Cours 12.2
05.12.24

NPO ! • ex 3 à rendre → 10 déc
• exas interm. 22 & 23 sur Moodle
• question ouverte 23 sur Moodle sem. 12

Rappel :

si $W \subseteq \mathbb{R}^n$, alors $\exists$ T.L

$proj_W : \mathbb{R}^n \to \mathbb{R}^n$

telle que $Im(proj_W) = W$
$Ker(proj_W) = W^\perp$

Thm proj orthogonale :

et $\forall y \in \mathbb{R}^n$ on a $y = \underbrace{proj_W(y)}_{\hat{y}} + z$ avec $\hat{y} \in W$, $z \in W^\perp$

← écriture unique

et $\forall w \in W \setminus \{\hat{y}\}$ on a $\|y - \hat{y}\| < \|y - w\|$ ♡

♡ càd $\hat{y}$ est le vecteur de W qui est le plus proche de y.

En effet, soit $w \in W \setminus \{\hat{y}\}$

$\left(\hat{y} = proj_W(y) \in W \right)$

écrivons

$$y - w \overset{\text{astuce}}{=} \underbrace{y - \hat{y}}_{\in W^\perp} + \underbrace{\hat{y} - w}_{\in W}$$

donc $y - \hat{y}$ et $\hat{y} - w$ sont orthogonaux

Thm de Pythagore

$$\|y-w\|^2 = \|y-\hat{y}\|^2 + \underbrace{\|\hat{y}-w\|^2}_{>0}$$

or $\hat{y} \neq w$

donc $\|\hat{y}-w\| \neq 0$

$$\|y-w\|^2 > \|y-\hat{y}\|^2$$

$\sqrt{\ }$

$$\|y-w\| > \|y-\hat{y}\|$$

§ 6.7 Méthode des moindres carrés

(Important)

Problème 6.7.1 :

Si $A \in M_{m\times n}(\mathbb{R})$, $b \in \mathbb{R}^m$

on sait si $Ax = b$ possède une solution

($\Leftrightarrow b \in \mathrm{Im}(A) = \mathrm{Col}(A)$)

mais qu'en est-il si $Ax \neq b$ $\forall x \in \mathbb{R}^n$?

(i.e $b \notin \mathrm{Im}(A)$, système incompatible)

<u>Idée</u> : Chercher le(s) $\hat{x} \in \mathbb{R}^n$

tel(s) que $A\hat{x} - b$ soit aussi "petit" que possible. Plus précisément

on cherche à minimiser la norme de $A\hat{x} - b$

càd on cherche les $\hat{x}$ t.q $\|A\hat{x} - b\|$ minimal

<u>Déf 6.7.2</u> : On dit que $\hat{x} \in \mathbb{R}^n$ est

une <u>pseudo-solution de $Ax = b$</u>

ou <u>solution au sens des moindres carrés de $Ax=b$</u>

si $\forall x \in \mathbb{R}^n$ on a

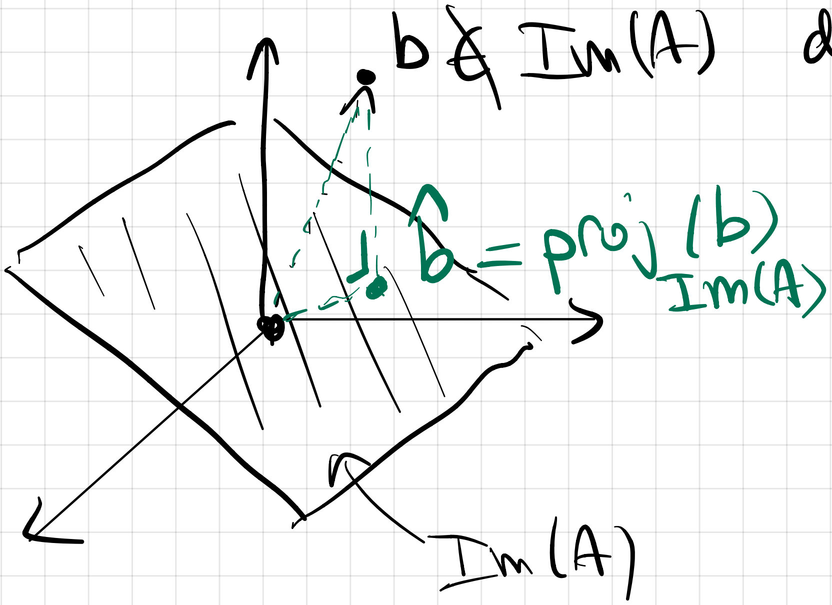
$$\|\underbrace{A\hat{x}}_{\in \operatorname{Im}(A)} - b\| \overset{\star}{\leq} \|\underbrace{Ax}_{\in \operatorname{Im}(A)} - b\|$$

