

République Algérienne Démocratique et Populaire
Ministère de l'Enseignement Supérieur et de la Recherche
Scientifique
Université 8 Mai 1945 Guelma

Faculté des Mathématiques et de l'Informatique
et des Sciences de la Matière
Département de Mathématiques



M/110 226

Mémoire

Présenté en vue de l'obtention du diplôme de

Master en Mathématiques

Option : EDP et Analyse numérique

Par :kerboub raounek



Intitulé

**Introduction sur le problème semi-Définie positive
(SDP)**

Dirigé par :Dr. Mousaab Bouafia

Devant le jury

PRESIDENT
RAPPORTEUR
EXAMINATEUR

Pr.chaou Abd razak
Dr: Mousaab Bouafia
Dr: Berahail Amel

Univ-Guelma
Univ-Guelma
Univ-Guelma

Session Juin 2018

Remerciement

Nous tenons tout d'abord à remercier Dieu le tout puissant et miséricordieux, qui nous a donné la force et la patience d'accomplir ce Modeste travail.

je souhaitais adresser mes remerciements les plus sincères aux personnes qui m'ont apporté leur aide et qui ont contribué à l'élaboration de ce mémoire ainsi qu'à la réussite de cette formidable année universitaire. Je tiens à remercier sincèrement Monsieur mouaab bouafia, qui, en tant que Directeur de mémoire, s'est toujours montré à l'écoute et très disponible tout au long de la réalisation de ce mémoire, ainsi pour l'inspiration, l'aide et le temps qu'il a bien voulu me consacrer et sans qui ce mémoire n'aurait jamais vu le jour. Je n'oublie pas ma famille pour leur contribution, leur soutien et leur patience. J'adresse mes plus sincères remerciements à tous mes proches et amis, qui m'ont toujours soutenue et encouragée au cours de la réalisation de ce mémoire. Merci à tous et à toutes.

Nos profonds remerciements pour les membres de jury qui ont accepté d'évaluer ce travail .

Dédicace

Je dédie ce mémoire spécialement à mon défunt père qui m'a toujours encouragé et m'avais toujours soutenu.

**Je le dédie particulièrement à ma chère mère et mes adorables sœurs et frères et à toute ma famille pour leur soutien tout au long de mon parcours universitaire, Que ce travail soit l'accomplissement de vos vœux tant allégués, et le fruit de votre soutien infailible,
Merci d'être toujours là pour moi.**

Enfin je le dédie pour tous mes amis et tous ceux qui me connaissent

Table des matières

Introduction générale	3
1 Préliminaires	5
1.1 Quelques notations	5
1.2 Rappel sur le gradient et le hessienne	6
1.2.1 Notion de la dérivée partielle	6
1.2.2 Rappel sur le gradient	7
1.2.3 Rappel sur le hessienne	7
1.3 Eléments d'analyse convexe	9
1.3.1 Ensembles et fonctions convexes	9
1.3.2 Formules de Taylor	10
1.4 Problème général d'optimisation	11
1.5 Conditions nécessaires de minimum	11
2 Programmation mathématique	14
2.1 Problème d'optimisation sans contrainte	14
2.1.1 L'existence et d'unicité	14
2.1.2 Condition d'optimalité	15
2.2 problème d'optimisation avec contrainte	18
2.3 programme mathématique	18
2.3.1 Classification et résolution d'un programme mathématique	18
2.3.2 Existence et Unicité de solution	18
2.3.3 Conditions d'optimalité	19
2.3.4 Dualité Lagrangienne	19
2.4 Algorithme d'optimisation	20
2.4.1 Description	20
2.4.2 Convergence	21
2.4.3 Taux de convergence	21
2.5 Programmation linéaire (<i>PL</i>)	22
3 La programmation semi-définie	25
3.1 Eléments de calcul matriciel	25
3.1.1 Matrices	25

3.1.2	Matrice symétrique semi-définies positives	26
3.1.3	Produit de Kronecker	28
3.1.4	le cone des matrices semi-définies positives	30
3.2	La programmation semi-définie	31
3.2.1	Les programmes semi-définis	31
3.2.2	Théorie de la dualité	34
3.3	Méthodes de points intérieurs	36
3.3.1	Introduction	36
3.3.2	Méthodes de trajectoire centrale	37
3.3.3	Etude de problème perturbé(SDP) $_{\mu}$	40
3.3.4	Condition d'optimalité pour (P_{μ})	42
	Conclusion générale	43
	Bibliographie	44

Introduction générale

Il existe de nombreuses sortes de problèmes d'optimisation. Tous ces problèmes possèdent des structures différentes et ne peuvent être traités de la même façon. Le présent projet a pour vocation de se focaliser sur les problèmes d'optimisation continues et non linéaires. Ceux-ci sont habituellement définis par un ensemble X appelé ensemble de solutions admissibles et une fonction objectif $f : X \rightarrow \mathbb{R}$ qui associe à chacun des éléments de X un nombre réel, ou un coût (nous désignerons d'ailleurs parfois également cette fonction par l'appellation fonction de coût ; au sujet de X , la dénomination d'ensemble admissible ou encore de domaine admissible sera utilisée tantôt, dans un souci d'abréviation). Nous considérons le cas où X est un sous-ensemble de \mathbb{R}^n . Ses éléments sont donc des vecteurs de nombres réels à n dimensions de la forme (x_1, x_2, \dots, x_n) , avec $x_i \in \mathbb{R} \forall i \in \{1, 2, \dots, n\}$. Le problème consiste à trouver l'élément de X dont le coût est minimal ou maximal, mais cela ne change rien à la difficulté ou aux types de méthodes employées. Cela peut être formulé de manière plus concise comme suit :

$$\begin{cases} \text{minimiser } f(x) \\ \text{sous contrainte } x \in X \end{cases}$$

X est généralement défini par une collection de contraintes exprimées sous forme d'égalités et d'inégalités. Nous admettons que f est continue et continuellement différentiable. Les dérivées seconde et première jouent un rôle important aussi bien dans la caractérisation des solutions optimales que dans les idées qui conduisent à la plupart algorithmes utilisés et implémentés au cours de ce projet. Par optimisation sans contraintes, nous désignons le cas particulier où $X = \mathbb{R}^n$.

L'optimisation semi définie positive a connu un essor important dans les années 1990 pour au moins trois raisons. D'abord, bien qu'ils soient non linéaires ; les problèmes d'optimisation (*SDP*) peuvent être résolus en un nombre d'itérations polynomial. En suite, un grand nombre de problèmes convexes ont pu être exprimés dans ce formalisme, ce qui montre par conséquent qu'ils sont polynomialement résolubles, certains problèmes non convexes ont une relation (*SDP*) précise, qui permet de leur trouver rapidement une solution approchée de qualité.

la programmation semi définie positive est l'un des problèmes de l'optimisation continue. L'étude de ce genre de problèmes a connu un fantastique regain d'intérêt depuis les années 90, entre autres parce que l'on a disposé depuis d'al-

algorithmes efficaces permettant venue est en fait devenu très récemment un axe de recherche mathématique à part entière. l'intérêt est justifié par ses différents champs d'applications compte tenu du nombre et de la variété de ces domaines d'applications. Cependant un nouveau type de méthode de résolution a fait son apparition en 1984 : Les méthodes de points intérieurs la plupart des idées sous-jacentes à ces méthodes proviennent du domaine de programmation non linéaire. Parmi leurs avantages, citons :

1. Efficacité théorique : il est possible de prouver que ces méthodes s'exécutent en temps polynomial (ce qui n'est pas le cas de l'algorithme de simplexe, de nature exponentielle (dans le problème linéaire)).
2. Efficacité pratique : les temps de calcul et de réponse de ces méthodes sont compétitifs (et battent l'algorithme du simplexe dans le cas de problèmes de grande taille (dans le problème linéaire)).
3. Traitement de très grands problèmes : ces méthodes permettent de résoudre des problèmes de très grande taille qu'aucun autre algorithme connu ne pourrait traiter en un temps acceptable.
4. Adaptation au cas non linéaire : il est possible d'adapter ces méthodes à la programmation non linéaire, plus particulièrement à la programmation convexe, ce qui permet de traiter de nouveaux types de problèmes pour les quels on ne connaissait jusqu'à présent aucune méthode de résolution efficace.

présentation de la mémoire :

Dans cette mémoire, on s'intéresse à le problème (*SDP*) en particulier, l'impact des méthodes de points intérieurs est sans précédent, dans cette mémoire on présente l'étude algorithmique de la méthode de trajectoire centrale.

Le mémoire est divisé en trois chapitres organisés comme suit :

Le premier chapitre contient un rappel introductif sur l'analyse convexe.

Le second chapitre s'intéresse la programmation mathématique.

Le troisième chapitre est consacré à la programmation semi-définie.

Chapitre 1

Préliminaires

Le but de ce premier chapitre est de présenter quelques notions fondamentales de calcul différentiel et l'analyse convexe, de la programmation linéaire et l'étude asymptotique de la convergence d'un algorithme d'optimisation. Ces notions sont utiles pour démontrer les résultats théoriques dans les chapitres suivants.

1.1 Quelques notations

1- Pour tout $n \in \mathbb{N}^*$, \mathbb{R}^n désigne l'espace euclidien $\mathbb{R} \times \mathbb{R} \times \dots \times \mathbb{R}$ ("produit n fois"). En général un vecteur $x \in \mathbb{R}^n$ sera noté $x = (x_1, x_2, \dots, x_n)^t$ (vecteur colonne).

2- On note e_1, e_2, \dots, e_n les éléments de la base canonique de \mathbb{R}^n , où e_i est le vecteur de \mathbb{R}^n donné par :

$$(e_i)_j = \delta_{ij} = \begin{cases} 0 & \text{si } i \neq j \\ 1 & \text{si } i = j \end{cases}, \quad \forall i, j \in \{1, 2, \dots, n\}.$$

(symbole de Kronecker).

3- Pour tous $x, y \in \mathbb{R}^n$ on note par $\langle x, y \rangle \in \mathbb{R}$ le produit scalaire de x et y , qui est donné par

$$\langle x, y \rangle = \sum_{i=1}^n x_i y_i.$$

4- Pour tout $x \in \mathbb{R}^n$ on note par $\|x\| \geq 0$ la norme euclidienne de x , donné par

$$\|x\| = \sqrt{\langle x, x \rangle} = \sqrt{\sum_{i=1}^n x_i^2}.$$

5- Pour tout $x \in \mathbb{R}^n$ et $r > 0$ on notera par $B(x, r)$ la boule ouverte du centre x et rayon r , donnée par

$$B(x, y) = \{y \in \mathbb{R}^n, \|y - x\| < r\}.$$

6- Si $a, b \in \mathbb{R}^n$ on note $[a, b]$ le sous-ensemble de \mathbb{R}^n donné par

$$[a, b] = \{a + t(b - a) \equiv (1 - t)a + tb, t \in [0, 1]\}.$$

Si $a, b \in \mathbb{R}$ avec $a < b$ alors on retrouve le fait que $[a, b]$ désigne l'intervalle des nombres $x \in \mathbb{R}$ tels que $a \leq x \leq b$.

7- Rappelons aussi l'inégalité de Cauchy-Schwarz :

$$|\langle x, y \rangle| \leq \|x\| \cdot \|y\| \quad \forall x, y \in \mathbb{R}^n.$$

1.2 Rappel sur le gradient et le hessienne

1.2.1 Notion de la dérivée partielle

Soit $\Omega \subset \mathbb{R}^n$ un ouvert et $f : \Omega \rightarrow \mathbb{R}$.

1- On dit que f est de classe C^m sur Ω ($f \in C^m(\Omega)$) si toutes les dérivées partielles jusqu'à l'ordre m existent et sont continues.

2- Pour tout $x \in \Omega$ et tout $i \in \{1, 2, \dots, n\}$ on note (quand \exists)

$$\frac{\partial f}{\partial x_i}(x) = \lim_{t \rightarrow 0} \frac{1}{t} [f(x + te_i) - f(x)].$$

(c'est la dérivée partielle de f en x de direction x_i).

3- Pour tout $x \in \Omega$ on note (quand \exists)

$$J_f(x) = \left(\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \dots, \frac{\partial f}{\partial x_n} \right) \in \mathbb{R}^n, \forall x \in \Omega$$

(la Jacobienne de f en x). On a

$$\nabla f = (J_f)^T$$

4- Pour tout $x \in \Omega$ et $h \in \mathbb{R}^n$ on note (quand \exists)

$$\frac{\partial f}{\partial h}(x) = \lim_{t \rightarrow 0} \frac{1}{t} [f(x + th) - f(x)] = g'(0).$$

(c'est la dérivée directionnelle de f en x de direction h) ou on a noté $\frac{\partial f}{\partial h}(x)$.

i) $\frac{\partial f}{\partial 0}(x) = 0$.

ii) $\frac{\partial f}{\partial x_i}(x) = \frac{\partial f}{\partial e_i}(x)$.

