COMPLEMENTARITY AND RELATED PROBLEMS

by

LIANGHAI XIAO

A thesis submitted to
The University of Birmingham
for the degree of
DOCTOR OF PHILOSOPHY

School of Mathematics
College of Engineering and Physical Sciences
The University of Birmingham
March 2020

UNIVERSITY OF
BIRMINGHAM

University of Birmingham Research Archive

e-theses repository

This unpublished thesis/dissertation is copyright of the author and/or third parties. The intellectual property rights of the author or third parties in respect of this work are as defined by The Copyright Designs and Patents Act 1988 or as modified by any successor legislation.

Any use made of information contained in this thesis/dissertation must be in accordance with that legislation and must be properly acknowledged. Further distribution or reproduction in any format is prohibited without the permission of the copyright holder.

Abstract

In this thesis, we present results related to complementarity problems.

We study the linear complementarity problems on extended second order cones. We convert a linear complementarity problem on an extended second order cone into a mixed complementarity problem on the non-negative orthant. We present algorithms for this problem, and exemplify it by a numerical example. Following this result, we explore the stochastic version of this linear complementarity problem. Finally, we apply complementarity problems on extended second order cones in a portfolio optimisation problem. In this application, we exploit our theoretical results to find an analytical solution to a new portfolio optimisation model.

We also study the spherical quasi-convexity of quadratic functions on spherically self-dual convex sets. We start this study by exploring the characterisations and conditions for the spherical positive orthant. We present several conditions characterising the spherical quasi-convexity of quadratic functions. Then we generalise the conditions to the spherical quasi-convexity on spherically self-dual convex sets. In particular, we highlight the case of spherical second order cones.

CONTENTS

1 Basic Concepts of Complementarity Problems **1**

1.1 Introduction 1

1.2 Preliminaries 4

1.3 Complementarity problems 8

1.4 Extended second order cone 12

1.5 Convex sets on the sphere 15

2 Linear Complementarity Problems on Extended Second Order Cones **26**

2.1 Problem formulation 26

2.2 Find the solution 34

2.2.1 Reformulate to a system of nonlinear equations 34

2.2.2 Reformulate to an unconstrained minimisation problem 38

2.3 A numerical example 47

2.4 Conclusions and comments 49

3 Stochastic Linear Complementarity Problems on Extended Second Order Cones **51**

3.1 Problem formulation 52

3.2 An algorithm 67

3.3 A numerical example 69

3.4 Conclusions and comments 71

4 Application: Portfolio Optimisation Problems **73**

4.1 Introduction to portfolio optimisation problem 73

4.1.1 Review of mean-variance model and mean-absolute deviation model . . . 74

4.1.2 Formulation of the mean-Euclidean norm model 79
4.2 Conclusions and comments 88

5 Spherically quasi-convex quadratic functions **90**
5.1 Motivation of this study 90
5.2 Spherically quasi-convex quadratic functions on spherically convex sets 93
5.3 Spherically quasi-convex quadratic functions on the spherical positive orthant . . 99
5.4 Spherically quasi-convex quadratic functions on the subdual convex sets 113
5.5 Spherically Quasi-Convex Quadratic Functions on the Spherical Lorentz Convex Set 128
5.6 Conclusions and comments 129

6 Final remarks **131**
6.1 Summary of results 131
6.2 Contributions 132

List of References **143**

LIST OF FIGURES

1.1 Second order cone and extended second order cone 13

1.2 The geodesic and the tangent hyperplane 19

1.3 Closed set $\mathcal{S}$ and the cone $\mathcal{K}_{\mathcal{S}}$ spanned by $\mathcal{S}$. 20

3.1 VaR and CVaR for $\theta(x, \omega) = \omega$, where $\omega \sim N(0, 1)$ 56

3.2 The minimum point of merit function varies $\theta_{FB}^{\text{SMixCP}}$ as $|\Omega|$ increases 61

4.1 The graph of the probability of inequality (4.24) to hold relates to its parameters 84

4.2 The graph of the probability of inequality (4.33) to hold relates to its parameters 87

LIST OF TABLES

2.1 Numerical Example: the iteration of $\{z_j\}$ 48

3.1 The result of the numerical example 71

CHAPTER 1

BASIC CONCEPTS OF COMPLEMENTARITY PROBLEMS

1.1 Introduction

The concept of complementarity is firstly introduced by Karush [53] and considered by Dantzig and Cottle in a technical report [18], for the non-negative orthant. In 1968, Cottle and Dantzig [16] connected the linear programming problem, the quadratic programming problem and the bimatrix game problem to the complementarity problem, which attracted many researchers' attentions to this field (see [2, 8, 19, 36, 70]).

As a cross-cutting problem, complementarity problem provides a powerful framework for the study of optimisation and equilibrium problems, and hence has a wide range of applications in engineering and economics. Earlier works in cone complementarity problems present the theory for a general cone and the practical applications merely for the non-negative orthant only (similarly to the books [19, 59]). These are related to equilibrium problems in economics, engineering, physics, finance and traffic. Examples in economics are Walrasian price equilibrium models, price oligopoly models, Nash-Cournot production/distribution models, models of invariant capital stock, Markov perfect equilibria, models of decentralised economy and perfect competition equilibrium, models with individual markets of production factors. Engineering and physics applications are frictional contact problems, elastoplastic structural analysis and nonlinear obstacle problems. An example in finance is the discretisation of the differential complementarity formulation of the Black-Scholes models for the American options [47]. An application to congested traffic networks is the prediction of steady-state traffic flows. In the

recent years several applications have emerged where the complementarity problems are defined by cones essentially different from the non-negative orthant such as positive semidefinite cones, second order cones and direct product of these cones (for mixed complementarity problems containing linear subspaces as well). Recent applications of second order cone complementarity problems are in elastoplasticity [105,109], robust game theory [66,85] and robotics [3]. All these applications come from the Karush-Kuhn-Tucker conditions of second order conic optimisation problems.

Németh and Zhang extended the concept of second order cone in [83] to the extended second order cone (ESOC). Their extension seems the most natural extension of second order cones. Sznajder showed that the extended second order cones in [83] are irreducible cones (i.e., they cannot be written as a direct product of simpler cones) and calculated the Lyapunov rank of these cones [99]. The applications of second order cones and the elegant way of extending them suggest that the extended second order cones will be important from both theoretical and practical point of view. Although conic optimisation problems with respect to ESOC can be reformulated as conic optimisation problems with respect to second order cones (SOC), we expect that for several such problems using the particular inner structure of the second order cones provides a more efficient way of solving them than solving the transformed conic optimisation problem with respect to second order cones. Indeed, such a particular problem is the projection onto an extended second order cone which is much easier to solve directly than solving the reformulated second order conic optimisation problem [26].

Until now the extended second order cones of Németh and Zhang were used as a working tool only for finding the solutions of mixed complementarity problems on general cones [83] and variational inequalities for cylinders whose base is a general convex set [84]. The applications above for second order cones show the importance of these cones and motivates considering conic optimization and complementarity problems on extended second order cones. In this thesis we develop an application to portfolio optimisation problems [72, 92] described in Chapter 4.

We further extend our study to the existance of the solution to a nonlinear complementarity problem. The existence of the solution to a nonlinear complementarity problem can be converted to a problem of minimising a quadratic function on the intersection between a cone and a sphere, according to [81, Theorem 18] and [45, Corollary 8.1]. We also study the spher-

ical convexity problem. It is started by exploring the characterisations and conditions for the spherical positive orthant, then it is extended to the spherical quasi-convexity on spherically self-dual convex sets.

The thesis is organised as follows: In the rest of this chapter, we illustrate the main terminologies and definitions used in this thesis. The terminologies, definitions and basic results of complementarity problem, extended second order cone,and convex sets on the sphere are in Section 1.3, Section 1.4, and Section 1.5, respectively.

In Chapter 2, we reformulate the linear complementarity problem as a mixed (implicit, mixed implicit) complementarity problem (MixCP) on the non-negative orthant. Our main result is Theorem 2.1.1, which discusses the connections between an linear complementarity problems on extended second order cones (ESOCLCP) and mixed (implicit, mixed implicit) complementarity problems on nonnegative orthant (MixCP). Based on the above, we use some algorithms to solve the MixCP. A solution to this MixCP is equivalent to a solution to the corresponding ESOCLCP. In the last section of this chapter, we provide an example of ESOCLCP corresponding to the cases in Item (iv) of Proposition 1.4.3.

In Chapter 3, we study the stochastic linear complementarity problems on extended second order cones (stochastic ESOCLCP). We first convert the problem to a stochastic mixed complementarity problem on the nonnegative orthant (SMixCP). Enlightened by the idea of Chen and Lin [15], we introduce the Conditional Value at Risk (CVaR) method to measure the loss of complementarity in the stochastic case. A CVaR - based minimisation problem is introduced to achieve a solution which is "good enough" for the complementarity requirement of the original SMixCP. Smoothing function and sample average approximation methods are introduced and the the problem is converted to a form which can be solved by Levenberg-Marquardt smoothing SAA algorithm. At the end of this chapter, a numerical example will be used to illustrates our results.

In Chapter 4, we present an application of extended second order cones to portfolio optimisation problems. Based on the mean-absolute deviation (MAD) model, we introduce the mean-Euclidean norm (MEN) model for portfolio optimisation. This new setting has advantages of low computational cost because we work out its analytical solution.

In Chapter 5, we turn to study the spherical convexity as we are motivated by the fact

that such questions are related to the existance of the solution of nonlinear complementarity problem. In Section 5.2 we characterise the quadratic spherically quasi-convexity of functions on a general spherically convex set. In Section 5.3 we study the conditions and the properties of spherically quasi-convex quadratic functions defined on the spherical positive orthant. The results of this chapter is published in our paper [28]. The results in Section 5.4 are based on the previous sections. It provides derivations of many useful properties of spherically quasi-convex functions on spherically subdual convex sets. In particular, the spherical positive orthant studied in Section 5.3 is also a spherically subdual convex set. The results obtained on subdual convex sets are applied on the second order cone (Lorentz cone) in Section 5.5. We also prove a condition partially characterising the spherical quasi-convexity of quadratic functions on spherically convex sets associated to the second order cone.

We conclude the study in this thesis by making final remarks in Chapter 6.

1.2 Preliminaries

In this section, the notations and auxiliary results used throughout this thesis will be presented. Let us start with the definitions of inner products and cones. For $\mathbb{R}^n$, the Euclidian space whose elements are column vectors, the definition of the canonical inner product $\langle \cdot, \cdot \rangle$ is given by

$$\langle x, y \rangle = \sum_{i=1}^{n} x_i y_i, \quad x, y \in \mathbb{R}^n,$$

and the definition of norm $\| \cdot \|$ is given by

$$\|x\| = \sqrt{\langle x, x \rangle}.$$

Denote by $\mathbb{R}^n_+$ the nonnegative orthant and by $\mathbb{R}^n_{++}$ the positive orthant , that is,

$$\mathbb{R}^n_+ = \{x = (x_1, \ldots, x_n)^\top : x_1 \geq 0, \ldots, x_n \geq 0\},$$

and

$$\mathbb{R}^n_{++} = \{x = (x_1, \ldots, x_n)^\top : x_1 > 0, \ldots, x_n > 0\}.$$

Denote by $\mathcal{L}$ the second order cone (Lorentz cone)

$$\mathcal{L} := \left\{ x = (x_1, \ldots, x_n)^\top \in \mathbb{R}^n : x_1 \geq \sqrt{x_2^2 + \cdots + x_n^2} \right\}. \tag{1.1}$$

It should be noted that the Lorentz cone $\mathcal{L}$, the nonnegative orthant and the positive orthant are self-dual cones.

Let k, l be positive integers. The inner product of pairs of vectors $\binom{x}{u}$, $\binom{y}{v} \in \mathbb{R}^k \times \mathbb{R}^\ell$, where $x \in \mathbb{R}^k$ and $u \in \mathbb{R}^\ell$, is defined by

$$\langle \binom{x}{u}, \binom{y}{v} \rangle = \langle x, y \rangle + \langle u, v \rangle.$$

Let $\mathbb{R}^n$ be a Euclidian space. A set $\mathcal{K} \subseteq \mathbb{R}^n$ is called a *convex cone* if for any $\alpha, \beta > 0$, and $x, y \in \mathcal{K}$, we have

$$\alpha x + \beta y \in \mathcal{K}.$$

In other words, a convex cone is a set which is invariant under multiplication of vectors with positive scalars and addition of vectors. The *dual cone* of cone $\mathcal{K} \subseteq \mathbb{R}^n$ is the convex cone $\mathcal{K}^* := \{x \in \mathbb{R}^n : \langle x, y \rangle \geq 0, \ \forall\, y \in \mathcal{K}\}$. A convex cone $\mathcal{K} \subseteq \mathbb{R}^n$ is called *pointed* if $\mathcal{K} \cap \{-\mathcal{K}\} \subseteq \{0\}$, or equivalently, if $\mathcal{K}$ does not contain straight lines through the origin. A convex cone which is a closed set is called a *closed convex cone*. Any pointed closed convex cone with nonempty interior will be called *proper cone*. The cone $\mathcal{K}$ is called *subdual* if $\mathcal{K} \subseteq \mathcal{K}^*$, *superdual* if $\mathcal{K}^* \subseteq \mathcal{K}$, and *self-dual* if $\mathcal{K}^* = \mathcal{K}$.

The matrix I_n denotes the $n \times n$ identity matrix. We denote by $\mathbb{R}^{k \times \ell}$ the set of matrices with k rows and ℓ columns with real elements. In particular $\mathbb{R}^k$ can be identified with $\mathbb{R}^{k \times 1}$.

Let $x \in \mathbb{R}^n$, then the *projection* $\mathrm{P}_{\mathcal{K}}(x)$ of the point x onto the cone $\mathcal{K}$ is defined by

$$\mathrm{P}_{\mathcal{K}}(x) := \mathrm{argmin}_y \{\|x - y\| : y \in \mathcal{K}\}.$$

For any $x \in \mathcal{K}$, we define the nonnegative part of x, nonpositive part of x and the absolute

value of x with respect to $\mathcal{K}$ by

$$x_+^{\mathcal{K}} := \mathrm{P}_{\mathcal{K}}(x), \qquad x_-^{\mathcal{K}} := \mathrm{P}_{\mathcal{K}^*}(-x), \qquad |x|^{\mathcal{K}} := x_+^{\mathcal{K}} + x_-^{\mathcal{K}}, \tag{1.2}$$

respectively. We recall from Moreau's decomposition theorem [79] (see also [41, Theorem 3.2.5]), that for a closed convex cone $\mathcal{K}$ there hold:

$$x = x_+^{\mathcal{K}} - x_-^{\mathcal{K}}, \qquad \left\langle x_+^{\mathcal{K}}, x_-^{\mathcal{K}} \right\rangle = 0, \qquad x \in \mathbb{R}^n. \tag{1.3}$$

For any $z \in \mathbb{R} \times \mathbb{R}^{n-1}$, let $z := (z_1, z^{(2)}) \in \mathbb{R} \times \mathbb{R}^{n-1}$, where $z^{(2)} := (z_2, z_3, \ldots, z_n)^\top$. An explicit formula for the projection mapping $\mathrm{P}_{\mathcal{L}}$ onto the Lorentz cone $\mathcal{L}$ is given in [33, Proposition 3.3], which is recalled for the case when $x \notin \mathcal{L} \cup -\mathcal{L}$ in the following lemma.

Lemma 1.2.1

Let $x = (x_1, x^{(2)}) \in \{(y_1, y^{(2)}) \in \mathbb{R} \times \mathbb{R}^{n-1} : \ |y_1| < \|y^{(2)}\|\}$ and $\mathcal{L}$ be the Lorentz cone. Then,

$$x_+^{\mathcal{L}} = \left(\frac{x_1 + \|x^{(2)}\|}{2\|x^{(2)}\|} \right) \left(\|x^{(2)}\|, x^{(2)} \right), \qquad x_-^{\mathcal{L}} = \left(\frac{-x_1 + \|x^{(2)}\|}{2\|x^{(2)}\|} \right) \left(\|x^{(2)}\|, -x^{(2)} \right)$$

and, as a consequence, the absolute value of x with respect to $\mathcal{L}$ is given by

$$|x|^{\mathcal{L}} = \frac{1}{\|x^{(2)}\|} \left(\|x^{(2)}\|^2, x_1 x^{(2)} \right).$$

For a general nonzero vector $x = \left(x_1, x^{(2)}\right) \in \mathbb{R} \times \mathbb{R}^{n-1}$ the absolute value of x with respect to $\mathcal{L}$ is given in the next lemma, which follows immediately from Lemma 1.2.1 and equations (1.3).

Lemma 1.2.2

Consider a nonzero vector $x = \left(x_1, x^{(2)}\right) \in \mathbb{R} \times \mathbb{R}^{n-1}$ and let $\mathcal{L}$ be the Lorentz cone. Then, the absolute value of x is given by

$$|x|^{\mathcal{L}} = \frac{1}{\|x^{(2)}\|} \Big(\max\left(|x_1|, \|x^{(2)}\|\right) \|x^{(2)}\|, \ \min(|x_1|, \|x^{(2)}\|) \operatorname{sgn}(x_1) x^{(2)} \Big),$$

where $\operatorname{sgn}(x_1)$ is equal to -1, 0 or 1 whenever x_1 is negative, zero or positive, respectively.

Definition 1.2.1 (S_0 matrix)

A matrix $A \in \mathbb{R}^{n\times n}$ is said to be an S_0 matrix if there exists a vector $x \in \mathbb{R}^n_+$ such that

$$Ax \geq 0.$$

Definition 1.2.2 (P_0 matrix)

[104, Definition 2.2] *A matrix $A \in \mathbb{R}^{n\times n}$ is said to be a P_0 matrix if every principal minor of A is non-negative. In particular, if every principal minor of A is positive, A is said to be a P matrix.*

Definition 1.2.3 (Schur complement)

[106] *The Schur complement for a matrix $M = \left(\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right)$ in nonsingular matrix D is*

$$(M/D) = A - BD^{-1}C.$$

In this study, both smooth function and semi-smooth function will be carefully reviewed. We will introduce relevant concepts about both continuously differentiable and Lipschitz continuous.

Definition 1.2.4 (Fréchet differentiable)

Let $\mathcal{K}$ be an open subset with $\mathcal{K} \subseteq \mathbb{R}^\ell$ and $f : \mathbb{R}^\ell \supseteq \mathcal{K} \to \mathbb{R}^k$, k is not necessary different from ℓ. We say that f is a differentiable function on $\mathcal{K}$, if there is a linear map $J : \mathcal{K} \to \mathbb{R}^k$ such that

$$\lim_{t\to 0} \frac{\|f(x+t) - f(x) - J(t)\|}{\|t\|} = 0,$$

for any $x \in \mathcal{K}$.

Definition 1.2.5 (Continuously differentiable function)

Let $\mathcal{K}$ be an open subset with $\mathcal{K} \subseteq \mathbb{R}^\ell$ and $f : \mathbb{R}^\ell \supseteq \mathcal{K} \to \mathbb{R}^k$, k is not necessary different from ℓ. We say that f is a continuously differentiable function on $\mathcal{K}$, if there is a linear map $J : \mathcal{K} \to \mathbb{R}^k$

$$\lim_{t\to 0} \frac{\|f(x+t) - f(x) - J(t)\|}{\|t\|} = 0,$$

such that, for any $x \in \mathcal{K}$ the map J is continuous.

Definition 1.2.6 (Lipschitz continuous function)

[97, Definition 4.6.2]

(i) Let $\mathcal{K}$ be an open subset with $\mathcal{K} \subseteq \mathbb{R}^\ell$ and $f : \mathbb{R}^\ell \supseteq \mathcal{K} \to \mathbb{R}^k$, k is not necessary different from ℓ. We say that f is a Lipschitz continuous function on $\mathcal{K}$, if there is a constant $\lambda > 0$ such that

$$\|f(x) - f(x')\| \leq \lambda \|x - x'\| \quad \forall x, x' \in \mathcal{K} \tag{1.4}$$

(ii) We say that f is locally Lipschitz continuous if for any $x \in \mathcal{K}$, there exists $\epsilon > 0$ such that f is Lipschitz on $\mathcal{K} \cap \bar{\boldsymbol{B}}(x, \epsilon)$, where $\bar{\boldsymbol{B}}(x, \epsilon) := \left\{ y \in \mathbb{R}^k : \|x - y\| \leq \epsilon \right\}$ is the closed ball centered at x.

Definition 1.2.7 (Semismooth function)

[78, Definition 1] A function $f(x)$ is semismooth at $x \in \mathbb{R}^n$ if

(i) $f(x)$ is a Lipschitz function on $\bar{\boldsymbol{B}}(x, \epsilon)$, the closed ball centered at x, and

(ii) for each $d \in \mathbb{R}^n$ and for any sequences $\{t_m\} \subseteq \mathbb{R}_+$, $\{\epsilon_m\} \subseteq \mathbb{R}^n$ and $\{g_m\} \subseteq \mathbb{R}^n$ such that $\{t_m\} \downarrow 0$, $\{\frac{\epsilon_m}{t_m}\} \to 0 \in \mathbb{R}^n$ and $g_m \in \partial f(x + t_m d + \epsilon_m)$, the sequence $\{\langle g_m, d \rangle\}$ has exactly one accumulation point.

1.3 Complementarity problems

This section briefly overviews the terminologies and definitions of complementarity problems before needed in following chapters. Some basic results about complementarity problems will be presented.

Definition 1.3.1 (Complementarity set)

Let $\mathcal{K} \subseteq \mathbb{R}^m$ be a nonempty closed convex cone and $\mathcal{K}^*$ its dual. The set

$$\mathcal{C}(\mathcal{K}) := \{(x, y) \in \mathcal{K} \times \mathcal{K}^* : \langle x, y \rangle = 0\}$$

is called the complementarity set *of cone $\mathcal{K}$.*

Definition 1.3.2 (Complementarity function)

A function $\phi(a, b)$ *is called* complementarity function (C-function) *if it satisfies:*

$$\phi(a, b) = 0 \quad \Leftrightarrow \quad a \geq 0, \quad b \geq 0, \quad ab = 0.$$

Definition 1.3.3 (Variational inequalities)

Let $\mathcal{K} \subseteq \mathbb{R}^n$ *be a nonempty closed convex cone and* $F : \mathbb{R}^n \to \mathbb{R}^n$ *be a mapping. The* variational inequality (VI) *defined by* F *and* $\mathcal{K}$ *is the problem*

$$VI(F, \mathcal{K}) \begin{cases} Find\ x \in \mathbb{R}^n,\ such\ that \\ \langle y - x, F(x) \rangle \geq 0, \quad \forall y \in \mathcal{K}. \end{cases} \tag{1.5}$$

Definition 1.3.4 (Complementarity problem)

Let $F : \mathbb{R}^n \to \mathbb{R}^n$ *be a mapping. Let* $\mathcal{K} \subseteq \mathbb{R}^n$ *be a nonempty closed convex cone and* $\mathcal{K}^*$ *its dual. Defined by* $\mathcal{K}$ *and* F *the* complementarity problem (CP) *is:*

$$CP(F, \mathcal{K}) \begin{cases} Find\ x \in \mathbb{R}^n,\ such\ that \\ (x, F(x)) \in \mathcal{C}(\mathcal{K}). \end{cases} \tag{1.6}$$

The solution set of $\mathrm{CP}(F, \mathcal{K})$ *is denoted by* $\mathrm{SOL\text{-}CP}(F, \mathcal{K})$*:*

$$\mathrm{SOL\text{-}CP}(F, \mathcal{K}) = \{x \in \mathbb{R}^n : (x, F(x)) \in \mathcal{C}(\mathcal{K})\}.$$

In particular, the definition of the *linear complementarity problem (LCP)* is:

$$LCP(F, \mathcal{K}) \begin{cases} Find\ x \in \mathbb{R}^n,\ such\ that \\ (x, F(x)) \in \mathcal{C}(\mathcal{K}). \end{cases}$$

where $F(x)$ defined by $F(x) = Tx + r$ is a linear function, where $T \in \mathbb{R}^{n \times n}$ is a matrix and $r \in \mathbb{R}^n$. The solution set of $\mathrm{LCP}(T, r, \mathcal{K})$ is denoted by $\mathrm{SOL\text{-}LCP}(T, r, \mathcal{K})$.

Definition 1.3.5 (Implicit complementarity problem)

Let $G, F : \mathbb{R}^n \to \mathbb{R}^n$ be mappings. The implicit complementarity problem (ICP) defined by G, F, and the cone $\mathcal{K}$ is:

$$\mathrm{ICP}(G, F, \mathcal{K}) \begin{cases} \textit{Find } x \in \mathbb{R}^n, \textit{ such that} \\ (G(x), F(x)) \in \mathcal{C}(\mathcal{K}). \end{cases}$$

The solution set of $\mathrm{ICP}(G, F, \mathcal{K})$ *is denoted by* $\text{SOL-ICP}(G, F, \mathcal{K})$*:*

$$\text{SOL-ICP}(G, F, \mathcal{K}) = \{x \in \mathbb{R}^n : (G(x), F(x)) \in \mathcal{C}(\mathcal{K})\}.$$

Definition 1.3.6 (Mixed complementarity problem)

Consider the mappings $F_1 : \mathbb{R}^k \times \mathbb{R}^\ell \to \mathbb{R}^k$ and $F_2 : \mathbb{R}^k \times \mathbb{R}^\ell \to \mathbb{R}^\ell$. Let $\mathcal{S} \in \mathbb{R}^k$ be a nonempty closed convex cone. The mixed complementarity problem (MixCP) defined by F_1, F_2, and $\mathcal{S}$ is:

$$\mathrm{MixCP}(F_1, F_2, \mathcal{S}) : \begin{cases} \textit{Find } \binom{x}{u} \in \mathbb{R}^k \times \mathbb{R}^\ell, \textit{ such that} \\ F_2(x, u) = 0, \textit{ and } (x, F_1(x, u)) \in \mathcal{C}(\mathcal{S}). \end{cases}$$

The solution set of $\mathrm{MixCP}(F_1, F_2, \mathcal{S})$ *is denoted by* $\text{SOL-MixCP}(F_1, F_2, \mathcal{S})$*:*

$$\text{SOL-MixCP}(F_1, F_2, \mathcal{S}) = \{\binom{x}{u} \in \mathbb{R}^k \times \mathbb{R}^\ell : F_2(x, u) = 0, (x, F_1(x, u)) \in \mathcal{C}(\mathcal{S})\}.$$

The mixed complementarity problem is one of the most important problem formulations in mathematical programming. Many well-studied optimisation problems can be converted into a mixed complementarity problem.

Definition 1.3.7 (Mixed implicit complementarity problem)

Consider the mappings F_1, $G_1 : \mathbb{R}^k \times \mathbb{R}^\ell \to \mathbb{R}^k$, $F_2 : \mathbb{R}^k \times \mathbb{R}^\ell \to \mathbb{R}^\ell$, and a proper cone $\mathcal{S} \in \mathbb{R}^k$. The mixed implicit complementarity problem (MixICP) defined by F_1, F_2, G_1, and $\mathcal{S}$ is

$$\mathrm{MixICP}(F_1, F_2, G_1, \mathcal{S}) : \begin{cases} \textit{Find } \binom{x}{u} \in \mathbb{R}^k \times \mathbb{R}^\ell, \textit{ such that} \\ F_2(x, u) = 0, \textit{ and } (G_1(x, u), F_1(x, u)) \in \mathcal{C}(\mathcal{S}). \end{cases}$$

The solution set of MixICP$(F_1, F_2, G_1, \mathcal{S})$ *is denoted by* SOL-MixICP$(F_1, F_2, G_1, \mathcal{S})$*:*

$$\text{SOL-MixICP}(F_1, F_2, G_1, \mathcal{S}) = \\ \{\binom{x}{u} \in \mathbb{R}^k \times \mathbb{R}^\ell : F_2(x,u) = 0, (G_1(x,u), F_1(x,u)) \in \mathcal{C}(\mathcal{S})\}.$$

Based on the definitions above, we get the following propositions straightforwardly.

Proposition 1.3.1

Let n, k, ℓ be nonnegative integers such that $n = k + l$, $\mathcal{S} \in \mathbb{R}^k$ be a nonempty closed convex cone and $\mathcal{K} = \mathcal{S} \times \mathbb{R}^\ell$. Denote by $\mathcal{S}^*$ the dual of $\mathcal{S}$ in $\mathbb{R}^k$ and by $\mathcal{K}^* \subset \mathbb{R}^k \times \mathbb{R}^\ell$ the dual of $\mathcal{K}$ in $\mathbb{R}^k \times \mathbb{R}^\ell$. Consider the mappings $F_1 : \mathbb{R}^k \times \mathbb{R}^\ell \to \mathbb{R}^k$, $F_2 : \mathbb{R}^k \times \mathbb{R}^\ell \to \mathbb{R}^\ell$. Define the mappings $F : \mathbb{R}^k \times \mathbb{R}^\ell \to \mathbb{R}^k \times \mathbb{R}^\ell$ by $F(x,u) = \begin{pmatrix} F_1(x,u) \\ F_2(x,u) \end{pmatrix}$. Then,

$$\binom{x}{u} \in \text{SOL-CP}(F, \mathcal{K}) \iff \binom{x}{u} \in \text{SOL-MixCP}(F_1, F_2, \mathcal{S}).$$

Proof:

It is easy to check that $\mathcal{K}^* = \mathcal{S}^* \times \{0\}$. The result follows immediately from $\mathcal{K}^* = \mathcal{S}^* \times \{0\}$ and the definitions of $CP(F, \mathcal{K})$ and MixCP$(F_1, F_2, \mathcal{S})$. □

Proposition 1.3.2

Let n, k, ℓ be nonnegative integers such that $n = k + l$, $\mathcal{S} \in \mathbb{R}^k$ be a nonempty closed convex cone and $\mathcal{K} = \mathcal{S} \times \mathbb{R}^\ell$. Denote by $\mathcal{S}^*$ the dual of $\mathcal{S}$ in $\mathbb{R}^k$ and by $\mathcal{K}^* \subset \mathbb{R}^k \times \mathbb{R}^\ell$ the dual of $\mathcal{K}$. Consider the mappings $F_1, G_1 : \mathbb{R}^k \times \mathbb{R}^\ell \to \mathbb{R}^k$, $F_2, G_2 : \mathbb{R}^k \times \mathbb{R}^\ell \to \mathbb{R}^\ell$. Define the mappings $F, G : \mathbb{R}^k \times \mathbb{R}^\ell \to \mathbb{R}^k \times \mathbb{R}^\ell$ by $F(x,u) = (F_1(x,u), F_2(x,u))$, $G(x,u) = (G_1(x,u), G_2(x,u))$, respectively. Then,

$$\binom{x}{u} \in \text{SOL-ICP}(F, G, \mathcal{K}) \iff \binom{x}{u} \in \text{SOL-MixICP}(F_1, F_2, G_1, \mathcal{S}).$$

Proof:

The result follows immediately from $\mathcal{K}^* = \mathcal{S}^* \times \{0\}$ and the definitions of ICP$(F, G, \mathcal{K})$ *and* MixICP$(F_1, F_2, G_1, \mathcal{S})$. □

1.4 Extended second order cone

The extended second order cones (ESOC), introduced by Németh and Zhang [83], are natural extensions of the second order cones (or Lorentz cones). The researches about ESOCs can be found in [26, 55, 84, 99]. The particular structure of ESOCs provides a more direct method for solving these problems without reformulations, such important examples are mixed complementarity problems on general cones [83] and variational inequalities on cylinders [55,84]. Moreover, from purely conic analysis point of view, the ESOCs cannot be trivially reduced to second order cones because the ESOCs are irreducible [99] (i.e., they cannot be written as a direct product of simpler cones). Employing the forerunners' results associated to ESOCs, this section offers some fundamental definitions and properties before needed in the later discussion of linear complementarity problems on ESOCs.

Let n, k, ℓ be nonnegative integers such that $n = k + l$. The definitions of the mutually dual extended second order cone $L(k, \ell)$ and $M(k, \ell)$ in $\mathbb{R}^n \equiv \mathbb{R}^k \times \mathbb{R}^\ell$ are:

$$L(k, \ell) = \{\left(\begin{smallmatrix} x \\ u \end{smallmatrix}\right) \in \mathbb{R}^k \times \mathbb{R}^\ell : x \geq \|u\|e\}, \tag{1.7}$$

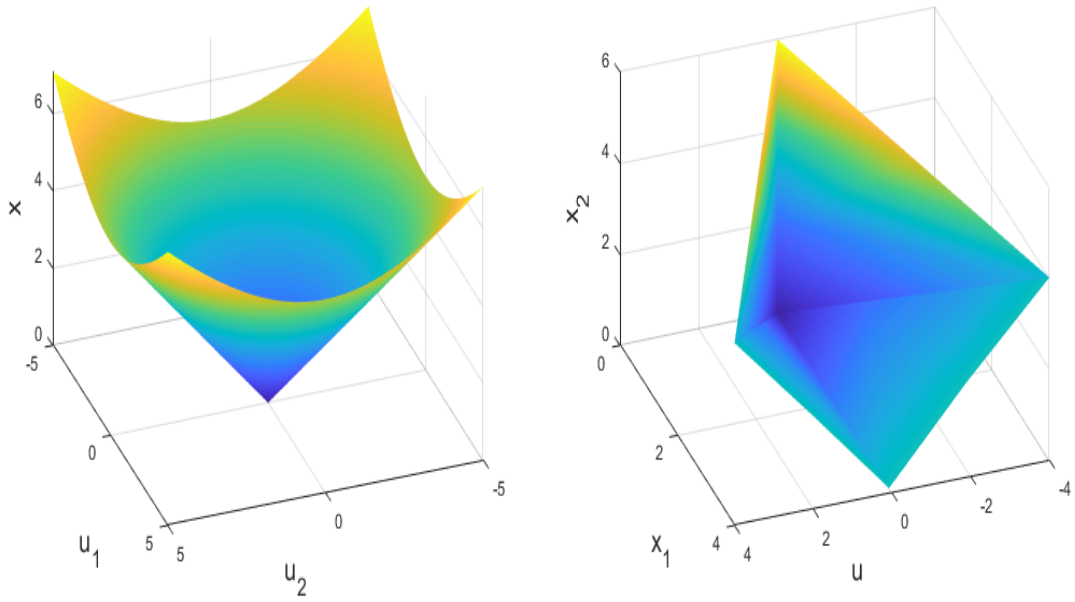
$$M(k, \ell) = \{\left(\begin{smallmatrix} x \\ u \end{smallmatrix}\right) \in \mathbb{R}^k \times \mathbb{R}^\ell : e^\top x \geq \|u\|,\ x \geq 0\}, \tag{1.8}$$

where $e = (1, \ldots, 1)^\top \in \mathbb{R}^k$. If there is no ambiguity about the dimensions, then we simply denote $L(k, \ell)$ and $M(k, \ell)$ by L and M, respectively.

We remark that when $k = 1$ the ESOC is equivalent to a second order cone (defined in (1.1) in $\mathbb{R} \times \mathbb{R}^\ell$. Hence in the following study we assume that the integer $k \geq 2$. Figure 1.1 provides the comparison of three-dimensional second order cone and three-dimensional extended second order cone. Both cones can be expressed as a pair of two vectors. Though both cones are in $\mathbb{R}^3$, the shapes of them are very different. This is because that the second order cone is in $\mathbb{R}^1 \times \mathbb{R}^2$, whilst the extended second order cone is in $\mathbb{R}^2 \times \mathbb{R}^1$.

Proposition 1.4.1

The extended second order cone is a pointed, closed convex cone with nonempty interior (henceforth it is a proper cone).



Note: This figure provides the comparison of second order cone (SOC) : $C = \{(x,u)^\top \in \mathbb{R}^{1+2} : x \geq \|u\|\}$ and extended second order cone (ESOC): $C = \{(x,u)^\top \in \mathbb{R}^{2+1} : x \geq \|u\|e\}$. It shows that the SOC is symmetric, whereas the ESOC is not symmetric. Both cones are in $\mathbb{R}^3$, but the shapes of them are very different.

Figure 1.1: Second order cone and extended second order cone

Proposition 1.4.2

Given any nonnegative integers k and ℓ with $k \geq 2$, we have $L \subset M$, and $M \not\subseteq L$.

Proposition 1.4.2 shows that an ESOC is subdual but not superdual. Hence, the ESOC is not self-dual.

Proposition 1.4.3

Let $x, y \in \mathbb{R}^k$ and $u, v \in \mathbb{R}^\ell \setminus \{0\}$.

(i) $(x,0,y,0) := (\binom{x}{0}, \binom{y}{0}) \in \mathcal{C}(L)$ *if and only if* $(x,y) \in \mathcal{C}(\mathbb{R}^k_+)$.

(ii) $(x,0,y,v) \in \mathcal{C}(L)$ *if and only if* $e^\top y \geq \|v\|$ *and* $(x,y) \in \mathcal{C}(\mathbb{R}^k_+)$.

(iii) $(x,u,y,0) \in \mathcal{C}(L)$ *if and only if* $x \geq \|u\|e$ *and* $(x,y) \in \mathcal{C}(\mathbb{R}^k_+)$.

(iv) $(x,u,y,v) \in \mathcal{C}(L)$ *if and only if there exists* $\lambda > 0$ *such that* $v = -\lambda u$, $e^\top y = \|v\|$ *and* $(x - \|u\|e, y) \in C(\mathbb{R}^k_+)$.

Proof:

Item (i) follows definition (1.7) *and* (1.8). *Let* $(x,0,y,0) \in \mathcal{C}(L)$, *it is trivial to have that* $x \geq 0$, $y \geq 0$ *and* $\langle x, y\rangle = 0$, *i.e.,* $(x,y) \in \mathcal{C}(\mathbb{R}^k_+)$.

Item (ii) follows definition (1.8). *Let* $(x, 0, y, v) \in \mathcal{C}(L)$, *naturally we have* $\langle \binom{x}{0}, \binom{y}{v} \rangle = \langle x, y \rangle + \langle 0, v \rangle = \langle x, y \rangle = 0$, *as well as* $e^\top y \geq \|v\|$ *because* $(y, v) \in M$, *so that we conclude* $(x, y) \in \mathcal{C}(\mathbb{R}^k_+)$.

Item (iii) follows definition (1.7). *Let* $(x, u, y, 0) \in \mathcal{C}(L)$, *an easy consequence is* $x \geq \|u\|e$; *and* $\langle \binom{x}{u}, \binom{y}{0} \rangle = \langle x, y \rangle = 0$, *then* $(x, y) \in \mathcal{C}(\mathbb{R}^k_+)$.

Item (iv)follows from [26, Proposition 1]. For the completeness of the results we will repeat its proof here. First assume that there exists $\lambda > 0$ *such that* $v = -\lambda u$, $e^\top y = \|v\|$ *and* $(x - \|u\|e, y) \in C(\mathbb{R}^k_+)$. *Thus,* $\binom{x}{u} \in L$ *and* $\binom{y}{v} \in M$. *On the other hand,*

$$\langle \binom{x}{u}, \binom{y}{v} \rangle = x^\top y + u^\top v = \|u\| e^\top y - \lambda \|u\|^2 = \|u\| \|v\| - \lambda \|u\|^2 = 0.$$

Thus, $(x, u, y, v) \in C(L)$.

Conversely, if $(x, u, y, v) \in C(L)$, *then* $\binom{x}{u} \in L$, $\binom{y}{v} \in M$ *and*

$$0 = \langle \binom{x}{u}, \binom{y}{v} \rangle = x^\top y + u^\top v \geq \|u\| e^\top y + u^\top v \geq \|u\| \|v\| + u^\top v \geq 0.$$

This implies the existence of a $\lambda > 0$ *such that* $v = -\lambda u$, $e^\top y = \|v\|$ *and* $(x - \|u\|e)^\top y = 0$. *It follows that* $(x - \|u\|e, y) \in C(\mathbb{R}^k_+)$. □

The following corollary generalises all the cases shown in Proposition 1.4.3:

Corollary 1.4.1

Let x, $y \in \mathbb{R}^k$, *and* u, $v \in \mathbb{R}^\ell$. *Then,* $(x, u, y, v) \in \mathcal{C}(L)$ *if and only if there exists* $\lambda > 0$ *such that*

(1) $\|u\| \|v\| (v + \lambda u) = 0$,

(2) $\|u\| (x - \|u\|e) \geq 0$,

(3) $(\|u\| \|v\|, e^\top y - \|v\|) \in \mathcal{C}(\mathbb{R}^2_+)$,

(4) $[1 - sgn(\|u\| \|v\|)] (x, y) \in C(\mathbb{R}^k_+)$, *and*

(5) $\|v\| (x - \|u\|e, y) \in \mathcal{C}(\mathbb{R}^k_+)$.

Proof:

Given the four cases in Proposition 1.4.3, we will examine each item in the corollary accordingly.

If $u = v = 0$, then the group of items (1)-(5) is equivalent to the group of items (3)-(4), because items (1)-(3) and (5) trivially hold. In turn the group of items (3)-(4) is equivalent to $(x, y) \in \mathcal{C}(\mathbb{R}^k_+)$. Hence, the result in this case follows from Proposition 1.4.3 item (i).

If $u = 0$ and $v \neq 0$, then the group of items (1)-(5) is equivalent to the group of items (3)-(5), because items (1) and (2) trivially hold. In turn the group of items (3)-(5) is equivalent to $e^\top y \geq \|v\|$ and $(x, y) \in \mathcal{C}(\mathbb{R}^k_+)$. Hence, the result in this case follows from Proposition 1.4.3 item (ii).

If $u \neq 0$ and $v = 0$, then the group of items (1)-(5) is equivalent to the group of items (2)-(4), because items (1) and (5) trivially hold. In turn the group of items (2)-(4) is equivalent to $x \geq \|u\|e$ and $(x, y) \in \mathcal{C}(\mathbb{R}^k_+)$. Hence, the result in this case follows from Proposition 1.4.3 item (iii).

If $u \neq 0$ and $v \neq 0$, then item (1) is equivalent to $v = -\lambda u$. In turn item (3) is equivalent to $e^\top y = \|v\|$, the group of items $\{$(2), (5)$\}$ is equivalent to $(x - \|u\|e, y) \in \mathcal{C}(\mathbb{R}^k_+)$ and item (4) trivially holds. In conclusion the group of items (1)-(5) is equivalent to $v = -\lambda u$ for some λ, $e^\top y = \|v\|$ and $(x - \|u\|e, y) \in \mathcal{C}(\mathbb{R}^k_+)$. Hence, the result in this case follows from Proposition 1.4.3 item (iv).

□

1.5 Convex sets on the sphere

This section gives the definitions about the convex sets on the sphere. Some results in this chapter are based on the results in [80], but we provide more explicit statements and proofs herein. We start this section with the definition of copositive matrix and Z-matrix.

Definition 1.5.1 ($\mathcal{K}$-Copositive matrix)

A matrix $A \in \mathbb{R}^{n\times n}$ is $\mathcal{K}$-copositive if

$$\langle Ax, x\rangle \geq 0$$

for any $x \in \mathcal{K}$.

Particularly, we give the following definition:

Definition 1.5.2 (Copositive matrix)

A matrix $A \in \mathbb{R}^{n\times n}$ *is copositive if*

$$\langle Ax, x\rangle \geq 0$$

for any $x \in \mathbb{R}^n_+$.

According to the two definitions above, the Definition 1.5.2 is equivalent to Definition 1.5.1 when $\mathcal{K} = \mathbb{R}^n_+$.

Let $\mathcal{K} \subseteq \mathbb{R}^n$ be a (not necessarily convex) cone. The Lorentz cone $\mathcal{L}$ can be written as

$$\mathcal{L} := \left\{x = (x_1, \ldots, x_n)^\top \in \mathbb{R}^n : \ x_1 \geq 0, \ \langle Jx, x\rangle \geq 0\right\},$$

where $J = \mathrm{diag}(1, -1, \ldots, -1) \in \mathbb{R}^{n\times n}$. It is easy to see that

$$\mathcal{L} \cup -\mathcal{L} = \left\{x = (x_1, \ldots, x_n)^\top \in \mathbb{R}^n : \langle Jx, x\rangle \geq 0\right\}.$$

This straightforwardly implies that $A \in \mathbb{R}^{n\times n}$ is $\mathcal{L}$-copositive if and only if it is $\mathcal{L} \cup -\mathcal{L}$-copositive. Hence, the S-Lemma (see [87, 103]) implies:

Lemma 1.5.1

$A \in \mathbb{R}^{n\times n}$ is $\mathcal{L}$-copositive if and only if there exist a $\rho \geq 0$ such that $A - \rho J$ is positive semidefinite.

The matrix I_n denotes the $n \times n$ identity matrix. We denote by $\mathbb{R}^{k\times \ell}$ the set of matrices with k rows and ℓ columns with real elements. In particular $\mathbb{R}^k$ can be identified with $\mathbb{R}^{k\times 1}$.

Recall that $A = (a_{ij}) \in \mathbb{R}^{n\times n}$ is *positive* if $a_{ij} > 0$ and *nonnegative* if $a_{ij} \geq 0$ for any $i, j = 1, \ldots, n$. A matrix $A \in \mathbb{R}^{n\times n}$ is *reducible* if there is permutation matrix $P \in \mathbb{R}^{n\times n}$ such that

$$P^T AP = \begin{bmatrix} B_{11} & B_{12} \\ 0 & B_{22} \end{bmatrix},$$

$$B_{11} \in \mathbb{R}^{m\times m}, \ B_{22} \in \mathbb{R}^{(n-m)\times(n-m)}, \ B_{12} \in \mathbb{R}^{m\times(n-m)}, \quad m < n.$$

A matrix $A \in \mathbb{R}^{n\times n}$ is *irreducible* if it not reducible. In the following we state a version of *Perron-Frobenius theorem* for both positive matrices and nonnegative irreducible matrices, its proof can be found in [44, Theorem 8.2.11] and [44, Theorem 8.4.4], respectively.

Theorem 1.5.1 (Perron-Frobenius Theorem)

Let $A \in \mathbb{R}^{n\times n}$ be either nonnegative and irreducible or positive. Then A has a dominant eigenvalue $\lambda_{max}(A) \in \mathbb{R}$ with associated eigenvector $v \in \mathbb{R}^n$ which satisfies the following properties:

i) The eigenvalue $\lambda_{max}(A) > 0$ and its associated eigenvector $v \in \mathbb{R}^n_{++}$;

ii) The eigenvalue $\lambda_{max}(A)$ has multiplicity one;

iii) Every other eigenvalue λ of A is less that $\lambda_{max}(A)$ in absolute value, i.e, $|\lambda| < \lambda_{max}(A)$;

iii) There are no other positive or non-negative eigenvectors of A except positive multiples of v.

Definition 1.5.3 (Z-matrix)

A matrix $A \in \mathbb{R}^{n\times n}$ is a Z-matrix if its off-diagonal elements are all nonpositive.

Definition 1.5.4 ($\mathcal{K}$-Z-property)

Let $\mathcal{K} \subseteq \mathbb{R}^n$ be a proper cone, the $\mathcal{K}$-Z-property of a matrix $A \in \mathbb{R}^{n\times n}$ means that

$$\langle Ax, y\rangle \leq 0, \qquad \forall (x, y) \in \mathcal{C}(\mathcal{K})$$

*where $\mathcal{C}(\mathcal{K})$ is the complementarity set defined as $\mathcal{C}(\mathcal{K}) := \{(x, y) \in \mathbb{R}^n \times \mathbb{R}^n : \ x \in \mathcal{K},\ y \in \mathcal{K}^*, \langle x, y\rangle = 0\}$ (see Definition 1.3.1 in **Part I**).*

The following theorem proves that when $\mathcal{K} = \mathbb{R}^n_+$, Definition 1.5.3 and Definition 1.5.4 are equivalent.

Theorem 1.5.2

The matrix $A \in \mathbb{R}^{n\times n}$ is a Z-matrix if and only if A has the $\mathbb{R}^n_+$-Z-property.

Proof:

Suppose that A has the $\mathbb{R}^n_+$-Z-property, take $x = e^i$, $y = e^j$ for any $i, j \in \{1, 2, \ldots, n\}$ with $i \neq j$, e^i and e^j are canonical vectors of $\mathbb{R}^n$. We have

$$\langle x, y\rangle = 0, \quad \langle Ae^i, e^j\rangle \leq 0,$$

which implies that $a_{ij} \leq 0$ for any $i \neq j$. Hence, A is a Z-matrix.

Conversely, suppose A is a Z-matrix, let $a = \max_i a_{ii}$, and denote $P = aI_n - A$. Then P is an entrywise nonnegative matrix. Arbitrarily take $(x, y) \in \mathcal{C}(\mathbb{R}^n_+)$, then

$$\begin{aligned}\langle Ax, y\rangle &= \langle (aI_n - P)x, y\rangle \\ &= a\langle x, y\rangle - \langle Px, y\rangle \\ &= -\langle Px, y\rangle \leq 0.\end{aligned}$$

Hence, A has the $\mathbb{R}^n_+$-Z-property.

Definition 1.5.5 (Euclidean sphere and its tangent hyperplane)

Denote by $\mathbb{S}^{n-1} := \{p = (p_1, \ldots, p_n) \in \mathbb{R}^n : \ \|p\| = 1\}$ the n-dimensional Euclidean sphere, the tangent hyperplane at point $x \in \mathbb{S}^{n-1}$ is

$$T_x\mathbb{S}^{n-1} := \left\{v \in \mathbb{R}^n : \ \langle x, v\rangle = 0, \ x \in \mathbb{S}^{n-1}\right\},$$

Definition 1.5.6 (Intrinsic distance on the sphere)

The intrinsic distance on the sphere between two arbitrary points $x, y \in \mathbb{S}^{n-1}$ is defined by

$$d(x, y) := \arccos\langle x, y\rangle. \tag{1.9}$$

By definition 1.5.6, it can be shown that the space $(\mathbb{S}^{n-1}, d)$ is a complete metric space, so that $d(x, y) \geq 0$ for any $x, y \in \mathbb{S}^{n-1}$, and $d(x, y) = 0$ if and only if $x = y$. It is also easy to check that $d(x, y) \leq \pi$ for any $x, y \in \mathbb{S}^{n-1}$, and $d(x, y) = \pi$ if and only if $x = -y$.

