On strong semismoothness and superlinear convergence of complementarity problems over homogeneous cones

Nguyen, Hai Ha

2018

https://doi.org/10.32657/10356/74464

NANYANG TECHNOLOGICAL UNIVERSITY

ON STRONG SEMISMOOTHNESS AND SUPERLINEAR CONVERGENCE OF COMPLEMENTARITY PROBLEMS OVER HOMOGENEOUS CONES

NGUYEN HAI HA

School of Physical and Mathematical Sciences

2017

On strong semismoothness and superlinear convergence of complementarity problems over homogeneous cones N.H.Ha

ON STRONG SEMISMOOTHNESS AND SUPERLINEAR CONVERGENCE OF COMPLEMENTARITY PROBLEMS OVER HOMOGENEOUS CONES

NGUYEN HAI HA

School of Physical and Mathematical Sciences

A thesis submitted to the Nanyang Technological University
in partial fulfillment of the requirement for the degree of
Master of Science

2017

Acknowledgements

I want to show my true gratefulness to Professor Chek Beng Chua. As my supervisor and instructor, he was always patient and careful in giving me guidance towards my study and also taught me precious lessons in life.

I want to give my special thank to Professor Chan Song Heng and Professor Toh Pee Choon, for their valuable advices which brings me more motivation to overcome the difficult times.

I appreciate Nanyang Technological University for providing me a chance as well as the financial support to study in Singapore.

I also want to give my thank to MAS division leaders and staffs, who always created a good working environment for me and provided me with helps in time.

I would like to acknowledge my friends: Ms. Mai Quynh Tram, Ms. Nguyen Thi Ngan Giang, Mr. Nguyen Huy Tiep, Mr. Vu Huy Hien and Mr. Trinh Ha Chuong, who always encouraged me and shared both joyful as well as miserable moments with me.

Neeless to say, when it feels like I lost all the direction, my home is the place to comeback, just like the lighthouse in the stormy night. I dedicate this thesis to my parents and my younger brother, who always be in my side, fill me with love, warmness and hope.

Contents

Acknowledgements **iii**

Abstract **viii**

Notation **xi**

1 Introduction and motivation **1**

1.1 Motivation 1

1.1.1 Motivation for the strong semismoothness of projection onto homogeneous cones 1

1.1.2 Motivation for the superlinear convergence of Newton's method solving complementarity problems 2

1.2 Concepts and notations for the space of symmetric matrices 3

1.3 Concepts and notations for Euclidean Jordan algebras and corresponding Symmetric Cones 4

1.4 Concepts and notations for second-order cone 7

1.5 Overview of contributions 8

2 On the strong semismoothness of the projection onto homogeneous cones **9**

2.1 Differentiability and semismoothness of mappings 9

2.2 One approach to consider the strong semismoothnesss of the projection onto homogeneous cones 11

2.3 Application on second-order cones 14

2.4 The importance of choosing the cone K and the mapping G 18

2.4.1 Projection onto the cone of positive semi-definite Hermitian matrices 18

2.4.2 Projection onto the SOC 22

3 On the superlinear convergence of smoothing Newton continuation algorithm 23

3.1 Stating the problem 23

3.2 Proof for the case of PSD cones 25

3.3 Proof for the case of symmetric cones 31

4 Conclusion and future work 40

Abstract

In Chapter 1, we first review several literature and relevant results that lead to the ideas of the main problems discussed within the thesis. The subsequent parts provide the basic notations and definitions for basic concepts regarding to the main classes of cones we consider in the thesis, including positive semi-definite (PSD) cones, symmetric cones and second-order cones (SOCs). Especially, for the class of symmetric cones, beside defining the symmetric cone via using the concept of homogeneous cone, we also introduce the closely related concepts like Euclidean Jordan algebra, Jordan frame, Pierce decomposition, etc. In the last section of this chapter, we take a glance over the main contributions, discussed in Chapers 2 and Chapter 3.

We start Chapter 2 by recalling several concepts about differentiability, semismoothness and strong semismoothness. In the next section, we revise the method of verifying the strong semismoothness of projection onto the closed convex cone $\tilde{K}$ in the vector space X given in the article "On the Semismoothness of Projection Mappings and Maximum Eigenvalues Function" by M. Goh and F. Meng, and divide the method into four steps. The next parts of Chapter 2 discuss the application of the method for adjusting the strong semismoothness of projection onto second-order cones, then give a couple of counter examples to see the important things we need to notice when doing this method.

Chapter 3 mentions the smoothing Newton continuation algorithm firstly given in the article"A combined smoothing and regularization method for monotone second-order cone complementarity problems" by S. Hayashi, N. Yamashita and M. Fukushima (Algorithm 2) to solve the SOC complementarity problems. C.B. Chua and L. T. K. Hien, in their article "A superlinearly convergent smoothing Newton continuation algorithm for variational inequalities over definable set", give the criterion for this algorithm to converge superlinearly when being applied to solve the smoothing natural map equation. The follow up sections of Chapter 3 give the proof for a lemma that ensure the sufficient

condition for one of the criterion, applied for the case of PSD cones, then generalize to symmetric cones (in the paper of Chua and Hien, the lemma is applied for the epigraph of nuclear norm). The method used for the proofs is based on the explicit formular for the smoothing approximations and application of Lowner's operator for the spectral decomposition.

Chapter 4 sums up the works of Chapter 2 and Chapter 3. It also points out the difficulties we may encounter for doing the method discussed in Chapter 2. Finally, we consider the possible way of generalize the lemma in Chapter 3 to the case of homogeneous cones, when we cannot get the implicit formula for the smoothing approximation, by using the graphical convergence of monotone mappings.

Notation

$\lvert a \rvert$	the absolute value of a scalar a
$\mathbb{R}$	set of real numbers
$\mathbb{R}^n$	the n-dimensional real vector space
$\mathbb{C}$	set of complex numbers
$\mathbb{C}^n$	the n-dimensional complex vector space
$\mathbb{R}^{m\times n}$	the space of $m \times n$ real matrices
$\mathbb{S}^n$	the space of $n \times n$ symmetric matrices.
$\mathbb{S}^n_+$	the cone of $n \times n$ symmetric positive semidefinite matrices
$\mathcal{O}^n$	set of orthogonal matrices in $\mathbb{S}^n$
$\mathrm{Diag}(\lambda_1, \ldots, \lambda_n)$	the $n \times n$ diagonal matrix with $(\mathrm{Diag}(\lambda_1, \ldots, \lambda_n))_{ii} = \lambda_i, i \in \{1, \ldots, n\}$
I	identity matrix of appropriate order
$\mathrm{tr}(A)$	trace of a matrix A
A^{-1}	the inverse of an operator or a matrix A

A^T	the transpose of a matrix A
$\det A$	the determinant of a matrix A
L_A	the Cholesky factor of a matrix A, if A is positive semidefinite
$A \succeq B$	means $A - B$ is positive semidefinite, i.e., $x^T(A-B)x \geq 0\ \forall x$, or all eigenvalues of $A - B$ are non-negative
$A \succ B$	means $A - B$ is positive definite, i.e., $x^T(A-B)x > 0\ \forall x \neq 0$, or all eigenvalues of $A - B$ are positive
$X^\sharp$	the dual set of a set X
X^-	the polar set of a set X
$\mathrm{int}(X)$	the interior of a set X
$\bar{X}$	the closure of a set X
$\mathrm{bd}(X)$	the boundary of a set X, i.e., $\mathrm{bd}(X) = \bar{X} \setminus \mathrm{int}(X)$
Π_X	the Euclidean projector onto a set X
∇f	the $\mathcal{F}$-derivative of a function f
$\mathbf{J}F_x(x,y)$	the partial derivative of F with respect to variable x
$-\ln\det(X)$	the barrier function defined on $\mathbb{S}^n_{++}$, given by $-\ln\det(X) := -\ln(\det X)$

Chapter 1

Introduction and motivation

1.1 Motivation

The initial studies for the *complementarity problems* (CPs) can be dated back to the 60s decade of the twentieth century, for instance, when Lemke and Howson, in their paper [19], gave the *linear complementarity problem* (LCP) form for calculating a *Nash equilibrium point*. In the same year (1964), Richard W. Cottle introduced the classical *nonlinear complementary problem* (NCP) in his thesis. These are also the first problems relating to the general problem of *variational inequalities problem* (VIP), until Hartman and Stampacchia gave the concept of VIP in 1966. One of the noticeable materials regarding to VIPs and CPs is the book [12] where we can find a great deal of formulations for those problems and related algorithms.

1.1.1 Motivation for the strong semismoothness of projection onto homogeneous cones

Among the problems that can be written in the form of CP, a problem of finding the projection onto homogeneous cones can be justified under many aspects. One aspect that we mentioned in Chapter 2 is considering the strong semismoothness of that projection. It is already proved for a class of symmetric cones, which is a subclass of homogeneous cones, that the projections onto those cones are strongly semismooth. The result given in [2] that a function that is locally Lipschitz and definable is also a semismooth one brings a great help in considering the semismooth. Since the projections onto classes of cones like homogeneous cones or power cones can be proved to be locally Lipschitz definable, they are semismooth. Howerver, the strong semismoothness of projection onto homogeneous cone in general is still a challenge. Wang and Xiu, in their paper [25], showed that the projection onto any slice of SOC is strong semismooth. However, the strong semismoothness of projection onto other nonpolyhedral symmetric cones, especially PSD cones, remains an unsolved question. That question draws our careful attention because of the fact mentioned in [8] that each homogeneous is actually isomorphic to a slice of a PSD cone. Motivated by those results, in this thesis, we will discuss one of the methods that possibly helps us carry out deeper study into the strong semismoothness of the projection onto homogeneous cones.

1.1.2 Motivation for the superlinear convergence of Newton's method solving complementarity problems

The Newton's approximation method for solving equations was invented and used from 17-th century. Several developments of the classical Newton's method can be listed. Pang [20] extended the classical Newton's method used to solve the continuously differentiable systems of nonlinear equations to the systems which is B-differentiable, which opened up a fundamental for nonlinear programming problems, complementarity problems, etc. Later on, a nonsmooth version of this method was created by Qi and Sun in [21], which is applied for locally Lipschitz and semismooth functions. One of the applications of Newton's method is solving complementarity problems via solving the corresponding natural map equations, which is essentially related to Euclidean projectors. To avoid the difficulty of using typical Newton-based method caused by the general nonsmoothness of the Euclidean projector, the Newton's method can be considered to be applied for the smoothing approximation of Eucliden projector. An example of this method is in [17], where the combined and smoothing regularization method was built to solve complementarity problems for second-order cones. Within the context of this thesis, we will consider one of the features that makes up the superlinear convergence of one algorithm mentioned in [17].

1.2 Concepts and notations for the space of symmetric matrices

In this thesis, we will let I_m denote the identity matrix of size $m \times m$ for some positive integer m. Besides, for simplicity, the notation I will also be used to denote the identity matrix if the size is already clearly determined and no confusion would happen (the general case that we consider is the space of $n \times n$ symmetric matrices, hence, I would always denote the corresponding matrix I_n). At the same time, we always let O denote the zero matrix.

Let $\mathbb{S}^n$ denote the space of real symmetric $n \times n$ matrices. $\mathbb{S}^n_+$ and $\mathbb{S}^n_{++}$ will be used to denote the subsets of $\mathbb{S}^n$, which are the cone of positive semi-definite matrices (from now on, this cone will be mentioned as the PSD cone) and the cone of positive definite matrices respectively, which are:

$$\mathbb{S}^n_+ = \{X \in \mathbb{S}^n : X \succeq 0\}$$
$$\mathbb{S}^n_{++} = \{X \in \mathbb{S}^n : X \succ 0\}.$$

Let $\Pi_{\mathbb{S}^n_+}$ denote the Euclidean projector onto $\mathbb{S}^n_+$, i.e.,

$$\Pi_{\mathbb{S}^n_+}(Z) = \arg\min_{X \in \mathbb{S}^n_+} \frac{1}{2}\|X - Z\|^2,$$

where $\|\cdot\|$ is the norm induced by the trace inner product $\langle \cdot, \cdot \rangle$ of $\mathbb{S}^n$ given by

$$\langle A, B \rangle = \mathrm{tr}(AB) \text{ for all } A, B \in \mathbb{S}^n.$$

During this thesis, the Euclidean projectors onto other cones in other spaces are also defined similarly.

Let $\mathbb{E}$ and $\mathbb{E}'$ denote two finite-dimensional vector spaces each equiped with an inner product and its corresponding induced norm. A *smoothing approximation* of a continuous map $G : \mathbb{E} \to \mathbb{E}'$ is a continuous map $H : \mathbb{E} \times \mathbb{R}$ such that $H(\cdot, \mu) = G$ when $\mu \leq 0$, and for each $\mu > 0$, $H(\cdot, \mu)$ is differentiable.

