Journal of Global Optimization (2006) 35: 653–673
DOI 10.1007/s10898-005-5880-3


# On the Semismoothness of Projection Mappings and Maximum Eigenvalue Functions

MARK GOH[1] and FANWEN MENG[2]
[1]*Department of Decision Sciences, National University of Singapore, 1 Business Link, Singapore 117592, Republic of Singapore (e-mail: bizgohkh@nus.edu.sg.)*
[2]*School of Mathematics, University of Southampton, Highfield Southampton SO17 1BJ, UK (e-mail: f.meng@soton.ac.uk.)*

(Received 13 September 2005; accepted in revised form 13 December 2005)

**Abstract.** This paper analyses the properties of the projection mapping over a set defined by a constraint function whose image is possibly a nonpolyhedral convex set. Under some nondegeneracy assumptions, we prove the (strong) semismoothness of the projection mapping. In particular, we derive the strong semismoothness of the projection mapping when the nonpolyhedral convex set under consideration is taken to be the second-order cone or the semidefinite cone. We also derive the semismoothness of the solution to the Moreau–Yosida regularization of the maximum eigenvalue function.

**Key words:** Projection mappings, semismooth functions, semidefinite cones, maximum eigenvalue functions, Moreau–Yosida regularization.

## 1. Introduction

Consider the following parametric optimization problem

$$\begin{aligned} \min\ & \tfrac{1}{2}\langle x-y, x-y\rangle, \quad y \in X, \\ \text{s.t.}\ & G(x) \in K, \\ & x \in X, \end{aligned} \tag{1}$$

where $G: X \to Y$ is a single-valued continuously differentiable mapping, $X$ and $Y$ are finite dimensional vector spaces each equipped with a scalar product denoted by $\langle \cdot, \cdot \rangle$ and its induced norm denoted by $\| \cdot \|$, $K \subseteq Y$ is a closed (possibly nonpolyhedral) convex set. Let $G^{-1}(K)$ denote the feasible set of (1), i.e.,

$$G^{-1}(K) = \{x \in X: G(x) \in K\}. \tag{2}$$

The solution mapping of (1) can be formulated as the projection mapping over $G^{-1}(K)$, namely,

$$\Pi_{G^{-1}(K)}(y) = \operatorname{argmin} \left\{ \tfrac{1}{2}\langle x-y, x-y\rangle \mid G(x) \in K, \quad x \in X \right\}, \tag{3}$$

where $\Pi_D(z)$ denotes the metric projection of $z$ onto $D$ for any closed convex set $D$ of the vector space $Z$ and $z \in Z$.

It is well known that strong regularity for generalized equations [19] and strong stability for the Karush–Kuhn–Tucker (KKT) system of nonlinear optimization problems [10] are two important concepts in perturbation analysis of optimization problems. Much progress on this topic have been made in the past few years, see [1, 4, 20] for instance. Recent work on the sensitivity analysis of generalized equations (GEs) with nonpolyhedral set constraints can be found in [12, 21, 24] where locally Lipschitz continuity [21] and semismoothness [12] of the solution mappings of perturbed GEs are derived.

In [13], Mifflin introduced an important subclass of Lipschitz functions – semismooth functions. In order to study the superlinear convergence of Newton's method for solving nondifferentiable equations, Qi and Sun [18] extended the definition of semismoothness to vector valued functions. After the work of Qi and Sun [18], semismoothness was extensively used to establish superlinear/quadratic convergence of Newton's method for solving the complementarity problem and variational inequalities, the convex best interpolation problem, and the inverse eigenvalue problem [26].

Properties, such as continuity and differentiability, of projection mappings have played an important role in optimization and are of its own interest as well [1, 2, 4, 11, 16, 20]. This paper investigates the properties of the projection mapping in form of (3). Under a blanket *nondegeneracy* assumption, which is introduced for abstract spaces [1], we derive the $G$-semismoothness (semismoothness, strong semismoothness) of this projection. In particular, we obtain the strong semismoothness of the projection mapping when $K$ is any one of the two important nonpolyhedral convex sets: (i) the second order cone (Lorentz cone); (ii) the cone of symmetric positive semidefinite matrices. We then apply the established results to the maximum eigenvalue problem, where the maximum eigenvalue function is convex and is usually not differentiable [15, 23]. It is known that the minimization of the nondifferentiable convex function can be converted to the minimization of its Moreau–Yosida regularization [8, 6, 14, 27]. Since the semismoothness of the gradient of the regularization function has played a key role in studying the superlinear convergence of the generalized Newton's method for solving the induced Moreau–Yosida regularization problem [6, 18], we investigate the semismoothness of the Moreau–Yosida regularization of the maximum eigenvalue function by studying the semismoothness of the projection mapping on the epigraph of the maximum eigenvalue function.

The rest of this paper is organized as follows. Section 2 presents some basic definitions and notations. Section 3 studies the semismoothness sensitivity analysis of projection mappings. Discussions on second-order cones are done in Section 4. Section 5 discusses the properties of the solution to Moreau–Yosida regularization of the maximum eigenvalue function.

## 2. Preliminaries

Let $X$ and $Y$ be finite dimensional vector spaces each equipped with a scalar product denoted by $\langle\cdot,\cdot\rangle$ and its induced norm denoted by $\|\cdot\|$. Let $\mathcal{O}$ be an open set in $X$ and $\Phi\colon\mathcal{O}\subseteq X\to Y$ be a locally Lipschitz continuous function on $\mathcal{O}$. By Rademacher's theorem, $\Phi$ is almost everywhere F-differentiable in $\mathcal{O}$. We denote by $\mathcal{D}_\Phi$ the set of points in $\mathcal{O}$ where $\Phi$ is F-differentiable. Let $J\Phi(x)$, a linear mapping from $X$ to $Y$, denote the derivative of $\Phi$ at $x\in\mathcal{O}$ if $\Phi$ is F-differentiable at $x$, and $J\Phi(x)^*\colon Y\to X$ the adjoint of $J\Phi(x)$. The B(ouligand)-subdifferential of $\Phi$ at $x\in\mathcal{O}$, denoted by $\partial_B\Phi(x)$, is then the set of $V$ such that

$$V=\lim_{k\to\infty}J\Phi(x^k),$$

where $\{x^k\}\in\mathcal{D}_\Phi$ is a sequence converging to $x$. The Clarke's generalized Jacobian of $\Phi$ at $x$ is the convex hull of $\partial_B\Phi(x)$ (see [2]), i.e.,

$$\partial\Phi(x)=\operatorname{conv}\{\partial_B\Phi(x)\}.$$

Semismoothness was originally introduced by Mifflin [13] for functionals. To study the superlinear convergence of Newton's method for solving nonsmooth equations, Qi and Sun [18] extended the definition of semismoothness to vector valued functions. There are several equivalent ways of defining semismoothness. Here we adopt the following definition.

DEFINITION 2.1. Let $\Phi\colon\mathcal{O}\subseteq X\to Y$ be a locally Lipschitz continuous function on the open set $\mathcal{O}$. $\Phi$ is semismooth at a point $x\in\mathcal{O}$ if

(i) $\Phi$ is directionally differentiable at $x$; and
(ii) for any $\Delta x\to 0$ and $V\in\partial\Phi(x+\Delta x)$,

$$\Phi(x+\Delta x)-\Phi(x)-V(\Delta x)=o(\|\Delta x\|). \tag{4}$$

Gowda [7] denotes a locally Lipschitz continuous function $\Phi$ "semismooth" at $x$ if (4) holds. To distinguish Gowda's definition on semismoothness of $\Phi$ at $x$, Pang et al. [17] defined $\Phi$ as $G$-semismooth at $x$ if (4) holds. In this paper, we will follow this terminology. A stronger notion than semismoothness is strong semismoothness. We say that $\Phi$ is strongly semismooth at $x$, if $\Phi$ is semismooth at $x$ and for any $\Delta x\to 0$ and $V\in\partial\Phi(x+\Delta x)$,

$$\Phi(x+\Delta x)-\Phi(x)-V(\Delta x)=O(\|\Delta x\|^2). \tag{5}$$

In particular, we state that $\Phi$ is strongly $G$-semismooth at $x$ if (5) holds. Likewise, $\Phi$ is $G$-semismooth (strongly $G$-semismooth, semismooth, strongly semismooth) on set $D \subseteq \mathcal{O}$ if $\Phi$ is $G$-semismooth (strongly $G$-semismooth, semismooth, strongly semismooth) at every point of $D$.

