

Chapitre 1

Rappels mathématiques

1.1 Introduction

Lorsqu'on parle de l'optimisation nous devons impérativement passer par les mathématiques qui présentent l'outil indispensable afin de Les techniques d'optimisation se basent sur

1.1.1 Notations

Dans ce qui suit, on considère la notation des variables suivantes :

- n, m, p, \dots , des scalaires qui indiquent souvent les dimensions,
- $x, y, t, \lambda, \alpha \dots$, des variables scalaires,
- $\mathbf{x}, \mathbf{y}, \mathbf{X}, \mathbf{Y}, \dots$, des vecteurs colonne,
- $\mathbf{A}, \mathbf{M}, \mathbf{P}, \dots$, des matrices.
- $\mathbf{E}, \mathbf{I}, \dots$, ensembles ou intervalles

1.1.2 Vecteur

Le vecteur réel \mathbf{X} de dimension n , se met sous la forme :

$$\mathbf{X} = \begin{bmatrix} x_1 \\ x_2 \\ \dots \\ x_n \end{bmatrix} \quad (1.1)$$

ou encore de la manière suivante :

$$\mathbf{X} = [x_1 \ x_2 \ \dots \ x_n]' \quad (1.2)$$

- deux vecteurs sont égaux ssi tous leurs éléments sont égaux.
- Le produit scalaire de deux vecteurs \mathbf{X} et \mathbf{Y} est défini par :

$$\mathbf{X}'\mathbf{Y} = \mathbf{Y}'\mathbf{X} = \sum_{i=1}^n x_i y_i \quad (1.3)$$

- Les deux vecteurs \mathbf{X} et \mathbf{Y} sont orthogonaux si :

$$\mathbf{X}'\mathbf{Y} = 0 \quad (1.4)$$

- La norme d'un vecteur \mathbf{X} est :

$$|\mathbf{X}| = \sqrt{\mathbf{X}'\mathbf{X}} \quad (1.5)$$

- L'inégalité de *Cauchy-Schwarz* est vérifiée pour les deux vecteurs \mathbf{X} et \mathbf{Y} :

$$|\mathbf{X}'\mathbf{Y}| \leq |\mathbf{X}| \cdot |\mathbf{Y}| \quad (1.6)$$

1.1.3 Matrice

Dans ce cours on ne considère que les matrices réelles. Une matrice \mathbf{A} de dimension $m \times n$ est donnée par :

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix} \quad (1.7)$$

La matrice transposée de la matrice \mathbf{A} qu'on note \mathbf{A}' ou \mathbf{A}^T , est de dimension $n \times m$ est la suivante :

$$\mathbf{A}' = \begin{bmatrix} a_{11} & a_{21} & \dots & a_{m1} \\ a_{12} & a_{22} & \dots & a_{m2} \\ \dots & \dots & \dots & \dots \\ a_{1n} & a_{2n} & \dots & a_{mn} \end{bmatrix} \quad (1.8)$$

La somme (différence) de deux matrices de dimensions identiques est une matrice de même dimension, dont les éléments sont égaux à la somme (différence) des éléments de même position.

Le produit de deux matrices \mathbf{A} ($m \times p$) et \mathbf{B} ($p \times n$) est une matrice \mathbf{C} ($m \times n$) dont les éléments sont :

$$c_{ij} = \sum_{k=1}^p a_{ik} b_{kj} \quad i = 1, \dots, m \quad j = 1, \dots, n \quad (1.9)$$

La multiplication des matrices requiert une attention particulière, le nombre de colonnes de la première doit être égal au nombre de lignes de la deuxième. Prenant l'exemple illustratif suivant : soit un vecteur \mathbf{X} de dimension n alors,

$$\begin{aligned} \mathbf{X}'\mathbf{X} &= \begin{bmatrix} x_1 & x_2 & \dots & x_n \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \dots \\ x_n \end{bmatrix} \\ &= x_1^2 + x_2^2 + \dots + x_n^2 \end{aligned}$$

Le résultat du produit $\mathbf{X}'\mathbf{X}$ est un scalaire.

$$\begin{aligned} \mathbf{X}\mathbf{X}' &= \begin{bmatrix} x_1 \\ x_2 \\ \dots \\ x_n \end{bmatrix} \begin{bmatrix} x_1 & x_2 & \dots & x_n \end{bmatrix} \\ &= \begin{bmatrix} x_1^2 & x_1x_2 & \dots & x_1x_n \\ x_2x_1 & x_2^2 & \dots & x_2x_n \\ \dots & \dots & \dots & \dots \\ x_nx_1 & x_nx_2 & \dots & x_n^2 \end{bmatrix} \end{aligned}$$

Le résultat du produit $\mathbf{X}\mathbf{X}'$ est une matrice.

