Séance 12 : compléments en algèbre et analyse numérique

Sujets abordés :

* Problèmes aux moindres carrés
* Décomposition en valeurs singulières
* Minimisation de fonctions quadratiques ou convexes
* Problèmes aux limites (EDO, EDP)

Textes illustratifs

* Une requête bibliographique
* Poids en équilibre sur une corde

1) Problèmes aux moindres carrés et fonctionnelle quadratique

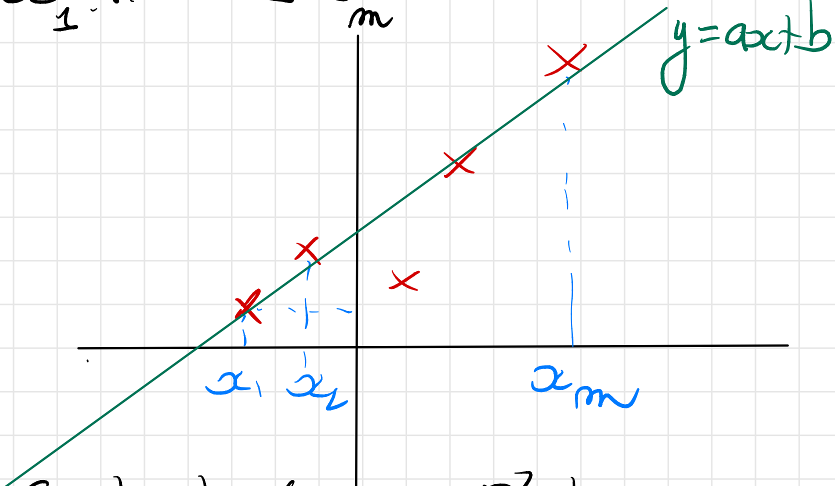
Déf : soit $n \leq m$ deux entiers ≥ 1 et $B \in \mathcal{M}_{m,n}(\mathbb{R})$, $c \in \mathbb{R}^m$.

On appelle solution du problème aux moindres carrés associé à B et c, tout élément $u \in \mathbb{R}^n$ tq

$$\boxed{\|Bu - c\|_2^2 = \inf_{x \in \mathbb{R}^n} \|Bx - c\|_2^2}$$

(on dispose de plus d'équations que d'inconnues)

Exemple : soit $m \geq 2$, on dispose de données $(x_i, y_i)_{1 \leq i \leq m}$ tq $x_1 < \dots < x_m$



On cherche $(a,b) \in \mathbb{R}^2$ tq $\sum_{i=1}^{m} (ax_i + b - y_i)^2$ soit minimal.

Il s'agit d'un problème de type moindres carrés avec $n = 2$ et

$$B = \begin{pmatrix} x_1 & 1 \\ \vdots & \vdots \\ x_m & 1 \end{pmatrix}, \quad c = \begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix}$$

($u = \begin{pmatrix} a \\ b \end{pmatrix}$ solution à rechercher)

On a le résultat suivant :

Théorème : une solution u du problème aux moindres carrés est exactement une solution du système $n \times n$:

$$\boxed{{}^tB B u = {}^tB c} \quad (1)$$

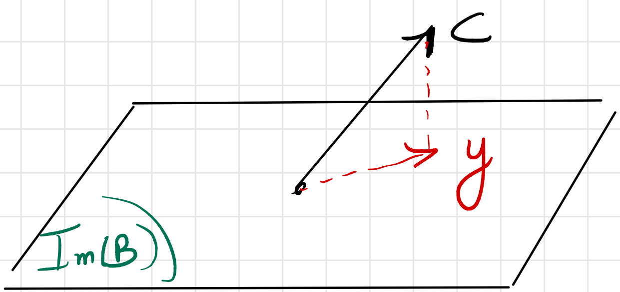
Une telle solution existe toujours.

Elle est unique ssi $\mathrm{Ker}(B) = \{0\}$

preuve (voir poly)

Remarque : on peut aussi invoquer le théorème de Riesz par l'existence et l'unicité de $y \in \mathrm{Im}(B)$ $(\subset \mathbb{R}^m)$ convexe, fermé de $\mathbb{R}^m$ tq

$$\|y-c\|^2 = \inf_{z \in \mathrm{Im}(B)} \|z-c\|$$



L'unicité de u tel que $Bu = y$ provient de l'injectivité de B.

D'un point de vue algorithmique, il n'est pas efficace de se ramener à la résolution de l'équation normale (1) par des raisons de coût et de robustesse.

