Following the Trail of the Operator Geometric Mean

Jimmie D. Lawson and Yongdo Lim

Abstract This paper traces the development of the theory of the matrix geometric mean in the cone of positive definite matrices and the closely related operator geometric mean in the positive cone of a unital C^*-algebra. The story begins with the two-variable matrix geometric mean, moves to the n-variable matrix setting, then to the extension to the positive cone of the C^*-algebra of operators on a Hilbert space, and even to general unital C^*-algebras, and finally to the consideration of barycentric maps on the space of integrable probability measures on the positive cone. Besides expected tools from linear algebra and operator theory, one observes a substantial interplay with operator monotone functions, geometrical notions in metric spaces, particularly the notion of nonpositive curvature, some probabilistic theory of random variables with values in a metric space of nonpositive curvature, and the appearance of related means such as the inductive and power means.

Keywords Geometric mean · Operator mean · Operator monotone function · Nonpositively curved metric spaces · Contractive barycentric map

Mathematics Subject Classification (2010) Primary 47A64; Secondary 46L05, 47B65, 47L07

1 Introduction

Positive definite matrices have become fundamental computational objects in many areas of engineering, statistics, quantum information, and applied mathematics.

J. D. Lawson (✉)
Department of Mathematics, Louisiana State University, Baton Rouge, LA, USA
e-mail: lawson@math.lsu.edu

Y. Lim
Department of Mathematics, Kyungpook National University, Taegu, Korea
e-mail: ylim@knu.ac.kr

P. Kielanowski et al. (eds.), *Geometric Methods in Physics XXXVIII*, Trends in Mathematics, https://doi.org/10.1007/978-3-030-53305-2_10

They appear as "data points" in a diverse variety of settings: covariance matrices in statistics, elements of the search space in convex and semidefinite programming, kernels in machine learning, observations in radar imaging, and diffusion tensors in medical imaging, to cite only a few. A variety of computational algorithms have arisen for approximation, interpolation, filtering, estimation, and averaging. Our interest focuses on the last named, the process of finding an average or mean, which is again positive definite.

A simple computation would be to take the arithmetic mean of a given finite set of positive definite matrices. However, researchers have learned that to find a mean or average that performs well and exhibits desirable properties, one needs to take into account the underlying geometric structure of Ω_n, the space of $n \times n$-positive definite matrices.

Formally a *mean of order n*, or *n-mean* for short, on a set X is a function $\mu : X^n \to X$ satisfying the idempotency condition $\forall x \in X,\ \mu(x, x, \dots, x) = x$. It is frequently assumed in the definition of a mean that a mean is invariant under any permutation of variables; we call these *symmetric means*. The mean $\mu : X^n \to X$ is *continuous* or a *topological mean* if X is a topological space and μ is continuous. Typically a mean represents some type of averaging operator.

The subject of (binary) means for positive numbers or line segments has a rich mathematical lineage dating back into antiquity. The Greeks, motivated by their interest in proportions and musical ratios, defined at least eleven different means (depending on how one counts), the arithmetic, geometric, harmonic, and golden being the best known. A geometric construction for the geometric mean $\sqrt{ab}$ of $a, b > 0$ is given by Euclid in Book II in the form of "squaring the rectangle," i.e., constructing a square of the same area as a given rectangle of sides a and b. The study of various means and their properties on the positive reals has remained an active area of investigation up to the present day.

2 Positive Definite Matrices

Let $\mathcal{M}_n(\mathbb{C})$, or simply $\mathcal{M}_n$, denote the set of $n \times n$ complex matrices. We may identify $\mathcal{M}_n$ with the set of linear operators on $\mathbb{C}^n$, where we consider $\mathbb{C}^n$ to be a complex Hilbert space of column vectors with the usual Hermitian inner product. Denoting the conjugate transpose of $A \in \mathcal{M}_n$ by A^*, we recall that A is *Hermitian* if $A = A^*$ and *unitary* if $A^* = A^{-1}$. The Hermitian matrix A is *positive definite* if $\forall u \neq 0,\ \langle u, Au\rangle > 0$. These notions readily generalize to $\mathcal{B}(H)$, the algebra of operators on an arbitrary Hilbert space.