Une matrice est dite carrée dans le cas particulier où $n = m$. On définit la matrice identité \mathbf{I} de la manière suivante :

$$\mathbf{I} = \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 1 \end{bmatrix} \quad (1.10)$$

Souvent on indique la dimension n de la matrice identité \mathbf{I}_n .

Exercice 1. *Opérations sur les matrices et les vecteurs*

. Soit les deux vecteurs $\mathbf{A} = [1, 2, 3]'$ et $\mathbf{B} = [-1, 0, 2]'$. Effectuer les opérations suivantes :

a. \mathbf{AB}' .

b. $\mathbf{A}'\mathbf{B}$.

c. $\mathbf{AB}'\mathbf{B}$

1.1.4 Valeurs propres

On appelle le scalaire λ valeur propre de la matrice \mathbf{A} s'il existe un vecteur non nul \mathbf{V} tel que :

$$\mathbf{AV} = \lambda\mathbf{V} \quad (1.11)$$

Le vecteur \mathbf{V} est appelé vecteur propre de la matrice \mathbf{A} associé à la valeur propre λ (**Exp.** $\mathbf{A} = \begin{bmatrix} 3 & 2 \\ 3 & -2 \end{bmatrix}$, $\mathbf{V} = [2 \ 1]'$, $\lambda = 4$).

Les valeurs propres présentent la solution de l'équation caractéristique suivante :

$$|\mathbf{A} - \lambda\mathbf{I}_n| = 0 \quad (1.12)$$

Le déterminant d'une matrice \mathbf{A} peut être donné par le produit de ses valeurs propres qu'on note $\{\lambda_i\}_{i=1}^n$.

$$|\mathbf{A}| = \prod_{i=1}^n \lambda_i \quad (1.13)$$

Si la matrice carrée \mathbf{A} est symétrique alors :

1. Les valeurs propres sont positifs.
2. Les vecteurs propres associées aux différentes valeurs propres sont orthogonaux.
3. Il existe une base orthogonale où chaque élément présente un vecteur propre de \mathbf{A} .

La matrice carrée \mathbf{A} est inversible si son déterminant est différent de zéro. La matrice inverse est notée \mathbf{A}^{-1} , on a alors :

$$\mathbf{AA}^{-1} = \mathbf{I}_n \quad (1.14)$$

La trace de \mathbf{A} est la somme de ses éléments diagonaux :

$$Tr(\mathbf{A}) = \sum_{i=1}^n a_{ii} \quad (1.15)$$

La trace d'une matrice carrée est égale à la somme de ces valeurs propres $Tr(\mathbf{A}) = \sum \lambda_i$.

Une matrice est dite symétrique si elle vérifie la condition :

$$\mathbf{A} = \mathbf{A}' \quad (1.16)$$

Dans ce qui suit on ne considère que le cas des matrices carrées et symétriques, sauf si on l'indique clairement.

Exercice 2.

Soit la matrice carrée A dont les valeurs propres sont : $\lambda_1 = 1, \lambda_2 = 2$ et $\lambda_3 = 5$.

1. Quelles sont les dimensions de la matrice A ?
2. Que peut on dire sur la positivité de la matrice A ?
3. Comment déterminer l'équation caractéristique de A ?
4. Calculer le déterminant de A .
5. La matrice A est elle inversible ? Pourquoi ?
6. Calculer la trace de la matrice A .

1.2 Positivité

1.2.1 Définition

- Une fonction $f(x)$ est dite positive ssi $f(x) > 0, \forall x \in \mathbb{R}$.
- Une matrice symétrique \mathbf{A} est définie positive ssi $\mathbf{X}'\mathbf{A}\mathbf{X} > 0, \forall \mathbf{X} \neq \mathbf{0}$.
- Une matrice symétrique \mathbf{A} de dimension n est définie positive si toutes les valeurs propres sont positives $\lambda_i > 0, \forall i = 1, \dots, n$.
- Une matrice \mathbf{A} de dimension n est définie positive si tous les mineurs principaux sont positifs.

1.2.2 Exemples

1. La fonction $f(x) = x^2 - 4x + 5$ est définie positive sur l'ensemble \mathbb{R} .
2. La matrice : $\mathbf{A} = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$ est définie positive, car ses valeurs propres sont positives. Les valeurs propres sont les solutions de l'équation caractéristique $|\mathbf{A} - \lambda \mathbf{I}_2| = 0$.

