TAIWANESE JOURNAL OF MATHEMATICS
Vol. 22, No. 1, pp. 79–94, February 2018
DOI: 10.11650/tjm/8117

The Order Properties and Karcher Barycenters of Probability Measures on the Open Convex Cone

Sejong Kim

Abstract. We study the probability measures on the open convex cone of positive definite operators equipped with the Loewner ordering. We show that two crucial push-forward measures derived by the congruence transformation and power map preserve the stochastic order for probability measures. By the continuity of two push-forward measures with respect to the Wasserstein distance, we verify several interesting properties of the Karcher barycenter for probability measures with finite first moment such as the invariant properties and the inequality for unitarily invariant norms. Moreover, the characterization for the stochastic order of uniformly distributed probability measures has been shown.

1. Introduction

K.-T. Sturm [22] has developed a theory of barycenters of probability measures on the Hadamard spaces, which are complete metric spaces satisfying the semi-parallelogram law. One of the important and typical examples of a Hadamard space is the open convex cone Ω of positive definite (Hermitian) matrices equipped with the Riemannian trace metric $d(A,B) = \left\|\log A^{-1/2}BA^{-1/2}\right\|_2$ (see [3, Chapter 6]). For a probability measure μ on Ω with finite first moment, the *Cartan barycenter* Γ is given by

$$\Gamma(\mu) = \underset{X\in\Omega}{\arg\min} \int_\Omega \left[d^2(X,A) - d^2(Y,A)\right] \, d\mu(A).$$

This barycenter is independent of Y and coincides with $\Lambda(\mu) = \arg\min_{X\in\Omega} \int_\Omega \delta^2(X,A)$ $d\mu(A)$ if μ has a finite second moment. Especially, for a uniformly distributed probability measure $\mu = \frac{1}{n}\sum_{j=1}^n \delta_{A_j}$, where δ_A is a point measure at $A \in \Omega$, the Cartan barycenter Γ is the unique minimzer of the sum of Riemannian distances to each point A_j. Moreover, H. Karcher [11] has proved that it is the unique solution X in Ω of the Karcher equation

$$\sum_{j=1}^n \log(X^{-1/2}A_jX^{-1/2}) = 0.$$

Received December 25, 2016; Accepted May 21, 2017.
Communicated by Xiang Fang.
2010 *Mathematics Subject Classification*. Primary: 47B65; Secondary: 15B48.
Key words and phrases. Loewner order, stochastic order, Wasserstein distance, max-flow and min-cut theorem, unitarily invariant norm, Karcher barycenter.

Since then we also call it the *Karcher mean* or *least squares mean*. Recently, many research topics about the Karcher mean such as finding properties and computing efficiently it have been arisen (see [9, 18, 19]).

On the setting of positive definite (bounded) operators on a Hilbert space, however, one has neither such Riemannian metric nor non-positive curvature metric. So a natural question to generalize the properties of the matrix Cartan barycenter to the Hilbert space setting is arisen. Lawson and Lim [16] have shown that operator power means are monotonically decreasing and have the limit as $t \to 0^+$, which satisfies the Karcher equation. It allows us to establish the existence of Karcher mean of positive definite operators on a Hilbert space. Recently, M. Pálfia [21] has generalized the operator Karcher mean on the setting of bounded probability measures. By the fundamental relationship between contractive intrinsic (symmetric and multiplicative) means and barycentric maps in a complete metric space appeared in [17], one can extend the Karcher mean to the barycenter on the probability measure space of positive definite operators equipped with the Thompson metric. We call in this article such a barycenter *Karcher barycenter*. By showing the continuity of push-forward measures constructed by congruence transformations and power maps with respect to the Wasserstein distance, we verify the invariant properties of the Karcher barycenters on the probability measure space of positive definite operators. Furthermore, we show a remarkable inequality of Karcher barycenters of positive definite matrices for log-majorization and unitarily invariant norms.

Furthermore, we investigate the stochastic order relation of probability measures on the open convex cone of positive definite operators. We show that two push-forward measures constructed by congruence transformations and power maps preserve the stochastic order, which generalize the Loewner-Heinz inequality. The relationship between the stochastic order of uniformly distributed probability measures and the Loewner order of supporting elements has been established by using the max-flow and min-cut theorem in the graph theory.

2. Probability measures on a metric space

Let (X, d) be a metric space with the algebra $\mathcal{B}(X)$ of Borel sets, the smallest σ-algebra containing the open sets. Let $\mathcal{P}(X)$ be the set of all probability measures on $(X, \mathcal{B}(X))$ with separable support. Let $\mathcal{P}_0(X)$ be the set of all discrete measures $\mu \in \mathcal{P}(X)$, of which form is the convex combination of point measures δ_x of mass 1 at $x \in X$. For $p \in [1, \infty)$ let $\mathcal{P}^p(X) \subseteq \mathcal{P}(X)$ be the set of all probability measures with finite p-th moment: for some (and hence, for all) $y \in X$

$$\int_X d^p(x, y)\, d\mu(x) < \infty.$$

For $p = \infty$, $\mathcal{P}^\infty(X)$ denotes the set of all probability measures with bounded separable support.

We introduce a pushforward measure that allows to get new measure from the given one. For metric spaces X and Y, a continuous function $f\colon X \to Y$ induces a *push-forward* map $f_*\colon \mathcal{P}(X) \to \mathcal{P}(Y)$ defined by

$$f_*(\mu)(B) = \mu(f^{-1}(B))$$

for any $\mu \in \mathcal{P}(X)$ and $B \in \mathcal{B}(Y)$. Note that $\mathrm{supp}(f_*(\mu)) = \overline{f(\mathrm{supp}(\mu))}$, the closure of the image of the support of μ under f. The following is called the *change of variables formula.*

Theorem 2.1. *Let $\mu \in \mathcal{P}(X)$ and let $f\colon X \to Y$ be a continuous function between metric spaces. A measurable function g on Y is integrable with respect to the push-forward measure $f_*(\mu)$ if and only if the composition $g \circ f$ is integrable with respect to μ. In this case,*

$$\int_Y g\, d(f_*(\mu)) = \int_X g \circ f\, d\mu.$$

We say that $\omega \in \mathcal{P}(X \times X)$ is a *coupling* of $\mu, \nu \in \mathcal{P}(X)$, or $\mu, \nu \in \mathcal{P}(X)$ are the marginals of $\omega \in \mathcal{P}(X \times X)$, if

$$\omega(B \times X) = \mu(B) \quad \text{and} \quad \omega(X \times B) = \nu(B)$$

for all $B \in \mathcal{B}(X)$. Equivalently, μ and ν are the push-forward measures of any coupling ω under the projection maps π_1 and π_2, respectively. Note that $\mathrm{supp}(\omega) \subseteq \mathrm{supp}(\mu) \times \mathrm{supp}(\nu)$ for any coupling ω. We denote the set of all couplings for μ and ν by $\Pi(\mu, \nu)$.