(exemple : $B = \begin{pmatrix} 1 & 1 & 1 \\ \varepsilon & 0 & 0 \\ 0 & \varepsilon & 0 \\ 0 & 0 & \varepsilon \end{pmatrix}$, par lequel $\,^tBB = \begin{pmatrix} 1+\varepsilon^2 & 1 & 1 \\ 1 & 1+\varepsilon^2 & 1 \\ 1 & 1 & 1+\varepsilon^2 \end{pmatrix}$)

On utilise plutôt une factorisation QR de la matrice B non carrée :

$$\underbrace{B}_{m\times n} = \underbrace{Q}_{m\times m} . \underbrace{R}_{m\times n} \left(\text{où } R = \begin{pmatrix} * \\ 0 \end{pmatrix}\right)$$

On a alors

$$\| Bx - c\|_2^2 = \| {}^tQ(Bx-c)\|_2^2$$

$$= \| Rx - {}^tQc\|_2^2$$

$$= \| \underbrace{({}^tQc)_n}_{\in \mathbb{R}^n} - \underbrace{R_1}_{n\times n} x\|^2 + \|({}^tQc)_{m-n}\|^2$$

(où ${}^tQc = \begin{pmatrix} ({}^tQc)_n \\ ({}^tQc)_{m-n} \end{pmatrix} \begin{array}{l} \updownarrow n \\ \updownarrow m-n \end{array}$

et $R = \begin{pmatrix} R_1 \\ \hline 0 \end{pmatrix}$)

Il ne reste alors qu'à résoudre le système triangulaire : $R_1 x = ({}^tQc)_n$.

Remarque : le problème aux moindres carrés consiste en la recherche du minimum de la fonction

$$J(x) = \|Bx-c\|^2 = \langle Bx-c, Bx-c\rangle$$

$$= \langle {}^tBBx, x\rangle - \langle x, 2\,{}^tBc\rangle + \langle c, c\rangle$$

$$= \frac{1}{2}\langle Ax, x\rangle - \langle b, x\rangle + cste$$

où $\begin{cases} A = 2\,{}^tBB \in \mathcal{S}_n(\mathbb{R}), \text{ positive} \\ b = 2\,{}^tBc \end{cases}$

(fonctionnelle quadratique si $A \gg 0$ ie si B injective)

2) Minimisation de fonctionnelles quadratiques ou convexes

On s'intéresse ici à la minimisation d'une fonctionnelle $f : \mathbb{R}^n \to \mathbb{R}$ supposée convexe (resp. s^r convexe) :

$\forall (x,y) \in \mathbb{R}^n$ $(x \neq y)$ $\forall \lambda \in [0,1]$ $(]0,1[)$

$$f(\lambda x + (1-\lambda) y) \leq (<) \lambda f(x) + (1-\lambda) f(y)$$

en particulier des fonctionnelles quadratiques :

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

$(A \gg 0$ et $b \in \mathbb{R}^n)$

Théorème : on suppose que f est une fonctionnelle strictement convexe et coercive, c'est à dire :

$$\lim_{\|x\| \to +\infty} f(x) = +\infty$$

Dans ce cas, la fonction f possède un unique minimum $x^* \in \mathbb{R}^n$.

Si de plus, f est C^1, x^* est l'unique solution du système (non linéaire)

$$\nabla f(x) = 0$$

Il est possible d'obtenir une approximation de x^* avec une suite itérative du type

$$\boxed{\begin{cases} x_0 \in \mathbb{R}^n \\ x_{k+1} = x_k - \alpha_k \nabla f(x_k) \end{cases}}$$

où $(\alpha_k) \in (\mathbb{R}_+^*)^{\mathbb{N}}$)

preuve (partielle)

→ coercivité ⇒ existence de x^* (par restriction sur un compact)

→ stricte convexité ⇒ unicité

→ $\nabla f(x) = 0$: point critique

→ x^* unique pt critique car f convexe

→ gradient (admis)

Cas particulier important : fonctionnelle quadratique :

$J(x) = \frac{1}{2} \langle Ax, x \rangle - \langle b, x \rangle$ avec $A \in S_n^{++}(\mathbb{R})$ et $b \in \mathbb{R}^n$.

→ J strictement convexe ($A \gg 0$)

→ J coercive car $J(x) \geq \frac{1}{2} \lambda_m \|x\|^2 - \|b\| \|x\| \nearrow +\infty$

(λ_m : plus petite valeur propre de A)

→ J est C^1 et $\nabla J(x) = Ax - b$

(en effet :

$J(x+h) = J(h) + \langle \overbrace{Ax - b}^{\nabla J(x)}, h \rangle + o(\|h\|)$)

La recherche du minimum de J revient à la résolution du système linéaire :

$$Ax = b$$

(par lequel on dispose de différentes méthodes : Cholesky, Jacobi, Gauss-Seidel et... gradient)

→ Méthode proposée ici :

$$\begin{cases} x_0 \in \mathbb{R}^n \\ x_{k+1} = x_k - \alpha_k \overbrace{(Ax_k - b)}^{\nabla J(x_k)} \end{cases}$$

On retrouve la méthode déjà vue du gradient. Celle-ci converge si $\alpha_k \in]m, M[$ où $0 < m$ et $M < \frac{2}{\rho(A)}$.