Let $\Omega \subseteq X$ be a closed convex set. Consider the projection mapping $\Pi_\Omega$: $X \to X$

$$\Pi_\Omega(y) = \operatorname*{argmin}_{x\in\Omega} \tfrac{1}{2}\langle x-y, x-y\rangle, \quad y \in X. \tag{6}$$

It is well known that $\Pi_\Omega(\cdot)$ is contractive, that is,

$$\|\Pi_\Omega(y_1) - \Pi_\Omega(y_2)\| \leqslant \|y_1 - y_2\|$$

for any $y_1, y_2 \in X$. Then, $\Pi_\Omega(\cdot)$ is globally Lipschitz continuous with module 1. Hence, by Rademacher's Theorem, $\Pi_\Omega$ is differentiable almost everywhere in $X$, and $\partial\Pi_\Omega(\cdot)$ is well defined on $X$. The following results can be found in [12, Proposition 1]:

PROPOSITION 2.1. *Let $\Omega \subseteq X$ be a closed convex set. Then, for any $x \in X$ and $V \in \partial\Pi_\Omega(x)$, the following statements hold:*

(i) *$V$ is self-adjoint*;
(ii) *$\langle d, Vd\rangle \geqslant 0$, for any $d \in X$*;
(iii) *$\langle Vd, d - Vd\rangle \geqslant 0$, for any $d \in X$.*

## 3. Semismoothness of Projection Mappings

We now investigate the parameterized optimization problem (1). In particular, we consider two cases: (i) $G$ is an affine mapping from $X$ to $Y$; (ii) $G$ is twice continuously differentiable from $X$ to $Y$. For a closed convex set $K \subset Y$, let $T_K(y)$ denote the tangent cone of $K$ at $y$ and $\mathrm{lin}(C)$ the lineality space of the closed convex cone $C$, i.e., $\mathrm{lin}(C) = C \cap (-C)$.

To study the semismoothness of the solution mapping of (1)

$$\Pi_{G^{-1}(K)}(y) = \operatorname*{argmin}_{G(x)\in K} \left\{\tfrac{1}{2}\langle x-y, x-y\rangle\right\},$$

we introduce the notion of *nondegeneracy* with respect to (1) taken from Bonnans and Shapiro [1], which is a blanket assumption in this paper.

DEFINITION 3.1. $\bar{x} \in G^{-1}(K)$ is nondegenerate, with respect to the mapping $G$ and set $K$, if

$$JG(\bar{x})X + \mathrm{lin}(T_K(\bar{z})) = Y, \tag{7}$$

where $\bar{z} := G(\bar{x})$.

It is known that the above nondegeneracy condition is reduced to the usual linear independent constraint qualification (LICQ) in nonlinear programming (NLP). Consider the Lagrangian function of problem (1)

$$L(x,\Lambda,y)=\tfrac{1}{2}\langle x-y,x-y\rangle+\Lambda^*(-G(x)), \tag{8}$$

where $(y,\Lambda,x)\in X\times Y\times X$. Under some standard constraint qualifications (such as Robinson's qualification [19]), we obtain the first order necessary condition of (1) below,

$$\begin{cases} \nabla_x L(x,\Lambda,y)=x-y-JG(x)^*\Lambda=0,\\ \Lambda\in K,\ -G(x)\in -K,\ \Lambda^*(-G(x))=0. \end{cases} \tag{9}$$

According to Eaves [3], the above equations can be written equivalently as:

$$\begin{cases} x-y-JG(x)^*\Lambda=0,\\ \Lambda-\Pi_K[\Lambda-G(x)]=0. \end{cases} \tag{10}$$

For convenience in description, we write (10) as the following generalized equations

$$\mathcal{H}(x,\Lambda;y):=\begin{bmatrix} x-JG(x)^*\Lambda-y\\ \Lambda-\Pi_K[\Lambda-G(x)] \end{bmatrix}=0. \tag{11}$$

By Shapiro [21], under the nondegeneracy condition (7), the multiplier $\Lambda$ satisfying (9) ((10), or (11)) is unique. Then, given $\bar{y}$, let $\bar{x}$ be the solution of (1) with the unique multiplier $\bar{\Lambda}$. Hence, $(\bar{x},\bar{\Lambda})$ is the unique solution of (11), i.e.,

$$\mathcal{H}(\bar{x},\bar{\Lambda};\bar{y})=0.$$

### 3.1. CASE I: AFFINE MAPPING $G$

We consider the case where $G$ is an affine mapping from $X$ to $Y$. We make the following assumption:

ASSUMPTION 3.1. 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 $Y$.

We now make some comments on this assumption. Consider two examples in NLP:

EXAMPLE 3.1. Consider

$$\begin{aligned} &\min \tfrac{1}{2}\|x-y\|^2 \\ &\text{s.t.} \quad x \geqslant 0, x \in \mathcal{R}^2. \end{aligned}$$

Here, $G(x)=x$, $K=\mathcal{R}^2_+$. When we choose $\bar{y}=0$, it is easy to see that the corresponding solution $\bar{x}=0$ and $G(\bar{x})=0$. First,

$$\nabla G_1(x)=\begin{pmatrix}1\\0\end{pmatrix}, \quad \nabla G_2(x)=\begin{pmatrix}0\\1\end{pmatrix}.$$

Clearly, LICQ holds at $\bar{x}$. In addition,

$$[\mathrm{lin}(T_K(G(\bar{x})))]^{\perp}=\mathcal{R}^2.$$

So, Assumption 3.1 holds in this case. This example shows that under LICQ (nondegeneracy condition), Assumption 3.1 is satisfied at $\bar{x}$ automatically. □

EXAMPLE 3.2. Consider

$$\begin{aligned} &\min \tfrac{1}{2}\|x-y\|^2 \\ &\text{s.t.} \quad 1-x_1-x_2 \geqslant 0, \\ &\qquad\; 1-x_1+x_2 \geqslant 0, \quad x \in \mathcal{R}^2. \end{aligned}$$

Here $K=\mathcal{R}^2_+$ and $G(x)=(1-x_1-x_2, 1-x_1+x_2)^T$. We now choose $\bar{y}=(3/2, 1/64)$. Note that LICQ holds at any $x \in \mathcal{R}^2$. After some operations, we obtain

$$\bar{x}=\begin{pmatrix}1\\0\end{pmatrix}, \quad \bar{\Lambda}=\begin{pmatrix}\frac{17}{32}\\ \frac{1}{2}\end{pmatrix}, \quad G(\bar{x})=\begin{pmatrix}0\\0\end{pmatrix},$$

and $\bar{\Lambda}-G(\bar{x})=\bar{\Lambda}\in \mathrm{int}\mathcal{R}^2_+$, which implies that $\partial\Pi_K(\bar{\Lambda}-G(\bar{x}))$ is reduced to a singleton with

$$\partial\Pi_K(\bar{\Lambda}-G(\bar{x}))=\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}.$$

Apparently, Assumption 3.1 does not hold at $\bar{x}$. □

According to the above examples, Assumption 3.1 is closely related to the nondegeneracy condition. In Sections 4 and 5, we show that Assumption 3.1 is redundant in the analysis if the set $K$ is chosen as the second order cone or the semidefinite cone.

