New deterministic approaches to the least square mean

Eduardo M. Ghiglioni * †
Depto. de Matemática, FCE-UNLP and IAM-CONICET, Argentina

Abstract

In this paper we presents new deterministic approximations to the least square mean, also called geometric mean or barycenter of a finite collection of positive definite matrices. Let $A_1, A_2, \ldots, A_m$ be any elements of $\mathcal{M}_d(\mathbb{C})^+$, where the set $\mathcal{M}_d(\mathbb{C})^+$ is the open cone in the real vector space of selfadjoint matrices $\mathcal{H}(n)$. We consider a sequence of blocks of m matrices, that is,

$$(A_1, \ldots, A_m, A_1, \ldots, A_m, A_1, \ldots A_m, \ldots).$$

We take a permutation on every block and then take the usual inductive mean of that new sequence. The main result of this work is that the inductive mean of this block permutation sequence approximate the least square mean on $\mathcal{M}_d(\mathbb{C})^+$. This generalizes a Theorem obtain by Holbrook. Even more, we have an estimate for the rate of convergence.

AMS subject classification: 15B48, 47A64, 53C20.
Keywords: Geometric mean, inductive mean, positive definite matrices, deterministic.

Contents

1 Introduction **1**
1.1 Setting of the problem 1
1.2 The problem 3
1.3 The main result of the paper 3

2 Proof of the main result **4**

3 Main theorem in Hadamard spaces **7**

1 Introduction

1.1 Setting of the problem

Let $\mathcal{M}_d(\mathbb{C})^+$ denote the set of positive definite matrices, which is an open convex cone in the real vector space of Hermitian matrices $\mathcal{H}(n)$. In particular, it inherits a differential structure where the tangent spaces can be identified with $\mathcal{H}(n)$. The manifold $\mathcal{M}_d(\mathbb{C})^+$ can be endowed with a natural Riemannian structure such that the natural action of the invertible matrices by congruence transformations $X \mapsto MXM^*$ becomes isometric. With respect to this metric structure, if $\alpha : [a, b] \to \mathcal{M}_d(\mathbb{C})^+$ is a piecewise smooth path, its length is defined by

$$L(\alpha) = \int_a^b \|\alpha^{-1/2}(t)\alpha'(t)\alpha^{-1/2}(t)\|_2 \, dt,$$

*Partially supported by CONICET (PIP 0150/14), FONCyT (PICT 1506/15) and FCE-UNLP (11X681), Argentina.
†e-mail addresses: eghiglioni@mate.unlp.edu.ar

where $\|\cdot\|_2$ denotes the Frobenius or Hilbert-Schmidt norm. As usual, a distance δ can be defined by

$$\delta(A,B) = \inf\{L(\alpha) : \ \alpha \text{ is a piecewise smooth path connecting } A \text{ with } B\}.$$

The infimum is actually a minimum, and the geodesic connecting two positive definite matrices A and B has the following simple expression

$$\gamma_{AB}(t) = A^{1/2}(A^{-1/2}BA^{-1/2})^t A^{1/2}.$$

It is usual in matrix analysis to use the notation $A\#_t B$ instead of $\gamma_{AB}(t)$. The midpoint $A\#_{\frac{1}{2}}B$ is called geometric mean or barycenter between A and B.

With the aforementioned Riemannian structure, $\mathcal{M}_d(\mathbb{C})^+$ becomes a Riemannian manifold with non-positive curvature. In particular, the distance function satisfies the so called semiparallelogram law.

Proposition 1.1. *Let $X, , Y$ be two matrices in $\mathcal{M}_d(\mathbb{C})^+$. Then for any Z in $\mathcal{M}_d(\mathbb{C})^+$*

$$\delta^2(X\#_{\frac{1}{2}}Y, Z) \leq \frac{1}{2}\delta^2(X,Z) + \frac{1}{2}\delta^2(Y,Z) - \frac{1}{4}\delta^2(X,Y). \tag{1}$$

Inductively for all dyadic rationals $t \in [0,1]$, and then by continuity we get the following inequality for other points in the geodesic:

$$\delta^2(X\#_t Y, Z) \leq (1-t)\delta^2(X,Z) + t\delta^2(Y,Z) - t(1-t)\delta^2(X,Y). \tag{2}$$

