République Algériènne Démocratique et populaire
Ministère de l'Enseignement Supérieur et de la Recherche Scientifique
Université des sciences et de la technologie d'Oran Mohammed Boudiaf
USTO-MB

Faculté de Génie Electrique
Département d'Automatique

Domaine : Sciences et Technologie
Filière : Automatique
Option : Automatique et informatique industrielle
Semestre : S1

Manuel de Cours

« Optimisation »

Par :

Dr. Meche Abdelkrim

Année universitaire 2021-2022

Introduction

Ce manuscrit traite les notions de base de l'optimisation et s'adresse essentiellement au étudiants de Master 1 spécialité Automatique et Informatique Industrielle.

L'optimisation est une branche des mathématiques, dont le but est de trouver analytiquement ou numériquement, la meilleur solution (*l'optimale*) à un problème donné. L'origine du mot optimale provient du Latin *optimum* qui signifie le meilleur. De nos jours, l'optimisation joue un rôle très important dans différents domaines de la vie.

Nous énumérons ci-dessous, quelques domaines d'application de l'optimisation :

- Transport et livraisons.
- Fabrication et production.
- Agriculture et génie civil.
- Finance, vente et marketing
- Gestion de stock
- Recherche et gestion des bases de données.
- ...

En optimisation, on parle de la fonction coût (objectif). C'est la fonction à minimiser/maximiser, après formulation mathématique du problème. On distingue deux grandes familles de techniques d'optimisation, et cela suivant le problème posé :

- Techniques d'optimisation sans contraintes.
- Techniques d'optimisation sous contraintes.

Le présent manuscrit traite les deux familles de problèmes. Nous allons considéré alors les fonctions : à une seule variable, à plusieurs variables, linéaires, non linéaires, continues et discrètes. Une étude particulière est consacrée aux fonctions convexes, ainsi, l'objectif sera de rechercher le minimum de cette fonction. Nous allons voir aussi, que ce type de fonction facilite énormément la résolution des problèmes d'optimisations.

Si dans un problème d'optimisation la formulation conduit à une fonction coût concave comme montrer dans la figure 1, alors nous allons la multiplier par -1, et le problème se trabnsforme en problème de minimisation d'une fonction convexe.

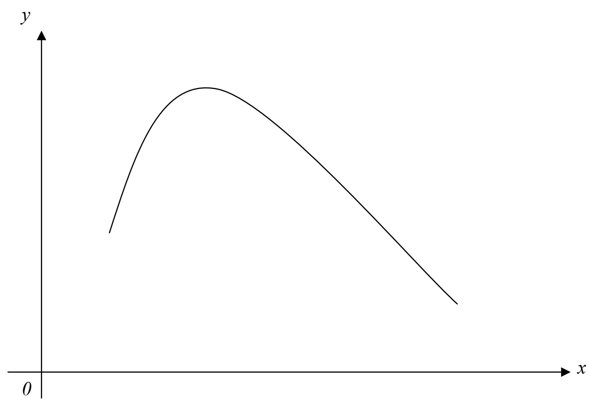


Figure 1 – Exemple d'une fonction coût concave

Ce manuel de cours comprend six chapitres : Le premier chapitre traite les notion de mathématique de base indispensables pour entamer le domaine de l'optimisation. Dans le second chapitre nous allons exposer le problème d'optimisation sans contraintes, et nous montrons par des exercices et des programme Matlab les techniques de résolution appelées de *descentes*. Le troisième chapitre considère le cas des problèmes avec contraintes, nous allons généraliser dans cette partie les techniques maitrisées dans le second chapitre. La programmation linéaire sera étudier dans le chapitre quatre avec des exercices montrant le principe de la méthode dite du *simplexe* et la méthode de résolution graphique. L'avant dernier chapitre considère le problème de la programmation non linéaire. Dans le chapitre six des exemples pratiques en économie, en commerce et en industrie sont présentés avec leurs solutions.

En espérant que ce modeste travail puisse être utile à nos étudiants. l'auteur est ouvert à toute suggestion pour enrichir et compléter ce manuel.

Table des matières

Table des figures

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 :

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

1.1.2 Vecteur

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

$$\mathbf{X} = \begin{bmatrix} x_1 \\ x_2 \\ \cdots \\ x_n \end{bmatrix} \tag{1.1}$$

ou encore de la manière suivante :

$$\mathbf{X} = \begin{bmatrix} x_1 & x_2 & \ldots & x_n \end{bmatrix}' \tag{1.2}$$

— deux vecteurs sont égaux ssi tous leurs éléments sont égaux.

— Le produit scalaire de deux vecteurs **X** et **Y** est défini par :

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

— Les deux vecteurs **X** et **Y** sont orthogonaux si :

$$\mathbf{X}'\mathbf{Y} = 0 \tag{1.4}$$

— La norme d'un vecteur **X** est :

$$|\mathbf{X}| = \sqrt{\mathbf{X}'\mathbf{X}} \tag{1.5}$$

— L'inégalité de *Cauchy-Schwarz* est vérifiée pour les deux vecteurs **X** et **Y** :

$$|\mathbf{X}'\mathbf{Y}| \leq |\mathbf{X}|.|\mathbf{Y} \tag{1.6}$$

1.1.3 Matrice

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

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

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

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

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

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

$$c_{ij} = \sum_{k=1}^{p} a_{ik} b_{kj} \qquad i = 1, ..., m \quad j = 1, ..., n \tag{1.9}$$

La multiplication des matrices requière une attention partculière, le nombre de colonne de la première doit être égale au nombre de ligne de la deuxième. Prenant l'exemple illsutratif suivant : soit un vecteur $\mathbf{X}$ de dimension n alors,

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

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

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

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

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

$$\mathbf{I} = \begin{bmatrix} 1 & 0 & ... & 0 \\ 0 & 1 & ... & 0 \\ ... & ... & ... & ... \\ 0 & 0 & ... & 1 \end{bmatrix} \tag{1.10}$$

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

$$\mathbf{AV} = \lambda\mathbf{V} \tag{1.11}$$

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

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

$$|\mathbf{A} - \lambda\mathbf{I}_n| = 0 \tag{1.12}$$

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

$$|\mathbf{A}| = \prod_{i=1}^{n} \lambda_i \tag{1.13}$$

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 :

$$\mathbf{AA}^{-1} = \mathbf{I}_n \tag{1.14}$$

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

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

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 :

$$\mathbf{A} = \mathbf{A}' \tag{1.16}$$

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 $f(x) > 0, \quad \forall x \in \Re$.
- Une matrice symétrique $\mathbf{A}$ est définie positive ssi $\mathbf{X}'\mathbf{A}\mathbf{X} > 0, \quad \forall \mathbf{X} \neq \mathbf{0}$.
- Une matrice symétrique $\mathbf{A}$ de dimension n est définie positive si toutes les valeurs propres sont positives $\lambda_i > 0, \quad \forall i = 1, ..., n$.
- 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 $\Re$.
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 $\Re^n$, 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 :

$$f(\alpha x_1 + (1 - \alpha) x_2) \leq \alpha f(x_1) + (1 - \alpha) f(x_2) \tag{1.17}$$

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

1.3.2 Ensemble convexe

Un ensemble $\Omega \subset \Re^n$ est dit convexe ssi :

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

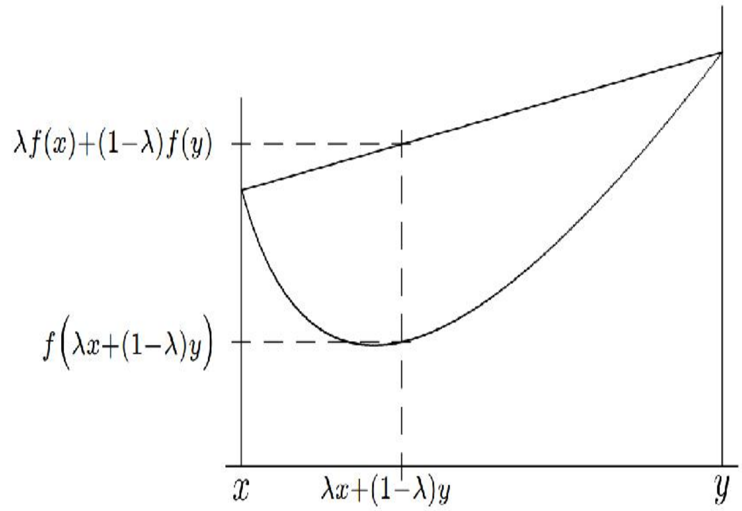


FIGURE 1.1 – Exemple d'une fonction convexe

Exercice 4.
Démontrer que l'ensemble définie par $\mathbf{\Omega} = \{\ \mathbf{X}\ \in\ \Re^n\ : \mathbf{a}'\mathbf{X}\ \geq\ \beta\}$, où $\mathbf{a} \in \Re^n$ et $\beta \in \Re$, est convexe, $\mathbf{\Omega}$ est appelé demi-espace *halfspace*.

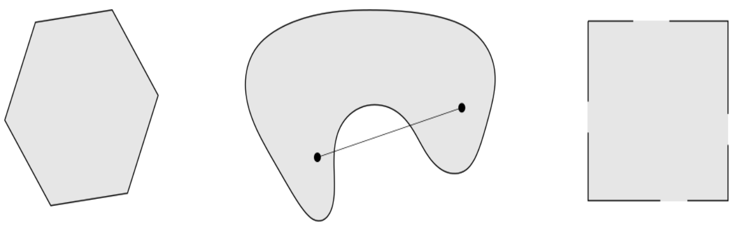


FIGURE 1.2 – Illustration d'ensembles convexes et non vonvexes

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 élements sont les dérivées partielles de la fonction f suivant chaque variable.

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

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 :

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

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 élements sont les dérivées partielles d'ordre deux de la fonction f.

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

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 :

$$\frac{\partial f}{\partial x} = 2x + y, \qquad \longrightarrow \frac{\partial^2 f}{(\partial x)^2} = 2$$
$$\longrightarrow \frac{\partial^2 f}{\partial x \partial y} = 1$$
$$\frac{\partial f}{\partial y} = x + 2y, \qquad \longrightarrow \frac{\partial^2 f}{(\partial y)^2} = 2$$
$$\longrightarrow \frac{\partial^2 f}{\partial y \partial x} = 1$$

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

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-definie positif pour tout $\mathbf{X} \in \Re^n$.

Exercice 5. *Gradient et Hessien*

. Considérons la fonction scalaire suivante :

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

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 Hessein 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 experience. 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 utlise la méthode des moindres carrées, et on choisis la relation suivante $y = ax + b$ avec $(a, b) \in \Re^2$. La fonction objective est la somme des carrées des erreurs (cette fonction est convexe) :

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

a. Claculer 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. Quelle 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 calcule 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 et positive car toutes les valeurs propres sont positives.
3. L'équation caractéristique de A est donnée par la résolution de l'équation :
$$det(A - \lambda I) = 0 \tag{1.21}$$
Dans notre cas, on a en plus $(\lambda - \lambda_1)(\lambda - \lambda_3)(\lambda - \lambda_3) = 0$
4. Le déterminant de A est égale à :
$$\det(A) = \prod_{i=1}^{3} \lambda_i = \lambda_1 \times \lambda_2 \times \lambda_3 = 10 \tag{1.22}$$
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.
$$trace(A) = \sum_{i=1}^{3} \lambda_i = \lambda_1 + \lambda_2 + \lambda_3 = 8 \tag{1.23}$$

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 $\mathbf{X'CX} > 0, \forall \mathbf{X} \in \Re^{*n}$.
(a) Pour la matrice $\mathbf{A}$
$$\mathbf{A} = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$$
$$\begin{aligned} \mathbf{X'AX} &= [x\ y] \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \\ &= 2x^2 + 2xy + 2y^2 \\ &= x^2 + y^2 + (x+y)^2 > 0, \quad \forall (x,y) \in \Re^{*2} \end{aligned}$$

Ainsi, la matrice **A** est définie positive.

(b) Pour la matrice **B**

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

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

Ainsi, la matrice **B** est définie positive.

2. **Utilisation des valeurs propres**

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

(a) Pour la matrice **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 **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 **A** et **B**, sont définies positives.

Solution Ex. 4.
Soit $\mathbf{Z} = \lambda\mathbf{X} + (1-\lambda)\mathbf{Y}$ pour $\lambda \in [0,1]$. Alors :

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

par conséquent, $\mathbf{\Omega}$ 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 :

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

— 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 $\Re$.

3. Les dérivées partielles d'ordre deux et le Hessien 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 fà 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 et donné par le système d'équation suivant :

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

3. Soit :

$$\begin{cases}
S_{11} = \sum_{i=1}^{n} x_i^2 \\
S_{12} = S_{21} = \sum_{i=1}^{n} x_i \\
S_{22} = n
\end{cases}$$

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

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 :

```
%% Construction d'un nuage de points
a=2;
b=3;
N=100;
x=3+10*rand(1,N);
y=a*x+b+3*randn(1,N);
plot(x,y,'.'), hold on
%% Calcul des parametres
S11=sum(x.^2);
S12=sum(x);
S21=S12;
S22=N;
Z1=sum(y.*x);
Z2=sum(y);
```

```

a=(S12*Z2-S22*Z1)/(S12^2-S11*S22);
b=(S21*Z1-S11*Z2)/(S12^2-S11*S22);
%% Dessin des resultats
x_min=min(x);
x_max=max(x);
pas=(x_max-x_min)/20;
x=(x_min:pas:x_max);
y_est=a*(x_min:pas:x_max)+b;

plot(x,y_est,'r','LineWidth',3)
xlabel('x')
ylabel('y')
legend('Nuage de points','Droite estimee')
```

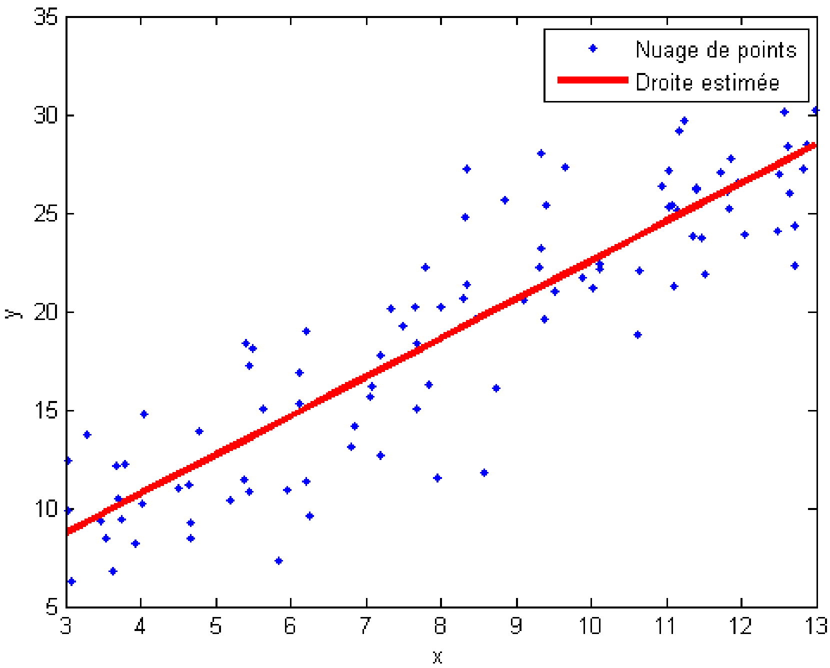


FIGURE 1.3 – Résultat de la régression linéaire

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 Hessien

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

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

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

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 :

$$\mathbf{J}_f = \nabla f' = \begin{bmatrix} \frac{\partial f}{\partial x} \\ \\ \frac{\partial f}{\partial y} \end{bmatrix} = \begin{bmatrix} 1 - 2xy + y^2 \\ \\ -x^2 + 2xy_2 \end{bmatrix} = \begin{bmatrix} 0 \\ \\ 0 \end{bmatrix} \quad \begin{matrix} (1) \\ \\ (2) \end{matrix}$$

De la deuxième équation on tire :

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

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 :

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

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

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 celle.

Chapitre 2

Optimisation sans contraintes : méthodes locales

2.1 Définitions

2.1.1 Définition 1

Soit $f(\mathbf{X})$ une fonction de $\Re^n \rightarrow \Re^n$; Le problème d'optimisation sans contraintes s'écrit de la façon suivante :

$$\text{Minimiser la fonction} f(\mathbf{X}) \text{ avec } \mathbf{X} \in \Re^n$$

La fonction $f(\mathbf{X})$ est souvent appelée :

* fonction coût,
* fonction objectif,
* critère d'optimisation.

2.1.2 Définition 2

On dit que $\mathbf{X}^*$ est un point qui minimise la fonction $f(\mathbf{X})$ sur l'ensemble $\Re^n$ ssi :

$$\forall \mathbf{X} \in \Re^n, \quad f(\mathbf{X}^*) \leq f(\mathbf{X})$$

2.2 Théorèmes d'existence de l'optimum

Avant d'entamer la recherche de optimum d'une fonction, nous devons s'assurer qu'il existe. A cet effet, on utilise une technique similaire à celle de la recherche du maximum ou du minimum d'une fonction scalaire. Ce qui nous amène à utiliser les dérivées partielles d'une fonction à plusieurs variables.

Théorème 1.
Soit la fonction $f(\mathbf{X}) : \Re^n \to \Re$ *différentiable sur* $\Re^n$*, on dit que* $f(\mathbf{X})$ *à un optimum au point* $\mathbf{X}^*$ *si :*

$$\nabla f(\mathbf{X}^*) = 0$$

Remarques

- $\mathbf{X}^*$ est appelé point stationnaire de la fonction $f(\mathbf{X})$.
- La relation $\nabla f(\mathbf{X}^*) = 0$ est aussi appelée équation d'*Euler*.
- Ce théorème n'a aucun sens si la fonction n'est pas différentiable (Exemple fonction en V, comme $f(x) = |x|$).

Théorème 2.
Soit la fonction $f(\mathbf{X}) : \Re^n \to \Re$ *convexe et continument différentiable à l'ordre deux sur* $\Re^n$*, on dit que* $f(\mathbf{X})$ *à un minimum global au point* $\mathbf{X}^*$ *ssi :*

$$\nabla f(\mathbf{X}^*) = 0$$

L'égalité précédente engendre un ensemble de n équations qui peuvent être, dans certains cas, résolues analytiquement.

