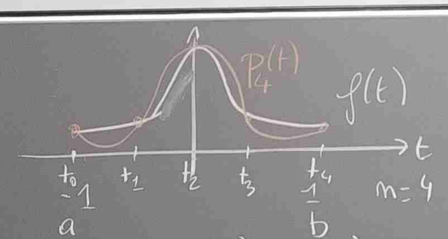
Chap I lime : interpolation

Rappel : interp. de lagrange

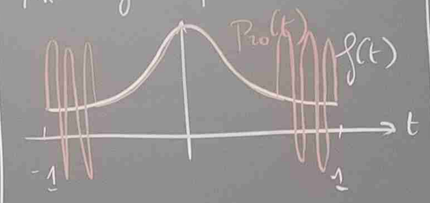
$f : [a,b] \to \mathbb{R}$ ex $[a,b] = [-1,1]$

$f(t) = \dfrac{1}{1+25t^2}$ $\quad t_j = a + \dfrac{b-a}{n} j$ $\quad j = 0,1,\dots,n$



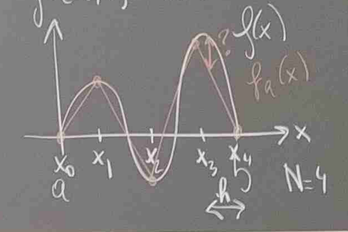
$P_n \in \mathbb{P}_n$ tq $P_n(t_j) = f(t_j)$ $\quad j = 0,1,\dots,n$

$P_n \xrightarrow{n\to\infty} f$? pas ici



Interpolation par intervalles :

$f : [a,b] \to \mathbb{R}$



$x_i = a + \underbrace{\dfrac{b-a}{N}}_{h} i$ $\quad i = 0,1,\dots,N$

f_h interp. degré 1 par intervalles

$f_h \in \mathcal{C}^0[a,b]$

$f_h(x_i) = f(x_i)$ $\quad i = 0,\dots,N$

$f_h|_{[x_i, x_{i+1}]} \in \mathbb{P}_1$ $\quad i = 0,\dots,N-1$

$f_h \xrightarrow{h\to 0} f$? oui

Thm 1.2 lime n=1 :

$\forall f \in \mathcal{C}^2[a,b]$ $\exists C > 0$ $\forall 0 < h < b-a$

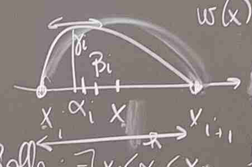
$\max_{a \le x \le b} |f(x) - f_h(x)| \le C h^2$

Interprétation :

- C dépend de f, mais pas de h
- Exp. num : on choisit un $f \in \mathcal{C}^2[a,b]$
  l'erreur est divisée par (au moins) 2^2 chaque fois que h est divisé par 2
  $\frac{2^2}{10^2}$ $\quad\quad$ $\frac{10}{10}$

Dem : Soit $x_i < x < x_{i+1}$ $\quad i = 0,\dots,N-1$

$w(x) = f(x) - f_h(x)$



Rolle : $\exists x_i < \alpha_i < x_{i+1}$ tq $w'(\alpha_i) = 0$

TAF : $w(x) = \underbrace{w(x_i)}_{0} + (x - x_i) w'(\beta_i)$ où $x_i < \beta_i < x$

or $w'(\beta_i) = w'(\alpha_i) + (\beta_i - x_i) w''(\gamma_i)$

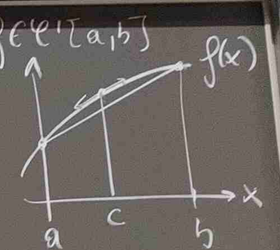
où $\min(\alpha_i, \beta_i) < \gamma_i < \max(\alpha_i, \beta_i)$

ie $w'(\beta_i) = (\beta_i - \alpha_i) f''(\gamma_i)$

et donc

$w(x) = (x - x_i)(\beta_i - \alpha_i) f''(\gamma_i)$

$f \in \mathcal{C}^1[a,b]$



$\exists a < c < b$ $\quad \dfrac{f(b) - f(a)}{b-a} = f'(c)$

$|w(x)| = |x - x_i| \cdot |\beta_i - \alpha_i| \cdot |f''(\gamma_i)|$

$|f(x) - f_h(x)| \le \underbrace{h}_{h^2} \underbrace{h}_{} \underbrace{\max_{a \le y \le b} |f''(y)|}_{C}$ $\quad x_i < x < x_{i+1}$ $\quad i = 0,\dots,N-1$

Si $x = x_i$ $\quad f(x_i) - f_h(x_i) = 0$

et donc $\forall x \in [a,b]$ on a $|f(x) - f_h(x)| \le C h^2$ c'est donc vrai pour le x qui réalise le max de $|f(x) - f_h(x)|$

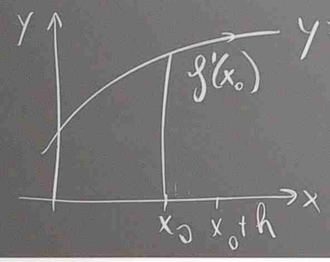
Thm (bis) :

$\exists C > 0$ $\forall f \in \mathcal{C}^2[a,b]$ $\forall 0 < h < b-a$

$\max_{a \le x \le b} |f(x) - f_h(x)| \le \underset{\le 1}{C} h^2 \max_{a \le y \le b} |f''(y)|$

Chap 2 bis - Formules de Diff. finies (FDF)

$f : \mathbb{R} \to \mathbb{R}$ $\mathcal{C}^1$
$x_0 \in \mathbb{R}$ $f'(x_0)$?



$$\left| f'(x_0) - \frac{f(x_0+h)-f(x_0)}{h} \right| = O(h) \quad \text{(FDF progressive } f \in \mathcal{C}^2\text{)}$$

$$\left| f'(x_0) - \frac{f(x_0)-f(x_0-h)}{h} \right| = O(h) \quad \text{(FDF rétrograde } f \in \mathcal{C}^2\text{)}$$

$$\left| f'(x_0) - \frac{f(x_0+\frac{h}{2})-f(x_0-\frac{h}{2})}{h} \right| = O(h^2) \quad \text{(FDF centrée } f \in \mathcal{C}^3\text{)}$$

$$\left| f''(x_0) - \frac{f(x_0-h)-2f(x_0)+f(x_0+h)}{h^2} \right| = O(h^2) \quad \text{(FDF centrée } f \in \mathcal{C}^4\text{)}$$

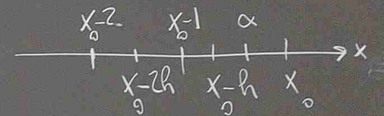
Aug : $\left| f'(x_0) - \frac{3f(x_0)-4f(x_0-h)+f(x_0-2h)}{2h} \right| = O(h^2)$ $(f \in \mathcal{C}^3)$

Thm. $\forall x_0 \in \mathbb{R}\ \forall f \in \mathcal{C}^3[x_0-2, x_0]\ \exists C>0\ \forall 0<h\leq 1$ on a :

$$\left| f'(x_0) - \frac{3f(x_0)-4f(x_0-h)+f(x_0-2h)}{2h} \right| \leq C h^2$$

Interpr :
- C dépend de x_0 et f mais pas de h.
- On choisit x_0 et f,
l'erreur est divisée (au moins) par 2^2 si h est divisé par 2
10^2 10

Dem : Soit $x_0 \in \mathbb{R}$, soit $f \in \mathcal{C}^3[x_0-2, x_0]$
soit $0<h\leq 1$



Attention : ne pas confondre dév. Taylor et série de Taylor

ex : $x \geq 0$ en 0 (série) $e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!} = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \dots + \frac{x^n}{n!} + \frac{x^{n+1}}{(n+1)!} + \dots$

développ $\to e^x = 1 + x + \frac{x^2}{2} e^{\alpha}$ où $0<\alpha<x$

$= 1 + x + \frac{x^2}{2} + \frac{x^3}{6} e^{\beta}$ où $0<\beta<x$

(*) : conclusion : on a atteint le thm avec $C = \max_{x_0-2 \leq x \leq x_0} |f'''(x)|$

$|a+b| \leq |a| + |b|$
$|-a+b| \leq \underbrace{|-a|}_{|a|} + |b|$

$$4\left(f(x_0-h) = f(x_0) - hf'(x_0) + \frac{h^2}{2}f''(x_0) - \frac{h^3}{6}f'''(\alpha)\right) \text{ où } x_0-h<\alpha<x_0$$

$$-\left(f(x_0-2h) = f(x_0) - 2hf'(x_0) + \frac{4h^2}{2}f''(x_0) - \frac{8h^3}{6}f'''(\beta)\right) \text{ où } x_0-2h<\beta<x_0$$

$$4f(x_0-h) - f(x_0-2h) = 3f(x_0) - 2hf'(x_0) - \frac{2}{3}h^3 f'''(\alpha) + \frac{4}{3}h^3 f'''(\beta)$$

$$f'(x_0) - \frac{3f(x_0)-4f(x_0-h)+f(x_0-2h)}{2h} = h^2\left(-\frac{1}{3}f'''(\alpha) + \frac{2}{3}f'''(\beta)\right)$$

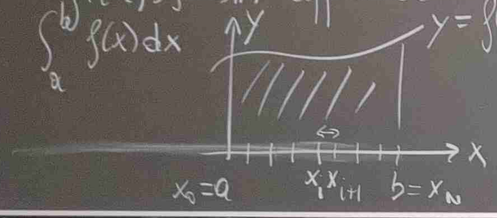
$$\left| f'(x_0) - \frac{3f(x_0)-4f(x_0-h)+f(x_0-2h)}{2h} \right| \leq h^2 \left| -\frac{1}{3}f'''(\alpha) + \frac{2}{3}f'''(\beta) \right|$$

$$\leq h^2 \left(\frac{1}{3}|f'''(\alpha)| + \frac{2}{3}|f'''(\beta)| \right)$$

$$\leq h^2 \left(\underbrace{\frac{1}{3}+\frac{2}{3}}_{1} \right) \max_{x_0-2 \leq x \leq x_0} |f'''(x)| \quad (*)$$

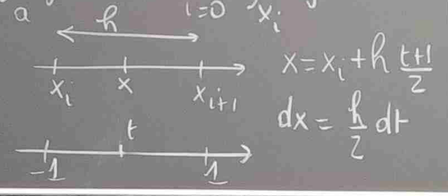
Chap 3. Formule de quadr.

$f:[a,b]\to\mathbb{R}$ approcher $\int_a^b f(x)dx$



$x_i = a + \frac{b-a}{N} i = a + ih$ $\quad 0\le i\le N$

$\int_a^b f(x)dx = \sum_{i=0}^{N-1} \int_{x_i}^{x_{i+1}} f(x)dx$



$x = x_i + h\frac{t+1}{2}$

$dx = \frac{h}{2}dt$

$\int_a^b f(x)dx = \frac{h}{2}\sum_{i=0}^{N-1}\int_{-1}^{1} f(x_i+h\frac{t+1}{2})dt$ $\quad g_i(t)$

Soit $g:[-1,1]\to\mathbb{R}$ on approche $\int_{-1}^{1} g(t)dt$ par $J(g) \overset{\text{Simpson}}{=} \frac{1}{3}\left(g(-1)+4g(0)+g(1)\right)$

$\forall p\in\mathbb{P}_3 \quad \int_{-1}^{1} p(t)dt = J(p)$

approx de $\int_a^b f(x)dx$: $L_h(f) = \frac{h}{2}\sum_{i=0}^{N-1} J(g_i)$

$L_h(f) = \frac{h}{3\cdot 2}\sum_{i=0}^{N-1}\left(g_i(-1)+4g_i(0)+g_i(1)\right)$

$= \frac{h}{6}\sum_{i=0}^{N-1}\left(f(x_i)+4f(x_i+\frac{h}{2})+f(x_{i+1})\right)$

Thm : 3 ! fonction $(r=3)$

Thm : $\forall f\in\mathcal{C}^4[a,b]\ \exists C>0\ \forall\, 0<h\le b-a$ on a $\left|\int_a^b f(x)dx - L_h(f)\right| \le Ch^4$

Interpr : • C dépend de f pas de h

• On choisit $f:[a,5]\to\mathbb{R}$ $\mathcal{C}^4$

L'erreur est (au moins) divisée par $2^4=16$ Chaque fois que h est divisée par 2.

Dem : $E = \int_a^b f(x)dx - L_h(f)$

$= \frac{h}{2}\sum_{i=0}^{N-1}\left(\int_{-1}^{1} g_i(t)dt - J(g_i)\right)$

Si on sait montrer que $i=0,\dots,N-1$ $-1\le t\le 1$

$(*)$ $g_i(t) = p_i(t) + r_i(t)$ où $p_i\in\mathbb{P}_3$

alors on a :

$E = \frac{h}{2}\sum_{i=0}^{N-1}\left(\int_{-1}^{1}(p_i(t)+r_i(t))dt - J(p_i+r_i)\right)$

$= \frac{h}{2}\sum_{i=0}^{N-1}\left(\int_{-1}^{1} r_i(t)dt - J(r_i)\right)$

Dem de $(*)$; $i=0,\dots,N-1$ $-1\le t\le 1$

$g_i(t) = g_i(0) + t g_i'(0) + \frac{t^2}{2}g_i''(0) + \frac{t^3}{6}g_i'''(0) + \frac{t^4}{24}g_i^{(4)}(\alpha_i)$ où $\min(0,t)<\alpha_i<\max(0,t)$

$g_i'(t) = f'(x_i+h\frac{t+1}{2})\frac{h}{2}$

$g_i''(t) = f''(\quad)\frac{h^2}{4}$

$g_i'''(t) = f'''(\quad)\frac{h^3}{8}$

$g_i^{(4)}(t) = f^{(4)}(\quad)\frac{h^4}{16}$

Donc $g_i(t) = \underbrace{f(x_i+\frac{h}{2}) + \frac{th}{2}f'(x_i+\frac{h}{2}) + \frac{t^2h^2}{8}f''(x_i+\frac{h}{2}) + \frac{t^3h^3}{48}f'''(x_i+\frac{h}{2})}_{p_i(t)} + \underbrace{\frac{t^4h^4}{16\cdot 24}f^{(4)}(x_i+h\frac{\alpha_i+1}{2})}_{r_i(t)}$

$r_i(t)\notin\mathbb{P}_4$ car α_i dépend de t et b

$|E| = \frac{h}{2}\left|\sum_{i=0}^{N-1}\left(\int_{-1}^{1} r_i(t)dt - J(r_i)\right)\right|$

$\le \frac{h}{2}\sum_{i=0}^{N-1}\left|\int_{-1}^{1} r_i(t)dt - J(r_i)\right| \le \frac{h}{2}\sum_{i=0}^{N-1}\left(\left|\int_{-1}^{1} r_i(t)dt\right| + \left|J(r_i)\right|\right)$

On a $\left|\int_{-1}^{1} r_i(t)dt\right| \le \int_{-1}^{1}|r_i(t)|dt \le 2\max_{-1\le t\le 1}|r_i(t)|$

$\le 2\,\frac{1^4h^4}{16\cdot 24}\max_{a\le x\le b}|f^{(4)}(x)|$

$|J(r_i)| = \left|\frac{1}{3}\left(r_i(-1)+4\underbrace{r_i(0)}_{0}+r_i(1)\right)\right|$

$\le \frac{1}{3}\left(|r_i(-1)| + |r_i(1)|\right)$

$\le \frac{2}{3}\,\frac{1^4h^4}{16\cdot 24}\max_{a\le x\le b}|f^{(4)}(x)|$

$|E| \le \frac{h}{2}\sum_{i=0}^{N-1}\frac{(2+\frac{2}{3})}{16\cdot 24}\max_{a\le x\le b}|f^{(4)}(x)|\,h^4$

$= \frac{(2+\frac{2}{3})}{2\cdot 16\cdot 24}\max_{a\le x\le b}|f^{(4)}(x)|\,h^4\,h\underbrace{\sum_{i=0}^{N-1}1}_{N}$

$\frac{b-a}{N}$

On a démontré le Thm avec $C = \frac{(b-a)(2+\frac{2}{3})}{2\cdot 16\cdot 24}\max_{a\le x\le b}|f^{(4)}(x)|$





Chap 4, 5, 6

Savoir résoudre un système linéaire

$A\bar{x} = \bar{b}$ $\quad A \in \mathbb{R}^{N\times N}$ $\quad \bar{b} \in \mathbb{R}^N$

par exemple si $A = \begin{pmatrix} 2 & -1 & (0) \\ -1 & \ddots & -1 \\ (0) & -1 & 2 \end{pmatrix}$

Aujourd'hui : lien entre $A\bar{x} = \bar{b}$ et un problème de minimisation (2e partie du cours)

Def. $A \in \mathbb{R}^{N\times N}$ sym. def positive

si $\forall \bar{x} \in \mathbb{R}^N \quad \bar{x}^T A \bar{x} \geqslant 0$

$\bar{x}^T A \bar{x} = 0 \Leftrightarrow \bar{x} = \bar{0}$

Notation : $\forall \bar{x}, \bar{y} \in \mathbb{R}^N$

produit scalaire $\bar{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_N \end{pmatrix}$ $\bar{y} = \begin{pmatrix} y_1 \\ \vdots \\ y_N \end{pmatrix}$ $\bar{x}^T \bar{y} = (x_1 \cdots x_N)\begin{pmatrix} y_1 \\ \vdots \\ y_N \end{pmatrix}$

$= \sum_{i=1}^N x_i y_i = \langle \bar{x}, \bar{y} \rangle$

$= \langle \bar{y}, \bar{x} \rangle = \bar{y}^T \bar{x}$

$\forall \bar{x} \in \mathbb{R}^N$

$\bar{x}^T A \bar{x} = \langle \bar{x}, A\bar{x} \rangle = \sum_{i=1}^N x_i (A\bar{x})_i$

$= \sum_{i=1}^N x_i \sum_{j=1}^N a_{ij} x_j$

$= \sum_{i=1}^N \sum_{j=1}^N a_{ij} x_i x_j = \sum_{i,j=1}^N a_{ij} x_i x_j$

Fonctionnelle $\mathcal{L}(\bar{y})$ quadratique

$\mathcal{L} : \mathbb{R}^N \to \mathbb{R}$, $\bar{y} \to \mathcal{L}(\bar{y})$

$\mathcal{L}(\bar{y}) = \frac{1}{2}\bar{y}^T A \bar{y} - \bar{b}^T \bar{y}$

où $A \in \mathbb{R}^{N\times N}$ et $b \in \mathbb{R}^N$ donnés

---

Thm : A sym. def positive $\bar{b} \in \mathbb{R}^N$

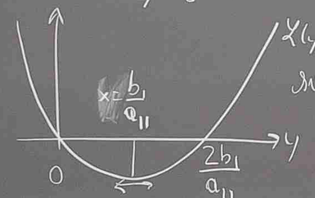
(Trouver $\bar{x} \in \mathbb{R}^N$ tel que $A\bar{x} = \bar{b}$)

$\Leftrightarrow$

(Trouver $\bar{x} \in \mathbb{R}^N$ tel que $\mathcal{L}(\bar{x}) \leqslant \mathcal{L}(\bar{y}) \ \forall \bar{y} \in \mathbb{R}^N$)

N = 1 $\quad \mathcal{L}(y) = \frac{1}{2} a_{11} y^2 - b_1 y \quad (y = y_1)$

$= y(\frac{1}{2} a_{11} y - b_1)$ $\quad a_{11} > 0$



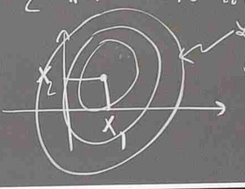
syst. linéaire $a_{11} x = b_1$

N = 2 $\quad A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$

$\mathcal{L}(\bar{y}) = \frac{1}{2} a_{11} y_1^2 + 2a_{12} y_1 y_2 + a_{22} y_2^2 - (b_1 y_1 + b_2 y_2)$

$\bar{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$

$A\bar{x} = \bar{b}$



grad $\mathcal{L}$: $\forall \bar{y} \in \mathbb{R}^N$ $\quad \bar{y} = \begin{pmatrix} y_1 \\ \vdots \\ y_N \end{pmatrix}$ $\quad \overrightarrow{\text{grad}}\,\mathcal{L}(\bar{y}) = \begin{pmatrix} \frac{\partial}{\partial y_1}\mathcal{L}(\bar{y}) \\ \vdots \\ \frac{\partial}{\partial y_N}\mathcal{L}(\bar{y}) \end{pmatrix}$

On a : $\overrightarrow{\text{grad}}\,\mathcal{L}(\bar{y}) = A\bar{y} - \bar{b}$

En effet montrons $\frac{\partial}{\partial y_k}\mathcal{L}(\bar{y}) = (A\bar{y} - \bar{b})_k$

$k = 1, \ldots, N$ $\quad = \sum_{i=1}^N a_{ki} y_i - b_k$

Soit $k = 1, \ldots, N$

$\frac{\partial}{\partial y_k}\mathcal{L}(\bar{y}) = \frac{\partial}{\partial y_k}\left(\frac{1}{2}\sum_{i=1}^N\sum_{j=1}^N a_{ij} y_i y_j - \sum_{i=1}^N b_i y_i\right)$

$= \frac{1}{2}\sum_{i=1}^N\sum_{j=1}^N a_{ij}\left(\frac{\partial y_i}{\partial y_k} y_j + y_i \frac{\partial y_j}{\partial y_k}\right) - \sum_{i=1}^N b_i \frac{\partial y_i}{\partial y_k}$

$\frac{\partial y_i}{\partial y_k} = 1$ si $i = k$, $= 0$ sinon $\quad \frac{\partial y_j}{\partial y_k} = 1$ si $j = k$, $= 0$ sinon

$\frac{\partial}{\partial y_k}\mathcal{L}(\bar{y}) = \frac{1}{2}\left(\sum_{j=1}^N a_{kj} y_j + \sum_{i=1}^N \underbrace{a_{ik}}_{a_{ki}\ (A = A^T)} y_i\right) - b_k$

$= \sum_{i=1}^N a_{ki} y_i - b_k = (A\bar{y} - \bar{b})_k$

Montrons le thm :

$\Leftarrow$: Supp $\mathcal{L}(\bar{x}) \leqslant \mathcal{L}(\bar{y}) \ \forall \bar{y} \in \mathbb{R}^N$

Analyse 2 : $\underbrace{\overrightarrow{\text{grad}}\,\mathcal{L}(\bar{x})}_{A\bar{x} - \bar{b}} = \bar{0}$

$\Rightarrow$: Supposons $A\bar{x} = \bar{b}$

$\forall \bar{y} \in \mathbb{R}^N \quad \mathcal{L}(\bar{y}) = \mathcal{L}(\bar{x} + \bar{y} - \bar{x})$

$= \frac{1}{2}(\bar{x} + \bar{y} - \bar{x})^T A(\bar{x} + \bar{y} - \bar{x}) - \bar{b}^T(\bar{x} + \bar{y} - \bar{x})$

$= \frac{1}{2}\left(\bar{x}^T A\bar{x} + \bar{x}^T A(\bar{y} - \bar{x}) + (\bar{y} - \bar{x})^T A\bar{x} + (\bar{y} - \bar{x})^T A(\bar{y} - \bar{x})\right)$
$\quad - \bar{b}^T\bar{x} + \bar{b}^T(\bar{y} - \bar{x})$

or $(\bar{y} - \bar{x})^T A\bar{x} = \langle \bar{y} - \bar{x}, A\bar{x} \rangle = \langle A\bar{x}, \bar{y} - \bar{x} \rangle = \bar{x}^T A(\bar{y} - \bar{x})$

Finalement $\mathcal{L}(\bar{y}) = \mathcal{L}(\bar{x}) + (\bar{y} - \bar{x})^T(A\bar{x} - \bar{b}) + \frac{1}{2}(\bar{y} - \bar{x})^T A(\bar{y} - \bar{x})$

$A\bar{x} - \bar{b} = 0$; $\geqslant 0$ car A s.d.p

et donc $\mathcal{L}(\bar{y}) \geqslant \mathcal{L}(\bar{x}) \ \forall \bar{y} \in \mathbb{R}^N$

Chap 8 bis ine : système d'éq. non linéaires

On cherche $x_1, x_2, \dots, x_N$ tels que

$$\begin{cases} 2x_1 - x_2 + e^{x_1} = 0 \\ 2x_2 - x_1 - x_3 + e^{x_2} = 0 \\ \vdots \\ 2x_{N-1} - x_{N-2} - x_N + e^{x_{N-1}} = 0 \\ 2x_N - x_{N-1} + e^{x_N} = 0 \end{cases}$$

On écrit ces eq sous la forme

Trouver $\vec{x} \in \mathbb{R}^N$ tel que $\vec{F}(\vec{x}) = \vec{0}$ $\quad (\vec{F} : \mathbb{R}^N \to \mathbb{R}^N)$

où $\vec{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_N \end{pmatrix}$ $\quad \vec{F}(\vec{x}) = \begin{pmatrix} F_1(x_1, \dots, x_N) \\ \vdots \\ F_N(x_1, \dots, x_N) \end{pmatrix} = \begin{pmatrix} 2x_1 - x_2 + e^{x_1} \\ \vdots \\ 2x_N - x_{N-1} + e^{x_N} \end{pmatrix}$

Newton : Etant donné $\vec{x}^k \in \mathbb{R}^N$

$\vec{x}^{k+1}$ satisfait :

$$(*) \quad \underbrace{D\vec{F}(\vec{x}^k)}_{\in \mathbb{R}^{N \times N}} \underbrace{(\vec{x}^k - \vec{x}^{k+1})}_{\in \mathbb{R}^N} = \underbrace{\vec{F}(\vec{x}^k)}_{\in \mathbb{R}^N}$$

calculable inconnue connu

$\forall \vec{x} \in \mathbb{R}^N$, $D\vec{F}(\vec{x})$ matrice jacobienne

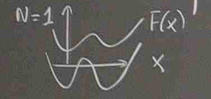
$$\left(D\vec{F}(\vec{x})\right)_{ij} = \frac{\partial F_i}{\partial x_j}(x_1, \dots, x_N) \qquad 1 \le i, j \le N$$

$$D\vec{F}(\vec{x}) = \begin{pmatrix} 2 + e^{x_1} & -1 & (0) & \\ -1 & \ddots & \ddots & \\ & \ddots & \ddots & -1 \\ (0) & & -1 & 2 + e^{x_N} \end{pmatrix}$$

sym. déf. positive donc inversible ($\forall \vec{b} \in \mathbb{R}^N \ \exists ! \vec{y} \in \mathbb{R}^N$ tq $D\vec{F}(\vec{x})\vec{y} = \vec{b}$)

2 questions :

- Existe-t-il $\vec{x} \in \mathbb{R}^N$ tq $\vec{F}(\vec{x}) = \vec{0}$?



- La solution, si elle existe, est unique ?

On va dem. l'unicité

Soit $\vec{x}$ et $\vec{y} \in \mathbb{R}^N$ tq $\vec{F}(\vec{x}) = \vec{0}$ et $\vec{F}(\vec{y}) = \vec{0}$

et montrons $\vec{x} = \vec{y}$.

On écrit $\vec{F}(\vec{x}) = A\vec{x} + \vec{e}^{\,x}$ où

$$A = \begin{pmatrix} 2 & -1 & & (0) \\ -1 & \ddots & \ddots & \\ & \ddots & \ddots & -1 \\ (0) & & -1 & 2 \end{pmatrix} \text{ et } \vec{e}^{\,x} = \begin{pmatrix} e^{x_1} \\ \vdots \\ e^{x_N} \end{pmatrix}$$

$$\begin{cases} A\vec{x} + \vec{e}^{\,x} = \vec{0} \\ A\vec{y} + \vec{e}^{\,y} = \vec{0} \end{cases}$$

$A(\vec{x} - \vec{y}) + \vec{e}^{\,x} - \vec{e}^{\,y} = \vec{0}$ produit scal avec $\vec{x} - \vec{y}$:

$$\left(\vec{x}^T\vec{y} = (x_1 \cdots x_N)\begin{pmatrix} y_1 \\ y_N \end{pmatrix} = \sum_{i=1}^N x_i y_i = \langle \vec{x}, \vec{y} \rangle\right)$$

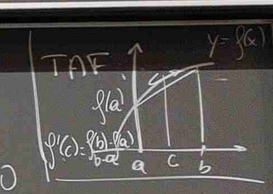
$$(\vec{x} - \vec{y})^T A(\vec{x} - \vec{y}) + \underbrace{(\vec{x} - \vec{y})^T(\vec{e}^{\,x} - \vec{e}^{\,y})}_{I} = 0$$

$$I = \sum_{i=1}^N (x_i - y_i)(e^{x_i} - e^{y_i})$$

$$= \sum_{i=1}^N (x_i - y_i)(x_i - y_i)e^{z_i} \ge 0$$

où $\min(x_i, y_i) \le z_i \le \max(x_i, y_i)$

Donc



$$(\vec{x} - \vec{y})^T A(\vec{x} - \vec{y}) \le (\vec{x} - \vec{y})^T A(\vec{x} - \vec{y}) + (\vec{x} - \vec{y})^T(\vec{e}^{\,x} - \vec{e}^{\,y}) = 0$$

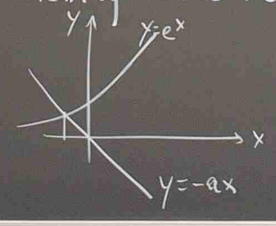
or A s.d.p : $(\vec{x} - \vec{y})^T A(\vec{x} - \vec{y}) \ge 0$

et donc $(\vec{x} - \vec{y})^T A(\vec{x} - \vec{y}) = 0$

et donc $\vec{x} - \vec{y} = \vec{0}$ ie $\vec{x} = \vec{y}$

Remarque : existence cas où $N = 1$

$x \in \mathbb{R}$ tq $ax + e^x = 0$ $\quad a > 0$



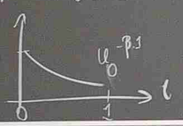
Chap 9. Eq. dif.

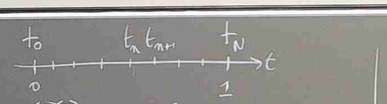
données : $u_0 \in \mathbb{R}$ $\beta > 0$

Cherché : $u(t)$ $0 < t \leq 1$

$$\begin{cases} u'(t) = -\beta u(t) & 0<t\leq 1 \\ u(0) = u_0 \end{cases}$$

$u(t) = u_0 e^{-\beta t}$





$t_n = nh$ $n = 0, 1, \dots, N$

On note u^n approx de $u(t_n)$

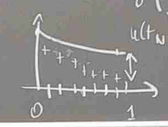
Schéma d'Euler rétrograde

$$\frac{u^{n+1} - u^n}{h} = -\beta u^{n+1} \qquad \begin{matrix} n = 0, \dots, N-1 \\ u^0 = u_0 \end{matrix}$$

$$u^{n+1} = \frac{u^n}{1+\beta h} = \frac{u_0}{(1+\beta h)^{n+1}}$$

erreur au temps final :

$$u(t_N) - u^N = u_0 \left(e^{-\beta N h} - \frac{1}{(1+\beta h)^N} \right)$$



Thm : $\exists C > 0 \ \forall u_0 \in \mathbb{R} \ \forall \beta > 0 \ \forall N > 0$

$$|u(t_N) - u^N| \leq C |u_0| \beta^2 \frac{1}{N} \qquad (\text{car } h = \tfrac{1}{N})$$

Interpret : on choisit $u_0 \in \mathbb{R}$, $\beta > 0$

l'erreur à $t = 1$ est divisée par 2 si N est mult. par 2 si h est divisé

Dem : $u(t_N) - u^N = u_0 \left(e^{-\beta h} - \frac{1}{1+\beta h} \right) \left(e^{-\beta h(N-1)} + \dots + \frac{1}{(1+\beta h)^{N-1}} \right)$

$$|u(t_N) - u^N| = |u_0| \cdot \left| e^{-\beta h} - \frac{1}{1+\beta h} \right| \cdot \left| e^{-\beta h(N-1)} + \dots + \frac{1}{(1+\beta h)^{N-1}} \right|$$

$$\leq \left| e^{-\beta h} - \frac{1}{1+\beta h} \right| \left(\left| e^{-\beta h(N-1)} \right| + \dots + \left| \frac{1}{(1+\beta h)^{N-1}} \right| \right)$$

Lemme : $\forall x > 0 \quad \left| e^{-x} - \frac{1}{1+x} \right| \leq \frac{3}{2} x^2$; d'autre part $x > 0$ $|e^{-x}| = e^{-x} \leq 1$, $\left| \frac{1}{1+x} \right| = \frac{1}{1+x} \leq 1$

Donc $|u(t_N) - u^N| \leq |u_0| \frac{3}{2} \beta^2 h^2 \ N$ (termes ≤ 1)

Mais $Nh = 1$; $|u(t_N) - u^N| \leq |u_0| \frac{3}{2} \beta^2 h$ ($h = \frac{1}{N}$)

On a dém le thm avec $C = 3/2$

Dem du lemme : Sat $x \geq 0$ $e^{-x} = 1 - x + \frac{x^2}{2} e^{-\alpha}$ $0 < \alpha < x$

$$\frac{1}{1+x} = 1 - x + \frac{x^2}{2} \frac{2}{(1+\gamma)^3} \qquad 0 < \gamma < x$$

$$e^{-x} - \frac{1}{1+x} = \frac{x^2}{2} \left(e^{-\alpha} - \frac{2}{(1+\gamma)^3} \right)$$

$$\left| e^{-x} - \frac{1}{1+x} \right| = \frac{x^2}{2} \left| e^{-\alpha} - \frac{2}{(1+\gamma)^3} \right|$$

$$\leq \frac{x^2}{2} \left(e^{-\alpha} + \left| \frac{2}{(1+\gamma)^3} \right| \right) \leq \frac{x^2}{2} (1+2) = \frac{3}{2} x^2$$

$a^N - b^N = (a-b)(a^{N-1} + a^{N-2} b + \dots + b^{N-1})$

$|a+b| \leq |a| + |b|$

$\left(\frac{1}{1+x} \right)' = -\frac{1}{(1+x)^2}$

$\left(\frac{1}{1+x} \right)'' = \frac{2}{(1+x)^3}$

$|a-b| \leq |a| + |-b| = |a| + |b|$

Que se passe-t-il si on utilise le schéma d'Euler progressif $\frac{u^{n+1} - u^n}{h} = -\beta u^n$

$u^{n+1} = u^n (1 - \beta h) = u^0 (1 - \beta h)^{n+1}$

$u^N = u_0 (1 - \beta h)^N$

$u(t_N) - u^N = u_0 \left(e^{-\beta h} - (1-\beta h) \right) \left(e^{-\beta h(N-1)} + \dots + (1-\beta h)^{N-1} \right)$

$|u(t_N) - u^N| \leq |u_0| \left| e^{-\beta h} - (1-\beta h) \right| \left(\left| e^{-\beta h(N-1)} \right| + \dots + \left| (1-\beta h)^{N-1} \right| \right)$

Lemme : $\forall x > 0 \quad \left| e^{-x} - (1-x) \right| \leq \frac{1}{2} x^2$

Condition de stabilité : $|1 - \beta h| \leq 1$

ie $-1 \leq 1 - \beta h \leq 1$

$\uparrow$ h toujours

mais si $\beta h \leq 2$ ie $h \leq \frac{2}{\beta}$

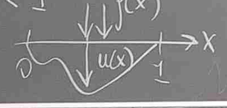
Si $\beta h \leq 2$ alors $|u(t_N) - u^N| \leq |u_0| \frac{1}{2} \beta^2 h^2 N = |u_0| \frac{1}{2} \beta^2 h$ ($hN = 1$)

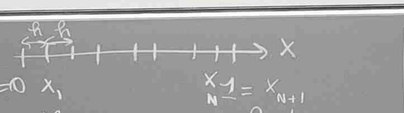
On a dém le thm avec $C = \frac{1}{2}$ et la condition $h \leq \frac{2}{\beta}$

Chap 10. Pbm aux limites 1D

Diff. finies (lundi éléments finis)

