Journal of Convex Analysis
Volume 33 (2026), No. 1&2, 267–291

Variational Analysis of Orthogonally Invariant Norm Cones of Symmetric Matrices

Yule Zhang
School of Science, Dalian Maritime University, Dalian, China
ylzhang@dlmu.edu.cn

Jihong Zhang
School of Science, Shenyang Ligong University, Shenyang, China
zjh7815040x@163.com

Liwei Zhang
(1) National Frontiers Science Center for Industrial Intelligence and Systems Optimization, Northeastern University, Shenyang, China;
(2) Key Laboratory of Data Analytics and Optimization for Smart Industry, Northeastern University, Ministry of Education, Shenyang, China
zhanglw@mail.neu.edu.cn

Received: May 11, 2024
Accepted: February 26, 2025

This paper is devoted to the study of variational analysis of the orthogonally invariant norm cone of symmetric matrices. For a general orthogonally invariant norm cone of symmetric matrices, formulas for the tangent cone, normal cone and second-order tangent set are established. The differentiability properties of the projection operator onto the orthogonally invariant norm cone are developed, including formulas for the directional derivative and the B-subdifferential. In particular, the directional derivative is characterized by the second-order derivative of the corresponding symmetric norm, which is convenient for computation.

Keywords: Orthogonally invariant norm cone, variational analysis, tangent cone, normal cone, outer second-order tangent set, projection, directional derivative, B-subdifferential.

2000 Mathematics Subject Classification: 90C30.

1. Introduction

We consider the space $\mathbb{S}^m$ of real $m \times m$ symmetric matrices endowed with the standard inner product $\langle \cdot, \cdot \rangle$ given by $\langle A, B \rangle = \mathrm{Tr}(A^T B)$ for $A, B \in \mathbb{S}^m$. The induced norm is the Frobenius norm, denoted by $\| \cdot \|_F$: $\|A\|_F = \sqrt{\langle A, A \rangle}$ for $A \in \mathbb{S}^m$. We say that a norm $\| \cdot \|$ in $\mathbb{S}^m$ is an orthogonally invariant norm if

$$\|P^T AP\| = \|A\|, \forall A \in \mathbb{S}^m, P \in \mathcal{O}^m,$$

where $\mathcal{O}^m$ denotes the set of all $m \times m$ orthogonal matrices.

Supported by the Major Program of National Natural Science Foundation of China (Nos. 72192830 and 72192831), National Natural Science Foundation of China (Nos. 12201097 and 12371298) and the 111 Project (B16009).

ISSN 0944-6532 / $ 2.50

Many nonlinear optimization problems can be formulated as the following form

$$\min f(x), \text{ s.t. } g(x) \in K, \ x \in E, \tag{1.1}$$

where g is a twice differentiable mapping not identically zero, $K = \operatorname{epi} \|\cdot\|$ is the epigraph of a norm $\|\cdot\|$ on a space of matrices, $E \subset \mathcal{X}$ is a closed set, and $\mathcal{X}$ is a finite dimensional Hilbert space. For example, some optimization problems considered in [4, 5, 17] and [6] can be reformulated or can be approximated by the mathematical model (1.1), where $\|\cdot\|$ is the nuclear norm or the Ky-Fan k-matrix norm.

In order to develop optimality conditions for Problem (1.1), we have to derive the variational geometry of the feasible region, see Chapter 3 of [2], [11] and [18]. This issue is closely related to the variational geometry of cone K.

The study about the variational analysis for the Ky Fan k-norm matrix cone, including spectral norm and nuclear norm matrix cones, is relatively mature, see Ding [7] and Ding et al. [8, 9, 10]. How to extend variational properties of the Ky Fan k-norm matrix cone to a general norm matrix cone? This motivates us to establish the variational analysis for the orthogonally invariant norm cone of symmetric matrices.

The main contributions can be summarized as follows.

- For a general orthogonally invariant norm cone of symmetric matrices, formulas for the tangent cone, normal cone and second-order tangent set are established. These formulas are significant in establishing the first-order necessary optimality conditions as well as the second-order sufficient optimality conditions.
- The differentiability properties of the projection operator onto the orthogonally invariant norm cone are developed, including formulas for the directional derivative and the B-subdifferential.

It should be pointed out that the above results are derived from variational geometry of the lower level set of a convex function by [2] and the technique, developed by [1] for projection operator on the epigraph of a convex function, the same methodology was adopted in [16] for developing variational geometry of norm cones in finite dimensional Euclidean space $\mathbb{R}^n$.

The paper is organized as follows. In Section 2, we give some preliminaries needed in the following sections. In Section 3, we develop variational analysis of the orthogonally invariant norm cone of symmetric matrices and directional derivative and B-subdifferential of the projection onto the orthogonally invariant norm cone. Section 4 focuses on simplifying the formulas for directional derivative and B-subdifferential of the projection operator on the orthogonally invariant norm cone of symmetric matrices. We conclude the paper in Section 5 with a number of conclusions.

2. Preliminaries

It follows from [19, Page 35] that all orthogonally invariant norms of symmetric matrices can be written as

$$\|A\| = \psi(\lambda(A)),$$

where $\psi : \mathbb{R}^m \to \mathbb{R}$ is a symmetric norm, i.e., such that $\psi(Py) = \psi(y)$ for all $y \in \mathbb{R}^m$ and all m-by-m permutation matrix P.

We use ψ_* to denote the dual norm of ψ, then $\psi_* : \mathbb{R}^m \to \mathbb{R}$ can be expressed as

$$\psi_*(u) = \max_{\psi(y)\le 1} \langle u, y\rangle.$$

Let $w_1, \dots, w_r$ be r distinct values of m eigenvalues of A, namely

$$\begin{aligned} &\lambda_1(A) = \cdots = \lambda_{k_1}(A) = w_1 \\ &> \lambda_{k_1+1}(A) = \ldots = \lambda_{k_2}(A) = w_2 \\ &> \cdots > \lambda_{k_{r-1}+1} = \cdots = \lambda_{k_r} = w_r, \end{aligned}$$

where $k_0 = 0, k_r = m$. Denote

$$\alpha_1 = \{1, \dots, k_1\},\ \alpha_2 = \{k_1+1, \dots, k_2\}, \dots,\ \alpha_r = \{k_{r-1}+1, \dots, k_r\}.$$

The following three results are from [15], which are about properties of a symmetric function.

Lemma 2.1. ([15, Lemma 2.1]) *Let $\psi : \mathbb{R}^m \to \mathbb{R}$ be a symmetric function, twice differentiable at the point $w \in \mathbb{R}^m$, and P be a permutation matrix such that $Pw = w$. Then* (i) $\nabla\psi(w) = P^T \nabla f(w)$ *and* (ii) $\nabla^2\psi(w) = P^T\nabla^2\psi(w)P$.
In particular we have the representation

$$\nabla^2\psi(w) = \begin{pmatrix} a_{11}\boldsymbol{1}_{|\alpha_1|}\boldsymbol{1}_{|\alpha_1|}^T + b_{k_1}I_{|\alpha_1|} & a_{12}\boldsymbol{1}_{|\alpha_1|}\boldsymbol{1}_{|\alpha_2|}^T & \cdots & a_{1r}\boldsymbol{1}_{|\alpha_1|}\boldsymbol{1}_{|\alpha_r|}^T \\ a_{21}\boldsymbol{1}_{|\alpha_2|}\boldsymbol{1}_{|\alpha_1|}^T & a_{22}\boldsymbol{1}_{|\alpha_2|}\boldsymbol{1}_{|\alpha_2|}^T + b_{k_2}I_{|\alpha_2|} & \cdots & a_{2r}\boldsymbol{1}_{|\alpha_2|}\boldsymbol{1}_{|\alpha_r|}^T \\ \vdots & \vdots & \vdots & \vdots \\ a_{r1}\boldsymbol{1}_{|\alpha_r|}\boldsymbol{1}_{|\alpha_1|}^T & a_{r2}\boldsymbol{1}_{|\alpha_r|}\boldsymbol{1}_{|\alpha_2|}^T & \cdots & a_{rr}\boldsymbol{1}_{|\alpha_r|}\boldsymbol{1}_{|\alpha_r|}^T + b_{k_r}I_{|\alpha_r|} \end{pmatrix},$$

where $(a_{ij})_{i,j=1}^r$ is a real symmetric matrix, $b := (b_1, \dots, b_m)^T$ is a vector which is block refined by w in the following sense

$$w_i = w_j \ \ \textit{implies}\ \ b_i = b_j \textit{ for all } i, j.$$

Let $b(w) \in \mathbb{R}^m$ with $b(w) = (b_1(w), \dots, b_m(w))^T$, be defined by

$$b_i(w) = \begin{cases} \psi''_{ii}(w) & \text{if } |\alpha_l| = 1, i \in \alpha_l, \\ \psi''_{pp}(w) - \psi''_{pq}(w) & \text{for any } p \ne q \in \alpha_l, i \in \alpha_l. \end{cases}$$

Define $\mathcal{A}(w) = [\mathcal{A}_{ij}(w)]$ by

$$\mathcal{A}_{ij}(w) = \begin{cases} 0 & \text{if } i = j, \\ b_i(w) & \text{if } i \ne j \text{ but } i, j \in \alpha_l, \\ \dfrac{\psi'_i(w) - \psi'_j(w)}{w_i - w_j} & \text{otherwise.} \end{cases} \tag{2.1}$$

Lemma 2.2. ([15, Lemma 3.1]) *Let $\psi : \mathbb{R}^m \to \mathbb{R}$ be a symmetric function and $A \in \mathbb{S}^m$. Then ψ is differentiable at the point $\lambda(A)$ if and only if $\psi\circ\lambda$ is differentiable at the point A. In that case we have the formula*

$$\nabla(\psi \circ \lambda)(A) = P(\mathrm{Diag}\,\nabla\psi(\lambda(A)))P^T$$

for an orthogonal matrix P satisfying $A = P(\mathrm{Diag}\lambda(A))P^T$.

Lemma 2.3. ([15, Theorem 3.3]) *Let $\psi : \mathbb{R}^m \to \mathbb{R}$ be a symmetric function and $A \in \mathbb{S}^m$. Then ψ is twice differentiable at the point $\lambda(A)$ if and only if $\psi \circ \lambda$ is twice differentiable at the point A. Moreover, in this case the Hessian of $\psi \circ \lambda$ at the matrix A is*

$$\nabla^2(\psi \circ \lambda)(A)[H] = P\left[\mathrm{Diag}\,(\nabla^2\psi(\lambda(A))\mathrm{diag}\,\widehat{H}) + \mathcal{A} \circ \widehat{H}\right] P^T$$

for orthogonal matrix P satisfying $A = P(\mathrm{Diag}\lambda(A))P^T$, where $\widehat{H} = P^T HP$ and $\mathcal{A} = \mathcal{A}(\lambda(A))$ is defined by (2.1)*, where*

$$\mathcal{A} \circ \widehat{H} = [\mathcal{A}_{ij}\widehat{H}_{ij}].$$

Fan's inequality is an important tool in the space of symmetric matrices, which will be used in the following discussions.

Lemma 2.4. (Fan's inequality [13]) *Let $Y, Z \in \mathbb{S}^m$. Then*

$$\langle Y, Z\rangle \leq \lambda(Y)^T\lambda(Z),$$

where the equality holds if and only if Y and Z admit a simultaneous ordered eigenvalue decomposition, i.e., there exists an orthogonal matrix $P \in \mathcal{O}^m$ such that

$$Y = P\mathrm{Diag}(\lambda(Y))P^T \quad \textit{and} \quad Z = P\mathrm{Diag}(\lambda(Z))P^T.$$

It follows from Lemma 2.4 that, for any $X, Y \in \mathbb{S}^m$,

$$\|\lambda(X) - \lambda(Y)\| \leq \|X - Y\|. \tag{2.2}$$

Let $\psi : \mathbb{R}^m \to \mathbb{R}$ be a symmetric norm. An orthogonally invariant norm of real symmetric matrices, denoted by N, is defined by

$$N : \mathbb{S}^m \to \mathbb{R}_+,\ N(A) = (\psi \circ \lambda)(A),\ A \in \mathbb{S}^m. \tag{2.3}$$

It is easy to obtain the following result from Theorem 5.2.2 of [3].

