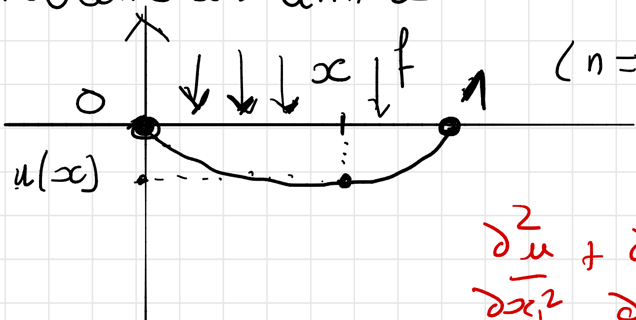
Séance 3 :

Normes et conditionnement de matrices

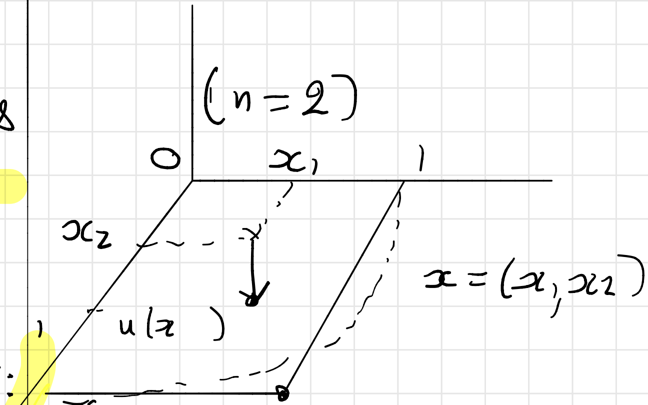
(Algèbre linéaire numérique)

1) Un exemple de matrice important : la matrice du laplacien discrétisé

Pb de départ : résolution d'un problème aux limites



$(n=1)$



* $x \mapsto u(x)$ est solution de l'EDO (ou EDP si $n=2$)

$$\begin{cases} -\underbrace{u''}_{\text{laplacien}}(x) + \underbrace{c}_{\text{rigidité}}\, u(x) = \underbrace{f(x)}_{\text{force}} \\ u(0) = u(1) = 0 \end{cases}$$

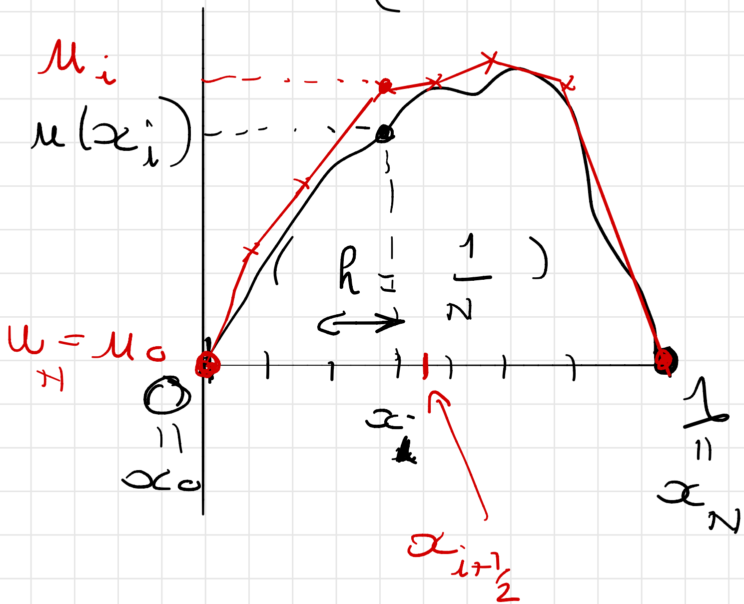
laplacien

$$\frac{\partial^2 u}{\partial x_1^2} + \frac{\partial^2 u}{\partial x_2^2}$$

$$\left(\text{ou } \begin{cases} -\Delta u + c\, u = f \\ u_{/\partial\Omega} = 0 \end{cases} \quad \text{si } n=2 \right)$$

* Cette EDO (ou EDP) n'admet pas de solution explicite exacte.