Cependant, on a souvent recours à résoudre cet ensemble d'équation numériquement et par des algorithmes itératifs. On construit ainsi, une suite de solutions qui converge vers la solution optimale

$$\mathbf{X}_0 \to \mathbf{X}_1 \to ... \to \mathbf{X}_\infty = \mathbf{X}^*$$

Le sens physique de l'infini signifie qu'à un instant donné ou après plusieurs itérations, nous devons accepter une certaine tolérance. Cela va engendrer un critère d'arrêt lorsqu'on atteint la tolérance acceptée ou bien arrêter les itérations dans le cas des boucles infinies.

2.3 Le développement en série de Taylor

Dans le cas unidimensionnel, la dérivée d'une fonction en un point, est comme suit :

$$\begin{aligned} f'(a) &= \lim_{x \to a} \frac{f(x) - f(a)}{x - a} \\ &= \lim_{h \to 0} \frac{f(a+h) - f(a)}{h} \end{aligned}$$

on peut faire l'approximation suivante :

$$f(x) - f(a) \approx f'(a)(x - a)$$

soit :

$$f(x) \approx f(a) + f'(a)(x - a)$$

ou encore

$$f(a + \Delta x) \approx f(a) + f'(a)\Delta x$$

pour plus de précision, on ajoute les termes :

$$\begin{aligned} f(x) &\approx f(a) + f'(a)(x - a) + \frac{1}{2} f''(a)(x - a)^2 + ... + \frac{1}{m!} f^{(m)}(a)(x - a)^m \\ &= \sum_{i=0}^{m} \frac{1}{i!} f^{(i)}(a)(x - a)^i \qquad (2.1) \end{aligned}$$

Cette dernière expression est appelée développement en série de *Taylor* d'ordre m au voisinage du point a.

Dans le cas multidimensionnel, le développement en série de *Taylor* du premier ordre de la fonction $f(\mathbf{X})$ au voisinage du vecteur $\mathbf{A}$, est donné par :

$$\begin{aligned} f(\mathbf{A} + \Delta\mathbf{X}) &\approx f(\mathbf{A}) + \nabla f(\mathbf{A})\Delta\mathbf{X} \qquad (2.2) \\ &= f(\mathbf{A}) + \sum_{i=1}^{n} \left.\frac{\partial f(X)}{\partial x_i}\right|_{x_i = a_i} (x_i - a_i) \qquad (2.3) \end{aligned}$$

Le développement en série de *Taylor* du second ordre est :

$$\begin{aligned} f(\mathbf{A}+\Delta\mathbf{X}) &\approx f(\mathbf{A}) + \nabla f(\mathbf{A})\Delta\mathbf{X} + \frac{1}{2!}\Delta\mathbf{X}'\mathbf{H}f(\mathbf{A})\Delta\mathbf{X} \\ &= f(\mathbf{A}) + \sum_{i=1}^{n} \left.\frac{\partial f(X)}{\partial x_i}\right|_{x_i=a_i} (x_i - a_i) \\ &\quad + \frac{1}{2!}\sum_{i,j=1}^{n} \left.\frac{\partial^2 f(X)}{\partial x_i \partial x_j}\right|_{\substack{x_i=a_i \\ x_j=a_j}} (x_i - a_i)(x_j - a_j) \end{aligned} \tag{2.4}$$

2.4 Résolution analytique (cas unidimensionnel)

Certains problèmes d'optimisation à plusieurs variables, peuvent être amenés à un problème équivalent à une seul variables uniquement. Dans cette partie nous allons examiner certains problèmes de ce type.

Ci-dessous les on présente la démarche pour résoudre ce type d'exercice.

1. Lire attentivement le problème
2. Faire des figures d'illustration
3. Identifier les variables
4. Définir la fonction à optimiser $f(x, y, z, ...)$
5. Reformuler la fonction suivant une seule variable $f(x)$
6. Résoudre l'équation $\frac{df(x)}{dx} = 0$
7. Réécrire la solution suivant les variables initiales.

Pour illustrer cela, nous présentons deux exercices, le premier concerne la détermination des dimension d'une boite et l'objectif étant la maximisation du volume. Le second concerne un problème typique d'une aire maximale.

Exercice 8. *Boite à volume maximal*

.

En utilisant une feuil de dimensions $L \times H$, on veut déterminer les dimensions correspondantes d'une boite (sans couvercle) dont le volume est maximal.

a. Dessiner sur une feuil le croquis de cette boite.

b. Déterminer la fonction à minimiser correspondante à ce problème suivant la hauteur h.

c. Déterminer les dimension de la boite à volume maximal.

d. Donner le résultat pour une page de dimension $A4$ (21×29.7).

e. Trouver le même résultat en utilisant Matlab (dessiner la fonction de volume ainsi que sa dérivée en fonction de la variable h.

Solution Voir le lien suivant 8.

Exercice 9. *Aire maximale*

.
La parabole d'équation :

$$y = -\frac{2}{9}x^2 + 8,$$

coupe l'axe des abscisses en A et B.

Le point $P(x, y)$ se déplace sur la parabole entre A et B. Soit A' la projection du point P sur la droite (AB).

a. Représenter graphiquement cet énoncé.

b. Déterminer les coordonnées du point P pour que l'aire du triangle rectangle $A'PB$ soit maximale.

Solution Voir le lien suivant 9.

2.5 Les méthodes de descente

Les méthodes de descente sont des techniques itératives permettant de trouver le minimum d'une fonction donnée. Le principe est de proposer un point initiale $\mathbf{X}_0$, ensuite on cherche à trouver un autre point $\mathbf{X}_1$ qui vérifie la condition $f(\mathbf{X}_1) < f(\mathbf{X}_0)$. Par la suite, $\mathbf{X}_1$ devient le point initiale pour la nouvelle itération et ainsi de suite jusqu'à l'obtention du point minimale. Le point $\mathbf{X}_1$ peut être mis sous la forme suivante :

$$\mathbf{X}_1 = \mathbf{X}_0 + \rho_0 \Delta \mathbf{X}_0$$

où $\rho_0 \in \Re^{*+}$ est appelé le *pas de descente* et $\Delta \mathbf{X}_0$ un vecteur de $\Re^n$ appelé *direction de descente*. Ainsi, à la première itération on doit trouver ρ_0 et $\Delta \mathbf{X}_0$ qui satisfont la condition $f(\mathbf{X}_1) < f(\mathbf{X}_0)$. Et en générale, à la $k^{ième}$ itération, on a :

$$\mathbf{X}_{k+1} = \mathbf{X}_k + \rho_k \Delta \mathbf{X}_k \tag{2.5}$$

et on doit trouver $\rho_k \in \Re^{*+}$ et $\Delta \mathbf{X}_k$ qui satisfont la condition $f(\mathbf{X}_{k+1}) < f(\mathbf{X}_k)$.

Remarque
Le pas de descente ρ est déterminé soit analytiquement en minimisant la fonction :

$$\underset{s \geq 0}{argmin} f(X + s\Delta X).$$

Soit itérativement en minimisant la fonction $f(\mathbf{X} + \rho \Delta \mathbf{X})$, suivant la variable ρ. L'algorithme correspondant dans ce cas est comme suit :

* Connaissant la direction $\Delta \mathbf{X}$ et soit deux réelles $0 < \alpha < 0.5$ et $0 < \beta < 1$.
* $\rho \leftarrow 1$.
* **Tant que** $f(\mathbf{X} + \rho \Delta \mathbf{X}) > f(\mathbf{X}) + \alpha \rho \nabla f(\mathbf{X}) \Delta \mathbf{X}$,
  faire $\rho :\, = \beta \rho$.

Remarque
En pratique, souvent, on prend :

$$\alpha \in [0.01\ ,\ 0.3]$$
$$\beta \in [0.1\ ,\ 0.8]$$

.

2.6 Méthode du gradient

2.6.1 Principe

La méthode du gradient est une technique de descente (itérative) permettant de trouver le minimum d'une fonction donnée. La façon la plus évidente de trouver une direction est l'utilisation de la dérivée. Ainsi, pour trouver la direction de descente, on utilise le sens inverse de la dérivée. A cet effet, on prend $\Delta\mathbf{X} = -\nabla f(\mathbf{X})'$. La relation itérative correspondante à la méthode du gradient est la suivante :

$$f(\mathbf{X}_{k+1}) \approx f(\mathbf{X}_k) + \rho_k \nabla f(\mathbf{X}_k)\Delta\mathbf{X}_k \tag{2.6}$$

Si ρ_k est constant, cette méthode est qualifiée de méthode du gradient à pas fixe. Lorsque ρ_k varie, on parle de la méthode du gradient à pas variable.

2.6.2 Algorithme

— connaissant la valeur initiale $\mathbf{X}_0$,

— Répéter

1. $\Delta\mathbf{X} \leftarrow -\nabla f(\mathbf{X})'$.
2. Choisir le pas ρ (s'il est fixe on ignore cette étape).
3. Mettre à jours $\mathbf{X} \leftarrow \mathbf{X} + \rho\Delta\mathbf{X}$.

— Arrêter lorsque la condition de convergence est satisfaite.

La condition de convergence est souvent pris comme : $\|\nabla f(\mathbf{X})\|_2 \leq \eta$ avec η un nombre petit et positif.

2.6.3 Exemple

Soit une fonction cout $f(\mathbf{X})$ quadratique en $\Re^2$ sous la forme :

$$f(\mathbf{X}) = x^2 + 3y^2,$$

avec $\mathbf{X} = \begin{bmatrix} x \\ y \end{bmatrix}$. Il est évident que l'origine est le point minimisant cette fonction. Cette fonction est illustrée sur les figures 2.3 et 2.2.

— **Résolution avec un pas fixe**
Dans ce cas on recherche un pas ρ optimale c'est celui qui minimise une certaine fonction, on a :

$$\begin{aligned} \nabla f(\mathbf{X}) &= [\ 2x\ ,\ 6y] \\ \xi(\rho) &= \mathbf{X} - \rho \nabla f(\mathbf{X})' \\ &= \begin{bmatrix} (1-2\rho)x \\ (1-6\rho)y \end{bmatrix} \end{aligned} \tag{2.7}$$

Pour l'obtention du paramètre ρ, considérons la fonction $g(\rho)$ qu'on doit minimiser suivant ρ :

$$\begin{aligned} g(\rho) &= f(\xi(\rho)) \\ &= (1-2\rho)^2 x^2 + (1-6\rho)^2 y^2, \end{aligned} \tag{2.8}$$

La dérivé de cette fonction suivant la variable ρ donne :

$$g'(\rho) = -4(1-2\rho)x^2 - 36(1-6\rho)y^2$$

la paramètre ρ qui minimise la fonction $g(\rho)$ est celui qui satisfait $g'(\rho) = 0$, soit :

$$\rho = \frac{x^2 + 9y^2}{2x^2 + 54y^2} \tag{2.9}$$

L'équation itérative de résolution de ce problème est donc :

— Initialisation $\mathbf{X} = \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 3 \\ 5 \end{bmatrix}$

— Répéter

1. $\Delta\mathbf{X} \leftarrow -\begin{bmatrix} 2x \\ 6y \end{bmatrix}$.
2. $\rho = \frac{x^2+9y^2}{2x^2+54y^2}$
3. $\mathbf{X} \leftarrow \mathbf{X} + \rho\Delta\mathbf{X}$.

— Arrêter lorsque $\|\Delta\mathbf{X}\|_2^2 \leq 10^{-3}$.

— **Résolution avec un pas itératif**
Lorsqu'on choisis un pas itératif alors, la deuxième étape dans l'algorithme est remplacée par la boucle suivante :
soit $\alpha = 0.25$, $\beta = 0.5$ $\rho = 1$,
alors :
pour une direction donnée $\Delta\mathbf{X} = -\begin{bmatrix} 2\mathbf{X}(1) \\ 6\mathbf{X}(2) \end{bmatrix}$,

— faire :

$$\mathbf{Y} = \mathbf{X} + \rho\Delta\mathbf{X},$$
$$f(\mathbf{Y}) = \mathbf{Y}(1)^2 + 3\mathbf{Y}(2)^2,$$
$$\nabla f(\mathbf{X}) = [2\mathbf{X}(1)\ ,\ 6\mathbf{X}(2)],$$

— si $f(\mathbf{Y}) < f(\mathbf{X}) + \alpha\rho\nabla f(\mathbf{X})\Delta\mathbf{X}$
fin
— sinon
$\rho = \beta\rho$
— fin

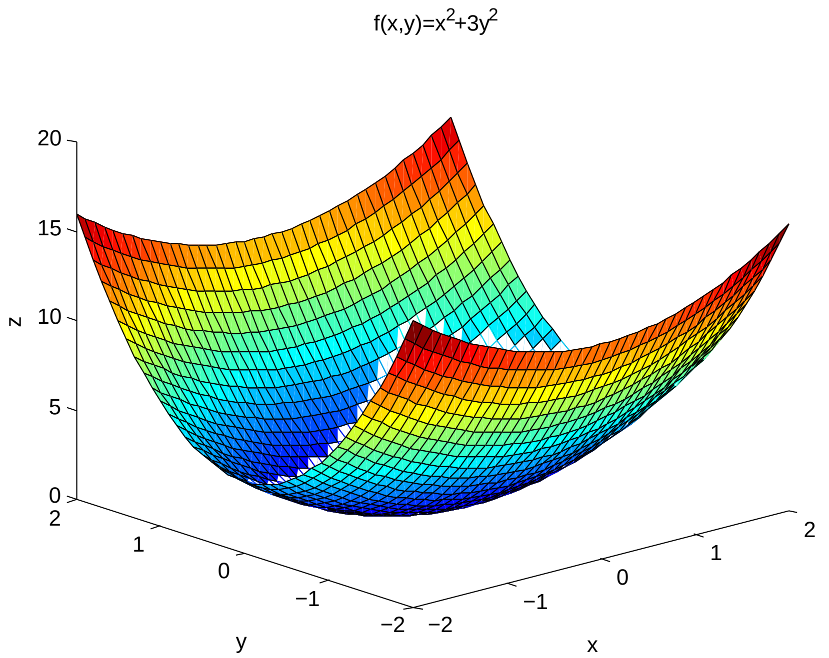


FIGURE 2.1 – Allure de la fonction $f(x, y) = x^2 + 3y^2$

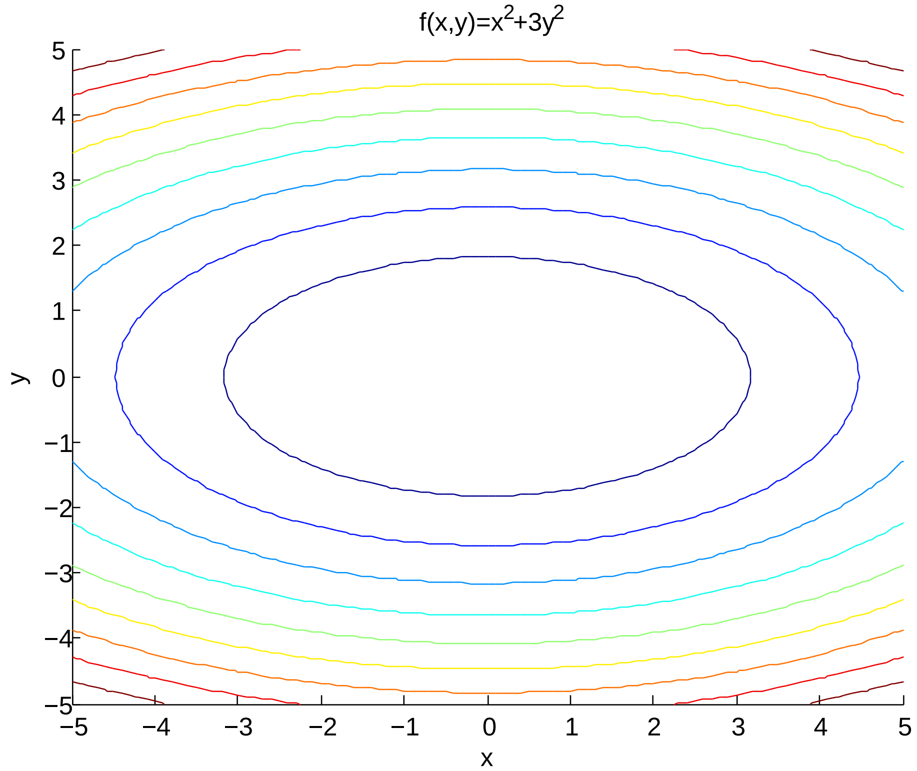


FIGURE 2.2 – Vue de haut de la fonction $f(x,y) = x^2 + 3y^2$

2.6.4 Etude de la convergence

On peut démontrer que si une fonction $f(\mathbf{X})$ est strictement convexe sur un ensemble Ω, alors il existe deux constantes $m, M \in \Re^{*+}$ qui satisfont :

$$m\mathbf{I} \leq \mathbf{H}f(\mathbf{X}) \leq M\mathbf{I} \tag{2.10}$$

Remarque :
La constante m est souvent considérée comme étant la valeur propre la plus petite, alors que la constante M est souvent considérée comme la valeur propre la plus grande.

Dans ce cas, on peut démontrer une limite quadratique supérieur par la relation suivante :

$$f(\mathbf{X}_{k+1}) \leq f(\mathbf{X}_k) - \rho_k \left\|\nabla f(\mathbf{X}_k)\right\|_2^2 + \frac{M\rho_k^2}{2} \left\|\nabla f(\mathbf{X}_k)\right\|_2^2 \tag{2.11}$$

Exercice 10.
Soit la fonction réelle à deux variables :

$$f(x,y) = x(2+x) + y^2 + 4$$

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 l'optimum de cette fonction, quelle est sa nature ?
8. Utiliser la méthode du gradient pour déterminer cet extremum.
9. Calculer le pas optimale.
10. Donner l'algorithme correspondant à cette fonction pour le calcul du pas de façon itérative.

