THE SMOOTHING FUNCTION OF THE NONSMOOTH MATRIX VALUED FUNCTION

ZHAO JINYE

A THESIS SUBMITTED

FOR THE DEGREE OF MASTER OF SCIENCE

DEPARTMENT OF MATHEMATICS

NATIONAL UNIVERSITY OF SINGAPORE

2004

This thesis is dedicated to

my parents

Acknowledgements

I would like to thank my advisor, Dr. Sun Defeng for all of his guidance, encouragement and support. He has inspired me through his own sincere interest in the areas of Operations Research. He taught me how to do research and did it well with great care and patience.

I would also like to acknowledge National University of Singapore for providing me the financial support and the pleasant environment for my study.

I am also grateful to many friends at National University of Singapore for their great help and support.

Last but not least, I would like to thank my family. I am very thankful to my mother who has kept her very best for me.

Contents

Summary

The smoothing functions of nonsmooth matrix valued functions play an important role in the smoothing Newton method. Recently the smoothing Newton method has been extensively studied to solve the semidefinite complementarity problems. In this thesis, we introduce a generalized smoothing function of the nonsmooth matrix valued function on the base of the smoothing function of the scalar valued function. The existence of such smoothing function can be obtained via convolution.

In Chapter 2, we first study the properties of the smoothing function of the vector valued function defined via convolution. We discuss the directional differentiability, semismoothness and strong semismoothness of the smoothing function of the vector valued function. Though we cannot prove that the smoothing function of the vector valued function inherits the strong semismoothness from the nonsmooth vector valued function corresponding to it, we can show that, under some conditions, the smoothing function corresponding to the piecewise LC^1 function is strongly semismooth.

The smoothing function of the scalar valued function is the one dimension case of the vector valued function function. By the results obtained in Chapter 2, we know that the smoothing function of the scalar valued function can have the properties of locally Lipschitz continuity, continuous differentiability, directional differentiability, semismoothness and strong semismoothness. Based on the relationship between the generalized smoothing function of the matrix valued function and the scalar valued function, we show in Chapter 3 and Chapter 4 that the smoothing function of the matrix valued function inherits the properties of the corresponding smoothing function of the scalar valued function.

We also extend the smoothing function for the second order cone complementarity problems and the matrix valued function defined by singular values in Chapter 5.

Chapter 1

Introduction

Let $\mathcal{S}^n$ denote the linear space of $n \times n$ symmetric matrices. Let $\mathcal{O}_n$ denote the set of $n \times n$ orthogonal matrices. For a matrix $X \in \mathcal{S}^n$, its eigenvalues are $\lambda_1(X), ..., \lambda_n(X)$ and it admits a spectral decomposition of the form:

$$X = Q(X)\text{diag}[\lambda_1(X), ..., \lambda_n(X)]Q(X)^T,$$

where $Q(X) \in \mathcal{O}_n$. For a continuous function $f : \mathcal{R} \mapsto \mathcal{R}$, we can define a corresponding matrix valued function $F : \mathcal{S}^n \mapsto \mathcal{S}^n$, by

$$F(X) := Q(X)\text{diag}[f(\lambda_1(X)), ..., f(\lambda_n(X))]Q(X)^T. \tag{1.1}$$

F is well defined (see [16, Sec. 6.2].) It is known that F inherits many properties from f (see [5],[7].) In particular, if f is a nonsmooth function, then F is also nonsmooth. The nonsmooth matrix valued function often arises from the semidefinite programs (SDP) and the semidefinite complementarity problems (SDCP), which include the nonlinear complementarity problems (NCP) as a special case. The smoothing Newton method based on the smoothing function has been widely investigated for NCP. See [24] and the references therein. In [7], Chen and Tseng extended the smoothing Newton method to SDCP. See [5], [19], [41] and the references therein for more discussion.

In [7], [40], [41], some specific smoothing functions of the matrix valued function $F(X) = X_+$ have been studied. However, in this thesis, we will focus on the study of the generalized smoothing functions of the nonsmooth matrix valued functions. In particular, we are interested in this kind of smoothing functions: $G(\varepsilon, X) : \mathcal{R} \times \mathcal{S}^n \mapsto \mathcal{S}^n$ such that G is continuously differentiable on $\mathcal{R} \times \mathcal{S}^n$ unless $\varepsilon = 0$ and $\lim_{(\varepsilon,Z)\to(0,X)} G(\varepsilon, Z) = F(X)$. We define a smoothing function of F by

$$G(\varepsilon, X) := Q(X)\text{diag}[g(\varepsilon, \lambda_1(X)), ..., g(\varepsilon, \lambda_n(X))]Q(X)^T, \tag{1.2}$$

where $g : \mathcal{R} \times \mathcal{R}$ is a smoothing function of f. For convenience of discussion, we always define $G(0, X) = F(X)$ and for any $\varepsilon < 0$ and $X \in \mathcal{S}^n$, $G(\varepsilon, X) = G(-\varepsilon, X)$.

Consider the locally Lipschitz continuous function $\mathcal{F} : \mathcal{R}^n \mapsto \mathcal{R}^m$. Let $\phi : \mathcal{R} \mapsto \mathcal{R}_+$ be a kernel function, i.e., ϕ is Lebesgue integrable and

$$\int_{\mathcal{R}} \phi(y)\, d\mu(y) = 1. \tag{1.3}$$

Here μ is Lebesgue measure (see [33, p.61] for the definition of Lebesgue measure.) Define $\text{supp}(\phi) := \{y \in \mathcal{R} | \phi(y) > 0\}$. Define $\Phi : \mathcal{R}^n \mapsto \mathcal{R}_+$ by

$$\Phi(x) := \prod_{i=1}^{n} \phi(x_i), x \in \mathcal{R}^n.$$

Define $\theta : \mathcal{R}_{++} \times \mathcal{R} \mapsto \mathcal{R}_+$ by

$$\theta(\varepsilon, x) := \varepsilon^{-1}\phi(\varepsilon^{-1}x), \quad (\varepsilon, x) \in \mathcal{R}_{++} \times \mathcal{R}.$$

Let $\Theta : \mathcal{R}_{++} \times \mathcal{R}^n \mapsto \mathcal{R}_+$,

$$\Theta(\varepsilon, x) := \varepsilon^{-n}\Phi(\varepsilon^{-1}x), \quad (\varepsilon, x) \in \mathcal{R}_{++} \times \mathcal{R}^n.$$

Then the smoothing function $\mathcal{G} : \mathcal{R} \times \mathcal{R}^n \mapsto \mathcal{R}^m$ corresponding to the vector valued function $\mathcal{F}$ can be defined by

$$\begin{aligned}\mathcal{G}(\varepsilon, x) &:= \int_{\mathcal{R}^n} \mathcal{F}(x-\varepsilon y)\Phi(y)d\mu(y) \\ &= \int_{\mathcal{R}^n} \mathcal{F}(x-y)\Theta(\varepsilon, y)d\mu(y) \\ &= \int_{\mathcal{R}^n} \mathcal{F}(y)\Theta(\varepsilon, x-y)d\mu(y),\end{aligned} \tag{1.4}$$

where $(\varepsilon, x) \in \mathcal{R}_{++} \times \mathcal{R}^n$, $\mathcal{G}(0, x) = \mathcal{F}(x)$ and for $\varepsilon < 0$, $\mathcal{G}(\varepsilon, x) = \mathcal{G}(-\varepsilon, x)$.

In [38], Sun and Qi investigated some properties, such as Lipschitz continuity, continuous differentiability, etc. of $\mathcal{G}$. Motivated by their results, the smoothing function $g : \mathcal{R} \times \mathcal{R} \mapsto \mathcal{R}$ corresponding to the nonsmooth function $f : \mathcal{R} \mapsto \mathcal{R}$, which is used in (1.2), can be defined by

$$\begin{aligned}g(\varepsilon, x) &:= \int_{\mathcal{R}} f(x-\varepsilon y)\phi(y)d\mu(y) \\ &= \int_{\mathcal{R}} f(x-y)\theta(\varepsilon, y)d\mu(y) \\ &= \int_{\mathcal{R}} f(y)\theta(\varepsilon, x-y)d\mu(y),\end{aligned} \tag{1.5}$$

where $(\varepsilon, x) \in \mathcal{R}_{++} \times \mathcal{R}$, $g(0, x) = f(x)$ and for $\varepsilon < 0$, $g(\varepsilon, x) = g(-\varepsilon, x)$.

In this thesis, before we discuss the smoothing function G, we first study the properties of $\mathcal{G}$ defined by (1.4) under two cases: supp(ϕ) is infinite and supp(ϕ) is bounded. We show that for a general vector valued function $\mathcal{F}$, under some conditions, the smoothing function $\mathcal{G}$ inherits the properties such as directional differentiability and semismoothness from $\mathcal{F}$. When supp(ϕ) is bounded, we can show that if $\mathcal{F}$ is strongly semismooth function, then so is the smoothing function $\mathcal{G}$. However, when supp(ϕ) is infinite, we have not obtained the analogous result for any strongly semismooth function $\mathcal{F}$. Fortunately, we can verify that if $\mathcal{F}$ is a piecewise LC^1 function, under some conditions, then $\mathcal{G}$ is strongly semismooth. These results are the main contributions of this thesis. Therefore, our study completes the analysis of the smoothing function of the vector valued function in [38].

Based on the study of the smoothing function of the vector valued function, we will show that the properties of Lipschitz continuity, continuous differentiability, directional differentiability and (strong) semismoothness are also inherited by G from g. Especially, the property of (strong) semismoothness of the matrix valued function plays an important role in the (quadratic) superlinear convergence analysis of the smoothing Newton method.

The organization of this thesis is as follows. In Subchapter 2.1, when $\text{supp}(\phi)$ is infinite, we prove the directional differentiability and semismoothness of the smoothing function of vector valued function. We also verify that the smoothing function of the piecewise LC^1 function is strongly semismooth under some conditions. In Subchapter 2.2, when $\text{supp}(\phi)$ is bounded, we do some analogous analysis of the vector valued smoothing function. These results are essential for establishing some properties of G. In Chapter 3, we will show that G inherits the properties of Lipschitz continuity, continuous differentiability and directional differentiability from g. In Chapter 4, we prove that if g is (strongly) semismooth, so is G. In Chapter 5, we apply our smoothing function to the vector valued function associated with the second order cone and extend the smoothing function to the matrix valued function over nonsymmetric matrices. The final remarks are stated in Chapter 6.

Chapter 2

Properties of the Smoothing Function $\mathcal{G}$

In this chapter, we will focus on the smoothing function $\mathcal{G}$ defined by (1.4). Some assumptions will be stated in the following subsections to make (1.4) well defined. Also see [25], [34], [36], [38] etc. for some discussion about $\mathcal{G}$. The following examples are three well-known smoothing functions of plus function $f(x) = x_+$.

Example 2.1. *The neural networks smoothing function ([2], [3])*

Let $\phi(x) = \dfrac{e^{-x}}{(1+e^{-x})^2}$, $x \in \mathcal{R}$. *Then the smoothing function of* $f(x) = x_+$ *is*

$$g(\varepsilon, x) = x + \varepsilon ln(1 + e^{-\frac{x}{\varepsilon}}), \ \textit{where } (\varepsilon, x) \in \mathcal{R}_{++} \times \mathcal{R}.$$

Example 2.2. *The uniform smoothing function ([10], [14], [44])*

Let $\phi(x) = \begin{cases} 1 & \textit{if } -\frac{1}{2} \le x \le \frac{1}{2}, \\ 0 & \textit{otherwise,} \end{cases}$ $x \in \mathcal{R}$. *Then the smoothing function of* $f(x) = x_+$ *is*

$$g(\varepsilon, x) = \begin{cases} x & \textit{if } x \ge \frac{\varepsilon}{2}, \\ \frac{1}{2\varepsilon}(x + \frac{\varepsilon}{2})^2 & \textit{if } -\frac{\varepsilon}{2} < x < \frac{\varepsilon}{2}, \\ 0 & \textit{if } x \le -\frac{\varepsilon}{2}, \end{cases} \tag{2.1}$$

where $(\varepsilon, x) \in \mathcal{R}_{++} \times \mathcal{R}$.

Example 2.3. *The Chen-Harker-Kanzow-Smale (CHKS) smoothing function ([1], [20], [37])*

Let $\phi(x) = \frac{2}{(x^2+4)^{3/2}}$, $x \in \mathcal{R}$. *Then the smoothing function of* $f(x) = x_+$ *is*

$$g(\varepsilon, x) = \frac{\sqrt{4\varepsilon^2 + x^2} + x}{2}, \text{ where } (\varepsilon, x) \in \mathcal{R}_{++} \times \mathcal{R}.$$

In the next example, we use the Weierstrass kernel function as ϕ.

