The Second Order Directional Derivative of Symmetric Matrix-valued Functions

Liwei Zhang†, Ning Zhang‡ and Xiantao Xiao§

April 24, 2011

Abstract. This paper focuses on the study of the second-order directional derivative of a symmetric matrix-valued function of the form

\[ F(X) = P \text{diag}[f(\lambda_1(X)), \cdots, f(\lambda_n(X))]P^T. \]

For this purpose, we first adopt a direct way to derive the formula for the second-order directional derivative of any eigenvalue of a matrix in Torki [13]; Second, we establish a formula for the (parabolic) second-order directional derivative of the symmetric matrix-valued function. Finally, as an application, the second-order derivative for the projection operator over the SDP cone is used to derive the formula for the second-order tangent set of the SDP cone in Bonnans and Shapiro [3], which is the key for the Sigma term in the second-order optimality conditions of nonlinear SDP problems.

Key words. the SDP cone; symmetric matrix-valued function; second-order directional derivative; second-order tangent set.


1 Introduction

Löwner [6] initiated the work on symmetric matrix-valued functions of the form

\[ F(X) = P \text{diag}[f(\lambda_1(X)), \cdots, f(\lambda_n(X))]P^T, \]

where \( f : \mathbb{R} \rightarrow \mathbb{R} \) is a real-valued function, \( X \in \mathcal{S}^n \) is a symmetric matrix with the spectral decomposition

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

Realizing the importance of matrix-valued functions in the study of semidefinite programming, Sun and Sun [12] discussed some important differential properties of matrix-valued functions. In particular, they gave the formula for the directional derivative of the
projection operator over the SDP cone and demonstrated the strong semi-smoothness of this operator.

Pang, Sun and Sun [7] established the formula for the B-subdifferential of the projection operator over the SDP cone and gave a different way to derive the formula for the directional derivative of the projection operator.

Recently, Ding, Sun and Toh [2] gave a throughout study of the first order differential properties of matrix-valued functions, including the directional derivative, B-subdifferential, (G-)semismoothness and the counterparts of smoothing matrix-valued functions.

It is well known that the second-order tangent set of the SDP cone can be expressed by the second-order directional derivative of the maximum eigenvalue of a symmetric matrix (denoted \( \lambda_{\text{max}}(\cdot) \)), which is the key to describe the second-order optimality conditions, see [10] and [3]. The second-order directional derivative of \( \lambda_{\text{max}}(\cdot) \) can be derived from the classical paper of Lancaster [5], or the perturbation theory of optimization of Shapiro [9]. Importantly, based on the matrix perturbation theory of Stewart and Sun [11], Torki [13] obtained the formula for the second-order directional derivative of any eigenvalue of a symmetric matrix.

Although the formulas for the second-order directional derivative of any eigenvalue and the first-order derivative of the symmetric matrix-valued function, are available, the formula for second-order derivative of the symmetric matrix-valued function is still unknown. The main purpose of this paper is to establish the second-order derivative of matrix-valued functions and discuss its applications.

The paper is organized as follows. First, in the next section, we use a direct way, different from that of Torki [13], to derive the second-order directional derivative of any eigenvalue of a matrix-valued function whose implications will be used in studying the second-order derivative of matrix-valued functions. In Section 3, we present the formula for second-order derivative of symmetric matrix-valued functions. In Section 4, as an application, the second-order derivative for the projection operator over the SDP cone is used to derive the second-order tangent set of the SDP cone.

We introduce some notations to end this section. Let \( S^n \) be the space of all \( n \) by \( n \) symmetric and real matrices. For any \( X \in S^n \), we use \( \lambda_1(X) \geq \lambda_2(X) \geq \ldots \geq \lambda_n(X) \) to denote the real eigenvalues of \( X \). Let \( \Lambda(X) \in S^n \) be the diagonal matrix whose \( i \)-th diagonal entry is given by \( \lambda_i(X) \), \( i = 1, \ldots, n \), i.e., \( \Lambda(X) = \text{diag}(\lambda(X)) \). Denote by \( O^n \) the set of all \( n \times n \) orthogonal matrices in \( \mathbb{R}^{n \times n} \).

Let \( X \in S^n \) be given. Then, there exists an orthogonal matrix \( P \in O^n \) such that

\[
X = P \, \text{diag}(\lambda(X)) \, P^T.
\]  

(1.1)

The set of such matrices \( P \) in the eigenvalue decomposition (1.1) is denoted by \( O^n(X) \). Let \( \mu_1 > \mu_2 > \ldots > \mu_r \) be the distinct eigenvalues of \( X \). Define the following subsets of \( \{1, \ldots, n\} \)

\[
\alpha_k := \{ i \mid \lambda_i(X) = \mu_k \}, \quad k = 1, \ldots, r.
\]  

(1.2)

Partition \( P \) as \( P = [P_{\alpha_1} \ P_{\alpha_2} \ldots \ P_{\alpha_r}] \) with \( P_{\alpha_k} = (p_i : i \in \alpha_k) \), where \( P_{\alpha_k} \in \mathbb{R}^{n \times |\alpha_k|} \) for \( k = 1, \ldots, r \). For \( k \in \{1, \ldots, r\} \), assume that \( P_{\alpha_k}^T H P_{\alpha_k} \in \mathbb{R}^{|\alpha_k| \times |\alpha_k|} \) admits the following spectral decomposition

\[
(Q^k)^T (P_{\alpha_k}^T H P_{\alpha_k}) Q^k = \text{diag}(\xi_1^k, \ldots, \xi_{|\alpha_k|}^k).
\]
where $Q^k \in \mathcal{O}^{\alpha_k}(P^T\alpha_k HP_{\alpha_k})$ and $\xi^k_i = \lambda_i(P^T\alpha_k HP_{\alpha_k}), \ i = 1, \ldots, |\alpha_k|$. Denote the distinct eigenvalues of $P^T\alpha_k HP_{\alpha_k}$ by $\eta^k_1, \ldots, \eta^k_{N_k}$. Define the following subsets of $\{1, \ldots, |\alpha_k|\}$

$$\beta^k_i := \{i | \xi^k_i = \eta^k_i, \ i = 1, \ldots, |\alpha_k|\}. \quad (1.3)$$

Let $\kappa_i := \sum_{j=1}^i |\alpha_j|$ and $\kappa^{(k)}_i = \sum_{j=1}^i |\beta^k_j|$, for notation simplicity, we define the following mappings:

$$m_a : \{1, \ldots, n\} \rightarrow \{1, \ldots, r\}, \ m_a(i) = k, \ if \ i \in \alpha_k$$

$$l : \{1, \ldots, n\} \rightarrow \mathbb{N}, \ \ l(i) = i - \kappa_{m_a(i)} - 1,$$

$$m_b : \{1, \ldots, n\} \rightarrow \mathbb{N}, \ m_b(i) = p, \ if \ l(i) \in \beta^m_{\eta(i)},$$

$$l' : \{1, \ldots, n\} \rightarrow \mathbb{N}, \ \ l'(i) = l(i) - \kappa_{m_b(i)} - 1.$$  

For $i' \in \beta^k_p$, the corresponding index in $\{1, \ldots, n\}$ is $i = \kappa^{(k)}_{p-1} + i' + \kappa_{k-1}$.

## 2 Second order directional derivatives of eigenvalues

For a mapping $G : \mathcal{X} \rightarrow \mathcal{Y}$, where $\mathcal{X}$ and $\mathcal{Y}$ are two finite dimensional Hilbert spaces, we say $G$ is second-order directionally differentiable at $\mathbf{x} \in \mathcal{X}$ (see [8] or [3]), if $G$ is directionally differentiable at $\mathbf{x}$ and for any $h \in \mathcal{X}, w \in \mathcal{X}$ the limit

$$\lim_{t \searrow 0} \frac{G(\mathbf{x} + th + \frac{1}{2}t^2w) - G(\mathbf{x}) - tG'(\mathbf{x}; h)}{t^2}$$

exists; and the above limit is the (parabolic) second-order directional derivative, denoted by $G''(\mathbf{x}; h, w)$.

In this section, we discuss the second-order directional derivative of the matrix function $\lambda_i(\cdot)$, where $\lambda_i(X)$ is the $i$-th eigenvalue of the symmetric matrix $X \in \mathcal{S}^n$. For $X, H, W \in \mathcal{S}^n$, define $Y(t) := X + th + \frac{t^2}{2}W$. Let $Y(t)$ admit the following spectral decomposition

$$Y(t) = U(t)\Xi(t)U(t)^T, \quad U(t) \in \mathcal{O}^n(Y(t)),$$  

(2.1)

where $\Xi(t) := \Lambda(Y(t))$.

For simplicity, we first consider the case that $X = \Lambda(X)$, i.e., $X$ is a diagonal matrix in $\mathcal{S}^n$. Obviously, $I \in \mathcal{O}^n(X)$. Without loss generality, we assume that $P = I$, then $P_{\alpha_i} = E_{\alpha_i} = (e_{\kappa_i-1+1} \cdots e_m) \in \mathbb{R}^{n \times |\alpha_i|}$, where $e_j$ is the $j$-th unit vector of $\mathbb{R}^n$. Thus, one has that $P^T_{\alpha_k} HP_{\alpha_i} = H_{\alpha_i \alpha_i}$.