$\star$ dit que la distance de $A\hat{x}$ à b

est la plus petite possible entre

les vecteurs de $\operatorname{Im}(A)$ et b.

Idée 6.7.3 : Comment trouver le(s) $\hat{x}$?



$b \notin Im(A)$ donc $Ax = b$ incompatible

donc on projette orthogonalement b sur $W = Im(A)$

Prenons $W = Im(A) = Col(A)$

et posons $\hat{b} = proj_W(b) \in Im(A)$

et l'on a $\|\hat{b} - b\| \leq \|Ax - b\| \quad \forall x \in \mathbb{R}^n$

Comme $\hat{b} \in Im(A)$, il existe au moins un $\hat{x} \in \mathbb{R}^n$ t.q $A\hat{x} = \hat{b}$ et $\hat{b}$ est le vect de $Im(A)$ le plus proche de b. On déduit que

$\hat{x}$ est solution au sens des moindres carrés de $Ax = b$ $\iff$ $A\hat{x} = \hat{b} = proj_{Im(A)}(b)$

But : Trouver tous les $\hat{x}$ grace à ↺

Thm 6.7.4 (Equation normale & moindres carrés) (important)

$A \in M_{m \times n}(\mathbb{R})$, $b \in \mathbb{R}^m$. Alors

L'ensemble des solutions au sens des moindres carrés de $Ax = b$

est égal à

$$S = \{ \hat{x} \in \mathbb{R}^n \mid A^T A \hat{x} = A^T b \}$$

← l'équation normale de $Ax = b$

dém idée :

supp. $\hat{x}$ sol. au sens des moindres □ de $Ax = b$

c.à.d $A\hat{x} = \hat{b} = proj_{Im(A)}(b)$

on a que $b - \hat{b} \in Im(A)^{\perp} = Col(A)^{\perp}$

donc si a_j est une colonne de A on aura

$$a_j \cdot (b - \hat{b}) = 0$$

↖ produit scalaire

et donc $a_j^T (b-\hat{b}) = 0$

produit matrice-ligne × matrice colonne

et les lignes de A^T sont les colonnes de A

donc $A^T(b-\hat{b}) = 0$

$\Rightarrow A^T b = A^T \hat{b}$ or $\hat{b} = A\hat{x}$

donc $A^T b = A^T A \hat{x}$

càd $\hat{x} \in S$. (fin démo 6.7.4)

<u>Déf 6.7.5</u> : on appel <u>écart quadratique</u> la distance $\|A\hat{x}-b\|$

$= \|\hat{b}-b\|$

<u>Ex 6.7.6</u>

$$A = \begin{pmatrix} 4 & 0 \\ 0 & 2 \\ 1 & 1 \end{pmatrix}_{3\times 2} \qquad b = \begin{pmatrix} 2 \\ 0 \\ 11 \end{pmatrix}$$

(exo : $b \notin \mathrm{Im}(A)$)

Thm-Recette 6.7.4 :

$$A^TA = \begin{pmatrix} 4 & 0 & 1 \\ 0 & 2 & 1 \end{pmatrix}\begin{pmatrix} 4 & 0 \\ 0 & 2 \\ 1 & 1 \end{pmatrix} \underset{\text{exo}}{=} \begin{pmatrix} 17 & 1 \\ 1 & 5 \end{pmatrix}$$

$$A^Tb = \begin{pmatrix} 4 & 0 & 1 \\ 0 & 2 & 1 \end{pmatrix}\begin{pmatrix} 2 \\ 0 \\ 11 \end{pmatrix} = \begin{pmatrix} 19 \\ 11 \end{pmatrix}$$