Example 2.4. *Let* $\phi(x) = \frac{1}{\sqrt{\pi}} e^{-x^2}$ *Then the smoothing function of* $f(x) = x_+$ *is*

$$g(\varepsilon, x) = \frac{x}{2} u(\frac{x}{\varepsilon}) + \frac{x}{2} + \frac{\varepsilon}{2\sqrt{\pi}} e^{-\frac{x^2}{\varepsilon^2}}, \tag{2.2}$$

where $(\varepsilon, x) \in \mathcal{R}_{++} \times \mathcal{R}$ *and*

$$u(x) := \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} dt.$$

We name this new smoothing function the Weierstrass smoothing function.

The smoothing function of the plus function $f(x) = x_+$ defined via convolution has been extensively studied recently. Based on the work of Chen and Mangasarian ([2], [3]), Tseng [43] introduced the $\mathcal{CM}$ function $\rho(x) : \mathcal{R} \mapsto \mathcal{R}$, which is a smooth convex function and satisfies $\lim_{x\to-\infty} \rho(x) = 0$, $\lim_{x\to\infty} \rho(x) - x = 0$ and $0 < \rho'(x) < 1$, for all $x \in \mathcal{R}$; and approximated the plus function $f(x) = x_+$ by the smoothing function $\varepsilon\rho(x/\varepsilon)$ where $\varepsilon > 0$. When ρ is twice continuous differentiable, $\varepsilon\rho(x/\varepsilon)$ can be written in the convolution form $\varepsilon\rho(x/\varepsilon) = \int_{-\infty}^{\infty} (x-t)_+ \frac{1}{\varepsilon}\phi(\frac{t}{\varepsilon}) dt$, where $\phi = \rho''$, which is actually a special case of the smoothing function defined in (1.4). Also see [7].

The smoothing function of the projection function, which is a more general case of the plus function was introduced in [13]. Qi, Sun and Zhou [27] discussed the strong semismoothness of this group of smoothing function. Also see [6].

In [29], Qi and Tseng studied the (strong) semismoothness of the recession function, which in fact is a kind of convolution (see [31].) Then the recession function is a smoothing function. But this recession function cannot cover the generalized smoothing function defined by (1.4).

Consider $\varphi(x) := \int_a^b [v(x,t)]_+ u(t)dt$, where $u(t) \geq 0$ for $t \in [a,b]$. [9], [26], [29], [30] studied the properties of differentiability and (strong) semismoothness of $\varphi(x)$.

However, all of these investigations are based on the explicit forms of the smoothing functions. In this chapter, we will generalize the results discussed in the above papers from a specific smoothing function to a general smoothing function. We will discuss the directional differentiability and (strong) semismoothness of $\mathcal{G}$ in this chapter, which complete the analysis of the smoothing function in [38].

For a general vector valued function $\Psi : \mathcal{R}^n \mapsto \mathcal{R}^m$, denote the set of points at which Ψ is differentiable by D_Ψ. Let $\partial_B \Psi(x)$ be the B-subdifferential of Ψ at $x \in \mathcal{R}^n$ defined by

$$\partial_B \Psi(x) = \{ \lim_{\substack{x^k \to x \\ x^k \in D_\Psi}} \Psi'(x^k) \}.$$

The generalized Jacobian $\partial \Psi(x)$ of Ψ at $x \in \mathcal{R}^n$ (in Clarke's sense) is defined as the convex hull of $\partial_B \Psi(x)$. These notions also hold for the matrix valued function.

The vector norm of $x \in \mathcal{R}^n$ is denoted by $\|x\|$. For any n-by-n matrix X, $\|X\|$ is the Frobenius norm of X. We denote the norm of the operator $M : \mathcal{S}^n \mapsto \mathcal{S}^n$ by $|||M||| := \max_{\|X\|=1} \|MX\|$.

(Strong) semismoothness plays an important role in the analysis of the (quadratic) superlinear convergence of generalized Newton methods for nonsmooth equations. Mifflin [23] and Qi and Sun [28] introduced the semismoothness and strong semismoothness for the vector valued function.

Definition 2.5. Suppose that $\Psi : \mathcal{R}^n \mapsto \mathcal{R}^m$ is locally Lipschitz continuous around $x \in \mathcal{R}^n$. Ψ is said to be semismooth at x if Ψ is directionally differentiable at x and for any $V \in \partial \Psi(x + \Delta x)$,

$$\Psi(x + \Delta x) - \Psi(x) - V(\Delta x) = o(\|\Delta x\|).$$

Ψ is said to be strongly semismooth at x if Ψ is semismooth at x and

$$\Psi(x + \Delta x) - \Psi(x) - V(\Delta x) = O(\|\Delta x\|^2).$$

We also use the following lemma to prove the semismoothness and strong semismoothness. This result was obtained in [40, Proposition 2.3].

Lemma 2.6. *Suppose that $\Psi : \mathcal{R}^n \mapsto \mathcal{R}^m$ is locally Lipschitz continuous around $x \in \mathcal{R}^n$. Then the following two statements are equivalent:*

(i) for any $V \in \partial\Psi(x + \Delta x)$,

$$\Psi(x + \Delta x) - \Psi(x) - V(\Delta x) = o(\|\Delta x\|)(\text{ respectively, } O(\|\Delta\|^2));$$

(ii) for any $x + \Delta x \in D_\Psi$,

$$\Psi(x + \Delta x) - \Psi(x) - \Psi'(x + \Delta x)(\Delta x) = o(\|\Delta x\|)(\text{ respectively, } O(\|\Delta\|^2)).$$

2.1 supp(ϕ) is infinite

Suppose that $\mathcal{F} : \mathcal{R}^n \mapsto \mathcal{R}^m$ is a locally Lipschitz continuous function. In this subchapter, we study the properties of the smoothing function $\mathcal{G}$ defined in (1.4), when supp(ϕ) is infinite. In order to guarantee some properties of $\mathcal{G}$ such as Lipschitz continuity and continuous differentiability, we introduce the following assumption in this subsection:

Assumption 2.7. *(i) $\mathcal{F}$ is globally Lipschitz continuous with Lipschitz constant $\mathcal{L}_\mathcal{F}$.*

(ii) $\displaystyle\int_{\mathcal{R}^n} \|y\|\Phi(y)d\mu(y) < \infty.$

(iii) Φ *is continuously differentiable on* $\mathcal{R}^n$*, with* $\int_{\mathcal{R}^n} \|y\|^2 \|\Phi'(y)\| d\mu(y) < \infty$*,*

for any $h \in \mathcal{R}^n$*,* $\sup_{t\in[0,1]} \int_{\mathcal{R}^n} \|y\| \|\Phi'(y+th) - \Phi'(y)\| d\mu(y) = O(\|h\|)$ *and*

for any $\tau \in \mathcal{R}_+$*,* $\sup_{t\in[0,1]} \int_{\mathcal{R}^n} \|y\|^2 \|\Phi'(\frac{1+\tau}{1+t\tau} y) - \Phi'(y)\| d\mu(y) = O(|\tau|)$*.*

(iv) $\sup_{y_i \in \mathcal{R}} |y_i|^3 \phi(y_i) < \infty$*,* $i = 1, ..., n$*.*

Here Assumption 2.7 (i) and (ii) are used to make $\mathcal{G}$ well defined. Assumption 2.7 (iii) is utilized to establish the continuous differentiability of $\mathcal{G}$. Since we need some results obtained in [38] in the following discussion, Lemma 2.10 shows that Assumption 2.7 can imply the assumptions used in [38, Theorem 3.7], which is stated by Assumption 2.8 and Assumption 2.9.

Assumption 2.8. *([38, Assumption 3.4])*

(i) $\mathcal{F}$ *is globally Lipschitz continuous with Lipschitz constant L.*

(ii) $\kappa := \int_{\mathcal{R}^n} \|y\| \Phi(y) d\mu(y) < \infty$*.*

(iii) Φ *is continuously differentiable and for any* $\varepsilon > 0$ *and* $x \in \mathcal{R}^n$*, the following integral:*

$$\int_{\mathcal{R}^n} \mathcal{F}(y) \Theta'_x(\varepsilon, x - y) dy$$

exists.

(iv) For any $\varepsilon > 0$*,* $x \in \mathcal{R}^n$ *and* $h \to 0$*, it holds that*

$$\sup_{t\in[0,1]} \int_{\mathcal{R}^n} \|y\| \|[\Theta'_x(\varepsilon, x + th - y) - \Theta'_x(\varepsilon, x - y)]h\| dy = o(\|h\|).$$

Assumption 2.9. *([38, Assumption 3.6])*

(i) For any $\varepsilon > 0$ *and* $x \in \mathcal{R}^n$*, the following integral:*

$$\int_{\mathcal{R}^n} \mathcal{F}(y) \Theta'_\varepsilon(\varepsilon, x - y) dy$$

exists.

(ii) For any $\varepsilon > 0$, $x \in \mathcal{R}^n$ and $\tau \in \mathcal{R}$ with $\tau \to 0$, we have

$$\sup_{t\in[0,1]} \int_{\mathcal{R}^n} \|y\| |[\Theta'_\varepsilon(\varepsilon + t\tau, x - y) - \Theta'_\varepsilon(\varepsilon + \tau, x - y)]\tau| dy = o(|\tau|).$$

(iii) For any $\varepsilon > 0$ and $x \in \mathcal{R}^n$, we have

$$\lim_{\tau\to\varepsilon, z\to x} \int_{\mathcal{R}^n} \|y\| |\Theta'_\varepsilon(\tau, z - y) - \Theta'_\varepsilon(\varepsilon, x - y)| dy = 0.$$

The following lemma shows the relationship between Assumption 2.7 and Assumption 2.8, Assumption 2.9.

Lemma 2.10. *(i) Assumption 2.7 (i), (ii) and (iii) imply Assumption 2.8.*

(ii) Assumption 2.7 (i), (ii) and (iii) imply Assumption 2.9 (i) and (ii).

(iii) Under Assumption 2.7 (i), (ii) and (iii), we have that $\mathcal{G}'_x(\varepsilon, x)$ and $\mathcal{G}'_\varepsilon(\varepsilon, x)$ are locally Lipschitz continuous on $\mathcal{R}_{++} \times \mathcal{R}^n$.

Proof. (i) For any $\varepsilon > 0$ and $x \in \mathcal{R}^n$, by the assumption $\int_{\mathcal{R}^n} \|y\|^2 \|\Phi'(y)\| d\mu(y) < \infty$, the integral

$$\int_{\mathcal{R}^n} \mathcal{F}(y)\Theta'_x(\varepsilon, x - y) d\mu(y) = \int_{\mathcal{R}^n} \mathcal{F}(y) \frac{1}{\varepsilon^{n+1}} \Phi'(\frac{x - y}{\varepsilon}) d\mu(y).$$

exists. For any $h \in \mathcal{R}^n$ with $h \to 0$,

$$\begin{aligned}
&\sup_{t\in[0,1]} \int_{\mathcal{R}^n} \|y\| \|[\Theta'_x(\varepsilon, x + th - y) - \Theta'_x(\varepsilon, x - y)]h\| d\mu(y) \\
&\quad = \sup_{t\in[0,1]} \int_{\mathcal{R}^n} \|y\| \|[\frac{1}{\varepsilon^{n+1}}\Phi'(\frac{x + th - y}{\varepsilon}) - \frac{1}{\varepsilon^{n+1}}\Phi'(\frac{x - y}{\varepsilon})]h\| d\mu(y) \\
&\quad = \frac{1}{\varepsilon^n} \sup_{t\in[0,1]} \int_{\mathcal{R}^n} \|\frac{x - y}{\varepsilon} - \frac{x}{\varepsilon}\| \|[\Phi'(\frac{x - y}{\varepsilon} + t\frac{h}{\varepsilon}) - \Phi'(\frac{x - y}{\varepsilon})]\| d\mu(y) \|h\| \\
&\quad \le \frac{\|x\|}{\varepsilon^{n+1}} \sup_{t\in[0,1]} \int_{\mathcal{R}^n} \|[\Phi'(\frac{x - y}{\varepsilon} + \frac{th}{\varepsilon}) - \Phi'(\frac{x - y}{\varepsilon})]\| d\mu(y) \|h\| \\
&\quad\quad + \frac{1}{\varepsilon^n} \sup_{t\in[0,1]} \int_{\mathcal{R}^n} \|\frac{x - y}{\varepsilon}\| \|[\Phi'(\frac{x - y}{\varepsilon} + \frac{th}{\varepsilon}) - \Phi'(\frac{x - y}{\varepsilon})]\| d\mu(y) \|h\| \\
&\quad = o(\|h\|),
\end{aligned}$$

where the last equation is followed by the assumption $\sup_{t\in[0,1]}\int_{\mathcal{R}^n}\|y\|\|\Phi'(y+th)-\Phi'(y)\|d\mu(y)=O(\|h\|)$. Then Assumption 2.8 holds.