Exercice 3. *Matrice positive*

. Déterminer parmi les matrices suivantes celles qui sont définies positives :

a.

$$\mathbf{A} = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$$

b.

$$\mathbf{B} = \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix}$$

Utiliser la définition ensuite les valeurs propres.

1.3 Convexité

1.3.1 Fonction convexe

Soit la fonction f , définie sur un intervalle \mathbf{I} de \mathbb{R}^n , telle que $\forall (x_1, x_2) \in \mathbf{I}^2$ et $\forall \alpha \in [0, 1]$, on dit que la fonction f est convexe (ou plus exactement strictement convexe) si elle vérifie l'inégalité suivante :

$$f(\alpha x_1 + (1 - \alpha)x_2) \leq \alpha f(x_1) + (1 - \alpha)f(x_2) \quad (1.17)$$

Remarque Si la fonction $f(x) \in \mathbb{R}$, elle est convexe si on a $f''(x) \geq 0$ sur le domaine de définition.

1.3.2 Ensemble convexe

Un ensemble $\Omega \subset \mathbb{R}^n$ est dit convexe ssi :

$$\forall u, v \in \Omega \quad \text{et} \quad \forall \alpha \in [0, 1], \quad \text{on a} \quad x = \alpha u + (1 - \alpha)v \in \Omega \quad (1.18)$$

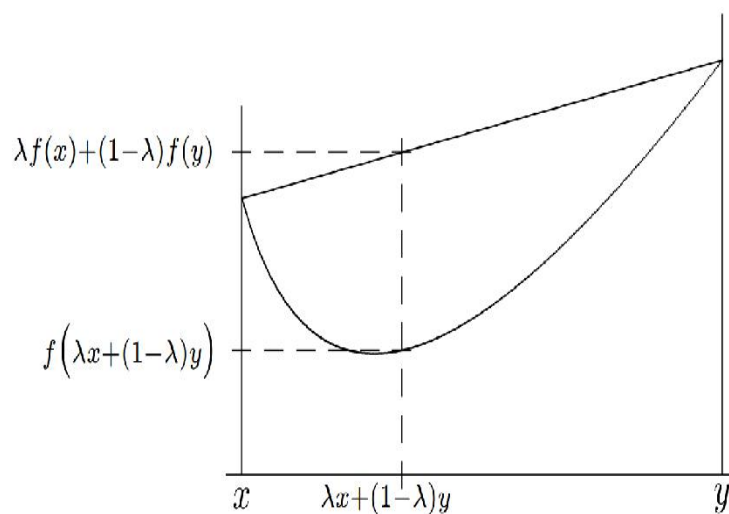


FIGURE 1.1 – Exemple d'une fonction convexe

Exercice 4.

Démontrer que l'ensemble définie par $\Omega = \{ \mathbf{X} \in \Re^n : \mathbf{a}'\mathbf{X} \geq \beta \}$, où $\mathbf{a} \in \Re^n$ et $\beta \in \Re$, est convexe, Ω est appelé demi-espace *halfspace*.

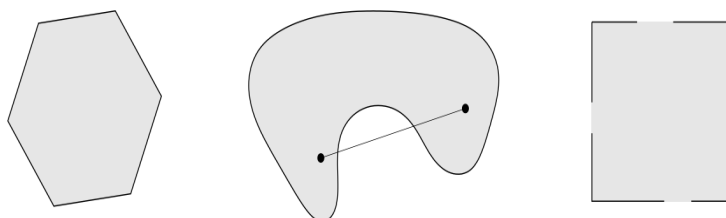


FIGURE 1.2 – Illustration d'ensembles convexes et non convexes

1.4 Le gradient

Soit $f(\mathbf{X})$ une fonction scalaire de n variables. Le gradient de la fonction f qu'on note ∇f est un vecteur ligne dont les éléments sont les dérivées partielles de la fonction f suivant chaque variable.

$$\nabla f = \frac{\partial f}{\partial \mathbf{X}} = \left[\frac{\partial f}{\partial x_1} \quad \frac{\partial f}{\partial x_2} \quad \dots \quad \frac{\partial f}{\partial x_n} \right] \quad (1.19)$$

Exemple Trouver le gradient de la fonction :

$$f(x, y) = x^2 + xy + y^2$$

Solution

Les dérivées partielles suivant les variables x et y sont :

$$\begin{aligned}\frac{\partial f}{\partial x} &= 2x + y, \\ \frac{\partial f}{\partial y} &= x + 2y,\end{aligned}$$

