

REPUBLIQUE ALGERIENNE DEMOCRATIQUE ET POPULAIRE
MINISTERE DE L'ENSEIGNEMENT SUPERIEUR ET DE LA RECHERCHE SCIENTIFIQUE
UNIVERSITE DES SCIENCES ET DE LA TECHNOLOGIE D'ORAN «MOHAMED BOUDIAF»

FACULTE DES SCIENCES
DEPARTEMENT D'INFORMATIQUE

THESE

PRESENTEE PAR
Ahmed CHIKHAOU

POUR L'OBTENTION DU GRADE DE

DOCTORAT EN SCIENCES

SPECIALITE : **INFORMATIQUE**

***Algorithmes d'Optimisation
des Problèmes de Formes Quadratiques***

MEMBRES DU JURY

Président :	Mr. AEK BENYETTOU	Professeur, U.S.T.O (Oran)
Rapporteur :	Mr. Bachir DJEBBAR	Professeur, U.S.T.O (Oran)
Examineurs :	Mr. Bouziane BELDJILLAL	Professeur, ESSENIA (Oran)
	Mr. AEK MOKHTARI	Professeur, Université de Laghouat
	Mme Yamina LASKRI	Maître de Conférences, Université de Annaba
	Mr. Kaddour SAADOUNI	Maître de Conférences, U.S.T.O (Oran)

Remerciements

Je tiens en premier lieu à exprimer ma sincère gratitude à Bachir Djebbar pour avoir accepté de me diriger, je le remercie aussi pour ses conseils, ses encouragements, et son enthousiasme, qui m'ont guidé et motivé tout au long de mon travail.

Je remercie sincèrement les membres du jury Bouziane Beldjillali, Yamina Laskri, Abdelkader Mokhtari et Kadour Sadouni pour avoir accepté de rapporter mon travail.

Que Mr Abdelkader Benyattou trouve ici l'expression de mon profond respect pour avoir accepté de présider mon jury de soutenance de thèse.

Je remercie également Youcef Bellabaci pour son apport tout au long des démonstrations et de la rédaction de la thèse.

Ma plus grande reconnaissance va aux personnes qui ont toujours été disponibles, pour leurs précieux conseils et l'encouragement constant qu'elles ont toujours su envelopper dans un style souple et mobilisateur.

Je remercie tout le personnel administratif du département d'informatique pour leur soutien logistique.

Table des matières

Introduction	4
1. Ensembles convexes	8
1.1 Notions fondamentales	8
1.1.1 Formulation	8
1.1.2 Caractéristiques des ensembles convexes	10
1.1.3 Cônes convexes	16
1.2 Ensembles convexes polyédriques	19
1.2.1 Combinaison linéaire convexe	19
1.2.2 Enveloppes Coniques	25
1.3 Structure des ensembles admissibles	29
1.3.1 Etude systématique	29
1.3.2 Systèmes homogènes	34
1.4 Equivalence de deux définitions d'un ensemble polyédrique Convexe	42
2. Fonctions convexes	47
2.1 Définitions	47
2.2 L'hyperplan tangent à l'hypersurface	50
2.3 Minimum et maximum d'une fonction convexe	51
3. Méthodes de programmation non linéaires	55
3.1 Optimisation d'une fonction convexe ψ	55
3.2 Cas où la fonction à optimiser est de la forme $f(x) = \sum_{i=1}^n \alpha_i x_i + \sum_{i,j}^n \beta_{ij} x_i^2$	62
3.3 Méthode de Franck-Wolfe	64
3.4 Application de la méthode du simplexe pour la résolution de certains problèmes quadratiques	67

4. Résolution d'un problème quadratique non convexe et sous contraintes linéaires.	76
4.1 Introduction	76
4.2 Décomposition de la fonction objectif	77
4.3 Optimisation de la fonction concave φ	78
4.4 Nouvelle forme de la fonction ψ	83
4.5 Méthode pour calculer la forme quadratique.	84
Conclusion	92
Annexe	96
Implémentation	110
Bibliographie	116

INTRODUCTION

L'optimisation est la science de la recherche de la meilleure décision parmi plusieurs solutions possibles dans un environnement complexe d'un domaine de la vie.

L'approche systématique de la prise de décision est composée de trois étapes.

L'objet de la première étape consiste à développer un modèle mathématique représentant le problème considéré.

On commence par identifier les variables de décision ; puis à recueillir les données pertinentes ; la formulation de la fonction objectif à optimiser ; et arranger les variables et les données dans un ensemble de relations mathématiques, telles des équations ou des inéquations, appelées contraintes.

Dans la seconde étape on analyse le modèle mathématique et on sélectionne une technique appropriée pour la recherche de la solution optimale.

L'analyse du programme non linéaire fournit de précieux renseignements sur la structure du problème et répond aux questions relatives à l'existence et la caractérisation de l'ensemble des solutions réalisables et des solutions optimales.

La troisième étape consiste en la recherche de la solution optimale.

Le processus se termine par l'interprétation de la solution optimale dans le sens de la prise de décision, et finalement l'évaluation des aspects pratiques de la convergence, l'efficacité, et l'analyse d'erreur de la technique utilisée.

Un nombre important des problèmes de l'industrie, de l'économie et des finances, des différents types de réseaux ... peuvent avoir une formulation de programmes quadratiques.

Dans les réseaux de télécommunication une fonction objectif couramment utilisée pour calculer des chemins optimaux au travers d'un réseau informatique est le nombre de routeurs traversés. La fonction objectif considérée associe à chaque chemin dans le réseau le nombre de routeurs que ce chemin traverse. Le but est de trouver un chemin qui minimise cette fonction.

Le transport aérien, joue un rôle majeur dans le développement économique globalisé que connaît actuellement le monde. Au cours des prochaines décennies, le trafic aérien mondial, s'il accompagne la tendance actuelle pourrait doubler. La programmation quadratique pourrait contribuer à l'amélioration des techniques de gestion du trafic des avions sur les plates-formes aéroportuaires aussi bien sur le moyen-long terme, que sur le court terme.

Développons deux exemples numériques simples pour illustrer notre propos.

Une entreprise produisant du savon tente de minimiser les coûts perdus dus aux retards de livraison car lors d'un retard de livraison l'entreprise est pénalisée de deux manières: les coûts dus aux manques à gagner, et les coûts dus aux pénalités de retards imposés par les clients concernés par ces retards.

On note: C les coûts totaux perdus lors de retards de livraison.

x_1 , la variable désignant les manques à gagner.

x_2 , la variable désignant les pénalités de retard.

Suite à une modélisation économétrique, le problème obtenu est formulé comme suit:

$$\begin{cases} \min f(x_1, x_2) = -3x_2 - 2x_1^2 + 2x_1x_2 \\ x_1 + x_2 \leq 2 \\ 2x_1 - x_2 \leq 1 \\ x_1, x_2 \geq 0 \end{cases} .$$

Un investisseur souhaite retenir P projets parmi n . Chaque projet a le même coût et a un rendement représenté par une variable aléatoire. L'espérance de rendement e_i ($i = 1, \dots, n$), la variance σ_i ($i = 1, 2, \dots, n$), et la covariance associée à chaque paire de projets, σ_{ij} ($i = 1, 2, \dots, n-1$; $j = 1, 2, \dots, n$) sont connues; elles sont estimées, par exemple, à partir d'un historique des rendements observés.

Le problème consiste à choisir P projets de façon à obtenir une espérance de rendement global supérieure ou égale à une valeur fixée ρ tout en minimisant le risque, c'est-à-dire la variance de ce rendement. Le programme mathématique convexe en variables bivalentes associé est :

$$(PI) : \min \left\{ r(x) = \frac{1}{p^2} \sum_{i=1}^n \sum_{j=1}^n \sigma_{ij} x_i x_j : \sum_{i=1}^n x_i = p, \sum_{i=1}^n e_i x_i \geq p\rho \right\};$$

La fonction objectif de (PI) est quadratique et les contraintes sont linéaires.

Schématiquement, l'investisseur souhaite constituer un portefeuille de projets et arbitrer entre les gains et les risques. Dans ce problème d'investissement, il doit s'assurer le risque minimal à rendement donné.

La matrice Hessienne de la fonction économique de (PI) n'est autre que la matrice de variance-covariance : il est donc semi-défini positif. Ainsi, $r(x)$ est une fonction convexe.

Les méthodes d'optimisation des fonctions quadratiques utilisées jusqu'ici sont soit exactes ou approximatives. [21] a transformé le problème quadratique initial en un problème linéaire

équivalent pour lui appliquer la méthode du simplexe. [1] donne seulement des solutions approximatives en construisant des séries convergentes qui approchent la solution réelle. [13] transforme le problème initial en un problème linéaire en ajoutant un nombre important de contraintes. [15] considère un algorithme du Lagrangien étendu. [20] a aussi commencé par rendre le problème linéaire en ajoutant beaucoup de contraintes afin de le résoudre par la méthode des tableaux du simplexe. [22] a implémenté plusieurs méthodes d'optimisation, il recherche l'optimum global en comparant différentes méthodes.

Le but de cette thèse, *la contribution à la résolution des problèmes quadratiques non convexes et sous contraintes linéaires* est de proposer une nouvelle méthode exacte, originale et efficace de résolution des programmes quadratiques sous contraintes linéaires. Ces problèmes sont généralement non convexes.

La méthode présentée dans cette thèse détermine la solution optimale exacte du problème, sans avoir à ajouter de variables supplémentaires. Le principe de cette technique est de projeter le point critique sur un nouveau convexe construit à partir de l'ensemble des solutions réalisables.

Un second ingrédient est l'étude de la fonction convexe. Les solutions optimales sont des points extrêmes qui sont toujours des sommets. Il suffit seulement de se déplacer d'un sommet à un autre et de choisir le sommet qui produit la solution optimale.

En se déplaçant d'un sommet à l'un de ses voisins, la forme de la fonction convexe change et de nouveaux termes contenant le double produit apparaissent dans l'expression de la fonction objectif voir [4]; nous donnons alors une technique simple à programmer, permettant de calculer ces coefficients.

Il est à mentionner ici que la méthode est générale pour n'importe quel problème d'optimisation quadratique.

Dans le chapitre 1, nous présentons le cadre général de l'optimisation quadratique et les fondements théoriques des ensembles réalisables. Il s'agit surtout de formuler et de démontrer des résultats importants relatifs aux ensembles convexes qui nous permettront par la suite d'étudier les problèmes de programmation quadratique.

Dans le chapitre 2 je résume les principaux résultats sur les fonctions convexes nécessaires aux démonstrations des résultats trouvés.

Le chapitre 3 fournit les fondements de la programmation non linéaire ; on applique notre méthode au cas d'une fonction convexe. Des méthodes exactes et approximatives de linéarisation sont développées.

Dans le chapitre 4, je développe le résultat original de ma contribution à la résolution des problèmes quadratiques non convexes et sous contraintes linéaires et qui consiste à trouver la solution exacte d'un problème de programmation quadratique avec des contraintes linéaires et une fonction objectif quadratique écrite sous sa forme canonique.

Cette étude décrit une nouvelle méthode basée sur la décomposition de la fonction objectif en une somme de deux fonctions l'une convexe et l'autre concave ; un nouvel ensemble réalisable est construit par une transformation homographique, de telle sorte que la projection du point critique de la fonction objectif sur cet ensemble, donne la solution exacte du problème étudié.

Notons qu'on n'a pas besoin de transformer le problème quadratique en un équivalent linéaire comme dans les méthodes numériques; la méthode est purement analytique et évite le problème épineux du choix de la solution initiale.

Enfin l'algorithme et des exemples sont traités en annexe.

Chapitre I

Ensembles Convexes

1.1 Notions fondamentales :

1.1.1 Formulation.

On va formuler et démontrer des résultats importants relatifs à un ensemble convexe quelconque qui nous permettront par la suite d'étudier les problèmes de programmation quadratique.

Définition 1.1. Un ensemble X est convexe si, chaque fois qu'il contient deux points x^1, x^2 quelconques, il contient également tous les points de la forme :

$$x = \alpha x^1 + (1 - \alpha)x^2, \quad 0 \leq \alpha \leq 1 \quad (1.1)$$

La droite passant par les points x^1, x^2 a pour équation paramétrique

$$x(\alpha) = x^2 + (x^1 - x^2)\alpha, \quad (1.2)$$

$x^1 - x^2$ étant son vecteur directeur. Lorsque $\alpha = 0$, on a $x(0) = x^2$, et $\alpha = 1$, on a $x(1) = x^1$.

Quand α varie entre 0 et 1, le point $x(\alpha)$ parcourt tout le segment compris entre x_1 et x_2 .

Définition 1.2. Un ensemble X 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.

Donnons quelques exemples d'ensembles convexes dans le plan.

Rectangle $X = \{(x_1, x_2) \mid 0 \leq x_1 \leq a, 0 \leq x_2 \leq b\}$ (Fig1.1).

Montrons que X est convexe. Soient $(x_1, x_2), (x'_1, x'_2)$ deux points du rectangle X , α un nombre arbitraire tel que $0 \leq \alpha \leq 1$.

Vérifions que le point de coordonnées $(\alpha x_1 + (1-\alpha)x'_1, \alpha x_2 + (1-\alpha)x'_2)$ appartient à X .

En d'autres termes, il faut tester les inégalités

$$0 \leq \alpha x_1 + (1-\alpha)x'_1 \leq a, \quad 0 \leq \alpha x_2 + (1-\alpha)x'_2 \leq b.$$

Comme $0 \leq x_1 \leq a, 0 \leq x'_1 \leq a$, on les multiplie par les nombres non négatifs α et $1-\alpha$ respectivement et on fait la somme, il vient

$$0 \leq \alpha x_1 + (1-\alpha)x'_1 \leq \alpha a + (1-\alpha)a = a.$$

On procède de même en ce qui concerne l'autre inégalité.

Disque $X = \{(x_1, x_2) \mid x_1^2 + x_2^2 \leq R^2\}$ (Fig 1.2)

Démontrons la convexité de X . Soit $(x_1, x_2) \in X, (x'_1, x'_2) \in X \quad 0 \leq \alpha \leq 1$.

Regardons si $(\alpha x_1 + (1-\alpha)x'_1, \alpha x_2 + (1-\alpha)x'_2)$ est un élément de X .

On effectue les transformations simples suivantes :

$$(\alpha x_1 + (1-\alpha)x'_1)^2 + (\alpha x_2 + (1-\alpha)x'_2)^2 = \alpha^2(x_1^2 + x_2^2) + 2\alpha(1-\alpha)(x_1x'_1 + x_2x'_2) + (1-\alpha)^2((x'_1)^2 + (x'_2)^2)$$

et on utilise l'inégalité fondamentale du produit scalaire

$$x_1x'_1 + x_2x'_2 \leq \sqrt{x_1^2 + x_2^2} \sqrt{(x'_1)^2 + (x'_2)^2}.$$

On substitue R^2 à $x_1^2 + x_2^2$ et $(x'_1)^2 + (x'_2)^2$ de toutes les expressions ci-dessus (ce qui renforce l'inégalité), il vient

$$(\alpha x_1 + (1-\alpha)x'_1)^2 + (\alpha x_2 + (1-\alpha)x'_2)^2 \leq R^2[\alpha^2 + 2\alpha(1-\alpha) + (1-\alpha)^2] = R^2(\alpha + (1-\alpha))^2 = R^2$$

i.e., le disque est un ensemble convexe.

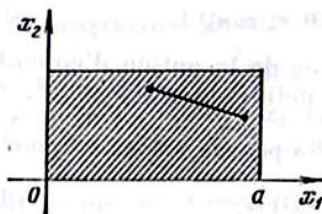


Fig. 1.1

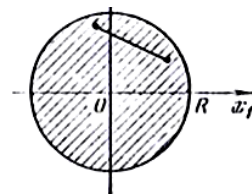


Fig. 1.2

Théorème 1.1. L'intersection d'ensembles convexes est convexe.

Démonstration.

Soit $X = X_1 \cap X_2$ avec X_1, X_2 des ensembles convexes, et soient x et 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 $X_1 \subset \mathbb{R}^n, X_2 \subset \mathbb{R}^n$ l'ensemble $X = \{x \mid x = x_1 + x_2, x_1 \in X_1, x_2 \in X_2\}$, obtenu par addition deux à deux des éléments de X_1 et X_2 .

Théorème 1.2. La somme formée d'ensembles convexes est convexe.

Démonstration.

Soient $x, \bar{x} \in X$, où $X = X_1 + X_2$.

Il existe dans X_1 et X_2 des éléments tels que $x = x^1 + x^2, \bar{x} = \bar{x}^1 + \bar{x}^2, x^1, \bar{x}^1 \in X_1, x^2, \bar{x}^2 \in X_2$.

Etant donné un nombre arbitraire $\alpha, 0 \leq \alpha \leq 1$, on a $\alpha x^1 + (1-\alpha)\bar{x}^1 \in X_1, \alpha x^2 + (1-\alpha)\bar{x}^2 \in X_2$.

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

$\alpha x^1 + (1-\alpha)\bar{x}^1 + \alpha x^2 + (1-\alpha)\bar{x}^2 = \alpha(x^1 + x^2) + (1-\alpha)(\bar{x}^1 + \bar{x}^2) = \alpha x + (1-\alpha)\bar{x} \in X$. (voir [23]).

Un ensemble convexe dans l'espace $\mathbb{R}^n$ est par exemple

$$\Omega = \{x \in \mathbb{R}^n : Ax \leq b, x \geq 0, b \geq 0, A \text{ une matrice}\}.$$

1.1.2 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 $\|x_0 - y_0\| = \inf_{y \in A} \|x_0 - y\|$

Preuve.

i) Existence :

Posons $d = \inf_{y \in A} \|y - x_0\| > 0$. Soit (y_n) , une suite de H telle que $\|x_0 - y_n\| \leq d + \frac{1}{n}$.

Montrons que (y_n) est une suite de Cauchy

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

D'après la loi du parallélogramme, on a

$$\begin{aligned} \|(y_n - x_0) - (y_m - x_0)\|^2 &= (\|(y_n - x_0)\|^2 + \|(y_m - x_0)\|^2) * 2 - \|(y_n - x_0) + (y_m - x_0)\|^2 \\ &= 2\|(y_n - x_0)\|^2 + 2\|(y_m - x_0)\|^2 - 4\left\|\frac{y_n + y_m}{2} - x_0\right\|^2 \end{aligned}$$

Mais $\frac{y_n + y_m}{2} \in A$ car A est convexe, donc $\left\|\frac{y_n + y_m}{2} - x_0\right\|^2 \geq d^2$ d'où

$$\|y_n - y_m\|^2 \leq 2\left(d + \frac{1}{m}\right)^2 + 2\left(d + \frac{1}{n}\right)^2 - 4d^2 = \frac{4}{m}d + \frac{2}{m^2} + \frac{4}{n}d + \frac{2}{n^2} < \varepsilon \quad \text{si } n, m > n_0(\varepsilon)$$

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

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

$$d \leq \|x_0 - y_0\| = \|x_0 - y_n + y_n - y_0\| \leq \|x_0 - y_n\| + \|y_n - y_0\| < d + \frac{1}{n} + \varepsilon \quad \text{car } y_n \rightarrow y_0 \quad y_0 \in A.$$

$$\Rightarrow \|x_0 - y_0\| < d + \frac{1}{n} + \varepsilon, \quad \forall n > n_0, \quad \forall \varepsilon > 0 \quad \text{d'où } \|x_0 - y_0\| = d.$$

iii) Unicité de y_0 . Supposons qu'il existe $z_0 \in A$; $d = \|x_0 - z_0\|$

$$\begin{aligned} \|z_0 - y_0\|^2 &= \|(z_0 - x_0) - (y_0 - x_0)\|^2 = 2\|(z_0 - x_0)\|^2 + 2\|(y_0 - x_0)\|^2 - \|z_0 + y_0 - 2x_0\|^2 \\ &= 4d^2 - 4\left\|\frac{z_0 + y_0}{2} - x_0\right\|^2 \leq 4d^2 - 4d^2 = 0 \end{aligned}$$

$$\Rightarrow \|z_0 - y_0\|^2 = 0 \quad \Rightarrow \quad z_0 = y_0 \quad \text{c.q.f.d.}$$

Remarque : si H est un espace de BANACH, l'unicité n'est pas vraie.

Théorème 1.4. Soit f une fonction convexe définie sur un convexe F de $\mathbb{R}^n$.

$f : F \rightarrow \mathbb{R}$; x son minimum local:

$f(x) = \inf_{z \in B(x,r)} f(z)$. Alors x est un minimum global. C'est-à-dire

$$f(x) = \inf_{z \in F} f(z).$$

En effet, supposons qu'il existe $y \in F$: $f(y) < f(x)$ pour $y \notin B(x,r)$

$$z = \lambda y + (1-\lambda)x \text{ où } \lambda = \frac{r}{2\|y-x\|} < 1 \quad \text{car } y \notin B(x,r)$$

$z \in F$ et F convexe.

$$f(z) < (1-\lambda)f(x) + \lambda f(y) = f(x) + \lambda[f(y) - f(x)] < f(x).$$

Mais $z \in B(x,r)$ ($\|z-x\| = \frac{r}{2} < r$) contradiction. (Voir [6])

Les ensembles convexes sont essentiellement caractérisés par la propriété de séparation.

Soit, dans le plan, un convexe fermé X et un point $d \notin X$ (fig. 1.3).

Il existe une droite $ax_1 + bx_2 = c$ telle que X et $d = (d_1, d_2)$ soient de part et d'autre de la droite, i.e. tout point $(x_1, x_2) \in X$ vérifie $ax_1 + bx_2 \leq c$, mais $ad_1 + bd_2 > c$.

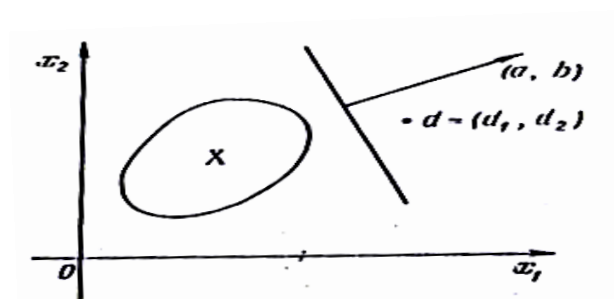


Fig. 1.3

S'agissant d'un ensemble convexe dans l'espace n -dimensionnel, on énonce et démontre cette propriété moyennant une notion analogue à la notion d'une droite dans le plan.

Définition 1.4. Un ensemble de points $x = (x^1, x^2, \dots, x^n)$ de coordonnées telles que $p_1 x^1 + p_2 x^2 + \dots + p_n x^n = c$, $p \neq 0$ ($\langle p, x \rangle = c$), s'appelle **hyperplan** de R^n .

Le vecteur $P = (p_1, p_2, \dots, p_n)$ est la normale à cet hyperplan (avec $P \neq 0$).

Théorème 1.5. (De séparation).

Soit X un ensemble convexe fermé, $a \notin X$, $a = (a_1, a_2, \dots, a_n)$. Il existe un hyperplan de normale $p \neq 0$ tel que $\langle p, a \rangle > c$ et on ait $\langle p, x \rangle \leq c$ pour tout $x \in X$.

Démonstration.

Soit la fonction $\varphi(x) = (x_1 - a_1)^2 + (x_2 - a_2)^2 + \dots + (x_n - a_n)^2$ de la variable indépendante $x = (x_1, x_2, \dots, x_n)$ sur l'ensemble X . Elle est évidemment continue. Aux termes 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 $\bar{x}$ quelconque de X et une boule U de centre a et dont le carré du rayon vaut $\varphi(\bar{x})$ (fig.1.4).

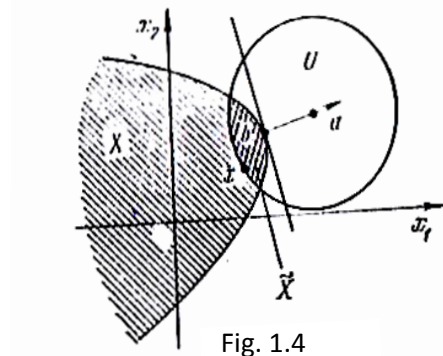


Fig. 1.4

L'intersection $\tilde{X} = U \cap X$ est un ensemble fermé et borné. La fonction $\varphi(x)$ atteint donc son minimum en un point $b \in \tilde{X}$. Il est clair que b réalise le minimum de $\varphi(x)$ sur X tout entier (voir le théorème 1.4).

Ainsi, $\varphi(x) \geq \varphi(b)$ quel que soit $x \in X$. On fixe un point quelconque $x \in X$ et on examine le segment $x(\alpha) = \alpha x + (1 - \alpha)b$, $0 \leq \alpha \leq 1$.

L'ensemble X étant convexe, on a $x(\alpha) \in X$ pour tout α , $0 \leq \alpha \leq 1$, donc $\varphi(x(\alpha)) \geq \varphi(b)$.

Par conséquent, $\frac{\varphi(x(\alpha)) - \varphi(b)}{\alpha} \geq 0$ pour tout α , $0 < \alpha \leq 1$. La fonction $\varphi(x(\alpha))$ est dérivable par rapport à α en tant que superposition de deux fonctions dérivables $\varphi(x)$ et $x(\alpha)$. Aussi la limite $\lim_{\alpha \rightarrow 0} \frac{\varphi(x(\alpha)) - \varphi(b)}{\alpha} = \varphi'_\alpha(x(0)) \geq 0$ existe.

On calcule directement la dérivée $\varphi'_\alpha(x(0))$ à l'aide de la règle de dérivation d'une fonction composée :

$$\varphi'_\alpha(x(\alpha)) = \left(\sum_{i=1}^n (x_i(\alpha) - a_i)^2 \right)' = 2 \sum_{i=1}^n (x_i(\alpha) - a_i) x'_i(\alpha).$$

Puisque $x'_i(\alpha) = (\alpha x_i + (1-\alpha)b_i)' = x_i - b_i$ et $x_i(0) = b_i$, on a

$$\varphi'_\alpha(x(\alpha)) = 2 \sum_{i=1}^n (x_i(\alpha) - a_i)(x_i - b_i),$$

$$\varphi'_\alpha(x(0)) = 2 \sum_{i=1}^n (b_i - a_i)(x_i - b_i).$$

D'après les résultats démontrés ci-dessus, cette dérivée est non négative, i.e.

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

On introduit les notations $p_i = a_i - b_i$, $i = 1, 2, \dots, n$; $p = (p_1, p_2, \dots, p_n)$

$$c = \sum_{i=1}^n b_i(a_i - b_i) \quad \text{et on récrit l'inégalité obtenue :}$$

$$\sum_{i=1}^n p_i x_i \leq c \quad (\langle p, x \rangle \leq c).$$

On note que le vecteur $p = (p_1, p_2, \dots, p_n)$ et le nombre c sont indépendants du vecteur choisi x de X . Aussi, il faut vérifier que $p \neq 0$ et $\langle p, a \rangle > c$ pour constater que toutes les affirmations du théorème sont justes pour p et c .

En effet, $p = a - b$ et $p \neq 0$ car dans le cas contraire $a = b$, ce qui est impossible puisque $b \in X$, $a \notin X$.

Enfin,

$$\langle p, a \rangle - c = \sum_{i=1}^n (a_i - b_i)a_i - \sum_{i=1}^n b_i(a_i - b_i) = \sum_{i=1}^n (a_i - b_i)^2 > 0$$

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

Il y a lieu de dire que l'hyperplan $\langle p, x \rangle = c$ construit lors de la démonstration du théorème présente la propriété de rencontrer l'ensemble X au point b : on vérifie aisément l'égalité $\langle p, b \rangle = c$.

On introduit donc la définition suivante.

Définition. 1.5. On dit que l'hyperplan $\langle p, x \rangle = c$ est un **hyperplan d'appui** à l'ensemble X en un point $b \in X$ si l'on a pour tout élément $x \in X$ les relations $\langle p, x \rangle \leq c$ et $\langle p, b \rangle = c$.

Ainsi, le théorème de séparation dit que l'ensemble convexe fermé X est séparé de tout point $a \notin X$ par un hyperplan d'appui.

Si le convexe X n'est pas fermé et si $a \notin X$ et $a \in \bar{X}$, $\bar{X}$ étant la fermeture de X , le théorème de séparation est en défaut. Il suffit cependant de corriger un peu l'énoncé, à savoir on exige $\langle p, a \rangle \geq c$ au lieu de $\langle p, a \rangle > c$. Cela découle du résultat auxiliaire suivant :

La fermeture d'un ensemble convexe est convexe.

En effet, soient $x, \bar{x} \in \bar{X}$, et $x = \lim_{k \rightarrow \infty} x^k, x^k \in X, \bar{x} = \lim_{k \rightarrow \infty} \bar{x}^k, \bar{x}^k \in X$.

Vu la convexité de X , on a $\alpha x^k + (1 - \alpha) \bar{x}^k \in X$,

et $\lim_{k \rightarrow \infty} (\alpha x^k + (1 - \alpha) \bar{x}^k) = \alpha x + (1 - \alpha) \bar{x}$ qui appartient à $\bar{X}$ du moment que $\bar{X}$ est un fermé.

Soit $a \notin X$, on démontre aisément qu'il existe alors une suite $a^k, k = 1, 2, \dots, a^k \notin \bar{X}, \lim_{k \rightarrow \infty} a^k = a$.

On applique le théorème de séparation à $\bar{X}$ et à a^k et on obtient l'hyperplan $\langle p^k, x \rangle = \langle p^k, b^k \rangle$.

Prenons l'hyperplan parallèle passant par a^k : $\langle p^k, x \rangle = \langle p^k, a^k \rangle$.

On estime sans restreindre la généralité que $\|p^k\| = 1, k = 1, 2, \dots$, et que par la suite des vecteurs p^k converge vers un $p \neq 0$. On a pour tout k et tout $x \in X$ l'inégalité $\langle p^k, x \rangle \leq \langle p^k, a^k \rangle$ car $\langle p^k, b^k \rangle \leq \langle p^k, a^k \rangle$.

On passe à la limite pour $k \rightarrow \infty$, il vient $\langle p, x \rangle \leq \langle p, a \rangle$ pour tout $x \in \bar{X}$, en particulier, pour $x \in X$.

On a donc le théorème de séparation pour les ensembles convexes non fermés.

Théorème 1.6. Soit X un ensemble convexe (non nécessairement fermé). On suppose que a n'est pas un point intérieur de X . Il existe un hyperplan $\langle p, x \rangle = c$, $p \neq 0$, tel que $\langle p, a \rangle = c$ et $\langle p, x \rangle \leq c$ pour tout $x \in X$. Voir [2])

On demande souvent que l'hyperplan sépare deux ensembles convexes. Le théorème ci-dessous en affirme la possibilité.

On dit que l'ensemble $R_+^n = \{x \mid x \in R^n, x \geq 0\}$ des vecteurs $x = (x_1, x_2, \dots, x_n)$ de coordonnées non négatives : $x_i \geq 0, i = 1, 2, \dots, n$, est l'**orthant positif** de l'espace R^n .