(ii) For any $\varepsilon>0$ and $x\in\mathcal{R}^n$, by the assumption $\int_{\mathcal{R}^n}\|y\|^2\|\Phi'(y)\|d\mu(y)<\infty$, the following integral

$$\begin{aligned}&\int_{\mathcal{R}^n}\mathcal{F}(y)\Theta'_\varepsilon(\varepsilon,x-y)d\mu(y)\\&\qquad=\int_{\mathcal{R}^n}\mathcal{F}(y)\frac{-1}{\varepsilon^{n+1}}\Phi(\frac{x-y}{\varepsilon})d\mu(y)+\int_{\mathcal{R}^n}\mathcal{F}(y)\frac{y-x}{\varepsilon^{n+2}}\Phi'(\frac{x-y}{\varepsilon})d\mu(y).\end{aligned}$$

exists. For any $\tau\in\mathcal{R}_+$ with $\tau\to0$, obviously,

$$\sup_{t\in[0,1]}\int_{\mathcal{R}^n}\|y\|\|[\frac{-1}{(\varepsilon+t\tau)^{n+1}}\Phi(\frac{x-y}{\varepsilon+t\tau})-\frac{-1}{(\varepsilon+\tau)^{n+1}}\Phi(\frac{x-y}{\varepsilon+\tau})]\tau|d\mu(y)=o(|\tau|).$$

Moreover, we have

$$\begin{aligned}&\sup_{t\in[0,1]}\int_{\mathcal{R}^n}\|y\|\|[\frac{y-x}{(\varepsilon+t\tau)^{n+2}}\Phi'(\frac{x-y}{\varepsilon+t\tau})-\frac{y-x}{(\varepsilon+\tau)^{n+2}}\Phi'(\frac{x-y}{\varepsilon+\tau})]\tau|d\mu(y)\\&\le\sup_{t\in[0,1]}\int_{\mathcal{R}^n}\|x-(\varepsilon+\tau)y\|\|[\frac{(\varepsilon+\tau)^2y}{(\varepsilon+t\tau)^{n+2}}\Phi'(\frac{\varepsilon+\tau}{\varepsilon+t\tau}y)-\frac{(\varepsilon+\tau)^2y}{(\varepsilon+\tau)^{n+2}}\Phi'(y)]\tau|d\mu(y)\\&\le\|x\||\varepsilon+\tau|^2\sup_{t\in[0,1]}\int_{\mathcal{R}^n}\|y\|\|[\frac{1}{(\varepsilon+t\tau)^{n+2}}\Phi'(\frac{\varepsilon+\tau}{\varepsilon+t\tau}y)-\frac{1}{(\varepsilon+\tau)^{n+2}}\Phi'(y)]\tau\|d\mu(y)\\&\quad+|\varepsilon+\tau|^3\sup_{t\in[0,1]}\int_{\mathcal{R}^n}\|y\|^2\|[\frac{1}{(\varepsilon+t\tau)^{n+2}}\Phi'(\frac{\varepsilon+\tau}{\varepsilon+t\tau}y)-\frac{1}{(\varepsilon+\tau)^{n+2}}\Phi'(y)]\tau\|d\mu(y)\\&=o(|\tau|),\end{aligned}$$

where the last equality is followed by the assumption $\sup_{t\in[0,1]}\int_{\mathcal{R}^n}\|y\|^2\|\Phi'(\frac{1+\tau}{1+t\tau}y)-\Phi'(y)\|d\mu(y)=O(|\tau|)$. Therefore, we have

$$\sup_{t\in[0,1]}\int_{\mathcal{R}^n}\|y\|\|[\Theta'_\varepsilon(\varepsilon+t\tau,x-y)-\Theta'_\varepsilon(\varepsilon+\tau,x-y)]\tau|d\mu(y)=o(|\tau|).$$

Then Assumption 2.9 (i) and (ii) follow.

(iii) By [38, Proposition 3.5], we have for $\varepsilon>0$ and $x\in\mathcal{R}^n$,

$$\mathcal{G}'_x(\varepsilon,x)=\int_{\mathcal{R}^n}\mathcal{F}(y)\frac{1}{\varepsilon^{n+1}}\Phi'(\frac{x-y}{\varepsilon})d\mu(y)=\frac{1}{\varepsilon^n}\int_{\mathcal{R}^n}\mathcal{F}(x-\varepsilon y)\Phi'(y)d\mu(y).$$

For (ε_1, x_1), $(\varepsilon_2, x_2) \in B(\varepsilon, x)$,

$$\begin{aligned}
&\| \int_{\mathcal{R}^n} \mathcal{F}(x_1 - \varepsilon_1 y)\Phi'(y)d\mu(y) - \int_{\mathcal{R}^n} \mathcal{F}(x_2 - \varepsilon_2 y)\Phi'(y)d\mu(y)\| \\
&\qquad \le \int_{\mathcal{R}^n} \|\mathcal{F}(x_1 - \varepsilon_1 y) - \mathcal{F}(x_2 - \varepsilon_2 y)\| \|\Phi'(y)\| d\mu(y) \\
&\qquad \le L \int_{\mathcal{R}^n} \|\Phi'(y)\| d\mu(y) \|x_1 - x_2\| + L \int_{\mathcal{R}^n} \|y\| \|\Phi'(y)\| d\mu(y) |\varepsilon_1 - \varepsilon_2| \\
&\qquad \le K \|(\varepsilon_1 - \varepsilon_2, x_1 - x_2)\|,
\end{aligned}$$

where K is a constant. Together with the locally lipschitz continuity of $\frac{1}{\varepsilon^n}$, $\mathcal{G}'_x(\varepsilon, x)$ is locally Lipschitz continuous on $\mathcal{R}_{++} \times \mathcal{R}^n$. By using the same way, we can also show that $\mathcal{G}'_\varepsilon(\varepsilon, x)$ is locally Lipschitz continuous on $\mathcal{R}_{++} \times \mathcal{R}^n$. □

There are lots of kernel functions satisfying these assumptions. For instance, the ones mentioned in Examples 2.1, 2.3 and 2.4. Next proposition shows the directional differentiability of $\mathcal{G}$.

Proposition 2.11. *Suppose that $\mathcal{F}$ is directionally differentiable at $x \in \mathcal{R}^n$ and Assumption 2.7 (i) and (ii) are satisfied. Then the directional derivative of $\mathcal{G}$ at $(0, x)$ exists and is given by*

$$\mathcal{G}'((0,x);(\tau,h)) = \int_{\mathcal{R}^n} \mathcal{F}'(x; h - |\tau| y)\Phi(y)d\mu(y) \tag{2.3}$$

for any $(\tau, h) \in \mathcal{R} \times \mathcal{R}^n$.

Proof. For any $(\tau, h) \in \mathcal{R} \times \mathcal{R}^n$, let

$$\begin{aligned}
\Delta_t \mathcal{G}((0,x);(\tau,h)) :=& \frac{\mathcal{G}(t|\tau|, x + th) - \mathcal{G}(0,x)}{t} \\
=& \int_{\mathcal{R}^n} \frac{\mathcal{F}(x + th - t|\tau| y) - \mathcal{F}(x)}{t} \Phi(y) d\mu(y).
\end{aligned}$$

By Assumption 2.7 (i), we have

$$\| \frac{\mathcal{F}(x + th - t|\tau|y) - \mathcal{F}(x)}{t} \Phi(y) \| \le L\|h\|\Phi(y) + L|\tau|\|y\|\Phi(y), \quad \forall t > 0.$$

Assumption 2.7 (ii) implies that $L\|h\|\Phi(y) + L|\tau|\|y\|\Phi(y)$ is Lebesgue integrable. Thus by Lebesgue's Dominated Convergence Theorem, we have

$$\begin{aligned}\lim_{t\downarrow 0} \Delta_t \mathcal{G}((0,x);(\tau,h)) &= \lim_{t\downarrow 0} \int_{\mathcal{R}^n} \frac{\mathcal{F}(x+th-t|\tau|y)-\mathcal{F}(x)}{t}\Phi(y)d\mu(y)\\ &= \int_{\mathcal{R}^n} \mathcal{F}'(x; h-|\tau|y)\Phi(y)d\mu(y),\end{aligned}$$

which, by the definition of directional derivative, proves (2.3).

□

Based on the results obtained in [38], the following two theorems show two important properties of $\mathcal{G}$, semismoothness and strong semismoothness, which haven't been proven completely in [38].

Theorem 2.12. *Suppose that Assumption 2.7 (i), (ii) and (iii) hold. If $\mathcal{F}$ is semismooth on $\mathcal{R}^n$, then $\mathcal{G}$ is semismooth on $\mathcal{R}_+ \times \mathcal{R}^n$.*

Proof. By [38, Theorem 3.7 (i)], $\mathcal{G}$ is continuously differentiable on $\mathcal{R}_{++} \times \mathcal{R}^n$. Then it implies that $\mathcal{G}$ is semismooth on $\mathcal{R}_{++} \times \mathcal{R}^n$. Thus, we only need to show that $\mathcal{G}$ is semismooth at $(0,x)$, $x \in \mathcal{R}^n$. By [38, Theorem 3.7 (vii)], we have for any $(\varepsilon, d) \in \mathcal{R}_{++} \times \mathcal{R}^n$ with $(\varepsilon, d) \to 0$,

$$\mathcal{G}(\varepsilon, x+d) - \mathcal{G}(0,x) - \mathcal{G}'(\varepsilon, x+d)\begin{pmatrix}\varepsilon\\ d\end{pmatrix} = o(\|(\varepsilon,d)\|).$$

It has been proved in [28] that $\mathcal{F}$ is semismooth at x if and only if all its component functions are. Then we have for $\forall\, \alpha_i \in \partial\mathcal{F}_i(x)$, $i = 1, ..., m$,

$$\mathcal{F}_i(x+d) - \mathcal{F}_i(x) - \alpha_i d = o(\|d\|).$$

Furthermore, by [22] and [38, Theorem 3.7 (iv)], we have $\pi_x \partial\mathcal{G}_i(0, x+d) \subseteq \partial\mathcal{F}_i(x+d)$, where

$$\pi_x \partial\mathcal{G}_i(0,x) := \{\alpha \in \mathcal{R}^n |\ \text{There exists } \beta \in \mathcal{R} \text{ such that } (\beta,\alpha) \in \partial\mathcal{G}_i(0,x)\}.$$

Therefore, for any $V_i = (\beta_i, \alpha_i) \in \partial\mathcal{G}_i(0, x+d)$, we have $\alpha_i \in \partial\mathcal{F}_i(x+d)$, then

$$\mathcal{G}_i(0, x+d) - \mathcal{G}_i(0, x) - V_i \begin{pmatrix} 0 \\ d \end{pmatrix} = \mathcal{F}_i(x+d) - \mathcal{F}_i(x) - \alpha_i d = o(\|d\|).$$

Therefore, $\mathcal{G}$ is semismooth at $(0, x)$. Our result follows. □

Next, we discuss the strong semismoothness of $\mathcal{G}$, when $\mathcal{F}$ is a piecewise LC^1 function. Before we prove the proposition, we review some properties of the piecewise LC^1 function.

Definition 2.13. A continuous function $\mathcal{F} : \mathcal{R}^n \mapsto \mathcal{R}^m$ is called piecewise LC^1 function, if there exist finitely many continuously differentiable functions $L_i : \mathcal{R}^n \mapsto \mathcal{R}^m$, whose derivatives are locally Lipschitz continuous, $i = 1, ..., k$, such that $\mathcal{F}(x) \in \{L_1(x), ..., L_k(x)\}$ holds for every $x \in \mathcal{R}^n$. L_i, $i = 1, ..., k$ are called the selection functions for $\mathcal{F}$.

The concept of the essentially active indices was introduced by Scholtes in [42].

Definition 2.14. The set of essentially active indices of piecewise LC^1 function $\mathcal{F}$ at x_0 is defined by

$$I^e_{\mathcal{F}}(x_0) = \{i \in \{1, ..., k\} | x_0 \in cl(\text{int}\{x \in R^n | \mathcal{F}(x) = L_i(x)\})\}.$$

A selection function L_i is called essentially active at x_0 if $i \in I^e_{\mathcal{F}}(x_0)$.

Let $\Sigma_i := \{x \in \mathcal{R}^n | \mathcal{F}(x) = L_i(x)\} \subseteq \mathcal{R}^n$, $i = 1, ..., k$. Since $\mathcal{F}$ is continuous, Σ_i, $i = 1, ..., k$ are closed sets. Lemma 2.15 shows the covering property of the union of Σ_i, $i = 1, ..., k$.