Proposition 2.5. *The dual norm of N, denoted by N_*, is expressed as*

$$N_*(A) = (\psi_* \circ \lambda)(A), \quad A \in \mathbb{S}^m.$$

The matrix norm cone in $\mathbb{S}^m$ associated with N, denoted by $\mathcal{K}$, is the epigraph of N, namely

$$\mathcal{K} = \{(A, s) \in \mathbb{S}^m \times \mathbb{R} : N(A) \leq s\}. \tag{2.4}$$

We use $\mathcal{K}_*$ to denote the matrix norm cone associated with N_*, namely the epigraph of N_*:

$$\mathcal{K}_* = \{(A, s) \in \mathbb{S}^m \times \mathbb{R} : N_*(A) \leq s\}. \tag{2.5}$$

The *polar* of $\mathcal{K}$, denoted by $\mathcal{K}^\circ$, is expressed as

$$\mathcal{K}^\circ = -\mathcal{K}_* = \{(A,s) \in \mathbb{S}^m \times \mathbb{R} : (\psi_* \circ \lambda)(A) + s \le 0\}. \tag{2.6}$$

In the next section, we will use Sherman-Morrison-Woodbury formula for linear operators in [12].

Lemma 2.6. ([12, Theorem 2.1]) *Let X and Y be vector spaces. Let $A \in \mathcal{B}(X)$, $U \in \mathcal{B}(Y,X)$ and $V \in \mathcal{B}(X,Y)$ such that A is invertible, where $\mathcal{B}(X)$ denotes the linear space of linear operators from X to X, $\mathcal{B}(X,Y)$ denotes the linear space of linear operators from X to Y. Then $A+UV$ is invertible if and only if $I_Y + VA^{-1}U$ is invertible. Furthermore, if $A+UV$ is invertible, then*

$$(A+UV)^{-1} = A^{-1} - A^{-1}U[I_Y + VA^{-1}U]^{-1}VA^{-1}.$$

3. Variational analysis of norm cone $\mathcal{K}$

Throughout the whole paper, we assume that $\psi : \mathbb{R}^m \to \mathbb{R}$ is a symmetric norm. In order to study the variational geometry of the norm cone $\mathcal{K}$, we introduce a Hilbert space and related definitions as follows. Let $\mathcal{H} = \mathbb{S}^m \times \mathbb{R}$, define the inner product $\langle \cdot, \cdot \rangle$ in $\mathcal{H}$:

$$\langle z, z' \rangle = \mathrm{Tr}(AA') + ss',\ z = (A,s) \in \mathcal{H}, z' = (A', s') \in \mathcal{H}.$$

The induced norm, denoted by $\|\cdot\|_{\mathcal{H}}$, is defined by

$$\|z\|_{\mathcal{H}} = \sqrt{\|A\|_F^2 + s^2}.$$

This section provides basic variational analysis on the convex cone $\mathcal{K}$ mostly through a study on the orthogonal projection onto $\mathcal{K}$:

$$\Pi_{\mathcal{K}}(z) := \operatorname{argmin}_{z' \in \mathcal{K}} \frac{1}{2}\|z' - z\|_{\mathcal{H}}^2.$$

Let $\mathcal{I}$ and $\mathcal{I}'$ be identity mapping in $\mathbb{S}^m$ and $\mathcal{H}$, respectively, namely

$$\mathcal{I}A = A,\ \mathcal{I}'z = z \text{ for } A \in \mathbb{S}^m \text{ and } z \in \mathcal{H}.$$

We usually denote $\mathcal{I}'$ by

$$\mathcal{I}' = \begin{bmatrix} \mathcal{I} & 0 \\ 0 & 1 \end{bmatrix}.$$

For a matrix $A \in \mathbb{S}^m$, we use A^* to denote the following linear operator

$$A^* : \mathbb{S}^m \to \mathbb{R},\ A^*H = \mathrm{Tr}(AH), \forall H \in \mathbb{S}^m.$$

3.1. The tangent cone and second-order tangent set of $\mathcal{K}$

In this subsection, we will develop the tangent cone, the normal cone and the outer second-order tangent set of a norm cone.

Lemma 3.1. *Let $\psi : \mathbb{R}^m \to \mathbb{R}$ be a symmetric differentiable norm. Let A have the spectral decomposition $A = P(\mathrm{Diag}\lambda(A))P^T$ for an orthogonal matrix P. Then the following results hold:*

(i) *For $(A,s) \in \mathcal{K}$, $\mathcal{T}_{\mathcal{K}}(A,s) =$*

$$= \begin{cases} \mathcal{H} & N(A) < s, \\ \mathcal{K} & (A,s) = (0,0), \\ \left\{(d_A, d_s) \in \mathbb{S}^m \times \mathbb{R} : \langle P\mathrm{Diag}\,(\nabla\psi(\lambda(A)))P^T, d_A\rangle \le d_s\right\} & N(A) = s > 0. \end{cases}$$

(ii) *For $(A,s) \in \mathcal{K}$, $\mathcal{N}_{\mathcal{K}}(A,s) =$*

$$= \begin{cases} \{0\} & N(A) < s, \\ \mathcal{K}^\circ & (A,s) = (0,0), \\ \left\{\alpha(Y,-1) \in \mathcal{H} : (\psi_* \circ \lambda)(Y) = 1, \langle Y, A/s\rangle = 1, \alpha \ge 0\right\} & N(A) = s > 0. \end{cases}$$

(iii) *Let $z \in \mathcal{K}$ and $d \in \mathcal{T}_{\mathcal{K}}(z)$ where $z = (A,s)$ and $d = (d_A, d_s)$. If ψ is a twice differentiable symmetric norm, then*

$$\mathcal{T}^2_{\mathcal{K}}(z,d) = \begin{cases} \mathcal{H} & d \in \mathit{int}\, \mathcal{T}_{\mathcal{K}}(z), \\ \mathcal{T}_{\mathcal{K}}(z) & z = 0, \\ \left\{ (\xi_A, \xi_s) : \begin{array}{c} \langle P\mathrm{Diag}\,(\nabla\psi(\lambda(A)))P^T, \xi_A\rangle - \xi_s \\ +\nabla^2\psi(\lambda(A))[\mathrm{diag}(d_A), \mathrm{diag}(d_A)] \\ +\langle \mathcal{A}, d_A \circ d_A\rangle \le 0 \end{array} \right\} & \mathit{otherwise}. \end{cases} \tag{3.1}$$

Proof. In view of Lemma 2.3, Assertions (i) and (iii) come from Proposition 2.61 and Proposition 3.30 of [2], respectively. We only need to prove the case when $(\psi \circ \lambda)(A) = s > 0$ in (ii). Since $\mathcal{K}$ is a closed convex cone, we have from (2.110) of [2] that the normal cone of $\mathcal{K}$ at z is expressed as

$$\mathcal{N}_{\mathcal{K}}(z) = \{(U,t) \in \mathcal{H} : (U,t) \in \mathcal{K}^\circ, \langle (U,t), (A,s)\rangle = 0\}. \tag{3.2}$$

Since $\mathcal{K} = \mathrm{epi}\, N = \mathrm{epi}\,(\psi \circ \lambda)$, $(A,s) \in \mathrm{epi}\, N$, one has from $(U,t) \in \mathcal{N}_{\mathcal{K}}(z)$ that $t < 0$, it is necessary to consider an element $(V,-1) \in \mathcal{N}_{\mathcal{K}}(z)$. It follows from (3.2) that

$$\begin{aligned} (V,-1) \in \mathcal{N}_{\mathcal{K}}(z) &\iff (V,-1) \in \mathcal{K}^\circ, \langle (V,-1), (A,s)\rangle = 0 \\ &\iff N_*(V) \le 1, \langle V, A\rangle = s. \end{aligned} \tag{3.3}$$

From the definition of N_*, and $N(A/s) = 1$ and $\langle V, A/s\rangle \le N(A/s)\cdot N_*(V) = N_*(V)$, we have $\langle V, W/s\rangle \le 1$. Thus $N_*(V) \le 1, \langle V, A/s\rangle = 1$ is equivalent to $\langle V, W/s\rangle = 1$ and $N_*(V) = 1$. The formula of $\mathcal{N}_{\mathcal{K}}(z)$ for the case $N(A/s) = 1$ follows from (3.3). The proof is completed. □

Corollary 3.2. *Let $\psi : \mathbb{R}^m \to \mathbb{R}$ be a symmetric differentiable norm. Assume that $(A, s) \in \mathcal{K}$. If A has the spectral decomposition $A = P(\mathrm{Diag}\lambda(A))P^T$ for an orthogonal matrix P, then*

$$\mathcal{N}_{\mathcal{K}}(A, s) = \begin{cases} \{0\} & N(A) < s, \\ \mathcal{K}^\circ & (A, s) = (0, 0), \\ \big\{\alpha(U\mathrm{Diag}(a)U^T, -1) \in \mathcal{H} : a \in \partial\psi(\lambda(A)), \alpha \geq 0\big\} & N(A) = s > 0. \end{cases}$$

Proof. We only need to prove the case where (A, s) satisfying $N(A) = s > 0$. In this case, we have that $(\psi \circ \lambda)(A) = s$, $(\psi_* \circ \lambda)(Y) = 1$, $\langle Y, A/s\rangle = 1$, implying

$$\langle Y, A/s\rangle = N(A/s)N_*(Y).$$

From Lemma 2.4, we have that Y and A admit a simultaneous ordered eigenvalue decomposition, i.e., $Y = P\mathrm{Diag}(\lambda(Y))P^T$. It follows from Lemma 3.1 that

$$\begin{aligned} \mathcal{N}_{\mathcal{K}}(A, s) &= \big\{\alpha(Y, -1) \in \mathcal{H} : (\psi_* \circ \lambda)(Y) = 1, \langle Y, A/s\rangle = 1, \alpha \geq 0\big\} \\ &= \big\{\alpha(P\mathrm{Diag}(\lambda(Y))P^T, -1) : \psi_*(\lambda(Y)) = 1, \langle \lambda(Y), \lambda(A/s)\rangle = 1, \alpha \geq 0\big\} \\ &= \big\{\alpha(P\mathrm{Diag}(a)P^T, -1) : (a, -1) \in \mathcal{N}_{\mathrm{epi}\,\psi}(\lambda(A), s), \alpha \geq 0\big\} \\ &= \big\{\alpha(P\mathrm{Diag}(a)P^T, -1) : a \in \partial\psi(\lambda(A)), \alpha \geq 0\big\}. \end{aligned}$$

The proof is complete. □

3.2. The formula for the projection operator $\Pi_{\mathcal{K}}(z)$

In this subsection, we will derive a formula of $\Pi_{\mathcal{K}}(z)$. In the proof of the next proposition, we relate $\Pi_{\mathcal{K}}(z)$ to $\Pi_{\mathrm{epi}\psi}(\lambda(A), s)$, where $\Pi_{\mathrm{epi}\psi}(\lambda(A), s)$ is the solution of the following problem

$$\min \frac{1}{2}\|(w, t) - (\lambda(A), s)\|^2 \quad \text{s.t. } \psi(w) - t \leq 0. \tag{3.4}$$

Let $\mathcal{L} : \mathbb{R}^{m+1} \times \mathbb{R} \to \mathbb{R}$ be the Lagrangian of Problem (3.4):

$$\mathcal{L}(w, t, \mu; z) = \frac{1}{2}\|(w, t) - (\lambda(A), s)\|^2 + \mu(\psi(w) - t).$$

Then (w_*, s_*) is a solution to Problem (3.4) if and only if there exists a real number $\mu_* \geq 0$ such that (w_*, s_*, μ_*) satisfies

$$\begin{aligned} &\mu\nabla\psi(w) + w - \lambda(A) = 0, \\ &t - s - \mu = 0, \\ &\mu(\psi(w) - t) = 0, \mu \geq 0. \end{aligned} \tag{3.5}$$

If $\mu_* > 0$, because otherwise $\Pi_{\mathcal{K}}(A, s) = (A, s)$, then (w_*, s_*, μ_*) satisfies

$$\begin{aligned} &\mu\nabla\psi(w) + w - \lambda(A) = 0, \\ &t - s - \mu = 0, \\ &\psi(w) - s - \mu = 0, \mu > 0. \end{aligned} \tag{3.6}$$

From (3.6), if for a given $z = (A, s) \in \mathcal{H}$ and any $(w, \mu) \in \mathbb{R}^m \times \mathbb{R}$ with $w \neq 0$, we define $F : \mathbb{R}^{m+1} \to \mathbb{R}^{m+1}$ by

$$F(w, \mu; z) = \begin{bmatrix} \mu \nabla \psi(w) + w - \lambda(A) \\ \psi(w) - \mu - s \end{bmatrix}, \tag{3.7}$$

then $F(w^*, \mu^*; z^*) = 0$. In the next proposition, F defined by (3.7) is used to characterize $\Pi_{\mathrm{epi}\psi}(\lambda(A), s)$.

