REPUBLIQUE ALGERIENNE DEMOCRATIQUE ET POPULAIRE

MINISTERE DE L'ENSEIGNEMENT SUPERIEUR

ET DE LA RECHERCHE SCIENTIFIQUE

UNIVERSITE IBN KHALDOUN - TIARET

FACULTE DES SCIENCES ET SCIENCES DE L'INGENIEUR

DEPARTEMENT DES SCIENCES EXACTES

MEMOIRE

Présenté en vue d'obtenir le diplôme de

MAGISTER

Spécialité : Maths

Option : Analyse fonctionnelle

Thème

Optimisation d'une Fonction Quadratique Sous Contraintes Linéaires

Par:

MAZZOUZ KADDA

Soutenu publiquement le …………….. devant le jury composé de :

Mc. Guedda Lahcene	**Université de Tiaret**	**Président**
Mc. Aïssani Mouloud	**Université de Tiaret**	**Examinateur**
Mc. Larabi Abderrahmane	**Université de Tiaret**	**Examinateur**
Mc. Bellabaci Youcef	**Université de Laghouat**	**Examinateur**
Pr. Mokhtari AEK	**Université de Laghouat**	**Rapporteur**

2010

Dédicaces

À mes parents

À toute ma famille

À tous mes frères et mes amis

À tous les enseignants et les éducateurs qui ont contribué à ma formation durant tout le parcours de mes études jusqu'à ce jour…

K.Mazzouz.

REMERCIEMENTS

Tout d'abord merci au bon Dieu le tout puissant, de m'avoir donné la force, la patience et la volonté afin de réaliser ce travail dans des meilleures circonstances et en bon état.

Je remercie infiniment monsieur Mokhtari Abdelkader professeur à l'université de Ammar Telidji (laghouat), qui m'a fait l'honneur de diriger ce travail. Je tiens à lui exprimer ma profonde reconnaissance, mes vifs remerciements les plus sincères pour son aide, ses conseils précieux, ses sacrifices, malgré ses obligations, ainsi que pour la confiance qu'il m'a prodiguée durant la réalisation de ce travail. J'adresse mes sincères remerciements à monsieur Guedda. Lahcene maître de conférences à l'université Ibn Khaldoun (Tiaret),
pour m'avoir honoré avec sa présence en acceptant de présider le jury de soutenance de ce mémoire.

Mes vifs remerciements vont également aux membres de jury de soutenance composé par :
Monsieur Bellabaci. Youcef maître de conférences à l'université Ammar Telidji (Laghouat). Monsieur Aissani Mouloud maître de conférences à l'université Ibn Khaldoun (Tiaret).
Monsieur Larabi Abderrahmane maître de conférences à l'université Ibn khaldoun (Tiaret).
Je les remercie chaleureusement pour leur présence et pour avoir accepté d'examiner le présent mémoire.

Je remercie infiniment monsieur Snouci Abdelkader maître de conférences et chef de projet.

Enfin, je voudrais associer à mes remerciements tous qui ont contribué de prés ou de loin à l'aboutissement de ce travail.

Table des matières

1 Préliminaires **5**
- 1.1 Notions fondamentales . . . 5
  - 1.1.1 Caractéristiques des ensembles convexes. . . . 6
- 1.2 Ensembles convexes polyédriques . . . 8
  - 1.2.1 Combinaison linéaire convexe . . . 8
- 1.3 Structure des ensembles admissibles . . . 15
  - 1.3.1 Etude systématique. . . . 15
  - 1.3.2 Systèmes homogènes. . . . 20
- 1.4 Fonctions convexes . . . 20
  - 1.4.1 Définitions. . . . 20
- 1.5 L'hyperplan tangent à l'hypersurface. . . . 23
- 1.6 Minimum et maximum d'une fonction convexe. . . . 25

2 Optimum de la fonction ψ **29**
- 2.1 Théorèmes fondamentaux de la programmation quadratique . . . 29
- 2.2 Variations de la fonction objectif lors d'un changement de base . . . 37
  - 2.2.1 Notations. . . . 37
  - 2.2.2 Variations de la fonction objectif lorsqu'on passe d'un sommet à un sommet voisin. . . . 38
  - 2.2.3 Variations de la fonction φ . . . 40
  - 2.2.4 L'accroissement de la fonction objectif. . . . 42
  - 2.2.5 Règle du plus grand gain marginal. . . . 46
  - 2.2.6 L'algorithme. . . . 46

3 Méthodes de programmation non linéaire **51**
- 3.1 Forme générale d'un programme non linéaire . . . 52
- 3.2 Fonction de Lagrange . . . 54
- 3.3 Théorie de Kuhn-Tucker . . . 55
- 3.4 Programmation quadratique . . . 60
- 3.5 La méthode de Franck-Wolfe . . . 69

4 Calcul de l'optimum de la fonction concave **75**
- 4.1 la recherche de la projection du point critique. . . . 79

4.2 la recherche de l'hyperplan séparant le point critique et le convexe fermé Ω. 83
4.3 Extrémum d'une fonction quadratique concave φ 88

5 décomposition d'une fonction quadratique sous sa forme canonique 97
5.1 optimisation de la fonction f. 97
5.2 Algorithme de recherche de l'optimum de f 100
5.3 Décomposition d'une fonction quadratique sous sa forme canonique 103
5.4 CONCLUSION GENERALE 106

Introduction

La notion d'optimisation est la recherche des valeurs de variables qui maximisent ou minimisent une fonction donnée, cet outil mathématique, une fois transposé au monde de l'enterprise, permet d'obtenir un rendement idéal. Cependant l'optimisation reste un sujet vaste de sorte qu'il est impossible de vraiment définir tous les domaines d'application. Ainsi l'optimisation peut représenter une aide pour respecter certaines contraintes du cahier des charges ou pour améliorer un produit (durée de vie, fiabilité, coût de fabrication). Cependant, pour qu'on puisse trouver une solution d'un problème d'optimisation il faut :
1)Exprimer un critère objectif d'optimmalité .
2) choisir les paramètres de conception.
3) définir un espace admissible pour les variables de conception.
De plus, il est important de noter que la qualité de l'optimisation dépend aussi du choix de la formulation de l'objectif et du choix de la formulation des contraintes.

Les domaines d'applications de l'optimisation sont nombreux citons commme exemples les problèmes économiques (production, transport, stockage des produits,...), problèmes du voyageurs, problèmes d'équilibre chimique,....

Le sujet de ce mèmoire a été consacré à l'optimisation des fonctions quadratiques quelconques sous contraintes linéaires, on s'intéresse à l'illustration de quelques méthodes utilisées pour résoudre les problèmes quadratiques, et la démonstration et l'éxécution de ces méthodes.Le contenu de ce mémoire est réparti en cinq chapitres :

Au premièr chapitre nous avons rapplé brièvement les définitions de quelques notions importantes, ainsi que quelques propriétés,comme la notion de conve- xité qui permet d'obtenir des résultats d'existence et d'unicité de solutions du problèmes d'optimisation à la fois dans le cas où ces *problèmes conduisent à la résolution du système linéaire et dans le cas de certains problèmes non linéaires,et nous envisageons particulièrement les théorèmes les plus utiles qui synthétisent les propriétés de minimisation(où maximisation) d'une fonction convexe (où concave).*

Dans le deuxième chapitre on s'intéresse à la résolution d'un problème quadratique convexe , on étudiera les variations de la fonction objectif lors du passage d'un sommet à un autre, les théorèmes fondamenteaux de la programation quadratique et l'algorithme de résolution .

Nous avons consacré le troisième chapitre au programmation non linéaire (sans que le mot de programmation ne fasse en quelque manière que ce soit référence à l'aspect informatique),en particulier la programmation quadratique et la théorie de Khun-tucker,en abordant les résultats et les conditions nécessaires d'optimalité de Khun-tucker,on termine ce chapitre par la méthode de Franck-wolfe, nous allons voir un type particulier d'algorithme qui permet de résoudre les problèmes d'optimisation non linéaire convexe.

Ma contribution est les deux dernièrs chapitres (4 et 5) ;
Le chapitre quatre est consacré au calcul de l'optimum d'une fonction concave ,en illustrant deux méthodes totalement différentes dont la première est plus éfficace et plus performante puisque elle donne une solution exacte du problème traité, cependant la deuxième est une méthode itérative qui est basée sur la construction des suites reccurentes qui convergent vers l'optimum qui est une solution approchée.

En fin,le dernièr chapitre est l'étude d'une fonction quadratique quelconque f qui sera décomposée en deux fonctions l'une est concave et l'autre est convexe.

Chapitre 1

Préliminaires

1.1 Notions fondamentales

On va formuler et démontrer des résultats importants relatifs aux ensembles convexes.

Définition 1.1 *Un ensemble est convexe si, chaque fois qu'il contient deux points x_1, x_2 quelconques, il contient également tous les points de la forme :*

$$x = \lambda x_1 + (1 - \lambda)x_2, \quad 0 \leqslant \lambda \leqslant 1. \tag{1.1}$$

La droite passant par les points x_1, x_2 a pour équation paramétrique

$$x(\lambda) = x_2 + (x_1 - x_2)\lambda \tag{1.2}$$

$(x_1 - x_2)$ étant son vecteur directeur. Lorsque $\lambda = 0$, on a $x(0) = x_2$,
et quand $\lambda = 1$, on a $x(1) = x_1$.
Quand λ varie entre 0 et 1, le point $x(\lambda)$ parcourt tout le segment $[x_1, x_2]$

Définition 1.2 *Un ensemble est convexe si, chaque fois qu'il contient deux points x_1, x_2 quelconques, il contient également le segment dont ils sont les extrémités.*

Théorème 1.1 *L'intersection de deux ensembles convexes est convexe.*

Preuve : Soit $X = X_1 \cap X_2$ avec X_1, X_2 des ensembles convexes, et soient x, x' deux points quelconques de X. Comme $X \subseteq X_1$, on a $x, x' \in X_1$. L'ensemble X_1 étant convexe, il en résulte que le segment $[x, x']$ tout entier est dans X_1.
De même $[x, x'] \subseteq X_2$. Ainsi $[x, x'] \subseteq X_1 \cap X_2 = X$.

Définition 1.3 *On appelle somme de deux ensembles convexes X_1, X_2 de $\mathbb{R}^n$ l'ensemble*
$X = \{x, x = x_1 + x_2,\ x_1 \in X_1,\ x_2 \in X_2\}$*, obtenu par addition deux à deux les éléments de X_1 et X_2.*

Théorème 1.2 *La somme de deux ensembles convexes est convexe.*

Preuve : Soient $x, x' \in X$, où $X = X_1 + X_2$. Il existe dans X_1 et X_2 des éléments tels que :

$$x = x_1 + x_2, \quad x' = x'_1 + x'_2, \quad x_1, x'_1 \in X_1,\ x_2, x'_2 \in X_2.$$

Etant donné in réel tel que $0 \leqslant \lambda \leqslant 1$, on a

$$\lambda x_1 + (1-\lambda)x'_1 \in X_1,\ \lambda x_2 + (1-\lambda)x'_2 \in X_2.$$

De la définition de X, il résulte que la somme de ces éléments est dans X :

$$\lambda x_1 + (1-\lambda)x'_1 + \lambda x_2 + (1-\lambda)x'_2 = \lambda(x_1 + x_2) + (1-\lambda)(x'_1 + x'_2) = \lambda x + (1-\lambda)x' \in X.$$

1.1.1 Caractéristiques des ensembles convexes.

Théorème 1.3 *Soit H un espace de Hilbert, $A \subset H$ un ensemble non vide convexe et fermé, $x_0 \in H \setminus A$. Alors il existe un et un seul élément $y_0 \in A$ tel que :* $\|x_0 - y_0\| = \inf\limits_{y \in A} \|x_0 - y\|$.

Preuve : $i)$ Existence ; Posons : $d = \inf\limits_{y \in A} \|y - x\| > 0$.
Soit (y_n) une suite de A telle que $\|x_0 - y_n\| \leqslant d + \frac{1}{n}$.

Montrons que (y_n) est une suite de Cauchy

$$\begin{array}{rcl} \|y_n - y_m\|^2 & = & \|(y_n - x_0) + (x_0 - y_m)\|^2 \\ & = & \|(y_n - x_0) - (y_m - x_0)\|^2 \end{array}$$

D'après l'identité du parallélogramme, on a :

$$\begin{array}{rcl} \|y_n - y_m\|^2 & = & 2\big(\|y_n - x_0\|^2 + \|y_m - x_0\|^2\big) - \|(y_n - x_0) + (y_m - x_0)\|^2 \\ & = & 2\|y_n - x_0\|^2 + 2\|y_m - x_0\|^2 - 4\|\dfrac{y_n + y_m}{2} - x_0\|^2 \end{array}$$

Mais $\dfrac{y_n + y_m}{2} \in A$ $\big($ car A est convexe $\big)$, donc $\|\dfrac{y_n + y_m}{2} - x_0\|^2 \geqslant d^2$ d'où

$\|y_n - y_m\|^2 \leqslant 2(d + \dfrac{1}{m})^2 + 2(d + \dfrac{1}{n})^2 - 4d^2 = \dfrac{4}{m}d + \dfrac{2}{m^2} + \dfrac{4}{n}d + \dfrac{2}{n^2} < \varepsilon$ si $n, m > n_0(\varepsilon)$

Donc (y_n) est de Cauchy, comme H est complet, alors $y_n \to y_0 \in A$ car A est fermé.

ii) Montrons que $\|x_0 - y_0\| = \inf\limits_{y \in A} \|x_0 - y\| = d$

$d \leqslant \|x_0 - y_0\| = \|x_0 - y_n + y_n - y_0\| \leqslant \|x_0 - y_n\| + \|y_n - y_0\| < d + \frac{1}{n} + \varepsilon$ car $y_n \to y_0 \in A$.

$\Rightarrow \|x_0 - y_0\| < d + \frac{1}{n} + \varepsilon, \forall \varepsilon > 0$ d'où $\|x_0 - y_0\| = d$.

iii)Prouvons l'unicité de y_0.

Supposons qu'il existe $z_0 \in A\,;\, d = \|x_0 - z_0\|$.

$$\begin{array}{rcl} \|z_0 - y_0\|^2 & = & \|(z_0 - x_0) - (y_0 - x_0)\|^2 \\ & = & 2\|(z_0 - x_0)\|^2 + 2\|(y_0 - x_0)\|^2 - \|z_0 + y_0 - 2x_0\|^2 \\ & = & 4d^2 - 4\|\dfrac{z_0 - y_0}{2} - x_0\|^2 \leqslant 4d^2 - 4d^2 \\ & = & 0. \end{array}$$

Ce qui implique que $z_0 = y_0$ d'où l'unicité de y_0.

Théorème 1.4 *Soit f une fonction convexe définie sur un convexe Ω de $\mathrm{I\!R}^n$.*
$f : \Omega \to \mathrm{I\!R}$; $\overline{x}$ son minimum local : $f(\overline{x}) = \inf\limits_{x \in B(\overline{x}, r)} f(x)$. Alors :
$\overline{x}$ est un minimum global. C'est-à-dire $f(\overline{x}) = \inf\limits_{x \in \Omega} f(x)$

Preuve : En effet, supposons qu'il existe ;
$y \in \Omega$ tel que $f(y) < f(\overline{x})$ pour tout $y \notin B(\overline{x}, r)$.
$x = \lambda y + (1-\lambda)\overline{x}$ où $\lambda = \dfrac{r}{2\|y-\overline{x}\|} < 1$ car $y \notin B(\overline{x}, r)$
$x \in \Omega$ et Ω est covexe.
$f(x) \leqslant (1-\lambda)f(\overline{x}) + \lambda f(y) = f(\overline{x}) + \lambda[f(y) - f(\overline{x})] < f(\overline{x})$.
Mais $x \in B(\overline{x}, r)$ ($\|x - \overline{x}\| = \frac{r}{2} < r$) contraduction.

1.2 Ensembles convexes polyédriques

1.2.1 Combinaison linéaire convexe

Dans ce paragraphe, on se propose d'analyser une classe spéciale d'ensembles convexes qui correspond géométriquement aux polyédres, voir [6].

Définition 1.4 *Soient $x_1, x_2, \ldots, x_k$ des points de $\mathbb{R}^n$. On appelle combinaison linéaire convexe de ces points la somme $\alpha_1 x_1 + \alpha_2 x_2 + \ldots + \alpha_k x_k$, avec :*
$\alpha_i \geqslant 0,\ i = 1, 2 \ldots, k$ *et* $\sum_{i=1}^{k} \alpha_i = 1$

Conformément à la définition 1.1, un ensemble convexe X est dit convexe s'il contient toute combinaison linéaire convexe des ses points.

Théorème 1.5 *Soient X un ensemble convexe et $x_1, x_2, \ldots, x_k$ des éléments de Ω.*
Alors l'ensemble Ω contient toute combinaison linéaire convexe de ces points.

Preuve. On démontre par récurrence sur le nombre k de points.
Si $k = 2$, l'affirmation du théorème coïncide avec la définition d'un ensemble convexe.

Supposons démontrer que toute combinaison linéaire convexe de $(k-1)$ points de Ω appartient à Ω.

Soient k points $x_1, \ldots, x_k$ et la combinaison linéaire convexe $x = \sum_{i=1}^{k} \alpha_i x_i$.
Si $\alpha_k = 1$, alors $\alpha_i = 0$ pour $1 \leqslant i \leqslant k-1$ et $x = x_k \in \Omega$.
Admettons que $\alpha_k < 1$, auquel $1-\alpha_k = \sum_{i=1}^{k-1} \alpha_i > 0$. On regroupe les termes dans l'expression de x :

$$x = (1-\alpha_k) \sum_{i=1}^{k-1} \frac{\alpha_i}{1-\alpha_k} x_i + \alpha_k x_k.$$

Les nombres $\frac{\alpha_i}{1-\alpha_k}$, $i = 1, \ldots, k-1$, sont non négatifs et leur somme vaut 1 :

$$\sum_{i=1}^{k-1} \frac{\alpha_i}{1-\alpha_k} = \frac{1}{1-\alpha_k} \sum_{i=1}^{k-1} \alpha_i = \frac{1}{1-\alpha_k}(1-\alpha_k) = 1$$

L'expression $x' = \sum_{i=1}^{k-1} \frac{\alpha_i}{1-\alpha_k} x_i$ est donc une combinaison linéaire convexe des points $x_1, \ldots, x_{k-1}$ de Ω. Par hypothèse de récurrence, $x' \in \Omega$, auquel cas le point
$x = (1-\alpha_k)x' + \alpha_k x_k$ est combinaison linéaire convexe de deux points de Ω, donc $x \in X$.

Définition 1.5 *Soit M un ensemble de points (fini ou infini) de IR^n. On appelle envloppe convexe de M, et on note $C(M)$, l'ensemble de toutes les combinaisons linéaires convexes des points de M.*

Théorème 1.6 *L'envloppe convexe $C(M)$ de l'ensemble M coïncide avec le plus petit ensemble convexe contenant M.*

Preuve. La démonstration est évidente car,d'une part, $C(M)$ est un ensemble convexe, et d'autre part Ω étant un convexe contenant M contient toutes les combinaisons linéaires convexes des points de M, i.e $\Omega \supseteq C(M)$.

Définition 1.6 *un point x d'un ensemble convexe Ω est dit extrême s'il n'admet pas la représentation $x = \frac{1}{2}x_1 + \frac{1}{2}x_2$ où $x_1 \neq x_2, x_1, x_2 \in \Omega$*

Remarque 1.1 *Tout ensemble convexe ne possède pas de points extrêmes.*

Exemple 1.1 *Le demi-plan supérieur fermé est un convexe, a des points frontières (que sont tous les points de la droite définissant le demi-plan), mais il ne possède pas de points extrèmes.*

Théorème 1.7 *Tout ensemble compact convexe de* $\mathbb{R}^n$ *est l'envloppe convexe de ses points extrêmes.voir [1]*

Remarque 1.2 *Cela signifie que tout point* $x \in \Omega$, *(*Ω *étant un compact convexe), s'écrit sous forme de combinaison linéaire convexe d'un nombre fini de points extrêmes de* Ω.

On sait qu'un compact convexe possède des points extêrmes.

De même, tout point x d'un compact convexe Ω est combinaison linéaire convexe de certains point frontières de Ω

Définition 1.7 *Un ensemble* Ω *s'appelle polyèdre convexe s'il est l'envloppe convexe des points en nombre fini.*

Autrement dit, un ensemble Ω *est un polyèdre convexe s'il contient un nombre fini de points* $x_1, \ldots, x_k$ *tel que tout élément* x *de* Ω *s'écrit*

$$x = \sum_{i=1}^{k} \alpha_i x_i, \ 0 \leqslant \alpha_i \leqslant 1, \ \sum_{i=1}^{k} = 1,$$

est que tous les points de cette forme soient dans Ω.

Théorème 1.8 *Un polyèdre convexe est compact, voir [6].*

Preuve. Soit X un polyèdre convexe et Ω l'enveloppe convexe de ses points $x_1, \ldots, x_k : \Omega = C(x_1, \ldots, x_k)$.

On démontre la compacité de Ω sous la condition suffisante qu'on extrait d'une sous suite $x(m) \in \Omega$, $m = 1, 2, \ldots$, une suite partielle convergente

vers un $\overline{x} \in \Omega$.
On a $x(m) \in \Omega$, $m = 1, 2, \ldots$, si bien que

$$x(m) = \sum_{i=1}^{k} \alpha_i(m)x_i, \ 0 \leqslant \alpha_i \leqslant 1, \quad \sum_{i=1}^{k} \alpha_i(m) = 1$$

Chaque suite $\alpha_i(m)$ $m = 1, 2, \ldots$, est bornée, on en choisit donc une sous-suite convergente.

Les suites $\{\alpha_i(m)\}$ étant en nombre fini $m = 1, 2, \ldots, k$, on peut former une suite partielle $m_1, m_2, \ldots$ telle que toutes les suites $\alpha_i(m_j), j = 1, 2, \ldots$, convergent.
Soit $\lim\limits_{j\to\infty} \alpha_i(m_j) = \alpha_i \geqslant 0$.
Evidemment $\sum\limits_{i=1}^{k} \alpha_i = 1$, auquel $\sum\limits_{i=1}^{k} \alpha_i(m_j)x_i$, $j = 1, 2, \ldots$, possède une limite et ;
$\overline{x} = \lim\limits_{j\to\infty} \sum\limits_{i=1}^{k} \alpha_i(m_j)x_i = \sum\limits_{i=1}^{k} \alpha_i x_i \in \Omega$.
D'où le théorème est démontré.

Remarque 1.3 *Les ensembles convexes sont caractérisés par la propriété de séparation*

Exemple 1.2 *Soit dans le plan un convexe fermé Ω et un point $d \notin \Omega$ alors il existe une droite $ax_1 + bx_2 = c$ tels que Ω et $d = (d_1, d_2)$ soient de par et d'autre de la droite i.e tout point $(x_1, x_2) \in \Omega$ vérifié $ax_1 + bx_2 \leqslant c$ mais $ad_1 + bd_2 > c$.*

Définition 1.8 *Un ensemble de points $x = (x_1, \ldots, x_n)$ tel que :*

$$a_1x_1 + \ldots + a_nx_n = c \text{ avex } a \in \mathbb{R}^n, \ a \neq 0, \ c \in \mathbb{R}.$$

S'appelle l'hyperplan de $\mathbb{R}^n$ de l'équation :

$$\langle a, x\rangle = c$$

le vecteur a $(a \neq 0)$ est la normale à cet hyperplan.

Théorème 1.9 *(de séparation). Soit Ω un ensemble fermé borné, $\alpha \in \Omega$, $\alpha = (\alpha_1, \ldots, \alpha_n)$. Il existe un hyperplan de normale a ($a \neq 0$) tel que $\langle a, \alpha \rangle > c$ et on ait $\langle a, x \rangle \leqslant c$ pour tout $x \in \Omega$.*

Preuve : Soit la fonction $\varphi(x) = (x_1 - \alpha_1)^2 + \ldots + (x_n - \alpha_n)^2$ de la variable indépendante $x = (x_1, \ldots, x_n)$ sur l'ensemble Ω. Elle est évidemment continue. Aux terme d'un théorème de Weierstrass, une fonction continue sur un ensemble fermé borné atteint sur cet ensemble sa borne inférieure. Pour qu'on puisse utiliser ce résultat, on considère un point $\overline{x}$ quelconque de Ω et une boule B de centre a et dont le carré de rayon vaut $\varphi(\overline{x})$.
L'intersection $\widetilde{\Omega} = B \cap \Omega$ est un ensemble fermé et borné. La fonction $\varphi(x)$ atteint donc son minimum en un point $b \in \widetilde{\Omega}$. Il est clair que b réalise le minimum de $\varphi(x)$ sur Ω tout entier.
Ainsi, $\varphi(x) \geqslant \varphi(b)$ quel que soit $x \in X$. On fixe un point quelconque $x \in \Omega$ et on pose : $x(\lambda) = \lambda x + (1 - \lambda) b, \quad 0 \leqslant \lambda \leqslant 1$.
L'ensemble Ω étant convexe, on a $x(\lambda) \in \Omega$ pour tout λ, $0 \leqslant \lambda \leqslant 1$,
donc $\quad \varphi(x(\lambda)) \geqslant \varphi(b)$.
Par conséquent, $\quad \dfrac{\varphi(x(\lambda)) - \varphi(b)}{\lambda} \geqslant 0$ pour tout $0 < \lambda \leqslant 1$.
La fonction $\varphi(x)$ est dérivable par rapport à λ en tant que la composée de deux fonctions dérivables $\varphi(x)$ et $x(\lambda)$ d'autre part la limite
$\lim\limits_{\lambda \to 0} \dfrac{\varphi(x(\lambda)) - \varphi(b)}{\lambda} \geqslant 0 = \varphi'_\lambda(x(0)) \geqslant 0$ existe.
On calcule directement la dérivée $\varphi'_\lambda(x(0))$ à l'aide de la règle de dérivation d'une fonction composée :

$$\varphi'_\lambda(x(\lambda)) = \left(\sum_{i=1}^{n} (x_i(\lambda) - \beta_i)^2 \right)' = 2 \sum_{i=1}^{n} (x_i(\lambda) - \alpha_i) x'_i(\lambda)$$

Puique $\quad x'_i(\lambda) = (\lambda x_i + (1 - \lambda)\lambda_i)' = x_i - \beta$ et $x_i(0) = \beta_i \quad$ on a

$$\varphi'_\lambda(x(\lambda)) = 2 \sum_{i=1}^{n} (x_i(\lambda) - \alpha_i)(x_i - \beta_i),$$

$$\varphi'_\lambda(x(0)) = 2 \sum_{i=1}^{n} (\beta_i - \alpha_i)(x_i - \beta_i),$$

D'après les résultats démontrés ci-dessus, cette dérivée est non négative, c'est-à-dire

$$\sum_{i=1}^{n}(b_i - \alpha_i)(x_i - b_i) \geqslant 0.$$

On introduit les notations $a_i = \alpha_i - \beta_i$ $i = 1, \ldots, n$, $a = (a_1, \ldots, a_n)$, $c = \sum_{i=1}^{n} \beta_i(\alpha_i - \beta_i)$ et on reécrit l'inégalité obtenue :

$$\sum_{i=1}^{n} a_i x_i \leqslant c, \quad \big(\langle a, x\rangle \leqslant c\big).$$

On note que le vecteur $a = (a_1, \ldots, a_n)$ et le nombre c sont indépendants du choix du vecteur x de Ω. Aussi, il faut vérifier que $a \neq 0$ et $\langle a, \alpha\rangle > c$ pour constater que toutes les affirmations du théorème sont justes pour a et c.
En effet, $a = \alpha - \beta$ et $a \neq 0$ car dans le cas contraire $\alpha = \beta$, ce qui est impossible puisque $\beta \in \Omega$, $\alpha \notin \Omega$.
Enfin,

$$\begin{aligned}\langle a, \alpha\rangle - c &= \sum_{i=1}^{n}(\alpha_i - \beta_i)\alpha_i - \sum_{i=1}^{n}(\alpha_i - \beta_i)\beta_i \\ &= \sum_{i=1}^{n}(\alpha_i - \beta_i)^2 > 0\end{aligned}$$

donc $\langle a, \alpha\rangle > c$, ce qui démontre le théorème.

Définition 1.9 *On dit que l'hyperplan $\langle a, x\rangle = c$ est un hyperplan d'appui à l'ensemble Ω en un point $b \in \Omega$ si l'on a pour tout élément $x \in \Omega$ les relations $\langle a, x\rangle \leqslant c$ et $\langle a, b\rangle = c$.*

Remarque 1.4 *Le théorème de séparation dit que tout ensemble convexe fermé X est séparé de tout point $a \in X$ par un hyperplan d'appui.*

Théorème 1.10 *La fermeture d'un ensemble convexe est convexe.*

Preuve : En effet, soient $x, x' \in \overline{\Omega}$, tel que $x = \lim\limits_{k\to\infty} x_n$, $x_n \in \Omega$, $x' = \lim\limits_{k\to\infty} x'_n$, $x'_n \in \Omega$.
Vue la convexité de Ω, on a :

$$[\lambda x_n + (1-\lambda)x'_n] \in \Omega \quad \text{et} \quad \lim_{k\to\infty}(\lambda x_n + (1-\lambda)x'_n) = \lambda x + (1-\lambda)x'$$

qui appartient à $\overline{\Omega}$ du moment que $\overline{\Omega}$ est un fermé.
Soit $\alpha \notin \Omega$, on démontre aisément qu'il existe une suite $(\alpha_n)_{n\in \mathbb{N}}$ de Ω $(\alpha_n) \in \overline{\Omega}$, $\lim\limits_{k\to\infty} \alpha_n = \alpha$. On applique le théorème de séparation à $\overline{\Omega}$ et à α_n, on obtient l'hyperplan $\langle a_n, x\rangle = \langle a_n, \beta_n\rangle$.
Prenons l'hyperplan parallèle passant par α_n : $\langle a_n, x\rangle = \langle a_n, \alpha_n\rangle$.
On estime sans restreindre la généralité que $\|\alpha_n\| = 1$, $n = 1, 2, \ldots$, et que par la suite des vecteurs a_n converge vers un a $(a \neq 0)$. On a pour tout n et tout $x \in \Omega$ l'inégalité :

$$\langle a_n, x\rangle \leqslant \langle a_n, \beta_n\rangle \text{ car } \langle a_n, \beta_n\rangle \leqslant \langle a_n, \alpha_n\rangle$$

.
On passe à la limite pour $n \to \infty$, il vient $\langle a, x\rangle \leqslant \langle a, \alpha\rangle$ pour tout $x \in \overline{\Omega}$, en particulier, pour $x \in \Omega$.
On a donc le théorème de séparation pour les ensembles convexes non fermés.

Théorème 1.11 *Soit Ω un ensemble convexe (non nécessairement fermé). On suppose que α n'est pas un point intérieur de X. Alors il existe un hyperplan*
d'equation : $\langle a, x\rangle = c$, $(a \neq 0)$, tel que :
$\langle a, \alpha\rangle = c$ et $\langle a, x\rangle \leqslant c$ pour tout $x \in \Omega$. Voir[1]

Définition 1.10 *On dit que l'ensemble $\mathbb{R}^n_+ = \{x/x \in \mathbb{R}^n, x \geqslant 0\}$ des vecteurs*
$x = (x_1, \ldots, x_n)$ de composantes non négatives $x_i \geqslant 0$, $i = 1, \ldots, n$, est l'orthant positif de l'espace $\mathbb{R}^n$.

