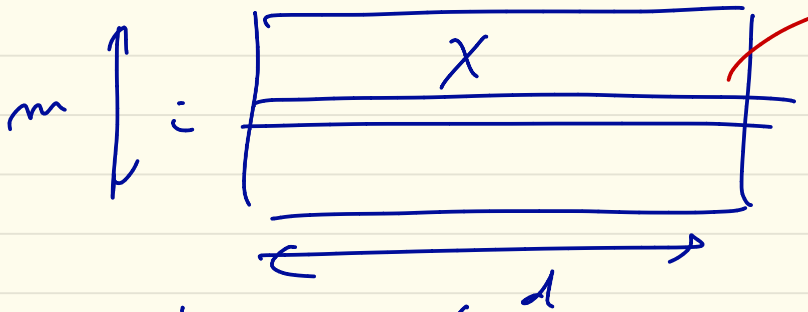
TD 4 – Projection, orthogonalité

Exo 1. Cadre matriciel $M(n,\mathbb{R}) = M(n,n,\mathbb{R})$
$\simeq \mathcal{L}(\mathbb{R}^n, \mathbb{R}^n)$,

$$A = \begin{array}{cc} & j \\ i & \begin{bmatrix} & \vdots & \\ \cdots & a_{ij} & \\ & & \end{bmatrix} \end{array} \; \updownarrow n$$
$$\overleftrightarrow{\quad n \quad}$$

Applications : ML (= Machine Learning)

"données structurées"



← i ème donnée
= vecteur ligne de $\mathbb{R}^d$

ex. : biologie / médecine, cohorte de taille
$n << d$; d = nombre de gènes ($\simeq 10^4 \ldots$)
réduction de dimension :
W tq $X \longrightarrow X \cdot W : n \times d'$,
$d' << d \ldots$

$\dim_{\mathbb{R}} M(n,\mathbb{R}) = n^2$

Base canonique de $M(n,\mathbb{R})$:

$$E_{ij} = \begin{array}{cc} & j \\ i & \begin{bmatrix} 0 & \vdots & 0 \\ \cdots 0 & 1 & 0 \end{bmatrix} \end{array}$$

$$A = (a_{ij})_{i,j=1,n} = \sum_i \sum_j a_{ij} . E_{ij}$$

$(X|Y) = \mathrm{tr}({}^tX.Y)$ ← prod. scalaire Frobenius

: mq on définit ainsi un produit scalaire sur E ;

i) <u>bilinéarité</u> :

$$(X,Y) \longmapsto ({}^tX, Y) \longmapsto {}^tX.Y \overset{\mathrm{tr}}{\longmapsto} \mathrm{tr}({}^tX.Y)$$

$\left({}^t(\lambda A + B) = \lambda {}^tA + {}^tB\right.$: transposition linéaire$\left.\right)$

$$\left(\mathrm{tr}(\lambda A + B) = \mathrm{tr}\left((\lambda a_{ij} + b_{ij})_{i,j}\right)\right.$$
$$= \sum_i \lambda a_{ii} + b_{ii}$$
$$= \lambda\, \mathrm{tr}(A) + \mathrm{tr}(B) : \text{linéaire}\left.\right)$$

D'où la linéarité par rapport à X et Y de cette application.

ii) <u>symétrie</u> : $(Y|X) = \mathrm{tr}({}^tY.X) \overset{?}{=} \mathrm{tr}({}^tX.Y)$

$$= \mathrm{tr}({}^t({}^tY.X))$$
$$(\text{cf. } \mathrm{tr}({}^tA) = \mathrm{tr}A)$$
$$= \mathrm{tr}({}^tX.Y)$$
$$= (X|Y).$$

iii) <u>définie positivité</u> :