Proposition 3.3. *Let $\psi : \mathbb{R}^m \to \mathbb{R}_+$ be a differentiable symmetric norm. Let $z = (A, s) \in \mathcal{H}$ be a given point with $A = P\mathrm{Diag}(\lambda(A))P^T$ for an orthogonal matrix P, then*

$$\Pi_{\mathcal{K}}(A, s) = \begin{cases} (A, s) & \text{if } (A, s) \in \mathcal{K}, \\ (0, 0) & \text{if } (A, s) \in \mathcal{K}^\circ, \\ (P\mathrm{Diag}(w(z))P^T, \mu(z) + s) & \text{otherwise}, \end{cases} \tag{3.8}$$

where $(w(z), \mu(z)) \in \mathbb{R}^m \times \mathbb{R}_{++}$ is a solution of the system of equations:

$$F(w, \mu; z) = 0.$$

Proof. We may express $\Pi_{\mathcal{K}}(A, s)$ as the solution of the following optimization problem

$$\min \|X - A\|_F^2 + (t - s)^2 + \delta_{\{(\psi \circ \lambda)(X) \le t\}}(X, t), \tag{3.9}$$

where δ_C denotes the indicator function of a set C. Noting that

$$\delta_{\{(\psi \circ \lambda)(X) \le t\}}(X, t) = \delta_{\mathrm{epi}\,\psi}(\lambda(X), t)$$

and from the inequality (2.2) that

$$\|X - A\|_F \ge \|\lambda(X) - \lambda(A)\|,$$

we obtain

$$\|X - A\|_F^2 + (t - s)^2 + \delta_{\{(\psi \circ \lambda)(X) \le t\}}(X, t) \ge \|\lambda(X) - \lambda(A)\|^2 + (t - s)^2 + \delta_{\mathrm{epi}\,\psi}(\lambda(X), t).$$

This implies

$$\mathrm{dist}^2((A, s), \mathcal{K}) \ge \mathrm{dist}^2((\lambda(A), s), \mathrm{epi}\,\psi) = \|\Pi_{\mathrm{epi}\,\psi}((\lambda(A), s)) - ((\lambda(A), s))\|^2.$$

Let $$(\widehat{\lambda}, \widehat{s}) = \Pi_{\mathrm{epi}\psi}(\lambda(A), s), \;\; \widehat{X} = P\mathrm{Diag}(\widehat{\lambda})P^T.$$

Then $(\widehat{X}, \widehat{s}) \in \mathcal{K}$, and

$$\|\widehat{X} - A\|_F^2 + (\widehat{s} - s)^2 = \|\widehat{\lambda} - \lambda(A)\|_2^2 + (\widehat{s} - s)^2 = \mathrm{dist}^2((\lambda(A), s), \mathrm{epi}\psi).$$

Therefore, $(\widehat{X}, \widehat{s})$ is the unique solution to Problem (3.9).

From the above discussions, we get that $(\widehat{X}, \widehat{s}) = \Pi_{\mathcal{K}}(A, s)$ if and only if

$$\widehat{X} = P\mathrm{Diag}(\widehat{\lambda})P^T \;\text{ with }\; (\widehat{\lambda}, \widehat{s}) = \Pi_{\mathrm{epi}\psi}(\lambda(A), s).$$

Therefore, we only need to consider to calculate $\Pi_{\mathrm{epi}\psi}(\lambda(A), s)$, namely the solution of Problem (3.4).

It is obvious that $\Pi_{\mathcal{K}}(A,s) = (A,s)$ when $(A,s) \in \mathcal{K}$ and $\Pi_{\mathcal{K}}(A,s) = (0,0)$ when $(A,s) \in [\mathcal{K}]^{\circ}$. If $(A,s) \notin (\mathcal{K} \cup [\mathcal{K}]^{\circ})$, then for $(A_*, s_*) = \Pi_{\mathcal{K}}(A,s)$, one has that $A_* \neq 0$ and $s_* > 0$. Obviously, one has that $(A_*, s_*) \in \operatorname{bdry} \mathcal{K}$, namely $N(A_*) - s_* = 0$. From the above analysis, we know that $\Pi_{\mathcal{K}}(A,s) = (P\mathrm{Diag}(w^*)P^T, s^*)$ where (w_*, s_*) is the unique solution to the convex programming problem (3.4). The point (w_*, s_*) is a solution to Problem (3.4) if and only if there exists a real number $\mu_* \geq 0$ such that (w_*, s_*, μ_*) satisfies (3.5), namely

$$\begin{aligned}
&\mu \nabla \psi(w) + w - \lambda(A) = 0,\\
&t - s - \mu = 0,\\
&\mu(\psi(w) - t) = 0, \mu \geq 0.
\end{aligned}$$

From the above system, we know $\mu_* > 0$, because otherwise $\Pi_{\mathcal{K}}(A,s) = (A,s)$, which contradicts the assumption that $(A,s) \notin (\mathcal{K} \cup [\mathcal{K}]^{\circ})$. Therefore, (w_*, s_*, μ_*) satisfies (3.5), namely

$$\begin{aligned}
&\mu \nabla \psi(w) + w - \lambda(A) = 0,\\
&t - s - \mu = 0,\\
&\psi(w) - s - \mu = 0, \mu > 0.
\end{aligned}$$

Namely $(w_*, s_*) = (w(z), \mu(z) + s)$, where $(w(z), \mu(z)) \in \mathbb{R}^m \times \mathbb{R}_{++}$ is a solution of the system of equations $F(w, \mu; z) = 0$. The proof is complete. $\square$

We use $\mathcal{K}_2$ to denote the second-order cone in $\mathbb{S}^m \times \mathbb{R}$, namely

$$\mathcal{K}_2 = \{(A,s) \in \mathbb{S}^m \times \mathbb{R} : \|A\|_F \leq s\}. \tag{3.10}$$

Then we can easily obtain, from Proposition 3.3, the formula of the projection operator onto $\mathcal{K}_2$.

Corollary 3.4. *Let $z = (A,s) \in \mathcal{H}$ be a given point, then*

$$\Pi_{\mathcal{K}_2}(A,s) = \begin{cases} (A,s) & \text{if } (A,s) \in \mathcal{K}_2,\\ (0,0) & \text{if } (A,s) \in [\mathcal{K}_2]^{\circ},\\ (\|A\|_F + s) \cdot \frac{1}{2}\left(\frac{A}{\|A\|_F}, 1\right) & \text{otherwise.} \end{cases} \tag{3.11}$$

Remark 3.5. It should be pointed out that formula (3.11) is true not just for $\mathcal{K}_2$, but also for the cone

$$K = \{(\xi, t) \in E \times \mathbb{R} : \|\xi\| \leq t\}$$

where $(E, \|\cdot\|)$ is any Euclidean space. The fact that $\|\cdot\|_F = \|\cdot\|_2 \circ \lambda$ is not essential for obtaining (3.11).

3.3. Directional derivative and B-subdifferential of $\Pi_{\mathcal{K}}(z)$

Since $\Pi_{\mathcal{K}}(z)$ is globally Lipschitz continuous with constant 1, it is differentiable almost everywhere. From the expression of $\Pi_{\mathcal{K}}(z)$, we know that it is differentiable at points in the following set

$$\operatorname{int} \mathcal{K} \cup \operatorname{int} [\mathcal{K}]^{\circ} \cup \operatorname{int} [\mathcal{H} \setminus (\mathcal{K} \cup \mathcal{K}^{\circ})].$$

Now we give the formulas for the derivatives of $\Pi_{\mathcal{K}}$ at points in the above set.

Proposition 3.6. *Let $\psi : \mathbb{R}^m \to \mathbb{R}_+$ be a twice differentiable symmetric norm. Let $z = (A, s) \in \mathcal{H}$ be a given point with $A = P\mathrm{Diag}(\lambda(A))P^T$ for , then the following results hold:*

(i) *If $z = (A, s) \in \mathrm{int}\mathcal{K}$, then $\mathrm{D}\Pi_{\mathcal{K}}(z)(d_A, d_s) = (d_A, d_s)$ for $(d_A, d_s) \in \mathcal{H}$;*

(ii) *If $z = (A, s) \in \mathrm{int}\mathcal{K}^\circ$, then $\mathrm{D}\Pi_{\mathcal{K}}(z)(d_A, d_s) = 0$ for $(d_A, d_s) \in \mathcal{H}$;*

(iii) *If $z = (A, s) \in \mathrm{int}[\mathcal{H} \setminus (\mathcal{K} \cup [\mathcal{K}]^\circ)]$, then for $(d_A, d_s) \in \mathcal{H}$,*

$$\mathrm{D}\Pi_{\mathcal{K}}(z) \begin{bmatrix} d_A \\ d_s \end{bmatrix} = \begin{bmatrix} G(z)^{-1} & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} d_A \\ d_s \end{bmatrix} \tag{3.12}$$

$$- \beta(z) \begin{bmatrix} G(z)^{-1}\nabla(\psi \circ \lambda)(X(z)) \\ -1 \end{bmatrix} [\nabla(\psi \circ \lambda)(X(z))^* G(z)^{-1} \quad -1] \begin{bmatrix} d_A \\ d_s \end{bmatrix},$$

where $$\beta(z) = (1 + \nabla(\psi \circ \lambda)(X(z))^* G(z)^{-1} \nabla(\psi \circ \lambda)(X(z)))^{-1},$$
$$G(z) = \mathcal{I} + \mu(z)\nabla^2(\psi \circ \lambda)(X(z)),$$

where $X(z) = P\mathrm{Diag}(u(z))P^T$ and $(u(z), \mu(z)) \in \mathbb{R}^m \times \mathbb{R}_{++}$ is a solution of the system of equations $F(u, \mu; z) = 0$.

Proof. From the expression of $\Pi_{\mathcal{K}}(z)$, the results for (i) and (ii) are obvious, we only need to verify the result for the case (iii).

Let $(X(z), t(z)) = \Pi_{\mathcal{K}}(z)$, then $(X(z), t(z)) = (P\mathrm{Diag}(u(z))P^T, t(z))$ with

$$(u(z), t(z)) = \Pi_{\mathrm{epi}\,\psi}((\lambda(A), s))$$

and $((u(z), t(z)), \mu(z))$ is the Karush-Kuhn-Tucker (KKT) pair of the following problem

$$\min \frac{1}{2}\|(u, t) - (\lambda(A), s)\|^2 \quad \text{s.t. } \psi(u) - t \le 0 \tag{3.13}$$

if and only if $((X(z), t(z)), \mu(z))$ is the KKT pair of the following problem

$$\begin{aligned} \min \quad & \frac{1}{2}\|X - A\|_F^2 + \frac{(t - s)^2}{2} \\ \text{s.t.} \quad & (\psi \circ \lambda)(X) \le t. \end{aligned} \tag{3.14}$$

It follows from Proposition 3.3 that $(P\mathrm{Diag}(u(z))P^T, t(z), \mu(z))$ is the KKT pair of Problem (3.14) if and only if $X(z) = P\mathrm{Diag}(u(z))P^T$,$t(z) = s + \mu(z)$ and $(u(z), \mu(z))$ is the solution of the system $F(u, \mu; z) = 0$.

