Journal of Convex Analysis
Volume 30 (2023), No. 4, 1307–1317

On the Differentiability of Symmetric Matrix Valued Functions

Alexander Shapiro[*]
School of Industrial and Systems Engineering, Georgia Institute of Technology, Atlanta, U.S.A.
ashapiro@isye.gatech.edu

Dedicated to Roger J-B Wets on the occasion of his 85th birthday.

Received: March 29, 2022
Accepted: February 19, 2023

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.

Keywords: Matrix function, eigenvalues and eigenvectors, directional derivatives, semismooth mappings.

2010 Mathematics Subject Classification: 15A18, 15A16, 90C30.

1. Introduction

Consider the linear space $\mathbb{S}^{p\times p}$ of $p \times p$ symmetric matrices. With a real valued function $f : \mathbb{R} \to \mathbb{R}$ can be associated a matrix valued function $F : \mathbb{S}^{p\times p} \to \mathbb{S}^{p\times p}$ defined as follows. For a matrix $X \in \mathbb{S}^{p\times p}$ 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[1]

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

Of course, the orthonormal eigenvectors $e_1(X), ..., e_p(X)$ are not defined uniquely. However, the right hand side of (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) have been studied extensively (see, e.g., [1, Chapter 6] and references therein). In particular, for the functions $f(x) := x^n$, $n = 1, ...,$ 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 semi-

[*] This research was partly supported by AFOSR grant FA9550-22-1-0244.

[1] By $A^\top$ we denote the transpose of matrix (vector) A.

ISSN 0944-6532 / $ 2.50

smoothness properties of the function $f(x)$ (by different techniques this was analysed in Chen, Qi and Tseng [3]). For the function $f(x) := |x|$, directional differentiability and semismoothness of the corresponding matrix function, denoted $|X|$, was shown in [12]. 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 $\mathbb{S}^{p\times p}_+$ of $p \times p$ positive semidefinite symmetric matrices. Semismoothness of that mapping (metric projection) was shown in Pang, Sun and Sun [8]. To some extend 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* is denoted by $Dg(x)h = g'(x, h)$.

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 [7], of a mapping $g : \mathbb{R}^n \to \mathbb{R}^m$. We use the following definition (cf., [10])

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

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

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

$$g(x + h) - g(x) = g'(x + h, h) + O(\|h\|^2). \qquad \square \tag{3}$$

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 $G\colon \mathbb{R}^k \to \mathbb{R}^n$ is a twice continuously differentiable mapping and $G(\bar{x}) = 0$, then the composite mapping $g \circ G\colon \mathbb{R}^k \to \mathbb{R}^m$ is strongly semismooth at $\bar{x}$ (see [4],[10] for a discussion of these properties).

We use the following notation throughout the paper. For a matrix $X \in \mathbb{S}^{p\times p}$ 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 $\bar{X} \in \mathbb{S}^{p\times p}$. 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

$$s_1 := 0,\ s_2 := r_1,\ \dots, s_q := r_1 + ... + r_{q-1}. \tag{4}$$

Also by $E_j(X)$ we denote the $p \times r_j$ matrix whose columns are formed by the eigenvectors $e_{s_j+1}(X), ..., e_{s_j+r_j}(X)$, i.e.,

$$E_j(X) := [e_{s_j+1}(X), ..., e_{s_j+r_j}(X)],\ j = 1, ..., q, \tag{5}$$

and define

$$P_j(X) := E_j(X)E_j(X)^\top. \tag{6}$$

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

$$\bar{E}_j := E_j(\bar{X}) \text{ and } \bar{P}_j := P_j(\bar{X}) = \bar{E}_j\bar{E}_j^\top. \tag{7}$$

That is, $\bar{P}_j$ is the orthogonal projection matrix onto the eigenvector space corresponding to the eigenvalue μ_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,

$$\bar{X} = \sum_{j=1}^q \mu_j \bar{P}_j \text{ and } F(\bar{X}) = \sum_{j=1}^q f(\mu_j)\bar{P}_j. \tag{8}$$

For a number $x \in \mathbb{R}$ we denote by $\text{sign}(x)$ its sign, i.e., $\text{sign}(x) = 1$ if $x > 0$, $\text{sign}(x) = -1$ if $x < 0$ and $\text{sign}(x) = 0$ 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 μ_j, $j = 1, ..., q$. For $z \neq \mu_j$, $j = 1, ..., q$, we have that the matrix $zI_p - \bar{X}$ is nonsingular and

$$(zI_p - \bar{X})^{-1} = \sum_{j=1}^q (z - \mu_j)^{-1}\bar{P}_j. \tag{9}$$

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

$$\frac{1}{2\pi i}\int_{C_j} f(z)(zI_p - \bar{X})^{-1}dz = f(\mu_j)\bar{P}_j. \tag{10}$$