The following are well-known equivalences for a Hermitian matrix A to be positive definite:

1. $\langle Ax, x\rangle > 0$ for all $0 \neq x$, where $\langle \cdot, \cdot \rangle$ is the Hilbert space inner product on $\mathbb{C}^n$.
2. $A = BB^*$ for some invertible B.
3. A has all positive eigenvalues.

4. $A = \exp B = \sum_{n=0}^{\infty} B^n/n!$ for some (unique) Hermitian B.
5. $A = UDU^*$ for some unitary U and diagonal D with positive diagonal entries.

The positive definite $n \times n$-matrices form an open cone in $\mathbb{H}_n$, the $n \times n$ Hermitian matrices, with closure the positive semidefinite matrices (equivalently, $\langle Ax, x\rangle \geq 0$ for all x). We denote the open cone of positive definite matrices by Ω (or Ω_n if we need to distinguish the dimension).

We define a partial order (sometimes called the *Loewner order*) on the vector space $\mathbb{H}_n$ of Hermitian matrices by $A \leq B$ if $B - A$ is positive semidefinite. We note $0 \leq A$ iff A is positive semidefinite and write $0 < A$ if $A \in \Omega$ iff A is positive definite. The matrix A is sometimes called *strictly positive* in this setting.

Every positive definite (Hermitian) matrix operator has a unique *spectral decomposition*

$$A = \sum_{i=1}^{n} \lambda_i E_i,$$

where $\lambda_i > 0$ ($\lambda_i \in \mathbb{R}$) is the i^{th}-eigenvalue and E_i is the orthogonal projection onto the eigenspace of λ_i. One then has

$$A^k = \sum_{i=1}^{n} \lambda_i^k E_i,$$

from which one can easily deduce that every positive definite matrix has a unique positive definite kth-root.

The arithmetic and harmonic means readily extend from $\mathbb{R}^{>0}$ to the set of positive definite matrices:

$$\mathcal{A}(A, B) = \frac{1}{2}(A + B); \qquad \mathcal{H}(A, B) = 2(A^{-1} + B^{-1})^{-1}.$$

The geometric mean is not so obvious (e.g., $\sqrt{AB}$ need not be positive definite for A, B positive definite). One approach is to rewrite the equation $x^2 = ab$ (which has positive solution the geometric mean of a and b) in its appropriate form in the noncommutative setting:

$$\begin{aligned} XA^{-1}X &= B \\ A^{-1/2}XA^{-1/2}A^{-1/2}XA^{-1/2} &= A^{-1/2}BA^{-1/2} \\ A^{-1/2}XA^{-1/2} &= (A^{-1/2}BA^{-1/2})^{1/2} \\ X &= A^{1/2}(A^{-1/2}BA^{-1/2})^{1/2}A^{1/2}. \end{aligned}$$

We write $A\#B(= A^{1/2}(A^{-1/2}BA^{-1/2})^{1/2}A^{1/2})$ for the matrix geometric mean. Other connections between the matrix geometric mean and the one for positive real numbers may be found in [8].

3 Operator Monotone Functions and the Kubo–Ando Theorem

For $M_1, M_2 \subseteq \mathbb{R}$ and a mapping $f : M_1 \to M_2$, we define a function on the set of all Hermitian A with spectrum contained in M_1 by $f(A) = \sum_{i=1}^{n} f(\lambda_i)E_i$, where $A = \sum_{i=1}^{n} \lambda_i E_i$ is the spectral decomposition (functions constructed in this way are called *primary matrix functions* and provide a simple example of the functional calculus). A continuous function $f : M_1 \to M_2$ is *operator monotone* if $f(A) \le f(B)$ whenever $A \le B$. Operator monotone functions defined on some interval are continuous, monotone (nondecreasing), and concave.

If $M\mu(A, B)M^* = \mu(MAM^*, MBM^*)$ for all invertible M, the mean $\mu : \Omega \times \Omega \to \Omega$ is said to be *invariant under congruence transformations*. The mean μ is *monotonic* if $A_1 \le A_2,\ B_1 \le B_2$ implies $\mu(A_1, B_1) \le \mu(A_2, B_2)$. The next result is a major 1980 result of F. Kubo and T. Ando [6].

