People's Democratic Republic of Algeria
Ministry of Higher Education and Scientific Research
Mohamed Kheider University, Biskra

Faculty of **EXACT SCIENCES**

Department of **MATHEMATICS**

Thesis *Presented by*

LAIDI SOUHEILA

In partial fulfillment of the requirements for the degree of

Master

Option **ANALYSIS**

STUDY OF INTERIOR-POINT METHOD FOR MONOTONE SEMIDEFINITE LINEAR COMPLEMENTARITY PROBLEMS

Examination Committee Members

BRAHIM REZKI	MAA	President
AMRANE HOUAS	MCA	Supervisor
NADJETTE BOUZIANE	MAA	Examiner

Academic Year

2024 / 2025

Dedication

I dedicate this humble work to my dear family, the source of love, support and patience, who have always encouraged and supported me to move forward.

A deep and eternal thought to my beloved brother, who left too soon, whose absence has deeply marked me, but whose memory has given me the strength to continue. May God grant him His mercy.

To my precious friends:**Aya**, **Najah**, **Oumaima** and **Imane**, for their sincere friendship, their reassuring presence and their constant encouragement.

And especially to my supervisor, Mr. Houas Amrane, for his valuable support, his wise advice, and his commitment throughout this work. I express my deepest gratitude and the most sincere thanks to him.

Acknowledgment

I extend my sincerest appreciation and gratitude to Dr. Houas Amrane, my supervisor, whose unwavering dedication and mentorship have been instrumental in shaping my professional growth and success. Their invaluable guidance, encouragement, and expertise have not only enriched my understanding but have also inspired me to strive for excellence in every endeavor.

Thank you, Dr. Houas Amrane, for your unwavering support and for being a beacon of inspiration in my professional journey.

I would like to thank all my teachers who have taught me throughout my academic journey. Additionally, I would also like to thank the jury members, Professor 1 and Dr. 2, for accepting to examine and evaluate this work. I also wish to thank all our teachers in the Department of Mathematics at Mohamed Khider University, who contributed to our education during the Bachelor's and Master's years.

To all those who have contributed to the realization of this memory,

Thank you

Notations and symbols

$\mathbb{R}^n$: The space of n-dimensional real vectors,

$\mathbb{R}^n_+$: The positive orthant of $\mathbb{R}^n$,

$\mathbb{R}^{m\times n}$: The vector space of real matrices of size $(m \times n)$,

$0_{m\times n}$: The zero matrix $(m \times n)$,

I : The identity matrix of order n,

A^T : The transpose of matrix $A \in \mathbb{R}^{m\times n}$,

A^{-1} : The inverse of a regular matrix A,

a_{ij} : Matrix element of $A \in \mathbb{R}^{m\times n}$,

$A^{\frac{1}{2}}$: The square root of the matrix $A \succ 0$,

$A \succeq 0$ $(A \succ 0)$: A is a positive semidefinite (positive definite) matrix,

$A \preceq 0$ $(A \prec 0)$: A is a negative semidefinite (negative definite) matrix,

$\lambda_i(A)$: The ith eigenvalue of $A \in \mathbb{R}^{n\times n}$,

$Sp(A)$: The spectrum of the matrix A,

$\mathrm{diag}(x)$: The diagonal matrix X with $X_{ii} = x_{ii}$,

$\mathrm{Tr}(A)$: Trace of matrix $A \in \mathbb{R}^{n\times n}$ $(\mathrm{Tr}(A) = \sum a_{ii} = \sum \lambda_i(A))$,

$\rho(A)$: The spectral radius of A $(\rho(A) = \max_i |\lambda_i(A)|)$,

$\det(A)$: The determinant of $A \in \mathbb{R}^{n\times n}$ $(\det(A) = \prod_i \lambda_i(A))$,

$$S^n = \left\{X : X \in \mathbb{R}^{n\times n}, X = X^T\right\},$$

$$S^n_+ = \{X : X \in S^n, X \succeq 0\},$$

$$S^n_{++} = \{X : X \in S^n, X \succ 0\},$$

$$A \bullet B = \mathrm{Tr}(A^T B) \quad \forall A, B \in \mathbb{R}^{n\times n}$$

Terminology

- LCP : Linear complementarity problem,
- SDP : Semidefinite programming,
- NT : Nesterov-Todd,
- TC : Trajectory center,
- inner : Inner iteration,
- outer : Outer iteration,
- SVD : Singular value decomposition,
- KKT : Karush-Kuhn-Tucker,
- IPMS : Interior point methods,
- SDLCP : Semidefinite linear complementarity problem,
- SDLS : Semidefinite least square,
- NS-SDLS : Non symmetric semidefinite least square,
- LMI-LS : Linear matrix inequalities least squares problem

Contents

Introduction

The field of mathematical optimization has witnessed remarkable progress over the past several decades, with interior-point methods (IPMs) emerging as one of the most significant developments in the late 20th century. Originally conceived for linear programming problems, these methods have since been extended to various classes of convex optimization problems, including semidefinite programming (SDP) and complementarity problems. This thesis focuses on the application and development of interior-point methods for solving monotone semidefinite linear complementarity problems (SDLCPs), a powerful framework that generalizes both classical linear complementarity problems and semidefinite programming.

Semidefinite linear complementarity problems arise naturally in numerous applications across engineering, economics, and operations research. They provide a unified framework for modeling problems involving matrix variables with semidefiniteness constraints, where the complementarity condition captures essential equilibrium properties. The monotonicity assumption, which ensures the problem's well-posedness, makes these problems particularly amenable to solution by interior-point methods. The growing importance of SDLCPs in modern optimization stems from their ability to model complex systems where traditional linear or nonlinear programming formulations prove inadequate.

The theoretical underpinnings of this work rest on three fundamental pillars: convex analysis, matrix optimization, and complementarity theory. Convex analysis provides the essential tools for understanding the geometric structure of the feasible region and the behavior of optimization algorithms. The theory of self-concordant barriers, developed by Nesterov and Nemirovskii, plays a crucial role in the complexity analysis of interior-point methods

for SDLCPs. This theory establishes that certain barrier functions allow the construction of polynomial-time algorithms for convex optimization problems, including those with semidefinite constraints.

Matrix optimization introduces additional layers of complexity compared to vector optimization problems. The spectral properties of symmetric matrices, their decomposition, and the various norms defined on matrix spaces all contribute to the rich structure of SDLCPs. The complementarity condition in SDLCPs, expressed through the trace inner product $\mathrm{Tr}(XY) = 0$ for matrix variables $X, Y \succeq 0$, generalizes the componentwise complementarity of traditional linear complementarity problems while maintaining crucial connections to optimality conditions in convex programming.

The core algorithmic contribution of this work lies in the development and analysis of primal-dual interior-point methods for SDLCPs. These methods follow the central path, a smooth trajectory through the interior of the feasible region that leads to optimal solutions as a barrier parameter approaches zero. The Nesterov-Todd (NT) direction, which preserves certain invariance properties crucial for efficient computation, serves as the foundation for our algorithmic framework.

A significant focus of this research is the investigation of kernel functions and their role in determining search directions. Kernel functions, which are strictly convex and barrier functions defined on the positive real line, can be extended to matrix spaces through spectral decomposition. Different choices of kernel functions lead to distinct algorithmic behaviors, with implications for both theoretical complexity and practical performance. We systematically analyze various kernel functions from the literature, comparing their properties and the resulting iteration complexity bounds.

The implementation of interior-point methods for SDLCPs presents unique computational challenges. Unlike linear programming where the main computational burden lies in solving linear systems, SDLCP algorithms require careful handling of matrix operations and eigenvalue computations. The symmetrization of the complementarity condition, necessary to maintain matrix symmetry throughout the iterations, introduces additional complexity in the linear algebra computations.

The practical significance of SDLCPs and their solution methods spans multiple disciplines. In control theory, they appear in linear matrix inequality (LMI) formulations of stability and performance analysis problems. Combinatorial optimization benefits from SDP relaxations of difficult problems, where SDLCP formulations often provide tighter bounds than linear programming relaxations. Eigenvalue optimization, a fundamental tool in numerical analysis and scientific computing, can be expressed naturally as an SDLCP. Particular applications include; Maximum cut problems in graph theory, Chebyshev approximation with logarithmic constraints, stability analysis of dynamical systems via Lyapunov equations, geometric problems involving quadratic forms and ellipsoidal approximations.

This document is organized as follows: Chapter 1 establishes the mathematical foundations, covering convex analysis, matrix calculus, and optimization theory. Chapter 2 focuses on semidefinite programming and complementarity problems, detailing their formulation, duality theory, and applications. Chapter 3 presents Primal-dual Inner point methods for SDLCP based on kernel functions.

Chapter 1

General Notions

The aim of this first chapter is to present some fundamental notions of differential calculus and convex analysis, linear programming, and the asymptotic study of the convergence of an optimization algorithm. These notions are useful for demonstrating theoretical results in the following chapters.

1.1 Some Notations

1. For any $n \in \mathbb{N}^*$, $\mathbb{R}^n$ denotes the Euclidean space $\mathbb{R} \times \mathbb{R} \times \ldots \times \mathbb{R}$ ("product n times"). In general, a vector $x \in \mathbb{R}^n$ will be denoted $x = (x_1, x_2, \ldots, x_n)^T$ (column vector).

2. We denote $e_1, e_2, \ldots, e_n$ the elements of the canonical basis of $\mathbb{R}^n$, where e_i is the vector of $\mathbb{R}^n$ given by

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

   (Kronecker symbol).

3. For all $x, y \in \mathbb{R}^n$, we denote by $\langle x, y \rangle \in \mathbb{R}$ the scalar product of x and y, which is given by

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

4. For any $x \in \mathbb{R}^n$, we denote by $\|x\| \geq 0$ the Euclidean norm of x, given by

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

5. For any $x \in \mathbb{R}^n$ and $r > 0$, $B(x, r)$ denotes the open ball centered at x with radius r, given by

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

6. If $a, b \in \mathbb{R}^n$, we denote $[a, b]$ the subset of $\mathbb{R}^n$ given by

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

   If $a, b \in \mathbb{R}^n$ with $a \neq b$, then $[a, b]$ denotes the interval of numbers $x \in \mathbb{R}$ such that $a \leq x \leq b$.

7. We also recall the Cauchy-Schwarz inequality

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

1.2 Gradient and Hessian

1.2.1 Notion of Partial Derivative

Let $\Omega \subset \mathbb{R}^n$ be an open set and $f : \Omega \to \mathbb{R}$.

1. We say that f is of class C^m on Ω ($f \in C^m(\Omega)$) if all partial derivatives up to order m exist and are continuous.

2. For any $x \in \Omega$ and any $i \in \{1, 2, \ldots, n\}$, we denote (when it exists)

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

   (This is the partial derivative of f at x in the direction x_i).

3. For any $x \in \Omega$, we denote (when it exists)

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

(The Jacobian of f at x). We have

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

4. For any $x \in \Omega$ and $h \in \mathbb{R}^n$, we denote (when it exists):

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

(This is the directional derivative of f at x in the direction h). We have

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

1.2.2 Hessian Matrix

Let $\Omega \subset \mathbb{R}^n$ be an open set and $f : \Omega \to \mathbb{R}$ a function of class C^2. The Hessian matrix of f at $x \in \Omega$ is the matrix $H(x)$ defined by

$$H(x) = \nabla^2 f(x) = \left(\frac{\partial^2 f}{\partial x_i \partial x_j} \right)_{1 \leq i,j \leq n}.$$

1.3 Convex Analysis

Convexity is a key mathematical tool for analyzing optimization problems. Here, we introduce some essential concepts frequently used in this context.

1.3.1 Convex Sets and Functions

1. A set C of $\mathbb{R}^n$ is said to be convex if:

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

2. C is said to be affine if:

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

3. C is a convex polyhedron if it is of the form:

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

where A_i is a non-zero vector of $\mathbb{R}^n$ and b_i is a scalar for $i = 1, \ldots, m$.

C can be written in the following matrix form:

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

where A is a matrix of $\mathbb{R}^{m \times n}$ and b is a vector of $\mathbb{R}^m$.

Let $f : C \to \mathbb{R}$ be a function and C a convex set of $\mathbb{R}^n$.

1. f is said to be mid-convex on C if:

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

2. f is said to be quasi-convex on C if:

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

3. f is said to be convex on C if the following inequality is satisfied:

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

4. f is said to be strictly convex on C if:

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

5. f is said to be strongly convex on C if there exists $\alpha > 0$ such that:

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

we have:

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

(We also say that f is α-convex).

6. f is convex on C if and only if f is mid-convex and quasi-convex on C.

7. If f is a continuous function on a convex set C, we have

   (a) f is convex on C if and only if f is mid-convex on C.