A mapping $\gamma : [x, y] \to \mathbb{S}^{n-1}$ is called a *geodesic*, if it is the intersection curve of a plane through the origin of $\mathbb{R}^n$ with the sphere $\mathbb{S}^{n-1}$. For any $x, y \in \mathbb{S}^{n-1}$ such that $y \neq x$ and $y \neq -x$,

there exists a unique segment of minimal geodesic from x to y. The definition of the *minimal geodesic* is as follows:

Definition 1.5.7 (Minimal geodesic)

For any $x, y \in \mathbb{S}^{n-1}$ such that $y \neq x$ and $y \neq -x$, the unique segment of minimal geodesic from to x to y is

$$\gamma_{xy}(t) = \left(\cos(td(x,y)) - \frac{\langle x, y\rangle \sin(td(x,y))}{\sqrt{1-\langle x, y\rangle^2}}\right) x + \frac{\sin(td(x,y))}{\sqrt{1-\langle x, y\rangle^2}}\, y, \qquad t \in [0,\ 1]. \quad (1.10)$$

In particular, let $x \in \mathbb{S}^{n-1}$ and $v \in T_x\mathbb{S}^{n-1}$ such that $\|v\| = 1$. The minimal geodesic connecting x to $-x$, starting at x with velocity v at x is given by

$$\gamma_{x\{-x\}}(t) := \cos(\pi t)\, x + \sin(\pi t)\, v, \qquad t \in [0,\ 1]. \quad (1.11)$$

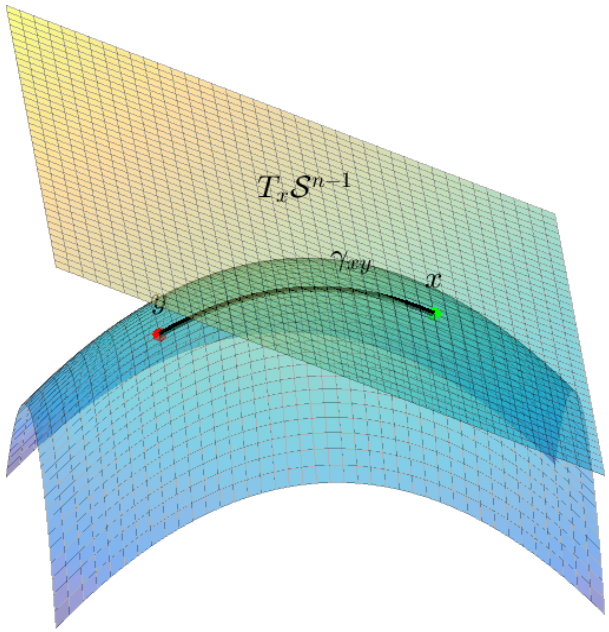


Figure 1.2: The geodesic and the tangent hyperplane

Definition 1.5.8 (Gradient on the sphere)

Let $\mathcal{S} \subseteq \mathbb{S}^{n-1}$ be a spherically open set (i.e., a set open with respect to the induced topology in $\mathbb{S}^{n-1}$). The gradient on the sphere of a differentiable function $f : \mathcal{S} \to \mathbb{R}$ at a point $x \in \mathcal{S}$ is the vector defined by

$$\operatorname{grad} f(x) := \left[I_n - xx^T\right] Df(x) = Df(x) - \langle Df(x), x\rangle\, x, \quad (1.12)$$

where $Df(x) \in \mathbb{R}^n$ *is the usual gradient of* f *at* $x \in \mathcal{S}$.

Let $\mathcal{D} \subseteq \mathbb{R}^n$ be an open set, $I \subseteq \mathbb{R}$ an open interval, $\mathcal{S} \subseteq \mathbb{S}^{n-1}$ a spherically open set and $\gamma : I \to \mathcal{S}$ a geodesic segment. If $f : \mathcal{D} \to \mathbb{R}$ is a differentiable function, then, since $\gamma'(t) \in T_{\gamma(t)}\mathbb{S}^{n-1}$ for any $t \in I$, we have $\langle \gamma'(t), \gamma(t) \rangle = 0$. The equality (1.12) implies

$$\frac{d}{dt} f(\gamma(t)) = \langle \operatorname{grad} f(\gamma(t)), \gamma'(t) \rangle = \langle Df(\gamma(t)), \gamma'(t) \rangle \,, \qquad \forall\, t \in I. \tag{1.13}$$

Definition 1.5.9 (Spherically convex set)

The set $\mathcal{S} \subseteq \mathbb{S}^{n-1}$ *is said to be* spherically convex *if for any* x, $y \in \mathcal{S}$, *the minimal geodesic segments from* x *to* y *are contained in* $\mathcal{S}$.

Example 1.5.1

The set $S_+ = \{(x_1, \ldots, x_n) \in \mathbb{S}^{n-1} : x_1 \geq 0, \ldots, x_n \geq 0\}$ *is a closed spherically convex set.*

We assume for convenience that from now on all spherically convex sets are *nonempty proper subsets of the sphere*. For any set $\mathcal{S} \subseteq \mathbb{S}^{n-1}$, we define $\mathcal{K}_\mathcal{S} \subseteq \mathbb{R}^n$ the *cone spanned by* $\mathcal{S}$, namely,

$$\mathcal{K}_\mathcal{S} := \{tx : x \in \mathcal{S},\ t \in [0, +\infty)\} \,. \tag{1.14}$$

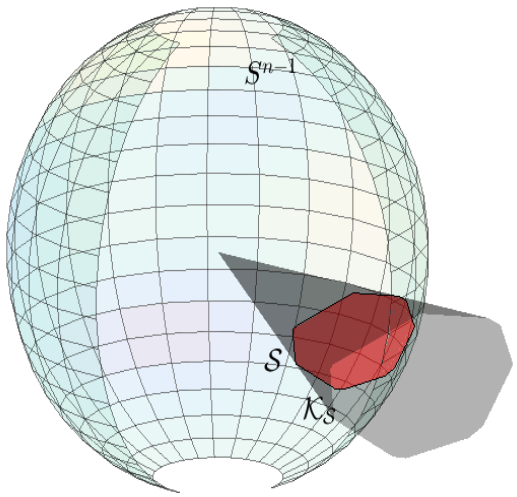


Figure 1.3: Closed set $\mathcal{S}$ and the cone $\mathcal{K}_\mathcal{S}$ spanned by $\mathcal{S}$.

Clearly, $\mathcal{K}_\mathcal{S}$ is the smallest closed cone which contains $\mathcal{S}$. The following proposition exhibits a relationship of spherically convex sets with the cones spanned by them.

Proposition 1.5.1

[24, Proposition 2] The set $\mathcal{S}$ is spherically convex if and only if the cone $\mathcal{K}_{\mathcal{S}}$ is convex and pointed.

Definition 1.5.10 (Spherically convex function)

Let $\mathcal{S} \subseteq \mathbb{S}^{n-1}$ be a spherically convex set. A function $f : \mathcal{S} \to \mathbb{R}$ is said to be (strictly) spherically convex *if for any minimal geodesic segment $\gamma : [0,1] \to \mathcal{S}$, the composition $f \circ \gamma : [0,1] \to \mathbb{R}$ is (strictly) convex in the usual sense.*

For a spherically convex set $\mathcal{S} \subseteq \mathbb{S}^{n-1}$, the *sub-level sets* of a function $f : \mathcal{S} \to \mathbb{R}$ are denoted by

$$[f \leq c] := \{x \in \mathcal{S} : \ f(x) \leq c\}, \qquad c \in \mathbb{R}. \tag{1.15}$$

Definition 1.5.11 (spherically quasi-convex function)

Let $\mathcal{S} \subseteq \mathbb{S}^{n-1}$ be a spherically convex set. A function $f : \mathcal{S} \to \mathbb{R}$ is said to be spherically quasi-convex (respectively, strictly spherically quasi-convex) if for any minimal geodesic segment $\gamma : [0,1] \to \mathcal{S}$, the composition $f \circ \gamma : [0,1] \to \mathbb{R}$ is quasi-convex (respectively, strictly quasi-convex) in the usual sense, i.e., $f(\gamma(t)) \leq \max\{f(\gamma(0)), f(\gamma(1))\}$ for any $t \in [0,1]$, (respectively, $f(\gamma(t)) < \max\{f(\gamma(0)), f(\gamma(1))\}$ for any $t \in [0,1]$).

From the above definition, it follows that (strictly) spherically convex functions are (strictly) spherically quasi-convex, but the converse is not true. It is worth to remark that the quasi-convexity concept generalises the convexity one, which was extensively studied in [25].

Proposition 1.5.2

Let $\mathcal{S} \subseteq \mathbb{S}^{n-1}$ be a spherically convex set. If function $f : \mathcal{S} \to \mathbb{R}$ is (strictly) spherically convex, then it is (strictly) spherical quasi-convex.

Proof:

We just prove spherically convex $\Rightarrow$ spherical quasi-convex. The strict case is omitted. Suppose function $f : \mathcal{S} \to \mathbb{R}$ is spherically convex, for any t_1, $t_2 \in [0,\ 1]$, we have

$$f(\gamma(\lambda t_1 + (1-\lambda)t_2)) \leq \lambda f(\gamma(t_1)) + (1-\lambda) f(\gamma(t_2))$$

for any $\lambda \in [0,\ 1]$. *Let* $t_1 = 0$, $t_2 = 1$, *we have*

$$f(\gamma(1-\lambda)) \leq \lambda f(\gamma(0)) + (1-\lambda) f(\gamma(1)) \leq \max\{f(\gamma(0)), f(\gamma(1))\}$$

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

For subsequent use only, we denote the *spherically open ball* with radius $\delta > 0$ and center in $x \in \mathbb{S}^{n-1}$ by

$$\mathbf{B}(x, \delta) := \{y \in \mathbb{S}^{n-1} : d(x, y) < \delta\},$$

and the *spherically closed ball* with radius $\delta > 0$ and center in $x \in \mathbb{S}^{n-1}$ by

$$\bar{\mathbf{B}}(x, \delta) := \{y \in \mathbb{S}^{n-1} : d(x, y) \leq \delta\}.$$

Proposition 1.5.3

Let $\mathcal{S} \subseteq \mathbb{S}^{n-1}$ *be a spherically convex set. A function* $f : \mathcal{S} \to \mathbb{R}$ *is spherically quasi-convex if and only if the sub-level sets* $[f \leq c]$ *(defined in* (1.15)*) are spherically convex for any* $c \in \mathbb{R}$.

Proof:

Suppose that f *is spherically quasi-convex and* $c \in \mathbb{R}$. *Arbitrarily take* $x, y \in [f \leq c]$, *and let* $\gamma_{xy} : [0, 1] \to \mathbb{S}^{n-1}$ *be the minimal geodesic from* x *to* y. *By* (1.10), *we have* $\gamma_{xy}(0) = x$ *and* $\gamma_{xy}(1) = y$. *Since* f *is a spherically quasi-convex function and* $x, y \in [f \leq c]$ *we have*

$$f(\gamma_{xy}(t)) \leq \max\{f(\gamma_{xy}(0)), f(\gamma_{xy}(1))\} \leq \max\{f(x), f(y)\} \leq c,$$

for any $t \in [0, 1]$, *which implies that* $\gamma(t) \in [f \leq c]$ *for any* $t \in [0, 1]$. *Hence we conclude that* $[f \leq c]$ *is a spherically convex set, for any* $c \in \mathbb{R}$.

Conversely, suppose that for any $c \in \mathbb{R}$, *the set* $[f \leq c]$ *is spherically convex. For any* c *with* $f(x) \leq c$ *and* $f(y) \leq c$, *we have* $x, y \in [f \leq c]$. *Without loss of generality, suppose* $c = \max\{f(x), f(y)\}$. *Let* $\gamma_{xy} : [0, 1] \to [f \leq c]$ *be the minimal geodesic from* x *to* y. *By Definition 1.5.9, we have* $\gamma_{xy}(t) \in [f \leq c]$ *for any* $t \in [0, 1]$, *which implies*

$$f(\gamma_{xy}(t)) \leq c = \max\{f(x), f(y)\} = \max\{f(\gamma(0)), f(\gamma(1))\},$$

for any $t \in [0,1]$. *Therefore,* f *is a spherically quasi-convex function.* □

Proposition 1.5.4

Let $\mathcal{S} \subseteq \mathbb{S}^{n-1}$ *be a spherically convex set and* $f : \mathcal{S} \to \mathbb{R}$ *be a spherically quasi-convex function. If* $x^* \in \mathcal{S}$ *is a strict local minimiser of* f*, then* x^* *is also a strict global minimiser of* f *in* $\mathcal{S}$.

Proof:

Suppose that there exists x^* *is a strict local minimiser of* f*, then there exists a scalar* $\delta > 0$ *such that*

$$f(x) > f(x^*), \qquad \forall\ x \in \boldsymbol{B}(x^*, \delta) \setminus \{x^*\} = \{y \in \mathcal{S}\ :\ 0 < d(y, x^*) < \delta\}. \tag{1.16}$$

Assume by contradiction that x^* *is not a strict global minimiser of* f *in* $\mathcal{S}$. *Thus, there exists* $\bar{x} \in \mathcal{S}$ *with* $\bar{x} \neq x^*$ *such that* $f(\bar{x}) \leq f(x^*)$. *Since* C *is spherically convex, we can take a minimal geodesic segment* $\gamma_{x^*\bar{x}} : [0,1] \to \mathcal{S}$ *joining* x^* *and* $\bar{x}$*, then we have*

$$\gamma_{x^*\bar{x}}(0) = x^*, \qquad \gamma_{x^*\bar{x}}(1) = \bar{x}.$$

Considering that f *is spherically quasi-convex, by Definition 1.5.11 we have*

$$f(\gamma_{x^*\bar{x}}(t)) \leq \max\{f(x^*), f(\bar{x})\} = f(x^*) \tag{1.17}$$

for any $t \in [0,1]$. *On the other hand, for* t *sufficiently small we have* $\gamma_{x^*\bar{x}}(t) \in \boldsymbol{B}(x^*, \delta)$*, which means*

$$f(\gamma_{x^*\bar{x}}(t)) > f(x^*)$$

Therefore, the inequality (1.17) *contradicts* (1.16). □

Proposition 1.5.5

Let $\mathcal{S} \subseteq \mathbb{S}^{n-1}$ *be a spherically convex set and* $f : \mathcal{S} \to \mathbb{R}$ *be a strictly spherically quasi-convex function. Then* f *has at most one local minimiser point which is also a global minimiser point of* f.

Proof:

Without loss of generality assume by contradiction that the strictly spherically quasi-convex $f(\gamma(t)) < \max\{f(x^*), f(\bar{x})\}$ *for any* $t \in [0,1]$. *Since we can take* t *sufficiently close to* 0 *or* 1, *the last inequality function* f *has two local minimiser* $x^*, \bar{x} \in \mathcal{S}$ *with* $\bar{x} \neq x^*$. *Thus,*

$$f(x) > f(x^*), \qquad \forall\ x \in \boldsymbol{B}(x^*, \delta) \setminus \{x^*\} = \{y \in \mathcal{S}\ :\ 0 < d(y, x^*) < \delta\},$$

and

$$f(x) > f(\bar{x}), \qquad \forall\ x \in \boldsymbol{B}(\bar{x}, \delta) \setminus \{\bar{x}\} = \{y \in \mathcal{S}\ :\ 0 < d(y, \bar{x}) < \delta\}.$$

we can take a minimal geodesic segment $\gamma_{x^*\bar{x}} : [0,1] \to \mathcal{S}$ *joining* x^* *and* $\bar{x}$, *then we have*

$$\gamma(0) = x^*, \quad \gamma(1) = \bar{x}.$$

Due to f *being strictly spherically quasi-convex, we have*

$$f(\gamma(t)) < \max\{f(x^*), f(\bar{x})\}$$

for any $t \in [0,1]$. *If we can take* t *sufficiently close to* 0 *or* 1, *it gives* $f(\gamma(t)) > f(x^*)$ *or* $f(\gamma(t)) > f(\bar{x})$, *respectively. The last two inequalities contradicts the assumption that* $x^*, \bar{x}$ *are two distinct local minimisers. Thus,* f *has at most one local minimiser point. Since* f *is strictly quasi-convex, the local minimiser is strict. Therefore, Proposition 1.5.4 implies that the local minimiser point is global.* □

Proposition 1.5.6

Let $\mathcal{S} \subseteq \mathbb{S}^{n-1}$ *be an open spherically convex set and* $f : \mathcal{S} \to \mathbb{R}$ *be a differentiable function. Then* f *is spherically quasi-convex if and only if*

$$f(x) \leq f(y) \Longrightarrow \langle Df(y), x\rangle - \langle x, y\rangle\langle Df(y), y\rangle \leq 0, \qquad \forall\ x, y \in \mathcal{S}. \tag{1.18}$$

Proof:

Let $\gamma : I \to \mathcal{S}$ *be a geodesic segment and consider the composition* $f \circ \gamma : I \to \mathbb{R}$. *The usual characterisation of scalar quasi-convex functions implies that* $f \circ \gamma$ *is quasi-convex if and*

only if

$$f(\gamma(t_1)) \leq f(\gamma(t_2)) \Longrightarrow \frac{d}{dt} f(\gamma(t_2))(t_1 - t_2) \leq 0, \qquad \forall\ t_1, t_2 \in I. \tag{1.19}$$

On the other hand, for each $x, y \in \mathcal{S}$ *with* $y \neq x$ *we have from* (1.10) *that* γ_{xy} *is the minimal geodesic segment from* $x = \gamma_{xy}(0)$ *to* $y = \gamma_{xy}(1)$ *and*

$$\gamma'_{xy}(1) = \frac{\arccos\langle x, y\rangle}{\sqrt{1 - \langle x, y\rangle^2}} \left(yy^T - I_n\right) x \in T_y\mathbb{S}^{n-1}, \qquad y \neq -x.$$

Note that letting $x = \gamma(t_1)$ *and* $y = \gamma(t_2)$ *we have that* $\gamma_{xy}(t) = \gamma(t_1 + t(t_2 - t_1))$. *Therefore, by using* (1.13) *we can rewrite the right hand side of* (1.19) *as*

$$\begin{aligned}
\frac{d}{dt}\left(f(\gamma(t_2))\right)(t_1 - t_2) &= \langle \operatorname{grad} f(\gamma(t_2)), \gamma'(t_2)\rangle (t_1 - t_2) \\
&= \left\langle \operatorname{grad} f(\gamma(t_2)), \frac{\gamma(t_2) - \gamma(t_1)}{t_2 - t_1} \right\rangle (t_1 - t_2) \\
&= \langle \operatorname{grad} f(\gamma(t_2)), \gamma(t_1)\rangle - 0 \\
&= \Big\langle Df(t_2) - \langle Df(t_2), \gamma(t_2)\rangle \gamma(t_2),\ \gamma(t_1) \Big\rangle \\
&= \langle Df(t_2), \gamma(t_1)\rangle - \langle Df(t_2), \gamma(t_2)\rangle \langle \gamma(t_2), \gamma(t_1)\rangle \leq 0
\end{aligned}$$

which is equivalent to (1.18). □

CHAPTER 2

LINEAR COMPLEMENTARITY PROBLEMS ON EXTENDED SECOND ORDER CONES

In this chapter we elaborate the formulation and the solution to linear complementarity problems on extended second order cones (ESOCLCP). We present the major characterisation of ESOCLCP in Theorem 2.1.1. Based on Theorem 2.1.1, an ESOCLCP can be converted into a mixed complementarity problem on the nonnegative orthant. We state necessary and sufficient conditions for a point to be a solution to the converted problem. We also present solution strategies for this problem, as well as some numerical examples. The results in this chapter are published in the paper [82], co-worked with my supervisor. In order to improve the readability of our results, in this chapter, we give more explicit proofs and more detailed explanations about these results.

2.1 Problem formulation

Let $T = \left(\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right)$, with $A \in \mathbb{R}^{k\times k}$, $B \in \mathbb{R}^{k\times \ell}$, $C \in \mathbb{R}^{\ell\times k}$ and $D \in \mathbb{R}^{\ell\times\ell}$. Let $r = \left(\begin{smallmatrix} p \\ q \end{smallmatrix}\right)$ with $p \in \mathbb{R}^k$, $q \in \mathbb{R}^\ell$. The linear complementarity problem defined by the extended second order cone L and a linear function $F(x,u) = T\left(\begin{smallmatrix} x \\ u \end{smallmatrix}\right) + r$ is:

$$\mathrm{LCP}(F,L)\begin{cases} Find\ \left(\begin{smallmatrix} x \\ u \end{smallmatrix}\right) \in L,\ such\ that \\ F(x,u) \in M\ and\ \left\langle \left(\begin{smallmatrix} x \\ u \end{smallmatrix}\right), F(x,u)\right\rangle = 0. \end{cases} \tag{2.1}$$

Based on the idea of Proposition 1.3.1 and Proposition 1.3.2, using Proposition 1.4.3, we developed the following theorem, which shows the equivalence of various complementarity

problems associated with ESOC. For convenience, let $n = k + \ell$.

Theorem 2.1.1

Denote $z = \binom{x}{u}$, $\hat{z} = \binom{\hat{x}}{u} := \binom{x - \|u\|e}{u}$, $\widetilde{z} = \begin{pmatrix} \tilde{x} \\ u \\ t \end{pmatrix} := \begin{pmatrix} x - t \\ u \\ t \end{pmatrix}$ *and* $r = \binom{p}{q}$ *with* $x, p \in \mathbb{R}^k$, $u, q \in \mathbb{R}^\ell$, *and* $t \in \mathbb{R}$. *Let* $T = \begin{pmatrix} A & B \\ C & D \end{pmatrix}$ *with* $A \in \mathbb{R}^{k \times k}$, $B \in \mathbb{R}^{k \times \ell}$, $C \in \mathbb{R}^{\ell \times k}$ *and* $D \in \mathbb{R}^{\ell \times \ell}$. *The square matrices* T, A *and* D *are nonsingular. Let* L *be the extended second order come.*

(i) Suppose $u = 0$. *We have*

$$z \in \text{SOL-LCP}(T, r, L) \iff x \in \text{SOL-LCP}(A, p, \mathbb{R}^k_+) \text{ and } e^\top (Ax + p) \geq \|Cx + q\|.$$

(ii) Suppose $Cx + Du + q = 0$. *Then,*

$$z \in \text{SOL-LCP}(T, r, L) \iff z \in \text{SOL-MixCP}(F_1, F_2, \mathbb{R}^k_+) \text{ and } x \geq \|u\|,$$

where $F_1(x, u) = Ax + Bu + p$ *and* $F_2(x, u) = 0$.

(iii) Suppose $u \neq 0$ *and* $Cx + Du + q \neq 0$. *We have*

$$z \in \text{SOL-LCP}(T, r, L) \iff z \in \text{SOL-MixICP}(G_1, F_1, F_2, \mathbb{R}^k_+),$$

where

$$F_2(x, u) = \left(\|u\|C + ue^\top A\right) x + ue^\top (Bu + p) + \|u\|(Du + q),$$

$G_1(x, u) = x - \|u\|e$ *and* $F_1(x, u) = Ax + Bu + p$.

(iv) Suppose $u \neq 0$ *and* $Cx + Du + q \neq 0$. *We have*

$$z \in \text{SOL-LCP}(T, r, L) \iff \hat{z} \in \text{SOL-MixCP}(F_1, F_2, \mathbb{R}^k_+),$$

where

$$F_2(\hat{x}, u) = \left(\|u\|C + ue^\top A\right) (\hat{x} + \|u\|e) + ue^\top (Bu + p) + \|u\|(Du + q)$$

and $F_1(\hat{x}, u) = A(\hat{x} + \|u\|e) + Bu + p$.

(v) Suppose $u \neq 0$, $Cx+Du+q \neq 0$ and $\|u\|C+u^\top eA$ is a nonsingular matrix. We have

$$z \in \text{SOL-LCP}(T,r,L) \iff \hat{z} \in \text{SOL-ICP}(F_1,F_2,\mathbb{R}^k_+),$$

where

$$F_1(u) = A\left(\left(\|u\|C+ue^\top A\right)^{-1}\left(ue^\top(Bu+p)+\|u\|(Du+q)\right)\right)+Bu+p$$

and

$$F_2(u) = \left(\|u\|C+ue^\top A\right)^{-1}\left(ue^\top(Bu+p)+\|u\|(Du+q)\right).$$

(vi) Suppose $u \neq 0$, $Cx+Du+q \neq 0$. We have

$$z \in \text{SOL-LCP}(T,r,L) \iff \exists t>0,$$

such that

$$\tilde{z} \in \text{MixCP}(\widetilde{F}_1,\widetilde{F}_2,\mathbb{R}^k_+),$$

where

$$\widetilde{F}_1(\tilde{x},u,t) = A(\tilde{x}+te)+Bu+p \tag{2.2}$$

and

$$\widetilde{F}_2(\tilde{x},u,t) = \begin{pmatrix} \left(tC+ue^\top A\right)(\tilde{x}+te)+ue^\top(Bu+p)+t(Du+q) \\ t^2-\|u\|^2 \end{pmatrix}. \tag{2.3}$$

Proof:

(i) We have that $z \in$ SOL-LCP(T,r,L) is equivalent to $(x,0,Ax+p,Cx+q) \in \mathcal{C}(L)$ or, by item (i) and (ii) of Proposition 1.4.3, equivalent to $(x,Ax+p) \in \mathcal{C}(\mathbb{R}^k_+)$ and $e^\top(Ax+p) \geq \|Cx+q\|$.

(ii) We have that $z \in$ SOL-LCP(T,r,L) is equivalent to $(x,u,Ax+Bu+p,0) \in \mathcal{C}(L)$ or, by item (i) and (iii) of Proposition 1.4.3, equivalent to $(x,Ax+Bu+p) \in \mathcal{C}(\mathbb{R}^k_+)$ and

$x \geq \|u\|$, *or equivalent to*

$$z \in \text{SOL-MixCP}(F_1, F_2, \mathbb{R}^k_+) \text{ and } x \geq \|u\|,$$

where $F_1(x,u) = Ax + Bu + p$ *and* $F_2(x,u) = 0$.

(iii) Suppose that $z \in \text{SOL-LCP}(T, r, L)$. *Then,* $(x, u, y, v) \in \mathcal{C}(L)$, *where* $y = Ax + Bu + p$ *and* $v = Cx + Du + q$. *Then, by item (iv) of Proposition 1.4.3 we obtain that* $\exists \lambda > 0$ *such that*

$$Cx + Du + q = v = -\lambda u, \tag{2.4}$$

$$e^\top (Ax + Bu + p) = e^\top y = \|v\| = \|Cx + Du + q\| = \lambda \|u\|, \tag{2.5}$$

$$(G_1(x,u), F_1(x,u)) = (x - \|u\|e, Ax + Bu + p) = (x - \|u\|e, y) \in \mathcal{C}(\mathbb{R}^k_+). \tag{2.6}$$

From equation (2.4) *we obtain* $\|u\|(Cx + Du + q) = -\lambda \|u\| u$, *which by equation* (2.5) *implies* $\|u\|(Cx + Du + q) = -ue^\top (Ax + Bu + p)$, *which after some algebra gives*

$$F_2(x,u) = 0. \tag{2.7}$$

From equations (2.6) *and* (2.7) *we conclude that* $z \in \text{SOL-MixICP}(F_1, F_2, G_1)$.

Conversely suppose that $z \in \text{SOL-MixICP}(F_1, F_2, G_1)$. *Then,*

$$\|u\| v + ue^\top y = \|u\|(Cx + Du + q) + ue^\top (Ax + Bu + p) = F_2(x,u) = 0 \tag{2.8}$$

and

$$(x - \|u\|e, y) = (x - \|u\|e, Ax + Bu + p) = (G_1(x,u), F_1(x,u)) \in \mathcal{C}(\mathbb{R}^k_+), \tag{2.9}$$

where $v = Cx + Du + q$ *and* $y = Ax + Bu + p$. *Equations* (2.9) *and* (2.8) *imply*

$$v = -\lambda u, \tag{2.10}$$

where

$$\lambda = (e^\top y)/\|u\| > 0. \tag{2.11}$$

Equations (2.10) *and* (2.11) *imply*

$$e^\top y = \|v\| \tag{2.12}$$

By item (iv) of Proposition 1.4.3, equations (2.10), (2.12) *and* (2.9) *imply* $(x, y, u, v) \in C(L)$ *and therefore* $z \in \text{SOL-LCP}(T, r, L)$.

(iv) It is a simple reformulation of item (iii) by using the change of variables

$$\begin{pmatrix} x \\ u \end{pmatrix} \to \begin{pmatrix} \hat{x} \\ u \end{pmatrix} := \begin{pmatrix} x - \|u\| e \\ u \end{pmatrix}.$$

(v) It is a simple reformulation of item (iv) by using that $\|u\|C + u^\top eA$ *is a nonsingular matrix.*

(vi) Suppose that $z \in \text{SOL-LCP}(T, r, L)$. *Then,* $(x, u, y, v) \in \mathcal{C}(L)$, *where* $y = Ax + Bu + p$ *and* $v = Cx + Du + q$. *Let* $t = \|u\|$, *Then, by item (iv) of Proposition 1.4.3 we have that* $\exists \lambda > 0$ *such that*

$$Cx + Du + q = v = -\lambda u, \tag{2.13}$$

$$e^\top (Ax + Bu + p) = e^\top y = \|v\| = \|Cx + Du + q\| = \lambda t, \tag{2.14}$$

$$(\tilde{x}, \widetilde{F}_1(\tilde{x}, u, t)) = (x - te, Ax + Bu + p) = (x - te, y) \in \mathcal{C}(\mathbb{R}^k_+) \tag{2.15}$$

where $\widetilde{z} = \begin{pmatrix} \tilde{x} \\ u \\ t \end{pmatrix} := \begin{pmatrix} x - t \\ u \\ t \end{pmatrix} \in \mathbb{R}^k \times \mathbb{R}^\ell \times \mathbb{R}$. *From equation* (2.13) *we obtain* $t(Cx + Du + q) = -t\lambda u$, *which by equation* (2.14) *implies* $t(Cx + Du + q) = -ue^\top (Ax + Bu + p)$, *which after some algebra gives*

$$\widetilde{F}_2(\tilde{x}, u, t) = 0. \tag{2.16}$$

Equations (2.15) *and* (2.16) *yield* $z \in \text{SOL-MixCP}(\widetilde{F}_1, \widetilde{F}_2, \mathbb{R}^k_+)$.

□

Comment: Many well-developed methods to a complementarity problem (1.6) are based

on a smooth function F [11, 12, 19]. Hence, by modifying the semi-smooth function $F_1(\hat{x}, u)$ in item (iv), we introduce item (vi) with a smooth function $\widetilde{F}_1(\tilde{x}, u, t)$.

Comment: Converting an ESOCLCP to a MixCP will very likely increase its complexity, because it converts a linear problem to a nonlinear one. However, due to lacking methods for solving an ESOCLCP, we have to use other available methods to solve it. Given the fact that there are many methods, especially the complementarity function (C-function) method, proposed for solving the complementarity problem on nonnegative orthant [30, 31, 51, 71], we are therefore motivated to implement such conversion. As it is converted to a MixCP, we will be able to solve the ESOCLCP by means of a C-function.

The scalar form of *Fischer-Burmeister (FB) C-function* [31] is defined as:

$$\psi_{FB}(a, b) = \sqrt{a^2 + b^2} - (a + b) \quad \forall (a, b) \in \mathbb{R}^2.$$

The equivalent FB-based formulation of MixCP is:

$$\mathbb{F}_{FB}^{\text{MixCP}}(x, u, t) := \begin{pmatrix} \psi_{FB}\left(x_1, (\widetilde{F}_1)_1(x, u, t)\right) \\ \vdots \\ \psi_{FB}\left(x_k, (\widetilde{F}_1)_k(x, u, t)\right) \\ \widetilde{F}_2(x, u, t) \end{pmatrix} \tag{2.17}$$

The FB-based formulation of MixCP is semi-smooth. Based on the property of FB C-function, if there is a point $\begin{pmatrix} x^* \\ u^* \\ t^* \end{pmatrix}$ such that

$$\mathbb{F}_{FB}^{\text{MixCP}}(x^*, u^*, t^*) = 0, \tag{2.18}$$

then $\begin{pmatrix} x^* \\ u^* \\ t^* \end{pmatrix}$ is a solution to MixCP. The equation (2.18) is semi-smooth, but it still can be solved by using semi-smooth Newton's method. Denote by $\partial \mathbb{F}_{FB}^{\text{MixCP}}(x, u, t)$ the generalised Jacobian set of $\mathbb{F}_{FB}^{\text{MixCP}}(x, u, t)$. Since $\mathbb{F}_{FB}^{\text{MixCP}}(x, u, t)$ is semi-smooth, we have that $\partial \mathbb{F}_{FB}^{\text{MixCP}}(x, u, t)$ satisfies

$$\partial \mathbb{F}_{FB}^{\text{MixCP}}(x, u, t) \subseteq \begin{pmatrix} \mathcal{D}_a(x, u, t) + \mathcal{D}_b(x, u, t) J_x \widetilde{F}_1(x, u, t) & \mathcal{D}_b(x, u, t) J_{\binom{u}{t}} \widetilde{F}_1(x, u, t) \\ J_x \widetilde{F}_2(x, u, t) & J_{\binom{u}{t}} \widetilde{F}_2(x, u, t) \end{pmatrix}.$$

where $\mathcal{D}_a$ and $\mathcal{D}_b$ are $k \times k$ diagonal matrices respectively denoted by $diag(a_1(x,u,t),\ \ldots,\ a_k(x,u,t))$ and $diag(b_1(x,u,t),\ldots,b_k(x,u,t))$, with $\bar{\mathbf{B}}(x,1)$ denoting a closed unit ball centered at the point x:

$$(a_i(x,u,t),b_i(x,u,t)) = \begin{cases} = \frac{(x_i,(\widetilde{F}_1)_i(x,u,t))}{\sqrt{x_i^2+(\widetilde{F}_1)_i^2(x,u,t)}} - (1,1) & if\ \left(x_i,(\widetilde{F}_1)_i(x,u,t)\right) \neq (0,0) \\ \in \bar{\mathbf{B}}\left((0,0),1\right) - (1,1) & if\ \left(x_i,(\widetilde{F}_1)_i(x,u,t)\right) = (0,0) \end{cases} \tag{2.19}$$

Specifically, for $i \in (1,\ldots,k)$ such that $(x_i,\widetilde{F}_1^i(x,u,t)) \neq (0,0)$ the i-th FB-based formulation $\left(\mathbb{F}_{FB}^{\text{MixCP}}\right)_i$ is differentiable at $\binom{x}{u}_t$. Take an element

$$\mathcal{A} \in \partial \mathbb{F}_{FB}^{\text{MixCP}}(x,u,t).$$

Denoting by $e^i = (0,0,\ldots,1,\ldots,0)^\top$ the i-th coordinate vector, and the i-th row of the Jacobian $\mathcal{A}_i$, which is the derivative of $\psi_{FB}\left(x_i,(\widetilde{F}_1)_i(x,u,t)\right)$ with respect to x, is shown as follows:

$$\begin{aligned}\left(\mathcal{A}_x\right)_i(x,u,t) =& \frac{\partial \psi_{FB}\left(x_i,(\widetilde{F}_1)_i(x,u,t)\right)}{\partial x} = a_i(x,u,t)e^i + b_i(x,u,t)J_x(\widetilde{F}_1)_i(x,u,t) \\ =& \left(\frac{x_i}{\sqrt{x_i^2+(\widetilde{F}_1)_i^2(x,u,t)}} - 1\right)e^i + \left(\frac{(\widetilde{F}_1)_i(x,u,t)}{\sqrt{x_i^2+(\widetilde{F}_1)_i^2(x,u,t)}} - 1\right)J_x(\widetilde{F}_1)_i(x,u,t)\end{aligned}$$

Similarly, for $i \in (1,\ldots,k)$ such that $(x_i,\widetilde{F}_1^i(x,u,t)) \neq (0,0)$, the i-th row of Jacobian $\mathcal{A}_i$ with respect to $\binom{u}{t}$ is:

$$\begin{aligned}\left(\mathcal{A}_{\binom{u}{t}}\right)_i(x,u,t) =& \frac{\partial \psi_{FB}\left(x_i,(\widetilde{F}_1)_i(x,u,t)\right)}{\partial \binom{u}{t}} = b_i(x,u,t)J_{\binom{u}{t}}(\widetilde{F}_1)_i(x,u,t) \\ =& \left(\frac{(\widetilde{F}_1)_i(x,u,t)}{\sqrt{x_i^2+(\widetilde{F}_1)_i^2(x,u,t)}} - 1\right)J_{\binom{u}{t}}(\widetilde{F}_1)_i(x,u,t).\end{aligned}$$

By (2.19), for $i \in (1,\ldots,k)$, if the pair $(x_i,\widetilde{F}_1^i(x,u,t)) = (0,0)$, since $\mathbb{F}_{FB}^{\text{MixCP}}$ is semi-smooth at origin, the Jacobian $\mathcal{A}_i$ at the origin will be a generalised Jacobian of a composite function provided that $\partial\|(0,0)\| = \bar{\mathbf{B}}\left((0,0),1\right)$. We have

$$\left(\mathcal{A}_x\right)_i(x,u,t) = \left\{\left(\tilde{a}_i e^i + \tilde{b}_i J(\widetilde{F}_1)_i(x,u,t)\right) : (\tilde{a},\tilde{b}) \in \bar{\mathbf{B}}\left((0,0),1\right) - (1,1)\right\},$$

for $i \in (1, \ldots, k)$ and

$$\left(\mathcal{A}_{\binom{u}{t}}\right)_i (x,u,t) = \left\{ \left(\tilde{a}_i \cdot 0 + \tilde{b}_i J(\widetilde{F}_1)_i(x,u,t) \right) : (\tilde{a},\tilde{b}) \in \bar{\mathbf{B}}\left((0,0),1\right) - (1,1) \right\},$$

for $i \in (k+1, \ldots, n+1)$.

For convenience, for $i \in (1, \ldots, k)$ satisfying $(x_i, \widetilde{F}_1^i(x,u,t)) = (0,0)$, we choose

$$\tilde{a}_i(x,u,t) = 0 - 1, \qquad \tilde{b}_i(x,u,t) = 0 - 1.$$

It is easy to prove that $(\tilde{a},\tilde{b}) \in \bar{\mathbf{B}}\left((0,0),1\right) - (1,1)$. Then we conclude

$$\begin{aligned} \left(\mathcal{A}_x\right)_i (x,u,t) &= \tilde{a}_i(x,u,t) e^i + \tilde{b}_i(x,u,t) J_x(\widetilde{F}_1)_i(x,u,t) \\ &= -\, e^i - J_x(\widetilde{F}_1)_i(x,u,t), \end{aligned}$$

and

$$\begin{aligned} \left(\mathcal{A}_{\binom{u}{t}}\right)_i (x,u,t) &= \tilde{b}_i(x,u,t) J_{\binom{u}{t}}(\widetilde{F}_1)_i(x,u,t) \\ &= -\, J_{\binom{u}{t}}(\widetilde{F}_1)_i(x,u,t). \end{aligned}$$

Moreover, by the continuous differentiability of $\widetilde{F}_2(x,u,t)$, for $i \in (k+1, \ldots, n+1)$, the Jacobian $\mathcal{A}_i$ is:

$$\mathcal{A}_i = \begin{pmatrix} J_x(\widetilde{F}_2)_i(x,u,t) & J_{\binom{u}{t}}(\widetilde{F}_2)_i(x,u,t) \end{pmatrix}.$$

Hence, the Jacobian matrix for $\mathbb{F}_{FB}^{\text{MixCP}}(x,u,t)$ can be written as:

$$\mathcal{A} = \begin{pmatrix} D_a + D_b J_x \widetilde{F}_1(x,u,t) & D_b J_{\binom{u}{t}} \widetilde{F}_1(x,u,t) \\ J_x \widetilde{F}_2(x,u,t) & J_{\binom{u}{t}} \widetilde{F}_2(x,u,t) \end{pmatrix}, \tag{2.20}$$

where D_a and D_b are nonpositive definite diagonal matrices:

$$(D_a)_{ii} := \begin{cases} \frac{x_i}{\sqrt{x_i^2 + (\widetilde{F}_1)_i^2(x,u,t)}} - 1, & if \ \left(x_i, (\widetilde{F}_1)_i(x,u,t)\right) \neq (0,0), \\ -1, & if \ \left(x_i, (\widetilde{F}_1)_i(x,u,t)\right) = (0,0), \end{cases} \qquad i \in \{1, \ldots, k\},$$

$$(D_b)_{ii} := \begin{cases} \frac{(\widetilde{F}_1)_i(x,u,t)}{\sqrt{x_i^2+(\widetilde{F}_1)_i^2(x,u,t)}} - 1, & if \ \Big(x_i, (\widetilde{F}_1)_i(x,u,t)\Big) \neq (0,0), \\ -1, & if \ \Big(x_i, (\widetilde{F}_1)_i(x,u,t)\Big) = (0,0), \end{cases} \qquad i \in \{1,\ldots,k\}.$$

Detailed methods of solving the semi-smooth equation (2.18) will be introduced in next section.

2.2 Find the solution

Many methods are proposed to solve the complementarity problem. Two of the most popular methods are: to reformulate the complementarity problem to a system of nonlinear equations; and, to reformulate it to an unconstrained minimisation problem. In this section, we will present the details for both methods.

2.2.1 Reformulate to a system of nonlinear equations

In the previous section, we provided the equivalent FB-based formulation of MixCP (2.17). Equation (2.18) enables us to solve the MixCP as an unconstrained system of nonlinear equation: For convenience, we restate the equation (2.18) here:

$$\mathbb{F}_{FB}^{\text{MixCP}}(x^*, u^*, t^*) = 0.$$

The study about solving the complementarity problem in such nonlinear equation system is abundant. Since the FB C-function is not continuously differentiable, many semi-smooth methods are developed for solving the complementarity problem. These includes Newton-type methods [4, 20, 40, 88], Levenberg-Marquardt methods [21, 67], etc. Theses methods are proved to have at least a linear rate of convergence under certain assumptions, among which the nonsingularity is the most important assumption to guarantee the convergence of these algorithms. We use the following proposition to obtain conditions for the nonsingularity of the generalised Jacobian of $\mathbb{F}_{FB}^{\text{MixCP}}(x^*, u^*, t^*)$.

Before stating the proposition, we define the following index sets:

$$\begin{aligned}
\alpha &:= \left\{ i : x_i = 0 < (\widetilde{F}_1)_i(x,u,t) \right\}, \\
\beta &:= \left\{ i : x_i = 0 = (\widetilde{F}_1)_i(x,u,t) \right\}, \\
\gamma &:= \left\{ i : x_i > 0 = (\widetilde{F}_1)_i(x,u,t) \right\}, \\
\delta &:= \{1,\ldots,k\} \setminus (\alpha \cup \beta \cup \gamma).
\end{aligned}$$

Proposition 2.2.1

[19, Proposition 9.4.2] If $\widetilde{F}_1(x,u,t)$ and $\widetilde{F}_1(x,u,t)$ are continuously differentiable, given $\begin{pmatrix} x \\ u \\ t \end{pmatrix} \in \mathbb{R}^{k+\ell+1}$. Let $\overline{\alpha} := \gamma \cup \beta \cup \delta$ be the complement of α in $\{1,\ldots,k\}$. Assume that

(i) the submatrices

$$\begin{pmatrix} J_{\binom{u}{t}}\widetilde{F}_2(x,u,t) & J_{x_{\tilde{\gamma}}}\widetilde{F}_2(x,u,t) \\ J_{\binom{u}{t}}(\widetilde{F}_1)_{\tilde{\gamma}}(x,u,t) & J_{x_{\tilde{\gamma}}}(\widetilde{F}_1)_{\tilde{\gamma}}(x,u,t) \end{pmatrix}$$

are nonsingular for all $\tilde{\gamma}$ satisfying

$$\gamma \subseteq \tilde{\gamma} \cup \gamma \cup \beta,$$

(ii) the Schur complement of

$$\begin{pmatrix} J_{\binom{u}{t}}\widetilde{F}_2(x,u,t) & J_{x_{\gamma}}\widetilde{F}_2(x,u,t) \\ J_{\binom{u}{t}}(\widetilde{F}_1)_{\gamma}(x,u,t) & J_{x_{\gamma}}(\widetilde{F}_1)_{\gamma}(x,u,t) \end{pmatrix}$$

in

$$\begin{pmatrix} J_{\binom{u}{t}}\widetilde{F}_2(x,u,t) & J_{x_{\overline{\alpha}}}\widetilde{F}_2(x,u,t) \\ J_{\binom{u}{t}}(\widetilde{F}_1)_{\overline{\alpha}}(x,u,t) & J_{x_{\overline{\alpha}}}(\widetilde{F}_1)_{\overline{\alpha}}(x,u,t) \end{pmatrix}$$

is a P_0 matrix,

then the Jacobian of $\mathbb{F}_{FB}^{\mathrm{MixCP}}(x,u,t)$ (2.20) is nonsingular.

First , we illustrate the semi-smooth inexact Newton's Method.

Algorithm 1 (Semi-smooth Inexact Newton's method) [20]:

Input: the initial point $z_0 := \begin{pmatrix} x_0 \\ u_0 \\ t_0 \end{pmatrix} \in \mathbb{R}^{k+\ell+1}$, and the tolerance $\eta_0 \in \mathbb{R}_+$.

Step 1: Set $j = 0$.

Step 2: If $\mathbb{F}_{FB}^{\text{MixCP}}(z_j) = 0$, then STOP.

Step 3: Select an element $\mathcal{A}$ in the generalised Jacobian set $\partial\mathbb{F}_{FB}^{\text{MixCP}}(x, u, t)$, and find a direction $d_j \in \mathbb{R}^{k+\ell+1}$ such that

$$\mathbb{F}_{FB}^{\text{MixCP}}(z_j) + \mathcal{A}^\top(z_j)d_j = r_j,$$

where the residual vector $r_j \in \mathbb{R}^{k+\ell+1}$ satisfying

$$\|r_j\| \leq \eta_j \|\mathbb{F}_{FB}^{\text{MixCP}}(z_j)\|.$$

Step 4: Choose $\eta_{j+1} \geq 0$; set $z_{j+1} := z_j + d_j$ and $j := j + 1$; go to **Step 2**.

The above algorithm is a modification of the semi-smooth algorithm introduced by Qi and Sun [89]. It is worth noting that there are many other Newton-type algorithms for solving a complementarity problem as a system of nonlinear equations. We refer interested readers to [13, 98, 107].

The following theorem is from [20, Theorem 3.2]. It proves that the semi-smooth inexact Newton's Method at least Q-linearly converges to a solution to (2.18).

Theorem 2.2.1

Let $\mathbb{F}_{FB}^{\text{MixCP}}(x, u, t)$ be semi-smooth in $\boldsymbol{B}(z^*, \delta)$, where $\delta > 0$, and $z^* := \begin{pmatrix} x^* \\ u^* \\ t^* \end{pmatrix}$ satisfies $\mathbb{F}_{FB}^{\text{MixCP}}(x^*, u^*, t^*) = 0$. If $\partial\mathbb{F}_{FB}^{\text{MixCP}}(x^*, u^*, t^*)$ is nonsingular. Then the following statements hold:

*(i) There exists $\bar{\eta} > 0$ such that, if $z_0 \in \boldsymbol{B}(z^*, \delta)$ and $\eta_j \leq \bar{\eta}$ for all j, then **Algorithm 1** is well defined and the sequence $\{z_j\}$ converges Q-linearly to z^*.*

(ii) If $\lim_{j\to\infty} \eta_j = 0$, then the sequence $\{z_j\}$ converges Q-superlinearly to z^*.

(iii) If $\mathbb{F}_{FB}^{\text{MixCP}}(x, u, t)$ is strongly semi-smooth at z^*, and if there exists $\tilde{\eta} > 0$ such that $\eta_j \leq \tilde{\eta}\|\mathbb{F}_{FB}^{\text{MixCP}}(x_j, u_j, t_j)$ for all j, then the sequence $\{z_j\}$ converges Q-quadratic to z^*.

As mentioned earlier, A disadvantage of employing Newton's method for finding solutions to (2.18) is that, it requires nonsingularity assumption of the Jacobian $\mathcal{A}$. Hence, it is worthy

to present the widely-used Levenberg-Marquardt (LM) algorithm [73]. LM algorithm have least a linear rate of convergence without requiring all the Jacobian matrices in the iteration to be nonsingular. LM algorithm approximates the Hessian matrix by:

$$\mathcal{H}(z) = \mathcal{A}^\top(z)\mathcal{A}(z),$$

and it approximates the gradient by:

$$\mathcal{G}(z) = \mathcal{A}^\top(z)\mathbb{F}_{FB}^{\text{MixCP}}(z).$$

Hence, its upgraded step will be

$$z_{j+1} = z_j - \left[\mathcal{A}^\top(z_j)\mathcal{A}(z_j) + \mu\mathbb{I}\right]^{-1}\mathcal{A}^\top(z_j)\mathbb{F}_{FB}^{\text{MixCP}}(z_j).$$

The parameter μ is used to prevent d_j from being too large when $\mathcal{A}^\top(z_j)\mathcal{A}(z_j)$ is nearly singular. When μ equals to zero, the upgrading step is just the same as a Newton's method which uses the Hessian matrix for approximation.