**Lemma 2.1** [2, Proposition 2.3] For any $H \in \mathcal{S}^n$, let $U$ be an orthogonal matrix such that

$$U^T(\Lambda(X) + H)U = \Lambda(\Lambda(X) + H).$$
Then, for any \( H \to 0 \), we have
\[
\begin{align*}
U_{\alpha_k \alpha_l} = & \ O(\|H\|), \quad k, l = 1, \ldots, r, k \neq l; \\
U_{\alpha_k \alpha_l} U_{\alpha_k \alpha_k}^T = & \ I_{\alpha_k} + O(\|H\|^2), \quad k = 1, \ldots, r; \\
\mathrm{dist}(U_{\alpha_k \alpha_k}, O^{\alpha_k}) = & \ O(\|H\|^2), \quad k = 1, \ldots, r,
\end{align*}
\]
and
\[
\lambda_i(\Lambda(X) + H) - \lambda_i(X) - \lambda_i(H_{\alpha_k \alpha_k}) = O(\|H\|^2), \quad i \in \alpha_k, \quad k = 1, \ldots, r.
\]
Hence, for any given direction \( H \in S^n \), the eigenvalue function \( \lambda_i(\cdot) \) is directionally differentiable at \( X \) with \( \lambda_i'(\Lambda(X); H) = \lambda_i(H_{\alpha_k \alpha_k}), i \in \alpha_k, k = 1, \ldots, r. \)

Directly from Lemma 2.1, we know that there exists \( Q^k(t) \in O^{\alpha_k} \) such that \( U_{\alpha_k \alpha_k}(t) = Q^k(t) + O(t^2) \), and
\[
(\Xi_{\alpha_k}(t) - \mu_I_{\alpha_k})^{-1} = \frac{1}{\mu_k - \mu_t} I_{\alpha_k} + \mathrm{diag}(O(t), \ldots, O(t)), \quad l \neq k. \tag{2.2}
\]

**Lemma 2.2** Let \( Y(t) \in S^n \) have the eigenvalue decomposition (2.1). For \( k, l \in \{1, \ldots, r\} \) with \( k \neq l \), there exists \( Q^l(t) \in O^{\alpha_l} \) such that
\[
\begin{align*}
U_{\alpha_k \alpha_l}(t) = & \ t \frac{H_{\alpha_k \alpha_l} Q^l(t)}{\mu_l - \mu_k} + O(t^2), \tag{2.3} \\
U_{\alpha_k \alpha_l}^T(t) U_{\alpha_k \alpha_k}(t) = & \ I_{\alpha_k} - t^2 \sum_{j \neq l} \frac{(Q^j(t))^T H_{\alpha_k \alpha_l}^T H_{\alpha_j \alpha_l} Q^j(t)}{(\mu_j - \mu_l)^2} + O(t^3). \tag{2.4}
\end{align*}
\]

**Proof.** From (2.1), we have \( Y(t) U(t) = U(t) \Xi(t) \), i.e.,
\[
\begin{pmatrix}
\mu_1 I_{\alpha_1} \\
\vdots \\
\mu_r I_{\alpha_r}
\end{pmatrix} + tH + \frac{t^2}{2} W =
\begin{pmatrix}
U_{\alpha_1 \alpha_1}(t) & \cdots & U_{\alpha_1 \alpha_k}(t) \\
\vdots & & \vdots \\
U_{\alpha_k \alpha_1}(t) & \cdots & U_{\alpha_k \alpha_k}(t)
\end{pmatrix}
\begin{pmatrix}
\Xi_1(t) \\
\vdots \\
\Xi_r(t)
\end{pmatrix}.
\]
Multiplying \( E_{\alpha_k}^T \) and \( E_{\alpha_l} \) to the left side and the right side of the above equation, respectively, yields that
\[
\mu_k U_{\alpha_k \alpha_l}(t) + t E_{\alpha_k}^T H U(t) E_{\alpha_l} + \frac{t^2}{2} E_{\alpha_k}^T W U(t) E_{\alpha_l} = U_{\alpha_k \alpha_l}(t) \Xi_{\alpha_l}(t).
\]
By the definition of \( E_{\alpha_k} \), we have
\[
E_{\alpha_k}^T H = [H_{\alpha_k \alpha_1} \cdots H_{\alpha_k \alpha_1} \cdots H_{\alpha_k \alpha_r}] \quad \text{and} \quad U(t) E_{\alpha_l} = \begin{pmatrix}
U_{\alpha_1 \alpha_l}(t) \\
\vdots \\
U_{\alpha_r \alpha_l}(t)
\end{pmatrix},
\]

4
and in turn,

$$\mu_k U_{\alpha_k\alpha_l}(t) + t \sum_{j=1}^{r} H_{\alpha_k\alpha_j} U_{\alpha_j\alpha_l}(t) + \frac{t^2}{2} E_{\alpha_k}^T W U(t) E_{\alpha_l} = U_{\alpha_k\alpha_l}(t) \Xi_{\alpha_l}(t).$$

This, together with Lemma 2.1 and (2.2) shows that

$$U_{\alpha_k\alpha_l}(t) = (t \sum_{j=1}^{r} H_{\alpha_k\alpha_j} U_{\alpha_j\alpha_l}(t) + \frac{t^2}{2} E_{\alpha_k}^T W U(t) E_{\alpha_l}) \Xi_{\alpha_l}(t) - \mu_k I_{|\alpha_l|}^{-1}$$

which proves (2.3). By using the equality $U_{\alpha_l}^T(t) U_{\alpha_l}(t) = I_{|\alpha_l|}$, we obtain (2.4) directly from (2.3).

**Proposition 2.1** For any $H, W \in S^n$, define

$$V_k(H, W) = E_{\alpha_k}^T [W - 2H(X - \mu_k I)^\top] E_{\alpha_k},$$

then, for any $i \in \alpha_k$, $k = 1, \ldots, r$,

$$\lambda_i(Y(t)) = \lambda_i(X) + t \lambda_i(H_{\alpha_k\alpha_k}) + \frac{t^2}{2} \lambda_{\nu(i)}((Q^k)^T V_k(H, W) Q^k_m(i)) + O(t^3),$$

(2.5)

where $Q^k \in \mathcal{O}^{\alpha_k}(P^T_{\alpha_k} H P_{\alpha_k})$. Namely,

$$\lambda_i(X; H) = \lambda_i(H_{\alpha_k\alpha_k})$$

(2.6)

and

$$\lambda''(X; H, W) = \lambda_{\nu(i)}((Q^k)^T V_k(H, W) Q^k_m(i)).$$

(2.7)
Proof. By Lemma 2.2, we have

\[ U_{\alpha_k}^T(t)A(X)U_{\alpha_k}(t) = \sum_{j=1}^{r} \mu_j U_{\alpha_j\alpha_k}^T(t)U_{\alpha_j\alpha_k}(t) \]

\[ = \mu_k U_{\alpha_k\alpha_k}^T(t)U_{\alpha_k\alpha_k}(t) + \sum_{j \neq k} \mu_j U_{\alpha_j\alpha_k}^T(t)U_{\alpha_j\alpha_k}(t) \]

\[ = \mu_k I_{|\alpha_k|} - t^2 \mu_k \sum_{j \neq k}^{r} \frac{(Q^k(t))^T H_{\alpha_j\alpha_k}^T H_{\alpha_j\alpha_k} Q^k(t)}{(\mu_j - \mu_k)^2} \]

\[ + t^2 \sum_{j \neq k}^{r} \mu_j \frac{(Q^k(t))^T H_{\alpha_j\alpha_k}^T H_{\alpha_j\alpha_k} Q^k(t)}{(\mu_j - \mu_k)^2} + O(t^3) \]

\[ = \mu_k I_{|\alpha_k|} + t^2 \sum_{j \neq k}^{r} \frac{(Q^k(t))^T H_{\alpha_j\alpha_k}^T H_{\alpha_j\alpha_k} Q^k(t)}{\mu_j - \mu_k} + O(t^3) \] (2.8)

and

\[ U_{\alpha_k}^T(t)H_{\alpha_k} U_{\alpha_k}(t) = \sum_{j \neq k}^{r} U_{\alpha_j\alpha_k}^T(t)H_{\alpha_k} U_{\alpha_k}(t) + \sum_{j=1, l \neq k}^{r} U_{\alpha_k\alpha_k}^T(t)H_{\alpha_k} U_{\alpha_k}(t) \]

\[ + \sum_{l=1}^{r} U_{\alpha_k\alpha_k}^T(t)H_{\alpha_k} U_{\alpha_k}(t) + U_{\alpha_k\alpha_k}^T(t)H_{\alpha_k} U_{\alpha_k}(t). \] (2.9)