A^Tb

résolvons

$$\underbrace{\begin{pmatrix} 17 & 1 \\ 1 & 5 \end{pmatrix}}_{B}\underbrace{\begin{pmatrix} \hat{x}_1 \\ \hat{x}_2 \end{pmatrix}}_{\hat{x}} = \begin{pmatrix} 19 \\ 11 \end{pmatrix}$$

B inversible $\qquad B^{-1} = \frac{1}{84}\begin{pmatrix} 5 & -1 \\ -1 & 17 \end{pmatrix}$

$$\Rightarrow \hat{x} = B^{-1}\begin{pmatrix} 19 \\ 11 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \end{pmatrix}.$$

(long) calculs

On a une <u>unique</u> solution au sens des <□

car $B = A^TA$ est inversible

mais ce n'est pas tjrs le cas → ★

écart quadratique : $\|A\hat{x}-b\| =$

$$= \left\| \begin{pmatrix} 4 & 0 \\ 0 & 2 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 2 \end{pmatrix} - \begin{pmatrix} 2 \\ 0 \\ 11 \end{pmatrix} \right\| = \left\| \begin{pmatrix} 4 \\ 4 \\ 3 \end{pmatrix} - \begin{pmatrix} 2 \\ 0 \\ 11 \end{pmatrix} \right\|$$

$$= \left\| \begin{pmatrix} 2 \\ 4 \\ -8 \end{pmatrix} \right\| = \sqrt{84} = 2\sqrt{21}.$$

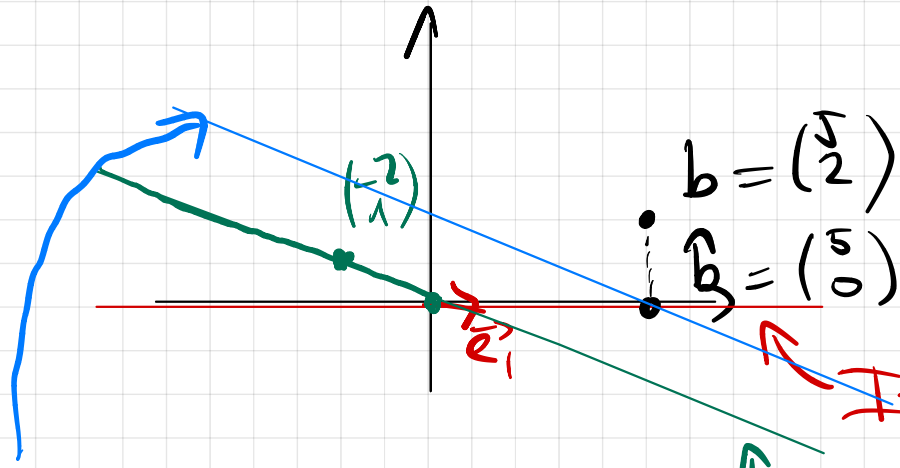
ex 6.7.6

* en effet il peut y avoir une infinité de solutions au sens des < □.

Ex 6.7.7 : $A = \begin{pmatrix} 1 & 2 \\ 0 & 0 \end{pmatrix}$ $\qquad b = \begin{pmatrix} 5 \\ 2 \end{pmatrix}$

A est une projection oblique de vecteur $\begin{pmatrix} -2 \\ 1 \end{pmatrix}$

$A^2 = A$.



