

RÉPUBLIQUE ALGÉRIENNE DÉMOCRATIQUE ET POPULAIRE

Ministère de L' Enseignement Supérieur et de la Recherche Scientifique

UNIVERSITE DE BATNA 2

Faculté de Mathématiques et de l'informatique

Département de Mathématiques



THÈSE

**Présentée en Vue de l' Obtention Du Diplôme de Doctorat 3^{eme} cycle
LMD en Mathématique**

Option : Optimisation et Contrôle

Par

Charif Roumaissa

Thème

**Sur l'étude de méthodes avancées pour la résolution des problèmes de
programmation Mathématique , Application**

Soutenue le :04/07/2024

Devant le jury composé de :

Djeffal Lakhdar	Prof	Président	Université de Batna 2
Djeffal El Amir	Prof	Rapporteur	Université de Batna 2
Yahia Djabrane	Prof	Examineur	Université de Biskra
Brahimi Mahmoud	M.C.A	Examineur	Université de Batna 2
Imad Eddine Lakhdari	M.C.A	Examineur	Université de Biskra

Dédicace

Je souhaite dédier cette

Thèse....

À mes chers parents

Aucune dédicace ne saurait exprimer adéquatement mon respect , mon amour éternel et ma profonde gratitude pour les sacrifices consentis en vue de mon éducation et de mon bien-être .

À ma mère , dont l'amour , le soutien inconditionnel , les nombreux sacrifices et les conseils précieux ont été des piliers essentiels dans ma vie , je lui adresse à travers ce modeste travail toute l'étendue de mes sentiments et de ma reconnaissance éternelle .

À mon père , dont les longues années de sacrifices et de privations ont contribué à mon parcours , je souhaite que ce travail soit le fruit de ses efforts . Je lui suis reconnaissant pour les nobles valeurs , l'éducation et le soutien constants qu'il m'a prodigués.

Je tiens également à remercier **mon mari** , pour son soutien , sa compréhension , son sacrifice de toutes parts et sa patience lors de l'élaboration de cette thèse.

A mes chères agréables petites filles Meriem et Nouha.

À mes chères sœurs et frère , qui ont toujours été pour moi des exemples de persévérance , de courage et de générosité , je souhaite exprimer toute mon affection fraternelle , ma tendresse et ma reconnaissance.

À ma belle-mère et à mon beau-père , ainsi que **mes belle sœurs** pour leur aide .

Enfin , à tous ceux qui m'aiment et à tous ceux que j'aime , je dédie le fruit de mes recherches en Doctorat.

Remerciements

Nous exprimons notre gratitude premièrement à Allah Tout-Puissant pour nous avoir accordé la volonté, la santé et la patience nécessaires tout au long de nos études, ainsi que le courage de réaliser ce modeste travail.

Nous remercions chaleureusement notre encadrant Djefal.El Amir pour la pertinence de ses commentaires et sa patience lors de la réalisation de ce travail. Nous admirons énormément son sérieux et sa façon de diriger qui ont été pour nous une grande source d'inspiration et de motivation.

Nous remercions sincèrement les membres du jury.

Nous souhaitons également exprimer notre reconnaissance à tous les professeurs qui ont contribué à notre formation, ainsi qu'à toute l'équipe du département de mathématiques.

Table des matières

Table des figures	3
Introduction	8
1 Définition et Préliminaire	11
1.1 Notion sur la convexité	11
1.1.1 Fonctions convexes	13
1.1.2 Les Cônes	16
1.1.3 Programmation mathématique	17
2 La programmation cônica et quelques méthode pour la résoudre	21
2.1 Classification des problèmes d'optimisation cônica	22
2.1.1 Les principaux aspects de l'algèbre de Jordan et ses propriétés fondamentales.	24
2.1.2 Les propriétés algébrique des (SOCOs)	31
2.1.3 Le produit cartésien du cône	36
2.2 Diverse méthodes pour résoudre un problème d'optimisation cônica	38
2.2.1 Les méthodes de projection	38
2.2.2 Les méthodes prédicteurs - correcteurs	53
3 La programmation cônica du second ordre et méthodes de points intérieurs	59
3.1 Position du problème	59
3.2 Dualité en Soco	61
3.2.1 Dualité faible	61
3.2.2 Dualité forte	62
3.2.3 Complémentarité en Soco	63
3.3 Méthodes de points intérieurs pour le problème Soco	64
3.3.1 La méthode de trajectoire central primal-dual pour la programmation du cône du second ordre basée sur les fonctions noyau	65
Conclusion	84

Perspectives	84
bibliographie	85

Table des figures

- Figure 1 : (Figure illustre la différence entre l' ensemble convexe et non convexe , (page 12)).
- Figure 2 : (Figure illustre la différence entre la fonction convexe et non convexe , (page 15)).
- Figure 3 : (Algorithme 3 , (page 59)).
- Figure 4 : (Algorithme 4 , (page 71)).

Notations

\mathbb{R}^n : L'ensemble des vecteurs réels de n composantes.

\mathbb{R}_+^n : L'orthant positif de \mathbb{R}^n .

$\mathbb{R}^{n \times m}$: L'espace vectoriel des matrices avec $n \times m$ composantes réelles.

$B(x, r)$: La boule ouverte de centre x et de rayon r.

$\|x\|_p = (\sum_{i=1}^n |x_i|^p)^{\frac{1}{p}}$ la p^{eme} norme.

$\|x\|_\infty = \max_{1 \leq i \leq n} |x_i|$ la norme infinie.

\mathcal{E}^n : L'espace vectoriel réel de n^{ieme} dimension $\mathbb{R} \times \mathbb{R}^{n-1}$.

ν : un espace vectoriel euclidien de dimension finie sur \mathbb{R}

\mathcal{K} : un cône de \mathbb{R}^n .

\mathcal{K}^* : Le cône dual de \mathcal{K} .

$int(\mathcal{K})$: L'ensemble de tout les points intérieur de \mathcal{K} .

$FR(\mathcal{K})$: L'ensemble des points frontières de \mathcal{K} .

$x \succeq_{\mathcal{K}} y = x - y \in \mathcal{K}$ un ordre partiel sur un cône \mathcal{K} .

$x \succ_{\mathcal{K}} y = x - y \in int\mathcal{K}$ un ordre partiel sur un cône \mathcal{K} .

$GL(n, \mathbb{R})$: Le groupe linéaire général de degré n sur \mathbb{R} .

$Aut(\mathcal{K})$: L'automorphisme groupe de \mathcal{K} .

\mathcal{J} : L'espace vectoriel de dimension fini sur \mathbb{R} .

(\mathcal{J}, \circ) : L'algèbre de Jordan.

\circ : un opérateur binaire.

$x \circ y$: La multiplication de Jordan entre x et y .

$\langle x, y \rangle = Tr(x \circ y) = 2x^T y$ Le produit scalaire de x et y .

$\lambda_k(x)$: Les valeurs propre d'un vecteur x .

$c_k(x)$: Les vecteurs propre d'un vecteur x .

$x^{-1} = \lambda_1^{-1}c_1 + \dots + \lambda_k^{-1}c_k \quad \forall \lambda_k \neq 0$ L'inverse de x .

$x^{-1} = \lambda_1^{\frac{1}{2}}c_1 + \dots + \lambda_k^{\frac{1}{2}}c_k \quad \forall \lambda_k \geq 0$ La racine carré de x .

$\|x\|_F = (\sum \lambda_i^2)^{\frac{1}{2}} = \sqrt{Tr(x^2)}$ La norme unitairement invariante sur des matrices carrées.

$L(x)y = x \circ y$ une carte linéaire sur \mathcal{J} .

$Q(x, z) = L(x)L(z) + L(z)L(x) - L(x \circ z)$ L'opérateur quadratique.

$Q(x) = 2L(x) - L(x^2)$ La représentation quadratique.

$Arw(x)$: La matrice associée au vecteur $x \in \mathcal{E}^n$.

e : L'élément unitaire de l'algèbre de Jordan.

$Tr(x)$: La trace de x associée au cône \mathcal{K} .

$det(x)$: Le déterminant de x associée au cône \mathcal{K} .

Terminologie

(PM) : La programmation mathématique.

KKT : Kurush-Kuhn-Tucker.

LO : La programmation linéaire.

SOCO : la programmation conique de second ordre.

SDP : La programmation semi-définie positive.

Q_i^n : Le cône de second ordre (ou cône quadratique) de dimension n .

\mathcal{L}^k : Le produit direct des cône de Lorentz.

\mathcal{S}_+^k : Le produit direct des cône semi-définies positives.

(P) : Le problème primal.

(D) : Le problème dual.

(IPC) : Condition de points intérieurs.

NT : Nesterov et Todd.

ملخص

في هذه الأطروحة، نركز على دراسة برمجة المخروط من الدرجة الثانية. نقدم التعريفات الأساسية وكذلك النتائج الرئيسية لهذا النوع من المخاريط، ثم نقدم جبر جوردان لمشاكل برمجة المخروط من الدرجة الثانية SOCP . في هذا العمل، هدفنا هو تقديم طريقة مسار مركزي أولي - ثنائي تعتمد على دالة نواة جديدة. تؤدي دراسة هذه الدالة إلى تعقيد متعدد الحدود لهذا النوع من الخوارزمية ذات الخطوات الكبيرة .

الكلمات المفتاحية : تحسين المخروط من الدرجة الثانية، دالة النواة، طريقة النقطة الداخلية، حدود التعقيد .

Résumé

Dans cette thèse nous nous intéressons à l'étude de la programmation de cône du second ordre , nous donnons les définitions fondamentales ainsi que les principaux résultats de ce type de cône puis nous introduisons l'algèbre de Jordan pour les problèmes SOCP . Dans ce travail , notre objectif est de présenter une méthode de type trajectoire centrale primal-dual basée sur une nouvelle fonction noyau , l'étude de cette fonction nous conduit à une complexité polynomiale pour ce type d'algorithme à grands pas .

Mots-clés : Optimisation du cône de second ordre , Fonction noyau , Méthode du point intérieur , Limite de complexité .

Abstract

In this thesis , we focus on studying second-order cone programming . We provide fundamental definitions as well as the main results of this type of cone , and then introduce Jordan algebra for second-order cone programming problems (SOCP). In this work , our objective is to present a primal-dual central trajectory method based on a new kernel function . The study of this function leads us to polynomial complexity for this type of large-step algorithm .

Keywords : Second-order Cone Optimization , Kernel function , interior point method , complexity bound

Introduction

La programmation mathématique (ou optimisation mathématique) est un domaine de recherche qui rassemble l'informatique et les mathématiques. Elle consiste à rechercher le minimum (ou le maximum) d'une fonction appelée coût ou objectif. Le problème d'optimisation a de nombreuses applications qui apparaissent dans l'ingénierie, la finance, le milieu industriel, la gestion de projet, etc., ce sont donc des outils puissants pour résoudre des problèmes complexes dans différentes branches.

Les problèmes de programmation de cônes du second ordre sont un domaine important de la programmation mathématique et de l'optimisation profonde, ces problèmes sont des parties essentielles de l'optimisation convexe où la fonction linéaire est minimisée sur l'intersection d'un sous-espace affine avec le produit cartésien des cônes du second ordre (cône de Lorentz, ice-cream cône ou bien un cône quadratique).

La forme standard de ces problèmes est la suivante :

$$\begin{cases} \min c^T x \\ \text{sc } Ax = b \\ x \in \mathcal{K} \end{cases}$$

où \mathcal{K} est le cône du second ordre (ou encore un produit cartésien de ce cône).

Le problème SOCP représente une extension de la programmation linéaire, la programmation quadratique convexe avec des contraintes quadratiques linéaires ou convexes, ce sera également le cas pour de nombreux autres problèmes qui n'entrent pas dans ces catégories. La programmation des cônes du second ordre est généralement considérée comme des cas particuliers de la programmation semi-définie, une autre raison pour laquelle les SOCP incluent dans la programmation semi-définies est que les SDPs sont liés à l'intersection d'un ensemble affine et des matrices semi-définies des cônes positifs, par suite ces récentes se situent entre la programmation linéaire (LP) , quadratique (QP) et SDP.

En outre, SOCP a été l'un des domaines de recherche les plus dynamiques en programmation mathématique, c'est pour cette raison que nous avons deux aspects principaux qui motivent cette étude, Ils interviennent dans une grande diversité de domaines pratiques et mathématiques tels que l'optimisation combinatoire, la conception de filtres FIR, l'optimisation de

portefeuille avec des contraintes de risque de perte, l'optimisation robuste le contrôle, l'ingénierie [42]... etc, de plus sa facilité de calcul et leur applicabilité dans une grande variété de problèmes méritent d'être étudiés. Une autre théorie qui établit un lien entre la programmation semi-définie positive (SDP), la programmation linéaire et la programmation semi-définie conique (SOCO) est l'algèbre de Jordan, le livre de (Koranyi et Faraut représente l'ouvrage le plus approprié pour cette étude, elle permet d'appliquer la méthode des points intérieurs pour résoudre certains problèmes coniques. Il est important de noter que les méthodes des points intérieurs (IPM) trouvent leur origine dans le domaine de la programmation non linéaire. Ces IPM présentent plusieurs avantages. Tout d'abord, ils sont conçus pour converger vers la solution optimale, ce qui en fait une méthode robuste pour résoudre des problèmes de grande taille et difficiles. De plus, ils conviennent à divers types de problèmes, tels que les problèmes convexes, les problèmes d'optimisation avec contraintes de complémentarité, les contraintes de semi-définition positive, etc. Enfin, ils présentent une complexité polynomiale.

La méthode de points intérieurs est définie comme une technique de résolution qui génère une séquence de points situés à l'intérieur du domaine réalisable (admissible) et qui converge vers une solution du problème traité.

Nous distinguons deux types de méthodes de points intérieurs, MPI réalisables ou non réalisables. Les MPI réalisables commencent par un point strictement réalisable et maintiennent la faisabilité pendant le processus de solution ; ils sont également difficiles à résoudre. Les MPI non réalisables dépendent fortement du choix du point initial. Den Hertog a déclaré qu'il existe quatre catégories de méthodes de points intérieurs :

- Les méthodes affines de mise à l'échelle.
- Les méthodes de réduction du potentiel.
- Les méthodes de trajectoire (chemin central).
- Les Méthodes projectives .

Récemment, plusieurs monographies ont été consacrées aux méthodes de points intérieurs. Nous citons les travaux de Roos[63], Terlaky [58], Nesterov et Nemirovski[71], Ye J... etc. Une excellente étude pour corroborer les affirmations sur l'efficacité des méthodes de points intérieurs pour la programmation linéaire a été réalisée par Gonzaga, par la suite plusieurs chercheurs ont étendu cette approche. Nous mentionnons Nesterov et al [36] qui ont montré que les (MPIs) sont capables de résoudre des problèmes généraux d'optimisation conique.

Actuellement, la programmation du cône du second ordre suscite un intérêt considérable dans la recherche, parmi lesquels on peut citer. Nesterov

et Nemirovski [71] qui ont prouvé que les algorithmes de points intérieurs basés sur les fonctions barrières d'un problème SOCP d'inégalité n du cône du second ordre conduisent à une complexité d'ordre \sqrt{n} en temps polynomial, Bai et Roos ([63],[64]) a proposé une classe d'algorithmes de point intérieur primal-dual pour LO, SDO et SOCO basés sur les fonctions noyaux éligibles . plus tard, Cai et Li [59]. ont également étendu les travaux antérieurs sur le développement de méthodes de points intérieurs primal-dual pour résoudre l'optimisation quadratique convexe, basées sur une fonction barrière finie et l'optimisation quadratique convexe sur un cône symétrique dans leur travail. Ensuite, Nemirovski et Scheinberg ont étendu avec succès les méthodes des points intérieurs primal-dual de PL à SOCP. La première extension de cette approche a été mise en œuvre par Nesterov et Todd [73] qui sont les initiateurs de la méthode NT, ils ont démontré que ces méthodes sont plus puissantes d'un point de vue numérique. D'autre part, Adler, Alizadeh, Schmieta [55] ont introduit la méthode $XZ + ZX$ méthode aux problèmes SOCP, ils ont ensuite développé les conditions de non-dégénérescence, ils ont ensuite étudié la relation entre les problèmes SDP et SOCP.

Les fonctions noyau occupent une place essentielle dans le développement de nouveaux algorithmes de points intérieurs primal-dual. C'est le sujet de plusieurs chercheurs tels que Y. Q. Bai et M. El Ghami [68], C. Roos [67] X. Z. Cai et G. Q. Wang [60]... etc

Notre travail est abouti à des contributions importantes dans la dimension théorique et algorithmique pour résoudre de manière appropriée des problèmes d'optimisation de la programmation des cônes de second ordre en se concentrant sur les fonctions du noyau qualifié.

Notre thèse est structurée en trois parties principales est organisé comme suit

Le chapitre 1 présente un rappel et les notions essentielles de l'analyse convexe, quelques résultats de programmation mathématique et les principales propriétés des cônes .

Le deuxième chapitre est dédié à la programmation conique en particulier à la programmation conique du second ordre. On donne tout d'abord, sa formulation générale, quelques propriétés algébrique puis nous décrivons quelques méthodes de résolution de programmation conique .

Dans le dernier chapitre, nous discutons de la dualité du SOCO , ensuite nous proposons un algorithme primal -dual de point intérieur pour SOCO basé sur une nouvelle fonction noyau qui doit être qualifiée. Nous analysons ainsi les résultats de complexité des méthodes à grands pas .

Chapitre 1

Définition et Préliminaire

1.1 Notion sur la convexité

Ensembles convexes

Le rôle principale de la convexité est d'être un outil essentiel à la fois dans la recherche des conditions d'optimalité nécessaires et suffisantes et dans la théorie classique de l'optimisation.

Définition 1.1.1 On dit qu'un ensemble \mathcal{C} de \mathbb{R}^n est convexe si

$$\forall x, y \in \mathcal{C}, \forall t \in [0, 1]; (1 - t)x + ty \in \mathcal{C}.$$

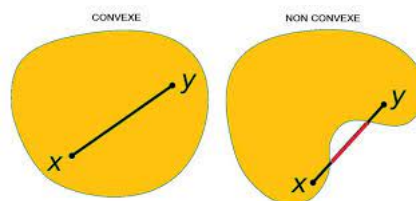


Figure 1 : (Figure illustre la différence entre l' ensemble convexe et non convexe)

propriétés

- Si \mathcal{C} est convexe et $t \in \mathbb{R}$ alors l'ensemble $\mathcal{S} = \{tx, x \in \mathcal{C}\}$ est convexe.
- Soit $\{\mathcal{C}_i\}_{i=1, \dots, n}$ une famille d'ensemble convexe alors :

1. $\mathcal{C} = \prod_{i=1}^n \mathcal{C}_i$ est un ensemble convexe .
2. $\mathcal{C} = \bigcap_{i=1}^n \mathcal{C}_i$ est un ensemble convexe .

On note par $\mathcal{C}(\mathbb{R}^n)$ l'ensemble de tous les sous ensemble convexe de \mathbb{R}^n

Théorème 1.1.1 . Soient $c \in \mathcal{C}(\mathbb{R}^n)$ et $\alpha \in \mathbb{R}$, alors $ac \in \mathcal{C}(\mathbb{R}^n)$.

- . Soient $c_1, c_2 \in \mathcal{C}(\mathbb{R}^n)$, alors $c_1 + c_2 \in \mathcal{C}(\mathbb{R}^n)$.
- . Soient n_1, n_2 deux entiers strictement positifs, Soient $c_1 \in \mathcal{C}(\mathbb{R}^{n_1}), c_2 \in \mathcal{C}(\mathbb{R}^{n_2})$, alors $c_1 \times c_2 \in \mathcal{C}(\mathbb{R}^{n_1+n_2})$.

Exemple 1.1 On cite quelques convexes usuels

- Soit $\mathcal{E} = \mathbb{R}^n$, $\mathcal{A} = \mathcal{E} \setminus \{0\}$ et $\alpha \in \mathbb{R}$
 - a) le demi espace ouvert $\mathcal{A}_{\mathcal{A}, \alpha} = \{\beta \in \mathcal{E}, \langle \beta, \mathcal{A} \rangle > \alpha\}$ est convexe.
 - b) le demi espace fermé $\bar{\mathcal{A}}_{\mathcal{A}, \alpha} = \{\beta \in \mathcal{E}, \langle \beta, \mathcal{A} \rangle \geq \alpha\}$ est convexe.
- L'union $\bigcup_{\lambda > 0} \lambda \mathcal{C}$ d'une partie convexe \mathcal{C} est un convexe .
- Soient $A \in \mathbb{R}^{n \times m}$, $b \in \mathbb{R}^m$ et soit l'application affine $Ax + b$, l'image direct ou inverse d'un convexe par cette application est convexe .
- Le disque unitaire $C = \{(x, y) \in \mathbb{R}^2 : x + y \leq 1\}$ est convexe .
- Soient $A \in \mathbb{R}^{n \times m}$, $b \in \mathbb{R}^m$ l'ensemble $\mathcal{P} = \{x \in \mathbb{R}^n : Ax = b\}$ est un ensemble convexe fermé appelé polyèdre .
- un polyèdre compact appelé polytope est convexe .
- Les boules ouverte $B_{a,r} = \{x \in E, \|x - a\| < r\}$ et fermées $\bar{B}_{a,r} = \{x \in E, \|x - a\| \leq r\}$ d'un espace vectoriel E sont convexes .
- L'intersection des demi plan $H = \{(x, y) \in \mathbb{R}^2, y \geq x\}$ est une partie convexe .

Définition 1.1.2 L'enveloppe convexe d'un ensemble $\mathcal{S} \subset \mathbb{R}^n$ est le plus petit sous ensemble convexe de \mathbb{R}^n qui contient \mathcal{S} (c'est-à-dire l'intersection de toutes les convexes contenant \mathcal{S}), et on la note par $\text{conv}(\mathcal{S})$

$$\text{conv}(\mathcal{S}) = \bigcap \{\mathcal{C}_i\}$$

aussi

$$\mathcal{S} \text{ convexe} \Leftrightarrow \mathcal{S} = \text{conv}(\mathcal{S})$$

Définition 1.1.3 Soit \mathcal{C} un convexe non vide de \mathbb{R}^n un point $x \in \mathcal{C}$ est dit extrémal si

$$\forall y, z \in \mathcal{C}, \forall t \in]0, 1[, x = ty + (1 - t)z, \text{ entraine } x = y = z.$$

Autrement dit, si le point $x \in \mathcal{C}$ n'est pas à l'intérieur d'un segment de droite contenu dans \mathcal{C} alors x est un point extrémal .

1.1.1 Fonctions convexes

On associe pour la fonction $F : \mathbb{R}^n \longrightarrow \mathbb{R}$, les ensembles suivants

L'épigraphe de F :

$$\text{epi}(F) := \{(x, a) \in \mathbb{R}^n \times \mathbb{R} : x \in \mathcal{C}, F(x) \leq a\}$$

Domaine effectif de F :

$$\text{dom}(F) := \{x \in \mathbb{R}^n : f(x) < +\infty\}$$

Ensembles de niveau $v \in \mathbb{R}$ de F :

$$\mathcal{S}_v = \begin{cases} \{x \in \mathbb{R}^n : f(x) \leq v\} & (\text{inférieur large}) \\ \{x \in \mathbb{R}^n : f(x) \geq v\} & (\text{supérieur large}) \end{cases}$$

Définition 1.1.4 Soit F est une fonction définie sur un ensemble \mathcal{C} fermé de \mathbb{R}^n on dit que F est inf-compact si pour tout $v \in \mathbb{R}$ ses ensembles de niveau inférieurs $\mathcal{S}_v(F)$ est compact .

Définition 1.1.5 une fonction F est dite propre si pour tout $x \in \mathbb{R}^n$

$$\text{dom}(F) \neq \emptyset \text{ et } F(x) > -\infty$$

Définition 1.1.6 on dit qu'une fonction F définie sur un ensemble convexe \mathcal{C} de \mathbb{R}^n est :

– Convexe si et seulement si

$$\forall x, y \in \mathcal{C}, \forall t \in [0, 1], F(tx + (1-t)y) \leq tF(x) + (1-t)F(y).$$

– Strictement convexe si et seulement si

$$\forall x, y \in \mathcal{C}, \forall t \in]0, 1[\text{ avec } x \neq y, F(tx + (1-t)y) < tF(x) + (1-t)F(y).$$

– Fortement convexe s'il existe $\alpha > 0$ tel que

$$\forall x, y \in \mathcal{C}, \forall t \in [0, 1], F(tx + (1-t)y) \leq tF(x) + (1-t)F(y) - \frac{\alpha}{2} t(1-t) \|x - y\|^2 .$$

Forte convexité \Leftrightarrow *stricte convexité* \Leftrightarrow *convexité* .

– Concave si $(-F)$ est convexe sur \mathcal{C} .

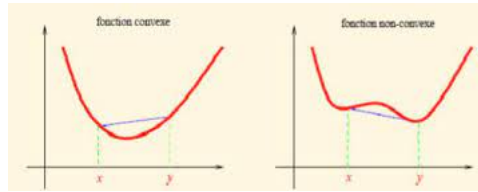


Figure 2 : (Figure illustre la différence entre la fonction convexe et non convexe)

Théorème 1.1.2 si F est une fonction deux fois continument différentiable alors les condition ci-dessous sont équivalent :

- F est convexe .
- $\forall x \in \mathcal{C}, \forall y \in \mathcal{C} \quad F(y) - F(x) \geq |\nabla F(x)|^T (y - x)$.
- $\forall x \in \mathcal{C}, \nabla^2 F(x)$ la hessien est une matrice semi définie positive.