Let the Lagrangian of Problem (3.14) be $L(X, t, \mu)$,

$$L(X, t, \mu) = \frac{1}{2}\|X - A\|_F^2 + \frac{(t - s)^2}{2} + \mu((\psi \circ \lambda)(X) - t).$$

The optimality conditions of Problem (3.14) can be expressed as

$$\begin{aligned} &\mathrm{D}_X L(X, t, \mu) = 0, \\ &\partial_t L(X, t, \mu) = 0, \\ &0 \le \mu, (\psi \circ \lambda)(X) - t \le 0, \mu[(\psi \circ \lambda)(X) - t] = 0. \end{aligned}$$

Since $z = (A, s) \in \mathrm{int}[\mathcal{H} \setminus (\mathcal{K} \cup [\mathcal{K}]^\circ)]$, we can easily check $\mu(z) > 0$, this yields

$$\begin{aligned} &X - A + \mu \nabla(\psi \circ \lambda)(X) = 0, \\ &t - s - \mu = 0, \\ &(\psi \circ \lambda)(X) - t = 0. \end{aligned}$$

or
$$\mathcal{F}(X, \mu; z) = \begin{pmatrix} X - A + \mu \nabla(\psi \circ \lambda)(X) \\ (\psi \circ \lambda)(X) - s - \mu \end{pmatrix} = 0. \tag{3.15}$$

It is easy to check

$$D_{X,\mu}\mathcal{F}(X, \mu; z) = \begin{bmatrix} \mathcal{I} + \mu \nabla^2(\psi \circ \lambda)(X) & \nabla(\psi \circ \lambda)(X) \\ \nabla(\psi \circ \lambda)(X)^* & -1 \end{bmatrix}.$$

Since the Schur complement matrix

$$D_{X,\mu}\mathcal{F}(X, \mu; z)/-1 = \mathcal{I} + \mu \nabla^2(\psi \circ \lambda)(X) + \nabla(\psi \circ \lambda)(X)\nabla(\psi \circ \lambda)(X)^*$$

is self-adjoint and positively definite, the operator $D_{X,\mu}\mathcal{F}(X, \mu; z)$ is nonsingular. It follows from $z = (A, s) \in \mathrm{int}[\mathbb{R}^{m+1} \setminus (\mathcal{K} \cup [\mathcal{K}]^\circ)]$ and the classical implicit function theorem that there are exists $\delta > 0$ and $\varepsilon > 0$ and a mapping

$$(\tilde{X}(\cdot), \tilde{\mu}(\cdot)) : \mathbf{B}_\delta(z) \to \mathbf{B}_\varepsilon(X(z), \mu(z))$$

such that $\tilde{X}(z) = X(z)$ and $\tilde{\mu}(z) = \mu(z)$, and for any $z' = (A', s') \in \mathbf{B}_\delta(z)$, $(\tilde{X}(z'), \tilde{t}(z'), \tilde{\mu}(z'))$ satisfies $\mathcal{F}(\tilde{X}(z'), \tilde{\mu}(z'); z') = 0$ and $\tilde{\mu}(z') = \tilde{t}(z') - s'$ and they are continuously differentiable at $z' = z$. Taking derivative on the both sides of $\mathcal{F}(\tilde{X}(z'), \tilde{\mu}(z'); z') = 0$ with respect to z' at $z' = z$ along $dz = (d_A, d_s)$, we obtain for $\tilde{X}' := \mathrm{D}\tilde{X}(z)dz$ and $\tilde{\mu}' := \mathrm{D}\tilde{\mu}(z)dz$ that

$$\begin{aligned} &X' - d_A + \mu(z)\nabla^2(\psi \circ \lambda)(X(z))[X'] + \mu' \nabla(\psi \circ \lambda)(X(z)) = 0, \\ &\langle \nabla(\psi \circ \lambda)(X(z)), X' \rangle - d_s - \mu' = 0. \end{aligned}$$

The above equations can be written as

$$\begin{bmatrix} \mathcal{I} + \mu(z)\nabla^2(\psi \circ \lambda)(X(z)) & \nabla(\psi \circ \lambda)(X(z)) \\ \nabla(\psi \circ \lambda)(X(z))^* & -1 \end{bmatrix} \begin{bmatrix} X' \\ \mu' \end{bmatrix} = \begin{bmatrix} d_A \\ d_s \end{bmatrix},$$

or equivalently
$$D_{X,\mu}\mathcal{F}(X(z), \mu(z); z) \begin{bmatrix} X' \\ \mu' \end{bmatrix} = \begin{bmatrix} d_A \\ d_s \end{bmatrix}.$$

Since $G(z)$ is self-adjoint and positively definite and the Schur complement matrix

$$D_{X,\mu}\mathcal{F}(X(z), \mu(z); z)/-1 = G(z) + \nabla(\psi \circ \lambda)(X(z))\nabla(\psi \circ \lambda)(X(z))^*$$

is self-adjoint and positively definite, the operator $D_{X,\mu}\mathcal{F}(X(z), \mu(z); z)$ is nonsingular. Thus we obtain

$$\begin{bmatrix} X' \\ \mu' \end{bmatrix} = \begin{bmatrix} G(z) & \nabla(\psi \circ \lambda)(X(z)) \\ \nabla(\psi \circ \lambda)(X(z))^* & -1 \end{bmatrix}^{-1} \begin{bmatrix} d_A \\ d_s \end{bmatrix}.$$

From the Sherman-Morrison-Woodbury formula of linear operators between Banach spaces in [12], we have

$$
\begin{aligned}
&[G(z)+\nabla(\psi\circ\lambda)(X(z))\nabla(\psi\circ\lambda)(X(z))^*]^{-1}\\
&=G(z)^{-1}-\beta(z)G(z)^{-1}\nabla(\psi\circ\lambda)(X(z))\nabla(\psi\circ\lambda)(X(z))^*G(z)^{-1},
\end{aligned}
$$

where $\beta(z)=(1+\nabla(\psi\circ\lambda)(X(z))^*G(z)^{-1}\nabla(\psi\circ\lambda)(X(z)))^{-1}$. Therefore we obtain from the formula for inverse of block matrix in [14] that

$$
\begin{aligned}
&\begin{bmatrix} X' \\ \mu' \end{bmatrix}=\begin{bmatrix} G(z) & \nabla(\psi\circ\lambda)(X(z)) \\ \nabla(\psi\circ\lambda)(X(z))^* & -1 \end{bmatrix}^{-1}\begin{bmatrix} d_A \\ d_s \end{bmatrix}\\
&=\begin{bmatrix} [G(z)+\nabla(\psi\circ\lambda)(X(z))\nabla(\psi\circ\lambda)(X(z))^*]^{-1} & \beta(z)G(z)^{-1}\nabla(\psi\circ\lambda)(X(z)) \\ \beta(z)\nabla(\psi\circ\lambda)(X(z))^*G(z)^{-1} & -\beta(z) \end{bmatrix}\begin{bmatrix} d_A \\ d_s \end{bmatrix}.
\end{aligned}
$$

Noting that $t(z)=\mu(z)+s$, we obtain

$$
\begin{aligned}
&\mathrm{D}\Pi_{\mathcal{K}}(z)d_z=\begin{bmatrix} X' \\ \mu' \end{bmatrix}+\begin{bmatrix} 0 \\ d_s \end{bmatrix}\\
&=\begin{bmatrix} G(z)^{-1}d_A \\ d_s \end{bmatrix}-\beta(z)\begin{bmatrix} G(z)^{-1}\nabla(\psi\circ\lambda)(X(z)) \\ -1 \end{bmatrix}(\nabla(\psi\circ\lambda)(X(z))^*G(z)^{-1}d_A-d_s)\\
&=\begin{bmatrix} G(z)^{-1} & 0 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} d_A \\ d_s \end{bmatrix}\\
&\quad-\beta(z)\begin{bmatrix} G(z)^{-1}\nabla(\psi\circ\lambda)(X(z)) \\ -1 \end{bmatrix}[\nabla(\psi\circ\lambda)(X(z))^*G(z)^{-1}\quad -1]\begin{bmatrix} d_A \\ d_s \end{bmatrix},
\end{aligned}
$$

which yields (3.12). $\square$

For the second-order cone in $\mathbb{S}^m\times\mathbb{R}$, we have the following result about the derivative of the projection operator onto $\mathcal{K}_2$.

Corollary 3.7. *Let $z=(A,s)\in\mathcal{H}$ be a given point with $A=P\mathrm{Diag}(\lambda(A))P^T$ for an orthogonal matrix P. If $z=(A,s)\in\mathrm{int}[\mathcal{H}\setminus(\mathcal{K}_2\cup[\mathcal{K}_2]^\circ)]$, then for $(d_A,d_s)\in\mathcal{H}$,*

$$
\mathrm{D}\Pi_{\mathcal{K}_2}(z)\begin{bmatrix} d_A \\ d_s \end{bmatrix}=\frac{1}{2}\begin{bmatrix} \mathcal{I}+\frac{s}{\|A\|_F}\mathcal{I}-\frac{s}{\|A\|_F}\frac{AA^*}{\|A\|_F^2} & \frac{A}{\|A\|_F} \\ \frac{A^*}{\|A\|_F} & 1 \end{bmatrix}\begin{bmatrix} d_A \\ d_s \end{bmatrix}. \tag{3.16}
$$

Proof. It is easy to obtain

$$
X(z)=\frac{1}{2}\cdot(\|A\|_F+s)\cdot\frac{A}{\|A\|_F},\quad \mu(z)=\frac{1}{2}\cdot(\|A\|_F-s)\quad \text{for } z=(A,s).
$$

Noting for $X(z) \neq 0$, ψ is twice continuously differentiable with

$$\nabla[\psi \circ \lambda](X(z)) = \frac{X(z)}{\|X(z)\|_F} = \frac{A}{\|A\|_F}.$$

For $\mathcal{A}$ defined by (2.1), it is easy to verify

$$\mathcal{A} = \mathbf{1}_m \mathbf{1}_m^T - I_m. \tag{3.17}$$

And for $u(z)$, we have

$$u(z)^T \begin{bmatrix} p_1^T H p_1 \\ \vdots \\ p_m^T H p_m \end{bmatrix} = \sum_{i=1}^m u_i(z) p_i^T H p_i = \sum_{i=1}^m \langle u_i(z) p_i p_i^T, H \rangle = \langle X(z), H \rangle. \tag{3.18}$$

Combing (3.17) and (3.18), we have from Lemma 2.3 that

$$\begin{aligned}
&\nabla^2(\psi \circ \lambda)(X(z))[H] \\
&= P \left[\mathrm{Diag}\left(\frac{1}{\|u(z)\|_2} \left(I - \frac{u(z)u(z)^T}{\|u(z)\|_2^2} \right) \begin{bmatrix} p_1^T H p_1 \\ \vdots \\ p_m^T H p_m \end{bmatrix} \right) + \mathcal{A} \circ \widehat{H} \right] P^T \\
&= \frac{1}{\|u(z)\|_2} \left[\mathcal{I} - \frac{X(z)X(z)^*}{\|u(z)\|^2} \right] H.
\end{aligned}$$

We obtain

$$G(z) = \mathcal{I} + \mu(z) \nabla^2(\psi \circ \lambda)(X(z)) = \frac{2\|A\|_F}{\|A\|_F + s} \left(\mathcal{I} - \frac{\|A\|_F - s}{2\|A\|_F} \cdot \frac{AA^*}{\|A\|_F^2} \right).$$

By using Sherman-Morrison-Woodbury formula, we obtain

$$G(z)^{-1} = \frac{1}{2} \left(1 + \frac{s}{\|A\|_F} \right) \mathcal{I} + \frac{1}{2} \left(1 - \frac{s}{\|A\|_F} \right) \frac{AA^*}{\|A\|_F^2}.$$

Therefore, we obtain

$$G(z)^{-1} \nabla(\psi \circ \lambda)(X(z)) = G(z)^{-1} \frac{A}{\|A\|_F} = \frac{A}{\|A\|_F}$$

and $$\nabla(\psi \circ \lambda)(X(z))^* G(z)^{-1} \nabla(\psi \circ \lambda)(X(z)) = 1,$$

implying $$(1 + \nabla(\psi \circ \lambda)(X(z))^* G(z)^{-1} \nabla(\psi \circ \lambda)(X(z)))^{-1} = 1/2.$$

Thus we get

$$\begin{aligned}
&G(z)^{-1} - \beta(z) G(z)^{-1} \nabla(\psi \circ \lambda)(X(z)) \nabla(\psi \circ \lambda)(X(z))^* G(z)^{-1} \\
&= \frac{1}{2} \left[\mathcal{I} + \frac{s}{\|A\|_F} \mathcal{I} - \frac{s}{\|A\|_F} \frac{AA^*}{\|A\|_F^2} \right].
\end{aligned}$$

Therefore we obtain

$$\mathrm{D}\Pi_{\mathcal{K}}(z) = \frac{1}{2}\begin{bmatrix} \mathcal{I} + \frac{s}{\|A\|_F}\mathcal{I} - \frac{s}{\|A\|_F}\frac{AA^*}{\|A\|_F^2} & \frac{A}{\|A\|_F} \\ \frac{A^*}{\|A\|_F} & 1 \end{bmatrix}.$$

Namely (3.16) holds. The proof is complete. □

From Proposition 3.6, for $z = (A, s), d_z = (d_A, d_s)$, it is easy to obtain the directional derivative $\Pi'_{\mathcal{K}}(z; d_z)$ when $z \in \mathrm{int}\mathcal{K} \cup \mathrm{int}[\mathcal{K}]^\circ \cup \mathrm{int}[\mathbb{S}^m \times \mathbb{R} \setminus (\mathcal{K} \cup [\mathcal{K}]^\circ)]$ and it is trivial to have $\Pi'_{\mathcal{K}}(z; d_z) = \Pi_{\mathcal{K}}(\mathrm{d}_z)$ when $z = (A, s) = (0, 0)$. The following theorem considers the case where $z = (A, s) \in \mathrm{bdry}\mathcal{K} \setminus \{0\}$.