For $p \in [1, \infty)$ the *p-Wasserstein distance* d_p^W, alternatively *Kantorovich-Rubinstein distance*, on $\mathcal{P}^p(X)$ is defined by

$$d_p^W(\mu, \nu) := \left(\inf_{\omega \in \Pi(\mu,\nu)} \int_{X \times X} d^p(x, y)\, d\omega(x, y) \right)^{1/p}.$$

It is known that d_p^W is a complete metric on $\mathcal{P}^p(X)$ whenever X is a complete metric space, and $\mathcal{P}_0(X)$ is d_p^W-dense in $\mathcal{P}^p(X)$ [5, 22]. The limit $\lim_{p \to \infty} d_p^W(\mu, \nu)$ is finite on $\mathcal{P}^\infty(X)$ and yields a complete metric space. So the ∞-Wasserstein distance is given by

$$d_\infty^W(\mu, \nu) = \inf_{\omega \in \Pi(\mu,\nu)} \sup\{d(x, y) : (x, y) \in \mathrm{supp}(\omega)\}.$$

Lemma 2.2. [17, Lemma 2.2] *Let $f\colon X \to Y$ be a Lipschitz map with Lipschitz constant C. Then $f_*\colon \mathcal{P}^p(X) \to \mathcal{P}^p(Y)$ is Lipschitz with Lipschitz constant C for $1 \le p \le \infty$.*

3. The open convex cone of positive definite operators

We switch our point of view to the open convex cone of positive definite operators. For a Hilbert space $\mathcal{H}$ with inner product $\langle \cdot, \cdot \rangle$, let $B(\mathcal{H})$ be the Banach space of all bounded linear operators on $\mathcal{H}$ equipped with the operator norm. Let $S(\mathcal{H}) \subseteq B(\mathcal{H})$ be the closed subspace of all self-adjoint operators, and let $\mathbb{P} \subseteq S(\mathcal{H})$ be the open convex cone of all positive definite operators. Note that $A \in \mathbb{P}$ means that $\langle x, Ax \rangle > 0$ for all nonzero $x \in \mathcal{H}$. For $X, Y \in S(\mathcal{H})$ we denote $X \leq Y$ if and only if $Y - X$ is positive semidefinite, and $X < Y$ if and only if $Y - X$ is positive definite.

We equip the open convex cone $\mathbb{P}$ with the Thompson metric

$$d(A, B) = \left\| \log(A^{-1/2} B A^{-1/2}) \right\|,$$

where $\| \cdot \|$ denotes the operator norm. It is known that d is a complete metric on $\mathbb{P}$ and that

$$d(A, B) = \max\{\log M(B/A), \log M(A/B)\},$$

where $M(B/A) = \inf\{\alpha > 0 : B \leq \alpha A\}$, the largest eigenvalue of $A^{-1/2} B A^{-1/2}$, see [6, 20, 23].

Lemma 3.1. [6, 15] *Basic properties of the Thompson metric on $\mathbb{P}$ include*

(i) $d(A, B) = d(A^{-1}, B^{-1}) = d(MAM^*, MBM^*)$ *for any invertible M, and*

(ii) $d(A^t, B^t) \leq t d(A, B)$, $t \in [0, 1]$.

For $p \in [1, \infty)$, $\mathcal{P}^p(\mathbb{P})$ is the set of all probability measures with finite p-th moment with respect to the Thompson metric d. We give two important push-forward measures obtained by the congruence transformation and power map. Let $\mathcal{O} \in \mathcal{B}(\mathbb{P})$ and $\mu \in \mathcal{P}(\mathbb{P})$. For $M \in \mathrm{GL}$, the general linear group, and $t \in \mathbb{R} \setminus \{0\}$,

$$(M.\mu)(\mathcal{O}) = \mu(M^{-1} \mathcal{O} (M^{-1})^*), \quad \mu^t(\mathcal{O}) = \mu(\mathcal{O}^{1/t}),$$

where

$$M \mathcal{O} M^* = \{MAM^* : A \in \mathcal{O}\}, \quad \mathcal{O}^t = \{A^t : A \in \mathcal{O}\}.$$

Note that $M.\mu, \mu^t \in \mathcal{P}^1(\mathbb{P})$ whenever $\mu \in \mathcal{P}^1(\mathbb{P})$. One can see that $M.\mu$ and μ^t are push-forward measures such that

$$M.\mu = f_*(\mu) \quad \text{and} \quad \mu^t = g_*(\mu),$$

where $f(X) = MXM^*$ and $g(X) = X^t$ are continuous functions on $\mathbb{P}$.

A set $\mathcal{U} \subseteq \mathbb{P}$ is called an upper set if whenever $A \in \mathcal{U}$ and $A \leq B$, then $B \in \mathcal{U}$ (see [12]). Note that $\mathcal{U} = \bigcup_{i=1}^n \mathcal{U}_i \subseteq \mathbb{P}$ is the upper set whenever $\mathcal{U}_i$'s are all upper sets in

$\mathbb{P}$. Indeed, if $A \in \mathcal{U}$ and $A \le B$, then $A \in \mathcal{U}_j$ for some $1 \le j \le n$. Since $\mathcal{U}_j$ is an upper set, $B \in \mathcal{U}_j \subset \mathcal{U}$. We define the order of probability measures such as $\mu \le \nu$ for probability measures μ and ν on $\mathbb{P}$ if $\mu(\mathcal{U}) \le \nu(\mathcal{U})$ for any upper set $\mathcal{U} \subseteq \mathbb{P}$. We see some interesting properties of the order relation for probability measures in the following.

