# On differentiability of symmetric matrix valued functions

**Alexander Shapiro***

School of Industrial and Systems Engineering,
Georgia Institute of Technology,
Atlanta, Georgia 30332-0205, USA
e-mail: ashapiro@isye.gatech.edu

**Abstract**

With every real valued function, of a real argument, can be associated a matrix function mapping a linear space of symmetric matrices into itself. In this paper we study directional differentiability properties of such matrix functions associated with directionally differentiable real valued functions. In particular, we show that matrix valued functions inherit semismooth properties of the corresponding real valued functions.

**Key words**: matrix function, eigenvalues and eigenvectors, directional derivatives, semismooth mappings

*Supported by the National Science Foundation under grant DMS-0073770.

# 1 Introduction

Consider the linear space $\mathcal{S}^p$ of $p \times p$ symmetric matrices. With a function $f : \mathbb{R} \to \mathbb{R}$ can be associate a matrix valued function $F : \mathcal{S}^p \to \mathcal{S}^p$ defined as follows. For a matrix $X \in \mathcal{S}^p$ denote by $\lambda_1(X) \geq ... \geq \lambda_p(X)$ its eigenvalues, arranged in the decreasing order, and by $e_1(X), ..., e_p(X)$ a set of the corresponding orthonormal eigenvectors. Then we define

$$F(X) := \sum_{i=1}^{p} f(\lambda_i(X))\, e_i(X) e_i(X)^T. \tag{1.1}$$

Of course, the orthonormal eigenvectors $e_1(X), ..., e_p(X)$ are not defined uniquely. However, the right hand side of (1.1) does not depend on a particular choice of the orthonormal eigenvector system of $X$, and hence the function $F(X)$ is well defined. Matrix functions of the form (1.1) have been studied extensively (see, e.g., [1, Chapter 6] and references therein). In particular, for $f(x) := x^n$, $n = 1, ...,$ we have $F(X) = X^n$ in the sense of the usual matrix multiplication. For analytic functions $f(x)$ one can then define $F(X)$ by using power series expansions.

In this paper we investigate differentiability properties of the function $F(X)$ in cases where $f(x)$ is directionally differentiable, but not necessarily differentiable everywhere. In particular, we show that the matrix function $F(X)$ inherits semismoothness properties of the function $f(x)$. For the function $f(x) := |x|$, directional differentiability and semismoothness of the corresponding matrix function, denoted $|X|$, was shown by Sun and Sun [13]. For $f(x) := x_+$, where $x_+ := \max\{0, x\}$, the corresponding mapping $X \mapsto F(X)$ represents the metric projection of $X$ onto the cone $\mathcal{S}^p_+$ of $p \times p$ positive semidefinite symmetric matrices. Semismoothness of that mapping (metric projection) was shown in Pang, Sun and Sun [8]. So this investigation is motivated by those papers.

Recall that a mapping $g : \mathbb{R}^n \to \mathbb{R}^m$ is said to be directionally differentiable, at a point $x \in \mathbb{R}^n$, if its directional derivative

$$g'(x, h) := \lim_{t \downarrow 0} \frac{g(x + th) - g(x)}{t}$$

exists in every direction $h \in \mathbb{R}^n$. If, moreover, the directional derivative $g'(x, \cdot)$ is linear, then $g(\cdot)$ is (Gâteaux) differentiable at $x$ and its differential $Dg(x)h = g'(x, h)$.

We assume throughout this paper that the function $f(\cdot)$ is directionally differentiable. Since here $x$ is real valued, directional differentiability of $f(\cdot)$ simply means that $f(\cdot)$ is right and left side differentiable. In particular, if $f(\cdot)$ is differentiable at a point $x$, then $f'(x,h) = f'(x)h$, where $f'(x) = df(x)/dx$ denotes the derivative of $f(\cdot)$ at $x$.

There are several equivalent ways of defining semismoothness, originally introduced in Mifflin [6], of a mapping $g : \mathbb{R}^n \to \mathbb{R}^m$. We use the following definition (cf., [7])

