Nonlinear Functional Analysis and Applications
Vol. 29, No. 4 (2024), pp. 1185-1198
ISSN: 1229-1595(print), 2466-0973(online)
https://doi.org/10.22771/nfaa.2024.29.04.16
http://nfaa.kyungnam.ac.kr/journal-nfaa

REMARKS ON THE PROPERTIES OF WASSERSTEIN MEDIANS ON SYMMETRIC CONES

Sangho Kum

Department of Mathematics Education, College of Education,
Chungbuk National University, Cheongju 28644, Korea
e-mail: shkum@cbnu.ac.kr

Abstract. The author studied a one-parameter family of divergences and the related median minimization problem of finite points over these divergences in a symmetric cone [14]. The unique solution of the minimization problem with a weight ω is called the ω-weighted Wasserstein median. Recently, in the special symmetric cone of the positive definite matrices, Hwang and Kim [8] explored several properties of the Wasserstein mean and found bounds for the Wasserstein mean with respect to Löwner order. Also Kim and Lee [10] presented some relations between the Wasserstein mean and other well-known matrix means such as the power mean, harmonic mean and Karcher mean. Motivated by these results, as an application of the previous work [14], we investigate several properties of the ω-weighted Wasserstein median which mainly extend the corresponding ones in [8, 10] into a general symmetric cone Ω with a purely Jordan-algebraic technique.

1. Introduction

A divergence, which measures discrepancy between two points, plays a crucial role in many problems such as information theory, statistics, optimization, computational vision, and neural networks [1, 2, 3]. For the definition of divergence, readers may refer to [14]. As is known, a divergence is almost a distance function except the symmetry with respect to its arguments and the triangle inequality. For instance, the square of a distance function is a (symmetric) divergence. Thus a divergence Φ gives rise to an important optimization problem

[0]Received March 25, 2024. Revised May 2, 2024. Accepted June 13, 2024.
[0]2020 Mathematics Subject Classification: 17C20, 15B48, 90C25.
[0]Keywords: Symmetric cone, Euclidean Jordan algebra, divergence, median, Wasserstein barycenter.

like a least squares problem on a Riemannian manifold M :

$$\arg\min_{x\in M}\sum_{j=1}^{m} w_j\Phi(a_j,x), \tag{1.1}$$

where $a_1,\ldots,a_m \in M$ and $\omega=(w_1,\ldots,w_m)\in\mathbb{R}^m$ is a positive probability vector. So a minimizer whenever it exists provides alternatively a barycenter or averaging on M, which is called the *ω-weighted Φ-median of* $a_1,\ldots,a_m$.

In a recent work [14], this median optimization problem on a special Riemannian manifold called symmetric cones is studied. The main result in [14] is briefly summarized as follow: Let V be a Euclidean Jordan algebra and let Ω be the symmetric cone (see section 2 for basic facts regarding Euclidean Jordan algebras and symmetric cones). Consider the function $\Phi_t:\Omega\times\Omega\to\mathbb{R}$ defined by

$$\Phi_t(a,b)=\mathrm{tr}\,((1-t)a+tb)-\mathrm{tr}\left(P(a^{\frac{1-t}{2t}})b\right)^t, \qquad 0<t<1, \tag{1.2}$$

where tr is the trace functional and P is the quadratic representation of V.

Theorem 1.1. ([14]) *For every* $0<t<1$, Φ_t *is a divergence on* Ω. *Moreover, the minimization problem*

$$\arg\min_{x\in\Omega}\sum_{j=1}^{m} w_j\Phi_t(a_j,x) \tag{1.3}$$

has a unique minimizer.

A meaningful reason to take the divergence (1.2) into account stems from the followings: The term $F_t(a,b)=\mathrm{tr}\left(P(a^{\frac{1-t}{2t}})b\right)^t$ in (1.2) is known as *sandwiched quasi-relative entropy* in the theory of quantum information; for positive (semi)definite matrices A and B,

$$F_t(A,B)=\mathrm{tr}\left(A^{\frac{1-t}{2t}}BA^{\frac{1-t}{2t}}\right)^t, \qquad t\in(0,1). \tag{1.4}$$

This is a parameterized version of the *fidelity* $F_{\frac{1}{2}}(A,B)=\mathrm{tr}\left(A^{\frac{1}{2}}BA^{\frac{1}{2}}\right)^{\frac{1}{2}}$. Fidelity and sandwiched quasi-relative entropies play an essential role in quantum information theory and quantum computation [7, 20, 21, 22]. In addition, the extended version d_W on the set of positive definite matrices of the *Bures distance* in quantum information is defined by

$$d_W(A,B)=\left[\mathrm{tr}\left(\frac{A+B}{2}\right)-\mathrm{tr}\left(A^{\frac{1}{2}}BA^{\frac{1}{2}}\right)^{\frac{1}{2}}\right]^{\frac{1}{2}},$$

which is also known as the *Wasserstein distance* in statistics and the theory of optimal transport [5, 12, 19]. Clearly, $d_W^2(A,B) = \Phi_{\frac{1}{2}}(A,B)$. This implies that the divergence (1.2) may have a rich background in various areas mentioned above even thought its square root is not a distance in general [14]. From now on, we call the unique minimizer of (1.3) the *ω-weighted Wasserstein median of* $a_1, \dots, a_m$.

