TAIWANESE JOURNAL OF MATHEMATICS Vol. 17, No. 6, pp. 1967-1982, December 2013 DOI: 10.11650/tjm.17.2013.2944 This paper is available online at http://journal.taiwanmathsoc.org.tw

NO DICE THEOREM ON SYMMETRIC CONES

Sangho Kum*, Hosoo Lee and Yongdo Lim

Abstract. The monotonicity of the least squares mean on the Riemannian manifold of positive definite matrices, conjectured by Bhatia and Holbrook and one of key axiomatic properties of matrix geometric means, was recently established based on the Strong Law of Large Number [14, 4]. A natural question concerned with the S.L.L.N is so called *the no dice conjecture*. It is a problem to make a construction of deterministic sequences converging to the least squares mean without any probabilistic arguments. Very recently, Holbrook [7] gave an affirmative answer to the conjecture in the space of positive definite matrices. In this paper, inspired by the work of Holbrook [7] and the fact that the convex cone of positive definite matrices is a typical example of a symmetric cone (self-dual homogeneous convex cone), we establish the no dice theorem on general symmetric cones.

1. INTRODUCTION

The open convex cone $\mathbb{P} = \mathbb{P}_m$ in the Euclidean space of $m \times m$ Hermitian matrices is a Cartan-Hadamard Riemannian manifold, a simply connected complete Riemannian manifold with non-positive sectional curvature (the canonical 2-tensor is non-negative), equipped with the trace Riemannian metric $ds = ||A^{-1/2}dA A^{-1/2}||_2 = (\operatorname{tr}(A^{-1}dA)^2)^{1/2}$, where $|| \cdot ||_2$ denotes the Frobenius norm. The Riemannian metric distance between A and B is given by $\delta(A, B) = ||\log A^{-1/2}BA^{-1/2}||_2$ and the curve $t \mapsto A \#_t B$ is the unique (up to parametrization) geodesic line containing A and B and its unique metric midpoint A # B is the geometric mean of A and B.

The least squares mean (Cartan mean, Karcher mean or Riemannian center of mass) of positive definite matrices A_1, \ldots, A_n is defined to be the unique minimizer of the sum of squares of the Riemannian trace metric distances to each of the A_i , i.e.,

$$\Lambda(A_1,\ldots,A_n) = \operatorname*{arg\,min}_{X \in \mathbb{P}} \sum_{i=1}^n \delta^2(X,A_i).$$

Received February 1, 2013, accepted May 8, 2013.

Communicated by Jen-Chih Yao.

Key words and phrases: Least squares mean, Symmetric cone, No dice theorem.

The first author was supported by Basic Science Research Program through NRF Grant No. 2012R1A1A2041136.

*Corresponding author.

²⁰¹⁰ Mathematics Subject Classification: 47A64, 17C50, 15B48, 53C20.

The least squares mean of positive definite matrices for the Riemannian trace metric plays an important role in many applied areas involving averaging of positive definite matrices. Actually, this mean was suggested independently by M. Moakher [19] and Bhatia and Holbrook [3], and is regarded as a suitable extension of the geometric mean of two positive definite matrices to *n*-variables. The monotonicity of the least squares mean, conjectured by Bhatia and Holbrook [3] and one of key axiomatic properties of matrix geometric means, was recently established by Lawson and Lim [14] via a probabilistic convergence of approximations and by Bhatia and Karandikar [4] using some probabilistic counting arguments, and both arguments depend heavily on basic inequalities for the Riemannian metric. These two important results for the least squares mean are based on Sturms' theorem of Strong Law of Large Number [21]: Assign to $k \in \{1, \ldots, n\}$ the probability 1/n and giving $\{1, 2, \ldots, n\}^{\mathbb{N}}$ the product probability, we have

$$\Lambda(A_1,\ldots,A_n) = \lim_{k \to \infty} S_k(A_{\sigma(1)},\ldots,A_{\sigma(k)})$$

for almost all $\sigma \in \{1, 2, \ldots, n\}^{\mathbb{N}}$, where $S_k(A_1, \ldots, A_k)$ is the inductive mean of A_1, \ldots, A_k defined by $S_1(A_1) = A_1, S_k(A_1, \ldots, A_k) = S_{k-1}(A_1, \ldots, A_{k-1}) #_{\frac{1}{k}} A_k$.

Sturm's result is theoretically important but it is of stochastic nature and so it does not provide a deterministic sequence of points converging to the least squares mean. A natural question concerned with the preceding S.L.L.N is the following, called *the no dice theorem*;

Problem. Is it true that $\Lambda(A_1, \ldots, A_n) = \lim_{k \to \infty} S_k(A_{\sigma(1)}, \ldots, A_{\sigma(k)})$, where σ is the periodic sequence defined by $\sigma(nj+i) = i, 1 \le i \le n$.

It is a problem to make a construction of deterministic sequences converging to the least squares mean without any probabilistic arguments. Very recently, Holbrook [7] gave an affirmative answer in the space of positive definite matrices \mathbb{P} .

Inspired by the work of Holbrook [7] and the fact that the convex cone of positive definite matrices is a typical example of a symmetric cone, a self-dual homogeneous open convex cone (it is well known that the class of symmetric cones is precisely that of Euclidean Jordan algebras), we establish the no dice theorem of the least squares mean on general symmetric cones. To do so, the method by Holbrook [7] is adopted. As is well-known [5, 22], the family of symmetric cones contains Lorenz cones (second order cones) and the cones of positive definite matrices over real and complex numbers. Moreover, the symmetric cone in a given Euclidean Jordan algebra is, in a unique way, a direct sum of symmetric cones in the five types of simple Euclidean Jordan algebras. (For the details, see Tao and Gowda [22].)

2. EUCLIDEAN JORDAN ALGEBRAS AND SYMMETRIC CONES

In this section, we briefly describe (following mostly [5]) some Jordan-algebraic concepts pertinent to our purpose. A *Jordan algebra* V over \mathbb{R} is a finite-dimensional

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 Proposition II.3.1 and Proposition II.3.3 of [5].

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: $P(x)^{-1} = P(x^{-1}).$
- (ii) 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).