Etant donné $f : [0,1] \to \mathbb{R}$ on cherche $u : [0,1] \to \mathbb{R}$ telle que $\begin{cases} -u''(x) = f(x) & 0<x<1 \\ u(0)=0 \\ u(1)=0 \end{cases}$





But : calculer u_i approx de $u(x_i)$

$-u''(x_i) = f(x_i)$ $i=1,\dots,N$

$\frac{-u(x_{i-1})+2u(x_i)-u(x_{i+1})}{h^2} = f(x_i) + r_i$ $i=1,\dots,N$

Si $u \in \mathcal{C}^4([0,1])$, $|r_i| \leq \frac{1}{12} \max_{0\leq x\leq 1} |u^{(4)}(x)| h^2$

(Chap 2)

Schéma $\begin{cases} \frac{-u_{i-1}+2u_i-u_{i+1}}{h^2} = f(x_i) & i=1,\dots,N \\ u_0 = 0 \quad u_{N+1}=0 \end{cases}$

$\vec{u} = \begin{pmatrix} u_1 \\ \vdots \\ u_N \end{pmatrix}$ $\vec{U} = \begin{pmatrix} u(x_1) \\ \vdots \\ u(x_N) \end{pmatrix}$ $\vec{f} = \begin{pmatrix} f(x_1) \\ \vdots \\ f(x_N) \end{pmatrix}$ $A = \frac{1}{h^2}\begin{pmatrix} 2 & -1 & & \\ -1 & \ddots & \ddots & \\ & \ddots & \ddots & -1 \\ & & -1 & 2 \end{pmatrix}$

$A\vec{u} = \vec{f}$ ($\vec{g}$ donné calcule $\vec{u}$)

$A\vec{U} = \vec{f} + \vec{r}$ ($\vec{r}$ et $\vec{U}$ ne sont pas calc)

$A(\vec{U}-\vec{u}) = \vec{r}$

On peut dém

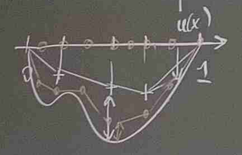
Thm : $\forall u \in \mathcal{C}^4([0,1])$ $\exists C>0$ $\forall 0<h\leq 1$

$\max_{1\leq i\leq N} |u(x_i) - u_i| \leq C h^2$

Interprétation :

On choisit f et donc u

L'erreur est divisée par $2^2=4$ si h est divisé par 2



Dem thm : facile si on admet le résultat suivant

Lemme : Soit $\vec{g} \in \mathbb{R}^N$ et soit $\vec{v}$ tq $A\vec{v} = \vec{g}$ on a :

$\|\vec{v}\|_\infty \leq \frac{1}{8}\|\vec{g}\|_\infty$

où on a noté $\|\vec{g}\|_\infty = \max_{1\leq i\leq N} |g_i|$

On applique ce lemme avec $\vec{g} = \vec{r}$ et $\vec{v} = \vec{U} - \vec{u}$

$\max_{1\leq i\leq N} |u(x_i) - u_i| = \|\vec{U}-\vec{u}\|_\infty \leq \frac{1}{8}\|\vec{r}\|_\infty \leq \frac{1}{8}\frac{1}{12} \max_{0\leq x\leq 1} |u^{(4)}(x)| h^2$

C dép de u pas de h

Dem lemme : on admet les 2 propriétés suivantes

1. $A\vec{w} = \vec{1}$ $w_i = x_i\frac{(1-x_i)}{2}$ $i=1,\dots,N$
2. $(A\vec{z} \geq \vec{0}) \Rightarrow (\vec{z} \geq \vec{0})$

force positive — def positive

On a : $|g_i| \leq \|\vec{g}\|_\infty$ $i=1,\dots,N$

$-\|\vec{g}\|_\infty \leq g_i \leq +\|\vec{g}\|_\infty$ ——

$-\vec{1}\,\|\vec{g}\|_\infty \leq \vec{g} \leq \vec{1}\,\|\vec{g}\|_\infty$

et donc

$A(\vec{w}\,\|\vec{g}\|_\infty) = \vec{1}\,\|\vec{g}\|_\infty$

$A\ \vec{N} = \vec{g}$

$A(\vec{w}\,\|\vec{g}\|_\infty - \vec{r}) = \vec{1}\,\|\vec{g}\|_\infty - \vec{g} \geq \vec{0}$

$A(\vec{w}\,\|\vec{g}\|_\infty + \vec{r}) = \vec{1}\,\|\vec{g}\|_\infty + \vec{g} \geq \vec{0}$

la propriété 2. donne

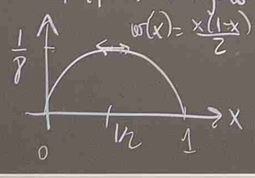
$\vec{w}\,\|g\|_\infty - \vec{r} \geq \vec{0}$

$\vec{w}\,\|g\|_\infty + \vec{r} \geq \vec{0}$

$-\vec{w}\,\|\vec{g}\|_\infty \leq \vec{r} \leq \vec{w}\,\|\vec{g}\|_\infty$

$-w_i\|\vec{g}\|_\infty \leq r_i \leq w_i\|\vec{g}\|_\infty$ $i=1,\dots,N$

$|r_i| \leq |w_i|\,\|\vec{g}\|_\infty$ ——



Donc $|v_i| \leq \frac{1}{8}\|\vec{g}\|_\infty$ $i=1,\dots,N$

et $\|\vec{v}\|_\infty = \max_{1\leq i\leq N} |v_i| \leq \frac{1}{8}\|\vec{g}\|_\infty$

on a dém le lemme.

Retour à la propriété 1 ;

Soit $\vec{w}(x)$ tq $\begin{cases} -w''(x) = 1 & 0<x<1 \\ w(0)=0 \quad w(1)=0 \end{cases}$

$w(x) = \frac{x(1-x)}{2}$

$-w''(x_i) = 1$ $i=1,\dots,N$

$\frac{-w(x_{i-1})+2w(x_i)-w(x_{i+1})}{h^2} = 1 + \text{reste}$

mais le reste vaut zéro car $w^{(4)}(x) = 0$ $0\leq x\leq 1$

On a donc bien

$A\vec{w} = \vec{1}$ où $w_i = x_i\frac{(1-x_i)}{2}$ $\widetilde{w(x_i)}$

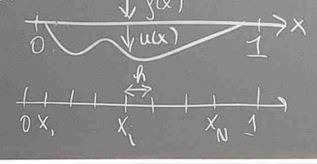
propriété 2 admise

Chap 10 : pbm aux limites 1d

leçon 7 MOOC : différences finies (1900)
(10.2 et 10.6 livre)
Aug 10.3 et 10.4 livre
Éléments finis (1950)

Rappel diff. finies

$$\begin{cases} -u''(x) = f(x) & 0<x<1 \\ u(0)=0 \quad u(1)=0 \end{cases}$$



$x_i = ih \quad i = 0, 1, \dots, N+1 \quad h = \frac{1}{N+1}$

u_i approx $u(x_i)$

$A\vec{u} = \vec{f}$

$$\frac{1}{h^2}\begin{pmatrix} 2 & -1 & & \\ -1 & 2 & -1 & \\ & \ddots & \ddots & \ddots \\ & & -1 & 2 \end{pmatrix}\begin{pmatrix} u_1 \\ \vdots \\ u_N \end{pmatrix} = \begin{pmatrix} f(x_1) \\ \vdots \\ f(x_N) \end{pmatrix}$$

Aug :
1_ Formulation variationnelle
2_ Méthode de Galerkin (Ritz)
3_ Méthode des éléments finis
pour aboutir au même syst. lin !
Très utilisée en pratique

1_ On multiplie (1) par $v(x)$
où $v : [0,1] \to \mathbb{R}$ est quelconque ($\mathcal{C}^1$)
et on intègre x entre 0 et 1

$$\int_0^1 -u''(x)v(x)\,dx = \int_0^1 f(x)\,v(x)\,dx$$

On intègre par parties

$(uv)' = u''v + uv'$

$[uv] = \int u'v + \int uv'$

$u \to u'$

$$[u'v] = \int u''v + \int u'v'$$

On obtient

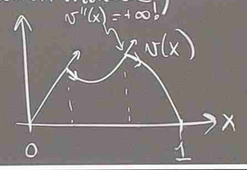
$$\int_0^1 u'(x)v'(x)\,dx - \underbrace{\left[u'(x)v(x)\right]_{x=0}^{x=1}}_{u'(1)v(1) - u'(0)v(0)} = \int_0^1 f(x)v(x)\,dx$$

Si on choisit v telle que $v(0)=0$ et $v(1)=0$, on obtient la formulation variationnelle :

On cherche $u \in V$ telle que $\int_0^1 u'(x)v'(x)\,dx = \int_0^1 f(x)v(x)\,dx \quad \forall v \in V$ (2)

(formulation variationnelle, faible, principe des travaux virtuels, v déformation virtuelle)

V ? (espace des déf admissibles)

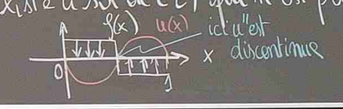


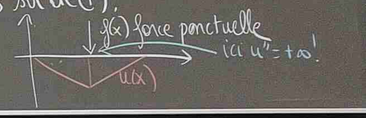
$$V = \{v : [0,1] \to \mathbb{R},\ v \text{ continue},\ v' \text{ cont. par morceaux},\ v(0)=0 \text{ et } v(1)=0\}$$

C'est un espace vectoriel, $\dim V = +\infty$ (pas implémentable)

Remarques :
- Si u est solution de (1) alors u est solution de (2)
- Il existe u sol de (2) qui n'est pas sol de (1) ;

Ex :



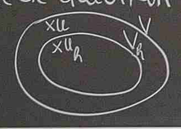


- Si on choisit $v = u$ dans (2), on obtient

$$\underbrace{\int_0^1 (u'(x))^2\,dx}_{\text{énergie de déformation}} = \underbrace{\int_0^1 f(x)u(x)\,dx}_{\text{travail des forces}}$$

2_ Méthode de Galerkin

$u_h \xrightarrow{h \to 0} u$?



$V_h \subset V \quad \dim V_h = N$
N grand
$h = \frac{1}{N+1}$ petit

Soit $\varphi_1, \varphi_2, \dots, \varphi_N \in V$ linéairement indépendantes (données)
Soit $V_h = \text{vec}(\varphi_1, \varphi_2, \dots, \varphi_N)$ (l'ens des C.L)
$\dim V_h = N$
On cherche $u_h \in V_h$ tq $\int_0^1 u_h'(x)v_h'(x)\,dx = \int_0^1 f(x)v_h(x)\,dx \quad \forall v_h \in V_h$ (3)

Trouver u_h solution de (3) est équivalent à résoudre un système linéaire

En effet, $u_h \in V_h$: $u_h(x) = \sum_{j=1}^{N} u_j \varphi_j(x) \quad 0 \le x \le 1$
(u_j : inconnues $u_1, \dots, u_N$ N inconnues)
On choisit $v_h = \varphi_i \quad i = 1, \dots, N$ dans (3), N équations

(3) devient :

$$\int_0^1 \sum_{j=1}^{N} u_j \varphi_j'(x)\varphi_i'(x)\,dx = \int_0^1 f(x)\varphi_i(x)\,dx \quad i=1,\dots,N$$

$$\sum_{j=1}^{N} u_j \underbrace{\int_0^1 \varphi_j'(x)\varphi_i'(x)\,dx}_{A_{ij}} = \underbrace{\int_0^1 f(x)\varphi_i(x)\,dx}_{f_i} \quad i=1,\dots,N$$

$$\sum_{j=1}^{N} A_{ij}u_j = f_i \quad i=1,\dots,N$$

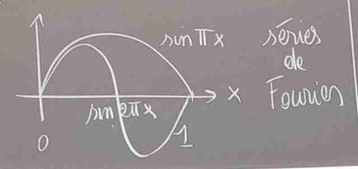
$A\vec{u} = \vec{f}$ (on peut montrer que A est symétrique définie positive)

u_j : Calculables, f_i connus

Reste à choisir $\varphi_1, \varphi_2, \dots, \varphi_N$

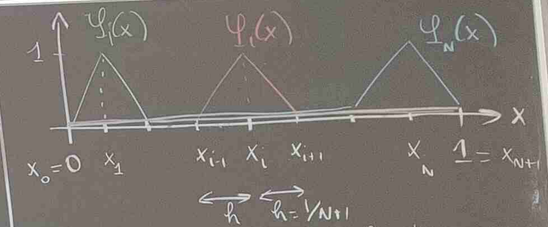
Plusieurs possibilités

- $\varphi_i(x) = \sin(i\pi x)$ $i=1,\dots,N$



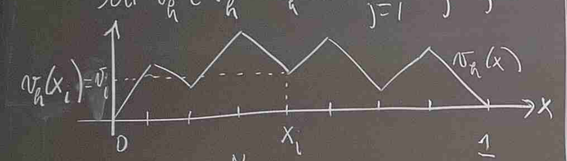
séries de Fourier

- $\varphi_i(x) = x^i(1-x)$ $i=1,\dots,N$
- Eléments finis, continus, de degré 1 par morceaux



$x_i = ih \quad i = 0,\dots,N+1 \quad h = \frac{1}{N+1}$

Soit $v_h \in V_h$ $\quad v_h(x) = \sum_{j=1}^{N} v_j \varphi_j(x)$



On a $v_h(x_i) = \sum_{j=1}^{N} v_j \underbrace{\varphi_j(x_i)}_{=1 \text{ si } j=i,\ 0 \text{ sinon}} = v_i$

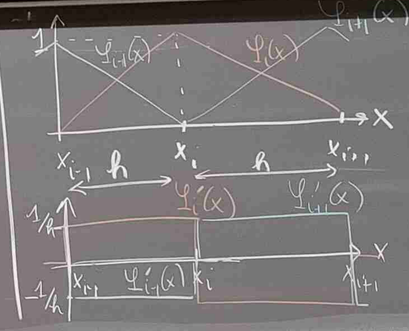
Reste à calculer A et $\vec{f}$

La matrice A est creuse (beaucoup de zéros)

Ex : $A_{N1} = \int_0^1 \varphi_1'(x)\varphi_N'(x)\,dx = 0$

la matrice A est symétrique, tridiagonale

Reste à calculer $A_{i,i}$ $A_{i,i-1}$ et f_i



$$A_{i,i} = \int_0^1 (\varphi_i'(x))^2\,dx = \int_{x_{i-1}}^{x_i} (\varphi_i'(x))^2\,dx + \int_{x_i}^{x_{i+1}} (\varphi_i'(x))^2\,dx$$

$$= h\left(\frac{1}{h}\right)^2 + h\left(-\frac{1}{h}\right)^2 = \frac{2}{h}$$

$$A_{i,i-1} = \int_0^1 \varphi_i'(x)\varphi_{i-1}'(x)\,dx = \int_{x_{i-1}}^{x_i} \varphi_i'(x)\varphi_{i-1}'(x)\,dx = h\,\frac{1}{h}\left(-\frac{1}{h}\right) = -\frac{1}{h}$$

$$f_i = \int_0^1 f(x)\varphi_i(x)\,dx = \int_{x_{i-1}}^{x_i} f(x)\varphi_i(x)\,dx + \int_{x_i}^{x_{i+1}} f(x)\varphi_i(x)\,dx$$

On utilise la formule de quadrature du trapèze

$$\int_{x_i}^{x_{i+1}} g(x)\,dx \simeq \frac{h}{2}\left(g(x_i) + g(x_{i+1})\right)$$

On obtient l'approximation suivante de f_i

$$\frac{h}{2}\Big(f(x_{i-1})\underbrace{\varphi_i(x_{i-1})}_{0} + f(x_i)\underbrace{\varphi_i(x_i)}_{1}\Big) + \frac{h}{2}\Big(f(x_i)\underbrace{\varphi_i(x_i)}_{1} + f(x_{i+1})\underbrace{\varphi_i(x_{i+1})}_{0}\Big)$$

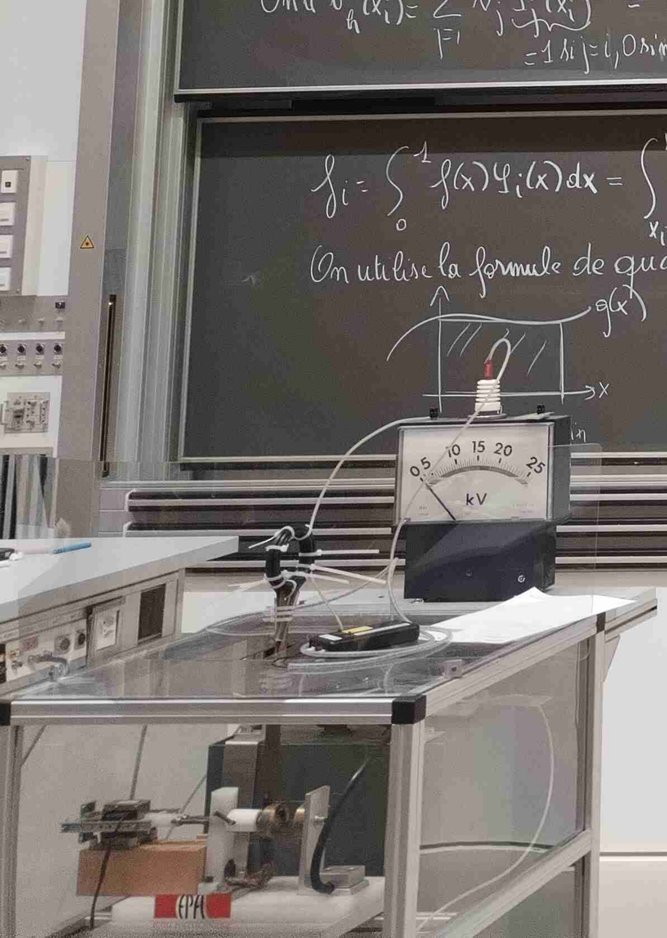
$$= h\,f(x_i)$$

Le système linéaire à résoudre est

$$\frac{1}{h}\begin{pmatrix} 2 & -1 & (0) & & \\ -1 & & \ddots & & \\ & \ddots & & \ddots & -1 \\ (0) & & -1 & & 2 \end{pmatrix}\begin{pmatrix} u_1 \\ \vdots \\ u_N \end{pmatrix} = h\begin{pmatrix} f(x_1) \\ \vdots \\ f(x_N) \end{pmatrix}$$

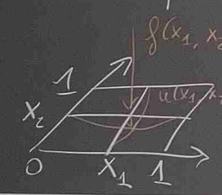
La solution est celle de la méthode des différence finies

dimension	équation	diff. finies	éléments finis
1	$-u''(x) = f(x)$	a b	a b
2	$-\Delta u(x,y) = f(x,y)$	1 1 x 0 1	
3	$-\Delta u(x,y,z) = f(x,y,z)$		



Chap II : pbm aux limites 2D

Remarque II.2



membrane élastique
Etant donnée
$f : [0,1]\times[0,1] \to \mathbb{R}$
$(x_1, x_2) \to f(x_1, x_2)$

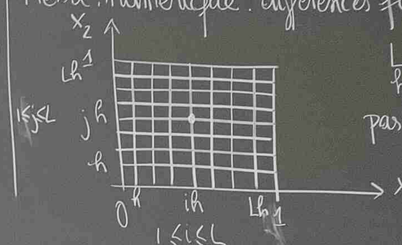
On cherche $u : [0,1]\times[0,1] \to \mathbb{R}$ telle que
$(x_1, x_2) \to u(x_1, x_2)$

$$\begin{cases} -\Delta u(x_1,x_2) = f(x_1,x_2) & \forall (x_1,x_2) \in \Omega =]0,1[^2 \\ u(x_1,x_2) = 0 & \forall (x_1,x_2) \in \partial\Omega \end{cases}$$

ici $\Delta u = \operatorname{div}(\operatorname{grad} u) = \vec{\nabla} \cdot \vec{\nabla} u \quad \left(\vec{\nabla} = \begin{pmatrix} \frac{\partial}{\partial x_1} \\ \frac{\partial}{\partial x_2} \end{pmatrix}\right)$
$= \frac{\partial^2 u}{\partial x_1^2} + \frac{\partial^2 u}{\partial x_2^2}$

Electrostatique :
$\operatorname{div} \vec{E} = \rho$ ← densité charge
champ élect
$\operatorname{rot} \vec{E} = \vec{0}$ $\quad \vec{E} = -\operatorname{grad} \varphi$ ← potentiel
$-\operatorname{div} \operatorname{grad} \varphi = \rho$

Résol. numérique : différences finies



$\frac{L}{h} = \frac{1}{L+1}$ pas d'espace

But. calculer $U_{i,j}$: approx $u(ih, jh)$ $\quad 1 \le i,j \le L$

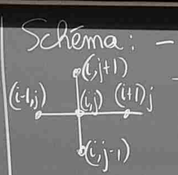
On écrit l'éq. au point (ih, jh) : $-\Delta u(ih, jh) = f(ih, jh)$

$$\frac{\partial^2 u}{\partial x_1^2}(ih, jh) = \frac{u((i-1)h, jh) - 2u(ih, jh) + u((i+1)h, jh)}{h^2} + O(h^2) \quad (u \in \mathcal{C}^4([0,1]^2))$$

$$\frac{\partial^2 u}{\partial x_2^2}(ih, jh) = \frac{u(ih, (j-1)h) - 2u(ih, jh) + u(ih, (j+1)h)}{h^2} + O(h^2)$$

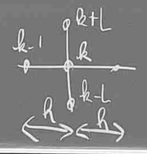
Pour obtenir le schéma, on remplace $u(ih, jh)$ par $U_{i,j}$ et on oublie les $O(h^2)$

Schéma : $\dfrac{-U_{i-1,j} - U_{i,j-1} + 4U_{i,j} - U_{i+1,j} - U_{i,j+1}}{h^2} = f(ih, jh) \quad 1 \le i,j \le L$

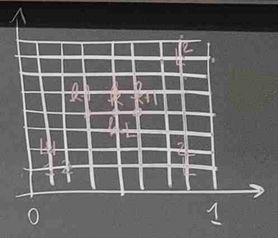


$U_{i,j} = 0$ si $i = 0$ ou $i = L+1$ ou $j = 0$ ou $j = L+1$

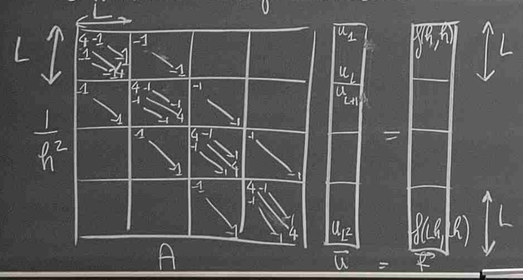
Numérotation par ligne



$U_{i,j} \to u_k$



On obtient le syst. linéaire suivant



On peut montrer que si $u \in \mathcal{C}^4([0,1]^2)$ alors

$$\max_{1 \le i,j \le L} |U_{i,j} - u(ih, jh)| = O(h^2)$$

Implémentation sur matlab :

$$A = \frac{1}{h^2}\left(D_L \otimes I_L + I_L \otimes D_L\right)$$

où $D_L = \begin{pmatrix} 2 & -1 & & (0) \\ -1 & \ddots & \ddots & \\ & \ddots & \ddots & -1 \\ (0) & & -1 & 2 \end{pmatrix}$ $\quad I_L = \begin{pmatrix} 1 & & (0) \\ & \ddots & \\ (0) & & 1 \end{pmatrix}$ (L × L)

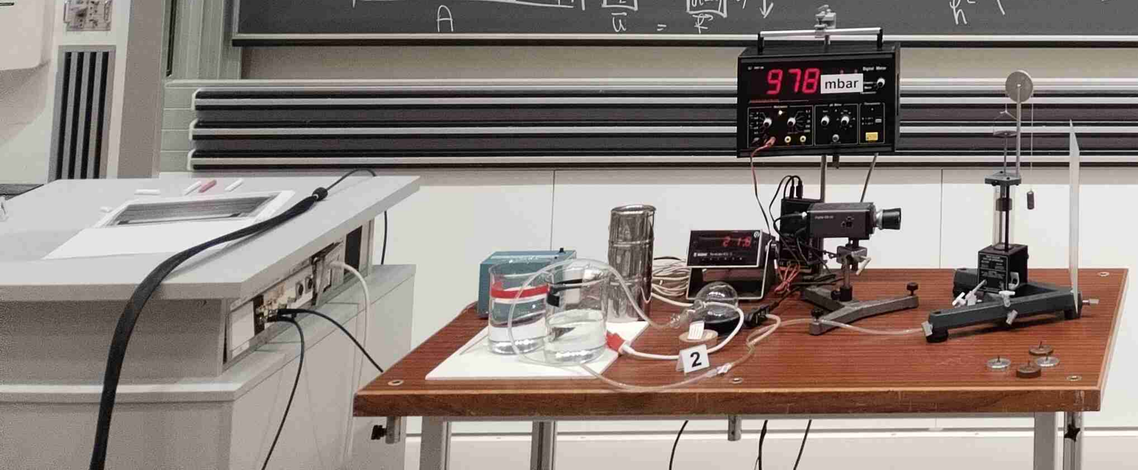
et $\otimes$ est le produit kronecker (kron)

$B, C \in \mathbb{R}^{L \times L}$

$$B \otimes C = \begin{pmatrix} b_{11}C & \cdots & b_{1L}C \\ \vdots & & \vdots \\ b_{L1}C & \cdots & b_{LL}C \end{pmatrix} \quad (L^2 \times L^2)$$

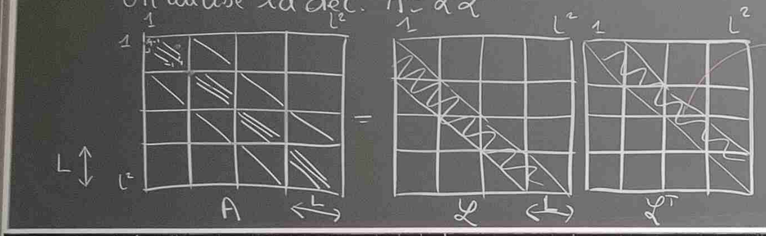
$$A = \frac{1}{h^2}\left(\begin{pmatrix} 2I_L & -I_L & & \\ -I_L & \ddots & \ddots & \\ & \ddots & \ddots & -I_L \\ & & -I_L & 2I_L \end{pmatrix} + \begin{pmatrix} D_L & & & \\ & D_L & & \\ & & \ddots & \\ & & & D_L \end{pmatrix} \right)$$

Mémoire ? Nombre d'opérations ? (complexité)





Pour résoudre le système linéaire $A\vec{u} = \vec{f}$
on utilise la déc. $A = \mathcal{L}\mathcal{L}^T$



la zone hachurée correspond aux éléments non-nuls de $\mathcal{L}$ et $\mathcal{L}^T$

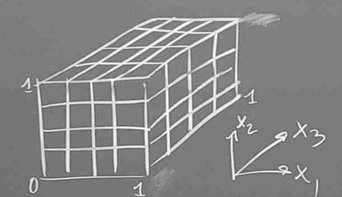
nombre d'inconnues : L^2
mémoire (pour $\mathcal{L}$) : $O(L^2 L) = O(L^3)$
nombre d'opérations : $O(L^2 L^2) = O(L^4)$
alg. $\mathcal{L}\mathcal{L}^T$: $i = 1, L^2$
$j = 1, L$
$k = 1, L$

$L = 10^2$	$L = 2 \cdot 10^2$	$L = 10^3$
$L^2 = 10^4$	$L^2 = 4\,10^4$	$L^2 = 10^6$
$L^3 = 10^6$	$L^3 = 8\,10^6$	$L^3 = 10^9$
$L^4 = 10^8$	$L^4 = 16\,10^8$	$L^4 = 10^{12}$
$\sim 1\,s$	$\sim 16\,s$	$10^4\,s$!

Calculer en dimension 3 :

$$\begin{cases} -\Delta u(x_1,x_2,x_3) = f(x_1,x_2,x_3) & (x_1,x_2,x_3) \in \Omega =]0,1[^3 \\ u(x_1,x_2,x_3) = 0 & \text{sur } \partial\Omega \end{cases}$$

$$\Delta u = \frac{\partial^2 u}{\partial x_1^2} + \frac{\partial^2 u}{\partial x_2^2} + \frac{\partial^2 u}{\partial x_3^2}$$



L points dans chaque direction
$h = \frac{1}{L+1}$

L^3 inconnues $U_{i,j,k}$ approx $u(ih, jh, kh)$ $1 \le i,j,k \le L$

Schéma $\dfrac{-U_{i-1,j,k} - U_{i,j-1,k} - U_{i,j,k-1} + 6U_{i,j,k} - U_{i+1,j,k} - U_{i,j+1,k} - U_{i,j,k+1}}{h^2} = f(ih, jh, kh)$ $1 \le i,j,k \le L$

$U_{i,j,k} = 0$ si $i = 0$ ou $i = L+1$
$j = 0$ $j = L+1$
$k = 0$ $k = L+1$

mémoire ? nombre d'opérations ?

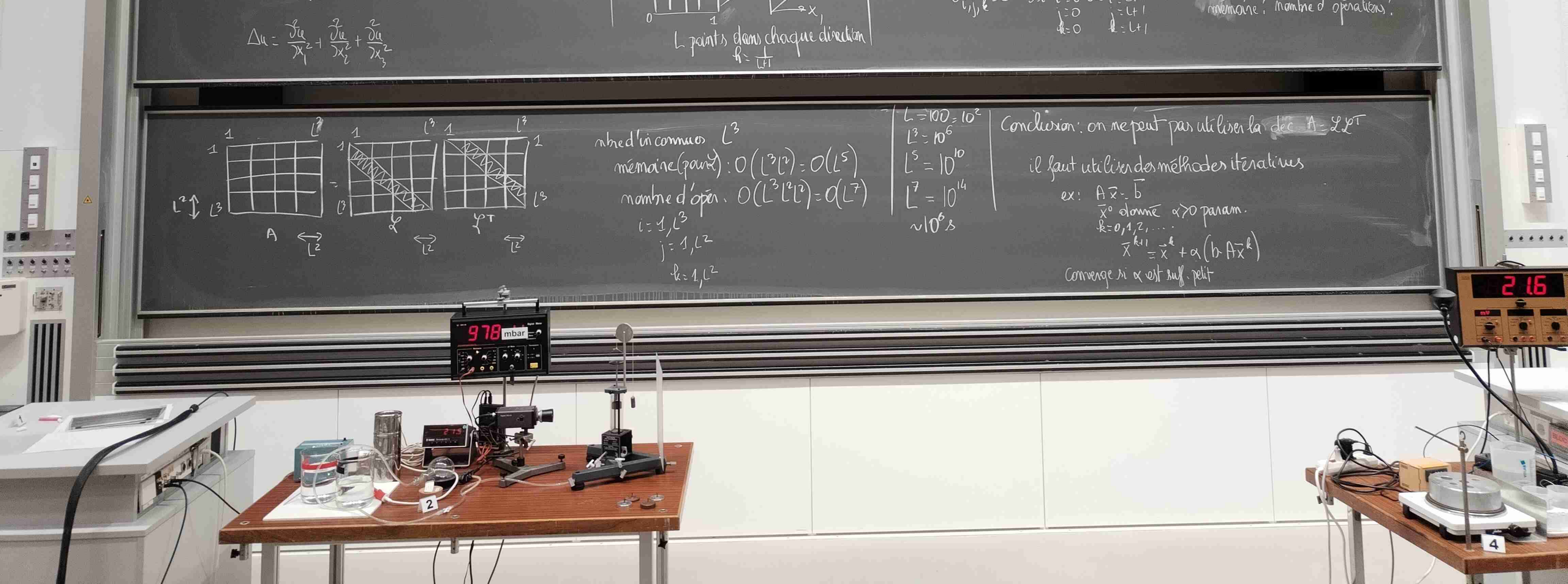
nbre d'inconnues L^3
mémoire (pour $\mathcal{L}$) : $O(L^3 L^2) = O(L^5)$
nombre d'opér. $O(L^3 L^2 L^2) = O(L^7)$
$i = 1, L^3$
$j = 1, L^2$
$k = 1, L^2$

$L = 100 = 10^2$
$L^3 = 10^6$
$L^5 = 10^{10}$
$L^7 = 10^{14}$
$\sim 10^6\,s$

Conclusion : on ne peut pas utiliser la déc. $A = \mathcal{L}\mathcal{L}^T$

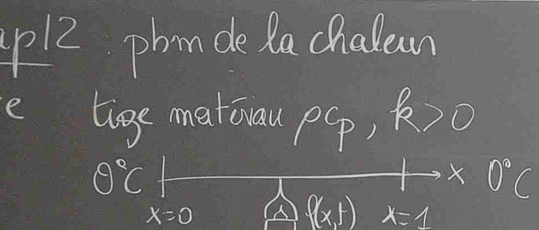
il faut utiliser des méthodes itératives

ex : $A\vec{x} = \vec{b}$
$\vec{x}^0$ donné $\alpha > 0$ param.
$k = 0, 1, 2, \ldots$
$\vec{x}^{k+1} = \vec{x}^k + \alpha(\vec{b} - A\vec{x}^k)$
Converge si α est suff. petit



Chap 12 : phm de la chaleur

Expérience : tige matériau ρc_p, $k>0$



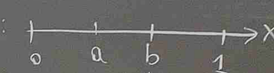
$t=0$ temp. $w(x)$ $\quad u(x,t)$ temp ?

On cherche $u(x,t)$ (température) telle que

$$\begin{cases} \rho c_p \dfrac{\partial u}{\partial t}(x,t) - k \dfrac{\partial^2 u}{\partial x^2}(x,t) = f(x,t) & 0<x<1,\ t>0 \\ u(0,t) = 0 \quad u(1,t) = 0 \ \text{(cond. aux limites)} & t>0 \\ u(x,0) = w(x) \ \text{(cond. initiale)} & 0<x<1 \end{cases}$$

Remarques : • L'éq. de la chaleur traduit un principe de conservation (thermique)

En effet :



intégrons entre $x=a$ et $x=b$

$$\int_a^b \rho c_p \frac{\partial u}{\partial t}(x,t)\,dx - \int_a^b k \frac{\partial^2 u}{\partial x^2}(x,t)\,dx = \int_a^b f(x,t)\,dx$$

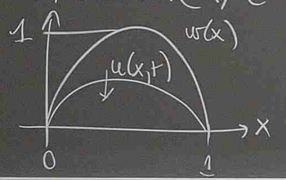
$$\frac{d}{dt}\int_a^b \rho c_p u(x,t)\,dx - \Big[k \frac{\partial u}{\partial x}(x,t)\Big]_{x=a}^{x=b} = \int_a^b f(x,t)\,dx$$

Variation au cours du temps : enthalpie ; flux de chaleur entrant ou sortant ; énergie thermique (source)

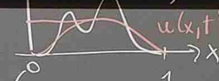
• Si $f(x,t)=0$, $\rho c_p = 1$ (dans la suite)

$w(x) = \sin(\pi x)$ alors $u(x,t)$ est donnée par

$$u(x,t) = \sin(\pi x)\, e^{-k\pi^2 t}$$

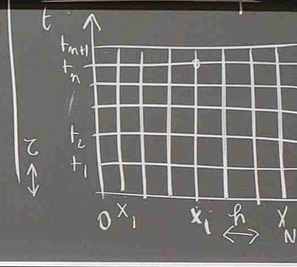


• Si $f(x,t)=0$ alors $u(x,t)$ satisfait les 2 propriétés suiv. (qu'on va essayer de reproduire num.)



