

REPUBLIQUE ALGERIENNE DEMOCRATIQUE ET POPULAIRE



MINISTRE DE L'ENSEIGNEMENT SUPERIEUR  
ET DE LA RECHERCHE SCIENTIFIQUE

UNIVERSITE FERHAT ABBAS SETIF-1  
FACULTE DES SCIENCES  
DEPARTEMENT DE MATHEMATIQUES



## THESE

Présentée par :

**TOUIL Imene**

Pour obtenir le diplôme de doctorat en Sciences

OPTION

MATHEMATIQUES APPLIQUEES

## THEME

*Etude théorique et numérique des méthodes de points intérieurs de type trajectoire centrale pour la programmation semi-définie linéaire*

Soutenue Le : 06/02/2017  
devant le jury composé de :

Président : Mr. M. ACHACHE

Prof. U. F. A. Sétif-1.

Rapporteurs : Mr. D. BENTERKI

Prof. U. F. A. Sétif-1.

Examineurs : Mr. R. BENZINE

Prof. E. S. T. I. Annaba.

: Mr. N. BENHAMIDOUCHE

Prof. U. M. B. M'sila.

Année Universitaire : 2016/ 2017

## Remerciements

(فَسُبْحَانَ اللَّهِ حِينَ تُمْسُونَ وَحِينَ تُصْبِحُونَ (17) وَلَهُ الْحَمْدُ فِي السَّمَاوَاتِ وَالْأَرْضِ وَعَشِيًّا وَحِينَ تُظْهِرُونَ (18)) (الروم)  
صدق الله العظيم

*Je tiens à remercier mon directeur de thèse Monsieur le Professeur Djamel Benterki qui m'a offert la possibilité de réaliser ce travail et m'a encadré durant ces années, mais aussi pour son relectures et commentaires.*

*Je remercie également les membres du jury :*

*Mr. M. Achache, professeur à l'Université de Ferhat Abbas de m'avoir fait l'honneur de présider ce jury et de juger mon travail ; Mr. R. Benzine, professeur à l'Ecole Supérieur de Technologie et d'Informatique d'avoir évalué mon travail et Mr. N. Benhamidouche professeur à l'Université Mohamed Boudiaf d'avoir accepté de faire partie de ce jury.*

*J'adresse mes remerciements aux personnes qui m'ont aidé dans la réalisation de ce travail.*

*Mes remerciements vont aussi à mes collègues à l'université Mohamed Seddik Ben Yahia et à tous les membres du Département de Mathématiques à l'Université de Ferhat Abbas.*

*Je ne pourrais jamais oublier le soutien et l'aide des personnes chères de ma nombreuse et merveilleuse famille. Je réserve une reconnaissance particulière à mes parents qui m'ont apporté la confiance dont j'avais besoin pour continuer à parcourir cette aventure et à mon mari pour son encouragement et soutien continu.*

*A tout les membres de ma grande famille.*

# Table des matières

<b>Introduction</b>	<b>4</b>
<b>1 Préliminaires et notions fondamentales</b>	<b>8</b>
1.1 Analyse matricielle . . . . .	8
1.1.1 Déterminant, spectre et norme spectrale . . . . .	8
1.1.2 Trace d'une matrice . . . . .	10
1.1.3 Produit scalaire et norme de Frobenius . . . . .	11
1.1.4 Matrices (semi-) définies positives . . . . .	11
1.2 Analyse convexe . . . . .	15
1.2.1 Ensembles et fonctions affines . . . . .	16
1.2.2 Ensembles et fonctions convexes . . . . .	17
1.2.3 Le cône des matrices symétriques semi-définies positives . . . . .	21
1.3 Programmation mathématique . . . . .	22
1.3.1 Définitions . . . . .	22
1.3.2 Principaux résultats d'existence et d'unicité . . . . .	24
1.3.3 Conditions d'optimalité . . . . .	25
1.3.4 Programme linéaire . . . . .	26
<b>2 La programmation semi-définie : Théorie, Applications et Résolution</b>	<b>27</b>
2.1 Les programmes semi-définis . . . . .	27
2.1.1 Exemples . . . . .	28
2.2 Théorie de la dualité . . . . .	30

2.2.1	Exemples pathologiques . . . . .	33
2.2.2	Relations primales-duales pour la programmation semi-définie . . . . .	35
2.3	Domaines d'applications en PSD . . . . .	36
2.3.1	Optimisation des valeurs propres . . . . .	37
2.3.2	Programmation quadratique avec contraintes quadratiques . . . . .	39
2.3.3	Approximation logarithmique de Tchebychev . . . . .	40
2.3.4	Problèmes géométriques en formes quadratiques . . . . .	42
2.3.5	Optimisation quadratique non convexe . . . . .	43
2.4	Méthode de points intérieurs pour résoudre (PSD) . . . . .	46
2.4.1	Méthodes affines . . . . .	46
2.4.2	Méthodes de réduction du potentiel . . . . .	47
2.4.3	Méthodes de trajectoire centrale . . . . .	49
<b>3</b>	<b>Méthode réalisable de trajectoire centrale pour la programmation semi-définie</b> . . . . .	<b>52</b>
3.1	Position du problème . . . . .	53
3.2	Existence et unicité de la solution optimale du problème $(PSD)_\mu$ et sa convergence vers une solution du (PSD) . . . . .	54
3.2.1	Existence et unicité de la solution optimale de $(PSD)_\mu$ . . . . .	54
3.2.2	Convergence de la solution optimale de $(PSD)_\mu$ vers la solution optimale de (PSD) . . . . .	56
3.3	Méthode de trajectoire centrale primale-duale . . . . .	57
3.3.1	Calcul du pas de déplacement . . . . .	60
3.4	L'algorithme prototype et l'analyse de sa complexité . . . . .	66
3.4.1	Algorithme de la méthode de trajectoire centrale . . . . .	66
3.4.2	Résultats de convergence . . . . .	67
3.4.3	Analyse de la complexité . . . . .	69
3.5	Mise en oeuvre de la méthode de trajectoire centrale pour (PSD) . . . . .	76

3.5.1	Exemples à taille fixe . . . . .	76
3.5.2	Exemple à taille variable . . . . .	80
3.5.3	Commentaires . . . . .	81
	<b>Conclusion</b>	<b>82</b>

# Introduction

La programmation mathématique, branche de l'optimisation, s'occupe de la minimisation (ou maximisation) sous contraintes (domaine réalisable) d'une fonction à plusieurs variables, schéma très général s'appliquant à de nombreuses situations pratiques dans beaucoup de domaines (minimisation de coûts, de durées,... etc).

La programmation semi-définie (*PSD*) est l'un des problèmes d'optimisation. 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'algorithmes efficaces permettant de les résoudre : il s'agit des algorithmes de méthodes de points intérieurs.

L'évolution rapide et le succès des méthodes de points intérieurs (*MPIs*) depuis leur relance par Karmarkar (1984) [23] dans le domaine de la programmation linéaire, ont incité les chercheurs du monde entier à développer tout un arsenal de méthodes permettant de traiter convenablement plusieurs classes de problèmes considérées jadis difficiles à résoudre, parmi lesquelles la programmation semi-définie (*PSD*). Grâce à ces méthodes, la programmation semi-définie a connu une évolution considérable sur tous les aspects : théorique, algorithmique et numérique.

Un tel succès constitue un bon stimulant pour d'autres développements. On parle déjà de (*PSD*) quadratique, non linéaire et programmation semi-infinie.

On désigne par méthode de points intérieurs (*MPI*), toute procédure de résolution générant une suite de points appartenant à l'intérieur (relatif) du domaine réalisable (admissible) et convergeant vers une solution optimale du programme considéré. Il y a

principalement trois grandes catégories de méthodes de points intérieurs : les méthodes affines, les méthodes de réduction du potentiel et les méthodes de trajectoire centrale. En 1990, Monteiro et al [36] ont montré la convergence polynomiale de la méthode affine primale-duale. La généralisation des méthodes de points intérieurs (*MPIs*) de la programmation linéaire (*PL*) à la programmation semi-définie (*PSD*) a commencé au début des années 1990. Les premiers travaux de (*MPIs*) pour (*PSD*) ont été développés indépendamment par Alizadeh [2] et Nesterov et Nemirousky [41]. En 1995, Alizadeh [2] a proposé une méthode de réduction du potentiel de Ye pour (*PSD*). Presque en même temps, en 1994, Nesterov et al [41] ont prouvé que (*MPIs*) sont capables de résoudre les problèmes d'optimisation conique généraux, en particulier les problèmes (*PSD*), en un temps polynomial. En 1996, Vanderbeghe et al [52] ont proposé un algorithme primal-dual de la méthode de réduction du potentiel. En 1997, Monteiro [38], a proposé une méthode de trajectoire centrale primale-duale. En 1999, Monteiro et al [39] ont proposé une méthode primale-duale de type trajectoire centrale et ont prouvé que la complexité de son algorithme est polynomiale. Au même temps, Ji et al [30] ont étudié la convergence de la méthode de trajectoire centrale de type prédicteur-correcteur. En 2002, Halicka et al [15] ont proposé une étude sur la convergence de la méthode de trajectoire centrale en programmation semi-définie. Une année après, Benterki et al [6] ont fait une étude aménagée théorique et numérique de l'algorithme proposé par Alizadeh [2]. Les travaux se poursuivent à présent, on trouve ainsi ceux de Koulaei et al [29] en (2007), qui se sont intéressés à l'extension de l'algorithme de type Mehrotra pour la programmation semi-définie. La même année, Benterki et al [7] ont proposé une méthode de points intérieurs réalisable pour résoudre le problème (*PSD*), l'avantage de cette dernière est qu'elle fournit une solution strictement réalisable initiale pour (*PSD*). En 2012, Liu et al [31] ont présenté un nouveau algorithme de type correcteur de second ordre pour (*PSD*) et ont prouvé sa convergence polynomiale pour la direction

de Nesterov et Todd (*NT*). En 2015, Kettab et al [27] ont proposé une relaxation de la méthode barrière logarithmique pour la programmation semi-définie.

La plupart de ces plus récents travaux sont concentrés sur les méthodes primales-duales.

Notre travail apporte sur des contributions théoriques, algorithmiques et numériques pour résoudre convenablement des problèmes d'optimisation de la programmation semi-définie linéaire. On s'intéresse plus particulièrement, aux méthodes de type trajectoire centrale primale-duale, à cause de leurs succès numérique vis-à-vis d'autres classes de méthodes de points intérieurs.

Cette thèse est répartie en trois chapitres. Le premier chapitre présente un rappel des notions fondamentales d'usage fréquent pour la suite, à savoir : l'analyse matricielle, l'analyse convexe, la notion des cônes des matrices symétriques semi-définies positives en donnant ses principales propriétés et quelques résultats de la programmation mathématique.

Le chapitre 2, est consacré à la programmation semi-définie. On donne alors la formulation générale de ce type de problème et sa dualité. On précise ainsi quelques domaines d'application et on termine par les méthodes de résolution d'un problème (*PSD*).

Le dernier chapitre est le centre de ce travail. Sur le plan théorique, nous montrons que le problème perturbé par le paramètre  $\mu$  admet une solution optimale unique pour chaque valeur de  $\mu$  fixé et converge vers une solution optimale du problème originel (*PSD*). Dans une deuxième partie, on présente d'une manière détaillée une méthode primale-duale de points intérieurs de type trajectoire centrale pour résoudre le problème (*PSD*) dans laquelle, en introduisant quatre nouvelles alternatives donnant lieu au calcul effectif du pas de déplacement intervenant dans l'algorithme. L'ensemble de ses résultats constituent l'originalité de notre travail. Ce chapitre est étayé par des



expérimentations numériques d'ordre comparatif sur quelques exemples connus dans la littérature.

# Chapitre 1

## Préliminaires et notions fondamentales

Dans ce chapitre, nous allons introduire quelques notions fondamentales et résultats bien connus sur les matrices carrées, symétriques et semi-définies positives, ainsi que des notions de base de l'analyse convexe. Nous présentons aussi certains éléments fondamentaux de la programmation mathématique.

### 1.1 Analyse matricielle

On notera  $\mathcal{M}_{m,n}(\mathbb{R})$  l'espace vectoriel des matrices réelles  $m \times n$  et  $\mathcal{M}_n(\mathbb{R})$  l'espace vectoriel des matrices carrées d'ordre  $n$ , (i.e.,  $\mathcal{M}_n(\mathbb{R}) = \mathcal{M}_{n,n}(\mathbb{R})$ ). L'espace vectoriel des matrices carrées symétriques d'ordre  $n$  est noté  $\mathcal{S}^n$ .

#### 1.1.1 Déterminant, spectre et norme spectrale

**Définition 1.1.1** Soit  $\sigma_n$  le groupe des permutations de  $\{1, \dots, n\}$ . On définit le déterminant de  $A \in \mathcal{M}_n(\mathbb{R})$  d'éléments  $(a_{ij})$  par

$$\det A = \sum_{\sigma \in \sigma_n} \epsilon_\sigma a_{\sigma(1)1} a_{\sigma(2)2} \dots a_{\sigma(n)n}$$

où  $\epsilon_\sigma \in \{-1, 1\}$  désigne la signature de  $\sigma$ .

**Proposition 1.1.1** [33] Soit la matrice  $A \in \mathcal{M}_n(\mathbb{R})$ , alors  $A$  est régulière (ou inversible) si et seulement si  $\det(A) \neq 0$ .

**Définition 1.1.2** Pour une matrice  $A$  de  $\mathcal{M}_n(\mathbb{R})$ , le polynôme à variable  $\lambda$ ,  $p_A(\lambda) = \det(\lambda I_n - A)$ , est appelé le polynôme caractéristique de  $A$ . Les racines de  $p_A(\lambda)$  sont dites les valeurs propres de  $A$ .

**Définition 1.1.3** Soient  $\lambda_1, \dots, \lambda_m$  les valeurs propres de la matrice  $A$  d'ordre  $n$ . Alors le polynôme caractéristique peut être représenté par

$$p_A(\lambda) = p_0 + p_1\lambda + \dots + p_{n-1}\lambda^{n-1} + \lambda^n = (\lambda - \lambda_1)^{n_1} \dots (\lambda - \lambda_m)^{n_m},$$

où les  $n_i$  sont des nombres entiers positifs avec  $\sum_{i=1}^m n_i = n$ . Le nombre  $n_i$  est le multiplicateur ou le multiplicateur algébrique de la valeur propre  $\lambda_i$ ,  $i = 1, \dots, m$ .

Une valeur propre est dite simple si son multiplicateur est égal à 1.

Il est important de signaler que le déterminant d'une matrice  $A$  est le produit de ses valeurs propres.

**Définition 1.1.4 (Spectre)** Le spectre de  $A \in \mathcal{M}_n(\mathbb{R})$ ,  $n \geq 1$ , est l'ensemble des valeurs propres de  $A$

$$\sigma(A) = \{\lambda \in \mathbb{R}, (\lambda I_n - A) \text{ est singulière (i.e., n'est pas inversible)}\}.$$

Le rayon spectral de  $A$  est la plus grande en valeur absolue des valeurs propres de  $A$

$$\rho(A) = \max_{\lambda \in \sigma(A)} |\lambda|.$$

**Théorème 1.1.2** Soit  $A$  une matrice de  $\mathcal{M}_{m,n}(\mathbb{R})$ . La norme spectrale de  $A$ , notée  $\|A\|$  est définie par

$$\|A\| = \sqrt{\max_{i=1}^n |\lambda_i|} = \sqrt{\rho(AA^T)}$$

où  $\lambda_i$ ,  $i = 1, \dots, n$ , sont les valeurs propres de la matrice  $A^T A$ .

Si  $A \in \mathbb{S}^n$ , alors  $\|A\| = \rho(A)$ .

Toutes les valeurs propres d'une matrice symétrique  $A \in \mathbb{S}^n$ , notées  $\lambda_i(A)$  sont réelles.

La décomposition spectrale d'une matrice symétrique  $A$  fait l'objet du théorème suivant

**Théorème 1.1.3 (Théorème Spectral)** Soit  $A$  une matrice de  $\mathbb{S}^n$ , alors il existe une matrice  $P$  dans  $\mathcal{M}_n(\mathbb{R})$  (dite de passage) vérifiant  $PP^T = I_n$  (i.e.,  $P$  orthogonale) qui diagonalise  $A$ , c'est-à-dire  $A = PDP^T$  où  $D$  est une matrice diagonale dont les éléments diagonaux sont les valeurs propres de  $A$  (dans ce cas les matrices  $A$  et  $D$  sont dites semblables).

Pour nos besoins, il convient d'écrire les valeurs propres dans l'ordre croissant,

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

**Lemme 1.1.4** [33] Soient  $A$  et  $B$  deux matrices de  $\mathbb{S}^n$ , alors

$$\lambda_{\min}(A + B) \geq \lambda_{\min}(A) + \lambda_{\min}(B),$$

$$\lambda_{\max}(A + B) \leq \lambda_{\max}(A) + \lambda_{\max}(B).$$

**Proposition 1.1.5** [33] Soit la matrice  $A \in \mathcal{M}_n(\mathbb{R})$ , alors on a

1. Si  $A$  est orthogonale, alors les valeurs propres de  $A$  sont 1 ou  $-1$ .
2.  $A$  est régulière (ou inversible) si et seulement si 0 n'est pas une valeur propre de  $A$ .

**Lemme 1.1.6** [33] Soit  $A$  une matrice inversible de  $\mathcal{M}_n(\mathbb{R})$ , alors

1.  $\lambda$  est une valeur propre de  $A$  si et seulement si  $\lambda$  est une valeur propre de  $A^T$ .
2.  $\lambda$  est une valeur propre de  $A$  si et seulement si  $\lambda^{-1}$  est une valeur propre de  $A^{-1}$ .
3.  $\lambda$  est une valeur propre de  $A$  si et seulement si  $\lambda + \beta$  est une valeur propre de  $A + \beta I_n$ .
4.  $\lambda$  est une valeur propre de  $A$  si et seulement si  $\beta\lambda$  est une valeur propre de  $\beta A$ .

## 1.1.2 Trace d'une matrice

**Définition 1.1.5** La trace d'une matrice  $A \in \mathcal{M}_n(\mathbb{R})$  est définie par la somme de ses éléments diagonaux

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

La trace est une application linéaire et pour  $A \in \mathcal{M}_n(\mathbb{R})$ , elle égale à la somme de ses valeurs propres.

**Proposition 1.1.7** [33] Soient  $A$  et  $B$  deux matrices de  $\mathcal{M}_n(\mathbb{R})$ , alors on a

1.  $\text{tr}(\alpha A + \beta B) = \alpha \text{tr}(A) + \beta \text{tr}(B)$ ,  $\forall \alpha, \beta \in \mathbb{R}$ . (Linéarité)
2.  $\text{tr}(A^T) = \text{tr}(A)$ .
3.  $\text{tr}(A^2) \leq \text{tr}(A^T A)$ . (Commutativité)
4.  $\text{tr}(AB) = \text{tr}(BA)$ . (Invariance par permutation)
5. Si  $A$  et  $B$  sont semblables, alors  $\text{tr}(B) = \text{tr}(A)$ .
6.  $\text{tr}(AB) \leq \frac{1}{2} \text{tr}(A^2 + B^2)$ , avec  $A$  et  $B \in \mathbb{S}^n$ .
7.  $\text{tr}(ABC) = \text{tr}(CAB) = \text{tr}(BCA) \neq \text{tr}(ACB)$ .

### 1.1.3 Produit scalaire et norme de Frobenius

L'espace vectoriel  $\mathcal{M}_{m,n}(\mathbb{R})$  est isomorphe à l'espace vectoriel  $\mathbb{R}^{m \times n}$ .

**Définition 1.1.6** *Le produit scalaire entre deux éléments  $A$  et  $B$  de l'espace vectoriel  $\mathcal{M}_{m,n}(\mathbb{R})$  est défini par*

$$\langle A, B \rangle = \text{tr}(B^T A) = \sum_{i=1}^m \sum_{j=1}^n a_{ij} b_{ij} = \text{tr}(A^T B). \quad (1.1)$$

*La norme, dite de Frobenius, associée à ce produit scalaire est défini comme suit*

$$\|A\|_F = \sqrt{\langle A, A \rangle} = \sqrt{\text{tr}(A^T A)}, \quad \forall A \in \mathcal{M}_n(\mathbb{R}).$$

Pour  $A, B \in \mathbb{S}^n$ , qui est isomorphe à l'espace vectoriel  $\mathbb{R}^{\frac{n(n+1)}{2}}$ , on a

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

La proposition suivante donne quelques propriétés sur la norme de Frobenius qui seront utilisés par la suite.

**Proposition 1.1.8** [33] *Pour toutes matrices  $A$  et  $B$  dans  $\mathcal{M}_n(\mathbb{R})$ , on a*

1.  $\|A\|_F = \|A^T\|_F$ .
2.  $\|A\| \leq \|A\|_F \leq \sqrt{n}\|A\|$ .
3.  $\|\lambda A\|_F = |\lambda| \|A\|_F, \quad \forall \lambda \in \mathbb{R}$ .
4.  $\|A + B\|_F \leq \|A\|_F + \|B\|_F$  (Inégalité triangulaire).
5.  $\|AB\|_F \leq \|A\|_F \cdot \|B\|_F$ .
6.  $\|A + B\|_F^2 + \|A - B\|_F^2 = 2(\|A\|_F^2 + \|B\|_F^2)$ . (Identité de Parallélogramme)
7. Si  $\langle A, B \rangle = 0$ , alors  $\|A + B\|_F^2 = \|A - B\|_F^2 = \|A\|_F^2 + \|B\|_F^2$ . (Théorème de Pythagore)
8.  $\langle A, B \rangle = \frac{1}{4}(\|A + B\|_F^2 - \|A - B\|_F^2)$ .
9.  $\langle A, B \rangle = \frac{1}{2}(\|A + B\|_F^2 - \|A\|_F^2 - \|B\|_F^2)$ .

### 1.1.4 Matrices (semi-) définies positives

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.

## Définitions et résultats fondamentaux

**Définition 1.1.7** On dit que la matrice  $A \in \mathbb{S}^n$  est semi-définie positive, et on écrira  $A \geq 0$ , si  $x^T A x \geq 0, \forall x \in \mathbb{R}^n$ . On note par  $\mathbb{S}_+^n$  l'ensemble des matrices symétriques semi-définies positives.

**Définition 1.1.8** On dit que la matrice  $A \in \mathbb{S}^n$  est définie positive, et on écrira  $A > 0$ , si  $x^T A x > 0, \forall x \in \mathbb{R}^n - \{0\}$ . On note par  $\mathbb{S}_{++}^n$  l'ensemble des matrices symétriques définies positives.

On énonce certaines conséquences immédiates de ces définitions qui nous serviront par la suite.

### **Théorème 1.1.9** [17](Caractérisations des matrices semi-définies positives)

Pour  $A \in \mathbb{S}^n$  les propriétés suivantes sont équivalentes

1.  $A \in \mathbb{S}_+^n$
2.  $\lambda_i(A) \geq 0, \forall i = 1, \dots, n$ .
3. Il existe  $C \in \mathcal{M}_{m,n}(\mathbb{R})$  tel que  $A = C^T C$  avec  $rg(C) = rg(A)$ .
4.  $\langle A, B \rangle \geq 0$ , pour toute matrice  $B \in \mathbb{S}_+^n$ .

### **Théorème 1.1.10** [17](Caractérisations des matrices définies positives)

Pour  $A \in \mathbb{S}^n$  les propriétés suivantes sont équivalentes

1.  $A \in \mathbb{S}_{++}^n$ .
2.  $\lambda_i(A) > 0, \forall i = 1, \dots, n$ .
3. Il existe  $C \in \mathcal{M}_n(\mathbb{R})$  avec  $rg(C) = n$  tel que  $A = C^T C$ .
4. Pour une suite arbitraire  $A_i \in \mathbb{S}^i, i = 1, \dots, n$ , de sous-matrices principales de  $A$  : on a  $\det(A_i) > 0$  pour  $i = 1, \dots, n$ , où  $\det(A_i), i = 1, \dots, n$ , sont appelés les mineurs principaux de la matrice  $A$ .

