

Devoir (en classe) n°11 CORRECTION

Problème (d'après CCP-PC (maths 1) 2011)

I. Droites des moindres carrées dans un espace euclidien

1. $F^\perp = \{\mathbf{x} \in E \mid \forall \mathbf{y} \in F, \langle \mathbf{x} | \mathbf{y} \rangle = 0\}$. On a toujours $F \oplus F^\perp$

Dans le cas où F est de dimension finie, ce qui est le cas si E est de dimension finie : $E = F \oplus F^\perp$

Pour $\mathbf{x} \in E$, on note $p_F(\mathbf{x})$ la projection orthogonale de $\mathbf{x}$ sur F .

2. $\{\|\mathbf{x} - \mathbf{z}\|, \mathbf{z} \in F\}$ est non vide, minorée par 0 donc admet un plus petit élément.

Notons $\mathbf{z} = p_F(\mathbf{x})$.

Pour tout $\mathbf{y} \in F$, $\|\mathbf{x} - \mathbf{y}\|^2 = \|(\mathbf{x} - \mathbf{z}) + (\mathbf{z} - \mathbf{y})\|^2 = \|\mathbf{x} - \mathbf{z}\|^2 + \|\mathbf{z} - \mathbf{y}\|^2$ (formule de Pythagore),
car $\mathbf{x} - \mathbf{z} \in F^\perp$ et $\mathbf{z} - \mathbf{y} \in F$, puisque $\mathbf{y}$ et $\mathbf{z} = p_F(\mathbf{x}) \in F$.

On a alors $\forall \mathbf{y} \in F$, $\|\mathbf{x} - \mathbf{y}\|^2 \geq \|\mathbf{x} - \mathbf{z}\|^2$, donc $\|\mathbf{x} - \mathbf{y}\| \geq \|\mathbf{x} - \mathbf{z}\|$.

Ainsi, $\|\mathbf{x} - \mathbf{z}\|$ est le minimum des $\|\mathbf{x} - \mathbf{y}\|$, $\mathbf{y} \in F$, minimum qui est donc atteint en $\mathbf{z} = p_F(\mathbf{x})$.

Il est unique, sinon, si $\mathbf{z}'$ était une autre solution alors avec $\mathbf{y} = \mathbf{z}'$,

on aurait $\|\mathbf{z}' - \mathbf{z}\|^2 = \|\mathbf{x} - \mathbf{z}'\|^2 - \|\mathbf{x} - \mathbf{z}\|^2 = 0$, et donc $\mathbf{z}' = \mathbf{z}$ car la norme est définie.

3. (a) Soit $(\mathbf{x}, \mathbf{y}) \mapsto \langle \mathbf{x} | \mathbf{y} \rangle_F$ un produit subordonné à F , alors

$$\begin{aligned} - \forall \mathbf{x}, \mathbf{y} \in E, \langle \mathbf{x}, \mathbf{y} \rangle_F &= \langle \mathbf{x}, \mathbf{y} - p_F(\mathbf{y}) + p_F(\mathbf{y}) \rangle_F = \langle \mathbf{x}, \mathbf{y} - p_F(\mathbf{y}) \rangle_F + \langle \mathbf{x}, p_F(\mathbf{y}) \rangle_F \text{ (linéarité)} \\ &= \langle \mathbf{x}, \mathbf{y} - p_F(\mathbf{y}) \rangle_F \text{ car } \langle \mathbf{x}, p_F(\mathbf{y}) \rangle_F = 0 \text{ puisque } p_F(\mathbf{y}) \in F \text{ par (iii)}. \\ &= \langle \mathbf{y} - p_F(\mathbf{y}), \mathbf{x} \rangle_F \text{ d'après (ii)} \\ &= \langle \mathbf{y} - p_F(\mathbf{y}), \mathbf{x} - p_F(\mathbf{x}) \rangle_F + \langle \mathbf{y} - p_F(\mathbf{y}), p_F(\mathbf{x}) \rangle_F \text{ d'après (i)} \\ &= \langle \mathbf{y} - p_F(\mathbf{y}), \mathbf{x} - p_F(\mathbf{x}) \rangle_F \text{ car } \langle \mathbf{y} - p_F(\mathbf{y}), p_F(\mathbf{x}) \rangle_F = 0 \text{ d'après (iii)} \end{aligned}$$

Enfin, par symétrie $\forall \mathbf{x}, \mathbf{y} \in E$, $\langle \mathbf{x}, \mathbf{y} \rangle_F = \langle (\mathbf{x} - p_F(\mathbf{x}) | \mathbf{y} - p_F(\mathbf{y})) \rangle$;

$$- \forall \mathbf{x} \in E, \langle \mathbf{x} | \mathbf{x} \rangle_F = \langle \mathbf{x} - p_F(\mathbf{x}) | \mathbf{x} - p_F(\mathbf{x}) \rangle_F = \|\mathbf{x} - p_F(\mathbf{x})\|^2 = (d(\mathbf{x}, F))^2;$$

$$- \text{Le carré de la distance étant positif, } \forall \mathbf{x} \in E, \langle \mathbf{x} | \mathbf{x} \rangle_F \geq 0;$$

$$- \forall \mathbf{x} \in E, \langle \mathbf{x} | \mathbf{x} \rangle_F = 0 \iff d(\mathbf{x}, F) = 0 \iff \mathbf{x} = p_F(\mathbf{x}) \iff \mathbf{x} \in F;$$

- (b) Pour tout $\mathbf{x}, \mathbf{y} \in E$, $\langle \mathbf{x}, \mathbf{y} \rangle_F = \langle \mathbf{x} - p_F(\mathbf{x}), \mathbf{y} - p_F(\mathbf{y}) \rangle$, ce qui assure une définition unique du produit subordonné à F , une fois F (et donc p_F) et le produit scalaire donné.