Théorème 1.12 *Soit Ω un ensemble convexe qui ne contient pas de points intérieurs de l'orthant positif $\mathbb{R}^n_+$ (i.e aucun point de X n'a toutes ses composantes positives). Il existe un hyperplan $\langle a, x\rangle = 0$ de normale $a > 0$ tel que $\langle a, x\rangle \leqslant 0$ pour tout $x \in \Omega$.*

Preuve : Soit l'esemble $M = \Omega - \mathbb{R}^n_+$. Il est clair que 0 n'est pas un point intérieur à M.
En effet, il existe dans le cas contraire un voisinage (une boule) d'origine des coordonnées, qui appartient à M. Ce voisinage contient un point u dont toutes les coordonnées sont positives : $(u > 0)$.
Comme $u \in M$ on a $u = x - v$, où $x \in \Omega$ et $v \in \mathbb{R}^n$, auquel $x = u + v > 0$ contradiction avec l'hypothèse du théorème.
Mais il existe un hyperplan $\langle a, x\rangle = 0$, $a \neq 0$ passant par l'origine tel que tout point $u \in M$ vérifie l'inégalité $\langle a, u\rangle \leqslant 0$.
On en conclut moyennant la définition de M que tout $x \in \Omega$ et $v \in \mathbb{R}^n_+$, satisfont à $\langle a, x\rangle \leqslant \langle a, v\rangle$.
On pose $a = 0$, il vient $\langle a, x\rangle \leqslant 0$ quel que soit $x \in \Omega$.
On fixe un point $x \in \Omega$ et on a $\langle a, v\rangle \leqslant \langle a, x\rangle$ pour tout $v \in \mathbb{R}^n_+$.
D'où évidemment $a \geqslant 0$.
En effet, si $a_i < 0$ et si l'on choisit pour v un point de la forme $(0, 0, \ldots, 0, \lambda, 0, \ldots, 0)$, avec $\lambda > 0$ à la place i, on obtient $\lambda a_i \geqslant \langle a, x\rangle$. Or on peut rendre λa_i négatif aussi petit que l'on veut, inférieur en tous cas à $\langle a, x\rangle$ fixé.

1.3 Structure des ensembles admissibles

1.3.1 Etude systématique.

On se propose d'étudier d'une manière systématique la structure des ensembles admissibles des problèmes de programmation linéaire, voir [6] :
Un ensemble admissible est décrit par le système d'inéquations linéaires de la forme :

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \leqslant b_1, \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n \leqslant b_2, \\ \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n \leqslant b_m, \end{cases} \tag{1.3}$$

On n'écrit pas à part les contraintes $x_i \geqslant 0$ car ces contraintes, si elles existent, sont incluses dans le système (1.3) sous forme $-x_i \leqslant 0$.
Le symbole a_i désignera dans la suite le vecteur $a_i = (a_{i1}, \ldots, a_{in})$ (la ligne i de la matrice A des coeficients de (1.3)).
a_j désignera le vecteur $a_j = (a_{j1}, \ldots, a_{jm})$ (la colonne j de la matrice A).

Le système d'inéquations (1.3) prend une forme plus compacte, à savoir

$$\langle a_1, x\rangle \leqslant b_1, \quad \langle a_2, x\rangle \leqslant b_2, \ldots, \langle a_m, x\rangle \leqslant b_m, \tag{1.4}$$

x étant comme toujours $x = (x_1, x_2, \ldots, x_n)$.
Si l'on fait recours à la notation $A = (a_{ij}),\ i = 1, 2, \ldots, m\ j = 1, \ldots, n$.
Notre système réecrit

$$Ax \leqslant b,\ \text{où } b = (b_1, b_2, \ldots, b_m). \tag{1.5}$$

Et on désignera par la suite l'ensemble des vecteurs réalisables (c'est-à-dire l'ensemble des vecteurs x vérifiant (1.3)) : par $\Omega = \{x \in \mathbb{R}^n : Ax \leqslant b\}$.

Remarque 1.5 *Il se peut que Ω soit vide.*

Exemple 1.3 *L'ensemble $\mathbb{R}$ défini par les contraintes $x \leqslant 1,\ -x \leqslant -2$.*

Théorème 1.13 *Tout ensemble Ω des vecteurs réalisables est convexe et fermé.*

Preuve : On montre que le demi-espace Ω_i défini par $\langle a_i, x\rangle \leqslant b_i$ est convexe.
En effet, soit $x, x' \in X_i$, on a $\langle a_i, x\rangle \leqslant b_i, \langle a_i, x'\rangle \leqslant b_i$.
Soit α un nombre quelconque tel que $0 \leqslant \alpha \leqslant 1$, auquel

$$\lambda\langle a_i, x\rangle b_i, \quad (1-\lambda)\langle a_i, x'\rangle \leqslant (1-\lambda)b_i.$$

On fait la somme et on utilise la linéarité du produit scalaire, il vient

$$\langle a_i, \lambda x + (1-\lambda)x'\rangle \leqslant \lambda b_i + (1-\lambda)b_i = b_i,$$

d'où la convexité de Ω.
On constate que l'ensemble Ω est l'intersection de demi-espaces Ω_i, i.e. $x \in \Omega$ si et seulement si x est élément de tous les Ω_i, $i = 1, 2, \ldots, m$.
Il ne reste qu' appliquer le le théorème 1.1 (l'intersection de deux ensembles convexes est convexes).
On peut utiliser l'écriture matricielle (1.5) du système d'inéquations donnant Ω pour démontrer le théorème 1.13.
Faisons-le, car dans la suite nous aurons souvent affaire aux matrices.

Si $x, x' \in \Omega$, alors $Ax \leqslant b$, $Ax' \leqslant b$, et par suite on obtient :
$\lambda Ax \leqslant \lambda b$, et $(1-\lambda)Ax' \leqslant (1-\lambda)b$ [on se sert de la non négativité des nombres λ et $(1-\lambda)$]. D'où $\lambda Ax + (1-\lambda)Ax' \leqslant b$.
L'opérateur de multiplication d'une matrice par un vecteur étant linéaire, d'où :

$$A(\lambda x + (1-\lambda)x') \leqslant b, \text{ i.e } \quad \lambda x + (1-\lambda)x' \in \Omega.$$

Démontrons la deuxième propriété (la ferméture de Ω).
Soit $(x_n)_{n\in \mathbb{N}}$, une suite convergente de Ω telle que, $\lim\limits_{n\to\infty} x_n = \overline{x}$.

On passe à la limite dans $Ax_n \leqslant b$ on obtient $A\overline{x} \leqslant b$.

Les points extrêmes de Ω sont des objets particulièrement intéressants. On établit les conditions qui garantissent l'existence des points extrêmes.

Définition 1.11 *Un point $x_0 \in \Omega$ est dit intérieur s'il vérifie strictement toutes les inéquations (1.3) (donc (1.4), (1.5)) : $\langle a_i, x_0\rangle < b_i$ pour $i = 1, 2, \ldots, m$. Tous les autres points de Ω seront qualifiés de points frontières.*

Remarque 1.6 *Cette définition est parfaitement cohérente avec la notion usuelle d'un point intérieur.*

En effet, si l'on utilise la continuité des fonctions $\langle a_i, x\rangle$ et le fait qu'elles sont en nombre fini dans notre cas ($i = 1, 2, \ldots, m$), on trouve facilement un nombre ε tel que si la longueur du vecteur $\delta = (\delta_1, \delta_2, \ldots, \delta_n)$ est plus petite que ε, alors on ait les inégalités $\langle a_i, x_0 + \delta\rangle \leqslant b_i$, $i = 1, 2, \ldots, m$ c'est-à-dire le point $(x_0 + \delta)$ appatient lui aussi à Ω.

En d'autres termes, tant que X contient un point intérieur x_0, il contient un $\varepsilon-$voisinage de ce point.

Le point intérieur x_0 de X ne peut évidemment être extrême.
En effet, il s'écrit $x_0 = \frac{1}{2}(x_0 + \delta) + \frac{1}{2}(x_0 - \delta)$ où : $x_0 + \delta$, $x_0 - \delta \in \Omega$: contradiction avec la définition (1.6) d'un point extrême.

Il est donc clair qu'on trouve les extrêmes parmi les points frontières, c'est-à-dire parmi les points pour lesquels au moins une condition (1.3) devient égalité.

Définition 1.12 *Le support d'un point frontière x_0 est l'ensemble des contraintes (1.3) vérifiées par x_0 sous forme d'égalités. Autrement dit, le support de x_0 est l'ensemble des hyperplans $\langle a_i, x\rangle = b_i$ qui définissent Ω et qui contiennent x_0.*

Soient $\sigma = (i_1, i_2, \ldots, i_k)$, $1 \leqslant i_j \leqslant m$, $j = 1, 2, \ldots, n$, les numéros des vecteurs-lignes a_i qui définissent le support de x_0 et $\overline{\sigma}$ l'ensemble des numéros $i \in \{1, 2, \ldots, m\}$ restants.

On note A_σ la matrice de vecteurs-lignes a_i, $i \in \sigma$ et b_σ le vecteur formé de coordonnées de b indicés par les éléménts de σ.

On introduit de même les notations $A_{\overline{\sigma}}$ et $b_{\overline{\sigma}}$.
Si on prend comme exemple :

$$A = \begin{bmatrix} 3 & 5 \\ -1 & 1 \\ 10 & 7 \\ -1 & -1 \\ -1 & 0 \\ 0 & -1 \end{bmatrix} \qquad b = \begin{bmatrix} 15 \\ 2 \\ 35 \\ -1 \\ 0 \\ 0 \end{bmatrix}, \qquad x = \left(\frac{70}{29}, \frac{45}{29}\right).$$

alors $\sigma = (1, 3)$, $\overline{\sigma} = (2, 4, 5, 6)$, $A_\sigma = \begin{pmatrix} 3 & 5 \\ 10 & 7 \end{pmatrix}$, $b_\sigma = \begin{pmatrix} 15 \\ 35 \end{pmatrix}$,

$$A_{\overline{\sigma}} = \begin{bmatrix} -1 & 1 \\ -1 & -1 \\ -1 & 0 \\ 0 & -1 \end{bmatrix} \qquad b_{\overline{\sigma}} = \begin{bmatrix} 2 \\ -1 \\ 0 \\ 0 \end{bmatrix}.$$

Les contraintes (1.3) (où (1.5)) deviennent pour x_0 :
$A_\sigma x_0 = b_\sigma, \quad A_{\overline{\sigma}} x_0 < b_{\overline{\sigma}}$.
On rappelle que le rang d'une matrice est le nombre maximum des lignes (ou colonnes) linéairement indépendantes.

Théorème 1.14 *Un point $x_0 \in \Omega$ est un point extrême si et seulement si le rang de A_σ est n.*

Preuve : L'équation $A_\sigma x_0 = 0$ n'a pas d'autre solution à part 0 si et seulement si le rang de A_σ est égal au nombre n de variables.

On montre en premier lieu que si x_0 est un point extrème, alors le rang de A_σ vaut n.

Supposons que ce n'est pas le cas. Le rang de A est alors inférieur à n, si bien que $A_\sigma x = 0$ a pour solution $\overline{x} \neq 0$. Comme $A_\sigma x \leqslant b_{\overline{\sigma}}$, on trouve $\alpha \neq 0$ suffisamment petit tel que :

$$A_{\overline{\sigma}}(x_0 + \alpha\overline{x}) < b_{\overline{\sigma}}\ ,\ \ A_{\overline{\sigma}}(x_0 - \alpha\overline{x}) < b_{\overline{\sigma}}$$

et

$$A_\sigma(x_0 + \alpha\bar{x}) = A_\sigma x_0 + \alpha A_\sigma \overline{x} = A_\sigma x_0 = b_\sigma.$$

De même $A_\sigma(x_0 - \alpha\overline{x}) = A_\sigma x_0 - A_\sigma \alpha \overline{x} = A_\sigma x_0 = b_\sigma$.
Il en résulte que $(x_0 + \alpha\overline{x}),\ (x_0 - \alpha\overline{x}) \in X$, ce qui contredit l'hypothèse x_0 est un point extrême $\left(\text{en effet, } x_0 = \frac{1}{2}(x_0 + \alpha\overline{x}) + \frac{1}{2}(x_0 - \alpha\overline{x}\right)$.
Inversement, soit n le rang de A_σ.
On suppose que x_0 n'est pas un point extrême de X.
Il existe deux points $x', x'' \in \Omega,\ x' \neq x''$, tel que $x_0 = \frac{1}{2}x' + \frac{1}{2}x''$, auquel on a :

$$b_\sigma = A_\sigma x_0 = \frac{1}{2}A_\sigma x' + \frac{1}{2}A_\sigma x''.$$

Du moment que $A_\sigma x' \leqslant b_\sigma$, $A_\sigma x'' \leqslant b_\sigma$, la dernière égalité entraine $A_\sigma x' = b_\sigma = A_\sigma x''$, d'où $A_\sigma(x' - x'') = 0$, ce qui contredit l'hypothèse faite sur le rang de A_σ (car $x' - x'' \neq 0$).
Donc le théorème est démontré.
Ce théorème nous donne deux résultats importants.

Théorème 1.15 *Les points extrêmes de l'ensemble Ω, s'ils en existent, sont en nombre fini.*

Preuve : D'aprés le théorème précédent, à tout point extrême ,il correspond une matrice carrée non dégénérée A_σ qui le définit complètement : $x_0 = A_\sigma^{-1} b_\sigma$.

Les points extrêmes de Ω sont donc en nombre fini (car il en est de même des sous-matrices de A).

Théorème 1.16 *Si l'ensemble Ω est borné, alors c'est un polyèdre convexe.*

On le prouve à l'aide de la définition (1.7) d'un polyèdre convexe et du théorème (1.5) $\big[$tout ensemble fermé et borné et (compact) est l'envloppe convexe de ses points extrêmes $\big]$.

1.3.2 Systèmes homogènes.

La démonstration du théorème 1.13 suggère l'idée qu'on ferait bien d'imiter le cas des systèmes d'équations linéaires et d'étudier le système homogène $Ax \leqslant 0$.

Définition 1.13 *Si le rang de la matrice A est égal à n, l'ensemble $\Omega = \{x : \quad Ax \leqslant b\}$ est dit régulier.*

Théorème 1.17 *L'ensemble Ω présente des points extrêmes si et seulement si il est régulier. Voir [1].*

1.4 Fonctions convexes

1.4.1 Définitions.

Soit Ω un ouvert de $\mathbb{R}^n$, et soient a, b deux points de Ω tels que le segment $(1-t)a + tb \in \Omega$ pour tout $t \in [0,1]$.

$$\begin{aligned} f:\ \Omega &\longrightarrow \mathbb{R} \\ x &\longrightarrow f(x). \end{aligned}$$

On suppose que f admet des dérivées partielles dans Ω.
Posons : $g(t) = f((1-t)a + tb)$.
Alors : $\left\{ \begin{aligned} g:\ [0,1] &\longrightarrow \mathbb{R} \\ t &\longrightarrow g(t) \end{aligned} \right.$ est dérivable dans $[0,1]$,

et on a $\quad g'(t) = \displaystyle\sum_{i=1}^{n} (b_i - a_i) \frac{\partial f}{\partial x_i}((1-t)a + tb)$, $\big($la dérivation d'une fonction composée. $\big)$

$$\sum_{i=1}^{n}(b_i-a_i)\frac{\partial f}{\partial x_i}((1-t)a+tb)=\Big\langle b-a,\nabla f((1-t)a+tb)\Big\rangle$$

où $b=(b_i)_{i=1,2,\ldots,n}$ et $a=(a_i)_{i=1,2,\ldots,n}$

Mais d'aprés le théorème de Lagrange , on a :

$$g(1)-g(0)=(1-0)g'(t_0) \text{ où } t_0\in[0,1],$$
$$g(1)=f(b) \quad g(0)=f(a) \quad t_0=\varepsilon, \text{ on a}$$

$$\begin{aligned} f(b)-f(a) &= \Big\langle b-a,\nabla f((1-\varepsilon)a+\varepsilon b)\Big\rangle \\ &= \sum_{i=1}^{n}(b_i-a_i)\frac{\partial f}{\partial x_i}((1-\varepsilon)a+\varepsilon b) \quad \text{où } \varepsilon\in[0,1]. \end{aligned}$$

Soit $f:\Omega\longrightarrow \mathbb{R}$, $\quad \begin{array}{ll}\varphi: & I\subset\mathbb{R}\longrightarrow\Omega \\ & t\longrightarrow\varphi(t)\end{array}$ $\quad \Omega$ un ouvert de $\mathbb{R}^n$.

Considérons l'hypersurface $f(x)=b$, et une courbe φ telle que $\varphi(t)=x$ et $f(x)=b$, φ est une courbe qui appartient à l'hypersuface $f(x)=b$.
Soit $t_0\in I \quad \varphi(t_0)=x_0; \quad f(x_0)=b.$

Définition 1.14 *On dit qu'une droite est tangente à une surface en un point x_0 si elle est tangente à n'importe quelle courbe φ passant par x_0.*

Définition 1.15 *Un point x_0 de la surface $f(x)=b$ est un point singulier si $\frac{\partial f}{\partial x_i}(x_i)=0$ $i=1,2,\ldots,n$. Et on dira que c'est un point simple s'il existe $i_0\in\{1,2,\ldots,n\}$ tel que : $\frac{\partial f}{\partial x_{i_0}}(x_0)\neq 0$.*

Toutes les droites tangentes à la surface en un point simple x_0 appartiennent à un même plan : applé plan tangent à la surface au point x_0.

$$\begin{array}{l} I\longrightarrow\Omega \\ t\longrightarrow\varphi(t)=(\varphi_1(t),\ldots,\varphi_n(t)) \end{array}$$

$$x_1 = \varphi_1(t), \quad \varphi_1'(t) = \frac{dx_1}{dt}$$
$$\vdots$$
$$x_n = \varphi_n(t), \quad \varphi_n'(t) = \frac{dx_n}{dt}$$

$(\varphi_1'(t_0), \ldots, \varphi_n'(t_0))$ est un vecteur tangent à la surface $f(x) = b$ au point x_0 $(\varphi(t_0) = x_0)$. Et comme $f(x) = b$ on a :
$f(x) - b = 0 \Rightarrow \frac{\partial f}{\partial x_1}.\frac{\partial x_1}{\partial t} + \frac{\partial f}{\partial x_2}.\frac{\partial x_2}{\partial t} + \ldots + \frac{\partial f}{\partial x_n}.\frac{\partial x_n}{\partial t} = 0.$
$N = (\frac{\partial f(x_0)}{\partial x_1}, \ldots, \frac{\partial f(x_0)}{\partial x_n})$ est le vecteur normal à la surface au point x_0.

Soit x_0 un point, $f(x) = b$, et soit x un point du plan tangent à la surface au point x_0.
Alors $(x - x_0)$ est un vecteur du plan tangent perpendiculaire à N donc $\left\langle \frac{\partial f(x_0)}{\partial x_i}, x - x_0 \right\rangle = 0$ ce qui implique que $\left\langle \nabla f(x_0), x - x_0 \right\rangle = 0$.
L'équation du plan tangent à la surface $f(x) = b$ est :

$$\left\langle \nabla f(x_0), x \right\rangle = \left\langle \nabla f(x_0), x_0 \right\rangle. \tag{1.6}$$

Définition 1.16 *Soit Ω un convexe de $\mathbb{R}^n$ et soit $f : \Omega \to \mathbb{R}$. On dit que f est convexe si pour tout x_1, x_2 de Ω, et pour tout $\lambda \in [0, 1]$, on a $f(\lambda x_2 + (1 - \lambda)x_1) \leq \lambda f(x_2) + (1 - \lambda)f(x_1)$.*

On dira que f est concave si $-f$ est convexe.

Théorème 1.18 *Soit f une fonction de classe C^1 dans un ouvert convexe Ω de $\mathbb{R}^n$. Si f est convexe dans Ω, alors on a :*
$\langle (x_2 - x_1), \nabla f(x_1) \rangle \leqslant f(x_2) - f(x_1)$ pour tout x_1, x_2 de Ω.

Preuve : En effet,

$$f(\lambda x_2 + (1 - \lambda)x_1) \leqslant \lambda f(x_2) + (1 - \lambda)f(x_1) \quad \forall \lambda \in [0, 1].$$

Pour tout $\lambda \in]0, 1]$, on a $\dfrac{f(\lambda x_2 + (1 - \lambda)x_1) - f(x_1)}{\lambda} \leqslant f(x_2) - f(x_1)$.
On pose $x_\lambda = \lambda x_2 + (1 - \lambda)x_1$

D'après (1),

$$\begin{array}{rcl} f(x_\lambda)-f(x_1) & = & \big\langle \lambda x_2+(1-\lambda)x_1-x_1, \nabla f[(1-\varepsilon)x_1+\varepsilon(\lambda x_2+(1-\lambda)x_1)]\big\rangle \\ & = & \big\langle \lambda(x_2+x_1), \nabla f(x_1+\varepsilon\lambda(x_2-x_1))\big\rangle \\ & = & \lambda\big\langle (x_2+x_1), \nabla f(x_1+\varepsilon\lambda(x_2-x_1))\big\rangle \\ & \Rightarrow & \big\langle (x_2+x_1), \nabla f(x_1+\varepsilon\lambda(x_2-x_1))\big\rangle \\ & \leqslant & f(x_2)-f(x_1)\ \forall\lambda\in]0,1] \end{array}$$

par passage à la limite quand $\lambda \to 0$, on obtient :

$$\big\langle (x_2+x_1), \nabla f(x_1)\big\rangle \leqslant f(x_2)-f(x_1).$$

1.5 L'hyperplan tangent à l'hypersurface.

Théorème 1.19 *Soit f une fonction convexe alors l'ensemble défini par : $\{x : f(x) \leqslant b, x \geqslant 0\}$ est convexe.*

Preuve : En effet, posons $\Omega = \{x : f(x) \leqslant b, x \geqslant 0\}$.
Si Ω contient un point le théorème est évident ;

Sinon, soient x_1, x_2 deux points de Ω, alors $\lambda x_1 + (1-\lambda)x_2 \in \Omega$ pour tout $\lambda \in [0,1]$,

$$f(\lambda x_1 + (1-\lambda)x_2) \leqslant \lambda f(x_1) + (1-\lambda) f(x_2) \leqslant b$$

ce qui implique que $\lambda x_1 + (1-\lambda)x_2 \in \Omega$ d'où la convexité de Ω.

Remarque 1.7 *Notons que $\{x : f(x) = b, x \geqslant 0\}$ n'est pas en général convexe car :*

$$f(\lambda x_1 + (1-\lambda)x_2) \leqslant \lambda f(x_1) + (1-\lambda) f(x_2) \leqslant b.$$

Cependant, l'application : $x \to \langle a, x\rangle = ax$ est convexe et
$\Omega = \{x : \langle a, x\rangle = b, x \geqslant 0\}$ est convexe.

Théorème 1.20 *Soit f une fonction convexe de classe C^1 pour tout $x \geqslant 0$. Alors l'hyperplan tangent à l'hypersurface $f(x) = b$ au point $x_0 \geqslant 0$ est l'hyperplan d'appui où $\Omega = \{x \geqslant 0 : f(x) \leqslant b\}$ au point x_0.*

Preuve : Ω est convexe ; comme f admet des dérivées partielles continues, elle est continue, donc Ω est un convexe fermé.

On sait qu'on peut construire un hyperplan à Ω au point x_0 voir [15] et voir [23].
Ici, on veut montrer que cet hyperplan à Ω au point x_0 ; $(f(x_0) = b)$ est le plan tangent à la surface $f(x) = b$ au point x_0.
Mais on sait que l'équation de l'hyperplan tangent à l'hypersrface $f(x) = b$ au point $x_0 \geqslant 0$ est

$$\langle \nabla f(x_0), x \rangle = \langle \nabla f(x_0), x_0 \rangle \text{ voir } (1.6).$$

Il est évident que l'hperplan $\langle \nabla f(x_0), x \rangle$ contient le point x_0.

Il reste à montrer qu'il sépare le convexe fermé $\Omega = \{x \geqslant 0 : f(x) \leqslant b\}$.
Soit $x_1 \in \Omega$; $x_1 \geqslant 0 : f(x_1) \leqslant b$;

$$\begin{aligned}\langle \nabla f(x_0), x_0 - x_1 \rangle &= \langle \nabla f(x_0), x_0 \rangle - \langle \nabla f(x_0), x_1 \rangle \\ &\leqslant f(x_0) - f(x_1) = b - f(x_1) \\ &\leqslant 0\end{aligned} \quad .$$

Ce qui implique que $\quad \langle \nabla f(x_0), x_0 \rangle \leqslant \langle \nabla f(x_0), x_1 \rangle$
Et par conséquent

$$\langle \nabla f(x_0), x \rangle \leqslant \langle \nabla f(x_0), x_1 \rangle \quad \forall x_1 \in \Omega$$

On peut de la même manière démontrer les résultats suivants pour les fonctions concaves.

Théorème 1.21 *Soit f une fonction concave de classe $\mathcal{C}^1$ dans un ouvert convexe $\Omega \subset \mathbb{R}^n$. Alors* $\quad \left\langle \nabla f(x_1), x_2 - x_1 \right\rangle \geqslant f(x_2) - f(x_1)$

Théorème 1.22 *Si f est une fonction concave pour tout $x_0 \geqslant 0$, alors l'ensemble $\Omega = \{x : f(x) \geqslant b ; x \geqslant 0\}$. Alors Ω est convexe.*

Théorème 1.23 *Si $f$est une fonction concave de classe $\mathcal{C}^1$ pour tout $x_0 \geqslant 0$ Alors l'hyperplan tangent à l'hypersurface $f(x) = b$ au point $x_0 \geqslant 0$ est l'hyperplan d'appui à $\Omega = \{x \geqslant 0 : f(x) \geqslant b\}$.*

1.6 Minimum et maximum d'une fonction convexe.

Théorème 1.24 *Soit f une fonction convexe définie sur un convexe fermé Ω de $\mathbb{R}^n$. Alors le minimum locale de f est un minimum global de f sur Ω.*

Preuve : On raisonne par l'absurde : soit $x_0 \in \Omega$ un minimum local de f, supposons qu'il existe un point $x^* \in \Omega$ tel que : $f(x^*) < f(x_0)$.

$$f(\lambda x^* + (1-\lambda)x_0) \leqslant \lambda f(x^*) + (1-\lambda)f(x_0) < \lambda f(x_0) + (1-\lambda)f(x_0) = f(x_0).$$

Mais $[\lambda x^* + (1-\lambda)x_0] \in \Omega$ et $f(\lambda x^* + (1-\lambda)x_0) < f(x_0),\ \forall \lambda \in [0,1]$ pour λ assez petit ; $(\lambda x^* + (1-\lambda)x^*) \in V(x_0)$ (voisinage de x_0), donc il existe un voisinage de $x_0,\ x \in V(x_0) \Rightarrow f(x) < f(x_0),\ x = \min(\lambda x^* + (1-\lambda)x_0)$. Contradiction avec le fait que x_0 est un minimum local.

Théorème 1.25 *Soit f une fonction convexe sur un convexe fermé Ω de $\mathbb{R}^n$. Alors l'ensemble des points minimaux locaux de f est convexe.*

Preuve : Soient x_1, x_2 deux points minimaux locaux $x_1 \neq x_2$; posons $x = \lambda x_1 + (1-\lambda)x_2, \quad \lambda \in [0,1]$

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

Mais : $f(x_1) = f(x_2) = f(\overline{x})$ est la valeur minimale de f dans Ω, donc

$$f(x) \leqslant \lambda f(\overline{x}) + (1-\lambda)f(x^*) = f(x^*) \Rightarrow f(x) = f(x^*) \quad \forall \lambda \in [0,1].$$

D'où x_1, x_2 deux points minimaux locaux $\Rightarrow [\lambda x_1 + (1-\lambda)x_2]$ est aussi un minimum local.

Remarque 1.8 $f(x) = \langle a, x \rangle$, *$f$ est convexe. Si x_1 est un minimum de f, et x_2 un autre alors le segment devient un minimum (Cas du simplexe), voir [17], et voir [24].*

Théorème 1.26 *Si f est strictement convexe sur un convexe fermé Ω de $\mathbb{R}^n$, alors le minimum local qui est global est unique.*

Preuve : Supposons qu'on ait deux points $x_1, x_2 \in \Omega$ tel que $x_1 \neq x_2$;

$f(\lambda x_1 + (1-\lambda)x_2) < \lambda f(x_1) + (1-\lambda) f(x_2) = f(x^*) \Rightarrow f(x) < f(x^*)$ contradiction.

Théorème 1.27 *Soit f une fonction de classe $\mathcal{C}^1$ sur un convexe Ω de $\mathbb{R}^n$. Si f est convexe, et si $\nabla f(x_0) = 0$, avec $x_0 \in \mathring{\Omega}$ (intérieur de Ω). Alors x_0 est le minimum global.*

Preuve : On a : $\langle x - x_0, \nabla f(x_0) \rangle \leqslant f(x) - f(x_0)$;

$$\nabla f(x_0) = 0 \Rightarrow f(x) - f(x_0) \geqslant 0 \Rightarrow f(x) \geqslant f(x_0) \text{ pour tout } x \in \Omega.$$

Donc x_0 est le minimum global de f

Théorème 1.28 *Soit f une fonction convexe sur un convexe fermé borné superieurement. Et Si le maximum global de f est fini, alors il est atteint en un ou plusieurs sommets (extrême).*

Preuve : $i)$ Supposons que $\max\limits_{x \in \Omega} f(x) = f(x^*)$ et que $x^* \in \mathring{\Omega}$.

Si f n'est pas constante, $\exists x_1 \in \Omega$ tel que $f(x_1) < f(x^*)$, mais $x^* \in \mathring{\Omega}$ donc $x^* = \lambda x_2 + (1-\lambda)x_1$ pour $\lambda \in]0, 1[$.

$$\begin{aligned} f(x^*) &\leqslant \lambda f(x_2) + (1-\lambda) f(x_1) \\ &< \lambda f(x_2) + (1-\lambda) f(x^*) \end{aligned}$$

D'où $\lambda[f(x^*) - f(x_2)] < 0 \Rightarrow f(x^*) < f(x_2)$ contradiction.

Donc $f(x) = c^{st}$ sur Ω.

Mais un ensemble Ω de $\mathbb{R}^n$ fermé et borné superieurement possède au moins un point extrème $\overline{x}$ et $f(\overline{x}) = c^{st}$.

d'où $\max\limits_{x \in \Omega} f(x) = c^{st} = f(\bar{x})$;

$ii)$ Supposons que x^* est un point frontière et que le maximum de f n'est pas un point intérieur de Ω. On doit montrer que x^* est point extrême. On sait qu'il existe un hyperpaln H, qui sépare x^* et le convexe fermé Ω, x^* est un point frontière donc l'hyperplan d'appui à Ω au point x^* est H_1.