Theorem *Every operator monotone function $f : \mathbb{R}^{>0} \to \mathbb{R}^{>0}$ with $f(1) = 1$ gives rise to a congruence-invariant, monotonic mean μ defined by*

$$\mu(A, B)[= A^{\frac{1}{2}}\mu(I, A^{-1/2}BA^{-1/2})A^{\frac{1}{2}}] = A^{\frac{1}{2}} f(A^{-1/2}BA^{-1/2})A^{\frac{1}{2}}.$$

The association $f \to \mu_f$ is a bijection between the operator monotone functions and the congruence-invariant, monotonic continuous means. (*For the converse, one defines f from $\mu(I, \lambda I) = f(\lambda)I$.)*

To illustrate we apply the Kubo–Ando roadmap for passing from numeric to matrix means for certain important examples:

1. The Geometric Mean $A\#B$ and Weighted Geometric Mean $A\#_t B$:

$$\gamma(a, b) = \sqrt{ab} \to f(x) = \gamma(1, x) = x^{1/2}$$
$$\to \mathcal{G}(A, B) = A\#B = A^{1/2}(A^{-1/2}BA^{-1/2})^{1/2}A^{1/2}$$
$$\gamma_t(a, b) = a^{1-t}b^t \to \mathcal{G}_t(A, B) = A\#_t B = A^{1/2}(A^{-1/2}BA^{-1/2})^t A^{1/2}$$

2. The Arithmetic Mean:

$$\alpha(a, b) = (a + b)/2 \quad \to \quad f(x) = \alpha(1, x) = (1/2)(1 + x)$$
$$\to \mathcal{A}(A, B) = A^{1/2}((1/2)(I + A^{-1/2}BA^{-1/2}))A^{1/2}$$
$$= (1/2)(A + B)$$

The Kubo–Ando Theorem provided the foundation for the rapid development of the theory of matrix and operator means of two variables. However, no such analogous theorem has been discovered for multivariable means, even when the extension is known for the case of the positive reals. Some means, such as the arithmetic and harmonic, admit rather obvious extensions to the n-variable matrix case. But the problem of extending the geometric mean to the multivariable matrix setting remained unsolved for a number of years. An important step along the way was the gradual realization that the geometric matrix mean of two variables had an important alternative geometric/metric characterization, apparently first appearing in print in an article of the authors in 2001 [8].

4 Means of Several Variables and NPC-Spaces

In 2004 Ando et al. [2] gave the first extension of the binary geometric mean to n-variables, which came to be called the ALM mean. They listed desirable axiomatic properties for such an n-variable extension g and showed they were satisfied by their extension. The proofs typically involved extending from the known case of $n = 2$ by induction.

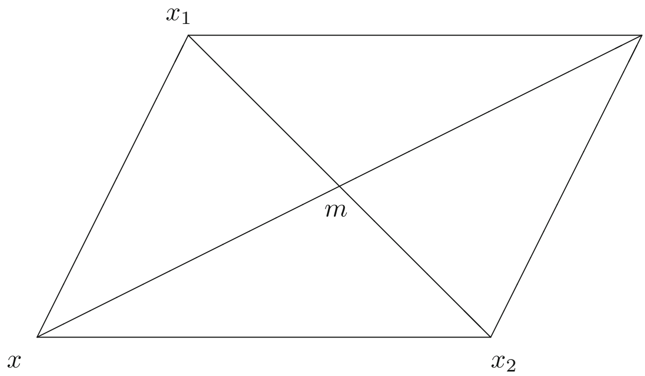
Let $\mathbb{A} = (A_1, \ldots, A_n), \mathbb{B} = (B_1, \ldots, B_n) \in \Omega^n$.