- positivité : si $(w(x)\geq 0,\ 0<x<1)$ alors $(u(x,t)\geq 0,\ 0<x<1)$
- principe du maximum : $\max_{0\leq x\leq 1} |u(x,t)| \leq \max_{0\leq x\leq 1} |w(x)|$

Méthode numérique :



But : calculer u_i^n approx $u(x_i, t_n)$

Notations : h pas d'espace

$h = \frac{1}{N+1}$ $\quad x_i = ih$ $\quad i = 0, 1, \dots, N, N+1$

τ pas de temps

$t_n = n\tau$ $\quad n = 0, 1, 2, \dots$

Marche en temps :

$n=0$ $\quad u_i^0 = w(x_i)$ $\quad i=1,\dots,N$ cond. initiale

$n = 0, 1, 2, \dots$

Etant donné les u_i^n, $i=1,\dots,N$

Comment calculer les u_i^{n+1} $i=1,\dots,N$?

Dans le livre, on discrétise d'abord en espace (x)

On aboutit à un syst. différentiel $\vec{u}'(t) + A\vec{u}(t) = \vec{f}(t)$

Puis on discrétise en temps (t) on aboutit au schéma

$$\frac{\vec{u}^{n+1} - \vec{u}^n}{\tau} + A\vec{u}^n = \vec{f}(t_n)$$

Ici on discrétise en x et t en une fois :

On écrit l'éq. de la chaleur en x_i, t_n

$$\underbrace{\frac{\partial u}{\partial t}(x_i,t_n)}_{\text{FDF progressive}} - k \underbrace{\frac{\partial^2 u}{\partial x^2}(x_i,t_n)}_{\text{FDF centrée}} = f(x_i,t_n)$$

$$\frac{u(x_i,t_{n+1}) - u(x_i,t_n)}{\tau} - k\,\frac{u(x_{i-1},t_n) - 2u(x_i,t_n) + u(x_{i+1},t_n)}{h^2} = f(x_i,t_n) + O(\tau) + O(h^2)$$

Pour obtenir le schéma, on remplace $u(x_i,t_n)$ par u_i^n et on oublie les $O(\tau)$ et $O(h^2)$

On obtient le schéma d'Euler progressif

$$\frac{u_i^{n+1} - u_i^n}{\tau} - k\,\frac{u_{i-1}^n - 2u_i^n + u_{i+1}^n}{h^2} = f(x_i,t_n) \qquad i=1,\dots,N$$

$n=0,1,2,\dots$ fixé $\quad u_0^n = 0 \quad u_{N+1}^n = 0$

Ce schéma est explicite

$u_i^{n+1} = \left(1 - 2\frac{\tau k}{h^2}\right) u_i^n + \frac{\tau k}{h^2}\left(u_{i-1}^n + u_{i+1}^n\right) + \tau f(x_i, t_n)$

On peut écrire ce schéma comme: $i = 1, \dots, N$

$\frac{\vec{u}^{n+1} - \vec{u}^n}{\tau} + A\vec{u}^n = \vec{f}(t_n)$

ou $\vec{u}^{n+1} = (I - \tau A)\vec{u}^n + \tau \vec{f}(t_n)$ où on a noté

$\vec{u}^n = \begin{pmatrix} u_1^n \\ \vdots \\ u_N^n \end{pmatrix}$ $\vec{f}(t) = \begin{pmatrix} f(x_1, t) \\ \vdots \\ f(x_N, t) \end{pmatrix}$ $A = \frac{k}{h^2}\begin{pmatrix} 2 & -1 & & \\ -1 & \ddots & \ddots & \\ & \ddots & \ddots & -1 \\ & & -1 & 2 \end{pmatrix}$

avantage : facile à programmer
inconvénient : cond. de stabilité

Schéma d'Euler rétrograde

On écrit l'eq. de la chaleur en x_i, t_{n+1} :

$\underbrace{\frac{\partial u}{\partial t}(x_i, t_{n+1})}_{\text{FDF rétrograde}} - k \underbrace{\frac{\partial^2 u}{\partial x^2}(x_i, t_{n+1})}_{\text{FDF centrée}} = f(x_i, t_{n+1})$

On obtient

$\begin{cases} \frac{u_i^{n+1} - u_i^n}{\tau} - k\frac{u_{i-1}^{n+1} - 2u_i^{n+1} + u_{i+1}^{n+1}}{h^2} = f(x_i, t_{n+1}) & i = 1, \dots, N \\ u_0^{n+1} = 0 \quad u_{N+1}^{n+1} = 0 \end{cases}$

On peut écrire ce schéma comme :

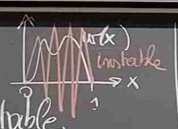
$\frac{\vec{u}^{n+1} - \vec{u}^n}{\tau} + A\vec{u}^{n+1} = \vec{f}(t_{n+1})$

$(I + \tau A)\vec{u}^{n+1} = \vec{u}^n + \tau \vec{f}(t_{n+1})$

inconvénient : inconnu — il faut résoudre un syst. linéaire — connu

avantage : pas de cond. de stabilité

Stabilité : $f(x,t) = 0$



Def : On dira que le schéma est stable si les 2 cond. suivantes sont satisfaites

(positivité) 1. si $(u_i^n \geq 0 \quad i = 1, \dots, N)$ alors $(u_i^{n+1} \geq 0 \quad i = 1, \dots, N)$

(principe du max.) 2. $\max_{1 \leq i \leq N} |u_i^{n+1}| \leq \max_{1 \leq i \leq N} |u_i^n|$

Lemme :

- le schéma d'Euler progressif $\left(\frac{\vec{u}^{n+1} - \vec{u}^n}{\tau} + A\vec{u}^n = 0\right)$ est stable si $\tau \leq \frac{h^2}{2k}$
- le schéma d'Euler rétrograde $\left(\frac{\vec{u}^{n+1} - \vec{u}^n}{\tau} + A\vec{u}^{n+1} = 0\right)$ est stable $\forall h, \tau > 0$

Remarques :

- les 2 schémas convergent en $O(\tau)$ et $O(h^2)$ (cf exercices)
- la condition de stabilité $\tau \leq \frac{h^2}{2k}$ est pénalisante :

Si N est multiplié par 2, h est divisé par 2
τ doit être divisé par 4, le nombre de pas de temps doit être multiplié par 4, $(2 \times 4 = 8$ fois plus d'opér.$)$

Dem. du lemme pour le schéma d'Euler explicite

Supposons $\tau \leq \frac{h^2}{2k}$ et montrons que les 2 propriétés 1 et 2 sont satisfaites

propriété 1 : supposons $u_i^n \geq 0 \quad i = 1, \dots, N$ et montrons $u_i^{n+1} \geq 0$

le schéma s'écrit $u_i^{n+1} = \left(1 - 2\frac{\tau k}{h^2}\right) u_i^n + \frac{\tau k}{h^2}\left(u_{i-1}^n + u_{i+1}^n\right)$

pour $i = 1, \dots, N$ avec $u_0^n = 0 = u_{N+1}^n$

Si $1 - 2\frac{\tau k}{h^2} \geq 0$ alors $u_i^{n+1} \geq 0$

On $1 - 2\frac{\tau k}{h^2} \geq 0 \iff 1 \geq 2\frac{\tau k}{h^2} \iff \tau \leq \frac{h^2}{2k}$ qui est bien satisfaite

propriété 2 : pour $i = 1, \dots, N$

$|u_i^{n+1}| = \left|\left(1 - 2\frac{\tau k}{h^2}\right) u_i^n + \frac{\tau k}{h^2}\left(u_{i-1}^n + u_{i+1}^n\right)\right|$

$\leq \left|\left(1 - 2\frac{\tau k}{h^2}\right) u_i^n\right| + \left|\frac{\tau k}{h^2}\left(u_{i-1}^n + u_{i+1}^n\right)\right|$

$\leq \underbrace{\left|1 - 2\frac{\tau k}{h^2}\right|}_{= 1 - 2\frac{\tau k}{h^2}} |u_i^n| + \underbrace{\left|\frac{\tau k}{h^2}\right|}_{= \frac{\tau k}{h^2}} \left(|u_{i-1}^n| + |u_{i+1}^n|\right)$

$|u_i^n|, |u_{i-1}^n|, |u_{i+1}^n| \leq \max_{1 \leq j \leq N} |u_j^n| \quad i = 1, \dots, N$

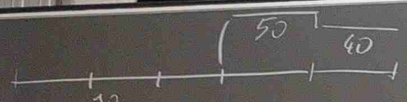
(si $i = 1$ $0 = |u_0^n| \leq$ ——— si $i = N$ $0 = |u_{N+1}^n| \leq$ ———)

Donc

$|u_i^{n+1}| \leq \underbrace{\left(1 - 2\frac{\tau k}{h^2} + \frac{\tau k}{h^2} + \frac{\tau k}{h^2}\right)}_{1} \max_{1 \leq j \leq N} |u_j^n| \quad i = 1, \dots, N$

$|a| = a$ si $a \geq 0$

A fortiori vrai pour l'indice i qui réalise le max des $|u_i^{n+1}|$



Chap 13 transport, ondes

13.1 Transport

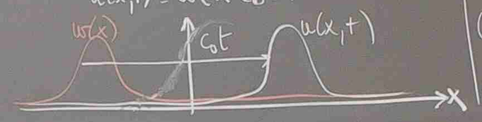
Exp : transport d'un polluant

$t=0$ concentration polluant $w(x)$ $\longrightarrow x \in \mathbb{R}$ (rivière)
Source de polluant $f(x,t)$ usine
vitesse de transport $c(x,t)$
on cherche $u(x,t)$ conc. polluant telle que

$$\begin{cases} \frac{\partial u}{\partial t}(x,t) + c(x,t)\frac{\partial u}{\partial x}(x,t) = f(x,t) & x \in \mathbb{R}, t>0 \\ u(x,0) = w(x) \end{cases}$$

Remarques :

- Si $f(x,t)=0$, $c(x,t)=c_0$ constante $\in \mathbb{R}$ la solution est donnée par $u(x,t) = w(x-c_0 t)$



vérification : $u(x,0) = w(x-0)$

$\frac{\partial u}{\partial t}(x,t) = w'(x-c_0 t)(-c_0)$

$\left(\frac{\partial u}{\partial x}(x,t) = w'(x-c_0 t)\,1\right) c_0$

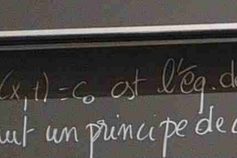
$\frac{\partial u}{\partial t} + c_0 \frac{\partial u}{\partial x} = 0$

On vérifie que u satisfait

- positivité : si $w(x) \geq 0 \ \forall x \in \mathbb{R}$ alors $u(x,t) \geq 0 \ \forall x \in \mathbb{R}$
- principe du maximum : $\sup_{x\in\mathbb{R}} |u(x,t)| \leq \sup_{x\in\mathbb{R}} |w(x)|$

On va essayer de reproduire ces 2 ppties numériquement

- Si $c(x,t)=c_0$ est l'éq. de transport traduit un principe de conservation



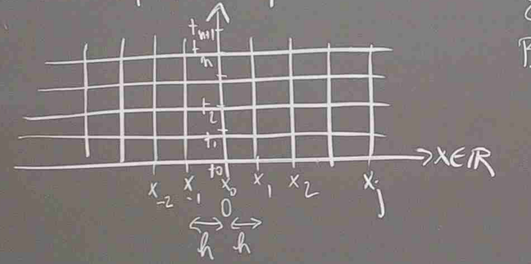
$$\int_a^b \frac{\partial u}{\partial t}(x,t)\,dx + c_0 \int_a^b \frac{\partial u}{\partial x}(x,t)\,dx = \int_a^b f(x,t)\,dx$$

source de polluant

$$\frac{d}{dt}\int_a^b u(x,t)\,dx + \left[c_0 u(x,t)\right]_{x=a}^{x=b} = \int_a^b f(x,t)\,dx$$

Variation du cours du temps — qté polluant — flux polluant entrant ou sortant

Grille espace temps



h pas d'espace $x_j = jh$ $j \in \mathbb{Z}$
τ pas de temps $t_n = n\tau$ $n = 0,1,2,\dots$
But : calculer u_j^n approx $u(x_j, t_n)$
marche en temps
$t=0$ $u_j^0 = w(x_j) = w(jh)$ $j \in \mathbb{Z}$
Etant donné u_j^n $j \in \mathbb{Z}$, comment calculer u_j^{n+1} $j \in \mathbb{Z}$?

2 schémas explicites

- centré : toujours instable ! $\forall h, \tau > 0$
- décentré : stable sous condition entre h et τ

- centré : on écrit l'éq en x_j, t_n

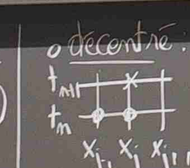
$$\underbrace{\frac{\partial u}{\partial t}(x_j,t_n)}_{\text{FDF progressive}} + c(x_j,t_n)\underbrace{\frac{\partial u}{\partial x}(x_j,t_n)}_{\text{FDF centrée}} = f(x_j,t_n)$$

$$\frac{u(x_j,t_{n+1}) - u(x_j,t_n)}{\tau} + c(x_j,t_n)\frac{u(x_{j+1},t_n) - u(x_{j-1},t_n)}{2h} = f(x_j,t_n) \quad j \in \mathbb{Z}$$
$$+ O(\tau) + O(h^2) \ (\text{si } f \text{ fixée})$$

Schéma : $\frac{u_j^{n+1} - u_j^n}{\tau} + c(x_j,t_n)\frac{u_{j+1}^n - u_{j-1}^n}{2h} = f(x_j,t_n)$

Instable : $f(x,t)=0$ $c(x,t)=c_0$ $u(x,t)=w(x-c_0t)$
$w(x) = \cos(mx)$ $|u(x,t)| = |\cos(m(x-c_0t))| \leq 1$
On peut montrer (cf ligne) que $|u_j^n| \xrightarrow{n\to+\infty} +\infty$ $\forall h, \tau > 0$

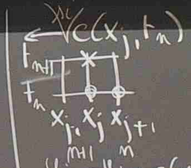
- décentré : $\times$: on veut calculer u_j^{n+1}



$\rightarrow$ si $c(x_j,t_n) \geq 0$
o : on va utiliser ces points

$$\frac{u_j^{n+1} - u_j^n}{\tau} + c(x_j,t_n)\frac{u_j^n - u_{j-1}^n}{h} = f(x_j,t_n) \quad j \in \mathbb{Z}$$

si $c(x_j,t_n) \leq 0$



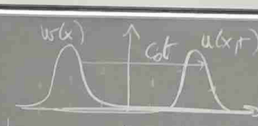
$$\frac{u_j^{n+1} - u_j^n}{\tau} + c(x_j,t_n)\frac{u_{j+1}^n - u_j^n}{h} = f(x_j,t_n) \quad j \in \mathbb{Z}$$

Ce schéma est

- explicite : $u_j^{n+1} = \dots$ $j \in \mathbb{Z}$
- stable sous condition entre τ et h
- sous cette condition, converge à l'ordre $O(h) + b(\tau)$ (si h et τ /2, l'erreur au temps final /2)

Stabilité : $f(x,t)=0$ $c(x,t)=c_0$ et $u(x,t)=w(x-c_0t)$

Déf : on dira que le schéma est stable si les 2 propriétés suivantes sont satisfaites
- positivité : si $u_j^n \geq 0$ $j\in\mathbb{Z}$ alors $u_j^{n+1} \geq 0$ $j\in\mathbb{Z}$
- principe du max : $\sup_{j\in\mathbb{Z}} |u_j^{n+1}| \leq \sup_{j\in\mathbb{Z}} |u_j^n|$



Lemme : le schéma décentré est stable si $\tau \leq \frac{h}{|c_0|}$

Dem : cas où $c_0>0$

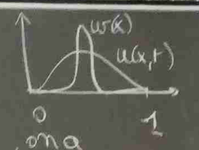
le schéma s'écrit

$$\frac{u_j^{n+1}-u_j^n}{\tau} + c_0 \frac{u_j^n - u_{j-1}^n}{h} = 0 \quad j\in\mathbb{Z}$$

i.e. $u_j^{n+1} = \left(1-\frac{c_0\tau}{h}\right) u_j^n + \frac{c_0\tau}{h} u_{j-1}^n$

- positivité : supposons $\tau \leq \frac{h}{c_0}$, supp $u_j^n \geq 0$ $j\in\mathbb{Z}$ et montrons $u_j^{n+1} \geqslant 0$ $j\in\mathbb{Z}$

$u_j^{n+1} \geqslant 0$ si $1-\frac{c_0\tau}{h} \geqslant 0 \Leftrightarrow 1 \geqslant \frac{c_0\tau}{h} \Leftrightarrow \tau \leq \frac{h}{c_0}$



- principe du max : supp $\tau \leqslant h/c_0$, on a

$$\begin{aligned} |u_j^{n+1}| &= \left|\left(1-\frac{c_0\tau}{h}\right) u_j^n + \frac{c_0\tau}{h} u_{j-1}^n\right| \quad j\in\mathbb{Z} \\ &\leqslant \left|1-\frac{c_0\tau}{h}\right| |u_j^n| + \left|\frac{c_0\tau}{h}\right| |u_{j-1}^n| \quad j\in\mathbb{Z} \\ &= \left(1-\frac{c_0\tau}{h}\right)|u_j^n| + \frac{c_0\tau}{h}|u_{j-1}^n| \quad j\in\mathbb{Z} \end{aligned}$$

$$\forall j\in\mathbb{Z} \quad |u_j^{n+1}| \leqslant \underbrace{\left(1-\frac{c_0\tau}{h}+\frac{c_0\tau}{h}\right)}_{1} \sup_{k\in\mathbb{Z}} |u_k^n|$$

$$\leqslant \sup_{k\in\mathbb{Z}} |u_k^n|$$

et donc $\sup_{j\in\mathbb{Z}} |u_j^{n+1}| \leqslant \sup_{k\in\mathbb{Z}} |u_k^n|$

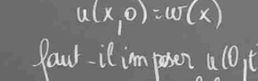
Si $c_0<0$ $u_j^{n+1} = \left(1+\frac{c_0\tau}{h}\right) u_j^n - \frac{c_0\tau}{h} u_{j+1}^n$ $j\in\mathbb{Z}$

Que se passe-t-il si $x\in[0,1]$?

$$\frac{\partial u}{\partial t}(x,t) + c(x,t)\frac{\partial u}{\partial x}(x,t) = f(x,t) \quad 0<x<1,\ t>0$$

$$u(x,0) = w(x)$$

$c(0,t)>0$ $0<x<1$ $c(1,t)<0$

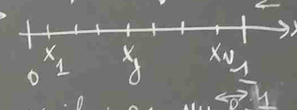


faut-il imposer $u(0,t)$? $u(1,t)$?

Si $c(0,t)>0$ il faut imposer $u(0,t)$ = qqchose

Si $c(1,t)<0$ il faut imposer $u(1,t)$ = autre chose

Sinon, on ne fait rien



$x_j = jh$ $j=0,1,\dots,N+1$ $\Leftrightarrow h=\frac{1}{N+1}$

schéma décentré si $c(x_1,t_n)>0$

$$\frac{u_1^{n+1}-u_1^n}{\tau} + c(x_1,t_n)\frac{u_1^n - u_0^n}{h} = f(x_1,t_n)$$

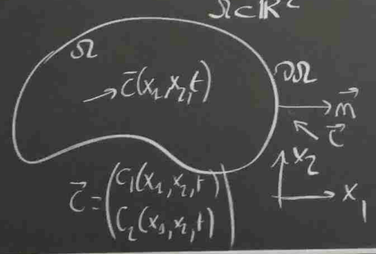
Il faut imposer u_0^n = par pouvoir calculer u_1^{n+1}

schéma décentré si $c(x_N,t_n)<0$

$$\frac{u_N^{n+1}-u_N^n}{\tau} + c(x_N,t_n)\frac{u_{N+1}^n - u_N^n}{h} = f(x_N,t_n)$$

il faut imposer u_{N+1}^n = par un moyen calculable

Transport 2D ?



$$\frac{\partial u}{\partial t}(x_1,x_2,t) + \vec{c}\cdot\vec{\nabla}u(x_1,x_2,t) = f(x_1,x_2,t)$$

$$\left(c_1\frac{\partial u}{\partial x_1} + c_2\frac{\partial u}{\partial x_2}\right)(x_1,x_2,t) = f(x_1,x_2,t) \quad (x_1,x_2)\in\Omega,\ t>0$$

$$u(x_1,x_2,0) = w(x_1,x_2) \quad (x_1,x_2)\in\Omega$$

Faut-il spécifier u sur $\partial\Omega$? Oui, sur le bord entrant

$$\{(x_1,x_2) \text{ telque } c(x_1,x_2,t)\cdot\vec{m}(x_1,x_2) < 0\}$$

Chap 13 phm hyperboliques

13.1 transport

13.2 eq. ondes

Exp : Corde vibrante

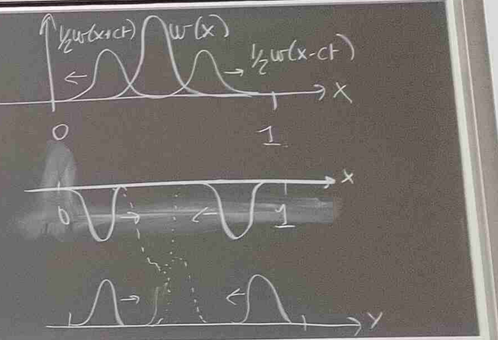


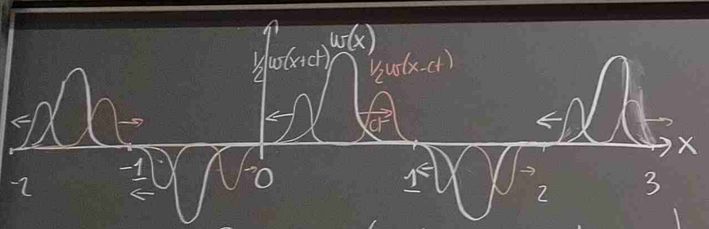
Etant donné $w(x)$, $v(x)$, $f(x,t)$ on cherche $u(x,t)$ la déformation telle que

$$\begin{cases} \frac{\partial^2 u}{\partial t^2}(x,t) - c^2 \frac{\partial^2 u}{\partial x^2}(x,t) = f(x,t) & 0<x<1,\ t>0 \\ u(0,t) = 0 \quad u(1,t) = 0 & \text{cond. limites} \quad t>0 \\ u(x,0) = w(x) & \text{cond. initiale} \quad 0<x<1 \\ \frac{\partial u}{\partial t}(x,0) = v(x) & \end{cases}$$

Remarques :

- Si $f(x,t) = 0$ et $v(x) = 0$, $u(x,t)$ est donnée par $u(x,t) = \frac{1}{2}\left(w(x-ct) + w(x+ct)\right)$ où w a été prolongée par imparité sur $(-1,0)$ et par 2-périodicité sur $\mathbb{R}$





Vérification : $\frac{\partial u}{\partial t}(x,t) = \frac{1}{2}\left(w'(x-ct)(-c) + w'(x+ct)c\right)$

$\frac{\partial u}{\partial t}(x,t) = \frac{c^2}{2}\left(w''(x-ct) + w''(x+ct)\right)$

$\frac{\partial u}{\partial x}(x,t) =$

$\frac{\partial^2 u}{\partial x^2}(x,t) =$

$\frac{\partial^2 u}{\partial t^2}(x,t) - c^2\frac{\partial^2 u}{\partial x^2} = 0$

$u(x,0) = w(x)$

$u(0,t) = \frac{1}{2}\left(w(-ct) + w(ct)\right)$

$= \frac{1}{2}\left(-w(ct) + w(ct)\right)$

$= 0$

$u(1,t) = \frac{1}{2}\left(w(1-ct) + w(1+ct)\right)$

$= \frac{1}{2}\left(w(-2+1-ct) + w(1+ct)\right)$

$= \frac{1}{2}\left(-w(1-ct) + w(1+ct)\right)$

$= 0$

$\frac{\partial u}{\partial t}(x,0) = \frac{1}{2}\left(w'(x)(-c) + w'(x)c\right)$

$= 0$

- L'eq. traduit un principe de conservation :

On mult. par $\frac{\partial u}{\partial t}(x,t)$ (vitesse) et on $\int_0^1 dx$

$$\underbrace{\int_0^1 \frac{\partial^2 u}{\partial t^2}(x,t)\frac{\partial u}{\partial t}(x,t)\,dx}_{\frac{1}{2}\frac{d}{dt}\int_0^1 \left(\frac{\partial u}{\partial t}(x,t)\right)^2 dx} - \underbrace{\int_0^1 c^2 \frac{\partial^2 u}{\partial x^2}(x,t)\frac{\partial u}{\partial t}(x,t)\,dx}_{\text{on intègre par parties (par rapport à } x)} = \underbrace{\int_0^1 f(x,t)\frac{\partial u}{\partial t}(x,t)\,dx}_{\text{Puissance apportée par } f}$$

$(uv)' = u'v + uv' \quad u(x)\ v(x)$

$[uv] = \int u'v\,dx + \int uv'\,dx$

$u \to u'$

$[u'v] = \int u''v\,dx + \int u'v'\,dx$

$$-c^2\int_0^1 \frac{\partial^2 u}{\partial x^2}(x,t)\frac{\partial u}{\partial t}(x,t)\,dx = c^2\int_0^1 \frac{\partial u}{\partial x}(x,t)\frac{\partial}{\partial x}\frac{\partial u}{\partial t}(x,t)\,dx - \left[\frac{\partial u}{\partial x}(x,t)\frac{\partial u}{\partial t}(x,t)\right]_{x=0}^{x=1}$$

$$= \frac{c^2}{2}\frac{d}{dt}\int_0^1 \left(\frac{\partial u}{\partial x}(x,t)\right)^2 dx - 0$$

car $\frac{\partial u}{\partial t}(0,t) = 0$ et $\frac{\partial u}{\partial t}(1,t) = 0$

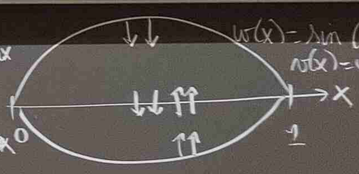
(en effet $\frac{\partial u}{\partial t}(0,t) = \lim_{\varepsilon\to 0}\frac{u(0,t+\varepsilon) - u(0,t)}{\varepsilon} = \lim_{\varepsilon\to 0}\frac{0-0}{\varepsilon} = 0$)

Finalement $\frac{d}{dt}\left(\underbrace{\int_0^1 \frac{1}{2}\left(\frac{\partial u}{\partial t}(x,t)\right)^2 dx}_{\text{en. cinétique}} + \underbrace{\int_0^1 \frac{c^2}{2}\left(\frac{\partial u}{\partial x}(x,t)\right)^2 dx}_{\text{en. déformation}}\right) = \int_0^1 f(x,t)\frac{\partial u}{\partial t}(x,t)\,dx$

Variation (en. cinétique + en. déformation) au cours du temps = puissance

Ex Si $f(x,t) = 0$ et $v(x) = 0$ et $w(x) = \sin(\pi x)$ la somme de l'en. cinétique + en. déformation est constante

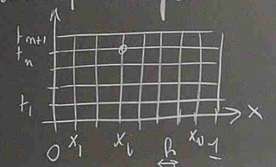
En. cin = 0
En. def max
En déf = 0
En. cin max



$u(x,t) = \sin(\pi x)\cos(\pi ct)$

Schéma numérique (Newmark)

Grille espace temps



$h = \frac{1}{N+1}$ pas d'espace $x_i = ih$ $i = 0, 1, \dots, N+1$

τ pas de temps $t_m = m\tau$ $m = 0, 1, 2, \dots$

But : calculer u_i^m approximation de $u(x_i, t_m)$

On écrit l'eq des ondes en (x_i, t_m) et on utilise des formules de diff finies centrée

$$\frac{\partial^2 u}{\partial t^2}(x_i, t_m) - c^2 \frac{\partial^2 u}{\partial x^2}(x_i, t_m) = f(x_i, t_m) \quad (m \text{ fixé } i = 1, \dots, N)$$

$$\frac{u(x_i, t_{m-1}) - 2u(x_i, t_m) + u(x_i, t_{m+1})}{\tau^2} - c^2 \frac{u(x_{i-1}, t_m) - 2u(x_i, t_m) + u(x_{i+1}, t_m)}{h^2} = f(x_i, t_m) + O(h^2) + O(\tau^2)$$

Schéma : m fixé
$$\begin{cases} \dfrac{u_i^{m-1} - 2u_i^m + u_i^{m+1}}{\tau^2} - c^2 \dfrac{u_{i-1}^m - 2u_i^m + u_{i+1}^m}{h^2} = f(x_i, t_m) & i = 1, \dots, N \\ u_0^m = 0 \text{ et } u_{N+1}^m = 0 \end{cases}$$

Ce schéma explicite permet de calculer u_i^{m+1} $i = 1, \dots, N$

$$u_i^{m+1} = -u_i^{m-1} + u_i^m \left(2 - 2\frac{\tau^2 c^2}{h^2}\right) + \left(u_{i-1}^m + u_{i+1}^m\right)\frac{\tau^2 c^2}{h^2} + \tau^2 f(x_i, t_m)$$

Etant donné u_i^0 et u_i^1, on peut calculer u_i^2, $i = 1, \dots, N$

$u_i^0 = w(x_i)$ $i = 1, \dots, N$ $(u(x,0) = w(x))$

Comment calculer u_i^1 ? (utiliser $\frac{\partial u}{\partial t}(x,0) = v(x)$ tout en gardant $O(\tau^2)$)

$\frac{\partial u}{\partial t}(x_i, t_0) = v(x_i)$ $i = 1, \dots, N$

$$\frac{u(x_i, t_1) - u(x_i, t_{-1})}{2\tau} = v(x_i) + O(\tau^2)$$

Schéma $m = 0$:
$$\begin{cases} \dfrac{u_i^{-1} - 2u_i^0 + u_i^1}{\tau^2} - c^2 \dfrac{u_{i-1}^0 - 2u_i^0 + u_{i+1}^0}{h^2} = f(x_i, t_0) & i = 1, \dots, N \\ \dfrac{u_i^1 - u_i^{-1}}{2\tau} = v(x_i) & i = 1, \dots, N \end{cases}$$

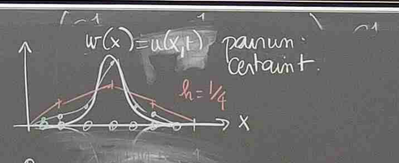
On élimine u_i^{-1} : $u_i^{-1} = u_i^1 - 2\tau v(x_i)$

Schéma pour calculer u_i^1
$$\begin{cases} \dfrac{2u_i^1 - 2\tau v(x_i) - 2u_i^0}{\tau^2} - c^2 \dfrac{u_{i-1}^0 - 2u_i^0 + u_{i+1}^0}{h^2} = f(x_i, t_0) & i = 1, \dots, N \\ u_0^0 = 0 \quad u_{N+1}^0 = 0 \end{cases}$$

On en déduit $u_i^1 = \dots$

On implémente ce schéma, on observe (et on peut démontrer que)

- Le schéma est stable si $\tau \leqslant \frac{h}{c}$ (idem 13.1 transport $\frac{\partial u}{\partial t} + c\frac{\partial u}{\partial x} = 0$) (cas où $f(x,t) = 0$ et $v(x) = 0$)
- Sous la cond. de stabilité, le schéma converge $O(h^2) + O(\tau^2)$ (l'erreur au temps final fixé est divisée par 4 si h et τ /2 sous la cond. $\tau \leqslant \frac{h}{c}$)



BLOC 2

Chap 7. Calcul de valeurs propres

$A \in \mathbb{R}^{N\times N}$ on cherche $\vec{\varphi} \neq \vec{0}$ $\vec{\varphi} \in \mathbb{C}^N$ (vecteur propre) et $\lambda \in \mathbb{C}$ (valeur propre) tq $A\vec{\varphi} = \lambda\vec{\varphi}$

Ex : $A = \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}$ $\begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} 1 \\ i \end{pmatrix} = (1-i)\begin{pmatrix} 1 \\ i \end{pmatrix}$

de plus $A\vec{\varphi} = \lambda\vec{\varphi} \Leftrightarrow (A-\lambda I)\vec{\varphi} = \vec{0}$ (avec $\vec{\varphi} \neq \vec{0}$)

$\Leftrightarrow \mathrm{Ker}(A-\lambda I) \neq \vec{0}$

$\Leftrightarrow \det(A-\lambda I) = 0$

$\Leftrightarrow \lambda$ racine du polynôme caract de A

Qu'il n'y a de formule explicite pour trouver les racines d'un polyn de degré $\leq N$

Il faut utiliser une méthode itérative

Si A est symmétrique

$A\vec{\varphi}_j = \lambda_j \vec{\varphi}_j \quad j = 1, \dots, N$

$\lambda_j \in \mathbb{R}$ $\vec{\varphi}_j \in \mathbb{R}^N$

$\vec{\varphi}_j^T \vec{\varphi}_i = (\longrightarrow\!|\!|) = \delta_{ij}$ ($=1$ si $i=j$, $=0$ sinon) $i,j = 1,\dots,N$

les valeurs et vecteurs propres sont réels et il existe une base orthonormée de vecteurs propres

Ex : $\begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix}\begin{pmatrix} 1 \\ 1 \end{pmatrix} = 1\begin{pmatrix} 1 \\ 1 \end{pmatrix}$, $\begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix}\begin{pmatrix} 1 \\ -1 \end{pmatrix} = 3\begin{pmatrix} 1 \\ -1 \end{pmatrix}$

$\lambda_1 = 1$ $\vec{\varphi}_1 = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ $\lambda_2 = 3$ $\vec{\varphi}_2 = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ -1 \end{pmatrix}$

Si A est sym. def. positive alors $0 < \lambda_j$ $j = 1, \dots, N$

Si on connaît les λ_j et $\vec{\varphi}_j$ tels que $A\vec{\varphi}_j = \lambda_j\vec{\varphi}_j$ $j = 1,\dots,N$ alors on peut résoudre les problèmes suivants

---

- Etant donné $A \in \mathbb{R}^{N\times N}$ inversible (sym) et $\vec{b} \in \mathbb{R}^N$, sat $\vec{x}$ tel que $A\vec{x} = \vec{b}$

On décompose $\vec{b}$ dans la base des vecteurs propres :

$\vec{b} = \sum_{j=1}^{N} \beta_j \vec{\varphi}_j \quad (\beta_j = \vec{b}^T\vec{\varphi}_j)$

($\vec{b} \in \mathbb{R}^2$: $\vec{b} = (\vec{b},\vec{e}_1)\vec{e}_1 + (\vec{b},\vec{e}_2)\vec{e}_2$)

idem pour $\vec{x}$: $\vec{x} = \sum_{j=1}^{N} \alpha_j \vec{\varphi}_j$ ← inconnu

$A\vec{x} = \vec{b}$

$A\sum_{j=1}^{N}\alpha_j\vec{\varphi}_j = \sum_{j=1}^{N}\beta_j\vec{\varphi}_j$

$\sum_{j=1}^{N}(\alpha_j\lambda_j - \beta_j)\vec{\varphi}_j = \vec{0}$

et donc $\alpha_j\lambda_j = \beta_j$ ie $\alpha_j = \frac{\beta_j}{\lambda_j}$

- On peut résoudre le syst différentiel

$\begin{cases} \vec{u}'(t) + A\vec{u}(t) = \vec{0} & t > 0 \\ \vec{u}(0) = \vec{u}_0 \end{cases}$

où $A \in \mathbb{R}^{N\times N}$ sym. def. positive, $\vec{u}_0 \in \mathbb{R}^N$

(Chap 12 problème Chaleur discrétisation en espace  $x_i = \frac{i}{N+1}$)

$\vec{u}(t) = \begin{pmatrix} u_1(t) \\ \vdots \\ u_N(t) \end{pmatrix}$ $u_i(t)$ approx de $u(x_i,t)$

On décompose $\vec{u}_0$ dans la base des vecteurs propre de A :

$\vec{u}_0 = \sum_{j=1}^{N}\alpha_j\vec{\varphi}_j \quad (\alpha_j = \vec{u}_0^T\vec{\varphi}_j)$

alors $\vec{u}(t) = \sum_{j=1}^{N}\alpha_j e^{-\lambda_j t}\vec{\varphi}_j$

Vérification : $\vec{u}(0) = \sum_j \alpha_j\vec{\varphi}_j = \vec{u}_0$

$\vec{u}'(t) = \sum_{j=1}^{N}\alpha_j(-\lambda_j)e^{-\lambda_j t}\vec{\varphi}_j$