Theorem 3.8. *Let ψ be a twice differentiable symmetric norm. Let $z = (A, s) \in \mathrm{bdry}\mathcal{K} \setminus \{0\}$ with $A = P\mathrm{Diag}(\lambda(A))P^T$, then*

$$\Pi'_{\mathcal{K}}(z; d_z) = \begin{bmatrix} d_A \\ d_s \end{bmatrix} - \frac{[\nabla(\psi \circ \lambda)(A)^* d_A - d_s]_+}{1 + \|\nabla(\psi \circ \lambda)(A)\|_F^2} \begin{bmatrix} \nabla(\psi \circ \lambda)(A) \\ -1 \end{bmatrix}. \tag{3.19}$$

Proof. Obviously one has that $z = \Pi_{\mathcal{K}}(z)$. If $d_z \in \mathcal{R}_{\mathcal{K}}(z)$,[1] then

$$\nabla(\psi \circ \lambda)(A)^* d_A - d_s \leq 0, \quad \Pi'_{\mathcal{K}}(z; d_z) = d_z$$

and (3.19) is satisfied for this case. If $d_z \in \mathcal{T}_{\mathcal{K}}(z)$, then there exists a sequence $d^k \in \mathcal{R}_{\mathcal{K}}(z)$ such that $d^k \to d_z$. Noting $\Pi'_{\mathcal{K}}(z; d^k) = d^k$ and $d \to \Pi'_{\mathcal{K}}(z; d)$ is Lipschitz continuous, we obtain $\Pi'_{\mathcal{K}}(z; d_z) = d_z$. Noting that $\mathrm{d}_z \in \mathcal{T}_{\mathcal{K}}(z)$ is characterized by $\nabla(\psi \circ \lambda)(A)^* d_A - d_s \leq 0$, we obtain that (3.19) is satisfied for this case. For any $z' = (A', s')$, denote

$$(X(z'), t(z')) = \Pi_{\mathcal{K}}(z') \quad \text{and} \quad \mu(z') = t(z') - s',$$

then $t(z') = (\psi \circ \lambda)(X(z')) = N(X(z'))$ when $z' \notin \mathrm{int}\,\mathcal{K}$. For the case when $z = (A, s) \in \mathrm{bdry}\mathcal{K} \setminus \{0\}$, let $d_z \notin \mathcal{T}_{\mathcal{K}}(z)$, then $\nabla(\psi \circ \lambda)(A)^* d_A - d_s \geq 0$ and

$$z + \gamma d_z \in \mathrm{int}[\mathbb{S}^m \times \mathbb{R} \setminus (\mathcal{K} \cup [\mathcal{K}]^\circ)]$$

for small $\gamma > 0$. From the definition of $(X(\cdot), \mu(\cdot))$, we have

$$\begin{aligned} &\mu(z + \gamma d_z)\nabla(\psi \circ \lambda)(X(z + \gamma d_z)) + X(z + \gamma d_z) - (A + \gamma d_A) = 0 \\ &(\psi \circ \lambda)(X(z + \gamma d_z)) - \mu(z + \gamma d_z) - (s + \gamma d_s) = 0. \end{aligned}$$

Thus we obtain

$$\begin{aligned} [(\psi \circ \lambda)(X(z + \gamma d_z)) - (s + \gamma d_s)]\nabla(\psi \circ \lambda)(X(z + \gamma d_z)) \\ + X(z + \gamma d_z) - (A + \gamma d_A) = 0. \end{aligned}$$

[1] $\mathcal{R}_C(a)$ denotes the radial cone of a nonempty convex set C at $a \in C$, which defined by

$$\mathcal{R}_C(a) = \bigcup_{\lambda \geq 0, a' \in C} \{\lambda(a' - a)\}.$$

Noting $A = X(z), s = (\psi \circ \lambda)(A)$, we obtain from the above relation that

$$\begin{aligned}[](\psi \circ \lambda)(X(z+\gamma d_z)) - (\psi \circ \lambda)(X(z)) - \gamma d_s]\nabla(\psi \circ \lambda)(X(z+\gamma d_z)) \\ +X(z+\gamma d_z) - (A+\gamma d_A) = 0. \end{aligned} \tag{3.20}$$

From the mean-value theorem, there exists $\beta_\gamma \in (0,1)$ such that

$$\begin{aligned} &(\psi \circ \lambda)(X(z+\gamma d_z)) - (\psi \circ \lambda)(X(z)) \\ &= \nabla(\psi \circ \lambda)((1-\beta_\gamma)X(z) + \beta_\gamma X(z+\gamma d_z))^*(X(z+\gamma d_z) - X(z)). \end{aligned} \tag{3.21}$$

Obviously we have $\beta_\gamma \to 0$ when $\gamma \searrow 0$. Combining (3.20) and (3.21), we obtain

$$\begin{aligned} &[\mathcal{I}+\nabla(\psi\circ\lambda)(X(z+\gamma d_z))\nabla(\psi\circ\lambda)((1-\beta_\gamma)X(z)+\beta_\gamma X(z+\gamma d_z))^*]\,\frac{X(z+\gamma d_z)-X(z)}{\gamma} \\ &= \nabla(\psi \circ \lambda)(X(z+\gamma d_z))d_s + d_A. \end{aligned}$$

When $\gamma > 0$ is small enough, then the operator

$$[\mathcal{I} + \nabla(\psi \circ \lambda)(X(z+\gamma d_z))\nabla(\psi \circ \lambda)((1-\beta_\gamma)X(z) + \beta_\gamma X(z+\gamma d_z))^*]$$

is nonsingular, and thus

$$\begin{aligned} \frac{X(z+\gamma d_z) - X(z)}{\gamma} = \; & [\mathcal{I} + \nabla(\psi \circ \lambda)(X(z+\gamma d_z))\nabla(\psi \circ \lambda)((1-\beta_\gamma)X(z)+ \\ & + \beta_\gamma X(z+\gamma d_z))^*]^{-1}\,(\nabla(\psi \circ \lambda)(X(z+\gamma d_z))d_s + d_A). \end{aligned}$$

Since the right-hand side of the above equation has the limit when $\gamma \searrow 0$, we have that $X(\cdot)$ is directionally differentiable at z along d_z, and

$$\begin{aligned} X'(z;d_z) &= [\mathcal{I} + \nabla(\psi \circ \lambda)(A)\nabla(\psi \circ \lambda)(A)^*]^{-1}(\nabla(\psi \circ \lambda)(A)d_s + d_A) \\ &= \left[\mathcal{I} - \frac{\nabla(\psi \circ \lambda)(A)\nabla(\psi \circ \lambda)(A)^*}{1 + \|(\psi \circ \lambda)(A)\|_F^2}\right](\nabla(\psi \circ \lambda)(A)d_s + d_A) \\ &= d_A + \frac{d_s - \nabla(\psi \circ \lambda)(A)^* d_A}{1 + \|\nabla(\psi \circ \lambda)(A)\|_F^2}\nabla(\psi \circ \lambda)(A). \end{aligned}$$

Noting $t(z) = (\psi \circ \lambda)(X(z))$, we obtain from the above expression of $X'(z,d_z)$ that

$$t'(z;d_z) = \nabla(\psi \circ \lambda)(A)^* X'(z;d_z) = d_s - \frac{d_s - \nabla(\psi \circ \lambda)(A)^* d_A}{1 + \|\nabla(\psi \circ \lambda)(A)\|_F^2}.$$

This implies the expression of (3.19). The proof is completed. □

Noting that $N = N_*$ when ψ is the l_2-norm of $\mathbb{R}^m$, N is a twice differentiable norm, and we have the following results about the second-order cone.

Corollary 3.9. *For the projection operator onto the second-order cone $\mathcal{K}_2$, we have*

(i) *If $z = (A,s) \in \mathrm{bdry}\mathcal{K}_2 \setminus \{0\}$, then*

$$\Pi'_{\mathcal{K}_2}(z;d_z) = \begin{bmatrix} d_A \\ d_s \end{bmatrix} - \frac{1}{2}\left[\frac{A^* d_A}{\|A\|_F} - d_s\right]_+ \cdot \begin{bmatrix} \frac{A}{\|A\|_F} \\ -1 \end{bmatrix}. \tag{3.22}$$

(ii) *If* $z = (A, s) \in \mathrm{bdry}[\mathcal{K}_2]^\circ \setminus \{0\}$, *then*

$$\Pi'_{\mathcal{K}_2}(z; d_z) = \frac{1}{2}\left[\frac{A^* d_A}{\|A\|_F} + d_s\right]_+ \cdot \begin{bmatrix} \frac{A}{\|A\|_F} \\ 1 \end{bmatrix}. \tag{3.23}$$

Proof. Noting $\nabla(\psi \circ \lambda)(A) = \frac{A}{\|A\|_F}$, $\|\nabla(\psi \circ \lambda)(A)\|_F = 1$, we obtain (i) and (ii) from Theorem 3.8. □

Let
$$\bar{G} = \begin{bmatrix} G(z) & 0 \\ 0 & 1 \end{bmatrix}, \quad v(z) = \begin{bmatrix} \nabla(\psi \circ \lambda)(X(z)) \\ -1 \end{bmatrix}.$$

Then when $z = (A, s) \in \mathrm{int}[\mathbb{R}^{m+1} \setminus (\mathcal{K} \cup [\mathcal{K}]^\circ)]$, we have that

$$\mathrm{D}\Pi_{\mathcal{K}}(z) = \bar{G}(z)^{-1} - \frac{\bar{G}(z)^{-1} v(z) v(z)^* \bar{G}(z)^{-1}}{v(z)^* \bar{G}_N(z)^{-1} v(z)}. \tag{3.24}$$

From Proposition 3.6, it is easy to obtain that $\partial_B \Pi_{\mathcal{K}}(z) = \{D\Pi_{\mathcal{K}}(z)\}$ whenever $z \in \mathrm{int}\mathcal{K} \cup \mathrm{int}[\mathcal{K}]^\circ \cup \mathrm{int}[\mathbb{S}^m \times \mathbb{R} \setminus (\mathcal{K} \cup [\mathcal{K}]^\circ)]$. The following theorem considers the case when $z \in \mathrm{bdry}\mathcal{K} \setminus \{0\} \cup \mathrm{bdry}[\mathcal{K}]^\circ \setminus \{0\}$.