A motivation of this paper is as follows. Recently, in the special symmetric cone of the positive definite matrices, Hwang and Kim [8] explored several properties of the Wasserstein mean and found bounds for the Wasserstein mean with respect to Löwner order. Also Kim and Lee [10] presented some relations between the Wasserstein mean and other well-known matrix means such as the power mean, harmonic mean and Karcher mean. Motivated by these results, as a continuation or an application of the previous work [14], in the present paper, we investigate several important properties of the ω-weighted Wasserstein median of $a_1, \dots, a_m$ which mainly extend the corresponding ones in [8, 10] into the setting of a general symmetric cone Ω with a purely Jordan-algebraic technique.

2. Euclidean Jordan algebras and symmetric cones

Before stating the motivation of this work, first we briefly describe (following mostly [6, 14]) some Jordan-algebraic concepts relevant to our purpose. A Jordan algebra V over $\mathbb{R}$ is a *non-associative* commutative algebra satisfying $x^2(xy) = x(x^2y)$ for all $x, y \in V$. For $x \in V$, let L_x be the linear operator defined by $L_x y = xy$, and let $P(x) = 2L_x^2 - L_{x^2}$. The map P is called the quadratic representation of V. An element $x \in V$ is said to be invertible if there exists an element y (denoted by $y = x^{-1}$) in the subalgebra generated by x and e (the Jordan identity) such that $xy = e$.

The following appears at [6, Propositions II.3.1, II.3.2].

Proposition 2.1. *Let V be a Jordan algebra.*

(i) *An element x in V is invertible if and only if $P(x)$ is invertible. In this case: $P(x)^{-1} = P(x^{-1})$.*
(2) *If x and y are invertible, then $P(x)y$ is invertible and $(P(x)y)^{-1} = P(x^{-1})y^{-1}$.*
(iii) *For any elements x and y, $P(P(x)y) = P(x)P(y)P(x)$. In particular, $P^m(x) = P(x^m), \quad m \geq 1$.*

An element $c \in V$ is called an idempotent if $c^2 = c$. We say that $c_1, \dots, c_k$ is a complete system of orthogonal idempotents if $c_i^2 = c_i$, $c_i c_j = 0$, $i \neq j$ and $c_1 + \cdots + c_k = e$. An idempotent is said to be primitive if it is non-zero and cannot be written as the sum of two non-zero idempotents. A Jordan frame is a complete system of orthogonal primitive idempotents.

A finite-dimensional Jordan algebra V with an identity element e is said to be *Euclidean* if there exists an inner product $\langle\cdot,\cdot\rangle$ such that $\langle xy, z\rangle = \langle y, xz\rangle$ for all $x, y, z \in V$.

Theorem 2.2. (Spectral theorem, first version [6, Theorem III.1.1]) *Let V be a Euclidean Jordan algebra. Then for $x \in V$, there exist real numbers $\lambda_1, \dots, \lambda_k$ all distinct and a unique complete system of orthogonal idempotents $c_1, \dots, c_k$ such that*

$$x = \sum_{i=1}^{k} \lambda_i c_i. \tag{2.1}$$

The numbers λ_i are called the eigenvalues and (2.1) *is called the spectral decomposition of x.*

Theorem 2.3. (Spectral theorem, second version [6, Theorem III.1.2]) *Any two Jordan frames in a Euclidean Jordan algebra V have the same number of elements (called the rank of V, denoted by* $\mathrm{rank}(V)$*). Given $x \in V$, there exists a Jordan frame $c_1, \dots, c_r$ and real numbers $\lambda_1, \dots, \lambda_r$, where r is the rank of V, such that $x = \sum_{i=1}^{r} \lambda_i c_i$. The numbers λ_i (with their multiplicities) are uniquely determined by x.*

Definition 2.4. Let V be a Euclidean Jordan algebra of $\mathrm{rank}(V) = r$. The spectral mapping $\lambda : V \to \mathbb{R}^r$ is defined by $\lambda(x) = (\lambda_1(x), \dots, \lambda_r(x))$, where $\lambda_i(x)$'s are eigenvalues of x (with multiplicities) as in Theorem 2.3 in non-increasing order $\lambda_{\max}(x) = \lambda_1(x) \geq \lambda_2(x) \geq \cdots \geq \lambda_r(x) = \lambda_{\min}(x)$. Furthermore, $\det(x) = \prod_{i=1}^{r} \lambda_i(x)$ and $\mathrm{tr}(x) = \sum_{i=1}^{r} \lambda_i(x)$.