Propriété

Soit F une fonction définie sur un ensemble convexe $\mathcal{C} \subset \mathbb{R}^n$, F est convexe si et seulement si son épigraphe est convexe.

Exemple 1.2 1) La fonction de Leontiev $U_{L,\alpha}(v_1, \dots, v(n)) = \max_{i=1}^n (\alpha_i v_i)$ est convexe pour tout $\alpha_1, \dots, \alpha_n$ des réels positifs.

2) La fonction $f(x) = x_{[1]} + x_{[2]} + \dots + x_{[r]}$ désigne la somme de r -plus grandes composantes de $x \in \mathbb{R}^n$ est convexe.

3) La fonction $H_A : x \in E \rightarrow \langle Ax, x \rangle$ où A est un opérateur symétrique est convexe, elle est strictement convexe si A est définie positif.

4) Toute norme $\| \cdot \|$ est convexe (particulièrement la fonction $|x|$, $x \in \mathbb{R}$ est convexe car c'est une conséquence de l'inégalité triangulaire et l'homogénéité de la norme).

5) La fonction log barrière $f : x \rightarrow \mathbb{R}_{++}^n$ tel que

$$f : x = \begin{cases} -\sum_{i=1}^m \log x_i & \text{si } x_i > 0 \\ +\infty & \text{sinon.} \end{cases}$$

est convexe .

6) Les fonctions sous linéaire i,e $\forall x, y \in \mathcal{C}, \forall \alpha > 0$

$$f(x_1 + x_2) \leq f(x_1) + f(x_2) \text{ et } f(\alpha x_1) = \alpha f(x_1)$$

est convexe .

7) La somme des fonctions convexes sont convexes .

8) Le sup et la multiplication par un scalaire positif de fonction convexe aussi convexe .

Définition 1.1.7 Soit C une partie convexe non borné et soit $x \in C$, on dit qu'une fonction F est coercive sur C si

$$\lim_{\|x\| \rightarrow +\infty} F(x) = +\infty$$

Définition 1.1.8 Soit F est une fonction défini sur C fermé de \mathbb{R}^n , on dit que F est inf-compact si pour tout $v \in \mathbb{R}$ ses ensembles de niveau inférieurs $S_v(F)$ sont compact .

Lemme 1.1.1 F est inf-compacte si et seulement si $\lim_{\|x\| \rightarrow +\infty} F(x) = +\infty$

Définition 1.1.9 Soit C un convexe fermé dont l'intérieur est non vide, on dit qu'une fonction $F : \text{int}C \rightarrow \mathbb{R}$ est une fonction barrière si :

- $F(x)$ est strictement convexe .
- $F(x)$ est fini lorsque x appartient à l'intérieur relatif du domaine réalisable .
- $F(x)$ tend vers $+\infty$ lorsque x tend vers la frontière du domaine réalisable .

Exemple 1.3 On prend l'ensemble convexe $C = \{x \in \mathbb{R}^n \mid f_i(x) \leq 0, i = 1, \dots, m\}$ avec f_i sont convexe car C est convexe

il possible de trouver une fonction barrière appropriée, en utilisant la fonction barrière logarithmique

$$F : \text{int}C \rightarrow \mathbb{R}, F(x) = \sum_{i=1}^m \ln(-f_i(x))$$

Exemple 1.4 La fonction barrière n'est pas toujours facile à trouver, en effet si $C = \mathbb{R}^n$, $f_i(x) = -x^x$ alors la fonction $F(x) = -\ln x^x = -x \ln x$ n'est pas une fonction barrière pour C .

Définition 1.1.10 On définit la fonction barrière logarithmique $F(\cdot)$ sur le cône du second ordre K_{m+}^n par

$$F(v) = \begin{cases} -\sum_{i=1}^m \ln \det(v_i) & \text{si } v_i \in \text{int}K_{m+}^n \\ +\infty & \text{ailleurs.} \end{cases}$$

Remarque 1.1.1 L'utilité de la fonction barrière est d'empêcher les itérés de sortir de l'ensemble C .

1.1.2 Les Cônes

Définition 1.1.11 Soient $x_1, \dots, x_m \in \mathbb{R}^n$ et $\lambda_1, \dots, \lambda_m \in \mathbb{R}_+$ la combinaison positive des points x_1, \dots, x_m est :

$$x := \sum_{j=1}^m \lambda_j x_j$$

généralement on dit que $x \in \mathbb{R}^n$ est une combinaison positive de point d'un sous ensemble $S \subseteq \mathbb{R}^n$, s'il existe un nombre fini de points de S tel que x est une combinaison positive .

Définition 1.1.12 On appelle l'enveloppe conique du $S \subseteq \mathbb{R}^n$ l'ensemble de toutes les combinaisons positives de points de S et on la note par $\text{cône}(S)$. l'enveloppe cônica est un cône convexe.

Définition 1.1.13 Un polyèdre dans \mathbb{R}^n est un ensemble de la forme $\mathcal{P} = \{x \in \mathbb{R}^n \mid Ax = b\}$ où $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$.

Définition 1.1.14 on dit que l'ensemble \mathcal{K} est un cône si $\forall \lambda > 0, \lambda \mathcal{K} \subset \mathcal{K}$.

Définition 1.1.15 Soit $\mathcal{K} \subseteq \mathbb{R}^n$ un sous ensemble non vide on dit que \mathcal{K} est un cône convexe . si est seulement si

$$\forall \alpha, \beta \in \mathbb{R}_+, \forall u, v \in \mathcal{K}, \alpha u + \beta v \in \mathcal{K}$$

. Autrement dit il est stable par rapport aux combinaisons linéaires positives.

Remarque 1.1.2 l'utilité des cônes convexes est de jouer un rôle important dans la structure des convexes .

Définition 1.1.16 Soit \mathcal{K} un sous ensemble de \mathbb{R}^n et soit $x \in \mathcal{K}$ on appelle cône tangente à \mathcal{K} on x et on la note $T_{\mathcal{K}}(x)$ l'ensemble des vecteur ω de \mathbb{R}^n pour lequel il existe une suite de vecteurs $\omega_K \in \mathbb{R}^n$ et une suite réels $\alpha_K > 0$ tels que

- $\alpha_K \rightarrow 0^+$ et $\omega_K \rightarrow \omega$.
- $\forall \mathcal{K}, x + \alpha_K \omega_K \in \mathcal{K}$.

Corollaire 1.1.1 Le cône tangent est bien un cône c'est-à-dire

$$\forall \omega \in T_{\mathcal{K}}(x), \forall \lambda \geq 0, \lambda \omega \in T_{\mathcal{K}}(x).$$

Remarque 1.1.3 Le cône tangent est un ensemble fermé .

Définition 1.1.17 Soit \mathcal{K} un sous ensemble convexe et fermé de \mathbb{R}^n on appelle le cône de récession de \mathcal{K} l'ensemble $\text{rec}(\mathcal{K})$ tel que :

$$\text{rec}(\mathcal{K}) := \{y \in \mathbb{R}^n \mid \forall x \in \mathcal{K}, \forall \lambda \geq 0, x + \lambda y \in \mathcal{K}\}.$$

Lemme 1.1.2 un cône de récession $\text{rec}(\mathcal{K})$ est un cône convexe fermé de \mathbb{R}^n si on a l'égalité

$$\text{rec}(\mathcal{K}) := \{y \in \mathbb{R}^n \mid \exists x \in \mathcal{K}, \forall \lambda \geq 0, x + \lambda y \in \mathcal{K}\}$$

de plus \mathcal{K} est borné si et seulement si $\text{rec}(\mathcal{K})$ est réduit à zéro .

Définition 1.1.18 On définit le cône dual de l'ensemble $\mathcal{K} \in \mathbb{R}^n$ par

$$\mathcal{K}^* := \{z \mid x^T z \geq 0, \forall x \in \mathcal{K}\}$$

\mathcal{K}^* est un cône convexe fermé .

Définition 1.1.19 On dit qu'un cône est finiment engendré dans \mathbb{R}^n si elle forme un enveloppe cônica d'un nombre fini de points de \mathbb{R}^n .

Corollaire 1.1.2 Les cônes finiment engendré de \mathbb{R}^n est fermé .

Définition 1.1.20 Le cône simplicial dans \mathbb{R}^n est l'enveloppe cônica d'un point linéairement indépendants et fini de \mathbb{R}^n .

Définition 1.1.21 Le cône polyédrique de \mathbb{R}^n est l'ensemble $\hat{K} := \{x \in \mathbb{R}^n \mid Ax \leq 0\}$ avec $A \in \mathbb{R}^{m \times n}$.

1.1.3 Programmation mathématique

Dans cette partie nous nous intéressons à donner quelques définitions fondamentales et d'unicité de la solution optimale ainsi que les condition d'optimalité.

Considérons le programme mathématique (PM) suivant :

$$(PM) \begin{cases} \min f(x) \\ g_i(x) \leq 0, \quad i = 1, \dots, r \\ h_j(x) = 0, \quad j = 1, \dots, p \\ x \in \mathcal{J} \end{cases}$$

où \mathcal{J} est une partie de \mathbb{R}^n

La fonction $f : \mathcal{J} \rightarrow \mathbb{R}$ est la fonction objective du problème.

Les fonction $g_i : \mathcal{J} \rightarrow \mathbb{R}$ sont les contraintes d'inégalités du problème.

Les fonction $h_j : \mathcal{J} \rightarrow \mathbb{R}$ sont les contraintes d'égalités du problème.

Soit l'ensemble des solutions admissibles \mathcal{M} tel que :

$$\mathcal{M} = \{x \in \mathcal{J}, g_i(x) \leq 0, i = 1, \dots, r, h_j(x) = 0, j = 1, \dots, p\}$$

On dit que le vecteur x est une solution réalisable de (PM) si $x \in \mathcal{M}$. L'ensemble de ces solutions réalisables forme le domaine réalisable de (PM) c-à-d :

$$D = \text{dom}f \cap \bigcap_{i=1}^r \text{dom}g_i \cap \bigcap_{j=1}^p \text{dom}h_j \cap \mathcal{J}.$$

Une solution réalisable $x^* \in \mathcal{M}$ est dite optimale de (PM), si elle minimise la fonction f sur \mathcal{M} où \mathcal{M} est l'ensemble de toutes les solutions réalisables.

Définition 1.1.22 (Minimum local) On dit que la fonction $f : \mathcal{J} \in \mathbb{R}^n \rightarrow \mathbb{R}$ admet un minimum local $x^* \in \mathcal{J}$ s'il existe $r > 0$ tel que :

$$\forall x \in B(x^*, \epsilon) \subset \mathcal{J}, f(x) \geq f(x^*), \quad \text{ou} \quad B((x^*, \epsilon) = \{x \in \mathcal{J}, \|x - x^*\| < \epsilon\}.$$

Définition 1.1.23 (Minimum global) On dit que la fonction $f : \mathcal{J} \in \mathbb{R}^n \rightarrow \mathbb{R}$ admet un minimum global (solution optimale global) $x^* \in \mathcal{J}$ si

$$\forall x \in \mathcal{J}, f(x) \geq f(x^*).$$

Remarque 1.1.4 La recherche du maximum de la fonction f se ramène à la minimisation du problème $-f$.

Classification des problèmes de programmation mathématique

Un programme mathématique (PM) est classifié à partir des propriétés fondamentales à savoir : la linéarité du fonction du problème, la convexité et la différentiabilité. En effet :

- (PM) est un problème linéaire si la fonction f est linéaire et les fonction g_i, h_j sont affines.
- (PM) est un problème convexe si les fonction f, g_i sont convexes et h_j sont affines.
- (PM) est un problème différentiable si les fonction f, g_i et h_j sont différentiables.

Qualification des contraintes

On distingue les cas suivants pour que la qualification des contraintes pour tout $x^* \in \mathcal{M}$ est satisfaite

- Toutes les contraintes sont affines (les contraintes qualifie en tout point réalisable).
- Les gradients des contraintes saturées en x^* sont libres.
- \mathcal{M} est convexe et $\text{int}(\mathcal{M}) \neq \emptyset$.

Résultats d'existence et d'unicité de la solution optimale du (PM)

Théorème 1.1.3 (Weistrass) Soit f une fonction continue sur $\mathcal{J} \subset \mathbb{R}^n$ compact (fermé et borné) alors (PM) admet au moins une solution optimale (global) $x^* \in \mathcal{J}$.

Corollaire 1.1.3 Si f est une fonction continue et coercive sur un ensemble \mathcal{J} fermé alors le problème (PM) admet au moins une solution optimale $x^* \in \mathcal{J}$.

Corollaire 1.1.4 si la fonction f est inf-compacte, alors le problème (PM) admet une solution optimale globale

Théorème 1.1.4 Soit \mathcal{J} est un ensemble convexe de \mathbb{R}^n et soit f est une fonction strictement convexe sur \mathcal{J} , alors le problème (PM) admet au plus une solution optimale.

Condition d'optimalité

Définition 1.1.24 Le lagrangien associe au problème (PM) est défini par $L(x, \lambda, \mu) = f(x) + \sum_{i=1}^r \lambda_i g_i(x) + \sum_{j=1}^p \mu_j h_j(x)$ où $\lambda_i \in \mathbb{R}^+$, $\mu_j \in \mathbb{R}$

Théorème 1.1.5 (Karush-Kuhn-Tucker (KKT)) Soit $x^* \in J$ est un minimum local du problème (PM) est soit $f : \mathbb{R}^n \rightarrow \mathbb{R}$ est une fonction différentiable sur J alors il existe des vecteurs (appelé multiplicateur de lagrange) $\lambda \in \mathbb{R}_+^r$, $\mu \in \mathbb{R}^p$ tel que

$$\begin{cases} \nabla f(x^*) + \sum_{i=1}^r \lambda_i \nabla g_i(x^*) + \sum_{j=1}^p \mu_j \nabla h_j(x^*) = 0 & \text{(condition d'optimalit)} \\ \lambda_i g_i(x^*) = 0, \quad i = 1, \dots, r & \text{(condition de complementarit)} \\ h_j(x^*) = 0, \quad j = 1, \dots, p \end{cases}$$

Remarque 1.1.5 *Si le problème (PM) est convexe, les conditions d'optimalité (KKT) sont à la fois nécessaires et suffisantes afin que x^* soit un minimum global.*

Dans le cas où les contraintes ne sont pas qualifiées en x^ , les conditions de (KKT) ne s'appliquent pas, alors x^* peut être un optimum sans vérifier ces conditions.*

Chapitre 2

La programmation cônica et quelques méthode pour la résoudre

D'abord nous allons intéresser à un genre de problème dont l'ensemble des contraintes \mathcal{C} est un cône convexe fermé et pointé dont l'intérieur est non vide . Cette structure particulière des problèmes est appelées la programmation cônica.

Beaucoup de problème pratique peuvent être modélisés sous cette forme aussi on peut prouver qu'un problème convexe est équivalent à un problème de programmation cônica , il existe trois types de cône dont nous aurai parler plus tôt , notre étude est basé sur la programmation cônica du second ordre , pour ces cône on peut construire une fonction barrière self-concordante et établir un algorithme primal-dual de suivi de chemin avec une telle complexité , à l'aide de ses fonctions barrière ayant des propriétés intéressantes on peut définir une trajectoire centrale et d'établir des méthodes de suivi de chemin central afin de calculer les solution optimales.

Ces méthode seront efficaces non seulement dans la théorie de complexité mais aussi pour résoudre des problèmes pratique . L'un des principaux avantages de la programmation cônica est la richesse de sa théorie de dualité.

2.1 Classification des problèmes d'optimisation cônica

un problème d'optimisation cônica est un problème de la forme suivante :

$$\min_x \{c^T x + d : Ax - b \in \mathcal{K}\}.$$

où, $x \in \mathbb{R}^n$, A est $m \times n$, $b \in \mathbb{R}^m$, $c \in \mathbb{R}^n$, $d \in \mathbb{R}$, et $\mathcal{K} \subset \mathbb{R}^m$ est un cône convexe à intérieur non vide, et l'application $x \mapsto Ax - b$ est de \mathbb{R}^n dans $\mathcal{K} \subset \mathbb{R}^m$.

La formulation cônica est la formulation universelle de la programmation convexe. Parmi les atouts de cette formulation particulière, on peut noter son pouvoir unificateur. Elle englobe une vaste éventail de problèmes d'optimisation convexe à travers trois types de cônes distincts :

- **Le produit directes quadrants non positifs** $\mathcal{K} = \mathbb{R}_+^m$: ces cônes donnent naissance aux problèmes d'optimisation linéaire

$$\min_x \{c^T x + d : a_i^T x - b_i \geq 0, i = 1, \dots, m\}.$$

- **Le produit directe des cônes de lorentz** $\mathcal{L}^k = \{x \in \mathbb{R}^k : x_k \geq \sqrt{x_1^2 + \dots + x_{k-1}^2}\}$: ces cônes engendrent des défis aux problèmes d'optimisation quadratique, la forme mathématique de ces fameux programmes est donnée comme suit :

$$\min_x \{c^T x + d : \|A_i x - b_i\|_2 \leq c_i^T x - d_i, i = 1, \dots, m\}.$$

ici la i^{eme} contrainte dite inégalité cônica quadratique. $I \subset \mathcal{K}$ exprime que l'application affine $x \mapsto [A_i x; c_i^T] - [b_i; d_i]$ est de \mathbb{R}^n dans un cône \mathcal{L}^i de dimension appropriée, tandis que le système de toutes les contraintes signifie que l'application affine $x \mapsto [[A_i x; c_i^T]; \dots; [A_m x; c_m^T]] - [[b_i; d_i]; \dots; [b_m; d_m]]$ est de \mathbb{R}^n dans $\mathcal{L}^1 \times \dots \times \mathcal{L}^m$

- **le produit direct de cône semi-définies** \mathcal{S}_+^k : il représente le cône des matrices semi-définies positives $\mathcal{K} \times \mathcal{K}$, il est contenu dans l'espace \mathcal{S}^k des matrices symétrique $\mathcal{K} \times \mathcal{K}$. \mathcal{S}^k considérée comme l'espace euclidien muni du produit de frobenius interne $\langle A, B \rangle = \text{trace}(AB)$. La famille des cônes semi-définie donne naissance aux problèmes

d'optimisation semi-définie qui ont la forme suivante :

$$\min_x \{c^T x + d : A_i - B_i \geq 0, \quad i = 1, \dots, m\}.$$

ce qui signifie que l'application affine $x \mapsto A_i - B_i = \sum_{j=0}^n x_j A^{ij} - B_i$ est de \mathbb{R}^n dans $\mathcal{S}_+^{k_i}$ telle que A^{ij} et B_i sont des matrices symétrique de dimension $\mathcal{K}_i \times \mathcal{K}_i$.

La contrainte $\sum_{j=0}^n x_j A^{ij} - B_i \geq 0$ est appelée LMA (Linear Matrix Inequality).

Remarque 2.1.1 : Notons que les trois type de cône au dessus admettent la propriété que le produit direct de cône appartenant a l'un des trois types est un cône de même type , Cela provient de la richesse d'un cône, une caractéristique présente dans ces trois types de cônes . C'est cette particularité qui justifier la classification précédente qui basée sur la représentation de sous produit directs de cône de l'ensemble des contraintes.

Récapitulation

Type de cône	\mathbb{R}_+^m	\mathcal{L}^k	\mathcal{S}_+^k
Dimension	m	m	$\frac{m(m+1)}{2}$
L'identité	$I = (1, \dots, 1)^T$	$(1, 0)^T$	I_m
Multiplication	$\begin{pmatrix} x_1 \\ \cdot \\ \cdot \\ x_m \end{pmatrix} \cdot \begin{pmatrix} y_1 \\ \cdot \\ \cdot \\ y_m \end{pmatrix} = \begin{pmatrix} x_1 y_1 \\ \cdot \\ \cdot \\ x_m y_m \end{pmatrix}$	$\begin{pmatrix} a \\ b \end{pmatrix} \cdot \begin{pmatrix} c \\ d \end{pmatrix} = \begin{pmatrix} ac + y^T d \\ cy + ad \end{pmatrix}$	$X \bullet Y = \frac{XY + YX}{2}$
Rang	m	2	m
Déterminant	$\prod_{i=1}^m x_i$	$\det \begin{pmatrix} a \\ b \end{pmatrix} = a^2 - \ b\ ^2$	$\det X$
Valeur propre	x_i	$a_{\pm} \ b\ ^2$	$\lambda_i(X)$
Inverse	$(x_1^{-1}, \dots, x_m^{-1})^T$	$\begin{pmatrix} a \\ b \end{pmatrix}^{-1} = \frac{1}{a^2 - \ b\ ^2} \begin{pmatrix} a \\ -b \end{pmatrix}$	X^{-1}
Les idempotents primitifs	les vecteurs de base e_i	$\frac{1}{2}(1, b^T)^T, \ b\ = 1$	$xx^T, \ x\ = 1$

(Tableau 1)

Quelques exemples sur la programmation cônica du second ordre

Exemple 2.1 Soit le problème impliquant un maximum de norme suivant :

$$\begin{cases} \min \sum_{i=1}^p t_i \\ \text{Sc } \| A_i x + b_i \| \leq t_i, i = 1, \dots, p \end{cases}$$

Exemple 2.2 considérons les problèmes robustes d'optimisation convexe dans lesquelles il y a une incertitude sur les données on peut utiliser la programmation cônica du second ordre pour résoudre ce type de problème premièrement on considère le problème de programmation linéaire suivant

$$\begin{cases} \min c^T x \\ \text{Sc } a_i^T x \leq b_i, i = 1, \dots, m \end{cases}$$

ainsi sont problème robuste est donné par

$$(PLR) = \begin{cases} \min c^T x \\ \text{Sc } a_i^T x \leq b_i, \forall a_i \in \varepsilon_i, i = 1, \dots, m \end{cases}$$

aussi la contrainte peut être également s'exprimer comme

$$\max\{a_i^T x \mid a_i \in \varepsilon_i\} = \bar{a}_i^T x + \| P_i x \| \leq b_i.$$

où P_i est une matrice symétrique semi-définie positive, ce qui est une contrainte sur le cône du second ordre.

Alors le problème de programmation linéaire robuste peut alors être exprimé comme un problème de programmation cônica du second ordre

$$\begin{cases} \min c^T x \\ \text{Sc } \bar{a}_i^T x + \| P_i x \| \leq b_i, i = 1, \dots, m \end{cases}$$

2.1.1 Les principaux aspects de l'algèbre de Jordan et ses propriétés fondamentales.

Définition de la norme

soit $x \in \mathbb{R}^n$ et $p \geq 1$ est un nombre réel, la p^{eme} norme de x est donnée par :

$$\| x \|_p = \left(\sum_{i=1}^n |x_i|^p \right)^{\frac{1}{p}}$$

Prenez en considération que pour $p = 1$ on obtient la norme 1 . Lorsque $p = 2$ on obtient la norme euclidienne , et lorsque p approche de l'infini la p^{eme} norme approche également vers l'infini ce qui est défini par

$$\|x\|_{\infty} = \max\{|x_1|, |x_2|, \dots, |x_n|\}.$$

Pour chaque vecteur $x \in \mathbb{R}^n$ dont la première entrée est indicée par 0 , on note par \bar{x} le sous-vecteur constitué des entrées de 1, jusqu'à $n - 1$ donc $x = (x_0, \bar{x}^T)^T \in \mathbb{R} \times \mathbb{R}^{n-1}$.

On note par \mathcal{E}^n l'espace vectoriel réel de n^{ieme} dimension $\mathbb{R} \times \mathbb{R}^{n-1}$ dont les vecteurs x sont indexés à partir de 0.

Maintenant , on définit la matrice associée à chaque vecteur $x \in \mathcal{E}^n$ et $p \in [1, \infty]$ par :

$$J_p(x) = \begin{pmatrix} 1 & 0^T \\ 0 & \frac{\|\bar{x}\|_p^2}{\|\bar{x}\|_2^2} I_{n-1} \end{pmatrix}$$

Soit \mathcal{K} un cône dans \mathbb{R}^n , on note par $int\mathcal{K}$ l'intérieur de \mathcal{K} et par $Fr(\mathcal{K})$ la frontière du \mathcal{K} .

On définit le cône d'ordre p de dimension n comme suit

$$\mathcal{P}_p = \{x \in \mathcal{E}^n : x_0 \geq \|\bar{x}\|_p\}.$$

On écrit par $x \succeq_p 0$ tout $x \in \mathcal{P}_p$ et par $x \succ_p 0$ tout $x \in int\mathcal{P}_p$, $i, e, (x_0 > \|\bar{x}\|_p)$ et nous écrivons $x \succeq_p y$ or $y \preceq_p x$ pour signifier que $x - y \succeq_p 0$ et $x \succ_p y$ ou bien $y \prec_p x$ pour signifier que $x - y \succ_p 0$.

Définition 2.1.1 un cône est dit fermer si est seulement s'il est fermé par rapport à la prise de limites .

Définition 2.1.2 un cône est dit solide si est seulement s'il a un intérieur non vide .