$$\mathrm{tr}({}^tX.Y) = \sum_{i=1}^{m} z_{ii} \quad \text{où } Z = (z_{ij})_{i,j} = {}^tX.Y$$
$$= \sum_{i=1}^{m} \sum_{k=1}^{m} ({}^tX)_{i,k} \cdot (Y)_{k,i} \quad (\text{lico})$$
$$= \sum_{i=1}^{m} \sum_{k=1}^{m} x_{ki} \cdot y_{ki}$$
$$= \sum_{i,j=1}^{m} x_{ij} \cdot y_{ij}$$

$$= \left(X(:) \,|\, Y(:) \right)_{\mathbb{R}^{n^2}}$$

où $X(:) := \begin{bmatrix} | \\ x_1 \\ | \\ \hline \vdots \\ \hline | \\ x_n \\ | \end{bmatrix}$ (n^2 lignes, chaque bloc de taille n) avec $X = \begin{bmatrix} | & & | \\ x_1 & \dots & x_n \\ | & & | \end{bmatrix}$

($x_j = j$-ème colonne de X)

Donc $\mathrm{tr}({}^tX.X) = 0 \Rightarrow \sum_{i,j} \overbrace{x_{i,j}^2}^{\geq 0} = 0$

$\Rightarrow (\forall i,j = 1,n) : x_{ij} = 0$

$\Rightarrow X = 0_{M(n,\mathbb{R})}$: défini ;

$\mathrm{tr}({}^tX.X) = \sum_{i,j} x_{i,j}^2 \geq 0$: positivité.

Remarque : $\mathrm{tr}({}^tX.X)$: ${}^tX.X$ sym. $\underbrace{\geq 0}_{\text{semi-déf. positive}}$

${}^t({}^tX.X) = {}^tX.{}^t({}^tX) = {}^tX.X$: sym.

On rappelle $A \in \mathrm{Sym}(n,\mathbb{R})$ est dite :

i) $A \geq 0$ (" semi-définie positive") :

$a \in \mathbb{R}$, $a \geq 0 \Leftrightarrow (\forall x \in \mathbb{R}) : a.x^2 \geq 0$

$$A \geq 0 \Leftrightarrow (\forall x \in \mathbb{R}^n) : (Ax \,|\, x)_{\mathbb{R}^n} \geq 0$$

ii) $A > 0$ ("définie positive") :

$a \in \mathbb{R}$, $a > 0 \Leftrightarrow (\forall x \neq 0) : a x^2 > 0$

$A > 0 \Leftrightarrow (\forall x \in \mathbb{R}^n, x \neq 0) : (Ax \mid x)_{\mathbb{R}^n} > 0.$

Ici, soit $x \in \mathbb{R}^n$,

$$({}^t X X . x \mid x) = (X.x \mid X.x)$$
$$= \|X.x\|^2 \geq 0 : {}^t X X \geq 0.$$

De plus, ${}^t XX$ étant symétrique réelle, elle se diagonalise sur $\mathbb{R}$ (dans une BON = Base Orthonormée de vp = vecteurs propres) ; on sait également

$$\begin{cases} A \geq 0 \Leftrightarrow \text{toutes les vp de } A \text{ sont } \geq 0 \\ A > 0 \Leftrightarrow \text{toutes les vp de } A \text{ sont } > 0 \end{cases}$$

Ici, $\operatorname{tr}({}^t XX) = \sum_{i=1}^{n} \lambda_i$, $\lambda_1 \to \lambda_n$ vp de ${}^t XX$

${}^t XX \geq 0 \Rightarrow \lambda_i \geq 0 \Rightarrow \operatorname{tr}({}^t XX) \geq 0.$

Si jamais $\operatorname{tr}({}^t X.X) = 0 \Rightarrow \lambda_1 = \dots = \lambda_n = 0$
$\Rightarrow {}^t XX = 0$
$\Rightarrow X = 0$.