As a consequence of this inequality we get that the function

$$f(t) = \delta(A\#_t A^{'}, B\#_t B^{'})$$

is convex on $[0,1]$ for any $A, A', B, B' \in \mathcal{M}_d(\mathbb{C})^+$, more precisely for $t \in [0,1]$

$$\delta(A\#_t A^{'}, B\#_t B^{'}) \leq (1-t)\delta(A,B) + t\delta(A^{'}, B^{'}). \tag{3}$$

From the semiparallelogram law, we obtain the following alternative characterization of the geometric mean

$$A\#_{\frac{1}{2}}B = \underset{C\in\mathcal{M}_d(\mathbb{C})^+}{\operatorname{argmin}} \Big(\delta^2(A,C) + \delta^2(B,C)\Big).$$

There is no reason to restrict our attention to only two matrices. The notion of geometric mean can be generalized for more than two matrices in the obvious way

$$G(A_1,\ldots,A_m) := \underset{C\in\mathcal{M}_d(\mathbb{C})^+}{\operatorname{argmin}} \Big(\sum_{j=1}^{m}\delta^2(A_j,C)\Big).$$

The solution of this least square minimization problem exists and is unique because of the convexity properties of the distance $\delta(\cdot,\cdot)$. A multivariable version of the semiparallelogram law holds because of the nonpositive curvature property (2).

Proposition 1.2. *Let $A_1, A_2, \ldots, A_m$ be any elements of $\mathcal{M}_d(\mathbb{C})^+$ and let $G = G(A_1,\ldots,A_m)$. Then for all $Z \in \mathcal{M}_d(\mathbb{C})^+$ we have*

$$\delta^2(Z,G) \leq \frac{1}{m}\left(\sum_{j=1}^{m}\delta^2(Z,A_j) - \delta^2(G,A_j)\right). \tag{4}$$

It is also known as the variance inequality, since $G(A_1,\dots,A_m)$ can be interpreted as a nonlinear expectation of the probability measure

$$\mu = \frac{1}{m}\sum_{j=1}^{m}\delta_{\{A_j\}}.$$

The geometric means naturally appear in many applied problems. For instance, they appear in the study of radar signals (see [7] and the references therein for more details). Another typical application of the geometric means is in problems related with the gradient or Newton like optimization methods (see [6],[19]).

1.2 The problem

The usual problem dealing with geometric means is that the geometric mean of three or more matrices does not have in general a closed formula (see [5], [8], [12] and [17]). In [8], Holbrook proved that they can be approximated by the so called inductive means.

Definition. *Given a sequence of positive definite matrices $A = (A_n)_{n\in\mathbb{N}}$, the inductive means are defined as follows:*

$$\begin{aligned} S_1(A) &= A_1 \\ S_n(A) &= S_{n-1}(A)\#_{\frac{1}{n}} A_n \qquad (n \geq 2). \end{aligned}$$

Let $A_0,\dots,A_{d-1}$ be positive definite matrices, and define the cyclic sequence $A_{cyclic} = (A_n)_{n\in\mathbb{N}}$, where for $n \geq d$ we define

$$A_n = A_k \quad \text{if } n \equiv k \mod d.$$

Then, the main result in [8] is the following

Theorem 1.3 (Holbrook)**.**

$$\lim_{n\to\infty} S_n(A_{cyclic}) = G(A_0,\dots,A_{d-1}).$$

1.3 The main result of the paper

The main advantage of Holbrook's result is that it is deterministic. The results in [12] and [5] say that, if we consider the uniform distribution in $\{0,\dots,d-1\}$ and we construct a sequence

$$A_{(r)} = \big(A_{r(1)}, A_{r(2)}, A_{r(3)}, A_{r(4)},\dots\big)$$

taking randomly the matrices $A_0,\dots,A_{d-1}$, then almost surely

$$\lim_{n\to\infty} S_n(A_{(r)}) = G(A_0,\dots,A_{d-1}).$$

Although this implies that there are plenty of such a sequences, except for the cyclic one A_{cyclic}, we do not know if given a specific sequence A we have that