Définition 2.1.3 un cône est dit pointé si est seulement s'il ne contient pas deux vecteurs opposés non nuls(donc l'origine est un point extrême).

Définition 2.1.4 un cône est dit régulier si est seulement si c'est un cône fermé, convexe, pointu, plein. Clairement, le cône d'ordre p est régulier .

Maintenant soit ν un espace vectoriel euclidien de dimension finie sur \mathbb{R} avec le produit scalaire " $\langle \cdot, \cdot \rangle$ ".

On note par \mathcal{K}^* le cône dual d'un cône régulier $\mathcal{K} \subset \nu$ est on définit par

$$\mathcal{K}^* = \{y \in \nu : \langle x, y \rangle \geq 0, \forall x \in \mathcal{K}\}.$$

Définition 2.1.5 un cône \mathcal{K} est dit auto-dual si est seulement s'il coïncide avec son cône dual \mathcal{K}^* , i.e, $\mathcal{K} = \mathcal{K}^*$.

Nous entendons par $GL(n, \mathbb{R})$ le groupe linéaire général de degré n sur \mathbb{R} (c-à-d l'ensemble des matrices inversibles $n \times n$ avec des entrées de \mathbb{R} ainsi que l'opération de multiplication matricielle ordinaire).

Pour un cône régulier $\mathcal{K} \subset v$ on note par $Aut(\mathcal{K})$ l'automorphisme groupe de \mathcal{K} , i.e $Aut(\mathcal{K}) = \{T \in GL(n, \mathbb{R}) : T(\mathcal{K}) = \mathcal{K}\}$.

Définition 2.1.6 Soit v un espace euclidien réel de dimension finie, Un cône régulier $\mathcal{K} \subset v$ est dit homogène si est seulement si pour tout $u, v \in \text{int}\mathcal{K}$, il existe une application linéaire inversible $F : v \rightarrow v$ telle que :

- $F(\mathcal{K}) = \mathcal{K}$, i.e, F est un automorphisme de \mathcal{K}
- $F(u) = v$

Définition 2.1.7 Un cône régulier \mathcal{K} est dit symétrique si est seulement s'il est à la fois auto-dual et homogène.

Soit \mathcal{J} un espace vectoriel de dimension fini sur \mathbb{R} L'espace vectoriel \mathcal{J} est appelé une algèbre sur \mathbb{R} si est seulement s'il existe une application bilinéaire $\circ : \mathcal{J} \times \mathcal{J} \rightarrow \mathcal{J}$.

Soit x un élément d'une algèbre \mathcal{J} , alors pour tout $n \geq 2$ on définit $x^{(n)}$ récursivement par $x^{(n)} = x \circ x^{(n-1)}$

Définition 2.1.8 Soit \mathcal{J} est une \mathbb{R} algèbre de dimension finie avec le produit bilinéaire $\circ : \mathcal{J} \times \mathcal{J} \rightarrow \mathcal{J}$ alors (\mathcal{J}, \circ) est appelée une algèbre de Jordan si est seulement si pour tout $x, y \in \mathcal{J}$ on a :

- $x \circ y = y \circ x$ (commutativité)
- $x \circ (x^2 \circ y) = x^2 \circ (x \circ y)$ (l'axiome de Jordan)

Cependant, si l'algèbre (\mathcal{J}, \circ) n'est pas commutative, alors on peut remplacer son opération binaire " \circ " par sa version commutative " $\bar{\circ}$ " où $x \bar{\circ} y = \frac{x \circ y + y \circ x}{2}$ pour tout $x, y \in \mathcal{E}^n$.

Remarque 2.1.2 - Le produit $x \circ y$ entre deux éléments x et y de l'algèbre de Jordan (\mathcal{J}, \circ) est appelée la multiplication de Jordan entre x et y .

- Une algèbre de Jordan (\mathcal{J}, \circ) a un élément d'identité si est seulement s'il existe un élément (nécessairement unique) $e \in \mathcal{J}$ telle que $x \circ e = e \circ x = x$ pour tous $x \in \mathcal{J}$.

- Une algèbre de Jordan (\mathcal{J}, \circ) n'est pas nécessairement associative c'est-à-dire, $x \circ (y \circ z) = (x \circ y) \circ z$. Il est possible que cela ne s'applique pas de manière générale, mais il s'agit plutôt du pouvoir d'association i.e, $x^n \circ x^m = x^{n+m}$ pour tout entier $n, m \geq 1$.

Exemple 2.3 On peut vérifier que l'espace \mathcal{E}^n avec la multiplication de Jordan

$$x \circ y = \begin{pmatrix} x^T y \\ x_0 \bar{y} + y_0 \bar{x} \end{pmatrix}$$

forme une algèbre de Jordan avec le vecteur identité

$$e := \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

Définition 2.1.9 Une algèbre de Jordan \mathcal{J} est appelée euclidienne si est seulement s'il existe un produit scalaire $\langle \cdot, \cdot \rangle$ sur \mathcal{J} telle que pour tout $x, y, z \in \mathcal{J}$ on a :

- $\langle x, x \rangle > 0, \forall x \neq 0$ (définie positive).
- $\langle x, y \rangle = \langle y, x \rangle$ (symétrique).
- $\langle x, y \circ z \rangle = \langle x \circ y, z \rangle$ (associative).

Autrement dit, \mathcal{J} admet une forme quadratique définie positive, symétrique, qui est également associative.

Exemple 2.4 Il est facile de vérifier que l'espace \mathcal{E} avec la multiplication de Jordan \circ définie au dessus, est une algèbre euclidienne de Jordan sous le produit scalaire standard.

Définition 2.1.10 On dit que deux élément x et y sont des opérateurs commutants si est seulement si pour tout $z \in \mathcal{J}$ on a : $x \circ (y \circ z) = y \circ (x \circ z)$.

Nous définissons maintenant le cône de carrés d'une algèbre euclidienne de Jordan

Définition 2.1.11 soit \mathcal{J} est une algèbre euclidienne de Jordan alors son cône de carrés est l'ensemble $\mathcal{J}_{\mathcal{J}} = \{x^2 : x \in \mathcal{J}\}$.

Exemple 2.5 Le cône de carrés de (\mathcal{E}^n, \circ) , avec \circ défini en exemple (2.3), est le cône du second ordre \mathcal{Q}^n .

Notons que

$$\mathcal{Q}^n = \{x = (x_0; \bar{x}) \in \mathbb{R}^n : x_0 \geq \|\bar{x}\|\}.$$

Le résultat fondamental suivant donne une correspondance entre les algèbres euclidiennes de Jordan et les cônes symétriques.

Théorème 2.1.1 (La caractérisation algébrique de Jordan des cônes symétriques) Un cône régulier \mathcal{K} est symétrique si et seulement si $\mathcal{K} = \mathcal{K}_{\mathcal{J}}$ pour une algèbre euclidienne de Jordan \mathcal{J} .

La décomposition spectrale

On définit la décomposition spectrale de $x \in \mathcal{J}$ comme la décomposition de x en vecteurs propres $(c_1(x), c_2(x))$ avec ses valeurs propres $(\lambda_1(x), \lambda_2(x))$ telle que $x = \lambda_1(x)c_1(x) + \lambda_2(x)c_2(x)$.

Les vecteurs propres c_1 et c_2 satisfont les propriétés suivantes :

- $c_1 \circ c_2 = 0$
- $c_1^2 = c_1, c_2^2 = c_2$
- $c_1 + c_2 = e$

Autrement dit une paire $\{c_1, c_2\}$ est un référentiel de Jordan si elle satisfait les propriétés ci-dessus.

Les éléments x et y sont décomposés simultanément s'ils partagent un référentiel de Jordan c'est-à-dire $x = \lambda_1 c_1 + \lambda_2 c_2$ et $y = \omega_1 c_1 + \omega_2 c_2$ pour un référentiel de Jordan $\{c_1, c_2\}$.

Exemple 2.6 La décomposition spectrale de $x \in \mathcal{E}^n$ associée au cône du second ordre est obtenue comme suit

$$x = \underbrace{(x_0 + \|\bar{x}\|)}_{\lambda_1(x)} \underbrace{\left(\frac{1}{2}\right) \begin{pmatrix} \frac{1}{\|\bar{x}\|} \\ \frac{\bar{x}}{\|\bar{x}\|} \end{pmatrix}}_{c_1(x)} + \underbrace{(x_0 - \|\bar{x}\|)}_{\lambda_2(x)} \underbrace{\left(\frac{1}{2}\right) \begin{pmatrix} \frac{1}{\|\bar{x}\|} \\ \frac{-\bar{x}}{\|\bar{x}\|} \end{pmatrix}}_{c_2(x)}$$

Il est possible d'étendre la définition de toute fonction continue à valeurs réelles $f(\cdot)$ aux éléments des algèbres de Jordan en utilisant leurs valeurs propres comme suit :

$$f(x) = f(\lambda_1)c_1 + \dots + f(\lambda_k)c_k.$$

Nous intéressons particulièrement aux éléments suivants :

- L'inverse de x est donnée par :
 $x^{-1} = \lambda_1^{-1}c_1 + \dots + \lambda_k^{-1}c_k$, pour tout $\lambda_i \neq 0$ si $\lambda_i = 0$ alors l'inverse est indéfini.
- La racine carrée de x est définie comme suit :
 $x^{\frac{1}{2}} = \lambda_1^{\frac{1}{2}}c_1 + \dots + \lambda_k^{\frac{1}{2}}c_k$, pour tout $\lambda_i \geq 0$ sinon la racine carrée est indéfinie.

Notons que $(x^{\frac{1}{2}})^2 = x$, si $x^{-1} \circ x = e$ alors x est dite inversible et son inverse est x^{-1} .

Soit $x \in \mathcal{J}$ on dit que x est semi-défini positive si toutes ses valeurs propres sont non négatives est on la note par $x \geq 0$. De même, elle est défini positif si toutes ses valeurs propres sont positives $x > 0$.

En d'autres termes, un élément est semi-défini positif si et seulement s'il réside dans le cône des carrés, et défini positive si et seulement s'il appartient à l'intérieur du cône de carrés.

Nous pouvons également définir divers norme sur \mathcal{J} comme des fonctions de valeurs propres de la même manière de définir des normes unitairement invariantes sur des matrices carrées

$$\|x\|_F = \left(\sum \lambda_i^2\right)^{\frac{1}{2}} = \sqrt{\text{tr}(x^2)}, \quad \|x\|_2 = \max |\lambda_i|.$$

Remarque 2.1.3 – pour chaque $x \in \mathcal{J}$, soit r le plus petit entier tel que l'ensemble $\{e, x, x^2, \dots, x^r\}$ soit linéairement indépendant, alors r est le degré de x que l'on note $\text{deg}(x)$, par ailleurs le rang de \mathcal{J} , $\text{rg}(\mathcal{J})$ est le plus grand $\text{deg}(x)$ de plusieurs nombres $x \in \mathcal{J}$.

- un élément x est dit régulier si son degré est égale au rang de l'algèbre de Jordan
- Observez que $\|e\|_F = \sqrt{r}$.

Exemple 2.7 (inverse, racine carrée, et la norme sur \mathcal{E}_{n+1}^+)

Soit $x = \lambda_1 c_1 + \lambda_2 c_2$ avec $\{c_1, c_2\}$ son référentiel de Jordan, alors :

- $x^{-1} = \frac{1}{\lambda_1} c_1 + \frac{1}{\lambda_2} c_2$.
- $x^{\frac{1}{2}} = \sqrt{\lambda_1} c_1 + \sqrt{\lambda_2} c_2$.
- $\|x\|_F^2 = \lambda_1^2 + \lambda_2^2 = 2x_0^2 + 2\|\bar{x}\|^2 = 2\|x\|^2$.
- $\|x\|_2 = \max\{\lambda_1, \lambda_2\} = |x_0| + \|\bar{x}\|$.
- $\langle x, y \rangle = \text{Tr}(x \circ y) = 2x^T y$.

Nous introduisons maintenant trois majeure cartes, qui jouent un rôle fondamental dans l'élaboration des méthodes de points intérieurs pour l'optimisation conique.

Définition 2.1.12 Soient x et y des éléments d'une algèbre de Jordan \mathcal{J} alors

1. La carte linéaire $L(x) : \mathcal{J} \rightarrow \mathcal{J}$ est définie par $L(x)y := x \circ y$ (de manière équivalente, est défini par $x \circ y = (L(x)y + L(y)x)/2$) pour tout $y \in \mathcal{J}$.

2. L'opérateur quadratique $Q(x, z) : \mathcal{J} \times \mathcal{J} \longrightarrow \mathcal{J}$ est défini par

$$Q(x, z) = L(x)L(z) + L(z)L(x) - L(x \circ z)$$

3. La représentation quadratique de x , $Q(x) : \mathcal{J} \longrightarrow \mathcal{J}$ qui est définie par

$$Q(x) = 2L(x)^2 - L(x^2) = Q(x, x)$$

Remarque 2.1.4 Noter que

$$L(x)e = x \quad \text{et} \quad L(x)x = x^2 \quad \text{aussi} \quad L(e) = Q(e) = I.$$

Exemple 2.8 A partir de l'exemple (2.5) la formule explicite de l'opérateur $L(\cdot)$ pour l'algèbre du cône de second ordre dans l'espace \mathcal{E}^n , peut être immédiatement donnée par :

$$L(x) = \text{Arw}(x) = \begin{pmatrix} x_0 & \bar{x}^T \\ \bar{x} & x_0 I \end{pmatrix}$$

Ici, $\text{Arw}(x)$ est la matrice associée au vecteur $x \in \mathcal{E}^n$, par conséquent l'opérateur quadratique pour l'algèbre du cône du second ordre est donné par :

$$\begin{aligned} Q(x, z) &= \text{Arw}(x)\text{Arw}(z) + \text{Arw}(z)\text{Arw}(x) - \text{Arw}(x \circ z) \\ &= \begin{pmatrix} x^T z & (x_0 \bar{z}^T + z_0 \bar{x}^T) \\ (x_0 \bar{z} + z_0 \bar{x}) & (\bar{x} \bar{z}^T + \bar{z} \bar{x}^T)(x_0 z_0 - \bar{x}^T \bar{z}) I_{n-1} \end{pmatrix} \end{aligned}$$

On peut aussi facilement vérifier que la représentation quadratique de l'algèbre du cône du second ordre est donnée par :

$$\begin{aligned} Q(x) &= 2\text{Arw}^2(x) - \text{Arw}(x^2) \\ &= \begin{pmatrix} \|x\|^2 & 2x_0 \bar{x}^T \\ 2x_0 \bar{x} & \det(x)I + 2\bar{x}(\bar{x})^T \end{pmatrix} \\ &= 2xx^T - \det(x)R. \end{aligned}$$

Enfin puisque " \circ " est bilinéaire et $\text{tr}(x \circ y)$ est une forme quadratique définie positive symétrique qui est associative, $\text{Tr}(x \circ (y \circ z)) = \text{Tr}((x \circ y) \circ z)$. Ainsi, il devient possible de définir le produit scalaire de la manière suivante :

$$\langle x, y \rangle = \text{Tr}(x \circ y).$$

Notez que puisque le produit scalaire $\langle x, y \rangle$ est associatif, il s'ensuit que $L(x)$ est symétrique par rapport à $\langle \cdot, \cdot \rangle$, car $\langle L(x)y, z \rangle = \langle x \circ y, z \rangle = \langle y, x \circ z \rangle = \langle y, L(x)z \rangle$.

Remarque 2.1.5 Soit $x = (x_1; x_2; \dots; x_r)$, $y = (y_1; y_2; \dots; y_r)$ avec $x_i, y_i \in \mathcal{E}^{n_i}$ $i = 1; \dots; r$ Alors ,

- $x^{-1} := (x_1^{-1}; x_2^{-1}; \dots; x_r^{-1})$.
- $x \circ y := (x_1 \circ y_1; x_2 \circ y_2; \dots; x_r \circ y_r)$.
- $x^T y := x_1^T y_1 + x_2^T y_2 + \dots + x_r^T y_r$.
- $Arw(x) := Arw(x_1) \oplus Arw(x_2) \oplus \dots \oplus Arw(x_r)$.
- $Q_x := Q_{x_1} \oplus Q_{x_2} \oplus \dots \oplus Q_{x_r}$.
- $f(x) := (f(x_1); f(x_2); \dots; f(x_r))$.
- $\|x\|_F^2 := \|x_1\|_F^2 + \|x_2\|_F^2 + \dots + \|x_r\|_F^2$.
- $e := (e_{n_1}; e_{n_2}; \dots; e_{n_r})$ le vecteur identité de \mathcal{E}_r^n .

2.1.2 Les propriétés algébrique des (SOCOs)

Dans cette sous section , nous avons besoin d'introduire quelques propriétés algébrique du cône de second ordre \mathcal{K} et son algèbre euclidienne de Jordan associée.

Pour tout vecteur $x, s \in \mathbb{R}^n$, l'opérateur bilinéaire \circ est donné par :

$$x \circ s = (x^T s; x_1 s_2 + s_1 x_2, \dots, x_1 s_n + s_1 x_n). \quad (2.1.1)$$

Evidemment le produit de Jordan \circ est commutatif de plus (\mathbb{R}^n, \circ) est une algèbre de Jordan .

Pour chaque vecteur $x \in \mathbb{R}^n$, nous associons une matrice en forme de flèche (symétrique) $L(x)$ comme suit :

$$L(x) := \begin{pmatrix} x_1 & x_{2:n}^T \\ x_{2:n} & x_1 E_1 \end{pmatrix} \quad (2.1.2)$$

où $x_{2:n} = (x_2; \dots; x_n)$

Propriétés

- 1) $x \in \mathcal{K}$ si et seulement si $L(x)$ est semi-définie positif.
- 2) $x \in \mathcal{K}$ si et seulement si $x = z \circ z$ pour certain $z \in \mathbb{R}^n$.
- 3) $x \circ s = L(x)x = s \circ x$.
- 4) $\forall x \in \mathbb{R}^n$, $e \circ x = x$ où $e = (1, 0, \dots, 0)$.
- 5) $x \in \mathcal{K}_+$ alors $L(x)e = x^{-1}$ (c'est-à-dire : $x \circ L^{-1}(x)e = 0$).

Remarque 2.1.6 – La première propriété montre que $x \rightarrow L(x)$ offre un prolongement naturel de \mathcal{K} dans le cône des matrices de taille $n \times n$ Semi définie positive.

- La deuxième propriétés implique que \mathcal{K} est l'ensemble des carres de l'algèbre de Jordan.
- Dans 4) : e désigne l'élément unitaire de l'algèbre de Jordan.
- De 5) on voit que x^{-1} est l'inverse de x dans cette algèbre.

Notons par λ_{max} , λ_{min} les valeurs propres maximales est minimales de $L(x)$ qui donnés par

$$\lambda_{max} := x_1 + \|x_{2:n}\| \quad , \quad \lambda_{min} := x_1 - \|x_{2:n}\| \quad (2.1.3)$$

il s'ensuit que

$$x \in \mathcal{K} \Leftrightarrow \lambda_{min} \geq 0 \quad , \quad x \in \mathcal{K}_+ \Leftrightarrow \lambda_{min}(x) > 0. \quad (2.1.4)$$

Définition 2.1.13 Soit $x \in \mathbb{R}^n$, la décomposition spectrale de x associée au cône du second ordre \mathcal{K} est donnée par

$$x := \lambda_{max}(x)z_1 + \lambda_{min}(x)z_2. \quad (2.1.5)$$

où

$$z_1 := \frac{1}{2} \left(1, \frac{x_{2:n}}{\|x_{2:n}\|} \right)^T \quad , \quad z_2 := \frac{1}{2} \left(1, \frac{-x_{2:n}}{\|x_{2:n}\|} \right)^T.$$

par convention

$$\frac{-x_{2:n}}{\|x_{2:n}\|} = 0 \quad \text{si} \quad x_{2:n} = 0.$$

il est facile de vérifie que

$$z_i^2 = z_i \quad \text{pour} \quad i = 1..2, \quad z_1 + z_2 = e \quad \text{et} \quad z_1 \circ z_2 = 0.$$

Lemme 2.1.1 Soit $x, s \in \mathbb{R}^n$ alors

$$\lambda_{min}(x) + \lambda_{min}(s) \leq \lambda_{min}(x+s) \leq \lambda_{max}(x+s) \leq \lambda_{max}(x) + \lambda_{max}(s).$$

Preuve 2.1.1 en utilisant l'inégalité triangulaire on obtient

$$\lambda_{min}(x+s) := x_1 + s_1 - \|(x+s)_{2:n}\| \geq x_1 + s_1 - \|x_{2:n}\| - \|s_{2:n}\| = \lambda_{min}(x) + \lambda_{min}(s).$$

et

$$\lambda_{max}(x+s) := x_1 + s_1 + \|(x+s)_{2:n}\| \leq x_1 + s_1 + \|x_{2:n}\| + \|s_{2:n}\| = \lambda_{max}(x) + \lambda_{max}(s).$$

d'où le résultat

Lemme 2.1.2 Soit $x \in \mathbb{R}^n$ alors

$$x - \lambda_{\min}(x)e \geq 0_{\mathcal{K}} \quad \text{et} \quad \lambda_{\max}(x)e - x \geq 0_{\mathcal{K}}.$$

Preuve 2.1.2 D'après la définition (2.1.13) et (4) et le fait que $z_1 + z_2 = e$

$$x - \lambda_{\min}(x)e = (\lambda_{\max}(x)z_1 + \lambda_{\min}(x)z_2) - \lambda_{\min}(x)(z_1 + z_2) = (\lambda_{\max}(x) - \lambda_{\min}(x))z_1.$$

aussi

$$\lambda_{\max}(x)e - x = \lambda_{\max}(x)(z_1 + z_2) - (\lambda_{\max}(x)z_1 + \lambda_{\min}(x)z_2) = (\lambda_{\max}(x) - \lambda_{\min}(x))z_2.$$

Corollaire 2.1.1 Soit $x \in \mathbb{R}^n$ et $\|x\|_F \leq 1$ alors

$$e - x \geq 0_{\mathcal{K}}.$$

Lemme 2.1.3 Soit $x \in \mathbb{R}^n$ alors

$$|\lambda_{\min}(x)| \leq \sqrt{2} \|x\| \quad \text{et} \quad |\lambda_{\max}(x)| \leq \sqrt{2} \|x\|.$$

Preuve 2.1.3 par la définition de $\lambda_{\min}(x)$ et $\lambda_{\max}(x)$ nous avons

$$\lambda_{\max}^2 + \lambda_{\min}^2 = 2 \|x\|^2.$$

ce qui complète la preuve

Lemme 2.1.4

$$\lambda_{\min}(x + s) \geq \lambda_{\min}(x) - \sqrt{\|s\|}.$$

Définition 2.1.14 La trace et le déterminant de $x \in \mathbb{R}^n$ associe a \mathcal{K} sont définis par

$$\text{Tr}(x) := \lambda_{\max}(x) + \lambda_{\min}(x) = 2x_1. \quad (2.1.6)$$

$$\det(x) := \lambda_{\max}(x)\lambda_{\min}(x) = x_1^2 - \|x_{2:n}\|^2. \quad (2.1.7)$$

Lemme 2.1.5 Pour tout $x, s \in \mathbb{R}^n$ on a

$$\text{Tr}(x \circ s) = 2x^T s \quad , \quad \text{Tr}(x \circ x) = 2 \|x\|^2. \quad (2.1.8)$$

il est facile de vérifier que

$$\text{Tr}((x \circ s) \circ t) = \text{Tr}(x \circ (s \circ t)).$$

Lemme 2.1.6 (lemme 6.2.3 dans [26]) Soit $x, s \in \mathbb{R}^n$ alors

$$\lambda_{\min}(x)\lambda_{\min}(s) + \lambda_{\min}(x)\lambda_{\max}(s) \leq \text{Tr}(x \circ s) \leq \lambda_{\max}(x)\lambda_{\max}(s) + \lambda_{\min}(x)\lambda_{\min}(s).$$

et

$$\det(x \circ s) \leq \det(x)\det(s).$$

La dernière inégalité devient égalité si et seulement si les vecteurs $x_{2:n}$ et $s_{2:n}$ sont linéairement dépendants

Corollaire 2.1.2 (corollaire 2.4 Dans [64]) Soit $x \in \mathbb{R}^n$ et $s \in \mathcal{K}$ alors

$$\lambda_{\min}(x)\text{Tr}(s) \leq \text{Tr}(x \circ s) \leq \lambda_{\max}(x)\text{Tr}(s).$$

Preuve 2.1.4 comme $s \in \mathcal{K}$ alors $\lambda_{\min}(s) \geq 0$ donc $\lambda_{\max}(s) \geq 0$ le lemme précédent implique que

$$\lambda_{\min}(x)(\lambda_{\max}(s) + \lambda_{\min}(s)) \leq \text{Tr}(x \circ s) \leq \lambda_{\max}(x)(\lambda_{\max}(s) + \lambda_{\min}(s)).$$

et comme $\text{Tr}(s) = \lambda_{\max}(s) + \lambda_{\min}(s)$ ce qui achève la preuve

Définition 2.1.15 Pour tout $x \in \mathbb{R}^n$ on a

$$P(x) := 2L(x)^2 - L(x^2).$$

par conséquent

$$P(x) = \begin{pmatrix} \|x\|^2 & 2x_1 x_{2:n}^T \\ 2x_1 x_{2:n} & \det(x)I + 2x_1 x_{2:n}^T \end{pmatrix}$$

L'importance de la décomposition spectrale est qu'elle nous permet d'étendre toute fonction continue à valeurs réelles à l'algèbre des cônes du second ordre.