Proposition 3.2. *Let $\mu, \nu \in \mathcal{P}^1(\mathbb{P})$ with $\mu \le \nu$. Then $M.\mu \le M.\nu$ and $\mu^t \le \nu^t$ for any $M \in \mathrm{GL}$ and $t \in (0,1]$.*

Proof. Let $\mathcal{U} \in \mathcal{B}(\mathbb{P})$ be any upper set of $\mathbb{P}$.

We first claim that MUM^* is an upper set of $\mathbb{P}$. Indeed, let $A \in M\mathcal{U}M^*$ and $A \le B$. Then $M^{-1}A(M^{-1})^* \in \mathcal{U}$, and $M^{-1}A(M^{-1})^* \le M^{-1}B(M^{-1})^*$ by Theorem 7.7.2 in [10]. Since U is the upper set, $M^{-1}B(M^{-1})^* \in \mathcal{U}$, that is, $B \in M\mathcal{U}M^*$. Thus, $(M.\mu)(\mathcal{U}) = \mu(M^{-1}\mathcal{U}(M^{-1})^*) \le \nu(M^{-1}\mathcal{U}(M^{-1})^*) = (M.\nu)(\mathcal{U})$, since $\mu \le \nu$.

Moreover, $\mathcal{U}^{1/t}$ is also an upper set of $\mathbb{P}$ for $t \in (0,1]$. Indeed, let $A \in \mathcal{U}^{1/t}$ and $A \le B$. Then $A^t \in \mathcal{U}$, and $A^t \le B^t$ by the Loewner-Heinz inequality in [3, Theorem 1.5.9]. Since $\mathcal{U}$ is the upper set, $B^t \in \mathcal{U}$, that is, $B \in \mathcal{U}^{1/t}$. Thus, $\mu^t(\mathcal{U}) = \mu(\mathcal{U}^{1/t}) \le \nu(\mathcal{U}^{1/t}) = \nu^t(\mathcal{U})$, since $\mu \le \nu$. □

Remark 3.3. The following is known as the *Loewner-Heinz inequality*: for $A, B \in \mathbb{P}$

$$A \le B \text{ implies } A^t \le B^t \text{ for any } t \in (0,1].$$

See [7] and the references therein. We assert that $\mu^t \le \nu^t$ whenever $\mu \le \nu$ for $\mu, \nu \in \mathcal{P}^1$ and $t \in (0,1]$. Let $\mu = \delta_A$ and $\nu = \delta_B$ satisfying $\mu \le \nu$, where δ_X is a point measure of mass 1 at $X \in \mathbb{P}$. Then $\mu \le \nu$ means that $A \le B$, and hence,

$$A^t = \mu^t(\mathbb{P}) \le \nu^t(\mathbb{P}) = B^t.$$

We now see the properties of probability measures obtained by a convex combination of given measures. From the definition of the order for probability measures, the following is obvious.

Proposition 3.4. *Let $\mu_1 = (1-t)\mu + t\nu_1$ and $\mu_2 = (1-t)\mu + t\nu_2$ for $t \in (0,1]$, where $\mu, \nu_1, \nu_2 \in \mathcal{P}(\mathbb{P})$. Then $\mu_1 \le \mu_2$ if and only if $\nu_1 \le \nu_2$.*

Corollary 3.5. *Let $\mu = (1-t)\delta_A + t\delta_B$ and $\nu = (1-t)\delta_A + t\delta_C$ for $t \in (0,1]$, where $A, B, C \in \mathbb{P}$. Then $\mu \le \nu$ if and only if $B \le C$.*

We show the convexity of the measure μ^t for $t \in [-1,1] \setminus \{0\}$ with respect to the Wasserstein distance.

Lemma 3.6. *For any $\mu, \nu \in \mathcal{P}^p(\mathbb{P})$, $1 \le p \le \infty$, and $t \in [-1,1] \setminus \{0\}$,*

$$d_p^W(\mu^t, \nu^t) \le |t| d_p^W(\mu, \nu). \tag{3.1}$$

Proof. Let $t \in [-1,1] \setminus \{0\}$. The power map $f\colon \mathbb{P} \to \mathbb{P}$ given by $f(A) = A^t$ for any $A \in \mathbb{P}$ gives us a push-forward measure μ^t for $\mu \in \mathcal{P}^p(\mathbb{P})$. Moreover, it is Lipschitz for the Thompson metric with Lipschitz constant $|t|$ by Lemma 3.1. By Lemma 2.2 $f_*\colon \mathcal{P}^p(\mathbb{P}) \to \mathcal{P}^p(\mathbb{P})$ is Lipschitz with Lipschitz constant $|t|$, that is, the property (3.1) holds. □

Remark 3.7. We have seen in Lemma 3.6 that the map $\mu \mapsto \mu^t$ on $\mathcal{P}^p(\mathbb{P})$ is continuous for $t \in [-1,1] \setminus \{0\}$ with respect to the p-Wasserstein metric. The continuity of the map $\mu \mapsto \mu^t$ on $\mathcal{P}^p(\mathbb{P})$ for $|t| > 1$ is an open problem.

Lemma 3.8. *For any $\mu, \nu \in \mathcal{P}^p(\mathbb{P})$, $1 \leq p \leq \infty$, and $M \in \mathrm{GL}$*

$$d_p^W(M.\mu, M.\nu) = d_p^W(\mu, \nu).$$

Proof. Let $\mu, \nu \in \mathcal{P}^p(\mathbb{P})$. By the change of variables in Theorem 2.1 and Lemma 3.1(1),

$$\begin{aligned} d_p^W(M.\mu, M.\nu) &= \left[\inf_{\omega \in \Pi(M.\mu, M.\nu)} \int_{\mathbb{P}\times\mathbb{P}} d^p(A,B)\, d\omega(A,B)\right]^{1/p} \\ &= \left[\inf_{M^{-1}.\omega \in \Pi(\mu,\nu)} \int_{\mathbb{P}\times\mathbb{P}} d^p(A,B)\, d\omega(A,B)\right]^{1/p} \\ &= \left[\inf_{\phi \in \Pi(\mu,\nu)} \int_{\mathbb{P}\times\mathbb{P}} d^p(A,B)\, d(M.\phi)(A,B)\right]^{1/p} \\ &= \left[\inf_{\phi \in \Pi(\mu,\nu)} \int_{\mathbb{P}\times\mathbb{P}} d^p(M.A, M.B)\, d\phi(A,B)\right]^{1/p} \\ &= \left[\inf_{\phi \in \Pi(\mu,\nu)} \int_{\mathbb{P}\times\mathbb{P}} d^p(A,B)\, d\phi(A,B)\right]^{1/p} = d_p^W(\mu,\nu). \end{aligned}$$ □