**Definition 1.1** Let $g : \mathbb{R}^n \to \mathbb{R}^m$ be a locally Lipschitz continuous and directionally differentiable mapping. It is said that $g(\cdot)$ is semismooth at a point $x \in \mathbb{R}^n$ if

$$g(x+h) - g(x) = g'(x+h,h) + o(\|h\|). \tag{1.2}$$

It is said that $g(\cdot)$ is strongly semismooth at $x$ if

$$g(x+h) - g(x) = g'(x+h,h) + O(\|h\|^2). \tag{1.3}$$

It is said that the mapping $g(x)$ is positively homogeneous if $g(tx) = tg(x)$ for any $x \in \mathbb{R}^n$ and $t \geq 0$. If $g(x)$ is positively homogeneous, then $g(0) = 0$, $g(x)$ it is directionally differentiable at $x = 0$ and $g'(h,h) = g(h)$, and hence $g(\cdot)$ is strongly semismooth at $x = 0$. Moreover, if $g(x)$ is locally Lipschitz continuous and $F : \mathbb{R}^k \to \mathbb{R}^n$ is a twice continuously differentiable mapping and $F(\bar{x}) = 0$, then the composite mapping $g \circ F : \mathbb{R}^k \to \mathbb{R}^m$ is strongly semismooth at $\bar{x}$ (see [3],[7] for a discussion of these properties).

We use the following notation throughout the paper. For a matrix $X \in \mathcal{S}^p$ we denote by $\sigma(X) := (\lambda_1(X), ..., \lambda_p(X))$ its spectrum. We study differentiability properties of $\sigma(X)$ and $F(X)$ in a neighborhood of a given matrix $\bar{X} \in \mathcal{S}^p$. By $r_1, ..., r_q$ we denote the multiplicities and by $\mu_1 > ... > \mu_q$ the distinct values of the eigenvalues $\lambda_1(\bar{X}), ..., \lambda_p(\bar{X})$, i.e., $\mu_j := \lambda_{s_j+1}(\bar{X}) = ... = \lambda_{s_j+r_j}(\bar{X})$, $j = 1, ..., q$, where

$$s_1 := 0,\ s_2 := r_1,\ \dots,\ s_q := r_1 + ... + r_{q-1}. \tag{1.4}$$

Also by $E_j(X)$ we denote the $p \times r_j$ matrix whose columns are formed by the eigenvectors $e_{s_j+1}(X), ..., e_{s_j+r_j}(X)$, $j = 1, ..., q$, and define $P_j(X) := E_j(X)E_j(X)^T$. Note that the matrix $P_j(X)$ represents the orthogonal projection onto the space generated by the eigenvectors $e_{s_j+1}(X), ..., e_{s_j+r_j}(X)$ and is independent of a particular choice of $E_j(X)$. In particular, denote

$$\bar{E}_j := E_j(\bar{X}) \ \text{ and } \ \bar{P}_j := P_j(\bar{X}) = \bar{E}_j\bar{E}_j^T.$$

That is, $\bar{P}_j$ is the orthogonal projection matrix onto the eigenvector space corresponding to the eigenvalue $\mu_j$. We have then that $\bar{P}_i\bar{P}_j = 0$ if $i \neq j$, $\bar{P}_i^2 = \bar{P}_i$ and $\sum_{j=1}^q \bar{P}_j = I_p$. Moreover,

$$\bar{X} = \sum_{j=1}^q \mu_j \bar{P}_j \quad \text{and} \quad F(\bar{X}) = \sum_{j=1}^q f(\mu_j)\bar{P}_j. \tag{1.5}$$

For a number $x \in \mathbb{R}$ we denote by $\operatorname{sign}(x)$ its sign, i.e., $\operatorname{sign}(x) = 1$ if $x > 0$, $\operatorname{sign}(x) = -1$ if $x < 0$ and $\operatorname{sign}(x) = 0$ if $x = 0$. By $I_p$ we denote the $p \times p$ identity matrix.

# 2 Analytic Functions

In this section we study the case where the function $f(x)$ is analytic in a neighborhood of every point $\mu_j$, $j = 1, ..., q$. For $z \neq \mu_j$, $j = 1, ..., q$, we have that the matrix $zI_p - \bar{X}$ is nonsingular and

$$(zI_p - \bar{X})^{-1} = \sum_{j=1}^q (z - \mu_j)^{-1}\bar{P}_j. \tag{2.1}$$

Therefore, if we take a curve $C_j$ in the complex plane to be a sufficiently small circle centered at $\mu_j$ and oriented in the anticlock direction, then by Cauchy's theorem,

$$\frac{1}{2\pi i}\int_{C_j} f(z)(zI_p - \bar{X})^{-1}dz = f(\mu_j)\bar{P}_j. \tag{2.2}$$

Consider the curve $C := \cup_{j=1}^q C_j$. It is well known that the eigenvalues $\lambda_1(X), ..., \lambda_p(X)$ are continuous, in fact even locally Lipschitz continuous, functions of $X \in \mathcal{S}^p$. Consequently, for all $X$ in a neighborhood of $\bar{X}$ the eigenvalues $\lambda_1(X), ..., \lambda_p(X)$ are contained inside the curve $C$. It follows that for all $X$ in a neighborhood of $\bar{X}$ the matrix function $F(X)$ can be represented in the following integral form

$$F(X) = \frac{1}{2\pi i}\int_C f(z)(zI_p - X)^{-1}dz. \tag{2.3}$$

The idea of using the above representation for studying matrix functions is going back to Poincaré [9] and is used extensively in Kato [4].