4. Pour tout $(\mathbf{x}, \mathbf{y}) \in E^2$, on a $\langle \mathbf{x} | \mathbf{y} \rangle_F = \langle \mathbf{x} - p_F(\mathbf{x}) | \mathbf{y} - p_F(\mathbf{y}) \rangle$.

Puis d'après l'inégalité de Cauchy-Schwarz : $|\langle (\mathbf{x} - p_F(\mathbf{x}) | \mathbf{y} - p_F(\mathbf{y})) \rangle| \leq \|\mathbf{x} - p_F(\mathbf{x})\| \cdot \|\mathbf{y} - p_F(\mathbf{y})\|$.

Or $\|\mathbf{x}\|_F = \sqrt{\langle \mathbf{x} | \mathbf{x} \rangle_F} = \sqrt{\langle \mathbf{x} - p_F(\mathbf{x}) | \mathbf{x} - p_F(\mathbf{x}) \rangle} = \|\mathbf{x} - p_F(\mathbf{x})\|$, de même $\|\mathbf{y}\|_F = \|\mathbf{y} - p_F(\mathbf{y})\|$.

Ainsi pour tout $(\mathbf{x}, \mathbf{y}) \in E^2$, on a $|\langle \mathbf{x} | \mathbf{y} \rangle_F| \leq \|\mathbf{x}\|_F \cdot \|\mathbf{y}\|_F$

On a l'égalité ssi l'inégalité de Cauchy-Schwarz est une égalité,

ssi $\mathbf{x} - p_F(\mathbf{x})$ et $\mathbf{y} - p_F(\mathbf{y})$ sont colinéaires

ssi $\exists (\lambda, \mu) \neq (0, 0)$ tels que $\lambda(\mathbf{x} - p_F(\mathbf{x})) + \mu(\mathbf{y} - p_F(\mathbf{y})) = 0$.

ssi $\exists (\lambda, \mu) \neq (0, 0)$ tels que $\lambda\mathbf{x} + \mu\mathbf{y} = \lambda p_F(\mathbf{x}) + \mu p_F(\mathbf{y})$.

Cela impose qu'il existe $(\lambda, \mu) \neq (0, 0)$ tels que $\lambda\mathbf{x} + \mu\mathbf{y} \in F$.

Réciproquement, si il existe $(\lambda, \mu) \neq (0, 0)$ tels que $\lambda\mathbf{x} + \mu\mathbf{y} \in F$, alors $p_F(\lambda\mathbf{x} + \mu\mathbf{y}) = \lambda\mathbf{x} + \mu\mathbf{y}$.

par linéarité de p_F , on a donc : $\lambda(\mathbf{x} - p_F(\mathbf{x})) + \mu(\mathbf{y} - p_F(\mathbf{y})) = 0$

et donc $\mathbf{x} - p_F(\mathbf{x})$ et $\mathbf{y} - p_F(\mathbf{y})$ sont colinéaires.

On a donc $|\langle \mathbf{x} | \mathbf{y} \rangle_F| = \|\mathbf{x}\|_F \cdot \|\mathbf{y}\|_F$ ssi il existe $(\lambda, \mu) \neq (0, 0)$ tels que $\lambda\mathbf{x} + \mu\mathbf{y} \in F$

5. (a) E n'est pas réduit à 0, si $\mathbf{a} (\neq 0) \in E$, alors $\mathbf{u} = \frac{1}{\|\mathbf{a}\|}\mathbf{a}$ vérifie $\mathbf{u} \in E$ et $\|\mathbf{u}\| = 1$.

- (b) On a la décomposition de $\mathbf{x}$ selon $E = D \oplus D^\perp$ si $\mathbf{x} = a\mathbf{u} + (\mathbf{x} - a\mathbf{u})$ avec $\mathbf{x} - a\mathbf{u} \in D^\perp$.

C'est le cas si et seulement si $\langle (\mathbf{x} - a\mathbf{u}) | \mathbf{u} \rangle = 0$ i.e. $a = \langle \mathbf{x} | \mathbf{u} \rangle$, car $\|\mathbf{u}\|^2 = \langle \mathbf{u} | \mathbf{u} \rangle = 1$.

Et donc pour tout $\mathbf{x} \in E$, $p_D(\mathbf{x}) = a\mathbf{u} = \langle \mathbf{x} | \mathbf{u} \rangle \mathbf{u}$.

- (c) On a pour tout $\mathbf{x}$, $\sigma_{\mathbf{x}} = \|\mathbf{x}\|_D = \|\mathbf{x} - p_D(\mathbf{x})\| = \|\mathbf{x} - m_{\mathbf{x}}\mathbf{u}\|$ d'après 3(a).