Nous rappelons aussi la formule :

$$\frac{\partial f}{\partial h}(x) = \langle \nabla f(x), h \rangle, \quad \forall x \in \Omega \quad \forall h \in \mathbb{R}^n.$$

1.2.2 Rappel sur le gradient

f admettant en un point x des dérivées partielles du premier ordre. On posera $x = (x_1, x_2, \dots, x_n)^t$ les éléments de Ω sont assimilés à des vecteurs colonnes on note et on appelle gradient de f au point x le vecteur colonne :

$$\nabla f(x) = \left(\frac{\partial f}{\partial x_1}(x), \dots, \frac{\partial f}{\partial x_n}(x) \right)^t.$$

Proposition 1.2.1 (*Gradient de la composée*)

Soient les deux fonctions $f : \Omega \rightarrow \mathbb{R}$ et $g : U \rightarrow \mathbb{R}$ et supposons qu'on deux ouverts $\Omega \subset \mathbb{R}^n$ et $U \subset \mathbb{R}$ avec en plus $f(\Omega) \subset U$ (on peut alors définir $g \circ f : \Omega \rightarrow \mathbb{R}$). Supposons que f, g sont de classe C^1 . Alors $g \circ f$ est aussi de classe C^1 avec en plus

$$\nabla (g \circ f)(x) = g'(f(x)) \nabla f(x) \quad \forall x \in \Omega.$$

1.2.3 Rappel sur le hessienne

Définition 1.2.1 *Si les dérivées de f possèdent à leur tour des dérivées partielles, on dit que f admet des dérivées partielles d'ordre 2, on pose :*

$$\nabla^2 f(\hat{x}) = \nabla (\nabla f(\hat{x})^t),$$

c'est à dire :

$$\nabla^2 f(\hat{x}) = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2}(\hat{x}) & \cdots & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_j}(\hat{x}) & \cdots & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n}(\hat{x}) \\ \frac{\partial^2 f}{\partial x_2 \partial x_1}(\hat{x}) & \cdots & \cdots & \frac{\partial^2 f}{\partial x_2 \partial x_j}(\hat{x}) & \cdots & \cdots & \frac{\partial^2 f}{\partial x_2 \partial x_n}(\hat{x}) \\ \vdots & \ddots & \cdots & \vdots & \ddots & \cdots & \vdots \\ \frac{\partial^2 f}{\partial x_i \partial x_1}(\hat{x}) & \cdots & \ddots & \frac{\partial^2 f}{\partial x_i \partial x_j}(\hat{x}) & \ddots & \cdots & \frac{\partial^2 f}{\partial x_i \partial x_n}(\hat{x}) \\ \vdots & \cdots & \ddots & \vdots & \ddots & \cdots & \vdots \\ \vdots & \ddots & \ddots & \vdots & \ddots & \cdots & \vdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1}(\hat{x}) & \cdots & \cdots & \frac{\partial^2 f}{\partial x_n \partial x_j}(\hat{x}) & \cdots & \cdots & \frac{\partial^2 f}{\partial x_n^2}(\hat{x}) \end{bmatrix}.$$

$\nabla^2 f$ s'appelle le hessien de f .

Remarque 1.2.1 *Si f est une fonction de classe C^2 (admet des dérivées partielles d'ordre 2 continues), le hessien de f est une matrice symétrique $H(\hat{x})$ d'ordre $n \times n$.*

Proposition 1.2.2 *soit $H(\hat{x})$ la matrice hessienne de la fonction f , si :*

a) La matrice $H(\hat{x})$ est dite semi-définie positive (*SDP*) ssi :

$$\forall x \in \mathbb{R}^n, x^t H(\hat{x}) x \geq 0.$$

où si les vecteurs propres de la matrice $H(\hat{x})$ sont positives.

b) La matrice $H(\hat{x})$ est dite définie positive (*DP*) ssi :

$$\forall x \in \mathbb{R}^n / \{0_{\mathbb{R}^n}\}, x^t H(\hat{x}) x > 0.$$

c-à-d les vecteurs propres de la matrice $H(\hat{x})$ sont strictement positives.

Quelques exemples importants :

1- Si $f : \mathbb{R}^n \rightarrow \mathbb{R}$ est une fonction constante alors

$$\nabla f = \nabla^2 f = 0$$

2- Soit $f : \mathbb{R}^n \rightarrow \mathbb{R}$ définie par

$$f(x) = \langle a, x \rangle \quad \forall x \in \mathbb{R}^n,$$

où $a \in \mathbb{R}^n$ est un vecteur donné (c'est à dire, f est une fonction linéaire). Alors on calcule facilement : $\frac{\partial f}{\partial x_k} = a_k$, donc

$$\nabla f = a$$

(le gradient est constant).

Ceci nous donne

$$\nabla^2 f = 0.$$

3- Soit $f : \mathbb{R}^n \rightarrow \mathbb{R}$ donnée par

$$f(x) = \langle Ax, x \rangle \quad \forall x \in \mathbb{R}^n,$$

où $A \in M_n(\mathbb{R})$ est une matrice carrée, réelle, de taille n (c'est à dire, f est une fonction quadratique associée à la matrice A). Alors on détermine le gradient :

$$\nabla f(x) = (A + A^T)x, \quad \forall x \in \mathbb{R}^n.$$

et la matrice hessienne :

$$\nabla^2 f = A + A^T, \quad \forall x \in \mathbb{R}^n.$$

Remarque 1.2.2 En particulier, si A est symétrique (c'est à dire $A = A^T$). Alors

$$\nabla \langle Ax, x \rangle = 2Ax, \quad \forall x \in \mathbb{R}^n.$$

$$\nabla^2 \langle Ax, x \rangle = 2A, \quad \forall x \in \mathbb{R}^n.$$

1.3 Eléments d'analyse convexe

La notion de convexité est un outil mathématique important pour l'étude théorique et numérique des problèmes d'optimisation. À ce propos, nous présentons dans ce paragraphe quelques notions de base d'usage courant.

1.3.1 Ensembles et fonctions convexes

- Un ensemble C de \mathbb{R}^n est dit convexe si

$$\lambda x + (1 - \lambda)y \in C, \forall x, y \in C, \forall \lambda \in [0, 1].$$

- C est dit affine si

$$\lambda x + (1 - \lambda)y \in C, \forall x, y \in C, \forall \lambda \in \mathbb{R}.$$

- C est un polyèdre convexe s'il est de la forme :

$$C = \{x \in \mathbb{R}^n : A_i^t x \leq b_i, i = 1, \dots, m\}$$

où A_i est un vecteur non nul de \mathbb{R}^n et b_i un scalaire pour $i = 1, \dots, m$.

C peut s'écrire sous la forme matricielle suivante :

$$C = \{x \in \mathbb{R}^n / Ax \leq b\},$$

où A est une matrice de $\mathbb{R}^{m \times n}$ et b un vecteur de \mathbb{R}^m .

Soit $f : C \rightarrow \mathbb{R}$ une fonction et C un ensemble convexe de \mathbb{R}^n .

• f est dite mid-convexe sur C si :

$$\forall x, y \in C, f\left(\frac{x+y}{2}\right) \leq \frac{f(x) + f(y)}{2}.$$

• f est dite quasi-convexe sur C si :

$$f(\lambda x + (1 - \lambda)y) \leq \max(f(x), f(y)), \forall \lambda \in [0, 1], \forall x, y \in C.$$

• f est dite convexe sur C si l'inégalité suivante est satisfaite :

$$f(\lambda x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda)f(y), \forall \lambda \in [0, 1], \forall x, y \in C.$$

• f est dite strictement convexe sur C si :

$$f(\lambda x + (1 - \lambda)y) < \lambda f(x) + (1 - \lambda)f(y), \forall \lambda \in]0, 1[, \forall x, y \in C \text{ et } x \neq y.$$

• f est dite fortement convexe sur C s'il existe $\alpha > 0$, tel que :

$$\forall \lambda \in]0, 1[, \forall x, y \in C \text{ et } x \neq y,$$

on a :

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

(on dit aussi que f est α -convexe).

- f est convexe sur C si seulement si f est mid-convexe et quasi-convexe sur C .
- Si f est une fonction continue sur un convexe C , on a :
 - f est convexe sur C si seulement si f est mid-convexe C .
 - f est α -convexe sur C si seulement si :

$$\forall x, y \in C, f\left(\frac{x + y}{2}\right) \leq \frac{f(x) + f(y)}{2} - \frac{\alpha}{8}\|x - y\|^2.$$

1.3.2 Formules de Taylor

Soit $\Omega \subset \mathbb{R}^n$ ouvert, $f : \Omega \rightarrow \mathbb{R}$, $a \in \Omega$ et $h \in \mathbb{R}^n$ tels que $[a, a + h] \subset \Omega$.
Alors :

1- Si $f \in C^1(\Omega)$ alors

i) Formule de Taylor à l'ordre 1 avec reste intégral

$$f(a + h) = f(a) + \int_0^1 \langle \nabla f(a + th), h \rangle dt.$$

ii) Formule de Taylor - Maclaurin à l'ordre 1

$$\text{il existe } \theta \in [0, 1] \text{ tel que } f(a + h) = f(a) + \langle \nabla f(a + \theta h), h \rangle.$$

iii) Formule de Taylor - Young à l'ordre 1

$$f(a + h) = f(a) + \langle \nabla f(a), h \rangle + o(\|h\|).$$

2- Si $f \in C^2(\Omega)$ alors

i) Formule de Taylor à l'ordre 2 avec reste intégral

$$f(a + h) = f(a) + \langle \nabla f(a), h \rangle + \int_0^1 (1 - t) \langle \nabla^2 f(a + th) h, h \rangle dt.$$

ii) Formule de Taylor - Maclaurin à l'ordre 2

$$\text{il existe } \theta \in [0, 1] \text{ telque } f(a + h) = f(a) + \langle \nabla f(a), h \rangle + \frac{1}{2} \langle \nabla^2 f(a + \theta h) h, h \rangle.$$

iii) Formule de Taylor - Young à l'ordre 2

$$f(a + h) = f(a) + \langle \nabla f(a), h \rangle + \frac{1}{2} \langle \nabla^2 f(a) h, h \rangle + o(\|h\|^2).$$

Remarque 1.3.1 Dans les formules précédentes, la notation $o(\|h\|^k)$ pour $k \in \mathbb{N}^*$ signifie une expression qui tend vers 0 plus vite que $\|h\|^k$ (c'est à dire, si on la divise par $\|h\|^k$, le résultat tend vers 0 quand h tend vers 0).

1.4 Problème général d'optimisation

L'optimisation a un vocabulaire particulier, pour cela nous allons introduire quelques notations et définitions classiques. Tout d'abord, nous donnons la formulation générale d'un problème d'optimisation, donc on a besoin :

- Une fonction objective ou fonction de coût ou critère à minimiser, noté $f : \mathbb{R}^n \rightarrow \mathbb{R}$ de plusieurs variables ($x = (x_1, x_2, \dots, x_n)$).

- Un ensemble $A \subset \mathbb{R}^n$ où l'on cherche la solution, on dit que A est l'ensemble des éléments admissibles du problème ou l'ensemble des contraintes.

On cherche à minimiser f sur A c'est à dire on cherche $\hat{x} \in A$ tel que :

$$f(\hat{x}) = \min_{x \in A} f(x) \iff f(\hat{x}) \leq f(x), \forall x \in A.$$

On a deux types d'optimisation : sans et avec contraintes, dans les deux cas, le but consiste à trouver les valeurs qui maximisent une fonction. Toutes dans l'optimisation avec contraintes, les solutions sont soumises à des restrictions.

1.5 Conditions nécessaires de minimum

Soit une fonction $f : \mathbb{R}^n \rightarrow \mathbb{R}$. Les minima locaux et globaux de f sur \mathbb{R}^n sont définis de la manière suivante :

Définition 1.5.1 :

1-On dit que la fonction f de problème (P) possède un *minimum globale* (ou *absolu*) en $\hat{x} \in \mathbb{R}^n$ ssi :

$$\forall x \in \mathbb{R}^n, f(\hat{x}) \leq f(x).$$

2-On dit que le point $\hat{x} \in \mathbb{R}^n$ est un *optimum locale* (ou *relatif*) du (P) s'il existe un voisinage V de \hat{x} tel que :

$$\forall x \in V(\hat{x}), f(x) \geq f(\hat{x}).$$

3-On dit que $\hat{x} \in \mathbb{R}^n$ est un *optimum local strict* de (P) ssi :

$$\exists V(\hat{x}) \text{ tel que } : \forall x \in V(\hat{x}), f(x) > f(\hat{x}).$$

4-On dit que $\hat{x} \in \mathbb{R}^n$ est un *optimum local strict isolé* du (P) ssi :

$$\exists V(\hat{x}) \text{ tel que } : \hat{x} \text{ est la seule solution optimale de } (P)$$

Remarque 1.5.1 : - Un minimum global est clairement un minimum local.

