國立臺灣師範大學數學系碩士班碩士論文

指導教授：陳界山博士

To What Extent are Second-Order Cone and Positive Semidefinite Cone Alike?

研 究 生：廖淳格

中 華 民 國 九 十 八 年 六 月

誌　謝

首先感謝家人對我的照顧與支持，尤其是我的父母，沒有你們不會有今天的我。在師大的這些日子裡，最感激的是我的指導教授：陳界山老師。在他的帶領之下我認識了這個以前不曾接觸過的領域，並順利完成我的碩士論文。除此之外，我也在老師的身上看到對於做學問應該要有的態度，更讓我反省自己，決定向老師看齊。再來感謝我的同學：子靖、晉宇，有他們的陪伴我才能愉快的度過碩士班的每一天。還有口試委員朱亮儒老師、柯春旭老師，他們在口試時的意見讓我受益良多。要感謝的人太多了，謝謝在我人生中出現的每一個人。

廖淳格　謹誌

2009 年 6 月

目次

To what extent are second-order cone and positive semidefinite cone alike?

June 5, 2009

Abstract. The cone of positive semidefinite matrices ($\mathcal{S}^n_+$) and second-order cone ($\mathcal{K}^n$) are both self-dual and special cases of symmetric cones. Each of them play an important role in semidefinite programming (SDP) and second-order cone programming (SOCP), respectively. It is known that an SOCP problem can be viewed as an SDP problem via certain relation between $\mathcal{S}^n_+$ and $\mathcal{K}^n$. Nonetheless, most analysis used for dealing SDP can not carried over to SOCP due to some difference, for instance, the matrix multiplication is associative for $\mathcal{S}^n_+$ whereas the Jordan product is not for $\mathcal{K}^n$. In this paper, we try to have a thorough study on the similarity and difference between these two cones which provide theoretical for further investigation of SDP and SOCP.

Key words. Second-order cone, convex function, monotone function, positive semidefinite matrix, spectral decomposition.

1 Introduction

The second-order cone(SOC) in $\mathbb{R}^n$, also called Lorentz cone, is defined by

$$\mathcal{K}^n = \{x = (x_1, x_2) \in \mathbb{R} \times \mathbb{R}^{n-1} : \|x_2\| \leq x_1\}, \tag{1}$$

where $\|\cdot\|$ denotes the Euclidean norm. If $n = 1$, let $\mathcal{K}^n$ denote the set of nonnegative reals $\mathbb{R}_+$. For any $x, y \in \mathbb{R}^n$, we write $x \succeq_{\mathcal{K}^n} y$ if $x - y \in \mathcal{K}^n$; and write $x \succ_{\mathcal{K}^n} y$ if $x - y \in \text{int}(\mathcal{K}^n)$. In other words, we have $x \succeq_{\mathcal{K}^n} 0$ if and only if $x \in \mathcal{K}^n$ and $x \succ_{\mathcal{K}^n} 0$ if and only if $x \in \text{int}(\mathcal{K}^n)$. The relation $\succeq_{\mathcal{K}^n}$ is a partial ordering, but not a linear ordering in $\mathcal{K}^n$, i.e., there exist $x, y \in \mathcal{K}^n$ such that neither $x \succeq_{\mathcal{K}^n} y$ nor $y \succeq_{\mathcal{K}^n} x$.

Let $\mathcal{S}^n$ be the set of $n \times n$ symmetirc matrices, we denote $\mathcal{S}^n_+$ the cone of all positive semidefinite matrices. It is well known that the nonnegative orthant $\mathbb{R}^n_+$, second-order cone $\mathcal{K}^n$, and positive semidefinite cone $\mathcal{S}^n_+$ are all self-dual, moreover, they all belong to symmetric cones under Euclidean Jordan algebra. A Euclidean Jordan algebra is a triple $(\mathbb{V}, \circ, \langle \cdot, \cdot \rangle_{\mathbb{V}})$, where $(\mathbb{V}, \langle \cdot, \cdot \rangle_{\mathbb{V}})$ is a finite dimensional inner product space over the real field $\mathbb{R}$ and $(x, y) \mapsto x \circ y : \mathbb{V} \times \mathbb{V} \to \mathbb{V}$ is a bilinear mapping satisfying the following three conditions:

(i) $x \circ y = y \circ x$ for all $x, y \in \mathbb{V}$;

(ii) $x \circ (x^2 \circ y) = x^2 \circ (x \circ y)$ for all $x, y \in \mathbb{V}$, where $x^2 := x \circ x$;

(iii) $\langle x \circ y, z \rangle_{\mathbb{V}} = \langle y, x \circ z \rangle_{\mathbb{V}}$ for all $x, y \in \mathbb{V}$.

In a Jordan algebra $(\mathbb{V}, \circ)$, $x \circ y$ is said to be the *Jordan product* of x and y. Note that a Jordan product is not associative, i.e., $x \circ (y \circ z) = (x \circ y) \circ z$ may not hold in general. We assume that there is an element $e \in \mathbb{V}$ such that $x \circ e = x$ for all $x \in \mathbb{V}$ and call e the unit element. Let $\zeta(x)$ be the degree of the minimal polynomial of $x \in \mathbb{V}$, which can be equivalently defined as $\zeta(x) := \min \left\{ k : \{e, x, x^2, \cdots, x^k\} \text{ are linearly dependent} \right\}$. Since $\zeta(x) \leq \dim(\mathbb{V})$, the rank of $(\mathbb{V}, \circ)$ is well defined by $q := \max\{\zeta(x) : x \in \mathbb{V}\}$. In a Euclidean Jordan algebra $\mathbb{A} = (\mathbb{V}, \circ, \langle \cdot, \cdot \rangle_{\mathbb{V}})$, we denote

$$\mathcal{K} := \{x^2 : x \in \mathbb{V}\} \tag{2}$$

by the set of squares. From [10, Theorem III.2.1], $\mathcal{K}$ is a symmetric cone. This means that $\mathcal{K}$ is a self-dual closed convex cone, that is,

$$\mathcal{K} = \mathcal{K}^* := \{y \in \mathbb{V} : \langle x, y \rangle \geq 0 \ \forall x \in \mathcal{K}\},$$

with nonempty interior $\text{int}(\mathcal{K})$, and *homogeneous*, i.e. for any $x, y \in \text{int}(\mathcal{K})$, there exists an invertible linear transformation $\mathcal{T} : \mathbb{V} \to \mathbb{V}$ such that $\mathcal{T}(\mathcal{K}) = \mathcal{K}$ and $\mathcal{T}(x) = y$.

Here are examples of symmetric cones. For $\mathbb{V} = \mathbb{R}^n$, let $\langle \cdot, \cdot \rangle_{\mathbb{V}}$ be the standard vector inner product, namely,

$$\langle x, y \rangle_{\mathbb{V}} := \sum_{i=1}^{n} x_i y_i;$$

and the Jordan product be defined as

$$x \circ y := (x_1 y_1, x_2 y_2, \cdots, x_n y_n).$$

Then, $\mathcal{K} = \mathbb{R}^n_+$ is the cone of squares under this Euclidean Jordan algebra. For $\mathbb{V} = \mathbb{R}^n$, we write $x = (x_1, x_2) \in \mathbb{R} \times \mathbb{R}^{n-1}$. Let $\langle \cdot, \cdot \rangle_{\mathbb{V}}$ be the standard vector inner product and the Jordan product be defined as

$$x \circ y := (\langle x, y \rangle, x_1 y_2 + y_1 x_2).$$

Then, the second-order cone $\mathcal{K}^n$ is the cone of squares under this Euclidean Jordan algebra. For $\mathbb{V} = \mathcal{S}^n$, let $\langle \cdot, \cdot \rangle_{\mathbb{V}}$ be the trace inner product, namely,

$$\langle X, Y \rangle_{\mathbb{V}} := \text{Trace}(XY);$$

and the Jordan product defined as

$$X \circ Y := \frac{1}{2}(XY + YX).$$

Then, the positive semedefinite cone $\mathcal{S}^n_+$ is the cone of squares under this Euclidean Jordan algebra.

Although symmetric cone provide a unified framework for $\mathbb{R}^n_+, \mathcal{K}^n$ and $\mathcal{S}^n_+$, there exist some differences among them. To name a few, (i) there are some "convex" merit functions associated with $\mathbb{R}^n_+$ become "non-convex" when associated with $\mathcal{K}^n$; (ii) different conditions are required to guarantee the coerciveness of some merit functions. Usually stronger conditions are required in $\mathcal{K}^n$ and $\mathcal{S}^n_+$ than in $\mathbb{R}^n_+$. To see these, we illustrate more as below. The Fischer-Burmeister (FB) merit function [8, 9]

$$\psi_{\rm FB}(a,b) := \frac{1}{2}\left|\sqrt{a^2+b^2}-(a+b)\right|^2$$

is known as convex. But the FB merit function associated with SOC

$$\psi_{\rm FB}(x,y) := \frac{1}{2}\left\|(x^2+y^2)^{1/2}-(x+y)\right\|^2$$

is not convex, see [7, Example 3.5]. Another example is the function $|(\phi_{\rm FB}(a,b))_+|^2$ where $(t)_+ := \max\{0,t\}\ \forall t \in \mathbb{R}$. In addition, the Mangasarian-Solodov merit function [15]

$$\phi_{\rm MS}(a,b) := ab + \frac{1}{2\alpha}\left\{[(a-\alpha b)_+]^2 - a^2 + [(b-\alpha a)_+]^2 - b^2\right\},$$

where $\alpha > 0$ ($\neq 1$) is a constant, has the following property [14, Lemma 6.2]:

If $(a \to -\infty)$ or $(b \to -\infty)$ or $(a \to \infty$ and $b \to \infty)$, then $|\phi_{\rm MS}(a,b)| \to \infty$.

However, the MS merit function associated with SOC

$$\phi_{\rm MS}(x,y) := x \circ y + \frac{1}{2\alpha}\left\{[(x-\alpha y)_+]^2 - x^2 + [(y-\alpha x)_+]^2 - y^2\right\},$$

where $\alpha > 0$ ($\neq 1$) is a constant, needs stronger condition to satisfy the same property [17, Prop. 4.2]: If $\{x^k\} \subset \mathbb{V}$ and $\{y^k\} \subset \mathbb{V}$ are the sequences satisfying one of the following conditions:

(i) $\lambda_{\min}(x^k) \to -\infty$;

(ii) $\lambda_{\min}(y^k) \to -\infty$;

(iii) $\lambda_{\min}(x^k)$, $\lambda_{\min}(y^k) > -\infty$, $\lambda_{\max}(x^k)$, $\lambda_{\max}(y^k) \to +\infty$ and $\frac{x^k}{\|x^k\|} \circ \frac{y^k}{\|y^k\|} \not\to 0$,

then $\|\phi_{\rm MS}(x^k,y^k)\| \to \infty$. It is remarked in [17] that the condition $\frac{x^k}{\|x^k\|} \circ \frac{y^k}{\|y^k\|} \not\to 0$ is required in $\mathcal{K}^n$ and $\mathcal{S}^n_+$ cases, though not needed in $\mathbb{R}^n_+$ case.

The above illustrations indicate that some properties associated with $\mathbb{R}^n_+$, $\mathcal{K}^n$ and $\mathcal{S}^n_+$ may vary or change a bit, even though these three cones all belong to symmetric cones. These various properties, due to the different structures of $\mathbb{R}^n_+$, $\mathcal{K}^n$ and $\mathcal{S}^n_+$, have great effect in analyzing optimization problems involved symmetric cones. It is well-known that the analysis for optimization problems involved $\mathbb{R}^n_+$ cannot be employed in analyzing optimization problems involved $\mathcal{K}^n$ and $\mathcal{S}^n_+$ since $\mathcal{K}^n$ and $\mathcal{S}^n_+$ are no longer polyhedral sets. What about any relation between $\mathcal{K}^n$ and $\mathcal{S}^n_+$? In fact, it is known that an SOCP problem can be viewed as an SDP problem via certain relation between $\mathcal{K}^n$ and $\mathcal{S}^n_+$, see [5, 19]. Nonetheless, most analysis techniques used in solution methods for semidefinite programming (SDP) and semidefinite complementarity problem (SDCP) cannot be carried over to second-order cone programming (SOCP) and second-order cone complementarity programming (SOCCP) because the following (see [6]):

(i) $(\mathcal{K}^n, \circ)$ is not closed, whereas $\mathcal{S}^n_+$ is.