$$\lim_{n\to\infty} S_n(A) = G(A_0,\dots,A_{d-1}). \tag{5}$$

The main result of this paper allows to enlarge the set of deterministic examples of sequences A such that (5) holds. So, fix positive definite matrices $A_0,\dots,A_{m-1}$, and consider a sequence of permutations of m-elements $\sigma = \{\sigma_j\}_{j\in\mathbb{N}}$. Then, define the sequence

$$\boldsymbol{A}_\sigma = (A_{\sigma_1(0)},\dots,A_{\sigma_1(m-1)}, A_{\sigma_2(0)},\dots,A_{\sigma_2(m-1)}, A_{\sigma_3(0)},\dots,A_{\sigma_3(m-1)},\dots),$$

Using this notation, the main theorem of this paper is the following:

Theorem 1.4. *If $n \geq 1$ and $G = G(A_0, \ldots, A_{m-1})$, then there exists $L > 0$ depending on these matrices such that*

$$\delta^2(S_n(\boldsymbol{A}_\sigma), G) \leq \frac{L}{n}. \tag{6}$$

In particular

$$\lim_{n\to\infty} S_n(\boldsymbol{A}_\sigma) = G.$$

Holbrook's rate of convergence is $1/n$ as $n \to \infty$. Actually, as it was pointed out by Lim and Pálfia [16], one cannot expect better convergence rates than $1/n$. For that purpose consider the case when $m = 2$. So our rate of convergence should be able to be improved.

Finally, note that the above result is also true if we permute blocks of length km matrices for some $k \in \mathbb{N}$. More precisely,

Corollary 1.5. *Let $A_1, \ldots, A_m \in \mathcal{M}_d(\mathbb{C})^+$ be positive definite matrices. Given $k \geq 1$, consider the sequence*

$$\boldsymbol{A}_\sigma^{(k)} = (\underbrace{A_{\sigma_1(1)}, A_{\sigma_1(2)}, \ldots, A_{\sigma_1(m)}, \ldots, A_{\sigma_1(1)}, A_{\sigma_1(2)}, \ldots, A_{\sigma_1(m)}}_{km\ matrices}, A_{\sigma_2(1)}, \ldots)$$

where each σ is a permutation of km-elements. Then, there exists $L > 0$ such that

$$\delta^2(S_n(\boldsymbol{A}_\sigma^{(k)}), G(A_1, \ldots, A_m)) \leq \frac{L}{n}.$$

Indeed, note that by definition

$$G(\underbrace{A_1, \ldots, A_m, A_1, \ldots, A_m, \ldots, A_1, \ldots, A_m}_{k\ times}) = G(A_1, \ldots, A_m).$$

2 Proof of the main result

In this section we are dedicated to prove Theorem 1.4. We begin with some basic fact about the inductive mean which is a direct consequence of (3).

Lemma 2.1. *Given two sequence $A = (A_i)_{i\in\mathbb{N}}, B = (B_i)_{i\in\mathbb{N}}$ in $\mathcal{M}_d(\mathbb{C})^+$, then*

$$\delta(S_n(A), S_n(B)) \leq \frac{1}{n} \sum_{i=1}^{n} \delta(A_i, B_i). \tag{7}$$

Now we will follow the work of Lim and Pálfia [18]. This first lemma is actually step 1 in their paper.

Lemma 2.2. *Given a sequence $A = (A_i)_{i\in\mathbb{N}}$ in $\mathcal{M}_d(\mathbb{C})^+$ and $Z \in \mathcal{M}_d(\mathbb{C})^+$, for every $k, m \in \mathbb{N}$*

$$\begin{aligned}\delta^2(S_{k+m}(A), Z) \leq\ & \frac{k}{k+m}\, \delta^2(S_k(A), Z) + \frac{1}{k+m} \sum_{j=0}^{m-1} \delta^2(A_{k+j+1}, Z) \\ & - \frac{k}{(k+m)^2} \sum_{j=0}^{m-1} \delta^2(S_{k+j}(A), A_{k+j+1}).\end{aligned}$$