La fonction barrière induite par la fonction noyau

Supposons maintenant que $\psi : \mathbb{R}_{++} \rightarrow \mathbb{R}_+$ est une fonction réelles et différentiable sur $]0, +\infty[$ telle que $\psi(t) > 0, \forall t > 0$.

Définition 2.1.16 (Définition 2.3 dans [25]) Soit $x \in \mathbb{R}^n$ avec la décomposition spectrale définie dans (2.1.5). La fonction vectorielle $\psi(x)$ est définie par

$$\psi(x) := \psi(\lambda_{\max}(x))z_1 + \psi(\lambda_{\min}(x))z_2. \quad (2.1.9)$$

De plus en remplaçant $\psi(\lambda_{\max}(x))$ et $\psi(\lambda_{\min}(x))$ dans (2.1.9) par $\psi'(\lambda_{\max}(x))$ et $\psi'(\lambda_{\min}(x))$ respectivement, on peut conclure que la fonction à valeur vectorielle

$$\psi'(x) := \psi'(\lambda_{\max}(x))z_1 + \psi'(\lambda_{\min}(x))z_2. \quad (2.1.10)$$

Si $\psi(t)$ est deux fois différentiable alors

$$\psi''(x) := \psi''(\lambda_{\max}(x))z_1 + \psi''(\lambda_{\min}(x))z_2. \quad (2.1.11)$$

Lemme 2.1.7 Soit $\psi : \mathbb{R}_{++} \rightarrow \mathbb{R}_+$ et $x \in \mathcal{K}_+$ alors $\psi(x) \in \mathcal{K}$.

Preuve 2.1.5 On sait que z_1 et z_2 appartient à \mathcal{K} et comme $x \in \mathcal{K}_+$ ses valeurs propres sont positives, donc $\psi(\lambda_{\max}(x))$ et $\psi(\lambda_{\min}(x))$ sont non négative, par conséquent $\psi(x) \in \mathcal{K}$.

Lemme 2.1.8 Soit $z_1, z_2 \in \mathcal{K}$ tel que $\|z_1\| = \|z_2\| = \frac{1}{\sqrt{2}}$ et $z_1^T z_2 = 0$. si $x_{2:n} \neq 0$ alors

$$\|\psi(x)\| = \frac{1}{\sqrt{2}} \sqrt{\psi(\lambda_{\max}(x))^2 + \psi(\lambda_{\min}(x))^2}. \quad (2.1.12)$$

L'expression (2.1.12) valable si $x_{2:n} \neq 0$ car $\lambda_{\max}(x) = \lambda_{\min}(x) = x_1$ d'où $\|\psi(x)\| = |\psi(x_1)|$ et $\psi(\lambda_{\max}(x))^2 + \psi(\lambda_{\min}(x))^2 = 2\psi(x_1)^2$.

On associe à chaque fonction noyau $\psi(t)$ une fonction à valeur réelle positive $\Psi(x)$ sur \mathcal{K}_+ comme suit

$$\Psi(x) := \text{Tr}(\psi(x)) = 2(\psi(x))_1 = \psi(\lambda_{\max}(x)) + \psi(\lambda_{\min}(x)), \quad x \in \mathcal{K}_+. \quad (2.1.13)$$

Notons que $(\psi(x))_1$ désigne la première coordonnée de $\psi(x)$.

Théorème 2.1.2 (proposition 6.2.9 dans [63]) $\Psi(x)$ est une fonction positive strictement convexe par rapport à $x \in \mathcal{K}_+$ et s'annule en son point minimal global $x = e = (1; 0; \dots; 0)$, c'est-à-dire

$$\Psi(x) = 0 \Leftrightarrow \psi(x) = 0 \Leftrightarrow \psi'(x) = 0 \Leftrightarrow x = e. \quad (2.1.14)$$

Autrement dit $\psi(t) \geq 0$ pour tout $t > 0$ et $\lambda_{\max}(x) \geq \lambda_{\min}(x) > 0$ donc $\Psi(x) \geq 0$ pour tout $x \in \mathcal{K}_+$.

De plus $\psi(t) = 0$ si et seulement si $t = 1$ donc $\Psi(x) = 0$ si et seulement si $\lambda_{\max}(x) = \lambda_{\min}(x) = 1$ c'est-à-dire si et seulement si $x = e$.

Remarque 2.1.7 Lorsque x appartient à \mathcal{K}_+ et tend vers la frontière du \mathcal{K} alors la valeur de $\lambda_{\min}(x)$ tend vers zéro, et donc $\Psi(x)$ tend vers l'infini, par suite $\Psi(x)$ est une fonction pour le cône \mathcal{K} .

2.1.3 Le produit cartésien du cône

Avant de terminer cette partie nous représentons la manière d'adapter les définitions et les propriétés des sous-section précédentes au cas général où \mathcal{K} le produit cartésien des cônes du second ordre

$$\mathcal{K} = \mathcal{K}^1 \times \mathcal{K}^2 \times \dots \times \mathcal{K}^N.$$

Pour commencer on partitionne tout vecteur $x \in \mathbb{R}^n$ selon la dimension des cônes successifs \mathcal{K}^j , par suite on a :

$$x = (x^1; \dots; x^N), \quad s = (s^1; \dots; s^N), \quad (2.1.15)$$

où $x^j, s^j \in \mathbb{R}^{n_j}$ avec $n = \sum_{j=1}^N n_j$

et on définit le produit direct de l'algèbre de Jordan comme

$$x \circ s = (x^1 \circ s^1; x^2 \circ s^2; \dots; x^N \circ s^N). \quad (2.1.16)$$

Notons par e^j l'élément unitaire du \mathcal{K}^j alors

$$e = (e^1; e^2; \dots; e^N). \quad (2.1.17)$$

est l'élément unitaire du cône \mathcal{K} , de plus $Tr(e) = 2N$.

Soit

$$L(x) = \text{diag}(L(x^1); L(x^2); \dots; L(x^N)). \quad (2.1.18)$$

et

$$P(x) = \text{diag}(P(x^1); P(x^2); \dots; P(x^N)). \quad (2.1.19)$$

Evidement le point w de x et s est donné par

$$w = (w^1; w^2; \dots; w^N)$$

et $P(w) = (P(w^1); P(w^2); \dots; P(w^N))$, cette matrice est symétrique définie positive, est $P(w)s = x$, c'est-à-dire un automorphisme de \mathcal{K} .

aussi on a

$$\lambda_{\max}(x) = \max\{\lambda_{\max}(x^j), \quad 1 \leq j \leq N\}. \quad (2.1.20)$$

$$\lambda_{\min}(x) = \min\{\lambda_{\min}(x^j), \quad 1 \leq j \leq N\}. \quad (2.1.21)$$

de plus

$$\begin{aligned} \text{Tr}(x) &= \sum_{j=1}^N \text{tr}(x^j) = \sum_{j=1}^N (\lambda_{\max}(x^j) + \lambda_{\min}(x^j)). \\ \det(x) &= \prod_{j=1}^N \det(x^j) = \prod_{j=1}^N \lambda_{\max}(x^j) \lambda_{\min}(x^j). \\ \|x\|_F^2 &= \sum_{j=1}^N \|x^j\|_F^2 = \sum_{j=1}^N (\lambda_{\max}(x^j)^2 + \lambda_{\min}(x^j)^2) = 2 \|x\|^2. \end{aligned}$$

La fonction $\psi(x)$ de l'algèbre du cône du second ordre est défini comme suit

$$\psi(x) := \sum_{j=1}^N (\psi(\lambda_{\max}(x^j))z_1 + \psi(\lambda_{\min}(x^j))z_2). \quad (2.1.22)$$

Si $\psi(t) \geq 0$ et $x \in \mathcal{K}$ alors l'équation (2.1.22) implique que $\psi(x) \in \mathcal{K}$.

de manière analogue

$$\psi'(x) := \sum_{j=1}^N (\psi'(\lambda_{\max}(x^j))z_1 + \psi'(\lambda_{\min}(x^j))z_2).$$

Les lemmes suivants seront utilisé dans l'analyse de l'algorithme

Lemme 2.1.9 (lemme 2.6 Dans [33]) Pour tout $x \in \mathcal{K}$ on a

$$\text{Tr}(x) \leq \sqrt{2} \|x\|_F$$

Preuve 2.1.6 on sait que $\|x\|_F = \sqrt{\langle x, x \rangle} = \sqrt{\text{Tr}(x^2)} = \sqrt{2} \|x\|$ aussi par (2.1.6) on a :

$$\text{Tr}(x) = 2x_1 \leq 2 \|x\| = \sqrt{2} \|x\|_F$$

Lemme 2.1.10 (lemme 2.13 Dans [33]) Pour tout $x, s \in \mathbb{R}^n$ on a

$$\|x^2\|_F \leq \|x\|_F^2.$$

On termine cette partie en introduisant les dérivées par rapport au paramètre réel t des fonction $(\psi(x(t)), \Psi(x(t)))$, où $x(t) = (x_1(t), \dots, x_n(t))$ on étend les concept usuels de dérivabilité, intégrabilité et continuité

Dénotant

$$x'(t) = (x'_1(t), x'_2(t), \dots, x'_n(t)).$$

on a alors

$$\frac{d}{dt}(x(t) \circ s(t)) = x'(t) \circ s(t) + x(t) \circ s'(t). \quad (2.1.23)$$

et

$$\frac{d}{dt} Tr(\psi(x(t))) = 2 \frac{d\psi(x(t)_1)}{dt} = 2\psi'(x(t))^T x'(t) = Tr(\psi'(x(t)) \circ x'(t)) \quad (2.1.24)$$

2.2 Diverse méthodes pour résoudre un problème d'optimisation conique

2.2.1 Les méthodes de projection

Certaines des approches de résolution que nous aurons à implémenter et à introduire dans cette thèse est la notion de projection. De nombreuses études ont été proposées pour arriver aux différents types de méthodes de projection, on peut citer les travaux récents de Bregman, Censor, Reich et Zepkowitz-Malachi. Afin de construire itérativement la projection de tout point sur l'intersection d'ensembles convexes fermés non vides, on distingue deux types de méthodes : les travaux de Bauschke et Combettes qui s'intéressent aux méthodes de projections parallèles relaxées et les méthodes de projections alternées dû à Boyle et Dykstra. Nous allons procéder à l'introduction des méthodes de projection alternée qui permettent de construire itérativement le point de l'intersection le plus proche d'un point initial donné. Cette approche a été présentée par Dykstra en 1983 dans le cas particulier où les convexes sont des cônes et où l'on est en dimension finie, en 1986 Boyle a repris le problème dans le cas des convexes quelconques puis par Bauschke et Borwein qui ont expliqué leurs propriétés de convergence, ensuite Glunt, Al et Escalante ont étendu cet algorithme à un autre type de problème de projection, en 1988 Han a découvert un contexte de dualisation d'un problème d'optimisation on lui a donné le nom de méthode des projections successives, En 1994, Borwein et Bauschke ont présenté une étude de cette méthode de projections alternées dans le cas de deux convexes.

Position du problème

Soit le problème conique suivant :

$$(P) \begin{cases} \min_x c^T x \\ Ax = b \\ x \in \mathcal{K} \end{cases} \quad (2.2.1)$$

et son dual

$$(D) \begin{cases} \max b^T y \\ A^T y + s = c \\ s \in \mathcal{K} \end{cases} \quad (2.2.2)$$

Le problème de projection conique

L'un des domaines de recherche les plus actifs qui à été introduit est la programmation conique linéaire qui stimule par de nombreuses méthodes développées et des applications efficaces Notre étude s'intéresse à la projection d'un point C qui appartient à l'espace \mathbb{R}^n menu du produit scalaire standard sur l'intersection $\mathcal{K} \cap \mathcal{A}$ où \mathcal{K} est un cône convexe fermé dans \mathbb{R}^n avec $\text{int}K \neq \emptyset$ (de plein dimension dans \mathbb{R}^n) et \mathcal{A} est un sous espace affine (un polyèdre convexe) définie par les égalités affines

$$\mathcal{A} := \{Ax = b, x \in \mathcal{K}\}. \quad (2.2.3)$$

Le fait que \mathcal{A} soit définie par des égalités alors on s'assure que notre problème et plus proche des problèmes de projection qui se présente en pratique, par conséquent l'expression explicite de la projection de x est :

$$\text{Proj}_{\mathcal{A}}(x) := x - A^T [AA^T]^{-1} (Ax - b). \quad (2.2.4)$$

L'idée de base de cette approche est de calculer un point réalisable par la projection sur l'intersection $\mathcal{K} \cap \mathcal{A}$ alors le problème de projection est soumis à des contraintes affines et des contraintes conique afin de minimiser la norme (au carré) de sorte qu'il ait une formulation analytique évidente en tant que problème de moindres carrés :

$$\begin{cases} \min_x \frac{1}{2} \|x - c\|^2 \\ Ax = b \\ x \in \mathcal{K} \end{cases} \quad (2.2.5)$$

où $c \in \mathbb{R}^n$ est le point à projeter .

Généralement on peut facilement projeter sur le sous espace affine \mathcal{A} ou sur le cône K , la difficulté réside dans la projection sur l'intersection $\mathcal{K} \cap \mathcal{A}$ (car elle pose deux contraintes l'une est affine et l'autre conique), de nombreuses méthodes numériques ont été proposées pour faciliter la tâche . Derrière le terme projections coniques découlent cinq différents types de méthodes :

- La méthode du point intérieur .
- La méthode de projection alternée .
- La méthode de direction alternative.
- Les méthodes basées sur l'utilisation de la programmation conique linéaire.
- Les méthodes dual.

1) La méthode du point intérieur

Puisque le problème (2.2.3) est convexe , on peut le résoudre à l'aide des méthodes de point intérieur la première méthode a été proposée dans [40] pour un type de problème de matrice de corrélation ensuite [47] pour les SDPs quadratiques général , notre problème (2.2.5) peuvent être attaqués en utilisant des machinerie de point intérieur [71] , supposons maintenant que le cône \mathcal{K} et son polaire sont équipés de barrière i.e :

$$\mathcal{K}^\circ := \{s \in \mathbb{R}^n : s^T x \leq 0, \text{ pour tout } x \in \mathcal{K}\}$$

et introduisons les conditions d'optimalité perturbée de (2.2.5)

$$\bar{x} \in \mathcal{K} \cap \mathcal{A} (c - \bar{x})^T (x - \bar{x}) \leq 0, \text{ pour tout } x \in \mathcal{K} \cap \mathcal{A}.$$

L'approche repose sur la résolution de ces conditions mieux que les conditions principales du problème , alors pour rendre les contraintes d'optimalité plus concrètes formulons les contraintes affines en utilisant $A_E \in \mathbb{R}^{n \times m_E}$ et $A_I \in \mathbb{R}^{n \times m_I}$ comme

$$\begin{cases} \min \|x - c\|^2 \\ A_E x = b_E, \quad A_I x \leq b_I \\ x \in \mathcal{K} \end{cases} \quad (2.2.6)$$

Sous une hypothèse de non-dégénérescence , les conditions (2.2.6) produisent le système de complémentarité suivant :

$$\begin{cases} x - c + \mu + A_E^T + A_I^T z = 0 \\ A_E x = b_E, y \in \mathbb{R}^{m_E} \\ A_I x \leq b_I, z \in \mathbb{R}_+^{m_I}, z^T (A_I x - b_I) = 0 \\ x \in \mathcal{K}, \mu \in \mathcal{K}^0, \mu^T x = 0 \end{cases}$$

En d'autres termes , l'approche des points intérieurs est basée sur la perturbation des équations complémentaires ci-dessus et le maintien des autres équations satisfaites .

2) La méthode de projection alternée

L'idée de base de cette approche intuitive est de projeter le point initial sur le premier ensemble \mathcal{A} ensuite à projeter le nouveau point sur le deuxième ensemble \mathcal{K} puis en projetant à nouveau le nouveau point sur le premier (\mathcal{A}) est continuer l'opération alternativement. Autrement dit il consiste à répéter

$$\begin{cases} x_{k+1} = Proj_{\mathcal{K}}(y_k) \\ y_{k+1} = Proj_{\mathcal{A}}(x_{k+1}) \end{cases} \quad (2.2.7)$$

On distingue deux phases différentes pour la convergence de l'algorithme

- La première phase : si l'intersection $\mathcal{K} \cap \mathcal{A}$ est régulière alors l'algorithme converge linéairement vers un point de l'intersection , de plus nous pouvons estimer la vitesse de convergence (pour le cas convexe voir [19] , et pour le cas général voir le résultat de [2])
- La deuxième phase consiste à ajouter une étape de correction pour chaque itération de (2.2.7) appelée correction de Dykstra[49]

$$\begin{cases} x_{k+1} = Proj_{\mathcal{K}}(z_k) \\ y_{k+1} = Proj_{\mathcal{A}}(x_{k+1}) \\ x_{k+1} = z_k - (x_{k+1} - y_{k+1}) \end{cases} \quad (2.2.8)$$

Cette correction assure la convergence de la suite $(x_k)_k$ non seulement vers un point de l'intersection $\mathcal{A} \cap \mathcal{K}$ mais exactement vers la projection $Proj_{\mathcal{A} \cap \mathcal{K}}(c)$.

Notez que les méthodes de direction alternée et les méthodes duales peuvent être traduites comme des variantes de cette méthode géométrique .

3) La méthode de direction alternative

La méthode des directions alternées a été proposer par [8] pour les problèmes de projection semi définie , ensuite par [31] pour les SDPs quadratique .

L'idée de base de cette méthode est d'utiliser les propriétés de la structure séparable du problème en répétant les variable pour obtenir le problème équivalent.

$$\begin{cases} \min \frac{1}{2} \|x - c\|^2 + \frac{1}{2} \|y - c\|^2 \\ x = y \\ x \in \mathcal{K}, y \in \mathcal{A} \end{cases} \quad (2.2.9)$$

d'ou l'approche appliqué à (2.2.9) donne ce schéma.

introduisons la fonction de lagrange augmentée suivante

$$L(x, y, z) := \frac{1}{2} \|x - c\|^2 + \frac{1}{2} \|y - c\|^2 - \langle z, x - y \rangle + \frac{\beta}{2} \|x - y\|^2 .$$

La décomposition de la minimisation de la fonction lagrangienne par rapport aux variable primal (x, y) se fait en deux étapes , afin qu'une itération de lagrangienne soit

$$\begin{cases} x_{k+1} = \operatorname{argmin}_{x \in \mathcal{K}} L(x, y_k, z_k) \\ y_{k+1} = \operatorname{argmin}_{y \in \mathcal{A}} L(x_{k+1}, y, z_k) \\ z_{k+1} = z_k - \beta(x_{k+1} - y_{k+1}) \end{cases}$$

nous résumons les minimisations ci-dessus (il est facile de prouver)aux projections suivantes :

plus précisément

$$x_{k+1} := \operatorname{Proj}_{\mathcal{K}}\left(\frac{\beta y_k + z_k + c}{1 + \beta}\right), \quad y_{k+1} := \operatorname{Proj}_{\mathcal{A}}\left(\frac{\beta x_{k+1} - z_k + c}{1 + \beta}\right).$$

ainsi pour calculer la projection sur l'intersection $\mathcal{K} \cap \mathcal{A}$ nous alter-nons les projections sur \mathcal{A} et \mathcal{K} .

Donc cette approche peut être considéré comme une modification du schéma de projection alternée (2.2.8) avec le même concept de Dykstra .

4) Les méthodes basées sur l'utilisation de la programmation conique linéaire

L'approche consiste à reformuler un problème de projection comme un problème de programmation conique usuel de sorte que nous pourrions utiliser les outils puissants développés pour ce cas, parmi les moyens approuvés est d'abaisser la fonction objectif avec une variable supplémentaire donc (2.2.5) est équivalent au problème suivant :

$$\begin{cases} \min t \\ x \in A \\ x - c = z \\ (x, (z, t)) \in \mathcal{K} \times L_{n+1} \end{cases}$$

ou la variable $z \in \mathbb{R}^n$ apparaît dans les contraintes désigne la contrainte supplémentaire de cône du second ordre, il existe plusieurs solveurs pour cette méthode par exemple les méthodes aux points intérieurs comme DDPT3 [48] ou SeDuMi [24]. Cette méthode n'est pas applicable en générale (voir [45], [50]) à l'aide de tests numériques, d'autre part l'ajout des variables (z, t) augmente le coût de calcul et l'espace mémoire affecté pour la méthode de point intérieur primal-dual.

comme le problème de projection (2.2.5) est un problème d'optimisation conique quadratique, il s'agit donc d'un cas particulier de la programmation conique non linéaire et pour le résoudre nous utilisons des logiciels dédiés à la programmation conique non linéaire par exemple la méthode de pénalisation.

L'inconvénient majeur de cette méthode est qu'elle n'est efficace que pour les problèmes de projection à petite taille aussi elle n'utiliserait pas la structure particulière de (2.2.5).

Notez que nous ne savons pas comment créer des algorithmes de pénalisation personnalisés pour le problème (2.2.5), donc cette méthode ne convient pas.

5) Les méthodes dual

Cette approche a été présentée pour les problèmes de moindre carrés conique dans [25] puis revisitée par [56] pour les cônes semi-définis positifs, puis étendue par [22] et [50] pour la projection sur des matrices de corrélation.

notez que le problème (2.2.5) semble plus compliqué qu'un pro-

blème conique avec une fonction objective linéaire au lieu d'une norme .

Il est clair que la forte convexité de la fonction objective aide à résoudre le problème de la programmation conique (2.2.5) avec cette approche dual ainsi qu'elle donne de bonnes propriétés , nous fournissons plus de détail dans la partie suivante .

Une comparaisons entre les méthodes de Projections alternées vs Points intérieurs

Pour terminer notre étude nous avons comparer entre la méthode de projection alternée avec l'approche du point intérieur un point de vue théorique l'algorithme des projections alternées ne nécessite que le calcul préalable des projections sur les convexes simples que l'on peut obtenir , il s'avère donc plus simple , au contraire l'algorithme des points intérieurs nécessite la maîtrise de l'algèbre linéaire numérique et quelques connaissances en analyse numérique .

d'un point de vue exécutif l'algorithme des points intérieurs a une vitesse de convergence rapide (quadratique) aussi une meilleure qualité de robustesse puis une grande précision dans les résultats , d'une part l'algorithme des projections alternées a une convergence sous-linéaire donc une difficulté dans la précision des résultats .

Comme nous l'avons annoncé précédemment , le problème (2.2.5) peut être résolu en utilisant uniquement les méthodes de programmation coniques habituelles mais ce n'est pas un bon concept , donc pour le résoudre, nous pouvons utiliser une succession de problèmes pour obtenir une estimation de la valeur optimale du problème .

Nous passons en revue les méthodes de résolution d'un problème de projection conique, nous exposons les idées principales et les points techniques ainsi que nous discutons les algorithmes, en plus des méthodes de point intérieur, nous posons les méthodes de projection alternée plus précisément les méthodes duales , le principe de ces approches est de séparer les deux ensembles de contraintes K et A et d'utiliser une succession de projection sur eux, donc l'approche que nous adoptons se généralise dans deux directions :

- premièrement ,nous pourrions remplacer le cône K par un ensemble convexe fermé et nous gardons les mêmes développement .
- deuxièmement on pourrait proposer des problèmes avec des fonction objectifs quadratiques .

En particulier les problèmes de projection sont très important, pour cela il faut concevoir des algorithmes spécifiquement pour eux .

L'approche de projections alternées

Cette approche directe consiste à faire une projection sur l'intersection de deux ensembles convexes "plus simples" sur lesquels on sait faire des projections , en d'autres termes , la solution adoptée consiste à approcher ces projections connues pour construire itérativement la projection que l'on recherche , elle peut donc être considérée comme une manière naturelle de traiter le problème .

Nous avons observé que le problème (2.2.5) a une solution unique pour tout c donné , si le domaine des réalisable n'est pas vide , donc le choix du point c dépend d'une connaissance préalable du problème de réalisabilité conique , on peut choisir $c = 0$ par défaut , ce choix est souvent intéressant.

On pratique le problème (2.2.5) est plus facile a résoudre que le problème (2.2.1) même s'il semble compliqué(sa fonction objectif est quadratique), il apparait que la coercivité de la fonction objective offre des propriétés intéressant au problème dual , qui peut être résolu par des projection successive et efficace sur le cône , c'est le cas pour les cône (polyédrique ,du second ordre).

L'approche dual

L'approche dual a été présenté par J . Malick dans [25] puis Boyd and Xiao fait des études profonde sur le cas du cône SDP , ensuite (Borsdorf , Higham and Q) [50] , revisité pour les problèmes de projection particulière .

L'algorithme dual

On utilise l'approche duale de [25] pour résoudre le problème (2.2.5) nous exploitons le mécanisme standard de la dualité et nous proposons l'algorithme de projection (dual) que nous illustrons plus tard .

Notons que le problème (2.2.5) a une solution unique si l'intersection de ces deux convexes fermés $\mathcal{K} \cap \mathcal{A}$ est non vide à savoir la projection du point c sur cette intersection .

Maintenant nous faisons les hypothèses suivantes :

- L'intersection $\mathcal{K} \cap \mathcal{A}$ est non vide.