(ii) $(\mathcal{K}^n, \circ)$ is not associative, whereas $\mathcal{S}^n_+$ is.

However, every $x = (x_1, x_2) \in \mathbb{R} \times \mathbb{R}^{n-1}$ has a spectral decomposition (see Section 2) while every $X \in \mathcal{S}^n$ has a decomposition $X = P^T DP$. This offers a parallel analysis concept under the sense that we have spectral values (vectors) of x v.s. eigenvalues (eigenvectors) of X. We want to know whether there any other similarities and difference between $\mathcal{K}^n$ and $\mathcal{S}^n_+$, which is the main purpose of this paper.

In what follows and throughout the paper, $\langle \cdot, \cdot \rangle$ denotes the Euclidean inner product and $\| \cdot \|$ is the Euclidean norm. The notation ":=" means "define". For any $f : \mathbb{R}^n \to \mathbb{R}$, $\nabla f(x)$ denotes the gradient of f at x and $\nabla^2 f(x)$ denotes the Hessian matrix of f at x. For any symmetric matrices $A, B \in R^{n\times n}$, we write $A \succeq B$ (respectively, $A \succ B$) to mean $A - B$ is positive semidefinite (respectively, positive definite). At last, $\|A\|_F$ is the Frobenius norm of matrix A.

2 Preliminary

In this section, we recall some concepts of Euclidean Jordan algebra that will be used in the subsequent analysis.

An element $c \in \mathbb{V}$ is said to be an idempotent if $c^2 = c$. Two idempotents c and d are said to be *orthogonal* if $c \circ d = 0$. One says that $\{c_1, c_2, \cdots, c_k\}$ is a complete system of orthogonal idempotents if

$$c_j^2 = c_j,\ c_j \circ c_i = 0 \text{ if } j \neq i \text{ for all } j, i = 1, 2, \cdots, k \text{ and } \sum_{j=1}^{k} c_j = e.$$

An idempotent is said to be *primitive* if it is nonzero and cannot be written as the sum of two other nonzero idempotents. We call a complete system of orthogonal primitive idempotents a *Jordan frame*. Then, we have the second version of the spectral decomposition theorem.

Theorem 2.1 *[10, Theorem III.1.2] Suppose that* $\mathbb{A} = (\mathbb{V}, \circ, \langle\cdot,\cdot\rangle_{\mathbb{V}})$ *is a Euclidean Jordan algebra with rank* q*. Then for each* $x \in \mathbb{V}$*, there exist a Jordan frame* $\{c_1, c_2, \cdots, c_q\}$ *and real numbers* $\lambda_1(x), \lambda_2(x), \cdots, \lambda_q(x)$ *such that* $x = \sum_{j=1}^{q} \lambda_j(x) c_j$*.*

The numbers $\lambda_j(x)$ (counting multiplicities), which are uniquely determined by x, are called the *eigenvalues* of x. In the sequel, we write the maximum eigenvalue and the minimum eigenvalue of x as $\lambda_{\max}(x)$ and $\lambda_{\min}(x)$, respectively. Furthermore, the *trace* and the *determinant* of x, denoted by $\mathrm{tr}(x)$ and $\det(x)$, respectively, are defined as $\mathrm{tr}(x) := \sum_{j=1}^{q} \lambda_j(x)$ and $\det(x) := \prod_{j=1}^{q} \lambda_j(x)$.

By [10, Proposition III.1.5], a Jordan algebra $(\mathbb{V}, \circ)$ over $\mathbb{R}$ with a unit element $e \in \mathbb{V}$ is Euclidean if and only if the symmetric bilinear form $\mathrm{tr}(x \circ y)$ is positive definite. Therefore, we may define an inner product $\langle\cdot,\cdot\rangle$ on $\mathbb{V}$ by

$$\langle x, y\rangle := \mathrm{tr}(x \circ y), \ \forall x, y \in \mathbb{V}.$$

In addition, we let $\|\cdot\|_{\mathbb{V}}$ be the norm on $\mathbb{V}$ induced by the inner product $\langle\cdot,\cdot\rangle$, i.e.,

$$\|x\|_{\mathbb{V}} := \sqrt{\langle x, x\rangle} = \left(\sum_{j=1}^{q} \lambda_j^2(x)\right)^{1/2}, \ \forall x \in \mathbb{V}.$$

The Jordan product "$\circ$" in Lorentz algebra is not associative, which is the main reason of complication in the analysis of SOC. For each $x = (x_1, x_2) \in \mathbb{R} \times \mathbb{R}^{n-1}$, define the matrix L_x by

$$L_x := \begin{bmatrix} x_1 & x_2^T \\ x_2 & x_1 I \end{bmatrix},$$

which can be viewed as a linear mapping from $\mathbb{R}^n$ to $\mathbb{R}^n$ with the following properties.

Property 2.1 **(a)** $L_x y = x \circ y$ *and* $L_{x+y} = L_x + L_y$ *for any* $y \in \mathbb{R}^n$*.*

(b) $x \in \mathcal{K}^n \Longleftrightarrow L_x \succeq O$ *and* $x \in \mathrm{int}(\mathcal{K}^n) \Longleftrightarrow L_x \succ O$*.*

(c) L_x *is invertible whenever* $x \in \mathrm{int}(\mathcal{K}^n)$ *with the inverse* L_x^{-1} *given by*

$$L_x^{-1} = \frac{1}{\det(x)} \begin{bmatrix} x_1 & -x_2^T \\ -x_2 & \dfrac{\det(x)}{x_1} I + \dfrac{x_2 x_2^T}{x_1} \end{bmatrix},$$

where $\det(x) := x_1^2 - \|x_2\|^2$ *denotes the determinant of* x*.*

We next recall from [11] that each $x = (x_1, x_2) \in \mathbb{R} \times \mathbb{R}^{n-1}$ admits a spectral factorization associated with $\mathcal{K}^n$, of the form

$$x = \lambda_1(x) \cdot u_x^{(1)} + \lambda_2(x) \cdot u_x^{(2)}, \tag{3}$$

where $\lambda_1(x), \lambda_2(x)$ and $u_x^{(1)}, u_x^{(2)}$ are the spectral values and the associated spectral vectors of x, respectively, defined by

$$\lambda_i(x) = x_1 + (-1)^i \|x_2\|,$$
$$u_x^{(i)} = \begin{cases} \frac{1}{2}\left(1, (-1)^i \frac{x_2}{\|x_2\|}\right), & \text{if } x_2 \neq 0, \\ \frac{1}{2}\left(1, (-1)^i w\right), & \text{if } x_2 = 0, \end{cases} \tag{4}$$

for $i = 1, 2$, with w being any vector in $\mathbb{R}^{n-1}$ satisfying $\|w\| = 1$. If $x_2 \neq 0$, the factorization is unique.

For any function $f : \mathbb{R} \to \mathbb{R}$, we define a function on $\mathbb{R}^n$ associated with $\mathcal{K}^n$ $(n \geq 1)$ by

$$f^{\text{soc}}(x) := f(\lambda_1(x))u_x^{(1)} + f(\lambda_2(x))u_x^{(2)}, \tag{5}$$

where $\lambda_1(x), \lambda_2(x), u_x^{(1)}, u_x^{(2)}$ are the spectral values and vectors of x in (4). The cases of $f(x) = x^{1/2}, x^2$ have some properties which are summarized as follows.

Property 2.2 *For any $x = (x_1, x_2) \in \mathbb{R} \times \mathbb{R}^{n-1}$, let $\lambda_1(x), \lambda_2(x)$ and $u_x^{(1)}, u_x^{(2)}$ be the spectral values and the associated spectral vectors. Then, the following results holds.*

(a) $x \in \mathcal{K}^n \iff 0 \leq \lambda_1(x) \leq \lambda_2(x)$ *and* $x \in \text{int}(\mathcal{K}^n) \iff 0 < \lambda_1(x) \leq \lambda_2(x)$.

(b) $x^2 = (\lambda_1(x))^2 \cdot u_x^{(1)} + (\lambda_2(x))^2 \cdot u_x^{(2)} \in \mathcal{K}^n$ *for any* $x \in \mathbb{R}^n$.

(c) *If* $x \in \mathcal{K}^n$, *then* $x^{1/2} = \sqrt{\lambda_1(x)} \cdot u_x^{(1)} + \sqrt{\lambda_2(x)} \cdot u_x^{(2)} \in \mathcal{K}^n$.

3 The convexity of function associated with SOC

The functions as below in Property 3.1 are all convex and usually employed as penalty and barrier functions when solving SDP [1, 2]. In this section, we wish to know whether such penalty and barrier functions are still convex in SOC case. Some of them were studied in [1], however, we provide different proofs here.

Property 3.1 *The following functions associated with $\mathcal{S}_+^n$ are convex.*

(a) $F_1(X) = -\ln(\det(-X))$ *for all* $X \prec O$.

(b) $F_2(X) = -\ln(\det(I - X))$ *for all* $X \prec O$.

(c) $F_3(X) = \text{tr}(exp(X))$ *for all* $X \in \mathcal{S}^n$.

(d) $F_4(X) = \text{tr}(-X^{-1})$ *for all* $X \prec O$.

(e) $F_5(X) = \mathrm{tr}((I - X)^{-1} \circ X)$ *for all* $X \prec I$.

(f) $F_6(X) = \ln(\det(I + exp(X)))$ *for all* $X \in \mathcal{S}^n$.

(g) $F_7(X) = \mathrm{tr}\left(\frac{X + (X^2 + 4I)^{1/2}}{2}\right)$ *for all* $X \in \mathcal{S}^n$.

Following are some tools that we will use in the proof of Proposition 3.1.

Lemma 3.1 *For any nonzero vector* $x \in \mathbb{R}^n$*, the matrix* xx^T *is positive semidefinite. Moreover, all eigenvalues of the matrix* xx^T *are* $\|x\|^2$ *and* 0 *with multiplicity* $n-1$.