Posons $T_1 = \Omega \cap H_1 \subset \Omega$, T_1 est fermé, convexe et borné inférieurement donc possède au moins un point extrême ; mais $T_1 \subset H_1$ (hyperpaln) donc

$$\dim T_1 = n - 1; \quad x^* \in H_1, \quad x^* \in \Omega \Rightarrow x^* \in T_1.$$

Si x^* est un point intérieur à T_1 alors d'après $i)$ $f = c^{st}$ donc on aura une contradiction ;
Sinon, $x^* \in T_1$ est un point frontière de T_1 (d'après ii) posons
$T_2 = T_1 \cap H_2, \ldots$ on arrive alors à $T_n = T_{n-1} \cap H_n$ et on a $\dim T_n = 0$ donc T_n contient un point unique.
Ce point unique est bien évidemment x^*, d'où $T_n = \{x^*\}$.
Un singleton est un point extrême car T_n est fermé et borné inférieurement donc il doit contenir un point extrème.
D'une manière analogue, on peut démontrer les théorèmes suivants,voir [25] :

Théorème 1.29 *Soit f une fonction concave définie sur un convexe fermé de $\mathbb{R}^n$. Alors le maximum local de f est global.*

Théorème 1.30 *Soit f une fonction concave définie sur un convexe fermé de $\mathbb{R}^n$. Alors l'ensemble des maximaux globaux est un convexe.*

Théorème 1.31 *Soit f une fonction concave définie sur un convexe fermé de $\mathbb{R}^n$. Si x_1 et x_2 sont deux maximaux alors $(\lambda x_1 + (1-\lambda)x_2)$ est aussi un maximum global pour tout $\lambda \in [0,1]$*

Théorème 1.32 *Soit f une fonction de classe C^1 dans un ouvert convexe Ω de $\mathbb{R}^n$. f est concave, et $\nabla f(x_0) = 0$, avec $x_0 \in \overset{\circ}{\Omega}$. Alors f atteint son maximum global au point x_0.*

Théorème 1.33 *Soit f une fonction de classe C^1 sur un convexe fermé et borné superieurement et si f est concave sur Ω, et si le maximum global de f existe, alors il est atteint en un point extrême du convexe Ω.*

Chapitre 2

Optimum de la fonction ψ

2.1 Théorèmes fondamentaux de la programmation quadratique

Considérons le problème canonique (P) suivant :

$$(P) : \begin{cases} \max z,\ \ z = \psi(x), \\ Ax \leqslant b, \\ x \geqslant 0, \end{cases}$$

ou $\psi(x) = \sum_{i=1}^{n} \alpha_i x_i + \sum_{i=}^{n} \beta_i x_i^2$; les nombres réels α_i sont quelconques, et $\forall i, \beta_i \geqslant 0$. A une matrice de type $m \times n$ a coefficients réels ; et b un vecteur de $\mathbb{R}_+^m$.

Définition 2.1 *Une solution réalisable x du probleme (P) est dite sommet de (P) si elle ne peut s'écrire sous la forme $x = \lambda x_0 + (1 - \lambda)x_1$ avec x_0, et x_1 deux solutions réalisables distinctes de (P) et $\lambda \in]0, 1[$*

Proposition 2.1 *Les deux assertions suivantes sont équivalentes :*

1. *x est un sommet de (P).*
2. *Pour toutes solutions réalisables x_0, et x_1 de (P) ; et pour tout λ, $0 < \lambda < 1$ $(x = \lambda x_0 + (1 - \lambda)x_1) \Rightarrow (x = x_0 = x_1)$.*

La preuve est immediate (il suffit d'utiliser la définition),.voir [7]
Pour tout $t \in [0, 1]$, posons $g(t) = \psi(tx_1 + (1-t)x_0)$, où $x_0 = (x_0^1, x_0^2, ..., x_0^n)$

et $x_1 = (x_1^1, x_1^2, ..., x_1^n)$ sont deux vecteurs de IR^n. Notons que $g(0) = \psi(x_0)$ et que $g(1) = \psi(x_1)$. Pour étudier le signe de $\psi(x_1) - \psi(x_0)$, utilisons le théorème de Lagrange : g est évidemment continue sur $[0, 1]$ et dérivable sur $]0, 1[$. Alors il existe $\xi \in]0, 1[$ tel que $g(1) - g(0) = g'(\xi)$.

On calcule $g'(t)$ en utilisant le théorème de la dérivée d'une fonction composée :

$$g'(t) = \sum_{i=1}^{n} (x_1^i - x_0^i) \frac{\partial \psi}{\partial x_i}(tx_1 + (1-t)x_0).$$

Comme $\psi(x) = \sum_{i=1}^{n} \alpha_i x_i + \sum_{i=1}^{n} \beta_i x_i^2$ alors $\frac{\partial \psi(x)}{\partial x_i} = \alpha_i + 2\beta_i x_i$ et donc

$$\begin{aligned} \frac{\partial \psi}{\partial x_i}(tx_1 + (1-t)x_0) &= \alpha_i + 2\beta_i(tx_1^i + (1-t)x_0^i) \\ &= \alpha_i + 2\beta_i t[x_1^i - x_0^i] + 2\beta_i x_0^i \end{aligned}$$

Par suite

$$g'(t) = \sum_{i=1}^{n} (x_1^i - x_0^i)\alpha_i + 2\sum_{i=1}^{n} (x_1^i - x_0^i)\beta_i x_0^i + 2t\sum_{i=1}^{n} (x_1^i - x_0^i)^2 \beta_i;$$

et donc $g''(t) = 2\sum_{i=1}^{n} (x_1^i - x_0^i)^2 \beta_i$. Par conséquent la fonction $g'(t)$ est monotone croissante puisque $\beta_i \geqslant 0$

$g'(t)$ est ainsi croissante : $g''(t) = 2\sum_{i=1}^{n} (x_1^i - x_0^i)^2 \beta_i \geqslant 0$.

Notons que si $g''(t) = 0$ alors $g'(t)$ est constante et g est monotone. Sinon, posons

$$t_0 = -\frac{\sum_{i=1}^{n} (x_1^i - x_0^i)\alpha_i + 2\sum_{i=1}^{n} (x_1^i - x_0^i)\beta_i x_0^i}{2\sum_{i=1}^{n} (x_1^i - x_0^i)^2 \beta_i}.$$

On a alors : le taleau de la variation de la fonction g

t	0		t_0		1
$g'(t)$		$-$	0	$+$	
$g(t)$	$g(0)$	$\searrow$	$g(t_0)$	$\nearrow$	$g(1)$

Remarquons que si $t_0 \notin]0,1[$ alors la fonction g est monotone croissante ou decroissante.
Les deux resultats suivants que nous avons établis sont essentiels :

Théorème 2.1 *Si (P) admet une solution optimale unique x, alors x est un sommet.*

Preuve : Supposons qu'il existe deux solutions réalisables distinctes x_0 et x_1. Soit une solution optimale x. Supposons que $x = (1 - t_1)x_0 + t_1 x_1$ où $t_1 \in]0,1[$. Deux cas sont possibles :

1. Si $t_1 \geqslant t_0$ alors $g(t_1) > g(t_0)$ puisque $g(t)$ est croissante (voir le tableau de variations ci-dessus) ; et $t_1 \in [t_0, 1[$. On a donc $g(t_1) < g(1)$.
2. Si $t_1 < t_0$ alors $g(t_1) > g(t_0)$ puisque $g(t)$ est décroissante sur $]0, t_0]$; et $t_1 \in]0, t_0]$. On a alors $g(t_1) < g(0)$.

Ceci montre que le maximum est atteint soit en $t = 0$ soit en $t = 1$. Comme x est unique alors ou bien $x = g(0) = x_0$ ou bien $x = g(1) = x_1$, donc x est un sommet.
Remarquons que si g est monotone, alors la solution optimale x est un sommet. Cela découle directement de la démonstration précédente.

Théorème 2.2 *Soit x_0, x_1 deux solutions optimales de (P). Alors il existe un nombre $\lambda \in]0,1[$ tel que la solution réalisable $x = (1 - \lambda)x_0 + \lambda x_1$ ne soit pas optimale.*

Preuve : Si x_0 et x_1 sont deux solutions optimales du problème (P) alors $\psi(x_0) = \psi(x_1)$ puisque chacune réalise le maximum de z, c'est-à-dire $g(0) = g(1)$. D'après le théorème de Rolle il existe $\lambda \in]0, 1[$ tel que $g'(\lambda) = 0$. On peut donc prendre $\lambda = t_0$.
D'après le tableau de variations précédent de la fonction g on a $g(\lambda) < g(1) = g(0)$ et donc $\psi(x) < \psi(x_0) = \psi(x_1)$. La solution x n'est donc pas optimale.

Remarque 2.1 *Le théorème (2.2) montre que si x_0 et x_1 sont deux solutions optimales, alors on peut trouver une solution x du segment $]x_0, x_1[$ qui ne soit pas optimale.*

Remarque 2.2 *Notons que si $\beta_i = 0$ pour tout i, $g''(t) = 0$. Ceci implique que $g'(t)$ est constante pour tout $t \in [0, 1]$ et, par suite, $g(0) = g(1) = g(t)$. Par conséquent tout le segment $[x_0, x_1]$, est constitue de solutions optimales. On retrouve le cas particulier du simplexe.*

Considérons maintenant le programme standard (standarisé de (P)) suivant :

$$(P), \begin{cases} \max z, \ \ z = \psi(x) \\ Ax = b \\ x \geqslant 0 \end{cases}$$

Notons par $A^1, ..., A^n$ les colonnes de A. On suppose que A est de rang $m (m < n)$. Alors $\{A^1, A^2, ..., A^n\}$ engendrent $\mathbb{R}^m$. Comme $m < n$, alors il existe au plus C_n^m bases $\{A^{i_1}, ..., A^{i_m}\}$ extraites de $\{A^1, A^2, ..., A^n\}$.

Théorème 2.3 *(caractérisation d'un sommet) Soit $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ une solution réalisable de (P), et, $i_1, i_2, ..., i_r$ les indices i tel que $x_i > 0$. Alors les deux assertions suivantes sont équivalentes :*

1. *x est un sommet*
2. *$\{A^{i_1}, ..., A^{i_r}\}$ est une partie libre de $\mathbb{R}^m$.*

Preuve : 1. Supposons que le système $\{A^{i_1}, ..., A^{i_r}\}$ est libre et prouvons que x est un sommet.
Soient x_0, x_1 deux solutions réalisables distinctes de (P). Supposons que $x = \lambda x_0 + (1-\lambda)x_1$ avec $\lambda \in]0,1[$.
L'équation $Ax = b$ s'écrit alors $\sum_{k=1}^{n} x_k A^k = b$ ou encore $\sum_{k=1}^{n}(\lambda x_0^k + (1-\lambda)x_1^k)A^k = b$. Pour $k \notin \{i_1, \dots, i_r\}$ $x_0^k = x_1^k = 0$ car $\lambda > 0$ et $(1-\lambda) > 0$ puisque tous les x_0^i et x_1^i sont positifs.
Comme x_0, x_1 sont deux solutions réalisables de (P) alors $\sum_{j=1}^{r} x_0^j A^{i_j} = b$ et $\sum_{j=1}^{r} x_1^j A^{i_j} = b$. Par suite $\sum_{j=1}^{r}(x_0^j - x_1^j)A^{i_j} = 0$ et comme $\{A^{i_1}, ..., A^{i_r}\}$ est libre alors $x_0^j - x_1^j = 0$, $\forall j \in \{1, 2, ..., r\}$.
Par conséquent $x_0 = x_1$ et donc $x = x_0 = x_1$ ce qui prouve que x est un sommet.
2. Inversement supposons x un sommet et montrons que $\{A^{i_1}, ..., A^{i_r}\}$ est libre.
Soit $\lambda_1, ..., \lambda_r$ des scalaires tels que $\sum_{j=1}^{r} \lambda_j A^{i_j} = 0$ comme x est un sommet alors $\sum_{j=1}^{r} x_j A^j = b$ et par suite, $\sum_{j=1}^{r} x_{i_j} A^{i_j} = b$ puisque les autres composantes de x sont nulles. Par conséquent, pour tout ν et tout μ, on voit que $\sum_{j=1}^{r}(x_{i_j} + \nu\lambda_j)A^{i_j} = b$ et $\sum_{j=1}^{r}(x_{i_j} - \mu\lambda_j)A^{i_j} = b$. Comme $x_{i_j} > 0$ alors on peut choisir ν et μ de sorte que les vecteurs y et z de composantes non nulles $x_{i_j} + \nu\lambda_i \geqslant 0$ et $x_{i_j} - \mu\lambda_i \geqslant 0$ soient solutions réalisables de (P). En effet si $\lambda_j > 0$, on peut choisir μ positif verifiant $0 < \mu \leqslant \frac{x_{i_j}}{\lambda_i}$ (on peut prendre $\mu = \nu$). Si $\lambda_j < 0$ on peut choisir ν positif quelconque et verifiant $0 < \nu \leqslant -\frac{x_{i_j}}{\lambda_i}$ (on peut prendre $\nu = \mu$). On voit donc qu'on peut choisir un nombre $\theta > 0$ assez petit tel que $x_{i_j} + \theta\lambda_j \geqslant 0$ et $x_{i_j} - \theta\lambda_j \geqslant 0$. Avec ces nouvelles composantes pour y et z on voit que $x = \frac{1}{2}(y+z)$ [on notera que les composantes y_j z_j, pour $j \notin \{i_1, i_2, ..., i_r\}$, sont nulles]. Donc, puisque

x est un sommet, $x = y = z$. Par conséquent le vecteur de composantes $\theta\lambda_j$ est un vecteur nul puisque $\theta > 0$. D'où $\lambda_j = 0$ pour tout $j = 1, 2, ..., r$. Le théorème est démontré. On rappelle les définitions suivantes

Définition 2.2 *On appelle base réalisable de (P) tout système B extrait des colonnes de la matrice A et tel que :*

1. *B soit une base de $\mathbb{R}^m$,*
2. *Les composantes de b suivant B soient toutes non négatives.*

Définition 2.3 *Une solution réalisable x de (P) est appelée solution réalisable de base s'il existe une base réalisable B de (P) telle que x soit la solution relativement à la base B.*

Corollaire 2.1 *Toute solution réalisable de base est un sommet.*

Preuve : Soit x une solution réalisable de base, et soit $B = \{A^{i_1}, ..., A^{i_m}\}$ une base réalisable de (P).
Par définition même B est une base de $\mathbb{R}^m$, donc libre. Et d'aprés le théorème (2.3) il résulte que x est un sommet.

Corollaire 2.2 *A tout sommet x on peut associer au moins une solution réalisable de base telle que $x = x_B$.*

En effet soit x un sommet $x =^t (x_1, ..., x_n)$. Notons par $I^+(x)$ l'ensemble de tous les indices i tels que $x_i > 0$, $I^+(x) = \{i_1, \ldots, i_n\}$. Alors, d'après le théorème (2.3), $\{A^{i_1}, ..., A^{i_r}\}$ est une partie libre de $\mathbb{R}^m$. Si $r = m$, $\{A^{i_1}, ..., A^{i_r}\}$ constitue donc une base B de $\mathbb{R}^m$ et on a bien $x = x_B$. Si $r < m$, on peut compléter cette famille $\{A^{i_1}, ..., A^{i_r}\}$ de façon à obtenir une base B de $\mathbb{R}^m$ et, là aussi, on a bien $x = x_B$.

Théorème 2.4 *Si le standarisé de (P) admet des solutions optimales de la forme $x =^t (x_1, ..., x_k, 0, 0, ..., 0)$ alors, au moins, un sommet est solution optimale.*

Preuve : Si $\{A^1, ..., A^k\}$ est libre alors x est un sommet, d'après le théorème (2.3). Sinon soit l le rang de ce système de vecteurs ($l < k$). Sans restreindre la généralité on peut supposer que la famille $\{A^1, ..., A^l\}$ est libre. Par suite $\{A^1, ..., A^l, A^{l+1}\}$ est liée. Il existe donc $\lambda_1, \lambda_2, ..., \lambda_l, \lambda_{l+1}$ non tous nuls tels que $\sum_{i=1}^{l+1} \lambda_i A^i = 0$, et donc $\sum_{i=1}^{l+1} \pm\theta\lambda_i A^i = 0$, pour tout réel θ. Posons $\lambda_{l+2} = \cdots = \lambda_n = 0$.

On voit alors que $\sum_{i=1}^{k} (x_i \pm \theta\lambda_i) A^i = b$, puisque $\sum_{i=1}^{k} \lambda_i A^i = b$. En particulier on peut choisir, par un raisonnement analogue à celui qui a été fait dans la démonstration du théorème (2.3), $\theta > 0$ suffisamment petit de sorte que $(x_i + \theta\lambda_i)$ et $(x_i - \theta\lambda_i)$ soient positifs. Les vecteurs x_θ et $x_{-\theta}$ de composantes $(x_i + \theta\lambda_i)$ et $(x_i - \theta\lambda_i)$ sont alors deux solutions réalisables, et on a $x = \frac{1}{2}(x_\theta + x_{-\theta})$.

Mais : $\psi(x_\theta) = \sum_{i=1}^{n} \alpha_i (x_i + \theta\lambda_i) + \sum_{i=1}^{n} \beta_i (x_i + \theta\lambda_i)^2$

et $\psi(x_{-\theta}) = \sum_{i=1}^{n} \alpha_i (x_i - \theta\lambda_i) + \sum_{i=1}^{n} \beta_i (x_i - \theta\lambda_i)^2$.

Alors : $\dfrac{\psi(x_\theta) + \psi(x_{-\theta})}{2} = \sum_{i=1}^{n} \alpha_i x_i + \sum_{i=1}^{n} \beta_i x_i^2 + \sum_{i=1}^{n} \beta_i \theta^2 \lambda_i^2$

d'où $\dfrac{\psi(x_\theta) + \psi(x_{-\theta})}{2} = \psi(x) + \theta^2 \sum_{i=1}^{n} \beta_i \lambda_i^2$.

Comme $\psi(x) \geqslant \psi(x_\theta)$ et $\psi(x) \geqslant \psi(x_{-\theta})$

alors on obtient :

$$\begin{aligned}
\psi(x_\theta) &= \sum_{i=1}^n \alpha_i x_i + \theta \sum_{i=1}^n \alpha_i \lambda_i + \sum_{i=1}^n \beta_i x_i^2 + 2\theta \sum_{i=1}^n \beta_i x_i \lambda_i + \theta^2 \sum_{i=1}^n \beta_i \lambda_i^2 \\
&\leqslant \sum_{i=1}^n \alpha_i x_i + \sum_{i=1}^n \beta_i x_i^2 \\
\psi(x_{-\theta}) &= \sum_{i=1}^n \alpha_i x_i - \theta \sum_{i=1}^n \alpha_i \lambda_i + \sum_{i=1}^n \beta_i x_i^2 - 2\theta \sum_{i=1}^n \beta_i x_i \lambda_i + \theta^2 \sum_{i=1}^n \beta_i \lambda_i^2 \\
&\leqslant \sum_{i=1}^n \alpha_i x_i + \sum_{i=1}^n \beta_i x_i^2
\end{aligned}$$

En simplifiant ces deux inéquations on trouve :

$\theta \left(\sum_{i=1}^n \alpha_i \lambda_i + 2 \sum_{i=1}^n \beta_i x_i \lambda_i + \theta \sum_{i=1}^n \beta_i \lambda_i^2 \right) \leqslant 0$

et $-\theta \left(\sum_{i=1}^n \alpha_i \lambda_i + 2 \sum_{i=1}^n \beta_i x_i \lambda_i - \theta \sum_{i=1}^n \beta_i \lambda_i^2 \right) \leqslant 0\,;$

et donc on a simultanément :

$\theta \left(\sum_{i=1}^n \alpha_i \lambda_i + 2 \sum_{i=1}^n \beta_i x_i \lambda_i \right) \leqslant - \left(\theta^2 \sum_{i=1}^n \beta_i \lambda_i^2 \right)$

et $\theta \left(\sum_{i=1}^n \alpha_i \lambda_i + 2 \sum_{i=1}^n \beta_i x_i \lambda_i \right) \geqslant \left(\theta^2 \sum_{i=1}^n \beta_i \lambda_i^2 \right)$ avec $\beta_i \geqslant 0$.

Par conséquent :

$\theta \left(\sum_{i=1}^n \alpha_i \lambda_i + 2 \sum_{i=1}^n \beta_i x_i \lambda_i \right) = 0,$

et par suite on a aussi : $\theta^2 \sum_{i=1}^n \beta_i \lambda_i^2 = 0,$

et donc $\psi(x_\theta) = \psi(x_{-\theta}) = \psi(x)$.

Les solutions x_θ, $x_{-\theta}$, x sont optimales on peut, par suite, choisir $\theta = \theta_0$ telle que l'un des vecteurs x_{θ_0} ou $x_{-\theta_0}$ a une composante nulle :

Posons : $x_\theta = (x_1 + \lambda_1\theta, x_2 + \lambda_2\theta, \ldots, x_{l+1} + \lambda_{l+1}\theta, x_{l+2}, \ldots, x_k, 0, \ldots, 0)^t$

et $x_{-\theta} = (x_1 - \lambda_1\theta, x_2 - \lambda_2\theta, \ldots, x_{l+1} - \lambda_{l+1}\theta, x_{l+2}, \ldots, x_k, 0, \ldots, 0)^t$. Pour

$\lambda_i > 0$, il faut choisir $\theta \leqslant \dfrac{x_i}{\lambda_i}$ pour que $x_{-\theta}$, soit solution réalisable, x_θ l'est déjà. De même, pour $\lambda_i < 0$, il faut choisir $\theta \leqslant -\dfrac{x_i}{\lambda_i}$ pour que x_θ soit solution réalisable, $x_{-\theta}$ l'est déjà. On voit donc qu'en choisissant $\theta = \theta_0 = \min\{\dfrac{x_i}{\lambda_i},\ \lambda > 0;\ -\dfrac{x_i}{\lambda_i}, \lambda_i < 0\}$ x_θ et $x_{-\theta}$ sont solutions réalisables. On remarquera que $\theta = \theta_0$ signifie qu'il existe un indice i_0 tel que $\theta_0 = \dfrac{x_{i_0}}{\lambda_{i_0}}$ (ici $\lambda_{i_0} > 0$) [ou bien $\theta_0 = -\dfrac{x_{i_0}}{\lambda_{i_0}}$ (ici $\lambda_{i_0} < 0$)]. Par conséquent une des composantes des solutions réalisables x_θ et $x_{-\theta}$ s'annule. Le nombre de composantes non nulles de ce vecteur est donc égal à $k-1$. Si ce vecteur (x_θ ou $x_{-\theta}$) est un sommet le théorème est complètement démontré. Sinon, on réitère le processus : on reprend avec x_θ (ou $x_{-\theta}$) à la place de x. Le nombre de composantes nulles augmente d'une unité à chaque fois, on parvient au bout d'un nombre fini d'applications du processus à une solution optimale qui est un sommet.

2.2 Variations de la fonction objectif lors d'un changement de base

2.2.1 Notations.

Considérons le programme standarisé de (P) suivant :

$$\begin{cases} \psi(x) = \displaystyle\sum_{i=1}^{n} \alpha_i x_i + \beta_i x_i^2 \\ Ax \leqslant b \\ x \geqslant 0 \end{cases}$$

Les nombres réels α_i sont de signes quelconques et $\beta_i \geqslant 0$, pour tout i.

On peut toujours supposer que $b = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}$, avec $b_i \geqslant 0$ pour tout i.

La matrice A s'écrit sous la forme : $A = (A_1\ A_2\ \ldots A_n\ A_{n+1}\ \ldots A_{n+m})$ où

$\{A_{n+1} \ \ldots A_{n+m}\} = B_0$ est la base canonique de (P).

Une solution de base (associée à B_0) s'écrit sous la forme : $x_0 = \begin{pmatrix} x_1 = 0 \\ \vdots \\ x_n = 0 \\ x_{n+1} = b_1 \\ \vdots \\ x_{n+m} = b_m \end{pmatrix}$

2.2.2 Variations de la fonction objectif lorsqu'on passe d'un sommet à un sommet voisin.

Soit $x(\theta)$ une solution réalisable voisine de x_0 (c'est-à-dire une solution réalisable obtenue à partir de x_0 en faisant croître l'une des variables hors base et en bloquant à 0 les autres variables hors base). Soit $x(\theta)$ le sommet voisin de x_0 :

$$x(\theta) = \begin{pmatrix} x_1 = 0 \\ \vdots \\ x_r = \theta \\ \vdots \\ x_n = 0 \\ x_{n+1} = x_{n+1}(\theta) \\ \vdots \\ x_{n+m} = x_{n+m}(\theta) \end{pmatrix}$$

Où r est l'indice de la variable sur laquelle on a choisi de faire porter l'accroissement $\theta \geqslant 0$ (r est la direction d'amélioration).

Si i est la direction d'amélioration, la solution réalisable $x(\theta)$ vérifie alors les contraintes de (P) :

$$\begin{cases} x(\theta) \geqslant 0 \\ \theta a_{ki} + x_{n+k}(\theta) = b_k \quad \text{pour } 1 \leqslant k \leqslant m \end{cases}$$

Pour que $x(\theta)$ soit solution réalisable, il faut que :
1. $x_{n+k}(\theta) = b_k - \theta a_{ki};\ 1 \leqslant k \leqslant m$,
2. $x_{n+k}(\theta) \geqslant 0$.

$x(\theta)$ est solution réalisable si et seulement si θ est solution du système :

$$(1)\left\{\begin{array}{l}\theta \geqslant 0\\ b_k \geqslant \theta a_{ki} + x_{n+k}(\theta) = b_k \quad \text{pour } 1 \leqslant k \leqslant m\end{array}\right.$$

On distingue deux cas :
1. $a_{ki} \leqslant 0$ pour tout i et pour tout k Alors $b_k - \theta a_{ki}$ pour tout $\theta \geqslant 0$, et pour tout nombre $\theta > 0$, $x_{n+k}(\theta) = b_k - \theta a_{ki}$ est positive. Donc pour tout $\theta > 0$ $x(\theta)$ est réalisable et la fonction objectif peut prendre une valeur aussi grande que l'on veut, et par suite l'extrémum n'existe pas.
2. Il existe des indices k et i tels que $a_{ki} > 0$
On choisit θ tel que $0 \leqslant \theta \leqslant \dfrac{b_k}{a_{ki}}$, et notons par $\theta_0 = \min\limits_{k}\{\dfrac{b_k}{a_{ki}}, a_{ki} > 0\} = \dfrac{b_{k_0}}{a_{k_0 i}}$.
On a alors $b_{k_0} - a_{k_0 i}\theta_0 = 0$; donc A_{n+k_0} est le vecteur sortant de la base.
Dans le cas 2, on note par :
$x(\theta) - x_0 =^t (0 \ldots \theta\, 0 \ldots\; x_{n+1}(\theta) \ldots 0\, x_{n+J+1}(\theta) \ldots x_{n+m}(\theta)) -^t (0 \ldots 0\, b_1 \;\ldots b_m) =^t (0 \ldots \theta\;\, 0 \ldots\;\, x_{n+1}(\theta) - b_1 \ldots - b_j\;\, x_{n+J+1}(\theta) - b_{j+1} \ldots x_{n+m}(\theta) - b_m)$ où $\theta = \min\limits_{a_{ki}>0}\{\dfrac{b_k}{a_{ki}}\}$; $x_{n+1}, \ldots, x_{n+M}$ sont les variables d'écarts, et $\alpha_k = \beta_k = 0$ pour $k = n+1, ..., n+m$.

$$\alpha = \begin{pmatrix}\alpha_1\\ \vdots\\ \alpha_n\\ 0\\ \vdots\\ 0\end{pmatrix}; \qquad \beta = \begin{pmatrix}\beta_1\\ \vdots\\ \beta_n\\ 0\\ \vdots\\ 0\end{pmatrix}; \qquad x(\theta) - x_0 = \begin{pmatrix}0\\ \vdots\\ \theta\\ 0\\ \vdots\\ x_{n+1}(\theta) - b_1\\ \vdots\\ -b_j\\ \vdots\\ x_{n+m}(\theta) - b_m\end{pmatrix}$$

La fonction $g(t)$ s'écrit alors :

$$g(t) = f((1-t)x_0 + tx(\theta))$$

et

$$g'(t) = A + 2Bt$$

où les expressions des constantes A et B sont les suivantes :

$$\begin{aligned} A &= \sum_{k=1}^{n+m}(x^k(\theta) - x_0^k)\alpha_k + 2\sum_{k=1}^{n+m}(x^k(\theta) - x_0^k)\beta_k x_0^k \\ &= \theta\alpha_i \text{ [car } x_0^k = 0 \text{ pour tout } k \leqslant n] \end{aligned}$$

$$\begin{aligned} B &= \sum_{k=1}^{n+m}(x_1^k - x_0^k)^2\beta_k \\ &= \sum_{k=1}^{n+m}(x^k(\theta) - x_0^k)\beta_k \\ &= \theta^2\beta_i. \end{aligned}$$

On voit alors que $g'(t) = \theta\alpha_i + 2\theta^2\beta_i t$ où i est la direction d'amélioration. Pour choisir i, on va étudier les variations de la fonction $\varphi(t) = \theta\alpha_i + \theta^2\beta_i$, car $g'(t) = \theta\alpha_i + \theta^2\beta_i t$ et donc $\displaystyle\int_0^1 g'(t)dt = g(1) - g(0) = \theta\alpha_i + \theta^2\beta_i$.

Remarque 2.3 *Un cas particulier simple est celui où tous les β_i sont nuls, et alors $\varphi(\theta) = \alpha_i\theta$. Dans ce cas, on applique la méthode du simplexe.*

2.2.3 Variations de la fonction φ .

On distingue deux cas.
1. $\alpha_i \geqslant 0$; $\beta_i > 0$ pour tout i.
Comme $\varphi'(\theta) = \alpha_i + 2\theta\beta_i$

le tableau de variations de la fonction $\varphi(\theta)$ est le suivant :

θ	$\frac{-\alpha_i}{\beta_i}$		$\frac{-\alpha_i}{2\beta_i}$		0
$\varphi'(\theta)$		$-$	0	$+$	
$\varphi(\theta)$	0	$\searrow$	$\varphi\left(\frac{-\alpha_i}{2\beta_i}\right)$	$\nearrow$	0

avec $\theta_0 = 0$ et $\theta_1 = \dfrac{-\alpha_i}{\beta_i}$; $\varphi(\theta_0) = \varphi(\theta_1) = 0$.

La plus grande valeur possible de θ donne le plus grand accroissement $g(1) - g(0)$ et par suite $x(\theta)$ est le meilleur sommet voisin de x_0 et, de plus, $g(1) - g(0) > 0$:

$\max\{g(1) - g(0)\} = \max\{f(x(\theta)) - f(x_0)\} = \max\{\varphi(\theta); \theta > 0\}$.

Soit i_1 une direction d'amélioration telle que $a_{ki_1} > 0$. Posons $\theta_{i_1} = \min\limits_{a_{ki_1}>0}\{\dfrac{b_k}{a_{ki_1}}\}$;

on passe ensuite à $i_2 \neq i_1$: $a_{ki} > 0$; $\theta_{i2} = \min\limits_{a_{ki_1}>0}\{\dfrac{b_k}{a_{ki_2}}\}, \ldots,$ on choisit $\theta = \max\limits_{j}\{\theta_{i_j}\}$.Alors $x(\theta)$ est le meilleur sommet voisin de x_0.