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

**Lemme 1.1.11 (Racine carrée d'une matrice)** Soit  $A \in \mathbb{S}_{++}^n$ . Alors, il existe une unique matrice de  $\mathbb{S}_+^n$  ( $\mathbb{S}_{++}^n$ ) telle que

$$A = BB = B^2,$$

Où  $B$  est appelée la racine carrée de  $A$ , et on note  $B = A^{\frac{1}{2}}$ . De plus, on a  $rg(A) = rg(B)$ .

**Proposition 1.1.12** [33] Soient  $A$  et  $B$  deux matrices de  $\mathcal{M}_n(\mathbb{R})$ , alors

1. Si  $B$  est une matrice inversible, alors

$\lambda$  est une valeur propre de  $A \Leftrightarrow \lambda$  est une valeur propre de  $B^{-1}AB$ .

2. Si  $B$  est une matrice orthogonale, alors

$\lambda$  est une valeur propre de  $A \Leftrightarrow \lambda$  est une valeur propre de  $B^T AB$ .

3. Si  $B$  est une matrice définie positive, alors

$\lambda$  est une valeur propre de  $BA \Leftrightarrow \lambda$  est une valeur propre de  $B^{1/2}AB^{1/2} \in \mathbb{S}^n$ .

Le lemme suivant montre que l'ensemble des matrices symétriques semi-définies positives a un angle d'ouverture égal à  $\frac{\pi}{2}$ .

**Lemme 1.1.13** [17] Soient  $A, B \in \mathbb{S}_+^n$ . Alors  $\langle A, B \rangle \geq 0$  et en plus  $\langle A, B \rangle = 0$  si et seulement si  $AB = 0$ .

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

**Corollaire 1.1.14** [18](Théorème de trace de Fejer)

$A \in \mathbb{S}_+^n$  si et seulement si  $\langle A, B \rangle \geq 0, \forall B \in \mathbb{S}_+^n$

**Proposition 1.1.15** Soit  $B \in \mathcal{M}_n(\mathbb{R})$  une matrice inversible. Alors

1.  $A \in \mathbb{S}^n$  si et seulement si  $B^T AB \in \mathbb{S}^n$ .
2.  $A \in \mathbb{S}_+^n$  si et seulement si  $B^T AB \in \mathbb{S}_+^n$ .
3.  $A \in \mathbb{S}_{++}^n$  si et seulement si  $B^T AB \in \mathbb{S}_{++}^n$ .

### Complément de Schur

**Théorème 1.1.16** [18] Soit  $M$  une matrice de dimension  $(p + q) \times (p + q)$  définie par

$$M = \begin{pmatrix} A & B \\ C & D \end{pmatrix}$$

où les blocs  $A, B, C$  et  $D$  sont des matrices de dimensions respectives  $(p \times p), (p \times q), (q \times p)$  et  $(q \times q)$ .

- Si  $D$  est inversible. Alors, le complément de Schur du bloc  $D$  de la matrice  $M$  est constitué par la matrice de dimension  $(p \times p)$  suivante

$$A - BD^{-1}C.$$

- De même, si  $A$  est inversible, le complément de Schur du bloc  $A$  de la matrice  $M$  est par définition  $D - CA^{-1}B$ .

Maintenant, si on suppose que la matrice  $M$  est symétrique i.e., les matrices  $A$  et  $D$  sont symétriques et  $C = B^T$ , alors la matrice  $M$  peut être exprimée par

$$M = \begin{pmatrix} A & B \\ B^T & D \end{pmatrix} = \begin{pmatrix} I & BD^{-1} \\ 0 & I \end{pmatrix} \begin{pmatrix} A - BD^{-1}B^T & 0 \\ 0 & D \end{pmatrix} \begin{pmatrix} I & BD^{-1} \\ 0 & I \end{pmatrix}^T.$$

La proposition suivante donne des caractérisations des matrices symétriques définies positives par l'utilisation du complément de Schur

**Proposition 1.1.17** [18] *Soit  $M$  une matrice symétrique de la forme*

$$M = \begin{pmatrix} A & B \\ B^T & D \end{pmatrix},$$

*si  $D$  est inversible, alors les propriétés suivantes sont satisfaites*

1.  *$M$  est définie positive si et seulement si  $D$  et  $(A - BD^{-1}B^T)$  sont définies positives.*
2. *Si  $D$  est définie positive, alors  $M$  est semi-définie positive si et seulement si  $(A - BD^{-1}B^T)$  est semi-définie positive.*

Une autre version de cette proposition, par l'utilisation du complément de Schur du bloc  $A$  à la place du complément de Schur du bloc  $D$ .

**Proposition 1.1.18** [18] *Soit  $M$  une matrice symétrique de la forme*

$$M = \begin{pmatrix} A & B \\ B^T & D \end{pmatrix},$$

*si  $A$  est inversible, alors les propriétés suivantes sont satisfaites*

1.  *$M$  est définie positive si et seulement si  $A$  et  $(D - B^T A^{-1}B)$  sont définies positives.*



2. Si  $A$  est définie positive, alors  $M$  est semi-définie positive si et seulement si  $(D - B^T A^{-1} B)$  est semi-définie positive.

**Théorème 1.1.19 (Factorisation de Cholesky)** Pour  $A \in \mathbb{S}_{++}^n$ , il existe une unique matrice triangulaire inférieure et inversible  $L$  telle que  $A = LL^T$ .

**Définition 1.1.9 (Matrice à diagonale (strictement) dominante)**

1. Une matrice  $A \in \mathcal{M}_n(\mathbb{R})$  est à diagonale dominante si

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

2. Une matrice  $A \in \mathcal{M}_n(\mathbb{R})$  est à diagonale strictement dominante si

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

**Théorème 1.1.20** Si  $A \in \mathbb{S}^n$  est à diagonale strictement dominante et si tous les éléments diagonaux sont strictement positifs. Alors  $A$  est définie positive.

**Définition 1.1.10** Soit  $x = (x_1, x_2, \dots, x_n)^T \in \mathbb{R}^n$ . On définit l'opérateur  $\text{Diag}$  par

$$\text{Diag}(x) = \begin{pmatrix} x_1 & 0 & \cdot & \cdot & \cdot & 0 \\ 0 & x_2 & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & 0 \\ 0 & \cdot & \cdot & \cdot & 0 & x_n \end{pmatrix}.$$

## 1.2 Analyse convexe

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

## 1.2.1 Ensembles et fonctions affines

### Ensembles affines

**Définition 1.2.1** *Un sous-ensemble  $A$  de  $\mathbb{R}^n$  est dit affine si*

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

Les ensembles affines élémentaires sont :  $\phi$ ,  $\mathbb{R}^n$ ,  $\{x\}$  avec ( $x \in \mathbb{R}^n$ ) et chaque sous-espace vectoriel de  $\mathbb{R}^n$ .

**Définition 1.2.2** *Soient  $x_1, \dots, x_m$ , des points quelconques de  $\mathbb{R}^n$ , on dit que  $x \in \mathbb{R}^n$  est une combinaison affine de  $x_1, \dots, x_m$ , s'il existe  $(\lambda_1, \lambda_2, \dots, \lambda_m) \in \mathbb{R}^m$  tels que*

$$x = \sum_{i=1}^m \lambda_i x_i \text{ avec } \sum_{i=1}^m \lambda_i = 1.$$

**Proposition 1.2.1** *Un sous-ensemble  $A \subseteq \mathbb{R}^n$  est affine si et seulement s'il contient toutes les combinaisons affines de ses éléments.*

**Définition 1.2.3** [26] *Etant donné  $A \subseteq \mathbb{R}^n$ , alors il existe une partie affine unique  $F \subseteq \mathbb{R}^n$  contenant  $A$  appelée enveloppe affine de  $A$  et notée  $aff(A)$ . Autrement dit  $aff(A)$  est la plus petite partie affine de  $\mathbb{R}^n$  contenant  $A$ .*

$$aff(A) = \cap \{F_A / F_A \supset A \text{ et } F_A \text{ affine} \}.$$

Nous avons par définition  $dim(A) = dim(aff(A))$ .

Si  $A \neq \phi$ , alors  $aff(A) \neq \phi$  (puisque  $A \subseteq aff(A)$ ).

Un sous-ensemble  $A$  de  $\mathbb{R}^n$  est affine si et seulement si  $A = aff(A)$ .

**Théorème 1.2.2** [26] *Soit  $A$  une partie non vide de  $\mathbb{R}^n$ , alors*

$$aff(A) = \{x \in \mathbb{R}^n : x = \sum_{i=1}^m \lambda_i x_i, x_i \in A, \sum_{i=1}^m \lambda_i = 1\}.$$

## Fonctions affines

**Définition 1.2.4** Une fonction  $f$  définie de  $\mathbb{R}^n$  dans  $\mathbb{R}^m$  est dite affine si

$$f[(1 - \lambda)x + \lambda y] = (1 - \lambda)f(x) + \lambda f(y), \forall x, y \in \mathbb{R}^n, \forall \lambda \in \mathbb{R}.$$

**Théorème 1.2.3** Toute fonction affine  $f$  de  $\mathbb{R}^n$  dans  $\mathbb{R}^m$  est de la forme  $f(x) = \mathbf{A}x + b$ , où  $\mathbf{A} \in \mathcal{M}_{m,n}(\mathbb{R})$  et  $b \in \mathbb{R}^m$ .

**Proposition 1.2.4** Une fonction affine  $f$  de  $\mathbb{R}^n$  dans  $\mathbb{R}^m$  est linéaire si et seulement si  $f(0) = 0$ , i.e.,  $b = 0$ .

## 1.2.2 Ensembles et fonctions convexes

### Ensembles convexes

**Définition 1.2.5** Un sous-ensemble  $C$  de  $\mathbb{R}^n$  est dit convexe si

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

Autrement dit, si le segment de droite joignant deux points quelconques  $x, y \in C$   $[x, y] = \{\lambda x + (1 - \lambda)y \mid 0 \leq \lambda \leq 1\}$  est entièrement inclus dans  $C$ .

**Définition 1.2.6** Soient  $x_1, \dots, x_m$ , des points quelconques de  $\mathbb{R}^n$ , on dit que  $x \in \mathbb{R}^n$  est une combinaison convexe de  $x_1, \dots, x_m$ , s'il existe  $(\lambda_1, \lambda_2, \dots, \lambda_m) \in \mathbb{R}_+^m$  tels que

$$x = \sum_{i=1}^m \lambda_i x_i \text{ avec } \sum_{i=1}^m \lambda_i = 1.$$

**Proposition 1.2.5** Un sous-ensemble  $C \subseteq \mathbb{R}^n$  est convexe si et seulement s'il contient toutes les combinaisons convexes de ses éléments.

**Définition 1.2.7** Etant donné  $C \subseteq \mathbb{R}^n$ , l'enveloppe convexe de  $C$ , noté  $\text{Conv}(C)$ , est l'intersection de tous les ensembles convexes contenant  $C$ . Autrement dit  $\text{Conv}(C)$  est l'ensemble de toutes les combinaisons convexes d'éléments de  $C$ .

Il découle de ce qui précède que

- Un sous-ensemble  $C$  de  $\mathbb{R}^n$  est convexe si et seulement si  $C = \text{Conv}(C)$ .
- Tout ensemble affine est convexe. La réciproque est fautive.

**Théorème 1.2.6** *L'intersection d'un nombre quelconque d'ensembles convexes est convexe.*

**Remarque 1.2.1** – *La boule  $B = \mathcal{B}(0, 1)$  unité euclidienne de  $\mathbb{R}^n$  est un convexe fermé et borné.*

–  $\forall a \in \mathbb{R}^n, \forall \varepsilon > 0$ , la boule de centre  $a$  et de rayon  $\varepsilon$  s'écrit

$$\{x : d(x, a) \leq \varepsilon\} = a + \{y : \|y\| \leq \varepsilon\} = a + \varepsilon B.$$

**Définition 1.2.8** *Soit  $C$  un convexe non vide de  $\mathbb{R}^n$  et  $x \in C$ . Le point  $x$  est dit extrémal (ou sommet de  $C$ ) s'il n'est pas à l'intérieur d'un segment de droite contenu dans  $C$ . Autrement dit si*

$$\forall x_1, x_2 \in C, \forall \lambda \in ]0, 1[, \quad x = \lambda x_1 + (1 - \lambda)x_2 \Rightarrow x = x_1 = x_2.$$

Tout point extrémal d'un convexe  $C$  est un point de la frontière de  $C$ .

**Proposition 1.2.7** *Soit  $C$  un convexe non vide de  $\mathbb{R}^n$ , alors  $x$  est un sommet de  $C$  si et seulement si  $C - \{x\}$  est un convexe.*

**Définition 1.2.9 (Intérieur relatif)** *L'intérieur relatif d'un sous-ensemble non vide  $C$  de  $\mathbb{R}^n$ , noté  $ri(C)$  est défini comme l'intérieur de  $C$  vu comme sous-ensemble de  $aff(C)$ .*

$$ri(C) = \{x \in aff(C) / \exists \varepsilon > 0 : (x + \varepsilon B) \cap aff(C) \subseteq C\}.$$

Si  $ri(C) = C$ , alors  $C$  est dit relativement ouvert.

**Théorème 1.2.8 (Caractérisation de  $ri(C)$ )**

*Soit  $C$  un convexe non vide de  $\mathbb{R}^n$ . Alors*

$$ri(C) = \{z / \forall x \in C, \exists \mu > 1 : (1 - \mu)x + \mu z \in C\}.$$

*Autrement dit, tout segment de droite dans  $C$  ayant  $z$  comme extrémité, peut être prolongé au-delà de  $z$  sans sortir de  $C$ .*

## Fonctions convexes

**Définition 1.2.10** *On dit que  $f : C \subset \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$  est convexe si l'une des deux propriétés équivalentes suivantes est vérifiée*

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

$$2. f\left(\sum_{i=1}^m \lambda_i x_i\right) \leq \sum_{i=1}^m \lambda_i f(x_i), \quad \forall m \in \mathbb{N}, \quad \forall \lambda_i \geq 0 \text{ tel que } \sum_{i=1}^m \lambda_i = 1, \quad \forall x_i \in \mathbb{R}^n.$$

Si l'inégalité (1) est stricte pour  $x \neq y$  et  $\lambda \in ]0, 1[$ ,  $f$  est dite strictement convexe.

L'inégalité (2) est appelée inégalité de **Jensen**.

**Théorème 1.2.9** Une combinaison linéaire à coefficients  $\lambda_i$  positifs des fonctions  $f_i$  convexes pour  $i = 1, \dots, m$  est une fonction convexe.

**Proposition 1.2.10** Soit  $f : \mathbb{R} \rightarrow \mathbb{R} \cup \{-\infty, +\infty\}$

1. Si  $f$  est dérivable sur un intervalle  $I \subseteq \mathbb{R}$ , alors  $f$  est convexe sur  $I$  si et seulement si  $f'$  est croissante sur  $I$ .
2. Si  $f$  est deux fois dérivable sur  $I$ , alors  $f$  est convexe sur  $I$  si et seulement si  $f''(t) \geq 0, \forall t \in I$ .

**Théorème 1.2.11** Soit  $f \in C^1(C)$ ,  $C$  est un convexe ouvert de  $\mathbb{R}^n$  (ou  $C = \mathbb{R}^n$ ).

Alors les affirmations suivantes sont équivalentes

1.  $f$  est convexe sur  $C$ .
2.  $f(y) - f(x) \geq \langle \nabla f(x), y - x \rangle, \quad \forall x, y \in C$ .
3.  $\langle \nabla f(y) - \nabla f(x), y - x \rangle \geq 0, \quad \forall x, y \in C$ .

**Théorème 1.2.12** Soit  $f \in C^2(C)$ ,  $C$  est un convexe ouvert de  $\mathbb{R}^n$  (ou  $C = \mathbb{R}^n$ ), Alors

1.  $f$  est convexe sur  $C$  si et seulement si la matrice Hessienne  $H(x) = \nabla^2 f(x)$  est semi-définie positive (i.e.,  $\langle H(x)y, y \rangle \geq 0, \quad \forall x, y \in C$ ).
2.  $f$  est strictement convexe sur  $C$  si et seulement si  $H(x)$  est définie positive (i.e.,  $\langle H(x)y, y \rangle > 0, \quad \forall y \neq 0, \quad \forall x, y \in C$ ).

**Définition 1.2.11** Une fonction  $f$  est dite concave si  $-f$  est convexe.

**Définition 1.2.12** Une fonction  $f$  est dite coercive sur un ensemble convexe  $C$  si

$$\lim_{\|x\| \rightarrow +\infty} f(x) = +\infty, \quad \forall x \in C.$$

**Définition 1.2.13** On appelle fonction barrière toute fonction  $f$  qui vérifie

1.  $f(x)$  finie, si  $x$  appartient à l'intérieur relatif du domaine réalisable (admissible).
2.  $f(x)$  tend vers l'infini quand  $x$  s'approche de la frontière du domaine réalisable.

**Définition 1.2.14** On appelle sous-gradient d'une fonction  $f$  convexe sur  $C$  au point  $x_0 \in C$  tout vecteur  $r = (r_1, r_2, \dots, r_n)^T \in \mathbb{R}^n$  vérifiant

$$f(x) \geq f(x_0) + r^T(x - x_0), \forall x \in C.$$

**Théorème 1.2.13** On appelle sous-différentiel de  $f$  au point  $x_0$ , l'ensemble de tous les sous-gradients de  $f$  en  $x_0$ . On le note  $\partial f(x_0)$ .

**Remarque 1.2.2** Si  $f$  est une fonction différentiable, alors  $\{\nabla f(x_0)\} = \partial f(x_0)$ .

**Lemme 1.2.14 (Direction admissible)**

Soit  $C$  une partie non vide de  $\mathbb{R}^n$ . Un vecteur  $d \neq 0$  de  $\mathbb{R}^n$  est dit direction de  $C$  en  $a \in C$  si

$$\exists \bar{\alpha} > 0 / a + \alpha d \in C, \forall \alpha \in [0, \bar{\alpha}].$$

Si la propriété est vraie  $\forall a \in C$ , on dit que  $d$  est une direction admissible pour  $C$ .

**Définition 1.2.15 (Dérivée directionnelle)**

Soit  $f$  une fonction quelconque de  $\mathbb{R}^n$  dans  $\mathbb{R} \cup \{-\infty, +\infty\}$  et  $x^0$  un point dans  $\mathbb{R}^n$  tel que  $f(x^0)$  est finie. Alors la dérivée directionnelle (**unilatérale**) de  $f$  en  $x^0$  dans la direction  $d \in \mathbb{R}^n$  est définie par

$$f'(x^0, d) = \lim_{\lambda \rightarrow 0^+} \frac{f(x^0 + \lambda d) - f(x^0)}{\lambda} \text{ si elle existe.}$$

( $\pm\infty$  sont des valeurs admises pour la limite).

Notons que

$$-f'(x^0, -d) = \lim_{\lambda \rightarrow 0^-} \frac{f(x^0 + \lambda d) - f(x^0)}{\lambda}.$$

- La dérivée directionnelle est dite **bilatérale** si  $f'(x^0, d) = -f'(x^0, -d)$ .
- Si  $f'(x^0, d)$  existe  $\forall d \in \mathbb{R}^n$ , on dira que  $f$  est directionnellement différentiable en  $x^0$ .
- S'il existe un vecteur constant  $g_0 \in \mathbb{R}^n$  tel que  $f'(x^0, d) = \langle g_0, d \rangle$ ,  $\forall d \in \mathbb{R}^n$ . On dit que  $f$  est **G-différentiable** (ou différentiable au sens de Gateaux) en  $x^0$  et on écrit :  
 $f'(x^0) = g_0$ .

### 1.2.3 Le cône des matrices symétriques semi-définies positives

**Définition 1.2.16** Un ensemble  $C \subset \mathbb{R}^n$  est appelé cône si  $\mathbb{R}_+^* C \subseteq C$ . Autrement dit

$$\forall x \in C, \forall \lambda > 0, \lambda x \in C.$$

C'est la réunion des demi-droites (positives) passant par l'origine, ce dernier peut (ou ne peut pas) appartenir à  $C$ .

- Un cône  $C$  est dit pointé ou saillant si  $C \cap (-C) = \{0\}$  (i.e.,  $C$  ne contenant aucune droite).
- Un cône  $C$  est dit convexe si l'ensemble  $C$  est convexe.

**Proposition 1.2.15** L'ensemble des matrices semi-définies positives  $\mathbb{S}_+^n$  est un cône (convexe), pointu et fermé de pleine dimension dans  $\mathbb{R}^{\frac{n(n+1)}{2}}$ .

La fonction

$$-\ln(\det(X)) = -\ln \prod_{i=1}^n \lambda_i(X) = -\sum_{i=1}^n \ln \lambda_i(X),$$

joue un rôle très important dans les méthodes de points intérieurs.

La fonction  $X \mapsto \det(X)$  est une fonction continue, qui est positive pour les matrices définies positives, et nulle pour les matrices semi-définies positives singulières. La fonction  $X \mapsto -\ln(\det(X))$  tend vers l'infini quand  $X \succ 0$  approche à la frontière du cône des matrices semi-définies positives, elle agit comme une barrière dans les itérations.

**Lemme 1.2.16** La fonction  $-\ln(\det(X))$  est strictement convexe sur l'ensemble des matrices définies positives  $\mathbb{S}_{++}^n$ .

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

**Définition 1.2.17 (La relation d'ordre partiel)** Soient  $A, B \in \mathbb{S}^n$ , alors

1.  $A \succeq B \iff (A - B) \in \mathbb{S}_+^n$ .
2.  $A \succ B \iff (A - B) \in \mathbb{S}_{++}^n$ .

**Proposition 1.2.17** Soit  $C$  un ensemble convexe non vide de  $\mathbb{R}^n$ , et  $a \in C$ , on pose

$$C_\infty(a) = \{d \in \mathbb{R}^n : a + \lambda d \in C, \forall \lambda > 0\},$$

alors  $C_\infty(a)$  est un cône convexe non vide.

**Définition 1.2.18** On appelle cône de récession (ou asymptote) de  $C$  l'ensemble

$$C_\infty = \bigcap_{a \in C} C_\infty(a).$$

Un élément  $d \in C_\infty$ , est appelé direction de récession.

**Proposition 1.2.18** Soit  $C$  un convexe, fermé et non vide de  $\mathbb{R}^n$ , alors

- $C_\infty(a) = C_\infty(b)$ ,  $\forall a, b \in C$ .
- $C$  est borné si et seulement si  $C_\infty = \{0\}$ .

**Définition 1.2.19** Une fonction  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  est dite inf-compacte si l'ensemble non vide des niveaux inférieurs de  $f$  est compact dans  $\mathbb{R}^n$ .

**Théorème 1.2.19** Une fonction  $f$  est dite inf-compacte si et seulement si  $C_\infty = \{0\}$ .

## 1.3 Programmation mathématique

### 1.3.1 Définitions

Dans cette partie, on donne les outils de base d'un problème d'optimisation.

On rappelle certaines définitions élémentaires, les principaux résultats d'existence et d'unicité de la solution optimale et les conditions d'optimalité.

#### Problème d'optimisation

**Définition 1.3.1** Un problème d'optimisation (minimisation) sans contraintes s'écrit comme suit

$$\begin{cases} \min f(x) = f(\bar{x}) \\ x \in \mathbb{R}^n. \end{cases}$$

autrement dit

$$\begin{cases} \text{Trouver } \bar{x} \in \mathbb{R}^n \text{ telle que} \\ f(\bar{x}) \leq f(x), \forall x \in \mathbb{R}^n. \end{cases}$$



Un problème d'optimisation (minimisation) sous contraintes s'écrit sous la forme

$$\begin{cases} \min f(x) = f(\bar{x}) \\ x \in D. \end{cases}$$

autrement dit

$$\begin{cases} \text{Trouver } \bar{x} \in D \text{ telle que} \\ f(\bar{x}) \leq f(x), \forall x \in D. \end{cases}$$

Où  $D \subseteq \mathbb{R}^n$  appelé ensemble des contraintes.