Proof. For any vector $d \in \mathbb{R}^n$, we have $d^T(xx^T)d = (d^Tx)(x^Td) = (d^Tx)^2 \geq 0$. Hence xx^T is positive semidefinite. Next we calculate the eigenvalues of the matrix xx^T.

$$
\begin{aligned}
\det\left(xx^T - \lambda I\right) &= \begin{vmatrix} x_1^2 - \lambda & x_1x_2 & x_1x_3 & \cdots & x_1x_n \\ x_2x_1 & x_2^2 - \lambda & x_2x_3 & \cdots & x_2x_n \\ x_3x_1 & x_3x_2 & x_3^2 - \lambda & \cdots & x_3x_n \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ x_nx_1 & x_nx_2 & x_nx_3 & \cdots & x_n^2 - \lambda \end{vmatrix} \\
&= \begin{vmatrix} x_1^2 - \lambda & x_2^2 & x_3^2 & \cdots & x_n^2 \\ x_1^2 & x_2^2 - \lambda & x_3^2 & \cdots & x_n^2 \\ x_1^2 & x_2^2 & x_3^2 - \lambda & \cdots & x_n^2 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ x_1^2 & x_2^2 & x_3^2 & \cdots & x_n^2 - \lambda \end{vmatrix} \\
&= \left((x_1^2 + x_2^2 + \cdots + x_n^2) - \lambda\right) \cdot \begin{vmatrix} 1 & x_2^2 & x_3^2 & \cdots & x_n^2 \\ 1 & x_2^2 - \lambda & x_3^2 & \cdots & x_n^2 \\ 1 & x_2^2 & x_3^2 - \lambda & \cdots & x_n^2 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_2^2 & x_3^2 & \cdots & x_n^2 - \lambda \end{vmatrix} \\
&= \left((x_1^2 + x_2^2 + \cdots + x_n^2) - \lambda\right) \cdot \begin{vmatrix} 1 & 0 & 0 & \cdots & 0 \\ 1 & -\lambda & 0 & \cdots & 0 \\ 1 & 0 & -\lambda & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & 0 & 0 & \cdots & -\lambda \end{vmatrix} \\
&= \left((x_1^2 + x_2^2 + \cdots + x_n^2) - \lambda\right) \cdot (-\lambda)^{n-1}.
\end{aligned}
$$

This shows that all eigenvalues of the matrix xx^T are $\|x\|^2$ and 0 with multiplicity $n-1$. $\square$

Lemma 3.2 *Let λ be an eigenvalue of the matrix M, then $\lambda + k$ is an eigenvalue of the matrix $(M + kI)$.*

Lemma 3.3 *[13, Theorem 7.7.6] Suppose that a symmetric matrix is partitioned as $\begin{bmatrix} A & B \\ B^T & C \end{bmatrix}$, where A and C are square. Then this matrix is positive definite if and only if $A \succ O$ and $C \succ B^T A^{-1} B$.*

Lemma 3.4 *Suppose that a symmetric matrix is partitioned as*

$$\begin{bmatrix} a & b\dfrac{x^T}{\|x\|} \\ b\dfrac{x}{\|x\|} & cI + (a-c)\dfrac{xx^T}{\|x\|^2} \end{bmatrix} = \begin{bmatrix} A & B \\ B^T & C \end{bmatrix}, \tag{6}$$

where $a, b, c \in \mathbb{R}, x \in \mathbb{R}^n, I \in \mathbb{R}^{n\times n}$. Then this matrix is positive definite if and only if $a > 0, c > 0$ and $a^2 - b^2 > 0$.

Proof. From Lemma 3.3, this matrix is positive definite if and only if $A \succ O$ and $C \succ B^T A^{-1} B$. Since $a \in \mathbb{R}$, $a > 0$ implies $A \succ O$ On the other hand,

$$\begin{aligned} AC - B^T B &= a\left(cI + (a-c)\frac{xx^T}{\|x\|^2}\right) - b^2\frac{xx^T}{\|x\|^2} \\ &= acI + \left(a^2 - ac\right)\frac{xx^T}{\|x\|^2} - b^2\frac{xx^T}{\|x\|^2} \\ &= acI + \left(a^2 - ac - b^2\right)\frac{xx^T}{\|x\|^2} \\ &= M, \end{aligned}$$

where we denote the whole matrix by M. From Lemma 3.1, we know that xx^T is positive semidefinite with only one nonzero eigenvalue $\|x\|^2$. From Lemma 3.2, all the eigenvalues of the matrix M are $ac + \dfrac{a^2 - ac - b^2}{\|x\|^2} \cdot \|x\|^2 = a^2 - b^2$ and ac with multiplicity of $n-1$, which are all positive. This shows $C \succ B^T A^{-1} B$ and the proof is complete. □

Lemma 3.5 *[11, Proposition 3.2]Let $x = (x_1, x_2) \in \mathbb{R} \times \mathbb{R}^{n-1}$ and $f : \mathbb{R} \to \mathbb{R}$.*

(a) *If $f(t) = \exp(t)$, then*

$$f^{\text{soc}}(x) = \exp(x) = \begin{cases} \exp(x_1)\left(\cosh(\|x_2\|), \sinh(\|x_2\|) \cdot \dfrac{x_2}{\|x_2\|}\right), & \text{if } x_2 \neq 0, \\ \exp(x_1)(1, 0), & \text{if } x_2 = 0, \end{cases}$$

where $\cosh(\alpha) = (\exp(\alpha) + \exp(-\alpha))/2$ and $\sinh(\alpha) = (\exp(\alpha) - \exp(-\alpha))/2$ for $\alpha \in \mathbb{R}$.

(b) *If $f(t) = \ln(t)$ and $x \in \text{int}(\mathcal{K}^n)$, then*

$$f^{\text{soc}}(x) = \ln(x) = \begin{cases} \frac{1}{2}\left(\ln(x_1^2 - \|x_2\|^2), \ln\left(\frac{x_1 + \|x_2\|}{x_1 - \|x_2\|}\right) \cdot \frac{x_2}{\|x_2\|}\right), & \text{if } x_2 \neq 0, \\ \ln(x_1)(1, 0), & \text{if } x_2 = 0. \end{cases}$$

Property 3.2 *Let S be a nonempty open convex set in $\mathbb{R}^n$, and let $f : S \to \mathbb{R}$.*

(1) (First-order condition) *[3, Theorem 3.3.3] Suppose f is differentiable on S. Then f is convex if and only if*

$$f(y) \geq f(x) + \nabla f(x)^T (y - x), \forall x, y \in S.$$

(2) (Second-order condition) *[3, Theorem 3.3.7] Suppose f is twice differentiable on S, that is, its Hessian or second derivative $\nabla^2 f(x)$ exists. Then f is convex if and only if*

$$\nabla^2 f(x) \succeq O, \forall x \in S.$$

Property 3.3 *[5, Proposition 4] For any $f : \mathbb{R} \to \mathbb{R}$, the following result hold:*

(a) *f^{soc} is differentiable at an $x = (x_1, x_2) \in \mathbb{R} \times \mathbb{R}^{n-1}$ with spectral values λ_1, λ_2 if and only if f is differentiable at λ_1, λ_2. Moreover,*

$$\nabla f^{\text{soc}}(x) = f'(x_1) I$$

if $x_2 = 0$, and otherwise

$$\nabla f^{\text{soc}}(x) = \begin{bmatrix} a & b\frac{x_2^T}{\|x_2\|} \\ b\frac{x_2}{\|x_2\|} & cI + (a - c)\frac{x_2 x_2^T}{\|x_2\|^2} \end{bmatrix},$$

where $a = \frac{1}{2}\left(f'(\lambda_2) + f'(\lambda_1)\right), b = \frac{1}{2}\left(f'(\lambda_2) - f'(\lambda_1)\right)$ and $c = \frac{f(\lambda_2) - f(\lambda_1)}{\lambda_2 - \lambda_1}$.

(b) *f^{soc} is differentiable if and only if f is differentiable.*

Lemma 3.6 *For any $x = (x_1, x_2) \in \mathbb{R} \times \mathbb{R}^{n-1}$, we define $w, z : \mathbb{R}^n \to \mathbb{R}^n$ by*

$$\begin{aligned} w &= (w_1, w_2) = (w_1(x), w_2(x)) = w(x) := x^2 + 4e, \\ z &= (z_1, z_2) = (z_1(x), z_2(x)) = z(x) := \left(x^2 + 4e\right)^{1/2}. \end{aligned} \tag{7}$$

Then $z(x)$ is differentiable. Moreover, $\nabla z(x) = L_x L_z^{-1}$, where $L_z^{-1} = (1/\sqrt{w_1})I$ if $w_2 = 0$, and otherwise

$$L_z^{-1} = \begin{bmatrix} a & b\frac{w_2^T}{\|w_2\|} \\ b\frac{w_2}{\|w_2\|} & cI + (a-c)\frac{w_2 w_2^T}{\|w_2\|^2} \end{bmatrix}$$

with

$$\begin{aligned} a &= \frac{1}{2}\left(\frac{1}{\sqrt{\lambda_2(w)}} + \frac{1}{\sqrt{\lambda_1(w)}}\right), \\ b &= \frac{1}{2}\left(\frac{1}{\sqrt{\lambda_2(w)}} - \frac{1}{\sqrt{\lambda_1(w)}}\right), \\ c &= \frac{2}{\sqrt{\lambda_2(w)} + \sqrt{\lambda_1(w)}}. \end{aligned}$$

Proof. We denote $g(x) = (x)^{1/2}$, which implies $z = w^{1/2} = g(w(x))$. Since $g(\cdot) = (\cdot)^{1/2}$ is continuously differentiable on $\mathbb{R}_{++}$, by Property 3.3(b), its corresponding SOC-function $g(x) = (x)^{1/2}$ is continuously differentiable on $\text{int}(\mathcal{K}^n)$. Then, for $w_2 \neq 0$, Property 3.3(a) and the chain rule give

$$\nabla z(x) = \nabla w(x) \cdot \nabla g(w) = 2L_x \cdot \frac{1}{2} L_{w^{1/2}}^{-1} = L_x \cdot L_z^{-1},$$

where $\nabla g(w) = \frac{1}{2} L_{w^{1/2}}^{-1}$ is from [11, page 454]. The case for $w_2 = 0$ is clear. Thus, the proof is complete. $\square$

Now, we are in position to prove the convexity of following functions. Our main method is to prove that the Hessian matrix of following functions are positive semidefinite.

Proposition 3.1 *The following functions associated with SOC are all convex.*

(a) $f_1(x) = -\ln(\det(-x))$, *for all* $x \prec_{\mathcal{K}^n} 0$.

(b) $f_2(x) = -\ln(\det(e - x))$, *for all* $x \prec_{\mathcal{K}^n} e$.

(c) $f_3(x) = \text{tr}(\exp(x))$, *for all* $x \in \mathbb{R}^n$.

(d) $f_4(x) = \text{tr}(-x^{-1})$, *for all* $x \prec_{\mathcal{K}^n} 0$.

(e) $f_5(x) = \text{tr}((e - x)^{-1} \circ x)$, *for all* $x \prec_{\mathcal{K}^n} e$.

(f) $f_6(x) = \ln(\det(e + \exp(x)))$, *for all* $x \in \mathbb{R}^n$.

(g) $f_7(x) = \text{tr}\left(\frac{x + (x^2 + 4e)^{1/2}}{2}\right)$, *for all* $x \in \mathbb{R}^n$.