Next, by virtue of Assumption 3.1, we derive the main result with respect to the projection mapping $\Pi_{G^{-1}(K)}(\cdot)$ as defined in (3).

THEOREM 3.1. *Given* $\bar{y} \in X$, *let* $(\bar{x}, \bar{\Lambda})$ *be the corresponding KKT solution of (9). Suppose (i)* $\bar{x}$ *is nondegenerate with respect to the mapping G and set K; (ii)* $G: X \to Y$ *is an affine mapping; (iii) Assumption 3.1 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* $\Pi_K$ *is G-semismooth (semismooth, strongly G-semismooth, strongly semismooth) around* $\bar{\Lambda} - G(\bar{x})$, *then* $(x(\cdot), \Lambda(\cdot))$ *is G-semismooth (semismooth, strongly G-semismooth, strongly semismooth) around* $\bar{y}$.

*Proof.* First, we show that $\partial_{(x,\Lambda)}\mathcal{H}(\bar{x}, \bar{\Lambda}; \bar{y})$ is nonsingular. Let $W$ be any element taken from $\partial_{(x,\Lambda)}\mathcal{H}(\bar{x}, \bar{\Lambda}; \bar{y})$. Let $(\Delta x, \Delta\Lambda) \in X \times Y$ be such that

$$W \begin{bmatrix} \Delta x \\ \Delta \Lambda \end{bmatrix} = 0.$$

According to Clarke [2], there exists $V \in \partial \Pi_K(\bar{\Lambda} - G(\bar{x}))$ such that

$$\begin{cases} \Delta x - JG(\bar{x})^* \Delta\Lambda = 0 \\ \Delta\Lambda - V(\Delta\Lambda - JG(\bar{x})\Delta x) = 0. \end{cases} \tag{12}$$

Set $\Delta H = \Delta\Lambda - JG(\bar{x})\Delta x$. It follows from the second equation of (12) that

$$\Delta H + JG(\bar{x})\Delta x = \Delta\Lambda = V(\Delta H). \tag{13}$$

So,

$$\Delta H - V(\Delta H) + JG(\bar{x})\Delta x = 0.$$

Hence, it yields that

$$\langle V(\Delta H), \Delta H - V(\Delta H)\rangle + \langle V(\Delta H), JG(\bar{x})\Delta x\rangle = 0.$$

By Proposition 2.1, we have

$$\langle V(\Delta H), \Delta H - V(\Delta H)\rangle \geqslant 0,$$

which implies

$$\langle V(\Delta H), JG(\bar{x})\Delta x\rangle \leqslant 0.$$

Thus, by virtue of (13) and the first equation in (12), we have

$$\begin{aligned}&\langle V(\Delta H), JG(\bar{x})\Delta x\rangle = \langle \Delta\Lambda, JG(\bar{x})\Delta x\rangle \\ &\quad = \langle JG(\bar{x})^*\Delta\Lambda, \Delta x\rangle = \langle JG(\bar{x})^*\Delta\Lambda, JG(\bar{x})^*\Delta\Lambda\rangle \\ &\quad \leqslant 0.\end{aligned}$$

Therefore,

$$\Delta x = JG(x)^*\Delta\Lambda = 0. \tag{14}$$

Next, we show that $\Delta\Lambda = 0$. For any $z \in X$, it follows from (14) that

$$\langle \Delta\Lambda, JG(\bar{x})z\rangle = \langle JG(\bar{x})^*\Delta\Lambda, z\rangle = 0.$$

So,

$$\Delta\Lambda \in [JG(\bar{x})X]^{\perp}.$$

However, by the second equation in (12) and (14), we have

$$(I - V)(\Delta\Lambda) = 0.$$

Hence, by Assumption 3.1

$$\Delta\Lambda \in [\mathrm{lin}(T_K(G(\bar{x}))]^{\perp}.$$

Therefore,

$$\Delta\Lambda \in [JG(\bar{x})X]^{\perp} \cap [\mathrm{lin}(T_K(G(\bar{x}))]^{\perp}.$$

Noticing that the nondegeneracy condition (7) is equivalent to

$$[JG(\bar{x})X]^{\perp} \cap [\mathrm{lin}(T_K(G(\bar{x}))]^{\perp} = \{0\},$$

hence,

$$\Delta\Lambda = 0. \tag{15}$$

Thus, by (14) and (15), $W$ is nonsingular. Consequently, $\partial_{(x,\Lambda)}\mathcal{H}(\bar{x}, \bar{\Lambda}; \bar{y})$ is nonsingular.

Part (i). According to the above arguments, the result follows from the implicit function theorem for locally Lipschitz continuous functions directly [2].

Part (ii). The result follows immediately from the established results in [7] and [25] of implicit function theorems for $G$-semismooth functions and semismooth functions. □

### 3.2. CASE II: TWICE CONTINUOUSLY DIFFERENTIABLE MAPPING $G$

We now consider the case where $G: X \to Y$ is a twice continuously differentiable mapping. Note that (11) can be rewritten as

$$\bar{\mathcal{H}}(x, \Lambda; y) := \begin{bmatrix} x - y + JG(x)^* \Lambda \\ G(x) - \Pi_K[G(x) + \Lambda] \end{bmatrix} = 0. \tag{16}$$

By using the cone reducibility notion, Shapiro [21] reduced the discussion on sensitivity analysis of $(x(\cdot), \Lambda(\cdot))$ to a new problem, making the sensitivity analysis simpler. We adopt the same idea here.

DEFINITION 3.2. A closed (not necessarily convex) set $C \subseteq Y$ is cone reducible at a point $y_0 \in C$ if there exists a neighborhood $\mathcal{V} \subseteq Y$ of $y_0$, a pointed closed convex cone $Q$ in a finite dimensional space $Z$ and a twice continuously differentiable mapping $\Theta: \mathcal{V} \to Z$ such that: (i) $\Theta(y_0) = 0 \in Z$, (ii) the derivative mapping $J\Theta(y_0): Y \to Z$ is onto, and (iii) $C \cap \mathcal{V} = \{y \in \mathcal{V} | \Theta(y) \in Q\}$. If $C$ is cone reducible at every point $y_0 \in C$ (possibly to a different cone $Q$), then $C$ is cone reducible.

It is known that many interesting sets such as the polyhedral convex set, the second-order cone, and the cone $S^n_+$ of positive semidefinite $n \times n$ symmetric matrices are all cone reducible [21]. In the subsequent analysis, we assume that the convex set $K$ is cone reducible at the point $\bar{z} := G(\bar{x})$ to a pointed closed convex cone $Q \subseteq Z$ by a mapping $\Theta$. Define the mapping $\mathcal{G}(x) := \Theta(G(x))$. Then, for all $(x, y)$ in a neighborhood of $(\bar{x}, \bar{y})$, (16) can be written as

$$\bar{\mathcal{H}}_{\mathcal{G}}(x, \mu; y) := \begin{bmatrix} x - y + J\mathcal{G}(x)^* \mu \\ \mathcal{G}(x) - \Pi_Q[\mathcal{G}(x) + \mu] \end{bmatrix} = 0 \tag{17}$$

in the sense that $(x(y), \Lambda(y))$ is a solution of (16) iff $(x(y), \mu(y))$ is a solution of (17) and

$$\Lambda(y) = [J\Theta(G(x(y)))]^* \mu(y). \tag{18}$$

Moreover, by Definition 3.2, we can derive that for $(x, y)$ sufficiently close to $(\bar{x}, \bar{y})$, the multiplier $\mu(y)$ is defined uniquely by (18). Hence, in what follows, we only need to study the sensitivity of the solution of (17) near $\bar{y}$. By Definition 3.2, we have $\mathcal{G}(\bar{x}, \bar{y}) = 0$ with the unique multiplier $\bar{\mu}$.