The integral representation (2.3) implies that the matrix function $F(X)$ is analytic, and hence is twice continuously differentiable, in a neighborhood of $\bar{X}$. It follows that $F(X)$ is strongly semismooth at $\bar{X}$. For example, the functions $f(x) := |x|$ and $f(x) := x_+$ are analytic at every point $x \neq 0$. Therefore, the corresponding matrix functions $F(X)$ are analytic, and hence strongly semismooth, at every nonsingular matrix $\bar{X}$.

It is also possible to use the integral representation (2.3) for calculation of derivatives of $F(X)$. We have that if the matrix $A := (zI_p - \bar{X})$ is nonsingular, then

$$(A - H)^{-1} = A^{-1} + A^{-1}HA^{-1} + o(\|H\|), \tag{2.4}$$

and hence

$$DF(\bar{X})H = \frac{1}{2\pi i}\int_C f(z)(zI_p - \bar{X})^{-1}H(zI_p - \bar{X})^{-1}dz. \tag{2.5}$$

Together with (2.1) this implies

$$DF(\bar{X})H = \frac{1}{2\pi i}\sum_{j,k,l=1}^{q}\int_{C_l} f(z)(z-\mu_j)^{-1}(z-\mu_k)^{-1}\bar{P}_j H \bar{P}_k dz. \tag{2.6}$$

We also have that

$$\frac{1}{2\pi i}\int_{C_l} f(z)(z-\mu_j)^{-1}(z-\mu_k)^{-1}dz = \begin{cases} 0, & \text{if } \ l \neq j \text{ and } l \neq k, \\ \frac{f(\mu_j)}{\mu_j - \mu_k}, & \text{if } \ l = j \text{ and } l \neq k, \\ f'(\mu_j), & \text{if } \ l = j \text{ and } l = k. \end{cases}$$

It follows that (cf., [4])

$$DF(\bar{X})H = \mathbb{A}_f(\bar{X}, H) + \sum_{j=1}^{q} f'(\mu_j)\bar{P}_j H \bar{P}_j, \tag{2.7}$$

where

$$\mathbb{A}_f(\bar{X}, H) := \tfrac{1}{2}\sum_{\substack{j\neq k \\ j,k=1}}^{q} \frac{f(\mu_j) - f(\mu_k)}{\mu_j - \mu_k}\left(\bar{P}_j H \bar{P}_k + \bar{P}_k H \bar{P}_j\right). \tag{2.8}$$

It also follows that, for any $j \in \{1, ..., q\}$, the mapping $X \mapsto P_j(X)$ is analytic in a neighborhood of $\bar{X}$ and

$$DP_j(\bar{X})H = \sum_{\substack{k \neq j \\ k=1}}^{q} (\mu_j - \mu_k)^{-1} \left( \bar{P}_j H \bar{P}_k + \bar{P}_k H \bar{P}_j \right). \tag{2.9}$$

The assumption that $f(x)$ is analytic at every $\mu_j$ was essential in the above analysis. In section 4 we discuss cases where $f(x)$ may be even non-differentiable at some points.

# 3 Differentiability of Eigenvalue Functions

By using the fact that the mappings $P_j(X)$ are analytic in a neighborhood of $\bar{X}$, it is possible to construct, for every $j \in \{1, ..., q\}$, a mapping $\Xi_j : \mathcal{S}^p \to \mathcal{S}^{r_j}$ with the following properties (a detail construction of this mapping is given in [11, p.558] and [2, example 3.98, p.211]).

**(i)** The mapping $\Xi_j(\cdot)$ is analytic in a neighborhood of $\bar{X}$,

**(ii)** $\Xi_j(\bar{X}) = \mu_j I_{r_j}$,

**(iii)** $D\Xi_j(\bar{X}) : \mathcal{S}^p \to \mathcal{S}^{r_j}$ is onto and

$$D\Xi_j(\bar{X})H = \bar{E}_j^T H \bar{E}_j, \tag{3.1}$$

**(iv)** For all $X \in \mathcal{S}^p$ in a neighborhood of $\bar{X}$ the following holds:

$$\lambda_{s_j+i}(X) = \nu_i(\Xi_j(X)), \quad i = 1, ..., r_j, \tag{3.2}$$

where $\nu_1(Y) \geq ... \geq \nu_{r_j}(Y)$ denote eigenvalues of matrix $Y \in \mathcal{S}^{r_j}$.