However, noting that the LM algorithm sacrifices calculation speed for the compatibility of singular Jacobian matrix, its number of iteration is probably greater than that of a Newton's method. It should be worth noting that the settings of parameters also influence the calculation speed. A greater value of parameter μ will lead to a longer calculation time and larger number of iteration. LM algorithm is demonstrated as follows:

Algorithm 2 (Semismooth Inexact Levenberg-Marquardt Method):

Input: the initial point $z_0 = \begin{pmatrix} x_0 \\ u \\ t \end{pmatrix} \in \mathbb{R}^{k+\ell+1}$, the LM parameter $\mu_0 \in \mathbb{R}_+$, and the tolerance $\eta_0 \in \mathbb{R}_+$.

Step 1: Set $k = 0$.

Step 2: If $\mathbb{F}_{FB}^{\text{MixCP}}(z_j) = 0$, stop.

Step 3: Select an element $\mathcal{A}$ in the generalised Jacobian set $\partial\mathbb{F}_{FB}^{\text{MixCP}}(x, u, t)$, and find a direction $d_j \in \mathbb{R}^{k+\ell+1}$ such that

$$\mathcal{A}(z_j)^\top\mathbb{F}_{FB}^{\text{MixCP}}(z_j) + \left[\mathcal{A}^\top(z_j)\mathcal{A}(z_j) + \mu\mathbb{I}\right] d_j = r_j,$$

where the residual vector $r_j \in \mathbb{R}^{k+\ell+1}$ satisfying

$$\|r_j\| \leq \eta_j \|\mathcal{A}^\top(z_j)\mathbb{F}_{FB}^{\text{MixCP}}(z_j)\|.$$

Step 4: Choose $\eta_{j+1} \geq 0$ and $\mu_{j+1} \geq 0$; set $z_{j+1} := z_j + d_j$ and $j := j + 1$; go to **Step 2**.

Admittedly, Levenberg-Marquardt algorithm is also an efficient algorithm for finding solutions, because it also converges at least quadratically to a numerical solution.

Theorem 2.2.2

[21] *Let $\mathbb{F}_{FB}^{\text{MixCP}}(x,u,t)$ be semi-smooth in $\boldsymbol{B}(z^*,\delta)$, where $\delta > 0$, and $z^* := \begin{pmatrix} x^* \\ u^* \\ t^* \end{pmatrix}$ satisfies $\mathbb{F}_{FB}^{\text{MixCP}}(x^*,u^*,t^*) = 0$. If $\partial\mathbb{F}_{FB}^{\text{MixCP}}(x^*,u^*,t^*)$ is nonsingular. Then the following statements hold:*

*(i) If $z_0 \in \boldsymbol{B}(z^*,\delta)$, then **Algorithm 2** is well defined and the sequence $\{z_j\}$ converges Q-linearly to z^*.*

(ii) If $\lim_{j\to\infty} \eta_j = 0$ and $\lim_{j\to\infty} \mu_j = 0$, then the sequence $\{z_j\}$ converges Q-superlinearly to z^*.

(iii) If $\mathbb{F}_{FB}^{\text{MixCP}}(x,u,t)$ is strongly semi-smooth at z^*, and if there exists $\tilde{\eta} > 0$ and $\tilde{\mu} > 0$ such that $\eta_j \leq \tilde{\eta}\|\mathbb{F}_{FB}^{\text{MixCP}}(x_j,u_j,t_j)$ and $\mu_j \leq \tilde{\mu}\|\mathbb{F}_{FB}^{\text{MixCP}}(x_j,u_j,t_j)$ for all j, then the sequence $\{z_j\}$ converges Q-quadratic to z^*.

2.2.2 Reformulate to an unconstrained minimisation problem

Another approach of solving the MixCP is to reformulate it to an unconstrained minimisation problem. We will investigate the associated merit function of MixCP:

$$\theta_{FB}^{\text{MixCP}}(x,u,t) := \frac{1}{2}\mathbb{F}_{FB}^{\text{MixCP}}(x,u,t)^\top \mathbb{F}_{FB}^{\text{MixCP}}(x,u,t). \tag{2.21}$$

Obviously, based on (2.18), if there is a point $\begin{pmatrix} x^* \\ u^* \\ t^* \end{pmatrix}$ such that

$$\theta_{FB}^{\text{MixCP}}(x^*,u^*,t^*) = 0, \tag{2.22}$$

then $\begin{pmatrix} x^* \\ u^* \\ t^* \end{pmatrix}$ is a solution to MixCP. FB C-function has its squared function $\psi_{FB}^2(a,b)$ to be continuously differentiable on $\mathbb{R}^2$ [50]. Hence, it is easy to verify that the merit function

$\theta_{FB}^{\text{MixCP}}(x,u,t)$ is continuously differentiable if both $\widetilde{F}_1(x,u,t)$ and $\widetilde{F}_2(x,u,t)$ are. Since the merit function $\theta_{FB}^{\text{MixCP}}(x,u,t)$ is nonnegative, if there exists a solution to (2.22), then the point $\begin{pmatrix} x^* \\ u^* \\ t^* \end{pmatrix}$ will be a global minimiser of (2.21). Hence, the MixCP can be reformulated as the following unconstrained minimisation problem:

$$\min_{\begin{pmatrix} x \\ u \\ t \end{pmatrix} \in \mathbb{R}^{k+\ell+1}} \theta_{FB}^{\text{MixCP}}(x,u,t). \tag{2.23}$$

In order to find a global minimiser of (2.23), the general gradient of the merit function $\theta_{FB}^{\text{MixCP}}(x,u,t)$ will be introduced:

$$\nabla \theta_{FB}^{\text{MixCP}}(x,u,t) = \mathcal{A}^\top \mathbb{F}_{FB}^{\text{MixCP}}(x,u,t),$$

for any $\mathcal{A} \in \partial \mathbb{F}_{FB}^{\text{MixCP}}(x,u,t)$. It is not complicate to find the gradient of the merit function, but the difficulty lies in the lack of efficient tools to work out the non-convexity of $\theta_{FB}^{\text{MixCP}}(x,u,t)$. For this purpose, we give the notion of stationary point. The point $\begin{pmatrix} x^* \\ u^* \\ t^* \end{pmatrix}$ is said to be a stationary point of (2.21) if it satisfies the following inequality:

$$\left\langle \begin{pmatrix} x - x^* \\ u - u^* \\ t - t^* \end{pmatrix}, \nabla \theta_{FB}^{\text{MixCP}}(x^*,u^*,t^*) \right\rangle = 0, \quad \forall \begin{pmatrix} x \\ u \\ t \end{pmatrix} \in \mathbb{R}_+^{k+\ell+1}. \tag{2.24}$$

The problem (2.24) is a variational inequality problem [34]. However, though $\begin{pmatrix} x^* \\ u^* \\ t^* \end{pmatrix}$ can be a stationary point of (2.21), it does not guarantee that $\begin{pmatrix} x^* \\ u^* \\ t^* \end{pmatrix}$ is a global minimiser of (2.21). The discussion below is associated with the Jacobian $\partial \mathbb{F}_{FB}^{\text{MixCP}}(x,u,t)$ at a stationary point of $\theta_{FB}^{\text{MixCP}}(x,u,t)$. Before introducing the notion of *FB regular [19]* point, we define the following index sets:

$$\begin{aligned}
&\mathcal{C} := \left\{ i : x_i \geq 0, (\widetilde{F}_1)_i \geq 0, x_i (\widetilde{F}_1)_i(x,u,t) = 0 \right\}, && \textit{complementarity index}, \\
&\mathcal{R} := \{1,\ldots,k\} \setminus \mathcal{C}, && \textit{residual index}, \\
&\mathcal{P} := \left\{ i \in \mathbb{R} : x_i > 0, (\widetilde{F}_1)_i(x,u,t) > 0 \right\}, && \textit{positive index}, \\
&\mathcal{N} := \mathcal{R} \setminus \mathcal{P}, && \textit{negative index}.
\end{aligned}$$

Comment: The FB regularity of the point $\begin{pmatrix} x^* \\ u^* \\ t^* \end{pmatrix}$ is defined by the Jacobian of $\widetilde{F}_1$ and $\widetilde{F}_2$ at $\begin{pmatrix} x^* \\ u^* \\ t^* \end{pmatrix}$. The motivation of introducing the notion of FB regularity is to avoid the case

when $\begin{pmatrix} x^* \\ u^* \\ t^* \end{pmatrix}$ is a stationary point but the Jacobian $\partial \mathbb{F}_{FB}^{\text{MixCP}}(x,u,t)$ is singular. The property of FB regularity is tailored to the FB C-function as suggested by its name. In the following we will introduce the definition of FB regularity, and use Theorem 2.2.3 to show the connection between FB regularity and the solution to (2.23).

Definition 2.2.1 (FB regular)

A point $\begin{pmatrix} x \\ u \\ t \end{pmatrix} \in \mathbb{R}^k \times \mathbb{R}^\ell \times \mathbb{R}$ *is called FB-Regular for the merit function* $\theta_{FB}^{\text{MixCP}}$ *if the Jacobian* $J_x\widetilde{F}_2(x,u,t)$ *is nonsingular, and if for any* $w \in \mathbb{R}^k$, $w \neq 0$ *with*

$$w_i \begin{cases} > 0, & if \;\; i \in \mathcal{P}, \\ < 0, & if \;\; i \in \mathcal{N}, \\ = 0, & if \;\; i \in \mathcal{C}\,. \end{cases} \qquad i \in \{1,\ldots,k\},$$

there exists a nonzero vector $v \in \mathbb{R}^k$ *such that*

$$v_i \begin{cases} \geq 0, & if \;\; i \in \mathcal{P}, \\ \leq 0, & if \;\; i \in \mathcal{N}, \\ = 0, & if \;\; i \in \mathcal{C}\,. \end{cases} \qquad i \in \{1,\ldots,k\}, \tag{2.25}$$

and

$$w^\top \left(M(x,u,t)/J_{\binom{u}{t}}\widetilde{F}_2(x,u,t) \right) v \geq 0, \tag{2.26}$$

where

$$M(x,u,t) := \begin{pmatrix} J_x\widetilde{F}_1(x,u,t) & J_{\binom{u}{t}}\widetilde{F}_1(x,u,t) \\ J_x\widetilde{F}_2(x,u,t) & J_{\binom{u}{t}}\widetilde{F}_2(x,u,t) \end{pmatrix} \in \mathbb{R}^{(k+\ell+1)\times(k+\ell+1)} \tag{2.27}$$

and $M(x,u,t)/J_{\binom{u}{t}}\widetilde{F}_2(x,u,t)$ *is the Schur complement of the block* $J_{\binom{u}{t}}\widetilde{F}_2(x,u,t)$ *of the matrix* $M(x,u,t)$.

Based on the original problem ESOCLCP, we can find the Jacobian of $\widetilde{F}_1$ and $\widetilde{F}_2$ for MixCP:

$$J\widetilde{F}_1(x,u,t) := \left(J_x\widetilde{F}_1(x,u,t) \quad J_{\binom{u}{t}}\widetilde{F}_1(x,u,t) \right) = \left(A \quad \widetilde{B} \right),$$

$$J\widetilde{F}_2(x,u,t) := \left(J_x\widetilde{F}_2(x,u,t) \quad J_{\binom{u}{t}}\widetilde{F}_2(x,u,t) \right) = \left(\widetilde{C} \quad \widetilde{D} \right),$$

where

$$\widetilde{B} := \begin{pmatrix} B & Ae \end{pmatrix}, \qquad \widetilde{C} := \begin{pmatrix} tC + ue^\top A \\ 0 \end{pmatrix},$$

and

$$\widetilde{D} := \begin{pmatrix} [A(x+te)+Bu+p]^\top eI_\ell + ue^\top B + tD & Cx + 2tCe + ue^\top Ae + Du \\ -2u^\top & 2t \end{pmatrix}.$$

In our case, if the Jacobian

$$J_{\binom{u}{t}}\widetilde{F}_2(x,u,t) = \widetilde{D}$$

is nonsingular, then the Schur complement exists:

$$\left(M(x,u,t)/J_{\binom{u}{t}}\widetilde{F}_2(x,u,t)\right) = A - \widetilde{B}\widetilde{D}^{-1}\widetilde{C}. \tag{2.28}$$

The following theorem is based on [19, Theorem 9.4.4]. We made a slight modification in order to fit the context of the problem MixCP$(\widetilde{F}_1, \widetilde{F}_2, \mathbb{R}^k)$. As there is no proof given in the source for this theorem, for the sake of completeness, a rewritten proof based on Definition 2.2.1 is provided as follows:

Theorem 2.2.3

Let $\widetilde{F}_1 : \mathbb{R}^k \times \mathbb{R}^\ell \times \mathbb{R} \to \mathbb{R}^k$ and $\widetilde{F}_2 : \mathbb{R}^k \times \mathbb{R}^\ell \times \mathbb{R} \to \mathbb{R}^\ell \times \mathbb{R}$ be continuously differentiable. If $\binom{x^*}{u^*}_{t^*} \in \mathbb{R}^k \times \mathbb{R}^\ell \times \mathbb{R}$ is a stationary point of $\theta_{FB}^{\text{MixCP}}$, then $\binom{x^*}{u^*}_{t^*}$ is a solution to (2.23) if and only if $\binom{x^*}{u^*}_{t^*}$ is an FB regular point of $\theta_{FB}^{\text{MixCP}}$.

Proof:

Suppose that $\binom{x^*}{u^*}_{t^*} \in$ SOL-MixCP$(\widetilde{F}_1, \widetilde{F}_2, \mathbb{R}^k)$. It then follows that $\binom{x^*}{u^*}_{t^*}$ is a global minimum and hence a stationary point of $\theta_{FB}^{\text{MixCP}}$. Thus, $(x^*, \widetilde{F}_1(z^*)) \in \mathcal{C}(\mathbb{R}^k_+)$, then we have $\mathcal{P} = \mathcal{N} = \emptyset$. Therefore the FB regularity of x^* holds since $x^* = x_{\mathcal{C}}$, and we cannot find a nonzero vector x satisfying conditions (2.25). Conversely, suppose that x^* is FB regular and $\binom{x^*}{u^*}_{t^*}$ is a stationary point of $\theta_{FB}^{\text{MixCP}}$. It follows that $\nabla\theta_{FB}^{\text{MixCP}} = 0$, i.e.:

$$\mathcal{A}^\top \mathbb{F}_{FB}^{\text{MixCP}} = \begin{pmatrix} D_a + \left(J_x\widetilde{F}_1(x^*,u^*,t^*)\right)^\top D_b & \left(J_x\widetilde{F}_2(x^*,u^*,t^*)\right)^\top \\ \left(J_{\binom{u}{t}}\widetilde{F}_1(x^*,u^*,t^*)\right)^\top D_b & \left(J_{\binom{u}{t}}\widetilde{F}_2(x^*,u^*,t^*)\right)^\top \end{pmatrix} \mathbb{F}_{FB}^{\text{MixCP}} = 0,$$

where D_a and D_b are diagonal matrices with

$$(D_a)_{ii} = \begin{cases} \frac{x_i}{\sqrt{x_i^2+(\widetilde{F}_1)_i^2(x,u,t)}} - 1, & if \ \left(x_i, (\widetilde{F}_1)_i(x,u,t)\right) \neq (0,0), \\ -1, & if \ \left(x_i, (\widetilde{F}_1)_i(x,u,t)\right) = (0,0), \end{cases} \qquad i \in \{1,\ldots,k\},$$

$$(D_b)_{ii} = \begin{cases} \frac{(\widetilde{F}_1)_i(x,u,t)}{\sqrt{x_i^2+(\widetilde{F}_1)_i^2(x,u,t)}} - 1, & if \ \left(x_i, (\widetilde{F}_1)_i(x,u,t)\right) \neq (0,0), \\ -1, & if \ \left(x_i, (\widetilde{F}_1)_i(x,u,t)\right) = (0,0), \end{cases} \qquad i \in \{1,\ldots,k\}.$$

Hence, for any $w \in \mathbb{R}^k \times \mathbb{R}^\ell \times \mathbb{R}$, we have

$$w^\top \begin{pmatrix} D_a + \left(J_x \widetilde{F}_1(x^*,u^*,t^*)\right)^\top D_b & \left(J_x \widetilde{F}_2(x^*,u^*,t^*)\right)^\top \\ \left(J_{\binom{u}{t}} \widetilde{F}_1(x^*,u^*,t^*)\right)^\top D_b & \left(J_{\binom{u}{t}} \widetilde{F}_2(x^*,u^*,t^*)\right)^\top \end{pmatrix} \mathbb{F}_{FB}^{\mathrm{MixCP}} = 0. \tag{2.29}$$

Suppose that $\binom{x^*}{u^*}_{t^*}$ is not a solution to MixCP*, we have that the index set $\mathcal{R}$ is not empty. Define $v := D_b \mathbb{F}_{FB}^{\mathrm{MixCP}}$, we have*

$$v_{\mathcal{C}} = 0, \qquad v_{\mathcal{P}} > 0, \qquad v_{\mathcal{N}} < 0.$$

Take w with

$$w_{\mathcal{C}} = 0, \qquad w_{\mathcal{P}} > 0, \qquad w_{\mathcal{N}} < 0.$$

From the definition of D_a and D_b, we know that $D_a \mathbb{F}_{FB}^{\mathrm{MixCP}}$ and $D_b \mathbb{F}_{FB}^{\mathrm{MixCP}}$ have the same sign. Therefore

$$w^\top (D_a \mathbb{F}_{FB}^{\mathrm{MixCP}}) = w_{\mathcal{C}}^\top (D_a \mathbb{F}_{FB}^{\mathrm{MixCP}})_{\mathcal{C}} + w_{\mathcal{P}}^\top (D_a \mathbb{F}_{FB}^{\mathrm{MixCP}})_{\mathcal{P}} + w_{\mathcal{N}}^\top (D_a \mathbb{F}_{FB}^{\mathrm{MixCP}})_{\mathcal{N}} > 0. \tag{2.30}$$

By regularity of $J\widetilde{F}_1(z)^\top$, we conclude

$$w^\top J\widetilde{F}_1(z)^\top (D_a \mathbb{F}_{FB}^{\mathrm{MixCP}}) = w^\top J\widetilde{F}_1(z)^\top w \geq 0. \tag{2.31}$$

The inequalities (2.30) *and* (2.31) *together contradict condition* (2.29)*. Hence $\mathcal{R} = \emptyset$. It means that z^* is the solution to* MixCP$(\widetilde{F}_1, \widetilde{F}_2, \mathbb{R}^k)$.

□

If the Schur complement $M(x,u,t)/J_{\binom{u}{t}}\widetilde{F}_2(x,u,t) \in \mathbb{R}^{k\times k}$ is a signed S_0 matrix, then $\binom{x}{u}{t}$ is an FB regular point of $\theta_{FB}^{\text{MixCP}}$. Before proving this assertion, the definition of signed S_0 matrix is given as follows:

Definition 2.2.2 (Signed S_0 matrix)

Let $\widetilde{F}_1 : \mathbb{R}^k \times \mathbb{R}^\ell \times \mathbb{R} \to \mathbb{R}^k$ and $\widetilde{F}_2 : \mathbb{R}^k \times \mathbb{R}^\ell \times \mathbb{R} \to \mathbb{R}^\ell \times \mathbb{R}$ be continuously differentiable. We say that the Schur Complement $M(x,u,t)/J_{\binom{u}{t}}\widetilde{F}_2(x,u,t)$ is a signed S_0 matrix if

$$\Xi := \Lambda \left(M(x,u,t)/J_{\binom{u}{t}}\widetilde{F}_2(x,u,t) \right) \Lambda \tag{2.32}$$

is a S_0 matrix (see Definition 1.2.1), where $\Lambda \in \mathbb{R}^{k\times k}$ is the diagonal matrix whose diagonal entries λ_i, $i \in \{1,\ldots,k\}$ satisfy

$$\lambda_i := \begin{cases} 1, & if \;\; i \in \mathcal{P}, \\ -1, & if \;\; i \in \mathcal{N}, \\ 0, & if \;\; i \in \mathcal{C}\,. \end{cases} \tag{2.33}$$

Theorem 2.2.4

Let $\widetilde{F}_1 : \mathbb{R}^k \times \mathbb{R}^\ell \times \mathbb{R} \to \mathbb{R}^k$ and $\widetilde{F}_2 : \mathbb{R}^k \times \mathbb{R}^\ell \times \mathbb{R} \to \mathbb{R}^\ell \times \mathbb{R}$ be continuously differentiable functions, and suppose the Jacobian matrix $J_{\binom{u}{t}}\widetilde{F}_2$ is nonsingular. If the Schur complement $M(x^*,u^*,t^*)/J_{\binom{u}{t}}\widetilde{F}_2(x^*,u^*,t^*)$ is a signed S_0 matrix, then $\binom{x^*}{u^*}{t^*}$ is an FB regular point of $\theta_{FB}^{\text{MixCP}}$.

Proof:

Suppose that $M(x^*,u^*,t^*)/J_{\binom{u}{t}}\widetilde{F}_2(x^*,u^*,t^*)$ is a signed S_0 matrix, denoted by

$$M_2 = M(x^*,u^*,t^*)/J_{\binom{u}{t}}\widetilde{F}_2(x^*,u^*,t^*),$$

then Definition 2.2.2 implies:

$$(\Lambda M_2 \Lambda)\, u \geq 0. \tag{2.34}$$

where Λ *is defined in* (2.32)*. It is clear that there exists* $0 \neq u \geq 0$ *such that* u *is a solution to* (2.34)*. Denote by* $(M_2)_i$ *the i-th row of matrix* M_2*, we can rewrite* (2.34) *as*

$$\begin{aligned} 1 \cdot (M_2)_i \Lambda u &\geq 0, \qquad \forall i \in \mathcal{P}, \\ -1 \cdot (M_2)_i \Lambda u &\geq 0, \qquad \forall i \in \mathcal{N}, \\ 0 \cdot (M_2)_i \Lambda u &= 0, \qquad \forall i \in \mathcal{C}\,. \end{aligned} \tag{2.35}$$

Take any $w \in \mathbb{R}^k \setminus \{0\}$ *with*

$$w_i \begin{cases} > 0, \quad if \;\; i \in \mathcal{P}, \\ < 0, \quad if \;\; i \in \mathcal{N}, \\ = 0, \quad if \;\; i \in \mathcal{C}, \end{cases} \qquad i \in \{1, \ldots, k\},$$

multiplies with inequalities/equality (2.35)*, then there exists a nonzero vector* $v = \Lambda u$ *such that*

$$\begin{aligned} w_i \cdot (M_2)_i v &\geq 0, \qquad \forall i \in \mathcal{P}, \\ w_i \cdot (M_2)_i v &\geq 0, \qquad \forall i \in \mathcal{N}, \\ w_i \cdot (M_2)_i v &= 0, \qquad \forall i \in \mathcal{C}\,. \end{aligned} \tag{2.36}$$

hold. Noting that

$$v := \begin{cases} u_i, \quad if \;\; i \in \mathcal{P}, \\ -u_i, \quad if \;\; i \in \mathcal{N}, \\ 0, \quad if \;\; i \in \mathcal{C}\,. \end{cases}$$

Summing up the three iequalities/equality above in (2.36)*, we get*

$$\sum_{i \in \mathcal{P}} w_i \left(M_2 v\right)_i + \sum_{i \in \mathcal{N}} w_i \left(M_2 v\right)_i + \sum_{i \in \mathcal{C}} w_i \left(M_2 v\right)_i \geq 0. \tag{2.37}$$

That is equivalent to

$$w^\top \left(M_2\right) v \geq 0.$$

Hence, $\begin{pmatrix} x^* \\ u^* \\ t^* \end{pmatrix}$ *is an FB regular point of* $\theta_{FB}^{\mathrm{MixCP}}$*.* □

Example 2.2.1

Taking the notation in Theorem 2.2.4. Suppose that M_2 *is a signed* S_0 *matrix. Let the*

Schur complement be

$$M_2 = \begin{pmatrix} m_{11} & m_{12} & m_{13} \\ m_{21} & m_{22} & m_{23} \\ m_{31} & m_{32} & m_{33} \end{pmatrix}.$$

with the index set $\mathcal{P} = \{1\}$, $\mathcal{N} = \{2\}$, *and* $\mathcal{C} = \{3\}$. *There exists a nonzero vector* $u = (u_1, u_2, u_3)^\top$ *with* $0 \neq u \geq 0$ *such that*

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} m_{11} & m_{12} & m_{13} \\ m_{21} & m_{22} & m_{23} \\ m_{31} & m_{32} & m_{33} \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix}$$
$$= \begin{pmatrix} 1 \cdot m_{11} & 1 \cdot m_{12} & 1 \cdot m_{13} \\ -1 \cdot m_{21} & -1 \cdot m_{22} & -1 \cdot m_{23} \\ 0 \cdot m_{31} & 0 \cdot m_{32} & 0 \cdot m_3 \end{pmatrix} \begin{pmatrix} u_1 \\ -u_2 \\ 0 \end{pmatrix} \geq \mathbf{0}.$$

Denote $v = (u_1, -u_2, 0)^\top$, *the above linear system can be written as:*

$$\begin{aligned} 1 \cdot (m_{11}, m_{12}, m_{13})\, v &\geq 0, \\ -1 \cdot (m_{21}, m_{22}, m_{23})\, v &\geq 0, \\ 0 \cdot (m_{31}, m_{32}, m_{33})\, v &= 0. \end{aligned}$$

Take $w = (w_1, -w_2, 0)$ *with* w_1, $w_2 \in \mathbb{R}_+$, *the following linear system holds:*

$$\begin{aligned} w_1 \cdot (m_{11}, m_{12}, m_{13})\, v &\geq 0, \\ -w_2 \cdot (m_{21}, m_{22}, m_{23})\, v &\geq 0, \\ 0 \cdot (m_{31}, m_{32}, m_{33})\, v &= 0. \end{aligned}$$

Hence, we can say for any vector $w = \begin{pmatrix} w_1 \\ -w_2 \\ 0 \end{pmatrix}$, *there exists a vector* $v = \begin{pmatrix} u_1 \\ -u_2 \\ 0 \end{pmatrix}$ *such that*

$$w^\top M_2 v \geq 0$$

holds. Hence,(x^*) *is a FB regular point of the merit function* $\theta(x)$.

The Theorem 2.2.4 has proved that the signed S_0 property of the Schur complement

$M(x,u,t)/J_{\binom{u}{t}}\widetilde{F}_2(x,u,t)$ is a sufficient condition for x to be an FB regular point. This condition outperforms the FB regularity condition, because that the verification of the S_0 property of the matrix (2.32) can be accomplished by simple linear programming. On the other hand, verifying FB regularity will be in general more complex and computationally expensive. Together with Theorem 2.2.3 and 2.2.4, the conclusion of the signed S_0 property enables us to find a solution to MixCP with algorithms.

Algorithm 3 (FB line search method):

Input: the initial point $z_0 = (x,u,t) \in \mathbb{R}^k \times \mathbb{R}^\ell \times \mathbb{R}$, $\rho > 0$, $\gamma \in (0,1)$, and the tolerance $r \in \mathbb{R}_{++}$.

Step 1: Set $k = 0$.

Step 2: If $\|\nabla\theta_{FB}^{\mathrm{MixCP}}(z_j)\| \leq r$, then STOP.

Step 3: Select an element $\mathcal{A} \in \partial\mathbb{F}_{FB}^{\mathrm{MixCP}}(x,u,t)$, and find a direction $d_j \in \mathbb{R}^k \times \mathbb{R}^\ell \times \mathbb{R}$ such that

$$\mathbb{F}_{FB}^{\mathrm{MixCP}}(z_j) + \mathcal{A}^\top(z_j)d_j = 0. \tag{2.38}$$

If the system (2.38) is not solvable or if the condition

$$\nabla\theta_{FB}^{\mathrm{MixCP}}(z_j)d_j \leq -\rho\|d_j\| \tag{2.39}$$

is not satisfied, reset $d_j := -\nabla\theta_{FB}^{\mathrm{MixCP}}(z_j)$.

Step 4: Find the smallest nonnegative integer i_j such that, with $i = i_j$, we have

$$\theta_{FB}^{\mathrm{MixCP}}(z_j + 2^{-i}d_j) \leq \theta_{FB}^{\mathrm{MixCP}}(z_j) + \gamma 2^{-i}\theta_{FB}^{\mathrm{MixCP}}(z_j)^\top d_j;$$

set $\tau_j := 2^{-i_j}$.

Step 5: Set $z_{j+1} := z_j + d_j$ and $j := j+1$, go to **Step 2**.

In general, matrix $\mathcal{A}$ in **Step 3** of this algorithm is not necessary to be a generalised Jacobian of $\mathbb{F}_{FB}^{\mathrm{MixCP}}$ at the iterate z_j, but it still is required to be a nonsingular matrix. In this algorithm, **Step 2** is to examine whether an iterate z_j is a stationary point of $\theta_{FB}^{\mathrm{MixCP}}$ or not; whereas the condition (2.39) and **Step 4** are used to ensure that the iterate z_j is a FB regular point of $\theta_{FB}^{\mathrm{MixCP}}$.

2.3 A numerical example

In this section, we will illustrate a numerical example corresponding to item (iv) of Proposition 1.4.3. Let $L(3,2)$ and $M(3,2)$ be an ESOC and its dual cone defined by (1.7) and (1.8), respectively. Denote

$$z = \left(\begin{smallmatrix} x \\ u \end{smallmatrix}\right) \in \mathbb{R}^3 \times \mathbb{R}^2, \quad \hat{z} = \left(\begin{smallmatrix} \hat{x} \\ u \end{smallmatrix}\right) := \left(\begin{smallmatrix} x - \|u\|e \\ u \end{smallmatrix}\right) \in \mathbb{R}^3 \times \mathbb{R}^2,$$

and

$$\tilde{z} = \begin{pmatrix} \tilde{x} \\ u \\ t \end{pmatrix} := \begin{pmatrix} x - t \\ u \\ t \end{pmatrix} \in \mathbb{R}^3 \times \mathbb{R}^2 \times \mathbb{R}.$$

Repeat for convenience an ESOCLCP defined by an extended second order cone $L \in \mathbb{R}^3 \times \mathbb{R}^2$ and a linear function $F : \mathbb{R}^3 \times \mathbb{R}^2 \to \mathbb{R}^3 \times \mathbb{R}^2$, $F(x,u) = T\left(\begin{smallmatrix} x \\ u \end{smallmatrix}\right) + r$, is:

$$\mathrm{LCP}(F, L) \begin{cases} Find\ \left(\begin{smallmatrix} x \\ u \end{smallmatrix}\right) \in L,\ such\ that \\ F(x,u) \in M\ and\ \langle \left(\begin{smallmatrix} x \\ u \end{smallmatrix}\right), F(x,u) \rangle = 0. \end{cases}$$

where $T = \left(\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right)$, $r = \left(\begin{smallmatrix} p \\ q \end{smallmatrix}\right)$, with $A \in \mathbb{R}^{3\times 3}$, $B \in \mathbb{R}^{3\times 2}$, $C \in \mathbb{R}^{2\times 3}$, $D \in \mathbb{R}^{2\times 2}$, $p \in \mathbb{R}^3$, and $q \in \mathbb{R}^2$. The solution to ESOCLCP is equivalent to the solution to a corresponding MixCP converted by employing item (vi) of Theorem 2.1.1. For convenience the $\mathrm{MixCP}(F_1, F_2, \mathbb{R}^3_+)$, defined by F_1, F_2, and $\mathbb{R}^3_+$, is provided:

$$\mathrm{MixCP}(F_1, F_2, \mathbb{R}^3_+) : \begin{cases} Find\ \begin{pmatrix} x \\ u \\ t \end{pmatrix} \in \mathbb{R}^3 \times \mathbb{R}^2 \times \mathbb{R}^1,\ such\ that \\ \widetilde{F}_2(x,u,t) = 0,\ and\ (x, \widetilde{F}_1(x,u)) \in \mathcal{C}(\mathbb{R}^3_+). \end{cases}$$

where

$$\widetilde{F}_1(x,u,t) = A(x + te) + Bu + p$$

and

$$\widetilde{F}_2(x,u,t) = \begin{pmatrix} \left(tC + ue^\top A\right)(x + te) + ue^\top(Bu + p) + t(Du + q) \\ t^2 - \|u\|^2 \end{pmatrix}.$$

As the propose of this section is not comparing the efficiency of the algorithms, we will only employ the Levenberg-Marquardt algorithm (**Algorithm 2**). Our objective is to solve the following FB-based equation formulation (2.17):

$$\mathbb{F}_{FB}^{\text{MixCP}}(x,u,t) = \begin{pmatrix} \psi_{FB}\left(x_1, (\widetilde{F}_1)_1(x,u,t)\right) \\ \vdots \\ \psi_{FB}\left(x_k, (\widetilde{F}_1)_k(x,u,t)\right) \\ \widetilde{F}_2(x,u,t) \end{pmatrix} = 0.$$

We set the tolerance parameter $r = 10^{-7}$, the initial LM parameter $\mu = 10^{-2}$ and multiply by 10^{-1} after each iteration.

Consider

$$T = \begin{pmatrix} A & B \\ C & D \end{pmatrix} = \begin{pmatrix} 41 & -3 & -31 & 18 & 19 \\ 28 & 22 & -33 & 25 & -29 \\ -23 & -29 & 11 & -21 & -43 \\ -9 & -31 & -20 & -12 & 47 \\ -8 & 46 & 50 & -22 & 21 \end{pmatrix}, \quad r = \begin{pmatrix} p \\ q \end{pmatrix} = \begin{pmatrix} -26 \\ 4 \\ 23 \\ 44 \\ -19 \end{pmatrix},$$

By using **Algorithm 2**, the sequence $\{z_j\}$ converges to a numerical solution in 11 iterations (Table 2.1).

Iteration	The value of $\|\nabla\theta_{FB}^{\text{MixCP}}(z_j)\|$	Optimality value of $\theta_{FB}^{\text{MixCP}}(z_j)$	μ	d
0	1.1760e+09	2.43e+08	0.01	
1	3.9793e+07	7.18e+06	0.001	86.1516
2	1.0348e+07	1.50e+06	1e-04	15.3125
3	640568.382	2.06e+05	1e-05	9.24632
4	37188.3244	2.96e+04	1e-06	4.70891
5	4932.01875	5.83e+03	1e-07	2.09599
6	394.966198	975.2398	1e-08	1.15061
7	19.3492349	115.2587	1e-09	0.51351
8	0.20392831	12.92491	1e-10	0.23590
9	1.5122e-06	0.034555	1e-11	0.03290
10	7.9130e-15	2.00e-06	1e-12	5.3e-05
11	8.2224e-15	3.95e-15	1e-13	1.1e-06

Note: This table shows the iteration of $\{z_j\}$ converges to a numerical solution by using **Algorithm 2**. It takes 11 iterations to reach a numerical solution with tolerance $r = 10^{-7}$.

Table 2.1: Numerical Example: the iteration of $\{z_j\}$

The solution to the MixCP is $\tilde{z}^* = (\tilde{x}^*, u^*, t^*)^\top = \left(\frac{781}{641}, 0, \frac{999}{1328}, \frac{333}{2693}, -\frac{619}{2428}\right)^\top$. Verifying the complementarity:

$$\tilde{x}^* = \left(\frac{781}{641}, 0, \frac{999}{1328}\right)^\top \geq 0, \qquad \widetilde{F}_1(\tilde{z}^*) = \left(0, \frac{8349}{292}, 0\right)^\top \geq 0,$$

$$\langle \tilde{x}^*, \widetilde{F}_1(\tilde{z}^*) \rangle = 0.$$

Therefore we have $(\tilde{x}^*, \widetilde{F}_1(\tilde{z}^*)) \in \mathcal{C}(\mathbb{R}^3_+)$. By the item (vi) of Theorem 2.1.1, we get the solution to LCP$(T, r, L(3,2))$ from $\tilde{z}^*$, that is $z^* = (\tilde{x}^* + t^*, u^*)^\top = \left(\frac{428}{285}, \frac{325}{1147}, \frac{1716}{1657}, \frac{333}{2693}, -\frac{619}{2428}\right)^\top$.

For verifying the complementarity, we show that $z^* \in L(3,2)$ because $\tilde{x}^* + t^* \geq \frac{325}{1147} = \sqrt{\frac{333}{2693}^2 + \frac{619}{2428}^2} = \|u^*\|$; and

$$F(x,u) = \begin{pmatrix} 41 & -3 & -31 & 18 & 19 \\ 28 & 22 & -33 & 25 & -29 \\ -23 & -29 & 11 & -21 & -43 \\ -9 & -31 & -20 & -12 & 47 \\ -8 & 46 & 50 & -22 & 21 \end{pmatrix} \begin{pmatrix} \frac{428}{285} \\ \frac{325}{1147} \\ \frac{1716}{1657} \\ \frac{333}{2693} \\ -\frac{619}{2428} \end{pmatrix} + \begin{pmatrix} -26 \\ 4 \\ 23 \\ 44 \\ -19 \end{pmatrix} = \begin{pmatrix} 0 \\ \frac{8349}{292} \\ 0 \\ -\frac{3943}{316} \\ \frac{4039}{157} \end{pmatrix}.$$

We obtain that $F(x,u) \in M(3,2)$ as $0 + \frac{8349}{292} + 0 \geq \sqrt{(-\frac{3943}{316})^2 + \frac{4039}{157}^2} = \frac{8349}{292}$, and hence $(z^*, F(x,u)) \in \mathcal{C}(L(3,2))$. Therefore, we can confirm that this is a solution to the problem LCP$(T, r, L(3,2))$.

2.4 Conclusions and comments

In this chapter, we study the linear complementarity problem on extended second order cones (ESOCLCP). Our main result is Theorem 2.1.1, which discusses the connections between an ESOCLCP and mixed (implicit, mixed implicit) complementarity problems. Under some mild conditions, we can rewrite an ESOCLCP to a mixed complementarity problems (MixCP) on the nonnegative orthant. In the new formulation, both $\widetilde{F}_1(\tilde{x}, u, t)$ and $\widetilde{F}_2(\tilde{x}, u, t)$ are smooth functions, which simplifies the process of finding solutions to MixCP. The conversion from an ESOCLCP to a MixCP on the nonnegative orthant reduces the complexity of finding solutions to the original problem.

The process of solving MixCP is straightforward. We introduced the FB C-function to reformulate the complementarity problem. By the FB C-function, the complementarity problem

can be reformulated to either a system of nonlinear equations or an unconstrained minimisation problem. For the nonlinear equation approach, we introduced proposition for the nonsingularity of the Jacobian. The semi-smooth inexact Newton method and the Levenberg-Marquardt method are illustrated. Further, we provided theorems to verify the rate of convergence of both algorithms. For the minimisation approach, we introduced and proved Theorem 2.2.3 and Theorem 2.2.4 for the difficulty of non-convexity. A point is sufficiently to be a solution to a MixCP if it satisfies specific conditions related to stationarity, FB regularity (Theorem 2.2.3), and Signed S_0 property (Theorem 2.2.4). These theorems can be used to determine whether a point is a solution to the MixCP or not. Based on the above, a solution to MixCP will be equivalent to a solution to the corresponding ESOCLCP. We use FB Line Search Algorithm for the minimisation approach. In the final section, we illustrate a numerical example corresponding to item (iv) of Proposition 1.4.3 and item (vi) of Theorem 2.1.1.

CHAPTER 3

STOCHASTIC LINEAR COMPLEMENTARITY PROBLEMS ON EXTENDED SECOND ORDER CONES

In many practical situations, uncertainty is a common and realistic problem that results from inaccurate measurement or stochastic variation of data such as price, capacities, loads, etc. In fact, the inaccuracy or uncertainty of these real-world data are inevitable. When these data are applied as parameters in mathematical models, the constraints of models may be violated because of their stochastic characters. These violations may finally cause some difficulties that the optimal solutions obtained from the stochastic data are no longer optimal, even infeasible. Amongst approaches proposed for modeling uncertain quantities, the stochastic models outstand because of their solid mathematical foundations, theoretical richness, and sound techniques of using real data. Complementarity problems imbedded with stochastic models occur in many areas such as finance, telecommunication and engineering. Hence, considering LCP with uncertainty will be meaningful for practical treatments. If partial or all of the coefficients in the LCP are uncertain, the LCP will be turned into a stochastic linear complementarity problem (SLCP), which is firstly introduced by Chen and Fukushima [14]. Articles about SLCP can be found in [15, 22, 37, 63].

Even though the fact that only limited number of results have been obtained on the stochastic complementarity problems, there are still some meaningful results. One of them is the CVaR (conditional value-at-risk, which is also called expected shortfall) minimisation reformulation of stochastic complementarity problem [102]. In this chapter, the stochastic linear complementarity problem on extended second order cones (S-ESOCLCP) will be studied. Based on the results in previous chapter, a method of finding solutions to S-ESOCLCP will be

elaborated, and a numerical example will be presented.

3.1 Problem formulation

Let $(\Omega, \mathcal{F}, \mathcal{P})$ be a probability space defined by:

1. $\Omega \subseteq \mathbb{R}^n$, the sample set of possible outcomes;

2. $\mathcal{F} \subseteq 2^{\Omega}$, a σ-algebra generated by Ω (all subsets of Ω); and

3. $\mathcal{P} : \mathcal{F} \to [0, 1]$, a function map from events to probabilities.

The following is the definition of a stochastic complementarity problem (SCP):

Definition 3.1.1 (Stochastic complementarity problem)

Given a random vector valued function $F(x, \omega) : \mathbb{R}^n \times \Omega \to \mathbb{R}^n$, where $\omega \in \Omega$ is an n-dimensional random vector. A stochastic complementarity problem is defined by

$$SCP(F, \mathbb{R}^n_+, \omega) \begin{cases} Find\ x \in \mathbb{R}^m_+,\ such\ that \\ F(x, \omega) \geq 0, \quad x^\top F(x, \omega) = 0, \quad \omega \in \Omega, \quad a.s. \end{cases} \tag{3.1}$$

The abbreviation almost surely (a.s.) means $F(x, \omega) \geq 0$ and $x^\top F(x, \omega) = 0$ hold almost surely for any $\omega \in \Omega$.

If $F(x, \omega)$ is a linear function of the form $F(x, \omega) = T(\omega)x + r(\omega)$, then we call problem (3.1) a *stochastic linear complementarity problem (SLCP)* , specifically:

$$SLCP(T(\omega), r(\omega), \mathbb{R}^n_+, \omega) \begin{cases} Find\ x \in \mathbb{R}^n_+,\ such\ that \\ T(\omega)x + r(\omega) \geq 0, x^\top (T(\omega)x + r(\omega)) = 0, \omega \in \Omega, \quad a.s. \end{cases} \tag{3.2}$$

In this chapter, we assume that the coefficients $T(\omega)$ and $r(\omega)$ are measurable functions of ω with the following property:

$$\mathbb{E}[\|T(\omega)^\top T(\omega)\|] < \infty \quad and \quad \mathbb{E}[\|r(\omega)\|] < \infty$$

where $\mathbb{E}[\cdot]$ represents the expected value of the random vector in the square bracket.

It should be mentioned that if the possible outcome set Ω contains only one single realisation (and this unique outcome definitely happens), problem (3.2) will degenerate to problem (2.1).

The stochastic linear complementarity problems are very useful in solving practical problems. However, because of the existence of the random vector ω in the function $F(x, \omega)$, it is very difficult and sometimes impossible to find a solution x satisfying all possible outcomes of $\omega \in \Omega$. One plausible idea to improve the viability of finding a solution to SLCP is to associate the problems with probability models, and then persuasive solutions to SLCP are obtainable by finding the solutions to the associated probability models.

Xu and Yu [102] summarised 6 different probability models for finding solutions to SLCP:

(i) **Expected value (EV) method, introduced by Gürkan et. al in [37]**. By using the expectation value $\mathbb{E}[F(x, \omega)]$ to replace the stochastic term $F(x, \omega)$, this method ultimately reformulates (3.2) to (2.1).

(ii) **Expected residual minimisation (ERM) method, introduced by Chen and Fukushima [14]**. This method minimises the expectation of the square norm of the residual $\Phi(x, \omega)$ defined by the following C-function:

$$\min_{x \in \mathbb{R}^n_+} \mathbb{E}\left[\|\Phi(x, \omega)\|^2\right] \tag{3.3}$$

where $\Phi : \mathbb{R}^n \times \Omega \to \mathbb{R}^n$ is a multi dimensional C-function defined as

$$\Phi(x, \omega) := \left(\phi\left(x_1, F_1(x, \omega)\right), \ldots, \phi\left(x_m, F_m(x, \omega)\right)\right)^\top.$$

where $\phi : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ can be any scalar C-function satisfying:

$$\phi(a, b) = 0 \quad \Leftrightarrow \quad a \geq 0, \quad b \geq 0, \quad ab = 0.$$

(iii) **Stochastic mathematical programs with equilibrium constraints (SMPEC) reformulation, introduced by Lin and Fukushima [63,64,75]**. This method highlights a recourse variate $z(\omega)$ to compensate the violation of complementarity in (3.2) for some

outcomes of $\omega \in \Omega$, then it reformulates (3.2) to the following model:

$$\begin{aligned} \min_x \quad & \mathbb{E}\left[\eta\left(z(\omega)\right)\right] \\ s.t. \quad & 0 \leq x \perp \left(F(x,\omega) + z(\omega)\right) \geq 0, \\ & z(\omega) \geq 0, \omega \in \Omega \quad a.s., \end{aligned} \tag{3.4}$$

where $\eta(z) = e^{tp}z$. Ambiguous solutions to SCP can be obtained by minimising the objective function in (3.4), i.e. the expected value of the compensation to the violation of complementarity in (3.2).

(iv) **Stochastic programming (SP) reformulation [100]**. Problem (3.2) is reformulated to the following:

$$\begin{aligned} \min_x \quad & \mathbb{E}\left[\| \left(x \circ F(x,\omega)\right)_+ \|^2\right] \\ s.t. \quad & F(x,\omega) \geq 0, \quad \omega \in \Omega \quad a.s. \\ & x \geq 0. \end{aligned}$$

where $x_+ := \max\{x, 0\}$, and $x \circ F(x,\omega)$ is the Hadamard product of x and $F(x,\omega)$.

(v) **Robust Optimisation [6, 7], which is a deterministic reformulation of** (3.2). And,

(vi) **CVaR minimisation (CM) reformulation [15]**. By using this method, (3.2) is reformulated to a problem that minimises the CVaR of the norm of the loss function $\theta(x,\omega)$, namely:

$$\min_{x\in\mathbb{R}^n} CVaR_\alpha\left(\|\theta(x,\omega)\|\right).$$

The reformulation in item (vi) uses the CVaR, a measure of risk widely applied in financial industry. CVaR was built based on Value at risk (VaR) [76, 91]. Let $\omega \in \Omega$ be a vector with random outcomes and let $\theta(x,\omega) : \mathbb{R}^n \times \Omega \to \mathbb{R}$ be a mapping, the VaR of ω for the loss function is defined as:

$$VaR_\alpha(\theta(x,\omega)) = min\{\Theta \in \mathbb{R} | \mathcal{P}[\theta(x,\omega) \geq \Theta] \leq \alpha\}. \tag{3.5}$$

where $\mathcal{P}[\cdot] \in [0,1]$ is the probability of the event in the square bracket. We call $\theta(x,\omega)$ the loss function. The probability (also called confidence level) $\alpha \in (0,1)$ quantifies the proportion of "worst cases" (that is, $\theta(x,\omega) \geq VaR_\alpha(\theta(x,\omega)) = \Theta$) in the group of all outcomes, and the

other outcomes ($\theta(x,\omega) < \Theta$) would happens with probability $1-\alpha$. Based on the definition of VaR, CVaR is defined as:

$$CVaR_\alpha(\theta(x,\omega)) = \frac{1}{\alpha}\mathbb{E}\left[\theta(x,\omega)\mathbf{1}_{[VaR_\alpha(\theta(x,\omega)),+\infty)}\left(\theta(x,\omega)\right)\right] \tag{3.6}$$

$$= \frac{1}{\alpha}\int_{\theta(x,\omega)\geq VaR_\alpha(\theta(x,\omega))} \theta(x,\omega)d\mathcal{P}(\omega)$$

$$= \frac{1}{\alpha}\int_0^\alpha VaR_\gamma(\theta(x,\omega))d\gamma, \tag{3.7}$$

where $\mathbf{1}_{[VaR_\alpha(\theta(x,\omega)),+\infty)}\left(\theta(x,\omega)\right)$ is an indicator function with

$$\mathbf{1}_{[VaR_\alpha(\theta(x,\omega)),+\infty)}\left(\theta(x,\omega)\right) = \begin{cases} 1 & if\ \theta(x,\omega) \in \left[VaR_\alpha\left(\theta(x,\omega)\right),+\infty\right) \\ 0 & otherwise. \end{cases}$$

$CVaR_\alpha(\theta(x,\omega))$ is the conditional expectation of all outcomes with $\theta(x,\omega) \geq VaR_\alpha(\theta(x,\omega))$. For better understanding the concept of VaR and CVaR, figure 3.1 gives a sample of a loss function $\theta(x,\omega) = \omega$ with one-dimensional normally distributed random value $\omega \sim N(0,1)$. This figure shows that when the confidence level $(1-\alpha)$ is set at 95%, the value of VaR equals to the horizontal coordinate of the red vertical line, and the value of CVaR with 95% confidence level equals the red area to the right of the line.