Théorème 1.7. Soit X un ensemble convexe qui ne contient pas de points intérieurs de l'orthant positif R_+^n (i.e. aucun point de X n'a toutes ses coordonnées positives). Il existe un hyperplan $\langle p, x \rangle = 0$ de normale $p \geq 0$, $p \neq 0$, tel que $\langle p, x \rangle \leq 0$ pour tout $x \in X$.

Démonstration.

Soit l'ensemble $M = X - R_+^n$. 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) de l'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 $x \in X, v \in R_+^n$, auquel cas $x = u + v > 0$ contradiction avec l'hypothèse du théorème.

D'après le théorème 1.4, on peut affirmer qu'il existe un hyperplan $\langle p, x \rangle = 0$, $p \neq 0$ passant par l'origine tel que tout point $u \in M$ vérifie l'inégalité $\langle p, u \rangle \leq 0$.

On en conclut moyennant la définition de M que tout $x \in X$ et $v \in R_+^n$, satisfont à $\langle p, x \rangle \leq \langle p, v \rangle$.

On pose $v = 0$, il vient $\langle p, x \rangle \leq 0$ quel que soit $x \in X$.

On fixe un point $x \in X$ et on a $\langle p, v \rangle \geq \langle p, x \rangle$ pour tout $v \in R_+^n$. D'où évidemment $p \geq 0$.

En effet, si $p_i < 0$ et si l'on choisit pour v un point de la forme $(0, 0, \dots, 0, \lambda, 0, \dots, 0)$, avec $\lambda > 0$ à la place i , on obtient $\lambda p_i \geq \langle p, x \rangle$. Or, on peut rendre λp_i négatif aussi petit que l'on veut, inférieur en tous cas à $\langle p, x \rangle$ fixé.

1.1.3 Cônes convexes :

Définition 1.6. Un ensemble $K \subseteq R^n$ s'appelle **cône convexe** si, toutes les fois qu'il contient des points x', x'' quelconques, il contient tous les points de la forme $x = \alpha x' + \beta x''$, $\alpha \geq 0, \beta \geq 0$.

Il en résulte de suite que, premièrement, le cône convexe K est un ensemble convexe et, deuxièmement, si K contient le vecteur x , il contient également la demi-droite $\alpha x, \alpha \geq 0$.

Evidemment, tout cône convexe K renferme l'origine des coordonnées.

Théorème 1.8. Tout hyperplan d'appui à un cône convexe K passe par l'origine des coordonnées.

Démonstration.

Supposons que l'hyperplan $\langle p, x \rangle = c$ est d'appui à K en un point $b \in K$: quel que soit $x \in K$, $\langle p, x \rangle \leq c$, $\langle p, b \rangle = c$. Prenons un point $\alpha b, \alpha \geq 0$.

On a $\alpha b \in K$ par définition d'un cône, donc $\langle p, \alpha b \rangle \leq c$.

Comme $\langle p, \alpha b \rangle = \alpha \langle p, b \rangle = \alpha c$, l'inégalité $\alpha c \leq c$ a lieu pour tout $\alpha \geq 0$.

D'où $(\alpha - 1)c \leq 0$. On a $c \geq 0$ et $c \leq 0$ pour $\alpha = 0$ et $\alpha = 2$ respectivement.

Ainsi $c = 0$, l'équation de l'hyperplan d'appui s'écrit $\langle p, x \rangle = 0$ et le point 0 est dans cet hyperplan, c.q.f.d.

La notion de cône convexe joue un rôle absolument exceptionnel dans les problèmes de programmation mathématique. Cette classe particulière d'ensembles convexes présente plusieurs caractères spécifiques.

Voyons trois exemples représentés par les fig. 1.5 à 1.7. Dans le premier cas, K est un cône convexe sans point commun avec l'intérieur de R_+^2 .

Le théorème 1.7 entraîne que K et R_+^2 peuvent être séparés par un hyperplan dont la normale a plusieurs coordonnées non négatives.

En outre, on peut choisir l'hyperplan séparant de façon que sa normale soit un vecteur positif.

Plaçons-nous dans le cas 1.6. L'ensemble convexe X est un disque de rayon unité et de centre $(-1,0)$. Le seul hyperplan qui sépare X et R_+^2 possède la normale $(1,0)$, si bien que l'hyperplan séparant de normale positive n'existe pas.

Le cône convexe de la fig. 1.7 est le demi-plan de gauche qui contient le point 0 sans contenir les autres points de l'axe $0x_2$. De même que K de la fig. 1.5, ce cône ne renferme aucun point de l'orthant non négatif R_+^2 à part le point 0 .

Mais il est impossible de le séparer de R_+^2 par l'hyperplan de normale positive. Le fait est qu'il n'est pas fermé.

L'étude des propriétés des cônes convexes utilise la définition suivante.

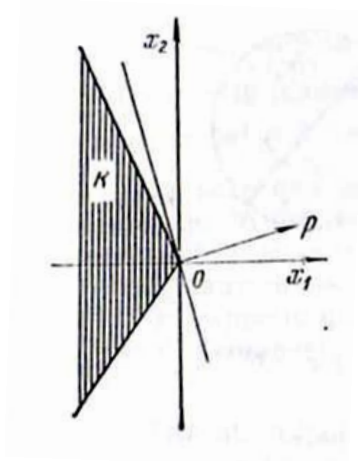


Fig. 1.5.

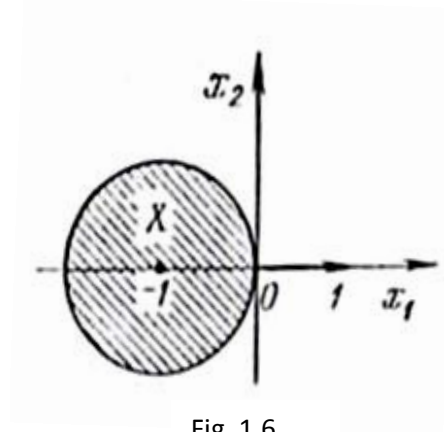


Fig. 1.6.

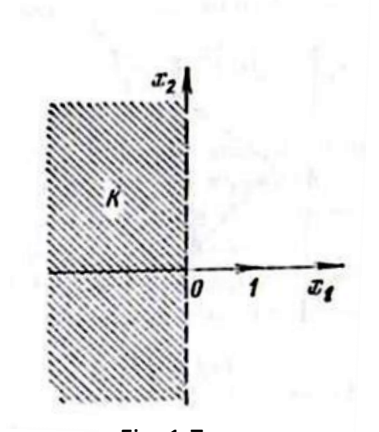


Fig. 1.7.

Définition 1.7. Soit K un cône convexe. L'ensemble $K^* = \{p \mid p \in R^n, \langle p, x \rangle \leq 0 \text{ pour tout } x \in K\}$ s'appelle **cône convexe dual de K**. (voir [7]).

On illustre cette notion par les Fig. 1.8 et 1.9, où les droites en pointillé représentent les générations des cônes K^* duals des cônes K (secteurs ombrés).

En d'autres termes, le cône K^* est formé de vecteurs p , chaque p faisant un angle non aigu avec tout vecteur de K . Aussi les générateurs des K^* des figures sont les demi-droites perpendiculaires aux générateurs de K .

Théorème 1.9. K^* est un cône fermé.

Démonstration.

On suppose que les vecteurs de la suite $p^1, p^2, \dots, p^k, \dots$ appartiennent à K^* et que la suite converge vers $\lim_{k \rightarrow \infty} p^k = p$.

Montrons que $p \in K^*$, i.e. $\langle p, x \rangle \leq 0$ pour tout vecteur $x \in K$.

On a $\langle p^k, x \rangle \leq 0$ parce que $p^k \in K^*$. On passe à la limite pour $k \rightarrow \infty$ et on utilise la propriété de continuité du produit scalaire (de la fonction linéaire dans notre cas), il vient $\langle p, x \rangle \leq 0$. Voir [21].

On peut associer à K^* un cône dual $(K^*)^*$ noté K^{**} . Ci-dessous un résultat géométriquement évident qui est d'une importance capitale.

Théorème 1.10. (Théorie de la dualité pour les cônes convexes).

- 1) $K^{**} \supseteq K$;
- 2) Si K est un cône convexe fermé, alors $K^{**} = K$.

Démonstration.

- 1) De la définition de K^* , il résulte que $\langle x, p \rangle \leq 0$ pour tout $x \in K$ et $p \in K^*$.

Donc, $x \in K^{**}$, i.e. $K \subseteq K^{**}$.

- 2) L'inclusion $K \subseteq K^{**}$ étant démontrée, il reste à vérifier $K \supseteq K^{**}$ pour K fermé.

On admet que $a \notin K$. Si $a \notin K^{**}$, le théorème se trouve démontré.

On applique au cône fermé K et au point $a \notin K$ le théorème de séparation et le résultat 1.8 (tout hyperplan d'appui à un cône passe par l'origine des coordonnées).

Soit p le vecteur de la normale à l'hyperplan qui sépare K et a , auquel cas $\langle x, p \rangle \leq 0$ pour tout point $x \in K$ et $\langle p, a \rangle > 0$.

D'où $p \in K^*$, donc $a \notin K^{**}$ parce que le produit scalaire des vecteurs a et $p \in K^*$ est plus grand que 0.

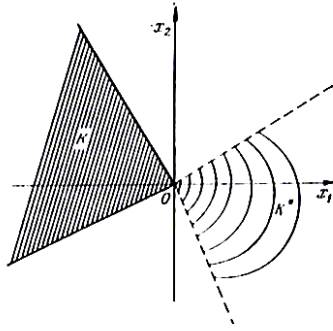


Fig. 1.8

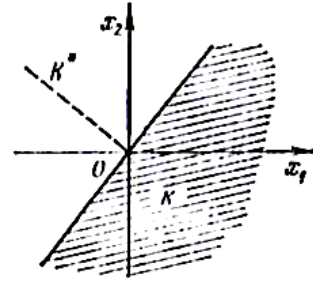


Fig. 1.19

Théorème 1.11. Si un cône convexe fermé K ne contient pas de vecteur non négatif $\neq 0$ (i.e. $K \cap R_+^n = \{0\}$), alors K^* contient un vecteur positif $p > 0$.

Démonstration.

Supposons par l'absurde que le cône K^* ne contient pas de vecteurs positifs. On recourt alors au théorème 1.7 (il existe un vecteur non négatif $p \geq 0$, $p \neq 0$, et $\langle p, x \rangle \leq 0$ pour tout $x \in K^*$).

Cela signifie à la lumière de la définition d'un cône dual que $p \in K^{**}$.

K étant fermé, $K = K^{**}$ par le théorème précédent, d'où $p \in K$, ce qui contredit l'hypothèse du théorème.

On est en mesure de préciser l'énoncé du théorème de séparation 1.7.

Théorème 1.12. Soit K un cône convexe fermé qui rencontre R_+^n le long du seul vecteur nul. Il existe un vecteur positif $p > 0$ tel que $\langle p, x \rangle \leq 0$ pour tout $x \in K$. voir [21].

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 polyèdres.

Définition 1.8. Soient $x^1, x^2, \dots, x^k$ des points quelconques de R^n . On appelle **combinaison linéaire convexe** de ces points la somme $\alpha_1 x^1 + \alpha_2 x^2 + \dots + \alpha_k x^k$, avec $\alpha_i, i = 1, 2, \dots, k$, des nombres non

négatifs arbitraires dont la somme est 1 : $\alpha_i \geq 0, i = 1, 2, \dots, k, \sum_{i=1}^k \alpha_i = 1$.

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

Théorème 1.13. Soient X un ensemble convexe et $x^1, x^2, \dots, x^k$ n'importe quels points de X . L'ensemble X contient toute combinaison linéaire convexe de ces points. Voir [22]

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 démontré que toute combinaison linéaire convexe de $k-1$ points de X appartient à X .

Soient k points $x^1, x^2, \dots, x^{k-1}, x^k$ et la combinaison linéaire convexe $x = \sum_{i=1}^k \alpha_i x^i$. Si $\alpha_k = 1$, alors

$\alpha_i = 0$ pour $1 \leq i \leq k-1$ et $x = x^k \in X$. Admettons que $\alpha_k < 1$, auquel cas $1 - \alpha_k = \sum_{i=1}^{k-1} \alpha_i > 0$. On groupe 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 $\frac{\alpha_i}{1 - \alpha_k}$, $i = 1, 2, \dots, k-1$, 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

$x^1, x^2, \dots, x^{k-1}$ de X . Par hypothèse de récurrence, $x' \in X$, auquel cas le point

$x = (1 - \alpha_k)x' + \alpha_k x^k$ est combinaison linéaire convexe de deux points de X , donc $x \in X$.

Définition 1.9. Soit M un ensemble ponctuel (fini ou infini) de R^n . On appelle **enveloppe convexe** de M , et on note $C(M)$, l'ensemble de toutes les combinaisons linéaires convexes des points de M .

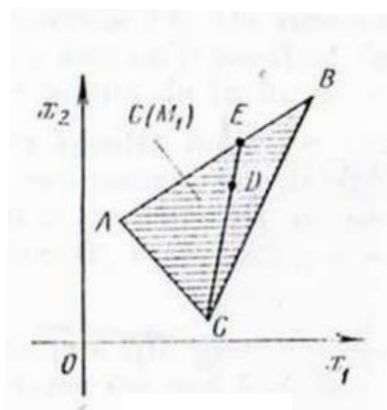


Fig. 1.10.

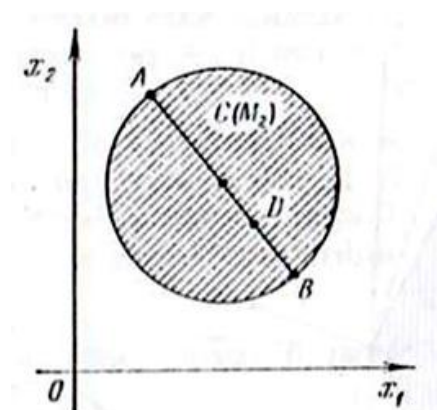


Fig. 1.11.

Considérons à titre d'exemples deux ensembles M_1 et M_2 des Fig. 1.10 et 1.11.

M_1 a pour éléments trois points différents (A, B, C) du plan, et M_2 est une circonférence.

On voit de suite que l'enveloppe convexe de M_1 est le triangle ABC (muni de ses points intérieurs).

En effet, soit D un point quelconque du triangle.

On mène par D et C une droite. Soit E le point où la droite rencontre le segment $[A, B]$.

Le point E est dans ce cas une combinaison linéaire convexe de A et B : $E = \alpha A + (1 - \alpha)B$, avec $0 \leq \alpha \leq 1$. Le point D est sur le segment $[E, C]$, et on écrit donc $D = \beta E + (1 - \beta)C$, $0 \leq \beta \leq 1$.

D'où $D = \beta(\alpha A + (1 - \alpha)B) + (1 - \beta)C = \alpha\beta A + (1 - \alpha)\beta B + (1 - \beta)C$.

Le second membre est combinaison linéaire convexe de A, B , et C parce que $\alpha\beta + (1 - \alpha)\beta + (1 - \beta) = 1$.

Tout point du triangle ABC appartient donc à l'enveloppe convexe de l'ensemble M_1 .

L'enveloppe convexe de l'ensemble M_2 est le disque limité par M_2 .

En effet, tout point D du disque s'écrit sous forme de combinaison linéaire convexe des points A et B de la circonférence qui sont les extrémités du diamètre mené par D .

La propriété des enveloppes convexes M_1 et M_2 de contenir ces points seulement découle du théorème suivant.

Théorème 1.14. L'enveloppe convexe $C(M)$ de l'ensemble M coïncide avec le plus petit ensemble convexe contenant M .

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

Deux exemples ci-dessus montrent que tout point des ensembles $C(M_1)$ et $C(M_2)$ peut être exprimé par les points extrêmes (ce sont (A, B, C) du premier exemple et les points de la circonférence dans le second).

Définition 1.10. Un point x d'un ensemble convexe X est dit extrême s'il n'admet pas la

représentation, $x = \frac{1}{2}x^1 + \frac{1}{2}x^2$ où $x^1 \neq x^2, x^1 \in X, x^2 \in X$.

On note que tout ensemble convexe ne possède pas de points extrêmes.

Ainsi, 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 extrêmes.

On rappelle qu'un ensemble fermé et borné est compact. Les ensembles compacts convexes ont une structure assez simple. (voir [10])

Théorème 1.15. Tout ensemble compact convexe de $\mathbb{R}^n$ est l'enveloppe convexe de ses points extrêmes. [1].

Cela signifie que tout point $x \in X$, X étant un compact convexe, s'écrit sous forme de combinaison convexe d'un nombre fini de points extrêmes de X .

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

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

En effet, soit s un vecteur quelconque dans $\mathbb{R}^n$.

Prenons la droite $l(x) = x + \lambda s, -\infty < \lambda < \infty$, qui passe par x pour (pour $\lambda = 0$). X étant un convexe borné, la droite $l(x)$ coupe sa frontière en deux points x^1, x^2 (peut être confondus). Dans ce cas, le point $x \in X$ appartient au segment $[x^1, x^2]$ et constitue donc une combinaison linéaire convexe des points frontières x^1, x^2 de X .

Le théorème 1.15 serait démontré si l'on montre que tout point frontière x^0 est combinaison linéaire convexe d'un nombre fini de points extrêmes de X .

On rappelle la définition d'une variété linéaire.

On appelle variété linéaire $Q \subseteq \mathbb{R}^n$ un ensemble de la forme $Q = x^0 + L$, avec L un sous-espace.

L'Algèbre linéaire montre que si $Q = x^0 + L$, L étant un sous-espace, alors $L = L$,
i.e. il correspond à la variété linéaire Q un seul sous-espace L dont Q est le translaté.

La dimension d'une variété linéaire $Q = x^0 + L$ est la dimension du sous-espace L .

La variété linéaire Q de la fig. 1.12 est obtenue par translation du sous-espace L : un point quelconque x de Q s'écrit $x = x^0 + l$, où $l \in L$.

Il est clair que le choix de x^0 n'est pas univoque. Dans l'exemple considéré, la dimension de Q est 1.

Comme R^n est une variété linéaire (de dimension n), $x^0 = 0$, $L = R^n$, tout ensemble convexe $X \subseteq R^n$ appartient à une variété linéaire de dimension n .

On appelle dimension d'un ensemble convexe $X \subseteq R^n$ la plus petite dimension des variétés linéaires contenant X .

Ainsi, la dimension du segment $[A, B]$ (fig. 1.10) est 1, celle de l'ensemble convexe réduit à A vaut 0, et le triangle ABC a pour dimension 2.

Si la dimension d'un ensemble convexe X est k , on dit que $X \subseteq R^k$; on prend, à cet effet, une variété linéaire Q , $X \subseteq Q$, k -dimensionnelle et, dans $Q = x^0 + L$, un système de coordonnées $x^0 + l_1, x^0 + l_2, \dots, x^0 + l_k$, avec $\{l_1, l_2, \dots, l_k\}$ une base de L .

La variété Q devient un espace vectoriel de dimension k . Ainsi, on choisit comme origine dans Q de la fig. 1.12 le point x^0 et comme base le vecteur l , auquel cas Q devient R^1 .

Voici un autre fait évident relatif à la notion de dimension des ensembles convexes.

Soit un ensemble convexe X de dimension k et $X \subseteq R^k$.

La dimension de l'intersection de X et de tout hyperplan

$$P = \{x \mid \langle p, x \rangle = d\},$$

où $p \in R^k$, $p \neq 0$, est au plus égale à $k-1$.

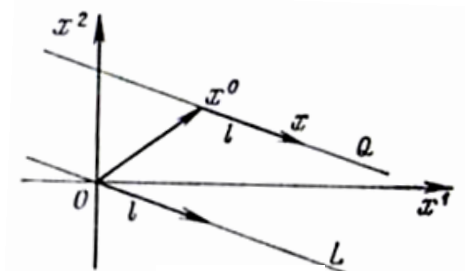


Fig. 1.12.

En effet, soit $\tilde{L} = \{x \mid \langle p, x \rangle = 0\}$, auquel cas la dimension de $\tilde{L}$ est $k-1$.

Comme $P \cap X \subseteq x^0 + \tilde{L}$, l'intersection $P \cap X$ est, par définition, au plus $k-1$ -dimensionnelle ($x^0 \in P \cap X$).

Démonstration du théorème 1.15.

On opère par récurrence sur la dimension k de X .

Si $k = 0$, alors X constitue un point, si bien que l'affirmation du théorème est triviale.

Supposons qu'elle soit vraie pour une dimension inférieure à k , et soit X un ensemble k -dimensionnelle. On note a un point frontière quelconque de X (il existe parce que X est un compact).

Aux termes du théorème 1.6, on trouve un hyperplan $\langle p, x \rangle = \langle p, a \rangle$ tel que $\langle p, x \rangle \leq \langle p, a \rangle$ pour tout $x \in X$.

L'intersection X_a de X et de l'hyperplan est un ensemble convexe compact non vide (l'hyperplan étant convexe et contenant le point $a \in X$).

L'ensemble X_a a une dimension plus petite que k et représente par hypothèse de récurrence l'enveloppe convexe de ses points extrêmes.

Cela étant, tout point extrême c de X_a est un point extrême de X .

On procède par l'absurde et on suppose que $c = \alpha x^1 + (1-\alpha)x^2$, $0 \leq \alpha \leq 1$, $x^1, x^2 \in X$.

On a $0 = \langle p, c - a \rangle = \alpha \langle p, x^1 - a \rangle + (1-\alpha) \langle p, x^2 - a \rangle$, du moment que c est un point de l'hyperplan.

Comme $\langle p, x^1 - a \rangle \leq 0$, $\langle p, x^2 - a \rangle \leq 0$,

on a les égalités $\langle p, x^1 - a \rangle = 0$, $\langle p, x^2 - a \rangle = 0$ i.e., $x^1 \in X_a$, $x^2 \in X_a$.

Or, c est un point extrême de X_a , donc $x^1 = c$, $x^2 = c$.

Ainsi, tout point frontière a de X est combinaison linéaire convexe des points extrêmes de X , c.q.f.d.

Définition 1.11. Un ensemble X s'appelle **polyèdre convexe** s'il est l'enveloppe convexe de points en nombre fini. Voir [21].

Autrement dit, un ensemble X est un polyèdre convexe s'il contient un nombre fini de points

$x^1, x^2, \dots, x^k$ tels que tout élément x de X s'écrit :
$$x = \sum_{i=1}^k \alpha_i x^i, \quad 0 \leq \alpha_i \leq 1, \quad \sum_{i=1}^k \alpha_i = 1$$

et que tous les points de cette forme soient dans X .

Théorème 1.16. Un polyèdre convexe est compact.

Démonstration.

Soit X un polyèdre convexe enveloppe convexe de ses points $x^1, x^2, \dots, x^k$: $X = C(x^1, x^2, \dots, x^k)$.

On démontre la compacité de X sous la condition suffisante qu'on extrait d'une sous suite arbitraire $x(m) \in X$, $m = 1, 2, \dots$, une suite partielle convergente vers un $\bar{x} \in X$.

On a $x(m) \in X$, $m = 1, 2, \dots$ si bien que

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

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

Les suites $\{\alpha_i(m)\}$ étant en nombre fini ($i = 1, 2, \dots, k$), on peut former une suite partielle de numéros $m_1, m_2, \dots$ telle que toutes les suites $\alpha_i(m_j)$, $j = 1, 2, \dots$, convergent.

Soit $\lim_{j \rightarrow \infty} \alpha_i(m_j) = \alpha_i \geq 0$.

Evidemment, $\sum_{i=1}^k \alpha_i = 1$, auquel cas $\sum_{i=1}^k \alpha_i(m_j) x^i$, $j = 1, 2, \dots$, possède une limite et

$$\bar{x} = \lim_{j \rightarrow \infty} \sum_{i=1}^k \alpha_i(m_j) x^i = \sum_{i=1}^k \alpha_i x^i \in X, \quad \text{ce qui démontre le théorème.}$$

Les deux derniers théorèmes entraînent qu'un polyèdre convexe est l'enveloppe convexe de l'ensemble de ses points extrêmes.

Si $M = C(x^1, x^2, \dots, x^k)$ est l'enveloppe convexe de l'ensemble ponctuel $S = \{x^1, x^2, \dots, x^k\}$ on voit clairement que seuls peuvent être des points extrêmes de M les (ou des) points de S .

Les résultats obtenus sont suffisamment intuitifs pour qu'on ait une idée nette de la structure des polyèdres convexes, i.e. des ensembles convexes bornés fermés avec les points extrêmes en nombre fini.

1.2.2. Enveloppes Coniques.

Définition 1.12. Soit $M \subseteq \mathbb{R}^n$ un ensemble quelconque. On appelle **enveloppe conique**, et on note $Co(M)$, l'ensemble de toutes les combinaisons linéaires non négatives $\sum_{i=1}^k \alpha_i x^i$ de points de M .

On montre que $Co(M)$ est le plus petit cône convexe contenant M . Le domaine ombré de la fig. 1.13 est $Co(M)$, avec M formé de trois vecteurs g^1, g^2, g^3 .

L'ensemble M de la fig. 1.14 est une circonférence.

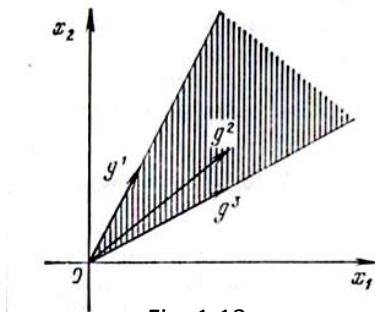


Fig. 1.13.

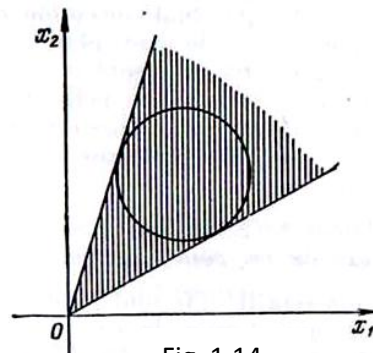


Fig. 1.14.

Définition 1.13 un cône convexe K est dit **polyédrique** s'il est l'enveloppe conique d'un nombre fini de ses points.

On comprend intuitivement qu'un cône polyédrique est fermé, mais la démonstration de cette propriété doit être précisée de celle du lemme suivant.

Lemme 1.1. On suppose qu'un vecteur $x \neq 0$ est combinaison linéaire non négative des vecteurs

$$x^1, x^2, \dots, x^k : x = \sum_{i=1}^k \alpha_i x^i, \alpha_i \geq 0, i = 1, 2, \dots, k.$$

Le vecteur x s'écrit comme combinaison linéaire non négative d'un certain nombre de vecteurs linéairement indépendants de l'ensemble $\{x^1, x^2, \dots, x^k\}$.

Démonstration.

On raisonne par récurrence sur k .

L'affirmation est triviale pour $k = 1$.

Supposons qu'elle soit vraie pour les nombres plus petits que k .

Si au moins l'un des α_i du lemme vaut 0, on applique l'hypothèse de récurrence.

On admet donc que $\alpha_i > 0, i = 1, 2, \dots, k$.

On suppose que $x^1, x^2, \dots, x^k$ sont linéairement indépendants. i.e. $\sum_{i=1}^k \lambda_i x^i = 0$, où λ_i sont non tous nuls.

On admet que certains λ_i sont positifs (dans le cas contraire, on les multiplie tous par -1).

Soit $\theta = \max_{1 \leq i \leq k} (\lambda_i / \alpha_i)$. Il est clair que $\theta > 0$.

Soit $\theta = \lambda_1 / \alpha_1$ pour fixer les idées, auquel cas

$$x = x - \frac{1}{\theta} \sum_{i=1}^k \lambda_i x^i = \sum_{i=1}^k \alpha_i x^i - \frac{1}{\theta} \sum_{i=1}^k \lambda_i x^i = \sum_{i=1}^k \left(\alpha_i - \frac{1}{\theta} \lambda_i \right) x^i.$$

Les coefficients $\alpha_i - \frac{\lambda_i}{\theta}$ sont non négatifs car $\theta \geq \frac{\lambda_i}{\alpha_i}, i = 1, 2, \dots, k$, et $\alpha_1 - \frac{\lambda_1}{\theta} = 0$.

Par conséquent, le vecteur x est combinaison linéaire non négative de $k-1$ vecteurs $x^2, x^3, \dots, x^k$.

On l'exprime, par hypothèse de récurrence, par un sous-système libre de ces vecteurs.

Théorème 1.17. Un cône polyédrique convexe est fermé.

Démonstration.

Soit $K = Co(x^1, x^2, \dots, x^k)$.

Plaçons-nous dans le cas où le système de vecteurs $x^i, i = 1, 2, \dots, k$, est linéairement indépendant.

Soit L un sous-espace vectoriel k -dimensionnelle de R^n engendré par ces vecteurs. Alors $\{x^1, x^2, \dots, x^k\}$ est une base de L .

On munit L d'un produit scalaire de façon que la base donnée soit orthonormée, ce qui permet d'identifier L et R^k , d'une part, K et R_+^k , de l'autre.

Il est évident que R_+^k est un ensemble fermé, ce qui entraîne la propriété de K d'être fermé pour la norme associée au produit scalaire choisi.

On utilise l'équivalence de normes. Soit le cas où l'ensemble de vecteurs $\{x^1, x^2, \dots, x^k\}$ est linéairement indépendant.

On considère un sous-système libre quelconque $(x^{i_1}, x^{i_2}, \dots, x^{i_s})$ de cet ensemble et le cône $K_\sigma = \text{Co}(x^{i_1}, x^{i_2}, \dots, x^{i_s})$, avec $\sigma = \{i_1, i_2, \dots, i_s\}$.

Comme nous l'avons vu, K_σ est fermé. On démontre que $K = \bigcup_{\sigma} K_\sigma$, la réunion étant prise par rapport à tous les jeux σ correspondant aux sous-systèmes libres de vecteurs $x^1, x^2, \dots, x^k$.

En effet, on a évidemment $K \supseteq \bigcup_{\sigma} K_\sigma$ et l'inclusion inverse découle du lemme 1.1.

La propriété, pour la réunion d'ensembles fermés en nombre fini, d'être fermée démontre le théorème.

Conséquences.

Tout cône polyédrique convexe K vérifie l'égalité $K = K^{**}$. [21]

L'affirmation résulte du caractère fermé du cône (on vient de le démontrer) et du théorème de dualité pour les cônes convexes.

Lorsqu'on étudie les cônes convexes, la notion de point extrême si utile pour préciser la structure des polyèdres convexes n'a plus aucune valeur.

On introduit donc un nouvel être mathématique.

Définition 1.14. Un vecteur $x \neq 0$, $x \in K$ est dit **vecteur extrémal** pour un cône convexe K si

$x = \frac{1}{2}x^1 + \frac{1}{2}x^2$ avec $x^1, x^2 \in K$ entraîne $x^1 = \lambda x$, $x^2 = \mu x$ pour certains

$\lambda \geq 0$, $\mu \geq 0$.

Cette définition a une signification claire. Ainsi, g^1, g^3 de la fig. 1.13 sont des exemples de vecteurs extrémaux.

Si K est le demi-plan supérieur fermé de R^2 : $K = \{x \mid x = (x_1, x_2), x_2 \geq 0\}$, il a pour vecteurs extrémaux $(1,0), (-1,0)$.