2. Il existe des $\alpha_i < 0$ ($\beta_i > 0$ pour tout i) :

Le tableau de variations de $\varphi(\theta)$ est alors le suivant :

θ	0		$\frac{-\alpha_i}{2\beta_i}$		$\frac{-\alpha_i}{\beta_i}$
$\varphi'(\theta)$		$-$	0	$+$	
$\varphi(\theta)$	0	$\searrow$	$\varphi\left(\frac{-\alpha_i}{2\beta_i}\right)$	$\nearrow$	0

où $\theta_0 = 0$ et $\theta_1 = \dfrac{-\alpha_i}{\beta_i}$; $\varphi(\theta_0) = \varphi(\theta_1) = 0$.

Soit $\theta = \max\limits_i \{\theta_{i_j}\}$ alors :

- si $\theta > \theta_1$ $\varphi(\theta) > 0$, et $x(\theta)$ est le meilleur sommet voisin de x_0 ;
- et si $\theta \leqslant \theta_1$ $\varphi(\theta) \leqslant 0$ x_0 est meilleur que $x(\theta)$.

2.2.4 L'accroissement de la fonction objectif.

Soit j_0 l'indice du vecteur entrant, et i_0 celui du vecteur sortant. D'après la ligne du pivot, on a :

$$x_{j_0} = \frac{b_{i_0}}{a_{i_0 j_0}} - \sum_{j \neq j_0} \frac{a_{i_0 j}}{a_{i_0 j_0}} x_j$$

Posons $\theta_0 = \dfrac{b_{i_0}}{a_{i_0 j_0}}$. On trouve donc :

$$\begin{aligned} x_{j_0} = \theta_0 - \sum_{j \neq j_0} \frac{1}{a_{i_0 j_0}} a_{i_0 j} x_j &= \theta_0 - \sum_{j \neq j_0} \frac{b_{i_0}}{b_{i_0} a_{i_0 j_0}} a_{i_0 j} x_j \\ &= \theta_0 - \theta_0 \sum_{j \neq j_0} \frac{1}{b_{i_0} a_{i_0 j_0}} a_{i_0 j} x_j \\ &= \theta_0 \left(1 - \sum_{j \neq j_0} \frac{1}{b_{i_0}} a_{i_0 j} x_j \right). \end{aligned}$$

Alors :

$$\begin{aligned} x_{j_0}^2 &= \theta_0^2 \left(1 - \sum_{j \neq j_0} \frac{1}{b_{i_0}} a_{i_0 j} x_j \right)^2 \\ &= \theta_0^2 + \frac{\theta_0^2}{b_{i_0}^2} \left(\sum_{j \neq j_0} a_{i_0 j} x_j \right)^2 - \frac{2\theta_0^2}{b_{i_0}} \sum_{j \neq j_0} a_{i_0 j} x_j \end{aligned}$$

Donc :

$\alpha_{j_0} x_{j_0} + \beta_{j_0} x_{j_0}^2 = \alpha_{j_0}\theta_0 - \dfrac{\alpha_{j_0}\theta_0}{b_{i_0}} \sum_{j \neq j_0} a_{i_0 j} x_j + \beta_{j_0}\theta_0^2 +$
$\beta_{j_0} \left(\dfrac{\theta_0}{b_{i_0}} \right)^2 \left(\sum_{j \neq j_0} a_{i_0 j} x_j \right)^2 - 2\dfrac{\beta_{j_0}\theta_0^2}{b_{i_0}} \sum_{j \neq j_0} a_{i_0 j} x_j$

D'où :

$\alpha_{j_0} x_{j_0} + \beta_{j_0} x_{j_0}^2 = (\alpha_{j_0}\theta_0 + \beta_{j_0}\theta_0^2) - \left(\dfrac{\alpha_{j_0}\theta_0}{b_{i_0}} + 2\dfrac{\beta_{j_0}\theta_0^2}{b_{i_0}} \right) \sum_{j \neq j_0} a_{i_0 j} x_j +$
$\beta_{j_0} \left(\dfrac{\theta_0}{b_{i_0}} \right)^2 \left(\sum_{j \neq j_0} a_{i_0 j} x_j \right)^2$

Alors la valeur de la fonction objectif devient :

$$\psi(x) = \alpha_{j_0} x_{j_0} + \beta_{j_0} x_{j_0}^2 + \sum_{j \neq j_0} \alpha_j x_j + \sum_{j \neq j_0} \beta_j x_j^2$$

On remplace $\left(\alpha_{j_0} x_{j_0} + \beta_{j_0} x_{j_0}^2 \right)$ par sa valeur, on trouve :

$$\begin{aligned} \psi(x) = \Delta_{j_0} - \left(\frac{\alpha_{j_0}\theta_0}{b_{i_0}} + 2\frac{\beta_{j_0}\theta_0^2}{b_{i_0}^2} \right) \sum_{j \neq j_0} a_{i_0 j} x_j + \\ \beta_{j_0} \left(\frac{\theta_0}{b_{i_0}} \right)^2 \left(\sum_{j \neq j_0} a_{i_0 j} x_j \right)^2 + \sum_{j \neq j_0} \alpha_j x_j + \sum_{j \neq j_0} \beta_j x_j^2 \\ = \Delta_{j_0} + \sum_{j \neq j_0} \left(\alpha_j - \left(\frac{\alpha_{j_0}\theta_0}{b_{i_0}} + 2\frac{\beta_{j_0}\theta_0^2}{b_{i_0}} \right) a_{i_0 j} \right) x_j + \\ \sum_{j \neq j_0} \beta_j x_j^2 + \beta_{j_0} \left(\frac{\theta_0}{b_{i_0}} \right)^2 \left(\sum_{j \neq j_0} a_{i_0 j} x_j \right)^2 \end{aligned}$$

Mais :

$$\begin{aligned}\left(\sum_{j\neq j_0} a_j x_j\right)^2 &= \sum_{j\neq j_0} a_{i_0 j}x_j \sum_{k\neq j_0} a_{j_0 k}x_k = \sum_{j\neq j_0}\sum_{k\neq j_0} a_{i_0 j}a_{i_0 k}x_j x_k \\ &= \sum_{k\neq j_0}(a_{i_0 k})^2 x_k^2 + 2\sum_{k\neq j_0, k\neq j, j\neq j_0}\sum a_{i_0 j}a_{i_0 k}x_j x_k\end{aligned}$$

D'où

$$\begin{aligned}\psi(x) &= \Delta_{j_0} + \sum_{j\neq j_0}\left(\alpha_j - \left(\frac{\alpha_{j_0}\theta_0}{b_{i_0}} + 2\frac{\beta_{j_0}\theta_0}{b_{i_0}}\right)a_{i_0 j}\right)x_j \\ &+ \sum_{j\neq j_0}\left(\beta_j + \beta_{j_0}\left(\frac{\theta_0}{b_{i_0}}\right)^2 (a_{i_0 j})^2\right)x_j^2 + 2\beta_{j_0}\left(\frac{\theta_0}{b_{i_0}}\right)^2 \sum_{k\neq j_0, k\neq j, j\neq j_0}\sum a_{i_0 j}a_{i_0 k}x_j x_k\end{aligned}$$

On obtient donc la forme des coefficients pour la première phase :

$$\alpha'_j = \alpha_j - (\alpha_{j_0} + 2\beta_{j_0}\theta_0)\frac{a_{i_0 j}}{a_{i_0 j_0}} \text{ si } j \neq j_0 \text{ et } \alpha'_{j_0} = 0$$

$$\beta'_j = \beta_j + \frac{\beta_{j_0}}{a_{i_0 j_0}^2}(a_{i_0 j})^2 \text{ si } j \neq j_0 \text{ et } \beta'_{j_0} = 0$$

$$\gamma'_{ij} = \beta_{j_0}\left(\frac{\theta_0}{b_{i_0}}\right)^2 a_{i_0 j}a_{i_0 i} \text{ pour } j \neq i, j \neq j_0, i \neq j_0, \text{ et } \gamma_{j_0 i} = \gamma_{i j_0} = \gamma_{ii} = 0, \forall i = 1, 2, ...$$

La forme de la fonction objectif devient alors :

$$\psi(x) = \Delta_{j_0} + \sum_{j\neq j_0}\alpha'_j x_j + \sum_{j\neq j_0}\beta'_j x_j^2 + \sum_j\sum_i \gamma_{ji}x_j x_i$$

L'expression $2\sum_{k\neq j_0, k\neq j, j\neq j_0}\sum a_{i_0 j}a_{i_0 k}x_j x_k$ avec $a_{i_0 j_0}a_{i_0 j} = 0 \ \forall j = 1, 2, ...$ peut être calculée de la manière suivante : Le produit cartésien de l'ensemble $\{a_{i_0 j}\}$ privé de la diagonale et de la ligne pivot.

	a_{i_01}	a_{i_02}			$a_{i_0j_0}$				$a_{i_0\ n+m}$
a_{i_01}	0	x_1x_2			0				x_1x_{n+m}
a_{i_02}	x_1x_2	0			0				x_2x_{n+m}
			0		0				
				0	0				
$a_{i_0j_0}$	0	0	0	0	0	0	0	0	0
					0	0			
					0		0		
					0			0	
					0				0
a_{i_0n+m}	$x_{n+m}x_1$				0				$x_{n+m}x_{n+m}$

L'expression $\sum_j \sum_i \gamma_{ji}x_jx_i$ est équivalente à l'expression suivante :

$$\gamma_{j_0j_0}x_{j_0}^2 + \sum_{i\neq j_0} \gamma_{j_0i}x_{j_0}x_i + \sum_{j\neq j_0} \gamma_{jj_0}x_jx_{j_0} + \sum \sum_{i\neq j_0, i\neq j_0} \gamma_{ji}x_jx_i$$

Comme $\gamma_{j_0j_0} = 0$ et l'expression $\sum\sum_{i\neq j_0, i\neq j_0} \gamma_{ji}x_jx_i$ ne contient pas x_{j_0} alors l'expression contenant x_{j_0} (vecteur entrant) est

$$\sum_{i\neq j_0} \gamma_{j_0i}x_{j_0}x_i + \sum_{j\neq j_0} \gamma_{jj_0}x_jx_{j_0}.$$

On obtient donc :

$$\sum_{i\neq j_0} \gamma_{j_0i}x_{j_0}x_i + \sum_{j\neq j_0} \gamma_{jj_0}x_jx_{j_0} = \frac{b_{i_0}}{a_{i_0j_0}} \sum_{j\neq j_0} \left(\gamma_{j_0j} + \gamma_{jj_0}\right) x_j$$
$$- \frac{1}{a_{i_0j_0}} \sum_{j\neq j_0} a_{i_0j_0} \left(\gamma_{j_0j} + \gamma_{jj_0}\right) x_j^2 - \frac{1}{a_{i_0j_0}} \sum_{i\neq j_0} \sum_{j\neq j_0, i\neq j} a_{i_0j}(\gamma_{ij_0} + \gamma_{j_0i})x_ix_j$$

La nouvelle forme de coefficients pour n'importe quelle phase est alors :

$$\alpha'_j = \alpha_j - (\alpha_{j_0} + 2\beta_{j_0}\theta_0)\frac{a_{i_0j}}{a_{i_0j_0}} + \frac{b_{i_0}}{a_{i_0j_0}}(\gamma_{j_0j} + \gamma_{jj_0}) \text{ si } j \neq j_0; \quad \alpha'_{j_0} = 0$$

$$\beta'_j = \beta_j + \frac{\beta_{j_0}}{(a_{i_0j_0})^2}(a_{i_0j})^2 - \frac{1}{a_{i_0j_0}} a_{i_0j}(\gamma_{j_0j} + \gamma_{jj_0}) \text{ si } j \neq j_0; \quad \beta'_{j_0} = 0$$

$$\gamma'_{ji} = \beta_{j_0}\left(\frac{\theta_0}{b_{i_0}}\right)^2 a_{i_0j}a_{i_0i} - \frac{a_{i_0j}}{a_{i_0j_0}}(\gamma_{j_0i} + \gamma_{ij_0}) \text{ pour } j \neq i, j \neq j_0, i \neq j_0;$$

$$\text{et} \quad \gamma_{j_0 i} = \gamma_{i j_0} = \gamma_{ii} = 0, \forall i = 1, 2, \dots$$

Les coefficients de la matrice des contraintes vont être changés de la même manière que ceux du simplexe :

- Si $i \neq i_0$
    1. $a'_{ij} = a_{ij} - \dfrac{a_{ij_0}}{a_{i_0 j_0}} a_{i_0 j}$
    2. $b'_i = b_i - \dfrac{a_{ij_0}}{a_{i_0 j_0}} b_{i_0}$
- Pour $i = i_0$
    1. $a'_{i_0 j} = \dfrac{a_{i_0 j}}{a_{i_0 j_0}}$
    2. $b'_{i_0} = \dfrac{b_{i_0}}{a_{i_0 j_0}}$

$a_{i_0 j_0}$ étant l'élément pivot).

2.2.5 Règle du plus grand gain marginal.

Pour chaque valeur possible de θ_j, on calcule le gain marginal Δ_j
$\Delta_j = \alpha_j \theta_j + \beta_j \theta_j^2$.
Posons $\Delta_{j_0} = \max\limits_j \{\alpha_j \theta_j + \beta_j \theta_j^2\} = \alpha_{j_0} \theta_{j_0} + \beta_{j_0} \theta_{j_0}^2$. Remarquons que, d'aprés les variations de la fonction φ, on a $\theta_{j_0} = \max\limits_j \{\theta_j\}$.

Ainsi, en passant du sommet x_0 à $x(\theta_{j_0})$, l'accroissement de la fonction est égal à $\alpha_{j_0} \theta_{j_0} + \beta_{j_0} \theta_{j_0}^2$. C'est donc le gain marginal de la fonction par rapport à sa j_0^{ieme} composante.

2.2.6 L'algorithme.

z représente la valeur de la fonction objectif initialisée à 0 ; γ_{ij} sont les coefficients des termes $x_i x_j$ initialisés à 0.

I. Dans le cas $\beta_i = 0$, pour tout $1 \leqslant i \leqslant n$, on applique l'algorithme du simplexe.

II. Si $\beta_i \geqslant 0$, pour tout $1 \leqslant i \leqslant n$, et α_i quelconque,alors deux cas sont possibles :

a) Si $a_{ki} < 0$ pour tout $1 \leqslant k \leqslant m$ et pour tout $1 \leqslant i \leqslant n$ alors on s'arrête : ce programme n'admet pas un optimum.

b) S'il existe k et i tels que $a_{ki} > 0$ on suit les étapes suivantes :

1. Calculer $\theta_k = \min\limits_{a_{kj}>0}\{\dfrac{b_k}{a_{kj}}\}$ pour tous les j t.q. $a_{kj} > 0$; choisir i_0 tel que $\theta_{i_0} = \dfrac{b_{i_0}}{a_{i_0 j}}$, x_{i_0} est alors le vecteur sortant.
2. Si $a_{kj} < 0$ pour tous les k, alors $\theta = \infty$.
3. Calculer $\Delta_j = \alpha_j \theta_j + \beta_j \theta_j^2$.
4. Choisir $\Delta_{j_0} = \max \Delta_j$.
5. Si $\Delta_{j_0} = \infty$ alors on s'arrête : ce programme n'admet pas un optimum.
6. Si $\Delta_{j_0} \leq 0$ alors ce programme est optimal et on s'arrête. Sinon on doit effectuer les opérations suivantes :
7. Poser $z = z + \Delta_{j_0}$ (j_0 est l'indice du vecteur entrant).
8. Remplacer x_{i_0} par x_{j_0}, dans la base et prendre $a_{i_0 j_0}$ comme pivot.
9. Calculer $\alpha'_j = \alpha_j - (\alpha_{j_0} + 2\beta_{j_0}\theta_0)\dfrac{a_{i_0 j}}{a_{i_0 j_0}} + \dfrac{b_{i_0}}{a_{i_0 j_0}}(\gamma_{j_0 j} + \gamma_{j j_0})$
10. Poser $\alpha'_{j_0} = 0$
11. Calculer $\beta'_j = \beta_j + \dfrac{\beta_{j_0}}{(a_{i_0 j_0})^2}(a_{i_0 j})^2 - \dfrac{1}{a_{i_0 j_0}} a_{i_0 k}(\gamma_{j_0 j} + \gamma_{j j_0})$
12. Poser $\beta'_{j_0} = 0$
13. Calculer $\gamma'_{ji} = \beta_j \left(\dfrac{\theta_0}{b_{i_0}}\right)^2 a_{i_0 j} a_{i_0 i} - \dfrac{a_{i_0 j}}{a_{i_0 j_0}}(\gamma_{j_0 j} + \gamma_{j j_0})$
14. Poser $\gamma_{j_0 i} = \gamma_{i j_0} = \gamma_{ii} = 0$ pour i
15. Calculer $a'_{ij} = a_{ij} - \dfrac{a_{i j_0}}{a_{i_0 j_0}} a_{i_0 j}$ si $i \neq i_0$.
16. Calculer $b'_i = b_i - \dfrac{a_{i j_0}}{a_{i_0 j_0}} b_{i_0}$ si $i \neq i_0$.

17. Calculer $a'_{i_0 j} = \dfrac{a_{i_0 j}}{a_{i_0 j_0}}$

18. Calculer $b'_{i_0} = \dfrac{b_{i_0}}{a_{i_0 j_0}}$.

Exemple 2.1

$$\max z; z = 2x_1 + x_2 + 2x_3 + x_2^2 + x_3^2$$

$$scq : \begin{cases} x_1 + 2x_2 + 4x_3 + x_4 = 4 \\ 3x_1 + x_2 + 5x_3 + x_5 = 9 \\ x_1 + 3x_2 + x_3 + x_6 \ = 2 \\ x \geqslant 0 \end{cases}$$

$$z = 2x_1 + x_2 + 2x_3 + x_2^2 + x_3^2$$

2	1	2	0	0	0	α
0	1	1	0	0	0	β
2	2/3	1				θ
4	10/9	1				Δ
1	2	4	1	0	0	$x_4 = 4$
3	1	5	0	1	0	$x_5 = 9$
1	3	1	0	0	1	$x_6 = 2$

$$z = -5x_2 - 2x_6 + x_2^2 + x_3^2 + 4$$

0	-5	0	0	0	-2	α
0	1	1	0	0	0	β
	2/3	2/3				θ
	$-26/9$	4/9				Δ
0	-1	3	1	0	-1	$x_4 = 2$
0	-8	2	0	1	-3	$x_5 = 3$
1	3	1	0	0	1	$x_1 = 2$

$$z = -\frac{41}{9}x_2 - \frac{4}{9}x_4 - \frac{14}{9}x_6 + \frac{10}{9}x_2^2 + \frac{1}{9}x_4^2 + \frac{1}{9}x_6^2 - \frac{4}{9}x_2x_4 + \frac{2}{9}x_2x_6 - \frac{2}{9}x_4x_6 + \frac{40}{9}$$

0	-41/9	0	$-4/9$	0	$-14/9$	α
0	10/9	0	$1/9$	0	$1/9$	β
	2/5		2		1	θ
	$-74/45$		-0.44		$-13/9$	Δ
0	$-1/3$	1	$1/3$	0	-13	$x_3 = 2/3$
0	$-22/3$	0	$-2/3$	1	$-7/3$	$x_5 = -5/3$
1	$10/3$	0	$-1/3$	0	$4/3$	$x_1 = 4/3$

Tous les Δ_j sont negatifs donc la solution optimale est la précédente.

Exemple 2.2

$$\max z; z = x_1 - x_2 + 4x_3 + x_1^2 + x_2^2 + x_3^2$$

$$scq: \begin{cases} x_1 + 2x_2 - 3x_3 + x_4 = 12 \\ 2x_1 + 3x_2 - x_3 + x_5 = 8 \\ 3x_1 + x_2 - 3x_3 + x_6 = 15 \\ x \geqslant 0 \end{cases}$$

$$z = 2x_1 + x_2 + 2x_3 + x_2^2 + x_3^2$$

1	-1	4	0	0	0	α
1	1	1	0	0	0	β
4	8/3	$+\infty$				θ
		$+\infty$				Δ
1	2	-3	1	0	0	$x_4 = 12$
2	3	-1	0	1	0	$x_5 = 8$
3	1	-3	0	0	1	$x_6 = 15$

La fonction n'admet aucun maximum

Exemple 2.3

$$\max z; z = 3x_1 - 2x_2 + x_1^2 + x_2^2$$

$$scq: \begin{cases} x_1 - x_2 + x_3 \quad = 2 \\ 2x_1 + x_2 + x_4 \quad = 4 \\ -3x_1 + 2x_2 + x_5 = 6 \\ x \geqslant 0 \end{cases}$$

$$z = 2x_1 + x_2 + 2x_3 + x_2^2 + x_3^2$$

3	-2	0	0	0	α
1	1	0	0	0	β
2	2				θ
10	0				Δ
1	1	1			$x_3 = 2$
2	1	0	1	0	$x_4 = 4$
-3	2	0	0	1	$x_5 = 6$

$$z = -9x_2 - 7x_3 + 2x_2^2 + 2x_2x_3 + 10$$

0	-9	-7	0	0	α
0	2	1	0	0	β
	2	2			θ
	-10	-10			Δ
1	1	1	0	0	$x_1 = 2$
0	-1	-2	1	0	$x_4 = 0$
0	5	3	0	1	$x_5 = 12$

Chapitre 3

Méthodes de programmation non linéaire

Si certains problèmes particuliers d'optimisation non linéaire peuvent se ramener à la résolution d'une suite de problèmes linéaires (comme les problèmes séparables) en général,il faudra prendre en compte explicitement le caractère non linéaire des problèmes pour les résoudre.

Beaucoup d'algorithmes de programmation non linéaire exploitent directement les conditions d'optimalité, ces conditios sont connues : annulation de la dérivée première pour une fonction d'une variable dans $\mathbb{R}$ (l'annulation du gradient dans $\mathbb{R}^n$).

Avant de passer aux méthodes de la programmation non linéaire,soulignons quelques différences fondamentales avec la programmation linéaire.
En programmation linéaire la solution optimale est toujours obtenue sur la frontière du domaine réalisable et même toute une arête du polygone peut être optimale, on peut toujours choisir un sommet de l'arête.
Autrement dit , la solution optimale est toujours en un sommet de la région réalisable.
Dés que l'objectif ou les contraintes deviennent non linéaire, cette propriété n'est plus vérifiée en général. Pire encore, on peut avoir un optimum local différent de l'optimum global. ce problème ne se produit cependant pas pour les problèmes convexes.
les modèles non linéaires sont beaucoup plus difficiles à résoudre en générale que les modèles linéaire.

Dans ce chapitre, nous allons introduire la méthode des points de selles et multiplicateurs de Lagrange qui permettront de résoudre pas mal de problèmes d'optimisation .

3.1 Forme générale d'un programme non linéaire

Considérons le problème non linéaire de la production. Notons

– m matières premières $M_1, M_2, ..., M_m$
– n produits $P_1, P_2, ..., P_n$
– x_j le nombre d'unités vendues du produit P_j $(j = 1, 2, ..., n)$
– $p_j(x_1, x_2, ..., x_n)$ le prix unitaire du produit P_j en fonction des quantités vendues de tous les produits
– k_j : coût fixe par unité du produit P_j
– a_{ij} : nombre d'unités de matière première M_i nécessaire pour la production d'une unité de P_j $(i = 1, 2, ..., m; j = 1, 2, ..., n)$
– b_i : nombre d'unités disponibles de la matière première M_i

Le but est de maximiser la recette réalisée par la vente de ces produits en tenant compte de la disponibilité des matières premières :

$$\text{maximiser } f(x_1, x_2, ..., x_n) = \sum_{j=1}^{n} (p_j(x_1, x_2, ..., x_n)x_j - k_j x_j)$$

$$\text{sous les contraintes } \sum_{j=1}^{n} a_{ij}x_j - b_i \leqslant 0 \, (i = 1, 2, ..., m) \, x_1, x_2, ..., x_n \geqslant 0$$

Exemple 3.1 $m = n = 2,\ p_j(x_1, x_2) = 7 - x_j,\ k_j = 2\ (j = 1, 2),$

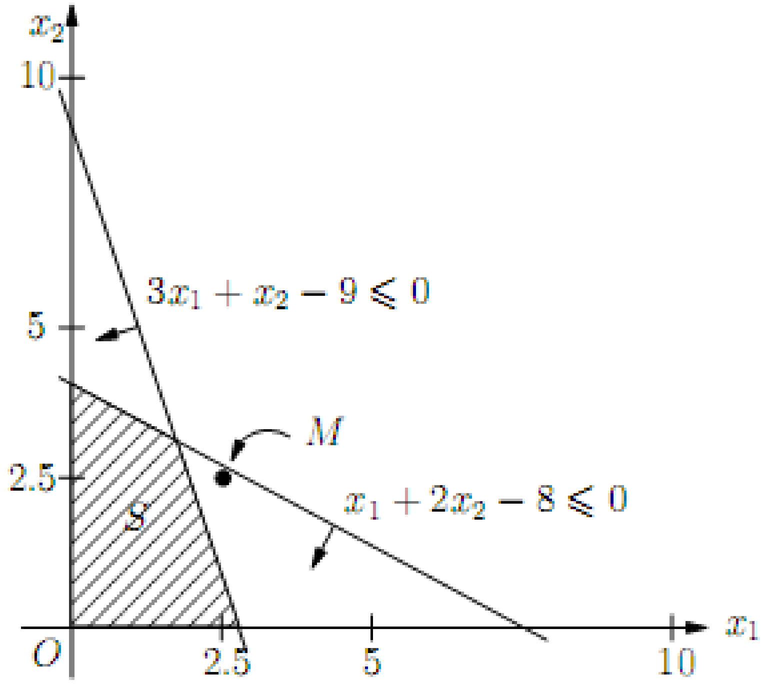
$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = \begin{pmatrix} 1 & 2 \\ 3 & 1 \end{pmatrix}, \quad \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} = \begin{pmatrix} 8 \\ 9 \end{pmatrix}.$$

Il s'agit alors de maximiser :

$$f(x_1, x_2) = (7 - x_1)x_1 + (7 - x_2)x_2 - 2x_1 - 2x_2 = 5x_1 + 5x_2 - x_1^2 - x_2^2$$

sous les contraintes :

$$\begin{cases} x_1 + 2x_2 - 8 \leqslant 0 \\ 3x_1 + x_2 - 9 \leqslant 0 \\ \qquad\quad x_1, x_2 \geqslant 0 \end{cases}$$



L'ensemble des solutions réalisables Ω correspond au domaine hachuré ci-dessus. f étant quadratique,
les ensembles $E_c = \{(x_1, x_2) : f(x_1, x_2) = c\}$ ne sont plus des droites comme en programmation linéaires, mais des cercles, parce que

$$5x_1 + 5x_2 - x_1^2 - x_2^2 = -(x_1 - 2.5)^2 - (x_2 - 2.5)^2 + 12.5.$$

$f(x_1, x_2) = c$ est donc l'équation d'un cercle centré en $M = (2.5, 2.5)$ et de rayon qui dépend de c.
Remarquons que M ne fait pas partie de Ω. La solution optimale est donc le point de Ω le plus proche de M, c'est-à-dire la projection orthogonale de M sur Ω. Il s'agit du point $P = (2.2, 2.4)$. On le trouve en construisant la droite qui est perpendiculaire à la droite $3x_1 + x_2 - 9 = 0$ et qui passe par M. P est alors le point d'intersection des deux droites.
La solution optimale est donnée par $x_1 = 2.2$ et $x_2 = 2.4$.

En termes généraux, un programme non linéaire se présente sous une des deux formes suivantes :

$(P_{\min})$ minimiser $f(x)$ sous les contraintes $\quad g_i(x) \leqslant 0 \quad (i = 1, 2, ..., m), x \in \mathrm{IR}_+^n$

$(P_{\max})$ maximiser$f(x)$ sous les contraintes $\quad g_i(x) \leqslant 0 \quad (i = 1, 2, ..., m), x \in \mathbb{R}^n_+$

Le but est de développer une méthode analytique qui permet de résoudre $(P_{\min})$ et $(P_{\max})$ pour une classe raisonnable de fonctions f et g_i en toutes dimensions n de x et quel que soit le nombre de contraintes.
Notons $\Omega = \{x : g_i(x) \leqslant 0$ pour $i = 1, 2, ..., m\}$ l'ensemble des solutions réalisables.

3.2 Fonction de Lagrange

Définition 3.1 *La fonction de Lagrange associée à* $(P_{\min})$ *resp.* $(P_{\max})$ *est donnée par*

$$L(x, \lambda) = f(x) + \sum_{i=1}^{m} \lambda_i g_i(x) \quad \textit{resp. } L'(x, \lambda) = f(x) - \sum_{i=1}^{m} \lambda_i g_i(x)$$

où $\lambda_1, \lambda_2, ..., \lambda_m \geqslant 0$ *sont les multiplicateurs de Lagrange.*

Remarque 3.1 *Maximiser* $f(x)$ *revient à minimiser* $-f(x)$. *Donc le lien entre* L *et* L' *est le suivant :*

$$L_{-f(x,\lambda)} = -f(x) + \sum_{i=1}^{m} \lambda_i g_i(x) = -((f(x) - \sum_{i=1} \lambda_i g_i(x)) = -L'_f(x, \lambda).$$

Nous allons voir que résoudre $(P_{\min})$ *où* $(P_{\max})$ *revient à trouver des points de selles de la fonction de Lagrange.*

Définition 3.2 $(x^0, \lambda^0) = (x^0_1, x^0_2, ..., x^0_n, \lambda^0_1, \lambda^0_2, ..., \lambda^0_m) \in \mathbb{R}^n \times \mathbb{R}^m_+$ *est un point de selle de* L *si* $L(x^0, \lambda) \leqslant L(x^0, \lambda^0) \leqslant L(x, \lambda^0)$ *pour tous* $x \in \mathbb{R}^n$ *et tous* $\lambda \in \mathbb{R}^m_+$.
En particulier : $L(x^0, \lambda^0) = \max_{\lambda} L(x^0, \lambda) = \min_{x} L(x, \lambda^0)$.

Remarque 3.2 *La définition ci-dessus signifie essentiellement qu'au point* (x^0, λ^0), L *est localement convexe en* x *et localement concave en* λ.
La définition analogue pour L' *se lit* $L'(x, \lambda^0) \leqslant L'(x^0, \lambda^0) \leqslant L'(x^0, \lambda)$ *pour tous* $x \in \mathbb{R}^n$ *et* $\lambda \in \mathbb{R}^m_+$.

Théorème 3.1 *Soit* $(x^0, \lambda^0) \in \mathbb{R}^n \times \mathbb{R}^m_+$ *un point de selle de* L *(resp.* L'*). Si* $x^0 \in \Omega$, x^0 *est solution optimale de* $(P_{\min})$, *resp. de* $(P_{\max})$.

Preuve : Soit (x^0, λ^0) un point de selle de L.
Il faut démontrer que : $f(x^0) \leqslant f(x)$ pour tous $x \in \Omega$.
Soit : $g = (g_1, g_2, ..., g_m)$.