Le vecteur du gradient est donnée par :

$$\frac{\partial f}{\partial \mathbf{X}} = [2x + y, x + 2y]$$

1.5 Le Hessien

Soit $f(\mathbf{X})$ une fonction scalaire de n variables. Le Hessien de la fonction f , qu'on note \mathbf{H}_f est une matrice de dimension $n \times n$, dont les éléments sont les dérivées partielles d'ordre deux de la fonction f .

$$\mathbf{H}_f = \frac{\partial^2 f(\mathbf{X})}{(\partial \mathbf{X})^2} = \begin{bmatrix} \frac{\partial^2 f(\mathbf{X})}{(\partial x_1)^2} & \frac{\partial^2 f(\mathbf{X})}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 f(\mathbf{X})}{\partial x_1 \partial x_n} \\ \frac{\partial^2 f(\mathbf{X})}{\partial x_2 \partial x_1} & \frac{\partial^2 f(\mathbf{X})}{(\partial x_2)^2} & \cdots & \frac{\partial^2 f(\mathbf{X})}{\partial x_2 \partial x_n} \\ \cdots & \cdots & \cdots & \cdots \\ \frac{\partial^2 f(\mathbf{X})}{\partial x_n \partial x_1} & \frac{\partial^2 f(\mathbf{X})}{\partial x_n \partial x_2} & \cdots & \frac{\partial^2 f(\mathbf{X})}{(\partial x_n)^2} \end{bmatrix} \quad (1.20)$$

Exemple

1. Trouver le Hessien de la fonction :

$$f(x, y) = x^2 + xy + y^2$$

2. Démontrer que cette matrice est définie positive.

Solution

1. Calcule du Hessien :

$$\begin{aligned}\frac{\partial f}{\partial x} &= 2x + y, & \longrightarrow \frac{\partial^2 f}{(\partial x)^2} &= 2 \\ & & \longrightarrow \frac{\partial^2 f}{\partial x \partial y} &= 1 \\ \frac{\partial f}{\partial y} &= x + 2y, & \longrightarrow \frac{\partial^2 f}{(\partial y)^2} &= 2 \\ & & \longrightarrow \frac{\partial^2 f}{\partial y \partial x} &= 1\end{aligned}$$

$$\mathbf{H}f = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$$

2. Pour démontrer que cette matrice est définie positive, on doit calculer les valeurs propres correspondantes.

L'équation caractéristique dans ce cas est donnée par :

$$\begin{aligned}(2 - \lambda)^2 - 1 &= 0 \\ (2 - \lambda - 1)(2 - \lambda + 1) &= 0 \\ (1 - \lambda)(3 - \lambda) &= 0\end{aligned}$$

Les deux valeurs propres correspondantes sont $\lambda_1 = 1$, $\lambda_2 = 3$ et sont positives, ainsi la matrice est définie positive.

Remarque

Une fonction $f(\mathbf{X})$ est convexe ssi $\mathbf{H}_f(\mathbf{X})$ est semi-définie positif pour tout $\mathbf{X} \in \mathbb{R}^n$.

Exercice 5. Gradient et Hessien

. Considérons la fonction scalaire suivante :

$$f(x, y) = x - y - x^2y + xy^2$$

- a. Calculer les dérivées partielles $f_x = \frac{\partial f}{\partial x}$ et $f_y = \frac{\partial f}{\partial y}$.
- b. Trouver les extrémums en résolvant les équations $f_x = 0$ et $f_y = 0$.
- c. Calculer le Hessien de la fonction f .
- d. Déduire la nature des extrémums.

1.6 Exercices de synthèse

Exercice 6. Régression linéaire

. Soit un ensemble de n points dans le plan $\{\mathbf{X}_i = [x_i, y_i]'\}_{i=1}^n$, ces points sont souvent les résultats d'une expérience. Les points ne sont pas sur la même ligne, Cependant on veut trouver la ligne approchant au mieux ce nuage de points. Pour cela on utilise la méthode des moindres carrés, et on choisit la relation suivante $y = ax + b$ avec $(a, b) \in \mathbb{R}^2$. La fonction objective est la somme des carrés des erreurs (cette fonction est convexe) :

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

- a.* Calculer le gradient de la fonction $f(a, b)$.
- b.* Ecrire le système d'équations qui donne les extremums.
- c.* Donner l'écriture matricielle correspondante à ce système.
- d.* Trouver les coefficients (a, b) qui minimisent cette fonction.

