Annals of Functional Analysis (2020) 11:194–207
https://doi.org/10.1007/s43034-019-00025-0

Tusi Mathematical Research Group

ORIGINAL PAPER

Inequalities of the Wasserstein mean with other matrix means

Sejong Kim[1] · Hosoo Lee[2]

Received: 21 March 2019 / Accepted: 10 June 2019 / Published online: 1 December 2019

Abstract

Recently, a new Riemannian metric and a least squares mean of positive definite matrices have been introduced. They are called the Bures–Wasserstein metric and Wasserstein mean, which are different from the Riemannian trace metric and Karcher mean. In this paper we find relationships of the Wasserstein mean with other matrix means such as the power means, harmonic mean, and Karcher mean.

Keywords Wasserstein mean · Power mean · Karcher mean · Harmonic mean

Mathematics Subject Classification 15B48 · 47B65

1 Introduction

Since Pusz and Woronowicz [17] have introduced two variable geometric mean of positive definite matrices, a variety of different schemes to construct multivariate geometric means have been developed in the settings of positive matrices and positive operators. The natural and canonical mean among those multivariate geometric means on the open convex cone $\mathbb{P}_k$ of $k \times k$ positive definite matrices is the least squares mean, denoted as $\Lambda(\omega; A_1, \dots, A_n)$, which is the unique point in $\mathbb{P}_k$ minimizing the weighted sum of squares of the Riemannian trace metrics to each variables $A_1, \dots, A_n \in \mathbb{P}_k$:

Communicated by Takeaki Yamazaki.

✉ Hosoo Lee
hosoo@jejunu.ac.kr

Sejong Kim
skim@chungbuk.ac.kr

[1] Department of Mathematics, Chungbuk National University, Cheongju 28644, South Korea

[2] Elementary Education Research Institute, Jeju National University, Jeju 63294, South Korea

$$\Lambda(\omega; A_1, \ldots, A_n) = \underset{X \in \mathbb{P}_k}{\arg\min} \sum_{j=1}^{n} w_j \delta^2(X, A_j),$$

where $\omega = (w_1, \ldots, w_n)$ is a positive probability vector and

$$\delta(A, B) = \| \log A^{-1/2} B A^{-1/2} \|_2$$

is the Riemannian trace metric between A and B. This has been anticipated by Élie Cartan [3, Section 6.1.5] and Karcher [13] has shown that $\Lambda(\omega; A_1, \ldots, A_n)$ coincides with the unique positive definite solution $X \in \mathbb{P}_k$ of the Karcher equation $\sum_{j=1}^{n} w_j \log(X^{-1/2} A_j X^{-1/2}) = O$. Many interesting properties including monotonicity and the extension theory of Karcher mean to positive definite operators have been derived from the Karcher equation [15], so we call $\Lambda(\omega; A_1, \ldots, A_n)$ the Karcher mean. Moreover, the theory of means has been developed by Sturm [18] to the setting of probability measures with finite first moment on the Hadamard space, which is a complete metric space satisfying the semi-parallelogram law.

Many research topics about the Karcher mean such as finding properties, computational algorithms, and extending to positive operators have been widely studied [5,11,13,15,16]. Among lots of interesting approaches to the Karcher mean, Lim and Pálfia [16] have introduced power means of positive definite matrices $P_t(\omega; A_1, \ldots, A_n)$ of order $t \in [-1, 1] \setminus \{0\}$ as the unique positive definite solution of the following non-linear equation:

$$X = \sum_{i=1}^{n} w_i (X \#_t A_i), \ \ t \in (0, 1]$$

$$X = \left[\sum_{i=1}^{n} w_i (X^{-1} \#_{-t} A_i^{-1}) \right]^{-1}, \ \ t \in [-1, 0),$$

where $A \#_t B = A^{1/2}(A^{-1/2} B A^{-1/2})^t A^{1/2}$ is the weighted geometric mean of positive definite matrices A and B. They have shown that the Karcher mean $\Lambda(\omega; A_1, \ldots, A_n)$ is the limit of the power means $P_t(\omega; A_1, \ldots, A_n)$ as $t \to 0$. This plays an important role in proving the monotonicity of Karcher mean and extending to positive invertible operators.

Motivated from barycenters in the Wasserstein space of Gaussian distributions [1,2], Bhatia, Jain and Lim [6] have developed the Wasserstein metric d and Wasserstein mean Ω on the open convex cone $\mathbb{P}_k$ with matrix analysis.

$$\Omega(\omega; A_1, \ldots, A_n) = \underset{X \in \mathbb{P}_k}{\arg\min} \sum_{j=1}^{n} w_j d^2(X, A_j),$$

where

$$d(A, B) = \left[\left(\frac{A + B}{2} \right) - (A^{1/2} B A^{1/2})^{1/2} \right]^{1/2}$$

is the Bures–Wasserstein distance between A and B. Although the Wasserstein mean does not satisfy the monotonicity, determinantal identity, and Wasserstein-harmonic mean inequality, it still gives interesting results such as the Bures–Wasserstein distance is a Riemannian metric and the Wasserstein mean satisfies the arithmetic-Wasserstein mean inequality and the Lie–Trotter–Kato formula. See [6,12] for more details.