THEOREM 3.2. *Given $\bar{y} \in X$. Let $\bar{x}$ be the unique solution of (1) associated with $\bar{\mu}$. Suppose $G$ is nondegenerate at $\bar{x} \in G^{-1}(K)$ with respect to $K$, and $K$*

*is cone reducible at the point $G(\bar{x})$ to a pointed closed convex cone $Q \subseteq Z$ by a mapping $\Theta$. Let*

$$\mathcal{D} := \left\{ d \in X \mid J\mathcal{G}(\bar{x})d \in \mathcal{C}_{\Pi_Q(\mathcal{G}(\bar{x})+\bar{\mu})} \right\}$$

*and*

$$\mathcal{C}_{\Pi_Q(\mathcal{G}(\bar{x})+\bar{\mu})} = \left\{ Vh \mid V \in \partial \Pi_Q(\mathcal{G}(\bar{x}) + \bar{\mu}), h \in Z \right\}.$$

*If for all $d \in \mathcal{D} \backslash \{0\}$,*

$$\left\langle d, (J^2\mathcal{G}(\bar{x})^* \bar{\mu})(d) \right\rangle > 0. \tag{19}$$

*Then, the following statements hold:*

(i) *there exists an open neighborhood $\mathcal{N}$ of $\bar{y}$ and a Lipschitz continuous function $(x(\cdot), \mu(\cdot))$ defined on $\mathcal{N}$ such that $\bar{\mathcal{H}}\mathcal{G}(x(y), \mu(y); y) = 0$ for every $y \in \mathcal{N}$; Moreover, we have $\Lambda(y)$ defined by*

$$\Lambda(y) = [J\Theta(G(x(y)))]^* \mu(y)$$

*such that*

$$\bar{\mathcal{H}}(x(y), \Lambda(y), y) = 0, \forall y \in \mathcal{N}.$$

(ii) *if $\Pi_Q$ is G-semismooth (semismooth) around $\mathcal{G}(\bar{x}) + \bar{\mu}$, then $(x(y), \mu(y))$ is G-semismooth (semismooth) around $\bar{y}$.*

(iii) *if $\Pi_Q$ is strongly G-semismooth (strongly semismooth) around $\mathcal{G}(\bar{x}) + \bar{\mu}$ and the second derivative of $\mathcal{G}$ are locally Lipschitz continuous near $\bar{x}$, then $(x(y), \mu(y))$ is strongly G-semismooth (strongly semismooth) around $\bar{y}$.*

*Proof.* We first investigate the nonsingularity of $\partial_{(x,\mu)}\bar{\mathcal{H}}_{\mathcal{G}}(\bar{x}, \bar{\mu}; \bar{y})$. Let $A \in \partial_{(x,\mu)}\bar{\mathcal{H}}_{\mathcal{G}}(\bar{x}, \bar{\mu}; \bar{y})$, and let $(\delta x, \delta\mu) \in X \times Z$ be such that

$$A \begin{bmatrix} \delta x \\ \delta \mu \end{bmatrix} = 0. \tag{20}$$

For a given $A$, there exists $V \in \partial \Pi_Q[\mathcal{G}(\bar{x}) + \bar{\mu}]$ such that

$$\begin{aligned} 0 &= \left(J^2\mathcal{G}(\bar{x})^* \bar{\mu}\right) \delta x + J\mathcal{G}(\bar{x})^* \delta\mu, \\ 0 &= J\mathcal{G}(\bar{x})\delta x - V(J\mathcal{G}(\bar{x})\delta x + \delta\mu). \end{aligned}$$

Let $q := J\mathcal{G}(\bar{x})\delta x + \delta\mu$. Then, $q - \delta\mu = J\mathcal{G}(\bar{x})\delta x = Vq$. Hence, $\langle Vq, q - Vq\rangle - \langle Vq, \delta\mu\rangle = 0$. By Proposition 2.1, $\langle Vq, q - Vq\rangle \geqslant 0$, then $\langle Vq, \delta\mu\rangle \geqslant 0$, from which we have

$$0 \leqslant \langle J\mathcal{G}(\bar{x})\delta x, \delta\mu\rangle = \left\langle \delta x, J\mathcal{G}(\bar{x})^*\delta\mu\right\rangle = -\left\langle \delta x, (J^2\mathcal{G}(\bar{x})^*\bar{\mu})\delta x\right\rangle, \tag{21}$$

which implies

$$\left\langle \delta x, (J^2\mathcal{G}(\bar{x})^*\bar{\mu})\delta x\right\rangle \leqslant 0. \tag{22}$$

Since $\delta x \in \mathcal{D}$, then by assumption (19), it follows that $\delta x = 0$. Consequently, we have

$$J\mathcal{G}(\bar{x})^*\delta\mu = 0.$$

Note that as $Q$ is a pointed closed convex cone, by assumption and from [21], $J\mathcal{G}(\bar{x})$ is now onto. Thus, $\delta\mu = 0$ and $\partial_{(x,\mu)}\bar{\mathcal{H}}_{\mathcal{G}}(\bar{x}, \bar{\mu}; \bar{y})$ is nonsingular.

Also, note that $\bar{\mathcal{H}}_{\mathcal{G}}(\bar{x}, \bar{\mu}, \bar{y}) = 0$. Hence, by the implicit function theorem and with the help of (18), statement (i) follows directly. Statements (ii) and (iii) follow directly by [25, Theorem 2.1]. □

## 4. Second-Order Cone

We now consider the case when $K$ in (1) is a second-order cone (SOC) in the form of

$$K = \mathcal{K}^n := \{(z, \tau) \in \mathcal{R}^n \times \mathcal{R} : \|z\| \leqslant \tau\}, \tag{23}$$

where $X = Y = \mathcal{R}^n \times \mathcal{R}$. We shall show that Assumption 3.1 holds. For $\bar{y} \in \mathcal{R}^n \times \mathcal{R}$, under the nondegeneracy condition, let $\bar{x}$ be the solution of (1) with the unique multiplier $\bar{\Lambda}$ as considered previously. We now investigate the following three cases:

(i) $G(\bar{x}) \in \operatorname{int} K$. Obviously, now, $\bar{\Lambda} = 0$. It follows that $\Pi_K(-G(\bar{x})) = 0$. Note that $-G(\bar{x}) \in \operatorname{int}(-K)$, we have $V = 0$ for any $V \in \partial\Pi_K(-G(\bar{x}))$, and $\operatorname{lin}(T_K(G(\bar{x}))] = \mathcal{R}^n \times \mathcal{R}$. Thus, Assumption 3.1 holds.

(ii) $G(\bar{x}) = 0$. We have $T_K(G(\bar{x})) = K$, which implies that $\operatorname{lin}(T_K(G(\bar{x}))) = \{0\}$. Thus, $[\operatorname{lin}(T_K(G(\bar{x}))]^{\perp} = \mathcal{R}^n \times \mathcal{R}$. Again, Assumption 3.1 holds.