Exercice 7.

Soit la fonction réelle à deux variables :

$$f(x, y) = x - x^2y + xy^2$$

1. Calculer le Jacobien,
2. Calculer le Hessien,
3. Déterminer l'équation caractéristique du Hessien,
4. Donner les valeurs propres de cette matrice,
5. Cette matrice est-elle définie positive, pourquoi ?
6. Calculer le déterminant en utilisant les valeurs propres.
7. Trouver analytiquement les optimums de cette fonction. Quelles sont leurs natures ?

1.7 Solutions des exercices

Solution Ex. 1.

$$\begin{aligned}\mathbf{AB}' &= \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} [-1 \ 0 \ 2] \\ &= \begin{bmatrix} -1 & 0 & 2 \\ -2 & 0 & 4 \\ -3 & 0 & 6 \end{bmatrix}\end{aligned}$$

$$\mathbf{A}'\mathbf{B} = [1 \ 2 \ 3] \begin{bmatrix} -1 \\ 0 \\ 2 \end{bmatrix} = 5$$

— Le calcul de l'expression $\mathbf{AB}'\mathbf{B}$ peut se faire de deux façon différentes, suivant l'ordre dans lequel nous calculons le produit :

•

$$\begin{aligned}(\mathbf{AB}')\mathbf{B} &= \begin{bmatrix} -1 & 0 & 2 \\ -2 & 0 & 4 \\ -3 & 0 & 6 \end{bmatrix} \begin{bmatrix} -1 \\ 0 \\ 2 \end{bmatrix} \\ &= \begin{bmatrix} 5 \\ 10 \\ 15 \end{bmatrix}\end{aligned}$$

•

$$\mathbf{A}(\mathbf{B}'\mathbf{B}) = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} \times 5 = \begin{bmatrix} 5 \\ 10 \\ 15 \end{bmatrix}$$

Remarque

Lors de calcul du produit de plusieurs matrices il est souvent préférable de commencer par le calcul des produits qui minimisent le nombre d'opérations de multiplications.

Solution Ex. 2.

La matrice carrée A dispose des 3 valeurs propres $\lambda_1 = 1, \lambda_2 = 2$ et $\lambda_3 = 5$. alors :

1. La matrice A est de dimension 3×3 (car elle est carrée et dispose de 3 valeurs propres).
2. La matrice A est positive car toutes les valeurs propres sont positives.
3. L'équation caractéristique de A est donnée par la résolution de l'équation :

$$\det(A - \lambda I) = 0 \quad (1.21)$$

Dans notre cas, on a en plus $(\lambda - \lambda_1)(\lambda - \lambda_2)(\lambda - \lambda_3) = 0$

4. Le déterminant de A est égale à :

$$\det(A) = \prod_{i=1}^3 \lambda_i = \lambda_1 \times \lambda_2 \times \lambda_3 = 10 \quad (1.22)$$

5. La matrice A est inversible, car le déterminant est non nul.
6. La trace de la matrice A est égale à la somme des éléments de sa diagonale. elle est égale aussi à la somme des valeurs propres.

$$\text{trace}(A) = \sum_{i=1}^3 \lambda_i = \lambda_1 + \lambda_2 + \lambda_3 = 8 \quad (1.23)$$

Solution Ex. 3.

Deux méthodes permettent de déterminer si une matrice réelle \mathbf{C} de dimension $(n \times n)$ est définie positive ou non :

1. Utilisation de la définition

La matrice \mathbf{C} est définie positive si on a $\mathbf{X}'\mathbf{C}\mathbf{X} > 0, \forall \mathbf{X} \in \mathbb{R}^{*n}$.

(a) Pour la matrice \mathbf{A}

$$\mathbf{A} = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$$

$$\begin{aligned} \mathbf{X}'\mathbf{A}\mathbf{X} &= [x \ y] \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \\ &= 2x^2 + 2xy + 2y^2 \\ &= x^2 + y^2 + (x + y)^2 > 0, \quad \forall (x, y) \in \mathbb{R}^{*2} \end{aligned}$$

Ainsi, la matrice \mathbf{A} est définie positive.

(b) Pour la matrice \mathbf{B}

$$\mathbf{B} = \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix}$$

$$\begin{aligned} \mathbf{X}'\mathbf{B}\mathbf{X} &= [x \ y] \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \\ &= 2x^2 - 2xy + 2y^2 \\ &= x^2 + y^2 + (x - y)^2 > 0 \quad \forall (x, y) \in \mathbb{R}^{*2} \end{aligned}$$

Ainsi, la matrice \mathbf{B} est définie positive.