Et pour tout $\mathbf{x}, \mathbf{y} \in E$, $\text{cov}(\mathbf{x}, \mathbf{y}) = \langle \mathbf{x} | \mathbf{y} \rangle_D = \langle \mathbf{x} - p_D(\mathbf{x}) | \mathbf{y} - p_D(\mathbf{y}) \rangle$

$\text{cov}(\mathbf{x}, \mathbf{y}) = \langle \mathbf{x} | \mathbf{y} - p_D(\mathbf{y}) \rangle - \langle p_D(\mathbf{x}) | \mathbf{y} - p_D(\mathbf{y}) \rangle = \langle \mathbf{x} | \mathbf{y} - p_D(\mathbf{y}) \rangle : p_D(\mathbf{x}) \in D, \mathbf{y} - p_D(\mathbf{y}) \in D^\perp$

$\text{cov}(\mathbf{x}, \mathbf{y}) = \langle \mathbf{x} | \mathbf{y} \rangle - \langle \mathbf{x} | p_D(\mathbf{y}) \rangle = \langle \mathbf{x} | \mathbf{y} \rangle - \langle \mathbf{x} - p_D(\mathbf{x}) | p_D(\mathbf{y}) \rangle - \langle p_D(\mathbf{x}) | p_D(\mathbf{y}) \rangle$

$\text{cov}(\mathbf{x}, \mathbf{y}) = \langle \mathbf{x} | \mathbf{y} \rangle - \langle p_D(\mathbf{x}) | p_D(\mathbf{y}) \rangle : p_D(\mathbf{y}) \in D, \mathbf{x} - p_D(\mathbf{x}) \in D^\perp$

Enfin, $\langle p_D(\mathbf{x}) | p_D(\mathbf{y}) \rangle = \langle m_{\mathbf{x}}\mathbf{u} | m_{\mathbf{y}}\mathbf{u} \rangle = m_{\mathbf{x}}m_{\mathbf{y}}\|\mathbf{u}\|$.

Et donc pour tout $\mathbf{x}, \mathbf{y} \in E$, $\text{cov}(\mathbf{x}, \mathbf{y}) = \langle \mathbf{x}, \mathbf{y} \rangle - m_{\mathbf{x}}m_{\mathbf{y}}$.

6. $\sigma_{\mathbf{x}} = \|\mathbf{x}\|_D \geq 0$, il est égal à 0 ssi $\mathbf{x} \in D$ d'après 3.(a).

il est donc nul ssi $\mathbf{x}$ et $\mathbf{u}$ sont colinéaires, ce qui est faux. Donc $\sigma_{\mathbf{x}} > 0$ et de même $\sigma_{\mathbf{y}} > 0$

7. (a) $\mathbf{x}^* = \frac{\mathbf{x} - m_{\mathbf{x}}\mathbf{u}}{\sigma_{\mathbf{x}}}$, donc $m_{\mathbf{x}^*} = \langle \mathbf{x}^* | \mathbf{u} \rangle = \frac{1}{\sigma_{\mathbf{x}}} (\langle \mathbf{x} | \mathbf{u} \rangle - m_{\mathbf{x}} \|\mathbf{u}\|) = 0$ (par définition)
 $\sigma_{\mathbf{x}^*} = \|\mathbf{x}^*\|_D = \|\mathbf{x}^* - m_{\mathbf{x}^*}\mathbf{u}\| = \|\mathbf{x}^*\| = \frac{1}{|\sigma_{\mathbf{x}}|} \|\mathbf{x} - m_{\mathbf{x}}\mathbf{u}\| = \frac{1}{\sigma_{\mathbf{x}}} \|\mathbf{x} - p_D(\mathbf{x})\| = \frac{\|\mathbf{x}\|_D}{\sigma_{\mathbf{x}}} = \frac{\sigma_{\mathbf{x}}}{\sigma_{\mathbf{x}}} = 1$.
Enfin, $\rho = \frac{\text{cov}(\mathbf{x}, \mathbf{y})}{\sigma_{\mathbf{x}}\sigma_{\mathbf{y}}} = \frac{\langle \mathbf{x} | \mathbf{y} \rangle_D}{\sigma_{\mathbf{x}}\sigma_{\mathbf{y}}}$ donc $|\rho| < \frac{\|\mathbf{x}\|_D \cdot \|\mathbf{y}\|_D}{\sigma_{\mathbf{x}}\sigma_{\mathbf{y}}} = 1$ d'après 4. ;
l'inégalité est stricte car $(\mathbf{x}, \mathbf{y}, \mathbf{u})$ est libre ; $\forall (\lambda, \mu) \neq (0, 0)$, $\lambda\mathbf{x} + \mu\mathbf{y} \in D \notin \text{vect}(\mathbf{u})$.
Donc $m_{\mathbf{x}^*} = 0$, que $\sigma_{\mathbf{x}^*} = 1$ et que $\rho \in]-1, 1[$.