$\mathrm{Im}(A) = \mathrm{Vect}\{\vec{e}_1\}$

$\mathrm{Ker}(A) = \mathrm{Vect}\left\{\begin{pmatrix} -2 \\ 1 \end{pmatrix}\right\}$

$$S = \{\hat{x} \in \mathbb{R}^2 \mid A\hat{x} = \hat{b}\}$$

$$= \left\{\hat{x} \in \mathbb{R}^2 \mid A\hat{x} = \begin{pmatrix} 5 \\ 0 \end{pmatrix}\right\}$$

$$= \left\{ \begin{pmatrix} 5 \\ 0 \end{pmatrix} + \underbrace{\lambda \begin{pmatrix} -2 \\ 1 \end{pmatrix}}_{\in \mathrm{Ker}(A) = \text{ens. des sol. syst } Ax=0} \;\middle|\; \lambda \in \mathbb{R} \right\}$$

Ex 6.7.8 :

$$A = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \end{pmatrix}_{4\times 3} \qquad b = \begin{pmatrix} 1 \\ 2 \\ 3 \\ 4 \end{pmatrix}$$

$b \notin \mathrm{Im}(A)$

Recette :

$$A^T A \overset{\text{exo}}{=} \begin{pmatrix} 4 & 2 & 2 \\ 2 & 2 & 0 \\ 2 & 0 & 2 \end{pmatrix} \qquad A^T b = \begin{pmatrix} 10 \\ 3 \\ 7 \end{pmatrix}$$

résoudre $A^T A \hat{x} = A^T b$:
$$\begin{pmatrix} 4 & 2 & 2 \\ 2 & 2 & 0 \\ 2 & 0 & 2 \end{pmatrix} \begin{pmatrix} \hat{x}_1 \\ \hat{x}_2 \\ \hat{x}_3 \end{pmatrix} = \begin{pmatrix} 10 \\ 3 \\ 7 \end{pmatrix}$$

$$\overset{\text{exo}}{\leadsto} \quad S = \left\{ \underbrace{\begin{pmatrix} 7/2 \\ -2 \\ 0 \end{pmatrix}}_{\text{solution particulière de } A^T A \hat{x} = A^T b} + t \begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix} \,\middle|\, t \in \mathbb{R} \right\}$$

solution du syst. homogène $A^T A \hat{x} = 0$ càd $\mathrm{Ker}(A^T A)$

(ex série 12 :
Mq $\mathrm{Ker}(A) = \mathrm{Ker}(A^T A)$.)

$\underline{\text{Thm 6.7.9}}$: Soit $A \in M_{\underline{m \times n}}(\mathbb{R})$

conditions equivalentes :

1) $\forall b \in \mathbb{R}^n$ l'equation $Ax = b$

possède une unique solution au sens des moindres carrés

2) les colonnes de A sont lin. indépendantes (càd A est injective)

3) $A^T A$ est inversible (NB : $A^T A \in M_{n \times n}(\mathbb{R})$ carrée.

Si 1, 2, 3 sont vraies, alors

$$\hat{x} = (A^T A)^{-1} A^T b$$

éq. normale $(A^T A \hat{x} = A^T b)$

est l'unique solution au sens des < □

$$(A^T A)^{-1} \neq A^{-1} A^{T^{-1}}$$

car A n'est pas forcément carrée.

$A, B \in M_{n \times n}(\mathbb{R})$
AB est inversible
$\Updownarrow$
A et B sont inversibles