2. Utilisation des valeurs propres

La matrice est positive si toutes ces valeurs propres sont positives.

(a) Pour la matrice \mathbf{A}

$$\begin{aligned} \det(\mathbf{A} - \lambda \mathbf{I}_2) &= \det \left(\begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} - \lambda \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \right) \\ &= \det \left(\begin{bmatrix} 2 - \lambda & 1 \\ 1 & 2 - \lambda \end{bmatrix} \right) \\ &= (2 - \lambda)^2 - 1 \end{aligned}$$

Le déterminant s'annule pour $\lambda_1 = 1$ et $\lambda_2 = 3$

(b) Pour la matrice \mathbf{B}

$$\begin{aligned} \det(\mathbf{B} - \lambda \mathbf{I}_2) &= \det \left(\begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix} - \lambda \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \right) \\ &= \det \left(\begin{bmatrix} 2 - \lambda & -1 \\ -1 & 2 - \lambda \end{bmatrix} \right) \\ &= (2 - \lambda)^2 - 1 \end{aligned}$$

Le déterminant s'annule pour $\lambda_1 = 1$ et $\lambda_2 = 3$

Les valeurs propres sont positives alors les deux matrices \mathbf{A} et \mathbf{B} , sont définies positives.

Solution Ex. 4.

Soit $\mathbf{Z} = \lambda\mathbf{X} + (1 - \lambda)\mathbf{Y}$ pour $\lambda \in [0, 1]$. Alors :

$$\begin{aligned}\mathbf{a}'\mathbf{Z} &= \mathbf{a}'(\lambda\mathbf{X} + (1 - \lambda)\mathbf{Y}) \\ &= \lambda\mathbf{a}'\mathbf{X} + (1 - \lambda)\mathbf{a}'\mathbf{Y} \\ &\geq \lambda\beta + (1 - \lambda)\beta = \beta \quad (\text{on a } \lambda \geq 0 \text{ et } (1 - \lambda) \geq 0) \\ \mathbf{a}'\mathbf{Z} \geq \beta &\Rightarrow \mathbf{Z} \in \Omega\end{aligned}$$

par conséquent, Ω est un ensemble convexe.

Solution Ex. 5. Soit la fonction

$$f(x, y) = x - y - x^2y + xy^2$$

1. Calcul des gradients

$$\begin{aligned}f_x &= \frac{\partial f}{\partial x} = 1 - 2xy + y^2 \\ f_y &= \frac{\partial f}{\partial y} = -1 - x^2 + 2xy\end{aligned}$$

2. En sommant les deux équations $f_x = 0$ et $f_y = 0$, on obtient :

$$y^2 - x^2 = 0 \Rightarrow y = \pm x$$

— Soit $y = x$ alors on a :

$$1 - 2x^2 + x^2 = 0 \rightarrow 1 - x^2 = 0 \rightarrow x = \pm 1$$

ainsi on a deux points : $(1, 1)$ et $(-1, -1)$

— Soit $y = -x$ alors on a :

$$1 + 2x^2 + x^2 = 1 + 3x^2 = 0$$

cette équation n'a pas de solution dans \mathbb{R} .

3. Les dérivées partielles d'ordre deux et le Hessian sont :

$$\frac{\partial^2 f}{(\partial x)^2} = -2y$$

$$\frac{\partial^2 f}{\partial x \partial y} = -2x + 2y$$

$$\frac{\partial^2 f}{(\partial y)^2} = 2x$$

$$\frac{\partial^2 f}{\partial y \partial x} = -2x + 2y$$

$$\mathbf{H}_f = \begin{bmatrix} -2y & -2x + 2y \\ -2x + 2y & 2x \end{bmatrix}$$

4. La nature des extrémums, pour les deux points sont :

— pour le point $(x, y) = (1, 1)$

$$\mathbf{H}_f(1, 1) = \begin{bmatrix} -2 & 0 \\ 0 & 2 \end{bmatrix} \longrightarrow |\mathbf{H}_f(1, 1)| = -4 < 0$$

ainsi ce point est un point de selle.

— pour le point $(x, y) = (-1, -1)$

$$\mathbf{H}_f(-1, -1) = \begin{bmatrix} 2 & 0 \\ 0 & -2 \end{bmatrix} \longrightarrow |\mathbf{H}_f(-1, -1)| = -4 < 0$$

ainsi ce point est aussi un point de selle.