Consider the curve

$$C := \bigcup_{j=1}^q C_j. \tag{11}$$

It is well known that the eigenvalues $\lambda_1(X), ..., \lambda_p(X)$ are continuous, in fact even locally Lipschitz continuous, functions of $X \in \mathbb{S}^{p\times p}$. Consequently, for all X in a sufficiently small 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

$$F(X) = \frac{1}{2\pi i}\int_C f(z)(zI_p - X)^{-1}dz. \tag{12}$$

The idea of using the above representation for studying matrix functions is going back to Poincaré [9] and is used extensively in Kato [5].

The integral representation (12) 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 (12) for calculation of derivatives of $F(X)$. We have that if the matrix $A := (zI_p - \bar{X})$ is nonsingular, then for $H \in \mathbb{S}^{p\times p}$ in a neighborhood of the null matrix,

$$(A - H)^{-1} = A^{-1} + A^{-1}HA^{-1} + o(\|H\|), \tag{13}$$

and hence
$$DF(\bar{X})H = \frac{1}{2\pi i}\int_C f(z)(zI_p - \bar{X})^{-1}H(zI_p - \bar{X})^{-1}dz. \tag{14}$$

Together with (9) and (11) this implies

$$DF(\bar{X})H = \sum_{j,k,v=1}^{q} \kappa^{(1)}_{j,k,v}\bar{P}_j H \bar{P}_k, \tag{15}$$

where
$$\kappa^{(1)}_{j,k,v} = \frac{1}{2\pi i}\int_{C_v} f(z)(z-\mu_j)^{-1}(z-\mu_k)^{-1}dz. \tag{16}$$

The coefficients $\kappa^{(1)}_{j,k,v}$ can be written in a closed form by applying Cauchy's theorem to $f(z)$, that is

$$\kappa^{(1)}_{j,k,v} = \begin{cases} 0, & \text{if } \ v \neq j \text{ and } v \neq k, \\ \frac{f(\mu_j)}{\mu_j - \mu_k}, & \text{if } \ v = j \text{ and } v \neq k, \\ \frac{f(\mu_k)}{\mu_k - \mu_j}, & \text{if } \ v \neq j \text{ and } v = k, \\ f'(\mu_j), & \text{if } \ v = j \text{ and } v = k. \end{cases} \tag{17}$$

It follows that (cf., [5])

$$DF(\bar{X})H = \mathbb{A}_f(\bar{X}, H) + \sum_{j=1}^{q} f'(\mu_j)\bar{P}_j H \bar{P}_j, \tag{18}$$

where
$$\mathbb{A}_f(\bar{X}, H) := \tfrac{1}{2}\sum_{\substack{j\neq k\\ j,k=1}}^{q} \frac{f(\mu_j) - f(\mu_k)}{\mu_j - \mu_k}\left(\bar{P}_j H \bar{P}_k + \bar{P}_k H \bar{P}_j\right). \tag{19}$$

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

$$DP_j(\bar{X})H = \sum_{\substack{k\neq j\\ k=1}}^{q} (\mu_j - \mu_k)^{-1}\left(\bar{P}_j H \bar{P}_k + \bar{P}_k H \bar{P}_j\right). \tag{20}$$

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

Remark 2.1. In principle it is possible to employ the representation (12) in order to compute the higher order terms in the expansion of $F(X)$. For example, by using

$$(A-H)^{-1} = A^{-1} + A^{-1}HA^{-1} + A^{-1}HA^{-1}HA^{-1} + o(\|H\|^2),$$

the second order term can be written as follows

$$\frac{1}{2\pi i}\int_C f(z)(zI_p-\bar{X})^{-1}H(zI_p-\bar{X})^{-1}H(zI_p-\bar{X})^{-1}dz = \sum_{j,k,\ell,v=1}^{q} \kappa^{(2)}_{j,k,\ell,v}\bar{P}_j H\bar{P}_k H\bar{P}_\ell$$

where $$\kappa^{(2)}_{j,k,\ell,v} = \frac{1}{2\pi i}\int_{C_v} f(z)(z-\mu_j)^{-1}(z-\mu_k)^{-1}(z-\mu_\ell)^{-1}dz.$$

