

RÉPUBLIQUE ALGERIENNE DÉMOCRATIQUE

ET POPULAIRE

MINISTRE DE L'ENSEIGNEMENT SUPÉRIEURE ET DE LA RECHERCHE

SCIENTIFIQUE



UNIVERSITÉ MOULOUD MAMMERIE DE TIZI OUZOU

DÉPARTEMENT MATHÉMATIQUE



MÉMOIRE DE MASTER

MATHÉMATIQUE APPLIQUÉE A LA GESTION

THÈME :

**RÉSOLUTION DES PROBLÈMES QUADRATIQUES
CONVEXES**

Présenté par:

M^{me}. BOUCHERIT Sara

M^{me}. BESSAOUD Cylia

Encadré et proposé par :

Mr. CHEBBAH Mohammed

U.M.M.T.O

Devant le jury d'examen :

M^{me}. LESLOUS Fadhila

U.M.M.T.O

Mr. GOUBI Mohammed

U.M.M.T.O

2019/2020

Remerciements

Je remercie mon Dieu le tout puissant pour m'avoir donné la patience de suivre mes études et de faire ce modeste travail.

Je remercie mes parents qui ont tout mis en œuvre pour me permettre d'y arriver dans mes études ainsi que dans ma vie.

Je remercie Monsieur Chebbah Mohammed pour l'intérêt qu'il a apporté à ce travail, pour son soutien, son aide, sa compréhension et ses conseils hautement fructueux, tout au long de ce travail. Qu'il trouve, dans cette étude l'expression sincère et concrète de ma reconnaissance

Je remercie également Madame Leslous Fadila qui nous fait l'honneur de présider ce jury.

J'exprime vivement ma gratitude à l'ensemble des enseignants qui nous ont suivis inlassablement durant tout notre cursus universitaire.

CYLIA & SARA

Dédicaces

J'aimerais dédier ce mémoire à mes parents qui seront toujours pour moi un symbole de courage et de résistance, à mes sœurs Wifak et Souhila, mon frère Abdelhak et mon beau-frère Nacir, qui ont été présents à chaque étape de ma vie.

Je dédie ce modeste travail à mon mari Massinissa qui m'a toujours soutenu et encouragé avec tous les moyens durant toute la réalisation de ce mémoire et aussi à toute ma belle famille.

Je dédie ce travail à ma grand-mère, mes tantes mes oncles et à ma cousine Tafath.

Je dédie aussi ce travail mes amies intimes en occurrences :

- *SADAOUI Noor El Houda*
- *AISSOUS Tanissia*
- *DEHIL Rachida*

Et surtout je dédie ce travail pour mon binôme Cylia.

Dédicaces

J'aimerais dédier ce mémoire à l'être que j'aime le plus au monde, ma source de bonheur yelli azizen Nélia, à mes tendres parents qui seront toujours pour moi un symbole de courage et de résistance, à mon adorable mari Lyes qui n'a jamais cessé de m'encourager, de m'aider et de me soutenir depuis que je l'ai connu, à ma sœur Vanessa et à mes frères Aghiles et Arezki, qui ont été présents à chaque étape de ma vie.

Je dédie ce travail pour ma deuxième maman djida ouerdia, mes tantes Tassadit et Nouara qui m'ont beaucoup appris dans la vie, pour mes oncles Ahcene et Mouh.

Je dédie ce modeste travail à mes grands parents paternels Djedi Ouali et Djida Fatima.

Je dédie aussi ce mémoire à toute ma belle famille.

Je dédie aussi ce travail pour mes cousines Fadhma, Saadha et Dihia je le dédie aussi pour mes copines Ouerdia et Zazou et sans oublié tous mes collègues du CEM Tizi Ntleta.

Et surtout je dédie ce travail pour mon binôme Sara.

SOMMAIRE

Sommaire

Introduction générale	1
-----------------------------	---

Chapitre I: Rappels sur l'algèbre linéaire et la programmation mathématique

Introduction.....	3
1.1 Bases de l'algèbre linéaire.....	3
1.1.1 Espaces vectoriels et vecteurs	3
1.1.2 Matrices et vecteurs.....	3
1.1.3 Matrices carrées et matrices particulières	3
1.1.4 Matrices symétriques.....	4
1.1.5 Critère de Sylvester	4
1.1.6 Critère des valeurs propre	5
1.1.7 Produit scalaire et normes vectorielles.....	5
1.1.8 Normes vectorielles et normes matricielles	5
1.2 Fonction de plusieurs variables	6
1.2.1 Convergence des suites dans \mathbb{R}^n où dans \mathbb{R}	6
1.2.2 Ensembles ouverts et ensembles fermés	7
1.2.3 Ensemble compacte.....	7
1.2.4 Fonctions continues.....	7
1.3 Différentiabilité et dérivées partielles	8
1.3.1 Différentiabilité : le premier ordre	8
1.3.2 Différentiabilité : le second ordre.....	10
1.4 Propriétés des formes quadratiques semi-définies positives.....	11
1.4.1 Formes quadratiques définies et non définies	12
1.4.2 Critère de Sylvester pour les formes quadratiques définies et semi-définies.....	12
1.4.3 Propriétés des matrices définies positives et- non négatives.....	14

Chapitre II: Ensemble et fonction convexes

Introduction.....	15
2.1 Ensembles convexes.....	15
2.2 Enveloppe convexe	17
2.2.1 Projection d'un point sur un ensemble.....	17
2.3 Convexité et concavité d'une fonction.....	18
2.3.1 Fonction convexe	18
2.3.2 Fonction concave.....	20
2.3.3 Inégalité de Jensen	20

Chapitre III: Généralité sur la programmation non linéaire

Introduction.....	21
3.1 Les problèmes d'optimisation	21
3.2 Les problèmes d'optimisation sans contraintes.....	24
3.2.1 Condition d'optimalité	24
3.3 Les problèmes d'optimisation avec contraintes égalités	27
3.4 Les problèmes d'optimisation avec contraintes inégalités	34
3.5 Les problèmes d'optimisations avec des contraintes mixtes.....	35
3.6 Problèmes quadratiques.....	38
3.6.1 Conditions d'optimalité.....	38
3.7 La dualité de Lagrange	40

Chapitre IV: Méthodes de résolutions d'un P.Q.C

Introduction	47
4.1 Méthode quadratique de Wolfe (1959)	47
4.1.1 Cas d'un problème quadratique standard (P.Q.C).....	47
4.1.2 Algorithme	49
4.1.3 Exemple numérique.....	50
4.2 Méthode direct de support pour la résolution d'un P.Q.C. standard	53
4.2.1 Position du problème et définitions.....	53
4.2.2 Formule d'Accroissement de la fonction objectif	54
4.2.3 Critère d'optimalité	56
4.2.4 Critère de sub-optimalité	58
4.2.5 Méthode de résolution	59
4.2.6 Algorithme de résolution.....	59
4.2.7 Finitude de la méthode	62
4.2.8 Algorithme de la méthode	62
4.2.9 Exemple numérique.....	63
Conclusion générale	67

**CHAPITRE I: RAPPELS
SUR L'ALGEBRE
LINEAIRE ET LA
PROGRAMMATION
MATHEMATIQUE**

Un problème d'optimisation mathématique est défini comme étant un problème de recherche, qui consiste à explorer un espace contenant l'ensemble de toutes les solutions potentielles réalisables, dans le but de trouver la solution optimale, sinon la plus proche possible de l'optimum, permettant de minimiser ou maximiser une fonction dite objectif, sous des contraintes linéaires ou non linéaires et de type inégalités et (ou) égalités.

Suivant la nature de la fonction objectif, et celles des fonctions intervenant dans les contraintes, on tombe dans une catégorie particulière de programmation mathématique; par exemple: lorsque la fonction objectif est linéaire et les contraintes sont linéaires, on obtient un problème de programmation linéaire. Si la fonction objectif est quadratique, on a affaire à un problème de programmation quadratique.

La programmation quadratique est considérée comme une transition naturelle de la programmation linéaire vers la programmation non linéaire. En effet la majorité des méthodes développées pour le cas quadratique sont des extensions directes de celle de la programmation Linéaire. De plus les algorithmes élaborés pour le cas de l'optimisation non linéaire reposent essentiellement sur l'approche quadratique.

L'optimisation quadratique trouve son application dans plusieurs domaines tels que l'économie, les sciences de l'ingénieur, la recherche opérationnelle et la commande optimale. elle s'adapte mieux à la réalité en modélisant par exemple le risque et la distance. En effet le problème de gestion de portefeuille de Markowitz est souvent utilisé par les économistes.

Historiquement, Barankin et Dorfman [2] furent les premiers à remarquer qu'en combinant les conditions d'optimalité de Lagrange avec celles du système original, la solution optimale d'un problème quadratique était une solution de base d'un système élargi ayant la propriété que seuls certains couples de variables figuraient dans la base. De son côté Markowitz montra qu'il est possible de modifier le système élargi et d'engendrer para métriquement une classe de solutions de base ayant la propriété particulière ci-dessus et convergeant vers l'optimum en un nombre fini d'itérations. Enfin, Wolf montra qu'en modifiant légèrement la méthode de simplexe d'une façon à ne pas autoriser l'introduction d'une variable dans la base si sa variable complémentaire s'y trouvait déjà, on parvenait aisément à l'optimum recherché.

D'autres méthodes ont été développées pour la résolution de ce type de problème, parmi les quelles on peut citer:

❖ **Les algorithmes de point intérieur** : Ces méthodes, étudiées maintenant depuis plus de 20 ans, ont d'abord été développées en 1984 par Karmarkar dans le cadre de problèmes linéaires, avant d'être généralisées à d'autres problèmes plus généraux, notamment la programmation quadratique et/ou convexe. Dans cette méthode, les itérés approchent la solution par l'intérieur de S et nécessitent un nombre d'itérations qui croît de façon polynomiale avec le nombre de variables.

❖ **La Méthode adaptée de support** : Cette méthode est développée par R Gabassov et F.M Killirova pour résoudre initialement le problème d'un contrôle optimal, puis étendre à la résolution des problèmes de la programmation linéaire et quadratique convexe sous forme générale.

Dans ce travail on s'intéresse à la résolution de problèmes de programmation quadratique dont l'importance provient du fait que plusieurs problèmes réels et académiques sont quadratiques c'est le cas de l'optimisation non linéaire qui se ramène à la résolution de problème séquentielle quadratique.

Donc ce travail commence par une introduction générale, suivie du premier chapitre dans le quel nous faisons quelques rappels algébriques.

Le deuxième chapitre est consacré à certains classiques et importants concernant les ensembles et les fonctions convexes.

Le troisième chapitre traite une généralité sur la programmation non linéaire.

Dans le dernier chapitre nous présenterons en détails les méthodes existantes pour la résolution des problèmes quadratiques convexe cités ci-dessus . Et en fin nous terminons notre travail avec une conclusion générale.

**CHAPITRE I: RAPPELS
SUR L'ALGEBRE
LINEAIRE ET LA
PROGRAMMATION
MATHEMATIQUE**

Introduction

Dans ce chapitre nous allons donner quelques définitions et propriétés des fonctions quadratiques et des matrices semi-définies positives. Ceci nous sera utile dans l'étude ultérieure de la programmation quadratique convexe. Les propriétés non démontrées dans ce chapitre sont pour la plupart classiques, et relevant d'un cours d'algèbre linéaire ou bilinéaire de base.

1.1 Bases de l'algèbre linéaire

1.1.1 Espaces vectoriels et vecteurs

On note \mathbb{R}^n l'espace vectoriel de dimension n .

La base canonique de \mathbb{R}^n est (e_1, \dots, e_n) .

Un vecteur $x \in \mathbb{R}^n$ de composantes x_1, \dots, x_n est noté $x = (x_i)$, $1 \leq i \leq n$.

Les vecteurs sont notés verticalement : $x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$

1.1.2 Matrices et vecteurs

Une matrice $A \in \mathbb{R}^{m \times n}$ est notée $A = (a_{ij})$, $1 \leq i \leq m$, $1 \leq j \leq n$.

Le $j^{\text{ème}}$ vecteur colonne de A est $a_j = Ae_j$ / $A = (a_1, a_2, \dots, a_n)$

Un système de n vecteurs u_1, \dots, u_n , un $\in \mathbb{R}^m$ est noté sous forme matricielle

$U = (u_1 \dots u_n) \in \mathbb{R}^{m \times n}$, avec u_j le $j^{\text{ème}}$ vecteur colonne de U .

1.1.3 Matrices carrées et matrices particulières

Une matrice $A \in \mathbb{R}^{m \times n}$ est carrée d'ordre n si $m = n$.

La trace d'une matrice carrée A d'ordre n est la somme des éléments diagonaux :

$$tr(A) = \sum_{i=1}^n a_{ii}.$$

Le déterminant d'une matrice carrée A d'ordre n est noté $det(A)$.

la matrice identité d'ordre k , dans $\mathbb{R}^{k \times k}$, vaut $I_k = (e_1 \dots e_k)$.

Une matrice carrée D est diagonale si les seuls éléments non nuls sont sur la diagonale ; elle est notée $D = diag(d_i)$, ou $d = (d_i)$ est le vecteur formé par les éléments diagonaux.

Une matrice carrée L triangulaire inférieure si les seuls éléments non nuls sont dans le triangle inférieur.

Exemple 1:

$$L = \begin{pmatrix} 1 & 0 & 0 \\ 3 & 1 & 0 \\ 4 & 5 & 0 \end{pmatrix} \text{ est une matrice triangulaire inferieur.}$$

Une matrice carrée L triangulaire supérieur si les seuls éléments non nuls sont dans le triangle supérieur.

Exemple 2:

$$L = \begin{pmatrix} 1 & 2 & 1 \\ 0 & 0 & 2 \\ 0 & 0 & 1 \end{pmatrix} \text{ est une matrice triangulaire supérieur.}$$

1.1.4 Matrices symétriques

La transposée A^t d'une matrice A est la matrice obtenue en interchangeant les lignes et les colonnes. Soit $A = (a_{ij})$ et $B = A^t = (b_{ij})$, on a donc $b_{ij} = a_{ji}$.

Une matrice carrée A est symétrique si $A = A^t$.

Les matrices A^t et A sont symétriques.

On a $(A^t)^t = A$ et $(AB)^t = B^t A^t$.

Définition 1.1.4 Soit A une matrice carrée symétrique d'ordre n .

1. A est dite définie positive si : $x^t Ax > 0 \quad \forall x \neq 0, x \in \mathbb{R}^n$.
2. A est dite semi définie positive si : $x^t Ax \geq 0, \quad x \in \mathbb{R}^n$.
3. A est dite définie négative si : $x^t Ax < 0 \quad \forall x \neq 0, x \in \mathbb{R}^n$.
4. A est dite semi définie négative si : $x^t Ax \leq 0, \quad x \in \mathbb{R}^n$.

1.1.5 Critère de Sylvester

Soit A une matrice carre symetrique :

- a) A est définie positive si et seulement si $\det A_k > 0, k = 1, \dots, n$.
- b) A est définie négative si et seulement si $(-1)^k \det A_k > 0$.

Où
$$A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix}$$

$$\det A_k = \begin{vmatrix} a_{11} & \cdots & a_{1k} \\ \vdots & \ddots & \vdots \\ a_{k1} & \cdots & a_{kk} \end{vmatrix}$$

Remarque : $\det A_k$ est aussi appelé mineur principal successif.

1.1.6 Critère des valeurs propre

Soit A une matrice carrée symétrique.

1. A est dite définie positive si toutes ses valeurs propres sont strictement positives
2. A est dite semi définie positive si toutes ses valeurs propres sont positives ou nulles dont l'une au moins est nulle.
3. A est dite définie négative si toutes ses valeurs propres sont strictement négatives.
4. A est dite semi définie négative si toutes ses valeurs propres sont négatives ou nulles dont l'une au moins est nulle.

1.1.7 Produit scalaire et normes vectorielles

le produit scalaire de deux vecteurs x et y est $x^t y = \sum_{i=1}^n x_i y_i$; $x, y \in \mathbb{R}^n$.

soit $x = (x_i)$, on a $x_i = e_i^t x$.

Soit $A = (a_{ij})$, on a $a_{ij} = e_i^t A e_j$.

Dans le cas de vecteurs complexes, le produit scalaire hermitien est définie par

$$(x, y) \in \mathbb{C}^n \times \mathbb{C}^n \rightarrow x^* y = \sum_{i=1}^n \bar{x}_i y_i$$

Où le sur-lignage d'une grandeur indique qu'on en considère le conjugué.

Il est donc possible de définir plusieurs normes dans l'espace vectoriel \mathbb{R}^n .

1.1.8 Normes vectorielles et normes matricielles

Définition 1.1.8.1 Soit V un espace vectoriel sur le corps \mathbb{K} (\mathbb{R} ou \mathbb{C}). Une norme sur V est une application.

$\| \cdot \| : V \rightarrow \mathbb{R}^+$ qui vérifie les propriétés suivantes :

- a. $\|v\| = 0 \Leftrightarrow v = 0$, et $\|v\| \geq 0$, $\forall v \in V$.
- b. $\|\alpha v\| = |\alpha| \|v\| \forall \alpha \in \mathbb{K}$, $\forall v \in V$.
- c. $\|v + w\| \leq \|v\| + \|w\|$, $\forall v, w \in V$. (inégalité triangulaire)

Exemple les trois normes usuelles sont :

1. $\|v\|_1 = \sum_{i=1}^n |v_i|$, $v = (v_1, v_2, \dots, v_n)^t$
2. $\|v\|_2 = \sqrt{\sum_{i=1}^n (v_i)^2}$, norme euclidienne
3. $\|v\|_\infty = \max_{i=1, \dots, n} |v_i|$

Définition 1.1.8.2 Une norme matricielle est une application $\|\cdot\| : M_n(\mathbb{K}) \rightarrow \mathbb{R}$ qui vérifie les propriétés suivantes :

1. $\|A\| = 0 \Leftrightarrow A = 0$, et $\|A\| \geq 0$, $\forall A \in M_n(\mathbb{K})$
2. $\|\alpha A\| = |\alpha| \|A\|$, $\forall \alpha \in \mathbb{K}$
3. $\|A + B\| \leq \|A\| + \|B\|$, $\forall A, B \in M_n(\mathbb{K})$
4. $\|A \cdot B\| \leq \|A\| \cdot \|B\|$, $\forall A, B \in M_n(\mathbb{K})$
5. $\|A \cdot x\| \leq \|A\| \cdot \|x\|$, $x \in \mathbb{R}^n$.

1.2 Fonction de plusieurs variables

1.2.1 Convergence des suites dans \mathbb{R}^n ou dans \mathbb{R}

Définition 1.2.1.1 On définit une suite dans \mathbb{R}^n par : $(p^k)_{k>0}$ avec :
 $P^k = (p_1^k, p_2^k, \dots, p_n^k)$ et $p_i^k \in \mathbb{R}$, $i = 1, 2, \dots, n$.

Exemple : $p^k = \left(\frac{1}{k+1}, \frac{2k}{1+k^2}\right)$ une suite dans \mathbb{R}^2 .

Définition 1.2.1.2 Soit $(p_k)_{k>0}$ une suite de points (ou vecteurs) de \mathbb{R}^n . on dit que cette suite converge vers une limite $p \in \mathbb{R}^n$ si et seulement si pour tout $\varepsilon > 0$, il existe k_0 tel que

$$k \geq k_0 \Rightarrow \|p^k - p\| \leq \varepsilon.$$

On note alors $\lim_{k \rightarrow \infty} p^k = p$ ou encore parfois $p^k \rightarrow p$. dans le cas contraire on dit que la suite diverge.

Dans le cas $n = 1$, on a $\|p^k - p\| = |p^k - p|$ et on retrouve la définition usuelle de la convergence des suites réelles.

Remarque : si (p^k) converge, sa limite est unique.

Définition 1.2.1.3 Une suite $(p^k)_{k>0}$ est dite "de Cauchy" si et seulement si pour tout $\varepsilon > 0$, il existe k_0 tel que

$$k, l \geq k_0 \Rightarrow \|p^k - p^l\| \leq \varepsilon. k > l$$

Remarque : dans \mathbb{R}^n (p^k) suite de Cauchy $\Leftrightarrow (p^k)$ convergente.

1.2.2 Ensembles ouverts et ensembles fermés

Définition 1.2.2.1(Point intérieur) Soit $Y \subset \mathbb{R}^n$ et $y \in Y$. on dit que y est intérieur à Y s'il existe un voisinage de y contenu dans Y . d'une manière équivalente, y est intérieur à Y s'il existe $\varepsilon > 0$ tel que

$$z \in Y, \quad \forall z \text{ tel que } \|z - y\| \leq \varepsilon.$$

L'ensemble des points intérieurs à Y est appelé *l'intérieur de Y* et noté $\text{int}(Y)$.

Exemple : l'ensemble $A = \{x/\|x\| \leq 1\}$ a pour intérieur la boule unité $B = \{x/\|x\| < 1\}$.

Définition 1.2.2.2 (Sous-ensemble ouverts) Un sous-ensemble S est dit *ouvert* s'il coïncide avec son intérieur c'est-à-dire si $S = \text{int}(S)$.