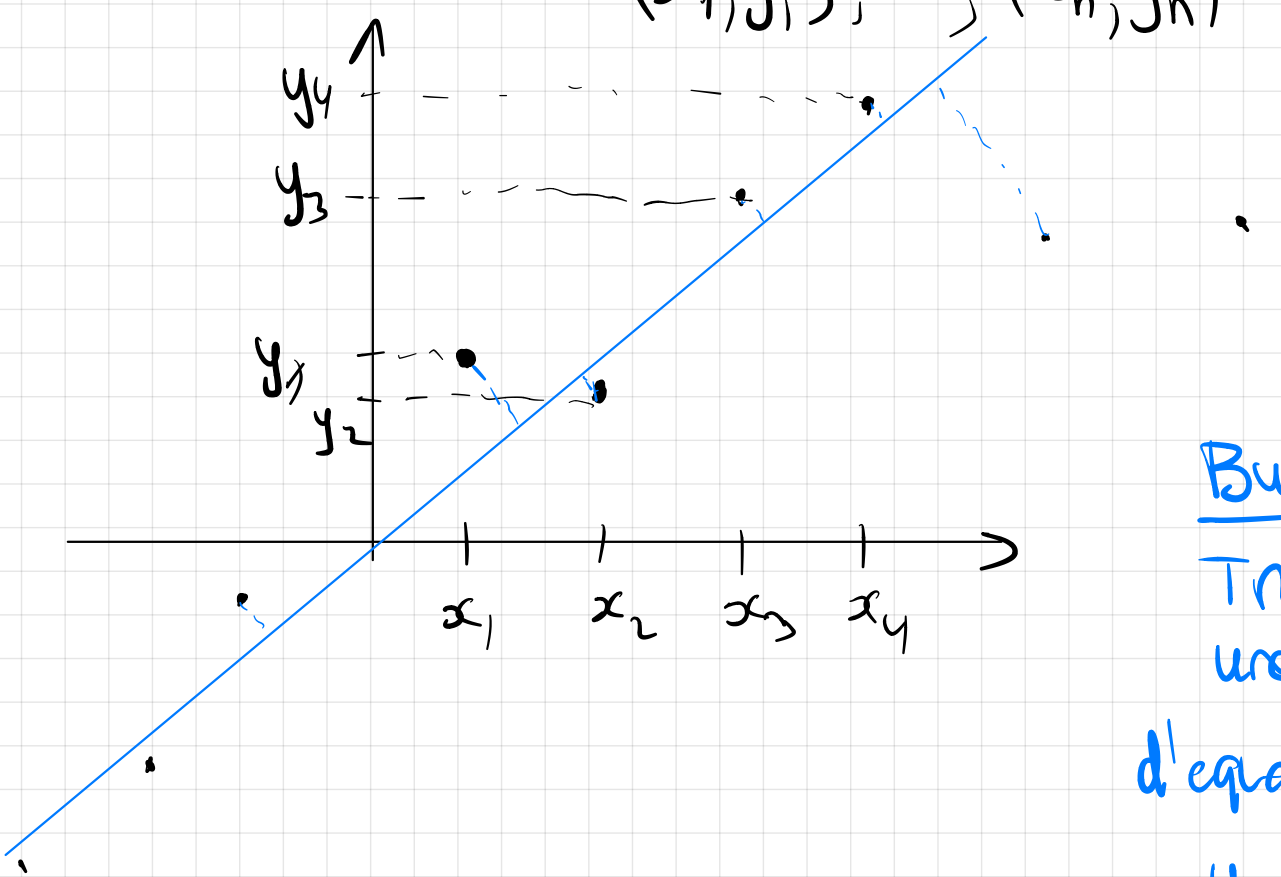
§6.8 Application(s) à des modèles linéaires

Droite de régression linéaire (important)

Discussion 6.8.1 : Donné : un "nuage" de points dans $\mathbb{R}^2$

$(x_1, y_1), \dots, (x_n, y_n)$

$x_i \neq x_j$



But : Trouver une droite d'equation $y = ax + b$ qui soit le plus près possible des points (x_i, y_i)

Fixons a, b pour l'instant.

y_i = la valeur observée en x_i

$ax_i + b$ = la valeur prédite (par la droite $ax + b$)

$r_i = y_i - (ax_i + b)$ = le ième résidu/écart.

Déf 6.8.2 : On appelle <u>droite des moindres □</u>
<u>ou droite de régression</u>

(des points (x_i, y_i) $i=1,\dots,n$)

la droite d'équation $y = ax+b$

qui minimise $\sum_{i=1}^{n} r_i^2$ (la somme des □ des résidus)

les coeff. a et b s'appellent

les coeff. de <u>régression linéaire</u>.

Construction 6.8.3

Valeur prédite	Val observée
ax_1+b	y_1
$\vdots$	$\vdots$
ax_n+b	y_n

si les (x_i, y_i) sont sur $y = ax+b$

alors le système linéaire suivant

est compatible :

$$\underbrace{\begin{pmatrix} x_1 & 1 \\ x_2 & 1 \\ \vdots & \vdots \\ x_n & 1 \end{pmatrix}}_{A} \; n\times 2 \begin{pmatrix} a \\ b \end{pmatrix} = \underbrace{\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}}_{y}$$

← vecteur d'observation

points alignés sur $ax+b$ $\Longleftrightarrow$ $A\begin{pmatrix} a \\ b \end{pmatrix} = y$ est compatible