Let $p(Z,\, \mu)$ denote the smoothing approximation of the corresponding Euclidean projector. For the case we consider the cone K being $\mathbb{S}^n_+$, p is corresponding to the barrier function defined on $\mathbb{S}^n_{++}$, given by $f(X) = -\ln\det(X) := -\ln(\det X)$.

For this thesis, we also use the formula $\mathrm{Diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)$ to denote the diagonal matrix with the main diagonal containing $\lambda_1, \lambda_2, \ldots, \lambda_n$ respectively. For a matrix $Z \in \mathbb{S}^n$, with $\lambda_1(Z) \geq \lambda_2(Z) \geq \cdots \geq \lambda_n(Z)$ being its eigenvalues with multiplicity, we let $\lambda_{f}(_Z) := (\lambda_1(Z), \ldots, \lambda_n(Z))^T$; and we use $\mathcal{O}^n(Z)$ to denote the set of orthogonal matrices such that $\mathrm{Diag}(\lambda_f(_Z)) = QZQ^T$.

We would also use the concept of Cholesky factor of a PSD matrix, which is given in [7]. For a matrix A which is positive semidefinite, we let L_A denote a unique lower triangular matrix satisfying $L_A L_A^T = A$ and if any entry on its main diagonal satisfies $(L_A)_{ii} = 0$, then all the entries on the corresponding i-th column equal zero. Such matrix L_A is called the Cholesky factor of A.

1.3 Concepts and notations for Euclidean Jordan algebras and corresponding Symmetric Cones

The concepts and notations we use in this part are mainly taken from the book [13], which is a cornerstone in symmetric cones study.

A *homogeneous cone* K is a full-dimensional convex pointed cone in a finite-dimensional space such that the group of linear automorphism of K acts transitively on it. A *symmetric cone* is a self-dual homogeneous cone. It is proved in [13] that any real symmetric cone is the interior of the cone of all squares of some real *Euclidean Jordan algebra*. Therefore, within the context of this thesis, we always consider the symmetric cone determined in this way from some Euclidean Jordan algebra. Now we take a closer look into the concept of Euclidean Jordan algebra.

A vector space V over $\mathbb{R}$ is an algebra if there exists a *product* on V, which is a bilinear mapping from $V \times V$ to V, denoted by $(x, y) \mapsto xy$. For $x \in V$, we have the linear mapping $L(x)$ defined by $L(x)y := xy$. The algebra V is called *Jordan algebra* if any $x, y \in V$ satisfy the following conditions:

$$xy = yx \tag{J1}$$

$$x(x^2y) = x^2(xy). \tag{J2}$$

A Jordan algebra V is said to be *Euclidean* if there exists an inner product $\langle \cdot, \cdot \rangle$ on V, which is a positive definite symmetric bilinear form from $V \times V$ to $\mathbb{R}$ satisfying $\langle L(x)u, v \rangle = \langle u, L(x)v \rangle$ for all x, u, v in V.

For an Euclidean Jordan algebra V, $c \in V$ is said to be an *idempotent* if $c^2 = c$. Two idempotents c and d in V is said to be *orthogonal* if $cd = 0$. An element c in V is called a primitive idempotent if it is a non-zero idempotent and we are not able to represent c as the sum of the other two non-zero idempotents. We say that the idempotents $c_1, \ldots, c_m$ in V make up a *Jordan frame*, if each c_j is a primitive one and if

$$c_j c_k = 0,\ j \neq k,$$
$$\sum_{j=1}^{m} c_j = e.$$

Let V be a real Euclidean Jordan algebra. It is pointed out in Theorem III.1.2 of [13] that for any $x \in V$, there exists a Jordan frame $c_1, \ldots, c_r$ and the scalars $\lambda_1, \ldots, \lambda_r \in \mathbb{R}$ such that

$$x = \sum_{j=1}^{r} \lambda_j c_j.$$

The scalars λ_i (with their multiplicities) are uniquely determined by x and called the *eigenvalues* of x. Besides, we use $\det x$ and $\operatorname{tr} x$ to denote the *determinant* and *trace* of x respectively, which are given by[1]

$$\det x := \prod_{j=1}^{r} \lambda_j,\ \ \operatorname{tr} x := \sum_{j=1}^{r} \lambda_j.$$

It is also pointed out in [13] that for $x, y \in V$, the bilinear form given by $(x, y) \mapsto \operatorname{tr}(xy)$ is an inner product of V.

Let V be a real Euclidean Jordan algebra with an identity element e and the inner product $\langle \cdot, \cdot \rangle$ defined by $\langle x, y \rangle = \operatorname{tr}(xy)$. Now we consider the set $K := \{x^2 \mid x \in V\}$. K

[1] Actually, in [13], the determinant and trace of x is defined via the *minimal polynomial* of x. However, within the context of this thesis, we just need to determine those concepts via the eigenvalues of x.

is the closed cone and therefore its closed dual

$$
\begin{aligned}
K^{\sharp} &= \{y \in V \mid \forall x \in V,\ \langle y, x^2 \rangle\ \geq 0\} \\
&= \{y \in V \mid \forall x \in V,\ \langle L(y)x, x \rangle\ \geq 0\} \\
&= \{y \in V \mid L(y) \text{ is positive semi-definite}\}
\end{aligned}
$$

is a closed convex cone. It is also proved in [13] that $K = K^{\sharp}$ and as a consequence, K is self-dual and a closed convex cone itself. As we already pointed out before, the interior Ω of K is a symmetric cone, the closure $\bar{\Omega}$ of Ω is exactly K itself and we also know that Ω is the connected component of e in the set $\mathfrak{I}$ of invertible elements in V.

Now we give the brief introduction to the Pierce decomposition. Let c be an idempotent element in an Euclidean Jordan algebra V. By Proposition III.1.3 of [13], $L(c)$ has only three possible values: 0, 1 and $\frac{1}{2}$. Therefore, V can be represented as the direct sum of the subspaces generated by those eigenvalues, which we denoted by $V(0, c)$, $V(1, c)$ and $V(\frac{1}{2}, c)$. The decomposition

$$
V = V(0, c) + V(1, c) + V(\frac{1}{2}, c)
$$

is called the *Pierce decomposition* of V with respect to the idempotent c.
Let V be a real Euclidean Jordan algebra and $c_1, \ldots, c_r$ be a Jordan frame in V. By the Theorem IV.2.1 in [13], V decomposes in the following orthogonal direct sum:

$$
V = \bigoplus_{1 \leq i \leq j \leq r} V_{ij},
$$

where,

$$
\begin{aligned}
&V_{ii} = V(1, c_i) = \mathbb{R}c_i, \\
&V_{ij} = V(\frac{1}{2}, c_i) \cap V(\frac{1}{2}, c_j) \qquad \text{when } i \neq j.
\end{aligned}
$$

The orthogonal projection onto V_{ij} is then denoted by Π_{ij} and given by

$$
\begin{aligned}
\Pi_{ii} &= L(c_i)(2L(c_i) - I), \\
\Pi_{ij} &= 4L(c_i)L(c_j).
\end{aligned}
$$

where I is an identity mapping in V. For an element $h \in V$, the reperesentation of h as the direct sum:

$$h = \sum_{1 \leq i \leq j \leq r} h_{ij} \qquad \text{for } h_{ij} \in V_{ij},$$

is called the *Pierce decomposition of* h.

1.4 Concepts and notations for second-order cone

Let $\mathcal{K}^n$ denote the second-order cone (SOC), or Lorentz cone in $\mathbb{R}^n$, which is defined as

$$\mathcal{K}^n = \{(x_1, x_2) \in \mathbb{R} \times \mathbb{R}^{n-1} \mid x_1 \geq \|x_2\|\}.$$

Actually, SOC belongs to the category of symmetric cones. Thus, a SOC $\mathcal{K}^n$ is a cone of squares of an Euclidean Jordan algebra. For any $x = (x_1, x_2) \in \mathbb{R} \times \mathbb{R}^{n-1}$ and $y = (y_1, y_2) \in \mathbb{R} \times \mathbb{R}^{n-1}$ their Jordan product is defined by

$$x \circ y = (\langle x, y \rangle, y_1 x_2 + x_1 y_2).$$

As being mentioned in [4], for each $x = (x_1, x_2) \in \mathbb{R} \times \mathbb{R}^{n-1}$, x has the spectral decomposition associated with $\mathcal{K}^n$ which has the form

$$x = \lambda_1(x) u_x^{(1)} + \lambda_2(x) u_x^{(2)} \tag{1.1}$$

where $\lambda_1(x)$, $\lambda_2(x)$ and $u_x^{(1)}$, $u_x^{(2)}$ are the spectral values and the corresponding spectral vectors of x, respectively, given by

$$\lambda_i(x) := x_1 + (-1)^i \|x_2\| \quad \text{and} \quad u_x^{(i)} = \begin{pmatrix} 1 \\ (-1)^i \bar{x}_2 \end{pmatrix}, \quad i = 1, 2, \tag{1.2}$$

where $\bar{x}_2 = x_2 / \|x_2\|$ if $x_2 \neq 0$, and otherwise $\bar{x}_2$ being any vector in $\mathbb{R}^{n-1}$ with $\|\bar{x}_2\| = 1$.

1.5 Overview of contributions

In Chapter 2, we discuss the method containing four steps, given by Goh and Meng in their paper [15], to justify the strong semismoothness of projection onto homogeneous cones and its application to the case of SOC cones, as well as the difficulties that may arise from this method.

In Chapter 3, the problem that we consider gives us the overview of the sufficient condition for one of the criteria to consider the superlinear convergence of an algorithm given in [17] by Hayashi, Yamashita and Fukushima. We then consider the superlinear convergence of this algorithm when being applied to case of symmetric cones, which is the subcase of a category of homogeneous cones. We will prove the lemma that provides a sufficient condition for one of the criteria to ensure the superlinear convergence.

Chapter 2

On the strong semismoothness of the projection onto homogeneous cones

We consider the problem of finding the Euclidean projector. For a closed convex subset K of a finite dimensional Euclidean vector space $\mathbb{E}$, the problem of finding the (Euclidean) projector $\Pi_K(y)$ of a vector $y \in \mathbb{E}$ is to find the unique solution of the following convex minimization problem:

$$\begin{aligned} &\text{minimize} \quad \frac{1}{2}(y-x)^T(y-x) \\ &\text{subject to} \quad x \in K. \end{aligned}$$

According to [12], this problem is in the form of the variational inequality, thus, it can be rewritten in the equivalent form of a *complementarity problem.*
In this section, we will discuss the problem of considering the strong semismoothness of the projection onto homogeneous cones, which is not yet fully solved. Within the context of the thesis, we will consider one of the possible approaches towards the problem and observe the application of that approach to a couple of subclasses of the class of homogeneous cones. Through this point of view, we have the initial view on the advantage and drawback of the method.

2.1 Differentiability and semismoothness of mappings

We begin with recalling several concepts that would be used later on.

Definition 2.1.1. Consider two finite-dimensional vector spaces X and Y which are equipped with the inner products $\langle\cdot,\cdot\rangle_X$, $\langle\cdot,\cdot\rangle_Y$ and their induced norms $\|\cdot\|_X$, $\|\cdot\|_Y$ respectively (from now on, for simplicity, we can use the common notation $\|\cdot\|$ for both $\|\cdot\|_X$ and $\|\cdot\|_Y$. Similarly, we use $\langle\cdot,\cdot\rangle$ to replace $\langle\cdot,\cdot\rangle_X$ and $\langle\cdot,\cdot\rangle_Y$). Let $f : X \to Y$ be a single-valued mapping and let $x, h \in X$. If $\nabla f(x) : X \to Y$ is a linear mapping that satisfied

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

then f is said to be *Fréchet differentiable* ($\mathcal{F}$-differentiable) and $\nabla f(x)$ is a $\mathcal{F}$-derivative of f at x.

We also give the definition of directional derivative.

Definition 2.1.2. Let $f : X \to Y$, $x, h \in X$ and τ be a positive real number. The *directional derivative* of f at x along h is the following limit (if it exists)

$$f'(x,h) = \lim_{\tau\downarrow 0} \frac{f(x+\tau h) - f(x)}{\tau}.$$

f is said to be *directionally diffentiable* at x if $f'(x,h)$ exists for all h.

Certainly, we have the following relation between $\mathcal{F}$-differentiability and directional differentiability.

Theorem 2.1.1. *If f is $\mathcal{F}$-differentiable at x, then it is directionally differentiable at x and for every $h \in X$,*

$$f'(x,h) = \nabla f(x)h.$$

We also need the concept of the Clarke generalized derivative, which plays an important role in stating the concept of semismoothness.