Exemple : $B_r(A) = \{p \in \mathbb{R}^n / \|p - A\| < r\}$ est un ensemble ouvert.

Définition 1.2.2.3 (Sous-ensemble fermé) Un sous-ensemble $S \subset \mathbb{R}^n$ est dit fermé si son complémentaire est ouvert.

Exemple : $B_r(A) = \{p \in \mathbb{R}^n / \|p - A\| \leq r\}$ est un ensemble fermé.

1.2.3 Ensemble compacte

Définition 1.2.3.1 On dit que $S \in \mathbb{R}^n$ est borné si :

$$\exists r > 0 \text{ tel que: } \forall x \in S \quad \|x\| < r \Leftrightarrow x \in B_r(0).$$

Définition 1.2.3.2 Un ensemble compact de \mathbb{R}^n est un ensemble fermé borné sur \mathbb{R}^n .

1.2.4 Fonctions continues

Définition 1.2.4.1 On appelle une fonction de plusieurs variables et à valeurs réelles une fonction :

$$f: \mathbb{R}^n \rightarrow \mathbb{R} : \text{ pour } x = (x_1, \dots, x_n)^t \in \mathbb{R}^n, \quad f(x) = f(x_1, \dots, x_n) \in \mathbb{R}.$$

Définition 1.2.4.2 Une fonction $f: \mathbb{R}^n \rightarrow \mathbb{R}$ est continue au point $x = (x_1, \dots, x_n)^t \in \mathbb{R}^n$ si et seulement si pour tout $\varepsilon > 0$, il existe $\alpha > 0$ tel que pour tout $y = (y_1, \dots, y_n)^t \in \mathbb{R}^n$,

$$\|y - x\| \leq \alpha \Rightarrow |f(y) - f(x)| \leq \varepsilon.$$

Une fonction f est continue sur un ensemble E si et seulement si elle est continue en tout point de cet ensemble.

Dans le cas de fonctions $f: \mathbb{R}^n \rightarrow \mathbb{R}^m$ à valeur vectorielle, il faut remplacer $|f(y) - f(x)|$ par $\|f(y) - f(x)\|$ dans la définition ci-dessus.

1.3 Différentiabilité et dérivées partielles

1.3.1 Différentiabilité : le premier ordre

Définition 1.3.1.1(Dérivée partielle) Soit $f : \mathbb{R}^n \rightarrow \mathbb{R}$ une fonction continue. La fonction notée $\nabla_i f(x) : \mathbb{R}^n \rightarrow \mathbb{R}$, également notée $\partial f(x)/\partial x_i$ est appelée $i^{\text{ième}}$ dérivée partielle de f et est définie par

$$\lim_{h \rightarrow 0} \frac{f(x_1, \dots, x_i + h, \dots, x_n) - f(x)}{h}$$

Cette limite peut ne pas exister.

Si les dérivées partielles $\partial f(x)/\partial x_i$ existent pour tout i , le gradient de f est défini de la façon suivante.

Définition 1.3.1.2(Gradient) Soit $f : \mathbb{R}^n \rightarrow \mathbb{R}$ une fonction continue. La fonction notée $\nabla f(x) : \mathbb{R}^n \rightarrow \mathbb{R}^n$, également notée $\partial f(x)/\partial x_i$ est appelée le gradient de f et est définie par

$$\nabla f(x) = \begin{pmatrix} \frac{\partial f(x)}{\partial x_1} \\ \vdots \\ \frac{\partial f(x)}{\partial x_n} \end{pmatrix}$$

elle peut ne pas exister pour certains $x \in \mathbb{R}^n$.

Remarque : Le gradient joue un rôle essentiel dans le développement et l'analyse des algorithmes d'optimisation.

Exemple : (gradient)

Soit $f(x_1, x_2, x_3) = x_1^3 x_3 + x_2 - x_1 x_2 x_3$. Le gradient de f est donné par

$$\nabla f(x_1, x_2, x_3) = \begin{pmatrix} 3x_1^2 x_3 - x_2 x_3 \\ 1 - x_1 x_3 \\ x_1^3 - x_1 x_2 \end{pmatrix}$$

Définition 1.3.1.3 (Dérivée directionnelle)

Soit $f : \mathbb{R}^n \rightarrow \mathbb{R}$ une fonction continue. Soient $x \in \mathbb{R}^n$ et $d \in \mathbb{R}^n$.

La dérivée directionnelle de f en x dans la direction d est donnée par :

$$\lim_{\substack{\alpha > 0 \\ \alpha \rightarrow 0}} \frac{f(x + \alpha d) - f(x)}{\alpha}$$

Si la limite existe. De plus, lorsque le gradient existe, la dérivée directionnelle est le produit scalaire entre le gradient de f et la direction d , c'est-à-dire

$$\nabla f(x)^t d.$$

Définition 1.3.1.4(fonction différentiable) Soit $f : \mathbb{R}^n \rightarrow \mathbb{R}$ une fonction continue. Si, pour tout $d \in \mathbb{R}^n$, la dérivée directionnelle de f dans la direction d existe, alors la fonction f est dite différentiable.

Exemple (dérivée directionnelle)

Soit $f(x_1, x_2, x_3) = x_1^3 x_3 + x_2 - x_1 x_2 x_3$ et soit

$$d = \begin{pmatrix} d_1 \\ d_2 \\ d_3 \end{pmatrix}.$$

La dérivée directionnelle de f dans la direction d est $\text{dtf}(x^*, d)$

$$(d_1 \quad d_2 \quad d_3) \nabla f(x_1, x_2, x_3) = d_1(3x_1^2 x_3 - x_2 x_3) + d_2(1 - x_1 x_3) + d_3(x_1^3 - x_1 x_2),$$

Définition 1.3.1.5(Matrice gradient) Soit $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ telle que Soit $f_i : \mathbb{R}^n \rightarrow \mathbb{R}$ est différentiable, pour $i = 1, \dots, m$. dans ce cas, f est différentiable, et la fonction $\nabla f(x) : \mathbb{R}^n \rightarrow \mathbb{R}^{n \times m}$ est appelée matrice gradient et est définie par

$$\begin{aligned} \nabla f(x) &= \begin{pmatrix} \left. \begin{array}{c} | \\ \nabla f_1 f(x) \\ | \end{array} \right. & \cdots & \left. \begin{array}{c} | \\ \nabla f_m(x) \\ | \end{array} \right. \end{pmatrix} \\ &= \begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_2}{\partial x_1} & \cdots & \frac{\partial f_m}{\partial x_1} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_1}{\partial x_n} & \frac{\partial f_2}{\partial x_n} & \cdots & \frac{\partial f_m}{\partial x_n} \end{pmatrix} \end{aligned}$$

Remarque : La matrice gradient est souvent utilisée dans sa forme transposée et est appelée *matrice jacobienne* de f

Définition 1.3.1.6(matrice jacobienne) Soit $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$. la fonction $J(x) : \mathbb{R}^n \rightarrow \mathbb{R}^{m \times n}$ est appelée *matrice jacobienne* et est définie par

$$J(x) = \nabla f(x)^t = \begin{pmatrix} \nabla f_1(x)^t \\ \vdots \\ \nabla f_m(x)^t \end{pmatrix} \\ = \begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} & \dots & \frac{\partial f_1}{\partial x_n} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} & \dots & \frac{\partial f_2}{\partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial x_1} & \frac{\partial f_m}{\partial x_2} & \dots & \frac{\partial f_m}{\partial x_n} \end{pmatrix}.$$

1.3.2 Différentiabilité : le second ordre

Nous pouvons effectuer la même analyse de différentiabilité faite sur la fonction f dans la section (1.3.1) pour chacune des fonctions $\nabla_i f(x)$ de la définition (1.3.1). La $j^{\text{ième}}$ dérivée partielle de $\nabla_i f(x)$ est la dérivée seconde de f par rapport aux variables i et j , car

$$\frac{\partial \nabla_i f(x)}{\partial x_j} = \frac{\partial (\partial f(x) / \partial x_i)}{\partial x_j} = \frac{\partial^2 f(x)}{\partial x_i \partial x_j}.$$

Il est courant d'organiser ces dérivées secondes dans une matrice $n \times n$ dont l'élément de la ligne i et la colonne j soit $\partial^2 f(x) / \partial x_i \partial x_j$. Cette matrice est appelée *matrice hessienne* ou *hessien*.

Définition 1.3.2. (*matrice hessienne*)

Soit $f : \mathbb{R}^n \rightarrow \mathbb{R}$ une fonction deux fois différentiable.

La fonction notée $\nabla^2 f(x) : \mathbb{R}^n \rightarrow \mathbb{R}^{n \times n}$ est appelée *matrice hessienne* ou *hessien* de f et est définie par

$$\nabla^2 f(x) = \begin{pmatrix} \frac{\partial^2 f(x)}{\partial x_1^2} & \frac{\partial^2 f(x)}{\partial x_1 \partial x_2} & \dots & \frac{\partial^2 f(x)}{\partial x_1 \partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f(x)}{\partial x_n \partial x_1} & \frac{\partial^2 f(x)}{\partial x_n \partial x_2} & \dots & \frac{\partial^2 f(x)}{\partial x_n^2} \end{pmatrix}$$

Remarque : la *matrice hessienne* est toujours symétrique.

Notons que le *hessien* de f est la matrice gradient de la *matrice jacobienne* de ∇f .

Exemple (*hessien*)

Soit $f(x_1, x_2, x_3) = x_1^3 x_3 + x_2 - x_1 x_2 x_3$

Le hessien de f est donné par

$$\nabla^2 f(x_1, x_2, x_3) = \begin{pmatrix} 6x_1 x_3 & -x_3 & 3x_1^2 - x_2 \\ -x_3 & 0 & -x_1 \\ 3x_1^2 - x_2 & -x_1 & 0 \end{pmatrix}.$$

1.4 Propriétés des formes quadratiques semi-définies positives

Définition 1.4.1 Une fonction réelle de la forme suivante :

$$F(x) = \sum_{i=1}^n \sum_{j=1}^n a_{ij} x_i x_j \quad (1.4.1)$$

Est dite forme quadratique de n variables x_1, x_2, \dots, x_n .

En posant $x^t = (x_1, x_2, \dots, x_n)$, et $A = (a_{ij}, 1 \leq i, j \leq n)$, la formule (1.4.1) s'écrit sous la forme suivante :

$$F(x) = x^t A x.$$

Exemple : soit une fonction à trois variables :

$$F(x) = a_{11}x_1^2 + a_{12}x_1x_2 + a_{13}x_1x_3 + a_{21}x_2x_1 + a_{22}x_2^2 + a_{23}x_2x_3 + a_{31}x_3x_1 + a_{32}x_3x_2 + a_{33}x_3^2.$$

$$= x_1(a_{11}x_1 + a_{12}x_2 + a_{13}x_3) + x_2(a_{21}x_1 + a_{22}x_2 + a_{23}x_3) + x_3(a_{31}x_1 + a_{32}x_2 + a_{33}x_3).$$

$$= (x_1, x_2, x_3) \begin{pmatrix} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 \\ a_{31}x_1 + a_{32}x_2 + a_{33}x_3 \end{pmatrix}$$

$$= (x_1, x_2, x_3) \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}.$$

D'où

$$F(x) = x^t A x.$$

En tenant compte de la commutativité de la multiplication dans \mathbb{R} , on peut écrire $F(x)$ de la manière suivante, avec $i \leq j$:

$$F(x) = a_{11}x_1^2 + a_{22}x_2^2 + a_{33}x_3^2 + (a_{12} + a_{21})x_1x_2 + (a_{13} + a_{31})x_1x_3 + (a_{23} + a_{32})x_2x_3.$$

Pour $i \neq j$, le coefficient devant le produit $x_i x_j$ est égal à $(a_{ij} + a_{ji})$. en vertu de cela, la matrice A est toujours supposée symétrique.

1.4.1 Formes quadratiques définies et non définies

Soit la forme quadratique $F(x) = x^t D x$. (D symétrique).

Définition 1.4.1.1 Une matrice symétrique D est dite matrice définie positive (non négative) et se note $D > 0$ ($D \geq 0$), si elle est associée à une forme quadratique définie positive (non négative).

Définition 1.4.1.2 Une forme quadratique $F(x)$ est dite non définie si $F(x)$ est positive pour certaines valeurs de x et négative pour d'autres.

1.4.2 Critère de Sylvester pour les formes quadratiques définies et semi-définies.

L'intérêt du critère de Sylvester est de caractériser une forme quadratique définie ou semi-définie. Pour cela, considérons la matrice symétrique suivante :

$$D = \begin{pmatrix} d_{11} & d_{12} & \cdots & d_{1n} \\ d_{21} & d_{22} & \cdots & d_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ d_{n1} & d_{n2} & \cdots & d_{nn} \end{pmatrix}.$$

Le mineur de la matrice D , formé des lignes i_1, i_2, \dots, i_p , et des colonnes j_1, j_2, \dots, j_p , sera noté comme suit :

$$D = \begin{pmatrix} i_1, i_2, \dots, i_p \\ j_1, j_2, \dots, j_p \end{pmatrix} = \begin{vmatrix} d_{i_1 j_1} & \cdots & d_{i_1 j_p} \\ d_{i_2 j_1} & \cdots & d_{i_2 j_p} \\ \vdots & \ddots & \vdots \\ d_{i_p j_1} & \cdots & d_{i_p j_p} \end{vmatrix}.$$

Ce mineur est principal si $i_1=j_1, i_2=j_2, \dots, i_p=j_p$, c'est-à-dire s'il est formé de lignes et de colonnes portant les mêmes numéros. Les mineurs suivants

$$D_1 = d_{11}, D_2 = \begin{vmatrix} d_{11} & d_{12} \\ d_{21} & d_{22} \end{vmatrix}, \dots, D_n = \begin{vmatrix} d_{11} & d_{12} & \cdots & d_{1n} \\ d_{21} & d_{22} & \cdots & d_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ d_{n1} & d_{n2} & \cdots & d_{nn} \end{vmatrix},$$

Sont appelés mineurs principaux successifs. Nous avons alors le critère de Sylvester :

Théorème 1.4.1 (Critère de Sylvester)

- i. Pour qu'une matrice D soit définie positive ($D > 0$), il est nécessaire et suffisant que les mineurs principaux successifs de D soient positifs :

$$D_1 > 0, D_2 > 0, \dots, D_n > 0; \quad (4.1.2)$$

- ii. Pour qu'une matrice D soit semi-définie positive ($D \geq 0$), il est nécessaire et suffisant que les mineurs principaux de D soient non négatifs :

$$D \begin{pmatrix} i_1, i_2, \dots, i_p \\ j_1, j_2, \dots, j_p \end{pmatrix} \geq 0, \quad 1 \leq i_1 < i_2 < \dots < i_p, p = 1, 2, \dots, n. \quad (4.1.3)$$

Remarque 1.4.1.1 La condition

$$D_1 \geq 0, D_2 \geq 0, \dots, D_n \geq 0$$

N'est pas suffisante pour que la matrice D soit définie non négative. En effet, pour la matrice

$$D = \begin{pmatrix} 0 & 0 \\ 0 & -1 \end{pmatrix} = \begin{pmatrix} d_{11} & d_{12} \\ d_{21} & d_{22} \end{pmatrix}$$

$$D_1 = 0, \quad D_2 = 0.$$

La forme quadratique associée à D s'écrit alors :

$$F(x) = d_{11}x_1^2 + 2d_{12}x_1x_2 + d_{22}x_2^2 = -x_2^2.$$

Cette forme n'est pas comme on le voudrait définie non négative (elle est définie non positive). La raison est que le mineur principal $D \begin{pmatrix} 2 \\ 2 \end{pmatrix} = -1 < 0$

Remarque 1.4.1.2 : Pour qu'une matrice D soit définie négative ou non positive, la condition (4.1.2) et (4.1.3) se reformulent ainsi :

i. $D < 0 \Leftrightarrow (-1)^p D_p > 0, p = 1, \dots, n;$

ii. $D \leq 0 \Leftrightarrow (-1)^p D \begin{pmatrix} i_1, i_2, \dots, i_p \\ j_1, j_2, \dots, j_p \end{pmatrix} \geq 0, p = 1, \dots, n.$

En effet, on a :

$$D < 0 \Leftrightarrow H = -D > 0 \Leftrightarrow H_p > 0, p = 1, \dots, n \Leftrightarrow (-1)^p D_p > 0, p = 1, \dots, n.$$

Remarque 1.4.1.3: Le critère de Sylvester n'est valable que pour les matrices symétriques. En effet, pour la matrice

$$D = \begin{pmatrix} -1 & -2 & -1 \\ 2 & -1 & 0 \\ -1 & 0 & -1 \end{pmatrix}$$

Le critère de négativité de Sylvester est vérifié :

$$D_1 = -1, \quad D_2 = \begin{vmatrix} -1 & -2 \\ 2 & -1 \end{vmatrix} = 5 > 0, \quad D_3 = |D| = -4 < 0.$$

Néanmoins, la matrice D n'est pas définie négative puisque pour $x = (1, 0, -1) \neq 0$, on a :

$$F(x) = x^t D x = (1, 0, -1) \begin{pmatrix} 0 \\ 2 \\ 0 \end{pmatrix} = 0.$$

1.4.3 Propriétés des matrices définies positives et- non négatives

Les matrices symétriques définies ont des propriétés très intéressantes. En voici quelques-unes :

Propriété 1.4.3.1

Soit une matrice symétrique $D = (d_{ij}, 1 \leq i, j \leq n)$. si D est définie positive (non négative),

alors on a :

$$d_{ii} > 0 (d_{ii} \geq 0), \forall i = 1, 2, \dots, n.$$

Propriété 1.4.3.2

Soit la matrice D partitionnée de la manière suivante :

$$D = \begin{pmatrix} & & m & & k \\ & & & & \\ & & D_{11} & D_{12} & \\ & & D_{21} & D_{22} & \\ & & & & k \end{pmatrix}, \quad m + k = n.$$

Si $D > 0$ ($D \geq 0$), alors les sous-matrices principales D_{11} et D_{22} sont aussi définies positives (non négatives). D'une manière générale, toute sous-matrice principale d'une matrice définie positive (non négative) est définie positive (non négative).

Propriété 1.4.3.3

Un élément diagonal d'une matrice symétrique D définie non négative ne peut s'annuler que si les autres éléments de la même ligne et colonne s'annulent aussi.

Propriété 1.4.3.4

Soit D une matrice symétrique définie non négative. si x est un point quelconque mais fixé de \mathbb{R}^n tel que $x^t D x = 0$, on a alors $D x = 0$.

CHAPITRE II:
ENSEMBLES ET
FONCTIONS CONVEXES

Introduction

Ce chapitre, est consacré a la convexité, elle joue un rôle extrêmement important dans les problèmes d'optimisation, avec ou sans contraintes, elle est aussi considérée comme un outil indispensable pour la recherche des conditions d'optimalité à la fois nécessaires et suffisantes.

2.1 Ensembles convexes

Définition 2.1.1 Un ensemble $S \subset \mathbb{R}^n$ est dit convexe si et seulement si:

$$\forall x, y \in S, \forall \lambda \in [0,1] \text{ On a } \lambda x + (1 - \lambda)y \in S.$$

Autrement dit S est convexe si pour tous les points de S , le segment $[x, y]$ est inclus dans S (C'est à dire, $\forall x, y \in S$ on a $[x, y] \subset S$).

Exemple 2.1.2

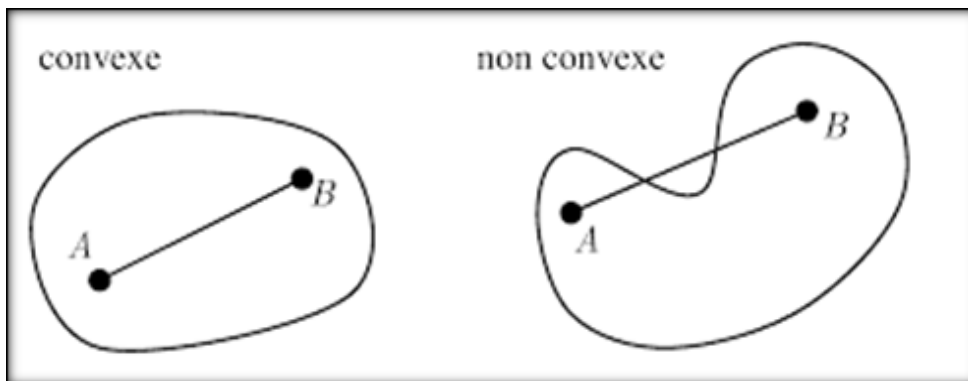


Figure 2.1- convexité d'un ensemble.

Proposition 2.1.3 On a la propriété élémentaire suivante:

Si S_1, S_2 sont deux convexe de \mathbb{R}^n et λ_1, λ_2 deux réels, alors $\lambda_1 S_1 + \lambda_2 S_2$ est un convexe de \mathbb{R}^n .