- (P1) (Consistency with scalars) $g(\mathbb{A}) = (A_1 \cdots A_n)^{1/n}$ if the A_i's commute;
- (P2) (Joint homogeneity) $g(a_1 A_1, \ldots, a_n A_n) = (a_1 \cdots a_n)^{1/n} g(\mathbb{A})$;
- (P3) (Permutation invariance) $g(\mathbb{A}_\sigma) = g(\mathbb{A})$, where $\mathbb{A}_\sigma = (A_{\sigma(1)}, \ldots, A_{\sigma(n)})$;
- (P4) (Monotonicity) If $B_i \le A_i$ for all $1 \le i \le n$, then $g(\mathbb{B}) \le g(\mathbb{A})$;
- (P5) (Continuity) g is continuous;
- (P6) (Congruence invariance) $g(M\mathbb{A}M^*) = Mg(\mathbb{A})M^*$ for M invertible, where $M(A_1, \ldots, A_n)M^* = (MA_1M^*, \ldots, MA_nM^*)$;
- (P7) (Joint concavity) $g(\lambda\mathbb{A} + (1-\lambda)\mathbb{B}) \ge \lambda g(\mathbb{A}) + (1-\lambda)g(\mathbb{B})$ for $0 \le \lambda \le 1$;
- (P8) (Self-duality) $g(A_1^{-1}, \ldots, A_n^{-1})^{-1} = g(A_1, \ldots, A_n)$;
- (P9) (Determinantal identity) $\operatorname{Det} g(\mathbb{A}) = \prod_{i=1}^n (\operatorname{Det} A_i)^{1/n}$; and
- (P10) (AGH mean inequalities) $n(\sum_{i=1}^n A_i^{-1})^{-1} \le g(\mathbb{A}) \le \frac{1}{n}\sum_{i=1}^n A_i$.

But a better candidate soon appeared. To understand it, we need some background. The parallelogram law in Hilbert spaces is given by

$$\text{sum of 2 diagonals squared} = \text{sum of 4 sides squared}$$

$$d^2(x_1, x_2) + 4d^2(x, m)(= (2d(x, m))^2) = 2d^2(x, x_1) + 2d^2(x, x_2)$$



Replacing the equality by an inequality in a general metric space yields the more general *semiparallelogram law*: for all $x_1, x_2 \in X$, there exists $m \in X$ such that for any $x \in X$,

$$d^2(x_1, x_2) + 4d^2(x, m) \le 2d^2(x, x_1) + 2d^2(x, x_2) \qquad \text{(NPC)}$$

One can show that $m = m(x_1, x_2)$ is unique and is the unique metric midpoint between x_1 and x_2.

(Global) NPC-spaces are complete metric spaces satisfying the semiparallelogram law (NPC). They have been intensely studied in recent years, often under alternative names such as CAT(0)-spaces or Hadamard spaces.

Condition (NPC) is a metric version of *NonPositive Curvature*, since the distance metric of a simply connected Riemannian manifold satisfies (NPC) iff the Riemannian metric has nonpositive curvature in the usual sense.

Example The open cone Ω_n of $n \times n$ positive definite matrices becomes a Riemannian manifold when equipped with the trace Riemannian metric: $\langle X, Y \rangle_A = \mathrm{tr} A^{-1} X A^{-1} Y$, where $A \in \Omega_n$ and X, Y are $n \times n$ Hermitian matrices. The corresponding distance metric on Ω_n is given by $d(A, B) = \| \log(A^{-1/2} B A^{-1/2} \|_2$, where $\| \cdot \|_2$ is the Frobenius (or Hilbert-Schmidt) norm. The cone Ω_n equipped with the metric d is an NPC-space. Furthermore, the unique midpoint between $A, B \in \Omega_n$ is the geometric mean $A \# B$.

Let (M, d) be a metric space. Given a *weight* $\mathbf{w} = (w_1, \ldots, w_n)$ (each $w_i \ge 0$ and $\sum_{i=1}^n w_i = 1$), the *weighted least squares mean* $\Lambda(\mathbf{w}; a_1, \ldots, a_n)$ of $(a_1, \ldots, a_n) \in M^n$ is defined as the solution to the optimization problem of minimizing the weighted sum of distances squared:

$$\operatorname*{argmin}_{x \in M} \sum_{i=1}^{n} w_i \delta^2(x, a_i),$$

provided the solution uniquely exists. This is the case for NPC spaces, since the function defined by $x \mapsto \sum_{i=1}^{n} w_i \delta^2(x, a_i)$ is uniformly convex in this case. E. Cartan considered such "barycenters" in the case of Riemannian manifolds, where they uniquely exist for the ones of nonpositive curvature, and M. Fréchet considered them in more general metric spaces. Thus the least squares mean is also called the Cartan mean or Fréchet mean.