Definition 2.1.3. Let $f : X \to Y$ be a locally Lipschitzian mapping. Let D_f denote the set of points at which f is $\mathcal{F}$-differentiable[1]. We let $\partial f(x)$ denote the *generalized derivative* of f at x, which is the set defined by

$$\partial f(x) := \mathrm{co}\left\{ \lim_{k\longrightarrow\infty} \nabla f(z_k) \mid \{z_k\} \subseteq D_f, z_k \to x \right\}.$$

[1]By Rademacher's theorem, a locally Lipschitzian mapping on X is $\mathcal{F}$-differentiable almost everywhere on X.

Notice that we use $\operatorname{co}(C)$ to denote the convex hull of a set C.

Now we give the definition of semismooth mapping. Though there are various ways of stating this concept, here we will use the similar definition with the one for semismooth matrix valued function given in [24].

Definition 2.1.4. Let $f : X \to Y$ be a locally Lipschitzian mapping. f is said to be *semismooth* at $x \in X$ if f is directionally differentiable at x and for any $V \in \partial f(x+h)$ and $h \to 0$,

$$f(x+h) - f(x) - Vh = o(\|h\|).$$

f is said to be p-order semismooth $(0 < p < \infty)$ if f is semismooth at x and

$$f(x+h) - f(x) - Vh = O(\|h\|^{1+p}).$$

In particular, for the case $p = 1$, f is said to be *strongly semismooth*.

2.2 One approach to consider the strong semismoothnesss of the projection onto homogeneous cones

Now we consider the method of proving the strong semismoothness for the projection mapping which is mention in [15]. We first state that problem. Let X and Y be two finite-dimensional vector spaces, each of them has an inner product $\langle \cdot, \cdot \rangle$ and its induced norm $\|\cdot\|$. $K \subseteq Y$ is a closed convex set. Let $G : X \to Y$ be a singled-valued continuously differentiable mapping. Consider the parametric optimization problem, parameterized by $y \in X$

$$\begin{aligned} &\min \frac{1}{2}\|x - y\|^2, \\ &\text{s.t. } G(x) \in K. \end{aligned} \tag{2.1}$$

Let $G^{-1}(K)$ be the set

$$G^{-1}(K) = \{x \in X \mid G(x) \in K\}. \tag{2.2}$$

The solution mapping of (2.1) can be represented as the following mapping

$$\Pi_{G^{-1}(K)}(y) := \operatorname{argmin}\left\{\frac{1}{2}\|x-y\|^2 \mid G(x) \in K\right\}. \tag{2.3}$$

which is the projection onto $G^{-1}(K)$.

For any vector $s \in Y$, we let $T_K(s)$ denote the tangent cone of K at s. For an arbitrary closed convex cone C, we use $\mathrm{lin}(C)$ to denote the lineality space of C, i.e., $\mathrm{lin}(C) = C \cap (-C)$. Now we bring out the concept of nondegeneracy, which always plays an important role in the approach that we will consider later on.

Definition 2.2.1. $\bar{x} \in G^{-1}(K)$ is said to be *nondegenerate*, with respect to the mapping G and the set K, if

$$\nabla G(\bar{x})X + \mathrm{lin}(T_K(G(\bar{x}))) = Y. \tag{2.4}$$

Next, we have the Lagrangian function of problem (2.1) to be

$$L(x,\Lambda,y) = \frac{1}{2}\langle x-y, x-y\rangle_X + \langle \Lambda, (-G(x))\rangle_Y, \tag{2.5}$$

where $(y,\Lambda,x) \in X \times Y \times X$. According to [15], the first order necessary condition of (2.1) can be written in the form

$$\mathcal{H}(x,\Lambda;y) = 0, \tag{2.6}$$

where $\mathcal{H}(x,\Lambda;y)$ is the vector-valued function given by

$$\mathcal{H}(x,\Lambda;y) = \begin{bmatrix} x - \nabla G(x)^*\Lambda - y \\ \Lambda - \Pi_{K^\sharp}[\Lambda - G(x)] \end{bmatrix},$$

with $\nabla G(x)^* : Y \to X$ is the adjoint of $\nabla G(x)$.
The solution of equation (2.6) is also called the Karush-Kuhn-Tucker solution, or KKT solution for simplicity.

In [15], it is pointed out that under the nondegeneracy condition (2.4), for any given $\bar{y} \in X$, equation (2.6) always has the unique corresponding solution $(\bar{x}, \bar{\Lambda})$. We also have the following assumption, which is considered under the condition that $\bar{x}$ is nondegenerate.

Assumption 2.2.1. If $(\mathcal{I} - V)(H) = 0$ for $V \in \partial\Pi_K(\bar{\Lambda} - G(\bar{x}))$ and $H \in Y$, then $H \in [\mathrm{lin}(T_K(G(\bar{x})))]^{\perp}$, where $\mathcal{I}$ denotes the identity mapping from Y to Y.

Now, we recall the Theorem 3.1 in [15] that we would use as a main approach for the method that we consider to prove the strong semismoothness of the projection onto homogeneous cones.

Theorem 2.2.1. *Given $\bar{y} \in X$, let $(\bar{x}, \bar{\Lambda})$ be the corresponding KKT solution of (2.6). Suppose (i) $\bar{x}$ is nondegenerate with respect to the mapping G and set K; (ii) The mapping $G : X \to Y$ is affine; (iii) Assumption 2.2.1 holds. Then,*
(i) there exists an open neighborhood $\mathcal{N}$ of $\bar{y}$ and a Lipschitz continuous function $(x(\cdot), \Lambda(\cdot))$ defined on $\mathcal{N}$ such that $\mathcal{H}(x(y), \Lambda(y); y) = 0$ for every $y \in \mathcal{N}$;
(ii) if Π_K is semismooth (strongly semismooth) around $\bar{\Lambda} - G(\bar{x})$, then $(x(\cdot), \Lambda(\cdot))$ is semismooth (strongly semismooth) around $\bar{y}$.

The above theorem gives us the idea of how to prove that the projection into a closed convex cone $\tilde{K} \subseteq X$ is strongly semismooth at some point $\bar{y} \in X$. The following are general steps of the process.

- **Step 1:** First, we need to find a closed convex cone K in the vector space Y that we already know the projection onto K is strongly semismooth.
- **Step 2:** The next step is to find a mapping $G : X \to Y$ that is affine and also satisfies $G^{-1}(K) = \tilde{K}$.
- **Step 3:** For a KKT solution $(\bar{x}, \bar{\Lambda}) = (x(\bar{y}), \Lambda(\bar{y}))$ corresponding to $\bar{y}$, we are able to verify that $\bar{x}$ is nondegenerate and Assumption 2.2.1 holds with $\bar{x}$ and $\bar{\Lambda}$.
- **Step 4:** Using Theorem 2.2.1, we conclude that $\Pi_{\tilde{K}}$ is strongly semismooth at $\bar{y}$.

The cone K that we put into consideration could be $\mathbb{S}^n_+$, the positive semi-definite cone. The reason for this choice is that the projection onto PSD cone is strongly semismooth, which is a result shown in Theorem 4.13 of [24]. Besides, it is already proved in Section 5 of [15] that under the nondegeneracy condition, Assumption 2.2.1 always holds if K is a cone of positive semi-definite matrices. Consequently, this way of chosing the cone K may reduce the works we need to carry out for the four steps mentioned above since we do not need to justify the validity of Assumption 2.2.1.

There is one more reason for us to draw our attention in using K as a PSD cone to prove the strong semismoothness of the projection onto homogeneous cones. It is proved in Corollary 4.3 of [8] that for any homogeneous cone $\tilde{K}$, there always exists an injective linear mapping M that maps $\tilde{K}$ into a slice of a PSD cone. This mapping M is obviously an affine one, hence, it could be a mapping G that we look for in **Step 2**. The cone K we may use here is certainly a PSD cone.

The notice in the relation between homogeneous cones and PSD cones could be the direction for the further study on proving the strong semismoothness of projection onto homogenous cones. However, within the context of this thesis, we only consider two simpler classes of homogeneous cones, that will be presented in the following parts.

2.3 Application on second-order cones

Now we proceed to use the method with four steps that we stated previously to verify the strong semismoothness of the projection onto second-order cones. It is convenient for us to choose the cone K in **Step 1** as a PSD cone. Now we consider the linear mapping $G : \mathbb{R} \times \mathbb{R}^{n-1} \to \mathbb{S}^n$ given by $G(x) = L_x$, where $x = (x_1, x_2) \in \mathbb{R} \times \mathbb{R}^{n-1}$ and L_x is defined as

$$L_x := \begin{bmatrix} x_1 & x_2^T \\ x_2 & x_1 I_{n-1} \end{bmatrix}.$$

L_x is an arrow-shaped matrix, which is sometimes called an arrow matrix and denoted by $\mathrm{Arw}(x)$. The mapping G in this case is obviously a linear mapping and hence an affine one. We use this mapping G because there is a close relation between elements in $\mathcal{K}^n$ and positive semi-definite arrow matrices. We can see that

$$x \in \mathcal{K}^n \Longleftrightarrow L_x = \begin{bmatrix} x_1 & x_2^T \\ x_2 & x_1 I_{n-1} \end{bmatrix} \succeq O,$$

Actually, this fact is verified in [1] with a very short argument. We have L_x is PSD if and only if either $x_1 = 0$, $x_2 = 0$ or $x_1 > 0$ and the Schur complement $x_1 - x_2^T (x_1 I_{n-1})^{-1} x_2 = \dfrac{1}{x_1}(x_1^2 - \|x_2\|^2) \geq 0$. Therefore, we easily see that $G(\mathcal{K}^n) \subset \mathbb{S}^n_+$ and actually, $G^{-1}(\mathbb{S}^n_+) = \mathcal{K}^n$.

According to [1], there are some known relations between the SOC and the PSD cone, for instance,

(a) $x \in \operatorname{int} \mathcal{K}^n \Leftrightarrow L_x \in \operatorname{int} \mathbb{S}^n_+$;

(b) $x = 0 \Leftrightarrow L_x = O$;

(c)$x \in \operatorname{bd} \mathcal{K}^n \setminus \{0\} \Leftrightarrow L_x \in \operatorname{bd} \mathbb{S}^n_+ \setminus \{O\}$.

Before moving on, we need to mention the following spectral decomposition for L_x.

Lemma 2.3.1. *Let $x = (x_1, x_2) \in \mathbb{R} \times \mathbb{R}^{n-1}$ have the spectral decomposition as being given in (1.1)-(1.2). Then L_x has the spectral decomposition:*

$$L_x = P \operatorname{Diag}(\lambda_1(x), \lambda_2(x), x_1, \ldots, x_1) P^T \tag{2.7}$$

where

$$P = \left[\sqrt{2} u_x^{(1)} \ \sqrt{2} u_x^{(2)} \ u_x^{(3)} \ \cdots \ u_x^{(n)} \right] \in \mathbb{R}^{n \times n} \tag{2.8}$$

is an orthogonal matrix, and $u_x^{(i)}$, for $i = 3, \ldots, n$, have the form of $(0, \bar{u}_i)$ with $\bar{u}_3, \ldots, \bar{u}_n$ forming an orthonomal basis for the linear subspace orthogonal to x_2. Needless to say, any matrix P of this form belongs to $\mathcal{O}^n(L_x)$.

The proof for Lemma 2.3.1 can be found in [6].

At this stage, we already find the cone K and a mapping G for **Step 1** and **Step 2** of the method. What left is only to verify the nondegeneracy property. Recall from previous arguments that to prove the strong semismoothness of a projection at a point $y \in \mathbb{R} \times \mathbb{R}^{n-1}$, it is sufficient for us to justify the nondegeneracy at $x = x(y) \in \mathcal{K}^n$, where $(x(y), \Lambda(y))$ is the solution of equation (2.6) corresponding to y. That means,

$$\nabla G(x)(\mathbb{R} \times \mathbb{R}^{n-1}) + \operatorname{lin}(T_{\mathbb{S}^n_+}(G(x))) = \mathbb{S}^n. \tag{2.9}$$

However, since G is a linear mapping, $\nabla G(x) = G$ for every $x \in \mathcal{K}^n$. By that fact together with $G(x) = L_x$, we can rewrite (2.7) to be

$$G(\mathbb{R} \times \mathbb{R}^{n-1}) + \operatorname{lin}(T_{\mathbb{S}^n_+}(L_x)) = \mathbb{S}^n. \tag{2.10}$$

Subsequently, we will consider the following cases of x.

- *Case 1.* $x \in \operatorname{int} \mathcal{K}^n$, corresponding to $y = x \in \operatorname{int} \mathcal{K}^n$.