Proof. By the inequality (2) applied to $S_{n+1}(A) = S_n(A) \#_{\frac{1}{n+1}} A_{n+1}$ we obtain

$$(n+1)\ \delta^2(S_{n+1}(A), Z) - n\ \delta^2(S_n(A), Z) \leq \delta^2(A_{n+1}, Z) - \frac{n}{(n+1)}\delta^2(S_n(A), A_{n+1}).$$

Summing these inequalities from $n = k$ until $n = k+m-1$ we get that the difference

$$(k+m)\ \delta^2(S_{k+m}(A), Z) - k\ \delta^2(S_k(A), Z),$$

obtained from the telescopic sum of the left hand side, is less or equal than

$$\sum_{j=0}^{m-1} \left(\delta^2(A_{k+j+1}, Z) - \frac{k+j}{(k+j+1)}\delta^2(S_{k+j}(A), A_{k+j+1}) \right).$$

Finally, using that $\frac{k+j}{k+j+1} \geq \frac{k}{k+m}$ for every $j \in \{0, \ldots, m-1\}$, this sum is bounded from the above by

$$\sum_{j=0}^{m-1} \left(\delta^2(A_{k+j+1}, Z) - \frac{k}{(k+m)}\delta^2(S_{k+j}(A), A_{k+j+1}) \right),$$

which completes the proof. □

Using the previous result we have this particular case:

Lemma 2.3. *Let $A_1, \ldots, A_m \in \mathcal{M}_d(\mathbb{C})^+$ fixed. For every sequence $\boldsymbol{A}_\sigma$,*

$$\delta^2(S_{(k+1)m}(\boldsymbol{A}_\sigma), G) \leq \frac{k}{k+1}\delta^2(S_{km}(\boldsymbol{A}_\sigma), G) + \frac{1}{k+1}\left(\frac{1}{m}\sum_{j=0}^{m-1} \delta^2((\boldsymbol{A}_\sigma)_{km+j+1}, G) \right) -$$

$$- \frac{k}{(k+1)^2}\left(\frac{1}{m}\sum_{j=0}^{m-1} \delta^2(S_{km+j}(\boldsymbol{A}_\sigma), (\boldsymbol{A}_\sigma)_{km+j+1}) \right).$$

Proof. Just change k with km, Z with G and A with $\mathbf{A}_\sigma$ in Lemma 2.2.

□

The next step is to find a lower bound for

$$\frac{1}{m}\sum_{j=0}^{m-1} \delta^2(S_{km+j}(\mathbf{A}_\sigma), (\mathbf{A}_\sigma)_{km+j+1}).$$

This is step 2 in [18] and here it is a little different.

Let $A_1, \ldots, A_m \in \mathcal{M}_d(\mathbb{C})^+$ fixed, we will denote

$$\Delta := \max_{1\leq i,j\leq m} \delta(A_i, A_j); \qquad \alpha := \frac{1}{m}\sum_{i=1}^{m} \delta^2(G, A_i).$$

Note that by (2), for every sequence $\mathbf{A}_\sigma$, all $k \in \mathbb{N}$ and all $j \in \mathbb{N}$,

$$\delta(S_{km}(\boldsymbol{A}_\sigma), (\boldsymbol{A}_\sigma)_j) \leq \Delta.$$

Lemma 2.4. *Let $A_1, \ldots, A_m \in \mathcal{M}_d(\mathbb{C})^+$ fixed. For every sequence $\boldsymbol{A}_\sigma$ and for all $k \in \mathbb{N}$ we have*

$$\frac{1}{m}\sum_{j=0}^{m-1} \delta^2(S_{km+j}(\boldsymbol{A}_\sigma), (\boldsymbol{A}_\sigma)_{km+j+1}) \geq \delta^2(S_{km}(\boldsymbol{A}_\sigma), G) + \alpha - \left(\frac{m^2}{(km+1)^2} + 2\frac{m}{km+1} \right)\Delta^2.$$

Proof. Let $0 \leq j \leq m-1$. Note that by (2) and all k,

$$\delta(S_{km+j}(\boldsymbol{A}_\sigma), S_{km+j+1}(\boldsymbol{A}_\sigma)) \leq \frac{\Delta}{km+j+1}.$$