First M. Moakher [13] and independently R. Bhatia and J. Holbrook [3, 4] studied the uniformly weighted least squares mean for the set of positive definite matrices Ω equipped with the trace metric as a multivariable generalization of the two-variable geometric mean. They established its (unique) existence and verified several of the axiomatic properties (P1)–(P10) satisfied by the Ando–Li–Mathias geometric mean: consistency with scalars, joint homogeneity, permutation invariance, congruence invariance, and self-duality (the last two being true since congruence transformations and inversion are isometries). Further, based on computational experimentation, Bhatia and Holbrook conjectured monotonicity for the least squares mean, but this was left as an open problem.

5 Monotonicity, Probability, and the Inductive Mean

One other mean will play an important role in what follows, one that we shall call the *inductive mean*, following the terminology of K.-T. Sturm [15]. It appeared elsewhere in the work of M. Sagae and K. Tanabe [14] and Ahn et al. [1]. It is defined inductively for NPC spaces (or more generally for metric spaces with weighted binary means $x\#_t y$) for each $k \geq 2$ by $S_2(x, y) = x\#y$ and for $k \geq 3$, $S_k(x_1, \ldots, x_k) = S_{k-1}(x_1, \ldots, x_{k-1})\#_{\frac{1}{k}} x_k$. (Here $x\#_t y$ is the unique point z such that $d(x, y) = (1-t)d(x, z) + td(y, z)$ for $0 \leq t \leq 1$.) Note that this mean at each stage is defined from the previous stage by taking the appropriate two-variable weighted mean, which is monotone. Thus the inductive mean is monotone.

Let (X, d) be an NPC metric space, $\{x_1, \ldots, x_m\} \subseteq X$. Set $\mathbb{N}_m = \{1, 2, \ldots, m\}$ and assign to $k \in \mathbb{N}_m$ the probability w_k, where $0 \leq w_k \leq 1$ and $\sum_{k=1}^{m} w_i = 1$. For each $\omega \in \prod_{n=1}^{\infty} \mathbb{N}_m$, define a sequence $\sigma = \sigma_\omega$ in X by $\sigma(1) = x_{\omega(1)}$, $\sigma(k) = S_k(x_{\omega(1)}, \ldots, x_{\omega(k)})$, where S_k is the inductive mean. (The sequence σ_ω may be viewed as a "walk" starting at $\sigma(1) = x_{\omega(1)}$ and obtaining $\sigma(k)$ by moving from $\sigma(k-1)$ toward $x_{\omega(k)}$ a distance of $(1/k)d(\sigma(k-1), x_{\omega(k)})$.) The following is a special case of Sturm's main results in [15].

Theorem 1 (Sturm's Theorem) *Giving $\prod_{n=1}^{\infty} \mathbb{N}_m$ the product probability, the set*

$$\{\omega \in \prod_{n=1}^{\infty} \mathbb{N}_m : \lim_n \sigma_\omega(n) = \Lambda(\mathbf{w}; x_1, \ldots, x_m)\}$$

has measure 1, i.e., $\sigma_\omega(n) \to \Lambda(\mathbf{w}; x_1, \ldots, x_m)$ for almost all ω.

More generally, Sturm establishes a version of the Strong Law of Large Numbers for random variables into an NPC metric space, with limit the least squares mean.

Using Sturm's Theorem, the authors were able to show (2011) [8]:

Theorem 2 *Let Ω be the open cone of positive definite matrices of some fixed dimension n.*

(1) *The least squares mean Λ on Ω is monotone: $A_i \leq B_i$ for $1 \leq i \leq n$ implies $\Lambda(A_1, \ldots, A_n) \leq \Lambda(B_1, \ldots B_n)$.*
(2) *All ten of the ALM axioms hold for Λ.*
(3) *In a natural way Λ can be extended to a weighted mean, and appropriate weighted versions of the ten properties hold.*

Note The ALM mean is typically distinct from the least squares mean for $n \geq 3$. Thus the ALM axioms do not characterize a mean. The latter fact had already been noted by Bini et al. [5], who introduced a much more computationally efficient variant of the ALM mean [5].

