PROBLÈME DES MOINDRES CARRÉS ET PSEUDO-SOLUTIONS

Dans ce problème, on s'intéresse aux systèmes linéaires n'ayant pas de solutions, et les résultats obtenus sont notamment utiles lorsqu'il y a plus d'équations linéaires à satisfaire que d'inconnues.
$\mathcal{M}_{n,1}(\mathbb{R})$ est muni de sa structure euclidienne canonique.

Soit $A \in \mathcal{M}_{n,p}(\mathbb{R})$ une matrice de **rang** $\boldsymbol{p}$, $X = \begin{pmatrix} x_1 \\ \vdots \\ x_p \end{pmatrix}$ une colonne de $\mathcal{M}_{p,1}(\mathbb{R})$ et B une colonne de $\mathcal{M}_{n,1}(\mathbb{R})$.
On s'intéresse au système linéaire d'inconnue X

$$(\mathcal{S}) : AX = B.$$

On dit que X est une **pseudo-solution** de $(\mathcal{S})$ lorsque

$$||AX - B|| = \min_{Y \in \mathcal{M}_{p,1}(\mathbb{R})} ||AY - B||\,,$$

autrement dit, lorsque AX est au plus près possible de B.
On montrera l'existence de $\displaystyle\min_{Y \in \mathcal{M}_{p,1}(\mathbb{R})} ||AY - B||$ au cours du problème.
On appelle enfin $(\mathcal{T})$ le système linéaire d'inconnue X

$$(\mathcal{T}) : {}^{t}AAX = {}^{t}AB.$$

1. Dans cette question, on suppose que $p = n$.
- **a)** Montrer que les systèmes $(\mathcal{S})$ et $(\mathcal{T})$ sont équivalents et admettent une unique solution.
- **b)** Justifier cette solution est aussi l'unique solution pseudo-solution de $(\mathcal{S})$.

On revient au cas général.

2.
- **a)** Justifier que $n \geqslant p$.
- **b)** Montrer que $AX = 0$ si, et seulement si, ${}^{t}AAX = 0$.
- **c)** En déduire que ${}^{t}AA$ est de rang p.
  Que peut-on dire du système $(\mathcal{T})$?

3. On note $(C_1, \dots, C_p)$ les p colonnes A et $F = \text{Vect}(C_1, \dots, C_p) = \text{Im}(A)$.
- **a)** Montrer que $\{AY - B / Y \in \mathcal{M}_{p,1}(\mathbb{R})\} = \{C - B / C \in F\}$.
- **b)** Justifier l'existence de $\displaystyle\min_{Y \in \mathcal{M}_{p,1}(\mathbb{R})} ||AY - B||$.

4.
- **a)** Montrer que $Z \in F^{\perp}$ si, et seulement si, ${}^{t}AZ = 0$.
- **b)** En déduire que si X est solution de $(\mathcal{T})$, alors AX est le projeté orthogonal $p_F(B)$ de B sur F.

5. Montrer finalement que $(\mathcal{S})$ admet une unique pseudo-solution, qui est l'unique solution de $(\mathcal{T})$.

6.
- **a)** Le système $\begin{cases} x + 2y = 3 \\ 2x - y = 1 \\ 3x + y = 2 \end{cases}$ admet-il une solution ?
- **b)** Justifier qu'il admet une unique pseudo-solution et la déterminer.

7. Application à la méthode de régression linéaire par les moindres carrés en y de Gauss
On considère une série statistique $(x_i, y_i)_{1 \leqslant i \leqslant n}$ pour laquelle les x_i (respectivement les y_i) ne sont pas tous égaux entre eux.
On chercher deux réels a et b tels que

$$\sum_{i=1}^{n} (y_i - (ax_i + b))^2$$

soit minimum.
- **a)** Montrer que, par un choix judicieux de A et B que l'on précisera, ce problème se ramène à la recherche de $X = \begin{pmatrix} a \\ b \end{pmatrix}$ réalisant le minimum de $||AX - B||$.
- **b)** Montrer que la solution au problème est

$$\begin{cases} a = \dfrac{\sigma_{xy}}{\sigma_x^2} \\ b = \overline{y} - a\overline{x} \end{cases}$$

où σ_{xy}, σ_x^2, $\overline{y}$ et $\overline{x}$ désignent respectivement la covariance de x et y, la variance de x, la moyenne de y et la moyenne de x.