Lemma 2.15. *If we remove those sets Σ_i with empty interior from Σ_i, $i = 1, .., k$, then the remaining collection of sets Σ_i still covers $\mathcal{R}^n$.*

Proof. By [42, Proposition 4.1.1], we know that for any $x_0 \in \mathcal{R}^n$, there exists a collection of selection functions for $\mathcal{F}$ at x_0, which are all essentially active, i.e. there exists $i_0 \in \{1, ..., k\}$, such that the interior of Σ_{i_0} is nonempty and $x_0 \in \Sigma_{i_0}$. □

Without loss of generality, in this thesis, we assume that the interior of each Σ_i, $i = 1, ..., k$ is nonempty.

Next proposition discusses the strong semismoothness of $\mathcal{G}$.

Theorem 2.16. *Suppose that Assumption 2.7 holds. Assume that $\mathcal{F}$ is a piecewise LC^1 function and there exist $K \geq 0$ and $\kappa > 0$, such that*

(i) $\|L_i'(z) - L_i'(z+y)\| \leq K\|y\|^{\kappa+1}$, $\forall\ z,\ y \in \mathcal{R}^n$, $i = 1, ..., k$,

(ii) $K \displaystyle\int_{\mathcal{R}^n} \|y\|^{2+\kappa}\Phi(y)d\mu(y) < \infty$.

Then $\mathcal{G}$ is strongly semismooth on $\mathcal{R}_+ \times \mathcal{R}^n$.

Proof. Since $\mathcal{F}$ is piecewise LC^1 function, $\mathcal{F}$ is strongly semismooth on $\mathcal{R}^n$. Then, by Proposition 2.12, $\mathcal{G}$ is semismooth on $\mathcal{R}_+ \times \mathcal{R}^n$. Lemma 2.10 (iii) implies that $\mathcal{G}'$ is locally Lipschitz continuous around any $(\varepsilon, x) \in \mathcal{R}_{++} \times \mathcal{R}^n$. Then $\mathcal{G}$ is strongly semismooth on $\mathcal{R}_{++} \times \mathcal{R}^n$. Next, we will focus on the strong semismoothness of $\mathcal{G}$ at $(0, x)$, $x \in \mathcal{R}^n$. For any $(\varepsilon, d) \in \mathcal{R}_{++} \times \mathcal{R}^n$ with $(\varepsilon, d) \to 0$, we have

$$
\begin{aligned}
&\|\mathcal{G}(\varepsilon, x+d) - \mathcal{G}(0, x) - \mathcal{G}'((\varepsilon, x+d); (\varepsilon, d))\| \\
&\quad = \| \int_{\mathcal{R}^n} [\mathcal{F}(x+d-\varepsilon y) - \mathcal{F}(x) - \mathcal{F}'(x+d-\varepsilon y; d-\varepsilon y)]\Phi(y)d\mu(y)\| \\
&\quad \leq \sum_{i=1}^{k} \int_{x+d-\varepsilon y \in \Sigma_i} \|\mathcal{F}(x+d-\varepsilon y) - \mathcal{F}(x) - \mathcal{F}'(x+d-\varepsilon y; d-\varepsilon y)\|\Phi(y)d\mu(y),
\end{aligned}
$$

where the last equality is followed by the covering property of the union of Σ_i, $i = 1, ..., k$. First, let us fix ε and d. Define $\mathcal{H} : \mathcal{R}^n \mapsto \mathcal{R}_+$ by

$$
\mathcal{H}(y) := \|\mathcal{F}(x+d-\varepsilon y) - \mathcal{F}(x) - \mathcal{F}'(x+d-\varepsilon y; d-\varepsilon y)\|\Phi(y), \quad y \in \mathcal{R}^n.
$$

Then $\mathcal{H}(\cdot)$ is an integrable function. By (i) of Assumption 2.7,

$$|\mathcal{H}(y)| \leq 2\mathcal{L}_{\mathcal{F}}(\|d\| + \varepsilon\|y\|)\Phi(y), \quad \forall y \in \mathcal{R}^n,$$

which, together with (ii) of Assumption 2.7 and (1.3), implies that there exists $M > 0$ such that

$$|\mathcal{H}(y)| \leq M(\|d\| + \varepsilon), \quad \forall y \in \mathcal{R}^n.$$

For each i, let

$$\Pi_i(\varepsilon, d) := \int_{x+d-\varepsilon y \in \Sigma_i} \mathcal{H}(y) d\mu(y).$$

Then $\|\mathcal{G}(\varepsilon, x+d) - \mathcal{G}(0, x) - \mathcal{G}'((\varepsilon, x+d); (\varepsilon, d))\| \leq \sum_{i=1}^{k} \Pi_i(\varepsilon, d)$. Since $\text{int}\Sigma_i$ is open, by [18, p.50, M7], $\text{int}\Sigma_i$ is measurable. Then for each $i \in \{1, ..., k\}$, by [18, p.50, M9], for any given $\delta > 0$, there exists a closed set $A_i(\delta)$, such that $A_i(\delta) \subseteq \text{int}\Sigma_i$ and $\mu(\Sigma_i - A_i(\delta)) < \delta$. Let

$$\pi_i(\delta; \varepsilon, d) := \int_{x+d-\varepsilon y \in A_i(\delta)} \mathcal{H}(y) d\mu(y), \quad i = 1, ..., k.$$

Therefore, we have

$$\begin{aligned} |\Pi_i(\varepsilon, d) - \pi_i(\delta; \varepsilon, d)| =& |\int_{x+d-\varepsilon y \in \Sigma_i - A_i(\delta)} \mathcal{H}(y) d\mu(y)| \\ \leq& \frac{M(\|d\| + \varepsilon)}{\varepsilon} \mu(\Sigma_i - A_i(\delta)) \\ <& \frac{M(\|d\| + \varepsilon)}{\varepsilon} \delta. \end{aligned}$$

Hence, $\lim_{\delta \to 0} \pi_i(\delta; \varepsilon, d) = \Pi_i(\varepsilon, d)$. For any $x \in \mathcal{R}^n$, there exists $\bar{j} \in \{1, ..., k\}$ such that $\mathcal{F}(x) = L_{\bar{j}}(x)$. Let i be an arbitrary index in $\{1, ..., k\}$. Then by the definition of $\mathcal{F}$, for $x + d - \varepsilon y \in A_i(\delta) \subseteq \text{int}\Sigma_i$,

$$\mathcal{F}'(x + d - \varepsilon y; d - \varepsilon y) = L_i'(x + d - \varepsilon y)(d - \varepsilon y)$$

and

$$\begin{aligned} &\pi_i(\delta; \varepsilon, d) \\ =& \int_{x+d-\varepsilon y \in A_i(\delta)} \|L_i(x + d - \varepsilon y) - L_{\bar{j}}(x) - L_i'(x + d - \varepsilon y)(d - \varepsilon y)\| \Phi(y) d\mu(y). \end{aligned}$$

We consider the following two cases.

Case i): $i = \bar{j}$. Then $x \in \Sigma_i$. Thus, we have

$$\begin{aligned}
&|\pi_i(\delta;\varepsilon,d)| \\
=&\int_{x+d-\varepsilon y\in A_i(\delta)} \|L_i(x+d-\varepsilon y) - L_{\bar{j}}(x) - L_i'(x+d-\varepsilon y)(d-\varepsilon y)\|\Phi(y)d\mu(y) \\
\leq&\int_{x+d-\varepsilon y\in \mathcal{R}^n} \|L_i(x+d-\varepsilon y) - L_i(x) - L_i'(x+d-\varepsilon y)(d-\varepsilon y)\|\Phi(y)d\mu(y) \\
=&\int_{\mathcal{R}^n} \| \int_0^1 [L_i'(x+\theta(d-\varepsilon y)) - L_i'(x+(d-\varepsilon y))](d-\varepsilon y)d\theta\|\Phi(y)d\mu(y) \\
\leq& K\int_{\mathcal{R}^n} \|d-\varepsilon y\|^{\kappa+2}\Phi(y)d\mu(y) \\
\leq& K2^{\kappa+2}\int_{\mathcal{R}^n} \max\{\|d\|^{\kappa+2}, \varepsilon^{\kappa+2}\|y\|^{\kappa+2}\}\Phi(y)d\mu(y) \\
\leq& O(\max\{\|d\|^{\kappa+2}, \varepsilon^{\kappa+2}\}) \\
=& O(\|(\varepsilon,d)\|^2),
\end{aligned}$$

where the second equality is followed by the Mean Value Theorem and the second inequality is followed by the condition (i). By the condition (ii), we get the last inequality.

Case ii): $i \neq \bar{j}$. Then $x \notin \Sigma_i$. Since Σ_i^c is open, there exists an open ball $B(x,r)$, such that $B(x,r) \subset \Sigma_i^c$. We can find a rectangle $I = [a_1, b_1] \times ... \times [a_n, b_n]$ with diameter less than $2r$, such that $x \in I \subset B(x,r)$. Then $A_i(\delta) \subset \Sigma_i \subset I^c$.

Hence, we have

$$\begin{aligned}
&\int_{x+d-\varepsilon y\in I^c}\Phi(y)d\mu(y)\\
=&1-\int_{x+d-\varepsilon y\in I}\phi(y_1)\cdots\phi(y_n)d\mu(y)\\
=&1-\left(\int_{x_1+d_1-\varepsilon y_1\in[a_1,b_1]}\phi(y_1)d\mu(y_1)\cdots\int_{x_{n-1}+d_{n-1}-\varepsilon y_{n-1}\in[a_{n-1},b_{n-1}]}\phi(y_{n-1})d\mu(y_{n-1})\right)\\
&\left(1-\int_{x_n+d_n-\varepsilon y_n\in[-\infty,a_n]\cup[b_n,\infty]}\phi(y_n)d\mu(y_n)\right)\\
\leq&1-\left(\int_{x_1+d_1-\varepsilon y_1\in[a_1,b_1]}\phi(y_1)d\mu(y_1)\cdots\int_{x_{n-1}+d_{n-1}-\varepsilon y_{n-1}\in[a_{n-1},b_{n-1}]}\phi(y_{n-1})d\mu(y_{n-1})\right)\\
&+\sup_{y_n\in\mathcal{R}}|y_n|^3\phi(y_n)\left(\int_{-\infty}^{\frac{x_n+d_n-b_n}{\varepsilon}}\frac{1}{|y_n|^3}d\mu(y_n)+\int_{\frac{x_n+d_n-a_n}{\varepsilon}}^{+\infty}\frac{1}{|y_n|^3}d\mu(y_n)\right)\\
=&1-\left(\int_{x_1+d_1-\varepsilon y_1\in[a_1,b_1]}\phi(y_1)d\mu(y_1)\cdots\int_{x_{n-1}+d_{n-1}-\varepsilon y_{n-1}\in[a_{n-1},b_{n-1}]}\phi(y_{n-1})d\mu(y_{n-1})\right)\\
&+O(|\varepsilon|^2),
\end{aligned}\tag{2.4}$$

where the inequality is implied by (1.3) and Assumption 2.7 (iv). Since we have

$$\begin{aligned}
&\int_{x_1+d_1-\varepsilon y_1\in[a_1,b_1]}\phi(y_1)d\mu(y_1)\cdots\int_{x_{n-1}+d_{n-1}-\varepsilon y_{n-1}\in[a_{n-1},b_{n-1}]}\phi(y_{n-1})d\mu(y_{n-1})\\
=&\int_{x_1+d_1-\varepsilon y_1\in[a_1,b_1]}\phi(y_1)d\mu(y_1)\cdots\int_{x_{n-2}+d_{n-2}-\varepsilon y_{n-2}\in[a_{n-2},b_{n-2}]}\phi(y_{n-2})d\mu(y_{n-2})\\
&\left(1-\int_{x_{n-1}+d_{n-1}-\varepsilon y_{n-1}\in[-\infty,a_{n-1}]\cup[b_{n-1},\infty]}\phi(y_{n-1})d\mu(y_{n-1})\right),
\end{aligned}$$

by repeating (2.4), we obtain that

$$\int_{x+d-\varepsilon y\in I^c}\Phi(y)d\mu(y)\leq\sum_{i=1}^{n}O(|\varepsilon|^2)=O(|\varepsilon|^2).\tag{2.5}$$

