A deterministic approach to the least squares mean on Hadamard spaces

Yongdo Lim

Kyungpook National University

Overview

There has recently been considerably interest in defining "means" (averaging, barycenters, centroids) on manifolds/ metric spaces. A natural and attractive candidate of an averaging procedure on a Cartan-Hadamard Riemannian manifold is the least squares mean. This mean has appeared under a variety of other designations: center of gravity, Frechet mean, Cartan mean, Riemannian center of mass, Riemannian geometric mean, or frequently, *Karcher mean*, the terminology we adopt. It plays a central role in image processing (subdivision schemes), medical imaging (DT-MRI), radar systems, and biology (DNA/Genomes), to cite a few.

Our purposes in this talk include the following for Hadamard spaces:

(1) A deterministic ("no dice") approach to the Karcher mean

(2) A limit theorem for contractive means

(3) A no dice conjecture and some geometry of contractive means

Contractive means

Let (M, δ) be a complete metric space and let $G : M^n \to M$.

(1) G is a mean if $G(a, \dots, a) = a$.

(2) G is symmetric if $G(a_{\sigma(1)}, \dots, a_{\sigma(n)}) = G(a_1, \dots, a_n)$.

(3) G is contractive if $\delta(G(a_1, \dots, a_n), G(b_1, \dots, b_n)) \leqslant \frac{1}{n} \sum_{i=1}^{n} \delta(a_i, b_i)$.

Contractive means

Let (M, δ) be a complete metric space and let $G : M^n \to M$.

(1) G is a mean if $G(a, \ldots, a) = a$.

(2) G is symmetric if $G(a_{\sigma(1)}, \ldots, a_{\sigma(n)}) = G(a_1, \ldots, a_n)$.

(3) G is contractive if $\delta(G(a_1, \ldots, a_n), G(b_1, \ldots, b_n)) \leqslant \frac{1}{n} \sum_{i=1}^{n} \delta(a_i, b_i)$.

For contractive n-means G and H, define

$$\delta(G, H) = \sup_{\substack{\mathbf{a} \in M^n \\ \Delta(\mathbf{a}) \neq 0}} \frac{\delta(G(\mathbf{a}), H(\mathbf{a}))}{\Delta(\mathbf{a})},$$

where $\Delta(a_1, \ldots, a_n) = \max_{1 \leqslant i,j \leqslant n} \delta(a_i, a_j)$.

• The set of all contractive n-means is a metric space with Diam $\leqslant 1$. Question: Which metric spaces support contractive n-means (for some or all $n \geqslant 2$) ?

Least Squares Mean

The Karcher mean of $a_1, \ldots, a_n \in M$, is defined as the unique minimizer (provided it exists) of the optimization problem

$$\Lambda(a_1, \ldots, a_n) = \underset{x \in M}{\operatorname{argmin}} \sum_{i=1}^{n} \delta^2(x, a_i).$$

This idea had been anticipated by Élie Cartan, who showed among other things that such a unique minimizer exists if the points all lie in a convex ball in a Riemannian manifold. Alternatively, it arises as the unique solution of the *Karcher equation*:

$$\sum_{i=1}^{n} \log_x a_i = 0.$$

• Recent research includes: Numerical methods for its approximation, e.g., Newton and gradient descent methods (exhibit local convergence, heavy dependence on initial points and step length); detailed investigation for the case of positive definite matrices. Question: What about its contractiveness?

Means and S.L.L.N.

For $\{1, 2, \ldots, n\}$ equipped with a probability measure, $P_n := \{1, 2, \ldots, n\}^{\mathbb{N}}$ is a probability space w.r.t. the product measure, e.g.,

$$w^* : \mathbb{N} \to \{1, 2, 3\}, \quad \omega^* : 1, 2, 3, 1, 2, 3, 1, 2, 3, 1, \ldots, .$$

Means and S.L.L.N.