Proposition 3.1.1

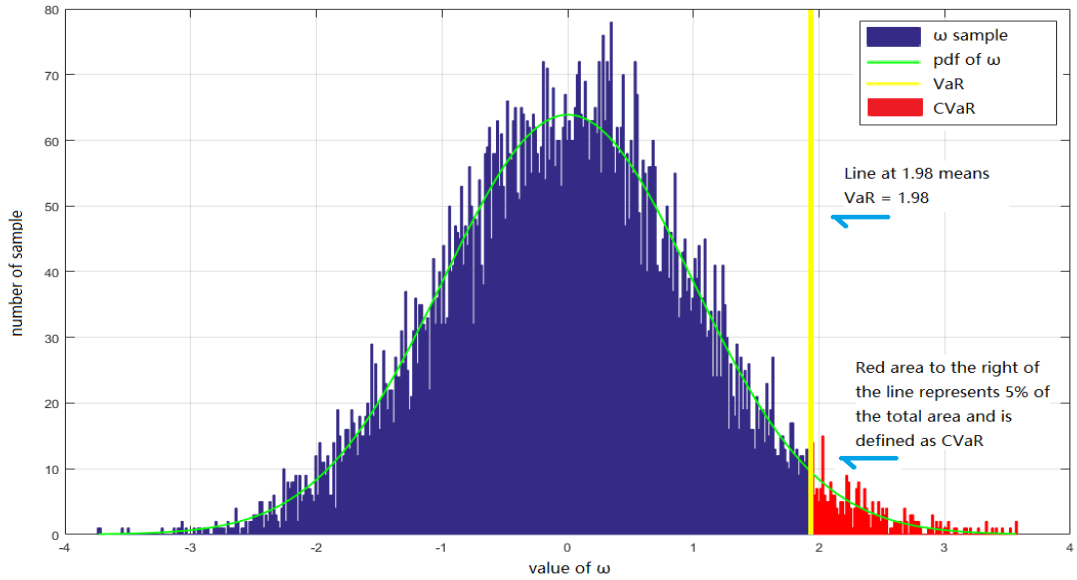
A risk measure $S\left(\theta(x,\omega)\right)$ can have more than one of the following properties:

1. *Positive homogeneity: $S\left(\lambda\theta(x,\omega)\right) = \lambda S\left(\theta(x,\omega)\right)$ for any $\lambda > 0$ and $\omega \in \Omega$,*
2. *Monotonicity: if $\theta(x_1,\omega) \geq \theta(x_2,\omega)$ for any $\omega \in \Omega$, we have $S\left(\theta(x_1,\omega)\right) \geq S\left(\theta(x_2,\omega)\right)$, and*
3. *Sub-additivity: $S\left(\theta(x_1,\omega)+\theta(x_2,\omega)\right) \leq S\left(\theta(x_1,\omega)\right) + S\left(\theta(x_2,\omega)\right)$ for any $\omega \in \Omega$.*

Proposition 3.1.2

[32] The risk measure VaR is

1. *Positive homogeneous, and*
2. *Monotonic.*



Note: For a normal distributed (stochastic) event $\omega \sim N(0,1)$, the distribution of this event shows that only 5% of the outcomes are above 1.98. If the confidence level is set at 95%, then the value of VaR equals 1.98(horizontal axis marked by yellow line), and the value of CVaR equals the integral of the area marked in red color.

Figure 3.1: VaR and CVaR for $\theta(x,\omega) = \omega$, where $\omega \sim N(0,1)$

We remark that VaR is not sub-additive. A counter example shown in [17] will be repeated here.

Example 3.1.1

Consider the function $\theta(x,\omega) = x + \omega$, *where*

$$\omega = \epsilon + \eta, \quad \epsilon \sim N(0,1), \quad \eta = \begin{cases} 0 & \textit{with probability } 0.991 \\ 10 & \textit{with probability } 0.009 \end{cases}$$

In the case when $\alpha = 0.01$, *we obtain*

$$VaR_\alpha\left(\theta(x,\omega) + \theta(y,\omega)\right) = x+y+9.8 > VaR_\alpha\left(\theta(x,\omega)\right)+VaR_\alpha\left(\theta(y,\omega)\right) = x+3.1+y+3.1.$$

Proposition 3.1.3

[1, 32] The risk measure CVaR is

1. *Positive homogeneous,*

2. *Monotonic, and*

3. sub-additive.

Consider SLCP(F, L, ω) defined by the function $F(x, u, \omega) = T(\omega)\binom{x}{u} + r(\omega)$ and the extended second order cone L, problem (3.2) becomes:

$$\text{SLCP}(T(\omega), r(\omega), L) \begin{cases} \textit{Find } (x, u) \in L, \textit{ such that} \\ F(x, u, \omega) \in M \textit{ and } \langle \binom{x}{u}, F(x, u, \omega) \rangle = 0, \ \omega \in \Omega, \quad a.s. \end{cases}$$

where $T(\omega) = \begin{pmatrix} A(\omega) & B(\omega) \\ C(\omega) & D(\omega) \end{pmatrix}$, with $A(\omega) \in \mathbb{R}^{k\times k}$, $B(\omega) \in \mathbb{R}^{k\times \ell}$, $C(\omega) \in \mathbb{R}^{\ell\times k}$ and $D(\omega) \in \mathbb{R}^{\ell\times\ell}$; $r(\omega) = \begin{pmatrix} p(\omega) \\ q(\omega) \end{pmatrix}$, with $p(\omega) \in \mathbb{R}^k$, $q(\omega) \in \mathbb{R}^\ell$, for $\omega \in \Omega$.

By using item (vi) of Theorem 2.1.1, we reformulate SLCP$(T(\omega), r(\omega), L\omega)$ to a stochastic mixed complementarity problem (SMixCP). The Stochastic mixed complementarity problem SMixCP$(\widetilde{F}_1, \widetilde{F}_2, \mathbb{R}^k_+, \omega)$ defined by $\widetilde{F}_1$, $\widetilde{F}_2$, and $\mathbb{R}^k_+$, is

$$\text{SMixCP}(\widetilde{F}_1, \widetilde{F}_2, \mathbb{R}^k_+, \omega) : \begin{cases} \textit{Find } \begin{pmatrix} x \\ u \\ t \end{pmatrix} \in \mathbb{R}^k \times \mathbb{R}^\ell \times \mathbb{R}, \textit{ such that} \\ \widetilde{F}_2(x, u, t, \omega) = 0, \textit{ and } (x, \widetilde{F}_1(x, u, t, \omega)) \in \mathcal{C}(\mathbb{R}^k_+), \ \omega \in \Omega, \quad a.s. \end{cases}$$

Theorem 3.1.1

Suppose $u \neq 0$, $Cx + Du + q \neq 0$. We have

$$z \in \text{SOL-SLCP}(T(\omega), r(\omega), L) \iff \exists t > 0,$$

such that

$$\tilde{z} \in \text{SOL-SMixCP}(\widetilde{F}_1, \widetilde{F}_2, \mathbb{R}^k_+, \omega),$$

where

$$\widetilde{F}_1(x, u, t, \omega) = A(\omega)(x + te) + B(\omega)u + p(\omega)$$

and

$$\widetilde{F}_2(x, u, t, \omega) = \begin{pmatrix} \big(tC(\omega) + ue^\top A(\omega)\big)(x + te) + ue^\top(B(\omega)u + p(\omega)) + t(D(\omega)u + q(\omega)) \\ t^2 - \|u\|^2 \end{pmatrix}. \quad (3.8)$$

The proof is omitted here, as it inherits the idea of Theorem 2.1.1. Theorem 3.1.1 provides an alternative way to find the solutions to the SLCP$(T(\omega), r(\omega), L, \omega)$, by converting it to the

SMixCP($\widetilde{F}_1, \widetilde{F}_2, \mathbb{R}^k_+, \omega$). Such conversion enables us to study SLCP($T(\omega), r(\omega), L, \omega$) through a C-function.

Similar to the process in Chapter 2, *Fischer-Burmeister C-function* will be associated with the problem SMixCP($\widetilde{F}_1, \widetilde{F}_2, \mathbb{R}^k_+, \omega$). The FB-based equation formulation of SMixCP($\widetilde{F}_1, \widetilde{F}_2, \mathbb{R}^k_+, \omega$) is:

$$\mathbb{F}_{FB}^{\text{SMixCP}}(x,u,t,\omega) = \begin{pmatrix} \psi_{FB}\left(x_1, (\widetilde{F}_1)_1(x,u,t,\omega)\right) \\ \vdots \\ \psi_{FB}\left(x_k, (\widetilde{F}_1)_k(x,u,t,\omega)\right) \\ \widetilde{F}_2(x,u,t,\omega) \end{pmatrix}. \tag{3.9}$$

where $\psi_{FB}(\cdot) : \mathbb{R}^2 \to \mathbb{R}$ is the scalar FB C-function stated in Chapter 2. It should be mentioned that the FB C-function is convex, but non-smooth on $\psi_{FB}(0,0)$. According to the definition of FB C-function, a point $\begin{pmatrix} x^* \\ u^* \\ t^* \end{pmatrix}$ is a solution to the stochastic mixed complementarity problem SMixCP($\widetilde{F}_1, \widetilde{F}_2, \mathbb{R}^k_+, \omega$), if and only if

$$\mathbb{F}_{FB}^{\text{SMixCP}}(x,u,t,\omega) = 0. \tag{3.10}$$

Based on the results in the previous chapter, of the SMixCP($\widetilde{F}_1, \widetilde{F}_2, \mathbb{R}^k_+, \omega$) the associated merit function is:

$$\theta_{FB}^{\text{SMixCP}}(x,u,t,\omega) = \frac{1}{2}\mathbb{F}_{FB}^{\text{SMixCP}}(x,u,t,\omega)^\top \mathbb{F}_{FB}^{\text{SMixCP}}(x,u,t,\omega). \tag{3.11}$$

Based on (3.9) and (3.11), the merit function can be written as:

$$\theta_{FB}^{\text{SMixCP}}(x,u,t,\omega) = \frac{1}{2}\sum_{i=1}^{k} \psi_{FB}^2\left(x_i, \widetilde{F}_1^i(x,u,t,\omega)\right) + \frac{1}{2}\sum_{j=1}^{\ell} \widetilde{F}_2^j(x,u,t,\omega).$$

By the definition of merit function, a point $\begin{pmatrix} x^* \\ u^* \\ t^* \end{pmatrix}$ is a solution to the stochastic mixed complementarity problem SMixCP($\widetilde{F}_1, \widetilde{F}_2, \mathbb{R}^k_+, \omega$), if

$$\theta_{FB}^{\text{SMixCP}}(x^*,u^*,t^*,\omega) = 0, \quad \omega \in \Omega \quad a.s.$$

Proposition 3.1.4

The associated merit function $\theta_{FB}^{\mathrm{SMixCP}}(x^*, u^*, t^*, \omega)$ *is continuously differentiable on* $\mathbb{R}^k \times \mathbb{R}^\ell \times \mathbb{R}$, *if* $\widetilde{F}_1(x^*, u^*, t^*, \omega)$ *and* $\widetilde{F}_2(x^*, u^*, t^*, \omega)$ *are continuously differentiable on* $\mathbb{R}^k$ *and* $\mathbb{R}^\ell \times \mathbb{R}$, *respectively.*

Proof:

First we prove that ψ_{FB}^2 *is continuously differentiable. We note that* ψ_{FB} *is continuously differentiable at every* $(a, b) \neq (0, 0)$. *It is easy to verify that* ψ_{FB}^2 *is continuously differentiable at every* $(a, b) \neq (0, 0)$. *Consider the following to limits at point* $(a, b) = (0, 0)$:

$$\lim_{\Delta x \to 0} \frac{\psi_{FB}^2(\Delta x, 0) - \psi_{FB}^2(0, 0)}{\Delta x} = \frac{2\,(\Delta x^2) - 2\sqrt{\Delta x^2} \cdot \Delta x}{\Delta x} = 2(\Delta x - |\Delta x|) = 0,$$

and

$$\lim_{\Delta y \to 0} \frac{\psi_{FB}^2(0, \Delta y) - \psi_{FB}^2(0, 0)}{\Delta y} = \frac{2\,(\Delta y^2) - 2\sqrt{\Delta y^2} \cdot \Delta y}{\Delta y} = 2(\Delta y - |\Delta y|) = 0.$$

where Δx, $\Delta y > 0$. *Both partial derivatives of* ψ_{FB}^2 *at* $(0, 0)$ *are continuous,* ψ_{FB}^2 *is continuously differentiable. Hence,* $\theta_{FB}^{\mathrm{SMixCP}}(x^*, u^*, t^*, \omega)$ *is continuously differentiable on* $\mathbb{R}^k \times \mathbb{R}^\ell \times \mathbb{R}$ *if and only if* $\widetilde{F}_1(x^*, u^*, t^*, \omega)$ *and* $\widetilde{F}_2(x^*, u^*, t^*, \omega)$ *are continuously differentiable on* $\mathbb{R}^k$ *and* $\mathbb{R}^\ell \times \mathbb{R}$, *respectively.* □

Next we focus on the convexity of the merit function. The function $\psi_{FB}^2(a, b)$ is not convex on $\mathbb{R}^2$, which implies that the merit function $\theta_{FB}^{\mathrm{SMixCP}}(x, u, t, \omega)$ is not convex on its feasible region.

In case the merit function is not convex, Theorem 2.2.3 will be helpful for finding the solution to the minimisation problem for a non-convex merit function. By Theorem 2.2.3, a FB regular point $\begin{pmatrix} x^* \\ u^* \\ t^* \end{pmatrix}$ is a solution to the stochastic mixed complementarity problem $\mathrm{SMixCP}(\widetilde{F}_1, \widetilde{F}_2, \mathbb{R}^k_+, \omega)$, if:

$$\nabla \theta_{FB}^{\mathrm{SMixCP}}(x^*, u^*, t^*, \omega) = 0 \quad \omega \in \Omega, \quad a.s. \tag{3.12}$$

That is

$$\mathcal{A}(\omega)^\top \mathbb{F}_{FB}^{\mathrm{SMixCP}}(x^*, u^*, t^*, \omega) = 0 \quad \omega \in \Omega, \quad a.s., \tag{3.13}$$

where

$$\mathcal{A} = \begin{pmatrix} D_a + D_b J_x \widetilde{F}_1(x^*, u^*, t^*, \omega) & D_b J_{\binom{u}{t}} \widetilde{F}_1(x^*, u^*, t^*, \omega) \\ J_x \widetilde{F}_2(x^*, u^*, t^*, \omega) & J_{\binom{u}{t}} \widetilde{F}_2(x^*, u^*, t^*, \omega) \end{pmatrix}$$

is a nonsingular matrix. Combining equation (3.13) with equation (3.10) implies that equation (3.12) is a necessary condition for $\begin{pmatrix} x^* \\ u^* \\ t^* \end{pmatrix}$ to be a solution to SMixCP$(\widetilde{F}_1, \widetilde{F}_2, \mathbb{R}^k_+, \omega)$.

The feasible set of SMixCP$(\widetilde{F}_1, \widetilde{F}_2, \mathbb{R}^k_+, \omega)$ shrinks as $|\Omega|$ (i.e., the size of the possible outcome set Ω) increases. When $|\Omega| = \infty$, we cannot generally find a solution to the problem SMixCP$(\widetilde{F}_1, \widetilde{F}_2, \mathbb{R}^k_+, \omega)$ such that system (3.12) holds almost surely for any $\omega \in \Omega$, because there will be a large number of equations in system (3.12). Figure 3.2 shows the situation when the size of Ω.

As it is introduced above, probability models provide appropriate deterministic reformulations of the stochastic complementarity problems. It can be associated with the stochastic complementarity problems to find persuasive solutions. These persuasive solutions to stochastic complementarity problems would make a proper trade-off between the satisfaction of infinite complementarity constraints and solvability of the problems.

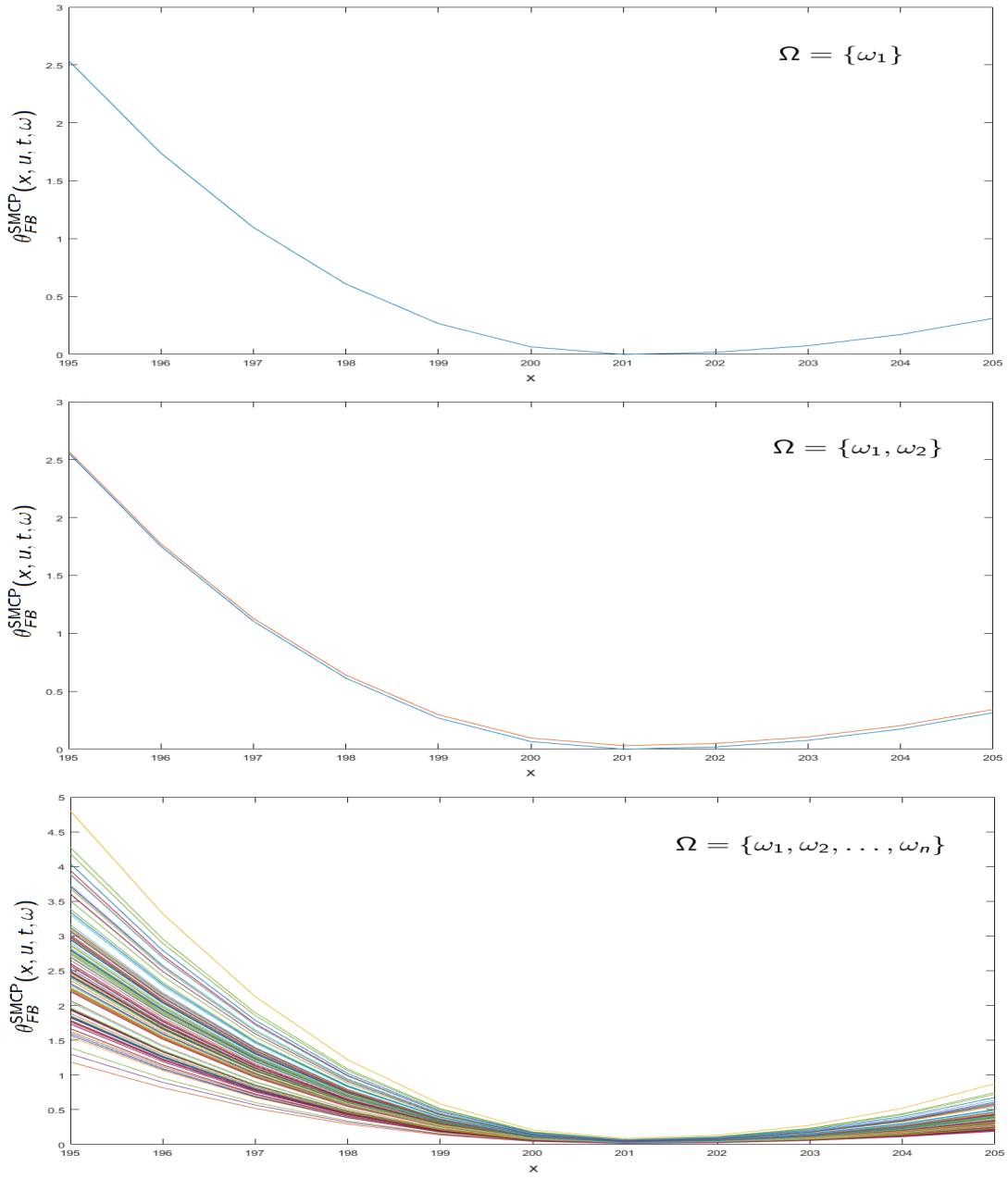
Since $\theta^{\text{SMixCP}}_{FB}(x, u, t, \omega) \geq 0$, given a confidence level $(1 - \alpha) \in (0, 1)$, a point $\begin{pmatrix} x^* \\ u^* \\ t^* \end{pmatrix}$ is a plausible solution to SMixCP$(\widetilde{F}_1, \widetilde{F}_2, \mathbb{R}^k_+, \omega)$ if

$$\begin{pmatrix} x^* \\ u^* \\ t^* \end{pmatrix} \in \arg\min_{x,u,t} \{\Theta | \mathcal{P}\{\theta^{\text{SMixCP}}_{FB}(x, u, t, \omega) \leq \Theta\} \geq 1 - \alpha\}, \tag{3.14}$$

This is a relaxation of problem (3.12). A small value of α means that the satisfaction of the complementarity constraints is preferred to solvability of the problem. A large value of α means that the solvability of the problem is preferred, rather than the satisfaction of the complementarity constraints. Note that the problem (3.14) can be written as:

$$\begin{pmatrix} x^* \\ u^* \\ t^* \end{pmatrix} \in \arg\min_{x,u,t} \{\Theta | \mathbb{E}[\mathbf{1}_{[0,+\infty)} \left(\theta^{\text{SMixCP}}_{FB}(x, u, t, \omega) - \Theta\right)] \leq \alpha\}. \tag{3.15}$$

However, the indicator function $\mathbf{1}_{[0,+\infty)}(\cdot)$ is neither convex nor continuously differentiable at the point 0. Hence, even though the function $\theta^{\text{SMixCP}}_{FB}(\cdot)$ is convex and continuously differentiable, the objective function (3.15) is non-smooth. If we use the indicator function in the objective function, difficulties occur when applying algorithms which are only viable for



Note: For a possible outcome set Ω, when the size of Ω equals 1, i.e. $|\Omega| = 1$ (figure 1), we can easily find a solution (the point when the merit function $\theta_{FB}^{\text{SMixCP}}(x, u, t) = 0$) to the problem by using the merit function. When $|\Omega|$ increases to 2 (figure 2), the solution for the first case is not longer suitable for both outcomes. As the size of $|\Omega|$ increases (figure 3), it become almost impossible to find a solution to the problem which is suitable for any outcomes.

Figure 3.2: The minimum point of merit function varies $\theta_{FB}^{\text{SMixCP}}$ as $|\Omega|$ increases

smooth objective functions. Addressing this concern, the CVaR method will be considered, which undertakes convex and continuously differentiable objective functions. It harmonises the incompatibility between the satisfaction of infinite number of complementarity constraints and solvability of the problems, as well as inherits convexity (it is not globally convex, but at least it may be convex on some neighbourhoods) and continuous differentiability from the merit function $\theta_{FB}^{\text{SMixCP}}(x, u, t, \omega)$. In the CVaR method, $(\theta_{FB}^{\text{SMixCP}}(x, u, t, \omega) - \Theta)$ will be used as the "loss

function" to measure the "loss" of complementarity. It should be emphasised that, the higher the value of the "loss function", the more complementarity constraints of this stochastic complementarity problem are lost. We will transform (3.14) into CVaR based objective function and then construct the stochastic programming model in the following context.

Rewritting (3.15) as Value-at-Risk (VaR) to measure of the loss of complementarity:

$$\begin{pmatrix} x^* \\ u^* \\ t^* \end{pmatrix} \in \{\begin{pmatrix} x \\ u \\ t \end{pmatrix})|VaR_\alpha \left(\theta_{FB}^{\mathrm{SMixCP}}(x,u,t,\omega) - \Theta\right) \leq 0\}.$$

VaR is a measure of complementarity loss defined in (3.5). However, the disadvantages of using VaR as the measure of complementarity loss is significant: VaR is not consistent, which means that it is neither convex nor smooth [5]. On the other hand, CVaR (defined in (3.7)) has superior mathematical properties outperforming VaR, as it inherits continuous differentiability and (local) convexity from the merit function. Moreover, CVaR is a more conservative measure of complementarity loss than VaR.

Theorem 3.1.2

If $\theta_{FB}^{\mathrm{SMixCP}}(x,u,t,\omega)$ is continuously differentiable on $\mathbb{R}^k \times \mathbb{R}^\ell \times \mathbb{R}$, then for any $0 < \alpha < 1$, the measure of complementarity loss $CVaR_\alpha(\theta_{FB}^{\mathrm{SMixCP}}(x,u,t,\omega))$ is continuously differentiable on $\mathbb{R}^k \times \mathbb{R}^\ell \times \mathbb{R}$.

Proof:

Immediate from the continuous differentiability of $\theta_{FB}^{\mathrm{SMixCP}}(x,u,t,\omega)$ and (3.7). □

Theorem 3.1.3

If $\theta_{FB}^{\mathrm{SMixCP}}(x,u,t,\omega)$ is convex on a neighbourhood $\mathcal{S} \subset \mathbb{R}^k \times \mathbb{R}^\ell \times \mathbb{R}$, then for any $0 < \alpha < 1$, the measure of complementarity loss $CVaR_\alpha(\theta_{FB}^{\mathrm{SMixCP}}(x,u,t,\omega))$ is also convex on $\mathcal{S}$.

Proof:

Denote z, $z' \in \mathcal{S} \subset \mathbb{R}^k \times \mathbb{R}^\ell \times \mathbb{R}$, suppose that $\theta_{FB}^{\mathrm{SMixCP}}(z,\omega)$ is convex on $\mathcal{S}$, we have

$$\theta_{FB}^{\mathrm{SMixCP}}(\lambda z + (1-\lambda)z', \omega) \leq \lambda\theta_{FB}^{\mathrm{SMixCP}}(z,\omega) + (1-\lambda)\theta_{FB}^{\mathrm{SMixCP}}(z',\omega),$$

where $\lambda \in [0,1]$. *Noting that*

$$\begin{aligned}
&CVaR_\alpha(\theta(\lambda z + (1-\lambda)z', \omega)) \\
&\quad = \frac{1}{\alpha}\int_0^\alpha VaR_\gamma\left(\theta(\lambda z + (1-\lambda)z', \omega)\right) d\gamma \\
&\quad \leq \frac{1}{\alpha}\int_0^\alpha VaR_\gamma\left(\lambda\theta_{FB}^{\text{SMixCP}}(z,\omega) + (1-\lambda)\theta_{FB}^{\text{SMixCP}}(z',\omega)\right) d\gamma \\
&\quad = \frac{1}{\alpha}\int_0^\alpha \left[VaR_\gamma\left(\lambda\theta_{FB}^{\text{SMixCP}}(z,\omega)\right) + VaR_\gamma\left((1-\lambda)\theta_{FB}^{\text{SMixCP}}(z',\omega)\right)\right] d\gamma \\
&\quad = \frac{\lambda}{\alpha}\int_0^\alpha VaR_\gamma\left(\theta_{FB}^{\text{SMixCP}}(z,\omega)\right) d\gamma + \frac{1-\lambda}{\alpha}\int_0^\alpha VaR_\gamma\left(\theta_{FB}^{\text{SMixCP}}(z',\omega)\right) d\gamma \\
&\quad = \lambda CVaR_\alpha(\theta(z,\omega)) + (1-\lambda)CVaR_\alpha(\theta(z',\omega)).
\end{aligned}$$

Hence, $CVaR_\alpha(\theta_{FB}^{\text{SMixCP}}(x,u,t,\omega))$ *is convex on* $\mathcal{S}$. □

Definition 3.1.2 (Conservativeness)

Suppose $S_1(\theta(x,\omega))$, $S_2(\theta(x,\omega)) : \mathbb{R}^n \to \mathcal{S}$ *are two risk measures. Given an outcome* $\omega \in \Omega$, *risk measure* $S_1(\theta(x,\omega))$ *is said to be* more conservative *than risk measure* $S_2(\theta(x,\omega))$ *if*

$$S_1(\theta(x,\omega)) \geq S_2(\theta(x,\omega))$$

for any $x \in \mathbb{R}^n$.

Proposition 3.1.5

For the measuring the complementarity loss of the merit function $\theta_{FB}^{\text{SMixCP}}(x,u,t,\omega)$, *the measure* $CVaR_\alpha(\cdot)$ *is more conservative than the measure* $VaR_\alpha(\cdot)$.

Proof:

By definition (3.7) *we have:*

$$\begin{aligned}
CVaR_\alpha\left(\theta_{FB}^{\text{SMixCP}}(x,u,t,\omega)\right) &= \alpha^{-1}\int_0^\alpha VaR_\tau\left(\theta_{FB}^{\text{SMixCP}}(x,u,t,\omega)\right) d\tau \\
&= \mathbb{E}[VaR_\tau\left(\theta_{FB}^{\text{SMixCP}}(x,u,t,\omega)\right) | 0 \leq \tau \leq \alpha] \\
&\geq \min\{VaR_\tau\left(\theta_{FB}^{\text{SMixCP}}(x,u,t,\omega)\right) | 0 \leq \tau \leq \alpha\} \\
&= VaR_\alpha\left(\theta_{FB}^{\text{SMixCP}}(x,u,t,\omega)\right)
\end{aligned}$$

Hence, we conclude

$$CVaR_\alpha(\theta_{FB}^{\text{SMixCP}}(x,u,t,\omega)) \geq VaR_\alpha(\theta_{FB}^{\text{SMixCP}}(x,u,t,\omega)).$$

□

Reformulate the problem (3.15) to the following CVaR based minimisation problem:

$$\min_{(x,u,t)\in\mathbb{R}^k\times\mathbb{R}^\ell\times\mathbb{R}} CVaR_\alpha(\theta_{FB}^{\text{SMixCP}}(x,u,t,\omega)), \tag{3.16}$$

where

$$CVaR_\alpha(\theta_{FB}^{\text{SMixCP}}(x,u,t,\omega)) = \alpha^{-1}\int_0^\alpha VaR_\gamma\left(\theta_{FB}^{\text{SMixCP}}(x,u,t,\omega)\right)d\gamma,$$

and

$$VaR_\alpha(\theta_{FB}^{\text{SMixCP}}) = min\{\Theta|\mathcal{P}[\theta_{FB}^{\text{SMixCP}}(x,u,t,\omega)\geq\Theta]\leq\alpha\}.$$

It means that a solution $\begin{pmatrix} x^* \\ u^* \\ t^* \end{pmatrix}$ to SMixCP should minimise the "loss" of complementarity from stochasticity.

Let

$$[t]_+ := \max\{0,t\},$$

$$\nu_{(\Theta,\alpha)}(x,u,t,\omega) := \Theta + \alpha^{-1}[\theta_{FB}^{\text{SMixCP}}(x,u,t,\omega)-\Theta]_+,$$

and define

$$\mathcal{N}_\alpha(x,u,t,\omega,\Theta) := \mathbb{E}\left[\nu_{(\Theta,\alpha)}(x,u,t,\omega)\right] = \Theta + \alpha^{-1}\mathbb{E}[\theta_{FB}^{\text{SMixCP}}(x,u,t,\omega)-\Theta]_+.$$

Lemma 3.1.1

The problem (3.16) *is equivalent to the following problem:*

$$\min_{(x,u,t)\in\mathbb{R}^k\times\mathbb{R}^\ell\times\mathbb{R}} \quad \mathcal{N}_\alpha(x,u,t,\omega,\Theta^*) \tag{3.17}$$

where Θ^* *is the optimal value satisfying:*

$$\Theta^* \in \arg\min_{\Theta\in\mathbb{R}}\left\{\mathcal{N}_\alpha(x,u,t,\omega,\Theta)\right\}.$$

Proof:

Immediate from the alternative definition of CVaR [90]:

$$CVaR_{\alpha}(\theta_{FB}^{\mathrm{SMixCP}}(x,u,t,\omega)) := \inf_{\Theta\in\mathbb{R}} \left\{\Theta + \alpha^{-1}\mathbb{E}[\theta_{FB}^{\mathrm{SMixCP}}(x,u,t,\omega) - \Theta]_{+}\right\}.$$

□

Problem (3.17) simplifies (3.16) because it does not contain integration, and inherits the local convexity from the merit function $\theta_{FB}^{\mathrm{SMixCP}}(x,u,t,\omega)$. However, since the presence of the operator $[\cdot]_{+}$, the objective function in problem (3.17) is not smooth at the point 0. Using mathematical techniques to smooth the objective function can make continuation method applicable on this problem [12]. Chen and Harker [11] summarised four palmary smoothing functions. They are provided as follows:

(i) Neural network smoothing function:

$$p(t,\mu) = t + \mu\log(1 + e^{-\frac{t}{\mu}}).$$

(ii) Interior point smoothing function:

$$p(t,\mu) = \frac{t + \sqrt{t^2 + 4\mu}}{2}.$$

(iii) Auto-scaling interior point smoothing function:

$$p(t,\mu) = \frac{t + \sqrt{t^2 + 4\mu^2}}{2} + \mu.$$

(iv) Chen-Harker-Kanzow-Smale (CHKS) smoothing function:

$$p(t,\mu) = \frac{t + \sqrt{t^2 + 4\mu^2}}{2}.$$

where $\mu \geq 0$ is the parameter of the approximation function p. It should be noted that:

$$\lim_{\mu\to+0} p(t,\mu) = [t]_{+}.$$

In this study, we choose Chen-Harker-Kanzow-Smale (CHKS) smoothing function and denote:

$$[t]_\mu = \frac{t + \sqrt{t^2 + 4\mu^2}}{2}.$$

We rewrite problem (3.17) as:

$$\min_{(x,u,t)\in\mathbb{R}^k\times\mathbb{R}^\ell\times\mathbb{R},\Theta\in\mathbb{R}} \quad \mathcal{N}_\alpha(x,u,t,\omega,\Theta) = \Theta + \alpha^{-1}\mathbb{E}[\theta_{FB}^{\mathrm{SMixCP}}(x,u,t,\omega) - \Theta]_\mu$$

The mathematical expectation is another difficulty that needs to be carefully treated. In many instances, the mathematical expectation $\mathbb{E}[\cdot]$ cannot be calculated with accuracy. A common treatment is using the Sample Average Approximation (SAA) method, which is based on the Law of large numbers. SAA method provides a persuasive result of measuring an expectation value [37, 48]. If the distribution of the random vector ω is known, then the Monte-Carlo approach can be used to generate a sample independently and identically distributed (i.i.d.) $\{\omega^1, \ldots, \omega^N\}$ with the distribution of ω. Let $\{\omega^1, \ldots, \omega^N\}$ be an i.i.d. sample set. The SAA method estimates the mathematical expectation $\mathbb{E}[\theta_{FB}^{\mathrm{SMixCP}}(x,u,t,\omega) - \Theta]_\mu$ using averaged value of all observations $[\theta_{FB}^{\mathrm{SMixCP}}(x,u,t,\omega^1) - \Theta]_\mu$, $[\theta_{FB}^{\mathrm{SMixCP}}(x,u,t,\omega^2) - \Theta]_\mu$, $\ldots$, $[\theta_{FB}^{\mathrm{SMixCP}}(x,u,t,\omega^N) - \Theta]_\mu$. That is,

$$\begin{aligned}\hat{\mathcal{N}_\alpha}^N(x,u,t,\Theta) &:= \frac{1}{N}\sum_{i=1}^N \mathcal{N}_\alpha(x,u,t,\omega^i,\Theta) \\ &= \Theta + \alpha^{-1}\frac{1}{N}\sum_{i=1}^N [\theta_{FB}^{\mathrm{SMixCP}}(x,u,t,\omega^i) - \Theta]_\mu.\end{aligned}$$

Then, problem (3.17) becomes

$$\begin{aligned}\min_{(x,u,t)\in\mathbb{R}^k\times\mathbb{R}^\ell\times\mathbb{R},\Theta\in\mathbb{R}} \quad & \hat{\mathcal{N}}_\alpha(x,u,t,\Theta) = \Theta + \alpha^{-1}\tfrac{1}{N}\sum_{i=1}^N [\theta_{FB}^{\mathrm{SMixCP}}(x,u,t,\omega^i) - \Theta]_\mu \\ s.t. \quad & (x,u,t) \in \mathbb{R}^k \times \mathbb{R}^\ell \times \mathbb{R}, \quad \Theta \in \mathbb{R}.\end{aligned} \tag{3.18}$$

The gradient of $\hat{\mathcal{N}_\alpha}(x,u,t,\Theta)$ is:

$$\nabla\hat{\mathcal{N}_\alpha}(x,u,t,\Theta) = \begin{pmatrix} \nabla_{x,u,t}\mathcal{N}_\alpha(x,u,t,\Theta) \\ \nabla_\Theta\mathcal{N}_\alpha(x,u,t,\Theta) \end{pmatrix}$$

where

$$\nabla_{x,u,t}\hat{\mathcal{N}}_\alpha(x,u,t,\Theta) = \alpha^{-1}\frac{1}{2N}\sum_{j=1}^{N}\left[1+\frac{\theta_{FB}^{\mathrm{SMixCP}}(x,u,t,\omega^j)-\Theta}{\sqrt{\left(\theta_{FB}^{\mathrm{SMixCP}}(x,u,t,\omega^j)-\Theta\right)^2+4\mu}}\right]\mathcal{A}_j^\top \mathbb{F}_{FB}^{\mathrm{SMixCP}}(x,u,t,\omega^j), \tag{3.19}$$

$$\mathcal{A}_j = \begin{pmatrix} D_{a,j}+D_{b,j}\widetilde{A}_j & D_{b,j}\widetilde{B}_j \\ \widetilde{C}_j & \widetilde{D}_j \end{pmatrix}, \tag{3.20}$$

$$D_{a,j} = diag\left(\frac{x_i}{\sqrt{(x_i)^2+\widetilde{F}_1^i(x,u,t,\omega^j)^2}}-1\right) \qquad D_{b,j} = diag\left(\frac{\widetilde{F}_1^i(x,u,t,\omega^j)}{\sqrt{(x_i)^2+\widetilde{F}_1^i(x,u,t,\omega^j)^2}}-1\right), \qquad i\in\{1,\dots,k\},$$

$$\widetilde{A}_j = A(\omega^j), \qquad \widetilde{B}_j = \begin{pmatrix} B(\omega^j) & A(\omega^j)e \end{pmatrix}, \qquad \widetilde{C}_j = \begin{pmatrix} tC(\omega^j)+ue^\top A(\omega^j) \\ 0 \end{pmatrix},$$

$$\widetilde{D}_j = \begin{pmatrix} \left[A(\omega^j)(x+te)+B(\omega^j)u+p(\omega^j)\right]^\top eI+ue^\top B(\omega^j)+tD(\omega^j) & C(\omega^j)x+2tC(\omega^j)e+ue^\top A(\omega^j)e+D(\omega^j)u \\ -2u^\top & 2t \end{pmatrix}$$

and

$$\nabla_\Theta\hat{\mathcal{N}}_\alpha(x,u,t,\Theta) = 1-\alpha^{-1}\frac{1}{N}\sum_{j=1}^{N}\left[\frac{1}{2}+\frac{\theta_{FB}^{\mathrm{SMixCP}}(x,u,t,\omega^j)-\Theta}{2\sqrt{\left(\theta_{FB}^{\mathrm{SMixCP}}(x,u,t,\omega^j)-\Theta\right)^2+4\mu}}\right]. \tag{3.21}$$

Since the objective function $\hat{\mathcal{N}}_\alpha(x,u,t,\Theta)$ is continuously differentiable, Problem (3.18) can be solved by finding some solutions $\begin{pmatrix} x^* \\ u^* \\ t^* \\ \Theta^* \end{pmatrix}$ to

$$\nabla\hat{\mathcal{N}}_\alpha(x,u,t,\Theta) \ = 0. \tag{3.22}$$

3.2 An algorithm

In the previous section, we have modified the SLCP$(T(\omega), r(\omega), L, \omega)$ to the problem (3.18) with a convex and continuously differentiable objective function. Furthermore, the solution to the SLCP$(T(\omega), r(\omega), L, \omega)$ can be obtained by finding some solution $\begin{pmatrix} x^* \\ u^* \\ t^* \\ \Theta^* \end{pmatrix}$ to equation (3.22). In this section, an algorithm will be developed to solve (3.22). Different from the algorithms in Chapter 2, the new algorithm designed in this section involves stochasticity. This algorithm contains Monte-Carlo approach to generate i.i.d. random vector sample sets. We denote $z := \begin{pmatrix} x \\ u \\ t \end{pmatrix} \in \mathbb{R}^{n+1}$. Given the tolerance $r > 0$, stop criterion is that the following

condition is satisfied:

$$\max_i \left\{ \left\| \frac{\partial \mathcal{N}_\alpha^{(N_j,\mu_t)}(x,u,t,\omega,\Theta)}{\partial z_i} \right\| \right\} \leq r, \quad i \in \{1,\ldots,n+1\}. \tag{3.23}$$

It is shown as follows:

Algorithm 3 (Line search smoothing SAA):

Input: initial point $z_0 := \begin{pmatrix} x_0 \\ u_0 \\ t_0 \end{pmatrix} \in \mathbb{R}^k \times \mathbb{R}^\ell \times \mathbb{R}$, $\Theta_0 \in \mathbb{R}$, confidence level α, LM parameter ν, the smoothing parameter μ, maximum iteration number j_{max} for j, k_{max} for k, the sequence of sample set sizes $N_1 < N_2 < \cdots < N_{j_{max}}$, parameters of the approximation ν, μ, the tolerance $r > 0$, $\varepsilon > 0$, and parameters for Wolfe conditions c_1, $c_2 \in (0,1)$.

Step 1: Set $j = 1$.

Step 2: Set the sample size $N = N_j$, and generate i.i.d samples $\{\omega^1,\ldots,\omega^N\}$.

Step 3: If $j > 1$, and $\|z^j - z^{j-1}\| < \varepsilon$, **Stop**.

Step 4: Set $k = 0$, and $y_0 = z_0$.

Step 5: If either (3.23) or $k = k_{max}$, then set $j = j+1$, $z_j = y_k$, and go to **Step 3**.

Step 6: Denote $\bar{\mathcal{A}}_j = \frac{1}{N}\sum_{i=1}^{N_j} \mathcal{A}_i$, and find a direction $d_k \in \mathbb{R}^k \times \mathbb{R}^\ell \times \mathbb{R}$ such that

$$\bar{\mathcal{A}}_j(y_k)^\top \mathbb{F}_{FB}^{\mathrm{SMixCP}}(y_k) + \left[\bar{\mathcal{A}}_j{}^\top(y_k) \bar{\mathcal{A}}_j(y_k) + \mu \mathbb{I} \right] d_k = 0. \tag{3.24}$$

If the system (3.24) is not solvable or if the condition

$$\nabla \hat{\mathcal{N}}_\alpha(y_k,\Theta)^\top d_k \leq -r\|d_k\|$$

is not satisfied, (re)set $d_k = -\nabla \hat{\mathcal{N}}_\alpha(y_k,\Theta)$.

Step 7: Find step length $s_k \in R_+$ such that

$$\hat{\mathcal{N}}_\alpha(y_k + s_k d_k,\Theta) \leq \hat{\mathcal{N}}_\alpha(y_k + s_k d_k,\Theta) + c_1 s_k \nabla \hat{\mathcal{N}}_\alpha(y_k,\Theta)^\top d_k,$$

and

$$\hat{\mathcal{N}}_\alpha(y_k + s_k d_k,\Theta)^\top d_k \geq c_2 \nabla \hat{\mathcal{N}}_\alpha(y_k,\Theta)^\top d_k.$$

Step 8: Set $y_{k+1} := y_k + s_k d_k$ and $k := k+1$, go to **Step 5**.

Comment: This algorithm requires the Monte-Carlo approach to generate i.i.d. random vector samples. For a $j \in \{1, \ldots, j_{max}\}$, if the value of N_j is large, the algorithm is anticipated to be more accurate, but it will sacrifice time and computing power. On the other hand, if the value of N_j's is small, the costs of finding result is relatively low, but the accuracy of the solution is sacrificed.

3.3 A numerical example

This section illustrates a numerical example for the stochastic ESOCLCP. Denote by $L(3,2)$ an extended second order cone in $\mathbb{R}^3 \times \mathbb{R}^2$. Let $x \in \mathbb{R}^3$ and $u \in \mathbb{R}^2$ be two real vectors. Denote

$$z = \binom{x}{u} \in \mathbb{R}^3 \times \mathbb{R}^2, \quad \hat{z} = \binom{x - \|u\| e}{u} \in \mathbb{R}^3 \times \mathbb{R}^2, \quad \textit{and } \tilde{z} = \begin{pmatrix} x - t \\ u \\ t \end{pmatrix} \in \mathbb{R}^3 \times \mathbb{R}^2 \times \mathbb{R}.$$

Repeat for convenience a stochastic ESOCLCP defined by the extended second order cone $L(3,2)$ and a stochastic linear function $F(x,u,\omega) = T(\omega)\binom{x}{u} + r(\omega)$ is:

$$SLCP(T(\omega), r(\omega), L(3,2)) \begin{cases} \textit{Find } x \in L(3,2), \textit{ such that} \\ T(\omega)x + r(\omega) \geq 0, x^\top (T(\omega)x + r(\omega)) = 0, \ \omega \in \Omega, \quad a.s., \end{cases}$$

where

$$T = \begin{pmatrix} A & B \\ C & D \end{pmatrix} = \begin{pmatrix} 41+\omega_1 & -3 & -31 & 18 & 19 \\ 28 & 22 & -33 & 25 & -29 \\ -23 & -29 & 11 & -21 & -43 \\ -9 & -31 & -20+2\omega_2 & -12 & 47 \\ -8 & 46 & 50 & -22 & 21 \end{pmatrix}, \quad r = \begin{pmatrix} p \\ q \end{pmatrix} = \begin{pmatrix} -26 \\ 4-\omega_3 \\ 23 \\ 44 \\ -19 \end{pmatrix},$$

with $A \in \mathbb{R}^{3\times 3}$, $B \in \mathbb{R}^{3\times 2}$, $C \in \mathbb{R}^{2\times 3}$, $D \in \mathbb{R}^{2\times 2}$, $p \in \mathbb{R}^3$, and $q \in \mathbb{R}^2$. $\omega = (\omega_1, \omega_2, \omega_3)^\top \in \Omega$ is a stochastic vector with i.i.d. random variables $\omega_i \sim N(0,1)$ for any $i \in \{1,2,3\}$. It is easy to verify that square matrices T, A and D are nonsingular for any outcome of ω_i in $\mathbb{R}$, $i \in \{1,2,3\}$.

By using Theorem 3.1.1, we reformulate $SLCP(T(\omega), r(\omega), L(3,2))$ to a SMixCP defined

by $\widetilde{F}_1$, $\widetilde{F}_2$, and $\mathbb{R}^3_+$:

$$\text{SMixCP}(\widetilde{F}_1, \widetilde{F}_2, \mathbb{R}^3_+, \omega): \begin{cases} Find \ \begin{pmatrix} x \\ u \\ t \end{pmatrix} \in \mathbb{R}^3 \times \mathbb{R}^2 \times \mathbb{R}, \ such \ that \\ \widetilde{F}_2(x,u,t,\omega) = 0, \ and \ (x, \widetilde{F}_1(x,u,t,\omega)) \in \mathcal{C}(\mathbb{R}^3_+), \ \omega \in \Omega, \quad a.s. \end{cases}$$

where

$$\widetilde{F}_1(x,u,t,\omega) = A(\omega)(x+te) + B(\omega)u + p(\omega)$$

and

$$\widetilde{F}_2(x,u,t,\omega) = \begin{pmatrix} \left[tC(\omega) + ue^\top A(\omega)\right](x+te) + ue^\top \left[B(\omega)u + p(\omega)\right] + t\left[D(\omega)u + q(\omega)\right] \\ t^2 - \|u\|^2 \end{pmatrix}.$$

We will convert this to the form of (3.18) and then (3.22). Given $\alpha = 0.05$, we rewrite problem (3.18) as:

$$\min_{(x,u,t)\in\mathbb{R}^3\times\mathbb{R}^2\times\mathbb{R}, \Theta\in\mathbb{R}} \Theta + 0.05^{-1}\frac{1}{N}\sum_{i=1}^{N}[\theta_{FB}^{\text{SMixCP}}(x,u,t,\omega^i) - \Theta]_\mu,$$

where

$$\theta_{FB}^{\text{SMixCP}}(x,u,t,\omega) = \frac{1}{2}\sum_{i=1}^{3}\psi_{FB}^2\left(x_i, \widetilde{F}_1^i(x,u,t,\omega)\right) + \frac{1}{2}\sum_{j=1}^{2}\widetilde{F}_2^j(x,u,t,\omega).$$

Since the distribution of the random vector ω is known, we use the Monte Carlo (MC) method to simulate j_{max} sample sets with number of observation $N_1, N_2, \ldots, N_{j_{max}}$. The solutions are shown in the following table:

The average loss of complementarity (ALoC) is calculated by:

$$ALoC = \frac{1}{N_j}\sum_{i=1}^{N_j}\|(x,u)^\top F(x,u,\omega_i)\|.$$

As it is shown in the table, the solution converges to $(1.546, 0.261, 1.059, 0.125, -0.254)^\top$ as the value of N_j increases. As the value of N_j increases, the computation time increases as well. However, the Average loss of complementarity and the value of threshold Θ remains unchanged. It means that it may not be necessary to set a large N_j for the algorithm to get a precise solution.

j	N_j	$\binom{x}{u}^\top$	$F(x,u,\omega)^\top$
1	10	(1.537, 0.273, 1.060, 0.136, -0.262)	(0.784, 29.054, -0.194, -13.466, 25.803)
2	100	(1.542, 0.263, 1.058, 0.127, -0.253)	(1.093, 28.552, -0.214, -12.609, 25.544)
3	1000	(1.549, 0.257, 1.060, 0.122, -0.252)	(1.277, 28.397, -0.162, -12.418, 25.477)
4	10000	(1.548, 0.262, 1.060, 0.125, -0.254)	(1.215, 28.605, -0.204, -12.701, 25.578)
5	100000	(1.546, 0.261, 1.059, 0.125, -0.254)	(1.186, 28.587, -0.176, -12.643, 25.516)
6	1000000	(1.546, 0.261, 1.059, 0.124, -0.254)	(1.200, 28.566, -0.177, -12.617, 25.514)

j	N_j	Computation time (sec)	Average loss of complementarity	Threshold Θ
1	10	0.090439	0.347	0.063
2	100	0.696431	0.893	0.095
3	1000	5.202383	1.179	0.090
4	10000	39.39705	1.060	0.087
5	100000	553.4596	1.054	0.088
6	1000000	4759.294	1.073	0.089

Note: The first table shows the solutions to $SLCP(T(\omega), r(\omega), L(3,2))$ and the value of the function $F(x,u,\omega)^\top$ with respect to different value of N. The value of solution does not variate significantly, while the value of the function differs but converges to around 1.200 as the value of N increase. The second table shows the computation time (in second), average loss of complementarity, and the value of threshold. The run time increases significantly along with the value N increases. On the other hand, the average loss of complementarity and the value of threshold remains relative constant no matter what change to the value of N.