It follows from Lemma 2.2 again that

\[ \sum_{j \neq k, l \neq k}^{r} U_{\alpha_j\alpha_k}^T(t)H_{\alpha_k} U_{\alpha_k}(t) = O(t^2), \] (2.10)

\[ \sum_{j=1}^{r} \sum_{j \neq k}^{r} U_{\alpha_j\alpha_k}^T(t)H_{\alpha_k} U_{\alpha_k}(t) = [Q^k(t) + O(t^2)]^T \sum_{j=1}^{r} H_{\alpha_j\alpha_j} (t) \frac{H_{\alpha_j\alpha_k} Q^k(t) + O(t^2)}{\mu_k - \mu_j} \]

\[ = t \sum_{j=1}^{r} \sum_{j \neq k}^{r} \frac{(Q^k(t))^T H_{\alpha_j\alpha_k}^T H_{\alpha_j\alpha_k} Q^k(t)}{\mu_k - \mu_j} + O(t^2). \] (2.11)

Similarly, the following relations hold:

\[ \sum_{l=1}^{r} \sum_{l \neq k}^{r} U_{\alpha_l\alpha_k}^T(t)H_{\alpha_k} U_{\alpha_k}(t) = t \sum_{l=1}^{r} \sum_{l \neq k}^{r} \frac{(Q^k(t))^T H_{\alpha_l\alpha_k}^T H_{\alpha_l\alpha_k} Q^k(t)}{\mu_k - \mu_l} + O(t^2), \] (2.12)

\[ U_{\alpha_k\alpha_k}^T(t)H_{\alpha_k\alpha_k} U_{\alpha_k\alpha_k}(t) = (Q^k(t) + O(t^2))^T H_{\alpha_k\alpha_k} (Q^k(t) + O(t^2)) \]

\[ = (Q^k(t))^T H_{\alpha_k\alpha_k} Q^k(t) + O(t^2). \] (2.13)
Combining (2.8) (2.9) (2.10) (2.11), (2.12) with (2.13), we get

\[ U_{\alpha_k}^T(t) H U_{\alpha_k}(t) = (Q^k(t))^T H_{\alpha_k,\alpha_k} Q^k(t) + 2 t \sum_{\substack{l=1 \atop l \neq k}}^r \frac{(Q^k(t))^T H_{\alpha_l,\alpha_k} H_{\alpha_l,\alpha_k} Q^k(t)}{\mu_l - \mu_k} + O(t^2). \]  

This, together with

\[ U_{\alpha_k}^T(t) W U_{\alpha_k}(t) = (Q^k(t))^T W_{\alpha_k,\alpha_k} Q^k(t) + O(t) \]

implies that

\[ \Xi_{\alpha_k}(t) = \mu_k I_{[\alpha_k]} + \frac{t^2}{2} (Q^k(t))^T [W_{\alpha_k,\alpha_k} - 2 \sum_{l \neq k} \frac{H_{\alpha_l,\alpha_k} H_{\alpha_l,\alpha_k}}{\mu_l - \mu_k}] Q^k(t) \]

\[ + t (Q^k(t))^T H_{\alpha_k,\alpha_k} Q^k(t) + O(t^3) \]

\[ = \mu_k I_{[\alpha_k]} + t (Q^k(t))^T H_{\alpha_k,\alpha_k} Q^k(t) + \frac{t^2}{2} (Q^k(t))^T V_k(H, W) Q^k(t) + O(t^3). \]  

(2.15)

It follows from (2.15) that

\[ H_{\alpha_k,\alpha_k} = \frac{1}{t} Q^k(t) [\Xi_{\alpha_k}(t) - \mu_k I_{[\alpha_k]}](Q^k(t))^T + O(t). \]  

(2.16)

Since for any \( t > 0 \), \( Q^k(t) \) is uniformly bounded, we may assume when \( t \searrow 0 \), \( \{Q^k(t)\} \) converges to an orthogonal matrix \( Q^k \). Then we have

\[ H_{\alpha_k,\alpha_k} = Q^k \text{ diag } (\lambda_i^k(X, H) : i \in \alpha_k)(Q^k)^T \]

\[ = Q^k \text{ diag } (\eta_1^k I_{[\beta_1]}^k, \ldots, \eta_N^k N_{N_k} I_{[\beta_N]}^k)(Q^k)^T. \]  

(2.17)

In fact, any cluster point \( Q^k \) of \( \{Q^k(t)\} \) when \( t \searrow 0 \) satisfies (2.17), or in other words, \( Q^k \in \mathcal{O}^{[\alpha_k]}(H_{\alpha_k,\alpha_k}). \) By (2.15), we have

\[ (Q^k)^T Q^k(t) [\Xi_{\alpha_k}(t)](Q^k(t))^T Q^k - \mu_k I_{[\alpha_k]} - t \text{ diag } (\eta_1^k I_{[\beta_1]}^k, \ldots, \eta_N^k I_{[\beta_N]}^k) \]

\[ = \frac{t^2}{2} (Q^k)^T V_k(H, W) Q^k + O(t^3). \]  

(2.18)

Then one has that

\[ \Lambda_{\alpha_k}(Y(t)) = \mu_k I_{[\alpha_k]} + t \Lambda(\text{ diag } (\eta_1^k I_{[\beta_1]}^k, \ldots, \eta_N^k I_{[\beta_N]}^k)) \]

\[ + \frac{t}{2} (Q^k)^T V_k(H, W) Q^k + O(t^3). \]  

(2.19)

Since the eigenvalue mapping \( \lambda(\cdot) \) is Lipschitz continuous, we obtain from (2.19) that, for any \( i \in \alpha_k, \)

\[ \lambda_i(Y(t)) = \lambda_i(X) + t \lambda_{i(i)}(H_{\alpha_k,\alpha_k}) + \frac{t^2}{2} \lambda_{i(i)}((Q^k_{m_0(i)})^T V_k(H, W) Q^k_{m_0(i)}) + O(t^3). \]
The proof is completed. 

Now we consider the general case in which $X$ is not necessary a diagonal matrix. Notice that, in this case $Y(t)$ can be expressed as

$$ Y(t) = P[\Lambda(X) + t\tilde{H} + \frac{1}{2}t^2\tilde{W}]P^T, \quad (2.20) $$

where

$$ \tilde{H} = P^T HP, \quad \tilde{W} = P^T WP, \quad (2.21) $$

we can easily obtain, from Proposition 2.1, the following formula of the second-order directional derivative of $\lambda_i(\cdot)$ at $X$ along $(H,W)$.

**Theorem 2.1** For any $H, W \in S^n$, define

$$ \tilde{V}_k(H, W) = P_{\alpha_k}^T[W - 2H(X - \mu_k I)^{\dagger}H]P_{\alpha_k}, \quad (2.22) $$

then, for any $i \in \alpha_k$, $k = 1, \ldots, r$,

$$ \lambda_i(Y(t)) = \lambda_i(X) + t\lambda_{l(i)}(\tilde{H}_{\alpha_k}) + \frac{t^2}{2}\lambda_{l(i)}((Q_{\alpha_k}^k(i))^{\dagger}\tilde{V}_k(H, W)Q_{\alpha_k}^k(i)) + O(t^3), \quad (2.23) $$

where $Q_k \in \mathcal{O}^{\alpha_k}(\tilde{H}_{\alpha_k})$. Namely,

$$ \lambda_i'(X; H) = \lambda_{l(i)}(\tilde{H}_{\alpha_k}) \quad (2.24) $$

and

$$ \lambda_i''(X; H, W) = \lambda_{l(i)}((Q_{\alpha_k}^k(i))^{\dagger}\tilde{V}_k(H, W)Q_{\alpha_k}^k(i)). \quad (2.25) $$

**Proof.** It is easy to check that

$$ V_k(\tilde{H}, \tilde{W}) = \tilde{V}_k(H, W), $$

and we obtain the conclusion from Proposition 2.1. 

**Remark 2.1** Formula (2.24) is derived in a similar way to [2]. Both (2.24) and (2.25) have been obtained by [13] by using the matrix perturbation theory of Stewart and Sun [11].

### 3 Second order directional derivative of the symmetric matrix-valued function

As the main part of the paper, this section is devoted to developing the formula for the second-order directional derivative of the symmetric matrix-valued function $F$ associated with a scalar function $f : \mathbb{R} \rightarrow \mathbb{R}$. Since $F$ is second-order directional derivative at $X \in S^n$ if and only if $f$ is second order directionally differentiable at each $\lambda_i(X)$ for $i = 1, \ldots, n$, we assume that $f$ is second-order directionally differentiable at each $\lambda_i(X)$ for $i = 1, \ldots, n$ through this section.
As in the previous section, we first consider the case that \( X \) is a diagonal matrix, namely, \( X = \Lambda(X) \). We choose \( P = I \in \mathcal{O}^n(X) \). Then \( P_{\alpha_k} = E_{\alpha_k} \) for \( k = 1, \ldots, r \), and \( P_{\alpha_k}^T H P_{\alpha_l} = H_{\alpha_k \alpha_l} \) for \( k, l = 1, \ldots, r \).