Si un vecteur $x \in K$ est extrémal, il en est évidemment de même pour tout $\mu x, \mu > 0$.

Nous dirons par convention que le vecteur x appartient à la demi-droite extrémale $\mu x, \mu > 0$.

On voudrait certes obtenir pour les cônes polyédriques un théorème analogue au théorème 1.13 relatif aux ensembles convexes compacts et dire qu'un cône polyédrique est l'enveloppe conique de ses vecteurs extrémaux.

Mais il existe des cônes polyédriques convexes dépourvus de vecteurs extrémaux. Citons, par exemple, le plan R^2 tout entier ou le cône défini par $x_3 \geq 0$ de l'espace R^3 .

Le dernier cône est polyédrique (il coïncide avec l'enveloppe conique des vecteurs $(0,0,1), (1,0,0), (-1,0,0), (0,1,0), (0,-1,0)$) sans posséder aucun vecteur extrémal.

On a néanmoins le résultat évident suivant :

Lemme 1.2. Soient $K = Co(x^1, x^2, \dots, x^k)$ et $\{c_1, c_2, \dots, c_s\}$ l'ensemble de tous ses vecteurs extrémaux appartenant à des demi-droites extrémales distinctes.

Parmi les vecteurs $x^j, j = 1, 2, \dots, k$, il y'en a nécessairement s appartenant à ces demi-droites.

Les exemples cités de cônes sans vecteurs extrémaux expliquent la chose ; ces cônes contiennent un sous-espace. On formalise ce fait à l'aide de la définition suivante.

Définition 1.15. Un cône convexe K est dit **pointé** si les relations d'appartenance $x \in K$ et $-x \in K$ impliquent $x = 0$.

Cette définition correspond exactement à l'exigence que K ne renferme aucun sous-espace vectoriel de R^n . Les deux cônes des fig. 1.13 et fig. 1.14 sont pointés.

Lemme 1.3. Soient des vecteurs $x^1, x^2, \dots, x^k$ des vecteurs non nuls. Si le cône $K = Co(x^1, x^2, \dots, x^k)$ est pointé, les vecteurs $x^1, x^2, \dots, x^k$ sont positivement indépendants,

i.e. les conditions $\sum_{j=1}^k \lambda_j x^j = 0, \lambda_j \geq 0$ entraînent $\lambda_1 = \lambda_2 = \dots = \lambda_k = 0$.

Démonstration.

On raisonne par récurrence sur k . L'affirmation du lemme est évidente pour $k = 1$ parce que x^1 est non nul, par hypothèse.

Supposons que le lemme est démontré pour des nombres plus petits.

Soit $\sum_{j=1}^k \lambda_j x^j = 0, \lambda_j \geq 0$ auquel cas $\lambda_1 x^1 = -\sum_{j=2}^k \lambda_j x^j$. Comme $\sum_{j=2}^k \lambda_j x^j \in K, \lambda_1 x^1 \in K$, on a par

définition d'un cône pointé $\lambda_1 x^1 = 0 = \sum_{j=2}^k \lambda_j x^j$. Alors $\lambda_1 = 0$, et $\lambda_2 = \lambda_3 = \dots = \lambda_k = 0$, par hypothèse de récurrence.

Théorème 1.16. Un cône polyédrique pointé K est l'enveloppe conique de ses vecteurs extrémaux en nombre fini.

Démonstration.

Soit $K = Co(x^1, x^2, \dots, x^k)$. On biffe celui des vecteurs $x^1, x^2, \dots, x^k$ qui s'exprime comme combinaison linéaire non négative des autres. Si la chose est possible, on examine $k-1$ vecteurs restants et on répète l'opération.

On aboutit finalement à un ensemble de vecteurs, disons $x^1, x^2, \dots, x^r$ ($r \leq k$), qui jouit des propriétés suivantes: $K = Co(x^1, x^2, \dots, x^r)$; aucun vecteur x^j , $1 \leq j \leq r$, n'est combinaison linéaire non négative des autres.

On montre que cet ensemble est formé des seuls vecteurs extrémaux. Faisons l'hypothèse de x^1 non

extrémal. Dans ce cas, $x^1 = \sum_{j=1}^r \alpha_j x^j$, où tous les nombres $\alpha_j \geq 0$.

D'où $(1 - \alpha_1)x^1 = \sum_{j=2}^r \alpha_j x^j$.

Si $\alpha_1 > 1$, on a $(1 - \alpha_1)x^1 \in K$ et $(\alpha_1 - 1)x^1 \in K$, i.e. $x^1 = 0$.

Si $\alpha_1 = 1$, on a $\sum_{j=2}^r \alpha_j x^j = 0$, résultat impossible parce que, premièrement, il contredit le lemme 1.3, et, deuxièmement, aucun vecteur x_j , $j = 1, 2, \dots, r$, n'est évidemment nul.

Soit maintenant $\alpha_1 < 1$. On a $x^1 = \frac{1}{1 - \alpha_1} \sum_{j=2}^r \alpha_j x^j$, i.e. x^1 est combinaison linéaire non négative des $x^1, x^2, \dots, x^r$, propriété exclue. Ainsi, le théorème se trouve démontré.

Faisons une remarque sur la différence entre la notion de point extrême d'un ensemble convexe (définition 1.10) et celle du vecteur extrémal d'un cône convexe (définition 1.14).

Ce qui les distingue essentiellement, c'est le fait que le vecteur 0 peut être un point extrême d'un cône sans en être un vecteur extrémal.

D'autre part, le point 0 est évidemment un point extrême d'un cône pointé K .

En effet, soit $0 = \frac{1}{2}x' + \frac{1}{2}x''$, où $x', x'' \in K$.

Il en découle que $x'' = -x'$, ce qui contredit la définition d'un cône pointé.

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.

Un ensemble admissible est décrit par le système d'inéquations linéaires de la forme :

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n \leq b_1, \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n \leq b_2, \\ \cdot \\ a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n \leq b_m \end{cases} \quad (1.3)$$

On n'écrit pas à part les contraintes $x_i \geq 0$ car ces contraintes, si elles existent, sont incluses dans le système (1.3) sous forme $-x_i \leq 0$.

Le symbole a_i désignera dans la suite le vecteur $a_i = (a_{i1}, a_{i2}, \dots, a_{in})$ (la ligne i de la matrice A des coefficients de (1.3)), et a^j désignera le vecteur $a^j = (a_{1j}, a_{2j}, \dots, a_{mj})$ (la colonne j de la matrice A).

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

$$\langle a_1, x \rangle \leq b_1, \quad \langle a_2, x \rangle \leq b_2, \quad \dots, \quad \langle a_m, x \rangle \leq b_m \quad (1.4)$$

x étant comme toujours $x = (x_1, x_2, \dots, x_n)$.

Si l'on fait recours à la notation $A = (a_{ij})$, $i = 1, 2, \dots, m$, $j = 1, 2, \dots, n$.

Notre système se récrit

$$Ax \leq b, \quad (1.5)$$

Où $b = (b_1, b_2, \dots, b_m)$.

La lettre X désignera l'ensemble des vecteurs réalisables,

i.e. l'ensemble des vecteurs x vérifiant (1.3) : $X = \{x \in R^n \mid Ax \leq b\}$.

Il se peut que l'ensemble X soit vide.

Un exemple simple en est l'ensemble R^1 défini par les contraintes $x_1 \leq 1, -x_1 \leq -2$.

On fait à ce propos la restriction suivante : dans ce paragraphe, toutes les affirmations relatives aux propriétés de X se rapportent au cas de X non vide.

Théorème 1.19. Tout ensemble X des vecteurs réalisables est convexe et fermé.

Démonstration.

On montre que le demi-espace X_i défini par $\langle a_i, x \rangle \leq b_i$ est convexe.

En effet, soit $x, x' \in X_i$. on a $\langle a_i, x \rangle \leq b_i, \langle a_i, x' \rangle \leq b_i$.

Soit α un nombre quelconque tel que $0 \leq \alpha \leq 1$, auquel cas

$$\alpha \langle a_i, x \rangle \leq \alpha b_i, \quad (1-\alpha) \langle a_i, x' \rangle \leq (1-\alpha) b_i.$$

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

$$\langle a_i, \alpha x + (1-\alpha)x' \rangle \leq \alpha b_i + (1-\alpha)b_i = b_i, \quad \text{d'où la convexité de } X_i.$$

On constate que l'ensemble X est l'intersection de demi-espaces X_i , i.e. $x \in X$ si et seulement si x est élément de tous les X_i , $i = 1, 2, \dots, m$.

Il ne reste qu'à appliquer le théorème 1.1 (l'intersection d'ensembles convexes est convexe).

On peut utiliser, pour démontrer le théorème 1.19, l'écriture matricielle (1.5) du système d'inéquations donnant X .

Faisons-le, car dans la suite nous aurons souvent affaire aux matrices.

Si $x, x' \in X$, alors $Ax \leq b$, $Ax' \leq b$, auquel cas $\alpha Ax \leq \alpha b$, $(1-\alpha)Ax' \leq (1-\alpha)b$ (on se sert de la non négativité des nombres α et $(1-\alpha)$).

D'où
$$\alpha Ax + (1-\alpha)Ax' \leq b.$$

L'opération de multiplication d'une matrice par un vecteur étant linéaire, on a

$$A(\alpha x + (1-\alpha)x') \leq b, \quad \text{i.e. } \alpha x + (1-\alpha)x' \in X.$$

Démontrons la deuxième propriété voulue (X est fermé).

Soit $x(k), k = 1, 2, \dots, x(k) \in X$, une suite convergente arbitraire, $\lim_{k \rightarrow \infty} x(k) = \bar{x}$.

On passe à la limite dans $Ax(k) \leq b$ quand $k \rightarrow \infty$.

La continuité de la fonction Ax fait que $Ax \leq b$, c.q.f.d.

Les points extrêmes de X sont des objets particulièrement intéressants.

Il se peut que X en soit dépourvu (certains exemples de ces X ont été vus dans le paragraphe précédent).

On établit les conditions qui garantissent l'existence des points extrêmes.

Définitions 1.16.

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

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 ($i = 1, 2, \dots, m$), on trouve facilement un nombre ε tel que si la longueur du vecteur $\Delta = (\Delta_1, \Delta_2, \dots, \Delta_n)$ est plus petite que ε , $\|\Delta\| < \varepsilon$, alors on ait les inégalités $\langle a_i, x_0 + \Delta \rangle \leq b_i, i = 1, 2, \dots, m$, i.e. le point $x_0 + \Delta$ appartient lui aussi à X .

En d'autres termes, toutes les fois que X contient un point intérieur x_0 , il contient un ε -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 : $x_0 + \Delta, x_0 - \Delta \in X$: contradiction avec la définition 1.10 d'un point extrême.

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

Définition 1.17. 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 X et qui contiennent x_0 .

Si l'on reprend la figure suivante, on voit que le point A a pour support deux droites $3x_1 + 5x_2 = 15$, $10x_1 + 7x_2 = 35$, et le point de coordonnées $(2,0)$ la droite $x_2 = 0$.

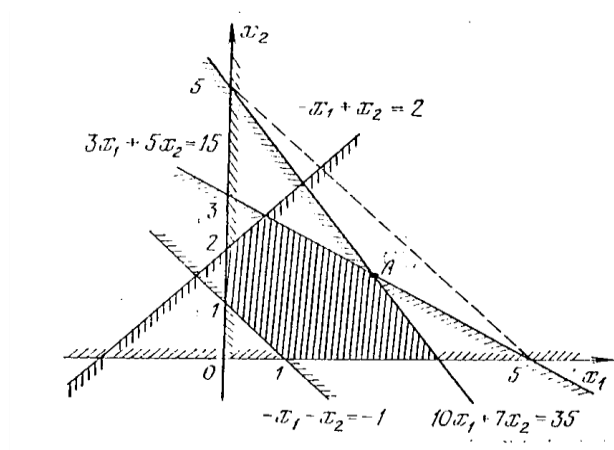


Fig. 1.14.2 Support du point A .

Soient $\sigma = (i_1, i_2, \dots, i_k)$, $1 \leq i_j \leq m$, $j = 1, 2, \dots, k$, les numéros des vecteurs a_i qui définissent le support de x_0 et $\bar{\sigma}$ l'ensemble des numéros $i \in \{1, 2, \dots, m\}$ 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 éléments de σ .

On introduit de même les notations $A_{\bar{\sigma}}$ et $b_{\bar{\sigma}}$.

Si l'on a par exemple

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

(c'est le cas de la fig. précédente), alors $\sigma = (1, 3)$, $\bar{\sigma} = (2, 4, 5, 6)$, $A_\sigma = \begin{pmatrix} 3 & 5 \\ 10 & 7 \end{pmatrix}$, $b_\sigma = \begin{pmatrix} 15 \\ 35 \end{pmatrix}$,

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

Les contraintes (1.3) (ou (1.5)) deviennent pour x_0 : $A_\sigma x_0 = b_\sigma$, $A_{\bar{\sigma}} x_0 < b_{\bar{\sigma}}$.

On rappelle que le rang d'une matrice est par définition le nombre maximum des ses lignes (ou colonnes) linéairement indépendantes.

Théorème 1.20. Un point $x_0 \in X$ est un point extrême si et seulement si le rang de A_σ est n .

Démonstration.

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 extrême, alors le rang de A_σ vaut n .

Supposons que tel n'est pas le cas. Le rang de A est alors inférieur à n , si bien que $A_\sigma x = 0$ a pour solution $\bar{x} \neq 0$.

Comme $A_{\bar{\sigma}} x_0 < b_{\bar{\sigma}}$, on trouve $\alpha \neq 0$ suffisamment petit tel que

$$A_{\bar{\sigma}}(x^0 + \alpha \bar{x}) < b_{\bar{\sigma}}, \quad A_{\bar{\sigma}}(x^0 - \alpha \bar{x}) < b_{\bar{\sigma}},$$

et

$$A_{\sigma}(x^0 + \alpha \bar{x}) = A_{\sigma} x^0 + \alpha A_{\sigma} \bar{x} = A_{\sigma} x^0 = b_{\sigma}.$$

$$\text{De même } A_{\sigma}(x^0 - \alpha \bar{x}) = A_{\sigma} x^0 - A_{\sigma} \alpha \bar{x} = A_{\sigma} x^0 = b_{\sigma}.$$

Il en résulte que $x^0 + \alpha \bar{x}, x^0 - \alpha \bar{x} \in X$, ce qui contredit l'hypothèse « x^0 est un point extrême » (en

$$\text{effet, } x^0 = \frac{1}{2}(x^0 + \alpha \bar{x}) + \frac{1}{2}(x^0 - \alpha \bar{x}).$$

Inversement, soit n le rang de A_σ .

On admet que x^0 n'est pas un point extrême de X .

Il existe deux points $x', x'' \in X$, $x' \neq x''$, tels que $x^0 = \frac{1}{2}x' + \frac{1}{2}x''$, auquel cas

$$b_\sigma = A_\sigma x^0 = \frac{1}{2}A_\sigma x' + \frac{1}{2}A_\sigma x''.$$

Du moment que $A_\sigma x' \leq b_\sigma$, $A_\sigma x'' \leq b_\sigma$, la dernière égalité entraîne $A_\sigma x' = b_\sigma = A_\sigma x''$, d'où $A_\sigma(x' - x'') = 0$, ce qui contredit l'hypothèse faite sur le rang de A_σ (car $x' - x'' \neq 0$).

Le théorème est démontré.

Ce théorème nous donne deux résultats importants.

Théorème 1.21. Les points extrêmes de l'ensemble X , s'il en existe, sont en nombre fini.

Démonstration.

Selon le théorème précédent, il correspond à tout point extrême une matrice carrée non dégénérée A_σ qui le définit complètement : $x^0 = A_\sigma^{-1}b_\sigma$.

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

Théorème 1.22. Si l'ensemble X est borné, c'est un polyèdre convexe.

On le prouve à l'aide de la définition 1.11 d'un polyèdre convexe et du théorème 1.13 (tout ensemble fermé et borné (compact) est l'enveloppe convexe de ses points extrêmes).

1.3.2 Systèmes homogènes.

La démonstration du théorème 1.18 suggère l'idée qu'on ferait bien d'imiter le cas des systèmes d'équations linéaires et d'étudier le système homogène $Ax \leq 0$.

On appellera K l'ensemble des vecteurs x vérifiant $Ax \leq 0$: $K = \{x \mid Ax \leq 0\}$.
L'affirmation suivante est d'une portée capitale.

Théorème 1.23. L'ensemble $K = \{x \mid Ax \leq 0\}$ est un cône polyédrique convexe.

Démonstration.

Soit, dans R^n , le cube $Q = \{x \mid -1 \leq x_i \leq 1, i=1,2,3,\dots,n, x=(x_1, x_2, \dots, x_n)\}$ et l'intersection $S = K \cap Q$.

L'ensemble S est un polyèdre convexe (th. 1.20) parce qu'il est défini par le système d'un nombre fini d'inéquations linéaires: $S = \{x \mid Ax \leq 0, x_j \leq 1, -x_j \leq 1, j=1,2,\dots,n\}$ et est évidemment borné.

Soient $x^1, x^2, \dots, x^k$ les points extrêmes de S , et soit le cône

$$\bar{K} = \left\{x \mid x = \sum_{i=1}^k \alpha_i x^i, \alpha_i \geq 0, i=1,2,\dots,k\right\} \text{ enveloppe conique de l'ensemble } \{x^1, x^2, \dots, x^k\}.$$

Montrons que $K = \bar{K}$.

En effet, $\bar{K} \subseteq K$ du moment que tous les $x^i, i=1,2,\dots,k$, appartiennent au cône K ($S \subseteq K$) et que $\bar{K}$ est le plus petit cône contenant ces points.

D'autre part, si l'on suppose que $x \in K$, il existe un nombre $\lambda > 0$ suffisamment petit tel que $\lambda x \in Q$, i.e. $\lambda x \in K \cap Q = S$.

On représente donc le vecteur λx comme combinaison linéaire convexe de points extrêmes de

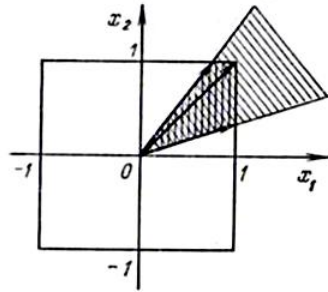
$$S : \lambda x = \sum_{i=1}^k \beta_i x^i, \beta_i \geq 0, i=1,2,\dots,k, \sum_{i=1}^k \beta_i = 1.$$

. Alors

$$x = \sum_{i=1}^k \alpha_i x^i, \quad \text{où } \alpha_i = \frac{\beta_i}{\lambda} \geq 0, i=1,2,\dots,k,$$

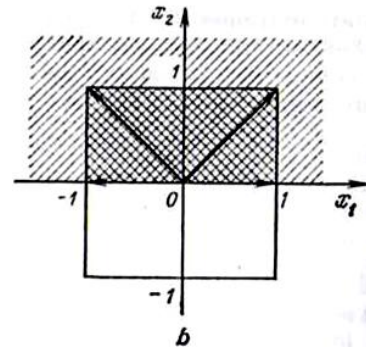
il en découle que $x \in \bar{K}$ i.e. $K \subseteq \bar{K}$.

Conclusion $K = \bar{K}$.



Le théorème est illustré

Les cônes K sont ici hachurés et les ensembles quadrillés.



On a marqué les vecteurs sous-tendant le cône K . Il y en a qui sont superflus, mais l'essentiel est que K est engendré par des vecteurs en nombre fini.

Le théorème 1.18 permet de dire que ce sont les cônes pointés qui possèdent une structure particulièrement simple.

Aussi il est important d'établir une condition nécessaire et suffisante pour que $K = \{x \mid Ax \leq 0\}$ soit pointé.

Théorème 1.24. Le cône $K = \{x \mid Ax \leq 0\}$ est pointé si et seulement si le rang de la matrice A est égal à n .

Démonstration.

On fait l'hypothèse que A est de rang n , K non pointé et $x \neq 0$ tel que $x \in K, -x \in K$.

Dans ce cas, $Ax \leq 0, -Ax \leq 0$, i.e. $Ax = 0$, égalité impossible parce que cette équation homogène possède une solution unique $x = 0$ (le rang de A est égal au nombre de variables).

Si l'on suppose par contre que le rang de A est inférieur à n , alors $Ax = 0$ a une solution non nulle $x \neq 0$ et $A(-x) = 0$ i.e. $x \in K, -x \in K$, si bien que K n'est pas pointé, ce qui démontre le théorème.

Le cas où A est de rang n se présente souvent dans des applications pratiques.

Ainsi, la matrice des contraintes du programme linéaire standard (voir §3) est de rang n .

En effet, si l'on se rappelle la convention faite dans ce chapitre, les contraintes

$x_1 \geq 0, x_2 \geq 0, \dots, x_n \geq 0$ s'incorporent dans la matrice générale des contraintes ; alors la matrice des coefficients du problème standard renferme une sous-matrice $-I$ de rang n .

Définition 1.18. Si le rang de la matrice A est égal à n , l'ensemble $X = \{x \mid Ax \leq b\}$ est dit **régulier**.

Théorème 1.25. L'ensemble X présente des points extrêmes si et seulement si il est régulier.

Démonstration.

Soit X un ensemble non régulier, i.e. le rang de A est plus petit que n .

La matrice A ne renferme donc pas de sous-matrices A_σ de rang n et partant, il n'y a pas de points extrêmes (voir th. 1.20).

La démonstration dans l'autre sens utilise un procédé qui permet de ramener l'étude d'un système quelconque d'inéquations à celle d'un système homogène.

Soit l'espace $R^{n+1} = R^n \times R^1$ de points (x, t) , $x \in R^n$, $t \in R^1$, et soit dans R^{n+1} , le système

d'inéquations linéaires $Ax - bt \leq 0$, $-t \leq 0$. La matrice des contraintes s'écrit $\tilde{A} = \begin{pmatrix} A & -b \\ 0 & -1 \end{pmatrix}$, et son rang est $n+1$ (vu que A est de rang n).

Les théorèmes 1.23 et 1.24 impliquent que l'ensemble $\tilde{K}$ des solutions est un cône polyédrique convexe pointé. Le cône $\tilde{K}$ possède des vecteurs extrémaux (th. 1.18).

On suppose que tous ces vecteurs extrémaux s'écrivent $(x, 0)$.

Comme $\tilde{K}$ coïncide avec l'enveloppe conique de l'ensemble de ses vecteurs extrémaux, tout vecteur de $\tilde{K}$ prend la forme $(x, 0)$. Cela signifie qu'aucun vecteur de $\tilde{K}$ n'est $(x, 1)$, i.e. le système d'inéquations $Ax \leq b$ est incompatible, cas qui ne nous intéresse pas.

On admet donc qu'il existe dans $\tilde{K}$ un vecteur extrémal (x, t) , $t > 0$.

Montrons que le point $\frac{1}{t}x$ est un point extrême de X .

Il est clair que $\frac{1}{t}x \in X$.

Soit $\frac{1}{t}x = \frac{1}{2}x' + \frac{1}{2}x''$, avec $x' \neq x''$, $x' \in X, x'' \in X$. On a $(x, t) = \frac{1}{2}(tx', t) + \frac{1}{2}(tx'', t)$.
De plus, $(tx', t) \in \tilde{K}, (tx'', t) \in \tilde{K}$.

Démontrons-le pour (tx', t) par exemple. Comme $x' \in X$, on a $Ax' \leq b$, auquel cas $A(tx') \leq bt$, c.q.f.d.

L'hypothèse « (x, t) est un vecteur extrémal de $\tilde{K}$ » entraîne
 $(tx', t) = \lambda(x, t)$, $\lambda \geq 0$, $(tx'', t) = \mu(x, t)$, $\mu \geq 0$.

Dans le cas, $\lambda = \mu = 1$, $tx' = x = tx''$, i.e. $x' = x''$, ce qui contredit l'hypothèse que $x' \neq x''$.

On constate une analogie entre la structure de l'ensemble des solutions des systèmes d'équations linéaires et celle de l'ensemble X des programmes réalisables du problème de programmation linéaire (i.e. l'ensemble des solutions du système 2.3).

On éclaircit cette analogie par recours au système $Ax = b$ et au système homogène associé $Ax = 0$.

Soient $\bar{X}$ et L les ensembles des solutions respectifs, i.e. $\bar{X} = \{x \mid Ax = b\}$, $L = \{x \mid Ax = 0\}$.

On note que L est un sous-espace, i.e. L est un cas particulier des cônes.

Il est classique en Algèbre linéaire que la solution générale du premier système s'écrit $x_0 + L$, x_0 étant tout vecteur de $\bar{X}$.

En d'autres termes, $\bar{X} = x_0 + L$.

Voyons ce qu'il en est du système d'inéquations linéaires de la forme (1.3).

On désigne par M l'enveloppe linéaire convexe de l'ensemble des points extrêmes de X .

On vient de démontrer le théorème 1.25 qui entraîne que si X est régulier, alors M n'est pas vide et est un polyèdre convexe.

Théorème 1.26. Si le rang de la matrice A est n (l'ensemble X est régulier), alors $X = M + K$.

Démonstration.

On utilise l'artifice qui a permis de démontrer le théorème 1.25. On a établi notamment que si

(x, t) , $t > 0$, est un vecteur extrémal de $\tilde{K}$, alors $\frac{1}{t}x$ est un point extrême de X .

La réciproque est également vraie : si x un point extrême de X , alors $(x, 1)$ est un vecteur extrémal dans $\tilde{K}$.

En effet, si l'on admet que $(x,1) = \frac{1}{2}(x',t_1) + \frac{1}{2}(x'',t_2)$, où $(x',t_1) \in \tilde{K}$, $(x'',t_2) \in \tilde{K}$, alors

$$x = \frac{1}{2}x' + \frac{1}{2}x'', \quad t_1 + t_2 = 2. \quad \text{Si } t_1, t_2 > 0,$$

on a $\frac{1}{t_1}x' \in X, \frac{1}{t_2}x'' \in X$.

$$x = \frac{t_1}{2} \left(\frac{1}{t_1} x' \right) + \frac{t_2}{2} \left(\frac{1}{t_2} x'' \right),$$

Le point x admet la représentation

$$\frac{1}{t_1}x' = \frac{1}{t_2}x'' = x, \quad \text{d'où } \frac{1}{t_1}x' = \frac{1}{t_2}x'' = x, \quad (x \text{ étant un point extrémal}).$$

i.e. $(x',t_1) = t_1(x,1)$, $(x'',t_2) = t_2(x,1)$, ce qui démontre que $(x,1)$ est un point extrémal de $\tilde{K}$.

On suppose que t_1 ou t_2 est nul, disons $t_2 = 0$.

On a $t_1 = 2, \frac{1}{2}x' \in X, \frac{1}{2}x'' \in X$.

Soit σ le support du point extrême x (voir déf. 1.17), auquel cas $b_\sigma = A_\sigma x = \frac{1}{2}A_\sigma x' + \frac{1}{2}A_\sigma x''$.

Etant donné que $\frac{1}{2}A_\sigma x' \leq b_\sigma, \frac{1}{2}A_\sigma x'' \leq 0$, il en découle en particulier que $A_\sigma x'' = 0$.

Le rang de la matrice A_σ est n . (th. 1.20), si bien que $x'' = 0$.

Par conséquent $x' = 2x$ et $(x',t_1) = 2(x,1)$, $(x'',t_2) = 0 \cdot (x,1)$.

On a donc établi une fois de plus que $(x,1)$ est un vecteur extrémal de $\tilde{K}$.

Soit $(x^1,1), (x^2,1), \dots, (x^k,1), (x^{k+1},0), \dots, (x^n,0)$ les vecteurs extrémaux appartenant aux différentes demi-droites extrémales du cône $\tilde{K}$, les k premiers vecteurs correspondant aux points extrêmes de X .

Il est évident que $x^j \in K$ pour $j = k+1, \dots, n$.

On passe au théorème à démontrer.

On a certes $X \supseteq M + K$.

En effet, si $x^0 \in M$, $z \in K$, alors $Ax^0 \leq b$ et $Az \leq 0$.

Montrons l'inclusion dans le sens opposé.

Soit $x \in X$, auquel cas $(x,1) \in \tilde{K}$ et

$$(x,1) = \sum_{j=1}^k \alpha_j (x^j, 1) + \sum_{j=k+1}^s \beta_j (x^j, 0) \quad \text{avec} \quad \alpha_j \geq 0, \beta_j \geq 0 \quad \text{pour tous les } j.$$

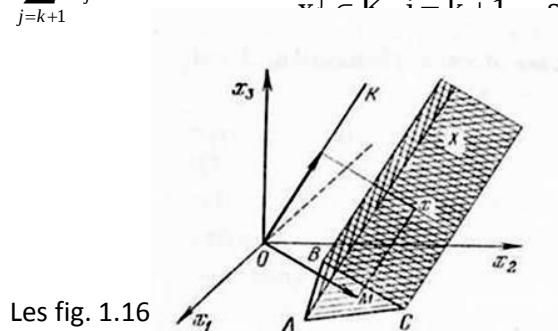
$$x = \sum_{j=1}^k \alpha_j x^j + \sum_{j=k+1}^s \beta_j x^j, \quad \sum_{j=1}^k \alpha_j = 1$$

D'où

Or,

$$\sum_{j=1}^k \alpha_j x^j \in M \quad \text{en tant que combinaison linéaire convexe des points extrêmes } x^j, j = 1, 2, \dots, k,$$

$$\sum_{j=k+1}^s \beta_j x^j \in K$$



Les fig. 1.16

Fig. 1.16.

L'ensemble X de la Fig. 1.16. est un cylindre incliné dans R^3 dont l'axe est parallèle à l'axe x_3 et dont la base est dans le plan $x_1 O x_2$.

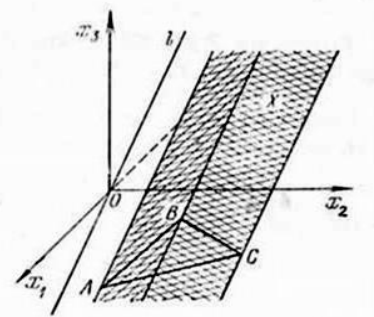


Fig. 1.17.

L'ensemble X de la Fig. 1.17. est un cylindre illimité de deux côtés. Les points extrêmes sont les points A, B, C, et le cône K est la demi-droite issue de O et parallèle à la génératrice du cylindre.

On voit que tout point $x \in X$ peut s'écrire comme somme de deux vecteurs, l'un étant dans K et l'autre dans M.

Quant à X de la fig. 1.17, c'est un cylindre illimité de deux côtés.

Les points extrêmes font défaut, mais on voit assez bien que X est la somme du triangle ABC et du sous-espace L.

Il se trouve que la situation est caractéristique du cas général de X non régulier.

Montrons-le.

Avec la notation a^j pour la colonne j de la matrice A , on voit de suite que le système (1.3) peut s'écrire

$$a^1 x_1 + a^2 x_2 + \dots + a^n x_n \leq b$$

$$(1.6)$$

Si le rang de A est $r < n$, il existe, parmi $a^1, a^2, \dots, a^n$, r colonnes linéairement indépendantes, les autres en étant combinaison linéaire.