The coefficients $\kappa^{(2)}_{j,k,\ell,v}$ can be written in a closed form by applying Cauchy's theorem to $f(z)$. In particular, for v different from j,k and ℓ, we have that $\kappa^{(2)}_{j,k,\ell,v} = 0$; for $v=j=k=\ell$, we have $\kappa^{(2)}_{j,k,\ell,v} = \frac{1}{2}f''(\mu_j)$; for $v=j=k$ and $v\neq\ell$, we have $\kappa^{(2)}_{j,k,\ell,v} = f'(\mu_j)/(\mu_j-\mu_\ell)$; for $v=j$ and $v\neq k$, $v\neq\ell$, we have $\kappa^{(2)}_{j,k,\ell,v} = \frac{f'(\mu_j)}{(\mu_j-\mu_k)(\mu_j-\mu_\ell)}$. $\square$

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 $\Xi_j : \mathbb{S}^{p\times p}\to\mathbb{S}^{r_j\times r_j}$ 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) $D\Xi_j(\bar{X}) : \mathbb{S}^{p\times p}\to\mathbb{S}^{r_j\times r_j}$ is onto and

$$D\Xi_j(\bar{X})H = \bar{E}_j^\top H\bar{E}_j, \tag{21}$$

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

$$\lambda_{s_j+i}(X) = \nu_i(\Xi_j(X)), \quad i=1,...,r_j, \tag{22}$$

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

Consider a matrix $H\in\mathbb{S}^{p\times p}$. 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 (21) 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 $\bar{E}_j^\top H\bar{E}_j$ arranged in the decreasing order.*

Of course the above result is known, e.g., it is a particular case of a more general result given in [6, 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

$$\lambda_{s_j+i}(X) = \mu_j + \nu_i\left(\Gamma_j(X)\right), \tag{23}$$

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 $\nu_i : \mathbb{S}^{r_j \times r_j} \to \mathbb{R}$ 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 $\lambda_i : \mathbb{S}^{p\times p} \to \mathbb{R}$, $i = 1, ..., p$, are strongly semismooth at every $X \in \mathbb{S}^{p\times p}$.*

The above proposition implies that the spectral mapping $X \mapsto \sigma(X)$ is strongly semismooth at every $X \in \mathbb{S}^{p\times p}$. By a different technique this was shown in Sun and Sun [13].

In the subsequent analysis we will need the following construction. For a given (direction) matrix $H \in \mathbb{S}^{p\times p}$ and $j \in \{1, ..., q\}$ consider matrix $\bar{E}_j^\top H \bar{E}_j$. Note that the matrix $\bar{E}_j$ is defined up to an orthogonal rotation, i.e., it can be replaced by $\bar{E}_j Q$, where Q is an $r_j \times r_j$ orthogonal matrix. Of course, the eigenvalues of the matrix $Q^\top \bar{E}_j^\top H \bar{E}_j Q$ coincide with the corresponding eigenvalues of $\bar{E}_j^\top H \bar{E}_j$. We can choose orthogonal matrix Q in such a way that $Q^\top \bar{E}_j^\top H \bar{E}_j Q$ is diagonal. Such rotation of $\bar{E}_j^\top H \bar{E}_j$ to diagonal form depends on H and may still be not defined uniquely.

We denote by $\bar{E}_j(H)$ a particular matrix $\bar{E}_j$ such that $\bar{E}_j^\top H \bar{E}_j$ is diagonal.

Lemma 3.3. *For $j \in \{1, ..., q\}$, any accumulation point of $E_j(\bar{X} + tH)$, as $t \downarrow 0$, is matrix $\bar{E}_j = \bar{E}_j(H)$.*

Proof. By passing to the limit in the definition of matrix $E_j(\cdot)$ we obtain that any accumulation point of $E_{jt} := E_j(\bar{X} + tH)$, as $t \downarrow 0$, is a matrix $\bar{E}_j$. We have that

$$E_{jt}^\top(\bar{X} + tH)E_{jt} = E_{jt}^\top \bar{X} E_{jt} + tE_{jt}^\top H E_{jt} \tag{24}$$

is diagonal. Since the limit of the first term in the right hand side of (24) is diagonal, it follows that the limit of the second term in the right hand side of (24) is also diagonal. This completes the proof. □