The main goal of this paper is to find relationships between the Wasserstein mean and other matrix means in terms of trace and Loewner order. In Sect. 2 we recall the Bures–Wasserstein distance and Wasserstein mean with known properties, and in Sect. 3 we show the relationships between the Wasserstein mean and power mean, and between the Wasserstein mean and Karcher mean. Finally, we see in Sect. 4 the order relation between the Wasserstein mean and the harmonic mean.

2 Wasserstein mean

Let $\mathbb{H}_k$ be the real vector space of all $k \times k$ Hermitian matrices. Let $\mathbb{P}_k \subset \mathbb{H}_k$ be the open convex cone of all $k \times k$ positive definite matrices. The general linear group GL_k of all $k \times k$ invertible matrices acts on $\mathbb{P}_k$ via congruence transformations $\Gamma_M(X) = MXM^*$ for $M \in GL_k$ and $X \in \mathbb{P}_k$. We denote as $\mathbb{U}_k$ the compact group of all $k \times k$ unitary matrices. For any $A, B \in \mathbb{H}_k$ we write $A \leq B$ if $B - A$ is positive semi-definite, and $A < B$ if $B - A$ is positive definite. This is indeed a partial order on $\mathbb{H}_k$, known as the Loewner order.

A new metric defined by

$$d(A, B) = \left[\left(\frac{A + B}{2} \right) - (A^{1/2} B A^{1/2})^{1/2} \right]^{1/2}$$

and the unique geodesic for this metric on the open convex cone $\mathbb{P}_m$ of positive definite matrices have been recently introduced in [6]. This metric is the matrix version of the Hellinger distance

$$d(\overrightarrow{p}, \overrightarrow{q}) = \left[\frac{1}{2} \sum_{i=1}^{n} (\sqrt{p_i} - \sqrt{q_i})^2 \right]^{1/2}$$

for two probability distributions $\overrightarrow{p} = (p_1, \ldots, p_n)$ and $\overrightarrow{q} = (q_1, \ldots, q_n)$. Moreover, it coincides with the Bures distance of density matrices in quantum information theory and the Wasserstein metric in statistics and the theory of optimal transport. The Bures–Wasserstein metric is a Riemannian metric induced by the inner product

$$\langle X, Y \rangle_A = \sum_{i,j=1}^{k} \frac{\alpha_i \mathrm{Re}(\overline{x_{ji}} y_{ji})}{(\alpha_i + \alpha_j)^2}$$

for any $X = [x_{ij}]$ and $Y = [y_{ij}]$ on the tangent space $T_A \mathbb{P}_k \equiv \mathbb{H}_k$ for each $A \in \mathbb{P}_k$, where $\alpha_1, \ldots, \alpha_k$ are positive eigenvalues of $A \in \mathbb{P}_k$. The unique geodesic connecting from A to B for the Bures–Wasserstein distance is given by

$$A \diamond_t B := (1-t)^2 A + t^2 B + t(1-t)\left[(AB)^{1/2} + (BA)^{1/2}\right], \; t \in [0,1].$$

Let $\mathbb{A} = (A_1, \ldots, A_n) \in \mathbb{P}_k^n$, and let $\omega = (w_1, \ldots, w_n) \in \Delta_n$, the simplex of all positive probability vectors in $\mathbb{R}^n$. We consider the following minimization problem

$$\underset{X \in \mathbb{P}_k}{\arg\min} \sum_{j=1}^{n} w_j d^2(X, A_j), \tag{2.1}$$

where d is the Bures–Wasserstein distance on $\mathbb{P}_k$. By using tools from non-smooth analysis, convex duality, and the optimal transport theory, it has been proved in [1, Theorem 6.1] that the above minimization problem has a unique solution in $\mathbb{P}_k$. On the other hand, it has been shown in [6] that the objective function $f(X) = \sum_{j=1}^{n} w_j d^2(X, A_j)$ is strictly convex on $\mathbb{P}_k$, by applying the strict concavity of the map $h : \mathbb{P}_k \to \mathbb{R}$, $h(X) = \mathrm{Tr}(X^{1/2})$. Therefore, we define such a unique minimizer of (2.1) as the *Wasserstein mean*, denoted by $\Omega(\omega; \mathbb{A})$. That is,

$$\Omega(\omega; \mathbb{A}) = \underset{X \in \mathbb{P}_k}{\arg\min} \sum_{j=1}^{n} w_j d^2(X, A_j).$$

To find the unique minimizer of objective function $f : \mathbb{P}_k \to \mathbb{R}$, we evaluate the derivative $Df(X)$ and set it equal to zero. By using matrix differential calculus, we have the following.

Theorem 2.1 [6, Theorem 8] *The Wasserstein mean $\Omega(\omega; \mathbb{A})$ is a unique solution $X \in \mathbb{P}_k$ of the nonlinear matrix equation*

$$I = \sum_{j=1}^{n} w_j (A_j \# X^{-1}),$$

equivalently,

$$X = \sum_{j=1}^{n} w_j (X^{1/2} A_j X^{1/2})^{1/2}.$$

Remark 2.2 It has been shown in [6] that

$$\Omega(1-t, t; A, B) = (1-t)^2 A + t^2 B + t(1-t)\left[(AB)^{1/2} + (BA)^{1/2}\right],$$

and it does not hold the monotonicity and the Wasserstein-harmonic mean inequality. So the Wasserstein mean $\Omega(1-t, t; A, B)$ is different from the usual geometric mean $A\#B$, even for multivariate cases.