On admet que la propriété d'être linéairement indépendantes est spéciale aux r premières colonnes, auquel cas

$$a^s = \sum_{j=1}^r \beta_{js} a^j, \quad s = r+1, \dots, n$$

On reporte dans (1.6) et on réduit les termes semblables, il vient

$$\sum_{j=1}^n a^j x_j = \sum_{j=1}^r a^j x_j + \sum_{s=r+1}^n a^s x_s = \sum_{j=1}^r a^j x_j + \sum_{s=r+1}^n \left(\sum_{j=1}^r \beta_{js} a^j \right) x_s = \sum_{j=1}^r \left(x_j + \sum_{s=r+1}^n \beta_{js} x_s \right) a^j.$$

Ainsi, le système (1.6) est équivalent à

$$\sum_{j=1}^r \left(x_j + \sum_{s=r+1}^n \beta_{js} x_s \right) a^j \leq b \quad (1.7)$$

On introduit la notation

$$l_j(x) = x_j + \sum_{s=r+1}^n \beta_{js} x_s, \quad j = 1, 2, \dots, r,$$

$$l(x) = (l_1(x), l_2(x), \dots, l_r(x), 0, 0, \dots, 0), \quad l(x) \in R^n.$$

On note une propriété évidente de la transformation $l : l(l(x)) \equiv l(x)$.

Soit, dans R^n , deux sous-espaces vectoriels

$$L = \{x \mid l(x) = 0\}, \quad R^r = \{x \mid x_{r+1} = x_{r+2} = \dots = x_n = 0\} \text{ et l'ensemble } Q = L \cap R^r.$$

On établit facilement que $x \in Q$ si et seulement si :

- 1) $l(x) = x$
- 2) $Ax \leq b$.

Les points de l'espace R^r sont de la forme $(x_1, x_2, \dots, x_r, 0, \dots, 0)$, si bien que le système $Ax \leq b$

$$\sum_{j=1}^r a^j x_j \leq b$$

s'écrit pour cet espace

Le rang de A étant r , l'ensemble Q s'écrit par le théorème 1.26 comme somme

$Q = M + K_0$, avec M un polyèdre convexe de R^r et K_0 un cône dans R^r ,
 $K_0 = \{x \mid x \in R^r, Ax \leq 0\}$.

Montrons que $X = Q + L$.

L'inclusion $X \supseteq Q + L$ est évidente.

En effet, si $x = y + z$, où $y \in Q$, $z \in L$, alors $Ax = Ay + Az \leq b$ (parce que $Ay \leq b$, $Az \leq 0$)
i.e. $x \in X$.

Inversement, soit $x \in X$.

On pose $y = x - l(x)$, auquel cas $l(y) = l(x) - l(l(x)) = 0$, i.e. $y \in L$.

Comme $x = l(x) + y$, on a $x \in Q + L$ parce que $l(x) \in Q$.

Ainsi, $X = M + K_0 + L$.

Montrons que $K_0 + L = K = \{x \mid Ax \leq 0\}$.

On a par suite de (1.7) que $Ay = 0$ pour tout $y \in L$.

La définition de K_0 entraîne $Az \leq 0$ quel soit $z \in K_0$.

Donc $K_0 + L \subseteq K$.

Inversement, on suppose que x est dans K et on fait $y = x - l(x)$.

Il est connu que dans ce cas $y \in L$. Cela étant, $l(x) \in R^r$ et $Ax \leq 0$, i.e. $l(x) \in K_0$.

Vu que $x = y + l(x)$, on a $x \in K_0 + L$.

On vient de démontrer le théorème suivant :

Théorème 1.27. L'ensemble X des programmes du problème de programmation linéaire s'écrit
 $X = M + K$, avec M un polyèdre convexe et $K = \{x \mid Ax \leq 0\}$.

1.4. Equivalence de deux définitions d'un ensemble polyédrique convexe.

Selon les résultats obtenus dans le paragraphe précédent, l'ensemble X des solutions du système d'inéquations linéaires $Ax \leq b$ s'écrit $X = M + K$, où M est un polyèdre convexe borné et K un cône polyédrique convexe.

Avec les définitions 1.11 et 1.12, on énonce ce fait comme suit.

L'ensemble X admet la représentation

$$X = \left\{ x \mid x = \sum_{i=1}^k \alpha_i x^i + \sum_{j=1}^s \beta_j y^j \right\} \quad (1.8)$$

où x^i , $i = 1, 2, \dots, k$, et, y^j , $j = 1, 2, \dots, s$ sont des vecteurs de X , $\alpha_i \geq 0$, $\beta_j \geq 0$, $\sum_{i=1}^k \alpha_i = 1$.

On formule naturellement la définition suivante.

Définition 1.19. L'ensemble X de la forme (1.8) s'appelle **ensemble polyédrique convexe**.

Cette définition aidant, les résultats principaux du §3 constituent une démonstration du fait que tout ensemble X défini par un système d'inéquations linéaires est un ensemble polyédrique convexe.

Il s'avère que la réciproque est vraie elle aussi, à savoir que tout ensemble polyédrique convexe s'écrit comme ensemble des solutions d'un système d'inéquations linéaires.

Cela autorise à énoncer la définition suivante.

Définition 1.20. On appelle **ensemble polyédrique** tout ensemble X qui coïncide avec l'ensemble des solutions d'un système d'inéquations linéaires.

Lorsque nous démontrerons l'équivalence des deux dernières définitions, nous terminerons par là même l'exposé de la théorie des ensembles polyédriques convexes.

Aux termes du théorème 1.19, l'ensemble des solutions du système est convexe et fermé.

En premier lieu, on prouve donc ces propriétés pour X de la forme (1.8).

On introduit les notations

$$M = \left\{ x \mid x = \sum_{i=1}^k \alpha_i x^i, \alpha_i \geq 0, i = 1, 2, \dots, k; \sum_{i=1}^k \alpha_i = 1 \right\},$$

$$K = \left\{ y \mid y = \sum_{j=1}^s \beta_j y^j, \beta_j \geq 0, j=1,2,\dots,s \right\},$$

auquel cas $X = M + K$.

Il est connu que M et K sont respectivement un polyèdre convexe et un cône polyédrique convexe, si bien que X est convexe en tant que somme de deux ensembles convexes (th. 1.2).

L'ensemble X est fermé en vertu du lemme suivant.

Lemme 1.4. Soient M un compact et K un fermé. La somme $X = M + K$ est un ensemble fermé.

Démonstration.

Soit $v^1, v^2, \dots, v^m$ une suite convergente de points de X , $\lim_{m \rightarrow \infty} v^m = v$.

Montrons que $v \in X$, ce qui entraîne la propriété demandée de X .

Il existe pour chaque v^m des éléments $x^m \in M$, $y^m \in K$ tel que $v^m = x^m + y^m$, $m=1,2,\dots$

Avec M un compact, on extrait de $\{x^m\}$ (qui peut ne pas converger) une suite partielle convergente.

On n'introduit pas de nouveaux indices (car cela alourdirait l'exposé), et on estime sans restreindre la

généralité que la suite $\{x^m\}$ converge. Soit $\lim_{m \rightarrow \infty} x^m = x$.

On a $x \in M$ du moment que M est fermé.

Etant donné $y^m = v^m - x^m$, on affirme à la lumière de certains résultats connus de la théorie des limites que la suite $\{y^m\}$, $m=1,2,\dots$, possède une limite qu'on note y .

Donc $y = \lim_{m \rightarrow \infty} y^m = \lim_{m \rightarrow \infty} v^m - \lim_{m \rightarrow \infty} x^m = v - x$, $y \in K$,

d'où $v = x + y$, i.e. $v \in X$.

On se propose de former le système d'inéquations linéaires dont X défini par (1.8) est l'ensemble des solutions.

Prenons l'espace R^{n+1} de points $w = (u_1, u_2, \dots, u_n, v)$ ou, plus brièvement, $w = (u, v)$ avec $u = (u_1, u_2, \dots, u_n)$ et v un nombre.

Soit dans R^{n+1} , le système d'inéquations linéaires

$$\langle u, x^1 \rangle - v \leq 0, \langle u, x^2 \rangle - v \leq 0, \dots, \langle u, x^k \rangle - v \leq 0, \quad (1.9)$$

$$\langle u, y^1 \rangle \leq 0, \langle u, y^2 \rangle \leq 0, \dots, \langle u, y^k \rangle \leq 0$$

$x^1, x^2, \dots, x^k$ et $y^1, y^2, \dots, y^k$ deux vecteurs de dimension n intervenant dans l'écriture (1.8) de X .

On désigne par W le sous ensemble de R^{n+1} qui coïncide avec l'ensemble des solutions du système (1.9).

D'après le théorème (1.23), W est un cône polyédrique convexe, si bien qu'on trouve des points $w^1, w^2, \dots, w^r \in W$ pour lesquels

$$W = \left\{ w \mid w = \sum_{i=1}^r \gamma_i w^i, \gamma_i \geq 0, i = 1, 2, \dots, r \right\}. \quad (1.10)$$

Chaque point w^t s'écrit $w^t = (u^t, v^t)$, $t = 1, 2, \dots, r$.

On construit moyennant les nombres v^t et les vecteurs u^t le système d'inéquations dans R^n :

$$\langle u^t, x \rangle \leq v^t, \quad t = 1, 2, \dots, r. \quad (1.11)$$

Soit $\bar{X}$ l'ensemble des solutions du système. On dit que $\bar{X}$ coïncide avec X défini par (1.8)

On commence par montrer que $X \subseteq \bar{X}$.

Soit $x \in X$ un point quelconque.

On a pour certains $\alpha_i \geq 0$, $i = 1, 2, \dots, k$, $\sum \alpha_i = 1$ et $\beta_j \geq 0$, $j = 1, 2, \dots, s$,

$$x = \sum_{i=1}^k \alpha_i x^i + \sum_{j=1}^s \beta_j y^j.$$

On reporte cette expression de x dans (1.11), il vient

$$\langle u^t, x \rangle = \sum_{i=1}^k \alpha_i \langle u^t, x^i \rangle + \sum_{j=1}^s \beta_j \langle u^t, y^j \rangle \leq v^t \sum_{i=1}^k \alpha_i = v^t.$$

On a utilisé le fait que le point $w^t = (u^t, v^t)$ est solution du système (1.9), i.e.

$$\langle u^t, x^i \rangle \leq v^t, \quad i=1,2,\dots,k \quad \text{et} \quad \langle u^t, y^j \rangle \leq 0, \quad j=1,2,\dots,s$$

L'inclusion $X \subseteq \bar{X}$ est donc établie.

Supposons que $\bar{X} \not\subset X$.

Il existe un point $x_0 \in \bar{X}$, $x_0 \notin X$.

On applique à x_0 et à l'ensemble convexe fermé X le théorème de séparation (voir § 1).

Résultat : il existe un vecteur p de dimension n et un nombre c tels que $\langle p, x \rangle \leq c$ pour tous les $x \in X$ et $\langle p, x_0 \rangle > c$.

Montrons que le point $(p, c) \in R^{n+1}$ vérifie les contraintes du système (1.9).

C'est vrai en ce qui concerne les k premières contraintes parce que

$$\langle p, x^i \rangle \leq c, \quad (x^i \in X, \quad i=1,2,\dots,k).$$

Pour voir s'il en est de même des s contraintes restantes, on prend l'ensemble de points

$$x_\beta = x + \beta y^j,$$

avec j fixé, $1 \leq j \leq s$,

β un nombre non négatif quelconque

et $x \in X$ tout point (fixe) de l'ensemble X .

Evidemment, $x_\beta \in X$ pour tout $\beta \geq 0$, auquel cas $\langle x_\beta, p \rangle = \langle x, p \rangle + \beta \langle y_j, p \rangle \leq c$.

Or, la dernière inégalité n'est possible que si $\langle y_j, p \rangle \leq 0$.

Donc $(p, c) \in W$.

Aussi on a pour certains $\gamma_t \geq 0$, $t=1,2,\dots,r$

$$(p, c) = \sum_{t=1}^r \gamma_t w^t = \sum_{t=1}^r \gamma_t (u^t, v^t)$$

$$\text{i.e.} \quad p = \sum_{t=1}^r \gamma_t u^t, \quad c = \sum_{t=1}^r \gamma_t v^t$$

Comme $x_0 \in \bar{X}$, i.e. x_0 satisfait aux conditions (1.11),

$$\langle p, x_0 \rangle = \sum_{t=1}^r \gamma_t \langle u^t, x_0 \rangle \leq \sum_{t=1}^r \gamma_t v^t = c$$

; Ce qui contredit $\langle p, x_0 \rangle > c$.

L'inclusion $\bar{X} \subseteq X$ a lieu et, partant, l'égalité $\bar{X} = X$.

Des développements ci-dessus on en déduit :

Théorème 1.28. Il y a équivalence entre les définitions 1.19 et 1.20 d'un ensemble polyédrique convexe.

Chapitre II

Fonctions Convexes

2.1 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 $t \in [0, 1]$.

$$f : \Omega \rightarrow \mathbb{R}$$

$$x \mapsto f(x)$$

On suppose que f admet des dérivées partielles continues dans Ω .

$$\text{Posons } g(t) = f((1-t)a + tb).$$

$$\text{Alors } g : [0, 1] \rightarrow \mathbb{R}$$

$$t \mapsto g(t) \text{ dérivable dans } [0, 1],$$

et on a $g'(t) = \sum_{i=1}^n (b_i - a_i) \frac{\partial f}{\partial x_i}((1-t)a + tb)$. (c'est le théorème de la dérivation d'une fonction composée).

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

où $b = (b_i)_{i=1,2,\dots,n}$; $a = (a_i)_{i=1,2,\dots,n}$.

Mais d'après le théorème de Lagrange (voir [16]), on a :

$$\begin{aligned} g(1) - g(0) &= (1-0)g'(t_0) \quad \text{où } t_0 \in [0,1] \\ g(1) &= f(b) \quad g(0) = f(a) \quad t_0 = \varepsilon, \text{ on a} \\ f(b) - f(a) &= \langle b - a, \nabla f((1-\varepsilon)a + \varepsilon b) \rangle \\ &= \sum_{i=1}^n (b_i - a_i) \frac{\partial f}{\partial x_i}((1-\varepsilon)a + \varepsilon b) \quad \text{où } \varepsilon \in [0,1] \end{aligned} \quad (1)$$

Soit $f : \Omega \rightarrow \mathbb{R}$, Ω un ouvert de $\mathbb{R}^n$.

$$\varphi : I \subset \mathbb{R} \rightarrow \Omega$$

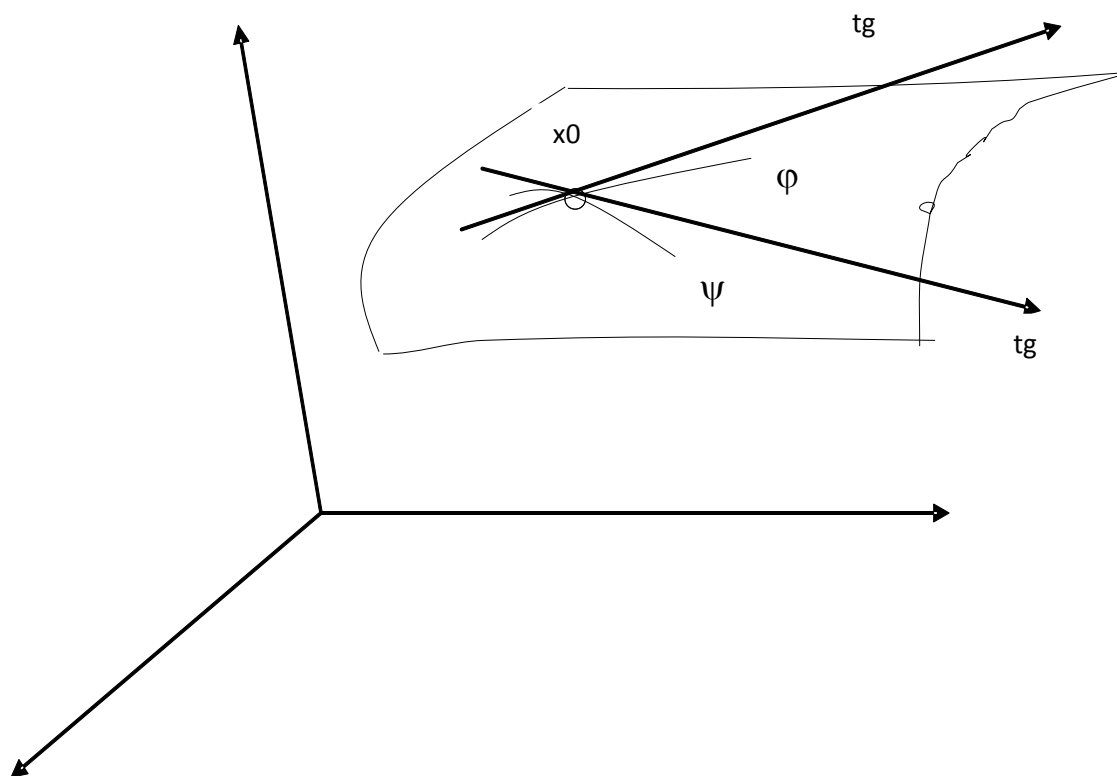
$$t \rightarrow \varphi(t)$$

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'hypersurface $f(x) = b$.

Soit $t_0 \in I$: $\varphi(t_0) = x_0$; $f(x_0) = b$.

Définition 2.1. 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 .



Comme il y'a une infinité de courbes passant par x_0 , on a une infinité de tangentes à cette surface au point x_0 (voir [2])

Définition 2.2.

Un point x_0 de la surface $f(x) = b$ est un **point singulier** si $\frac{\partial f}{\partial x_{i_0}}(x_0) = 0 \quad \forall i = 1, 2, \dots, n$. S'il existe

$i_0 \in \{1, 2, \dots, n\} : \frac{\partial f}{\partial x_{i_0}}(x_0) \neq 0$ on dira que c'est un **point simple** de la surface.

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

$$I \rightarrow \Omega$$

$$t \rightarrow \varphi(t) = (\varphi_1(t), \dots, \varphi_n(t))$$

$$x_1 = \varphi_1(t) ; \quad \varphi_1'(t) = \frac{dx_1}{dt}$$

·
·

$$x_n = \varphi_n(t) ; \quad \varphi_n'(t) = \frac{dx_n}{dt}$$

$(\dot{\varphi}_1(t_0), \dots, \dot{\varphi}_n(t_0))$ est un vecteur tangent à la surface $f(x) = b$ au point x_0 ($\varphi(t_0) = x_0$).

Comme $f(x) = b \Rightarrow f(x) - b = 0 \Rightarrow \frac{\partial f}{\partial x_1} \cdot \frac{\partial x_1}{\partial t} + \frac{\partial f}{\partial x_2} \cdot \frac{\partial x_2}{\partial t} + \dots + \frac{\partial f}{\partial x_n} \cdot \frac{\partial x_n}{\partial t} = 0$.

$N = \frac{\partial f(x_0)}{\partial x_1}, \dots, \frac{\partial f(x_0)}{\partial x_n}$ 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 point du plan tangent et ce vecteur est $x_0 \perp N$ donc

$$\left\langle \left(\frac{\partial f(x_0)}{\partial x_i} \right)_i, x - x_0 \right\rangle = 0 \quad \text{d'où} \quad \langle \nabla f(x_0), x - x_0 \rangle = 0$$

$\Rightarrow$ l'équation du plan tangent à la surface $f(x) = b$ est :

$$\langle \nabla f(x_0), x \rangle = \langle \nabla f(x_0), x_0 \rangle \quad (2).$$

Définition 2.3. Soit Ω un convexe de R^n et soit $f : \Omega \rightarrow R$. 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 2.1. Soit f une fonction de classe C^1 dans un ouvert convexe Ω de R^n .

Si f est convexe dans Ω , alors on a :

$$\langle (x_2 - x_1), \nabla f(x_1) \rangle \leq f(x_2) - f(x_1) \quad \text{pour tout } x_1, x_2 \text{ de } \Omega$$

En effet,

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

$$\text{Pour tout } \lambda \in]0,1], \text{ on a } \frac{f(\lambda x_2 + (1-\lambda)x_1) - f(x_1)}{\lambda} \leq f(x_2) - f(x_1).$$

Mais d'après (1),

$$\begin{aligned} f(\lambda x_2 + (1-\lambda)x_1) - f(x_1) &= \langle \lambda x_2 + (1-\lambda)x_1 - x_1, \nabla f[(1-\varepsilon)x_1 + \varepsilon(\lambda x_2 + (1-\lambda)x_1)] \rangle \\ &= \langle \lambda(x_2 - x_1), \nabla f(x_1 + \varepsilon\lambda(x_2 - x_1)) \rangle \\ &= \lambda \langle (x_2 - x_1), \nabla f(x_1 + \varepsilon\lambda(x_2 - x_1)) \rangle \end{aligned}$$

$$\Rightarrow \langle (x_2 - x_1), \nabla f[x_1 + \varepsilon\lambda(x_2 - x_1)] \rangle \leq f(x_2) - f(x_1) \quad \forall \lambda \in]0,1]$$

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

$$\langle (x_2 - x_1), \nabla f(x_1) \rangle \leq f(x_2) - f(x_1).$$

2.2 L'hyperplan tangent à l'hypersurface.

Théorème 2.2. Soit f une fonction convexe pour tout $x \geq 0$. Alors $\{x: f(x) \leq b, x \geq 0\}$ est convexe.

En effet, posons $\Omega = \{x: f(x) \leq b, x \geq 0\}$.

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 $\lambda \in [0,1]$.

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

$$\Rightarrow \lambda x_1 + (1-\lambda)x_2 \in \Omega.$$

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

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

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

Théorème 2.3. Soit f une fonction convexe pour tout $x \geq 0$ et de classe C^1 . Alors l'hyperplan tangent à l'hypersurface $f(x) = b$ au point $x_0 \geq 0$ est l'hyperplan d'appui où $\Omega = \{x \geq 0: f(x) \leq b\}$ 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 Théorème 1.6).

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'hypersurface $f(x) = b$ au point $x_0 \geq 0$ est $\langle \nabla f(x_0), x \rangle = \langle \nabla f(x_0), x_0 \rangle$ (voir 2).

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

Il reste à montrer qu'il sépare le convexe fermé $\Omega = \{x \geq 0 : f(x) \leq b\}$.

Soit $x_1 \in \Omega; x_1 \geq 0 : f(x_1) \leq b$;

$$\langle \nabla f(x_0), -x_1 + x_0 \rangle = \langle \nabla f(x_0), x_0 \rangle - \langle \nabla f(x_0), x_1 \rangle \leq f(x_0) - f(x_1) = b - f(x_1) \leq 0$$

$$\Rightarrow \langle \nabla f(x_0), x_0 \rangle \leq \langle \nabla f(x_0), x_1 \rangle \text{ . C.q.f.d.}$$

$$\langle \nabla f(x_0), x \rangle \leq \langle \nabla f(x_0), x_1 \rangle \quad \forall x_1 \in \Omega \text{ .}$$

On peut de la même manière démontrer les résultats suivants pour le cas d'une fonction concave.

Théorème 2.4. Soit f une fonction de classe C^1 dans un ouvert convexe $\Omega \subset R^n$, et f concave dans X . Alors $\langle \nabla f(x_1), x_2 - x_1 \rangle \geq f(x_2) - f(x_1)$

Théorème 2.5. Si f est une fonction concave pour tout $x_0 \geq 0$, et soit $\Omega = \{x : f(x) \geq b; x \geq 0\}$. Alors Ω est convexe.

Théorème 2.6. Si f est une fonction concave pour tout $x \geq 0$, et de classe C^1 .

Alors l'hyperplan tangent à l'hypersurface $f(x) = b$ au point $x_0 \geq 0$ est l'hyperplan d'appui à $\Omega = \{x \geq 0 : f(x) \geq b\}$.

2.3 Minimum et maximum d'une fonction convexe.

Théorème 2.7. Soit f une fonction convexe sur un convexe fermé Ω de R^n .

Alors le minimum local de f est un minimum global de f sur Ω .

On raisonne par l'absurde : soit $x_0 \in \Omega$ un minimum local de f .

Supposons alors qu'il existe un point $x^* \in \Omega : f(x^*) < f(x_0)$.

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

Mais

$$\lambda x^* + (1-\lambda)x_0 \in \Omega \text{ et } f(\lambda x^* + (1-\lambda)x_0) < f(x_0), \quad \forall \lambda \in [0,1] \text{ pour } \lambda \text{ assez petit ;}$$

$$\lambda x^* + (1-\lambda)x^* \in v(x_0) \text{ (voisinage de } x_0 \text{),}$$

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.

Remarque 2.1. Ici, on suppose que le minimum local existe. f est localement bornée donc bornée sur le fermé Ω et par suite elle est continue d'après le théorème 2.7.

Théorème 2.8. Soit f une fonction convexe sur un convexe fermé Ω de R^n .

Alors l'ensemble des points minimaux locaux de f est convexe.

En effet, s'il y'a un point minimum local, alors il est global.

Soient x_1, x_2 deux points minimaux locaux

$$x_1 \neq x_2 ; \text{ posons } x = \lambda x_1 + (1-\lambda)x_2, \quad \lambda \in [0,1]$$

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

$$\text{Mais } f(x_1) = f(x_2) = f(x^*)$$

est la valeur minimale de f dans Ω , donc

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

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

Remarque 2.2. $f(x) = \langle a, x \rangle$, f est convexe. Si x_1 est un minimum de f , x_2 un autre alors le segment devient un minimum (Cas du simplexe).

Théorème 2.9. Si f est strictement convexe sur un convexe fermé Ω de R^n .

Alors le minimum local qui est global est unique.

En effet, 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 2.10. Soit f une fonction de classe C^1 sur un convexe Ω de R^n .

Si f est convexe, et si $\nabla f(x_0) = 0$, avec $x_0 \in \dot{\Omega}$ intérieur de Ω .

Alors le minimum global est $x = x_0$.

En effet, on a

$$\langle x - x_0, \nabla f(x_0) \rangle \leq f(x) - f(x_0);$$

$$\nabla f(x_0) = 0 \Rightarrow f(x) - f(x_0) \geq 0$$

$$\Rightarrow f(x) \geq f(x_0) \text{ pour tout } x \in \Omega.$$

Théorème 2.11. Soit f une fonction convexe sur un convexe fermé et borné inférieurement. Si le maximum global de f est fini, alors il est atteint en un ou plusieurs sommets (extrêmes).

i) Supposons que $\max_{x \in \Omega} f(x) = f(x^*)$ et que $x^* \in \dot{\Omega}$. Si f n'est pas constante,
 $\exists x_1 \in \Omega: f(x_1) < f(x^*)$, mais $x^* \in \dot{\Omega} \Rightarrow x^* = \lambda_0 x_2 + (1 - \lambda_0) x_1$ pour $\lambda_0 \in]0, 1[$

$$f(x^*) \leq \lambda_0 f(x_2) + (1 - \lambda_0) f(x_1) < \lambda_0 f(x_2) + (1 - \lambda_0) f(x^*)$$

$$\Rightarrow \lambda_0 [f(x^*) - f(x_2)] < 0$$

$$\Rightarrow f(x^*) < f(x_2) \text{ contradiction.}$$

Donc $f(x) = \text{cst sur } \Omega$

Mais un ensemble Ω de R^n fermé et borné inférieurement possède au moins un point extrême $\bar{x}$ et $f(\bar{x}) = \text{cst}$.

D'où $\max_{x \in \Omega} f(x) = \text{cst} = f(\bar{x})$;

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 hyperplan H_1 , 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 (voir [6]).

Posons $T_1 = \Omega \cap H_1 \subset \Omega$, T_1 est fermé, convexe et borné inférieurement ($\subset \Omega$) donc possède au moins un point extrême ; mais $T_1 \subset H_1$ (hyperplan) 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 d'après i) $\Rightarrow f = cste$

et on aura une contradiction ;

Sinon $x^* \in T_1$ est un point frontière de T_1 d'après ii) posons $T_2 = T_1 \cap H_2, \dots$

on arrive alors à $T_n = T_{n-1} \cap H_n$, et maintenant $\dim T_n = 0 \Rightarrow T_n$ contient un point unique.

Ce point unique est bien évidemment $x^*, \quad T_n = \{x^*\}.$

Un singleton est un point extrême car T_n est fermé et borné inférieurement donc il doit contenir un point extrême. C.q.f.d.

De la même manière, on peut alors démontrer :

Théorème 2.12. Soit f une fonction concave définie sur un convexe fermé de R^n .

Alors le maximum local de f est global.

Théorème 2.13. Soit f une fonction concave définie sur un convexe fermé de R^n .

Alors l'ensemble des maximaux « globaux » est un convexe.

Théorème 2.14. Soit f une fonction concave définie sur un convexe fermé de 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 2.15. Soit f une fonction de classe C^1 dans un ouvert convexe Ω de R^n .

f est concave, et $\nabla f(x_0) = 0$, avec $x_0 \in \dot{\Omega}$ intérieur de Ω .

Alors f atteint son maximum global au point x_0 .

Théorème 2.16. Soit f une fonction de classe C^1 sur un convexe fermé et borné

inférieurement et si f est concave sur Ω , et si le minimum global de

f existe, alors il est atteint en un point extrême du convexe Ω .

Chapitre III

Méthodes de Programmation non Linéaires

3.1. Optimisation d'une fonction convexe ψ :

$$(P) = \begin{cases} x \geq 0; \\ Ax \leq b; \\ \text{Max}_{x \in \Omega} Z, Z = f(x) \end{cases}$$

Considérons le problème canonique **(P)** suivant :

Où

$$f(x) = \sum_{i=1}^n \alpha_i x_i + \sum_i \beta_i x_i^2, \text{ les nombres réels } \alpha_i \text{ sont quelconques, et } \forall i, \beta_i \geq 0.$$

A est une matrice $m \times n$ à coefficients réels, et $b \in R^m$ est un vecteur réel.

Définition 3.1. Une solution réalisable x du problème (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 $\lambda \in]0,1[$, c'est-à-dire c'est un point extrême.

La proposition suivante est bien connue :

Proposition 3.1. 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 $\lambda \in]0,1[$,
 $\{(x = \lambda x_0 + (1-\lambda)x_1) \Rightarrow (x = x_0 = x_1)\}$

La preuve est immédiate :

Pour tout $t \in [0,1]$, posons $g(t) = f(tx_1 + (1-t)x_0)$, où $x_0 = (x_0^1, x_0^2, \dots, x_0^n)$ et $x_1 = (x_1^1, x_1^2, \dots, x_1^n)$ sont deux vecteurs de R^n . Notons que $g(0) = f(x_0)$ et que $g(1) = f(x_1)$.