3) Décomposition en valeurs singulières (SVD)

On a le résultat suivant relatif à la réduction d'une matrice par matrices équivalentes :

Théorème : soit $A \in \mathcal{M}_{n,m}(\mathbb{R})$ et de rang $r \leq \min(m, n)$.

Il existe $(U, V) \in O_n(\mathbb{R}) \times O_m(\mathbb{R})$ telles que :

$$A = U \Sigma_r {}^t V$$ où

$$\Sigma_r = \begin{pmatrix} \sigma_1 & & 0 & \vdots & \\ 0 & \ddots & & \vdots & \\ & & \sigma_r & \vdots & \\ \cdots & \cdots & \cdots & 0 & \\ & & & & \ddots \\ & & & & & 0 \end{pmatrix}$$

pour une unique famille de coefficients :
$\sigma_1 \geq \sigma_2 \cdots \geq \sigma_r > 0$

On dit que les coefficients (σ_i) sont les valeurs singulières de A.

preuve : admis ([Ciarlet], ...)

Les σ_i représentent les racines carrées des valeurs propres de ${}^tA\,A$ (et $A\,{}^tA$) et les r colonnes de U et V sont des vecteurs propres associés de ces 2 matrices respectives.

Remarque : à nouveau, on peut interpréter ce résultat en termes de minimisation :

$$|||A - U\Sigma_k{}^tV|||_2 = \operatorname{Inf}_{\begin{cases} B \in \mathcal{M}_{mn}(\mathbb{R}) \\ rg(B) = k \end{cases}} |||A - B|||_2$$

où $\Sigma_k = \begin{pmatrix} \sigma_1 & & & \vdots & \\ & \ddots & & \vdots & \\ & & \sigma_k & \vdots & \\ \cdots & \cdots & \cdots & 0 & \\ & 0 & & \vdots & \ddots \\ & & & \vdots & & 0 \end{pmatrix}$ $(k \leq r)$

(meilleure approximation de rang k de A)

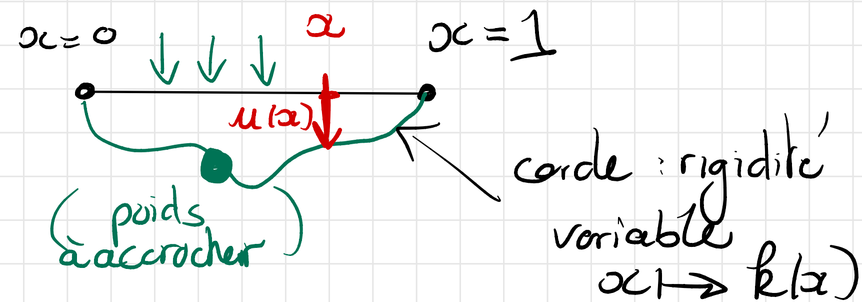
Cette propriété est utilisée dans

de nombreuses applications (voir texte : requête bibliographique)

La recherche de (U, Σ_r, V) peut se faire efficacement à l'aide d'un algorithme basé sur la factorisation QR (voir texte requête bibliographique).

4) Problèmes aux limites

On aborde ce dernier thème sur l'exemple du texte : poids en équilibre sur une corde.



La corde se déforme en x (déformation $x \mapsto u(x)$). Cette déformation est solution du problème différentiel aux limites suivant :

$$(1)\quad \begin{cases} -\dfrac{\partial}{\partial x}\left(k(x)\dfrac{\partial u(x)}{\partial x}\right) = f(x) \\ u(0) = u(1) = 0 \end{cases}$$

Qn : où positionner le poids ponctuel pour limiter le risque de rupture ?

On peut montrer que pour tout fonction $k \in \mathcal{C}([0,1], \mathbb{R}_+^*)$, et $f \in \mathcal{C}([0,1], \mathbb{R})$, il existe une unique solution $u \in \mathcal{C}^2([0,1], \mathbb{R})$ (méthode de tir).

* Expression du risque de rupture aux extrémités :

$$R = \frac{1}{2}\left|k(0)\frac{\partial u}{\partial x}(0)\right|^2 + \frac{1}{2}\left|k(1)\frac{\partial u}{\partial x}(1)\right|^2$$

à minimiser en fonction du choix de la position $a \in]0,1[$ du poids ponctuel.

* Discrétisation du problème :

→ problème aux limites : méthode des différences finies (voir texte)

→ fonction de rupture

→ gradient de la fonction de rupture (voir texte)

→ Implémentation : (voir page web)

on utilise une méthode de gradient pour minimiser la fonctionnelle convexe $a \mapsto R(a)$

(à noter que si k = cste dans le cas d'une corde homogène, la fonctionnelle est quadratique)