Journal of Nonlinear and Convex Analysis
Volume 23, Number 2, 2022, 241–250

Yokohama Publishers
ISSN 1880-5221 ONLINE JOURNAL

A SEMI-DISTANCE AND PROXIMAL DISTANCE ASSOCIATED WITH SYMMETRIC CONE

XINHE MIAO* AND JEIN-SHAN CHEN†

Abstract. Recently, there was a distance function $\mathbf{d}$ studied in [16], which was shown a semi-distance in the setting of symmetric cone. This result was indeed verified under the assumption of operator commute. It was further proved that it could become a proximal distance in the setting of second-order cone. In this short paper, we improve these results by showing that the distance function $\mathbf{d}$ studied in [16] is not only a semi-distance, but also a proximal distance in the setting of symmetric cone without assuming operator commute.

1. Motivation and background materials

We are interested in constructing distance-like functions, which can be used in proximal-like algorithm [4,6,7,13,17,18] for nonlinear symmetric cone programming (SCP):

$$\begin{array}{ll} \min & f(x) \\ \text{s.t.} & x \in \mathcal{K}, \end{array} \tag{1.1}$$

where $\mathbb{V}$ is a Euclidean Jordan algebra, $\mathcal{K} \subset \mathbb{V}$ denotes a symmetric cone, $f : \mathbb{V} \to \mathbb{R} \cup \{+\infty\}$ is a proper lower semicontinuous (l.s.c. in short) convex function. That is, the problem (1.1) is a convex programming problem [19].

A popular approach to deal with SCP is the proximal point algorithm, which generates a sequence $\{x^k\}$ via the following iterative scheme:

$$x^{k+1} = \arg\min_{x \in \mathcal{K}} \{f(x) + \lambda_k D(x, x^k)\}. \tag{1.2}$$

Here $D(\cdot,\cdot)$ is a certain function satisfying some desirable properties and $\{\lambda_k\}_{k\in\mathbb{N}}$ a positive sequence. The choice of $D(\cdot,\cdot)$ is important and examples of $D(\cdot,\cdot)$ are the distances induced by Euclidean norm, quasi-distance, Bregman distance, φ-divergence, and proximal distance, etc., see [2,3,8,9,15,20,21,23]. The below chart shows the relationship among these distances.

A *Euclidean Jordan algebra* [5, 10] is a finite dimensional inner product space $(\mathbb{V}, \langle\cdot,\cdot\rangle)$ ($\mathbb{V}$ for short) over the field of real numbers $\mathbb{R}$ equipped with a bilinear map $(x, y) \mapsto x \circ y : \mathbb{V} \times \mathbb{V} \to \mathbb{V}$, which satisfies the following conditions:

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

2020 *Mathematics Subject Classification.* 26B05, 26B35, 90C33.
Key words and phrases. Symmetric cone, semi-distance, proximal distance, quadratic representation.
*The author's work is supported by National Natural Science Foundation of China (No. 11471241).
†Corresponding author. The author's work is supported by Ministry of Science and Technology, Taiwan.

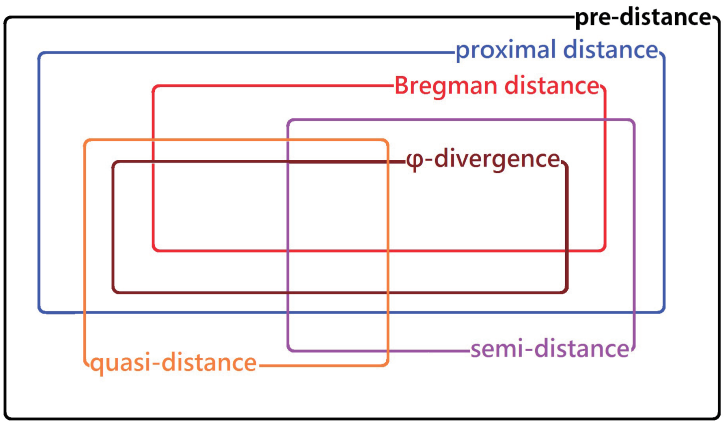

FIGURE 1. Relationship among various distances.

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

Here $x \circ y$ is called the *Jordan product* of x and y. If a Jordan product only satisfies the conditions (i) and (ii) in the above definition, the algebra $\mathbb{V}$ is said to be a *Jordan algebra*. Moreover, if there is an (unique) element $e \in \mathbb{V}$ such that $x \circ e = x$ for all $x \in \mathbb{V}$, the element e is called the *identity element* in $\mathbb{V}$. Note that a Jordan algebra does not necessarily have an identity element. Throughout this paper, we assume that $\mathbb{V}$ is a Euclidean Jordan algebra with an identity element e.