Pour étudier le signe de $f(x_1) - f(x_0)$, 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:

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

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

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

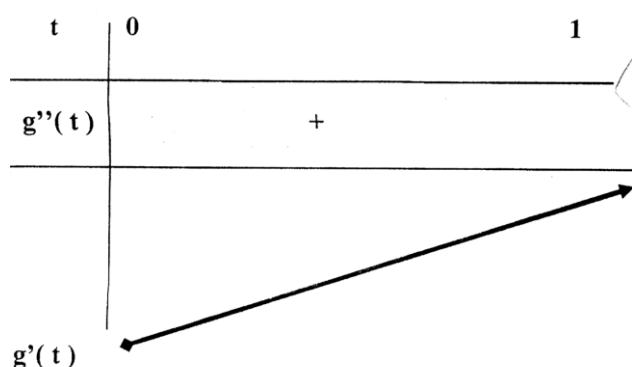
Par suite :

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

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 $\beta_i \geq 0$.



$g'(t)$ est ainsi croissante, linéaire et de coefficient angulaire :

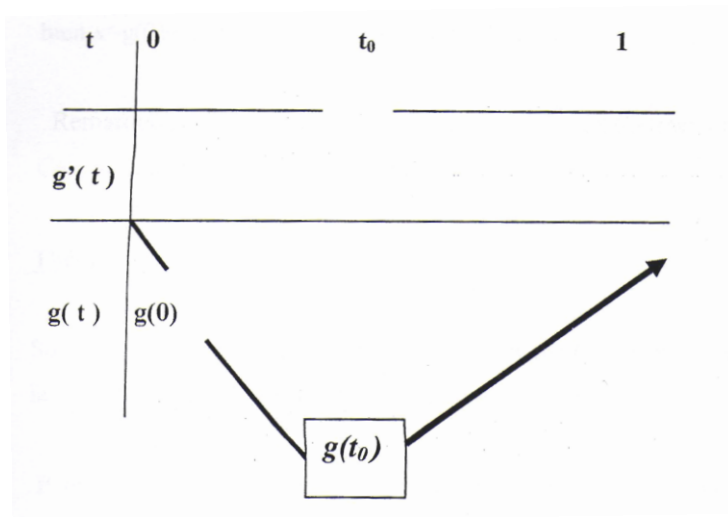
$$g''(t) = 2 \sum_{i=1}^n (x_1^i - x_0^i)^2 \beta_i \geq 0.$$

Notons que si $g''(t) = 0$ alors $g'(t)$ est croissante 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 théorème suivant bien connu, caractérise un sommet :

Théorème 3.1 :

$$x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

Soit une solution réalisable de (P), et $i_1, i_2, \dots, i_r$ les indices i tel que $x_i > 0$.

Alors les deux assertions suivantes sont équivalentes:

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

Démonstration :

1. Supposons le système $\{A^{i_1}, \dots, A^{i_r}\}$ 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 x_0^i et x_1^i sont positifs (voir [10]).

Comme x_0 et 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 $\{A^{i_1}, \dots, A^{i_r}\}$ est libre
alors $x_0^j - x_1^j = 0 \forall j \in \{1, \dots, r\}$.

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 est un sommet et montrons que $\{A^{i_1}, \dots, A^{i_r}\}$ est libre.

Soient comme $\lambda_1, \dots, \lambda_r$ des scalaires telle que $\sum_{j=1}^r \lambda_j A^{i_j} = 0$.

Comme x est un sommet alors $\sum_{j=1}^n x_j A^j = b$ 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 μ on voit que :

$$\sum_{j=1}^r (x_{i_j} + \nu \lambda_j) A^{i_j} = b \quad \text{et} \quad \sum_{j=1}^r (x_{i_j} - \mu \lambda_j) A^{i_j} = b$$

Comme $x_{i_j} > 0$, on peut alors choisir ν et μ de sorte que les vecteurs y et z de composantes respectives non nulles $x_{i_j} + \nu \lambda_j \geq 0$ et $x_{i_j} - \mu \lambda_j \geq 0$ soient solutions réalisables de (P). En effet,

si $\lambda_j > 0$, on peut choisir μ vérifiant $0 < \mu \leq \frac{x_{i_j}}{\lambda_j}$ (on peut prendre $\nu = \mu$).

Si $\lambda_j < 0$ on peut choisir ν positif quelconque et vérifiant : $0 < \nu \leq \frac{x_{i_j}}{\lambda_j}$ (on peut prendre $\mu = \nu$).

On voit donc qu'on peut choisir un nombre $\theta > 0$ assez petit tel que

$$x_{i_j} + \theta \lambda_j \geq 0 \quad \text{et} \quad x_{i_j} - \theta \lambda_j \geq 0$$

Avec ces nouvelles composantes pour les vecteurs y et z , on voit que $x = \frac{1}{2}(y + z)$.

On notera que les composantes y_j et z_j pour $j \notin \{i_1, \dots, i_r\}$ sont nulles.

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 $j \in \{1, 2, \dots, r\}$. Le théorème est démontré.

Définition 3.2.

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 R^m .
2. Les composantes de b suivant B soient toutes non négatives.

Définition 3.3.

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 3.1. Toute solution réalisable de base est un sommet.

Démonstration.

Soit x une solution de base réalisable, et soit $B = \{A^{i_1}, \dots, A^{i_m}\}$ une base réalisable de (P).

Par définition même B est une base de R^m , donc libre.

Du théorème 3.1 résulte que x est un sommet.

Corollaire 3.2. 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 $x = {}^t(x_1, x_2, \dots, x_n)$.

Notons par $I^+(x)$ l'ensemble de tous les indices i tels que $x_i > 0$, $I^+(x) = \{i_1, \dots, i_r\}$.

Alors, d'après le théorème 3.1, $\{A^{i_1}, \dots, A^{i_r}\}$ est une partie libre de R^m .

Si $r = m$, $\{A^{i_1}, \dots, A^{i_r}\}$ constitue donc une base B de R^m et on a bien $x = x_B$.

Si $r < m$, on peut compléter cette famille $\{A^{i_1}, \dots, A^{i_r}\}$ de façon à obtenir une base B de R^m et, là on a bien $x = x_B$.

Théorème 3.2 : Si le standardisé de (P) admet des solutions optimales de la forme:

$$x = {}^t(x_1, x_2, \dots, x_k, 0, 0, \dots, 0).$$

Alors au moins un sommet est solution optimale.

Démonstration : (voir [11])

Si $\{A^1, \dots, A^k\}$ est libre alors x est un sommet, d'après le théorème 3.1.

Sinon soit l le rang de ce système de vecteurs $l < k$.

Sans restreindre de la généralité, on peut supposer que la famille $\{A^1, \dots, A^l\}$ est libre.

Par suite $\{A^1, \dots, A^l, A^{l+1}\}$ est liée. Il existe donc $\lambda_1, \lambda_2, \dots, \lambda_l, \lambda_{l+1}$ non tous nuls tels que :

$$\sum_{i=1}^{l+1} \lambda_i A^i = 0$$

$$\text{et donc } \sum_{i=1}^{l+1} \pm \theta \lambda_i A^i = 0, \text{ pour tout réel } \theta.$$

Posons $\lambda_{l+2} = \dots = \lambda_n = 0$.

On voit alors que $\sum_{i=1}^k (x_i \pm \theta \lambda_i) A^i = b$, puisque $\sum_{i=1}^k x_i A^i = b$.

En particulier on peut choisir, $\theta > 0$ 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 des solutions réalisables,

$$x = \frac{1}{2}(x_\theta + x_{-\theta}).$$

et on a

Mais,

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

et

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

Alors

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

d'où

$$\frac{f(x_\theta) + f(x_{-\theta})}{2} = f(x) + \theta^2 \sum_{i=1}^n \beta_i \lambda_i^2$$

Comme

$$f(x) \geq f(x_\theta) \text{ et } f(x) \geq f(x_{-\theta})$$

on obtient alors :

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

et

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

en simplifiant ces deux inéquations on trouve :

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

et

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

et donc on a simultanément :

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

et

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

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

Par conséquent :

$$\theta^2 \sum_{i=1}^n \beta_i \lambda_i^2 = 0.$$

et par suite

$$\text{Et donc } f(x_\theta) = f(x_{-\theta}) = f(x).$$

Les solutions $x, x_\theta, x_{-\theta}$ sont optimales.

On peut par suite choisir $\theta = \theta_0$ telle que l'un des vecteurs $x_{\theta_0}, x_{-\theta_0}$ a une composante nulle.

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

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

Pour $\lambda_i > 0$ il faut choisir $\theta \leq \frac{x_i}{\lambda_i}$ pour que $x_{-\theta}$ soit solution réalisable, x_θ l'est déjà.

De même, pour $\lambda_i < 0$ il faut choisir $\theta \leq -\frac{x_i}{\lambda_i}$ pour que x_θ soit solution réalisable, $x_{-\theta}$ l'est déjà.

$$\theta = \theta_0 = \min \left\{ \frac{x_i}{\lambda_i}, \lambda_i > 0; -\frac{x_i}{\lambda_i}, \lambda_i < 0 \right\};$$

On voit donc qu'en choisissant

x_θ et $x_{-\theta}$ sont solutions réalisables.

On remarquera que $\theta = \theta_0$ signifie qu'il existe un indice i_0 tel que $\theta_0 = \frac{x_{i_0}}{\lambda_{i_0}}$ (ici $\lambda_{i_0} > 0$)

$$\theta_0 = -\frac{x_{i_0}}{\lambda_{i_0}} \quad (\text{ici } \lambda_{i_0} < 0) \quad .$$

ou bien

Par conséquent une des composantes des solutions réalisables x_θ ou $x_{-\theta}$ s'annule.

Le nombre de composantes non nulles de ce vecteur est égal à $k - 1$.

Si ce vecteur (x_{θ_0} ou $x_{-\theta_0}$) est un sommet le théorème est complètement démontré.

Sinon ,

on réitère le processus : on reprend avec x_{θ_0} (ou $x_{-\theta_0}$) à la place de x .

Le nombre de composantes nulles augmente d'une unité à chaque fois, et on parvient au bout d'un nombre fini d'applications du processus à une solution optimale qui est un sommet.

3.2 Cas où la fonction à optimiser est de la forme $f(x) = \sum_{i=1}^n \alpha_i x_i + \sum_{i,j} \beta_{ij} x_i^2$, $\beta_i \in R$.

$$(\Omega) = \begin{cases} a_{11}x_1 + a_{12}x_2 + \dots + a_{1p}x_p \dots + a_{1n}x_n \leq b_1, \\ a_{n1}x_1 + a_{n2}x_2 + \dots + a_{np}x_p \dots + a_{nn}x_n \leq b_n; \quad x_i \geq 0. \end{cases}$$

Posons

$$A_1 = \begin{pmatrix} a_{11} \\ a_{21} \\ . \\ a_{n1} \end{pmatrix}, \dots, A_n = \begin{pmatrix} a_{1n} \\ a_{2n} \\ . \\ a_{nn} \end{pmatrix} \quad (\Omega) \quad \text{s'écrit :} \quad \sum_{i=1}^n A_i x_i \leq b = \begin{pmatrix} b_1 \\ . \\ b_n \end{pmatrix}$$

$$\sum_{i=1}^p A_i x_i + \sum_{i=p+1}^n A_i x_i \leq b.$$

ou encore

Supposons que $\max_{x \in \Omega} f(x) = f(\bar{x}) \Rightarrow$

$$\bar{x} = \bar{x}_1 + \bar{x}_2 = (\bar{x}_{11}, \bar{x}_{12}, \dots, \bar{x}_{1p}, 0, 0, \dots, 0) + (0, \dots, 0, \bar{x}_{2p+1}, \bar{x}_{2p+2}, \dots, \bar{x}_{2n})$$

On peut toujours supposer que $\bar{x}$ s'écrit sous cette forme.

Ici $\beta_1, \beta_2, \dots, \beta_p$ sont positives ou nuls

et $\beta_{p+1}, \beta_{p+2}, \dots, \beta_n$ sont strictement négatifs. (voir [4])

Théorème 3.3 : Si $\{A_1, A_2, \dots, A_p\}$ est libre, alors $\bar{x}_1$ est un sommet de Ω ;

Preuve :

Supposons que $\{A_1, A_2, \dots, A_p\}$ n'est pas libre.

Pour tout $\theta \in \mathbb{R}, \theta \geq 0$, on a $\pm \theta \sum_{i=1}^p \lambda_i A_i = 0 \Rightarrow \sum_{i=1}^p A_i [\bar{x}_{1i} \pm \theta \lambda_i] + \sum_{i=p+1}^n A_i [\bar{x}_{2i}] \leq b$,

car $\sum_{i=1}^p A_i \bar{x}_{1i} + \sum_{i=p+1}^n A_i \bar{x}_{2i} \leq b$ et $\pm \theta \sum_{i=1}^p \lambda_i A_i = 0$.

Donc $\bar{x}_{1i} > 0$ pour tout $i = 1, 2, \dots, p \Rightarrow \bar{x}_{1i} \pm \theta \lambda_i \geq 0$ pour un θ assez petit,

et par suite $[\bar{x}_{1i} \pm \theta \lambda_i]_{i=1, \dots, p} + (\bar{x}_{2i})_{i=p+1, \dots, n}$ est une solution réalisable.

Mais, $f([\bar{x}_{1i} \pm \theta \lambda_i] + (\bar{x}_{2i})) = \psi(\bar{x}_{1i} \pm \theta \lambda_i) + \varphi(\bar{x}_{2i})$

$$f(\bar{x}_{1i} + \bar{x}_{2i}) = \psi(\bar{x}_{1i}) + \varphi(\bar{x}_{2i}) \geq \psi(\bar{x}_{1i} \pm \theta \lambda_i) + \varphi(\bar{x}_{2i})$$

$$\Rightarrow \psi(\bar{x}_{1i}) \geq \psi(\bar{x}_{1i} \pm \theta \lambda_i) \quad (1)$$

$$(\bar{x}_{1i})_{i=1, 2, \dots, p} = \frac{1}{2} [(\bar{x}_{1i} + \theta \lambda_i) + (\bar{x}_{1i} - \theta \lambda_i)] \quad \text{avec}$$

$$\psi(\bar{x}_{1i}) \geq \psi(\bar{x}_{1i} + \theta \lambda_i)$$

$$\text{et} \quad \psi(\bar{x}_{1i}) \geq \psi(\bar{x}_{1i} - \theta \lambda_i)$$

$$\psi(\bar{x}_{1i} + \theta \lambda_i) = \sum_{i=1}^p \alpha_i [\bar{x}_{1i} + \theta \lambda_i] + \sum_{i=1}^p \beta_i [\bar{x}_{1i} + \theta \lambda_i]^2$$

$$\begin{aligned}
\psi(\bar{x}_{1i} - \theta\lambda_i) &= \sum_{i=1}^p \alpha_i [\bar{x}_{1i} - \theta\lambda_i] + \sum_{i=1}^p \beta_i [\bar{x}_{1i} - \theta\lambda_i]^2 \\
\psi(\bar{x}_{1i} + \theta\lambda_i) &= \sum_{i=1}^p \alpha_i \bar{x}_{1i} + \theta \sum_{i=1}^p \alpha_i \lambda_i + 2\theta \sum_{i=1}^p \beta_i \bar{x}_{1i} \lambda_i + \theta^2 \sum_{i=1}^p \beta_i \lambda_i^2 + \sum_{i=1}^p \beta_i \bar{x}_{1i}^2 \\
&\leq \psi(\bar{x}_{1i}) = \sum_{i=1}^p \alpha_i \bar{x}_{1i} + \sum_{i=1}^p \beta_i \bar{x}_{1i}^2 \\
\Rightarrow \theta \left[\sum_{i=1}^p \alpha_i \lambda_i + 2 \sum_{i=1}^p \beta_i \bar{x}_{1i} \lambda_i + \theta \sum_{i=1}^p \beta_i \lambda_i^2 \right] &\leq 0
\end{aligned}$$

$$\text{de même} \quad -\theta \left[\sum_{i=1}^p \alpha_i \lambda_i + 2 \sum_{i=1}^p \beta_i \bar{x}_{1i} \lambda_i - \theta \sum_{i=1}^p \beta_i \lambda_i^2 \right] \leq 0$$

$$\Rightarrow \theta \left[\sum_{i=1}^p \alpha_i \lambda_i + 2 \sum_{i=1}^p \beta_i \bar{x}_{1i} \lambda_i \right] \leq -\theta^2 \sum_{i=1}^p \beta_i \lambda_i^2 \quad \beta_i \geq 0$$

$$\Rightarrow -\theta \left[\sum_{i=1}^p \alpha_i \lambda_i + 2 \sum_{i=1}^p \beta_i \bar{x}_{1i} \lambda_i \right] \leq -\theta^2 \sum_{i=1}^p \beta_i \lambda_i^2 \quad \beta_i \geq 0$$

$$\Rightarrow \theta \left[\sum_{i=1}^p \alpha_i \lambda_i + 2 \sum_{i=1}^p \beta_i \bar{x}_{1i} \lambda_i \right] \geq \theta^2 \sum_{i=1}^p \beta_i \lambda_i^2 = k$$

$$0 \leq k \leq \theta \left[\sum_{i=1}^p \alpha_i \lambda_i + 2 \sum_{i=1}^p \beta_i \bar{x}_{1i} \lambda_i \right] \leq -k$$

D'où

$$\Rightarrow \theta \left[\sum_{i=1}^p \alpha_i \lambda_i + 2 \sum_{i=1}^p \beta_i \bar{x}_{1i} \lambda_i \right] = 0 \text{ et } k = 0$$

$$\psi(\bar{x}_{1i} + \theta\lambda_i) = \psi(\bar{x}_{1i} - \theta\lambda_i) = \sum_{i=1}^p \alpha_i \bar{x}_{1i} + \sum_{i=1}^p \beta_i \bar{x}_{1i}^2 = \psi(\bar{x}_{1i})$$

D'où

$\Rightarrow (\bar{x}_{1i} + \theta\lambda_i)_{i=1,2,\dots,p} + (\bar{x}_{2i})_{i=p+1,\dots,n}$ et $(\bar{x}_{1i} - \theta\lambda_i)_{i=1,\dots,p} + (\bar{x}_{2i})_{i=p+1,\dots,n}$ sont aussi deux **solutions optimales**.

On peut, par suite, choisir $\theta = \theta_0$ telle que l'un des vecteurs

$(\bar{x}_{1i} + \theta\lambda_i)_{i=1,2,\dots,p}$ ou $(\bar{x}_{1i} - \theta\lambda_i)_{i=1,\dots,p}$ ait une composante nulle, et le nombre de composantes non nulles de l'un de ces deux vecteurs se réduit à $k-1$, s'il est sommet le théorème est démontré.

Sinon on réitère le processus.

Au bout d'un nombre fini d'itérations, on parvient à une solution qui est un sommet et

$(\bar{x}_{1i} - \theta\lambda_i)_{i=1,\dots,p} + (\bar{x}_{2i})_{i=p+1,\dots,n}$ est solution optimale de f . c.q.f.d. (voir [17])

3.3. 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.

Son principe est particulièrement simple dans le cas de problèmes avec contraintes linéaires.

En effet, cet algorithme combine des approximations linéaires de l'objectif pour trouver la **direction de recherche** avec une recherche unidimensionnelle afin de déterminer le **pas** suivant cette direction.

L'idée de la méthode est la suivante. Il s'agit d'un algorithme **d'approximations linéaires successives**. Nous allons l'appliquer au problème suivant dont les contraintes sont purement linéaires :

$$\max f(x) = 5x_1 - x_1^2 + 8x_2 - 2x_2^2$$

$$\text{s.c.q.} \begin{cases} 3x_1 + 2x_2 \leq 6 \\ x_1 \geq 0 \\ x_2 \geq 0. \end{cases}$$

Ce type de problème est souvent rencontré pour la résolution de problèmes 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 d'ordre 1:

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

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

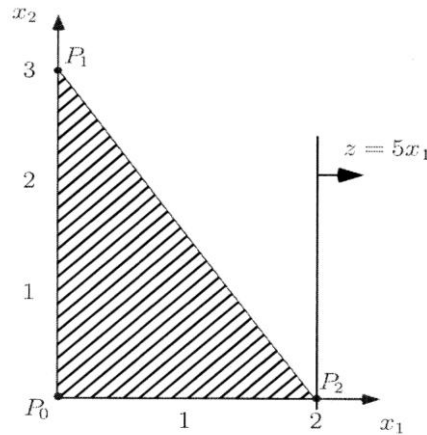
$$\frac{\partial f}{\partial x_1}(0,0) = 5 - 2(0) = 5 \qquad \frac{\partial f}{\partial x_2}(0,0) = 8 - 4(0) = 8$$

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 :

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

par l'algorithme du Simplexe en général. Comme il n'y a que deux variables, ceci est fait graphiquement à la figure 3.1. La **solution optimale** est le point



$$x_{LP}^1 = (0,3)$$

Fig. 3.1.

3. 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_{LP}^1 , cependant en se dirigeant dans la direction de x_{LP}^1 au début du moins, on pourra réduire $f(x)$.

On va déterminer le pas dans la direction : $x^1 = x^0 + \alpha(x_{LP}^1 - x^0)$

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} + \alpha \begin{pmatrix} 0-0 \\ 3-0 \end{pmatrix} = \begin{pmatrix} 0 \\ 3\alpha \end{pmatrix}$$

avec $\alpha \in [0,1]$ qui minimise cette fois la vraie fonction : $h(\alpha) = f(x^0 + \alpha(x_{LP}^1 - x^0))$.

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 : $\alpha = \frac{2}{3}$. D'où $x^1 = (0,0) + \frac{2}{3}[(0,3) - (0,0)] = (0,2)$

Itération 2 :

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

$$\frac{\partial f}{\partial x_1}(0,2) = 5 - 2(0) = 5$$

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

$$2. \text{ Résoudre le problème linéaire : } \max g(x) = 5x_1 + 0x_2 \quad \text{s.c.q.} \begin{cases} 3x_1 + 2x_2 \leq 6 \\ x_1 \geq 0 \\ x_2 \geq 0. \end{cases}$$

Ceci est fait à la figure 3.3. La solution optimale est $x_{LP}^2 = (2,0)$.

$$3. \text{ Recherche unidimensionnelle : } x^2 = x^1 + \alpha(x_{LP}^2 - x^1)$$

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 2 \end{pmatrix} + \alpha \begin{pmatrix} 2-0 \\ 0-2 \end{pmatrix} = \begin{pmatrix} 2\alpha \\ 2-2\alpha \end{pmatrix}$$

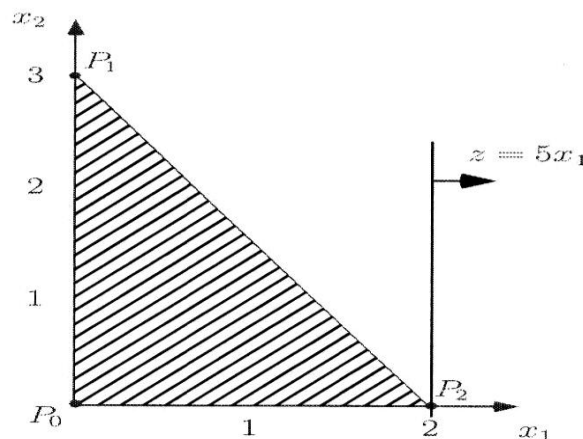
On obtient donc le problème de minimisation unidimensionnelle:

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

$$\text{qui est le maximum pour } \alpha = \frac{5}{12} \text{ d'où : } x^2 = (0,2) + \frac{5}{12}[(2,0) - (0,2)] = \left(\frac{5}{6}, \frac{7}{6}\right)$$

$$x_{LP}^1 = (0,3)$$

Fig. 3.2.



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 $j = 1, 2, \dots, n$, évaluer

$$c_j = \frac{\partial f(x^{k-1})}{\partial x_j}$$

3. Calcul de la direction. Trouver, par application de l'algorithme du Simplexe, x_{LP}^k la solution optimale du problème linéaire suivant :

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

$$s.c.q. \begin{cases} Ax \leq b \\ x \geq 0 \end{cases}$$

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

$$\max f[x^{k-1} + \alpha(x_{LP}^k - x^{k-1})]$$

$$s.c.q. \ 0 \leq \alpha \leq 1$$

et mettre
$$x^k = x^{k-1} + \alpha_k(x_{LP}^k - x^{k-1})$$

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

3.4 Application de la méthode du simplexe pour la résolution de certains problèmes quadratiques.

Le but de cette section est d'étudier le cas où f est quadratique et $g_1, g_2, \dots, g_m$ sont des contraintes 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.

Exemple 3.1. Optimisation de portefeuille en finance. Considérons deux valeurs financières (actions, obligations, ...) 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,
- $\mu_i = E r_i$ l'espérance de r_i ,
- $\sigma_i^2 = Var r_i$ la variance de r_i ,
- $\sigma_{12} = Cov(r_1, r_2)$ la covariance de r_1 et r_2 .

$\mu_1, \mu_2, \sigma_1^2, \sigma_2^2$ et σ_{12} sont supposés connus.

Le but dans cet exemple est de trouver un portefeuille composé de ces 2 titres, tel que son rendement est optimal en un sens que nous préciserons maintenant. 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

- $r_i = x_1 r_{1i} + x_2 r_{2i}$ la rentabilité annuelle $P(x_1, x_2)$,
- $\mu_p = Er_p = x_1 \mu_1 + x_2 \mu_2$ l'espérance de r_p ,
- $\sigma_p^2 = Var r_p = Var(x_1 r_1 + x_2 r_2) = x_1^2 \sigma_1^2 + x_2^2 \sigma_2^2 + 2x_1 x_2 \sigma_{12}$ la variance de r_p ,

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 :

$$\begin{aligned} \min f(x_1, x_2) &= \sigma_p^2(x_1, x_2) = x_1^2 \sigma_1^2 + x_2^2 \sigma_2^2 + 2x_1 x_2 \sigma_{12} \\ \text{s.c.q.} \quad \mu_p(x_1, x_2) &= \mu_1 x_1 + \mu_2 x_2 \geq \mu_0, \quad x_1 + x_2 = 1, \quad x_1, x_2 \geq 0. \end{aligned}$$

Remarque 3.2. On supprime parfois la contrainte $x_1, x_2 \geq 0$ 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 problème quadratique. La fonction f devient alors une forme quadratique.

$$(P_{\min}) \min f(x) = cx + \frac{1}{2} x^t Q x$$

$$\text{s.c.q.} \begin{cases} Ax - b \leq 0 \\ x \in R_+^n \end{cases}$$

$$(P_{\max}) \max f(x) = cx - \frac{1}{2} x^t Q x$$

$$\text{s.c.q.} \begin{cases} Ax - b \leq 0 \\ x \in R_+^n \end{cases}$$

Ici $c = (c_1, c_2, \dots, c_n) \in R^n$, $Q = (q_{ij}; i, j = 1, 2, \dots, n)$ est une matrice symétrique,

$A = (a_{ij}; i = 1, 2, \dots, m; j = 1, 2, \dots, n)$ est une matrice $m \times n$ et $b = (b_1, b_2, \dots, b_m) \in R^m$.

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 3.3 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 définie positive si, pour tout $x = (x_1, x_2, \dots, x_n) \in R^n$,

$${}^t x Q x = \sum_{i=1}^n \sum_{j=1}^n q_{ij} x_i x_j \geq 0.$$

Ceci revient à dire que le déterminant de Q ainsi que tous les sous-déterminants sont non-négatifs.

Exemple 3.2.

Problème non linéaire de la production

$$f(x_1, x_2) = 5x_1 + 5x_2 - x_1^2 - x_2^2 = cx - \frac{1}{2} {}^t x Q x$$

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

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

Proposition 3.3. La fonction $\left\{ \begin{array}{l} f(x) = cx + \frac{1}{2} {}^t x Q x \\ f(x) = cx - \frac{1}{2} {}^t x Q x \end{array} \right\}$ est $\left\{ \begin{array}{l} \text{convexe} \\ \text{concave} \end{array} \right\}$ si et seulement si Q est est symétrique et semi définie positive.

La preuve se fait par un calcul qui vérifie directement les inégalités de la définition de la convexité et de la concavité.

Nous évaluons maintenant le Théorème de Kuhn pour un programme quadratique général.
La fonction de Lagrange est donnée par

$$L(x, \lambda) = f(x) + \lambda g(x) = cx + \frac{1}{2} {}^t x Q x + \lambda(Ax - b)$$

$$L'(x, \lambda) = f(x) - \lambda g(x) = cx - \frac{1}{2} {}^t x Q x - \lambda(Ax - b)$$

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

$$(1) \quad \frac{\partial}{\partial x_j} L(x, \lambda) = c_j + (Qx)_j + (\lambda A)_j \geq 0, \quad j = 1, 2, \dots, n$$

$$\Leftrightarrow {}^t c + Qx + {}^t(\lambda A) \geq 0$$

$$\frac{\partial}{\partial x_j} L'(x, \lambda) = c_j - (Qx)_j - (\lambda A)_j \leq 0, \quad j = 1, 2, \dots, n$$

$$\Leftrightarrow {}^t c - Qx - {}^t(\lambda A) \geq 0$$

$$(2) \quad \left. \begin{array}{l} \frac{\partial}{\partial \lambda_i} L(x, \lambda) \leq 0 \quad i = 1, 2, \dots, m \\ \frac{\partial}{\partial \lambda_i} L'(x, \lambda) \geq 0 \quad i = 1, 2, \dots, m \end{array} \right\} \Leftrightarrow Ax - b \leq 0$$

$$(3) \quad \left\{ \begin{array}{l} \sum_{j=1}^n x_j \frac{\partial}{\partial x_j} L(x, \lambda) = 0 \Leftrightarrow {}^t x \cdot ({}^t c + Qx + {}^t(\lambda A)) = 0 \\ \sum_{j=1}^n x_j \frac{\partial}{\partial x_j} L'(x, \lambda) = 0 \Leftrightarrow {}^t x \cdot ({}^t c - Qx - {}^t(\lambda A)) = 0 \end{array} \right.$$

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

(4)

$$(5) \quad x \in R_+^n, \quad \lambda \in R_+^m$$

(x, λ) est un point selle de L (resp. L') si et seulement si

$$(1) \quad {}^t c + Qx + {}^t(\lambda A) \geq 0 \quad \text{resp.} \quad {}^t c - Qx - {}^t(\lambda A) \leq 0$$

$$(2) \quad Ax - b \leq 0$$

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

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

$$(5) \quad x \in R_+^n, \quad \lambda \in R_+^m$$

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

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

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

$$\begin{aligned}
 (3) \quad & \left. \begin{aligned} {}^t x \cdot u &= 0 \\ \text{resp. } {}^t x \cdot (-u) &= 0 \end{aligned} \right\} \Leftrightarrow {}^t x \cdot u = 0 \\
 (4) \quad & \left. \begin{aligned} \lambda \cdot (-v) &= 0 \Leftrightarrow \lambda \cdot v = 0 \end{aligned} \right\} \Leftrightarrow {}^t x \cdot u + \lambda \cdot v = 0 \\
 (5) \quad & x, u \in R_+^n, \lambda, v \in R_+^m
 \end{aligned}$$

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

$$\begin{aligned}
 & \left| \begin{aligned} Qx + {}^t(\lambda A) - u &= \pm {}^t c \\ Ax &+ v = b \end{aligned} \right| \quad (- \text{ pour } L, + \text{ pour } L') \\
 & x, u \in R_+^n, \lambda, v \in R_+^m \text{ et } {}^t x u + \lambda v = 0 \quad (\text{seul terme non linéaire})
 \end{aligned}$$

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

Maximiser $-My$ (M est une constant $M > 0$)

$$\begin{aligned}
 & \left| \begin{aligned} Qx + {}^t(\lambda A) - u &+ y = \pm {}^t c \\ Ax &+ v = b \end{aligned} \right| \\
 \text{S.c.q.} \quad & y, x, u \in R_+^n, \lambda, v \in R_+^m \text{ et } {}^t x u + \lambda v = 0
 \end{aligned}$$

En appliquant le simplexe pour résoudre ce problème auxiliaire, il faut à chaque étape veiller à ce que les variables entrantes soient telles que ${}^t x u + \lambda v = 0$ reste valable. Ceci modifie légèrement les règles du simplexe.