Proof. **(a)** $f_1(x) = -\ln(\det(-x)) = -\ln(\det(x))$, from [7, Proposition 2.4], $\ln(\det(x))$ is concave for all $x \succ_{\mathcal{K}^n} 0$. This implies that $f_1(x)$ is convex for all $x \prec_{\mathcal{K}^n} 0$.
(b) Since $f_2(x) = f_1(x-e)$, $f_2(x)$ is convex when $x \prec_{\mathcal{K}^n} e$.
(c) From Lemma 3.5(a), we have

$$f_3(x) = \begin{cases} \exp(x_1)\cosh(\|x_2\|) & \text{if } x_2 \neq 0, \\ \exp(x_1) & \text{if } x_2 = 0. \end{cases}$$

Case(1): $x_2 \neq 0$. Since $\mathbb{R}^n$ is a convex set, it suffices to show that $\nabla^2 f_3(x)$ is positive definite for all $x \in \mathbb{R}^n$. From direct computation, we have

$$\begin{aligned} \nabla f_3(x) &= \begin{bmatrix} 2\exp(x_1)\cosh(\|x_2\|) \\ 2\exp(x_1)\sinh(\|x_2\|)\dfrac{x_2}{\|x_2\|} \end{bmatrix} \\ &= 2\exp(x_1) \begin{bmatrix} \cosh(\|x_2\|) \\ \sinh(\|x_2\|)\dfrac{x_2}{\|x_2\|} \end{bmatrix}, \end{aligned}$$

and

$$\begin{aligned} \frac{\partial^2 f_3}{\partial x_1^2} &= 2\exp(x_1)\cosh(\|x_2\|), \\ \frac{\partial^2 f_3}{\partial x_1 \partial x_2} &= 2\exp(x_1)\sinh(\|x_2\|)\frac{x_2^T}{\|x_2\|}, \\ \frac{\partial^2 f_3}{\partial x_2 \partial x_1} &= 2\exp(x_1)\sinh(\|x_2\|)\frac{x_2}{\|x_2\|}, \\ \frac{\partial^2 f_3}{\partial x_2^2} &= 2\exp(x_1)\left(\frac{\cosh(\|x_2\|)\frac{x_2}{\|x_2\|}\cdot\|x_2\| - \sinh(\|x_2\|)\frac{x_2}{\|x_2\|}}{\|x_2\|^2}x_2^T + \frac{\sinh(\|x_2\|)}{\|x_2\|}I\right) \\ &= 2\exp(x_1)\left(\frac{\sinh(\|x_2\|)}{\|x_2\|}I + \left(\cosh(\|x_2\|) - \frac{\sinh(\|x_2\|)}{\|x_2\|}\right)\frac{x_2x_2^T}{\|x_2\|^2}\right). \end{aligned}$$

Therefore,

$$\begin{aligned} \nabla^2 f_3(x) &= 2\exp(x_1)\cdot \begin{bmatrix} a & b\dfrac{x_2^T}{\|x_2\|} \\ b\dfrac{x_2}{\|x_2\|} & cI + (a-c)\dfrac{x_2x_2^T}{\|x_2\|^2} \end{bmatrix} \\ &= 2\exp(x_1)\cdot M, \end{aligned}$$

where $a = \cosh(\|x_2\|), b = \sinh(\|x_2\|)$ and $c = \dfrac{\sinh(\|x_2\|)}{\|x_2\|}$. Since $a > 0, c > 0$ and $a^2 - b^2 = \cosh^2(\|x_2\|) - \sinh^2(\|x_2\|) = 1 > 0$, from Lemma 3.4, we have M is positive definite. Also notice that $\exp(x_1) > 0$, thus $\nabla^2 f_3(x)$ is positive definite for all $x \in \mathbb{R}^n$.
Case(2): $x_2 = 0$. It is not hard to compute

$$\nabla^2 f_3(x) = \begin{bmatrix} 2\exp(x_1) & 0^T \\ 0 & O \end{bmatrix},$$

which is positive semidefinite. Hence $f_3(x)$ is convex for all $x \in \mathbb{R}^n$.

(d) From direct computation, we have $-x^{-1} = \dfrac{-1}{x_1^2 - \|x_2\|^2}(x_1, -x_2)$. Hence, $f_4(x) = \mathrm{tr}(-x^{-1}) = \dfrac{-2x_1}{x_1^2 - \|x_2\|^2}$. Since $\mathrm{int}(-\mathcal{K}^n)$ is a convex set, it suffices to show that $\nabla^2 f_4(x)$ is positive definite for all $x \in \mathrm{int}(-\mathcal{K}^n)$. From direct computation, we have

$$
\begin{aligned}
\nabla f_4(x) &= -2 \begin{bmatrix} \dfrac{(x_1^2 - \|x_2\|^2) - 2x_1 x_1}{(x_1^2 - \|x_2\|^2)^2} \\ \dfrac{-x_1(-2x_2)}{(x_1^2 - \|x_2\|^2)^2} \end{bmatrix} \\
&= \frac{2}{(x_1^2 - \|x_2\|^2)^2} \begin{bmatrix} x_1^2 + \|x_2\|^2 \\ -2x_1 x_2 \end{bmatrix},
\end{aligned}
$$

and

$$
\begin{aligned}
\frac{\partial^2 f_4}{\partial x_1^2} &= \frac{2}{(x_1^2 - \|x_2\|^2)^4} \big((x_1^2 - \|x_2\|^2)^2 \cdot 2x_1 - (x_1^2 + \|x_2\|^2) \cdot 2(x_1^2 - \|x_2\|^2) \cdot 2x_1\big) \\
&= \frac{-4}{(x_1^2 - \|x_2\|^2)^3} \cdot x_1(x_1^2 + 3\|x_2\|^2), \\
\frac{\partial^2 f_4}{\partial x_1 \partial x_2} &= \frac{2}{(x_1^2 - \|x_2\|^2)^4} \big((x_1^2 - \|x_2\|^2)^2 \cdot (-2x_2^T) - (-2x_1 x_2^T) \cdot 2(x_1^2 - \|x_2\|^2) \cdot 2x_1\big) \\
&= \frac{4}{(x_1^2 - \|x_2\|^2)^3} (3x_1^2 + \|x_2\|^2) x_2^T, \\
\frac{\partial^2 f_4}{\partial x_2 \partial x_1} &= \frac{4}{(x_1^2 - \|x_2\|^2)^3} (3x_1^2 + \|x_2\|^2) x_2, \\
\frac{\partial^2 f_4}{\partial x_2^2} &= \frac{2}{(x_1^2 - \|x_2\|^2)^4} \big((x_1^2 - \|x_2\|^2)^2 \cdot (-2x_1 I) - (-2x_1 x_2) \cdot 2(x_1^2 - \|x_2\|^2) \cdot (-2x_2^T)\big) \\
&= \frac{-4}{(x_1^2 - \|x_2\|^2)^3} \cdot x_1 \big((x_1^2 - \|x_2\|^2) I + 4x_2 x_2^T\big).
\end{aligned}
$$

Therefore,

$$
\nabla^2 f_4(x) = \frac{4}{(x_1^2 - \|x_2\|^2)^3} \cdot \begin{bmatrix} A & B \\ B^T & C \end{bmatrix},
$$

where

$$
\begin{aligned}
A &= -x_1(x_1^2 + 3\|x_2\|^2), \\
B &= (3x_1^2 + \|x_2\|^2) x_2^T, \\
C &= -x_1 \big((x_1^2 - \|x_2\|^2) I + 4x_2 x_2^T\big).
\end{aligned}
$$

Since $x \in \mathrm{int}(-\mathcal{K}^n)$, we have $\dfrac{4}{(x_1^2 - \|x_2\|^2)^3} > 0$. From Lemma 3.3, it suffices to show that $A \succ O$ (here A is a scalar) and $C \succ B^T A^{-1} B$. First, $A > 0$ since $x \prec_{\mathcal{K}^n} 0 \Rightarrow x_1 < 0$.

Second, we show that $C \succ B^T A^{-1} B$.

$$\begin{aligned} AC - B^T B &= x_1^2(x_1^2 + 3\|x_2\|^2)\left((x_1^2 - \|x_2\|^2)I + 4x_2x_2^T\right) - \left(3x_1^2 + \|x_2\|^2\right)^2 x_2x_2^T \\ &= (x_1^2 - \|x_2\|^2)\left(x_1^2(x_1^2 + 3\|x_2\|^2)I - (5x_1^2 - \|x_2\|^2)x_2x_2^T\right) \\ &= (x_1^2 - \|x_2\|^2) \cdot M, \end{aligned}$$

where we denote the whole matrix by M. From Lemma 3.1, we know that xx^T is positive semidefinite with only one nonzero eigenvalue $\|x\|^2$. From Lemma 3.2, all the eigenvalues of the matrix M are $(x_1^2 + 3\|x_2\|^2)$ with multiplicity $n-2$ and

$$\begin{aligned} & x_1^2(x_1^2 + 3\|x_2\|^2) - (5x_1^2 - \|x_2\|^2) \cdot \|x_2\|^2 \\ =\ & x_1^4 + 3x_1^2\|x_2\|^2 - 5x_1^2\|x_2\|^2 + \|x_2\|^4 \\ =\ & (x_1^2 - \|x_2\|^2)^2, \end{aligned}$$

they are all positive. Thus, $M \succ O$. This implies $AC - B^T B$ is positive definite and hence $C \succ B^T A^{-1} B$. Thus, $f_4(x)$ is (strictly) convex for all $x \prec_{\mathcal{K}^n} 0$.

(e) From direct computation, we have

$$\begin{aligned} (e-x)^{-1} \circ x &= (1 - x_1, -x_2)^{-1} \circ (x_1, x_2) \\ &= \frac{1}{(1-x_1)^2 - \|-x_2\|^2}(1 - x_1, x_2) \circ (x_1, x_2) \\ &= \frac{1}{(1-x_1)^2 - \|x_2\|^2}\left(x_1(1-x_1) + \|x_2\|^2, x_1x_2 + (1-x_1)x_2\right) \\ &= \frac{1}{(1-x_1)^2 - \|x_2\|^2}\left(x_1 - x_1^2 + \|x_2\|^2, x_2\right). \end{aligned}$$

Hence,

$$\begin{aligned} f_5(x) &= \operatorname{tr}\left((e-x)^{-1} \circ x\right) \\ &= 2 \cdot \frac{x_1 - x_1^2 + \|x_2\|^2}{(1-x_1)^2 - \|x_2\|^2} \\ &= 2 \cdot \frac{(\|x_2\|^2 - (x_1 - 1)^2) + (1 - x_1)}{(1-x_1)^2 - \|x_2\|^2} \\ &= 2 \cdot \left(\frac{1 - x_1}{(1-x_1)^2 - \|x_2\|^2} - 1\right). \end{aligned}$$

Let $S = \{x \mid x \prec_{\mathcal{K}^n} e, x \in \mathbb{R}^n\}$. Since S is a convex set, it suffices to show that $\nabla^2 f_5(x)$ is positive definite for all $x \in S$. From direct computation, we have

$$\begin{aligned} \nabla f_5(x) &= 2 \begin{bmatrix} \dfrac{((1-x_1)^2 - \|x_2\|^2) \cdot (-1) - (1-x_1) \cdot (-2(1-x_1))}{((1-x_1)^2 - \|x_2\|^2)^2} \\ \dfrac{-(1-x_1)(-2x_2)}{((1-x_1)^2 - \|x_2\|^2)^2} \end{bmatrix} \\ &= \frac{2}{((1-x_1)^2 - \|x_2\|^2)^2} \begin{bmatrix} (1-x_1)^2 + \|x_2\|^2 \\ 2(1-x_1)x_2 \end{bmatrix}, \end{aligned}$$

and

$$
\begin{aligned}
\frac{\partial^2 f_5}{\partial x_1^2} &= \frac{2}{((1-x_1)^2-\|x_2\|^2)^4}\Big(\left((1-x_1)^2-\|x_2\|^2\right)^2 \cdot (-2)(1-x_1) \\
&\quad - \left((1-x_1)^2+\|x_2\|^2\right)\cdot 2\left((1-x_1)^2-\|x_2\|^2\right)\cdot(-2)(1-x_1)\Big) \\
&= \frac{4}{((1-x_1)^2-\|x_2\|^2)^3}(1-x_1)((1-x_1)^2+3\|x_2\|^2), \\
\frac{\partial^2 f_5}{\partial x_1 \partial x_2} &= \frac{2}{((1-x_1)^2-\|x_2\|^2)^4}\Big(\left((1-x_1)^2-\|x_2\|^2\right)^2 \cdot (-2x_2^T) \\
&\quad - 2(1-x_1)x_2^T\cdot 2\left((1-x_1)^2-\|x_2\|^2\right)\cdot(-2)(1-x_1)\Big) \\
&= \frac{4}{((1-x_1)^2-\|x_2\|^2)^3}\left(3(1-x_1)^2+\|x_2\|^2\right)x_2^T, \\
\frac{\partial^2 f_5}{\partial x_2 \partial x_1} &= \frac{4}{((1-x_1)^2-\|x_2\|^2)^3}\left(3(1-x_1)^2+\|x_2\|^2\right)x_2, \\
\frac{\partial^2 f_5}{\partial x_2^2} &= \frac{2}{((1-x_1)^2-\|x_2\|^2)^4}\Big(\left((1-x_1)^2-\|x_2\|^2\right)^2 \cdot 2(1-x_1)I \\
&\quad - 2(1-x_1)x_2\cdot 2\left((1-x_1)^2-\|x_2\|^2\right)\cdot(-2x_2^T)\Big) \\
&= \frac{4}{((1-x_1)^2-\|x_2\|^2)^3}(1-x_1)\left(\left((1-x_1)^2-\|x_2\|^2\right)I+4x_2x_2^T\right).
\end{aligned}
$$