Let Q be the set of all square elements of V. Then Q is a closed convex cone of V with $Q \cap -Q = \{0\}$, and is the set of elements $x \in V$ such that $L(x)$ is positive semi-definite. It turns out that Q has non-empty interior $\Omega := \mathrm{int}(Q)$, and Ω is a symmetric cone, that is, the group $G(\Omega) = \{g \in \mathrm{GL}(V) \mid g(\Omega) = \Omega\}$ acts transitively on it and Ω is a self-dual cone with respect to the inner product $\langle\cdot,\cdot\rangle$ (see [6]). We say that $g \in \mathrm{GL}(V)$ is a Jordan automorphism if $g(xy) = g(x)g(y)$ for all $x, y \in V$. For an element $g \in G(\Omega)$, g is a Jordan automorphism if and only if $g(e) = e$ [6, Propositions VIII.2.4, 2.8] or [11]).

Note that $\overline{\Omega} = \{x \in V \mid \lambda_i(x) \geq 0, i = 1, \dots, r\}$. For $x, y \in V$, we define the Löwner order by

$$x \leq y \quad \text{if} \quad y - x \in \overline{\Omega}$$

and $x < y$ if $y - x \in \Omega$. Clearly $\overline{\Omega} = \{x \in V \mid x \geq 0\}$ and $\Omega = \{x \in V \mid x > 0\}$. Furthermore, for any $a \in \Omega$, $P(a) \in G(\Omega)$ so that $P(a)$ is an order preserving linear operator.

On the other hand, the symmetric cone Ω in V admits a $G(\Omega)$-invariant Riemannian metric defined by

$$\langle u, v\rangle_x = \langle P(x)^{-1}u, v\rangle, \ x \in \Omega, \ u, v \in V. \tag{2.2}$$

So Ω is a Riemannian manifold [6]. It is shown in [16, Proposition 2.6] that the unique geodesic joining a and b is

$$t \mapsto a\#_t b := P(a^{1/2})(P(a^{-1/2})b)^t, \tag{2.3}$$

where $a^t = \sum_{j=1}^r \lambda_j(a)^t c_j$ for the spectral decomposition $a = \sum_{j=1}^r \lambda_j(a)c_j$ in Theorem 2.3. *The geometric mean of a and b* is defined to be $a\#b := a\#_{1/2}b$, which is a unique geodesic middle between a and b. It is well-known that $a\#b$ is the unique solution on Ω of the following quadratic equation, called the Riccati equation:

$$P(x)a^{-1} = b. \tag{2.4}$$

3. Main results

For simplicity, in what follows, we restrict our attention only to the case $t = 1/2$. In this case, we denote by $W(\omega; a_1, \dots, a_m)$ the unique minimizer of (1.3), that is, the ω-weighted Wasserstein median of $a_1, \dots, a_m$,

$$\begin{aligned} W(\omega; a_1, \dots, a_m) &= \underset{x\in\Omega}{\arg\min} \sum_{j=1}^m w_j \, \Phi_{\frac{1}{2}}(a_j, x) \\ &= \underset{x\in\Omega}{\arg\min} \sum_{j=1}^m w_j \left(\operatorname{tr}\left(\frac{a_j + x}{2}\right) - \operatorname{tr}\left(P(a_j^{\frac{1}{2}})x\right)^{\frac{1}{2}} \right). \end{aligned} \tag{3.1}$$

Let $\mathbb{A} = (a_1, \dots, a_m) \in \Omega^m$ and let σ be a permutation on $\{1, \dots, m\}$. We need the following notations:

$$\begin{aligned} W(\omega; \mathbb{A}) &= W(\omega; a_1, \dots, a_m), \\ \mathbb{A}_\sigma &= (a_{\sigma(1)}, \dots, a_{\sigma(m)}), \\ \mathbb{A}^k &= (\underline{a_1, \dots, a_m}, \dots, \underline{a_1, \dots, a_m}) \in \Omega^{km}, \\ \omega_\sigma &= (w_{\sigma(1)}, \dots, w_{\sigma(m)}), \\ \omega^k &= \frac{1}{k}(\underline{w_1, \dots, w_m}, \dots, \underline{w_1, \dots, w_m}) \in \mathbb{R}^{km}, \end{aligned}$$

where the number of blocks of the above expressions is a positive integer k, and $\omega = (w_1, \dots, w_m) \in \mathbb{R}^m$ is a positive probability vector as before. We begin with simple observations.