2.7 Méthode de Newton

2.7.1 Principe

La méthode de Newton fait partie des méthodes de descente. Cette technique cherche à minimiser le développement en série de Fourier du second ordre de la fonction $f(\mathbf{X}) \in C^2$ suivant la quantité $\Delta\mathbf{X}$:

$$f(\mathbf{X}+\Delta\mathbf{X}) \approx f(\mathbf{X}) + \nabla f(\mathbf{X})\Delta\mathbf{X} + \frac{1}{2!}\Delta\mathbf{X}'\mathbf{H}_f(\mathbf{X})\Delta\mathbf{X}$$

le $\Delta\mathbf{X}$ qui minimise la partie droite est comme suit :

$$\Delta\mathbf{X} = -\mathbf{H}_f(\mathbf{X})^{-1}\nabla f(\mathbf{X})'$$

pour la convergence cette méthode on calcule la quantité positive :

$$\frac{1}{2}\nabla f(\mathbf{X})'\mathbf{H}_f(\mathbf{X})^{-1}\nabla f(\mathbf{X}),$$

et on vérifie qu'elle est inférieur ou égale à une certaine précision ϵ.

2.7.2 Algorithme

— Connaissant la valeur initiale $\mathbf{X}_0$, et la précision ϵ

— Répéter

1. $\Delta\mathbf{X} \leftarrow -\mathbf{H}_f(\mathbf{X})^{-1}\nabla f(\mathbf{X})'$,
2. on arrêter si $\frac{1}{2}\nabla f(\mathbf{X})'\mathbf{H}_f(\mathbf{X})^{-1}\nabla f(\mathbf{X}) \leq \epsilon$.
3. Choisir le pas ρ (s'il est fixe on ignore cette étape).
4. Mettre à jours $\mathbf{X} \leftarrow \mathbf{X} + \rho\Delta\mathbf{X}$.

2.7.3 Exemple

Reprenant l'exemple de la section précédente, soit une fonction coût $f(\mathbf{X})$ quadratique en $\Re^2$:

$$f(\mathbf{X}) = x^2 + 3y^2,$$

avec $\mathbf{X} = \begin{bmatrix} x \\ y \end{bmatrix}$.

Donner la forme quadratique correspondante, puis déterminer le minimum par la méthode de Newton.

Solution

Les différentes termes utilisés dans l'algorithme sont :

$$\nabla f(\mathbf{X}) = [2x \ , \ 6y]$$

$$\mathbf{H}_f = \begin{bmatrix} 2 & 0 \\ 0 & 6 \end{bmatrix} \quad \longrightarrow \quad \mathbf{H}_f^{-1} = \begin{bmatrix} \frac{1}{2} & 0 \\ 0 & \frac{1}{6} \end{bmatrix}$$

$$\Delta\mathbf{X} = -\mathbf{H}_f^{-1}\nabla f(\mathbf{X})' = -\begin{bmatrix} x \\ y \end{bmatrix}$$

$$\frac{1}{2}\nabla f(\mathbf{X})'\mathbf{H}_f(\mathbf{X})^{-1}\nabla f(\mathbf{X}) = \frac{1}{2}\begin{bmatrix} 2x & 6y \end{bmatrix}\begin{bmatrix} \frac{1}{2} & 0 \\ 0 & \frac{1}{6} \end{bmatrix}\begin{bmatrix} 2x \\ 6y \end{bmatrix}$$

$$= \begin{bmatrix} x^2 \\ 3y^2 \end{bmatrix}$$

2.8 La méthode de Levenberg-Marquardt

2.8.1 Introduction

Crée en 1944 par *Kenneth Levenberg* pour la résolution d'un problème des moindres carrés non linéaires, voir description plus bas 2.8.4. En 1963, *Donald Marquardt* a mis à niveau cette technique. Elle est souvent utilisée pour résoudre le problème de l'ajustement de courbe (curve-fitting).

2.8.2 Principe

Cette méthode présente une combinaison entre la méthode du gradient et la méthode de Newton. La méthode de Levenberg-Marquardt se comporte comme le gradient lorsque les paramètres se trouvent loin de la valeurs optimale et agit comme la méthode de Newton lorsqu'on s'approche de la solution optimale.

Cette technique se base sur l'expression itérative suivante :

$$\mathbf{X}_{k+1} = \mathbf{X}_{k+1} - (\mathbf{H}_f(\mathbf{X}_k) + \rho_k \mathbf{I})^{-1} \nabla f(\mathbf{X}_{k+1})$$

2.8.3 Algorithme

1. $\nu = 2$;
   $\mathbf{X} = \mathbf{X}_0$
   $\mathbf{H} = \mathbf{J}(\mathbf{X})'\mathbf{J}(\mathbf{X})$;
   $\mathbf{g} = \mathbf{J}(\mathbf{X})' r(\mathbf{X})$;
   $sol = (\|\mathbf{g}\| \leq \epsilon_1)$, résultat logique 0 ou 1 ;
   $\rho = \nu \times max(H_{ii})$, on prend l'élément le plus grand de la matrice.
2. **Tant que** $(sol == 0)$
   — Résoudre $(\mathbf{H} + \rho \mathbf{I})\delta \mathbf{X} = -\mathbf{g}$
   — **si**
   $\|\delta \mathbf{X}\| \leq \epsilon_2(\|\mathbf{X}\| + \epsilon_2)$
   $sol = 1$
   — **sinon**
   $\mathbf{X}_{new} = \mathbf{X} + \delta \mathbf{X}$
   $\kappa = \frac{f(\mathbf{X}) - f(\mathbf{X}_{new})}{l(\mathbf{0}) - l(\delta \mathbf{X})}$

— **si** $\kappa > 0$
  — $\mathbf{X} = \mathbf{X}_{new}$
  — $\mathbf{H} = \mathbf{J}(\mathbf{X})'\mathbf{J}(\mathbf{X})$;
    $\mathbf{g} = \mathbf{J}(\mathbf{X})'r(\mathbf{X})$
  — $sol = (\|\mathbf{g}\| \leq \epsilon_1)$
  — $\rho = \rho * max\{\frac{1}{3}, 1-(2\kappa-1)^3\}$;
    $\nu = 2;$
— **sinon**
  $\rho = \rho\nu$;
  $\nu = 2 * \nu$

3. **fin**

2.8.4 Problème des moindres carrées

Le problème des moindres carrées est formulé comme suit :

$$f(\mathbf{X}) = \frac{1}{2}\sum_{i=1}^{m} r_i^2(\mathbf{X})$$

où les fonctions $r_i(\mathbf{X})$ représentent les résidus et $m \geq n$. La fonction $f(\mathbf{X})$ peut être mise sous la forme :

$$f(\mathbf{X}) = \frac{1}{2}\left\|r(\mathbf{X})\right\|^2$$

Remarque
Dans le cas linéaire la solution optimale peut être donnée par la relation :

$$\mathbf{X}^* = -(J'J)^{-1}J'r(\mathbf{0})$$

2.9 La méthode Quasi-Newton

Problème des matrices mal conditionnées
Soit les matrices $\mathbf{A}$ et $\mathbf{B}$:

$$\mathbf{A} = \begin{bmatrix} 1.2644 & 1.8426 \\ 2.7640 & 4.0284 \end{bmatrix} \longrightarrow \mathbf{A}^{-1} = \begin{bmatrix} 7725.7 & -3533.8 \\ -5300.8 & 2424.9 \end{bmatrix}$$

$$\mathbf{B} = \begin{bmatrix} 1.2645 & 1.8426 \\ 2.7640 & 4.0284 \end{bmatrix} \longrightarrow \mathbf{B}^{-1} = \begin{bmatrix} 4358.5 & -1993.6 \\ -2990.4 & 1368.1 \end{bmatrix}$$

Nous remarquons que $\mathbf{B}$ ne diffère de $\mathbf{A}$ que sur l'élément de position (1,1) et cette différence est de l'ordre de 10^{-4}.

Si on a les deux problèmes linéaires à résoudre, $AX = [1\ 1]'$ et $BX = [1\ 1]'$, cela donne respectivement les solutions suivantes :$X = [4191.8\ \ -2.8759]'$ et $X = [2364.8\ \ -1622.3]'$.

On remarque que pour une petit erreur de 10^{-4} la solution varie énormément. La cause principale de cette différence peut être expliqué par rapport aux valeurs propres correspondantes des deux matrices $\mathbf{A}$ et $\mathbf{B}$ qui sont $[0.0001\ 5.2927]'$ et $[0.0002\ 5.2927]'$, respectivement.

On peut dire qu'il y a un certain déséquilibre entre les valeurs propres de ces matrices. On dit qu'on a des matrices mal conditionnées.

2.9.1 Introduction

La méthode de Newton requière l'inversion du Hessien. L'inversion des matrices est une opération qui est souvent lourde et peut être la cause d'une instabilité numérique si la matrice est mal conditionnée. Dans ces conditions, on applique la méthode Quasi-Newton.

2.9.2 Principe

Le principe de cette méthode est de remplacer la matrice inverse du Hessien $\mathbf{H}_f(\mathbf{X}_k)^{-1}$ par une matrice symétrique et semi-définie positive qu'on appel $\mathbf{Q}_f(\mathbf{X}_k)$ qu'on note $\mathbf{Q}_k$. Le Hessien peut être approximé par :

$$\mathbf{H}_f(\mathbf{X}_{k+1})\,(\mathbf{X}_{k+1} - \mathbf{X}_k) = \nabla f(\mathbf{X}_{k+1}) - \nabla f(\mathbf{X}_k)$$

Si on multiplie cette équation par l'inverse du Hessien, on s'aperçois que la matrice $\mathbf{Q}_{k+1}$ doit satisfaire la condition suivante :

$$\mathbf{Q}_{k+1}\,(\nabla f(\mathbf{X}_{k+1}) - \nabla f(\mathbf{X}_k)) = \mathbf{X}_{k+1} - \mathbf{X}_k$$

De cette façon, on a éviter le passage par l'inversion d'une matrice. mais on a recours à la résolution d'un système d'équations linéaires.

Remarques
Cette technique est très utilisée dans beaucoup de problèmes, cependant, elle nécessite un nombre d'itérations très important si le problème est mâle conditionné ou si une mauvaise estimée initiale du Hessien est utilisée.

La première méthode quasi-Newton est appelée DFP, elle a été découverte en 1959 par Davidson, ensuite Fletcher et Powell ont exploité ces propriétés durant la décennie suivante. Cette méthode exploite la variation du gradient entre deux itérations. L'une des formes récursives les plus utilisées est celle nommée BFGS (Broyden, Fletcher, Goldfarb et Shanno) développée en 1970. Dans cette méthode on ne cherche pas à résoudre un système d'équations linéaires mais on utilise des formules algébriques (multiplication de vecteurs et de matrices). alors on à :

$$\begin{aligned}
\mathbf{Q}_{k+1} &= (I - \rho_k d\mathbf{X}_k \mathbf{Y}'_k)\,\mathbf{Q}_k\,(I - \rho_k d\mathbf{X}_k \mathbf{Y}'_k)' + d\mathbf{X}_k \rho_k d\mathbf{X}'_k \\
d\mathbf{X}_k &= \mathbf{X}_{k+1} - \mathbf{X}_k \\
\mathbf{Y}_k &= \nabla f(\mathbf{X}_{k+1}) - \nabla f(\mathbf{X}_k)
\end{aligned}$$

Lorsqu'on traite un problème de grande dimension, il n'est pas possible de sauvegarder toutes les données à cet effet, il existe une version limitée en espace de cette méthode dont le nom est L-BFGS (Limited-memory BFGS).

2.10 Méthodes du gradient conjugué

L'inconvénient de la méthode du gradient, c'est quelle effectue, souvent, beaucoup de zigzags avant d'atteindre la solution souhaité, voir figure 2.3 . Afin d'éliminer ce problème, la méthode du gradient conjugué, cherche une nouvelle direction à chaque itération.

2.10.1 Principe

A l'itération k, on cherche l'inverse du gradient et on lui ajoute une combinaison linéaire des anciens vecteurs de direction, afin d'obtenir un nouveau vecteur conjugué de direction.

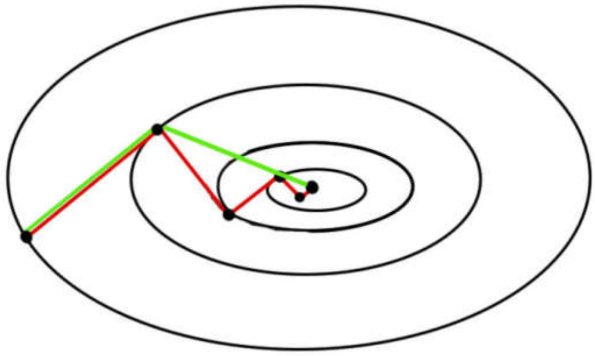


FIGURE 2.3 – Problème de rapidité de convergence de la méthode du gradient

2.10.2 Problème introductif

- **Question** : Comment résoudre l'équation à une seule variable :

$$ax - b = 0,$$

en utilisant les techniques de descente ?
- **Indication** La réponse passe par l'utilisation de l'équation :

$$\frac{1}{2}ax^2 - bx - c = 0$$

- **Réponse** : Cette équation peut être vue comme étant un problème de minimisation de la forme quadratique :

$$\frac{1}{2}ax^2 - bx - c = 0.$$

Effectivement, la dérivée de cette équation donne notre équation $ax - b = 0$.

2.10.3 Définition

Deux vecteurs $\mathbf{U}$ et $\mathbf{V}$ dans $\Re^n$ sont conjugués par rapport à une forme bilinéaire de matrice symétrique définie positive $\mathbf{Q}$ (on dit aussi qu'ils sont $\mathbf{Q}$-orthogonaux), si :

$$\mathbf{U}'\mathbf{Q}\mathbf{V} = 0$$

Remarques

1. Pour $\mathbf{Q} = \mathbf{I}$ on retrouve la définition classique de l'orthogonalité.
2. Un ensemble finie de vecteurs $\mathbf{U}_i\}_{i=1}^{k}$ est dit $\mathbf{Q}$-Orthogonal si :

$$\mathbf{U}_i \mathbf{Q} \mathbf{U}_j = 0 \quad \forall i \neq j.$$

Ces vecteurs sont en plus linéairement indépendants.

2.10.4 Résolution d'un système d'équations linéaire

La résolution d'un système d'équation linéaire de la forme $\mathbf{QX} = \mathbf{B}$ peut être vue comme un problème d'optimisation (recherche du minimum) d'une fonction quadratique de la forme :

$$\min \left\{ \frac{1}{2} \mathbf{X}^T \mathbf{Q} \mathbf{X} - \mathbf{B}^T \mathbf{X} \right\}$$

On note que la solution optimale de ce problème d'optimisation n'est autre que la solution du système d'équations linéaire.

Chaque solution peut être mis sous la forme d'une combinaison linéaire d'un ensemble de vecteurs $\mathbf{Q}$-orthogonal $\{\mathbf{U}_i\}_{i=1}^{n}$, alors :

$$\mathbf{X}^* = \alpha_1 \mathbf{U}_1 + \alpha_1 \mathbf{U}_1 + \ldots + \alpha_n \mathbf{U}_n$$

pour déterminer le coefficient α_i, on multiplie à gauche par la quantité $\mathbf{U}_i^T \mathbf{Q}$. Cela nous donne :

$$\alpha_i = \frac{\mathbf{U}_i^T \mathbf{Q} \mathbf{X}^*}{\mathbf{U}_i^T \mathbf{Q} \mathbf{U}_i} = \frac{\mathbf{U}_i^T \mathbf{B}}{\mathbf{U}_i^T \mathbf{Q} \mathbf{U}_i}$$

Cela montre que la solution $\mathbf{X}^*$ peut être obtenue par :

$$\mathbf{X}^* = \sum_{i=1}^{n} \frac{\mathbf{U}_i^T \mathbf{B}}{\mathbf{U}_i^T \mathbf{Q} \mathbf{U}_i} \mathbf{U}_i$$

Suivant cette expression, on peut dire que la solution $\mathbf{X}^*$ peut être considérer comme le résultat d'un processus itératif de n étapes où à chaque étape on ajoute un terme $\alpha_i \mathbf{U}_i$.

2.10.5 Théorème des directions conjuguées

Soit l'ensemble $\{\mathbf{U}_i\}_{i=0}^{n-1}$, $\mathbf{Q}$-orthogonale. Pour un vecteur $\mathbf{X}_0$ quelconque, la séquence $\{\mathbf{X}_i\}$ générée suivant l'expression :

$$\mathbf{X}_{k+1} = \mathbf{X}_k + \alpha_k \mathbf{U}_k$$

où

$$\alpha_k = -\frac{g_k^T \mathbf{U}_k}{\mathbf{U}_k^T \mathbf{Q} \mathbf{U}_k}$$

et

$$g_k = \mathbf{Q}\mathbf{X}_k - \mathbf{B}$$

converge vers la solution unique $\mathbf{X}^*$ de $\mathbf{QX} = \mathbf{B}$ après n itérations (c-à-d $\mathbf{X}_n = \mathbf{X}^*$).

2.10.6 Algorithme

1. Pour $\mathbf{X}_0$ donné, on prend $\mathbf{U}_0 = -g_0 = \mathbf{B} - \mathbf{Q}\mathbf{X}_0$
2. $\alpha_k = -\frac{g_k^T \mathbf{U}_k}{\mathbf{U}_k^T \mathbf{Q} \mathbf{U}_k}$
3. $\mathbf{X}_{k+1} = \mathbf{X}_k + \alpha_k \mathbf{U}_k$
4. $g_{k+1} = \mathbf{Q}\mathbf{X}_{k+1} - \mathbf{B}$
5. $\beta_k = \frac{g_{k+1}^T \mathbf{Q} \mathbf{U}_k}{\mathbf{U}_k^T \mathbf{Q} \mathbf{U}_k}$
6. $\mathbf{U}_{k+1} = -g_{k+1} + \beta_k \mathbf{U}_k$

La direction à l'itération k représente une combinaison linéaire des anciennes directions de descente.

2.11 Solutions des exercices

Solution Ex. 8.