Therefore,

$$
\nabla^2 f_5(x) = \frac{4}{((1-x_1)^2-\|x_2\|^2)^3}\cdot\begin{bmatrix} A & B \\ B^T & C \end{bmatrix},
$$

where

$$
\begin{aligned}
A &= (1-x_1)((1-x_1)^2+3\|x_2\|^2), \\
B &= \left(3(1-x_1)^2+\|x_2\|^2\right)x_2^T, \\
C &= (1-x_1)\left(\left((1-x_1)^2-\|x_2\|^2\right)I+4x_2x_2^T\right).
\end{aligned}
$$

Since $x \in S$, we have $\frac{4}{((1-x_1)^2-\|x_2\|^2)^3} > 0$. From Lemma 3.3, it suffices to show that $A \succ O$ (here A is a scalar) and $C \succ B^TA^{-1}B$ First, $A > 0$ since $x \prec_{\mathcal{K}^n} e \Rightarrow 1-x_1 > 0$. Second, we show that $C \succ B^TA^{-1}B$.

$$
\begin{aligned}
AC - B^TB &= (1-x_1)^2((1-x_1)^2+3\|x_2\|^2)\left(\left((1-x_1)^2-\|x_2\|^2\right)I+4x_2x_2^T\right) \\
&\quad - \left(3(1-x_1)^2+\|x_2\|^2\right)^2 x_2x_2^T \\
&= \left((1-x_1)^2-\|x_2\|^2\right)\Big((1-x_1)^2((1-x_1)^2+3\|x_2\|^2)I \\
&\quad - (5(1-x_1)^2-\|x_2\|^2)x_2x_2^T\Big) \\
&= \left((1-x_1)^2-\|x_2\|^2\right)\cdot M,
\end{aligned}
$$

where we denote the whole matrix by M. From Lemma 3.1, we know that xx^T is positive semidefinite with only one nonzero eigenvalue $\|x\|^2$. From Lemma 3.2, all the eigenvalues of the matrix M are $((1-x_1)^2 + 3\|x_2\|^2)$ with multiplicity $n-2$ and

$$\begin{aligned}
&(1-x_1)^2((1-x_1)^2 + 3\|x_2\|^2) - (5(1-x_1)^2 - \|x_2\|^2) \cdot \|x_2\|^2 \\
=\ &(1-x_1)^4 + 3(1-x_1)^2\|x_2\|^2 - 5(1-x_1)^2\|x_2\|^2 + \|x_2\|^4 \\
=\ &\left((1-x_1)^2 - \|x_2\|^2\right)^2,
\end{aligned}$$

they are all positive. Thus, M is positive definite. This implies $AC - B^TB$ is positive definite and hence $C \succ B^TA^{-1}B$. Thus, $f_5(x)$ is (strictly) convex for all $x \prec_{\mathcal{K}^n} e$.

(f) Similarly as (c), we consider following two cases:

Case(1): $x_2 \neq 0$. From Lemma 3.5(a), we have

$$e + \exp(x) = \left(1 + \exp(x_1)\cosh(\|x_2\|), \exp(x_1)\sinh(\|x_2\|)\frac{x_2}{\|x_2\|}\right).$$

Here we denote $e^\alpha = \exp(\alpha), \forall \alpha \in \mathbb{R}$ for convenience. Therefore,

$$\begin{aligned}
\det(e + \exp(x)) &= (1 + e^{x_1}\cosh(\|x_2\|))^2 - \left\| e^{x_1}\sinh(\|x_2\|)\frac{x_2}{\|x_2\|}\right\|^2 \\
&= \left(1 + e^{x_1 + \|x_2\|}\right)\left(1 + e^{x_1 - \|x_2\|}\right),
\end{aligned}$$

and

$$\begin{aligned}
f_6(x) &= \ln(\det(e + \exp(x))) \\
&= \ln\left(\left(1 + e^{x_1+\|x_2\|}\right)\left(1 + e^{x_1-\|x_2\|}\right)\right) \\
&= \ln\left(1 + e^{x_1+\|x_2\|}\right) + \ln\left(1 + e^{x_1-\|x_2\|}\right).
\end{aligned}$$

Since $\mathbb{R}^n$ is a convex set, it suffices to show that $\nabla^2 f_6(x)$ is positive definite for all $x \in \mathbb{R}^n$. From direct computation, we have

$$\begin{aligned}
\nabla f_6(x) &= \begin{bmatrix} \dfrac{e^{x_1+\|x_2\|}}{1 + e^{x_1+\|x_2\|}} + \dfrac{e^{x_1-\|x_2\|}}{1 + e^{x_1-\|x_2\|}} \\ \left(\dfrac{e^{x_1+\|x_2\|}}{1 + e^{x_1+\|x_2\|}} - \dfrac{e^{x_1-\|x_2\|}}{1 + e^{x_1-\|x_2\|}}\right)\dfrac{x_2}{\|x_2\|} \end{bmatrix} \\
&= \begin{bmatrix} 2 - \dfrac{1}{1 + e^{x_1+\|x_2\|}} - \dfrac{1}{1 + e^{x_1-\|x_2\|}} \\ \left(\dfrac{1}{1 + e^{x_1-\|x_2\|}} - \dfrac{1}{1 + e^{x_1+\|x_2\|}}\right)\dfrac{x_2}{\|x_2\|} \end{bmatrix},
\end{aligned}$$

and

$$
\begin{aligned}
\frac{\partial^2 f_6}{\partial x_1^2} &= \frac{e^{x_1+\|x_2\|}}{\left(1+e^{x_1+\|x_2\|}\right)^2} + \frac{e^{x_1-\|x_2\|}}{\left(1+e^{x_1-\|x_2\|}\right)^2}, \\
\frac{\partial^2 f_6}{\partial x_1 \partial x_2} &= \left(\frac{e^{x_1+\|x_2\|}}{\left(1+e^{x_1+\|x_2\|}\right)^2} - \frac{e^{x_1-\|x_2\|}}{\left(1+e^{x_1-\|x_2\|}\right)^2} \right) \frac{x_2^T}{\|x_2\|}, \\
\frac{\partial^2 f_6}{\partial x_2 \partial x_1} &= \left(\frac{e^{x_1+\|x_2\|}}{\left(1+e^{x_1+\|x_2\|}\right)^2} - \frac{e^{x_1-\|x_2\|}}{\left(1+e^{x_1-\|x_2\|}\right)^2} \right) \frac{x_2}{\|x_2\|}, \\
\frac{\partial^2 f_6}{\partial x_2^2} &= \left(\frac{1}{1+e^{x_1-\|x_2\|}} - \frac{1}{1+e^{x_1+\|x_2\|}} \right) \frac{1}{\|x_2\|} I \\
&\quad + \frac{1}{\|x_2\|^2} \left(\|x_2\| \left(\frac{e^{x_1+\|x_2\|}}{\left(1+e^{x_1+\|x_2\|}\right)^2} + \frac{e^{x_1-\|x_2\|}}{\left(1+e^{x_1-\|x_2\|}\right)^2} \right) \frac{x_2}{\|x_2\|} \right. \\
&\quad \left. - \left(\frac{1}{1+e^{x_1-\|x_2\|}} - \frac{1}{1+e^{x_1+\|x_2\|}} \right) \frac{x_2}{\|x_2\|} \right) x_2^T \\
&= \frac{1}{\|x_2\|} \left(\frac{1}{1+e^{x_1-\|x_2\|}} - \frac{1}{1+e^{x_1+\|x_2\|}} \right) I \\
&\quad + \left(\left(\frac{e^{x_1+\|x_2\|}}{\left(1+e^{x_1+\|x_2\|}\right)^2} + \frac{e^{x_1-\|x_2\|}}{\left(1+e^{x_1-\|x_2\|}\right)^2} \right) \right. \\
&\quad \left. - \frac{1}{\|x_2\|} \left(\frac{1}{1+e^{x_1-\|x_2\|}} - \frac{1}{1+e^{x_1+\|x_2\|}} \right) \right) \frac{x_2 x_2^T}{\|x_2\|^2}.
\end{aligned}
$$

Therefore,

$$
\nabla^2 f_6(x) = \begin{bmatrix} a & b \dfrac{x_2^T}{\|x_2\|} \\ b \dfrac{x_2}{\|x_2\|} & cI + (a-c) \dfrac{x_2 x_2^T}{\|x_2\|^2} \end{bmatrix},
$$

with

$$
\begin{aligned}
a &= \frac{e^{x_1+\|x_2\|}}{\left(1+e^{x_1+\|x_2\|}\right)^2} + \frac{e^{x_1-\|x_2\|}}{\left(1+e^{x_1-\|x_2\|}\right)^2}, \\
b &= \frac{e^{x_1+\|x_2\|}}{\left(1+e^{x_1+\|x_2\|}\right)^2} - \frac{e^{x_1-\|x_2\|}}{\left(1+e^{x_1-\|x_2\|}\right)^2}, \\
c &= \frac{1}{\|x_2\|} \left(\frac{1}{1+e^{x_1-\|x_2\|}} - \frac{1}{1+e^{x_1+\|x_2\|}} \right).
\end{aligned}
$$

From Lemma 3.4, it suffices to show that $a > 0, c > 0$ and $a^2 - b^2 > 0$. Note that $a > 0$ is obviously, now we proof $c > 0$ and $a^2 - b^2 > 0$. From direct computation, we have

$$\begin{aligned} c &= \frac{1}{\|x_2\|}\left(\frac{1}{1+e^{x_1-\|x_2\|}} - \frac{1}{1+e^{x_1+\|x_2\|}}\right) \\ &= \frac{e^{x_1}\left(e^{\|x_2\|} - e^{-\|x_2\|}\right)}{\|x_2\|\left(1+e^{x_1+\|x_2\|}\right)\left(1+e^{x_1-\|x_2\|}\right)} \\ &> 0, \end{aligned} \tag{8}$$

and

$$\begin{aligned} a^2 - b^2 &= (a+b)(a-b) \\ &= \frac{4\cdot e^{2x_1}}{\left(1+e^{x_1+\|x_2\|}\right)^2\left(1+e^{x_1-\|x_2\|}\right)^2} \\ &> 0, \end{aligned}$$

where the last inequality in (8) is because of the property of exponential function. Thus M is positive definite and $\nabla^2 f_6(x)$ is positive definite for all $x \in \mathbb{R}^n$.
Case(2): $x_2 = 0$. From Lemma 3.5, we have $e + \exp(x) = (1 + e^{x_1}, 0)$. Therefore, $f_6(x) = \ln(\det(e + \exp(x)) = 2\ln(1 + e^{x_1})$. It is not hard to compute that

$$\nabla^2 f_6(x) = \begin{bmatrix} \frac{2e^{x_1}}{(1+e^{x_1})^2} & 0^T \\ 0 & O \end{bmatrix},$$

which is positive semidefinite Hence $f_6(x)$ is convex for all $x \in \mathbb{R}^n$.
(g) Let w, z be defined as in (7) and $\lambda_1 = \lambda_1(w), \lambda_2 = \lambda_2(w)$ for convenience. Then $f_7(x) = \operatorname{tr}\left(\frac{x+z(x)}{2}\right)$. From direct computation, we have $w = (w_1, w_2)$ with $w_1 = x_1^2 + \|x_2\|^2 + 4$ and $w_2 = 2x_1x_2$.
Case(1): $w_2 \neq 0$. Since $\mathbb{R}^n$ is a convex set, it suffices to show that $\nabla^2 f_7(x)$ is positive definite for all $x \in \mathbb{R}^n$. Note that $\operatorname{tr}\left(\frac{x+z(x)}{2}\right) = \operatorname{tr}\left(\frac{x}{2}\right) + \operatorname{tr}\left(\frac{z(x)}{2}\right)$, and $\nabla^2 \operatorname{tr}\left(\frac{x}{2}\right) = 0$, we only consider the Hessian of $\operatorname{tr}\left(\frac{z(x)}{2}\right)$. Moreover, we assume $x_1 \geq 0$. The case $x_1 < 0$ is similar as above, we omit here. Therefore, from direct computation, we have

$$\begin{aligned} \lambda_1 = \lambda_1(w) = w_1 - \|w_2\| = (x_1 - \|x_2\|)^2 + 4, \\ \lambda_2 = \lambda_2(w) = w_1 + \|w_2\| = (x_1 + \|x_2\|)^2 + 4, \end{aligned}$$

and

$$\begin{aligned} \nabla_{x_1}(\sqrt{\lambda_1}) = \frac{x_1 - \|x_2\|}{\sqrt{\lambda_1}}, \quad \nabla_{x_2}(\sqrt{\lambda_1}) = \frac{-(x_1 - \|x_2\|)}{\sqrt{\lambda_1}}\frac{x_2}{\|x_2\|}, \\ \nabla_{x_1}(\sqrt{\lambda_2}) = \frac{x_1 + \|x_2\|}{\sqrt{\lambda_2}}, \quad \nabla_{x_2}(\sqrt{\lambda_2}) = \frac{x_1 + \|x_2\|}{\sqrt{\lambda_2}}\frac{x_2}{\|x_2\|}. \end{aligned} \tag{9}$$