### Programme mathématique

Un programme mathématique noté (PM) est un problème d'optimisation sous contraintes dans  $\mathbb{R}^n$  qui peut s'écrire de la façon suivante

$$(PM) \begin{cases} \min f(x) \\ g_i(x) \leq 0, i = 1, \dots, n, \\ h_j(x) = 0, j = 1, \dots, m, \\ x \in D. \end{cases}$$

Où

- $D$  est une partie de  $\mathbb{R}^n$  et  $x$  un vecteur appelé variable, ses  $n$  composantes sont dites les inconnues du problème (PM).
- La fonction  $f : D \rightarrow \mathbb{R}$  est appelée fonction objectif ou économique.
- Les fonctions  $g_i : D \rightarrow \mathbb{R}, i = 1, \dots, n$ , qui forment des inégalités sont appelées les contraintes inégalités du problème.
- Les fonctions  $h_j : D \rightarrow \mathbb{R}, j = 1, \dots, m$ , qui forment des équations sont appelées les contraintes égalités du problème.
- Si  $D = \mathbb{R}^n$ , on dit que (PM) est un problème d'optimisation sans contraintes.
- Un vecteur  $x$  vérifiant les contraintes de (PM), c'est-à-dire  $g_i(x) \leq 0, (i = 1, \dots, n), h_j(x) = 0, (j = 1, \dots, m)$  et  $x \in D$  est dit solution réalisable de (PM), l'ensemble de ces solutions réalisables forme le domaine de définition (ou domaine réalisable)

de (PM), i.e.,

$$D = \text{dom} f \bigcap_{i=1}^n \text{dom} g_i \bigcap_{j=1}^m \text{dom} h_j.$$

- Une solution réalisable  $x^*$  est dite optimale (ou optimum global) de (PM), si elle minimise la fonction  $f(x)$  sur l'ensemble de toutes les solutions réalisables.
- Un vecteur  $x^0$  est dit optimum local de (PM) si et seulement s'il existe un voisinage  $V$  de  $x^0$  tel que  $x^0$  soit optimum global du problème

$$\left\{ \begin{array}{l} \min f(x) \\ g_i(x) \leq 0, i = 1, \dots, n, \\ h_j(x) = 0, j = 1, \dots, m, \\ x \in D \cap V. \end{array} \right.$$

- la recherche du maximum d'une fonction  $f$  se ramène immédiatement au problème de minimisation de  $-f$ .

**Remarque 1.3.1** – L'ensemble des minimas globaux de (PM) est noté par  $\arg \min_D f(x)$ .

- L'ensemble des minimas locaux de (PM) est noté par  $\text{loc} \min_D f(x)$ .
- On a toujours :  $\arg \min_D f(x) \subseteq \text{loc} \min_D f(x)$ , avec  $D$  un ouvert non vide.

### Classification des problèmes mathématiques

On classifie le problème (PM) à partir de deux propriétés fondamentales à savoir la convexité et la différentiabilité des fonctions du problème. En effet

- (PM) est convexe si les fonctions  $f$  et  $g_i$  sont convexes et les fonctions  $h_j$  sont affines.
- (PM) est différentiable si les fonctions  $f$ ,  $g_i$  et  $h_j$  sont différentiables.

### 1.3.2 Principaux résultats d'existence et d'unicité

**Théorème 1.3.1** (Weierstrass) Si  $f$  est une fonction réelle continue sur  $D \subset \mathbb{R}^n$  compact ( $D$  fermé borné), alors le problème (PM) admet au moins une solution optimale globale  $x^* \in D$ .

**Théorème 1.3.2** Si  $D$  est convexe et  $f$  est strictement convexe alors, il existe au plus une solution optimale de (PM).

**Remarque 1.3.2** L'unicité d'une éventuelle solution optimale est en souvent une conséquence de la stricte convexité de la fonction objectif et de la convexité du domaine réalisable du problème d'optimisation.

**Théorème 1.3.3** Si  $f$  est une fonction inf-compacte, alors le problème (PM) admet au moins une solution optimale.

### 1.3.3 Conditions d'optimalité

**Définition 1.3.2** Le lagrangien d'un programme mathématique (PM) est défini par

$$L(x, \lambda, y) = f(x) + \sum_{i=1}^n \lambda_i g_i(x) + \sum_{j=1}^m y_j h_j(x), \quad \lambda_i \in \mathbb{R}_+, y_j \in \mathbb{R}$$

**Théorème 1.3.4 (Karush-Kuhn-Tucker(KKT))**

Soit  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  une fonction différentiable sur  $D \subset \mathbb{R}^n$ , si  $x^*$  est un minimum local du problème (PM) alors, il existe un vecteur  $y \in \mathbb{R}^m$  et  $\lambda \in \mathbb{R}_+^n$  tel que

$$\begin{cases} \nabla f(x^*) + \sum_{i=1}^n \lambda_i \nabla g_i(x^*) + \sum_{j=1}^m y_j \nabla h_j(x^*) = 0 & (\text{condition d'optimalité}), \\ \lambda_i g_i(x^*) = 0, \quad i = 1, \dots, n & (\text{condition de complémentarité}), \\ h_j(x^*) = 0, \quad j = 1, \dots, m. \end{cases}$$

Les  $\lambda_i$  et les  $y_j$  sont appelés multiplicateurs de Karush-Kuhn-Tucker.

**Définition 1.3.3** On dit qu'un problème de programmation mathématique est convexe s'il consiste à minimiser une fonction convexe (ou maximiser une fonction concave) sur un domaine convexe.

Ainsi le problème (PM) est un problème de programmation convexe (ou simplement un programme convexe) si  $f$  est convexe, les fonctions  $g_i$  ( $i = 1, \dots, n$ ) et  $h_j$  ( $j = 1, \dots, m$ ) sont convexes et  $D \subset \mathbb{R}^n$  est aussi convexe.

**Remarque 1.3.3** Si (PM) est convexe, alors les conditions d'optimalité (KKT) sont à la fois nécessaires et suffisantes.

La propriété fondamentale des programmes convexes apparaît alors dans le résultat suivant

**Théorème 1.3.5** Pour un programme convexe, tout optimum local est un optimum global.

### 1.3.4 Programme linéaire

Un problème de programmation linéaire est un problème de programmation mathématique convexe dans lequel

- L'ensemble des contraintes (ou des solutions réalisables) est défini par un système d'équations (et/ou) d'inéquations linéaires.
- La fonction objectif (ou économique) est linéaire.

On considère des programmes linéaires sous forme standard du type

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

où

$n$  = nombre de variables

$m$  = nombre de contraintes

$A$  = matrice réelle  $m \times n$  (matrice des contraintes)

$c = (c_1, \dots, c_n)^T$  vecteur des coûts

$b = (b_1, \dots, b_m)^T$  vecteur du second membre

$c^T x = \sum_{i=1}^n c_i x_i$  fonction à minimiser (fonction objectif ou fonction économique).

On peut toujours mettre un programme linéaire quelconque sous forme standard en introduisant des variables supplémentaires, dites variables d'écart.

Le dual de (PL) est donné par

$$(DL) \begin{cases} \max b^T y \\ A^T y \leq c \\ y \in \mathbb{R}^m. \end{cases}$$

Pour plus de détails sur l'analyse matricielle voir [18, 33], sur l'analyse convexe voir [26, 46] et sur la programmation mathématique voir [35].

# Chapitre 2

## La programmation semi-définie : Théorie, Applications et Résolution

### 2.1 Les programmes semi-définis

Rappelons que  $\mathcal{S}^n$  est l'ensemble de  $(n \times n)$  matrices symétriques.

Dans ce chapitre, on désigne par

- $\mathcal{S}_+^n$  l'ensemble de  $(n \times n)$  matrices symétriques semi-définies positives

$$\mathcal{S}_+^n = \{A \in \mathcal{S}^n : A \text{ est semi-définie positive (ou } A \geq 0)\}.$$

- $\mathcal{S}_{++}^n$  l'ensemble de  $(n \times n)$  matrices symétriques définies positives

$$\mathcal{S}_{++}^n = \{A \in \mathcal{S}^n : A \text{ est définie positive (ou } A > 0)\} = \text{int}(\mathcal{S}_+^n).$$

La programmation semi-définie (*PSD*) est une généralisation de la programmation linéaire (*PL*). En comparant avec la programmation linéaire standard le vecteur de variables  $x \in \mathbb{R}_+^n$  est remplacé par une matrice variable  $X \in \mathcal{S}_+^n$ . Autrement dit, le cône de l'orthant positif  $x \geq 0$  est remplacé par le cône des matrices semi-définies positives  $X \geq 0$ .

**Définition 2.1.1** *Le problème primal de la programmation semi-définie sous forme standard,*

s'écrit comme suit

$$(PSD) \begin{cases} \min \langle C, X \rangle \\ \langle A_i, X \rangle = b_i, \quad i = 1, \dots, m, \\ X \in \mathbb{S}_+^n. \end{cases} \quad (2.1)$$

où

$b \in \mathbb{R}^m$ ,  $C$  et  $A_i$ ,  $i = 1, \dots, m$ , sont des matrices dans  $\mathbb{S}^n$ .

$$\langle C, X \rangle = \text{trace}(CX) = \sum_{i,j} C_{ij} X_{ij}.$$

**Proposition 2.1.1** *Un programme semi-défini est un programme convexe.*

Pour la démonstration, il est facile de voir que la fonction objectif est une fonction linéaire, donc démontrer qu'un programme semi-défini est convexe, revient à démontrer que l'ensemble des contraintes du problème (PSD) est un ensemble convexe. Pour plus de détails voir [17].

**Remarque 2.1.1** *Si  $X$  est une matrice diagonale, le problème (PSD) se réduit à un problème de programmation linéaire, néanmoins, plusieurs problèmes de programmation non linéaire peuvent se formuler sous la forme (PSD), tel que la programmation quadratique convexe, la programmation combinatoire, les problèmes de coupure maximale dans la théorie des graphes, ainsi que les problèmes de min-max des valeurs propres.*

### 2.1.1 Exemples

Nous présentons ici quelques problèmes de programmation mathématique qui peuvent se convertir en (PSD).

#### Problème de programmation non linéaire

On donne le programme non linéaire suivant

$$\begin{cases} \min(x_1^3 + x_2) \\ x_1^3 x_2 \geq 1, \\ x_1 \geq 0, x_2 \geq 0. \end{cases}$$

Ce problème s'écrit sous la forme d'un problème (PSD) comme suit

$$\begin{cases} \min \langle C, X \rangle \\ \langle A_1, X \rangle = b_1 \\ X \in \mathbb{S}_+^2, \end{cases}$$

avec

$$C = I, \quad X = \begin{pmatrix} x_1^2 & x_3 \\ x_3 & x_2 \end{pmatrix}, \quad A_1 = \begin{pmatrix} 0 & \frac{1}{2} \\ \frac{1}{2} & 0 \end{pmatrix} \text{ et } b_1 = 1.$$

### Problème de coupure maximale

Dans la théorie de graphe, le problème de coupure maximale s'écrit sous la forme suivante

$$(CM) \begin{cases} \max \frac{1}{4} x^T L x \\ x \in \{-1, +1\}^n. \end{cases} \quad (2.1)$$

Où

$x$  est un vecteur de  $\mathbb{R}^n$  qui représente une coupure  $S$  du graphe.

$L = \text{Diag}(Ae) - A$ , est une matrice  $(n \times n)$  laplacienne associée au graphe avec  $e = (1, \dots, 1)^T \in \mathbb{R}^n$  et  $A$  est une matrice  $(n \times n)$  adjacente du graphe.

Le problème (CM) s'écrit sous forme d'un problème semi-défini (PSD) comme suit

$$(PSD) \begin{cases} \max \langle L, X \rangle \\ \langle A_i, X \rangle = \frac{1}{4}, \quad i = 1, \dots, n, \\ X \in \mathbb{S}_+^n. \end{cases} \quad (2.2)$$

Où  $X = xx^T$  et  $A_i = e_i e_i^T$  tel que  $e_i$  est le vecteur de  $\mathbb{R}^n$  dont la  $i^{\text{ème}}$  composante est égale à 1 et les autres sont nulles.

Le problème (PSD) obtenu est un problème continu.

Donc, on passe d'un cas discret (CM) à un cas continu (PSD) ce qui représente un avantage considérable en théorie et même en pratique.

## 2.2 Théorie de la dualité

Avant de définir le dual de (PSD), on a besoin d'introduire quelques notations pour simplifier les résultats.

On définit l'opérateur linéaire  $\mathcal{A} : \mathbb{S}^n \rightarrow \mathbb{R}^m$  avec  $\mathcal{A}X = (\langle A_i, X \rangle)_{i=1}^m$ .

L'adjoint de  $\mathcal{A}$ , par définition est l'opérateur  $\mathcal{A}^* : \mathbb{R}^m \rightarrow \mathbb{S}^n$  vérifiant

$$\langle \mathcal{A}X, y \rangle = \langle X, \mathcal{A}^*y \rangle \text{ pour tout } X \in \mathbb{S}^n \text{ et } y \in \mathbb{R}^m.$$

Ainsi

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

on obtient

$$\mathcal{A}^*y = \sum_{i=1}^m y_i A_i.$$

Pour obtenir le problème dual de (PSD), on considère la fonction lagrangienne associée à ce dernier comme suit

$$L(X, y) = \langle C, X \rangle + \langle b - \mathcal{A}X, y \rangle, X \in \mathbb{S}_+^n, y \in \mathbb{R}^m$$

et on calcule la fonction duale associée  $H(y)$

$$\begin{aligned} H(y) &= \min L(X, y) \\ &= \min_{X \in \mathbb{S}_+^n} \langle C, X \rangle + \langle b - \mathcal{A}X, y \rangle \\ &= \min_{X \in \mathbb{S}_+^n} \langle C, X \rangle + \sum_{i=1}^m (b_i - \langle A_i, X \rangle) y_i \\ &= \min_{X \in \mathbb{S}_+^n} \langle C - \sum_{i=1}^m y_i A_i, X \rangle + \sum_{i=1}^m b_i y_i \\ &= \langle b, y \rangle + \min_{X \in \mathbb{S}_+^n} [\langle C - \mathcal{A}^*y, X \rangle]. \end{aligned}$$

Il est facile de voir que

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



D'où

$$\max_{y \in \mathbb{R}^m} H(y) = \begin{cases} \max \langle b, y \rangle & \text{si } C - \mathcal{A}^* y \geq 0 \\ -\infty & \text{sinon.} \end{cases}$$

Donc le problème dual (*SDD*) de (*PSD*) s'écrit sous la forme suivante

$$(SDD) \begin{cases} \max \langle b, y \rangle \\ \mathcal{A}^* y + S = C \\ y \in \mathbb{R}^m, S \geq 0. \end{cases} \quad (2.3)$$

**Théorème 2.2.1** (*Dualité faible*)

Si  $X$  et  $(y, S)$  sont des solutions réalisables de (*PSD*) et (*SDD*) respectivement, alors on a toujours

$$\langle C, X \rangle - b^T y = \langle S, X \rangle \geq 0.$$

Cette différence est appelée *saut de dualité*.

**Preuve.** On a

$$\begin{aligned} \langle C, X \rangle - b^T y &= \langle C, X \rangle - \sum_{i=1}^m b_i y_i \\ &= \langle C, X \rangle - \sum_{i=1}^m y_i \langle A_i, X \rangle \\ &= \langle C - \sum_{i=1}^m y_i A_i, X \rangle = \langle S, X \rangle. \end{aligned}$$

Puisque les matrices  $S$  et  $X$  sont toutes les deux semi-définies positives et d'après le Lemme 1.1.13, on déduit que  $\text{tr}(XS) \geq 0$ .

Ce qui donne

$$\langle C, X \rangle - b^T y = \langle S, X \rangle \geq 0. \quad \blacksquare$$

Contrairement à la programmation linéaire, il n'est pas toujours vrai que l'optimalité de deux problèmes (*PSD*) et (*SDD*) implique que  $\langle S, X \rangle = 0$ . Pour illustrer ce phénomène, on donne l'exemple de Vandenberghe et Boyd [53]

**Exemple 2.2.1** (Contre exemple)

Considérons le problème semi-défini suivant

$$(PSD) \begin{cases} \min \langle C, X \rangle \\ \langle A_i, X \rangle = b_i, \quad i = 1, \dots, 4, \\ X \in \mathbb{S}_+^3. \end{cases} \quad (2.3)$$

où

$$C = \begin{pmatrix} 0 & \frac{1}{2} & 0 \\ \frac{1}{2} & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad A_1 = \begin{pmatrix} 0 & -\frac{1}{2} & 0 \\ -\frac{1}{2} & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad A_2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix},$$

$$A_3 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \quad A_4 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \quad \text{et} \quad b = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}.$$

Qui est équivalent à

$$\begin{cases} \min & x_{12} \\ & X = \begin{pmatrix} 0 & x_{12} & 0 \\ x_{12} & x_{22} & 0 \\ 0 & 0 & 1 + x_{12} \end{pmatrix} \geq 0. \end{cases}$$

En écrivant le problème dual associé, on obtient

$$\begin{cases} \max & y_1 \\ & S = C - y_1 A_1 - y_2 A_2 - y_3 A_3 - y_4 A_4 \geq 0. \end{cases}$$

Le problème dual peut s'écrire sous la forme suivante

$$\begin{cases} \max & y_1 \\ & S = \begin{pmatrix} -y_2 & \frac{1+y_1}{2} & -y_3 \\ \frac{1+y_1}{2} & 0 & -y_4 \\ -y_3 & -y_4 & -y_1 \end{pmatrix} \geq 0. \end{cases}$$

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 pour le problème dual, on obtient  $s_{12} = 0$  car  $s_{22} = 0$ , i.e.  $y_1 = -1$ . Le saut de dualité à l'optimalité est égal à 1 qui est différent de 0.

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

**Définition 2.2.1** (Faisabilité stricte)

Un point  $X$  est dit strictement réalisable pour (PSD) s'il est réalisable pour (PSD) et vérifie  $X > 0$ .

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

**Théorème 2.2.2** [17] (Dualité forte)

Supposons qu'il existe une solution strictement réalisable  $(y_0, S_0)$  pour (SDD). Soient

$$p^* = \min\{\langle C, X \rangle : \mathcal{A}X = b, X \geq 0\},$$

et

$$q^* = \max\{\langle b, y \rangle : \mathcal{A}^*y + S = C, S \geq 0\},$$

alors  $p^* = q^*$  et si  $p^*$  est une valeur finie, elle est atteinte pour une certaine matrice  $X \in \{X \geq 0 : \mathcal{A}X = b\}$ .

On donne dans le corollaire suivant le résultat de toutes les combinaisons primales-duales possibles.

**Corollaire 2.2.3** [17]

Soient  $p^*$  et  $q^*$  définis comme dans le Théorème 2.2.2

1. Si (PSD) est strictement réalisable avec  $p^*$  finie, alors  $p^* = q^*$  et cette valeur est atteinte pour (SDD).
2. Si (SDD) est strictement réalisable avec  $q^*$  finie, alors  $p^* = q^*$  et cette valeur est atteinte pour (PSD).
3. Si (PSD) et (SDD) sont strictement réalisables, alors  $p^* = q^*$  et ces deux valeurs sont atteintes pour les deux problèmes.

**2.2.1 Exemples pathologiques**

On présente dans cette partie deux exemples de problème de programmation semi-définie de Vandenberghe et Boyd [53], où on montre comment la dualité forte n'est pas atteinte.

**Minimum dual n'est pas atteint**

Considérons le problème de programmation semi-définie suivant

$$p^* = \begin{cases} \max \left\langle \begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix}, X \right\rangle \\ \left\langle \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, X \right\rangle = 1, \\ \left\langle \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}, X \right\rangle = 0, X \geq 0, \end{cases}$$

et son dual est donné par

$$q^* = \begin{cases} \min y_1 \\ S = y_1 \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + y_2 \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} - \begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix} = \begin{pmatrix} y_1 & 1 \\ 1 & y_2 \end{pmatrix} \geq 0. \end{cases}$$

Dans cet exemple,  $p^* = q^* = 0$  et le maximum est atteint dans le primal mais le minimum n'est pas atteint dans le dual. En effet, le primal n'est pas strictement réalisable (puisque  $x_{22} = 0$  pour toute solution réalisable).

**Saut de dualité positif**

A l'optimalité, il peut y avoir un saut de dualité positif entre le primal et le dual des programmes mathématiques. Considérons à titre d'exemple le problème de programmation semi-définie primal avec les matrices de données suivantes

$$C = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}, A_1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, A_2 = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 2 \end{pmatrix},$$

et  $b_1 = 0, b_2 = 2$ . Donc le problème (PSD) s'écrit

$$p^* = \min \{x_{33} : x_{11} = 0, x_{12} + x_{33} = 1, X \geq 0\},$$

et son dual est

$$q^* = \max \left\{ 2y_2 : S = y_1 A_1 + y_2 A_2 - C = \begin{pmatrix} -y_1 & -y_2 & 0 \\ -y_2 & 0 & 0 \\ 0 & 0 & 1 - 2y_2 \end{pmatrix} \geq 0 \right\}.$$

La solution optimale primale est

$$X^* = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

donc la valeur optimale primale est égale à  $p^* = 1$ ,

De même pour le dual, on a la valeur optimale duale est égale à  $q^* = 0$ , pour la solution optimale duale  $y^* = (0, 0)^T$ . D'où, à l'optimalité il y a un saut de dualité positif :  $p^* - q^* = 1$ .

### 2.2.2 Relations primales-duales pour la programmation semi-définie

Comme pour la programmation linéaire, le problème (PSD) peut s'écrire sous plusieurs formes, le tableau suivant présente les différents types de problèmes duaux

correspondants

Minimisation	Maximisation
Variables	Contraintes
matrice ou scalaire $\geq 0$	matrice ou scalaire $\leq$
matrice ou scalaire $\leq 0$	matrice ou scalaire $\geq$
matrice $\geq 0$	matrice $\leq$
matrice $\leq 0$	matrice $\geq$
matrice ou scalaire non astreint	matrice ou scalaire $=$
Contraintes	Variables
matrice ou scalaire $\geq$	matrice ou scalaire $\geq 0$
matrice ou scalaire $\leq$	matrice ou scalaire $\leq 0$
matrice $\geq$	matrice $\geq 0$
matrice $\leq$	matrice $\leq 0$
matrice ou scalaire $=$	matrice ou scalaire non astreint

## 2.3 Domaines d'applications en PSD

L'intérêt pour (*PSD*) s'est encore accru ces dernières années lorsque de nombreuses applications ont été identifiées dans des domaines variés tels que le contrôle, les statistiques, la finance, la localisation, l'optimisation robuste, l'ingénierie, etc. Pour ne citer que quelques exemples, les problèmes (*PSD*) apparaissent naturellement en contrôle (stabilité), analyse des séries temporelles (complétion de covariances), graphes (mélange de chaînes de Markov), etc. Quelques unes de ces applications sont présentées dans la partie suivante.

### 2.3.1 Optimisation des valeurs propres

#### a) Recherche des valeurs propres extrêmes

Il s'agit de plus anciens problèmes traités à l'aide de la programmation semi-définie.