For a given \( Y \in \mathcal{S}^n \) near to \( X \), let \( Y \) admit the following eigenvalue decomposition,

\[
Y = U \Lambda(Y) U^T, \quad U \in \mathcal{O}^n(Y),
\]

define \( \mathcal{P}_k(Y) := \sum_{i \in \alpha_k} u_i u_i^T \).

For each \( k \in \{1, \ldots, r\} \), there exists \( \delta_k > 0 \) such that \(|\mu_l - \mu_k| > \delta_k, \forall l \neq k\). Define a continuously scalar function \( g_k(\cdot) : \mathbb{R} \rightarrow \mathbb{R} \) by (see for instance [2])

\[
g_k(t) = \begin{cases} 
-\frac{1}{6 \delta_k} (t - \mu_k - \frac{\delta_k}{2}) & t \in [\mu_k + \frac{\delta_k}{3}, \mu_k + \frac{\delta_k}{2}], \\
1 & t \in [\mu_k - \frac{\delta_k}{3}, \mu_k + \frac{\delta_k}{3}], \\
\frac{1}{6 \delta_k} (t - \mu_k + \frac{\delta_k}{2}) & t \in [\mu_k - \frac{\delta_k}{2}, \mu_k - \frac{\delta_k}{3}], \\
0 & \text{otherwise},
\end{cases}
\]

then, there exists \( \varepsilon > 0 \) such that

\[
\mathcal{P}_k(Y) = G_k(Y) \equiv \sum_{i=1}^n g_k(\lambda_i(Y)) u_i u_i^T, \quad \forall Y \in \mathcal{B}_\varepsilon(X).
\]

From [1, Exercise V3.9], we know that for any \( H, W \in \mathcal{S}^n \),

\[
\mathcal{P}_k^2(X)(H, W) = \sum_{1 \leq i,j,l \leq n} g_{k}^{[2]}(\lambda_i, \lambda_j, \lambda_l)(p_ip_i^T H p_j p_j^T W p_l p_l^T + p_l p_l^T W p_j p_j^T H p_i p_i^T), \quad (3.1)
\]

where \( g_k^{[2]}(\lambda_i, \lambda_j, \lambda_l) \) is the second divided differences (see e.g. [1]).

By the definition of the second divided differences , we get

\[
g_{k}^{[2]}(\lambda_i, \lambda_j, \lambda_l) = \begin{cases} 
\frac{1}{(\mu_k - \mu_p)(\mu_k - \mu_q)}, & i \in \mathcal{J}_k, j \in \mathcal{J}_p, l \in \mathcal{J}_q, k \neq p \neq q, \\
\frac{1}{(\mu_q - \mu_k)^2}, & i \in \mathcal{J}_k, j \in \mathcal{J}_q, l \in \mathcal{J}_q, k \neq q, \\
-\frac{1}{(\mu_q - \mu_k)^2}, & i \in \mathcal{J}_k, j \in \mathcal{J}_k, l \in \mathcal{J}_q, k \neq q, \\
0, & \text{otherwise},
\end{cases}
\]

where \( \mathcal{J}_i = \alpha_i, i = 1, \ldots, r \).

For each \( k \in \{1, \ldots, r\} \), \( \mathcal{P}_k(\cdot) \) is twice continuously differentiable and the second order derivative of \( \mathcal{P}_k(\cdot) \) is given by

\[
\mathcal{P}_k''(X; H, W) = \mathcal{P}_k'(X)W + \mathcal{P}_k''(X)(H, H),
\]

9
where

\[ P_k^r(X)W = \sum_{\substack{l=1 \atop l \neq k}}^r \frac{1}{\mu_k - \mu_l} (P_k(X)WP_l(X) + P_l(X)WP_k(X)), \quad k = 1, \ldots, r, \]

\[ P_k^r(X)(H, H) = 2 \sum_{\substack{l=1 \atop l \neq k}}^r \frac{1}{(\mu_k - \mu_l)^2} [\hat{H}_{kl} \hat{H}_{lk} + \hat{H}_{kl} \hat{H}_{kl} + \hat{H}_{kl} \hat{H}_{kl}] \]

\[ -2 \sum_{\substack{l=1 \atop l \neq k}}^r \frac{1}{(\mu_k - \mu_l)^2} [\hat{H}_{kk} \hat{H}_{kl} + \hat{H}_{kl} \hat{H}_{lk} + \hat{H}_{kl} \hat{H}_{kk}] \]

\[ +2 \sum_{\substack{l \leq j \leq r \atop l \neq k, j \neq k}}^r \frac{1}{(\mu_k - \mu_l)(\mu_k - \mu_j)} [\hat{H}_{kl} \hat{H}_{lj} + \hat{H}_{lk} \hat{H}_{lk} + \hat{H}_{lj} \hat{H}_{jl}]. \]

where \( \hat{P}_k = P_k(X)H_P(X) = P_{\alpha \kappa} P_{\alpha \kappa}^T H P_{\alpha \ell} P_{\alpha \ell}^T. \)

**Lemma 3.1** For \( l, k \in \{1, 2, \ldots, r\} \) and \( l \neq k \), it holds that

(i) \( (\Xi_{\alpha l}(t) - \mu_k I_{|\alpha l|})^{-1} = \frac{I_{|\alpha l|}}{\mu_l - \mu_k} - \frac{t}{(\mu_l - \mu_k)^2} \left( Q^1(t) H_{\alpha l \alpha l} Q^1(t) \right) + O(t^2), \)

(ii) \( U_{\alpha l \alpha l}(t) = \frac{1}{\mu_l - \mu_k} \left[ t H_{\alpha l \alpha l} - \frac{t^2}{(\mu_l - \mu_k)^2} \frac{H_{\alpha l \alpha l} H_{\alpha l \alpha l}}{\mu_l - \mu_k} - \sum_{j=1}^{r} H_{\alpha l \alpha j} H_{\alpha j \alpha l} - \frac{1}{2} W_{\alpha l \alpha l} \right] Q^1(t) + O(t^3). \)

**Proof.** From (2.16), we know that

\[ \Xi_{\alpha l}(t) = \mu_l I_{|\alpha l|} + t Q^1(t) P_{\alpha l \alpha l} Q^1(t) + O(t^2), \]

then

\[ (\Xi_{\alpha l}(t) - \mu_k I_{|\alpha l|})^{-1} = \frac{1}{\mu_l - \mu_k} (Q^1(t) J_{|\alpha l|} + \frac{t H_{\alpha l \alpha l}}{\mu_l - \mu_k})^{-1} Q^1(t) + O(t^2) \]

\[ = \frac{1}{\mu_l - \mu_k} (Q^1(t) J_{|\alpha l|} - \frac{t H_{\alpha l \alpha l}}{\mu_l - \mu_k} + \frac{t^2 H_{\alpha l \alpha l}^2}{(\mu_l - \mu_k)^2} + O(t^3)) Q^1(t) + O(t^2) \]

\[ = \frac{1}{\mu_l - \mu_k} I_{|\alpha l|} - \frac{t (Q^1(t) J_{|\alpha l|} Q^1(t))}{(\mu_l - \mu_k)^2} + O(t^2), \quad (3.2) \]

which proves (i).

Since \( E_{\alpha k}^T (X + tH + \frac{t^2}{2} W) E_{\alpha l} = U_{\alpha k \alpha l}(t) \Xi_{\alpha l}(t) U_{\alpha l \alpha l}^T(t) \), we get

\[ \mu_k U_{\alpha k \alpha l}(t) + t H_{\alpha k \alpha l} U_{\alpha l \alpha l} + t \sum_{j \neq l} H_{\alpha k \alpha j} U_{\alpha j \alpha l}(t) + \frac{t^2}{2} E_{\alpha k} W U_{\alpha l}(t) = U_{\alpha k \alpha l}(t) \Xi_{\alpha l}(t), \]

which together with the fact that \( U_{\alpha k \alpha l}(t) = O(t) \ (k \neq l) \), implies that

\[ U_{\alpha k \alpha l}(t) = (t H_{\alpha k \alpha l} U_{\alpha l \alpha l} + t \sum_{j \neq l} H_{\alpha k \alpha j} U_{\alpha j \alpha l}(t) + \frac{t^2}{2} E_{\alpha k} W U_{\alpha l}(t)) (\Xi_{\alpha l}(t) - \mu_k I_{|\alpha l|})^{-1} + O(t^3). \]
By Lemma 2.2 and (i), we have

\[ U_{\alpha \alpha_l}(t) = \left( \frac{t H_{\alpha_l \alpha_l}}{\mu_l - \mu_k} - \frac{t^2 H_{\alpha_l \alpha_l} H_{\alpha_l \alpha_l}}{(\mu_l - \mu_k)^2} + t^2 \sum_{j \neq l} \frac{H_{\alpha_l \alpha_j} H_{\alpha_l \alpha_j}}{(\mu_l - \mu_k)(\mu_l - \mu_j)} + \frac{t^2 W_{\alpha \alpha_l}}{2(\mu_k - \mu_l)} \right) Q^l(t) + O(t^3). \]

The proof is completed. \( \square \)

Before presenting the main result which is about the second order directional derivative of the symmetric matrix-valued function, some analysis is given in advance.