As we mentioned above, $x \in \operatorname{int} \mathcal{K}^n \Leftrightarrow L_x \in \operatorname{int} \mathbb{S}^n_+$. Thus, we see that $L_x \in \operatorname{int} \mathbb{S}^n_+$ in this case. We have $T_{\mathbb{S}^n_+}(L_x)$ is the set of all matrices T in $\mathbb{S}^n$ such that there exist a sequence of matrix $\{T_\nu\} \subset \mathbb{S}^n_+$ and a sequence of positive scalars $\{\tau_\nu\} \in$ such that

$$\lim_{\nu\to\infty} T_\nu = L_x, \quad \lim_{\nu\to\infty} \tau_\nu = 0, \quad \text{and} \quad \lim_{\nu\to\infty} \frac{T_\nu - L_x}{\tau_\nu} = T. \tag{2.11}$$

By the assumption, $L_x \in \operatorname{int} \mathbb{S}^n_+$, then for every $T \in \mathbb{S}^n$, there exists a sequence $\{\tau_\nu\}$ of small enough positive scalars converging to 0 that satisfies: for a sequence $\{T_\nu\}$ which is defined by $T_\nu := L_x + \tau_\nu T$, we have $T_\nu \in \mathbb{S}^n_+$ for every ν. It then follows that

$$\lim_{\nu\to\infty} \frac{T_\nu - L_x}{\tau_\nu} = \lim_{\nu\to\infty} \frac{(L_x + \tau_\nu T) - L_x}{\tau_\nu} = T.$$

Consequently, for every $T \in \mathbb{S}^n$, we also have $T \in T_{\mathbb{S}^n_+}(L_x)$. It shows that $T_{\mathbb{S}^n_+}(L_x) = \mathbb{S}^n$ and obviously $\operatorname{lin}(T_{\mathbb{S}^n_+}(L_x)) = \mathbb{S}^n$. This fact helps us see that the nondegeneracy condition (2.8) holds at x.

• *Case 2.* $x \in \operatorname{bd} \mathcal{K}^n \setminus \{0\}$, corresponding to $y \in \mathbb{R} \times \mathbb{R}^n \setminus (\operatorname{int} \mathcal{K}^n \cup (\mathcal{K}^n)^-)$, where $(\mathcal{K}^n)^-$ is a polar cone of $\mathcal{K}^n$).

In this case, we will use the result from [3] that

$$T_{\mathbb{S}^n_+}(L_x) = \{A \in \mathbb{S}^n \mid (u_x^{(1)})^T A u_x^{(1)} \geq 0\}.$$

Therefore, we have,

$$\operatorname{lin}(T_{\mathbb{S}^n_+}(L_x)) = \{A \in \mathbb{S}^n \mid (u_x^{(1)})^T A u_x^{(1)} = 0\}. \tag{2.12}$$

For a matrix $P \in \mathcal{O}^n(L_x)$ of the form $P = \begin{bmatrix} \sqrt{2}u_x^{(1)} & \sqrt{2}u_x^{(2)} & u_x^{(3)} & \cdots & u_x^{(n)} \end{bmatrix}$ given in (2.8), from (2.12) we see that

$$(P^T A P)_{11} = 0 \quad \text{for any } A \in \operatorname{lin}(T_{\mathbb{S}^n_+}(L_x)), \tag{2.13}$$

where we use $(P^T A P)_{ij}$ to denote the entry at i-th row and j-th column of the matrix $(P^T A P)$. This time, instead of proving the condition (2.8) directly, we will verify its following equivalent form:

$$(G(\mathbb{R} \times \mathbb{R}^{n-1}))^\perp \cap (\operatorname{lin}(T_{\mathbb{S}^n_+}(L_x)))^\perp = \{O\}. \tag{2.14}$$

Suppose that $H \in (G(\mathbb{R} \times \mathbb{R}^{n-1}))^{\perp} \cap (\mathrm{lin}(T_{\mathbb{S}^n_+}(L_x)))^{\perp}$, then for any $A \in \mathrm{lin}(T_{\mathbb{S}^n_+}(L_x))$, we have:

$$\langle H, A \rangle = \mathrm{tr}(HA) = 0. \tag{2.15}$$

Since P is an orthogonal matrix, it is easy to derive that

$$\langle H, A \rangle = \langle P^T HP, P^T AP \rangle = 0 \quad \text{for every } A \in \mathrm{lin}(T_{\mathbb{S}^n_+}(L_x)). \tag{2.16}$$

With the form of A shown in (2.12)-(2.13), in the equality (2.16) we can choose $P^T AP$ to be the symmetric matrix of the form $O(i, j)$ for any $i, j = 1, \dots, n$ and $(i, j) \neq (1, 1)$, where $O(i, j)$ is a symmetric matrix in $\mathbb{S}^n$, defined as

$$O(i,j)_{kl} = O(i,j)_{lk} = \begin{cases} 1 & \text{if } (k,l) = (i,j) \text{ or } (l,k) = (i,j) \\ 0 & \text{if } (k,l) \neq (i,j) \text{ and } (l,k) \neq (i,j) \end{cases}$$

Substituting $P^T AP = O(i, j)$ into the latter equality in (2.16), we can easily get that

$$(P^T HP)_{ij} = 0 \quad \text{for every } (i, j) \neq (1, 1). \tag{2.17}$$

On the other hand, we also have $H \in (G(\mathbb{R} \times \mathbb{R}^{n-1}))^{\perp}$ while

$$G(\mathbb{R} \times \mathbb{R}^{n-1}) = \{L_z \mid z \in \mathbb{R} \times \mathbb{R}^{n-1}\}.$$

Therefore, we have

$$\langle H, L_z \rangle = \langle P^T HP, P^T L_z P \rangle = 0 \quad \text{for every } z \in \mathbb{R} \times \mathbb{R}^{n-1}. \tag{2.18}$$

Due to (2.17), we can rewrite (2.18) as

$$\langle P^T HP, P^T L_z P \rangle = (P^T HP)_{11}(P^T L_z P)_{11} = 0 \quad \text{for every } z \in \mathbb{R} \times \mathbb{R}^{n-1}. \tag{2.19}$$

We can simply let $z = (1, 0, \cdots, 0) \in \mathbb{R} \times \mathbb{R}^{n-1}$. With that choice, we have $L_z = I_n$.

Substituting $L_z = I_n$ into (2.19), we have

$$(P^T HP)_{11}(P^T I_n P)_{11} = (P^T HP)_{11}(I_n)_{11} = (P^T HP)_{11} = 0. \tag{2.20}$$

From (2.17) and (2.20), we see that

$$\begin{aligned} & P^T HP = O, \\ \Leftrightarrow\ & P(P^T HP)P^T = O, \\ \Leftrightarrow\ & H = O. \end{aligned}$$

And we can conclude that the condition (2.14) holds.

• *Case 3.* $x = 0$, corresponding to y belonging to the polar cone of $\mathcal{K}^n$.

We will leave this case for the discussion in the subsequent part.

2.4 The importance of choosing the cone K and the mapping G

Now we will give some counter examples where we cannot use the method, to see the necessity of choosing the appropriate cone K and the mapping G in **Step 1** and **Step 2** of the method.

2.4.1 Projection onto the cone of positive semi-definite Hermitian matrices

Consider the vector spaces $\mathcal{H}^n$ containing all Hermitian matrices of the form $M = A+iB$ where $A \in \mathbb{S}^n$ and B is an $n \times n$ real skew-symmetric matrix, which means $B^T = -B$. Each Hermitian matrix M has the spectral decomposition

$$M = P^\dagger \mathrm{Diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)P$$

where P is an $n \times n$ complex unitary matrix and $P^\dagger$ is its Hermitian conjugate. We also let λ_1,. . . , λ_n denote the eigenvalues of M. Notice that all the eigenvalues of Hermitian matrices are *real numbers*.

Now we consider the cone $\mathcal{H}_+^n$ of positive semi-definite Hermitian matrices, which is determined by

$$\mathcal{H}_+^n = \{M = P^\dagger \mathrm{Diag}(\lambda_1, \lambda_2, \ldots, \lambda_n) P \in \mathcal{H}^n \mid \lambda_1, \lambda_2, \ldots, \lambda_n \geq 0\},$$

and we try applying our method to justify the strong semismoothness of the projection onto $\mathcal{H}_+^n$. First, we can observe that if $M = A + iB$ belongs to $\mathcal{H}_+^n$ if and only if for any complex vector $u + iv \in \mathbb{C}^n$ $(u, v \in \mathbb{R}^n)$, we must have:

$$\begin{aligned} & (u+iv)^\dagger (A+iB)(u+iv) \geq 0 \\ \Leftrightarrow\ & (u^T - iv^T)(A+iB)(u+iv) \geq 0 \\ \Leftrightarrow\ & u^T A u + v^T A v - u^T B v - v^T B^T u \geq 0 \\ \Leftrightarrow\ & \begin{pmatrix} u^T & v^T \end{pmatrix} \begin{bmatrix} A & -B \\ -B^T & A \end{bmatrix} \begin{pmatrix} u \\ v \end{pmatrix} \geq 0. \end{aligned} \tag{2.21}$$

The inequality (2.21) gives us the idea of taking the cone K as $\mathbb{S}_+^{2n}$ and the linear mapping $G : \mathcal{H}^n \to \mathbb{S}^{2n}$ given by:

$$G(A+iB) = \begin{bmatrix} A & -B \\ -B^T & A \end{bmatrix}.$$

However, now we will bring out a counter example to show that under this kind of choosing a cone K and a mapping G, our method does not work. More exactly, for that specific case, the nondegeneracy condition (equivalent form)

$$(\nabla G(A+iB)\mathcal{H}^n)^\perp \cap \Big(\mathrm{lin}(T_{\mathbb{S}_+^{2n}}(G(A+iB)))\Big)^\perp = \{O\} \tag{2.22}$$

does not hold.

We consider the space $\mathcal{H}^2$. The matrix $M = A + iB \in \mathcal{H}^2$ in this case will simply be the real symmetric matrix, with $B = 0$ and A given by,

$$A = P \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} P^T,$$

where,

$$P = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}.$$

We consider $H \in (\nabla G(A+iB)\mathcal{H}^2)^{\perp}$. At first, since G is a linear mapping, then the fact that $H \in (\nabla G(A+iB)\mathcal{H}^2)^{\perp}$ is actually the same with $H \in (G(\mathcal{H}^2))^{\perp}$. That means, for any matrices $C \in \mathbb{S}^n$ and 2×2 real skew-symmetric matrix D, we always have:

$$\begin{aligned} &\langle H, G(C+iD)\rangle = 0 \\ \Leftrightarrow\ &\operatorname{tr}\left(H \begin{bmatrix} C & -D \\ -D^T & C \end{bmatrix} \right) = 0. \end{aligned}$$

From this inequality, it is not difficult for us to verify the fact that $H \in (\nabla G(A+iB)\mathcal{H}^2)^{\perp}$ if and only if it is of the following form

$$H = \begin{bmatrix} O & S \\ S & O \end{bmatrix} \tag{2.23}$$

where O is the 2×2 zero matrix and S is some 2×2 real symmetric matrix.

Now we turn our attention to the tangent cone $T_{\mathbb{S}^4_+}(G(A+iB))$ and its corresponding lineality space. With our choice of A and B, $G(A+iB)$ has the form $G(A+iB) = Q\,\mathrm{Diag}(1,1,0,0)\,Q^T$, where,

$$Q = \begin{bmatrix} \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} & 0 \\ -\frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} & 0 \\ 0 & \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \\ 0 & -\frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \end{bmatrix},$$

which is an orthogonal matrix. Actually, Q is easily built up from the matrix P in the spectral decomposition of A.