(iii) $G(\bar{x}) \in \operatorname{bd} K \setminus \{0\}$. For convenience in description, let $(\bar{z}, \bar{\tau}) = G(\bar{x})$ where $\bar{\tau} = ||\bar{z}||$ and $\bar{\Lambda} = (u, t) \in \mathcal{R}^n \times \mathcal{R}$. Then,

$$\operatorname{lin} T_K(G(\bar{x})) = \left\{ \alpha \begin{pmatrix} \bar{z} \\ \bar{\tau} \end{pmatrix} : \alpha \in \mathcal{R} \right\}. \tag{24}$$

Since

$$\begin{bmatrix} u \\ t \end{bmatrix} = \Pi_K\left[\begin{pmatrix} u \\ t \end{pmatrix} - \begin{pmatrix} \bar{z} \\ \bar{\tau} \end{pmatrix}\right],$$

then according to [5], there exists $\beta \in \mathcal{R}$ such that $u = \beta\bar{z}$ and $t = -\beta||\bar{z}||$. Thus,

$$\Pi_K\left[\begin{pmatrix} u \\ t \end{pmatrix} - \begin{pmatrix} \bar{z} \\ \bar{\tau} \end{pmatrix}\right] = \Pi_K\begin{bmatrix} (\beta-1)\bar{z} \\ -(1+\beta)||\bar{z}|| \end{bmatrix} = \begin{bmatrix} \beta\bar{z} \\ -\beta||\bar{z}|| \end{bmatrix},$$

and it is not difficult to verify that

$$\beta = -1.$$

So, it follows that

$$\Pi_K\left[\begin{pmatrix} u \\ t \end{pmatrix} - \begin{pmatrix} \bar{z} \\ \bar{\tau} \end{pmatrix}\right] = \Pi_K\begin{bmatrix} -2\bar{z} \\ 0 \end{bmatrix} = \begin{bmatrix} -\bar{z} \\ ||\bar{z}|| \end{bmatrix}.$$

We have

$$\Pi_K'\left[\begin{pmatrix} -2\bar{z} \\ 0 \end{pmatrix}\right] = \frac{1}{2}\begin{bmatrix} I + 2\bar{z}\bar{z}^T/||\bar{z}||^2 & -\bar{z}/||\bar{z}|| \\ -\bar{z}^T/||\bar{z}|| & 1 \end{bmatrix}.$$

Now, consider $h = (h_1, h_2) \in \mathcal{R}^n \times \mathcal{R}$ satisfying

$$\left(I_{n+1} - \Pi_K'\left[\begin{pmatrix} -2\bar{z} \\ 0 \end{pmatrix}\right]\right) h = 0.$$

Then,

$$\begin{bmatrix} h_1 \\ h_2 \end{bmatrix} = \frac{1}{2}\begin{bmatrix} h_1 \\ h_2 \end{bmatrix} + \frac{1}{2}\begin{bmatrix} 2\bar{z}(\bar{z}^T h_1) - (\bar{z}/||\bar{z}||)h_2 \\ -\bar{z}^T h_1/||\bar{z}|| + h_2 \end{bmatrix},$$

which leads to

$$\begin{bmatrix} h_1 \\ h_2 \end{bmatrix} = \begin{bmatrix} \bar{z}(2\bar{z}^T h_1 - h_2/||\bar{z}||) \\ \\ h_2 - \bar{z}^T h_1/||\bar{z}|| \end{bmatrix}.$$

Thus,

$$\begin{cases} \bar{z}^T h_1 = 0 \\ h_1 = -(h_2/||\bar{z}||)\bar{z}. \end{cases}$$

Hence,

$$\begin{aligned}\langle h, (\bar{z}, \|\bar{z}\|)\rangle &= h_1^T \bar{z} + h_2\|\bar{z}\| \\ &= -h_2/\|\bar{z}\|\|\bar{z}\|^2 + h_2\|\bar{z}\| = 0.\end{aligned}$$

By virtue of (24), we have $h \in [\mathrm{lin} T_K(\bar{z}, \bar{\tau})]^{\perp}$. Again, Assumption 3.1 holds. Based on the above arguments, we obtain the following proposition.

PROPOSITION 4.1. *Let $K$ be the second-order cone of $\mathcal{R}^n \times \mathcal{R}$ in* (1). *For any given $\bar{y} \in R^n \times \mathcal{R}$, let $(\bar{x}, \bar{\Lambda})$ be the corresponding KKT solution of* (9). *If*

$$(I_{n+1} - V)h = 0$$

*for $h \in R^{n+1}$ and $V \in \partial\Pi_K(\bar{\Lambda} - G(\bar{x}))$, then $h \in [\mathrm{lin}(T_K(G(\bar{x})))]^{\perp}$.*

By virtue of Theorem 3.1 and Proposition 4.1, we derive the following conclusion.

THEOREM 4.1. *Let $K$ be the second-order cone $\mathcal{K}^n$ of $\mathcal{R}^n \times \mathcal{R}$ and $G: \mathcal{R}^n \times \mathcal{R} \to \mathcal{R}^n \times \mathcal{R}$ an affine mapping in* (1). *Let $\bar{y} \in \mathcal{R}^n \times \mathcal{R}$ be given and $(\bar{x}, \bar{\Lambda})$ the KKT solution of* (9). *Under the nondegeneracy condition, the projection mapping $x(\cdot)$ is strongly semismooth near $\bar{y}$.*

*Proof.* $\Pi_{\mathcal{K}^n}$ is strongly semismooth everywhere by [16]. Hence, the proof follows immediately from Proposition 4.1 and Theorem 3.1. □

## 5. Maximum Eigenvalue Problem

Consider the maximum eigenvalue function

$$\bar{f}(x) = \lambda_{\max}(A_0 + \mathcal{A}(x)), \tag{25}$$

where $\mathcal{A}(x) = \sum_{i=1}^m x_i A_i$ and $A_0, A_1, \ldots, A_m \in S^n$, $S^n$ denotes the set of symmetric matrices in $\mathcal{R}^{n\times n}$. It is known that $\bar{f}: \mathcal{R}^m \to \mathcal{R}$ is a convex and usually nondifferentiable function. Hence, minimizing the maximum eigenvalue is a nonsmooth optimization problem in general. We now consider the Moreau [14] and Yosida [27] regularization of $\bar{f}$ defined by

$$\begin{aligned}\hat{f}_\epsilon(y) := \min \ & \left\{\bar{f}(x) + \frac{\epsilon}{2}\langle y - x, y - x\rangle\right\} \\ \text{s.t. } & x \in \mathcal{R}^m,\end{aligned} \tag{26}$$

where $\epsilon$ is a positive number. Then, by [8, 20], minimizing the maximum eigenvalue problem

$$\min\{\bar{f}(x) \mid x \in \mathcal{R}^m\}. \tag{27}$$

is equivalent to solving the regularized problem

$$\min\{\hat{f}_\epsilon(x) \mid x \in \mathcal{R}^m\}. \tag{28}$$

It is well known that $\hat{f}_\epsilon$ is continuously differentiable [8] on $\mathcal{R}^m$ with

$$\nabla \hat{f}_\epsilon(y) = \epsilon(y - x(y)), \quad y \in \mathcal{R}^m, \tag{29}$$

where $x(\cdot)$ represents the unique optimal solution of (26). Also, $x(\cdot)$ is globally Lipschitz continuous, which implies that $\nabla \hat{f}_\varepsilon$ is globally Lipschitz continuous [20, p. 546]. Here, we are interested in the semismoothness of $x(\cdot)$ near a given point $\bar{y} \in \mathcal{R}^m$, which is a key condition for the superlinear convergence of algorithms based on generalized Newton methods as designed in Fukushima and Qi [6] for solving nonsmooth convex optimization problems.