4. Properties of matrix functions

In this section we study continuity and differentiability properties of the matrix functions $F(X)$, of the form (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 $P_j(X) = \sum_{k=s_j+1}^{s_j+r_j} e_k(X)e_k(X)^\top$, we can write

$$F(X) = \sum_{j=1}^{q} f(\mu_j)P_j(X) + \sum_{j=1}^{q}\sum_{k=s_j+1}^{s_j+r_j} \left[f(\lambda_k(X)) - f(\mu_j)\right] e_k(X)e_k(X)^\top. \tag{25}$$

It follows that $\|F(X) - F(\bar{X})\| \leq$

$$\leq \sum_{j=1}^{q} |f(\mu_j)|\, \|P_j(X) - P_j(\bar{X})\| + \sum_{j=1}^{q}\sum_{k=s_j+1}^{s_j+r_j} |f(\lambda_k(X)) - f(\mu_j)|\, \|e_k(X)e_k(X)^\top\|.$$

Since $\|e_k(X)e_k(X)^\top\|$ 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 $\psi_j(\cdot) := f'(\mu_j, \cdot)$, and let $\Psi_j : \mathbb{S}^{r_j\times r_j} \to \mathbb{S}^{r_j\times r_j}$ be the corresponding matrix functions, $j = 1, ..., q$.

Theorem 4.2. *Suppose that the function $f(x)$ is directionally differentiable at every point μ_j, $j = 1, ...q$. Then the corresponding matrix function $F(x)$ is directionally differentiable at $\bar{X}$ and*

$$F'(\bar{X}, H) = \mathbb{A}_f(\bar{X}, H) + \mathbb{B}_f(\bar{X}, H), \tag{26}$$

where $\mathbb{A}_f(\bar{X}, H)$ is defined in (19)*, $\bar{E}_j = \bar{E}_j(H)$ are such that $\bar{E}_j^\top H \bar{E}_j$ are diagonal, $j = 1, ..., q$, and*

$$\mathbb{B}_f(\bar{X}, H) := \sum_{j=1}^{q} \bar{E}_j \left[\Psi_j\left(\bar{E}_j^\top H \bar{E}_j\right)\right] \bar{E}_j^\top. \tag{27}$$

Proof. Consider the decomposition (25) with $X := \bar{X} + tH$. We have

$$\lim_{t\downarrow 0} t^{-1} \sum_{j=1}^{q} f(\mu_j)\left[P_j(\bar{X} + tH) - P_j(\bar{X})\right] = \sum_{j=1}^{q} f(\mu_j) DP_j(\bar{X})H, \tag{28}$$

and by (20):

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

$$= \sum_{1\leq j<k\leq q} \frac{f(\mu_j) - f(\mu_k)}{\mu_j - \mu_k}\left(\bar{P}_j H \bar{P}_k + \bar{P}_k H \bar{P}_j\right) = \mathbb{A}_f(\bar{X}, H). \tag{29}$$

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

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

By Lemma 3.3 we have that any accumulation point of $E_1(\bar{X} + tH)$ is a matrix $\bar{E}_1 = \bar{E}_1(H)$. 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: $f'\left(\mu_1, \lambda_k'(\bar{X}, H)\right) = f'\left(\mu_1, \bar{e}_k^\top H \bar{e}_k\right)$.

It follows that

$$\lim_{t\downarrow 0}\Delta_1(t)=\sum_{k=s_1+1}^{s_1+r_1} f'\left(\mu_1,\bar e_k^\top H\bar e_k^\top\right)\,\bar e_k\bar e_k^\top=\bar E_1\left[\Psi_1\left(\bar E_1^\top H\bar E_1\right)\right]\bar E_1^\top. \tag{30}$$

Note that the right hand side of (30) does not depend on a particular choice of the columns $e_1(\bar X),...,e_{r_1}(\bar X)$ of the matrix $\bar E_1=\bar E_1(H)$.

The same calculations can be performed for every $j\in\{1,...,q\}$. The required result then follows by (29) and (30). □

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

$$\bar E_j\left[\Psi_j\left(\bar E_j^\top H\bar E_j\right)\right]\bar E_j^\top=f'(\mu_j)\bar P_j H\bar P_j. \tag{31}$$

It follows then that $F'(\bar X,H)$ is linear in H, and is given by the right hand side of (18), 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 4.3. *Suppose that the function $f(x)$ is differentiable at every point μ_j, $j=1,...,q$. Then $F(X)$ is (Gâteaux) differentiable at $\bar X$ and its differential is given by formula* (18)*. If, moreover, $f(x)$ is continuously differentiable at every point μ_j, $j=1,...,q$, then $F(X)$ is continuously differentiable at $\bar X$.*