   (b) f is α-convex on C if and only if

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

1.4 Taylor Formulas

Let $\Omega \subset \mathbb{R}^n$ be an open set, $f : \Omega \to \mathbb{R}$, $a \in \Omega$, and $h \in \mathbb{R}^n$ such that $[a, a+h] \subset \Omega$. Then

1. If $f \in C^1(\Omega)$, then

(a) Taylor's formula of order 1 with integral remainder:

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

(b) Taylor-Maclaurin formula of order 1:

There exists $\theta \in [0,1]$ such that $f(a+h) = f(a) + \langle \nabla f(a+\theta h), h \rangle$.

(c) *iii)* Taylor-Young formula of order 1:

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

2. If $f \in C^2(\Omega)$, then

(a) Taylor's formula of order 2 with integral remainder:

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

(b) Taylor-Maclaurin formula of order 2:

There exists $\theta \in [0,1]$ such that $f(a+h) = f(a) + \langle \nabla f(a), h \rangle + \frac{1}{2} \langle \nabla^2 f(a+\theta h)h, h \rangle$.

(c) Taylor-Young formula of order 2

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

Remark 1.4.1 *In the previous formulas, the notation $o(\|h\|^k)$ for $k \in \mathbb{N}^*$ means an expression that tends to 0 faster than $\|h\|^k$ (i.e., if we divide it by $\|h\|^k$, the result tends to 0 as h tends to 0).*

1.5 General Optimization Problem

Optimization has a specific vocabulary, so we will introduce some classical notations and definitions. First, we give the general formulation of an optimization problem, for which we need

1. An objective function or cost function or criterion to minimize, denoted $f : \mathbb{R}^n \to \mathbb{R}$, which depends on several variables $x = (x_1, x_2, \ldots, x_n)$.
2. A set $A \subset \mathbb{R}^n$ where we seek the solution. We say that A is the set of feasible elements of the problem or the set of constraints.

We seek to minimize f on A, i.e., we seek $\tilde{x} \in A$ such that

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

There are two types of optimization unconstrained and constrained. In both cases, the goal is to find the values that minimize a function. However, in constrained optimization, the solutions are subject to restrictions.

1.5.1 Necessary Conditions for a Minimum

Let $f : \mathbb{R}^n \to \mathbb{R}$ be a function and let the problem

$$\min_{x \in R^n} f(x)$$

. The local and global minima of f on $\mathbb{R}^n$ are defined as follows

Definition 1.5.1 *1. We say that the function f of problem (P) has a global (or absolute) minimum at $\widetilde{x} \in \mathbb{R}^n$ if and only if*

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

2. We say that the point $\widetilde{x} \in \mathbb{R}^n$ is a local (or relative) optimum of (P) if there exists a

neighborhood V of $\widetilde{x}$ such that

$$\forall x \in V(\widetilde{x}), f(\widetilde{x}) \leq f(x).$$

3. *We say that $\widetilde{x} \in \mathbb{R}^n$ is a strict local optimum of (P) if and only if*

$$\exists V(\widetilde{x}) \text{ such that } \forall x \in V(\widetilde{x}), f(\widetilde{x}) < f(x).$$

4. *We say that $\widetilde{y} \in \mathbb{R}^n$ is a revised strict local optimum of (P) if*

$$\exists V(\widetilde{x}) \text{ such that } \widetilde{x} \text{ is the only optimal solution of } (P).$$

Remark 1.5.1 *1. A global minimum is clearly a local minimum.*

2. *If we simply say minimum, we mean global minimum.*

3. *For maximization, minimize the function $(-f)$*

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

1.5.2 Extremum Points

Let $f : \mathbb{R}^n \to \mathbb{R}$ be a function, and $\widehat{x} \in \mathbb{R}^n$.

1. We say that $\widehat{x}$ is a point of absolute (respectively relative) maximum of f on $\mathbb{R}^n$ if $\widehat{x}$ is a point of absolute (respectively relative) minimum of $-f$ on $\mathbb{R}^n$.

2. We say that $\widehat{x}$ is a point of absolute (respectively relative) extremum of f on $\mathbb{R}^n$ if $\widehat{x}$ is either a point of absolute (respectively relative) minimum of $-f$ on $\mathbb{R}^n$ or a point of absolute (respectively relative) maximum of f on U.

Lemma 1.5.1 *Let $U \in \mathbb{R}^n$, $a \in \mathbb{R}^n$, and u^* an element belonging to the interior of U ($u^* \in \overset{\circ}{U}$). Then the following two assertions are equivalent*

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

2. $a = 0$.

Definition 1.5.2 *Let $U \subset \mathbb{R}^n$ be a set and $u^* \in U$. We say that $w \in \mathbb{R}^n$ is an admissible direction for u^* in U if there exists $t_0 > 0$ such that $u^* + tw \in U$ for all $t \in [0, t_0]$.*

Example 1.5.1

1. *If $u^* \in \overset{\circ}{U}$, then any vector $w \in \mathbb{R}^n$ is an admissible direction for u^* in U.*
2. *If U is convex, then for any $\nu \in U$, the vector $\nu - u^*$ is an admissible direction for u^* in U.*

Lemma 1.5.2 *Let $\Omega \subset \mathbb{R}^n$ be an open set, $U \subset \Omega$, $f : \Omega \to \mathbb{R}$ a function of class C^1, and $u^* \in U$ a point of local minimum of f on U. Let $w \in \mathbb{R}^n$ be an admissible direction for u^* in U. Then*

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

Lemma 1.5.3 *Let $\Omega \subset \mathbb{R}^n$ be an open set, $U \subset \Omega$ a convex set, and $f : \Omega \to \mathbb{R}$ a function of class C^1. Let $u^* \in U$ be a point of local minimum of f on U. Then*

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

   (This is the Fermat condition.)

2. *If in addition u^* is in the interior of U ($u^* \in \overset{\circ}{U}$), then the condition*

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

   is equivalent to

$$\nabla f(u^*) = 0 \quad \textit{(This is the Euler equation).}$$

1.5.3 Optimality Conditions

Theorem 1.5.1 *Let $\Omega \subset \mathbb{R}^n$ be an open set, $U \subset \Omega$ a convex set, $f : \Omega \to \mathbb{R}$ a function of class C^1 and convex, and $u^* \in U$. Then the following three assertions are equivalent*

1. *u^* is a point of global minimum of f on U.*

2. u^* *is a point of local minimum of* f *on* U.

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

Remark 1.5.2 *In the case where* $u^* \in \overset{\circ}{U}$*, then assertion 3 of Theorem 1.5.1 can be replaced, thanks to Lemma 1.5.3, by the Euler equation:*

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

1.6 Mathematical Programming

Mathematical programming is a broad and rich field in numerical analysis. It addresses several mathematical models and important practical problems.

1.6.1 Unconstrained Optimization Problem

We will study the unconstrained optimization problem in which we minimize the function $f : \mathbb{R}^n \to \mathbb{R}$ over the entire space $\mathbb{R}^n$. We consider the problem formulated as follows

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

which can be written as

$$\{\text{find } \hat{x} \in \mathbb{R}^n \text{ such that } f(\hat{x}) \leq f(x); \quad \forall x \in \mathbb{R}^n\}$$

1.6.1.1 Existence and Uniqueness

We are now interested in the question of the existence of minima for unconstrained optimization problems.

Theorem 1.6.1 (Existence) *Let* $f : \mathbb{R}^n \to \mathbb{R}$ *be an application such that*

1. f *is continuous.*

2. f *is coercive (i.e.,* $\lim_{||x|| \to \infty} f(x) = +\infty$*).*

Then, there exists $\widehat{x} \in \mathbb{R}^n$ such that $f(\widehat{x}) \leq f(y)$ for all $y \in \mathbb{R}^n$.

Theorem 1.6.2 (Uniqueness) *Let $f : \mathbb{R}^n \to \mathbb{R}$ be strictly convex. Then there exists at most one $\widehat{x} \in \mathbb{R}^n$ such that:*

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

Theorem 1.6.3 *Let f be a function of class C^1. Suppose there exists $\alpha > 0$ (called the ellipticity constant) such that for all $(x, y) \in \mathbb{R}^n \times \mathbb{R}^n$, we have:*

$$(\nabla f(x) - \nabla f(y), x - y) \geq \alpha \|x - y\|^2 \quad (\alpha - elliptic)$$

Then f is strictly convex and coercive.

1.6.1.2 Optimality Conditions

In this section, we will seek to obtain the necessary and sometimes sufficient conditions for minimality, since these optimality conditions will most often be used to facilitate the calculation of a minimum. Therefore, these conditions will be expressed using the first or second derivative.

The two necessary conditions for optimality are as follows:

First-Order Necessary Conditions

Given a vector $\widehat{x} \in \mathbb{R}^n$, we would like to be able to determine if this vector is a local or global minimum of the function f. The property of continuous differentiability of f provides a first way to characterize an optimal solution.

Theorem 1.6.4 *Let $f : \mathbb{R}^n \to \mathbb{R}$ be a differentiable function at the point $\widehat{x} \in \mathbb{R}^n$. If $\widehat{x}$ is a local optimum of (P), then*

$$\nabla f(\widehat{x}) = 0 \qquad (*)$$

Remark 1.6.1 *1. A point $\widehat{x}$ in $\mathbb{R}^n$ that satisfies $\nabla f(\widehat{x}) = 0$ is called a critical point or*

stationary point.

2. *The previous theorem does not apply if the function f is not differentiable.*

3. *Condition $(*)$ is a first-order condition because it involves only the first derivative of the function f.*

There are situations where the relation $(*)$ is a necessary and sufficient condition.

Theorem 1.6.5 *Let $f : \mathbb{R}^n \to \mathbb{R}$ be a convex function C^1. A point $\widehat{x}$ achieves a minimum over $\mathbb{R}^n$ if and only if $\nabla f(\widehat{x}) = 0$.*

We now give a necessary condition to further specify the possible minima. This condition will involve the second derivative of f.

Second-Order Necessary Conditions

Theorem 1.6.6 *Let $f : \mathbb{R}^n \to \mathbb{R}$ be of class C^2. If f has a local minimum at $\widehat{x}$, then*

1. *$\nabla f(\widehat{x}) = 0$, and*

2. *The Hessian matrix $H(\widehat{x})$ is positive semi-definite.*

Second-Order Sufficient Conditions

The conditions given above are necessary, meaning that they must be satisfied for any local minimum. However, any vector that satisfies these conditions is not necessarily a local minimum. The following theorem establishes a sufficient condition for a vector to be a local minimum if f is continuously differentiable twice.

Theorem 1.6.7 *Let $f : \mathbb{R}^n \to \mathbb{R}$ be of class C^2. If*

1. *$\nabla f(\widehat{x}) = 0$, and*

2. *$H(\widehat{x})$ is positive definite,*

then f has a local minimum at $\widehat{x}$.

1.6.2 Constrained Optimization Problem

A constrained optimization problem is defined as follows.

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

$f : \mathbb{R}^n \to \mathbb{R}$ is continuous, $C \subseteq \mathbb{R}^n$ is the set of constraints. If $(C = \mathbb{R}^n)$, (P) is called an unconstrained optimization problem.

1.6.2.1 Mathematical Program

In general, a mathematical program is defined as follows

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

where

$$C = \begin{cases} x \in \mathbb{R}^n / g_i(x) \leq 0, i = 1, \ldots, m \\ h_j(x) = 0, j = 1, \ldots, p \end{cases}$$

and f, g_i, h_j are given functions from $\mathbb{R}^n$ to $\mathbb{R}$.

f is called the objective function and C the set of feasible solutions or the constraint set or simply "the domain".

A feasible solution of (PM) is any point x^* satisfying the constraints (i.e., $(x^* \in C)$).

1.6.2.2 Classification and Resolution of a Mathematical Program

The classification of (PM) and its numerical treatment are established based on the fundamental properties of the functions f, g_i, h_j, namely convexity, differentiability, and linearity. Among the most studied special cases, we note

- Linear programming (f linear, g_i, h_j affine).

- Convex programming (f, g_i convex, h_j affine, C convex).
- Integer programming (C is a discrete set, i.e., the variables are integers).

1.6.2.3 Existence and Uniqueness of Solution

In this paragraph, we give two existence theorems and the uniqueness theorem.

Theorem 1.6.8 *If C is a non-empty compact subset of $\mathbb{R}^n$ and if f is continuous on C, then (PM) admits at least one optimal solution $x^* \in C$.*

Theorem 1.6.9 *If C is a closed non-empty subset of $\mathbb{R}^n$, f is continuous and coercive on C (i.e., $\lim_{||x||\to\infty} f(x) = +\infty$), then (PM) admits at least one optimal solution.*

Theorem 1.6.10 *If C is a convex non-empty subset of $\mathbb{R}^n$, f is strictly convex on C, then (PM) admits at most one optimal solution.*

1.6.2.4 Optimality Conditions

Before giving the optimality conditions of (PM), we require that the constraints satisfy certain criteria called "qualification criteria".
A constraint g_i is said to be active (or saturated) at $\bar{x} \in C$ if $g_i(\bar{x}) = 0$.