- Si on dit simplement minimum on comprend minimum global.

- Pour la maximisation, faire la minimisation de la fonction $(-f)$:

$$\max f(x) = - \min f(-x)$$

Soit une fonction $f : \mathbb{R}^n \rightarrow \mathbb{R}$, et $\hat{x} \in \mathbb{R}^n$

1- On dit que \hat{x} est un point de maximum absolu (respectivement relatif) de f sur \mathbb{R}^n si \hat{x} est un point de minimum absolu (respectivement relatif) de $-f$ sur \mathbb{R}^n .

2- On dit que \hat{x} est un point d'extremum absolu (respectivement relatif) de f sur \mathbb{R}^n si \hat{x} : soit un point de minimum absolu (respectivement relatif) de $-f$ sur \mathbb{R}^n , soit un point de maximum absolu (respectivement relatif) de f sur U .

Lemme 1.5.1 Soit $U \in \mathbb{R}^n$, $a \in \mathbb{R}^n$ et u^* un élément appartenant à l'intérieur de U ($u^* \in \overset{\circ}{U}$). Alors les deux assertions suivantes sont équivalentes :

- 1- $\langle a, u - u^* \rangle \geq 0, \quad \forall u \in U.$
- 2- $a = 0.$

Définition 1.5.2 Soit $U \subset \mathbb{R}^n$ un ensemble et $u^* \in U$. on dit que $w \in \mathbb{R}^n$ est une direction admissible pour u^* en U s'il existe $t_0 > 0$ tel que $u^* + tw \in U \forall t \in [0, t_0]$.

Exemples :

1. Si $u^* \in \overset{\circ}{U}$ alors tout vecteur $w \in \mathbb{R}^n$ est une direction admissible pour u^* en U .
2. Si U est convexe alors pour tout $v \in U$ le vecteur $v - u^*$ est une direction admissible pour u^* en U .

Lemme 1.5.2 Soit $\Omega \subset \mathbb{R}^n$ un ouvert, $U \subset \Omega$, $f : \Omega \rightarrow \mathbb{R}$ une fonction de classe C^1 et $u^* \in U$ un point de minimum relatif de f sur U . Soit $w \in \mathbb{R}^n$ une direction admissible pour u^* en U . Alors

$$\langle \nabla f(u^*), w \rangle \geq 0.$$

Lemme 1.5.3 Soit $\Omega \subset \mathbb{R}^n$ un ouvert, $U \subset \Omega$ un ensemble convexe et $f : \Omega \rightarrow \mathbb{R}$ une fonction de classe C^1 , $u^* \in U$ un point de minimum relatif de f sur U . Alors

1. $\langle \nabla f(u^*), u - u^* \rangle \geq 0, \quad \forall u \in U.$
(c'est l'inéquation d'Euler)

2. Si en plus u^* est dans l'intérieur de U ($u^* \in \overset{\circ}{U}$) alors la condition

$$\langle \nabla f(u^*), u - u^* \rangle \geq 0, \forall u \in U$$

est équivalente au

$$\nabla f(u^*) = 0 \quad (\text{c'est l'équation d'Euler}).$$

Théorème 1.5.1 Soit $\Omega \subset \mathbb{R}^n$ un ouvert, $U \subset \Omega$ un ensemble convexe, $f : \Omega \rightarrow \mathbb{R}$ une fonction de classe C^1 et convexe et $u^* \in U$. Alors les trois assertions suivantes sont équivalentes :

- 1- u^* est un point de minimum global de f sur U .
- 2- u^* est un point de minimum local de f sur U .
- 3- $\langle \nabla f(u^*), u - u^* \rangle \geq 0, \quad \forall u \in U$.

Remarque 1.5.2 Dans le cas où $u^* \in \overset{\circ}{U}$ alors l'assertion 3 du (théorème 1.5.1) peut être remplacée grâce au (Lemme 1.5.3) par l'équation d'Euler :

$$\nabla f(u^*) = 0.$$

Chapitre 2

Programmation mathématique

La programmation mathématique constitue un domaine vaste et riche dans l'analyse numérique. Elle traite plusieurs modèles mathématiques et problèmes pratiques importants.

2.1 Problème d'optimisation sans contrainte

Nous allons étudier le problème d'optimisation sans contraintes où on effectue la minimisation de la fonction $f : \mathbb{R}^n \rightarrow \mathbb{R}$ sur tout l'espace \mathbb{R}^n , Nous considérons donc le problème formulé de la façon suivante :

$$(P) \{ \min f(x); x \in \mathbb{R}^n \}.$$

qui peut s'écrire sous la forme

$$\left\{ \text{trouver } \hat{x} \in \mathbb{R}^n \text{ tel que } f(\hat{x}) \leq f(x); \quad \forall x \in \mathbb{R}^n \right\}.$$

2.1.1 L'existence et d'unicité

Nous allons maintenant nous intéresser à la question de l'existence de minima pour des problèmes d'optimisation sans contraintes.

Théorème 2.1.1 (Existence) Soit $f : \mathbb{R}^n \rightarrow \mathbb{R}$ une application telle que :

(i) f est continue.

(ii) f est coercive (c'est-à-dire $\lim_{\|x\| \rightarrow +\infty} f(x) = +\infty$).

Alors, il existe $\hat{x} \in \mathbb{R}^n$ tel que $f(\hat{x}) \leq f(y)$ pour tout $y \in \mathbb{R}^n$.

Preuve. Soit (x_n) une suite minimisante de \mathbb{R}^n c-à-d : $\lim_{n \rightarrow \infty} f(x_n) = \inf f(x)$.

la fonction f est coercive donc la suite (x_n) est bornée, on peut extraire une sous suite notée (x_n) qui converge vers un point $\hat{x} \in \mathbb{R}^n$, et par la continuité de f on a :

$$\lim_{n \rightarrow \infty} f(x_n) = f(\hat{x}),$$

d'où

$$f(\hat{x}) = \inf_{x \in \mathbb{R}^n} f(x),$$

donc \hat{x} est une solution du (P). ■

Toutefois, on n'a pas forcément l'unicité. Nous donnons ci-dessus un critère pour l'unicité.

Théorème 2.1.2 : (Unicité)

Soit $f : \mathbb{R}^n \rightarrow \mathbb{R}$ strictement convexe alors il existe au plus un $\hat{x} \in \mathbb{R}^n$ tel que :

$$f(\hat{x}) \leq f(y), \quad \forall y \in \mathbb{R}^n.$$

Preuve. Supposons \hat{x}_1 et \hat{x}_2 deux solutions distinctes du problème (P) c-à-d :

$$f(\hat{x}_1) = f(\hat{x}_2) = \min_{x \in \mathbb{R}^n} f(x).$$

On a $x^* = \frac{(\hat{x}_1 + \hat{x}_2)}{2} \in \mathbb{R}^n$, et par la stricte convexité de la fonction, on obtient :

$$f(x^*) = f\left(\frac{\hat{x}_1 + \hat{x}_2}{2}\right) < \frac{1}{2}(f(\hat{x}_1) + f(\hat{x}_2)) = \min_{x \in \mathbb{R}^n} f(x).$$

cela est contredit, donc : $\hat{x}_1 = \hat{x}_2$. ■

Nous pouvons maintenant énoncer un deuxième résultat d'existence et d'unicité dans le cas particulier où f est α -elliptique.

Théorème 2.1.3 : Soit une fonction f de classe C^1 , on suppose qu'il existe $\alpha > 0$ (appelée la constante d'ellipticité), tel que $\forall (x, y) \in \mathbb{R}^n \times \mathbb{R}^n$, on a :

$$(\nabla f(x) - \nabla f(y), x - y) \geq \alpha \|x - y\|^2 \quad (\alpha - \text{elliptique}).$$

alors f est strictement convexe et coersive.

On va chercher à montrer que cette solution est une solution de certaines équations, de sorte qu'il sera plus facile de calculer.

2.1.2 Condition d'optimalité

Dans cette section, nous allons chercher à obtenir des conditions nécessaires et parfois suffisantes de minimalité puisque ces conditions d'optimalité seront le plus souvent utilisées pour tenter de calculer un minimum. Ces conditions vont donc s'exprimer à l'aide de la dérivée première ou seconde.

Les deux conditions nécessaires d'optimalité sont les suivants :

Condition nécessaires du premier ordre

Etant donné un vecteur $\hat{x} \in \mathbb{R}^n$, nous souhaiterions être capables de déterminer si ce vecteur est un minimum local ou global de la fonction f . La propriété de la différentiabilité continue de f fournit une première manière de caractériser une solution optimale.

Théorème 2.1.4 : Soit $f : \mathbb{R}^n \rightarrow \mathbb{R}$ une fonction différentiel au point $\hat{x} \in \mathbb{R}^n$. Si \hat{x} est un optimum locale du (P) alors :

$$\nabla f(\hat{x}) = 0 \quad (\text{a})$$

Remarque 2.1.1 :

1- Un point \hat{x} de \mathbb{R}^n vérifiant $\nabla f(\hat{x}) = 0$ est appelé point critique ou point stationnaire.

2- Le théorème précédente n'a pas de sens si la fonction f n'est pas différentiable.

3- La condition (a) est une condition du premier ordre car elle ne fait intervenir que la dérivée première de la fonction f .

Il existe des situations où la relation (a) est une condition nécessaire et suffisante :

Théorème 2.1.5 : Soit $f : \mathbb{R}^n \rightarrow \mathbb{R}$ une fonction de classe C^1 et convexe. Un point \hat{x} réalise un minimum def sur \mathbb{R}^n ssi $\nabla f(\hat{x}) = 0$.

Preuve. On a vu que la condition est toujours nécessaire, montrons qu'elle est suffisante. Soit $\hat{x} \in \mathbb{R}^n$ tel que $\nabla f(\hat{x}) = 0$. Comme f est convexe, donc

$$\forall x \in \mathbb{R}^n, f(x) \geq f(\hat{x}) + (\nabla f(\hat{x}), x - \hat{x}) = f(\hat{x}).$$

On a donc immédiatement le fait que \hat{x} réalise un minimum de f . ■

Nous donnons maintenant une condition nécessaire permettant de préciser encore les éventuels minima. Ce condition va intervenir la seconde dérivée de f .

Condition nécessaires du seconde ordre

Théorème 2.1.6 : Soit $f : \mathbb{R}^n \rightarrow \mathbb{R}$ de classe C^2 . Si f possède un min local en \hat{x} alors :

1- $\nabla f(\hat{x}) = 0$, et

2- La matrice Hessien $H(\hat{x})$ est une matrice semi définie positive.

Condition suffisante du seconde ordre

Les conditions données précédemment sont nécessaires, c'est-à-dire qu'elle doivent être satisfaites pour tout minimum local, cependant, tout vecteur vérifiant ces conditions n'est pas nécessairement un minimum local. Le théorème suivant établit une condition suffisante pour qu'un vecteur soit un minimum local, si f est deux fois continûment différentiable.

Théorème 2.1.7 : Soit $f : \mathbb{R}^n \rightarrow \mathbb{R}$ de classe C^2 , si :

- 1- $\nabla f(\hat{x}) = 0$. et
- 2- $H(\hat{x})$ est une matrice définie positive, alors f possède un minimum local en \hat{x} .