$a)$ D'après la définition, on a : $L(x^0, \lambda) \leqslant L(x^0, \lambda^0)$.
Donc $f(x^0) + \lambda g(x^0) \leqslant f(x^0) + \lambda^0 g(x^0)$,
et par suite $\lambda g(x^0) \leqslant \lambda^0 g(x^0)$.
Comme $x^0 \in \Omega$ on a $g(x^0) \leqslant 0$; d'où $\lambda^0 g(x^0) \leqslant 0$.
En posant $\lambda = 0$ on obtient aussi $\lambda^0 g(x^0) \geqslant 0$. Donc $\lambda^0 g(x^0) = 0$.

$b)$ D'après la définition, on a aussi $L(x^0, \lambda^0) \leqslant L(x, \lambda^0)$.
Donc $f(x^0) = f(x^0) + \lambda^0 g(x^0) \leqslant f(x) + \lambda^0 g(x) \leqslant f(x)$ pour tous $x \in \Omega$, car $\lambda^0 g(x) \leqslant 0$ pour $x \in \Omega$.

La démonstration est analogue pour L'.

Il reste la question de comment trouver le (ou les) point de selle de L, resp. de L'.

3.3 Théorie de Kuhn-Tucker

Théorème 3.2 ***(Kuhn, Tucker)***
a) Les conditions suivantes sont nécessaires pour que (x, λ) soit un point de selle de L (resp. de L') dans $\Omega \times \mathrm{I\!R}^m_+$:

$$(1) \qquad \begin{cases} \dfrac{\partial}{\partial x} L(x, \lambda) \geqslant 0 \\ resp, \ \dfrac{\partial}{\partial x} L'(x, \lambda) \leqslant 0. \end{cases}$$

C'est-à-dire :

$$\begin{cases} \dfrac{\partial}{\partial x_j} f(x) + \displaystyle\sum_{i=1}^{m} \lambda_i \dfrac{\partial}{\partial x_j} g_i(x, \lambda) \geqslant 0 \\ \dfrac{\partial}{\partial x_j} f(x) - \displaystyle\sum_{i=1}^{m} \lambda_i \dfrac{\partial}{\partial x_j} g_i(x, \lambda) \leqslant 0. \end{cases}$$

(2)
$$\begin{cases} \dfrac{\partial}{\partial \lambda} L(x,\lambda) \leqslant 0 \\ resp,\ \dfrac{\partial}{\partial \lambda} L'(x,\lambda) \geqslant 0. \end{cases}$$

C'est-à-dire :

$$g_i(x) \leqslant 0, \quad i = 1, 2, \ldots, m$$

(3)
$$\begin{cases} {}^t x . \dfrac{\partial}{\partial x} L(x,\lambda) = 0 \\ resp,\ {}^t x . \dfrac{\partial}{\partial x} L'(x,\lambda) = 0. \end{cases}$$

C'est-à-dire :

$$\begin{cases} \displaystyle\sum_{j=1}^{n} x_j \left(\frac{\partial}{\partial x_j} f(x) + \sum_{i=1}^{m} \lambda_i \frac{\partial}{\partial x_j} g_i(x) \right) = 0 \\ \displaystyle\sum_{j=1}^{n} x_j \left(\frac{\partial}{\partial x_j} f(x) - \sum_{i=1}^{m} \lambda_i \frac{\partial}{\partial x_j} g_i(x) \right) = 0. \end{cases}$$

(4)
$$\begin{cases} \lambda \dfrac{\partial}{\partial \lambda} L(x,\lambda) = 0 \\ resp,\ \lambda \dfrac{\partial}{\partial \lambda} L'(x,\lambda) = 0. \end{cases}$$

C'est-à-dire :

$$\sum_{i=1}^{m} \lambda_i g_i(x) = 0$$

(5) $x \geqslant 0, \lambda \geqslant 0$

b) Si f est convexe (resp. concave) et si g_i est convexe pour $i = 1, 2, ..., m$, les conditions (1) à (5) en partie a) sont nécessaires et suffisantes pour que (x, λ) soit un point de selle de L (resp. L').

Rappel :

$f : \mathbb{R}^n \longrightarrow \mathbb{R}$ est convexe (resp. concave) si

$$\begin{cases} f(\lambda x + (1-\lambda) y) \leqslant \lambda f(x) + (1-\lambda) f(y) \\ resp. f(\lambda x + (1-\lambda) y) \geqslant \lambda f(x) + (1-\lambda) f(y) \end{cases}$$

pour tous $x, y \in \mathbb{R}^n$ et tous $\lambda \in [0, 1]$.

Remarque 3.3 *1) Le théorème (3.1) assure que x dans $(x, \lambda) \in \Omega \times \mathrm{I\!R}_+^m$ en partie b) du théoreme (3.2) est solution optimale de (P_{min}) resp. de (P_{max}). Les conditions sur f et g_i impliquent que L est convexe et que L' est concave en x.*
2) Pour être tout à fait exact il faudrait ajouter en partie a) la condition dite de Slater qui est la suivante :

$$\text{il existe } x \in \mathrm{I\!R}^n \text{ tel que } g_i(x) < 0 \text{ pour } i = 1, 2, ..., m.$$

Cette condition assure alors l'existence d'un λ tel que (x, λ) est point de selle de L resp. de L'.

Exemple 3.2 *L'exemple suivant montre que cette condition n'est pas superflue :*

$$\begin{gathered}\text{minimiser } f(x,y) \text{ (} f \text{ n'est pas spécifiée davantage)} \\ \text{sous la contrainte } x^2 + y^2 = 0 \ (\Leftrightarrow x^2 + y^2 \leqslant 0).\end{gathered}$$

La solution optimale est évidemment $(x, y) = (0, 0)$ et

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

donc :

$$\frac{\partial}{\partial x} L(x, y, \lambda) \mid_{(0,0)} = \frac{\partial}{\partial x} f(x, y) \mid_{(0,0)} \geqslant 0$$

par la condition 1) de la partie a)du théorème (3.2). Il n'existe donc pas, en général, $\lambda \geqslant 0$ tel que cette condition est satisfaite. la condition de Slater n'est pas vérifiée.

Exemple 3.3 *Problème nonlinéaire de la production (suite) :*

$$L'(x_1, x_2, \lambda_1, \lambda_2) = 5x_1 + 5x_2 - x_1^2 - x_2^2 - \lambda_1(x_1 + 2x_2 - 8) - \lambda_2(3x_1 + x_2 - 9).$$

Les conditions du théorème de Kuhn-Tucker s'écrivent alors :

$$(1) \qquad \begin{cases} \dfrac{\partial}{\partial x_1} L'(x_1, x_2, \lambda_1, \lambda_2) = 5 - 2x_1 - \lambda_1 - 3\lambda_2 \leqslant 0 \\ \dfrac{\partial}{\partial x_2} L'(x_1, x_2, \lambda_1, \lambda_2) = 5 - 2x_2 - 2\lambda_1 - \lambda_2 \leqslant 0 \end{cases}$$

$$(2) \qquad x_1 + 2x_2 \leqslant 8, 3x_1 + x_2 \leqslant 9$$

$$(3) \qquad x_1(5 - 2x_1 - \lambda_1 - 3\lambda_2) + x_2(5 - 2x_2 - 2\lambda_1 - \lambda_2) = 0$$

$$(4) \qquad \lambda_1(x_1 + 2x_2 - 8) + \lambda_2(3x_1 + x_2 - 9) = 0$$

$$(5) \qquad x_1, x_2, \lambda_1, \lambda_2 \geqslant 0$$

On remarque que, dans (3) et (4), dés que le premier facteur des produits est positif, le second doit s'annuler. De même, dés que le deuxième est négatif, le premier doit s'annuler. Ceci nous fait penser de chercher une solution qui satisfait à (1) avec l'inégalité remplacée par une égalité. Dans cet exemple, on trouve la solution effectivement de cette manière.
1^{er} *cas* : $(\lambda_1 = 0, \lambda_2 \neq 0)$.
On obtient le système d'équations :

$$\begin{cases} 5 - 2x_1 - 3\lambda_2 = 0 \\ 5 - 2x_2 - \lambda_2 \ = 0 \\ 3x_1 + x_2 - 9 \ = 0. \end{cases}$$

Il s'ensuit que :

$$\begin{cases} x_1 = 2.2 \\ x_2 = 2.4 \\ \lambda_2 = 0.2 \\ \lambda_1 = 0 \end{cases}$$

Il s'agit de la solution trouvée auparavant.
Pour une étude systématique, il faut distinguer les cas :
a) $x_1 \neq 0$ *et* $x_2 \neq 0$ *(égalités en (1), cas ci-dessus)*
b) $x_1 = 0$ *et* $x_2 \neq 0$ $(3x_1 + x_2 = 9 \Rightarrow x_2 = 9, 5 - 2x_2 - \lambda_2 = 0 \Rightarrow \lambda_2 < 0)$
c) $x_1 \neq 0$ *et* $x_2 = 0$ $(3x_1 + x_2 = 9 \Rightarrow x_1 = 3, 5 - 2x_1 - 3\lambda_2 = 0 \Rightarrow \lambda_2 < 0)$
2^{eme} *cas* : $(\lambda_1 \neq 0, \lambda_2 = 0)$.

$$\begin{cases} 5 - 2x_1 - \lambda_1 &= 0 \\ 5 - 2x_2 - 2\lambda_1 &= 0 \\ x_1 + 2x_2 - 8 &= 0 \end{cases}$$

L'étude systématique analogue à celle du premièr cas ne donne pas de solutions avec $\lambda_1, \lambda_2 \geqslant 0$.

<u>3^{eme} *cas*</u> : $(\lambda_1 \neq 0, \lambda_2 \neq 0)$.

Par (4) : $\begin{cases} x_1 + 2x_2 = 8 \\ 3x_1 + x_2 = 9 \end{cases} \Longrightarrow \begin{cases} x_1 = 2 \\ x_2 = 3 \end{cases}$

Par (3) : $\begin{cases} 1 - \lambda_1 - 3\lambda_2 = 0 \\ -1 - 2\lambda_1 - \lambda_2 = 0 \end{cases} \Longrightarrow \begin{cases} \lambda_1 = -4/5 \\ \lambda_2 = 3/5 \end{cases}$

On n'obtient donc pas de solutions avec $\lambda_1, \lambda_2 \geqslant 0$.
Pour démontrer que $(x_1, x_2) = (2.2, 2.4)$ *est solution optimale, il suffit de vérifier que* f *est concave et d'appliquer la partie b) du Théorème (3.2).* g_1 *et* g_2 *sont linéaires, donc convexes.*

Remarque 3.4 *Considérons le cas où certaines contraintes sont des égalités. Pour* (P_{min}) *considérons le programme : minimiser* $f(x)$ $(x \in \mathbb{R}^n)$
sous les contraintes $\quad g_i(x) \leqslant 0 \quad (i = 1, ..., m_1), \quad h_i(x) = 0 \quad (i = 1, ..., m_2)$
Notons que :

$$h_i(x) = 0 \Longleftrightarrow h_i(x) \leqslant 0 \quad \text{et} \quad -h_i(x) \leqslant 0 \quad (i = 1, ..., m_2).$$

La fonction de Lagrange est donc donnée par

$$\begin{aligned} L(x, \lambda) &= f(x) + \sum_{i=1}^{m_1} \lambda_i g_i(x) + \sum_{i=1}^{m_2} \lambda'_i h_i(x) - \sum_{i=1}^{m_2} \lambda''_i h_i(x) \\ &= f(x) + \sum_{i=1}^{m_1} \lambda_i g_i(x) + \sum_{i=1}^{m_2} \mu_i h_i(x), \end{aligned}$$

où $\lambda_i \geqslant 0 \quad (i = 1, ..., m_1) \quad$ *et* $\quad \mu_i = \lambda'_i - \lambda''_i \in \mathbb{R} \quad (i = 1, ..., m_2)$.
On a donc une fonction de Lagrange du type $L(x, \lambda, \mu)$ *avec des multiplicateurs de Lagrange* $\lambda_i \geqslant 0$ *et* $\mu_j \in \mathbb{R}$ $(i = 1, ..., m_1, \quad j = 1, ..., m_2)$.

3.4 Programmation quadratique

Le but de cette section est d'étudier le cas où la fonction f est quadratique et $g_1, g_2, ..., g_m$ sont linéaires et de montrer que la théorie de Kuhn-Tucker aboutit, dans ce cas particulier, à l'algorithme du simplexe appliqué à un programme auxiliaire, voir [18].

Exemple 3.4 $\big($*Optimisation de portefeuilles en finance*$\big)$.
Considérons deux valeurs financières (action, obligation,...) dont le rendement n'est pas fixe, mais soumis à des fluctuations non prévisibles. De façon plus précise, soit :
• r_i *la rentabilité annuelle du titre* i $(i = 1, 2)$; r_i *est une variable aléatoire,*
• $\mu_i = Er_i$ $\big($*l'espérance de* $r_i\big)$,
• $\sigma_i^2 = Var\ r_i$ $\big($*la variance de* $r_i\big)$,
• $\sigma_{12} = Cov(r_1, r_2)$ $\big($*la covariance de* r_1 *et* $r_2\big)$.
μ_1, μ_2, σ_1^2, σ_2^2 *et* σ_{12} *sont supposés connus.*
Le but dans cet exemple est de trouver un portefeuille composé de ces deux titres, tel que son rendement est optimal en un sens que nous précisons maintenant,voir [22] . Soit $P(x_1, x_2)$ *le portefeuille composé de* x_1 *parts du titre 1 et de* x_2 *parts du titre 2. On a donc* $x_1 + x_2 = 1$.
Soit :
• $r_p = x_1r_1 + x_2r_2$ $\big($*la rentabilité annuelle de* $P(x_1, x_2)\big)$,
• $\mu_p = Er_p = x_1\mu_1 + x_2\mu_2$ $\big($*l'espérence de* $r_p\big)$,
• $\sigma_p^2 = Varr_p = Var(x_1r_1 + x_2r_2)$
$= x_1^2\sigma_1^2 + x_2^2\sigma_2^2 + 2x_1x_2\sigma_{12}$ $\big($*la variance de* $r_p\big)$, *voir [4], et [10].*
La composition optimale d'un portefeuille est alors caractérisée par le fait que le risque, mesuré par σ_p^2, *est minimal, mais qu'un bénéfice moyen, supposé donné et noté* μ_0, *est garanti. Le portefeuille optimal est alors solution du programme quadratique suivant :*
minimiser $f(x_1, x_2) = \sigma_p^2(x_1, x_2) = x_1^2\sigma_1^2 + x_2^2\sigma_2^2 + 2x_1x_2\sigma_{12}$

$$\text{sous les contraintes :} \begin{cases} \mu_p(x_1, x_2) & = & x_1\mu_1 + x_2\mu_2 \\ x_1 + x_2 & = & 1 \\ x_1, x_2 & \geqslant & 0 \end{cases}$$

Remarque 3.5 *On supprime parfois la contrainte* $x_1, x_2 \geqslant 0$ *et on admet alors une partie négative d'un titre. Ceci veut dire qu'on fait une vente à*

découvert.
On passe maintenant à la forme générale d'un programme quadratique. La fonction f devient alors une forme quadratique.
$P_{(min)}$ *minimiser* $f(x) = cx + \frac{1}{2}\ {}^t xQx$
sous les contraintes : $Ax - b \leqslant 0$, $x \in \mathbb{R}^n_+$
$P_{(max)}$ *maximiser* $f(x) = cx - \frac{1}{2}\ {}^t xQx$
sous les contraintes : $Ax - b \leqslant 0$, $x \in \mathbb{R}^n_+$
Ici $c = (c_1, c_2, ..., c_n) \in \mathbb{R}^n, Q = (q_{ij};\ i,j = 1,2,...,n)$ *est une matrice symétrique,* $A = (a_{ij};\ i = 1,2,...,m, j = 1,2,...,n)$ *est une matrice* $m \times n$ *et* $b = {}^t(b_1, b_2, ..., b_m) \in \mathbb{R}^m$.
Notre but étant de résoudre ces programmes quadratiques à l'aide de la partie b) du théorème de Kuhn-Tucker, nous supposons en plus que Q est définie positive. En effet la proposition I.3 ci-dessous confirme que la fonction f, et donc la fonction de Lagrange, possède la propriété de convexité resp. de concavité requise dans la partie b) du théorème (3.2).

Rappel.
Q est semmi-définie positive si, pour tous $x = (x_1, x_2, ..., x_n) \in \mathbb{R}^n$,
${}^t xQx = \displaystyle\sum_{i=1}^{n}\sum_{j=1}^{n} q_{ij}x_i x_j \geqslant 0$. Ceci revient à dire que le déterminant de Q ainsi que tous les sous-déterminants sont non-négatifs.

Exemple 3.5 *Problème non linéaire de la production (suite).*

$$\begin{aligned} f(x_1, x_2) &= 5x_1 + 5x_2 - x_1^2 - x_2^2 \\ &= cx - \tfrac{1}{2}{}^t xQx \end{aligned}$$

où

$$x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \quad c = (5,5), \quad Q = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}.$$

Q est définie positive parce que ${}^t xQx = 2(x_1^2 + x_2^2) > 0$ *si* $x \neq 0$.

Proposition 3.1 *[18] La fonction* $\left\{ \begin{array}{l} f(x) = cx + \frac{1}{2}\ {}^t xQx \\ f(x) = cx - \frac{1}{2}\ {}^t xQx \end{array} \right\}$ *est* $\left\{ \begin{array}{c} \textit{convexe} \\ \textit{concave} \end{array} \right\}$ *si et seulement si Q est symétrique et semi-définie positive.*

La preuve se fait par un calcul simple (admis) qui vérifie directement les inégalités de la définition de la convexité et de la concavité, voir [18].
Nous évaluons maintenant le théorème (3.2) pour un programme quadratique général. La fonction de Lagrange est donnée par :

$$\begin{aligned} L(x,\lambda) &= f(x) + \lambda g(x) = cx + \tfrac{1}{2}\,{}^t xQx + \lambda(Ax-b) \\ L'(x,\lambda) &= f(x) - \lambda g(x) = cx - \tfrac{1}{2}\,{}^t xQx + \lambda(Ax-b) \end{aligned}$$

Les conditions (1) à (5) de la partie a) du théorème (3.2) se lisent :

(1)
$$\begin{aligned} &\frac{\partial}{\partial x_j}L(x,\lambda) = c_j + (Qx)_j + (\lambda A)_j \geqslant 0,\ j=1,2,...,n \\ &\qquad \Longleftrightarrow {}^t c + Qx + {}^t(\lambda A) \geqslant 0 \\ &\frac{\partial}{\partial x_j}L'(x,\lambda) = c_j - (Qx)_j - (\lambda A)_j \leqslant 0,\ j=1,2,...,n \\ &\qquad \Longleftrightarrow {}^t c - Qx - {}^t(\lambda A) \leqslant 0 \end{aligned}$$

(2)
$$\left.\begin{aligned} &\frac{\partial}{\partial \lambda_i}L(x,\lambda) \leqslant 0, \quad \text{i=1,2,...,m} \\ &\frac{\partial}{\partial \lambda_i}L'(x,\lambda) \geqslant 0, \quad \text{i=1,2,...m} \end{aligned}\right\} \Longleftrightarrow Ax - b \leqslant 0$$

(3)
$$\begin{aligned} &\sum_{j=1}^{n}\frac{\partial}{\partial x_j}L(x,\lambda) = 0 \Longleftrightarrow {}^t x.({}^t c + Qx + {}^{(}\lambda A)) = 0 \\ &\sum_{j=1}^{n}\frac{\partial}{\partial x_j}L'(x,\lambda) = 0 \Longleftrightarrow x.(c - Qx - (\lambda A)) = 0 \end{aligned}$$

(4)
$$\left.\begin{aligned} &\sum_{i=1}^{m}\lambda_i\frac{\partial}{\partial \lambda_i}L(x,\lambda) = 0 \\ &\sum_{i=1}^{m}\lambda_i\frac{\partial}{\partial \lambda_i}L'(x,\lambda) = 0 \end{aligned}\right\} \Longleftrightarrow \lambda.(Ax-b) = 0$$

(5) $x \in \mathbb{R}_+^n, \lambda \in \mathbb{R}_+^m$

par le théorème (3.2), (x,λ) est un point de selle de $L(resp.L')$ si et seulement si

(1) ${}^t c + Qx + {}^t(\lambda A) \geqslant 0$ resp. ${}^t c - Qx - {}^t(\lambda A) \leqslant 0$

(2) $Ax - b \leqslant 0$

(3) ${}^t x.({}^t c + Qx + {}^t(\lambda A)) = 0$ resp. ${}^t x.({}^t c - Qx - {}^t(\lambda A)) = 0$

(4) $\lambda.(Ax-b) = 0$

(5) $\quad x \in \mathbb{R}_+^n, \lambda \in \mathbb{R}_+^m$

Pour obtenir des égalités dans (1) et (2) on ajoute des variables d'écart $u \in \mathbb{R}_+^n$ et $v \in \mathbb{R}_+^m$:

(1) $\quad {}^tc + Qx + {}^t(\lambda A) - u = 0 \quad$ resp. $\quad {}^tc - Qx - {}^t(\lambda A) + u = 0$

(2) $\quad Ax - b + v = 0$

$$(3)\quad \left.\begin{array}{l} \left.\begin{array}{r} {}^tx.u = 0 \\ resp.\quad {}^tx.(-u) = 0 \end{array}\right\} \Longleftrightarrow {}^tx.u = 0 \\ \\ (4)\qquad \lambda.(-v) = 0 \Longleftrightarrow \lambda v = 0 \end{array}\right\} \Longleftrightarrow {}^tx.u + \lambda.v = 0$$

(5) $\quad x, u \in \mathbb{R}_+^n, \lambda, v \in \mathbb{R}_+^m$

Le système d'équations à résoudre est donc le suivant :

$$\left\{\begin{array}{ll} Qx + {}^t(\lambda A) - u = \mp {}^tc \\ Ax \qquad\qquad\quad + v = b \end{array}\right. \qquad (-pour\mathrm{L}, +pour\mathrm{L'})$$

$$x, u \in \mathbb{R}_+^n, \lambda, v \in \mathbb{R}_+^m \text{ et } {}^txu + \lambda v = 0 \qquad \text{(seul terme nonlinéaire)}$$

La résolution de ce système se fait à l'aide du programme linéaire suivant qu'on obtient en ajoutant la variable auxiliaire $y \in \mathbb{R}_+^n$ dans (1) :

maximiser $-My$ (M est une constante $M >> 0$)

sous les contraintes :

$$\left\{\begin{array}{ll} Qx + {}^t(\lambda A) - u & + y = \mp {}^tc \\ Ax & + v = b \\ y, x, \in \mathbb{R}_+^n, \lambda, v\mathbb{R}_+^m \; et \; {}^txu + \lambda v = 0 \end{array}\right.$$

En appliquant le simplexe pour résoudre ce programme auxiliaire il faut à chaque étape veiller à ce que les variables entrantes sont telles que ${}^txu + \lambda v = 0$ reste valable. Ceci modifie légèrement les règles du simplexe,voir [7].

Exemple 3.6 *Si $x_1 > 0$ est variable de base, il n'est pas possible de faire entrer u_1 en base, parce qu'on aurait alors $x_1u_1 > 0$ et donc* ${}^txu = \sum_{j=1}^{n} x_ju_j > 0$. *Il faut choisir i' tel que $x_{i'} = 0$ afin que $x_{i'}u_{i'} = 0$ et* ${}^txu = 0$.

Exemple 3.7 *Problème de la production (suite et fin)*

$$f(x_1, x_2) = \begin{pmatrix} 5 & 5 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} - \frac{1}{2} \begin{pmatrix} x_1 & x_2 \end{pmatrix} \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

sous $\begin{pmatrix} 1 & 2 \\ 3 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \leqslant \begin{pmatrix} 8 \\ 9 \end{pmatrix}, x_1, x_2 \geqslant 0$

Le programme auxiliaire se lit :

maximiser $\left(-My_1 - My_2\right)$ *sous les contraintes :*

$$\begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + \begin{pmatrix} 1 & 2 \\ 3 & 1 \end{pmatrix} \begin{pmatrix} \lambda_1 \\ \lambda_2 \end{pmatrix} - \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} + \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} 5 \\ 5 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 2 \\ 3 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} 8 \\ 9 \end{pmatrix}$$

$$y_1, y_2, x_1, x_2, \lambda_1, \lambda_2, u_1, u_2, v_1, v_2 \geqslant 0 \quad et \quad {}^t x u + \lambda v = 0$$

Les contraintes s'écrivent

$$\begin{cases} 2x_1 + \lambda_1 + 3\lambda_2 - u_1 + y_1 = 5 \\ 2x_2 + 2\lambda_1 + \lambda_2 - u_2 + y_2 = 5 \\ x_1 + 2x_2 \qquad\qquad + v_1 = 8 \\ 3x_1 + x_2 \qquad\qquad + v_2 = 9 \end{cases}$$

Donc on obtient :

$$\begin{cases} y_1 = 5 - 2x_1 - \lambda_1 - 3\lambda_2 + u_1 \\ y_2 = 5 - 2x_2 - 2\lambda_1 - \lambda_2 + u_1 \\ x_1 + 2x_2 \qquad\qquad + v_1 = 8 \\ 3x_1 + x_2 \qquad\qquad + v_2 = 9 \end{cases}$$

Rappelons qu'il faut substituer y_1 *et* y_2 *dans la fonction à maximiser :*

$-My_1 - My_2 = -10M + 2Mx_1 + 2Mx_2 + 3M\lambda_1 + 4M\lambda_2 - Mu_1 - Mu_2.$

Le premier tableau du simplexe se lit :

			$2M$	$2M$	$3M$	$4M$	$-M$	$-M$	0	0	0	0
c^j			x_1	x_2	λ_1	λ_2	u_1	u_2	v_1	v_2	y_1	y_2
0	$x_1^j = y_1$	5	2	0	1	3	-1	0	0	0	1	0
0	$x_2^j = y_2$	5	0	2	2	1	0	-1	0	0	0	1
0	$x_3^j = v_1$	8	1	2	0	0	0	0	1	0	0	0
0	$x_4^j = v_2$	9	3	1	0	0	0	0	0	1	0	0
$-10M$		0	$-2M$	$-2M$	$-3M$	$-4M$	M	M	0	0	0	0

voir [**?**]*Aprés trois étapes du simplexe on trouve la solution optimale :*

$$(x_1, x_2, \lambda_1, \lambda_2) = (2.2, 2.4, 0, 0.23).$$

Exemple 3.8 *minimiser* $f(x_1, x_2) = -20x_1 + 10x_2 + 3x_1^2 + 2x_2^2$

sous les contraintes : $\begin{cases} 2x_1 - x_2 \leqslant 6 \\ -x_1 + x_2 \leqslant 10 \\ 2x_1 + 3x_2 \geqslant 8 \\ x_1, x_2 \geqslant 0 \end{cases}$

En forme matricielle, ce programme s'écrit

$$f(x) = cx + \frac{1}{2}{}^t xQx = \begin{pmatrix} -20 & 10 \end{pmatrix} x + \frac{1}{2}{}^t x \begin{pmatrix} 6 & 0 \\ 0 & 4 \end{pmatrix} x$$

$$Ax - b \leqslant 0 \Longleftrightarrow \begin{pmatrix} 2 & -1 \\ -1 & 1 \\ -2 & -3 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} - \begin{pmatrix} 6 \\ 10 \\ -8 \end{pmatrix} \leqslant 0$$

Le programme auxuliaire s'écrit :
maximiser $\big(-My_1 - My_2\big)$ *sous les contraintes :*

$$\begin{cases} (1) & Qx + {}^t(\lambda A) - u + y = -{}^t c \\ (2) & Ax + v = b \\ & y, x, u \in \mathrm{I\!R}_+^3, \lambda, v \in \mathrm{I\!R}_+^3 \ \ et \ \ {}^t xu + \lambda v = 0 \end{cases}$$

$$(1) \iff \begin{pmatrix} 6 & 0 \\ 0 & 4 \end{pmatrix} x + \begin{pmatrix} 2 & -1 & -2 \\ -1 & 1 & -3 \end{pmatrix} \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \lambda_3 \end{pmatrix} - \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} + \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} 20 \\ -10 \end{pmatrix}$$

$$\iff \begin{cases} 6x_1 + 2\lambda_1 - \lambda_2 - 2\lambda_3 - u_1 + y_1 = 20 \\ -4x_2 + \lambda_1 - \lambda_2 + 3\lambda_3 + u_2 + y_2 = 10 \end{cases}$$

Remarquons que y_2 *est superflu, parce que* u_2 *peut faire partie de la solution de base initiale.*

$$(2) \iff \begin{cases} 2x_1 - x_2 + v_1 = 6 \\ -x_1 + x_2 + v_2 = 10 \\ 2x_1 + 3x_2 - v_3 = 8 \end{cases}$$

$$\iff \begin{cases} 2x_1 - x_2 + v_1 = 6 \\ -x_1 + x_2 + v_2 = 10 \\ 2x_1 + 3x_2 - v_3 + \widetilde{y_2} = 8 \end{cases}$$

On en déduit :

$$\begin{cases} y_1 = 20 - 6x_1 - 2\lambda_1 + \lambda_2 + 2\lambda_3 + u_1 \\ \widetilde{y_2} = 8 - 2x_1 - 3x_2 + v_3. \end{cases}$$

La fonction à maximiser devient
$-My_1 - M\widetilde{y_2} = -28M + 8Mx_1 + 3Mx_2 + 2M\lambda_1 - M\lambda_2 - 2M\lambda_3 - Mu_1 - Mv_3.$
La solution de base initiale est donnée par :
$(y_1, u_2, v_1, v_2, \widetilde{y_2}) = (20, 10, 6, 10, 8)$
La résolution du programme par le simplexe se fait par les tableaux suivants (Tableaux 3.1 à 3.5). On obtient que $(x_1, x_2) = (3.25, 0.5)$ *est point de selle de L (notons que* f *est convexe).*
Par le théorème (3.1), ce point est solution optimale , voir [18]

Résumons cette section par la remarque suivante :

Remarque 3.6 *La solution optimale du programme quadratique étudié dans cette section est donnée par la solution optimale du programme linéaire suivant :*
maximiser $-My$ *(*$M > 0$ *une constante,* $y = {}^t(y_1, ..., y_n) \in \mathbb{R}^n_+$*)*
sous les contraintes :

$$\begin{cases} Qx + {}^t(\lambda A) - u + y = \mp\, {}^tc \quad \big(-\,pour(P_{min}), +pour(P_{max})\big) \\ Ax + y = b, \quad {}^txu + \lambda v = 0, \quad x, u \in \mathbb{R}^n_+, \quad \lambda, v \in \mathbb{R}^m_+ \end{cases}$$

	c_i	x_1	x_2	λ_1	λ_2	λ_3	u_1	u_2	v_1	v_2	v_3	y_1	$\tilde{y}_2$	b
y_1	$-M$	6	0	2	-1	-2	-1					$\boxed{1}$		20
u_2	0	0	-4	1	-1	3		$\boxed{1}$						10
v_1	0	2	-1						$\boxed{1}$					6
v_2	0	-1	1							$\boxed{1}$				10
$\tilde{y}_2$	$-M$	2	3								-1		$\boxed{1}$	8
		$-8M$	$-3M$	$-2M$	M	$2M$	M	0	0	0	M	0	0	

TAB. 3.1 - Les variables de base sont encadrées. Variable qui entre en base : x_1*. Variable qui sor de la base :* v_1*. La condition* ${}^txu + \lambda v = 0$ *est toujours satisfaite parce que* $u_1 = 0$ *et donc* $x_1u_1 = 0$*.*

	c_i	x_1	x_2	λ_1	λ_2	λ_3	u_1	u_2	v_1	v_2	v_3	y_1	$\tilde{y}_2$	b
y_1	$-M$		3	2	-1	-2	-1		-3			1		2
u_2	0		-4	1	-1	3		1						10
x_1	0	1	-1/2						1/2					3
v_2	0		1/2						1/2	1				13
$\tilde{y}_2$	$-M$		4						-1		-1		1	2
		0	$-7M$	$-2M$	M	$2M$	M	0	$4M$	0	M	0	0	

TAB.3.2 - Variable entrante : x_2. La variable sortante serait y_2. Mais $u_2 > 0$, donc $x_2 u_2 > 0$. La variable entrante est donc λ_1. La variable sortante est y_1. La contrainte ${}^t xu + \lambda v = 0$ reste satisfaite.

	c_i	x_1	x_2	λ_1	λ_2	λ_3	u_1	u_2	v_1	v_2	v_3	$\tilde{y}_2$	b
λ_1	0		3/2	1	-1/2	-1	-1/2		-3/2				1
u_2	0		-11/2		-1/2	4	1/2	1	3/2				9
x_1	0	1	-1/2						1/2				3
v_2	0		1/2						1/2	1			13
$\tilde{y}_2$	$-M$		4						-1		-1	1	2
		0	$-4M$	0	0	0	0	0	M	0	M	0	

TAB. 3.3 -Variable entrante : x_2. La variable sortante serait y_2. Mais $u_2 > 0$, donc $x_2 u_2 > 0$. La variable entrante est donc λ_3. La variable sortante est u_2. La contrainte ${}^t xu + \lambda v = 0$ reste satisfaite.

	c_i	x_1	x_2	λ_1	λ_2	λ_3	u_1	u_2	v_1	v_2	v_3	$\tilde{y}_2$	b
λ_1	0		1/8	1	-5/8		-3/8	1/4	-9/8				13/4
λ_3	0		-11/8		-1/8	1	1/8	1/4	3/8				9/4
x_1	0	1	-1/2						1/2				3
v_2	0		1/2						1/2	1			13
$\tilde{y}_2$	$-M$		4								-1	1	2
		0	$-4M$	0	0	0	0	0	M	0	M	0	

TAB. 3.4 -Variable entrante : x_2. Variable sortante $\widetilde{y}_2$. La contrainte ${}^t xu + \lambda v = 0$ reste satisfaite, parce que $u_2 = 0$ (sorti de base à l'étape précédente).

	c_i	x_1	x_2	λ_1	λ_2	λ_3	u_1	u_2	v_1	v_2	v_3	b
λ_1	0			1	-5/8		-3/8	-1/4	-35/32		1/32	3.18
λ_3	0				-1/8	1	1/8	1/4	1/32		-11/32	2.93
x_1	0	1							3/8		-1/8	3.25
v_2	0								5/8	1	1/8	12.75
x_2	0		1						-1/4		-1/4	0.5
		0	0	0	0	0	0	0	0	0	0	

TAB. 3.5 - Solution optimale du programme auxuliaire avec $y_1 = y_2 = 0$. *On a* $(x_1, x_2, \lambda_1, \lambda_2, \lambda_3) = (3.25, 0.5, 3.18, 0, 2.93)$.