Solution Ex. 6.

Nous avons la fonction f à deux variables a et b :

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

1. Les gradients suivant les variables a et b sont comme suit :

$$\begin{aligned} \frac{\partial f}{\partial a} &= \sum_{i=1}^n 2(-y_i + ax_i + b)x_i \\ &= 2a \sum_{i=1}^n x_i^2 + 2b \sum_{i=1}^n x_i - 2 \sum_{i=1}^n y_i x_i \\ \frac{\partial f}{\partial b} &= \sum_{i=1}^n 2(-y_i + ax_i + b) \\ &= 2a \sum_{i=1}^n x_i + 2nb - 2 \sum_{i=1}^n y_i \end{aligned}$$

2. Sachant que cette fonction est convexe, l'extremum est donné par le système d'équation suivant :

$$\begin{aligned} \sum_{i=1}^n x_i^2 a + \sum_{i=1}^n x_i b &= \sum_{i=1}^n y_i x_i \\ \sum_{i=1}^n x_i a + n b &= \sum_{i=1}^n y_i \end{aligned}$$

3. Soit :

$$\left\{ \begin{array}{l} S_{11} = \sum_{i=1}^n x_i^2 \\ S_{12} = S_{21} = \sum_{i=1}^n x_i \\ S_{22} = n \end{array} \right.$$

$$\begin{cases} Z_1 = \sum_{i=1}^n y_i x_i \\ Z_2 = \sum_{i=1}^n y_i \end{cases}$$

L'écriture matricielle correspondante est :

$$\begin{bmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} Z_1 \\ Z_2 \end{bmatrix}$$

4. Les coefficients a et b sont données en résolvant le système d'équation précédent, on obtient alors :

$$a = \frac{S_{12}Z_2 - S_{22}Z_1}{S_{12}^2 - S_{11}S_{22}}$$

$$b = \frac{S_{12}Z_1 - S_{11}Z_2}{S_{12}^2 - S_{11}S_{22}}$$

Programme Matlab

La figure 1.3 montre le résultat après exécution du programme suivant :

```

1 %% Construction d'un nuage de points
2 a=2;
3 b=3;
4 N=100;
5 x=3+10*rand(1,N);
6 y=a*x+b+3*randn(1,N);
7 plot(x,y, '. ' ), hold on
8 %% Calcul des parametres
9 S11=sum(x.^2);
10 S12=sum(x);
11 S21=S12;
12 S22=N;
13 Z1=sum(y.*x);
14 Z2=sum(y);

```

```

15
16 a=(S12*Z2-S22*Z1)/(S12^2-S11*S22);
17 b=(S21*Z1-S11*Z2)/(S12^2-S11*S22);
18 %% Dessin des resultats
19 x_min=min(x);
20 x_max=max(x);
21 pas=(x_max-x_min)/20;
22 x=(x_min:pas:x_max);
23 y_est=a*(x_min:pas:x_max)+b;
24
25 plot(x,y_est,'r','LineWidth',3)
26 xlabel('x')
27 ylabel('y')
28 legend('Nuage de points','Droite estimee')

```

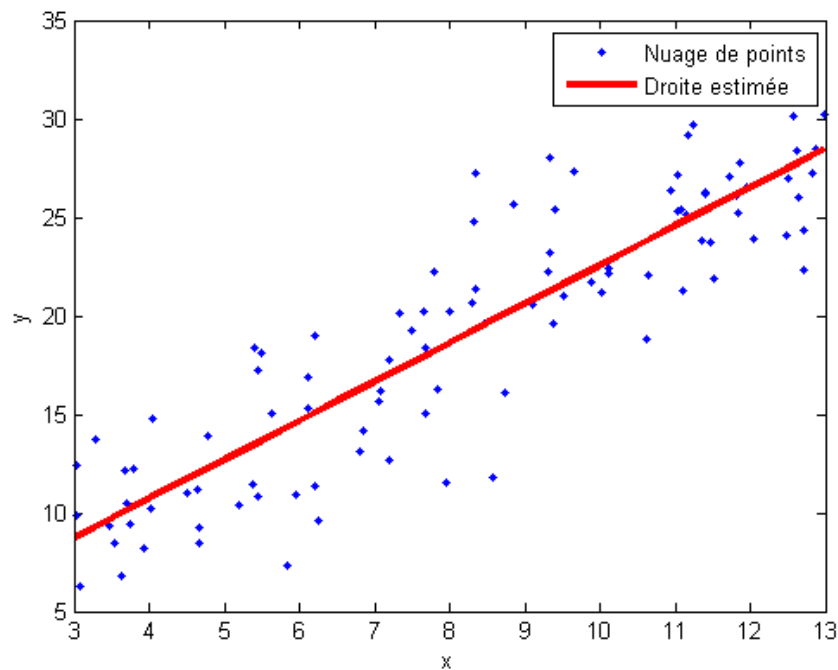


FIGURE 1.3 – Résultat de la régression linéaire

Solution Ex. 7.