A Jordan algebra V 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 $x, y, z \in V$. An element $c \in V$ is called an idempotent if $c^2 = c \neq 0$. We say that c_1, \ldots, c_k is a complete system of orthogonal idempotents if $c_i^2 = c_i, c_i c_j = 0, i \neq j, c_1 + \cdots + c_k = e$. An idempotent is 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.

Theorem 2.2. (Spectral theorem, first version [5, Theorem III.1.1]). Let V be a Euclidean Jordan algebra. Given $x \in V$, there exist real numbers $\lambda_1, \ldots, \lambda_k$ all distinct and a unique complete system of orthogonal idempotents c_1, \ldots, c_k such that

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

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 [5, 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 rank(V)). Given $x \in V$, there exists a Jordan frame c_1, \ldots, c_r and real numbers $\lambda_1, \ldots, \lambda_r$ 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 $\operatorname{rank}(V) = r$. The spectral mapping $\lambda : V \to \mathbb{R}^r$ is defined by $\lambda(x) = (\lambda_1(x), \ldots, \lambda_r(x))$, where $\lambda_i(x)$'s are eigenvalues of x (with multiplicities) as in Theorem 2.3 in non-increasing order $\lambda_{\max}(x) = \lambda_1(x) \ge \lambda_2(x) \ge \cdots \ge \lambda_r(x) = \lambda_{\min}(x)$. Furthermore, $\det(x) = \prod_{i=1}^r \lambda_i(x)$ and $\operatorname{tr}(x) = \sum_{i=1}^r \lambda_i(x)$.

Throughout this paper, we assume that V is a Euclidean Jordan algebra of rank r and equipped with the trace inner product $\langle x, y \rangle = tr(xy)$. With the norm induced by the trace inner product, we have the following non-expansive (Lipschitz) property of the spectral mapping appearing in [1, 6] which is an immediate consequence of a Jordan-algebraic version of von Neumann-Theobald inequality [2, 18, 15].

Theorem 2.5. For $x, y \in V$,

$$\|\lambda(x) - \lambda(y)\|_2 \le \|x - y\|$$

where $||x|| = \sqrt{\operatorname{tr}(x^2)}$ and $||\cdot||_2$ denotes the usual Euclidean norm in \mathbb{R}^r .

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 element $x \in V$ such that L(x) is positive semi-definite. It turns out that Q has non-empty interior Ω , and Ω is a symmetric cone, that is, the group $G(\Omega) = \{g \in \operatorname{GL}(V) \mid g(\Omega) = \Omega\}$ acts transitively on it and Ω is a self-dual cone with respect to the inner product $\langle \cdot, \cdot \rangle$. Furthermore, for any a in Ω , $P(a) \in G(\Omega)$ and is positive definite.

As mentioned in the introduction, any symmetric cone (self-dual, homogeneous open convex cone) can be realized as an interior of squares in an appropriate Euclidean Jordan algebra [5].

Proposition 2.6. The symmetric cone Ω has the following properties:

- (i) $\Omega = \{x^2 \mid x \text{ is invertible}\};$
- (*ii*) $\Omega = \{x \in V \mid L(x) \text{ is positive definite}\}; and$
- (iii) $\Omega = \{ x \in V \mid \lambda_{\min}(x) > 0 \}.$

We note that $\overline{\Omega} = \{x \in V \mid \lambda_i(x) \ge 0, i = 1, \dots, r\}$. For $x, y \in V$, we define

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

and x < y if $y - x \in \Omega$. Clearly $\overline{\Omega} = \{x \in V \mid x \ge 0\}$ and $\Omega = \{x \in V \mid x > 0\}$.

Proposition 2.7. The inversion $x \to x^{-1}$ on Ω is order reverting. That is, $a \leq b$ if and only if $b^{-1} \leq a^{-1}$ for any $a, b \in \Omega$.

3. KARCHER MEANS ON SYMMETRIC CONES

It turns out [5] that the symmetric cone Ω admits a $G(\Omega)$ -invariant Riemannian metric defined by

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

The inversion $j(x) = x^{-1}$ is an involutive isometry fixing *e*. It is a symmetric Riemannian space of non-compact type with respect to its distance metric [10, 11, 16]. The

1970

unique geodesic curve joining a and b is $t \mapsto a \#_t b := P(a^{1/2})(P(a^{-1/2})b)^t$ and the Riemannian distance $\delta(a, b)$ is given by

$$\delta(a,b) = \left(\sum_{i=1}^{r} \log^2 \lambda_i (P(a^{-1/2})b)\right)^{1/2}$$

The geodesic middle $a\#b := a\#_{1/2}b = P(a^{1/2})(P(a^{-1/2})b)^{1/2}$ is called the *geometric* mean of a and b. An important property of the metric δ is the semiparallelogram law

$$\delta^{2}(z, x \# y) \leq \frac{1}{2}\delta^{2}(z, x) + \frac{1}{2}\delta^{2}(z, y) - \frac{1}{4}\delta^{2}(x, y)$$

and its general form is for any $t \in [0, 1]$

(3.1)
$$\delta^2(z, x \#_t y) \le (1-t)\delta^2(z, x) + t\delta^2(z, y) - t(1-t)\delta^2(x, y).$$

The metric space (Ω, δ) is an important example of a Hadamard space, a complete metric space satisfying the semiparallelogram law.

The Karcher mean, also called the *least squares mean*, of n elements a_1, \ldots, a_n on Ω is defined as the unique minimizer of the sum of squares of the Riemannian trace metric distances to each of the a_i , i.e.,

(3.2)
$$\Lambda(a_1,\ldots,a_n) = \operatorname*{arg\,min}_{x\in\Omega} \sum_{i=1}^n \delta^2(x,a_i).$$

Observe that the least squares mean exists and unique for any NPC spaces.