- Le cône \mathcal{K} est de dimension plein ($\text{int}\mathcal{K} \neq \emptyset$).
- Les contraintes affines sont linéairement indépendante (rang $A = m$).

L'idée de l'approche du [25] consiste à utiliser le principe de la dualité lagrangienne basé sur la dualisation des contraintes affines (voir chapitre XII de [23]) et bénéficier aux caractéristique géométrique du problème .
Le lagrangienne du variable primale $x \in \mathcal{K}$ et du variable dual $y \in \mathbb{R}^m$ est le suivant :

$$L(x, y) := \frac{1}{2} \|c - x\|^2 - y^T(Ax - b). \quad (2.2.10)$$

et la fonction concave dual correspondante qui doit être maximisée est :

$$\theta(y) := \min_{x \in \mathcal{K}} L(x, y). \quad (2.2.11)$$

Observez qu'il n'y a plus de contrainte affine dans l'expression (2.2.11)
Notons par $P_{\mathcal{K}}$ l'opérateur de projection sur \mathcal{K} . On peut prouver [25 ,théorème 3,1] que (2.2.11) ayant un unique point qui atteint le minimum à savoir

$$x(y) := P_{\mathcal{K}}(c + A^T y). \quad (2.2.12)$$

donc nous avons

$$\theta(y) := -\frac{1}{2} \|x(y)\|^2 + b^T y. \quad (2.2.13)$$

Il est facile de montrer [25, théorème 3,2] que la fonction θ est différentiable sur \mathbb{R}^m et que son gradient

$$\nabla\theta(y) := Ax(y) + b \quad (2.2.14)$$

est lipschitz continue, comme toutes les fonctions qui à un gradient de lipschitz , θ est deux fois différentiable presque partout (par propriété de la dérivabilité du projection)
donc le problème dual est le suivant

$$\begin{cases} \min \theta(y) \\ y \in \mathbb{R}^m \end{cases} \quad (2.2.15)$$

sous l'hypothèse d'optimisation convexe standard , disons l'hypothèse (faible) le problème (2.2.15) a une forte dualité c'est -à-dire $\exists \bar{x} \in \mathcal{A} \cap \text{int}\mathcal{K}$ donc elle donne en plus qu'il existe des solution à (2.2.15) .

Autrement dit pour calculer la projection de c sur $\in \mathcal{A} \cap \mathcal{K}$, il suffit de résoudre le problème dual (2.2.15), plus profondément, les contraintes de simples contraintes de positivité et la fonction dual est concave, différentiable à gradient de lipschitz.

L'importance pratique de cette régularité est d'ouvrir la voie à l'utilisation d'algorithmes d'optimisation sans contraintes.

par conséquent nous pouvons utiliser l'une des méthodes numériques suivantes :

- Les méthodes de gradient : gradient standard, gradient évoluées par exemple méthode de Nesterov
- Les méthodes de type Newton : quasi Newton, le Newton généralisé tranqué, Newton inexact Newton - CG, les méthodes semi-lisses de Newton

Nous pouvons donc attaqué le problème dual (2.2.15) avec des outils classique ou avec des technique plus évoluées.

Nous mentionnons également que le choix des méthodes de résolution dépend du niveau de sophistication ou de la structure du problème.

Notre travaille est concentré sur la méthode du gradient pour considérer l'interprétation des projections alterné.

d'autre côté la solution du problème dual y^* peut être obtenue avec des outils classique car $x^* = x(y^*)$ ainsi la solution primal x^* de (2.2.5), donc pour résoudre le problème nous avons la méthode dual suivante :

Algorithme 1

Nous présentons l'algorithme qui calcule la projection sur l'intersection, cet algorithme maximise la fonction θ sur \mathbb{R}^m et ceci en utilisant des fonctions à gradient de lipschitz afin de générer une séquence maximisant $(y_k)_k$ ainsi que

- . $x_k = x(y_k)$.
- . $\nabla\theta(y_k) = -Ax_k + b$.
- . $\theta(y_k) = -\frac{1}{2} \|x\|^2 + y_k^T b$.

L'algorithme calcule la solution approchée y_k à l'itération k (si la solution dual y_k existe), donc x_k est une approximation de l'unique solution du projection (2.2.5) dans le sans ou

$$\text{Si } \|y - y_k\| \leq \epsilon, \text{ alors } \|x - x_k\| \leq \|A\| \epsilon$$

La résolution du problème (2.2.5) nécessite d'une hypothèse qui assure l'existence de la solution dual qui détermine l'intérêt de l'algorithme ci dessus, cette hypothèse est connue par la condition (primal) de Slater

supposons qu'il existe un point strictement réalisable à l'intersection, i.e

$$\exists \bar{x} \in \mathbb{R}^n : A\bar{x} = b, \bar{x} \in \text{int}\mathcal{K}$$

Notons que cette hypothèse assure le bon comportement de l'algorithme aussi un attrait géométrique naturel dans le contexte des méthodes de projection, alors que la forte condition de complémentarité ou la condition de Slater n'a pas un tel rôle

L'intersection régulière et la condition de Slater

Dans cette partie nous étudions le rôle de la condition de Slater dans la programmation conique (voir le manuel [57]. et la revue [36] pour plus d'information), ainsi que dans les méthodes projective, puis en présentant des résultats alternatifs de type Farkus et les relier avec des algorithmes de projection et de régularité, enfin en énonçant la signification géométrique de cette condition

L'intersection régulière

Nous avons d'abord besoin des notations suivantes

Étant donné un cône \mathcal{K} dans \mathbb{R}^n , le cône polaire de \mathcal{K} noté par \mathcal{K}° est l'ensemble des points de \mathbb{R}^n tels que la projection de ceux-ci sur \mathcal{K} soit 0, on terme mathématique

$$\mathcal{K}^\circ := \{y \in \mathbb{R}^n : y^T x \leq 0, x \in \mathcal{K}\}.$$

Soit $x \in \mathcal{K}$, le cône normal de \mathcal{K} en x est noté par $N_{\mathcal{K}}(x)$ et le cône tangent de \mathcal{K} en x est $T_{\mathcal{K}}(x)$. ainsi

$$N_{\mathcal{K}}(x) = (T_{\mathcal{K}}(x))^\circ.$$

On dit que \mathcal{K} et \mathcal{A} peuvent être séparés si

$$\exists v \neq 0 \in \mathbb{R}^n, \alpha \in \mathbb{R} \forall k \in \mathcal{K}, a \in \mathcal{A} : v^T k \leq \alpha \leq v^T a.$$

est strictement séparés si l'inégalité ci-dessus est strict.

Nous soulignons que la condition de Slater garantit le bon comportement de l'algorithme dual, elle réalise pour le problème de réalisabilité que l'on considère un sens géométrique intrinsèque.

Proposition 2.2.1 (*l'hypothèse de slater*)

Les propriétés suivantes sont équivalentes

- 1) $\text{int}\mathcal{K} \cap \mathcal{A} \neq \emptyset$.
- 2) $T_{\mathcal{K}}(x) + \text{Ker}\mathcal{A} = \mathbb{R}^n, \forall x \in \mathcal{K} \cap \mathcal{A}$.
- 3) $N_{\mathcal{K}}(x) \cap \text{span}\mathcal{A}^T = \{0\} \forall x \in \mathcal{K} \cap \mathcal{A}$.
- 4) $0 \in \text{int}(\mathcal{K} - \mathcal{A})$.

\mathcal{K} et \mathcal{A} ne peut pas être séparé

Preuve 2.2.1 voir [37][9], et [51] section 16.4.2 (*projection methods for conic problems*)

Un point de vue géométrique, les propriétés ci-dessus signifient que le problème de réalisabilité est bien posé, d'autre part elle joue un rôle important dans les problème de réalisabilité convexe.

Lewis et al [2] ils relient la régularité métrique et la distance au mal-posé à l'intersection qui disent qu'elle est linéairement régulière.

Dans notre position la condition de slater est équivalent à la régularité linéaire de l'intersection $\mathcal{K} \cap \mathcal{A}$, vue que \mathcal{K} et \mathcal{A} ne peuvent pas être séparé par un hyperplan, cette régularité s'avère avoir des effets algorithmiques intéressantes pour les méthode de projection de base, donc c'est un outil pour avoir une convergence linéaire entre les projection alternées et moyennées, comme nous l'expliquerons plus tard.

Projection Alternative

Pour confirmer que l'hypothèse de slater est naturelle dans notre situation nous créons un liens entre l'algorithme de projection duale et le schéma de projection alterné, plusieurs travaux ont été réalisé pour cette méthode en citant les travaux de Helmke, Ait Rami et Moove [37] qui présentent un algorithme de projection alterné pour certains problèmes réalisabilité conique.

La méthode de projection alternée

Cette méthode consiste à projeter un point initial sur le premier ensemble disons \mathcal{A} , ensuite elle projette le nouveau point sur le second ensemble \mathcal{K} , puis reprojeter le nouveau point sur le premier c'est à -dire continue à projeter alternativement, afin de trouver un point dans l'intersection de deux ensemble.

Donc la méthode de projection alternée est un schéma algorithmique intuitif .

Autrement dit , il consiste à répéter l'opération

$$P_{\mathcal{K}}(P_{\mathcal{A}}(X)) \rightarrow X. \quad (2.2.16)$$

Cette algorithme converge linéairement vers un point de $\mathcal{K} \cap \mathcal{A}$ si les ensembles précédent est une intersection régulière , et avec cette condition nous estimons la vitesse de convergence.

Pour le cas général voir [2] , et pour le cas des ensembles convexe voir [19]

Dans notre cas , nous avons une projection explicite car l'un des sous - ensemble est affine , alors la projection de x sur \mathcal{A} est la suivante

$$P_{\mathcal{K}}(x) = x - A^T[AA^T]^{-1}(Ax - b). \quad (2.2.17)$$

Comme $\text{rang}(\mathcal{A}) = m$ (par hypothèse),dés qu'on connaîtra projeter sur le cône \mathcal{K} , on appliquera le schéma de projection alternée ,c'est particulièrement le cas pour le cône du second ordre , l'orthante positive et le cône SDP.

Or la convergence de cette méthode est linéaire avec un taux proche de 1 , c'est -à- dire qu'elle est lente , avec la correction de Dykstra [16] qui rajoute avant le projeté sur le cône \mathcal{K} on affirme la convergence vers la projection de c sur l'intersection , afin de compense la courbure de \mathcal{K} .

Algorithme 2 : (La projection alternée corrigée)

posons $\delta_0 = 0$, $a_0 = c$, $x_0 = P_{\mathcal{K}}(c)$, alors
pour $k = 1 , 2 , \dots$
 $\bar{a}_k = a_{k-1} - \delta_{k-1}$ (Dykstra's correction)
 $x_k = P_{\mathcal{K}}(\bar{a}_k)$
 $\delta_k = x_k - \bar{a}_k$
 $a_k = P_{\mathcal{K}}(x_k)$
Fin

L'intersection duale de la méthode de projection alternée est qu'il s'agit d'un algorithme d'optimisation du gradient plus évoluée avec une taille de pas constante pour maximiser la fonction dual .

Nous pouvons dire alors que l'algorithme 1 est une généralisation de la méthode de projection alterné.

La proposition suivante généralise le résultat de [25] sur le problème le plus proche de la matrice de corrélation .

Proposition 2.2.2 ([37])

La suite $(x_k)_k$ générée par l'algorithme 2 est la même que la suite $(x(y_k))_k$ de

l'algorithme 1 appliquée à la suite $(y_k)_k$ produite par le schéma

$$y_0 = 0, y_{k+1} = y_k + [AA^T]^{-1} \nabla \theta(y_k). \quad (2.2.18)$$

Preuve 2.2.2 (pour plus de détail voir [37])

il suffit de prouver par récurrence que $\bar{a}_k = c + A^T y_k$ est vrai pour $k = 0$
supposons qu'elle vrai pour k est on démontre $k + 1$ alors

$$\begin{aligned} \bar{a}_{k+1} &= a_k - \delta_k \quad (\text{Dijkstra's correction}) \\ &= a_k - x_k + \bar{a}_k \quad (\text{definition de } \delta_k) \\ &= -A^T [AA^T]^{-1} - (Ax_k - b) + \bar{a}_k \quad (\text{utilisant la definition de } \bar{a}_k \text{ et (2.2.17)}) \\ &= -A^T [AA^T]^{-1} - (Ax(y_k) - b) + c + A^T y_k \quad (\text{hypothese de reccurence}) \\ &= c + A^T (y_k - [AA^T]^{-1} (Ax(y_k) - b)) \end{aligned}$$

d'ou

$$x_{k+1} = P_{\mathcal{K}}(\delta_{k+1}) = P_{\mathcal{K}}(c + A^T y_{k+1}) = x(y_{k+1}).$$

cela est complété la preuve

Il faut maintenant savoir si l'intersection $\mathcal{K} \cap \mathcal{A}$ est vide pendant l'exécution de l'algorithme 1, le problème de réalisabilité conique suivant

$$A^T Y \in \mathcal{K}^\circ, b^T y > 0. \quad (2.2.19)$$

assure l'irréalisabilité du problème (2.2.1). Si le problème (2.2.19) possède une solution réalisable, alors le problème (2.2.1) n'a pas de solution réalisable comme nous écrivons dans la proposition suivante

Proposition 2.2.3 (proposition 3.4 dans [37])

- (i) Si $\mathcal{K} \cap \mathcal{A} \neq \emptyset$ alors il n'existe aucune solution réalisable de (2.2.19).
- (ii) Supposons que $c = 0$, s'il existe y tel que $x(y) = 0$, alors $\mathcal{K} \cap \mathcal{A} = \emptyset$.

Preuve 2.2.3 pour (i) : soit $x \in \mathcal{K} \cap \mathcal{A}$ et soit $y \in \mathbb{R}^n$ tel que $Ay \in \mathcal{K}^\circ$, on savait que le fait que x et Ay soient dans \mathcal{K}° implique

$$b^T y = (Ax)^T y = x^T A^T y \leq 0.$$

(ii) : supposons que $c = 0$ observe qu'on a

$$A^T y \in \mathcal{K} \Leftrightarrow P_{\mathcal{K}}(A^T y) = 0.$$

$$\Leftrightarrow x(y) = 0.$$

par suite (ii) est l'énoncé inverse de (i)

Cette proposition a deux résultats importants

-le premier résultat : un vecteur y du problème (2.2.19) affirme l'infaisabilité du problème (2.2.1).

-le deuxième résultat : puisque la modélisation du problème (2.2.1) ne donne aucun indice sur le choix de c on prend $c = 0$ car la vérification de ce choix pousse l'exécution de l'algorithme 1 à un coût numérique nul.

par conséquent si $c = 0$ on a besoin de calculer une projection supplémentaire sur le cône pour vérifier si $P_{\mathcal{K}}(Ay) = 0$ (nous évaluons la fonction dual (2.2.19) par le calcul de la norme $\|x(y)\|^2$).

La proposition suivante montre que si \mathcal{K} et \mathcal{A} ne se coupent pas alors il existe toujours un vecteur y satisfaisant (2.2.19)

Proposition 2.2.4 Si \mathcal{K} et \mathcal{A} peuvent être strictement séparés alors il existe une solution réalisable à (2.2.19).

Preuve 2.2.4 La démonstration se fait en deux étapes :

D'abord montrons que $u \in \mathcal{K}^\circ$

Soit $\alpha \in \mathbb{R}$ et $u \in \mathbb{R}^n$, comme $0 \in \mathcal{K}$ alors $\alpha \geq 0$, utilisant la définition de la séparation de \mathcal{K} et \mathcal{A}

$$\text{on a } \forall k \in \mathcal{K}, a \in \mathcal{A} : u^T k \leq \alpha \leq u^T a$$

pour $x \in \mathcal{K}$ et $t > 0$, on a $tx \in \mathcal{K}$ donc $tu^T x \leq \alpha$, on fait tendre t vers $+\infty$ on obtient $u^T x \leq 0$ ce qui prouve que $u \in \mathcal{K}^\circ$.

Il reste à prouver que $u \in \text{span } \mathcal{A}^T$

décomposez u de sorte que $u = Ay + z$ avec $Ay \in \text{span } \mathcal{A}^T$ et $z \in \text{Ker } A$ pour tout $a \in \mathcal{A}$ on a

$$\alpha \leq u^T a = y^T b + z^T a$$

ainsi pour tout $t > 0$ et $a_0 \in \mathcal{A}$ fixé, utilisant l'inégalité ci-dessus on obtient

$$\alpha - y^T b - z^T a_0 < -t \|z\|^2.$$

donc $z = 0$ et $u = Ay \in \text{span } \mathcal{A}^T$

par conséquent $u = Ay$ réside dans \mathcal{K}° et $0 \leq \alpha$

Remarque 2.2.1 *Dans cette partie, nous avons introduit un algorithme pour résoudre des problèmes de faisabilité coniques qui tire parti des caractéristiques géométriques ainsi que des propriétés intéressantes lorsque l'hypothèse de Slater est vérifiée. Nous nous intéressons aux algorithmes de projection suivant le schéma de l'algorithme 1 qui sont encore en cours de développement. Cette approche par projection pourrait être considérée comme une alternative à l'approche des points intérieurs afin d'obtenir des solutions algorithmiques assez rapides.*

2.2.2 Les méthodes prédicteurs - correcteurs

Cette approche est une partie essentielle des méthodes primales-duales, elle a d'abord été proposée par Kheirfam [3] qui a présenté la méthode du point intérieur prédicteur-correcteur pour un problème de complémentarité horizontale en utilisant la technique de Darvay, puis il a introduit un algorithme correcteur-prédicteur basé sur la mesure de proximité suivant un chemin intérieur pour SDO.

dans l'autre sens, Kheirfam [4] a modifié les directions NT pour construire un algorithme prédicteur-correcteur pour l'optimisation du cône symétrique quadratique convexe (CQSCO) et a extrait la limite de complexité. ces méthodes se concentrent sur deux objectifs principaux pour atteindre leurs buts, elles ont utilisé deux types de pas, de ce fait les itérations de l'approche prédicteur correcteur alternent entre ces deux types :

- Les pas prédicteurs : réduire la mesure de dualité (pour réduire μ).
- Les pas correcteurs : rendre l'itération près du voisinage du chemin centrale (pour se centrer).

Avant de représenter ces deux itération nous allons examiner le choix particulier du paramètre σ

- cas $\sigma = 0$

pour cette condition nous avons seulement réduit la valeur de μ , c'est-à-dire réduit la mesure de dualité pour atteindre l'optimum (sans s'inquiéter de la centralité des itérés).

L'inconvénient de ce choix est d'arriver plus vite aux bords de la zone admissible, ce qui réduit la taille des pas de Newton pouvant être effectuer .

- cas $\sigma = 1$

dans cette situation notre but est de centrer les itérés pour se rapprocher de la trajectoire centrale, c'est à dire ne pas avoir besoin de progresser vers la solution optimal.

En d'autres termes, nous essayons de calculer la direction du pas correcteur afin qu'il n'y ait aucune chance de converger vers l'optimum. Il faut cependant noter que le choix du paramètre σ correspond aux itérations proposées par l'approche prédicteur-correcteur donc il faut être prudent dans le choix de la valeur de σ .

Etude détaillée de la méthode

Considérons le problème primale (P) et son duale (D) avec le domaine des solutions réalisables du primale et du duale $F_P * F_D$, et son intérieur $F_P^\circ * F_D^\circ$

où

$$F^\circ(P) := \{x \in \mathcal{J} : Ax = b, x \in \text{int}\mathcal{K}\}.$$

et

$$F^\circ(D) := \{(y, s) \in \mathbb{R}^m \times \mathcal{J} : A^T y + s = c, s \in \text{int}\mathcal{K}\}.$$

supposons que l'ensemble de solutions réalisables $F^\circ(P) * F^\circ(D)$ est non vide, cette hypothèse est notée par la condition de (IPC) (condition de point intérieur).

Il est clair que la recherche de la solution optimale du (P) et (D) est équivalent à la résolution du système suivant

$$\begin{cases} Ax = b, & x \in \mathcal{K} \\ A^T y + s = c, & s \in \mathcal{K} \\ x \circ s = 0 \end{cases} \quad (2.2.20)$$

le principe général des méthodes des points intérieurs primaux- duaux (IPMS) est de remplacer la troisième équation $x \circ s = 0$ appelée condition de complémentarité dans (2.2.20) par l'équation paramétrée $x \circ s = \mu e$ avec $\mu > 0$ Par conséquent le système devient

$$\begin{cases} Ax = b, & x \in \mathcal{K} \\ A^T y + s = c, & s \in \mathcal{K} \\ x \circ s = \mu e \end{cases} \quad (2.2.21)$$

le fait que les conditions de (IPC) soient vérifiées, alors pour tout $\mu > 0$ le système (2.2.21) admet une unique solution $(x(\mu), y(\mu), s(\mu))$ (voir [28]) qui est appelé le μ -centre, l'ensemble de ces μ -centres forme la trajectoire centrale pour (P) et (D), qui se rapproche de la solution optimale de la paire du problème donné lorsque μ tend vers zéro.

Lemme 2.2.1 (lemme 28 Dans [55]) Soient x, s et w des éléments dans l'algèbre euclidienne de Jordan \mathcal{J} , tels que $x, s \in \text{int}\mathcal{K}$ et w inversible, alors

$$x \circ s = \mu e \quad \text{ssi} \quad P(w^{-\frac{1}{2}})x \circ P(w^{\frac{1}{2}})s = \mu e.$$

La façon la plus courante d'imposer cette approche est de supposer que l'itération actuelle (x, y, s) est une solution réalisable primale-duale, alors remplaçons l'équation paramétrée $x \circ s = \mu e$ par $P(w^{-\frac{1}{2}})x \circ P(w^{\frac{1}{2}})s = \mu e$ et appliquons la méthode de Newton, on obtient le système de Newton comme suit

$$\begin{cases} A \Delta x = 0 \\ A^T \Delta y + \Delta s = 0 \\ P(w^{-\frac{1}{2}}) \Delta x \circ P(w^{\frac{1}{2}})s + P(w^{-\frac{1}{2}})x \circ P(w^{\frac{1}{2}}) \Delta s = \mu e - P(w^{-\frac{1}{2}})x \circ P(w^{\frac{1}{2}})s \end{cases} \quad (2.2.22)$$

On choisit

$$w = P(x^{\frac{1}{2}})(P(x^{\frac{1}{2}})s)^{-\frac{1}{2}} = P(s^{-\frac{1}{2}})(P(s^{\frac{1}{2}})x)^{\frac{1}{2}}$$

donc nous avons

$$P(w^{-\frac{1}{2}})x = P(w^{\frac{1}{2}})s$$

comme la direction Δs est symétrique, le problème est imposé sur la direction Δx qui ne peut pas être symétrique, de nombreuses méthodes de symétrisation ont été proposées, nous utilisons dans cette partie la symétrisation NT qui donne la direction suivante [72, 73]

$$v = \frac{P(w^{-\frac{1}{2}})x}{\sqrt{\mu}} := \frac{P(w^{\frac{1}{2}})s}{\sqrt{\mu}} := \left(\frac{P(w^{-\frac{1}{2}})x \circ P(w^{\frac{1}{2}})s}{\mu} \right)^{\frac{1}{2}}. \quad (2.2.23)$$

et

$$dx := \frac{P(w^{-\frac{1}{2}}) \Delta x}{\sqrt{\mu}}, \quad ds := \frac{P(w^{\frac{1}{2}}) \Delta s}{\sqrt{\mu}}. \quad (2.2.24)$$

Alors le système (2.2.22) devient

$$\begin{cases} \bar{A}dx = 0 \\ \bar{A}^T \frac{\Delta y}{\mu} + ds = 0 \\ d_x + d_s = v^{-1} - v \end{cases} \quad (2.2.25)$$

où

$$\bar{A} := \sqrt{\mu}AP(w^{\frac{1}{2}})$$

Définissons maintenant la mesure de proximité $\sigma(x, s, \mu)$ comme suit

$$\sigma(x, s, \mu) := \sigma(v) := \|e - v^2\|_F$$

Alors

$$\sigma(v) = 0 \Leftrightarrow v^2 = e \Leftrightarrow P(w^{-\frac{1}{2}})x \circ P(w^{\frac{1}{2}})s = \mu e \Leftrightarrow x \circ s = \mu e$$

notez que nous pouvons utiliser la forme approximative $\sigma(v)$ pour mesurer la distance entre l'itération actuelle (x, y, s) et le μ -centre $(x(\mu), y(\mu), s(\mu))$ aussi une simple borne après une étape prédictive-corrective donc cela nous aide d'analyser notre algorithme.