We take the notice that the following matrix

$$W = \begin{bmatrix} \frac{1}{\sqrt{2}} & 0 \\ \frac{1}{\sqrt{2}} & 0 \\ 0 & \frac{1}{\sqrt{2}} \\ 0 & \frac{1}{\sqrt{2}} \end{bmatrix},$$

is the one containing two columns of Q which correspond to the eigevalue 0 in the eigenvalue decomposition of $G(A + iB)$. Therefore, the tangent cone $T_{\mathbb{S}^4_+}(G(A + iB))$ is determined by

$$T_{\mathbb{S}^4_+}(G(A + iB)) = \{M \in \mathbb{S}^4 \mid W^T M W \succeq O\}.$$

And certainly, its lineality space has the form

$$\mathrm{lin}(T_{\mathbb{S}^4_+}(G(A + iB))) = \{M \in \mathbb{S}^4 \mid W^T M W = O\}. \tag{2.24}$$

We let a_{ij}, $i, j = 1, \ldots, 4$ denote the entries of a matrix M, with a notice that $a_{ij} = a_{ji}$ for every $i, j = 1, \ldots, 4$. Then after a several basic computations, we see that the equality (2.24) is equivalent to the fact that

$$\begin{aligned} &a_{11} + 2a_{12} + a_{22} = 0, \\ &a_{33} + 2a_{34} + a_{44} = 0, \\ &a_{13} + a_{23} + a_{14} + a_{24} = 0. \end{aligned} \tag{2.25}$$

Next, we only need to take the matrix H_0, where

$$H_0 = \begin{bmatrix} 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \end{bmatrix},$$

then we can easily compute the inner product between H and any matrix $M = [a_{ij}]_{4\times 4} \in \operatorname{lin}(T_{\mathbb{S}^4_+}(G(A+iB)))$ to get

$$\langle H_0, M\rangle = 2(a_{13} + a_{23} + a_{14} + a_{24}) = 0.$$

By this result, this matrix H_0 belongs to $\Big(\operatorname{lin}(T_{\mathbb{S}^4_+}(G(A+iB)))\Big)^{\perp}$. At the same time, by (2.23), H_0 also belongs to $(\nabla G(A+iB)\mathcal{H}^2)^{\perp}$, consequently, we can conclude that

$$\big(\nabla G(A+iB)\mathcal{H}^2\big)^{\perp} \cap \Big(\operatorname{lin}(T_{\mathbb{S}^4_+}(G(A+iB)))\Big)^{\perp} \neq \{O\}.$$

Therefore, the nondegeneracy does not hold for all the matrices in $\mathcal{H}^2_+$, and we cannot use our method in this case.

2.4.2 Projection onto the SOC

Now we are back to the *case 3* that remains unconsidered from the usage of the method for the projection onto second order cones. When $x = 0$, according to the result from [3], we have

$$T_{\mathbb{S}^n_+}(L_x) = \mathbb{S}^n_+,$$

and the corresponding lineality space $\operatorname{lin}(T_{\mathbb{S}^n_+}(L_x))$ is certainly $\mathbb{S}^n_+ \cap (-\mathbb{S}^n_+) = \{O\}$. Therefore, we have

$$G(\mathbb{R}\times\mathbb{R}^{n-1}) + \operatorname{lin}(T_{\mathbb{S}^n_+}(L_x)) = G(\mathbb{R}\times\mathbb{R}^{n-1}) \neq \mathbb{S}^n.$$

The reason for the inequality here is that $G(\mathbb{R}\times\mathbb{R}^{n-1})$ contains only arrow-shaped matrices hence cannot cover all the matrices in $\mathbb{S}^n$. Thus, unfortunately, we also cannot apply our method in this case.

Chapter 3

On the superlinear convergence of smoothing Newton continuation algorithm

During this chapter, we will consider one criterion for the superlinear convergence of the Algorithm 4.2 of [9]. This algorithm is initially stated in [17] (Algorithm 2).

3.1 Stating the problem

We consider the finite-dimensional vector space $\mathbb{E}$ which has $\langle \cdot, \cdot \rangle$ as it inner product. Let $X \subseteq \mathbb{E}$ be closed convex such that $\mathrm{int}(X) \neq \emptyset$. Let Ω be a subset of $\mathbb{E}$ that contains X. Let $F : \Omega \to \mathbb{E}$ be a continuous mapping which is differentiable in $\mathrm{int}(\Omega)$. The variational inequality problem $VI(X, F)$ is the one requiring us to find $x \in X$ satisfying

$$\langle F(x), y - x \rangle \geq 0 \quad \text{for all } y \in X. \tag{3.1}$$

This problem can be solved, using the natural map equation: $G^{\mathrm{nat}}(x, y) = 0$, where G^{nat} is the *natural map* defined by

$$(x, y) \in \Omega \times \mathbb{E} \mapsto (x - \Pi_X(x - y), F(x) - y)^T.$$

The smoothing approximation H^{nat} of G^{nat} is defined by

$$(x, y, \mu, \varepsilon) \in \mathbb{E}^2 \times \mathbb{R}^2 \mapsto (x - p(x - y, \mu), F(x) + \Pi_{\mathbb{R}_+} x - y)^T,$$

where $f : \text{int}(X) \to \mathbb{R}$ is a differentiable barrier function and p the smoothing approximation of Π_X corresponding to f.

From now on, for the finite-dimensional vector spaces $\mathbb{E}$, $\mathbb{E}'$, $\mathbb{E}''$ and a mapping $T : \mathbb{E} \times \mathbb{E}' \to \mathbb{E}''$, we will use the notation $\mathbf{J}F_x(x, y)$ to denote the partial derivative of F with respect to variable x (if it exists). We consider the following algorithm which aims at approximating the solution of $VI(X, F)$.

Algorithm 3.1.1. (Algorithm 4.2 of [9])

Inputs. Initial data $w_0 = (x_0, y_0) \in \mathbb{E} \times \mathbb{E}$ and parameters $\beta > 0$, α, $\eta \in (0, 1)$, $\bar{\eta} \in (0, 1/2)$ and $\kappa > 0$. Set $k = 0$ and $\mu_0 = \varepsilon_0 = \|G^{\text{nat}}(w_0)\|$ and we repeat the steps mentioned below until $\|G^{\text{nat}}(w_k)\| = 0$.

Step 1. Set $j = 0$ and $v_{k0} = w_k$.

Step 1a. Find $d_{kj} \in \mathbb{E}^2$ such that

$$H^{\text{nat}}(v_{kj}, \mu_k, \varepsilon_k) + \text{J}_w H^{\text{nat}}(v_{kj}, \mu_k, \varepsilon_k) d_{kj} = 0.$$

Step 1b. If $\|H^{\text{nat}}(v_{kj} + d_{kj}, \mu_k, \varepsilon_k)\| \leq \beta\eta^k$, set $w_{k+1} = v_{kj} + d_{kj}$ and proceed to Step 2.

Step 1c. Otherwise, find the largest $\lambda_{kj} \in \{1, \alpha, \alpha^2, \ldots\}$ such that

$$\|H^{\text{nat}}(v_{kj} + \lambda_{kj} d_{kj}, \mu_k, \varepsilon_k)\|^2 - \|H^{\text{nat}}(v_{kj} + d_{kj}, \mu_k, \varepsilon_k)\|^2 \leq -2\sigma\lambda_{kj} \|H^{\text{nat}}(v_{kj} + d_{kj}, \mu_k, \varepsilon_k)\|^2$$

and set $v_{k,j+1} = v_{kj} + \lambda_{kj} d_{kj}$.

Step 1d. If $\|H^{\text{nat}}(v_{k,j+1} + d_{kj}, \mu_k, \varepsilon_k)\| \leq \beta\eta^k$, set $w_{k+1} = v_{k,j+1}$ and proceed to Step 2.

Otherwise, update $j = j + 1$ and return to Step 1a.

Step 2. Set $\mu_{k+1} = \min\{\kappa \|G^{\text{nat}}(w_{k+1})\|^2, \mu_0\, \bar{\eta}^{k+1}\}$ and $\varepsilon_{k+1} = \min\{\kappa \|G^{\text{nat}}(w_{k+1})\|^2, \varepsilon_0\, \bar{\eta}^{k+1}\}$ and update $k = k + 1$. $\square$

Theorem 4.3 of [9] proved that if F is locally Lipschitz continuous and monotone, and the solution set of $VI(X, F)$ is non-empty and bounded, then Algorithm 3.1.1 generates a bounded sequence $\{w_k = (x_k, y_k)\}$ with an accumulation point $w^* = (x^*, y^*)$ that is a zero of G^{nat}. This theorem also states the conditions under which the x-component of the sequence $\{x_k\}$ converges superlinearly to the solution x^* of $VI(X, F)$. One of these conditions is that $\{\mathbf{J}_w H^{\text{nat}}(w_{k_l}, \mu_{k_l}, \varepsilon_{k_l})\}$ is uniformly nonsingular for any subsequence $\{w_{k_l}\}$ that converges to w^*. Subsequently, in Theorem 4.4 of that paper, it is pointed out the sufficient condition for the sequence $\{\mathbf{J}_w H^{\text{nat}}(w_{k_l}, \mu_{k_l}, \varepsilon_{k_l})\}$ to be uniformly nonsingular is that for a sequence $\{(x_k, y_k, \mu_k, z_k)\}$ generated by Algorithm 3.1.1 that converges to $\{(x^*, y^*, 0, 0)\}$ and any limit point J of $\{\mathbf{J}_z p(x_k - y_k, \mu_k)\}$, $L^\perp \cap (\text{nullspace}(J) \times \text{nullspace}(I - J)) = 0$, where $L = \{(h, k) \mid \mathbf{J}F(x^*)h = k\}$. Later on, when we let X to be a closed convex cone K of $\mathbb{E}$, it is shown in section 6 of [9] that nondegeneracy[1] implies this sufficient condition, and the cone considered in that situation is the epigraph of the nuclear norm. The key result for the proof is Lemma 6.1 of that paper. Within the context of this thesis, we will state and prove the same result for the cases of PSD cones and symmetric cones.

3.2 Proof for the case of PSD cones

Consider the Lemma 6.1 of [9]. In the paper, the statement of that lemma holds in the case of the cone K being the epigraph of nuclear norm. In this thesis, we will restate the lemma in the case of cone K being the PSD cone, prove it and generalize the proof for the case that K being the symmetric cone $\mathbb{S}^n$.

To avoid any confusion, we will keep the notations from the original paper. Now let K be a PSD cone $\mathbb{S}^n_+$, which is the closed convex one. Notice that since the PSD cone is self-dual, the dual cone $K^\sharp$ of K is K itself, hence, we do not need to state the part of the lemma for the dual cone. The smoothing approximation p of the projection onto PSD cone mentioned in the lemma is defined by the barrier function $f : \mathbb{S}^n_{++} \to \mathbb{R}$ given by $f(X) = -\ln\det(X)$.

Lemma 3.2.1. *(Statement of lemma 6.1 of [9] for the PSD cone)*

If $\{Z_k\}$ is a convergent sequence in $\mathbb{S}^n$ and μ_k is a positive sequence coverging to 0*, then*

[1] Once again, we mention the concept of *nondegeneracy*. However, in this context, a vector $x \in K$ is defined to be nondegenerate for L if it satisfies $L^\perp \cap \text{span}(F^\Delta) = \{0\}$ with F being the smallest face of K containing x and F^Δ being the complementary face of F.

every limit point J of the sequence $\{\mathbf{J}_z p(Z_k,\, \mu_k)\}$ satisfies

$$nullspace(J) \subseteq span(F_{X^*}^{\triangle}),$$

where F_{X^*} is the smallest face of $\mathbb{S}_+^n$ containing the limit point

$$X^* := \lim_{k\to\infty} p(Z_k,\, \mu_k)$$

and $F_{X^*}^{\triangle}$ is its complementary face.

Proof. First, we find the formula of the smoothing approximation $p(Z,\, \mu)$ for the given matrix Z and the positive number μ. We recall that p is defined by the barrier function $f(X) = -\ln\det(X)$. Thus, for $Z \in \mathbb{S}^n$ and $\mu > 0$, we have

$$p(Z,\, \mu) + \mu^2 \nabla f(p(Z, \mu)) = Z. \tag{3.2}$$

By the calculation given in [16], we have $\nabla f(X) = -X^{-1}$ and $\nabla f(X)(H) = -\mathrm{tr}(X^{-1}H)$. Now we get back to the equation (3.2). We make some transformation for this equation based on the $\nabla f(X)$ we calculated above.

$$\begin{aligned} & p(Z,\, \mu) + \mu^2 \nabla f(p(Z)) = Z \\ \Leftrightarrow\ & p(Z,\, \mu) - \mu^2 p(Z,\, \mu)^{-1} = Z. \end{aligned} \tag{3.3}$$

We will consider the spectral decomposition

$$Z = Q^T D Q$$

for $Q \in \mathcal{O}^n(Z)$ and $D = \mathrm{Diag}(\lambda_f(_Z))$.