6 The Karcher Equation

The uniform convexity of the trace metric d on Ω yields that the least squares mean is the unique critical point for the function $X \mapsto \sum_{k=1}^{n} d^2(X, A_k)$. The least squares mean is thus characterized by the vanishing of the gradient, which is equivalent to its being a solution of the following *Karcher equation*:

$$\sum_{i=1}^{n} w_i \log(X^{-1/2} A_i X^{-1/2}) = 0. \tag{1}$$

The Karcher equation (1) can be used to *define* a mean on the cone Ω of positive invertible bounded operators on an infinite-dimensional Hilbert space (where one no longer has an NPC-space), called the *Karcher mean*. As we just previously noted, restricted to the matrix case it yields the least squares mean.

Power means for positive definite matrices were introduced by Lim and Palfia [11].

Theorem 3 *Let $A_1, \ldots, A_n \in \Omega$ and let $\mathbf{w} = (w_1, \ldots, w_n)$ be a weight. Then for each $t \in (0, 1]$, the following equation has a unique positive definite solution $X = P_t(\mathbf{w}; A_1, \ldots, A_n)$, called the weighted power mean:*

$$X = \sum_{i=1}^{n} w_i (X \#_t A_i).$$

When restricted to the positive reals, the power mean reduces to the usual power mean

$$P_t(\mathbf{w}; a_1, \ldots, a_n) = \left(w_1 a_1^t + \cdots + w_n a_n^t\right)^{\frac{1}{t}}.$$

In 2014 the authors showed [9] that the preceding notion of power mean extended to the setting of bounded operators on a Hilbert space [9] and established that the power means are decreasing, $s < t$ implies $P_s(\cdot;\ \cdot) \le P_t(\cdot;\ \cdot)$. Using power means we were able to establish the existence and uniqueness of the Karcher mean in the C^*-algebra of bounded operators on a Hilbert space.

Theorem 4 *In the strong operator topology*

$$\Lambda(\cdot;\ \cdot) = \lim_{t\to 0^+} P_t(\cdot;\ \cdot) = \inf_{t>0} P_t(\cdot;\ \cdot),$$

where Λ *is the Karcher mean, the unique solution of the Karcher equation*

$$X = \Lambda_n(A_1, \ldots, A_n) \Leftrightarrow \sum_{i=1}^{n} \log(X^{-1/2} A_i X^{-1/2}) = 0.$$

Via this machinery many of the axiomatic properties of the least squares mean in the finite-dimensional setting were extended to the corresponding Karcher mean in the infinite-dimensional setting.

Recent work by Lim and Palfia [12] and independently by Lawson [7] shows that the preceding constructions and results remain valid for the open cone of positive invertible elements in any unital C^*-algebra.

7 Barycenters

A *Borel probability measure* on a metric space (X, d) is a countably additive nonnegative measure μ on the Borel algebra $\mathcal{B}(X)$, the smallest Σ-algebra containing the open sets, such that $\mu(X) = 1$. We denote the set of all probability measures on $(X, \mathcal{B}(X))$ by $\mathcal{P}(X)$. Let $\mathcal{P}_0(X)$ be the set of all uniform finitely supported probability measures, i.e., all $\mu \in \mathcal{P}(X)$ of the form $\mu = \frac{1}{n}\sum_{j=1}^{n} \delta_{x_j}$ for some $n \in \mathbb{N}$, where δ_x is the point measure of mass 1 at x.

A measure $\mu \in \mathcal{P}(X)$ is said to be *integrable* if

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

The set of integrable measures is denoted by $\mathcal{P}^1(X)$.

The *Wasserstein distance* (alternatively Kantorovich–Rubinstein distance) d^W on $\mathcal{P}^1(X)$ is a standard metric for probability measures. It is known that d^W is a complete metric on $\mathcal{P}^1(X)$ whenever X is a complete metric space and that $\mathcal{P}_0(X)$ is d^W-dense in $\mathcal{P}^1(X)$.