For $\{1,2,\ldots,n\}$ equipped with a probability measure, $P_n := \{1,2,\ldots,n\}^{\mathbb{N}}$ is a probability space w.r.t. the product measure, e.g.,

$$w^* : \mathbb{N} \to \{1,2,3\}, \quad \omega^* : 1,2,3,1,2,3,1,2,3,1,\ldots,.$$

Let $\mu = \{G_k\}_{k=1}^{\infty}$ be a sequence of means $G_k : M^k \to M$. The Strong Law of Large Numbers for μ can be stated as follows: for any n-tuple $\mathbf{a} = (a_1,\ldots,a_n) \in M^n$,

$$\exists \lim_{k\to\infty} G_k(a_{\omega(1)},\ldots,a_{\omega(k)}), \quad a.e.\ \omega \in P_n.$$

- Define $\mu^*(a_1,\ldots,a_n)$ to be the “common” limit.
- Find such a $\mu = \{G_k\}_{k=1}^{\infty}$ and describe the mean μ^*.
- “No Dice” Conjecture:
  $\lim_{k\to\infty} G_k(a_{\omega^*(1)},\ldots,a_{\omega^*(k)}) = \mu^*(a_1,\ldots,a_n).$

Contractive means on Busemann spaces

There are several types of contractive means on Busemann NPC spaces. A *metric space with midpoint selector* is a triple $(M, \delta, \#)$ such that (M, δ) is a metric space and $(x, y) \mapsto x\#y$ is a binary operation that assigns to each point a metric midpoint in such a way that $x\#y = y\#x$. Menger proved that a complete metric space with midpoint selector is a geodesic metric space; for all distinct $x, y \in M$, there exists a minimal geodesic $\gamma_{x,y}; [0, 1] \to M$ from x to y. Denote

$$x\#_t y := \gamma_{x,y}(t), \qquad (x\#y = x\#_{\frac{1}{2}} y)$$

the t-*weighted geometric mean of* a *and* b.

Definition

A complete metric space with midpoint selector $(M, \delta, \#)$ is said to satisfy the *Busemann NPC-inequality*, if for all a, b, c, d

$$\delta(a\#b, c\#d) \leqslant \frac{1}{2}[\delta(a, c) + \delta(b, d)].$$

Lemma

Let $(M, \delta, \#)$ *be a complete metric space with midpoint selector satisfying the Busemann NPC-inequality. Then for* $0 \leqslant t \leqslant 1$,

$$\delta(a\#_t b, c\#_t d) \quad \leqslant \quad (1-t)\delta(a, c) + t\delta(b, d).$$

• Hadamard spaces. • Symmetric cones equipped with Thompson

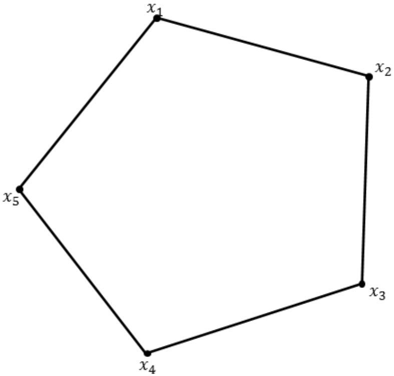
metric: Domains of positivity, Jordan-Banach algebras (e.g. the convex cone of positive definite operators on a Hilbert space, forward light cones)

• Finite dim. symmetric cones: self-dual homogeneous cones, Euclidean Jordan algebras, having rich (Finsler) geometric structures inherited from symmetric gauge norms (equivalently, unitarily invariant norms; Schatten p-norms and spectral norm).

Inductive mean and Birkhoff shortening

Let $(M, \delta, \#)$ be a complete metric space with midpoint selector satisfying the *Busemann NPC-inequality*.
The inductive mean (convex combin.) of $a_1, \ldots, a_n$ is defined as

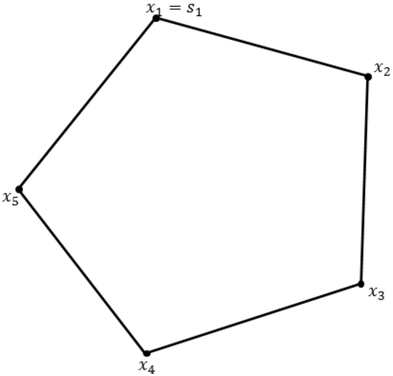
$$S(a_1) = a_1, \quad S(a_1, \ldots, a_n) = S(a_1, \ldots, a_{n-1}) \#_{\frac{1}{n}} a_n.$$



Inductive mean and Birkhoff shortening

Let $(M, \delta, \#)$ be a complete metric space with midpoint selector satisfying the *Busemann NPC-inequality*.
The inductive mean (convex combin.) of $a_1, \ldots, a_n$ is defined as

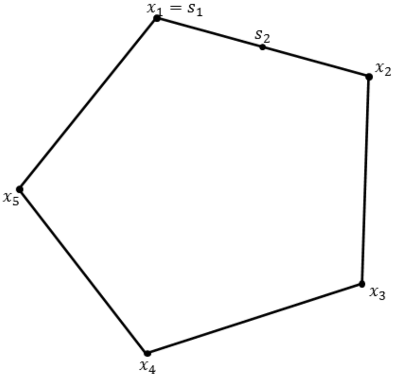
$$S(a_1) = a_1, \quad S(a_1, \ldots, a_n) = S(a_1, \ldots, a_{n-1}) \#_{\frac{1}{n}} a_n.$$