4. Order relations of uniformly distributed probability measures

We here see the characterization of order relation for uniformly distributed probability measures $\mu = \frac{1}{n}\sum_{i=1}^n \delta_{A_i}$ and $\nu = \frac{1}{n}\sum_{i=1}^n \delta_{B_i}$ satisfying $\mu \leq \nu$. The following shows the relationship for $n = 2$.

Proposition 4.1. *Let $\mu = \frac{1}{2}(\delta_{A_1} + \delta_{A_2})$ and $\nu = \frac{1}{2}(\delta_{B_1} + \delta_{B_2})$, where $A_1, A_2, B_1, B_2 \in \mathbb{P}$. Then $\mu \leq \nu$ if and only if at least one of the following holds:*

(i) $A_1 \leq B_1$ *and* $A_2 \leq B_2$,

(ii) $A_1 \leq B_2$ *and* $A_2 \leq B_1$.

Proof. Assume that $\mu \leq \nu$ for $\mu = \frac{1}{2}(\delta_{A_1}+\delta_{A_2})$ and $\nu = \frac{1}{2}(\delta_{B_1}+\delta_{B_2})$, where $A_1, A_2, B_1, B_2 \in \mathbb{P}$. Note that

$$\frac{1}{2} \leq \mu([A_i, \infty)) \leq \nu([A_i, \infty))$$

for all $i = 1, 2$, where $[A, \infty) := \{X \in \mathbb{P} : A \leq X\}$ is the upper set in $\mathbb{P}$. So we have the following four cases:

(1) If $\nu([A_1, \infty)) = \nu([A_2, \infty)) = 1$, then $B_1, B_2 \in [A_i, \infty)$ for all $i = 1, 2$. That is, $A_i \leq B_1, B_2$ for all $i = 1, 2$, and hence, both (i) and (ii) hold.

(2) If $\nu([A_1, \infty)) = 1$ and $\nu([A_2, \infty)) = 1/2$, then $A_1 \leq B_j$ for all $j = 1, 2$, and either B_1 or B_2 belongs to the upper set $[A_2, \infty)$. That is, either $A_2 \leq B_1$ or $A_2 \leq B_2$, and hence, one of (i) and (ii) holds.

(3) If $\nu([A_1, \infty)) = 1/2$ and $\nu([A_2, \infty)) = 1$, then one of (i) and (ii) holds by the similar proof of the case (2).

(4) If $\nu([A_1, \infty)) = 1/2$ and $\nu([A_2, \infty)) = 1/2$, then either B_1 or B_2 belongs to the upper set $[A_i, \infty)$ for each $i = 1, 2$. So we have the following cases:

   (a) If $B_1 \in [A_i, \infty)$ and $B_2 \notin [A_i, \infty)$ for all $i = 1, 2$, then

$$\mu([A_1, \infty) \cup [A_2, \infty)) = 1 > \frac{1}{2} = \nu([A_1, \infty) \cup [A_2, \infty))$$

   for the upper set $[A_1, \infty) \cup [A_2, \infty)$, and it is a contradiction to $\mu \leq \nu$.

   (b) If $B_1 \notin [A_i, \infty)$ and $B_2 \in [A_i, \infty)$ for all $i = 1, 2$, it contradicts by the similar proof of (a).

   (c) If $B_1 \in [A_1, \infty)$, $B_1 \notin [A_2, \infty)$ and $B_2 \notin [A_1, \infty)$, $B_2 \in [A_2, \infty)$, then (i) holds.

   (d) If $B_1 \notin [A_1, \infty)$, $B_1 \in [A_2, \infty)$ and $B_2 \in [A_1, \infty)$, $B_2 \notin [A_2, \infty)$, then (ii) holds.

It is obvious that if at least one of the items (i) and (ii) holds, then $\mu \leq \nu$. □

By the similar proof of Proposition 4.1, we may prove the order relation of uniformly distributed probability measures μ and ν supporting on $\{A_1, \dots, A_n\}$ and $\{B_1, \dots, B_n\}$ for $n > 2$, respectively. However, the cases are very complicated. In order to generalize Proposition 4.1 we recall the *Max-Flow and Min-Cut Theorem* in the graph theory. Consider a directed graph with two distinguished vertices, a source s with only outgoing arrows and a sink t with only incoming arrows. We associate with all directed edges a number in $[0, \infty)$, called the *capacity* of the edge. For example, we may think of s as a water provider, t as a water consumer, the edges as pipes that can carry up to their individual capacities of water, and the direction of the edge as the direction of water flow.

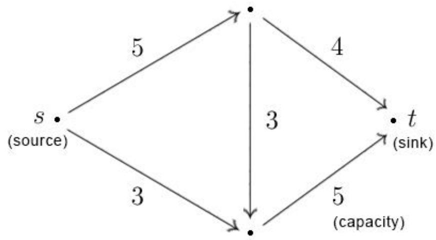

Figure 4.1: Directed graph.

Let V be the set of all vertices on a directed graph. A *cut* is a partition of the vertices into subsets P and Q such that $s \in P$ and $t \in Q$. We add up the capacities of all directed edges that start in P and end in Q, and we denote the sum as $\text{cut}(P, Q)$. A *minimal cut* is the cut having a minimum value of $\text{cut}(P, Q)$ for all partitions (P, Q), which we call the value of the minimal cut. On the directed graph in Figure 4.1 the value of the minimal cut is 8 when $(P, Q) = (\{s\}, V \setminus \{s\})$.