* Il est nécessaire de rechercher une solution approchée en discrétisant le domaine Ω $(= [0,1]$ ou $[0,1]^2)$



* On cherche à calculer une approximation u_i de $u(x_i)$ en discrétisant le laplacien :

$$\boxed{u''(x_i) \simeq \frac{\underbrace{\frac{u(x_{i+1}) - u_i}{h}}_{\simeq u'(x_{i+\frac{1}{2}})} - \underbrace{\frac{u(x_i) - u(x_{i-1})}{h}}_{\simeq u'(x_{i-\frac{1}{2}})}}{h} \simeq \frac{u(x_{i+1}) - 2u(x_i) + u(x_{i-1})}{h^2}}$$

* En substituant cette approximation dans l'EDO, on se ramène à la résolution du système :

$$\begin{cases} -\dfrac{u_{i-1}+2u_i-u_{i+1}}{h^2} + c\,u_i = f_i \quad (1 \leq i \leq N-1) \\ u_0 = u_N = 0 \end{cases}$$

$(f_i = f(x_i))$

Soit :

$\boxed{AX + cX = b}$ (ou $(A + cI)X = b$)

avec $A = \dfrac{1}{h^2}\begin{pmatrix} 2 & -1 & & 0 \\ -1 & \ddots & \ddots & \\ & \ddots & \ddots & -1 \\ 0 & & -1 & 2 \end{pmatrix}$ $\updownarrow N-1$

$b = \begin{pmatrix} f_1 \\ \vdots \\ f_{N-1} \end{pmatrix}$

(matrice du laplacien discrétisé en dim. 1)

* Cette matrice est symétrique, tridiagonale par bloc ; on peut montrer que A est définie positive (et même calculer explicitement ses espaces propres). (voir TD à venir).

* A est également bien conditionnée (voir TD)

2) Normes de matrices

Def : E espace vectoriel. $N : E \to \mathbb{R}$ est une norme si :

$$\begin{cases} (i)\ \forall x \in E,\ N(x) \geq 0 \text{ et} \\ N(x) = 0 \iff x = 0 \\ (ii)\ \forall (x,y) \in E^2,\ N(x+y) \leq N(x) + N(y) \\ (iii)\ \forall x \in E,\ \forall \lambda \in \mathbb{R},\ N(\lambda x) = |\lambda| N(x) \end{cases}$$

Exemple : $E = \mathbb{R}^n$

* $\|x\|_1 = \sum_{i=1}^{n} |x_i|$
* $\|x\|_2 = \sqrt{\sum_{i=1}^{n} x_i^2}$
* $\|x\|_\infty = \underset{1 \leq i \leq n}{\text{Max}} |x_i|$

(idem sur $\mathcal{M}_n(\mathbb{R})$:
$\|A\|_1 = \sum_{i,j} |A_{i,j}|$, ...)

Th : si E est un ev de dimension finie, toutes les normes sur E sont équivalentes : soit N_1 et N_2 2 normes sur E, il existe $C_1 > 0$, $C_2 > 0$ tq $\forall x \in E$,

$$C_1 N_1(x) \leq N_2(x) \leq C_2 N_1(x)$$

(preuve : comparer N_1 et N_2 avec)
$\|x\|_\infty = \underset{1 \leq i \leq n}{\text{Max}} |x_i|$ où
$x = \sum_{i=1}^{n} x_i e_i$

* Déf : si E est une algèbre, on dit que N est une norme d'algèbre si $\forall (x,y) \in E^2$,

$$N(xy) \leq N(x)N(y)$$

Qn : $\| \ \|_1, \| \ \|_2, \| \ \|_\infty$ normes d'algèbre sur $\mathcal{M}_n(\mathbb{R})$? (voir TD)

* Déf : si $E = \mathcal{M}_n(\mathbb{R})$, on dit que N est une norme subordonnée à la norme $\| \ \|$ sur $\mathbb{R}^n$ si $\forall A \in \mathcal{M}_n(\mathbb{R})$

$$\boxed{N(A) = \sup_{x \neq 0} \frac{\|Ax\|}{\|x\|}} \quad (*)$$

(ou $N(A) = \underset{\|x\|=1}{\cancel{\sup}\ \text{Max}} \|Ax\|$)

On la note $N(A) = |||A|||$)

Remarque : la relation $(*)$ définit bien une norme sur $\mathcal{M}_n(\mathbb{R})$.