Inductive mean and Birkhoff shortening

Let $(M, \delta, \#)$ be a complete metric space with midpoint selector satisfying the *Busemann NPC-inequality*.
The inductive mean (convex combin.) of $a_1, \ldots, a_n$ is defined as

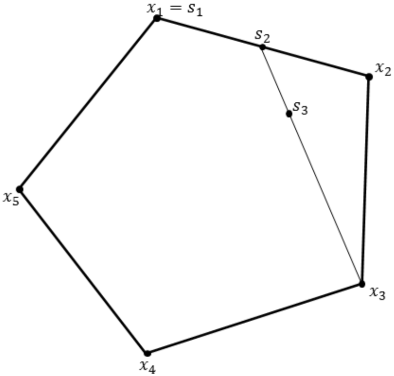
$$S(a_1) = a_1, \quad S(a_1, \ldots, a_n) = S(a_1, \ldots, a_{n-1}) \#_{\frac{1}{n}} a_n.$$



Inductive mean and Birkhoff shortening

Let $(M, \delta, \#)$ be a complete metric space with midpoint selector satisfying the *Busemann NPC-inequality*.
The inductive mean (convex combin.) of $a_1, \ldots, a_n$ is defined as

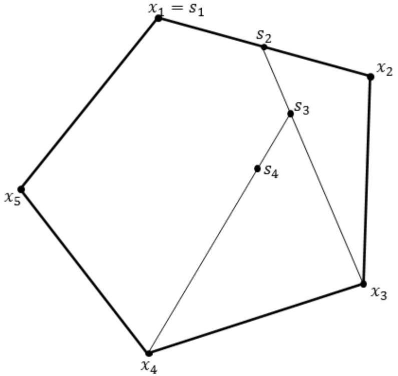
$$S(a_1) = a_1, \quad S(a_1, \ldots, a_n) = S(a_1, \ldots, a_{n-1}) \#_{\frac{1}{n}} a_n.$$



Inductive mean and Birkhoff shortening

Let $(M, \delta, \#)$ be a complete metric space with midpoint selector satisfying the *Busemann NPC-inequality*.
The inductive mean (convex combin.) of $a_1, \ldots, a_n$ is defined as

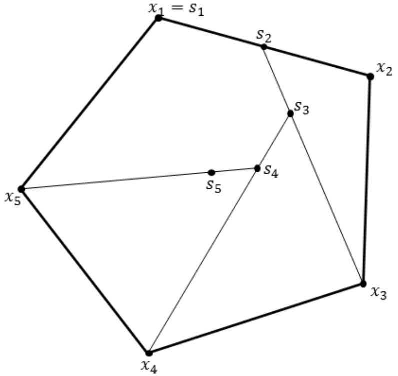
$$S(a_1) = a_1, \quad S(a_1, \ldots, a_n) = S(a_1, \ldots, a_{n-1}) \#_{\frac{1}{n}} a_n.$$



Inductive mean and Birkhoff shortening

Let $(M, \delta, \#)$ be a complete metric space with midpoint selector satisfying the *Busemann NPC-inequality*.
The inductive mean (convex combin.) of $a_1, \ldots, a_n$ is defined as

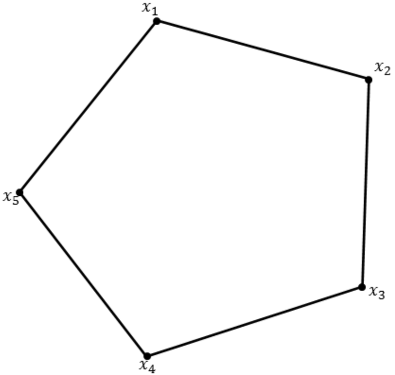
$$S(a_1) = a_1, \quad S(a_1, \ldots, a_n) = S(a_1, \ldots, a_{n-1}) \#_{\frac{1}{n}} a_n.$$



Inductive mean and Birkhoff shortening

The Birkhoff shortening of $a_1, \ldots, a_n$ is the common limit of the the following sequences