Définition 2.1.4 Un point x s'appelle combinaison convexe des points x_1, x_2, \dots, x_N s'il existe des nombres $\alpha_1, \alpha_2, \dots, \alpha_i$ ($\alpha_i \geq 0$ ($i = \overline{1, N}$)) tels que :

$$x = \sum_{i=1}^N \alpha_i x_i \quad \text{et} \quad \sum_{i=1}^N \alpha_i = 1.$$

Plus généralement, si $S \subseteq \mathbb{R}^n$ est un sous-ensemble quelconque, on dit que $x \in \mathbb{R}^n$ est une combinaison convexe de points S s'il existe un nombre fini de points de S dont x soit une combinaison convexe.

Dans le cas particulier de deux points x_1 et x_2 , toute combinaison convexe de x_1 et x_2

Peut s'écrire sous la forme:

$$x = \lambda x_1 + (1 - \lambda)x_2$$

Exemple 2.1.5 La figure 2.2 montre que tout point x d'un triangle est une combinaison convexe de ses sommets x_1, x_2 , et x_3 :

$$\begin{aligned} x &= \alpha x_1 + (1 - \alpha)x_4, & 0 \leq \alpha \leq 1, \\ x_4 &= \beta x_2 + (1 - \beta)x_3, & 0 \leq \beta \leq 1, \end{aligned}$$

D'où

$$x = \alpha x_1 + (1 - \alpha)\beta x_2 + (1 - \alpha)(1 - \beta)x_3$$

Et

$$\alpha + (1 - \alpha)\beta + (1 - \alpha)(1 - \beta) = 1.$$

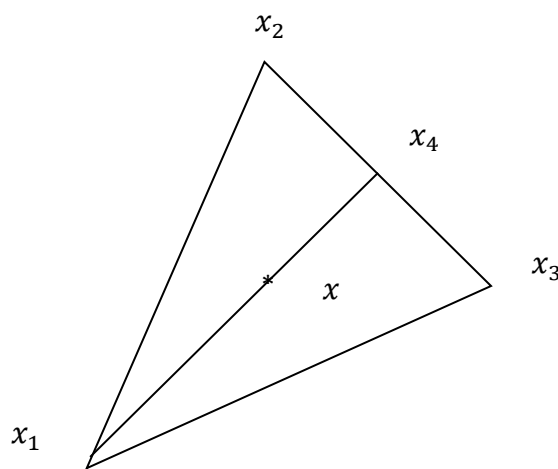


Figure 2.2 - combinaison convexe

Théorème 2.1.6 Soient S_1 et S_2 deux ensembles convexes, alors $S_1 \cap S_2$ est convexe.

Démonstration :

$$\begin{aligned} x_1, x_2 \in S_1 \cap S_2 &\Rightarrow x_1, x_2 \in S_1 \text{ et } x_1, x_2 \in S_2. \\ \left. \begin{array}{l} S_1 \text{ convexe} \Rightarrow \lambda x_1 + (1 - \lambda)x_2 \in S_1 \\ S_2 \text{ convexe} \Rightarrow \lambda x_1 + (1 - \lambda)x_2 \in S_2 \end{array} \right\} &\Rightarrow \lambda x_1 + (1 - \lambda)x_2 \in S_1 \cap S_2, \quad \lambda \in [0,1] \\ &\Rightarrow S_1 \cap S_2 \text{ est convexe.} \end{aligned}$$

Remarque : $\bigcap_{i=1}^n S_i$ est convexe si S_i convexe.

2.2 Enveloppe convexe

Définition 2.2.1 Soit $S \subset \mathbb{R}^n$. On appelle enveloppe convexe de S , et on note $\text{conv } S$, le plus petit ensemble convexe (au sens de l'inclusion) contenant S .

Il ya deux manière de construire $\text{conv } S$:

$\text{conv } S =$ Intersection de tous les convexes contenant S .

$\text{conv } S =$ Ensemble de toutes les combinaisons convexes d'éléments de S .

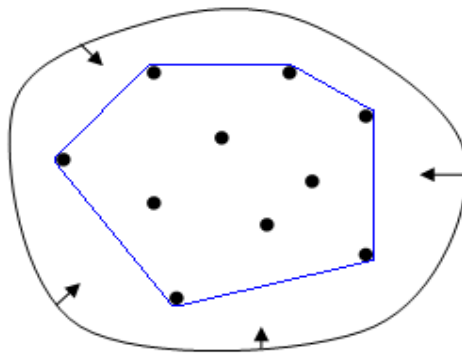


Figure 2.3 –enveloppe convexe

Définition 2.2.2 Un point x d'un ensemble X convexe est dit point extrémal s'il n'existe pas dans X des points.

x_1 et $x_2, x_1 \neq x_2$ Tel que $x = \lambda x_1 + (1 - \lambda)x_2. \quad \lambda \in]0,1[$

2.2.1 Projection d'un point sur un ensemble

Définition 2.2.1.1 On appelle projection d'un point x_0 sur un ensemble convexe fermé S un point p_{x_0} tel que

$$\| p_{x_0} - x_0 \| = \min_{x \in S} \| x - x_0 \|.$$

On dit que $\| p_{x_0} - x_0 \|$ est la distance de x_0 à S .

Définition 2.2.1.2 On appelle hyperplan dans \mathbb{R}^n un ensemble de la forme :

$$\pi = \{ X \in \mathbb{R}^n / \langle C, X \rangle = \alpha \}$$

Où $C \neq 0$. dans l'espace \mathbb{R}^n un hyperplan définit deux demi-espaces

$$\{ X \in \mathbb{R}^n / \langle C, X \rangle \leq \alpha \} \quad \{ X \in \mathbb{R}^n / \langle C, X \rangle \geq \alpha \}.$$

2.3 Convexité et concavité d'une fonction

2.3.1 Fonction convexe

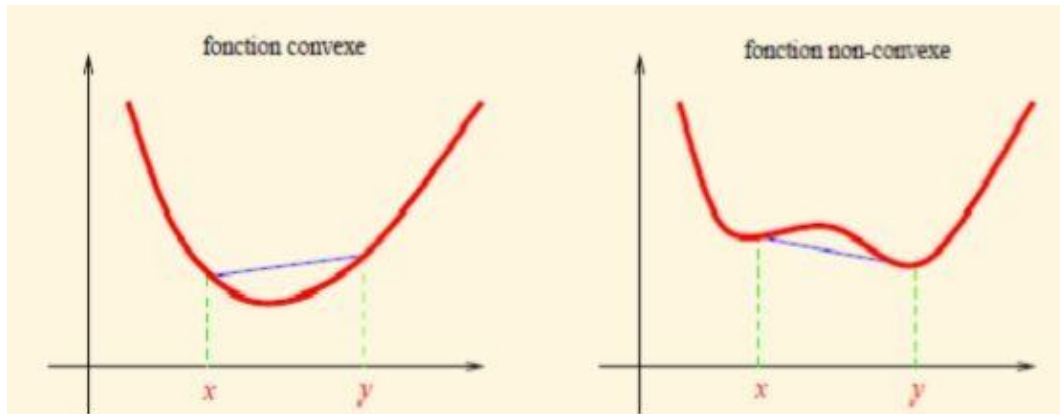


Figure 2.4-graphe de fonction convexe et de fonction non convexe

Définition 2.3.1.1

Une fonction $f: \mathbb{R}^n \rightarrow \mathbb{R}$ est dite convexe si, pour tout x et $y \in \mathbb{R}^n$ et pour tout $\lambda \in [0,1]$, on a

$$f(\lambda x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda)f(y).$$

La notion de convexité d'une fonction peut être renforcée pour obtenir la convexité stricte.

Définition 2.3.1.2

Une fonction $f: \mathbb{R}^n \rightarrow \mathbb{R}$ est dite strictement convexe si, pour tout x et $y \in \mathbb{R}^n$, $x \neq y$, et pour tout $\lambda \in]0,1[$, on a:

$$f(\lambda x + (1 - \lambda)y) < \lambda f(x) + (1 - \lambda)f(y).$$

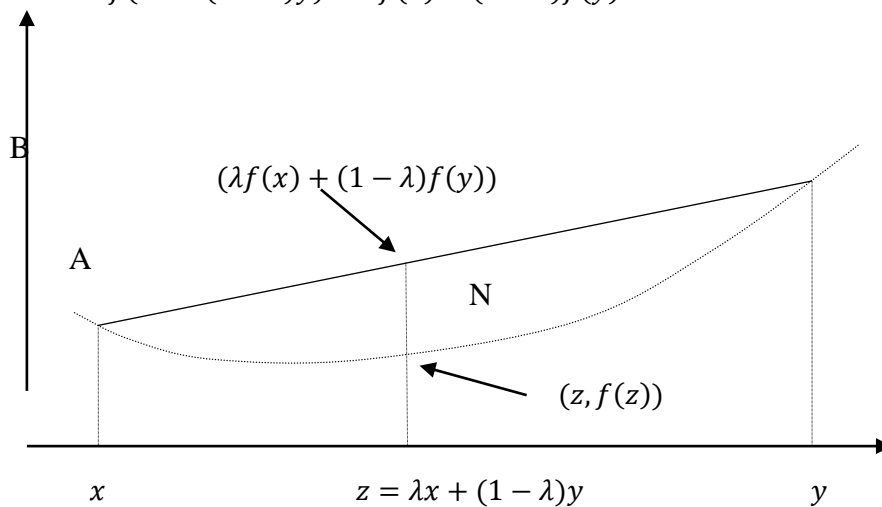


Figure 2.5- Illustration de la définition 2.3.1

On voit ici l'arc de la courbe N située entre deux points A, B de N et la corde $[A, B]$ correspondante.

Proposition 2.3.1.3

Une fonction $f: \mathbb{R}^n \rightarrow \mathbb{R}$ est dite convexe si et seulement si :

Tout arc de la courbe N est sous la corde correspondante.

Définition 2.3.1.4 Une fonction $f: \mathbb{R}^n \rightarrow \mathbb{R}$ est dite fortement convexe (α -convexe) s'il existe un $\alpha > 0$ tel que:

Pour tout x et $y \in \mathbb{R}^n$ et pour tout $\lambda \in [0,1]$, on a :

$$f(\lambda x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda)f(y) - \alpha\lambda(1 - \lambda) \| y - x \|^2$$

Remarque

f Est α -convexe $\Rightarrow f$ est strictement convexe $\Rightarrow f$ est convexe (la réciproque n'est pas vraie)

Définition 2.3.1.5 Si la fonction $(-f)$ est convexe, on dit que f est concave

Définition 2.3.1.6 On appelle épigraphe d'une fonction f convexe sur c convexe, l'ensemble noté :

$$epi(f) = \{(x, r) \in c \times \mathbb{R}^n / f(x) \leq r\}$$

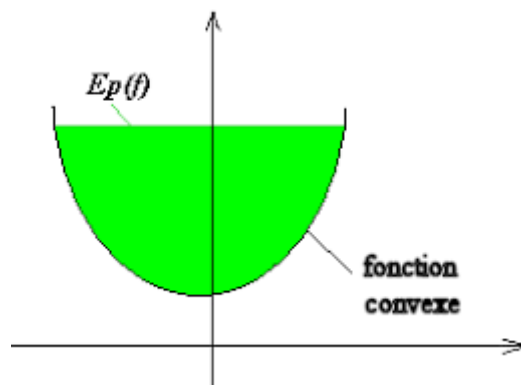


Figure 2.6 - épigraphe de la fonction f

(L'épigraphe de la fonction est la zone grisée au-dessus du graphe de la fonction)

Propriété 2.3.1.7 Soit f une fonction définie sur un ensemble convexe $c \subset \mathbb{R}^n$. alors f est si et seulement si son épigraphe

$$epi(f) = \{(x, r) \in c \times \mathbb{R}^n / f(x) \leq r\} \quad \text{est un ensemble convexe.}$$

2.3.2 Fonction concave

Définition 2.3.2.1 Une fonction $f: \mathbb{R}^n \rightarrow \mathbb{R}$ est dite concave si $-f$ est une fonction convexe, c'est-à-dire si pour tout $x, y \in \mathbb{R}^n$ et pour tout $\lambda \in [0,1]$, on a

$$f(\lambda x + (1 - \lambda)y) \geq \lambda f(x) + (1 - \lambda)f(y).$$

Remarque :

La convexité et la concavité ne sont pas des propriétés complémentaires. Une fonction peut n'être ni convexe ni concave. C'est le cas de la fonction représentée dans la figure suivante

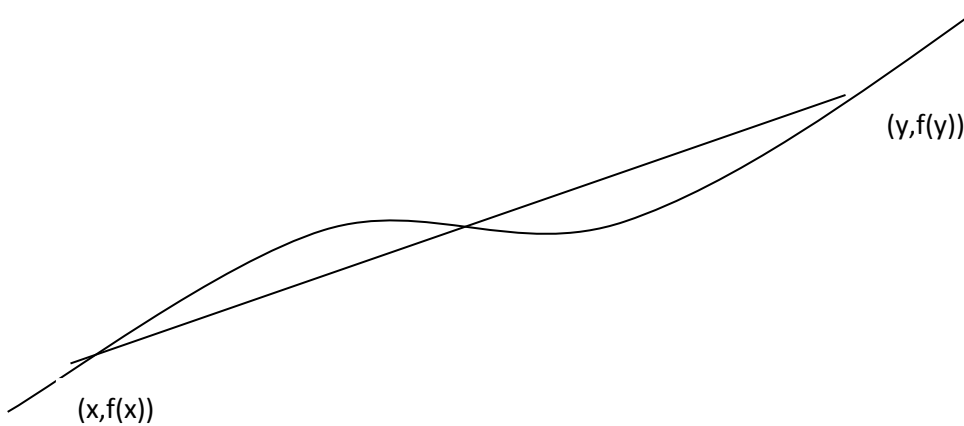


Figure 2.7 - fonction ni convexe ni concave

2.3.3 Inégalité de Jensen

Soit C un convexe de \mathbb{R}^n et $f: C \rightarrow \mathbb{R}$ une fonction convexe sur C . Pour tous $x_1, \dots, x_n \in C$ et tous $\lambda_1, \dots, \lambda_n \in \mathbb{R}^+$ tels que $\sum_{i=1}^n \lambda_i = 1$, on a l'inégalité suivante :

$$f\left(\sum_{i=1}^n \lambda_i x_i\right) \leq \sum_{i=1}^n \lambda_i f(x_i)$$

Démonstration: procédons par récurrence

Pour $n=2$ on a : $f(\lambda_1 x_1 + \lambda_2 x_2) \leq \lambda_1 f(x_1) + \lambda_2 f(x_2) \Rightarrow f$ convexe (définition).

Supposons qu'il soit vrai pour $n = k$. prouvons-le pour $n = k+1$

1er cas: $\lambda_{k+1} = 1$ on a $\lambda_1 = \lambda_2 = \dots = \lambda_k = 0$.

$$f\left(\sum_{i=1}^{k+1} \lambda_i x_i\right) = f(\lambda_{k+1} x_{k+1}) \leq \lambda_{k+1} f(x_{k+1}).$$

2e cas: $0 \leq \lambda_{k+1} < 1$, la convexité et l'hypothèse de récurrence entraînent

$$\begin{aligned} f\left(\sum_{i=1}^{k+1} \lambda_i x_i\right) &= f\left[\left(1 - \lambda_{k+1}\right) \sum_{i=1}^k \frac{\lambda_i}{1 - \lambda_{k+1}} x_i + \lambda_{k+1} x_{k+1}\right] \leq \\ &\leq (1 - \lambda_{k+1}) f\left(\sum_{i=1}^k \frac{\lambda_i}{1 - \lambda_{k+1}} x_i\right) + \lambda_{k+1} f(x_{k+1}) \leq \\ &\leq (1 - \lambda_{k+1}) \sum_{i=1}^k \frac{\lambda_i}{1 - \lambda_{k+1}} f(x_i) + \lambda_{k+1} f(x_{k+1}) = \sum_{i=1}^{k+1} \lambda_i f(x_i). \end{aligned}$$

Propositions 2.3.3.1

- Soient $f_1, f_2 : C \rightarrow \mathbb{R}$ deux fonctions convexes sur l'ensemble convexe C . Alors l'application $f_1 + f_2$ est convexe.

Démonstration :

Posons $F(x) = f_1(x) + f_2(x)$ et $x_1, x_2 \in C$, $\lambda \in [0,1]$.

$$\begin{aligned} F(\lambda x_1 + (1 - \lambda)x_2) &= f_1(\lambda x_1 + (1 - \lambda)x_2) + f_2(\lambda x_1 + (1 - \lambda)x_2) \leq \\ &\leq \lambda f_1(x_1) + (1 - \lambda)f_1(x_2) + \lambda f_2(x_1) + (1 - \lambda)f_2(x_2) \\ &\leq \lambda(f_1(x_1) + f_2(x_1)) + (1 - \lambda)(f_1(x_2) + f_2(x_2)) \\ &\leq \lambda F(x_1) + (1 - \lambda)F(x_2) \\ &\Rightarrow F \text{ est convexe} \Rightarrow f_1 + f_2 \text{ convexe.} \end{aligned}$$

- Soient $f_1, f_2 : C \rightarrow \mathbb{R}$ deux fonctions convexes sur l'ensemble convexe C . Alors l'application $\max\{f_1, f_2\}$ est convexe.

Démonstration : posons $F(x) = \max\{f_1(x); f_2(x)\}$ et $x_1, x_2 \in C$, $\lambda \in [0,1]$.

$$\begin{aligned} F(\lambda x_1 + (1 - \lambda)x_2) &= \max\{f_1(\lambda x_1 + (1 - \lambda)x_2), f_2(\lambda x_1 + (1 - \lambda)x_2)\} \leq \\ &\leq \max\{\lambda f_1(x_1) + (1 - \lambda)f_1(x_2), \lambda f_2(x_1) + (1 - \lambda)f_2(x_2)\} \\ &\leq \max\{\lambda(f_1(x_1), f_2(x_1)) + (1 - \lambda)(f_1(x_2), f_2(x_2))\} \\ &\leq \lambda \max\{f_1(x_1), f_2(x_1)\} + (1 - \lambda) \max\{f_1(x_2), f_2(x_2)\} \\ &\leq \lambda F(x_1) + (1 - \lambda)F(x_2) \\ &\Rightarrow F \text{ est convexe} \Rightarrow \max\{f_1, f_2\} \text{ convexe.} \end{aligned}$$

Propriété 2.3.3.2 Soit f une fonction de classe C^1 sur un ensemble C . Alors f est convexe si et seulement si $f(x_1) \geq f(x_2) + \nabla f(x_2)(x_1 - x_2) \quad \forall x_1, x_2 \in C$.

Exemple : $f(x) = x^2$

$$f(x_1) = x_1^2, \quad f(x_2) = x_2^2, \quad \nabla f(x_2) = f'(x_2) = 2x_2.$$

$$\begin{aligned} f(x_1) - f(x_2) - f'(x_2)(x_1 - x_2) &= x_1^2 - x_2^2 - 2x_2(x_1 - x_2). \\ &= x_1^2 + x_2^2 - 2x_1x_2 = (x_1 - x_2)^2 \geq 0 \quad \forall x_1, x_2 \in \mathbb{R}. \end{aligned}$$

$$\Rightarrow f(x_1) \geq f(x_2) + \nabla f(x_2)(x_1 - x_2) \Rightarrow f \text{ est convexe.}$$

Propriété 2.3.3.3 Soit f une fonction de classe C^2 . Alors f est convexe si et seulement si

La matrice hessienne de f est semi-définie positive.

Exemple : $f(x_1, x_2) = x_1^2 + x_2$.

$$Hf(x, y) = \begin{pmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 \partial x_2} \\ \frac{\partial^2 f}{\partial x_2 \partial x_1} & \frac{\partial^2 f}{\partial x_2^2} \end{pmatrix} = \begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix}$$

Les valeurs propres de la matrice hessienne sont :

$$\lambda_1 = 2, \lambda_2 = 0 \Rightarrow Hf \text{ est s.d.p} \Rightarrow f \text{ est convexe.}$$

Théorème 2.3.3.4 Si la matrice hessienne est définie positive, alors f est strictement convexe.

Exemple : $f(x_1, x_2) = x_1^2 + x_2^2$.

$$Hf(x, y) = \begin{pmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 \partial x_2} \\ \frac{\partial^2 f}{\partial x_2 \partial x_1} & \frac{\partial^2 f}{\partial x_2^2} \end{pmatrix} = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$$

On applique le critère de Sylvester pour les mineurs principaux successifs :

$$A_1 = 2 > 0, \quad A_2 = 4 > 0 \Rightarrow Hf \text{ est d.p} \Rightarrow f \text{ est strictement convexe.}$$

CHAPITRE III

GENERALITE SUR LA

PROGRAMMATION NON LINEAIRE

Introduction

La programmation mathématique, permet de résoudre des problèmes de décision très complexes, impliquant plusieurs variables inter-reliées. Des contraintes peuvent exister limitant le choix des valeurs des variables. La programmation mathématique est une approche quantitative où l'on s'intéresse à maximiser ou minimiser une fonction objective qui mesure la performance ou la "qualité" de notre décision.