- (b) On sait que $\|\mathbf{u}\| = 1$ et $\|\mathbf{x}^*\| = 1$ (dans la chaîne d'égalité de $\sigma_{\mathbf{x}^*}$, question précédente).
De plus, on a vu que $\langle \mathbf{u} | \mathbf{x}^* \rangle (= m_{\mathbf{x}^*}) = 0$.
Donc la famille $(\mathbf{u}, \mathbf{x}^*)$ est orthonormée, elle est donc libre et $\dim(\text{vect}(\mathbf{u}, \mathbf{x}^*)) = 2$.
En outre, $\mathbf{u} \in F$ et $\mathbf{x}^* \in F$ aussi car $\mathbf{x}^* = \frac{1}{\sigma_{\mathbf{x}}}\mathbf{x} - \frac{m_{\mathbf{x}}}{\sigma_{\mathbf{x}}}\mathbf{u}$.
Donc $\text{vect}(\mathbf{u}, \mathbf{x}^*) \subset F$, et comme ces deux espaces ont la même dimension 2 : $\text{vect}(\mathbf{u}, \mathbf{x}^*) = F$.
Donc $(\mathbf{u}, \mathbf{x}^*)$ est une base orthonormale de $F = \text{vect}(\mathbf{u}, \mathbf{x})$.

- (c) On applique à $F = \text{vect}(\mathbf{u}, \mathbf{x})$, le résultat de 2. Le minimum est atteint en $\mathbf{z} = p_F(\mathbf{y})$.
Puis on a $\inf_{(a,b) \in \mathbb{R}^2} \|\mathbf{y} - a\mathbf{x} - b\mathbf{u}\| = \|\mathbf{y} - p_F(\mathbf{y})\| = d(\mathbf{y}, F)$

- (d) Or comme $(\mathbf{u}, \mathbf{x}^*)$ est une base orthonormée de F , on a $p_F(\mathbf{y}) = \langle \mathbf{y} | \mathbf{u} \rangle \mathbf{u} + \langle \mathbf{y} | \mathbf{x}^* \rangle \mathbf{x}^*$.
On remplace : $\inf_{(a,b) \in \mathbb{R}^2} \|\mathbf{y} - a\mathbf{x} - b\mathbf{u}\| = \|\mathbf{y} - m_{\mathbf{y}}\mathbf{u} - \langle \mathbf{y} | \mathbf{x}^* \rangle \mathbf{x}^*\|$ (car $\langle \mathbf{y} | \mathbf{u} \rangle = m_{\mathbf{y}}$).

- (e) Puis $\sigma_{\mathbf{y}}\mathbf{y}^* = \mathbf{y} - m_{\mathbf{y}}\mathbf{u}$, par définition de $\mathbf{y}^*$,
alors que $\sigma_{\mathbf{y}}\rho = \frac{\text{cov}(\mathbf{x}, \mathbf{y})}{\sigma_{\mathbf{x}}\sigma_{\mathbf{y}}} = \frac{1}{\sigma_{\mathbf{x}}} \langle \mathbf{x} | \mathbf{y} \rangle_D = \frac{1}{\sigma_{\mathbf{x}}} \langle \mathbf{x} - p_D(\mathbf{x}) | \mathbf{y} - p_D(\mathbf{y}) \rangle = \frac{\langle \frac{\mathbf{x} - m_{\mathbf{x}}\mathbf{u}}{\sigma_{\mathbf{x}}} | \mathbf{y} - p_D(\mathbf{y}) \rangle}{\sigma_{\mathbf{x}}}$
 $= \langle \mathbf{x}^* | \mathbf{y} - p_D(\mathbf{y}) \rangle = \langle \mathbf{x}^* | \mathbf{y} \rangle - \langle \mathbf{x}^* | p_D(\mathbf{y}) \rangle = \langle \mathbf{x}^* | \mathbf{y} \rangle$ car $\underbrace{\langle \mathbf{x}^* | p_D(\mathbf{y}) \rangle}_{\substack{\in D^\perp \\ \in D}} = 0$

Donc (en remplaçant) $\inf_{(a,b) \in \mathbb{R}^2} \|\mathbf{y} - a\mathbf{x} - b\mathbf{u}\| = \sigma_{\mathbf{y}} \|\mathbf{y}^* - \rho \mathbf{x}^*\|$

- (f) $\sigma_{\mathbf{y}}(\mathbf{y}^* - \rho \mathbf{x}^*) = \mathbf{y} - m_{\mathbf{y}}\mathbf{u} - \frac{\rho\sigma_{\mathbf{y}}}{\sigma_{\mathbf{x}}}\mathbf{x} + \frac{\rho\sigma_{\mathbf{y}}}{\sigma_{\mathbf{x}}}m_{\mathbf{x}}\mathbf{u}$.
Donc $a_0 = \frac{\rho\sigma_{\mathbf{y}}}{\sigma_{\mathbf{x}}}$ et $b_0 = m_{\mathbf{y}} - \frac{\rho\sigma_{\mathbf{y}}}{\sigma_{\mathbf{x}}}m_{\mathbf{x}} = \frac{\sigma_{\mathbf{x}}m_{\mathbf{y}} - \rho\sigma_{\mathbf{y}}m_{\mathbf{x}}}{\sigma_{\mathbf{x}}}$

8. On a donc $y = \frac{\rho\sigma_{\mathbf{y}}}{\sigma_{\mathbf{x}}}x + m_{\mathbf{y}} - \frac{\rho\sigma_{\mathbf{y}}m_{\mathbf{x}}}{\sigma_{\mathbf{x}}}$, donc $\mathcal{D}_0$ a pour équation $\frac{y - m_{\mathbf{y}}}{\sigma_{\mathbf{y}}} = \rho \frac{x - m_{\mathbf{x}}}{\sigma_{\mathbf{x}}}$.