By chain rule and Lemma 3.6, we have

$$
\begin{aligned}
\nabla f_7(x) &= \nabla z(x) \cdot \nabla \mathrm{tr}\left(\frac{z(x)}{2}\right) \\
&= L_x \cdot L_z^{-1} \cdot \begin{bmatrix} 1 \\ 0 \end{bmatrix} \\
&= \begin{bmatrix} x_1 & x_2^T \\ x_2 & x_1 I \end{bmatrix} \cdot \frac{1}{2} \begin{bmatrix} \dfrac{1}{\sqrt{\lambda_2}} + \dfrac{1}{\sqrt{\lambda_1}} \\ \left(\dfrac{1}{\sqrt{\lambda_2}} - \dfrac{1}{\sqrt{\lambda_1}}\right) \dfrac{w_2}{\|w_2\|} \end{bmatrix} \\
&= \frac{1}{2} \begin{bmatrix} x_1 \left(\dfrac{1}{\sqrt{\lambda_2}} + \dfrac{1}{\sqrt{\lambda_1}}\right) + \left(\dfrac{1}{\sqrt{\lambda_2}} - \dfrac{1}{\sqrt{\lambda_1}}\right) \dfrac{x_2^T w_2}{\|w_2\|} \\ \left(\dfrac{1}{\sqrt{\lambda_2}} + \dfrac{1}{\sqrt{\lambda_1}}\right) x_2 + x_1 \left(\dfrac{1}{\sqrt{\lambda_2}} - \dfrac{1}{\sqrt{\lambda_1}}\right) \dfrac{w_2}{\|w_2\|} \end{bmatrix} \\
&= \frac{1}{2} \begin{bmatrix} x_1 \left(\dfrac{1}{\sqrt{\lambda_2}} + \dfrac{1}{\sqrt{\lambda_1}}\right) + \left(\dfrac{1}{\sqrt{\lambda_2}} - \dfrac{1}{\sqrt{\lambda_1}}\right) \|x_2\| \\ \left(\dfrac{1}{\sqrt{\lambda_2}} + \dfrac{1}{\sqrt{\lambda_1}}\right) x_2 + x_1 \left(\dfrac{1}{\sqrt{\lambda_2}} - \dfrac{1}{\sqrt{\lambda_1}}\right) \dfrac{x_2}{\|x_2\|} \end{bmatrix} \\
&= \frac{1}{2} \begin{bmatrix} \dfrac{x_1 + \|x_2\|}{\sqrt{\lambda_2}} + \dfrac{x_1 - \|x_2\|}{\sqrt{\lambda_1}} \\ \left(\dfrac{x_1 + \|x_2\|}{\sqrt{\lambda_2}} - \dfrac{x_1 - \|x_2\|}{\sqrt{\lambda_1}}\right) \dfrac{x_2}{\|x_2\|} \end{bmatrix}.
\end{aligned}
$$

Using (9), we can compute

$$
\begin{aligned}
\frac{\partial^2 f_7}{\partial x_1^2} &= \frac{\sqrt{\lambda_2} - (x_1 + \|x_2\|) \cdot \frac{x_1 + \|x_2\|}{\sqrt{\lambda_2}}}{\lambda_2} + \frac{\sqrt{\lambda_1} - (x_1 - \|x_2\|) \cdot \frac{x_1 - \|x_2\|}{\sqrt{\lambda_1}}}{\lambda_1} \\
&= \frac{4}{\lambda_2 \sqrt{\lambda_2}} + \frac{4}{\lambda_1 \sqrt{\lambda_1}}, \\
\frac{\partial^2 f_7}{\partial x_1 \partial x_2} &= \left(\frac{\sqrt{\lambda_2} - (x_1 + \|x_2\|) \cdot \frac{x_1 + \|x_2\|}{\sqrt{\lambda_2}}}{\lambda_2} - \frac{\sqrt{\lambda_1} - (x_1 - \|x_2\|) \cdot \frac{x_1 - \|x_2\|}{\sqrt{\lambda_1}}}{\lambda_1} \right) \frac{x_2^T}{\|x_2\|} \\
&= \left(\frac{4}{\lambda_2 \sqrt{\lambda_2}} - \frac{4}{\lambda_1 \sqrt{\lambda_1}} \right) \frac{x_2^T}{\|x_2\|}, \\
\frac{\partial^2 f_7}{\partial x_2 \partial x_1} &= \left(\frac{4}{\lambda_2 \sqrt{\lambda_2}} - \frac{4}{\lambda_1 \sqrt{\lambda_1}} \right) \frac{x_2}{\|x_2\|}, \\
\frac{\partial^2 f_7}{\partial x_2^2} &= \left(\frac{\sqrt{\lambda_2} - (x_1 + \|x_2\|) \cdot \frac{x_1 + \|x_2\|}{\sqrt{\lambda_2}}}{\lambda_2} + \frac{\sqrt{\lambda_1} - (x_1 - \|x_2\|) \cdot \frac{x_1 - \|x_2\|}{\sqrt{\lambda_1}}}{\lambda_1} \right) \frac{x_2 x_2^T}{\|x_2\|^2} \\
&\quad + \left(\frac{x_1 + \|x_2\|}{\sqrt{\lambda_2}} - \frac{x_1 - \|x_2\|}{\sqrt{\lambda_1}} \right) \cdot \frac{1}{\|x_2\|^2} \left(\|x_2\| I - \frac{1}{\|x_2\|} x_2 x_2^T \right) \\
&= \frac{1}{\|x_2\|} \left(\frac{x_1 + \|x_2\|}{\sqrt{\lambda_2}} - \frac{x_1 - \|x_2\|}{\sqrt{\lambda_1}} \right) I + \left(\frac{4}{\lambda_2 \sqrt{\lambda_2}} + \frac{4}{\lambda_1 \sqrt{\lambda_1}} \right. \\
&\quad \left. - \frac{1}{\|x_2\|} \left(\frac{x_1 + \|x_2\|}{\sqrt{\lambda_2}} - \frac{x_1 - \|x_2\|}{\sqrt{\lambda_1}} \right) \right) \frac{x_2 x_2^T}{\|x_2\|^2}.
\end{aligned}
$$

Therefore,

$$
\nabla^2 f_7(x) = \begin{bmatrix} a & b \frac{x_2^T}{\|x_2\|} \\ b \frac{x_2}{\|x_2\|} & cI + (a - c) \frac{x_2 x_2^T}{\|x_2\|^2} \end{bmatrix},
$$

with

$$
\begin{aligned}
a &= \frac{4}{\lambda_2 \sqrt{\lambda_2}} + \frac{4}{\lambda_1 \sqrt{\lambda_1}}, \\
b &= \frac{4}{\lambda_2 \sqrt{\lambda_2}} - \frac{4}{\lambda_1 \sqrt{\lambda_1}}, \\
c &= \frac{1}{\|x_2\|} \left(\frac{x_1 + \|x_2\|}{\sqrt{\lambda_2}} - \frac{x_1 - \|x_2\|}{\sqrt{\lambda_1}} \right).
\end{aligned}
$$

From Lemma 3.4, it suffices to show that $a > 0, c > 0$ and $a^2 - b^2 > 0$. Note that $a > 0$ is obviously, now we proof $c > 0$ and $a^2 - b^2 > 0$. From direct computation, we have

$$\begin{aligned} c &= \frac{1}{\|x_2\|}\left(\frac{x_1 + \|x_2\|}{\sqrt{\lambda_2}} - \frac{x_1 - \|x_2\|}{\sqrt{\lambda_1}}\right) \\ &= \frac{\sqrt{\lambda_1}(x_1 + \|x_2\|) - \sqrt{\lambda_2}(x_1 - \|x_2\|)}{\|x_2\| \cdot \sqrt{\lambda_1 \lambda_2}} \\ &> 0, \end{aligned}$$

where the last inequality is because

$$\left(\sqrt{\lambda_1}(x_1 + \|x_2\|)\right)^2 - \left(\sqrt{\lambda_2}(x_1 - \|x_2\|)\right)^2 = 4\left((x_1 + \|x_2\|)^2 - (x_1 - \|x_2\|)^2\right) > 0,$$

and

$$a^2 - b^2 = (a+b)(a-b) = \frac{64}{\lambda_1 \lambda_2 \sqrt{\lambda_1 \lambda_2}} > 0.$$

Thus M is positive definite and $\nabla^2 f_7(x)$ is positive definite for all $x \in \mathbb{R}^n$.

Case(2): $w_2 = 0$. Since $w_2 = 2x_1x_2 = 0$, we consider following two subcases:

Subcase(i): $x_2 = 0$. From Lemma 3.6 and direct computation, we have

$$\begin{aligned} \nabla f_7(x) &= \nabla z(x) \cdot \nabla \mathrm{tr}\left(\frac{z(x)}{2}\right) \\ &= L_x \cdot L_z^{-1} \cdot \begin{bmatrix} 1 \\ 0 \end{bmatrix} \\ &= \frac{1}{\sqrt{w_1}} \begin{bmatrix} x_1 & x_2^T \\ x_2 & x_1 I \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 0 \end{bmatrix} \\ &= \frac{1}{\sqrt{x_1^2 + 4}} \begin{bmatrix} x_1 \\ 0 \end{bmatrix}, \end{aligned}$$

and

$$\nabla^2 f_7(x) = \begin{bmatrix} \dfrac{4}{(x_1^2+4)^{3/2}} & 0^T \\ 0 & O \end{bmatrix}.$$

Thus $\nabla^2 f_7(x)$ is positive semidefinite for all $x \in \mathbb{R}^n$.