A point $\bar{x} \in C$ is said to be regular (or the constraints are qualified at $\bar{x}$) if the gradient components corresponding to the saturated constraints at $\bar{x}$ are linearly independent. There are also two usual qualification criteria at any point of C, namely

- If all constraints are affine.
- If C is defined only by inequalities, we have Slater's condition: $g_i(x)$ is convex for all $i = 1, \ldots, m$ and there exists a point x^0 such that $g_i\left(x^0\right) < 0$, (int $(C) \neq \varnothing$).

1.6.2.5 Lagrangian Duality

Let

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

and consider the primal problem

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

The Lagrangian associated with this problem is the function $L : D \times [0, +\infty]^k \times \mathbb{R}^m \longrightarrow \mathbb{R}$, defined by

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

We set

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

Thus

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

The dual problem associated with the primal problem is

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

We have the duality inequality

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

Theorem 1.6.11 *(Karush-Kuhn-Tucker): Let $\bar{x} \in C$ satisfy one of the qualification conditions and suppose that f, g_i, h_j are $C^1(\mathbb{R}^n)$. We have:*

If $\bar{x}$ is a local optimum for (PM), then there exist real numbers called Lagrange multipliers:

$\mu_i \in \mathbb{R}^+, i = 1, \ldots, m$ *and* $\lambda_j \in \mathbb{R}, j = 1, \ldots, p$ *such that:*

$$\begin{cases} \nabla f(\bar{x}) + \sum_{i=1}^{m} \mu_i \nabla g_i(\bar{x}) + \sum_{j=1}^{p} \lambda_j \nabla h_j(\bar{x}) = 0 \\ \mu_i g_i(\bar{x}) = 0, \\ i = 1, \ldots, m \\ h_j(\bar{x}) = 0, \\ j = 1, \ldots, p \end{cases}$$

(optimality conditions) (complementarity conditions)

If, in addition, f, g_i, h_j *are convex, the above conditions are both necessary and sufficient for* $\bar{x}$ *to be a global optimum for* (PM).

1.6.3 Optimization Algorithm

We will present an algorithm to converge to an optimal solution of the problem (PM). Most constrained optimization algorithms exploit the optimality conditions to determine local minima. We will give some definitions here.

1.6.3.1 Description

An algorithm is defined by an application A, from C to C, where C is the set of feasible solutions, allowing the generation of a sequence of elements of C by the formula

$$\begin{cases} x_0 \in C \text{ given}, k = 0 \quad \text{Initialization step} \\ x_{k+1} = A\left(x_k\right), k = k + 1 \quad \text{Iteration} \end{cases}$$

If we replace C with its interior, assuming that $\text{int}(C) \neq \varnothing$, the algorithm is called an interior point algorithm.

Defining an algorithm is nothing more than constructing a sequence $(x_k)_{k \in \mathbb{N}}$ of C and conducting a study to show its convergence.

1.6.3.2 Convergence

Definition 1.6.1 *We say that the algorithm A is convergent if the sequence $(x_k)_{k\in\mathbb{N}}$ generated by the algorithm converges to a limit x^*.*

1.6.3.3 Rate of Convergence

A criterion for measuring the speed (or rate) of convergence is the evolution of the error made at each iteration ($e_k = \|x_k - x^*\|$).

Before giving the notations of convergence, we give the definitions of the following asymptotic notations

Definition 1.6.2 (Notation O) *Let $f, g : \mathbb{N} \to \mathbb{R}^+$. We write $f(n) = \mathrm{O}(g(n))$ when there exist integers c and n_0 such that for all $n \geq n_0$,*

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

Intuitively, this means that the value of the function f is less than that of g up to a multiplicative constant, for sufficiently large instances (data). Similarly, we define

Definition 1.6.3 ((Notations $\mathbf{o}, \mathbf{\Omega}, \Theta$)) *Let $f, g : \mathbb{N} \to \mathbb{R}^+$.*

- *We write $f(n) = \mathbf{o}(g(n))$ when for all real c, there exists an integer n_0 such that for all $n \geq n_0$,*

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

- *We write $f(n) = \mathbf{\Omega}(g(n))$ when there exist integers c and n_0 such that for all $n > n_0$,*

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

- *We write $f(n) = \Theta(g(n))$ when $f(n) = \mathbf{O}(g(n))$ and $f(n) = \mathbf{\Omega}(g(n))$.*

Let $(x_k)_{k\in\mathbb{N}}$ be a sequence given by the algorithm A and convergent to x^*. The classification of the speed of convergence of a sequence is based on the notions of comparison of functions in the neighborhood of $+\infty$.

Indeed, if we assume that the error e_k does not vanish, the speed of convergence can be

- Linear: If $\|e_k\| = \Omega\left(\|e_{k+1}\|\right)$ and $\left(\frac{\|e_{k+1}\|}{\|e_k\|}\right) < 1$, for sufficiently large k. We also say that the error e_k decreases linearly, i.e.,

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

- Superlinear: If $\|e_{k+1}\| = \mathbf{o}\left(\|e_k\|\right)$, where the error decreases as follows $\exists \alpha_k$ a positive sequence that converges to 0 such that $e_{k+1} \leq \alpha_k e_k$.

- Of order γ with $\gamma > 1$: If $\|e_{k+1}\| = \mathrm{O}\left(\|e_k\|^\gamma\right)$ and $\left(\frac{\|e_{k+1}\|}{\|e_k\|^\gamma}\right) < 1$, for sufficiently large k, where the error decreases as follows

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

  In the case $\gamma = 2$, the convergence is said to be quadratic.

1.6.4 Linear Programming (PL)

A linear program (PL) is an optimization problem that consists of maximizing (or minimizing) a linear objective function of n decision variables subject to a set of constraints expressed as linear equations or inequalities.

1.6.4.1 General Form

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

where $A \in \mathbb{R}^{m\times n}$ and $D \in \mathbb{R}^{p\times n}$ are given matrices, $c \in \mathbb{R}^n, b \in \mathbb{R}^m$ and $e \in \mathbb{R}^p$ are given vectors.

We can show that any linear program can be reduced to one of the following two forms

1.6.4.2 Canonical Form

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

1.6.4.3 Standard Form

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

where A is a real matrix of type (m, n) assumed to be of full rank (i.e., $\text{rg}(A) = m < n$), b is a vector in $\mathbb{R}^m$.

In the following, we are interested in the linear programming problem (PL) in the following standard form

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

The dual of the linear program (PL) is a linear program defined by

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

We denote by

- $F_{(PL)} = \{x \in \mathbb{R}^n : Ax = b, x \geq 0\}$, the set of feasible primal solutions of (PL).
- A vector $x \in F_{(PL)}$ is called a feasible solution of (PL).
- A vector $x^* \in F_{(PL)}$ minimizing the objective function of (PL) is called an optimal solution of (PL).
- A feasible linear program (PL) is bounded if the objective function is bounded on $F_{(PL)}$.
- $\overset{\circ}{F}_{(PL)} = \{x \in \mathbb{R}^n : Ax = b, x > 0\}$, the set of strictly feasible primal solutions of (PL).
- $F_{(DL)} = \left\{y \in \mathbb{R}^m : A^t y + s = c, s \geq 0\right\}$, the set of feasible dual solutions of (DL).
- A vector $y^* \in F_{(DL)}$ maximizing the objective function of (DL) is called an optimal solution of (DL).
- $\overset{\circ}{F}_{(DL)} = \left\{y \in \mathbb{R}^m : A^t y + s = c, s > 0\right\}$, the set of strictly feasible dual solutions of (DL).

- $\overset{\circ}{F}' = \overset{\circ}{F}_{(PL)} \times \overset{\circ}{F}_{(DL)}$, the set of strictly feasible primal-dual solutions of (PL) and (DL).

We give some fundamental results of duality in linear programming

- If one of the problems (PL) and (DL) admits an optimal solution, the same is true for the other, and their corresponding optimal values are equal.
- If one of the problems has an infinite optimal value, the other has no optimal solution.

Theorem 1.6.12 (Weak Duality) *If x and (y, s) are respectively feasible solutions for (PL) and (DL), then*

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

Theorem 1.6.13 (Strong Duality) *If $\bar{x}$ and $(\bar{y}, \bar{s})$ are respectively feasible solutions corresponding to a finite optimal value for (PL) and (DL) such that*

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

then $\bar{x}$ is an optimal primal solution of (PL) and $\bar{y}$ is an optimal dual solution of (DL).

Remark 1.6.2 *We can easily remark that if $\bar{x}$ and $(\bar{y}, \bar{s})$ are feasible solutions of (PL) and (DL) respectively, then we have the following property:*

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

1.7 Matrices calculation

Matrix analysis provides essential tools for optimization problems. For instance, the second-order partial derivatives make up the Hessian matrix, which is essential in assessing the nature of critical points (i.e., minima, maxima, or saddle points) of amultivariable function. This is an important topic in optimization, where the eigenvalues of the Hessian matrix

specify whether the function is convex or concave. (Semi-definite matrices come into play inoptimization. Thepositive semi-definiteness of the Hessian carries specific implications on the convexity of the function within a region. Convex functions have the nice property that every local minimum is also a global minimum. Positive semi-definite matrices are fundamental to the theory andalgorithms of semi-definite programming.
In this section, we present some known results for norms, symmetric matrices, and semidefinite matrices.

1.7.1 Scalar product and norms

We start with the definition of the scalar product of two vectors.

Definition 1.7.1 *The usual scalar product of two vectors X and Y of $\mathbb{R}^n$ is defined by:*

$$\langle X, Y\rangle = \sum_{i=1}^{n} x_i y_i = X^T Y.$$

Similarly, we define a scalar product on the set of real square matrices.

Definition 1.7.2 *Let $A, B \in \mathbb{R}^{n\times n}$, the scalar product of A and B denoted $A \bullet B$ is defined by:*

$$A \bullet B = \mathrm{Tr}(A^T B) = \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij} b_{ij} = B \bullet A.$$

We recall that $\mathrm{Tr}(\cdot)$ *is the trace of* $\mathbb{R}^{n\times n}$ *matrix.*

It is important to note that the trace is a linear function; moreover, it verifies the following properties

1. $\forall A, B \in \mathbb{R}^n : Tr(AB) = Tr(BA)$,
2. $\forall A \in \mathbb{R}^{m\times n} : Tr(A) = Tr(A^T)$,
3. $\forall A, B \in \mathbb{R}^n : A \sim B \Rightarrow Tr(A) = Tr(B)$.

Let us now state the notion of vector norm.

Definition 1.7.3 *The vector norm is an application from* $\mathbb{R}^n$ *to* $\mathbb{R}_+$ *denoted by* $\|\cdot\|$ *that satisfies the following conditions:*

1. $\forall x \in \mathbb{R}^n : \|x\| = 0 \Leftrightarrow x = 0$,
2. $\forall x \in \mathbb{R}^n, \forall \alpha \in \mathbb{R} : \|\alpha x\| = |\alpha| \|x\|$,
3. $\forall x, y \in \mathbb{R}^n : \|x + y\| \leq \|x\| + \|y\|$.

The matrix norm associated with the scalar product of two matrices is defined by

Definition 1.7.4 *Let* $A, B \in \mathbb{R}^{n \times m}$, *the application* $\|\cdot\| : \mathbb{R}^{n \times m} \to \mathbb{R}_+$ *is called matrix norm if it verifies the following conditions:*

1. $\|A\| = 0 \Leftrightarrow A = 0, \ \forall A \in \mathbb{R}^{n \times m}$,
2. $\|\alpha A\| = |\alpha| \|A\|, \ \forall \alpha \in \mathbb{R}$,
3. $\|A + B\| \leq \|A\| + \|B\|, \ \forall A, B \in \mathbb{R}^{n \times m}$,
4. $\|AB\| \leq \|A\| \|B\|, \ \forall A, B \in \mathbb{R}^{n \times m}$.