For given $\mathbb{A} = (A_1, \ldots, A_n) \in \mathbb{P}_k^n$, any permutation σ on $\{1, \ldots, n\}$, and any invertible matrix M, we denote as

$$\begin{aligned} \mathbb{A}_\sigma &= (A_{\sigma(1)}, \ldots, A_{\sigma(n)}) \in \mathbb{P}_k^n, \\ M\mathbb{A}M^* &= (MA_1M^*, \ldots, MA_nM^*) \in \mathbb{P}_k^n, \\ \mathbb{A}^r &= (\underline{A_1, \ldots, A_n}, \ldots, \underline{A_1, \ldots, A_n}) \in \mathbb{P}_k^{nr}, \end{aligned}$$

where the number of blocks in the last expression is r. For given $\omega = (w_1, \ldots, w_n) \in \Delta_n$, we also denote as

$$\begin{aligned} \omega_\sigma &= (w_{\sigma(1)}, \ldots, w_{\sigma(n)}) \in \Delta_n, \\ \omega^r &= \frac{1}{r}(\underline{w_1, \ldots, w_n}, \ldots, \underline{w_1, \ldots, w_n}) \in \Delta_{nr}. \end{aligned}$$

The following are some properties of the Wasserstein mean: see [6,12].

Lemma 2.3 *Let $\mathbb{A} = (A_1, \ldots, A_n) \in \mathbb{P}_k^n$, and let $\omega = (w_1, \ldots, w_n) \in \Delta_n$. Then the following are satisfied.*

(1) (*Consistency with scalars*) $\Omega(\omega; \mathbb{A}) = \left[\sum_{j=1}^n w_j A_j^{1/2}\right]^2$ *if the A_j's commute.*
(2) (*Homogeneity*) $\Omega(\omega; \alpha\mathbb{A}) = \alpha\Omega(\omega; \mathbb{A})$ *for any* $\alpha > 0$.
(3) (*Permutation invariancy*) $\Omega(\omega_\sigma; \mathbb{A}_\sigma) = \Omega(\omega; \mathbb{A})$ *for any permutation σ on* $\{1, \ldots, n\}$.
(4) (*Repetition invariancy*) $\Omega(\omega^k; \mathbb{A}^k) = \Omega(\omega; \mathbb{A})$ *for any* $k \in \mathbb{N}$.
(5) (*Unitary congruence invariancy*) $\Omega(\omega; U\mathbb{A}U^*) = U\Omega(\omega; \mathbb{A})U^*$ *for any* $U \in \mathbb{U}_m$.
(6) (*Arithmetic-Wasserstein mean inequality*) $\Omega(\omega; \mathbb{A}) \leq \sum_{j=1}^n w_j A_j$.

Moreover, $X = \Omega(\omega; A_1, \ldots, A_{n-1}, X)$ if and only if $X = \Omega(\hat{\omega}; A_1, \ldots, A_{n-1})$, where $\hat{\omega} = \frac{1}{1-w_n}(w_1, \ldots, w_{n-1}) \in \Delta_{n-1}$.

Proof All items (1)-(6) have been proved in [6,12].

Let $X = \Omega(\omega; A_1, \ldots, A_{n-1}, X)$. Then it is equivalent by Theorem 2.1 to

$$I = \sum_{j=1}^{n-1} w_j(A_j \# X^{-1}) + w_n(X \# X^{-1}) = \sum_{j=1}^{n-1} w_j(A_j \# X^{-1}) + w_n I.$$

So it is simplified to $I = \sum_{j=1}^{n-1} \frac{w_j}{1-w_n}(A_j \# X^{-1})$, and thus, $X = \Omega(\hat{\omega}; A_1, \ldots, A_{n-1})$ by Theorem 2.1. □

Lemma 2.4 *Let $\omega = (w_1, \ldots, w_n) \in \Delta_n$ and $\mathbb{A} = (A_1, \ldots, A_n) \in \mathbb{P}_k^n$ with $0 < mI \leq A_j \leq MI$ for all j and some positive scalars m, M. Then $mI \leq \Omega(\omega; \mathbb{A}) \leq MI$.*

Proof Assume that $0 < mI \leq A_j \leq MI$ for all $j = 1, \ldots, n$. Let $X = \Omega(\omega; \mathbb{A})$. Since the congruence transformation and square root map preserve the Loewner order, we have $mX \le X^{1/2} A_j X^{1/2} \le MX$, and $\sqrt{m} X^{1/2} \le (X^{1/2} A_j X^{1/2})^{1/2} \le \sqrt{M} X^{1/2}$. Then

$$\sqrt{m} X^{1/2} \le \sum_{j=1}^{n} w_j (X^{1/2} A_j X^{1/2})^{1/2} \le \sqrt{M} X^{1/2}.$$

So $\sqrt{m} X^{1/2} \le X \le \sqrt{M} X^{1/2}$ by Theorem 2.1, and $\sqrt{m} I \le X^{1/2} \le \sqrt{M} I$. Hence, $mI \le X \le MI$. □

An iteration approach to the Wasserstein mean has been recently shown in [2] by using the map $K : \mathbb{P}_k \to \mathbb{P}_k$ defined as

$$K(A) = A^{-1/2} \left[\sum_{j=1}^{n} w_j (A^{1/2} A_j A^{1/2})^{1/2} \right]^2 A^{-1/2}. \tag{2.2}$$

for each $A \in \mathbb{P}_k$.