Preuve. La fonction f est deux fois différentiable en \hat{x} , alors :

$$\forall x \in \mathbb{R}^n, f(x) = f(\hat{x}) + (\nabla f(\hat{x}))^t (x - \hat{x}) + \frac{1}{2} (x_k - \hat{x})^t H(\hat{x}) (x - \hat{x}) + \|x - \hat{x}\|^2 \alpha(\hat{x}, x - \hat{x}).$$

Supposons que \hat{x} n'est pas un min local strict :

$$\forall V(\hat{x}), \exists \bar{x} \in V(\hat{x}), f(\bar{x}) < f(\hat{x}).$$

Notons

$$d_k = \frac{x_k - \hat{x}}{\|x_k - \hat{x}\|}, f(x_k) \xrightarrow[k \rightarrow \infty]{} f(\hat{x}) \text{ et } f(x_k) \leq f(\hat{x}) \quad \forall k \in \mathbb{N}.$$

donc $\|d_k\| = 1$, d'après le théorème de **Bolsano weistrass** il' existe sous suite converge $\{d_k\}$

$$d_k \xrightarrow[k \rightarrow \infty]{} \bar{d} \text{ et } \|\bar{d}\| = 1.$$

$$f(x_k) - f(\hat{x}) = \nabla f(\hat{x})^t (x_k - \hat{x}) + \frac{1}{2} (x_k - \hat{x})^t H(\hat{x}) (x_k - \hat{x}) + \|x_k - \hat{x}\|^2 \alpha(\hat{x}, x_k - \hat{x}).$$

donc

$$\frac{f(x_k) - f(\hat{x})}{\|x_k - \hat{x}\|} = \frac{1}{2} d_k^t H(\hat{x}) d_k + \alpha(\hat{x}, x_k - \hat{x}).$$

or

$$f(x_k) < f(\hat{x}) \implies f(x_k) - f(\hat{x}) \leq 0.$$

donc

$$\forall k, \frac{1}{2} d_k^t H(\hat{x}) d_k + \alpha(\hat{x}, x_k - \hat{x}) \leq 0.$$

passant à la limite

$$k \longrightarrow +\infty \implies \frac{1}{2} \bar{d}^t H(\hat{x}) \bar{d} \leq 0, \bar{d} \neq 0.$$

car $d_k \rightarrow \bar{d}$ et $\|\bar{d}\| = 1$, donc $H(\hat{x})$ n'est pas définie positive. ■

2.2 problème d'optimisation avec contrainte

On définit un problème d'optimisation avec contraintes comme suit :

$$(P) \quad \begin{cases} \min f(x) \\ x \in C \end{cases}$$

$f : \mathbb{R}^n \rightarrow \mathbb{R}$ continue, $D \subseteq \mathbb{R}^n$ l'ensemble des contraintes
Si ($C = \mathbb{R}^n$), (P) est appelé problème d'optimisation sans contraintes.

2.3 programme mathématique

En général, un programme mathématique est défini comme suit :

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

et f, g_i, h_j sont des fonctions données de \mathbb{R}^n vers \mathbb{R} .

On appelle f la fonction objectif et C l'ensemble des solutions réalisables ou ensemble des contraintes ou tout simplement "le domaine".

On appelle solution réalisable de (PM) tout point x^0 vérifiant les contraintes (i.e., ($x^0 \in D$)).

2.3.1 Classification et résolution d'un programme mathématique

La classification de (PM) et son traitement numérique sont établis à partir des propriétés fondamentales des fonctions f, g_i, h_j à savoir la convexité, la différentiabilité et la linéarité.

parmi les cas particuliers les plus étudiés on note :

- La programmation linéaire (f linéaire, g_i, h_j affines).
- La programmation convexe (f, g_i convexes, h_j affines, C convexe).
- La programmation en nombres entiers (C est un ensemble discret, c'est à dire les variables sont entières).

2.3.2 Existence et Unicité de solution

Dans ce paragraphe, nous donnons les deux théorèmes d'existence et le théorème d'unicité

Théorème 2.3.1 *Si C est compact non vide de \mathbb{R}^n et si f est continue sur C alors (PM) admet au moins une solution optimale $x^* \in C$.*

Théorème 2.3.2 Si C est fermé non vide de \mathbb{R}^n , f est continue et coercive sur C (c'est-à-dire $\lim_{\|x\| \rightarrow +\infty} f(x) = +\infty$) alors (PM) admet au moins une solution optimale.

Théorème 2.3.3 Si C est convexe non vide de \mathbb{R}^n , f est strictement convexe sur C alors (PM) admet une solution optimale au plus.

2.3.3 Conditions d'optimalité

Avant de donner les conditions d'optimalité de (PM), on exige que les contraintes doivent satisfaire certains critères dits "critères de qualification".

Une contrainte g_i est dite active (ou saturée) en $\bar{x} \in C$ si $g_i(\bar{x}) = 0$.

• Un point $\bar{x} \in C$ est dit régulier (on dit également que les contraintes sont qualifiées en \bar{x}) si les composantes de gradient, correspondant aux contraintes saturées en \bar{x} , sont linéairement indépendantes.

Il existe aussi deux critères usuels de qualification en tout point de C , à savoir :

- Si toutes les contraintes sont affines.
- Si C est défini uniquement par des inégalités, on a le critère de Slater suivant : $g_i(x)$ est convexe pour tout $i = 1, \dots, m$ et qu'il existe un point x^0 tel que $g_i(x^0) < 0$, ($\text{int}(C) \neq \emptyset$).

2.3.4 Dualité Lagrangienne

Soit :

$$S = \{x \in D \subset \mathbb{R}^n : g_i(x) \leq 0, \quad i = 1, \dots, k, \quad h_j(x) = 0, \quad j = 1, \dots, m\}$$

et on considère le problème primal :

$$m = \inf_x [f(x), \quad x \in S]$$

Le Lagrangien associé à ce problème est la fonction $L : D \times [0, +\infty[^k \times \mathbb{R}^m \rightarrow \mathbb{R}$, définie par :

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

On pose :

$$\alpha(x) = \sup_{\lambda, \mu} [L(x, \lambda, \mu) : \lambda \geq 0] = \begin{cases} f(x) & \text{si } g_i(x) \leq 0 \text{ et } h_j(x) = 0 \\ +\infty & \text{sinon} \end{cases}$$

donc :

$$\bar{\alpha} = \inf_{x \in D} \alpha(x) = \inf_x [f(x) : x \in S]$$

Le problème dual associé au problème primal est :

$$\bar{\beta} = \sup_{(\lambda, \mu)} \inf_{x \in D} [L(x, \lambda, \mu)]$$

On a l'inégalité de dualité

$$-\infty \leq \bar{\beta} \leq \bar{\alpha}$$

Théorème 2.3.4 (Karush - Kuhn - Tucker)[13] :

Soit $\bar{x} \in C$ satisfaisant l'une des conditions de qualification et supposons que f, g_i, h_j sont $C^1(\mathbb{R}^n)$, on a :

Si \bar{x} est un optimum local pour (PM), alors il existe des réels dits multiplicateurs de Lagrange :

$\mu_i \in \mathbb{R}^+, i = 1, \dots, m$ et $\lambda_j \in \mathbb{R}, j = 1, \dots, p$ tels que :

$$\begin{cases} \nabla f(\bar{x}) + \sum_{i=1}^m \mu_i \nabla g_i(\bar{x}) + \sum_{j=1}^p \lambda_j \nabla h_j(\bar{x}) = 0 & (\text{conditions d'optimalité}) \\ \mu_i g_i(\bar{x}) = 0, \quad i = 1, \dots, m. & (\text{conditions de complémentarité}) \\ h_j(\bar{x}) = 0, \quad j = 1, \dots, p. & \end{cases}$$

Si de plus, f, g_i, h_j sont convexes, les conditions précédentes sont à la fois nécessaires et suffisantes pour que \bar{x} soit un optimum global pour (PM).

2.4 Algorithme d'optimisation

Nous allons présenter un algorithme permettant de converger vers une solution optimale du problème (PM). La plupart des algorithmes d'optimisation avec contraintes exploitent les conditions d'optimalité pour déterminer des minima locaux. Nous donnerons ici quelques définitions.

2.4.1 Description

Un algorithme est défini par une application A , de C dans C , où C est l'ensemble des solutions réalisables, permettant la génération d'une suite d'éléments de C par la formule :

$$\begin{cases} x_0 \in C \text{ donné}, k = 0 & \text{Etape d'initialisation} \\ x_{k+1} = A(x_k), k = k + 1 & \text{Itération} \end{cases} ;$$

Si on remplace C par son intérieur, en supposant que $\text{int}(C) \neq \emptyset$, l'algorithme est dit un algorithme de points intérieurs.

Définir un algorithme n'est autre que construire une suite $(x_k)_{k \in \mathbb{N}}$ de C et réaliser une étude pour montrer sa convergence.

2.4.2 Convergence

Définition 1 :

On dit que l'algorithme A est convergent si la suite $(x_k)_{k \in \mathbb{N}}$ engendrée par l'algorithme converge vers une limite x^* .

2.4.3 Taux de convergence

Un critère de mesure de la vitesse (ou le taux) de convergence est l'évolution de l'erreur commise à chaque itération ($e_k = \|x_k - x^*\|$).

Avant de donner les notations des convergences, nous donnons les définitions des notations asymptotiques suivantes :

Définition 2 (Notation \mathbf{O}) Soient deux fonctions $f, g : \mathbb{N} \rightarrow \mathbb{R}^+$. On note $f(n) = \mathbf{O}(g(n))$ lorsqu'il existe des entiers c et n_0 tels que pour tout $n \geq n_0$,

$$f(n) \leq cg(n)$$

Intuitivement, cela signifie que la valeur de la fonction f est inférieure à celle de g à une constante multiplicative près, pour les instances (données) de tailles suffisamment grandes. De même on définit :

Définition 3 (Notations \mathbf{o} , $\mathbf{\Omega}$, $\mathbf{\Theta}$) Soient deux fonctions $f, g : \mathbb{N} \rightarrow \mathbb{R}^+$

- On note $f(n) = \mathbf{o}(g(n))$ lorsque pour tout réel c , il existe un entier n_0 tel que pour tout $n \geq n_0$,

$$f(n) \leq cg(n)$$

- On note $f(n) = \mathbf{\Omega}(g(n))$ lorsqu'il existe des entiers c et n_0 tels que pour tout $n \geq n_0$,

$$cg(n) \leq f(n)$$

- On note $f(n) = \mathbf{\Theta}(g(n))$ lorsque $f(n) = \mathbf{O}(g(n))$ et $f(n) = \mathbf{\Omega}(g(n))$.

Soit $(x_k)_{k \in \mathbb{N}}$ une suite donnée par l'algorithme A et convergente vers x^* .

La classification de la vitesse de convergence d'une suite est basée sur les notions de comparaison des fonctions au voisinage de $+\infty$.

En effet, si on suppose que l'erreur e_k ne s'annule pas, la vitesse de la convergence pourra être :

- **Linéaire** : Si $\|e_k\| = \mathbf{\Omega}(\|e_{k+1}\|)$ et $\left(\frac{\|e_{k+1}\|}{\|e_k\|}\right) < 1$, pour k assez grand. On dit aussi que l'erreur e_k décroît linéairement c'est-à-dire :

$$\exists c \in [0, 1[, \exists k_0 \in \mathbb{N}, \forall k \geq k_0, e_{k+1} \leq ce_k.$$

- **Superlinéaire** : Si $\|e_{k+1}\| = \mathbf{o}(\|e_k\|)$, où l'erreur décroît de la manière suivante :

$$\exists \alpha_k \text{ une suite positive qui converge vers } 0 \text{ tel que } e_{k+1} \leq \alpha_k e_k.$$

-D'ordre γ avec $\gamma > 1$: Si $\|e_{k+1}\| = O\left(\|e_k\|^\gamma\right)$ et $\left(\frac{\|e_{k+1}\|}{\|e_k\|^\gamma}\right) < 1$, pour k assez grand, où l'erreur décroît de la manière suivante :

$$\exists c \in [0, 1[, \exists k_0 \in \mathbb{N}, \forall k \geq k_0, e_{k+1} \leq c(e_k)^\gamma.$$

Dans le cas $\gamma = 2$, la convergence est dite quadratique.

2.5 Programmation linéaire (PL)

Un programme linéaire (PL) est un problème d'optimisation qui consiste à maximiser (ou minimiser) une fonction objectif linéaire de n variables de décision sujet à un ensemble

de contraintes exprimées sous forme d'équations ou d'inéquations linéaires.

a) **Forme général**

$$(PL) \begin{cases} \min c^t x \\ s.c \\ Ax = b \\ Dx \geq e \\ x \in \mathbb{R}^n. \end{cases}$$

où $A \in \mathbb{R}^{m \times n}$ et $D \in \mathbb{R}^{p \times n}$ sont des matrices données, $c \in \mathbb{R}^n, b \in \mathbb{R}^m$ et $e \in \mathbb{R}^p$ sont des vecteurs donnés.

On peut montrer que tout programme linéaire peut se ramener à l'une des deux formes suivantes :

b) **Forme canonique :**

$$(PLC) \begin{cases} \min c^t x \\ s.c \\ Ax \geq b \text{ (ou } \leq) \\ x > 0. \end{cases}$$

c) **Forme standard :**

$$(PLS) \begin{cases} \min c^t x \\ s.c \\ Ax = b \\ x \geq 0. \end{cases}$$

où A est une matrice réelle de type (m, n) supposée de plein rang (c'est-à-dire : $\text{rg}(A) = m < n$), b un vecteur de \mathbb{R}^m .