Theorem 3.1. *The following properties of the ω-weighted Wasserstein median of $a_1, \ldots, a_m$ are satisfied.*

(i) $W(\omega; \alpha\mathbb{A}) = \alpha W(\omega; \mathbb{A})$ *for any* $\alpha > 0$ *(Positive homogeniety).*

(ii) $W(\omega_\sigma; \mathbb{A}_\sigma) = W(\omega; \mathbb{A})$ *for any permutation* σ *on* $\{1, \ldots, m\}$ *(Permutation invariancy).*

(iii) $W(\omega^k; \mathbb{A}^k) = W(\omega; \mathbb{A})$ *for any* $k \in \mathbb{N}$ *(Repetition invariancy).*

Proof. First observe that according to [14, (5.3)], the ω-weighted Wasserstein median $x_0 = W(\omega; \mathbb{A})$ is characterized by the equivalence:

$$x_0 = \sum_{j=1}^m w_j \left(P(x_0^{1/2})a_j\right)^{\frac{1}{2}}. \tag{3.2}$$

This tells us that $x_0 = W(\omega; \mathbb{A})$ is the unique solution on Ω of the nonlinear equation

$$x = \sum_{j=1}^m w_j \left(P(x^{1/2})a_j\right)^{\frac{1}{2}}. \tag{3.3}$$

(i) By (3.2), we have

$$\begin{aligned} \alpha x_0 &= \sum_{j=1}^m w_j \left(\alpha^2 P(x_0^{1/2})a_j\right)^{\frac{1}{2}} \\ &= \sum_{j=1}^m w_j \left(P(\alpha^{1/2}x_0^{1/2})(\alpha a_j)\right)^{\frac{1}{2}} \\ &= \sum_{j=1}^m w_j \left(P(\alpha x_0)^{1/2})(\alpha a_j)\right)^{\frac{1}{2}}. \end{aligned}$$

This implies that αx_0 satisfies (3.3) for $\alpha a_1, \ldots, \alpha a_m$. Hence, $W(\omega; \alpha\mathbb{A}) = \alpha W(\omega; \mathbb{A})$.

(ii) This is clear from the definition of $W(\omega; \mathbb{A})$ by (3.1).

(iii) Put $x_* = W(\omega^k; \mathbb{A}^k)$. From (3.2), we obtain that

$$\begin{aligned} x_* &= \left[\sum_{j=1}^m \frac{w_j}{k} \left(P(x_*^{1/2})a_j\right)^{\frac{1}{2}}\right] + \cdots + \left[\sum_{j=1}^m \frac{w_j}{k} \left(P(x_*^{1/2})a_j\right)^{\frac{1}{2}}\right] \\ &= k \cdot \frac{1}{k}\left[\sum_{j=1}^m w_j \left(P(x_*^{1/2})a_j\right)^{\frac{1}{2}}\right] \\ &= \sum_{j=1}^m w_j \left(P(x_*^{1/2})a_j\right)^{\frac{1}{2}}, \end{aligned}$$

which implies that $x_* = W(\omega; \mathbb{A})$ by (3.2). □

Theorem 1.1 tells us the existence of $W(\omega;\mathbb{A})$. However, there does not exist a closed form of $W(\omega;\mathbb{A})$ in general. Only in two variable case, the exact formula of $W(\omega;\mathbb{A})$ is calculated in [14, Theorem 8.1] on Ω. Nonetheless, as in the case of positive definite matrices when the matrices A_j's commute (see [4]), an analogous result can be derived in the symmetric cone Ω. To this end, we recall that elements x and y are said to operator commute if L_x and L_y commute, that is,, $L_xL_y = L_yL_x$. It is known that x and y operator commute if and only if x and y have their spectral decompositions with respect to a common Jordan frame [6, Lemma X.2.2]. Equivalently, we say that x and y are simultaneously diagonalizable. In this case, x and y lie in an associative subalgebra of V. With this property, we obtain the following.

Theorem 3.2. (Consistency with scalars) *Assume that a_j's operator commute. Then we have*

$$W(\omega;\mathbb{A}) = \left(\sum_{j=1}^m w_j\, a_j^{1/2}\right)^2.$$

Proof. Due to (3.2), it suffices to show that

$$\left(\sum_{j=1}^m w_j\, a_j^{1/2}\right)^2 = \sum_{j=1}^m w_j \left(P\left(\sum_{i=1}^m w_i\, a_i^{1/2}\right)a_j\right)^{\frac{1}{2}}. \tag{3.4}$$