Theorem 2.5 [2] *Let $\omega \in \Delta_n$ and $\mathbb{A} \in \mathbb{P}_k^n$. For every $S_0 \in \mathbb{P}_k$ the sequence $S_{r+1} = K(S_r)$ constructed iteratively from the map K in (2.2) converges to $\Omega(\omega; \mathbb{A})$, and for all natural numbers r*

$$S_r \le S_{r+1} \le \Omega(\omega; \mathbb{A}).$$

3 Power means and Karcher mean

It is well known from [18] that the least squares mean, the Cartan mean or Karcher mean, uniquely exists in the Hadamard space. The Karcher mean of positive definite matrices $A_1, \ldots, A_n$ with a positive probability vector $\omega = (w_1, \ldots, w_n)$ is defined as the unique point $\Lambda(\omega; A_1, \ldots, A_n)$ that minimizes the variance function $f(X) = \sum_{j=1}^{n} w_j \delta^2(X, A_j)$.

$$\Lambda(\omega; A_1, \ldots, A_n) = \underset{X \in \mathbb{P}_m}{\arg\min} \sum_{j=1}^{n} w_j \delta^2(X, A_j).$$

Vanishing the gradient of objective function $f(X) = \sum_{j=1}^{n} w_j \delta^2(X, A_j)$ we obtain that the Karcher mean $\Lambda(\omega; A_1, \ldots, A_n)$ coincides with the unique solution $X \in \mathbb{P}_k$ of nonlinear matrix equation, called the Karcher equation,

$$\sum_{j=1}^{n} w_j \log(X^{-1/2} A_j X^{-1/2}) = O.$$

The Karcher mean of positive definite matrices is a currently active research topic in many areas such as matrix analysis, optimization, numerical computation, and operator theory. Among a variety of approaches to the Karcher mean, Lim and Pálfia have introduced in [16] a powerful theory of power means. The matrix power mean $P_t(\omega; A_1, \ldots, A_n)$ for $t \in (0, 1]$ is defined as the unique solution $X \in \mathbb{P}_m$ of the following equation

$$X = \sum_{j=1}^{n} w_j X \#_t A_j.$$

Indeed, the map $g : \mathbb{P}_k \to \mathbb{P}_k,\ g(X) = \sum_{j=1}^{n} w_j X \#_t A_j$ for $t \in (0, 1]$ is a strict contraction for the Thompson metric $d_T(A, B) = \| \log A^{-1/2} B A^{-1/2} \|$, where $\| \cdot \|$ denotes the operator norm. Therefore,

$$\lim_{k \to \infty} g^k(Z) = P_t(\omega; A_1, \ldots, A_n) \quad \text{for any } Z \in \mathbb{P}_k.$$

For $t \in [-1, 0)$ we define $P_t(\omega; A_1, \ldots, A_n) = P_{-t}(\omega; A_1^{-1}, \ldots, A_n^{-1})^{-1}$. The most remarkable consequence of matrix power means is that matrix power means $P_t(\omega; A_1, \ldots, A_n)$ converges to the Karcher mean $\Lambda(\omega; A_1, \ldots, A_n)$ as $t \to 0$. This plays an important role to construct the Karcher mean of positive invertible operators: see [15]. Furthermore, for $0 \le s \le t \le 1$

$$\begin{aligned} &\left[\sum_{j=1}^{n} w_j A_j^{-1} \right]^{-1} \\ &= P_{-1} \le P_{-t} \le P_{-s} \le \cdots \le \Lambda \le \cdots \le P_s \le P_t \le P_1 = \sum_{j=1}^{n} w_j A_j. \end{aligned} \tag{3.1}$$

In this section we investigate the relationship of Wasserstein mean with power means and Karcher mean of positive definite matrices.

Theorem 3.1 *Let $\omega = (w_1, \ldots, w_n) \in \Delta_n$ and $\mathbb{A} = (A_1, \ldots, A_n) \in \mathbb{P}_k^n$. For any $0 < t \le 1/2$,*

$$P_t(\omega; \mathbb{A}) \le \Omega(\omega; \mathbb{A}).$$

There are two different approaches of the proof for Theorem 3.1. One method is the use of monotonicity of matrix power means P_t for parameters, and another method is the use of iteration approach for the Wasserstein mean in Theorem 2.5.