The inductive mean $S_n(a_1, \ldots, a_n)$ is defined by

$$S_1(a_1) = a_1, \quad S_n(a_1, \dots, a_n) = S_{n-1}(a_1, \dots, a_{n-1}) #_{\frac{1}{n}} a_n.$$

For a fixed $n \in \mathbb{N}$, equipped $\{1, 2, ..., n\}$ with the point mass measure, the set $P_n := \{1, 2, ..., n\}^{\mathbb{N}}$ of all functions from \mathbb{N} to $\{1, 2, ..., n\}$ is a probability space with measure 1. By the Strong Law of Large Number in Sturm [21] (see also [14]), we have

Theorem 3.1. Let $a_1, \ldots, a_n \in \Omega$. Then for almost all $\sigma \in \{1, 2, \ldots, n\}^{\mathbb{N}}$,

$$\Lambda(a_1,\ldots,a_n) = \lim_{k \to \infty} S(a_{\sigma(1)},\ldots,a_{\sigma(k)}).$$

As stated in the introduction, a natural question concerned with the preceding S.L.L.N is the following, called the no dice theorem:

Problem. Is it true that $\Lambda(a_1, \ldots, a_n) = \lim_{k \to \infty} S_k(a_{\sigma(1)}, \ldots, a_{\sigma(k)})$, where $\sigma(nj+i) = i, \ 1 \le i \le n$.

In the next section, we will prove the no dice theorem on symmetric cones.

4. A PROOF OF NO DICE THEOREM

Proposition 4.1. Let V be a finite dimensional inner product space and $T: V \to V$ be a positive (semi)definite linear operator. Then the (operator) norm ||T|| is the maximum eigenvalue of T.

Proposition 4.2. Let V be a Euclidean Jordan algebra and D be a bounded subset of V. Then the quadratic representation $P : D \rightarrow \mathcal{L}(V, V)$ is Lipschtz continuous. Here $\mathcal{L}(V, V)$ denotes the normed space of every (continuous) linear operator from V to V.

Proof. First observe that for all $x, y \in D$,

(4.1)
$$\begin{aligned} \|L(x)L(x) - L(y)L(y)\| &= \|(L(x) - L(y))L(y) + L(x)(L(x) - L(y))\| \\ &\leq (\|L(x)\| + \|L(y)\|)\|L(x) - L(y)\| \\ &\leq \|L\|^2(\|x\| + \|y\|)\|x - y\| \\ &\leq M_1\|x - y\| \text{ for some } M_1 \end{aligned}$$

where ||L|| denotes the norm of the linear operator $L : V \to \mathcal{L}(V, V)$ defined by $x \mapsto L(x)$. Moreover, for all $x, y \in D$,

(4.2)
$$\begin{aligned} \|L(x^2) - L(y^2)\| &= \|L(x^2 - y^2)\| \le \|L\| \|x^2 - y^2\| \le \|L\|^2 \|x + y\| \|x - y\| \\ &\le M_2 \|x - y\| \text{ for some } M_2. \end{aligned}$$

The second inequality follows from the observation $||xy|| = ||L(x)y|| \le ||L|| ||x|| ||y||$ for all $x, y \in V$. The definition of $P(x) = 2L^2(x) - L(x^2)$ together with (4.1) and (4.2) yields the conclusion.

Definition 4.3. Let $x, y \in \Omega \subset V$. The geodesic distance between x and y is defined to be $\delta(x, y) = \|\log P(x^{-\frac{1}{2}})y\|$. Given $a_k \in \Omega$, $(k = 1, \dots, m)$, we define, for each $x \in \Omega$ and each positive integer n,

$$ss(x) = \sum_{k=1}^{m} \delta^{2}(x, a_{k}),$$

$$\varphi_{n}(x) = x \#_{\frac{1}{n}} a_{k} = P(x^{\frac{1}{2}}) (P(x^{-\frac{1}{2}}) a_{k})^{\frac{1}{n}} \text{ with } k \equiv n \pmod{m},$$

$$L_{k}(x) = \log P(x^{-\frac{1}{2}}) a_{k}, \quad S(x) = \sum_{k=1}^{m} L_{k}(x).$$

In addition, the Karcher mean of a_i 's is defined to be the unique minimizer of $ss(x) = \sum_{k=1}^{m} \delta^2(x, a_k)$ over Ω .

Proposition 4.4. [9, 17]. We have the following gradient formula:

$$\nabla ss(x) = 2P(x^{-\frac{1}{2}}) \left(\sum_{k=1}^{m} \log P(x^{\frac{1}{2}}) a_k^{-1} \right) = -2P(x^{-\frac{1}{2}})S(x).$$

Lemma 4.5. For two positive numbers α and β , let $D = \{x \in V \mid \alpha e \leq x \leq \beta e\}$ contain $\{a_1, \dots, a_m\}$. Then there exists $M < \infty$ such that for all $x \in D$ and all n, we have

(4.3)
$$\|\varphi_n(x) - x - \frac{1}{n} P(x^{\frac{1}{2}}) L_k(x)\| \le \frac{M}{n^2},$$

where $k \equiv n \pmod{m}$ and

(4.4)
$$\|\varphi_n(x) - x\| \le \frac{M}{n}.$$

Proof. Let $x = \sum_{i=1}^{r} \lambda_i c_i \in D$. Since $\frac{1}{n} \log a = \log a^{\frac{1}{n}}$ for all $a \in \Omega$, we get

$$\exp\left(\frac{1}{n}\log P(x^{-\frac{1}{2}})a_k\right) = \exp\left(\log(P(x^{-\frac{1}{2}})a_k)^{\frac{1}{n}}\right) = (P(x^{-\frac{1}{2}})a_k)^{\frac{1}{n}}$$

and

$$\varphi_n(x) = P(x^{\frac{1}{2}})(P(x^{-\frac{1}{2}})a_k)^{\frac{1}{n}}$$

= $P(x^{\frac{1}{2}})\exp\left(\frac{1}{n}\log P(x^{-\frac{1}{2}})a_k\right) = P(x^{\frac{1}{2}})\exp\left(\frac{L_k(x)}{n}\right).$

This implies that

$$\varphi_n(x) - x = P(x^{\frac{1}{2}}) \left(\exp\left(\frac{L_k(x)}{n}\right) - e \right)$$