Define \( \hat{U}^k(t) := (Q^k)^T Q^k(t) \), then \( \hat{U}^k(t) \in \mathcal{O}[\alpha_k] \). By (2.18), we get

\[
\begin{align*}
&\text{diag} \left( \eta^k_I \eta^k_I, \ldots, \eta^k_{N_k} \eta^k_{N_k} \right) + \frac{t}{2} (Q^k)^T V_k(H, W) Q^k + O(t) \\
&= \hat{U}^k(t) \Lambda \left( \text{diag} \left( \eta^k_I \eta^k_I, \ldots, \eta^k_{N_k} \eta^k_{N_k} \right) + \frac{t}{2} [(Q^k)^T V_k(H, W) Q^k + O(t)] \right) (\hat{U}^k(t))^T.
\end{align*}
\]

Similar to Lemma 2.2, there exists \( \hat{Q}^k_{\beta_i \beta_j}(t) \in \mathcal{O}[\alpha_k] \) such that

\[
\hat{U}^k_{\beta_i \beta_j}(t) = \begin{cases} 
\hat{Q}^k_{\beta_i \beta_i}(t) + O(t^2), & \text{if } i = j; \\
\frac{t}{2(\eta^k_i - \eta^k_j)} (Q^k)^T V_k(H, W) Q^k \hat{Q}^k_{\beta_i \beta_j}(t) + O(t^2), & \text{if } i \neq j.
\end{cases}
\]

Define \( \hat{Q}^k(t) \in \mathbb{R}[\alpha_k \times \alpha_k] \),

\[
[\hat{Q}^k(t)]_{\beta_i \beta_j} = \begin{cases} 
\hat{Q}^k_{\beta_i \beta_i}(t), & \text{if } i = j; \\
\frac{t}{2(\eta^k_i - \eta^k_j)} (Q^k)^T V_k(H, W) Q^k \hat{Q}^k_{\beta_i \beta_j}(t), & \text{if } i \neq j.
\end{cases}
\]

It follows that

\[ Q^k(t) = Q^k \hat{U}^k(t) = Q^k \hat{Q}^k(t) + O(t^2). \]

Lemma 3.1 implies that

\[ U_{\alpha_l \alpha_k}(t) = \frac{1}{\mu_k - \mu_l} \left( t H_{\alpha_l \alpha_k} + \frac{t^2}{2} \Delta_{lk} \right) Q^k(t) + O(t^3), \quad l \neq k, \]

where

\[ \Delta_{lk} := W_{\alpha_l \alpha_k} + 2 \sum_{j \neq k} \frac{H_{\alpha_l \alpha_j} H_{\alpha_j \alpha_k}}{\mu_k - \mu_j} - 2 \frac{H_{\alpha_l \alpha_k} H_{\alpha_k \alpha_k}}{\mu_k - \mu_l}. \]

To simplify the notation, we use \( U_{\alpha_l \beta_k}(t) \) instead of \( U_{\alpha_l \alpha_k}(t) \), then

\[
\begin{align*}
U_{\alpha_l \beta_k}(t) &= \frac{1}{\mu_k - \mu_l} \left( t H_{\alpha_l \alpha_k} + \frac{t^2}{2} \Delta_{lk} \right) Q^k \hat{Q}^k_{\beta_k \beta_k}(t) + O(t^3) \\
&= \frac{1}{\mu_k - \mu_l} \left( t H_{\alpha_l \alpha_k} + \frac{t^2}{2} \Delta_{lk} \right) \left( Q^k \hat{Q}^k_{\beta_k \beta_k}(t) + \frac{t}{2} \sum_{j \neq k} Q^k \hat{Q}^k_{\beta_j \beta_j}(Q^k)^T V_k(H, W) Q^k \hat{Q}^k_{\beta_j \beta_j}(t) \right) + O(t^3) \\
&= \frac{1}{\mu_k - \mu_l} t H_{\alpha_l \alpha_k} Q^k \hat{Q}^k_{\beta_k \beta_k}(t) + O(t^2).
\end{align*}
\]
Therefore, we have

\[
Q^k_{\beta^i} (t) = \frac{Q^k_{\beta^i} (t) + O(t^2)}{\eta_t^k - \eta_j^k}
\]

Consequently, for each \(i \in \{1, \ldots, N_k\}\),

\[
U_{\beta^i} (t) = \left[ \frac{tH_{\alpha_i \alpha_k}}{\mu_k - \mu_1} \right]^{T} V_k (H, W) Q^k_{\beta^i} (t) + O(t^2).
\]

Define

\[
g(t) = \sum_{k=1}^{r} g_k(t) f(\mu_k).
\]

For any \(k, l \in \{1, \ldots, r\}\), careful calculation yields that

\[
\left( \sum_{i=1}^{r} f(\mu_i) P''(X)(H, H) \right)_{\alpha_k \alpha_l} = P^T_{\alpha_k} \left( \sum_{i=1}^{r} f(\mu_i) P''(X)(H, H) \right) P_{\alpha_l}
\]

\[
= \begin{cases} 
P^T_{\alpha_k} \left( \sum_{j \neq k} \frac{f(\mu_j) - f(\mu_k)}{(\mu_j - \mu_k)^2} \hat{H}_{kj} \hat{H}_{jk} \right) P_{\alpha_l}, & \text{if } l = k, \\
\sum_{j \neq k \neq l} g^{[2]}(\mu_k, \mu_j, \mu_l) \hat{H}_{kj} \hat{H}_{jl} + \frac{f(\mu_j) - f(\mu_k)}{(\mu_j - \mu_k)^2} [\hat{H}_{kl} \hat{H}_{il} - \hat{H}_{kk} \hat{H}_{kl}] P_{\alpha_l}, & \text{if } k > l.
\end{cases}
\]

Therefore, we have

\[
\left( \sum_{i=1}^{r} f(\mu_i) P''(X)(H, H) \right)_{\alpha_k \alpha_l} = P^T_{\alpha_k} \left( \sum_{j=1}^{r} g^{[2]}(\mu_k, \mu_j, \mu_l) \hat{H}_{kj} \hat{H}_{jl} \right) P_{\alpha_l}
\]

\[
= \sum_{j=1}^{r} g^{[2]}(\mu_k, \mu_j, \mu_l) P^T_{\alpha_k} H P_{\alpha_j} P^T_{\alpha_j} H P_{\alpha_l}
\]

\[
= P^T_{\alpha_k} H G_{kl}^{[2]} (X) H P_{\alpha_l}
\]
and
\[ \sum_{i=1}^{r} f(\mu_i)P_{ii}(X)(H, H) = \sum_{k=1}^{r} \sum_{l=1}^{r} P_{\alpha_k} P_{\alpha_l} H G_{kl}^{[2]}(X) P_{\alpha_l} P_{\alpha_l}^{T}, \]

where \( G_{kl}^{[2]}(X) := \sum_{j=1}^{r} g^{[2]}(\mu_k, \mu_j, \mu_l) P_{\alpha_j} P_{\alpha_l}^{T} \) is the Löwner operator defined by the function \( g_{kl}^{[2]}(\cdot) := g^{[2]}(\mu_k, \cdot, \mu_l). \)

From the above formulas, we have
\[
\sum_{i=1}^{n} f(\lambda_i(X))u_i(t)u_i^{T}(t) = \sum_{k=1}^{r} f(\mu_k)P_{k}(Y(t)) \\
= \sum_{k=1}^{r} f(\mu_k)[P_{k}(X) + tP_{k}(X)H + t^2(\frac{1}{2}P_{k}(X)W + P_{k}(X)(H, H)))] + O(t^3) \\
= F(X) + tP[\sigma](\Lambda(X)) \circ P^{T}HPP^{T} + t^2 \frac{1}{2} P[\sigma](\Lambda(X)) \circ P^{T}WP^{T} \\
+ t^2 \sum_{k=1}^{r} \sum_{l=1}^{r} P_{\alpha_k} P_{\alpha_l}^{T} H G_{kl}^{[2]}(X) H P_{\alpha_l} P_{\alpha_l}^{T} + O(t^3). \tag{3.5}
\]