3.1 Les problèmes d'optimisation

Un programme mathématique est un problème d'optimisation d'une fonction pour des contraintes possédant la forme générale :

$$(p) \quad \begin{cases} \text{"optimiser"} f(x) \\ h_j(x) = 0 \quad j = 1, \dots, p \\ g_i(x) \leq 0 \quad i = 1, \dots, m \\ x \in \mathbb{R}^n \end{cases} \quad (3.1)$$

Où $x = (x_1, x_2, \dots, x_n) \in \mathbb{R}^n$ est appelé la variable d'optimisation, la fonction $f: \mathbb{R}^n \rightarrow \mathbb{R}$ est appelée la fonction objectif (ou fonction coût). Les inégalités $g_i(x) \leq 0$ sont appelées les contraintes inégalités et les fonctions correspondantes $g_i: \mathbb{R}^n \rightarrow \mathbb{R}$ sont appelées les fonctions des contraintes inégalités. Les équations $h_j(x) = 0$ sont appelées les contraintes égalités et les fonctions $h_j: \mathbb{R}^n \rightarrow \mathbb{R}$ sont les fonctions des contraintes égalités.

Remarque : optimiser f signifie maximiser f ou minimiser f nous avons alors $\max f = - \min(-f)$ Donc on peut toujours se ramener à un problème de minimisation à partir d'un problème de maximisation et vice-versa

Définition 3.1.1. (Point admissible) On appelle solution réalisable du problème (3.1). Tout vecteur $x \in \mathbb{R}^n$ vérifiant :

$$Y = \begin{cases} h_j(x) = 0 \quad j = 1, \dots, p \text{ Contrainte d'égalité} \\ g_i(x) \leq 0 \quad i = 1, \dots, m. \text{ Contrainte d'inégalité} \\ x \in \mathbb{R}^n \end{cases}$$

Notons que dans la littérature, ce concept est parfois appelé solution admissible ou vecteur admissible.

Définition 3.1.2 (solution optimale) On appelle solution optimale du problème (3.1) une solution réalisable $x^* \in Y$ Tel que :

$$f(x^*) \geq f(x), \quad \forall x \in Y \text{ Problème de max}$$

Où

$$f(x^*) \leq f(x), \quad \forall x \in Y \text{ Problème de min}$$

$f(x^*)$ est dite valeur maximale ou minimale de $f(x)$

Définition 3.1.3 (Optimum local) On dit que $x^* \in Y$ est un optimum local du problème (3.1) S'il existe un voisinage $v(x)$ de x^* :

$$f(x^*) \geq f(x), \quad \forall x \in v(x) \cap Y \text{ (Problème de max)}$$

$$f(x^*) \leq f(x), \quad \forall x \in v(x) \cap Y \text{ (Problème de min)}$$

Avec $v(x) = \{x \in \mathbb{R}^n \mid \|x - x^*\| < \varepsilon, \varepsilon > 0\}$

Définition 3.1.4 (Minimum global) Appelons $Y = \{ x \in \mathbb{R}^n / g_i(x) \leq 0 \quad i = 1, \dots, m, h_j(x) = 0 \quad j = 1, \dots, p\}$ l'ensemble des vecteurs vérifiant toutes les contraintes. Le vecteur $x^* \in Y$ est appelé minimum global du problème (1.1) si

$$f(x^*) \leq f(x), \quad \forall x \in Y.$$

Définition 3.1.5 (Maximum global) Appelons $Y = \{ x \in \mathbb{R}^n / g_i(x) \leq 0 \quad i = 1, \dots, m, h_j(x) = 0 \quad j = 1, \dots, p\}$ l'ensemble des vecteurs vérifiant toutes les contraintes. Le vecteur $x^* \in Y$ est appelé maximum global du problème (1.1) si

$$f(x^*) \geq f(x), \quad \forall x \in Y.$$

Définition 3.1.6 (Point critique) Soit $f: \mathbb{R}^n \rightarrow \mathbb{R}$ une fonction différentiable. Tout vecteur $x \in \mathbb{R}^n$ tel que $\nabla f(x) = 0$ est appelé point critique ou point stationnaire de f .

Parmi tous les points stationnaires on peut avoir des minimums, des maximums et des points selles

Exemple :

$$f(x) = 3e^{-x^2} + e^{-(x-3)^2}$$

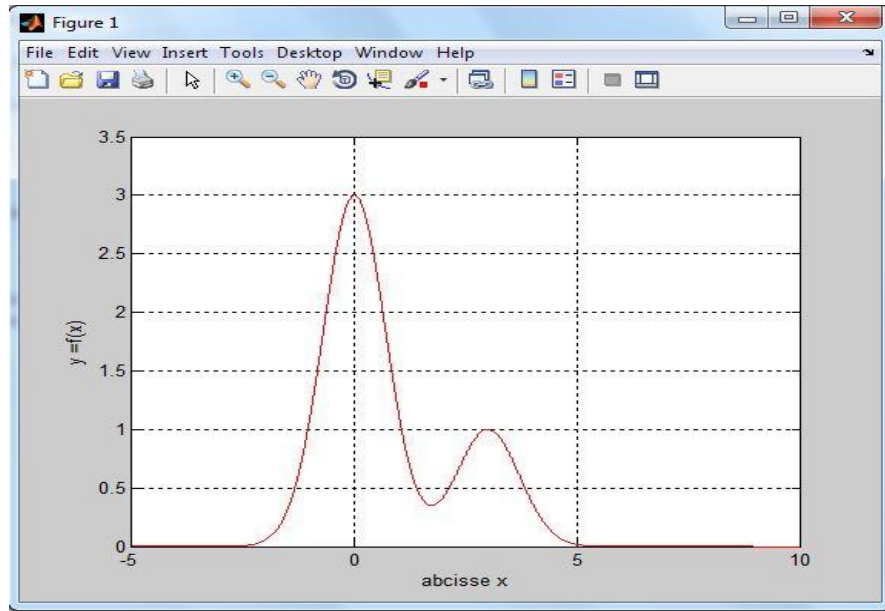


Figure 3.1 : minimum et maximum de la fonction $f(x)$

x_1 : Maximum global

x_2 : Minimum local

x_3 : Maximum local

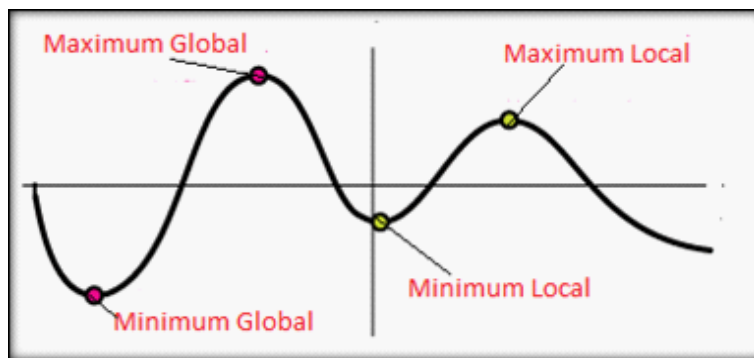


Figure 3.2 infinité de minimums et maximums globaux

3.2 Les problèmes d'optimisation sans contraintes

Définition 3.2.1. On définit un problème d'optimisation sans contraintes par :

$$(p) \quad \begin{cases} \text{minimiser } f(x) \\ x \in \mathbb{R}^n \end{cases} \quad (3.2)$$

3.2.1 Condition d'optimalité

3.2.1.1 Conditions nécessaire d'optimalité

Théorème 3.3.1.1 Soit $f: \mathbb{R}^n \rightarrow \mathbb{R}$ différentiable, si x^* un minimum local de f alors

a. $\nabla f(x^*) = 0$.

Si de plus, f est deux fois différentiable, alors

b. $\nabla^2 f(x^*)$ est semi-définie positive.

La condition (a) est appelée condition nécessaire de premier ordre,

La condition (b) est appelée condition nécessaire de second ordre.

Preuve :

a. On écrit la formule de Taylor avec $t > 0$ et $d \in \mathbb{R}^n$.

$$f(x^* + td) = f(x^*) + \nabla f(x^*)^t td + o(t|d|).$$

$$f(x^* + td) - f(x^*) = \nabla f(x^*)^t td + o(t|d|).$$

$$0 \leq \frac{f(x^* + td) - f(x^*)}{t} = \nabla f(x^*)^t d + \frac{o(t|d|)}{t}.$$

On passe à la limite dans les deux membres,

$$0 \leq \lim_{t \rightarrow 0} \frac{f(x^* + td) - f(x^*)}{t} = \lim_{t \rightarrow 0} \left(\nabla f(x^*)^t d + \frac{o(t|d|)}{t} \right) = \nabla f(x^*)^t d.$$

$$\Rightarrow \nabla f(x^*)^t d \geq 0 \quad \forall d \in \mathbb{R}^n.$$

En particulier c'est vrai pour $d = -(\nabla f(x^*))$.

$$\Rightarrow \nabla f(x^*)^t (-\nabla f(x^*)) = -\|\nabla f(x^*)\|^2 \geq 0$$

On aura donc $\nabla f(x^*) = 0$

b. On écrit la formule de Taylor à l'ordre 2 avec $t > 0$ et $d \in \mathbb{R}^n$.

$$f(x^* + td) = f(x^*) + \nabla f(x^*)^t td + \frac{1}{2} td^t Hf(x^*) td + o(t^2 \|d\|^2).$$

Comme $\nabla f(x^*) = 0$

$$0 \leq \frac{f(x^* + td) - f(x^*)}{t^2} = \frac{1}{2} d^t Hf(x^*) d + \frac{o(t^2 \|d\|^2)}{t^2}.$$

On passe à la limite dans les deux membres,

$$0 \leq \lim_{t \rightarrow 0^+} \frac{f(x^* + td) - f(x^*)}{t^2} = \frac{1}{2} d^t Hf(x^*) d + \lim_{t \rightarrow 0^+} \left(\frac{o(t^2 \|d\|^2)}{t^2} \right).$$

$$\Rightarrow \frac{1}{2} d^t Hf(x^*) d \geq 0.$$

$\Rightarrow Hf(x^*)$ est semi définie positive.

Exemple 1 :

$$f(x) = -x^4$$

a. $\nabla f(x) = f'(x) = -4x^3 = 0 \Rightarrow x^3 = 0 \Rightarrow x^* = 0.$

b. $\nabla^2 f(x) = Hf(x^*) = f''(x^*) = -12x^2$ Au point $x^* = 0$

Les conditions a et b sont vérifiées pourtant $x^* = 0$ n'est pas un minimum local car $Hf(x^*)$ n'est pas définie positive. (N'est pas aussi semi définie positive)

Exemple 2 :

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

a. $\nabla f(x, y) = (2x, -2y) = (0, 0) \Rightarrow x = y = 0.$

b. Au point $x^* = (0, 0), \Rightarrow Hf(x^*) = \begin{pmatrix} 2 & 0 \\ 0 & -2 \end{pmatrix}$

On applique le critère des valeurs propre : $\lambda_1 = -2, \lambda_2 = 2.$ Donc $x^* = (0, 0)$ n'est pas un minimum local car $Hf(x^*)$ n'est pas définie positive. (N'est pas aussi semi définie positive)

Remarque :

Les conditions nécessaires d'optimalité du premier ordre et second ordre ne sont pas suffisante pour l'optimalité,

Un point x^* qui vérifie la condition $\nabla f(x^*) = 0$ est appelée point stationnaire

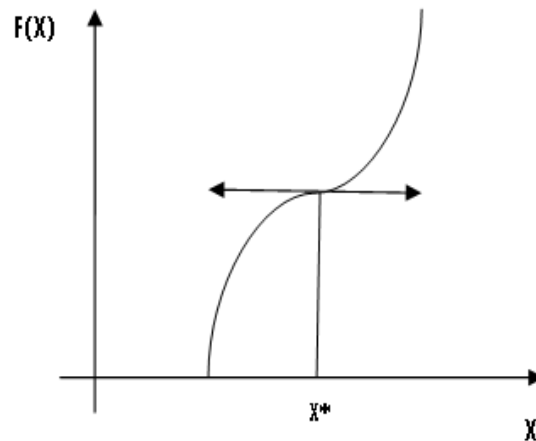


Figure 3.3 : x^* est un point stationnaire, mais ce n'est ni un minimum local ni maximum local.

Théorème 3.2.1.2 (Condition suffisante d'optimalité locale)

Soit $f: \mathbb{R}^n \rightarrow \mathbb{R}$ une fonction deux fois continument différentiable .soit x^* vérifiant :

a. $\nabla f(x^*) = 0.$

b. *la matrice hessienne $Hf(x^*)$ est définie positive.*

Alors x^* est un minimum local strict.

Remarque :

Pour un maximum local, la condition b. devient *$Hf(x^*)$ est définie négative.*

Exemple :

$$f(x) = 1 - e^{-x^2} \quad x \in \mathbb{R}$$

a. $\nabla f(x) = 2xe^{-x^2} = 0 \Rightarrow x = 0.$

b. $Hf(x^*) = 2e^{-x^2} - 4x^2e^{-x^2}$ Au point $x^* = (0), \Rightarrow 2 > 0$

donc $x^* = (0)$ est un minimum local car $Hf(x^*)$ est définie positive.

3.3 Les problèmes d'optimisation avec contraintes égalités

Définition 3.3.1 On définit un problème d'optimisation avec contraintes égalités par :

$$(p) \quad \begin{cases} \text{minimiser } f(x) \\ h_i(x) = 0 \quad i = 1, \dots, m \\ x \in \mathbb{R}^n \end{cases} \quad (3.3)$$

Où $f: \mathbb{R}^n \rightarrow \mathbb{R}, h_i: \mathbb{R}^n \rightarrow \mathbb{R}, i = 1, \dots, m,$ sont des fonctions continûment dérivables.

Pour simplifier on notera $H(x) = 0$ avec $H: \mathbb{R}^n \rightarrow \mathbb{R}^m.$

Définition 3.3.2 Soit $H(x) = (h_1(x), h_2(x), \dots, h_m(x))^t.$

$$h_i: \mathbb{R}^n \rightarrow \mathbb{R}, \quad i = 1, \dots, m,$$

$H(x) : \mathbb{R}^n \rightarrow \mathbb{R}^m$ supposons que $m < n.$

$$JH(x) = \begin{pmatrix} \nabla h_1^t(x) \\ \vdots \\ \nabla h_m^t(x) \end{pmatrix} = \begin{pmatrix} \frac{\partial h_1}{\partial x_1} & \dots & \frac{\partial h_1}{\partial x_m} & \dots & \frac{\partial h_1}{\partial x_n} \\ \frac{\partial h_2}{\partial x_1} & \dots & \frac{\partial h_2}{\partial x_m} & \dots & \frac{\partial h_2}{\partial x_n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \frac{\partial h_m}{\partial x_1} & \dots & \frac{\partial h_m}{\partial x_m} & \dots & \frac{\partial h_m}{\partial x_n} \end{pmatrix}$$

On dit que la matrice $JH(x)$ est de rang plein si son rang est égal à m , et on note $rg(JH(x)) = m$.

Définition 3.3.3. On dit que x^* est un point régulier si $rg(JH(x^*)) = m$.

Proposition 3.3.4 Soit x^* un minimum local régulier pour le problème (3.1).

Soit $x^* \in \mathbb{R}^n$ tel que : $H(x^*) = 0$

Alors les m premières composantes de x s'exprime en fonction de $(n - m)$ autre composantes de x

Preuve :

$$JH(x^*) = \begin{pmatrix} \frac{\partial h_1(x^*)}{\partial x_1} & \dots & \frac{\partial h_1(x^*)}{\partial x_m} & \frac{\partial h_1(x^*)}{\partial x_{m+1}} & \dots & \frac{\partial h_1(x^*)}{\partial x_n} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \frac{\partial h_m(x^*)}{\partial x_1} & \dots & \frac{\partial h_m(x^*)}{\partial x_m} & \frac{\partial h_m(x^*)}{\partial x_{m+1}} & \dots & \frac{\partial h_m(x^*)}{\partial x_n} \end{pmatrix}$$

Soit la formule de Taylor appliquer à H au voisinage de x^* .

$$H(x) = H(x^*) + JH(x^*)(x - x^*) + o(\|x - x^*\|)$$

Comme $H(x) = H(x^*) = 0$ on a forcément $JH(x^*)(x - x^*) = 0$ et $JH(x^*)(x) = JH(x^*)(x^*)$

$$\text{Soit } JH(x^*)(x) = JH(x^*)(x^*) = \alpha$$

x^* est régulier $\Rightarrow JH(x^*)$ est de rang plein $m \Rightarrow$ il y a m colonnes linéairement indépendantes

Supposons que les m premières colonnes soient linéairement indépendantes (si c'est pas le cas on permute les colonnes)

$$JH(x^*) = (JH_{m \times m}(x^*) \quad \vdots \quad JH_{m \times n-m}(x^*))$$

$$JH(x^*)x = (JH_{m \times m}(x^*) \quad \vdots \quad JH_{m \times n-m}(x^*)) \begin{pmatrix} x_m \\ \vdots \\ x_{n-m} \end{pmatrix}$$

$$= JH_{m \times m}(x^*)x_m + JH_{m \times n-m}(x^*)x_{n-m} = \alpha.$$

$$\Rightarrow JH_{m \times m}(x^*)x_m = \alpha - JH_{m \times n-m}(x^*)x_{n-m}.$$

$$x_m = [JH_{m \times m}(x^*)]^{-1}(\alpha - JH_{m \times n-m}(x^*)x_{n-m}) = \xi(x_{n-m}).$$

$$x = (x_m, x_{n-m})^t = (\xi(x_{n-m}), x_{n-m}).$$

Théorème 3.3.5 : (théorème de Lagrange-condition nécessaire de premier ordre)

Soit x^* un minimum local pour le problème (3.3) sous les contraintes $H(x) = 0$, et supposons que les gradients des contraintes $\nabla h_1(x^*), \dots, \nabla h_m(x^*)$ sont linéairement indépendants. alors il existe un vecteur unique.

$\lambda^* = (\lambda_1^*, \dots, \lambda_m^*)$ appelé vecteur des multiplicateurs de Lagrange, tel que :

$$\nabla f(x^*) + \sum_{i=1}^m \lambda_i^* \nabla h_i(x^*) = 0$$

Preuve : (indication)

D'après la proposition précédente :

$$H(x) = H(\xi(x_{n-m}), x_{n-m})$$

$$(H(x))_{x_{n-m}}^t = JH_{m \times m}(x)J\xi + JH_{m \times n-m}(x) = 0$$

$$\Rightarrow J\xi(x^*) = -(JH_{m \times m}(x^*))^{-1}JH_{m \times n-m}(x^*)$$

On pose :

$$\lambda^* = -\nabla f_m(x^*)(JH_{m \times m}(x^*))^{-1}.$$

$$\lambda^*JH_{m \times m}(x^*) = -\nabla f_m(x^*) \Rightarrow \nabla f_m(x^*) + \lambda^*JH_{m \times m}(x^*) = 0.$$

$$\nabla f_m(x^*) + \lambda^*JH_{m \times m}(x^*) = 0 \Leftrightarrow \frac{\partial f(x^*)}{\partial x_k} + \sum_{i=1}^m \lambda_i^* \frac{\partial h_i(x^*)}{\partial x_k} = 0, \quad k = 1, \dots, m.$$

Il reste les $(n - m)$ autres équations.

C'est-à-dire : $\frac{\partial f(x^*)}{\partial x_k} + \sum_{i=1}^m \lambda_i \frac{\partial h_i(x^*)}{\partial x_k}$, $k = m, m+1, \dots, n$.

On a :

$$f(x) = f(\xi(x_{n-m}), x_{n-m}) = F(x_{n-m}).$$

f admet un minimum local $x^* = (x_m^*, x_{n-m}^*)$ sous contraintes égalités.

Mais F admet un minimum local en x_{n-m}^* sans contraintes.

On a la condition nécessaire de premier ordre pour F :

$$\nabla F(x_{n-m}^*) = 0.$$

$$\nabla F(x_{n-m}^*) = \nabla f(\xi(x_{n-m}^*), x_{n-m}^*)$$

$$= \nabla f_m(x^*) J\xi(x^*) + \nabla f_{n-m}(x^*) = 0.$$

On remplace $J\xi(x^*)$ par sa valeur :

$$\nabla f_m(x^*) \left(- (JH_{m \times m}(x^*))^{-1} JH_{m \times n-m}(x^*) \right) + \nabla f_{n-m}(x^*) = 0.$$

On obtient :

$$\lambda^* JH_{m \times m-m}(x^*) + \nabla f_{n-m}(x^*) = 0.$$

$$\sum_{i=1}^m \lambda_i \frac{\partial h_i(x^*)}{\partial x_k} + \frac{\partial f(x^*)}{\partial x_k} = 0, \quad k = m+1, \dots, n.$$

Définition 3.3.6 (Fonction lagrangienne) Soit le problème d'optimisation :

$$(p) \quad \begin{cases} \text{minimiser } f(x) \\ h_i(x) = 0 \quad i = 1, \dots, m \\ x \in \mathbb{R}^n \end{cases}$$

Et soit le vecteur $\lambda \in \mathbb{R}^m$. La fonction $L: \mathbb{R}^{n+m} \rightarrow \mathbb{R}$ définie par

$$L(x, \lambda) = f(x) + \sum_{i=1}^m \lambda_i h_i(x)$$

Est appelée lagrangien ou fonction lagrangienne du problème (p).

