Séance 22a : compléments en algèbre linéaire

On présente deux notions indépendantes présentes au programme d'écrit : problèmes aux moindres carrés et décomposition en valeurs singulières.

I) Problèmes de type moindres carrés

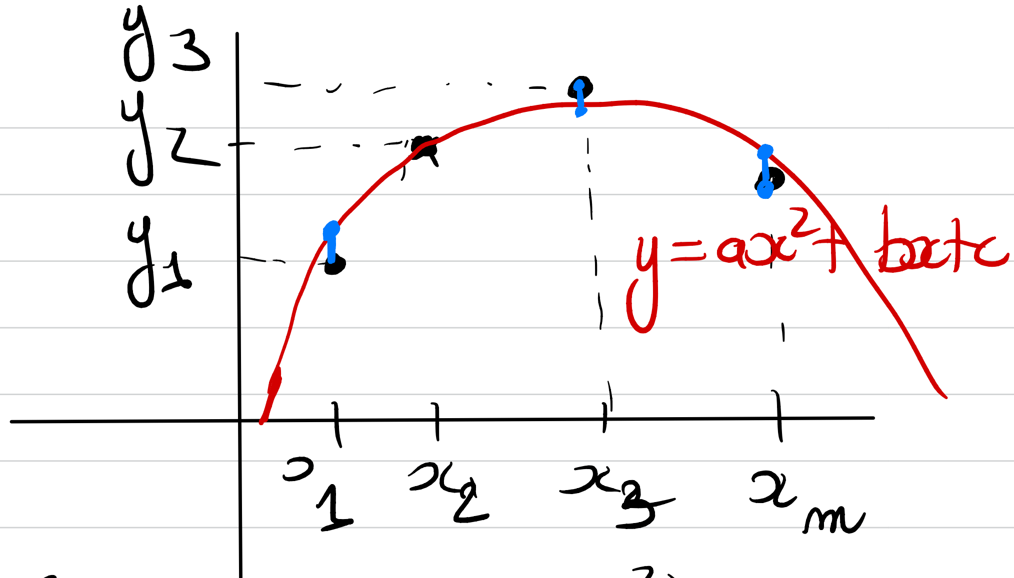
On se trouve ici dans la résolution d'un problème de type linéaire contenant plus d'équations que d'inconnues : "$Bx = c$"

où $B \in \mathcal{M}_{m,n}(\mathbb{R})$ avec $m > n$ (en général $rg(B) = n$) et $c \in \mathbb{R}^m$. On cherche à définir une unique solution $x \in \mathbb{R}^n$ "au sens des moindres carrés" : $x \in \mathbb{R}^n$ est tel que

$$\boxed{\|Bx - c\|_2^2 = \underset{y \in \mathbb{R}^n}{\cancel{\mathrm{Inf}}\ (\mathrm{Min})} \|By - c\|_2^2 \quad (1)}$$

($\| \ \|_2$: norme euclidienne de $\mathbb{R}^m$)

Exemple : approximation polynomiale d'un jeu de données $(x_i, y_i)_{1 \leq i \leq m}$



Ici $B = \begin{pmatrix} 1 & x_1 & x_1^2 \\ \vdots & & \\ 1 & x_m & x_m^2 \end{pmatrix}$

$c = \begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix}$

et x (à déterminer) $= \begin{pmatrix} c \\ b \\ a \end{pmatrix}$

On a le résultat d'existence et d'unicité de x suivant :

2

Théorème : le problème (1) admet une solution x tel que

$$\boxed{{}^tBBx = {}^tBc}$$ (système $n \times n$)

Cette solution est unique si $rg(B)=n$ (ou de manière équivalente $Ker(B)=\{0\}$)

preuve : il existe plusieurs démonstrations possibles :

→ en étudiant le problème d'optimisation associé :

- Minimiser

$$y \longmapsto \|By-c\|_2^2 = \langle {}^tBBy, y\rangle - 2\langle {}^tBc, y\rangle + {}^tcc$$

→ tous les minimum de cette

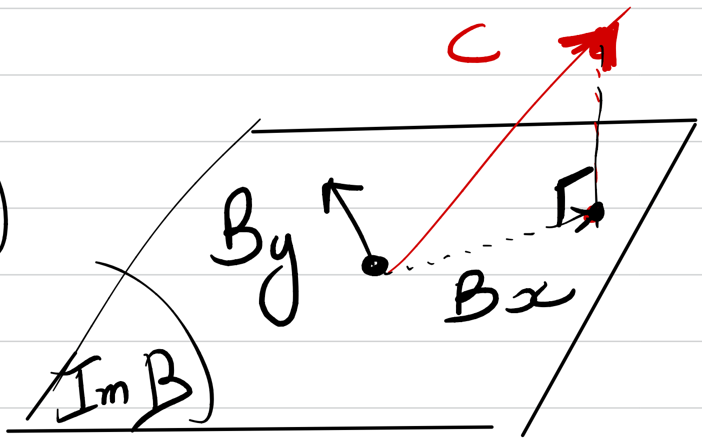
fonction correspondant aux points critiques tels que

$$\nabla J(y) = 2\,{}^tBBy - 2\,{}^tBc = 0$$

L'unicité s'obtient exactement lorsque ${}^tBB$ est symétrique définie positive, c'est à dire lorsque $\ker B = \{0\}$