Hence

$$\begin{aligned}
\delta(S_{km}(\boldsymbol{A}_\sigma), (\mathbf{A}_\sigma)_{km+j+1}) &\leq \delta(S_{km}(\boldsymbol{A}_\sigma), S_{km+j}(\boldsymbol{A}_\sigma)) + \delta(S_{km+j}(\boldsymbol{A}_\sigma), (\mathbf{A}_\sigma)_{km+j+1}) \\
&\leq \sum_{h=1}^{j} \frac{\Delta}{km+h} + \delta(S_{km+j}(\boldsymbol{A}_\sigma), (\mathbf{A}_\sigma)_{km+j+1}) \\
&\leq \frac{m}{km+1}\Delta + \delta(S_{km+j}(\boldsymbol{A}_\sigma), (\mathbf{A}_\sigma)_{km+j+1}).
\end{aligned}$$

Therefore, for every $j \leq m$,

$$\delta^2(S_{km}(\boldsymbol{A}_\sigma), (\mathbf{A}_\sigma)_{km+j+1}) \leq \left(\frac{m^2}{(km+1)^2} + 2\frac{m}{km+1}\right)\Delta^2 + \delta^2(S_{km+j}(\boldsymbol{A}_\sigma), (\mathbf{A}_\sigma)_{km+j+1})$$

where we have used that $\delta(S_{km+j}(\boldsymbol{A}_\sigma), (\mathbf{A}_\sigma)_{km+j+1}) \leq \Delta$ for every $k, j \in \mathbb{N}$. Summing up these inequalities and dividing by m, we get

$$\begin{aligned}
\frac{1}{m}\sum_{j=0}^{m-1} \delta^2(S_{km}(\boldsymbol{A}_\sigma), (\mathbf{A}_\sigma)_{km+j+1}) \leq \left(\frac{m^2}{(km+1)^2} + 2\frac{m}{km+1}\right)\Delta^2 + \\
+ \frac{1}{m}\sum_{j=0}^{m-1} \delta^2(S_{km+j}(\boldsymbol{A}_\sigma), (\mathbf{A}_\sigma)_{km+j+1}).
\end{aligned} \tag{8}$$

By the variance inequality (4),

$$\delta^2(S_{km}(\boldsymbol{A}_\sigma), G) \leq \frac{1}{m}\sum_{i=1}^{m} \delta^2(S_{km}(\boldsymbol{A}_\sigma), (\mathbf{A}_\sigma)_i) - \alpha. \tag{9}$$

Note that

$$\frac{1}{m}\sum_{j=0}^{m-1} \delta^2(S_{km}(\boldsymbol{A}_\sigma), (\mathbf{A}_\sigma)_{km+j+1}) = \frac{1}{m}\sum_{i=1}^{m} \delta^2(S_{km}(\boldsymbol{A}_\sigma), (\mathbf{A}_\sigma)_i). \tag{10}$$

So, combining (8) and (9) we get the desired result. □

Remark 2.5. The equality (10) follows from the fact that the matrices that appear in the first block $A_{\sigma_1(1)}, \ldots, A_{\sigma_1(m)}$ are the same as those that appear in the $(k+1)$-th block (except by the order). On the other hand, note that this result can not be extended to weighted means as those consider in [18]. Indeed, in that setting, the weights do not allow to consider permutations.

Now we prove Theorem 1.4 and the rate of convergence for a special subsequence.

Lemma 2.6. *Let $A_1, \ldots, A_m \in \mathcal{M}_d(\mathbb{C})^+$ fixed. For every sequence $\boldsymbol{A}_\sigma$ and for all $k \in \mathbb{N}$,*

$$\delta^2(S_{km}(\boldsymbol{A}_\sigma), G) \leq \frac{L}{k},$$

where $L = \alpha + 3\Delta^2$.