1. La hauteur est donnée par la variable h. suivant la figure 2.4.

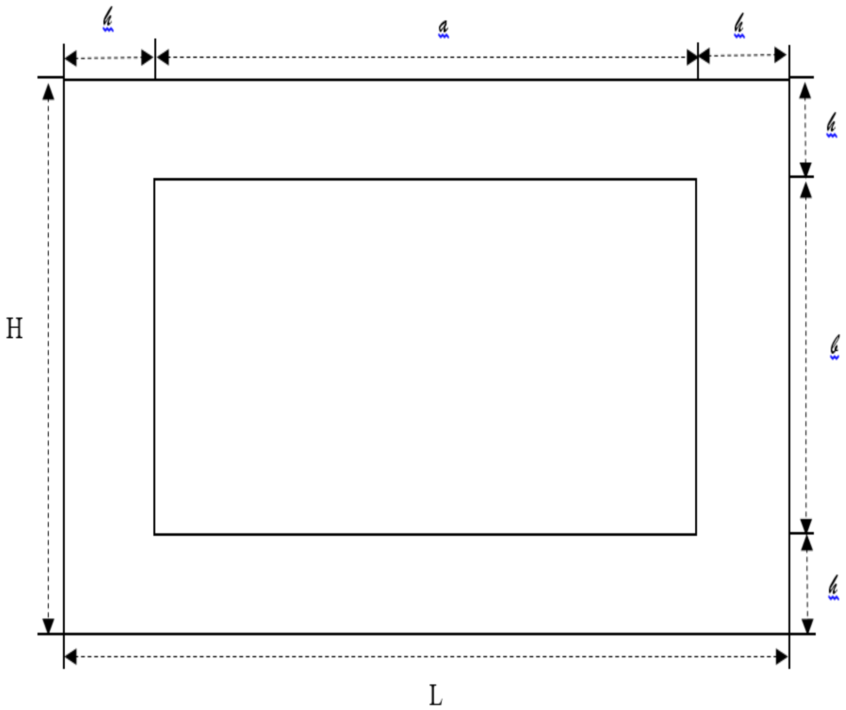


FIGURE 2.4 – Croquis de la boite

2. Le volume de la boite est donné par la relation :

$$\begin{aligned} V &= a \times b \times h \\ a &= L - 2h \\ b &= H - 2h \\ V(h) &= (L-2h)(H-2h)h \\ V(h) &= 4h^3 - 2(L+H)h^2 + LHh \end{aligned}$$

3. La dérivation de $V(h)$ donne :

$$\frac{dV(h)}{dh} = 12h^2 - 4(L+H)h + LH$$

Le maximum de cette fonction vérifie la solution de l'équation :

$$\frac{dV(h)}{dh} = 0.$$

Les solutions de cette équation sont $h_1 = 40.4mm$ et $h_2 = 128.58mm$. Suivant le problème posé, la hauteur h ne peut pas être plus grande que $\frac{1}{2}min(H, L)$.

Les dimensions de la boites sont alors :

- La hauteur $h = 40.4\ mm$.
- La longueur $a = H - 2h = 216.2\ mm$.
- La largeur $b = L - 2h = 129.2\ mm$.

Le volume maximale correspondant est alors $V = 1.1285 \times 10^6\ mm^3$. Sachant quen $1L = 1000cm^3$ le volume en litre vaut $V = 1.128\ L$

Programme Matlab

La figure 2.5, présente le tracé de la fonction $V(h)$ pour h allant de 0 à $\frac{1}{2}min(H, L)$. Cette figure est le résultats du programme MATLAB . .

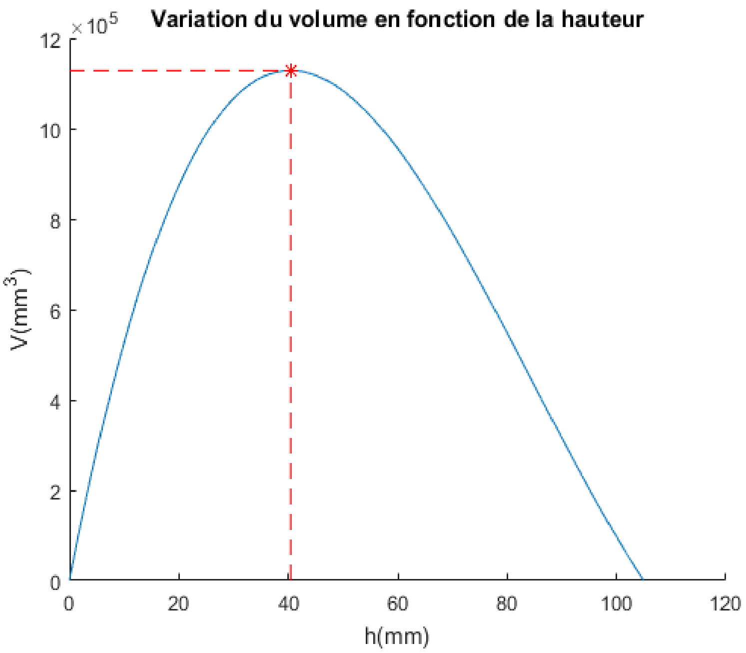


FIGURE 2.5 – Volume en fonction de la hauteur

```
L=210;
H=297;
a=min(H,L)/2;

h=0:a;
V=4*h.^3-2*(L+H)*h.^2+L*H*h;
figure,hold on
plot(h,V)
xlabel('h(mm)')
ylabel('V(mm^3)')
title('Variation du volume en fonction de la auteur')
[V_max, i_max]=max(V);
h_max=h(i_max);
plot(h_max,V_max,'*r')
line([h_max h_max],[0 V_max],'LineStyle','--','Color','r')
line([0 h_max],[V_max V_max],'LineStyle','--','Color','r')
Vp=12*h.^2-4*(L+H)*h+L*H;
h_zero=find(abs(Vp)<1)
Vp_max= Vp(i_max);
subplot(2,1,2),hold on, plot(h, Vp)
xlabel('h(mm)'),ylabel('V'''),
title('Variation de la derivee en fonction de la hauteur')
```

Solution Ex. 9. 1. l'Aire du triangle rectangle est égal à :

$$S(x,y) = (\frac{1}{2}AB + x) \times y$$

. Les coordonnées des points A et B sont $A(6,0)$ et $B(-6,0)$ et suivant la fonction on obtient la surface :

$$S = (-\frac{2}{9}x^2 + 8) \times (6 + x)$$
$$= -\frac{2}{9}x^3 - \frac{4}{3}x^2 + 8x + 48$$

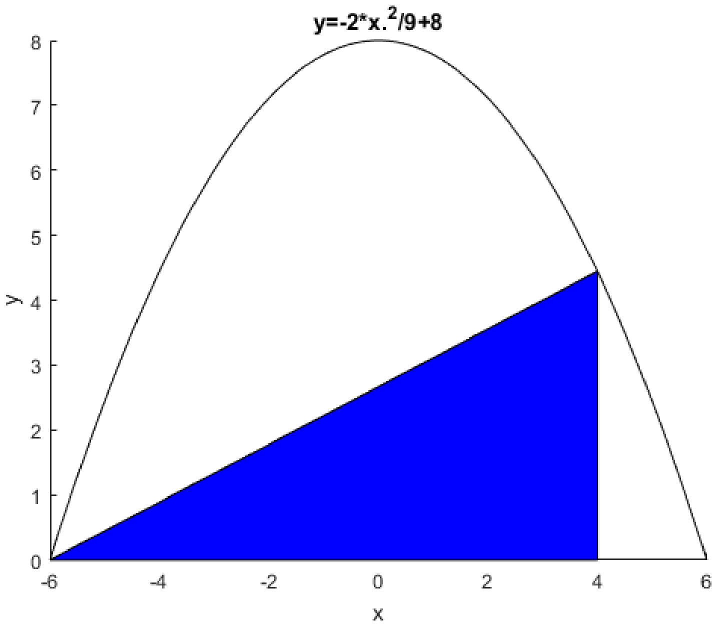


FIGURE 2.6 – Dessin du triangle rectangle

La dérivé de la surface suivant la variable x donne :

$$\frac{dS(x)}{dx} = -\frac{2}{3}x^2 - \frac{8}{3}x + 8$$

Le maximum de cette fonction est atteint pour $x = 2$. Les deux figures 2.6 et 2.7, sont les résultats du programme MATLAB ci-dessus. Suivant l'allure de la dérivée figure 2.7, il est claire que le maximum est atteint au voisinage du point dont l'abscisse est 2.

```
figure , hold on
x=-6:0.01:6;
y=-2*x.^2/9+8;
plot(x,y,'k');
xlabel('x')
```

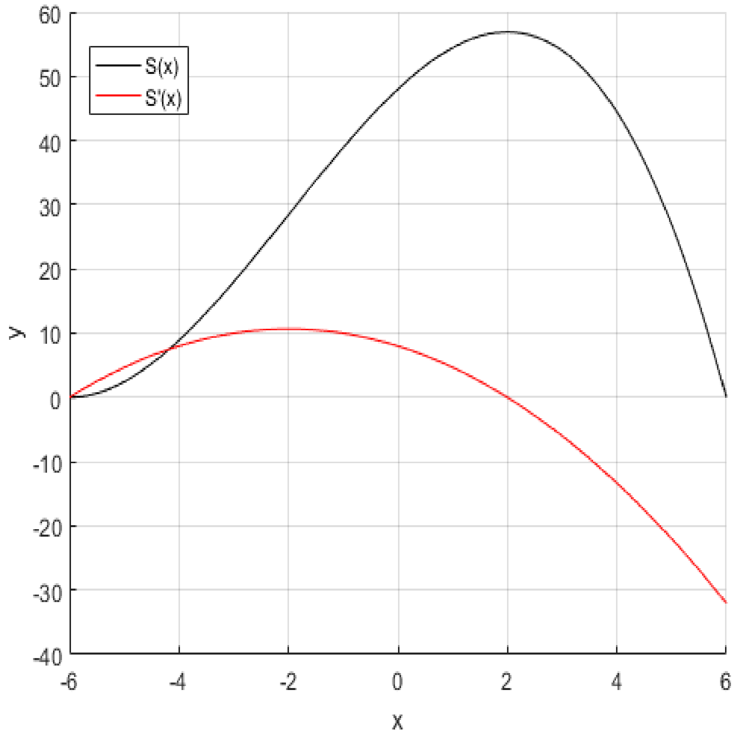


FIGURE 2.7 – Allure de la surface ainsi que sa dérivée suivant x.

```
ylabel('y')
title('y=-2*x.^2/9+8')

x1=4;
y1=-2*x1^2/9+8;
plot([x1, x1],[0 y1],'b')
plot([-6 x1],[0 y1],'b')

x2=4;
y2=-2*x2^2/9+8;
fill([-6 x2 x2 -6],[0 0 y2 0],'b')

figure, hold on
S=-2*x.^3/9-4*x.^2/3+8*x+48;
Sp=-2*x.^2/3-8*x/3+8;
plot(x,S,'k')
```

```
plot(x,Sp,'r')
legend('S(x)','S''(x)')
grid on
xlabel('x')
ylabel('y')
```

Solution Ex. 10.
Soit la fonction réelle à deux variables :

$$f(x,y) = x^2 + 2x + y^2 + 4$$

1. Calcule du Jacobien,

$$\mathbf{J}_f = \nabla f' = \begin{bmatrix} \frac{\partial f}{\partial x} \\ \\ \frac{\partial f}{\partial y} \end{bmatrix} = \begin{bmatrix} 2x+2 \\ \\ 2y \end{bmatrix}$$

2. Le Hessien

$$\mathbf{H}_f = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$$

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

$$\det(H - \lambda I) = 0 \quad \rightarrow \quad (2-\lambda)^2 = 0$$

4. Les valeurs propres de la matrice $\mathbf{H}_f$ sont les solutions de l'équation caractéristique de cette matrice, alors :

$$(2-\lambda)^2 = 0 \quad \rightarrow \quad \lambda_1 = \lambda_2 = 2$$

5. La matrice est définie positive, car les valeurs propres sont positives.
6. Le déterminant est donné par le produit des valeurs propres :

$$\det(\mathbf{H}_f) = \lambda_1 \times \lambda_2 = 4$$

7. L'optimum de cette fonction est obtenu analytiquement lorsqu'on résous le système d'équations suivant :

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

soit :

$$x = 1 \quad et \quad y = 0$$

l'optimum est $f(-1, 0) = 3$, et présente un minimum car on peut facilement démonter l'une des propriétés suivantes :

— $\det(\mathbf{H}_f) > 0$ et $H_{11} > 0$
— Toutes les valeurs propres sont positives
— La matrice est définie positive

8. Par la suite on déterminera la méthode de résolution à pas fixe uniquement. Dans ce cas on recherche un pas ρ optimale qui minimise une certaine fonction, on a :

$$\nabla f(\mathbf{X}) = [\ 2x+2\ ,\ 2y]$$
$$\xi(\rho) = \mathbf{X} - \rho \nabla f(\mathbf{X})' = \begin{bmatrix} x - 2\rho(x+1) \\ \\ y - 2\rho y \end{bmatrix}$$

Pour l'obtention du paramètre ρ, considérons la fonction $g(\rho)$ qu'on doit minimiser suivant ρ :

$$\begin{aligned} g(\rho) &= f(\xi(\rho)) \\ &= [x - 2\rho(x+1)]^2 + 2[x - 2\rho(x+1)] + [y - 2\rho y]^2 + 4 \qquad (2.12) \\ g'(\rho) &= -4(x+1)[x - 2\rho(x+1)] - 4(x+1) - 4y[y - 2\rho y] \end{aligned}$$

la paramètre ρ qui minimise la fonction $g(\rho)$ est celui qui satisfait $g'(\rho) = 0$, soit :

$$\rho = \frac{1}{2} \times \frac{x^2 + 2x + y^2 + 1}{(x+1)^2 + y^2} = \frac{1}{2} \qquad (2.13)$$

Ci-dessous, l'algorithme pour résoudre numériquement ce problème avec une précision de $\epsilon = 10^{-3}$ et en prenant comme point initial $\mathbf{X}_0 = [3\ ,\ 5]'$:

— Initialisation
  $\mathbf{X} = [x\ ,\ y]' = [3\ ,\ 5]'$
  $\rho = \frac{1}{2}$
— Répéter

(a) $\Delta\mathbf{X} \leftarrow -[\ 2x + 2\ ,\ 2y]'$.

(b) $\mathbf{X} \leftarrow \mathbf{X} + \rho\Delta\mathbf{X}$.

(c) Arrêter lorsque $\|\Delta\mathbf{X}\|_2^2 = (2x+2)^2 + 4y^2 \leq 10^{-3}$.

Chapitre 3

Optimisation avec contraintes : méthodes globales

3.1 Définitions

Un problème d'optimisation avec ou sous contrainte est donné sous la forme suivante :

$$\begin{cases} min\ f(\mathbf{X}) \quad \mathbf{X} \in \Re^n \\ \qquad h_i(\mathbf{X}) = 0 \quad i = 1, ..., p \\ \qquad g_j(\mathbf{X}) = 0 \quad j = 1, ..., q \end{cases}$$

Cette forme est appelée forme standard.

Exemple

—

$$\begin{cases} \min f(x) = x^2 \\ \qquad \text{sous la contrainte} \\ \qquad x \geq 1, \end{cases}$$

—

$$\begin{cases} \min f(x, y) = x^2 + y^2 \\ \qquad \text{sous la contrainte} \\ \qquad x + y = 5 \end{cases}$$

3.1.1 Le Lagrangien du problème

Le lagrangien d'un problème d'optimisation sous contraintes se met sous la forme :

$$L(\mathbf{X}, \lambda, \gamma) = f(\mathbf{X}) + \sum_{i=1}^{p} \lambda_i h_i(\mathbf{X}) + \sum_{j=1}^{q} \gamma_i g_i(\mathbf{X}) \tag{3.1}$$

où les vecteurs $\lambda = [\lambda_1,\ \lambda_2,\ ...,\ \lambda_p]'$ et $\gamma = [\gamma_1,\ \gamma_2,\ ...,\ \gamma_q]'$ sont appelés les multiplicateurs de Lagrange.

Remarque
L'intérêt du Lagrangien est de ramener un problème d'optimisation sous contraintes en un problème d'optimisation sans contraintes.

Afin de rendre simple le développement théorique, nous allons examiné les contraintes d'une façon isoler. Ainsi on traite les problèmes d'optimisation avec contrainte de type inégalité puis en considère le cas des contraintes égalité. La généralisation au problème contenant les deux types de contraintes ne présente par la suite aucune difficulté.

3.2 Conditions d'optimalité

Les conditions d'optimalité sont les mêmes que pour des problèmes sans contrainte avec une légère modification.

3.2.1 Cas d'une contrainte d'inégalité

Soit le problème suivant :

$$\begin{cases} \min f(\mathbf{X}) \quad \mathbf{X} \in \Re^n \\ g(\mathbf{X}) \leq 0 \end{cases}$$

avec :

$$g(\mathbf{X}) = \begin{bmatrix} g_1(\mathbf{X}) \\ g_2(\mathbf{X}) \\ \vdots \\ g_q(\mathbf{X}) \end{bmatrix}$$

Théorème 3.
Si $f(\mathbf{X})$ et $g(\mathbf{X})$ sont des fonctions différentiables de $\Re^n$ dans $\Re$, et si $\mathbf{X}^*$ est un minimum local de f sur l'ensemble $\{\mathbf{X}/g(\mathbf{X}) \leq 0\}$. Si on suppose de plus que $\nabla g(\mathbf{X}^*) \neq 0$ alors, il existe $\gamma \geq 0$ tel que :

$$\begin{aligned}&\nabla f(\mathbf{X}^*) + \gamma \nabla g(\mathbf{X}^*) = 0\\&\gamma g(\mathbf{X}^*) = 0\end{aligned}$$

de plus, si f et g sont convexes alors ces deux égalités sont suffisantes pour assurer que $\mathbf{X}^*$ est un minimum local.