Dans toute la suite, on s'intéresse au problème (PL) sous forme standard suivant :

$$(PL) \begin{cases} \min c^t x \\ s.c \\ Ax = b \\ x \geq 0. \end{cases}$$

Le dual du programme linéaire (PL) est un programme linéaire défini par :

$$(DL) \begin{cases} \max b^t y \\ s.c \\ A^t y + s = c \\ s \geq 0, s \in \mathbb{R}^n \\ y \in \mathbb{R}^m. \end{cases}$$

On note par :

- $F_{(PL)} = \{x \in \mathbb{R}^n : Ax = b, x \geq 0\}$, l'ensemble des solutions primales réalisables de (PL) .
- Un vecteur $x \in F_{(PL)}$ est appelé solution réalisable de (PL) .
- Un vecteur $x^* \in F_{(PL)}$ minimisant la fonction objectif de (PL) s'appelle solution optimale de (PL) .
- Un programme linéaire (PL) réalisable est borné si la fonction objectif est bornée sur $F_{(PL)}$.
- $\overset{\circ}{F}_{(PL)} = \{x \in \mathbb{R}^n : Ax = b, x > 0\}$, l'ensemble des solutions primales strictement réalisables de (PL) .
- $F_{(DL)} = \{y \in \mathbb{R}^m : A^t y + s = c, s \geq 0\}$, l'ensemble des solutions duales réalisables de (DL) .
- Un vecteur $y^* \in F_{(DL)}$ maximisant la fonction objectif de (DL) s'appelle solution optimale de (DL) .
- $\overset{\circ}{F}_{(DL)} = \{y \in \mathbb{R}^m : A^t y + s = c, s > 0\}$, l'ensemble des solutions duales strictement réalisables de (DL) .
- $\overset{\circ}{H} = \overset{\circ}{H}_{(PL)} \times \overset{\circ}{H}_{(DL)}$, l'ensemble des solutions primales-duales strictement réalisables de (PL) et (DL) .

Donnons quelques résultats fondamentaux de dualité en programmation linéaire :

- Si l'un des problèmes (PL) et (DL) admet une solution optimale, il en est de même pour l'autre, et leurs valeurs optimales correspondantes sont égales.
- Si l'un des problèmes a une valeur optimale non finie, l'autre n'a pas de solution optimale.

Théorème 2.5.1 (*Dualité faible*) Si x et (y, s) sont respectivement des solu-

tions réalisables pour (PL) et (DL) alors,

$$c^t x \geq b^t y.$$

Théorème 2.5.2 (*Dualité forte*) Si \bar{x} et (\bar{y}, \bar{s}) sont respectivement des solutions réalisables correspondant une valeur optimale finie pour (PL) et (DL) telles que

$$c^t \bar{x} = b^t \bar{y},$$

alors \bar{x} est une solution primale optimale de (PL) et \bar{y} est une solution duale optimale de (DL).

Remarque 2.5.1 On peut remarquer facilement que si \bar{x} et (\bar{y}, \bar{s}) sont respectivement des solutions réalisables de (PL) et (DL), alors on a la propriété suivante :

$$c^t \bar{x} = b^t \bar{y} \Leftrightarrow \bar{x} \bar{s} = 0 \Leftrightarrow \bar{x}^t \bar{s} = 0.$$

à la place de M_n on travaillera habituellement dans S_n qui est isomorphe à $\mathbb{R}^{t(n)}$, où $t(n) = C_{n+1}^2 = \frac{n(n+1)}{2}$.

Pour $A, B \in S_n$, on a

$$\langle A, B \rangle = \text{tr}(AB) \text{ et } \|A\| = \sqrt{\sum_{i=1}^n \lambda_i(A)^2}$$

Toutes les valeurs propres d'une matrice symétrique $A \in S_n$ sont réelles et en plus il existe une matrice orthogonale $P \in M_n$ qui diagonalise A , c'est-à-dire

$P^T A P = D_A$ où D_A est une matrice diagonale dont les éléments diagonaux sont les valeurs propres de A . Ainsi les valeurs propres sont les solutions du polynôme caractéristique $p_n(\lambda) = \det(A - \lambda I)$. On note par A_i la $i^{\text{ième}}$ ligne de A et par $A_{.,j}$ la $j^{\text{ième}}$ colonne de A . Pour nos besoins, il convient d'écrire les valeurs propres dans l'ordre croissant,

$$\lambda_{\min}(A) = \lambda_1(A) \leq \lambda_2(A) \leq \dots \leq \lambda_n(A) = \lambda_{\max}(A)$$

pour une matrice $A \in S_n$ avec $\text{rang}(A) = k$, la décomposition spectrale de A , $A = P D_A P^T$

est donnée par une matrice diagonale $D_A \in S_k$, dont les éléments diagonaux sont les valeurs propres non nulles de A sur sa diagonale principale, et une matrice $P \in M_{n,k}$ telle que $P^T P = I_k$

On étudie maintenant les matrices semi-définies positives, bien qu'il est possible de définir ce terme pour des matrices carrées quelconques, on va l'utiliser exclusivement pour des matrices symétriques.

3.1.2 Matrice symétrique semi-définies positives

Définition 3.1.1 $A \in S_n$ est semi-définie positive ($A \in S_n^+$ ou $A \geq 0$) si $x^T A x \geq 0 \forall x \in \mathbb{R}^n$

$A \in S_n$ est définie positive ($A \in S_n^{++}$ ou $A > 0$) si $x^T A x > 0 \forall x \in \mathbb{R}^n \setminus \{0\}$

On énonce certaines conséquences immédiates de ces définitions qui seront utilisées par la suite.

- Toute sous-matrice principale d'une matrice semi-définie (resp. définie) positive est aussi semi-définie (resp. définie) positive. En particulier, tous les éléments diagonaux d'une matrice semi-définie (resp. définie) positive doivent être positifs (resp. strictement positifs).

- Pour $A \in S_n^+$, il existe $i \in \{1, \dots, n\}$; $a_{ij} = \max |a_{ij}|$; $i, j \in \{1, \dots, n\}$, si $A \in S_n^+$ et $a_{ii} = 0$ pour certain $i \in \{1, \dots, n\}$ alors $a_{ij} = 0$ pour tout $j \in \{1, \dots, n\}$

Proposition 3.1.1 Soit $B \in M_n$ une matrice non singulière. Alors $A \in S_n^+$ si et seulement si $B^T A B \in S_n^+$ et $A \in S_n^{++}$ si et seulement si $B^T A B \in S_n^{++}$.

Preuve. Pour $x \in \mathbb{R}^n$ et $y = B^{-1}x$, on obtient $x^T Ax = y^T B^T A B y$ ■

Théorème 3.1.1 (*caractérisation des matrices définies positives*)
pour $A \in S_n$ les propriétés suivantes sont équivalentes :

1. $A \in S_n^{++}$
2. $\lambda_i(A) > 0 \forall i = 1, \dots, n$
3. Il existe $C \in M_n$ avec $\text{rang}(C) = n$ tel que $A = C^T C$
4. Pour une suite arbitraire $A_i \in S_i, i = 1, \dots, n$; de sous-matrices principales de A

$\det(A_i) > 0$ pour $i = 1, \dots, n$

Il est facile de constater qu'une matrice $A \in S_n^{++}$ si et seulement si $A^{-1} \in S_n^{++}$, car les valeurs propres de A^{-1} sont $\frac{1}{\lambda_i(A)}$ pour tout $i = 1, \dots, n$.

Théorème 3.1.2 (*complément de Schur*)

soient $A \in S_n, B \in M_{n,m}$ et $C \in S_m$ avec A inversible. Alors on a :

Si $A \in S_n^{++}$, la matrice $\begin{pmatrix} A & B \\ B^T & C \end{pmatrix}$ est définie (semi-définie) positive si et seulement si $(C - B^T A^{-1} B)$ est définie (semi-définie) positive.

Théorème 3.1.3 (Factorisation de Cholesky) Pour $A > 0$, il y a une seule matrice triangulaire inférieure et inversible L telle que $A = LL^T$.

Définition 3.1.2 Une matrice $A \in M_n$ est à diagonale strictement dominante si

$$|a_{ii}| > \sum_{i \neq j} |a_{ij}| \text{ pour tout } i = 1, \dots, n$$

Théorème 3.1.4 Si $A \in S_n$ est à diagonale strictement dominante si et seulement si tous les éléments diagonaux sont strictement positifs. Alors A est définie positive.

Théorème 3.1.5 (*caractérisations des matrices semi-définies positives*)
Pour $A \in S_n$ les propriétés suivantes sont équivalentes :

1. $A \in S_n^+$
2. $\lambda_i(A) \geq 0 \forall i = 1, \dots, n$
3. Il existe $C \in M_n$ tel que $A = C^T C$ avec $\text{rang}(C) = \text{rang}(A)$.

3.1.3 Produit de Kronecker

le produit de Kronecker est une application $\otimes : M_{m,n} \times M_{k,l} \longrightarrow M_{mk,nl}$ définie par

$$A \otimes B = \begin{pmatrix} a_{11}B & \cdots & a_{1j}B & \cdots & a_{1n}B \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ a_{i1}B & \cdots & a_{ij}B & \cdots & a_{in}B \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ a_{m1}B & \cdots & a_{mj}B & \cdots & a_{mn}B \end{pmatrix}$$

En relation avec le produit de Kronecker on doit transformer souvent une matrice $A \in M_{m,n}$ en un vecteur dans $\mathbb{R}^{m \cdot n}$, en utilisant l'opérateur $Vec : M_{m,n} \longrightarrow \mathbb{R}^{m \cdot n}$, définie par :

$$Vec(A) = \begin{pmatrix} A_{\cdot,1} \\ \vdots \\ A_{\cdot,n} \end{pmatrix}$$

On liste certaines propriétés importantes du produit de Kronecker

Proposition 3.1.2 *soient A, B, C et D des matrices des tailles appropriées*

$$(A \otimes B) = A^T \otimes B^T \quad (1,2)$$

$$A \otimes (B \otimes C) = (A \otimes B) \otimes C \quad (1,3)$$

$$(A + B) \otimes (C + D) = (A \otimes C) + (A \otimes D) + (B \otimes C) + (B \otimes D) \quad (1,4)$$

$$(A \otimes B) (C \otimes D) = (AC) \otimes (BD) \quad (1,5)$$

$$tr(A \otimes B) = tr(A) tr(B) \quad (1,6)$$

$$Vec(ABC) = (C^T \otimes A) Vec(B) \quad (1,7)$$

$$Vec(AB + BC) = (I \otimes A + C^T \otimes I) Vec(B) \quad (1,8)$$

Proposition 3.1.3 *Soient λ_i et μ_i pour $i = 1, \dots, n$ les valeurs propres de A et B respectivement avec x_i et $y_i \in \mathbb{C}^n$ les valeurs propres orthogonales correspondants. Alors toutes les valeurs propres de $A \otimes B$ sont données par $\lambda_i \mu_i$ avec les vecteurs propres orthogonaux correspondants $x_i \otimes y_i$.*

Preuve

$$\begin{aligned} (A \otimes B)(x_i \otimes y_j) &= (Ax_i) \otimes (By_j) \quad (\text{d'après (1.5)}) \\ &= (\lambda_i x_i) \otimes (\mu_j y_j) \\ &= \lambda_i \mu_j (x_i \otimes y_j) \end{aligned}$$

$$\begin{aligned} (x_i \otimes y_j)^\dagger (x_h \otimes y_k) &= (x_i^\dagger \otimes y_j^\dagger) (x_h \otimes y_k) \quad (\text{d'après (1.2)}) \\ &= (x_i^\dagger x_h) \otimes (y_j^\dagger y_k) \\ &= 0 \end{aligned}$$

■

Nous nous intéressons à l'étude des fonctions qui dépendent des matrices et leurs dérivées. Initialement, les dérivées des matrices variables peuvent causer des confusions dues aux problèmes d'arrangement des termes, dans ce cas on utilise l'opérateur Vec . Cependant on définira seulement le gradient pour les fonctions dépendantes des vecteurs, mais on réarrangera aussi le jacobien résultant dans certaines autres formes matricielles.