Définition 3.3.7 On appelle sous espace tangent aux contraintes au point x^* noté

$$T(x^*) = \{d \in \mathbb{R}^n \mid JH(x^*)d = 0\}.$$

Proposition 3.3.8: soit x^* un point régulier admissible alors : $T(x^*)$ contient l'ensemble des directions admissibles au point x^* .

Preuve :

On a la formule de Taylor :

$$H(x^* + td) = H(x^*) + JH(x^*)td + o(\|x - x^*\|)$$

On a :

$$H(x^* + td) - H(x^*) = JH(x^*)td + o(\|x - x^*\|).$$

Pour que x^* et $x^* + td$ soient admissible (d est une direction admissible)

$$\Rightarrow JH(x^*)d + \frac{o(\|td\|)}{t} = 0$$

$$\Rightarrow JH(x^*)d = 0 \Rightarrow d \in T(x^*).$$

Exemple :

Soit $\{(x, y) \in \mathbb{R}^2 \mid x + y = 1\}$.

$x = y = \frac{1}{2}$, $(\frac{1}{2}, \frac{1}{2})$ est admissible.

$$T(x^*) = \{d \in \mathbb{R}^2 \mid JH(x^*)d = 0\}.$$

$$= \left\{ d \in \mathbb{R}^2 \mid (1, 1) \begin{pmatrix} d_1 \\ d_2 \end{pmatrix} = 0 \right\}.$$

$$= \{(d_1, d_2) \in \mathbb{R}^2 \mid d_1 + d_2 = 0\} \Rightarrow d_2 = -d_1.$$

$$(d_1, d_2) = (d_1, -d_1) = d_1(1, -1).$$

On a une direction admissible $(1, -1)$.

$(\frac{1}{2}, \frac{1}{2}) + t(1, -1) = (\frac{1}{2} + t, \frac{1}{2} - t)$ est admissible car $\frac{1}{2} + t + \frac{1}{2} - t = 1$.

Théorème 3.3.9 (Condition nécessaire du deuxième ordre)

On suppose que f et $g_i: g_i(x) = 0 \quad i = 1, \dots, m$ sont de classe c^2 .

Soit x^* un minimum régulier alors :

- a. $\exists \lambda_i^*$ tel que : $\nabla f(x^*) + \sum_{i=1}^m \lambda_i^* \nabla g_i(x^*) = 0$.
- b. $H_x L$ est semi définie positive sur $T(x^*)$

Preuve :

Pour plus de clarté, on suppose qu'on a le problème suivant :

$$\begin{cases} \min f(x, y) \\ g(x, y) = 0 \\ x, y \in \mathbb{R} \end{cases}$$

- a. Est déjà démontrée (condition nécessaire du premier ordre)
- b. On a : $X^* = (x^*, y^*)$ régulier. $JG(x^*, y^*) = 1$.
On peut exprimer $x = x(y)$.

$$g(x, y) = g(x(y), y) = 0.$$

$$g'(x(y), y) = g'_x \cdot x' + g'_y = 0 \quad \Rightarrow \quad x' = -(g'_x)^{-1} \cdot g'_y$$

$$g''(x(y), y) = (g'_x \cdot x' + g'_y)' = (g''_x \cdot x' + g''_{xy})x' + g'_x \cdot x'' + g''_{xy}x' + g''_{yy} = 0.$$

$$x'' = -(g'_x)^{-1} [x' g''_{xx} x' + 2x' g''_{xy} + g''_{yy}]$$

$$D'autre part : F(y) = f(x(y), y) \Rightarrow F'(y) = f'(x) \cdot x' + f'(y).$$

$$F''(y) = (f'_x \cdot x' + f'_y)' = (f''_{xx} \cdot x' + f''_{xy}) \cdot x' + f'_x \cdot x'' + f''_{xy} \cdot x' + f''_{yy}.$$

$$= x' f''_{xx} \cdot x' + 2x' f''_{xy} + f''_{yy} + f'_x [(-g'_x)^{-1} (x' g''_{xx} x' + 2x' g''_{xy} + g''_{yy})]$$

$$= x' (f''_{xx} + \lambda^* g''_{xx}) \cdot x' + 2x' (f''_{xy} + \lambda^* g''_{xy}) + (f''_{yy} + \lambda^* g''_{yy})$$

La condition nécessaire du deuxième ordre pour $F(y)$ sans contrainte :

$\xi F'' \xi \geq 0$ (on peut écrire $F'' \geq 0$ car $\xi \cdot \xi = \xi^2 \geq 0$).

$$\xi F'' \xi = \xi x' (f''_{xx} + \lambda^* g''_{xx}) \cdot x' \xi + 2\xi x' (f''_{xy} + \lambda^* g''_{xy}) \xi + \xi (f''_{yy} + \lambda^* g''_{yy}) \xi \geq 0.$$

$$= (\xi x' \quad \xi) \begin{pmatrix} L''_{xx} & L''_{xy} \\ L''_{yx} & L''_{yy} \end{pmatrix} \begin{pmatrix} \xi x' \\ \xi \end{pmatrix} \geq 0.$$

$$g(x, y) = 0.$$

$$JG(x, y) = (g'_x, g'_y)$$

$$JG(x, y) \begin{pmatrix} \xi x' \\ \xi \end{pmatrix} = (g'_x, g'_y) \begin{pmatrix} \xi x' \\ \xi \end{pmatrix}.$$

$$g'_x \xi x' + g'_y \cdot \xi = g'_x \xi (-g'_x)^{-1} \cdot g'_y + g'_y \cdot \xi$$

$$= -\xi g'_y + g'_y \xi = 0 \quad \Rightarrow (\xi x' \quad \xi) \in T(x^*) \Rightarrow$$

$H_x L$ est semi définie positive sur $T(x^*)$.

Théorème 3.3.10 condition suffisante du deuxième ordre

On suppose que f et $g_i : g_i(x) = 0 \quad i = 1, \dots, m$ sont de classe c^2 .

Soit x^* un point regulier :

- $\exists \lambda_i^*$ tel que : $\nabla f(x^*) + \sum_{i=1}^m \lambda_i^* \nabla g_i(x^*) = 0$.
- $H_x L$ est définie positive sur $T(x^*)$ alors x^* est un minimum local.

3.4 Les problèmes d'optimisation avec contraintes inégalités

Définition 3.4.1 On définit un problème d'optimisation avec contraintes inégalités par :

$$(p) \quad \begin{cases} \text{minimiser } f(x) \\ g_i(x) \leq 0 \quad i = 1, \dots, m \\ x \in \mathbb{R}^n \end{cases} \quad (3.4)$$

On note $G(x) \leq 0, G: \mathbb{R}^n \rightarrow \mathbb{R}^m$.

Définition 3.4.2 La contrainte $g_{i_0}(x) \leq 0$ est dite active ou saturée au point x^* si $g_{i_0}(x^*) = 0$.

On note l'ensemble des contraintes actives en x^* par $E = \{i/g_i(x^*) = 0\}$.

Lemme 3.4.3 (Qualification des contraintes)

Pour que la qualification des contraintes soit vérifiée en tout point $x \in S$ il suffit que l'une des conditions (i), (ii) et (iii) soit vérifiée :

- i. Toute les fonctions g_i sont linéaires (**Karlin 1959**)
- ii. Toutes les fonctions g_i sont convexes et S a un intérieur non vide (**Slater 1950**)
- iii. Pour que la qualification des contraintes soit vérifiée en tout point $x^* \in X$, il suffit que l'on ait : les gradients $\nabla g_i(x^*)$, $i \in E$ des contraintes saturées en x^* sont linéairement indépendantes (**Fiacco et McCormick**).

Théorème 3.4.4 (Théorème Karush, Kuhn et Tucker)

(Condition nécessaire de premier ordre)

Supposons que les fonctions f et $g_i, i = 1, \dots, m$ sont de classe C^1 . supposons que x^* est un minimum régulier pour le problème (1.4) alors il existe $\lambda_1^*, \dots, \lambda_m^* \in \mathbb{R}^+$ appelés multiplicateurs de **Karush, Kuhn et Tucker** tel que

$$\nabla f(x^*) + \sum_{i=1}^m \lambda_i^* \nabla g_i(x^*) = 0.$$

$$\lambda_i^* \nabla g_i(x^*) = 0, \quad i = 1, \dots, m,$$

$$\lambda_i \geq 0$$

Preuve :

On ramène le problème (3.4) à un problème avec contraintes égalités par l'ajout des variables d'écart, on aura donc

$$(p) \quad \begin{cases} \text{minimiser } f(x) \\ g_i(x) + z_i = 0 \quad i = 1, \dots, m \\ x \in \mathbb{R}^n, z \geq 0 \end{cases}$$

Où le lagrangien $L(x, z, \lambda) = f(x) + \sum_{i=1}^m \lambda_i (g_i(x) + z_i)$

D'après le théorème de Lagrange $\exists \lambda_i^* \in \mathbb{R}$ tel que

- i. $\nabla L_x(x^*, z^*, \lambda^*) = 0.$
- ii. $\nabla L_z(x^*, z^*, \lambda^*) = 0.$
- iii. $\nabla L_\lambda(x^*, z^*, \lambda^*) = 0.$

De (i) on a $\nabla f(x^*) + \sum_{i=1}^m \lambda_i^* \nabla g_i(x^*) = 0$

$$(ii) \quad \begin{cases} \lambda_i^* \geq 0 \text{ si } z_i^* = 0 \\ \lambda_i^* = 0 \text{ si } z_i^* > 0 \end{cases}$$

$$\Leftrightarrow \lambda_i^* z_i^* = 0, \lambda_i^* \geq 0, z_i^* \geq 0$$

$$\Leftrightarrow \lambda_i^* (-g_i(x^*)) = 0, \lambda_i^* \geq 0, z_i^* \geq 0$$

$$\Leftrightarrow \lambda_i^* g_i(x^*) = 0, \lambda_i^* \geq 0, z_i^* \geq 0.$$

Donc il existe $\lambda_i^* \geq 0$ tel que $\nabla f(x^*) + \sum_{i=1}^m \lambda_i^* \nabla g_i(x^*) = 0.$

$$\lambda_i^* g_i(x^*) = 0, \quad i = 1, \dots, m.$$

3.5 Les problèmes d'optimisations avec des contraintes mixtes [10]

Définition 3.5.1 On définit un problème d'optimisation avec contraintes égalités et inégalités par :

$$(p) \quad \begin{cases} \text{minimiser } f(x) \\ g_i(x) \leq 0 \quad i = 1, \dots, m \\ h_j(x) = 0 \quad j = 1, \dots, p \\ x \in \mathbb{R}^n \end{cases} \quad (3.5)$$

Où $f: \mathbb{R}^n \rightarrow \mathbb{R}$, $g_i: \mathbb{R}^n \rightarrow \mathbb{R}$, $i = 1 \dots m$, $h_j: \mathbb{R}^n \rightarrow \mathbb{R}$, $j = 1, \dots, p$.

Définition 3.5.2 Un point x^0 admissible pour le problème (3.5) est dit régulier si les gradients $\nabla g_i(x^0)$ et $\nabla h_j(x^0)$ sont linéairement indépendants i.e. ;

$$rg \begin{pmatrix} \nabla g_i(x^0) \\ \nabla h_j(x^0) \end{pmatrix} = |E| + P \quad i \in E.$$

Lemme 3.5.3 (Qualification des contraintes)

Pour que les qualifications des contraintes soient vérifiées en tout point $x \in S$ il suffit que l'une des contraintes (i) et (ii) soit vérifiée :

- i. Si les fonctions $g_i (i = 1, \dots, m)$ sont convexes et les fonctions $h_j (j = 1, \dots, p)$ sont linéaires et il existe \bar{x} tel que $g_i(\bar{x}) < 0, i = 1, \dots, m$ et $h_j(\bar{x}) = 0, j = 1, \dots, p$ alors la qualification des contraintes est vérifiée partout sur le domaine admissible (**Slater**).
- ii. $rg \begin{pmatrix} \nabla g_i(x^0) \\ \nabla h_j(x^0) \end{pmatrix} = |E| + P \quad i \in E \quad \text{(Fiacco et McCormick)}$.

Théorème 3.5.4 : (théorème de KKT)

(Condition nécessaire de premier ordre) [10]

Supposons que f , et les fonctions $g_i, i = 1, \dots, m$, et les $h_j, j = 1, \dots, p$ sont de classe C^1 .

Supposons que x^* est un minimum local régulier pour le problème (3.5), alors ils existent

$\lambda_1^*, \dots, \lambda_m^* \in \mathbb{R}^+$ et $\mu_1^*, \dots, \mu_p^* \in \mathbb{R}$ appelés multiplicateurs de Lagrange tel que :

$$\nabla f(x^*) + \sum_{i=1}^m \lambda_i^* \nabla g_i(x^*) + \sum_{j=1}^p \mu_j^* \nabla h_j(x^*) = 0.$$

$$\lambda_i^* g_i(x^*) = 0.$$

Preuve :

On ramène le problème (3.5) à un problème avec contraintes égalités en ajoutant des variables d'écart, on obtient le problème suivant :

$$(p) \quad \begin{cases} \text{minimiser } f(x) \\ g_i(x) + z_i = 0 & i = 1, \dots, m \\ h_j(x) = 0 & j = 1, \dots, p \\ x \in \mathbb{R}^n & z_i \geq 0. \end{cases}$$

On a le lagrangien

$$L(x, z, \lambda, \mu) = f(x) + \sum_{i=1}^m \lambda_i (g_i(x) + z_i) + \sum_{j=1}^p \mu_j h_j(x)$$

D'après le théorème de Lagrange il existe des $\lambda_i^* \in \mathbb{R} i = 1, \dots, m$ et $\mu_j^* \in \mathbb{R}, j = 1, \dots, p$ tel que :

- a. $\nabla_x L(x^*, z^*, \lambda^*, \mu^*) = 0.$
 b. $\nabla_z L(x^*, z^*, \lambda^*, \mu^*) = 0.$

c. $\nabla_{\lambda} L(x^*, z^*, \lambda^*, \mu^*) = 0.$

d. $\nabla_{\mu} L(x^*, z^*, \lambda^*, \mu^*) = 0.$

De (a) on a $\nabla f(x^*) + \sum_{i=1}^m \lambda_i^* \nabla g_i(x^*) = 0$

$$(b) \begin{cases} \lambda_i^* \geq 0 \text{ si } z_i^* = 0 \\ \lambda_i^* = 0 \text{ si } z_i^* > 0 \end{cases}$$

$$\Leftrightarrow \lambda_i^* z_i^* = 0, \lambda_i^* \geq 0, z_i^* \geq 0$$

$$\Leftrightarrow \lambda_i^* (-g_i(x^*)) = 0, \lambda_i^* \geq 0, z_i^* \geq 0$$

$$\Leftrightarrow \lambda_i^* g_i(x^*) = 0, \lambda_i^* \geq 0, z_i^* \geq 0.$$

(d) $\nabla_{\mu} L(x^*, z^*, \lambda^*, \mu^*) = h_j(x^*), j = 1, \dots, p$ sont admissibles

Conclusion :

$\lambda_1^*, \dots, \lambda_m^* \in \mathbb{R}^+$ et $\mu_1^*, \dots, \mu_p^* \in \mathbb{R}$ Tel que

$$\nabla f(x^*) + \sum_{i=1}^m \lambda_i^* \nabla g_i(x^*) + \sum_{j=1}^p \nabla h_j(x^*) = 0.$$

Et

$$\lambda_i^* g_i(x^*) = 0, i = 1, \dots, m.$$

Remarque :

Si la fonction f et les $g_i, i = 1, \dots, m$ sont convexes et les $h_j, j = 1, \dots, p$ sont convexes alors la condition devient suffisante et on obtient un minimum global.

3.6 Problèmes quadratiques

Définition 3.6.1 Un problème d'optimisation avec une fonction quadratique et des contraintes linéaires est appelé un programme quadratique. Le problème quadratique général peut s'écrire comme suit :

$$\text{minimiser } f(x) = \frac{1}{2} x^t D x + c^t x + g$$

$$Ax = b$$

$$Qx \leq h \tag{3.6.1}$$

$$x \in \mathbb{R}^n,$$

Où D est une (n, n) matrice symétrique, c, x sont des n -vecteurs, A est une (p, n) matrice avec $\text{rang}(A) = p$, Q est une $(m \times n)$ matrice ; $\text{rang}(Q) = m$, b est un p -vecteur, et h est un m -vecteur, $g \in \mathbb{R}$.

Remarque:

Si la matrice D est semi définie positive, on dit que le problème (3.6.1) est un problème quadratique convexe (le problème est dit strictement convexe si D est définie positive).

3.6.1 Conditions d'optimalité

3.6.1.1 Problème sans contraintes

Théorème 3.6.2.1 Considérons le problème

$$\min_{x \in \mathbb{R}^n} f(x) = \frac{1}{2} x^t D x + c^t x + g, \quad (3.6.2.1.1)$$

Où D est une matrice symétrique $n \times n$, $c \in \mathbb{R}^n$ et $g \in \mathbb{R}$.

1. Si D n'est pas semi-définie positive, alors le problème (3.6.2.1.1) ne possède pas de solution, c'est-à-dire qu'il n'existe aucun $x \in \mathbb{R}^n$ qui soit un minimum local de (3.6.2.1.1).
2. Si D est définie positive, alors

$$x^* = -D^{-1}c \quad (3.6.2.1.2)$$

Est l'unique minimum global de (3.6.2.1.1).

Preuve :

Nous avons $\nabla f(x) = Dx + c$ et $\nabla^2 f(x) = D$.

Supposons par l'absurde qu'il existe x^* minimum local de (3.6.2.1.1).

D'après le théorème (3.3.1), $\nabla^2 f(x) = D$ est semi-définie positive, ce qui contredit l'hypothèse.

Comme D est définie positive, le point x^* dans (3.6.2.1.2) est bien défini et

$$\nabla f(x^*) = -DD^{-1}c + c = 0.$$

Les condition suffisantes d'optimalité (3.1.2) sont vérifiées et x^* est un minimum local de f .

De plus comme D est définie positive f est strictement convexe donc x^* est l'unique minimum global.

3.6.1.2 Problème avec contraintes

Soit le problème d'optimisation quadratique avec contraintes d'égalité :

$$\min_{x \in \mathbb{R}^n} f(x) = \frac{1}{2} x^t D x + c^t x \quad (3.6.2.2.1)$$

Sous contrainte

$$Ax = b, \quad (3.6.2.2.2)$$

Où D une matrice symétrique $\in \mathbb{R}^{n \times n}$, $c \in \mathbb{R}^n$, $A \in \mathbb{R}^{m \times n}$ et $b \in \mathbb{R}^m$.

La fonction lagrangienne est

$$L(x, \lambda) = \frac{1}{2} x^t D x + c^t x + \lambda^t (b - Ax),$$

Avec $\lambda \in \mathbb{R}^m$. en appliquant directement le théorème (3.3.5) la condition nécessaire d'optimalité du premier ordre s'écrit

$$\nabla_x L(x, \lambda) = Dx + c - A^t \lambda = 0. \quad (3.6.2.2.3)$$

En combinant (3.6.2.2.2) et (3.6.2.2.3), on obtient le système linéaire

$$\begin{pmatrix} D & -A^t \\ A & 0 \end{pmatrix} \begin{pmatrix} x \\ \lambda \end{pmatrix} = \begin{pmatrix} -c \\ b \end{pmatrix}. \quad (3.6.2.2.4)$$

Montrons dans quel cas ce système possède une solution unique.

Lemme 3.6.2.3 Soit le problème quadratique (3.6.2.2.1)-(3.6.2.2.2), avec A de rang plein. Soit $Z \in \mathbb{R}^{n \times (n-m)}$ une matrice dont les colonnes forment une base de l'espace nul de A , c'est-à-dire $AZ = 0$, et Z de rang plein. Si la matrice hessienne réduite $Z^t D Z$ est définie positive, alors le système (3.6.2.2.4) est non singulier et possède une solution (x^*, λ^*) unique.

Preuve : Soit x et λ tels que

$$\begin{pmatrix} D & -A^t \\ A & 0 \end{pmatrix} \begin{pmatrix} x \\ \lambda \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix},$$

C'est-à-dire $Dx = A^t \lambda$ et $Ax = 0$. montrons que x et λ sont nuls pour prouver que la matrice est non singulière. Comme $Ax = 0$, on a

$$0 = (x^t \quad \lambda^t) \begin{pmatrix} D & -A^t \\ A & 0 \end{pmatrix} \begin{pmatrix} x \\ \lambda \end{pmatrix} = x^t D x.$$

Comme Z est de rang plein, il existe y tel que $x = Zy$. Dés lors,

$$y^t Z^t D Z y = 0.$$

Puisque $Z^t D Z$ est définie positive, alors $y = 0$. en conséquence, $x = Zy = 0$ et la première équation s'écrit

$$Dx - A^t \lambda = -A^t \lambda = 0.$$

Comme A est de rang plein, alors $\lambda = 0$.