Soit la fonction réelle à deux variables :

$$f(x, y) = x - x^2y + xy^2$$

1. Calcule du Jacobien,

$$\begin{aligned} \mathbf{J}_f = \nabla f' &= \begin{bmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \end{bmatrix} \\ &= \begin{bmatrix} 1 - 2xy + y^2 \\ -x^2 + 2xy \end{bmatrix} \end{aligned}$$

2. Le Hessian

$$\mathbf{H}_f = \begin{bmatrix} -2y & -2x + 2y \\ -2x + 2y & 2x \end{bmatrix}$$

3. L'équation caractéristique du Hessian est donnée par :

$$\begin{aligned} \det(H - \lambda I) &= 0 \\ \rightarrow (-2y - \lambda)(2x - \lambda) - (-2x + 2y)^2 &= 0 \\ \lambda^2 + 2(y - x)\lambda - (-2x + 2y)^2 - 4xy &= 0 \end{aligned}$$

4. Les valeurs propres sont les solutions de l'équation caractéristique de cette matrice, c'est une équation du deuxième ordre alors :

$$\begin{aligned} \Delta &= 4(y - x)^2 + 16(y - x)^2 + 16xy \\ &= 20(y - x)^2 + 16xy \end{aligned}$$

Les solution sont alors :

$$\begin{aligned} \lambda_1 &= \frac{-2(y - x) + \sqrt{\Delta}}{2} \\ &= x - y + \sqrt{5(y - x)^2 + 4xy} \\ \lambda_2 &= \frac{-2(y - x) - \sqrt{\Delta}}{2} \\ &= x - y - \sqrt{5(y - x)^2 + 4xy} \end{aligned}$$

5. Pour que cette matrice soit définie positive, il faut que $\lambda_1 > 0$ et $\lambda_2 > 0$. Ainsi la positivité de dépend des valeurs attribuées aux variables x et y .
6. Le déterminant est donné par le produit des valeurs propres :

$$\begin{aligned}\det(\mathbf{H}_f) &= \lambda_1 \times \lambda_2 \\ &= \left(x - y + \sqrt{5(y - x)^2 + 4xy}\right) \left(x - y - \sqrt{5(y - x)^2 + 4xy}\right) \\ &= -4(x^2 + y^2 - xy)\end{aligned}$$

7. L'optimum de cette fonction est obtenu analytiquement comme suit :

$$\mathbf{J}_f = \nabla f' = \begin{bmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \end{bmatrix} = \begin{bmatrix} 1 - 2xy + y^2 \\ -x^2 + 2xy \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad (1)$$

$$(2)$$

De la deuxième équation on tire :

$$x^2 - 2xy = 0 \rightarrow x(x - 2y) = 0$$

on élimine la solution $x = 0$ car elle ne vérifie pas la première équation. Alors on prend $x = 2y$ et on le remplace dans la première équation on trouve :

$$1 - 4y^2 + y^2 = 0 \rightarrow y = \pm \sqrt{\frac{1}{3}}$$

$$\begin{aligned}y = \sqrt{\frac{1}{3}} &\rightarrow x = \frac{2}{\sqrt{3}} \rightarrow X_1 = \left[\frac{2}{\sqrt{3}}, \sqrt{\frac{1}{3}}\right]' \\ y = -\sqrt{\frac{1}{3}} &\rightarrow x = -\frac{2}{\sqrt{3}} \rightarrow X_2 = \left[\frac{-2}{\sqrt{3}}, -\sqrt{\frac{1}{3}}\right]'\end{aligned}$$

La nature des points dépend de la valeur de la Hessienne à ces points :

$$\det(H) = -4(x^2 + y^2 - xy)$$

— pour X_1 on a $\det(H(X = X_1)) = -\frac{4}{3}$

— pour X_2 on a $\det(H(X = X_2)) = -\frac{4}{3}$

pour les deux points on a $\det(H) < 0$ alors on a deux points de selle.