Devoir (en classe) n°11

Problème

Soit $\mathcal{P} = \mathbb{R}^2$ le plan muni du produit scalaire canonique et du repère orthonormé $\mathcal{R} = (O, \mathbf{i}, \mathbf{j})$.
Les vecteurs sont notés en gras : $\mathbf{x}$, $\mathbf{y}$...afin de les différenciés des scalaires $x, y \dots \in \mathbb{R}$.

Le but de ce problème est, un certain nombre de points étant donnés $A_1(x_1, f(x_1)), \dots A_n(x_n, f(x_n))$, de trouver la droite qui est la « plus proche »(au sens des moindres carrés) de ce nuage de points. C'est par exemple la méthode employée dans le logiciel Regressi.
Dans la première partie, nous appliquons la méthode classique de minimisation d'une norme, par projection orthogonale. Nous en répercutons les résultats pour le calcul des moindres carrés. Dans la seconde partie, nous étudions l'impact d'un changement de base orthogonale pour redéfinir le produit scalaire comme on le souhaite (on parle d'une base adaptée). Dans la dernière partie, nous exploitons les résultats des deux parties précédentes pour généraliser les calculs dans le cas d'une norme particulière : celle qui donne les droites de moindres carrés.

Pour a et b deux réels donnés, on définit $\mathcal{D}_{a,b}$ la droite d'équation dans $\mathcal{R}$: $y = ax + b$.
Si $M \in \mathcal{P}$ a pour coordonnées $(x, y) \in \mathcal{R}$, on note $p_{a,b}(M)$ l'unique point de $\mathcal{D}_{a,b}$ ayant dans $\mathcal{R}$ la même abscisse x que M.
On définit de même $\mathcal{D}'_{a,b}$ la droite d'équation dans $\mathcal{R}$: $x = ay + b$ et $p'_{a,b}(M)$ l'unique point de $\mathcal{D}'_{a,b}$ ayant dans $\mathcal{R}$ la même ordonnée y que M.

I. Droites des moindres carrées dans un espace euclidien

Soit E un espace préhilbertien réel non réduit à $\{0\}$ et F un sous-espace vectoriel de dimension finie de E. On note $\langle \cdot | \cdot \rangle$ le produit scalaire sur E et $\| \cdot \|$ la norme associée à ce produit scalaire.

1. Donner la définition de $F^\perp$. Énoncer (sans démonstration) une propriété vérifiée par F et $F^\perp$ valable en général. Dans le cas où E est de dimension finie, que peut-on dire de plus ?

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

2. Démontrer que $\inf\limits_{\mathbf{z} \in F} \|\mathbf{x} - \mathbf{z}\|$ est bien défini et que cette borne inférieure est atteinte en un unique élément de $\mathbf{z}$ de F défini par $\mathbf{z} = p_F(\mathbf{x})$.

Cette borne inférieure est notée $d(\mathbf{x}, F)$. On a donc $d(\mathbf{x}, F) = \|\mathbf{x} - p_F(\mathbf{x})\|$.

On dit qu'une application $(\mathbf{x}, \mathbf{y}) \mapsto \langle \mathbf{x} | \mathbf{y} \rangle_F$ de E^2 dans $\mathbb{R}$ est un <u>produit subordonné à F</u> si elle vérifie les 4 propriétés suivantes :

- **(i)** $\forall\, \mathbf{x} \in E$, l'application $\mathbf{y} \mapsto \langle \mathbf{x} | \mathbf{y} \rangle_F$ est une forme linéaire sur E ;
- **(ii)** $\forall\, \mathbf{x} \in E, \forall\, \mathbf{y} \in E$, $\langle \mathbf{x} | \mathbf{y} \rangle_F = \langle \mathbf{y} | \mathbf{x} \rangle_F$;
- **(iii)** $\forall\, \mathbf{x} \in E, \forall\, \mathbf{y} \in F$, $\langle \mathbf{x} | \mathbf{y} \rangle_F = 0$;
- **(iv)** $\forall\, \mathbf{x} \in F^\perp, \forall\, \mathbf{y} \in F^\perp$, $\langle \mathbf{x} | \mathbf{y} \rangle_F = \langle \mathbf{x} | \mathbf{y} \rangle$.