One can view the Karcher mean $(A_1, \ldots, A_n) \mapsto \Lambda(A_1, \ldots, A_n)$ on Ω, the open cone of positive invertible operators, alternatively as yielding a barycenter for the probability measure with weight $1/n$ at each A_k. It turns out that this barycentric map is contractive from $\mathcal{P}_0(\Omega)$ to Ω, and hence extends uniquely to a contractive barycentric map $\Lambda : \mathcal{P}^1(\Omega) \to \Omega$. We call this extended map the *Karcher barycentric map*. It is characterized by

$$X = \Lambda(\mu) \Leftrightarrow \int_\Omega \log(X^{-1/2} A X^{-1/2}) d\mu(A) = 0.$$

The existence and basic theory and properties of the Karcher barycentric map can be found in [10]. We note that from its definition it extends the Karcher mean.

8 Summary

In the preceding we have attempted to trace out how the matrix/operator geometric mean has strikingly developed over the past 15 years from a two-variable mean to a multivariable matrix mean (the least squares mean) to an operator mean in unital C^*-algebras (the Karcher mean) to a barycentric map on integrable Borel probability measures. Whatever future developments may hold, it is clear that a substantial theory has already emerged.

References

1. Ahn, E., Kim, S., Lim, Y.: An extended Lie–Trotter formula and its applications. Linear Algebra Appl. **427**(2–3), 190–196 (2007). MR 2351352
2. Ando, T., Li, C.-K., Mathias, R.: Geometric means. Linear Algebra Appl. **385**, 305–334 (2004). MR 2063358
3. Bhatia, R., Holbrook, J.: Noncommutative geometric means. Math. Intell. **28**(1), 32–39 (2006). MR 2202893
4. Bhatia, R., Holbrook, J.: Riemannian geometry and matrix geometric means. Linear Algebra Appl. **413**(2–3), 594–618 (2006). MR 2198952
5. Bini, D.A., Meini, B., Poloni, F.: An effective matrix geometric mean satisfying the Ando–Li–Mathias properties. Math. Comp. **79**(269), 437–452 (2010). MR 2552234
6. Kubo, F., Ando, T.: Means of positive linear operators. Math. Ann. **246**(3), 205–224 (1979/80). MR 563399
7. Lawson, J.D.: Existence and uniqueness of the Karcher mean on unital c^*-algebras. J. Math. Anal. Appl. **483**(2), 123625 (2020)

8. Lawson, J.D., Lim, Y.: The geometric mean, matrices, metrics, and more. Am. Math. Mon. **108**(9), 797–812 (2001). MR 1864051
9. Lawson, J., Lim, Y.: Karcher means and Karcher equations of positive definite operators. Trans. Am. Math. Soc. Ser. B **1**, 1–22 (2014). MR 3148817
10. Lawson, J.D., Lim, Y.: Contractive barycentric maps. J. Oper. Theory **77**(1), 87–107 (2017). MR 3614507
11. Lim, Y., Pálfia, M.: Matrix power means and the Karcher mean. J. Funct. Anal. **262**(4), 1498–1514 (2012). MR 2873848
12. Lim, Y., Pálfia, M.: Existence and uniqueness of the l^1-Karcher mean (2017). arXiv:1703.04292
13. Moakher, M.: A differential geometric approach to the geometric mean of symmetric positive-definite matrices. SIAM J. Matrix Anal. Appl. **26**(3), 735–747 (2005). MR 2137480
14. Sagae, M., Tanabe, K.: Upper and lower bounds for the arithmetic-geometric-harmonic means of positive definite matrices. Linear Multilinear Algebra **37**(4), 279–282 (1994). MR 1310971
15. Sturm, K.-T.: Probability measures on metric spaces of nonpositive curvature. In: Heat Kernels and Analysis on Manifolds, Graphs, and Metric Spaces (Paris, 2002). Contemporary Mathematics, vol. 338, pp. 357–390. American Mathematical Society, Providence (2003). MR 2039961