9. L'unicité est assuré par le fait que ce minimum est atteint en $z = p_{F'}(\mathbf{x})$ où $F' = \text{vect}(\mathbf{u}, \mathbf{y})$.
La même série de calcul qu'en 7. mais avec une inversion $\mathbf{x} \leftrightarrow \mathbf{y}$ donne

$$a_1 = \frac{\rho\sigma_{\mathbf{x}}}{\sigma_{\mathbf{y}}} \text{ et } b_1 = m_{\mathbf{x}} - \frac{\rho\sigma_{\mathbf{x}}}{\sigma_{\mathbf{y}}}m_{\mathbf{y}} = \frac{\sigma_{\mathbf{y}}m_{\mathbf{x}} - \rho\sigma_{\mathbf{x}}m_{\mathbf{y}}}{\sigma_{\mathbf{y}}}$$

10. On a donc (de même avec l'inversion $x \leftrightarrow y$) : $\mathcal{D}_1$ a pour équation $\frac{x - m_{\mathbf{x}}}{\sigma_{\mathbf{x}}} = \rho \cdot \frac{y - m_{\mathbf{y}}}{\sigma_{\mathbf{y}}}$

11. Les droites se coupent en un seul point ssi les coefficients directeurs sont différents.
Celui de $\mathcal{D}_0$ vaut $\frac{\rho\sigma_{\mathbf{y}}}{\sigma_{\mathbf{x}}}$, celui de $\mathcal{D}_1$ vaut $\frac{\sigma_{\mathbf{y}}}{\rho\sigma_{\mathbf{x}}}$. Ce sont les mêmes si $\rho^2 = 1$.

Ce qui est faux car $\rho \in]-1, 1[$. Les droites se coupent donc en un seul point.
Soit $M(m_{\mathbf{x}}, m_{\mathbf{y}})$, alors les coordonnées de M vérifient les équations de $\mathcal{D}_0$ et de $\mathcal{D}_1$.
 M de coordonnées $(m_{\mathbf{x}}, m_{\mathbf{y}})$ est donc l'unique point d'intersection de $\mathcal{D}_0$ et de $\mathcal{D}_1$

12. Les droites sont orthogonales

ssi le produit des coefficients directeurs vaut -1 ou si l'un est nul et l'autre infini.

Or ce produit vaut : $\frac{\rho\sigma_{\mathbf{y}}}{\sigma_{\mathbf{x}}} \times \frac{\sigma_{\mathbf{y}}}{\rho\sigma_{\mathbf{x}}} = \frac{\sigma_{\mathbf{y}}^2}{\sigma_{\mathbf{x}}^2} > 0$, donc jamais égal à -1 .

ou bien il suffit que $\rho = 0$ (le premier coefficient serait nul et le second infini).

ou bien $\sigma_{\mathbf{x}} = 0$ (premier coefficient infini), mais le second ne serait pas nul.

Par conséquent, les droites $\mathcal{D}_0$ et $\mathcal{D}_1$ sont orthogonales $\iff \rho = 0$

$$\iff \text{cov}(\mathbf{x}, \mathbf{y}) = 0 \iff \langle \mathbf{x} | \mathbf{y} \rangle = m_{\mathbf{x}}m_{\mathbf{y}} \text{ car } \text{cov}(\mathbf{x}, \mathbf{y}) = \langle \mathbf{x} | \mathbf{y} \rangle - m_{\mathbf{x}}m_{\mathbf{y}}.$$

II. Base adaptée à un produit scalaire dans un espace euclidien

1. Il suffit de comparer les deux résultats.

On a $\langle \mathbf{x} | \mathbf{y} \rangle = \langle \sum_{i=1}^n x_i \mathbf{e}_i | \sum_{j=1}^n y_j \mathbf{e}_j \rangle = \sum_{i=1}^n \sum_{j=1}^n x_i y_j \langle \mathbf{e}_i | \mathbf{e}_j \rangle$ par bilinéarité.

Par ailleurs, SY produit d'une matrice carrée par une matrice colonne est une matrice colonne,
en notant i , l'indice d'un nombre située en i -ième position (ligne), on a $(SY)_i = \sum_{k=1}^n s_{i,k} y_k$.

puis $\mathbf{X}SY$ est un nombre qui vaut $\sum_{i=1}^n x_i (SY)_i = \sum_{i=1}^n \sum_{j=1}^n x_i y_j s_{i,j}$.

Or pour tout i, j , $s_{i,j} = \langle \mathbf{e}_i | \mathbf{e}_j \rangle$, donc pour tout $(\mathbf{x}, \mathbf{y}) \in E_n^2$, $\langle \mathbf{x} | \mathbf{y} \rangle = \mathbf{X}SY$ si $S = (\langle \mathbf{e}_i | \mathbf{e}_j \rangle)_{(i,j) \in \{1, \dots, n\}^2}$.