Proof. Gâteaux differentiability follows by the above discussion. So let us suppose that $f(x)$ is continuously differentiable at every point μ_j, $j=1,...,q$. Let $X\in\mathbb{S}^{p\times p}$ be sufficiently close to $\bar X$ and consider the terms corresponding to $j=1$ in (18). Suppose for the moment that $\lambda_1(X)>...>\lambda_{r_1}(X)$. We have then by the Mean Value Theorem that

$$\tfrac12\sum_{\substack{i\neq k\\ i,k=1}}^{r_1}\frac{f(\lambda_i(X))-f(\lambda_k(X))}{\lambda_i(X)-\lambda_k(X)}\,\Pi_{ik}(X,H)=\tfrac12\sum_{\substack{i\neq k\\ i,k=1}}^{r_1} f'(\eta_{ik})\,\Pi_{ik}(X,H), \tag{32}$$

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

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

We also have that as X tends to $\bar X$, the eigenvalues $\lambda_1(X),...,\lambda_{r_1}(X)$ tend to μ_1 and hence the derivatives $f'(\eta_{ik})$ tend to $f'(\mu_1)$. Moreover,

$$\tfrac12\sum_{\substack{i\neq k\\ i,k=1}}^{r_1}\Pi_{ik}(X,H)=P_1(X)HP_1(X)-\sum_{i=1}^{r_1}e_i(X)e_i(X)^\top He_i(X)e_i(X)^\top.$$

Consequently

$$\lim_{X\to\bar X}\left\{\begin{array}{l}\tfrac12\displaystyle\sum_{\substack{i\neq k\\ i,k=1}}^{r_1}\frac{f(\lambda_i(X))-f(\lambda_k(X))}{\lambda_i(X)-\lambda_k(X)}\,\Pi_{i,k}(X,H)\\ +\displaystyle\sum_{i=1}^{r_1}f'(\lambda_i(X))e_i(X)e_i(X)^\top He_i(X)e_i(X)^\top\end{array}\right\}=f'(\mu_1)\bar P_1H\bar P_1.$$

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 4.4. *Suppose that the function $f(x)$ is locally Lipschitz continuous, directionally differentiable in a neighborhood of every μ_j, and (strongly) semismooth at every point μ_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 4.2 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 (25) with $X := \bar{X} + H$. Since $P_j(\cdot)$ are twice continuously differentiable near $\bar{X}$ we have

$$\sum_{j=1}^{q} f(\mu_j) \left[P_j(X) - P_j(\bar{X})\right] = \sum_{j=1}^{q} f(\mu_j) DP_j(X)H + O(\|H\|^2). \tag{33}$$

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

$$f(\lambda_k(X)) - f(\mu_j) = f'\big(\lambda_k(X), \lambda_k'(X, H)\big) + O(\|H\|^2). \tag{34}$$

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

$$F(X) - F(\bar{X}) =$$

$$= \sum_{j=1}^{q} f(\mu_j) DP_j(X)H + \sum_{i=1}^{p} f'\big(\lambda_i(X), \lambda_i'(X, H)\big) e_i(X)e_i(X)^\top + O(\|H\|^2). \tag{35}$$

We also have by the proof of Theorem 4.2 (see equations (29) and (30) in particular) that

$$\sum_{j=1}^{q} f(\mu_j) DP_j(X)H = \mathbb{A}_f(X, H) + O(\|H\|^2), \tag{36}$$

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

$$\sum_{i=1}^{p} f'\big(\lambda_i(X), \lambda_i'(X, H)\big) e_i(X)e_i(X)^\top = \mathbb{B}_f(X, H). \tag{37}$$

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 4.5. Consider function $f(x) := x^2$. Then by (18) we have

$$DF(\bar{X})H = 2 \sum_{j=1}^{q} \mu_j \bar{P}_j H \bar{P}_j + \tfrac{1}{2} \sum_{\substack{j \neq k \\ j,k=1}}^{q} \frac{\mu_j^2 - \mu_k^2}{\mu_j - \mu_k} \left(\bar{P}_j H \bar{P}_k + \bar{P}_k H \bar{P}_j\right). \tag{38}$$