Exemple :

$$|||A|||_{1,2,\infty} = \sup_{x \neq 0} \frac{\|Ax\|_{1,2,\infty}}{\|x\|_{1,2,\infty}}$$

Proposition : toute norme subordonnée est une norme d'algèbre : en effet

Si A et $B \in \mathcal{M}_n(\mathbb{R})$

$$|||AB||| = \sup_{x \neq 0} \frac{\|ABx\|}{\|x\|}$$

$$\leq |||A||| \cdot \sup \frac{\|Bx\|}{\|x\|}$$

(car $\|Ay\| \leq |||A|||\ \|y\|$)

$$\leq |||A|||.\ |||B|||$$

On peut expliciter les 3 normes subordonnées classiques sur $\mathcal{M}_n(\mathbb{R})$

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

(i)

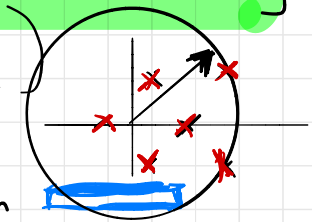
$$|||A|||_1 = \underset{1 \leq j \leq n}{\mathrm{Max}} \left(\sum_{i=1}^{n} |A_{ij}| \right)$$

(ii)

$$|||A|||_2 = \sqrt{\rho({}^tA A)} \quad \text{où}$$

$$\rho(M) = \mathrm{Max}\{|\lambda|,\ \lambda \in Sp_{\mathbb{C}}(M)\}$$

(rayon spectral de M)



(iii)

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

preuve :

(i)

$$\|Ax\|_1 = \left\| \begin{pmatrix} \sum A_{1,j} x_j \\ \vdots \\ \sum A_{n,j} x_j \end{pmatrix} \right\|_1$$

$$= \sum_{i=1}^{n} \left| \sum_{j=1}^{n} A_{i,j} x_j \right|$$

$$\leq \sum_{i,j} |A_{i,j}|\,|x_j|$$

$$= \sum_{j} |x_j| \underbrace{\left(\sum_{i} |A_{i,j}| \right)}$$

$$\leq \max_{j} \left(\sum_{i} |A_{i,j}| \right) \|x\|_1$$

$$\Rightarrow |||A|||_1 \leq \max_{j} \left(\sum_{i} |A_{i,j}| \right)$$

Réciproquement, soit

$$X = \begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix} \leftarrow j_0$$ où j_0 est tel que

$\left(\sum_{i=1}^{n} |A_{i,j}| \right)$ est maximal en j_0.

Dans ce cas

$$\|AX\|_1 = \left\| \begin{pmatrix} A_{1,j_0} \\ \vdots \\ A_{n,j_0} \end{pmatrix} \right\| = \max_{j} \left(\sum_{i} |A_{i,j}| \right)$$

et on a bien la relation cherchée.

(ii)

$$\frac{\|Ax\|_2^2}{\|x\|_2^2} = \frac{\langle Ax, Ax\rangle}{\langle x, x\rangle} = \frac{\langle {}^tAAx, x\rangle}{\langle x, x\rangle}$$

Or, ${}^tAA$ est une matrice symétrique, positive. Ses valeurs propres sont réelles et positives : $0 \leq \lambda_1 \leq \dots \leq \lambda_n$

De plus,

$$\langle {}^tAAx, x\rangle = \sum_{i=1}^{n} \lambda_i |x_i|^2$$

si $x = \sum_{i=1}^{n} x_i e_i$ (dans une base associée aux valeurs propres de ${}^tAA$)

On a donc :

$$\max_{x \neq 0} \frac{\|Ax\|^2}{\|x\|^2} = \lambda_n = \rho({}^tAA)$$

(iii) à faire en exercice.