In a given Euclidean Jordan algebra $\mathbb{V}$, the set of squares $\mathcal{K} := \{x^2 \,|\, x \in \mathbb{V}\}$ is a *symmetric cone* [10, Theorem III.2.1]. This means that $\mathcal{K}$ is a self-dual closed convex cone and, for any two elements $x, y \in \mathrm{int}(\mathcal{K})$, there exists an invertible linear transformation $\Gamma : \mathbb{V} \to \mathbb{V}$ such that $\Gamma(x) = y$ and $\Gamma(\mathcal{K}) = \mathcal{K}$. It is well known that second-order cone is a special symmetric cone, which is defined as follows in $\mathbb{R}^n$:

$$\mathcal{K}^n := \left\{x = (x_0, \bar{x}) \in \mathbb{R} \times \mathbb{R}^{n-1} \mid x_0 \geq \|\bar{x}\|\right\},$$

and the corresponding Jordan product of x and y in $\mathbb{R}^n$ with $x = (x_0, \bar{x})$, $y = (y_0, \bar{y}) \in \mathbb{R} \times \mathbb{R}^{n-1}$ is given by

$$x \circ y := \begin{bmatrix} x^T y \\ x_0\bar{y} + y_0\bar{x} \end{bmatrix}.$$

In particular, in the setting of the second-order cone $\mathcal{K}^n$, the identity element $e = (1, 0) \in \mathbb{R} \times \mathbb{R}^{n-1}$, where 0 denotes the zero vector in $\mathbb{R}^{n-1}$.

For $x \in \mathbb{V}$, we denote $m(x)$ the *degree* of the minimal polynomial of x, that is,

$$m(x) := \min\left\{k > 0 \,|\, \{e, x, \ldots, x^k\} \text{ is linearly dependent}\right\},$$

and the *rank* of $\mathbb{V}$ is well-defined by $r := \max\{m(x) \,|\, x \in \mathbb{V}\}$. In Euclidean Jordan algebra $\mathbb{V}$, an element $e^i \in \mathbb{V}$ is an *idempotent* if $(e^i)^2 = e^i$, and it is a *primitive idempotent* if it is nonzero and cannot be written as a sum of two nonzero idempotents. The idempotents e^i and e^j are said to be *orthogonal* if $e^i \circ e^j = 0$. In addition, we say that a finite set $\{e^1, e^2, \ldots, e^r\}$ of primitive idempotents in $\mathbb{V}$ is a *Jordan frame* if

$$e^i \circ e^j = 0 \quad \text{for} \quad i \neq j, \quad \text{and} \quad \sum_{i=1}^{r} e^i = e.$$

Note that $\langle e^i, e^j \rangle = \langle e^i \circ e^j, e \rangle$ whenever $i \neq j$.

With the above, there have the spectral decomposition and Peirce decomposition of an element x in $\mathbb{V}$.

Theorem 1.1 ((The Spectral Decomposition Theorem) [10, Theorem III.1.2])**.** *Let $\mathbb{V}$ be a Euclidean Jordan algebra. Then, there is a number r such that, for every $x \in \mathbb{V}$, there exists a Jordan frame $\{e^1, \ldots, e^r\}$ and real numbers $\lambda_1(x), \ldots, \lambda_r(x)$ with*

$$x = \lambda_1(x)e^1 + \cdots + \lambda_r(x)e^r.$$

Here, the numbers $\lambda_i(x)$ $(i = 1, \ldots, r)$ are the eigenvalues of x, the expression $\lambda_1(x)e^1 + \cdots + \lambda_r(x)e^r$ is the spectral decomposition of x. Moreover, $\mathrm{tr}(x) := \sum_{i=1}^{r} \lambda_i(x)$ *is called the trace of x, and* $\det(x) := \lambda_1(x) \ldots \lambda_r(x)$.

We point out that different elements x, y have their own Jordan frames in the spectral decomposition, which are not easy to handle when we need to do operations for x and y. Thus, we need another so-called Peirce decomposition to conquer such difficulty. In other words, in the Peirce decomposition, two different elements x, y share the same Jordan frame. We elaborate them more as below.