Note that the usual matrix norms for any $A \in \mathbb{R}^{n \times m}$ are:

$$\|A\|_1 = \max_{1 \leq j \leq m} \sum_{i=1}^{n} |a_{ij}|,$$

$$\|A\|_\infty = \max_{1 \leq i \leq n} \sum_{j=1}^{m} |a_{ij}|.$$

and

$$\|A\|_2 = \sqrt{\rho(AA^T)}.$$

This latter norm is called the spectral norm. We will also use the Frobenius norm.

$$\|A\|_F = \sqrt{\mathrm{Tr}(A^* A)} = \sqrt{\sum_{i=1}^{n} \sum_{j=1}^{m} (a_{ij})^2}.$$

If $A = A^T$, we have

$$\|A\|_F = \sqrt{\mathrm{Tr}(A^T A)} = \sqrt{\mathrm{Tr}(A)^2} = \sqrt{\sum_{k=1}^{n} (\lambda_k)^2},$$

and

$$\|A\|_2 = \sqrt{\rho(A^T)} = \max_{i=1}^{n} |\lambda_i|, \quad \|A\|_2 \leq \|A\|_F \leq \sqrt{n}\|A\|_2.$$

For any matrix norm, we have

$$\rho(A) \leq \|A\|.$$

Let us recall that S^n denotes the set of real symmetric matrices of order n. Here are some important properties

1. $\forall X, Y \in S^n$:

$$\mathrm{Tr}(XY) = \mathrm{Tr}(YX),$$

2. If A is an invertible matrix, then

$$\mathrm{Tr}(AXA^{-1}) = \mathrm{Tr}(X).$$

Definition 1.7.5 *Let $X \in S^n$, then*

1. *X is a positive semi-definite matrix ($X \in S^n_+$, or $X \succeq 0$) if for all $u \in \mathbb{R}^n$, $u^T X u \geq 0$.*
2. *X is a positive definite matrix ($X \in S^n_{++}$, or $X \succ 0$) if for all $u \in \mathbb{R}^n$, $u \neq 0$, $u^T X u > 0$.*

Theorem 1.7.1 *Let $A \in S^n$, the following conditions are equivalent*

1. *$A \in S^n_+$ (resp. $A \in S^n_{++}$),*
2. *$\lambda_{\min}(A) \geq 0$ (resp. $\lambda_{\min}(A) > 0$),*

3. $\exists P \in \mathbb{R}^{n\times n} : A = PP^T$ *(resp.* $\exists P \in \mathbb{R}^{n\times n} : \mathrm{rg}(P) = n, A = PP^T$*).*

Theorem 1.7.2 *Let* $A \in S^n_{++}$*, then there exists a unique matrix* $B \in S^n_{++}$ *such that* $A = B^2$*. Moreover,* B *is called the square root of* A*.*

Lemma 1.7.1 *Let* $A, B \in S^n_{++}$*, the following assertions are equivalent:*

1. $A \bullet B = 0$,
2. $AB = 0$,
3. $\frac{1}{2}(AB + BA) = 0$.

Lemma 1.7.2 *Let* $A, B \in S^n_{++}$*, then*

$$\lambda_{\min}(A)\lambda_{\min}(B) \leq \lambda_{\min}(A)\mathrm{Tr}(B) \leq A \bullet B \leq \lambda_{\max}(A)\mathrm{Tr}(B) \leq n\lambda_{\max}(A)\lambda_{\max}(B).$$

Theorem 1.7.3 *Let* $A, B \in S^n$*, we have:*

1. $\mathrm{Tr}(AB) \geq 0$ *for all matrices* $A, B \succeq 0$,
2. $\forall A, B \succeq 0, \mathrm{Tr}(AB) = 0$ *if and only if* $AB = 0$,
3. $\forall A \in S^n, \mathrm{Tr}(AB) \geq 0$ *for all* $B \succeq 0$ *implies* $A \succeq 0$.

Now, we present some definitions and results of matrix calculation.

Definition 1.7.6

1. *$A \in \mathbb{R}^{n\times n}$ is a **normal** matrix if $AA^T = A^TA$. Moreover, if $AA^T = I$ then A is said to be **orthogonal**, where I is the identity matrix of order n.*
2. *A normal matrix A is said to be **positively stable** if and only if A is positive definite.*
3. *Any matrix $A \in S^n$ is diagonalizable, i.e., there exists an orthogonal matrix U such that:*

$$A = UDU^T \quad with \quad D = \mathrm{Diag}(\lambda_1, \lambda_2, \ldots, \lambda_n),$$

where λ_i are the eigenvalues of A.

4. *If $A \succeq 0$ then $UAU^T \succeq 0$ for any orthogonal matrix U.*

5. *For $A, B \in S^n$, if $AB = BA$ (i.e., A and B are commuting matrices), then there exists an orthogonal matrix U and two diagonal matrices Λ and Σ such that:*

$$A = U\Lambda U^T \quad \textit{and} \quad B = U\Sigma U^T.$$

Definition 1.7.7 *Let $A \in S^n$. If $A \succeq 0$ with spectral decomposition $A = U\Lambda U^T$, the matrix square root $\sqrt{A}$ is defined by:*

$$\sqrt{A} = U \begin{pmatrix} \sqrt{\lambda_1} & 0 & \cdots & 0 \\ 0 & \sqrt{\lambda_2} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \sqrt{\lambda_n} \end{pmatrix} U^T,$$

where λ_i are the eigenvalues of A (with U orthogonal and $\Lambda = \mathrm{Diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)$). The matrix $\sqrt{A}$ is called the principal square root of A and is itself a positive semidefinite symmetric matrix.

1.7.2 Singular Value Decomposition

The Singular Value Decomposition (SVD) generalizes the notion of eigenvalues to rectangular $(m \times n)$ matrices and enables the construction of generalized inverses [**?**, **?**].

Lemma 1.7.3 *Let $A \in \mathbb{R}^{m\times n}$, then the matrix $AA^T \in S^n_+$.*

Definition 1.7.8 *Let $A \in \mathbb{R}^{m\times n}$. The singular values of A are the square roots of the eigenvalues of AA^T.*

Theorem 1.7.4 *Let $A \in \mathbb{R}^{m\times n}$ with at least one nonzero eigenvalue of AA^T. There exist unitary matrices $U \in \mathbb{R}^{m\times m}$, $V \in \mathbb{R}^{n\times n}$, and a diagonal matrix $W \in \mathbb{R}^{m\times n}$ such that*

(i) $W = U^T AV$

(ii) The matrix W has the form:

$$W = \begin{pmatrix} D & 0_{r\times(n-r)} \\ 0_{(m-r)\times r} & 0_{(m-r)\times(n-r)} \end{pmatrix},$$

where $D = \mathrm{diag}(\sigma_1, \sigma_2, \ldots, \sigma_r)$ with $r = \min(m, n)$, and $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0$ are the nonzero singular values of A.

Chapter 2

Semidefinite Programming problem and complementarity

In this chapter, we will study minimization problems of linear functions with linear constraints, where the variable is a symmetric positive semi-definite matrix $X \in \mathbb{S}^n$.

2.1 Problem Formulation

A semidefinite program is a mathematical program with

- A set of variables x_{ij}.
- A linear objective function.
- A set of linear constraints.
- A positive semi-definiteness constraint.

Definition 2.1.1 *A semidefinite program in standard form is written as follows:*

$$(SDP)\begin{cases} \min & C \bullet X = \langle C, X\rangle = Tr(CX) \\ s.t. & A_i \bullet X = b_i, \quad i = 1, \ldots, m \\ & X \succeq 0, \end{cases}$$

where: $b = (b_1, b_2, \ldots, b_m) \in \mathbb{R}^m$, *and* C, X *and* A_i $(i = 1, \ldots, m)$ *are matrices in* $\mathbb{S}^n$.

Proposition 2.1.1 *A semidefinite program is a convective program. For the proof, it is easy to see that the objective function is a linear function, so demonstrating that a semidefinite program is convective comes down to demonstrating that the set of constraints of the problem (SDP) is a convective set.*

Definition 2.1.2 *1) A matrix* $X \in \mathbb{S}^n$ *is called feasible for* (SDP) *if*

$$X \succeq 0, \quad A_i \bullet X = b_i, \quad i = 1, \ldots, m.$$

The set of primal feasible solutions of (SDP) *is denoted by* $\mathcal{F}_p = \{X \in \mathbb{S}^n \mid X \text{ is feasible}\}$.

2) A matrix $X \in \mathbb{S}^n$ *is called strictly feasible for* (SDP) *if*

$$X \succ 0, \quad A_i \bullet X = b_i, \quad i = 1, \ldots, m.$$

The set of strictly primal feasible solutions of (SDP) *is denoted by* $\mathcal{F}_p^* = \{X \in \mathbb{S}^n \mid X \text{ is strictly feasible}\}$.

Definition 2.1.3 *1) The primal optimal value of* (SDP) *is defined by:*

$$p^* = \inf \{C \bullet X \mid X \in \mathbb{S}^n_+,\ A_i \bullet X = b_i,\ i = 1, \ldots, m\}.$$

2) X^* *is a primal optimal solution of* (SDP) *if*

$$X^* \in \mathcal{F}_p \text{ and } C \bullet X^* = p^*.$$

Several mathematical programs can be formulated as an (SDP) problem. We present a few of them here.

2.1.1 Linear Programming

In practice, transforming a linear program into a semidefinite program (SDP) is not numerically advantageous, even though a linear program can theoretically be written as an (SDP). Specifically

A linear program (LP) is a problem of the form:

$$(LP)\begin{cases} \min & c^T x \\ \text{s.t.} & Ax = b \\ & x \geq 0, \end{cases}$$

where $x, c \in \mathbb{R}^n$, $b \in \mathbb{R}^m$, and $A \in \mathbb{R}^{m \times n}$.

To convert a linear program (LP) into a semidefinite program (SDP), follow these steps

$$X = \text{diag}(x), \quad C = \text{diag}(c), \text{ and } A_i = \text{diag}(a_i) \text{ for } i \in \{1, \ldots, m\},$$

where a_i denotes the i-th row of the matrix A. This preserves the linear objective function in the form

$$\langle C, X \rangle = c^T x,$$

and for the constraints, we have

$$A_i \bullet X = b_i, \text{ where } A_i \bullet X = a_i^T x, \quad i = 1, \ldots, m.$$

Additionally,

$$x \geq 0 \implies X \succeq 0.$$

Thus, the (LP) in the variable x can be rewritten in the standard form of a (SDP) in the

matrix variable X as follows

$$(SDP)\begin{cases} \min & \langle C, X\rangle \\ \text{s.t.} & A_i \bullet X = b_i, \quad i = 1, \ldots, m \\ & X \succeq 0. \end{cases}$$

2.2 Application Domains

The preprocessing described here can be applied to any problem that can be written in the form of a quadratic constraint. This includes problems such as quadratic programming, quadratic assignment, and, more generally, any problem that can be expressed in the form of a quadratic constraint.

2.2.1 Eigenvalue search

It is about the oldest problems addressed using semi-definite programming.

2.2.1.1 Search for Extreme Progress Variables

Considering a symmetric matrix A, to search for the extreme value of the progress variable, we must find the minimum eigenvalue of A. This is given by

$$\lambda_{\min}(A) = \min_{||x||=1} x^T A x = \min_{||x||=1} \frac{x^T A x}{x^T x}$$

Let $X = xx^T$, then

$$\mathrm{Tr}(AX) = \sum_{i=1}^{n} \lambda_i - ||x||^2 = 1$$

The problem $(V.E)$ becomes

$$\min \lambda$$

$$\text{s.t. } \mathrm{Tr}(AX) = 1$$

$$X \succeq 0$$

In reality, the problem of finding the minimum value of the progress variable is written as

$$\lambda_{\min}(A) = \max_{||x||=1} x^T Ax = \max_{||x||=1} \frac{x^T Ax}{x^T x}$$

and its formulation in (SDP) is

$$\begin{cases} \max \lambda \\ \mathrm{Tr}(AX) = 1 \\ X \succeq 0 \end{cases}$$

Problems of Min-Max of Progress Variables

The problem is studied in the sense of DD0 in linear algebra as follows.

Search for the optimal value

$$\lambda^* = \min_{x \in \mathbb{R}^n} \lambda_{\min}(C + A(x))$$

Given $C \in M_n$ and

$$A : \mathbb{R}^n \to M_n, \quad x \mapsto A_0 + x_1 A_1 + \ldots + x_n A_n$$

The problem is written in a weak form (in the sense of the problem (SDP)) as follows.

$$\begin{cases} \min \lambda \\ \lambda_{\min}(C + A(x)) \leq \lambda \\ x \in \mathbb{R}^n \end{cases} \iff \begin{cases} \min \lambda \\ \lambda I - C - A(x) \succeq 0 \\ x \in \mathbb{R}^n \end{cases}$$

2.2.1.2 Spectral Norm of a Matrix

Given $A \in M_n$, we consider its spectral norm

$$||A||_2 = \sqrt{\lambda_{\max}(A^T A)} = \sqrt{\lambda_{\max}(AA^T)}$$

This norm can be calculated using semidefinite programming (SDP) as follows

$$\begin{cases} ||A||_2 = \min \gamma \\ \begin{pmatrix} \gamma I_n & A \\ A^T & \gamma I_m \end{pmatrix} \succeq 0 \end{cases}$$

We use the Schur complement theorem here.