Theorem 3.10. *Let* $\psi : \mathbb{R}^m \to \mathbb{R}_+$ *be a twice symmetric norm. Let* $z = (A, s) \in \mathcal{H}$ *be a given point with* $A = P\mathrm{Diag}(\lambda(A))P^T$, *then the following results hold:*

(i) *If* $z = (A, s) \in \mathrm{bdry}\mathcal{K} \setminus \{0\}$, *then*

$$\partial_B \Pi_{\mathcal{K}}(z) = \left\{\mathcal{I}', \begin{bmatrix} \mathcal{I} & 0 \\ 0 & 1 \end{bmatrix} - \begin{bmatrix} \nabla(\psi \circ \lambda)(A) \\ -1 \end{bmatrix} \frac{[\nabla(\psi \circ \lambda)(A)^* - 1]}{1 + \|\nabla(\psi \circ \lambda)(A)\|_F^2}\right\}. \tag{3.25}$$

(ii) *If* $z = (A, s) \in \mathrm{bdry}[\mathcal{K}]^\circ \setminus \{0\}$, *then*

$$\partial_B \Pi_{\mathcal{K}}(z) = \left\{0, \begin{bmatrix} \nabla(\psi_* \circ \lambda)(A) \\ 1 \end{bmatrix} \frac{[\nabla(\psi_* \circ \lambda)(A)^* \ 1]}{1 + \|\nabla(\psi_* \circ \lambda)(A)\|_F^2}\right\}. \tag{3.26}$$

Proof. When $z = (A, s) \in \mathrm{bdry}\mathcal{K} \setminus \{0\}$, for constructing the B-subdifferential of $\Pi_{\mathcal{K}}$ at z, we have two ways of z' approaches z, one is $z' \xrightarrow{\mathrm{int}\mathcal{K}} z$ and the other is $z' \xrightarrow{\mathcal{D}} z$, where $\mathcal{D} = \mathrm{int}[\mathbb{S}^m \times \mathbb{R} \setminus (\mathcal{K} \cup [\mathcal{K}]^\circ)]$. For $z' \xrightarrow{\mathcal{D}} z$, we obtain $\mathcal{I}' \in \partial_B \Pi_{\mathcal{K}}(z)$. For $z' \xrightarrow{\mathcal{D}} z$, one has

$$\lim_{z' \xrightarrow{\mathcal{D}} z} X(z') = A, \ \lim_{z' \xrightarrow{\mathcal{D}} z} \mu(z') = 0,$$

which implies $\lim_{z' \xrightarrow{\mathcal{D}} z} G(z') = \mathcal{I}$, $\lim_{z' \xrightarrow{\mathcal{D}} z} G(z)^{-1} \nabla(\psi \circ \lambda)(X(z)) = \nabla(\psi \circ \lambda)(A)$.

Thus we obtain $\lim_{z' \xrightarrow{\mathcal{D}} z} \beta(z') = (1 + \|\nabla(\psi \circ \lambda)(A)\|_F^2)^{-1}$ and

$$\lim_{z' \xrightarrow{\mathcal{D}} z} \mathrm{D}\Pi_{\mathcal{K}}(z') =$$

$$= \begin{bmatrix} \mathcal{I} & 0 \\ 0 & 1 \end{bmatrix} - (1 + \|\nabla(\psi \circ \lambda)(A)\|_F^2)^{-1} \begin{bmatrix} \nabla(\psi \circ \lambda)(A) \\ -1 \end{bmatrix} [\nabla(\psi \circ \lambda)(A)^* \ \ -1].$$

Combining the above two limits, we obtain (3.25).

When $z = (A, s) \in \mathrm{bdry}[\mathcal{K}]^\circ \setminus \{0\}$, for constructing the B-subdifferential of $\Pi_{\mathcal{K}}$ at z, we have two ways of z' approaches z, one is $z' \xrightarrow{[\mathcal{K}]^\circ} z$ and the other is $z' \xrightarrow{\mathcal{D}} z$, where $\mathcal{D} = \mathrm{int}[\mathbb{S}^m \times \mathbb{R} \setminus (\mathcal{K} \cup [\mathcal{K}]^\circ)]$. For $z' \xrightarrow{\mathcal{D}} z$, we obtain $0 \in \partial_B \Pi_{\mathcal{K}}(z)$. For $z' \in \mathcal{D}$, one has

$$\mathrm{D}\Pi_{\mathcal{K}}(z') = \mathcal{I}' - \mathrm{D}\Pi_{[\mathcal{K}]^\circ}(z').$$

Applying Proposition 3.6 to the dual norm, we have for $z' \in \mathcal{D}$, that

$$\begin{aligned}\mathrm{D}\Pi_{\mathcal{K}^\circ}(z') = &\begin{bmatrix} G_*(z')^{-1} & 0 \\ 0 & 1 \end{bmatrix} \qquad (3.27)\\ &- \beta_*(z') \begin{bmatrix} G_*(z')^{-1}\nabla(\psi_* \circ \lambda)(X(z')) \\ 1 \end{bmatrix} [\nabla(\psi_* \circ \lambda)(X(z'))^* G_*(z')^{-1} \;\; 1],\end{aligned}$$

where $\quad \beta_*(z') = (1 + \nabla(\psi_* \circ \lambda)(X(z'))^* G_*(z')^{-1} \nabla(\psi_* \circ \lambda)(X(z')))^{-1},$

$$G_*(z') = \mathcal{I} + \mu(z')\nabla^2(\psi_* \circ \lambda)(X(z')),$$

where $X(z') = P[\mathrm{Diag} u_*(z')]P^T$ with $(u_*(z'), \mu_*(z')) \in \mathbb{R}^m \times \mathbb{R}_{++}$ being a solution of the system of equations $F_*(u, \mu; z') = 0$. For $z' \xrightarrow{\mathcal{D}} z$, one has

$$\lim_{z' \xrightarrow{\mathcal{D}} z} u_*(z') = \lambda(A), \; \lim_{z' \xrightarrow{\mathcal{D}} z} \mu_*(z') = 0,$$

which implies

$$\lim_{z' \xrightarrow{\mathcal{D}} z} G_*(z') = \mathcal{I}, \; \lim_{z' \xrightarrow{\mathcal{D}} z} G_*(z')^{-1}\nabla(\psi_* \circ \lambda)(X(z')) = \nabla(\psi_* \circ \lambda)(A).$$

Thus we obtain $\lim_{z' \xrightarrow{\mathcal{D}} z} \beta_*(z') = (1 + \|\nabla(\psi_* \circ \lambda)(A)\|_F^2)^{-1}$ and

$$\begin{aligned}\lim_{z' \xrightarrow{\mathcal{D}} z} \mathrm{D}\Pi_{\mathcal{K}}(z') &= \mathcal{I}' - \lim_{z' \xrightarrow{\mathcal{D}} z} \mathrm{D}\Pi_{[\mathcal{K}]^\circ}(z') \\ &= (1 + \|\nabla(\psi_* \circ \lambda)(A)\|_F^2)^{-1} \begin{bmatrix} \nabla(\psi_* \circ \lambda)(A) \\ 1 \end{bmatrix} [\nabla(\psi_* \circ \lambda)(A)^* \;\; 1].\end{aligned}$$

Combining the above two limits, we obtain (3.26). □

For the second-order cone, we have the following results about the B-subdifferential of $\Pi_{\mathcal{K}_2}$.

Corollary 3.11. *For the second-order cone $\mathcal{K}_2$ in $\mathbb{S}^m \times \mathbb{R}$, we have*

(i) *If $z = (A, s) \in \mathrm{bdry}\mathcal{K}_2 \setminus \{0\}$, then*

$$\partial_B \Pi_{\mathcal{K}_2}(z) = \left\{ \mathcal{I}', \frac{1}{2} \begin{bmatrix} 2\mathcal{I} - \frac{AA^*}{\|A\|_F^2} & \frac{A}{\|A\|_F} \\ \frac{A^*}{\|A\|_F} & 1 \end{bmatrix} \right\}. \qquad (3.28)$$

(ii) *If* $z = (A, s) \in \mathrm{bdry}[\mathcal{K}_2]^\circ \setminus \{0\}$, *then*

$$\partial_B \Pi_{\mathcal{K}_2}(z) = \left\{ 0,\ \frac{1}{2} \begin{bmatrix} \dfrac{AA^*}{\|A\|_F^2} & \dfrac{A}{\|A\|_F} \\ \dfrac{A^*}{\|A\|_F} & 1 \end{bmatrix} \right\}. \tag{3.29}$$

(iii) *If* $z = (A, s) = (0, 0)$, *we have* $\partial_B \Pi_{\mathcal{K}_2}(0) =$

$$= \{0, \mathcal{I}'\} \cup \left\{ \frac{1}{2} \begin{bmatrix} 2a(\mathcal{I} - YY^*) + YY^* & Y \\ Y^* & 1 \end{bmatrix} : a \in [0,1], \|Y\|_F = 1 \right\}. \tag{3.30}$$

Proof. For (i), $z = (A, s) \in \mathrm{bdry}\mathcal{K}_2 \setminus \{0\}$, $A \neq 0$, $\nabla(\phi \circ \lambda)(A) = A/\|A\|_F$ and $\|\nabla(\phi \circ \lambda)(A)\|_F = 1$. The formula (3.28) comes from (3.25). For (ii), we have $z = (A, s) \in \mathrm{bdry}[\mathcal{K}_2]^\circ \setminus \{0\}$, noting that $\psi = \psi_*$ and in this case $\|\nabla(\phi_* \circ \lambda)(A)\|_2 = 1$, we obtain (3.29) from (3.26).

For $\mathcal{D} = \mathrm{int}[\mathbb{S}^m \times \mathbb{R} \setminus (\mathcal{K}_2 \cup [\mathcal{K}_2]^\circ)]$, if $z' = (w', s') \in \mathcal{D}$, then we have from Corollary 3.7(iii) that $\mathrm{D}\Pi_{\mathcal{K}_2}(z')$ is given by (3.16). Note that, when $z' \xrightarrow{\mathcal{D}} z$, the outer limit of

$$\left\{ \frac{s'}{\|A'\|_F}, \frac{A'}{\|A'\|_F} \right\}$$

coincides with $\{(a, Y) : a \in [0,1], \|Y\|_F = 1\}$, and we obtain the formula (3.30) from (3.16). The proof is complete. □

4. Calculating $\nabla^2(\psi \circ \lambda)$ and $G(z)^{-1}$

In this section, in order to simplify formulas for directional derivatives and B-subdifferentials of $\Pi_{\mathcal{K}}$ when ψ is twice continuously differentiable, we develop formulas for $\nabla^2(\psi \circ \lambda)(A)$ and $G(z)^{-1}H$, which are easy to calculate.

Proposition 4.1. *Let* A *have the spectral decomposition* $A = P\mathrm{Diag}\lambda(A)P^T$ *and assume that* $w_1, \dots, w_r$ *are* r *distinct values of* m *eigenvalues of* A *with*

$$w_1 = \lambda_1(A) = \cdots = \lambda_{k_1}(A), \quad w_2 = \lambda_{k_1+1}(A) = \cdots = \lambda_{k_2}(A),$$
$$w_r = \lambda_{k_{r-1}+1}(A) = \cdots = \lambda_m(A).$$

Define $\alpha_1 = \{1, \dots, k_1\}$, $\alpha_2 = \{k_1 + 1, \dots, k_2\}, \dots, \alpha_r = \{k_{r-1} + 1, \dots, k_r\}$.

Let $\psi : \mathbb{R}^m \to \mathbb{R}$ *be a symmetric function, twice differentiable at the point* $w \in \mathbb{R}^m$, *and* P *be a permutation matrix such that* $Pw = w$. *Then*

(1) *there are real numbers* $b_1, \dots, b_r$ *and a symmetric matrix* $(a_{ij})_{i,j=1}^r$ *such that* $\nabla^2\psi(w)$ *is equal to*

$$\begin{bmatrix} b_1 I_{|\alpha_1|} & & \\ & \ddots & \\ & & b_r I_{|\alpha_r|} \end{bmatrix} + \begin{bmatrix} a_{11}\boldsymbol{1}_{|\alpha_1|}\boldsymbol{1}_{|\alpha_1|}^T & \cdots & a_{1r}\boldsymbol{1}_{|\alpha_1|}\boldsymbol{1}_{|\alpha_r|}^T \\ \vdots & \vdots & \vdots \\ a_{r1}\boldsymbol{1}_{|\alpha_r|}\boldsymbol{1}_{|\alpha_1|}^T & \cdots & a_{rr}\boldsymbol{1}_{|\alpha_r|}\boldsymbol{1}_{|\alpha_r|}^T \end{bmatrix}. \tag{4.1}$$