The Peirce decomposition: Fix a Jordan frame $\{e^1, e^2, \ldots, e^r\}$ in a Euclidean Jordan algebra $\mathbb{V}$. For $i, j \in \{1, 2, \ldots, r\}$, we define the following eigen-spaces

$$\mathbb{V}_{ii} := \{x \in \mathbb{V} \,|\, x \circ e^i = x\} = \mathbb{R}e^i$$

and

$$\mathbb{V}_{ij} := \left\{ x \in \mathbb{V} \,\middle|\, x \circ e^i = \frac{1}{2}x = x \circ e^j \right\} \quad \text{for} \quad i \neq j.$$

Theorem 1.2 ([10, Theorem IV.2.1])**.** *The space $\mathbb{V}$ is the orthogonal direct sum of spaces $\mathbb{V}_{ij} (i \leq j)$. Furthermore,*

$$\begin{aligned}
&\mathbb{V}_{ij} \circ \mathbb{V}_{ij} \subset \mathbb{V}_{ii} + \mathbb{V}_{jj}, \\
&\mathbb{V}_{ij} \circ \mathbb{V}_{jk} \subset \mathbb{V}_{ik}, \quad \text{if} \quad i \neq k, \\
&\mathbb{V}_{ij} \circ \mathbb{V}_{kl} = \{0\}, \quad \text{if} \quad \{i, j\} \cap \{k, l\} = \emptyset.
\end{aligned}$$

Hence, given any Jordan frame $\{e^1, e^2, \ldots, e^r\}$, we can write any element $x \in \mathbb{V}$ as

$$x = \sum_{i=1}^{r} x_i e^i + \sum_{i<j} x_{ij},$$

where $x_i \in \mathbb{R}$ and $x_{ij} \in \mathbb{V}_{ij}$. The expression $\sum_{i=1}^{r} x_i e^i + \sum_{i<j} x_{ij}$ is called the Peirce decomposition of x.

Theorem 1.3 ([22, Theorem 4.6]). *Suppose that $\mathbb{V}$ is simple and $\{e^1, \dots, e^r\}$ is any fixed Jordan frame in $\mathbb{V}$. Let $z = \sum_{i=1}^r z_i e^i + \sum_{i<j} z_{ij} \in \mathcal{K}$. Then, we have*

$$\sum_{i=1}^r z_i^p \le \operatorname{tr}(z^p) \quad \text{for} \quad p > 1 \quad \text{and} \quad \sum_{i=1}^r z_i^p \ge \operatorname{tr}(z^p) \quad \text{for} \quad 0 < p < 1,$$

where the equalities hold if and only if $z = \sum_{i=1}^r z_i e^i$.

In a Euclidean Jordan algebras $\mathbb{V}$, for any $x \in \mathbb{V}$, a linear transformation $L(x) : \mathbb{V} \to \mathbb{V}$ is called *Lyapunov transformation*, which is defined as $L(x)(y) := x \circ y$ for all $y \in \mathbb{V}$. The so-called *quadratic representation* $P(x)$ is define by $P(x) := 2L^2(x) - L(x^2)$. For any $x \in \mathbb{V}$, the endomorphisms $L(x)$ and $P(x)$ are self-adjoint. We say that two elements x and y of a Euclidean Jordan algebra $\mathbb{V}$ *operator commute* if $x \circ (y \circ z) = y \circ (x \circ z)$ for all $z \in \mathbb{V}$, which is equivalent to stating that $L(x)L(y) = L(y)L(x)$. For the quadratic representation $P(x)$, if x is invertible, then we have

$$P(x)\mathcal{K} = \mathcal{K} \quad \text{and} \quad P(x)\operatorname{int}(\mathcal{K}) = \operatorname{int}(\mathcal{K}).$$

Below is a useful property regarding the quadratic representation $P(x)$, which is needed for our subsequent analysis.

Theorem 1.4 ([12, Proposition 2.5]). *Suppose that $\{e^1, e^2, \dots, e^r\}$ is Jordan frame in $\mathbb{V}$ and the spectral decomposition of x can be expressed as $x = \lambda_1(x)e^1 + \cdots + \lambda_r(x)e^r$. For any $z \in \mathbb{V}$, if the Peirce decomposition of z is $z = \sum_{i=1}^r z_i e^i + \sum_{i<j} z_{ij}$, we have*

$$P(x)z = \sum_{i=1}^r \lambda_i(x)^2 z_i e^i + \sum_{i<j} \lambda_i(x)\lambda_j(x) z_{ij}.$$