→ en utilisant le théorème de projection de Riesz

(dans $\mathbb{R}^m$)



le problème revient exactement à rechercher la projection de c sur Im(B). Celle-ci existe, est unique (sans que x soit forcément unique, vrai si B injective).

La caractérisation de cet élément s'écrit : $\forall y \in \mathbb{R}^n$

$$\langle By, c - Bx \rangle = 0, \text{ soit :}$$

$$\langle y, {}^tBc - {}^tBBx \rangle = 0$$

$$\Rightarrow {}^tBc = {}^tBBx \;//$$

D'un point de vue algorithmique, la résolution du problème (1) via l'équation normale n'est pas satisfaisante d'un point de vue stabilité numérique et coût de calcul (si $m >> n$). On utilise plutôt une factorisation QR de la matrice B :

$$B = \underset{(m\times m)}{Q}\ \underbrace{\begin{bmatrix} R_1 \\ 0 \end{bmatrix}}_{\substack{R \\ (m\times n)}} \qquad \begin{array}{l} \updownarrow n \\ \updownarrow m \end{array}$$

Le problème aux moindres carrés se récrit alors :

$$\|By-c\|_2^2 \overset{\text{isométrie}}{=} \|{}^tQ\,(By-c)\|_2^2$$

$$= \|Ry - \underbrace{{}^tQc}_{m\times 1}\|_2^2$$

$$= \|R_1 y - d_1\|_2^2 + \underbrace{\|d_2\|^2}_{\text{cste}} \quad \text{où}$$

$${}^tQc = \begin{bmatrix} d_1 \\ d_2 \end{bmatrix} \begin{array}{l} \updownarrow n \\ \\ \end{array}$$

On prend alors x tel que $R_1 x = d_1$ (par remontée)

II] Décomposition en valeurs singulières

Il existe un résultat de réduction d'une matrice en une matrice diagonale équivalente, possédant des propriétés intéressantes :

Théorème : soit $A \in \mathcal{M}_{m,n}(\mathbb{R})$ de rang $r \le \min(m,n)$. Alors il existe un couple (U,V) tq $U \in O_m(\mathbb{R})$, $V \in O_n(\mathbb{R})$ et une unique famille de réels $\sigma_1 \geq \sigma_2 \ldots \geq \sigma_r > 0$ tels que

$$A = \overset{m\times m}{U} \overset{m\times n}{\left(\begin{array}{ccc|c} \sigma_1 & & 0 & \\ & \ddots & & 0 \\ 0 & & \sigma_r & \\ \hline & 0 & & 0 \end{array}\right)} \overset{n\times n}{{}^tV}$$

(matrice de taille m lignes, n colonnes)

preuve : (voir [Ciarlet]). On a en particulier :

* $\sigma_1^2 \ldots \sigma_r^2$: valeurs propres de ${}^tAA$

$$\left({}^tAA = V \Sigma_r \underbrace{{}^tU U}_{Id} \Sigma_r {}^tV\right)$$

* V : vecteurs propres de ${}^tAA$ $\big)_{n\times n}$
* U : vecteurs propres de $A\,{}^tA$ $\big)_{m\times m}$

D'un point de vue algorithmique, il existe des algorithmes de SVD (singular value decomposition) basé sur la factorisation QR (voir Ciarlet).

Cette décomposition contient en elle un résultat intéressant de projection de A sur des espaces plus simple :

$$\| A - U \begin{pmatrix} \overbrace{\begin{matrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_k \end{matrix}}^{k} & & 0 \\ & 0 & \\ 0 & & \ddots \\ & & & 0 \end{pmatrix}^r V \|^2 = \dots$$

$$\dots = \operatorname{Inf}_{\begin{cases} B \in \mathcal{M}_{m,n}(\mathbb{R}) \\ rg(B) = k \end{cases}} \| A - B \|^2$$