Proof. We will prove it by induction. If $k = 1$ the result is trivial because

$$\delta^2(S_m(\boldsymbol{A}_\sigma), G) \leq \Delta^2 \leq L.$$

Let's suppose that it's true for k. Combining Lemmas 2.3 and 2.4 we get

$$\begin{aligned}
\delta^2(S_{(k+1)m}(\boldsymbol{A}_\sigma), G) &\leq \frac{k}{k+1}\delta^2(S_{km}(\boldsymbol{A}_\sigma), G) + \frac{1}{k+1}\alpha \\
&\quad - \frac{k}{(k+1)^2}\left[\delta^2(S_{km}(\boldsymbol{A}_\sigma), G) + \alpha - \left(\frac{m^2}{(km+1)^2} + 2\frac{m}{km+1}\right)\Delta^2\right] \\
&\leq \frac{k^2}{(k+1)^2}\delta^2(S_{km}(\boldsymbol{A}_\sigma), G) + \frac{1}{(k+1)^2}\alpha + 3\frac{1}{(k+1)^2}\Delta^2 \\
&\leq \frac{k}{(k+1)^2}L + \frac{1}{(k+1)^2}L \\
&\leq \frac{L}{(k+1)}.
\end{aligned}$$

□

To conclude the proof of Theorem 1.4. let $n = km + d$ such that $d \in \{1, \ldots, m-1\}$, $k \in \mathbb{N}$. Since $X \#_t X = X$ for all $X \in \mathcal{M}_d(\mathbb{C})^+$ and all $t \in [0,1]$, using Lemma 2.1 with the sequences

$$(\ (\mathbf{A}_\sigma)_1, \ldots, (\mathbf{A}_\sigma)_{km}, \underbrace{S_{km}(\boldsymbol{A}_\sigma), \ldots, S_{km}(\boldsymbol{A}_\sigma)}_{\text{d times}}\ )$$

and

$$(\ (\mathbf{A}_\sigma)_1, \ldots, (\mathbf{A}_\sigma)_{km},\ (\mathbf{A}_\sigma)_{km+1}\ ,\ \ldots\ ,\ (\mathbf{A}_\sigma)_{km+d}\ ),$$

and taking into account that $\delta(S_{km}(\boldsymbol{A}_\sigma), (\mathbf{A}_\sigma)_{km+j}) \leq \Delta$ for every $j \in \{1, \ldots, d\}$, we get

$$\begin{aligned}
\delta(S_{km}(\boldsymbol{A}_\sigma), S_n(\boldsymbol{A}_\sigma)) &\leq \frac{1}{km+d}\sum_{j=1}^{d}\delta(S_{km}(\boldsymbol{A}_\sigma), (\mathbf{A}_\sigma)_{km+j}) \\
&\leq \frac{d}{km+d}\Delta \leq \frac{1}{k}\Delta \xrightarrow[k\to\infty]{} 0.
\end{aligned}$$

Combining this with Lemma 2.6 we obtain that for n big enough $\delta(S_n(\boldsymbol{A}_\sigma), G) < \varepsilon$. □

3 Main theorem in Hadamard spaces

In this section we will mention how all the previous result can be generalized to a much more general context. If we look in detail all the previous proofs, we can note that they are based mainly on the semiparallelogram law and the variance inequality (and their consequences). So, with the necessary chances, we can generalized the main theorem to non-positively curved (NPC) spaces, also called Hadamard spaces or (global) CAT(0) spaces. A complete metric space (M, δ) is called a Hadamard space if it satisfies the semiparallelogram law, i.e., for each $x, y \in M$ there exists $m \in M$ satisfying

$$\delta^2(m, z) \leq \frac{1}{2}\delta^2(x, z) + \frac{1}{2}\delta^2(y, z) - \frac{1}{4}\delta^2(x, y) \tag{11}$$

for all $z \in M$. The point m is called (metric) midpoint between x and y. Taking $z = x$ and $z = y$ in the inequality (11), it is easy to conclude that $\delta(x, m) = \delta(m, y) = \frac{1}{2}\delta(x, y)$. Moreover, this inequality also implies that the midpoint is unique. The existence and uniqueness of midpoints give rise to a unique (metric) geodesic $\gamma_{a,b} : [0,1] \to M$ connecting any given two points a and b, that we denote as before as $a\#_t b$. The inequality (11) also extends to arbitrary points on geodesics.

$$\delta^2(x\#_t y, z) \leq (1-t)\delta^2(x,z) + t\delta^2(y,z) - t(1-t)\delta^2(x,y). \tag{12}$$

As in the case of positive definite matrices, the inductive mean is defined in general Hadamard spaces in the same way (using the geodesic mention before). Also, the notion of barycenter and all the results that we use related to the barycenter (Existence and uniqueness - Variance Inequality) can be extend to this spaces. We refer the reader to [20], [21].