Considérons  $A$  une matrice symétrique réelle, le problème de recherche de la valeur propre minimale d'une matrice symétrique peut se formuler comme suit

$$\lambda_{\min}(A) = \min_{\|x\|=1} x^T A x = \min_{x \in \mathbb{R}^n, x \neq 0} \frac{x^T A x}{\|x\|^2}. \quad (VE)$$

Posons  $X = x x^T$ , alors

$$\text{Tr}(X) = \sum_{i=1}^n x_i^2 = \|x\|^2 = 1.$$

Le problème (VE) devient alors

$$\lambda_{\min}(A) = \begin{cases} \min \langle A, X \rangle \\ \text{Tr}(X) = 1 \\ X \geq 0. \end{cases}$$

De même, le problème de la plus grande valeur propre peut s'écrire comme suit

$$\lambda_{\max}(A) = \max_{\|x\|=1} x^T A x = \max_{x \in \mathbb{R}^n, x \neq 0} \frac{x^T A x}{\|x\|^2},$$

et sa formulation en (PSD) est

$$\lambda_{\max}(A) = \begin{cases} \max \langle A, X \rangle \\ \text{Tr}(X) = 1 \\ X \geq 0. \end{cases}$$

#### b) Problèmes de Min-Max des valeurs propres

On suppose que la matrice symétrique  $A(x)$  dépend affinement d'un paramètre  $x \in \mathbb{R}^k$ , i.e.,  $A(x) = A_0 + x_1 A_1 + \dots + x_k A_k$ , avec  $A_i \in \mathbb{S}^n$ ,  $i = 0, \dots, k$ .

Le problème d'optimisation est de minimiser la valeur propre maximale de  $A(x)$  sur  $\mathbb{R}^k$ , i.e., il s'agit de résoudre le problème convexe et non différentiable ci-dessous

$$\min_{x \in \mathbb{R}^k} \lambda_{\max}(A(x)). \quad (2.4)$$

Ce problème peut se transformer en un problème de (PSD). Pour cela, on a besoin de la propriété suivante

**Propriété** Soient  $M \in \mathbb{S}^n$  et  $t \in \mathbb{R}$ . Alors

$$M \leq tI_n \iff \lambda_{\max}(M) \leq t.$$

**Preuve.** D'après le Théorème 1.1.3 et comme  $M \in \mathbb{S}^n$ . On a

$$tI_n - M = tI_n - PDP^T = P(tI_n - D)P^T.$$

En utilisant la Proposition 1.1.15 et la caractérisation des matrices symétriques semi-définies positives (Théorème 1.1.9), il vient

$$\begin{aligned} tI_n \geq M &\iff tI_n - M \geq 0 \\ &\iff tI_n - D \geq 0 \\ &\iff \lambda_{\min}(tI_n - D) \geq 0 \\ &\iff t \geq \lambda_{\max}(M). \end{aligned}$$

■

Revenons au problème (2.4) et introduisons une variable réelle  $t$  de façon à obtenir une fonction objectif linéaire i.e., il s'agit de résoudre le problème suivant

$$\begin{cases} \min t \\ \lambda_{\max}(A(x))_{x \in \mathbb{R}^k} \leq t. \end{cases}$$

Ce dernier est équivalent au problème (PSD) suivant

$$\begin{cases} \min t \\ tI_n - A(x) \geq 0 \\ x \in \mathbb{R}^k. \end{cases}$$

Les problèmes de ce type surviennent dans la théorie de contrôle, l'optimisation structurelle, la théorie de graphe et l'optimisation combinatoire. Pour de plus amples développements, voir [34, 43, 56].



### c) Minimisation de la somme des $r$ plus grandes valeurs propres

Pour minimiser la somme des  $r$  plus grandes valeurs propres de la matrice symétrique  $A(x)$  dépend affinement de  $x$  i.e.,  $A(x) = A_0 + x_1A_1 + \dots + x_kA_k$ , avec  $A_i \in \mathbb{S}^n$ ,  $i = 0, \dots, k$ , on résout le problème de programmation semi-définie suivant

$$\begin{cases} \min rt + \text{tr}(X) \\ tI_n + X - A(x) \geq 0 \\ X \geq 0. \end{cases}$$

Où  $X \in \mathbb{S}^n$  et  $t \in \mathbb{R}$ .

Pour la démonstration et plus de détails voir [42].

### d) Minimisation de la norme spectrale

Pour minimiser la norme  $\|A(x)\|$  de la matrice  $A(x) = A_0 + x_1A_1 + \dots + x_kA_k \in \mathcal{M}_n(\mathbb{R})$ , où les matrices  $A_i$ ,  $i = 0, \dots, k$ , sont quelconques (i.e., pas nécessairement symétriques), on résout le problème (PSD) suivant

$$\begin{cases} \|A(x)\| = \min t \\ \begin{pmatrix} tI_n & A(x) \\ A(x)^T & tI_n \end{pmatrix} \geq 0. \end{cases}$$

Avec les variables  $x \in \mathbb{R}^k$  et  $t \in \mathbb{R}$ .

On utilise ici le théorème du complément de Schur.

## 2.3.2 Programmation quadratique avec contraintes quadratiques

Soit la contrainte  $(Ax + b)^T(Ax + b) - c^T x - d \leq 0$ , avec  $x \in \mathbb{R}^k$ . Il s'agit en fait de la forme générale d'une contrainte quadratique convexe. On montre qu'elle peut s'écrire comme suit

$$\begin{pmatrix} I & Ax + b \\ (Ax + b)^T & c^T x + d \end{pmatrix} \geq 0.$$

Le côté gauche de cette inégalité dépend affinement du vecteur  $x$ , c'est-à-dire l'inégalité peut être exprimée par

$$F(x) = F_0 + x_1 F_1 + \dots + x_k F_k \geq 0,$$

où

$$F_0 = \begin{pmatrix} I & b \\ b^T & d \end{pmatrix}, \quad F_i = \begin{pmatrix} 0 & a_i \\ a_i^T & c_i^T \end{pmatrix}, \quad i = 1, \dots, k,$$

avec  $A = [a_1, \dots, a_k]$ . Donc un programme quadratique convexe avec contraintes quadratiques de type

$$\begin{cases} \min f_0(x) \\ f_i(x) \leq 0, \quad i = 1, \dots, L, \end{cases}$$

où  $f_i$ ,  $i = 0, \dots, L$ , sont des fonctions quadratiques convexes

$$f_i = (A_i x + b)^T (A_i x + b) - c_i^T x - d_i,$$

peut s'écrire sous la forme du programme semi-défini suivant

$$\begin{cases} \min t \\ \begin{pmatrix} I & A_0 x + b \\ (A_0 x + b)^T & c_0^T + d_0 + t \end{pmatrix} \geq 0, \\ \begin{pmatrix} I & A_i x + b \\ (A_i x + b)^T & c_i^T + d_i \end{pmatrix} \geq 0, \quad i = 1, \dots, L. \end{cases}$$

Avec les variables  $x \in \mathbb{R}^k$  et  $t \in \mathbb{R}$ . Ce programme a une dimension de  $m \times n$ , telle que  $m = k + 1$  et  $n = n_0 + \dots + n_L$ , où  $A_i \in \mathbb{R}^{n_i \times k}$ .

### 2.3.3 Approximation logarithmique de Tchebychev

Supposons qu'on cherche à résoudre approximativement le système d'équations  $a_i^T x = b_i$  pour  $i = 1, \dots, m$ .

Le terme approximativement a été utilisé par ce que la résolution exacte est très difficile (trop de contraintes). On peut dès lors choisir plusieurs critères pour déterminer la

meilleure solution. Le plus courant est la minimisation de la somme des carrés des écarts  $(a_i^T x - b_i)$  ; cependant, dans l'approximation de Chebychev, on minimise la norme du résidu  $l_\infty$ , i.e., on minimise l'écart maximal en valeur absolue, selon

$$\min \max_i |a_i^T x - b_i|. \quad (2.5)$$

Ce problème peut se mettre sous la forme du programme linéaire suivant

$$\begin{cases} \min t \\ -t \leq a_i^T x - b_i \leq t, \quad i = 1, \dots, m, \end{cases}$$

où  $x$  est une variable et  $t$  est une variable auxiliaire.

Dans certaines applications les  $b_i$  sont des quantités s'exprimant dans une unité de puissance ou d'intensité et généralement de logarithme. Dans ce cas, il est préférable d'un point de vue purement physique de minimiser le maximum des écarts entre les logarithmes de  $a_i^T x$  et  $b_i$ , ce qui se traduit par

$$\min \max_i |\log a_i^T x - \log b_i|. \quad (2.6)$$

On suppose que  $b_i > 0$  et  $a_i^T x > 0, \forall i = 1, \dots, m$ . Ce problème est appelé problème d'approximation logarithmique de Chebychev et peut être transformé en un programme semi-défini. On transforme alors les contraintes en

$$-\log t \leq \log \frac{a_i^T x}{b_i} \leq \log t, \quad \text{d'où } \frac{1}{t} \leq \frac{a_i^T x}{b_i} \leq t.$$

Le problème (2.6) est équivalent au problème semi-défini suivant

$$\begin{cases} \min t \\ \frac{1}{t} \leq \frac{a_i^T x}{b_i} \leq t, \quad i = 1, \dots, m, \end{cases}$$

où

$$\begin{cases} \min t \\ \begin{pmatrix} t - \frac{a_i^T x}{b_i} & 0 & 0 \\ 0 & \frac{a_i^T x}{b_i} & 1 \\ 0 & 1 & t \end{pmatrix} \geq 0, \quad i = 1, \dots, m. \end{cases}$$

Le premier terme de la diagonale est responsable de l'inégalité  $\frac{a_i^T x}{b_i} \leq t$ , tandis que  $\frac{a_i^T x}{b_i} \geq \frac{1}{t}$  se traduit par la matrice  $2 \times 2$  inférieure droite. En effet, pour une matrice  $2 \times 2$ , le fait d'être semi-définie positive implique la non-négativité du déterminant, d'où on déduit  $\frac{a_i^T x}{b_i} t - 1 \geq 0$  et finalement l'inégalité recherchée.

Donc, cet exemple illustre deux points importants. Premièrement, il montre que la programmation semi-définie inclut beaucoup de problèmes d'optimisation qui ne semblent pas pour la première vue. Deuxièmement, il montre que le problème est plus général qu'un programme linéaire, malgré la proche analogie.

### 2.3.4 Problèmes géométriques en formes quadratiques

Plusieurs problèmes géométriques impliquant des fonctions quadratiques peuvent être exprimés sous forme des programmes semi-définis. On présente ici un exemple simple.

Supposons, qu'on donne  $k$  ellipsoïdes  $\mathcal{E}_1, \dots, \mathcal{E}_k$  décrivent comme l'ensemble des sous-niveaux des fonctions quadratiques

$$f_i(x) = x^T A_i x + 2b_i^T x + c_i, \quad i = 1, \dots, k,$$

i.e.,  $\mathcal{E}_i = \{x / f_i(x) \leq 0\}$ . Le but est de trouver la plus petite sphère contenant l'ensemble des  $k$  ellipsoïdes donnés. La condition qu'un ellipsoïde contient un autre, peut être exprimée en termes d'une inégalité matricielle.

Supposons que les ellipsoïdes  $\mathcal{E} = \{x : f(x) \leq 0\}$  et  $\tilde{\mathcal{E}} = \{x : \tilde{f}(x) \leq 0\}$ , avec

$$f(x) = x^T A x + 2b^T x + c, \quad \tilde{f}(x) = x^T \tilde{A} x + 2\tilde{b}^T x + \tilde{c},$$

ont un intérieur non vide. Alors, on peut montrer que  $\mathcal{E}$  contient  $\tilde{\mathcal{E}}$  si et seulement s'il existe un  $\tau \geq 0$  tel que

$$\begin{pmatrix} A & b \\ b^T & c \end{pmatrix} \leq \tau \begin{pmatrix} \tilde{A} & \tilde{b} \\ \tilde{b}^T & \tilde{c} \end{pmatrix}.$$

Revenons à notre problème, considérons la sphère  $S$  représentée par

$$f(x) = x^T x - 2x_c^T x + \gamma \leq 0.$$

$S$  contient les ellipsoïdes  $\mathcal{E}_1, \dots, \mathcal{E}_k$  si et seulement s'il existe  $\tau_1, \dots, \tau_k$  positifs tels que

$$\begin{pmatrix} I & -x_c \\ -x_c^T & \gamma \end{pmatrix} \preceq \tau_i \begin{pmatrix} A_i & b_i \\ b_i^T & c_i \end{pmatrix}, \quad i = 1, \dots, k.$$

Notre but est de minimiser le rayon de la sphère  $S$  qui est  $r = \sqrt{x_c^T x_c - \gamma}$ . Pour cela, on exprime la condition  $r^2 \leq t$  par l'inégalité matricielle suivante

$$\begin{pmatrix} I & x_c \\ x_c^T & t + \gamma \end{pmatrix} \succeq 0,$$

et on minimise la variable  $t$ .

Donc, la recherche de la plus petite sphère qui contient les ellipsoïdes  $\mathcal{E}_1, \dots, \mathcal{E}_k$ , revient à la résolution du problème de programmation semi-défini suivant

$$\begin{cases} \min t \\ \begin{pmatrix} I & -x_c \\ -x_c^T & \gamma \end{pmatrix} \preceq \tau_i \begin{pmatrix} A_i & b_i \\ b_i^T & c_i \end{pmatrix}, \quad i = 1, \dots, k, \quad \tau_i \geq 0, \\ \begin{pmatrix} I & x_c \\ x_c^T & t + \gamma \end{pmatrix} \succeq 0, \end{cases}$$

où  $x_c, \tau_1, \dots, \tau_k$  et  $t$  sont les variables du problème.

Cet exemple montre encore l'ampleur des problèmes qui peuvent être reformulés en programmation semi-définie. Et que cette reformulation n'est pas simple.

### 2.3.5 Optimisation quadratique non convexe

La programmation semi-définie joue un rôle très utile dans l'optimisation non convexe ou combinatoire. Considérons par exemple, le problème d'optimisation qua-

dratique suivant

$$\begin{cases} \min f_0(x) \\ f_i(x) \leq 0, \quad i = 1, \dots, L, \end{cases} \quad (2.7)$$

où  $f_i(x) = x^T A_i x + 2b_i^T x + c_i, i = 0, \dots, L$ . Les matrices  $A_i$  peuvent être indéfinies, dans ce cas le problème (2.7) est un problème d'optimisation non convexe très difficile à résoudre. Par exemple, il inclut tous les problèmes d'optimisation de fonction objectif polynomiale et contraintes polynomiales (voir [42, 47]).

Dans la pratique, par exemple l'algorithme de branch-and-bound, il est important d'avoir des bonnes et moins couteuses estimations inférieures de la valeur optimale de (2.7) qui soient calculables efficacement. Shor [47] et Poljak et al [44] ont proposé le calcul des estimations inférieures par la résolution du programme semi-défini suivant

$$\begin{cases} \max t \\ \begin{pmatrix} A_0 & b_0 \\ b_0^T & c_0 - t \end{pmatrix} + \tau_1 \begin{pmatrix} A_1 & b_1 \\ b_1^T & c_1 \end{pmatrix} + \dots + \tau_L \begin{pmatrix} A_L & b_L \\ b_L^T & c_L \end{pmatrix} \geq 0, \\ \tau_i \geq 0, \quad i = 1, \dots, L. \end{cases} \quad (2.8)$$

Où  $t$  et  $\tau_i$  sont les variables du problème. On peut vérifier facilement que ce programme semi-défini (2.8) offre des estimations inférieures pour (2.7).

Supposons que  $x$  est réalisable (satisfait les contraintes) pour le problème non convexe (2.7), i.e.,

$$f_i(x) = \begin{pmatrix} x \\ 1 \end{pmatrix}^T \begin{pmatrix} A_i & b_i \\ b_i^T & c_i \end{pmatrix} \begin{pmatrix} x \\ 1 \end{pmatrix} \leq 0,$$

Alors, si les variables  $t$  et  $\tau_1, \dots, \tau_L$ , satisfont les contraintes du programme semi-défini (2.8), on déduit que

$$\begin{aligned} 0 &\leq \begin{pmatrix} x \\ 1 \end{pmatrix}^T \left[ \begin{pmatrix} A_0 & b_0 \\ b_0^T & c_0 - t \end{pmatrix} + \tau_1 \begin{pmatrix} A_1 & b_1 \\ b_1^T & c_1 \end{pmatrix} + \dots + \tau_L \begin{pmatrix} A_L & b_L \\ b_L^T & c_L \end{pmatrix} \right] \begin{pmatrix} x \\ 1 \end{pmatrix} \\ &= f_0(x) - t + \tau_1 f_1(x) + \dots + \tau_L f_L(x) \\ &\leq f_0(x) - t. \end{aligned}$$

Par conséquent  $t \leq f_0(x)$  pour tout point  $x$  réalisable dans (2.7). Le problème (2.8) peut être aussi obtenu via la dualité du lagrangien ; pour plus de détails, voir Shor [47] ou Poljak et al [44].

Par ailleurs, le problème dual associé au programme semi-défini (2.8) est donné par

$$\begin{cases} \min tr(A_0X) + 2b_0^T x + c_0 \\ tr(A_iX) + 2b_i^T x + c_i \leq 0, \quad i = 1, \dots, L \\ \begin{pmatrix} X & x \\ x^T & 1 \end{pmatrix} \geq 0, \end{cases} \quad (2.9)$$

où les variables sont  $X \in \mathbb{S}^n$  et  $x \in \mathbb{R}^n$ .

On voit que les deux problèmes (2.8) et (2.9) donne la même estimation.

Notons que la contrainte

$$\begin{pmatrix} X & x \\ x^T & 1 \end{pmatrix} \geq 0, \quad (2.10)$$

est équivalente à  $X \geq xx^T$ . Par conséquent, le problème de programmation semi-défini (2.9) peut directement se considéré comme une relaxation du problème original (2.7), qui peut s'écrire comme suit

$$\begin{cases} \min tr(A_0X) + 2b_0^T x + c_0 \\ tr(A_iX) + 2b_i^T x + c_i \leq 0, \quad i = 1, \dots, L \\ X = xx^T. \end{cases} \quad (2.11)$$

La seule différence entre les deux problèmes (2.11) et (2.9) est le remplacement de la contrainte non convexe  $X = xx^T$  par la relaxation de la contrainte convexe  $X \geq xx^T$ .

Il est intéressant de noter que la formulation (2.11) est équivalente au problème (2.7).

Par exemple, nous considérons le programme  $(-1, 1)$ -quadratique

$$\begin{cases} \min x^T Ax + 2b^T x \\ x_i^2 = 1, \quad i = 1, \dots, k, \end{cases} \quad (2.12)$$

qui est un problème non convexe difficile. La contrainte entière  $x_i \in \{-1, 1\}$  peut être écrite comme une contrainte d'égalité quadratique  $x_i^2 = 1$  ou l'équivalence de deux inégalités  $x_i^2 \leq 1$  et  $x_i^2 \geq 1$ . En appliquant (2.9), on trouve que le programme semi-défini

$$\begin{cases} \min tr(AX) + 2b^T x \\ X_{ii} = 1, \quad i = 1, \dots, k, \\ \begin{pmatrix} X & x \\ x^T & 1 \end{pmatrix} \succeq 0, \end{cases} \quad (2.13)$$

donne une estimation inférieure de la valeur optimale du problème (2.12), avec  $X = X^T$  et  $x$  sont des variables.

## 2.4 Méthode de points intérieurs pour résoudre (PSD)

La méthode de points intérieurs est l'une des méthodes les plus utilisées et les plus efficaces pour la résolution des problèmes (PSD), elle est relativement nouvelle et s'apparente à la méthode projective de Karmarkar pour la programmation linéaire. Derrière le terme points intérieurs découlent trois différents types de méthodes

1. Les méthodes affines.
2. Les méthodes de réduction du potentiel.
3. Les méthodes de trajectoire centrale.

### 2.4.1 Méthodes affines

L'idée remonte à Dikin (1967), Puis reprise et développée par plusieurs chercheurs au milieu des années 80. Il s'agit pratiquement de l'algorithme de Karmarkar sans fonction potentielle et sans transformation projective, on utilise une transformation affine puis on remplace la contrainte de non négativité par un ellipsoïde qui contient le nouvel itéré.



L'algorithme correspondant à cette méthode est d'une structure simple, malheureusement, il n'est pas facile de démontrer la polynomialité. À l'égard de cette dernière, Dikin [12], en 1967 a prouvé la convergence de la méthode affine primale sous des hypothèses de non dégénérescence. En 1990 Monteiro et al [36] ont démontré que la méthode affine primale-duale est de complexité polynomiale ( $O(nL^2)$ ), où  $L$  représente le nombre de bits requis pour stocker (et traiter) les données, par contre, la convergence pour les méthodes affines primales ou duales est restée une question ouverte.

### 2.4.2 Méthodes de réduction du potentiel

Ces méthodes sont le fruit direct d'une grande partie des études acharnées menées par plusieurs chercheurs vers la fin des années 80. En 1995, Alizadeh [2] a proposé une méthode projective primale-duale pour résoudre les problèmes (PSD). En 2003, Benterki et al [6] ont fait l'étude théorique et numérique de l'algorithme proposé par Alizadeh [2]. En 2007, Benterki et al [7] ont proposé une méthode primale réalisable de réduction du potentiel. L'algorithme de ces méthodes est basé sur ces deux principaux éléments

1. La transformation projective  $T_k$ .
2. La réduction de la fonction potentielle primale-duale  $\psi(X, S)$ .

#### Transformation projective

Dans ces méthodes, à chaque itération, on utilise la transformation projective  $T_k$  qui nous permet de ramener le problème (PSD) à une forme réduite plus maniable. En effet, supposons à l'itération  $k$ , qu'on dispose d'une solution strictement réalisable  $X_k \in \mathbb{S}_{++}^n$ . La factorisation de  $X_k$  donne une matrice triangulaire inférieure  $L_k$  telle que  $X_k = L_k L_k^T$ . Il existe plusieurs choix de  $L_k$  telle que la factorisation de Cholesky de  $X_k$  ou bien la racine carrée de  $X_k$ , i.e.,  $L_k = X_k^{\frac{1}{2}}$ .

Soit  $r$  un entier fixé, on définit la transformation projective comme suit

$$\begin{aligned} T_k : \mathbb{S}^n &\rightarrow \mathbb{S}^n \times \mathbb{R}^r \\ X &\mapsto T_k(X) = (\bar{X}, \bar{x}) \end{aligned}$$

où

$$\bar{X} = \frac{(n+r)L_k^{-1}XL_k^{-T}}{r + \langle X_k^{-1}, X \rangle}, \quad \bar{x} = \frac{(n+r)}{r + \langle X_k^{-1}, X \rangle} e_r, \quad e_r = (1, \dots, 1)^T \in \mathbb{R}^r.$$

Aussi, la transformation inverse est donnée par

$$X = T_k^{-1}(\bar{X}, \bar{x}) = r \frac{L_k \bar{X} L_k^T}{e_r^T \bar{x}}.$$

Sous la transformation  $T_k$ , le problème (PSD) est transformé en un problème de programmation semi-définie linéaire (TPSD). Puis, comme dans la programmation linéaire, on remplace les contraintes des inégalités  $\bar{X} \geq 0$  et  $\bar{x} \geq 0$  de (TPSD) par la boule du centre  $(I, e_r)$  et de rayon  $\beta$  avec  $0 < \beta < 1$ . Donc le problème (TPSD) devient

$$\begin{cases} \min[\langle \bar{C}, \bar{X} \rangle - \frac{z_k e_r^T \bar{x}}{r}] \\ \langle A_i^k, \bar{X} \rangle - \frac{b_i e_r^T \bar{x}}{r} = 0, \quad i = 1, \dots, m \\ \text{tr}(\bar{X}) + e_r^T \bar{x} = n + r \\ \|\bar{X} - I\|^2 + \|\bar{x} - e_r\|^2 \leq \beta^2 < 1, \end{cases}$$

où

$$\bar{C} = L_k^T C L_k, \quad z_k = \langle C, X_k \rangle \text{ et } A_i^k = L_k^T A_i L_k.$$

Ce problème est un programme mathématique convexe et différentiable, donc les conditions de **K.K.T** sont nécessaires et suffisantes. La solution optimale obtenue est

$$\begin{cases} \bar{X}^* = I - \beta P_k \\ \bar{x}^* = (1 - \beta p_k) e_r, \end{cases}$$

où

$$P_k = V_k / (\|V_k\|^2 + r\alpha_k^2)^{\frac{1}{2}}, \quad p_k = \alpha_k / (\|V_k\|^2 + r\alpha_k^2)^{\frac{1}{2}},$$

avec

$$V_k = C_k + \sum_{i=1}^m y_i A_i^k, \quad \alpha_k = -\frac{1}{r} \left( \sum_{i=1}^m b_i y_i + z_k \right),$$

$y \in \mathbb{R}^m$  est la solution du système linéaire symétrique défini positif suivant

$$My = d, \quad M_{i,j} = \langle A_i^k, A_j^k \rangle + \frac{b_i b_j}{r}, \quad d_i = -\frac{b_i z_k}{r} - \langle C_k, A_i^k \rangle, \quad i, j = 1, \dots, m.$$

Le nouveau itéré est donc  $X_{k+1} = T_k^{-1}(\bar{X}, \bar{x})$ .