Algorithme prédictive-corrective

Nous devons maintenant introduire un voisinage de la trajectoire centrale et proposons un algorithme prédicteur-correcteur de point intérieur pour le problème pour le cône du second ordre .
définir le voisinage du chemin central tel que $0 < \tau < 1$ soit un paramètre positif

$$N(\tau, \mu) := \{(x, s) : Ax = b, x \in \text{int}K, A^T y + s = c; s \in \text{int}K, y \in \mathbb{R}^m, \sigma(x, s, \mu) \leq \tau\}$$

soit (x, y, s) un point qui appartient au voisinage $N(\tau, \mu)$, notre but est de générer un nouveau point (x^p, y^p, s^p) en utilisant la nouvelle direction NT. pour calculer la direction de correction $(dx, \Delta y, ds)$ à l'itération courante (x, s, y) avec $N\mu > \epsilon$, un des moyens utilisés est de résoudre le système linéaire suivant

$$\begin{cases} \bar{A}d_x = 0 \\ \bar{A}^T \frac{\Delta y}{\mu} + d_s = 0 \\ d_x + d_s = v^{-1} - v \end{cases} \quad (2.2.26)$$

tels que :

$$v = \frac{P(w^{-\frac{1}{2}})x}{\sqrt{\mu}} := \frac{P(w^{\frac{1}{2}})s}{\sqrt{\mu}}$$

et

$$\bar{A} := \sqrt{\mu}AP(w^{\frac{1}{2}})$$

on pose aussi

$$\Delta x := \sqrt{\mu}P(w^{\frac{1}{2}})d_x$$

et

$$\Delta y := \sqrt{\mu}P(w^{-\frac{1}{2}})d_s$$

.

Notons que la nouvelle itération $x_+ := x + \Delta x$, $y_+ := y + \Delta y$, $s_+ := s + \Delta s$ est obtenue en faisant un pas de Newton complet le long de la direction corrective.

de plus on a

$$v = \frac{P(w_+^{-\frac{1}{2}})x_+}{\sqrt{\mu}} := \frac{P(w_+^{\frac{1}{2}})s_+}{\sqrt{\mu}}, \quad \text{et,} \quad \bar{A}_+ := \sqrt{\mu}AP(w_+^{\frac{1}{2}})$$

Par conséquent, la direction prédictive $(d_x^p, \Delta^p y, d_s^p)$ peut être obtenue par la solution du système suivant

$$\begin{cases} \bar{A}_+ d_x^p = 0 \\ \bar{A}_+^T \frac{\Delta y}{\mu} + d_s^p = 0 \\ d_x^p + d_s^p = -v_+ \end{cases} \quad (2.2.27)$$

avec

$$\Delta^p x := \sqrt{\mu}P(w^{\frac{1}{2}})d_x, \quad \text{et,} \quad \Delta^p y := \sqrt{\mu}P(w^{-\frac{1}{2}})d_s$$

Donc on a le point prédicteur

$$\begin{aligned} x^p &:= x_+ + \theta \Delta^p x \\ y^p &:= y_+ + \Delta^p y \\ s^p &:= s_+ + \Delta^p s \\ \mu^p &:= (1 - \theta)\mu \end{aligned}$$

Il s'ensuit que le pas du prédicteur réduit la valeur de μ d'un facteur $(1 - \theta)$ qui doit être choisi de manière convenable. L'algorithme proposé se répète jusqu'à ce que nous ayons une valeur inférieure que le paramètre de précision.

L'algorithme générique correcteur-prédicteur de point intérieur pour SOCO est donné par

L'algorithme prédicteur-correcteur pour l'optimisation du cône de second ordre (SOCO)

Début d'algorithme

Données :

Un paramètre de seuil $0 < \tau < 1$ (par défaut $\tau = \frac{1}{8}$).

un paramètre de précision $\epsilon > 0$;

Un paramètre de mise à jour fixe $\theta \in (0, 1)$;

Initialisation : soit $x := x^0$; $y := y^0$ $s := s^0$; telle que $\sigma(x^0, s^0, \mu^0) \leq \tau$;

début

$x := x^0$; $y := y^0$; $s := s^0$; $\mu := \mu^0$;

tant que $N\mu \geq \epsilon$ **faire**

début

l'étape correcteur

résoudre le système (2.2.26) puis

$x_+ := x + \Delta x$;

$y_+ := y + \Delta y$;

$s_+ := s + \Delta s$;

l'étape prédicteur

résoudre le système (2.2.27) puis

$x^p := x_+ + \theta \Delta^p x$;

$y^p := y_+ + \theta \Delta^p y$;

$s^p := s_+ + \theta \Delta^p s$;

mise à jour

$\mu^p := (1 - \theta)\mu$;

$x := x^p$; $y := y^p$; $s := s^p$; $\mu := \mu^p$;

Fin tant que.

Fin.

Figure 3 : *Algorithm 3.*

Chapitre 3

La programmation cônica du second ordre et méthodes de points intérieurs

Notre étude est s'intéresse aux problèmes de minimisation d'une composition d'une fonction objectif linéaire soumise à l'intersection d'un ensemble affine avec le produit cartésien d'un nombre fini de cônes du second ordre (de Lorentz ou de glace) Puis , on développe un algorithme de chemin central de type primal-dual pour résoudre un problème SOCO et on termine par l'étude de la convergence de cet algorithme et l'analyse de la complexité.

3.1 Position du problème

Problème primal

Soient deux entiers naturel m et n avec $m \leq n$ on partitionne les vecteurs $x, c \in \mathcal{K}$ et la matrice $A \in \mathcal{K}$ en conséquence comme $x = (x^1; x^2; \dots; x^N)$ $s = (s^1; s^2; \dots; s^N)$, $c = (c^1; c^2; \dots; c^N)$ et $A = (A^1, A^2, \dots, A^N)$, où $x^j, s^j, c^j \in \mathbb{R}^{n_j}$ $A^j \in \mathbb{R}^{m \times n_j}$ et $b \in \mathbb{R}^m$.

Un programme du cône du second ordre sous forme primal est un problème d'optimisation donné par :

$$(P) \begin{cases} \min_x c^T x \\ Ax = b \\ x \in \mathcal{K} \end{cases}$$

définissons maintenant l'ensemble des solutions réalisables (respectivement strictement réalisables) de (P)

$$F_p = \{x \in \mathcal{K} : Ax = b, x \succeq_{\mathcal{K}} 0\}.$$

$$F_p^0 = \{x \in \mathcal{K} : Ax = b, x \succ_{\mathcal{K}} 0\}$$

La valeur optimale primal du problème (P) est donnée par :

$$val(P) = \inf_x \{c^T x : Ax = b, x \in \mathcal{K}\}.$$

On dit que x^* est une solution optimale primal de (P) si :

$$x^* \in F_p \quad \text{et} \quad val(P) = c^T x^*.$$

Problème dual

Le problème dual (D) de (P) est considéré en fonction de la lagrangienne L de (P) comme suit :

$$L(x, y) = \langle c, x \rangle + \langle b - Ax, y \rangle.$$

ainsi la fonction dual associée est donnée par :

$$\{\max_{y \in \mathbb{R}^m} \min_{x \in \mathcal{K}} L(x, y) = \max_{y \in \mathbb{R}^m} H(y)\}.$$

où

$$\begin{aligned} H(y) &= \min_{x \in \mathcal{K}} L(x, y), \quad y \in \mathbb{R}^m \\ &= \min_{x \in \mathcal{K}} \{\langle c, x \rangle + \langle b - Ax, y \rangle\}, \quad y \in \mathbb{R}^m \\ &= \min_{x \in \mathcal{K}} \{\langle c, x \rangle + \langle -Ax, y \rangle + \langle b, y \rangle\}, \quad y \in \mathbb{R}^m \\ &= \min_{x \in \mathcal{K}} \{\langle c - A^T y, x \rangle + \langle b, y \rangle\}, \quad y \in \mathbb{R}^m \end{aligned}$$

Finalement, on obtient :

$$(H) \begin{cases} \max_y \langle b, y \rangle & \text{si } c - A^T y \succeq 0 \\ -\infty & \text{ailleurs} \end{cases}$$

Donc le problème dual de (P) s'écrit sous la forme suivante

$$(D) \begin{cases} \max_y b^T y \\ A^T y + s = c \\ s \in \mathcal{K} \quad \text{et} \quad y \in \mathbb{R}^m \end{cases}$$

On note par la suite

$$F_D = \{(y, s) \in \mathbb{R}^m \times \mathcal{K} : A^T y + s = c, s \succeq_{\mathcal{K}} 0\}.$$

$$F_D^0 = \{(y, s) \in \mathbb{R}^m \times \mathcal{K} : A^T y + s = c, s \succ_{\mathcal{K}} 0\}.$$

l'ensemble des solutions réalisables (respectivement strictement réalisables) pour (D)

On définit la valeur optimale dual du problème (D) comme suit :

$$val(D) = \sup_{(y,s)} \{b^T y : A^T y + s = c; s \in \mathcal{K} \text{ et } y \in \mathbb{R}^m\}.$$

On dit que $(y^*; s^*)$ est une solution optimale dual de (D) si :

$$(y^*, s^*) \in F_D \quad \text{et} \quad val(D) = b^T y^*.$$

3.2 Dualité en Soco

La notion de dualité recouvre différents objets et concepts en mathématique tel que l'optimisation, elle permet d'associer à un problème primal qui est parfois difficile à résoudre pour différents raisons un autre problème liés et qui permet de le résoudre plus facilement, cette notion n'est pas valable seulement pour les fonction convexes même s'il convient à la résolution de problèmes convexe mais aussi pour d'autres types de problèmes.

3.2.1 Dualité faible

Théorème 3.2.1 Soit y une solution admissible pour le problème conique dual et x une solution admissible pour le problème conique primal, alors on a toujours $c^T x \geq b^T y$.

Preuve 3.2.1 1)-on a

$$\begin{aligned} c^T x - b^T y &= c^T x - (Ax)^T y \\ &= x^T c - x^T A^T y \\ &= x^T (c - A^T y) \geq 0 \end{aligned}$$

et le fait que cette dernière quantité est positive découle de la définition du cône dual K^* et des inclusions $x \in K$, $c - A^T y \in K^*$

2)-Nous avons

$$\begin{aligned} c^T x &\geq b^T y && \forall x \in F_P \text{ et } \forall y \in F_D \\ \Rightarrow c^T x^* &\geq b^T y && \forall (y, s) \in F_D \\ \Rightarrow p^* = c^T x^* &\geq b^T y^* = d^* \\ \Rightarrow p^* &\geq d^* \end{aligned}$$

On a aussi

$$p^* = -\infty \Rightarrow d^* = -\infty \text{ et } d^* = +\infty \Rightarrow p^* = +\infty$$

3.2.2 Dualité forte

Définition 3.2.1 Une solution admissible x du problème conique primal est dite strictement admissible si elle appartient à l'intérieur du cône primal K c'est-à-dire si $Ax = b$ et $x \succ_K 0$.

Une solution admissible y du problème conique dual est dite strictement admissible si elle satisfait strictement l'inégalité impliquant le cône dual K^* c'est-à-dire si $A^T y \prec_{K^*} c$.

Définition 3.2.2 Un problème conique (primal ou dual) est dit strictement admissible si il admet au moins une solution strictement admissible.

avant de donner la notion de dualité forte on a les deux théorèmes suivants

Théorème 3.2.2 Si le problème dual est strictement admissible, on a soit

- **un problème dual non borné :**

dans ce cas $d^* = +\infty = p^*$ et le primal n'est pas admissible.

- **un problème dual borné :**

dans ce cas le problème primal est forcément borné, soluble et satisfait $p^* = d^*$ (il existe donc au moins une solution primale admissible x^* telle que $c^T x^* = p^* = d^*$).

Théorème 3.2.3 Si le problème primal est strictement admissible, on a soit

- **un problème primal non borné :**
dans ce cas $p^* = -\infty = d^*$ et le dual n'est pas admissible .

- **un problème primal borné :**
dans ce cas le problème dual est forcément borné, soluble et satisfait $p^* = d^*$ (il existe donc au moins une solution dual admissible y^* telle que $b^T y^* = d^* = p^*$).

Enfin , on peut en déduire aisément le corollaire suivant :

Corollaire 3.2.1 Si le problème primal et le problème dual sont tous deux strictement admissibles, les deux problèmes sont bornés, solubles et satisfont $p^* = d^*$ (il existe donc des solutions admissibles x^* et y^* telles que $b^T y^* = d^* = p^* = c^T x^*$).

Exemple 3.1 considérons le problème conique suivant :

$$\mathcal{K} = \mathcal{L}^2, A = \begin{pmatrix} -1 & 0 & 1 \\ 0 & -1 & 0 \end{pmatrix}, b = \begin{pmatrix} 0 & -1 \end{pmatrix}^T \text{ et } c = \begin{pmatrix} 0 & 0 & 0 \end{pmatrix}^T$$

Le problème primal peut s'écrire

$$\begin{cases} \min 0 \\ -x_1 - x_3 = 0 \\ -x_2 = -1 \\ (x_1, x_2, x_3) \in \mathcal{L}^2 \end{cases}$$

mais cette dernière contrainte ne peut être satisfaite, en utilisant, $x_2 = -1$ $x_3 = -x_1$ la contrainte conique $x_2^2 + x_3^2 \leq x_1^2$ devient $1 + x_1^2 \leq x_1^2$, une contradiction. Par conséquent, le problème primal est non admissible et $p^* = +\infty$, Le problème dual peut s'écrire

$$\begin{cases} \max y_2 \\ (x_1, x_2, x_3) \in \mathcal{L}^2 \end{cases}$$

Toutefois, l'inclusion dans \mathcal{L}^2 implique $y_1^2 + y_2^2 \leq y_2^2$ ce qui implique $y_1^2 = 0$ et par conséquent la fonction objectif du dual est identiquement égale à 0 pour toutes les solutions admissibles : $d^* = 0$. En conclusion, ce problème possède un dual admissible et borné et pourtant son primal est impossible.

3.2.3 Complémentarité en Soco

On peut exprimer la condition pour x^* et y^* d'être des solutions optimales de (Soco) et (DSoco), sous forme d'une condition dite de complémentarité.

Théorème 3.2.4 Soient $x^* \in F_p$ et $(y, s) \in F_D$ le saut de dualité

$$\langle x^*, s^* \rangle = \langle c, x^* \rangle - b^T y^*$$

Alors x^* et (y^*, s^*) sont des solutions optimales pour (Soco) et (DSoco) respectivement si et seulement si $x^* s^* = 0$.

Le problème de complémentarité est le suivant :

$$(Pc) \begin{cases} Ax = b, x \succ_{\mathcal{K}} 0 \\ A^T y + s = c, s \succ_{\mathcal{K}} 0 \\ xs = 0 \end{cases}$$

3.3 Méthodes de points intérieurs pour le problème Soco

Les méthodes de points intérieurs forme une classe d'algorithme efficaces et plus utilisable pour la résolution de type de problème de cône de second ordre , cette méthode permet de résoudre une série de problème approchés c-à-d génère une suite de points appartenant à l'intérieur relatif du domaine réalisable et converge vers la solution optimale . Cette méthode se décline en trois type :

- Les méthodes affines.
- Les méthodes de réduction du potentiel.
- Les méthodes de la trajectoire centrale.

1) Les méthodes affines

Les algorithmes affines sont des algorithmes de point intérieur, c'est-à-dire des algorithmes qui progressent tout en demeurant strictement à l'intérieur du domaine réalisable . Ces méthodes ont été introduites pour la première fois par Dikin en puis par Barnes, al. Kortanek et Shi . Ils ont ensuite été prouvé que les algorithmes convergent sous l'hypothèse de non dégénérescence . Récemment , Tsuchiya, Terlaky et Muramatsu , Tseng et Luo ont pu montre que la méthode converge même dans le cas de la dégénéré , Cependant , aucune borne de polynomialité n'a été démontrée jusqu'à présent ces méthodes ont été mises en oeuvre par de nombreux chercheurs, notamment Morton et al.

2) Les méthodes de réduction du potentiel

Le premier algorithme de ce type a été proposé par Gonzaga puis une extension de ceci a été proposé par Freund , le résultat est entièrement semblable à la forme standard de l'algorithme projectif (basé sur le fait, qu'on progresse tout en restant strictement à l'intérieur du domaine réalisable d'où la connotation " intérieur"). Indépendamment, Ye a dérivé un algorithme similaire qui utilise une fonction de potentiel primale-duale pour obtenir un algorithme de complexité par un facteur de \sqrt{n} . Dernièrement, Anstreicher et Bosch ont proposé un algorithme dont la complexité est meilleure que celle des algorithmes mentionnés précédemment.

3) Les méthodes de la trajectoire centrale

Ces méthodes attirent une grande attention par plusieurs chercheurs , la méthode de la trajectoire central a été développée la première fois par Barnes au début des années 90 . Ensuite en 1997 Monteiro présente une méthode de trajectoire central primal-dual et prouve la convergence polynomiale de l'algorithme à long et à petit pas puis les deux chercheurs Monteiro et al montre la convergence polynomiale de la méthode du trajectoire central primal-dual , en 1999 Ji et al ont introduit la convergence de type prédicteur correcteur de la méthode trajectoire central , les recherches sur cette méthode sont continuées jusqu'à maintenant , on trouve en 2002 Koulaï et al ont étendue l'algorithme prédicteur correcteur de Mohrotra basé sur la direction de Nesterov et Todd (NT) , ensuite l'année 2009 Wang et Bai ont proposé un algorithme de point intérieur primal-dual pour l'optimisation du cône de second ordre avec une direction de Nesterov – Todd complète.

L'idée principale de cette méthode est de suivre une trajectoire central c-à-d restreintes le plus max des itérés dans un voisinage du chemin central , afin d'obtenir une complexité polynomiale est une convergence super-linéaire).

3.3.1 La méthode de trajectoire central primal-dual pour la programmation du cône du second ordre basée sur les fonctions noyau

Considérons le problème primal de (SOCO) et son dual comme suit

$$(P) \quad : \min\{c^T x : Ax = b, x \geq 0_{\mathcal{K}}\}$$

$$(D) : \max\{b^T y : A^T y + s = c, s \geq 0_{\mathcal{K}}\}$$

où $b = (b_1, \dots, b_m)$, $y \in \mathbb{R}^m$, $c = (c_1, \dots, c_n)$, $x = (x_1, \dots, x_n)$ et $A = A_i$ ($\forall i = 1, \dots, m$) sont des matrices dans \mathcal{K} .

Tout au long de cette partie nous formulons les hypothèses suivantes :

- 1) La matrice A à un rang plein
- 2) les problèmes (P) et (D) satisfant la condition de point intérieur (IPC)

$$(i.e) \quad \exists x \in F_P^\circ, (y, s) \in F_D^\circ$$

Nous supposons que (P) et (D) satisfont la condition du point intérieur (IPC) c'est-à-dire qu'il existe (x^0, y^0, s^0) tel que

$$\begin{cases} Ax^0 = b, & x^0 \geq 0_{\mathcal{K}} \\ A^T y + s^0 = c, & s^0 \geq 0_{\mathcal{K}} \end{cases} \quad (3.3.1)$$

Il est bien connu que la recherche d'une solution optimale (X^*, Y^*, S^*) de (P) et (D) est équivalent à la recherche du système suivant

$$\begin{cases} Ax = b, & x > 0_{\mathcal{K}} \\ A^T y + s = c, & s > 0_{\mathcal{K}} \\ L(x)s = 0 \end{cases} \quad (3.3.2)$$

L'idée de base des méthode de point intérieur (IPMs) est de remplacer la troisième équation (la condition de complémentarité) dans le système (3.3.2) par l'équation paramétrisée $L(x)s = \mu e$ avec $\mu > 0$. On considère donc le système suivant

$$\begin{cases} Ax = b, & x > 0_{\mathcal{K}} \\ A^T y + s = c, & s > 0_{\mathcal{K}} \\ L(x)s = \mu e \end{cases} \quad (3.3.3)$$

Si les conditions de (IPC) sont satisfait par les problèmes (P) et (D) alors pour chaque $\mu > 0$ il existe une solution unique $(X(\mu), Y(\mu), S(\mu))$ pour le système (3.3.3) que l'on appelle μ -centre, l'ensemble des μ -centre (avec μ prenant plusieurs valeurs décroissantes qui convergent vers 0) construit la trajectoire centrale de (P) et (D) appelé chemin centra.

Notez qu'au μ -centre nous avons

$$x(\mu)^T s(\mu) = \frac{1}{2} Tr(x \circ s) = \frac{1}{2} Tr(L(x)s) = \frac{1}{2} Tr(\mu e) = N\mu \quad (3.3.4)$$

Si $\mu \rightarrow 0$ alors la limite du chemin central existe et puisque les points limites satisfont la condition de complémentarité ,la limite donne des solutions optimales de (P) et (D).

En fait pour simplifier les contributions théorique , en utilisant la technique d'intégration auto-dual nous pouvons supposer que $x^0 = s^0 = e$.

Une nouvelle direction de recherche pour SOCO

Les IPMs suivent le chemin central de manière appropriée et trouvent une solution approximative de (P) et (D) lorsque μ tend vers le zéro , généralement la direction de recherche (de descente) est obtenue à partir d'un certain système de type Newton ,le système de Newton pour(3.3.3) revient à résoudre le système suivant

$$\begin{cases} A \Delta x = b, & x > 0_{\mathcal{K}} \\ A^T \Delta y + \Delta s = c, & s > 0_{\mathcal{K}} \\ L(x) \Delta s + L(s) \Delta x = \mu e - L(x)s \end{cases} \quad (3.3.5)$$

Ce système a une solution unique si et seulement si la matrice $AL(s)^{-1}L(x)A^T$ est non singulière . Malheureusement cela peut ne pas être le cas , même si A à un rang complet cela est dû à la matrice $L(s)^{-1}L(x)$ peut ne pas être symétrique .

Différentes méthodes de symétrisation de la troisième equation (3.3.5) sont proposées afin que le nouveau système symétrisé ait une solution unique , dans ce chapitre nous basons notre attention sur le schéma de symétrisation NT [73].

Notons que W est l'unique automorphisme de \mathcal{K} qui satisfait $Wx = W^{-1}s$ pour tout $x, s > 0_{\mathcal{K}}$ et $j \in J$.

On note par u_j le vecteur de NT et W_{NT}^j la matrice de NT telle que

$$u_j = \left(\frac{(sj)^T Q^j sj}{(xj)^T Q^j xj} \right)^{\frac{1}{4}}, \quad w_j = \frac{u_j^{-1} s^j + u_j Q^j x^j}{\sqrt{2} \sqrt{(xj)^T s^j + \sqrt{(xj)^T Q^j x^j (sj)^T Q^j sj}}}$$

$$W_{NT}^j = \begin{pmatrix} w_1^j & (w_{2:nj}^j)^T \\ w_{2:nj}^j & E_{nj-1} \frac{1}{1+w_1^j} w_{2:nj}^j (w_{2:nj}^j)^T \end{pmatrix}$$

Où $Q^j := \text{diag}(1; -1, \dots, -1) \in \mathbb{R}^{n_j \times n_j}$, $E_{n_j} := \text{diag}(1, 1, \dots, 1) \in \mathbb{R}^{n_j \times n_j}$

pour les choix ci-dessus, on a ([voir [15]])

$$\bar{x}^j := u_j W_{NT}^j x^j = u_j^{-1} (W_{NT}^j)^{-1} s^j = \bar{s}^j, j \in J$$

Lemme 3.3.1 (proposition 6.3.3 dans [26])

pour tout $j \in J$ on a

- 1) $\text{Tr}(x^j \circ s^j) = \text{Tr}(\bar{x}^j \circ \bar{s}^j)$
- 2) $u_j^2 \det(x^j)$, $u_j^{-2} \det(\bar{s}^j)$
- 3) $x^j \geq 0_K (x^j > 0_K) \Leftrightarrow \bar{x}^j \geq 0_K (\bar{x}^j > 0_K)$

nous définissons

$$v := \frac{Wx}{\sqrt{\mu}} = \left(\frac{W^{-1}s}{\sqrt{\mu}} \right).$$

et

$$\bar{A} := \frac{1}{\sqrt{\mu}} A W^{-1}, d_x := \frac{W \Delta x}{\sqrt{\mu}}, d_s := \frac{W^{-1} \Delta s}{\sqrt{\mu}}.$$

dans l'espace réduit de v , le système de Newton devient

$$\begin{cases} \bar{A} d_x = b \\ \bar{A}^T \Delta y + d_s = c \\ d_x + d_s = v^{-1} - v \end{cases} \quad (3.3.6)$$

puisque la matrice $(A^{-1})^T A^{-1}$ est définie positive, ce système a une solution unique.

Nous décrivons l'approche pour obtenir la direction de recherche classique définie uniquement pour Soco, l'approche consiste à remplacer le côté droite de la troisième équation dans (3.3.6) par $\psi'(t)$ où $\psi(t)$ est la fonction noyau.

Nous utiliserons alors le système suivant pour définir notre nouvelle direction de recherche

$$\begin{cases} \bar{A} d_x = b \\ \bar{A}^T \Delta y + d_s = c \\ d_x + d_s = -\psi'(v) \end{cases} \quad (3.3.7)$$

En revenant dans l'espace x et s respectivement , on obtient la direction de recherche dans l'espace d'origine avec

$$\Delta x = \sqrt{\mu}W^{-1}d_x , \Delta s = \sqrt{\mu}Wd_s$$

comme x et s sont orthogonaux , d_x et d_s sont également orthogonaux , on a donc :

$$d_x = d_s = 0 \Leftrightarrow \psi'(v) = 0 \Leftrightarrow \psi(v) = 0 \Leftrightarrow \Psi(v) = 0 \Leftrightarrow v = e$$

c-à-d si et seulement si et seulement si $L(x)s = \mu e$, c'est à dire ,si et seulement si $x = x(\mu)$ et $s = s(\mu)$, sinon $\Psi(v) > 0$ donc si $(x, y, s) \neq (x(\mu), y(\mu), s(\mu))$, $(\Delta x, \Delta y, \Delta s)$ alors (x, y, s) est non nul.