3. (a) Montrer que si $(\mathbf{x}, \mathbf{y}) \mapsto \langle \mathbf{x} | \mathbf{y} \rangle_F$ est un produit subordonné à F, alors
   — $\forall\, \mathbf{x}, \mathbf{y} \in E$, $\langle \mathbf{x}, \mathbf{y} \rangle_F = \langle (\mathbf{x} - p_F(\mathbf{x}) | \mathbf{y} - p_F(\mathbf{y}) \rangle$;
   — $\forall\, \mathbf{x} \in E$, $\langle \mathbf{x} | \mathbf{x} \rangle_F = \big(d(\mathbf{x}, F)\big)^2$;
   — $\forall\, \mathbf{x} \in E$, $\langle \mathbf{x} | \mathbf{x} \rangle_F \geqslant 0$;
   — $\forall\, \mathbf{x} \in E$, $\langle \mathbf{x} | \mathbf{x} \rangle_F = 0 \iff \mathbf{x} \in F$;

   (b) Vérifier qu'il existe unique produit subordonné à F.

On note alors $\langle \cdot | \cdot \rangle_F$ ce produit subordonné à F et pour $\mathbf{x} \in E$, on pose $\|\mathbf{x}\|_F = \sqrt{\langle \mathbf{x} | \mathbf{x} \rangle_F}$.

4. Montrer que pour tout $(\mathbf{x}, \mathbf{y}) \in E^2$, $|\langle \mathbf{x} | \mathbf{y} \rangle_F| \leqslant \|\mathbf{x}\|_F \cdot \|\mathbf{y}\|_F$; à quelle condition sur $\mathbf{x}$ et $\mathbf{y}$ peut-on dire que $|\langle \mathbf{x} | \mathbf{y} \rangle_F| = \|\mathbf{x}\|_F \cdot \|\mathbf{y}\|_F$?

5. (a) Montrer l'existence d'un élément de E noté $\mathbf{u}$, tel que $\|\mathbf{u}\| = 1$.

On note $D = \mathrm{vect}(\mathbf{u})$, la droite vectoriel engendré par $\mathbf{u}$ et p_D, la projection orthogonale sur D.

   (b) Vérifier que pour tout $\mathbf{x} \in E$, $p_D(\mathbf{x}) = \langle \mathbf{x} | \mathbf{u} \rangle \mathbf{u}$.

Pour tout élément $\mathbf{x} \in E$, on pose $m_{\mathbf{x}} = \langle \mathbf{x} | \mathbf{u} \rangle$, $\sigma_{\mathbf{x}} = \|\mathbf{x}\|_D$.
Pour tout couple $(\mathbf{x}, \mathbf{y}) \in E^2$, on pose $\mathrm{cov}(\mathbf{x}, \mathbf{y}) = \langle \mathbf{x} | \mathbf{y} \rangle_D$.

(c) Montrer que $\sigma_{\mathbf{x}} = \|\mathbf{x} - m_{\mathbf{x}}\mathbf{u}\|$ et que $\text{cov}(\mathbf{x}, \mathbf{y}) = \langle \mathbf{x}, \mathbf{y} \rangle - m_{\mathbf{x}} m_{\mathbf{y}}$.

On suppose dans la suite de cette partie que $\mathbf{x}$ et $\mathbf{y}$ sont deux éléments de E tels que la famille $(\mathbf{u}, \mathbf{x}, \mathbf{y})$ soit libre.

6. Montrer que $\sigma_{\mathbf{x}}$ et $\sigma_{\mathbf{y}}$ sont deux réels strictement positifs.

On pose alors $\mathbf{x}^* = \dfrac{\mathbf{x} - m_{\mathbf{x}}\mathbf{u}}{\sigma_{\mathbf{x}}}$, $\mathbf{y}^* = \dfrac{\mathbf{y} - m_{\mathbf{y}}\mathbf{u}}{\sigma_{\mathbf{y}}}$ et $\rho = \dfrac{\text{cov}(\mathbf{x}, \mathbf{y})}{\sigma_{\mathbf{x}}\sigma_{\mathbf{y}}}$.