2.3 Duality

The Duality in linear (SDP) is very similar to the duality in classical (LP), with a few key differences.

Consider the standard linear (SDP) problem

$$(SDP) \begin{cases} \min C \bullet X \\ A_i \bullet X = b_i, \quad i = 1, \ldots, m \\ X \succeq 0 \end{cases}$$

To obtain the dual problem of (SDP), we consider the Lagrangian function

$$\begin{aligned} q(y) &= \min_{X \in \mathbb{S}^n_+} \left[C \bullet X + \sum_{i=1}^{m} (b_i - A_i \bullet X) y_i, \quad y \in \mathbb{R}^m \right] \\ &= \min_{X \in \mathbb{S}^n_+} \left[\left(C - \sum_{i=1}^{m} y_i A_i \right) \bullet X + b^T y, \quad y \in \mathbb{R}^m \right] \end{aligned}$$

We then have

$$\max_{y\in\mathbb{R}^m} q(y) = \begin{cases} \max b^T y & \text{if } C - \sum_{i=1}^m y_i A_i \succeq 0 \\ -\infty & \text{otherwise.} \end{cases}$$

Therefore, by convention, the dual of (SDP) is

$$(DSDP) \begin{cases} \max b^T y \\ C - \sum\limits_{i=1}^m y_i A_i \succeq 0, \quad i = 1, \ldots, m \\ y \in \mathbb{R}^m. \end{cases}$$

This is a semi-definite program that can also be written as

$$(DSDP) \begin{cases} \max b^T y \\ S + \sum\limits_{i=1}^m y_i A_i = C, \quad i = 1, \ldots, m \\ y \in \mathbb{R}^m, \quad S \succeq 0. \end{cases}$$

Definition 2.3.1 *A pair $(y, S) \in \mathbb{R}^m \times \mathbb{S}^n$ is called a feasible solution of $(DSDP)$ if*

$$S \succeq 0, \quad \sum_{i=1}^m y_i A_i + S = C, \quad i = 1, \ldots, m.$$

We denote by $\mathcal{F}_d = \{(y, S) \in \mathbb{R}^m \times \mathbb{S}^n \mid (y, S) \text{ is feasible}\}$ the set of dual feasible solutions of $(DSDP)$.

Similarly, (y, S) is called a strictly feasible solution of $(DSDP)$ if

$$S \succ 0, \quad \sum_{i=1}^m y_i A_i + S = C, \quad i = 1, \ldots, m.$$

We denote by $\mathcal{F}_d^+ = \{(y, S) \in \mathbb{R}^m \times \mathbb{S}^n \mid (y, S) \text{ is strictly feasible}\}$ the set of strictly feasible dual solutions of $(DSDP)$.

Definition 2.3.2 *The optimal value of $(DSDP)$ is defined by*

$$d^* = \sup_y \left\{ b^T y \;\middle|\; C - \sum_{i=1}^m y_i A_i \in \mathbb{S}_+^n, \; i = 1, \ldots, m, \; y \in \mathbb{R}^m \right\}.$$

A pair (y^*, S^*) *is an optimal dual solution of* $(DSDP)$ *if*

$$(y^*, S^*) \in \mathcal{F}_d \text{ and } b^T y^* = d^*.$$

2.3.1 Primal-Dual Relationships in Semi-Definite Programming

As in linear programming, the (SDP) problem can be written in several forms. The following table presents the different types of corresponding dual problems.

2.3.2 Weak Duality

Proposition 2.3.1 *Let* $X \in \mathcal{F}_p$ *and* $(y, S) \in \mathcal{F}_d$. *Then*

$$C \bullet X - b^T y = S \bullet X \geq 0, \text{ and } p^* \geq d^*.$$

Proof.

1) We have

$$C \bullet X - b^T y = C \bullet X - \sum_{i=1}^{m} b_i y_i = C \bullet X - \sum_{i=1}^{m} y_i (A_i \bullet X) = \left(C - \sum_{i=1}^{m} y_i A_i \right) \bullet X = S \bullet X \geq 0,$$

since $S \succeq 0$ and $X \succeq 0$.

Therefore,

$$C \bullet X - b^T y \geq 0 \quad \forall X \in \mathcal{F}_p \text{ and } \forall (y, S) \in \mathcal{F}_d.$$

2) We have

$$C \bullet X \geq b^T y \quad \forall X \in \mathcal{F}_p \text{ and } \forall y \in \mathcal{F}_d$$

$$\Rightarrow C \bullet X^* \geq b^T y \quad \forall (y, S) \in \mathcal{F}_d$$

$$\Rightarrow p^* = C \bullet X^* \geq b^T y^* = d^*$$

$$\Rightarrow p^* \geq d^*$$

Hence weak duality holds.

■

2.3.3 Strong Duality

Although linear programming and (SDP) have very similar structures, some duality results from (LP) do not hold for (SDP). In particular, strong duality is not guaranteed unless strict feasibility is preserved for one of the two problems, as shown in the following theorem

Theorem 2.3.1 *1) If the (SDP) problem is strictly feasible, i.e., $\exists X \in \mathcal{F}_p^+$, then $p^* = d^*$. Moreover, if p^* is finite, then the set of optimal dual solutions for $(DSDP)$ is non-empty and compact.*

2) If the $(DSDP)$ problem is strictly feasible, i.e., $\exists (y, S) \in \mathcal{F}_d^+$, then $p^* = d^*$. Moreover, if d^* is finite, then the set of optimal primal solutions for (SDP) is non-empty and compact.

2.4 Complementarity in SDP

Analogous to linear programming, we can express the condition for X^* and y^* to be optimal solutions of (SDP) and $(DSDP)$, respectively, through a complementarity condition.

Theorem 2.4.1 *Let $X^* \in \mathcal{F}_P$ and $(y^*, Z^*) \in \mathcal{F}_D$ with duality gap*

$$\langle C, X^* \rangle - b^T y^* = \langle X^*, Z^* \rangle.$$

Then X^* and (y^*, Z^*) are optimal solutions for (SDP) and $(DSDP)$ respectively if and only if $X^* Z^* = 0$.

The complementarity problem can then be written as

$$(P_c) \begin{cases} A_i \bullet X = b_i & i = 1, \ldots, m, \quad X \succeq 0 \\ Z + \sum_{i=1}^{m} y_i A_i = C, & Z \succeq 0 \\ XZ = 0. \end{cases}$$

Remark 2.4.1 *Since* (SDP) *is a convex problem, the solution to the complementarity problem is a global optimum for* (SDP).

2.5 Solution Methods for SDP

The similarity between SDP and linear programming (LP) has motivated researchers to apply proven LP techniques, particularly primal-dual interior-point methods of the central path variety.

The generalization of interior-point methods from LP to SDP dates back to the early 1990s. The first algorithms in this direction were introduced independently by Alizadeh and Nesterov & Nemirovskii. Alizadeh extended Ye's potential reduction projective algorithm from LP to SDP and argued that many known interior-point algorithms for LP could be adapted to solve SDP. Nesterov and Nemirovskii developed the profound theory of interior-point methods based on self-concordant barrier functions.

2.5.1 Interior-Point Methods

Interior-point methods are among the most widely used and efficient approaches for solving SDP problems. These relatively new methods are analogous to Karmarkar's projective method for linear programming.

The term "interior-point" encompasses three distinct types of methods; affine methods, potential reduction methods and central trajectory methods. Our focus will be on central trajectory methods, also known as path-following methods, which have proven particularly effective for semidefinite programming

2.5.2 Primal-dual central trajectory methods

The principle of these methods is to minimize the duality gap by solving the following system

$$\begin{cases} A_i \bullet X = b_i & i = 1, \ldots, m, \\ Z + \sum_{i=1}^{m} y_i A_i = C, & \\ XZ = \mu I, & X, Z \succeq 0 \end{cases} \tag{(2.8)}$$

which represents the parameterized system of optimality conditions for the (SDP) problems and $(DSDP)$. The system (2.8) admits a unique solution under the assumptions that

- A_i, $i = 1, \ldots, n$ are linearly independent
- $\mathcal{F}^0 \neq \emptyset$ (the strictly feasible set is non-empty)

The solution of the system is denoted by $(X(\mu), y(\mu), Z(\mu))$ for a fixed $\mu > 0$.

The set of solutions $(X(\mu), y(\mu), Z(\mu))$ for all $\mu > 0$ defines the *central path*, which converges to the optimal solution as μ approaches 0.

To solve the nonlinear system (2.8), we use Newton's method. The goal is to compute primal and dual directions ΔX, Δy, and ΔZ respectively by solving the linear system

$$\begin{cases} A_i \bullet X = 0 & i = 1, \ldots, m, \\ \sum_{i=1}^{m} \Delta y_i A_i + \Delta Z = 0, & \\ X\Delta Z + Z\Delta X = \mu I - XZ, & X, Z \succ 0. \end{cases} \tag{2.9}$$

The complete Newton iteration is defined by

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

where $\alpha \in (0, 1]$ is a step length parameter, and must maintain strict feasibility (i.e. $X^+ \in \mathcal{F}_p^0$ and $(y^+, Z^+) \in \mathcal{F}_D^0$).

Unfortunately, X^+ is not always symmetric. To address this issue, several researchers have proposed symmetric directions in the literature. Notable contributions include the works

of Zhang, Halemeyer et al., Kojima et al. and Monteiro, Alizadeh-Haeberly-Overton, and Nesterov-Todd.

Currently, several researchers have extended the kernel function approach from linear programming to semidefinite programming to develop new directions and improve algorithm complexity. Among the various symmetric direction proposals, the Nesterov-Todd (NT) direction has emerged as particularly significant for semidefinite programming. A detailed study of this method for Semidefinite Linear Complementarity Problems (SDLCP) will be presented in the next chapter.

Chapter 3

Primal-dual Inner point methods for SDLCP based on kernel functions

3.1 Definitions and formulation

Given a linear transformation $L : S^n \to S^n$ and a symmetric matrix $Q \in S^n$, the semidefinite linear complementarity problem (SDLCP) seeks to find matrices $X, Y \in S^n_+$ such that

$$Y = L(X) + Q \quad \text{and} \quad \text{Tr}(XY) = 0,$$

where $\text{Tr}(XY)$ denotes the trace of the matrix product XY.

Definition 3.1.1 *Let $L : S^n \to S^n$ be a linear transformation. Then*

1. *L is a P-transformation if*

$$X \in S^n, \quad XL(X) = L(X)X, \quad \textit{and} \quad XL(X) \leq 0 \implies X = 0.$$

2. *L is a P_0-transformation if $(L + \epsilon I)$ is a P-transformation for all $\epsilon > 0$, where I is*

the identity transformation on S^n.

3. *L is a P_2-transformation if*

$$X \geq 0, \quad Y \geq 0, \quad (X-Y)[L(X)-L(Y)](X+Y) \leq 0 \implies X = Y.$$

4. *L is a monotone transformation (strictly monotone) if*

$$L(X) \bullet X \geq 0 \quad (L(X) \bullet X > 0), \quad \textit{for all matrices} \quad X \in S^n \quad (X \neq 0).$$

Definition 3.1.2 *1. A transformation L is strictly semi-monotone (S.S.M) or an E^--transformation if it satisfies:*

$$X \geq 0, \quad XL(X) = L(X)X, \quad \textit{and} \quad XL(X) \leq 0 \implies X = 0.$$

2. *A transformation L is semi-monotone (S.M) or an E_0-transformation if $L + \epsilon I$ is strictly semi-monotone for all $\epsilon > 0$.*

Definition 3.1.3 *Let $L : S^n \to S^n$ be a linear transformation.*

1. *L has the R_0-property if the only solution to $(SDLCP)(L, 0)$ is the zero matrix.*
2. *L has the Q-property if for every $Q \in S^n$, the problem $(SDLCP)(L, Q)$ has a solution.*

Definition 3.1.4 *A linear transformation $L : S^n \to S^n$ is said to have the globally uniquely solvable (GUS) property if, for every $Q \in S^n$, the SDLCP problem (L, Q) has a unique solution.*

Definition 3.1.5 *For a linear transformation $L : S^n \to S^n$, we say that L has*

1. *the **column sufficiency** property if, for every $Q \in S^n$, the problem $(SDLCP)(L, Q)$ has a convex solution set (which may be empty).*

2. *the* ***cross commutative*** *property if, for all* $Q \in S^n$ *and any solutions* X_1 *and* X_2 *of* $(SDLCP)(L, Q)$*, the commutativity conditions*

$$X_1 Y_2 = Y_2 X_1 \quad \textit{and} \quad X_2 Y_1 = Y_1 X_2,$$

hold, where $Y_i = L(X_i) + Q$ *for* $i = 1, 2$.