Exp.

$$\begin{aligned}&f(x) = x^2 \qquad\qquad\qquad\qquad f(x) = x^2\\&x \geq 1 \qquad\qquad\qquad g(x) = 1 - x \leq 0\end{aligned}$$

alors :
$f'(x) = 2x$ et,
$g(x) = 1 - x$ d'où $g'(x) = -1 \neq 0$
les conditions de convergence peuvent s'écrire :

$$\begin{aligned}&2x^* + \gamma(-1) = 0 \quad \gamma = 0 \Rightarrow x^* = 0 \quad \text{non admissible}\\&\gamma(1 - x^*) = 0 \quad x^* = 1 \Rightarrow \gamma = 2\end{aligned}$$

3.2.2 Cas d'une contrainte d'égalité

Un problème d'optimisation statique qui fait intervenir un ensemble de p contraintes d'égalité est formulé comme suit :

$$\begin{cases}\min f(\mathbf{X})\\ h(\mathbf{X}) = 0\end{cases}$$

avec :

$$h(\mathbf{X}) = \begin{bmatrix} h_1(\mathbf{X})\\ h_2(\mathbf{X})\\ \vdots\\ h_p(\mathbf{X})\end{bmatrix}$$

Les conditions nécessaires pour l'existence d'un point stationnaire sont données par :

$$\begin{cases} f(\mathbf{X}) = 0 \\ f_{\mathbf{X}} + \lambda' h_{\mathbf{X}} = 0 \end{cases}$$

λ est appelé vecteur adjoint, ses composantes sont appelées les coefficients de Lagrange. Ce sont des variables intermédiaires dont l'introduction permet de simplifier les calcul.

Ainsi on a un système linéaire de $n+p$ équations et $n+p$ inconnues $\mathbf{X}$ et λ. On défini la fonction L par :

$$L(\mathbf{X}) = f + \lambda' h$$

les conditions d'optimalité s'écrivent alors $L_{\mathbf{X}} = 0$.

Exemple

Trouver le vecteur $\mathbf{X}$ qui minimise la fonction

$$L(\mathbf{X}) = x_1^2 + x_2^2$$

avec la condition

$$2x_1 + x_2 = 4$$

.

Solution

$$L = f + \lambda' h = x_1^2 + x_2^2 + \lambda(2x_1 + x_2 - 4)$$

$L_{\mathbf{X}} = 0$ donne :

$$L_{\mathbf{X}} = 0 \longrightarrow \begin{cases} \frac{\partial L}{\partial \mathbf{x_1}} = 2x_1 + 2\lambda = 0 \\ \\ \frac{\partial L}{\partial \mathbf{x_2}} = 2x_2 + \lambda = 0 \end{cases}$$

On considérons la contrainte $2x_1 + x_2 = 4$, on obtient un système à 3 équations et 3 inconnus.

Exercice 11.

. Quel est le minimum de fil de fer nécessaire pour clôturer une zone rectangulaire de surface S.

Exercice 12.

.
On veut résoudre le problème d'optimisation suivant :

$$f(x, y) = (x - 5)^2 + (y - 2)^2$$

Sous la contrainte suivante :

$$y - x = -3$$

a. Trouver graphiquement la solution.

b. En utilisant les multiplicateurs de Lagrange, trouver les solutions analytiquement.

Exercice 13.
On veut Déterminer les dimensions d'une boîte en carton parallélépipédique (comme une boite d'allumette) et dépourvue de couvercle dont la construction demande le moins de carton possible tout en ayant une contenance V déterminée.

a. Définir les variables correspondantes à ce problème.

b. Donner une formulation mathématique de ce problème d'optimisation sous contraintes(En déterminant la fonction coût et la contrainte).

c. En éliminant une variable (la hauteur), trouver les points stationnaires puis résoudre le problème en tant que problème d'optimisation sans contrainte. (**Indication** : éliminer la contrainte)

d. Démontrer que l'optimum obtenu présente bien un minimum.

3.3 Méthode du gradient projeté

3.3.1 Introduction

Cette méthode est semblable à celle vue au chapitre précédent alors on a :

$$\mathbf{X}_{k+1} = \mathbf{X}_k - \rho_k \nabla f(\mathbf{X}_k) \tag{3.2}$$

Soit Ω l'ensemble des vecteurs $\mathbf{X}$ qui vérifient les contraintes. Si $\mathbf{X}_k$ vérifie les contraintes $h(\mathbf{X}_k) = 0$ et $g(\mathbf{X}_k) \leq 0$, rien ne garantie que $\mathbf{X}_{k+1}$ vérifie à son tour ces contraintes.

La méthode du gradient projeté consiste alors à la projection du point $\mathbf{X}_k$ sur l'ensemble Ω.

3.3.2 Projection sur un convexe

Soit Ω un convexe fermé de $\Re^n$, la projection d'un point $\mathbf{X} \in \Re^n$ sur Ω, notée $P_\Omega(\mathbf{X})$ est définie comme l'unique solution de :

$$\max \frac{1}{2}\|\mathbf{X} - \mathbf{Y}\|_2^2 \qquad : \forall \mathbf{Y} \in \Omega \tag{3.3}$$

3.3.3 Algorithme

Cette méthode se base sur un algorithme identique à celui du gradient, sauf qu'on doit ajouter la projection du point sur l'ensemble Ω :

$$\mathbf{Y}_k = \mathbf{X}_k - \rho_k \nabla f(\mathbf{X}_k) \tag{3.4}$$

pour obtenir le point

$$\mathbf{X}_{k+1} = P_\Omega(\mathbf{Y}_k) \tag{3.5}$$

Le vecteur $\mathbf{Y}_k$ peut être obtenu en minimisant la fonction :

$$\min \frac{1}{2}\|\mathbf{X}_k - \rho_k \nabla f(\mathbf{X}_k) - \mathbf{Y}\|_2^2 \quad \forall \mathbf{Y} \in \Re^n \tag{3.6}$$

sous les contraintes donc c'est un problème d'optimisation sur un convexe.

Exp.

$$\begin{aligned} f(x, y) &= \frac{1}{2}(x^2 + 7y^2) \\ y - x &= 1 \end{aligned}$$

Sol. $\mathbf{X}^* = [-0.875\ 0.125]'$

3.4 Solutions des exercices

Solution Ex. 11.

— Soit un rectangle dont la longueur a et la largeur b voir figure ci-contre. La surface de ce rectangle est $S = a \times b$.

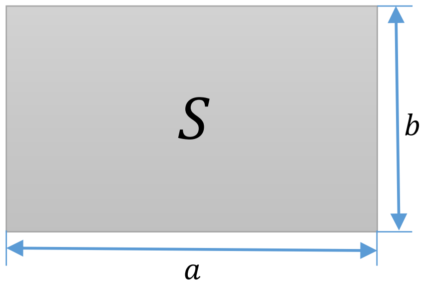


FIGURE 3.1 – Rectangle de dimension $a \times b$

— Pour clôturer ce rectangle, nous avons besoins d'un fil de longueur

$$f(a, b) = 2(a + b).$$

— La contrainte correspondante à ce problème est la surface

$$a \times b = S.$$

Tandis que la fonction à minimisé est la longueur f.

— Ainsi le Lagrangien correspondant est donné par :

$$L = f - \lambda g = 2a + 2b - \lambda(S - ab)$$

— Calcul des dérivées

$$L_X = 0 \longrightarrow \begin{cases} \frac{\partial L}{\partial a} = 2 - \lambda b = 0 \\ \\ \frac{\partial L}{\partial b} = 2 - \lambda a = 0 \end{cases}$$

de ces deux équations on tire $a = b$.

— En utilisant la contrainte, on obtient :

$$ab = S \quad \rightarrow \quad a^2 = S \quad \rightarrow a = b = \sqrt{S}$$

Ainsi nous avons un cas particulier du rectangle où $a = b$, ce qui forme un carré.

Solution Ex. 12.

.

On à le problème d'optimisation suivant :

$$\begin{cases} \min f(x,y) = (x-5)^2 + (y-2)^2 \\ \text{avec la contrainte} \\ \;y - x = -3 \end{cases}$$

Sous la forme standard on obtient :

$$\begin{cases} \min f(x,y) = (x-5)^2 + (y-2)^2 \\ \\ \;g(x,y) = y - x + 3 = 0 \end{cases}$$

Le Lagrangien s'écrit sous la forme :

$$L = f + \lambda g$$

$$L = (x-5)^2 + (y-2)^2 + \lambda(y - x + 3)$$

Le système d'équations pour la résolution de ce problème est :

$$\begin{cases} L_x = 2(x-5) - \lambda = 0 \ldots\ldots (1) \\ \\ L_y = 2(y-2) + \lambda = 0 \ldots\ldots (2) \\ \\ y - x + 3 = 0 \ldots\ldots\ldots\ldots\ldots (3) \end{cases}$$

$$(1) + (2) \rightarrow \{\; x - 5 + y - 2 = 0 \ldots\ldots (4)$$

$$(4) \text{ et } (2) \rightarrow \begin{cases} y = 7 - x \\ y = x - 3 \end{cases}$$

on tire alors

$$7 - x = x - 3 \implies x = 5 \text{ et } y = 2$$

le Hessien correspondant au point $(5,2)$ est :

$$H_f(5,2) = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$$

Suivant les exercices précédents, le Hessien est défini positif alors en ce point on à un minimum.

Solution Ex. 13.

. On montre via ce problème d'optimisation sous contraintes de type égalité qu'il est possible (dans certaines situations), de le ramener à un problème d'optimisation sans contraintes.

— Considérons les trois variables x, y et z représentant les dimensions de la boite.

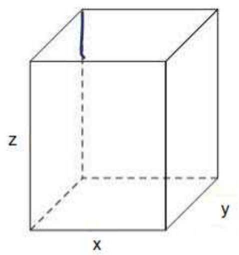


FIGURE 3.2 – Boite sans couvercle.

— Le volume de la boite est donné par la relation (Représente la contrainte) :

$$V = xyz$$

— La boite étant sans couvercle alors, la surface correspondante à la somme des 5 faces est donnée par :

$$S = xy + 2(xz + yz)$$

Cette équation représente la fonction à minimisé, car on cherche le minimum de carton possible pour couvrir les surfaces.

— L'élimination de la contrainte est possible car on à :

$$z = \frac{V}{xy}$$

— alors, $f(x, y) = xy + \frac{2V}{y} + \frac{2V}{x}$

— le point stationnaire est donné pour :

$$\begin{cases} \frac{\partial f}{\partial x} = y - \frac{2V}{x^2} = 0 \quad \to \ 2V = yx^2 \\ \\ \frac{\partial f}{\partial y} = x - \frac{2V}{y^2} = 0 \quad \to \ 2V = xy^2 \end{cases}$$

alors

$$xy^2 = yx^2 \implies x = y$$

— Enfin on obtient $V = 2x^3 \quad \rightarrow \; x = y = \sqrt[3]{2V}$ et $z = \sqrt[3]{\frac{V}{4}}$

— Pour déterminer la nature de l'optimum on doit calculer les dérivées partielles :

$$\begin{cases} \frac{\partial^2 f}{\partial x^2} = \frac{4V}{x^3} & \frac{\partial^2 f}{\partial x \partial y} = 1 \\ \\ \frac{\partial f}{\partial y \partial x} = 1 & \frac{\partial^2 f}{\partial y^2} = \frac{4V}{y^3} \end{cases}$$

alors le Hessien est donné par :

$$\mathbf{H} = \begin{pmatrix} \frac{4V}{x^3} & 1 \\ \\ 1 & \frac{4V}{y^3} \end{pmatrix}$$

— au point optimum $\mathbf{H}(\sqrt[3]{2V}, \sqrt[3]{2V})$ on obtient :

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

on a $\det \mathbf{H} = 3 > 0$ et $\mathbf{H}_{11} = 2 > 0$ alors ce point présente un minimum.

Chapitre 4

Programmation linéaire

4.1 Exemples introductifs

— Représenter graphiquement les solutions des inégalités suivantes :

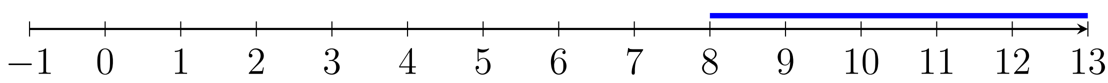
$$x - 8 \geq 0$$
$$x^2 + y^2 \leq 1.5$$

— Résoudre graphiquement le système d'équations suivant :

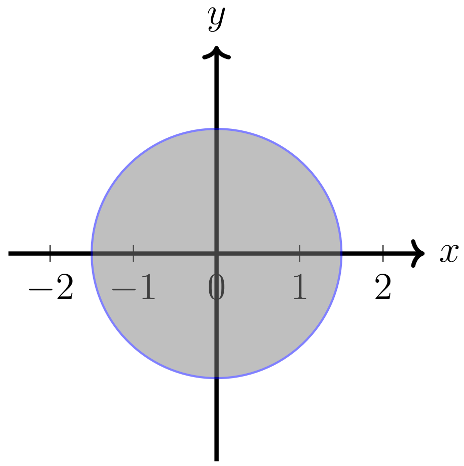
$$\begin{cases} x - y & = 1 \\ 6x + 8y & = 2 \end{cases}$$

Réponses

— pour $x - 8 \geq 0$, nous avons la demi-droite en bleu :

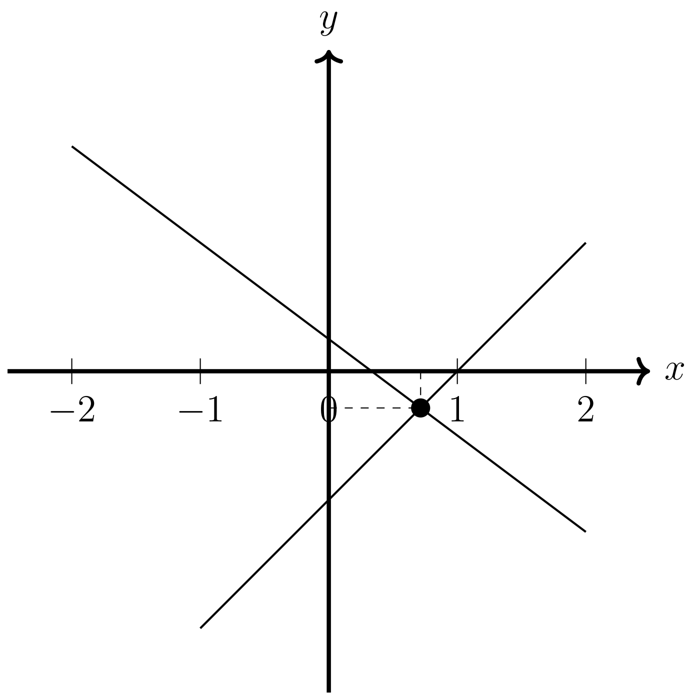


— pour $x^2 + y^2 \leq 1.5$, l'ensemble des solutions est donné par le surface à l'intérieur du cercle de rayon 1.5 :



— La solution dans ce cas est donnée par l'intersection des deux droites :

$$y = x - 1$$
$$y = -\frac{3}{4}x + \frac{1}{4}$$



4.2 Définition 1

La programmation linéaire, est une méthode d'optimisation, dont la fonction coût (objective) et les contraintes sont linéaires. Ainsi dans ce chapitre on ne considère que la résolution des équations linéaires.

La forme standard et compacte d'un problème de programmation linéaire (**PL**) est donnée par :

$$\begin{cases} \min \quad \mathbf{C'X} \quad \text{(fonction coût)} \\ \text{sous les contraintes} \\ \mathbf{AX} = \mathbf{B} \\ \mathbf{X} \geq 0 \end{cases}$$

Où : $\mathbf{X}$, $\mathbf{B}$ et $\mathbf{C}$ sont des vecteurs de dimensions n et $\mathbf{A}$ une matrice de dimension $m \times n$.
Si la contrainte 1 est donnée sous la forme d'une inégalité $\mathbf{AX} \leq \mathbf{B}$, elle peut être remis sous la forme d'une égalité comme suit : $\mathbf{AX} + \mathbf{Y} = \mathbf{B}$ où $\mathbf{Y} \geq 0$
Ainsi le problème devient :

$$\begin{cases} \min \quad \mathbf{C'X} \\ \mathbf{AX} + \mathbf{Y} = \mathbf{B} \\ \mathbf{X} \geq 0 \\ \mathbf{Y} \geq 0 \end{cases}$$

$\mathbf{Y}$ est appelée variable d'écart (*slack variable*).

Exemple 1

Considérons le problème suivant :

$$\begin{cases} \min \quad z = x_1 - x_2 \\ 2x_1 + x_2 \geq 2 \\ x_1 + 3x_2 \leq 3 \\ x_1, \; x_2 \geq 0 \end{cases}$$

Ce problème se transforme sous la forme standard comme suit :

$$\begin{cases} \min \quad z = x_1 - x_2 \\ 2x_1 + x_2 - x_3 = 2 \\ x_1 + 3x_2 + x_4 = 3 \\ x_1, \; x_2, \; x_3, \; x_4 \geq 0 \end{cases}$$

4.3 Définitions 2

On considère le domaine des solutions réalisable comme étant l'intersection d'un sous-espace avec l'orthant positif (quadrant en 2D, Octant en 3D, et orthant en dimension n). Les points extrêmes des solutions réalisables sont appelées solutions de base.

4.4 Le théorème fondamentale de la programmation linéaire

Étant donné un programme linéaire (**PL**) sous la forme standard

$$\begin{cases} \min \quad \mathbf{C}'\mathbf{X} \\ \mathbf{AX} = \mathbf{B} \\ \mathbf{X} \geq 0 \end{cases}$$

où $\mathbf{A}$ est une matrice $m \times n$ de rang m, alors :

1. S'il existe une solution réalisable au programme linéaire, alors il possède une solution de base réalisable.
2. S'il existe une solution réalisable optimale au programme linéaire, alors il existe une solution de base réalisable optimale.

4.5 Relation avec la convexité

4.5.1 Définition

Un point $\mathbf{X}$ appartenant à un ensemble convexe Ω, est dit *point extrême*, s'il n'existe pas deux points distincts $\mathbf{X}_1$ et $\mathbf{X}_2$ dans Ω qui satisfont :

$$\mathbf{X} = \alpha\mathbf{X}_1 + (1-\alpha)\mathbf{X}_2$$

pour $\alpha \in]0, 1[$.