3.5 La méthode de Franck-Wolfe

Nous allons voir un type particulier d'algorithme qui permet de résoudre les problèmes d'optimisation non linéaires convexes. Il est du type séquence d'approximations linéaires. Il s'agit de **L'algorithme de Franck-Wolfe**, voir [14].
Son principe est particulièrement simple dans le cas de problèmes avec contraintes linéaires. En effet, cette procédure combine des approximations linéaires de l'objectif pour trouver la **direction de recherche** avec une recherche unidimensionnelle pour trouver **le pas** suivant cette direction.
L'idée de la méthode est la suivante. Il s'agit d'un algorithme **d'approximationslinéaires successives.** Appliquons cette idée au problème suivant dont les contraintes sont purement linéaires :

$$\max f(x);\ \ f(x) = 5x_1 - x_1^2 + 8x_2 - 2x_2^2$$
$$s.c.q. \qquad \begin{cases} 3x_1 + 2x_2 \leqslant 6 \\ x_1 \qquad\quad \geqslant 0 \\ \qquad\quad x_2 \geqslant 0 \end{cases}$$

En effet, c'est le type de problème que nous allons rencontrer pour la solution du problème d'affectation du trafic.
Une **itération** de la méthode comporte les étapes suivantes :
a) Calcul de la direction :
On calcule la direction de recherche en résolvant un problème linéaire obtenu en remplaçant la fonction objectif par son approximation linéaire (points 1 et 2 ci-dessous).
b) Calcul du pas :
On calcule le pas à faire dans la direction en minimisant la fonction $f(x)$ le long de la direction déterminée en **a)** (points 3 et 4 ci-dessous).
Nous partons du point initial $x^0 = (0, 0)$.
Itération 1 :
1. Comme seul l'objectif est non linéaire, on va remplacer $f(x)$ par son approximation de Taylor limitée à l'ordre 1 :

$$f(x) \sim f(0,0) + \frac{\partial f}{\partial x_1}(0,0)(x_1 - 0) + \frac{\partial f}{\partial x_2}(0,0)(x_2 - 0)$$

Evaluons le gradient en $x^0 = (0, 0)$

$$\frac{\partial f}{\partial x_1} = 5 - 2(0) = 5$$
$$\frac{\partial f}{\partial x_2} = 8 - 4(0) = 8$$

D'où l'approximation linéaire de la fonction en $(0, 0)$:

$$g(x) = 5x_1 + 8x_2$$

2. On peut donc résoudre le problème linéaire :

$$\max g(x);\ \ g(x) = 5x_1 + 8x_2$$
$$s.c.q. \quad \begin{cases} 3x_1 + 2x_2 \leqslant 6 \\ x_1 \geqslant 0 \\ x_2 \geqslant 0 \end{cases}$$

par l'algorithme du simplexe en général. Comme il n'a que deux variables, ceci est fait graphiquement à la figure (3.1). La **solution optimale** est le point $x^1 = (0, 3)$

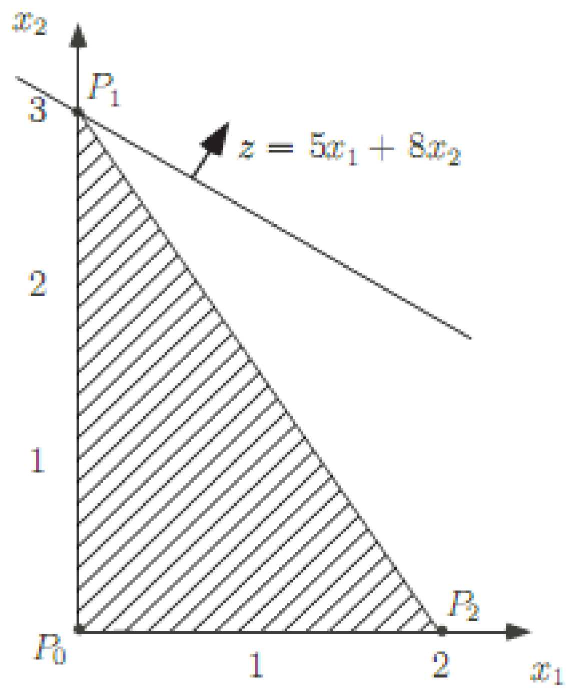


Figure (3.1) :méthode de Franck-Wolfe : itération 1

Comme, plus on s'éloigne du point de départ x^0, plus l'approximation linéaire $g(x)$ s'éloigne de la fonction $f(x)$, on a probablement été trop loin en allant jusqu'au point x^1_{LP}. Cependant, en se dirigeant dans la direction de x^1_{LP}, et on va déterminer le pas dans la direction :

$$x^1 = x^0 + \alpha(x^1_{LP} - x^0)$$
$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} + \alpha \begin{pmatrix} 0-0 \\ 3-0 \end{pmatrix} = \begin{pmatrix} 0 \\ 3\alpha \end{pmatrix}$$

avec $\alpha \in [0,1]$ qui minimise cette fois la vraie fonction :

$$h(\alpha) = f(x^0 + \alpha(x^1_{LP} - x^0)).$$

On obtient donc le problème de minimisation unidimensionnelle :

$$max\ h(\alpha) = 24\alpha - 18\alpha^2$$

dont le maximum est obtenu en annulant la dérivée : $\alpha = 2/3$. D'où

$$x^1 = (0,0) + 2/3[(0,3) - (0,0)] = (0,2)$$

Itération 2 :
1.Évaluation du gradient en $x^1 = (0,2)$

$$\frac{\partial f}{\partial x_1} = 5 - 2(0) = 5$$
$$\frac{\partial f}{\partial x_2} = 8 - 4(2) = 0$$

2. Résoudre le problème linéaire, voir [14] :

$$\max f(x);\ f(x) = 5x_1 + 0x_2$$
$$s.c.q. \qquad \begin{cases} 3x_1 + 2x_2 \leqslant 6 \\ x_1 \geqslant 0 \\ x_2 \geqslant 0 \end{cases}$$

Ceci est fait à la figure (3.2). La **solution optimale** est $x^2_{LP} = (2,0)$.
3. Recherche unidimensionnelle :

$$x^2 = x^1 + \alpha(x^2_{LP} - x^1)$$
$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 2 \end{pmatrix} + \alpha \begin{pmatrix} 2-0 \\ 0-2 \end{pmatrix} = \begin{pmatrix} 2\alpha \\ 2-2\alpha \end{pmatrix}$$

On obtient donc le problème de minimisation unidimensionnelle :

$$\max h(\alpha);\ \ h(\alpha) = 8 + 10\alpha - 12\alpha^2$$

qui est maximum pour $\alpha = 5/12$ d'où :

$$x^2 = (0,2) + 5/12[(2,0) - (0,2)] = (5/6, 7/6)$$

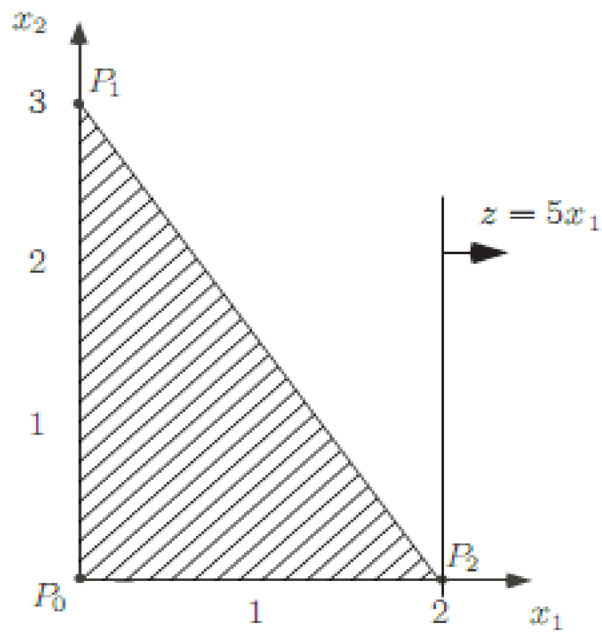


Figure (3.2) méthode de Franck-Wolfe : itération 2

On peut résumer l'algorithme de Franck-Wolfe comme suit :
Algorithme de Franck-Wolfe :
1. Initialisation :
Soit x^0 une solution initiale réalisable. Posons $k = 1$.
2. Itération k :
Pour $j = 1, 2, ..., n$, évaluer $c_j = \dfrac{\partial f(x^{k-1})}{\partial x_j}$
3. Calcul de la direction :
Trouver, par application de l'algorithme du simplexe, x^k_{LP} la solution optimale du problème linéaire suivant, voir [14] :

$$max\ g(x) = \sum_{j=1}^{n} c_j x_j$$

$$scq \qquad \begin{cases} Ax \leqslant b \\ x \ \ \geqslant 0 \end{cases}$$

4. Calcul du pas :
Trouver α_k la solution optimale de

$$\begin{array}{ll} max & f\left[x^{k-1} + \alpha(x^k_{LP} - x^{k-1})\right] \\ scq & 0 \leqslant \alpha \leqslant 1 \end{array}$$

et mettre

$$x^k = x^{k-1} + \alpha_k(x^k_{LP} - x^{k-1})$$

5. Critère d'arrêt :
Si x^{k-1} et x^k sont suffisamment proches, stop. Sinon retour en 2.

Chapitre 4

Calcul de l'optimum de la fonction concave

Soit la fonction concave définit sur $\mathbb{R}^n$ par :

$$\varphi(x) = \sum_{i=1}^{n} \alpha_i x_i + \beta_i x_i^2 \textit{ avec } \alpha_i \in \mathbb{R},\ \textit{et } \beta_i < 0 \textit{ pour tout indice } i$$

et soit Ω le convexe fermé de $\mathbb{R}^n$ tel que :

$$\Omega = \{x \in \mathbb{R}^n : Ax \leqslant 0,\ x \geqslant 0\}.$$

Posons :

$$x^* = (x_i^*)_i = \left(\frac{-\alpha_i}{2\beta_i}\right)_i, \quad y^* = (y_i^*)_i = \left(\frac{\alpha_i}{2\sqrt{-\beta_i}}\right)_i$$

les points critiques de la fonction φ.

$$y = (y_i)_i = \left(\sqrt{-\beta_i} x_i\right)_i$$

$\overline{x}$ la projection du point critique x^* sur Ω
$\overline{y}$ la projection du point critique y^* sur $\grave{\Omega}$ où

$$\grave{\Omega} = \{y \in \mathbb{R}^n : \grave{A}y \leqslant 0,\ y \geqslant 0\}.$$

$$\max_{x \in \Omega} \varphi(x) = \varphi(\overline{x}),\ \overline{x} = (\overline{x_i})_i \in \Omega$$

Théorème 4.1 *Il existe un ensemble convexe fermé borné $\grave{\Omega}$ de $\mathbb{R}^n$, et un vecteur $\overline{y} = (\overline{y}_i)_i \in \grave{\Omega}$, tel que les conditions (propriétés) suivantes soient satisfaites :*

1. $\max\limits_{x\in\Omega} \varphi(x) = \varphi(x^*) - \|y^* - \overline{y}\|^2$
2. $\|y^* - \overline{y}\| = \inf\limits_{y\in\grave{\Omega}} \|y^* - y\|$
3. $\overline{x}_i = \dfrac{\overline{y}_i}{\sqrt{-\beta_i}}$ *pour tout* $i,\ i = 1, 2, ..., n.$

Preuve,voir [13] : Pour tout $x \in \Omega$, soit :

$$\begin{aligned}
\Delta\varphi_i &= (\alpha_i x_i^* + \beta_i x_i^*) - (\alpha_i x_i + \beta_i x_i^2) \\
&= \alpha_i(\frac{-\alpha_i}{2\beta_i}) + \beta_i(\frac{-\alpha_i}{2\beta_i})^2 - \alpha_i x_i - \beta_i x_i^2 \\
&= (\frac{-\alpha_i^2}{2\beta_i}) + \frac{\alpha_i^2}{4\beta_i}) - \alpha_i x_i - \beta_i x_i^2 \\
&= (\frac{-\alpha_i^2}{4\beta_i}) - \alpha_i x_i - \beta_i x_i^2 \\
&= -\beta_i\Big[x_i^2 + \frac{\alpha_i}{\beta_i}x_i + (\frac{\alpha_i}{2\beta_i})^2\Big] \\
&= -\beta_i\Big(x_i + \frac{\alpha_i}{2\beta_i}\Big)^2 \\
&= -\beta_i\Big(x_i - x_i^*\Big)^2.
\end{aligned}$$

D'une autre part on a :

$$\varphi(x^*) - \varphi(x) = \sum_{i=1}^{n} \Delta\varphi_i.$$

Donc :

$$\varphi(x^*) - \varphi(x) = \sum_{i=1}^{n} -\beta_i(x_i - x_i^*)^2 \textit{ pour tout } x \in \Omega.$$

Par conséquent :

$$\inf_{x\in\Omega}\Big[\varphi(x^*) - \varphi(x)\Big] = \inf_{x\in\Omega}\Big[\sum_{i=1}^{n} -\beta_i(x_i - x_i^*)^2\Big]$$

Peut être écrite sous la forme suivante :

$$\begin{aligned}\varphi(x^*) - \max_{x\in\Omega}\varphi(x) &= \inf_{x\in\Omega}\sum_{i=1}^{n}\left(\sqrt{-\beta_i}(x_i - x_i^*)\right)^2\\ &= \inf_{y\in\grave{\Omega}}\sum_{i=1}^{n}\left(y_i^* - y_i\right)^2\\ &= \|y^* - \overline{y}\|^2\end{aligned}$$

parce que : $\inf_{y\in\grave{\Omega}}\sum_{i=1}^{n}(y_i^* - y_i)^2 = \|y^* - \overline{y}\|^2,\ \overline{y} \in \grave{\Omega}$ avec

$$\grave{\Omega} = \{y = (y_i)_i \in \mathbb{R}^n,\ y_i = \sqrt{-\beta_i}x_i,\ x = (x_i)_i \in \Omega\}.$$

Donc :

$$\varphi(x^*) - \max_{x\in\Omega}\varphi(x) = \|y^* - \overline{y}\|^2$$

D'où

$$\max_{x\in\Omega}\varphi(x) = \varphi(x^*) - \|y^* - \overline{y}\|^2$$

d'où la propriété (1)
Prouvons la propriété (2) on sait que :

$$\inf_{y\in\grave{\Omega}}\|y^* - y\|^2 = \|y^* - \overline{y}\|^2$$

Et donc on a :

$$\|y^* - \overline{y}\|^2 \leqslant \|y^* - y\|^2 \textit{ pour tout } y \in \grave{\Omega}$$

Par conséquent nous avons :

$$\|y^* - \overline{y}\| \leqslant \inf_{y\in\grave{\Omega}}\|y^* - y\| \tag{4.1}$$

D' autre part puisque $y \in \grave{\Omega}$ on a :

$$\|y^* - \overline{y}\| \geqslant \inf_{x\in\grave{\Omega}}\|y^* - y\| \tag{4.2}$$

D'après (4.1) et (4.2) on obtient : $\|y^* - \overline{y}\| = \inf_{y\in\grave{\Omega}}\|y^* - y\|$

d'où la propriété (2)

Prouvons maintenant la propriété (3)
Puisque $\bar{y}$ est la projection du y^* sur le nouveau convexe $\grave{\Omega}$ nous avons :

$$\begin{aligned}
\max_{x\in\Omega}\varphi(x) &= \varphi(\overline{x}) \\
&= \varphi(x^*) - \|y^* - \overline{y}\|^2 \\
&= \varphi(x^*) - \inf_{x\in\Omega}\sum_{i=1}^{n}\left(y_i^* - \sqrt{-\beta_i}x_i\right)^2 \\
&= \varphi(x^*) - \sum_{i=1}^{n}\left(y_i^* - \sqrt{-\beta_i}(\overline{x}_i\right)^2
\end{aligned}$$

D'où la propriété (3)

Remarque 4.1 *Si $x^* \in \Omega$ alors :*

$$\inf_{x\in\Omega}\sum_{i=1}^{n} -\beta_i(x_i^* - x_i)^2 = 0 \qquad et \quad \max_{x\in\Omega}\varphi(x) = \varphi(x^*)$$

.

Remarque 4.2 *Soit toujours Ω le convexe fermé borné.*
On peut se restreindre au cas $\alpha_i \geq 0$ pour tout i.
En effet supposons que nous avons :

$$\varphi(x) = \sum_i \alpha_i x_i + \beta_i x_i^2 \text{ avec } \alpha_i < 0 \text{ et } \beta_i < 0 \text{ pour tout } i.$$

Soit $\max(x_i) = \delta_i^*,\ y_i = \delta_i^* - x_i \geqslant 0.$
Par conséquent :

$$\begin{aligned}
\varphi(x) &= \sum_i \alpha_i(\delta_i^* - y_i) + \beta_i(\delta_i^* - y_i)^2 \\
&= \sum_i \alpha_i\delta_i^* - \alpha_i y_i + \beta_i(\delta_i^{*2} + y_i^2 - 2\delta_i^* y_i) \\
&= \sum_i (-\alpha_i - 2\beta_i\delta_i^*)y_i + \beta_i y_i^2 + (\alpha_i\delta_i^* + \beta_i\delta_i^{*2}) \\
&= \sum_i \grave{\alpha}_i y_i + \beta_i y_i^2 + \gamma_i
\end{aligned}$$

où

$$\grave{\alpha}_i = -\alpha_i - 2\beta_i\delta_i^* \geqslant 0 \quad et \quad \gamma_i = \alpha_i\delta_i^* + \beta_i\delta_i^{*2}.$$

Donc il suffit de remplacer :
x_i *par* $(\delta_i^* - y_i)$ *et* $(\alpha_i x_i + \beta_i x_i^2)$ *par* $(\alpha_i y_i + \beta_i x_i^2 + \gamma_i)$.

4.1 la recherche de la projection du point critique.

Soit H l'hyperplan de l'équation : $< a, x >= b$ et soit : $x^* \notin H$ c'est-à-dire : $< a, x^* > -b > 0$
Considérons le point x_0 de $\mathbb{R}^n$ tel que $x^* - x_0 = \lambda a$ $\big[(x^* - x_0)$ est parallèle à $a\big]$ où λ est une constante positive et tel que $x_0 \in H$. Pour cela on doit avoir :

$$x_0 = x^* - \lambda a \tag{4.3}$$

Alors :

$$\begin{aligned} < a, x_0 >= b &\iff < x^* - \lambda a, a >= b \\ &\iff < x^*, a > -\lambda < a, a >= b \\ &\iff \lambda < a, a >=< x^*, a > -b \\ &\iff \lambda = \frac{< x^*, a > -b}{< a, a >} \\ &\iff \lambda = \frac{< x^*, a > -b}{\|a\|^2} \end{aligned} \tag{4.4}$$

où $\|a\|^2 =< a, a >$. Pour cette valeur de λ et l'appartenance de x_0 à H on a :

$$\|x^* - x_0\| = \lambda\|a\| = \frac{< x^*, a > -b}{\|a\|}$$

Posons :$d(x^*, H) = d(x^*, \overline{x})$ où $\overline{x}$ est la projection orthogonale du point x^* sur H.
Alors :

$$\|x^* - x_0\| \geqslant d(x^*, \overline{x}) \ car \ x_0 \in H$$

d'où

$$\frac{< x^*, a > -b}{\|a\|} \geqslant d(x^*, \overline{x}) = \|x^* - \overline{x}\|. \tag{4.5}$$

Mais d'une autre part l'application :

$$\begin{array}{rcl} \mathbb{R}^n & \longrightarrow & \mathbb{R} \\ x & \longmapsto & < a, x >= ax \end{array}$$

est linéaire et continue donc :

$$\begin{aligned} | < a, x^* > - < a, \overline{x} > | &= \|ax^* - a\overline{x}\| \\ &= |ax^* - b| \\ &= ax^* - b \end{aligned}$$

car : $\overline{x} \in H$ et $ax^* - b > 0$

$$\begin{aligned} \|ax^* - a\overline{x}\| &= \|a(x^* - \overline{x})\| \leqslant \|a\|\|x^* - \overline{x}\| \\ &= ax^* - b \leqslant \|a\|\|x^* - \overline{x}\| \end{aligned}$$

Ce qui implique que :

$$\|x^* - \overline{x}\| \geqslant \frac{ax^* - b}{\|a\|} = \frac{< x^*, a > -b}{\|a\|} \tag{4.6}$$

D'après les inégalités (4.5) et (4.6) on obtient :

$$\|x^* - \overline{x}\| = \frac{< x^*, a > -b}{\|a\|}$$

Remarquons que d'après : (4.3) et (4.4) ona :

$$x_0 = x^* - \frac{< x^*, a > -b}{\|a\|^2}.a \tag{4.7}$$

Conclusion,voir [13]

On cherche $\inf\{\frac{< x^*, a_i > -b_i}{\|a_i\|^2}\},\ i = 1, 2, ...$

Soit $\frac{< x^*, a_{i_0} > -b_{i0}}{\|a_{i0}\|^2}$ et par suite :

$x_0 = x^* - \frac{< x^*, a_{i_0} > -b_{i0}}{\|a_{i0}\|^2}.a_{i0}$ si $x^* \in H_{j_0} \Longrightarrow < a_{j_0}, x^* >= b_{j_0}$

ce qui implique :

$$\begin{aligned} x_0 &= x^* - \frac{0}{\|a_{j_0}\|^2}.a_{j_0} \\ &= x^* \end{aligned}$$

Donc si x^* appartient au convexe alors x^* est l'éxtrémum, si non , $x^* \notin H$ c-à-d :

$$< a, x^* > -b > 0 \ et \ x_0 = \overline{x}$$

Remarque 4.3 *Dans le cas où* $\|x^* - x_0\| = d(x^*, \overline{x})$ *on a :* $x_0 = \overline{x}$ *et d'après l'égalité (4.7) on trouve*

$$\overline{x} = x^* - \frac{< x^*, a > -b}{\|a\|^2}.a.$$

Exemple 4.1 *Soit le problème quadratique suivant :*

$$\begin{cases} \max \varphi(x_1, x_2) = 5x_1 + 5x_2 - x_1^2 - x_2^2, \\ 3x_1 + x_2 - 9 \leqslant 0, \\ x_1 + 2x_2 - 8 \leqslant 0, \\ x_1, x_2 \geqslant 0. \end{cases}$$

on pose : $\Omega = \{x \in \mathbb{R}^2;\ 3x_1 + x_2 - 9 \leqslant 0,\ x_1 + 2x_2 - 8 \leqslant 0,\ x \geqslant 0\}$
on a : $\beta_1 = \beta_2 = -1$ *donc :* $\Omega = \grave{\Omega}$ *cherchons le point critique*

$$\begin{cases} \dfrac{\partial \varphi}{\partial x_1}(x_1, x_2) = 5 - 2x_1, \\ \dfrac{\partial \varphi}{\partial x_2}(x_1, x_2) = 5 - 2x_2. \end{cases}$$

$$\begin{cases} \dfrac{\partial \varphi}{\partial x_1}(x_1, x_2) = 0, \\ \dfrac{\partial \varphi}{\partial x_2}(x_1, x_2) = 0. \end{cases}$$

Et par suite on obtient :

$$\begin{cases} x_1^* = \dfrac{5}{2}, \\ x_2^* = \dfrac{5}{2}. \end{cases}$$

D'où
$x^* = (\dfrac{5}{2}, \dfrac{5}{2})$ *qui n'appartient pas à* Ω *parceque* $3 \times \frac{5}{2} + \frac{5}{2} = 1 > 0$
Calculons maintenant $\overline{x}$

$$\overline{x} = P_\Omega(x^*) = x^* - \frac{< x^*, a > -b}{\|a\|^2}.a$$

où $a = (3,1)$, $b = 9$

$$< x^*, a >= \frac{5}{2} \times 3 + \frac{5}{2} \times 1 = 10 \; et \; \|a\|^2 = 10$$

$$\begin{aligned} \overline{x} &= (\frac{5}{2}, \frac{5}{2}) - \frac{1}{10}(3,1) \\ &= (\frac{5}{2} - \frac{3}{10}, \frac{5}{2} - \frac{1}{10}) \\ &= (\frac{22}{10}, \frac{24}{10}) \\ &= (2,2, \; 2,4). \end{aligned}$$

et par conséquent on trouve :

$$\begin{aligned} \max_{(x_1,x_2)\in\Omega} \varphi(x_1, x_2) &= \varphi(\overline{x}) \\ &= 5 \times 2,2 + 5 \times 2,4 - (2,2)^2 - (2,4)^2 \\ &= 12,4 \end{aligned}$$

Exemple 4.2 *Soit le problème :*

$$\begin{cases} \max \varphi(x_1, x_2) = 5x_1 + 8x_2 - x_1^2 - 2x_2^2, \\ 3x_1 + 2x_2 - 8 \leqslant 6, \\ x_1 \geqslant 0, x_2 \geqslant 0. \end{cases}$$

posons $\Omega = \{x \in \mathbb{R}^2; \; 3x_1 + 2x_2 \leqslant 6, \; x \geqslant 0\}$ *Dans ce cas :*
$\beta_1 = -1$ *et* $\beta_2 = -2$ *donc* $\Omega \neq \grave{\Omega}$ *cherchons le point critique :*
$\frac{\partial \varphi}{\partial x_1}(x_1, x_2) = 5 - 2x_1 = 0 \Longrightarrow x_1^* = \frac{5}{2}$,
$\frac{\partial \varphi}{\partial x_2}(x_1, x_2) = 8 - 4x_2 = 0 \Longrightarrow x_2^* = 2$,
Donc le point critique est $x^* = (\frac{5}{2}, 2) \notin \Omega$ *parce que* $3 \times \frac{5}{2} + 2 \times 2 > 6$
Trouvons $\overline{x}$

$$\overline{x} = P_\Omega(x^*) = x^* - \frac{< x^*, a > -b}{\|a\|^2}.a$$

où $a = (3,2)$, $b = 6$

$$< x^*, a >= \frac{5}{2} \times 3 + 2 \times 2 = \frac{23}{2} \; et \; \|a\|^2 = 13$$

$$\begin{aligned}\overline{x} &= (\frac{5}{2}, 2) - \frac{11}{2 \times 13}(3, 2) \\ &= (\frac{5}{2} - \frac{33}{26}, 2 - \frac{11}{13}) \\ &= (\frac{16}{13}, \frac{15}{13}).\end{aligned}$$

$$\grave{\Omega} = \{(y_1, y_2) \in \mathbb{R}^2 : 3y_1 + \frac{2}{\sqrt{2}}y_2 \leqslant 6,\ y_1 \geqslant 0,\ y_2 \geqslant 0\}$$

De plus $y^* = (\frac{5}{2}, 2\sqrt{2})$.
La projection de ce point sur $\grave{\Omega}$ *donne le point suivant* $\overline{y} = (1, \frac{3}{\sqrt{2}})$
ce qui implique que : $\overline{x} = (1, \frac{3}{2})$ *et* $\max_{x \in \Omega} \varphi(x) = \varphi(1, \frac{3}{2}) = 11,5$.
Notons que la projection du point x^* *sur* Ω *donne le point* $\overline{x} = (\frac{16}{13}, \frac{15}{13})$ *et* $\varphi(\frac{16}{13}, \frac{15}{13}) = 11,207$.
Evidemment $\max\limits_{x \in \Omega} \varphi(x) \leq \varphi(x^*) = \varphi(\frac{5}{2}, 2) = 14,25$.
Le point $\overline{x} = (\frac{16}{13}, \frac{15}{13})$ *de l'ensemble convexe* Ω *est le point plus proche du point critique* x^* *mais ce n'est pas la solution optimale de* φ

4.2 la recherche de l'hyperplan séparant le point critique et le convexe fermé Ω.