3.7 La dualité de Lagrange

Définition 3.7.1 Soit le problème primal P

$$(p) \quad \begin{cases} \text{minimiser } f(x) \\ g_i(x) \leq 0 \quad i = 1, \dots, m \\ h_j(x) = 0 \quad j = 1, \dots, p \\ x \in \mathbb{R}^n \end{cases} \quad (3.7.1)$$

On définit la fonction de Lagrange associée au problème P par :

$$L(x, \lambda, \mu) = f(x) + \sum_{i=1}^m \lambda_i g_i(x) + \sum_{j=1}^p \mu_j h_j(x),$$

On définit le problème dual de p par :

$$D \quad \begin{cases} \text{Maximiser } \theta(\lambda, \mu) \\ \lambda \geq 0 \end{cases}$$

Où $\theta(\lambda, \mu) = \inf_x \{f(x) + \sum_{i=1}^m \lambda_i g_i(x) + \sum_{j=1}^p \mu_j h_j(x)\}$ est la fonction duale

Théorème 3.7.2 (Théorème de dualité faible)

Soit x une solution admissible de (p), soit encore (λ, μ) une solution admissible de D, $\lambda \geq 0$, alors $f(x) \geq \theta(\lambda, \mu)$.

Preuve :

Par définition de θ , on a :

$$\begin{aligned}\theta(\lambda, \mu) &= \inf_x \{f(x) + \lambda^t g(x) + \mu^t h(x)\}. \\ &\leq f(x) + \lambda^t g(x) + \mu^t h(x) \leq f(x)\end{aligned}$$

Puisque $\lambda \geq 0, g(x) \leq 0$ et $h(x) = 0$.

Corollaire 3.7.3

Si $\inf_x \{f(x) + \lambda^t g(x) + \mu^t h(x)\} = -\infty$ alors $\theta(\lambda, \mu) = -\infty$ pour certain $\lambda \geq 0$.

Corollaire 3.7.4.

Si $\sup \{\theta(\lambda, \mu) : \lambda \geq 0\} = \infty$, alors le problème primal n'admet pas de solution admissible.

Théorème 3.7.5

Soit $f: \mathbb{R}^n \rightarrow \mathbb{R}$, $g: \mathbb{R}^n \rightarrow \mathbb{R}^m$, des fonction convexes et soit $h: \mathbb{R}^n \rightarrow \mathbb{R}^p$ une fonction affine de forme $Ax - b = 0$. Supposons que les qualifications des contraintes sont vérifiées. Alors il existe $\bar{x} \in \mathbb{R}^n$ tel que $g(\bar{x}) < 0$ et $h(\bar{x}) = 0$, alors

$$\inf_x \{f(x) + \lambda^t g(x) + \mu^t h(x)\} = \sup \{\theta(\lambda, \mu)\}. \quad \lambda \geq 0.$$

Preuve :

Pour tout x admissible et tout $\lambda \geq 0, \mu \in \mathbb{R}^p$ on a $g(x) \leq 0$ et $h(x) = 0$ et du théorème (dualité faible) on obtient

$$\begin{aligned}\theta(\lambda, \mu) &= \inf_x \left\{ f(x) + \sum_{i=1}^m \lambda_i g_i(x) + \sum_{j=1}^p \mu_j h_j(x) \right\}. \\ &\leq f(x) + \sum_{i=1}^m \lambda_i g_i(x) + \sum_{j=1}^p \mu_j h_j(x) \leq f(x)\end{aligned}$$

En prenant le minimum sur tout x admissible on obtient

$$\theta(\lambda, \mu) \leq f(x^*) \quad \forall \lambda \geq 0, \mu \in \mathbb{R}^p \quad (3.7.5.1)$$

Où x^* est la solution optimale du primale.

D'après le théorème de **KKT** il existe $\lambda^* \in \mathbb{R}_+^m$, $\mu^* \in \mathbb{R}^p$ tel que

$$\lambda_i^* g_i(x^*) = 0, \quad i = 1, \dots, m,$$

En utilisant la définition de la fonction duale on obtient

$$\begin{aligned} \theta(\lambda^*, \mu^*) &= \inf_x \left\{ f(x) + \sum_{i=1}^m \lambda_i^* g_i(x) + \sum_{j=1}^p \mu_j^* h_j(x) \right\}. \\ &= f(x^*) + \sum_{i=1}^m \lambda_i^* g_i(x^*) + \sum_{j=1}^p \mu_j^* h_j(x^*) \\ &= f(x^*) \end{aligned} \quad (3.7.5.2)$$

Donc de (3.7.5.1) et (3.7.5.2) on voit que (λ^*, μ^*) est la solution optimale du problème dual et que $\theta(\lambda^*, \mu^*) = f(x^*)$

Définition 3.7.6 On définit le point selle du lagrangien associée à (3.4), un point $(\bar{x}, \bar{\lambda})$ tel que $\lambda \geq 0$

$$L(\bar{x}, \lambda) \leq L(\bar{x}, \bar{\lambda}) \leq L(x, \bar{\lambda}).$$

Où

$$\begin{aligned} L(\bar{x}, \bar{\lambda}) &= \min_x L(x, \bar{\lambda}) \\ &= \max_{\lambda \geq 0} L(\bar{x}, \lambda) \end{aligned}$$

Théorème 3.7.7 $(\bar{x}, \bar{\lambda})$ est un point selle pour $L(x, \lambda)$, avec $(\lambda \in \mathbb{R}_+^m)$ si et seulement si :

- a. $L(\bar{x}, \bar{\lambda}) \leq L(x, \bar{\lambda})$
- b. $g_i(\bar{x}) \leq 0 \quad i = 1, \dots, m.$
- c. $\bar{\lambda}_i g_i(\bar{x}) = 0 \quad i = 1, \dots, m.$

Preuve :

$(\bar{x}, \bar{\lambda})$ est un point selle :

- a) Est vérifiée par la définition du point selle .

b) On a : $\forall \lambda \geq 0 \quad L(\bar{x}, \lambda) \leq L(\bar{x}, \bar{\lambda})$

$$f(\bar{x}) + \sum_{i=1}^m \lambda_i g_i(\bar{x}) \leq f(\bar{x}) + \sum_{i=1}^m \bar{\lambda}_i g_i(\bar{x})$$

$$\sum_{i=1}^m (\lambda_i - \bar{\lambda}_i) g_i(\bar{x}) \leq 0 \quad (*)$$

Par l'absurde :

Si b) n'est pas vérifiée : $\exists i_0$ tq $g_{i_0}(\bar{x}) > 0$

Si on prend λ_{i_0} assez grand, on doit avoir $\sum_{i=1}^m (\lambda_i - \bar{\lambda}_i) g_i(\bar{x}) \geq 0$

Donc on a une contradiction avec (*) donc ce qu'on a supposé est faux

$$\Rightarrow g_i(\bar{x}) \leq 0, \quad i = 1, \dots, m.$$

b) est vérifiée.

c) si on prend $\lambda_i = 0, \quad i = 1, \dots, m.$ dans (*).

$$\sum_{i=1}^m -\bar{\lambda}_i g_i(\bar{x}) \leq 0 \quad \Rightarrow \quad \sum_{i=1}^m \bar{\lambda}_i g_i(\bar{x}) \geq 0 \quad (i)$$

On a:

$$g_i(\bar{x}) \leq 0 \quad i = 1, \dots, m.$$

$$\bar{\lambda}_i \geq 0 \quad i = 1, \dots, m.$$

$$\sum_{i=1}^m \bar{\lambda}_i g_i(\bar{x}) \leq 0 \quad (ii).$$

$$\text{De (i) et (ii)} \Rightarrow \sum_{i=1}^m \bar{\lambda}_i g_i(\bar{x}) = 0 \Rightarrow \bar{\lambda}_i g_i(\bar{x}) = 0, \quad i = 1, \dots, m.$$

c) est donc vérifiée.

Réciproquement :

$$\text{De a) } L(\bar{x}, \bar{\lambda}) \leq L(x, \bar{\lambda}) \quad \forall x \quad (1)$$

$$(c) \Rightarrow L(\bar{x}, \bar{\lambda}) = f(\bar{x}) + \sum_{i=1}^m \bar{\lambda}_i g_i(\bar{x}) = f(\bar{x}).$$

$$L(\bar{x}, \lambda) = f(\bar{x}) + \sum_{i=1}^m \lambda_i g_i(\bar{x}) \leq f(\bar{x}) = L(\bar{x}, \bar{\lambda}) \quad (2).$$

(1) et (2) $\Rightarrow (\bar{x}, \bar{\lambda})$ est un point selle.

Proposition 3.7.8 Si $(\bar{x}, \bar{\lambda})$ est un point selle du lagrangien associée au problème (3.4) alors \bar{x} est solution du problème (p).

Preuve :

$$(\bar{x}, \bar{\lambda}) \text{ est un point selle} \Rightarrow \begin{cases} a) & L(\bar{x}, \bar{\lambda}) \leq L(x, \bar{\lambda}) \\ b) & g_i(\bar{x}) \leq 0 & i = 1, \dots, m. \\ c) & \bar{\lambda}_i g_i(\bar{x}) = 0 & i = 1, \dots, m. \end{cases}$$

$$a) \quad L(\bar{x}, \bar{\lambda}) \leq L(x, \bar{\lambda})$$

$$f(\bar{x}) + \sum_{i=1}^m \bar{\lambda}_i g_i(\bar{x}) \leq f(x) + \sum_{i=1}^m \bar{\lambda}_i g_i(x) \leq f(x) \quad \forall x \text{ Verifies } g_i(x) \leq 0.$$

$$f(\bar{x}) \leq f(x) \forall x \text{ Vérifiant } g_i(x) \leq 0 \Leftrightarrow \bar{x} \text{ est un minimum global pour (p).}$$

Remarque:

C'est un résultat très général qui s'applique dans le cas général, mais pour certains problèmes, il peut ne pas exister de point selle.

C'est le cas en général pour les problèmes non convexes en particulier les problèmes discrets (problèmes en nombres entiers).

CHAPITRE IV

METHODES DE RESOLUTION EN
PROGRAMMATION QUADRATIQUE
CONVEXE

Introduction

Dans ce chapitre, nous introduisons les différentes approches développées pour la résolution des problèmes de programmation quadratique convexe avec contraintes d'inégalités linéaires. Dans la littérature, plusieurs méthodes ont été développées pour la réalisation des problèmes de programmation quadratique. Elles sont généralement des extensions de celles développées pour la programmation linéaire.

4.1 Méthode quadratique de Wolfe (1959)

Beaucoup d'algorithmes ont été développés pour la résolution du problème de programmation quadratique convexe, mais il serait intéressant de connaître la méthode la plus classique de Wolfe, qui n'est autre que la méthode du simplexe légèrement modifiée.

Le principe de cette méthode est la résolution du système de Kuhn-Tucher et consiste à trouver une solution réalisable pour un système linéaire avec une condition supplémentaire du type $x_j \delta_j = 0$, Où x et δ sont des vecteurs de même dimension. En d'autre terme c'est trouver une solution réalisable basique initiale en résolvant un problème de programmation linéaire, assujetti à la nouvelle condition. Voyons les cas suivant :

4.1.1 Cas d'un problème quadratique standard (P.Q.C)

Formulons le théorème de K.K.T pour le problème suivant :

$$\left\{ \begin{array}{l} F(x) = \frac{1}{2} x^t D x + c^t x \rightarrow \min \\ Ax = b, \\ x \geq 0, \end{array} \right. \quad (4.1.1.1)$$

Où $D^t = D \geq 0$, $c \in \mathbb{R}^n$, $b \in \mathbb{R}^m$, $\text{rang} A = m < n$.

Ce problème peut encore s'écrire sous la forme équivalente suivante :

$$\left\{ \begin{array}{l} F(x) = \frac{1}{2} x^t D x + c^t x \rightarrow \min \\ Ax - b \leq 0, \\ -Ax + b \leq 0, \\ -x \leq 0, \end{array} \right. \quad (4.1.1.2)$$

Le point de minimum x^* est alors caractérisé par les équations et les inégalités suivantes :

Il existe deux m -vecteurs $\lambda_1^* \geq 0$, $\lambda_2^* \geq 0$, ainsi qu'un n -vecteur $\delta^* \geq 0$ tels que

$$\begin{cases} \frac{\partial \mathcal{L}}{\partial x}(x^*, \lambda_1^*, \lambda_2^*, \delta^*) = 0, \\ \lambda_1^{*t}(Ax^* - b) = 0, \\ \lambda_2^{*t}(-Ax^* + b) = 0, \\ \delta^{*t}x^* = 0, \\ Ax^* - b = 0, \\ x^* \geq 0, \\ \delta^* \geq 0, \\ \lambda_1^* \geq 0, \\ \lambda_2^* \geq 0. \end{cases}$$

Où

$$\mathcal{L}(x, \lambda_1, \lambda_2, \delta) = \frac{1}{2}x^t D x + c^t x + \lambda_1^t(Ax - b) + \lambda_2^t(-Ax + b) - \delta^t x,$$

$$\frac{\partial \mathcal{L}}{\partial x}(x, \lambda_1, \lambda_2, \delta) = D x + c + A^t \lambda_1 - A^t \lambda_2 - \delta.$$

En posant $\lambda = \lambda_1 - \lambda_2$, $\lambda \in \mathbb{R}^m$, le point x^* est défini par le système suivant :

$$\begin{cases} D x^* + A^t \lambda^* - \delta^* = -c, & (L_1) \\ A x^* = b, & (L_2) \\ \delta^{*t} x^* = 0, & (L_3) \\ \lambda^{*t}(A x^* - b) = 0 & (L_4) \\ x^* \geq 0, \lambda^* \in \mathbb{R}^m, \delta^* \geq 0. \end{cases}$$

Comme l'équation (L_4) est tout le temps vérifiée, le système se réduit ainsi :

$$\begin{cases} D x^* + A^t \lambda^* - \delta^* = -c, & (L_1) \\ A x^* = b, & (L_2) \\ \delta^{*t} x^* = 0, & (L_3) \\ x^* \geq 0, \lambda^* \in \mathbb{R}^m, \delta^* \geq 0. \end{cases}$$

Un tel système n'est pas linéaire par rapport au multi-vecteur $(x^*, \lambda^*, \delta^*)$ à cause de l'équation (L_3) . On obtient donc un système linéaire de $(n + m)$ équations à $(n + m + n)$ inconnues, constitué des équations (L_1) et (L_2) , avec en plus n équations non linéaires $\delta_j x_j = 0, j = \overline{1, n}$. Pour trouver une solution telle que $\delta_j x_j = 0, j = \overline{1, n}$, il suffit d'obtenir une solution réalisable basique du système linéaire, avec x_j basique et δ_j non basique ou vice-versa. Pour cela, on appliquera la première phase du simplexe et on choisira l'indice j_0 du vecteur qui entre dans la base de telle sorte que les vecteurs-colonnes correspondant à x_{j_0} et δ_{j_0} ne se retrouvent pas en même temps dans la base. Pour appliquer la méthode du simplexe,

il faut alors écrire le système $(L_1), (L_2)$ sous forme standard, à savoir que le second membre doit être positif ou nul, ainsi que le vecteur λ^* qui doit être réécrit sous la forme :

$$\lambda_i^* = \alpha_i^* - \alpha_{m+i}^*, \quad \alpha_{m+i}^* \geq 0, \quad \alpha_i^* \geq 0, \quad i = \overline{1, m}.$$

4.1.2 Algorithme

Algorithme

Début

1. Introduire les données, D, A, b, c ;
 Appliquer les conditions de K.K.T au problème ;
 Déterminer les équations de K.K.T ;
2. Détermination des paramètres du programme linéaire ;
 Introduire les variables artificielle v_i ;
 Construire la matrice des contraintes A
 Construire le vecteur du second membre b
 Construire le vecteur des coûts c ;
3. Initialiser le vecteur solution (x, λ, δ, v) ;
 Déterminer l'ensemble des indices J_B et J_N ;
 Extraire les éléments de base x_B, c_B, A_B ;
4. Calculer le vecteur des potentiels $u^t = c_B^t A_B^{-1}$;
 Calculer le vecteur des estimations $E_N^t = u^t A_N - c_N^t$;
 Si $E_N^t \geq 0$ Alors la solution actuelle est optimale ;

Sinon aller à (5) ;

FinSi ;

5. Déterminer la variable qui entre en base tout en vérifiant la condition

$$\delta_j x_j = 0 \quad j = \overline{1, n} ;$$

Déterminer la variable qui sort de la base ;

Mettre à jour $(A_B, x_B, c_B, J_B, J_N)$ et **aller en 4** ;

Fin.

4.1.3 Exemple numérique

Soit le problème quadratique suivant ;

$$F(x) = x_1^2 + x_1x_2 + 6x_2^2 - 2x_1 + 8x_2 \rightarrow \min,$$

$$\text{S.c} \begin{cases} x_1 + 2x_2 \leq 4 \\ 2x_1 + x_2 \leq 5 \\ x_1, x_2 \geq 0. \end{cases} \quad (4.1.3.1)$$

Le problème (4.1.3.1) se formule ainsi :

$$F(x) = x_1^2 + x_1x_2 + 6x_2^2 - 2x_1 + 8x_2 \rightarrow \min,$$

$$\text{S.c} \begin{cases} x_1 + 2x_2 + x_3 = 4 \\ 2x_1 + x_2 + x_4 = 5 \\ x_1, x_2, x_3, x_4 \geq 0. \end{cases} \quad (4.1.3.2)$$

Si (x_1, x_2, x_3, x_4) est un point de minimum de la fonction $F(x)$, alors $\exists \lambda \in \mathbb{R}^2; \delta_1, \delta_2, \delta_3, \delta_4 \geq 0$ Tels que :

$$\left\{ \begin{array}{l} \frac{\partial \mathcal{L}}{\partial x_1}(x, \lambda, \delta) = 0, \\ \frac{\partial \mathcal{L}}{\partial x_2}(x, \lambda, \delta) = 0, \\ \frac{\partial \mathcal{L}}{\partial x_3}(x, \lambda, \delta) = 0, \\ \frac{\partial \mathcal{L}}{\partial x_4}(x, \lambda, \delta) = 0, \\ \text{et} \\ x_1 + 2x_2 + x_3 = 4, \\ 2x_1 + x_2 + x_4 = 5, \\ \delta_j x_j = 0 \\ x_j \geq 0, \delta_j \geq 0, j = \overline{1,4}, \end{array} \right. \quad (4.1.3.3)$$

Où

$$L(x, \lambda, \delta) = x_1^2 + x_1x_2 + 6x_2^2 - 2x_1 + 8x_2 + \lambda_1(x_1 + 2x_2 + x_3 - 4) + \lambda_2(2x_1 + x_2 + x_4 - 5) - \delta_1x_1 - \delta_2x_2 - \delta_3x_3 - \delta_4x_4.$$

Alors on obtient le système suivant :

$$\left\{ \begin{array}{l} \frac{\partial \mathcal{L}}{\partial x_1} = 2x_1 + x_2 - 2 + \lambda_1 + 2\lambda_2 - \delta_1 = 0, \\ \frac{\partial \mathcal{L}}{\partial x_2} = x_1 + 12x_2 + 8 + 2\lambda_1 + \lambda_2 - \delta_2 = 0, \\ \frac{\partial \mathcal{L}}{\partial x_3} = \lambda_1 - \delta_3 = 0 \Leftrightarrow \lambda_1 = \delta_3, \\ \frac{\partial \mathcal{L}}{\partial x_4} = \lambda_2 - \delta_4 = 0 \Leftrightarrow \lambda_2 = \delta_4, \\ \text{et} \\ x_1 + 2x_2 + x_3 = 4, \\ 2x_1 + x_2 + x_4 = 5, \\ \delta_j x_j = 0 \\ x_j \geq 0, \delta_j \geq 0, j = \overline{1,4}. \end{array} \right. \quad (4.1.3.4)$$

C'est-à-dire

$$\left\{ \begin{array}{ll} 2x_1 + x_2 + \delta_3 + 2\delta_4 - \delta_1 & = 2 & (L_1) \\ x_1 + 12x_2 + 2\delta_3 + \delta_4 - \delta_2 & = -8 & (L_2) \\ x_1 + 2x_2 + x_3 & = 4 & (L_3) \\ 2x_1 + x_2 + x_4 & = 5 & (L_4) \\ \delta_j x_j = 0 & & (L_5) \\ x_j \geq 0, \delta_j \geq 0, j = \overline{1,4} & & \end{array} \right.$$

En multipliant (L_2) par (-1) le système (4.1.3.4) s'écrit donc :

$$\left\{ \begin{array}{ll} 2x_1 + x_2 - \delta_1 + \delta_3 + 2\delta_4 & = 2 \\ -x_1 - 12x_2 + \delta_2 - 2\delta_3 - \delta_4 & = -8 \\ x_1 + 2x_2 + x_3 & = 4 \\ 2x_1 + x_2 + x_4 & = 5 \\ \delta_j x_j = 0 & \\ x_j \geq 0, \delta_j \geq 0, j = \overline{1,4} & \end{array} \right.$$

Nous avons obtenu un système linéaire de 4 équations avec 8 inconnues et 4 équations non linéaire. Pour trouver une solution telle que $\delta_j x_j = 0$ avec $j = \overline{1,4}$, il suffit d'obtenir une solution réalisable basique (x, δ) du système linéaire avec x_j basique et δ_j hors base ou vice-versa. Pour cela, il faut choisir l'indice J_0 entrant dans la base de telle sorte que les vecteurs-colonne correspondant à x_{j_0} et δ_{j_0} ne se trouvent pas en même temps dans la base.