4.5.2 Équivalence entre point extrême et solution de base

Soit $\mathbf{A}$ une matrice de dimension $m \times n$ et de rang m, et $\mathbf{B} = \{\mathbf{B}_i\}_{i=1}^{m}$, m vecteurs. Soit K le polytope convexe constitué des n vecteurs $\mathbf{X}$ qui satisfont

$$\begin{cases} \mathbf{AX} = \mathbf{B} \\ \mathbf{X} \geq 0 \end{cases}$$

Un vecteur $\mathbf{X}$ est un point extrême de K si et seulement si $\mathbf{X}$ est une solution de base réalisable.

4.5.3 Exemple

Considérons le problème de PL suivant :

$$\begin{cases} x_1 + x_2 + x_3 = 1 \\ x_1,\ x_2,\ x_3 \geq 0 \end{cases}$$

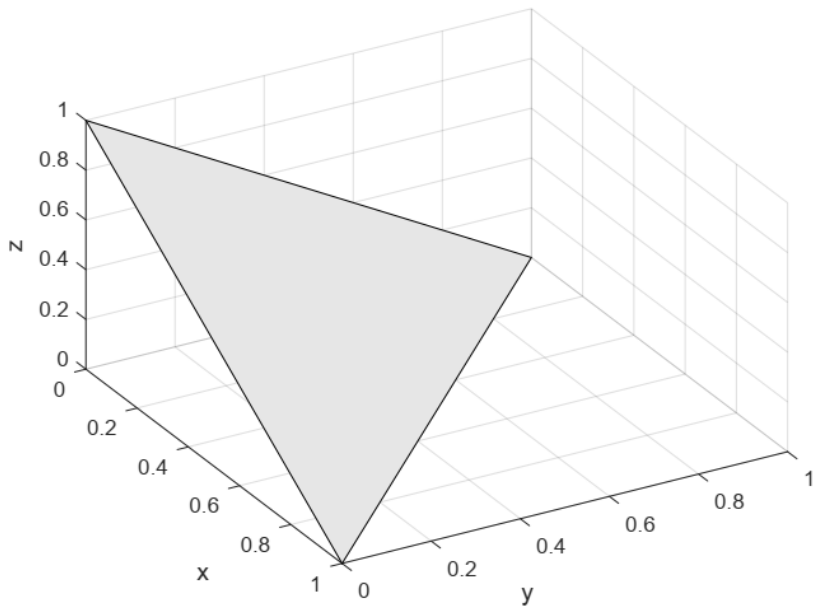


FIGURE 4.1 – Ensemble des solutions possibles pour le problème précédent

Cette figure montre la surface (le triangle), où appartient la solution admissible. Les trois sommets du triangle présentent les solutions de base (1,0,0), (0,1,0) et (0,0,1).

4.6 Résolution graphique

Certains PL peuvent être résolu de façon graphique, c'est le type de problèmes où le nombre de variables ne dépassent pas trois.

Exemple

Une entreprise fabrique des vêtements pour femme. Les profits réalisés sont de 50$ par robe et de 30$ par chemisier vendu. Combien de robes et de chemisiers doit-elle vendre pour réaliser un profit maximum en respectant les conditions suivantes :

— On ne peut pas fabriquer plus de 80 items de vêtement par mois.

— Il faut deux heures pour coudre une robe et une heure pour coudre un chemisier, alors que la machine à coudre n'est disponible que pendant cent heures par mois.

Résoudre graphiquement ce problème.

Réponse

Formulation du problème :

$$\begin{cases} \max f(x_1, x_2) = 50x_1 + 30x_2 \\ x_1 + x_2 \leq 80 \\ 2x_1 + x_2 \leq 100 \\ x_1, \ x_2 \geq 0 \end{cases}$$

La solution à ce problème correspond à l'un des sommets du polygones $[0\ \ 0], [0\ \ 50], [20\ \ 60]$ et $[0\ \ 80]$, alors on remplace ces coordonnées dans la fonction $f(x_1, x_2)$. on trouve que le maximum de profit correspond au point $[20\ \ 60]$, figure 4.2.

Exercice 14.

Soit le modèle de programmation linéaires suivant :

$$\text{M}aximiser \quad z = 4x_1 + 10x_2$$

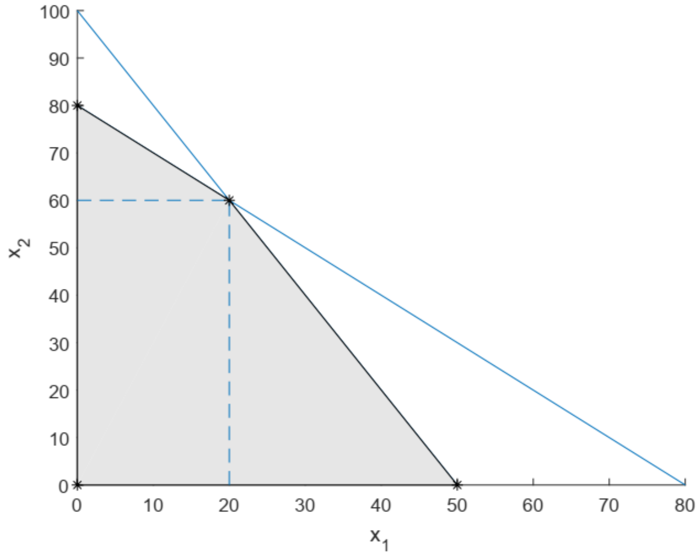


FIGURE 4.2 – Résolution géométrique du programme linéaire

Les variables x_1 et x_2 sont soumises aux contraintes suivantes :

$$\begin{cases} 3x_1 + 4x_2 \geq 12 \\ 9x_1 + 12x_2 \leq 108 \\ x_1 \geq 2 \\ x_1 \leq 10 \\ x_2 \leq 4 \\ x_1,\ x_2 \geq 0 \end{cases}$$

a. Résoudre graphiquement ce programme linéaire.

b. Quelle est la valeur maximale de la fonction objectif ?

Solution 14.

4.7 La méthode du simplexe

4.7.1 Introduction

L'idée de la méthode du simplexe est de démarrer avec une solution réalisable de base (l'un des points extrêmes) de l'ensemble des contraintes d'un problème sous sa forme standard et de passer à un autre de telle façon à minimiser la fonction coût jusqu'à l'obtention du minimum. On note qu'il est suffisant de se limité au solutions réalisable de base dans notre recherche du point optimale réalisable.

4.7.2 Algorithme (Étapes)

Les étapes constituants cette méthode sont comme suit :

1. Mise sous la forme standard
2. Ajout des variables d'écart
3. Création du tableau
4. Recherche du pivot
5. Création du nouveau tableau
6. arrêter si toutes les valeurs de la dernière ligne sont positives
7. Identification de la valeur optimale

Remarque

Les variables d'écart ont un sens physique suivant le problème traité. ainsi ils peuvent être considérés comme des activités fictives.

4.7.3 Exemple 1

Pour illustrer cette méthode on considère le programme linéaire suivant : On veut maximiser la fonction coût suivante :

$$f(\mathbf{X}) = 12x_1 + 16x_2$$

sous les contraintes suivantes :

$$\begin{cases} 10x_1 + 20x_2 \leq 120 \\ 8x_1 + 8x_2 \leq 80 \\ x_1,\ x_2 \geq 0 \end{cases}$$

Afin d'appliquer l'algorithme du simplexe, on doit mettre le problème sous sa forme standard. Ainsi, il faut transformer le problème de maximisation en un problème de minimisation. Cela est fait simplement en multipliant la fonction $f(\mathbf{X})$ par -1, et en introduisant les 2 variables non négatives d'écart x_3 et x_4. le problème se reformule comme suit :

On veut minimiser la fonction coût suivante :

$$h(\mathbf{X}) = -12x_1 - 16x_2$$

sous les contraintes suivantes :

$$\begin{cases} 10x_1 + 20x_2 + x_3 = 120 \\ 8x_1 + 8x_2 + x_4 = 80 \\ x_1,\ x_2,\ x_3,\ x_4 \geq 0 \end{cases}$$

On construit ainsi le tableau suivant :

	x_1	x_2	x_3	x_4	b	
l_1	1	2	1	0	12	(12/2=6)
l_2	1	1	0	1	10	(10/1=10)
l_3	-12	-16	0	0		

-16 est la plus petite valeur, on choisi alors la deuxième colonne, on divise les éléments de la colonne b par ceux de la deuxième colonne. Ensuite on choisis le pivot qui donne le rapport le plus petit, soit 2 dans ce cas.

	x_1	x_2	x_3	x_4	b
l_1	1/2	1	1/2	0	6
l_2	1	1	0	1	10
l_3	-12	-16	0	0	

Il faut annuler tous les éléments de la deuxième colonne sauf le pivot, alors on obtient les équations des nouvelles lignes comme suit :

$$l_2 = l_2 - l_1$$
$$l_3 = l_3 + 16l_1$$

Le tableau devient :

	x_1	x_2	x_3	x_4	b	
l_1	1/2	1	1/2	0	6	(6/0.5=12)
l_2	1/2	0	-1/2	1	4	(4/0.5=8)
l_3	-4	0	8	0		

On répète les mêmes opérations, on obtient :

	x_1	x_2	x_3	x_4	b
l_1	1/2	1	1/2	0	6
l_2	1	0	-1	2	8
l_3	-4	0	8	0	

pour annuler les éléments de la première colonne, on met à jours les nouvelles lignes comme suit :

$$l_1 = l_1 - \frac{1}{2} l_2$$

$$l_3 = l_3 + 4l_2$$

	x_1	x_2	x_3	x_4	b
l_1	0	1	5/2	-1	2
l_2	1	0	-1	2	8
l_3	0	0	4	8	

Il n'y-a pas d'éléments négatifs dans la dernière ligne alors les solutions sont :

$$\begin{cases} x_1 = 8 \\ x_2 = 2 \end{cases}$$

Ainsi le maximum de la fonctions sous les contraintes est :

$$\max f(x_1, x_2) = 152.$$

4.7.4 Exemple 2

On veut maximiser la fonction coût suivante [3] :

$$f(\mathbf{X}) = 3x_1 + x_2 + 3x_3$$

sous les contraintes suivantes :

$$\begin{cases} 2x_1 + x_2 + x_3 & \leq 2 \\ x_1 + 2x_2 + 3x_3 & \leq 5 \\ 2x_1 + 2x_2 + x_3 & \leq 6 \\ x_1,\ x_2,\ x_3 \geq 0 \end{cases}$$

Afin d'appliquer l'algorithme du simplexe, on doit mettre le problème sous sa forme standard. Ainsi, il faut transformer le problème de maximisation en un problème de minimisation. Cela est fait simplement en multipliant la fonction $f(\mathbf{X})$ par -1, et en introduisant les 3 variables non négatives d'écart x_4, x_5 et x_6. le problème se reformule comme suit :

On veut minimiser la fonction coût suivante :

$$h(\mathbf{X}) = -3x_1 - x_2 - 3x_3$$

sous les contraintes suivantes :

$$\begin{cases} 2x_1 + x_2 + x_3 + x_4 = 2 \\ x_1 + 2x_2 + 3x_3 + x_5 = 5 \\ 2x_1 + 2x_2 + x_3 + x_6 = 6 \\ x_1,\ x_2,\ x_3,\ x_4,\ x_5,\ x_6 \geq 0 \end{cases}$$

On construit ainsi le tableau suivant :

	x_1	x_2	x_3	x_4	x_5	x_6	b
	2	1	1	1	0	0	2
	1	2	3	0	1	0	5
	2	2	1	0	0	1	6
r^T	-3	-1	-3	0	0	0	0

Le premier tableau étant construit, nous cherchons le pivot. qui est le 1 encerclé dans le tableau suivant :

	x_1	x_2	x_3	x_4	x_5	x_6	b
	2	1	1	1	0	0	2
	-3	0	1	-2	1	0	1
	-2	0	-1	-2	0	1	2
r^T	-1	0	-2	1	0	0	2

Le deuxième pivot et les opérations correspondantes sont résumé sur le tableau suivant :

	x_1	x_2	x_3	x_4	x_5	x_6	b
	5	1	0	3	-1	0	1
	-3	0	1	-2	1	0	1
	-5	0	0	-4	1	1	3
r^T	-7	0	0	-3	2	0	4

Les mêmes opérations sont réalisées pour le troisième pivot ce qui donne le tableau :

	x_1	x_2	x_3	x_4	x_5	x_6	b
	1	1/5	0	3/5	-1/5	0	1/5
	0	3/5	1	-1/5	2/5	0	8/5
	0	1	0	-1	0	1	4
r^T	0	7/5	0	6/5	3/5	0	27/5

Dans ce tableau, on ne trouve pas d'éléments négatives dans la dernière ligne. Alors nous arrêtons les calculs. Les solutions optimale correspondantes sont :

$$x_1 = \frac{1}{5},\ x_2 = 0,\ x_3 = \frac{8}{5}$$

Les valeurs correspondantes aux variables auxiliaires qui ne sont pas pris en compte dans la fonction coût, sont :

$$x_4 = 0,\ x_5 = 0,\ x_6 = 4$$

La valeur maximale de la fonction coût est :

$$f(x_1,\ x_2,\ x_3) = \frac{27}{5}$$

Exercice 15. *Application du Simplexe*
. En utilisant la méthode du simplexe, déterminer la solution du programme linéaire suivant :

$$\begin{cases} \max\ z = 20x_1 + 15x_2 + 18x_3 \\ \text{sous les contraintes} \\ 5x_1 + 10x_2 + 4x_3 \leq 80 \\ 15x_1 + 12x_2 + 5x_3 \leq 120 \\ 7x_1 + 21x_2 + 3x_3 \leq 84 \\ x_1,\ x_2,\ x_3 \geq 0 \end{cases}$$

Solution 16

Exercice 16. *Modèle d'un programme linéaire.* [1]
Une entreprise fabrique quatre produits P_1, P_2, P_3 et P_4. Le processus de fabrication peut exiger jusqu'à cinq machines différentes pour la fabrication de ces produits M_1, M_2, M_3, M_4 et M_5. Les temps de fabrication sont indiqués sur le tableau suivant :

	P_1	P_2	P_3	P_4
M_1	0.25	2	1.5	-
M_2	1	1.2	2	1.4
M_3	1.5	2	0.25	2
M_4	2	0.4	0.5	1.5
M_5	0.5	1	1	2

TABLE 4.1 – Temps de fabrication en heures

Chaque machine peut opérer 40 heures par semaine, et la quantité de machines pour la fabrication est indiquée ci-après : Les bénéfices obtenus de

Machine	M_1	M_2	M_3	M_4	M_5
Quantité	5	6	10	5	8

la vente des produits sont : L'objectif étant de maximiser le bénéfice, formuler

Produit	P_1	P_2	P_3	P_4
Bénéfice ($)	8	6	9	5

le modèle de programmation linéaire pour ce problème.

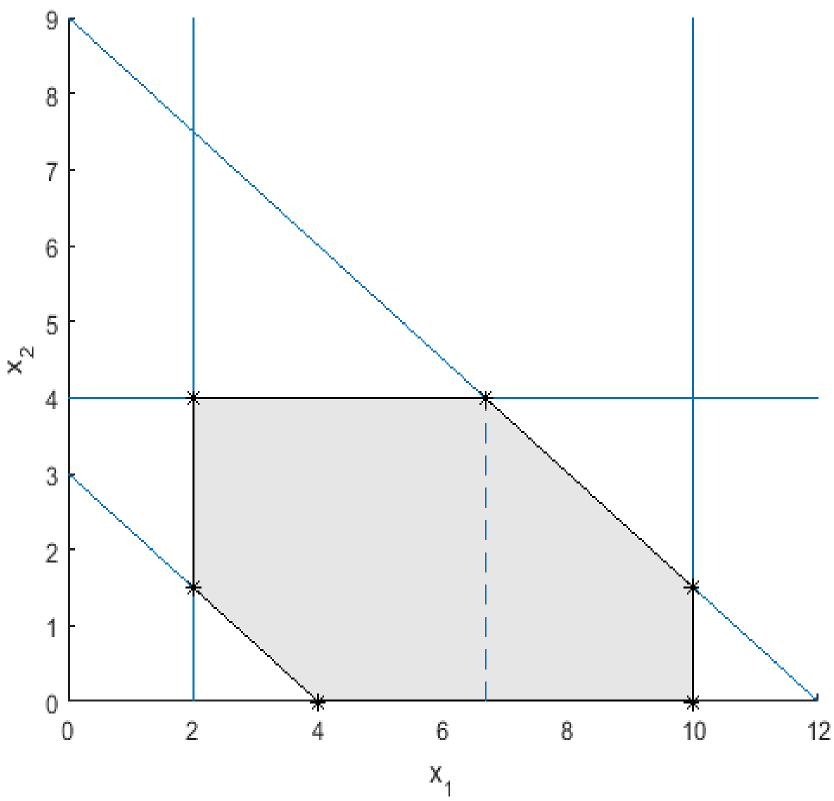
4.8 Solutions des exercices

Solution Ex. 14. énoncé de l'exercice 14.

La première étape consiste à tracer les lignes dont les équations sont :

$$\begin{cases} 3x_1 + 4x_2 = 12 \\ 9x_1 + 12x_2 = 108 \\ x_1 = 2 \\ x_1 = 10 \\ x_2 = 4 \\ x_1 = 0 \\ x_2 = 0 \end{cases}$$

Cela donne le résultat montré dans la figure suivante :



Le programme MATLAB correspondant à cette figure est :

```

figure, hold on

% Tracer des lignes
line([0 4],[3 0])
line([12 0],[0 9])
line([2 2],[0 9])
line([10 10],[0 9])
line([0 12],[4 4])

% Aire correpondante aux solutions
X=[4 10 10 20/3 2 2];
Y=[0 0 1.5 4 4 1.5];
fill(X,Y,[0.5 0.5 0.5]+0.4)

line([20/3 20/3],[0 4],'LineStyle','--')

plot(X,Y,'*k')

xlabel('x_1')
ylabel('x_2')
```

Nous devons par la suite identifier les points qui présentent les coins du polygone soit six points dans ce cas :

Point	P1	P2	P3	P4	P5	P6
x_1	2	$\frac{20}{3}$	10	10	4	2
x_2	4	4	1.5	0	0	1.5
$z = 4x_1 + 10x_2$	56	$\frac{200}{3}$	55	40	16	23

La valeur de la fonction coût correspondante à la solutions est

$$z = 4 \times \frac{20}{3} + 10 \times 4 = \frac{200}{3}$$

Solution Ex. 15. *Formulation*

.