A *flow* is an assignment to each edge of a value from $[0, \infty)$ that

(i) does not exceed the capacity of that edge, and

(ii) has the property that the sum of the values flowing into any vertex, excluding the source s and sink t, is equal to the sum of the values flowing out.

A *maximal flow* is one having a maximum sum of the values flowing out from the source, which we call the value of the maximal flow. Note that the value of the maximal flow is equal to the maximum sum of the values flowing into the sink.

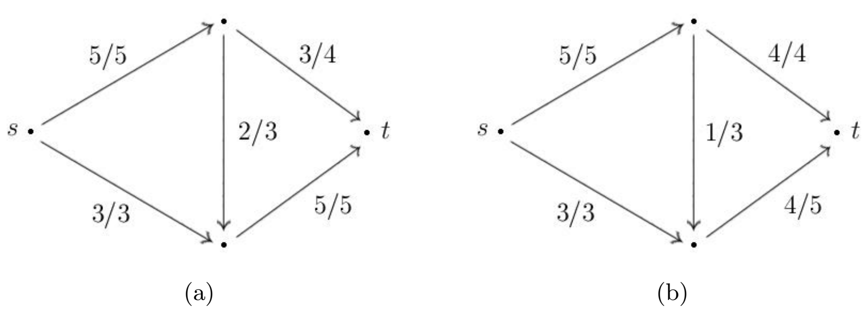

Figure 4.2: Maximal flows.

The following is called the *Max-Flow and Min-Cut Theorem*, known as a special case of the duality theorem for linear programs and used to derive Menger's theorem and the König-Egerváry theorem.

Theorem 4.2. [13, Section 4.5] *On a directed graph the value of the maximal flow is equal to the value of the minimal cut.*

Figure 4.2 shows that the value of the maximal flow is 8, which is the same as the value of the minimal cut, and a maximal flow is not unique.

Assume that $\mu = \frac{1}{n}\sum_{i=1}^{n}\delta_{A_i}, \nu = \frac{1}{n}\sum_{i=1}^{n}\delta_{B_i} \in \mathcal{P}_0(\mathbb{P})$ with $\mu \leq \nu$. Construct a directed graph with vertices labeled $A_1, \ldots, A_n$ from the source s and vertices labeled $B_1, \ldots, B_n$ into the sink t of capacity 1 on each arrow.

Remark 4.3. Note that for each i, $A_i \leq B_j$ for some $j \in \{1, \ldots, n\}$ since $\nu([A_i, \infty)) \geq \mu([A_i, \infty)) \geq 1/n$.

We put a directed edge $A_i \to B_j$ with capacity n on the graph if $A_i \leq B_j$. We denote the directed graph constructed in this way as G. See Figure 4.3.

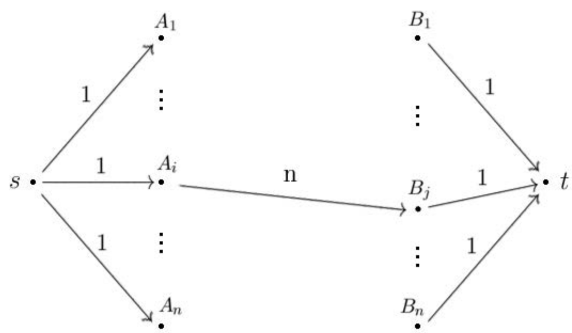

Figure 4.3: Measure order and graph G.

Lemma 4.4. *On the directed graph G in Figure* 4.3, $\mathrm{cut}(P,Q) \geq n$ *for any partition (P,Q) satisfying that $(P,Q) \neq (\{s\}, V \setminus \{s\})$ and $(P,Q) \neq (V \setminus \{t\}, \{t\})$, where $V = \{s, A_1, \ldots, A_n, B_1, \ldots, B_n, t\}$.*

Proof. We have four cases of the partition (P,Q) satisfying that $(P,Q) \neq (\{s\}, V \setminus \{s\})$ and $(P,Q) \neq (V \setminus \{t\}, \{t\})$.

(1) If a subset P contains a nontrivial subset of $\{A_1, \ldots, A_n\}$ and no elements of $\{B_1, \ldots, B_n\}$, then clearly $\mathrm{cut}(P,Q) \geq n$ since there is an arrow $A_i \to B_j$ with capacity n for some $j \in \{1, \ldots, n\}$ by Remark 4.3.

(2) If a subset P contains a nontrivial proper subset X of $\{A_1, \ldots, A_n\}$ and a nontrivial proper subset Y of $\{B_1, \ldots, B_n\}$, then we consider two cases. Let $|X| = p$ and $|Y| = q$, where $|X|$ denotes the cardinality of the set X.

(a) If there is an arrow from an element of X to an element of $\{B_1, \ldots, B_n\} \setminus Y$, then obviously $\mathrm{cut}(P,Q) \geq n$.

(b) If there are no arrows from any element of X to any element of $\{B_1, \ldots, B_n\} \setminus Y$, then $\mathrm{cut}(P, Q) = n - p + q$ since we add up all capacities of arrows from s to $\{A_1, \ldots, A_n\} \setminus X$ and from $\{B_1, \ldots, B_n\} \setminus Y$ to t. Suppose $p - q > 0$, that is, $p > q$. Let $X = \{A_{i_1}, \ldots, A_{i_p}\}$. Then

$$\mu\left(\bigcup_{l=1}^{p}[A_{i_l}, \infty)\right) \geq \frac{p}{n} > \frac{q}{n} = \nu\left(\bigcup_{l=1}^{p}[A_{i_l}, \infty)\right).$$

The equality follows from the assumption. It is a contradiction to $\mu \leq \nu$. So $p \leq q$, and hence, $\mathrm{cut}(P, Q) = n - p + q \geq n$.

(3) If a subset P contains $\{A_1, \ldots, A_n\}$ and a nontrivial proper subset of $\{B_1, \ldots, B_n\}$, then there is an element $B_k \notin P$ for some $k \in \{1, \ldots, n\}$. Since

$$\nu\left(\bigcup_{i=1}^{n}[A_i, \infty)\right) \geq \mu\left(\bigcup_{i=1}^{n}[A_i, \infty)\right) = 1,$$

$\nu\left(\bigcup_{i=1}^{n}[A_i, \infty)\right) = 1$. It means that there must exist an arrow with capacity n pointing into B_k from some of $\{A_1, \ldots, A_n\}$. So $\mathrm{cut}(P, Q) \geq n$.