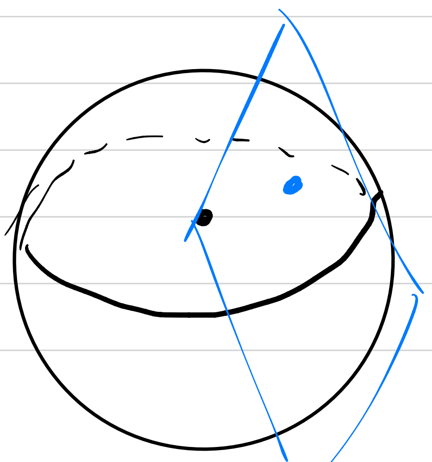
$(k \leq r)$

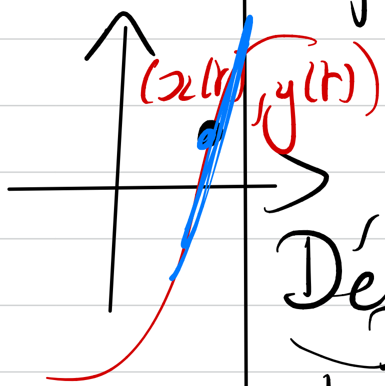
Applications : recherche biblio, etc...

Séance 22b : introduction à la géométrie différentielle

Il s'agit de présenter ici le vocabulaire (et les notions principales) permettant d'étendre la notion de courbe sur $\mathbb{R}^2$ ou $\mathbb{R}^3$ ou de surface sur $\mathbb{R}^3$.

$$\sum_{i=1}^{n} x_i^2 = 1$$





Notation : on se place dans $\mathbb{R}^n$ et si nécessaire, on identifie $\mathbb{R}^n$ et $\mathbb{R}^p \times \mathbb{R}^{n-p}$
$$(x) \equiv (y, z)$$

On donne ici la définition initiale d'une sous variété qui sera réinterprétée de manière plus "facilement représentable" ultérieurement.

Déf M est une sous variété de dimension p, de classe C^k

si pour tout $x_0 \in M$, il existe un voisinage U de x_0 et

$\varphi : U \to \mathbb{R}^n$, C^k difféo tel que

$$\varphi(M \cap U) = \varphi(U) \cap (\mathbb{R}^p \times \{0\}_{n-p})$$

(M est C^k difféomorphe localement à un sev de dimension p)

Il existe 3 définitions équivalentes (dont la justification repose sur les théorèmes des fonctions implicites et de l'inversion locale)

Déf équivalente 1 (équation locale)

M est une sous variété de dimension p et de classe C^k, si pour tout $x_0 \in M$, il existe U voisinage ouvert et

$f : U \to \mathbb{R}^{n-p}$ C^k avec $df(x_0) \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^{n-p})$ surjective, telle que

$$\boxed{M \cap U = \{x \in U \,/\, f(x) = 0\}}$$

3

Déf. équivalente 2 (paramétrisation locale)

M est une sous variété de $\mathbb{R}^n$ de dimension p, de classe C^k, si pour tout $x_0 \in M$, il existe U voisinage de x_0 et $j : V \to \mathbb{R}^n$ (V voisinage de $\{0\}$ dans $\mathbb{R}^p$) de classe C^k, $j(0) = x_0$, $dj(0) \in \mathcal{L}(\mathbb{R}^p, \mathbb{R}^n)$ injective, tels que :

$$\boxed{M \cap U = j(V)}$$

Définition équivalente 3 (graphe local)

M est une sous variété de dimension p, de classe C^k, si pour tout $x_0 = (\underbrace{y_0}_{\mathbb{R}^p}, \underbrace{z_0}_{\mathbb{R}^{n-p}})$, il existe V voisinage ouvert de y_0 et W voisinage ouvert de z_0 et une fonction $g : V \to W$, C^k, et $A \in GL_n(\mathbb{R})$ tels que

rotation

$$\boxed{M \cap (V \times W) = \left\{ A \begin{pmatrix} y \\ g(y) \end{pmatrix}, y \in V \right\}}$$

4

Exemple : la sphère unité S^1 de $\mathbb{R}^n$ est une sous variété de dimension $n-1$, de classe C^∞ ;

$S^1 \cap \mathbb{R}^n = \{x \in \mathbb{R}^n \text{ tq } f(x) = 0\}$

où $f : \begin{cases} \mathbb{R}^n \to \mathbb{R}^{n-(n-1)} \\ x \mapsto \sum_{i=1}^n x_i^2 - 1 \end{cases}$

$(df(x_0).h = 2\langle x_0, h\rangle)$

si U_p ouvert de $\mathbb{R}^p$

$U_p \times \{0\}_{\mathbb{R}^{n-p}}$ est une sous variété de $\mathbb{R}^n$: avec