Proof 1 By Theorem 2 in [7] with the Schatten 1-norm, we have

$$Q_{1/2}(\omega; \mathbb{A}) \le \Omega(\omega; \mathbb{A}),$$

where $Q_p(\omega; \mathbb{A}) = \left(\sum_{j=1}^n w_j A_j^p\right)^{1/p}$ for any $p \in (-\infty, \infty)$ is the weighted quasi-arithmetic mean of $A_1, \ldots, A_n$. Moreover, it has been proved in [8] that

$$P_t(\omega; \mathbb{A}) \le Q_t(\omega; \mathbb{A})$$

for any $t \in (0, 1)$. Since the power means P_t is monotone for parameters by (3.1), we conclude that $P_t \le P_{1/2} \le Q_{1/2} \le \Omega$ for $0 < t \le 1/2$. □

Proof 2 Let $X_0 = P_t(\omega; \mathbb{A})^{-1}$ for $0 < t \le 1/2$. Then by the affine property of parameters: $(A\#_s B)\#_t(A\#_u B) = A\#_{(1-t)s+tu}B$ for any $s, t, u \in [0, 1]$, and by the two-variable arithmetic-geometric mean inequality,

$$\begin{aligned} X_0^{-1} &= \sum_{j=1}^n w_j X_0^{-1}\#_t A_j = \sum_{j=1}^n w_j \left[X_0^{-1}\#_{2t}(X_0^{-1}\# A_j)\right] \\ &\le \sum_{j=1}^n w_j \left[(1-2t)X_0^{-1} + 2t(X_0^{-1}\# A_j)\right] = (1-2t)X_0^{-1} + 2t\sum_{j=1}^n w_j X_0^{-1}\# A_j. \end{aligned}$$

A simple calculation yields that $X_0^{-1} \le \sum_{j=1}^n w_j X_0^{-1}\# A_j$. By applying the congruence transformation via $X_0^{1/2}$, we obtain

$$I \le \sum_{j=1}^n w_j (X_0^{1/2} A_j X_0^{1/2})^{1/2}.$$

Taking square on both sides yields $I \le \left[\sum_{j=1}^n w_j (X_0^{1/2} A_j X_0^{1/2})^{1/2}\right]^2$. By applying the congruence transformation via $X_0^{-1/2}$ we obtain

$$X_0^{-1} \le X_0^{-1/2}\left[\sum_{j=1}^n w_j (X_0^{1/2} A_j X_0^{1/2})^{1/2}\right]^2 X_0^{-1/2} = K(X_0).$$

Therefore, by Theorem 2.5

$$P_t(\omega; \mathbb{A}) = X_0^{-1} \le K(X_0) = X_1 \le \Omega(\omega; \mathbb{A}).$$

□

Corollary 3.2 *Let $\omega = (w_1, \ldots, w_n) \in \Delta_n$ and $\mathbb{A} = (A_1, \ldots, A_n) \in \mathbb{P}_k^n$. For any $-1/2 \le t < 0$,*

$$P_t(\omega; \mathbb{A}) > k\left[\Omega(\omega; \mathbb{A}^{-1})\right]^{-1},$$

where $\mathbb{A}^{-1} = (A_1^{-1}, \ldots, A_n^{-1}) \in \mathbb{P}_k^n$.

Proof Note that $P_t(\omega; \mathbb{A}) = P_{-t}(\omega; \mathbb{A}^{-1})^{-1}$ for any $-1/2 \le t < 0$. By Theorem 3.1 we have

$$\left[P_t(\omega; \mathbb{A})^{-1}\right] = P_{-t}(\omega; \mathbb{A}^{-1}) \le \Omega(\omega; \mathbb{A}^{-1}).$$

Since $[A^{-1}] > k[A]^{-1}$ for any $A \in \mathbb{P}_k$, the above inequality implies that

$$k\,[P_t(\omega; \mathbb{A})]^{-1} < \Omega(\omega; \mathbb{A}^{-1}).$$

Therefore, we obtain the desired inequality. □

Remark 3.3 Let

$$A = \begin{bmatrix} 1 & 2 \\ 2 & 5 \end{bmatrix}, \qquad B = \begin{bmatrix} 4 & 4 \\ 4 & 5 \end{bmatrix}.$$

One can see easily that A, B are positive definite and $AB \ne BA$. The Wasserstein mean $\Omega\left(\frac{1}{2}, \frac{1}{2}; A, B\right) = A \diamond B$ and the Karcher mean $\Lambda\left(\frac{1}{2}, \frac{1}{2}; A, B\right) = A\#B$ of positive definite matrices A and B, respectively, are

$$\Omega\left(\frac{1}{2}, \frac{1}{2}; A, B\right) = \frac{1}{4}\begin{bmatrix} 9 & 12 \\ 12 & 20 \end{bmatrix}, \qquad \Lambda\left(\frac{1}{2}, \frac{1}{2}; A, B\right) = \begin{bmatrix} 1.6641 & 2.2188 \\ 2.2188 & 4.1603 \end{bmatrix}.$$

One can see that there is no order relation between the Wasserstein mean $\Omega\left(\frac{1}{2}, \frac{1}{2}; A, B\right)$ and the Karcher mean $\Lambda\left(\frac{1}{2}, \frac{1}{2}; A, B\right)$.