In light of the trace function $\operatorname{tr}(\cdot)$, a semi-distance function associated with symmetric cone was proposed in [16]:

$$\mathrm{d}(x, y) := \operatorname{tr}(x + y) - 2\operatorname{tr}\left(P(x^{\frac{1}{2}})y\right)^{\frac{1}{2}}, \quad \text{for } x, y \in \mathcal{K}. \tag{1.3}$$

When the symmetric cone reduces to the second-order cone $\mathcal{K}^n$, the function $\mathrm{d}(x, y)$ was modified a bit as below distance function (1.4) and is further proved a proximal distance in [16]. When the symmetric cone reduces to the semi-definite positive matrix cone, the function $\mathrm{d}(x, y)$ corresponds to the matrix distance proposed by Givens and Shortt [11]:

$$\mathrm{d}(A, B) := \operatorname{tr}(A + B) - 2\operatorname{tr}\left(A^{\frac{1}{2}} B A^{\frac{1}{2}}\right)^{\frac{1}{2}}.$$

Theorem 1.5 ([16, Theorem 2.3]). *Let $\mathrm{d}(\cdot, \cdot)$ be defined as in (1.3). For any $x, y \in \mathcal{K}$, assume that x and y operator commute. Then, the function $\mathrm{d}(\cdot, \cdot)$ is a semi-distance, i.e.,*

(a) $\mathrm{d}(x, y) \ge 0$;
(b) $\mathrm{d}(x, y) = 0$ *if and only if* $x = y$;
(c) $\mathrm{d}(x, y) = \mathrm{d}(y, x)$.

As mentioned earlier, the function $\mathrm{d}(x,y)$ was modified a bit in the setting of second-order cone in [16], which could become a proximal distance. In particular, for any $x, y \in \mathbb{R}^n$, there defines $\mathbf{d} : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}_+ \cup \{+\infty\}$ by

$$\mathbf{d}(x,y) := \begin{cases} \operatorname{tr}(x+y) - 2\operatorname{tr}\left(P(x^{\frac{1}{2}})y\right)^{\frac{1}{2}} & \forall x \in \operatorname{int}(\mathcal{K}^n),\ y \in \mathcal{K}^n, \\ +\infty & \text{otherwise.} \end{cases} \tag{1.4}$$

This modified function $\mathbf{d}(x,y)$ is a proximal distance on $\operatorname{int}(\mathcal{K}^n)$, see [16, Theorem 3.7].

Theorem 1.6 ([16, Theorem 3.7])**.** *Let the function $\mathbf{d}(\cdot,\cdot)$ be defined by (1.4). Then, the function $\mathbf{d}(\cdot,\cdot)$ is a proximal distance with respect to $\overline{\operatorname{int}(\mathcal{K}^n)}$, i.e.,*

(a) *$\mathbf{d}(\cdot,y)$ is proper, l.s.c., convex, continuously differentiable on $\operatorname{int}(\mathcal{K}^n)$;*
(b) *$\operatorname{dom}\mathbf{d}(\cdot,y) \subset \overline{\operatorname{int}(\mathcal{K}^n)}$ and $\operatorname{dom}\partial_1\mathbf{d}(\cdot,y) = \operatorname{int}(\mathcal{K}^n)$, where the symbol $\partial_1\mathbf{d}(\cdot,y)$ denotes the classical subgradient map of the function $\mathbf{d}(\cdot,y)$ with respect to the first variable;*
(c) *$\mathbf{d}(\cdot,y)$ is level bounded on $\mathbb{R}^n$ i.e., $\lim_{\|u\|\to\infty}\mathbf{d}(u,y) = +\infty$;*
(d) *$\mathbf{d}(y,y) = 0$.*

In this short paper, we improve these two results by showing that without assuming operator commute, the function $\mathrm{d}(\cdot,\cdot)$ is a semi-distance, and the function $\mathbf{d}(\cdot,\cdot)$ is a proximal distance in the setting of symmetric cone. These generalizations enable them applicable to proximal-like algorithm for nonlinear symmetric cone programming.

2. Main results

In this section, without assuming operator commute, we show our main results. Indeed, there exists a difficulty that the same Jordan frame is not available for any two elements x and y in $\mathbb{V}$, when there is no condition of operator commute. Our novel idea to tackle with it is using the spectral decomposition of x, whereas employing the Peirce decomposition of y. These together with the quadratic representation $P(x)$ paves a way to do the analysis.