(4) If a subset P contains a proper subset of $\{A_1, \ldots, A_n\}$ and all elements of $\{B_1, \ldots, B_n\}$, then each arrow $B_j \to t$ has the capacity 1, and so $\mathrm{cut}(P, Q) \geq n$.

□

Remark 4.5. By Lemma 4.4, $\mathrm{cut}(P, Q) = n$ is the value of the minimal cut when the partition $(P, Q) = (\{s\}, V \setminus \{s\})$ or $(P, Q) = (V \setminus \{t\}, \{t\})$. By Theorem 4.2 the value of the maximal flow is n, and thus, the value of each flow $s \to A_i$ and $B_i \to t$ for all i must be 1.

Theorem 4.6. *Let $\mu = \frac{1}{n}\sum_{i=1}^{n}\delta_{A_i}$ and $\nu = \frac{1}{n}\sum_{i=1}^{n}\delta_{B_i}$ be discrete measures, where $A_i, B_i \in \mathbb{P}$ for all $i = 1, 2, \ldots, n$. Then $\mu \leq \nu$ if and only if there exists a permutation σ on $\{1, 2, \ldots, n\}$ such that*

$$A_i \leq B_{\sigma(i)}, \quad i = 1, \ldots, n.$$

Proof. Assume that there is a permutation σ on $\{1, \ldots, n\}$ such that $A_i \leq B_{\sigma(i)}$ for all i. Then we can easily see that $\mu(\mathcal{U}) \leq \nu(\mathcal{U})$ for any upper set $\mathcal{U} \subseteq \mathbb{P}$, and hence, $\mu \leq \nu$.

Conversely, we use the directed graph G constructed as above. Since the flow value from the source s into A_i is 1, there do not exist two different arrows with flow value 1 from A_i into two different vertices, so each A_i has exactly one arrow leaving it. Suppose that there are arrows with flow value 1 from two different vertices A_i, A_k into the same vertex B_j for some $j \in \{1, \ldots, n\}$. Then the value of the flow from B_j into the sink t

must be at least 2. Since $B_j \to t$ has capacity 1, it contradicts the definition of a flow. So there are different arrows from two distinguished vertices A_i and A_k into two distinguished vertices B_j and B_l. Since

$$\nu\left(\bigcup_{i=1}^{n}[A_i,\infty)\right) = 1,$$

there must exist an arrow pointing into each B_j from $\{A_1,\ldots,A_n\}$. Therefore, there is a permutation σ on $\{1,\ldots,n\}$ such that $A_i \to B_{\sigma(i)}$, i.e., $A_i \le B_{\sigma(i)}$ for all i. □

5. Properties of Karcher barycenter

The Karcher mean $\Lambda = \{\Lambda_n\}_{n\ge 2}$ on the open convex cone $\mathbb{P}$ of positive definite operators is defined as the unique solution in $\mathbb{P}$ of the Karcher equation

$$X = \Lambda_n(A_1,\ldots,A_n) \quad \Longleftrightarrow \quad \sum_{j=1}^{n}\log(X^{-1/2}A_jX^{-1/2}) = 0.$$

It has been shown in [16] that the Karcher equation has a unique solution in $\mathbb{P}$ and the Karcher mean Λ_n for $n \ge 2$ satisfies

(i) Λ_n is invariant under permutation, that is, for any permutation σ on $\{1,\ldots,n\}$

$$\Lambda_n(A_{\sigma(1)},\ldots,A_{\sigma(1)}) = \Lambda_n(A_1,\ldots,A_n),$$

(ii) Λ_n is multiplicative, that is,

$$\Lambda_{nk}(A_1,\ldots,A_n,\ldots,A_1,\ldots,A_n) = \Lambda_n(A_1,\ldots,A_n),$$

where the number of blocks is k (≥ 2),

(iii) Λ_n is invariant under congruence transformation and inversion, that is, for any $M \in \mathrm{GL}$

$$\Lambda_n(MA_1M^*,\ldots,MA_nM^*) = M\Lambda_n(A_1,\ldots,A_n)M^*,$$
$$\Lambda_n(A_1^{-1},\ldots,A_n^{-1}) = \Lambda_n(A_1,\ldots,A_n)^{-1},$$

(iv) Λ_n is monotonic, that is,

$$\Lambda_n(A_1,\ldots,A_n) \le \Lambda_n(B_1,\ldots,B_n)$$

whenever $A_i \le B_i$ for all $i = 1,\ldots,n$,

(v) Λ_n is contractive, that is,

$$d(\Lambda_n(A_1,\dots,A_n),\Lambda_n(B_1,\dots,B_n)) \le \frac{1}{n}\sum_{j=1}^{n} d(A_j,B_j),$$

where d is the Thompson metric.

Since the Karcher mean Λ is an intrinsic (invariant under permutation and multiplicative) mean in a complete metric space $\mathbb{P}$ with the Thompson metric d, we have the following by Proposition 3.7 in [17].

Proposition 5.1. *A barycentric map $\beta_\Lambda\colon \mathcal{P}^1(\mathbb{P}) \to \mathbb{P}$ determined by*

$$\beta_\Lambda\left(\frac{1}{n}\sum_{j=1}^{n}\delta_{A_j}\right) = \Lambda_n(A_1,\dots,A_n)$$

uniquely exists and satisfies the contraction property: for any $\mu,\nu \in \mathcal{P}^1(\mathbb{P})$

$$d(\beta_\Lambda(\mu),\beta_\Lambda(\nu)) \le d_1^W(\mu,\nu). \tag{5.1}$$

Definition 5.2. The barycentric map β_Λ in Proposition 5.1 determined by the Karcher mean Λ is called the *Karcher barycenter* on $\mathcal{P}^1(\mathbb{P})$.

The following shows the invariant properties of the Karcher barycenter, which generalize those of the Karcher mean Λ.