Proposition 3.4 *Let $\omega = (w_1, \ldots, w_n) \in \Delta_n$ and $\mathbb{A} = (A_1, \ldots, A_n) \in \mathbb{P}_k^n$. Then*

$$k\left[\Omega(\omega; \mathbb{A}^{-1})\right]^{-1} \le \Lambda(\omega; \mathbb{A}) \le \Omega(\omega; \mathbb{A}).$$

Proof Note that

$$\lim_{t \to 0} P_t(\omega; \mathbb{A}) = \Lambda(\omega; \mathbb{A}).$$

Taking the limit as $t \to 0^+$ in the result of Theorem 3.1 and applying the continuity of the trace map, we obtain $\Lambda(\omega; \mathbb{A}) \le \Omega(\omega; \mathbb{A})$. Moreover, taking the limit as $t \to 0^-$ in the result of Corollary 3.2 yields that $\Lambda(\omega; \mathbb{A}) \ge k\left[\Omega(\omega; \mathbb{A}^{-1})\right]^{-1}$. □

The following shows the relation between the Wasserstein mean and the matrix power mean under certain assumption.

Theorem 3.5 *Let $\omega = (w_1, \ldots, w_n) \in \Delta_n$ and $\mathbb{A} = (A_1, \ldots, A_n) \in \mathbb{P}_k^n$. For any $1/2 \le t \le 1$, $\Omega(\omega; \mathbb{A}) \ge I$ implies $\Omega(\omega; \mathbb{A}) \ge P_t(\omega; \mathbb{A})^{-1}$.*

Proof Let $X = \Omega(\omega; \mathbb{A})^{-1}$. Then $X \leq I$ by assumption. By Theorem 2.1, the affine property of parameters, and the two-variable arithmetic-geometric mean inequality, we have

$$\begin{aligned} I &= \sum_{j=1}^{n} w_j (X \# A_j) = \sum_{j=1}^{n} w_j \left[X \#_{\frac{1}{2t}} (X \#_t A_j) \right] \\ &\leq \sum_{j=1}^{n} w_j \left[\left(1 - \frac{1}{2t} \right) X + \frac{1}{2t} (X \#_t A_j) \right] \\ &= \left(1 - \frac{1}{2t} \right) X + \frac{1}{2t} \sum_{j=1}^{n} w_j (X \#_t A_j). \end{aligned}$$

Since $X \leq I$, we have

$$X \leq 2tI - (2t - 1)X \leq \sum_{j=1}^{n} w_j (X \#_t A_j).$$

Since the map $f : \mathbb{P}_k \to \mathbb{P}_k$ defined by $f(Z) = \sum_{j=1}^{n} w_j (Z \#_t A_j)$ is monotone increasing, we get $X \leq f(X) \leq f^2(X) \leq \cdots \leq f^r(X)$ for all $r \geq 1$. Note that f is a strict contraction with respect to the Thompson metric, and as $r \to \infty$ $f^r(Z)$ converges to a unique fixed point, which is the power mean, for any $Z \in \mathbb{P}_k$ by the Banach fixed point theorem. So we obtain

$$X \leq \lim_{r \to \infty} f^r(X) = P_t(\omega; \mathbb{A}).$$

Therefore, $\Omega(\omega; \mathbb{A}) = X^{-1} \geq P_t(\omega; \mathbb{A})^{-1}$. □

Corollary 3.6 *Let* $\omega = (w_1, \ldots, w_n) \in \Delta_n$ *and* $\mathbb{A} = (A_1, \ldots, A_n) \in \mathbb{P}_k^n$. *For any* $-1 \leq t \leq -1/2$, $\Omega(\omega; \mathbb{A}) \geq I$ *implies* $\Omega(\omega; \mathbb{A}) \geq P_t(\omega; \mathbb{A}^{-1}) \geq \left[\sum_{j=1}^{n} w_j A_j \right]^{-1}$.

Proof Assume that $\Omega(\omega; \mathbb{A}) \geq I$. Since $P_{-t}(\omega; \mathbb{A})^{-1} = P_t(\omega; \mathbb{A}^{-1})$ for any $-1 \leq t \leq -1/2$, we obtain from Theorem 3.5 and the monotonicity of power means for parameters in (3.1) that

$$\Omega(\omega; \mathbb{A}) \geq P_t(\omega; \mathbb{A}^{-1}) \geq P_{-1}(\omega; \mathbb{A}^{-1}) = \left[\sum_{j=1}^{n} w_j A_j \right]^{-1}.$$

□

4 Inequalities with harmonic means

In this section we investigate the inequality relationship between the Wasserstein mean and the harmonic mean $\mathcal{H}(\omega; A_1, \ldots, A_n) = \left[\sum_{j=1}^n w_j A_j^{-1}\right]^{-1}$.

For $1 \leq i, j \leq n$ let $A_{ij} \in M_k$, the set of all $k \times k$ matrices with entries in the field of complex numbers. We define a map $\Phi : M_n(M_k) \to M_k$ as

$$\Phi\left(\begin{bmatrix} A_{11} & \cdots & A_{1n} \\ \vdots & \ddots & \vdots \\ A_{n1} & \cdots & A_{nn} \end{bmatrix}\right) = \sum_{j=1}^n w_j A_{jj}.$$

Then one can easily see that Φ is a unital positive linear map.