Thus, we obtain that

$$
\begin{aligned}
&|\pi_i(\delta;\varepsilon,d)| \\
=&\int_{x+d-\varepsilon y\in A_i(\delta)} \|L_i(x+d-\varepsilon y)-L_{\bar{j}}(x)-L_i'(x+d-\varepsilon\varrho y)(d-\varepsilon y)\|\Phi(y)d\mu(y) \\
\leq&\int_{x+d-\varepsilon y\in A_i(\delta)} \|L_i(x+d-\varepsilon y)-L_i(x)-L_i'(x+d-\varepsilon y)(d-\varepsilon y)\|\Phi(y)d\mu(y) \\
&+\|L_i(x)-L_{\bar{j}}(x)\|\int_{x+d-\varepsilon y\in I^c}\Phi(y)d\mu(y) \\
=&O(\|(\varepsilon,d)\|^2)+\|L_i(x)-L_{\bar{j}}(x)\|\sum_{j=1}^{n}O(|\varepsilon|^2) \\
=&O(\|(\varepsilon,d)\|^2),
\end{aligned}
$$

where the second equality is followed by the result obtained in the case (i) and (2.5). Hence, we have $\Pi_i(\varepsilon,d)=O(\|(\varepsilon,d)\|^2)$, which implies that $\mathcal{G}(\varepsilon,x+d)-\mathcal{G}(0,x)-\mathcal{G}'((\varepsilon,x+d);(\varepsilon,d))=O(\|(\varepsilon,d)\|^2)$. When $\varepsilon=0$, similar to the discussion in Proposition 2.12, for any $V_i=(\beta_i,\alpha_i)\in\partial\mathcal{G}_i(0,x+d)$, $i=1,...,m$, we have $\alpha_i\in\partial\mathcal{F}_i(x+d)$. Then

$$\mathcal{G}_i(0,x+d)-\mathcal{G}_i(0,x)-V_i\begin{pmatrix}0\\d\end{pmatrix}=\mathcal{F}_i(x+d)-\mathcal{F}_i(x)-\alpha_i d=O(\|d\|^2),$$

if $\mathcal{F}$ is strongly semismooth on $\mathcal{R}^n$. Consequently, $\mathcal{G}$ is strongly semismooth on $(0,x)$, $x\in\mathcal{R}^n$.

□

Remark: in Theorem 2.16, whenever each L_i is a linear function, condition (i) and (ii) are not required.

2.2 supp(ϕ) is bounded

In this subchapter, we assume that supp(ϕ) is bounded. We can get some results analogous to those in Subchapter 2.1.

Proposition 2.17. *If $\mathcal{F}$ is directionally differentiable at $x \in \mathcal{R}^n$, then $\mathcal{G}$ is directionally differentiable at $(0, x)$ and for any $(\tau, h) \in \mathcal{R} \times \mathcal{R}^n$,*

$$\mathcal{G}'((0, x); (\tau, h)) = \int_{\mathcal{R}^n} \mathcal{F}'(x; h - |\tau| y)\Phi(y) d\mu(y).$$

The proof is similar to the proof of Proposition 2.11.

Theorem 2.18. *Suppose that Φ is continuously differentiable on $\mathcal{R}^n$. If $\mathcal{F}$ is semismooth on $\mathcal{R}^n$, then $\mathcal{G}$ is semismooth on $\mathcal{R}_+ \times \mathcal{R}^n$.*

The proof is similar to the proof of Theorem 2.12.

Theorem 2.19. *Suppose that Φ is continuously differentiable on $\mathcal{R}^n$. If $\mathcal{F}$ is strongly semismooth on $\mathcal{R}^n$, then $\mathcal{G}$ is strongly semismooth on $\mathcal{R}_+ \times \mathcal{R}^n$.*

Proof. Since $\mathcal{G}$ is continuously differentiable on $\mathcal{R}_{++} \times \mathcal{R}^n$ and $\mathcal{G}'(\varepsilon, x)$ is locally Lipschitz continuous on $\mathcal{R}_{++} \times \mathcal{R}^n$, $\mathcal{G}$ is strongly semismooth on $\mathcal{R}_{++} \times \mathcal{R}^n$. Then we only need to show that $\mathcal{G}$ is strongly semismooth at $(0, x)$, $x \in \mathcal{R}^n$. By [38, Theorem 3.3 (vii)], for any $(\varepsilon, d) \in \mathcal{R}_{++} \times \mathcal{R}^n$ with $(\varepsilon, d) \to 0$, we have

$$\mathcal{G}(\varepsilon, x + d) - \mathcal{G}(0, x) - \mathcal{G}'(\varepsilon, x + d)\begin{pmatrix} \varepsilon \\ d \end{pmatrix} = o(\|(\varepsilon, d)\|^2).$$

Moreover, for any $V_i = (\beta_i, \alpha_i) \in \partial \mathcal{G}_i(0, x + d)$, $i = 1, ..., m$, by [38, Theorem 3.3 (iv)], we have $\alpha_i \in \partial \mathcal{F}_i(x + d)$. Then

$$\mathcal{G}_i(0, x + d) - \mathcal{G}_i(0, x) - V_i\begin{pmatrix} 0 \\ d \end{pmatrix} = \mathcal{F}_i(x + d) - \mathcal{F}_i(x) - \alpha_i d = O(\|d\|^2),$$

if $\mathcal{F}$ is strongly semismooth on $\mathcal{R}^n$. Therefore, $\mathcal{G}$ is strongly semismooth on $(0, x)$, $x \in \mathcal{R}^n$. Then our result follows.

□

Chapter 3

Lipschitz Continuity, Continuous Differentiability and Directional Differentiability

In this chapter, we study locally Lipschitz continuity, continuous differentiability and directional differentiability of the smoothing function G defined by (1.2). Recall that

$$G(\varepsilon, X) = Q(X)\text{diag}[g(\varepsilon, \lambda_1(X)), ..., g(\varepsilon, \lambda_n(X))]Q(X)^T,$$

where $g : \mathcal{R} \times \mathcal{R} \mapsto \mathcal{R}$ is a smoothing function corresponding to $f : \mathcal{R} \mapsto \mathcal{R}$. The existence of such a smoothing function g for the locally Lipschitz continuous function f can be obtained via convolution, which has been discussed in Chapter 2.

It was shown in [5] that the matrix valued function F defined by (1.1) inherits many properties from the scalar valued function f. Shaprio [35] also obtained the similar results in a concise method. In this thesis, we will make an analogous study of the properties of G. According to the results obtained in Chapter 2, we know that under some conditions, the smoothing function g can have the properties of locally Lipschitz continuity, continuous differentiability, directional differentiability, semismoothness and strong semismoothness. We will prove that if g has these properties, then so does G. In particular, we will show that G inherits the properties of Lipschitz continuity, continuous differentiability and directional differentiability from g in this chapter. The properties of semismoothness and strong semismoothness of G will be discussed in the next chapter. Before we begin our discussion, let us first introduce some notations and review some properties.

For a matrix $X \in \mathcal{S}^n$, denote by $q_1(X), ..., q_n(X)$ a set of the orthonormal eigenvectors corresponding to the eigenvalues $\lambda_1(X), ..., \lambda_n(X)$. For a given matrix $\bar{X} \in \mathcal{S}^n$, we denote the multiplicities by $r_1, ..., r_q$. $\mu_1 > ... > \mu_q$ are the distinct values of the eigenvalues $\lambda_1(\bar{X}), ..., \lambda_n(\bar{X})$, i.e. $\mu_j := \lambda_{s_j+1}(\bar{X}) = ... = \lambda_{s_j+r_j}(\bar{X})$, $j = 1, .., q$, where $s_1 := 0, s_2 := r_1, ..., s_q := r_1 + ... + r_{q-1}$. $Q_j(X)$ is the $n \times r_j$ matrix whose columns are formed by the eigenvectors $q_{s_j+1}(X), ..., q_{s_j+r_j}(X)$, $j = 1, ..., q$, and define

$$P_j(X) := Q_j(X)Q_j(X)^T. \tag{3.1}$$

In particular, denote $\bar{q}_j := q_j(\bar{X})$, $\bar{Q}_j := Q_j(\bar{X})$ and $\bar{P}_j := P_j(\bar{X})$. Obviously, $\bar{P}_i\bar{P}_j = 0$ if $i \neq j$, $\bar{P}_i^2 = \bar{P}_i$ and $\sum_{j=1}^q \bar{P}_j = I_n$. Then the smoothing function $G(\varepsilon, \bar{X})$ defined in (1.2), can be written as

$$G(\varepsilon, \bar{X}) = \sum_{j=1}^{q} g(\varepsilon, \mu_j)\bar{P}_j.$$

It is known that the function $P_j(\cdot)$ defined by (3.1) is analytic in a neighborhood of $\bar{X} \in \mathcal{S}^n$, which is given as the following lemma.

Lemma 3.1. *(Shapiro [35]) The mapping $X \mapsto P_j(X)$ for any $j \in \{1, ..., q\}$ is analytic in a neighborhood of $\bar{X}$ and for any $H \in \mathcal{S}^n$,*

$$P_j'(\bar{X})H = \sum_{\substack{k \neq j \\ k=1}}^{q} (\mu_j - \mu_k)^{-1}(\bar{P}_j H \bar{P}_k + \bar{P}_k H \bar{P}_j). \tag{3.2}$$

The next proposition shows that $\lambda_i(X)$, $i = 1, ..., n$ are directionally differentiable, which is a particular case of [21, Theorem 7].

Proposition 3.2. *Given $\bar{X} \in \mathcal{S}^n$, for any $H \in \mathcal{S}^n$, the directional derivatives $\lambda'_{s_j+i}(\bar{X}; H)$, $i = 1, ..., r_j$, exist and coincide with the corresponding eigenvalues of the matrix $\bar{Q}_j^T H \bar{Q}_j$ arranged in the decreasing order.*

The following proposition shows the locally Lipschitz continuity of G. The conclusion of this proposition can be obtained by combining the techniques used in [5, Proposition 4.6] and [41, Lemma 2.3].

Proposition 3.3. *If g is locally Lipschitz continuous on $\mathcal{R} \times \mathcal{R}$, then G is locally Lipschitz continuous on $\mathcal{R} \times \mathcal{S}^n$.*

Proof. Consider $(\bar{\varepsilon}, \bar{X}) \in \mathcal{R} \times \mathcal{S}^n$. Since g is locally Lipschitz continuous on $\mathcal{R} \times \mathcal{R}$, by [32, Theorem 9.67], there exist continuously differentiable function $g_n : \mathcal{R} \times \mathcal{R} \mapsto \mathcal{R}$, $n = 1, 2, ...$, converging uniformly to g and satisfying

$$\|g'_n(\varepsilon, \xi)\| \leq \kappa, \quad \forall (\varepsilon, \xi) \in C, \tag{3.3}$$

where $\kappa > 0$ is a constant and $C := \cup_{i=1}^n [\lambda_i(\bar{X}) - \delta_i, \lambda_i(\bar{X}) + \delta_i] \times [\bar{\varepsilon} - \eta, \bar{\varepsilon} + \eta]$, for some $\delta_i > 0$, $i = 1, ..., n$ and $\eta > 0$.

Let $G_n(\varepsilon, X) := Q(X)\text{diag}[g_n(\varepsilon, \lambda_1(X)), ..., g_n(\varepsilon, \lambda_n(X))]Q(X)^T$.

Obviously, $\{G_n\}_{n=1}^\infty$ converges uniformly to G on $B(\bar{\varepsilon}, \bar{X})$. Fix (τ, Y), $(\nu, Z) \in B(\bar{\varepsilon}, \bar{X})$. Therefore, for any $M > 0$, there exists $N > 0$ such that for $n > N$, we have

$$\|G_n(\varepsilon, X) - G(\varepsilon, X)\| \leq M \|(\tau, Y) - (\nu, Z)\|, \text{ for all } (\varepsilon, X) \in B(\bar{\varepsilon}, \bar{X}).$$

By Proposition 3.4, for any $(\varepsilon, X) \in B(\bar{\varepsilon}, \bar{X})$ with $\varepsilon \neq 0$, $G_n(\varepsilon, X)$ is continuously differentiable for all n. Moreover, because of (3.3), there exists a scalar L such that $|||G'_n(\varepsilon, X)||| \leq L$, for all n.

Then we have, for any $(\tau, Y), (\nu, Z) \in B(\bar{\varepsilon}, \bar{X})$, where $\tau \neq 0$ and $\nu \neq 0$,

$$
\begin{aligned}
&\|G(\tau, Y) - G(\nu, Z)\| \\
=&\|G(\tau, Y) - G_n(\tau, Y) + G_n(\tau, Y) - G_n(\nu, Z) + G_n(\nu, Z) - G(\nu, Z)\| \\
\leq&\|G(\tau, Y) - G_n(\tau, Y)\| + \|G_n(\tau, Y) - G_n(\nu, Z)\| + \|G_n(\nu, Z) - G(\nu, Z)\| \\
\leq&2M\|(\tau, Y) - (\nu, Z)\| + \|\int_0^1 G_n'(\nu + t(\tau - \nu), Z + t(Y - Z))(\tau - \nu, Y - Z)dt\| \\
\leq&(2M + L)\|(\tau - \nu, Y - Z)\|.
\end{aligned}
$$

By a limiting process, the above inequality also holds for $\tau\nu = 0$. Hence, G is locally Lipschitz continuous on $\mathcal{R} \times \mathcal{S}^n$.