Consider a matrix $H \in \mathcal{S}^p$. We have that

$$\Xi_j(\bar{X} + tH) = \Xi_j(\bar{X}) + tD\Xi_j(\bar{X})H + o(t) = \mu_j I_{r_j} + tD\Xi_j(\bar{X})H + o(t).$$

Since the eigenvalues of a symmetric matrix are locally Lipschitz continuous, it follows that for $i = 1, ..., r_j$ and $t \geq 0$ small enough,

$$\lambda_{s_j+i}(\bar{X} + tH) = \nu_i \left( \mu_j I_{r_j} + tD\Xi_j(\bar{X})H \right) + o(t) = \mu_j + t\nu_i(D\Xi_j(\bar{X})H) + o(t).$$

Together with (3.1) this implies the following.

**Proposition 3.1** *The directional derivatives* $\lambda'_{s_j+i}(\bar{X}, H)$, $i = 1, ..., r_j$, *exist and coincide with the corresponding eigenvalues of the matrix* $\bar{E}_j^T H \bar{E}_j$ *arranged in the decreasing order.*

The above result, of course, is known, e.g., it is a particular case of a more general result given in [5, Theorem 7].

For $i \in \{1, ..., r_j\}$ the function $\lambda_{s_j+i}(X)$ can be represented in a neighborhood of $\bar{X}$ in the form

$$\lambda_{s_j+i}(X) = \mu_j + \nu_i \left( \Gamma_j(X) \right), \tag{3.3}$$

where $\Gamma_j(X) := \Xi_j(X) - \mu_j I_{r_j}$. We have that the mapping $\Gamma_j(X)$ is analytic at $\bar{X}$, $\Gamma_j(\bar{X}) = 0$, and the function $\nu_i : \mathcal{S}^{r_j} \to \mathbb{R}$ is positively homogeneous. As it was mentioned in the Introduction, composition of the positively homogeneous and locally Lipschitz continuous function $\nu_i(\cdot)$ and the twice continuously diferentiable mapping $\Gamma_j(\cdot)$ is strongly semismooth at a point $\bar{X}$ such that $\Gamma_j(\bar{X}) = 0$. We obtain therefore the following.

**Proposition 3.2** *The eigenvalue functions* $\lambda_i : \mathcal{S}^p \to \mathbb{R}$, $i = 1, ..., p$, *are strongly semismooth at every* $X \in \mathcal{S}^p$.

The above proposition implies that the spectral mapping $X \mapsto \sigma(X)$ is strongly semismooth at every $X \in \mathcal{S}^p$. By a different technique this was shown in Sun and Sun [12].

In the subsequent analysis we will need the following construction. For a given (direction) matrix $H \in \mathcal{S}^p$ and $j = 1$ consider the matrix $\bar{E}_1^T H \bar{E}_1$. Note that the matrix $\bar{E}_1$ is defined up to an orthogonal rotation, i.e., it can be replaced by $\bar{E}_1 Q$, where $Q$ is an $r_1 \times r_1$ orthogonal matrix. Of course, the eigenvalues of the matrix $Q^T \bar{E}_1^T H \bar{E}_1 Q$ coincide with the corresponding eigenvalues of $\bar{E}_1^T H \bar{E}_1$. Now let $\tilde{e}_1 = e_1(\bar{X}), ..., \tilde{e}_{r_1} = e_{r_1}(\bar{X})$ be chosen in such a way that:

**(a)** $\tilde{e}_i^T H \tilde{e}_j = 0$ for $i \neq j \in \{1, ..., r_1\}$,

**(b)** $\tilde{e}_1^T H \tilde{e}_1$,...,$\tilde{e}_{r_1}^T H \tilde{e}_{r_1}$ form the eigenvalues of the $r_1 \times r_1$ matrix $\bar{E}_1^T H \bar{E}_1$ arranged in the decreasing order.

That is, $\bar{E}_1^T H \bar{E}_1$ is rotated to a diagonal form. Denote by $\tilde{E}_1$ the $p \times r_1$ matrix whose columns are formed by vectors $\tilde{e}_1, ..., \tilde{e}_{r_1}$. The matrix $\tilde{E}_1$ may still be not defined uniquely, so we use this as a generic notation for such matrices.

We have then that any accumulation point of $E_1(\bar{X}+tH)$, as $t \downarrow 0$, is a matrix $\tilde{E}_1$ satisfying the above construction. This follows from Proposition 3.1 by passing to a limit (cf., [10]). Of course, the same procedure can be performed for every $j \in \{1, ..., q\}$.