$\varphi : \mathbb{R}^n \xrightarrow{Id} \mathbb{R}^n$, on a

$M \cap U = U \cap (\mathbb{R}^p \times \{0\}_{\mathbb{R}^{n-p}})$

(ouvert de $\mathbb{R}^n$)

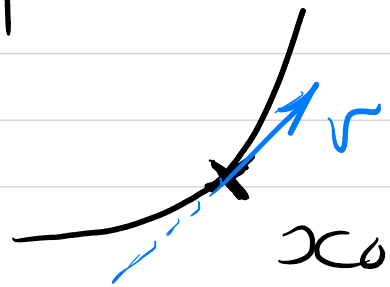
On peut associer en tout $x_0 \in M$, la notion d'espace tangent :

Déf : on note

$T_{x_0}M = \{v \in \mathbb{R}^n, \exists \gamma : I \xrightarrow{C^1} M,$

I intervalle ouvert de $\mathbb{R}^n$ contenant 0,

tel que $\gamma(0) = x_0$ et $\gamma'(0) = v\}$



On l'appelle espace tangent à M en x_0.

On a la représentation suivante de $T_{x_0}M$ en fonction des 4 définitions :

Théorème : soit M sous variété de $\mathbb{R}^n$, de dimension p, de classe C^k. $T_{x_0}M$ est un sous espace vectoriel de $\mathbb{R}^n$ caractérisé par :

$\rightarrow T_{x_0}M = d\varphi(x_0)^{-1}(\mathbb{R}^p \times \{0\}_{\mathbb{R}^{n-p}})$

$\rightarrow T_{x_0}M = \mathrm{Ker}(df(x_0))$

$\rightarrow T_{x_0}M = \mathrm{Im}(dj(0))$

$\rightarrow T_{x_0}M = \left\{ A \begin{pmatrix} h \\ dg(x_0).h \end{pmatrix}, h \in \mathbb{R}^p \right\}$

(preuve : admis)

On présente enfin la méthode des extremas liés par rechercher le minimum d'une fonction sous contrainte (c'est à dire sur une variété).

Théorème : soit $f : U \rightarrow \mathbb{R}$ où f C^1 sur U ouvert de $\mathbb{R}^n$.

Soit $M = \left\{ x \in U \,/\, \begin{cases} g_1(x) = 0 \\ \vdots \\ g_p(x) = 0 \end{cases} \right\}$

où $g_1, \dots, g_p : U \rightarrow \mathbb{R}^n$, C^1.

Soit x_0 un point de minimum local de f sur M.

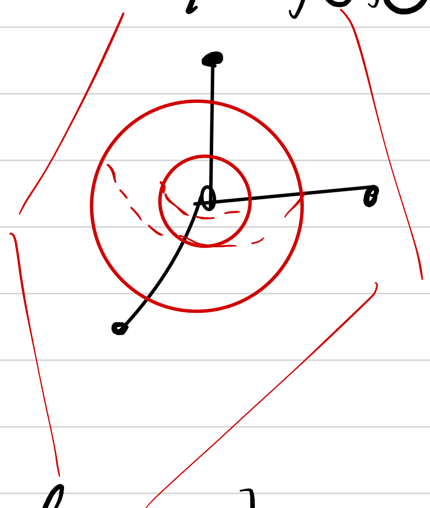
alors, il existe $\lambda_1, \dots, \lambda_p \in \mathbb{R}$
(appelés multiplicateurs de Lagrange)
tels que

$$\boxed{\nabla f(x_0) = \sum_{i=1}^{p} \lambda_i \nabla g_i(x_0)}$$

(Condition nécessaire)

preuve : admis (on utilise un lemme d'Algèbre et la caractérisation "équation locale" de $T_{x_0}M$.

Exemple : minimiser
$$f(x,y,z) = x^2+y^2+z^2$$
sur $M = \{(x,y,z),\ x+y+z=1\}$ 6



Par les extremas liés, si $(x_0,y_0,z_0) \in M$ minimise f sur M localement, il existe $\lambda \in \mathbb{R}$ tq

$$\left.\begin{array}{l} 2x_0 = \lambda 1 \\ 2y_0 = \lambda . 1 \\ 2z_0 = \lambda . 1 \\ \text{et } x_0+y_0+z_0 = 1 \end{array}\right\} \Rightarrow \left\{\begin{array}{l} x_0 = \frac{1}{3} \\ y_0 = \frac{1}{3} \\ z_0 = \frac{1}{3} \\ \lambda = \frac{2}{3} \end{array}\right.$$

Réciproquement, f possède un minimum sur M car M est fermé et f coercive :

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

Il s'agit donc forcément de $(x_0, y_0, z_0) = \left(\frac{1}{3}, \frac{1}{3}, \frac{1}{3}\right)$