By the result given in Example 3.2 of [10], the smoothing approximation $p(z, \mu)$ that satisfies equation (3.3) has the form

$$\begin{aligned} p(Z,\, \mu) &= \frac{1}{2} Q^T \lambda_f(Z) Q + \frac{1}{2} Q^T \left(\sqrt{\lambda_1(Z)^2 + 4\mu^2}, \cdots, \sqrt{\lambda_n(Z)^2 + 4\mu^2} \right) Q \\ &= \frac{1}{2} Z + \frac{1}{2} Q^T \left(\sqrt{\lambda_1(Z)^2 + 4\mu^2}, \ldots, \sqrt{\lambda_n(Z)^2 + 4\mu^2} \right) Q. \end{aligned} \tag{3.4}$$

For the next part of the proof, we need the concept of "spectral function" that is defined in [23]. Based on the formula for $p(Z,\, \mu)$ we just found, we consider the real function

$$f_\mu : u \in \mathbb{R} \mapsto f_\mu(u) = u + \sqrt{u^2 + \mu^2},$$

then $f'_\mu(u) = 1 + \dfrac{u}{\sqrt{u^2 + 4\mu^2}}$. Based on the formula given in [10], we now define the spectral function $f_\mu^{[1]} : \mathbb{R}^n \to \mathbb{S}^n$ given by

$$(f_\mu^{[1]}(d))_{ij} = 1 + \frac{d_i + d_j}{\sqrt{d_i^2 + 4\mu^2} + \sqrt{d_j^2 + 4\mu^2}}. \tag{3.5}$$

Then we have the formula $\mathbf{J}_z p(Z,\, \mu)[H] = \frac{1}{2} Q^T \left[f_\mu^{[1]}(\lambda_f(Z)) \circ (QHQ^T) \right] Q$ for $Z, H \in \mathbb{S}^n$ and $Q \in \mathcal{O}^n(Z)$. The symbol $\circ$ stands for the Hadamard product.

Now we turn our attention to the main statement of the lemma. We consider the sequences $\{Z_k\}$ and $\{\mu_k\}$ that we already mentioned in the statement of the lemma. Let $Z^* \in \mathbb{S}^n$ be the limit of $\{Z_k\}$. Consider J being a limit point of the sequence $\mathbf{J}_z p(Z_k,\, \mu_k)$, then there exist a subsequence $\{Z_{k_l}\}$ such that $\mathbf{J}_z p(Z_{k_l},\, \mu_{k_l})$ converges to J. For simplicity, we still use $\{Z_k, \mu_k\}$ in the place of that subsequence, without any loss of generality.

For each k, let $Z_k = Q_k^T\, \lambda_f(Z_k)\, Q_k$ being the spectral decomposition of Z_k. Obviously, $\{\lambda_f(Z_k)\}$ converges to $\lambda_f(Z^*)$. Since $\{Q_k\}$ is the sequence of orthogonal matrices, it is bounded. Thus, it has a converging subsequence $\{Q_{k_m}\}$. Let Q^* be the limit of $\{Q_{k_m}\}$, we then have

$$\begin{aligned} Q_{k_m}^T\, \lambda_f(Z_{k_m})\, Q_{k_m} &\longrightarrow (Q^*)^T\, \lambda_f(Z^*)\, Q^* \\ \Leftrightarrow\;\; Z_{k_m} &\longrightarrow (Q^*)^T\, \lambda_f(Z^*)\, Q^*. \end{aligned}$$

However, since $\{Z_{k_m}\}$ converges to Z^*, we have $Z^* = (Q^*)^T\, \lambda_f(Z^*)\, Q^*$. It implies that $Q^* \in \mathcal{O}^n(Z^*)$. For the following steps of this proof, we can still use $\{Z_k\}$ in the place of $\{Z_{k_m}\}$ without changing its generality. Under that procedure, we see that

$$J[H] = \lim_{k\to\infty} \frac{1}{2} Q_k^T \left[f_{\mu_k}^{[1]}(\lambda_f(Z_k)) \circ (Q_k H Q_k^T) \right] Q_k \qquad \forall H \in \mathbb{S}^n. \tag{3.6}$$

Because $Q_k \longrightarrow Q^*$ and $\lambda_f(Z_k) \longrightarrow \lambda_f(Z^*)$, we see that the right hand side of (3.6) equals to $\frac{1}{2}(Q^*)^T \left[f_0^{[1]}(\lambda_f(Z^*)) \circ (Q^* H (Q^*)^T)\right] Q^*$ and so does $J[H]$.

We now consider $X^* = \lim_{k\to\infty} p(Z_k, \mu_k)$. According to (3.4) we see that for each k, $p(Z_k, \mu_k)$ has the eigenvalues being

$$\frac{1}{2}\left(\lambda_1(Z_k) + \sqrt{\lambda_1(Z_k)^2 + 4\mu_k^2}\right), \ldots, \frac{1}{2}\left(\lambda_n(Z_k) + \sqrt{\lambda_n(Z_k)^2 + 4\mu_k^2}\right).$$

Hence, as k tends to infinity, we see that the eigenvalues of X^* will be $\frac{1}{2}\left(\lambda_1(Z^*) + |\lambda_1(Z^*)|\right), \ldots$, $\frac{1}{2}\left(\lambda_n(Z^*) + |\lambda_n(Z^*|\right)$. Since $\lambda_1(Z^*) \geq \cdots \geq \lambda_n(Z_*)$, we easily see that X^* has the spectral decomposition

$$X^* = (Q^*)^T \operatorname{Diag}\left(\lambda_1(X^*), \ldots, \lambda_m(X^*), 0, \ldots, 0\right) Q^*$$

with $\lambda_i(X^*) = \lambda_i(Z^*)$ for $i = 1, 2, \ldots, m$. Here $\lambda_1(Z^*) \geq \cdots \geq \lambda_m(Z^*)$ are the first m eigenvalues of Z^* which are positive while the remaining $n-m$ ones are non-positive. This spectral decomposition of X^* actually shows that X^* is $\Pi_{\mathbb{S}_+^n}(Z^*)$. This result agrees with the fact that X^* is the limit of the sequence $\{p(Z_k, \mu_k)\}$ of smoothing approximations.

Now, according to formula (169) from [11], we see that the smallest face of X^* in $\mathbb{S}_+^n$ has the form

$$\begin{aligned} F_{X^*} &= \left\{Y \in \mathbb{S}_+^n \mid \langle (Q^*(I - \Lambda\Lambda^\ddagger)\,(Q^*)^T, Y\rangle = 0\right\} \\ &= \left\{Y \in \mathbb{S}_+^n \mid \operatorname{tr}\left((Q^* Y (Q^*)^T)(I - \Lambda\Lambda^\ddagger)\right) = 0\right\} \end{aligned} \tag{3.7}$$

with Λ being the diagonal matrix $\operatorname{Diag}(\lambda_1(X^*), \ldots, \lambda_m(X^*), 0, \ldots, 0)$ in the spectral decomposition of X^* and $\Lambda^\ddagger$ stands for the *Moore-Penrose pseudoinverse matrix* of Λ. We see that

$$I - \Lambda\Lambda^\ddagger = \begin{pmatrix} 0 & & & \\ & \ddots & & \\ & & 0 & \\ & & & I_{n-m} \end{pmatrix}.$$

We recall the result from Lemma A.2 of [18] that two PSD matrices W_1 and W_2 satisfy $\operatorname{tr}(W_1, W_2) = 0$ if and only if their matrix product satisfies $W_1 W_2 = O$. We combine that fact with (3.7) and the formula of $I - \Lambda\Lambda^\ddagger$ to see that

$$Y \in F_{X^*} \Leftrightarrow Y = \left(\begin{array}{c|c} B & A \\ \hline A & O \end{array}\right), \qquad B \in \mathbb{S}^m_+, A \in \mathbb{R}^{m\times(n-m)}.$$

However, we notice that, for a PSD matrix X, if one of its entries on the main diagonal X_{ii} equals 0, then all of the corresponding entries on the i-th row and i-th column all equal 0. This fact is easily seen by using the Cholesky factor of X. When $X_{ii} = 0$, then by the simple computation, we also have $(L_X)_{ii} = 0$. By the property of the Cholesky factor given in [7], all the entries in the i-th column of L_X are equal 0. Consequently, by the direct computation of the entries in the i-th column and i-th row of X via $X = L_X L_X^T$, we are able to confirm the above statement that $X_{ik} = X_{ki} = 0$ for every $k = 1, \cdots, n$. With the notice we just give, all the matrix Y belongs to F_{X^*} if and only if

$$Q^* Y (Q^*)^T = \left(\begin{array}{c|c} B & O \\ \hline O & O \end{array}\right).$$

We then have F_{X_*} to be:

$$F_{X_*} = \left\{ Y \in \mathbb{S}^m : Y = (Q^*)^T \left(\begin{array}{c|c} B & O \\ \hline O & O \end{array}\right) Q^*, \qquad B \in \mathbb{S}^m_+ \right\}.$$

Thus, the complementary face $F^\triangle_{X^*}$ has the form

$$\begin{aligned} F^\triangle_{X^*} &= \left\{ Y \in \mathbb{S}^n_+ \ : \left\langle Y, (Q^*)^T \left(\begin{array}{c|c} B & O \\ \hline O & O \end{array}\right) Q^* \right\rangle = 0, \qquad \forall B \in \mathbb{S}^m_+ \right\} \\ &= \left\{ Y \in \mathbb{S}^n_+ \ : \left\langle Q^* Y (Q^*)^T, \left(\begin{array}{c|c} B & O \\ \hline O & O \end{array}\right) \right\rangle = 0, \qquad \forall B \in \mathbb{S}^m_+ \right\}, \end{aligned}$$

or equivalently, using the Lemma A.2 of [18] once again, we have:

$$F^\triangle_{X^*} = \left\{ Y \in \mathbb{S}^n_+ \ : Q^* Y (Q^*)^T = \left(\begin{array}{c|c} O & O \\ \hline O & C \end{array}\right), \qquad C \in \mathbb{S}^{n-m}_+ \right\}.$$

Consequently, with the notice that the spanning space of $\mathbb{S}^{n-m}_+$ is $\mathbb{S}^{n-m}$, the spanning space of $F^{\triangle}_{X^*}$ has the following form

$$\mathrm{span}(F^{\triangle}_{X^*}) = \left\{ Y \in \mathbb{S}^n \ : Q^* Y (Q^*)^T = \left(\begin{array}{c|c} O & O \\ \hline O & C \end{array} \right), \qquad C \in \mathbb{S}^{n-m} \right\}. \tag{3.8}$$

On the other hand, assume that $H \in \mathrm{nullspace}(J)$. As $J[H]$ is already calculated above, we have

$$J[H] = \frac{1}{2}(Q^*)^T \left[f_0^{[1]}(\lambda_f(Z^*)) \circ (Q^* H (Q^*)^T) \right] Q^* = 0. \tag{3.9}$$

From (3.5), for each k, we get

$$(f_{\mu_k}^{[1]}(\lambda_f(Z_k)))_{ij} = 1 + \frac{\lambda(Z_k)_i + \lambda(Z_k)_j}{\sqrt{\lambda(Z_k)_i^2 + 4\mu_k^2} + \sqrt{\lambda(Z_k)_j^2 + 4\mu_k^2}} \ ,$$

and when $k \to \infty$ we have

$$(f_0^{[1]}(\lambda_f(Z^*)))_{ij} = 1 + \frac{\lambda_i^* + \lambda_j^*}{|\lambda_i^*| + |\lambda_j^*|} \ . \tag{3.10}$$

Here, for simplicity in notation, we use λ_i^* in the place of $\lambda(Z^*)_i$. Recall that the first m eigenvalues of Z^* are positive. Clearly, for every couple of indexes $1 \le i, j \le m$, from (3.10) we see that

$$(f_0^{[1]}(\lambda_f(Z^*)))_{ij} = 1 + \frac{\lambda_i^* + \lambda_j^*}{\lambda_i^* + \lambda_j^*} > 0 \ .$$

For every couple of indexes $1 \le i \le m < j$, we have

$$(f_0^{[1]}(\lambda_f(Z^*)))_{ij} = 1 + \frac{\lambda_i^* + \lambda_j^*}{\lambda_i^* - \lambda_j^*} > 0 \ \text{ with } \lambda_i^* > 0, \lambda_j^* \le 0.$$

By the same argument, we also see that $(f_0^{[1]}(\lambda_f(Z^*)))_{ij} > 0$ when $1 \le j \le m < i$ and therefore $(f_0^{[1]}(\lambda_f(Z^*)))_{ij} > 0$ when either $i \le m$ or $j \le m$.

For $H \in \mathrm{nullspace}(J)$, from (3.9), we have

$$\begin{aligned} (f_0^{[1]}(\lambda_f(Z^*)))_{ij} \left(Q^* H (Q^*)^T \right)_{ij} &= 0 \qquad \forall\, i, j \text{ such that } 1 \le \min\{i,j\} \le m \\ \Rightarrow \ \left(Q^* H (Q^*)^T \right)_{ij} &= 0 \qquad \forall\, i, j \text{ such that } 1 \le \min\{i,j\} \le m. \end{aligned} \tag{3.11}$$

From (3.11), we see that the matrix $Q^*H(Q^*)^T$ has the form

$$Q^*H(Q^*)^T = \left(\begin{array}{c|c} O & O \\ \hline O & \hat{H} \end{array}\right), \tag{3.12}$$

with $\hat{H} \in \mathbb{S}^{n-m}$.