Appliquons la première phase du simplexe en considérant le problème de programmation linéaire suivant :

$$\begin{aligned}
 & Z = -v_1 \rightarrow \max \\
 SC \left\{ \begin{array}{l}
 2x_1 + x_2 - \delta_1 + \delta_3 + 2\delta_4 + v_1 = 2 \\
 -x_1 - 12x_2 + \delta_2 - 2\delta_3 - \delta_4 = 8 \\
 x_1 + 2x_2 + x_3 = 4 \\
 2x_1 + x_2 + x_4 = 5 \\
 x_j \geq 0, \delta_j \geq 0, j = \overline{1,4} \quad v_1 \geq 0
 \end{array} \right. \quad (4.1.3.5)
 \end{aligned}$$

Le vecteur $\bar{x} = (x, \delta, v_1) = (0, 0, 4, 5, 0, 8, 0, 0, 2)$ est une solution réalisable initiale basique du problème (4.1.3.5), avec $A_B = (a_9, a_6, a_3, a_4) = I_4$. Dressons alors les tableaux simplexes suivants :

		\bar{x}	x_1	x_2	x_3	x_4	δ_1	δ_2	δ_3	δ_4	v_1		
		c	0	0	0	0	0	0	0	0	-1		
c_B	base	b	a_1	a_2	a_3	a_4	a_5	a_6	a_7	a_8	a_9	θ	
-1	a_9	2	2	1	0	0	-1	0	1	2	1	1	$\rightarrow j_1 = 9$
0	a_6	8	-1	-12	0	0	0	1	-2	-1	0	∞	
0	a_3	4	1	2	1	0	0	0	0	0	0	4	
0	a_4	5	2	1	0	1	0	0	0	0	0	5/2	
$z = -2$		E	-2	-1	0	0	1	0	-1	-2	0		
				$\uparrow j_0 = 2$									

Tab. 1. Phase 1 du simplexe

		\bar{x}	x_1	x_2	x_3	x_4	δ_1	δ_2	δ_3	δ_4	v_1	
		c	0	0	0	0	0	0	0	0	-1	
c_B	base	b	a_1	a_2	a_3	a_4	a_5	a_6	a_7	a_8	a_9	θ
0	a_9	1	1	1/2	0	0	-1/2	0	1/2	1	1/2	
0	a_6	9	0	0	0	0	-1/2	1	-3/2	0	1/2	
0	a_3	3	0	23/12	1	0	1/2	0	-1/2	-1	-1/2	
0	a_4	3	0	3/2	1	1	1	0	-1	-2	-1	
				0								
$z = 0$		E	0	0	0	0	0	0	0	0	1	

Tab. 2. Phase 2 du simplexe

Le critère d'optimalité étant vérifié, il s'ensuit que le vecteur $\bar{x} = (1,0,3,3,0,9,9,0,0,0)$ est une solution optimale du problème linéaire. Par conséquent, la solution réalisable basique du système d'optimalité du problème quadratique est le vecteur $(\hat{x}, \hat{\delta})$ tels que

$$\hat{x} = (\hat{x}_1, \hat{x}_2, \hat{x}_3, \hat{x}_4) = (1,0,3,3),$$

$$\hat{\delta} = (\hat{\delta}_1, \hat{\delta}_2, \hat{\delta}_3, \hat{\delta}_4) = (0,9,0,0).$$

Donc le point de minimum du problème (4.1.3.1) est le vecteur $x^* = (1,0)$, avec $F(x^*) = -1$.

4.2 Méthode direct de support pour la résolution d'un P.Q.C. standard

La méthode du simplexe quadratique de Wolfe est la méthode la plus classique pour la résolution d'un problème de programmation quadratique convexe. Dans ce qui suit, on applique la méthode de R. Gasbassov et F.M. Kirillova appelée Méthode Direct de support, pour la construction d'un algorithme de résolution d'un problème de programmation quadratique convexe, donné sous forme standard.

Le principe de cette méthode est le suivant : partant d'un plan de support initial, formé d'une solution réalisable et de deux matrices non dégénérées correspondant respectivement aux contraintes et à la fonction objectif, chaque itération consiste à trouver une direction d'amélioration et un pas maximal le long de cette direction de façon à améliorer la valeur de la fonction objectif, tout en s'assurant de ne pas sortir du domaine admissible déterminé par les contraintes du problème.

4.2.1 Position du problème et définitions

Le problème de la programmation quadratique convexe se présente sous la forme standard suivante :

$$\begin{cases} F(x) = \frac{1}{2} x^t D x + c^t x \rightarrow \min, & (4.2.1.1) \\ Ax = b, & (4.2.1.2) \\ x \geq 0, & (4.2.1.3) \end{cases}$$

Où $D^t = D \geq 0$, c et x sont des n -vecteurs, b est un m -vecteur, A une matrice d'ordre $m \times n$, avec $\text{rang} A = m < n$; $I = \{1,2, \dots, m\}$: l'ensemble des indices des lignes de A ,

$J = \{1,2, \dots, n\}$: l'ensemble des indices des colonnes de A .

Soit une partition de l'ensemble J telle que $J = \{1,2, \dots, n\} = J_B \cup J_N$, avec

$$J_B \cap J_N = \emptyset, |J_B| = m.$$

On peut alors écrire et fractionner les vecteurs et la matrice A de la manière suivante :
 $x = x(J) = (x_j, j \in J)$,

$$x = \begin{pmatrix} x_B \\ - \\ x_N \end{pmatrix}, \quad x_B = x(J_B) = (x_j, j \in J_B), \quad x_N = x(J_N) = (x_j, j \in J_N),$$

$$c = \begin{pmatrix} c_B \\ - \\ c_N \end{pmatrix}, \quad c_B = c(J_B) = (c_j, j \in J_B), \quad c_N = c(J_N) = (c_j, j \in J_N),$$

$$A = A(I, J) = (a_{ij}, \quad 1 \leq i \leq m, \quad 1 \leq j \leq n); \quad A = (a_j, j \in J), \quad a_j = \begin{pmatrix} a_{1j} \\ \vdots \\ a_{mj} \end{pmatrix},$$

$$A = (A_B/A_N), \quad A_B = A(I, J_B), \quad A_N = A(I, J_N).$$

Définition 4.2.1.1 Un vecteur x vérifiant les contraintes (4.2.1.2)-(4.2.1.3) est appelé *plan ou solution réalisable* du problème (4.2.1.1)-(4.2.1.3).

Définition 4.2.1.2 Un plan x^0 est dit *optimal* si

$$F(x^0) = \frac{1}{2} x^{0t} D x^0 + c^t x^0 = \min \frac{1}{2} x^t D x + c^t x,$$

Où x est pris parmi tous les plans du problème (4.2.1.2)-(4.2.1.3)

Définition 4.2.1.3 Un plan x^ϵ est appelé *ϵ -optimal* ou *sub-optimal* si

$$F(x^\epsilon) - F(x^0) = \frac{1}{2} (x^\epsilon)^t D x^\epsilon + c^t x^\epsilon - \frac{1}{2} (x^0)^t D x^0 - c^t x^0 \leq \epsilon,$$

Où x^0 est une solution optimale du problème (4.2.1.1)-(4.2.1.3) et ϵ est un nombre positif ou nul, donné à l'avance.

Définition 4.2.1.4 L'ensemble $J_B \subset J, |J_B| = m$, est appelé *support* des contraintes si

$$\det A_B = \det A(I, J_B) \neq 0.$$

Définition 4.2.1.5 le couple $\{x, J_B\}$, formé du plan x et du support J_B , est appelé *plan de support des contraintes*.

Définition 4.2.1.6 Le plan de support est dit *non dégénéré* si

$$x_j > 0, \quad \forall j \in J_B. \quad (4.2.1.4)$$

4.2.2 Formule d'Accroissement de la fonction objectif

Soit $\{x, J_B\}$ un plan de support des contraintes du problème (4.2.1.1)-(4.2.1.3). Considérons un autre plan quelconque $\bar{x} = x + \Delta x$.

L'accroissement de la fonction objectif s'écrit alors :

$$\begin{aligned}
 F(\bar{x}) - F(x) &= \frac{1}{2} \bar{x}^t D \bar{x} + c^t \bar{x} - \frac{1}{2} x^t D x - c^t x \\
 &= \frac{1}{2} (x + \Delta x)^t D (x + \Delta x) + c^t (x + \Delta x) - \frac{1}{2} x^t D x - c^t x \\
 &= (\Delta x)^t (D x + c) + \frac{1}{2} (\Delta x)^t D \Delta x,
 \end{aligned}$$

$$F(\bar{x}) - F(x) = g^t(x) \Delta x + \frac{1}{2} (\Delta x)^t D \Delta x \quad (4.2.2.1)$$

Où $g(x) = D x + c$ est le gradient de la fonction (4.2.1.1), avec $g(x) = g(J) = (g_j, j \in J)$, $g_B = g(J_B)$, $g_N = g(J_N)$.

Par ailleurs, on a $\begin{cases} A x = b \\ A \bar{x} = b \end{cases} \Leftrightarrow A \bar{x} = A(\Delta x + x) = A \Delta x + A x = b \Rightarrow A \Delta x = 0$.

En posant $\Delta x = \begin{pmatrix} \Delta x_B \\ - \\ \Delta x_N \end{pmatrix}$, $\Delta x_B = \Delta x(J_B)$, $\Delta x_N = \Delta x(J_N)$,

L'égalité $A \Delta x = 0$ s'écrit $A_B \Delta x_B + A_N \Delta x_N = 0$; d'où

$$\Delta x_B = -A_B^{-1} A_N \Delta x_N. \quad (4.2.2.2)$$

La formule (4.2.2.1) devient alors :

$$F(\bar{x}) - F(x) = g_B^t \Delta x_B + g_N^t \Delta x_N + \frac{1}{2} \begin{pmatrix} \Delta x_B \\ \Delta x_N \end{pmatrix}^t D \begin{pmatrix} \Delta x_B \\ \Delta x_N \end{pmatrix}.$$

En vertu de (4.2.2.2), on obtient donc

$$\begin{aligned}
 F(\bar{x}) - F(x) &= g_B^t (-A_B^{-1} A_N \Delta x_N) + g_N^t \Delta x_N \\
 &\quad + \frac{1}{2} (-A_B^{-1} A_N \Delta x_N, \Delta x_N)^t D (-A_B^{-1} A_N \Delta x_N, \Delta x_N) \\
 &= [-g_B^t A_B^{-1} A_N + g_N^t] \Delta x_N + \frac{1}{2} (\Delta x_N)^t \begin{pmatrix} -A_B^{-1} A_N \\ I_N \end{pmatrix}^t D \begin{pmatrix} -A_B^{-1} A_N \\ I_N \end{pmatrix} \Delta x_N,
 \end{aligned}$$

Où $I_N = I(J_N, J_N)$ est une matrice identité d'ordre $(n - m)$.

Posons

$$Z = Z(J, J_N) = \begin{pmatrix} -A_B^{-1} A_N \\ I_N \end{pmatrix}, M(J_N, J_N) = Z^t D Z, \quad (4.2.2.3)$$

Et définissons le vecteur des potentiels u ainsi que le vecteur des estimations E comme suit :

$$u^t = g_B^t A_B^{-1}, \quad E^t = E^t(J) = g^t - u^t A = (E_B^t, E_N^t),$$

Où

$$E_B^t = 0, \quad E_N^t = g_N^t - u^t A_N.$$

Finalement l'accroissement (4.2.2.) aura la forme suivante :

$$F(\bar{x}) - F(x) = E_N^t \Delta x_N + \frac{1}{2} (\Delta x_N)^t Z^t D Z (\Delta x_N).$$

$$F(\bar{x}) - F(x) = E_N^t \Delta x_N + \frac{1}{2} (\Delta x_N)^t M (\Delta x_N) \quad (4.2.2.4)$$

Soit J_{N0} et J_{N+} deux sous ensemble de J_N tels que

$$J_{N0} = \{j \in J_N / x_j = 0\}, \quad J_{N+} = \{j \in J_N / x_j > 0\}.$$

Les sous-vecteurs x_N et E_N peuvent alors être fractionnés sous la forme suivante :

$$x_N = \begin{pmatrix} x_{N0} \\ x_{N+} \end{pmatrix}, \quad x_{N0} = (x_j, j \in J_{N0}), \quad x_{N+} = (x_j, j \in J_{N+}).$$

$$E_N = \begin{pmatrix} E_{N0} \\ E_{N+} \end{pmatrix}, \quad E_{N0} = (E_j, j \in J_{N0}), \quad E_{N+} = (E_j, j \in J_{N+}).$$

La formule d'accroissement (4.2.2.4) peut s'écrire :

$$F(\bar{x}) - F(x) = E_{N0}^t \Delta x_{N0} + E_{N+}^t \Delta x_{N+} + \frac{1}{2} (\Delta x_N)^t M \Delta x_N. \quad (4.2.2.5)$$

4.2.3 Critère d'optimalité

Théorème. Soit $\{x, J_B\}$ un plan de support des contraintes du problème (4.2.1.1) – (4.2.1.3).

Alors les relations :

$$\begin{cases} E_j \geq 0 & \text{si } j \in J_{N0}, \\ E_j = 0 & \text{si } j \in J_{N+}, \end{cases} \quad j \in J_N \quad (4.2.3.1)$$

Sont suffisantes pour l'optimalité du plan x .

Ces mêmes relations sont aussi nécessaires, si le plan de support des contraintes est non-dégénéré.

Démonstration. Condition suffisante :

Soit $\{x, J_B\}$ un plan de support des contraintes vérifiant les relations (4.2.1.2.). Pour tout plan \bar{x} du problème (4.2.1.1)-(4.2.1.3), la formule d'accroissement (4.2.2.5) nous permet d'écrire :

$$\Delta F = F(\bar{x}) - F(x) \geq E_{N0}^t \Delta x_{N0} + E_{N+}^t \Delta x_{N+},$$

Car la matrice M est semi-définie positive.

D'où

$$F(\bar{x}) - F(x) \geq \sum_{j \in J_{N_0}} E_j(\bar{x}_j - x_j) + \sum_{j \in J_{N^+}} E_j(\bar{x}_j - x_j).$$

En vertu des relations (4.2.3.1), on aura

$$F(\bar{x}) - F(x) \geq \sum_{j \in J_{N_0}} E_j \bar{x}_j \geq 0,$$

Car \bar{x} est un plan du problème (4.2.1.1)-(4.2.1.3). D'où $F(\bar{x}) \geq F(x)$.

Le vecteur x est par conséquent une solution optimale du problème (4.2.1.1)-(4.2.1.3).

Condition nécessaire : Soit $\{x, J_B\}$ un plan de support optimal non-dégénéré du problème (4.2.1.1)-(4.2.1.3) et supposons que les relations (4.2.3.1) ne sont pas vérifiées, c.-à-d., qu'il existe au moins un indice $j_0 \in J_N$, tel que

$$E_{j_0} < 0, \text{ pour } j_0 \in J_{N_0} \text{ ou bien } E_{j_0} \neq 0, \text{ pour } j_0 \in J_{N^+}.$$

On construit alors un autre plan $\bar{x} = x + \theta l$, où θ est un nombre réel positif, et $l = l(J)$ est un vecteur de direction que l'on construit comme suit :

$$\begin{cases} l_{j_0} = -\text{sign}E_{j_0}, \\ l_j = 0, j \neq j_0, j \in J_N. \end{cases}$$

D'autre part, on doit avoir

$$A\bar{x} = Ax + \theta Al = b \Leftrightarrow Al = A_B l_B + A_N l_N = 0.$$

$$\text{D'où } l_B = -A_B^{-1} A_N l_N = A_B^{-1} a_{j_0} \text{sign}E_{j_0}.$$

On a donc

$$\begin{cases} l_{j_0} = -\text{sign}E_{j_0}, \\ l_j = 0, j \neq j_0, j \in J_N, \\ l_B = A_B^{-1} a_{j_0} \text{sign}E_{j_0} \end{cases}$$

Le vecteur \bar{x} vérifie la contrainte principale $A\bar{x} = b$. Pour que \bar{x} soit un plan du problème (4.2.1.1)-(4.2.1.3), il doit en plus vérifier l'inégalité

$$\bar{x} \geq 0 \Leftrightarrow x + \theta l \geq 0.$$

Soit, en écrivant composante par composante

$$\begin{cases} \theta l_j \geq -x_j, & j \in J_N, \\ \theta l_{j_0} \geq -x_{j_0}. \end{cases} \Leftrightarrow \begin{cases} \theta l_j \in [-x_j, +\infty[, \\ \theta \text{sign}E_{j_0} \leq x_{j_0} \end{cases}$$

Pour $j_0 \in J_{N_0}$, on a $E_{j_0} < 0$ et $x_{j_0} = 0 \Rightarrow \theta \geq 0$.

Pour $j_0 \in J_{N^+}$, on a $E_{j_0} \neq 0 \Rightarrow E_{j_0} > 0$ ou $E_{j_0} < 0$.

Si $E_{j_0} > 0$ comme $x_{j_0} > 0$, on a $\theta \text{sign} E_{j_0} \leq x_{j_0} \Rightarrow \theta \leq x_{j_0}$.

Si $E_{j_0} < 0$ comme $x_{j_0} > 0$, on a $\theta \text{sign} E_{j_0} \leq x_{j_0} \Rightarrow \theta \geq -x_{j_0}$.

Comme $x_j > 0$, $j \in J_B$, le vecteur \bar{x} sera alors un plan du problème (4.2.1.1)-(4.2.1.3) pour un nombre θ positif assez petit.

La formule d'accroissement (4.2.2.4) nous donne :

$$\begin{aligned}
 F(\bar{x}) - F(x) &= E_N^t \Delta x_N + \frac{1}{2} \Delta^t x_N M \Delta x_N \\
 &= \theta E_N^t l_N + \frac{1}{2} \theta^2 l_N^t M l_N \\
 &= \theta (-E_{j_0} \text{sign} E_{j_0} + \frac{1}{2} \theta l_N^t M l_N) \\
 &= \theta \left(-|E_{j_0}| + \frac{1}{2} \theta l_N^t M l_N \right) = \varphi(\theta). \tag{4.2.3.2}
 \end{aligned}$$

Pour θ et l ainsi choisis on aura $-|E_{j_0}| + \frac{1}{2} \theta l_N^t M l_N < 0$. D'où $F(\bar{x}) - F(x) < 0$, ce qui contredit l'optimalité de x .

Par conséquent, les relations (4.2.3.1) sont suffisantes, et aussi nécessaires pour l'optimalité du plan x dans le cas où x est non dégénéré.

4.2.4 Critère de sub-optimalité

Pour estimer l'écart qui existe entre la valeur optimale $F(x^0)$ et une autre valeur $F(x)$ d'un plan de support des contraintes quelconque $\{x, J_B\}$, il suffit de remplacer dans la formule d'accroissement (4.2.8). le vecteur \bar{x} par x^0 et de minorer l'expression. On aura donc :

$$F(x^0) - F(x) \geq \sum_{j \in J_N} E_j (x_j^0 - x_j).$$

D'où

$$F(x) - F(x^0) \leq \sum_{j \in J_N} E_j (x_j - x_j^0).$$

Puisque la solution optimale x^0 est admissible et en supposant que $E_N \geq 0$, nous aurons

$$\langle x_j^0 \geq 0, j \in J \rangle \Rightarrow \langle x_j - x_j^0 \leq x_j - 0 \rangle \Rightarrow \langle E_j (x_j - x_j^0) \leq E_j x_j \rangle.$$

Par conséquent, on obtient la majoration suivante :

$$F(x) - F(x^0) \leq \sum_{j \in J_N} E_j x_j = \beta(x, J_B). \quad (4.2.3.3)$$

Le nombre $\beta(x, J_B)$ est appelé *estimation de sub-optimalité*. On a alors le théorème suivant :

Théorème 4.2.4 (Condition suffisante de sub-optimalité)

Soit $\{x, J_B\}$ un plan de support des contraintes du problème (4.2.1.1)-(4.2.1.3) et ϵ un nombre positif ou nul arbitraire.

Si $E_N \geq 0$ et si $\beta(x, J_B) \leq \epsilon$, alors le plan x est ϵ -optimal.

Démonstration.

En vertu de (4.2.3.3), nous avons

$$F(x) - F(x^0) \leq \sum_{j \in J_N} E_j x_j = \beta(x, J_B) \leq \epsilon \Rightarrow F(x) - F(x^0) \leq \epsilon.$$

Alors le plan x est donc ϵ -optimal.