Theorem 2.1. *Let $\mathrm{d}(\cdot,\cdot)$ be defined as in (1.3). For any $x, y \in \mathcal{K}$, the function $\mathrm{d}(\cdot,\cdot)$ is a semi-distance, i.e., there hold*

(a) *$\mathrm{d}(x,y) \geq 0$;*
(b) *$\mathrm{d}(x,y) = 0$ if and only if $x = y$;*
(c) *$\mathrm{d}(x,y) = \mathrm{d}(y,x)$.*

Proof. (a) Suppose that $\{e^1, e^2, \ldots, e^r\}$ is a Jordan frame in $\mathbb{V}$. With this, we write out the spectral decomposition of x and the Peirce decomposition of y, respectively, as below:

$$\begin{aligned} x &= \lambda_1(x)\, e^1 + \cdots + \lambda_r(x)\, e^r, \\ y &= y_1\, e^1 + \cdots + y_r\, e^r + \sum_{i<j} y_{ij}. \end{aligned}$$

Based on the spectral decomposition of x, it follows that

$$x^{\frac{1}{2}} = \sqrt{\lambda_1(x)}\, e^1 + \cdots + \sqrt{\lambda_r(x)}\, e^r.$$

Combining with Theorem 1.5, this implies that

$$P(x^{\frac{1}{2}})y = \lambda_1(x)y_1\, e^1 + \cdots + \lambda_r(x)y_r\, e^r + \sum_{i<j} \sqrt{\lambda_i(x)\lambda_j(x)}\, y_{ij}.$$

Then, applying Theorem 1.3, we have

$$\operatorname{tr}\left(P(x^{\frac{1}{2}})y\right)^{\frac{1}{2}} \le \sum_{i=1}^{r} \sqrt{\lambda_i(x)y_i}.$$

According to this, for any $x, y \in \mathcal{K}$, we achieve

$$\begin{aligned} \mathrm{d}(x,y) &= \operatorname{tr}(x+y) - 2\operatorname{tr}\left(P(x^{\frac{1}{2}})y\right)^{\frac{1}{2}} \\ &\ge \operatorname{tr}(x) + \operatorname{tr}(y) - 2\sum_{i=1}^{r} \sqrt{\lambda_i(x)y_i} \\ &\ge \sum_{i=1}^{r} \lambda_i(x) + \sum_{i=1}^{r} y_i - 2\sum_{i=1}^{r} \sqrt{\lambda_i(x)y_i} \\ &= \sum_{i=1}^{r} \left(\sqrt{\lambda_i(x)} - \sqrt{y_i}\right)^2 \\ &\ge 0, \end{aligned}$$

where the second inequality follows from [12, Corollary 4.6]. Hence, we prove that $\mathrm{d}(x,y) \ge 0$.

(b) From the proof of part (a), we know that

$$\begin{aligned} \mathrm{d}(x,y) &= \operatorname{tr}(x+y) - 2\operatorname{tr}\left(P(x^{\frac{1}{2}})y\right)^{\frac{1}{2}} \ \ge\ \operatorname{tr}(x) + \operatorname{tr}(y) - 2\sum_{i=1}^{r} \sqrt{\lambda_i(x)y_i} \\ &\ge \sum_{i=1}^{r} \left(\sqrt{\lambda_i(x)} - \sqrt{y_i}\right)^2 \ \ge\ 0. \end{aligned}$$

Hence, it follows from $\mathrm{d}(x,y) = 0$ that

$$\operatorname{tr}(x+y) - 2\operatorname{tr}\left(P(x^{\frac{1}{2}})y\right)^{\frac{1}{2}} = \operatorname{tr}(x) + \operatorname{tr}(y) - 2\sum_{i=1}^{r} \sqrt{\lambda_i(x)y_i}$$

and

$$\sum_{i=1}^{r} \left(\sqrt{\lambda_i(x)} - \sqrt{y_i}\right)^2 = 0.$$

These lead that

$$\operatorname{tr}\left(P(x^{\frac{1}{2}})y\right)^{\frac{1}{2}} = \sum_{i=1}^{r} \sqrt{\lambda_i(x)y_i} \quad \text{and} \quad \sqrt{\lambda_i(x)} = \sqrt{y_i} \quad \forall i = 1, \ldots, r.$$

In addition, applying Theorem 1.3 yields

$$x = \sum_{i=1}^{r} \lambda_i(x) e^i = \sum_{i=1}^{r} y_i e^i = y.$$

Therefore, it is clear to see that $\mathrm{d}(x, y) = 0$ if and only if $x = y$.