Theorem 4.1 *Let $\omega = (w_1, \ldots, w_n) \in \Delta_n$ and $\mathbb{A} = (A_1, \ldots, A_n) \in \mathbb{P}_k^n$. If there exist positive scalars M and m such that $0 < mI \leq A_j \leq MI$ for all j, then*

$$\frac{4Mm}{(M+m)^2} I \leq \mathcal{H}(\omega; X^{-1}\#A_1, \ldots, X^{-1}\#A_n) \leq I,$$

where $X = \Omega(\omega; A_1, \ldots, A_n)$.

Proof Let $X = \Omega(\omega; A_1, \ldots, A_n)$. Then by the arithmetic-harmonic mean inequality

$$\mathcal{H}(\omega; X^{-1}\#A_1, \ldots, X^{-1}\#A_n) \leq \sum_{j=1}^n w_j X^{-1}\#A_j = I.$$

If there exist positive scalars M and m such that $0 < mI \leq A_j \leq MI$ for all j, then $mI \leq X = \Omega(\omega; A_1, \ldots, A_n) \leq MI$ by Lemma 2.4, and $mI \leq (X^{1/2}A_jX^{1/2})^{1/2} \leq MI$ for all j. So

$$mI \leq \begin{bmatrix} (X^{1/2}A_1X^{1/2})^{1/2} & \cdots & O \\ \vdots & \ddots & \vdots \\ O & \cdots & (X^{1/2}A_nX^{1/2})^{1/2} \end{bmatrix} \leq MI.$$

Applying Proposition 2.7.8 in [4] to the unital positive linear map Φ, we obtain

$$\Phi\left(\begin{bmatrix} (X^{1/2}A_1X^{1/2})^{1/2} & \cdots & O \\ \vdots & \ddots & \vdots \\ O & \cdots & (X^{1/2}A_nX^{1/2})^{1/2} \end{bmatrix}\right)$$
$$\leq \frac{(M+m)^2}{4Mm} \Phi\left(\begin{bmatrix} (X^{1/2}A_1X^{1/2})^{1/2} & \cdots & O \\ \vdots & \ddots & \vdots \\ O & \cdots & (X^{1/2}A_nX^{1/2})^{1/2} \end{bmatrix}^{-1}\right)^{-1}.$$

Equivalently, by Theorem 2.1

$$X = \sum_{j=1}^{n} w_j (X^{1/2} A_j X^{1/2})^{1/2} \leq \frac{(M+m)^2}{4Mm} \left[\sum_{j=1}^{n} w_j (X^{1/2} A_j X^{1/2})^{-1/2} \right]^{-1}.$$

Taking the congruence transformation by $X^{-1/2}$ on both sides, we obtain

$$\begin{aligned} I &\leq \frac{(M+m)^2}{4Mm} X^{-1/2} \left[\sum_{j=1}^{n} w_j (X^{1/2} A_j X^{1/2})^{-1/2} \right]^{-1} X^{-1/2} \\ &= \frac{(M+m)^2}{4Mm} \left[\sum_{j=1}^{n} w_j X \# A_j^{-1} \right]^{-1} = \frac{(M+m)^2}{4Mm} \mathcal{H}(\omega; X^{-1} \# A_1, \ldots, X^{-1} \# A_n). \end{aligned}$$

□

Remark 4.2 Note that the constant $\frac{(M+m)^2}{4Mm}$ appeared in Theorem 4.1 is known as the Kantorovich constant. It plays an important role in the reverse inequalities of the weighted arithmetic, geometric and harmonic means: see [9,14].

The notions of operator convexity and concavity are characterized by Jensen type inequalities in [10]. For every contraction X we have

$$(X^* A X)^p \leq X^* A^p X \quad \text{if } 1 \leq p \leq 2,$$

and

$$(X^* A X)^p \geq X^* A^p X \quad \text{if } 0 \leq p \leq 1. \tag{4.1}$$

For $X \in GL_k$ such that its inverse X^{-1} is a contraction,

$$(X^* A X)^p \leq X^* A^p X \quad \text{if } 0 \leq p \leq 1. \tag{4.2}$$

Indeed, applying (4.1) to the contraction X^{-1} and $X^* A X$, we obtain

$$A^p = ((X^{-1})^* (X^* A X) X^{-1})^p \geq (X^{-1})^* (X^* A X)^p X^{-1},$$

and hence, $(X^* A X)^p \leq X^* A^p X$.

Theorem 4.3 *Let* $\omega = (w_1, \ldots, w_n) \in \Delta_n$ *and* $\mathbb{A} = (A_1, \ldots, A_n) \in \mathbb{P}_k^n$.

(1) *If* $\Omega(\omega; A_1, \ldots, A_n) \geq I$, *then* $\Omega(\omega; A_1, \ldots, A_n) \geq \mathcal{H}(\omega; A_1^{-1/2}, \ldots, A_n^{-1/2})$.
(2) *If* $\Omega(\omega; A_1, \ldots, A_n) \leq I$, *then* $\Omega(\omega; A_1, \ldots, A_n) \leq \mathcal{H}(\omega; A_1^{-1/2}, \ldots, A_n^{-1/2})$.

Proof Let $X = \Omega(\omega; A_1, \ldots, A_n)$.