Theorem 5.3. *The Karcher barycenter $\beta_\Lambda\colon \mathcal{P}^1(\mathbb{P}) \to \mathbb{P}$ satisfies that for any $\mu \in \mathcal{P}^1(\mathbb{P})$, $M \in \mathrm{GL}$*

$$\beta_\Lambda(M.\mu) = M.\beta_\Lambda(\mu), \quad \beta_\Lambda(\mu^{-1}) = \beta_\Lambda(\mu)^{-1}.$$

Proof. Let $\mu \in \mathcal{P}^1(\mathbb{P})$. Since $\mathcal{P}_0(\mathbb{P})$ is d_1^W-dense in $\mathcal{P}^1(\mathbb{P})$ [5, 22], there exists a sequence $\{\mu_n\} \subset \mathcal{P}_0(\mathbb{P})$ of the form $\mu_n = \frac{1}{n}\sum_{j=1}^{n}\delta_{A_j}$ such that $d_1^W(\mu_n,\mu) < \epsilon$ for any $\epsilon > 0$. Then $M.\mu_n, \mu_n^{-1} \in \mathcal{P}_0(\mathbb{P})$, and by the contraction property (5.1), Lemmas 3.8 and 3.6

$$\begin{aligned} d(\beta_\Lambda(M.\mu_n),\beta_\Lambda(M.\mu)) &\le d_1^W(M.\mu_n, M.\mu) = d_1^W(\mu_n,\mu) < \epsilon,\\ d(\beta_\Lambda(\mu_n^{-1}),\beta_\Lambda(\mu^{-1})) &\le d_1^W(\mu_n^{-1},\mu^{-1}) \le d_1^W(\mu_n,\mu) < \epsilon. \end{aligned}$$

Since the Karcher barycenter β_Λ satisfies $\beta_\Lambda(\mu_n) = \Lambda_n(A_1,\dots,A_n)$ by Proposition 5.1 and (iii) the Karcher mean Λ is invariant under congruence transformation and inversion by Theorem 6.8 in [16], we obtain the following:

$$\begin{aligned} \beta_\Lambda(M.\mu) &= \lim_{n\to\infty}\beta_\Lambda(M.\mu_n) = \lim_{n\to\infty} M.\beta_\Lambda(\mu_n) = M.\beta_\Lambda(\mu),\\ \beta_\Lambda(\mu^{-1}) &= \lim_{n\to\infty}\beta_\Lambda(\mu_n^{-1}) = \left[\lim_{n\to\infty}\beta_\Lambda(\mu_n)\right]^{-1} = \beta_\Lambda(\mu)^{-1}. \end{aligned}$$

□

Remark 5.4. The monotonicity of Karcher barycenter corresponding to the intrinsic means on a metric space with a closed partial order has been shown in [14]. Thus, the Karcher barycenter $\beta_\Lambda\colon \mathcal{P}^1(\mathbb{P}) \to \mathbb{P}$ determined by the Karcher mean Λ is monotonic, i.e.,

$$\mu \le \nu \quad \Longrightarrow \quad \beta_\Lambda(\mu) \le \beta_\Lambda(\nu).$$

Since the Karcher mean Λ satisfies many properties such as the joint concavity and the arithmetic-Karcher-harmonic mean inequalities, many interesting problems for the Karcher barycenter β_Λ on the setting of probability measures remain still open.

As application of the Karcher barycenter of positive definite matrices, we see the remarkable property with the log-majorization and unitarily invariant norms. Let A be an $m \times m$ positive semidefinite (Hermitian) matrix with eigenvalues $\lambda_j(A)$ for $j = 1, \dots, m$ arranged in decreasing order, i.e., $\lambda_1(A) \ge \cdots \ge \lambda_m(A)$. Let $A, B \ge 0$. We say that B weakly log-majorizes A, written as $A \prec_{(w \log)} B$, if and only if

$$\prod_{i=1}^k \lambda_i(A) \le \prod_{i=1}^k \lambda_i(B) \quad \text{for } k = 1, 2, \dots, m.$$

If the equality holds for $k = m$, in addition, then we say that B log-majorizes A, written as $A \prec_{(\log)} B$.

A norm $|||\cdot|||$ on M_m, the set of all $m \times m$ matrices with complex entries, is said to be *unitarily invariant* if and only if $|||UAV||| = |||A|||$ for any $A \in M_m$ and unitary matrices $U, V \in M_m$. A typical example of unitarily invariant norms is the Ky Fan k-norm: see [10, Chapter 5, Chapter 7] for more information. It is well known that $A \prec_{(\log)} B$ implies $|||A||| \le |||B|||$ for all unitarily invariant norms $|||\cdot|||$.

Theorem 5.5. *Let $\mu \in \mathcal{P}^1(\mathbb{P}_m)$, where $\mathbb{P}_m$ is the open convex cone of all $m \times m$ positive definite matrices. Then*

$$|||\beta_\Lambda(\mu)^{1/p}||| \le |||\beta_\Lambda(\mu^{1/p})|||$$

for all $p \ge 1$ and all unitarily invariant norms $|||\cdot|||$.

Proof. Let $\mu \in \mathcal{P}^1(\mathbb{P}_m)$. Since $\mathcal{P}_0(\mathbb{P}_m)$ is d_1^W-dense in $\mathcal{P}^1(\mathbb{P}_m)$, there exists a sequence $\{\mu_n\} \subset \mathcal{P}_0(\mathbb{P}_m)$ of the form $\mu_n = \frac{1}{n}\sum_{j=1}^n \delta_{A_j}$ such that $\mu_n \to \mu$ as $n \to \infty$.