$A\vec{u}(t) = \sum_{j=1}^{N}\alpha_j e^{-\lambda_j t}A\vec{\varphi}_j$

$= \sum_{j=1}^{N}\alpha_j\lambda_j e^{-\lambda_j t}\vec{\varphi}_j$

et donc $\vec{u}'(t) + A\vec{u}(t) = \vec{0}$

- On peut aussi résoudre le syst. diff. suivant

$\begin{cases} \vec{u}''(t) + A\vec{u}(t) = \vec{0} & t > 0 \\ \vec{u}(0) = \vec{u}_0 \\ \vec{u}'(0) = 0 \end{cases}$

où $A \in \mathbb{R}^{N\times N}$ sym. def. pos., $\vec{u}_0 \in \mathbb{R}^N$

$\vec{u}_0 = \sum_{j=1}^{N}\alpha_j\vec{\varphi}_j \quad (\alpha_j = \vec{u}_0^T\vec{\varphi}_j)$

alors $\vec{u}(t) = \sum_{j=1}^{N}\alpha_j\cos(\sqrt{\lambda_j}t)\vec{\varphi}_j$

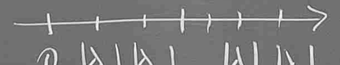
(Chap 13. eq. ondes, $\vec{u}(t) = \begin{pmatrix} u_1(t) \\ \vdots \\ u_N(t) \end{pmatrix}$ $u_i(t)$ approx $u(x_i,t)$)

Dans la suite $A \in \mathbb{R}^{N\times N}$ sym. définie

On ordonne les valeurs propres

$A\vec{\varphi}_j = \lambda_j\vec{\varphi}_j \quad j = 1,\dots,N$

de la manière suivante :



λ_1 est la plus grande valeur propre en valeur absolue

méthode de la puissance (power)

$\vec{x}^0 \in \mathbb{R}^N$ donné

$n = 1, 2, \dots$

$\vec{x}^n = A\vec{x}^{n-1}$

$\mu^n = \frac{\vec{x}^{nT}A\vec{x}^n}{\vec{x}^{nT}\vec{x}^n}$

quotient de Rayleigh

Thm (hyp) : $A \in \mathbb{R}^{N\times N}$ sym. $A\vec{\varphi}_j = \lambda_j\vec{\varphi}_j$ $j = 1,N$ $\vec{\varphi}_j^T\vec{\varphi}_i = \delta_{ij}$

- $0 \leq |\lambda_N| \leq |\lambda_{N-1}| \leq \dots \leq |\lambda_2| < |\lambda_1|$ ($|\lambda_1| > |\lambda_j|$ $j = 2,\dots,N$)
- $\vec{x}^{0T}\vec{\varphi}_1 \neq 0$

Concl : $\frac{\vec{x}^n}{\|\vec{x}^n\|} \xrightarrow[n\to\infty]{} \pm\vec{\varphi}_1$ et $\mu^n \xrightarrow[n\to\infty]{} \lambda_1$

Dem : on déc. $\vec{x}^0$ dans la base des vecteurs propres :

$$\vec{x}^0 = \sum_{j=1}^{N} \alpha_j \vec{\varphi}_j \quad (\alpha_j = \vec{x}^{0T}\vec{\varphi}_j)$$

$$\vec{x}^1 = A\vec{x}^0 = A\sum_{j=1}^{N} \alpha_j \vec{\varphi}_j = \sum_{j=1}^{N} \alpha_j \underbrace{A\vec{\varphi}_j}_{\lambda_j\vec{\varphi}_j} = \sum_{j=1}^{N} \alpha_j \lambda_j \vec{\varphi}_j$$

$$\vec{x}^n = \sum_{j=1}^{N} \alpha_j (\lambda_j)^n \vec{\varphi}_j$$

$$\|\vec{x}^n\|^2 = \vec{x}^{nT}\vec{x}^n = \left(\sum_{j=1}^{N} \alpha_j \lambda_j^n \vec{\varphi}_j\right)^T \left(\sum_{k=1}^{N} \alpha_k \lambda_k^n \vec{\varphi}_k\right)$$

$$= \sum_{j=1}^{N}\sum_{k=1}^{N} \alpha_j \alpha_k \lambda_j^n \lambda_k^n \underbrace{\vec{\varphi}_j^T \vec{\varphi}_k}_{=\delta_{jk}\ (\substack{1 \text{ si } j=k \\ 0 \text{ sinon}})}$$

$$= \sum_{j=1}^{N} (\alpha_j)^2 (\lambda_j)^{2n}$$

$$\frac{\vec{x}^n}{\|\vec{x}^n\|} = \frac{\sum_{j=1}^{N} \alpha_j (\lambda_j)^n \vec{\varphi}_j}{\sqrt{\sum_{j=1}^{N} (\alpha_j)^2 (\lambda_j)^{2n}}}$$

or $|\lambda_j| < |\lambda_1| \quad j = 2, \dots, N$

$\left|\frac{\lambda_j}{\lambda_1}\right| < 1 \quad (|\alpha|<1,\ a^{n \to \infty} \to 0)$

$\left(\frac{\lambda_j}{\lambda_1}\right)^n \xrightarrow{n\to\infty} 0 \quad j = 2, \dots, N$

donc $\frac{\vec{x}^n}{\|\vec{x}^n\|} = \pm \frac{\sum_{j=1}^{N} \alpha_j \left(\frac{\lambda_j}{\lambda_1}\right)^n \vec{\varphi}_j}{\sqrt{\sum_{j=1}^{N} (\alpha_j)^2 \left(\frac{\lambda_j}{\lambda_1}\right)^{2n}}}$

$\vec{x}^{0T}\vec{\varphi}_1 \neq 0$; $\sum_{j=1}^{N} \alpha_j \underbrace{\vec{\varphi}_j^T \vec{\varphi}_1}_{=\alpha_1} \neq 0$

$$\frac{\vec{x}^n}{\|\vec{x}^n\|} \xrightarrow{n\to\infty} \pm \frac{\alpha_1 \vec{\varphi}_1}{\sqrt{\alpha_1^2}} = \pm \vec{\varphi}_1 \quad (\alpha_1 \neq 0)$$

μ^n ? $\quad A\vec{x}^n = A\sum_{j=1}^{N} \alpha_j (\lambda_j)^n \vec{\varphi}_j = \sum_{j=1}^{N} \alpha_j (\lambda_j)^n \underbrace{A\vec{\varphi}_j}_{\lambda_j\vec{\varphi}_j}$

$$= \sum_{j=1}^{N} \alpha_j (\lambda_j)^{n+1} \vec{\varphi}_j$$

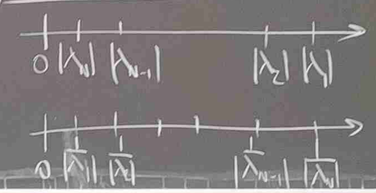
$$\mu^n = \frac{\vec{x}^{nT} A \vec{x}^n}{\vec{x}^{nT}\vec{x}^n} = \frac{\sum_{j=1}^{N} (\alpha_j)^2 (\lambda_j)^{2n+1}}{\sum_{j=1}^{N} (\alpha_j)^2 (\lambda_j)^{2n}} = \frac{\sum_{j=1}^{N} (\alpha_j)^2 \frac{(\lambda_j)^{2n+1}}{(\lambda_1)^{2n+1}} \lambda_1}{\sum_{j=1}^{N} (\alpha_j)^2 \left(\frac{\lambda_j}{\lambda_1}\right)^{2n}} \xrightarrow{n\to\infty} \frac{(\alpha_1)^2 \lambda_1}{(\alpha_1)^2} = \lambda_1$$

Méthode de la puissance inverse (calculer λ_N)

$A\vec{\varphi}_j = \lambda_j \vec{\varphi}_j$ $\quad A$ inversible $(\lambda_j \neq 0)$

$\vec{\varphi}_j = \lambda_j A^{-1}\vec{\varphi}_j$

$A^{-1}\vec{\varphi}_j = \frac{1}{\lambda_j}\vec{\varphi}_j$ (les vecteurs propres sont les mêmes)



$\frac{1}{|\lambda_N|}$ valeur propre la plus grande de A^{-1} en valeur absolue

Idée : appliquer la méth de la puissance à A^{-1}

$\vec{x}^0 \in \mathbb{R}^N$ donné

$n = 1, 2, \dots$

$\vec{x}^n = A^{-1}\vec{x}^{n-1}$ (ie connaissant $A\vec{x}^n = \vec{x}^{n-1}$)

$\mu^n = \frac{\vec{x}^{nT} A \vec{x}^n}{\vec{x}^{nT}\vec{x}^n}$ (intuitivement si $\frac{\vec{x}^n}{\|\vec{x}^n\|} \xrightarrow{n\to\infty} \vec{\varphi}_N$)

$$\mu^n \xrightarrow{n\to\infty} \frac{\vec{\varphi}_N^T A \vec{\varphi}_N}{\vec{\varphi}_N^T \vec{\varphi}_N} = \frac{\vec{\varphi}_N^T \vec{\varphi}_N \lambda_N}{1} = \lambda_N$$

Thm : hyp : $0 < |\lambda_N| < |\lambda_{N-1}| \leq |\lambda_{N-2}| \leq \dots \leq |\lambda_2| \leq |\lambda_1|$

$\vec{x}_0^T \vec{\varphi}_N \neq 0$

Concl : $\frac{\vec{x}^n}{\|\vec{x}^n\|} \xrightarrow{n\to\infty} \pm \vec{\varphi}_N$

$\mu^n \xrightarrow{n\to\infty} \lambda_N$

Il existe une variante qui permet de calculer les autres valeurs propres

SVD (Singular value decomposition)

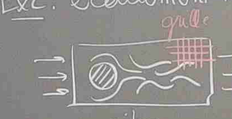
$A = \begin{bmatrix} | & | & & | \\ \vec{a}_1 & \vec{a}_2 & \cdots & \vec{a}_n \\ | & | & & | \end{bmatrix} \in \mathbb{R}^{m \times n}$ (1 à n en haut, 1 à m à gauche)

Ex1 : image



Utilité : compresser l'information

Ex2 : écoulement turbulent derrière un cyl.



$\vec{a}_i$: champ de vitesse sur la grille au temps t_i

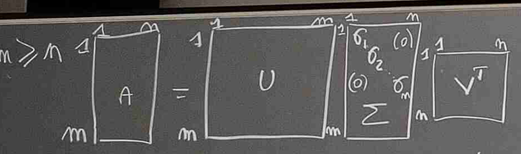
Thm : $A \in \mathbb{R}^{m \times n}$, il existe

- $U \in \mathbb{R}^{m \times m}$ orthonormale $(UU^T = U^T U = I_m)$
- $V \in \mathbb{R}^{n \times n}$ $(VV^T = V^T V = I_n)$
- Σ diagonale $\in \mathbb{R}^{m \times n}$ de coeff. diagonaux positifs

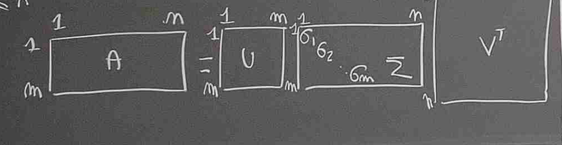
$$\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_{\min(m,n)} \geq 0$$

telles que $A = U \Sigma V^T$

Si $m \geq n$:



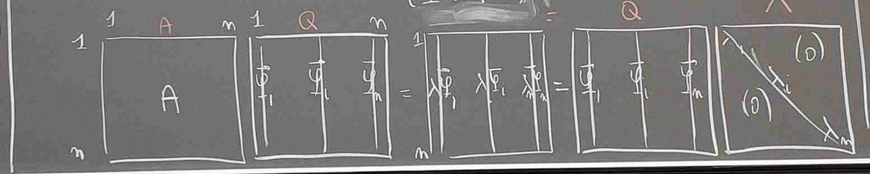
Si $m \leq n$:



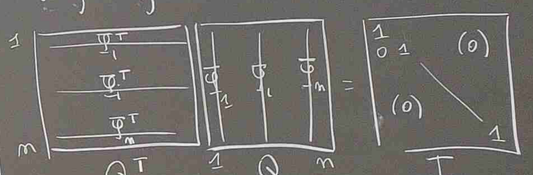
Ex : $A \in \mathbb{R}^{n \times n}$ sym définie positive

$A\vec{\varphi}_i = \lambda_i \vec{\varphi}_i \quad \lambda_i > 0 \qquad \vec{\varphi}_i^T \vec{\varphi}_j = \delta_{ij} \quad i,j = 1,\dots,n$

$(1 \text{ si } i=j, 0 \text{ sinon})$



$\vec{\varphi}_i^T \vec{\varphi}_j = \delta_{ij}$:



$$Q^T Q = I = Q Q^T$$

$(AQ = Q\Lambda)\, Q^T$

$AQQ^T = Q\Lambda Q^T$

$A = Q\Lambda Q^T$

$= U \Sigma V^T$ (thm)

On a aussi : $Q^T(AQ = Q\Lambda)$

$Q^T A Q = \Lambda$

matlab, help svd

$$A = \begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} 3 & 0 \\ 0 & 1 \end{pmatrix} \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}$$

$$A = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

(U, Σ, V^T)

pas de base de vecteurs propres (pas diagonalisable)

$$\begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 0 \\ \alpha \end{pmatrix} = 0 \begin{pmatrix} 0 \\ \alpha \end{pmatrix} \qquad \forall \alpha \in \mathbb{R}$$

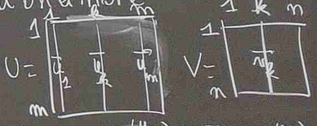
Def : Soit r tel que $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0 = \sigma_{r+1} = \cdots = \sigma_{\min(m,n)}$

On dit que la matrice est de rang r et on a

$$A = U \Sigma V^T = \sum_{k=1}^{r} \sigma_k \vec{u}_k \vec{v}_k^T$$

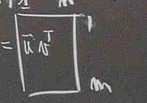
$(\vec{u}\vec{v}^T)_{ij} = u_i v_j \quad i = 1,\dots,m \quad j = 1,\dots,n$

où on a noté



et pour $\vec{u} = \begin{pmatrix} u_1 \\ \vdots \\ u_m \end{pmatrix} \in \mathbb{R}^m$ $\vec{v} = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} \in \mathbb{R}^n$

$\vec{u}\vec{v}^T =$ ( ) ( ) $=$



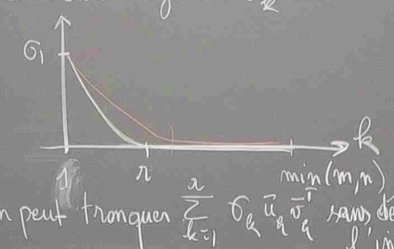
Utilité de (*) :

mémoire nécessaire pour $A \in \mathbb{R}^{m\times n}$: $m \times n$ coeff.

$\sum_{k=1}^{r} \sigma_k \vec{u}_k \vec{v}_k^T$ $r(m+n)$ coeff.

Si r est petit devant m et n, (*) est intéressant

Si les valeurs singulières σ_k décroissent rapidement



on peut tronquer $\sum_{k=1}^{r} \sigma_k \vec{u}_k \vec{v}_k^T$ sans détériorer l'information

Dem de (*) :

$$A_{ij} = (U\Sigma V^T)_{ij} = \sum_{k=1}^{r} \sigma_k (\vec{u}_k \vec{v}_k^T)_{ij} \qquad i=1,\dots,m \quad j=1,\dots,n$$

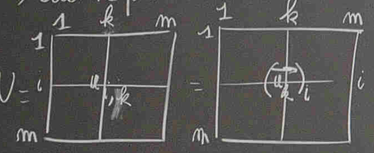
$$(U\Sigma V^T)_{ij} = \sum_{k=1}^{m} (U\Sigma)_{ik} (V^T)_{kj} = \sum_{k=1}^{m}\sum_{\ell=1}^{m} U_{i\ell}\, \Sigma_{\ell k}\, V_{jk}$$

$$= \sum_{k=1}^{m} u_{ik}\, \sigma_k\, v_{jk} = \sum_{k=1}^{r} \sigma_k u_{ik} v_{jk}$$

$\Sigma_{\ell k} = \sigma_k$ si $\ell = k$

$= 0$ sinon

D'autre part



donc $u_{i,k} = (\vec{u}_k)_i$

et donc $(U\Sigma V^T)_{ij} = \sum_{k=1}^{r} \sigma_k (\vec{u}_k)_i (\vec{v}_k)_j = \sum_{k=1}^{r} \sigma_k (\vec{u}_k \vec{v}_k^T)_{ij}$

et on a dem (*)

Calcul de U, Σ, V :

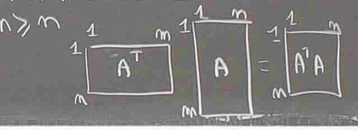
$A = U\Sigma V^T$, $A^T = (U\Sigma V^T)^T = V\Sigma^T U^T$ et donc

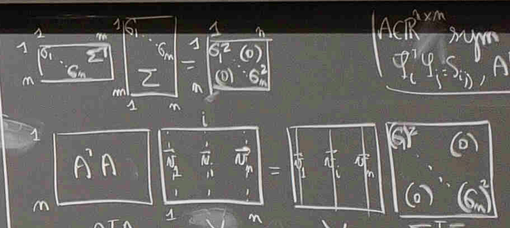
$A^TA = V\Sigma^T \underbrace{U^T U}_{I} \Sigma V^T$

$(A^TA = V\Sigma^T\Sigma V^T)V$

$A^TAV = V\Sigma^T\Sigma$

Si $m \geq n$





$A^TA\ \vec{v}_i = (\sigma_i)^2 \vec{v}_i \quad i=1,\dots,n$

$A \in \mathbb{R}^{m\times n}$ sym $AQ = Q\Lambda$

$\varphi_i^T \varphi_j = \delta_{ij}$, $A\varphi_i = \lambda_i \varphi_i$

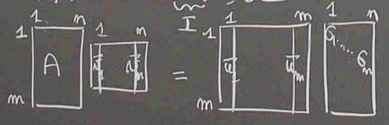
Conclusion : pour trouver V, il faut trouver les valeurs propres et vecteurs propres de A^TA.

La méthode la puissance permet de trouver σ_1^2 et $\vec{v}_1$

la méth de la puissance inverse σ_n^2 et $\vec{v}_n$

Et les $\vec{u}_i$ (ou U)

$AV = U\Sigma \underbrace{V^T V}_{I} = U\Sigma$



et donc $A\vec{v}_i = \sigma_i \vec{u}_i \quad i=1,\dots,n$

ce qui permet de déterminer $\vec{u}_i$

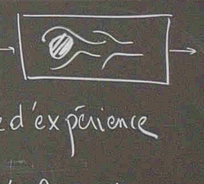
On obtient le même genre de formules avec AA^T

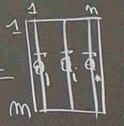
$AA^T = U\Sigma V^T (U\Sigma V^T)^T = U\Sigma V^T V \Sigma^T U^T$

$(AA^T = U\Sigma\Sigma^T U^T)U$

$AA^TU = U\Sigma\Sigma^T$

Retour à l'exemple 2 :



$A =$  matrice d'expérience

Comment représenter l'information de manière plus concise ?

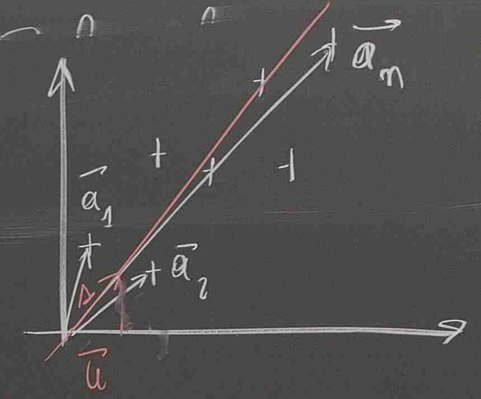
$\vec{a}_i \in \mathbb{R}^m$ $\vec{u}_i$ base orthonormée de $\mathbb{R}^m$ et donc

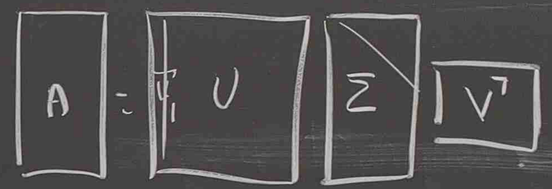
$\vec{a}_i = (\vec{a}_i^T \vec{u}_1)\vec{u}_1 + (\vec{a}_i^T \vec{u}_2)\vec{u}_2 + \dots + (\vec{a}_i^T \vec{u}_m)\vec{u}_m \qquad i = 1, \dots, n$

On cherche $\vec{u} \in \mathbb{R}^m$ $\|\vec{u}\| = 1$ de sorte à minimiser

$\sum_{i=1}^{n} \|\vec{a}_i - (\vec{a}_i^T \vec{u})\vec{u}\|$ On peut montrer que $\vec{u} = \vec{u}_1$

$m = 2$





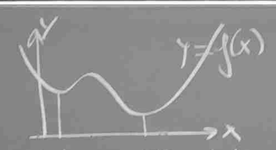
Partie 2 : optimisation

Optim sans contraintes : théorie

Etant donné $f : \mathbb{R}^n \to \mathbb{R}$, $\vec{x} \to f(\vec{x})$, $(x_1, x_2, \dots, x_n) \to f(x_1, x_2, \dots, x_n)$

on cherche $\vec{x}^* \in \mathbb{R}^n$ tel que $f(\vec{x}^*) \leq f(\vec{x}) \ \forall \vec{x} \in \mathbb{R}^n$ (1)

$n=1$



$n=2$



On note aussi $\min_{\vec{x} \in \mathbb{R}^n} f(\vec{x})$ ou $\vec{x}^* = \arg\min_{\vec{x} \in \mathbb{R}^n} f(\vec{x})$

Soit $\vec{x}^* \in \mathbb{R}^n$ tq $f(\vec{x}^*) \geq f(\vec{x}) \ \forall \vec{x} \in \mathbb{R}^n$

alors $-f(\vec{x}^*) \leq -f(\vec{x})$ —

et on se ramène à (1) en posant $\tilde{f}(\vec{x}) = -f(\vec{x})$

Notations : $f : \mathbb{R}^n \to \mathbb{R}$ $\mathcal{C}^1$

$\vec{\nabla} f(\vec{x}) = \begin{pmatrix} \frac{\partial f}{\partial x_1}(\vec{x}) \\ \vdots \\ \frac{\partial f}{\partial x_n}(\vec{x}) \end{pmatrix}$ gradient

et si $f : \mathbb{R}^n \to \mathbb{R}$ $\mathcal{C}^2$

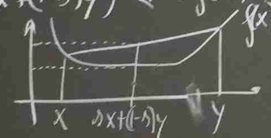
$H(\vec{x})$ la matrice hessienne $\in \mathbb{R}^{n \times n}$

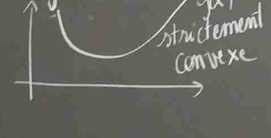
déf. par $\left(H(\vec{x})\right)_{i,j} = \frac{\partial^2 f}{\partial x_i \partial x_j}(\vec{x})$ $1 \leq i,j \leq n$

Déf : $f : \mathbb{R}^n \to \mathbb{R}$ est convexe si $\forall \vec{x}, \vec{y} \in \mathbb{R}^n \ \forall 0 < s < 1$

$f(s\vec{x} + (1-s)\vec{y}) \leq s f(\vec{x}) + (1-s) f(\vec{y})$

$n=1$





Théorème :

a. Si $f : \mathbb{R}^n \to \mathbb{R}$ continue, infinie à l'infini ($\text{si } \|\vec{x}\| \to +\infty$ alors $f(\vec{x}) \to +\infty$)

alors il existe $\vec{x}^*$ sol de (1)

b. Si $f : \mathbb{R}^n \to \mathbb{R}$ $\mathcal{C}^1$, si $\vec{x}^*$ est solution de (1)

alors $\vec{\nabla} f(\vec{x}^*) = \vec{0}$

c. Si $f : \mathbb{R}^n \to \mathbb{R}$ $\mathcal{C}^2$, si $\vec{x}^* \in \mathbb{R}^n$ tq $\vec{\nabla} f(\vec{x}^*) = \vec{0}$

si $H(\vec{x}^*)$ est sym définie positive

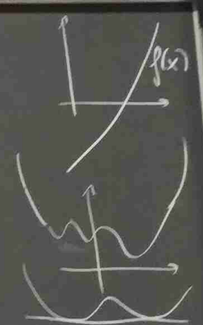
alors $f(\vec{x}^*) \leq f(\vec{x}) \ \forall \vec{x}$ dans un voisinage de $\vec{x}^*$ (min local)

d. Si $f : \mathbb{R}^n \to \mathbb{R}$ $\mathcal{C}^1$ convexe

Trouver $\vec{x}^* \in \mathbb{R}^n$ tq $f(\vec{x}^*) \leq f(\vec{x}) \ \forall \vec{x} \in \mathbb{R}^n$

$\Leftrightarrow$

Trouver $\vec{x}^* \in \mathbb{R}^n$ tq $\vec{\nabla} f(\vec{x}^*) = \vec{0}$

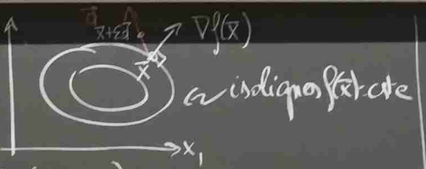


Dem b. dans $\mathbb{R}^n$ les définitions suivantes de la dérivée (gradient) sont équiv.

$\vec{\nabla} f(\vec{x}) = \begin{pmatrix} \frac{\partial f}{\partial x_1}(\vec{x}) \\ \vdots \\ \frac{\partial f}{\partial x_n}(\vec{x}) \end{pmatrix}$ $\frac{\partial f}{\partial x_1}(x_1, x_2, \dots, x_n) = \lim_{\varepsilon \to 0} \frac{f(x_1 + \varepsilon, x_2, \dots, x_n) - f(x_1, x_2, \dots, x_n)}{\varepsilon}$

dérivée directionnelle $\forall \vec{x} \in \mathbb{R}^n \ \forall \vec{d} \in \mathbb{R}^n$ $\vec{\nabla} f(\vec{x})^T \vec{d} = \lim_{\varepsilon \to 0} \frac{f(\vec{x} + \varepsilon \vec{d}) - f(\vec{x})}{\varepsilon}$

$n=2$



Si $\vec{d} = \vec{e}_1 = (1, 0, \dots, 0)$

$\nabla f(\vec{x})^T \vec{d} = \lim_{\varepsilon \to 0} \frac{f(\vec{x} + \varepsilon \vec{e}_1) - f(\vec{x})}{\varepsilon}$

$= \frac{\partial f}{\partial x_1}(\vec{x})$

Soit $\vec{x}^*$ tq $f(\vec{x}^*) \leq f(\vec{x}) \forall \vec{x} \in \mathbb{R}^n$

montrons $\vec{\nabla} f(\vec{x}^*) = \vec{0}$

Soit $\vec{d} \in \mathbb{R}^n$ $\varepsilon > 0$ on a : $f(\vec{x}^*) \leq f(\vec{x}^* + \varepsilon \vec{d})$

et donc $\frac{f(\vec{x}^* + \varepsilon \vec{d}) - f(\vec{x}^*)}{\varepsilon} \geq 0$ on prend $\lim \varepsilon \to 0$:

$\vec{\nabla} f(\vec{x}^*)^T \vec{d} \geqslant 0 \quad \forall \vec{d} \in \mathbb{R}^n$ on choisit $\vec{d} = -\vec{\nabla} f(\vec{x}^*)$

$-\|\vec{\nabla} f(\vec{x}^*)\|^2 \geqslant 0$ donc $\|\vec{\nabla} f(\vec{x}^*)\| = 0 \Leftrightarrow \vec{\nabla} f(\vec{x}^*) = \vec{0}$

$\vec{a}, \vec{b} : \begin{pmatrix} \\ \\ \end{pmatrix} \in \mathbb{R}^n$

$\vec{a}^T \vec{b} = (\quad)(\quad)$

c. Soit $\bar{x}^* \in \mathbb{R}^n$ tq $\bar{\nabla} f(\bar{x}^*) = \bar{0}$ et $H(\bar{x}^*)$ s.d.p.
montrons $f(\bar{x}^*) \leq f(\bar{x})$ $\forall \bar{x}$ dans un voisinage de $\bar{x}^*$

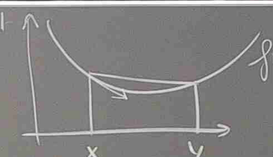
$\forall \bar{x} \in \mathbb{R}^n \quad f(\bar{x}) = f(\bar{x}^*) + \bar{\nabla} f(\bar{x}^*)^T (\bar{x} - \bar{x}^*) + \frac{1}{2} (\bar{x} - \bar{x}^*)^T H(s\bar{x} + (1-s)\bar{x}^*)(\bar{x} - \bar{x}^*)$

$(n=1 \quad f(x) = f(x^*) + f'(x^*)(x - x^*) + \frac{1}{2} f''(sx + (1-s)x^*)(x - x^*)^2$

où $0 < s < 1$

$H(\bar{x}^*)$ s.d.p, cont, reste s.d.p dans un voisinage de $\bar{x}^*$
et donc $f(\bar{x}) \geq f(\bar{x}^*)$ $\forall \bar{x}$ dans un voisinage de $\bar{x}^*$

d. Soit $\bar{x}^* \in \mathbb{R}^n$ tq $\bar{\nabla} f(\bar{x}^*) = \bar{0}$
Si f est convexe montrons $f(\bar{x}^*) \leq f(\bar{x})$ $\forall \bar{x} \in \mathbb{R}^n$



$f'(x) \leq \frac{f(y) - f(x)}{y - x} \quad \forall x, y \in \mathbb{R}$

n quelconque $\forall \bar{x}, \bar{y} \in \mathbb{R}^n$, $\bar{\nabla} f(\bar{x})^T (\bar{y} - \bar{x}) \leq f(\bar{y}) - f(\bar{x})$

Si $\bar{x}^*$ tq $\bar{\nabla} f(\bar{x}^*) = \bar{0}$ alors
$\forall \bar{x} \in \mathbb{R}^n$ on a $(\bar{y} \to \bar{x}, \bar{x} \to \bar{x}^*)$
$f(\bar{x}) - f(\bar{x}^*) \geq \underbrace{\bar{\nabla} f(\bar{x}^*)^T}_{0} (\bar{x} - \bar{x}^*)$

ie $f(\bar{x}) \geq f(\bar{x}^*)$ $\forall \bar{x} \in \mathbb{R}^n$

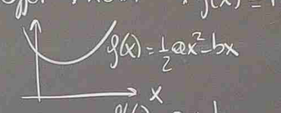
2 Exemples :
1. Fonctionnelle quadratique
2. Moindres carrés, régression

1. Soit $A \in \mathbb{R}^{n \times n}$ sym. def. pos, $\bar{b} \in \mathbb{R}^n$
Soit $f : \mathbb{R}^n \to \mathbb{R}$ def. $\forall \bar{x} \in \mathbb{R}^n$ par
$f(\bar{x}) = \frac{1}{2} \bar{x}^T A \bar{x} - \bar{b}^T \bar{x}$

On a dém. (MOOC Chap 4, 5, 6 ligne
leçon 4, autres notations)
(Trouver $\bar{x}^* \in \mathbb{R}^n$ tq $f(\bar{x}^*) \leq f(\bar{x})$ $\forall x \in \mathbb{R}^n$)
$\Leftrightarrow$
(Trouver $\bar{x}^* \in \mathbb{R}^n$ tq $A\bar{x}^* = b$)

En effet $\forall \bar{x} \in \mathbb{R}^n$ $\bar{\nabla} f(\bar{x}) = A\bar{x} - \bar{b}$

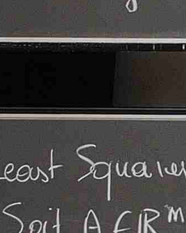
$n=1$



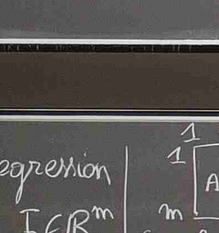
$k = 1, \dots, n$: $\frac{\partial}{\partial x_k} f(\bar{x}) = \frac{\partial}{\partial x_k} \left(\frac{1}{2} \sum_{i=1}^n \left(\sum_{j=1}^n a_{ij} x_i x_j - b_i x_i \right) \right)$
$= (A\bar{x} - \bar{b})_k = \sum_{i=1}^n a_{ki} x_i - b_k$

D'après Thm b. si $f(\bar{x}^*) \leq f(\bar{x})$ $\forall \bar{x} \in \mathbb{R}^n$
alors $\bar{\nabla} f(\bar{x}^*) = \bar{0} = A\bar{x}^* - \bar{b}$
De plus f est strictement convexe car $H(\bar{x})$ est sym. def. positive $\forall \bar{x} \in \mathbb{R}^n$
$H(\bar{x}) = A$
$\left((H(\bar{x}))_{ij} = \frac{\partial^2 f}{\partial x_i \partial x_j}(\bar{x}) = a_{ij} \right) \left(\frac{\partial}{\partial x_j} (Ax - b)_i = a_{ij} \right)$

2. Least Squares, regression
Soit $A \in \mathbb{R}^{m \times n}$ $\bar{b} \in \mathbb{R}^m$
Soit f def $\forall \bar{x} \in \mathbb{R}^n$ par
$f(\bar{x}) = \frac{1}{2} \| A\bar{x} - \bar{b} \|^2 = \frac{1}{2} \sum_{i=1}^m (A\bar{x} - \bar{b})_i^2$



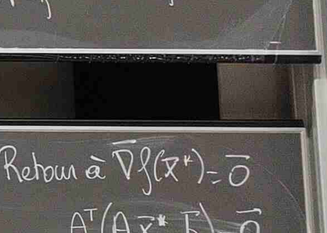
On cherche $\bar{x}^* \in \mathbb{R}^n$ tq $f(\bar{x}^*) \leq f(\bar{x})$ $\forall x \in \mathbb{R}^n$
D'après thm b. $\bar{\nabla} f(\bar{x}^*) = \bar{0}$
On a : $\forall \bar{x} \in \mathbb{R}^n$ $\bar{\nabla} f(\bar{x}) = A^T (A\bar{x} - \bar{b})$



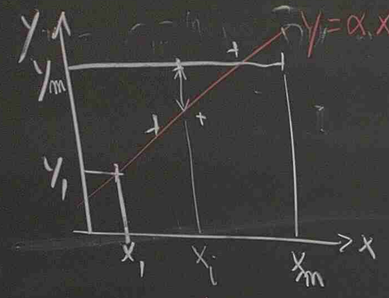
En effet montrons que $k = 1, \dots, n$
$\frac{\partial}{\partial x_k} f(\bar{x}) = \left(A^T (A\bar{x} - \bar{b}) \right)_k$
On a : $\left(A^T (A\bar{x} - \bar{b}) \right)_k = \sum_{i=1}^m (A^T)_{ki} (A\bar{x} - \bar{b})_i$
$= \sum_{i=1}^m a_{ik} (A\bar{x} - \bar{b})_i$

D'autre part :
$\frac{\partial}{\partial x_k} f(\bar{x}) = \frac{\partial}{\partial x_k} \frac{1}{2} \sum_{i=1}^m (A\bar{x} - b)_i^2$
$= \sum_{i=1}^m (A\bar{x} - \bar{b})_i \frac{\partial}{\partial x_k} \left(\sum_{j=1}^n a_{ij} x_j \right)$
$= \sum_{i=1}^m (A\bar{x} - \bar{b})_i \, a_{ik}$

Retour à $\bar{\nabla} f(\bar{x}^*) = \bar{0}$
$A^T (A\bar{x}^* - \bar{b}) = \bar{0}$
$A^T A \bar{x}^* = A^T \bar{b}$



Ex:

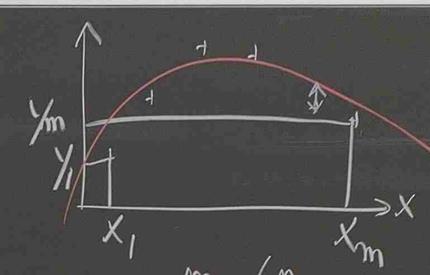


$$\min_{\alpha,\beta} \frac{1}{2} \sum_{i=1}^{m} \left(\alpha x_i + \beta - y_i\right)^2$$

$$A = \begin{bmatrix} x_1 & 1 \\ \vdots & \vdots \\ x_m & 1 \end{bmatrix}_{m \times 2} \quad \vec{x} = \begin{bmatrix} \alpha \\ \beta \end{bmatrix}_{2} \quad \vec{b} = \begin{bmatrix} y_1 \\ \vdots \\ y_m \end{bmatrix}_{m}$$

$$A^T A \begin{bmatrix} \alpha \\ \beta \end{bmatrix} = A^T \vec{b}$$

Ex:



$$\sum_{j=1}^{n} \alpha_j \varphi_j(x)$$

φ_j données (ex $\sin(j\pi x)$)

inconnues

$$\min_{\alpha_1,\dots,\alpha_n} \frac{1}{2} \sum_{i=1}^{m} \left(\sum_{j=1}^{n} \alpha_j \varphi_j(x_i) - y_i \right)^2$$

$$\left(\left(A\vec{\alpha} - \vec{b} \right)_i \right)^2$$

Optimisation sans contraintes

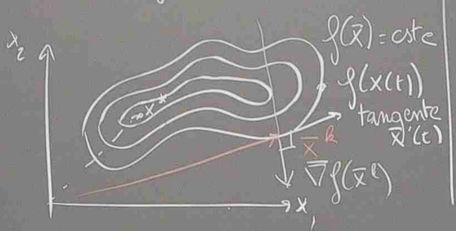
Méthode du gradient

Etant donné $f : \mathbb{R}^n \to \mathbb{R}$ continue
$\vec{x} \to f(\vec{x}) = f(x_1, x_2, \dots, x_n)$

on cherche $\vec{x}^* \in \mathbb{R}^n$ tel que $f(\vec{x}^*) \leq f(\vec{x}) \quad \forall \vec{x} \in \mathbb{R}^n$

Si f est $\mathcal{C}^1$ alors $\vec{\nabla} f(\vec{x}^*) = \vec{0}$

$n=2$:



Comment calculer $\vec{x}^{k+1}$ à partir de $\vec{x}^k$?