1.2. Soient X et $Y \in M(n, \mathbb{R})$,

(a) $({}^t X \mid {}^t Y) = \operatorname{tr}({}^t({}^t X).{}^t Y)$
$= \operatorname{tr}(X.{}^t Y)$
$= \operatorname{tr}({}^t Y.X) = (Y \mid X) = (X \mid Y).$

Cf. $\operatorname{tr}(B.A) = \operatorname{tr}(A.B)$

Rappel : deux matrices semblables ont même trace (cf. la trace est l'un des coeffs du polynôme caractéristique :

$$\begin{aligned}\chi_A(X) &= \det(\lambda I - A) \quad \text{mat. semblable à } A \\ &= \det(\lambda I - (PAP^{-1})) \quad (P \text{ inv.}) \\ &= \det(P(\lambda I - A)P^{-1}) \\ &= \cancel{\det}(P) \cdot \chi_A(X) \cdot \underbrace{\det(P^{-1})}_{(\cancel{\det} P)^{-1}} \quad)\end{aligned}$$

Or, si B est inversible, AB et BA sont semblables puisque :

$$AB = B^{-1} \cdot (BA) \cdot B$$

Comme toute matrice $B \in M(n, \mathbb{R})$ est limite d'une suite de matrices inversibles (cf. $\overline{GL(n, \mathbb{R})} = M(n, \mathbb{R})$), on a :

$$\begin{aligned}\operatorname{tr}(B.A) &= \operatorname{tr}(\overbrace{\lim_n B_n}^{B} . A) \quad B_n \in GL(n, \mathbb{R}) \\ &= \lim_n \underbrace{\operatorname{tr}(B_n . A)}_{\operatorname{tr}(A . B_n)} \quad (\operatorname{tr} \text{ continue}) \\ &= \operatorname{tr}(A.B) \quad (\text{continuité bis}).\end{aligned}$$

(2) Soit $A \in M(n, \mathbb{R})$, mq

$$(AX|Y) = (X|{}^tA.Y) \quad ?$$

On a $(AX|Y) = \mathrm{tr}({}^t(AX).Y)$
$= \mathrm{tr}(({}^tX\,{}^tA).Y)$
$= \mathrm{tr}({}^tX.({}^tA.Y))$ (associativité)
$= (X \mid {}^tA.Y)$

③ Soit $O \in O(n) = \{A \in M(n,\mathbb{R}) \mid {}^tA.A = A.{}^tA = I\}$

$\cap$ $GL(n,\mathbb{R})$ (ie $A^{-1} = {}^tA$)

groupe orthogonal

Mq $X \overset{\varphi}{\longmapsto} O.X$ isométrie de $(M(n,\mathbb{R}), (.|.))$

Rappel: une isométrie est une application qui préserve la norme: on doit mq

$$\underset{M(n,\mathbb{R})}{\|O.X\|} = \|X\|.$$

On sait que c'est en fait équivalent à préserver le produit scalaire (ie ici à :

$(O.X \mid O.Y) = (X \mid Y)$); en effet, le produit scalaire d'un e.v. s'exprime en fonction de la norme ("polarisation"):

$x, y \in \underbrace{(E, (.|.))}_{\text{e.v.}}$,

$$(x|y) = \frac{\overbrace{\|x+y\|^2}^{\|x\|^2 + 2(x|y) + \|y\|^2} - \|x-y\|^2}{4}.$$

Ici, soit $X \in M(n, \mathbb{R})$,

$$\|OX\|^2 = (OX | OX)$$
$$= \operatorname{tr}\left({}^t(OX) \cdot OX\right)$$
$$= \operatorname{tr}\left({}^tX \cdot \underbrace{({}^tO \cdot O)}_{I} \cdot X\right)$$
$$= \|X\|^2.$$

Exo 2. 1) Déterminer le projeté orthogonal de $x_0 : [0, 2\pi] \longrightarrow \mathbb{R}$, $t \longmapsto x_0(t)$ $\in L^2([0, 2\pi])$