The following analysis is based on the structure of the epigraph of the maximum eigenvalue function $\bar{f}$. The epigraph of $\bar{f}$, denoted by $\Omega$, can be written as

$$\begin{aligned}\Omega := \mathrm{epi}(\bar{f}) &= \{(x,t) \in \mathcal{R}^m \times \mathcal{R} : t \geqslant \lambda_{\max}(A_0 + \mathcal{A}(x))\}\\ &= \{(x,t) \in \mathcal{R}^m \times \mathcal{R} : tI - [A_0 + \mathcal{A}(x)] \succeq 0\},\end{aligned}$$

where for $M \in \mathcal{R}^{n\times n}$, $M \succeq 0$ represents a symmetric positive semidefinite matrix $M$. Then, the projection $\Pi_\Omega[(y,\tau)]$ is the unique optimal solution of the following parameterized problem:

$$\begin{array}{ll}\min & \frac{1}{2}\left\{||x-y||^2 + (t-\tau)^2\right\}\\ \text{s.t.} & tI - [A_0 + \mathcal{A}(x)] \succeq 0,\\ & (x,t) \in \mathcal{R}^m \times \mathcal{R}.\end{array} \tag{30}$$

Let $S_+^n$ denote the set of symmetric positive semidefinite $n \times n$ matrices. For simplicity in description, set

$$\begin{aligned}&K := S_+^n,\ G(x,t) := tI - [A_0 + \mathcal{A}(x)],\\ &f((x,t);(y,\tau)) := \tfrac{1}{2}\left[||x-y||^2 + (t-\tau)^2\right].\end{aligned}$$

Then, (30) can be rewritten as

$$\begin{array}{ll}\min & \frac{1}{2}\langle (x,t)-(y,\tau), (x,t)-(y,\tau)\rangle,\\ \text{s.t.} & G(x,t) \in K,\ (x,t) \in \mathcal{R}^m \times \mathcal{R}.\end{array} \tag{31}$$

which is as the same form as (1). Evidently, $G$ is an affine mapping from $\mathcal{R}^m \times \mathcal{R} \to S^n$.

Next, we show that, under the nondegeneracy condition, Assumption 3.1 holds automatically. For any two matrices $A, B \in S^n$, let $\langle A, B\rangle$ denote the matrix Frobenius inner product between $A$ and $B$ as follows:

$$\langle A, B\rangle = A \bullet B = \mathrm{trace}(AB) = \mathrm{trace}(BA),$$

where trace$(\cdot)$ denotes the trace of a matrix. Note that for any orthogonal matrix $Q$,

$$\langle Q^T A Q, Q^T B Q\rangle = \langle A, B\rangle.$$

LEMMA 5.1. For $(y,\tau) \in \mathcal{R}^m \times \mathcal{R}$, let $(x,t) = \Pi_\Omega[(y,\tau)]$. If $(y,\tau) \notin \Omega, t = \lambda_{\max}(A_0 + \mathcal{A}(x))$.

*Proof.* Consider the KKT system of (30)

$$\begin{cases} \begin{bmatrix} x \\ t \end{bmatrix} - \begin{bmatrix} y \\ \tau \end{bmatrix} + \begin{bmatrix} \langle D\mathcal{A}, \Lambda\rangle \\ \langle -I, \Lambda\rangle \end{bmatrix} = 0, \\ G(x,t) \succeq 0,\ \Lambda \succeq 0,\ \langle \Lambda, G(x,t)\rangle = 0, \end{cases} \tag{32}$$

where

$$\langle D\mathcal{A}, \cdot\rangle := \begin{bmatrix} \langle A_1, \cdot\rangle \\ \langle A_2, \cdot\rangle \\ \vdots \\ \langle A_m, \cdot\rangle \end{bmatrix}.$$

If the desired result is not true, namely,

$$t > \lambda_{\max}(A_0 + \mathcal{A}(x)),$$

which implies that

$$G(x,t) = A_0 + \mathcal{A}(x) - tI \prec 0.$$

According to the second inequalities in (32), we have $\Lambda = 0$. From the first equation of (32), it follows that

$$(y,\tau) = (x,t).$$

Since $(x,t) = \Pi_\Omega[(y,\tau)]$, so $(y,\tau) \in \Omega$, which leads to a contradiction with the assumption. This completes the proof. $\square$

Consider the Lagrangian function of (31):

$$L((x,t,\Lambda;(y,\tau))) = f((x,t);(y,\tau)) + \langle \Lambda, G(x,t)\rangle.$$

Its KKT system can be expressed as the following generalized equations:

$$\begin{cases} \nabla_{(x,t)} L((x,t,\Lambda;(y,\tau))) = \begin{bmatrix} x \\ t \end{bmatrix} - \begin{bmatrix} y \\ \tau \end{bmatrix} + \begin{bmatrix} \langle D\mathcal{A}, \Lambda \rangle \\ \langle -I, \Lambda \rangle \end{bmatrix} = 0, \\ \Lambda = \Pi_K[\Lambda - G(x,t)]. \end{cases} \tag{33}$$

We define the mapping $\mathcal{H}: \mathcal{R}^m \times \mathcal{R} \times S^n \times \mathcal{R}^m \times \mathcal{R} \to \mathcal{R}^m \times \mathcal{R} \times S^n$ by

$$\mathcal{H}((x,t),\Lambda;(y,\tau)) = \begin{bmatrix} \begin{bmatrix} x \\ t \end{bmatrix} - \begin{bmatrix} y \\ \tau \end{bmatrix} + \begin{bmatrix} \langle D\mathcal{A}, \Lambda \rangle \\ \langle -I, \Lambda \rangle \end{bmatrix} \\ \Lambda - \Pi_K[\Lambda - G(x,t)] \end{bmatrix}.$$

Then, the KKT system (33) can be written as

$$\mathcal{H}((x,t),\Lambda;(y,\tau)) = 0. \tag{34}$$

Given $(\bar{y},\bar{\tau}) \in \mathcal{R}^m \times \mathcal{R}$, let $(\bar{x},\bar{t})$ be the unique solution of (31). In this case, recall that the nondegeneracy condition (7) is stated as follows: the nondegeneracy condition holds at $(\bar{x},\bar{t})$, with respect to $G$ and $K$, if

$$JG(\bar{x},\bar{t})(\mathcal{R}^m \times \mathcal{R}) + \mathrm{lin}(T_K(G(\bar{x},\bar{t}))) = S^n. \tag{35}$$

Here $JG: \mathcal{R}^m \times \mathcal{R} \to S^n$ denotes the derivative mapping of $G$.

Under condition (35), the corresponding Lagrangian multiplier $\bar{\Lambda}$ is unique. By Lemma 5.1 and the second equation of (33), the eigenvalues of both $G(\bar{x},\bar{t})(=\bar{t}I - (A_0 + \mathcal{A}(\bar{x})))$ and $\bar{\Lambda}$ are all nonnegative. Then, there exists an orthogonal matrix $P$ such that

$$P^T G(\bar{x},\bar{t}) P = \begin{bmatrix} 0 & & \\ & 0 & \\ & & -\Sigma_\gamma \end{bmatrix},$$

and

$$P^T \bar{\Lambda} P = \begin{bmatrix} \Sigma_\alpha & & \\ & \Sigma_\beta & \\ & & 0 \end{bmatrix},$$

where $\Sigma_\alpha > 0$ are the positive eigenvalues of $\bar{\Lambda}$ with total number $\alpha$, $-\Sigma_\gamma > 0$ are the positive eigenvalues of $G(\bar{x},\bar{t})$ with total number $\gamma$, $\Sigma_\beta = 0$ with total number $\beta$. $P$ is then the corresponding orthogonal matrix consisting of orthonormal eigenvectors. So,

$$P^T[\bar{\Lambda} - G(\bar{x},\bar{t})]P = \begin{bmatrix} \Sigma_\alpha & & \\ & \Sigma_\beta & \\ & & \Sigma_\gamma \end{bmatrix}.$$

Clearly, $\alpha, \beta, \gamma$ represent the number of positive, zero, and negative eigenvalues of matrix $\bar{\Lambda} - G(\bar{x}, \bar{t})$, respectively. For convenience, we set $\Sigma_\alpha \cup \Sigma_\beta \cup \Sigma_\gamma = \{\lambda_1, \ldots, \lambda_n\}$ and define the matrix $\mathcal{U} \in S^n$ with entries

$$u_{ij} = \frac{\max\{\lambda_i, 0\} + \max\{\lambda_j, 0\}}{|\lambda_i| + |\lambda_j|}, \quad i,\ j = 1, 2, \ldots, n,$$

where 0/0 is defined as 1.