On cherche l'hyperplan $< a, x >= b$ qui sépare x^* et Ω c'est-à-dire $< a, x^* >> b$
On sait que $\overline{x}$ est un élément de l'hyperplan d 'equation $< a, x >= b$
Alors :

$$a_{11}x_1 + a_{12}x_2 + ... + a_{1n}x_n = b \tag{4.8}$$

Supposons que $a_{11} \neq 0$ par suit :

$$\begin{aligned}(4.8) \iff & x_1 \frac{a_{12}}{a_{11}}x_2 + ... + \frac{a_{1n}}{a_{11}}x_n = \frac{b}{a_{11}} \\ \iff & \begin{cases} \grave{a}_{11}x_1 + \grave{a}_{12}x_2 + ... + \grave{a}_{1n}x_n = \grave{b}, \\ \grave{a}_{12} = \frac{a_{12}}{a_{11}}, ..., \grave{a}_{1n} = \frac{a_{1n}}{a_{11}},\ \grave{b} = \frac{b}{a_{11}}. \end{cases}\end{aligned}$$

Soit alors le problème :

$$\begin{cases} \min \varphi(x), \varphi(x) = \sum_{i=1}^{n}(x_i - x_i^*)^2 \\ x_1 + \grave{a}_{12}x_2 + ... + \grave{a}_{1n}x_n = \grave{b}. \end{cases}$$

D'après le théorème de Lagrange voir [13], et [18]

$$L(x,\lambda) = \sum_{i=1}^{n}(x_i - x_i^*)^2 + \lambda\{x_1 + \grave{a}_{12}x_2 + ... + \grave{a}_{1n}x_n - \grave{b}\}$$

$$\begin{aligned} \frac{\partial L}{\partial x_i}(x,\lambda) = 0 \iff & 2(x_i - x_i^*) + \lambda\grave{a}_{1i} = 0 \; pour \; i \neq 1 \\ \iff & x_i - x_i^* = -\frac{\lambda}{2}\grave{a}_{1i}, \; i \neq 1 \end{aligned}$$

$$\begin{aligned} \frac{\partial L}{\partial x_1}(x,\lambda) = 0 \iff & 2(x_i - x_1^*) + \lambda = 0 \\ \iff & \frac{\lambda}{2} = x_1^* - x_1. \end{aligned}$$

où

$$\begin{cases} x_i - x_i^* = -\frac{\lambda}{2}\grave{a}_{1i}, \; si \; i \neq 1 \\ \frac{\lambda}{2} = x_1^* - x_1. \end{cases}$$

où $x_i = x_i^* - \grave{a}_{1i}(x_1^* - x_1)$.
Calculons x_1, on a :

$$\begin{aligned} x_1 &= \grave{b} - \sum_{i=2}^{n}\grave{a}_{1i}x_i \\ &= \grave{b} - \sum_{i=2}^{n}\grave{a}_{1i}[x_i^* - \grave{a}_{1i}(x_1^* - x_1)] \\ x_1 &= \grave{b} - \sum_{i=2}^{n}\grave{a}_{1i}x_i^* + \sum_{i=2}^{n}\grave{a}_{1i}^2(x_i^* - x_1) \end{aligned}$$

$$x_1 + \sum_{i=2}^{n}\grave{a}_{1i}x_i = \grave{b} - \sum_{i=2}^{n}\grave{a}_{1i}x_i^* + \sum_{i=2}^{n}\grave{a}_{1i}^2x_i^*$$

$$x_1\big(1+\sum_{i=2}^{n}\grave{a}_{1i}^2\big) = \grave{b} + x_1^*\sum_{i=2}^{n}\grave{a}_{1i}^2 - \sum_{i=2}^{n}\grave{a}_{1i}x_i^*$$

ce qui implique :

$$x_1 = \frac{1}{1+\sum_{i=2}^{n}\grave{a}_{1i}^2}\Big\{\grave{b} + x_1^*\sum_{i=2}^{n}\grave{a}_{1i}^2 - \sum_{i=2}^{n}\grave{a}_{1i}x_1^*\Big\}$$

et

$$x_i = x_i^* - \grave{a}_{1i}(x_1^* - x_1),\ \forall i = 1, 2, ...$$

Théorème 4.2 *Si* $\overline{x} \in \Omega$, *alors :*

$$\inf_{x\in\Omega}\|x^* - x\| = \|x^* - \overline{x}\|$$

Preuve : Soit H l'hyperplan séparant x^* et Ω de l'équation $< a, x >= b$. On sait que : $< a, x^* >> b$ et $< a, x >\leqslant b,\ \forall x \in \Omega$. pour $x_0 \in \Omega\backslash H$, $< a, x_0 >< b$ posons :

$$\|\overline{x} - x^*\| = \inf_{x\in H}\|x - x^*\|$$

Alors :

$$\|\overline{x} - x^*\| \leqslant \|x - x^*\|,\ \forall x \in H \tag{4.9}$$

Soit $x_0 \in \Omega\backslash H$, il existe $z \in H$ tel que :

$$z = \lambda x_0 + (1-\lambda)x^*$$

En effet :

$$\begin{aligned}
< a, z >= b &\iff < a, \lambda x_0 + (1-\lambda)x^* >= b\\
&\iff < a, \lambda x_0 > + < a, (1-\lambda)x^* >= b\\
&\iff \lambda < a, x_0 > + < a, x^* > + < a - \lambda x^* >= b\\
&\iff \lambda < a, x_0 > -\lambda < a, x^* >= b- < a, x^* >\\
&\iff -\lambda[< a, x^* > - < a, x_0 >] = b- < a, x^* >\\
&\iff \lambda[< a, x^* > - < a, x_0 >] =< a, x^* > -b
\end{aligned}$$

D'autre part on a : $< a, x^* >> b$ implique $< a, x^* > -b > 0$
et

$$\begin{cases} < a, x^* >> b \\ < a, x_0 >< b. \end{cases}$$

implique

$$\begin{cases} < a, x^* >> b \\ - < a, x_0 >> -b. \end{cases}$$

D'où :

$$< a, x^* > - < a, x_0 >> 0$$

Donc on trouve que :

$$0 < \lambda = \frac{< a, x^* > -b}{< a, x^* > - < a, x_0 >} < 1$$

Or $0 < \lambda < 1$ du fait que :

$$< a, x^* > -b << a, x^* > - < a, x_0 >$$

car : $< a, x_0 >< b$ et $x_0 \in \Omega \backslash H$
Comme :

$$\begin{aligned} z - x^* &= \lambda x_0 + (1 - \lambda) x^* - x^* \\ &= \lambda x_0 - \lambda x^* \\ &= \lambda (x_0 - x^*) \end{aligned}$$

ça implique que :

$$\| z - x^* \| < \| x^* - x_0 \| \; car \; 0 < \lambda < 1$$

et d'après l'inégalité (4.9) on trouve que

$$\| \overline{x} - x^* \| \leqslant \| z - x^* \| \text{ parceque } z \in H$$

D'où $\| \overline{x} - x^* \| \leqslant \| z - x^* \| < \| x^* - x_0 \|$
Alors :

$$\forall x_0 \in \Omega \backslash H \quad \| \overline{x} - x^* \| < \| x^* - x_0 \|.$$

Théorème 4.3 *S'il existe deux hyperplans H_1 et H_2 tels que :*
$\inf\limits_{x \in H_1} \|x^* - x\| = \|x^* - \overline{x}_1\|$ *et* $\overline{x}_1 \in \Omega$
$\inf\limits_{x \in H_2} \|x^* - x\| = \|x^* - \overline{x}_2\|$ *et* $\overline{x}_2 \in \Omega$
Aors $\overline{x}_1 = \overline{x}_2$ et $\overline{x}_1$ est un point extrème de Ω

Preuve :
En effet l'application :$x \longrightarrow \|x^* - x\|$ est convexe, et le minimum est unique D'après le théorème (4.2) on a :

$$\|x^* - \overline{x}_1\| = \|x^* - \overline{x}_2\|$$

d'où $\overline{x}_1 = \overline{x}_2$ et $\overline{x}_1 \in H_1 \cap H_2$,
ici H_1 et H_2 séparant x^*.

4.3 Extrémum d'une fonction quadratique concave φ

Dans cette section on va décrire une méthode numérique , qui nous donne une valeur approchée de l'optimum de la fonction φ.
Soit : $\varphi = \sum_{i=1}^{i=n} \alpha_i x_i + \beta_i x_i^2$ $\alpha_i \in \mathbb{R}, \beta_i < 0$, et Ω un convexe fermé défini par : $\Omega = \{x \in \mathbb{R}^n; Ax \leqslant 0, x \geqslant 0\}$.
On va maximiser la fonction φ ; posons : $\psi(x) = -\varphi(x)$;
Donc on obtient :

$$\max_{x \in \Omega} \varphi(x) = -\min_{x \in \Omega} \psi(x).$$

Alors au lieu de maximiser φ il suffit de minimiser la fonction ψ .
Remarquons que la fonction ψ est convexe.
Soit : $\psi(x) = \sum_{i=1}^{i=n} \alpha_i x_i + \beta_i x_i^2, \text{avec } \alpha_i \in \mathbb{R}; et\ \beta_i > 0$
Construisons une suite récurrente $(x_n)_{n \in \mathbb{N}}$:

$$x_{n+1} = P_\Omega(x_n - \alpha \nabla \psi(x));\ et\ x_0 \in \Omega\ (x_0 \text{un point donné})$$

Où α est un réel vérifiant : $0 < \alpha < \dfrac{1}{\gamma}$ où $\gamma = \max_i \beta_i$
On va démontrer que cette suite converge vers le minimum local de la fonction objectif ψ.

Théorème 4.4 *Soit la suite* $(x_n)_{n \in \mathbb{N}}$ *définie par :*

$$\begin{cases} x_0 \text{ un point donné de } \Omega. \\ x_{n+1} = P_\Omega\big(x_n - \alpha \nabla \psi(x_n)\big)\ n \geqslant 0 \end{cases}$$

Où $0 < \alpha < \dfrac{1}{\gamma}$, $\beta_i > 0$ *et* $\gamma = \max_i \beta_i$
Alors (x_n) *converge vers le minimum local de la fonction* ψ.

La démonstation de ce théorème passe par quelques lemmes.

Lemme 4.1 *soit* ψ *une fonction définie sur un ouvert* U *d'un espace de Hilbert et* Ω *une partie convexe de* U, *si la fonction* ψ *est différentiable en un point* $x \in \Omega$ *et si elle admet en* x *un minimum local, alors :*

$$\langle \nabla \psi(x), y - x \rangle \geqslant 0 \quad \textit{pour tout} \quad y \in \Omega$$

Preuve : Soit $y = x + h$ un point qulconque de Ω, les points de la forme : $(x + \theta h)$ avec $0 \leqslant \theta \leqslant 1$ sont encore dans Ω,
car :$\big((x + \theta h) = (1-\theta)x + \theta y$ et Ω est convexe$\big)$.
D'une autr part :
la différentiabilité de la fonction ψ en x permet d 'écrire :

$$\psi(x + \theta h) - \psi(x) = \theta[\langle \nabla\psi(x), h\rangle + \varepsilon(\theta)]$$

où $\lim_{\theta\to 0} \varepsilon(\theta) = 0$ pour tout $\theta \in [0, 1]$
(le vecteur h étant fixé),alors le nombre $\langle \nabla\psi(x), h\rangle$ est nécessairement positif, sans quoi la différance
$(\psi(x + \theta h) - \psi(x))$ serait strictement négative pour $\theta > 0$ suffisament petit, puisque : $\psi(x + \theta h) > \psi(x)$, car : ($\psi$ admet un minimum en x), donc : $\langle \nabla\psi(x), h\rangle \geqslant 0$
d'où : $\langle \nabla\psi(x), y - x\rangle \geqslant 0$ pour tout $y \in \Omega$

Lemme 4.2 *soit Ω un convexe fermé non vide de $\mathbb{R}^n$ et $P_\Omega(x)$ la projection orthogonale du point x sur Ω Alors ; $P_\Omega(x)$ est caractérisée par :*
$\langle x - P_\Omega(x), y - P_\Omega(x)\rangle \leqslant 0$ *pour tout $y \in \Omega$.*

Preuve : Soit y un point de Ω, le point $y_t = [(1-t)P_\Omega(x) + ty]$ appartient à Ω pour tout $t \in [0, 1]$ (car Ω est convexe).
Donc d'aprés la définition de la projetée d'un point sur un conexe fermé.
On a :

$$\|x - P_\Omega(x)\|^2 \leqslant \|x - y_t\|^2 \text{ pour tout } y_t \in \Omega.$$

Mais :

$$\begin{aligned}
\|x - y_t\|^2 &= \|x - (1-t)P_\Omega(x) - ty)\|^2 \\
&= \|x - P_\Omega(x) - t(y - P_\Omega(x))\|^2 \\
Car:\ y_t &= P_\Omega(x) + t(y - P_\Omega(x)) \\
D'ou: & \\
\|x - y_t\|^2 &= \|x - P_\Omega(x)\|^2 + t^2\|y - P_\Omega(x)\|^2 - 2t\langle x - P_\Omega(x), y - P_\Omega(x)\rangle.
\end{aligned}$$

Puisque : $\|x - P_\Omega(x)\|^2 \leqslant \|x - y_t\|^2$, $\forall\, y_t \in \Omega$

On obtient :

$$2t\langle x - P_\Omega(x), y - P_\Omega(x)\rangle \leqslant t^2\|y - P_\Omega(x)\|^2$$

Ce qui implique :

$$\langle x - P_\Omega(x), y - P_\Omega(x)\rangle \leqslant t/2\|y - P_\Omega(x)\|^2$$

D'une autre part cette inégalité est vérifie pour tout $t, 0 \leqslant t \leq 1$
ce qui donne : $\langle x - P_\Omega(x), y - P_\Omega(x)\rangle \leqslant 0$.

Lemme 4.3 *Soit $x_0 \in \Omega$, (Ω étant un convexe fermé), supposons que ψ est différentiable en x_0 et qu'il existe un réel positif α tel que :*

$$P_\Omega(x_0 - \alpha\nabla\psi(x_0)) = x_0.$$

Alors la fonction $\psi(x)$ admet un minimum local en x_0.

Preuve : posons :

$$\begin{cases} x = x_0 - \alpha\nabla\psi(x_0) \\ y = x_0 + v \end{cases}$$

avec v un vecteur quelconque tel que $y \in \Omega$.
Appliquons lemme (4.2) .
C'est-à-dire : $\langle x - P_\Omega(x), y - P_\Omega(x)\rangle \leqslant 0$.
Donc on a :

$$\langle(x_0 - \alpha\nabla\psi(x_0) - P_\Omega(x), x_0 + v - P_\Omega(x_0 - \alpha\nabla\psi(x_0))\rangle \leqslant 0. \tag{4.10}$$

Puisque : $P_\Omega(x_0 - \alpha\nabla\psi(x_0)) = x_0$.
Alors (4.10) devient

$$\langle -\alpha\nabla\psi(x_0), v\rangle \leqslant 0$$

$$-\alpha\langle\nabla\psi(x_0), v\rangle \leqslant 0$$

D'où $\langle\nabla\psi(x_0), v\rangle \geqslant 0$, puisque $\alpha > 0$ et par suite $\langle\nabla\psi(x_0), y - x_0\rangle$ et d'aprés lemme (4.1), la fonction ψ admet un minimum local en x_0.

Lemme 4.4 *Soit* $\widetilde{\Omega} = \{x \in \Omega; u \leqslant \psi(x) \leqslant v\}$.
Alors

$$\langle \nabla\psi(x), v_x \rangle \leqslant -\frac{1}{\alpha}\|v_x\|^2$$

avec

$$v_x = P_\Omega(x - \alpha\nabla\psi(x)) - x$$

si de plus si Ω *est borné et* $x \in \widetilde{\Omega}$ *alors :*
il existe γ_0 $(\gamma_0 > 0)$ *tel que :* $\|v_x\| \geqslant \gamma_0$.

Preuve : Soit $x \in \Omega$, utilisons l'inégalité du lemme (4.2) on substutions x par : $\big(x - \alpha\nabla\psi(x)\big)$ et y par x, on obtient :

$$\langle x - \alpha\nabla\psi(x) - P_\Omega(x - \alpha\nabla\psi(x), x - P_\Omega(x - \alpha\nabla\psi(x)) \rangle \leqslant 0$$

D'où

$$\langle -v_x - \alpha\nabla\psi(x), -v_x \rangle \leqslant 0$$

Soit

$$\langle v_x + \alpha\nabla\psi(x), v_x \rangle \leqslant 0$$

ce qui est équivalent à

$$\langle v_x, v_x \rangle + \langle \alpha\nabla\psi(x), v_x \rangle \leqslant 0$$

et par conséquent :

$$\langle \alpha\nabla\psi(x), v_x \rangle + \|v_x\|^2 \leqslant 0$$
$$\langle \alpha\nabla\psi(x), v_x \rangle \leqslant -\|v_x\|^2$$

Et par suite

$$\langle \nabla\psi(x), v_x \rangle \leqslant -\frac{1}{\alpha}\|v_x\|^2 \qquad (\alpha > 0)$$

Donc la première partie est démontrée.
Considérons :

$$x \longmapsto \|P_\Omega(x - \alpha\nabla\psi(x)) - x\| = \|v_x\|$$

C'est une fonction continue sur Ω et donc aussi sur $\widetilde{\Omega}$, qui est un compact si Ω est borné ; donc elle atteint sa borne inferieure γ_0 $(\gamma_0 > 0)$ sur $\widetilde{\Omega}$.
Et de plus si $\gamma_0 = 0$, alors il existe $x_0 \in \widetilde{\Omega}$ tel que $\|v_x\| = 0$ c'est-à-dire

$$P_\Omega(x_0 - \alpha\nabla\psi(x_0)) = x_0$$

Donc d'aprés lemme 4.3, ψ admet en x_0 un minimum local dans Ω. Or par définition même de $\widetilde{\Omega}$ ce minimum est hors $\widetilde{\Omega}$.

Lemme 4.5 *Soit Ω un convexe fermé et α un réel vérifiant* $0 < \alpha < \dfrac{1}{\gamma}$,

$$\textit{tel que} \quad v_x = P_\Omega(x - \alpha\nabla\psi(x)) - x \quad \textit{pour tout } x \textit{ de } \Omega.$$

1) Alors le passage de x à $\overline{x} = P_\Omega(x - \alpha\nabla\psi(x))$, ne se fait par croitre la fonction $\psi(x)$ de plus :

$$\psi(x) - \psi(\overline{x}) \geqslant (\frac{1}{\alpha} - \gamma)\|v_x\|^2$$

$$\big(\gamma = \max_i B_i\big)$$

2) Si en outre, Ω est borné et $x \in \widetilde{\Omega}$ $\big($ $\widetilde{\Omega}$ étant le convexe du lemme (4.4)$\big)$, l'inégalité précédente se transforme en :

$$\psi(x) - \psi(\overline{x}) \geqslant (\frac{1}{\alpha} - \gamma)\gamma_0^2$$

où γ_0 est la constante du lemme (4.4).

Preuve : 1) Prouvons la première partie.
Examinons d'abord le cas où $x \in \overset{\circ}{\Omega}$, on a alors :
i) $\overline{x} \in \Omega$
ii)$\psi(\overline{x}) = \psi(x) + \langle\nabla\psi(x), \overline{x} - x\rangle + \dfrac{1}{2}\langle\nabla^2\psi(x + \theta h), (\overline{x} - x)^2\rangle, \quad 0 \leqslant \theta \leqslant 1$
(dévloppement de Taylor avec reste de Lagrange).
Par suite :

$$\overline{x} - x = P_\Omega(x - \alpha\nabla\psi(x)) - x = v_x$$

D'où :

$$\begin{aligned}\langle\nabla\psi(x), \overline{x} - x\rangle &= \langle\nabla\psi(x), v_x\rangle \\ &\leqslant -\frac{1}{\alpha}\|v_x\|^2 \quad \text{voir lemme(4.4).}\end{aligned}$$

Et par conséquent :

$$\psi(\overline{x}) - \psi(x) \leqslant -\frac{1}{\alpha}\|v_x\|^2 + \frac{1}{2}(2\gamma\|v_x\|^2)$$

et par suite

$$\psi(\overline{x}) - \psi(x) \leqslant (\gamma - \frac{1}{\alpha})\|v_x\|^2$$

D'où :

$$\psi(x) - \psi(\overline{x}) \geqslant (\frac{1}{\alpha} - \gamma)\|v_x\|^2 \tag{4.11}$$

Donc la première partie est démontré.
2) c'est évident car d'aprés (4.2) et l'inégalité

$$\|v_x\| \geqslant \gamma_0 \quad \big(\textit{voir lemme } (4.4)\big)$$

on obtient :

$$\psi(x) - \psi(\overline{x}) \geqslant (\frac{1}{\alpha} - \gamma)\gamma_0^2$$

Preuve du théorème (4.4) : Premier cas $\big(\Omega$ est borné $\big)$.
Le convexe Ω est borné donc compact (Ω étant toujours le convexe fermé de $\mathbb{R}^n$).
Soit $x_1 \in \Omega$, α un réel vérifiant la condition $0 < \alpha < \dfrac{1}{\gamma}$.
Construisons la suite (x_n), alors d'aprés lemme (4.5)on a :

$$\psi(x_n) - \psi(x_{n+1}) \geqslant (\frac{1}{\alpha} - \gamma)\|x_n - x_{n+1}\|^2$$

Donc $(\psi(x_n))_n$ est décroissante et comme elle est minorée par $m = \inf\limits_{x\in\Omega} \psi(x)$ elle est convergente vers un nombre u, évidemment $u \geqslant m$ et par suite $\psi(x_n) \geqslant u \quad \forall n$
• Montrons que $u = m$
supposons que $u > m$
D'aprés les lemmes (4.4) et (4.5) :

$$\psi(x_k) - \psi(x_{k+1}) \geqslant (\frac{1}{\alpha} - \gamma)\gamma_0^2, \quad \forall k \geqslant 0$$

Ecrivons ces inégalités pour $k = 0, 1, ..., n$ et en faisant les additioner membre à membre on obtient :

$$\psi(x_0) - \psi(x_n) \geqslant n(\frac{1}{\alpha} - \gamma)\gamma_0^2, \quad \forall n \geqslant 0$$

C'est-à-dire

$$\psi(x_n) \leqslant \psi(x_0) - n(\frac{1}{\alpha} - \gamma)\gamma_0^2, \quad \forall n \geqslant 0$$

Pour n assez grand $\psi(x_0) - n(\frac{1}{\alpha} - \gamma)\gamma_0^2$ devient strictement inférieur à u ce qui contredit le fait que $\psi(x_n) \geq u$ Par conséquent ; $u = m$
c'est-à-dire : $\psi(x_n)$ tend vers $m = \inf\limits_{x\in\Omega} \psi(x) = \psi(\overline{x}_0)$ quand n tend vers $+\infty$

• Prouvons maintenant que : $x_n \longrightarrow x_0$
Soit $\varepsilon > 0$, $B(\overline{x}_0, \varepsilon)$ (la boule ouverte de centre $\overline{x}_0$ et de rayon ε)
Posons :

$$\Omega_\varepsilon = \Omega \cap B^C(\overline{x}_0, \varepsilon)$$

où $B^C(\overline{x}_0, \varepsilon)$ est le complémentaire de $B(\overline{x}_0, \varepsilon)$
Ω_ε est fermé comme intersection de deux fermés, et borné car il est contenu dans Ω qui est borné, et donc Ω_ε est compact.
Soit $m_\varepsilon = \inf\limits_{x\in\Omega_\varepsilon} \psi(x)$
Donc on a $m_\varepsilon > m$
Par suite $\psi(x_n) \geq m_\varepsilon > m$. Donc Ω_ε ne peut contenir qu'un nombre fini d'éléments de la suite (x_n) $\big($ si non $\psi(x_n)$ ne peut pas avoir m comme limite$\big)$
Par conséquent :
tous les x_n (à partir d'un certain rang) sont dans $B(\overline{x}_0, \varepsilon)$, ce qui signifier que

$$x_n \longrightarrow \overline{x}_0$$

Deuxième cas : $\big($$\Omega$ est un convexe fermé non borné $\big)$
Nous avons vu dans ce cas (démonstration th1) que
le min $\psi(x) = \psi(\overline{x})$ est atteint dans Ω_r $\big(\Omega_r = \Omega \cap C_r\big)$ avec $\big[C_r = B(x_0, r), r > \|\nabla\psi(x_0)\|\big]$ Alors le convexe Ω_r est un compact,
et si $x \notin \Omega_r$ $\psi(x) \geqslant \psi(x_1) \geqslant \psi(x_0)$ donc d'aprés lemme(4.5) ; la suite $(\psi(x_n))$ est décroissante ;donc $\forall k$ $x_k \in \Omega_r$,
d'où :

$$\begin{aligned} x_{k+1} &= P_\Omega\big(x_k - \alpha\nabla\psi(x)\big) \\ &= P_{\Omega_r}\big(x_k - \alpha\nabla\psi(x)\big) \end{aligned}$$

On se retrouve dans le premier cas avec Ω_r convexe fermé et borné et la suite (x_n) défini par :

$$\begin{cases} x_0 \text{ un point donnu de } \Omega \\ x_{n+1} = P_\Omega\big(x_n - \alpha \nabla \psi(x_n)\big) \ n \geqslant 0 \end{cases}$$

Chapitre 5

décomposition d'une fonction quadratique sous sa forme canonique

5.1 optimisation de la fonction f.

Soit a résoudre le problème suivant :

$$\begin{cases} f(x) = \displaystyle\sum_{i=1}^{n} \alpha_i x_i + \beta_i x_i^2, \\ \displaystyle\max_{x\in\Omega} f(x); \quad \alpha_i, \beta_i \in \mathbb{R} \quad \text{pour tout} i. \end{cases} \tag{5.1}$$

où $\Omega = \{x \in \mathbb{R}^n : x \geqslant 0 \ et \ Ax \leqslant b\}$, A une matrice de type $m \times n$ et b un vecteur de $\mathbb{R}_+^m$.
On peut se restreindre a l'étude du cas suivant :

$$f(x) = \sum_{i=1}^{n} \alpha_i x_i + \beta_i x_i^2, \text{ où}$$

les coefficients α_i et β_i sont des réels quelconques et la fonction f n'est convexe ni concave.
commençons par voir comment optimiser la fonction f pour cela posons :
$\varphi(x) = \sum \alpha_i x_i + \beta_i x_i^2$, où la somme porte sur tous les indices i pour $\beta_i < 0$;
et
$\psi(x) = \sum \alpha_i x_i + \beta_i x_i^2$; où la somme porte sur tous les indices i pour $\beta_i \geqslant 0$

Aprés décomposition ,la fonction f devient comme suit :

$$f(x) = \varphi(x) + \psi(x)$$

φ est strictement concave, et ψ est convexe.
On va montrer que la résolution du problème (5.1) revient à résoudre les problèmes (5.2) et (5.3) suivant :

$$\begin{cases} \psi(x) = \sum_i \alpha_i x_i + \beta_i x_i^2, \\ \quad \max_{x\in\Omega} \psi(x); \ \alpha_i \in \mathbb{R}, \beta_i \in \mathbb{R}_+. \end{cases} \tag{5.2}$$

$$\begin{cases} \varphi(x) = \sum_i \alpha_i x_i + \beta_i x_i^2, \\ \quad \max_{x\in\Omega} \varphi(x); \ \alpha_i \in \mathbb{R}, \beta_i < 0. \end{cases} \tag{5.3}$$

On sait déjà résoudre le problème (5.2) et (5.3) voir [chapitres 2 et 4] Posons :

$$\begin{cases} \max_{x\in\Omega} \psi(x) & = & \psi(\overline{x}_\psi); \\ \max_{x\in\Omega} \varphi(x) & = & \varphi(\overline{x}_\varphi) \end{cases}$$

Lemme 5.1 *Soit toujours Ω le convexe fermé défini par :*

$$\Omega = \{x \in \mathbb{R}^n : x \geqslant 0 \text{ et } Ax \leqslant b\}$$

Si $(\overline{x}_\varphi + \overline{x}_\psi) \in \Omega$, *alors :*

$$\max_{x\in\Omega} f(x) = f(\overline{x}_\varphi + \overline{x}_\psi)$$

Preuve : En effet : $f(x) = \varphi(x) + \psi(x)$,
Donc on a :

$$\max_{x\in\Omega} f(x) \leqslant \max_{x\in\Omega} \varphi(x) + \max_{x\in\Omega} \psi(x) = \varphi(\overline{x}_\varphi) + \psi(\overline{x}_\psi)$$

Mais $(\overline{x}_\varphi + \overline{x}_\psi) \in \Omega$ et par suite

$$f(\overline{x}_\varphi + \overline{x}_\psi) = \varphi(\overline{x}_\varphi) + \psi(\overline{x}_\psi)$$

d'où $\max_{x\in\Omega} f(x) = f(\overline{x}_\varphi + \overline{x}_\psi)$

Remarque 5.1 *pour la suite on pose :* $\max_{x\in\Omega} f(x) = f(\overline{x})$, *et* $\overline{x} = (\overline{x}_1, \overline{x}_2, ..., \overline{x}_n)$

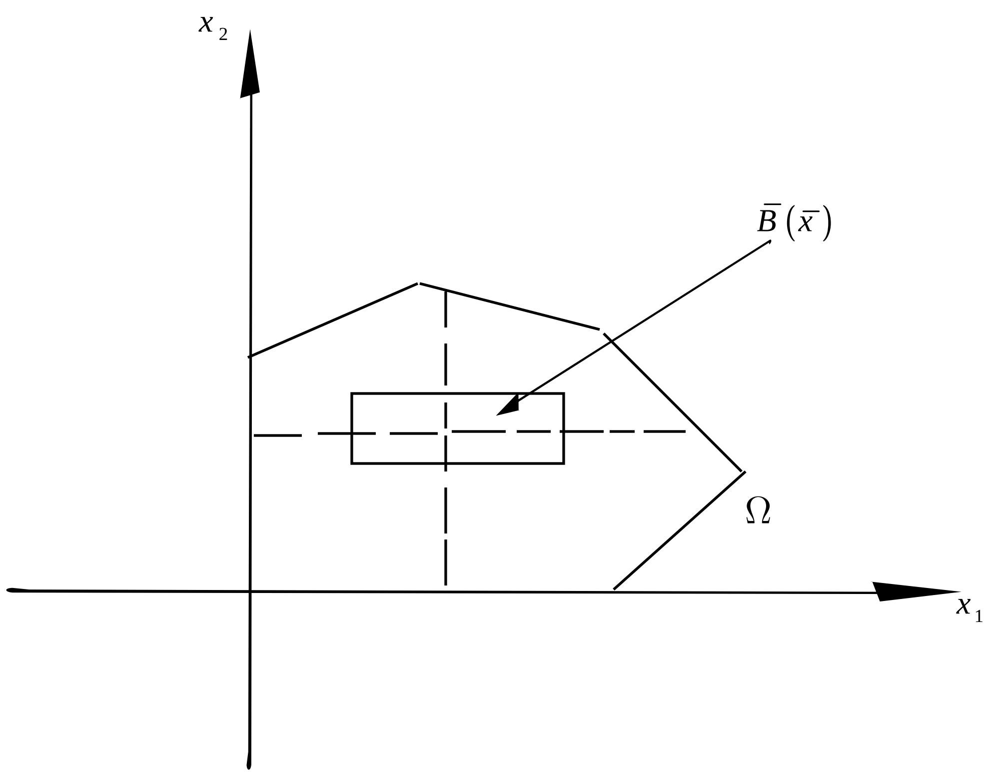
Lemme 5.2 *Si* $\overline{x} \in \overset{\circ}{\Omega}$, *alors :* $\max_{x\in\Omega} f(x) = \varphi(\overline{x}_\varphi)$

Preuve : Si $\overline{x} \in \overset{\circ}{\Omega}$, alors il existe une boule fermée de centre $\overline{x}$ et incluse dans $\overset{\circ}{\Omega}$:
$\overline{B}(\overline{x}) = \{x \in \Omega : |\overline{x}_i - x_i| \leqslant \varepsilon_i, \quad \varepsilon_i > 0 \ \ \textit{pour tout } i = 1, 2, ..., n\}$.
Supposons qu'il existe un $\beta_{i_0} \geqslant 0$
Alors l'application qui à $x_i \to \psi_{i_0}(x_i) = \alpha_{i_0} x_i + \beta_{i_0} x_i^2$ est convexe pour tout $x_i \in \mathbb{R}_+$.
Posons $I_{i_0} = [\overline{x}_{i_0} - \varepsilon_{i_0}, \overline{x}_{i_0} + \varepsilon_{i_0}]$

$$\psi_{i_0} : I_{i_0} \to \mathbb{R}$$

$$x_i \to \psi_{i_0}(x_i),$$

$\max_{I_{i_0}} \psi_i(x_i) = \max\{\psi_{i_0}(\overline{x}_{i_0} - \varepsilon_{i_0}), \psi_{i_0}(\overline{x}_{i_0} + \varepsilon_{i_0})\}$



donc

$$\psi_{i_0}(\overline{x}_{i_0}) < \max\{\psi_{i_0}(\overline{x}_{i_0} - \varepsilon_{i_0}), \psi_{i_0}(\overline{x}_{i_0} + \varepsilon_{i_0})\} = \psi_{i_0}(\overline{x}_{i_0}(\varepsilon_{i_0}))$$

où $\overline{x}_{i_0}(\varepsilon_{i_0}) = \overline{x}_{i_0} - \varepsilon_{i_0}$ où bien $\overline{x}_{i_0}(\varepsilon_{i_0}) = \overline{x}_{i_0} + \varepsilon_{i_0}$
avec $\overline{x}_{i_0}(\varepsilon_{i_0}) \in [\overline{x}_{i_0} - \varepsilon_{i_0}, \overline{x}_{i_0} + \varepsilon_{i_0}]$
Posons :
$\widetilde{x} = (\overline{x}_1, \overline{x}_2, ..., \overline{x}_{i_0-1}, \overline{x}_{i_0}(\varepsilon_{i_0}), \overline{x}_{i_0+1}, ..., \overline{x}_n) \in \overline{B}(\overline{x})$ et $f(\widetilde{x}) = \varphi(\widetilde{x}) + \psi(\widetilde{x})$.
Comme $\psi_{i_0}(\overline{x}_i) < \psi(\overline{x}_{i_0}(\varepsilon_{i_0}))$, *on a* $\psi(\overline{x}) < \psi(\widetilde{x})$.
D'où la contradiction $f(\widetilde{x}) > f(\overline{x}) = \max\limits_{x\in\Omega} f(x)$.
Donc la supposition est fausse et par suite on a :
$\beta_i < 0 \quad i = 1, 2, ..., n,\ f(x) = \varphi(x)$ et $\max\limits_{x\in\Omega} f(x) = \max\limits_{x\in\Omega} \varphi(x) = \varphi(\overline{x}_\varphi)$.

