(x+\delta x)=b+\delta b$$

On a alors :

$$\underbrace{\frac{N(\delta x)}{N(x)}}_{\text{erreur sur la solution}} \le \overbrace{\underset{|||\ |||_N}{\mathrm{cond}}(A)}^{\text{amplification}} \cdot \overbrace{\frac{N(\delta b)}{N(b)}}^{\text{erreur sur la donnée } b}$$

preuve :

$$Ax = b \text{ et } A(x+\delta x) = b + \delta b$$

implique :

$$\delta x = A^{-1}(\delta b) \text{ puis :}$$

$$(0 \leq)\, N(\delta x) \leq |||A^{-1}|||_N \, N(\delta b)$$

De même

$$0 \leq N(b) = N(Ax) \leq |||A|||_N \, N(x)$$

En multipliant, on trouve le résultat.

Remarques : 1) on montre de même :

$$\left(\frac{N(\delta x)}{N(\delta x + x)}\right) \leq \mathrm{cond}_{|||\cdot|||_N}(A) \left(\frac{|||\delta A|||_N}{|||A|||_N}\right)$$

erreur sur la donnée A

2) $\mathrm{cond}(A) \geq 1$

(car $AA^{-1} = Id \Rightarrow$

$$1 = |||Id||| \leq |||A|||\,.\,|||A^{-1}|||)$$

et un "bon conditionnement" correspond à une valeur peu élevée de $\mathrm{cond}(A)$

3) Il existe des techniques d'estimation rapide de $\mathrm{cond}_2(A)$ sans avoir à calculer A^{-1}.

4) Certaines matrices, (partant simples) ont des conditionnement très élevés.

Par exemple, la matrice de Hilbert

$$H_n = \begin{pmatrix} 1/2 & 1/3 & \cdots & \frac{1}{n+1} \\ 1/3 & & & \\ \vdots & & & \\ \frac{1}{n+1} & \cdots & & \frac{1}{2n} \end{pmatrix} \left(\left[\frac{1}{i+j}\right]\right)$$

est très mal conditionnée :

$$\mathrm{cond}_2(H_{10}) \sim 10^{19}$$

alors que cette matrice est sym. définie, positive à coefficients dans $]0,1[$.

(autres exemples : poly par exemple).

TD / TP n° 3 :

Ex 1

$A = [A_{ij}]$

$$N(A) = \sum_{i,j} |A_{ij}|$$

i) norme sur $\mathcal{M}_n(\mathbb{R})$? immédiat (on identifie $\mathcal{M}_n(\mathbb{R})$ et $\mathbb{R}^{n^2}$)

ii) norme subordonnée ? non car $N(I_d) = n \neq 1$ (par toute norme subordonnée, $|||I_d||| = 1$).

iii) norme matricielle ? oui (après avoir étudié le cas $n=2$). En effet,

$$N(AB) = \sum_{i,j} \left(\left| \sum_k a_{i,k} b_{k,j} \right| \right)$$

$$\leq \sum_{i,j} \left| \sum_k |a_{i,k}| \, |b_{k,j}| \right|$$

et

$$N(A)\,N(B) = \left(\sum_{i,k} |a_{i,k}| \right) \left(\sum_{k',j} |b_{k',j}| \right)$$

$$= \sum_{i,j,k,k'} |a_{i,k}| \, |b_{k',j}|$$

Ainsi $N(AB) \leq N(A)N(B)$

(norme d'algèbre non subordonnée)

Ex 2 $N(A) = \max_{i,j} |A_{i,j}|$

i) norme sur $\mathcal{M}_n(\mathbb{R})$? oui (cf Ex 1)

ii) $\exists A \in \mathcal{M}_n(\mathbb{K})$ tq $\rho(A) > N(A)$

On prend

$$J = \begin{pmatrix} 1 & \cdots & 1 \\ \vdots & & \vdots \\ 1 & \cdots & 1 \end{pmatrix}$$

On a $\rho(J) = n$ (vp : 0 et n)

$N(J) = 1$

iii) A subordonnée ? non

(ici $N(I_d) = 1$: pas de contradiction)

car sinon on aurait toujours

$\rho(A) \leq N(A)$.

Ex3) On note

$$\| A \|_s = \sqrt{\sum_{i,j} |A_{ij}|^2}$$

(norme de Schur, euclidienne)

B) c) non subordonnée car

$$\| Id \|_s = \sqrt{n} \neq 1$$

d) norme matricielle car :

$$\| AB \|_s^2 = \sum_{i,j} \Big(\sum_k |A_{ik} B_{kj}| \Big)^2$$

$$\| A \|_s^2 \| B \|_s^2 = \Big(\sum_{i,j} |A_{i,j}|^2 \Big) \Big(\sum_{k,l} |B_{k,l}|^2 \Big)$$

On a, avec Cauchy Schwarz :

$$\| AB \|_s^2 \leq \sum_{i,j} \Big(\Big(\sum_k |A_{i,k}|^2 \Big) \Big(\sum_k |B_{k,j}|^2 \Big) \Big)$$

puis

$$\| AB \|_s^2 \leq \sum_{i,j,k} |A_{i,k}|^2 |B_{k,j}|^2$$

$$\leq \underbrace{\sum_{i,j,k,k'} |A_{i,k}|^2 |B_{k',j}|^2}_{\| A \|_s \ \| B \|_s}$$

e) $\rho(A) \leq \| A \|_s$

$\lambda \in Sp_{\mathbb{C}}(A) \iff \exists X \in \mathbb{C}^n$ tq

$$\left.\begin{array}{l} A_{11} X_1 + \cdots + A_{1,n} X_n = \lambda X_1 \\ \\ A_{n1} X_1 + \cdots + A_{n,n} X_n = \lambda X_n \end{array}\right\}$$

A-t-on $|\lambda|^2 \leq \| A \|_s^2$?

En élevant au carré et en sommant :

$$|\lambda|^2 \sum_i |X_i|^2 = \sum_i \left|\sum_k A_{i,k} X_k\right|^2$$

D'où, en supposant $\sum_i |X_i|^2 = 1$, on a

$$|\lambda|^2 \leqslant \underbrace{\sum_i \left(\sum_k |A_{i,k}|^2\right)}_{\|A\|_S} \underbrace{\left(\sum_k |X_k|^2\right)}_{1}$$

par Cauchy-Schwarz.

Reste : f) et Ex 4)