Table 3.1: The result of the numerical example

3.4 Conclusions and comments

In this chapter, we study the stochastic linear complementarity problem on extended second order cones (stochastic ESOCLCP) which is a stochastic extension of ESOCLCP studied in Chapter 2. Based on Theorem 2.1.1 we derive Theorem 3.1.1, then we can rewrite an stochastic ESOCLCP to a stochastic mixed complementarity problems (stochastic MixCP) on the nonnegative orthant. Enlightened by the idea from [15], we introduce the CVaR method to measure the loss of complementarity in the stochastic case. In contrast to the merit function in the deterministic case (2.22), the merit function (3.12) is not required to equal zero almost surely for any $\omega \in \Omega$. Instead, a CVaR-based minimisation problem (3.16) is introduced to obtain a solution which is "good enough" for the complementarity requirement of the original SMixCP. For solving the CVaR-based minimisation problem derived from the original SMixCP, smoothing function and sample average approximation methods are introduced and finally converted to the form in (3.18). Finally, a line search smoothing SAA algorithm is provided for finding the solution to this CVaR-based minimisation problem and it is illustrated by a numerical example.

Stochastic methods on complementarity problems were pioneered by Chen and Fukushima

[14]. They introduced the idea of minimising the square norm of the merit function to solve a stochastic complementarity problem (SCP). This approach is commonly used in many researches [15, 68, 102, 111]. However, this approach led to non-convexity and consequently increased the difficulty of solving SCP by algorithms. Our algorithm introduced in this chapter only guarantees a stationary point rather than a solution to the problem. The improvement of the process of finding solutions to a stochastic ESOCLCP will be considered as a good topic of our future research.

CHAPTER 4

APPLICATION: PORTFOLIO OPTIMISATION PROBLEMS

4.1 Introduction to portfolio optimisation problem

In this chapter, we will apply the results obtained from Chapter 2 on the portfolio optimisation problem. The foundation of the mathematical formulation of portfolio optimisation problem is established by the pioneering paper of Markowitz [72]. His mean variance model (MV) model is a typical quadratic optimisation programming problem. Also, the Karush-Kuhn-Tucker (KKT) condition of this problem is a complementarity problem on nonnegative orthant. The mean-variance (MV) scheme of portfolio optimisation sets up a classic framework for the research of portfolio optimisation. Among numerous models developed based on the MV model, the mean-absolute deviation (MAD) model attracted our particular interest. The MAD model is introduced by [58] as a route to solve large-scale portfolio optimisation problems. Instead of using covariance matrix, the MAD model uses the absolute-deviation of the rate of return as a measure of the risk. From a mathematical point of view, using absolute-deviation as the measurue of the risk is almost equivalent to the way of using covariance. However, the MAD model significantly reduce the computational cost [56, 57]. The KKT condition of the MAD model is a complementarity problem on second order cone.

Both of these two models contribute to the development of the research of portfolio optimisation problem. However, multiple papers challenge their major drawbacks. Though the MV model requires only the mean values and the covariance matrix of asset return, it still become very computationally expensive when a large number of assets are considered in the

optimisation. In addition, the optimal solution derived from MV model turns out to be highly concentrated in just a few assets, which usually means they are not sufficiently diversified. Parameter sensitivity is another drawback of the MV model. The optimal solution is highly sensitive to its parameter, i.e. the asset returns' mean values and its correlation matrix [60]. The influences of parameter sensitivity on the final result are hard to be rescued because of the inevitable appearance of estimation noises and measurement errors. Hence, optimising a portfolio with the basic MV model will be undiversified and inaccurate, which brings considerable limitations to the actual application of the model. Purely using this optimal allocation on investment decision without amendment often causes a poor out-of-sample portfolio performance. Compare to the MV model, the MAD model has a lower computational cost. However, the modulus in the absolute deviation still lead to some difficulty in computation. The Lagrange function of MAD model is semi-smooth. Unlike the MV model, the MAD model does not have an analytical solution.

Enlightened by the theorems developed in previous chapters, we introduce a portfolio optimisation model based on the MV model and the MAD model: the *Mean-Euclidean Norm (MEN) model*. We find the analytical solution to this model.

4.1.1 Review of mean-variance model and mean-absolute deviation model

The MV model considers a single-period investment. Assume that an investor is in a market with n assets to be considered. Let $\tilde{r} \in \mathbb{R}^n$ denote the random column vector of asset returns in a certain period. We use $r = \mathbb{E}[\tilde{r}] \in \mathbb{R}^n$ to denote the mean return vector of the assets, where $\mathbb{E}[\cdot]$ represents the expected value of the random variable in the square bracket. Suppose that this investor has wealth X at the beginning of the period and he would like to invest all his wealth in these n assets. Let $x \in \mathbb{R}^n$ denote the vector of wealth the investor put in these n assets and $\sum_{i=1}^{n} x_i = X$. At the end of the period, the expectation of final wealth X' of the investor will be:

$$\mathbb{E}[X'] = (e + r)^\top x,$$

where $e = (1, 1, \ldots, 1)^\top \in R^n$. $w_i = \frac{x_i}{X}, i \in \{1, \ldots, n\}$ denote the weight of wealth invested in asset i, so $e^\top w = 1$. The expected rate of return R_p of investor's portfolio P will be:

$$R_p = (e + r)^\top w - 1 = r^\top w. \tag{4.1}$$

Since the rate of return vector on assets is random, the investor cannot be sure that how much his wealth will be at the end of the investment period. Hence, he needs to measure the risk of portfolio returns. The variance of the asset rate of return is commonly used as a surrogate for risk. Let $\sigma_{ij} = cov(R_i, R_j), i, j \in \{1, \ldots, n\}$ denote the estimated covariance between ith and jth asset returns in a certain period. Hence, the covariance matrix is $\Sigma = (\sigma_{ij})$. The risk σ_p^2 of the portfolio P is

$$\sigma_p^2 = w^\top \Sigma w \tag{4.2}$$

For each unit of wealth he invested, the investor either hopes to earn at least γ profit (i.e., the constraint of minimum rate of return), or can only tolerate a risk not exceeding σ (i.e., the constraint of maximum risk). Hence, we obtain the following two equivalent mean variance optimisation formulations according to Markowitz's [72]:

$$\begin{aligned} \max_w \quad & r^\top w \\ s.t. \quad & w^\top \Sigma w \leq \sigma \\ & e^\top w = 1 \end{aligned}$$

$$\begin{aligned} \min_w \quad & w^\top \Sigma w \\ s.t. \quad & r^\top w \geq \gamma \\ & e^\top w = 1. \end{aligned} \tag{4.3}$$

The problem (4.3) is a quadratic optimisation problem. The Karush-Kuhn-Tucker (KKT)

conditions of it can be written as:

$$\begin{cases} 2\Sigma w - \lambda r - \mu e = 0 \\ r^\top w - \gamma \geq 0, \lambda \geq 0, e^\top w = 1, \\ \lambda^\top (r^\top w - \gamma) = 0 \end{cases}$$

for some λ and μ. Noting that $(\lambda, r^\top w - \gamma) \in \mathcal{C}(\mathbb{R}_+)$. So this KKT condition is a mixed complementarity problem on nonnegative orthant. Since the covariance matrix Σ is a symmetric and positive semi-definite matrix, if there is a vector $(w^*, \lambda^*, \mu^*)^\top$ satisfies the KKT condition above, w^* will be a solution to problem (4.3).

An alternative formulation of (4.3) is to include both risk and return in the objective function by using the Arrow-Pratt absolute risk -aversion index $c_0 > 0$ [49]. Such inclusion can be deemed as trading risks against return. Problem (4.3) is reformulated as:

$$\begin{aligned} \min_{y,w} \quad & c_0 y - r^\top w \\ s.t. \quad & y \geq w^\top \Sigma w \\ & e^\top w = 1. \end{aligned} \tag{4.4}$$

The solution to problem (4.4) is:

$$w = (2c_0\Sigma)^{-1} \left(r - \frac{e^\top \Sigma^{-1} r}{e^\top \Sigma^{-1} e} e\right) + \frac{\Sigma^{-1} e}{e^\top \Sigma^{-1} e}, \tag{4.5}$$

$$y = w^\top \Sigma w.$$

Kallberg and Ziemba [49] showed that different coefficients c_0 can reflects different risk-preference attitudes of an investor. When $c_0 \geq 6$, the objective function reveals strong risk-aversion; when $2 \leq c_0 < 6$, it shows a moderate risk-aversion; whilst $0 < c_0 < 2$, the function reflects a risk-seeking situation.

One major dispute about the MV model is its computational inefficiency. It requires $n(n+1)/2$ covariance coefficients σ_{ij} to be calculated based on the historical data or some

results of scenario models. The calculation will be tedious when solving a large-scale portfolio optimisation problem, say, a portfolio with 500 securities included. This is a reason why MV model has not been extensively applied in practises. A good way to alleviate the computation difficulty is to switch the risk measure from covariance matrix to absolute deviation. Konno and Yamazaki [58] introduced the mean-absolute deviation (MAD) model to reduce the computational cost of the MV model. According to Konno and Yamazaki, if the return is multivariate normally distributed, the MAD model provides similar results with the MV model. The introduction of risk aversion coefficient c_0 makes the models equivalent [93]. The MAD model demonstrated a stronger computability than the MV model as the computational time for solving a linear programming problem will not be substantially increased comparing with that of a quadratic programming problem.

The MAD model outperforms the MV model in many other aspects. The MV model quantifies the portfolio selection into a form with only two criteria: expected returns measured by means and risks measured by covariance. This simple quantification is also criticized as not consistent with any degrees of stochastic dominance [62, 101]. On the other hand, the MAD model depends on a relation of second degree stochastic dominance (SSD) (A portfolio is said to be second-order stochastic dominant of another if this portfolio involves less risk and has at least as high return).

Assume that the rates of returns of assets $\tilde{r} = (\tilde{r}_1, \tilde{r}_2, \ldots, \tilde{r}_n)^\top \in \mathbb{R}^n$ are distributed over a finite (discrete) sequence of points $\{R_j\} = \{(R_{1j}, R_{2j}, \ldots, R_{nj})^\top\} \in \mathbb{R}^n,\ j \in \{1, 2, \ldots, T\}$. That is, there are T different scenarios leads to different outcomes of asset returns. Let $f_j,\ j \in \{1, 2, \ldots, T\}$ denote the probability distribution of the outcomes of the rates of returns of assets:

$$f_j = Pr\{(\tilde{r}_1, \tilde{r}_2, \ldots, \tilde{r}_n)^\top = (R_{1j}, R_{2j}, \ldots, R_{nj})^\top\},\ t = 1, 2, \ldots, T. \tag{4.6}$$

The sequences $\{R_j\}$ and $\{f_j\}$ are acquired through historical data or some techniques of future projection. By definition, it is clear that $\sum_{j=1}^{T} f_j = 1$ and $f_j \geq 0$ for any $j \in \{1, \ldots, T\}$. In particular,

$$r = \mathbb{E}[\tilde{r}] = \sum_{j=1}^{T} f_j R_j \in \mathbb{R}^n. \tag{4.7}$$

Denote $U = (U_1, U_2, \ldots, U_T)^\top$, where $U_j = (R_j - r),\ j \in \{1, \ldots, T\}$. The MAD model is

the following linear programming problem:

$$\begin{aligned} \min_{y,w} \quad & c_0 f^\top y - r^\top w \\ s.t. \quad & y_j \geq |U_j^\top w|, \quad j \in \{1, \ldots, T\} \\ & e^\top w = 1. \end{aligned} \tag{4.8}$$

Denote by $\mathcal{L}$ the second order cone:

$$\mathcal{L} := \{(x, y) \in \mathbb{R} \times \mathbb{R}^n : x \geq \|y\|\}.$$

The KKT condition of Problem (4.8) can be written as the following complementarity problem on second order cone:

$$\mathcal{L} \ni \begin{pmatrix} y_j \\ U_j^\top w \end{pmatrix} \perp \begin{pmatrix} c_0 f_j - \theta_j \\ \left(U_j^{-1}\right)^\top \left(-r + \mu e + \sum_{i=1}^{T} \theta_i \frac{U_i U_i^\top}{|U_i^\top w|} w\right) \end{pmatrix} \in \mathcal{L}, \quad j \in \{1, \ldots, T\},$$

$$e^\top w - 1 = 0,$$

where $\mu \in \mathbb{R}$, and $\theta \in \mathbb{R}^T$ are Lagrangian multipliers. Noting that this KKT condition is a mixed complementarity problem on second order cone $\mathcal{L}$.

Unlike the MV model, the MAD model does not have an analytical solution [9]. If we try to solve for w, we have:

$$w = c_0^{-1} B \left(r - \frac{e^\top B r}{e^\top B e} e\right) + \frac{Be}{e^\top B e} \tag{4.9}$$

where

$$B = \left(\sum_{j=1}^{T} \frac{f_j}{|U_j^\top w|} U_j U_j^\top\right)^{-1}$$

The existence of the modulus in the term $|U_j^\top w|$ implies that (4.9) is not an analytic solution the MAD model. We emphasis that $T \geq n$ is a necessary condition for the matrix B to be nonsingular. It can be easily proved by some basic linear algebra knowledge.

4.1.2 Formulation of the mean-Euclidean norm model

If we slightly modify the constraint of problem (4.8), we get the mean-Euclidean norm model (MEN):

$$\begin{aligned} \min_{y,w} \quad & c_0 f^\top y - r^\top w \\ s.t. \quad & y_j \geq \|U_j\|\|w\| \quad j \in \{1,\ldots,T\} \\ & e^\top w = 1, \end{aligned} \tag{4.10}$$

where $\|\cdot\|$ is the Euclidean norm and $\|w\| = \sqrt{\langle w, w\rangle}$. As a modification of problem (4.8), problem (4.10) has a different feasible set comparing to problem (4.8). The feasible set of problem (4.8) is:

$$\mathcal{F}_3 = \left\{(y,w) : y_j \geq |\left(R_j - r\right)^\top w|,\ and\ e^\top w = 1,\ j \in \{1,\ldots,T\}\right\}, \tag{4.11}$$

whereas the feasible set of problem (4.10) is

$$\mathcal{F}_3' = \left\{(y,w) : y_j \geq \|U_j\|\|w\|,\ and\ e^\top w = 1,\ j \in \{1,\ldots,T\}\right\}. \tag{4.12}$$

The following corollary shows the relationship between these two feasible sets.

Proposition 4.1.1

The feasible set $\mathcal{F}_3'$ (4.12) is a subset of $\mathcal{F}_3$ (4.11).

Proof:

For any $j \in \{1,2,\ldots,T\}$, we have by Cauchy's inequality

$$|U_j^\top w| \leq \|U_j\|\|w\|.$$

Hence, we have $\mathcal{F}' \subseteq \mathcal{F}$. The equation holds ($\mathcal{F}' = \mathcal{F}$) only if U_j and w are linearly dependent. □

Since $\|U_j\| > 0$ for any $j \in \{1,\ldots,T\}$, $\mathcal{F}_3'$ can be written as:

$$\mathcal{F}_3' = \left\{(y,w) : \frac{y_j}{\|U_j\|} \geq \|w\|,\ e^\top w = 1,\ j \in \{1,\ldots,T\}\right\}$$

Comment: The MEN model (4.10) and the MAD model (4.8) have the same objective function and similar feasible sets, but the feasible set of MEN model is the subset of that of MAD model. Unfortunately, the optimal solution to MAD model may be excluded from the feasible set of MEN model. On the other hand, the advantage of MEN model over MAD model is that the former, by applying the Proposition 1.4.3, provides possibility of finding analytical solution. In subsequent we will show in Proposition 4.1.2 how can we use Proposition 1.4.3 to work out an analytical solution to the MEN model.

Recall the definitions of the mutually dual extended second order cone $L(T,n), M(T,n) \in \mathbb{R}^T \times \mathbb{R}^n$ introduced in (1.7), (1.8):

$$L(T,n) = \{\binom{y}{w} \in \mathbb{R}^T \times \mathbb{R}^n : y \geq \|w\|e,\, y \geq 0\},$$

$$M(T,n) = \{\binom{y}{w} \in \mathbb{R}^T \times \mathbb{R}^n : e^\top y \geq \|w\|,\, y \geq 0\},$$

Denoting $U_{\|\cdot\|} = (\|U_1\|, \|U_2\|, \ldots, \|U_T\|)^\top$. We use "$\circ$" to represent the Hadamard product [43]. The KKT condition of Problem (4.10) is:

$$L \ni \begin{pmatrix} y \circ U_{\|\cdot\|}^{-1} \\ w \end{pmatrix} \perp \begin{pmatrix} c_0 U_{\|\cdot\|} \circ f - \theta \\ -r + \mu e + \frac{w}{\|w\|} \sum_{j=1}^{T} \theta_j \end{pmatrix} \in M, \tag{4.13}$$

$$e^\top w - 1 = 0, \tag{4.14}$$

where, $\mu \in \mathbb{R}$, and $\theta \in \mathbb{R}^T$ are Lagrangian multipliers.

KKT condition of the M2LN model is a nonlinear complementarity problem on ESOC. Applying the item (iv) of Proposition 1.4.3 on condition (4.13) with $x = y \circ U_{\|\cdot\|}^{-1}$, $u = w$, $z = c_0 U_{\|\cdot\|} \circ f - \theta$, and $v = -r + \mu e + \frac{w}{\|w\|} \sum_{j=1}^{T} \theta_j$, we obtain the following proposition:

Proposition 4.1.2

If $-r + \mu e + \frac{w}{\|w\|} \sum_{j=1}^{T} \theta_j \neq 0$, *then there exists a parameter* $\lambda \in \mathbb{R}_+$ *such that*

$$-r + \mu e + \frac{w}{\|w\|} \sum_{j=1}^{T} \theta_j = -\lambda w,$$

$$e^\top \left(c_0 U_{\|\cdot\|} \circ f - \theta\right) = \left\| -r + \mu e + \frac{w}{\|w\|} \sum_{j=1}^{T} \theta_j \right\|,$$

and

$$\mathbb{R}_+^T \ni \left(y \circ U_{\|\cdot\|}^{-1} - \|w\| e\right) \perp \left(c_0 U_{\|\cdot\|} \circ f - \theta\right) \in \mathbb{R}_+^T.$$

Remark 4.1.1

Item (i) and item (ii), and item (iii) of Proposition 1.4.3 are inapplicable in the circumstance of finding the solution to problem (4.10)*. In Proposition 1.4.3, item (i) and item (ii) state that* $w = 0$*, which contradicts to condition* (4.14) *as the later requires* $w \neq 0$*. Therefore, item (i) and (ii) are inapplicable.*

Item (iii) and item (iv) are applicable in finding the solution to problem (4.10)*. However, both items have its own limitations. In Proposition 1.4.3, item (iii) assume that*

$$-r + \mu e + \frac{w}{\|w\|} \sum_{j=1}^{T} \theta_j = 0 \tag{4.15}$$

We will use the following proposition to show that the conjecture in (4.15) *is not always appropriate.*

Proposition 4.1.3

With conjecture $-r + \mu e + \frac{w}{\|w\|} \sum_{j=1}^{T} \theta_j = 0$ *we cannot always find a solution to problem* (4.10) *for any* $\{R_j\} \in \mathbb{R}^n$*,* $j \in \{1, \ldots, T\}$*.*

Proof:

Suppose that (4.15) *holds, it can be rewritten to:*

$$w = (r - \mu e) \frac{\|w\|}{\sum_{j=1}^{T} \theta_j}, \tag{4.16}$$

adding (4.14)*, we get*

$$1 = e^\top w = \left(e^\top r - n\mu\right) \frac{\|w\|}{\sum_{j=1}^{T} \theta_j}. \tag{4.17}$$

Combine (4.16) *and* (4.17)*, we conclude that:*

$$w = \frac{r - \mu e}{e^\top r - n\mu}. \tag{4.18}$$

By KKT condition (4.13)*, we have:*

$$\mathbb{R}_+^T \ni \left(y \circ U_{\|\cdot\|}^{-1}\right) \perp \left(c_0 U_{\|\cdot\|} \circ f - \theta\right) \in \mathbb{R}_+^T. \tag{4.19}$$

On the other hand, by the complementarity in (4.19)*, and* $y \circ U_{\|\cdot\|}^{-1} > 0$ *implied by* (4.13) *and* (4.14)*, we have*

$$0 = c_0 \|U_j\| f_j - \theta_j, \quad j \in \{1, \ldots, T\},$$

that is

$$\theta_j = c_0 \|U_j\| f_j, \quad j \in \{1, \ldots, T\}. \tag{4.20}$$

Substitute w *and* θ_j *in* (4.16) *by* (4.18) *and* (4.20)*, respectively, we have*

$$(r - \mu e)\left(1 - \frac{c_0 U_{\|\cdot\|}^\top f}{\|r - \mu e\|}\right) = 0. \tag{4.21}$$

Apparently, $r \neq \mu e$ *because of the nature of asset returns. Let the term in the right bracket of* (4.21) *equals zero. Recalling* $U_{\|\cdot\|}^\top = (\|R_1 - r\|, \ldots, \|R_T - r\|)$*, we get:*

$$\|r - \mu e\| = c_0 U_{\|\cdot\|}^\top f = c_0 \sum_{j=1}^{T} \|R_j - r\| f_j.$$

Solving above equation for μ*, we have*

$$\|r\|^2 - 2\bar{r}\mu + n\mu^2 = \left(c_0 \sum_{j=1}^{T} \|R_j - r\| f_j\right)^2,$$

$$\mu = \bar{r} \pm \sqrt{\frac{\bar{r} - \|r\|^2 + \left(c_0 \sum_{j=1}^{T} \|R_j - r\| f_j\right)^2}{n}}. \tag{4.22}$$

Recall that $\bar{r} = \frac{1}{n} e^\top r$*. From* (4.22) *we can observe that if we pick up some* $\{R_j\} \in \mathbb{R}^n$*,* $j \in \{1, \ldots, T\}$ *such that*

$$\bar{r} + \left(c_0 \sum_{j=1}^{T} \|R_j - r\| f_j\right)^2 \leq \|r\|^2 \tag{4.23}$$

then μ *is not a real number.* □

Example 4.1.1

This example gives a numerical case to show that inequality (4.23) *does not hold for some* $\{R_j\} \in \mathbb{R}^n$, $j \in \{1, \ldots, T\}$. *Given the absolute risk-preference index* $c_0 = 4$, *number of asset class* $n = 3$, *and the number of scenarios* $T = 5$*:*

$$f = \begin{pmatrix} 0.01 \\ 0.14 \\ 0.27 \\ 0.12 \\ 0.46 \end{pmatrix}, \qquad R = \begin{pmatrix} 0.10 & 0.70 & 0.80 & 0.80 & 1.00 \\ 0.30 & 0.80 & 0.60 & 0.40 & 0.70 \\ 0.50 & 0.60 & 0.50 & 0.00 & 0.60 \end{pmatrix},$$

then we can calculate $r = (0.8710, 0.6470, 0.5000)^\top$ *and* $\bar{r} = 0.6727$. *Hence,*

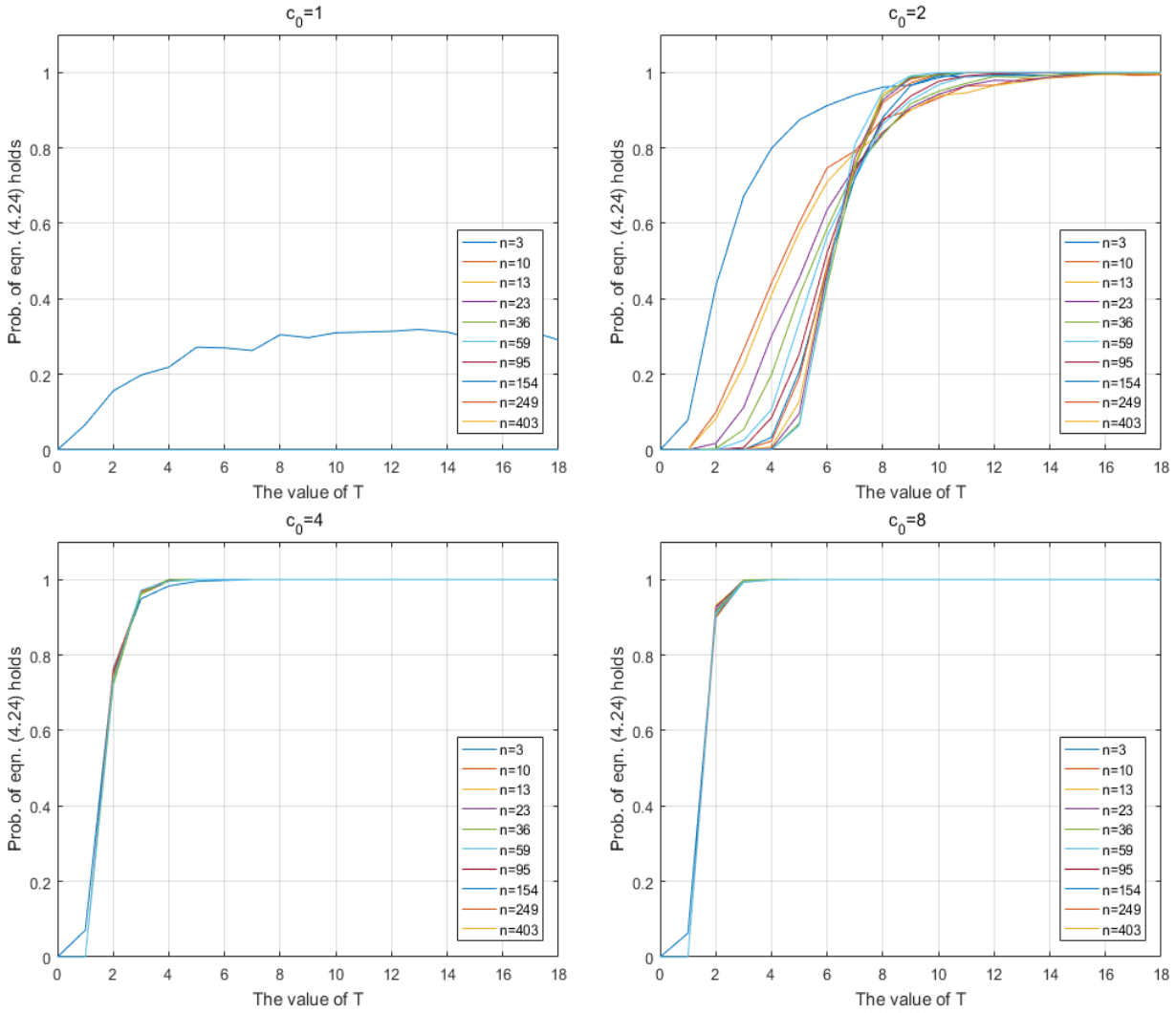
$$\begin{aligned} &\bar{r} + \left(c_0 \sum_{j=1}^{T} \|R_j - r\| f_j \right)^2 - \|r\|^2 \\ =& 0.6727 + 0.7251 - 1.4273 \\ =& -0.0294 < 0 \end{aligned}$$

Proposition 4.1.3 and Example 4.1.1 reveal that the assumptions in Proposition 1.4.3 item (iii) does not always hold. Using this item of Proposition 1.4.3 is appropriate only if:

$$\bar{r} + \left(c_0 \sum_{j=1}^{T} \|R_j - r\| f_j \right)^2 - \|r\|^2 \geq 0 \tag{4.24}$$

For any $\{R_j\} \in \mathbb{R}^n$, $j \in \{1, \ldots, T\}$, whether inequality (4.24) holds or not depends on the parameters such as the number of assets (n), the size of sample set (T), and absolute risk-preference index (c_0). The numerical experiment is implemented by generating a series of random return data R and its distribution f, then test whether (4.24) holds with the data. Hence, we can calculate the probability of inequality (4.24) to hold. By doing some numerical experiments, we discover that the probability of inequality (4.24) to hold is positively correlated to n and T. Also, a large absolute risk-preference index c_0 also leads to a higher probability that (4.24) holds. If the value of n and T are large enough, inequality (4.24) holds with a probability almost equals 1. Figure 4.1 can give an impression to readers about the correlations

between the probability of inequality (4.24) to hold and its parameters. In Figure 4.1, each curves represent a result with different number of asset (n). The sample number of asset (n) is selected based on a Fibonacci sequence.



Note: This figure shows the correlation between probability and the size of sample set. When $c_0 = 1$, no matter what the values of n and T are, the inequality (4.24) is less likely to hold. Excluding the case when $c_0 = 1$, the probability is approaching 1 as the size of sample set increases. It can be observed that the parameter n does not have a linear correlation with the probability. The correlation between them seems more likely quadratic. In the case when $c_0 > 0$, as the value of n increases, the probability decreases first and then increases after $n \geq 154$. Curves converge to 1 as T increases.

Figure 4.1: The graph of the probability of inequality (4.24) to hold relates to its parameters

Proposition 4.1.2 item (iv) is also an appropriate case for problem (4.10). Suppose $-r + \mu + \frac{w}{\|w\|} \sum_{j=1}^{T} \theta_j \neq 0$, by item (iv) of Proposition 4.1.2, there exists a parameter $\lambda \in \mathbb{R}_+$ such that

$$-r + \mu e + \frac{w}{\|w\|} \sum_{j=1}^{T} \theta_j = -\lambda w, \tag{4.25}$$

$$e^\top \left(c_0 U_{\|\cdot\|} \circ f - \theta\right) = \left\| -r + \mu e + \frac{w}{\|w\|} \sum_{j=1}^{T} \theta_j \right\|, \tag{4.26}$$

and

$$\mathbb{R}_+^T \ni \left(y \circ U_{\|\cdot\|}^{-1} - \|w\| e\right) \perp \left(c_0 U_{\|\cdot\|} \circ f - \theta\right) \in \mathbb{R}_+^T. \tag{4.27}$$

From (4.25) and (4.26) , we get

$$w = \frac{\|w\| \left(r - \mu e\right)}{\sum\limits_{j=1}^{T} \theta_j + \lambda \|w\|}, \tag{4.28}$$

$$\lambda \|w\| = e^\top \left(c_0 U_{\|\cdot\|} \circ f - \theta\right). \tag{4.29}$$

Combine (4.28), (4.29) and (4.14), we have

$$\mu = \bar{r} - \frac{c_0 \sum_{j=1}^{T} \|R_j - r\| f_j}{n \|w\|}$$

Substitute μ in (4.28), we obtain

$$w = \frac{(r - \bar{r} e)}{c_0 \sum_{j=1}^{T} \|R_j - r\| f_j} \|w\| + \frac{e}{n}. \tag{4.30}$$

The modulus $|U_j^\top w|$ in (4.9) is an obstacle against finding analytical solution to problem (4.8), but the norm $\|w\|$ in (4.30) is removable therefore leading to the analytical solution to problem (4.10). Since

$$\|w\|^2 = \langle w, w \rangle, \tag{4.31}$$

noting that $\langle r - \bar{r} e, e \rangle = 0$, we substitute w in the right-hand side of the equation (4.31) by (4.30), then we have:

$$
\begin{aligned}
\|w\|^2 &= \left\langle \frac{(r-\bar{r}e)}{c_0 \sum_{j=1}^{T} \|R_j - r\| f_j} \|w\| + \frac{e}{n}, \frac{(r-\bar{r}e)}{c_0 \sum_{j=1}^{T} \|R_j - r\| f_j} \|w\| + \frac{e}{n} \right\rangle \\
&= \frac{\|w\|^2}{\left(c_0 \sum_{j=1}^{T} \|R_j - r\| f_j\right)^2} \left\langle (r-\bar{r}e) - \frac{e}{n}, (r-\bar{r}e) - \frac{e}{n} \right\rangle \\
&= \frac{\|w\|^2}{\left(c_0 \sum_{j=1}^{T} \|R_j - r\| f_j\right)^2} \|r-\bar{r}e\|^2 + 1,
\end{aligned}
$$

then make some transformations, we have the following equation

$$
\left(1 - \frac{\|r-\bar{r}e\|^2}{\left(c_0 \sum_{j=1}^{T} \|R_j - r\| f_j\right)^2}\right) \|w\|^2 = \frac{1}{n}. \tag{4.32}
$$

Since $\|w\| > 0$, $\|w\|$ in (4.32) is a real number only if:

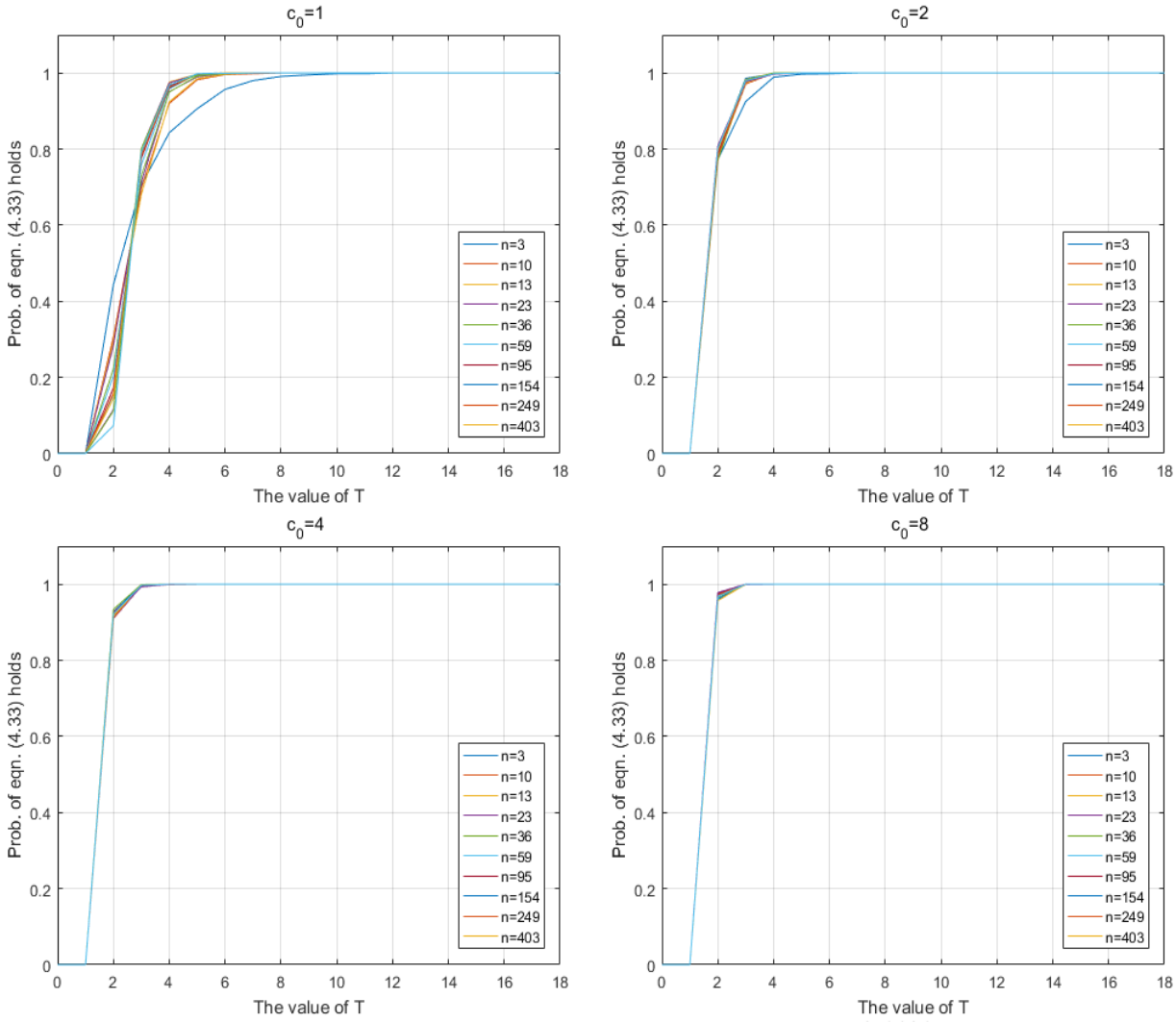
$$
1 - \frac{\|r-\bar{r}e\|^2}{\left(c_0 \sum_{j=1}^{T} \|R_j - r\| f_j\right)^2} > 0. \tag{4.33}
$$

Figure 4.2 shows the correlations between inequality (4.33) and its parameters. Similar to inequality (4.24), inequality (4.33) is not always hold. However, inequality (4.33) is somehow more plausible than inequality (4.24). When the size of sample set (T) is large enough (greater than 12 in the $c_0 = 1$ case), the inequality (4.33) has very high probability to hold. Unlike inequality (4.24), the probability is less related to the risk-preference index c_0. It means that it is suitable for more scenarios.

Suppose that inequality (4.33) holds, we can solve for w by (4.32) and (4.30):

$$
w = \frac{(r-\bar{r}e)}{\sqrt{n\left(\left(c_0 \sum_{j=1}^{T} \|R_j - r\| f_j\right)^2 - \|r-\bar{r}e\|^2\right)}} + \frac{e}{n}. \tag{4.34}
$$

The result (4.34) indicates that the weight of asset i relates to its excess return $(r_i - \bar{r}e)$, total risk of the market $U_{\|\cdot\|}^{\top} f$, and the absolute risk-preference index c_0. MEN conservatively consider the total market risk instead of the risk of single asset in the process of determining



Note: Compare to Proposition 1.4.3 item (iii), item (iv) is more applicable. Unlike the case in Figure 4.1, when $c_0 = 1$, the probability of inequality (4.33) to hold will be very close to 1 if $T \geq 12$. For $c_0 \geq 1$, curves converge to 1 as T increases. It also shows a quadratic correlations between probability and parameter n.

Figure 4.2: The graph of the probability of inequality (4.33) to hold relates to its parameters

the weight of each asset.

From (4.27), we get

$$\left(y_j \|U_j\|^{-1} - \|w\|\right)\left(c_0 \|U_j\| f_j - \theta_j\right) = 0, \quad j \in \{1, \ldots, T\}, \tag{4.35}$$

and

$$y_j = \|U_j\| \|w\|, \quad j \in \{1, \ldots, T\},$$

always hold if the objective function is minimised.

4.2 Conclusions and comments

In this chapter, we introduced a modified version of portfolio selection model based on the mean-variance model (MV) and the mean-absolute deviation model (MAD): the mean-Euclidean norm (MEN) model. The MV model has its analytical solution. However, this analytical solution requires the estimation of the covariance matrix, and the estimation of the covariance matrix of the MV model is computationally expensive. In addition, the result of such estimation is always negatively influenced by the estimation noises. Based on the MV model, Konno and Yamazaki [58] introduced the MAD model. The advantage of MAD model is that does not require to calculate covariance matrix of asset returns, therefore its computational costs is less than that of the MV model. However, given the modulus term in the constraint of MAD model, it is almost impossible to work out an analytical solution to the MAD model. Hence, when the number of assets is large, it also becomes computationally expensive in the process of finding the optimal weight of assets. Hence, the results from both the MV model and MAD model are criticised by many scholars [60, 61, 77, 86].

The MEN model considers a similar way of finding the optimal portfolio to both the mean-variance (MV) model and the MAD model: maximising the risk adjusted return. The objective of introducing this model is to find a new method based on the historical return data of assets. This model is designed for calculating the optimal weight of a portfolio with low computational cost. It is also designed to help in reducing the negative impacts of the inaccuracy of optimal solutions caused by estimation noises. The idea of introducing the MEN model is based on Corollary 4.1.1, which conservatively tighten the feasible set of the MAD model from (4.11) to (4.12). Given the fact that the KKT condition of the MEN model is a nonlinear complementarity problem on extended second order cone, we innovatively obtained the analytical solution to the MEN model based on Proposition 4.1.2 in Chapter 2.

The major advantage of the MEN model is has analytical solutions. However, due to the insufficient time and effort, this study fails to provide an empirical evaluations for comparing these three portfolio optimisation approaches. The empirical evaluations usually contain constraints such as short-sales constraints, lower-bound/upper-bound-on-individual-asset constraints, sector constraints, etc. These constraints are commonly occurred in practices. Without considering these constraints, the empirical evaluation results always looks absurd. Hence, our

future work is to explore the MEN models with linear constraints and evaluate this new model empirically.

Our current study can be extended and improved from practical aspects. The empirical evaluations for the MEN model is the suggested direction. Also, the complementarity problem on extended second order cone has potential on the application of solving the asset allocation problem for the defined-contribution pension plan [10].

CHAPTER 5

SPHERICALLY QUASI-CONVEX QUADRATIC FUNCTIONS

5.1 Motivation of this study

In this chapter, we study the spherical quasi-convexity of quadratic functions on spherically convex sets, which is related to the problem of finding their minimiser. The spherically convex set is a natural extension of the concepts and techniques of convexity of mathematical programming problem. The original intention of this study is to explicitise certain fixed point theorems, surjectivity theorems, and existence theorems for complementarity problems and variational inequalities. Parts of the results of this chapter are published in [28], but we provide more detailed definitions and proofs in this thesis.

Recall the definition of the nonlinear complementarity problem:

Definition 5.1.1 (Nonlinear complementarity problem)

Let $F : \mathbb{R}^n \to \mathbb{R}^n$ be a mapping. Let $\mathcal{K} \subseteq \mathbb{R}^n$ be a nonempty closed convex cone and $\mathcal{K}^*$ its dual. Defined by $\mathcal{K}$ and F the nonlinear complementarity problem (NCP) *[19] is:*

$$NCP(F,\mathcal{K}) \begin{cases} Find\ x \in \mathbb{R}^n,\ such\ that \\ (x, F(x)) \in \mathcal{C}(\mathcal{K}). \end{cases} \tag{5.1}$$

The solution set of NCP$(F,\mathcal{K})$ *is denoted by* SOL-NCP$(F,\mathcal{K})$*:*

$$\text{SOL-NCP}(F,\mathcal{K}) = \{x \in \mathbb{R}^n : (x, F(x)) \in \mathcal{C}(\mathcal{K})\}.$$

To explain the motivation of this study in detail, we start with presenting the following three definitions:

Definition 5.1.2 (Inversion)

[81, Definition 5] The operator

$$i : \mathbb{R}^n \setminus \{0\} \to \mathbb{R}^n \setminus \{0\};\ i(x) := \frac{x}{\|x\|^2}$$

is called inversion (of pole 0).

It is easy to see that i is a one-to-one mapping, and $i^{-1} = i$.

Definition 5.1.3 (Inversion of a mapping)

[81, Definition 6] The inversion (of pole 0) of the mapping $F : \mathbb{R}^n \to \mathbb{R}^n$, is the mapping $\mathcal{I}(F) : \mathcal{K} \to \mathbb{R}^n$ defined by:

$$\mathcal{I}(F)(x) := \begin{cases} \|x\|^2 (F \circ i)(x) \ \ if\ x \neq 0, \\ 0 \ \ if\ x = 0. \end{cases}$$

Definition 5.1.4 (Lower scalar derivative)

[46, Definition 1.6] Consider the mapping $F : \mathcal{K} \to \mathbb{R}^n$. The limit

$$\underline{F}^{\#}(x_0, \mathcal{K}) := \liminf_{x \to x_0, x - x_0 \in \mathcal{K}} \frac{\langle F(x) - F(x_0), x - x_0 \rangle}{\|x - x_0\|^2}$$

is called the lower scalar derivative of F at x_0.

The lower scalar derivative is highly related to the minimising questions on spherically convex sets. We can observe the clues from [81, Theorem 18]. This theorem gives a more explicit expression of the lower scalar derivative if the mapping F is Fréchet differentiable (see definition 1.2.4):

Theorem 5.1.1

[81, Theorem 18] Let $\mathcal{K} \subseteq \mathbb{R}^n$ be a closed convex cone with non-empty interior and x an interior point of $\mathcal{K}$. If $F : \to \mathbb{R}^n$ is Fréchet differentiable in x, with the differential $JF(x)$, then

$$\underline{F}^{\#}(x, \mathcal{K}) = \min_{\|u\|=1, u \in \mathcal{K}} \langle JF(x)(u), u \rangle.$$

Theorem 5.1.1 leads to the study about minimising questions on the intersection between a cone and the sphere. The Corollary [45, Corollary 8.1] is the trigger of this study. We state this corollary here for convenience.

Corollary 5.1.1

[45, Corollary 8.1] Let $\mathcal{K} \subseteq \mathbb{R}^n$ be a closed convex cone, and $F : \mathcal{K} \to \mathbb{R}^n$ be a continuous mapping such that its inversion (of pole 0) is differentiable at 0. Denote by $J\mathcal{I}(F)(0)$ the Jacobian matrix of the inversion of mapping F at 0. Then if

$$\underline{\mathcal{I}(F)}^{\#}(0) := \min_{\|x\|=1, u\in\mathcal{K}} \langle J\mathcal{I}(F)(0)u, u\rangle > 0.$$

then the nonlinear complementarity problem NCP *has a solution.*

By [81, Theorem 18] and [45, Corollary 8.1], the question of the existence of the solution to a nonlinear complementarity problem can be converted to a problem of minimising a quadratic function on the intersection between a cone and the sphere. These sets are exactly the spherically convex sets (see [25]), which leads to the problem of minimising quadratic functions on spherically convex sets.

Apart from the above, the motivation of this study is much wider. For instance, consider the quadratic constrained optimisation problem on the sphere

$$\min\{\langle Qx, x\rangle \ : \ x \in \mathcal{S} \cap \mathcal{K}\}, \tag{5.2}$$

where $\mathcal{K}$ is a cone on the sphere, and $\mathcal{S} \subseteq \mathbb{S}^{n-1} := \{x \in \mathbb{R}^n \ : \ \|x\| = 1\}$ is a sphere. Suppose the matrix Q is a symmetric matrix, then the problem (5.2) is a minimum eigenvalue problem in $\mathcal{S}$. In particular, this problem includes the problem of finding the spectral norm of the matrix $-Q$ when $\mathcal{S} = \mathbb{S}^{n-1}$ (see, e.g., [95]). We focus on the cases when $\mathcal{S}$ is an intersection of a subdual convex set with the sphere. Also, it is important to highlight that the special case when $\mathcal{S}$ is the intersection of the nonnegative orthant with the shpere is of particular interest because the nonnegativity of the minimum value is equivalent to the copositivity of the matrix Q [42, Proposition 1.3] and to the nonnegativity of all Pareto eigenvalues of Q [42, Theorem 4.3]. As far as we are aware there are no methods for finding the Pareto spectra by using the intrinsic geometrical properties of the sphere, hence our study is expected to open new perspectives for

detecting the copositivity of a symmetric matrix. Another important special case is when $\mathcal{S}$ is the intersection of the Lorentz cone with the sphere. We pay attention to this case because the minimum eigenvalue of Q in $\mathcal{S}$ is nonnegative if and only if the matrix Q is Lorentz copositive, see [35, 65]. In general, changing the Lorentz cone by an arbitrary closed convex cone $\mathcal{K}$ would lead to a more general concept of $\mathcal{K}$ copositivity, thus our study is anticipated to initialise new perspectives of investigating the general copositivity of a symmetric matrix. More problems that deals with "spherical" constraint can be found in [69].

Optimisation problems posed on the sphere have a specific underlying algebraic structure that could be exploited to greatly reduce the cost of obtaining the solutions; see [38, 39, 95, 96, 108, 110]. It is worth to point out that when a quadratic function is spherically quasi-convex, then a spherical strict local minimiser is equal to a spherical strict global minimiser. Therefore, it is natural to consider the problem of determining the spherically quasi-convex quadratic functions on spherically convex sets. The goal of the study is to present necessary conditions and sufficient conditions for quadratic functions which are spherically quasi-convex on spherical convex sets. As a particular case, we exhibit several such results for both the spherical positive orthant and, more general, the spherical subdual convex set.

Apart from the questions about the existence of a solution to the nonlinear complementarity problem, this study also related to many other questions. For example, minimising a quadratic function defined on spherical nonnegative orthant is equivalent to finding the minimum Pareto eigenvalues of the quadratic function. Hence, if the minimum value of this quadratic function is nonnegative, then the minimum Pareto eigenvalues of the quadratic function will also be nonnegative. In this chapter, we present several conditions that characterise the spherical quasi-convexity of quadratic functions. The study can be considered as a first spherical analogue for the study of quasi-convexity of quadratic functions. Without the aim of completeness, we list here some of the main papers about the quasi-convexity of quadratic functions: [23,52,54,74,94].

5.2 Spherically quasi-convex quadratic functions on spherically convex sets

In this section our aim is to present some conditions characterising quadratic spherically quasi-convex functions on a general spherically convex set. We assume for convenience that

from now on the cone $\mathcal{K} \subseteq \mathbb{R}^n$ is a proper subdual cone. Define

$$\mathcal{S} = \mathbb{S}^{n-1} \cap \operatorname{int}(\mathcal{K}), \quad \bar{\mathcal{S}} = \mathbb{S}^{n-1} \cap \mathcal{K}, \tag{5.3}$$

and assume that $\mathcal{S}$ is an open spherically convex set.