2. (a) Soit S est associée à $\mathcal{B}$

le produit scalaire est symétrique, i.e. pour tout i, j , $s_{i,j} = \langle \mathbf{e}_i | \mathbf{e}_j \rangle = \langle \mathbf{e}_j | \mathbf{e}_i \rangle = s_{j,i}$.

Donc S est une matrice symétrique. Elle est évidemment carrée d'ordre n .

S est symétrique réelle (E_n euclidien donc réel), donc toute ses valeurs propres sont réels.
Soit $\lambda \in \mathbb{R}$, une valeur propre de S et X un vecteur propre (réel) associé (non nul).

Alors $SX = \lambda X$ et donc $\|x\|^2 = {}^tXSX = \lambda {}^tXX = \lambda n(X)$ où $n(X) = \sum_{i=1}^n x_i^2 > 0$.

Donc $\lambda = \frac{\|x\|^2}{n(X)} > 0$. Ainsi le spectre de S dans $\mathbb{C}$ est inclus dans $\mathbb{R}_+^*$.

(b) S est diagonale ssi pour tout $i \neq j$, $s_{i,j} = 0 = \langle \mathbf{e}_i | \mathbf{e}_j \rangle$.

Donc S est diagonale ssi la base $\mathcal{B}$ est orthogonale (mais non nécessairement orthonormée).

3. Si $X = (x_i)_i$, $Y = (y_j)_j$ et $A = (a_{i,j})$, on a ${}^tXAY = \sum_{k,h=1}^n x_k a_{k,h} y_h$.

Donc avec $X = E_i$ et $Y = E_j$ (matrice colonnes avec le 1 en i -ième position resp. j -ième),

on a ${}^tXAY = {}^tE_i A E_j = a_{i,j}$ car $x_k = 0$ si $k \neq i$ et $y_h = 0$ si $h \neq j$.

et de même ${}^tXBY = {}^tE_i B E_j = b_{i,j}$.

Donc comme pour tout $X \in M_{n,1}(\mathbb{R})$ et $Y \in M_{n,1}(\mathbb{R})$, ${}^tXAY = {}^tXBY$,

on a $\forall i, j \in \mathbb{N}_n$, $a_{i,j} = b_{i,j}$ i.e. $A = B$.

4. (a) Formule du cours à connaître : $M_B(\mathbf{x}) = P_B^{\mathcal{B}'} \times M_{B'}(\mathbf{x})$ donc $X = P \times X'$.

(b) On a exactement la même relation qu'en 1. :

pour tout $(\mathbf{x}, \mathbf{y}) \in E_n^2$, $\langle \mathbf{x} | \mathbf{y} \rangle = {}^tX' S' Y'$ si $S = (\langle \mathbf{e}'_i | \mathbf{e}'_j \rangle)_{(i,j) \in \{1, \dots, n\}^2}$.

On a donc ${}^tX' S' Y' = {}^tXSY$.

Or $X = P \times X'$, donc ${}^tX = {}^tX' \times {}^tP$ et de même $Y = P \times Y'$.

Donc pour tout X', Y' , ${}^tX' S' Y' = {}^tX' ({}^tPSP) Y'$ et donc d'après 3. : $S' = {}^tPSP$.

(c) Il s'agit exactement de la condition $\mathcal{B}'$ est orthonormée.

Or tout espace euclidien admet une base orthonormée. Donc il existe bien une base $\mathcal{B}'$ de E_n

telle que, pour tout $(\mathbf{x}, \mathbf{y}) \in E_n^2$, $\langle \mathbf{x}, \mathbf{y} \rangle = \sum_{i=1}^n x'_i y'_i$ si $\mathbf{x} = \sum_{i=1}^n x'_i \mathbf{e}'_i$ et $\mathbf{y} = \sum_{i=1}^n y'_i \mathbf{e}'_i$

(d) Une matrice P est orthogonale si elle transforme toute base orthonormée en une autre.

Or $\mathcal{B}'$ est orthonormée. Donc P est orthogonale ssi $\mathcal{B}$ est orthonormée.

5. Soit $(\mathbf{e}_i)_{i \in \mathbb{N}_n}$ une base orthogonale telle que $\langle \mathbf{e}_i | \mathbf{e}_i \rangle = d_i$.

Une telle base existe, il suffit de prendre une base $(\mathbf{e}'_i)$ orthonormée de E_n

puis de considérer $\mathbf{e}_i = \sqrt{d_i} \mathbf{e}'_i$, alors $\langle \mathbf{e}_i | \mathbf{e}_j \rangle = 0$ si $i \neq j$ et $\langle \mathbf{e}_i | \mathbf{e}_i \rangle = (\sqrt{d_i})^2 \langle \mathbf{e}'_i | \mathbf{e}'_i \rangle = d_i$.

Alors M_1 est associée à la base $(\mathbf{e}_i)_{i \in \mathbb{N}_n}$

6. Soit $M_3 = M_B(f_3) = \begin{pmatrix} 3 & -1 & -1 \\ -1 & 3 & -1 \\ -1 & -1 & 3 \end{pmatrix}$.

(a) Notons que M_3 est symétrique réelle donc diagonalisable.