(2) *there are real numbers* $c_1, \dots, c_r$ *such that*

$$\nabla\psi(w) = \begin{bmatrix} c_1 \mathbf{1}_{|\alpha_1|} \\ \vdots \\ c_r \mathbf{1}_{|\alpha_r|} \end{bmatrix},$$

and for any $H \in \mathbb{S}^m$, $\quad \nabla^2(\psi \circ \lambda)(A)[H] =$

$$= P \begin{bmatrix} \left(\sum_{j=1}^{r} a_{1j} \mathrm{Tr}\left(P_{\alpha_j}^T H P_{\alpha_j}\right)\right) I_{|\alpha_1|} & & \\ & \ddots & \\ & & \left(\sum_{j=1}^{r} a_{rj} \mathrm{Tr}\left(P_{\alpha_j}^T H P_{\alpha_j}\right)\right) I_{|\alpha_r|} \end{bmatrix} P^T$$

$$+P \begin{bmatrix} \delta_{11} P_{\alpha_1}^T H P_{\alpha_1} & \cdots & \delta_{1r} P_{\alpha_1}^T H P_{\alpha_r} \\ \vdots & \ddots & \vdots \\ \delta_{r1} P_{\alpha_r}^T H P_{\alpha_1} & \cdots & \delta_{rr} P_{\alpha_r}^T H P_{\alpha_r} \end{bmatrix} P^T \tag{4.2}$$

and $$\nabla^2(\psi \circ \lambda)(A)[H, H] = \tag{4.3}$$

$$= \sum_{i=1}^{r}\sum_{j=1}^{r} a_{ij} \mathrm{Tr}(P_{\alpha_i}^T H P_{\alpha_i}) \mathrm{Tr}(P_{\alpha_j}^T H P_{\alpha_j}) + \sum_{i=1}^{r}\sum_{j=1}^{r} \delta_{ij} \sum_{i' \in \alpha_i, j' \in \alpha_j} (p_{i'}^T H p_{j'})^2,$$

where $$\begin{cases} \delta_{ii} = b_i, i = 1, \dots, r, \\ \delta_{ij} = \dfrac{c_i - c_j}{w_i - w_j}, i, j = 1, \dots r, i \neq j. \end{cases}$$

Proof. The formula (4.1) is an equivalent version of the formula in Lemma 2.1(ii). In view of Lemma 2.1, under the assumptions here we can express $\mathcal{A}$ as

$$\mathcal{A} = - \begin{bmatrix} \delta_{11} I_{|\alpha_1|} & & \\ & \ddots & \\ & & \delta_{rr} I_{|\alpha_r|} \end{bmatrix} + \begin{bmatrix} \delta_{11} \mathbf{1}_{|\alpha_1|} \mathbf{1}_{|\alpha_1|}^T & \cdots & \delta_{1r} \mathbf{1}_{|\alpha_1|} \mathbf{1}_{|\alpha_r|}^T \\ \vdots & \vdots & \vdots \\ \delta_{r1} \mathbf{1}_{|\alpha_r|} \mathbf{1}_{|\alpha_1|}^T & \cdots & \delta_{rr} \mathbf{1}_{|\alpha_r|} \mathbf{1}_{|\alpha_r|}^T \end{bmatrix}. \tag{4.4}$$

For $\widehat{H} = P^T H P$, define

$$h_1 = \begin{pmatrix} p_1^T H p_1 \\ \vdots \\ p_{|\alpha_1|}^T H p_{|\alpha_1|} \end{pmatrix}, h_2 = \begin{pmatrix} p_{|\alpha_1|+1}^T H p_{|\alpha_1|+1} \\ \vdots \\ p_{|\alpha_2|}^T H p_{|\alpha_2|} \end{pmatrix}, \dots, h_r = \begin{pmatrix} p_{|\alpha_{r-1}|+1}^T H p_{|\alpha_{r-1}|+1} \\ \vdots \\ p_{|\alpha_r|}^T H p_{|\alpha_r|} \end{pmatrix}.$$

We have from (4.1) and (4.4) that

$$\nabla^2(\psi\circ\lambda)(A)[H] = P\left(\mathrm{Diag}(\nabla^2\psi(\lambda(A))\mathrm{diag}\,\widehat{H}) + \mathcal{A}\circ\widehat{H}\right)P^T$$

$$= P\begin{bmatrix} \left(\sum_{j=1}^{r} a_{1j}\mathrm{Tr}\left(P_{\alpha_j}^T H P_{\alpha_j}\right)\right) I_{|\alpha_1|} & & \\ & \ddots & \\ & & \left(\sum_{j=1}^{r} a_{rj}\mathrm{Tr}\left(P_{\alpha_j}^T H P_{\alpha_j}\right)\right) I_{|\alpha_r|} \end{bmatrix} P^T$$

$$+P\begin{bmatrix} \delta_{11}P_{\alpha_1}^T H P_{\alpha_1} & \cdots & \delta_{1r}P_{\alpha_1}^T H P_{\alpha_r} \\ \vdots & \ddots & \vdots \\ \delta_{r1}P_{\alpha_r}^T H P_{\alpha_1} & \cdots & \delta_{rr}P_{\alpha_r}^T H P_{\alpha_r} \end{bmatrix} P^T,$$

which verifies (4.6).

The second-order directional derivative of $(\psi\circ\lambda)$ at A along H is

$$\nabla^2(\psi\circ\lambda)(A)[H,H] = \nabla^2\psi(w)[\mathrm{diag}\widehat{H},\mathrm{diag}\widehat{H}] + \langle\mathcal{A},\widehat{H}\circ\widehat{H}\rangle$$

$$= \sum_{i=1}^{r}\sum_{j=1}^{r} a_{ij}\mathrm{Tr}(P_{\alpha_i}^T H P_{\alpha_i})\mathrm{Tr}(P_{\alpha_j}^T H P_{\alpha_j}) + \sum_{i=1}^{r}\sum_{j=1}^{r}\delta_{ij}\sum_{i'\in\alpha_i, j'\in\alpha_j}(p_{i'}^T H p_{j'})^2.$$

The proof is complete. □

Corollary 4.2. *Let $w_1(z) := u_{k_1(z)}(z),\ldots,w_{r(z)}(z) := u_{k_r(z)}(z)$ be $r(z)$ distinct values of m values of $u(z)$, and assume that $X(z)$ has the spectral decomposition $X(z) = P\mathrm{Diag}u(z)P^T$ with $u(z) = \lambda(X(z))$, namely*

$$u_1(z) = \cdots = u_{k_1(z)}(z) > u_{k_1(z)+1}(z) = \cdots = u_{k_2(z)}(z) > u_{k_2(z)+1}(z)\cdots u_{k_r(z)}(z),$$

where $k_0 = 0$, $k_r(z) = m$. Define $\alpha_1(z) = \{1,\ldots,k_1(z)\}$,

$$\alpha_2(z) = \{k_1(z)+1,\ldots,k_2(z)\},\ldots,\ \alpha_{r(z)}(z) = \{k_{r-1}(z)+1,\ldots,k_{r(z)}\}.$$

Let $\psi:\mathbb{R}^m\to\mathbb{R}$ be a symmetric function, twice differentiable at the point $w\in\mathbb{R}^m$, and P be a permutation matrix such that $Pw = w$. Then

(1) *there are real numbers $b_1(z),\ldots,b_{r(z)}(z)$ and a symmetric matrix $(a_{ij}(z))_{i,j=1}^{r(z)}$ such that*

$$\nabla^2\psi(u(z)) = \begin{bmatrix} b_1(z)I_{|\alpha_1(z)|} & & \\ & \ddots & \\ & & b_{r(z)}(z)I_{|\alpha_{r(z)}(z)|} \end{bmatrix}$$

$$+\begin{bmatrix} a_{11}(z)\boldsymbol{1}_{|\alpha_1(z)|}\boldsymbol{1}_{|\alpha_1(z)|}^T & \cdots & a_{1r(z)}(z)\boldsymbol{1}_{|\alpha_1(z)|}\boldsymbol{1}_{|\alpha_{r(z)}(z)|}^T \\ \vdots & \vdots & \vdots \\ a_{r(z)1}(z)\boldsymbol{1}_{|\alpha_{r(z)}(z)|}\boldsymbol{1}_{|\alpha_1(z)|}^T & \cdots & a_{r(z)r(z)}(z)\boldsymbol{1}_{|\alpha_{r(z)}(z)|}\boldsymbol{1}_{|\alpha_{r(z)}(z)|}^T \end{bmatrix}. \quad (4.5)$$

(2) *there are real numbers* $c_1(z), \ldots, c_{r(z)}(z)$ *such that*

$$\nabla\psi(u(z)) = \begin{bmatrix} c_1(z)\boldsymbol{1}_{|\alpha_1(z)|} \\ \vdots \\ c_r(z)\boldsymbol{1}_{|\alpha_{r(z)}(z)|} \end{bmatrix},$$

then for any $H \in \mathbb{S}^m$, *we have* $\nabla^2(\psi\circ\lambda)(X(z))[H] =$

$$\begin{aligned} &= P \begin{bmatrix} \Delta_1(z) I_{|\alpha_1(z)|} & & \\ & \ddots & \\ & & \Delta_r(z) I_{|\alpha_{r(z)}(z)|} \end{bmatrix} P^T \\ &+ P \begin{bmatrix} \delta_{11}(z) P^T_{\alpha_1(z)} H P_{\alpha_1(z)} & \cdots & \delta_{1r}(z) P^T_{\alpha_1(z)} H P_{\alpha_{r(z)}(z)} \\ \vdots & \ddots & \vdots \\ \delta_{r(z)1} P^T_{\alpha_{r(z)}(z)} H P_{\alpha_1(z)} & \cdots & \delta_{r(z)r(z)} P^T_{\alpha_{r(z)}(z)} H P_{\alpha_{r(z)}(z)} \end{bmatrix} P^T \end{aligned} \tag{4.6}$$

and $\nabla^2(\psi\circ\lambda)(A)[H,H] =$

$$\begin{aligned} &= \sum_{i=1}^{r(z)}\sum_{j=1}^{r(z)} a_{ij}(z)\mathrm{Tr}(P^T_{\alpha_i(z)} H P_{\alpha_i(z)})\mathrm{Tr}(P^T_{\alpha_j(z)} H P_{\alpha_j(z)}) \\ &\quad + \sum_{i=1}^{r(z)}\sum_{j=1}^{r(z)} \delta_{ij}(z) \sum_{i'\in\alpha_i(z), j'\in\alpha_j(z)} (p^T_{i'} H p_{j'})^2, \end{aligned} \tag{4.7}$$

where
$$\Delta_i(z) = \left(\sum_{j=1}^{r(z)} a_{ij}(z) \mathrm{Tr}\left(P^T_{\alpha_j(z)} H P_{\alpha_j(z)} \right) \right)$$

and
$$\begin{cases} \delta_{ii}(z) = b_i(z), i = 1, \ldots, r(z), \\ \delta_{ij}(z) = \dfrac{c_i(z) - c_j(z)}{w_i(z) - w_j(z)}, i, j = 1, \ldots r(z), i \neq j. \end{cases}$$

Proposition 4.3. *Under the setting of Proposition* 4.1, *for any* $H \in \mathbb{S}^m$, *one has* $Y = G(z)^{-1}H$ *can be characterized by*

$$P^T_{\alpha_i(z)} Y P_{\alpha_j(z)} = \frac{1}{1+\mu(z)\delta_{ij}(z)} P^T_{\alpha_i(z)} H P_{\alpha_j(z)}, \quad i, j = 1, \ldots, r(z),\ i \neq j, \tag{4.8}$$

for $k = 1, \ldots, r(z)$, $i, j \in \alpha_k$, $i \neq j$,

$$P_i^T Y P_j = \frac{1}{1+\mu(z)\delta_{kk}(z)} P_i^T H P_j, \tag{4.9}$$

and
$$\mathrm{diag}(P^T Y P) = [I_m + \mu(z)\nabla^2\psi(u(z))]^{-1}\mathrm{diag}(P^T H P). \tag{4.10}$$

Proof. For $Y = G(z)^{-1}H$, we have for

$$\Delta_i(z) = \left(\sum_{j=1}^{r(z)} a_{ij}(z) \mathrm{Tr}\left(P^T_{\alpha_j(z)} H P_{\alpha_j(z)} \right) \right)$$

that $H = Y + \mu(z)\nabla^2(\varphi_p \circ \lambda)(X(z))[Y] =$

$$= Y + \mu(z)P \begin{bmatrix} \Delta_1(z)I_{|\alpha_1(z)|} & & \\ & \ddots & \\ & & \Delta_r(z)I_{|\alpha_{r(z)}(z)|} \end{bmatrix} P^T$$

$$+\mu(z)P \begin{bmatrix} \delta_{11}(z)P^T_{\alpha_1(z)}YP_{\alpha_1(z)} & \cdots & \delta_{1r}(z)P^T_{\alpha_1(z)}YP_{\alpha_{r(z)}(z)} \\ \vdots & \ddots & \vdots \\ \delta_{r(z)1}P^T_{\alpha_{r(z)}(z)}YP_{\alpha_1(z)} & \cdots & \delta_{r(z)r(z)}P^T_{\alpha_{r(z)}(z)}YP_{\alpha_{r(z)}(z)} \end{bmatrix} P^T.$$