4.2.5 Méthode de résolution

Avant d'entamer la méthode de résolution donnons quelques définitions essentielles.

Définition 4.2.5 On appelle support de la fonction (4.2.1.1) l'ensemble des indices $J_S \subset J_N$ tel que $\det M_S = \det M(J_S, J_S) \neq 0$, où M est la matrice (4.2.2.3). on posera $J_{NN} = J_N \setminus J_S$.

- On appelle support du problème (4.2.1.1)-(4.2.1.3) l'ensemble $J_P = \{J_B, J_S\}$ formé du support des contraintes J_B et de celui de la fonction objectif J_S .
- On appelle plan de support du problème (4.2.1.1)-(4.2.1.3) la paire $\{x, J_P\}$ formé du plan x et du support J_P , il est dit accordé si $E(J_S) = 0$.
- La direction l est dite admissible si $Al = 0$.

Elle est dite d'amélioration si en outre $E^t l < 0$ ($\varphi(\theta) \leq 0$).

4.2.6 Algorithme de résolution

Étant donné un nombre réel positif ou nul quelconque ϵ et un plan de support initial $\{x, J_P\}$, le but de l'algorithme est alors de construire un plan ϵ -optimal x^ϵ ou carrément un plan optimal x^0 .

L'itération de l'algorithme consiste à faire le passage de $\{x, J_P\}$ vers $\{\bar{x}, \bar{J}_P\}$ tel que

$F(\bar{x}) \leq F(x)$. Pour cela, construisons le nouveau plan \bar{x} de la manière suivante :

$\bar{x} = x + \theta l$, où l est un n -vecteur appelé direction d'amélioration ; θ est le pas le long de cette direction.

Dans cet algorithme, on choisira la métrique du simplexe. On ne fera donc varier qu'une seule composante parmi celles qui ne vérifient pas les relations (4.2.3.1).

Pour que l'accroissement soit maximal, il faut choisir l'indice J_0 tel que :

$$|E_{j_0}| = \max(|E_j|, j \in J_{NNO}),$$

Où $J_{NNO} \subset J_N$ est l'ensemble des indices non optimaux. On calcule l_{NN} de manière à assurer $E^t l < 0$, on posera alors :

$$l_{j_0} = -\text{sign}E_{j_0}, \quad l_j = 0, j \neq j_0, \quad j \in J_{NN}.$$

On calculera la composante l_S de manière à assurer $\bar{E}_j = E_j(x + \theta l) = 0, j \in J_S$. En vertu de(4.2.2.3) nous avons :

$$E_N^t = (g_B^t, g_N^t) \begin{pmatrix} -A_B^{-1}A_N \\ I_N \end{pmatrix} = g^t(x)Z.$$

Comme $l = Zl_N$, on aura donc

$$\bar{E}_N^t = g^t(x + \theta^0 l)Z = E_N^t + \theta^0 l^t DZ = E_N^t + \theta^0 l_N^t Z^t DZ = E_N^t + \theta^0 l_N^t M.$$

Finalement, on a

$$\bar{E}_N = E_N + \theta^0 M l_N.$$

Puisque $E(J_S) = 0$, alors l'équation $\bar{E}(J_S) = 0$ est équivalente à

$$M(J_S, J_S)l(J_S) + M(J_S, J_{NN})l(J_{NN}) = 0.$$

D'où

$$l(J_S) = -M_S^{-1}M(J_S, J_{NN})l(J_{NN}).$$

Ensuite, on calcule l_B de manière à avoir $Al = 0$:

$$l_B = l(J_B) = -A_B^{-1}(A_S l_S + A_{NN} l_{NN}).$$

Alors pour l'indice j_0 , nous avons les formules suivantes pour la construction de la direction d'amélioration $l = (l_j, j \in J) = (l(J_B), l(J_S), l(J_{NN}))$:

$$\begin{cases} l_{j_0} = -\text{sign}E_{j_0}, \\ l_j = 0, j \neq j_0, j \in J_{NN}, \\ l(J_S) = +M_S^{-1}M(J_S, J_0)\text{sign}E_{j_0}, \\ l(J_B) = -A_B^{-1}[A(I, J_S)l(J_S) - A(I, J_0)\text{sign}E_{j_0}]. \end{cases} \quad (4.2.6)$$

D'autre part, le pas θ doit vérifier les relations suivantes :

1. $-\theta l_j \leq x_j, j \in J_B.$
2. $-\theta l_j \leq x_j, j \in J_S.$
3. $\theta \text{sign} E_{j_0} \leq x_{j_0}.$

En calculons les différentes valeurs maximales que peut prendre le pas θ dans ces relations, on aura

$$\theta_{j_1} = \min_{j \in J_B}(\theta_j), \quad \text{où } \theta_j = \begin{cases} -x_j/l_j & \text{si } l_j < 0, \\ \infty, & \text{si } l_j \geq 0. \end{cases}$$

$$\theta_{j_S} = \min_{j \in J_S}(\theta_j), \quad \text{où } \theta_j = \begin{cases} -x_j/l_j & \text{si } l_j < 0, \\ \infty, & \text{si } l_j \geq 0. \end{cases}$$

$$\theta_{j_0} = \begin{cases} +x_{j_0} & \text{si } E_{j_0} > 0, \\ \infty, & \text{si } E_{j_0} < 0. \end{cases}$$

Ensuite, déterminons le pas θ_F pour lequel le passage de x à \bar{x} doit assurer une diminution maximale de la fonction objectif, i.e., d'après (4.2.3.2) nous devons avoir :

$$\frac{\partial \varphi(\theta)}{\partial \theta} = 0 \Leftrightarrow -|E_{j_0}| + \theta l'_N M l_N = 0.$$

On en déduit que la valeur de θ_F est égale à :

$$\theta_F = \begin{cases} \frac{|E_{j_0}|}{l'_N M l_N}, & \text{si } l'_N M l_N > 0, \\ \infty, & \text{si } l'_N M l_N = 0. \end{cases}$$

Par conséquent, le pas maximal θ^0 Le long de la direction l vaut :

$$\theta^0 = \min(\theta_{j_0}, \theta_{j_1}, \theta_{j_S}, \theta_F).$$

Le nouveau plan s'écrit $\bar{x} = x + \theta^0 l$. si $\bar{E}_N \geq 0$ et $\beta(x, J_B) \leq \epsilon$, alors le plan \bar{x} est ϵ optimal et on peut arrêter l'algorithme.

Sinon, on changera J_P de la manière suivante :

- Si $\theta^0 = \theta_{j_0}$, alors $\bar{J}_B = J_B, \bar{J}_S = J_S, \bar{J}_P = J_P$;
- Si $\theta^0 = \theta_{j_1}$, alors $\bar{J}_B = (J_B \setminus j_1) \cup j_0, \bar{J}_S = J_S, \bar{J}_P = \{\bar{J}_B, \bar{J}_S\}$;
- Si $\theta^0 = \theta_{j_S}$, alors $\bar{J}_B = J_B, \bar{J}_S = J_S \setminus j_S, \bar{J}_P = \{\bar{J}_B, \bar{J}_S\}$;
- Si $\theta^0 = \theta_F$, alors $\bar{J}_B = J_B, \bar{J}_S = J_S \cup j_0, \bar{J}_P = \{\bar{J}_B, \bar{J}_S\}$;

On recommencera alors une nouvelle itération avec le nouveau plan de support accordé $\{\bar{x}, \bar{J}_P\}$.

Remarque : le passage de $\{\mathbf{x}, J_P\} \rightarrow \{\bar{\mathbf{x}}, \bar{J}_P\}$ assure les conditions :

$$\det \bar{A}_B \neq 0, \quad \det \bar{M}_S \neq 0, \quad \text{et} \quad \bar{E}(\bar{J}_S) = 0.$$

4.2.7 Finitude de la méthode

La finitude de la méthode de résolution présentée ci-dessus est garantie pour peu qu'une certaine règle soit observée, à savoir qu'il faut commencer par changer en premier lieu les composantes non optimales, ne correspondant pas aux composantes critiques du plan courant \mathbf{x}

Soit J_{NNO} l'ensemble des indices non optimaux, formé des sous-ensembles suivant :

$$J_{NNO} = J_E \cup J_{cr},$$

Avec $J_E = \{j \in J_{NN} / x_j > 0, E_j \neq 0\}$ et $J_{cr} = \{j \in J_{NN} / x_j = 0, E_j < 0\}$

Théorème 4.2.7 Tant que $J_E \neq \emptyset$, l'indice j_0 est toujours choisi dans J_E .

4.2.8 Algorithme de la méthode

La méthode est résumée dans l'algorithme suivant :

Algorithme

Début

0. Soit un nombre réel positif ou nul quelconque ϵ et un plan de support initial $\{\mathbf{x}, J_P\}$ tel que $J_P = \{J_B, J_S\}$, avec $J_S \neq \emptyset$ pour plus de facilité ;
1. Calculer le vecteur des estimations $E_N^t = g_N^t - u^t A_N$;
2. Test d'optimalité du plan de support $\{\mathbf{x}, J_P\}$;
 - Si $E_N \geq 0$ alors
 - Calculer la valeur de sub-optimalité $\beta(\mathbf{x}, J_B)$;
 - Si $\beta(\mathbf{x}, J_B) = 0$; Alors
 - Le processus de résolution s'arrête avec $\{\mathbf{x}, J_P\}$ plan de support optimal ;
 - Fin Si
 - Si $\beta(\mathbf{x}, J_B) \leq \epsilon$; Alors
 - Le processus de résolution s'arrête avec $\{\mathbf{x}, J_P\}$ plan de support ϵ - optimal ;
 - Fin Si
 - Si $\beta(\mathbf{x}, J_B) > \epsilon$ Alors aller en (3) ;
 - Fin Si
 - Sinon ($E_N \not\geq 0$) Aller directement en (3) ;

Fin Si

3. Changement du plan x en $\bar{x} : \bar{x} = x + \theta^0 l$.
 - Choisir l'indice j_0 tel que $|E_{j_0}| = \max(|E_j|, j \in J_{NNO})$, où J_{NNO} est l'ensemble des indices non optimaux ;
 - Calculer la direction d'amélioration l ;
 - Calculer le pas $\theta^0 = \min(\theta_{j_0}, \theta_{j_1}, \theta_{j_S}, \theta_F)$.
 - Calculer $\bar{x} = x + \theta^0 l$;
4. Test d'optimalité du nouveau plan \bar{x} ;
 - Si $\bar{E}_N \geq 0$ Alors
 - Si $\beta(\bar{x}, J_B) \leq \epsilon$, Alors
 - Le processus de résolution s'arrête avec $\{\bar{x}, J_P\}$ plan du support ϵ -optimal;
 - Sinon aller en 5 ;
 - Fin Si
 - Sinon ($\bar{E}_N \not\geq 0$) Aller directement en 5 ;
 - Fin Si
 - 5. Changement de support J_P en \bar{J}_P ;
 - Si $\theta^0 = \theta_{j_0}$ Alors

$$\bar{J}_B = J_B; \quad \bar{J}_S = J_S; \quad \bar{J}_P = J_P;$$
 - Fin Si
 - Si $\theta^0 = \theta_{j_1}$ Alors

$$\bar{J}_B = (J_B \setminus j_1) \cup j_0, \quad \bar{J}_S = J_S, \quad \bar{J}_P = \{\bar{J}_B, \bar{J}_S\};$$
 - Fin Si
 - Si $\theta^0 = \theta_{j_S}$ Alors

$$\bar{J}_B = J_B, \quad \bar{J}_S = J_S \setminus j_S, \quad \bar{J}_P = \{\bar{J}_B, \bar{J}_S\};$$
 - Fin Si
 - Si $\theta^0 = \theta_{j_F}$ Alors

$$\bar{J}_B = J_B, \quad \bar{J}_S = J_S \cup j_0, \quad \bar{J}_P = \{\bar{J}_B, \bar{J}_S\};$$
 - Fin Si
 - Aller en 1 avec le nouveau plan de support (\bar{x}, \bar{J}_P) .

Fin.

4.2.9 Exemple numérique

Soit le problème de programmation quadratique suivant :

$$F(x) = x_1^2 + x_1 x_2 + 6x_2^2 - 2x_1 + 8x_2 \rightarrow \min,$$

$$\begin{cases} x_1 + 2x_2 + x_3 = 4 \\ 2x_1 + x_2 + x_4 = 5 \\ x_1, x_2, x_3, x_4 \geq 0. \end{cases}$$

On a

$$D = \begin{pmatrix} 2 & 1 & 0 & 0 \\ 1 & 12 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, c = \begin{pmatrix} -2 \\ 8 \\ 0 \\ 0 \end{pmatrix} b = \begin{pmatrix} 4 \\ 5 \end{pmatrix} A = \begin{pmatrix} 1 & 2 & 1 & 0 \\ 2 & 1 & 0 & 1 \end{pmatrix}.$$

Soit $x = (0,0,4,5)$ un plan initial du problème. Posons $J_B = \{3, 4\}$, $J_N = \{1, 2\}$.

Nous avons alors $A_B = (a_3, a_4) = I_2$ et $A_N = (a_1, a_2) = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}$, $A_B^{-1}A_N = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}$.

Déterminant la matrice $M = Z^t D Z$ avec

$$Z = Z(J, J_N) = \begin{pmatrix} -A_B^{-1}A_N \\ I_N \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ -1 & -2 \\ -2 & -1 \end{pmatrix}$$

D'où

$$M = Z^t D Z = \begin{pmatrix} 1 & 0 & -1 & -2 \\ 0 & 1 & -2 & -1 \end{pmatrix} \begin{pmatrix} 2 & 1 & 0 & 0 \\ 1 & 12 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ -1 & -2 \\ -2 & -1 \end{pmatrix} = \begin{pmatrix} 2 & 1 \\ 1 & 12 \end{pmatrix}.$$

Calculons le vecteur gradient $g(x) = (g_B, g_N)$:

$$g(x) = D x + c = \begin{pmatrix} 2 & 1 & 0 & 0 \\ 1 & 12 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 0 \\ 0 \\ 4 \\ 5 \end{pmatrix} + \begin{pmatrix} -2 \\ 8 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} -2 \\ 8 \\ 0 \\ 0 \end{pmatrix},$$

$$g_B = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \quad \text{et} \quad g_N = \begin{pmatrix} -2 \\ 8 \end{pmatrix}.$$

D'où le vecteur des estimations :

$$E_N = g_N - (g_B^t A_B^{-1} A_N)^t = \begin{pmatrix} -2 \\ 8 \end{pmatrix}.$$

Posons $J_S = \emptyset$, $J_{NN} = J_N - J_S = \{1, 2\}$. La paire $\{x, J_P\}$, avec $J_P = \{J_B, J_S\}$, est alors le plan de support du problème considéré.

Itération 1

Le critère d'optimalité n'étant pas vérifié pour l'indice $j_0 = 1$,

-Calculons alors la direction d'amélioration l :

$$\begin{cases} l_{j_0} = l_1 = -\text{sign}E_1 = 1, & \text{car } E_1 < 0 \\ l_{j_0} = 0, & \text{si } j \neq j_0 \end{cases}$$

$$l(J_{NN}) = \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

$$l(J_S) = (l_j, j \in J_S) = 0, \text{ car } (J_S = \emptyset)$$

$$l(J_B) = (l_j, j \in J_B) = -A_B^{-1}A(I, 1) = \begin{pmatrix} -1 \\ -2 \end{pmatrix}$$

-Calculons le pas θ^0 le long de cette direction :

$$\theta_{j_0} = \theta_1 = \infty, \quad \text{car } E_1 < 0,$$

$$\theta_{j_1} = \min_{j \in J_B} \theta_j = \min\{\theta_3, \theta_4\}, \quad \text{avec } \begin{cases} \theta_3 = 4 \\ \theta_4 = 5/2 \end{cases}$$

$$\theta_{j_1} = \theta_4 = 5/2 \Rightarrow j_1 = 4$$

$$\theta_{j_S} = \infty, \quad \text{car } J_S = \emptyset,$$

$$\theta_F = \frac{|E_1|}{\alpha}, \quad \text{avec } \alpha = l_N^t M l_N = 2,$$

$$\theta_F = 2/2 = 1.$$

Donc $\theta^0 = \min(\theta_{j_0}, \theta_{j_1}, \theta_{j_S}, \theta_F) = \theta_F = 1$. on a alors le nouveau plan :

$$\bar{x} = x + \theta^0 l = \begin{pmatrix} 1 \\ 0 \\ 3 \\ 3 \end{pmatrix}$$

Donc

$$\bar{J}_B = J_B = \{3,4\}; \quad \bar{J}_S = \{1\}; \quad \bar{J}_{NN} = \{2\}; \quad \bar{J}_P = \{\bar{J}_B, \bar{J}_S\}.$$

On recommence alors une nouvelle itération avec le plan de support $\{\bar{x}, \bar{J}_p\}$:
 Calculons alors le nouveau vecteur des estimations :

$$g(\bar{x}) = D\bar{x} + c = \begin{pmatrix} 2 & 1 & 0 & 0 \\ 1 & 12 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ 3 \\ 3 \end{pmatrix} + \begin{pmatrix} -2 \\ 8 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 9 \\ 0 \\ 0 \end{pmatrix},$$

$$\bar{E}_N = \begin{pmatrix} 0 \\ 9 \end{pmatrix}.$$

Le critère d'optimalité (4.29) étant vérifié, le vecteur $x^0 = (1, 0, 3, 3)$ est alors un plan optimal avec $F(x^0) = -1$.

CONCLUSION GENERALE

Conclusion générale

Ce travail est une synthèse de deux méthodes de la résolution des problèmes de programmation quadratique convexe (la méthode Wolfe et la méthode adaptée) ; pour cela nous avons d'abord rappelé les notions fondamentales sur l'algèbre linéaire, la programmation quadratique et la convexité ; ensuite nous avons présenté les résultats classiques sur l'optimisation non linéaire , sans contrainte où on a fait un petit rappel sur les méthodes de minimisation et avec contraintes , les conditions d'optimalité de Karush-Kuhn-Tucker (KKT), ainsi que les théorèmes de dualité en programmation quadratique convexe .

Bibliographie :

- [01] A.Faradji. Algorithmes de Minimisation d'une fonctionnelle Quadratique. Master's thesis, université de Tizi-ouzou, 1998.
- [02] André Fortin ; Analyse numerque, Ed Hermann, 1996.
- [03] Haim Brezis : Analyse fonctionnelle : théorie et applications. Masson, Paris, 1983.
- [04] Jean-Louis Merrien, ANALYSE NUMERIQUE AVEC MATLAB, édition DUNOD, 2007.
- [05] Jean-Pierre Aubin: Analyse fonctionnelle appliqué, tome1. Presse universitaire de France, 1987.
- [06] Jean-Thierry LAPRESTE, Introduction à MATLAB, édition Ellipses,1999.
- [07] Laurent Schwartz : Analyse : Topologie générale et analyse fonctionnelle. Hermann, 1993.
- [08] M.aidene et B.oukacha ; programmation lineaire, 2005.
- [09] Michel Bierlaire Introduction à l'optimisation différentiable, presse polytechnique et universitaire romandes 2006.
- [10] Michel Minoux Programmation mathématique, théorie et algorithmes 2^{eme} édition, TEC et DOC 2009.
- [11] Moise SYBONY ET Jean-Claude MARDON ANALYSE NUMERIQUE, Systèmes linéaires et non linéaires. HERMANN, EDITEURS DES SCIENCES ET DES ARTS 1988.
- [12] Mokhtar S.Bazarra and C.M Shetty. Nonlinear Pogramming, volume 2. Theory and Algorithmes. John wiley and sons, inc edition, 1979.
- [13] Mustapha LAKRIB Cours d'Analyse Numérique, office des publications universitaire 2008.
- [14] P.wolfe. the simplex method for quadratic programming. Econometrica, 1959

- [15] Pierre Faure Analyse Numérique : note d'optimisation Ecole Polytechnique, Ellipse 1988.
- [16] Philippe G. Ciarlet. Introduction à l'analyse numérique matricielle et à l'optimisation DUNOD 1998.
- [17] Pierre-Jean Laurent : Approximation et optimisation. Hermann, 1972.
- [18] R. Gabassov and F.M. Kirillova. O.I Kostyukova. And V.M Raketskii. Methodes constructive d'optimisation, volume 4: problems convexes. Université de Minsk,1987.
- [19] R. Gabassov and F.M. Kirillova. Méthodes de programmation linéaire, volume 1,2 et 3. Edition de l'université, Minsk, 1977, 1978 et 1980.
- [20] R. Gabassov and F.M. Kirillova. Méthodes d'optimisation. Edition de l'université, Minsk, 1981.
- [21] R. Gabassov and F.M. Kirillova, and Raketskii V.M. On methods for solving the general problem of convex quadratic programming. Soviet Math. Dokl, 1981.
- [22] Xavier Antoine ; yannick Privat ; Introduction à l'optimisation : Aspects théoriques et Algorithmes Ed Eyrolles, 2007.
- [23] Jean-Christophe CULIOLI , Introduction à l'optimisation, ellipses 1994.