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

$$\frac{1}{2}\sum_{\substack{j\neq k\\ j,k=1}}^{q} \frac{\mu_j^2-\mu_k^2}{\mu_j-\mu_k}\left(\bar P_j H \bar P_k + \bar P_k H \bar P_j\right) = \frac{1}{2}\sum_{\substack{j\neq k\\ j,k=1}}^{q} (\mu_j+\mu_k)\left(\bar P_j H \bar P_k + \bar P_k H \bar P_j\right)$$

$$= \sum_{j=1}^{q}\left[\mu_j \bar P_j H (I_p - \bar P_j) + (\bar X - \mu_j \bar P_j) H \bar P_j\right] = \bar X H + H \bar X - 2\sum_{j=1}^{q}\mu_j \bar P_j H \bar P_j.$$

It follows that

$$DF(\bar X)H = \bar X H + H \bar X. \tag{39}$$

Of course, we have here that $F(X) = X^2$ in the sense of the usual matrix multiplication, and the above formula (39) can be easily derived in a straightforward way. □

Example 4.6. Consider function $f(x) := |x|$ and the corresponding matrix function $F(X) = |X|$. We have here that $f(x)$ is differentiable (in fact even analytic) everywhere except $x = 0$, and $f'(x) = \operatorname{sign}(x)$ for $x \neq 0$ and $f'(0,h) = |h|$. So suppose that $\mu_s = 0$ for some $s \in \{1,...,q\}$. Of course this implies that the corresponding matrix $\bar X$ is singular, and the respective eigenvalues $\mu_1,...,\mu_{s-1}$ are positive and $\mu_{s+1},...,\mu_q$ are negative. Then we have by (26) that

$$F'(\bar X, H) = \frac{1}{2}\sum_{\substack{j\neq k\\ j,k=1}}^{q} \frac{|\mu_j|-|\mu_k|}{\mu_j-\mu_k}\left(\bar P_j H \bar P_k + \bar P_k H \bar P_j\right) + \sum_{j=1}^{q}\operatorname{sign}(\mu_j)\bar P_j H \bar P_j + \bar E_s\left|\bar E_s^\top H \bar E_s\right|\bar E_s^\top, \tag{40}$$

where $\bar E_s = \bar E_s(H)$. It follows from (40) that

$$|\bar X| F'(\bar X, H) + F'(\bar X, H)|\bar X| = 2\sum_{j=1}^{q}\mu_j \bar P_j H \bar P_j + \frac{1}{2}\sum_{\substack{j\neq k\\ j,k=1}}^{q} \frac{|\mu_j|-|\mu_k|}{\mu_j-\mu_k}\left(|\mu_j|+\mu_k|\right)\left(\bar P_j H \bar P_k + \bar P_k H \bar P_j\right). \tag{41}$$

Note that the right hand side of (41) coincides with the right hand side of (38), and hence

$$|\bar X| F'(\bar X, H) + F'(\bar X, H)|\bar X| = \bar X H + H \bar X. \tag{42}$$

If the matrix $\bar X$ is nonsingular, i.e., all its eigenvalues are different from zero, then $F(\cdot)$ is differentiable at $\bar X$ with the derivative given by (18) - (19). That formula was obtained in Sun and Sun [12, Theorem 4.6]. A general expression for $F'(\bar X, H)$ was also given in [12, Theorem 4.7] and strong semismoothness of $F(X) = |X|$ was shown in [12, Theorem 4.13]. □

References

[1] R. Bellman: *Introduction to Matrix Analysis*, McGraw-Hill, New York (1960).

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

[3] X. Chen, H. Qi, P. Tseng: *Analysis of nonsmooth symmetric matrix valued functions with applications to semidefinite complementarity problems*, SIAM J. Optimization 4 (2003) 960–985.

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

[5] T. Kato: *Perturbation Theory for Linear Operators*, Sringer, New York (1984).

[6] P. Lancaster: *On eigenvalues of matrices dependent on a parameter*, Numer. Math. 6 (1964) 377–387.

[7] R. Mifflin: *Semismooth and semiconvex functions in constrained optimization*, SIAM J. Control Optim. 33 (1977) 957–972.

[8] J. S. Pang, D. Sun, J. Sun: *Semismooth homeomorphisms and strong stability of semidefinite and Lorentz complementarity problems*, Math. Oper. Research 28 (2003) 39–63.

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

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

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

[12] D. Sun, J. Sun: *Semismooth matrix-valued functions*, Math. Oper. Research 27 (2002) 150–169.

[13] D. Sun, J. Sun: *Strong semismoothness of eigenvalues of symmetric matrices and its application to inverse eigenvalue*, SIAM J. Numer. Analysis 6 (2003) 2352–2367.