$\vec{\nabla} f(\vec{x}^k) \perp$ ligne $f(\vec{x}) = cste$ passant par $\vec{x}^k$

Soit $\vec{x}(t)$ une param de cette courbe

$$f(\vec{x}(t)) = cste$$

$$\frac{d}{dt} f(\vec{x}(t)) = 0$$

$$\frac{d}{dt} f(x_1(t), \dots, x_n(t)) = \frac{\partial f}{\partial x_1}(\vec{x}(t)) x_1'(t) + \dots + \frac{\partial f}{\partial x_n}(\vec{x}(t)) x_n'(t)$$

$$= \vec{\nabla} f(\vec{x}(t))^T \vec{x}'(t)$$

$$= 0$$

donc $\vec{\nabla} f(\vec{x}(t)) \perp \vec{x}'(t)$

méthode du gradient :

$$\vec{x}^{k+1} = \vec{x}^k - \alpha \vec{\nabla} f(\vec{x}^k) \qquad \alpha > 0$$

Alg : $\vec{x}^0 \in \mathbb{R}^n$ donné, $\varepsilon > 0$ petit donné, $\alpha > 0$ donné

$k = 0, 1, 2 \dots$

$\vec{p}^k = -\vec{\nabla} f(\vec{x}^k)$

Si $\|\vec{p}^k\| \leq \varepsilon$ stop

Sinon

$\vec{x}^{k+1} = \vec{x}^k + \alpha \vec{p}^k$

Questions :

- $\vec{x}^k$ converge vers $\vec{x}^*$ si $k \to \infty$?
- Pour quelles valeurs de $\alpha > 0$?

Si l'alg. converge $\lim_{k\to\infty} \vec{x}^k = \vec{x}^*$ alors

$$\lim_{k\to\infty} \vec{x}^{k+1} = \lim_{k\to\infty} \vec{x}^k - \alpha \lim_{k\to\infty} \vec{\nabla} f(\vec{x}^k)$$

et donc $\vec{x}^* = \vec{x}^* - \alpha \vec{\nabla} f(\vec{x}^*)$

et donc $\vec{\nabla} f(\vec{x}^*) = \vec{0}$

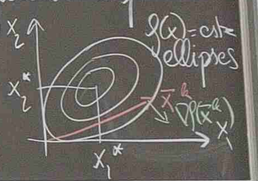
On peut répondre à ces 2 questions dans le cas

- quadratique : $f(\vec{x}) = \frac{1}{2} \vec{x}^T A \vec{x} - \vec{b}^T \vec{x}$, $A \in \mathbb{R}^{n \times n}$ sym déf. pos. $\vec{b} \in \mathbb{R}^n$
- il existe une théorie similaire dans le cas où f est convexe ($\mathcal{C}^2$)

- pas de théorie dans le cas général

Cas quadratique

$n=2$.



Les ellipses sont alignées sur les vecteurs propres

$$A \vec{\varphi}_i = \lambda_i \vec{\varphi}_i \quad i = 1, \dots, n$$

$$\vec{\varphi}_i^T \vec{\varphi}_j = \delta_{ij} \quad i,j = 1, \dots, n \qquad 0 < \lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_n$$

Ex : $A = \begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix}$ $\begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix}\begin{pmatrix} 1 \\ 1 \end{pmatrix} = 1\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ $\begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix}\begin{pmatrix} 1 \\ -1 \end{pmatrix} = 3\begin{pmatrix} 1 \\ -1 \end{pmatrix}$

Thm : $f(\vec{x}) = \frac{1}{2} \vec{x}^T A \vec{x} - \vec{b}^T \vec{x}$ A s.d.p. $\vec{b} \in \mathbb{R}^n$

la méthode du gradient à pas constant converge ($\lim_{k\to\infty} \vec{x}^k = \vec{x}^*$) $\forall \vec{x} \in \mathbb{R}^n \ \forall \vec{b} \in \mathbb{R}^n$

si $\alpha < \frac{2}{\lambda_n}$

Rappel : $\vec{\nabla} f(\vec{x}) = A\vec{x} - \vec{b}$ (MOOC leçon 4)

$\vec{\nabla} f(\vec{x}^*) = A\vec{x}^* - \vec{b} = \vec{0}$

Dem : on a : $\vec{x}^{k+1} = \vec{x}^k - \alpha \vec{\nabla} f(\vec{x}^k)$

$= \vec{x}^k + \alpha(\vec{b} - A\vec{x}^k)$

Soit $\vec{x}^* \in \mathbb{R}^n$ tel que $A\vec{x}^* = \vec{b}$

$\vec{x}^* = \vec{x}^* - \alpha \vec{\nabla} f(\vec{x}^*)$

$= \vec{x}^* + \alpha(\vec{b} - A\vec{x}^*)$

On a $\vec{x}^* - \vec{x}^{k+1} = \vec{x}^* - \vec{x}^k - \alpha A(\vec{x}^* - \vec{x}^k)$

$= (I - \alpha A)(\vec{x}^* - \vec{x}^k)$

Si on peut démontrer que :

$\|\vec{x}^* - \vec{x}^{k+1}\| \leq C \|\vec{x}^* - \vec{x}^k\|$ (*)

avec $C < 1$ alors

$\|\vec{x}^* - \vec{x}^{k+1}\| \leq C^{k+1} \|\vec{x}^* - \vec{x}^0\|$

$\xrightarrow[k\to\infty]{} 0$

On va décomposer $\vec{x}^* - \vec{x}^k$ dans la base des vecteurs pour montrer (*)

$\bar{x}^* - \bar{x}^k = \sum_{j=1}^{n} \beta_j \bar{\varphi}_j$ $\left(\beta_j = (\bar{x}^* - \bar{x}^k)^T \bar{\varphi}_j\right)$

$\|\bar{x}^* - \bar{x}^k\|^2 = (\bar{x}^* - \bar{x}^k)^T (\bar{x}^* - \bar{x}^k)$

$= \left(\sum_{j=1}^{n} \beta_j \bar{\varphi}_j\right)^T \left(\sum_{i=1}^{n} \beta_i \bar{\varphi}_i\right)$

$= \sum_{i=1}^{n}\sum_{j=1}^{n} \beta_i \beta_j \underbrace{\bar{\varphi}_j^T \bar{\varphi}_i}_{\delta_{ij}} = \sum_{j=1}^{n} \beta_j^2$

D'autre part :

$\bar{x}^* - \bar{x}^{k+1} = (I - \alpha A) \sum_{j=1}^{n} \beta_j \bar{\varphi}_j$

$= \sum_{j=1}^{n} \beta_j (I - \alpha A) \bar{\varphi}_j$

$= \sum_{j=1}^{n} \beta_j (1 - \alpha\lambda_j) \bar{\varphi}_j$

et donc $\|\bar{x}^* - \bar{x}^{k+1}\|^2 = \sum_{j=1}^{n} \beta_j^2 (1 - \alpha\lambda_j)^2$

On a donc

$\|\bar{x}^* - \bar{x}^{k+1}\|^2 = \sum_{j=1}^{n} \beta_j^2 (1 - \alpha\lambda_j)^2$

$\leq \max_{1\leq i\leq n} (1 - \alpha\lambda_i)^2 \underbrace{\sum_{j=1}^{n} \beta_j^2}_{\|\bar{x}^* - \bar{x}^k\|^2}$

On prend la $\sqrt{\ }$:

$\|\bar{x}^* - \bar{x}^{k+1}\| \leq \max_{1\leq i\leq n} |1 - \alpha\lambda_i| \, \|\bar{x}^* - \bar{x}^k\|$

Si $\max_{1\leq i\leq n} |1 - \alpha\lambda_i| < 1$ alors la méthode converge

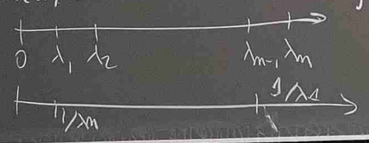
$|1 - \alpha\lambda_i| < 1 \Leftrightarrow$ (toujours le cas car $\alpha > 0$, $\lambda_i > 0$)

$-1 < 1 - \alpha\lambda_i < 1 \Leftrightarrow$

$\alpha\lambda_i < 2 \Leftrightarrow \alpha < \frac{2}{\lambda_i}$

$n=2$: $\vec{a} = \frac{\vec{a}\cdot\vec{e}_1}{\|\vec{e}_1\|^2}\vec{e}_1 + \frac{\vec{a}\cdot\vec{e}_2}{\|\vec{e}_2\|^2}\vec{e}_2$

Si $\alpha < \frac{2}{\lambda_i}$ $i = 1, \dots, n$ alors $\max_{1\leq i\leq n} |1 - \alpha\lambda_i| < 1$, on a démontré $(*)$ et la méthode converge



Si $\alpha < \frac{2}{\lambda_n}$ alors $\alpha < \frac{2}{\lambda_i}$ $i = 1, \dots, n$

Dans le cas quadratique, on peut calculer le α optimal

En effet, soit $\alpha^k \in \mathbb{R}$ tel que $(\bar{x}^k$ donné, $\bar{p}^k$ donné$)$

$f(\bar{x}^k + \alpha^k \bar{p}^k) \leq f(\bar{x}^k + \alpha \bar{p}^k)$ $\forall \alpha \in \mathbb{R}$

$g(\alpha^k) \leq g(\alpha)$

Puisque $g : \mathbb{R} \to \mathbb{R}$ est $\mathcal{C}^1$ on a

$\alpha \to g(\alpha)$

$g'(\alpha^k) = 0$

$\forall \alpha \in \mathbb{R}$, calculons $g'(\alpha)$

$g(\alpha) = f(\bar{x}^k + \alpha \bar{p}^k) = f(x_1^k + \alpha p_1^k, \dots, x_n^k + \alpha p_n^k)$

$g'(\alpha) = \sum_{j=1}^{n} \frac{\partial f}{\partial x_j}(\bar{x}^k + \alpha\bar{p}^k) p_j^k = (\nabla f(\bar{x}^k + \alpha\bar{p}^k))^T \bar{p}^k$

$= (A(\bar{x}^k + \alpha\bar{p}^k) - \bar{b})^T \bar{p}^k$

α^k est donc tel que

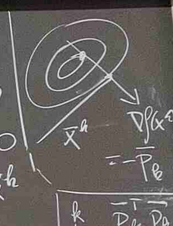
$(A(\bar{x}^k + \alpha^k \bar{p}^k) - \bar{b})^T \bar{p}^k = 0$

$(A\bar{x}^k - \bar{b} + \alpha^k A\bar{p}^k)^T \bar{p}^k = 0$

Puisque $\bar{p}^k = \bar{b} - A\bar{x}^k$

$-\bar{p}^{kT}\bar{p}^k + \alpha^k \bar{p}^{kT} A \bar{p}^k = 0$

$\alpha^k = \frac{\bar{p}^{kT}\bar{p}^k}{\bar{p}^{kT} A \bar{p}^k}$



Alg du gradient pas optimal

$\bar{x}^0$ donné $\varepsilon > 0$

$k = 0, 1, 2, \dots$

$\bar{p}^k = \bar{b} - A\bar{x}^k$

Si $\|\bar{p}^k\| \leq \varepsilon$ stop

Sinon $\alpha^k = \|\bar{p}^k\|^2 / \bar{p}^{kT} A \bar{p}^k$ $(\alpha^k > 0)$

$\bar{x}^{k+1} = \bar{x}^k + \alpha^k \bar{p}^k$

Nombre d'itérations k ?

On peut montrer que le nombre d'itérations $\sim O\left(\frac{\lambda_n}{\lambda_1}\right)$

Par exemple si $A = \begin{pmatrix} 2 & -1 & (0) \\ -1 & \ddots & \ddots \\ (0) & \ddots & 2 \end{pmatrix}$ on a $\lambda_i = 2 - 2\cos\frac{i\pi}{n+1}$ $(i = 1, \dots, n)$

$\lambda_n \to 4$ $\lambda_1 \to 0$



et $O\left(\frac{\lambda_n}{\lambda_1}\right) = O(n^2)$

nombre d'opérations total pour minimiser $f(\bar{x}) = \frac{1}{2}\bar{x}^T A\bar{x} - \bar{b}^T\bar{x}$ ou résoudre $A\bar{x} = \bar{b}$:

- $O(n^2)$ itérations et à chaque iter il faut calculer $\bar{p}^k = \bar{b} - A\bar{x}^k$ ($O(n)$ opérations)

  Au total $O(n^2 n) = O(n^3)$ opérations !

- Elimination ou Cholesky $A = LL^T$ $O(n)$ opérations !

Si on résout $-\Delta u = f$ dans $[0,1]^d$ avec une méthode diff. finies et n points dans chaque direction on a $O(n^d)$ inconnues et il faut faire $O(n^2)$ iterations et donc $O(n^{d+2})$ opérations au total. Pour $d = 3$ $O(n^5)$

Cholesky $O(n^7)$ | gradient à pas optimal $O(n^5)$ | gradient conjugué $O(n^4)$

Optim avec contraintes (1/2)

Rappel : sans contraintes

$f : \mathbb{R}^n \to \mathbb{R}$ $\mathcal{C}^1$

On cherche $\vec{x}^* \in \mathbb{R}^n$ tq $f(\vec{x}^*) \leq f(\vec{x}) \ \forall \vec{x} \in \mathbb{R}^n$

$\vec{\nabla} f(\vec{x}^*) = \vec{0}$

$\vec{x}^{k+1} = \vec{x}^k - \alpha \vec{\nabla} f(\vec{x}^k)$

$f(\vec{x}) = \frac{1}{2} \vec{x}^T A \vec{x} - \vec{b}^T \vec{x}$

A s.d.p, converge si $\alpha < \frac{2}{\lambda_n}$

Avec contraintes :

On cherche $\vec{x}^* \in \Omega$ tq $f(\vec{x}^*) \leq f(\vec{x}) \ \forall \vec{x} \in \Omega$

où $\Omega \subset \mathbb{R}^n$ défini à partir de

$\vec{g} : \mathbb{R}^n \to \mathbb{R}^m$

$\vec{x} \to \vec{g}(\vec{x}) = \begin{pmatrix} g_1(x_1, x_2, \dots, x_n) \\ \vdots \\ g_m(x_1, x_2, \dots, x_n) \end{pmatrix}$

qui correspond aux contraintes d'inégalités

et $\vec{h} : \mathbb{R}^n \to \mathbb{R}^p$

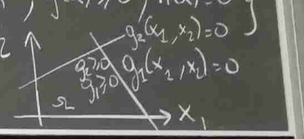
$\vec{x} \to \vec{h}(\vec{x}) = \begin{pmatrix} h_1(x_1, x_2, \dots, x_n) \\ h_p(x_1, x_2, \dots, x_n) \end{pmatrix}$

qui correspond aux contraintes d'égalité

et $\Omega = \{ \vec{x} \in \mathbb{R}^n ; g_i(\vec{x}) \geqslant 0 \ i = 1, \dots, m \text{ et } h_i(\vec{x}) = 0 \ i = 1, \dots, p \}$

$= \{ \vec{x} \in \mathbb{R}^n ; \vec{g}(\vec{x}) \geqslant \vec{0}, \vec{h}(\vec{x}) = 0 \}$

$n=2$; $m=2$; $p=0$



Thm : conditions d'optimalité

KKT (Karush, Kuhn, Tucker) 1939 — 1953

Soit $f, g, h \in \mathcal{C}^1$, soit $\vec{x}^* \in \Omega$ tq $f(\vec{x}^*) \leq f(\vec{x}) \ \forall \vec{x} \in \Omega$

Sous des hypothèses raisonnables entre $\vec{x}^*, g, h$

$\exists \vec{\lambda}^* \in \mathbb{R}^m$ $\exists \vec{\mu}^* \in \mathbb{R}^p$ tels que

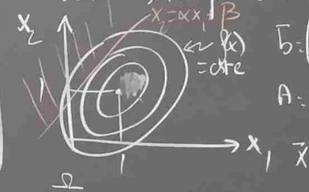
(appelés mult de lagrange)

$$\begin{cases} \vec{\nabla} f(\vec{x}^*) = \sum_{i=1}^{m} \lambda_i^* \vec{\nabla} g_i(\vec{x}^*) + \sum_{i=1}^{p} \mu_i^* \vec{\nabla} h_i(\vec{x}^*) \\ h_i(\vec{x}^*) = 0 \quad i = 1, \dots, p \\ g_i(\vec{x}^*) \geqslant 0 \quad i = 1, \dots, m \\ \lambda_i^* \geqslant 0 \quad \text{——} \\ \lambda_i^* \, g_i(\vec{x}^*) = 0 \quad \text{——} \end{cases}$$

← cond. de complémentarité

Ex : $n=2$ $f(\vec{x}) = \frac{1}{2} \vec{x}^T A \vec{x} - \vec{b}^T \vec{x}$

$A \in \mathbb{R}^{n \times n}$ s.d.p $\quad \vec{b} \in \mathbb{R}^n$



$\vec{b} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$

$A = \begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix}$

$\vec{x}^* = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$

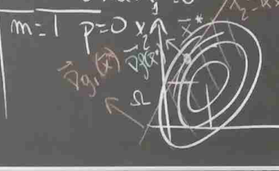
$m=1$ $p=0$ $x_2 \leqslant \alpha x_1 + \beta$ $\quad \alpha x_1 + \beta - x_2 \geqslant 0$

$\vec{x}^* = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ $\quad \vec{\nabla} f(\vec{x}^*) = \vec{0}$ $\quad g_1(x_1, x_2)$

$= \lambda_1^* \vec{\nabla} g_1(x_1^*, x_2^*) = \lambda_1^* \begin{pmatrix} \alpha \\ -1 \end{pmatrix}$

on a $\lambda_1^* = 0$

$m=1$ $p=0$



$x_2 \geqslant \alpha x_1 + \beta$

$x_2 - \alpha x_1 - \beta \geqslant 0$

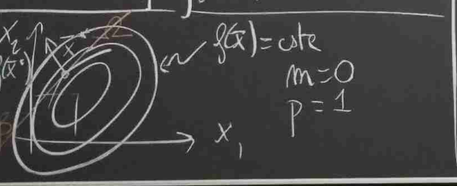
$g_1(x_1, x_2)$

$\vec{\nabla} g_1(x_1, x_2) = \begin{pmatrix} -\alpha \\ 1 \end{pmatrix}$

On a bien $\vec{\nabla} f(\vec{x}^*) = \lambda_1^* \vec{\nabla} g_1(\vec{x}^*)$

avec $\lambda_1^* \geqslant 0$ de plus $g_1(\vec{x}^*) = 0$

et donc $\lambda_1^* g_1(\vec{x}^*) = 0$



$\Omega = \{ \vec{x} \in \mathbb{R}^2 ; \ x_2 = \alpha x_1 + \beta \}$

$h_1(x_1, x_2) = x_2 - \alpha x_1 - \beta$

ou

$= \alpha x_1 + \beta - x_2$

$\vec{\nabla} h_1(x_1, x_2) = \begin{pmatrix} -\alpha \\ 1 \end{pmatrix}$ ou $\begin{pmatrix} \alpha \\ -1 \end{pmatrix}$

On a bien $\vec{\nabla} f(\vec{x}^*) = \mu_1^* \vec{\nabla} h_1(\vec{x}^*)$ avec $\mu_1^* \geqslant 0$ ou $\leqslant 0$

$\mu_1^* \in \mathbb{R}$

On peut écrire la première condition KKT plus simplement : on introduit le lagrangien défini par

$\mathcal{L} : \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^p \to \mathbb{R}$

$(\vec{x}, \vec{\lambda}, \vec{\mu}) \to \mathcal{L}(\vec{x}, \vec{\lambda}, \vec{\mu})$ défini par

$\mathcal{L}(\vec{x}, \vec{\lambda}, \vec{\mu}) = f(\vec{x}) - \sum_{i=1}^{m} \lambda_i g_i(\vec{x}) + \sum_{i=1}^{p} \mu_i h_i(\vec{x})$

On a $\partial_{\vec{x}} \mathcal{L}(\vec{x}, \vec{\lambda}, \vec{\mu}) = \vec{\nabla} f(\vec{x}) + \sum_{i=1}^{m} \lambda_i \vec{\nabla} g_i(\vec{x}) + \sum_{i=1}^{p} \mu_i \vec{\nabla} h_i(\vec{x})$

(ie $\frac{\partial}{\partial x_k} \mathcal{L}(\vec{x}, \vec{\lambda}, \vec{\mu}) = \frac{\partial f}{\partial x_k}(\vec{x}) + \sum_{i=1}^{m} \lambda_i \frac{\partial g_i}{\partial x_k}(\vec{x}) + \sum_{i=1}^{p} \mu_i \frac{\partial h_i}{\partial x_k}(\vec{x}) \Big)_{k=1,\dots,n}$

La première condition KKT s'écrit

$\partial_{\vec{x}} \mathcal{L}(\vec{x}^*, \vec{\lambda}^*, \vec{\mu}^*) = \vec{0}$

Ex 1 : Contraintes d'égalité

$f(\vec{x}) = \frac{1}{2}\vec{x}^T A\vec{x} - \vec{b}^T\vec{x}$ $\quad b \in \mathbb{R}^n$ $\quad A \in \mathbb{R}^{n\times n}$ s.d.p

$\vec{h}(\vec{x}) = B\vec{x} - \vec{c} = \begin{matrix}1\\p\end{matrix}\boxed{\begin{matrix}\\B\\\\\end{matrix}}\begin{matrix}n\\\boxed{\vec{x}}\\m\end{matrix}^1 = \begin{matrix}1\\\boxed{\vec{c}}\\p\end{matrix}$

$m = 0$

On cherche $\vec{x}^* \in \Omega$ tq $f(\vec{x}^*) \leq f(\vec{x})\ \forall \vec{x} \in \Omega$

où $\Omega = \{\vec{x} \in \mathbb{R}^n$ tq $\vec{h}(\vec{x}) = B\vec{x} - \vec{c} = 0\}$

Conditions KKT : on introduit le lagrangien

$\mathcal{L} : \mathbb{R}^n \times \mathbb{R}^p \to \mathbb{R}$

$(\vec{x}, \vec{\mu}) \to \mathcal{L}(\vec{x}, \vec{\mu})$

$\mathcal{L}(\vec{x}, \vec{\mu}) = f(\vec{x}) - \sum_{i=1}^{p} \mu_i h_i(\vec{x})$

$= \frac{1}{2}\vec{x}^T A\vec{x} - \vec{b}^T\vec{x} - \sum_{i=1}^{p} \mu_i (B\vec{x} - \vec{c})_i$

$\forall \vec{x} \in \mathbb{R}^n$ $\quad \forall \vec{\mu} \in \mathbb{R}^p$

On doit calculer $\partial_x \mathcal{L}(\vec{x}, \vec{\mu})$

On a : $\partial_x \mathcal{L}(\vec{x}, \vec{\mu}) = A\vec{x} - \vec{b} - B^T\vec{\mu}$

$= \boxed{A}\boxed{\vec{x}} - \boxed{\vec{b}} - \begin{matrix}1\\n\end{matrix}\boxed{\begin{matrix}\\B^T\\\\\end{matrix}}\begin{matrix}p\\\boxed{\vec{\mu}}\\p\end{matrix}^1$

En effet $k = 1, \dots, n$

$\frac{\partial}{\partial x_k}\mathcal{L}(\vec{x}, \vec{\mu}) = \frac{\partial}{\partial x_k}\left(\frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} A_{ij}x_i x_j - \sum_{i=1}^{n} b_i x_i\right) - \frac{\partial}{\partial x_k}\left(\sum_{i=1}^{p}\mu_i \sum_{j=1}^{n} B_{ij}x_j - c_i\right)$

$= (A\vec{x} - \vec{b})_k - \sum_{i=1}^{p} \mu_i B_{ik}$

$= (A\vec{x} - \vec{b})_k - (B^T\vec{\mu})_k$

La première condition KKT s'écrit

$\partial_x \mathcal{L}(\vec{x}^*, \vec{\mu}^*) = \vec{0}$

i.e.

$A\vec{x}^* - \vec{b} - B^T\vec{\mu}^* = 0$

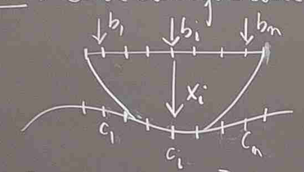
Conclusion : on cherche $\vec{x}^* \in \mathbb{R}^n$ $\vec{\mu}^* \in \mathbb{R}^p$ tels que

$$\begin{cases} A\vec{x}^* - \vec{b} - B^T\vec{\mu}^* = \vec{0} & n \text{ eq} \\ B\vec{x}^* - \vec{c} = \vec{0} & p \text{ eq} \end{cases} \quad n+p \text{ inc.}$$

Il s'agit de résoudre le syst. lin

$$\begin{pmatrix} A & -B^T \\ B & 0 \end{pmatrix}\begin{pmatrix} \vec{x}^* \\ \vec{\mu}^* \end{pmatrix} = \begin{pmatrix} \vec{b} \\ \vec{c} \end{pmatrix}$$

Ex 2 : corde élastique avec obstacle



$f(\vec{x}) = \frac{1}{2}\vec{x}^T A\vec{x} - \vec{b}^T\vec{x}$

$m = n$ contraintes $x_i \geq c_i$ $\quad i = 1, \dots, n$

$g_i(x_1, x_2, \dots, x_n) = x_i - c_i \geq 0$

$\Omega = \{\vec{x} \in \mathbb{R}^n,\ \vec{g}(\vec{x}) \geqslant \vec{0}\}$

$= \{$ —— $g_i(\vec{x}) \geqslant 0$ $\quad i = 1, \dots, n\}$

On cherche $\vec{x}^* \in \Omega$ tq $f(\vec{x}^*) \leq f(\vec{x})\ \forall \vec{x} \in \Omega$

On introduit le lagrangien

$\mathcal{L} : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$

$(\vec{x}, \vec{\lambda}) \to \mathcal{L}(\vec{x}, \vec{\lambda}) = f(\vec{x}) - \sum_{i=1}^{m} \lambda_i (x_i - c_i)$

$\partial_x \mathcal{L}(\vec{x}, \vec{\lambda}) = A\vec{x} - \vec{b} - \vec{\lambda} \in \mathbb{R}^n$

Les conditions KKT sont :

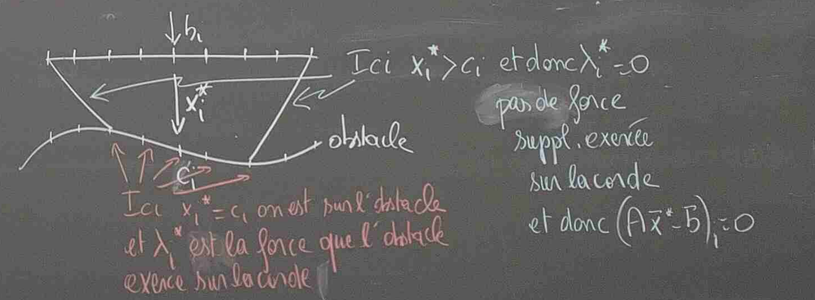
trouver $\vec{x}^* \in \mathbb{R}^n$, $\vec{\lambda}^* \in \mathbb{R}^m$ tels que

$\partial_x \mathcal{L}(\vec{x}^*, \vec{\lambda}^*) = \vec{0}$ ie $\begin{cases} A\vec{x}^* - \vec{b} - \vec{\lambda}^* = \vec{0} \\ x_i^* - c_i \geqslant 0 \quad i = 1, \dots, n \quad (\vec{x}^* - \vec{c} \geqslant \vec{0}) \\ \lambda_i^* \geqslant 0 \quad \text{——} \\ \lambda_i^*(x_i^* - c_i) = 0 \quad \text{——} \end{cases}$

Il s'agit d'un système d'eq. non linéaires avec des inégalités

Dans la suite on utilisera la méthode des points intérieurs pour approcher $\vec{x}^*$ et $\vec{\lambda}^*$

Interprétation physique



Ici $x_i^* > c_i$ et donc $\lambda_i^* = 0$ pas de force suppl. exercée sur la corde et donc $(A\vec{x}^* - \vec{b})_i = 0$

Ici $x_i^* = c_i$ on est sur l'obstacle et λ_i^* est la force que l'obstacle exerce sur la corde

Optim. avec contraintes (2/2)

Données : $f : \mathbb{R}^n \to \mathbb{R}$, $\vec{x} \to f(\vec{x})$; $\vec{g} : \mathbb{R}^n \to \mathbb{R}^m$, $\vec{x} \to \vec{g}(\vec{x}) = \begin{pmatrix} g_1(\vec{x}) \\ \vdots \\ g_m(\vec{x}) \end{pmatrix}$; $\vec{h} : \mathbb{R}^n \to \mathbb{R}^p$, $\vec{x} \to \vec{h}(\vec{x}) = \begin{pmatrix} h_1(\vec{x}) \\ \vdots \\ h_p(\vec{x}) \end{pmatrix}$

On définit : $\Omega = \{\vec{x} \in \mathbb{R}^n ;\ g_i(\vec{x}) \geqslant 0\ i=1,\dots,m,\ h_i(\vec{x}) = 0,\ i=1,\dots,p\}$

On cherche $\vec{x}^* \in \Omega$ tq $f(\vec{x}^*) \leqslant f(\vec{x})\ \forall \vec{x} \in \Omega$ (1)

Les conditions KKT sont :
trouver $\vec{x}^* \in \mathbb{R}^n$, $\vec{\lambda}^* \in \mathbb{R}^m$, $\vec{\mu}^* \in \mathbb{R}^p$ tq

$$(2)\ \begin{cases} \vec{\nabla} f(\vec{x}^*) = \sum_{i=1}^{m} \lambda_i^* \vec{\nabla} g_i(\vec{x}^*) + \sum_{i=1}^{p} \mu_i^* \vec{\nabla} h_i(\vec{x}^*) \\ h_i(\vec{x}^*) = 0 \quad i = 1,\dots,p \\ g_i(\vec{x}^*) \geqslant 0 \quad i = 1,\dots,m \\ \lambda_i^* \geqslant 0 \\ \lambda_i^* g_i(\vec{x}^*) = 0 \end{cases}$$

La première cond. KKT peut s'écrire $\partial_{\vec{x}} \mathcal{L}(\vec{x}^*, \vec{\lambda}^*, \vec{\mu}^*) = 0$

où $\mathcal{L}(\vec{x}, \vec{\lambda}, \vec{\mu}) = f(\vec{x}) - \sum_{i=1}^{m} \lambda_i g_i(\vec{x}) - \sum_{i=1}^{p} \mu_i h_i(\vec{x})$

$\forall \vec{x} \in \mathbb{R}^n\ \forall \vec{\lambda} \in \mathbb{R}^m\ \forall \vec{\mu} \in \mathbb{R}^p$

On va dem que la sol. de (1) est sol. de (2) pour les 2 exemples suivants
Ex 1. contr. d'égalité ($m=0$)
Ex 2. —— d'inégalité ($p=0$)

Ex 1 : $f(\vec{x}) = \frac{1}{2} \vec{x}^T A \vec{x} - \vec{b}^T \vec{x}$

où $A \in \mathbb{R}^{n \times n}$ sym. def. pos. et $\vec{b} \in \mathbb{R}^n$

$m = 0$ $\vec{h}(\vec{x}) = B\vec{x} - \vec{c} = \begin{matrix} p \\ \{ \end{matrix} \overbrace{\boxed{B}}^{n} \overbrace{\begin{bmatrix} \vec{x} \\ \vdots \\ \vec{x} \end{bmatrix}}^{n} - \overbrace{\begin{bmatrix} \vec{c} \\ \vdots \\ \vec{c} \end{bmatrix}}^{p}$

$B \in \mathbb{R}^{p \times n}$ $\vec{c} \in \mathbb{R}^p$

$\forall \vec{x} \in \mathbb{R}^n\ \forall \vec{\mu} \in \mathbb{R}^p\ \mathcal{L}(\vec{x}, \vec{\mu}) = f(\vec{x}) - \sum_{i=1}^{p} \mu_i (B\vec{x} - \vec{c})_i$

$\partial_{\vec{x}} \mathcal{L}(\vec{x}, \vec{\mu}) = \vec{\nabla} f(\vec{x}) - B^T \vec{\mu}$

les cond KKT (2) s'écrivent
trouver $\vec{x}^* \in \mathbb{R}^n$, $\vec{\mu}^* \in \mathbb{R}^p$ tq

$$A\vec{x}^* - \vec{b} - B^T \vec{\mu}^* = \vec{0}$$
$$B\vec{x}^* - \vec{c} = \vec{0}$$

$$\begin{matrix} n \\ p \end{matrix} \overbrace{\begin{pmatrix} A & -B^T \\ B & 0 \end{pmatrix}}^{n} \begin{pmatrix} \vec{x}^* \\ \vec{\mu}^* \end{pmatrix} = \begin{pmatrix} \vec{b} \\ \vec{c} \end{pmatrix} \quad (3)$$

Ex 2 : $f(\vec{x}) = \frac{1}{2} \vec{x}^T A \vec{x} - \vec{b}^T \vec{x}$

$A \in \mathbb{R}^{n \times n}$ s.d.p. $\vec{b} \in \mathbb{R}^n$

$p = 0$ $\vec{g}(\vec{x}) = \vec{x} - \vec{c}$

$g_i(\vec{x}) = x_i - c_i \quad i = 1,\dots,n = m$

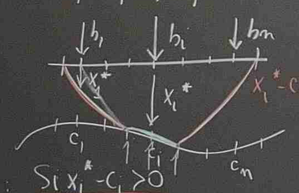
$\forall \vec{x} \in \mathbb{R}^n\ \forall \vec{\lambda} \in \mathbb{R}^m\ \mathcal{L}(\vec{x}, \vec{\lambda}) = f(\vec{x}) - \sum_{i=1}^{m} \lambda_i (x_i - c_i)$