(1) By assumption $I \le X = \sum_{j=1}^n w_j (X^{1/2} A_j X^{1/2})^{1/2}$. By the inequality (4.2) we have

$$\begin{aligned} I &\le \sum_{j=1}^n w_j (X^{1/2} A_j X^{1/2})^{1/2} \\ &\le \sum_{j=1}^n w_j X^{1/2} A_j^{1/2} X^{1/2} = X^{1/2} \left[\sum_{j=1}^n w_j A_j^{1/2} \right] X^{1/2}. \end{aligned}$$

Taking the congruence transformation by $X^{-1/2}$ on both sides, we obtain $X^{-1} \le \sum_{j=1}^n w_j A_j^{1/2}$. Thus, by taking inverse on both sides we get

$$X \ge \left[\sum_{j=1}^n w_j A_j^{1/2} \right]^{-1} = \mathcal{H}(\omega; A_1^{-1/2}, \ldots, A_n^{-1/2}).$$

(2) By assumption $I \ge X = \sum_{j=1}^n w_j (X^{1/2} A_j X^{1/2})^{1/2}$. By the inequality (4.1) we have

$$I \ge \sum_{j=1}^n w_j X^{1/2} A_j^{1/2} X^{1/2} = X^{1/2} \left[\sum_{j=1}^n w_j A_j^{1/2} \right] X^{1/2}.$$

Taking the congruence transformation by $X^{-1/2}$ on both sides, we obtain $X^{-1} \ge \sum_{j=1}^n w_j A_j^{1/2}$. Thus,

$$X \le \left[\sum_{j=1}^n w_j A_j^{1/2} \right]^{-1} = \mathcal{H}(\omega; A_1^{-1/2}, \ldots, A_n^{-1/2}).$$

□

Remark 4.4 Theorem 4.3 (2) can not be proved directly by Theorem 4.3 (1), since the Wasserstein mean Ω is not invariant under inversion.

Acknowledgements The work of S. Kim was supported by the National Research Foundation of Korea (NRF) Grant funded by the Korea government (MSIT) (No. NRF-2018R1C1B6001394). The work of H. Lee was supported by the National Research Foundation of Korea (NRF) Grant funded by the Korea government (MIST) (No. NRF-2018R1D1A1B07049948).

References

1. Agueh, M., Carlier, G.: Barycenters in the Wasserstein space. SIAM J. Math. Anal. Appl. **43**, 904–924 (2011)

2. Alvarez-Esteban, P.C., del Barrio, E., Cuesta-Albertos, J.A., Matran, C.: A fixed point approach to barycenters in Wasserstein spaces. J. Math. Anal. Appl. **441**, 744–762 (2016)
3. Berger, M.: A Panoramic View of Riemannian Geometry. Springer, Berlin (2003)
4. Bhatia, R.: Positive Definite Matrices, Princeton Series in Applied Mathematics. Princeton University Press, Princeton (2007)
5. Bhatia, R., Holbrook, J.: Riemannian geometry and matrix geometric means. Linear Algebra Appl. **413**, 594–618 (2006)
6. Bhatia, R., Jain, T., Lim, Y.: On the Bures-Wasserstein distance between positive definite matrices. Expo. Math. **37**(2), 165–191 (2019)
7. Bhatia, R., Jain, T., Lim, Y.: Inequalities for the Wasserstein mean of positive definite matrices. Linear Algebra Appl. **576**, 108–123 (2019)
8. Bhatia, R., Lim, Y., Yamazaki, T.: Some norm inequalities for matrix means. Linear Algebra Appl. **501**, 112–122 (2016)
9. Fujii, J.I., Fujii, M., Nakamura, M., Pečarić, J., Seo, Y.: A reverse inequality for the weighted geometric mean due to Lawson–Lim. Linear Algebra Appl. **427**, 272–284 (2007)
10. Hansen, F., Pedersen, G.K.: Jensens inequality for operators and Löwners theorem. Math. Ann. **258**, 229241 (1982)
11. Hiai, F., Lim, Y.: Geometric mean flows and the Cartan barycenter on the Wasserstein space over positive definite matrices. Linear Algebra Appl. **533**, 118–131 (2017)
12. Hwang, J., Kim, S.: Bounds for the Wasserstein mean with applications to the Lie–Trotter mean. J. Math. Anal. Appl. **475**, 1744–1753 (2019)
13. Karcher, H.: Riemannian center of mass and mollifier smoothing. Commun. Pure Appl. Math. **30**, 509–541 (1977)
14. Kim, S., Lim, Y.: A converse inequality of higher order weighted arithmetic and geometric means of positive definite operators. Linear Algebra Appl. **426**, 490–496 (2007)
15. Lawson, J., Lim, Y.: Karcher means and Karcher equations of positive definite operators. Trans. Am. Math. Soc. Ser. B **1**, 1–22 (2014)
16. Lim, Y., Pálfia, M.: Matrix power mean and the Karcher mean. J. Funct. Anal. **262**, 1498–1514 (2012)
17. Pusz, W., Woronowicz, S.L.: Functional calculus for sesquilinear forms and the purification map. Rep. Math. Phys. **8**, 159–170 (1975)
18. Sturm, K.-T.: Probability measures on metric spaces of nonpositive curvature, Heat kernels and analysis on manifolds, graphs, and metric spaces (Paris, 2002), pp. 357–390, Contemp. Math., vol. 338. Amer. Math. Soc., Providence (2003)