It follows from (3.3) that
\[
t \sum_{i=1}^{n} f'(\lambda_i(X); X'(X; H))u_i(t)u_i^{T}(t) = t \sum_{k=1}^{r} \sum_{i=1}^{N_k} f'(\mu_k; \eta^k_i)U_{\beta_i}^{T}(t)U_{\beta_i}^{T}(t) = t \sum_{k=1}^{r} \Omega^k + O(t^3),
\]

where \( \Omega^k \in \mathbb{R}^{a_k \times |a_k|} \) with
\[
\Omega_{lk}^{k} = (\Omega_{kl}^{k})^{T} = \frac{tH_{\alpha_k \alpha_l}}{\mu_k - \mu_l} \sum_{i=1}^{N_k} f'(\mu_k, \eta^k_i)Q_{\beta_i}^{k} (Q_{\beta_i}^{k})^{T}, \quad l \neq k,
\]
\[
\Omega_{kk}^{k} = \sum_{i=1}^{N_k} f'(\mu_k, \eta^k_i)Q_{\beta_i}^{k} (Q_{\beta_i}^{k})^{T} + \frac{t}{2} \sum_{i=1}^{N_k} \sum_{j \neq i} Q_{\beta_i}^{k} (Q_{\beta_j}^{k})^{T} \frac{V_k(H, W)}{\eta^k_i - \eta^k_j} f'(\mu_k, \eta^k_i)Q_{\beta_i}^{k} (Q_{\beta_j}^{k})^{T} \\
+ \frac{t}{2} \sum_{i=1}^{N_k} \sum_{j \neq i} Q_{\beta_j}^{k} (Q_{\beta_j}^{k})^{T} \frac{V_k(H, W)}{\eta^k_i - \eta^k_j} f'(\mu_k, \eta^k_i)Q_{\beta_j}^{k} (Q_{\beta_j}^{k})^{T} + O(t^2),
\]
\[
\Omega_{lj}^{k} = O(t^2), \quad l, j \neq k.
\]

For each \( k \in \{1, \ldots, r\}, j = 1, \ldots, N_k \), there exists \( \delta^k_j > 0 \) such that \( |\eta^k_j - \eta^k_l| > \delta^k_j \), for \( l \neq j, l \in \{1, \ldots, N_k\} \). Define a continuously scalar function \( \phi^k_j(\cdot) : \mathbb{R} \rightarrow \mathbb{R} \) by
\[
\phi^k_j(t) = \begin{cases} 
-\frac{1}{6\delta^k_j}(t - \eta^k_j + \frac{\delta^k_j}{2}) & t \in [\eta^k_j + \frac{\delta^k_j}{3}, \eta^k_j + \frac{\delta^k_j}{2}], \\
1 & t \in [\eta^k_j - \frac{\delta^k_j}{3}, \eta^k_j + \frac{\delta^k_j}{3}], \\
\frac{1}{6\delta^k_j}(t - \eta^k_j + \frac{\delta^k_j}{2}) & t \in [\eta^k_j - \frac{\delta^k_j}{2}, \eta^k_j - \frac{\delta^k_j}{3}], \\
0 & \text{otherwise},
\end{cases}
\]
Define $\phi_k(t) := \sum_{j=1}^{N_k} f'(\mu_k, \eta_j^k) \phi_j^k(t)$, then we have from

$$H_{\alpha_k \alpha_k} = Q^k \begin{bmatrix} \eta_1^k I_{[\beta_1]} \\ \vdots \\ \eta_{N_k}^k I_{[\beta_{N_k}]} \end{bmatrix} (Q^k)^T$$

that

$$\sum_{i=1}^{N_k} f'(\mu_k, \eta_i^k) Q_{\beta_i^k}^k (Q_{\beta_i^k}^k)^T = Q^k \begin{bmatrix} \phi_k(\eta_1^k) I_{[\beta_1]} \\ \vdots \\ \phi_k(\eta_{N_k}^k) I_{[\beta_{N_k}]} \end{bmatrix} (Q^k)^T := \Phi_k(H_{\alpha_k \alpha_k}),$$

where $Q^k \in \mathcal{O}^{[\alpha_k]}(H_{\alpha_k \alpha_k})$. This together with the fact that

$$\sum_{i=1}^{N_k} \sum_{j \neq i} \left\{ Q_{\beta_i^k}^k (Q_{\beta_i^k}^k)^T \frac{\phi_k(\eta_i^k)}{\eta_i^k - \eta_j^k} V_k(H, W) Q_{\beta_i^k}^k (Q_{\beta_i^k}^k)^T + Q_{\beta_i^k}^k (Q_{\beta_i^k}^k)^T \frac{V_k(H, W) \phi_k(\eta_i^k)}{\eta_i^k - \eta_j^k} Q_{\beta_i^k}^k (Q_{\beta_i^k}^k)^T \right\}$$

implies that

$$t \sum_{k=1}^{r} \sum_{i=1}^{N_k} f'(\mu_k; \eta_i^k) U_{\beta_i^k}(t) U_{\beta_i^k}^T(t)$$

$$= t \begin{bmatrix} \Phi_1(H_{\alpha_1 \alpha_1}) \\ \Phi_2(H_{\alpha_2 \alpha_2}) \\ \vdots \\ \Phi_r(H_{\alpha_r \alpha_r}) \end{bmatrix} + \frac{t^2}{2} \Theta + O(t^3),$$

where, for $k, l = 1, \ldots, r$,

$$\Theta_{\alpha_i \alpha_k} = \begin{cases} \frac{2H_{\alpha_i \alpha_k}}{\mu_k - \mu_l} \Phi_k(H_{\alpha_i \alpha_k}) + \Phi_l(H_{\alpha_i \alpha_i}) \frac{2H_{\alpha_i \alpha_k}}{\mu_l - \mu_k}, & \text{if } l \neq k, \\
Q^k[\phi_l^k](\Lambda(H_{\alpha_i \alpha_k})) \circ (Q^k)^T V_k(H, W) Q^k(Q^k)^T, & \text{if } l = k, \end{cases}$$

and $h^{[1]}(\cdot)$ is the first divided difference matrix for any differentiable scalar function $h$ (see e.g. [1]).

From the definition $\hat{U}^k(t) = (Q^k)^T Q^k(t)$, (2.18) can be rewritten as

$$\hat{U}^k(t) \Xi_{\alpha_k}(k) (\hat{U}^k(t))^T - \mu_k I_{[\alpha_k]} - t \text{ diag } (\eta_1^k I_{[\beta_1]}, \ldots, \eta_{N_k}^k I_{[\beta_{N_k}]}),$$

$$= \frac{t^2}{2} (Q^k)^T V_k(H, W) Q^k + O(t^3),$$

(3.8)
and consequently,

\[
(Q_k^{(k)})^T V_k(H, W) Q_k^{(k)} = \frac{2}{t^2} ((\tilde{U}^k(t)\Xi_{\alpha_k}(t)\tilde{U}^k(t))^T \beta_i^k \beta_i^k - \mu_k I_{[\beta_i^k]} - t\eta_k I_{[\beta_i^k]}) + O(t). \tag{3.9}
\]

Let \( B_k(t) := \tilde{U}^k(t)\Xi_{\alpha_k}(t)(\tilde{U}^k(t))^T \), since for any \( i, j, l \in \{1, \ldots, N_k\} \),

\[
\tilde{U}^k_{\beta_i^k \beta_i^k}(t) = O(t), \quad i \neq j;
\]

\[
\tilde{U}^k_{\beta_i^k \beta_i^k}(t)^T \tilde{U}^k_{\beta_i^k \beta_i^k}(t) = I_{[\beta_i^k]} + O(t^2),
\]

and

\[
[\Xi_{\alpha_k}(t)]_{\beta_i^k} = \mu_k I_{[\beta_i^k]} + t\eta_k I_{[\beta_i^k]} + \frac{t^2}{2} (\Lambda_{\alpha_k})_{\beta_i^k}''(X; H, W) + O(t^3),
\]

we have

\[
(B_k(t))_{\beta_i^k \beta_i^k} = \tilde{U}^k_{\beta_i^k \beta_i^k}(t)\Xi_{\alpha_k}(t)_{\beta_i^k} (\tilde{U}^k_{\beta_i^k \beta_i^k}(t))^T + \sum_{j \neq i} \tilde{U}^k_{\beta_i^k \beta_j^k}(t)\Xi_{\alpha_k}(t)_{\beta_j^k} (\tilde{U}^k_{\beta_i^k \beta_j^k}(t))^T
\]

\[
= \mu_k I_{[\beta_i^k]} + t\eta_k I_{[\beta_i^k]} + \frac{t^2}{2} \tilde{U}^k_{\beta_i^k \beta_i^k}(t)(\Lambda_{\alpha_k})_{\beta_i^k}''(X; H, W)(\tilde{U}^k_{\beta_i^k \beta_i^k}(t))^T + O(t^3).
\]

Consequently,

\[
(Q_k^{(k)})^T V_k(H, W) Q_k^{(k)} = \tilde{U}^k_{\beta_i^k \beta_i^k}(t)(\Lambda_{\alpha_k})_{\beta_i^k}''(X; H, W)(\tilde{U}^k_{\beta_i^k \beta_i^k}(t))^T + O(t).
\]