7. (a) Montrer que $m_{\mathbf{x}^*} = 0$, que $\sigma_{\mathbf{x}^*} = 1$ et que $\rho \in]-1, 1[$.
   (b) Vérifier alors que $(\mathbf{u}, \mathbf{x}^*)$ est une base orthonormale de $F = \text{vect}(\mathbf{u}, \mathbf{x})$.
   (c) Montrer que $\inf\limits_{(a,b)\in\mathbb{R}^2} \|\mathbf{y} - a\mathbf{x} - b\mathbf{u}\|$ est bien défini et vaut $d(\mathbf{y}, F)$.
   (d) Établir que $\inf\limits_{(a,b)\in\mathbb{R}^2} \|\mathbf{y} - a\mathbf{x} - b\mathbf{u}\| = \|\mathbf{y} - m_{\mathbf{y}}\mathbf{u} - \langle \mathbf{y}|\mathbf{x}^* \rangle \mathbf{x}^*\|$.
   (e) Vérifier que $\inf\limits_{(a,b)\in\mathbb{R}^2} \|\mathbf{y} - a\mathbf{x} - b\mathbf{u}\| = \sigma_{\mathbf{y}} \|\mathbf{y}^* - \rho \mathbf{x}^*\|$
   (f) Déterminer, en fonction de $\mathbf{x}$, $\mathbf{y}$, $\mathbf{u}$ (ou plutôt ρ, $\sigma_{\mathbf{x}}$, $\sigma_{\mathbf{y}}$, $m_{\mathbf{x}}$ et $m_{\mathbf{y}}$), l'unique couple de réels (a_0, b_0) tel que :
$$\inf_{(a,b)\in\mathbb{R}^2} \|\mathbf{y} - a\mathbf{x} - b\mathbf{u}\| = \|\mathbf{y} - a_0\mathbf{x} - b_0\mathbf{u}\|$$

Dans le plan $\mathcal{P}$, on définit $\mathcal{D}_0$ comme étant la droite dont l'équation dans $\mathcal{R}$ est : $y = a_0 x + b_0$.

8. Montrer que $\mathcal{D}_0$ a pour équation dans $\mathcal{R}$: $\dfrac{y - m_{\mathbf{y}}}{\sigma_{\mathbf{y}}} = \rho \cdot \dfrac{x - m_{\mathbf{x}}}{\sigma_{\mathbf{x}}}$.
9. Montrer de même qu'il existe un unique couple de réels (a_1, b_1) tel que
$$\inf_{(a,b)\in\mathbb{R}^2} \|\mathbf{x} - a\mathbf{y} - b\mathbf{u}\| = \|\mathbf{x} - a_1\mathbf{y} - b_1\mathbf{u}\|$$
   Dans le plan $\mathcal{P}$, on définie $\mathcal{D}_0$ comme étant la droite dont l'équation dans $\mathcal{R}$ est : $x = a_1 y + b_1$
10. Montrer que $\mathcal{D}_1$ a pour équation dans $\mathcal{R}$: $\dfrac{x - m_{\mathbf{x}}}{\sigma_{\mathbf{x}}} = \rho \cdot \dfrac{y - m_{\mathbf{y}}}{\sigma_{\mathbf{y}}}$ avec le même réel ρ défini précédemment.
11. Vérifier que $\mathcal{D}_0$ et $\mathcal{D}_1$ se coupent en un unique point $M \in \mathcal{P}$ de coordonnées dans $\mathcal{R}$: $(m_{\mathbf{x}}, m_{\mathbf{y}})$.
12. Montrer que les droites $\mathcal{D}_0$ et $\mathcal{D}_1$ sont orthogonales si et seulement si $\langle \mathbf{x}|\mathbf{y} \rangle = m_{\mathbf{x}} m_{\mathbf{y}}$

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

Soit E_n un espace euclidien de dimension n, avec $n \geqslant 1$.
On note $\langle \cdot|\cdot \rangle$ le produit scalaire sur E_n et $\|\cdot\|$ la norme associée à ce produit scalaire.
Soit $\mathcal{B} = (\mathbf{e}_1, \mathbf{e}_2, \ldots \mathbf{e}_n)$ une base de E_n. Pour tout élément $\mathbf{z} \in E_n$, on notera $Z\big(= M_{\mathcal{B}}(\mathbf{z})\big)$ la matrice (colonne) de $\mathbf{z}$ dans la base $\mathcal{B}$. Ainsi si $\mathbf{z} = \sum\limits_{i=1}^{n} z_i \mathbf{e}_i$, on a $Z = \begin{pmatrix} z_1 \\ \vdots \\ z_n \end{pmatrix}$.

1. Vérifier que pour tout $(\mathbf{x}, \mathbf{y}) \in E_n^2$, $\langle \mathbf{x}|\mathbf{y} \rangle = {}^tXSY$ si $X = M_{\mathcal{B}}(\mathbf{x})$, $Y = M_{\mathcal{B}}(\mathbf{y})$ et $S = \big(\langle \mathbf{e}_i|\mathbf{e}_j \rangle\big)_{(i,j)\in\{1,\ldots n\}^2}$.