Indeed, by the definition of the quadratic representation P, we have

$$\begin{aligned} P\left(\sum_{i=1}^m w_i\, a_i^{1/2}\right)a_j &= \left(2L_{\sum_{i=1}^m w_i\, a_i^{1/2}}L_{\sum_{i=1}^m w_i\, a_i^{1/2}} - L_{(\sum_{i=1}^m w_i\, a_i^{1/2})^2}\right)(a_j)\\ &= 2\left(\sum_{i=1}^m w_i\, a_i^{1/2}\right)\left(\left(\sum_{i=1}^m w_i\, a_i^{1/2}\right)a_j\right) - \left(\sum_{i=1}^m w_i\, a_i^{1/2}\right)^2 a_j\\ &= 2\left(\sum_{i=1}^m w_i\, a_i^{1/2}\right)^2 a_j - \left(\sum_{i=1}^m w_i\, a_i^{1/2}\right)^2 a_j\\ &= \left(\sum_{i=1}^m w_i\, a_i^{1/2}\right)^2 a_j, \end{aligned}$$

where third equality comes from the assumption that a_j's operator commute so that they have a common Jordan frame and lie in an associative subalgebra of V as mentioned before. Thus, we get

$$\left(P\left(\sum_{i=1}^m w_i\, a_i^{1/2}\right)a_j\right)^{\frac{1}{2}} = \left(\left(\sum_{i=1}^m w_i\, a_i^{1/2}\right)^2 a_j\right)^{\frac{1}{2}} = \left(\sum_{i=1}^m w_i\, a_i^{1/2}\right)a_j^{1/2}. \tag{3.5}$$

The second equality is repeatedly obtained from the fact that a_j's have a common Jordan frame. Summing up (3.5) on both sides yields that

$$\sum_{j=1}^{m} w_j\left(P\left(\sum_{i=1}^{m} w_i\, a_i^{1/2}\right)a_j\right)^{\frac{1}{2}} = \sum_{j=1}^{m} w_j\left(\left(\sum_{i=1}^{m} w_i\, a_i^{1/2}\right)a_j^{1/2}\right) = \left(\sum_{i=1}^{m} w_i\, a_i^{1/2}\right)^2,$$

which entails the claim (3.4), as desired. This completes the proof. □

Remark 3.3. In [4, 10] the authors just mentioned and quoted the consistency with scalars for the case of positive definite matrices which is a special case of the symmetric cone Ω. We provided a basic and elegant proof of Theorem 3.2 in Ω. So Theorem 3.2 extends the corresponding one in [4, 10] into the general setting of Ω.

In [8, Proposition 2.3], *the unitary congruence invariancy* of $W(\omega;\mathbb{A})$, $W(\omega;U\mathbb{A}U^*) = UW(\omega;\mathbb{A})U^*$ holds true when U is an unitary matrix and $\mathbb{A} = (A_1,\ldots,A_m)$ is an m-tuple of positive definite matrices A_j's. This can be generalized in Ω as follows:

Theorem 3.4. (Automorphism invariancy) $W(\omega;g(\mathbb{A})) = g(W(\omega;\mathbb{A}))$ *for any Jordan automorphism* g*, where* $g(\mathbb{A}) = (g(a_1),\ldots,g(a_m))$.

Proof. Claim 1. $P(g(x)) = gP(x)g^{-1}$ for $x \in V$.
First observe that $gL_xg^{-1} = L_{g(x)}$. Indeed, for $y \in V$, we have

$$(gL_xg^{-1})(y) = g(xg^{-1}(y)) = g(x)g(g^{-1}(y)) = g(x)y = L_{g(x)}(y).$$

Thus, we get

$$\begin{aligned} gP(x)g^{-1} &= g(2L_x^2 - L_{x^2})g^{-1} = 2gL_xL_xg^{-1} - gL_{x^2}g^{-1} \\ &= 2(gL_xg^{-1})(gL_xg^{-1}) - L_{g(x^2)} \\ &= 2L_{g(x)}^2 - L_{g(x)^2} = P(g(x)). \end{aligned}$$

Claim 2. $P(a\#b)a^{-1} = b$ for $a, b \in \Omega$.
In fact, a proof of this claim is presented implicitly in [17, Proposition 6 (i)]. For reader's convenience, we provide a detailed one.

$$\begin{aligned} P(a\#b)a^{-1} &= P(a\#b)P(a^{-1/2})e \\ &= (P(a)\#P(b))P(a^{-1/2})e \\ &= (P(a^{1/2})[P(a^{-1/2})P(b)P(a^{-1/2})]^{1/2}P(a^{1/2}))P(a^{-1/2})e \\ &= P(a^{1/2})[P(P(a^{-1/2})b)]^{1/2}e \\ &= P(a^{1/2})P([P(a^{-1/2})b]^{1/2})e \\ &= P(a^{1/2})P(a^{-1/2})b = b, \end{aligned}$$

where the second equality is from [16, Proposition 2.5], and the third, fourth and fifth ones come from Proposition 2.1.