(c) First, from [14, Proposition 3.2], for any $x, y \in \mathcal{K}$, we have

$$\lambda_i \left(P(x^{\frac{1}{2}}) y \right) = \lambda_i \left(P(y^{\frac{1}{2}}) x \right)$$

for $i = 1, \ldots, r$. This leads to

$$\lambda_i \left(P(x^{\frac{1}{2}}) y \right)^{\frac{1}{2}} = \lambda_i \left(P(y^{\frac{1}{2}}) x \right)^{\frac{1}{2}} \quad \forall i = 1, \ldots, r.$$

Hence, it follows that $\mathrm{tr} \left(P(x^{\frac{1}{2}}) y \right)^{\frac{1}{2}} = \mathrm{tr} \left(P(y^{\frac{1}{2}}) x \right)^{\frac{1}{2}}$, which implies that

$$\begin{aligned} \mathrm{d}(x, y) &= \mathrm{tr}(x + y) - 2 \,\mathrm{tr} \left(P(x^{\frac{1}{2}}) y \right)^{\frac{1}{2}} \\ &= \mathrm{tr}(y + x) - 2 \,\mathrm{tr} \left(P(y^{\frac{1}{2}}) x \right)^{\frac{1}{2}} = \mathrm{d}(y, x). \end{aligned}$$

Then, the proof is complete. □

Theorem 2.2. *Let* $\mathrm{d}(\cdot, \cdot)$ *be defined as in (1.3). Then, the function* $\mathrm{d}(x, y)$ *is convex, for any a fixed* $x \in \mathcal{K}$ *or* $y \in \mathcal{K}$.

Proof. The proof is the similar to [16, Theorem 2.4]. Hence, we omit it. □

From Theorem 2.1 and Theorem 2.2, we have shown that the function $\mathrm{d}(\cdot, \cdot)$ defined as in (1.3) is a convex semi-distance associated with symmetric cone. However, as indicated in [16], it can be verified by using the convexity of $\mathrm{d}(\cdot, \cdot)$ that the triangle inequality fails. To see this, for given any $x, y \in \mathcal{K}$, taking $z = \lambda x + (1 - \lambda) y$ and $0 < \lambda < 1$, there have

$$\begin{aligned} \mathrm{d}(x, z) &= \mathrm{d}\left(x, \lambda x + (1 - \lambda) y\right) \\ &\leq \lambda \mathrm{d}(x, x) + (1 - \lambda) \mathrm{d}(x, y) = (1 - \lambda) \mathrm{d}(x, y), \end{aligned} \tag{2.1}$$

$$\begin{aligned} \mathrm{d}(z, y) &= \mathrm{d}\left(\lambda x + (1 - \lambda) y, y\right) \\ &\leq \lambda \mathrm{d}(x, y) + (1 - \lambda) \mathrm{d}(y, y) = \lambda \mathrm{d}(x, y). \end{aligned} \tag{2.2}$$

Then, adding (2.1) and (2.2) together yields

$$\mathrm{d}(x, z) + \mathrm{d}(z, y) \leq \mathrm{d}(x, y).$$

In other words, the semi-distance $\mathrm{d}(\cdot, \cdot)$ defined as in (1.3) could not become a "distance function" (metric function). Thus, we turn our attention to the possibility of $\mathrm{d}(\cdot, \cdot)$ becoming a proximal distance.

In order to prove $\mathrm{d}(\cdot, \cdot)$ could become a proximal distance, we need to modify it a bit. For any $x, y \in \mathbb{R}^n$, we define $\mathbf{d} : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}_+ \cup \{+\infty\}$ by

$$\mathbf{d}(x, y) := \begin{cases} \mathrm{tr}(x + y) - 2 \,\mathrm{tr} \left(P(x^{\frac{1}{2}}) y \right)^{\frac{1}{2}} & \forall x \in \mathrm{int}\mathcal{K}, \ y \in \mathcal{K}, \\ +\infty & \text{otherwise}. \end{cases} \tag{2.3}$$

The above function $\mathbf{d}(\cdot,\cdot)$ is different from the ones given in [1]. To our best knowledge, it may be the only proximal distance which is not induced from Bregman distance or φ-divergence. This function, as will be shown below, is a proximal distance on $\mathrm{int}\mathcal{K}$.