Definition 3.1.6 *Let* $L : S^n \to S^n$ *be a linear transformation. Its transpose,* $L^T : S^n \to S^n$*, satisfies*

$$\langle L(X), Y \rangle = \langle X, L^T(Y) \rangle \quad \textit{for all} \quad X, Y \in S^n.$$

A linear transformation L *is self-adjoint on* S^n *if* $L = L^T$*, and it is normal if* $LL^T = L^T L$.

Theorem 3.1.1 *For a linear transformation* $L : S^n \to S^n$*, consider the following properties*

(a) L *is monotone* ($L(X) \bullet X \geq 0$ *for all* $X \in S^n$)*,*

(b) L *is column sufficient,*

(c) L *satisfies the cross commutative property.*

1. 2. 3.

The implications between these properties are

$$(a) \Rightarrow (b) \Rightarrow (c).$$

3.2 Existence and Uniqueness

Theorem 3.2.1 *Consider a linear transformation* $L : S^n \to S^n$*. If both* $(SDLCP)(L, 0)$ *and* $(SDLCP)(L, E)$ *for some positive definite matrix* $E \in S^n$ *have unique solutions (specifically zero), then* $(SDLCP)(L, Q)$ *has a solution for every* $Q \in S^n$*.*

Theorem 3.2.2 *Let* $L : S^n_+ \to S^n_+$ *have the P-property. Then, for every* $Q \in S^n$*, the SDLCP with data* (L, Q) *has a solution.*

Theorem 3.2.3 *For a linear transformation $L : S^n \to S^n$, the following assertions are equivalent*

1. *For all $Q \in S^n$, $(SDLCP)(L, Q)$ has at most one solution,*
2. *L has the P- and cross commutative properties,*
3. *L has the GUS-property.*

Theorem 3.2.4 *Given a linear transformation $L : S^n \to S^n$, if L is monotone, then for every $Q \in S^n$, the problem $(SDLCP)(L, Q)$ has a convex (possibly empty) solution set.*

Corollary 3.2.1 *If L is a linear monotone transformation, then*

$$L \text{ has GUS-property} \iff L \text{ has } P\text{-property.}$$

Theorem 3.2.5 *If L is a strictly monotone transformation, then L has the GUS-property.*

Theorem 3.2.6 *For a linear transformation $L : S^n \to S^n$, we have the following assertions:*

$$\text{strictly monotone property} \Rightarrow P_2\text{-property} \Rightarrow \text{GUS-property.}$$

3.3 Well-known transformations

3.3.1 Lyapunov transformation

The Lyapunov transformation associated with square matrix A is defined by

$$L_A(X) = AX + XA^T.$$

Theorem 3.3.1 *Consider the Lyapunov transformation L_A with a given square matrix A. The following properties hold:*

(a) The following assertions are equivalent:

(i) L_A *has the GUS-property,*

(ii) A *is positive stable and positive semi-definite.*

(b) L_A *has the* P_2*-property if and only if* A *is positive definite.*

(c) If A *is a normal matrix, then the following are equivalent:*

(i) L_A *has the strictly monotone property,*

(ii) L_A *has the GUS-property,*

(iii) L_A *has the* P*-property.*

3.3.2 Stein transformation

The Stein transformation S_A associated with a square matrix A is defined as

$$S_A(X) = A^T X A$$

Theorem 3.3.2 *Consider the Stein transformation* S_A*, with a given square matrix* A*. The following properties hold*

(a) The following assertions are equivalent:

(i) S_A *has the GUS-property (globally uniquely solvable),*

(ii) A *is Schur stable (all eigenvalues lie inside the unit disk) and* $I - A^T a)$ A *is positive semi-definite.*

(b) S_A *has the* P_2*-property if and only if* A *is such that* $I - A^T A$ *is positive definite.*

(c) If A *is a normal matrix (*$AA^T = A^T A$*), then the following are equivalent*

(i) S_A *has the strictly monotone property,*

(ii) S_A *has the GUS-property,*

(iii) S_A *has the* P*-property.*

Definition 3.3.1 *Let A be a square matrix. The two-sided transformation associated with matrices A and A^T is defined by*

$$L_{A,A^T}(X) = AXA^T,$$

where X is a matrix of compatible dimensions.

1. **Symmetry Preservation:** If X is symmetric, then $L_{A,A^T}(X)$ is also symmetric. This is because

$$L_{A,A^T}(X)^T = (AXA^T)^T = AX^TA^T = AXA^T = L_{A,A^T}(X).$$

2. **Positive Definiteness:** If A is invertible and X is positive definite, then $L_{A,A^T}(X)$ is positive definite. This follows from the fact that for any non-zero vector v,

$$v^T L_{A,A^T}(X)v = v^T AXA^T v = (A^T v)^T X(A^T v) > 0,$$

since X is positive definite and $A^T v \neq 0$ for non-zero v.

3. **Eigenvalue Relationship:** The eigenvalues of $L_{A,A^T}(X)$ are related to the eigenvalues of X and A. Specifically, if λ is an eigenvalue of X with eigenvector u, then Au is an eigenvector of $L_{A,A^T}(X)$ with the same eigenvalue λ.

4. **Trace Preservation:** The trace of the transformation is given by

$$\mathrm{Tr}(L_{A,A^T}(X)) = \mathrm{Tr}(AXA^T) = \mathrm{Tr}(XA^TA).$$

If A is orthogonal (i.e., $A^TA = I$), then

$$\mathrm{Tr}(L_{A,A^T}(X)) = \mathrm{Tr}(X).$$

Theorem 3.3.3 *1. $L_{A,A^T}(X)$ has the GUS-property,*

2. A is positive definite or negative definite.

3.4 Primal-dual central path

Consider the following semidefinite linear complementarity problem (SDLCP): Seek a pair of matrices $(X, Y) \in S^n \times S^n$ satisfying the following conditions.

$$X,\ Y \in S^n_+,\ Y\ =\ L(X) + Q,\ \text{and}\ X \bullet Y\ =\ \mathrm{Tr}(XY)\ =\ 0, \tag{3.1}$$

where $L : S^n \to S^n$ is a linear transformation and $Q \in S^n$.

The feasible set, strict feasible set, and solution set of System (3.1) are subsets of $\mathbb{R}^{n\times n}$, denoted by $\mathcal{F}$, $\mathcal{F}^0$, and $\mathcal{S}$ respectively.

$$\begin{aligned}
\mathcal{F} &= \{\ (\ X,\ Y\ ) \in S^n \times S^n,\ Y\ =\ L(X) + Q :\ X\ \succeq\ 0,\ Y\ \succeq\ 0\}\,,\\
\mathcal{F}^0 &= \{\ (\ X,\ Y\ ) \in \mathcal{F}\ :\ X\ \succ\ 0, Y\ \succ\ 0\}\,,\\
\mathcal{S} &= \{(\ X,\ Y\ ) \in \mathcal{F}\ :\ \mathrm{Tr}(XY)\ =\ 0\}\,.
\end{aligned}$$

We assume that the (SDLCP) satisfies the following conditions

1. $\mathcal{F}^0$ is not empty,
2. L is monotone ($\langle L(X), X\rangle \geq 0, \forall X \in S^n$),
3. L is self-adjoint ($L = L^T$, such that: $\langle L(X), Y\rangle = \langle X, L^T(Y)\rangle \forall X, Y \in S^n$).

Since for $X, Y \in S^n_+, X \bullet Y = 0$ if and only if $XY = 0$. Then, finding a solution of (3.1) is equivalent to solving the following system:

$$\begin{cases} Y = L(X) + Q, \\ XY = 0, \\ X \succeq 0, Y \succeq 0. \end{cases} \tag{3.2}$$

Theorem 3.4.1 ([44]) *Under conditions **(1), (2)**, the set of solution of (SDLCP)*

$$\mathcal{S} = \{(X, Y) \in \mathcal{F} : \mathrm{Tr}(XY) = 0\},$$

is non-empty and convex.

Proposition 3.4.1 *The function $f(X) = \mathrm{Tr}(X(L(X) + Q))$ is convex, if conditions **(2), (3)** are satisfied.*

Theorem 3.4.2 *Assume that the conditions **(1), (2)** and **(3)** are satisfied, then the problem (SDLCP) is equivalent to the following convex semidefinite problem*

$$(PO) \begin{cases} \min_X \ \mathrm{Tr}(X(L(X) + Q)), \\ X \succeq 0, \\ Y = L(X) + Q \succeq 0. \end{cases}$$

Hence, finding the solution of (3.1) is equivalent to finding the minimizer of (PO).

3.4.1 Logarithmic barrier function

To introduce the method of interior point (primal-dual central path) to solve the system (PO), we associate it with the following nonlinear barrier minimization problem

$$(PO)_\mu \begin{cases} \min \left[\psi_\mu(X, Y) = XY - \mu \log\det(XY)\right], \quad \mu > 0, \\ Y = L(X) + Q, \\ X \succ 0, Y \succ 0. \end{cases}$$

where the barrier parameter $\mu > 0$.

In the following theorems, it is assumed that conditions **(1), (2), and (3)** hold.

Theorem 3.4.3 *The function $\psi_\mu(X, Y)$, $\mu > 0$ is strictly convex.*

Theorem 3.4.4 *There exists $(X^0, Y^0) \in \mathcal{F}^0$, the set*

$$\Omega_\mu = \left\{X \in S^n : \psi_\mu(X) \leq \psi_\mu(X^0)\right\}$$

is compact.

The problem $(PO)_\mu$ has a unique solution if the function $\psi_\mu(X,Y)$ is strictly convex for $\mu > 0$. Since $(PO)_\mu$ is a convex semidefinite optimization problem, the KKT conditions are necessary and sufficient. These conditions can be written as follows

$$\begin{cases} XY - \mu I = 0, \\ Y = L(X) + Q, \\ X \succ 0, Y \succ 0. \end{cases} \tag{3.3}$$

where $\mu > 0$ and I is the identity matrix.

The resolution of system (3.3) is equivalent to that of $(PO)_\mu$. Assuming a strictly feasible pair $(X^0, Y^0) \in \mathcal{F}^0$ exists satisfying the interior point condition (IPC) for (SDLCP) and that L is a linear monotone transformation, both system (3.3) and $(PO)_\mu$ have a unique solution for a fixed $\mu > 0$.

Definition 3.4.1 *The solution of the problem $(PO)_\mu$ for $\mu > 0$ is $(X(\mu), Y(\mu))$ and*

$$C = \{(X(\mu), Y(\mu)) : \mu > 0\}.$$

is the set of all solutions of the system $(PO)_\mu$ and is called the ***central path****.*

As $\mu \to 0$, the sequence $(X(\mu), Y(\mu))$ approaches the solution (X, Y) of problem (SDLCP).

Lemma 3.4.1 *Under the conditions* ***(1), (2)****, the set*

$$\{(X(\mu), Y(\mu)) : 0 \leq \mu \leq \mu^-\},$$

is bounded for all $\mu^- > 0$.

Theorem 3.4.5 $\lim_{\mu \to 0}(X(\mu), Y(\mu)) = (X^*, Y^*)$*, where (X^*, Y^*) is a solution of (SDLCP).*

3.4.2 Classical search direction

To solve system (3.3), we employ primal-dual path-following algorithms. These algorithms approximate the central path, guiding us toward the solution of (3.3) through Newton steps, where each iteration updates the variables as follows

$$X^{k+1} = X^k + \alpha\Delta X, \quad Y^{k+1} = Y^k + \alpha\Delta Y.$$

We observe that (X, Y) is a solution of (SDLCP) if and only if it solves (3.3) with $\mu = 0$. The first equation in system (3.3) is nonlinear, making direct resolution generally infeasible. Specifically,

$$F_\mu(X, Y) = 0 \quad \text{and} \quad (X, Y) \in S^n_{++} \times S^n_{++}, \quad \mu \in \mathbb{R}^*_+, \tag{3.4}$$

where $F_\mu : S^n \times S^n \to S^n \times \mathbb{R}^{n\times n}$ is defined by

$$F_\mu(X, Y) = \begin{pmatrix} XY - \mu I \\ L(X) + Q - Y \end{pmatrix} = 0.$$

Applying Newton's method to equation (3.4) yields

$$F_\mu(X, Y) + \nabla F_\mu(X, Y)(\Delta X, \Delta Y)^T = 0,$$

which simplifies to the system of linear equations

$$\begin{cases} L(\Delta X) &= \Delta Y \\ X\Delta Y + \Delta XY &= \mu I - XY \end{cases} \tag{3.5}$$

This system can be rewritten as

$$\begin{cases} L(\Delta X) &= \Delta Y \\ \Delta X + X\Delta YY^{-1} &= \mu Y^{-1} - X \end{cases} \tag{3.6}$$

Let $\mu > 0$ and (X, Y) be a strictly primal-dual feasible point. The Newton direction $(\Delta X, \Delta Y)$ at this point is the unique solution to system (3.5).