Claim 3. $g(a\#b) = g(a)\#g(b)$ for $a, b \in \Omega$.
Clearly $g \in G(\Omega)$. Then we adopt the technique in [17, Proposition 6 (vi)]. By the Riccati equation (2.4), we have

$$P(g(a)\#g(b))g(a)^{-1} = g(b).$$

On the other hand, by Claims 1 and 2, we also obtain that

$$\begin{aligned} P(g(a\#b))g(a)^{-1} &= gP(a\#b)g^{-1}g(a^{-1}) \\ &= gP(a\#b)a^{-1} \\ &= g(b). \end{aligned}$$

As $g(a\#b)$ belongs to Ω, this implies that $g(a\#b)$ is a solution of the same Riccati equation $P(x)g(a)^{-1} = g(b)$. Hence, we get $g(a\#b) = g(a)\#g(b)$.

Claim 4. $W(\omega; g(\mathbb{A})) = g(W(\omega; \mathbb{A}))$.
Let $x_0 = W(\omega; \mathbb{A})$. Note that from [14, (5.2)] we have

$$x_0 = W(\omega; \mathbb{A}) \Leftrightarrow e = \sum_{j=1}^{m} w_j\, (a_j \# x_0^{-1}). \tag{3.6}$$

Since g is a Jordan automorphism, it follows from Claim 3 that

$$\begin{aligned} e = g(e) &= g\left(\sum_{j=1}^{m} w_j\, (a_j \# x_0^{-1})\right) = \sum_{j=1}^{m} w_j\, g(a_j \# x_0^{-1}) \\ &= \sum_{j=1}^{m} w_j\, (g(a_j) \# g(x_0^{-1})) \\ &= \sum_{j=1}^{m} w_j\, (g(a_j) \# g(x_0)^{-1}). \end{aligned}$$

This means that $g(x_0) = W(\omega; g(\mathbb{A}))$ by (3.6). Therefore, we conclude that

$$g(W(\omega; \mathbb{A})) = g(x_0) = W(\omega; g(\mathbb{A})).$$

This finishes the proof. □

Also we get two basic inequalities as below.

Theorem 3.5. (Arithmetic-Wasserstein median inequality)

$$W(\omega; \mathbb{A}) \le \sum_{j=1}^{m} w_j\, a_j.$$

.

Proof. First, the map $f(x) = x^2$ on V is operator-convex in the sense that

$$f((1-t)a + tb) \leq (1-t)f(a) + tf(b) \text{ for all } t \in [0,1],\ a,\, b \in V.$$

In fact,

$$\begin{aligned}(1-t)f(a) + tf(b) - f((1-t)a + tb) &= (1-t)a^2 + tb^2 - \big((1-t)a + tb\big)^2 \\ &= t(1-t)(a-b)^2 \\ &\geq 0.\end{aligned}$$

Put $x_0 = W(\omega; \mathbb{A})$. Then by (3.2) and the operator-convexity of f, we get

$$\begin{aligned}x_0^2 &= \left(\sum_{j=1}^m w_j \left(P(x_0^{1/2})a_j\right)^{\frac{1}{2}}\right)^2 \\ &\leq \sum_{j=1}^m w_j \left(P(x_0^{1/2})a_j\right) \\ &= P(x_0^{1/2})\left(\sum_{j=1}^m w_j\, a_j\right).\end{aligned}$$

Taking the order preserving linear operator $P(x_0^{-1/2})$ on the both sides yields that

$$\begin{aligned}x_0 &= P(x_0^{-1/2})x_0^2 \\ &\leq P(x_0^{-1/2})P(x_0^{1/2})\left(\sum_{j=1}^m w_j\, a_j\right) \\ &= \sum_{j=1}^m w_j\, a_j,\end{aligned}$$

as desired. □

Remark 3.6. Theorem 3.5 is an extension of corresponding result in [4, 8] into the symmetric cone Ω.

Theorem 3.7. (Determinantal inequality)

$$\det W(\omega; \mathbb{A}) \geq \prod_{j=1}^m (\det a_j)^{w_j}.$$

Proof. Put $x_0 = W(\omega; \mathbb{A})$. As is well known, the function $g(x) = -\log\det x$ is strictly convex on Ω because the Hessian $\nabla^2 g(x) = P(x^{-1})$ is positive definite

[15]. Thus

$$\log\det\left(\sum_{j=1}^{m} w_j\, a_j\right) \geq \sum_{j=1}^{m} w_j\, \log\det a_j.$$

Appealing to [14, (5.2)] and the argument in [9], we get

$$\begin{aligned}
0 &= \log\det e \\
&= \log\det\left(\sum_{j=1}^{m} w_j\left(a_j\# x_0^{-1}\right)\right) \\
&\geq \sum_{j=1}^{m} w_j \log\det\left(a_j\# x_0^{-1}\right) \\
&= \frac{1}{2}\sum_{j=1}^{m} w_j \log(\det a_j)(\det x_0^{-1}) \\
&= \frac{1}{2}\sum_{j=1}^{m} \log(\det a_j)^{w_j} - \frac{1}{2}\log\det x_0 \\
&= \frac{1}{2}\log\left(\prod_{j=1}^{m}(\det a_j)^{w_j}\right) - \frac{1}{2}\log\det x_0,
\end{aligned}$$

which entails the desired inequality. □

Remark 3.8. Theorem 3.7 is a generalization of [8, Proposition 2.3 (5)] under the circumstances of symmetric cone Ω.