(6.8.3)

Si les points ne sont pas alignés

on cherche $\begin{pmatrix} a \\ b \end{pmatrix} \in \mathbb{R}^2$ t.q $\left\| A\begin{pmatrix} a \\ b \end{pmatrix} - y \right\|$ soit minimale

car $$\left\| \begin{pmatrix} ax_1+b \\ \vdots \\ ax_n+b \end{pmatrix} - \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \right\|^2 = (\overbrace{ax_1+b-y_1}^{r_1})^2 + \cdots + (\underbrace{ax_n+b-y_n}_{r_n})^2$$

$$= r_1^2 + \cdots + r_n^2$$

Prop 6.8.4 : Calculer les solution(s) de $A\binom{a}{b} = y$ au sens des moindres $\square$ équivaut à trouver la droite de régression $y = ax + b$

Ex 6.8.5 : Trouver droite de régression du nuage

$(2,1), (5,2), (7,3), (8,3)$

$$A = \begin{pmatrix} 2 & 1 \\ 5 & 1 \\ 7 & 1 \\ 8 & 1 \end{pmatrix} \qquad y = \begin{pmatrix} 1 \\ 2 \\ 3 \\ 3 \end{pmatrix}$$

$$\begin{pmatrix} x_i^1 & x_i^0 \end{pmatrix}$$

on résoud $A\binom{a}{b} = y$ au sens des $< \square$

grâce à l'éq. normale :

$$A^T A \binom{a}{b} = A^T y$$

calculs : $A^T A = \begin{pmatrix} 142 & 22 \\ 22 & 4 \end{pmatrix}$ $\quad A^T y = \begin{pmatrix} 57 \\ 9 \end{pmatrix}$

exo

inversible

$$\begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} 142 & 22 \\ 22 & 4 \end{pmatrix}^{-1} \begin{pmatrix} 57 \\ 9 \end{pmatrix}$$

$$= \frac{1}{84} \begin{pmatrix} 4 & -22 \\ -22 & 142 \end{pmatrix} \begin{pmatrix} 57 \\ 9 \end{pmatrix}$$

$$= \frac{1}{84} \begin{pmatrix} 30 \\ 24 \end{pmatrix} = \begin{pmatrix} 5/14 \\ 2/7 \end{pmatrix}$$

$$y = \frac{5}{14}x + \frac{2}{7} \iff 5x - 14y + 4 = 0$$

Modèle linéaire général 6.8.6

Approximer un nuage de points
par un polynôme de degré $d \geq 1$
(on a fait $d = 1$)

<u>ex avec $d=2$</u> : on cherche $y = ax^2 + bx + c$
qui approxime au mieux
le nuage.

Recette :

Poser $A = \begin{pmatrix} x_1^2 & x_1 & 1 \\ x_2^2 & x_2 & 1 \\ \vdots & \vdots & \vdots \\ x_n^2 & x_n & 1 \end{pmatrix}_{n\times 3}$ $\qquad y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$

$$\begin{pmatrix} x_1^2 \\ \vdots \\ x_n^2 \end{pmatrix} \quad \begin{pmatrix} x_1^1 \\ \vdots \\ x_n^1 \end{pmatrix} \quad \begin{pmatrix} x_1^0 \\ \vdots \\ x_n^0 \end{pmatrix}$$

et on cherche à résoudre $A\begin{pmatrix} a \\ b \\ c \end{pmatrix} = y$

au sens des moindres $\square$

Ici : $\begin{pmatrix} 0^2 & 0 & 1 \\ 1^2 & 1 & 1 \\ 2^2 & 2 & 1 \\ 3^2 & 3 & 1 \end{pmatrix}\begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} -1 \\ 1 \\ 0 \\ 1 \end{pmatrix}$

longs calculs $\rightsquigarrow$ $\qquad \hat{x} = \begin{pmatrix} -1/4 \\ 5/4 \\ -3/4 \end{pmatrix}$

$A^T A \hat{x} = A^T y$

$$y = -\frac{x^2}{4} + \frac{5}{4}x - \frac{3}{4}$$

Fin 6.6.7.