Acknowledgements:

The author wish to express his gratitude to Professor Jorge A. Antezana and Professor Demetrio Stojanoff for their valuable comments and suggestions. The author also thanks the anonymous referees for their careful reading of the paper and helpful comments.

References

[1] T. Ando, C. K. Li, R. Mathias, Geometric means, Linear Algebra Appl. 385 (2004), 305-334.

[2] J. Antezana, E. Ghiglioni and D. Stojanoff, Ergodic theorem in Hadamard spaces in terms of inductive means, preprint (2018), arXiv:1808.02060, 22 pages.

[3] M. Berger, A Panoramic View of Riemannian Geometry, Springer, (2003).

[4] R. Bhatia, Positive Definite Matrices, Princeton Series in Applied Mathematics, Princeton University Press, (2007).

[5] R. Bhatia, R. Karandikar, Monotonicity of the matrix geometric mean. Math. Ann. 353(4)(2012) 1453-1467.

[6] D. Bini, B. Iannazzo, Computing the Karcher mean of symmetric positive definite matrices, Linear Algebra Appl. 438 (2013) 1700-1710.

[7] F. Barbaresco, Interactions between symmetric cone and information geometries: Bruhat-Tits and Siegel spaces models for higher resolution autoregressive Doppler imagery, Emerging Trends in Visual Computing, Lecture Notes in Computer Science 5416 (2009) 124-163.

[8] J. Holbrook, No dice: a deterministic approach to the Cartan centroid, J. Ramanujan Math. Soc. 27 (2012) 509-521.

[9] S. Lang, Fundamentals of Differential Geometry, Graduate Texts in Mathematics, Springer, (1999).

[10] J. Lawson, Y. Lim, Contractive barycentric maps, J. Operator Theory 77 (2017) 87-107.

[11] J. Lawson, Y. Lim, Karcher means and Karcher equations of positive definite operators, Trans. Amer. Math. Soc. Ser. B1 (2014) 1-22.

[12] J. Lawson, Y. Lim, Monotonic properties of the least squares mean, Math. Ann. 351 (2011) 267-279.

[13] J. Lawson, Y. Lim, The geometric mean, matrices, metrics, and more, Amer. Math. Monthly 108 (2001).

[14] J. Lawson, Y. Lim, Weighted means and Karcher equations of positive operators, Proc. Natl. Acad. Sci. USA 110 (2013) 15626-15632.

[15] Y. Lim, Riemannian distances between Geometric means. SIAM J. Matrix Anal. Appl. 34 (2013), 932-945.

[16] Y. Lim, M. Pálfia, Approximations to the Karcher mean on Hadamard spaces via geometric power means, Forum Math. 27 (2015) 2609-2635.

[17] Y. Lim, M. Pálfia, Matrix power mean and the Karcher mean, J. Funct. Anal. 262 (2012) 1498-1514.

[18] Y. Lim, M. Pálfia, Weighted deterministic walks and no dice approach for the least squares mean on Hadamard spaces, Bull. Lond. Math. Soc. 46 (2014) 561-570.

[19] M. Moakher, M. Zerai, The Riemannian geometry of the space of positive-definite matrices and its application to the regularization of positive-definite matrix-valued data, J. Math. Imaging Vision 40 (2011) 171-187.

[20] M. Pálfia, Means in metric spaces and the center of mass, J. Math. Anal. Appl. 381 (2011) 383-391.

[21] K. T. Sturm, Probability measures on metric spaces of nonpositive curvature, in: P. Auscher, et al. (Eds.), Heat Kernels and Analysis on Manifolds, Graphs, and Metric Spaces, in: Contemp. Math., vol. 338, Amer. Math. Soc. (AMS), Providence, (2003).