Definition 5.2.1 (Quadratic function)

The associated quadratic function $q_A : \mathcal{S} \to \mathbb{R}$ defined by the symmetric matrix $A = A^T \in \mathbb{R}^{n\times n}$ is

$$q_A(x) := \langle Ax, x\rangle. \tag{5.4}$$

We remark that q_A can be extended to $\bar{\mathcal{S}}$. For the simplicity of notations we will denote the extended values by $q_A(x)$ too, but the spherical quasi-convexity of q_A will always be understood as a function defined on $\mathcal{S}$. To proceed we need the following definition:

Definition 5.2.2 (Rayleigh quotient function)

The Rayleigh quotient function $\varphi_A : \operatorname{int}(\mathcal{K}) \to \mathbb{R}$ restricted on $\operatorname{int}(\mathcal{K})$ defined by matrix A is

$$\varphi_A(x) := \frac{\langle Ax, x\rangle}{\|x\|^2}. \tag{5.5}$$

In the following proposition we present some equivalent characterisations of the convexity of the associated quadratic function on spherically convex sets q_A defined in (5.4).

Proposition 5.2.1

Let q_A and φ_A be the functions defined in (5.4) *and* (5.5)*, respectively. The following statements are equivalent:*

(a) The quadratic function q_A is spherically quasi-convex;

(b) $\langle Ax, y\rangle \leq \langle x, y\rangle \max\{q_A(x),\ q_A(y)\}$ for any $x, y \in \mathbb{S}^{n-1} \cap \mathcal{K}$;

(c) $\dfrac{\langle Ax, y\rangle}{\langle x, y\rangle} \leq \max\{\varphi_A(x),\ \varphi_A(y)\}$, for any $x, y \in \mathcal{K}$ with $\langle x, y\rangle \neq 0$.

Proof:

(a)⇒(b): First of all, we assume that item (a) holds. Arbitrarily take $x, y \in \mathcal{S}$. *Thus, either* $q_A(x) \leq q_A(y)$ *or* $q_A(y) \leq q_A(x)$ *holds. By using Proposition 1.5.6 we conclude that*

$$\begin{aligned} & q_A(x) \leq q_A(y) \\ \Rightarrow\ & \langle Dq_A(y), x\rangle - \langle x, y\rangle \langle Dq_A(y), y\rangle \leq 0 \\ \Rightarrow\ & \langle Ay, x\rangle \leq \langle x, y\rangle \langle Ay, y\rangle \\ \Rightarrow\ & \langle Ay, x\rangle \leq \langle x, y\rangle q_A(y) \end{aligned}$$

Similarly,

$$q_A(y) \leq q_A(x) \Rightarrow \langle Ax, y\rangle \leq \langle x, y\rangle q_A(x)$$

Thus, the symmetric matrix A *implies* $\langle Ax, y\rangle = \langle Ay, x\rangle$, *taking into account that* $\mathcal{S} = \mathbb{S}^{n-1} \cap \operatorname{int}(\mathcal{K})$ *and* $\mathcal{K}$ *is a subdual cone and hence* $\langle x, y\rangle > 0$, *we have*

$$\langle Ax, y\rangle \leq \max\{\langle x, y\rangle q_A(x), \langle x, y\rangle q_A(y)\} = \langle x, y\rangle \max\{q_A(x), q_A(y)\},\ \forall\ x, y \in \mathcal{S}.$$

Therefore, by continuity the above inequality can be extended to all $x, y \in \mathbb{S}^{n-1} \cap \mathcal{K}$ *and, then item (b) holds.*

(b)⇒(a): Conversely, we assume that item (b) holds. Take $x, y \in \mathcal{S}$ *satisfying* $q_A(x) \leq q_A(y)$. *Then, by the inequality in item (b) and the fact that* $\mathcal{K}$ *is a subdual cone, we have*

$$q_A(x) \leq q_A(y) \ \Rightarrow\ \langle Ax, y\rangle - \langle x, y\rangle q_A(y) \leq 0.$$

Hence, by using Proposition 1.5.6 we conclude that q_A *is a spherically quasi-convex function.*

(b) ⇒ (c): To establish the equivalence between (b) and (c), we firstly assume that item (b) holds. Let $x, y \in \mathcal{K}$ *with* $\langle x, y\rangle \neq 0$. *Then,* $x \neq 0$ *and* $y \neq 0$. *Moreover, we have by the property of a cone:*

$$u := \frac{x}{\|x\|} \in \mathbb{S}^{n-1} \cap \mathcal{K}, \qquad v := \frac{y}{\|y\|} \in \mathbb{S}^{n-1} \cap \mathcal{K}.$$

Hence, by using the inequality in item (b) with $x = u$ and $y = v$, we obtain the inequality in item (c).

(c) $\Rightarrow$ (b): Conversely, suppose that (c) holds. Take $x, y \in \mathbb{S}^{n-1} \cap \mathcal{K}$ with $\langle x, y\rangle \neq 0$. We have $\|x\| = \|y\| = 1$ as $x, y \in \mathbb{S}^{n-1}$. From the inequality in item (c) we conclude that

$$\frac{\langle Ax, y\rangle}{\langle x, y\rangle} \leq \max\left\{q_A(x),\ q_A(y)\right\}.$$

Due to $\mathcal{K}$ being a subdual cone, we have $\langle x, y\rangle \geq 0$, and hence the last inequality is equivalent to the inequality in item (b).

Now, assume that $\langle x, y\rangle = 0$. Then, take two sequences $\{x_k\}, \{y_k\} \subseteq \mathcal{S}$ such that $\lim_{k\to+\infty} x_k = x$, $\lim_{k\to+\infty} y_k = y$ and $\langle x_k, y_k\rangle \neq 0$. Since $\mathcal{K}$ is a subdual cone, we have $\langle x_k, y_k\rangle > 0$ for any $k \in \{1, 2, \ldots\}$. Therefore, considering that $\|x_k\| = \|y_k\| = 1$ for any $k = 1, 2, \ldots$, we can apply again the inequality in item (c) to conclude

$$\langle Ax_k, y_k\rangle \leq \langle x_k, y_k\rangle \max\left\{q_A(x_k),\ q_A(y_k)\right\}, \qquad k = 1, 2, \ldots.$$

By tending with k to infinity, we conclude that the inequality in item (b) also holds for $\langle x, y\rangle = 0$ and the proof of the equivalence between (b) and (c) is complete. □

Corollary 5.2.1

Assume that $\mathcal{K}$ is a self-dual cone. If the quadratic function q_A is spherically quasi-convex, then A has the $\mathcal{K}$-Z-property (see Definition 1.5.4).

Proof:

To prove A has the $\mathcal{K}$-Z-property, we need to prove that

$$\langle Ax, y\rangle \leq 0$$

for any $(x, y) \in \mathcal{C}(\mathcal{K})$. Take $x, y \in \mathbb{R}^n$ such that $(x, y) \in \mathcal{C}(\mathcal{K})$. If either $x = 0$ or $y = 0$, we have $\langle Ax, y\rangle = 0$. Thus, assume that $x \neq 0$ and $y \neq 0$. Considering that $\mathcal{K}$ is a self-dual cone, we have $\frac{x}{\|x\|},\ \frac{y}{\|y\|} \in \mathbb{S}^{n-1} \cap \mathcal{K}$. Suppose that q_A is spherically quasi-convex and by the items (a) and (b) of Proposition 5.2.1, we have

$$\left\langle A\frac{x}{\|x\|}, \frac{y}{\|y\|}\right\rangle \leq \left\langle \frac{x}{\|x\|}, \frac{y}{\|y\|}\right\rangle \max\left\{q_A\left(\frac{x}{\|x\|}\right), q_A\left(\frac{y}{\|y\|}\right)\right\}, \quad \forall x, y \in \mathcal{C}(\mathcal{K}).$$

By fact that $\left\langle \frac{x}{\|x\|}, \ \frac{y}{\|y\|} \right\rangle = 0$, *we obtain* $\langle Ax, y\rangle \leq 0$. □

Theorem 5.2.1

The function q_A *defined in* (5.4) *is spherically quasi-convex if and only if* φ_A *defined in* (5.5) *is quasi-convex.*

Proof:

For any $c \in \mathbb{R}$, *let* $[q_A \leq c] := \{x \in \mathcal{S} \ : \ q_A(x) \leq c\}$ *and* $[\varphi_A \leq c] := \{x \in \operatorname{int}(\mathcal{K}) \ : \ \varphi_A(x) \leq c\}$ *be the sublevel sets of* q_A *and* φ_A, *respectively. Let* $\mathcal{K}_{[q_A \leq c]}$ *be the cone spanned by* $[q_A \leq c]$. *Since* $\mathcal{S} = \mathbb{S}^{n-1} \cap \operatorname{int}(\mathcal{K})$, *we conclude that* $x \in \operatorname{int}\mathcal{K}$ *if and only if* $x/\|x\| \in \mathcal{S}$. *Hence, by the definitions of* $[q_A \leq c]$ *and* $[\varphi_A \leq c]$ *we obtain:*

$$\begin{aligned}
\mathcal{K}_{[q_A \leq c]} &= \left\{ tx \ : \ x \in \mathcal{S}, q_A(x) \leq c, t \in [0, +\infty) \right\} \\
&= \left\{ x \in \operatorname{int}(\mathcal{K}) \ : \ q_A\left(\frac{x}{\|x\|}\right) \leq c \right\} \\
&= \{x \in \operatorname{int}(\mathcal{K}) \ : \ \varphi_A(x) \leq c\}.
\end{aligned}$$

That is:

$$\mathcal{K}_{[q_A \leq c]} = [\varphi_A \leq c]. \tag{5.6}$$

Suppose that the quadratic function q_A *is spherically quasi-convex. Thus, from Poposition 1.5.3 we conclude that* $[q_A \leq c]$ *is spherically convex for any* $c \in \mathbb{R}$. *Hence, it follows from Proposition 1.5.1 that the cone* $\mathcal{K}_{[q_A \leq c]}$ *is convex and pointed, which implies from* (5.6) *that* $[\varphi_A \leq c]$ *is convex for any* $c \in \mathbb{R}$. *Therefore, again by Poposition 1.5.3 we conclude that* φ_A *is quasi-convex.*

Conversely, suppose that φ_A *is quasi-convex. Thus,* $[\varphi_A \leq c]$ *is convex for any* $c \in \mathbb{R}$. *On the other hand, given* $\mathcal{K}$ *is a proper subdual cone,* $\operatorname{int}\mathcal{K}$ *is pointed. Thus, we conclude that* $[\varphi_A \leq c] \subseteq \operatorname{int}\mathcal{K}$ *is also a pointed cone. Hence, from* (5.6) *it follows that* $\mathcal{K}_{[q_A \leq c]}$ *is, again, a pointed convex cone. Hence, Proposition 1.5.1 implies that* $[q_A \leq c]$ *is spherically convex for any* $c \in \mathbb{R}$. *Therefore, by using Proposition 1.5.3, we conclude that* q_A *is a spherically quasi-convex function.* □

Let $c \in \mathbb{R}$, recall the definition (1.15) with $f \equiv \varphi_A$:

$$\begin{aligned}[\varphi_A \leq c] :=& \{x \in \mathcal{S} : \ \varphi_A(x) \leq c\} \\ =& \{x \in \mathcal{S} : \ \langle A_c x, x\rangle \leq 0\}, \qquad A_c := A - cI_n.\end{aligned}$$

Corollary 5.2.2

The function q_A is spherically quasi-convex if and only if for any $c \in \mathbb{R}$ the set $[\varphi_A \leq c]$ is convex.

Proof:

Suppose that the quadratic function q_A is spherically quasi-convex. Hence Theorem 5.2.1 implies that φ_A is quasi-convex, and the sub-level set $[\varphi_A \leq c]$ is convex for any $c \in \mathbb{R}$ by Proposition 1.5.3. Since $\{x \in \text{int}(\mathcal{K}) : \ \langle A_c x, x\rangle < 0\} \neq \emptyset$, we conclude that

$$cl\Big(\big\{x \in \text{int}(\mathcal{K}) \ : \ \langle A_c x, x\rangle \leq 0\big\}\Big) = \{x \in \mathcal{K} \ : \ \langle A_c x, x\rangle \leq 0\}.$$

where "$cl(\cdot)$" is the topological closure operator of a set. Thus, considering that

$$\begin{aligned}[\varphi_A \leq c] &= \Big\{x \in \text{int}(\mathcal{K}) \ : \ \frac{\langle Ax, x\rangle}{\|x\|^2} \leq c\Big\} \\ &= \{x \in \text{int}(\mathcal{K}) \ : \ \langle Ax, x\rangle - c\|x\|^2 \leq 0\} \\ &= \{x \in \text{int}(\mathcal{K}) \ : \ \langle Ax, x\rangle - c\langle x, x\rangle \leq 0\} \\ &= \{x \in \text{int}(\mathcal{K}) \ : \ \langle Ax - cx, x\rangle \leq 0\} \\ &= \Big\{x \in \text{int}(\mathcal{K}) \ : \ \big\langle (A - cI_n)x, x\big\rangle \leq 0\Big\},\end{aligned}$$

we obtain that

$$cl\,([\varphi_A \leq c]) = \{x \in \mathcal{K} \ : \ \langle A_c x, x\rangle \leq 0\},$$

Taking into account that $[\varphi_A \leq c]$ is convex, the set $cl\big([\varphi_A \leq c]\big)$ is also convex. □

5.3 Spherically quasi-convex quadratic functions on the spherical positive orthant

In this section we present some properties of a quadratic function defined in the spherical positive orthant, which corresponds to $\mathcal{K} = \mathbb{R}^n_+$ (therefore $\operatorname{int} K = \mathbb{R}^n_{++}$). If A has only one eigenvalue, it is easy to conclude that q_A is spherically quasi-convex. However, suppose that the only eigenvalue of A is λ, then $q_A(x) = \lambda$ for any $x \in \mathcal{S}$ (we say q_A is constant in this case), which is meaningless to discuss. Therefore, *throughout this section we assume that A has at least two distinct eigenvalues.* By the definitions in (5.4) and (5.5), the domains of q_A and φ_A ($\mathcal{S}$ and $\operatorname{int}(\mathcal{K})$, respectively) are given by

$$\mathcal{S} := \mathbb{S}^{n-1} \cap \mathbb{R}^n_{++}, \qquad \operatorname{int}(\mathcal{K}) := \mathbb{R}^n_{++}, \tag{5.7}$$

Next we present a technical lemma which will be useful in the sequel.

Lemma 5.3.1

Let $n \geq 2$ and $V = [v^1\ v^2\ v^3\ \cdots\ v^n] \in \mathbb{R}^{n\times n}$ be an orthogonal matrix, $A = V\Lambda V^\top$ and $\Lambda = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$. Assume that $\lambda_1 < \lambda_2 \leq \ldots \leq \lambda_n$. If $v^1 \in \mathbb{R}^n_+$, then the sublevel set $[\varphi_A \leq c]$ is convex for any $c \notin (\lambda_2, \lambda_n)$.

Proof:

By using that $V^\top V = I_n$ and $A = V\Lambda V^\top$ we obtain from the definition (5.5) that

$$\begin{aligned}
[\varphi_A \leq c] &= \left\{ x \in \mathbb{R}^n_{++} :\ \frac{\langle Ax, x\rangle}{\|x\|^2} \leq c \right\} \\
&= \left\{ x \in \mathbb{R}^n_{++} :\ \langle Ax, x\rangle - c\|x\|^2 \leq 0 \right\} \\
&= \left\{ x \in \mathbb{R}^n_{++} :\ \langle (A - cI_n)x, x\rangle \leq 0 \right\} \\
&= \left\{ x \in \mathbb{R}^n_{++} :\ x^\top V\left(\Lambda - cI_n\right)V^\top x \leq 0 \right\} \\
&= \left\{ x \in \mathbb{R}^n_{++} :\ \sum_{i=1}^n (\lambda_i - c)\left(x^\top v^i\right)^2 \leq 0 \right\} \\
&= \left\{ x \in \mathbb{R}^n_{++} :\ \sum_{i=1}^n (\lambda_i - c)\langle v^i, x\rangle^2 \leq 0 \right\}
\end{aligned} \tag{5.8}$$

In the following we will show that $[\varphi_A \leq c]$ is convex for any $c \notin (\lambda_2, \lambda_n)$.

If $c < \lambda_1$, then since $v^1, v^2, \ldots, v^n$ are linearly independent, we conclude from (5.8) *that $[\varphi_A \leq c] = \{0\}$ and therefore it is convex.*

If $c = \lambda_1$, then from (5.8) *we conclude that $[\varphi_A \leq c] = \mathcal{O} \cap \mathbb{R}^n_{++}$, where $\mathcal{O} := \{x \in \mathbb{R}^n : \langle v^2, x\rangle = 0, \ \ldots, \ \langle v^n, x\rangle = 0\}$ is a convex cone. Hence $[\varphi_A \leq c]$ is convex.*

If $\lambda_1 < c \leq \lambda_2$, letting $y = V^\top x$, i.e., $y_i = \langle v^i, x\rangle$, for $i = 1, \ldots, n$. Since $v^1 \in \mathbb{R}^n_{++}$ and $x \in \mathbb{R}^n_{++}$, we have $y_1 = \langle v^1, x\rangle > 0$. From (5.8) *we obtain*

$$[\varphi_A \leq c] = \left\{ x \in \mathbb{R}^n_{++} : \ (c - \lambda_1)\langle v^1, x\rangle^2 \geq \sum_{i=2}^{n} (\lambda_i - c)\langle v^i, x\rangle^2 \right\}$$
$$= \left\{ x \in \mathbb{R}^n_{++} : \ \langle v^1, x\rangle^2 \geq \sum_{i=2}^{n} \theta_i \langle v^i, x\rangle^2 \right\}$$

where $\theta_i = \frac{\lambda_i - c}{c - \lambda_1}$ for $i = 2, \ldots, n$. Denote the cone

$$\mathcal{L} := \left\{ y = (y_1, \ldots, y_n) \in \mathbb{R}^n : \ y_1 \geq \sqrt{\theta_2 y_2^2 + \ldots + \theta_n y_n^2} \right\},$$

we have $[\varphi_A \leq c] = \mathcal{L} \cap V^\top \mathbb{R}^n_{++}$. Since both $\mathcal{L}$ and $V^\top \mathbb{R}^n_{++}$ are convex sets, we conclude that $[\varphi_A \leq c]$ is convex.

If $c \geq \lambda_n$, then $[\varphi_A \leq c] = \mathbb{R}^n_{++}$ is convex. □

It should be mentioned that if $\lambda_1 < \lambda_2 = \cdots = \lambda_n$, then the sublevel set $[\varphi_A \leq c]$ is convect for any $c \in \mathbb{R}$. This will be a useful fact in a proof of a following theorem.

Lemma 5.3.2

Let λ be an eigenvalue of A. If $\lambda I_n - A$ is copositive and $\lambda \leq c$, then

$$[\varphi_A \leq c] = \mathbb{R}^n_{++}$$

and consequently it is a convex set.

Proof:

Let $c \in \mathbb{R}$ and $[\varphi_A \leq c] = \{x \in \mathbb{R}^n_{++} : \ \langle Ax, x\rangle - c\|x\|^2 \leq 0\}$. Suppose that $\lambda \leq c$, for any $x \in \mathbb{R}^n_{++}$ we have

$$\langle Ax, x\rangle - c\|x\|^2 \leq \langle Ax, x\rangle - \lambda\|x\|^2 = \langle (A - \lambda I_n)x, x\rangle,$$

and $\lambda I_n - A$ is copositive, that is

$$\langle Ax, x\rangle - c\|x\|^2 \leq \langle (A - \lambda I_n)x, x\rangle \leq 0,$$

hence $\langle Ax, x\rangle - c\|x\|^2 \leq 0$ holds for any $x \in \mathbb{R}^n_{++}$, which implies that $[\varphi_A \leq c] = \mathbb{R}^n_{++}$. $\square$

The next theorem exhibits a series of implications and, in particular, conditions which imply that the quadratic function q_A is spherically quasi-convex.

Theorem 5.3.1

Let $A \in \mathbb{R}^{n\times n}$ be a symmetric matrix and let $\lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_n$ its eigenvalues. Consider the following statements:

(i) q_A *is a spherically quasi-convex function.*

(ii) A *is a Z-matrix.*

(iii) A *is a Z-matrix and* $\lambda_2 \geq a_{ii}$ *for any* $i \in \{1, 2, \ldots, n\}$.

(iv) A *is a Z-matrix,* $\lambda_1 < \lambda_2$ *and* $\lambda_2 \geq a_{ii}$ *for any* $i \in \{1, 2, \ldots, n\}$.

(v) A *is an irreducible Z-matrix and* $\lambda_2 \geq a_{ii}$ *for any* $i \in \{1, 2, \ldots, n\}$.

(vi) $\lambda_2 I_n - A$ *is copositive and there exists an eigenvector* $v^1 \in \mathbb{R}^n_{++}$ *corresponding to the eigenvalue* λ_1 *of* A.

Then the following implications hold:

$$\begin{array}{ccccccc} & & & (iv) & & & \\ & & & \Downarrow & & & \\ (iii) & \Leftarrow & (vi) & \Rightarrow & (i) & \Rightarrow & (ii) \\ & & & \Uparrow & & & \\ & & & (v) & & & \end{array}$$

Proof:

(iv)⇒(vi): Suppose that A is a Z-matrix, $\lambda_1 < \lambda_2$ and $\lambda_2 \geq a_{ii}$ for any $i \in \{1, 2, \ldots, n\}$. It is easy to verify that $\lambda_2 I_n - A$ is a nonnegative matrix, and hence $\lambda_2 I_n - A$ is copositive. Moreover, applying item (i) of Perron-Frobenius theorem (see Theorem 1.5.1) to the matrix $\lambda_2 I_n - A$, we obtain that there exists an eigenvector $v^1 \in \mathbb{R}^n_{++}$ of $\lambda_2 I_n - A$ corresponding to the largest eigenvalue $\lambda_2 - \lambda_1$ (v^1 is also the eigenvector of A corresponding to λ_1).

(v)⇒(vi): Suppose that A is an irreducible Z-matrix and $\lambda_2 \geq a_{ii}$ for any $i \in \{1, 2, \ldots, n\}$. Again, by applying item (i) of Perron-Frobenius theorem we conclude that there exists an eigenvector $v^1 \in \mathbb{R}^n_{++}$ of $\lambda_2 I_n - A$ corresponding to the largest eigenvalue $\lambda_2 - \lambda_1$ (v^1 is also the eigenvector of A corresponding to λ_1).

(vi)⇒(i): If $c \leq \lambda_2$, then Lemma 5.3.1 implies that $[\varphi_A \leq c]$ is convex. If $c \geq \lambda_2$, then from Lemma 5.3.2 we have $[\varphi_A \leq c] = \mathbb{R}^n_{++}$, which is convex. Hence, $[\varphi_A \leq c]$ is convex for any $c \in \mathbb{R}$. Therefore, by using Theorem 5.2.1, we conclude that q_A is spherically quasi-convex function.

(i)⇒(ii): Suppose that q_A is spherically quasi-convex. From Corollary 5.2.1, it follows that A has the $\mathbb{R}^n_+$-Z-property. By applying Theorem 1.5.2, we obtain that A is a Z-matrix.

(vi)⇒(iii): Suppose (vi) holds, by above proof (vi) $\Longrightarrow$ (i) $\Longrightarrow$ (ii), it follows that A is a Z-matrix. Arbitrarily take $x = e^i$ with $i \in \{1, 2, \ldots, n\}$, e^i are canonical vectors of $\mathbb{R}^n$. Since $\lambda_2 I_n - A$ is copositive, it follows that

$$\langle (\lambda_2 I_n - A)x, x \rangle \geq 0 \Rightarrow \lambda_2 - a_{ii} \geq 0$$

for any $i \in \{1, 2, \ldots, n\}$.

□

Corollary 5.3.1

Let $n \geq 2$ and $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$ be the eigenvalues of A. Assume that $-A$ is an entrywise positive matrix, $\lambda_1 < \lambda_2 \leq \ldots \leq \lambda_n$ and $0 < \lambda_2$. Then q_A is spherically quasi-convex.

Proof:

Suppose that $-A$ is an entrywise positive matrix, $\lambda_1 < \lambda_2 \leq \ldots \leq \lambda_n$ and $0 < \lambda_2$, then the matrix $\lambda_2 I_n - A$ is an entrywise positive matrix with $\lambda_2 - \lambda_1 > 0$ to be its largest eigenvalue. Thus, Theorem 1.5.1 implies that the eigenvalue $\lambda_2 - \lambda_1$ has the associated

eigenvector $v^1 \in \mathbb{R}^n_{++}$. *Since by the definition of eigenvector*

$$(\lambda_2 I_n - A)v^1 = (\lambda_2 - \lambda_1)v^1 \Rightarrow Av^1 = \lambda_1 v^1,$$

we conclude that v^1 *is also an eigenvector of* A *associated to* λ_1. *Therefore, considering that* A *is a entrywise negative matrix, then* A *is also a Z-matrix. Since* $v^1 \in \mathbb{R}^n_{++}$, $\lambda_1 < \lambda_2$ *and* $\lambda_2 \geq a_{ii}$ *for any* $i \in \{1, 2 \ldots, n\}$, *it follows from Theorem 5.3.1 (iv)⇒(i) that* q_A *is spherically quasi-convex.* □

In the following two examples we use Theorem 5.3.1 (vi)⇒(i) to illustrate a class of quadratic quasi-convex functions defined in the spherical positive orthant.

Example 5.3.1

Let $n \geq 3$ *and* $V = [v^1 \ v^2 \ v^3 \ \cdots \ v^n] \in \mathbb{R}^{n\times n}$ *be an orthogonal matrix,* $A = V\Lambda V^\top$ *and* $\Lambda := \operatorname{diag}(\lambda, \mu, \ldots, \mu, \nu)$, *where* $\lambda, \mu, \nu \in \mathbb{R}$. *Then* q_A *is a spherically quasi-convex if*

$$v^1 - \sqrt{\frac{\nu - \mu}{\mu - \lambda}}|v^n| \in \mathbb{R}^n_+, \qquad \lambda < \mu < \nu, \tag{5.9}$$

where $|v^n| := (|v^n_1|, \ldots, |v^n_n|)$. *Indeed, by using that* $V^\top V = I_n$ *and* $A = V\Lambda V^\top$, *then*

$$\begin{aligned}
\mu\|x\|^2 - \langle Ax, x\rangle &= \mu\|x\|^2 - \langle V\Lambda V^\top x, x\rangle \\
&= \mu\|x\|^2 - \left(x^\top v^1\right)^2 \lambda + \sum_{i=2}^{n-1} \left(x^\top v^i\right)^2 \mu + \left(x^\top v^n\right)^2 \nu \\
&= \mu\langle I_n x, x\rangle - \lambda\langle v^1, x\rangle^2 + \sum_{i=2}^{n-1} \mu\langle v^i, x\rangle^2 + \nu\langle v^n, x\rangle^2 \\
&= \mu\langle V^\top V x, x\rangle - \lambda\langle v^1, x\rangle^2 + \sum_{i=2}^{n-1} \mu\langle v^i, x\rangle^2 + \nu\langle v^n, x\rangle^2 \\
&= \sum_{i=1}^{n} \mu\langle v^i, x\rangle^2 - \lambda\langle v^1, x\rangle^2 + \sum_{i=2}^{n-1} \mu\langle v^i, x\rangle^2 + \nu\langle v^n, x\rangle^2 \\
&= (\mu - \lambda)\left[\langle v^1, x\rangle^2 - \frac{\nu - \mu}{\mu - \lambda}\langle v^n, x\rangle^2\right].
\end{aligned} \tag{5.10}$$

Thus, using the condition in (5.9) *and arbitrarily take* $x \in \mathbb{R}^n_{++}$, *we have*

$$\langle v^1, x\rangle^2 - \frac{\nu - \mu}{\mu - \lambda}\langle v^n, x\rangle^2 \geq \frac{\nu - \mu}{\mu - \lambda}\left[\langle |v^n|, x\rangle^2 - \langle v^n, x\rangle^2\right] \geq 0.$$

Hence, by combining the last inequality with (5.10)*, we conclude that*

$$\langle(\mu I_n - A)x, x\rangle \geq 0$$

for any $x \in \mathbb{R}^n_{++}$*. Hence* $\mu I_n - A$ *is copositive. Therefore, since* $v^1 \in \mathbb{R}^n_+$ *we can apply Theorem 5.3.1 (vi)⇒(i) with* $\lambda_2 = \mu$ *to conclude that* q_A *is a spherically quasi-convex function.*

For instance, taking $\lambda < (\lambda+\nu)/2 < \mu < \nu$*, and the vectors* $v^1 = (e^1 + e^n)/\sqrt{2}, v^2 = e^2, \ldots, v^{n-1} = e^{n-1}, v^n = (e^1 - e^n)/\sqrt{2}$*, satisfy* (5.9)*. We can conclude that* q_A *is spherically quasi-convex.*

Example 5.3.2

Let $n \geq 3$ *and* $V = [v^1\ v^2\ v^3\ \cdots\ v^n] \in \mathbb{R}^{n\times n}$ *be an orthogonal matrix,* $\Lambda = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$ *and* $A = V\Lambda V^\top$ *. Then* q_A *is a spherically quasi-convex function, if*

$$v^1 = (v^1_1, \ldots, v^1_n)^\top \in \mathbb{R}^n_{++}, \qquad \lambda_1 < \lambda_2 \leq \cdots \leq \lambda_n \leq \lambda_2 + \frac{\alpha^2}{(n-2)}(\lambda_2 - \lambda_1), \tag{5.11}$$

where $\alpha := \min\left\{v^1_i : \ i \in \{1, \ldots, n\}\right\}$ *is the minimum (nonzero) element of eigenvector* v^1*. Indeed, by using* $V^\top V = I_n$ *and the definition of the matrix A, we obtain*

$$\begin{aligned}
\lambda_2\|x\|^2 - \langle Ax, x\rangle &= \lambda_2\langle V^\top Vx, x\rangle - \langle V\Lambda V^\top x, x\rangle \\
&= \sum_{i=1}^n \lambda_2\langle v^i, x\rangle^2 - \sum_{i=1}^n \lambda_i\langle v^i, x\rangle^2 \\
&= (\lambda_2 - \lambda_1)\langle v^1, x\rangle^2 + \sum_{i=3}^n (\lambda_2 - \lambda_i)\langle v^i, x\rangle^2.
\end{aligned} \tag{5.12}$$

Since $\lambda_2 - \lambda_1 > 0$ *and* $\lambda_2 - \lambda_n \leq \lambda_2 - \lambda_j \leq 0$*, for any* $j \in \{3, \ldots, n\}$*, from* (5.12) *we have*

$$\begin{aligned}
\lambda_2\|x\|^2 - \langle Ax, x\rangle &= (\lambda_2 - \lambda_1)\langle v^1, x\rangle^2 + \sum_{i=3}^n (\lambda_2 - \lambda_i)\langle v^i, x\rangle^2 \\
&= (\lambda_2 - \lambda_1)\left[\langle v^1, x\rangle^2 + \sum_{i=3}^n \frac{(\lambda_2 - \lambda_i)}{(\lambda_2 - \lambda_1)}\langle v^i, x\rangle^2\right] \\
&\geq (\lambda_2 - \lambda_1)\left[\langle v^1, x\rangle^2 + \sum_{i=3}^n \frac{\lambda_2 - \lambda_n}{\lambda_2 - \lambda_1}\langle v^i, x\rangle^2\right].
\end{aligned} \tag{5.13}$$

On the other hand, by using that $v_i^1 \in \mathbb{R}_{++}$ *and* $v_i^1 \geq \alpha$ *for any* $i \in \{1, \ldots, n\}$*, we conclude that*

$$\begin{aligned}\langle v^1, x\rangle^2 &= (v_1^1 x_1 + \cdots + v_n^1 x_n)^2 \\ &\geq \alpha^2 (x_1 + \cdots + x_n)^2 \geq \alpha^2 (x_1^2 + \cdots + x_n^2) = \alpha^2 \|x\|^2, \end{aligned} \tag{5.14}$$

for any $x \in \mathbb{R}^n_+$*. Moreover, taking into account that* $\|v^j\| = 1$ *for any* $j \in \{3, \ldots, n\}$*, applying Cauchy-Schwarz ineqauality, it follows that*

$$\langle v^3, x\rangle^2 + \cdots + \langle v^n, x\rangle^2 \leq \|v^3\|^2 \|x\|^2 + \cdots + \|v^n\|^2 \|x\|^2 \leq (n-2)\|x\|^2,$$

for any $x \in \mathbb{R}^n_+$*. Thus, combining the last inequalities with* (5.13) *and* (5.14) *and considering that the last inequality in* (5.11) *is equivalent to* $-\alpha^2 + (n-2)(\lambda_n - \lambda_2)/(\lambda_2 - \lambda_1) \leq 0$*, we have*

$$\begin{aligned}\lambda_2 \|x\|^2 - \langle Ax, x\rangle &\geq (\lambda_2 - \lambda_1) \left[\langle v^1, x\rangle^2 + \sum_{i=3}^{n} \frac{\lambda_2 - \lambda_n}{\lambda_2 - \lambda_1} \langle v^i, x\rangle^2 \right] \\ &\geq (\lambda_2 - \lambda_1) \left[\alpha^2 \|x\|^2 + \sum_{i=3}^{n} \frac{\lambda_2 - \lambda_n}{\lambda_2 - \lambda_1} \langle v^i, x\rangle^2 \right] \\ &\geq (\lambda_2 - \lambda_1) \left[\alpha^2 + (n-2) \frac{\lambda_2 - \lambda_n}{\lambda_2 - \lambda_1} \right] \|x\|^2 \geq 0,\end{aligned}$$

for any $x \in \mathbb{R}^n_+$*. Hence, we conclude that* $\lambda_2 I_n - A$ *is copositive. Therefore, since* $v^1 \in \mathbb{R}^n_{++}$ *is the eigenvector of* A *corresponding to the eigenvalue* λ_1*, we apply Theorem 5.3.1 (vi)⇒(i), to conclude that* q_A *is a spherically quasi-convex function.*

For instance, $n \geq 3$*,* $A = V \Lambda V^\top$*,* $\Lambda = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$*,* $V = [v^1 \; v^2 \; v^3 \; \cdots \; v^n] \in \mathbb{R}^{n \times n}$*, and* $\alpha = 1/\sqrt{n}$*,*

$$v^1 := \frac{1}{\sqrt{n}} \sum_{i=1}^{n} e^i, \quad v^j := \frac{1}{\sqrt{(n+1-j) + (n+1-j)^2}} \left[e^1 - (n+1-j) e^j + \sum_{i>j}^{n} e^i \right],$$

for $j \in \{2, \ldots, n\}$ *and* $\lambda_1 < \lambda_2 \leq \ldots \leq \lambda_n < \lambda_2 + (1/[n(n-2)])(\lambda_2 - \lambda_1)$*, satisfy the orthogonality of* V *and the condition* (5.11)*, therefore we conclude that* q_A *is a spherically quasi-convex.*

In the next theorem we establish the characterisation for a quasi-convex quadratic function q_A on the spherical positive orthant, where A is a symmetric matrix having only two distinct eigenvalues.

Theorem 5.3.2

Let $n \geq 3$ and $A \in \mathbb{R}^{n\times n}$ be a symmetric matrix with only two distinct eigenvalues, such that its smallest eigenvalue has multiplicity one. Then, q_A is spherically quasi-convex if and only if there is an eigenvector of A corresponding to the smallest eigenvalue with all components nonnegative.

Proof:

Let $A := (a_{ij}) \in \mathbb{R}^{n\times n}$, $\lambda_1, \lambda_2, \ldots, \lambda_n$ be the eigenvalues of A corresponding to an orthonormal set of eigenvectors $v^1, v^2, \ldots, v^n$, respectively. Then, we can assume with no loss of generality that $\lambda_1 =: \lambda < \mu := \lambda_2 = \cdots = \lambda_n$. Thus, we have

$$A = V\Lambda V^T, \qquad V := [v^1\ v^2\ \ldots\ v^n] \in \mathbb{R}^{n\times n}, \qquad \Lambda := \operatorname{diag}(\lambda, \mu, \ldots, \mu) \in \mathbb{R}^{n\times n}. \quad (5.15)$$

First we suppose that q_A is a spherically quasi-convex function. The matrix Λ can be equivalently written as follows

$$\Lambda = \mu I_n + (\lambda - \mu)D, \quad (5.16)$$

where $D := \operatorname{diag}(1, 0, \ldots, 0) \in \mathbb{R}^{n\times n}$. Then (5.16) and (5.15) imply

$$a_{ij} = (\lambda - \mu)v_i^1 v_j^1, \qquad i \neq j. \quad (5.17)$$

Since q_A is spherically quasi-convex and $e^i \in \mathcal{S} = \mathbb{S}^{n-1} \cap \mathbb{R}^n_{++}$ for any $i \in \{1, \ldots, n\}$, by using Theorem 5.3.1 (ii)⇒(i) we conclude that the matrix A is a Z-matrix, that is $a_{ij} \leq 0$ for any $i, j \in \{1, \ldots, n\}$ with $i \neq j$. Thus, owing that $\lambda < \mu$, we obtain form (5.17) that $0 \leq v_i^1 v_j^1$ for any $i \neq j$, which implies $v^1 \in \mathbb{R}^n_+$ (or $-v^1 \in \mathbb{R}^n_+$, they are the same because v^1 is an eigenvector). Therefore, there is an eigenvector corresponding to the smallest eigenvalue with all components nonnegative.

Conversely, suppose that $v^1 \in \mathbb{R}^n_+$. Then, applying Lemma 5.3.1 with $\lambda = \lambda_1 < \mu = \lambda_2 = \cdots = \lambda_n$ we conclude that $[\varphi_A \leq c]$ is convex for any $c \in \mathbb{R}$, and hence φ_A is quasi-convex by Proposition 1.5.3. Therefore, by using Theorem 5.2.1, we conclude that q_A is

spherically quasi-convex. □

In the following examples we present two classes of matrices satisfying the assumptions of Theorem 5.3.2.

Example 5.3.3

Let $v \in \mathbb{R}^n_+$ and define the Householder matrix $H := I_n - 2vv^\top/\|v\|^2$. The matrix H is nonsingular and symmetric. Moreover, the Householder matrix has a property that $Hv = -v$. Given that $\frac{vv^\top}{\|v\|^2}$ is a Rank-1 matrix, the characteristic polynomial of H is

$$\begin{aligned} p(\lambda) = |\lambda I_n - H| &= (\lambda-1)^n + Tr\left(2\frac{vv^\top}{\|v\|^2}\right)(\lambda-1)^{n-1} \\ &= |\lambda I_n - H| = (\lambda-1)^n + Tr\left(2\frac{v^\top v}{\|v\|^2}\right)(\lambda-1)^{n-1} \\ &= |\lambda I_n - H| = (\lambda-1)^n + 2(\lambda-1)^{n-1} \\ &= |\lambda I_n - H| = (\lambda+1)(\lambda-1)^{n-1}, \end{aligned}$$

we conclude that -1 and 1 are eigenvalues of H with multiplicities one and $n-1$, respectively.

Furthermore, the eigenvector corresponding to the smallest eigenvalue of H has all components nonnegative. Therefore, Theorem 5.3.2 implies that the quadratic function $q_H(x) = \langle Hx, x\rangle$ is spherically quasi-convex.

Example 5.3.4

Let $u \in \mathbb{R}^n$, $\alpha \in \mathbb{R}_{++}$. Define the matrix $A := I_n - \alpha uu^\top$. The eigenvalues of A are $\{\lambda_1, \lambda_2, \ldots, \lambda_n\} = \{-\alpha\langle u, u\rangle + 1, 1, \cdots, 1\}$. Thus, we have

$$A = V\Lambda V^T, \qquad V := [v^1\ v^2\ \ldots\ v^n] \in \mathbb{R}^{n\times n}, \qquad \Lambda := \mathrm{diag}(1-\alpha\langle u, u, \rangle, \mu, \ldots, \mu) \in \mathbb{R}^{n\times n}. \tag{5.18}$$

The matrix Λ can be equivalently written as follows

$$\Lambda = I_n - \alpha\langle u, u\rangle D, \tag{5.19}$$

where $D := \mathrm{diag}(1, 0, \ldots, 0) \in \mathbb{R}^{n\times n}$. *Then* (5.19) *and* (5.18) *imply*

$$a_{ij} = \begin{cases} -\alpha\langle u, u\rangle v_i^1 v_j^1, & i \neq j, \\ -\alpha\langle u, u\rangle (v_i^1)^2 + 1, & i = j. \end{cases}$$

Therefore, if $v^1 \in \mathbb{R}^n_+$, *Theorem 5.3.2 implies that the function* $q_A = \langle Ax, x\rangle$ *is spherically quasi-convex.*

Also, if $v^1 \in \mathbb{R}^n_+$, *then* A *is a Z-matrix, we have* $\lambda_1 < \lambda_2$ *and* $\lambda_2 \geq a_{ii}$ *for any* $i \in \{1, 2, \ldots, n\}$. *Theorem 5.3.1 (iv)⇒ (i) implies that the function* $q_A = \langle Ax, x\rangle$ *is spherically quasi-convex.*

In order to give a *complete characterisation of the spherical quasi-convexity of* q_A *for the case when* A *is diagonal*, in the following result we start with a necessary condition for q_A to be spherically quasi-convex on the spherical positive orthant.

Lemma 5.3.3

Let $n \geq 3$, $\mathcal{S} = \mathbb{S}^{n-1} \cap \mathbb{R}^n_{++}$ *and* $A \in \mathbb{R}^{n\times n}$ *be a nonsingular diagonal matrix. If* q_A *is spherically quasi-convex, then* A *has only two distinct eigenvalues, such that its smallest one has multiplicity one.*

Proof:

The proof will be made by contradiction. First we suppose that A *has at least three distinct eigenvalues, among which exactly two are negative, or at least two distinct eigenvalues, among which exactly one is negative and has multiplicity greater than one, i.e.,*

$$Ae^1 = -\lambda_1 e^1, \qquad Ae^2 = -\lambda_2 e^2, \quad Ae^3 = \lambda_3 e^3, \qquad \lambda_1, \lambda_2, \lambda_3 > 0 \tag{5.20}$$

with either $-\lambda_1 < -\lambda_2 < 0 < \lambda_3$ *or* $-\lambda_1 = -\lambda_2 < 0 < \lambda_3$ *and* e^1, e^2, e^3 *are canonical vectors of* $\mathbb{R}^n$. *Define the following two auxiliary vectors*

$$v^1 := e^1 + t_1 e^3, \qquad v^2 := e^2 + t_2 e^3, \qquad t_i = \sqrt{\frac{\lambda_i}{\lambda_3}}, \qquad i \in \{1, 2\}. \tag{5.21}$$

Hence, (5.20) *and* (5.21) *implies that*

$$\langle Av^1, v^1\rangle = \langle Ae^1, e^1\rangle + t_1^2\langle Ae^3, e^3\rangle = -\lambda_1\|e^1\| + \frac{\lambda_1}{\lambda_3}\lambda_3\|e^3\| = 0,$$

and

$$\langle Av^2, v^2\rangle = \langle Ae^2, e^2\rangle + t_2^2\langle Ae^3, e^3\rangle = -\lambda_2\|e^2\| + \frac{\lambda_2}{\lambda_3}\lambda_3\|e^3\| = 0.$$

Since $v^1, v^2 \in \mathbb{R}^n_+$, *we conclude that* $v^1, v^2 \in \left\{x \in \mathbb{R}^n_+ \ : \ \langle Ax, x\rangle \leq 0\right\}$. *However, using again* (5.20) *and* (5.21) *we obtain that*

$$\langle A(v^1+v^2), v^1+v^2\rangle = 2\langle Av^1, v^2\rangle = 2\frac{\sqrt{\lambda_1\lambda_2}}{\lambda_3}\lambda_3 = 2\sqrt{\lambda_1\lambda_2} > 0,$$

and therefore $v^1+v^2 \notin \left\{x \in \mathbb{R}^n_+ \ : \ \langle Ax, x\rangle \leq 0\right\}$. *We conclude that* $\left\{x \in \mathbb{R}^n_+ \ : \ \langle Ax, x\rangle \leq 0\right\}$ *is not a convex cone.*

Finally, suppose that A *has at least three distinct eigenvalues or at least two distinct ones with the smallest one having multiplicity greater than one. Let* λ, μ, ν *be eigenvalues of* A *such that either* $\lambda < \mu < \nu$ *or* $\lambda = \mu < \nu$. *Take a constant* $c \in \mathbb{R}$ *such that* $\mu < c < \nu$. *Letting* $A_c := A - cI_n$ *we conclude that* $\lambda - c, \mu - c, \nu - c$ *are eigenvalues of* A_c *and satisfying*

$$\lambda - c < \mu - c < 0 < \nu - c$$

or

$$\lambda - c = \mu - c < 0 < \nu - c.$$

Thus, by the first part of the proof, with A_c *in the role of* A, *we conclude that*

$$\left\{x \in \mathbb{R}^n_+ \ : \ \langle A_c x, x\rangle \leq 0\right\}$$

is not a convex cone. On the other hand, due to $e^i \in \mathbb{R}^n_+$ *and* $\langle Ae^i, e^i\rangle = \lambda - c < 0$, *for some* i, *we obtain that*

$$\left\{x \in \mathbb{R}^n_{++} \ : \ \langle A_c x, x\rangle < 0\right\} \neq \emptyset.$$

Henceforth, applying Corollary 5.2.2 with $\mathcal{K} = \mathbb{R}^n_+$ *and taking into account that both*

$$\left\{x \in \mathbb{R}^n_+ \ : \ \langle Ax, x\rangle \leq 0\right\}$$

in the first part of the proof, and

$$\left\{x \in \mathbb{R}^n_+ \ : \ \langle A_c x, x\rangle \leq 0\right\}$$

in the second part of the proof are not convex, we conclude that q_A is not spherically quasi-convex. □

To make the study self-contained we state the result of [27, Theorem 1] explicitly here:

Theorem 5.3.3

[27, Theorem 1] Let $\mathcal{S} = \mathbb{S}^{n-1} \cap \mathbb{R}^n_{++}$ and $A \in \mathbb{R}^{n\times n}$ be a symmetric matrix. Then, q_A is spherically convex if and only if there exists $\lambda \in \mathbb{R}$ such that $A = \lambda I_n$. In this case q_A is a constant function.

The next result gives a full characterisation for q_A to be spherically quasi-convex quadratic function on the spherical positive orthant, where *A is a diagonal matrix*. The proof of this result is a combination of Theorem 5.3.2, Lemma 5.3.3 and Theorem 5.3.3. Before presenting the result we need the following definition:

Definition 5.3.1 (Merely spherically quasi-convex)

A function is called merely spherically quasi-convex *if it is spherically quasi-convex, but it is not spherically convex.*

Theorem 5.3.4

Let $n \geq 3$ and $A \in \mathbb{R}^{n\times n}$ be a nonsingular diagonal matrix. Then q_A is merely spherically quasi-convex if and only if A has only two eigenvalues, such that its smallest one has multiplicity one and has a corresponding eigenvector with all components nonnegative.

Proof:

Given that $A \in \mathbb{R}^{n\times n}$ is a nonsingular diagonal matrix. We suppose q_A is a merely spherically quasi-convex function defined on $\mathcal{S} = \mathbb{S}^{n-1} \cap \mathbb{R}^n_{++}$. By applying Lemma 5.3.3, we conclude that A has only two eigenvalues, such that its smallest one has multiplicity one and has a corresponding eigenvector with all components nonnegative.

Conversely, suppose that A is a nonsingular diagonal (therefore, symmetric) matrix and has only two eigenvalues, such that its smallest one has multiplicity one and has a

corresponding eigenvector with all components nonnegative. By Theorem 5.3.3 we conclude that q_A is not spherically convex because $A \neq \lambda I_n$ for any $\lambda \in \mathbb{R}$. Using Theorem 5.3.2, we obtain that q_A is spherically quasi-convex. Therefore q_A is merely spherically quesi-convex.

□

We end this section by showing that, if a symmetric matrix A has three eigenvectors in the nonnegative orthant associated to at least two distinct eigenvalues, then the associated quadratic function q_A cannot be spherically quasi-convex.

Lemma 5.3.4

Let $n \geq 3$ and $v^1, v^2, v^3 \in \mathbb{R}^n$ be distinct eigenvectors of a symmetric matrix A associated to the eigenvalues $\lambda_1, \lambda_2, \lambda_3 \in \mathbb{R}$, respectively, among which at least two are distinct. If q_A is spherically quasi-convex, then $v^i \notin \mathbb{R}^n_+$ for some $i \in \{1, 2, 3\}$.