(cf. x_0 continue sur le compact $[0, 2\pi]$, donc bornée : $x_0 \in L^\infty(\underbrace{[0, 2\pi]}_{\mu < \infty}) \subset L^2([0, 2\pi])$:

$$\int_0^{2\pi} \underbrace{|x(t)|}_{\leq M}{}^2 \, dt \leq 2\pi \cdot M^2 < \infty \quad)$$

sur $F = \operatorname{Vect}\{\sin, \cos\} \subset L^2([0, 2\pi])$.
↑
sev
(= sous espace vectoriel)

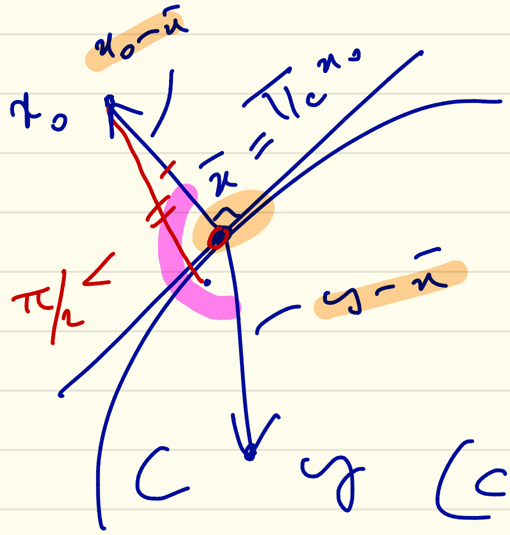
Ie : trouver $\bar{x} \in F$ tq

$$\|x_0 - \bar{x}\|_{L^2} = \inf_{y \in F} \|x_0 - y\|.$$
$$=: d(x_0, F).$$

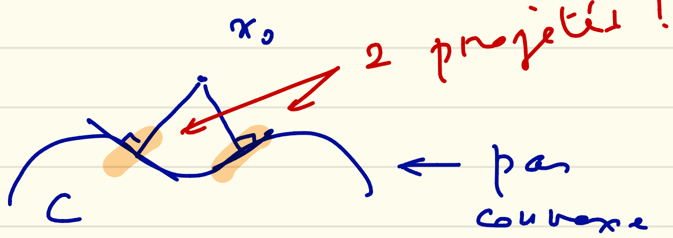
On est donc en train de résoudre le problème d'approximation (au sens L^2) suivant :

$$\begin{cases} \sqrt{\displaystyle\int_0^{2\pi} \Big|\underbrace{e^t}_{x_0(t)} - \underbrace{(a\sin t + b\cos t)}_{\in F}\Big|^2} \to \min \\ (a,b) \in \mathbb{R}^2 \quad \leftarrow \text{2 inconnues} \end{cases}$$

Dans l'e.h. $(L^2([0,2\pi]), \langle \cdot | \cdot \rangle)$, il suffit que les hypothèses du th. de la projection soient remplies pour garantir l'existence et l'unicité de ce projeté (qui est dans ce cas le "projeté orthogonal $\Pi_F x_0$" de x_0 sur F).

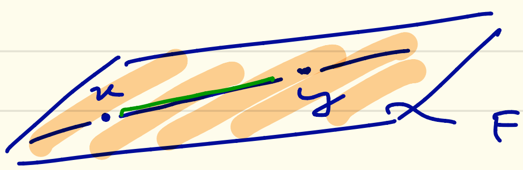


(C doit être convexe, fermé $\neq \emptyset$)



Ici, F est un sev, $F = \mathrm{Vect}\{\sin, \cos\}$:

- F sev $\Rightarrow 0 \in F \Rightarrow F \neq \emptyset$
- F sev $\Rightarrow F$ convexe, cf. une combinaison convexe est une combinaison linéaire !