Soient x un vecteur dans \mathbb{R}^n et $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$; $x \mapsto (f_1(x), \dots, f_m(x))^T$ une fonction continument différentiable. Le jacobien de $f(x)$ par rapport à x est la matrice de type $n \times m$ suivante :

$$f'(x) = \begin{pmatrix} \frac{\partial f_1(x)}{\partial x_1} & \dots & \frac{\partial f_i(x)}{\partial x_1} & \dots & \frac{\partial f_m(x)}{\partial x_1} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ \frac{\partial f_1(x)}{\partial x_j} & \dots & \frac{\partial f_i(x)}{\partial x_j} & \dots & \frac{\partial f_m(x)}{\partial x_j} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ \frac{\partial f_1(x)}{\partial x_n} & \dots & \frac{\partial f_i(x)}{\partial x_n} & \dots & \frac{\partial f_m(x)}{\partial x_n} \end{pmatrix}$$

L'approximation du premier ordre de $f(x_0 + \Delta x)$ est

$$f(x_0 + \Delta x) \simeq f(x_0) + f'(x_0)\Delta x$$

Proposition 3.1.4 (Règle de composition) *étant donnée*

$$f : \mathbb{R}^n \rightarrow \mathbb{R}^m \quad \text{et} \quad g : \mathbb{R}^m \rightarrow \mathbb{R}^k$$

deux fonctions continument différentiables alors

$$(g \circ f)'(x) = g'(f(x))f'(x)$$

En particulier si A est une matrice de type $m \times n$ et $f(x) = Ax$ alors

$$(g \circ f)'(x) = g'(Ax)A$$

On illustre l'application de la dernière proposition aux matrices pour le produit AX et XA où $X, A \in S_n$. pour AX on a $Vec(AX) = (I \otimes A)Vec(X)$, et d'après (1.7)

$$\nabla_X Vec(AX) = (I \otimes A)$$

et par analogie

$$\nabla_X Vec(XA) = (A \otimes I).$$

pour obtenir des représentations agréables des linéarisations on observe que $(I \otimes A)Vec(\Delta X) = Vec(A\Delta X)$ et par conséquence :

$$[\nabla_X Vec(AX)]^T Vec(\Delta X) = (I \otimes A)Vec(\Delta X) = Vec(A\Delta X)$$

et

$$[\nabla_X Vec(XA)]^T Vec(\Delta X) = (A \otimes I)Vec(\Delta X) = Vec(\Delta XA).$$

Plus intuitivement, on écrira $A\Delta X$ et ΔXA les linéarisations de AX et XA respectivement.

3.1.4 le cône des matrices semi-définies positives

Nous allons maintenant considérer le cône des matrices semi-définies positives comme un sous-ensemble de S_n .

Définition 3.1.3 *Un ensemble $C \subset \mathbb{R}^n$ est un cône s'il est stable pour la multiplication par des scalaires positifs et l'addition. ($x, y \in C \Rightarrow x + y \in C$ et $\lambda x \in C \forall \lambda \geq 0$) un cône C est pointu si $C \cap (-C) = \{0\}$.*

Remarquons, que cette définition implique qu'un cône est un ensemble convexe.

Proposition 3.1.5 S_n^+ est un cône (convexe), pointu et fermé de pleine dimension dans $\mathbb{R}^{t(n)}$.

L'ensemble des matrices définies-positives S_n^{++} n'est pas un cône car $0 \notin S_n^{++}$. Il est facile de voir que S_n^{++} est l'intérieur du cône S_n^+ et que la frontière de S_n^+ est constituée par les matrices semi-définies positives ayant au moins une valeur propre nulle.

le lemme suivant montre que le cône des matrices semi-définies à un angle d'ouverture égale à $\frac{\pi}{2}$.

Lemme 3.1.1 *Soient $A, B \in S_n^+$. Alors $\langle A, B \rangle \geq 0$ et en plus $\langle A, B \rangle = 0$ si et seulement si $AB = 0$.*

Définition 3.1.4 *le cône polaire C^* d'un cône C est l'ensemble $\{y : \langle x, y \rangle \geq 0 \text{ pour tout } x \in C\}$.*

Pour un cône C , le cône C^* peut être vu comme l'ensemble des inégalités strictes valides pour C ou comme l'ensemble des plans tangents à C . Il est donc naturel de parler de C^* comme le cône dual de C . un cône qui vérifie $C = C^*$ est appelé identité-polaires (self-polar) ou identité-dual (self-dual).

Lemme 3.1.2 $S_n^+ = S_n^{+*}$.

Preuve. $S_n^+ \subset S_n^{+*}$ d'après le lemme 11 pour démontrer que $S_n^{+*} \subset S_n^+$, notant que pour tout $x \in \mathbb{R}^n$ la matrice xx^T est semi-définie positive. soit $A \in S_n^{+*}$, $0 \leq \langle A, xx^T \rangle = x^T A x \forall x \in \mathbb{R}^n$, d'où $A \in S_n^+$. ■

Le lemme 12 est équivalent au théorème de trace de Fejer (Fejer's trace Theorem), qu'on le formule comme un corollaire.

Corollaire 3.1.1 (théorème de trace de Fejer) $A \in S_n^+$ si et seulement si $\langle A, B \rangle \geq 0 \forall B \in S_n^+$.

le cône des matrices semi-définies positives induit une relation d'ordre partiel sur l'ensemble des matrices symétriques.

Définition 3.1.5 (la relation d'ordre partielle)

Soient $A, B \in S_n$, $A \geq B \iff (A - B) \in S_n^+$.

C'est l'origine de la notation $A \geq 0$ dans la définition 3 pour $A \in S_n^+$.

Définition 3.1.6 Soient $x = (x_1, x_2, \dots, x_n)^T \in \mathbb{R}^n$ et $X = (x_{ij}) \in M_n$. on définit les deux opérateurs *diag* et *diag par* :

$$\begin{aligned} \text{diag}(X) &= (x_{11}, x_{22}, \dots, x_{nn})^T \in \mathbb{R}^n. \\ \text{diag}(x) &= \begin{pmatrix} x_{11} & 0 & \cdots & 0 \\ 0 & x_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & x_{nn} \end{pmatrix} \in M_n \end{aligned}$$

le produit d'Hadamard ou produit de Schur, pour $A, B \in M_{n,m}$, est défini par :

$$A \circ B = [a_{ij} \cdot b_{ij}]$$

Théorème 3.1.6 (théorème du produit de Schur)

si $A, B \in S_n^+$ alors $A \circ B \in S_n^+$. si $A, B \in S_n^{++}$ alors $A \circ B \in S_n^{++}$

3.2 La programmation semi-définie

3.2.1 Les programmes semi-définis

Définition 3.2.1 un programme semi-défini, dit primal, sous forme standard s'écrit comme suit :

$$(PSDP) \begin{cases} \text{Min } \langle C, X \rangle \\ \text{s.c } AX = b \\ X \geq 0 \end{cases} \quad (2,1)$$

où $b \in \mathbb{R}^m$ et A est un opérateur linéaire de S_n tel que

$$AX = \begin{pmatrix} \langle A_1, X \rangle \\ \vdots \\ \langle A_m, X \rangle \end{pmatrix}$$

C et A_i , $i = 1, \dots, m$, sont des matrices de M_n . sans perte de généralité on peut supposer que C et les A_i sont symétriques.

Proposition 3.2.1 un programme semi-défini est un programme convexe .

Preuve. En effet , le problème (2,1) n'est autre que

$$(PSDP) \begin{cases} \text{Min} & \langle C, X \rangle \\ \text{s.c} & X \in \{Y \in S_n^+ : AY = b\} \end{cases} \quad (2,2)$$

Il est facile de voir que la fonction objectif est une fonction linéaire ,donc démontrer qu'un programme semi-définie est convexe ,revient à démontrer que l'ensemble $\{Y \in S_n^+ : AY = b\}$ est un ensemble convexe.En effet ,soient X_1 et X_2 deux éléments quelconques de $\{Y \in S_n^+ : AY = b\}$ et $\lambda \in [0, 1]$.

Il faut démontrer que $\lambda X_1 + (1 - \lambda) X_2 \in S_n^+$ car pour tout $x \in \mathbb{R}^n$ on a

$$x^T(\lambda X_1 + (1 - \lambda) X_2)x = \lambda x^T X_1 x + (1 - \lambda) x^T X_2 x \geq 0$$

En plus

$$\begin{aligned} A(\lambda X_1 + (1 - \lambda) X_2) &= \begin{pmatrix} \langle A_1, \lambda X_1 + (1 - \lambda) X_2 \rangle \\ \vdots \\ \langle A_m, \lambda X_1 + (1 - \lambda) X_2 \rangle \end{pmatrix} \\ &= \begin{pmatrix} \lambda \langle A_1, X_1 \rangle + (1 - \lambda) \langle A_1, X_2 \rangle \\ \vdots \\ \lambda \langle A_m, X_1 \rangle + (1 - \lambda) \langle A_m, X_2 \rangle \end{pmatrix} \\ &= \begin{pmatrix} \lambda b_1 + (1 - \lambda) b_1 \\ \vdots \\ \lambda b_m + (1 - \lambda) b_m \end{pmatrix} \\ &= b. \end{aligned}$$

La programmation semi-définie est la programmation linéaire sur le cone des matrices semi-définies positives.En comparaison avec la programmation linéaire standard le vecteur de variables $x \in \mathbb{R}_+^n$ sera remplacé par une matrice variable $X \in S_n^+$. Autrement dit ,le cone de l'orthant positif $x \geq 0$ sera remplacé par le cone des matrices semi-définies positives $X \geq 0$. Pour déclarer cette similitude,on utilise l'opérateur Vec , on obtient

$$(PSDP) \begin{cases} \text{Min} & Vec(C)^T Vec(X) \\ \text{s.c} & A Vec(X) = b \\ & X \geq 0 \end{cases}$$

où $A = \begin{pmatrix} A_{1,\cdot} \\ \vdots \\ A_{m,\cdot} \end{pmatrix}$, $A_{i,\cdot} = Vec(A_i)^T$ pour $i = 1, \dots, m$.

En posant $c = Vec(C)$, $x = Vec(X)$ donc $(PSDP)$ s'écrit

$$\begin{cases} \text{Min} & c^T x \\ \text{s.c} & Ax = b \\ & X \geq 0 \end{cases}$$

où $c, x \in \mathbb{R}^{n^2}$, $b \in \mathbb{R}^m$ et $A \in M_{m,n^2}$. Afin de définir le dual de (*PSDP*) nous avons besoin de l'opérateur adjoint de A .

Par définition, c'est l'opérateur $A^T : \mathbb{R}^m \rightarrow S_n$ vérifiant $\langle AX, y \rangle = \langle X, A^T y \rangle$ pour tout $X \in S_n$ et $y \in \mathbb{R}^m$. Ainsi

$$\langle AX, y \rangle = \sum_{i=1}^m y_i \text{tr}(A_i X) = \text{tr}(X \sum_{i=1}^m y_i A_i) = \langle X, A^T y \rangle$$

on obtient

$$A^T y = \sum_{i=1}^m y_i A_i$$

Pour écrire le dual de (*PSDP*), on considère la fonction lagrangienne associée à ce dernier

$$L(X, y) = \langle C, X \rangle + \langle b - AX, y \rangle, \quad X \in S_n^+, \quad y \in \mathbb{R}^m$$

et on calcule la fonctionnelle duale associée $H(y)$:

$$\begin{aligned} H(y) &= \text{Min} \langle C, X \rangle + \langle b - AX, y \rangle && X \in S_n^+ \\ &= \text{Min} \langle C, X \rangle + \sum_{i=1}^m (b_i - \langle A_i, X \rangle) y_i && X \in S_n^+ \\ &= \text{Min} \langle C - \sum y_i A_i, X \rangle + \sum b_i y_i && X \in S_n^+ \\ &= \langle b, y \rangle + \text{Min}[\langle C - A^T y, X \rangle] && X \in S_n^+ \end{aligned}$$

Il facile de voir que

$$\min_{X \geq 0} [\langle C - A^T y, X \rangle] = \begin{cases} 0 & \text{si } C - A^T y \geq 0 \\ -\infty & \text{sinon} \end{cases}$$

D'où

$$\max_{H(y)} = \begin{cases} \text{Max} \langle b, y \rangle & \text{si } C - A^T y \geq 0 \\ -\infty & \text{sinon} \end{cases}$$

Ainsi le problème dual (*DSDP*) de (*PSDP*) s'écrit sous la forme suivante :

$$(PSDP) \begin{cases} \text{Max} & \langle b, y \rangle \\ \text{s.c} & A^T y + Z = C \\ & y \in \mathbb{R}^m, Z \geq 0 \end{cases} \quad (2,3)$$

L'utilisation de la variable libre y dans (*DSDP*), peut créer un doute si (*DSDP*) est un programme semi-défini. Pour éliminer ce doute, on donne une présentation légèrement différente de (*PSDP*) et (*DSDP*), soulignée par Nesterov et Nemirovskii[71]. On suppose que le système $AX = b$ est vérifié, i.e, il existe un $\hat{X} \in S_n$ satisfait $A\hat{X} = b$. dans ce cas on peut éliminer le variable y dans (*DSDP*). On exprime tout d'abord la fonction de cout en fonction de Z

$$\langle b, y \rangle = \langle A\hat{X}, y \rangle = \langle \hat{X}, A^T y \rangle = \langle \hat{X}, C - Z \rangle$$

Notons maintenant $\text{Im}(A^T)$ l'image de A^T , et $\text{Ker}(A)$ est le noyau de A , il est facile de voir que ces deux sous-espaces sont orthogonaux, $\text{Im}(A^T) = (\text{Ker}(A))^{\perp}$, ce qui nous permet de remplacer les programmes (*PSDP*) et (*DSDP*) respectivement par les représentations suivantes :

$$\begin{cases} \min \langle C, X \rangle = \max \langle \hat{X}, C - Z \rangle \\ \text{s.c.} \begin{cases} X \in S_n^+ \cap (\hat{X} + \text{Ker}A) \\ Z \in S_n^+ \cap (C + (\text{Ker}A)^{\perp}) \end{cases} \end{cases}$$

par conséquent, (*PSDP*) et (*DSDP*), sont des programmes semi-définis et n'importe quelle propriété prise dans la formulation primale a son analogue dans la formulation duale. ■

3.2.2 Théorie de la dualité

En déduisant le problème dual de (*PSDP*) par une approche lagrangienne, on obtient le problème (2,3).