Lemme 5.3 *Si $\overline{x} \notin \overset{\circ}{\Omega}$, alors l'hyperplan contenant $\overline{x}$ contient $\overline{x}_\psi$.*

Preuve : Si $\overline{x} \notin \overset{\circ}{\Omega}$, alors il existe un hyperplan $\langle a_{i_0}, x\rangle = b_{i_0}$ contenant $\overline{x} : \langle a_{i_0}, \overline{x}\rangle = b_{i_0}$.
$\overline{x}_\psi$ est un point extrême voir [25] ;
si $\langle a_{i_0}, \overline{x}_\psi\rangle < b_{i_0}$ posons $\Omega_{i_0} = \{x \in \mathbb{R}^n_+ : \langle a_{i_0}, x\rangle \leqslant b_{i_0}\}$; Ω_{i_0} est un convexe fermé et $\overline{x}_\psi$ est un point intérieur de Ω_{i_0}.
Le même raisonnement que le lemme (5.2) montre qu'il existe $\widetilde{x}_\psi$ tel que $\psi(\overline{x}_\psi) < \psi(\widetilde{x}_\psi)$. D'où la contradiction avec $\psi(\overline{x}_\psi) = \max\limits_{x\in\Omega} \psi(x)$.
Donc La supposition est fausse, et on a $\langle a_{i_0}, \overline{x}_\psi\rangle = b_{i_0}$.
On a alors le théorème suivant :

Théorème 5.1 *Sous les hypothèses ci dessus, on a, voir [13] :*
(1) Si $(\overline{x}_\varphi + \overline{x}_\psi) \in \Omega$, alors $\overline{x} = \overline{x}_\varphi + \overline{x}_\psi$;
(2) Si $(\overline{x}_\varphi + \overline{x}_\psi) \notin \Omega$, et $\overline{x} \in \overset{\circ}{\Omega}$, alors $\overline{x} = x_\varphi$;
(3) Si $(\overline{x}_\varphi + \overline{x}_\psi) \notin \Omega$, et $\overline{x} \notin \overset{\circ}{\Omega}$, alors $\overline{x}$ est un point frontière de Ω.

5.2 Algorithme de recherche de l'optimum de f

(1) On calcule $\max\limits_{x\in\Omega} \varphi(x)$;
(2) On calcule $\max\limits_{x\in\Omega} \psi(x)$;
(3) Si $(\overline{x}_\varphi + \overline{x}_\psi) \in \Omega$, alors $\overline{x} = x_\varphi + x_\psi$;

(4) Sinon, au sommet x_ψ, f s'écrit sous la forme :
$f(x) = \sum_i \alpha'_i x_i + \beta'_i x_i^2 + \sum \sum_{i \neq j} \gamma'_{ij} x_i x_j + \psi(x_\psi)$.
Soit j_0 un vecteur entrant $\beta_{j_0} < 0$.
($\star$) Si $\triangle_{j_0} = \alpha'_{j_0} x_{j_0} + \beta'^2_{j_0} x^2_{j_0} > 0$, alors :
$f(x) = \sum_i \alpha''_i x_i + \beta'_i x_i^2 + \sum \sum_{i \neq j} \gamma''_{ij} x_i x_j + \psi(x_\psi) + \triangle_{j_0}$.
on pose :
$x(\lambda) = \lambda x_\psi + (1-\lambda) x(j_0)$ où $x(j_0)$ est le sommet voisin de x_ψ, on calcule $\max\limits_{0 \leqslant \lambda \leqslant 1} f(x(\lambda))$, et on pose :

$$f(\overline{x}) = \max\{\max_{0 \leqslant \lambda \leqslant 1} f(x_\lambda), \psi(x_\psi) + \triangle_{j_0}\},$$

s'il existe un j_1 un vecteur entrant : $\beta_{j_1} < 0$ implique ($\star$)
Sinon stop.

Exemple 5.1

$$f(x_1, x_2, x_3) = 2x_1 + 3x_2 + 3x_3 + x_1^2 + 2x_2^2 - x_3^2,$$

$$s.q.c : \qquad \begin{cases} x_1 + 3x_2 \leqslant 18 \\ x_1 + x_2 + x_3 \leqslant 8 \\ 2x_1 + x_2 + 2x_3 \leqslant 14. \end{cases}$$

D'aprés des calculs simples on trouve que :

$$\begin{cases} x_\varphi = 3/2 \\ x_\psi = (0.6) \end{cases}$$

D' autre part on a $(x_\psi, x_\varphi) \in \Omega$ *ce qui implique que*

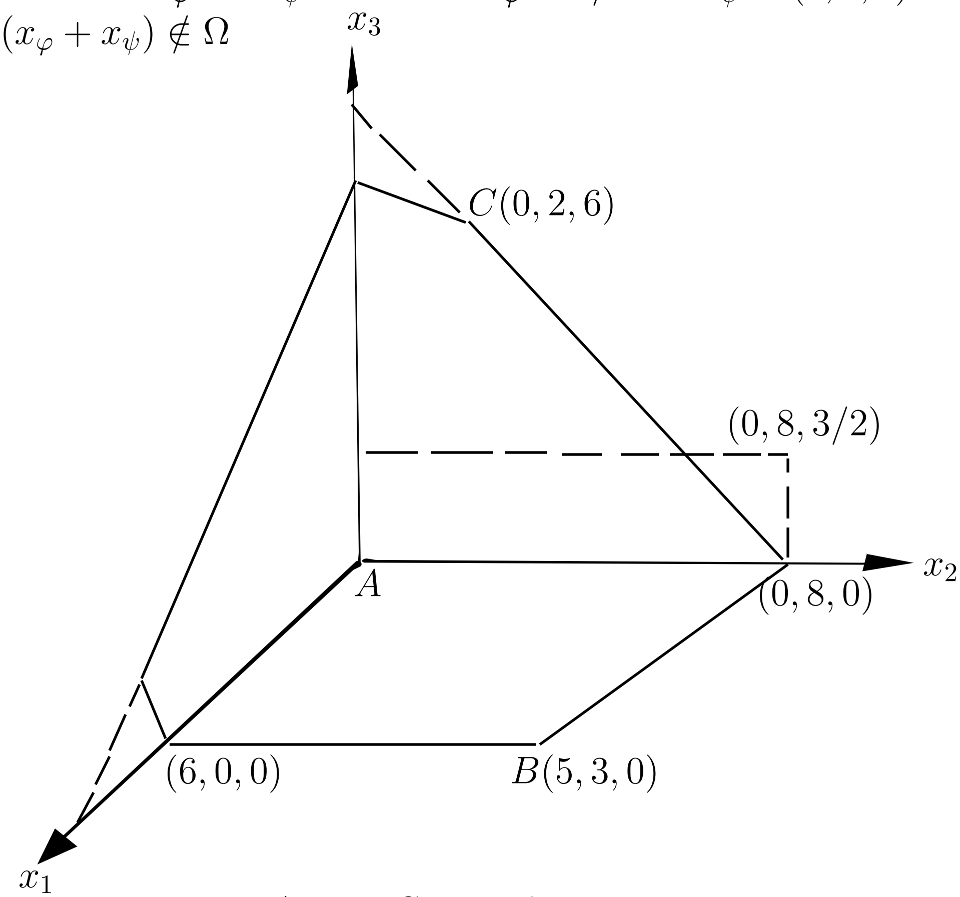
$$\max f(x) = f(0, 6, 3/2) = 92.25$$

Exemple 5.2

$$f(x_1, x_2, x_3) = 2x_1 + 3x_2 + 3x_3 + x_1^2 + 2x_2^2 - x_3^2,$$

$$s.q.c : \qquad \begin{cases} 3x_1 + x_2 \leqslant 18 \\ x_1 + x_2 + x_3 \leqslant 8 \\ 2x_1 + x_2 + 2x_3 \leqslant 14. \end{cases}$$

Aprés le calcul de x_φ *et* x_ψ *on trouve* $x_\varphi = 3/2$ *et* $x_\psi = (0,8,0)$ *et par suite* $(x_\varphi + x_\psi) \notin \Omega$



Les sommets voisins sont A,B et C ;mais le vecteur entrant est x_3 *et la direction d'amélioration est le sommet* $C = (0,2,6)$
$\lambda(0,2,6) + (1-\lambda)(0,8,0) = (0, 8-6\lambda, 6\lambda); 0 \leqslant \lambda \leqslant 1$

$$\begin{aligned} f(0,8-6\lambda,6\lambda) &= \psi(0,8-6\lambda) + \varphi(6\lambda) \\ &= 3(8-6\lambda) + 2(8-6\lambda)^2 + 3(6\lambda) - 3(6\lambda)^2 \\ &= 24 + 128 - 18\lambda - 96\lambda + 18\lambda - 108\lambda^2 + 72\lambda^2 \\ &= 152 - \varepsilon(\lambda) \ \ \textit{avec}\ \varepsilon(\lambda) \geqslant 0 \end{aligned}$$

Donc $\quad \max_{x \in \Omega} f(x) = f(x_\psi) = f(0,8,0) = \psi(x_\psi) = 152$
Avec $\Omega = \{3x_1 + x_2 \leqslant 18, x_1 + x_2 + x_3 \leqslant 8, 2x_1 + x_2 + 2x_3 \leqslant 14\}$

Remarque 5.2 *On a* $\quad P_\Omega(0,8,3/2) = (29/4, 3/4, 0)$
Donc $\quad f(0,3/4,29/4) < 152.$
Le point $(0,3/4,29/4)$ *est procche du point* $x_\psi + x_\varphi = (0,8,3/2)$ *,mais ça ne donne pas le maximum de la fonction objectif* f*.*

5.3 Décomposition d'une fonction quadratique sous sa forme canonique

Considérons le problème suivant :

$$(P): \begin{cases} f(x) = \sum_{i=1}^{n} \alpha_i x_i + \beta_i x_i^2 + \sum_{i \neq j} \sum \gamma_{ij} x_i x_j, \\ \alpha_i \in \mathbb{R},\ \beta_i \in \mathbb{R},\ \gamma_{ij} \in \mathbb{R} \\ Ax \leqslant b, \qquad x \geqslant 0, \end{cases}$$

Soit $\gamma_{i_0 j_0}$ une valeur de $\mathbb{R}$ non nulle,
$\big(x_{i_0} + x_{j_0}\big)^2 = x_{i_0}^2 + x_{i_0}^2 + 2x_{i_0} x_{j_0}$.
Alors $\quad x_{i_0} x_{j_0} = \dfrac{1}{2}\big\{\big(x_{i_0} + x_{j_0}\big)^2 - x_{i_0}^2 - x_{i_0}^2\big\}$
$\gamma_{i_0 j_0} x_{i_0} x_{j_0} = \dfrac{1}{2}\big\{\gamma_{i_0 j_0}\big(x_{i_0} + x_{j_0}\big)^2 - \gamma_{i_0 j_0} x_{i_0}^2 - \gamma_{i_0 j_0} x_{i_0}^2\big\}$
comme $x_{i_0}, x_{j_0} \geqslant 0$, alors $x_{i_0} + x_{j_0} \geqslant 0\,$;
posons alors $x_{i_0} + x_{j_0} = x_{i_0 j_0}$
$\gamma_{i_0 j_0} x_{i_0 j_0} = \frac{1}{2}\gamma_{i_0 j_0} x_{i_0 j_0}^2 - \frac{1}{2}\gamma_{i_0 j_0} x_{i_0}^2 - \frac{1}{2}\gamma_{i_0 j_0} x_{j_0}^2$.
f devient alors :

$$f(x) = \sum_{i=1}^{n} \alpha_i x_i + \sum_{i=1}^{n} \beta_i x_i^2 + \sum_{i \neq j} \sum \frac{1}{2}\gamma_{ij} x_{ij}^2 - \sum_{i \neq j} \sum \frac{1}{2}\gamma_{ij} x_i^2 - \sum_{i \neq j} \sum \frac{1}{2}\gamma_{ij} x_j^2$$

Donc le problème (p) sera remplacé par le problème (p') équivalent :

$$f(x) = \sum_{i=1}^{n} \alpha_i x_i + \sum_{i=1}^{n} \beta_i x_i^2 + \frac{1}{2} \sum_{i \neq j} \sum \big(\gamma_{ij} x_{ij}^2 - \gamma_{ij} x_i^2 - \gamma_{ij} x_j^2\big)$$

$$(P'): \begin{cases} f(x) = \sum_{i=1}^{n} \alpha_i x_i + \beta_i x_i^2 + \frac{1}{2} \sum_{i \neq j} \sum \gamma_{ij} \big(x_{ij}^2 - x_i^2 - x_j^2\big) \\ \alpha_i\ \beta_i\ \gamma_{ij}\ \in \mathbb{R}; \\ Ax \leqslant 0 \quad x \geqslant 0 \\ x_i + x_j - x_{ij} = 0 \end{cases}$$

Exemple 5.3 *Maximisons $f(x)$ tel que :*

$$f(x) = 2x_1 + 3x_2 + x_1^2 + 2x_2^2 - 2x_1 x_2$$

Sous les contraintes :

$$\begin{cases} x_1 + 3x_2 \leqslant 18 \\ x_1 + x_2 \quad \leqslant 8 \\ 2x_1 + x_2 \leqslant 14 \end{cases}$$

Soit $\Omega = \{x_1 + 3x_2 \leqslant 18,\ x_1 + x_2 \quad \leqslant 8\ 2x_1 + x_2 \leqslant 14\}$
Posons $x_1 + x_2 = x_{1,2} = x_{12} = x_3$
$f(x_1, x_2, x_3) = 2x_1 + 3x_2 + x_1^2 + 2x_2^2 - \frac{2}{2}x_3^2 - \frac{2}{2}x_1^2 - \frac{2}{2}x_2^2$
$= 2x_1 + 3x_2 + x_2^2 - x_3^2$
Donc on va maximiser :

$$f(x_1, x_2, x_3) = 2x_1 + 3x_2 + x_2^2 - x_3^2$$

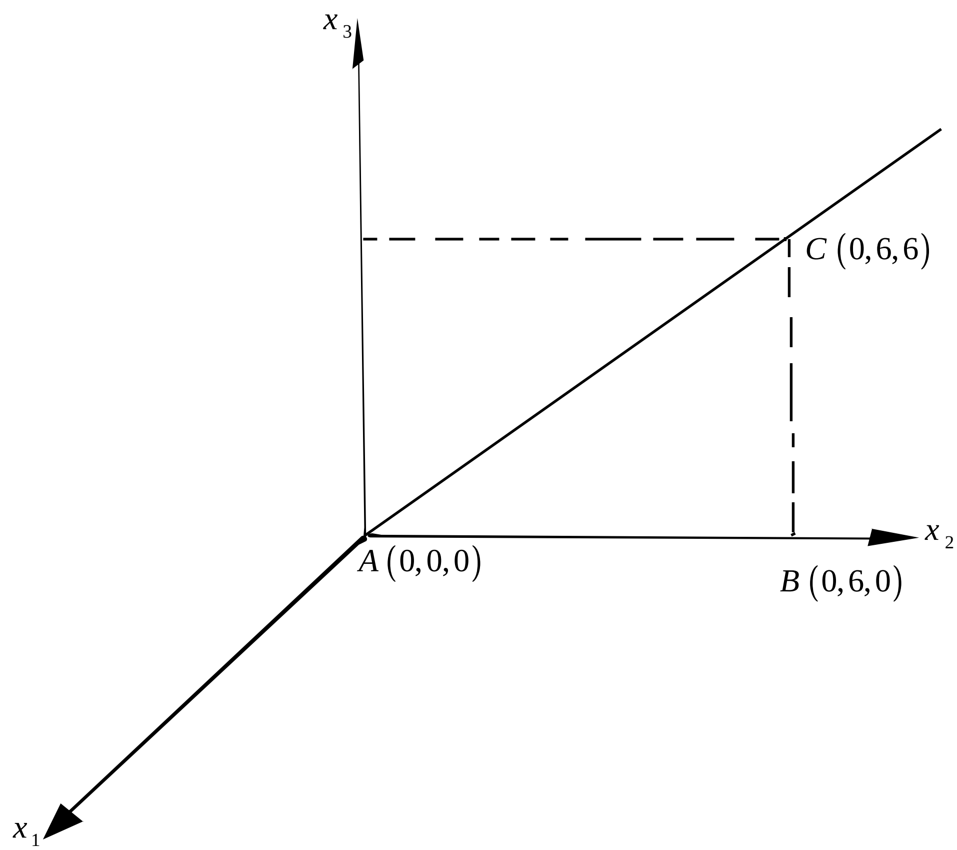
Sous les contraintes :

$$\begin{cases} x_1 + 3x_2 \leqslant 18 \\ x_1 + x_2 \quad \leqslant 8 \\ 2x_1 + x_2 \leqslant 14 \\ x_1 + x_2 \quad = x_3 \end{cases}$$

Soit $\Omega' = \{x_1 + 3x_2 \leqslant 18,\ x_1 + x_2 \quad \leqslant 8\ 2x_1 + x_2 \leqslant 14, x_1 + x_2 \quad = x_3\}$
On a $\varphi(x_3) = -x_3^2$ *ce qui implique que* $\max \varphi(x) = \varphi(0) = 0$
$\psi(x_1, x_2) = 2x_1 + 3x_2 + x_2^2$ *on voit que* $\max \psi(x_1, x_2) = \psi(0, 6) = 54$

Puisque $\Omega' \subset \Omega$ *il résulte :*

$$\max_{x \in \Omega'} \psi(x_1, x_2) \leqslant \max_{x \in \Omega} \psi(x_1, x_2)$$



Et comme

$$\begin{cases} (0, 6, 6) \in \Omega' \\ x_1 = 0, x_2 = 6, x_3 = 6 \end{cases}$$

ça implique $\max_{x \in \Omega'} \psi(x) = \psi(0, 6, 6)$

Mais $x_\psi + x_\varphi = (0, 6, 0) + (0, 0, 0) \notin \Omega'$

Donc $f\left[\lambda(0, 6, 6) + (1 - \lambda)0\right] = 18\lambda$

D'où $\max_{0 \leqslant \lambda \leqslant 1} f(\lambda) = 18$ *et* $\overline{x} = 1(0, 6, 6) + (1 - 1)0 = (0, 6, 6)$

5.4 CONCLUSION GENERALE

Ce rapport de mémoire a été l'occasion de présenter un cadre général de la programmation quadratique. Ce cadre a ensuite donné lieu à la présentation détaillée d'une nouvelle méthode de résolution des problèmes quadratiques dont la fonction objectif est convexe ou non convexe.

Nous avons présenté certaines méthodes de calcul de la solution optimale d'un programme quadratique.

Les problèmes des flots en réseau à coût minimum peuvent être modélisés par des fonctions quadratiques. Parmi les problèmes classiques associés à cette formulation on peut citer : l'équilibre du trafic en réseaux de communication (de trafic urbain, télécommunication), le routage en réseau, l'affectation du trafic...

Les algorithmes de résolution des problèmes d'optimisation quadratiques sont trés complexes et trés lourds à mettre en oeuvre pour les problèmes de grande taille ou pour la résolution en temps réel.

Des techniques de linéarisation, des versions spécialisées de l'algorithme de Franck-Wolfe, les méthodes Newton ou quasi-Newton, sont les algorithmes les plus employés jusqu'a maintenant pour la résolution de ces problèmes quadratiques.

Des formes particulières de ce problème ont été étudiées pour réduire le temps du calcul. L'un de ces cas est le problème des flots à coût minimum avec des coût quadratiques, plus précisément avec des coût quadratiques séparables.

En développant plusieurs exemples à l'aide de notre récente, nous avons montré la simplicité des calculs et surtout la précision des résultats obtenus aprés un nombre d'itérations trés limité.

Nous avons explicité les conditions d'utilisation de cette méthode en démontrant l'existence des solutions et les techniques permettant de les calculer :

a) *Aprés la décomposition de la fonction objectif en deux fonctions, l'une concave* φ *et l'autre convexe* ψ *; si le point critique appartient à* Ω *(ensemble des solutions réalisables), la solution optimale de* f *est la solution optimale de* φ.
b) Si le point critique n'appartient pas à Ω *alors la solution optimale de* f*, est la projection du point critique sur* Ω' .
c) L'extrémum de la fonction convexe ψ *est un sommet ; la technique du calcul est développée dans le chapitre (2).*
d) Nous avons aussi donné la technique du calcul de la valeur de la fonction ψ *lors du passage d'un sommet à un sommet voisin.*
e) En passant d'un sommet à un sommet voisin, le double produit apparaît dans l'expression de la fonction, et ceci nécessite une technique pour le calculer. Celle-ci a été présentée dans le chapitre(2).

Le théorème reste valable si le convexe Ω *est remplacé par un convexe fermé et borné inférieurement dans* $\mathbb{R}^n$*, il permet de résoudre les problèmes quadratiques dont les contraintes ne sont pas nécessairement linéaires.*
f) Contrairement aux autres méthodes analytiques donnant la solution exacte en passant par le processus de linéarisation ou les méthodes numériques fournissant des solutions approchées voir [14].

Les exemples suivants ont été développés par d'autres méthodes et celle présentée dans ce mémoire pour mieux apprécier les différences.Prennons l'exemple suivant :

$$\begin{array}{ll} max & f(x_1, x_2) = 5x_1 + 8x_2 - x_1^2 - 2x_2^2 \\ s.c & \begin{cases} 3x_1 + 2x_2 \leqslant 6 \\ x_1, x_2 \geqslant 0 \end{cases} \end{array}$$

qui a été traité par la méthode de Franck-Wolf et par cette méthode.
La complexité de la méthode de Franck-Wolf :

- *Initialisation de la méthode,*
- *Itération k : boucle d'évaluation des coefficients* c_j
- *Détermination de la direction d'amélioration,*
- *Calcul du pas,*
- *Exécution du critère d'arrêt.*

De plus la solution obtenue n'est qu'approximative. Dans notre cas $x^* = \left(\frac{5}{6}, \frac{7}{6}\right)$ *et la valeur de la fonction à l'optimum est :*
$max\ f(x_1, x_2) = f\left(\frac{5}{6}, \frac{7}{6}\right) = 10.94295 \leq 11.50$

Quand à la méthode du chapitre (4) , on calcule la valeur du point extrémal x^* *s'il appartient à* Ω *stop ; c'est la solution optimale.*
Sinon, la solution exacte est la projection de x^* *sur* Ω *qui donne :*

$$x^* = \left(\frac{5}{2}, 2\right) \in \Omega, \text{ le calcul de } \overline{x} \text{ nous donne } \overline{x} = \left(1, \frac{3}{2}\right) \text{ d'où}$$
$$\max_{x \in \Omega} \varphi(x) = 11.50.$$

Concernant la méthode de linéarisation ; l'exemple ci-aprés :

$$max\ f(x_1, x_2) = 5x_1 + 5x_2 - x_1^2 - x_2^2$$
$$s.c \quad \begin{cases} x_1 + 2x_2 - 8 \leqslant 6 \\ 3x_1 + x_2 - 9 \leqslant 6 \\ x_1, x_2 \geqslant 0 \end{cases}$$

La résolution de ce problème a nécessité l'ajout de 8 variables, 3 tableaux du simplexe de cinq lignes et de 12 colonnes, on obtient la solution exacte $\overline{x} = (2.2, 2.4) \qquad f(\overline{x}) = f(2.2, 2.4) = 12.4.$
Quand à méthode précédente, on calcule la valeur du point extrémal x^* *s'il appartient à* Ω *stop ; c'est la solution optimale. Sinon, la solution exacte est la projection de* x^* *sur* Ω *qui donne :*

$$x^* = \left(\frac{5}{2}, \frac{5}{2}\right) \in \Omega, \text{ le calcul de } \overline{x} \text{ nous donne } \overline{x} = (2.2, 2.4) \text{ d'où}$$
$$\max_{x \in \Omega} \varphi(x) = 12.40.$$

De plus la méthode présentée dans ce mémoire donne une interprétation géométrique de la solution optimale $\overline{x}$ *: c'est la projection de* x^* *sur* Ω *alors que la méthode de linéarisation ne situe pas la solution optimale par rapport à l'ensemble convexe des solutions réalisables.*

Notons enfin que pour le cas convexe suivant :

$$min\ f(x_1, x_2) = -20x_1 + 10x_2 + 3x_1^2 + 2x_2^2$$
$$s.c \quad \begin{cases} 2x_1 - x_2 \leqslant 6 \\ -x_1 + x_2 \leqslant 10 \\ 2x_1 + 3x_2 \leqslant 8 \\ x_1, x_2 \geqslant 0 \end{cases}$$

La méthode de linéarisation a nécessité l'ajout de 10 variables, 5 tableaux du simplexe, pour seulement 2 variables de décision, ce qui montre la lourdeur de cette méthode.

Ceci nous donne une brève idée ,imaginons la lourdeur des calculs si le problème comporte 50 variables de décision.

D'un point de vue algorithmique, notons que la recherche de la solution est simple et dispense des lourds calculs tels les conditions de Kuhn, du lagrangien et le calcul des matrice hessienne.

Les travaux présentés dans ce document permettent de traiter toutes les situations modélisables sous forme de problèmes de programmation quadratique.

Bibliographie

[1] S. Achmanov, *Programmation linéaire*, Edition Mir, Moscou, 1984.

[2] G. Allaire, *Analyse numérique et optimisation*, Ellipse, 2005.

[3] G. Allaire, *Introduction à l'analyse numérique matricielle et à l'optimisation*, Masson, 1998.

[4] Pierre-André cornillon, *Régression Théorie et applications* , Spring, 2006.

[5] J. P. Aubin, *Explicit Methods of Optimization*, BORDAS, Paris, 1984.

[6] A. Auslender, *Optimisation Méthodes numériques*, Masson, 1979.

[7] A. Belabbaci, *La programmation quadratique et son application à la programmation séparable*, Mémoire de magister, 2007.

[8] Frédéric. Bonnans, *Optimisation continue , cours et problèmes corrigés*, Dunod, 2006.

[9] Stephen. Boyed and Liven. Vandenberghe, *Convex optimization*, Cambridge.

[10] Pierre. Brémaud, *Initiation aux probabilités et aux chaines de Markov*, Springer, 1984.

[11] Haim. Brezis, *Analyse fonctionnelle, Théorie et applications*, Dunod , 1999.

[12] Phlippe. Caron, Alain. juhel, *Programmation linéaire Méthodes et Applications*,Dunod.

[13] A. Chikhaoui, B.Djebbar, A. belabbaci, A.Mokhtar, *Optimization of a quadratic function under its canonical form*, Jornal of Applied Sciences, 499-510 ISSN 1996-3343, 2009.

[14] Daniel De Wolf, *Recherche opérationnelle*, 2006.

[15] Jean. Dieudonné, *Algébre linéaire et géométrie élémentaire*, Herman, 1964.

[16] Jean. Dieudonné, *Calcul infinitésimal*, Hermane, 1997.

[17] P. Dimitri. Bertsekas, *Convex analysis and optimization*, Athena Scientific, 2002.

[18] Marco Dozzi, *Méthodes non linéaires et stochastiques de la recherche opérationnelle*, 2004.

[19] J. F. Hêche, Thomas.M.Liebling, *Recherche opérationnelle pour ingénieurs*, 1 Presses Polythechnique et Universitaire, Romandes.

[20] J. F. Hêche, Thomas. M. Liebling, *Recherche opérationnelle pour ingénieurs*, 2 Presses Polythechnique et Universitaire, Romandes.

[21] Lars. Hormander, *Notions of Convexity*, Birkhauser Boston, 1994.

[22] Robert. Faure , *Précis de recherche opérationnelle Méthodes et exercices d'application*, Dunod, 1996.

[23] M. Simmonard, *Programmation linéaire, 1 Fondements*, Dunod, Paris, 1972.

[24] M. Simmonard, *Programmation linéaire, 2 Extentions*, Dunod, Paris, 1973.

[25] G. R. Walsh, *Methods of Optimization*, Printed page Bros, Ltd Mile Crois Lavie, Norwich, 1979.