Ie : Soient $x, y \in F$, il faut que $[x, y] \subset F$,

ie que $(\forall \lambda \in [0,1]) : \underbrace{(1-\lambda) x + \lambda y}_{\text{CL}} \in F$

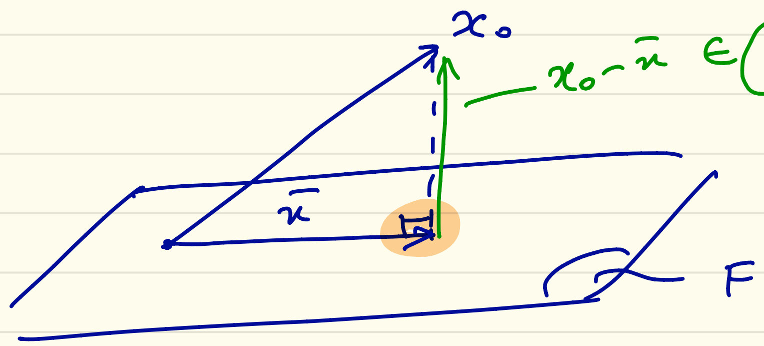
$\in F$

de vecteurs de F, donc $\in F$!

– F sev de dim finie $\Longrightarrow$ F fermée
(cf. $\dim_{\mathbb{R}} F \leq 2$) (cf. aussi exo 3)

Le th. de la projection s'applique : $\exists ! \, \bar{x} \in F$ caractérisé par l'inéquation variationnelle suivante :

$$(\forall y \in F) : \quad (x_0 - \bar{x} \mid y - \bar{x}) \leq 0 \qquad (*)$$



$x_0 - \bar{x} \in (F^{\perp})$: $(x_0 - \bar{x} \mid y) = 0$, $\forall y \in F$

cf. § III. 3

Soit $y \in F$; $\bar{x} \in F$, $y + \bar{x} \in F$

$$\underset{(*)}{\Longrightarrow} \quad (x_0 - \bar{x} \mid \underbrace{(y + \cancel{\bar{x}}) - \cancel{\bar{x}}}_{y}) \leq 0 \qquad (**)$$

De plus, F sev $\Longrightarrow$ si $y \in F$, $-y \in F$,

$(**)$ $\quad (x_0 - \bar{x} \mid (-y)) \leq 0$

$\Longrightarrow (x_0 - \bar{x} \mid y) \geq 0$: $(x_0 - \bar{x} \mid y) = 0$

Donc, $(x_0 - \bar{x} \mid y) = 0$, $\forall y \in F$

$$\Longleftrightarrow \begin{cases} (x_0 - \bar{x} \mid \sin) = 0 \\ (x_0 - \bar{x} \mid \cos) = 0 \end{cases} \quad \text{2 équations}$$

2 inconnues

De plus, $\bar{x} \in F \Rightarrow (\exists (a, b) \in \mathbb{R}^2)$:

$$\bar{x} = a.\sin + b.\cos$$

$$\Longleftrightarrow \begin{cases} (x_0 - (a.\sin + b.\cos) \mid \sin) = 0 \\ (x_0 - (a.\sin + b.\cos) \mid \cos) = 0 \end{cases} \quad \text{SL } 2 \times 2$$

$$\Longleftrightarrow \quad G.\begin{bmatrix} a \\ b \end{bmatrix} = c \in \mathbb{R}^2 \quad \text{avec}$$

$$G = \begin{bmatrix} (\sin \mid \sin) & (\cos \mid \sin) \\ (\sin \mid \cos) & (\cos \mid \cos) \end{bmatrix}$$

matrice de Gram de {sin, cos}

$$c = \begin{bmatrix} (x_0 \mid \sin) \\ (x_0 \mid \cos) \end{bmatrix}, \quad x_0 = e^t$$

$$(\sin \mid \sin) = \int_0^{2\pi} \sin^2(t)\, dt = \ldots$$

$$(e^t \mid \sin) = \int_0^{2\pi} e^t . \sin t\, dt = \ldots$$

ipp ou :

$$\operatorname{Im}\left(\int_0^{2\pi} e^t . e^{it}\right) = \operatorname{Im} \int_0^{2\pi} e^{(1+i)t}\, dt$$