### Fonction potentielle primale-duale

Pour tout  $X, S \in \mathbb{S}_{++}^n$  la fonction potentielle primale-duale est définie par

$$\psi(X, S) = (n + r) \ln \langle X, S \rangle - \ln \det(XS).$$

Cette fonction joue un rôle très important dans le développements des algorithmes de réduction du potentiel après 1988. Les algorithmes correspondants à ces méthodes possèdent une complexité polynomiale, ils nécessitent  $O(\sqrt{n} |\ln \epsilon|)$  itérations pour réduire le saut de dualité ( $\langle X, S \rangle \leq \epsilon$  où  $\epsilon$  est une précision donnée ). (Voir Alizadeh [2]).

### 2.4.3 Méthodes de trajectoire centrale

Les méthodes de trajectoire centrale primale-duale ont été introduites à la même époque que les méthodes de réduction du potentiel et pleinement développées au début des années 90. Elles ont attiré une grande attention de la part des chercheurs dans le monde entier et elles montrent en général un excellent comportement pratique et théorique (une complexité polynomiale et une convergence super linéaire). On trouve en 1996, les travaux de Helmberg et al [16] qui ont proposé un algorithme primal-dual de trajectoire centrale pour (PSD). En 1997, Monteiro [38] a proposé une méthode primale-duale de trajectoire centrale et a montré la convergence polynomiale de l'algorithme à court et long pas. En 1999, Ji et al [30] ont étudié la convergence de la méthode de trajectoire centrale de type prédicteur-correcteur. Aussi, Monteiro et al [39] ont proposé une méthode primale-duale de trajectoire centrale et ont montré la convergence

polynomiale vers une solution optimale. Ces travaux se poursuivent à ce jour, on trouve ainsi ceux de Halicka et al (2002) [15] qui ont proposé une étude sur la convergence de la méthode de trajectoire centrale en optimisation semi-définie. En 2007 Koulaei et al [29] ont proposé une extension de l'algorithme de Mehrotra prédicteur-correcteur basé sur la direction de Nesterov et Todd (NT). En 2012, Liu et al [31] ont présenté un nouveau algorithme de type correcteur de second ordre pour (PSD) et ont prouvé la convergence polynomiale de ce dernier pour la direction de Nesterov et Todd (NT). En 2015, Kettab et al [27] ont proposé une relaxation de la méthode barrière logarithmique pour la programmation semi-définie.

Les méthodes de trajectoire centrale reposent sur les principes suivants :

On associe au problème initial (PSD) un problème perturbé  $(PSD)_\mu$ , puis on applique les conditions de (KKT) sur ce problème, on obtient le système non linéaire suivant

$$\begin{cases} \text{tr}(A_i X) = b_i, i = 1, \dots, m, X > 0, \\ \sum_{i=1}^m y_i A_i + S = C, S > 0, \\ XS = \mu I, \quad \mu > 0. \end{cases} \quad (2.14)$$

Où, on suppose que

- $(H_1)$  Les matrices  $A_i, i = 1, \dots, m$ , sont linéairement indépendentes.
- $(H_2)$   $\overset{0}{\mathcal{F}}(PSD) \times \overset{0}{\mathcal{F}}(SDD) \neq \phi$ .

tels que  $\overset{0}{\mathcal{F}}(PSD)$  et  $\overset{0}{\mathcal{F}}(SDD)$  désignent les ensembles des solutions strictement réalisables de (PSD) et (SDD) respectivement, i.e.,

$$\overset{0}{\mathcal{F}}(PSD) = \{X \in \mathbb{S}_{++}^n : \mathcal{A}X = b\},$$

$$\overset{0}{\mathcal{F}}(SDD) = \{(y, S) \in \mathbb{R}^m \times \mathbb{S}_{++}^n : \mathcal{A}^*y + S = C\},$$

Sous les hypothèses  $(H_1)$  et  $(H_2)$ , le système (2.14) admet pour chaque  $\mu > 0$ , une solution unique  $(X(\mu), y(\mu), S(\mu))$ . Cette solution s'appelle  $\mu$ -centre de (PSD) et de (SDD). L'ensemble des  $\mu$ -centres,  $\Lambda = \{X(\mu), y(\mu), S(\mu), \mu > 0\}$ , construit la trajectoire centrale

qui converge vers la solution optimale primale-duale de (PSD) et (SDD) quand  $\mu$  tend vers zéro.

Le système (2.14) doit être résolu par la méthode de Newton, cette dernière consiste à trouver la direction  $\Delta W = (\Delta X, \Delta y, \Delta S)$ , en résolvant le système linéaire suivant

$$\begin{cases} \mathcal{A}\Delta X = b, & X > 0 \\ \mathcal{A}^*\Delta y + \Delta S = C, & S > 0 \\ S\Delta X + X\Delta S = \mu I - XS, & \mu > 0. \end{cases} \quad (2.15)$$

Le nouvel itéré est donné par

$$X^+ = X + \Delta X$$

$$y^+ = y + \Delta y$$

$$S^+ = S + \Delta S.$$

Signalant que la symétrie des itérés  $X^+$  et  $S^+$  n'est pas préservée pour toutes les directions trouvées dans (2.15). Pour cela, plusieurs aménagements ont été proposés par des chercheurs afin de résoudre ce problème. On cite par exemple les travaux de Zhang [59], Helmsberg et al, Kojima et al et Monteiro [16, 28, 38], Alizadeh-Heaberly-Overton [3] et Nesterov-Todd [48], qui ont proposé des opérateurs symétrisant les directions de déplacement.

Actuellement, plusieurs chercheurs parmi eux, Y. Q. Bai et autres [5, 8, 10, 13, 55, 56] ont introduit la notion des fonctions noyaux pour trouver une nouvelle classe de directions, à travers laquelle les auteurs ont pu améliorer la complexité algorithmique de méthodes à court et long pas.

Dans le chapitre suivant, on présente d'une manière détaillée une étude théorique et numérique sur une méthode de trajectoire centrale. Et on met en œuvre l'algorithme correspondant pour la résolution d'un programme semi-défini (PSD).

# Chapitre 3

## Méthode réalisable de trajectoire centrale pour la programmation semi-définie

Dans ce chapitre, on considère une méthode réalisable primale-duale de points intérieurs pour la programmation semi-définie linéaire (PSD) basée sur la direction de Alizadeh-Haeberly-Overton (AHO) [3, 40]. En premier lieu et par une nouvelle et simple technique, on établit l'existence et l'unicité de la solution optimale du problème perturbé  $(PSD)_\mu$  et sa convergence vers la solution optimale du problème (PSD). Ensuite, on présente quatre nouvelles différentes alternatives pour calculer le pas de déplacement. Après, on établit la convergence de l'algorithme obtenu et on montre que sa complexité est d'ordre  $\mathcal{O}\left(\sqrt{n} \ln \left[\varepsilon^{-1}(\langle X^0, S^0 \rangle)\right]\right)$ . Finalement, on présente quelques tests numériques qui montrent l'efficacité de l'algorithme développé dans ce chapitre.

### 3.1 Position du problème

Rappelons que le problème de programmation semi-définie, considéré comme primal est défini par

$$(PSD) \begin{cases} \min \langle C, X \rangle \\ \mathcal{A}X = b, \\ X \in \mathbb{S}_+^n, \end{cases}$$

où  $b \in \mathbb{R}^m$ , les matrices  $C$  et  $A_i, i = 1, \dots, m$ , sont dans  $\mathbb{S}^n$ .

Le problème dual associé à  $(PSD)$  est défini par

$$(SDD) \begin{cases} \max b^T y \\ \mathcal{A}^* y + S = C, \\ S \in \mathbb{S}_+^n, \end{cases}$$

Rappelons aussi que les ensembles des solutions strictement réalisables de  $(PSD)$  et  $(SDD)$  sont respectivement

$$\overset{0}{\mathcal{F}}(PSD) = \{X \in \mathbb{S}_{++}^n : \mathcal{A}X = b\},$$

$$\overset{0}{\mathcal{F}}(SDD) = \{(y, S) \in \mathbb{R}^m \times \mathbb{S}_{++}^n : \mathcal{A}^* y + S = C\},$$

Tout au long de ce chapitre, on suppose que les hypothèses  $(H_1)$  et  $(H_2)$  sont satisfaites, i.e., les matrices  $A_i, i = 1, \dots, m$ , sont linéairement indépendantes et que  $\overset{0}{\mathcal{F}}(PSD) \times \overset{0}{\mathcal{F}}(SDD)$  est non vide. Puisque sous la deuxième hypothèse, il est bien connu que les deux problèmes  $(PSD)$  et  $(SDD)$  ont des solutions optimales  $\bar{X}$  et  $(\bar{S}, \bar{y})$  telles que  $\langle C, \bar{X} \rangle = b^T \bar{y}$ , i.e., les valeurs optimales de  $(PSD)$  et  $(SDD)$  coïncident. La dualité forte, et donc établie et peut être exprimé par  $\langle \bar{X}, \bar{S} \rangle = 0$  ou  $\bar{X} \bar{S} = 0$ .

Pour étudier le problème  $(PSD)$ , on le remplace par son problème perturbé équivalent

$$(PSD)_\mu \begin{cases} \min [f_\mu(X) = \langle C, X \rangle + \mu g(X) + n\mu \ln \mu], & \mu > 0, \\ \mathcal{A}X = b, \end{cases}$$

3.2. Existence et unicité de la solution optimale du problème  $(PSD)_\mu$  et sa convergence vers une solution du  $(PSD)$

---

où

$$g(X) = \begin{cases} -\ln(\det X) & \text{si } X \in \mathbb{S}_{++}^n, \\ +\infty & \text{ailleurs.} \end{cases}$$

On peut aussi étudier  $(PSD)$  à travers son problème dual perturbé

$$(SDD)_\mu \begin{cases} \max[g_\mu(y, S) = b^T y - \mu g(S) - n\mu \ln \mu], & \mu > 0, \\ \mathcal{A}^* y + S = C. \end{cases}$$

L'avantage principal de  $(PSD)_\mu$  réside dans la stricte convexité de sa fonction objectif et son domaine réalisable. Par conséquent, les conditions d'optimalité sont nécessaires et suffisantes. Ceci stimule les études théoriques et numériques du problème.

Commençons d'abord par montrer que  $(PSD)_\mu$  admet au moins une solution optimale.

## 3.2 Existence et unicité de la solution optimale du problème $(PSD)_\mu$ et sa convergence vers une solution du $(PSD)$

### 3.2.1 Existence et unicité de la solution optimale de $(PSD)_\mu$

Comme la fonction  $f_\mu$  est strictement convexe, alors si une solution optimale du problème  $(PSD)_\mu$  existe elle est unique. Pour montrer que le problème  $(PSD)_\mu$  a une solution, il suffit de montrer que la fonction  $f_\mu$  est inf-compacte, ce qui revient en particulier à montrer que le cône de récession se réduit à zéro. Pour cette raison, on a besoin du lemme suivant.

**Lemme 3.2.1** Soient  $X$  une matrice de  $\mathbb{S}_{++}^n$  et  $t \in \mathbb{R}$ , alors

$$\lim_{t \rightarrow +\infty} \frac{f_\mu(X + t\Delta X) - f_\mu(X)}{t} = \langle C, \Delta X \rangle.$$

Où  $\Delta X$  est la direction de descente qui sera définie par la suite.

**Preuve.** On pose  $\xi(t) = \frac{f_\mu(X+t\Delta X) - f_\mu(X)}{t}$ , alors

$$\xi(t) = \langle C, \Delta X \rangle - \mu t^{-1} [\ln \det(X + t\Delta X) - \ln \det(X)].$$



3.2. Existence et unicité de la solution optimale du problème  $(PSD)_\mu$  et sa convergence vers une solution du  $(PSD)$

---

La factorisation de Cholesky de  $X \in \mathbb{S}_{++}^n$  donne une matrice triangulaire inférieure  $L_X$  avec des éléments diagonaux positifs telle que  $X = L_X L_X^T$ . Alors, on peut écrire

$$\begin{aligned}\xi(t) &= \langle C, \Delta X \rangle - \mu t^{-1} \left[ \ln \det(X) + \ln \det(I + t L_X^{-1} \Delta X L_X^{-T}) - \ln \det(X) \right] \\ &= \langle C, \Delta X \rangle - \mu t^{-1} \left[ \ln \det(I + t L_X^{-1} \Delta X L_X^{-T}) \right],\end{aligned}$$

donc

$$\xi(t) = \begin{cases} \langle C, \Delta X \rangle - \mu t^{-1} \left[ \ln \det(I + t L_X^{-1} \Delta X L_X^{-T}) \right] & \text{si } I + t L_X^{-1} \Delta X L_X^{-T} \in \mathbb{S}_{++}^n, \\ +\infty & \text{ailleurs.} \end{cases}$$

Passons à la limite, on trouve

$$\begin{aligned}\lim_{t \rightarrow +\infty} \xi(t) &= \lim_{t \rightarrow +\infty} [\langle C, \Delta X \rangle - \mu t^{-1} \ln \det(I + t L_X^{-1} \Delta X L_X^{-T})] \\ &= \langle C, \Delta X \rangle - \mu \lim_{t \rightarrow +\infty} \frac{1}{t} \left[ \ln \det(I + t L_X^{-1} \Delta X L_X^{-T}) \right] \\ &= \langle C, \Delta X \rangle - \mu \lim_{t \rightarrow +\infty} \sum_{i=1}^n \frac{[\ln(1 + t \lambda_i(\Delta X))]}{t}.\end{aligned}$$

Puisque

$$\det(I + t L_X^{-1} \Delta X L_X^{-T}) = \prod_{i=1}^n (1 + t \lambda_i(\Delta X)),$$

où  $\lambda_i(\Delta X)$ ,  $i = 1, \dots, n$ , sont les valeurs propres de  $L_X^{-1} \Delta X L_X^{-T}$ .

Comme  $\lim_{t \rightarrow +\infty} \frac{\ln(1+t)}{t} = 0$ , on obtient

$$\lim_{t \rightarrow +\infty} \xi(t) = \lim_{t \rightarrow +\infty} \frac{f_\mu(X + t \Delta X) - f_\mu(X)}{t} = \langle C, \Delta X \rangle.$$

Ce qui complète la preuve. ■

Rappelons que le cône de récession est défini par

$$C_\infty(f_\mu) = \left\{ \Delta X \in \mathbb{S}_+^n : (f_\mu)_\infty(\Delta X) = \lim_{t \rightarrow +\infty} [\xi(t) = \frac{f_\mu(X + t \Delta X) - f_\mu(X)}{t}] \leq 0 \right\},$$

qui s'écrit de manière équivalente d'après le lemme précédent

$$C_\infty(f_\mu) = \{ \Delta X \in \mathbb{S}_+^n : \langle C, \Delta X \rangle \leq 0 \}.$$

3.2. Existence et unicité de la solution optimale du problème  $(PSD)_\mu$  et sa convergence vers une solution du  $(PSD)$

---

**Proposition 3.2.2**  $C_\infty(f_\mu) = \{\Delta X \in \mathbb{S}_+^n : \langle C, \Delta X \rangle \leq 0\} = \{0\}$ . Par conséquent, le problème  $(PSD)_\mu$  admet une solution unique.

**Preuve.** Supposons que  $C_\infty(f_\mu) \neq \{0\}$ , alors il existe  $\Delta X \neq 0$  tel que  $\Delta X \in C_\infty(f_\mu)$  i.e.,  $\langle C, \Delta X \rangle \leq 0$  et d'après l'hypothèse  $(H_2)$ , i.e.,  $\overset{0}{\mathcal{F}}(PSD) \times \overset{0}{\mathcal{F}}(SDD) \neq \emptyset$ , il existe  $(y, \hat{S}) \in \overset{0}{\mathcal{F}}(SDD)$  tel que

$$\begin{aligned} 0 < \langle \hat{S}, \Delta X \rangle &= \left\langle C - \sum_{i=1}^m y_i A_i, \Delta X \right\rangle \\ &= \langle C, \Delta X \rangle - \sum_{i=1}^m y_i \langle A_i, \Delta X \rangle. \end{aligned}$$

Comme  $\Delta X$  est une direction de descente, i.e.,  $\langle A_i, \Delta X \rangle = 0, \forall i = 1, \dots, m$ , alors on obtient  $\langle C, \Delta X \rangle > 0$  contradiction. ■

### 3.2.2 Convergence de la solution optimale de $(PSD)_\mu$ vers la solution optimale de $(PSD)$

**Lemme 3.2.3** Soit  $\bar{X}_\mu$  une solution optimale primale de  $(PSD)_\mu$ , alors  $\bar{X} = \lim_{\mu \rightarrow 0} \bar{X}_\mu$  est une solution optimale de  $(PSD)$ .

**Preuve.** On pose  $f_\mu(X) = f(X, \mu)$  et  $f(X) = f(X, 0)$ . Comme  $f_\mu(X)$  est différentiable et convexe, alors il existe une solution optimale  $\bar{X}_\mu$  de  $(PSD)_\mu$  telle que

$$\nabla_X f_\mu(\bar{X}_\mu) = \nabla_X f(\bar{X}_\mu, \mu) = 0.$$

D'où pour tout  $X \in \overset{0}{\mathcal{F}}(PSD)$ , on a

$$\begin{aligned} f(X) &\geq f(\bar{X}_\mu, \mu) + \langle X - \bar{X}_\mu, \nabla_X f(\bar{X}_\mu, \mu) \rangle + (0 - \mu) \frac{df}{d\mu} f(\bar{X}_\mu, \mu) \\ &\geq f(\bar{X}_\mu, \mu) + \mu \ln \det \bar{X}_\mu - n\mu \ln \mu - n\mu \\ &\geq \langle C, \bar{X}_\mu \rangle - \mu \ln \det \bar{X}_\mu + n\mu \ln \mu + \mu \ln \det \bar{X}_\mu - n\mu \ln \mu - n\mu \\ &\geq \langle C, \bar{X}_\mu \rangle - n\mu, \end{aligned}$$

ce qui implique d'une part

$$\min_{X \in \overset{0}{\mathcal{F}}(PSD)} [f(X) = f(X, 0)] \geq \langle C, \bar{X}_\mu \rangle - n\mu,$$

d'autre part, on a

$$\langle C, \bar{X}_\mu \rangle = f(\bar{X}_\mu) \geq \min_{X \in \mathcal{F}^0(PSD)} [f(X) = f(X, 0)].$$

Donc quand  $\mu$  tend vers 0, on conclut que

$$f(\bar{X}) = \min_{X \in \mathcal{F}^0(PSD)} f(X),$$

où  $\bar{X} = \lim_{\mu \rightarrow 0} \bar{X}_\mu$ .

Par conséquent,  $\bar{X}$  est une solution optimale de (PSD). ■

**Remarque 3.2.1** On sait que si les valeurs des fonctions objectifs des problèmes (PSD) et (SDD) sont égales et finies, alors si l'un de ces problèmes admet une solution optimale, l'autre l'admet aussi.

### 3.3 Méthode de trajectoire centrale primale-duale

Rappelons que  $(PSD)_\mu$  est strictement convexe, donc les conditions de **KKT** sont nécessaires et suffisantes. Par conséquent,  $(\bar{X}_\mu, \bar{y}_\mu, \bar{S}_\mu)$  est une solution optimale primale-duale de  $(PSD)_\mu$  et  $(SDD)_\mu$  lorsqu'elle satisfait le système non linéaire suivant

$$\begin{cases} \mathcal{A}X = b, X > 0, \\ \mathcal{A}^*y + S = C, S > 0, \\ XS = \mu I, \quad \mu > 0. \end{cases} \quad (3.1)$$

Pour tout  $\mu > 0$ , notons par  $(X_\mu, y_\mu, S_\mu)$  la solution du système (3.1), ce triplet est appelé *point centre*. L'ensemble de toutes ces solutions (points centres) définit *le chemin central* ou *la trajectoire centrale*.

On dit qu'un point  $(X, y, S)$  est proche de la trajectoire centrale, c'est-à-dire dans le voisinage de longueur  $\theta$  de la trajectoire centrale s'il appartient à l'ensemble suivant

$$T_\theta(\mu) = \{(X, y, S) \in \mathcal{F}^0(PSD) \times \mathcal{F}^0(SDD) : \|\frac{XS + SX}{2} - \mu I\|_F \leq \theta\mu\},$$

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

Comme  $X, S \in \mathcal{S}^n$ , le produit  $XS$  n'est pas généralement dans  $\mathcal{S}^n$ , alors le côté gauche du système (3.1) est une application de  $\mathcal{S}^n \times \mathbb{R}^m \times \mathcal{S}^n$  dans  $\mathbb{R}^{n \times n} \times \mathbb{R}^m \times \mathcal{S}^n$ . Donc, on doit modifier le côté gauche du système (3.1) par une application de  $\mathcal{S}^n \times \mathbb{R}^m \times \mathcal{S}^n$  dans lui même. Pour cela, on utilise l'opérateur de *symétrisation similaire*  $H_P : \mathbb{R}^{n \times n} \rightarrow \mathcal{S}^n$  introduit par Zhang [59] défini comme suit

$$H_P(M) = \frac{1}{2}[PMP^{-1} + (PMP^{-1})^T], \quad \forall M \in \mathbb{R}^{n \times n},$$

où  $P \in \mathbb{R}^{n \times n}$  est une matrice régulière appelée matrice mise à l'échelle (Scaling matrix). Zhang [59] a aussi observé que si  $P$  est inversible et  $M$  est semblable à une matrice (symétrique) définie positive, alors

$$H_P(M) = \mu I \iff M = \mu I.$$

Donc, pour toute matrice inversible donnée, le système (3.1) est équivalent à

$$\begin{cases} \mathcal{A}X = b, X \succ 0, \\ \mathcal{A}^*y + S = C, S \succ 0, \\ H_P(XS) = \mu I, \quad \mu > 0. \end{cases} \quad (3.2)$$

En appliquant la méthode de Newton sur le système (3.2), on obtient l'itération

$$X^+ = X + \Delta X, \quad y^+ = y + \Delta y, \quad S^+ = S + \Delta S,$$

où  $(\Delta X, \Delta y, \Delta S)$ , est une solution du système linéaire suivant

$$\begin{cases} \mathcal{A}\Delta X & = 0, \\ \mathcal{A}^*\Delta y + \Delta S & = 0, \\ H_P(X\Delta S + \Delta XS) & = \sigma\mu I - H_P(XS), \end{cases} \quad (3.3)$$

Notons que  $\Delta W = (\Delta X, \Delta y, \Delta S) \in \mathcal{S}^n \times \mathbb{R}^m \times \mathcal{S}^n$  est la direction cherchée,  $\sigma \in (0, 1)$  est le paramètre de centralité et  $\mu = \frac{\langle X, S \rangle}{n}$  est le saut de dualité qui correspond à  $(X, y, S)$ . Le système (3.3) peut être écrit comme suit

$$\nabla F(X, y, S)\Delta W = -F(X, y, S), \quad (3.4)$$

avec

$$F(X, y, S) = \begin{bmatrix} \mathcal{A}X - b \\ \mathcal{A}^*y + S - C \\ H_P(XS) - \sigma\mu I \end{bmatrix}.$$

Les directions obtenues par la résolution du système précédent sont appelées la famille de Monteiro-Zhang (**MZ**). Pour  $\mathbf{P} = \mathbf{I}$ , on obtient la direction de Alizadeh-Haeberly-Overton (**AHO**) [3, 40], pour  $\mathbf{P} = [\mathbf{X}^{\frac{1}{2}}(\mathbf{X}^{\frac{1}{2}}\mathbf{S}\mathbf{X}^{\frac{1}{2}-\frac{1}{2}}\mathbf{X}^{\frac{1}{2}})^{\frac{1}{2}}]$ , on obtient la direction de Nesterov et Todd (**NT**) [55] et pour  $\mathbf{P} = \mathbf{X}^{-\frac{1}{2}}$  ou  $\mathbf{P} = \mathbf{S}^{\frac{1}{2}}$ , on obtient la famille des directions de Helmberg et al, Kojima et al et Monteiro (**HKM**) [16, 28, 38].