The following theorem in [14] is concerned with the location of $W(\omega;\mathbb{A})$, which is sort of a sharpened extension of [10, Lemma 2.4] on Ω. For the sake of readers, we give a proof.

Theorem 3.9. *Let* $\alpha := \min\{\lambda_{\min}(a_j) \mid j = 1,\dots,m\}$, $\beta := \max\{\lambda_{\max}(a_j) \mid j = 1,\dots,m\}$, *where* $\lambda_{\min}(a_j)$ *and* $\lambda_{\max}(a_j)$ *denote the minimum and maximum eigenvalues of* a_j*, respectively. Then*

$$W(\omega;\mathbb{A}) \in [\alpha e, \beta e] = \{x \mid \alpha e \leq x \leq \beta e\}.$$

Proof. First note that $a_j \in [\alpha e, \beta e],\ j = 1,\dots,m$. Define a mapping $F : [\alpha e, \beta e] \to [\alpha e, \beta e]$ by

$$F(x) = \sum_{i=1}^{m} w_j\left(P(x^{1/2})a_j\right)^{\frac{1}{2}}.$$

To see that F is a self-map, let $x \in [\alpha e,\ \beta e]$. By the order preserving property of $P(x^{1/2})$, we have

$$\begin{aligned}\alpha^2 e &= \alpha(\alpha e)\\ &\le \alpha x = \alpha P(x^{1/2})e = P(x^{1/2})(\alpha e)\\ &\le P(x^{1/2})a_j \le P(x^{1/2})(\beta e) = \beta P(x^{1/2})e = \beta x\\ &\le \beta(\beta e) = \beta^2 e.\end{aligned}$$

Therefore,

$$\begin{aligned}\alpha e &= \sum_{j=1}^m w_j \alpha e = \sum_{j=1}^m w_j \left(\alpha^2 e\right)^{\frac{1}{2}}\\ &\le \sum_{j=1}^m w_j \left(P(x^{1/2})a_j\right)^{\frac{1}{2}} \ (= F(x))\\ &\le \sum_{j=1}^m w_j \left(\beta^2 e\right)^{\frac{1}{2}} = \sum_{j=1}^m w_j \beta e = \beta e.\end{aligned}$$

By Brouwer's fixed point theorem, there exists a point $x_0 \in [\alpha e, \beta e]$ such that $x_0 = F(x_0)$. It follows from (3.3) that x_0 is the unique solution of (3.3), that is, $x_0 = W(\omega; \mathbb{A})$. This completes the proof. □

As noted in [4], the Wasserstein-harmonic mean inequality does not hold. However, a lower and an upper bound for $W(\omega; \mathbb{A})$ is derived in [8]. Using the same argument, we get an extension of [8, Theorem 3.4] in the symmtric cone Ω.

Theorem 3.10. *$W(\omega; \mathbb{A})$ satisfies the following order relation.*

$$2e - \sum_{j=1}^m w_j a_j^{-1} \ \le W(\omega; \mathbb{A}) \ \le \left(2e - \sum_{j=1}^m w_j a_j\right)^{-1}$$

as long as $2e - \sum_{j=1}^m w_j a_j$ has an inverse.

Proof. Let $x_0 = W(\omega; \mathbb{A})$. According to the arithmetic-geometric-harmonic mean inequalities [16, Theorem 2.8] in Ω, we have, for $j = 1, \dots, m$,

$$\left(\frac{a_j^{-1} + x_0}{2}\right)^{-1} \ \le \ a_j \# x_0^{-1} \ \le \ \frac{a_j + x_0^{-1}}{2}.$$

Hence, we obtain

$$\sum_{j=1}^m w_j \left(\frac{a_j^{-1} + x_0}{2}\right)^{-1} \ \le \ e = \sum_{j=1}^m w_j (a_j \# x_0^{-1}) \ \le \ \frac{1}{2}\sum_{j=1}^m w_j a_j + \frac{1}{2} x_0^{-1},$$

where the equality $e = \sum_{j=1}^{m} w_j(a_j \# x_0^{-1})$ comes from (3.6). From the second inequality, we get easily the upper bound for $x_0 = W(\omega; \mathbb{A})$. By the order-reversing property of the inversion [18, Lemma 11], taking inverse on the both sides of the first inequality together with the arithmetic-harmonic mean inequalities [16, Theorem 2.8] yields the following inequalities:

$$\begin{aligned} e &\leq \left[\sum_{j=1}^{m} w_j \left(\frac{a_j^{-1} + x_0}{2}\right)^{-1}\right]^{-1} \\ &\leq \sum_{j=1}^{m} w_j \left(\frac{a_j^{-1} + x_0}{2}\right) \\ &= \frac{1}{2}\sum_{j=1}^{m} w_j a_j^{-1} + \frac{1}{2} x_0. \end{aligned}$$