3.4.2.1 Algorithm

Algorithm approach

1. We start with a strictly primal-dual feasible point (X^0, Y^0) in a neighborhood of the central path, with a known initial parameter $\mu^0 = \frac{\mathrm{Tr}(X^0 Y^0)}{n}$;
2. Using Newton directions ΔX, ΔY, we construct the new pair (X^{k+1}, Y^{k+1});
3. We check if X^{k+1} and Y^{k+1} remain strictly feasible for all $\mu > 0$, and consider two cases:
   - If $n\mu^k < \epsilon$ (where ϵ is a specified precision), then (X^{k+1}, Y^{k+1}) are approximate solutions to system (3.3).
   - Otherwise, if $n\mu^k \geq \epsilon$, (X^{k+1}, Y^{k+1}) are not yet approximate solutions. In this case, we reduce μ to $\mu^+ = (1-\theta)\mu$ with $0 < \theta < 1$, ensuring proximity to the central trajectory.

Remark 3.4.1 *If θ depends on n, in particular $\theta = \Theta\left(\frac{1}{\sqrt{n}}\right)$ then the algorithm is called a **small-step** algorithm. If θ is a constant, $\theta = \mathcal{O}(1)$, then the algorithm is called a **large-step** algorithm.*

Symmetrization technique

Under the given conditions, the system admits a unique, generally nonsymmetric solution $(\Delta X, \Delta Y)$. To address this, replacing the last equation in (3.5) with

$$H_P(X\Delta Y + \Delta X Y) = \mu I - H_P(XY),$$

where H_P is a linear transformation defined by

$$H_P(M) = \frac{1}{2}\left(PMP^{-1} + P^{-T}M^T P^T\right),$$

for an invertible matrix P and a real square matrix M of order n. Applying this method, the linearized second Newton equation in (3.5) becomes

$$\begin{cases} L(\Delta X) = \Delta Y, \\ \Delta X + P\Delta Y P^T = \mu Y^{-1} - X, \end{cases}$$

The several choices for the matrix P include I, $Y^{\frac{1}{2}}$, $X^{\frac{1}{2}}$, X^{-1}, and Y.

3.5 Nesterov-Todd direction

We apply the Nesterov-Todd symmetrization scheme to define

$$P = X^{\frac{1}{2}}(X^{\frac{1}{2}}YX^{\frac{1}{2}})^{-\frac{1}{2}}X^{\frac{1}{2}} = Y^{-\frac{1}{2}}(Y^{\frac{1}{2}}XY^{\frac{1}{2}})^{\frac{1}{2}}Y^{-\frac{1}{2}}.$$

Let $D = P^{\frac{1}{2}}$, where $P^{\frac{1}{2}}$ denotes the symmetric square root of P.

The matrix D is used to scale X and Y to the same matrix V, defined by

$$V = \frac{1}{\sqrt{\mu}}D^{-1}XD^{-1} = \frac{1}{\sqrt{\mu}}DYD,$$

$$V^2 = \frac{1}{\mu}D^{-1}XYD.$$

Both matrices V and D are symmetric and positive definite. Using (3.9), the system (3.6) becomes

$$\begin{cases} \widetilde{L}(\Delta_X) = \Delta_Y, \\ \Delta_X + \Delta_Y = V^{-1} - V, \end{cases} \tag{3.11}$$

The equations are defined as

$$\Delta_X = \frac{1}{\sqrt{\mu}}D^{-1}\Delta X D^{-1}, \quad \Delta_Y = \frac{1}{\sqrt{\mu}}D\Delta Y D, \quad \text{and} \quad \widetilde{L}(\Delta_X) = DL(D\Delta_X D)D. \tag{3.12}$$

The linear transformation $\widetilde{L}$ is also monotone on S^n.

According to the given hypotheses, the new linear system (3.11) has a unique symmetric solution (Δ_X, Δ_Y). These directions are not orthogonal, as

$$\Delta_X \bullet \Delta_Y = \mathrm{Tr}(\Delta_Y \Delta_X) = \tfrac{1}{\mu}\Delta X \bullet L(\Delta X) \geq 0.$$

This non-orthogonality complicates the analysis compared to standard SDO problems.
To maintain the feasibility of the algorithm's iterations and their proximity to the central trajectory, we define a proximity measure

$$\delta(V) = \tfrac{1}{2}\|V^{-1} - V\|. \tag{3.13}$$

This measure satisfies

$$\delta(V) = 0 \quad \text{if and only if} \quad V = V^{-1} \quad \text{or equivalently} \quad XY = \mu I.$$

This confirms that the points are on the central trajectory.

3.6 The generic primal-dual Interior point algorithm

This section outlines the framework of a generic interior-point primal-dual algorithm for SDLCP.

- **Initialization:** We select a threshold value $\tau \geq 1$ and assume the existence of a strictly feasible initial point (X^0, Y^0) satisfying $\delta(X^0, Y^0, \mu^0) \leq \tau$ for some known μ^0.
- **Iteration Step:** Using the computed directions $(\Delta X, \Delta Y)$ and a step size $\alpha \in (0, 1)$, the algorithm generates new iterates

$$(X^{k+1}, Y^{k+1}) = (X^k + \alpha \Delta X, Y^k + \alpha \Delta Y).$$

- **Duality Gap Update:** The search for an optimal primal-dual solution corresponds to driving $X \bullet Y$ to zero, achieved by updating the parameter μ

$$\mu^{k+1} = (1-\theta)\mu^k \quad \text{with} \quad 0 < \theta < 1.$$

The generic form of the large-update primal-dual interior point algorithm for solving SDLCP is stated as follows

Algorithm 1 Primal-Dual Interior Point Algorithm for SDLCP

Require:
- Threshold parameter $\tau \geq 1$
- Accuracy parameter $\epsilon > 0$
- Barrier update parameter $0 < \theta < 1$
- Strictly feasible pair $(X^0 \succ 0, Y^0 \succ 0)$
- Initial $\mu^0 = \text{Tr}(X^0Y^0)/n$ with $\delta(X^0, Y^0, \mu^0) \leq \tau$

```
Initialize: X ← X^0, Y ← Y^0, μ ← μ^0
while nμ ≥ ε do
    μ^{k+1} ← (1 − θ)μ^k
    while δ(X, Y, μ) > τ do
        Solve System (3.11) and use (3.12) to obtain (ΔX, ΔY)
        Determine step size α ∈ (0, 1)
        Update (X^{k+1}, Y^{k+1}) ← (X^k + αΔX, Y^k + αΔY)
    end while
end while
```

3.7 Kernel function and its qualification

In recent years, kernel functions have gained significant popularity, proving crucial and highly advantageous in various domains of mathematical programming research. A kernel function is defined as follows

Definition 3.7.1 *Let $\Psi(t) : \mathbb{R}_{++} \to \mathbb{R}_+$ be a twice continuously differentiable function. Then Ψ is called a kernel function if it satisfies the following conditions*

1. $\Psi'(1) = \Psi(1) = 0$,
2. $\Psi''(t) > 0$,

3. $\lim_{t\to 0^+} \Psi(t) = \lim_{t\to+\infty} \Psi(t) = +\infty$.

The first two conditions imply that Ψ is strictly convex and attains its minimum at $t = 1$ with $\Psi(1) = 0$. This allows $\Psi(t)$ to be expressed as

$$\Psi(t) = \int_1^t \left(\int_1^x \Psi''(y)\, dy \right) dx. \tag{3.1}$$

Condition (3) indicates that Ψ acts as a barrier function.

Lemma 3.7.1 *Let $\Psi(t)$ be a twice-differentiable function, then the following properties are equivalent*

1. $\Psi(\sqrt{t_1 t_2}) \leq \frac{\Psi(t_1)+\Psi(t_2)}{2}$, *for all* $t_1, t_2 > 0$.
2. $t\Psi"(t) + \Psi\prime(t) \geq 0, t > 0$.
3. $\Psi(e^{\epsilon})$ *is convex.*

In the following table, we provide the different known kernel functions in the literature, and the complexity of their algorithm for small- and large-step interior point methods

	Kernel function	algorithmic complexity	algorithmic complexity
	$\Psi_i(t)$	Small step	large step
1	$\frac{1}{2}(t^2-1)-logt$	$O(\sqrt{n}\log\frac{n}{\epsilon})$	$O(n\log\frac{n}{\epsilon})$
2	$\frac{t^2-1}{2}+\frac{t^{1-q}-1}{q(q-1)}-\frac{q-1}{q(q-1)}(t-1), q>1$	$O(q\sqrt{n}\log\frac{n}{\epsilon})$	$O(qn^{\frac{q+1}{2q}}\log\frac{n}{\epsilon})$
3	$\frac{t^2-1}{2}+\frac{(e-1)^2}{e}\frac{1}{e-1}-\frac{e-1}{e}$	$O(\sqrt{n}\log\frac{n}{\epsilon})$	$O(n^{3/4}\log\frac{n}{\epsilon})$
4	$\frac{1}{2}(t^2-\frac{1}{t})^2$	$O(\sqrt{n}\log\frac{n}{\epsilon})$	$O(n^{2/4}\log\frac{n}{\epsilon})$
5	$\frac{t^2-1}{2}+e^{\frac{1}{t-1}}-1$	$O(\sqrt{n}\log\frac{n}{\epsilon})$	$O(\sqrt{n}(\log n)^2\log\frac{n}{\epsilon})$
6	$\frac{t^2-1}{2}-\int_1^t e^{\frac{1}{\xi}}d\xi$	$O(\sqrt{n}\log\frac{n}{\epsilon})$	$O(\sqrt{n}\log^2 n\log\frac{n}{\epsilon})$
7	$\frac{t^2-1}{2}+\frac{t^{1-q}-1}{q-1}, q>1$	$O(q^2\sqrt{n}\log\frac{n}{\epsilon})$	$O(qn^{\frac{q+1}{2q}}\log\frac{n}{\epsilon})$
8	$t-1+\frac{t^{1-q}-1}{q-1}, q>1$	$O(q^2\sqrt{n}\log\frac{n}{\epsilon})$	$O(q\log^2 n\log\frac{n}{\epsilon})$
9	$\frac{t^{1-p}-1}{p-1}-\log t, p\in[0,1]$	$O(\sqrt{n}\log\frac{n}{\epsilon})$	$O(n\log\frac{n}{\epsilon})$
10	$\frac{t^{1-p}-1}{p-1}+\frac{t^{1-q}-1}{q-1}, p\in[0,1], q>1$	$O(\sqrt{n}\log\frac{n}{\epsilon})$	$O(qn^{\frac{q+p}{q(1+p)}}\log\frac{n}{\epsilon})$
11	$t^2-1+\frac{t^{1-q}-1}{q-1}-\log t, p>1, q>1$	$O(\sqrt{n}\log\frac{n}{\epsilon})$	$O(qn^{\frac{q+1}{2q}}\log\frac{n}{\epsilon})$
12	$(m+1)t^2-(m+2)t+\frac{1}{t^m}, m>4$	$O(q^2\sqrt{n}\log\frac{n}{\epsilon})$	$O(m^{\frac{2m+1}{2m}}\log\frac{n}{\epsilon})$
13	$t^2-2t+\frac{1}{\sin(\frac{\pi t}{1+t})}, t>0$	$O(\sqrt{n}\log\frac{n}{\epsilon})$	$O(n^{3/4}\log\frac{n}{\epsilon})$
14	$\frac{t^2-1}{2}+\frac{4}{\pi p}[\tan^p h(t)-1], t>0$	$O(p^2\sqrt{n}\log\frac{n}{\epsilon})$	$O(pn^{\frac{p+2}{2(p+1)}}\log\frac{n}{\epsilon})$

Table 3.1: Different known kernel functions.

3.8 Spectral decomposition for symmetric matrices

We present results on the spectral decomposition for symmetric matrices, enabling us to extend the definition of a function $\psi:\mathbb{R}\to\mathbb{R}$ to a matrix function $\psi: S^n\to S^n$.

Theorem 3.8.1 (Spectral Theorem for Symmetric Matrices) *A real $n\times n$ matrix A is symmetric if and only if there exists an orthogonal matrix $Q\in\mathbb{R}^{n\times n}$ such that $Q^\top AQ$ is a diagonal matrix Λ containing the eigenvalues of A.*

Definition 3.8.1 *Let V be a symmetric matrix with spectral decomposition*

$$V = Q^\top \operatorname{diag}(\lambda_1(V), \lambda_2(V), \ldots, \lambda_n(V))Q,$$

where Q is an orthogonal matrix that diagonalizes V. The matrix-valued function $\psi : S^n \to S^n$ is defined by

$$\psi(V) = Q^\top \operatorname{diag}\big(\psi(\lambda_1(V)), \psi(\lambda_2(V)), \ldots, \psi(\lambda_n(V))\big)Q.$$

We note that $\psi(V)$ is well-defined when $\psi(t)$ is well-defined at each eigenvalue of V.