$\partial_{\vec{x}} \mathcal{L}(\vec{x}, \vec{\lambda}) = \vec{\nabla} f(\vec{x}) - \vec{\lambda} = A\vec{x} - \vec{b} - \vec{\lambda}$

les cond. KKT (2) s'écrivent :
trouver $\vec{x}^* \in \mathbb{R}^n$, $\vec{\lambda}^* \in \mathbb{R}^n$ tq

$$(4)\ \begin{cases} A\vec{x}^* - \vec{b} = \vec{\lambda}^* \\ x_i^* - c_i \geqslant 0 \quad i = 1,\dots,n \\ \lambda_i^* \geqslant 0 \\ \lambda_i^* (x_i^* - c_i) = 0 \end{cases}$$

Interpr. physique :



$i = 1,\dots,n$: Si $x_i^* - c_i > 0$ alors $\lambda_i^* = 0$ et donc $(A\vec{x}^* - \vec{b})_i = 0$

Si $\lambda_i^* > 0$ alors $x_i^* - c_i = 0$
on est sur l'obstacle
$(A\vec{x}^* - \vec{b})_i = \lambda_i^* > 0$
λ_i^* peut être interprété comme la force qu'il faut appliquer pour que la corde reste sur l'obstacle

Rappel : dérivée directionnelle $f : \mathbb{R}^n \to \mathbb{R}$ quelconque

$\forall \vec{x} \in \mathbb{R}^n\ \forall \vec{d} \in \mathbb{R}^n$ on def.

$$\vec{\nabla} f(\vec{x})^T \vec{d} = \lim_{\varepsilon \to 0} \frac{f(\vec{x} + \varepsilon \vec{d}) - f(\vec{x})}{\varepsilon}$$

$\vec{d} = \vec{e}_i = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$ $\vec{\nabla} f(\vec{x})^T \vec{e}_j = \frac{\partial f}{\partial x_i}(\vec{x})$

Lemme : Soit Ω convexe

$(\forall \vec{x}, \vec{y} \in \Omega, \forall 0 \leqslant s \leqslant 1,\ s\vec{x} + (1-s)\vec{y} \in \Omega)$

Soit $\vec{x}^*$ sol. de (1) ($\vec{x}^* \in \Omega$, $f(\vec{x}^*) \leqslant f(\vec{x})$ $\forall \vec{x} \in \Omega$)

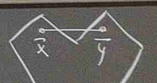
On a : $\vec{\nabla} f(\vec{x}^*)^T (\vec{x} - \vec{x}^*) \geqslant 0\ \forall \vec{x} \in \Omega$

(f augmente lorsqu'on passe de $\vec{x}^*$ à $\vec{x}$)

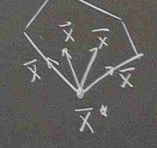


convexe





pas convexe



Dem : Soit $\bar{x} \in \Omega$, soit $\bar{x}^*$ sol de (1)

Ω convexe $0 \leq s \leq 1$ $s\bar{x} + (1-s)\bar{x}^* \in \Omega$

et donc $f(s\bar{x} + (1-s)\bar{x}^*) \geq f(\bar{x}^*)$

i.e. $f(\bar{x}^* + s(\bar{x} - \bar{x}^*)) - f(\bar{x}^*) \geq 0$

on prend la limite $s \to 0$ et on obtient $\vec{\nabla} f(\bar{x}^*)^T (\bar{x} - \bar{x}^*) \geq 0$

Ex 1 : Soit $\bar{x}^*$ sol de (1)

avec $f(\bar{x}) = \frac{1}{2} \bar{x}^T A \bar{x} - \vec{b}^T \bar{x}$

et $\Omega = \{ \bar{x} \in \mathbb{R}^n,\ B\bar{x} - \bar{c} = \bar{0} \}$

Montrons $\exists \bar{\mu}^* \in \mathbb{R}^p$ tq

$A\bar{x}^* - \vec{b} = B^T \bar{\mu}^*$

Utilisons le lemme

Ω est convexe :

$\forall \bar{x}, \bar{y} \in \mathbb{R}^n$ $B\bar{x} - \bar{c} = \bar{0}$, $B\bar{y} - \bar{c} = \bar{0}$

$\forall 0 \leq s \leq 1$ $B(s\bar{x} + (1-s)\bar{y}) = sB\bar{x} + (1-s)B\bar{y}$

$= s\bar{c} + (1-s)\bar{c}$

$= \bar{c}$

donc $s\bar{x} + (1-s)\bar{y} \in \Omega$

D'après le lemme : $\vec{\nabla} f(\bar{x}^*)^T (\bar{x} - \bar{x}^*) \geq 0 \quad \forall \bar{x}$ tq $B\bar{x} - \bar{c} = \bar{0}$

$(A\bar{x}^* - \vec{b})^T (\bar{x} - \bar{x}^*) \geq 0$

On choisit $\bar{x} = \bar{x}^* + \bar{y}$

où $\bar{y} \in \ker B$ $(B\bar{y} = \bar{0})$

$\bar{x} = \bar{x}^* + \bar{y} \in \Omega$?, oui

$B\bar{x} = B\bar{x}^* + B\bar{y} = \bar{c} + \bar{0} = \bar{c}$

On a donc $(A\bar{x}^* - \vec{b})^T \bar{y} \geq 0$

Si on choisit $\bar{x} = \bar{x}^* - \bar{y}$ on obtient

$(A\bar{x}^* - \vec{b})^T (-\bar{y}) \geq 0$

et donc $(A\bar{x}^* - \vec{b})^T \bar{y} = 0 \quad \forall \bar{y} \in \ker B$

Donc $(A\bar{x}^* - \vec{b}) \in (\ker B)^\perp = \operatorname{Im} B^T$ (cf exercice $(\ker B = (\operatorname{Im} B^T)^\perp)$)

$\exists \bar{\mu}^* \in \mathbb{R}^p$ tq $A\bar{x}^* - \vec{b} = B^T \bar{\mu}^*$

Ex 2 : Soit $\bar{x}^*$ sol de (1)

avec $f(\bar{x}) = \frac{1}{2} \bar{x}^T A \bar{x} - \vec{b}^T \bar{x}$

$\Omega = \{ \bar{x} \in \mathbb{R}^n ;\ \bar{x} - \vec{c} \geq \vec{0} \}$

Montrons $\exists \bar{\lambda}^* \in \mathbb{R}^n$ tq (4) :

$$\begin{cases} A\bar{x}^* - \vec{b} = \bar{\lambda}^* \\ x_i^* - c_i \geq 0 \quad i = 1, \dots, n \\ \lambda_i^* \geq 0 \quad \text{——} \\ \lambda_i^* (x_i^* - c_i) = 0 \quad \text{——} \end{cases}$$

Eliminons $\bar{\lambda}^*$:

$$\begin{cases} A\bar{x}^* - \vec{b} \geq \vec{0} \\ \bar{x}^* - \vec{c} \geq \vec{0} \\ (A\bar{x}^* - \vec{b})_i (x_i^* - c_i) = 0 \quad i = 1, \dots, n \end{cases}$$

Ω est convexe :

$\forall \bar{x}, \bar{y} \in \mathbb{R}^n$ $\bar{x} - \vec{c} \geq \vec{0}$ $\bar{y} - \vec{c} \geq \vec{0}$

$\forall 0 \leq s \leq 1$ $s\bar{x} + (1-s)\bar{y} \geq s\vec{c} + (1-s)\vec{c} = \vec{c}$

donc $s\bar{x} + (1-s)\bar{y} \in \Omega$

D'après le lemme : $(A\bar{x}^* - \vec{b})^T (\bar{x} - \bar{x}^*) \geq 0 \quad \forall \bar{x} \in \mathbb{R}^n$ tq $\bar{x} \geq \vec{c}$

Montrons $(A\bar{x}^* - \vec{b})_i \geq 0 \quad i = 1, \dots, n$

Choisissons $\bar{x} = \bar{x}^* + \vec{e}_i$ —— $(\bar{x} = \bar{x}^* + \vec{e}_i \geq \vec{c} + \vec{e}_i \geq \vec{c})$ donc $\bar{x} \in \Omega$

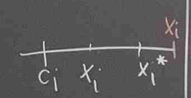
On obtient $(A\bar{x}^* - \vec{b})^T \vec{e}_i \geq 0 \quad i = 1, \dots, n$

i.e. $(A\bar{x}^* - \vec{b})_i \geq 0$ ——

Reste à montrer que soit $(A\bar{x}^* - \vec{b})_i = 0$ soit $x_i^* - c_i = 0 \quad i = 1, \dots, n$

Soit $i = 1, \dots, n$ fixé

- Si $x_i^* - c_i = 0$ (on est sur l'obstacle) on a bien $(A\bar{x}^* - \vec{b})_i (x_i^* - c_i) = 0$
- Si $x_i^* - c_i > 0$ on choisit $\bar{x}$ tq $c_i < x_i < x_i^*$ et $x_j = x_j^*$ $j \neq i$



$\bar{x} = (x_1^*, \dots, x_{i-1}^*, x_i, x_{i+1}^*, \dots, x_n^*)$

D'après le lemme $(A\bar{x}^* - \vec{b})^T (\bar{x} - \bar{x}^*) \geq 0$

i.e. $(A\bar{x}^* - \vec{b})_i (x_i - x_i^*) \geq 0$ ($x_i - x_i^* \leq 0$)

et donc $(A\bar{x}^* - \vec{b})_i \leq 0$, ≥ 0 $\Rightarrow (A\bar{x}^* - \vec{b})_i = 0$

Optim avec contraintes

Alg. des points intérieurs

Ex 1. quadratic programming

$\bar{x}^* \in \Omega$ tq $f(\bar{x}^*) \leqslant f(\bar{x}) \ \forall \bar{x} \in \Omega$

où $f(\bar{x}) = \frac{1}{2} \bar{x}^T A \bar{x} - \bar{b}^T \bar{x}$ où $A \in \mathbb{R}^{n \times n}$ s.d.p et $b \in \mathbb{R}^n$

et $\Omega = \{ \bar{x} \in \mathbb{R}^n \text{ tq } B\bar{x} - \bar{c} \geqslant \bar{0} \}$

où $B \in \mathbb{R}^{m \times n}$ et $\bar{c} \in \mathbb{R}^m$

$\forall \bar{x} \in \mathbb{R}^n \ \forall \bar{\lambda} \in \mathbb{R}^m$

$\mathcal{L}(\bar{x}, \bar{\lambda}) = \frac{1}{2} \bar{x}^T A \bar{x} - \bar{b}^T \bar{x} - \sum_{i=1}^{m} \lambda_i (B\bar{x} - \bar{c})_i$

$\nabla_x \mathcal{L}(\bar{x}, \bar{\lambda}) = A\bar{x} - \bar{b} - B^T \bar{\lambda}$

Conditions KKT : trouver $\bar{x}^* \in \mathbb{R}^n, \bar{\lambda}^* \in \mathbb{R}^m$ tq

$$\begin{cases} A\bar{x}^* - \bar{b} - B^T \bar{\lambda}^* = \bar{0} & \updownarrow \text{n eq.} \\ B\bar{x}^* - \bar{c} \geqslant \bar{0} & \updownarrow \text{m ineq.} \\ \bar{\lambda}^* \geqslant \bar{0} & \updownarrow \\ \lambda_i^* (B\bar{x}^* - \bar{c})_i = 0 \quad i = 1, \dots, m & \updownarrow \text{m eq.} \end{cases}$$

On introduit une "slack variable" $\bar{s} \in \mathbb{R}^m$

$$\begin{cases} A\bar{x}^* - \bar{b} - B^T \bar{\lambda}^* = 0 & \updownarrow \text{n eq.} \\ B\bar{x}^* - \bar{c} = \bar{s} & \updownarrow \text{m eq.} \\ \lambda_1^* s_1 = 0 \\ \vdots \\ \lambda_m^* s_m = 0 & \updownarrow \text{m eq.} \\ \bar{\lambda}^* \geqslant \bar{0} \\ \bar{s} \geqslant \bar{0} \end{cases} \quad \bar{F}(\bar{x}^*, \bar{\lambda}^*, \bar{s}) = \bar{0}$$

n+m+m inconnues $\bar{x}^*, \bar{\lambda}^*, \bar{s}$

Rappel Newton $\bar{F}(\bar{x}) = \bar{0}$

$\bar{0} \simeq \bar{F}(\bar{x}^{k+1}) = \bar{F}(\bar{x}^k) + DF(\bar{x}^k)(\bar{x}^{k+1} - \bar{x}^k) + \dots$

$DF(\bar{x}^k)(\bar{x}^{k+1} - \bar{x}^k) = -\bar{F}(\bar{x}^k)$

matrice connue si $\bar{x}^k$ l'est — $\Delta\bar{x}$ inconnue — connu

Alg. Etant donné $\bar{x}^k, \bar{\lambda}^k, \bar{s}^k$

On résoud $DF(\bar{x}^k, \bar{\lambda}^k, \bar{s}^k) \begin{pmatrix} \Delta\bar{x} \\ \Delta\bar{\lambda} \\ \Delta\bar{s} \end{pmatrix} = \bar{F}(\bar{x}^k, \bar{\lambda}^k, \bar{s}^k)$

Puis on pose $\bar{x}^{k+1} = \bar{x}^k - \Delta\bar{x}$

$\bar{\lambda}^{k+1} = \bar{\lambda}^k - \Delta\bar{\lambda}$ $\quad \lambda_i^{k+1} = \max(0, \lambda_i^k - \Delta\lambda_i)$

$\bar{s}^{k+1} = \bar{s}^k - \Delta\bar{s}$ $\quad s_i^{k+1} = \max(0, s_i^k - \Delta s_i)$

$i = 1, \dots, m$

Reste à calculer $DF(\bar{x}, \bar{\lambda}, \bar{s})$

$\forall \bar{x} \in \mathbb{R}^n \ \forall \bar{\lambda} \in \mathbb{R}^m \ \forall \bar{s} \in \mathbb{R}^m$

$$\bar{F}(\bar{x}, \bar{\lambda}, \bar{s}) = \begin{pmatrix} A\bar{x} - \bar{b} - B^T\bar{\lambda} \\ B\bar{x} - \bar{c} - \bar{s} \\ \lambda_1 s_1 \\ \vdots \\ \lambda_m s_m \end{pmatrix}$$

$\frac{1}{2} \bar{x}^T A \bar{x}$ —gradient→ $A\bar{x}$

$A\bar{x}$ —dériver par rapport à $\bar{x}$→ A

$\frac{\partial}{\partial x_k} (A\bar{x})_i = \frac{\partial}{\partial x_k} \sum_{j=1}^{n} a_{ij} x_j = \sum_{j=1}^{n} a_{ij} \frac{\partial x_j}{\partial x_k}$ ($\frac{\partial x_j}{\partial x_k} = 1$ si $j = k$, 0 sinon)

$= a_{ik}$

$$DF(\bar{x}, \bar{\lambda}, \bar{s}) = \begin{pmatrix} A & -B^T & 0 \\ B & 0 & -I \\ 0 & s_1 \ (0) & \lambda_1 \ (0) \\ & (0) \ s_m & (0) \ \lambda_m \end{pmatrix} \begin{matrix} \updownarrow n \\ \updownarrow m \\ \updownarrow m \end{matrix}$$

$\leftrightarrow n \quad \leftrightarrow m \quad \leftrightarrow m$

le syst linéaire $DF(\bar{x}^k, \bar{\lambda}^k, \bar{s}^k) \begin{pmatrix} \Delta\bar{x} \\ \Delta\bar{\lambda} \\ \Delta\bar{s} \end{pmatrix} = \bar{F}(\bar{x}^k, \bar{\lambda}^k, \bar{s}^k)$ s'écrit :

$$\begin{pmatrix} A & -B^T & 0 \\ B & 0 & -I \\ 0 & \cancel{s_1^k} & \cancel{\lambda_1^k} \\ & & \lambda_m^k \ s_m^k \end{pmatrix} \begin{pmatrix} \Delta\bar{x} \\ \Delta\bar{\lambda} \\ \Delta\bar{s} \end{pmatrix} = \begin{pmatrix} A\bar{x}^k - \bar{b} - B^T\bar{\lambda}^k \\ B\bar{x}^k - \bar{c} - \bar{s}^k \\ \lambda_1^k s_1^k \\ \lambda_m^k s_m^k \end{pmatrix}$$

On peut montrer que si $s_i^k > 0$ et $\lambda_i^k > 0$ alors $DF(\bar{x}^k, \bar{\lambda}^k, \bar{s}^k)$ est inversible ($\ker = \bar{0}$)

Par conséquent on pose (ε > 0 paramètre petit)

$\bar{x}^{k+1} = \bar{x}^k - \Delta\bar{x}$

$\lambda_i^{k+1} = \max(\varepsilon, \lambda_i^k - \Delta\lambda_i)$

$s_i^{k+1} = \max(\varepsilon, s_i^k - \Delta s_i)$

$i = 1, \dots, m$

Ex 2. linear programming

Ex2. linear programming

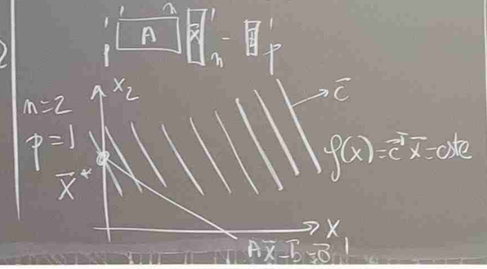
$\bar{x}^* \in \Omega$ tq $f(\bar{x}^*) \leq f(\bar{x})$ $\forall \bar{x} \in \Omega$

où $f(\bar{x}) = \bar{c}^T \bar{x}$ où $\bar{c} \in \mathbb{R}^n$

et $\Omega = \{ \bar{x} \in \mathbb{R}^n ;\ \bar{x} \geq \bar{0}$ et $A\bar{x} - \bar{b} = \bar{0} \}$

$m = n$

où $A \in \mathbb{R}^{p \times n}$ et $b \in \mathbb{R}^p$



$\forall \bar{x} \in \mathbb{R}^n$ $\forall \bar{\lambda} \in \mathbb{R}^n$ $\forall \bar{\mu} \in \mathbb{R}^p$

$\mathcal{L}(\bar{x}, \bar{\lambda}, \bar{\mu}) = \sum_{i=1}^{n} c_i x_i - \sum_{i=1}^{n} \lambda_i x_i - \sum_{i=1}^{p} \mu_i (A\bar{x} - \bar{b})_i$

$\partial_x \mathcal{L}(\bar{x}, \bar{\lambda}, \bar{\mu}) = \bar{c} - \bar{\lambda} - A^T \bar{\mu} \in \mathbb{R}^n$

les cond KKT s'écrivent :

trouver $\bar{x}^* \in \mathbb{R}^n$, $\bar{\lambda}^* \in \mathbb{R}^n$, $\bar{\mu}^* \in \mathbb{R}^p$ tq

$$\left\{ \begin{array}{ll} \bar{c} - \bar{\lambda}^* - A^T \bar{\mu}^* = \bar{0} & \updownarrow n \text{ eq} \\ A\bar{x}^* - \bar{b} = \bar{0} & \updownarrow p \text{ eq} \\ x_1^* \lambda_1^* = 0 & \\ \vdots & \updownarrow n \text{ eq} \\ x_n^* \lambda_n^* = 0 & \\ \bar{x}^* \geq \bar{0} & \\ \bar{\lambda}^* \geq \bar{0} & \end{array} \right\} \quad \bar{F}(\bar{x}^*, \bar{\lambda}^*, \bar{\mu}^*) = \bar{0}$$

$n + n + p$ inconnues $\bar{x}^*, \bar{\lambda}^*, \bar{\mu}^*$

Etant donné $\bar{x}^k, \bar{\lambda}^k, \bar{\mu}^k$, 1 pas de la méth. de Newton et on projette

$$DF(\bar{x}^k, \bar{\lambda}^k, \bar{\mu}^k) \begin{pmatrix} \Delta\bar{x} \\ \Delta\bar{\lambda} \\ \Delta\bar{\mu} \end{pmatrix} = F(\bar{x}^k, \bar{\lambda}^k, \bar{\mu}^k) \qquad (*)$$

et on pose

$$x_i^{k+1} = \max(0, x_i^k - \Delta x_i) \qquad i = 1, \dots, n$$
$$\lambda_i^{k+1} = \max(0, \lambda_i^k - \Delta\lambda_i)$$
$$\bar{\mu}^{k+1} = \bar{\mu}^k - \Delta\bar{\mu}$$

Reste à calculer $DF(\bar{x}, \bar{\lambda}, \bar{\mu})$

$\forall \bar{x} \in \mathbb{R}^n$, $\forall \bar{\lambda} \in \mathbb{R}^n$, $\forall \bar{\mu} \in \mathbb{R}^p$

$$\bar{F}(\bar{x}, \bar{\lambda}, \bar{\mu}) = \begin{pmatrix} \bar{c} - \bar{\lambda} - A^T \bar{\mu} \\ A\bar{x} - \bar{b} \\ x_1 \lambda_1 \\ \vdots \\ x_n \lambda_n \end{pmatrix}$$

$$DF(\bar{x}, \bar{\lambda}, \bar{\mu}) = \left(\begin{array}{c|c|c} 0 & -I & -A^T \\ \hline A & 0 & 0 \\ \hline \lambda_1 \ (0) \ \searrow \ x_1 & \cdots & 0 \end{array} \right)$$

le système linéaire $(*)$ s'écrit

$$\left(\begin{array}{c|c|c} 0 & -I & -A^T \\ \hline A & 0 & 0 \\ \hline x_i^k & \cdots & 0 \end{array} \right) \begin{pmatrix} \Delta\bar{x} \\ \Delta\bar{\lambda} \\ \Delta\bar{\mu} \end{pmatrix} = \begin{pmatrix} \bar{c} - \bar{\lambda}^k - A^T \bar{\mu}^k \\ A\bar{x}^k - \bar{b} \\ x_1^k \lambda_1^k \\ \vdots \\ x_n^k \lambda_n^k \end{pmatrix}$$

n | n | p

n | p | n

On peut montrer que si $x_i^k > 0$ et $\lambda_i^k > 0$ $i = 1, \dots, n$

alors $DF(\bar{x}^k, \bar{\lambda}^k, \bar{\mu}^k)$ est inversible. Par conséquent on pose

$$x_i^{k+1} = \max(\varepsilon, x_i^k - \Delta x_i) \qquad i = 1, \dots, n$$
$$\lambda_i^{k+1} = \max(\varepsilon, \lambda_i^k - \Delta\lambda_i)$$
$$\bar{\mu}^{k+1} = \bar{\mu}^k - \Delta\bar{\mu}$$

Contrôle optimal (Problème inverse)

Modèle $A\bar{x} = \bar{b}$

Mesures : observations $\bar{x}^0$

Compromis entre le modèle et mesures

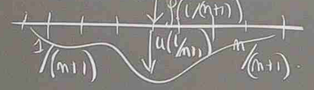
Comment modifier le modèle de sorte qu'il corresponde aux mesures ?

Modèle $A\bar{x} = \bar{b}$

Corde élastique (chap 10 Lime)

$q - u''(x) = f(x) \quad 0 < x < 1$

$u(0) = 0 \quad u(1) = 0$



$A = (n+1)^2 \begin{pmatrix} 2 & -1 & & \\ -1 & \ddots & \ddots & \\ & \ddots & \ddots & -1 \\ & & -1 & 2 \end{pmatrix}$

$\bar{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ x_i approx de $u(i/(n+1))$

$\bar{b} = \begin{pmatrix} f(1/(n+1)) \\ \vdots \\ f(n/(n+1)) \end{pmatrix}$

dans la suite A est inversible ($\det A \neq 0$, $\ker A = \bar{0}$, $\operatorname{Im} A = \mathbb{R}^n$)

Mesures : soit $\bar{x}^0$ les observations $\begin{pmatrix} x_{i_1}^0 \\ \vdots \\ x_{i_m}^0 \end{pmatrix}$

M^0 est la matrice des observations

$M^0 \in \mathbb{R}^{m \times n}$ m : nombre d'observations

$m = 1 \quad M^0 = (1_{i_1} 0 \cdots 0) \quad M^0 \bar{x}^0 = x_{i_1}^0$ (1 obs. pour la $i_1$ème comp de $\bar{x}^0$)

$m = 2 \quad M^0 = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \; 1 \end{pmatrix} \quad M^0 \bar{x}^0 = \begin{pmatrix} x_{i_1}^0 \\ x_{i_2}^0 \end{pmatrix}$

Dans la suite $\alpha > 0$ est un paramètre qui pondère modèle et mesures

3 problèmes équivalents (on peut passer de l'un à l'autre)

Pbm 1 (primal, physique) :

On cherche $\bar{x}^* \in \mathbb{R}^n$ tq $f(\bar{x}^*) \leq f(\bar{x}) \; \forall \bar{x} \in \mathbb{R}^n$

où $f(\bar{x}) = \frac{\alpha}{2} \|A\bar{x} - \bar{b}\|^2 + \frac{1}{2} \|M^0(\bar{x} - \bar{x}^0)\|^2$

On introduit une variable (de contrôle) $\bar{q} = A\bar{x} - \bar{b}$

(c'est la force qu'il faut ajouter au modèle de sorte à ce qu'il corresponde aux mesures)

Pbm 2 ($\bar{x}$ variable physique, primale ; $\bar{q}$ — de contrôle, duale)

On cherche $(\bar{x}^*, \bar{q}^*) \in \Omega$ tq $f(\bar{x}^*, \bar{q}^*) \leq f(\bar{x}, \bar{q}) \; \forall (\bar{x}, \bar{q}) \in \Omega$

où $\Omega = \{(\bar{x}, \bar{q}) \in \mathbb{R}^n \times \mathbb{R}^n$ tq $A\bar{x} - \bar{b} - \bar{q} = \bar{0}\}$

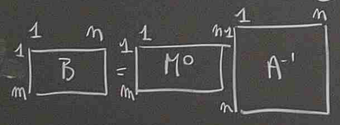
et $f(\bar{x}, \bar{q}) = \frac{\alpha}{2} \|\bar{q}\|^2 + \frac{1}{2} \|M^0(\bar{x} - \bar{x}^0)\|^2$

On élimine $\bar{x}$ dans le Pbm 2 : $\bar{x} = A^{-1}(\bar{b} + \bar{q})$

Pbm 3 : (problème dual, plus de $\bar{x}$, seulement $\bar{q}$)

On cherche $\bar{q}^* \in \mathbb{R}^n$ tq $f(\bar{q}^*) \leq f(\bar{q}) \; \forall \bar{q} \in \mathbb{R}^n$

où $f(\bar{q}) = \frac{\alpha}{2} \|\bar{q}\|^2 + \frac{1}{2} \|\underbrace{M^0}_{B}(\underbrace{A^{-1}}_{\bar{q}}(\bar{b} + \bar{q}) - \bar{x}^0)\|^2 - \underbrace{(M^0(\bar{x}^0 - A^{-1}\bar{b}))}_{\bar{c}}$



Solution du Pbm 1 :

$\bar{x}^*$ sol du Pbm 1 satisfait $\bar{\nabla} f(\bar{x}^*) = \bar{0}$

$\forall \bar{x} \in \mathbb{R}^n \quad \bar{\nabla} f(\bar{x}) = \alpha A^T(A\bar{x} - \bar{b}) + M^{0T} M^0 (\bar{x} - \bar{x}^0)$

$M^{0T} M^0 = \boxed{M^{0T}}_{n \times m} \boxed{M^0}_{m \times n} = \boxed{M^{0T} M^0}_{n \times n}$

par exemple si $m = 2$ $M^0 = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & \cdots & 0 & 1 \end{pmatrix}$

$M^{0T} M^0 = \begin{pmatrix} 1 & 0 & & (0) \\ & \ddots & & \\ (0) & & 0 & \\ & & & 1 \end{pmatrix}$

Donc $\bar{x}^*$ sol du pbm 1 est tq :

$\alpha A^T(A\bar{x}^* - \bar{b}) + M^{0T} M^0 (\bar{x}^* - \bar{x}^0) = \bar{0}$ (1)

(1) admet une unique solution.

En effet, montrons $\ker(\alpha A^T A + M^{0T} M^0)$

Soit $\bar{x} \in \mathbb{R}^n$ tq $(\alpha A^T A + M^{0T} M^0)\bar{x} = \bar{0}$

montrons $\bar{x} = \bar{0}$:

$\bar{x}^T(\alpha A^T A + M^{0T} M^0)\bar{x} = 0$

$\alpha \bar{x}^T A^T A \bar{x} + \bar{x}^T M^{0T} M^0 \bar{x} = 0$ (A inversible)

$\alpha \|A\bar{x}\|^2 + \|M^0 \bar{x}\|^2 = 0 \Rightarrow \|A\bar{x}\| = 0 \Rightarrow A\bar{x} = 0 \Rightarrow \bar{x} = 0$

(1) s'écrit : $\underbrace{(\alpha A^T A + M^{0T} M^0)}_{\text{connue}} \underbrace{\bar{x}^*}_{\text{inconnue}} = \underbrace{\alpha A^T \bar{b} + M^{0T} M^0 \bar{x}^0}_{\text{connu}}$

Solution du Pbm 2 :

pbm d'optim. avec contraintes d'égalité :

Lagrangien : $\mathcal{L}(\bar{x}, \bar{q}, \bar{\mu}) = f(\bar{x}, \bar{q}) - \sum_{i=1}^{m} \mu_i (A\bar{x} - \bar{b} - \bar{q})_i$

$\bar{x} \in \mathbb{R}^n$, $\bar{q} \in \mathbb{R}^m$, $\bar{\mu} \in \mathbb{R}^m$

$= \frac{\alpha}{2} \|\bar{q}\|^2 + \frac{1}{2} \|M^0(\bar{x} - \bar{x}^0)\|^2 - \sum_{i=1}^{m} \mu_i (A\bar{x} - \bar{b} - \bar{q})_i$

$\partial_{\bar{x}} \mathcal{L}(\bar{x}, \bar{q}, \bar{\mu}) = M^{0T} M^0 (\bar{x} - \bar{x}^0) - A^T \bar{\mu}$

$\partial_{\bar{q}} \mathcal{L}(\bar{x}, \bar{q}, \bar{\mu}) = \alpha \bar{q} + \bar{\mu}$

Conditions KKT : $\bar{x}^* \in \mathbb{R}^n, \bar{q}^* \in \mathbb{R}^m, \bar{\mu}^* \in \mathbb{R}^m$ tq

$$\begin{cases} M^{0T} M^0 (\bar{x}^* - \bar{x}^0) - A^T \bar{\mu}^* = \vec{0} \\ \alpha \bar{q}^* + \bar{\mu}^* = 0 \\ A\bar{x}^* - \bar{b} = \bar{q}^* \end{cases}$$

On élimine $\bar{\mu}^*$:

$$\begin{cases} M^{0T} M^0 (\bar{x}^* - \bar{x}^0) + \alpha A^T \bar{q}^* = \vec{0} \\ A\bar{x}^* - \bar{b} = \bar{q}^* \end{cases} \quad (2)$$

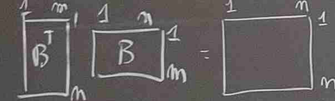
Si on pose $\bar{q}^* = A\bar{x}^* - \bar{b}$ dans (1) on obtient bien (2)

Solution du Pbm (3)

$\bar{q}^*$ sol du Pbm 3 satisfait $\nabla f(\bar{q}^*) = \bar{0}$

$\forall \bar{q} \in \mathbb{R}^m \quad f(\bar{q}) = \frac{\alpha}{2} \|\bar{q}\|^2 + \frac{1}{2} \|B\bar{q} - \bar{c}\|^2$

$\nabla f(\bar{q}) = \alpha \bar{q} + B^T (B\bar{q} - \bar{c})$

$B^T B =$ 

$\bar{q}^*$ sol du Pbm 3 satisfait $\alpha \bar{q}^* + B^T (B\bar{q}^* - \bar{c}) = \vec{0}$ (3)

(3) s'écrit :

$\alpha \bar{q}^* + A^{-T} M^{0T} M^0 (A^{-1}(\bar{b} + \bar{q}^*) - \bar{x}^0)$ (4)

$(\bar{x}^*, \bar{q}^*)$ sol de (2), on remplace $\bar{x}^*$ par $A^{-1}(\bar{b} + \bar{q}^*)$

on obtient $M^{0T} M^0 (A^{-1}(\bar{b} + \bar{q}^*) - \bar{x}^0) + \alpha A^T \bar{q}^* = \bar{0}$

on multiplie par $A^{-T} = (A^{-1})^T$ et on obtient (4)

Implémentation avec matlab :

Par exemple (2) :

$$\begin{pmatrix} A & -I \\ M^{0T} M^0 & \alpha A^T \end{pmatrix} \begin{pmatrix} \bar{x}^* \\ \bar{q}^* \end{pmatrix} = \begin{pmatrix} \bar{b} \\ M^{0T} M^0 \bar{x}^0 \end{pmatrix} \quad \text{ou} \quad \begin{pmatrix} M^{0T} M^0 & \alpha A^T \\ A & -I \end{pmatrix} \begin{pmatrix} \bar{x}^* \\ \bar{q}^* \end{pmatrix} = \begin{pmatrix} M^{0T} M^0 \bar{x}^0 \\ \bar{b} \end{pmatrix}$$

$(AB)^T = B^T A^T$

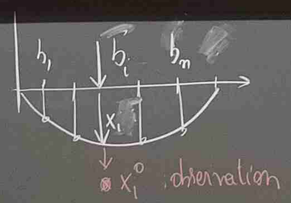
$(ABC)^T = (B^T C^T) A^T = B^T C^T A^T$

Pbm 3 : On cherche $\bar{q}^* \in \mathbb{R}^m$ tq $f(\bar{q}^*) \leq f(\bar{q}) \ \forall \bar{q} \in \mathbb{R}^m$

$f(\bar{q}) = \frac{\alpha}{2} \|\bar{q}\|^2 + \frac{1}{2} \|B\bar{q} - \bar{c}\|^2$

Pbm 4 : On cherche à minimiser $\|\bar{q}\|^2$ sous la contrainte $B\bar{q} = \bar{c}$

(les mesures sont exactement satisfaites)



Contrôle optimal (suite et fin)

Pbm 1 : $\min_{\vec{x}} \frac{\alpha}{2}\underbrace{\|A\vec{x}-\vec{b}\|^2}_{\text{modèle}} + \frac{1}{2}\underbrace{\|M^0(\vec{x}-\vec{x}^0)\|^2}_{\text{observations}}$

$M^0 = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 0 & & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{pmatrix}$ (1 ligne, n colonnes) $\quad M^{0T}M^0 = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 0 & & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{pmatrix}$ (n lignes, n colonnes)

$\alpha A^T(A\vec{x}^*-\vec{b}) + M^{0T}M^0(\vec{x}^*-\vec{x}^0) = \vec{0}$

$(\alpha A^TA + M^{0T}M^0)\vec{x}^* = \alpha A^T\vec{b} + M^{0T}M^0\vec{x}^0$

Pbm 2 : $\min_{\vec{x},\vec{q}} \frac{\alpha}{2}\|\vec{q}\|^2 + \frac{1}{2}\|M^0(\vec{x}-\vec{x}^0)\|^2$ sous la contrainte $A\vec{x}-\vec{b} = \vec{q}$

$\begin{cases} \alpha A^T\vec{q}^* + M^{0T}M^0(\vec{x}^*-\vec{x}^0) = \vec{0} \\ A\vec{x}^*-\vec{b} = \vec{q}^* \end{cases}$

Pbm 3 : (probl. dual)

$\vec{x} = A^{-1}(\vec{b}+\vec{q})$

$\min_{\vec{q}} \frac{\alpha}{2}\|\vec{q}\|^2 + \frac{1}{2}\|\underbrace{M^0(A^{-1}(\vec{b}+\vec{q})-\vec{x}^0)}_{B\vec{q}-\vec{c}}\|^2$

$\alpha\vec{q}^* + B^T(B\vec{q}^*-\vec{c}) = \vec{0}$

$\alpha\vec{q}^* + A^{-T}M^{0T}(M^0(A^{-1}(\vec{q}^*+\vec{b})-\vec{x}^0)) = \vec{0}$

$B = M^0A^{-1} \qquad \vec{c} = M^0(\vec{x}^0 - A^{-1}\vec{b})$

Pbm 4 : $\min_{\vec{q}} \frac{1}{2}\|\vec{q}\|^2$ sous la contrainte $B\vec{q}-\vec{c} = \vec{0}$ (les mesures sont satisfaites exactement)

On cherche donc $\vec{q}^* \in \Omega$ tq $f(\vec{q}^*) \le f(\vec{q})\ \forall \vec{q} \in \Omega$

où $\forall \vec{q} \in \mathbb{R}^n \quad f(\vec{q}) = \frac{1}{2}\|\vec{q}\|^2$

et $\Omega = \{\vec{q} \in \mathbb{R}^n \text{ tq } B\vec{q}-\vec{c} = \vec{0}\}$

$B = M^0A^{-1} =$ 

$\vec{c} \in \mathbb{R}^m \quad \vec{c} = M^0(\vec{x}^0 - A^{-1}\vec{b})$

On introduit le lagrangien $\mathcal{L}(\vec{q},\vec{\mu})$

$\forall \vec{q} \in \mathbb{R}^n\ \forall \vec{\mu} \in \mathbb{R}^m$

$\mathcal{L}(\vec{q},\vec{\mu}) = \frac{1}{2}\|\vec{q}\|^2 - \sum_{i=1}^{m} \mu_i (B\vec{q}-\vec{c})_i$

$\frac{\partial}{\partial \vec{q}}\mathcal{L}(\vec{q},\vec{\mu}) = \vec{q} - B^T\vec{\mu}$

les conditions KKT sont : trouver $\vec{q}^* \in \mathbb{R}^n$, $\vec{\mu}^* \in \mathbb{R}^m$ tq

$\begin{cases} \vec{q}^* - B^T\vec{\mu}^* = \vec{0} \\ B\vec{q}^* - \vec{c} = \vec{0} \end{cases}$

$\begin{pmatrix} I & -B^T \\ B & 0 \end{pmatrix}\begin{pmatrix} \vec{q}^* \\ \vec{\mu}^* \end{pmatrix} = \begin{pmatrix} \vec{0} \\ \vec{c} \end{pmatrix}$ (n, m ; n, m)

Matrice $\begin{pmatrix} I & -B^T \\ B & 0 \end{pmatrix}$ inversible ?