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

Il existe un lien entre rayon spectral et norme subordonnée :

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

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

De plus, soit $\varepsilon > 0$, il existe une norme subordonnée, $||| \ |||_s$ telle que

($\rho(A) \leq$)

$$|||A|||_s \leq \rho(A) + \varepsilon$$

preuve : soit $\lambda \in Sp_{\mathbb{C}}(A)$ et X un vecteur propre associé ($\in \mathbb{C}^n$)

$$||AX|| = |\lambda| \, ||X|| \leq |||A|||.||X||$$

$$\Longrightarrow \quad |\lambda| \leq |||A|||$$

"Réciproquement", on triangularise A sur $\mathcal{M}_n(\mathbb{C})$:

$$A = U \underbrace{\begin{pmatrix} \lambda_1 & & * \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}}_{T} U^{-1}$$

Soit $\delta > 0$, on note

$$D_\delta = \mathrm{diag}(1, \delta, \ldots, \delta^{n-1})$$

et

$$T_\delta = (UD_\delta)^{-1} A (UD_\delta)$$

$$= \begin{pmatrix} \lambda_1 & \delta t_{2,1} & \cdots & \delta^{n-1} t_{1,n} \\ & \ddots & & \vdots \\ & & \ddots & \delta t_{n-1,n} \\ 0 & & & \lambda_n \end{pmatrix} \quad \leq \varepsilon$$

On choisit $\delta > 0$ tel que

$$\sum_{j=i+1}^{n} \delta^{j-1} |t_{ij}| \leq \varepsilon$$

L'application

$$B \longmapsto ||| (UD_\delta)^{-1} B (UD_\delta) |||_\infty$$

est une norme subordonnée. De plus

$$||| A |||_s = ||| T_\delta |||_\infty$$

$$= \max_{1 \leq i \leq n} \left(\sum_j |(T_\delta)_{ij}| \right)$$

$$\leq \underbrace{\max_{1 \leq i \leq n} |\lambda_i| + \varepsilon}_{\rho(A) + \varepsilon}$$

3) Suites et séries de matrices

Même s'il n'y a qu'une notion de convergence, il est intéressant de bien choisir la norme utilisée en fonction du contexte.

La convergence des suites géométriques de matrices $i \mapsto A^i$ peut se faire de différentes manières :

Proposition On a :

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

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

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

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

preuve : (i) $\Rightarrow$ (ii)

$$\|A^i x\| \leq \underbrace{|||A^i|||}_{\downarrow \, 0} \|x\|$$

(ii) $\Rightarrow$ (iii) Par l'absurde, soit X tq

$AX = \rho(A) X$

On a $A^i X = \underbrace{(\rho(A))^i}_{|\ | \geq 1} X \not\to 0$.

(iii) $\Rightarrow$ (iv) Soit $\varepsilon > 0$ tq $\rho(A) + \varepsilon < 1$ et on conclut avec le théorème précédent et $|||\ |||_s$.

(iv) $\Rightarrow$ (i)

$$|||A^i|||_s \leq |||A|||_s^i \xrightarrow[i \to +\infty]{} 0.$$

On peut aussi exploiter le caractère complet de $\mathcal{M}_n(\mathbb{R})$ par la convergence de séries de matrices :

Proposition : soit la série entière $\sum a_i z^i$ de rayon de convergence R. Soit $A \in \mathcal{M}_n(\mathbb{R})$ telle que $\rho(A) < R$. La série $\sum a_i A^i$ est absolument convergente, donc convergente dans $\mathcal{M}_n(\mathbb{R})$.

preuve : $$\left|\left|\left| \sum_{i=1}^{N} a_i A^i \right|\right|\right|_\Delta \leq \sum_{i=1}^{N} |a_i| \, \underbrace{|||A|||_\Delta}_{< R}^{\,i}$$

En particulier, soit $A \in \mathcal{M}_n(\mathbb{R})$ tq $\rho(A) < 1$. La série $\sum_{i=0}^{\infty} A^i$ est convergente et de plus
$$\sum_{i=0}^{\infty} A^i = (I - A)^{-1}$$

preuve : $\sum A^i$ est convergente car $\rho\left(\sum_{i=0}^{\infty} z^i\right) = 1$

De plus
$$\underbrace{\left(\sum_{i=0}^{N} A^i\right)}_{\downarrow \ \sum_{i=0}^{\infty} A^i} (I - A) = I - \underbrace{A^{N+1}}_{\downarrow \ I}$$