In the following analysis, for brevity, we also use $\alpha, \beta$, and $\gamma$ to denote the three index sets of the corresponding positive, zero, and negative eigenvalues of $\bar{\Lambda} - G(\bar{x}, \bar{t})$. Next, we partition the matrix $P$ into

$$P = [W_\alpha \quad Z \quad W_\gamma], \tag{36}$$

where $W_\alpha \in R^{n\times|\alpha|}$, $Z \in R^{n\times|\beta|}$, and $W_\gamma \in R^{n\times|\gamma|}$.

Let $\bar{U} = \bar{\Lambda} - G(\bar{x}, \bar{t})$. For any $V \in \partial_B \Pi_K(\bar{U})$, according to [17, Lemma 11], there exists two index sets $\alpha'$ and $\gamma'$ that partition $\beta$ and a matrix $\Gamma_{\alpha'\times\gamma'}$ with entries in [0,1] such that for any $H \in S^n$,

$$V(H) = P \begin{bmatrix} W_\alpha^T H W_\alpha & W_\alpha^T H Z & \mathcal{U}_{\alpha\gamma} \circ W_\alpha^T H W_\gamma \\ Z^T H W_\alpha & S(Z^T H Z) & 0 \\ W_\gamma^T H W_\alpha \circ (\mathcal{U}_{\alpha\gamma})^T & 0 & 0 \end{bmatrix} P^T, \tag{37}$$

where $\circ$ denotes the Hadamard product, and

$$S(Z^T H Z) = \begin{bmatrix} (Z^T H Z)_{\alpha'\alpha'} & \Gamma_{\alpha'\gamma'} \circ (Z^T H Z)_{\alpha'\gamma'} \\ (Z^T H Z)^T_{\alpha'\gamma'} \circ (\Gamma_{\alpha'\gamma'})^T & 0 \end{bmatrix}.$$

According to (37), for any $H \in S^n$ and any $V \in \partial \Pi_K(\bar{U})$, there exists $V^i \in \partial_B \Pi_K(\bar{U})$, $\nu_i \geqslant 0, i = 1, 2, \ldots, \kappa,$ with $\sum_{i=1}^{\kappa} \nu_i = 0$, such that

$$\begin{aligned} V(H) &= \sum_{i=1}^{\kappa} \nu_i V^i(H) \\ &= P \begin{bmatrix} W_\alpha^T H W_\alpha & W_\alpha^T H Z & \mathcal{U}_{\alpha\gamma} \circ W_\alpha^T H W_\gamma \\ Z^T H W_\alpha & \sum_{i=1}^{\kappa} \nu_i S^i(Z^T H Z) & 0 \\ W_\gamma^T H W_\alpha \circ (\mathcal{U}_{\alpha\gamma})^T & 0 & 0 \end{bmatrix} P^T. \end{aligned} \tag{38}$$

Next, to check Assumption 3.1, we consider $V \in \partial \Pi_K(\bar{U})$ and $H \in S^n$ satisfying

$$(I - V)(H) = 0. \tag{39}$$

It follows from (38) that

$$P^T HP = P^T V(H)P$$
$$= \begin{bmatrix} W_\alpha^T HW_\alpha & W_\alpha^T HZ & \mathcal{U}_{\alpha\gamma} \circ W_\alpha^T HW_\gamma \\ Z^T HW_\alpha & \sum_{i=1}^{\kappa} \nu_i S^i(Z^T HZ) & 0 \\ W_\gamma^T HW_\alpha \circ (\mathcal{U}_{\alpha\gamma})^T & 0 & 0 \end{bmatrix}. \tag{40}$$

Recasting the expression of $P^T HP$ by using (36) and comparing both sides of (40), we have

$$\begin{cases} Z^T HW_\gamma = 0, \\ W_\gamma^T HW_\gamma = 0, \\ \mathcal{U}_{\alpha\gamma} \circ W_\alpha^T HW_\gamma = W_\alpha^T HW_\gamma, \\ Z^T HZ = \sum_{i=1}^{\kappa} \nu_i S^i(Z^T HZ). \end{cases} \tag{41}$$

As

$$T_K(G(\bar{x}, \bar{t})) = \{A \in S^n : [W_\alpha \ Z]^T A[W_\alpha \ Z] \succeq 0\}, \tag{42}$$

this leads to

$$\begin{aligned} \operatorname{lin}(T_K(G(\bar{x}, \bar{t}))) &= \{A \in S^n : [W_\alpha \ Z]^T A[W_\alpha \ Z] = 0\} \\ &= \left\{ A \in S^n : \begin{bmatrix} W_\alpha^T AW_\alpha & W_\alpha^T AZ \\ Z^T AW_\alpha & Z^T AZ \end{bmatrix} = 0 \right\}. \end{aligned} \tag{43}$$

For $H$ to satisfy (39), according to Assumption 3.1, we only need to show that

$$H \in (\operatorname{lin}(T_K(G(\bar{x}, \bar{\tau}))))^\perp.$$

For any $A \in \operatorname{lin}(T_K(G(\bar{x}, \bar{t})))$, by (41) and (43), it is easy to derive that

$$\langle H, A \rangle = \langle P^T HP, P^T AP \rangle$$
$$= 2\langle \mathcal{U}_{\alpha\gamma} \circ W_\alpha^T HW_\gamma, W_\alpha^T AW_\gamma \rangle.$$

Since $0 < \mathcal{U}_{i,j} < 1$ for any $(i, j) \in \alpha \times \gamma$, so by the third equation in (41),

$$W_\alpha^T H W_\gamma = 0,$$

which implies that

$$\langle H, A \rangle = 0.$$

Thus

$$H \in (\mathrm{lin}(T_K(G(\bar{x}, \bar{t}))))^\perp,$$

hence, Assumption 3.1 holds.

Based on the above arguments, we derive the following result:

PROPOSITION 5.1. *Let* $(\bar{y}, \bar{\tau}) \in \mathcal{R}^m \times \mathcal{R}$ *and* $(\bar{x}, \bar{t})$ *be the corresponding solution of* (31). *Assume the nondegeneracy condition holds and* $\bar{\Lambda}$ *is the associated unique multiplier of KKT system* (33). *For* $H \in S^n$ *and* $V \in \partial \Pi_{S_+^n}(\bar{\Lambda} - G(\bar{x}, \bar{t}))$, *if* $(I - V)(H) = 0$, *then* $H \in (\mathrm{lin}(T_{S_+^n}(G(\bar{x}, \bar{\tau}))))^\perp$, *where* $G(\bar{x}, \bar{t}) = \bar{t} I - [A_0 + \mathcal{A}(\bar{x})]$.

By applying Theorem 3.1 and Proposition 5.1, we derive the following result.

THEOREM 5.1. *Let* $(\bar{x}, \bar{t})$ *be the solution of* (30) *at the given point* $(\bar{y}, \bar{\tau}) \in \mathcal{R}^m \times \mathcal{R}$. *Then, under the nondegeneracy condition* (35) *at* $(\bar{x}, \bar{t})$, *the projection mapping on the epigraph of the maximum eigenvalue function is strongly semismooth near* $(\bar{y}, \bar{\tau})$.