Theorem 2.3. *Let $\mathbf{d}(\cdot,\cdot)$ be defined as in (2.3) in the setting of symmetric cone. Then, the function $\mathbf{d}(\cdot,\cdot)$ is a proximal distance, i.e., it satisfies*

(a) $\mathbf{d}(\cdot,y)$ *is proper, l.s.c., convex, continuously differentiable on* $\mathrm{int}\mathcal{K}$*;*

(b) $\mathrm{dom}\,\mathbf{d}(\cdot,y)\subset\overline{\mathrm{int}\mathcal{K}}$ *and* $\mathrm{dom}\,\partial_1\mathbf{d}(\cdot,y)=\mathrm{int}\mathcal{K}$*, where* $\partial_1\mathbf{d}(\cdot,y)$ *denotes the classical subgradient map of the function* $\mathbf{d}(\cdot,y)$ *with respect to the first variable;*

(c) $\mathbf{d}(\cdot,y)$ *is level bounded on* $\mathbb{R}^n$ *i.e.,* $\lim_{\|u\|\to\infty}\mathbf{d}(u,y)=+\infty$*;*

(d) $\mathbf{d}(y,y)=0$.

Proof. (a) The proof is similar to [5, Lemma 3.1], we omit the details.

(b) The arguments are similar to [16, Proposition 3.5], due to only the general cone structure is used. We also omit them.

(c) Suppose $y\in\mathrm{int}\mathcal{K}$. For any $x\in\mathrm{int}\mathcal{K}$, as what we do in Theorem 2.1, we write out the spectral decomposition of x and the Peirce decomposition of y, respectively, i.e.,

$$x=\lambda_1(x)\,e^1+\cdots+\lambda_r(x)\,e^r\quad\text{and}\quad y=y_1\,e^1+\cdots+y_r\,e^r+\sum_{i<j}y_{ij}.$$

Note that $\|x\|^2=\lambda_1(x)^2\|e^1\|^2+\cdots+\lambda_r(x)^2\|e^r\|^2\leq r\lambda_1(x)^2\|e^1\|^2$, where the inequality holds because $\|e^i\|$ is a constant on $\mathbb{V}$ for any primitive idempotent e^i $(i=1,\ldots,r)$ and $\lambda_1(x)\geq\cdots\geq\lambda_r(x)\geq 0$. Hence, it is easy to check that $\lambda_1(x)\to\infty$ as $\|x\|\to\infty$. From this and the proof of part (a) in Theorem 2.1, we have

$$\mathbf{d}(x,y)\geq\sum_{i=1}^{r}\left(\sqrt{\lambda_i(x)}-\sqrt{\mu_i}\right)^2\geq\left(\sqrt{\lambda_1(x)}-\sqrt{\mu_1}\right)^2\to\infty.$$

It follows that $\mathbf{d}(x,y)\to\infty$ as $\|x\|\to\infty$ for any $x\in\mathrm{int}\mathcal{K}$. Moreover, $\mathbf{d}(x,y)=\infty$ when $x\notin\mathrm{int}\mathcal{K}$. Then, we prove that $\mathbf{d}(x,y)\to\infty$ as $\|x\|\to\infty$ for any $x\in\mathbb{R}^n$. Thus, we conclude that $\mathbf{d}(\cdot,y)$ is level bounded on $\mathbb{R}^n$.

(d) This property is trivial.

To sum up, the function $\mathbf{d}(\cdot,\cdot)$ defined as in (2.3) is a proximal distance in the setting of symmetric cone. $\square$

Remark 2.4. We say a few words about Theorem 2.1 and Theorem 2.3. In fact, when the symmetric cone $\mathcal{K}$ reduces to the second-order cone $\mathcal{K}^n$, the conclusions of Theorem 2.1 and Theorem 2.3 correspond to the contents of Theorem 2.5 and theorem 3.7 in [16], respectively. In other words, our results are generalizations of Theorem 2.5 and theorem 3.7 in [16] in a broader framework.

3. Concluding remarks

In this paper, we study a semi-distance associated with symmetric cone $\mathcal{K}$. Furthermore, based on it, we construct a proximal distance on int$\mathcal{K}$, which also answers a question raised in [16]. Again, we would like to point out some possible future directions as mentioned in [16].

- Can the function $\mathbf{d}(\cdot,\cdot)$ further become a Bregman distance or φ-divergence?
- Can the function $\mathbf{d}(\cdot,\cdot)$ be extended to nonsymmetric cone setting? In particular, for circular cone $\mathcal{L}_\theta$, we have already known one type of spectral decomposition of x and some differentiabilities of $\lambda_i(x)$, see [24]. By using these facts, we may consider to construct an analogous distance function $\mathbf{d}(\cdot,\cdot)$ in the setting of circular cone.

References

[1] A. Auslender and M. Teboulle, *Interior gradient and proximal methods for convex and conic optimization*, SIAM J. Optim. **16** (2006), 697–725.