Le polynôme caractéristique de M_3 (donc de f_3) est

$$\chi(\lambda) = \det(M_3 - \lambda I_3) = (3 - \lambda)^3 - 1 - 1 - 3 \times (3 - \lambda) \times (-1)^2 = 16 - 24\lambda + 9\lambda^2 - \lambda^3$$

$$\chi(\lambda) = (1 - \lambda)(16 - 8\lambda + \lambda^2) = (1 - \lambda)(4 - \lambda)^2.$$

Donc les valeurs propres de f_3 sont 1 et 4 (double).

Cherchons les espaces propres associés.

— Soit $(x, y, z) \in E_1$ alors $\begin{cases} 2x & -y & -z & = & 0 \\ -x & +2y & -z & = & 0 \\ -x & -y & +2z & = & 0 \end{cases} \Leftrightarrow \begin{cases} y = x \\ z = x \end{cases}$ Donc $E_1 = \text{vect}((1, 1, 1))$.

Il faut l'orthonormaliser : $\|(1, 1, 1)\| = \sqrt{3}$, donc $E_1 = \text{vect}((\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}))$

— Soit $(x, y, z) \in E_4$ alors $\begin{cases} -x & -y & -z & = & 0 \\ -x & -y & -z & = & 0 \\ -x & -y & -z & = & 0 \end{cases} \Leftrightarrow \begin{cases} x = -y - z \end{cases}$ Donc $E_4 = \text{vect}((-1, 1, 0), (-1, 0, 1))$.

Il faut l'orthonormaliser : $\|(-1, 1, 0)\| = \sqrt{2}$, puis pour second vecteur, on peut prendre le produit vectoriel de $(\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}})$ par $(\frac{-1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0)$ i.e. $(\frac{-1}{\sqrt{6}}, \frac{-1}{\sqrt{6}}, \frac{2}{\sqrt{6}})$.

Ainsi $E_4 = \text{vect}((\frac{-1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0), (\frac{-1}{\sqrt{6}}, \frac{-1}{\sqrt{6}}, \frac{2}{\sqrt{6}}))$

(b) Notons $D_3 = \text{diag}(1, 4, 4)$, alors d'après 5, D_3 est une matrice associée à une base de E_3 : la base $\mathcal{B}_1 = ((\mathbf{e}_1, 2\mathbf{e}_2, 2\mathbf{e}_3))$ où $\mathbf{e}_1 = (\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}})$, $\mathbf{e}_2 = (\frac{-1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0)$ et $\mathbf{e}_3 = (\frac{-1}{\sqrt{6}}, \frac{-1}{\sqrt{6}}, \frac{2}{\sqrt{6}})$.

Puis : $M_3 = P D_3 {}^tP$ où P est la matrice de passage orthogonale. D'où ${}^tX M_2 Y = {}^t({}^tP X) D_3 ({}^tP Y)$.

Ainsi, l'écriture $X' = {}^tP \times X$ correspond à celle d'une base adaptée pour D_3 .

Par conséquent : la transformation par $P = ({}^tP)^{-1}$ de la base $\mathcal{B}_1$ donne une base adaptée à M_3 .

7. (a) Il s'agit d'une famille libre, car orthogonale .

De plus, elle est composée de n éléments, donc toute famille adaptée est une base de E_n .

(b) En fait, une base adaptée est une base associée à la matrice $\frac{1}{n} I_n$.

Comme pour la question 5, une base adaptée existe donc toujours

- (c) Si $\mathcal{B}_1 = (\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n)$ est une base adaptée, alors $\mathcal{B}_2 = (-\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n)$ est aussi adaptée.
Donc E_n n'admet pas une unique base adaptée

(d) Par bilinéarité, on a $\langle \mathbf{x} | \mathbf{y} \rangle = \sum_{i=1}^n \sum_{j=1}^n x_i y_j \langle \mathbf{e}_i | \mathbf{e}_j \rangle = \sum_{i=1}^n x_i y_i \langle \mathbf{e}_i | \mathbf{e}_i \rangle = \frac{1}{n} \sum_{i=1}^n x_i y_i$.

(e) Notons $\mathbf{x} = \sum_{i=1}^n \mathbf{e}_i$, alors $\|\mathbf{x}\| = \langle \mathbf{x} | \mathbf{x} \rangle = \frac{1}{n} \sum_{i=1}^n x_i^2 = \frac{1}{n} \sum_{i=1}^n 1^2 = 1$

III. Droites des moindres carrés dans le cas général

1. Il s'agit de l'existence de la base adaptée de la partie précédente.

2. Rappelons que les points (A_i) ne sont pas alignés.

Soit $\alpha, \beta, \gamma \in \mathbb{R}$ tels que $\alpha \mathbf{u} + \beta \mathbf{x} + \gamma \mathbf{y} = 0$.

donc pour tout $i \in \mathbb{N}_n$, $\alpha 1 + \beta x_i + \gamma y_i = 0$.

Et par conséquent, tous les points A_i se trouvent sur la droite d'équation $\beta x + \gamma y + \alpha = 0$, ils sont donc tous alignés (si cette droite n'est pas vide).

Ce qui est faux, donc nécessairement, il faut que $\alpha = \beta = \gamma = 0$ (« droite vide »).