Soving this for $x_0 = W(\omega; \mathbb{A})$, we obtain the desired lower bound. This completes the proof. □

4. Final remark

As mentioned in the introduction, $\Phi_{\frac{1}{2}}(a,b)^{\frac{1}{2}}$ is not a metric on Ω [14]. Nonetheless, it shares many interesting properties with usual means as seen in the current work. In the theory of matrix mean, Karcher mean or multivariate geometric mean is the most significant one [10]. However, its symmetric cone version has not been investigated up to now except only in [13]. In [10], an order relation between the Wasserstein mean and Karcher mean by way of the power mean is presented. But, the power mean is not defined on Ω, yet. So with the theory development of Karcher mean and power mean on the symmetric cone Ω, a comparison study of the Wasserstein median with Karcher mean and power mean becomes an interesting future research.

Acknowledgments: This work was conducted during the research year of Chungbuk National University in 2024.

References

[1] S. Amari, *Information Geometry and its Applications*, Springer (Tokyo), 2016.
[2] S. Amari, *Divergence function, information monotonicity and information geometry*, Workshop on Information Theoretic Meth. Sci. Eng., (WITMSE), 2009.
[3] S. Amari and A. Cichocki, *Information geometry of divergence functions*, Bull. Pol. Acad. Sci.Tech. Sci., **58** (2010), 183-195.
[4] R. Bhatia, T. Jain and Y. Lim, *On the BuresWasserstein distance between positive definite matrices*, Expo. Math., **37** (2019), 165-191.

[5] D. Dowson and B. Landau, *The Fréchet distance between multivariate normal distributions*, J. Math. Anal., **12** (1982), 450-455.
[6] J. Faraut and A. Korányi, *Analysis on symmetric cones*, Clarendon Press, Oxford, 1994.
[7] R. Frank and E. Lieb, *Monotonicity of a relative Rényi entropy*, J. Math. Phys., **54** (2013), 122201.
[8] J. Hwang and S. Kim, *Bounds for the Wasserstein mean with applications to the Lie-Trotter mean*, J. Math. Anal. Appl., **475** (2019), 1744-1753.
[9] S. Kim, *Parameterized Wasserstein means*, J. Math. Anal. Appl., **525** (2023), 127272.
[10] S. Kim and H. Lee, *Inequalities of the Wasserstein mean with other matrix means*, Ann. Funct. Anal., **11** (2020), 194-207.
[11] J. Kim and Y. Lim, *Jordan automorphic generators of Euclidean Jordan algebras*, J. Kor. Math. Soc., **43** (2006), 507-528.
[12] M. Knott and C. Smith, *On the optimal mapping of distributions*, J. Optim. Theory Appl., **43** (1984), 39-49.
[13] S. Kum, H. Lee and Y. Lim, *No dice theorem on symmetric cones*, Taiwanese J. Math., **17** (2013), 1967-1982.
[14] S. Kum, Y. Lim and S. Yun, *Divergences on symmetric cones and medians*, Taiwanese J. Math., **26** (2022), 867-886.
[15] H. Lee and Y. Lim, *Metric and spectral geometric means on symmetric cones*, Kyungpook. Math. J., **47** (2007), 133-150.
[16] Y. Lim, *Geometric means on symmetric cones*, Arch. der Math., **75** (2000), 39-45.
[17] Y. Lim, *Applications of geometric means on symmetric cones*, Math. Ann., **319** (2001), 457-468.
[18] Y. Lim, *Birkhoff formula for conformal compressions of symmetric cones*, Amer. J. Math., **125** (2003), 167-182.
[19] I. Olkin and F. Pukelsheim, *The distance between two random vectors with given dispersion matrices*, Linear Algebra Appl., **48** (1982), 257-263.
[20] A. Uhlmann, *Density operators as an arena for differential geometry*, Rep. Math. Phys., **33** (1993), 255-263.
[21] A. Uhlmann, *Transition probability (fidelity) and its relatives*, Found. Phys., **41** (2011), 288-298.
[22] M. Wilde, A. Winter and D. Yang, *Strong converse for the classical capacity of entanglement-breaking and Hadamard channels via a sandwiched Renyi relative entropy,* Comm. Math. Phys., **331** (2014), 593-622.