Subcase(ii): $x_1 = 0$. From Lemma 3.6 and direct computation, we have

$$\begin{aligned} \nabla f_7(x) &= \nabla z(x) \cdot \nabla \mathrm{tr}\left(\frac{z(x)}{2}\right) \\ &= L_x \cdot L_z^{-1} \cdot \begin{bmatrix} 1 \\ 0 \end{bmatrix} \\ &= \frac{1}{\sqrt{w_1}} \begin{bmatrix} x_1 & x_2^T \\ x_2 & x_1 I \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 0 \end{bmatrix} \\ &= \frac{1}{\sqrt{\|x_2\|^2 + 4}} \begin{bmatrix} 0 \\ x_2 \end{bmatrix}, \end{aligned}$$

and

$$\nabla^2 f_7(x) = \begin{bmatrix} 0 & 0^T \\ 0 & \dfrac{1}{\sqrt{\|x_2\|^2+4}} I - \dfrac{1}{(\|x_2\|^2+4)^{3/2}} x_2 x_2^T \end{bmatrix}.$$

Thus $\nabla^2 f_7(x)$ is positive semidefinite for all $x \in \mathbb{R}^n$ and $f_7(x)$ is convex for all $x \in \mathbb{R}^n$. $\square$

4 Equalities and Inequalities associated with SOC

Property 4.1 *Let $A, B \in \mathcal{S}^n$.*

(a) *If $B \succeq O$, then $\lambda_i(A) \le \lambda_i(A+B)$ for all $i = 1, 2, \cdots, n$.*

(b) *$\lambda_i(A) + \lambda_{\min}(B) \le \lambda_i(A+B) \le \lambda_i(A) + \lambda_{\max}(B)$ for all $i = 1, 2, \cdots, n$.*

(c) *If $A \succeq O$, $B \succeq O$, then $\sum_{i=1}^n \lambda_i(A)\lambda_{n-i+1}(B) \le \sum_{i=1}^n \lambda_i(AB) \le \sum_{i=1}^n \lambda_i(A)\lambda_i(B)$ for all $i = 1, 2, \cdots, n$.*

(d) *If $A \succ O$, $B \succ O$, then $\dfrac{\lambda_i^2(AB)}{\lambda_{\max}(A)\lambda_{\max}(B)} \le \lambda_i(A)\lambda_i(B) \le \dfrac{\lambda_i^2(AB)}{\lambda_{\min}(A)\lambda_{\min}(B)}$ for all $i = 1, 2, \cdots, n$.*

(e) *If $\lambda_i(A)$ and $\lambda_i(B)$ are both arranged in increasing or decreasing order, then*

$$\left(\sum_{i=1}^n (\lambda_i(A) - \lambda_i(B))^2 \right)^{1/2} \le \|A - B\|_F.$$

Proof. These are all well-known results in matrix analysis, see [4, 13, 18]. In particular, part(b) is known as Weyl's Theorem. $\square$

Proposition 4.1 *Let $x = (x_1, x_2), y = (y_1, y_2) \in \mathbb{R} \times \mathbb{R}^{n-1}$.*

(a) *If $y \succeq_{\mathcal{K}^n} 0$, then $\lambda_i(x) \le \lambda_i(x+y)$ for all $i = 1, 2$.*

(b) *$\lambda_i(x) + \lambda_1(y) \le \lambda_i(x+y) \le \lambda_i(x) + \lambda_2(y)$, for all $i = 1, 2$.*

(c) *If $x \succeq_{\mathcal{K}^n} 0$, $y \succeq_{\mathcal{K}^n} 0$, then $\lambda_1(x)\lambda_2(y) + \lambda_2(x)\lambda_1(y) \le \mathrm{tr}(x \circ y) \le \lambda_1(x)\lambda_1(y) + \lambda_2(x)\lambda_2(y)$.*

(d) *When $n = 2$, if $x \succ_{\mathcal{K}^n} 0$, $y \succ_{\mathcal{K}^n} 0$, then $\dfrac{\lambda_i^2(x \circ y)}{\lambda_2(x)\lambda_2(y)} \le \lambda_i(x)\lambda_i(y) \le \dfrac{\lambda_i^2(x \circ y)}{\lambda_1(x)\lambda_1(y)}$ for all $i = 1, 2$.*

(e) *If* $\lambda_i(x)$ *and* $\lambda_i(y)$ *are both arranged in increasing or decreasing order, then*

$$\left(\sum_{i=1}^{2}(\lambda_i(x)-\lambda_i(y))^2\right)^{1/2} \le \|x-y\|_{\mathbb{V}}.$$

Proof. (a) First, we prove $\lambda_1(x+y) \ge \lambda_1(x)$. From direct computation and triangle's inequality, we have

$$\begin{aligned}
\lambda_1(x+y)-\lambda_1(x) &= (x_1+y_1-\|x_2+y_2\|)-(x_1-\|x_2\|)\\
&= y_1-\|x_2+y_2\|+\|x_2\|\\
&\ge y_1-(\|x_2\|+\|y_2\|)+\|x_2\|\\
&= y_1-\|y_2\|\\
&\ge 0,
\end{aligned}$$

where the last inequality is because of $y \succeq_{\mathcal{K}^n} 0$. Second, we prove $\lambda_2(x+y) \ge \lambda_2(x)$. From direct computation and triangle's inequality, we have

$$\begin{aligned}
\lambda_2(x+y)-\lambda_2(x) &= (x_1+y_1+\|x_2+y_2\|)-(x_1+\|x_2\|)\\
&= y_1+\|x_2+y_2\|-\|x_2\|\\
&\ge y_1+(\|x_2\|-\|y_2\|)-\|x_2\|\\
&= y_1-\|y_2\|\\
&\ge 0.
\end{aligned}$$

(b) See [1, Prop. 3.1].
(c) See [7, Prop. 2.3].
(d) We prove the inequality by separating to four parts. First, we prove $\dfrac{\lambda_1^2(x\circ y)}{\lambda_2(x)\lambda_2(y)} \le \lambda_1(x)\lambda_1(y)$. Since $x \succ_{\mathcal{K}^n} 0$, $y \succ_{\mathcal{K}^n} 0$, we have $\lambda_1(x)=x_1-|x_2|>0$, $\lambda_2(x)=x_1+|x_2|>0$, $\lambda_1(y)=y_1-|y_2|>0$ and $\lambda_2(y)=y_1+|y_2|>0$. Thus,

$$\frac{\lambda_1^2(x\circ y)}{\lambda_2(x)\lambda_2(y)} \le \lambda_1(x)\lambda_1(y) \Longleftrightarrow \lambda_1^2(x\circ y) \le \det(x)\det(y).$$

From direct computation, we have

$$\begin{aligned}
&\det(x)\det(y)-\lambda_1^2(x\circ y)\\
&= (x_1^2-|x_2|^2)\cdot(y_1^2-|y_2|^2)-(x_1y_1+x_2y_2-|x_1y_2+y_1x_2|)^2\\
&= 2(x_1y_1+x_2y_2)\cdot|x_1y_2+y_1x_2|\\
&\ge 0,
\end{aligned}$$

where the last inequality is due to $x_1y_1 > |x_2y_2| > -x_2y_2$. Second, we prove $\dfrac{\lambda_2^2(x\circ y)}{\lambda_2(x)\lambda_2(y)} \le \lambda_2(x)\lambda_2(y)$. Since $\lambda_2(x\circ y)=x_1y_1+x_2y_2+|x_1y_2+y_1x_2|>0$, we have

$$\frac{\lambda_2^2(x\circ y)}{\lambda_2(x)\lambda_2(y)} \le \lambda_2(x)\lambda_2(y) \Longleftrightarrow \lambda_2(x\circ y) \le \lambda_2(x)\lambda_2(y).$$

From direct computation, we have

$$\begin{aligned}
&\lambda_2(x)\lambda_2(y) - \lambda_2(x \circ y) \\
&= (x_1 + |x_2|) \cdot (y_1 + |y_2|) - (x_1 y_1 + x_2 y_2 + |x_1 y_2 + y_1 x_2|) \\
&= x_1 |y_2| + y_1 |x_2| - |x_1 y_2 + y_1 x_2| + |x_2 y_2| - x_2 y_2 \\
&\geq 0.
\end{aligned}$$

Third, we prove $\lambda_2(x)\lambda_2(y) \leq \dfrac{\lambda_2^2(x \circ y)}{\lambda_1(x)\lambda_1(y)}$. Since $\lambda_1(x) > 0$ and $\lambda_1(y) > 0$,

$$\lambda_2(x)\lambda_2(y) \leq \frac{\lambda_2^2(x \circ y)}{\lambda_1(x)\lambda_1(y)} \Longleftrightarrow \det(x)\det(y) \leq \lambda_2^2(x \circ y).$$

From direct computation, we have

$$\begin{aligned}
&\lambda_2^2(x \circ y) - \det(x)\det(y) \\
&= (x_1 y_1 + x_2 y_2 + |x_1 y_2 + y_1 x_2|)^2 - (x_1^2 - |x_2|^2) \cdot (y_1^2 - |y_2|^2) \\
&= 2\left((x_1 y_2 + y_1 x_2)^2 + (x_1 y_1 + x_2 y_2) \cdot |x_1 y_2 + y_1 x_2|\right) \\
&\geq 0.
\end{aligned}$$

Final, we prove $\lambda_1(x)\lambda_1(y) \leq \dfrac{\lambda_1^2(x \circ y)}{\lambda_1(x)\lambda_1(y)}$. Since $\lambda_1(x \circ y) = x_1 y_1 + x_2 y_2 - |x_1 y_2 + y_1 x_2|$ and $(x_1 y_1 + x_2 y_2)^2 - |x_1 y_2 + y_1 x_2|^2 = (x_1^2 - x_2^2)(y_1^2 - y_2^2) > 0$, therefore, $\lambda_1(x \circ y) > 0$ and

$$\lambda_1(x)\lambda_1(y) \leq \frac{\lambda_1^2(x \circ y)}{\lambda_1(x)\lambda_1(y)} \Longleftrightarrow \lambda_1(x)\lambda_1(y) \leq \lambda_1(x \circ y).$$

From direct computation, we have

$$\lambda_1(x \circ y) - \lambda_1(x)\lambda_1(y) = x_1|y_2| + y_1|x_2| - |x_1 y_2 + y_1 x_2| + x_2 y_2 - |x_2 y_2|. \tag{10}$$

Here we consider two cases:
Case(1): $x_2 y_2 \geq 0$. Obviously, (10) is equal to 0, which is our desired result.
Case(2): $x_2 y_2 < 0$. Without loss of generality, we assume $x_2 > 0, y_2 < 0$. If $x_1 y_2 + y_1 x_2 \geq 0$, from (10) we have

$$\begin{aligned}
\lambda_1(x \circ y) - \lambda_1(x)\lambda_1(y) &= -x_1 y_2 + y_1 x_2 - (x_1 y_2 + y_1 x_2) + x_2 y_2 + x_2 y_2 \\
&= 2y_2(x_2 - x_1) \\
&\geq 0.
\end{aligned}$$

Otherwise,

$$\begin{aligned}
\lambda_1(x \circ y) - \lambda_1(x)\lambda_1(y) &= -x_1 y_2 + y_1 x_2 + (x_1 y_2 + y_1 x_2) + x_2 y_2 + x_2 y_2 \\
&= 2x_2(y_2 + y_1) \\
&\geq 0.
\end{aligned}$$

Hence we have $\lambda_1(x\circ y) - \lambda_1(x)\lambda_1(y) \geq 0$.
(e) To see this, we calculate

$$\begin{aligned}
\sum_{i=1}^{2}(\lambda_i(x)-\lambda_i(y))^2 &= \lambda_1^2(x)+\lambda_2^2(x)+\lambda_1^2(y)+\lambda_2^2(y)-2(\lambda_1(x)\lambda_1(y)+\lambda_2(x)\lambda_2(y)) \\
&= 2(x_1^2+\|x_2\|^2)+2(y_1^2+\|y_2\|^2)-4(x_1y_1+\|x_2\|\|y_2\|) \\
&\leq 2(x_1^2+\|x_2\|^2)+2(y_1^2+\|y_2\|^2)-4(x_1y_1+\langle x_2,y_2\rangle) \\
&= 2\left(\|x\|^2+\|y\|^2-2\langle x,y\rangle\right) \\
&= \|x-y\|_{\mathbb{V}}^2,
\end{aligned}$$

where the inequality is due to Cauchy's inequality. Thus, the proof is complete. □

Remark: Unlike Property 4.1(d) for matrix case, Proposition 4.1(d) does not hold for general $n \geq 3$. We give a counterexample below.