Par conséquent, $(\mathbf{u}, \mathbf{x}, \mathbf{y})$ est une famille libre de E_n .

3. $\mathcal{D}_{a,b}$ est la droite d'équation $y = ax + b$.

Et $p_{a,b}(A_i)$ a la même abscisse que A_i c'est-à-dire x_i et donc ordonnée $y_i = ax_i + b$.

Puis $\|\overrightarrow{p_{a,b}(A_i)A_i}\|_2^2 = (x_i - x_i)^2 + (y_i - ax_i - b)^2$.

$$\text{Donc } f_0(a, b) = \sum_{i=1}^n \|\overrightarrow{p_{a,b}(A_i)A_i}\|_2^2 = \sum_{i=1}^n (y_i - ax_i - b)^2.$$

$$\text{Alors que } \|\mathbf{y} - a\mathbf{x} - b\mathbf{u}\|^2 = \frac{1}{n} \sum_{i=1}^n (y_i - ax_i - b)^2.$$

$$\text{Finalement, on a bien pour tout } (a, b) \in \mathbb{R}^2, f_0(a, b) = \sum_{i=1}^n (y_i - ax_i - b)^2 = n \|\mathbf{y} - a\mathbf{x} - b\mathbf{u}\|^2.$$

$$\text{Et de même } f_1(a, b) = \sum_{i=1}^n (x_i - ay_i - b)^2 = n \|\mathbf{x} - a\mathbf{y} - b\mathbf{u}\|^2.$$

4. (a) n est indépendant de (a, b) , donc pour minimiser f_0 , il suffit de minimiser $\|\mathbf{y} - a\mathbf{x} - b\mathbf{u}\|^2$ et donc de faire la projection orthogonale de $\mathbf{y}$ sur $F = \text{vect}(\mathbf{u}, \mathbf{x})$.

Il existe alors une unique solution (a_0, b_0) qui vérifie $p_F(\mathbf{y}) = a_0 \mathbf{x} + b_0 \mathbf{u}$.

De même il existe une unique solution (a_1, b_1) qui minimise f_1 et qui vérifie $p_G(\mathbf{x}) = a_1 \mathbf{y} + b_1 \mathbf{u}$ où $G = \text{vect}(\mathbf{u}, \mathbf{y})$ - cf. partie I.

- (b) Nous avons vu (I.11) que les droites se coupent en un unique point de coordonnées $(m_{\mathbf{x}}, m_{\mathbf{y}})$.

$$\text{Or } m_{\mathbf{x}} = \langle \mathbf{x} | \mathbf{u} \rangle = \frac{1}{n} \sum_{i=1}^n x_i \text{ et } m_{\mathbf{y}} = \langle \mathbf{y} | \mathbf{u} \rangle = \frac{1}{n} \sum_{i=1}^n y_i.$$

Il s'agit exactement des coordonnées de l'isobarycentre des points $(A_0, A_1, \dots, A_n)$.

- (c) D'après la fin de la partie I, ces droites sont orthogonales ssi $\langle \mathbf{x} | \mathbf{y} \rangle = m_{\mathbf{x}} m_{\mathbf{y}}$.

$$\text{Or } \langle \mathbf{x} | \mathbf{y} \rangle = \frac{1}{n} \sum_{i=1}^n x_i y_i \text{ et } m_{\mathbf{x}} = \langle \mathbf{x} | \mathbf{u} \rangle = \frac{1}{n} \sum_{i=1}^n x_i \text{ (rappel : } u_i = 1).$$

$$\text{Donc ces droites sont orthogonales ssi } \frac{1}{n} \sum_{i=1}^n x_i y_i = \frac{1}{n^2} \sum_{i=1}^n x_i \sum_{i=1}^n y_i.$$

Et nous avons vu que dans ce cas, aux équations des droites sont associés les nombres $a_0 = \frac{\rho_{\mathbf{x}\mathbf{y}}}{\sigma_{\mathbf{x}}} = 0$ et $\frac{1}{a_1} = \infty$ donc $a_1 = 0$.

Finalement, on obtient les droites $\mathcal{D}_0$ d'équation $y = b_0 = m_{\mathbf{y}} = \frac{1}{n} \sum_{i=1}^n y_i$ et $\mathcal{D}_1$ d'équation

$$x = b_1 = m_{\mathbf{x}} = \frac{1}{n} \sum_{i=1}^n x_i, \text{ à la condition nécessaire et suffisante : } \sum_{i=1}^n x_i y_i = \frac{1}{n} \sum_{i=1}^n x_i \sum_{i=1}^n y_i$$

- (d) Prenons $A_1 = (-1, -1)$, $A_2 = (-1, 1)$, $A_3 = (1, 1)$ et $A_4 = (1, -1)$.

$$\sum_{i=1}^n x_i y_i = 1 + (-1) + 1 + (-1) = 0 \text{ et } \frac{1}{n} \sum_{i=1}^n x_i \sum_{i=1}^n y_i = \frac{1}{n} (-1 - 1 + 1 + 1) (-1 + 1 + 1 - 1) = 0.$$

La condition est vérifiée.

Les droites des moindres carrées ont pour équations $\mathcal{D}_0 : y = 0$ et $\mathcal{D}_1 : x = 0$, elles sont bien orthogonales.