Le saut de dualité entre une solution réalisable duale (y, Z) et une solution réalisable primale X est

$$\langle C, X \rangle - \langle b, y \rangle - \langle A^T y + Z, X \rangle - \langle AX, y \rangle - \langle Z, X \rangle \geq 0$$

d'après le lemme 11, la propriété que la valeur objective de chaque solution réalisable primale est plus grande ou égale à la valeur objective de chaque solution réalisable duale et la différence positive est appelée la dualité faible.

Contrairement à la programmation linéaire, il n'est pas toujours vrai que l'optimalité de (*PSDP*) et (*DSDP*) implique que $\langle Z, X \rangle = 0$

Exemple 3.2.1 *Considérons le problème semi-défini suivant :*

$$\left\{ \begin{array}{l} \min \left\langle \left(\begin{array}{ccc} 0 & \frac{1}{2} & 0 \\ & 0 & 0 \\ & & 0 \end{array} \right), X \right\rangle \\ \text{s.c.} \left\{ \begin{array}{l} \left\langle \left(\begin{array}{ccc} 0 & -\frac{1}{2} & 0 \\ & 0 & 0 \\ & & 1 \end{array} \right), X \right\rangle = 1 \\ \left\langle \left(\begin{array}{ccc} 1 & 0 & 0 \\ & 0 & 0 \\ & & 0 \end{array} \right), X \right\rangle = 0 \\ \left\langle \left(\begin{array}{ccc} 0 & 0 & 1 \\ & 0 & 0 \\ & & 0 \end{array} \right), X \right\rangle = 0 \\ \left\langle \left(\begin{array}{ccc} 0 & 0 & 0 \\ & 0 & 1 \\ & & 0 \end{array} \right), X \right\rangle = 0 \\ X \geq 0. \end{array} \right. \end{array} \right.$$

qui est équivalent à :

$$\left\{ \begin{array}{l} \min x_{12} \\ \text{s.c.} X = \left(\begin{array}{ccc} 0 & x_{12} & 0 \\ & x_{22} & 0 \\ & & 1 + x_{12} \end{array} \right) \geq 0 \end{array} \right.$$

En écrivant le problème dual associé, on obtient .

$$\left\{ \begin{array}{l} \max y_1 \\ \text{s.c.} Z = C - y_1 A_1 - y_2 A_2 - y_3 A_3 - y_4 A_4 \geq 0 \end{array} \right.$$

Le problème dual peut s'écrire sous la forme suivante :

$$\left\{ \begin{array}{l} \max y_1 \\ \text{s.c.} Z = \left(\begin{array}{ccc} y_2 & \frac{1+y_1}{2} & y_3 \\ & 0 & -y_4 \\ & & -y_1 \end{array} \right) \geq 0 \end{array} \right.$$

Une condition nécessaire pour que la matrice primale soit semi-définie positive est que $x_{12} = 0$ car $x_{11} = 0$, de la même manière obtient que $x_{22} = 0$ car $x_{22} = 0$, i.e. $y_1 = -1$, dans le problème dual. le saut de dualité entre deux solutions optimales primale et duale est égal à 1.

On retourne maintenant aux conditions qui assurent la dualité forte, et l'existence des solutions primales et duales.

Définition 3.2.2 (Réalibilité stricte) *Un point X est dit strictement réalisable pour (PSDP) s'il est réalisable pour (PSDP) et vérifie $X > 0$.*

Un point (y, Z) est dit strictement réalisable pour (DSDP) s'il est réalisable pour (DSDP) et vérifie $Z > 0$.

Théorème 3.2.1 (La dualité forte[15]) *Supposons qu'il existe une solution strictement réalisable (y_0, Z_0) pour (DSDP). Soient*

$$p^* = \inf \{ \langle C, X \rangle : AX = b, X \geq 0 \}$$

et

$$q^* = \sup \{ \langle b, y \rangle : A^T y + Z = C, Z \geq 0 \}$$

alors $p^ = q^*$ et si p^* est une valeur finie, elle est atteinte pour une certaine matrice*

$$X \in \{ X \geq 0 : AX = b \}.$$

Corollaire 3.2.1 *Soient p^* et q^* définis comme dans le théorème de dualité forte*

1- Si (PSDP) est strictement réalisable avec p^ finie, alors $p^* = q^*$ et cette valeur est atteinte pour (DSDP).*

2- Si (DSDP) est strictement réalisable avec q^ finie, alors $p^* = q^*$ et cette valeur est atteinte pour (PSDP).*

3- Si (PSDP) et (DSDP) sont strictement réalisables, alors $p^ = q^*$ et ces deux valeurs sont atteintes pour les deux problèmes.*

3.3 Méthodes de points intérieurs

3.3.1 Introduction

Dans le cadre de la programmation mathématique avec contraintes, on désigne par méthode de points intérieur, toute procédure de résolution générant une suite de points appartenant à l'intérieur relatif du domaine réalisable et convergeant vers une solution optimale.

De même, on appelle fonction barrière toute fonction f qui vérifie,

- f est à valeurs finies à l'intérieur relatif du domaine réalisable.
- $f(x) \rightarrow \infty$ quand x s'approche de la frontière.

Ces méthodes sont réputées pour leur convergence polynomiale, leur rapidité et efficacité et se sont révélées comme de véritables concurrentes des méthodes classiques (simplexe, pivotage de Lemk, etc). La littérature sur ces méthodes a connu une grande expansion et s'est enrichie plusieurs classes et variantes dans le but de réduire la complexité et améliorer la convergence et l'efficacité. Le lecteur peut se reporter, par exemple, les livres de Ye [19] et Wright [16] qui retracent l'évolution des méthodes de points intérieurs.

Il y a pratiquement trois catégories des méthodes de points intérieurs : les méthodes affines, les méthodes de réduction de potentiel, et les méthodes de trajectoire centrale.

En ce qui concerne le problème (*SDP*) en particulier, l'impact des méthodes de points intérieurs est sans précédent, dans cette mémoire on présente l'algorithme de méthode de trajectoire centrale.

3.3.2 Méthodes de trajectoire centrale

Ces méthodes sont le fruit direct d'une grande partie des études acharnées menées par plusieurs chercheurs vers la fin des années 80, et pleinement développées au début des années 90. Elles possèdent les propriétés théoriques les plus esthétiques : complexité polynomiale, convergence superlinéaire et caractère algorithmique newtonien. Ces qualités de confort placent cette classe de méthodes au centre de l'intérêt primordial des chercheurs pour résoudre effectivement des programmes mathématiques avec contraintes.

Revenons à notre problème (*SDP*), en rappelant ses formulations primale et duale :

$$(PSDP) \begin{cases} \text{Min } \langle C, X \rangle \\ \text{s.c } AX = b \\ X \geq 0 \end{cases} \quad (DSDP) \begin{cases} \text{Max } \langle b, y \rangle \\ \text{s.c } A^T y + Z = C \\ y \in \mathbb{R}^m, Z \geq 0 \end{cases}$$

D'après le théorème de la dualité forte et son corollaire une condition suffisante pour atteindre des solutions optimales primales et duales et pour assurer la dualité forte est l'existence des solutions primales et duales strictement réalisables.

Proposition 3.3.1 *Supposons qu'il existe une solution strictement réalisable X^0 pour (*PSDP*) et une solution strictement réalisable (y^0, Z^0) pour (*DSDP*).*

comme les algorithmes de points intérieurs démarrent avec un point appartenant à l'intérieur du cône des matrices semi-définies positives, on remplace les deux problèmes initiaux (*PSDP*) et (*DSDP*) par deux suites de problèmes de barrières auxiliaires, (P_μ) et (D_μ) paramétrées par le scalaire $\mu > 0$, pour rester à l'intérieur de S_n^+ .

$$(P_\mu) \begin{cases} \text{Min } \langle C, X \rangle - \mu \log \det(X) \\ \text{s.c } AX = b \\ X > 0 \end{cases} \quad (D_\mu) \begin{cases} \text{Max } \langle b, y \rangle + \mu \log \det(Z) \\ \text{s.c } A^T y + Z = C \\ y \in \mathbb{R}^m, Z > 0 \end{cases}$$

Comme les fonctions objectives de (P_μ) et (D_μ) sont respectivement strictement convexe et concave, alors les solutions X_μ de (P_μ) et (y_μ, Z_μ) de (D_μ) existent et sont uniques.

Ici $\mu > 0$ est appelé le paramètre de barrière et $\log(\det(Y))$ ($Y \in S_n^{++}$) est la fonction barrière. Cette fonction tend vers l'infini quand une valeur propre de la matrice Y tend vers zero, i.e. quand Y s'approche de la frontière de S_n^+ .

On transforme les problèmes de barrières en des problèmes sans contraintes, en utilisant les multiplicateurs de lagrange pour les contraintes d'égalité :

$$\begin{aligned} L_{p\mu}(X, y) &= \langle C, X \rangle - \mu \log \det(X) + \langle b - AX, y \rangle \\ L_{D\mu}(X, y, Z) &= \langle b, y \rangle + \mu \log \det(Z) + \langle X, C - A^T y - Z \rangle \end{aligned}$$

En écrivant les conditions de Karuch-Tucker (KKT), on obtient

$$\begin{aligned} 0 = \nabla L_{p\mu}(X, y) &= \begin{pmatrix} C - \mu X^{-1} - A^T y \\ b - AX \end{pmatrix} \\ 0 = \nabla L_{D\mu}(X, y, Z) &= \begin{pmatrix} C - Z - A^T y \\ b - AX \\ \mu Z^{-1} - X \end{pmatrix} \end{aligned}$$

On peut réécrire ces conditions comme :

$$(S_\mu) \begin{cases} AX = b, X > 0 \\ A^T y + Z = C, Z > 0 \\ XZ = \mu I \end{cases} \quad (2,5)$$

les deux premières conditions correspondent au réalisabilité primale et duale. Pour $\mu = 0$, la troisième condition correspond à la condition de complémentarité $XZ = 0$.

On note par (X_μ, y_μ, Z_μ) la solution de (S_μ) pour $\mu > 0$ fixé.

L'ensemble des solutions (X_μ, y_μ, Z_μ) pour tout $\mu > 0$ forme la trajectoire centrale qui converge vers une solution optimale (X^*, y^*, Z^*) quand μ tend vers 0.

si on résout (S_μ) par la méthode de Newton, on obtient le système linéaire suivant :

$$\begin{cases} A\Delta X = -(AX - b) = R_p \\ A^T \Delta y + \Delta Z = -(A^T y + Z - C) = R_d \\ \Delta XZ + \Delta ZX - \mu I - XZ = R_c \end{cases} \quad (2,6)$$

La résolution de ce système donne une matrice symétrique ΔZ , par contre la matrice ΔX n'est pas nécessairement symétrique. Comme l'itéré suivant doit être symétrique (défini positif), cela pose un vrai problème. Pour cela on introduit l'opérateur de symétrisation H_p pour une matrice inversible donnée $P \in S_n$, soit $M \in M_n$, on définit

$$H_p = \frac{1}{2} \left(PMP^{-1} + (PMP^{-1})^T \right)$$

Différents choix de P coresspondent à différentes directions de recherche. Nous citons ici les trois approches les plus célèbres.

La première approche de (H.R.V.W. [4]), (K.S.H.[8]), et (M.[9]) correspond à $P = Z^{\frac{1}{2}}$.

la deuxième approche de (A.H.O.[1]) correspond à $P = I$.

La troisième approche de (N.T.[11]) correspond à P vérifiant $P^T P = W = X^{\frac{1}{2}} \left(X^{\frac{1}{2}} Z X^{\frac{1}{2}} \right)^{\frac{1}{2}} X^{\frac{1}{2}}$.