Let \( \tilde{Q}_k^{(k)} \) be the cluster of \( \tilde{U}^k_{\beta_i^k \beta_i^k}(t) \) as \( t \downarrow 0 \), then \( \tilde{Q}_k^{(k)} \in O[\beta_i^k] \) and

\[
(Q_k^{(k)})^T V_k(H, W) Q_k^{(k)} = \tilde{Q}_k^{(k)} (\Lambda_{\alpha_k})_{\beta_i^k}''(X; H, W)(\tilde{Q}_k^{(k)})^T \beta_i^k \beta_i^k ,
\]

i.e., \( \tilde{Q}_k^{(k)}(t) \in O[\beta_i^k]((Q_k^{(k)})^T V_k(H, W) Q_k^{(k)}), \Lambda((Q_k^{(k)})^T V_k(H, W) Q_k^{(k)}) = (\Lambda_{\alpha_k})_{\beta_i^k}''(X; H, W).\)

Define \( \psi_{k,p}() := f''(\mu_k; \eta_{k-p}; \cdot) \) and \( j' := \kappa_{k-1} + \kappa_{p-1} + j \), then by Theorem 2.1,

\[
\frac{t^2}{2} \sum_{i=1}^{n} f''(\lambda_i(X); \lambda_i'(X, H), \lambda_i''(X, H, W))u_i(t)u_i^T(t)
\]

\[
= \frac{t^2}{2} \sum_{k=1}^{r} \sum_{p=1}^{N_k} \sum_{j' \in \beta_p} \psi_{k,p}(\lambda_j((Q_k^{(k)})^T V_k(H, W) Q_k^{(k)})) u_{j'}(t)u_{j'}^T(t). \tag{3.11}
\]

Observing that

\[
\sum_{p=1}^{N_k} \sum_{j' \in \beta_p} \psi_{k,p}(\lambda_j((Q_k^{(k)})^T V_k(H, W) Q_k^{(k)})) u_{j'}(t)u_{j'}^T(t) = \begin{bmatrix} O(t^2) & O(t) & O(t^2) \\
O(t) & R & O(t) \\
O(t^2) & O(t) & O(t^2) \end{bmatrix},
\]

15
with
\[
R = \sum_{p=1}^{N_k} Q_{\beta_p}^k \left[ \sum_{j \in \beta_p} \psi_{k,p}(\lambda_j [(Q_{\beta_p}^k)^T V_k(H,W)Q_{\beta_p}^k]) (\tilde{Q}_{\beta_p}^k)^T (Q_{\beta_p}^k)^T \right] + o(1)
\]
\[
= \sum_{p=1}^{N_k} Q_{\beta_p}^k \tilde{\Psi}_{k,p} ((Q_{\beta_p}^k)^T V_k(H,W)Q_{\beta_p}^k) (Q_{\beta_p}^k)^T + o(1)
\]
\[
= Q^k \text{diag} \left( \tilde{\Psi}_{k,1} ((Q_{\beta_p}^k)^T V_k(H,W)Q_{\beta_p}^k), \ldots, \tilde{\Psi}_{k,N_k} ((Q_{\beta_p}^k)^T V_k(H,W)Q_{\beta_p}^k) \right) (Q^k)^T + o(1),
\]
we get
\[
\frac{t^2}{2} \sum_{i=1}^{n} f''(t) = \frac{t^2}{2} A + o(t^2), \quad (3.12)
\]
where
\[
A_{\alpha \lambda \alpha \lambda} = \begin{cases} 
0, & \text{if } k \neq l, \\
Q^k \text{diag} \left( \tilde{\Psi}_{k,1} ((H,W), \ldots, \tilde{\Psi}_{k,N_k} ((H,W)) \right) (Q^k)^T, & \text{if } k = l.
\end{cases}
\]
with \( \tilde{\Psi}_{k}^i (H,W) := (Q_{\beta_p}^k)^T V_k(H,W)Q_{\beta_p}^k, i = 1, \ldots, |N_k| \).

For any \( X \in \mathbb{R}^{m \times n} \) which is not necessary a diagonal matrix, in this case \( Y(t) \) can be expressed as (2.20), where \( \bar{H} \) and \( \bar{W} \) are defined by (2.21), namely \( \bar{H} = P^T HP \) and \( \bar{W} = P^T WP \). Based on the above analysis for the case that \( X \) is a diagonal matrix, we are able to establish the following formula of the second-order directional derivative of \( F(\cdot) \) at \( X \) along \( (H,W) \).

**Theorem 3.1** Let \( X \in \mathbb{S}^n \) be given (it is not necessary a diagonal matrix) and have the eigenvalue decomposition (1.1). The matrix valued function \( F(\cdot) \) is second order directionally differentiable at \( X \) if and only if the corresponding Löwner operator \( f(\cdot) \) is second order directionally differentiable at every point \( \lambda_i(X), i = 1, \ldots, n \). In this case, the first and the second order directional derivative of \( F(\cdot) \) at \( X \) along \( H \) is

\[
F'(X,H) = P \begin{bmatrix} 
\Phi_1(\bar{H}_{\alpha_1 \alpha_2}) & f[1](\mu_1, \mu_2) \bar{H}_{\alpha_1 \alpha_2} & \cdots & f[1](\mu_1, \mu_r) \bar{H}_{\alpha_1 \alpha_r} \\
\vdots & \ddots & \ddots & \vdots \\
f[1](\mu_r, \mu_1) \bar{H}_{\alpha_r \alpha_1} & f[1](\mu_r, \mu_2) \bar{H}_{\alpha_r \alpha_2} & \cdots & \Phi_r(\bar{H}_{\alpha_r \alpha_r})
\end{bmatrix} P^T. \quad (3.13)
\]

And the second order directional derivative of \( F(\cdot) \) at \( X \) along \( (H,W) \) is

\[
F''(X;H,W) = PBP^T, \quad (3.14)
\]
where
\[ B_{\alpha k\alpha k} = Q^k[\partial^1_k(\Lambda(\tilde{H}_{\alpha k\alpha k})) \circ (Q^k)^T \tilde{V}_k(H, W)Q^k](Q^k)^T \]
\[ + Q^k \text{diag} (\Psi_{k, 1}((Q^k)^T \tilde{V}_k(H, W)Q^k_{\beta \beta}), \ldots, \Psi_{k, N_k((Q^k)^T \tilde{V}_k(H, W)Q^k_{\beta \beta})))(Q^k)^T, \]
\[ B_{\alpha k\alpha l} = g^{[1]}(\Lambda(X))_{\alpha k\alpha l} \circ \tilde{W}_{\alpha k\alpha l} + P_{\alpha k} \tilde{H} G_{\kappa}^{[2]}(X) \tilde{H} P_{\alpha l} \]
\[ + \frac{2\tilde{H}_{\alpha k\alpha l} \Phi_1(\tilde{H}_{\alpha k\alpha l})}{\mu_k - \mu_l} + \frac{2\Phi_k(\tilde{H}_{\alpha k\alpha l})\tilde{H}_{\alpha k\alpha l}}{\mu_k - \mu_l}, \text{ for } k, l = 1, \ldots, r \text{ and } k \neq l, \]
and \( \tilde{V}_k(H, W) \) is defined by (2.22) for \( k = 1, \ldots, r \).

**Proof.** "\( \Leftarrow \)" For any \( H, W \in S^n \) and \( t > 0 \), let \( Z(t) = X + tH + \frac{t^2}{2}W \) and \( Y(t) = P^T Z(t) P \). Then \( Y(t) = \Lambda(X) + t\tilde{H} + \frac{t^2}{2}\tilde{W} \) and \( F(Z(t)) = PF(Y(t))P^T \). Let \( Y(t) \) have the eigenvalue decomposition \( Y(t) = U(t) \text{diag}(\lambda(Y(t)))U(t)^T \), then \( Y(t) = U(t) \text{diag}(\lambda(Z(t)))U(t)^T \). Since \( f \) is second order directionally differentiable at every point \( \lambda_i(X) \), noting that \( \lambda'_i(\Lambda(X); \tilde{H}) = \lambda'_i(X; H) \) and \( \lambda''_i(\Lambda(X); \tilde{H}, \tilde{W}) = \lambda''_i(X; H, W) \), we have
\[
F(Y(t)) = \sum_{i=1}^{n} f(\lambda_i(Y(t)))u_i(t)u_i^T(t)
\]
\[ = \sum_{i=1}^{n} f(\lambda_i(X) + t\lambda'_i(\Lambda(X); \tilde{H}) + \frac{t^2}{2}\lambda''_i(\Lambda(X); \tilde{H}, \tilde{W}) + O(t^3))u_i(t)u_i^T(t)
\]
\[ = \sum_{i=1}^{n} f(\lambda_i(X))u_i(t)u_i^T(t) + t \sum_{i=1}^{n} f'(\lambda_i(X); \lambda'_i(\Lambda(X); \tilde{H}))u_i(t)u_i^T(t)
\]
\[ + \frac{t^2}{2} \sum_{i=1}^{n} f''(\lambda_i(X); \lambda'_i(\Lambda(X); \tilde{H}); \lambda''_i(\Lambda(X); \tilde{H}, \tilde{W}))u_i(t)u_i^T(t) + O(t^3).\]
From this expression, (3.5), (3.7) and (3.12), we get the conclusions.