Proof:

Assume by contradiction that $v^i \in \mathbb{R}^n_+$ for any $i \in \{1, 2, 3\}$. Without loss of generality we can also assume that $\|v^i\| = 1$, for $i \in \{1, 2, 3\}$. Given that at least two eigenvalues are distinct, we have three possibilities: $\lambda_1 < \lambda_2 < \lambda_3$, $\lambda_1 = \lambda_2 < \lambda_3$ or $\lambda_1 < \lambda_2 = \lambda_3$. We start by analysing the possibilities $\lambda_1 < \lambda_2 < \lambda_3$ or $\lambda_1 = \lambda_2 < \lambda_3$. First we assume that either $\lambda_1 < \lambda_2 < 0 < \lambda_3$ or $\lambda_1 = \lambda_2 < 0 < \lambda_3$. Define the following auxiliary vectors

$$w^1 := v^1 + t_1 v^3, \qquad w^2 := v^2 + t_2 v^3, \qquad t_1 := \sqrt{\frac{-\lambda_1}{\lambda_3}}, \qquad t_2 := \sqrt{\frac{-\lambda_2}{\lambda_3}}. \tag{5.22}$$

We have $\langle v^i, v^j \rangle = 0$ for any $i, j \in \{1, 2, 3\}$ with $i \neq j$, and since

$$Av^1 = \lambda_1 v^1, \qquad Av^2 = \lambda_2 v^2, \quad Av^3 = \lambda_3 v^3, \qquad v^1, v^2, v^3 \in \mathbb{R}^n_+, \tag{5.23}$$

we conclude from (5.22) that

$$\langle Aw^1, w^1 \rangle = \langle Av^1, v^1 \rangle + \frac{-\lambda_1}{\lambda_3} \langle Av^3, v^3 \rangle = \lambda_1 \|v^1\| + \frac{-\lambda_1}{\lambda_3} \lambda_3 \|v^3\| = 0,$$

and

$$\langle Aw^2, w^2 \rangle = \langle Av^2, v^2 \rangle + \frac{-\lambda_2}{\lambda_3} \langle Av^3, v^3 \rangle = \lambda_2 \|v^2\| + \frac{-\lambda_2}{\lambda_3} \lambda_3 \|v^3\| = 0.$$

Moreover, since $v^1, v^2, v^3 \in \mathbb{R}^n_+$ *we conclude that* $w^1, w^2 \in \left\{x \in \mathbb{R}^n_+ \ : \ \langle Ax, x\rangle \leq 0\right\}$. *On the other hand, by using* (5.23) *and* (5.22), *we obtain that*

$$\langle A(w^1 + w^2), w^1 + w^2\rangle = 2\langle Aw^1, w^2\rangle = 2t_1t_2\langle Av^3, v^3\rangle = 2\frac{\sqrt{\lambda_1\lambda_2}}{\lambda_3}\lambda_3 = 2\sqrt{\lambda_1\lambda_2} > 0,$$

hence $w^1 + w^2 \notin \left\{x \in \mathbb{R}^n_+ \ : \ \langle Ax, x\rangle \leq 0\right\}$. *Thus,* $\left\{x \in \mathbb{R}^n_+ \ : \ \langle Ax, x\rangle \leq 0\right\}$ *is not a convex cone.*

For the general case, take $c \in \mathbb{R}$ *such that* $\lambda_2 < c < \lambda_3$. *Letting* $A_c := A - cI_n$ *we conclude that* $\lambda_1 - c, \lambda_2 - c, \lambda_3 - c$ *are eigenvalues of* A_c *and satisfying*

$$\lambda_1 - c < \lambda_2 - c < 0 < \lambda_3 - c$$

or

$$\lambda_1 - c = \lambda_2 - c < 0 < \lambda_3 - c$$

with the three corresponding orthonormal eigenvectors $v^1, v^2, v^3 \in \mathbb{R}^n_+$. *Hence, by the first part of the proof, with* A_c *in the role of* A, *we conclude that the cone* $\left\{x \in \mathbb{R}^n_+ \ : \ \langle A_cx, x\rangle \leq 0\right\}$ *is not convex. On the other hand, due to* $v^1 \in \mathbb{R}^n_+$ *and* $\langle Av^1, v^1\rangle = \lambda_1 - c < 0$, *we have* $\left\{x \in \mathbb{R}^n_{++} \ : \ \langle A_cx, x\rangle < 0\right\} \neq \emptyset$. *Thus, applying Corollary 5.2.2 with* $\mathcal{K} = \mathbb{R}^n_+$ *and taking into account that* $\left\{x \in \mathbb{R}^n_+ \ : \ \langle A_cx, x\rangle \leq 0\right\}$ *is not convex, we conclude that* q_A *is not spherically quasi-convex.*

To analyse the possibility $\lambda_1 < \lambda_2 = \lambda_3$, *first assume that* $\lambda_1 < 0 < \lambda_2 = \lambda_3$ *and define the vectors*

$$w^1 := t_1v^1 + v^3, \qquad w^2 := t_2v^1 + v^3, \qquad t_1 = \sqrt{\frac{\lambda_2}{-\lambda_1}}, \qquad t_2 = \sqrt{\frac{\lambda_3}{-\lambda_1}},$$

and then proceed as above to obtain again a contradiction. Therefore, $v^i \notin \mathbb{R}^n_+$ *for some* $i \in \{1, 2, 3\}$. □

5.4 Spherically quasi-convex quadratic functions on the subdual convex sets

In this chapter we present a condition characterising the spherical quasi-convexity of quadratic functions on spherically subdual convex sets associated to subdual cones. The results obtained generalise the corresponding ones obtained in previous chapter (or in our published paper [28, Section 4.1]). We also summarised the results of this chapter in [29]. *Throughout this chapter we follow the conventions used before and assume that the cone $\mathcal{K}$ is a subdual (i.e., $\mathcal{K} \subseteq \mathcal{K}^*$)* and proper cone. A closed set $\mathcal{S} \subseteq \mathbb{S}^{n-1}$ is called a *spherically subdual convex set* if the associated cone $\mathcal{K}_{\mathcal{S}}$ (defined in (1.14)) is subdual. It is clear that if $A = A^\top \in \mathbb{R}^{n\times n}$ has only one eigenvalue, then q_A is constant and, consequently, it is spherically quasi-convex. Henceforth, throughout this chapter *we assume that A has at least two distinct eigenvalues.* We remind that q_A and φ_A are defined in (5.4) and (5.5), respectively. Two technical lemmas, which are useful in the following text, will be presented. They are generalisations of Lemma 5.3.1 and 5.3.2 (or, Lemmas 14 and 15 of [28]), respectively. For stating the next lemma, denote by $\{v^1, v^2, \ldots, v^n\}$ a orthonormal system of eigenvectors of A corresponding to the eigenvalues $\lambda_1 < \lambda_2 \leq \ldots \leq \lambda_n$. Given $c \in (\lambda_1, \lambda_2]$, we define the following convex cone

$$\mathcal{L}_c := \left\{ x \in \mathbb{R}^n : \ \langle v^1, x\rangle \geq \sqrt{\sum_{i=2}^{n} \theta_i(c)\langle v^i, x\rangle^2} \right\}, \qquad \theta_i(c) := \frac{\lambda_i - c}{c - \lambda_1}, \tag{5.24}$$

for $i \in \{2, \ldots, n\}$. Note that if $\lambda_1 < c < \lambda_2$, then $\theta_i(c) > 0$, for $i \in \{2, \ldots, n\}$, and both $\mathcal{L}_c$ and $-\mathcal{L}_c$ are proper cones (recall that a proper cone is a closed, convex, pointed cone with nonempty interior). We also need to consider the following cone

$$\mathcal{W} := (\mathcal{L}_{\lambda_2} \cup -\mathcal{L}_{\lambda_2}) \cap \operatorname{int}(\mathcal{K}). \tag{5.25}$$

Considering that $\mathcal{K}$, $\mathcal{L}_c$, and $-\mathcal{L}_c$ are proper cones, as a conclusion the cone $\mathcal{W}$ is also a proper cone, and $\operatorname{int}(\mathcal{W}) \neq \emptyset$. The following lemma is a general version of Lemma 5.3.1:

Lemma 5.4.1

Let $n \geq 2$, $A = A^\top \in \mathbb{R}^{n\times n}$ and $\{v^1, v^2, \ldots, v^n\}$ be an orthonormal system of eigenvectors of A corresponding to the eigenvalues $\lambda_1 < \lambda_2 \leq \ldots \leq \lambda_n$, respectively. Then, the sublevel set $[\varphi_A \leq c]$ is convex for any $c \notin (\lambda_2, \lambda_n)$ if and only if $v^1 \in \mathcal{W}^* \cup -\mathcal{W}^*$. In particular if

$v^1 \in \mathcal{K}^*$, *then* $[\varphi_A \leq c]$ *is convex for any* $c \notin (\lambda_2, \lambda_n)$.

Proof:

By using the spectral decomposition of A, *we have* $A = V\Lambda V^\top = \sum_{i=1}^n \lambda_i v^i (v^i)^\top$. *From the definition* (5.5) *we have*

$$\begin{aligned}
[\varphi_A \leq c] &= \left\{ x \in \operatorname{int}(\mathcal{K}) : \ \frac{\langle Ax, x\rangle}{\|x\|^2} \leq c \right\} \\
&= \{ x \in \operatorname{int}(\mathcal{K}) : \ \langle (A - cI_n)x, x\rangle \leq 0 \} \\
&= \left\{ x \in \operatorname{int}(\mathcal{K}) : \ \sum_{i=1}^n (\lambda_i - c)\langle v^i, x\rangle^2 \leq 0 \right\} \qquad (5.26)
\end{aligned}$$

If $\lambda_1 < c \leq \lambda_2$, *then by using* (5.24), *the equality* (5.26) *can be completed as follows*

$$\begin{aligned}
\mathcal{W} = [\varphi_A \leq \lambda_2] \supseteq [\varphi_A \leq c] &= (\mathcal{L}_c \cup -\mathcal{L}_c) \cap \operatorname{int}(\mathcal{K}) \\
&= \left\{ x \in \operatorname{int}(\mathcal{K}) : \ \langle v^1, x\rangle^2 \geq \sum_{i=2}^n \theta_i(c)\langle v^i, x\rangle^2 \right\}. \qquad (5.27)
\end{aligned}$$

Sufficiency of the first statement:

Let $v^1 \in \mathcal{W}^*$ *(a similar argument holds for* $v^1 \in -\mathcal{W}^*$*).*

If $c < \lambda_1$, *then considering that* $v^1, v^2, \ldots, v^n$ *are linearly independent and* $0 \notin \operatorname{int}(\mathcal{K})$, *we obtain from* (5.26) *that* $\sum_{i=1}^n (\lambda_i - c)\langle v^i, x\rangle^2 > 0$ *for any* $x \in \operatorname{int}(\mathcal{K})$ *and hence* $[\varphi_A \leq c] = \emptyset$ *is convex.*

If $c = \lambda_1$, *then* (5.26) *implies that* $[\varphi_A \leq c] = \mathcal{S} \cap \operatorname{int}(\mathcal{K})$, *where*

$$\begin{aligned}
\mathcal{S} &:= \left\{ x \in \mathbb{R}^n : \ \sum_{i=2}^n (\lambda_i - c)\langle v^i, x\rangle^2 = 0 \right\} \\
&= \left\{ x \in \mathbb{R}^n \ : \langle v^i, x\rangle = 0, \ i \in \{2, \ldots, n\} \right\}.
\end{aligned}$$

Thus, due to $\operatorname{int}(\mathcal{K})$ *and* $\mathcal{S}$ *being convex, we conclude that* $[\varphi_A \leq c]$ *is also convex.*

If $\lambda_1 < c \leq \lambda_2$. *Since* $v^1 \in \mathcal{W}^*$, *for any* $x \in \mathcal{W}$ *we obtain that* $\langle v^1, x\rangle \geq 0$ *and from* (5.27) *we have* $-\mathcal{L}_c \cap \operatorname{int}(\mathcal{K}) = \emptyset$ *and hence* $[\varphi_A \leq c] = \mathcal{L}_c \cap \operatorname{int}(\mathcal{K})$. *Due to the convexity of the cones* $\mathcal{L}_c$ *and* $\operatorname{int}(\mathcal{K})$, *we obtain that* $[\varphi_A \leq c]$ *is convex.*

Finally, if $c \geq \lambda_n$, *then* (5.26) *implies that* $\sum_{i=1}^n (\lambda_i - c)\langle v^i, x\rangle^2 \leq 0$ *for any* $x \in \operatorname{int}(\mathcal{K})$ *and hence* $[\varphi_A \leq c] = \operatorname{int}(\mathcal{K})$ *is convex.*

Necessity of the first statement:

We will show that $v^1 \notin \mathcal{W}^* \cup -\mathcal{W}^*$ *implies* $[\varphi_A \leq c]$ *is not convex, for some* $c \in (\lambda_1, \lambda_2)$. *Suppose that* $v^1 \notin \mathcal{W}^* \cup -\mathcal{W}^*$. *Thus, considering that* $\operatorname{int}(\mathcal{W}) \neq \emptyset$, *there exist* $y, z \in \operatorname{int}(\mathcal{W})$ *such that* $\langle v^1, y\rangle > 0$ *and* $\langle v^1, z\rangle < 0$. *Thus,* (5.24) *and* (5.25) *implies that*

$$y \in \operatorname{int}(\mathcal{K}) \cap \operatorname{int}(\mathcal{L}_{\lambda_2}), \qquad z \in \operatorname{int}(\mathcal{K}) \cap \operatorname{int}(-\mathcal{L}_{\lambda_2}). \tag{5.28}$$

We claim that there exists a $\bar{c} \in (\lambda_1, \lambda_2)$ *such that* $y \in \operatorname{int}(\mathcal{K}) \cap \operatorname{int}(\mathcal{L}_{\bar{c}})$ *and* $z \in \operatorname{int}(\mathcal{K}) \cap \operatorname{int}(-\mathcal{L}_{\bar{c}})$. *In order to simplify the notations, for* $x \in \mathbb{R}^n$ *and* $c \in (\lambda_1, \lambda_2]$, *we define the following function*

$$\xi(x, c) := \sqrt{\sum_{i=2}^{n} \theta_i(c) \langle v^i, x\rangle^2}. \tag{5.29}$$

Note that ξ *is a continuous function and, from the definition of* θ_i *in* (5.24), *it is also decreasing with respect to the second variable* c. *By using* (5.24) *and* (5.29) *we have*

$$\operatorname{int}(\mathcal{K}) \cap \operatorname{int}(\mathcal{L}_c) = \left\{x \in \operatorname{int} \mathcal{K} : \ \langle v^1, x\rangle > \xi(x, c)\right\}, \qquad \forall\ c \in (\lambda_1, \lambda_2]. \tag{5.30}$$

Thus, taking into account the first inclusion in (5.28) *we conclude, by setting* $c = \lambda_2$ *in* (5.30), *that*

$$\lim_{c \to \lambda_2} \xi(y, c) = \xi(y, \lambda_2) < \langle v^1, y\rangle.$$

Hence, there exists a $\hat{c} \in (\lambda_1, \lambda_2)$ *sufficiently close to* λ_2 *such that* $\xi(y, \hat{c}) < \langle v^1, y\rangle$. *Similarly, we can also prove that there exists a* $\tilde{c} \in (\lambda_1, \lambda_2)$ *sufficiently close to* λ_2 *such that* $\xi(z, \tilde{c}) < -\langle v^1, z\rangle$. *Thus, letting* $\bar{c} = \max\{\hat{c}, \tilde{c}\}$ *we conclude that* $\xi(y, \bar{c}) < \langle v^1, y\rangle$ *and* $\xi(z, \bar{c}) < -\langle v^1, z\rangle$, *which by* (5.29) *and* (5.30) *yields*

$$y \in \operatorname{int}(\mathcal{L}_{\bar{c}}), \qquad z \in \operatorname{int}(-\mathcal{L}_{\bar{c}}). \tag{5.31}$$

We know by (5.28) *that* $y \in \operatorname{int}(\mathcal{K})$ *and* $z \in \operatorname{int}(\mathcal{K})$, *which together with* (5.31) *yields* $y \in \operatorname{int}(\mathcal{K}) \cap \operatorname{int}(\mathcal{L}_{\bar{c}})$ *and* $z \in \operatorname{int}(\mathcal{K}) \cap \operatorname{int}(-\mathcal{L}_{\bar{c}})$ *and the claim is concluded. Therefore, there exist* $r_y > 0$ *and* $r_z > 0$ *such that* $\boldsymbol{B}(y, r_y) \subset \operatorname{int}(\mathcal{K}) \cap \operatorname{int}(\mathcal{L}_{\bar{c}})$ *and* $\boldsymbol{B}(z, r_z) \subset \operatorname{int}(\mathcal{K}) \cap \operatorname{int}(-\mathcal{L}_{\bar{c}})$, *where* $\boldsymbol{B}(y, r_y)$ *and* $\boldsymbol{B}(z, r_z)$ *denote the open balls with centers* y, z *and radius* $r_y > 0$,

$r_z > 0$, *respectively. Hence, by dimensionality reasons, we can take* $u_y \in \text{int}(\mathcal{K}) \cap \text{int}(\mathcal{L}_{\bar{c}})$ *and* $u_z \in \text{int}(\mathcal{K}) \cap \text{int}(-\mathcal{L}_{\bar{c}})$ *such that* v^1, u_y *and* u_z *are linearly independent (l.i.). Thus, in particular, we have* $0 \notin [u_y, u_z]$, *where* $[u_y, u_z]$ *denotes the straight line segment joining* u_y *to* u_z. *Since* $\text{int}(\mathcal{L}_{\bar{c}}) \cap \text{int}(-\mathcal{L}_{\bar{c}}) = \emptyset$ *and* $0 \notin [u_y, u_z]$, *the segment* $[u_y, u_z]$ *is intersecting, at the distinct points* $w_y \neq 0$ *and* $w_z \neq 0$, *the boundaries of the sets* $\text{int}(\mathcal{L}_{\bar{c}})$ *and* $\text{int}(-\mathcal{L}_{\bar{c}})$, *respectively. Moreover, due to* u_y *and* u_z *being l.i.,* $0 \notin [u_y, u_z]$ *and* $w_y, w_z \in [u_y, u_z]$, *we conclude that the vectors* v^1, w_y *and* w_z *are also l.i.. Our next task is to prove that*

$$\frac{1}{2}(w_y + w_z) \notin \mathcal{L}_{\bar{c}} \cup -\mathcal{L}_{\bar{c}}. \tag{5.32}$$

First, due to w_y *and* w_z *belonging to the boundaries of* $\mathcal{L}_{\bar{c}}$ *and* $-\mathcal{L}_{\bar{c}}$, *respectively, we obtain from* (5.24) *that*

$$\langle v^1, w_y \rangle = \sqrt{\sum_{i=2}^{n} \theta_i(\bar{c}) \langle v^i, w_y \rangle^2}, \qquad \langle v^1, w_z \rangle = -\sqrt{\sum_{i=2}^{n} \theta_i(\bar{c}) \langle v^i, w_z \rangle^2}. \tag{5.33}$$

On the other hand, by using the two equalities in (5.33), *we obtain after some algebraic manipulations that*

$$\begin{aligned}\sum_{i=2}^{n} \theta_i(\bar{c}) \left\langle v^i, \frac{1}{2}(w_y + w_z) \right\rangle^2 &= \sum_{i=2}^{n} \theta_i(\bar{c}) \left(\left\langle v^i, \frac{1}{2} w_y \right\rangle + \left\langle v^i, \frac{1}{2} w_z \right\rangle \right)^2 \\ &= \left\langle v^1, \frac{1}{2} w_y \right\rangle^2 + \left\langle v^1, \frac{1}{2} w_z \right\rangle^2 + 2 \sum_{i=2}^{n} \theta_i(\bar{c}) \left\langle v^i, \frac{1}{2} w_y \right\rangle \left\langle v^i, \frac{1}{2} w_z \right\rangle.\end{aligned}$$

Thus, considering that

$$\left\langle v^1, \frac{1}{2}(w_y + w_z) \right\rangle^2 = \left\langle v^1, \frac{1}{2} w_y \right\rangle^2 + \left\langle v^1, \frac{1}{2} w_z \right\rangle^2 + 2 \left\langle v^1, \frac{1}{2} w_y \right\rangle \left\langle v^1, \frac{1}{2} w_z \right\rangle,$$

we have

$$\begin{aligned}\sum_{i=2}^{n} \theta_i(\bar{c}) \left\langle v^i, \frac{1}{2}(w_y + w_z) \right\rangle^2 &= \left\langle v^1, \frac{1}{2}(w_y + w_z) \right\rangle^2 - 2 \left\langle v^1, \frac{1}{2} w_y \right\rangle \left\langle v^1, \frac{1}{2} w_z \right\rangle \\ &\quad + 2 \sum_{i=2}^{n} \theta_i(\bar{c}) \left\langle v^i, \frac{1}{2} w_y \right\rangle \left\langle v^i, \frac{1}{2} w_z \right\rangle. \end{aligned} \tag{5.34}$$

Applying Cauchy-Schwarz inequality and then, using again both equalities in (5.33), *we*

conclude that

$$-\sum_{i=2}^{n}\theta_i(\bar{c})\left\langle v^i,\frac{1}{2}w_y\right\rangle\left\langle v^i,\frac{1}{2}w_z\right\rangle \leq \sqrt{\sum_{i=2}^{n}\theta_i(\bar{c})\langle v^i,w_y\rangle^2}\sqrt{\sum_{i=2}^{n}\theta_i(\bar{c})\langle v^i,w_z\rangle^2} \tag{5.35}$$
$$= -\left\langle v^1,\frac{1}{2}w_y\right\rangle\left\langle v^1,\frac{1}{2}w_z\right\rangle.$$

We are going to prove that the inequality (5.35) *is strict. For that, assume the contrary, i.e., that the last inequality holds as equality. In this case, there exists* $\alpha \neq 0$ *such that*

$$\Big(\sqrt{\theta_2(\bar{c})}\left\langle v^2,\frac{1}{2}w_y\right\rangle,\ldots,\sqrt{\theta_n(\bar{c})}\left\langle v^n,\frac{1}{2}w_y\right\rangle\Big)$$
$$=\alpha\Big(\sqrt{\theta_2(\bar{c})}\left\langle v^2,-\frac{1}{2}w_z\right\rangle,\ldots,\sqrt{\theta_n(\bar{c})}\left\langle v^n,-\frac{1}{2}w_z\right\rangle\Big),$$

which implies that $w_y+\alpha w_z$ *is orthogonal to the set of vectors* $\{v^2,\ldots,v^n\}$*. Thus, since the set* $\{v^1,v^2,\ldots,v^n\}$ *is an orthonormal system,* $w_y+\alpha w_z$ *is parallel to the vector* v^1*, which is absurd due to vectors* v^1*,* w_y *and* w_z *being l.i.. Hence,* (5.35) *holds strictly and combining it with* (5.34) *we conclude that*

$$\sum_{i=2}^{n}\theta_i(\bar{c})\left\langle v^i,\frac{1}{2}(w_y+w_z)\right\rangle^2 > \left\langle v^1,\frac{1}{2}(w_y+w_z)\right\rangle^2,$$

and (5.32) *holds. Therefore, considering that* $\frac{1}{2}(w_y+w_z)\in(u_y,u_z)$*, we conclude that* $(u_y,u_z)\not\subset\mathcal{L}_{\bar{c}}\cup-\mathcal{L}_{\bar{c}}$*. Thus, using notation* (5.27)*, we also have* $(u_y,u_z)\not\subset(\mathcal{L}_{\bar{c}}\cup-\mathcal{L}_{\bar{c}})\cap\operatorname{int}(\mathcal{K})=[\varphi_A\leq\bar{c}]$*, and due to* $u_y,u_z\in(\mathcal{L}_{\bar{c}}\cup-\mathcal{L}_{\bar{c}})\cap\operatorname{int}(\mathcal{K})=[\varphi_A\leq\bar{c}]$*, it follows that* $[\varphi_A\leq\bar{c}]$ *is not convex.*

Proof of second statement: It follows from $\mathcal{K}^*\subseteq\mathcal{W}^*$*.* □

Remark 5.4.1

It is easy to check that for any two cones $\mathcal{A}\subset\mathbb{R}^n$ *and* $\mathcal{B}\subset\mathbb{R}^n$*, we have* $(\mathcal{A}\cap\mathcal{B})=(\mathcal{A}^*+\mathcal{B}^*)$*. The dual of* $\mathcal{W}$ *in* (5.25)*can be expressed as*

$$\begin{aligned}\mathcal{W}^*=[(\mathcal{K}\cap\mathcal{L}_{\lambda_2})\cup(\mathcal{K}\cap-\mathcal{L}_{\lambda_2})]^* &= (\mathcal{K}\cap\mathcal{L}_{\lambda_2})^*\cap(\mathcal{K}\cap-\mathcal{L}_{\lambda_2})^*\\ &= (\mathcal{K}^*+\mathcal{L}^*_{\lambda_2})\cap(\mathcal{K}^*-\mathcal{L}^*_{\lambda_2}).\end{aligned} \tag{5.36}$$

Corollary 5.4.1

Suppose that $n \geq 3$ and $\lambda_2 \leq (\lambda_1 + \lambda_3)/2$. If either $\mathcal{K} \cap -\mathcal{L}_{\lambda_2} = \{0\}$ or $\mathcal{K} \cap \mathcal{L}_{\lambda_2} = \{0\}$, then $[\varphi_A \leq c]$ is convex for any $c \notin (\lambda_2, \lambda_n)$.

Proof:

First note that if $n \geq 3$ and $\lambda_2 \leq (\lambda_1 + \lambda_3)/2$, then $\theta_i(\lambda_2) \geq 1$ for any $i \geq 3$. Define the cone

$$\mathcal{L}_{[v^2]^\perp} := \left\{ x \in \mathbb{R}^n : \ \langle v^1, x \rangle \geq \sqrt{\sum_{i=3}^{n} \langle v^i, x \rangle^2} \right\}.$$

Note that $\mathcal{L}_{[v^2]^\perp}$ is a self-dual Lorentz cone as a subset of the subspace $[v^2]^\perp$. Moreover, considering that $\theta_i(\lambda_2) \geq 1$ for any $i \geq 3$, we conclude

$$\mathcal{L}_{\lambda_2} \cap [v^2]^\perp = \left\{ x \in \mathbb{R}^n : \ \langle v^1, x \rangle \geq \sqrt{\sum_{i=3}^{n} \theta(\lambda_2) \langle v^i, x \rangle^2} \right\} \subseteq \mathcal{L}_{[v^2]^\perp}.$$

Consequently, taking into account that $\mathcal{L}_{[v^2]^\perp}$ is a self-dual cone, the cone $\mathcal{L}_{\lambda_2} \cap [v^2]^\perp$ is subdual as a subset of the subspace $[v^2]^\perp$. To simplify the notation, denote by upper star (i.e., *) the dual of a cone in $\mathbb{R}^n$ and by lower star (i.e., $_*$) the dual of a cone in $[v^2]^\perp$. Thus, using this notation we will prove

$$\mathcal{L}^*_{\lambda_2} = (\mathcal{L}_{\lambda_2} \cap [v^2]^\perp)_* \tag{5.37}$$

Indeed, since $v^2, -v^2 \in \mathcal{L}_{\lambda_2}$, for any $z \in \mathcal{L}^*_{\lambda_2}$, we have $\langle z, v^2 \rangle = 0$ and hence $\mathcal{L}^*_{\lambda_2} \subseteq [v^2]^\perp$, which implies $\mathcal{L}^*_{\lambda_2} \subseteq (\mathcal{L}_{\lambda_2} \cap [v^2]^\perp)_*$.

Conversely, arbitrarily take $u \in (\mathcal{L}_{\lambda_2} \cap [v^2]^\perp)_*$, and take $w \in \mathcal{L}_{\lambda_2} \cap [v^2]^\perp$ then for any $t \in \mathbb{R}$ we have $v = w + tv^2 \in \mathcal{L}_{\lambda_2}$. Hence, $\langle u, v \rangle = \langle u, w \rangle \geq 0$, which implies that $u \in \mathcal{L}^*_{\lambda_2}$. Hence, we conclude that $(\mathcal{L}_{\lambda_2} \cap [v^2]^\perp)_* \subseteq \mathcal{L}^*_{\lambda_2}$, and (5.37) is proved. Suppose $\mathcal{K} \cap -\mathcal{L}_{\lambda_2} = \{0\}$, by using the first equality in (5.36) we obtain $\mathcal{W}^* = (\mathcal{K} \cap \mathcal{L}_{\lambda_2})^*$. Therefore, considering that $\mathcal{L}_{\lambda_2} \cap [v^2]^\perp$ is subdual and (5.37), we obtain

$$v^1 \in \mathcal{L}_{\lambda_2} \cap [v^2]^\perp \subseteq (\mathcal{L}_{\lambda_2} \cap [v^2]^\perp)_* = \mathcal{L}^*_{\lambda_2} \subseteq (\mathcal{K} \cap \mathcal{L}_{\lambda_2})^* = \mathcal{W}^*.$$

Hence, following Lemma 5.4.1 we conclude that $[\varphi_A \leq c]$ is convex for any $c \notin (\lambda_2, \lambda_n)$. The case $\mathcal{K} \cap \mathcal{L}_{\lambda_2} = \{0\}$ can be proved similarly. □

Lemma 5.4.2

Let $n \geq 3$ and $B = B^\top \in \mathbb{R}^{n\times n}$. Let $\mu_1 \leq \mu_2 \leq \ldots \leq \mu_n$ be eigenvalues of the matrix B. Assume that one of the following two conditions holds:

(a) $\mu_1 = \mu_2 < 0 < \mu_n$;

(b) $\mu_1 < \mu_2 < 0 < \mu_n$.

Then, for any $\bar{x} \in \mathbb{R}^n \setminus \{0\}$ such that $B\bar{x} \neq 0$ and $\langle B\bar{x}, \bar{x}\rangle = 0$, and any number $\delta > 0$, the set $\Xi(B, \bar{x}, \delta) := \{x \in \mathbb{R}^n : \|x - \bar{x}\| \leq \delta,\ \langle Bx, x\rangle \leq 0\}$ is not convex.

Proof:

Since $\mu_1 = \min_{x\in\mathbb{S}^{n-1}} q_B(x) < \max_{x\in\mathbb{S}^{n-1}} q_B(x) = \mu_n$, we can take $\bar{x} \in \mathbb{R}^n \setminus \{0\}$ such that $B\bar{x} \neq 0$ and $\langle B\bar{x}, \bar{x}\rangle = 0$. Define the following vector subspace of $\mathbb{R}^n$:

$$\mathcal{N} := [\{u \in \mathbb{R}^n :\ Bu = \mu u,\ \text{for some } \mu < 0\}].$$

It follows from assumption (a) or (b) that $\dim(\mathcal{N}) \geq 2$. For simplifying the notation we set

$$\bar{y} := B\bar{x} \neq 0. \tag{5.38}$$

To proceed with the proof, we first need to prove that $\mathcal{N} \neq [\bar{y}]^\perp$. Assume to the contrary that $\mathcal{N} = [\bar{y}]^\perp$. In this case, due to (5.38) and $B = B^\top$, the definition of $[\bar{y}]^\perp$ implies that

$$\langle Bv, \bar{x}\rangle = \langle B\bar{x}, v\rangle = \langle \bar{y}, v\rangle = 0, \quad \forall v \in \mathcal{N}. \tag{5.39}$$

Thus, it follows from the definition of $\mathcal{N}$ that $\langle Bv, \bar{x}\rangle = \langle v, \bar{x}\rangle = 0$, for any $v \in \mathcal{N}$, which implies $\mathcal{N} \subset [\bar{x}]^\perp := \{v \in \mathbb{R}^n :\ \langle v, \bar{x}\rangle = 0\}$. Moreover, considering that $\langle \bar{y}, \bar{x}\rangle = 0$, we also have $\bar{y} \in [\bar{x}]^\perp$. Hence, we conclude that $[\bar{y}] + \mathcal{N} \subset [\bar{x}]^\perp$. Since by definition (5.38) we have$\bar{y} \neq 0$, then we conclude $\bar{y} \notin [\bar{y}]^\perp = \mathcal{N}$. Due to $\bar{y} \neq 0$ and $\mathcal{N} = [\bar{y}]^\perp$ we have $\dim([\bar{y}] + \mathcal{N}) = n$. Combining with the fact that $[\bar{y}] + \mathcal{N} \subset [\bar{x}]^\perp$, we obtain $\bar{x} = 0$, which contradicts the assumption $\bar{x} \neq 0$. Therefore, $\mathcal{N} \neq [\bar{y}]^\perp$. Thus, we have

$$\dim(\mathcal{N} \cap [\bar{y}]^\perp) \geq \dim \mathcal{N} + \dim [\bar{y}]^\perp - \dim \mathbb{R}^n \geq 2 + (n-1) - n = 1.$$

Hence, there exist a unit vector $a \in \mathcal{N} \cap [\bar{y}]^{\perp}$, *so that* $\langle a, \bar{y} \rangle = 0$. *Since* $\mathcal{N} \neq [\bar{y}]^{\perp}$, *we can choose a sequence of vectors* $\{a^n\} \subset \mathcal{N}$ *such that* $\lim_{n\to\infty} a^n = a$ *and* $\langle a^n, \bar{y} \rangle \neq 0$. *Let* $\{u^1, u^2, \ldots, u^n\}$ *be an orthonormal system of eigenvectors of* B *corresponding to the eigenvalues* $\mu_1, \mu_2, \ldots, \mu_n$, *respectively. Note that the spectral decomposition of* B *implies* $B = \sum_{i=1}^{n} \mu_i u^i (u^i)^{\top}$. *Since* $\{a^n\} \subset \mathcal{N}$, *we can write* $a^n = \sum_{i=1}^{\ell} \alpha_{n,i} u^i$, *where* $2 \leq \ell = \dim(\mathcal{N}) < n$ *and* $\mu_1, \ldots, \mu_\ell$ *are the negative eigenvalues of* B. *Thus,*

$$\langle Ba^n, a^n \rangle = \sum_{i=1}^{\ell} \sum_{j=1}^{\ell} \langle B\alpha_{n,i} u^i, \alpha_{n,j} u^j \rangle = \sum_{i=1}^{\ell} \alpha_{n,i}^2 (u^i)^{\top} B u^i = \sum_{i=1}^{\ell} \alpha_{n,i}^2 \mu_i < 0.$$

For proceeding with the proof, we define

$$p^n := \bar{x} + t_n a^n, \qquad t_n := -2 \frac{\langle a^n, \bar{y} \rangle}{\langle Ba^n, a^n \rangle}.$$

Then, $\langle Bp^n, p^n \rangle = 0$ *and, due to* $\langle a, \bar{y} \rangle = 0$ *and* $\lim_{n\to\infty} a^n = a$, *we have* $\lim_{n\to\infty} p^n = \bar{x}$. *Hence, if* n *is sufficiently large, then for any* $\delta > 0$ *arbitrary but fixed, we have* $p^n \in \Xi(B, \bar{x}, \delta)$. *For such an* n, *after some simple algebraic manipulations we conclude*

$$\left\langle B\left(\frac{\bar{x} + p^n}{2}\right), \frac{\bar{x} + p^n}{2} \right\rangle = -\frac{\langle a^n, \bar{y} \rangle^2}{\langle Ba^n, a^n \rangle} > 0.$$

Hence, $\bar{x}, p^n \in \Xi(B, \bar{x}, \delta)$, *but* $(\bar{x} + p^n)/2 \notin \Xi(B, \bar{x}, \delta)$. *Therefore,* $\Xi(B, \bar{x}, \delta)$ *is not convex.* □

Proposition 5.4.1

Let $n \geq 3$ *and* $A = A^{\top} \in \mathbb{R}^{n\times n}$ *is a nonsingular matrix. Suppose that* q_A *is not constant and* $\lambda_1 \leq \lambda_2 \leq \ldots \leq \lambda_n$ *are eigenvalues of* A. *If* q_A *is quasi-convex, then the following conditions hold:*

(i) $\lambda_1 < \lambda_2$;

(ii) either $\lambda_2 \leq \min_{x\in\bar{\mathcal{C}}} q_A(x)$ *or* $\max_{x\in\bar{\mathcal{C}}} q_A(x) \leq \lambda_2$.

Proof:

Suppose by contradiction that one of the following two conditions holds:

(a) $\lambda_1 = \lambda_2$;

(b) $\min_{x\in\bar{\mathcal{C}}} q_A(x) < \lambda_2 < \max_{x\in\bar{\mathcal{C}}} q_A(x)$.

First of all, note that due to q_A *not being constant, we have* $\lambda_1 \leq \min_{x\in\bar{\mathcal{C}}} q_A(x) < \max_{x\in\bar{\mathcal{C}}} q_A(x) \leq \lambda_n$, *where* $\bar{\mathcal{C}}$ *is defined in* (5.3). *If the condition (a) holds, we can take a scalar* $\mu \in \mathbb{R}$ *such that* $\mu \neq \lambda_i$ *for any* $i \in \{1, \ldots, n\}$ *and satisfying*

$$\lambda_1 = \lambda_2 \leq \min_{x\in\bar{\mathcal{C}}} q_A(x) < \mu < \max_{x\in\bar{\mathcal{C}}} q_A(x) \leq \lambda_n. \tag{5.40}$$

Otherwise, if the condition (b) holds, we take $\mu \in \mathbb{R}$ *satisfying*

$$\lambda_1 \leq \min_{x\in\bar{\mathcal{C}}} q_A(x) < \lambda_2 < \mu < \max_{x\in\bar{\mathcal{C}}} q_A(x) \leq \lambda_n. \tag{5.41}$$

Then, either the conditions (5.40) *or* (5.41) *implies that* $\pm(A - \mu I_n)$ *is not* $\mathcal{K}$*-copositive. Since the matrix* $A - \mu I_n$ *is not* $\mathcal{K}$*-copositive, we can find a point* $p \in \mathcal{K}$ *such that* $\langle Ap, p\rangle < \mu\|p\|^2$. *Hence, we can find* $u \in \operatorname{int}(\mathcal{K})$ *sufficiently close to* p *such that* $\langle Au, u\rangle < \mu\|u\|^2$.

Similarly, since $-(A - \mu I_n) = \mu I_n - A$ *is not* $\mathcal{K}$*-copositive, we can find* $v \in \operatorname{int}(\mathcal{K})$ *such that* $\langle Av, v\rangle > \mu\|v\|^2$. *Therefore, we take* $t \in (0, 1)$ *and define*

$$\operatorname{int}(\mathcal{K}) \ni \bar{x} := (1 - t)u + tv.$$

By continuity, we have $\langle A\bar{x}, \bar{x}\rangle = \mu\|\bar{x}\|^2$. *Denoting* $B = A - \mu I$, *the eigenvalues of* B *are given by* $\mu_i := \lambda_i - \mu$, *for* $i \in \{1, 2, \ldots, n\}$. *Thus, we conclude from* (5.40) *and* (5.41) *that either*

$$\mu_1 = \mu_2 < 0 < \mu_n, \quad \textit{or} \qquad \mu_1 < \mu_2 < 0 < \mu_n, \tag{5.42}$$

if either the condition (a) or (b) holds, respectively. Considering that $B\bar{x} \neq 0$ *and* $\langle B\bar{x}, \bar{x}\rangle = 0$, *we conclude from Lemma 5.4.2 that, for any* $\delta > 0$, *the set*

$$\Xi(B, \bar{x}, \delta) := \{x \in \mathbb{R}^n : \|x - \bar{x}\| \leq \delta,\ \langle Bx, x\rangle \leq 0\},$$

is not convex. Hence, there exists an $s \in (0, 1)$ *and* $a^0, a^1 \in \Xi(B, \bar{x}, \delta)$ *such that* $a^s := (1 - s)a^0 + sa^1 \notin \Xi(B, \bar{x}, \delta)$. *Thus, since the closed ball centered at* $\bar{x}$ *and radius* δ *is convex,* $a^s \notin \Xi(B, \bar{x}, \delta)$ *implies* $\langle Aa^s, a^s\rangle - \mu\|a^s\|^2 = \langle Ba^s, a^s\rangle > 0$. *On the other hand,*

since $a^0, a^1 \in \Xi(B, \bar{x}, \delta)$, *we have* $\langle Aa^i, a^i\rangle - \mu\|a^i\|^2 = \langle Ba^i, a^i\rangle \leq 0$, *for* $i \in \{0,1\}$. *Furthermore, if* δ *is sufficiently small, then since* $\bar{x} \in \operatorname{int}(\mathcal{K})$, *we have* $a^0, a^1 \in \operatorname{int}\mathcal{K}$. *Hence,* $a^0, a^1 \in [\varphi_A \leq \mu]$ *and* $a^s \notin [\varphi_A \leq \mu]$. *By using Corollary 5.2.2, this contradicts the spherical quasi-convexity of A.* □

The proof of following lemma is based on Lemma 5.3.2.

Lemma 5.4.3

Let $A \in \mathbb{R}^{n\times n}$ *and* $\lambda, c \in \mathbb{R}$ *such that* $\lambda \leq c$. *If* $\lambda I_n - A$ *is* $\mathcal{K}$*-copositive, then* $[\varphi_A \leq c] = \operatorname{int}(\mathcal{K})$. *As a consequence, the set* $[\varphi_A \leq c]$ *is convex.*

Proof:

Let $c \in \mathbb{R}$ *and* $[\varphi_A \leq c] = \{x \in \operatorname{int}(\mathcal{K}) : \langle Ax, x\rangle - c\|x\|^2 \leq 0\}$. *Suppose that* $\lambda \leq c$, *for any* $x \in \operatorname{int}(\mathcal{K})$ *we have*

$$\langle Ax, x\rangle - c\|x\|^2 \leq \langle Ax, x\rangle - \lambda\|x\|^2 = \langle (A - \lambda I_n)x, x\rangle,$$

and considering that $\lambda I_n - A$ *is* $\mathcal{K}$*-copositive, that is*

$$\langle Ax, x\rangle - c\|x\|^2 \leq \langle (A - \lambda I_n)x, x\rangle \leq 0, \quad \forall x \in \operatorname{int}(\mathcal{K}),$$

hence $\langle Ax, x\rangle - c\|x\|^2 \leq 0$ *holds for any* $x \in \mathcal{K}$, *which implies that* $[\varphi_A \leq c] = \operatorname{int}(\mathcal{K})$. □

The following theorem combines the results from Proposition 5.4.1, Lemma 5.4.1, Lemma 5.4.3 and Corollary 5.2.2:

Theorem 5.4.1

Let $n \geq 3$, $k \geq 1$, $A = A^\top \in \mathbb{R}^{n\times n}$ *and* $\{v^1, v^2, \ldots, v^n\}$ *be an orthonormal system of eigenvectors of A corresponding to the eigenvalues* $\lambda_1 = \cdots = \lambda_j < \lambda_{k+1} \leq \ldots \leq \lambda_n$, *respectively. Then, we have the following statements:*

(i) If q_A *is quasi-convex and not constant, then* $k = 1$.

(ii) If q_A *is quasi-convex and not constant, then either* $\lambda_2 \leq \min_{x\in\bar{\mathcal{C}}} q_A(x)$ *or* $\max_{x\in\bar{\mathcal{C}}} q_A(x) \leq \lambda_2$.

(iii) Suppose that $k = 1$ *and* $\lambda_2 I_n - A$ *is* $\mathcal{K}$*-copositive. Then,* q_A *is spherically quasi-convex if and only if* $v^1 \in \mathcal{W}^* \cup -\mathcal{W}^*$. *In particular if* $v^1 \in \mathcal{K}^*$, *then* q_A *is spherically quasi-convex.*

Proof:

Items (i) and (ii) follow from Proposition 5.4.1. Item (iii) follows from Lemma 5.4.1, Lemma 5.4.3 and Corollary 5.2.2. □

The next corollary follows by combining Lemma 5.4.3 and Corollary 5.4.1.

Corollary 5.4.2

Let $n \geq 3$, $A = A^\top \in \mathbb{R}^{n\times n}$ and $\lambda_1 < \lambda_2 \leq \ldots \leq \lambda_n$ the eigenvalues of A. Suppose that $\lambda_2 \leq (\lambda_1 + \lambda_3)/2$ and $\lambda_2 I_n - A$ is $\mathcal{K}$-copositive. If either $\mathcal{K} \cap -\mathcal{L}_{\lambda_2} = \{0\}$ or $\mathcal{K} \cap \mathcal{L}_{\lambda_2} = \{0\}$, then q_A is spherically quasi-convex.

Proof:

Let $n \geq 3$, $A = A^\top \in \mathbb{R}^{n\times n}$ and $\lambda_1 < \lambda_2 \leq \ldots \leq \lambda_n$ the eigenvalues of A. If we have that $\lambda_2 \leq (\lambda_1 + \lambda_3)/2$ and suppose either $\mathcal{K} \cap -\mathcal{L}_{\lambda_2} = \{0\}$ or $\mathcal{K} \cap \mathcal{L}_{\lambda_2} = \{0\}$ by Corollary 5.4.1 we obtain that $[\varphi_A \leq c]$ is convex for any $c \notin (\lambda_2, \lambda_n)$.

Suppose that $\lambda_2 I_n - A$ is $\mathcal{K}$-copositive, by Lemma 5.4.3 we have $[\varphi_A \leq c] = \operatorname{int}(\mathcal{K})$ is convex for $c \geq \lambda_2$. Therefore, by Corollary 5.2.2 we conclude that q_A is spherically quasi-convex. □

In the following two theorems we present classes of quadratic quasi-convex functions defined in spherically subdual convex sets, which include as particular instances in Example 5.3.1 and 5.3.2.

Theorem 5.4.2

Let $n \geq 3$, $A = A^\top \in \mathbb{R}^{n\times n}$ and $\{v^1, v^2, \ldots, v^n\}$ be an orthonormal system of eigenvectors of A corresponding to the eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$, respectively. Assume that $\lambda := \lambda_1$, $\mu := \lambda_2 = \ldots = \lambda_{n-1}$, $\eta := \lambda_n$ and

$$v^1 - \sqrt{\frac{\eta - \mu}{\mu - \lambda}} |v^n|^{\mathcal{K}} \in \mathcal{K}^*, \qquad \lambda < \mu < \eta, \tag{5.43}$$

where $|\cdot|^{\mathcal{K}}$ is defined in (1.2). Then, the quadratic function q_A is spherically quasi-convex.

Proof:

By using the spectral decomposition of A, we have

$$A = \sum_{i=1}^{n} \lambda_i v^i (v^i)^\top = \lambda v^1 (v^1)^\top + \mu \sum_{j=2}^{n-1} v^j (v^j)^\top + \eta v^n (v^n)^\top. \tag{5.44}$$

Hence, for any $x \in \mathcal{K}$, *by using* $\|x\|^2 = \sum_{i=1}^{n} \langle v^i, x\rangle^2$ *and* (5.44), *we obtain*

$$\begin{aligned}\langle Ax, x\rangle - \mu\|x\|^2 &= \left(\lambda\langle v^1, x\rangle^2 + \mu\sum_{i=2}^{n-1}\langle v^i, x\rangle^2 + \eta\langle v^n, x\rangle^2\right) - \mu\sum_{i=1}^{n}\langle v^i, x\rangle^2 \\ &= (\mu-\lambda)\left[\frac{\eta-\mu}{\mu-\lambda}\langle v^n, x\rangle^2 - \langle v^1, x\rangle^2\right]. \end{aligned} \tag{5.45}$$

From (5.43) *we conclude that*

$$0 \leq \langle v^1 - \sqrt{\frac{\eta-\mu}{\mu-\lambda}}|v^n|^{\mathcal{K}}, x\rangle, \qquad \forall\ x \in \mathcal{K}. \tag{5.46}$$

To procced with the proof we note that (1.2) *implies that* $|v^n|^{\mathcal{K}} \in \mathcal{K} + \mathcal{K}^*$ *and, owing to* $\mathcal{K} \subseteq \mathcal{K}^*$, *we conclude that* $|v^n|^{\mathcal{K}} \in \mathcal{K}^*$. *Combing* (5.46) *we have*

$$0 \leq \sqrt{\frac{\eta-\mu}{\mu-\lambda}}\langle |v^n|^{\mathcal{K}}, x\rangle \leq \langle v^1, x\rangle, \qquad \forall\ x \in \mathcal{K}.$$

Hence, for any $x \in \mathcal{K}$, *the last inequality yields*

$$\begin{aligned}\frac{\eta-\mu}{\mu-\lambda}\langle v^n, x\rangle^2 - \langle v^1, x\rangle^2 &\leq \frac{\eta-\mu}{\mu-\lambda}\left[\langle v^n, x\rangle^2 - \langle |v^n|^{\mathcal{K}}, x\rangle^2\right] \\ &= \frac{\eta-\mu}{\mu-\lambda}\langle v^n + |v^n|^{\mathcal{K}}, x\rangle\langle v^n - |v^n|^{\mathcal{K}}, x\rangle. \end{aligned} \tag{5.47}$$

On the othet hand, by using $|v^n|^{\mathcal{K}} = \mathrm{P}_{\mathcal{K}}(v^n) + \mathrm{P}_{\mathcal{K}^*}(-v^n)$, $v^n = \mathrm{P}_{\mathcal{K}}(v^n) - \mathrm{P}_{\mathcal{K}^*}(-v^n)$, $P_{\mathcal{K}}(v^n) \in K \subseteq K^*$, *we obtain* $\langle v^n + |v^n|^{\mathcal{K}}, x\rangle\langle v^n - |v^n|^{\mathcal{K}}, x\rangle = -4\langle P_{\mathcal{K}}(v^n), x\rangle\langle P_{\mathcal{K}^*}(-v^n), x\rangle \leq 0$, *for any* $x \in \mathcal{K}$. *Thus, due to* $\lambda < \mu < \eta$, *the previous inequality together* (5.47) *implies*

$$\frac{\eta-\mu}{\mu-\lambda}\langle v^n, x\rangle^2 - \langle v^1, x\rangle^2 \leq 0, \qquad \forall\ x \in \mathcal{K}. \tag{5.48}$$

Thus, considering that $\lambda < \mu$, *the combination of* (5.45) *with* (5.48), *implies that*

$$\langle Ax, x\rangle - \mu\|x\|^2 \leq 0, \quad \forall x \in \mathcal{K},$$

which means $\mu\mathrm{I_n} - A$ *is* $\mathcal{K}$*-copositive. Taking into account that* $|v^n|^{\mathcal{K}} \in \mathcal{K}^*$, (5.43) *implies* $v^1 \in \mathcal{K}^*$. *Therefore, we can apply the item (iii) of Theorem 5.4.1 to conclude that* q_A *is spherically quasi-convex.* □

The following example satisfies the assumptions of Theorem 5.4.2.