= $P(x^{\frac{1}{2}}) \left(e + \frac{L_k(x)}{n} + \frac{1}{2!} \left(\frac{L_k(x)}{n}\right)^2 + \dots - e \right)$
= $P(x^{\frac{1}{2}}) \left(\frac{L_k(x)}{n}\right) + E = \frac{1}{n} P(x^{\frac{1}{2}}) L_k(x) + E$

where

$$E = P(x^{\frac{1}{2}}) \left(\sum_{j=2}^{\infty} \frac{1}{j!} \left(\frac{L_k(x)}{n} \right)^j \right).$$

So

(4.5)
$$||E|| = \left\| P(x^{\frac{1}{2}}) \left(\sum_{j=2}^{\infty} \frac{1}{j!} \left(\frac{L_k(x)}{n} \right)^j \right) \right\| \le ||P(x^{\frac{1}{2}})|| \left\| \sum_{j=2}^{\infty} \frac{1}{j!} \left(\frac{L_k(x)}{n} \right)^j \right\|.$$

As $P(x^{\frac{1}{2}})$ is a positive definite linear operator from V to V itself, by Proposition 4.1, $||P(x^{\frac{1}{2}})||$ is $\kappa = \max\{\sqrt{\lambda_i\lambda_j} \mid \lambda_i, \lambda_j \text{ are eigenvalues of } x\}$ because the eigenvalues of P(x) are of the form $\lambda_i\lambda_j$ [20, Theorem 3.1]. Since $\forall i, \alpha \leq \lambda_i \leq \beta$, we may assume that there is a constant M_1 such that

(4.6)
$$||P(x)||, ||P(x^{\frac{1}{2}})||, ||P(x^{-\frac{1}{2}})|| \le M_1 \text{ for all } x \in D.$$

On the other hand, $\alpha e \leq a_k \leq \beta e$ so that

$$\alpha P(x^{-\frac{1}{2}}) e \le P(x^{-\frac{1}{2}}) a_k \le \beta P(x^{-\frac{1}{2}}) e$$
, that is, $\alpha x^{-1} \le P(x^{-\frac{1}{2}}) a_k \le \beta x^{-1}$.

Hence, we have

$$\alpha \frac{e}{\beta} \le \alpha x^{-1} \le P(x^{-\frac{1}{2}})a_k \le \beta x^{-1} \le \beta \frac{e}{\alpha}$$

which, in turn, leads to

$$\log \frac{\alpha}{\beta} e \le \log P(x^{-\frac{1}{2}})a_k \le \log \frac{\beta}{\alpha} e.$$

So

$$-\log\frac{\beta}{\alpha} \le \log\mu_i \le \log\frac{\beta}{\alpha} \iff |\log\mu_i| \le \log\frac{\beta}{\alpha}$$

where μ_i 's are the eigenvalues of $P(x^{-\frac{1}{2}})a_k$. Therefore, we get

$$\sum_{i=1}^{r} (\log \mu_i)^2 \le r \left(\log \frac{\beta}{\alpha} \right)^2$$

where $r = \operatorname{rank} V$. This implies that

(4.7)
$$||L_k(x)|| = ||\log P(x^{-\frac{1}{2}})a_k|| = \left(\sum_{i=1}^r (\log \mu_i)^2\right)^{\frac{1}{2}} \le \sqrt{r} \log \frac{\beta}{\alpha}$$

Moreover,

(4.8)
$$\sum_{j=2}^{\infty} \frac{1}{j!} \left(\frac{\|L_k(x)\|}{n} \right)^j \le \frac{1}{2} \left(\frac{\|L_k(x)\|}{n} \right)^2 \exp\left(\frac{\|L_k(x)\|}{n} \right).$$

By (4.5), (4.6) and (4.8), we see

$$||E|| \le M_1 \sum_{j=2}^{\infty} \frac{1}{j!} \left(\frac{||L_k(x)||}{n}\right)^j \le \frac{M_1}{2} \left(\frac{||L_k(x)||}{n}\right)^2 \exp\left(\frac{||L_k(x)||}{n}\right)$$
$$\le M_1 \frac{r \log^2 \frac{\beta}{\alpha}}{2n^2} \left(\frac{\beta}{\alpha}\right)^{\frac{\sqrt{r}}{n}}$$
$$\le \frac{M_1 r \log^2 \frac{\beta}{\alpha}}{2n^2} \left(\frac{\beta}{\alpha}\right)^{\sqrt{r}} = \frac{M}{n^2}$$

where the first inequality comes from the fact that $\forall a \in V$, $||a||^n \ge ||a^n||$. Consequently, by (4.3), (4.6) and (4.7), we have

$$\|\varphi_n(x) - x\| \le \frac{M}{n^2} + \frac{1}{n} \|P(x^{\frac{1}{2}})\| \|L_k(x)\| \le \frac{M}{n^2} + \frac{M}{n} \le \frac{M}{n}.$$

This completes the proof.

Lemma 4.6. Let D be as in Lemma 4.5. Then there exists $M < \infty$ such that for all $x, y \in D$, we have

(4.9)
$$||x^{\frac{1}{2}} - y^{\frac{1}{2}}|| \le M ||x - y||,$$

(4.10)
$$||L_k(x) - L_k(y)|| \le M ||x - y||,$$

(4.11)
$$||P(x^{\frac{1}{2}})L_k(x) - P(y^{\frac{1}{2}})L_k(y)|| \le M ||x - y||.$$

Proof. By Sun and Sun [20, Theorem 13], the map $x \to x^{\frac{1}{2}}$ is continuously differentiable on D because the real valued function $f(x) = \sqrt{x}$ is continuously differentiable on the interval $[\alpha, \beta]$. Since D is a compact convex set, we see that the map $x \to x^{\frac{1}{2}}$ is Lipschtz continuous on D by the mean value theorem. This entails (4.9). In addition,

$$\begin{aligned} \|L_k(x) - L_k(y)\| &= \|\log P(x^{-\frac{1}{2}})a_k - \log P(y^{-\frac{1}{2}})a_k\| \\ &\leq M_1 \|P(x^{-\frac{1}{2}})a_k - P(y^{-\frac{1}{2}})a_k\| \text{ for some } M_1 \\ &\leq M_2 \|P(x^{\frac{1}{2}})a_k^{-1} - P(y^{\frac{1}{2}})a_k^{-1}\| \text{ for some } M_2 \\ &\leq M_2 \|P(x^{\frac{1}{2}}) - P(y^{\frac{1}{2}})\|\|a_k^{-1}\| \\ &\leq M_3 \|x^{\frac{1}{2}} - y^{\frac{1}{2}}\| \text{ for some } M_3 \\ &\leq M \|x - y\| \text{ for some } M. \end{aligned}$$