$\begin{pmatrix} I & -B^T \\ B & 0 \end{pmatrix}\begin{pmatrix} \vec{q} \\ \vec{\mu} \end{pmatrix} = \begin{pmatrix} \vec{0} \\ \vec{0} \end{pmatrix} \overset{?}{\Rightarrow} \vec{q} = \vec{0}$ et $\vec{\mu} = \vec{0}$

$\begin{cases} \vec{q} - B^T\vec{\mu} = \vec{0} \\ B\vec{q} = \vec{0} \end{cases}$ on élimine $\vec{q}$: $\vec{q} = B^T\vec{\mu}$, $BB^T\vec{\mu} = \vec{0}$

et donc $B^T\vec{\mu} = \vec{0} \Rightarrow \underbrace{A^{-T}M^{0T}}_{\text{inv. si } M^0 \text{ de rang } m}\vec{\mu} = \vec{0} \Rightarrow \vec{\mu} = \vec{0} \Rightarrow \vec{q} = \vec{0}$

panne si $M^0 = (1\ 0\ \cdots\ 0)$

$M^{0T} = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$ $\quad M^{0T}\vec{\mu} = \mu_1 = 0 \Rightarrow \mu_1 \in \mathbb{R}$

donc $\vec{\mu}^T BB^T\vec{\mu} = 0$, $\|B^T\vec{\mu}\|^2$

$m = 1$, $M^{0T} = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}_m$ $\quad \vec{\mu} \in \mathbb{R}^1$ $\quad \vec{\mu} = (\mu_1)$

Résumé moodle

Pbm 1 à 4

Pbm 5 $\min_{\vec{q}} \sum_{i=1}^{m} |q_i|$ sous la contrainte $B\vec{q}-\vec{c} = \vec{0}$

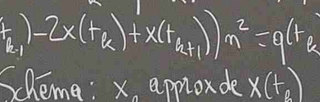
Pbm 7 : $x(t)$ position verticale d'un satellite de masse 1 soumis à une force $q(t)$

A certains temps t_i, on souhaite que $x(t_i)$ soit proche d'une objectif $x^0(t_i)$

Quelle est la force $q(t)$ correspondante ?

Eq. de Newton : $\begin{cases} x''(t) = q(t) & 0 \le t \le 1 \\ x(0) = 0 \\ x'(0) = 0 \end{cases}$

Méthode de différences finies



$t_k = \frac{k}{m}$ $\quad k = 0, 1, \ldots, m+1$

$x''(t_k) = q(t_k) \quad k = 0, \ldots, m-1$

$\left(x(t_{k-1}) - 2x(t_k) + x(t_{k+1})\right)m^2 = q(t_k) + O\left(\frac{1}{m^2}\right)$

Schéma : x_k approx de $x(t_k)$ $\quad k = 1, \ldots, m$

q_k $\quad q(t_k)$ $\quad k = 0, \ldots, m-1$

$m^2(x_{k-1} - 2x_k + x_{k+1}) = q_k \quad k = 0, \ldots, m-1$

$\begin{cases} m^2(x_{-1} - 2x_0 + x_1) = q_0 \\ m^2(x_0 - 2x_1 + x_2) = q_1 \\ \vdots \\ m^2(x_{m-2} - 2x_{m-1} + x_m) = q_{m-1} \end{cases}$

$x_{-1} \overset{?}{=} x_1$ $\quad x'(0) = 0$

$(x_1 - x_{-1})m = 0$

i.e. $x_{-1} = x_1$

$m^2(2x_1) = q_0$

Résumé moodle

Pbm 1 à 4

Pbm 5 $\min_{\vec{q}} \sum_{i=1}^{m} |q_i|$ sous la contrainte $B\vec{q} - \vec{c} = \vec{0}$

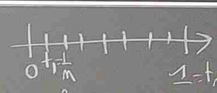
Pbm 7 : $x(t)$ position verticale d'un satellite de masse 1 soumis à une force $q(t)$

A certains temps t_i on souhaite que $x(t_i)$ soit proche d'une objectif $x^0(t_i)$

Quelle est la force $q(t)$ correspondante ?

Eq. de Newton : $\begin{cases} x''(t) = q(t) & 0 \le t \le 1 \\ x(0) = 0 \\ x'(0) = 0 \end{cases}$

Méthode de différences finies



$t_k = \frac{k}{m}$ $k = 0, 1, \dots, m+1$

$x''(t_k) = q(t_k)$ $k = 0, \dots, m-1$

$\left(x(t_{k-1}) - 2x(t_k) + x(t_{k+1})\right) m^2 = q(t_k) + O\left(\frac{1}{m^2}\right)$

Schéma : x_k approx de $x(t_k)$ $k = 1, \dots, m$

q_k $q(t_k)$ $k = 0, \dots, m-1$

$$m^2(x_{k-1} - 2x_k + x_{k+1}) = q_k \qquad k = 0, \dots, m-1$$

$$\begin{cases} m^2(x_{-1} - 2x_0 + x_1) = q_0 \\ m^2(x_0 - 2x_1 + x_2) = q_1 \\ \vdots \\ m^2(x_{m-2} - 2x_{m-1} + x_m) = q_{m-1} \end{cases}$$

x_{-1} ? $x'(0) = 0$

$(x_1 - x_{-1})m = 0$

i.e. $x_{-1} = x_1$

$m^2(2x_1) = q_0$

On écrit ces relations sous la forme $B\vec{x} = \vec{q}$

où $\vec{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_m \end{pmatrix}$ $\vec{q} = \begin{pmatrix} q_0 \\ \vdots \\ q_{m-1} \end{pmatrix}$ $B = m^2 \begin{pmatrix} -2 & 1 & & (0) \\ 1 & \ddots & \ddots & \\ & \ddots & \ddots & 1 \\ (0) & & 1 & -2 \end{pmatrix}$

On écrit les positions objectifs sous la forme $M^0 \vec{x}^0$

$M^0_{i,j} = 1$ si le i^e objectif est x_j^0

$i = 1, \dots, m$ nombre d'objectifs

Si à $t = t_j$ on souhaite $x_j = x_j^0$

alors $M^0 = (1\,0 \dots 0)$

Si $t = t_1$ objectif $x_1 = x_1^0$

$t = t_m$ —— $x_m = x_m^0$

$M^0 = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & \cdots & 0 & 1 \end{pmatrix}$ $\begin{matrix} 1 \\ 2 \end{matrix}$

$\begin{matrix} 1 & & m \end{matrix}$

On veut $\min_{\vec{x}, \vec{q}} \frac{\alpha}{2}\|\vec{q}\|^2 + \frac{1}{2}\|M^0(\vec{x} - \vec{x}^0)\|^2$ sous la contrainte $B\vec{x} = \vec{q}$

On cherche $(\vec{x}^*, \vec{q}^*) \in \Omega$ tq $f(\vec{x}^*, \vec{q}^*) \le f(\vec{x}, \vec{q})$ $\forall (\vec{x}, \vec{q}) \in \Omega$

où $\forall \vec{x} \in \mathbb{R}^n \ \forall \vec{q} \in \mathbb{R}^n$ $f(\vec{x}, \vec{q}) = \frac{\alpha}{2}\|\vec{q}\|^2 + \|M^0(\vec{x} - \vec{x}^0)\|^2$

et $\Omega = \{(\vec{x}, \vec{q}) \in \mathbb{R}^n \times \mathbb{R}^n \text{ tq } B\vec{x} - \vec{q} = \vec{0}\}$

$\forall \vec{x} \in \mathbb{R}^n \ \forall \vec{q} \in \mathbb{R}^n \ \forall \vec{\mu} \in \mathbb{R}^m$ $\mathcal{L}(\vec{x}, \vec{q}, \vec{\mu}) = \frac{\alpha}{2}\|\vec{q}\|^2 + \frac{1}{2}\|M^0(\vec{x} - \vec{x}^0)\|^2 - \sum_{i=1}^{m} \mu_i (B\vec{x} - \vec{q})_i$

$\partial_{\vec{x}} \mathcal{L}(\vec{x}, \vec{q}, \vec{\mu}) = M^{0T} M^0 (\vec{x} - \vec{x}^0) - B^T \vec{\mu}$

$\partial_{\vec{q}} \mathcal{L}(\vec{x}, \vec{q}, \vec{\mu}) = \alpha \vec{q} + \vec{\mu}$

Conditions KKT : on cherche $\vec{x}^*, \vec{q}^*, \vec{\mu}^* \in \mathbb{R}^n$ tq

$$\begin{cases} M^{0T} M^0 (\vec{x}^* - \vec{x}^0) - B^T \vec{\mu}^* = \vec{0} \\ \alpha \vec{q}^* + \vec{\mu}^* = \vec{0} \\ B\vec{x}^* = \vec{q}^* \end{cases}$$

On élimine $\vec{\mu}^* = -\alpha \vec{q}^*$

$$\begin{cases} \alpha B^T \vec{q}^* + M^{0T} M^0 (\vec{x}^* - \vec{x}^0) = \vec{0} \\ B\vec{x}^* = \vec{q}^* \end{cases}$$

On élimine $\vec{q}^*$;

$\alpha B^T B \vec{x}^* + M^{0T} M^0 (\vec{x}^* - \vec{x}^0) = \vec{0}$

(idem Pbm 1 avec $B = A$)

la matrice $\alpha B^T B + M^{0T} M^0$ est inversible

Car $\ker(\alpha B^T B + M^{0T} M^0) = \vec{0}$

En effet, soit $\vec{x} \in \mathbb{R}^n$ tq $(\alpha B^T B + M^{0T} M^0)\vec{x} = \vec{0}$

montrons $\vec{x} = \vec{0}$, on effectue le produit scalaire avec $\vec{x}$:

$\vec{x}^T (\alpha B^T B + M^{0T} M^0)\vec{x} = 0$

$\alpha \vec{x}^T B^T B \vec{x} + \vec{x}^T M^{0T} M^0 \vec{x} = 0$

Donc $\alpha \|B\vec{x}\|^2 + \|M^0 \vec{x}\|^2 = 0$

donc $\|B\vec{x}\| = 0$ $\|M^0\vec{x}\| = 0$

donc $B\vec{x} = \vec{0}$

donc $\vec{x} = \vec{0}$ (B inversible)

Ex : $m = 100$, $m = 7$, $\alpha = 10^{-7}$

Neural Networks

2012 →
algorithme pour interpoler
$f : \mathbb{R}^{d_{in}} \to \mathbb{R}^{d_{out}}$
avec d_{in} et d_{out} grands (10^6 !)

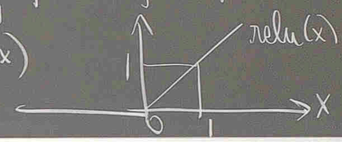
Ex :
- reconnaissance d'images
  image → chat, chien, arbre, ...
  image : 1000 × 1000 pixels $d_{in} = 10^6$
- deepl : traduction
  phrase en anglais → français
  phrase : mots
- chatgpt :
  question → réponse
  question, réponse : mots

aujourd'hui : 2 layers NN
mercredi : deep NN

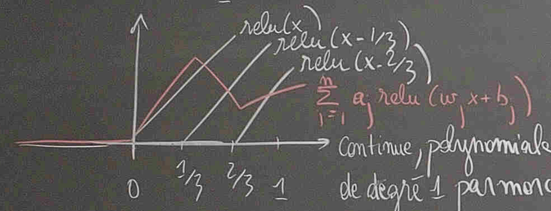
On se donne $f : [0,1] \to \mathbb{R}$
$x \to f(x)$ continue
qu'on veut approcher par

$$\sum_{j=1}^{m} a_j \, \mathrm{relu}(w_j x + b_j) \qquad a_j, w_j, b_j \in \mathbb{R}$$

$\mathrm{relu}(x) = \max(0, x)$



Ex : $n = 3$ $\quad a_1 \, \mathrm{relu}(x) + a_2 \, \mathrm{relu}(x - 1/3) + a_3 \, \mathrm{relu}(x - 2/3)$



Continue, polynomiale de degré 1 par morceaux

(cf éléments finis, chap 0 ligne interpolation, chap 1 ligne)

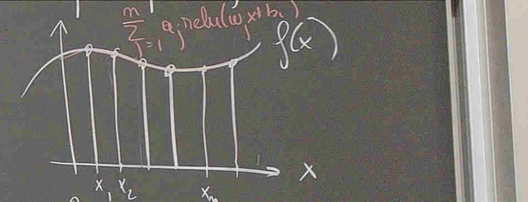
Question : (training)
Comment choisir les a_j, w_j, b_j de sorte que $\sum_{j=1}^{m} a_j \, \mathrm{relu}(w_j x + b_j)$ approche correctement $f(x)$?

On se donne des points d'interpolation x_i, $i = 1, \dots, n$ et on veut minimiser

$$\frac{1}{2} \sum_{i=1}^{n} \left(\sum_{j=1}^{m} a_j \, \mathrm{relu}(w_j x_i + b_j) - f(x_i) \right)^2$$

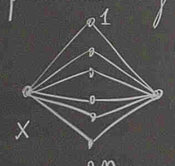
pour trouver des a_j, w_j, b_j $\quad j = 1, \dots, m$

le mieux qu'on puisse faire ($x \in \mathbb{R}$)



(on ne peut pas faire ce graphique si $x \in \mathbb{R}^{d_{in}}$ d_{in} grand)

Représentation graphique



$\sum_{j=1}^{n} a_j \, \mathrm{relu}(w_j x + b_j)$

NN 1 couche cachée (hidden layer)
2 couches

Écrivons ce pbm sous forme d'un pbm de min. sans contraintes

$$\vec{a} = \begin{pmatrix} a_1 \\ \vdots \\ a_m \end{pmatrix} \quad \vec{w} = \begin{pmatrix} w_1 \\ \vdots \\ w_n \end{pmatrix} \quad \vec{b} = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix}$$

On cherche $(\vec{a}^*, \vec{w}^*, \vec{b}^*) \in \mathbb{R}^m \times \mathbb{R}^n \times \mathbb{R}^m$ tq

$$F(\vec{a}^*, \vec{w}^*, \vec{b}^*) \leqslant F(\vec{a}, \vec{w}, \vec{b}) \quad \forall (\vec{a}, \vec{w}, \vec{b}) \in \mathbb{R}^m \times \mathbb{R}^n \times \mathbb{R}^m$$

où $F(\vec{a}, \vec{w}, \vec{b}) = \frac{1}{2} \sum_{i=1}^{n} \left(\sum_{j=1}^{m} a_j \, \mathrm{relu}(w_j x_i + b_j) - f(x_i) \right)^2$

$\forall \vec{a}, \vec{w}, \vec{b} \in \mathbb{R}^m$

La solution satisfait

$$\frac{\partial}{\partial \vec{a}} F(\vec{a}^*, \vec{w}^*, \vec{b}^*) = \vec{0}$$
$$\frac{\partial}{\partial \vec{w}} F(\quad) = \vec{0}$$
$$\frac{\partial}{\partial \vec{b}} F(\quad) = \vec{0}$$

où on a noté

$$\frac{\partial}{\partial \vec{a}} F(\vec{a}, \vec{w}, \vec{b}) = \begin{pmatrix} \frac{\partial}{\partial a_1} F(\vec{a}, \vec{w}, \vec{b}) \\ \vdots \\ \frac{\partial}{\partial a_m} F(\vec{a}, \vec{w}, \vec{b}) \end{pmatrix}$$

idem pour $\frac{\partial}{\partial \vec{w}}$ et $\frac{\partial}{\partial \vec{b}}$

$$\frac{\partial}{\partial a_k} F(\vec{a},\vec{w},\vec{b}) \quad k=1,\dots,n$$

$$= \sum_{i=1}^{m} \left(\sum_{j=1}^{n} a_j \, \underbrace{\text{relu}(w_j x_i + b_j)}_{A_{ij}} - \underbrace{f(x_i)}_{f_i} \right) \times \frac{\partial}{\partial a_k} \left(\sum_{j=1}^{n} a_j \, \text{relu}(w_j x_i + b_j) \right)$$

$$= \sum_{i=1}^{m} \left(\sum_{j=1}^{n} A_{ij} a_j - f_i \right) \underbrace{\text{relu}(w_k x_i + b_k)}_{A_{ik}}$$

$$= \sum_{i=1}^{m} (A\vec{a} - \vec{f})_i \, A_{ik}$$

$$= \left(A^T (A\vec{a} - \vec{f}) \right)_k \qquad \frac{\partial}{\partial \vec{a}} F(\vec{a},\vec{w},\vec{b}) = A^T(A\vec{a} - \vec{f})$$

(on a déjà $\min \frac{1}{2}\|A\vec{x} - \vec{b}\|^2 = f(\vec{x})$)

$\vec{\nabla} f(\vec{x}) = A^T(A\vec{x} - \vec{b})$

$\vec{\nabla} f(\vec{x}^*) = \vec{0} = A^T(A\vec{x}^* - \vec{b})$

ici $\vec{x} \longrightarrow \vec{a}, \vec{w}, \vec{b}$ ⚠ A dépend de $\vec{w}$ et $\vec{b}$

Si $\vec{w}$ et $\vec{b}$ sont donnés, moindres carrés

$$\frac{\partial}{\partial w_k} F(\vec{a},\vec{w},\vec{b}) \quad k=1,\dots,n$$

$$= \sum_{i=1}^{m} \left(\sum_{j=1}^{n} a_j \, \text{relu}(w_j x_i + b_j) - f(x_i) \right)$$

$$\frac{\partial}{\partial w_k} \left(\sum_{j=1}^{n} a_j \, \text{relu}(w_j x_i + b_j) \right)$$

$$= \sum_{i=1}^{m} (A\vec{a} - \vec{f})_i \, \underbrace{a_k \, \text{relu}'(w_k x_i + b_k) x_i}_{B_{ik}}$$

$$\frac{\partial}{\partial \vec{w}} F(\vec{a},\vec{w},\vec{b}) = B^T(A\vec{a} - \vec{f})$$

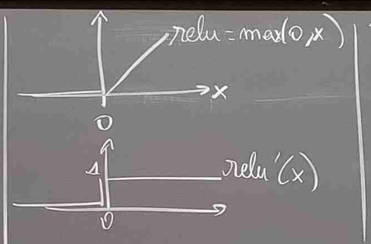
B dépend de $\vec{a}, \vec{w}$ et $\vec{b}$

$$\frac{\partial}{\partial b_k} F(\vec{a},\vec{w},\vec{b}) = \sum_{i=1}^{m} (A\vec{a} - \vec{f})_i \, \frac{\partial}{\partial b_k} \left(\sum_{j=1}^{n} a_j \, \text{relu}(w_j x_i + b_j) \right)$$

$$= \sum_{i=1}^{m} (A\vec{a} - \vec{f})_i \, \underbrace{a_k \, \text{relu}'(w_k x_i + b_k) 1}_{C_{ik}}$$

$$= \left(C^T (A\vec{a} - \vec{f}) \right)_k$$

et donc $\frac{\partial}{\partial \vec{b}} F(\vec{a},\vec{w},\vec{b}) = C^T(A\vec{a} - \vec{f})$



Donc $\vec{a}^*, \vec{w}^*, \vec{b}^*$ sont tels que

$A^T(A\vec{a}^* - \vec{f}) = \vec{0}$

$B^T(\text{——}) = \vec{0}$

$C^T(\text{——}) = \vec{0}$

Système non linéaire car A, B, C dépde $(\vec{a}^*, \vec{w}^*, \vec{b}^*)$

2 questions math :

- Est-ce que $\sum_{j=1}^{n} a_j \, \text{relu}(w_j x - b_j)$ approche bien $f(x)$? (expressivité) → pbm résolu (cf chap 1)
- Comment trouver $\vec{a}^*, \vec{w}^*, \vec{b}^*$?

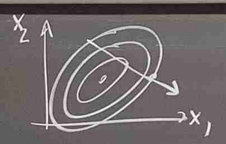
  méthode du gradient (convergence question ouverte)

Méthode du gradient : $(\eta > 0)$

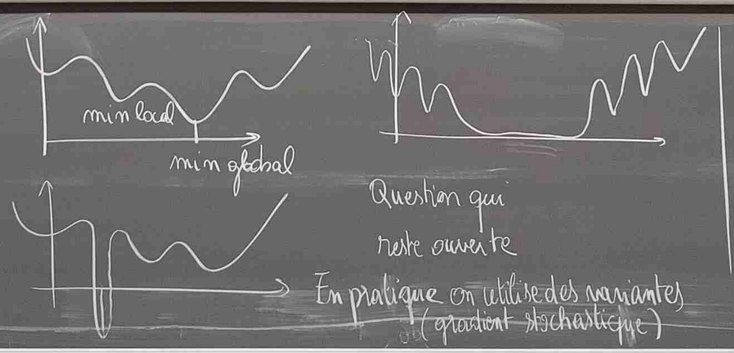
$\vec{a}^k, \vec{w}^k, \vec{b}^k$ sont donnés

"training"

$$\begin{cases} \vec{a}^{k+1} = \vec{a}^k - \eta \frac{\partial F}{\partial \vec{a}}(\vec{a}^k, \vec{w}^k, \vec{b}^k) \\ \vec{w}^{k+1} = \vec{w}^k - \eta \frac{\partial F}{\partial \vec{w}}(\text{——}) \\ \vec{b}^{k+1} = \vec{b}^k - \eta \frac{\partial F}{\partial \vec{b}}(\text{——}) \end{cases}$$



Pour quelles valeurs de η la méthode converge ?



Question qui reste ouverte

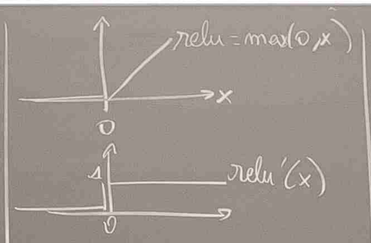
En pratique on utilise des variantes (gradient stochastique)

$$\frac{\partial}{\partial b_k} F(\bar a, \bar w, \bar b) = \sum_{i=1}^{m} (A\bar a - \bar f)_i \frac{\partial}{\partial b_k}\left(\sum_{j=1}^{n} a_j \,\mathrm{relu}(w_j x_i + b_j)\right)$$

$$= \sum_{i=1}^{m} (A\bar a - \bar f)_i \; a_k \,\mathrm{relu}'(w_k x_i + b_k) \cdot 1$$

$$= \left(C^T (A\bar a - \bar f)\right)_k \qquad C_{ik}$$

et donc $\frac{\partial}{\partial \bar b} F(\bar a, \bar w, \bar b) = C^T(A\bar a - \bar f)$



Donc $\bar a^*, \bar w^*, \bar b^*$ sont tels que

$A^T(A\bar a^* - \bar f) = \bar 0$

$B^T(\quad) = \bar 0$

$C^T(\quad) = \bar 0$

Système non linéaire car A, B, C dépde $(\bar a^*, \bar w^*, \bar b^*)$

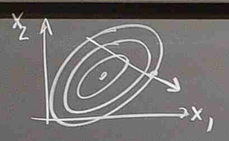
2 questions math :

- Est-ce que $\sum_{j=1}^{n} a_j \,\mathrm{relu}(w_j x - b_j)$ approche bien $f(x)$? (expressivité) → pbm résolu (cf chap 1)
- Comment trouver $\bar a^*, \bar w^*, \bar b^*$? méthode du gradient (convergence question ouverte)

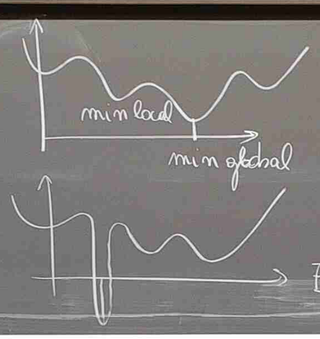
Méthode du gradient : $(\eta > 0)$

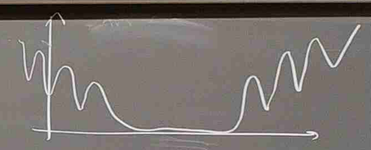
$\bar a^k, \bar w^k, \bar b^k$ sont donnés

"training" $\begin{cases} \bar a^{k+1} = \bar a^k - \eta \frac{\partial F}{\partial \bar a}(\bar a^k, \bar w^k, \bar b^k) \\ \bar w^{k+1} = \bar w^k - \eta \frac{\partial F}{\partial \bar w}(\quad) \\ \bar b^{k+1} = \bar b^k - \eta \frac{\partial F}{\partial \bar b}(\quad) \end{cases}$



Pour quelles valeurs de η la méthode converge ?

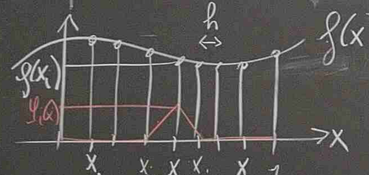




Question qui reste ouverte

En pratique on utilise des variantes (gradient stochastique)

Remarque : lien avec les éléments finis



$h = \max_{1 \le i \le n-1} |x_{i+1} - x_i|$

$n = m$

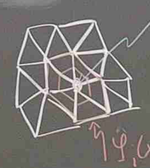
Interpolant de degré 1 par intervalle

$$\sum_{i=1}^{n} f(x_i)\,\varphi_i(x) = f_h(x)$$

$$f_h(x_j) = \sum_{i=1}^{n} f(x_i)\underbrace{\varphi_i(x_j)}_{\substack{1 \text{ si } i=j \\ 0 \text{ sinon}}} = f(x_j)$$

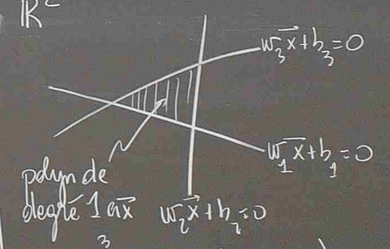
$f_h|_{[x_i, x_{i+1}]}$ est un polynôme de degré 1

On peut faire cette construction si $x \in \mathbb{R}^2$ ou $\mathbb{R}^3$



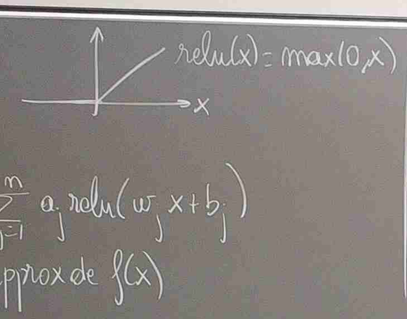
$\bar x \in \mathbb{R}^{10}$ on ne peut plus construire un maillage (1 sommet avec 10^6 voisins)

$\mathbb{R}^2$



$$\sum_{i=1}^{3} a_i \,\mathrm{relu}(w_i \bar x + b_i)$$

Neural Networks (2/2)



$f : [0,1] \to \mathbb{R}$, $x \to f(x)$

- NN 1 couche cachée 1960

$\sum_{j=1}^{m} a_j \, \mathrm{relu}(w_j x + b_j)$ approx de $f(x)$

x_i, $i = 1, \dots, m$ points d'interpolation

$$\min_{a_j, w_j, b_j} \underbrace{\frac{1}{2} \sum_{i=1}^{m} \left(\sum_{j=1}^{n} a_j \, \mathrm{relu}(w_j x_i + b_j) - f(x_i) \right)^2}_{F(\vec{a}, \vec{w}, \vec{b})}$$

$\vec{a}^*, \vec{w}^*, \vec{b}^* \in \mathbb{R}^n$ tq $F(\vec{a}^*, \vec{w}^*, \vec{b}^*) \leq F(\vec{a}, \vec{w}, \vec{b})$ $\forall \vec{a}, \vec{w}, \vec{b} \in \mathbb{R}^n$

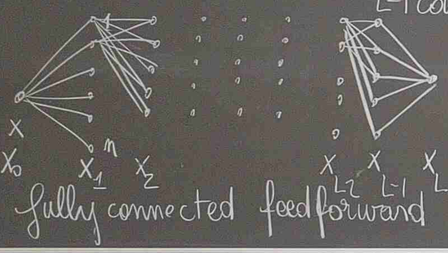
$$\begin{cases} \frac{\partial}{\partial \vec{a}} F(\vec{a}^*, \vec{w}^*, \vec{b}^*) = \vec{0} \\ \frac{\partial}{\partial \vec{w}} \quad \text{——} \quad = \vec{0} \\ \frac{\partial}{\partial \vec{b}} \quad \text{——} \quad = \vec{0} \end{cases} \quad \begin{cases} \vec{a}^{k+1} = \vec{a}^k - \eta \frac{\partial F}{\partial \vec{a}}(\vec{a}^k, \vec{w}^k, \vec{b}^k) \\ \vec{w}^{k+1} = \vec{w}^k - \eta \frac{\partial}{\partial \vec{w}} \quad \text{——} \\ \vec{b}^{k+1} = \vec{b}^k - \eta \frac{\partial}{\partial \vec{b}} \quad \text{——} \end{cases}$$

training

$\sum_{j=1}^{m} a_j^* \, \mathrm{relu}(w_j^* x + b_j^*)$ bonne approx de $f(x)$

- Deep NN 2012 (L couches à n neurones)

L−1 couches cachées



approx de $f(x)$

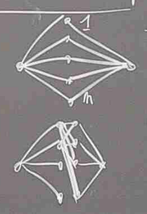
fully connected feedforward NN

$$\begin{aligned} x_1 &= \mathrm{relu}(W_1 x_0 + b_1) \\ x_2 &= \mathrm{relu}(W_2 x_1 + b_2) \\ &\vdots \\ x_{L-1} &= \mathrm{relu}(W_{L-1} x_{L-2} + b_{L-1}) \\ x_L &= \mathrm{relu}(W_L x_{L-1} + b_L) \end{aligned}$$

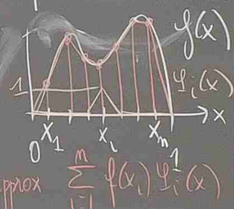
$(x_1)_j = \mathrm{relu}((w_1)_j x_0 + (b_1)_j)$

$L = 2$: 1 couche cachée

$L = 3$: 2 couches cachées

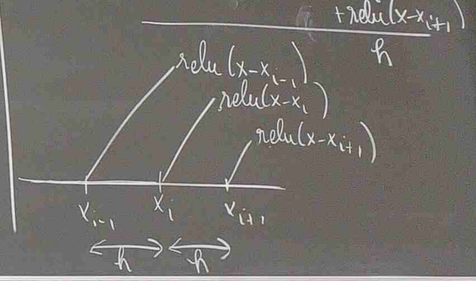


Remarque : les éléments finis sont un cas particulier $L = 3$



approx $\sum_{i=1}^{m} f(x_i) \varphi_i^1(x)$

$\varphi_i(x) = \mathrm{relu}(x - x_{i-1}) - 2\,\mathrm{relu}(x - x_i) + \mathrm{relu}(x - x_{i+1})$ / h



Comment trouver les $W_1, b_1, \dots, W_L, b_L$ paramètres du NN ?

On note $x_L = f_{deep}(x, \theta)$ où $\theta = (W_1, b_1, \dots, W_L, b_L)$ weights, poids

On se donne des points d'interpolation (data) $\tilde{x}_i$, $i = 1, \dots, m$

et on veut $\min_{\theta} \underbrace{\frac{1}{2} \sum_{i=1}^{m} \left(f_{deep}(\tilde{x}_i, \theta) - f(\tilde{x}_i) \right)^2}_{F(\theta)}$

i.e θ^* tq $F(\theta^*) \leq F(\theta)$ $\forall \theta$

finalement $f_{deep}(x, \theta^*)$ approx de $f(x)$

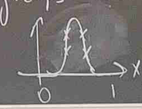
gradient $\theta^{k+1} = \theta^k - \underbrace{\eta}_{>0} \nabla F(\theta^k)$

pour calculer $\nabla F(\theta)$: back propagation

Chain rule,

Exp 1 : $f : [0,1] \to \mathbb{R}$

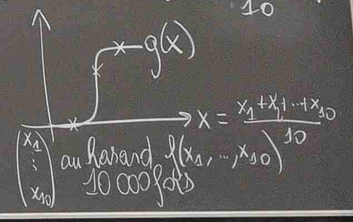
$n = 100$, $m = 20$, $\eta = 1$, 10 000 iter



$L = 2$ (1 couche cachée, n, m, η, nombre d'iter.)

Exp 2 : $f : [0,1]^{10} \to \mathbb{R}$

$f(x_1, x_2, \dots, x_{10}) = g\left(\frac{x_1 + x_2 + \dots + x_{10}}{10}\right)$



Chap 14 : Diffusion-convection

Chap 12 : eq. chaleur (diffusion)

Chap 13 : eq. transport / convection

Expérience : diffusion-convection polluant le long d'une rivière

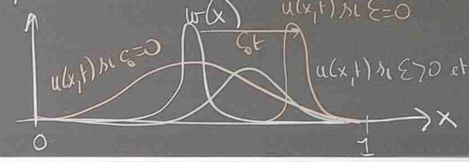
Etant donné $w(x)$ $0 \le x \le 1$ (conc. init.)