If $\psi(t)$ is differentiable, then its matrix derivative $\psi'(V)$ is given by

$$\psi'(V) = Q^\top \operatorname{diag}\big(\psi'(\lambda_1(V)), \psi'(\lambda_2(V)), \ldots, \psi'(\lambda_n(V))\big)Q.$$

A matrix function is a matrix $M(t)$ where each element $M_{ij}(t)$ is a function of t, forming the matrix $[M_{ij}(t)]_{i,j=1}^n$. Standard properties like continuity, differentiability, and integrability apply naturally to matrix functions, interpreted component-wise.

Let $H(t)$ and $M(t)$ be differentiable matrix functions. We have the following derivatives

$$\begin{aligned}
\frac{d}{dt}\operatorname{Tr}(\psi(H(t))) &= \operatorname{Tr}(\psi'(H(t))H'(t)),\\
\frac{d}{dt}(H(t)M(t)) &= H'(t)M(t) + H(t)M'(t),\\
\frac{d}{dt}\operatorname{Tr}(H(t)) &= \operatorname{Tr}(H'(t)),\\
\frac{d}{dt}H(t) &= \left[\frac{d}{dt}H_{ij}(t)\right]_{i,j=1}^n = H'(t).
\end{aligned}$$

For any function (t), the divided difference (t) is defined as

$$\Delta\psi(t_1, t_2) = \begin{cases} \dfrac{\psi(t_1) - \psi(t_2)}{t_1 - t_2}, & \text{when } t_1 \neq t_2,\\[2ex] \psi'(t), & \text{when } t_1 = t_2. \end{cases}$$

Lemma 3.8.1 *Suppose $H(t)$ is a matrix function that is positive definite with eigenvalues*

$\lambda_1(t) \geq \lambda_2(t) \geq \cdots \geq \lambda_n(t) > 0$. *If* $H(t)$ *is twice differentiable with respect to* $t \in (l_t, u_t)$ *and* $\psi(t)$ *is twice continuously differentiable on a domain containing all eigenvalues of* $H(t)$*, then*

$$\begin{aligned}
&\frac{d}{dt}\operatorname{Tr}(\psi(H(t))) = \operatorname{Tr}(\psi'(H(t))H'(t)),\\
&\frac{d^2}{dt^2}\operatorname{Tr}(\psi(H(t))) \leq \omega\|H'(t)\|^2 + \operatorname{Tr}(\psi'(H(t))H''(t)),
\end{aligned}$$

where

$$\omega = \max\Big\{\big|\Delta\psi'(\lambda_j(t), \lambda_k(t))\big| : t \in (l_t, u_t),\ j,k = 1,2,\ldots,n\Big\}$$

is a constant depending on $H(t)$ *and* $\psi(t)$*, with the divided difference of the derivative defined as:*

$$\Delta\psi'(t_1, t_2) = \frac{\psi'(t_1) - \psi'(t_2)}{t_1 - t_2}, \quad \forall t_1, t_2 \in [l_t, u_t],\ t_1 \neq t_2.$$

The following theorem is crucial for interior point methods based on the kernel function for SDLCP.

Theorem 3.8.2 *Let* V_1 *and* V_2 *be symmetric positive definite matrices, and let* Ψ *be the real-valued matrix function induced by the scalar function* ψ*. Then the following inequality holds:*

$$\Psi\left(\left(V_1^{1/2}V_2V_1^{1/2}\right)^{1/2}\right) \leq \frac{1}{2}\left(\Psi(V_1) + \Psi(V_2)\right).$$

3.9 The search directions determined by kernel functions

Revising the second equation in system (3.11), we substitute $-\psi'(V)$ for its right-hand side. This substitution results in a new system for determining search directions.

$$\begin{cases} \widetilde{L}(D_X) = D_Y, \\ D_X + D_Y = -\psi'(V). \end{cases} \tag{4.10}$$

Here, $\psi(t)$ is a given kernel function, and $\psi(V)$, $\psi'(V)$ are the associated matrix functions. The search directions D_X and D_Y are obtained by solving system (4.10), from which ΔX and ΔY are computed via (3.12). Note that these search directions satisfy the non-orthogonality condition

$$D_X \bullet D_Y \geq 0.$$

Using these directions with a step size α determined by appropriate line search rules, we construct new iterates

$$X_+ = X + \alpha \Delta X \quad \text{and} \quad Y_+ = Y + \alpha \Delta Y.$$

To analyze an interior-point algorithm, we define the proximity measure $\delta(V) : S^n_{++} \to \mathbb{R}_+$ as follows

$$\delta(V) = \frac{1}{2}\| - \varphi'(V)\| = \frac{1}{2}\sqrt{\mathrm{Tr}(\varphi'(V)^2)} = \frac{1}{2}\|D_x + D_r\|. \tag{4.11}$$

The value of $\Psi(V)$ $\forall(V)$, is strictly convex and attains its minimum value zero at $V = I$. Therefore, we have

$$\delta(V) = 0 \Leftrightarrow V = I \Leftrightarrow \forall(V) = 0.$$

Therefore, the value of $\Psi(V)$ provides a measure of the distance between a given iterate (X, Y) and the central path point $(X(\mu), Y(\mu))$.

3.10 The Generic Primal–Dual IPM for SDLCP

The general framework of the primal–dual kernel function-based interior-point method (IPM) for SDLCP operates as follows:

- **Initialization:** We assume the existence of a strictly feasible initial point (X^0, Y^0)

in a τ-neighborhood, i.e., $\Psi(V) \leq \tau$.

- **Outer Iteration:** The barrier parameter μ is reduced by the factor $(1-\theta)\mu$ with fixed $0 < \theta < 1$, which changes V and yields a new μ-center $(X(\mu), Y(\mu))$. Consequently, $\Psi(V)$ exceeds the threshold τ.

- **Inner Iteration:**

  - Solve the Newton system (15) using (11) to obtain the unique search direction
  - Introduce a step size to reduce $\Psi(V)$
  - Repeat until the iterate returns to the τ-neighborhood of the current μ-center

- **Termination:** The process repeats until $n\mu \leq \epsilon$ for given accuracy ϵ.

Algorithm 2 Generic Interior-Point Algorithm for SDLCP

Require:
- Threshold parameter $\tau \geq 1$
- Accuracy parameter $\epsilon > 0$
- Barrier update parameter $0 < \theta < 1$
- Strictly feasible initial pair (X^0, Y^0)
- Initial $\mu_0 = 1$ satisfying $\Psi(X^0, Y^0; \mu_0) \leq \tau$

Initialize: $X \leftarrow X^0$, $Y \leftarrow Y^0$, $\mu \leftarrow \mu_0$
while $n\mu \geq \epsilon$ **do**
  $\mu \leftarrow (1-\theta)\mu$ {Update barrier parameter}
  while $\Psi(V) > \tau$ **do**
    Solve system (15) and use (11) to obtain ΔX, ΔY
    Determine step size α
    Update: $X \leftarrow X + \alpha\Delta X$, $Y \leftarrow Y + \alpha\Delta Y$
  end while
end while

Algorithm 2 is well-defined and achieves the currently best-known iteration bound of $\mathcal{O}\left(\sqrt{n}\log n \log \frac{n}{\epsilon}\right)$ for large-update methods.

Bibliography

[1] Achache, Mohamed, and Nesrine Tabchouche. "Complexity analysis and numerical implementation of large-update interior-point methods for SDLCP based on a new parametric barrier kernel function." *Optimization* 67.8 (2018): 1211–1230.

[2] Abdessemed, Nabila, Rachid Benacer, and Naima Boudiaf. "Role of New Kernel Function in Complexity Analysis of an Interior Point Algorithm for Semidefinite Linear Complementarity Problem." *arXiv preprint arXiv:2108.07325* (2021).

[3] Cottle, Richard W., Jong-Shi Pang, and Richard E. Stone. *The Linear Complementarity Problem.* Society for Industrial and Applied Mathematics, 2009.

[4] Deklerk, E. *Interior Point Methods for Semidefinite Programming.* Thesis, Technische University Delft / University of Pretoria, 1997.

[5] Frommer, Andreas, and Valeria Simoncini. "Matrix Functions." *Model Order Reduction: Theory, Research Aspects and Applications.* Berlin, Heidelberg: Springer, 2008. 275–303.

[6] Glineur, François. "Etude des méthodes de point intérieur appliquées à la programmation linéaire et à la programmation semidéfinie." Document, https://perso.uclouvain.be/francois.glineur/oldwww/courses/inma2471, 1997.

[7] Gondzio, Jacek, and Tamás Terlaky. *A Computational View of Interior-Point Methods for Linear Programming.* Univ. de Genève/HEC, 1994.

[8] Gowda, M. Seetharama, and Yoon Song. "On Semidefinite Complementarity Problems." *Mathematical Programming, Series A* 88 (2000): 575–587.

[9] Gowda, M. Seetharama, and Yoon Song. "Some New Results for the Semidefinite Linear Complementarity Problem." *SIAM Journal on Matrix Analysis and Applications* 24.1 (2003): 25–39.

[10] Horn, Roger A., and Charles R. Johnson. *Matrix Analysis.* Cambridge University Press, 2012.

[11] Kojima, Masakazu, Masayuki Shida, and Susumu Shindoh. "Search Direction in the SDP and the Monotone SDLCP, Generalization and Inexact Computation." Research Report B-327, Dept. of Mathematical and Computing Sciences, Tokyo Institute of Technology, 1997.

[12] Malek, Abderraouf, and Debbar Khouloud. "Etude théorique sur les fonctions noyau pour la programmation linéaire." Masters Thesis, Université de Guelma, 2019.

[13] Monteiro, Renato DC. "Primal-Dual Path-Following Algorithms for Semidefinite Programming." *SIAM Journal on Optimization* 7 (1997): 663–678.

[14] Rump, Siegfried M. "Verification of Positive Definiteness." *BIT Numerical Mathematics* 46 (2006): 433–452.

[15] Shapiro, Alexander. "On Differentiability of Symmetric Matrix Valued Functions." School of Industrial and Systems Engineering, Georgia Institute of Technology, 2002.

ABSTRACT

This thesis investigates interior-point methods for solving monotone semidefinite linear complementarity problems (SDLCPs). It presents theoretical foundations of convex analysis, matrix optimization, and complementarity theory, with a focus on primal-dual interior-point algorithms based on kernel functions. The study analyzes the central path, Nesterov-Todd directions, and various kernel functions' impact on algorithmic complexity. Applications in optimization, control theory, and eigenvalue problems are discussed. The work provides complexity bounds for large and small update methods, demonstrating their polynomial-time convergence. Key results include equivalence theorems between SDLCP and related problems, as well as spectral decomposition techniques for symmetric matrices.

RESUME

Cette thèse étudie les méthodes de points intérieurs pour résoudre les problèmes de complémentarité linéaire semi-définis monotones (SDLCP). Elle présente les fondements théoriques de l'analyse convexe, de l'optimisation matricielle et de la théorie de la complémentarité, en se concentrant sur les algorithmes primaux-duaux basés sur des fonctions noyau. L'étude analyse le chemin central, les directions de Nesterov-Todd et l'impact de diverses fonctions noyau sur la complexité algorithmique. Les applications en optimisation, théorie du contrôle et problèmes aux valeurs propres sont discutées. Le travail fournit des bornes de complexité pour les méthodes à grands et petits pas, démontrant leur convergence en temps polynomial. Les résultats clés incluent des théorèmes d'équivalence entre SDLCP et problèmes connexes, ainsi que des techniques de décomposition spectrale pour matrices symétrique

ملخص

تدرس هذه الأطروحة طرق النقطة الداخلية بالنسبة لمسائل الأمثلة لحل مسائل التكميل الخطي شبه الموجبة (SDLCP). ويقدم الأسس النظرية للتحليل المحدب، وخواص ونتائج نظرية حول المصفوفات المعرفة موجبا، ونظرية التكميل، مع التركيز على خوارزميات الثنائية الأولية القائمة على وظائف النواة. تحلل الدراسة المسار المركزي واتجاهات Nesterov-Todd وتأثير وظائف النواة المختلفة على التعقيد الخوارزمي. يتم مناقشة التطبيقات في التحسين ونظرية التحكم ومشاكل القيمة الذاتية. يقدم العمل حدود التعقيد للطرق ذات الخطوات الكبيرة والصغيرة، مما يوضح تقاربها في وقت متعدد الحدود. تتضمن النتائج الرئيسية نظريات التكافؤ بين SDLCP والمشاكل ذات الصلة، بالإضافة إلى تقنيات التحلل الطيفي للمصفوفات المتماثلة.