The first and second inequalities come from Sun and Sun [20, Theorem 13] and the mean value theorem repeatedly to the maps $x \to \log x$ and $x \to x^{-1}$ (the corresponding real-valued functions are $f(x) = \log x$ and $g(x) = \frac{1}{x}$ on the positive interval $[\frac{\alpha}{\beta}, \frac{\beta}{\alpha}]$). The fourth and the last follow from Proposition 4.2 and (4.9), respectively. For (4.11), note that

$$\begin{aligned} &\|P(x^{\frac{1}{2}})L_{k}(x) - P(y^{\frac{1}{2}})L_{k}(y)\| \\ &= \|P(x^{\frac{1}{2}})(L_{k}(x) - L_{k}(y)) + (P(x^{\frac{1}{2}}) - P(y^{\frac{1}{2}}))L_{k}(y)\| \\ &\leq \|P(x^{\frac{1}{2}})\|\|L_{k}(x) - L_{k}(y)\| + \|(P(x^{\frac{1}{2}}) - P(y^{\frac{1}{2}}))L_{k}(y)\| \\ &\leq M_{4}\|x - y\| + M_{5}\|x^{\frac{1}{2}} - y^{\frac{1}{2}}\|\|L_{k}(y)\| \quad \exists M_{4} \text{ and } M_{5} \\ &\leq M\|x - y\|. \end{aligned}$$

The second inequality is due to (4.6), (4.10) and Proposition 4.2. The third is from (4.9) and (4.7). This completes the proof.

Lemma 4.7. Let D be as in Lemma 4.5. Then there exists $M < \infty$ such that, whenever x and x + ty ($0 \le t \le 1$) belong to D,

$$(4.12) |ss(x+y) - ss(x) - \langle \nabla ss(x), y \rangle| \le M ||y||^2.$$

Proof. First note that

$$\frac{d}{dt}ss(x+ty) = \langle \nabla ss(x+ty), y \rangle$$

Thus

(4.13)
$$ss(x+y) - ss(x) = \int_0^1 \langle \nabla ss(x+ty), y \rangle \, dt = \langle \nabla ss(x), y \rangle + E$$

where

$$E = \int_0^1 \langle \nabla ss(x+ty) - \nabla ss(x), y \rangle \, dt.$$

By Proposition 4.4, we have $\nabla ss(x) = -2P(x^{-\frac{1}{2}})S(x)$. Hence for any $z \in D$,

$$\begin{aligned} \|\nabla ss(z) - \nabla ss(x)\| \\ &= 2\|P(z^{-\frac{1}{2}})S(z) - P(x^{-\frac{1}{2}})S(x)\| \\ &= 2\|P(z^{-\frac{1}{2}})(S(z) - S(x)) + (P(z^{-\frac{1}{2}}) - P(x^{-\frac{1}{2}}))S(x)\| \\ &\leq 2\|P(z^{-\frac{1}{2}})\|\|S(z) - S(x)\| + 2\|(P(z^{-\frac{1}{2}}) - P(x^{-\frac{1}{2}}))S(x)\| \\ &\leq M_1\|z - x\| + M_2\|z^{-\frac{1}{2}} - x^{-\frac{1}{2}}\| \quad \text{for some } M_1 \text{ and } M_2 \\ &\leq M\|z - x\|. \end{aligned}$$

The second inequality follows from (4.6), (4.10), (4.7) and Proposition 4.2. Also the third one is derived from using Sun and Sun [20, Theorem 13] and the mean value theorem to the map $x \to x^{-\frac{1}{2}}$ (the corresponding real-valued function is $f(x) = \frac{1}{\sqrt{x}}$ on the positive interval $[\alpha, \beta]$). So, by (4.14), when we put $z = x + ty \in D$,

$$|E| \le \max_{0 \le t \le 1} \|\nabla ss(x+ty) - \nabla ss(x)\| \|y\| \le M \|y\|^2.$$

Therefore, we obtain from (4.13) that

$$|E| = |ss(x+y) - ss(x) - \langle \nabla ss(x), y \rangle| \le M ||y||^2.$$

This completes the proof.

For positive integers a, d, let

$$\varphi_{a,d} = \varphi_{a+d-1} \circ \varphi_{a+d-2} \circ \cdots \circ \varphi_a.$$

Lemma 4.8. Given $x, a_1, \dots, a_m \in \Omega$, let

$$r = \max{\{\delta(\Lambda, x), \delta(\Lambda, a_1), \cdots, \delta(\Lambda, a_m)\}}$$
 and

$$\mathcal{K}_x = \{ y \in \Omega \mid \delta(\Lambda, y) \le r \} \supseteq \{ x, a_1, \cdots, a_m \}$$

where Λ is the Cartan centroid of a_i 's. Then there exists $M < \infty$ such that for all $y \in \mathcal{K}_x$ and $n \in \mathbb{N}$,

(4.15)
$$\|\varphi_{n,m}(y) - y - \frac{1}{n}P(y^{\frac{1}{2}})S(y)\| \le \frac{M}{n^2}.$$

Proof. Since \mathcal{K}_x is compact, there exist two positive numbers α and β such that $D = \{y \in \Omega \mid \alpha e \leq y \leq \beta e\}$ contains \mathcal{K}_x . Note that for any $y \in \mathcal{K}_x$ and $n, d \in \mathbb{N}$, $\varphi_n(y)$ and $\varphi_{n,d}(y)$ always belong to \mathcal{K}_x by definition. So replacing n by n+1 and y by $\varphi_n(y)$ in (4.3) yields that