Example 5.4.1

Letting $\mathcal{K} = \mathbb{R}^n_+$ *and* $\lambda < (\lambda+\eta)/2 < \mu < \eta$, *the unit vectors* $v^1 = (e^1 + e^n)/\sqrt{2}, v^2 = e^2, \ \ldots, \ v^{n-1} = e^{n-1}, v^n = (e^1 - e^n)/\sqrt{2}$ *are pairwise orthogonal and satisfy the condition* (5.43). *Now, taking* $\mathcal{K} = \mathcal{L}$ *and denoting* $v^n = ((v^n)_1, (v^n)^2)$, *by using Lemma 1.2.2, condition* (5.43) *can be written as*

$$v^1 - \sqrt{\frac{\eta-\mu}{\mu-\lambda}}\frac{1}{\|(v^n)^2\|}\Big(\max\big(|(v^n)_1|, \|(v^n)^2\|\big)\,\|(v^n)^2\|,\ \min\big(|(v^n)_1|, \|(v^n)^2\|\big)\,\mathrm{sgn}((v^n)_1)(v^n)^2\Big) \in \mathcal{K},$$

and $\lambda < \mu < \eta$. *The vectors* $v^1 = (e^1 + e^n)/\sqrt{2}, v^2 = e^2, \ \ldots, \ v^{n-1} = e^{n-1}, v^n = (-e^1 + e^n)/\sqrt{2}$ *are pairwise orthogonal and satisfy the last inclusion.*

Theorem 5.4.3

Let $n \geq 3$, $A = A^\top \in \mathbb{R}^{n\times n}$ *and* $\{v^1, v^2, \ldots, v^n\}$ *be an orthonormal system of eigenvectors of* A *corresponding to the eigenvalues* $\lambda_1, \lambda_2, \ldots, \lambda_n$, *respectively, such that* $v^1 \in \mathrm{int}(\mathcal{K}^*)$. *Let*

$$\alpha := \min\{\langle v^1, y\rangle^2 : \ y \in \mathbb{S}^n \cap \mathcal{K}\} > 0,$$

$$\eta := \max\left\{\frac{\sum_{i=3}^n \langle v^i, y\rangle^2}{\langle v^1, y\rangle^2} : \ y \in S^n \cap \mathcal{K}\right\} > 0. \tag{5.49}$$

Assume that

$$\lambda_1 < \lambda_2 \leq \cdots \leq \lambda_n \leq \lambda_2 + \delta(\lambda_2 - \lambda_1), \qquad \delta \in \{\alpha, 1/\eta\}. \tag{5.50}$$

Then, $\lambda_2 \mathrm{I_n} - A$ *is* $\mathcal{K}$*-copositive. Consequently, the quadratic function* q_A *is spherically quasi-convex.*

Proof:

Note that the spectral decomposition of A *implies* $A = \sum_{i=1}^n \lambda_i v^i (v^i)^\top$. *Thus, considering that* $\|x\|^2 = \sum_{i=1}^n \langle v^i, x\rangle^2$, *for any* $x \in \mathcal{K}$, *we conclude that*

$$\langle Ax, x\rangle - \lambda_2 \|x\|^2 = \sum_{i=1}^n (\lambda_i - \lambda_2)\langle v^i, x\rangle^2. \tag{5.51}$$

Since (5.50) *implies* $\lambda_2 - \lambda_1 > 0$ *and* $0 \leq \lambda_j - \lambda_2 \leq \lambda_n - \lambda_2$, *for any* $j \in \{3, \ldots, n\}$, *it*

follows from (5.51) *that*

$$\langle Ax, x\rangle - \lambda_2\|x\|^2 \leq (\lambda_2 - \lambda_1)\left[\frac{\lambda_n - \lambda_2}{\lambda_2 - \lambda_1}\sum_{i=3}^{n}\langle v^i, x\rangle^2 - \langle v^1, x\rangle^2\right]. \tag{5.52}$$

Since (5.49) *implies* $\sum_{i=3}^{n}\langle v^i, x\rangle^2 \leq \eta\langle v^1, x\rangle^2$, *the inequality* (5.52) *becomes*

$$\langle Ax, x\rangle - \lambda_2\|x\|^2 \leq (\lambda_2 - \lambda_1)\left[\left(\eta\frac{\lambda_n - \lambda_2}{\lambda_2 - \lambda_1} - 1\right)\langle v^1, x\rangle^2\right]. \tag{5.53}$$

First we suppose that $\delta = 1/\eta$. *Thus, the last inequality in* (5.50) *implies* $\eta(\lambda_n - \lambda_2)/(\lambda_2 - \lambda_1) \leq 1$, *which combined with* (5.53) *yields*

$$\langle Ax, x\rangle - \lambda_2\|x\|^2 \leq 0, \quad \forall x \in \mathcal{K}. \tag{5.54}$$

Next, suppose that $\delta = \alpha$. *First of all, noting that for any* $y \in S^n$ *we have* $\sum_{i=3}^{n}\langle v^i, y\rangle^2 \leq \sum_{i=1}^{n}\langle v^i, y\rangle^2 = \|y\|^2 = 1$. *Thus, using* (5.49), *we conclude that*

$$\eta = \max\left\{\frac{\sum_{i=3}^{n}\langle v^i, y\rangle^2}{\langle v^1, y\rangle^2} : \ y \in S^n \cap K\right\} \leq \max\left\{\frac{1}{\langle v^1, y\rangle^2} : y \in S^n \cap K\right\} = \frac{1}{\alpha}.$$

Hence, it follows from (5.53) *that*

$$\langle Ax, x\rangle - \lambda_2\|x\|^2 \leq (\lambda_2 - \lambda_1)\left[\left(\frac{1}{\alpha}\frac{\lambda_n - \lambda_2}{\lambda_2 - \lambda_1} - 1\right)\langle v^1, x\rangle^2\right]. \tag{5.55}$$

Due to $\delta = \alpha$, *the last inequality in* (5.50) *implies* $(\lambda_n - \lambda_2)/[\alpha(\lambda_2 - \lambda_1)] \leq 1$, *which together with* (5.55) *also implies* (5.54). *Hence, we conclude that* $\lambda_2 \mathrm{I_n} - A$ *is* $\mathcal{K}$*-copositive. Therefore, since* $v^1 \in \mathcal{K}^*$ *and it is an eigenvector of* A *corresponding to the eigenvalue* λ_1, *by applying item (iii) of Theorem 5.4.1, we can conclude that the function* q_A *is spherically quasi-convex.* □

In the following we present an example satisfying the assumptions of Theorem 5.4.3.

Example 5.4.2

Let $\mathcal{L}$ *be the Lorentz cone,* $v^i = e^i$, *for any* $i \in \{1, \ldots, n\}$, *and* $\lambda_1 < \lambda_2 \leq \ldots \leq \lambda_n < \lambda_2 + (1/2)(\lambda_2 - \lambda_1)$ *satisfy condition* (5.50). *Note that in this case* $\alpha = 1/2$.

Theorem 5.4.4

Let $n \geq 3$ and $A = A^\top \in \mathbb{R}^{n\times n}$. Suppose that A has only two distinct eigenvalues, and the smaller one has multiplicity one. If there exists an eigenvector of A corresponding to the smaller eigenvalue belonging to $\mathcal{K}^*$, then q_A is spherically quasi-convex.

Proof:

Let $\{v^1, v^2, \ldots, v^n\}$ be an orthonormal system of eigenvectors of A corresponding to the eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$, respectively. Without loss of generality, we assume that $\lambda_1 =: \lambda < \mu := \lambda_2 = \cdots = \lambda_n$ and $v^1 \in \mathcal{K}^*$. Thus, using the spectral decomposition of A, we have

$$A = \lambda v^1 (v^1)^\top + \sum_{j=2}^{n} \mu v^j (v^j)^\top. \tag{5.56}$$

Since $\|x\|^2 = \sum_{i=1}^n \langle v^i, x\rangle^2$, for any $x \in \mathbb{R}^n$, by using (5.56) and $\lambda < \mu$, we obtain that

$$\mu\|x\|^2 - \langle Ax, x\rangle = (\mu - \lambda)\langle v^1, x\rangle^2 \geq 0, \qquad \forall x \in \mathbb{R}^n. \tag{5.57}$$

In particular, (5.57) implies that $\mu I_n - A$ is $\mathcal{K}$-copositive. Thus, since $v^1 \in \mathcal{K}^*$, by applying item (iii) of Theorem 5.4.1 with $\lambda_2 = \mu$ we can conclude that the function q_A is spherically quasi-convex. □

In the next example we show how to generate matrices satisfying the assumptions of Theorem 5.4.4 and consequently generate spherically quasi-convex functions on spherically subdual convex sets.

Example 5.4.3

The Householder matrix associated to $v \in \operatorname{int}(\mathcal{K}^*)$ is defined by $H := \mathrm{I_n} - 2vv^T/\|v\|^2$. We know that H is a symmetric and nonsingular matrix. Furthermore, $Hv = -v$ and $Hu = u$ for any $u \in \mathcal{S}$, where $\mathcal{S} := \{u \in \mathbb{R}^n \ : \ \langle v, u\rangle = 0\}$. It is easy to verify that the dimension of $\mathcal{S}$ is $n-1$, then we have that 1 and -1 are eigenvalues of H with multiplicities $n-1$ and 1, respectively. Moreover, considering that $v \in \operatorname{int}(\mathcal{K}^*)$, Theorem 5.4.4 implies that $q_H(x) = \langle Hx, x\rangle$ is spherically quasi-convex.

5.5 Spherically Quasi-Convex Quadratic Functions on the Spherical Lorentz Convex Set

In this section we present a condition partially characterising the spherical quasi-convexity of quadratic functions on spherically convex sets associated to the second order cone (Lorentz cone). We remark that for the second order cone $\mathcal{L}$, since by Lemma 1.5.1, we have a characterisation of $\mathcal{L}$-copositive matrices. By using the idea of item (iii) of Theorem 5.4.1,we can provide a more general result than Theorem 5.4.3:

Theorem 5.5.1

Let $\mathcal{L}$ be the second order cone, $n \geq 2$, $A = A^\top \in \mathbb{R}^{n\times n}$, $\lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_n$ be the eigenvalues of A, v^1 be an eigenvector of A corresponding to λ_1 and $J = \operatorname{diag}(1, -1, \ldots, -1) \in \mathbb{R}^{n\times n}$. If $v^1 \in \mathcal{L}$ and there exists an $\rho \geq 0$ such that $\lambda_2 I_n - A - \rho J$ is positive semidefinite, then q_A is spherically quasi-convex.

Proof:

If there exists an $\rho \geq 0$ such that $\lambda_2 I_n - A - \rho J$ is positive semidefinite, then it follows from Lemma 1.5.1 that $\lambda_2 I_n - A$ is a $\mathcal{L}$-copositive matrix. Therefore, considering that $v^1 \in \mathcal{L} = \mathcal{L}^*$ and it is an eigenvector of A corresponding to the eigenvalue λ_1, by applying item (iii) of Theorem 5.4.1, we conclude that q_A is spherically quasi-convex. □

The next result is a version of Theorem 5.3.2 for the second order cone.

Theorem 5.5.2

Let $n \geq 3$ and $A = A^\top \in \mathbb{R}^{n\times n}$. Assume that A has only two distinct eigenvalues and the smaller one has multiplicity one. Then, q_A is a spherically quasi-convex function if and only if the second order cone $\mathcal{L}$ contains an eigenvector of A corresponding to the smaller eigenvalue.

Proof:

If there exists an eigenvector of A corresponding to the smaller eigenvalue belonging to $\mathcal{L}$, then Theorem 5.4.4 implies that q_A is spherically quasi-convex. Conversely, suppose that q_A is spherically quasi-convex. Let $\lambda_1, \lambda_2, \ldots, \lambda_n$ be the eigenvalues of A corresponding to an orthonormal system of eigenvectors $\{v^1, v^2, \ldots, v^n\}$, respectively. Then, without loss of generality, we assume that $\lambda_1 =: \lambda < \mu := \lambda_2 = \cdots = \lambda_n$. Thus, by using the spectral

decomposition of A, we have

$$A = \lambda v^1 (v^1)^\top + \sum_{j=2}^{n} \mu v^j (v^j)^\top. \tag{5.58}$$

We can also assume, without loss of generality, that $v_1^1 \geq 0$. Let $x \in \partial\mathcal{L} \setminus \{0\}$ and note that $y = 2x_1 e^1 - x \in \partial\mathcal{L} \setminus \{0\}$. Since $\sum_{i=1}^{n} v^i (v^i)^\top = I_n$ (i.e., the spectral decomposition of I_n) and $\langle x, y\rangle = 0$, (5.58) implies that

$$\langle Ax, y\rangle = \left\langle \left[\mu \sum_{i=1}^{n} v^i (v^i)^\top + (\lambda - \mu)(v^1)(v^1)^\top \right] x, y \right\rangle = (\lambda - \mu)\langle v^1, x\rangle \langle v^1, y\rangle. \tag{5.59}$$

Since $x, y \in \mathcal{L}$, $\langle x, y\rangle = 0$ and $\mathcal{L}$ is a self-dual cone, it follows from Corollary 5.2.1 that $\langle Ax, y\rangle \leq 0$. Thus, considering that $\lambda < \mu$ and $y = 2x_1 e^1 - x$, equation (5.59) yields

$$0 \leq \langle v^1, x\rangle \langle v^1, y\rangle = \langle v^1, x\rangle [2v_1^1 x_1 - \langle v^1, x\rangle]. \tag{5.60}$$

On the other hand, due to $x \in \mathcal{L}$, we have $x^1 \geq 0$. Thus, since $v_1^1 \geq 0$, if $\langle v^1, x\rangle < 0$, then $\langle v^1, x\rangle [(2v_1^1 x_1 - \langle v^1, x\rangle] < 0$, which contradicts (5.60). Hence $\langle v^1, x\rangle \geq 0$, where x can be chosen arbitrarily in $\partial\mathcal{L} \setminus \{0\}$. Therefore, we proved that $v^1 \in \mathcal{L}$. □

5.6 Conclusions and comments

In this chapter, we present several conditions characterising the spherical quasi-convexity of quadratic functions. In Section 5.2, we started from finding conditions characterising quadratic spherically quasi-convex function on a general spherically convex set. Then we move forward to the researches about the properties and the conditions implying spherical quasi-convexity of quadratic function defined on the spherical positive orthant. Minimising a quadratic function on the spherical nonnegative orthant is of particular interest because the nonnegativity of the minimum value is equivalent to the copositivity of the corresponding matrix [42, Proposition 1.3] and to the nonnegativity of its Pareto eigenvalues [42, Theorem 4.3].

In Section 5.4, we extended our results obtained in Section 5.2 to the general subdual convex sets. We studied the sufficient conditions for spherical quasi-convex functions on spherically

subdual convex sets. Also, we proved a condition for the partial characterisation of spherical quasi-convexity on spherical second order sets (Lorentz sets) in Section 5.5.

There are still many interesting questions to be answered in this topic:

1. First of all, we note that for all our classes of spherically quasi-convex quadratic functions q_A on the spherically subdual convex set $\mathcal{C} = \mathbb{S}^{n-1} \cap \text{int}(\mathcal{K})$, the matrix A has the smallest eigenvalue with multiplicity one and the associated eigenvector belongs to the dual $\mathcal{K}^*$ of the subdual cone $\mathcal{K}$. We conjecture that this condition is necessary and sufficient to characterise spherically quasi-convex quadratic functions.

2. We also remark that, in Theorem 5.5.2 we present a partial characterisations of spherically quasi-convex quadratic functions on the spherical Lorentz convex set. However, the general question remains open even for this specific set.

3. An even more challenging problem is to develop efficient algorithms for constrained quadratic optimisation problems on spherically convex sets. Minimising a quadratic function on the intersection of the second order cone with the sphere is a particularly relevant, related topic, since the nonnegativity of the minimum value is equivalent to the Lorentz-copositivity of the corresponding matrix, see [35, 65].

4. In general, replacing the second order cone with an arbitrary closed convex cone K leads to the more general concept of $\mathcal{K}$-copositivity. By considering the intrinsic geometrical properties of the sphere, interesting perspectives for detecting the general copositivity of matrices emerge.

CHAPTER 6

FINAL REMARKS

In this thesis, we studied the complementarity and it related problems. We now flash through the results of this thesis, summarise our contributions, and present possible future works.

6.1 Summary of results

We start this study from solving a linear complementarity problem on an extended second order cone. In Chapter 2, we convert an ESOCLCP to a MixCP on nonnegative orthant, therefore we can exploit complementarity function methods to solve the original problem. Then we provide two approaches: reformulate to a system of nonlinear equations; and, reformulate to an unconstrained minimisation problems. Based on these two approaches, algorithms are provided for solving the problem. It must be recognised that such conversion of a linear complementarity problem to a mixed complementarity problem may increase the complexity of the original problem. Finally, we furnish this chapter by providing a numerical example.

In Chapter 3, we studied the stochastic linear complementarity problems on extended second order cones. Similarly, we convert the stochastic ESOCLCP to a corresponding stochastic mixed complementarity problem on nonnegative orthant. We introduced the Conditional Value-at-Risk method to measure the loss of complementarity in the stochastic case. Unlike the ESOCLCP discussed in Chapter 2, we can only use the unconstrained minimisation approach to solve this problem. We also introduce an algorithm and provide a numerical example for this problem.

In Chapter 4, we introduced the mean-Euclidean norm (MEN) model for portfolio optimi-

sation problem. This new model is based on the mean-absolute deviation (MAD) model. The KKT condition of the MEN model is a nonlinear complementarity problem on extended second order cone. Based on the results obtained in Chapter 1 and Chapter 2, we find an analytical solution to the MEN model.

Chapter 5 presents several conditions characterising the spherical quasi-convexity of quadratic functions. We started from finding conditions characterising quadratic spherically quasi-convex function on a general spherically convex set. Then we move forward to the researches about the properties and the conditions implying spherical quasi-convexity of quadratic function defined on the spherical positive orthant. Further, we extended our results on to the general subdual convex sets. The sufficient conditions for spherical quasi-convex functions on spherically subdual convex sets are presented. In addition, a condition for the partial characterisation of spherical quasi-convexity on spherical Lorentz sets is provided.

6.2 Contributions

We now outline our main contributions:

1. The study about the linear complementarity problem on extended order cone. We successfully find an solution to this problem. Algorithms and numerical examples are provided. These results are published on our paper [82].

2. The study about the stochastic linear complementarity problem on extended order cone. The CVaR-based minimisation problem is used on the merit function to measure the loss of complementarity in the stochastic case. Since the objective function of the CVaR-based minimisation problem is convex in some neighbourhoods (it is not globally convex), given an initial point that close enough to the optimal point, we can still solve this problem by an algorithm. Hence, we introduced an algorithm for solving this problem and provide a numerical example.

3. The introduction of the mean-Euclidean norm (MEN) model. Since the KKT condition of this model is a nonlinear complementarity problem on extended second order cone, we innovatively used the results in Chapter 1 and Chapter 2 to find the analytical solution to this model.

4. Chapter 5 is a continuation of [24, 25, 27], where intrinsic properties of the spherically quasi-convex sets and functions were studied. As far as we know this is the pioneering study of spherically quasi-convex quadratic functions on spherically subdual convex sets. The results about the conditions characterising spherical quasi-convexity of quadratic function defined on the spherical positive orthant are published on [28].

ACRONYMS

Notation	Description	Page List
a.s.	almost surely	52
ALoC	average loss of complementarity	70
C-function	complementarity function	9
CHKS	Chen-Harker-Kanzow-Smale	66
CM	CVaR minimisation	54
CP	complementarity problem	9
CVaR	Conditional Value at Risk	3
ERM	Expected residual minimisation	53
ESOC	extended second order cone	2
ESOCLCP	linear complementarity problems on extended second order cones	26
EV	Expected value	53
FB	Fischer-Burmeister	31
ICP	implicit complementarity problem	10
KKT	Karush-Kuhn-Tucker	73
l.i.	linearly independent	116

Notation	Description	Page List
LCP	linear complementarity problem	9
LM	Levenberg-Marquardt	37
MEN	mean-Euclidean norm model	79
MixCP	mixed complementarity problem	10
MixICP	mixed implicit complementarity problem	10
MV	mean variance model	73
NCP	nonlinear complementarity problem	90
S-ESOCLCP	the stochastic linear complementarity problem on extended second order cones	51
SAA	Sample Average Approximation	66
SCP	stochastic complementarity problem	52
SLCP	stochastic linear complementarity problem	51
SMixCP	stochastic mixed complementarity problem	57
SMPEC	Stochastic mathematical programs with equilibrium constraints	53
SP	Stochastic programming	54
SSD	second degree stochastic dominance	77
VaR	Value at risk	54
VI	variational inequality	9

NOMENCLATURE

Functions

$[t]_+$	$:= \max\{0, t\}$ the nonnegative operator
$\mathbb{E}[\cdot]$	the expected value of the random vector
γ	$: [x, y] \to \mathbb{S}^{n-1}$ the geodesic
$\gamma_{xy}(t)$	$: [0, 1] \to \mathbb{S}^{n-1}$ the minimal geodesic from x to y
$\operatorname{grad} f(x)$	the gradient on the sphere of a differentiable function $f : \Omega \to \mathbb{R}$ at a point $x \in \Omega$
$\mathbb{F}_{FB}^{\mathrm{MixCP}}(x, u, t)$	FB-based C-function of MixCP problem
$\mathbf{1}_S(x)$	$: S \subseteq \mathbb{R}^n \to \{0, 1\}$ the indicator function with $\mathbf{1}_S(x) = 1$ if $x \in S$ and $\mathbf{1}_S(x) = 0$ otherwise
$\mathcal{I}(F)$	$: \mathcal{K} \to \mathbb{R}^n$ the inversion (of pole 0) of the mapping $F : \mathbb{R}^n \to \mathbb{R}^n$
$\mathcal{P}$	$: \mathcal{F} \to [0, 1]$, a function maps from events to probabilities.
$\partial\mathbb{F}_{FB}^{\mathrm{MixCP}}(x, u, t)$	the generalised Jacobian set of $\mathbb{F}_{FB}^{\mathrm{MixCP}}(x, u, t)$
Φ	$: \mathbb{R}^n \times \Omega \to \mathbb{R}^n$ the multi dimensional C-function
ϕ	$: \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ a scalar C-function such that $\phi(a, b) = 0 \;\Leftrightarrow\; a \geq 0,\; b \geq 0,\; ab = 0$
$\psi_{FB}(a, b)$	scalar form of FB C-function
$\theta_{FB}^{\mathrm{MixCP}}(x, u, t)$	the associated merit function based on FB C-function

$\underline{F}^{\#}(x_0, \mathcal{K})$	the lower scalar derivative of the mapping $F : \mathcal{K} \to \mathbb{R}^n$ at x_0
φ_A	: int($\mathcal{K}$) $\to \mathbb{R}$ the Rayleigh quotient function restricted on cone int($\mathcal{K}$) defined by matrix A
$Df(x)$	the gradient a differentiable function $f : \Omega \to \mathbb{R}$ at a point x
$F(x)$	the linear function $F(x) = Tx + r : \mathbb{R}^n \to \mathbb{R}^n$
$f \circ g$	the composition of function f and g
$J_x F(x, y)$	$:= \left(\frac{\partial F_i}{\partial x_j}\right)$ the $m \times n$ partial Jacobian of a mapping $F : \mathbb{R}^k \times \mathbb{R}^\ell \to \mathbb{R}^m$
JF	$\equiv \frac{\partial F_i}{\partial x_j}$ the $m \times n$ Jacobian of a mapping $F : \mathbb{R}^n \to \mathbb{R}^m$
$p(\cdot, \mu)$	a smoothing function of the nonnegative operator
q_A	: $\mathcal{S} \to \mathbb{R}$ the associated quadratic function defined by the symmetric matrix $A = A^T \in \mathbb{R}^{n \times n}$
$\nabla\theta$	$:= \left(\frac{\partial \theta}{\partial x_j}\right)$ the gradient of a function $\theta : \mathbb{R}^n \to \mathbb{R}$

Matrices

$\mathcal{D}_a$	denoted by $\text{diag}\big(a_1(x,u,t),\ \ldots,\ a_k(x,u,t)\big)$ the $k \times k$ diagonal matrices
$\mathcal{D}_b$	denoted by $\text{diag}\big(b_1(x,u,t),\ \ldots,\ b_k(x,u,t)\big)$ the $k \times k$ diagonal matrices
$\lambda_{max}(A)$	the largest eigenvalue of a matrix A
$\mathcal{A}$	$\partial \mathbb{F}_{FB}^{\text{MixCP}}(x,u,t)$ an element in the generalised Jacobian set of $\mathbb{F}_{FB}^{\text{MixCP}}(x,u,t)$
A	$:= (a_{ij})$ a matrix with entries a_{ij}
$A^\top$	the transpose of a matrix A
A^{-1}	the inverse of a matrix A

I_n	the $n \times n$ identity matrix
MA	the Schur complement of A in M
S_0	the S_0 matrix
$T = \left(\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right)$	the T matrix with component matrices A, B, C, and D

Problem Classes and Fundamental Objects

$\mathrm{CP}(F, \mathcal{K})$	the complementarity problem defined by function F and cone $\mathcal{K}$
$\mathrm{ICP}(G, F, \mathcal{K})$	the implicit complementarity problem defined by G, F, and $\mathcal{K}$
$\mathrm{MixCP}(F_1, F_2, \mathcal{S})$	the mixed complementarity problem defined by function F_1, F_2, and set $\mathcal{S}$
$\mathrm{SOL\text{-}CP}(F, \mathcal{K})$	the solution set of $\mathrm{CP}(F, \mathcal{K})$
$\mathrm{SOL\text{-}ICP}(G, F, \mathcal{K})$	the solution set of $\mathrm{ICP}(G, F, \mathcal{K})$
$\mathrm{SOL\text{-}LCP}(T, r, \mathcal{K})$	the solution set of $\mathrm{LCP}(T, r, \mathcal{K})$
$\mathrm{SOL\text{-}MixCP}(F_1, F_2, \mathcal{S})$	the solution set of $\mathrm{MixCP}(F_1, F_2, \mathcal{S})$
$\mathrm{SOL\text{-}MixICP}(F_1, F_2, G_1, \mathcal{S})$	the solution set of the mixed complementarity problem MixICP $(F_1, F_2, G_1, \mathcal{S})$
$MixICP(F_1, F_2, G_1, \mathcal{S})$	the mixed implicit complementarity problem defined by function F_1, F_2, and G_1, as well as the set $\mathcal{S}$

Spaces

$\mathbb{V}^{\perp}$	the orthogonal complement of a subspace $\mathbb{V}$ in $\mathbb{R}^n$
$\mathbb{R}^n$	the n dimensional Euclidian space
$\mathbb{R}^n_+$	the nonnegative orthant
$\mathbb{R}^n_{++}$	the positive orthant

$T_x\mathbb{S}^{n-1}$	$:= \{v \in \mathbb{R}^n : \ \langle x, v\rangle = 0, \ x \in \mathbb{S}^{n-1}\}$ the tangent hyperplane at point $x \in \mathbb{S}^{n-1}$
$[f \leq c]$	the sub-level set of a function $f : \ \mathcal{S} \to \mathbb{R}$
$\arg\min_a F(x)$	the set of constrained minimisers of F on a
$\bar{\mathbf{B}}(x, \delta)$	a spherically closed ball center in point x
$\mathcal{C}(\mathcal{K})$	the complementarity set of $\mathcal{K}$
$\cup, \ \cap$	set union, intersection
$\emptyset$	the empty set
$\in, \ \notin$	element membership, non-membership in a set
$\operatorname{int}\mathcal{S}$	the interior of a set $\mathcal{S}$
$\mathcal{C}$	the complementarity index set
$\mathcal{F}$	$\subseteq 2^{\Omega}$, a σ-algebra generated by Ω (or a collection of all subsets of Ω)
$\mathcal{N}$	the negative index set
$\mathcal{P}$	the positive index set
$\mathcal{R}$	the residual index set
Ω	the sample set of possible outcomes
$\subseteq, \ \subset$	subset inclusion, subset
$\mathbf{B}(x, \delta)$	a spherically open ball center in point x
$L(k, \ell)$, L	the extended second order cone
$M(k, \ell)$, M	the dual cone of extended second order cone
$S_1 \setminus S_2$	set S_1 minus set S_2
$\mathcal{K}$	a cone, a convex cone, or a convex set

$\mathcal{K}^*$	the dual cone of $\mathcal{K}$
$\mathcal{K}_{\mathcal{S}}$	a cone spanned by $\mathcal{S}$
$\mathcal{L}_c$	$:= \left\{x \in \mathbb{R}^n : \ \langle v^1, x\rangle \geq \sqrt{\sum_{i=2}^n \theta_i(c)\langle v^i, x\rangle^2}\right\}$ the convex cone defined by c
$\mathcal{N}_{\overline{x}}$	the neighbourhood of $\overline{x}$

Vectors

$\binom{x}{u}$	the pairs of vectors of x and u
$\langle \cdot, \cdot \rangle$	the inner product of two vectors
ω	$\in \Omega$ an n-dimensional random vector
$x \circ y$	$:= (x_i y_i)_{i=1}^n$ the Hadamard product of x and y
$x \perp y$	x and y are perpendicular

INDEX

σ-algebra, 52

Canonical inner product, 4

Complementarity

- Complementarity function, 9
- Complementarity set, 8
- Variational inequalities, 9

Complementarity problem, 9

- Implicit complementarity problem, 10
- Linear complementarity problem, 9
- Mixed complementarity problem, 10
- Mixed implicit complementarity problem, 11
- Nonlinear complementarity problem, 90
- Stochastic complementarity problem, 52
- Stochastic linear complementarity problem, 52

Conditional Value at risk, CVaR, 55

Cone

- Closed cone, 5
- Closed convex cone, 5
- Convex cone, 5
- Dual cone, 5
- Pointed cone, 5
- Proper cone, 5
- Second order cone, Lorentz cone, 5
- Self-dual cone, 5
- Spherically convex set, 20
- Subdual cone, 5
- Superdual cone, 5

Continuously differentiable function, 8

Copositive, 16

- $\mathcal{K}$-Copositive, 16

Euclidean sphere, 18

Extended second order cone, 12

FB regular, 40

Fréchet differentiable, 7

Geodesic, 19

Gradient on the sphere, 20

Homogeneity, 55

Householder matrix, 107, 127

Identity matrix, 5, 16

Intrinsic distance on the sphere, 18

Inversion, 91

Inversion of a mapping, 91

Lipschitz function, 8

Lower scalar derivative, 91

Mean-absolute deviation model, MAD model, 78
Mean-Euclidean norm model, MEN model, 79
Mean-variance model, MV model, 75, 76
merit function, 38
Minimal geodesic, 19
Monotonicity, 55
Moreau's decomposition, 6

Nonnegative orthant, 4
Positive orthant, 4
Norm, 4

Probability space, 52
Projection, projection mapping, 6

Quadratic function, 94

Rayleigh quotient function, 94

Sample set of possible outcomes, 52
Schur complement, 7
Second order cone, Lorentz cone, 16
Semismooth function, 8
Signed S_0 matrix, 43
Spherically convex
Merely spherically quasi-convex, 110
Spherically convex function, 21
Spherically quasi-convex function, 21
Sub-additivity, 55

Tangent hyperplane, 18

Value at risk, VaR, 54

LIST OF REFERENCES

[1] C. Acerbi and D. Tasche. On the coherence of expected shortfall. *Journal of Banking and Finance*, 26(7):1487–1503, 2002.

[2] F. Alizadeh and D. Goldfarb. Second-order cone programming. *Mathematical Programming*, 95(1):3–51, 2003.

[3] R. Andreani, A. Friedlander, M. P. Mello, and S. A. Santos. Box-constrained minimization reformulations of complementarity problems in second-order cones. *Journal of Global Optimization*, 40(4):505–527, 2008.

[4] Fischer Andreas. On the local superlinear convergence of a newton-type method for lcp under weak conditions. *Optimization Methods and Software*, 6(2):83–107, 1995.

[5] P. Artzner, F. Delbaen, J. M. Eber, and D. Heath. Coherent measures of risk. *Mathematical Finance*, 9(3):203–228, 1999.

[6] A. Ben-Tal, L. E. Ghaoui, and A. Nemirovski. *Robust optimization*. Princeton University Press, 2009.

[7] A. Ben-Tal and A. Nemirovski. Robust optimization–methodology and applications. *Mathematical Programming*, 92(3):453–480, 2002.

[8] J. M. Borwein and M. A. Dempster. The linear order complementarity problem. *Mathematics of Operations Research*, 14(3):534–558, 1989.

[9] B. Bower and P. Wentz. Portfolio optimization: Mad vs. markowitz. *Rose-Hulman Undergraduate Mathematics Journal*, 6(2):3, 2005.

[10] A. J. Cairns, D. Blake, and K. Dowd. *Optimal dynamic asset allocation for defined-contribution pension plans*. Pensions Institute, 2000.

[11] B. Chen and P. T. Harker. Smooth approximations to nonlinear complementarity problems. *SIAM Journal on Optimization*, 7(2):403–420, 1997.

[12] Chunhui Chen and Olvi L Mangasarian. A class of smoothing functions for nonlinear and mixed complementarity problems. *Computational Optimization and Applications*, 5(2):97–138, 1996.

[13] Jein-Shan Chen and Shaohua Pan. A regularization semismooth newton method based on the generalized fischer–burmeister function for p0-ncps. *Journal of Computational and Applied Mathematics*, 220(1-2):464–479, 2008.

[14] X. Chen and M. Fukushima. Expected residual minimization method for stochastic linear complementarity problems. *Mathematics of Operations Research*, 30(4):1022–1038, 2005.

[15] X. Chen and G. H. Lin. CVaR-based formulation and approximation method for stochastic variational inequalities. *Numerical Algebra, Control & Optimization*, 1(1):35–48, 2011.

[16] R. W. Cottle and G. B. Dantzig. Complementary pivot theory of mathematical programming. *Linear Algebra and its Applications*, 1(1):103–125, 1968.

[17] J. Danielsson, B. N. Jorgensen, S. Mandira, G. Samorodnitsky, and C. G. De Vries. Subadditivity re-examined: the case for value-at-risk. Technical report, Cornell University Operations Research and Industrial Engineering, 2005.

[18] G. B. Dantzig and R. W. Cottle. Positive (semi-) definite matrices and mathematical programming. Technical report, California Univ Berkeley Operations Research Center, 1963.

[19] F. Facchinei and J. S. Pang. *Finite-dimensional variational inequalities and complementarity problems. Vol. II.* Springer-Verlag, New York, 2003.

[20] Francisco Facchinei, Andreas Fischer, and Christian Kanzow. Inexact newton methods for semismooth equations with applications to variational inequality problems. In *Nonlinear Optimization and Applications*, pages 125–139. Springer, 1996.

[21] Francisco Facchinei and Christian Kanzow. A nonsmooth inexact newton method for the solution of large-scale nonlinear complementarity problems. *Mathematical Programming*, 76(3):493–512, 1997.

[22] H. Fang, X. Chen, and M. Fukushima. Stochastic r_0 matrix linear complementarity problems. *SIAM Journal on Optimization*, 18(2):482–506, 2007.

[23] J. A. Ferland. Maximal domains of quasi-convexity and pseudo-convexity for quadratic functions. *Mathematical Programming*, 3(1):178–192, 1972.

[24] O. P. Ferreira, A. N. Iusem, and S. Z. Németh. Projections onto convex sets on the sphere. *Journal of Global Optimization*, 57(3):663–676, 2013.

[25] O. P. Ferreira, A. N. Iusem, and S. Z. Németh. Concepts and techniques of optimization on the sphere. *TOP*, 22(3):1148–1170, 2014.

[26] O. P. Ferreira and S. Z. Németh. How to project onto extended second order cones. *Journal of Global Optimization*, 70(4):707–718, 2018.

[27] O. P. Ferreira and S. Z. Németh. On the spherical convexity of quadratic functions. *Journal of Global Optimization*, 73(3):537–545, 2019.

[28] O. P. Ferreira, S. Z. Németh, and L. Xiao. On the spherical quasi-convexity of quadratic functions. *Linear Algebra and its Applications*, 562:205–222, 2019.

[29] O. P. Ferreira, S. Z. Németh, and L. Xiao. On the spherical quasi-convexity of quadratic functions on spherically subdual convex sets. *arXiv preprint arXiv:1905. 06891*, 2019.

[30] A. Fischer. A special Newton-type optimization method. *Optimization*, 24(3-4):269–284, 1992.

[31] A. Fischer. A Newton-type method for positive-semidefinite linear complementarity problems. *Journal of Optimization Theory and Applications*, 86(3):585–608, 1995.

[32] Rüdiger Frey and Alexander J McNeil. Var and expected shortfall in portfolios of dependent credit risks: conceptual and practical insights. *Journal of banking & finance*, 26(7):1317–1334, 2002.

[33] M. Fukushima, Z. Q. Luo, and P. Tseng. Smoothing functions for second-order-cone complementarity problems. *SIAM Journal on Optimization*, 12(2):436–460, 2002.

[34] Masao Fukushima. Equivalent differentiable optimization problems and descent methods for asymmetric variational inequality problems. *Mathematical programming*, 53(1-3):99–110, 1992.

[35] P. Gajardo and A. Seeger. Solving inverse cone-constrained eigenvalue problems. *Numerische Mathematik*, 123(2):309–331, 2013.

[36] C. B. Garcia. Some classes of matrices in linear complementarity theory. *Mathematical Programming*, 5(1):299–310, 1973.

[37] G. Gürkan, A. Y. Özge, and S. M. Robinson. Sample-path solution of stochastic variational inequalities. *Mathematical Programming*, 84(2):313–333, 1999.

[38] W. W. Hager. Minimizing a quadratic over a sphere. *SIAM Journal on Optimization*, 12(1):188–208, 2001.

[39] W. W. Hager and S. Park. Global convergence of SSM for minimizing a quadratic over a sphere. *Mathematics of Computation*, 74(251):1413–1423, 2005.

[40] Patrick T Harker and Jong-Shi Pang. Finite-dimensional variational inequality and nonlinear complementarity problems: a survey of theory, algorithms and applications. *Mathematical programming*, 48(1-3):161–220, 1990.

[41] J. B. Hiriart-Urruty and C. Lemaréchal. *Convex analysis and minimization algorithms: Fundamentals. I*, volume 305 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, 1993.

[42] J. B. Hiriart-Urruty and A. Seeger. A variational approach to copositive matrices. *SIAM Review*, 52(4):593–629, 2010.

[43] R. A. Horn. The hadamard product. In *Proc. Symp. Appl. Math*, volume 40, pages 87–169, 1990.

[44] R. A. Horn and C. R. Johnson. *Matrix analysis*. Cambridge University Press, Cambridge, 1985.

[45] G. Isac and S. Z. Németh. Duality in nonlinear complementarity theory by using inversions and scalar derivatives. *Mathematical Inequalities and Applications*, 9(4):781–795, 2006.

[46] G. Isac and S. Z. Németh. *Scalar and asymptotic scalar derivatives: theory and applications*, volume 13. Springer Science & Business Media, 2008.

[47] P. Jaillet, D. Lamberton, and B. Lapeyre. Variational inequalities and the pricing of American options. *Acta Applicandae Mathematicae*, 21(3):263–289, 1990.

[48] H. Jiang and H. Xu. Stochastic approximation approaches to the stochastic variational inequality problem. *IEEE Transactions on Automatic Control*, 53(6):1462–1475, 2008.

[49] J. G. Kallberg and W. T. Ziemba. Mis-specifications in portfolio selection problems. In *Risk and Capital*, pages 74–87. Springer, 1984.

[50] C. Kanzow. An unconstrained optimization technique for large-scale linearly constrained convex minimization problems. *Computing*, 53(2):101–117, 1994.

[51] Christian Kanzow. Nonlinear complementarity as unconstrained optimization. *Journal of optimization Theory and Applications*, 88(1):139–155, 1996.

[52] S. Karamardian, S. Schaible, and J. P. Crouzeix. Characterizations of generalized monotone maps. *Journal of Optimization Theory and Applications*, 76(3):399–413, 1993.

[53] W. Karush. Minima of functions of several variables with inequalities as side conditions. pages 217–245, 2014.

[54] S. Komlósi. Generalized convexity of a certain class of quadratic functions. *Izvestiya Vysshikh Uchebnykh Zavedenii. Matematika*, (9):38–43, 1984.

[55] D. Kong, L. Liu, and Y. Wu. Isotonicity of the metric projection by lorentz cone and variational inequalities. *Journal of Optimization Theory and Applications*, 173(1):117–130, 2017.

[56] H. Konno and T. Koshizuka. Mean-absolute deviation model. *Iie Transactions*, 37(10):893–900, 2005.

[57] H. Konno and A. Wijayanayake. Mean-absolute deviation portfolio optimization model under transaction costs. *Journal of the Operations Research Society of Japan*, 42(4):422–435, 1999.

[58] H. Konno and H. Yamazaki. Mean-absolute deviation portfolio optimization model and its applications to tokyo stock market. *Management Science*, 37(5):519–531, 1991.

[59] I. Konnov. *Equilibrium models and variational inequalities*, volume 210. Elsevier, 2007.

[60] L. Laloux, P. Cizeau, J. P. Bouchaud, and M. Potters. Noise dressing of financial correlation matrices. *Physical Review Letters*, 83(7):1467, 1999.

[61] L. Laloux, P. Cizeau, M. Potters, and J. P. Bouchaud. Random matrix theory and financial correlations. *International Journal of Theoretical and Applied Finance*, 3(03):391–397, 2000.

[62] H. Levy. Stochastic dominance and expected utility: survey and analysis. *Management Science*, 38(4):555–593, 1992.

[63] G. H. Lin. Combined Monte Carlo sampling and penalty method for stochastic nonlinear complementarity problems. *Mathematics of Computation*, 78(267):1671–1686, 2009.

[64] G. H. Lin, X. Chen, and M. Fukushima. Solving stochastic mathematical programs with equilibrium constraints via approximation and smoothing implicit programming with penalization. *Mathematical Programming*, 116(1-2):343–368, 2009.

[65] R. Loewy and H. Schneider. Positive operators on the n-dimensional ice cream cone. *Journal of Mathematical Analysis and Applications*, 49(2):375–392, 1975.

[66] G. M. Luo, X. An, and J. Y. Xia. Robust optimization with applications to game theory. *Applicable Analysis*, 88(8):1183–1195, 2009.

[67] Changfeng Ma, Jia Tang, and Xiaohong Chen. A globally convergent levenberg–marquardt method for solving nonlinear complementarity problem. *Applied mathematics and computation*, 192(2):370–381, 2007.

[68] H. Ma and N. Huang. CVaR–based formulation and approximation method for a class of stochastic variational inequality problems. *Mathematical Inequalities & Applications*, 16(4):981–998, 2013.

[69] J. Malick. The spherical constraint in boolean quadratic programs. *Journal of Global Optimization*, 39(4):609–622, 2007.

[70] O. L. Mangasarian. Linear complementarity problems solvable by a single linear program. *Mathematical Programming*, 10(1):263–270, 1976.

[71] Olvi L Mangasarian. Equivalence of the complementarity problem to a system of nonlinear equations. *SIAM Journal on Applied Mathematics*, 31(1):89–92, 1976.

[72] H. Markowitz. Portfolio selection. *The Journal of Finance*, 7(1):77–91, 1952.

[73] D. W. Marquardt. An algorithm for least-squares estimation of nonlinear parameters. *Journal of the society for Industrial and Applied Mathematics*, 11(2):431–441, 1963.

[74] B. Martos. Subdefinite matrices and quadratic forms. *SIAM Journal on Applied Mathematics*, 17:1215–1223, 1969.

[75] S. Mataramvura and B. Øksendal. Risk minimizing portfolios and HJBI equations for stochastic differential games. *Stochastics An International Journal of Probability and Stochastic Processes*, 80(4):317–337, 2008.

[76] A. Meucci. *Risk and asset allocation*. Springer Science and Business Media, 2009.

[77] R. O. Michaud and R. O. Michaud. *Efficient asset management: a practical guide to stock portfolio optimization and asset allocation*. Oxford University Press, 2008.

[78] R. Mifflin. Semismooth and semiconvex functions in constrained optimization. *SIAM Journal on Control and Optimization*, 15(6):959–972, 1977.

[79] J. J. Moreau. Décomposition orthogonale d'un espace hilbertien selon deux cônes mutuellement polaires. *C. R. Acad. Sci. Paris*, 255:238–240, 1962.

[80] S. Z. Németh. Five kinds of monotone vector fields. *Pure Mathematics and Applications*, 9(3-4):417–428, 1998.

[81] S. Z. Németh. Scalar derivatives in Hilbert spaces. *Positivity*, 10(2):299–314, 2006.

[82] S. Z. Németh and L. Xiao. Linear complementarity problems on extended second order cones. *Journal of Optimization Theory and Applications*, 176(2):269–288, 2018.

[83] S. Z. Németh and G. Zhang. Extended Lorentz cones and mixed complementarity problems. *Journal of Global Optimization*, 62(3):443–457, 2015.

[84] S. Z. Németh and G. Zhang. Extended Lorentz cones and variational inequalities on cylinders. *Journal of Optimization Theory and Applications*, 168(3):756–768, 2016.

[85] R. Nishimura, S. Hayashi, and M. Fukushima. Robust Nash equilibria in N-person non-cooperative games: uniqueness and reformulation. *Pacific Journal of Optimization*, 5(2):237–259, 2009.

[86] S. Pafka and I. Kondor. Noisy covariance matrices and portfolio optimization II. *Physica A: Statistical Mechanics and its Applications*, 319:487–494, 2003.

[87] I. Pólik and T. Terlaky. A survey of the s-lemma. *SIAM review*, 49(3):371–418, 2007.

[88] Liqun Qi. Convergence analysis of some algorithms for solving nonsmooth equations. *Mathematics of operations research*, 18(1):227–244, 1993.

[89] Liqun Qi and Jie Sun. A nonsmooth version of newton's method. *Mathematical programming*, 58(1-3):353–367, 1993.

[90] R. T. Rockafellar and S. Uryasev. Conditional value-at-risk for general loss distributions. *Journal of Banking and Finance*, 26(7):1443–1471, 2002.

[91] R. T. Rockafellar, S. Uryasev, et al. Optimization of conditional value-at-risk. *Journal of Risk*, 2:21–42, 2000.

[92] A. D. Roy. Safety first and the holding of assets. *Econometrica: Journal of the Econometric Society*, pages 431–449, 1952.

[93] M. Rudolf, H. Wolter, and H. Zimmermann. A linear model for tracking error minimization. *Journal of Banking and Finance*, 23(1):85–103, 1999.

[94] S. Schaible. Quasiconvex, pseudoconvex, and strictly pseudoconvex quadratic functions. *Journal of Optimization Theory and Applications*, 35(3):303–338, 1981.

[95] S. T. Smith. Optimization techniques on Riemannian manifolds. In *Hamiltonian and gradient flows, algorithms and control*, volume 3 of *Fields Inst. Commun.*, pages 113–136. Amer. Math. Soc., Providence, RI, 1994.

[96] A. M. C. So. Deterministic approximation algorithms for sphere constrained homogeneous polynomial optimization problems. *Math. Program.*, 129(2, Ser. B):357–382, 2011.

[97] H. H. Sohrab. *Basic real analysis*, volume 231. Springer, 2003.

[98] Defeng Sun. A regularization newton method for solving nonlinear complementarity problems. *Applied Mathematics and Optimization*, 40(3):315–339, 1999.

[99] R. Sznajder. The Lyapunov rank of extended second order cones. *Journal of Global Optimization*, 66(3):585–593, 2016.

[100] M. Wang and M. M. Ali. Stochastic nonlinear complementarity problems: stochastic programming reformulation and penalty-based approximation method. *Journal of Optimization Theory and Applications*, 144(3):597–614, 2010.

[101] G. A. Whitmore and M. C. Findlay. *Stochastic dominance: an approach to decision-making under risk*. Lexington Books, 1978.

[102] L. Xu and B. Yu. CVaR-constrained stochastic programming reformulation for stochastic nonlinear complementarity problems. *Computational Optimization and Applications*, 58(2):483–501, 2014.

[103] V. A. Yakubovich. S-procedure in nonlinear control theory. *Vestnick Leningrad Univ. Math.*, 4:73–93, 1997.

[104] Nobuo Yamashita and Masao Fukushima. Modified newton methods for solving a semismooth reformulation of monotone complementarity problems. *Mathematical Programming*, 76(3):469–491, 1997.

[105] K. Yonekura and Y. Kanno. Second-order cone programming with warm start for elastoplastic analysis with von Mises yield criterion. *Optimization and Engineering*, 13(2):181–218, 2012.

[106] F. Zhang. *The Schur complement and its applications*, volume 4. Springer Science & Business Media, 2006.

[107] Hongchao Zhang and William W Hager. A nonmonotone line search technique and its application to unconstrained optimization. *SIAM journal on Optimization*, 14(4):1043–1056, 2004.

[108] L. Zhang. On the convergence of a modified algorithm for the spherical facility location problem. *Operations Research Letters*, 31(2):161–166, 2003.

[109] L. L. Zhang, J. Y. Li, H. W. Zhang, and S. H. Pan. A second order cone complementarity approach for the numerical solution of elastoplasticity problems. *Computational Mechanics*, 51(1):1–18, 2013.

[110] X. Zhang, C. Ling, and L. Qi. The best rank-1 approximation of a symmetric tensor and related spherical optimization problems. *SIAM Journal on Matrix Analysis and Applications*, 33(3):806–821, 2012.

[111] G. L. Zhou and L. Caccetta. Feasible semismooth newton method for a class of stochastic linear complementarity problems. *Journal of Optimization Theory and Applications*, 139(2):379, 2008.