# 4 Properties of Matrix Functions

In this section we study continuity and differentiability properties of the matrix functions $F(X)$, of the form (1.1), for possibly nondifferentiable functions $f(x)$.

**Proposition 4.1** *Suppose that the function $f(x)$ is locally Lipschitz continuous. Then the corresponding matrix function $F(X)$ is locally Lipschitz continuous.*

**Proof.** Since $P_j(X) = \sum_{k=s_j+1}^{s_j+r_j} e_k(X)e_k(X)^T$, we can write

$$F(X) = \sum_{j=1}^{q} f(\mu_j)P_j(X) + \sum_{j=1}^{q}\sum_{k=s_j+1}^{s_j+r_j} [f(\lambda_k(X)) - f(\mu_j)]\, e_k(X)e_k(X)^T. \tag{4.1}$$

It follows that

$$\begin{aligned} \|F(X) - F(\bar{X})\| &\le \sum_{j=1}^{q} |f(\mu_j)|\, \|P_j(X) - P_j(\bar{X})\| \\ &+ \sum_{j=1}^{q}\sum_{k=s_j+1}^{s_j+r_j} |f(\lambda_k(X)) - f(\mu_j)|\, \|e_k(X)e_k(X)^T\|. \end{aligned}$$

Since $\|e_k(X)e_k(X)^T\|$ are uniformly bounded, the result then follows from locally Lipschitz continuity of the eigenvalue functions $\lambda_k(\cdot)$ and of $P_j(\cdot)$. ∎

We assume in the remainder of this section that the function $f(x)$ is directionally differentiable. Consider functions $\psi_j(\cdot) := f'(\mu_j, \cdot)$, $j = 1, ..., q$, and let $\Psi_j : \mathcal{S}^{r_j} \to \mathcal{S}^{r_j}$ be the corresponding matrix functions.

**Theorem 4.1** *Suppose that the function $f(x)$ is directionally differentiable at every point $\mu_j$, $j = 1, ...q$. Then the corresponding matrix function $F(x)$ is directionally differentiable at $\bar{X}$ and*

$$F'(\bar{X}, H) = \mathbb{A}_f(\bar{X}, H) + \mathbb{B}_f(\bar{X}, H), \tag{4.2}$$

*where $\mathbb{A}_f(\bar{X}, H)$ is defined in (2.8) and*

$$\mathbb{B}_f(\bar{X}, H) := \sum_{j=1}^{q} \bar{E}_j \left[\Psi_j \left(\bar{E}_j^T H \bar{E}_j\right)\right] \bar{E}_j^T. \tag{4.3}$$

**Proof.** Consider the decomposition (4.1) with $X := \bar{X} + tH$. We have

$$\lim_{t \downarrow 0} t^{-1} \sum_{j=1}^{q} f(\mu_j) \left[P_j(\bar{X} + tH) - P_j(\bar{X})\right] = \sum_{j=1}^{q} f(\mu_j) D P_j(\bar{X}) H, \tag{4.4}$$

and by (2.9)

$$\begin{aligned} \sum_{j=1}^{q} f(\mu_j) D P_j(\bar{X}) H &= \sum_{j=1}^{q} \sum_{k \neq j} \frac{f(\mu_j)}{\mu_j - \mu_k} \left(\bar{P}_j H \bar{P}_k + \bar{P}_k H \bar{P}_j\right) \\ &= \sum_{1 \leq j < k \leq q} \frac{f(\mu_j) - f(\mu_k)}{\mu_j - \mu_k} \left(\bar{P}_j H \bar{P}_k + \bar{P}_k H \bar{P}_j\right) = \mathbb{A}_f(\bar{X}, H). \end{aligned} \tag{4.5}$$

Now for $t > 0$ and $j = 1$ consider

$$\Delta_1(t) := t^{-1} \sum_{k=s_1+1}^{s_1+r_1} \left[f\left(\lambda_k(\bar{X} + tH)\right) - f(\mu_1)\right] e_k(\bar{X} + tH) e_k(\bar{X} + tH)^T.$$

As it was discussed at the end of section 3, any accumulation point of $E_1(\bar{X} + tH)$ is a matrix $\tilde{E}_1$ whose columns $\tilde{e}_1, ..., \tilde{e}_{r_1}$ satisfy the corresponding conditions (a) and (b). We also have, for $k = 1, ..., r_1$, that

$$\lim_{t \downarrow 0} t^{-1} \left[f\left(\lambda_k(\bar{X} + tH)\right) - f(\mu_1)\right] = f'\left(\mu_1, \lambda_k'(\bar{X}, H)\right),$$

and by Proposition 3.1

$$f'\left(\mu_1, \lambda_k'(\bar{X}, H)\right) = f'\left(\mu_1, \tilde{e}_k^T H \tilde{e}_k\right).$$