Dans ce qui suit, on utilise dans notre algorithme la direction de **AHO** vu sa simplicité pratique.

**Remarque 3.3.1** *La définie positivité des matrices  $X^+ = X + \Delta X$  et  $S^+ = S + \Delta S$  n'est pas toujours garantie. Pour surmonter cette difficulté, on introduit un paramètre  $\alpha > 0$  appelé pas de déplacement et on pose*

$$X^+ = X + \alpha\Delta X, \quad y^+ = y + \alpha\Delta y \quad \text{et} \quad S^+ = S + \alpha\Delta S.$$

Le calcul du pas de déplacement par les méthodes de recherche linéaire classiques est indésirable et en général impossible.

Dans ce sens, J. P. Crouzeix et B. Merikhi [11] ont utilisé les notions des fonctions majorantes pour le problème dual de (*PSD*), cela permet de calculer des pas de déplacement par une technique simple. Indépendamment, en se basant sur des propriétés du calcul matriciel sur la définie positivité, on propose quatre différentes alternatives qui offrent des pas de déplacement variables à chaque itération. L'efficacité d'une alternative par rapport à l'autre peut être mesurée par des tests numériques qu'on présentera à la fin de ce chapitre.

### 3.3.1 Calcul du pas de déplacement

L'itéré de la méthode de Newton modifié est décrit par la formule

$$X^+ = X + \alpha \Delta X, y^+ = y + \alpha \Delta y, S^+ = S + \alpha \Delta S,$$

où  $\alpha$  est le pas de déplacement choisi tel que  $X^+, S^+ \in \mathbb{S}_{++}^n$ .

Avant de donner l'expression de  $\alpha$ , on a besoin d'appliquer sur une matrice carrée le résultat de H. Wolkowicz et G. P. H. Styan [58] que nous le rappelons dans la proposition suivante.

**Proposition 3.3.1** [58] *Soit  $A$  une  $(n \times n)$  matrice et soient  $\lambda_1, \lambda_2, \dots, \lambda_n$ , ses valeurs propres, on a*

$$\bar{\lambda} - \delta \sqrt{n-1} \leq \min_{i=1}^n \lambda_i \leq \bar{\lambda} - \frac{\delta}{\sqrt{n-1}},$$

$$\bar{\lambda} + \frac{\delta}{\sqrt{n-1}} \leq \max_{i=1}^n \lambda_i \leq \bar{\lambda} + \delta \sqrt{n-1},$$

où

$$\bar{\lambda} = \frac{1}{n} \sum_{i=1}^n (A)_{ii} \text{ et } \delta^2 = \frac{1}{n} \sum_{i=1}^n \sum_{j=1}^n (A)_{ij}^2 - \bar{\lambda}^2.$$

En se basant sur cette proposition, on donne dans les lemmes suivants, quatre différentes alternatives pour calculer le pas de déplacement  $\alpha$ .

#### Première alternative

**Lemme 3.3.2** *Soit  $(X, y, S) \in \mathcal{F}(PSD) \times \mathcal{F}(SDD)$ . Si  $\alpha = \rho \min(\alpha_X, \alpha_S)$  avec  $0 < \rho < 1$ , alors*

$$(X^+, S^+) \in \mathbb{S}_{++}^n \times \mathbb{S}_{++}^n,$$

où pour  $A = X$  ou  $A = S$ , on a

$$\alpha_A = \begin{cases} \frac{-1}{\bar{\lambda}_A - \delta_A \sqrt{n-1}} - \varepsilon & \text{si } \left( \frac{-1}{\bar{\lambda}_A - \delta_A \sqrt{n-1}} > 0 \text{ et } \min_{i=1}^n \lambda_i(\Delta A) < 0 \right), \\ \varepsilon & \text{si } \left( \frac{-1}{\bar{\lambda}_A - \delta_A \sqrt{n-1}} < 0 \text{ et } \min_{i=1}^n \lambda_i(\Delta A) < 0 \right), \\ 1 & \text{si } \min_{i=1}^n \lambda_i(\Delta A) > 0, \end{cases}$$

tels que  $\bar{\lambda}_A = \frac{1}{n} \sum_{i=1}^n (L_A^{-1} \Delta A L_A^{-T})_{ii}$ ,  $\delta_A^2 = \frac{1}{n} \sum_{i=1}^n \sum_{j=1}^n (L_A^{-1} \Delta A L_A^{-T})_{ij}^2 - \bar{\lambda}_A^2$ ,  $\lambda_i(\Delta A)$ ,  $i = 1, \dots, n$ , sont les valeurs propres de  $L_A^{-1} \Delta A L_A^{-T}$ ,  $\varepsilon$  est un petit réel positif et  $A = L_A L_A^T$  avec  $L_A$  est la factorisation de Cholesky de  $A$ .

**Preuve.** On sait que  $X = L_X L_X^T$ . Par conséquent

$$\begin{aligned} X^+ &= X + \alpha \Delta X \\ &= L_X L_X^T + \alpha \Delta X \\ &= L_X (I + \alpha L_X^{-1} \Delta X L_X^{-T}) L_X^T. \end{aligned}$$

Or d'après la Proposition 1.1.15, la matrice  $X^+$  est définie positive si et seulement si la matrice  $(I + \alpha L_X^{-1} \Delta X L_X^{-T})$  l'est, ce qui est équivalent d'après le Lemme 1.1.9 à

$$1 + \alpha \lambda_i(\Delta X) > 0, \quad \forall i = 1, \dots, n,$$

où  $\lambda_i(\Delta X)$ ,  $i = 1, \dots, n$ , sont les valeurs propres de  $L_X^{-1} \Delta X L_X^{-T}$ .

Alors, il suffit de trouver  $\alpha > 0$  vérifiant

$$1 + \alpha \min_{i=1}^n \lambda_i(\Delta X) > 0,$$

ou de manière équivalente, trouver  $\alpha > 0$  tel que

$$\alpha \min_{i=1}^n \lambda_i(\Delta X) > -1.$$

Si  $\min_{i=1}^n \lambda_i(\Delta X) > 0$  (ce qui se traduit par le fait que toutes les valeurs propres de  $\Delta X$  sont strictement positives), alors l'inégalité précédente est vérifiée pour toute valeur de

$\alpha > 0$ . Par contre, si  $\min_{i=1}^n \lambda_i(\Delta X) < 0$ , il suffit de choisir  $0 < \alpha < \frac{-1}{\min_{i=1}^n \lambda_i(\Delta X)}$ .

D'autre part, d'après la Proposition 3.3.1, on en déduit que

$$\frac{-1}{\bar{\lambda}_X - \delta_X \sqrt{n-1}} \leq \frac{-1}{\min_{i=1}^n \lambda_i(\Delta X)},$$

ce qui donne l'expression de  $\alpha_X$ .

De la même manière, on obtient l'expression de  $\alpha_S$ .

Finalement, on prend

$$\alpha = \rho \min(\alpha_X, \alpha_S), \quad 0 < \rho < 1,$$

ce qui complète la preuve. ■

### Deuxième alternative

**Lemme 3.3.3** Soit  $(X, y, S) \in \mathcal{F}^0(PSD) \times \mathcal{F}^0(SDD)$ . Si  $\alpha = \rho \min(\alpha_X, \alpha_S)$  avec  $0 < \rho < 1$ , alors

$$(X^+, S^+) \in \mathbb{S}_{++}^n \times \mathbb{S}_{++}^n,$$

où pour  $A = X$  ou  $A = S$ , on a

$$\alpha_A = \begin{cases} \frac{-1}{\bar{\lambda}_A - \delta_A \sqrt{n-1}} - \varepsilon & \text{si } \left( \frac{-1}{\bar{\lambda}_A - \delta_A \sqrt{n-1}} > 0 \text{ et } \min_{i=1}^n \lambda_i(A^{-1}\Delta A) < 0 \right), \\ \varepsilon & \text{si } \left( \frac{-1}{\bar{\lambda}_A - \delta_A \sqrt{n-1}} < 0 \text{ et } \min_{i=1}^n \lambda_i(A^{-1}\Delta A) < 0 \right), \\ 1 & \text{si } \min_{i=1}^n \lambda_i(A^{-1}\Delta A) > 0, \end{cases}$$

tels que  $\bar{\lambda}_A = \frac{1}{n} \sum_{i=1}^n (A^{-1}\Delta A)_{ii}$ ,  $\delta_A^2 = \frac{1}{n} \sum_{i=1}^n \sum_{j=1}^n (A^{-1}\Delta A)_{ij}^2 - \bar{\lambda}_A^2$ ,  $\lambda_i(A^{-1}\Delta A)$ ,  $i = 1, \dots, n$ , sont les valeurs propres de  $A^{-1}\Delta A$  et  $\varepsilon$  est un petit réel positif.

**Preuve.** On a

$$\begin{aligned} X^+ &= X + \alpha \Delta X \\ &= X(I + \alpha X^{-1} \Delta X). \end{aligned}$$

Comme  $X \in \mathbb{S}_+^n$  et d'après le Lemme 1.1.12, 3)

$$X^+ = X(I + \alpha X^{-1} \Delta X) \text{ et } X^{\frac{1}{2}}(I + \alpha X^{-1} \Delta X)X^{\frac{1}{2}}$$

ont les mêmes valeurs propres, or d'après la Proposition 1.1.15, si  $(I + \alpha X^{-1} \Delta X)$  est définie positive, alors  $X^+$  l'est aussi.

Par le Lemme 1.1.9, la matrice  $(I + \alpha X^{-1} \Delta X)$  est définie positive si et seulement si

$$1 + \alpha \lambda_i(X^{-1} \Delta X) > 0, \quad \forall i = 1, \dots, n,$$



où  $\lambda_i(X^{-1}\Delta X), i = 1, \dots, n$ , sont les valeurs propres de  $X^{-1}\Delta X$ .

Alors, il suffit de trouver  $\alpha > 0$  vérifiant

$$1 + \alpha \min_{i=1}^n \lambda_i(X^{-1}\Delta X) > 0,$$

ou de manière équivalente, trouver  $\alpha > 0$  tel que

$$\alpha \min_{i=1}^n \lambda_i(X^{-1}\Delta X) > -1.$$

Si  $\min_{i=1}^n \lambda_i(X^{-1}\Delta X) > 0$  (ce qui se traduit par le fait que toutes les valeurs propres de  $\Delta X$  sont strictement positives), alors l'inégalité précédente est vérifiée pour toute valeur de  $\alpha > 0$ . Par contre, si  $\min_{i=1}^n \lambda_i(X^{-1}\Delta X) < 0$ , il suffit de choisir  $0 < \alpha < \frac{-1}{\min_{i=1}^n \lambda_i(X^{-1}\Delta X)}$ .

D'autre part, d'après la Proposition 3.3.1, on déduit que

$$\frac{-1}{\bar{\lambda}_X - \delta_X \sqrt{n-1}} \leq \frac{-1}{\min_{i=1}^n \lambda_i(X^{-1}\Delta X)},$$

ce qui donne l'expression de  $\alpha_X$ .

De la même manière, on obtient l'expression de  $\alpha_S$ .

Finalement, on prend

$$\alpha = \rho \min(\alpha_X, \alpha_S), \quad 0 < \rho < 1,$$

ce qui complète la preuve. ■

### Troisième alternative

**Lemme 3.3.4** Soit  $(X, y, S) \in \mathcal{F}^0(\text{PSD}) \times \mathcal{F}^0(\text{SDD})$ . Si  $\alpha = \rho \min(\alpha_X, \alpha_S)$  avec  $0 < \rho < 1$ , alors

$$(X^+, S^+) \in \mathbb{S}_{++}^n \times \mathbb{S}_{++}^n,$$

où pour  $A = X$  ou  $A = S$ , on a

$$\alpha_A = \begin{cases} -\frac{(\bar{\lambda}_A - \delta_A \sqrt{n-1})}{(\bar{\beta}_A - \delta_{\Delta A} \sqrt{n-1})} - \varepsilon & \text{si } \left( -\frac{(\bar{\lambda}_A - \delta_A \sqrt{n-1})}{(\bar{\beta}_A - \delta_{\Delta A} \sqrt{n-1})} > 0 \text{ et } \min_{i=1}^n \lambda_i(\Delta A) < 0 \right), \\ \varepsilon & \text{si } \left( -\frac{(\bar{\lambda}_A - \delta_A \sqrt{n-1})}{(\bar{\beta}_A - \delta_{\Delta A} \sqrt{n-1})} < 0 \text{ et } \min_{i=1}^n \lambda_i(\Delta A) < 0 \right), \\ 1 & \text{si } \min_{i=1}^n \lambda_i(\Delta A) > 0, \end{cases}$$

tels que  $\bar{\lambda}_A = \frac{1}{n} \sum_{i=1}^n (A)_{ii}$ ,  $\delta_A^2 = \frac{1}{n} \sum_{i=1}^n \sum_{j=1}^n (A)_{ij}^2 - \bar{\lambda}_A^2$ ,  $\bar{\beta}_A = \frac{1}{n} \sum_{i=1}^n (\Delta A)_{ii}$ ,  $\delta_{\Delta A}^2 = \frac{1}{n} \sum_{i=1}^n \sum_{j=1}^n (\Delta A)_{ij}^2 - \bar{\beta}_A^2$ ,  $\lambda_i(\Delta A)$ ,  $i = 1, \dots, n$ , sont les valeurs propres de  $\Delta A$  et  $\varepsilon$  est un petit réel positif.

**Preuve.** On sait que  $X^+ = X + \alpha \Delta X$  est définie positive si et seulement si  $\min_{i=1}^n \gamma_i > 0$ , où  $\gamma_i$ ,  $i = 1, \dots, n$ , sont les valeurs propres de  $X^+$ .

D'après le Lemme 1.1.4, on a

$$\min_{i=1}^n \gamma_i \geq \min_{i=1}^n \beta_i + \min_{i=1}^n \lambda_i(\Delta X),$$

où  $\beta_i$  et  $\lambda_i(\Delta X)$ ,  $i = 1, \dots, n$ , sont les valeurs propres de  $X$  et  $\Delta X$  respectivement.

Alors, il suffit de trouver  $\alpha$  tel que

$$\min_{i=1}^n \gamma_i \geq \min_{i=1}^n \beta_i + \alpha \min_{i=1}^n \lambda_i(\Delta X) > 0.$$

Si  $\min_{i=1}^n \lambda_i(\Delta X) > 0$  (ce qui se traduit par le fait que toutes les valeurs propres de  $\Delta X$  sont strictement positives), alors l'inégalité précédente est vérifiée pour toute valeur de  $\alpha > 0$ . Par contre, si  $\min_{i=1}^n \lambda_i(\Delta X) < 0$ , il suffit de choisir  $0 < \alpha < -\frac{\min_{i=1}^n \beta_i}{\min_{i=1}^n \lambda_i(\Delta X)}$ .

D'autre part, d'après la Proposition 3.3.1, on déduit que

$$-\frac{(\bar{\lambda}_X - \delta_X \sqrt{n-1})}{(\bar{\beta}_X - \delta_{\Delta X} \sqrt{n-1})} \leq -\frac{\min_{i=1}^n \beta_i}{\min_{i=1}^n \lambda_i(\Delta X)}.$$

Ce qui donne l'expression de  $\alpha_X$ .

De la même manière, on obtient l'expression de  $\alpha_S$ .

Finalement, on prend

$$\alpha = \rho \min(\alpha_X, \alpha_S), \quad 0 < \rho < 1.$$

Ce qui achève la preuve. ■

#### Quatrième alternative

**Lemme 3.3.5** Soit  $(X, y, S) \in \mathcal{F}^0(\text{PSD}) \times \mathcal{F}^0(\text{SDD})$ . Si  $\alpha = \rho \min(\alpha_X, \alpha_S)$  avec  $0 < \rho < 1$ , alors

$$(X^+, S^+) \in \mathbb{S}_{++}^n \times \mathbb{S}_{++}^n,$$

où pour  $A = X$  ou  $A = S$ , on a

$$\alpha_A = \begin{cases} \min \frac{\left( \sum_{i \neq j=1}^n |A_{ij}| - A_{ii} \right)}{\left( \Delta A_{ii} - \sum_{i \neq j=1}^n |\Delta A_{ij}| \right)} & \text{si } I_A \neq \phi \\ +\infty & \text{si } I_A = \phi, \end{cases}$$

tel que

$$I_A = \left\{ i \in \{1, \dots, n\} : \Delta A_{ii} - \sum_{i \neq j=1}^n |\Delta A_{ij}| < 0 \right\}.$$

**Preuve.** D'après la Définition 1.1.9,  $X^+ = X + \alpha \Delta X$  est définie positive si

$$X_{ii}^+ > \sum_{i \neq j=1}^n |X_{ij}^+|, \quad \forall i = 1, \dots, n,$$

qui est équivalent à

$$X_{ii} + \alpha \Delta X_{ii} > \sum_{i \neq j=1}^n |X_{ij} + \alpha \Delta X_{ij}|, \quad \forall i = 1, \dots, n, \quad (3.5)$$

comme

$$\sum_{i \neq j=1}^n (|X_{ij}| + \alpha |\Delta X_{ij}|) \geq \sum_{i \neq j=1}^n |X_{ij} + \alpha \Delta X_{ij}|.$$

Alors, il suffit de trouver  $\alpha$  tel que

$$X_{ii} + \alpha \Delta X_{ii} > \sum_{i \neq j=1}^n (|X_{ij}| + \alpha |\Delta X_{ij}|), \quad \forall i = 1, \dots, n,$$

(3.5) est satisfaite dès que

$$\alpha \left( \Delta X_{ii} - \sum_{i \neq j=1}^n |\Delta X_{ij}| \right) > \sum_{i \neq j=1}^n |X_{ij}| - X_{ii}, \quad \forall i = 1, \dots, n,$$

d'où l'expression de  $\alpha_X$ .

En appliquant le même principe sur la matrice  $S^+ = S + \alpha \Delta S$ , on obtient l'expression de

$\alpha_S$ .

Finalement, on prend

$$\alpha = \rho \min(\alpha_X, \alpha_S), \quad 0 < \rho < 1.$$

D'où la preuve. ■

## 3.4 L'algorithme prototype et l'analyse de sa complexité

Dans cette section, on présente l'algorithme de notre approche, puis on donne quelques résultats qui servent pour l'étude de la convergence et la complexité de ce dernier mesurée respectivement par la monotonie de l'objectif et le nombre d'itérations total produit pour obtenir une solution optimale.

### 3.4.1 Algorithme de la méthode de trajectoire centrale

#### Début algorithme

**Initialisation :**  $\varepsilon > 0$  est une précision donnée,  $\sigma, \theta \in (0, 1)$ ,  $(X^0, y^0, S^0) \in T_\theta(\mu)$  et  $k = 0$ ;

**Tant que**  $\langle X^k, S^k \rangle \geq \varepsilon$  faire

1. Prendre

- $\mu^k = \frac{\langle X^k, S^k \rangle}{n}$ ;
- $H^k = 2\sigma\mu^k I - (X^k S^k + S^k X^k)$ ;

2. Calculer  $\Delta W = (\Delta X^k, \Delta y^k, \Delta S^k)$ ;

$$\begin{cases} \Delta X^k &= \mathcal{E}^{-1}(H^k - \mathcal{F} \Delta S^k); \\ \Delta y^k &= -(B^{-1} \mathcal{A} \mathcal{E}^{-1} H^k); \\ \Delta S^k &= -\mathcal{A}^* \Delta y^k; \end{cases}$$

3. Poser

$$\begin{cases} X^{k+1} &= X^k + \alpha^k \Delta X^k; \\ y^{k+1} &= y^k + \alpha^k \Delta y^k; \\ S^{k+1} &= S^k + \alpha^k \Delta S^k; \end{cases}$$

$\alpha^k$  est obtenu par l'une des quatres alternatives.

4. Prendre  $k = k + 1$ ;

**Fin tant que**

**Fin algorithme**

Dans l'algorithme,  $\mathcal{E}$  et  $\mathcal{F}$  sont des opérateurs de  $\mathbb{S}^n$  dans  $\mathbb{S}^n$  définis par

$$\mathcal{E}M = \frac{1}{2}(SM + MS) \text{ et } \mathcal{F}M = \frac{1}{2}(XM + MX).$$

### 3.4.2 Résultats de convergence

**Lemme 3.4.1** Soient  $(X, y, S) \in \mathbb{S}_{++}^n \times \mathbb{R}^m \times \mathbb{S}_{++}^n$  et  $\Delta W = (\Delta X, \Delta y, \Delta S)$  une solution du système (3.3), alors

i)  $\langle \Delta X, \Delta S \rangle = 0.$

ii)  $\langle X, \Delta S \rangle + \langle S, \Delta X \rangle = \text{tr}(H),$  où  $H = \sigma\mu I - \left(\frac{XS + SX}{2}\right)$  tels que  $\sigma \in (0, 1)$  et  $\mu = \frac{\langle X, S \rangle}{n}.$

iii)  $\langle (X + \alpha\Delta X), (S + \alpha\Delta S) \rangle = (1 - \alpha(1 - \sigma)) \langle X, S \rangle, \forall \alpha \in \mathbb{R}.$

**Preuve.**

i) Par les deux premières équations du système (3.3), on obtient

$$\langle \Delta X, \Delta S \rangle = \langle \Delta S, \Delta X \rangle = \left\langle -\left(\sum_{i=1}^m \Delta y_i A_i\right), \Delta X \right\rangle = -\sum_{i=1}^m \Delta y_i \langle A_i, \Delta X \rangle = 0.$$

ii) On a

$$\begin{aligned} 2\text{tr}(H) &= \text{tr}(2H) = \text{tr}(2\sigma\mu I - (XS + SX)) \\ &= \text{tr}(X\Delta S + S\Delta X + \Delta XS + \Delta SX) \\ &= \text{tr}((X\Delta S + \Delta XS) + (S\Delta X + \Delta SX)) \\ &= 2\text{tr}(X\Delta S + \Delta XS) \\ &= 2(\langle X, \Delta S \rangle + \langle \Delta X, S \rangle). \end{aligned}$$

iii) En utilisant i), ii) avec les égalités

$$H = \sigma\mu I - \left(\frac{XS + SX}{2}\right) \text{ et } \text{tr}(XS) = n\mu,$$

on obtient

$$\begin{aligned}
 \langle (X + \alpha\Delta X), (S + \alpha\Delta S) \rangle &= \langle X, S \rangle + \alpha(\langle X, \Delta S \rangle + \langle \Delta X, S \rangle) + \alpha^2 \langle \Delta X, \Delta S \rangle \\
 &= \langle X, S \rangle + \alpha \operatorname{tr} \left( \frac{X\Delta S + S\Delta X + \Delta X S + \Delta S X}{2} \right) \\
 &= \langle X, S \rangle + \alpha \operatorname{tr} \left( (\sigma\mu I) - \left( \frac{XS + SX}{2} \right) \right) \\
 &= \langle X, S \rangle + \alpha \operatorname{tr}(\sigma\mu I) - \alpha \operatorname{tr} \left( \frac{XS + SX}{2} \right) \\
 &= \langle X, S \rangle + \alpha \sigma n \mu - \alpha \langle X, S \rangle \\
 &= (1 - \alpha(1 - \sigma)) \langle X, S \rangle, \forall \alpha \in \mathbb{R}.
 \end{aligned}$$

■

**Lemme 3.4.2** Soit  $(X^+, y^+, S^+)$  une solution strictement réalisable, telle que

$$\begin{aligned}
 X^+ &= X + \alpha\Delta X \\
 y^+ &= y + \alpha\Delta y \\
 S^+ &= S + \alpha\Delta S,
 \end{aligned}$$

alors

$$\langle X^+, S^+ \rangle < \langle X, S \rangle.$$

**Preuve.** Du lemme précédent, on a

$$\begin{aligned}
 \langle X^+, S^+ \rangle &= \langle (X + \alpha\Delta X), (S + \alpha\Delta S) \rangle \\
 &= (1 - \alpha(1 - \sigma)) \langle X, S \rangle \\
 &< \langle X, S \rangle,
 \end{aligned}$$

puisque  $(1 - \alpha(1 - \sigma)) < 1$  ( $\alpha > 0$  et  $\sigma \in (0, 1)$ ). ■

Ce qui complète la preuve.