Solution Ex. 16. *Application du Simplexe*
. Pour utiliser la méthode du simplexe nous devons ajouter 3 variables d'écart x_4, x_5 et x_6 afin de satisfaire les égalités. Écriture sous forme standard :

$$\begin{cases} \min\ z = -20x_1 - 15x_2 - 18x_3 - 0x_4 - 0x_5 - 0x_6 \\ \text{sous les contraintes} \\ 5x_1 + 10x_2 + 4x_3 + x_4 = 80 \\ 15x_1 + 12x_2 + 5x_3 v + x_5 = 120 \\ 7x_1 + 21x_2 + 3x_3 + x_6 = 84 \\ \text{avec : } x_1,\ x_2,\ x_3;\ x_4,\ x_5,\ x_6 \geq 0 \end{cases}$$

La solution de base initiale est mis sur le tableau suivant :

	x_1	x_2	x_3	x_4	x_5	x_6	b
	5	10	4	1	0	0	80
	15	12	5	0	1	0	120
	7	21	3	0	0	1	84
r^T	-20	-15	-18	0	0	0	0

donc nous allons choisir la colonne 1 et on prend comme pivot élément de la deuxième colonne

	x_1	x_2	x_3	x_4	x_5	x_6	b
	5	10	4	1	0	0	80
	15	12	5	0	1	0	120
	7	21	3	0	0	1	84
r^T	-20	-15	-18	0	0	0	0

division de la deuxième ligne par le pivot

	x_1	x_2	x_3	x_4	x_5	x_6	b
	5	10	4	1	0	0	80
	1	4/5	1/3	0	1/15	0	8
	7	21	3	0	0	1	84
r^T	-20	-15	-18	0	0	0	0

afin d'éliminer le premier élément de la première ligne, on multiplie la 2ème ligne par -5 et on l'additionne à la première ligne.

	x_1	x_2	x_3	x_4	x_5	x_6	b
	0	6	7/3	1	-1/3	0	40
	1	4/5	1/3	0	1/15	0	8
	7	21	3	0	0	1	84
r^T	-20	-15	-18	0	0	0	0

afin d'éliminer le premier élément de la 3ème ligne, on multiplie la 2ème ligne par -7 et on l'additionne à la 3ème ligne.

	x_1	x_2	x_3	x_4	x_5	x_6	b
	0	6	7/3	1	-1/3	0	40
	1	4/5	1/3	0	1/15	0	8
	0	77/5	2/3	0	-7/15	1	28
r^T	-20	-15	-18	0	0	0	0

afin d'éliminer le premier élément de la 4ème ligne, on multiplie la 2ème ligne par 20 et on l'additionne à la 4ème ligne.

	x_1	x_2	x_3	x_4	x_5	x_6	b
	0	6	7/3	1	-1/3	0	40
	1	4/5	1/3	0	1/15	0	8
	0	77/5	2/3	0	-7/15	1	28
r^T	0	1	-34/3	0	4/3	0	160

La valeur minimale dans la 4ème ligne est -34/3, alors on prend la colonne 3 et en identifie le pivot qui est 7/3.

	x_1	x_2	x_3	x_4	x_5	x_6	b
	0	6	7/3	1	-1/3	0	40
	1	4/5	1/3	0	1/15	0	8
	0	77/5	2/3	0	-7/15	1	28
r^T	0	1	-34/3	0	4/3	0	160

On divise la première ligne par le pivot.

	x_1	x_2	x_3	x_4	x_5	x_6	b
	0	18/7	1	3/7	-1/7	0	120/7
	1	4/5	1/3	0	1/15	0	8
	0	77/5	2/3	0	-7/15	1	28
r^T	0	1	-34/3	0	4/3	0	160

pour annuler le 3ème élément de la ligne 2, on multiplie la première ligne par -1/3 et on l'ajoute à la 3ème ligne.

	x_1	x_2	x_3	x_4	x_5	x_6	b
	0	18/7	1	3/7	-1/7	0	120/7
	1	-2/35	0	-1/7	4/35	0	16/7
	0	77/5	2/3	0	-7/15	1	28
r^T	0	1	-34/3	0	4/3	0	160

pour annuler le 3ème élément de la ligne 3ème ligne, on multiplie la première ligne par -3/2 et on l'ajoute à la 3ème ligne.

	x_1	x_2	x_3	x_4	x_5	x_6	b
	0	18/7	1	3/7	-1/7	0	120/7
	1	-2/35	0	-1/7	4/35	0	16/7
	0	479/35	0	-2/7	-13/35	1	116/7
r^T	0	1	-34/3	0	4/3	0	160

pour annuler le 3ème élément de la ligne 4ème ligne, on multiplie la première ligne par 34/3 et on l'ajoute à la 4ème ligne.

	x_1	x_2	x_3	x_4	x_5	x_6	b
	0	18/7	1	3/7	-1/7	0	120/7
	1	-2/35	0	-1/7	4/35	0	16/7
	0	479/35	0	-2/7	-13/35	1	116/7
r^T	0	211/7	0	34/7	-2/7	0	2480/7

La solution optimale n'est pas encore obtenue tant qu'il y-a une valeur négative dans la quatrèle ligne. Donc on cherche le pivot dans la 5ème colonne, il se touve à la 2ème ligne. donc on dévise ette ligne sur le pivot.

	x_1	x_2	x_3	x_4	x_5	x_6	b
	0	18/7	1	3/7	-1/7	0	120/7
	35/4	-1/2	0	-5/4	1	0	20
	0	479/35	0	-2/7	-13/35	1	116/7
r^T	0	211/7	0	34/7	-2/7	0	2480/7

pour annuler le 5ème élément de la première ligne, on multiplie la 2ème ligne par 1/7 et on l'ajoute à la première ligne.

	x_1	x_2	x_3	x_4	x_5	x_6	b
	5/4	5/2	1	1/4	0	0	20
	5/2	-1/2	0	-5/4	1	0	20
	0	479/35	0	-2/7	-13/35	1	116/7
r^T	0	211/7	0	34/7	-2/7	0	2480/7

pour annuler le 5ème élément de la 3ème ligne, on multiplie la 2ème ligne par 13/35 et on l'ajoute à la 3ème ligne.

	x_1	x_2	x_3	x_4	x_5	x_6	b
	5/4	5/2	1	1/4	0	0	20
	5/2	-1/2	0	-5/4	1	0	20
	13/4	27/2	0	-3/4	0	1	24
r^T	0	211/7	0	34/7	-2/7	0	2480/7

pour annuler le 5ème élément de la 4ème ligne, on multiplie la 2ème ligne par 2/7 et on l'ajoute à la 4ème ligne.

	x_1	x_2	x_3	x_4	x_5	x_6	b
	5/4	5/2	1	1/4	0	0	20
	5/2	-1/2	0	-5/4	1	0	20
	13/4	27/2	0	-3/4	0	1	24
r^T	5/2	30	0	9/2	0	0	360

A cette étape on arrêt les calcul car dans la quatrième ligne on ne trouve pas de valeur négative. Ainsi la solution de base réalisable est :

$$\begin{cases} x_3 = 20 \\ x_5 = 20 \\ x_6 = 24 \\ x_1 = 0 \\ x_2 = 0 \\ x_4 = 0 \end{cases}$$

La solution optimale est :

$$\begin{cases} x_1 = 0 \\ x_2 = 0 \\ x_3 = 20 \end{cases}$$

Chapitre 5

Programmation Non linéaire

5.1 Introduction

La forme générale d'un programme non linéaire est donnée par :

$$\begin{cases} min \quad f(X) \quad \text{(fonction coût)} \\ \text{sous les contraintes} \\ h(X) = 0 \\ g(X) \leq 0 \\ X \in \Omega \end{cases} \tag{5.1}$$

où :

$$h(X) = \begin{bmatrix} h_1(X) \\ h_2(X) \\ \vdots \\ h_m(X) \end{bmatrix}, \qquad g(X) = \begin{bmatrix} g_1(X) \\ g_2(X) \\ \vdots \\ g_p(X) \end{bmatrix}$$

$h(X) = 0$ et $g(X) \geq 0$ sont appelés les contraintes fonctionnelles, alors que $X \in \Omega$ est appelée la contrainte d'ensemble. On note aussi que $m \leq n$ et que $h_i,\ g_j \in C^0$.

5.2 Définitions

5.2.1 Contrainte active

Une contrainte d'inégalité $g_i(X) \leq 0$ est dite active en un point faisable X si $g_i(X) = 0$et inactive à X si $g_i(X) < 0$. Par convention on dit que $h_i(X)$

est active à tout point faisable.
La contrainte active au point X, définie le domaine de faisabilité au voisinage de X, alors que les autres contraintes inactives, n'ont aucune influence sur le voisinage de X. Alors l'étude des propriétés se limites au voisinage des contraintes actives. dans la figure ci-dessous la contrainte $g_1(X) \leq 0$ est acive alors que les contraintes $g_2(X) \leq 0$ et $g_3(X) \leq 0$ sont inactives.

si en plus on connait apriori qu'une contrainte est active au voisinage de X^*, alors la solution devient un minimum local. le problème peut être vu comme ayant des égalités uniquement.

5.2.2 Le plan tangent

L'ensemble des contraintes $h(X) = 0$ un hyperplan de dimension $n - m$. Si $h_i \in C^1$ alors cette surface est dite lisse.

Le plan tangent est associé à une surface lisse à un point donné.

une courbe sur une surface S est une famille de points $X(t) \in S$ paramétrique et continue suivant la variable $t \in [a, b]$.

Une courbe est différentiable si $\dot{X} \equiv \frac{d}{dt}X(t)$ existe, si de plus $\ddot{X}$ alors elle est deux fois différentiable.

Une courbe $X(t)$ passe par un point X^* si $X^* = X(t^*)$ pour $t^* \in [a, b]$. $\dot{X}(t^*)$ définie un vecteur sur E^n.

soit l'ensemble des courbes différentiables sur S et qui passent par le point X^*. Le plan tangent au point X^* présente l'ensemble des dérivées au point X^* de toutes ces courbes différentiables, c'est un sous espace de E^n.

Le plan tangent sera exprimé par les dérivées de h_i ce qui forme une surface.

soit le sous espace :

$$M = \{Y : \nabla h(X^*)Y = 0\} \tag{5.2}$$

On cherche alors dans quelles conditions, M sera égale au plan tangent au point X^*.

5.2.3 Définition

Un point X^* tel que, $h(X^*) = 0$ est dit régulier *regular point* pour les contraintes si les vecteurs gradients $\nabla h_1(X^*)$, $\nabla h_2(X^*)$, ..., $\nabla h_m(X^*)$ sont linéairement indépendants.

Au point régulier, il est possible de définir le plan tangent en terme des gradients correspondants aux contraintes.

Théorème 4. *Au point régulier X^* de la surface S définie par $h(X) = 0$, le plan tangent est égale à :*

$$M = \{Y : \nabla h(X^*)y = 0\} \tag{5.3}$$

Pour la démonstration veuillez consulté [3].

5.3 Conditions nécessaires de premier ordre (Contraintes égalité)

Connaissant maintenant la représentation du plan tangent. la dérivation des conditions de nécessité qu'un point présente un minimum local est simple.

Lemme 5. Soit X^* un point régulier de la contraintes $h(X) = 0$ et un extrémum de $f(.)$ satisfaisant les contraintes.
Alors $\forall Y \in E^n$ qui satisfait :

$$\nabla h(X^*)Y = 0 \tag{5.4}$$

doit satisfaire :

$$\nabla f(X^*)Y = 0 \tag{5.5}$$

On dit alors que $\nabla f(X^*)$ est orthogonale à Y. Par la suite, on va conclure que $\nabla f(X^*)$ présente une combinaison linéaire des gradient de $h(.)$ au point X^*. Cela nous amène à introduire les multiplicateurs de LAGRANGE .

Théorème 6. *Soit X^* un extrémum local de $f(.)$ sous les contraintes $h(X) = 0$. Soit X^* un point régulier de ces contraintes. Alors il existe un $\lambda \in E^m$ tel que :*

$$\nabla f(X^*) + \lambda^T \nabla h(X^*) = 0 \tag{5.6}$$

La condition nécessaire 5.6 avec les contraintes $h(X^*) = 0$, forment un système de $n+m$ équations à $n+m$ variables X^* et λ.

Le Lagrangien associer au problème est défini par :

$$\mathcal{L}(X, \lambda) = f(X) + \lambda^T h(X) \tag{5.7}$$

La condition nécessaire peut être formulée comme suit :

$$\begin{cases} \nabla_X \mathcal{L}(X, \lambda) = 0 \\ \nabla_\lambda \mathcal{L}(X, \lambda) = 0 \end{cases} \tag{5.8}$$

La deuxième équation de cette ensemble présente les contraintes.

Exercice 17.
Soit le problème :

$$\begin{cases} min\ x_1x_2 + x_2x_3 + x_1x_3 \\ \text{sous la contrainte} \\ x_1 + x_2 + x_3 = 3 \end{cases} \tag{5.9}$$

a. Exprimer le Lagrangien du problème.

b. Donner les conditions nécessaires de l'existence d'un extrémum.

c. Trouver l'extrémum.

Solution Ex. 17.

Le Lagrangien du problème est donné par :

$$\begin{aligned} \mathcal{L}(X, \lambda) &= f(X) + \lambda^T h(X) \\ &= x_1x_2 + x_2x_3 + x_1x_3 + \lambda(x_1 + x_2 + x_3 - 3) \end{aligned} \tag{5.10}$$

Les conditions nécessaires de l'existence d'un extrémum sont :

$$\begin{aligned}
\frac{\partial}{\partial x_1}\mathcal{L}(X,\lambda) &= 0\\
\frac{\partial}{\partial x_2}\mathcal{L}(X,\lambda) &= 0\\
\frac{\partial}{\partial x_3}\mathcal{L}(X,\lambda) &= 0\\
\frac{\partial}{\partial \lambda}\mathcal{L}(X,\lambda) &= 0
\end{aligned}$$

Cela donne le système suivant :

$$\begin{aligned}
x_2 + x_3 + \lambda &= 0\\
x_1 + x_3 + \lambda &= 0\\
x_1 + x_2 + \lambda &= 0\\
x_1 + x_2 + x_3 - 3 &= 0
\end{aligned}$$

La résolution du système précédent donne $x_1 = x_2 = x_3 = 1$ et $\lambda = -2$.

Exercice 18. Soit le problème du volume maximale d'une boite. On veut Déterminer les dimensions d'une boîte en carton dont la construction demande le moins de carton possible tout en ayant une contenance ou volume V déterminée.

Exprimer mathématiquement ce problème.

Donner le Lagrangien du problème.

Déterminer les conditions nécessaires de l'existence d'un extrémum.

Donner la solution de ce problème.

Solution Ex. 18. Soit x, y et z les dimensions de la boite, et c sa surface. Le problème d'optimisation est formulé comme suit :

$$\begin{cases}
max\ xyz\\
\text{sous la contrainte}\\
2(xy + yz + xz) = c
\end{cases} \tag{5.11}$$

Le Lagrangien est :

$$\begin{cases} yz + \lambda(y+z) = 0 \\ xz + \lambda(x+z) = 0 \\ xy + \lambda(x+y) = 0 \\ 2(xy+yz+xz) - c = 0 \end{cases}$$

La résolution de ce système nous conduit au solutions suivantes $x = y = z = \sqrt{\frac{c}{6}}$.

Exercice 19. *Entropie*. La caractérisation da la présence d'une densité de probabilité comme une distribution du maximum d'entropie. Considérons une densité de probabilité discrète d'une valeur mesurée donnant n valeurs $x_1,\ x_2,\ ...,\ x_n$.

La probabilité associé à x_i est p_i, avec $p_i \geq 0$ et

$$\sum_{i=1}^{n} p_i = 1 \tag{5.12}$$

L'entropie d'une telle densité est donnée par :

$$\epsilon = -\sum_{i=1}^{n} p_i \log(p_i) \tag{5.13}$$

La moyenne correspondante est

$$m = \sum_{i=1}^{n} x_i p_i \tag{5.14}$$

Le principe du maximum d'entropie suggère que la densité doit satisfaire le problème suivant :

$$\begin{cases} max \ -\sum_{i=1}^{n} p_i \log(p_i) \\ \quad \text{sous les contraintes} \\ \qquad \sum_{i=1}^{n} p_i = 1 \\ \qquad \sum_{i=1}^{n} x_i p_i = m \\ \qquad p_i \geq 0, \ i = 1, ..., n \end{cases} \tag{5.15}$$

Solution Ex. 19. *Maximum d'entropie*
. Le Lagrangien du problème est donné par :

$$\mathcal{L} = \sum_{i=1}^{n} \{-p_i \log(p_i) + \lambda p_i + \mu x_i p_i\} - \lambda - \mu m \tag{5.16}$$

Les conditions nécessaires de convergence présentent un ensemble de $n+2$ équations.

$$\begin{cases} -\log(p_1) - 1 + \lambda + \mu x_1 = 0 \\ -\log(p_2) - 1 + \lambda + \mu x_2 = 0 \\ \qquad\qquad\qquad \vdots \\ -\log(p_n) - 1 + \lambda + \mu x_n = 0 \end{cases} \tag{5.17}$$

Ce qui donne :

$$p_i = e^{\{\mu x_i + \lambda - 1\}}, \quad i = 1, 2, ..., n \tag{5.18}$$

Ainsi les valeurs des probabilités p_i dépendent des coefficients de LAGRANGE , il est nécessaire de les choisir de telle sorte de garantir les contraintes.