*Proof.* By [17], the projection mapping $\Pi_K(\cdot)(= \Pi_{S_+^n}(\cdot))$ is strongly semismooth everywhere. Hence, the result follows immediately from Theorem 3.1 and Proposition 5.1. □

### 5.1. SEMISMOOTHNESS OF SOLUTION MAPPINGS

Let $\phi: X \to \mathcal{R}$ be a lower semicontinuous convex function, where $X$ is a finite dimensional vector space. Let $\bar{\phi}_\epsilon$ be its Moreau–Yosida regularization, namely,

$$\bar{\phi}_\epsilon(y) = \min \left\{ \phi(x) + \frac{\epsilon}{2} \langle y - x, y - x \rangle \mid x \in X \right\}, \quad \epsilon > 0, \tag{44}$$

which is continuously differentiable with

$$\nabla \bar{\phi}_\epsilon(y) = \epsilon(y - x(y)), \qquad y \in X,$$

where $x(\cdot)$ denotes the unique solution mapping of (44) on $X$. Let epi$(\phi)$ denote the epigraph of $\phi$, i.e., epi$(\phi) := \{(x, u) \in X \times \mathcal{R} \mid u \geqslant \phi(x)\}$ and define $u(\cdot) := \phi(x(\cdot))$. Then, we have the following result with respect to the solution mapping of the Moreau–Yosida regularization.

PROPOSITION 5.2. *Given* $\bar{y} \in X$, *let* $\bar{x} = x(\bar{y})$ *and* $\bar{u} = \phi(x(\bar{y}))$. *Then,* $(x(\cdot), u(\cdot))$ *and* $\nabla\bar{\phi}_\epsilon(\cdot)$ *are G-semismooth (strongly G-semismooth, semismooth, strongly semismooth) near* $\bar{y}$ *if the projection mapping* $\Pi_{\mathrm{epi}(\phi)}(\cdot)$ *is G-semismooth (strongly G-semismooth, semismooth, strongly semismooth) near* $(\bar{y}, \bar{u} - 1/\epsilon)$.

*Proof.* The proof is similar to [12, Theorem 4]. First, we get Clarke's nonsingularity condition by [12, Proposition 4]. The desired result follows immediately by [25, Theorem 2.1]. □

By virtue of Theorem 5.1 and Proposition 5.2, we derive the semismoothness of the gradient of the Moreau–Yosida regularization of the maximum eigenvalue function.

THEOREM 5.2. *For* $(\bar{y}, \bar{\tau}) \in \mathcal{R}^m \times \mathcal{R}$, *let* $(\bar{x}, \bar{t})$ *be the solution of* (31). *Suppose the nondegeneracy condition holds at* $(\bar{x}, \bar{t})$. *Then, the solution mapping* $x(\cdot)$ *to the Moreau–Yosida regularization* (26) *of the maximum eigenvalue function is strongly semismooth near* $\bar{y}$. *Further,* $\nabla \hat{f}_\epsilon(\cdot)$ *is strongly semismooth near* $\bar{y}$, *where* $\nabla \hat{f}_\epsilon(\cdot)$ *is defined in* (29).

*Proof.* $\bar{f}$ is finite valued everywhere, i.e., $\mathrm{dom}(\bar{f}) = \mathcal{R}^m$. By Theorem 5.1, $\Pi_{\mathrm{epi}(\bar{f})}(\cdot)$ is strongly semismooth near $(\bar{y}, \bar{\tau})$ with $\bar{\tau} = \bar{f}(\bar{x}) - 1/\epsilon$. Hence, the result follows from Proposition 5.2. □

**Acknowledgments**

We would like to thank Professor Alexander Shapiro for his constructive comments, in particular Assumption 3.1, on an earlier version of this paper [22].

## References

1. Bonnans, J.F. and Shapiro, A. (2000), *Perturbation Analysis of Optimization Problems*, Springer-Verlag, New York.
2. Clarke, F.H. (1983), *Optimization and Nonsmooth Analysis*, John Wiley and Sons, New York.
3. Eaves, B.C. (1971), On the basic theorem of complementarity, *Mathematical Programming* 1, 68–75.
4. Facchinei, F. and Pang J.S., (2003), *Finite-Dimensional Variational Inequalities and Complementarity Problems*, Vols. I and II. Springer-Verlag, New York.
5. Fukushima, M., Luo, Z.Q. and Tseng, P. (2002), Smoothing functions for second-order cone complementarity problems, *SIAM Journal on Optimization* 12, 436–460.

6. Fukushima, M. and Qi, L. (1996), A global and superlinear convergent algorithm for nonsmooth convex minimization, *SIAM Journal on Optimization* 6, 1106–1120.
7. Gowda, M.S. (2004), Inverse and implicit function theorems for H-differentiable and semismooth functions, *Optimization Methods and Software* 19, 443–461.
8. Hiriart-Urruty, J.B. and Lemaréchal, C. (1993), *Convex Analysis and Minimization Algorithms*, Springer, Berlin.
9. Klatte, D. and Tammer, K. (1990), Strong stability of stationary solutions and Karush–Kuhn–Tucker points in nonlinear optimization, *Annals of Operations Research* 27, 285–308.
10. Kojima, M. (1980), Strongly stable stationary solutions in nonlinear programs. In: Robinson, S.M. (ed.), *Analysis and Computation of Fixed Points,* Academic Press, New York, pp. 93–138.
11. Luo, Z.Q., Pang, J.-S. and Ralph, D. (1996), *Mathematical Programs with Equilibrium Constraints*, Cambridge University Press, Cambridge.
12. Meng, F., Sun, D. and Zhao, G. (2005), Semismoothness of Solutions to Generalized Equations and the Moreau–Yosida Regularization, *Mathematical Programming, Series B*, 104, 561–581.
13. Mifflin, R. (1977), Semismooth and semiconvex functions in constrained optimization, *SIAM Journal on Control and Optimization* 15, 957–972.
14. Moreau, J.J. (1965), Proximite et dualite dans un espace hilbertien, *Bulletin de la Societe Mathematique de France* 93, 273–299.
15. Overton, M.L. (1988), On minimizing the maximum eigenvalue of a symmetric matrix, *SIAM Journal of Matrix Analysis and Applications* 9, pp. 256–268.
16. Pang, J.-S. and Ralph, D. (1996), Piecewise smoothness, local invertibility, and parametric analysis of normal maps, *Mathematics of Operations Research* 21, 401–426.
17. Pang J.-S., Sun, D. and Sun, J. (2003), Semismooth Homeomorphisms and Strong Stability of Semidefinite and Lorentz Cone Complementarity Problems, *Mathematics of Operations Research* 28, 39–63.
18. Qi, L. and Sun, J. (1993), A nonsmooth version of Newton's method, *Mathematical Programming* 58, 353–367.
19. Robinson, S.M. (1980), Strongly regular generalized equations, *Mathematics of Operations Research* 5, 43–62.
20. Rockafellar, R.T. and Wets, R.J.-B. (1998), *Variational Analysis*, Springer-Verlag, New York.
21. Shapiro, A. (2003), Sensitivity analysis of generalized equations, *Journal of Mathematical Sciences* 115, 2554–2565.
22. Shapiro, A. (2005), Private Communication. May.
23. Shapiro, A. and Fan, M.K.H. (1995), On eigenvalue optimization, *SIAM Journal on Optimization* 5, 552–568.
24. Sun, D. (2006), *The Strong Second Order Sufficient Condition and Constraint Nondegeneracy in Nonlinear Semidefinite Programming and their Implications*, to appear in *Mathematics of Operations Research.*
25. Sun, D. (2001), A further result on an implicit function theorem for locally Lipschitz functions, *Operations Research Letters* 28, 193–198.
26. Sun, D. and Sun, J. (2002), Semismooth matrix valued functions, *Mathematics of Operations Research* 27, 150–169.
27. Yosida, K. (1964), *Functional Analysis*, Springer Verlag, Berlin.