It follows that

$$\lim_{t\downarrow 0} \Delta_1(t) = \sum_{k=s_1+1}^{s_1+r_1} f'\left(\mu_1, \tilde{e}_k^T H \tilde{e}_k\right) \tilde{e}_k \tilde{e}_k^T = \bar{E}_1 \left[\Psi_1\left(\bar{E}_1^T H \bar{E}_1\right)\right] \bar{E}_1^T. \tag{4.6}$$

Note that the right hand side of (4.6) does not depend on a particular choice of the columns $e_1(\bar{X}), ..., e_{r_1}(\bar{X})$ of the matrix $\bar{E}_1$.

The same calculations can be performed for every $j \in \{1, ..., q\}$. The required result then follows by (4.5) and (4.6). ∎

In particular, suppose that $f(\cdot)$ is differentiable at every point $\mu_j$, $j = 1, ..., q$. Then $f'(\mu_j, h) = f'(\mu_j)h$ and hence

$$\bar{E}_j \left[\Psi_j\left(\bar{E}_j^T H \bar{E}_j\right)\right] \bar{E}_j^T = f'(\mu_j) \bar{P}_j H \bar{P}_j. \tag{4.7}$$

It follows then that $F'(\bar{X}, H)$ is linear in $H$, and is given by the right hand side of (2.7), and hence $F(X)$ is Gâteaux differentiable at $\bar{X}$. Suppose further that $f(x)$ is locally Lipschitz continuous. Then, by Proposition 4.1, $F(X)$ is also locally Lipschitz continuous, and hence Fréchet differentiability of $F(X)$ at $\bar{X}$ follows.

**Proposition 4.2** *Suppose that the function $f(x)$ is differentiable at every point $\mu_j$, $j = 1, ..., q$. Then $F(X)$ is (Gâteaux) differentiable at $\bar{X}$ and its differential is given by formula (2.7). If, moreover, $f(x)$ is continuously differentiable at every point $\mu_j$, $j = 1, ..., q$, then $F(X)$ is continuously differentiable at $\bar{X}$*

**Proof.** Gâteaux diferentiability follows by the above discussion. So let us suppose that $f(x)$ is continuously differentiable at every point $\mu_j$, $j = 1, ..., q$. Let $X \in \mathcal{S}^p$ be sufficiently close to $\bar{X}$ and consider the terms corresponding to $j = 1$ in (2.7). Suppose for the moment that $\lambda_1(X) > ... > \lambda_{r_1}(X)$. We have then by the Mean Value Theorem that

$$\tfrac{1}{2} \sum_{\substack{i \neq k \\ i,k=1}}^{r_1} \frac{f(\lambda_i(X)) - f(\lambda_k(X))}{\lambda_i(X) - \lambda_k(X)} \Pi_{ik}(X, H) = \tfrac{1}{2} \sum_{\substack{i \neq k \\ i,k=1}}^{r_1} f'(\eta_{ik}) \Pi_{ik}(X, H), \tag{4.8}$$

where $\eta_{ik}$ is a point between the numbers $\lambda_i(X)$ and $\lambda_k(X)$ and

$$\Pi_{ik}(X, H) := e_i(X) e_i(X)^T H e_k(X) e_k(X)^T + e_k(X) e_k(X)^T H e_i(X) e_i(X)^T.$$

We also have that as $X$ tends to $\bar{X}$, the eigenvalues $\lambda_1(X), ..., \lambda_{r_1}(X)$ tend to $\mu_1$ and hence the derivatives $f'(\eta_{ik})$ tend to $f'(\mu_1)$. Moreover,

$$\frac{1}{2}\sum_{\substack{i \neq k \\ i,k=1}}^{r_1} \Pi_{ik}(X,H) = P_1(X)HP_1(X) - \sum_{i=1}^{r_1} e_i(X)e_i(X)^T He_i(X)e_i(X)^T.$$

Consequently

$$\lim_{X \to \bar{X}} \left\{ \begin{array}{l} \frac{1}{2}\sum_{\substack{i \neq k \\ i,k=1}}^{r_1} \dfrac{f(\lambda_i(X)) - f(\lambda_k(X))}{\lambda_i(X) - \lambda_k(X)} \Pi_{i,k}(X,H) \\ + \sum_{i=1}^{r_1} f'(\lambda_i(X))e_i(X)e_i(X)^T He_i(X)e_i(X)^T \end{array} \right\} = f'(\mu_1)\bar{P}_1 H \bar{P}_1.$$