We combine (3.12) and (3.8) to see that $H \in \mathrm{span}(F_{X^*}^{\triangle})$ for every $H \in \mathrm{nullspace}(J)$. Consequently, $\mathrm{nullspace}(J) \subseteq \mathrm{span}(F_{X^*}^{\triangle})$ and we complete our proof. $\square$

3.3 Proof for the case of symmetric cones

Now we will state and prove the similar lemma for the case that K is the symmetric cone, which is the generalized case of the one in the previous part on PSD cones. First, we need the barrier function $f : \Omega \to \mathbb{R}$ given by $f(x) = -\ln \det x := -\ln(\det x)$. The smoothing approximation of the Euclidean projector that is generated from f is still denoted by $p(z, \mu)$ for $z \in V$ and $\mu \in \mathbb{R}$. Now we give the lemma, which is actually the generalized one for the case of PSD cone, because the PSD cone is a subcase of a symmetric cone.

Lemma 3.3.1. *Let V be an Eulcidean Jordan algebra of rank r. K is the set of all squares in V and $\Omega = \mathrm{Int}(K)$ is the corresponding symmetric cone. If $\{z_k\}$ is a convergent sequence in V and μ_k is a positive sequence coverging to 0, then every limit point J of the sequence $\{\mathbf{J}_z p(z_k, \mu_k)\}$ satisfies*

$$nullspace(J) \subseteq span(F_{x^*}^{\triangle}),$$

where F_{x^*} is the smallest face of K containing the limit point

$$x^* := \lim_{k\to\infty} p(Z_k, \mu_k)$$

and $F_{x^*}^{\triangle}$ is its complementary face.

Proof. By the result given in Example 3.3 of [10], we get the formula for ∇f as

$\nabla f :\ x \mapsto -x^{-1}$ for every $x \in \Omega$ and the formula for for p as

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

where we use $\sqrt{z^2 + 4\mu^2 e}$ to denote the element of V which has the spectral decomposition as

$$\sqrt{z^2 + 4\mu^2 e} = \sum_{j=1}^{r} \sqrt{\lambda(z)_j^2 + 4\mu^2}\, c_j.$$

The following concepts are taken from [23]. For $\phi : \mathbb{R} \to \mathbb{R}$ be a scalar valued function, we define a vector valued function ϕ_V associated with V. For $z \in V$ which has the spectral decomposition $z = \sum_{j=1}^{r} \lambda(z)_j c_j$, we have

$$\phi_V(z) = \sum_{j=1}^{r} \phi(\lambda(z)_j) c_j.$$

ϕ_V is called the Löwner's operator. When $\phi(t) = t_+ := \max(t, 0)$, $t \in \mathbb{R}$, Löwner's operator becomes the metric projection operator

$$\phi_V(z) = z_+ = (\lambda(z)_1)_+ c_1 + \cdots + (\lambda(z)_r)_+ c_r.$$

Next, we consider the scalar function

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

which is corresponding to the formula of $p(z,\, \mu)$ that we just found. Similar to the proof for the case of PSD cone, we also define the map $\phi_\mu^{(1)} : \mathbb{R}^r \to \mathbb{S}^r$ given by

$$
\begin{aligned}
(\phi_\mu^{(1)}(\tau))_{ij} &= \begin{cases} \dfrac{\phi_\mu(\tau_i) - \phi_\mu(\tau_j)}{\tau_i - \tau_j} & \text{if} \quad \tau_i \neq \tau_j \\ \phi'_\mu(\tau_i) & \text{if} \quad \tau_i = \tau_j \end{cases} \\
&= \begin{cases} 1 + \dfrac{\sqrt{\tau_i^2 + 4\mu^2} - \sqrt{\tau_j^2 + 4\mu^2}}{\tau_i - \tau_j} & \text{if} \quad \tau_i \neq \tau_j \\ 1 + \dfrac{\tau_i}{\sqrt{\tau_i^2 + 4\mu^2}} & \text{if} \quad \tau_i = \tau_j \end{cases} \\
&= 1 + \frac{\tau_i + \tau_j}{\sqrt{\tau_i^2 + 4\mu^2} + \sqrt{\tau_j^2 + 4\mu^2}}.
\end{aligned} \tag{3.13}
$$

By the formula (33) from [23], any $h \in V$ is able to be represented as the Pierce decomposition

$$
h = \sum_{j=1}^{r} \langle c_j, h \rangle c_j + \sum_{1 \le j < l \le r} 4 c_j (c_l h)
$$

and also from Lemma 3.1 of [23], we have

$$
((\phi_\mu)_V)'(z)h = \sum_{j=1}^{r} (\phi_\mu^{(1)}(\lambda(z)))_{jj} \langle c_j, h \rangle c_j + \sum_{1 \le j < l \le r} 4(\phi_\mu^{(1)}(\lambda(z)))_{jl} c_j (c_l h) \tag{3.14}
$$

with $\lambda(z) = (\lambda_1(z), \cdots, \lambda_r(z))$ and $\lambda_1(z) \ge \cdots \ge \lambda_r(z)$ are the eigenvalues of z.
Now we get back to the context of the lemma. For each k, we have $z_k = \sum_{j=1}^{r} \lambda(z_k)_j c_j^k$ being the spectral decomposition of z_k. For the limit z^* of $\{z_k\}$, we have the spectral decomposition $z^* = \sum_{j=1}^{r} \lambda(z^*)_j c_j$. It is obvious that $\lambda(z_k)_j \to \lambda(z^*)_j$ for each $j = 1, \ldots, r$. Besides, similar to the PSD cone case, without loss of generality, using the subsequence if necessary, we can assume that $c_j^k \to c_j$ for each j. We take $m \le r$ such that $\lambda(z^*)_1 \ge \lambda(z^*)_2 \ge \cdots \ge \lambda(z^*)_m > 0$ and $0 \ge \lambda(z^*)_{m+1} \ge \cdots \ge \lambda(z^*)_r$.
Next, for x^* taken from the lemma, by the same argument with the case of PSD cone, we easily find that

$$
\begin{aligned}
x^* = z^*_+ &= \sum_{j=1}^{r} (\lambda(z^*)_j)_+ c_j \\
&= \sum_{j=1}^{m} \lambda(z^*)_j c_j.
\end{aligned}
$$

We can easily see that $c_1, \dots, c_r$ is a Jordan frame.

Now we consider the following cases.

- *Case 1:* $x^* \in \bar{\Omega} \setminus \operatorname{Int} \bar{\Omega}$.

According to Theorem 2 of [14], the face F_{x^*} is

$$F_{x^*} = \bar{\Omega} \cap V(1, c_1 + \cdots + c_m),$$

so its complementary face is

$$F_{x^*}^{\triangle} = \{y \in \bar{\Omega} \mid \langle y, z \rangle = 0 \ \ \forall z \in \bar{\Omega} \cap V(1, c_1 + \cdots + c_m)\}.$$

And we then imply that

$$F_{x^*}^{\triangle} \supseteq \{y \in \bar{\Omega} \mid \langle y, z \rangle = 0 \ \ \forall z \in V(1, c_1 + \cdots + c_m)\}. \tag{3.15}$$

Therefore, from (3.15), we may imply that

$$\operatorname{span}(F_{x^*}^{\triangle}) \supseteq \operatorname{span}(\{y \in \bar{\Omega} \mid \langle y, z \rangle = 0 \ \ \forall z \in V(1, c_1 + \cdots + c_m)\}). \tag{3.16}$$

As being pointed out in Section 1.3, we know that V has the following orthogonal decomposition corresponding to the Jordan frame $c_1, \dots, c_r$

$$V = \bigoplus_{1 \le i \le j \le r} V_{ij}$$

where $V_{jj} = V(1, c_j) = \mathbb{R}c_j$ and $V_{ij} = V(\frac{1}{2}, c_i) \cap V(\frac{1}{2}, c_j)$.

Now we consider one arbitrary limit point J of the sequence $\{\mathbf{J}_z p(z_k,\, \mu_k)\}$. Using the subsequence if necessary, for $h \in V$ we may consider $J(h)$ as the the following limit

$$\begin{aligned}
J(h) &= \lim_{k\to\infty} \mathbf{J}_z p(z_k,\ \mu_k)(h) \\
&= \lim_{k\to\infty} ((\phi_{\mu_k})_V)'(z_k)h \\
&= \lim_{k\to\infty} \sum_{j=1}^{r} (\phi_{\mu_k}^{(1)}(\lambda(z_k)))_{jj} \langle c_j^k, h\rangle c_j^k + \lim_{k\to\infty} \sum_{1\le j<l\le r} 4(\phi_{\mu_k}^{(1)}(\lambda(z_k)))_{jl} c_j^k (c_l^k h) \\
&= \sum_{j=1}^{r} (\phi_0^{(1)}(\lambda(z^*)))_{jj} \langle c_j, h\rangle c_j + \sum_{1\le j<l\le r} 4(\phi_0^{(1)}(\lambda(z^*)))_{jl} c_j (c_l h). \qquad (3.17)
\end{aligned}$$

The equality in (3.17) gives us the representation of $J(h)$ as the direct sum of elements in the subspaces V_{ij} for $1 \le i \le j \le r$. Recall that this representation is called the *Pierce decomposition* of $J(h)$. From formula (3.17), we see that for any $h \in$ nullspace(J) and each j such that $1 \le j \le m$, the projection of $J(h)$ into the subspaces V_{jj} must equal zero. Recall that we let $\Pi_{V_{jj}}$ denote that projection, then

$$\Pi_{V_{jj}}(J(h)) = (\phi_0^{(1)}(\lambda(z^*)))_{jj} \langle c_j, h\rangle c_j = 0. \qquad (3.18)$$

Therefore, for each j such that $1 \le j \le m$, we have

$$\begin{aligned}
\phi_0^{(1)}(\lambda(z^*)))_{jj} &= \lim_{k\to\infty} \phi_{\mu_k}^{(1)}(\lambda(z_k)))_{jj} \\
&= \lim_{k\to\infty} \left(1 + \frac{(\lambda(z_k)_j}{\sqrt{(\lambda(z_k)_j^2 + 4\mu_k^2}} \right) = 2 > 0, \qquad (3.19)
\end{aligned}$$

here we already used formula (3.13) for the second equality. From (3.18) and (3.19) we imply that $\langle c_j, h\rangle = 0$ for every $j = 1, \ldots, m$.

On the other hand, for each couple $i < j$ such that $i \le m$, we have

$$\Pi_{V_{ij}}(J(h)) = \phi_0^{(1)}(\lambda(z^*)))_{ij} c_i (c_j h) = 0. \qquad (3.20)$$

And we also see that

$$\begin{aligned}\phi_0^{(1)}(\lambda(z^*)))_{ij} &= \lim_{k\to\infty}\phi_{\mu_k}^{(1)}(\lambda(z_k)))_{ij}\\ &= \lim_{k\to\infty}\left(1+\frac{\lambda(z_k)_i+\lambda(z_k)_j}{\sqrt{\lambda(z_k)_i^2+4\mu_k^2}+\sqrt{\lambda(z_k)_j^2+4\mu_k^2}}\right)\\ &= 1+\frac{\lambda(z^*)_i+\lambda(z^*)_j}{|\lambda(z^*)_i|+|\lambda(z^*)_j|}>0. \end{aligned} \tag{3.21}$$

Where we get the last inequality in (3.21) by the same argument we did in the proof for the case of PSD cone. From (3.20) and (3.21), we see that $c_i(c_jh)=0$ for every couple $i<j$ such that $i\le m$.