For $1 \le k \le m$, let Γ^k be the k-th asymmetric tensor power: see [2, 4] for basic properties of Γ^k. Then for $A, A_1, \dots, A_n > 0$

$$\Lambda_n(\Gamma^k A_1, \dots, \Gamma^k A_n) = \Gamma^k \Lambda_n(A_1, \dots, A_n),$$

$$\lambda_1(\Gamma^k A) = \prod_{i=1}^k \lambda_i(A),$$

$$\Gamma^k(A^p) = (\Gamma^k A)^p, \quad p > 0.$$

Using these properties, one can see from [8] that for $\mu_n \in \mathcal{P}_0(\mathbb{P}_m)$ and $p \geq 1$

$$\Lambda_n(A_1^p, \ldots, A_n^p)^{1/p} \prec_{(\log)} \Lambda_n(A_1, \ldots, A_n).$$

Substituting A_j by $A_j^{1/p}$ for all j, we have

$$\Lambda_n(A_1, \ldots, A_n)^{1/p} \prec_{(\log)} \Lambda_n(A_1^{1/p}, \ldots, A_n^{1/p}),$$

and thus, $|||\Lambda_n(A_1, \ldots, A_n)^{1/p}||| \leq |||\Lambda_n(A_1^{1/p}, \ldots, A_n^{1/p})|||$. By continuity of the norm function, the contraction property (5.1), and Lemma 3.6, we conclude

$$\begin{aligned} |||\beta_\Lambda(\mu)^{1/p}||| &= |||\lim_{n\to\infty} \beta_\Lambda(\mu_n)^{1/p}||| = \lim_{n\to\infty} |||\Lambda_n(A_1, \ldots, A_n)^{1/p}||| \\ &\leq \lim_{n\to\infty} |||\Lambda_n(A_1^{1/p}, \ldots, A_n^{1/p})||| = |||\lim_{n\to\infty} \beta_\Lambda(\mu_n^{1/p})||| = |||\beta_\Lambda(\mu^{1/p})|||. \quad \square \end{aligned}$$

Remark 5.6. It has been shown in [1,2] that

$$(A\#_t B)^\alpha \prec_{(\log)} A^\alpha \#_t B^\alpha$$

for any $A, B \in \mathbb{P}_m$, $0 < t < 1$ and $0 < \alpha < 1$. This implies that

$$|||(A\#_t B)^\alpha||| \leq |||A^\alpha \#_t B^\alpha||| \tag{5.2}$$

for all unitarily invariant norm $|||\cdot|||$. The inequality (5.2) is the special case of Theorem 5.5 when $\mu = (1-t)\delta_A + t\delta_B$ for $0 < t < 1$ and $\alpha = 1/p$ for $p \geq 1$. In other words, Theorem 5.5 is a generalization of the result (5.2).

Acknowledgments

I would like to thank to Professor Jimmie Lawson for a useful idea to solve the order relation of uniformly distributed probability measures, and also thank to referee for valuable comments. This work was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Science, ICT and Future Planning (2015R1C1A1A02036407).

References

[1] T. Ando, *Majorizations and inequalities in matrix theory*, Linear Algebra Appl. **199** (1994), 17–67.

[2] T. Ando and F. Hiai, *Log majorization and complementary Golden-Thompson type inequalities*, Linear Algebra Appl. **197** (1994), 113–131.

[3] R. Bhatia, *Positive Definite Matrices*, Princeton Series in Applied Mathematics, Princeton University Press, Princeton, NJ, 2007.

[4] R. Bhatia and R. Karandikar, *Monotonicity of the matrix geometric mean*, Math. Ann. **353** (2012), no. 4, 1453–1467.

[5] F. Bolley, *Separability and completeness for the Wasserstein distance*, in *Séminaire de Probabilités XLI*, 371–377, Lecture Notes in Mathematics **1934**, Springer, Berlin, 2008.

[6] G. Corach, H. Porta and L. Recht, *Convexity of the geodesic distance on spaces of positive operators*, Illinois J. Math. **38** (1994), no. 1, 87–94.

[7] M. Fujii and T. Furuta, *Löwner-Heinz, Cordes and Heinz-Kato inequalities*, Math. Japon. **38** (1993), no. 1, 73–78.

[8] F. Hiai and D. Petz, *Riemannian metrics on positive definite matrices related to means II*, Linear Algebra Appl. **436** (2012), no. 7, 2117–2136.

[9] J. Holbrook, *No dice: a deterministic approach to the Cartan centroid*, J. Ramanujan Math. Soc. **27** (2012), no. 4, 509–521.

[10] R. A. Horn and C. R. Johnson, *Matrix Analysis*, Second edition, Cambridge University Press, Cambridge, 2013.

[11] H. Karcher, *Riemannian center of mass and mollifier smoothing*, Comm. Pure Appl. Math. **30** (1977), no. 5, 509–541.

[12] K. Keimel and W. Roth, *Ordered Cones and Approximation*, Lecture Notes in Mathematics **1517**, Springer-Verlag, Berlin, 1992.

[13] E. Lawler, *Combinatorial Optimization: Networks and Matroids*, Dover, 2011.

[14] J. Lawson, *Ordered probability spaces*, J. Math. Anal. Appl. **455** (2017), no. 1, 167–179.

[15] J. Lawson and Y. Lim, *Metric convexity of symmetric cones*, Osaka J. Math. **44** (2007), no. 4, 795–816.

[16] ———, *Karcher means and Karcher equations of positive definite operators*, Trans. Amer. Math. Soc. Ser. B **1** (2014), 1–22.

[17] ———, *Contractive barycentric maps*, J. Operator Theory **77** (2017), no. 1, 87–107.

[18] Y. Lim and M. Pálfia, *Matrix power means and the Karcher mean*, J. Funct. Anal. **262** (2012), no. 4, 1498–1514.

[19] M. Moakher, *A differential geometric approach to the geometric mean of symmetric positive-definite matrices*, SIAM J. Matrix Anal. Appl. **26** (2005), no. 3, 735–747.

[20] R. D. Nussbaum, *Finsler structures for the part metric and Hilbert's projective metric and applications to ordinary differential equations*, Differential Integral Equations **7** (1994), no. 5-6, 1649–1707.

[21] M. Pálfia, *Operator means of probability measures and generalized Karcher equations*, Adv. Math. **289** (2016), 951–1007.

[22] K.-T. Sturm, *Probability measures on metric spaces of nonpositive curvature* in *Heat Kernels and Analysis on Manifolds, Graphs, and Metric Spaces* (Paris, 2002), 357–390, Contemp. Math. **338**, Amer. Math. Soc., Providence, RI, 2003.

[23] A. C. Thompson, *On certain contraction mappings in a partially ordered vector space*, Proc. Amer. Math. Soc. **14** (1963), 438–443.

Sejong Kim
Department of Mathematics, Chungbuk National University, Cheongju 28644, Korea
E-mail address: `skim@chungbuk.ac.kr`