4) Conditionnement d'une matrice

* On s'intéresse au problème de la résolution numérique d'un système linéaire de Cramer :

$$\boxed{Ax = b} \quad (A \in GL_n(\mathbb{R}))$$

* On cherche à quantifier le risque d'une erreur commise sur la matrice A (par exemple si $A = \begin{pmatrix} 1/3 & 1/7 \\ \pi & \sqrt{2} \end{pmatrix}$) sur le résultat, ici la valeur de x.
* Si le risque est faible (l'amplification d'erreur est faible), on dit que la matrice A est bien conditionnée pour la résolution des systèmes linéaires (et mal conditionnée sinon).
* Dans le cas présent, on peut quantifier cette notion de conditionnement :

Déf : soit $A \in GL_n(\mathbb{R})$ et $||| \ |||$ une norme subordonnée. On note

$$\boxed{\text{cond}(A) = |||A||| \cdot |||A^{-1}|||}$$

Proposition (i) Soit $A \in GL_n(\mathbb{R})$, $b \in \mathbb{R}^n$, $\delta b \in \mathbb{R}^n$

On note $\begin{cases} x : \text{solution de } Au = b \\ x+\delta x \text{ ---- de } Au = b+\delta b \end{cases}$

Alors $$\boxed{\underbrace{\frac{\|\delta x\|}{\|x\|}}_{\text{erreur relative sur } x} \leq \underbrace{\text{cond}(A)}_{\text{(amplification)}} \underbrace{\frac{\|\delta b\|}{\|b\|}}_{\text{erreur relative sur } b}}$$

(ii) Soit $A \in GL_n(\mathbb{R})$. On note $A+\delta A \in GL_n(\mathbb{R})$

$\begin{cases} x \text{ solution de } Au = b \\ x+\delta x \text{ -- de } (A+\delta A)u = b \end{cases}$

On a

$$\boxed{\underbrace{\frac{\|\delta x\|}{\|x+\delta x\|}}_{\text{erreur sur } x} \leq \underbrace{\text{cond}(A)}_{\text{(amplification)}} \underbrace{\frac{|||\delta A|||}{|||A|||}}_{\text{erreur sur } A}}$$

preuve :

(i) $Ax = b$ et $A(x+\delta x) = b+\delta b$ implique $A\delta x = \delta b$ soit $\delta x = A^{-1}\delta b$. On a donc

$\alpha \leq \|\delta x\| \leq |||A^{-1}||| \, \|\delta b\|$

et

$0 \leq \|b\| \leq |||A||| \, \|x\|$ //

(ii) voir TD

Remarques :

* il s'agit des constantes d'amplification optimales.
* il existe des méthodes de calcul approché de $\text{cond}_2(A)$ (associé à $|||\ |||_2$) sans avoir à calculer A^{-1} !
* On a toujours $\text{cond}(A) \geq 1$ car

$$|||Id||| = 1$$
$$= |||AA^{-1}||| \leq \underbrace{|||A|||.|||A^{-1}|||}_{\text{cond}(A)}$$

* Il existe des matrices pour lesquelles $\text{cond}(A)$ est très grand

Par exemple, la matrice de Hilbert $H_n = \begin{pmatrix} 1 & 1/2 & 1/3 & \cdots & 1/(n+1) \\ 1/2 & & & & \vdots \\ \vdots & & & & \vdots \\ 1/(n+1) & & \cdots & & 1/(2n) \end{pmatrix}$

($\in \mathcal{S}_n^{++}(\mathbb{R})$) est très mal conditionnée ($\text{cond}_2(H_{10}) \sim 10^{13}$)

(mais aussi : Vandermonde, etc. ....)

(voir TP)

TD3 :

1) ii) $\|A\|_1 = \sum_{i,j} |A_{i,j}|$

n'est pas subordonnée car :

$\|Id\|_1 = n \neq 1$ (si $n>1$)

2) iii) $\|A\|_\infty = \operatorname{Max}_{i,j} |A_{i,j}|$

n'est pas subordonnée car :

$\|A\|_\infty < \rho(A)$ pour certaines matrices A.

A faire (par groupes éventuellement)

Ex 1, 2, 3, 4 à compléter

→ Mercredi 30/09