□

By [38, Theorem 3.3 (i) and Theorem 3.7 (i)], we know that under some assumptions, the smoothing function g is continuously differentiable around $(\varepsilon, x) \in \mathcal{R}_{++} \times \mathcal{R}$. Based on this fact, we will prove the continuous differentiability of G in the following proposition. Our proof uses the idea from Shapiro [35], in which the derivative of matrix valued function was given in a nice form.

Proposition 3.4. *Given $(\varepsilon, \bar{X}) \in \mathcal{R}_{++} \times \mathcal{S}^n$, if g is continuously differentiable around (ε, μ_j), $j = 1, ..., q$, then G is continuously differentiable around $(\varepsilon, \bar{X})$. Moreover, for any $(\tau, H) \in \mathcal{R}_{++} \times \mathcal{S}^n$, the derivative of G is given by*

$$
\begin{aligned}
G'(\varepsilon, \bar{X})\begin{pmatrix}\tau \\ H\end{pmatrix} &= G_X'(\varepsilon, \bar{X})H + G_\varepsilon'(\varepsilon, \bar{X})\tau \\
&= \mathbb{A}(\varepsilon, \bar{X}, H) + \mathbb{B}(\varepsilon, \bar{X}, H) + \mathbb{C}(\varepsilon, \bar{X}, \tau),
\end{aligned}
$$

where

$$
\begin{aligned}
\mathbb{A}(\varepsilon, \bar{X}, H) =& \sum_{j=1}^{q} g(\varepsilon, \mu_j)\bar{P}_j'(\bar{X})H \\
=&\frac{1}{2}\sum_{\substack{j \neq k \\ j,k=1}}^{q} \frac{g(\varepsilon, \mu_j) - g(\varepsilon, \mu_k)}{\mu_j - \mu_k}(\bar{P}_j H \bar{P}_k + \bar{P}_k H \bar{P}_j),
\end{aligned}
\tag{3.4}
$$

$$\mathbb{B}(\varepsilon,\bar{X},H)=\sum_{j=1}^{q}g'_x(\varepsilon,\mu_j)\bar{P}_jH\bar{P}_j, \tag{3.5}$$

and

$$\mathbb{C}(\varepsilon,\bar{X},\tau)=\sum_{j=1}^{q}g'_\varepsilon(\varepsilon,\mu_j)\tau\bar{P}_j. \tag{3.6}$$

Proof. First fix $\varepsilon > 0$. By using [35, Proposition 4.2], we know that $G(\varepsilon,\cdot)$ is continuously differentiable around $\bar{X}\in\mathcal{S}^n$ and for any $H\in\mathcal{S}^n$,

$$G'_X(\varepsilon,\bar{X})H=\mathbb{A}(\varepsilon,\bar{X},H)+\mathbb{B}(\varepsilon,\bar{X},H),$$

where $\mathbb{A}(\varepsilon,\bar{X},H)$ and $\mathbb{B}(\varepsilon,\bar{X},H)$ are given by (3.4) and (3.5) respectively.

For fixed $\bar{X}\in\mathcal{S}^n$, since $g(\cdot,\mu(j))$, $j=1,...,q$, are continuously differentiable on $\mathcal{R}_{++}$, $G(\cdot,\bar{X})$ is continuously differentiable on $\mathcal{R}_{++}$ and for any $\tau\in\mathcal{R}$, we have

$$G'_\varepsilon(\varepsilon,\bar{X})\tau=\mathbb{C}(\varepsilon,\bar{X},\tau),$$

where $\mathbb{C}(\varepsilon,\bar{X},\tau)$ is given by (3.6).

By Lemma 2.10 (iii) and locally Lipschitz continuity of λ_i, we can show that $g'_\varepsilon(\nu,\lambda_i(Z))\to g'_\varepsilon(\varepsilon,\lambda_i(\bar{X}))$, as $\nu\to\varepsilon$, $Z\to\bar{X}$. Then $|||G'_\varepsilon(\nu,Z)-G'_\varepsilon(\varepsilon,\bar{X})|||\to 0$, as $\nu\to\varepsilon$, $Z\to\bar{X}$. Hence, by the definition, we can get

$$G(\varepsilon+\tau,\bar{X}+H)-G(\varepsilon,\bar{X})-G'_\varepsilon(\varepsilon,\bar{X})\tau-G'_X(\varepsilon,\bar{X})H=o(\|(\tau,H)\|).$$

Thus, the derivative of G exists at $(\varepsilon,\bar{X})\in\mathcal{R}_{++}\times\mathcal{S}^n$. The continuity of $G'(\varepsilon,\bar{X})$ follows from the continuity of $G'_X(\varepsilon,\cdot)$ and $G'_\varepsilon(\cdot,X)$. Therefore, G is continuously differentiable on $\mathcal{R}_{++}\times\mathcal{S}^n$. $\square$

Proposition 3.4 not only verifies the continuous differentiability of G, but gives us the form of the derivative of G as well.

The following proposition shows that if g is directionally differentiable, so is G. Moreover, it also provides the explicit form of this directional derivative.

Proposition 3.5. *Given* $(\varepsilon, \bar{X}) \in \mathcal{R}_+ \times \mathcal{S}^n$*, if* g *is directionally differentiable at* $(\varepsilon, \lambda_j(\bar{X}))$*,* $j = 1, ..., n$*, then the directional derivative of* G *at* $(\varepsilon, \bar{X})$ *exists and is given by*

$$G'((\varepsilon, \bar{X}); (\tau, H)) = \mathbb{A}(\varepsilon, \bar{X}, H) + \mathbb{D}(\varepsilon, \bar{X}, \tau, H),$$

for any $(\tau, H) \in \mathcal{R} \times \mathcal{S}^n$ *and where* $\mathbb{A}(\varepsilon, \bar{X}, H)$ *is given by (3.4) and* $\mathbb{D}(\varepsilon, \bar{X}, \tau, H)$ *is defined by*

$$\mathbb{D}(\varepsilon, \bar{X}, \tau, H) := \sum_{j=1}^{n} g'\left((\varepsilon, \lambda_j(\bar{X})); (\tau, \lambda_j'(\bar{X}; H))\right) \bar{q}_j \bar{q}_j^T. \tag{3.7}$$

Proof. For any $(\nu, X) \in \mathcal{R}_+ \times \mathcal{S}^n$, consider the decomposition,

$$G(\nu, X) = \sum_{j=1}^{q} g(\varepsilon, \mu_j) P_j(X) + \sum_{j=1}^{n} \left(g(\nu, \lambda_j(X)) - g(\varepsilon, \lambda_j(\bar{X}))\right) q_j(X) q_j(X)^T. \tag{3.8}$$

Then for any $t > 0$, $(\tau, H) \in \mathcal{R} \times \mathcal{S}^n$, we have

$$t^{-1}\left(G(t|\tau| + \varepsilon, \bar{X} + tH) - G(\varepsilon, \bar{X})\right) = S + T,$$

where

$$S = t^{-1} \sum_{j=1}^{q} g(\varepsilon, \mu_j) \left(P_j(\bar{X} + tH) - \bar{P}_j\right)$$

and

$$T = t^{-1} \sum_{j=1}^{n} \left(g(t|\tau| + \varepsilon, \lambda_j(\bar{X} + tH)) - g(\varepsilon, \lambda_j(\bar{X}))\right) q_j(\bar{X} + tH) q_j(\bar{X} + tH)^T.$$

By Lemma 3.1, $\lim_{t \downarrow 0} S = \mathbb{A}(\varepsilon, \bar{X}, H)$, where $\mathbb{A}(\varepsilon, \bar{X}, H)$ is defined by (3.4).

Since g is directionally differentiable at $(\varepsilon, \lambda_j(\bar{X}))$, $j = 1, ..., n$, together with the directional differentiability of $\lambda_j(\bar{X})$, $j = 1, ..., n$, we obtain that for each j,

$$\begin{aligned} &\lim_{t \downarrow 0} t^{-1}\left(g(t|\tau| + \varepsilon, \lambda_j(\bar{X} + tH)) - g(\varepsilon, \lambda_j(\bar{X}))\right) \\ =&g'\left((\varepsilon, \lambda_j(\bar{X})); (\tau, \lambda_j'(\bar{X}; H))\right). \end{aligned}$$

Thus,

$$\lim_{t \downarrow 0} T = \mathbb{D}(\varepsilon, \bar{X}, \tau, H).$$

Hence, G is directionally differentiable at $(\varepsilon, \bar{X})$ and the directional derivative of G is given by

$$G'\big((\varepsilon, \bar{X}); (\tau, H)\big) = \mathbb{A}(\varepsilon, \bar{X}, H) + \mathbb{D}(\varepsilon, \bar{X}, \tau, H).$$

□

Chapter 4

Semismoothness and Strong Semismoothness

In Theorem 2.12, Theorem 2.16, Theorem 2.18 and Theorem 2.19, we have proved that the smoothing function g can be semismooth (strongly semismooth) on $\mathcal{R}_+ \times \mathcal{R}$ under some conditions. In this chapter, we will show that semismoothness and strong semismoothness are inherited by G from g.

Theorem 4.1. *Given $(\varepsilon, \bar{X}) \in \mathcal{R}_+ \times \mathcal{S}^n$, if g is semismooth at (ε, μ_j), $j = 1, ..., p$, then G is semismooth at $(\varepsilon, \bar{X})$.*

The proof of the semismoothness of G is quite similar to the proof of strong semismoothness. We omit the detail here. Please refer to the proof of Theorem 4.2.

Theorem 4.2. *Given $(\varepsilon, \bar{X}) \in \mathcal{R}_+ \times \mathcal{S}^n$, if g is strongly semismooth at (ε, μ_j), $j = 1, ..., p$, then G is strongly semismooth at $(\varepsilon, \bar{X})$.*

Proof. Since g is strongly semismooth at (ε, μ_j), $j = 1, ..., p$, by the definition, we know that g is locally Lipschitz continuous in the neighborhood of (ε, μ_j), $j = 1, ..., p$ and g is directionally differentiable at (ε, μ_j), $j = 1, ..., p$. Then by Proposition 3.3 and Proposition 3.5, G is locally Lipschitz continuous around $(\varepsilon, \bar{X})$

and G is directionally differentiable at $(\varepsilon, \bar{X})$. Next, we will show that for any $(\tau, H) \in \mathcal{R}_+ \times \mathcal{S}^n$ with $(\tau, H) \to 0$,

$$G(\varepsilon+\tau, \bar{X}+H) - G(\varepsilon, \bar{X}) - G'((\varepsilon+\tau, \bar{X}+H); (\tau, H)) = O(\|(\tau, H)\|^2).$$

By the decomposition as (3.8), we have

$$\begin{aligned} &G(\varepsilon+\tau, \bar{X}+H) - G(\varepsilon, \bar{X}) \\ &\quad= \sum_{j=1}^{q} g(\varepsilon, \mu_j)(P_j(\bar{X}+H) - \bar{P}_j) \\ &\quad\quad+ \sum_{j=1}^{n} \left(g(\varepsilon+\tau, \lambda_j(\bar{X}+H)) - g(\varepsilon, \lambda_j(\bar{X}))\right) q_j(\bar{X}+H) q_j(\bar{X}+H)^T. \end{aligned}$$

Since P_j, $j = 1, ..., q$, are twice continuously differentiable at $\bar{X}$, we have

$$P_j(\bar{X}+H) - \bar{P}_j = P_j'(\bar{X}+H)H + O(\|H\|^2), \quad j = 1, ..., n.$$

Since g is strongly semismoonth at (ε, μ_j), $j = 1, ..., p$, together with the strong semismoothness of λ_j, $j = 1, ..., n$, the composition functions $g(\varepsilon, \lambda_j(X))$, $j = 1, ..., n$ are also strongly semismooth at $(\varepsilon, \bar{X})$ (see [11] and [23, Theorem 5].) Then

$$\begin{aligned} &g\left(\varepsilon+\tau, \lambda_j(\bar{X}+H)\right) - g\left(\varepsilon, \lambda_j(\bar{X})\right) \\ &\quad= g'\left((\varepsilon+\tau, \lambda_j(\bar{X}+H)); (\tau, \lambda_j'(\bar{X}+H; H))\right) + O(\|(\tau, H)\|^2). \end{aligned}$$