[2] A. Banerjee, I. Dhillon and J. Ghosh, *Clustering with Bregman divergences*, J. Mach. Learn. Res. **6** (2005), 1705–1749.

[3] A. Banerjee, I. Dhillon, J. Ghosh, S. Merugu and D. Modha, *A generalized maximum Entropy approach to Bregman co-clustering and matrix approximations*, J. Mach. Learn. Res. **8** (2007), 1919–1986.

[4] Y. Censor and A. Zenios, *Proximal minimization algorithms with D-functions*, J. Optim. Theory Appl. **73** (1992), 451–464.

[5] Y.-L. Chang and J.-S. Chen, *Convexity of symmetric cone trace functions in Euclidean Jordan algebras*, J. Nonlinear Convex Anal. **14** (2013), 53–61.

[6] J.-S. Chen and S. Pan, *Proximal-like algorithm using the quasi D-function for convex second-order cone programming*, J. Optim. Theory Appl. **138** (2008), 95–113.

[7] G. Chen and M. Teboulle, *Convergence analysis of proximal-like minimization algorithm using Bregman functions*, SIAM J. Optim. **3** (1993), 538–543.

[8] I. Dhillon and S. Sra, *Generalized nonnegative matrix approximations with Bregman divergences*, in: NIPS'05: Proceedings of the 18th International Conferencen on Neural Information Processing Systems, 2005, pp. 283–290.

[9] S. Dhillon and J.A. Tropp, *Matrix nearness problems with Bregman divergences*, SIAM J. Matrix Anal. Appl. **29** (2007), 1120–1146.

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

[11] C. R. Givens and R. M. Shortt, *A class of Wasserstein metrics for probability distributions*, Mich. Math. J. **31** (1984), 231–240.

[12] M. S. Gowda and J. Tao, *The Cauchy interlacing theorem in simple Euclidean Jordan algebras and some consequences*, Linear Multilinear Algebra, **59** (2011), 65–86.

[13] K. C. Kiwiel, *Proximal minimization methods with generalized Bregman functions*, SIAM J. Control Optim. **35** (1997), 1142–1168.

[14] J. Kim and Y. Lim, *Jordan automorphic generators of Euclidean Jordan algebras*, J. Korean Math. Soc. **43** (2006), 507–528.

[15] B. Kulis, M. Sustik and I. Dhillon, *Low-Rank Kernel Learning with Bregman Matrix Divergences*, J. Mach. Learn. Res. **10** (2009), 341–376.

[16] X.-H. Miao, C.-H. Huang, Y. Lim and J.-S. Chen, *A semi-distance associated with symmetric cone and a new proximal distance function on second-order cone*, Linear and Nonlinear Anal. **5** (2019), 421–437.

[17] S. Pan and J.-S. Chen, *A class of interior proximal-like algorithms for convex second-order cone programming*, SIAM J. Optim. **19** (2008), 883–910.

[18] E. A. Papa Quiroz, *An extension of the proximal point algorithm with Bregman distances on Hadamard manifolds*, J. Global Optim. **56** (2013), 43–59.
[19] R. T. Rockafellar, *Convex Analysis*, Princeton University Press, Princeton, New Jersey, 1970.
[20] W. Stummer and I. Vajda, *On Bregman distances and divergences of probability measures*, IEEE Trans. Inf. Theory. **58** (2012), 1277–1288.
[21] M. Teboulle, *Entropic proximal mappings with applications to nonlinear programming*, Math. Oper. Res. **17** (1992), 670–690.
[22] J. Tao, L. C. Kong, Z. Y. Luo and N. H. Xiu, *Some majorization inequalities in Euclidean Jordan algebras*, Linear Algebra Appl. **461** (2014), 92–122.
[23] K. Tsuda, G. Ratsch and M. K. Warmuth, *Matrix exponentiated gradient updates for on-line learning and Bregman projection*, J. Mach. Learn. Res. **6** (2005), 995–1018.
[24] J. C. Zhou, J. Y. Tang and J.-S. Chen, *Parabolic second-order directional differentiability in the Hadamard sense of the vector-valued functions associated with circular cones*, J. Optim. Theory Appl. **172** (2017), 802–823.

Manuscript received December 25, 2020
revised December 19, 2021

X. Miao
School of Mathematics, Tianjin University, Tianjin 300072, P.R. China
E-mail address: `xinhemiao@tju.edu.cn`

J.-S. Chen
Department of Mathematics, National Taiwan Normal University, Taipei 11677, Taiwan
E-mail address: `jschen@math.ntnu.edu.tw`