On dit alors que <u>S est associée à $\mathcal{B}$</u>.

2. (a) Vérifier que si S est associée à $\mathcal{B}$, alors S est une matrice carrée symétrique réelle d'ordre n et que le spectre de S dans $\mathbb{C}$ est inclus dans $\mathbb{R}_+^*$.
   (b) A quelle condition sur $\mathcal{B}$, la matrice S associée à $\mathcal{B}$ est-elle diagonale ?
3. On considère deux matrices A et B carrées d'ordre n telles que pour tout $X \in M_{n,1}(\mathbb{R})$ et $Y \in M_{n,1}(\mathbb{R})$, ${}^tXAY = {}^tXBY$. Montrer que $A = B$.

Notons $\mathcal{B}' = (\mathbf{e}'_1, \ldots, \mathbf{e}'_n)$, une base de E_n et P, la matrice de passage de $\mathcal{B}$ à $\mathcal{B}'$.

4. (a) Pour $\mathbf{x} \in E_n$, on note $X = M_{\mathcal{B}}(\mathbf{x})$ et $X' = M_{\mathcal{B}'}(\mathbf{x})$.
   Donner la relation entre X, X' et P.
   (b) On note $S' = (\langle \mathbf{e}'_1, \mathbf{e}'_2 \rangle)_{i,j\in\{1,\ldots n\}^2}$. Pour $(\mathbf{x}, \mathbf{y}) \in E_n^2$, $X' = M_{\mathcal{B}'}(\mathbf{x})$ et $Y' = M_{\mathcal{B}'}(\mathbf{y})$, donner l'expression de $\langle \mathbf{x}|\mathbf{y} \rangle$ en fonction de X', Y' et S'. En déduire que $S' = {}^tPSP$.
   (c) Montrer qu'il existe une base $\mathcal{B}'$ de E_n telle que, pour tout $(\mathbf{x}, \mathbf{y}) \in E_n^2$,
$$\langle \mathbf{x}, \mathbf{y} \rangle = \sum_{i=1}^{n} x'_i y'_i \qquad \text{si} \quad \mathbf{x} = \sum_{i=1}^{n} x'_i \mathbf{e}'_i \quad \text{et} \quad \mathbf{y} = \sum_{i=1}^{n} y'_i \mathbf{e}'_i$$
   (d) A quelle condition sur $\mathcal{B}$, la matrice de passage P de $\mathcal{B}$ à la base précédente $\mathcal{B}'$ est-elle orthogonale ?
5. Étant donné une matrice $M_1 \in \mathcal{M}_n(\mathbb{R})$ diagonale avec des réels $(d_1, \ldots, d_n)$ strictement positifs sur la diagonale, M_1 est-elle la matrice associée à une base de E_n ?

6. Soit $M_3 = \begin{pmatrix} 3 & -1 & -1 \\ -1 & 3 & -1 \\ -1 & -1 & 3 \end{pmatrix}$ et $\mathcal{B} = (\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$ une base orthonormale de E_3. On note f_3 l'endomorphisme de E_3 dont la matrice dans $\mathcal{B}$ est M_3.
   (a) Déterminer le spectre de f_3 et une base orthonormale de chaque sous-espace propre de f_3.
   (b) M_3 est-elle la matrice associée à une base de E_3 ?

On dit qu'une famille $\mathcal{B} = (\mathbf{e}_1, \dots, \mathbf{e}_n)$ d'éléments de E_n est adaptée si :

$$\forall\, i \neq j \in \mathbb{N}_n \quad \langle \mathbf{e}_i | \mathbf{e}_j \rangle = 0 \qquad \text{et} \qquad \forall\, i \in \mathbb{N}_n \langle \mathbf{e}_i | \mathbf{e}_i \rangle = \frac{1}{n}$$

7. (a) Montrer qu'une famille adaptée est une base de E_n
   (b) Montrer l'existence d'une base adaptée
   (c) E_n admet-il une unique base adaptée ?
   (d) On suppose que $\mathcal{B}$ est une base adaptée. Pour $(\mathbf{x}, \mathbf{y}) \in E_n^2$, déterminer l'expression de $\langle \mathbf{x} | \mathbf{y} \rangle$ en fonction des coordonnées de $\mathbf{x}$ et de $\mathbf{y}$ dans la base $\mathcal{B}$.
   (e) Calculer alors la norme du vecteur $\displaystyle\sum_{i=1}^{n} \mathbf{e}_i$.