By noticing the fact that $\|q_j(\bar{X}+H) q_j(\bar{X}+H)^T\|$ are uniformly bounded, we get

$$\begin{aligned} &G(\varepsilon+\tau, \bar{X}+H) - G(\varepsilon, \bar{X}) \\ &\quad= \sum_{j=1}^{q} g(\varepsilon, \mu_j) P_j'(\bar{X}+H)H \\ &\quad\quad+ \sum_{j=1}^{n} g'\left((\varepsilon+\tau, \lambda_j(\bar{X}+H)); (\tau, \lambda_j'(\bar{X}+H; H))\right) q_j(\bar{X}+H) q_j(\bar{X}+H)^T \\ &\quad\quad+ O(\|(\tau, H)\|^2). \end{aligned}$$

According to Proposition 3.5, we have

$$G'\left((\varepsilon+\tau,\bar{X}+H);(\tau,H)\right)=\mathbb{A}(\varepsilon+\tau,\bar{X}+H,H)+\mathbb{D}(\varepsilon+\tau,\bar{X}+H,\tau,H),$$

where $\mathbb{A}(\varepsilon+\tau,\bar{X}+H,H)$ and $\mathbb{D}(\varepsilon+\tau,\bar{X}+H,\tau,H)$ are given by (3.4) and (3.7) respectively. Moreover, it is known that

$$\sum_{j=1}^{q} g(\varepsilon,\mu_j)P_j'(\bar{X}+H)H=\mathbb{A}(\varepsilon+\tau,\bar{X}+H,H)+O(\|(\tau,H)\|^2),$$

and

$$\begin{aligned}&\sum_{j=1}^{n} g'\left((\varepsilon+\tau,\lambda_j(\bar{X}+H));(\tau,\lambda_j'(\bar{X}+H;H))\right)q_j(\bar{X}+H)q_j(\bar{X}+H)^T\\ =&\mathbb{D}(\varepsilon+\tau,\bar{X}+H,\tau,H).\end{aligned}$$

Then it follows that

$$G(\varepsilon+\tau,\bar{X}+H)-G(\varepsilon,\bar{X})-G'\left((\varepsilon+\tau,\bar{X}+H);(\tau,H)\right)=O(\|(\tau,H)\|^2).$$

The proof is completed. □

Example 4.3. *For the CHKS smoothing function,* $\phi(x)=\dfrac{2}{(x^2+4)^{3/2}}$, $x\in\mathcal{R}$. *Then we have* $G(\varepsilon,X)=\dfrac{\sqrt{4\varepsilon^2I+X^2}+X}{2}$, $(\varepsilon,X)\in\mathcal{R}_{++}\times\mathcal{S}^n$, *with infinity supp*$(\phi)$. *Obviously,* f *and* ϕ *satisfy the assumptions and the conditions in Theorem 2.16. Then* g *corresponding to* $f(x)=x_+$ *is strongly semismooth on* $\mathcal{R}_+\times\mathcal{R}$. *Thus, by Theorem 4.2,* $G(\varepsilon,X)$ *is strongly semismooth on* $\mathcal{R}_+\times\mathcal{S}^n$. *Sun, Sun and Qi [41] have used another method to prove the strong semismoothness of this smoothing function. Also see [7], [40].*

Example 4.4. *For the extreme value smoothing function, since* $\phi(x)=\dfrac{1}{\sqrt{\pi}}e^{-x^2}$, $x\in\mathcal{R}$, *for any* $(\varepsilon,X)\in\mathcal{R}_{++}\times\mathcal{S}^n$, *we have* $G(\varepsilon,X)=\dfrac{1}{2}X+\dfrac{\varepsilon}{2\sqrt{\pi}}e^{-\frac{X^2}{\varepsilon^2}}+U(\varepsilon,X)$, *where* $e^{-\frac{X^2}{\varepsilon^2}}$ *is the matrix exponential of* $-\dfrac{X^2}{\varepsilon^2}$ *and*

$$U(\varepsilon,X)=\sum_{i=1}^{n}\frac{\lambda_i(X)}{2}u(\frac{\lambda_i(X)}{\varepsilon})q_i(X)q_i(X)^T,$$

where u is defined by (2.4). Here supp(ϕ) is also infinity. Since f and ϕ satisfy the assumptions and conditions in Theorem 4.2, g corresponding to $f(x) = x_+$ is strongly semismooth on $\mathcal{R}_+ \times \mathcal{R}$. Then by using Theorem 4.2, G is strongly semismooth on $\mathcal{R}_+ \times \mathcal{S}^n$.

Chapter 5

Some Applications of the Smoothing Function

In this chapter, we will extend the smoothing function to the vector valued function with second order cone and the matrix valued function over nonsymmetric matrices.

The second order cone (SOC), also called the Lorentz cone, in $\mathcal{R}^n$, is defined by

$$\mathcal{K}^n = \{(x_1, x_2^T)^T | x_1 \in \mathcal{R}, x_2 \in \mathcal{R}^{n-1}, \text{and} \quad x_1 \geq \|x_2\|\}.$$

For convenience, we write $x = (x_1, x_2)$ instead of $x = (x_1, x_2^T)^T$. Denote $e = (1, 0, ..., 0)^T \in \mathcal{R}^n$.

Any $x = (x_1, x_2) \in \mathcal{R} \times \mathcal{R}^{n-1}$ can be decomposed as

$$x = \lambda_1 \mu^{(1)} + \lambda_2 \mu^{(2)},$$

where λ_1, λ_2 and $\mu^{(1)}$, $\mu^{(2)}$ are *spectral values* and the associated *spectral vectors* of x, with respect to $\mathcal{K}^n$, given by

$$\lambda_i = x_1 + (-1)^i \|x_2\|$$

and

$$\mu^{(i)} = \begin{cases} \frac{1}{2}(1, (-1)^i \frac{x_2}{\|x_2\|}) & \text{if } x_2 \neq 0 \\ \frac{1}{2}(1, (-1)^i \omega) & \text{otherwise} \end{cases}$$

for $i = 1, 2$ and ω is any vector in $\mathcal{R}^{n-1}$, satisfying $\|\omega\| = 1$. In [12], for any function $f : \mathcal{R} \to \mathcal{R}$, the following vector valued function associated with $\mathcal{K}^n$ is introduced:

$$f^{soc}(x) := f(\lambda_1)\mu^{(1)} + f(\lambda_2)\mu^{(2)}.$$

The interest of this function is stemmed from the second-order-cone complementarity problem (SOCCP). See [4], [8], [12], [15] and the references therein for the study of the smoothing method for solving SOCCP. Analogous to the matrix valued function, we will construct a smoothing function of $f^{soc}(x)$.

Assume that f is locally Lipschitz continuous. Let $e = (1, 0, ..., 0)^T \in \mathcal{R}$. For any $x = (x_1, x_2) \in \mathcal{R} \times \mathcal{R}^{n-1}$, we define $L(x)$ and $\tilde{L}(x_2)$ by

$$L(x) = \begin{bmatrix} x_1 & x_2^T \\ x_2 & x_1 I \end{bmatrix}$$

$$\tilde{L}(x_2) = \begin{bmatrix} 0 & 0^T \\ 0 & I - x_2 x_2^T / \|x_2\|^2 \end{bmatrix}.$$

In [4, Lemma 4.1], Chen, Chen and Tseng showed that for any $t \in \mathcal{R}$,

$$f^{soc}(x) = F(L(x) + t\tilde{L}(x_2))e,$$

where F is the matrix valued function defined by (1.1) and the eigenvalues of $L(x) + t\tilde{L}(x_2)$ are λ_1, λ_2 and $x_1 + t$ of multiplicity $n - 2$. This result is the key to relating f^{soc} to F. [4] showed that $f^{soc}(x)$ inherits from f the properties of Lipschitz continuity, continuous differentiability, (strong) semismoothness etc. We define the smoothing function $g^{soc}(\varepsilon, x) : \mathcal{R}_{++} \times \mathcal{R}^n \mapsto \mathcal{R}^n$ corresponding to f^{soc} by

$$g^{soc}(\varepsilon, x) := G(\varepsilon, L(x) + t\tilde{L}(x_2))e, \tag{5.1}$$

where G is the smoothing function of F defined by (1.2), $g^{soc}(0, x) = f^{soc}(x)$, for any $\varepsilon < 0$, $g^{soc}(\varepsilon, x) = g^{soc}(-\varepsilon, x)$. Chapter 3 and Chapter 4 have shown that the properties of locally Lipschitz continuity , continuous differentiability, directional differentiability, semismoothness and strong semismoothness are inherited by G from g. Based on the relationship between G and g^{soc}, Theorem 5.1 proves that g^{soc} also inherits these properties from g.

Theorem 5.1. *(a) If g is locally Lipschitz continuous on $\mathcal{R} \times \mathcal{R}$, then g^{soc} is locally Lipschitz continuous on $\mathcal{R} \times \mathcal{R}^n$;*

(b) If g is continuously differentiable on $\mathcal{R}_{++} \times \mathcal{R}$, then g^{soc} is continuously differentiable on $\mathcal{R}_{++} \times \mathcal{R}^n$;

(c) If g is directionally differentiable at (ε, λ_1), (ε, λ_2) with $\varepsilon \geq 0$, then g^{soc} is directionally differentiable at $(\varepsilon, x) \in \mathcal{R}_+ \times \mathcal{R}^n$;

(d) If g is semismooth on $\mathcal{R}_+ \times \mathcal{R}$, then g^{soc} is semismooth on $\mathcal{R}_+ \times \mathcal{R}^n$;

(e) If g is strongly semismooth on $\mathcal{R}_+ \times \mathcal{R}$, then g^{soc} is strongly semismooth on $\mathcal{R}_+ \times \mathcal{R}^n$.

Proof. (a) Let $t = 0$. By Proposition 3.3, G is locally Lipschitz continuous around $(\varepsilon, L(x))$ with $(\varepsilon, x) \in \mathcal{R} \times \mathcal{R}$. Since $L(x)$ is locally Lipschitz continuous around x, the relationship (5.1) yields that g^{soc} is locally Lipschitz continuous on $\mathcal{R} \times \mathcal{R}^n$.

(b) Let $t = 0$. Analogous to the proof of (a), this result follows directly form Proposition 3.4, the fact that $L(x)$ is continuously differentiable at $x \in \mathcal{R}^n$ and the relationship (5.1) .

(c) Let $t = \|x_2\|$. Then by [4, Lemma 4.1 (a)], for any $x \in \mathcal{R}^n$, the eigenvalues of $L(x) + \|x_2\|\tilde{L}(x_2)$ is λ_1 and λ_2 of multiplicity $n-1$. Then by proposition 3.5, G is directionally differentiable at $(\varepsilon, L(x) + \|x_2\|\tilde{L}(x_2))$ with $(\varepsilon, x) \in \mathcal{R}_+ \times \mathcal{R}^n$, if g is directionally differentiable at (ε, λ_1) and (ε, λ_2). Since $L(x) + \|x_2\|\tilde{L}(x_2)$ is continuously differentiable at x, the result follows.

(d)-(e) Let $t = 0$. By (Theorem 4.2) Theorem 4.1, G is (strongly) semismooth at $(\varepsilon, L(x))$ with $(\varepsilon, x) \in \mathcal{R}_+ \times \mathcal{R}^n$. Since $L(x)$ is (strongly) semismooth, by [23, Theorem 5] and [11], the composite function $g^{soc}(\varepsilon, x) = G(\varepsilon, L(x))e$ is still (strongly) semismooth.

$\square$

Let $\mathcal{M}_{n,m}$ be the space of $n \times m$ real matrices. Without loss of generality, we assume $n \leq m$. For any scalar valued function $f : \mathcal{R} \mapsto \mathcal{R}$, in [40], Sun and Sun introduced a matrix valued function over nonsymmetric matrices defined by singular values. $N : \mathcal{M}_{n,m} \mapsto \mathcal{S}^n$ is defined by

$$N(A) := U\text{diag}(f(\sigma_1(A)), ..., f(\sigma_n(A)))U^T, \ A \in \mathcal{M}_{n,m},$$

where $\sigma_1(A) \geq \sigma_2(A) \geq ... \geq \sigma_n(A) \geq 0$ are the singular values of A, and

$$U[\text{diag}(\sigma_1(A), ..., \sigma_n(A)) \ 0]V^T$$

is the singular value decomposition of A, where $V \in \mathcal{O}_m$, $U \in \mathcal{O}_n$(see [17, p 415].)