(4.16)
$$\|\varphi_{n,2}(y) - \varphi_n(y) - \frac{1}{n+1} P(\varphi_n(y)^{\frac{1}{2}})(L_{k+1}(\varphi_n(y)))\| \le \frac{M}{(n+1)^2}$$

where k + 1 is computed modulo m. Moreover, by (4.11), we see, for some $M < \infty$,

(4.17)
$$\begin{aligned} \|\frac{1}{n+1}P(\varphi_n(y)^{\frac{1}{2}})(L_{k+1}(\varphi_n(y))) \\ -\frac{1}{n+1}P(y^{\frac{1}{2}})(L_{k+1}(y))\| \le \frac{M}{n+1}\|\varphi_n(y) - y\|. \end{aligned}$$

Hence we have, for some $M < \infty$,

$$\begin{split} \|\varphi_{n,2}(y) - y - \frac{1}{n}P(y^{\frac{1}{2}})(L_{k}(y) + L_{k+1}(y))\| \\ &\leq \|\varphi_{n,2}(y) - \varphi_{n}(y) - \frac{1}{n}P(y^{\frac{1}{2}})(L_{k+1}(y))\| + \|\varphi_{n}(y) - y - \frac{1}{n}P(y^{\frac{1}{2}})(L_{k}(y))\| \\ &\leq \frac{M}{n^{2}} + \|\varphi_{n,2}(y) - \varphi_{n}(y) - \frac{1}{n+1}P(\varphi_{n}(y)^{\frac{1}{2}})(L_{k+1}(\varphi_{n}(y)))\| \\ &\quad + \|\frac{1}{n+1}P(\varphi_{n}(y)^{\frac{1}{2}})(L_{k+1}(\varphi_{n}(y))) - \frac{1}{n}P(y^{\frac{1}{2}})(L_{k+1}(y))\| \\ &\leq \frac{M}{n^{2}} + \frac{M}{(n+1)^{2}} + \|\frac{1}{n+1}P(\varphi_{n}(y)^{\frac{1}{2}})(L_{k+1}(\varphi_{n}(y))) - \frac{1}{n+1}P(y^{\frac{1}{2}})(L_{k+1}(y))\| \\ &\quad + \left(\frac{1}{n} - \frac{1}{n+1}\right)\|P(y^{\frac{1}{2}})(L_{k+1}(y))\| \\ &\leq \frac{M}{n^{2}} + \frac{M}{(n+1)^{2}} + \frac{M}{n^{2}} + \frac{1}{n^{2}}\|P(y^{\frac{1}{2}})(L_{k+1}(y))\| \leq \frac{M}{n^{2}} \end{split}$$

by means of (4.3), (4.16), (4.17), (4.6) and (4.7). Proceeding in this way for m-steps, we get the conclusion (4.15).

Lemma 4.9. Let \mathcal{K}_x and D be as in Lemma 4.8. Then there exists $M < \infty$ such that for all $z \in \mathcal{K}_x$ and $n \in \mathbb{N}$,

(4.18)
$$ss(\varphi_{n,m}(z)) - ss(z) + \frac{1}{n} \|S(z)\|^2 \le -\frac{1}{n} \|S(z)\|^2 + \frac{M}{n^2}.$$

Proof. Put x = z, $y = \varphi_{n,m}(z) - z$ in Lemma 4.7. Clearly $z + ty = (1 - t)z + t\varphi_{n,m}(z) \in D$, $\forall t \in [0, 1]$ because D is convex. Thus Lemma 4.7 is available so that

$$ss(\varphi_{n,m}(z)) - ss(z) - \langle \nabla ss(z), \varphi_{n,m}(z) - z \rangle | \le M ||y||^2$$
 for some $M < \infty$.

Since (4.15) implies $||y|| \leq \frac{1}{n} ||P(z^{\frac{1}{2}})S(z)|| + \frac{M}{n^2} \leq \frac{M}{n}$ for some $M < \infty$, by (4.6) and (4.7), we obtain

$$|ss(\varphi_{n,m}(z)) - ss(z) - \langle \nabla ss(z), \varphi_{n,m}(z) - z \rangle| \le \frac{M}{n^2}$$
 for some $M < \infty$.

Hence, for some $M < \infty$,

$$|ss(\varphi_{n,m}(z)) - ss(z) - \langle \nabla ss(z), \frac{1}{n} P(z^{\frac{1}{2}}) S(z) \rangle| \\ \leq \frac{M}{n^2} + \|\nabla ss(z)\| \|\varphi_{n,m}(z) - z - \frac{1}{n} P(z^{\frac{1}{2}}) S(z) \rangle\| \leq \frac{M}{n^2} + \frac{M}{n^2} \leq \frac{M}{n^2}$$

by (4.15), (4.6) and (4.7). That is,

(4.19)
$$|ss(\varphi_{n,m}(z)) - ss(z) + \frac{2}{n} ||S(z)||^2| \le \frac{M}{n^2}$$

because

$$\begin{split} \langle \nabla ss(z), \frac{1}{n} P(z^{\frac{1}{2}}) S(z) \rangle &= -\frac{2}{n} \langle P(z^{-\frac{1}{2}}) S(z), P(z^{\frac{1}{2}}) S(z) \rangle \\ &= -\frac{2}{n} \langle S(z), P(z^{-\frac{1}{2}}) P(z^{\frac{1}{2}}) S(z) \rangle = -\frac{2}{n} \| S(z) \|^2. \end{split}$$

It follows immediately from (4.19) that (4.18) holds.

Theorem 4.10. We have

$$\varphi_{1,d}(x) \to \Lambda \text{ as } d \to \infty.$$

Proof. Given $\epsilon > 0$, let

$$\mathcal{K}_{\epsilon} = \{ z \in \mathcal{K}_x \mid ss(z) \ge ss(\Lambda) + \epsilon \}$$

and

$$\mu = \inf\{\|S(z)\|^2 \mid z \in \mathcal{K}_{\epsilon}\}.$$

Since $\nabla ss(\cdot)$ vanishes only at Λ and \mathcal{K}_{ϵ} is compact, $\mu > 0$ is attained by Proposition 4.4. Let $z \in \mathcal{K}_x$.