Exemple 3.3.

Si $x_1 > 0$ est variable de base, il n'est pas possible de faire rentrer u_1 en base, parce que on aurait

$$\begin{aligned}
 {}^t x u &= \sum_{j=1}^n x_j u_j > 0. \\
 \text{alors } x_1 u_1 &> 0 \text{ et donc } \quad \text{Il faut choisir } i' \text{ tel que } x_{i'} = 0 \text{ afin que } x_{i'} u_{i'} = 0 \text{ et } {}^t x u = 0.
 \end{aligned}$$

Exemple 3.4 .

Problème de la production (suite et fin)

$$\begin{aligned}
 \min \quad & f(x_1, x_2) = (5 \ 5) \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} - \frac{1}{2} (x_1 \ x_2) \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \\
 \text{sous} \quad & \begin{pmatrix} 1 & 2 \\ 3 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \leq \begin{pmatrix} 8 \\ 9 \end{pmatrix}, \quad x_1, x_2 \geq 0
 \end{aligned}$$

Le programme auxiliaire se lit :

$$\max -My_1 - My_2$$

$$\begin{aligned} ss.c.q. \quad & \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + {}^t \begin{pmatrix} 1 & 2 \\ 3 & 1 \end{pmatrix} \begin{pmatrix} \lambda_1 \\ \lambda_2 \end{pmatrix} - \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} + \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} 5 \\ 5 \end{pmatrix} \\ & \begin{pmatrix} 1 & 2 \\ 3 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} 8 \\ 9 \end{pmatrix} \\ & y_1, y_2, x_1, x_2, \lambda_1, \lambda_2, u_1, u_2, v_1, v_2 \geq 0 \text{ et } {}^t x u + \lambda v = 0 \end{aligned}$$

Les contraintes s'écrivent

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

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 :

	c		2M	2M	3M	4M	-M	-M	0	0	0	0
c^j	Variable de base		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

Après trois étapes du simplexe on trouve la solution optimale :

$$(x_1, x_2, \lambda_1, \lambda_2) = (2.2, 2.4, 0, 0.2).$$

Exemple 3.4.

$$\min f(x_1, x_2) = -20x_1 + 10x_2 + 3x_1^2 + 2x_2^2$$

$$s.c.q. \quad \begin{cases} 2x_1 - x_2 \leq 6, \\ -x_1 + x_2 \leq 10, \\ 2x_1 + 3x_2 \geq 8, \\ x_1, x_2 \geq 0 \end{cases}$$

En forme matricielle, ce programme s'écrit

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

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

Le programme auxiliaire s'écrit :

$$\text{Maximiser } -My_1 - My_2$$

$$(1) \quad s.c.q. \quad Qx + {}^t(\lambda A) - u + y = -{}^t c$$

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

$$y, x, u \in R_+^3, \lambda, v \in R_+^3 \quad \text{et} \quad {}^t x u + \lambda v = 0$$

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

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

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

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

$$\Leftrightarrow \begin{pmatrix} 2x_1 - x_2 & + v_1 & = 6 \\ -x_1 + x_2 & + v_2 & = 10 \\ 2x_1 + 3x_2 & - v_3 + \tilde{y}_2 & = 8 \end{pmatrix}$$

On en déduit :

$$y_1 = 20 - 6x_1 - 2\lambda_1 + \lambda_2 + 2\lambda_3 + u_1,$$

$$\tilde{y}_2 = 8 - 2x_1 - 3x_2 + v_3.$$

La fonction à maximiser est donnée par

$$-My_1 - M\tilde{y}_2 = -28M + 8Mx_1 + 3Mx_2 + 2M\lambda_1 - M\lambda_2 - 2M\lambda_3 - Mu_1 - Mv_3.$$

La solution de base initiale est donnée par $(y_1, u_2, v_1, v_2, \tilde{y}_2) = (20, 10, 6, 10, 8)$.

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 un point selle de L (notons que f est convexe), puisque $(x_1, x_2) \in S$. Ce point est solution optimale.

Résumons cette section par le théorème suivant :

Théorème 3.4. La solution optimale du programme quadratique étudié dans cette section est donnée par la solution optimale du programme linéaire suivant :

$$\max -My \quad (M > 0 \text{ une constante, } y = {}^t(y_1, \dots, y_n) \in \mathbb{R}_+^n)$$

$$s.c.q. \quad Qx + {}^t(\lambda A) - u + y = {}^t c \quad (-\text{pour } (P_{\min}) + \text{pour } (P_{\max}))$$

$$Ax + y = b, \quad {}^t x u + \lambda v = 0, \quad x, u \in \mathbb{R}_+^n, \lambda, v \in \mathbb{R}_+^m$$

	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					1		20
u_2	0	0	-4	1	-1	3		1						10
v_1	0	2	-1						1					6
v_2	0	-1	1							1				10
$\tilde{y}_2$	-M	2	3								-1		1	8
		-8M	-3M	-2M	M	2M	M	0	0	0	M	0	0	

Tab.I.1.- Les variables de base sont encadrées. Variable entrante x_1 . Variable sortante v_1 . La condition ${}^t x u + \lambda v = 0$ est toujours satisfaite parce que $u_1 = 0$ et donc $x_1 u_1 = 0$.

	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.I.2.- La variable entrante est x_2 . La variable sortante est 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 condition ${}^t x u + \lambda v = 0$ 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.I.3.- La variable entrante est x_2 . La variable sortante est 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 condition ${}^t x u + \lambda v = 0$ 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	1	2
		0	-4M	0	0	0	0	0	M	0	M	0	

Tab.I.4.- La variable entrante est x_2 . La variable sortante est $\tilde{y}_2$. Mais $u_2 = 0$, donc $x_2 u_2 = 0$.

La condition ${}^t x u + \lambda v = 0$ reste satisfaite.

	c_i	x_1	x_2	λ_1	λ_2	λ_3	u_1	u_2	v_1	v_2	v_3	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		-1/32	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.I.5.- Solution optimale du programme auxiliaire avec $y_1 = y_2 = 0$. On a $(x_1, x_2, \lambda_1, \lambda_2, \lambda_3) = (3.25, 0.5, 3.18, 0, 2.93)$.

Chapitre IV

Résolution d'un problème quadratique non convexe et sous contraintes linéaires.

4.1. Introduction.

L'objet de ce chapitre consiste à trouver la solution exacte d'un problème de programmation quadratique avec des contraintes linéaires et une fonction objectif quadratique écrite sous sa forme canonique.

Cette étude décrit une nouvelle méthode basée sur la décomposition de la fonction objectif en une somme de deux fonctions l'une convexe et l'autre concave ; un nouvel ensemble réalisable est construit par une transformation homographique, de telle sorte que la projection du point critique de la fonction objectif sur cet ensemble, donne la solution exacte du problème étudié.

Notons qu'on n'a pas besoin de transformer le problème quadratique en un équivalent linéaire comme dans les méthodes numériques ; la méthode est purement analytique et évite le choix de la solution initiale.

La technique est simple et nous permet de trouver les coefficients de la fonction convexe lors du passage d'un sommet à un sommet voisin.

Le théorème fournis reste valable pour un domaine convexe fermé et borné ; il peut être appliqué à un grand nombre de problèmes d'optimisation.

Les résultats obtenus sont d'une grande importance pour la résolution des cas de programmation séparable.

[5] a traité l'exemple suivant :

$$(P) \begin{cases} \max f(x, y) = 5x + 5y - x^2 - y^2 \\ x + 2y - 8 \leq 0 \\ 3x + y - 9 \leq 0 \\ x, y \geq 0 \end{cases}$$

La méthode utilisée pour résoudre ce problème nécessite au moins dix variables. Dans la méthode présentée dans cette thèse, sans ajouter de variable, la solution exacte est atteinte dans la première tentative.

Cette technique analytique détermine la solution optimale exacte du problème. Il est à mentionner ici que la méthode est générale pour n'importe quel problème d'optimisation quadratique.

Le principe de cette technique est de projeter le point critique sur un nouveau convexe construit à partir de l'ensemble des solutions réalisables.

Un second ingrédient est l'étude de la fonction convexe. Les solutions optimales sont des points extrêmes qui sont toujours des sommets. Donc il est seulement nécessaire de se déplacer d'un sommet à un autre et de choisir le sommet qui produit la solution optimale.

En se déplaçant d'un sommet à l'un de ses voisins, la forme de la fonction convexe change et de nouveaux termes contenant le double produit apparaissent [4]; alors une technique simple à programmer, calcule ces coefficients est présentée.

Enfin un exemple illustratif de la méthode est traité.

4.2. DECOMPOSITION DE LA FONCTION OBJECTIF

Considérons le problème suivant:

$$(P) = \begin{cases} f(x) = \sum_{i=1}^n \alpha_i x_i + \sum_{i,j} \gamma_{ij} x_i x_j \\ \max_{x \in \Omega} f(x) \end{cases} \quad (1)$$

où $\Omega = \{ x \in R^n : x \geq 0 \text{ et } Ax \leq b \}$, A une matrice $m \times n$ et b un vecteur de R_+^m .

Il est facile de voir qu'on peut se restreindre à l'étude du cas suivant $f(x) = \sum_{i=1}^n \alpha_i x_i + \beta_i x_i^2$, où les coefficients α_i et β_i sont des réels quelconques, et la fonction $f(x)$ n'est ni convexe ni concave.

Commençons par voir comment optimiser la fonction $f(x)$.

Pour cela, soit $\varphi(x) = \sum_{i=1}^n \alpha_i x_i + \beta_i x_i^2$, où la somme porte sur tous les indices i pour $\beta_i < 0$; et

$\psi(x) = \sum_{i=1}^n \alpha_i x_i + \beta_i x_i^2$, où la somme porte sur tous les indices i pour $\beta_i \geq 0$.

Après décomposition la fonction $f(x)$ devient comme suit:

$$f(x) = \varphi(x) + \psi(x).$$

Nous allons commencer par étudier la fonction concave $\varphi(x)$.

$$x^* = \left(\frac{-\alpha_i}{2\beta_i} \right)_i$$

Nous donnons ici la solution optimale exacte en projetant le point critique nouveau convexe Ω' . Si le point critique x^* est à l'intérieur du convexe Ω , alors la solution optimale se trouve exactement en ce point (voir la remarque 2.1). Sinon, la projection directe du point x^* sur Ω donne seulement une solution approchée comme le font les autres méthodes (voir Exemple 4.2).

Nous montrons dans le théorème 4.1, que $\text{Max}_{x \in \Omega} \varphi(x) = \varphi(\bar{x})$, où $\bar{x}$ est la projection du point x^* sur le convexe Ω' .

Soit $\Omega \subset R^n \rightarrow \Omega' \subset R^n$ l'application qui transforme le convexe Ω en un convexe Ω' .

$T(x) = \Lambda x$, $\Lambda = (\sqrt{-\beta_1}, \dots, \sqrt{-\beta_n})$ n'est pas nulle, parce que les $\beta_i < 0$ pour tout i . Alors elle est conforme.

4.3. OPTIMISATION DE LA FONCTION CONCAVE φ

Soit Ω un ensemble convexe fermé borné de R^n . Posons :

$$\varphi(x) = \sum_{i=1}^n \alpha_i x_i + \beta_i x_i^2, \quad \alpha_i \in R, \quad \beta_i < 0, \quad \text{pour tout } i.$$

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

$$y = (y_i)_i = (\sqrt{-\beta_i} x_i)_i, \quad \text{Max}_{x \in \Omega} \varphi(x) = \varphi(\bar{x}); \quad \bar{x} = (\bar{x}_i)_i \in \Omega.$$

Théorème 4.1. Il existe un ensemble convexe fermé borné Ω' de R^n , et un vecteur $y_0 = (y_{0i})_i \in \Omega'$, tel que les conditions suivantes soient satisfaites :

1. $\text{Max}_{x \in \Omega} \varphi(x) = \varphi(x^*) - \|y^* - y_0\|^2;$ (Propriété 1)
2. $\|y^* - y_0\| = \inf_{y \in \Omega'} \|y^* - y\|;$ (Propriété 2)
3. $\bar{x}_i = \frac{y_{0i}}{\sqrt{-\beta_i}},$ pour tout $i, i = 1, 2, \dots, n.$ (Propriété 3)

Preuve :

Pour tout $x \in \Omega$, soit

$$\Delta \varphi_i = (\alpha_i x_i^* + \beta_i x_i^{*2}) - (\alpha_i x_i + \beta_i x_i^2). \text{ Alors}$$

$$\Delta \varphi_i = \alpha_i \left(\frac{-\alpha_i}{2\beta_i} \right) + \beta_i \left(\frac{-\alpha_i}{2\beta_i} \right)^2 - \alpha_i x_i - \beta_i x_i^2$$

$$= -\beta_i \left(x_i + \frac{\alpha_i}{2\beta_i} \right)^2$$

$$= -\beta_i (x_i - x_i^*)^2.$$

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

Mais

$$\varphi(x^*) - \varphi(x) = \sum_{i=1}^n -\beta_i (x_i - x_i^*)^2; \text{ Pour tout } x \in \Omega,$$

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

Par conséquent,

peut être écrite sous la forme suivante :

$$\varphi(x^*) - \max_{x \in \Omega} \varphi(x) = \inf_{x \in \Omega} \sum_{i=1}^n \left(\sqrt{-\beta_i} (x_i^* - x_i) \right)^2 = \inf_{y \in \Omega'} \sum_{i=1}^n (y_i^* - y_i)^2.$$

$$\text{Soit } \Omega' = \left\{ y = (y_i)_i \in R^n : y_i = \sqrt{-\beta_i} x_i, x = (x_i)_i \in \Omega \right\},$$

$$\inf_{y \in \Omega'} \sum_{i=1}^n (y_i^* - y_i)^2 = \|y^* - y_0\|^2, y_0 \in \Omega'.$$

parce que

$$\max_{x \in \Omega} \varphi(x) = \varphi(x^*) - \|y^* - y_0\|^2, \quad \text{d'où la propriété (1).}$$

$$\inf_{y \in \Omega'} \|y^* - y\|^2 = \|y^* - y_0\|^2, \text{ alors } \|y^* - y_0\|^2 \leq \|y^* - y\|^2, \text{ pour tout } y \in \Omega'.$$

$$\text{Ceci implique que } \|y^* - y_0\| \leq \|y^* - y\|, \text{ pour tout } y \in \Omega'.$$

$$\|y^* - y_0\| \leq \inf_{y \in \Omega'} \|y^* - y\|.$$

Nous avons par conséquent

Puisque $y \in \Omega'$ alors

$$\|y^* - y_0\| \geq \inf_{y \in \Omega'} \|y^* - y\| ; \text{ donc } \|y^* - y_0\| = \inf_{y \in \Omega'} \|y^* - y\|. \text{ D'où la propriété (2).}$$

Le vecteur y_0 est la projection du vecteur y^* sur le nouveau convexe Ω' .

$$\text{Nous avons } \underset{x \in \Omega}{\text{Max}} \varphi(x) = \varphi(\bar{x}) = \varphi(x^*) - \|y^* - y_0\|^2$$

$$\begin{aligned} &= \varphi(x^*) - \inf_{x \in \Omega} \sum_{i=1}^n (y_i^* - \sqrt{-\beta_i} x_i)^2 \\ &= \varphi(x^*) - \sum_{i=1}^n (y_i^* - \sqrt{-\beta_i} \bar{x}_i)^2 \end{aligned} \text{ . D'où la propriété (3).}$$

Remarque 4.1. Si $x^* \in \Omega$, alors $\inf_{x \in \Omega} \sum_{i=1}^n -\beta_i (x_i^* - x_i)^2 = 0$ et $\underset{x \in \Omega}{\text{Max}} \varphi(x) = \varphi(x^*)$.

Remarque 4.2. La transformation $T: \Omega \subset \mathbb{R}^n \rightarrow \Omega' \subset \mathbb{R}^n$, qui pour tout $x \in \Omega$, associe $T(x) = \Lambda x$, $\Lambda = (\sqrt{-\beta_1}, \dots, \sqrt{-\beta_n})$ a comme matrice Jacobéenne.

$$\begin{bmatrix} \sqrt{-\beta_1} & 0 & \dots & 0 \\ 0 & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots \\ 0 & \dots & 0 & \sqrt{-\beta_n} \end{bmatrix}$$

Son déterminant est $\prod_{i=1}^n \sqrt{-\beta_i} \neq 0$. Donc elle est conforme.

Remarque 4.3.

Le convexe Ω est borné. On peut se restreindre au cas $\alpha_i \geq 0$ pour tout i .

En effet, supposons que nous avons $\alpha_i x_i + \beta_i x_i^2$ avec $\alpha_i < 0$.

Soit $\text{Max}(x_i) = \delta_i^*$, $y_i = \delta_i^* - x_i \geq 0$.

Par conséquent

$$\begin{aligned} \alpha_i x_i + \beta_i x_i^2 &= \alpha_i x_i + \beta_i x_i^2 - \alpha_i \delta_i^* + \alpha_i \delta_i^* \\ &= \beta_i y_i^2 + (-\alpha_i - 2\beta_i \delta_i^*) y_i + \alpha_i \delta_i^* + \beta_i \delta_i^{*2} \\ &= \beta_i y_i^2 + \alpha_i' y_i + K_i, \text{ où } \alpha_i' = -\alpha_i - 2\delta_i^* \beta_i \geq 0 \text{ et } K_i = \alpha_i \delta_i^* + \beta_i \delta_i^{*2}. \end{aligned}$$

Par conséquent, il suffit de remplacer : x_i par $\delta_i^* - y_i$,

et $\alpha_i x_i + \beta_i x_i^2$ par $\beta_i y_i^2 + \alpha_i' y_i + K_i$.

Exemple 4.1.

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

Dans ce cas $\beta_1 = \beta_2 = -1$, donc $\Omega = \Omega'$.

$$\frac{\partial f}{\partial x_1}(x_1, x_2) = 5 - 2x_1 = 0 \Rightarrow x_1^* = 5/2$$

$$\frac{\partial f}{\partial x_2}(x_1, x_2) = 5 - 2x_2 = 0 \Rightarrow x_2^* = 5/2$$

Le point critique $x^* = \left(\frac{5}{2}, \frac{5}{2}\right) \notin \Omega$, parce que $3\frac{5}{2} + \frac{5}{2} - 9 = 1 > 0$,

$$x_0 = P_{\Omega}(x^*) = x^* - \frac{\langle x^*, a \rangle - b}{\|a\|^2} a \quad \text{où } a = (3, 1), \quad b = 9.$$

$$\langle x^*, a \rangle = \frac{5}{2} * 3 + \frac{5}{2} * 1 = 10, \quad \|a\|^2 = 10.$$

$$x_0 = \left(\frac{5}{2}, \frac{5}{2}\right) - \frac{1}{10}(3, 1) = (2.2, 2.4)$$

$$\max_{(x_1, x_2) \in \Omega} \varphi(x_1, x_2) = \varphi(\bar{x}) = 12.40$$

Ici nous avons

$$\Omega = \{(x_1, x_2) \in R^2 / 3x_1 + x_2 - 9 \leq 0, x_1 + 2x_2 - 8 \leq 0, x_1 \geq 0, x_2 \geq 0\}.$$

Exemple 4.2.

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

Dans ce cas $\beta_1 = -1$ et $\beta_2 = -2$,

donc $\Omega \neq \Omega'$.

$$\frac{\partial f}{\partial x_1}(x_1, x_2) = 5 - 2x_1 = 0 \Rightarrow x_1^* = 5/2$$

$$\frac{\partial f}{\partial x_2}(x_1, x_2) = 8 - 4x_2 = 0 \Rightarrow x_2^* = 2$$

Le point critique $x^* = \left(\frac{5}{2}, 2\right) \notin \Omega$, parce que $3\frac{5}{2} + 2.2 > 6$,

$$x_0 = P_{\Omega'}(x^*) = x^* - \frac{\langle x^*, a \rangle - b}{\|a\|^2} a \quad \text{où } a = (3, 2), \quad b = 6.$$

$$\langle x^*, a \rangle = \frac{5}{2} * 3 + 2 * 2 = \frac{23}{2}, \quad \|a\|^2 = 13.$$

$$x_0 = \left(\frac{5}{2}, 2\right) - \frac{11}{2 * 13} (3, 2) = \left(\frac{5}{2} - \frac{33}{26}, 2 - \frac{11}{13}\right) = \left(\frac{16}{13}, \frac{15}{13}\right).$$

$$\Omega' = \left\{ (y_1, y_2) \in \mathbb{R}^2 / 3y_1 + \frac{2}{\sqrt{2}} y_2 \leq 6, y_1 \geq 0, y_2 \geq 0 \right\}.$$

De plus $y^* = \left(\frac{5}{2}, 2\sqrt{2}\right).$

La projection de ce point sur Ω' donne le point suivant y_0 :

$$y_0 = \left(1, \frac{3}{\sqrt{2}}\right) \Rightarrow \bar{x} = \left(1, \frac{3}{2}\right) \quad \text{et} \quad \underset{x \in \Omega}{\text{Max}} \varphi(x) = \varphi\left(1, \frac{3}{2}\right) = 11,500.$$

Notons que la projection du point x^* sur Ω donne le point $\left(\frac{16}{13}, \frac{15}{13}\right)$, et $\varphi\left(\frac{16}{13}, \frac{15}{13}\right) = 11,207$.

Evidemment $\underset{x \in \Omega}{\text{Max}} \varphi(x) \leq \varphi(x^*) = \varphi\left(\frac{5}{2}, 2\right) = 14,250$.

Le point $x_0 = \left(\frac{16}{13}, \frac{15}{13}\right)$ de l'ensemble convexe Ω est le plus proche point du point critique x^* , mais il n'est pas la solution optimale de φ .

$$\begin{cases} \max f(x_1, x_2, x_3) = 2x_1 + 3x_2 + 3x_3 + x_1^2 + 2x_2^2 - x_3^2 \\ x_1 + 3x_2 \leq 18 \\ x_1 + x_2 + x_3 \leq 8 \\ 2x_1 + x_2 + 2x_3 \leq 14 \\ x_1, x_2, x_3 \geq 0 \end{cases}$$

Exemple 4.3.

Nous avons $\varphi(x_3) = 3x_3 - x_3^2$,

le point critique est $x^* = \left(\frac{3}{2}\right)$ et $\max_{x_3 \in \Omega} \varphi(x_3) = \varphi\left(\frac{3}{2}\right) = \frac{9}{4}$.

D'autre part $\psi(x_1, x_2) = 2x_1 + 3x_2 + x_1^2 + 2x_2^2$.

Ψ est une fonction convexe, et sa solution optimale est un point extrême (sommet de Ω).

$$\max_{(x_1, x_2) \in \Omega} \psi(x_1, x_2) = \psi(0, 6) = 90 \quad [4].$$

Mais
$$\max_{x \in \Omega} f(x) \leq \max_{x \in \Omega} \psi(x) + \max_{x \in \Omega} \varphi(x).$$

Par conséquent $\max_{x \in \Omega} f(x) \leq 92.25$.

Aussi
$$\left(0, 6, \frac{3}{2}\right) \in \Omega, \quad f\left(0, 6, \frac{3}{2}\right) \leq \max_{x \in \Omega} f(x).$$

alors

$$\max_{x \in \Omega} f(x) = f\left(0, 6, \frac{3}{2}\right) = 92.25.$$

Remarque 4.4.

Si nous posons
$$\psi(x) = \sum_{\beta_i > 0} \alpha_i x_i + \beta_i x_i^2$$

alors
$$\psi(x^*) - \psi(x) = \sum_{i=1}^n -\beta_i (x_i^* - x_i)^2$$
 pour tout $x \in \Omega$; et

$$\psi(x^*) - \max_{x \in \Omega} \psi(x) = -\max_{x \in \Omega} \sum_{i=1}^n \left(\sqrt{\beta_i} x_i^* - \sqrt{\beta_i} x_i \right)^2.$$

Donc
$$\max_{x \in \Omega} \psi(x) = \psi(x^*) + \max_{x \in \Omega} \sum_{i=1}^n \left(\sqrt{\beta_i} x_i^* - \sqrt{\beta_i} x_i \right)^2.$$

Soit
$$y^* = (y_i^*)_i = (\alpha_i / 2\sqrt{\beta_i})_i \quad \text{et} \quad y = (y_i)_i = (\sqrt{\beta_i} x_i)_i.$$

Notons que y_0 est le point le plus éloigné du convexe

$$\Omega' = \left\{ y \in R^n : y_i = \sqrt{\beta_i} x_i, \quad x = (x_i)_i \in \Omega \right\}.$$

4.4. LA NOUVELLE FORME DE LA FONCTION ψ

Calculons les nouveaux coefficients α'_i, β'_i de la fonction ψ , lors du passage d'un sommet à un sommet voisin.

Soit i_0 l'indice du vecteur sortant, et j_0 celui du vecteur entrant.

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

De la colonne du pivot, nous avons

$$\theta_0 = \frac{b_{i_0}}{a_{i_0 j_0}} \quad x_{j_0} = \theta_0 - \sum_{j \neq j_0} \frac{1}{a_{i_0 j_0}} a_{i_0 j} x_j.$$

Posons $\frac{b_{i_0}}{a_{i_0 j_0}}$, par conséquent

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

Ici

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

Par conséquent

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

La valeur de la fonction devient donc

Nous remplaçons $\alpha_{j_0} x_{j_0} + \beta_{j_0} x_{j_0}^2$ par sa valeur:

$$\psi(x) = \Delta_0 + \sum_{j \neq j_0} \left(\alpha_j - \left(\frac{\alpha_{j_0} \theta_0}{b_{i_0}} + \frac{2\beta_{j_0} \theta_0^2}{b_{i_0}} \right) a_{i_0 j} \right) x_j + \sum_{j \neq j_0} \beta_j x_j^2 + \beta_{j_0} \left(\frac{\theta_0}{b_{i_0}} \right)^2 \left(\sum_{j \neq j_0} a_{i_0 j} x_j \right)^2,$$

$$\left(\sum_{j \neq j_0} a_{i_0 j} x_j \right)^2 = \sum_{k \neq j_0} (a_{i_0 k})^2 x_k^2 + 2 \sum_{\substack{j \neq j_0 \\ k \neq j \\ k \neq j_0}} a_{i_0 j} a_{i_0 k} x_j x_k$$

Où $\Delta_0 = \alpha_{j_0} x_{j_0} + \beta_{j_0} x_{j_0}^2$; mais

$$\psi(x) = \Delta_0 + \sum_{j \neq j_0} \left(\alpha_j - \left(\frac{\alpha_{j_0} \theta_0}{b_{i_0}} + \frac{2\beta_{j_0} \theta_0^2}{b_{i_0}} \right) a_{i_0 j} \right) x_j + \sum_{j \neq j_0} \left(\beta_j + \beta_{j_0} \left(\frac{\theta_0}{b_{i_0}} \right)^2 (a_{i_0 j})^2 \right) x_j^2$$

donc

$$+ 2\beta_{j_0} \left(\frac{\theta_0}{b_{i_0}} \right)^2 \sum_{\substack{j \neq j_0 \\ k \neq j \\ k \neq j_0}} \sum a_{i_0 j} a_{i_0 k} x_j x_k$$

Nous obtenons la forme des coefficients pour la première phase

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

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

$$\gamma_{ji} = \beta_{j_0} \left(\frac{\theta_0}{b_{i_0}} \right)^2 a_{i_0 j} a_{i_0 i} \text{ pour } j \neq i, j \neq j_0, i \neq j_0 \text{ et } \gamma_{j_0 i} = \gamma_{i j_0} = \gamma_{ii} = 0 \quad \forall i = 1, 2, \dots$$

La forme de la fonction devient alors

$$\psi(x) = \Delta_0 + \sum_{j \neq j_0} \alpha'_j x_j + \sum_{j \neq j_0} \beta'_j x_j^2 + 2 \sum_{\substack{j \neq j_0 \\ i \neq j \\ i \neq j_0}} \sum \gamma_{ji} x_j x_i$$

4.4.1. Méthode pour calculer l'expression $\sum_j \sum_i \gamma_{ji} x_j x_i$.

$$2 \sum_{\substack{j \neq j_0 \\ i \neq j \\ i \neq j_0}} \sum a_{i_0 j} a_{i_0 i} x_j x_i$$

L'expression avec $a_{i_0 j_0} a_{i_0 j} = 0 \quad \forall j = 1, 2, \dots$ peut être calculée de la manière

suivante : le produit cartésien de l'ensemble $\{a_{i_0 j}\}$ excluant la diagonale et la ligne du pivot, en utilisant la table suivante :

Table 1. Calcul des valeurs de y^* .

	$a_{i_0 1}$	$a_{i_0 2}$			$a_{i_0 i_0}$				$a_{i_0 n+m}$
$a_{i_0 1}$	0	$a_{i_0 1} a_{i_0 2} x_1 x_2$			0				$a_{i_0 1} a_{i_0 n+m} x_1 x_{n+m}$
$a_{i_0 2}$	$a_{i_0 2} a_{i_0 1} x_2 x_1$	0			0				$a_{i_0 2} a_{i_0 n+m} x_2 x_{n+m}$
			0		0				

				0	0				
$a_{i_0 j_0}$	0	0	0	0	0	0	0	0	0
					0	0			
					0		0		
					0			0	
$a_{i_0 n+m} a_{i_0 n+m} a_{i_0 1} x_{n+m} x_1$					0				0

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

$$\frac{b_{i_0}}{a_{i_0 j_0}} \sum_{j \neq j_0} (\gamma_{j_0 j} + \gamma_{j j_0}) x_j - \frac{1}{a_{i_0 j_0}} \sum_{j \neq j_0} a_{i_0 j} (\gamma_{j_0 j} + \gamma_{j j_0}) x_j^2 - \frac{1}{a_{i_0 j_0}} \sum_{i \neq j_0} \sum_{i \neq j} a_{i_0 j} (\gamma_{i j_0} + \gamma_{j_0 i}) x_i x_j,$$

d'où la nouvelle forme des coefficients pour n'importe quelle phase :

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

$$\beta'_j = \beta_j + \frac{\beta_{j_0}}{(a_{i_0 j_0})^2} (a_{i_0 j})^2 - \frac{1}{a_{i_0 j_0}} a_{i_0 j} (\gamma_{j_0 j} + \gamma_{j j_0}), \text{ si } j \neq j_0 \text{ et } \beta'_{j_0} = 0 \text{ sinon}$$

$$\gamma_{ji} = \gamma_{ij}^* - \frac{a_{i_0 j}}{a_{i_0 j_0}} (\gamma_{j_0 j} + \gamma_{j j_0}),$$

pour $j \neq i, j \neq j_0, i \neq j_0$ et $\gamma_{j_0 i} = \gamma_{i j_0} = \gamma_{ii} = 0 \quad \forall i = 1, 2, \dots$, où $\gamma_{ij}^* = \beta_{j_0} \left(\frac{\theta_0}{b_{i_0}} \right)^2 a_{i_0 j} a_{i_0 i}$.