In [40], Sun and Sun obtained the relationship between the matrix valued function N defined over nonsymmetric matrices and the matrix valued function F defined over symmetric matrices:

$$N(A) = \pi(F(\Xi(A))),$$

where $\pi : \mathcal{S}^{n+m} \mapsto \mathcal{S}^n$ is defined by $(\pi(X))_{ij} := X_{ij}, \ i, j = 1, ..., n, \ X \in \mathcal{S}^{n+m}$ and $\Xi : \mathcal{M}_{n,m} \mapsto \mathcal{S}^{n \times m}$ is defined by $\Xi(A) := \begin{bmatrix} 0 & A \\ A^T & 0 \end{bmatrix}$, $A \in \mathcal{M}_{n,m}$. By [17, Theorem 7.3.7], we know that the eigenvalues of $\Xi(A)$ are $\pm\sigma_i(A)$, $i = 1, ..., n$ and 0 of multiplicity $m - n$.

Then the smoothing function of N can be constructed as $M : \mathcal{R} \times \mathcal{M}_{n,m} \mapsto \mathcal{S}^n$

$$M(\varepsilon, A) = \pi(G(\varepsilon, \Xi(A))), \quad (\varepsilon, A) \in \mathcal{R}_{++} \times \mathcal{M}_{n,m}, \tag{5.2}$$

where G is the smoothing function corresponding to F defined by (1.2), $M(0,A) = N(A)$, for any $\varepsilon < 0$, $M(\varepsilon, X) = M(-\varepsilon, X)$.

We combine the relationship (5.2) between M and G and the relationship between g and G discussed in Chapter 3 and Chapter 4. Then in the following theorem, we make an analogous study for the smoothing function M.

Theorem 5.2. *(a) If g is locally Lipschitz continuous on $\mathcal{R} \times \mathcal{R}$, then M is locally Lipschitz continuous on $\mathcal{R} \times \mathcal{M}_{n,m}$;*

(b) If g is continuously differentiable on $\mathcal{R}_{++} \times \mathcal{R}$, then M is continuously differentiable on $\mathcal{R}_{++} \times \mathcal{M}_{n,m}$;

(c) If g is directionally differentiable at $(\varepsilon, \pm\sigma_i(A))$, $i = 1, ..., n$ and $(\varepsilon, 0)$ with $\varepsilon \geq 0$, then M is directionally differentiable at (ε, A) with $(\varepsilon, A) \in \mathcal{R}_+ \times \mathcal{M}_{n,m}$;

(d) If g is semismooth on $\mathcal{R}_+ \times \mathcal{R}$, then M is semismooth on $\mathcal{R}_+ \times \mathcal{M}_{n,m}$;

(e) If g is strongly semismooth on $\mathcal{R}_+ \times \mathcal{R}$, then M is strongly semismooth on $\mathcal{R}_+ \times \mathcal{M}_{n,m}$.

Proof. (a)For (τ, B) and (μ, C) in the neighborhood of $B(\varepsilon, A) \in \mathcal{R} \times \mathcal{M}_{n,m}$. Then we have

$$\begin{aligned}
\|M(\tau,B) - M(\mu,C)\| &= \|\pi[G(\tau,\Xi(B)) - G(\mu,\Xi(C))]\| \\
&\leq \|G(\tau,\Xi(B)) - G(\mu,\Xi(C))\| \\
&\leq L(|\tau-\mu| + \|\Xi(B) - \Xi(C)\|) \\
&= L(|\tau-\mu| + \sqrt{2\|B-C\|^2}) \\
&\leq 2L(|\tau-\mu| + \|B-C\|),
\end{aligned}$$

where the second inequality is followed by the locally Lipschitz continuity of G (see Proposition 3.3.) Then the result follows.

(b) By Proposition 3.4, G is continuously differentiable at $(\varepsilon, \Xi(A))$, with $\varepsilon > 0$ and $A \in \mathcal{M}_{n,m}$. Since the linear mapping is continuously differentiable, by the relationship (5.2), M is continuously differentiable around $(\varepsilon, A) \in \mathcal{R}_{++} \times \mathcal{M}_{n,m}$.

(c) For any $(\varepsilon, A) \in \mathcal{R}_{+} \times \mathcal{M}_{n,m}$, it is known that the eigenvalues of $\Xi(A)$ are $\pm\sigma_i(A)$, $i = 1, ..., n$ and 0 of multiplicity $m - n$. By Proposition 3.5, G is directionally differentiable at $(\varepsilon, \Xi(A))$. By relationship (5.2), M is directionally differentiable at $(0, A)$.

(d)-(e) By (Theorem 4.2) Theorem 4.1, G is (strongly) semismooth on $\mathcal{R}_{+} \times \mathcal{M}_{n,m}$. By using the fact that linear mapping is (strongly) semismooth, the composition function $M(\varepsilon, A) = \pi(G(\varepsilon, \Xi(A)))$ is (strongly) semismooth on $\mathcal{R}_{+} \times \mathcal{M}_{n,m}$.

$\square$

Chapter 6

Final Remarks

We have completed the analysis of the properties of the generalized smoothing function of the vector valued function. In particular, we studied the (strong) semismoothness of the smoothing function. Based on theses properties, we studied the smoothing function of the matrix valued function. We found out that the smoothing function of the matrix valued function inherits some properties from the scalar valued function corresponding to it. With these useful results, we extend the smoothing function to the vector valued function with second order cone and matrix valued function over nonsymmetric matrices. However, in this thesis, we cannot prove that when $\text{supp}(\phi)$ is infinite, $\mathcal{G}$ can inherit the strong semismoothness from $\mathcal{F}$, if $\mathcal{F}$ is strongly semismooth. Further study can be done in this aspect.

Bibliography

[1] B. Chen and P. T. Harker, "A non-interior-point continuation method for linear complementarity problems," *SIAM Journal on Matrix Analysis and Applications*, **14** (1993) 1168-1190.

[2] C. Chen and O. L. Mangasarian, "Smoothing methods for convex inequalities and linear complementarity problems," *Mathematical Programming,* **71** (1995) 51-69.

[3] C. Chen and O. L. Mangasarian, "A class of smoothing functions for nonlinear and mixed complementarity problems," *Computational Optimization and Applications,* **5** (1996) 97-138.

[4] J. Chen, X. Chen and P. Tseng, "Analysis of nonsmooth vector-valued functions associated with second- order cones," *Mathematical Programming,* **101** (2004) 95-117.

[5] X. Chen, H. Qi and P. Tseng, "Analysis of nonsmooth symmetric-matrix-valued functions with applications to semidefinite complementarity problems," *SIAM Journal on Optimization,* **13** (2003) 960-985.

[6] X. Chen, L. Qi and D. Sun, "Global and superlinear convergence of the smoothing Newtown method and its application to general box constrained variational inequalities," *Mathematics of Computation,* **67** (1998) 519-540.

[7] X. Chen and P. Tseng, "Non-interior continuation methods for solving semidefinite complementarity problems," *Mathematical Programming,* **95** (2003) 431-474.

[8] X. D. Chen, D. Sun and J. Sun, "Complementarity functions and numetical experiments for second-order-cone complementarity problems," *Computational Optimization and Applications,* **25** (2003) 39-56.

[9] A. L. Dontchev, H. Qi and L. Qi, "Convergence of Newton's method for convex best interpolation," *Numerische Mathematik,* **87** (2001) 435-456.

[10] Y. M. Ermoliev, V. I. Norkin and R. J.-B. Wets, "The minimization of semicontinuous functions: mollifier subgradients," *SIAM Journal on Control and Optimization,* **33** (1995) 149-167.

[11] A. Fischer, "Solution of monotone complementarity problems with locally Lipschitzian functions," *Mathematical Programming,* **76** (1997) 513-532.

[12] M. Fukushima, Z. Q. Luo and P. Tseng, "Smoothing functions for second-order-cone complementarity problem," *SIAM Journal on Optimization,* **12** (2002) 436-460.

[13] S. A. Gabriel and J. J. Moré, "Smoothing of mixed complementarity problems," *Complementarity and Variational Problems: State of the Art*, 105-116, M. C. Ferries and J. S. Pang, eds. SIAM, Philadelphia, PA, 1997.

[14] A. M. Gupal, "On a method for the minimization of almost differentiable functions," *Kibernetika*, (1997) 114-116.

[15] S. Hayashi, N. Yamashita and M. Fukushima, "A Combined Smoothing and Regularization Method for Monotone Second-Order-Cone Complementarity Problems," to appear in SIAM Journal on Optimization.

[16] R. A. Horn and C. R. Johnson, *Topics in Matrix Analysis*, Cambridge Universtiy Press, Cambridge, United Kindom, 1991.

[17] R. A. Horn and C. R. Johnson, *Matrix Analysis*, Cambridge Press, Cambridge, United Kingdom, 1985.

[18] F. Jones, *Lebesgue Integration on Euclidean Space*, Jones and Bartlett Publishers, London, 1993.

[19] C. Kanzow and C. Nagel, "Semidefinite programs: new search directions, smoothing-type methods, and numerical results," *SIAM Journal on Optimization,* **13** (2002) 1-23.

[20] C. Kanzow, "Some noninterior continuation methods for linear complementarity problems," *SIAM Journal on Matrix Analysis and Applications*, **17** (1996) 851-868.

[21] P. Lancaster, " On eigenvalues of matrices dependent on a parameter," *Numerische Mathematik,* **6** (1964) 377-387.

[22] D. Q. Mayne and E. Polak, "Nondifferential optimization via adaptiv smoothing," *Journal of Optimization Theory and Applications,* **43** (1984) 601-613.

[23] R. Mifflin, "Semismooth and semiconvex function in constrained optimization," *SIAM Journal on Control and Optimiztion,* **15** (1977) 957-972.

[24] J. S. Pang and F. Facchinei, *Finite-Dimensional Variational Inequalities and Complementarity Problems*, Volumes I and II, Springer-Verlag, New York, 2003.

[25] L. Qi and X. Chen, "A globally convergent successive approximation method for severely nonsmooth euqations," *SIAM Journal on Control and Optimization,* **33** (1995) 402-418.

[26] L. Qi, A. Shapiro and C. Ling, "Differentiability and semismoothness properites of integral functions and their applications," Mathematical Programming, to appear.

[27] L. Qi, D. Sun and G. Zhou, "A new look at smoothing Newton methods for nonlinear complementarity problems and box constrained variational inequalities," *Mathematical Programming,* **87** (2002) 1-35.

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

[29] L. Qi and P. Tseng, "An analysis of piecewise smooth functions and almost smooth functions," Technical Report, Department of Mathematics, University of Washington, July, 2002.

[30] L. Qi and H. Yin, "A strongly semismooth integral function and its application," *Computational Optimization and Applications*, **25** (2003) 223-246.

[31] R. T. Rockafellar, *Convex Analysis,* Princetion University Press, Princeton, New Jersey, 1970.

[32] R. T. Rockafellar and R. J. -B. Wets, *Variational Analysis*, Springer-Verlag, Berlin, 1998.

[33] H. L. Royden, *Real Analysis,* 3rd Edition, Collier-Macmillan Limited, London, 1988.

[34] L. Schwartz, *Théorie des Distributions,* Hermann, Paris, 1966.

[35] A. Shapiro,"On differentiability of symmetric matrix valued functions," Published electronically in *Optimization Online*, 2002.

[36] H. S. Shapiro, *Smoothing and Approximation of Functions,* Van Nostrand Reinhold Company, New York, 1969.

[37] S. Smale, "Algorithms for solving equations," in *Proceeding of the International Congress of Mathematicians*, Berkely, California, 1986, 172-195.

[38] D. Sun and L. Qi, "Solving variational inequality problems via smoothing-nonsmooth reformulations," Nonlinear programming and variational inequalities (Kowloon, 1998), *Journal of Computational and Applied Mathematics,* **129** (2001) 37-62.

[39] D. Sun and J. Sun, "Semismooth matrix valued functions," *Mathematics of Operations Research,* **27** (2002) 150-169.

[40] D. Sun and J. Sun, "Nonsmooth matrix valued functions defined by singular values", Technical Report, Department of Mathematics, National University of Singapore, 2002.

[41] J. Sun, D. Sun and L. Qi, "A squared smoothing Newton method for nonsmooth matrix equations and its applications in semidefinite optimization problems," *SIAM Journal on Optimization,* **14** (2004) 783-806.

[42] S. Scholtes, Introduction to piecewise differentiable equations, Habilitation thesis, University of Karlsruhe, Germany, 1994.

[43] P. Tseng, "Analysis of a non-interior continuation method based on Chen-Mangasarian smoothing functions for complementarity problems," in

Reformulation-Nonsmooth, Piecewise Smooth, Semismooth and smoothing Methods edited by M.Fukushima and L.Qi, Kluwer Academic Publishers, Boston, 1999, 381-404.

[44] I. Zang, "A smoothing-out technique for min-max optimization," *Mathematical Programming,* **19** (1980) 61-77.