5.4 Les conditions nécessaires d'ordre deux

Supposant que X^* est un minimum local satisfaisant $h(X) = 0$ et que X^* est un point régulier de ces contraintes. Alors il, existe un $\lambda \in E^m$ tel que :

$$\nabla f(X^*) + \lambda^T \nabla h(X^*) = 0 \tag{5.19}$$

soit M Le plan tangent : $M = \{y : \nabla h(X^*)Y = 0\}$, alors la matrice :

$$\mathcal{L} = F(X^*) + \lambda^T H(X^*) \tag{5.20}$$

est semi-définie positive sur M c-à-d :

$$\forall Y \in M, Y^T L(X^*) Y \geq 0 \tag{5.21}$$

Chapitre 6

Applications de l'optimisation

6.1 La modélisation

Exercice 20. *Modélisation quadratique*
. Une entreprise a compilé les données suivantes concernant la demande d'un certain produit pour différents prix de vente :

Demande (pièce)	Prix de vente (DA)
4000	10
3500	15
3000	20
2500	25
2000	30

a. Représenter graphiquement la demande en fonction du prix de vente.

b. Commenter la courbe.

c. Donner l'expression de la demande en fonction du prix de vente.

Solution Ex. 20.
Voir chapitre quatre.

6.2 Problème de la livraison

Exercice 21. *Transport à coût minimal.* [1]

une entreprise dispose de trois usines localisées à différents endroits. La production annuelle de chaque usine pour un certains type d'appareils est comme suit (table 6.1) :

Usine	Production annuelle
U1	15 000
U2	12 000
U3	23 000

TABLE 6.1 – Tableau 1

Ces usines alimentent quatre points de vente dont la demande annuelle comme illustrer sur la table 6.2 :

Points de vente	Demande annuelle
A	10 000
B	5 000
C	20 000
D	15 000

TABLE 6.2 – Tableau 2

Le coût unitaire de transport de chaque point de vente sont indiqués dans le tableau 6.3 :

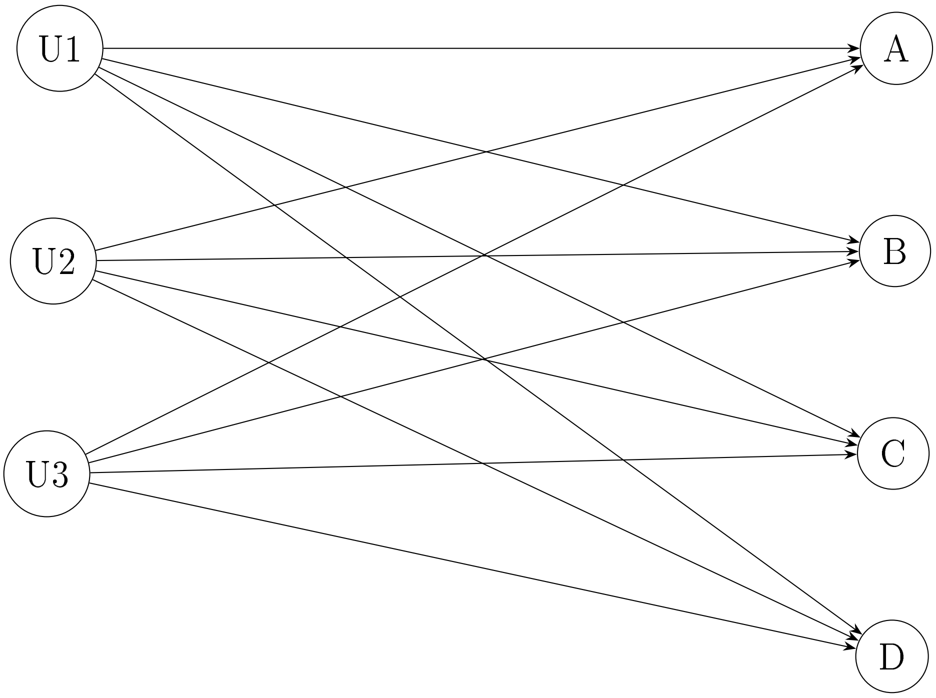
	A	B	C	D
U1	5	6	6	8
U2	11	9	4	7
U3	12	7	8	5

TABLE 6.3 – Tableau 3

1. Formuler le modèle de programmation linéaire qui permettrait d'obtenir un plan de transport à un coût minimum.

Solution Ex. 21.

Nous pouvons schématisé ce problème comme suit, les flèches représentent le déplacement de l'usine vers le point de vente :



***a.* variables de décision**

Pour ce genre de problèmes, il est préférable d'utiliser des variables à double indice. soit x_{ij}, le nombre d'unités d'appareils à expédier de l'usine $i(i = 1, 2, 3)$ au point de vente $j(j = 1(A), 2(B), 3(C), 4(D))$. nous obtenons 12 variables.

***b.* Contraintes**

Les contraintes sont résumées sur le tableau 6.4 :

	A	B	C	D	Capacité annuelle de production
U1	x_{11}	x_{12}	x_{13}	x_{14}	15 000
U2	x_{21}	x_{22}	x_{23}	x_{24}	12 000
U3	x_{31}	x_{32}	x_{33}	x_{34}	23 000
Demande annuelle	10 000	5 000	20 000	15 000	

TABLE 6.4 – Tableau 4

Les équations correspondantes aux tableau 6.4 peuvent être regroupées sous deux catégories comme suit :

Selon la capacité annuelle de production

$$\begin{aligned} x_{11} + x_{12} + x_{13} + x_{14} &<= 15000 \\ x_{21} + x_{22} + x_{23} + x_{24} &<= 12000 \\ x_{31} + x_{32} + x_{33} + x_{34} &<= 23000 \end{aligned} \tag{6.1}$$

Selon la demande annuelle de chaque zone

$$\begin{aligned} x_{11} + x_{21} + x_{31} &= 10000 \\ x_{12} + x_{22} + x_{32} &= 5000 \\ x_{13} + x_{23} + x_{33} &= 20000 \\ x_{14} + x_{24} + x_{34} &= 15000 \end{aligned} \tag{6.2}$$

Remarque

A ces deux ensembles de contraintes, s'ajoute la contrainte de positivité :

$$x_{ij} \geq 0;\ i \in \{1, 2, 3\},\ j \in \{1, 2, 3, 4\} \tag{6.3}$$

c. La fonction objective

En utilisant le tableau 6.3, et sachant que l'objectif est de minimiser le coût global, nous pouvons établir la fonction objective :

$$\begin{aligned} f(X) = &5x_{11} + 6x_{12} + 6x_{13} + 8x_{14} + 11x_{21} + 9x_{22} + \\ &4x_{23} + 7x_{24} + 12x_{31} + 7x_{32} + 8x_{33} + 5x_{34} \end{aligned} \tag{6.4}$$

Les équations 6.1, 6.2, 6.3 et 6.4 présentent la formulation du problème d'optimisation.

6.3 Problèmes en économie

Exercice 22. *Maximisation du prix de vente [2]*
. Une entreprise a compilé les données suivantes concernant la demande d'un certain produit pour différents prix de vente :

Demande (pièce)	Prix de vente (DA)
4000	10
3500	15
3000	20
2500	25
2000	30
1500	35
1000	40

TABLE 6.5 – Demande et prix de vente correspondant.

Les coûts de fabrication sont donnés sur le tableau suivant :

Aménagement de l'équipement	matière première et main-d'oeuvre par pièce (DA)
2800	8

TABLE 6.6 – Coût de fabrication.

En tant qu'analyste de l'entreprise, on vous demande de maximiser le profit.

Questions

a. Représenter graphiquement la demande n (Nombre de pièce) en fonction du prix de vente d'une pièce p.

b. Commenter la courbe.

c. Donner l'expression de la fonction $n = f(p)$.

d. En utilisant le tableau 6.6, élaborer le modèle du coût de fabrication.

e. Établir le modèle de revenu global, qui est donné par le produit du prix de vente par le nombre de pièces vendues.

f. Déterminer le modèle du bénéfice

g. Déterminer le prix de vente qui maximise le profit.

Solution Ex. 22.

a. Représentation graphique des données (figure 6.1) :

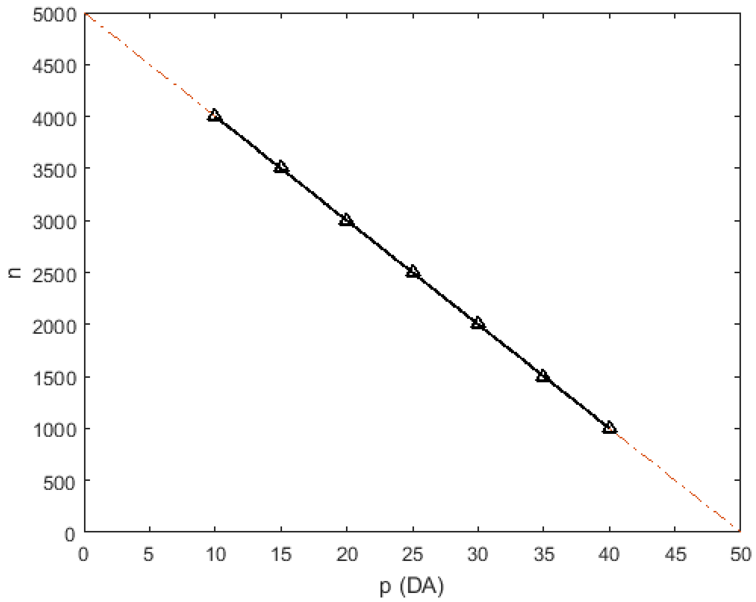


FIGURE 6.1 – Représentation graphique des données.

b. Le prix de vente diminue lorsque la quantité demandée augmente, cette attitude est bien connue dans le monde du commerce. Nous remarquons aussi que cette relation est linéaire.

c. Cette courbe est une droite. En utilisant deux points de cette droite, on peut déterminer la fonction linéaire suivante :

$$n = 5000 - 100 \times p \tag{6.5}$$

où :
n : est le nombre de pièces commandées,

p : le prix de vente d'une pièce.
Suivant cette fonction, nous remarquons que lorsque le prix augmente de 1 DA, la demande diminue par 100 unités.

d. Suivant le tableau 6.6, nous pouvons tracer le coût de production en fonction de la quantité vendue figure 6.2. Il faut remarquer que le coût de fabrication comporte :
— Une quantité fixe qui est le coût d'aménagement de l'équipement 2800 DA.
— Une quantité variable qui dépend du nombre de pièces vendus.

$$\begin{aligned} C &= 2800 + 8 \times n \\ &= 42800 - 800 \times p \end{aligned} \tag{6.6}$$

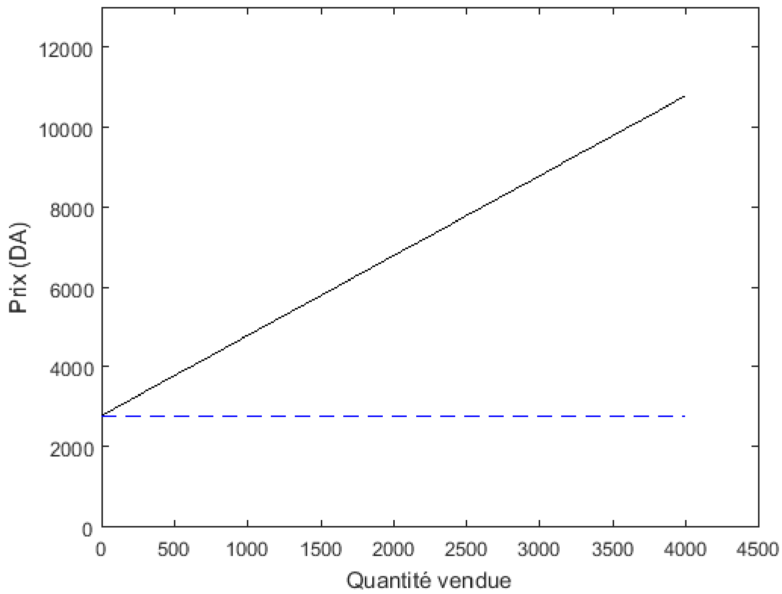


FIGURE 6.2 – Prix de vente en fonction de la quantité.

e. Le modèle du revenu global est :

$$\begin{aligned} R &= p \times n \\ &= p \times (5000 - 100 \times p) \end{aligned} \tag{6.7}$$

La fonction 6.7 n'est pas linéaire suivant le prix de vente p, la courbe correspondante est illustrée sur la figure 6.3. Cette figure montre que le revenu global est atteint pour un prix de vente égale à 25 DA.

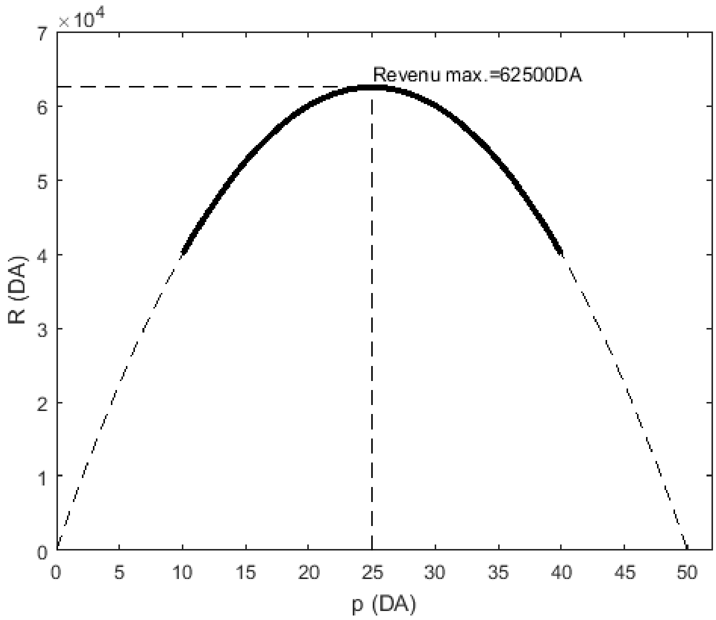


FIGURE 6.3 – Revenu en fonction du prix

f. Le bénéfice représente le revenu global d'où on retranche le coût, soit :

$$\begin{aligned} B &= R - C \\ &= p \times (5000 - 100 \times p) - (42800 - 800 \times p) \\ &= -42800 + 5800 \times p - 100 \times p^2 \end{aligned} \tag{6.8}$$

C'est un modèle quadratique figure 6.4. Pour trouver analytiquement le maximum, on prend la dérivé du bénéfice suivant la variable p soit : $B' = 5800 - 200 \times p$, pour $B' = 0$, on trouve $p = \frac{5800}{200} = 29$.

Stratégie optimale

Prix de vente suggéré :	29DA.
Volume éventuel des ventes :	2100 unité.

Revenu :	60900DA
Coût de fabrication :	
coût fixe :	2800 DA
coût variable :	16800 DA
Bénéfice :	41300 DA

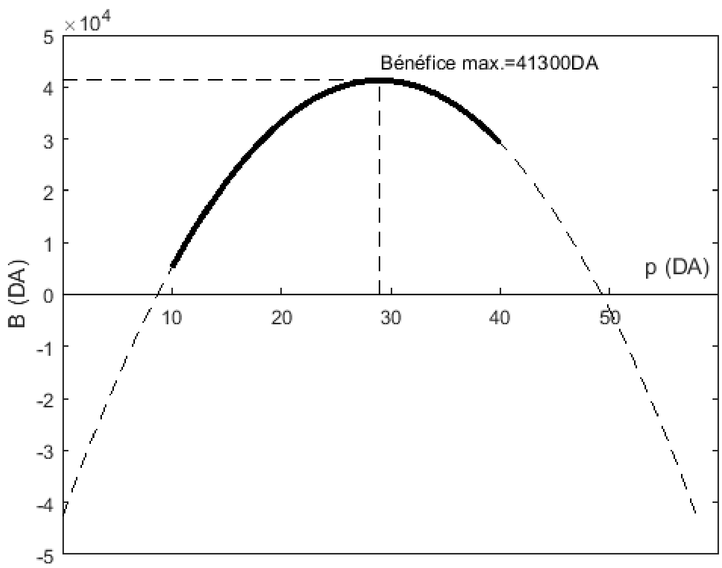


FIGURE 6.4 – Bénéfice en fonction du prix

6.4 Problème industriel

Exercice 23. *Processus Chimique.* [1]
Une raffinerie, raffine différents mélanges d'essences à l'aide de quatre pétroles bruts. Les caractéristiques techniques de ces pétroles sont les suivants :

Pétrole	Indice d'Octane	indice de pression	indice de viscosité	Disponibilité en barils
P1	72	24	3.0	10 000
P2	85	30	4.5	18 000
P3	90	28	3.8	24 000
P4	94	29	4.2	22 000

La distillation de ces pétroles permet d'obtenir les essences suivant : ordinaire, sans plomb et super. Les normes exigées pour ces essences sont les suivantes :

Essence	Indice d'Octane	indice de pression	indice de viscosité
Ordinaire	≥ 80	≤ 30	≤ 4
Sans plomb	≥ 84	≤ 29	≤ 3.8
Super	≥ 92	≤ 28.5	≤ 3.5

Une étude de marché révéla les frais suivants :

Essence	Potentiel max des ventes (barils)	bénéfice
Ordinaire	.	4.30\$
Sans plomb	20 000	5.10\$
Super	16 000	5.60\$

a. Formuler le modèle de programmation linéaire qui permettrait une répartition optimale des pétroles bruts pour obtenir les différents mélange d'essences d'après les normes indiquées et ceci tout en maximisant les bénéfices.

Remarque On suppose que les caractéristiques des pétroles bruts se mélangent d'une façon proportionnelle. Ainsi, mélanger 150 barils d'un pétrole ayant un indice d'octane de 75 et 250 barils d'un autre pétrole ayant un indice d'octane de 80, donne un mélange de 400 barils dont l'indice d'octane est

$$\frac{150 \times 75 + 250 \times 80}{400} = 78.13$$

chaque caractéristique est pondérée selon le volume de pétrole bruts mélangé. Cette remarque est valable pour les autres caractéristiques.

b. Mettez les contraintes sous forme linéaire.

Bibliographie

[1] Gérald Baillargeon (1996), Programmation linéaire appliquée, Les éditions SMG.

[2] Gérald Baillargeon (1977), Modélisation et optimisation, avec applications en gestion et en économie, deuxième édition, Les Editions SMG.

[3] David G. Luenberger, Linear and Nonlinear Programming, Second Edition. Addision-Wesley Publishing Company 1989.