"\( \Rightarrow \)" Suppose that \( F \) is second order differentially differentiable at \( X \). Fix any \( P \in O^n(X) \). For each \( i \in \{1, \ldots, n\} \), \( h_i \in \mathbb{R} \) and \( w_i \in \mathbb{R} \), let \( H = P \text{diag}(0, \ldots, h_i, \ldots, 0)P^T \) and \( W = P \text{diag}(0, \ldots, w_i, \ldots, 0)P^T \). It follows from Proposition 2.6 in [2] that \( f \) is directionally differentiable at \( \lambda_i(X) \) and
\[ F'(X; H) = P \text{diag}(0, \ldots, f'(\lambda_i(X); h_i), \ldots, 0)P^T.\]
Consequently, we have
\[
F''(X; H, W) = \lim_{t \downarrow 0} \frac{F(Y(t)) - F(X) - tF'(X; H)}{t^2} \]
\[ = P \text{diag}(0, \ldots, \lim_{t \downarrow 0} \frac{f(\lambda_i(Y(t))) - f(\lambda_i(X)) - tf'(\lambda_i(X); h_i)}{t^2} , \ldots, 0)P^T \]
\[ = P \text{diag}(0, \ldots, \lim_{t \downarrow 0} \frac{f(\lambda_i(X) + th_i + \frac{1}{2}t^2w_i) - f(\lambda_i(X)) - tf'(\lambda_i(X); h_i)}{t^2} , \ldots, 0)P^T \]
\[ = \frac{1}{2} \sum_{i=1}^{n} \frac{f''(\lambda_i(X); \lambda'_i(\Lambda(X); \tilde{H}); \lambda''_i(\Lambda(X); \tilde{H}, \tilde{W}))}{\mu_i - \mu_i}.\]
which implies that \( f''(\lambda_i(X); h_i, w_i) \) exists and

\[
f''(\lambda_i(X); h_i, w_i) = p_i^T F''(X; H, W)p_i.
\]

The proof is completed. \( \square \)

4 Applications

In this section, we apply Theorem 3.1 to metric project operator \( \Pi_{S^\alpha} \) and get the expressions of tangent cone \( T_{S^\alpha}(X) \) and second order tangent set \( T^2_{S^\alpha}(X, H) \) in a different way from that used in [3]. Without loss generality, we assume that \( X \in S^\alpha \) and \( X \) have the eigenvalue decomposition (1.1). Let \( \alpha_k, \beta_k, k = 1, \ldots, r, i = 1, \ldots, |\alpha_k|, \) be the corresponding subsets given by (1.2) and (1.3), respectively.

It follows from Moreau decomposition theorem that for any \( Y \in S^\alpha \),

\[
Y = \Pi_{S^\alpha}(Y) + \Pi_{S^\alpha}(Y).
\]

By the definition of tangent cone, we know that for any \( X \in S^\alpha \),

\[
T_{S^\alpha}(X) = \{ H \in S^\alpha : \text{dist}(X + tH, S^\alpha) = o(t) \}
= \{ H \in S^\alpha : ||\Pi_{S^\alpha}(X + tH) - (X + tH)|| = o(t) \}
= \{ H \in S^\alpha : \frac{||\Pi_{S^\alpha}(X + tH) - \Pi_{S^\alpha}(X)||}{t} = o(1) \}
= \{ H \in S^\alpha : \Pi'_{S^\alpha}(X; H) = 0 \}
\]

Assume that \( H \in T_{S^\alpha}(X) \), then

\[
T^2_{S^\alpha}(X; H) = \{ W \in S^\alpha : \text{dist}(X + tH + \frac{t^2}{2} W, S^\alpha) = o(t^2) \}
= \{ W \in S^\alpha : ||\Pi_{S^\alpha}(X + tH + \frac{t^2}{2} W) - (X + tH + \frac{t^2}{2} W)|| = o(t^2) \}
= \{ W \in S^\alpha : \frac{||\Pi_{S^\alpha}(X + tH + \frac{t^2}{2} W) - \Pi_{S^\alpha}(X) - \Pi_{S^\alpha}(X, H)||}{\frac{t^2}{2}} = o(1) \}
= \{ W \in S^\alpha : \Pi''_{S^\alpha}(X; H, W) = 0 \}.
\]

Let \( F(X) = \Pi_{S^\alpha}(X) \), the corresponding L"owner operator is \( f(x) = \max\{0, x\} \). An elementary calculation shows that

\[
f'(x; h) = \begin{cases} h, & \text{if } x > 0; \\ f(h), & \text{if } x = 0; \\ 0, & \text{if } x < 0, \end{cases}
\] and

\[
f''(x; y, z) = \begin{cases} z, & \text{if } x > 0 \text{ or } x = 0, y > 0; \\ 0, & \text{if } x < 0 \text{ or } x = 0, y < 0; \\ f(z), & \text{if } x = y = 0, \end{cases}
\]

and consequently, for \( k \in \{1, \ldots, r\}, i, j \in \{1, \ldots, |\alpha_k|\}, \)

\[
\phi_k(\eta^k; \mu^k) = \begin{cases} f(\eta^k), & \text{if } \mu^k = 0; \\ 0, & \text{if } \mu^k < 0, \end{cases}
\]
and

\[
\phi^{[1]}_k(\eta^k, j^k) = 0. \quad (4.1)
\]
We assume that \( \mu_1 = 0 \), then (4.1) and \( f^{[1]}(\mu_k, \mu_l) = 0 \) \((k \neq l)\) imply that \[
\Pi_{S^n_+}(X; H) = P \begin{bmatrix} Q^1 \Pi_{S^n_+}^{[1]}(\tilde{H}_{\alpha_1 \alpha_1})(Q^1)^T & 0 \\ 0 & 0 \end{bmatrix} P^T.
\]
This means that \( \Pi_{S^n_+}(X; H) = 0 \) is equivalent to \( \tilde{H}_{\alpha_1 \alpha_1} \leq 0 \). Obviously, the above conclusion also holds when \( \mu_1 < 0 \). Therefore,
\[
\mathcal{T}_{S^n_+}(X) = \{ H \in S^n : P^T_{\alpha_1} H P_{\alpha_1} \preceq 0 \}.
\]
Let \( \tilde{V}_k^p(H, W) := (Q^k_{\beta_p})^T \tilde{V}_k(H, W) Q^k_{\beta_p}, \) \( p = 1, \ldots, N_k \), then, for each \( i \in \{1, \ldots, |\beta_p^k|\} \)
\[
\psi_{k,p}(\lambda_i(\tilde{V}_k^p(H, W))) = \begin{cases} 0, & \text{if } \mu_k < 0 \text{ or } \mu_k = 0, \eta^k_p < 0; \\ f(\lambda_i(\tilde{V}_k^p(H, W))), & \text{if } \mu_k = \eta^k_p = 0, \end{cases}
\]
and
\[
\Psi_{k,p}(\tilde{V}_k^p(H, W)) = \tilde{Q}^k_{\beta_p \beta_p} \text{diag} \left( \psi_{k,p}(\lambda_1(\tilde{V}_k^p(H, W))), \ldots, \psi_{k,p}(\lambda_{|\beta_p^k|}(\tilde{V}_k^p(H, W))) \right) (\tilde{Q}^k_{\beta_p \beta_p})^T.
\]
The definition of \( g(t) \) and \( X \in S^n \) imply that
\[
g^{[1]}(\Lambda(X))_{\alpha_k \alpha_l} = 0, \quad \forall k, l = 1, \ldots, r; \\
g^{[2]}_{kl}(X) = 0, \quad \forall k \neq l.
\]
Assume that \( \mu_1 = 0 \), then
\[
\Pi_{S^n_+}^\prime(X; H, W) = P \begin{bmatrix} Q^1 M(Q^1)^T & 0 \\ 0 & 0 \end{bmatrix} P^T,
\]
where
\[
M = \begin{bmatrix} \tilde{Q}^1_{\beta_1 \beta_1} \Pi_{S^n_+}^{[1]}(\tilde{V}_1^1(H, W)) (\tilde{Q}^1_{\beta_1 \beta_1})^T & 0 \\ 0 & 0 \end{bmatrix}.
\]
This means that \( \Pi_{S^n_+}^\prime(X; H, W) = 0 \) is equivalent to \( \tilde{V}_1^1(H, W) \preceq 0 \). Obviously the above conclusion also holds when \( \mu_1 < 0 \). Therefore,
\[
\mathcal{T}_{S^n_+}^\prime(X) = \{ H \in S^n : \tilde{V}_1^1(H, W) \preceq 0 \},
\]
or equivalently,
\[
(Q^1_{\beta_1})^T P_{\alpha_1}^T W P_{\alpha_1} Q^1_{\beta_1} \preceq 2(Q^1_{\beta_1})^T P_{\alpha_1}^T H X^1 H P_{\alpha_1} Q^1_{\beta_1}, \tag{4.2}
\]
which is nothing but the formula for the second-order tangent set of \( S^n \) in (5.165) of [3].
References