Now, based on the Pierce decomposition of V, we have the following Pierce decomposition of h:

$$h=\sum_{1\le i\le j}^{r}h_{ij}\qquad \text{for } h_{ij}\in V_{ij}.$$

We observe that $h_{ij}=\Pi_{V_{ij}}(h)$. For each couple $i<j$ such that $i\le m$, according to Theorem IV.2.1 of [13], we have $\Pi_{V_{ij}}=4L(c_i)L(c_j)$. Therefore, we see that

$$h_{ij}=4L(c_i)L(c_j)(h)=4c_i(c_jh)=0.$$

Now, for each j such that $1\le j\le m$, since $V_{jj}=\mathbb{R}c_j$, we let $h_{jj}=\alpha_{jj}c_j$ where α_{jj} is some scalar. From Theorem IV.2.1 of [13], if $\{i,s\}\cap\{k,l\}=\emptyset$, then $V_{is}\cdot V_{kl}=\{0\}$. Therefore, for every couple k,l such that $k\ne j$ and $l\ne j$, we have

$$\langle c_j,h_{kl}\rangle=\langle c_jc_j,h_{kl}\rangle=\langle c_j,c_jh_{kl}\rangle=0,$$

since $c_j\in V_{jj}$ and $h_{kl}\in V_{kl}$. We combine that result with the fact that $h_{jk}=0$ for every $k\ne j$, we can transform $\langle c_j,h\rangle=0$ as

$$\begin{aligned}&\langle c_j,\sum_{1\le i\le j}^{r}h_{ij}\rangle=\sum_{1\le i\le j}^{r}\langle c_j,h_{ij}\rangle=0\\ \Leftrightarrow\ &\langle c_j,h_{jj}\rangle=\langle c_j,\alpha_{jj}c_j\rangle=0\\ \Leftrightarrow\ &\alpha_{jj}=0.\end{aligned}$$

Consequently, $h_{jj} = 0$ for each for each j such that $1 \le j \le m$. The Pierce decomposition of h then becomes

$$h = \sum_{m<i\le k}^{r} h_{ik} \qquad \text{for } h_{ik} \in V_{ik}.$$

Now we will show that $h \in V(0, c_1 + \cdots + c_m)$. For each $j = 1, \ldots, m$ we have $c_j \in V_{jj}$ and each i, k such that $m < i \le k$, we have $h_{ik} \in V_{ik}$. Since $\{j, j\} \cap \{i, k\} = \emptyset$ and so $V_{jj} \cup V_{ik} = 0$, we see that

$$c_j h_{ik} = 0 \qquad \forall\, j \le m,\ \forall\, m < i \le k.$$

Therefore,

$$c_j h = c_j \sum_{m<i\le k}^{r} h_{ik} = \sum_{m<i\le k}^{r} c_j h_{ik} = 0 \qquad \forall\, j \le m.$$

Consequently,

$$L(c_1 + \cdots + c_m)h = (c_1 + \cdots + c_m)h = c_1 h + \cdots + c_m h = 0.$$

Thus, we have $h \in V(0, c_1 + \cdots + c_m)$. Because $c_1 + \cdots + c_m$ is an idempotent of V, as it is pointed out in [14], $V(0, c_1 + \cdots + c_m)$ is a subalgebra of V and therefore an Euclidean Jordan algebra itself. For that reason, we can represent h_{ik} as a linear combination of the idempotents in $V(0, c_1 + \cdots + c_m)$. Assume that $\tilde{c}$ is one of those idempotents. Let z be one arbitrary element in $V(1, c_1 + \cdots + c_m)$, we see that

$$\begin{aligned}
\langle z, \tilde{c} \rangle &= \langle (c_1 + \cdots + c_m) z, \tilde{c} \rangle \\
&= \langle z, (c_1 + \cdots + c_m)\tilde{c} \rangle \\
&= \langle z, 0 \rangle = 0.
\end{aligned}$$

And we also notice that $\tilde{c} \in \bar{\Omega}$. All in all, we finally see that h can be represented as the linear combination of the elements in $\bar{\Omega}$, each of them is orthogonal to every z in $V(1, c_1 + \cdots + c_m)$. By that argument and (3.16), we imply that h belongs to $\text{span} F_{x^*}^{\triangle}$. That helps us observe that nullspace$(J) \subseteq \text{span}(F_{x^*}^{\triangle})$.

• *Case 2:* $x^* \in \operatorname{Int} \bar{\Omega}$.

Assume that a convex subset $T \subseteq \bar{\Omega}$ is the face of $\bar{\Omega}$ which contains x^*. For any $z \in \bar{\Omega}$, since $x^* \in \operatorname{Int} \bar{\Omega}$ there exist a scalar $\beta > 0$ such that $x^* + \beta z \in \bar{\Omega}$. Sice $\bar{\Omega}$ is a convex cone and $\dfrac{1}{1+\beta} > 0$, we have:

$$\begin{aligned} & \frac{1}{1+\beta}(x^* + \beta z) \in \bar{\Omega} \\ \Leftrightarrow \quad & \frac{1}{1+\beta} x^* + \frac{\beta}{1+\beta} z \in \bar{\Omega}. \end{aligned}$$

The above relation and the fact that $\dfrac{1}{1+\beta} + \dfrac{\beta}{1+\beta} = 1$ imply that $z \in T$, since T is a face of $\bar{\Omega}$. Then we see that $z \in T$ for any $z \in \bar{\Omega}$, hence $T = \bar{\Omega}$ for any face T containing x^*. Since F_{x^*} is the intersection of all the faces of $\bar{\Omega}$ containing x^*, $F_{x^*} = \bar{\Omega}$. Thus, its complementary face is

$$F_{x^*}^{\triangle} = \{y \in \bar{\Omega} : \langle y, z \rangle = 0 \quad \forall z \in \bar{\Omega}\} = \{0\},$$

and certainly, $\operatorname{span}(F_{x^*}^{\triangle}) = \{0\}$.

Now we consider an arbitrary element h in the null space of J. According to the results from Chapter III of [13], the interior Ω of $\bar{\Omega}$ is the set of elements with positive eigenvalues. Since $x^* \in \Omega$, all of its eigenvalues $(\lambda(z^*)_1)_+, \dots, (\lambda(z^*)_r)_+$ are positive, or $\lambda(z^*)_1, \dots, \lambda(z^*)_r$ are all positive. Hence, for every i, j such that $1 \le i \le j \le r$, by the same calculation that we did in Case 1, we see that

$$\begin{aligned} \phi_0^{(1)}(\lambda(z^*)))_{ij} &= \lim_{k\to\infty} \phi_{\mu_k}^{(1)}(\lambda(z_k)))_{ij} \\ &= 1 + \frac{\lambda(z^*)_i + \lambda(z^*)_j}{\lambda(z^*)_i + \lambda(z^*)_j} = 2 > 0. \end{aligned}$$

And by the same argument that we did in Case 1, we imply that $\langle c_j, h \rangle = 0$ for $j = 1, \cdots, r$ and for any $1 \le i < j \le r$, we have $c_i(c_j h) = 0$. Thus, when we represent h in the form

$$h = \sum_{1 \le i \le j}^{r} h_{ij} \qquad \text{for } h_{ij} \in V_{ij},$$

we can repeat the argument in Case 1 and show that $h_{ij} = 0$ for all $1 \le i < j \le r$, and hence h is zero element and obviously belongs to $\mathrm{span}(F_{x^*}^{\triangle})$. It leads to the fact that $\mathrm{nullspace}(J) \subseteq \mathrm{span}(F_{x^*}^{\triangle})$.

To sum up, the statement of the lemma holds for both cases that we consider, and we complete our proof. $\square$

Chapter 4

Conclusion and future work

Throughout Chapter 2, we already consider the usage of one method to justify the strong semismoothness of projections onto some subclasses of the class of homogeneous cones. The method helps us prove that the projection onto the SOC is actually strongly semismooth in several cases. However, in the case that $x(y) = 0$, which corresponds to the fact that y is in the polar cone of $\mathcal{K}^n$, with the choice of K and G we presented, the nondegeneracy does not hold. The same result happens when we consider the cone of positive semidefinite Hermitian matrices. These facts reflect the importance of the **Step 1** and **Step 2** of the method, which is how we choose the cone K and an appropriate affine mapping G. The good choice may help us prove the most important thing: the nondegeneracy property.

Indeed, the second order cones and the cone of positive semi-definite Hermitian matrices are the subcases of the category of symmetric cones. The projection onto symmetric cones is already proved to be strongly semismooth. What we did here is not something new. However, we got the initial insight into how we proceed one of the possible methods to consider the strong semismoothness of the projection onto homogeneous cone and the difficulties that we may encounter. We hope that in the future, this method would help to solve the problem.

In the paper [17], the Algorithm 2 was used to solve the second order cone complementarity problem (SOCCP). Chapter 3 of the thesis focuses on proving a lemma that allows us to build up the superlinear convergence for this Algorithm. However, the cone K in the lemma that we focus on in the context of this thesis is a symmetric cone, where the smoothing approximation can be calculated explicitly. Therefore, we will encounter many obstacles when generalize the statement of the lemma for bigger classes, like homogeneous cones and even hyperbolic cones, since we are not able to find the exact formula for the smoothing approximation. One possible direction that we may follow in the future is to utilize the Theorem 12.32 in [22] about the graphical convergence of monotone mappings to consider the convergence of the sequence of Jacobians of the smoothing approximation without knowing the formula of the smoothing approximation.

Bibliography

[1] F. ALIZADEH AND D. GOLDFARB, *Second-order cone programming*, Mathematical Programming, 95 (2003), No. 1, pp. 3-51.

[2] J. BOLTE, A. DANILIDIS AND A. LEWIS, *Tame functions are semismooth*, Mathematical Programming, 117 (2009), pp. 5-19.

[3] J.F. BONNANS AND H. RAMÍRES C., *Pertubation analysis of second-order cone programming problems*, Mathematical Programming, 104 (2005), pp. 205-227.

[4] J.-S. CHEN, X. CHEN AND P. TSENG, *A survey on SOC complementarity functions and solution methods for SOCPs and SOCCPs*, Pacific Journal of Optimization, 8 (2012), pp. 33-74.

[5] J.-S. CHEN AND S.-H. PAN, *Analysis of nonsmooth vector-valued functions associated with second-order cone*, Mathematical Programming, 101 (2004), pp. 95-117.

[6] J.-S. CHEN AND S.-H. PAN, *A survey on SOC complementarity functions and solution methods for SOCPs and SOCCPs*, Pacific Journal of Optimization, 8 (2012), no.1, pp. 33-74.

[7] C.B. CHUA, *A New Notion of Weighted Centers for Semidefinite Progaramming*, SIAM J. Optim, 16 (2006), pp. 1092-1109.

[8] C.B. CHUA, *Relating homogeneous cone and positive definite cones via T-algebra*, SIAM J. Optim, 14 (2003), pp. 500-506.

[9] C.B. CHUA AND L.T.K. HIEN, *A Superlinearly Convergent Smoothing Newton Continuation Algorithm for Variational Inequalities over Definable Sets*, SIAM J. Optim, 25 (2015), pp. 1034-1063.

[10] C.B. Chua and Z. Li, *A barrier-based smoothing proximal point algorithm for NCPs over closed convex cones*, SIAM J. Optim, 23 (2013), no.2, pp. 745-769.

[11] J. Dattoro, *Convex Optimization & Euclidean Distance Geometry* , Meboo Publishing USA, Version 05.04.2015.

[12] F. Facchinei and J.-S. Pang, *Finite-Dimensional Variational Inequalities and Complementarity Problems*, Springer Series in Operations Research (2003).

[13] J. Faraut and A. Korányi, *Analysis on Symmetric Cones*, Oxford Mathematical Monographs (1994).

[14] L. Faybusovich, *Jordan-Algebraic Approach to Convexity Theorems for Quadratic Mappings*, SIAM J. Optim, 17 (2006), pp. 558-576.

[15] M. Goh and F. Meng, *Semismooth Matrix-Valued Functions*, Mathematics of Operations Research, 27 (2002), pp 150-169.

[16] J. Gondzio , *Interior Point Methods: Semidefinite Programming and Second-Order Cone Programming*, NATCOR Convex Optimization course, The University of Edinburgh (2016).

[17] S. Hayashi, N. Yamashita and M. Fukushima, *A combined smoothing and regularization method for monotone second- order cone complementarity problem*, SIAM J. Optim, 15 (2005), pp. 593-615

[18] E. de Klerk, *Aspects of semidefinite programming*, Kluwer Academic Publishers (2002)

[19] C.E. Lemke and J.T. Howson, *Equilibrium points of bimatrix games*, SIAM J. Appl. Math., 12 (1964), pp. 413-423

[20] J.-S. Pang, *Newton's method for B-differentiable equations*, Mathematics of Operations Research, 15(1990), no. 2

[21] L. Qi and J. Sun, *A nonsmooth version of Newton's method*, Mathematical Programming, 58 (1993), pp. 353-367.

[22] R.T. Rockafellar and R.J.-T. Wets, *Variational Analysis*, Springer, A series of Comprehensive Studies in Mathematics, Volume 317 (2009).

[23] D. Sun and J. Sun, *Löwner's operator and spectral functions in Euclidean Jordan algebras*, Mathematics of Operations Research, 33 (2008).

[24] D. Sun and J. Sun, *Semismooth Matrix-Valued Functions*, Mathematics of Operations Research, 27 (2002), No.1, pp 150-169.

[25] Y. Wang and N. Xiu, *Strong Semismoothness of Projection onto Slices of Second-Order Cone*, Journal of Optimization Theory and Applications, 150 (2011), pp 599-614.