Par conséquent , on peut utiliser $\Psi(v)$ comme fonction de proximité pour mesurer la distance entre l'itération actuelle et le μ -centre correspondant.

L'algorithme qui proposé est déroule comme suit : en prenant un pas par défaut approprié de taille $\alpha \in (0, 1)$ ces directions de recherche construisent un nouveau triplet $(x + \alpha \Delta x, y + \alpha \Delta y, s + \alpha \Delta s)$ cette procédure est répétée jusqu'à ce que nous trouvions une itération dans certain voisinage de $(x(\mu), y(\mu), s(\mu))$ ensuite nous diminuons μ à $\mu = (1 - \theta)\mu$ pour $\theta \in (0, 1)$ et on appliquons la méthode de Newton ciblant les nouveaux μ -centre , ce processus est répété jusqu'à ce que μ est suffisamment petit et à ce stade nous avons trouvé une solution aux problèmes (P) et (D).

Les paramètres τ, θ et le pas de déplacement α doivent être choisi de telle sorte que l'algorithme est optimisé dans le sens que le nombre d'itérations produites par l'algorithme sera minimal que possible. Avant de présenter notre algorithme, nous définissons une fonction barrière en fonction des variables d'origine x, s et le paramètre barrière en fonction de

$$\Psi(x; s; \mu) := \Psi(v) = \sum_{j=1}^N Tr(\psi(v^j))$$

Algorithme de point intérieur primal-dual pour l'optimisation du cône de

Début d'algorithme

Données :

Un paramètre de seuil $\tau > 1$

un paramètre de précision $\epsilon > 0$;

Un paramètre de mise à jour fixe $\theta \in (0, 1)$;

Initialisation : soit $x := x^0$; $s := s^0$; $\mu := \mu_0$ une solution initial vérifiant (IPC);

début

$x := x^0$; $y := y^0$; $\mu := \mu_0$;

tant que $N\mu \geq \epsilon$ **faire**

début

$\mu := (1 - \theta)\mu$;

tant que $\Psi(x, s, \mu) > \tau$ **faire**

Résoudre le système (3.3.7) pour déterminer $(\Delta x; \Delta y; \Delta s)$;

mise à jour

$x := x + \alpha \Delta x$;

$y := y + \alpha \Delta y$;

$s := s + \alpha \Delta s$;

Fin tant que .

Fin tant que.

Fin.

Figure 4 : *Algorithm 4.*

Fonction noyau et ses propriétés

Définition 3.3.1 *On dit que $\psi :]0, +\infty[\longrightarrow [0, +\infty[$ est une fonction noyau si elle est différentiable et vérifié les condition suivantes :*

- 1) $\psi(1) = \psi'(1) = 0$;
- 2) $\psi''(t) > 0$, pour $t > 0$;
- 3) $\lim_{t \rightarrow 0} \psi(t) = \lim_{t \rightarrow 0} \psi'(t) = +\infty$

Les deux premières conditions signifient que $\psi(t)$ est strictement convexe et si $t = 1$ alors elle admet une valeur minimal avec $\psi(1) = 0$. A partir de sa dérivée second $\psi(t)$ s'écrit de la façon suivante

$$\psi(t) = \int_1^t \int_1^y \psi''(x) dx dy.$$

Tandis que la troisième condition indique que $\psi(t)$ est une fonction barrière .

Définissons la nouvelle fonction noyau $\psi(t)$ par

$$\psi(t) = t - 2 + e^{\left(\frac{1}{t}-1\right)}. \quad (3.3.8)$$

Les trois premières dérivées de $\psi(t)$ par rapport à t sont exprimées de la manière suivante :

$$\begin{aligned} \psi'(t) &= 1 - \frac{1}{t^2}e^{\left(\frac{1}{t}-1\right)}, \\ \psi''(t) &= \left(\frac{2}{t^3} + \frac{1}{t^4}\right)e^{\left(\frac{1}{t}-1\right)}, \\ \psi'''(t) &= -\left(\frac{1}{t^6} + \frac{6}{t^5} + \frac{6}{t^4}\right)e^{\left(\frac{1}{t}-1\right)} \end{aligned}$$

Il est clair que

$$\begin{cases} \psi(1) = \psi'(1) = 0 \\ \lim_{t \rightarrow 0} \psi(t) = \lim_{t \rightarrow 0} \psi'(t) = +\infty \end{cases}$$

d'après(3.3.8) on a $\psi''(t) > 0 \ t > 0$ ainsi $\psi''(t)$ est strictement décroissante pour tout $t \in (0, \infty)$, par conséquent $\psi(t)$ est strictement convexe .

Qualification(éligibilité) d'une fonction noyau

Le lemme suivant présente les conditions d'éligibilité d'une fonction noyau $\psi(t)$.

Lemme 3.3.2 Soit $\psi(t)$ définie dans (3.3.8) avec $t > 0$, on dit que $\psi(t)$ est qualifiée si elle vérifie les propriétés suivantes :

- (i) $t\psi''(t) + \psi'(t) > 0 \ t > 1$.
- (ii) $t\psi''(t) - \psi'(t) > 0 \ t < 1$.
- (iii) $\psi'''(t) < 0 \ t > 0$.
- (iv) $2[\psi''(t)]^2 - \psi'(t)\psi'''(t) > 0 \ t < 1$.
- (v) $\psi''(t)\psi'(\beta t) - \beta\psi'(t)\psi''(\beta t) > 0 \ t > 1, \beta > 1$.

Preuve 3.3.1 Pour(i) utilisant $\psi'(t)$ et $\psi''(t)$ on trouve :

$$t\psi''(t) + \psi'(t) = \left(\frac{1}{t^2} + \frac{1}{t^3}\right)e^{\left(\frac{1}{t}-1\right)} + 1 > 0 \quad \forall t > 0$$

. Pour (ii) on a :

$$t\psi''(t) - \psi'(t) = \left(\frac{3}{t^2} + \frac{1}{t^4}\right)e^{\left(\frac{1}{t}-1\right)} - 1 > 0 \quad \forall t < 1$$

. (iii) est vérifiée à partir de $\psi'''(t)$:

$$\psi'''(t) = -\left(\frac{1}{t^6} + \frac{6}{t^5} + \frac{6}{t^4}\right)e^{\left(\frac{1}{t}-1\right)} < 0$$

La condition (iv) est satisfait à partir de :

$$2\psi''(t)^2 - \psi'(t)\psi'''(t) = \left(\frac{1}{t^6} + \frac{6}{t^5} + \frac{6}{t^4}\right)e^{2\left(\frac{1}{t}-1\right)} + \left(\frac{2}{t^6} + \frac{2}{t^7} + \frac{1}{t^8}\right)e^{2\left(\frac{1}{t}-1\right)} > 0$$

Finalement pour (v) il est facile à vérifier que

$$\begin{aligned} \psi''(t)\psi'(\beta t) - \beta\psi'(t)\psi''(\beta t) &= 2\left(e^{\left(\frac{1}{t}-1\right)} - \frac{e^{\left(\frac{1}{\beta t}-1\right)}}{\beta^2}\right)\frac{1}{t^3} + \left(e^{\left(\frac{1}{t}-1\right)} - \frac{e^{\left(\frac{1}{\beta t}-1\right)}}{\beta^3}\right)\frac{1}{t^4} \\ &+ \left(\frac{1}{\beta} - 1\right)\frac{e^{\left(\frac{1}{t}-1\right)}e^{\left(\frac{1}{\beta t}-1\right)}}{\beta^2 t^6} > 2\left(1 - \frac{1}{\beta^2}\right)\frac{e^{\left(\frac{1}{t}-1\right)}e^{\left(\frac{1}{\beta t}-1\right)}}{t^6} > 0 \end{aligned}$$

par suite $\psi(t)$ est une fonction noyau qualifiée ce qui achève la preuve

Maintenant, nous présentons un lemme important qui sera utilisé plus tard

Lemme 3.3.3 On a

1)

$$\psi(t) \leq \frac{1}{2}\psi''(1)(t-1)^2 \quad \text{pour tout } t \geq 1$$

2)

$$\psi(t) \leq t$$

3)

$$\frac{1}{2}(t-1)^2 \leq \psi(t) \leq \frac{1}{2}\psi'(t)^2$$

Preuve 3.3.2 -1) utilisons le développement de Taylor avec $\psi(1) = \psi'(1) = 0$ et

$\psi''(1) = \frac{3}{2}$ et le fait que $\psi'''(t) < 0$ on trouve :

$$\begin{aligned}\psi(t) &= \psi(1) + \psi'(1)(t-1) + \frac{1}{2}\psi''(1)(t-1)^2 + \frac{1}{3!}\psi'''(\theta)(\Theta)^3 \\ &= \frac{1}{2}\psi''(1)(t-1)^2 + \frac{1}{3!}\psi'''(\theta)(\Theta)^3 \\ &< \frac{1}{2}\psi''(1)(t-1)^2 = \frac{3}{2}(t-1)^2\end{aligned}$$

-Pour (2) la preuve est immédiate

-En utilisant (3.3.8) nous obtenons :

$$\psi(t) = \int_1^t \int_1^x \psi''(y) dy dx \geq \int_1^t \int_1^x 1 dy dx = \frac{1}{2}(t-1)^2.$$

La seconde inégalité est obtenue de la manière suivante

$$\begin{aligned}\psi(t) &= \int_1^t \int_1^x \psi''(y) dy dx \leq \int_1^t \int_1^x \psi''(x)\psi''(y) dy dx = \int_1^t \psi''(x)\psi'(x) dx \\ &= \int_1^t \psi'(x) d\psi'(x) = \frac{1}{2}\psi'(t)^2\end{aligned}$$

Le lemme suivant montre que toute fonction noyau éligible $\psi(t)$ est exponentiellement convexe .

Lemme 3.3.4 (lemme 2.1.2 [26]) les propriétés suivantes sont équivalentes

- $\psi(\sqrt{t_1 t_2}) \leq \frac{1}{2}(\psi(t_1) + \psi(t_2)) \quad \forall t_1, t_2 > 0.$
- $t\psi''(t) + \psi'(t) > 0, \quad t > 0.$
- $\psi(e^\epsilon)$ est convexe.

Lemme 3.3.5 soit $\rho \in [0, +\infty[\rightarrow [1, +\infty[$ la fonction inverse de $\psi(t)$ pour tout $t \geq 1$. Alors on a

$$s + 1 \leq \rho(s) \leq 1 + s + \sqrt{s^2 + 2s}, s \geq 0.$$

Preuve 3.3.3 Soit $\psi(t) = s$ pour tout $t \geq 1$ alors $s = \psi(t) = t - 1 + \psi_b(t) \leq t - 1, \forall t \geq 1$

donc

$$t = \rho(s) \geq s + 1$$

Pour montrer la deuxième inégalité , Il est simplement nécessaire de démontrer que

$$s = \psi(t) \geq \frac{(t-1)^2}{2t} \quad t \geq 1 \quad \text{ainsi à prouver que } 2t\psi(t) \geq (t-1)^2.$$

Soit $g(t) = 2t\psi(t) - (t-1)^2$, $t \geq 1$ nous avons $g(1) = 2\psi(1) = 0$ et $g'(t) = 2\psi(t) + 2(1 - \frac{e^t}{t}) \geq 0$ pour tout $t \geq 1$.

comme $s = \psi(t)$ alors $2ts - (t-1)^2 \geq 0 \Rightarrow t^2 - 2(1+s)t + 1 \leq 0$ par conséquent

$$1 + s - \sqrt{s^2 + 2s} \leq \rho(s) \leq 1 + s + \sqrt{s^2 + 2s}, s \geq 0$$

d'où

$$s + 1 \leq \rho(s) \leq 1 + s + \sqrt{s^2 + 2s}, s \geq 0$$

Une borne supérieure de $\Psi(v)$ après chaque itération externe

Il convient de remarquer qu'au début de chaque itération externe de l'algorithme , juste avant la mise à jour du paramètre μ , on observe que $\Psi(v) \leq \tau$. Après avoir réduit le paramètre μ en le multipliant par un facteur $(1-\theta)$: $\mu_+ = (1-\theta)\mu$, avec $0 < \theta < 1$, le vecteur v est mise à jour selon la relation suivante : $v_+ = \frac{v}{\sqrt{1-\theta}}$ Cela entraîne généralement une augmentation de la valeur de $\Psi(v)$ au cours des itération externes, ensuite , pendant les itérations internes suivantes , $\Psi(v)$ diminue jusqu'à ce qu'elle atteigne la première valeur inférieure ou égale à τ . Pendant l'exécution de l'algorithme , les valeurs maximales de $\Psi(v)$ se produisent juste après les mises à jour de μ .

Lemme 3.3.6 (lemme 6.2.9 [26]) si x, s et $v \in \mathcal{K}_+$ vérifier : $d(v \diamond v) = \det(x)\det(s)$ et $Tr(v \diamond v) = Tr(x \diamond s)$ alors :

$$\Psi(v) \leq \frac{1}{2}(\Psi(x) + \Psi(s)).$$

Utilisons maintenant la proximité basé sur la norme suivante $\delta(v)$ pour v

$$\delta(v) = \frac{1}{\sqrt{2}} \sqrt{\sum_{i=1}^N \|\psi'(v^j)\|^2}. \quad (3.3.9)$$

Notons que $\delta(v) \geq 0$ et $\delta(v) = 0$ si et seulement si $\Psi(v) = 0$.

Théorème 3.3.1 (théorème 3.11 Dans [2]) On a

$$\delta(v) \geq \frac{1}{2} \psi'(\rho(\Psi(v)))$$

Nous devons présenter la décroissance de $\Psi(v)$ pour cela nous définissons une borne inférieure sur $\delta(v)$ en terme de $\Psi(v)$.

Le lemme suivant est obtenu en utilisant la reformulation du théorème 4.8 Dans [63] dont la preuve dépend de la condition (iii) du lemme 3.3.2)

Lemme 3.3.7 Pour tout vecteur positive v de \mathbb{R}^{\times} on a

$$\sqrt{\sum_{i=1}^P \psi'(v_i)^2} \geq \psi'(\rho(\sum_{i=1}^P (\psi(v_i))))$$

utilisons (3.3.9) et (2.1.12) on a

$$\delta(v) := \frac{1}{2} \sqrt{\sum_{i=1}^N [\psi'(\lambda_{max}(v^i))^2 + \psi'(\lambda_{min}(v^i))^2]}$$

et par (2.1.13) on obtient

$$\Psi(v) := \sum_{i=1}^N [\psi(\lambda_{max}(v^i)) + \psi(\lambda_{min}(v^i))]$$

Notez que $\delta(v)$ et $\Psi(v)$ ne dépendent que des valeurs propres $\lambda_{max}(v^i)$ et $\lambda_{min}(v^i)$ du vecteur v^i , $i = 1 \dots N$ ce qui amène a :

$$\delta(v) \geq \frac{1}{2} \psi'(\rho(\Psi(v))).$$

Illustrons l'augmentation de la fonction $\Psi(v)$ lors de la mise à jour
Le lemme suivant est représenté en [63] et cela dépend de la condition(v)

Lemme 3.3.8 Soit $v \in \mathbb{R}$ et $\beta \geq 1$ alors

$$\Psi(\beta v) \leq 2N \psi(\beta \rho(\frac{\Psi(v^i)}{2N})).$$

Corollaire 3.3.1 Soit $0 \leq \theta \leq 1$, $\frac{v}{\sqrt{1-\theta}}$ si $\Psi(v) \leq \tau$ alors

$$\Psi(v_+) \leq L_\psi(N, \theta, \tau) \leq 2N \psi(\frac{\rho(\frac{\tau}{2N})}{\sqrt{1-\theta}}) = L.$$

L'analyse de la décroissance de la fonction barrière et la complexité

Dans cette partie on va calculer le pas de déplacement α qui réduit la valeur de proximité et assure la stricte faisabilité de nouveau itéré durant les itérations internes .

Pour μ fixé, nous commençons avec une paire primale-dual (x,s) puis nous obtenons les nouvelles itérations avec α est le pas de déplacement

$$x^+ = x + \alpha \Delta x, y^+ = y + \alpha \Delta y, s^+ = s + \alpha \Delta s.$$

donc α est réalisable si et seulement si $x + \alpha \Delta x \in \mathcal{K}_+$ ($x^j + \alpha \Delta x^j \in \mathcal{K}_+$) et $s + \alpha \Delta s \in \mathcal{K}_+$ ($s^j + \alpha \Delta s^j \in \mathcal{K}_+$).

Soit W^j l'automorphisme de \mathcal{K}^j qui satisfait $W^j x^j = (W^{j-1})s^j$ et $v^j = \frac{W^j x^j}{\sqrt{\mu}}$ donc nous avons :

$$W^j((x^j + \alpha \Delta x^j) = \sqrt{\mu}(v^j + \alpha d_x^j).$$

$$W^{j-1}((s^j + \alpha \Delta s^j) = \sqrt{\mu}(v^j + \alpha d_s^j).$$

comme W^j est un automorphisme alors le pas de déplacement α est réalisable si et seulement si $v^j + \alpha d_x^j \in \mathcal{K}_+$ et $v^j + \alpha d_s^j \in \mathcal{K}_+$.

utilisons le lemme 2.8 Dans [63] avec $\bar{x} = \sqrt{\mu}(v^j + \alpha d_x^j)$ et $\bar{s} = \sqrt{\mu}(v^j + \alpha d_s^j)$ on obtient

$$\mu^2 \det(v^j + \alpha d_x^j) \det(v^j + \alpha d_s^j) = \det(x^j + \alpha \Delta x^j) \det(s^j + \alpha \Delta s^j).$$

$$\mu \text{Tr}((v^j + \alpha d_x^j) \circ (v^j + \alpha d_s^j)) = \text{Tr}((x^j + \alpha \Delta x^j) \circ (s^j + \alpha \Delta s^j)).$$

$$\mu^2 \det((v_+^j)^2) = \det(x^j + \alpha \Delta x^j) \det(s^j + \alpha \Delta s^j).$$

$$\mu \text{Tr}((v_+^j)^2) = \text{Tr}((x^j + \alpha \Delta x^j) \circ (s^j + \alpha \Delta s^j)).$$

on déduit que ,pour chaque j

$$\det((v_+^j)^2) = \det(x^j + \alpha \Delta x^j) \det(s^j + \alpha \Delta s^j) \quad (3.3.10)$$

$$\text{Tr}((v_+^j)^2) = \text{Tr}((x^j + \alpha \Delta x^j) \circ (s^j + \alpha \Delta s^j)).$$

utilisant le lemme 3.2 Dans [63] , prendre la somme sur tout j tel que $1 \leq j \leq N$ on obtient

$$\Psi(v_+) \leq \frac{1}{2}(\Psi(v + \alpha d_x) + \Psi(v + \alpha d_s))$$

Pour tout α , on considère la fonction $f(\alpha)$ qui mesure la différence de proximité entre l'ancien itéré et le nouveau itéré pour tout μ fixé

$$f(\alpha) = \Psi(v_+) - \Psi(v).$$

nous avons aussi

$$f(\alpha) \leq f_1(\alpha)$$

avec

$$f_1(\alpha) = \frac{1}{2}(\Psi(v + \alpha d_x) + \Psi(v + \alpha d_s)) - \Psi(v).$$

Evidement

$$f(0) = f_1(0) = 0.$$

considérons maintenant les dérivées première et seconde de $f_1(\alpha)$ par rapport à α pour estimer la décroissance de la fonction de proximité dans le cadre d'un pas

$$f_1'(\alpha) = \frac{1}{2}Tr(\psi'(v + \alpha d_x) \diamond d_x + \psi'(v + \alpha d_s) \diamond d_s). \quad (3.3.11)$$

$$f_1''(\alpha) = \frac{1}{2}Tr((d_x \diamond d_x) \diamond \psi''(v + \alpha d_x) + (d_s \diamond d_s) \psi''(v + \alpha d_s)). \quad (3.3.12)$$

utilisant (3.3.11) et (3.3.9) et la troisième inégalité du système (3.3.7) on trouve

$$f_1'(0) = \frac{1}{2}Tr(\psi'(v) \diamond (d_x + d_s)) = -\frac{1}{2}Tr(\psi'(v) \diamond v'(0)) = -2\delta(v)^2.$$

Lemme 3.3.9 (lemme 3.3 Dans [63]) si $\psi''(t)$ est décroissante de manière monotone dans t alors on a

$$f_1''(\alpha) \leq 2\delta(v)^2 \psi''(\lambda_{\min}(v) - 2\alpha\delta(v)).$$

Preuve 3.3.4 Comme d_x et d_s sont orthogonale et $d_s + d_s = -\psi'(v)$ alors

$$\|d_s + d_s\|^2 = \|d_x\|^2 + \|d_s\|^2 = 2\delta(v)^2.$$

par conséquent

$$\|d_x\| \leq \sqrt{2}\delta(v) \quad \text{et} \quad \|d_s\| \leq \sqrt{2}\delta(v).$$

d'après le lemme (2.1.4) on trouve

$$\begin{aligned}\lambda_{\max}(v + \alpha d_x) &\geq \lambda_{\min}(v + \alpha d_x) \geq \lambda_{\min}(v) - 2\alpha\delta(v). \\ \lambda_{\max}(v + \alpha d_s) &\geq \lambda_{\min}(v + \alpha d_s) \geq \lambda_{\min}(v) - 2\alpha\delta(v).\end{aligned}$$

d'après (iii) et la condition 2 de la définition 3.3.1 et le fait que $\psi''(t)$ est strictement décroissante, on obtient

$$\begin{aligned}0 < \psi''(\lambda_{\max}(v + \alpha d_x)) &\leq \psi''(\lambda_{\min}(v) - 2\alpha\delta(v)). \\ 0 < \psi''(\lambda_{\max}(v + \alpha d_s)) &\leq \psi''(\lambda_{\min}(v) - 2\alpha\delta(v)).\end{aligned}$$

En appliquant le corollaire (2.1.2), et en utilisant le fait que $\text{Tr}(z \circ z) = 2z^T z = \|z\|^2$ alors on a

$$\begin{aligned}\text{Tr}((d_x \diamond d_s) \diamond \psi''(v + \alpha d_x)) &\leq \psi''(\lambda_{\min}(v) - 2\alpha\delta(v)). \\ \text{Tr}((d_x \diamond d_s)) &= 2\psi''(\lambda_{\min}(v) - 2\alpha\delta(v)) \|d_x\|^2.\end{aligned}$$

et

$$\begin{aligned}\text{Tr}((d_x \diamond d_s) \diamond \psi''(v + \alpha d_s)) &\leq \psi''(\lambda_{\min}(v) - 2\alpha\delta(v)). \\ \text{Tr}((d_x \diamond d_s)) &= 2\psi''(\lambda_{\min}(v) - 2\alpha\delta(v)) \|d_s\|^2.\end{aligned}$$

la substitution dans (3.3.12) donne que

$$\begin{aligned}f_1''(\alpha) &\leq \psi''(\lambda_{\min}(v) - 2\alpha\delta(v))(\|d_x\|^2 + \|d_s\|^2). \\ &= 2\delta(v)^2 \psi''(\lambda_{\min}(v) - 2\alpha\delta(v)).\end{aligned}$$

ce qui complète la preuve

Lemme 3.3.10 (lemme 4.2 Dans [68]) Si α satisfait l'inégalité

$$-\psi'(\lambda_{\min}(v) - 2\alpha\delta(v)) + \psi'(\lambda_{\min}(v)) \leq 2\delta(v). \quad (3.3.13)$$

alors $f_1'(0) \leq 0$

Lemme 3.3.11 (lemme 4.3 Dans [68]) La plus grande valeur de α vérifiant (3.3.13) est donné par

$$\bar{\alpha} := \frac{\rho(\delta(v)) - \rho(2\delta(v))}{2\delta(v)}. \quad (3.3.14)$$

Lemme 3.3.12 (lemme 4.4 Dans [68]) Soient ρ et $\bar{\alpha}$ définis dans (3.3.14) alors

$$\bar{\alpha} \geq \frac{1}{\psi''(\rho(2\delta(v)))}.$$

par la suite on note par

$$\tilde{\alpha} := \frac{1}{\psi''(\rho(2\delta(v)))}. \quad (3.3.15)$$

et on utilise α comme taille de pas par défaut d'après le lemme(3.3.12) nous avons $\tilde{\alpha} \leq \bar{\alpha}$

Le lemme suivant montre que la proximité Ψ est décroissante, pour plus de détail voir[24]

Lemme 3.3.13 (lemme 3.12 Dans [24]) On suppose que $h(t)$ est une fonction convexe et deux fois différentiable telle que

$$h(0) = 0, \quad h'(0) < 0$$

et atteint son minimum globale en $t^* > 0$ et $h''(t)$ est croissante sur l'intervalle $[0; t^*]$ alors on a

$$h(t) \leq \frac{th'(0)}{2} \quad \text{pour tout } t \in [0, t^*].$$

Par suite le lemme suivant présente une borne supérieure pour la valeur de décroissance de la proximité à une itération interne

Lemme 3.3.14 (lemme 4.5 Dans [68]) Si le pas de déplacement α vérifie $\alpha \leq \bar{\alpha}$ alors

$$f(\alpha) \leq -\alpha\delta^2.$$

En combinant le résultat du lemme(3.3.12) et (3.3.14) on obtient le théorème suivant

Théorème 3.3.2 avec α étant la taille de pas par défaut donné par (3.3.15) on a

$$f(\tilde{\alpha}) \leq -\frac{\delta^2(v)}{\psi''(\rho(2\delta(v)))}. \quad (3.3.16)$$

Corollaire 3.3.2 Le membre droite de l'expression (3.3.16) est strictement décroissante par rapport à δ

La borne d'itération

Nous devons compter le nombre d'itération internes nécessaires pour revenir à la situation où $\Psi(v) \leq \tau$

Nous notons par Ψ_0 la valeur de Ψ d'après la mise à jour de μ et par $\Psi_k : k = 1, \dots, K$ les valeurs dans la même itération externe avec K désigne le nombre total d'itérations internes produites durant une itération externe

Lemme 3.3.15 (proposition 1,2,3 Dans [25]) Supposons qu'une séquence $t_k > 0, k = 1, \dots, K$ satisfaisant l'inégalité suivante

$$t_{k+1} \leq t_k - \beta t_k^{1-\gamma}, \beta > 0, \gamma \in [0, 1[, k = 1, \dots, K - 1.$$

Alors

$$K \leq \left\lceil \frac{t_0^\gamma}{\beta\gamma} \right\rceil.$$

Lemme 3.3.16 Soit K le nombre d'itération internes entre deux itérations externes successive alors

$$K \leq \frac{l\gamma}{\beta\gamma}.$$

Le lemme suivant présente le nombre d'itération externes à partir duquel l'algorithme 4 converge

Lemme 3.3.17 Soit K le nombre d'itération externes nécessaire pour que l'algorithme 4 converge vers une solution optimale primale-duale alors on a

$$K \geq \frac{1}{\theta} \log\left(\frac{2N}{\epsilon}\right).$$

Le nombre total d'itérations produites par l'algorithme 2 est obtenue en multipliant le nombre d'itérations internes par le nombre d'itérations externes comme suit

$$\frac{\psi_0^\gamma}{\theta\gamma\beta} \log \frac{2N}{\epsilon}.$$

Le schéma pour analyser l'algorithme basé sur une nouvelle fonction noyau

Pour analyser l'algorithme générique basé sur une fonction noyau nous avons besoin de suivre le schéma donné par Bai et al [68]

Etape 1. Résoudre l'équation $-\frac{1}{2}\psi'(t) = s$ pour obtenir $\rho(s)$, la fonction inverse de $-\frac{1}{2}\psi'(t)$ pour $t \leq 1$. Si l'équation est difficile à résoudre, trouver une borne inférieure pour $\rho(s)$.

