# Feuille d'exercices n°5
# Introduction à la méthode des moindres carrés

## Démonstrations de cours

*Les exercices de cette section* ***ne seront pas*** *traités en TD, les corrigés se trouvant dans le polycopié. Les exercices marqués ♣ sont exigibles au partiel et à l'examen.*

### ♣ Exercice 1 – Gradient et matrice hessienne d'une fonction quadratique généralisée

Module B2 – Propositions 1 et 2

Soit $E = \mathbb{R}^n$. Soient $A \in \mathcal{M}_{n,n}(\mathbb{R})$ une matrice **symétrique** réelle, $b \in E$ un vecteur et $c \in \mathbb{R}$ un scalaire. On considère la fonction quadratique généralisée

$$\left\{\begin{array}{ccc} E & \to & \mathbb{R} \\ x & \mapsto & \dfrac{1}{2}\,\langle A\,x, x\rangle - \langle b, x\rangle + c \end{array}\right.$$

**(a)** Soit $(x, h) \in E^2$. Calculer $f(x+h)$.

**(b)** Montrer que
$$\forall\, h \neq 0, \qquad \frac{1}{2}\,\frac{\langle A\,h, h\rangle}{\|h\|_2} \leq \frac{1}{2}\,|||A|||\cdot\|h\|_2$$
En déduire que
$$\frac{1}{2}\,\frac{\langle A\,h, h\rangle}{\|h\|_2} = o(\|h\|_2)$$

**(c)** Déduire de ce qui précède que $f$ est différentiable en $x$ et que
$$\forall\, h \in E, \qquad df(x)\cdot h = \frac{1}{2}\,\langle A\,x, h\rangle + \frac{1}{2}\,\langle A\,h, x\rangle - \langle b, x\rangle$$
puis que
$$\nabla f(x) = A\,x - b$$

**(d)** Montrer que $\nabla f : E \to E$ est différentiable, et calculer sa matrice jacobienne. En déduire que
$$\forall\, x \in E, \qquad \operatorname{Hess} f(x) = A$$

### ♣ Exercice 2 – Caractérisation de la convexité d'une fonction quadratique généralisée

Module B2 – Proposition 3

Soit $E = \mathbb{R}^n$. Soient $A \in \mathcal{M}_{n,n}(\mathbb{R})$ une matrice **symétrique** réelle, $b \in E$ un vecteur et $c \in \mathbb{R}$ un scalaire. On considère la fonction quadratique généralisée

$$\left\{\begin{array}{ccc} E & \to & \mathbb{R} \\ x & \mapsto & \dfrac{1}{2}\,\langle A\,x, x\rangle - \langle b, x\rangle + c \end{array}\right.$$

**(a)** Justifier que $f$ est convexe si et seulement si $A$ est définie positive.

**(b)**
1. Justifier que, si $A$ est définie positive, alors $f$ est strictement convexe.
2. *On suppose dans cette question que $f$ est strictement convexe.* Soit $x \in E$. Justifier que
$$\forall\, h \in E \setminus \{0\}, \qquad \langle \nabla f(x+h) - \nabla f(x), (x+h) - x\rangle > 0$$
En déduire que $A$ est définie positive.

## ♣ Exercice 3 – Minimisation d'une fonction quadratique généralisée convexe

Module B2 – Propositions 4–5

Soit $E=\mathbb{R}^n$. Soient $A\in\mathcal{M}_{n,n}(\mathbb{R})$ une matrice **symétrique** réelle, $b\in E$ un vecteur et $c\in\mathbb{R}$ un scalaire. On considère la fonction quadratique généralisée

$$\left\{\begin{array}{ccc} E & \to & \mathbb{R} \\ x & \mapsto & \dfrac{1}{2}\langle A\,x,x\rangle-\langle b,x\rangle+c \end{array}\right.$$

On suppose que $A$ est **semi-définie positive**.