Similar calculations can be performed if some of the eigenvalues $\lambda_i(X)$, $i \in \{1, ..., r_1\}$, are equal to each other, and for all $j \in \{1, ..., q\}$. Since continuity of the cross terms of $\mathbb{A}_f(X,H)$, for different values of the summation index $j$, clearly holds, continuity of $DF(X)$ then follows. ∎

**Theorem 4.2** *Suppose that the function $f(x)$ is locally Lipschitz continuous, directionally differentiable in a neighborhood of every $\mu_j$, and (strongly) semismooth at every point $\mu_j$, $j = 1, ..., q$. Then the corresponding matrix function $F(x)$ is (strongly) semismooth at $\bar{X}$.*

**Proof.** By Proposition 4.1 we have that $F(X)$ is locally Lipschitz continuous, and by Theorem 4.1 that $F(X)$ is directionally differentiable at $\bar{X}$. We give below derivations for the strong semismooth case, semismooth case can be derived in a similar way. Consider the decomposition (4.1) with $X := \bar{X} + H$. Since $P_j(\cdot)$ are twice continuously differentiable near $\bar{X}$ we have

$$\sum_{j=1}^{q} f(\mu_j)\left[P_j(X) - P_j(\bar{X})\right] = \sum_{j=1}^{q} f(\mu_j)DP_j(X)H + O(\|H\|^2). \tag{4.9}$$

Now since the eigenvalue functions $\lambda_k(\cdot)$ are strongly semismooth and $f(\cdot)$ are strongly semismooth at $\mu_j$ we have, for $k \in \{s_j + 1, ..., s_j + r_j\}$ and $j \in \{1, ..., q\}$, that

$$f(\lambda_k(X)) - f(\mu_j) = f'\big(\lambda_k(X), \lambda_k'(X,H)\big) + O(\|H\|^2). \tag{4.10}$$

Since $\|e_k(X)e_k(X)^T\|$ are uniformly bounded it follows that

$$\begin{aligned} F(X) - F(\bar{X}) \ &= \sum_{j=1}^{q} f(\mu_j) DP_j(X)H \\ &+ \sum_{i=1}^{p} f'\big(\lambda_i(X), \lambda_i'(X,H)\big) e_i(X)e_i(X)^T + O(\|H\|^2). \end{aligned} \tag{4.11}$$

We also have by the proof of Theorem 4.1 (see equations (4.5) and (4.6) in particular) that

$$\sum_{j=1}^{q} f(\mu_j) DP_j(X)H = \mathbb{A}_f(X,H) + O(\|H\|^2), \tag{4.12}$$

and, for an appropriate choice of $e_i(X)$,

$$\sum_{i=1}^{p} f'\big(\lambda_i(X), \lambda_i'(X,H)\big) e_i(X)e_i(X)^T = \mathbb{B}_f(X,H). \tag{4.13}$$

It follows that

$$F(X) - F(\bar{X}) = F'(X,H) + O(\|H\|^2),$$

Strong semismoothness of $F(X)$ at $\bar{X}$ then follows. ∎

**Example 4.1** Consider function $f(x) := x^2$. Then by (2.7) we have

$$DF(\bar{X})H = 2\sum_{j=1}^{q} \mu_j \bar{P}_j H \bar{P}_j + \tfrac{1}{2} \sum_{\substack{j \neq k \\ j,k=1}}^{q} \frac{\mu_j^2 - \mu_k^2}{\mu_j - \mu_k} \left(\bar{P}_j H \bar{P}_k + \bar{P}_k H \bar{P}_j\right). \tag{4.14}$$

Now, since $\sum_{j=1}^{q} \bar{P}_j = I_p$ and $\sum_{j=1}^{q} \mu_j \bar{P}_j = \bar{X}$, we have

$$\tfrac{1}{2} \sum_{\substack{j \neq k \\ j,k=1}}^{q} \frac{\mu_j^2 - \mu_k^2}{\mu_j - \mu_k} \left(\bar{P}_j H \bar{P}_k + \bar{P}_k H \bar{P}_j\right) = \tfrac{1}{2} \sum_{\substack{j \neq k \\ j,k=1}}^{q} (\mu_j + \mu_k) \left(\bar{P}_j H \bar{P}_k + \bar{P}_k H \bar{P}_j\right) =$$
$$\sum_{j=1}^{q} \left[\mu_j \bar{P}_j H (I_p - \bar{P}_j) + (\bar{X} - \mu_j \bar{P}_j) H \bar{P}_j\right] = \bar{X}H + H\bar{X} - 2\sum_{j=1}^{q} \mu_j \bar{P}_j H \bar{P}_j.$$

It follows that

$$DF(\bar{X})H = \bar{X}H + H\bar{X}. \tag{4.15}$$

Of course, we have here that $F(X) = X^2$ in the sense of the usual matrix multiplication, and the above formula (4.15) can be easily derived in a straightforward way.