Etape 2. Calculer la décroissance de $\Psi(v)$ en terme de $\delta = \delta(v)$ pour la taille de pas par défaut $\tilde{\alpha}$ tel que :

$$f(\tilde{\alpha}) \leq -\frac{\delta^2}{\psi''(\rho(2\delta))}$$

Etape 3. Résoudre l'équation $\psi(t) = s$ pour obtenir $\rho(s)$, la fonction inverse de $\psi(t)$, $t \geq 1$. Si l'équation est difficile à résoudre, trouver une borne supérieure pour $\rho(s)$.

Etape 4. Dédire une borne inférieure pour $\delta(v)$ en terme de $\Psi(v)$ de la forme :

$$\delta(v) \geq \frac{1}{2}\psi'(\rho(\Psi(v))).$$

Etape 5 : Utiliser les resultats des étapes 3 et 4 pour obtenir les constantes positives β et γ , avec $\gamma \in (0, 1]$, et telle que :

$$f(\tilde{\alpha}) \leq -\beta\Psi(v)^{1-\gamma}.$$

Etape 6 : Calculer une borne supérieure Ψ_0 pour $\Psi(v)$ sachant que

$$\Psi_0 \leq L_\psi(N, \theta, \tau) = 2N\psi\left(\frac{\rho(\frac{\tau}{2N})}{\sqrt{1-\theta}}\right).$$

Etape 7. Dédire une borne supérieure pour le nombre d'itérations totales sous la forme :

$$\frac{\Psi_0^\gamma}{\theta\beta\gamma} \log \frac{N}{\epsilon}.$$

Etape 8 : Si $\tau = (n)$ et $\theta = \Theta(1)$ on obtient une borne d'itérations pour les algorithmes à grand pas. Si $\tau = O(1)$ et $\theta = \Theta(n)$ on obtient une borne d'itérations pour les algorithmes à petit pas.

Preuve 3.3.5 Etape 1 : soit $s = -\frac{1}{2}\psi'(t)$ pour $t \in]0, 1]$, par la définition de $\rho(s) = t$, $t \in]0, 1]$, et $s \geq 0$ on a :

$$\begin{aligned} 2s &= -\psi'(t) = 1 + \psi'_b(t) \\ \implies -2s &= \psi'(t) = 1 + \psi'_b(t) \end{aligned}$$

ensuite nous avons

$$-\psi'_b(t) = 1 + 2s = \psi'_b(\bar{\rho}(1 + 2s))$$

comme

$-\psi''_b(t) = -\frac{1+2t}{t^4}e^{\frac{1}{t}}$, est strictement croissante avec $t \in]0,1]$. A partir du définition de $t = \rho(s) \geq \bar{\rho}(1 + 2s)$ nous obtenons

$$e^{\frac{1}{t}} = st^2 < s \implies \bar{\rho}(s) = t \geq \frac{1}{\log(s)}.$$

Cela nous mène à

$$\rho(s) \geq \bar{\rho}(1 + 2s) \geq \frac{1}{\log(1 + 2s)}. \quad (3.3.17)$$

Etape 2 : La fonction $\psi''(t)$ est strictement décroissante, alors

$$f(\bar{\alpha}) \leq -\frac{\delta^2}{\psi''(\rho(2\delta))} \leq \frac{-\delta^2}{\left(\frac{1+2t}{t^4}\right)(1+4\delta)t^2} \leq \frac{1}{(1+2t)(1+4\delta)(1+\log(1+4\delta))^2}.$$

donc

$$f(\bar{\alpha}) \leq \frac{1}{3(1+4\delta)(1+\log(1+4\delta))^2}. \quad (3.3.18)$$

Etape 3 : D'après le Lemme 4.2 la fonction inverse de $\psi(t)$ pour $t \in [1, \infty)$ satisfait

$$\psi(t) + 1 \leq \rho(\psi(t)) \leq 1 + \psi(t) + \sqrt{\psi^2(t) + 2\psi(t)}. \quad (3.3.19)$$

Etape 4 : En utilisant $\delta(v) \geq \frac{1}{2}\psi'(\rho(\Psi(v)))$, nous avons

$$\delta(v) \geq \frac{1}{2}\psi'(\rho(\Psi(v))) \geq \frac{1}{2}\left(1 - \frac{e^{\frac{1}{\rho(\Psi)}}}{\rho(\Psi)^2}\right) \geq \frac{1}{2}\left(1 - \frac{1}{\rho(\Psi)^2}\right) \geq \frac{1}{2}\left(1 - \frac{1}{\rho(\Psi)}\right).$$

comme $\rho(s) \geq 1 + s$ alors

$$\delta(v) \geq \frac{1}{2}\left(1 - \frac{1}{1+\Psi}\right) \geq \frac{1}{4}. \quad (3.3.20)$$

Etape 5 : $\Psi_0 \geq \Psi \geq \tau \geq 1$ de l'expression (3.3.20), on déduit que $1 \leq 4\delta$. et de (3.3.18), nous avons

$$f(\bar{\alpha}) \leq -\frac{\delta^2}{15\delta(1+\log(1+4\delta))^2} \leq -\frac{\delta}{15(1+\log(1+4\delta))^2} \leq -\frac{1}{60(1+\log(1+4\delta))^2}.$$

par conséquent

$$f(\bar{\alpha}) \leq -\frac{1}{60(1 + \log(2))^2}. \quad (3.3.21)$$

Il s'ensuit donc que

$$\Psi_{k+1} \leq \Psi_k - \beta(\Psi_k)^{1-\gamma}, \quad k = 1, \dots, K-1$$

où

$$\beta = \frac{1}{60(1 + \log(2))^2}, \quad \gamma = 1$$

Etape 6 : D'après le lemme 4.2, on a $\rho(\frac{\tau}{2N}) \leq 1 + \frac{\tau}{2N} + \sqrt{\frac{\tau^2}{4N} + \frac{\tau}{N}}$ Par conséquent

$$\Psi_0 \leq L_\psi(2N, \theta, \tau) = 2N\psi\left(\frac{\rho(\frac{\tau}{2N})}{\sqrt{1-\theta}}\right) \leq 2N\psi\left(\frac{1 + \frac{\tau}{2N} + \sqrt{\frac{\tau^2}{4N} + \frac{\tau}{N}}}{\sqrt{1-\theta}}\right)$$

Since $\psi(t) \leq \frac{3}{2}(t-1)^2$ for $t > 1$,

$$\Psi_0 \leq 3N\left(\frac{\rho(\frac{\tau}{2N})}{\sqrt{1-\theta}} - 1\right)^2 = \frac{3}{N(1-\theta)}\left(N + \frac{\tau}{2} + \sqrt{\frac{\tau^2}{4} + N\tau - N\theta}\right)^2. \quad (3.3.22)$$

Etape 7 : Par l'inégalité (3.3.21), le nombre total d'itérations internes est donné par :

$$K \leq \frac{\Psi_0^\gamma}{\beta\gamma} = (60(1 + \log(2))^2)\Psi_0.$$

En substituant (3.3.22) dans cette inégalité, on obtient

$$K \leq \frac{(180(1 + \log(2))^2)}{N(1-\theta)}\left(N + \frac{\tau}{2} + \sqrt{\frac{\tau^2}{4} + N\tau - N\theta}\right)^2 \frac{1}{\theta} \log \frac{N}{\epsilon}. \quad (3.3.23)$$

Etape 8 : Pour les algorithmes à grand pas, on considère $\tau = O(n)$ et $\theta = \Theta(1)$. Par conséquent

$$\Psi_0 \leq \frac{3}{N(1-\theta)}\left(N + \frac{\tau}{2} + \sqrt{\frac{\tau^2}{4} + N\tau - N\theta}\right)^2.$$

ainsi la borne d'itération devient

$$O\left(N \log\left(\frac{N}{\epsilon}\right)\right).$$

Conclusion

Dans ce travail , on s'est intéressé à la résolution d'un problème de programmation cônica du second ordre par la méthode des points intérieurs de suivi de chemin central (primales-duales) . Nous nous concentrons sur la nouvelle technique de Nesterov et Todd (schéma de mise à l'échelle NT) afin de trouver les directions de recherche , ce choix peut être justifié en soulignant que jusqu'à présent , les méthodes primales-duales par itérations intérieures (IPM) à grands pas , basées sur la direction de recherche (NT) , a donné les meilleurs résultats. Notre objectif dans cette étude est de présenter une nouvelle fonction noyau , nous avons établi que la complexité de cet algorithme à grands pas est d'ordre $O(N \log(\frac{N}{\epsilon}))$.

Perspectives

Notre prochain travail pourraient se concentrer sur deux aspects. La première consiste à concevoir une méthode du point intérieur (IPM) irréalisable basée sur la nouvelle fonction du noyau. Le deuxième aspect concerne la réalisation des tests numériques, qui constituent un travail intéressant pour étudier le comportement des algorithmes afin de les comparer à d'autres existants.

Bibliographie

- [1] A. Shapiro. *On duality theory of conic linear problems*. Springer, Boston, MA, volume 57, 135–165, (2001).
- [2] A. Lewis , D. Luke , J. Malick. *Local linear convergence for alternating and averaged nonconvex projections*. *Foundations of Computational Mathematics*, 9 :485-513, (2009).
- [3] B. Kheirfam. *A predictor-corrector interior-point algorithm for $P_*(k)$ horizontal linear complementarity problem*. *Numer. Algorithms*, 66(2), 349-361,(2014).
- [4] B. Kheirfam. *A corrector-predictor path-following method for convex quadratic symmetric cone optimization*, *J. Optim. Theory Appl.* 164(1), 246-260,(2015).
- [5] B. Kheirfam. *A corrector-predictor path-following method for second-order cone optimization*.*International Journal of Computer Mathematics*,(2015).
- [6] B. Alzalg. *The Algebraic Structure of the Arbitrary-Order Cone*.*J Optim Theory Appl*, Springer Science+Business Media New York ,(2016).
- [7] B. Alzalg. *Primal interior-point decomposition algorithms for two-stage stochastic extended second-order cone programming*.*Optimization*,(2018).
- [8] B. He, M. Xu , X. Yuan. *Solving large-scale least-squares covariance matrix problems by alternating direction methods*. *Technical report*, (2009).
- [9] B.N. Mordukhovich. *Variational Analysis and Generalized Differentiation* .Springer-Verlag, (2006).
- [10] B. Kyung Choi , G. Myung Lee. *On complexity analysis of the primal–dual interior-point method for semidefinite optimization problem based on a new proximity function*.*Nonlinear Analysis* 71(12), (2009).
- [11] B. Kyung Choi , G. Myung Lee. *On complexity analysis of the primal–dual interior point method for second-order cone optimization problem*.*J. KSIAM Vol.14, No.2, 93–111, (2010)*.

- [12] B. Kyung Choi , G. Myung Lee. *New complexity analysis for primal-dual interior-point methods for self-scaled optimization problems.* Choi and Lee *Fixed Point Theory and Applications* ,(2012).
- [13] C. El Boustani , J. Bleyera, M. Arquier, M. Ferradi, K. Saba. *Dual finite-element analysis using second-order cone programming for structures including contact.* *Engineering Structures*,(2019).
- [14] D. Henrion , J. Malick. *Projection methods for conic feasibility problems applications to polynomial sum-of-squares decompositions .* *Optimization Methods and Software*,26 :1, 23-46,(2011).
- [15] E. Andersen · C. Roos · T. Terlaky. *On implementing a primal-dual interior-point method for conic quadratic optimization.* *Math. Program, Ser* 249–277 ,(2003).
- [16] E. Djeffal , L. Djeffal . *A path following interior-point algorithm for semidefinite optimization problem based on new kernel function.* *Journal of Mathematical Modeling*,Vol. 4, No. 1, 35-58, (2016).
- [17] E. Djeffal , M. Laouar. *A primal-dual interior-point method based on a new kernel function for linear complementarity problem.* *Asian-European Journal of Mathematics* Vol. 13, No. 1, (2020).
- [18] F. Alizadeh . *An Introduction to Formally Real Jordan Algebras and Their Applications in Optimization.* *Handbook on Semidefinite, Conic and Polynomial Optimization* , 297–337,(2011).
- [19] F. Deutsch. *Best Approximation in Inner Product Spaces.* Springer, (2001).
- [20] G. Wang , Y. Bai. *A primal-dual interior-point algorithm for second-order cone optimization with full Nesterov–Todd step.* *Applied Mathematics and Computation* ,215, 1047–1061,(2009) .
- [21] H. Yamashita , H. Yabeb. *A primal-dual interior point method for non-linear optimization over second-order cones.* *Optimization Methods and Software* ,Vol. 24,407–426 , (2009).
- [22] H. Qi , D. Sun. *Quadratic convergence and numerical experiments of Newton’s method for computing the nearest correlation matrix.* *SIAM Journal on Matrix Analysis and Applications*,28 :360-385, (2006).
- [23] J. Hiriart-Urruty , C. Lemaréchal. *Convex Analysis and Minimization Algorithms*,Vols. 1 and 2, Springer-Verlag, (1993).
- [24] J. Sturm. *Using Sedumi 1.02, a Matlab toolbox for optimization over symmetric cones.* *Optimization Methods and Softwares*, 11-12 :625-653, (1999).

- [25] J. Malick. *A dual approach to semidefinite least-squares problems*. *SIAM Journal on Matrix Analysis and Applications*, 26, Number 1 :272-284, (2004).
- [26] J. Peng , C. Roos , T. Terlaky. *Self-regularity. A new paradigm for Primal-Dual Interior Point Algorithm*. Princeton University Press. Princeton, (2002).
- [27] J. Peng , C. Roos , T. Terlaky. *Self-regular functions and new search directions for linear and semi-definite optimization*. *Mathematical Programming*, 93 : 129-171, (2002).
- [28] J. Peng , C. Roos , T. Terlaky. *A new class of polynomial primal-dual interior-point methods for second-order cone optimization based on self-regular proximities*, *SIAM Journal on Optimization* 13 (1), 179–203, (2002) .
- [29] J. Dahl , D. Andersen. *A primal-dual interior-point algorithm for nonsymmetric exponential-cone optimization*. *Mathematical Programming*, (2021).
- [30] J. Sturm. *Implementation of interior point methods for mixed semidefinite and second order cone optimization problems*. *Optimization Methods and Software*, 1105-1154, (2010).
- [31] J. Sun , S. Zhang. *A modified alternating direction method for convex quadratically constrained quadratic semidefinite programs*. *European Journal of Operational Research*, 207(3) :1210 -1220, (2010).
- [32] L. Dong , J. Tang. *Interior-Point Method for Second-order Cone Programming Based on a Simple Kernel Function*. *Second International Conference on Computational Intelligence and Natural Computing*, (2010).
- [33] L. Zhang , Y. Bai , Y. Xu . *A full-step interior-point algorithm for second-order cone optimization based on a simple locally kernel function*. *Optimization Methods and Software*, 28 :3, 619-639, (2013).
- [34] L. Guerra , M. Achache. *A Parametric Kernel Function Generating the best Known Iteration Bound for Large-Update Methods for CQSDO*. *Statistics, Optimization, and Information Computing*, vol 8, 876–889, (2020).
- [35] L. Tunçel. *Generalization of primal-dual interior-point methods to convex optimization problems in conic form*. *Foundations of Computational Mathematics*, 1 :229-254, (2001).
- [36] L. Tunçel , H. Wolkowicz. *Strong duality and minimal representation for cone optimization*, *Research Report CORR , Department of Combinatorics and Optimization*, (2008).

- [37] M. Ait Rami , U. Helmke, J. Moore. *A finite step algorithm for solving convex feasibility problems*. *J. Global Optim* , 143–160,(2007) .
- [38] B. Kyung Choi , G. Myung Lee. *On complexity analysis of the primal-dual interior point method for second-order cone optimization problem*.*J. KSIAM Vol.14, No.2, 93–111, (2010)*.
- [39] M. Zangiabadi · G. Gu · C. Roos. *A Full Nesterov–Todd Step Infeasible Interior-Point Method for Second-Order Cone Optimization*.*Journal of Optimization Theory and Applications*,816–858 ,(2013).
- [40] M. Anjos , N. Higham , P. Takouda , H. Wolkowicz. *A semidefinite programming approach for the nearest correlation matrix problem*. *Technical report, Combinatorics and Optimization, University of Waterloo, N2L 3G1 , (2003)*.
- [41] M. Sarowar Morshed , C. Vogiatzis , M. Noor-E-Alam . *A primal-dual interior point method for a novel type-2 second order cone optimization problem* .*Control and optimization*,volume4,(2019).
- [42] M. Soma Lobo , L. Vandenberghe , S. Boyd , H. Le Bret . *Applications of second Order cone programming* .*Linear Algebra and its Applications*, 284, 193-3228,(1998).
- [43] M. Bouafia , D. Benterki , A. Yassine. *An efficient parameterized logarithmic kernel function for linear optimization* .*Springer-Verlag GmbH Germany* ,(2017).
- [44] M. Achache. *A new parameterized kernel function for LO yielding the best known iteration bound for a large-update interior point algorithm* .*African Mathematical Union and Springer-Verlag Berlin Heidelberg* ,(2015).
- [45] N. Higham. *Computing a nearest symmetric correlation matrix a problem from finance*. *IMA Journal of Numerical Analysis*,22(3) :329-343, (2002).
- [46] P. Xie , Jinwei Su , X. Peng. *Interior-point methods for second-order cone optimization based on the generalized trigonometric barrier function*.*The Second International Conference on Physics, Mathematics and Statistics*,(2019).
- [47] R. Tutuncu , K. Toh , M. Todd. *Inexact primal-dual path following algorithms for a special class of convex quadratic SDP and related problems*. *Pacific Journal on Optimization*, (2006).
- [48] R. Tutuncu , K. Toh , M. Todd. *Solving semidefinite quadratic-linear programs using SDPT3*. *Mathematical Programming B*, 95 :189-217, (2003).
- [49] R. Dykstra. *An algorithm for restricted least-square regression*.*Journal of the American Statistical Association*, 78 :837-842, (1983).

- [50] R. Borsdorf , N. Higham. *A preconditioned Newton algorithm for the nearest correlation matrix*, IMAJ. Numer.Anal,(2008).
- [51] R. Rockafellar. *Convex Analysis*. Princeton University Press, (1970).
- [52] S. Asadi , H. Mansouri , M. Zangiabadi. *A Primal-dual Interior-point Algorithm for Symmetric Cone Convex Quadratic Programming Based on the Commutative Class Directions*.Acta Mathematicae Applicatae Sinica, Vol. 35, 359–373 , (2019) .
- [53] S. Boutaghane , O. Ghelima . Imre Polik .*La méthode primale-duale du point intérieur basée sur plusieurs sélections du pas de déplacement pour la programmation du cône de second ordre (PCSO)* ,(2021).
- [54] S. Hafshejani , H. Mansouri , M. Reza Peyghami. *A large-update primal–dual interior-point algorithm for second-order cone optimization based on a new proximity function*.Optimization,A Journal of Mathematical Programming and Operations Research,1029-4945,(2016).
- [55] S. Schmieta , F. Alizadeh. *Extension of primal-dual interior point algorithms to symmetric cones*.Math. Program, Ser, A 96, 409–438, (2003).
- [56] S. Boyd , L. Xiao. *Least-squares covariance matrix adjustment*.SIAM Journal on Matrix Analysis and Applications, 27(2), (2005).
- [57] S. Boyd , L. Vandenberghe. *Convex Optimization*. Cambridge University Press, (2005).
- [58] T. Terlaky , Z. Wang. *On Identification of the Optimal Partition of Second Order Cone Optimization Problems*.SIAM J. OPTIM ,Vol. 24, No. 1, 385-414 ,(2014).
- [59] X. Cai , Linwu , Y. Yue , M. Li , G. Wang. *Kernel-function-based primal-dual interior-point methods for convex quadratic optimization over symmetric cone*.Journal of Inequalities and Applications , (2014).
- [60] X. Cai , G. Wang , M. El Ghami , Y. Yue. *Complexity Analysis of Primal-Dual Interior-Point Methods for Linear Optimization Based on a New Parametric Kernel Function with a Trigonometric Barrier Term*.Hindawi Publishing Corporation.Abstract and Applied Analysis,Volume 2014,(2014).
- [61] Y. Bai , P. Ma , J. Zhang. *A polynomial-time interior-point method for circular cone programming based on kernel functions*.Journal of industrial and management optimization,Volume 12,(2016).
- [62] Y. Hoon Lee , J. Jin , G . Cho. *Kernel function based interior–point algorithms for Semidefinite optimization*.Mathematical Inequalities and Applications,Volume 16, 1279–1294 , (2013).

- [63] Y. Bai , M. El Ghami , C. Roos. *A comparative study of kernel functions for primal-dual interior point algorithms in linear optimization*.SIAM Journal on Optimization,15 (1) :101-128,(2004).
- [64] Y. Bai , G. Wanga , C. Roos. *Primal-dual interior-point algorithms for second-order cone optimization based on kernel functions*,Nonlinear Analysis, 70 :3584-3602, (2009).
- [65] Y. Bai , G. Wang, C. Roos. *A new Primal-Dual Interior-Point Algorithm for Second-Order Cone Optimization*.(2004).
- [66] Y. Bai , G. Wang . *Primal-dual Interior point Algorithms for Second-order Cone Optimization Based on a New Parametric Kernel Function* .Acta Mathematica Sinica,2027–2042 ,(2007).
- [67] Y. Bai , G. Wang , C. Roos. *Primal–dual interior-point algorithms for second-order cone optimization based on kernel functions*.Nonlinear Analysis : Theory, Methods and Applications,Volume 70, Issue 10, 3584-3602, (2009).
- [68] Y. Bai , M. El Ghami , C. Roos. *A camparative study of new barrier functions for primal-dual interior-point algorithms in linear optimization*,(2003).
- [69] Y. Nesterov , M. Todd. *Primal-dual interior-point methods for self-scaled cones*.SIAM J. OPTIM,Vol. 8, 324-364,(1998).
- [70] Y-Ju Kuo. *Interior Point Methods for Second-Order Cone Programming and OR Applications*.Computational Optimization and Applications, 28, 255–285, (2004).
- [71] Y. Nesterov , A. Nemirovski. *Interior-Point Polynomial Algorithms in Convex Programming*. Number 13 in SIAM Studies in Applied Mathematics, (1994).
- [72] Y. Nesterov, M. Todd. *Self-scaled barriers and interior-point methods for convex programming*. Math. Oper. Res, 22(1) , 1-42,(1997).
- [73] Y. Nesterov, M. Todd. *Primal-dual interior-point methods for self-scaled cones*. SIAM J. Optim, 8(2), 324-364,(1998).
- [74] Z. Feng. *A New Iteration Large-Update Primal-Dual Interior-Point Method for Second-Order Cone Programming*.Numerical Functional Analysis and Optimization,33 :4, 397-414,(2012).