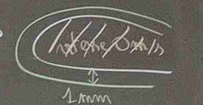
c_0 vitesse transport

$\varepsilon > 0$ coeff diffusion

on cherche $u(x,t)$ telle que

$$\begin{cases} \frac{\partial u}{\partial t}(x,t) - \varepsilon \frac{\partial^2 u}{\partial x^2}(x,t) + c_0 \frac{\partial u}{\partial x}(x,t) = 0 & 0<x<1, t>0 \\ u(0,t) = 0 \quad u(1,t) = 0 & t>0 \\ u(x,0) = w(x) & 0<x<1 \end{cases}$$

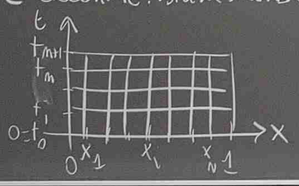




2 schémas

- explicite centré : stable sans conditions
- implicite décentré : stable sans conditions

notations :



$h = \frac{1}{N+1}$ $x_i = ih$ $i = 0, 1, \dots, N+1$

τ pas de temps $t_n = n\tau$ $n = 0, 1, 2, \dots$

But : étant donné u_i^n $i = 1, \dots, N$,

comment calculer u_i^{n+1} $i = 1, \dots, N$?

Cond. initiale $u_i^0 = w(x_i)$ $i = 1, \dots, N$

- explicite centré : on écrit l'eq. x_i, t_n
  - on utilise une formule diff. finie progressive pour $\frac{\partial u}{\partial t}(x_i, t_n)$
  - ——— centrée — $\frac{\partial^2 u}{\partial x^2}(x_i, t_n)$
  - ——— — $\frac{\partial u}{\partial x}(x_i, t_n)$

schéma :

$$\frac{u_i^{n+1} - u_i^n}{\tau} - \varepsilon \frac{u_{i-1}^n - 2u_i^n + u_{i+1}^n}{h^2} + c_0 \frac{u_{i+1}^n - u_{i-1}^n}{2h} = 0$$

$i = 1, \dots, N$

$u_0^n = 0$ $u_{N+1}^n = 0$

Ce schéma est explicite : $u_i^{n+1} = \dots$

——— stable sous conditions

——— converge $O(\tau) + O(h^2)$

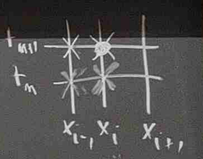
(l'erreur au temps final, sous les cond. de stabilité, est divisée par 4 si h est divisé par 2, et τ par 4.)

- implicite décentré ($c_0 > 0$) :
  on écrit l'eq x_i, t_{n+1}
  - on utilise une F DF rétrograde pour $\frac{\partial u}{\partial t}(x_i, t_{n+1})$
  - ——— centrée — $\frac{\partial^2 u}{\partial x^2}(x_i, t_{n+1})$
  - ——— décentrée — $\frac{\partial u}{\partial x}(x_i, t_{n+1})$

schéma :

$$\begin{cases} \frac{u_i^{n+1} - u_i^n}{\tau} - \varepsilon \frac{u_{i-1}^{n+1} - 2u_i^{n+1} + u_{i+1}^{n+1}}{h^2} + c_0 \frac{u_i^{n+1} - u_{i-1}^{n+1}}{h} = 0 \\ i = 1, \dots, N \\ u_0^{n+1} = 0 \quad u_{N+1}^{n+1} = 0 \end{cases}$$

(c_0)



Ce schéma est implicite

$$\begin{pmatrix} 1+\frac{2\tau\varepsilon}{h^2}+\frac{\tau c_0}{h} & -\frac{\tau\varepsilon}{h^2} & & \\ -\frac{\tau\varepsilon}{h^2}-\frac{c_0\tau}{h} & \ddots & \ddots & \\ & \ddots & \ddots & -\frac{\tau\varepsilon}{h^2} \\ & & -\frac{\tau\varepsilon}{h^2}-\frac{c_0\tau}{h} & \end{pmatrix} \begin{pmatrix} u_1^{n+1} \\ \vdots \\ u_N^{n+1} \end{pmatrix} = \begin{pmatrix} u_1^n \\ \vdots \\ u_N^n \end{pmatrix}$$

stable sans conditions

converge en $O(\tau)+O(h)$

Stabilité des schémas?

Rappel : la solution u satisfait les 2 propriétés suivantes

- Si $u(x) \geq 0$ $0<x<1$ alors $u(x,t) \geq 0$ $0<x<1$ (positivité)
- principe du max : $\max_{0<x<1} |u(x,t)| \leq \max_{0\leq x\leq 1} |u(x)|$

Déf. on dira que le schéma est stable si les 2 propriétés suivantes sont satisfaites

1. Si $u_i^n \geq 0$ $i=1,\dots,N$ alors $u_i^{n+1} \geq 0$ ——
2. $\max_{1\leq i\leq N} |u_i^{n+1}| \leq \max_{1\leq i\leq N} |u_i^n|$

1

Lemme :

- le schéma explicite centré est stable si $\tau \leq \frac{h^2}{2\varepsilon}$ et $h \leq \frac{2\varepsilon}{c_0}$
- le schéma implicite centré est stable $\forall h>0$ $\forall \tau>0$

Attention : $h, \tau \to 0$ pour obtenir la convergence

Dem :

- explicite centré : supposons $\tau \leq \frac{h^2}{2\varepsilon}$ et $h \leq \frac{2\varepsilon}{c_0}$ et montrons que les cond. 1 et 2 sont satisfaites

$$u_i^{n+1} = \left(1-\frac{2\tau\varepsilon}{h^2}\right) u_i^n + \left(\frac{\tau\varepsilon}{h^2}+\frac{\tau c_0}{2h}\right) u_{i-1}^n + \left(\frac{\tau\varepsilon}{h^2}-\frac{\tau c_0}{2h}\right) u_{i+1}^n \quad i=1,\dots,N$$

cond 1 : Supposons $u_i^n \geq 0$ $i=1,\dots,N$ et montrons $u_i^{n+1} \geq 0$ ——

Si $\begin{cases} 1-\frac{2\tau\varepsilon}{h^2} \geq 0 \\ \frac{\tau\varepsilon}{h^2}+\frac{\tau c_0}{2h} \geq 0 \\ \frac{\tau\varepsilon}{h^2}-\frac{\tau c_0}{2h} \geq 0 \end{cases} \iff \begin{cases} \tau \leq \frac{h^2}{2\varepsilon} \\ \text{toujours car } c_0>0 \\ h \leq \frac{2\varepsilon}{c_0} \end{cases}$

alors $u_i^{n+1} \geq 0$ $i=1,\dots,N$

cond 2 :

$$i=1,\dots,N : \quad |u_i^{n+1}| = \left| \left(1-\frac{2\tau\varepsilon}{h^2}\right) u_i^n + \left(\frac{\tau\varepsilon}{h^2}+\frac{\tau c_0}{2h}\right) u_{i-1}^n + \left(\frac{\tau\varepsilon}{h^2}-\frac{\tau c_0}{2h}\right) u_{i+1}^n \right|$$

$$\leq \left|1-\frac{2\tau\varepsilon}{h^2}\right| \cdot |u_i^n| + \left|\frac{\tau\varepsilon}{h^2}+\frac{\tau c_0}{2h}\right| \cdot |u_{i-1}^n| + \left|\frac{\tau\varepsilon}{h^2}-\frac{\tau c_0}{2h}\right| |u_{i+1}^n| \quad |a|=a \text{ si } a\geq 0$$

$$\leq \left(1-\frac{2\tau\varepsilon}{h^2}\right)|u_i^n| + \left(\frac{\tau\varepsilon}{h^2}+\frac{\tau c_0}{2h}\right)|u_{i-1}^n| + \left(\frac{\tau\varepsilon}{h^2}-\frac{\tau c_0}{2h}\right)|u_{i+1}^n|$$

$$|u_i^n|, |u_{i-1}^n|, |u_{i+1}^n| \leq \max_{1\leq j\leq N} |u_j^n|$$

(mais si $i=1$ et $i=N$ car $u_0^n = u_{N+1}^n = 0 \leq \max_{1\leq j\leq N} |u_j^n|$)

Donc $i=1,\dots,N$:

$$|u_i^{n+1}| \leq \underbrace{\left(1-\frac{2\tau\varepsilon}{h^2}+\frac{2\tau\varepsilon}{h^2}+\frac{\tau c_0}{2h}-\frac{\tau c_0}{2h}\right)}_{1} \max_{1\leq j\leq N} |u_j^n|$$

et donc $\max_{1\leq i\leq N} |u_i^{n+1}| \leq \max_{1\leq j\leq N} |u_j^n|$

- implicit décentré : montrons cond. 1

Soit k l'indice tel que $u_k^{n+1} \leq u_j^{n+1}$ $j=1,\dots,N$

(i.e. $u_k^{n+1} = \min_{1\leq j\leq N} u_j^{n+1}$)

Supposons $u_i^n \geq 0$ $i=1,\dots,N$ et montrons $u_k^{n+1} \geq 0$ $\left(\text{car } 0 \leq u_k^{n+1} \leq u_j^{n+1} \text{ } j=1,\dots,N\right)$ (ce qui implique)

2

Lemme :

- le schéma explicite centré est stable si $\tau \leq \frac{h^2}{2\varepsilon}$ et $h \leq \frac{2\varepsilon}{c_0}$
- le schéma implicite centré est stable $\forall h > 0 \ \forall \tau > 0$

Attention : $h, \tau \to 0$ pour observer la convergence

Dem :

- explicite centré : supposons $\tau \leq \frac{h^2}{2\varepsilon}$ et $h \leq \frac{2\varepsilon}{c_0}$ et montrons que les cond. 1 et 2 sont satisfaites

$$u_i^{n+1} = \left(1 - \frac{2\tau\varepsilon}{h^2}\right) u_i^n + \left(\frac{\tau\varepsilon}{h^2} + \frac{\tau c_0}{2h}\right) u_{i-1}^n + \left(\frac{\tau\varepsilon}{h^2} - \frac{\tau c_0}{2h}\right) u_{i+1}^n, \quad i = 1, \dots, N$$

cond 1 : Supposons $u_i^n \geq 0$ $i = 1, \dots, N$ et montrons $u_i^{n+1} \geq 0$ ——

Si $\begin{cases} 1 - \frac{2\tau\varepsilon}{h^2} \geq 0 \\ \frac{\tau\varepsilon}{h^2} + \frac{\tau c_0}{2h} \geq 0 \\ \frac{\tau\varepsilon}{h^2} - \frac{\tau c_0}{2h} \geq 0 \end{cases} \iff \begin{cases} \tau \leq \frac{h^2}{2\varepsilon} \\ \text{toujours car } c_0 > 0 \\ h \leq \frac{2\varepsilon}{c_0} \end{cases}$

alors $u_i^{n+1} \geq 0 \quad i = 1, \dots, N$

cond 2 :

$$i = 1, \dots, N : \quad |u_i^{n+1}| = \left| \left(1 - \frac{2\tau\varepsilon}{h^2}\right) u_i^n + \left(\frac{\tau\varepsilon}{h^2} + \frac{\tau c_0}{2h}\right) u_{i-1}^n + \left(\frac{\tau\varepsilon}{h^2} - \frac{\tau c_0}{2h}\right) u_{i+1}^n \right|$$

$$\leq \left|1 - \frac{2\tau\varepsilon}{h^2}\right| \cdot |u_i^n| + \left|\frac{\tau\varepsilon}{h^2} + \frac{\tau c_0}{2h}\right| \cdot |u_{i-1}^n| + \left|\frac{\tau\varepsilon}{h^2} - \frac{\tau c_0}{2h}\right| |u_{i+1}^n|$$

$|a| = a$ si $a \geq 0$

$$\leq \left(1 - \frac{2\tau\varepsilon}{h^2}\right) |u_i^n| + \left(\frac{\tau\varepsilon}{h^2} + \frac{\tau c_0}{2h}\right) |u_{i-1}^n| + \left(\frac{\tau\varepsilon}{h^2} - \frac{\tau c_0}{2h}\right) |u_{i+1}^n|$$

$|u_i^n|, |u_{i-1}^n|, |u_{i+1}^n| \leq \max_{1 \leq j \leq N} |u_j^n|$

(mais si $i = 1$ et $i = N$ car $0 = |u_0^n| \leq \max_{1 \leq j \leq N} |u_j^n|$)

Donc $i = 1, \dots, N$:

$$|u_i^{n+1}| \leq \underbrace{\left(1 - \frac{2\tau\varepsilon}{h^2} + \frac{2\tau\varepsilon}{h^2} + \frac{\tau c_0}{2h} - \frac{\tau c_0}{2h}\right)}_{1} \max_{1 \leq j \leq N} |u_j^n|$$

et donc $\max_{1 \leq i \leq N} |u_i^{n+1}| \leq \max_{1 \leq j \leq N} |u_j^n|$

- implicit décentré : montrons cond. 1

Soit k l'indice tel que $u_k^{n+1} \leq u_j^{n+1} \quad j = 1, \dots, N$

$\left(\text{i.e. } u_k^{n+1} = \min_{1 \leq j \leq N} u_j^{n+1}\right)$

Supposons $u_i^n \geq 0$ $i = 1, \dots, N$ et montrons $u_k^{n+1} \geq 0$ $\left(\text{ce qui implique } 0 \leq u_k^{n+1} \leq u_j^{n+1}, \ j = 1, \dots, N\right)$

Sur la k^e ligne le schéma s'écrit :

$$\left(1 + \frac{2\tau\varepsilon}{h^2} + \frac{c_0\tau}{h}\right) u_k^{n+1} = \left(\underbrace{\frac{\tau\varepsilon}{h^2} + \frac{c_0\tau}{h}}_{\geq 0}\right) u_{k-1}^{n+1} + \underbrace{\frac{\tau\varepsilon}{h^2}}_{\geq 0} u_{k+1}^{n+1} + u_k^n \quad \text{si } k+1 \neq N$$

$$\geq \left(\frac{\tau\varepsilon}{h^2} + \frac{c_0\tau}{h}\right) u_{k-1}^{n+1} + \frac{\tau\varepsilon}{h^2} u_{k+1}^{n+1} + u_k^n$$

$u_k^{n+1} \leq u_j^{n+1}, \ j = 1, \dots, N$

et donc $u_k^{n+1} \geq u_k^n \geq 0$

Si $k = 1$: $\left(1 + \frac{2\tau\varepsilon}{h^2} + \frac{c_0\tau}{h}\right) u_1^{n+1} = \frac{\tau\varepsilon}{h^2} u_2^{n+1} + u_1^n$

$$\geq \frac{\tau\varepsilon}{h^2} u_1^{n+1} + u_1^n$$

$$\underbrace{\left(1 + \frac{\tau\varepsilon}{h^2} + \frac{c_0\tau}{h}\right)}_{> 0} u_1^{n+1} \geq u_1^n \geq 0$$

et donc $u_1^{n+1} \geq 0$

Idem si $k = N$

exam2024.pdf ... Displays

Course: Analyse | exam2024.pdf

https://moodle.epfl.ch/pluginfile.php/3444390/mod_resource/content/1/exam2024.pdf

5 of 16 — 160%

+1/5/56+

Soit $g : [-1, 1] \to \mathbb{R}$ une fonction continue, on approche $\int_{-1}^{1} g(t)\mathrm{d}t$ à l'aide de la formule de quadrature:

$$J(g) = g(-\alpha) + g(\alpha), \text{ où } 0 < \alpha < 1. \tag{5}$$

Question 8 On a :

- ☐ Pour $\alpha = \frac{1}{\sqrt{3}}$, on a $\int_{-1}^{1} p(t)\mathrm{d}t = J(p) \quad \forall\, p \in \mathbb{P}_4$.
- ☐ Pour $\alpha = \frac{1}{\sqrt{3}}$, on a $\int_{-1}^{1} p(t)\mathrm{d}t = J(p) \quad \forall\, p \in \mathbb{P}_3$.
- ☐ Pour $\alpha = \frac{1}{\sqrt{3}}$, on a $\int_{-1}^{1} p(t)\mathrm{d}t = J(p) \quad \forall\, p \in \mathbb{P}_5$.
- ☐ Pour tout $0 < \alpha < 1$ on a $\int_{-1}^{1} p(t)\mathrm{d}t = J(p) \quad \forall\, p \in \mathbb{P}_1$.
- ☐ Pour $\alpha = \frac{1}{\sqrt{3}}$, on a $\int_{-1}^{1} p(t)\mathrm{d}t = J(p) \quad \forall\, p \in \mathbb{P}_2$.

Soit $f : [0, 1] \to \mathbb{R}$ une fonction continue, N un entier positif, $h = 1/N$, $x_i = ih$, $i = 0, 1, \ldots, N$. On a

$$\int_0^1 f(x)\mathrm{d}x = \frac{h}{2} \sum_{i=0}^{N-1} \int_{-1}^{1} f(x_i + h\frac{t+1}{2})dt. \tag{6}$$

Question 9 Vrai ou faux ?

☐ VRAI ☐ FAUX

On approche les intégrales de -1 à 1 dans (7) [illegible] de quadrature (6) avec $\alpha = \frac{1}{\sqrt{3}}$. On obtient ainsi l'approximation $L_h(f)$.

$\vec{\nabla} f(\vec{x}^*) = \sum_{i=1}^{m} \lambda_i^* \vec{\nabla} g_i(\vec{x}^*) = \sum_{i=1}^{n} \lambda_i^* \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} = \vec{\lambda}^*$

$m = n$

$\forall \vec{x} \in \mathbb{R}^n \quad g_i(\vec{x}) = x_i + c \geq 0 \quad i = 1, \ldots, n$

$\nabla g_i(\vec{x}) = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} i \qquad \vec{\nabla} f(\vec{x}) = A\vec{x} - \vec{b} + \begin{pmatrix} (x_1)^3 \\ \vdots \\ (x_n)^3 \end{pmatrix}$

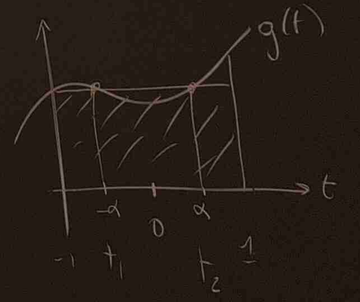
(1): $A\vec{x}^* - \vec{b} + \begin{pmatrix} (x_1^*)^3 \\ \vdots \\ (x_n^*)^3 \end{pmatrix} - \vec{\lambda}^* = \vec{0}$

(4): $\lambda_i^* (\underbrace{x_i^* + c}_{\geq 0}) = 0 \quad i = 1, \ldots, n$

$\vec{F}(\vec{x}, \vec{\lambda}, \vec{s}) = \begin{pmatrix} A\vec{x} - \vec{b} + \begin{pmatrix} (x_1)^3 \\ \vdots \\ (x_n)^3 \end{pmatrix} - \vec{\lambda} \\ \vec{x} + c\vec{1} - \vec{s} \\ \lambda_1 s_1 \\ \vdots \\ \lambda_n s_n \end{pmatrix}$

$DF(\vec{x}, \vec{\lambda}, \vec{s}) = \begin{pmatrix} A + 3\begin{pmatrix} (x_1)^2 & & (0) \\ & \ddots & \\ (0) & & (x_n)^2 \end{pmatrix} & -I & 0 \\ I & 0 & -I \\ 0 & \begin{pmatrix} s_1 & & (0) \\ & \ddots & \\ (0) & & s_n \end{pmatrix} & \begin{pmatrix} \lambda_1 & & (0) \\ & \ddots & \\ (0) & & \lambda_n \end{pmatrix} \end{pmatrix}$

$\int_{-1}^{1} g(t)dt \sim J(g) = g(-\alpha) + g(\alpha) = w_1 g(t_1) + w_2 g(t_2)$



exam2024.pdf ... Displays

Course: Analyse | exam2024.pdf

https://moodle.epfl.ch/pluginfile.php/3444390/mod_resource/content/1/exam2024.pdf

5 of 16 160%

+1/5/56+

Soit $g : [-1, 1] \to \mathbb{R}$ une fonction continue, on approche $\int_{-1}^{1} g(t)\mathrm{d}t$ à l'aide de la formule de quadrature:

$$J(g) = g(-\alpha) + g(\alpha), \text{ où } 0 < \alpha < 1. \tag{5}$$

Question 8 On a :

- ☐ Pour $\alpha = \frac{1}{\sqrt{3}}$, on a $\int_{-1}^{1} p(t)\mathrm{d}t = J(p) \quad \forall\, p \in \mathbb{P}_4$.
- ☐ Pour $\alpha = \frac{1}{\sqrt{3}}$, on a $\int_{-1}^{1} p(t)\mathrm{d}t = J(p) \quad \forall\, p \in \mathbb{P}_3$.
- ☐ Pour $\alpha = \frac{1}{\sqrt{3}}$, on a $\int_{-1}^{1} p(t)\mathrm{d}t = J(p) \quad \forall\, p \in \mathbb{P}_5$.
- ☐ Pour tout $0 < \alpha < 1$ on a $\int_{-1}^{1} p(t)\mathrm{d}t = J(p) \quad \forall\, p \in \mathbb{P}_1$.
- ☐ Pour $\alpha = \frac{1}{\sqrt{3}}$, on a $\int_{-1}^{1} p(t)\mathrm{d}t = J(p) \quad \forall\, p \in \mathbb{P}_2$.

Soit $f : [0, 1] \to \mathbb{R}$ une fonction continue, N un entier positif, $h = 1/N$, $x_i = ih$, $i = 0, 1, \ldots, N$. On a

$$\int_0^1 f(x)\mathrm{d}x = h \sum_{i=0}^{N-1} \int_{-1}^{1} f(x_i + h\frac{t+1}{2})dt. \tag{6}$$

Question 9 Vrai ou faux ?

FAUX

On approche les intégrales de -1 à 1 [illegible] la formule de quadrature (6) avec $\alpha = \frac{1}{\sqrt{3}}$; On obtient ainsi l'approximation $L_h(f)$

---

la formule de quadrature est exacte pour les polyn de degré ≤ 1 ssi

$w_i = \int_{-1}^{1} \ell_i(t)\mathrm{d}t \quad i = 1, 2$

$\int_{-1}^{1} p(t)\mathrm{d}t = J(p) \quad \forall\, p \in \mathbb{P}_1$

$w_1 = \int_{-1}^{1} \frac{t-\alpha}{-2\alpha}\mathrm{d}t = \int_{-1}^{1} \frac{\alpha}{2\alpha}\mathrm{d}t = \frac{2\alpha}{2\alpha} = 1$

$p(t) = t^2 \quad \int_{-1}^{1} p(t)\mathrm{d}t \overset{?}{=} J(p)$

Il suffit de vérifier avec $p(t) = t^2$

$\int_{-1}^{1} t^2\mathrm{d}t = \left[\frac{t^3}{3}\right]_{-1}^{1} = \frac{2}{3}$

$J(p) = (-\alpha)^2 + \alpha^2 = 2\alpha^2$

$2\alpha^2 = \frac{2}{3} \quad \alpha^2 = \frac{1}{3} \quad \alpha = \frac{1}{\sqrt{3}}$

$p(t) = t^3$

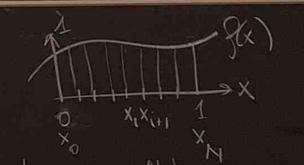
$\int_{-1}^{1} t^3\mathrm{d}t = 0$

$J(p) = (-\alpha)^3 + \alpha^3 = 0$

$p(t) = t^4$

$\int_{-1}^{1} t^4\mathrm{d}t = \left[\frac{t^5}{5}\right]_{-1}^{1} = \frac{2}{5} \neq 1$

$J(p) = (-\alpha)^4 + \alpha^4 = 2\alpha^4 = \frac{2}{9}$



$x = x_i + h\frac{t+1}{2}$

$dx = \frac{h}{2}dt$

$\int_0^1 f(x)dx = \sum_{i=0}^{N-1} \int_{x_i}^{x_{i+1}} f(x)dx$

$= \frac{h}{2} \sum_{i=0}^{N-1} \int_{-1}^{1} f(x_i + h\frac{t+1}{2})dt$

exam2024.pdf ... Displays

Course: Analyse | exam2024.pdf

https://moodle.epfl.ch/pluginfile.php/3444390/mod_resource/content/1/exam2024.pdf

8 of 16 — 160%

+1/8/53+

Etant donné $f : [0,1] \times [0,+\infty[\to \mathbb{R}$, on cherche $u : [0,1] \times [0,+\infty[\to \mathbb{R}$ telle que:

$$\begin{cases} \dfrac{\partial u}{\partial t}(x,t) - \dfrac{\partial^2 u}{\partial x^2}(x,t) - 10\dfrac{\partial u}{\partial x}(x,t) = f(x,t), & 0 < x < 1, t > 0, \\ u(0,t) = u(1,t) = 0, & t > 0, \\ u(x,0) = \sin(\pi x), & 0 < x < 1. \end{cases}$$

Soit N un entier positif, $h = 1/(N+1)$ le pas d'espace, $x_i = ih, i = 0, 1, ..., N+1$. Soit τ le pas de temps, $t_n = n\tau, n = 0, 1, 2, ...$ On note u_i^n une approximation de $u(x_i, t_n)$ obtenue à l'aide du schéma de différences finies suivant :
on écrit l'équation au point x_i et temps t_{n+1} et on approche $\partial u/\partial t$ par une formule de différences finies rétrograde (schéma implicite), $\partial^2 u/\partial x^2$ par une formule de différence finies centrée et $\partial u/\partial x$ par une formule de différence décentrée. A chaque pas de temps, il s'agit de résoudre le système linéaire

$$A\vec{u}^{n+1} = \vec{b}$$

où $b_i = u_i^n + \tau f(x_i, t_{n+1}), i = 1, ..., N$. Le fichier matlab/octave `exam1.m` implémente ce schéma dans le cas où $f(x,t) = -10\pi\cos(\pi x)e^{-\pi^2 t}$.

Question 14 A la ligne `a(i) = 1+2*tau/(h*h) ???;`, il faut remplacer les `???` par

- ☐ `+10*tau/h`
- ☐ `-10*tau/h`
- ☐ `+10*tau/(2h)`
- ☐ `-10*tau/(2h)`

Question 15 A la ligne `c(i)= ...` remplacer les `???` par

- ☐ `+10*tau/(2h)`
- ☐ `+10*tau/(h*h)`

$J(g) = g(-\frac{1}{\sqrt{3}}) + g(\frac{1}{\sqrt{3}})$

$L_h(f) = \frac{h}{2}\sum_{i=0}^{N-1}\left(f(x_i + h\frac{-\frac{1}{\sqrt{3}}+1}{2}) + f(x_i + h\frac{\frac{1}{\sqrt{3}}+1}{2})\right)$

$\frac{\sqrt{3}-1}{2\sqrt{3}}$ $\frac{\sqrt{3}+1}{2\sqrt{3}}$

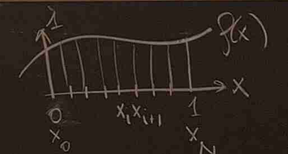
$x(=u^{n+1})$ tq $F(x) = 0 = x - u^n - h f(x, t_{n+1})$

$x^0 = u^n$

$m = 0, 1, 2, ...$

$x^{m+1} = x^m - \frac{F(x^m)}{F'(x^m)} = x^m - \frac{x^m - u^n - h f(x^m, t_{n+1})}{1 - h\frac{\partial f}{\partial x}(x^m, t_{n+1})}$

$f(x,t) = -x - x^5$



$x = x_i + h\frac{t+1}{2}$

$dx = \frac{h}{2}dt$

$\int_0^1 f(x)dx = \sum_{i=0}^{N-1}\int_{x_i}^{x_{i+1}} f(x)dx$

$= \frac{h}{2}\sum_{i=0}^{N-1}\int_{-1}^{1} f(x_i + h\frac{t+1}{2})dt$

$g(t)$

exam2024.pdf ... Displays

Course: Analyse | exam2024.pdf

https://moodle.epfl.ch/pluginfile.php/3444390/mod_resource/content/1/exam2024.pdf

10 of 16 — 160%

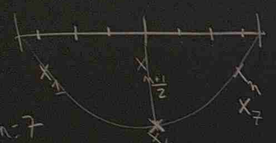
```
 end
 %
 % remplissage de la matrice  A
 %
 for i=1:N
   a(i) = 1+2*tau/(h*h) ???;
 end
 for i=1:N-1
   d(i) = ??? ;
 end
 for i=1:N-1
   c(i) = -tau/(h*h) ???;
 end
 %
 % decomposition A = LU
 %
 c(1)=???;
 for i=2:N-1
   a(i) = a(i)-???;
   c(i) = c(i)/???;
 end
   a(N) = a(N)-???;
 %
 % schema d'Euler retrograde
 %
 for n=1:M
   t=t+tau;
 %
 % second membre du systeme lineaire Ax=b
 %
 for i=1:N
   b(i) = b(i)-tau*10*pi*cos(pi*i*h)*exp(-pi*t)^2
 end
```

$$\frac{u_i^{n+1}-u_i^n}{\tau} - \frac{u_{i-1}^{n+1}-2u_i^{n+1}+u_{i+1}^{n+1}}{h^2} - 10\,\frac{u_{i+1}^{n+1}-u_i^{n+1}}{h} = f(x_i,t_{n+1})$$

t_{n+1}, t_n — $O(\tau)+O(h^2)$ — -10 $\Leftarrow$ — x_{i-1} x_i x_{i+1}

$$\begin{pmatrix} 1+\frac{2\tau}{h^2}+\frac{10\tau}{h} & -\frac{\tau}{h^2}-\frac{10\tau}{h} & & \\ -\frac{\tau}{h^2} & \ddots & \ddots & \\ & \ddots & \ddots & \\ & & & \end{pmatrix} \begin{pmatrix} u_1^{n+1} \\ \vdots \\ \vdots \\ u_N^{n+1} \end{pmatrix} = \begin{pmatrix} u_1^n+\tau f(x_1,t_{n+1}) \\ \vdots \\ \vdots \\ u_N^n+\tau f(x_N,t_{n+1}) \end{pmatrix}$$

$$A = \begin{pmatrix} a_1 & c_1 & & \\ d_1 & a_2 & c_2 & \\ & d_2 & \ddots & \ddots \\ & & \ddots & c_{N-1} \\ & & d_{N-1} & a_N \end{pmatrix}$$



$\forall x \in \mathbb{R}^n, \nabla f(x) = \alpha A^T(Ax-b) + \begin{pmatrix} 0 \\ \vdots \\ x_{\frac{n+1}{2}}-1 \\ \vdots \\ 0 \end{pmatrix}$

$\nabla f(\bar{x}^*) = \bar{0} = \alpha A^T(\underbrace{A\bar{x}^* - \bar{b}}_{\bar{q}^*}) + \begin{pmatrix} 0 \\ \vdots \\ x^*_{\frac{n+1}{2}}-1 \\ \vdots \\ 0 \end{pmatrix}$

$A\bar{x}^* - \bar{b} = \bar{q}^*$

$\alpha A^T\bar{q}^* + \begin{pmatrix} 0 \\ \vdots \\ x^*_{\frac{n+1}{2}}-1 \\ \vdots \\ 0 \end{pmatrix} = \bar{0}$

$\begin{pmatrix} A & -I \\ 0 \cdots 0 & \alpha A^T \end{pmatrix} \begin{pmatrix} \bar{x}^* \\ \bar{q}^* \end{pmatrix} = \begin{pmatrix} \bar{b} \\ 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix}$

$\begin{pmatrix} A & -I \\ 0 \cdots 0 & \alpha A^T \end{pmatrix} \begin{pmatrix} \bar{x} \\ \bar{q} \end{pmatrix} = \begin{pmatrix} \bar{0} \\ \bar{0} \end{pmatrix}$

$A\bar{x} - \bar{q} = \bar{0}$

$(\quad)\bar{x} + \alpha A^T\bar{q} = \bar{0}$

$(\quad)\bar{x} + \alpha A^T A\bar{x} = \bar{0}$

$\left((\quad) + \alpha A^T A \right)\bar{x} = \bar{0}$

$\bar{x}^T \left((\quad) + \alpha A^T A \right)\bar{x} = 0$

$\underbrace{\bar{x}^T (\quad)\bar{x}}_{\geq 0} + \bar{x}^T A^T A\bar{x} = 0$

$\bar{x}^T A^T A\bar{x} = 0$

$\|A\bar{x}\|^2 = 0 \Rightarrow A\bar{x} = \bar{0} \Rightarrow \bar{x} = \bar{0}$

- ☐ $\|\vec{x}^0\| = \sum_{j=1}^{n} \beta_j(\lambda_j)$
- ☐ $\vec{x}^k = \sum_{j=1}^{n} \beta_j(\lambda_j)^k \vec{\varphi}_j$

Question 25 On suppose $\beta_n > 0$ et $\lambda_{n-1} < \lambda_n$, on a

- ☐ $\lim_{k\to\infty} \frac{\vec{x}^{k\,T} A \vec{x}^k}{\|\vec{x}^k\|^2} = \lambda_n$
- ☐ $\frac{\vec{x}^{k\,T} A \vec{x}^k}{\|\vec{x}^k\|^2} = \frac{\sum_{j=1}^{n} (\beta_j)^2 (\lambda_j)^{2k+1}}{\sum_{j=1}^{n} (\beta_j)^2 (\lambda_j)^{2k}}$
- ☐ $\lim_{k\to\infty} \frac{\vec{x}^k}{\|\vec{x}^k\|} = \vec{\varphi}_n$

Montrer, que $\forall x \in \mathbb{R}, \forall u \in C^4([x-2, x+2]), \exists C > 0, \forall h \in]0,1]$ on a

$$\left| u'(x) - \frac{u(x-2h) - 8u(x-h) + 8u(x+h) - u(x+2h)}{12h} \right| \leq Ch^3.$$

On demande d'expliciter C.

$0<h\leq 1$

$8\cdot\left(u(x+h) = u(x) + hu'(x) + \frac{h^2}{2}u''(x) + \frac{h^3}{6}u'''(x) + \frac{h^4}{24}u^{(4)}(\alpha)\right)$ $\quad x<\alpha<x+h$

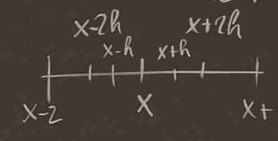
$-1\cdot\left(u(x+2h) = u(x) + 2hu'(x) + \frac{4h^2}{2} + \frac{8h^3}{6} + \frac{16h^4}{24}u^{(4)}(\beta)\right)$ $\quad x<\beta<x+2h$

$-8\cdot\left(u(x-h) = u(x) - hu'(x) + \frac{h^2}{2}u''(x) - \frac{h^3}{6}u'''(x) + \frac{h^4}{24}u^{(4)}(\gamma)\right)$ $\quad x-h<\gamma<x$

$+1\cdot\left(u(x-2h) = u(x) - 2hu'(x) + \frac{4h^2}{2}u''(x) - \frac{8h^3}{6}u'''(x) + \frac{16h^4}{24}u^{(4)}(\delta)\right)$ $\quad x-2h<\delta<x$

$$8u(x+h) - u(x+2h) - 8u(x-h) + u(x-2h) = 12hu'(x) + \frac{h^4}{24}\left(8u^{(4)}(\alpha) - 16u^{(4)}(\beta) - 8u^{(4)}(\gamma) + 16u^{(4)}(\delta)\right)$$

$$\left|u'(x) - \frac{8u(x+h) - 8u(x-h) - u(x+2h) + u(x-2h)}{12h}\right| = \frac{h^3}{12\cdot 24}\left|8u^{(4)}(\alpha) - 16u^{(4)}(\beta) - 8u^{(4)}(\gamma) + 16u^{(4)}(\delta)\right|$$



$$\leq \frac{h^3}{12\cdot 24}\left(8\left|u^{(4)}(\alpha)\right| + 16\left|u^{(4)}(\beta)\right| + 8\left|u^{(4)}(\gamma)\right| + 16\left|u^{(4)}(\delta)\right|\right)$$

$$\leq \frac{h^3}{12\cdot 24}\left(8+16+8+16\right)\max_{x-2\leq y\leq x+2}\left|u^{(4)}(y)\right| \qquad \forall\, 0<h\leq 1$$