Step 1. For *n* with $n \ge \frac{M}{\mu}$ (*M* in (4.18)), the sequence $\{\varphi_{n,d}(z) \mid d \in \mathbb{N}\}$ has infinitely many elements not contained in \mathcal{K}_{ϵ} .

Case (i)
$$z \in \mathcal{K}_{\epsilon}$$
.
As $n \ge \frac{M}{\mu}$, from (4.18) we obtain

(4.20)
$$ss(\varphi_{n,m}(z)) - ss(z) \le -\frac{1}{n} \|S(z)\|^2 \le -\frac{\mu}{n}$$

Suppose that $\varphi_{n,d}(z) \in \mathcal{K}_{\epsilon}$ for all $d \in \mathbb{N}$. As $\varphi_{n,lm}(z) = \varphi_{n+(l-1)m,m}(\varphi_{n,(l-1)m}(z))$ for all $l \ge 2$, from (4.20) we get

$$ss(\varphi_{n,lm}(z)) - ss(\varphi_{n,(l-1)m}(z)) = ss(\varphi_{n+(l-1)m,m}(\varphi_{n,(l-1)m}(z))) - ss(\varphi_{n,(l-1)m}(z)) \\ \leq -\frac{\mu}{n+(l-1)m}.$$

Hence

(4.21)

$$ss(\varphi_{n,lm}(z)) - ss(\Lambda)$$

$$= ss(\varphi_{n,lm}(z)) - ss(\varphi_{n,(l-1)m}(z)) + \dots + ss(\varphi_{n,m}(z)) - ss(\Lambda)$$

$$\leq -\mu \left(\frac{1}{n+m} + \dots + \frac{1}{n+(l-1)m}\right) + ss(\varphi_{n,m}(z)).$$

The RHS of (4.21) tends to $-\infty$ as $l \to \infty$, which implies that for sufficiently large $l \in \mathbb{N},$

$$ss(\varphi_{n,lm}(z)) < ss(\Lambda).$$

This is a contradiction to the definition of Λ . Thus, there is d_0 such that $\varphi_{n,d_0}(z)$ is not contained in \mathcal{K}_{ϵ} . Assume that there are only finitely many $\varphi_{n,d_0}(z), \cdots, \varphi_{n,d_p}(z)$ not contained in \mathcal{K}_{ϵ} . Put $h = \max\{d_0, \cdots, d_p\}$. Then $\varphi_{n,h+d}(z) \in \mathcal{K}_{\epsilon}$ for all $d \in \mathbb{N}$. Taking d = lm and repeating the similar argument above, we see that for sufficiently large $l \in \mathbb{N}$,

$$ss(\varphi_{n,h+lm}(z)) < ss(\Lambda),$$

which contradicts the definition of Λ . Hence, there are infinitely many elements of $\{\varphi_{n,d}(z) \mid d \in \mathbb{N}\}$ not contained in \mathcal{K}_{ϵ} .

Case (ii) $z \in \mathcal{K}_{\epsilon}^{c}$. Then either $\varphi_{n,1}(z) \in \mathcal{K}_{\epsilon}$ or $\varphi_{n,1}(z) \in \mathcal{K}_{\epsilon}^{c}$. If $\varphi_{n,1}(z) \in \mathcal{K}_{\epsilon}$, the sequence $\{\varphi_{n,d+1}(z) = \varphi_{n+1,d}(\varphi_{n,1}(z)) \mid d \in \mathbb{N}\}$ has infinitely many elements not contained

in \mathcal{K}_{ϵ} by the same argument as Case (i). When $\varphi_{n,1}(z) \in \mathcal{K}_{\epsilon}^{c}$, either $\varphi_{n,2}(z) \in \mathcal{K}_{\epsilon}$ or $\varphi_{n,2}(z) \in \mathcal{K}_{\epsilon}^{c}$. Repeating the same procedure to $\varphi_{n,2}(z)$, $\varphi_{n,3}(z)$,... as in the case of $\varphi_{n,1}(z)$, we can conclude that the sequence $\{\varphi_{n,d}(z) \mid d \in \mathbb{N}\}$ has infinitely many elements not contained in \mathcal{K}_{ϵ} .

Step 2. The sequence $\{\varphi_{1,d}(z) \mid d \in \mathbb{N}\}$ has infinitely many elements not contained in \mathcal{K}_{ϵ} .

Indeed, we may assume that $d \ge n \ge \frac{M}{\mu}$. Since $\varphi_{1,d}(z) = \varphi_{n,d-n+1}(\varphi_{1,n-1}(z))$, by Step 1, the sequence $\{\varphi_{1,d}(z) \mid d \in \mathbb{N}\}$ has infinitely many elements not contained in \mathcal{K}_{ϵ} .

Step 3. For sufficiently large k, we have $ss(\varphi_{1,k}(z)) < ss(\Lambda) + 2\epsilon$.

Let $n \geq \max\{\frac{M}{\mu}, \frac{mM^2}{\epsilon}\}$ (*M* satisfyng (4.4), (4.18)). By Step 2, we can choose positive integers $k_1 > k_0 > n$ such that $\varphi_{1,k_0}(z), \varphi_{1,k_1}(z) \in \mathcal{K}^c_{\epsilon}$, (i.e., $ss(\varphi_{1,k_i}(z)) < ss(\Lambda) + \epsilon$ for i = 1, 2) and $\varphi_{1,k}(z) \in \mathcal{K}_{\epsilon}$ for all $k_0 < k < k_1$. We will show that

(4.22)
$$ss(\varphi_{1,k}(z)) < ss(\Lambda) + 2\epsilon$$
, for all $k_0 < k < k_1$.