Avant de résoudre le système (2,6), nous faisons un changement de variables pour faciliter le calcul, En posant

$$c = \text{Vec}(C), x = \text{Vec}(X), z = \text{Vec}(Z) \text{ et } A = \begin{pmatrix} \text{Vec}(A_1)^T \\ \vdots \\ \text{Vec}(A_m)^T \end{pmatrix},$$

le système (2,6) peut s'écrire comme

$$\begin{aligned} A\Delta X &= r_p \\ A^T \Delta y + \Delta z &= r_d \\ E\Delta x + F\Delta z &= r_c \end{aligned}$$

où $E = Z \otimes I, F = I \otimes X$, et $r = \text{Vec}(R)$. Le système (2,6) est résolu en exprimant Δx en fonction Δz dans (2,9), $\Delta x = E^{-1}(r_c - F\Delta z)$, et Δz en fonction de Δy dans (2,8), $\Delta z = r_d - A^T \Delta y$. En remplaçant $\Delta x = E^{-1}(r_c - F(r_d - A^T \Delta y))$ dans (2,7) on obtient

$$M\Delta y = r_p + AE^{-1}(Fr_d - r_c)$$

où $M = AE^{-1}FA^T, M \in S_m^{++}$ si on suppose que A est de plein rang en lignes. Le travail le plus coûteux est la formation et la factorisation de M à chaque itération.

Définition 3.3.1 (Voisinage de la trajectoire centrale) On dit que (X, Z) est dans le voisinage de la trajectoire centrale (noté $V_F(\theta)$) s'il vérifie :

$$\|H_P(XZ) - \mu(X, Z)I\|_F \leq \theta \mu(X, Z)$$

pour certain $\theta \in (0, 1)$.

Nous présentons maintenant la description générale de l'algorithme de points intérieurs pour résoudre le problème (SDP).

Algorithme

Début algorithme**Initialisation** $k = 0;$ Entrée : A, b, C et un point initial $(X^{(0)}, y^{(0)}, Z^{(0)}) \in V_F(\theta)$ $l = 0;$ Pour un choix convenable de $0 < \theta < 1, 0 < \sigma < 1$ Un paramètre de précision $\epsilon > 0;$

$$\mu_k = \frac{\langle X^{(k)}, Z^{(k)} \rangle}{n}$$

Tant que $n\mu_k > \epsilon$ faire :**Etape 1 :** On résout le système suivant pour trouver $(\Delta X, \Delta y, \Delta Z)$.

$$A\Delta X = 0$$

$$A^T \Delta y + \Delta Z = 0$$

$$H_P(\Delta X Z^{(k)} + X^{(k)} \Delta Z) = \sigma \mu_k I - H_P(X^{(k)} Z^{(k)})$$

Etape 2 : $(X^{(k)}, y^{(k)}, Z^{(k)}) = (X^{(k)} + \Delta X, y^{(k)} + \Delta y, Z^{(k)} + \Delta Z)$ **Etape 3 :** Calculer $\mu_k = (1 - \theta) \mu_k$ $l = l + 1;$ **Fin tant que;****Etape 4** La solution optimale est $(X^*, y^*, Z^*) = (X^{(k)}, y^{(k)}, Z^{(k)})$ Le nombre des itération nécessaire d'exécution cet algorithme est : l ;**Fin algorithme.**

3.3.3 Étude de problème perturbé $(SDP)_\mu$

On commence par le lemme suivant.

Lemme 3.3.1 $\{Y \in S_n^+ : \langle C, Y \rangle \leq 0, \langle A_i, Y \rangle = 0, i = 1 \dots m\} = \{0\}$.**Preuve.** Nous savons que l'ensemble des solutions optimales de SDP ($Sol(SDP)$) est convexe fermé borné et non vide. Par conséquent, son cône de récession $Sol(SDP)^\infty$ est réduit aux singleton $\{0\}$. Mais

$$Sol(SDP)^\infty = \{Y \in S_n^+ : \langle C, Y \rangle \leq 0, \langle A_i, Y \rangle = 0, i = 1 \dots m\} = \{0\}.$$

■

Théorème 3.3.1 *Le problème $(SDP)_\mu$ admet une solution optimale unique.***Preuve.**

1. Existence : Il suffit de prouver que le cône de récession de l'ensemble convexe fermé

$$S_\lambda = \{X : \langle A_i, X \rangle = b_i \forall i, f_\mu(x) \leq \lambda\} \lambda \in \mathbb{R}.$$

est réduit à $\{0\}$. On a :

Calculons la fonction de récession f_μ^∞ . Soit $X \in S_n^{++}$, alors pour $Y \in S_n$,

$$f_\mu^\infty(Y) = \lim_{t \rightarrow +\infty} \frac{f_\mu(X + tY) - f_\mu(X)}{t}$$

$$S_\lambda^\infty = \{Y : \langle A_i, Y \rangle = 0, i = 1, \dots, m\} \cap \{f_\mu^\infty(y) \leq 0\}.$$

Il est clair que si $Y \notin S_n^+$, pour t assez grand, $X + tY$ n'appartient pas à S_n^+ et donc $f_\mu(Y) = +\infty$. Supposons que $Y \in S_n^+$. Alors :

$$f_\mu^\infty(Y) = \langle C, Y \rangle - \mu \lim_{t \rightarrow +\infty} \frac{\ln \det(X + tY) - \ln \det(X)}{t}$$

Puisque X est définie positive alors il existe une matrice symétrique définie positive Z telle que $Y = Z^2$. Il existe une matrice P et une matrice diagonale semi-définie D telle que

$$Z^{-1}YZ^{-1} = P^T P = I.$$

Il résulte que :

$$\det(X + tY) = \det(ZP^t(I + tD)PZ) = \det \prod_{i=1}^n (I + td_i),$$

où les d_i sont les éléments diagonaux de la matrice D . Donc

$$f_\mu^\infty(Y) = \langle C, Y \rangle - \mu \sum_{i=1}^n \lim_{t \rightarrow +\infty} \frac{\ln(1 + td_i)}{t} = \langle C, Y \rangle.$$

Comme

$$S_\lambda^\infty = \{Y : \langle A_i, Y \rangle = 0, i = 1 \dots m\} \cap \{f_\mu^\infty(y) \leq 0\}.$$

il s'ensuit :

$$S_\lambda^{\infty} = \{Y \in S_n^+ : \langle C, Y \rangle \leq 0, \langle A_i, Y \rangle = 0, i = 1 \dots m\}$$

En appliquant le lemme 3.3.1, nous obtenons

$$S_\lambda^\infty = \{0\}$$

2. Unicité : Comme f_μ est strictement convexe, il en résulte que la solution optimale de (P_μ) est unique.

■

3.3.4 Condition d'optimalité pour (P_μ)

Le problème (P_μ) est convexe, donc les conditions de K.K.T sont nécessaires et suffisantes. $X \in S_n^{++}$ est une solution optimale pour le problème (P_μ) si et seulement s'il existe $y \in \mathbb{R}^m$ tel que :

$$\begin{cases} C - \mu X^{-1} - \sum_{i=1}^m y_i A_i = 0 \\ \langle A_i, X \rangle = b_i, i = 1 \dots m \end{cases} \quad (a)$$

Posons $Z = \mu X^{-1}$, le système (a) devient :

$$\begin{cases} \sum_{i=1}^m y_i A_i + Z = C, \quad Z \in S_n^{++} \\ \langle A_i, X \rangle = b_i, \quad i = 1, \dots, m, \quad X \in S_n^{++} \\ XZ = \mu I, \quad \mu > 0. \end{cases} \quad (b)$$

Qui est le système paramétrisé du système (a). On note par $(X(\mu), y(\mu), Z(\mu))$ une solution du système (b). Nous savons déjà que $X(\mu)$ (et aussi $Z(\mu)$) est définie de façon unique. En supposons que les matrices A_i ($i = 1, \dots, m$) sont linéairement indépendantes, alors, $y(\mu)$ est aussi définie de manière unique.

Théorème 3.3.2 [8] Si $(X(\mu), y(\mu), Z(\mu))$ est une solution optimale de (P_μ) avec $\mu > 0$, alors :

$$\lim_{\mu \rightarrow 0} (X(\mu), y(\mu), Z(\mu)) = (X^*, y^*, Z^*),$$

est une solution optimale de (SDP).

Pour faciliter l'étude, nous considérons dans la suite la notation (X, y, Z) au lieu de $(X(\mu), y(\mu), Z(\mu))$.

Conclusion générale

Les méthodes de points intérieurs sont connues par leur efficacité, rapidité de convergence, simplicité algorithmique et capacité de résoudre des problèmes de grandes tailles. L'inconvénient principal dans ce type de méthodes est l'initialisation, c'est-à-dire la détermination d'un point initial qui se trouve à l'intérieur du domaine. La plupart des résultats théoriques supposent que ce point initial est connu, mais numériquement l'obtention de ce point initial prend beaucoup de temps.

Dans cette mémoire nous avons présenté une partie de l'introduction sur les problèmes semi-définis positifs (*SDP*), nous avons commencé par donner quelques rappels sur les matrices : Élément d'analyse convexe, problème général d'optimisation,.... ; puis la programmation mathématique : problème d'optimisation avec contrainte, programme mathématique, algorithme d'optimisation,.... ; et enfin la programmation semi-définie positive qui a connu grâce aux méthodes de points intérieurs une évolution considérable sur tous les aspects : théorique, algorithmique et numérique et on a fait une présentation algorithmique avec une méthode pratique et connue (*TC*) qui se base sur la méthode de Newton qui a comme méfait la difficulté de choisir (X^0, y^0, Z^0) , et comme bienfait lorsque (X^0, y^0, Z^0) appartient à un certain voisinage V elle converge théoriquement.

Les recherches futures pourraient se concentrer sur l'extension des autres types d'optimisation et les tests numériques pour étudier le comportement d'algorithme afin d'être comparée avec d'autres approches existantes.

Bibliographie

- [1] F. Alizadeh, J.-P.A. Haeberly, and M.L. Overton. *Primal-dual interior-point methods for Semidefinite programming : Convergence rates, stability and numerical results*. SIAM J. Optim., 8(3) :746-768, (1998).
- [2] M. Bouafia, *Étude asymptotique des méthodes de points intérieurs pour la programmation linéaire*, Thèse de doctorat en Mathématiques appliquées et application des mathématiques, Université du Havre, (2016).
- [3] G. Ciarlet, *Introduction à l'analyse numérique matricielle et à l'optimisation*, Masson, Paris, (1985).
- [4] C. Helmberg, F. Rendl, R. J. Vanderbei, and H. Wolkowicz. *An interior-point method for Semidefinite programming*. SIAM J. Optim., 6(2) :342-361, (1996).
- [5] A. Dehghani, J.L. Goffin, D. Orban, *A primal-dual regularized interior-point method for semidefinite programming*, Optimization Methods and Software, 30(1), 193–219, (2017).
- [6] N. K. Karmarkar, *A new polynomial-time algorithm for linear programming*, in : Proceedings of the 16th Annual ACM Symposium on Theory of Computing, 4, 373–395, (1984).
- [7] L. G. Khachiyan, *A polynomial algorithm in linear programming*, Soviet Mathematics Doklady, 20 191–194, (1979).
- [8] M. Kojima, S. Shindoh, and S. Hara. *Interior-point methods for the monoton linear complementarity problem in symmetric matrices*. SIAM J. Optim., 7(1) :86-125, (1997).
- [9] R.D.C. Monteiro. *Primal-dual path -following algorithms for Semidefinite programming*. SIAM J. Optim., 7(3) :663-678, (1997).
- [10] Y. E. Nesterov, A. S. Nemirovski, *Interior-point polynomial methods in convex programming*, SIAM Studies in Applied Mathematics. SIAM Publications, Philadelphia, (1994).
- [11] Y. Nesterov and M.J. Todd. *Self-scaled barriers and interior-point methods for convex programming*. Math. Oper. Res., 22 :1-42, (1997).
- [12] J. Renegar, *A polynomial-time algorithm, based on Newton's method, for linear programming*, Mathematical Programming, 40 59–93, (1988).

Résumé

Dans le présent travail, on s'intéresse à l'étude algorithmique de la méthode de trajectoire centrale (TC) de type primal-dual pour résoudre les problèmes de la programmation semi définie positive (SDP). Cette méthode est basée sur une nouvelle classe de direction de Newton. Plus précisément, on donne l'étude algorithmique de l'algorithme de la méthode (TC).

Mots Clés : Programmation Linéaire, Programmation Semi Définie positive, Méthode de Points Intérieurs, méthode de trajectoire centrale.

Abstract

In the present work, we are interested in the study algorithmic of the method of central trajectory (TC) of primal-dual to solve the problems of the semi defined positive programming (SDP). This method is based on a new class of direction of Newton. More exactly, we give the study algorithmics of the algorithm of the method (TC).

Key Words: Linear Programming, Semi Defined Positive programming, Interior Points Methods, Method of central trajectory.