Remarque 4.5. Généralement $\gamma_{ij} \neq \gamma_{ji}$, mais $\gamma_{ij}^* = \gamma_{ji}^*$.

γ_{ij}^* , est pris à partir du tableau du produit cartésien de la ligne du pivot multipliée par $\beta_{j_0} \left(\frac{\theta_0}{b_{i_0}} \right)^2$.

Pour plus de détails voir [4]. Pour la matrice des contraintes, ces coefficients seraient changés de la même manière que celle de la méthode du simplexe [12].

Exemple 4.4.:

Soit le problème quadratique suivant (standardisé):

$$(P) = \begin{cases} \max \psi(x) = x_1 + 2x_2 + x_1^2 + 3x_2^2 \\ -x_1 + x_2 + x_3 = 3 \\ x_1 - x_2 + x_4 = 6 \\ x_1 + 2x_2 + x_5 = 12 \\ x_1, x_2, x_3 \geq 0 \end{cases}$$

La fonction est $\psi(x) = x_1 + 2x_2 + x_1^2 + 3x_2^2 + 0$

Nous passerons du point extrême (0, 0, 0) au second (6, 0, 0).

La table 2.1 contient les valeurs initiales de α , et β , prises du problème (P) et nous calculons les valeurs de θ , Δ à l'aide des expressions:

$$\theta_{i_0} = \min_{a_{ij} > 0} \left\{ \frac{b_i}{a_{ij}} \right\}, \quad \Delta_j = \alpha_j \theta_j + \beta_j \theta_j^2,$$

$$\theta_1 = \min \left\{ \infty, \frac{6}{1}, \frac{12}{1} \right\} = 6; \quad \theta_2 = \min \left\{ \frac{3}{1} \infty, \frac{12}{2} \right\} = 3$$

$$\Rightarrow \theta_{i_0} = \theta_2 = 3; \quad i_0 = 2; \quad x_4 \text{ est la variable qui sort de la base.}$$

$$\Delta_1 = \alpha_1 \theta_1 + \beta_1 \theta_1^2 = 1 * 6 + 1 * (6)^2 = 6 + 36 = 42$$

$$\Delta_2 = \alpha_2 \theta_2 + \beta_2 \theta_2^2 = 2 * 3 + 3 * (3)^2 = 6 + 27 = 33$$

$$\Delta_{j_0} = \max \{ \Delta_1, \Delta_2 \} = \max \{ 42, 33 \} = 42 \Rightarrow j_0 = 1 \Rightarrow x_1 \text{ est la variable entrante.}$$

Posons $z = z + \Delta_{j_0}$; $z = 0 + 42$ qui est la valeur de la fonction objectif à cette itération.

Le pivot est alors $a_{i_0 j_0} = a_{21}$.

On va calculer les nouveaux coefficients de la fonction objectif α', β', γ' , les nouvelles valeurs de la matrice A, les nouvelles valeurs du vecteur des ressources et donc la nouvelle solution de base.

$$\theta_0 = \frac{b_{i_0}}{a_{i_0 j_0}} = \frac{6}{1} = 6$$

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

Table 2.1: valeurs initialise de α , β , θ , Δ .

x_1	x_2	x_3	x_4	x_5	
1	2	0	0	0	α

Table 2.2: calcul de γ_{24}^*

1	3	0	0	0	β
6	3				θ
42	33				Δ
-1	1	1	0	0	$x_3=3$
1	-1	0	1	0	$x_4=6$
1	2	0	0	1	$x_5=12$

	1	-1	0	1	0
1	0	0	0	0	0
-1	0	0	0	-1	0
0	0	0	0	0	0
1	0	-1	0	0	0
0	0	0	0	0	0

La table 2.2 contient le produit cartésien de l'ensemble $\{a_{i_0j}\}_j = \{1, -1, 0, 1, 0\}$ excluant la diagonale et la ligne du pivot. (i_0 est l'indice de la variable sortante ; j_0 est l'indice de la variable entrante ; on commence par mettre à des zéros dans la diagonale, dans la ligne et la colonne du pivot).

Appliquons la formule de calcul des γ_{ij}^* des variables hors bases :

$$\gamma_{ij}^* = a_{i_0j} a_{i_0i} \beta_{j_0} \left(\frac{\theta_0}{b_{i_0}} \right)^2 \quad \text{où} \quad \theta_0 = \frac{b_{i_0}}{a_{i_0j_0}} = \frac{b_2}{a_{21}} = \frac{6}{1} = 6$$

$$\gamma_{24}^* = a_{24} a_{22} \beta_1 \left(\frac{\theta_0}{b_2} \right)^2 = (-1) * \left(\frac{6}{6} \right)^2 * 1 = -1; \gamma_{42}^* = \gamma_{24}^*; \gamma_{24}' = \gamma_{24}^* - 0 \text{ parce que } \gamma_{14} = \gamma_{41} = 0;$$

$$\alpha_2' = \alpha_2 - \left(\alpha_1 + 2\beta_1 \theta_0 \right) \frac{a_{22}}{a_{21}} = 2 - (1 + 2 * 1 * 6) \frac{-1}{1} = 2 + (13) = 15$$

$$\beta_2' = \beta_2 + \frac{\beta_1 (a_{22})^2}{(a_{21})^2} = 3 + \frac{1 * (-1)^2}{(1)^2} = 4$$

$$\alpha_4' = \alpha_4 - \left(\alpha_1 + 2\beta_1 \theta_0 \right) \frac{a_{24}}{a_{21}} = 0 - (1 + 2 * 1 * 6) \frac{1}{1} = -13$$

$$\beta_4' = \beta_4 + \frac{\beta_1 (a_{24})^2}{(a_{21})^2} = 0 + \frac{1 * (1)^2}{(1)^2} = 1$$

$$\alpha_2' = 15 \quad \alpha_4' = -13 \quad \beta_2' = 4 \quad \beta_4' = 1$$

$$\gamma_{42} = \beta_1 \left(\frac{\theta_0}{b_2} \right)^2 a_{22} a_{24} = 1 * \left(\frac{6}{6} \right)^2 * (-1) * 1 = -1 \quad \gamma_{24} = \beta_1 \left(\frac{\theta_0}{b_2} \right)^2 a_{24} a_{22} = 1 * \left(\frac{6}{6} \right)^2 * 1 * (-1) = -1$$

Ici nous avons $\gamma_{24} = \gamma_{42} = -1$ et $\gamma_{24} x_2 x_4 + \gamma_{42} x_4 x_2 = -2x_2 x_4$.

La fonction devient $\psi(x) = 15x_2 - 13x_4 + 4x_2^2 + x_4^2 - 2x_2 x_4 + 42$.

Calculons les valeurs des b_1, b_2, b_4 avec l'expression $b_i' = b_i - \frac{a_{ij_0}}{a_{i_0j_0}} b_{i_0}$.

$$b_1' = b_1 - \frac{a_{11}}{a_{21}} b_2 = 3 - \frac{-1}{1} 6 = 9$$

$$b_3' = b_3 - \frac{a_{31}}{a_{21}} b_2 = 12 - \frac{1}{1} 6 = 6 \quad b_2' = \theta_0 = \frac{b_{i_0}}{a_{i_0j_0}} = \frac{b_2}{a_{21}} = \frac{6}{1} = 6 \quad (\text{car } i_0 = 2)$$

$$a_{ij}' = a_{ij} - \frac{a_{ij_0}}{a_{i_0j_0}} a_{i_0j}$$

Calculons les coefficients a_{ij} avec

$$a_{11}' = a_{11} - \frac{a_{11}}{a_{21}} a_{21} = -1 - (-1) = 0 \quad a_{12}' = a_{12} - \frac{a_{11}}{a_{21}} a_{22} = 1 - \frac{-1}{1} (-1) = 0$$

$$a_{13}' = a_{13} - \frac{a_{11}}{a_{21}} a_{23} = 1 - \frac{-1}{1} * 0 = 1 \quad a_{14}' = a_{14} - \frac{a_{11}}{a_{21}} a_{24} = 0 - \frac{-1}{1} * 1 = 1$$

$$a_{15}' = a_{15} - \frac{a_{11}}{a_{21}} a_{25} = 0 - \frac{-1}{1} 0 = 0$$

Pour $i_0 = 2$ on appliqué la formule $a_{i_0j}' = \frac{a_{i_0j}}{a_{i_0j_0}}$

$$a_{21}' = \frac{a_{21}}{a_{21}} = 1 \quad a_{22}' = \frac{a_{22}}{a_{21}} = \frac{-1}{1} = -1 \quad a_{23}' = \frac{a_{23}}{a_{21}} = \frac{0}{1} = 0 \quad a_{24}' = \frac{a_{24}}{a_{21}} = \frac{1}{1} = 1 \quad a_{25}' = \frac{a_{25}}{a_{21}} = \frac{0}{1} = 0$$

$$a_{31}' = a_{31} - \frac{a_{31}}{a_{21}} a_{21} = 1 - 1 = 0 \quad a_{32}' = a_{32} - \frac{a_{31}}{a_{21}} a_{22} = 2 - \frac{1}{1} * (-1) = 3$$

$$a_{33}' = a_{33} - \frac{a_{31}}{a_{21}} a_{23} = 0 - \frac{1}{1} * 0 = 0 \quad a_{34}' = a_{34} - \frac{a_{31}}{a_{21}} a_{24} = 0 - \frac{1}{1} * 1 = -1$$

$$a_{35}' = a_{35} - \frac{a_{31}}{a_{21}} a_{25} = 1 - \frac{1}{1} * 0 = 1$$

Table 3.1: calcul des valeurs de

$\alpha, \beta, \theta, \Delta$ pour la deuxième phase.

x_1	x_2	x_3	x_4	x_5	
0	15	0	-13	0	α
0	4	0	1	0	β
	2		6		θ
	46		-42		Δ
0	0	1	1	0	$x_3=9$
1	-1	0	1	0	$x_1=6$
0	3	0	-1	1	$x_5=6$

La fonction

$$\psi(x) = \frac{-20}{3} x_4 - \frac{31}{3} x_5 + \frac{7}{9} x_4^2 + \frac{4}{9} x_5^2 - \frac{2}{9} x_4 x_5 + 88.$$

Table 3.2: calcul de γ_{45}^*

	0	3	0	-1	1
0	0	0	0	0	0
3	0	0	0	0	0
0	0	0	0	0	0
-1	0	0	0	0	-1
1	0	0	0	-1	0

devient

Pour passer au point extrême suivant (8, 2, 0), la table 3.1 donne les nouvelles valeurs de α , β , θ , Δ .

La table 3.2 contient le produit cartésien de l'ensemble $(a_{i_0j}) = (0, 3, 0, -1, 1)$ excluant la diagonale et

la ligne du pivot. $\gamma_{54}^* = \gamma_{45}^* = (-1) * \left(\frac{\beta'_{i_0}}{a'_{i_0j_0}} \right) = (-1) * \left(\frac{2}{3} \right)^2 * 1 = -\frac{4}{9}$;

donc $\gamma_{45} = -\frac{4}{9} - \frac{a_{34}}{3} [\gamma_{25} + \gamma_{52}] = -\frac{4}{9}$ parce que $\gamma_{25} = \gamma_{52} = 0$ dans l'expression précédente de ψ ;

$$\gamma_{54} = -\frac{4}{9} - \frac{a_{35}}{3} [\gamma_{24} + \gamma_{42}] = -\frac{4}{9} - \frac{1}{3} [-1 - 1] = \frac{2}{3} - \frac{4}{9} = \frac{2}{9} \quad \gamma_{45} = -\frac{4}{9} \text{ et } \gamma_{54} = \frac{2}{9} ;$$

nous constatons que $\gamma_{45} \neq \gamma_{54}$ mais $\gamma_{54}^* = \gamma_{45}^*$.

$$\gamma_{45} x_4 x_5 + \gamma_{54} x_5 x_4 = -\frac{4}{9} x_4 x_5 + \frac{2}{9} x_4 x_5 = -\frac{2}{9} x_4 x_5$$

$$\alpha'_4 = -13 - [15 + 2 * 2 * 4] \frac{a_{34}}{3} + \frac{6}{3} (\gamma_{24} + \gamma_{42}) = -13 - \frac{31}{3} (-1) + \frac{6}{3} (-1 - 1) = -\frac{20}{3} ;$$

$$\alpha'_5 = 0 - \left[\frac{31}{3} \right] a_{35} + \frac{6}{3} (\gamma_{25} + \gamma_{52}) = -\frac{31}{3} (1) + \frac{6}{3} (0) = -\frac{31}{3} .$$

$$\beta'_4 = 1 + \frac{4}{9} (a_{34})^2 - \frac{1}{3} a_{34} (\gamma_{24} + \gamma_{42}) = 1 + \frac{4}{9} (-1)^2 - \frac{1}{3} (-1) (-1 - 1) = 1 + \frac{4}{9} - \frac{2}{3} = \frac{7}{9}$$

$$\beta'_5 = 0 - \left[\frac{4}{9} \right] (a_{35})^2 - \frac{1}{3} a_{35} (\gamma_{25} + \gamma_{52}) \text{ mais } \gamma_{25} = \gamma_{52} = 0$$

$$\text{donc } \beta'_5 = \left[\frac{4}{9} \right] (1)^2 = \frac{4}{9} . \quad \text{Ici } \gamma_{45} = \gamma_{54} = -\frac{1}{9} .$$

Passage au point extrême suivant (2, 5, 9)

La table 4.2 contient le produit cartésien de l'ensemble $(a_{i_0j}) = (0, 0, 1, 1, 0)$ excluant la diagonale et la

ligne du pivot. $\gamma_{35}^* = \gamma_{53}^* = 0$;

$$\text{donc } \gamma_{35} = 0 - \frac{a_{13}}{1} [\gamma_{45} + \gamma_{54}] = -\frac{1}{1} \left[-\frac{4}{9} + \frac{2}{9} \right] = \frac{2}{9} \quad \gamma_{53} = 0 - \frac{a_{15}}{1} [\gamma_{43} + \gamma_{34}] = 0 ,$$

parce que $\gamma_{43} = \gamma_{34} = 0$ dans l'expression précédente de ψ et $a_{15} = 0$.

Notons que $\gamma_{35} = \frac{2}{9}$ et $\gamma_{53} = 0$ donc $\gamma_{35} \neq \gamma_{53}$.

$$\alpha'_3 = 0 - \left[\frac{20}{3} + 2 * 9 * \frac{7}{9} \right] \frac{a_{13}}{1} + \frac{9}{1} (\gamma_{43} + \gamma_{34}) = - \left[\frac{20}{3} + 14 \right] (1) + \frac{9}{1} (0) = -\frac{22}{3} ;$$

$$\alpha'_5 = -\frac{31}{3} - \left[\frac{20}{3} + 14 \right] \frac{a_{15}}{1} + \frac{9}{1} (\gamma_{45} + \gamma_{54}) = -\frac{31}{3} + \left[-\frac{20}{3} + 14 \right] * 0 + 9 \left[-\frac{1}{9} - \frac{1}{9} \right] = -\frac{37}{3} .$$

$$\beta'_3 = 0 + \frac{7}{9} * \frac{1}{(1)^2} (a_{13})^2 - \frac{1}{1} a_{13} (\gamma_{43} + \gamma_{34}) = \frac{7}{9} (1) = \frac{7}{9}$$

$$\beta'_5 = \frac{4}{9} + \left[\frac{7}{9} \right] \frac{(a_{15})^2}{1} - \frac{1}{1} a_{15} (\gamma_{45} + \gamma_{54}) = \frac{4}{9} + \frac{7}{9} * 0 = \frac{4}{9}$$

Table 4.1: calcul des valeurs de

$\alpha, \beta, \theta, \Delta$ pour la troisième phase.

x_1	x_2	x_3	x_4	x_5	
0	0	0	-	-	α
			20/4	31/3	
0	0	0	7/9	4/9	β
			9	6	θ
			3	-46	Δ
0	0	1	1	0	$x_3=9$
1	0	0	2/3	1/3	$x_1=8$
0	1	0	-1/3	1/3	$x_2=2$

La table 4.1
nouvelles
 Δ .

La forme finale de la fonction ψ devient:

$$\psi(x) = -\frac{22}{3} x_3 - \frac{37}{3} x_5 + \frac{7}{9} x_3^2 + \frac{4}{9} x_5^2 + \frac{2}{9} x_3 x_5 + 91.$$

Table 4.2: calcul de γ_{35}^*

	0	0	1	1	0
0	0	0	0	0	0
0	0	0	0	0	0
1	0	0	0	0	0
1	0	0	0	0	0
0	0	0	0	0	0

donne les
valeurs de $\alpha, \beta, \theta,$

Passage au point extrême suivant.

La table 5.1 donne les nouvelles valeurs de $\alpha, \beta, \theta, \Delta$.

Tous les $\Delta < 0$ **stop**. La dernière valeur de $\psi(x)$ est la solution optimale.

Table 5.1: calcul des valeurs de

$\alpha, \beta, \theta, \Delta$ pour la dernière phase

x_1	x_2	x_3	x_4	x_5	
0	0	-22/3	0	-37/3	α
0	0	7/9	0	4/9	β
		9		6	θ

		-3		-58	Δ
0	0	1	1	0	$x_4=9$
1	0	$-2/3$	0	$1/3$	$x_1=2$
0	1	$1/3$	0	$1/3$	$x_2=5$

Remarque 4.5: Nous utilisons la même technique pour calculer les nouveaux coefficients de f lors du passage d'un point extrême à un point extrême voisin.

CONCLUSION

Ce rapport de thèse a été l'occasion de présenter un cadre général de la programmation quadratique. Ce cadre a ensuite donné lieu à la présentation détaillée de la nouvelle méthode de résolution des problèmes quadratiques à fonction objectif convexe ou non convexe. Nous avons commencé par rappeler les outils des ensembles convexes et des fonctions convexes pour les appliquer dans un contexte d'optimisation non linéaire. Ceci nous a permis de nous placer dans un contexte de programmation quadratique.

Nous avons ensuite présenté certaines méthodes de calcul de la solution optimale en programmation 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 problèmes de classification.

Les algorithmes de résolution des problèmes d'optimisation quadratiques sont très complexes et très lourds à mettre en œuvre pour les problèmes de grande taille ou pour la résolution en temps réel.

Des techniques de linéarisation, des versions spécialisées de l'algorithme de Frank-Wolfe, les méthodes Newton ou quasi-Newton, sont les algorithmes les plus employés jusque là pour la résolution de ces problèmes quadratiques.

Des formes particulières de ce problème ont été étudiées pour réduire les temps de calcul. Un de ces cas est le problème des flots à coût minimum avec des coûts quadratiques, plus précisément avec des coûts quadratiques séparables,

L'algorithme polynomial proposé dans [21], consiste à faire des approximations successives de la fonction coût par des fonctions linéaires par morceaux jusqu'à ce que l'optimum ou un quasi-optimum soit trouvé.

La méthode de résolution est basée sur la dualité entre les flots non linéaires séparables dans un réseau et le diagramme des arcs non conformes.

En développant plusieurs exemples à l'aide de notre méthode, nous avons montré la simplicité des calculs et surtout la précision des résultats obtenus après un nombre d'itérations très limitées. Nous avons explicité les conditions d'utilisation de notre méthode en démontrant l'existence de solutions et les techniques permettant de les calculer:

- a. Après 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' (nouveau convexe obtenu par la translation T).
- c. L'extrémum de la fonction convexe ψ est un sommet ; la technique de calcul est développée dans le chapitre 3.
- d. Nous avons aussi donné la technique de 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 3.

Le théorème reste valable si le convexe Ω est remplacé par un convexe fermé et borné inférieurement dans R^n , il permet de résoudre les problèmes quadratiques dont les contraintes ne sont pas nécessairement linéaires.

Les exemples suivant ont été développés par d'autres méthodes et celle présentée dans cette thèse pour mieux apprécier les différences.

- L'exemple suivant :
$$\begin{cases} \max f(x_1, x_2) = 5x_1 + 8x_2 - x_1^2 - 2x_2^2 \\ 3x_1 + 2x_2 \leq 6 \\ x_1, x_2 \geq 0 \end{cases}$$

a été traité par la méthode de Franck-Wolf et par notre 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.

$x^* = \left(\frac{5}{6}, \frac{7}{6}\right)$ 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 notre méthode, 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 :

$x^* = \left(\frac{5}{2}, 2\right) \notin \Omega$, le calcul de $\bar{x}$ nous donne $\bar{x} = \left(1, \frac{3}{2}\right)$ d'où

$$\max_{x \in \Omega} \varphi(x) = 11.50$$

- Concernant la méthode de linéarisation ; l'exemple ci-après :

$$\begin{cases} \max f(x_1, x_2) = 5x_1 + 5x_2 - x_1^2 - x_2^2 \\ x_1 + 2x_2 - 8 \leq 0 \\ 3x_1 + x_2 - 9 \leq 0 \\ x_1, x_2 \geq 0 \end{cases}$$

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 $\bar{x} = (2.2, 2.4)$ $f(\bar{x}) = f(2.2, 2.4) = 12.4$.

Vu le nombre important de tableaux et de variables d'écart ajouter on ne peut se permettre de présenter cette méthode

Quand à notre méthode, 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 :

$x^* = \left(\frac{5}{2}, \frac{5}{2}\right) \notin \Omega$, le calcul de $\bar{x}$ nous donne $\bar{x} = (2.2, 2.4)$ d'où

$$\max_{x \in \Omega} \varphi(x) = 12.40.$$

Nous vous invitons à méditer sur la complexité des deux méthodes.

De plus la méthode présentée dans cette thèse donne une interprétation géométrique de la solution optimale $\bar{x}$: 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 :

$$\begin{cases} \min f(x_1, x_2) = -20x_1 + 10x_2 + 3x_1^2 + 2x_2^2 \\ 2x_1 - x_2 \leq 6 \\ -x_1 + x_2 \leq 10 \\ 2x_1 + 3x_2 \leq 8 \\ x_1, x_2 \geq 0 \end{cases}$$

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 idée du temps de calcul si le problème comportait 50 variables de décision.

D'un point de vu algorithmique, notons que la recherche de la solution est simple et dispense des lourds calculs tels les conditions de Kuhn, du lagrangien et autre matrice hesienne.

Les travaux présentés dans cette thèse permettent de traiter toutes les situations modélisables sous forme de problèmes de programmation quadratique.

ANNEXE.

A. Algorithme.

Notons par Z la valeur de la fonction objectif ; elle est initialisée à 0 ;

γ_{ij} sont les coefficients des termes $x_i x_j$ initialisés à 0 .

I. Si $\beta_i = 0$ pour tout $1 \leq i \leq n$. On applique l'algorithme du simplexe. **Stop.**

Si non

début

Si $\beta_i \geq 0$ pour tout $1 \leq i \leq n$ et α_i quelconque, alors deux cas sont possibles.

début

Si $a_{ki} < 0$ pour tout $1 \leq k \leq m$ et $1 \leq i \leq n$ alors on s'arrête : ce programme n'admet pas d'optimum. **Stop.**

Si non

début

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

1. Calculer $\theta_j = \min_{a_{kj}} \left\{ \frac{b_k}{a_{kj}} \right\}$ pour tous les k tel que $a_{kj} > 0$.

Choisir i_0 tel que $\theta_{i_0} = \frac{b_{i_0}}{a_{i_0 j}}$, x_{i_0} , est alors 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 $\Delta_{j_0} = \max_j \Delta_j$.
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 **on s'arrête** : ce programme est **optimal**.
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 $\alpha'_j = \alpha_j - (\alpha_{j_0} + 2\beta_{j_0} \theta_{j_0}) \frac{a_{i_0 j}}{a_{i_0 j_0}} + \frac{b_{i_0}}{a_{i_0 j_0}} (\gamma_{j_0 j} + \gamma_{j j_0})$.
10. Poser $\alpha'_{j_0} = 0$.
11. $\beta'_j = \beta_j + \frac{\beta_{j_0}}{(a_{i_0 j_0})^2} - \frac{1}{a_{i_0 j_0}} a_{i_0 k} (\gamma_{j_0 j} + \gamma_{j j_0})$.
12. Poser $\beta'_{j_0} = 0$.
13. Calculer $\gamma'_{ji} = \beta_{j_0} \left(\frac{\theta_{i_0}}{b_{i_0}} \right)^2 a_{i_0 j} a_{i_0 i} - \frac{a_{i_0 j}}{a_{i_0 j_0}} (\gamma_{j_0 i} + \gamma_{i j_0})$
14. Poser $\gamma_{j_0 i} = \gamma_{i j_0} = \gamma_{ii} = 0$ pour tout i .
15. Calculer $a'_{ij} = a_{ij} - \frac{a_{ij_0}}{a_{i_0 j_0}} a_{i_0 j}$, si $i \neq i_0$.
16. Calculer $b'_i = b_i - \frac{a_{ij_0}}{a_{i_0 j_0}} b_{i_0}$, si $i \neq i_0$.
17. Calculer $a'_{i_0 j} = \frac{a_{i_0 j}}{a_{i_0 j_0}}$.

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

Fin

Fin

Fin

B. Exemples.

Exemple 1:

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

$$\text{s.c.q.} \begin{cases} x_1 + 2x_2 - 3x_3 + x_4 = 12 \\ 2x_1 + 3x_2 - x_3 + x_5 = 8 \\ 3x_1 + x_2 - 3x_3 + x_6 = 15 \\ x \geq 0 \end{cases}$$

1	-1	4	0	0	0	α
1	1	1	0	0	0	β
4	8/3	$+\infty$				θ
20	4.44	$+\infty$				Δ
1	2	-3	1	0	0	$x_4 = 12$
2	3	-1	0	1	0	$x_5 = 8$
3	1	-1	0	0	1	$x_6 = 15$

Notre système n'admet pas de solution.

Exemple 2 :

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

$$\text{s.c.q.} \quad \begin{cases} x_1 - x_2 + x_3 = 2 \\ 2x_1 + x_2 + x_4 = 4 \\ -3x_1 + 2x_2 + x_5 = 6 \\ x \geq 0 \end{cases}$$

$$z = 3x_1 - 2x_2 + x_1^2 + x_2^2 + 0$$

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

$$z = 5x_2 - 7x_3 + 2x_2^2 + x_3^2 - 2x_2x_3 + 10$$

0	5	-7	0	0	α
0	2	1	0	0	β
	0	2			θ
	0	-10			Δ

1	-1	1	0	0	$x_1 = 2$
0	3	-2	1	0	$x_4 = 0$
0	1	3	0	1	$x_5 = 12$

Tous les Δ_j sont négatifs ou nuls la solution optimale est la précédente.

Exemple 3 :

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

$$\text{s.c.q.} \quad \begin{cases} x_1 + 2x_2 + 4x_3 + x_4 & = 4 \\ 3x_1 + x_2 + 5x_3 + x_5 & = 9 \\ x_1 + 3x_2 + x_3 + x_6 & = 2 \\ x \geq 0 \end{cases}$$

$$z = 2x_1 + 2x_3 + x_1^2 + 2x_2^2 + x_3^2 + 0$$

2	0	2	0	0	0	α
---	---	---	---	---	---	----------

1	1	1	0	0	0	β
2	2/3	1				θ
8	4/9	3				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	$x_6 = 2$

$$z = -18x_2 - 4x_3 - 6x_6 + 2x_2^2 + 2x_3^2 + x_6^2 + 6x_2x_3 + 6x_2x_6 + 2x_3x_6 + 8$$

0	-18	-4	0	0	-6	α
0	10	2	0	0	1	β
	2/3	2/3			2	θ
	-11.11	-1.78			-8	Δ
0	-1	3	0	0	-1	$x_4 = 2$
0	-8	2	1	1	-3	$x_5 = 3$
1	3	1	0	0	1	$x_1 = 2$

Tous les Δ_j sont négatifs la solution optimale est la précédente.

Exemple 4 :

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

$$\text{s.c.q.} \quad \begin{cases} x_1 + 2x_2 + 4x_3 + x_4 & = 4 \\ 3x_1 + x_2 + 5x_3 + x_5 & = 9 \\ x_1 + 3x_2 + x_3 + x_6 & = 2 \\ & x \geq 0 \end{cases}$$

$$z = 2x_1 + x_2 + 2x_3 + x_2^2 + x_3^2 + 0$$

2	1	2	0	0	0	α
0	1	1	0	0	0	β
2	2/3	1				θ
4	10/9	3				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			2	θ
	-26/9	4/9			-4	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	$x_1 = 2$

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

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 négatifs la solution optimale est la précédente.

Exemple 5 :

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

$$\text{s.c.q.} \quad \begin{cases} x_1 + 3x_2 \leq 18 \\ x_1 + x_2 \leq 8 \\ 2x_1 + x_2 \leq 14 \\ x \geq 0 \end{cases}$$

$$z = 2x_1 + 3x_2 + x_1^2 + 2x_2^2 + 0$$

2	3	0	0	0	α
1	2	0	0	0	β
7	6				θ
63	90				Δ
1	3	1	0	0	$x_3 = 18$
1	1	0	1	0	$x_4 = 8$

2	1	0	0	1	$x_5 = 14$
---	---	---	---	---	------------

$$z = -7x_1 - 9x_3 + 11/9x_1^2 + 2/9x_3^2 + 4/9x_1x_3 + 90$$

-7	0	-9	0	0	α
11/9	0	2/3	0	0	β
3		18			θ
-10		-90			Δ
1/3	1	1/3	0	0	$x_2 = 6$
2/3	0	-1/3	1	0	$x_4 = 4$
5/3	0	-1/3	0	1	$x_5 = 8$

Tous les Δ_j sont négatifs la solution optimale est la précédente.

Exemple 6 :

$$\max z = 2x_1 - 3x_2 + 2x_1^2 + 4x_2^2$$

$$\text{s.c.q.} \quad \begin{cases} x_1 - x_2 + x_3 = 0 \\ -x_1 + 2x_2 + x_4 = 2 \\ x \geq 0 \end{cases}$$

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

$$z = 2x_1 - 3x_2 + 2x_1^2 + 4x_2^2 + 0$$

9/2	0	0	-5/2	α
3	40	0	1	β
2			2	θ
21			-1	Δ
1/2	0	1	1/2	$x_3 = 1$
-1/2	1	0	1/2	$x_2 = 1$

$$z = 9/2x_1 - 5/2x_4 + 3x_1^2 + x_4^2 - 2x_1x_4 + 1$$

0	0	-33	-23	α
0	0	12	6	β
		1	2	θ
		-21	-22	Δ
1	0	2	1	$x_1 = 2$
0	1	1	1	$x_2 = 2$

$$z = -33x_3 - 23x_4 + 12x_3^2 + 6x_4^2 + 16x_3x_4 + 27$$

Tous les Δ_j sont négatifs la solution optimale est la précédente.