**Lemme 3.4.3** Soient  $X^+$  et  $X$  deux solutions strictement réalisables de  $(\text{PSD})_\mu$  avec  $X^+ = X + \alpha\Delta X$ , où  $\alpha$  est le pas de déplacement et  $\Delta X$  est la direction de Newton, alors  $f_\mu(X^+) < f_\mu(X)$ .

**Preuve.** On a

$$f_\mu(X^+) \simeq f_\mu(X) + \langle \nabla f_\mu(X), X^+ - X \rangle.$$

Donc

$$f_\mu(X^+) - f_\mu(X) \simeq \langle \nabla f_\mu(X), \alpha \Delta X \rangle.$$

Comme

$$\nabla f_\mu(X) = -\nabla^2 f_\mu(X) \Delta X,$$

on déduit que

$$\begin{aligned} f_\mu(X^+) - f_\mu(X) &\simeq \alpha \langle -\nabla^2 f_\mu(X) \Delta X, \Delta X \rangle \\ &\simeq -\alpha \langle \nabla^2 f_\mu(X) \Delta X, \Delta X \rangle < 0, \end{aligned}$$

tant que  $f_\mu$  est strictement convexe. D'où  $f_\mu(X^+) < f_\mu(X)$ . ■

### 3.4.3 Analyse de la complexité

Dans cette partie, on donne une preuve simple de la convergence polynomiale de la méthode de trajectoire centrale à court pas, en suivant les mêmes étapes de la preuve de Monteiro [38] où on prouve que la complexité de notre algorithme est bornée par  $\mathcal{O}\left(\sqrt{n} \ln\left[\varepsilon^{-1}(\langle X^0, S^0 \rangle)\right]\right)$ . Tout au long de cette partie, on donne quelques lemmes techniques qui seront fréquemment utilisés pendant l'analyse.

L'algorithme sélectionne une suite de pas de déplacement  $\{\alpha_k\}$  et de paramètres de centralité  $\{\sigma_k\}$ , d'après la règle suivante pour tout  $k \geq 0$ , soit  $\alpha_k = 1$  et  $\sigma_k = 1 - \frac{\delta}{\sqrt{n}}$ , où  $\delta > 0$  est une constante qui est spécifiée dans le théorème ci-dessous ; le résultat suivant analyse le comportement d'une itération de la méthode de trajectoire centrale à court pas.

**Lemme 3.4.4** Soit  $(X, y, S) \in \overset{0}{\mathcal{F}}(\text{PSD}) \times \overset{0}{\mathcal{F}}(\text{SDD})$  et soit  $(\Delta X, \Delta y, \Delta S)$  la solution du système (3.3) avec  $H = \sigma \mu I - \left(\frac{XS + SX}{2}\right)$ . Pour tout  $\alpha \in \mathbb{R}$ , on pose

$$(X(\alpha), y(\alpha), S(\alpha)) = (X, y, S) + \alpha (\Delta X, \Delta y, \Delta S), \quad (3.6)$$

$$\mu(\alpha) = \frac{\langle X(\alpha), S(\alpha) \rangle}{n} \quad (3.7)$$

$$Q(\alpha) = \frac{X(\alpha)S(\alpha) + S(\alpha)X(\alpha)}{2} - \mu(\alpha)I. \quad (3.8)$$

Alors,

$$Q(\alpha) + Q(\alpha)^T = 2(1 - \alpha) \left( \frac{XS + SX}{2} - \mu I \right) + \alpha^2 (\Delta X \Delta S + \Delta S \Delta X). \quad (3.9)$$

**Preuve.** Soit  $\alpha \in \mathbb{R}$  donné. Par le Lemme 3.4.1, *iii*), on a

$$\langle X(\alpha), S(\alpha) \rangle = (1 - \alpha + \alpha\sigma) \langle X, S \rangle,$$

alors

$$\mu(\alpha) = (1 - \alpha + \alpha\sigma) \mu.$$

D'où,

$$\begin{aligned} Q(\alpha) + Q(\alpha)^T &= X(\alpha)S(\alpha) + S(\alpha)X(\alpha) - 2\mu(\alpha)I \\ &= (X + \alpha\Delta X)(S + \alpha\Delta S) + (S + \alpha\Delta S)(X + \alpha\Delta X) \\ &\quad - 2(1 - \alpha + \alpha\sigma)\mu I \\ &= (1 - \alpha)(XS + SX - 2\mu I) + \underbrace{\alpha(XS + SX - 2\sigma\mu I)}_{-2\alpha H} \\ &\quad + \underbrace{2\alpha(\mathcal{E}\Delta X + \mathcal{F}\Delta S)}_{2\alpha H} + \alpha^2(\Delta X \Delta S + \Delta S \Delta X) \\ &= 2(1 - \alpha) \left( \frac{XS + SX}{2} - \mu I \right) + \alpha^2(\Delta X \Delta S + \Delta S \Delta X). \end{aligned}$$

■

Le lemme suivant fournit des bornes sur les directions  $\Delta X X^{-1}$  et  $X \Delta S$  pour  $(X, y, S) \in \overset{0}{\mathcal{F}}(PSD) \times \overset{0}{\mathcal{F}}(SDD)$ .

Dans la suite, on aura souvent besoin d'utiliser les deux inégalités suivantes décrites dans [33]

$$\|A_1 A_2\|_F \leq \|A_1\| \|A_2\|_F \text{ et } \|A_1 A_2\|_F \leq \|A_1\|_F \|A_2\|, \quad \forall A_1, A_2 \in \mathcal{M}_n(\mathbb{R}).$$

**Lemme 3.4.5** Soit  $(X, y, S) \in \overset{0}{\mathcal{F}}(PSD) \times \overset{0}{\mathcal{F}}(SDD)$  telle que  $\|XS - \mu I\| \leq \theta\mu$ , pour  $\theta \in [0, 1)$  et  $\mu > 0$ .



### 3.4. L'algorithme prototype et l'analyse de sa complexité

Supposons que  $(\Delta X, \Delta y, \Delta S) \in \mathbb{S}^n \times \mathbb{R}^m \times \mathbb{S}^n$  est une solution du système (3.3) pour  $H \in \mathbb{R}^{n \times n}$  et soit  $\delta_X = \mu \|\Delta X X^{-1}\|_F$ ,  $\delta_S = \|X \Delta S\|_F$ . Alors,

$$\delta_X \delta_S \leq \frac{1}{2} (\delta_X^2 + \delta_S^2) \leq \frac{\|H\|_F}{2(1-\theta)^2}.$$

**Preuve.** En utilisant la dernière équation du système (3.3), on obtient

$$\begin{aligned} H &= \frac{\Delta S X + X \Delta S + \Delta X S + S \Delta X}{2} \\ &= \frac{\Delta S X + X \Delta S}{2} + \frac{\mu(\Delta X X^{-1} + X^{-1} \Delta X)}{2} + \frac{1}{2} \Delta X X^{-1} (X S - \mu I) \\ &\quad + \frac{1}{2} (S X - \mu I) X^{-1} \Delta X, \end{aligned}$$

il s'ensuit que

$$\begin{aligned} \|H\|_F &\geq \left\| \frac{\Delta S X + X \Delta S}{2} + \frac{\mu(\Delta X X^{-1} + X^{-1} \Delta X)}{2} \right\|_F - \|X S - \mu I\| \|\Delta X X^{-1}\|_F \\ &\geq \left( \left\| \frac{\Delta S X + X \Delta S}{2} \right\|_F^2 + \left\| \frac{\mu(\Delta X X^{-1} + X^{-1} \Delta X)}{2} \right\|_F^2 \right)^{\frac{1}{2}} - \theta \mu \left( \frac{\delta_X}{\mu} \right) \\ &\geq \left( \|X \Delta S\|_F^2 + \mu^2 \|\Delta X X^{-1}\|_F^2 \right)^{\frac{1}{2}} - \theta \delta_X \\ &= (\delta_X^2 + \delta_S^2)^{\frac{1}{2}} - \theta \delta_X \geq (1-\theta) (\delta_X^2 + \delta_S^2)^{\frac{1}{2}}, \end{aligned}$$

où la deuxième inégalité est une conséquence de l'hypothèse  $\|X S - \mu I\| \leq \theta \mu$ . En utilisant le Lemme 3.4.1, *i*), on montre que  $\langle \Delta S X + X \Delta S, \Delta X X^{-1} + X^{-1} \Delta X \rangle = \langle \Delta X, \Delta S \rangle = 0$ .

Le résultat se déduit facilement de la dernière inégalité. ■

**Théorème 3.4.6** Soient  $\theta \in (0, 1)$  et  $\delta \in [0, \sqrt{n})$  des constantes satisfaisant

$$\frac{\theta^2 + \delta^2}{2(1-\theta)^2 \left(1 - \frac{\delta}{\sqrt{n}}\right)} \leq \theta, \text{ avec } \theta \leq \frac{1}{2}. \quad (3.10)$$

Supposons que  $(X, y, S) \in T_\theta(\mu)$  et soit  $(\Delta X, \Delta y, \Delta S)$  la solution du système (3.3) avec  $H = \sigma \mu I - \left(\frac{X S + S X}{2}\right)$  et  $\sigma = 1 - \frac{\delta}{\sqrt{n}}$ . Alors,

(a)  $(X^+, y^+, S^+) = (X, y, S) + (\Delta X, \Delta y, \Delta S) \in T_\theta(\mu^+),$

(b)  $\langle X^+, S^+ \rangle = \left(1 - \frac{\delta}{\sqrt{n}}\right) \langle X, S \rangle.$

**Preuve.**

(b) est une conséquence immédiate du Lemme 3.4.1, iii) avec  $\alpha = 1$  et  $\sigma = 1 - \frac{\delta}{\sqrt{n}}$ . D'où,

$$\mu^+ = \frac{\langle X^+, S^+ \rangle}{n} = \left(1 - \frac{\delta}{\sqrt{n}}\right) \mu. \quad (3.11)$$

(a) Comme  $\langle \frac{XS+SX}{2} - \mu I, I \rangle = 0$  et  $(X, y, S) \in T_\theta(\mu)$ , on obtient

$$\begin{aligned} \left\| \sigma \mu I - \frac{XS + SX}{2} \right\|_F^2 &= \|(\sigma - 1)\mu I\|_F^2 + \left\| \mu I - \frac{XS + SX}{2} \right\|_F^2 \\ &\leq [(1 - \sigma)^2 n + \theta^2] \mu^2 = (\delta^2 + \theta^2) \mu^2. \end{aligned} \quad (3.12)$$

Tant que  $\|XS - \mu I\| \leq \theta \mu$ , il s'ensuit du lemme 3.4.5 avec  $H = \sigma \mu I - \left(\frac{XS+SX}{2}\right)$ , que

$$\|\Delta X X^{-1}\|_F \leq \frac{\left\| \sigma \mu I - \frac{XS+SX}{2} \right\|_F}{(1 - \theta) \mu}, \quad (3.13)$$

et

$$\|\Delta X X^{-1}\|_F \|X \Delta S\|_F \leq \frac{\left\| \sigma \mu I - \frac{XS+SX}{2} \right\|_F^2}{2(1 - \theta)^2 \mu}. \quad (3.14)$$

Soit  $Q^+ = Q(1) = \frac{X^+ S^+ + S^+ X^+}{2} - \mu^+ I$ . En utilisant (3.9) avec  $\alpha = 1$ , (3.14), (3.12), (3.10) et

(3.11), on obtient

$$\begin{aligned} \frac{1}{2} \|Q^+ + (Q^+)^T\|_F &= \|Q^+\|_F = \frac{1}{2} \|\Delta X \Delta S + \Delta S \Delta X\|_F \leq \|\Delta X \Delta S\|_F \\ &\leq \|\Delta X X^{-1}\|_F \|X \Delta S\|_F \leq \frac{\left\| \sigma \mu I - \frac{XS+SX}{2} \right\|_F^2}{2(1 - \theta)^2 \mu} \\ &\leq \frac{(\delta^2 + \theta^2) \mu}{2(1 - \theta)^2} \leq \theta \left(1 - \frac{\delta}{\sqrt{n}}\right) \mu = \theta \mu^+, \end{aligned}$$

donc, on conclut que

$$\left\| \frac{X^+ S^+ + S^+ X^+}{2} - \mu^+ I \right\|_F \leq \theta \mu^+. \quad (3.15)$$

En utilisant (3.13), (3.12) et (3.10), on obtient

$$\begin{aligned} \|\Delta X X^{-1}\|_F &\leq \frac{\left\| \sigma \mu I - \frac{XS+SX}{2} \right\|_F}{(1 - \theta) \mu} \leq \frac{(\delta^2 + \theta^2)^{\frac{1}{2}}}{(1 - \theta)} \\ &\leq \left[ 2\theta \left(1 - \frac{\delta}{\sqrt{n}}\right) \right]^{\frac{1}{2}} < 1, \end{aligned}$$

puisque  $\theta \leq \frac{1}{2}$  et  $0 < 1 - \frac{\delta}{\sqrt{n}} < 1$ .

Il est facile de voir que la dernière relation implique que  $I + \Delta X X^{-1} > 0$ , d'où  $X^+ = X + \Delta X = (I + \Delta X X^{-1}) X > 0$ . En utilisant la première équation du système (3.3), on obtient

$$\mathcal{A}X^+ = \mathcal{A}(X + \Delta X) = \mathcal{A}X + \mathcal{A}\Delta X = b.$$

Et par conséquent  $X^+ \in \overset{0}{\mathcal{F}}(PSD)$ .

L'inégalité (3.15) implique que

$$\lambda_{\min}\left(\frac{X^+S^+ + S^+X^+}{2}\right) = \lambda_{\min}(X^+S^+) \geq (1 - \theta)\mu^+ > 0,$$

d'où  $X^+S^+ = (X^+)^{\frac{1}{2}}(X^+)^{\frac{1}{2}}S^+(X^+)^{\frac{1}{2}}(X^+)^{-\frac{1}{2}} > 0$ , ce qui donne  $(X^+)^{\frac{1}{2}}S^+(X^+)^{\frac{1}{2}} > 0$ , donc  $S^+ > 0$ . D'après la deuxième équation du système (3.3), on a

$$\mathcal{A}^*y^+ + S^+ = \mathcal{A}^*(y + \Delta y) + (S + \Delta S) = \mathcal{A}^*y + S + \mathcal{A}^*\Delta y + \Delta S = C,$$

ce qui implique  $(y^+, S^+) \in \overset{0}{\mathcal{F}}(SDD)$ .

Donc, on conclut que  $(X^+, y^+, S^+) \in T_\theta(\mu^+)$ . ■

**Théorème 3.4.7** Soient  $\theta \in (0, 1)$ ,  $\theta' = (1 + \alpha)\theta$  et  $\delta \in [0, \sqrt{n})$ .

Supposons que  $(X, y, S) \in T_\theta(\mu)$  et soit  $(\Delta X, \Delta y, \Delta S)$  la solution du système (3.3) avec  $H = \sigma\mu I - \left(\frac{XS+SX}{2}\right)$  et  $\sigma = 1 - \frac{\delta}{\sqrt{n}}$ . Alors,

$$(a) (X^+, y^+, S^+) = (X, y, S) + \alpha(\Delta X, \Delta y, \Delta S) \in T_{\theta'}(\mu),$$

$$(b) \langle X^+, S^+ \rangle = \left(1 - \alpha\frac{\delta}{\sqrt{n}}\right)\langle X, S \rangle.$$

**Preuve.**

(b) est une conséquence immédiate du Lemme 3.4.1, iii) avec  $\sigma = 1 - \frac{\delta}{\sqrt{n}}$ . D'où,

$$\mu^+ = \frac{\langle X^+, S^+ \rangle}{n} = \left(1 - \alpha\frac{\delta}{\sqrt{n}}\right)\mu. \quad (3.16)$$

(a) On a  $X^+ = X + \alpha\Delta X$  et  $S^+ = S + \alpha\Delta S$ , donc

$$\begin{aligned}
\left\| \frac{X(\alpha)S(\alpha) + S(\alpha)X(\alpha)}{2} - \mu(\alpha)I \right\|_F &= \left\| \frac{(X + \alpha\Delta X)(S + \alpha\Delta S) + (S + \alpha\Delta S)(X + \alpha\Delta X)}{2} - (1 - \alpha + \alpha\sigma)\mu I \right\|_F \\
&= \left\| \frac{XS + SX}{2} + \alpha \left( \frac{X\Delta S + \Delta XS + S\Delta X + \Delta SX}{2} \right) - (1 - \alpha + \alpha\sigma)\mu I \right\|_F \\
&= \left\| \frac{XS + SX}{2} - \mu I + \alpha \left( \frac{X\Delta S + \Delta XS + S\Delta X + \Delta SX}{2} + (1 - \sigma)\mu I \right) \right\|_F \\
&= \left\| \frac{XS + SX}{2} - \mu I + \alpha \left( \sigma\mu I - \frac{XS + SX}{2} + (1 - \sigma)\mu I \right) \right\|_F && \text{En utilisant} \\
&= \left\| \frac{XS + SX}{2} - \mu I + \alpha \left( \mu I - \frac{XS + SX}{2} \right) \right\|_F \\
&= \left\| \frac{XS + SX}{2} - \mu I \right\|_F + \alpha \left\| \mu I - \frac{XS + SX}{2} \right\|_F \\
&= (1 + \alpha) \left\| \mu I - \frac{XS + SX}{2} \right\|_F \\
&\leq (1 + \alpha)\theta\mu = \theta'\mu.
\end{aligned}$$

la première équation du système (3.3), on obtient

$$\mathcal{A}X^+ = \mathcal{A}(X + \alpha\Delta X) = \mathcal{A}X + \alpha\mathcal{A}\Delta X = b.$$

Et par conséquent  $X^+ \in \overset{0}{\mathcal{F}}(PSD)$ .

D'après la deuxième équation du système (3.3), on a

$$\mathcal{A}^*y^+ + S^+ = \mathcal{A}^*(y + \alpha\Delta y) + (S + \alpha\Delta S) = \mathcal{A}^*y + S + \alpha(\mathcal{A}^*\Delta y + \Delta S) = C,$$

ce qui implique  $(y^+, S^+) \in \overset{0}{\mathcal{F}}(SDD)$ .

Donc, on conclut que  $(X^+, y^+, S^+) \in T_{\theta'}(\mu)$ . ■

**Corollaire 3.4.8** Soient  $\theta$  et  $\delta$  donnés comme dans le Théorème 3.4.7 et  $(X^0, y^0, S^0) \in T_\theta(\mu)$ . Alors la méthode de trajectoire centrale à court pas génère une suite de points  $\{(X^k, y^k, S^k)\} \subset T_\theta(\mu)$  telle que

$$\langle X^k, S^k \rangle = \left(1 - \alpha \frac{\delta}{\sqrt{n}}\right)^k \langle X^0, S^0 \rangle \quad \forall k \geq 0.$$

De plus, pour une précision donnée  $\varepsilon > 0$ , la méthode de trajectoire centrale à court pas trouve un point  $(X^k, y^k, S^k)$  vérifiant  $\langle X^k, S^k \rangle \leq \varepsilon$  dans au plus  $\mathcal{O}\left(\sqrt{n} \ln \left[\varepsilon^{-1}(\langle X^0, S^0 \rangle)\right]\right)$  itérations.

**Preuve.** Déterminer  $k$  qui vérifie  $\langle X^k, S^k \rangle \leq \varepsilon$ , équivalent à trouver  $k$  tel que

$$\left(1 - \alpha \frac{\delta}{\sqrt{n}}\right)^k \langle X^0, S^0 \rangle \leq \varepsilon, \forall \alpha > 0. \tag{3.17}$$

1. Pour  $0 < \alpha < 1$ , l'inégalité (3.10) est équivalente à

$$k \ln \left( 1 - \alpha \frac{\delta}{\sqrt{n}} \right) \leq \ln \left( \frac{\varepsilon}{\langle X^0, S^0 \rangle} \right),$$

et comme

$$\ln(1 - x) \geq -x, \quad 0 < x = \alpha \frac{\delta}{\sqrt{n}} < 1,$$

cela est vérifié si

$$k \left( -\alpha \frac{\delta}{\sqrt{n}} \right) \leq \ln \left( \frac{\varepsilon}{\langle X^0, S^0 \rangle} \right),$$

ou, de manière équivalente

$$k \geq \left[ (\alpha \delta)^{-1} \sqrt{n} \ln \left( \frac{\langle X^0, S^0 \rangle}{\varepsilon} \right) \right].$$

2. Pour  $\alpha \geq 1$ , de l'inégalité (3.10), on a

$$\left( 1 - \alpha \frac{\delta}{\sqrt{n}} \right)^k \langle X^0, S^0 \rangle \leq \left( 1 - \frac{\delta}{\sqrt{n}} \right)^k \langle X^0, S^0 \rangle \leq \varepsilon,$$

donc trouver  $k$  qui vérifie cette inégalité, équivalent à trouver  $k$  tel que

$$k \ln \left( 1 - \frac{\delta}{\sqrt{n}} \right) \leq \ln \left( \frac{\varepsilon}{\langle X^0, S^0 \rangle} \right),$$

et comme

$$\ln(1 - x) \geq -x, \quad 0 < x = \frac{\delta}{\sqrt{n}} < 1,$$

cela est vérifié si

$$k \left( -\frac{\delta}{\sqrt{n}} \right) \leq \ln \left( \frac{\varepsilon}{\langle X^0, S^0 \rangle} \right),$$

ou, de manière équivalente

$$k \geq \left[ \delta^{-1} \sqrt{n} \ln \left( \frac{\langle X^0, S^0 \rangle}{\varepsilon} \right) \right].$$

Ce qui complète la preuve. ■

## 3.5 Mise en oeuvre de la méthode de trajectoire centrale pour (PSD)

Les exemples suivants sont pris de la littérature (voir par exemple [11, 25]) et implémentés sur MATLAB R2008b sur Intel® Core i3 (1.80 GHz) avec 4.00 Go RAM. On a pris  $\varepsilon = 1.35e - 006$ ,  $\sigma = 0.1$  et  $\rho = 0.99$ .