III. Droites des moindres carrés dans le cas général

Soit n un entier supérieur ou égal à 3.
On considère $A_1, A_2, \dots, A_n$, n points distincts du plan $\mathcal{P}$ qui ne sont pas alignés. On note $(x_1, y_1), \dots, (x_n, y_n)$ leurs coordonnées respectives dans $\mathcal{R}$.
On définit deux applications f_0 et f_1 de $\mathbb{R}^2$ dans $\mathbb{R}$ en posant : pour tout $(a, b) \in \mathbb{R}^2$,

$$f_0(a,b) = \sum_{i=1}^{n} \|\overrightarrow{p_{a,b}(A_i)A_i}\|_2^{\,2} \qquad \text{et} \qquad f_1(a,b) = \sum_{i=1}^{n} \|\overrightarrow{p'_{a,b}(A_i)A_i}\|_2^{\,2}$$

On considère dans toute la suite du problème un espace euclidien E_n de dimension n, dont le produit scalaire et la norme associée sont notés respectivement $\langle \cdot | \cdot \rangle$ et $\| \cdot \|$.

1. Justifier l'existence d'une base $\mathcal{B} = (\mathbf{e}_1, \dots \mathbf{e}_n)$ de E_n telle que :

$$\forall\, (\mathbf{z}, \mathbf{t}) \in E_n^2, \quad \langle \mathbf{z} | \mathbf{t} \rangle = \frac{1}{n} \sum_{i=1}^{n} z_i t_i \qquad \text{si} \quad \mathbf{z} = \sum_{i=1}^{n} z_i \mathbf{e}_i \quad \text{et} \quad \mathbf{t} = \sum_{i=1}^{n} t_i \mathbf{e}_i$$

On pose $\mathbf{u} = \displaystyle\sum_{i=1}^{n} \mathbf{e}_i$, si bien que $\|\mathbf{u}\| = 1$.
On définit alors, à partir des points $A_1(x_1, y_1), \dots, A_n(x_n, y_n)$, deux éléments $\mathbf{x}$ et $\mathbf{y}$ dans E_n en posant $\mathbf{x} = \displaystyle\sum_{i=1}^{n} x_i \mathbf{e}_i$ et $\mathbf{y} = \displaystyle\sum_{i=1}^{n} y_i \mathbf{e}_i$.

2. Montrer que $(\mathbf{u}, \mathbf{x}, \mathbf{y})$ est une famille libre de E_n
3. Montrer que pour tout $(a, b) \in \mathbb{R}^2$, $f_0(a, b) = n\|\mathbf{y} - a\mathbf{x} - b\mathbf{u}\|^2$ et $f_1(a, b) = n\|\mathbf{x} - a\mathbf{y} - b\mathbf{u}\|^2$.
4. (a) En déduire que f_0 admet un minimum sur $\mathbb{R}^2$ qui est atteint en un unique couple de réels, noté (a_0, b_0), et qu'il en est de même de f_1 avec un unique couple de réels noté (a_1, b_1).

Dans le plan $\mathcal{P}$, on définit alors les droites $\mathcal{D}_0$ et $\mathcal{D}_1$ d'équation dans $\mathcal{R}$: $y = a_0 x + b_0$ et $x = a_1 y + b_1$.
On les appelle les droites des moindres carrés associées à $A_1, \dots A_n$.

   (b) Montrer que les droites $\mathcal{D}_0$ et $\mathcal{D}_1$ se coupent en un unique point $M \in \mathcal{P}$ qui est l'isobarycentre de $(A_0, A_1, \dots A_n)$.
   (c) A quelle condition sur les $x_1, \dots, x_n, y_1, \dots y_n$ les droites $\mathcal{D}_0$ et $\mathcal{D}_1$ sont-elles orthogonales ? Donner dans ce cas les équations dans $\mathcal{R}$ de $\mathcal{D}_0$ et $\mathcal{D}_1$.
   (d) Donner un exemple de quatre points distincts et non alignés $A_1, \dots, A_4$ de $\mathcal{P}$ tels que les droites des moindres carrés $\mathcal{D}_0$ et $\mathcal{D}_1$ associées à $A_1, \dots, A_4$ soient orthogonales et donner dans ce cas les équations dans $\mathcal{R}$ de $\mathcal{D}_0$ et $\mathcal{D}_1$.