Exemple 7 :

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

$$\text{s.c.q.} \quad \begin{cases} -x_1 + x_2 + x_3 = 3 \\ x_1 - x_2 + x_4 = 6 \\ x_1 + 2x_2 + x_5 = 12 \\ x \geq 0 \end{cases}$$

$$z = x_1 + 2x_2 + x_1^2 + 3x_2^2 + 0$$

1	2	0	0	α
1	3	0	0	β
6	3			θ
42	33			Δ
-1	1	1	0	$x_3 = 3$
1	-1	0	1	$x_4 = 6$
1	2	0	0	$x_5 = 12$

$$z = 15x_2 - 13x_4 + 4x_2^2 + x_4^2 - 2x_2x_4 + 42$$

0	15	0	-13	0	α
0	4	0	1	0	β
	2		6		θ
	46		-42		Δ
0	0	1	1	0	$x_3 = 9$
1	-1	0	1	0	$x_1 = 6$
0	3	0	-1	1	$x_5 = 6$

$$z = -20/3x_4 - 31/3x_5 + 7/9x_4^2 + 4/9x_5^2 - 2/9x_4x_5 + 88$$

0	0	0	-20/4	-31/3	α
0	0	0	7/9	4/9	β
			9	6	θ
			3	-46	Δ
0	0	1	1	0	$x_3 = 9$
1	0	0	2/3	1/3	$x_1 = 8$
0	1	0	-1/3	1/3	$x_2 = 2$

$$z = -22/3x_3 - 37/3x_5 + 7/9x_3^2 + 4/9x_5^2 + 2/9x_3x_5 + 91$$

0	0	0	0	-37/3	α
0	0	0	0	4/9	β
		9		6	θ
		-3		-58	Δ
0	0	1	1	0	$x_4 = 9$
1	0	-2/3	0	1/3	$x_1 = 2$
0	1	1/3	0	1/3	$x_2 = 5$

Tous les Δ_j sont négatifs la solution optimale est la précédente.

Exemple 8:

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

Dans ce cas $\beta_1 = \beta_2 = -1$, donc $\Omega = \Omega'$.

$$\frac{\partial f}{\partial x_1}(x_1, x_2) = 5 - 2x_1 = 0 \Rightarrow x_1^* = 5/2$$

$$\frac{\partial f}{\partial x_2}(x_1, x_2) = 5 - 2x_2 = 0 \Rightarrow x_2^* = 5/2$$

Le point critique $x^* = \left(\frac{5}{2}, \frac{5}{2}\right) \notin \Omega$, parce que $3\frac{5}{2} + \frac{5}{2} - 9 = 1 > 0$,

$$\bar{x} = P_{\Omega}(x^*) = x^* - \frac{\langle x^*, a \rangle - b}{\|a\|^2} a \quad \text{où } a = (3, 1), \quad b = 9.$$

$$\langle x^*, a \rangle = \frac{5}{2} * 3 + \frac{5}{2} * 1 = 10, \quad \|a\|^2 = 10.$$

$$\bar{x} = \left(\frac{5}{2}, \frac{5}{2}\right) - \frac{1}{10}(3, 1) = (2.2, 2.4)$$

$$\max_{(x_1, x_2) \in \Omega} \varphi(x_1, x_2) = \varphi(\bar{x}) = 12.40$$

Ici

$$\Omega = \{(x_1, x_2) \in \mathbb{R}^2 / 3x_1 + x_2 - 9 \leq 0, x_1 + 2x_2 - 8 \leq 0, x_1 \geq 0, x_2 \geq 0\}.$$

Exemple 9:

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

dans ce cas $\beta_1 = -1$ et $\beta_2 = -2$,

donc $\Omega \neq \Omega'$.

$$\frac{\partial f}{\partial x_1}(x_1, x_2) = 5 - 2x_1 = 0 \Rightarrow x_1^* = 5/2$$

$$\frac{\partial f}{\partial x_2}(x_1, x_2) = 8 - 4x_2 = 0 \Rightarrow x_2^* = 2$$

Le point critique $x^* = \left(\frac{5}{2}, 2\right) \notin \Omega$, parce que $3\frac{5}{2} + 2 \cdot 2 > 6$,

$$\bar{x} = P_{\Omega}(x^*) = x^* - \frac{\langle x^*, a \rangle - b}{\|a\|^2} a \quad \text{où} \quad a = (3, 2), \quad b = 6.$$

$$\langle x^*, a \rangle = \frac{5}{2} \cdot 3 + 2 \cdot 2 = \frac{23}{2}, \quad \|a\|^2 = 13.$$

$$\bar{x} = \left(\frac{5}{2}, 2\right) - \frac{11}{2 \cdot 13} (3, 2) = \left(\frac{5}{2} - \frac{33}{26}, 2 - \frac{11}{13}\right) = \left(\frac{16}{13}, \frac{15}{13}\right).$$

$$\Omega' = \left\{ (y_1, y_2) \in \mathbb{R}^2 / 3y_1 + \frac{2}{\sqrt{2}} y_2 \leq 6, y_1 \geq 0, y_2 \geq 0 \right\}.$$

De plus $y^* = \left(\frac{5}{2}, 2\sqrt{2} \right).$

La projection de ce point sur Ω' donne le point suivant y_0 :

$$y_0 = \left(1, \frac{3}{\sqrt{2}} \right) \Rightarrow \bar{x} = \left(1, \frac{3}{2} \right) \quad \text{Et} \quad \underset{x \in \Omega}{\text{Max}} \varphi(x) = \varphi \left(1, \frac{3}{2} \right) = 11,500.$$

Notons que la projection du point x^* sur Ω donne le point $\left(\frac{16}{13}, \frac{15}{13} \right),$

et $\varphi \left(\frac{16}{13}, \frac{15}{13} \right) = 11,207.$

Evidemment $\underset{x \in \Omega}{\text{Max}} \varphi(x) \leq \varphi(x^*) = \varphi \left(\frac{5}{2}, 2 \right) = 14,250.$

Le point $\left(\frac{16}{13}, \frac{15}{13} \right)$ de l'ensemble convexe Ω est le plus proche point du point critique x^* ,

mais il n'est pas la solution optimale de φ .

Exemple 10:

$$\begin{cases} \max f(x_1, x_2, x_3) = 2x_1 + 3x_2 + 3x_3 + x_1^2 + 2x_2^2 - x_3^2 \\ x_1 + 3x_2 \leq 18 \\ x_1 + x_2 + x_3 \leq 8 \\ 2x_1 + x_2 + 2x_3 \leq 14 \\ x_1, x_2, x_3 \geq 0 \end{cases}$$

Nous avons $\varphi(x_3) = 3x_3 - x_3^2$, le point critique est $x^* = \left(\frac{3}{2}\right)$

et $\max_{x_3 \in \Omega} \varphi(x_3) = \varphi\left(\frac{3}{2}\right) = \frac{9}{4}$.

D'autre part $\psi(x_1, x_2) = 2x_1 + 3x_2 + x_1^2 + 2x_2^2$.

Ψ est une fonction convexe, et sa solution optimale est un point extrême (sommet de Ω).

$$\max_{(x_1, x_2) \in \Omega} \psi(x_1, x_2) = \psi(0, 6) = 90$$

Mais

$$\max_{x \in \Omega} f(x) \leq \max_{x \in \Omega} \psi(x) + \max_{x \in \Omega} \varphi(x).$$

Par conséquent $\max f(x) \leq 92.25$.

Aussi $\left(0, 6, \frac{3}{2}\right) \in \Omega$, $f\left(0, 6, \frac{3}{2}\right) \leq \max_{x \in \Omega} f(x)$.

Alors $\max_{x \in \Omega} f(x) = f\left(0, 6, \frac{3}{2}\right) = 92,25$.

Exemple 11:

$$\begin{cases} \max f(x_1, x_2, x_3) = 2x_1 + 3x_2 + 3x_3 + x_1^2 + 2x_2^2 - 3x_3^2 \\ x_1 + 3x_2 \leq 18 \\ x_1 + x_2 + x_3 \leq 8 \\ 2x_1 + x_2 + 2x_3 \leq 14 \\ x_1, x_2, x_3 \geq 0 \end{cases}$$

Nous avons $\varphi(x_3) = 3x_3 - 3x_3^2$, le point critique est $x^* = \left(\frac{3}{2}\right)$

et
$$\max_{x_3 \in \Omega} \varphi(x_3) = \varphi\left(\frac{3}{2}\right) = \frac{9}{4}.$$

D'autre part $\psi(x_1, x_2) = 2x_1 + 3x_2 + x_1^2 + 2x_2^2.$

Ψ est une fonction convexe, et sa solution optimale est un point extrême (sommet de Ω).

$$\max_{(x_1, x_2) \in \Omega} \psi(x_1, x_2) = \psi(0, 6) = 90$$

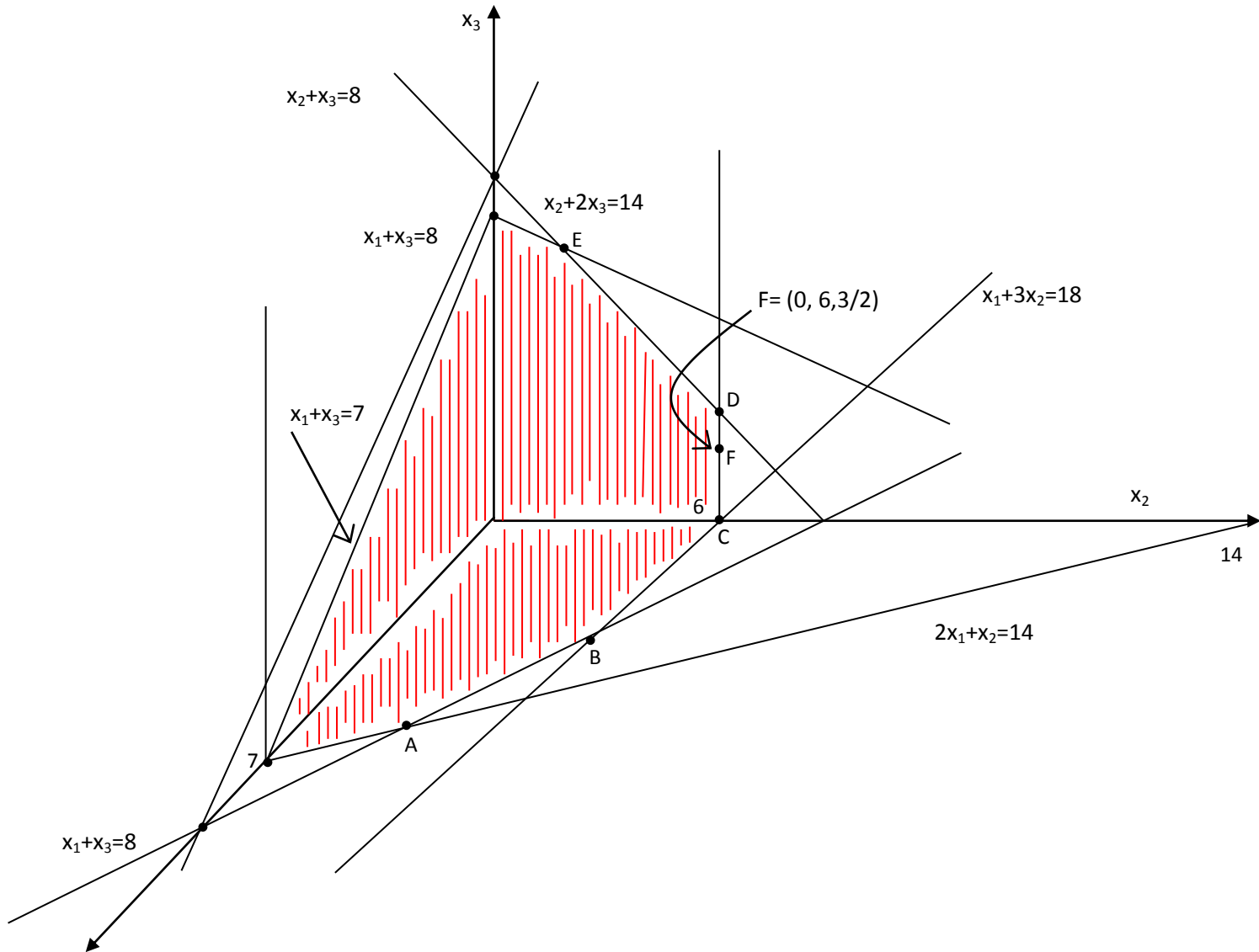
Mais
$$\max_{x \in \Omega} f(x) \leq \max_{x \in \Omega} \psi(x) + \max_{x \in \Omega} \varphi(x).$$

Par conséquent

$$\max_{x \in \Omega} f(x) \leq 92.25. \quad \text{aussi} \quad \left(0, 6, \frac{3}{2}\right) \in \Omega, \quad f\left(0, 6, \frac{3}{2}\right) \leq \max_{x \in \Omega} f(x).$$

Alors

$$\max_{x \in \Omega} f(x) = f\left(0, 6, \frac{3}{2}\right) = 92,25.$$



Exemple 12:

$$\begin{cases} \max f(x_1, x_2) = 2x_1 + 3x_2 - x_1^2 + 2x_2^2 \\ x_1 + 3x_2 \leq 18 \\ x_1 + x_2 \leq 8 \\ 2x_1 + x_2 \leq 14 \\ x_1, x_2 \geq 0 \end{cases}$$

Nous avons

$$\psi(x_2) = 3x_2 + x_2^2.$$

$$\varphi(x_1) = 2x_1 - x_1^2,$$

sommets dont la première composante est nulle ; $(0,0); (0,8); (0,6); (0,14);$

mais $(0,8)$ et $(0,14) \notin \Omega$

donc ils restent $(0,0)$ et $(0,6)$. $\bar{x}_2 = (0,0) \Rightarrow \bar{x}_1 = (x_1, 0)$

$$\left. \begin{array}{l} x_1 + 18 \leq 18 \\ x_1 + 6 \leq 8 \\ 2x_1 + 6 \leq 14 \end{array} \right\} \Rightarrow x_1 = 0 \Rightarrow \bar{x}_1 = (0,6) + (0,0) = (0,6)$$

$$f(\bar{x}) = 18 + 36 = 54$$

Les sommets de Ω sont : $(0,0); (0,6); (3,5); (6,2)$ et $(7,0)$.

Le point $(0,6)$ est un sommet.

Exemple 13:

$$\left\{ \begin{array}{l} \max f(x_1, x_2) = 2x_1 + 3x_2 - x_1^2 + x_2^2 \\ x_1 + x_2 \leq 8 \\ 2x_1 + x_2 \leq 14 \\ x_2 \leq 5 \\ x_1, x_2 \geq 0 \end{array} \right.$$

Nous avons

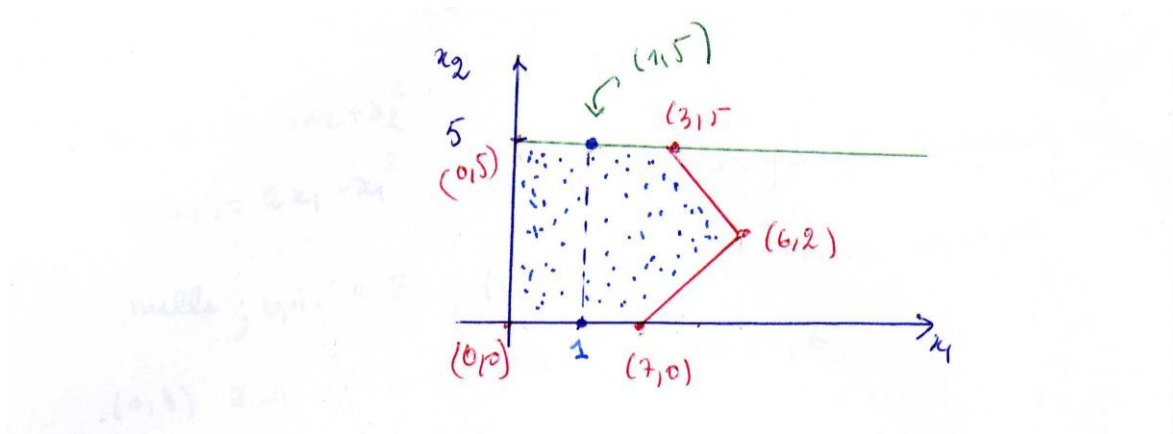
$$\psi(x_2) = 2x_2 + x_2^2, \quad \varphi(x_1) = 2x_1 - x_1^2,$$

$$\bar{x}_2 = (0,5) \Rightarrow \bar{x}_1 = (x_1, 0)$$

$$\left. \begin{array}{l} x_1 + 5 \leq 8 \\ 2x_1 + 5 \leq 14 \end{array} \right\} \Rightarrow \left\{ \begin{array}{l} x_1 \leq 3 \\ x_1 \leq \frac{9}{2} \end{array} \right. ; \varphi(x) = 2x_1 - x_1^2, \quad \frac{-\alpha_i}{2\beta_i} = \frac{-2}{-2} = 1$$

donc $\bar{x}_1 = (1,0) \Rightarrow \bar{x} = (1,5)$; $f(1,5) = 2 - 1 + 15 + 25 = 41$.

Le point $(1,5)$ n'est pas un sommet du fait qu'on a $0 \cdot x_1 + x_2 \leq 5$.



Remarque :

$$f(3,5) = 6 - 9 + 15 + 25 = 37 \quad f(0,5) = 15 + 25 = 40$$

$$t(0,5) + (1-t)(3,5) = (3-3t, 5)$$

$$f(3-3t, 5) = 6 - 6t + 15 - (3-3t)^2 + 25$$

$$f'(3-3t, 5) = 0 \Rightarrow -6 + 6(3-3t) = 0$$

$$\Rightarrow t = \frac{2}{3} \text{ et } \bar{x} = (3 - 3 \cdot \frac{2}{3}, 5) = (1, 5).$$

IMPLEMENTATION DU LOGICIEL

Pour son implémentation nous avons utilisé le langage de développement WinDev.

wo Programmation Quadratique

Saisies Données Editions P S I Editions P H I Divers Aide

Nom **Mat_A_exemple_6** Variable Décision **2** Nbre_Contrain **2** Nbre_Var_tota **4**

Matrice A

X1	X2	X3	X4	X5	X6	X7	X8	X9	X10
1,00	-1,00	1,00							
-1,00	2,00		1,00						

Vecteur b

b1	b2	b3	b4	b5	b6	b7	b8	b9
	2							

Alpha

1	2	3	4	5	6	7	8	9
2,000	-3,000							

Béta

1	2	3	4	5	6	7	8
2,0000	4,000						

Liste des Programmes Quadratiques

- Mat_A_exemple_1
- Mat_A_exemple_2
- Mat_A_exemple_3
- Mat_A_exemple_4
- Mat_A_exemple_5
- Mat_A_exemple_6**
- Mat_A_exemple_7
- Mat_A_exemple_mixte_1
- Mat_A_phi_exemple_1
- Mat_A_phi_exemple_2

Nouveau
Enregistrer
Charger
Calculs P S I
Calculs P H I
Imprimer PSI

Cette fenêtre sert à saisir les données du problème quadratique : Matrice A, Vecteur b, les coefficients Alpha, les coefficients Béta.

L'aide intégrée au logiciel est suffisamment documentée pour vous guider dans chaque étape.

Programmation Quadratique									
Saisies Données Editions P S I Editions P H I Divers Aide									
Résultats des calculs									
	X1	X2	X3	X4	X5	X6	X7		
Problème initial :									
Alpha	2,00	-3,00							
Béta	2,00	4,00							
x3 = 0	1,00	-1,00	1,00						
x4 = 2	-1,00	2,00		1,00					
Itération : 1									
Alpha	2,00	-3,00							
Béta	2,00	4,00							
Gama x1									
x2									
x3									
x4									
Téta		1,00							
Delta		1,00							
x3 = 0	1,00	-1,00	1,00						
x2 = 2	-1,00	2,00		1,00					
Cumul Delta = 1									
Itération : 2									
Alpha	4,50			-2,50					
Béta	3,00			1,00					
Gama x1				-1,00					
x2									

Programmation Quadratique									
Traitement de la fonction PHI									
Problème initial :									
$3X1 + X2 - 9 \leq 0$ $X1 + 2X2 - 8 \leq 0$ Alpha : $+ 5X1 + 5X2$ Béta : $- 1X1 - 1X2$ *****									
--- les Béta sont égaux à -1 donc : $\Omega = \Omega'$ --- Coordonnées du point Critique $X^* = \{ 2.5 , 2.5 \}$ --- Le point critique appartient à Ω . --- $X_{\text{barre}} = \{ 2.2 , 2.4 \}$									
--- Ω' : $3X1 + X2 - 9 \leq 0$ $X1 + 2X2 - 8 \leq 0$ --- $Y^* = \{ 2.5 , 2.5 \}$									
La projection de Y^* sur Ω' donne Y_0									
--- $Y_0 = \{ 2.2 , 2.4 \}$ --- $Y_{\text{barre}} = \{ 2.2 , 2.4 \}$									
--- $\Phi(Y_{\text{barre}}) = 12.4$									
--- Notons que la projection du point X^* sur Ω donne le point $X_{\text{barre}} = \{ 2.2 , 2.4 \}$ et que $\Phi(2.2 , 2.4) = 12.4$									

Prévisualisation de l'état Etat_Resultats_Prg_Quad

90 % 1 / 3

Word Excel HTML PDF XML Email Email PDF

Résultats du traitement du programme quadratique **Mat_A_exemple_7** 24/11/2009

	x1	x2	x3	x4	x5	x6	x7	x8	x9
Problème initial :									
Alpha	1	2							
Bêta	1	3							
x3 = 3	-1	1	1						
x4 = 6	1	-1		1					
x5 = 12	1	2			1				
Itération : 1									
Alpha	1	2							
Bêta	1	3							
Gama x1									
x2									
x3									
x4									
x5									
Téta	6	3							
Delta	42	33							
x3 = 3	-1	1	1						

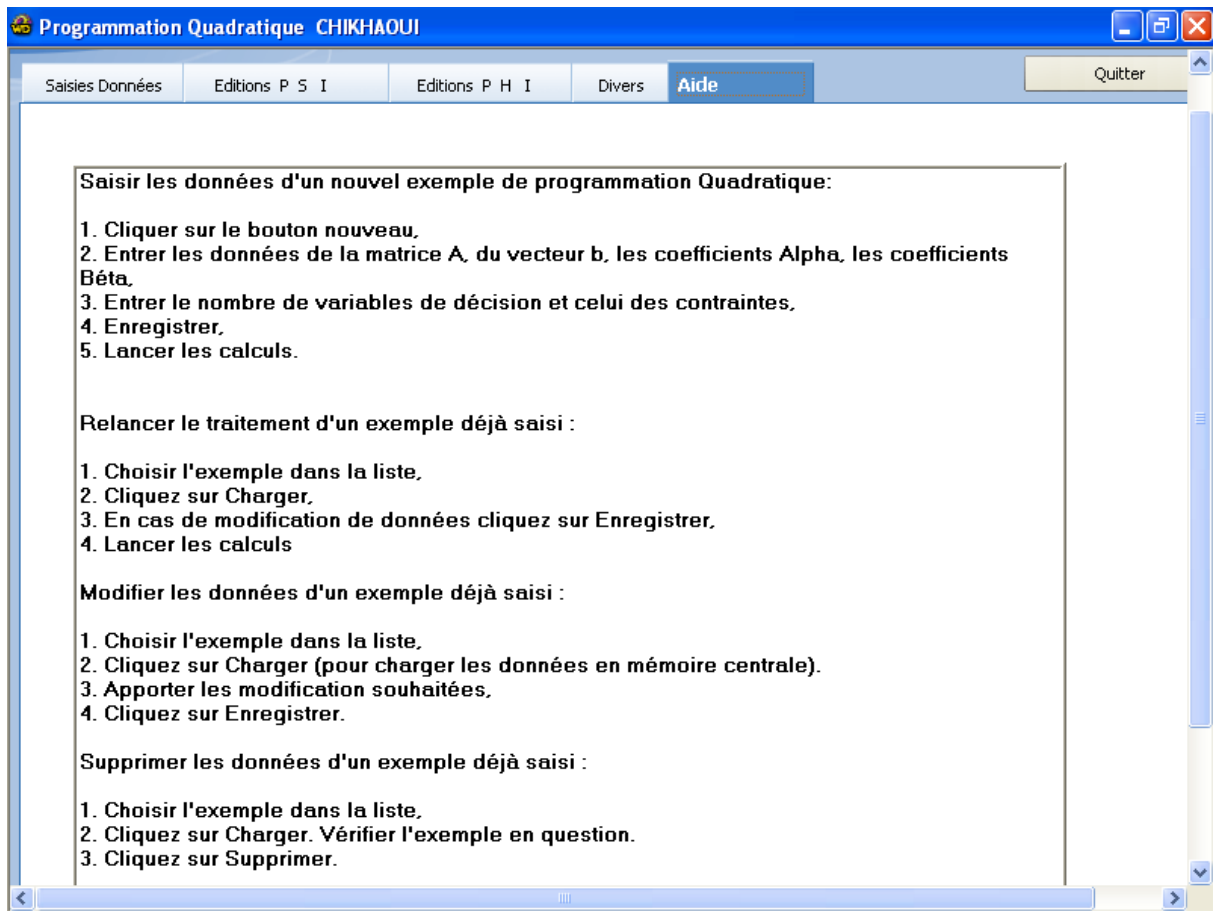
Prévisualisation de l'état Etat_Resultats_Prg_Quad

90 % 2 / 3

Word Excel HTML PDF XML Email Email PDF

	x1	x2	x3	x4	x5	x6	x7	x8	x9
Bêta		4		1					
Gama x1									
x2				-1					
x3									
x4		-1							
x5									
Téta		2		6					
Delta		46		-42					
x3 = 9			1	1					
x1 = 6	1	-1		1					
x2 = 6		3		-1	1				
Cumul Delta = 88									
Itération : 3									
Alpha				-6,66667	-10,33333				
Bêta				0,77778	0,44444				
Gama x1									
x2									
x3									
x4					0,22222				

Prévisualisation de l'état Etat_Resultats_Prg_Quad									
90 % 3 / 3									
Word	Excel	HTML	PDF	XML	Email	Email PDF			
	X1	X2	X3	X4	X5	X6	X7	X8	X9
Cumul Delta = 91									
Itération : 4									
Alpha					-10,33333				
Béta					0,44444				
Gama x1									
x2									
x3									
x4									
x5									
Téta	8				6				
Delta					-46				
x3 = 9			1	1					
x1 = 8	1			0,66667	0,33333				
x2 = 2		1		-0,33333	0,33333				
Cumul Delta = 91									
Prg OPTIMAL									
Fin de l'impression									



Aide des instructions essentielles.

BIBLIOGRAPHIE

- [1] S. Achmanov, 1984, « Programmation linéaire», Traduction française.
Edition MIR. Moscou.
- [2] V. Arnold, 2008, « Equations différentielles ordinaires», Dunod, 2^{ème} Edition
- [3] J. Bass, 1977, « Cours de mathématiques», Tome 1 fascicule 1. Masson (Paris)
- [4] A. Belabbaci, 2007, « La programmation quadratique et son application à la programmation séparable», Mémoire de magistère.
Université de Laghouat.
- [5] M. Bierlaire, 2006, «Introduction à l'optimisation différentiable »,
Press polytechniques et universitaires rometes. Lausanne.
ISBN-10 :2-88074-669-8 ; ISBN13 :978-2-88074-669-8 1^{ère} édition.
- [6] S.Boyd, V. Lieven, 1999, «Convex Optimization», Course Reader for EECS290N
UC Berkeley .
- [7] P. Caron,A. Juhel, F. Vandeveld, 1988, «Programmation linéaire. Méthodes et Applications »,Bordas. Paris.
- [8] A. Chikhaoui, 2003, « Méthode Adaptative. Mémoire de magistère », Université des Sciences & Technologie Oran.
- [9] A.Chikhaoui, B.Djebbar, A.Belabbaci, A.Mokhtari, 2009.
« Optimization of a quadratic function under its Canonical form»,
Asian Journal of Applied Sciences. 499-510 ISSN 1996-3343.
- [10] A.Chikhaoui, B.Djebbar, R.Mekki, 2010.
“ New Method for finding an optimal solution to quadratic
Programming Problem”
Journal of Applied Sciences. 10(15):1627-1631 ISSN 1812-5654.
- [11] G. Chilov, 1973, « Analyse mathématique. Fonctions d'une variable »,
3^{ème} édition. Moscou : Edition Mir, 1973.
- [12] G. Cullma, 1970, « Recherche opérationnelle», Théorie et Pratique.
Masson & Cie/Eyrolles. Paris
- [13] G.B. Dantzig, 1963, «Linear programming et extensions », Princeton university press.
Princeton, New Jersey.

- [14] D. De Werra et al, 2003, « Recherche opérationnelle pour ingénieur I », Press polytechniques et universitaires romandes. Lausanne.
- [15] E. De Klerk & D.V. Pasechnik, 2007, « A linear programming reformulation of the standard quadratic optimization problem », *Journal of Global Optimization (J Glob Optim)*, 37(1) 75-84.
DOI 10.1007/s10898-006-9037-9
- [16] F. Delbos et J. Ch. Gilbert, 2003: « Global linear convergence of an augmented Lagrangian algorithm for solving convex quadratic optimization problems », *Journal of convex analysis*, 12, 45-69.
ISSN 0249-6399 ISRN INRIA/RR--5028--FR+ENG
- [17] J. Dieudonné, 1968, « Calcul infinitésimal », Gallimard, Paris
- [18] R. Faure, 1979 « Précis de recherche opérationnelle », Dunod. Paris.
- [19] A. Henry-Labordère, 1995, « Cours de Recherche Opérationnelle », Tome 1. Presse de l'école des Ponts et chaussées. Paris.
- [20] S. Hillier et G. Lieberman, 2005: « Introduction to operations research », Eighth Edition. Chapter 12 (547-602). ISBN 007-123828-X
McGraw-Hill International Edition (Asia).
- [21] S. Ketabchi, H. Navidi et N. Mollaverdi, 2009, « Modelized Lagrangian function and Armijo rule normal solution of the LP problem », Far East Journal of Applied Sciences, 35 : 103-111. ISSN 1812-5654.
- [22] Marco Dozzi, 2004, « Méthodes non linéaires et stochastiques de la recherche opérationnelle », Chapitre I (5-22).
- [23] M. Simonnard, 1973, « Programmation linéaire. Techniques du calcul économique 2 », Extensions, 40(1) 32-45. 2^{ème} édition. Dunod. Paris.
ISBN 2-04-002689-4.
- [24] Shuonan Dong, 2006, « Methods for Constrained Optimization », 18.086 Project 2, Spring 2006.
- [25] Vo-Khac-Khoan, 1972, « Distributions analyse de Fourier, opérateurs aux dérivées partielles », Tome 1. Vuibert. Paris.
- [26] G. R. Walsh, 1979, « Methods of optimization. Printed Page Bros », Ltd Mile Crois Lavie, Norwich. ISBN 0471919225.