Dans le tableau des résultats, (ex  $(m, n)$ ) représente la taille de l'exemple, (**Itr**) représente le nombre d'itérations nécessaire pour obtenir une solution optimale et (**Pas**) représente l'intervalle des pas de déplacement pour chaque alternative (notée par (**Alt**)). On compare nos résultats obtenus avec le code CVX standard noté par (**CVX**). Rappelons que le problème primal considéré est

$$(PSD) \begin{cases} \min \langle C, X \rangle \\ \langle A_i, X \rangle = b_i, \quad i = 1, \dots, m, \\ X \in \mathbb{S}_+^n, \end{cases}$$

où, son dual associé est

$$(SDD) \begin{cases} \max b^T y \\ C - \sum_{i=1}^m y_i A_i = S, \\ y \in \mathbb{R}^m, S \in \mathbb{S}_+^n. \end{cases}$$

### 3.5.1 Exemples à taille fixe

#### Exemple 3.5.1

$$C = \begin{pmatrix} -1 & -1 \\ -1 & -1 \end{pmatrix}, \quad A_1 = \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}, \quad A_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \text{et } b = (1, 1)^T.$$

La solution primale-duale strictement réalisable initiale prise est

$$X^0 = \begin{pmatrix} 0.5 & 0 \\ 0 & 1.5 \end{pmatrix}, \quad y^0 = (0, -3)^T \quad \text{et } S^0 = \begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix}.$$

La solution optimale primale-duale trouvée est

$$X^* = \begin{pmatrix} 1 & 0 \\ 0 & 0.99999 \end{pmatrix}, \quad y^* = \begin{pmatrix} 1 \\ -2 \end{pmatrix} \text{ et } S^* = 10^{-7} \times \begin{pmatrix} 0.10558 & 0 \\ 0 & 0.10558 \end{pmatrix}.$$

La valeur optimale est  $\langle C, X^* \rangle = b^T y^* = -2$ .

### Exemple 3.5.2

$$C = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad A_1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad A_2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad b = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}.$$

La solution primale-duale strictement réalisable initiale prise est

$$X^0 = \begin{pmatrix} 0.25 & 0 & 0 \\ 0 & 0.25 & 0 \\ 0 & 0 & 0.5 \end{pmatrix}, \quad S^0 = \begin{pmatrix} 3 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad y^0 = \begin{pmatrix} -1 \\ -1 \end{pmatrix}.$$

La solution optimale primale-duale trouvée est

$$X^* = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0.99999 \end{pmatrix}, \quad S^* = \begin{pmatrix} 1.39603 & 0 & 0 \\ 0 & 0.60397 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad y^* = \begin{pmatrix} -0.39603 \\ 0 \end{pmatrix}.$$

La valeur optimale est  $\langle C, X^* \rangle = b^T y^* = 0$ .

### Exemple 3.5.3

$$C = \begin{pmatrix} 5 & 0 & 0 & 0 \\ 0 & 8 & 0 & 0 \\ 0 & 0 & 8 & 0 \\ 0 & 0 & 0 & 5 \end{pmatrix}, \quad A_4 = I, \quad b = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 2 \end{pmatrix},$$

et les matrices  $A_k, k = 1, \dots, 3$ , sont définies comme suit

$$A_k(i, j) = \begin{cases} 1 & \text{si } i = j = k \text{ ou } i = j = k + 1, \\ -1 & \text{si } i = k, j = k + 1 \text{ ou } i = k + 1, j = k, \\ 0 & \text{ailleurs.} \end{cases}$$

La solution primale-duale strictement réalisable initiale prise est

$$X^0 = \frac{1}{2}I, \quad y^0 = \begin{pmatrix} 1.5 \\ 1.5 \\ 1.5 \\ 1.5 \end{pmatrix} \quad \text{et} \quad S^0 = \begin{pmatrix} 2 & 1.5 & 0 & 0 \\ 1.5 & 3.5 & 1.5 & 0 \\ 0 & 1.5 & 3.5 & 1.5 \\ 0 & 0 & 1.5 & 2 \end{pmatrix}.$$

La solution optimale primale-duale trouvée est

$$X^* = \begin{pmatrix} 0.75 & 0 & 0 & 0 \\ 0 & 0.25 & -0.25 & 0 \\ 0 & -0.25 & 0.25 & 0 \\ 0 & 0 & 0 & 0.75 \end{pmatrix}, \quad y^* = \begin{pmatrix} 0 \\ 1.5 \\ 0 \\ 5 \end{pmatrix} \quad \text{et} \quad S^* = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 1.5 & 1.5 & 0 \\ 0 & 1.5 & 1.5 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

La valeur optimale est  $\langle C, X^* \rangle = b^T y^* = 11.5$ .

#### Exemple 3.5.4

$$C = \begin{pmatrix} -4 & 0 & 0 & 0 & 0 & 0 \\ 0 & -2 & 0 & 0 & 0 & 0 \\ 0 & 0 & -2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}, \quad A_1 = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix},$$

$$A_2 = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}, \quad A_3 = \begin{pmatrix} 2 & 0 & 0 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \quad \text{et} \quad b = \begin{pmatrix} 6 \\ 2 \\ 4 \end{pmatrix}.$$

La solution primale-duale strictement réalisable initiale prise est

$$X^0 = \begin{pmatrix} 1.467 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0.087 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0.36 & 0 & 0 & 0 \\ 0 & 0 & 0 & 4.26 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0.086 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0.532 \end{pmatrix}, \quad y^0 = \begin{pmatrix} -1 \\ -1 \\ -2 \end{pmatrix}$$



et

$$S^0 = \begin{pmatrix} 2 & 0 & 0 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 2 \end{pmatrix}.$$

La solution optimale primale-duale trouvée est

$$X^* = \begin{pmatrix} 1.99415 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 4.00251 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}, \quad y^* = \begin{pmatrix} 0 \\ -2.41354 \\ -0.79323 \end{pmatrix}$$

et

$$S^* = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1.20677 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 2.41354 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0.79323 \end{pmatrix}.$$

La valeur optimale est  $\langle C, X^* \rangle = b^T y^* = -8$ .

Le nombre d'itérations et l'intervalle des pas de déplacement pour chaque alternative dans les exemples précédents sont résumés dans les tableaux suivants

ex ( $m, n$ )	1 <sup>ère</sup> Alt		2 <sup>ème</sup> Alt	
	Itr	Pas	Itr	Pas
3.5.1 (2, 2)	5	[1.05, 1.06]	5	[1.05, 1.06]
3.5.2 (2, 3)	11	[0.69, 0.82]	3	[0.29, 1.34]
3.5.3 (4, 4)	14	[0.7, 0.78]	7	0.99
3.5.4 (3, 6)	18	[0.51, 0.7]	8	0.99

3 <sup>ème</sup> Alt		4 <sup>ème</sup> Alt		CVX	
Itr	Pas	Itr	Pas	Itr	Pas
5	[0.99, 1.09]	5	[0.99, 1.09]	6	[0.989, 1]
7	0.99	7	[0.99, 1.08]	5	[0.93, 1]
7	0.99	7	[0.99, 1.08]	5	[0.93, 1]
8	0.99	6	[0.99, 1.1]	7	[0.92, 1]

### 3.5.2 Exemple à taille variable

$n = 2m$ ,  $C = -I$ ,  $b(i) = 2$ ,  $i = 1, \dots, m$ , et les matrices  $A_k$ ,  $k = 1, \dots, m$ , sont définies comme suit

$$A_k(i, j) = \begin{cases} 1 & \text{si } i = j = k, \\ 1 & \text{si } i = j \text{ et } i = m + k, \\ 0 & \text{ailleurs,} \end{cases}$$

La solution primale-duale strictement réalisable initiale prise est

$$X^0(i, j) = \begin{cases} 1.5 & \text{si } i \leq j \\ 0.5 & \text{si } i > j \end{cases}, y^0(i) = -2, i = 1, \dots, m, \text{ et } S^0 = I.$$

La solution optimale primale-duale trouvée pour  $m \leq 150$  est

$$X^* = I, y^* = -1 \text{ et } S^* = 0_{\mathbb{R}^{n \times n}}.$$

La valeur optimale pour  $m \leq 150$  est

$$\langle C, X^* \rangle = b^T y^* = -n = -2m.$$

Les tableaux suivants résument les résultats obtenus qui concernent le nombre d'itérations et l'intervalle des pas de déplacement pour chaque alternative en considérant différentes

tailles

taille ( $m, n$ )	1 <sup>ère</sup> Alt		2 <sup>ème</sup> Alt	
	Itr	Pas	Itr	Pas
(10, 20)	4	1.099	4	1.099
(20, 40)	4	1.099	4	1.099
(50, 100)	4	1.099	4	1.099
(100, 200)	5	1.099	5	1.099
(150, 300)	6	[0.88, 1.099]	6	[0.99, 1.099]

3 <sup>ème</sup> Alt		4 <sup>ème</sup> Alt		CVX	
Itr	Pas	Itr	Pas	Itr	Pas
8	[0.99, 1.099]	4	1.1	6	[0.989, 1]
8	0.99	4	1.1	6	[0.989, 1]
9	0.99	4	1.1	6	[0.989, 1]
9	0.99	5	1.1	7	[0.989, 1]
9	0.99	5	1.1	7	[0.989, 1]

### 3.5.3 Commentaires

À travers les tests numériques réalisés, les quatre alternatives offrent une solution optimale de (PSD) et (SDD) avec un nombre d'itérations raisonnable.

On constate aussi que la quatrième alternative est la meilleure et que les résultats numériques comparatifs obtenus la favorisent vis-à-vis du logiciel classique CVX. C'est probablement dû à la nature du logiciel CVX qui base sur l'utilisation d'un point initial non réalisable.

# Conclusion

Dans ce travail, on s'est intéressé à résoudre le problème de programmation semi-définie ( $PSD$ ) par la méthode de trajectoire centrale. On a associé à ( $PSD$ ) un problème perturbé, noté  $(PSD)_\mu$ . En premier lieu, on a montré l'existence et l'unicité de la solution optimale du problème perturbé  $(PSD)_\mu$  et que celle-ci converge vers la solution optimale du problème originel ( $PSD$ ) quand  $\mu$  tend vers zéro. Puis, on a prouvé le décroissement de la fonction objectif sur la suite déterminée par notre algorithme.

Comme le problème  $(PSD)_\mu$  est strictement convexe, les conditions de **KKT** sont nécessaires et suffisantes. Ceci nous a permis d'utiliser la méthode de Newton pour calculer la direction de descente et déterminer un nouveau itéré mieux que l'ancien itéré.

Pour calculer le pas de déplacement, plusieurs méthodes ont été proposées. À titre d'exemple celles de recherche linéaire qui sont très coûteuses. Pour remédier ce problème, on a proposé une nouvelle approche : on donne quatre nouvelles alternatives pour calculer le pas de déplacement par une technique simple, facile et moins coûteuse. Pour en finir, on a analysé la convergence de l'algorithme à court pas obtenu et prouvé que sa complexité est bornée par  $O\left(\sqrt{n} \ln\left[\varepsilon^{-1}(\langle X^0, S^0 \rangle)\right]\right)$  itérations.

Pour valoriser notre contribution, illustrer l'efficacité de notre approche et confirmer la convergence des quatre alternatives vers la solution optimale du problème ( $PSD$ ), on a présenté des simulations numériques pour illustrer. Les résultats obtenus confirment

les propos théoriques et montrent la supériorité de la quatrième alternative par rapport aux autres, mesurée par le nombre d'itérations.

L'ensemble de ces résultats sont publiés dans "Journal of Computational and Applied Mathematics" [51].

# Bibliographie

- [1] F. Alizadeh, Combinatorial optimization with interior point methods and semidefinite matrices, Ph. D. thesis, University of Minnesota, Minneapolis, MN, Oct. 1991.
- [2] F. Alizadeh, Interior point methods in semidefinite programming with applications to combinatorial optimization, *SIAM J. Optim.* 5 (1995) 13–51.
- [3] F. Alizadeh, J. P. A. Haeberly, M. L. Overton, Primal-dual interior-point methods for semidefinite programming : Convergence rates, stability and numerical results, *SIAM J. Optim.* 8 (1998) 746–768.
- [4] M. S. Babynin, V. G. Zhadan, A primal interior point method for linear semidefinite programming problem, *Comput. Math. Phy.* 48 (2008) 1746–1767.
- [5] Y. Q. Bai, M. El Ghami, C. Roos, A comparative study of kernel function for primal-dual interior-point algorithms in linear optimization, *SIAM J. Optim.* 15 (2004) 101–128.
- [6] D. Benterki, J. P. Crouzeix, B. Merikhi, A numerical implementation of an interior point method for semidefinite programming, *Pesquisa Operacional.* 23 (1) Janeiro (2003) 49–59.
- [7] D. Benterki, J. P. Crouzeix, B. Merikhi, A numerical feasible interior point method for linear semidefinite programs, *RAIRO – Oper. Res.* 41 (2007) 49–59.

- 
- [8] M. Bouafia, D. Benterki, A. Yassine, An efficient primal-dual interior point method for linear programming problems based on a new kernel function with a trigonometric barrier term, *J. Optim. Theory Appl.* 170 (2) (2016) 528–545.
- [9] N. Brixius, F. A. Potra, R. Sheng, Nonsymmetric search directions for semidefinite programming, *SIAM J. Optim.* 9 (1999) 863–876.
- [10] B. K. Choi, G. M. Lee, On complexity analysis of the primal-dual interior-point method for semidefinite optimization problem based on a new proximity function. *Nonlinear Anal.* 71 (2009) 2628–2640.
- [11] J. P. Crouzeix, B. Merikhi, A logarithm barrier method for semidefinite programming, *RAIRO – Oper. Res.* 42 (2008) 123–139.
- [12] I. I. Dikin, Iterative solution of problems of linear and quadratic programming, *Doklady Akademii Nauk SSSR.* 174 (1967) 747–748.
- [13] M. ElGhami, C. Roos, T. Steihauga, A generic primal-dual interior point method for semidefinite optimization based on a new class of kernel functions, *Optimization Methods Software.* 25 (3) (2010) 387–403.
- [14] M. Grötschel, L. Lovász, A. Schrijver, *Geometric algorithms and combinatorial optimization*, vol. 2, Algorithms and combinatorics, Springer-Verlag, New York, Berlin, 1998.
- [15] M. Halicka, E. D. Klerk, C. Roos, On the convergence of the central path in semidefinite optimization, *SIAM J. Optim.* 12 (2002) 1090–1099.
- [16] C. Helmberg, F. Rendl, R. J. Vanderbei, H. Wolkowicz, An interior-point method for semidefinite programming, *SIAM J. Optim.* 6 (1996) 342–361.
- [17] C. Helmberg, *Semidefinite programming for combinatorial optimization*, Konrad-Zuse-Zentrum für Informationstechnik Berlin, Takustrabe 7, D-14195 Berlin, Germany, 2000.

- [18] R. A. Horn, C. R. Johnson, *Matrix analysis*, Cambridge University Press, New York, 1985.
- [19] B. Jansen, C. Roos, T. Terlaky, A family of polynomial affine scaling algorithms for positive semidefinite linear complementarity problems, *SIAM J. Optim.* 7 (1997) 126–140.
- [20] J. Ji, F. A. Potra and R. Sheng, On the local convergence of a predictor-corrector method for semidefinite programming, *SIAM J. Optim.* 10 (1999) 195–210.
- [21] C. Kanzow, C. Nagel, Technical note some structural properties of a Newton-type method for semidefinite programs, *J. Optim. Theory Appl.* 122 (2004) 219–226.
- [22] C. Kanzow, C. Nagel, Quadratic convergence of a nonsmooth Newton-type method for semidefinite programs without strict complementarity, *SIAM J. Optim.* 15 (2005) 654–672.
- [23] N. K. Karmarkar, A new polynomial time algorithm for linear programming, *Combinatorica.* 4 (1984) 373–395.
- [24] Z. Kebbiche, D. Benterki, A weighted path-following method for linearly constrained convex programming, *Revue Roumaine de Mathématiques Pures et appliquées.* 57 (2012) 245–256.
- [25] A. Keraghel, *Etude adaptative et comparative des principales variantes dans l’algorithme de Karmarkar* (Ph.D. thesis), Joseph Fourier University, Grenoble, France, 1989.
- [26] A. Keraghel, *Analyse convexe, Théorie fondamentale et exercices*, Université de Sétif, Faculté des sciences, Laboratoire de Mathématique Fondamentale et Numérique, 2001.
- [27] S. Kettab, D. Benterki, A relaxed logarithmic barrier method for semidefinite programming, *RAIRO-Oper. Res.* 49 (2015) 555–568.



- [28] M. Kojima, S. Shindoh, S. Hara, Interior-point methods for the monotone semidefinite linear complementarity problem in symmetric matrices, *SIAM J. Optim.* 7 (1997) 86–125.
- [29] M. H. Koulaei, T. Terlaky, On the extension of a Mehrotra-type algorithm for semidefinite optimization, Technical Report 2007/4, Advanced optimization Lab. Department of Computing and Software, McMaster University, Hamilton, Ontario, Canada.
- [30] I. Krim, Optimization sous contraintes de semi définie positivité, Thèse de Magister, Université d’Oran, Algérie, 2014.
- [31] C. Liu, H. Liu, A new second-order corrector interior-point algorithm for semidefinite programming, *Math Meth Oper Res.* 75 (2012) 165–183.
- [32] I. J. Lustig, R. E. Marsten, D. F. Shanno, Interior point methods : Computational state of the art, *ORSA Journal on Computing*, 6 (1994) 1–15.
- [33] H. Lütkepohl, *Handbook of Matrices*, Humboldt-Universität zu Berlin, Germany, 1996.
- [34] B. Mohar, S. Poljak, D. F. Shanno, Interior point methods for linear programming : Computational state of the art, *ORSA J. Comput.* 6 (1994) 1–14.
- [35] M. Minoux, *Programmation Mathématique, Théorie et Algorithmes*, Bordas, 1983.
- [36] R. D. C. Monteiro, I. Adler, M. G. C. Resende, A polynomial time primal dual affine scaling algorithm for linear and convex quadratic programming and its power series extension, *Mathematics of Operations Research.* 15 (1990) 191–214.
- [37] R. D. C. Monteiro, Polynomial convergence of primal-dual algorithms for semidefinite programming based on the Monteiro and Zhang family of directions, *SIAM J. Optim.* 8 (1998) 797–812.

- [38] R. D. C. Monteiro, Primal-dual path-following for semidefinite programming, *SIAM J. Optim.* 7 (3) (1997) 663–678.
- [39] R. D. C. Monteiro, T. Tsuchiya, Polynomial convergence of a new family of primal-dual algorithms for semidefinite programming, *SIAM J. Optim.* 9 (1999) 551–577.
- [40] R. D. C. Monteiro, P. R. Zanjácomo, A note on the existence of alizadeh-Heaberly-Overton direction for semidefinite programming, *Math. Program.* 78 (1997) 393–396.
- [41] Y. E. Nesterov and A. S. Nemirovsky, Interior point methods in convex programming : theory and applications, SIAM, Philadelphia, PA, 1994.
- [42] Y. E. Nesterov, A. S. Nemirovsky, Interior-polynomial methods in convex programming, vol. 13, Studies in applied mathematics, Society for Industrial and applied mathematics, Philadelphia, PA, 1994.
- [43] M. Overton, Large-scale optimization of eigenvalues, *SIAM J. Optim.* 2 (1992) 88–120.
- [44] S. Poljak, F. Rendl, H. Wolkowicz, A recipe for best semidefinite relaxation for (0,1)-quadratic programming, Tech. report CORR 94-7, University
- [45] F. A. Potra, R. Sheng, Superlinear convergence of interior-point algorithms for semidefinite programming, *J. Optim. Theory Appl.* 99 (1998) 103–119.
- [46] R. T. Rockafellar, *Convex Analysis*, Princeton University Press, Princeton, New Jersey, 1970.
- [47] N. Z. Shor, Quadratic optimization problems, *Soviet J. Circuits Systems Sciences.* 25 (1987) 1–11.
- [48] M. J. Todd, K. C. Toh, Tütüncü, On the Nesterov-Todd direction in semidefinite programming, *SIAM J. Optim.* 8 (1998) 769–796.
- [49] M. J. Todd, *Semidefinite optimization*, Cambridge University Press, *Acta Numerica.* 10 (2001) 515–560.

- [50] K. C. Toh, Some new search directions for primal-dual interior point methods in semidefinite programming, *SIAM J. Optim.* 11 (2000) 223–242.
- [51] I. Touil, D. Benterki, A. Yassine, A feasible primaldual interior point method for linear semidefinite programming, *J. Comput. Appl. Math.* 312 (2017) 216–230.
- [52] L. Vandenberghe, S. Boyd, Primal-dual potential reduction method for problems involving matrix inequalities, *Math. Programming, Series B*, 69 (1995) 205–236.
- [53] L. Vandenberghe, S. Boyd, Semidefinite programming, *SIAM Review.* 38 (1) (1996) 49–95.
- [54] L. Vandenberghe, S. Boyd, Applications of semidefinite programming, *Applied Numerical Mathematics.* 29 (3) (1998) 283–299.
- [55] G. Q. Wang, Y. Q. Bai, C. Roos, Primal-dual interior-point algorithms for semidefinite optimization based on a simple kernel function, *Journal of Mathematical Modelling and Algorithms* 4 (2005) 409–433.
- [56] G. Q. Wang, Y. Q. Bai, A new primal-dual path-following interior-point algorithm for semidefinite optimization, *J. Math. Anal. Appl.* 353 (2009) 339–349.
- [57] H. Wolkowicz, R. Saigal, L. Vandenberghe (Eds.), *Handbook of semidefinite programming : Theory, algorithms and applications*, MA. Kluwer Academic Publishers, Boston, 2000.
- [58] H. Wolkowicz, G. P. H. Styan, Bounds for eigenvalues using traces, *Linear Algebra Appl.* 29 (1980) 471–506.
- [59] Y. Zhang, On extending some primal-dual interior-point algorithms from linear programming to semidefinite programming, *SIAM J. Optim.* 8 (1998) 365–386.

## ملخص

في هذه الأطروحة، اهتمنا بحل مسألة البرمجة الخطية نصف المعرفة (PSD) باستعمال طريقة المسار المركزي . أرفقنا المسألة (PSD) بالمسألة التقريبية  $(PSD)_\mu$  ذات الوسيط  $\mu$  . في البداية، برهنا وجود ووحداية الحل الأمثل للمسألة  $(PSD)_\mu$  ، ثم بينا أن هذا الحل يتقارب إلى الحل الأمثل للمسألة الأصلية (PSD) لما  $\mu$  يؤول إلى الصفر. لتعيين خطوة الانتقال، اقترحنا طريقة تعطي أربعة بدائل جديدة تكون بسيطة، سهلة وغير مكلفة لحساب خطوة الانتقال.

في النهاية قمنا بتجارب عددية لتوضيح فعالية وكذا تقارب البدائل الأربعة نحو الحل الأمثل للمسألة المعطاة (PSD) .

## كلمات مفتاحية:

البرمجة الخطية نصف المعرفة، طريقة المسار المركزي، طريقة النقط الداخلية الأولية- الثنوية، التحليل العقدي.

## Abstract

In this thesis, we are interested in solving the linear semidefinite programming problem (PSD) by a central trajectory method. We associated to (PSD) a perturbed problem, denoted  $(PSD)_\mu$ . First, we showed the existence and uniqueness of the optimal solution of the problem  $(PSD)_\mu$  and we showed that this solution converges to the optimal solution of the original problem (PSD) when  $\mu$  goes to zero.

Then, to calculate the displacement step, we proposed four new alternatives to calculate the displacement step by a simple, easy and much less expensive technique. At last, to enhance our contribution, we presented numeric simulations to illustrate the efficiency and the convergence of the four alternatives toward the optimal solution of the considered problem (PSD).

## Keywords:

Linear semidefinite programming, Central trajectory methods, Primal–dual interior point methods, Complexity analysis.

## Résumé

Dans cette thèse, on s'est intéressé à la résolution du problème de programmation semi-définie (PSD) par la méthode de trajectoire centrale. On a associé à (PSD) un problème perturbé, noté  $(PSD)_\mu$ . En premier lieu, on a montré l'existence et l'unicité de la solution optimale du problème  $(PSD)_\mu$  , ensuite on a montré que cette solution converge vers la solution optimale du problème originel (PSD) quand  $\mu$  tend vers zéro.

Puis, on a proposé quatre nouvelles alternatives pour calculer le pas de déplacement par une technique simple, facile et moins couteuse. Enfin, pour valoriser notre contribution, on a présenté des simulations numériques pour illustrer l'efficacité et la convergence des quatre alternatives vers la solution optimale du problème considéré (PSD).

## Mots clés :

Programmation semi-définie linéaire, Méthode de trajectoire centrale, Méthode de points intérieurs primale–duale, Analyse de la complexité.