(i) $k_0 < k \le k_0 + m$. Then, by (4.4), we have

$$\begin{aligned} \|\varphi_{1,k}(z) - \varphi_{1,k_0}(z)\| &\leq \sum_{j=1}^{k-k_0} \|\varphi_{1,k_0+j}(z) - \varphi_{1,k_0+j-1}(z)\| \\ &= \sum_{j=1}^{k-k_0} \|\varphi_{k_0+j}(\varphi_{1,k_0+j-1}(z)) - \varphi_{1,k_0+j-1}(z)\| \\ &\leq \frac{mM}{n} \end{aligned}$$

We may assume $M \ge \max_{z \in D} \|\nabla ss(z)\|$ (D in Lemma 4.8). By the mean value theorem,

$$|ss(\varphi_{1,k}(z)) - ss(\varphi_{1,k_0}(z))| \le M \|\varphi_{1,k}(z) - \varphi_{1,k_0}(z)\| \le \frac{mM^2}{n}$$

so that

$$ss(\varphi_{1,k}(z)) \le ss(\varphi_{1,k_0}(z)) + \frac{mM^2}{n} \le ss(\Lambda) + \epsilon + \frac{mM^2}{n} \le ss(\Lambda) + 2\epsilon.$$

(ii) $k = k_0 + m + 1 < k_1$. From (4.20) and (i) above we obtain

$$ss(\varphi_{1,k}(z)) = ss(\varphi_{k_0+2,m}(\varphi_{1,k_0+1}(z))) \le ss(\varphi_{1,k_0+1}(z)) \le ss(\Lambda) + 2\epsilon.$$

Using this argument in a similar way, we get the assertion (4.22). As $k_i \to \infty$ increasingly by Step 2, Step 3 holds true.

$$\lim_{d \to \infty} ss(\varphi_{1,d}(x)) = ss(\Lambda),$$

equivalently,

$$\varphi_{1,d}(x) \to \Lambda \text{ as } d \to \infty.$$

This completes the proof.

Remark. Since the work of Lawson and Lim [14], the Karcher mean has sprung up in time as the most attractive averaging in many applied areas among other multivariable geometric means. In fact, a currently active research topic in linear algebra is understanding, finding properties of, and computing efficiently the least squares mean. The ALM and BMP means on symmetric cones [12, 13] are not effective and less interesting in computational aspects because they are inductively constructed and need to compute their means in each step via symmetrization procedures. Various numerical methods for the solution of the Karcher equation have been introduced in the literature: optimization algorithms like Newton's method or a gradient descent method, and iterative methods where the choice of an initial point close to the Karcher mean with geometric mean properties plays a key role [8].

References

- 1. M. Baes, Convexity and differentiability properties of spectral functions and spectral mappings on Euclidean Jordan algebras, *Linear Algebra Appl.*, **422** (2007), 664-700.
- 2. H. H. Bauschke, O. Güler, A. S. Lewis and H. S. Sendov, Hyperbolic polynomials and convex analysis, *Can. J. Math.*, **53** (2001), 470-488.
- 3. R. Bhatia and J. Holbrook, Riemannian geometry and matrix geometric means, *Linear Algebra Appl.*, **413** (2006), 594-618.
- 4. R. Bhatia and R. Karandikar, Monotonicity of the matrix geometric mean, *Math. Ann.*, **353** (2012), 1453-1467.
- 5. J. Faraut and A. Korányi, Analysis on Symmetric Cones, Clarendon Press, Oxford, 1994.
- 6. L. Faybusovich, Several Jordan-algebraic aspects of optimization, *Optimization* **57** (2008), 379-393.
- 7. J. Holbrook, No dice: a deterministic approach to the Cartan centroid, J. Ramanujan Math. Soc. 27 (2012), 509-521.
- 8. B. Jeuris, R. Vandebril and B. Vandereycken, A survey and comparison of contemporary algorithms for computing the matrix geometric mean, *Electronic Transactions on Numerical Analysis*, to appear.
- 9. H. Karcher, Riemannian center of mass and mollifier smoothing, *Comm. Pure Appl. Math.*, **30** (1977), 509-541.
- 10. S. Lang, *Fundamentals of Differential Geometry*, Graduate Texts in Mathematics, Springer, 1999.

- 11. J. D. Lawson and Y. Lim, The geometric mean, matrices, metrics, and more, *Amer. Math. Monthly*, **108** (2001), 797-812.
- 12. J. D. Lawson and Y. Lim, A general framework for extending means to higher orders, *Colloq. Math.*, **113** (2008), 191-221.
- 13. J. Lawson, H. Lee and Y. Lim, Weighted geometric means, Forum Math., 24 (2012), 1067-1090.
- 14. J. Lawson and Y. Lim, Monotonic properties of the least squares mean, *Math. Ann.*, **351** (2011), 267-279.
- 15. A. S. Lewis, The mathematics of eigenvalue optimization, *Math. Program., Ser. B*, **97** (2003), 155-176.
- 16. Y. Lim, Geometric means on symmetric cones, Arch. der Math., 75 (2000), 39-45.
- 17. Y. Lim, The Riemannian center of mass on symmetric cones of finite ranks, preprint.
- 18. Y. Lim, J. Kim and L. Faybusovich, Simultaneous diagonalization on simple Euclidean Jordan algebras and its applications, *Forum Math.*, **15** (2003), 639-644.
- 19. M. Moakher, A differential geometric approach to the geometric mean of symmetric positive-definite matrices, *SIAM J. Matrix Anal. Appl.*, **26** (2005), 735-747.
- D. Sun and J. Sun, Löwner's operator and spectral functions on Euclidean Jordan algebras, *Math. Oper. Res.*, 33 (2008), 421-445.
- K.-T. Sturm, Probability measures on metric spaces of nonpositive curvature, in: *Heat Kernels and Analysis on Manifolds, Graphs, and Metric Spaces*, (P. Auscher et. al. Eds.), Contemp. Math. 338, Amer, Math. Soc. (AMS), Providence, 2003.
- 22. J. Tao and M. S. Gowda, Some *P*-properties for nonlinear transformations on Euclidean Jordan algebras, *Math. Oper. Res.*, **30** (2005), 985-1004.

Sangho Kum Department of Mathematics Education Chungbuk National University Cheongju 361-763, Korea E-mail: shkum@cbnu.ac.kr

Hosoo Lee Department of Mathematics Kyungpook National University Taegu 702-701, Korea E-mail: thislake@knu.ac.kr

Yongdo Lim Department of Mathematics Sungkyunkwan University Suwon 440-746, Korea E-mail: ylim@skku.edu.kr