Example 4.1 Let $x = \begin{pmatrix} 3 \\ -1 \\ 2 \end{pmatrix}$ and $y = \begin{pmatrix} 4 \\ 1 \\ 1 \end{pmatrix}$, then $x \succ_{\mathcal{K}^n} 0$, $y \succ_{\mathcal{K}^n} 0$. From direct computation, we have $x\circ y = \begin{pmatrix} 13 \\ -1 \\ 11 \end{pmatrix}$. It is easy to verify that $\lambda_1(x) = 3-\sqrt{5}$, $\lambda_2(x) = 3+\sqrt{5}$, $\lambda_1(y) = 4-\sqrt{2}$, $\lambda_2(y) = 4+\sqrt{2}$, $\lambda_1(x\circ y) = 13-\sqrt{122}$ and $\lambda_2(x\circ y) = 13+\sqrt{122}$. Therefore, $\dfrac{\lambda_1^2(x\circ y)}{\lambda_1(x)\lambda_1(y)} \doteqdot 1.93 < 1.97 \doteqdot \lambda_1(x)\lambda_1(y)$.

Property 4.2 **(a)** *If $A \succ O$, then* $\det(A) = \exp(\mathrm{tr}(\ln A))$.

(b) *If $A \succ O, B \succ O$, then* $\det(A+B)^{1/n} \geq \det(A)^{1/n} + \det(B)^{1/n}$ *for any $n \in \mathbb{N}$.*

(c) *If $A \succ O$, then* $\det(A)^{1/m} = \min\left\{\dfrac{\mathrm{tr}(AB)}{m} : B \succeq O \text{ and } \det(B) = 1\right\}$.

(d) *If $A \succ O, B \succeq O$, then* $\det(A+B) \geq \det(A)$ *with equality if and only if $B = O$.*

(e) *If $A \succ O, B \succ O$ and $A - B \succeq O$, then* $\det A \geq \det B$ *with equality if and only if $A = B$.*

Proof. Again, these are all well-known results in matrix analysis, see [4, 13, 18]. In addition, part(b) is known as Minkowski inequality. □

Proposition 4.2 **(a)** *If $x \succ_{\mathcal{K}^n} 0$, then* $\det(x) = \exp(\mathrm{tr}(\ln x))$.

(b) *If* $x \succ_{\mathcal{K}^n} 0, y \succ_{\mathcal{K}^n} 0$, *then* $\det(x+y)^{1/n} \geq \frac{4^{1/n}}{2}\left(\det(x)^{1/n} + \det(y)^{1/n}\right)$ *for any* $n \in \mathbb{N}, n \geq 2$.

(c) *If* $x \succ_{\mathcal{K}^n} 0$, *then* $\det(x)^{1/2} = \min\left\{\frac{\operatorname{tr}(x \circ y)}{2} : y \succeq_{\mathcal{K}^n} 0 \text{ and } \det(y) = 1\right\}$.

(d) *If* $x \succ_{\mathcal{K}^n} 0, y \succeq_{\mathcal{K}^n} 0$, *then* $\det(x+y) \geq \det(x)$ *with equality if and only if* $y = 0$.

(e) *If* $x \succ_{\mathcal{K}^n} 0, y \succ_{\mathcal{K}^n} 0$ *and* $x - y \succeq_{\mathcal{K}^n} 0$, *then* $\det(x) \geq \det(y)$ *with equality if and only if* $x = y$.

Proof. (a) From Lemma 3.5(b) and the definition of $\operatorname{tr}(x)$, we have $\det(x) = \exp(\operatorname{tr}(\ln x))$ for all $x \succ_{\mathcal{K}^n} 0$.
(b) See [12, Prop. 3.2].
(c) Let $x = (x_1, x_2, \cdots, x_n), y = (y_1, y_2, \cdots, y_n) \in \mathbb{R}^n$. Note that $\frac{\operatorname{tr}(x \circ y)}{2} = x_1y_1 + x_2y_2 + \cdots + x_ny_n$. Consider the minimization problem

$$\begin{array}{rl} \min: & x_1y_1 + x_2y_2 + \cdots + x_ny_n \\ \text{s.t.} & y_1 > 0 \\ & y_1^2 - (y_2^2 + \cdots + y_n^2) = 1. \end{array}$$

Use the method of Lagrange multiplier, we can change the minimization problem to system of equations:

$$\begin{cases} y_1 > 0 \\ x_1 = 2\lambda y_1 \\ x_2 = -2\lambda y_2 \\ \vdots \\ x_n = -2\lambda y_n \\ y_1^2 - (y_2^2 + \cdots + y_n^2) = 1 \end{cases} \tag{11}$$

Solve (11) by substitution, we have $\lambda = \frac{\sqrt{x_1^2 - (x_2^2 + \cdots + x_n^2)}}{2} = \frac{\det(x)^{1/2}}{2}$ and $y_1 = \frac{x_1}{\det(x)^{1/2}}, y_i = \frac{-x_i}{\det(x)^{1/2}}$ for $i = 2, \cdots, n$. Therefore, the optimal solution is $\det(x)^{1/2}$, which is the desired result.
(d) Let $x = (x_1, x_2), y = (y_1, y_2) \in \mathbb{R} \times \mathbb{R}^{n-1}$. Since $x \succ_{\mathcal{K}^n} 0, y \succeq_{\mathcal{K}^n} 0$, we have $x_1 > \|x_2\|, y_1 \geq \|y_2\|$ and

$$x_1y_1 \geq \|x_2\|\|y_2\| \geq |\langle x_2, y_2\rangle|. \tag{12}$$

Note that the first equality hold in (12) if and only if $y_1 = \|y_2\| = 0$, i.e. $y = 0$. Therefore,

$$\begin{aligned} \det(x+y) - \det(x) &= (x_1 + y_1)^2 - \|x_2 + y_2\|^2 - (x_1^2 - \|x_2\|^2) \\ &= 2(x_1y_1 - \langle x_2, y_2\rangle) \\ &\geq 0, \end{aligned}$$

where the last equality hold in if and only if $y = 0$.
(e) Since $x \succ_{\mathcal{K}^n} 0, y \succ_{\mathcal{K}^n} 0$ and $x - y \succeq_{\mathcal{K}^n} 0$, we have $x_1 > \|x_2\|, y_1 > \|y_2\|$ and

$$\begin{array}{l} x_1 + y_1 > \|x_2\| + \|y_2\|, \\ x_1 - y_1 \geq \|x_2 - y_2\| \geq \|x_2\| - \|y_2\|. \end{array} \tag{13}$$

Therefore,

$$\begin{aligned} \det(x) - \det(y) &= (x_1^2 - \|x_2\|^2) - (y_1^2 - \|y_2\|^2) \\ &= (x_1 + y_1)(x_1 - y_1) - (\|x_2\| + \|y_2\|)(\|x_2\| - \|y_2\|) \\ &\geq 0, \end{aligned}$$

where the last equality hold in if and only if $x = y$. $\square$

Remark: The inequality in Property 4.2(b) is the famous Minkowski inequality in matrix analysis. However, such inequality has slightly different form when it is extended to SOC case as shown in Proposition 4.2(b). More specifically, it is not true that

$$\det(x+y)^{1/n} \geq \left(\det(x)^{1/n} + \det(y)^{1/n}\right) \ \forall x \succ_{\mathcal{K}^n} 0, y \succ_{\mathcal{K}^n} 0.$$

On the other hand, the formula in Property 4.2(c) does not hold for general $m \neq 2$, when it is considered in SOC case as seen in Proposition 4.2(c). In fact, when $m = 2$, we have the minimization problem

$$\begin{array}{rl} \min: & \dfrac{2}{m}(x_1y_1 + x_2y_2 + \cdots + x_ny_n) \\ \text{s.t.} & y_1 > 0 \\ & y_1^2 - (y_2^2 + \cdots + y_n^2) = 1. \end{array}$$

Using same method, we can get the optimal solution $\dfrac{2}{m}\det(x)^{1/2} \neq \det(x)^{1/m}$.

5 Final Remarks

In this paper, we have investigated: to what extent are positive semidefinite cone and second-order cone like? We show that they share many similarities, but still have some differences. We believe that such study will be helpful for designing solution methods for SDP and SOCP. There have some interesting directions to be explored along this topic. It is well known that the adjoint X^* of a symmetric matrix X plays an important role in matrix analysis. In fact, there are many matrix inequalities and matrix equations which involve X^*. What is the corresponding role of x^* (adjoint of x) in SOC case? On the other hand, two matrices A and B are said to be similar if there exists an invertible matrix S such that $A = S^{-1}BS$. Such concept is crucial in the classification of matrices. Can we define analogous concept of similarity of two vectors x and y associated with SOC? We leave these as future research topics.

References

[1] A. Auslender, *Variational inequalities over the cone of semidefinete positive symmetric matrices and over the lorentz cone*, Optimization Methods and Software, vol. 18, pp. 359-376, 2003.

[2] A. Auslender, *Penalty and barrier methods: a unified framework*, SIAM Journal on Optimization, vol. 10, pp.211-230, 1999.

[3] M. S. Bazaraa, H. D. Sherali and C. M. Shetty, *Nonlinear Programming*, John Wiley and Sons, 3rd edition, 2006.

[4] R. Bhatia, *Matrix Analysis*, Springer-Verlag, New York, 1997.

[5] J.-S. Chen, X. Chen and P. Tseng, *Analysis of nonsmooth vector-valued functions associated with second-oredr cone*, Mathmatical Programming, vol.101, pp. 95-117, 2004.

[6] J.-S. Chen and P. Tseng, *An unconstrained smooth minimization reformulation of the second-order cone complementarity problem*, Mathematical Programming, vol. 104, pp. 293-327, 2005.

[7] J.-S. Chen, *The convex and monotone functions associated with second-order cone*, Optimization, vol. 55, pp. 363-385, 2006.

[8] A. Fischer *A special Newton-type optimization methods*, Optimization, vol. 24, pp. 269-284, 1992.

[9] A. Fischer *Solution of the monotone complementarity problem with locally Lipschitzian functions*, Mathematical Programming, vol. 76, pp.513-532, 1997.

[10] J. Faraut and A. Korányi, *Analysis on symmetric Cones*, Oxford Mathematical Monographs, Oxford University Press, New York, 1994.

[11] M. Fukushima, Z.-Q Luo and P. Tseng, *Smoothing functions for second-order cone complementarity problems*, SIAM Journal on Optimazation, vol. 12, pp. 436-460, 2002.

[12] M.-L. Hsu, *Some results on functions associated with second-order cone*, Master thesis, National Taiwan Normal University, 2007.

[13] R. Horn and C. Johnson, *Matrix Analysis*, Cambridge University Press, 1985.

[14] C. Kanzow and M. Fukushima, *Equivalence of the generalized complementarity problem to differentiable unconstrained minimization*, Journal of Optimization Theory and Applications, vol. 90, pp. 581-603, 1996.

[15] O. L. Mangasarian and M. V. Solodov, *Nonlinear complementarity as unconstrained and constrained minimization*, Mathematical Programming, vol. 62, pp. 277-297, 1993.

[16] S.-H. Pan and J.-S. Chen, *A damped Gauss-Newton method for the second-order cone complementarity problem*, Mathematics and Optimization, vol. 59, pp. 293-318, 2009.

[17] S.-H. Pan and J.-S. Chen, *Growth behavior of two classes of merit functions for symmetric cone complementarity problems*, Journal of Optimization Theory and Applications , vol. 141, pp. 167-191, 2009.

[18] J. Schott, *Matrix Analysis for Statistics*, John Wiley and Sons, 2nd edition, 2005.

[19] C-K. Sim and G. Zhao, *A note on treating a second-order cone program as a special case of a semidefinite program*, Mathematical Programming, vol. 102, pp. 609-613, 2005.