**(a)** Montrer que les minimiseurs de $f$ sont exactement les solutions du système linéaire

$$A\,x=b$$

**(b)** En déduire le nombre de minimiseurs pour $f$.

**(c)** Donner un exemple où $f$ n'admet aucune minimiseur et un exemple où $f$ admet une infinité de minimiseurs.

**(d)** Dans quel cas $f$ admet-elle une unique solution ?

## ⋆ Exercice 4 – Minimisation d'une fonction quadratique généralisée strictement convexe

Module B2 – Propositions 6 et 7 et Corollaire 1

Soit $E=\mathbb{R}^n$. Soient $A\in\mathcal{M}_{n,n}(\mathbb{R})$ une matrice **symétrique** réelle, $b\in E$ un vecteur et $c\in\mathbb{R}$ un scalaire. On considère la fonction quadratique généralisée

$$\left\{\begin{array}{ccc} E & \to & \mathbb{R} \\ x & \mapsto & \dfrac{1}{2}\langle A\,x,x\rangle-\langle b,x\rangle+c \end{array}\right.$$

On suppose que $A$ est **définie positive**.

**(a)** Justifier que $f$ admet au plus un minimiseur.

**(b)**
1. Justifier que $A$ est diagonalisable, et que ses valeurs propres (que l'on note $\lambda_i$) sont strictement positives. Justifier par ailleurs qu'il existe une base de vecteurs propres orthonormée $(V_i)_{i\in[\![1\,;\,n]\!]}$.
2. Soit $x\in E$. En posant $P=(V_1,\ldots,V_n)^\top\in\mathcal{M}_{n,n}(\mathbb{R})$, montrer que
$$\langle A\,x,x\rangle=\sum_{i=1}^n\lambda_i\,(P\,x)_i^2$$
En déduire que
$$\langle A\,x,x\rangle\geq\lambda_1\sum_{i=1}^n(P\,x)_i^2 \qquad \text{où } \lambda_1=\min_{1\leq i\leq n}\lambda_i,$$
puis que
$$\langle A\,x,x\rangle\geq\lambda_1\,\|x\|_2^2$$
Justifier que $f$ est infinie à l'infini.
3. En déduire que $f$ admet au moins un minimiseur.
4. En déduire que $f$ admet un unique minimiseur.

## ♣ Exercice 5 – Minimisation d'une fonction quadratique généralisée non convexe

Module B2 – Propositions 4 et 5

Soit $E=\mathbb{R}^n$. Soient $A\in\mathcal{M}_{n,n}(\mathbb{R})$ une matrice **symétrique** réelle, $b\in E$ un vecteur et $c\in\mathbb{R}$ un scalaire. On considère la fonction quadratique généralisée

$$\left\{\begin{array}{ccc} E & \to & \mathbb{R} \\ x & \mapsto & \dfrac{1}{2}\langle A\,x,x\rangle-\langle b,x\rangle+c \end{array}\right.$$

On suppose que $A$ **n'est pas semi-définie positive**.

**(a)** Montrer qu'aucun $x\in E$ n'est minimiseur local de $f$.

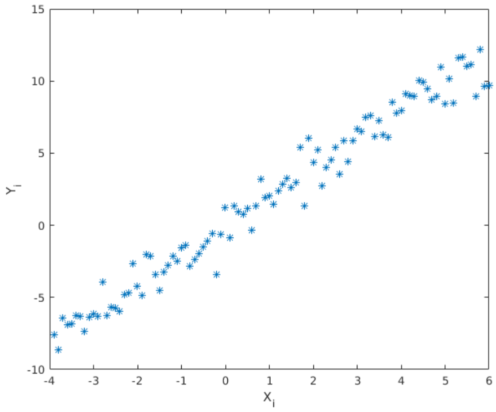
**(b)** En déduire que $f$ n'admet aucun minimiseur.

# Exercices fondamentaux

**Exercice 6 – Fonction quadratique** Soient $a \in \mathbb{R}$ et $(b_i)_{1\leq i\leq n} \in \mathbb{R}^n$. Montrer que les fonction suivantes sont des fonctions quadratiques généralisées. Sont-elles convexes ?

**(a)** $f : \left\{ \begin{array}{ccc} \mathbb{R} & \to & \mathbb{R} \\ t & \mapsto & -2\,t^2 - t + 2 \end{array} \right.$ **(b)** $f : \left\{ \begin{array}{ccc} \mathbb{R} & \to & \mathbb{R} \\ t & \mapsto & \displaystyle\frac{1}{2}\sum_{i=1}^{n}(b_i - t\,a_i)^2 \end{array} \right.$ **(c)** $f : \left\{ \begin{array}{ccc} \mathbb{R}^2 & \to & \mathbb{R} \\ (x,y) & \mapsto & x^2 - 2\,x\,y + y^2 \end{array} \right.$