**Example 4.2** Consider function $f(x) := |x|$ and the corresponding matrix function $F(X) = |X|$. We have here that $f(x)$ is differentiable everywhere except $x = 0$, and $f'(x) = \operatorname{sign}(x)$ for $x \neq 0$ and $f'(0, h) = |h|$. Suppose that the eigenvalues $\mu_1, ..., \mu_{n-1}$ are positive, $\mu_n = 0$ and $\mu_{n+1}, ..., \mu_q$ are negative. Then we have by (4.2) that

$$\begin{aligned} F'(\bar{X}, H) &= \sum_{j=1}^{q} \operatorname{sign}(\mu_j)\bar{P}_j H \bar{P}_j + \bar{E}_n \left|\bar{E}_n^T H \bar{E}_n\right| \bar{E}_n^T \\ &+ \tfrac{1}{2} \sum_{\substack{j \neq k \\ j,k=1}}^{q} \frac{|\mu_j| - |\mu_k|}{\mu_j - \mu_k} \left(\bar{P}_j H \bar{P}_k + \bar{P}_k H \bar{P}_j\right). \end{aligned} \tag{4.16}$$

If all eigenvalues of $\bar{X}$ are nonzero, i.e., $\bar{X}$ is nonsingular, then the term $\bar{E}_n \left|\bar{E}_n^T H \bar{E}_n\right| \bar{E}_n^T$ in the right hand side of (4.16) should be omitted.

It follows from (4.16) that

$$\begin{aligned} &|\bar{X}|F'(\bar{X}, H) + F'(\bar{X}, H)|\bar{X}| = \\ &2\sum_{j=1}^{q} \mu_j \bar{P}_j H \bar{P}_j + \tfrac{1}{2} \sum_{\substack{j \neq k \\ j,k=1}}^{q} \frac{|\mu_j| - |\mu_k|}{\mu_j - \mu_k} (|\mu_j| + \mu_k|) \left(\bar{P}_j H \bar{P}_k + \bar{P}_k H \bar{P}_j\right). \end{aligned} \tag{4.17}$$

Note that the right hand side of (4.17) coincides with the right hand side of (4.14), and hence

$$|\bar{X}|F'(\bar{X}, H) + F'(\bar{X}, H)|\bar{X}| = \bar{X}H + H\bar{X}. \tag{4.18}$$

If the matrix $X$ is nonsingular, i.e., all its eigenvalues are different from zero, then the linear system (4.18) defines $F'(\bar{X}, H)$ uniquely. This gives a formula for $F'(\bar{X}, H)$ by using the inverse of the corresponding linear operator. That formula was obtained in Sun and Sun [13, Theorem 4.6]. A general expression for $F'(\bar{X}, H)$ was also given in [13, Theorem 4.7] and strong semismoothness of $F(X) = |X|$ was shown in [13, Theorem 4.13].

## References

[1] R. Bellman, *Introduction to Matrix Analysis*, McGraw-Hill Book Company, New York, 1960.

[2] J.F. Bonnans and A. Shapiro, *Perturbation Analysis of Optimization Problems*, Springer, New York, 2000.

[3] F. Facchinei and J.S. Pang, *Finite-Dimensional Variational Inequalities and Complementarity Problems*, Springer, New York, 2002, forthcoming.

[4] T. Kato, *Perturbation Theory for Linear Operators*, Sringer-Verlag, New York, 1984.

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

[6] R. Mifflin, Semismooth and semiconvex functions in constrained optimization, *SIAM Journal on Control and Optimization*, 33 (1977), 957–972.

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

[8] J.S. Pang, D. Sun and J. Sun, Semismooth homeomorphisms and strong stability of semidefinite and Lorentz complementarity problems, *preprint*, 2002.

[9] H. Poincaré, Sur les groupes continus, *Trans. Cambridge Phil. Soc.*, 18 (1899), 220–225.

[10] A. Shapiro, Perturbation theory of nonlinear programs when the set of optimal solutions is not a singleton, *Applied Mathematics and Optimization*, 18 (1988), 215–229.

[11] A. Shapiro and M.K.H. Fan, On eigenvalue optimization, *SIAM J. Optimization*, 5 (1995), 552–569.

[12] D. Sun and J. Sun, Strong semismoothness of eigenvalues of symmetric matrices and its application to inverse eigenvalue, *preprint*, 2001.

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