This implies, for $i, j = 1, \dots, r(z),\ i \neq j$,

$$P^T_{\alpha_i(z)}YP_{\alpha_j(z)} = \frac{1}{1+\mu(z)\delta_{ij}(z)}P^T_{\alpha_i(z)}HP_{\alpha_j(z)}.$$

And for $k = 1, \dots, r(z)$, we have $P^T_{\alpha_k(z)}HP_{\alpha_k(z)} =$

$$= (1+\mu(z)\delta_{kk}(z))P^T_{\alpha_k(z)}YP_{\alpha_k(z)} + I_{\alpha_k(z)}\mu(z)\Big(\textstyle\sum_{j=1}^{r(z)} a_{kj}(z)\mathrm{Tr}\left(P^T_{\alpha_j(z)}YP_{\alpha_j(z)}\right)\Big)$$
$$= (1+\mu(z)\delta_{kk}(z))P^T_{\alpha_k(z)}YP_{\alpha_k(z)} + \mu(z)\Big(\textstyle\sum_{j=1}^{r(z)} a_{kj}(z)I_{\alpha_k(z)}I^*_{\alpha_j(z)}\left(P^T_{\alpha_j(z)}YP_{\alpha_j(z)}\right)\Big)$$

which implies for $k = 1, \dots, r(z)$, $i, j \in \alpha_k(z)$, $i \neq j$,

$$P_i^TYP_j = \frac{1}{1+\mu(z)\delta_{kk}(z)}P_i^THP_j.$$

For $k = 1, \dots, r(z)$,

$$\begin{aligned} \mathrm{diag}[P^T_{\alpha_k(z)}HP_{\alpha_k(z)}] &= (1+\mu(z)\delta_{kk}(z))\mathrm{diag}[P^T_{\alpha_k(z)}YP_{\alpha_k(z)}] \\ &\quad + \mu(z)\mathbf{1}_{|\alpha_k(z)|}\Bigg(\sum_{j=1}^{r(z)} a_{ij}(z)\mathbf{1}^T_{|\alpha_j(z)|}\mathrm{diag}\left(P^T_{\alpha_j(z)}YP_{\alpha_j(z)}\right)\Bigg). \end{aligned} \tag{4.11}$$

Define
$$\begin{pmatrix} \widehat{y}_1 \\ \widehat{y}_2 \\ \vdots \\ \widehat{y}_{r(z)} \end{pmatrix} = \begin{pmatrix} \mathrm{diag}\left(P^T_{\alpha_1(z)}YP_{\alpha_1(z)}\right) \\ \mathrm{diag}\left(P^T_{\alpha_2(z)}YP_{\alpha_2(z)}\right) \\ \vdots \\ \mathrm{diag}\left(P^T_{\alpha_{r(z)}(z)}YP_{\alpha_{r(z)}(z)}\right) \end{pmatrix}$$

and
$$\begin{pmatrix} \widehat{h}_1 \\ \widehat{h}_2 \\ \vdots \\ \widehat{h}_{r(z)} \end{pmatrix} = \begin{pmatrix} \mathrm{diag}\left(P^T_{\alpha_1(z)}HP_{\alpha_1(z)}\right) \\ \mathrm{diag}\left(P^T_{\alpha_2(z)}HP_{\alpha_2(z)}\right) \\ \vdots \\ \mathrm{diag}\left(P^T_{\alpha_{r(z)}(z)}HP_{\alpha_{r(z)}(z)}\right) \end{pmatrix}.$$

Then the solution $(\widehat{y}_1, \ldots, \widehat{y}_r)$ of (4.11) is the solution to the following system of linear equations:

$$\begin{pmatrix} \widehat{h}_1 \\ \vdots \\ \widehat{h}_r \end{pmatrix} = M(z) \begin{pmatrix} \widehat{y}_1 \\ \vdots \\ \widehat{y}_{r(z)} \end{pmatrix}, \quad \text{where}$$

$$M(z) = \left\{ \begin{pmatrix} \widehat{\delta}_{11}(z) I_{|\alpha_1(z)|} & \cdots & 0 \\ \vdots & \vdots & \vdots \\ 0 & \cdots & \widehat{\delta}_{rr}(z)) I_{|\alpha_{r(z)}(z)|} \end{pmatrix} + \begin{pmatrix} \widehat{a}_{11}(z) \mathbf{1}_{|\alpha_1(z)|} \mathbf{1}^T_{|\alpha_1(z)|} & \cdots & \widehat{a}_{1r(z)}(z) \mathbf{1}_{|\alpha_1(z)|} \mathbf{1}^T_{|\alpha_{r(z)}(z)|} \\ \vdots & \vdots & \vdots \\ \widehat{a}_{r1}(z) \mathbf{1}_{|\alpha_{r(z)}(z)|} \mathbf{1}^T_{|\alpha_1(z)|} & \cdots & \widehat{a}_{rr}(z)) \mathbf{1}_{|\alpha_{r(z)}(z)|} \mathbf{1}^T_{|\alpha_{r(z)}(z)|} \end{pmatrix} \right\}$$

with $\widehat{\delta}_{ii}(z) = 1 + \mu(z)\delta_{ii}(z), i = 1, \ldots, r(z)$, $\widehat{a}_{ij}(z) = \mu(z) a_{ij}(z), i, j = 1, \ldots, r(z)$. It is not difficult to verify that

$$M(z) = I_m + \mu(z) \nabla^2 \psi(u(z)). \tag{4.12}$$

Therefore, the formulas (4.8), (4.9) and (4.10) hold. □

Define $\pi_{ij}(z) = (1 + \mu(z)\delta_{ij}(z))^{-1}, i, j = 1, \ldots, r(z)$ and

$$\mathcal{B}(z) = - \begin{bmatrix} \pi_{11}(z) I_{|\alpha_1(z)|} & & \\ & \ddots & \\ & & \pi_{r(z)r(z)}(z) I_{|\alpha_r(z)|} \end{bmatrix} \tag{4.13}$$

$$+ \begin{bmatrix} \pi_{11}(z) \mathbf{1}_{|\alpha_1(z)|} \mathbf{1}^T_{|\alpha_1(z)|} & \cdots & \pi_{1r(z)}(z) \mathbf{1}_{|\alpha_1(z)|} \mathbf{1}^T_{|\alpha_{r(z)}(z)|} \\ \vdots & \vdots & \vdots \\ \pi_{r(z)1}(z) \mathbf{1}_{|\alpha_{r(z)}(z)|} \mathbf{1}^T_{|\alpha_1(z)|} & \cdots & \pi_{r(z)r(z)}(z) \mathbf{1}_{|\alpha_{r(z)}(z)|} \mathbf{1}^T_{|\alpha_{r(z)}(z)|} \end{bmatrix}.$$

Then, for any $H \in \mathbb{S}^m$,

$$G^{-1}(z)H = P \left[\mathrm{Diag} \left(M(z)^{-1} \mathrm{diag}(P^T H P) \right) + \mathcal{B}(z) \circ P^T H P \right] P^T. \tag{4.14}$$

Proposition 4.4. *Let $\psi : \mathbb{R}^m \to \mathbb{R}_+$ be a twice differentiable symmetric gauge function. Let $z = (A, s) \in \mathcal{H}$ be a given point with $A = P\mathrm{Diag}(\lambda(A))P^T$. Let $X(z) = P\mathrm{Diag}(u(z))P^T$ and $(u(z), \mu(z)) \in \mathbb{R}^m \times \mathbb{R}_{++}$ be a solution of the system of equations $F(u, \mu; z) = 0$.*

Then for $\quad \beta(z) = (1 + \nabla(\psi \circ \lambda)(X(z))^* G(z)^{-1} \nabla(\psi \circ \lambda)(X(z)))^{-1},$

$$G(z) = \mathcal{I} + \mu(z) \nabla^2 (\psi \circ \lambda)(X(z)),$$

we have

$$G(z)^{-1}\nabla(\psi\circ\lambda)(X(z)) = P\mathrm{Diag}\left[(I_m+\mu(z)\nabla^2\psi(u(z)))^{-1}\nabla\psi(u(z))\right]P^T \quad (4.15)$$

and

$$\beta(z) = 1+\nabla\psi(u(z))^T(I_m+\mu(z)\nabla^2\psi(u(z)))^{-1}\nabla\psi(u(z)). \quad (4.16)$$

5. Conclusion

This paper developed variational analysis of orthogonally invariant norm cone of symmetric matrices including the computation of the tangent cone, the normal cone, and the second-order tangent set of an orthogonally invariant norm cone, as well as the formulas for directional derivative and B-subdifferential of the projection operator onto a norm cone. An important topic left is the variational analysis of an orthogonally invariant norm cone when the associated symmetric norm is nonsmooth. Another topic is the study on optimality and stability theories of Problem (1.1), including the first-order and second-order necessary optimality conditions, and the second-order sufficient optimality conditions, the strong regularity and isolated calmness of the KKT system of Problem (1.1). Therefore, there are many theoretical applications of the variational analysis of orthogonally invariant norm cones to the orthogonally invariant norm conic optimization.

Acknowledgments. The authors are so grateful to the reviewers for their valuable comments and suggestions, for example including Remark 3.5, which helped to improve the quality of this paper.

References

[1] A. Beck: *First-Order Methods in Optimization*, Society for Industrial and Applied Mathematics, Philadelphia (2017).

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

[3] J. M. Borwein, A. S. Lewis: *Convex Analysis and Nonlinear Optimization: Theory and Examples*, Springer, New York (2006).

[4] S. Boyd, P. Diaconis, L. Xiao: *Fastest mixing Markov chain on a graph*, SIAM Review 46 (2004) 667–689.

[5] E. J. Candes, B. Recht: *Exact matrix completion via convex optimization*, Found. Comput. Math. 9 (2009) 717–772.

[6] Y. Cui, C. Ding, X. Zhao: *Quadratic growth conditions for convex matrix optimization problems associated with spectral functions*, SIAM J. Optimization 27 (2017) 2332–2355.

[7] C. Ding: *Variational analysis of the Ky Fan k-norm*, Set-Valued Var. Analysis 25 (2017) 265–296.

[8] C. Ding, D. F. Sun, J. Sun, K. C. Toh: *Spectral operator of matrices*, Math. Programming, Ser. B 168/1-2 (2018) 509–531.

[9] C. Ding, D. F. Sun, J. Sun, K. C. Toh: *Spectral operators of matrices: semismoothness and characterizations of the generalized Jacobian*, SIAM J. Optimization 30 (2020) 630–659.

[10] C. Ding, D. F. Sun, K. C. Toh: *An introduction to a class of matrix cone programming*, Math. Programming, Ser. A 144/1-2 (2014) 141–179.

[11] C. Ding, D. F. Sun, L. W. Zhang: *Characterization of the robust isolated calmness for a class of conic programming problems*, SIAM J. Optimization 27 (2017) 67–90.

[12] J. Ding, A. H. Zhou: *A spectrum theorem for perturbed bounded linear operators*, Appl. Math. Comput. 201/1 (2008) 723–728.

[13] K. Fan: *On a theorem of Weyl concerning eigenvalues of linear transformations*, Proc. Nat. Acad. Sci. U.S.A. 35 (1949) 652–655.

[14] G. H. Golub, C. F. Van Loan: *Matrix Computations*, 3rd ed., The Johns Hopkins University Press, Baltimore (1996).

[15] A. S. Lewis, H. S. Sendov: *Twice differentiable spectral functions*, SIAM J. Matrix Anal. Appl. 23/2 (2001) 368–386.

[16] H. Y. Liu, J. Wu, L. W. Zhang: *Variational analysis of norm cones in finite dimensional Euclidean spaces*, Set-Valued Var. Analysis 31/3 (2023), art.no. 27, 28p.

[17] B. Recht, M. Fazel, P. A. Parrilo: *Guaranteed minimum-rank solutions of linear matrix equations via nuclear norm minimization*, SIAM Review 52/3 (2010) 471–501.

[18] D. F. Sun: *The strong second order sufficient condition and constraint nondegeneracy in nonlinear semidefinite programming and their implications*, Math. Oper. Res. 31 (2006) 761–776.

[19] G. A. Watson: *Characterization of the subdifferential of some matrix norms*, Linear Algebra Appl. 170 (1992) 33–45.