**Exercice 7 – Régression linéaire** On dispose du nuage de points $(X,Y) = \{(X_i, Y_i)\}_{i=1,\dots,p} \in \mathbb{R}^{2\,p}$ suivant :




On cherche à estimer la fonction **linéaire**

$$\left\{ \begin{array}{ccc} \mathbb{R} & \to & \mathbb{R} \\ t & \mapsto & a\,t \end{array} \right.$$

approchant ces données au sens des moindres carrés, c'est-à-dire minimiser la fonction suivante

$$J : \left\{ \begin{array}{ccc} \mathbb{R} & \to & \mathbb{R} \\ a & \mapsto & \displaystyle\frac{1}{2}\sum_{i=1}^{p}(Y_i - a\,X_i)^2 \end{array} \right.$$

**(a)** Montrer que $J$ est une fonction quadratique généralisée.

**(b)** Justifier que $J$ est dérivable et calculer sa dérivée.

**(c)** Justifier que $J$ est deux fois différentiable et calculer sa dérivée seconde.

**(d)** Montrer que $J$ est convexe.

**(e)** Justifier l'existence de minimiseurs pour $J$.

**(f)** Utiliser la condition d'optimalité du premier ordre pour caractériser les minimiseurs de $J$. *On distinguera le cas où $X = 0$ et $X \neq 0$.*

**(g)** Donner une condition nécessaire et suffisante pour que $J$ admette un unique minimiseur. Que vaut-elle dans ce cas ?

**Exercice 8 – Régression affine** Traiter les questions de l'**Exercice 7** lorsque l'on cherche à estimer la fonction **affine** $t \mapsto a\,t + b$ approchant les données du nuage de points au sens des moindres carrés, c'est-à-dire minimiser la fonction suivante

$$J : \left\{ \begin{array}{ccc} \mathbb{R} & \to & \mathbb{R} \\ t & \mapsto & a \mapsto \displaystyle\frac{1}{2}\sum_{i=1}^{p}(Y_i - a\,X_i - b)^2 \end{array} \right.$$

# Compléments

**Exercice 9 – Méthode de Gauss–Seidel** On s'intéresse à la résolution du système linéaire suivant :

$$\left\{ \begin{array}{l} x + \dfrac{y}{2} = 1 \\ \dfrac{x}{2} + \dfrac{y}{3} = 1 \end{array} \right. \qquad (\mathcal{S})$$

**(a)** Résoudre le système linéaire $(\mathcal{S})$.

**(b)** Justifier que l'unique solution de $(\mathcal{S})$ est l'unique minimiseur de la fonction

$$f : \left\{ \begin{array}{ccc} \mathbb{R}^2 & \to & \mathbb{R} \\ (x,y) & \mapsto & \dfrac{1}{2}\left(x^2 + x\,y + \dfrac{1}{3}\,y^2\right) - x - y \end{array} \right.$$

**(c)** On considère l'algorithme suivant :

$$y_0 \in \mathbb{R} \qquad \text{et} \qquad \forall\, k \in \mathbb{N}, \quad \begin{cases} x_{k+1} = \arg\min\limits_{x\in\mathbb{R}} f(x, y_k) \\ y_{k+1} = \arg\min\limits_{y\in\mathbb{R}} f(x_{k+1}, y) \end{cases}$$

**1.** Justifier que l'algorithme considéré est bien défini, c'est-à-dire que les réels $x_{k+1}$ et $y_{k+1}$ sont définis de manière unique pour tout $k \in \mathbb{N}$.

**2.** Soit $k \in \mathbb{N}$. Donner l'expression de $x_{k+1}$ et $y_{k+1}$ en fonction de $y_k$. Vérifier que

$$\begin{cases} x_{k+1} + \dfrac{y_k}{2} & = 1 \\ \dfrac{x_{k+1}}{2} + \dfrac{y_{k+1}}{3} & = 1 \end{cases}$$

**3.** Justifier que $\qquad \forall\, k \in \mathbb{N}, \qquad y_k = \left(\dfrac{3}{4}\right)^k (y_0 - 6) + 6$

puis que $\qquad \forall\, k \in \mathbb{N}, \qquad x_k = -2 - \dfrac{2}{3}\left(\dfrac{3}{4}\right)^k (y_0 - 6)$

**4.** En déduire que la suite $((x_k, y_k))_{k\in\mathbb{N}}$ converge vers la solution de $(\mathcal{S})$.