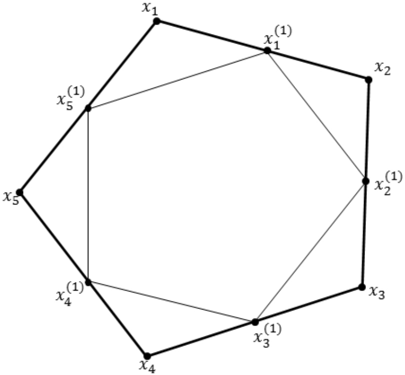
$$\begin{aligned} \gamma(a_1, \ldots, a_n) &= (a_1 \# a_2, a_2 \# a_3, \ldots, a_n \# a_1) \\ \lim_{k \to \infty} \gamma^k(a_1, \ldots, a_n) &= (x^*, x^*, \ldots, x^*) \end{aligned}$$



Inductive mean and Birkhoff shortening

The Birkhoff shortening of $a_1, \ldots, a_n$ is the common limit of the the following sequences

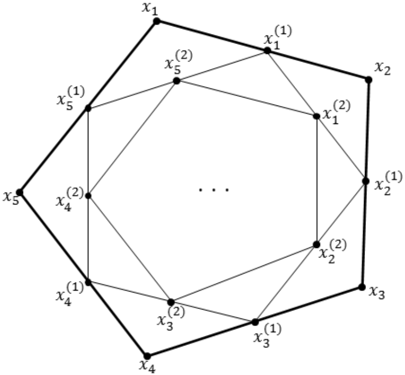
$$\begin{aligned} \gamma(a_1, \ldots, a_n) &= (a_1 \# a_2, a_2 \# a_3, \ldots, a_n \# a_1) \\ \lim_{k\to\infty} \gamma^k(a_1, \ldots, a_n) &= (x^*, x^*, \ldots, x^*) \end{aligned}$$



Inductive mean and Birkhoff shortening

The Birkhoff shortening of $a_1, \ldots, a_n$ is the common limit of the the following sequences

$$\begin{aligned} \gamma(a_1, \ldots, a_n) &= (a_1 \# a_2, a_2 \# a_3, \ldots, a_n \# a_1) \\ \lim_{k\to\infty} \gamma^k(a_1, \ldots, a_n) &= (x^*, x^*, \ldots, x^*) \end{aligned}$$



Inductive mean and Birkhoff shortening

Let $(M, \delta, \#)$ be a complete metric space with midpoint selector satisfying the *Busemann NPC-inequality*.
The inductive mean (convex combin.) of $a_1, \ldots, a_n$ is defined as

$$S(a_1) = a_1, \quad S(a_1, \ldots, a_n) = S(a_1, \ldots, a_{n-1}) \#_{\frac{1}{n}} a_n.$$

The Birkhoff shortening of $a_1, \ldots, a_n$ is the common limit of the following sequences

$$\begin{aligned} \gamma(a_1, \ldots, a_n) &= (a_1 \# a_2, a_2 \# a_3, \ldots, a_n \# a_1) \\ \lim_{k \to \infty} \gamma^k(a_1, \ldots, a_n) &= (x^*, x^*, \ldots, x^*) \end{aligned}$$

• Indeed, the common limit exists.

• Inductive mean and Birkhoff shortening are contractive, but non-symmetric (can vary with permutations).

ALM and BMP Means

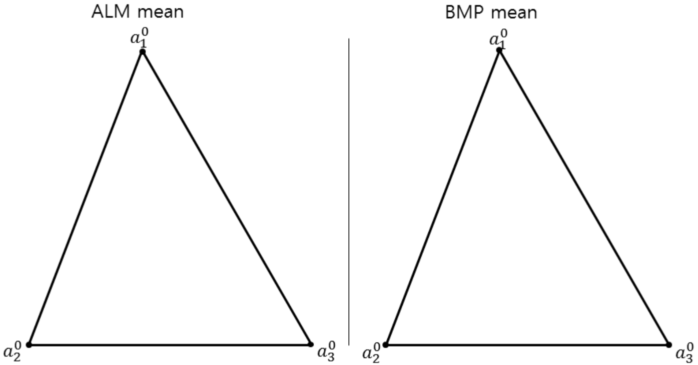
There are alternative symmetric and contractive means, namely ALM $= \{\mathrm{Alm}_n\}_{n=2}^{\infty}$ and BMP $= \{\mathrm{Bmp}_n\}_{n=2}^{\infty}$ means, on a complete metric space with midpoint selector satisfying the Busemann NPC-inequality via "symmetrization procedures" and "induction".

• See Ando-Li-Mathias for positive definite matrices, Es-Sahib and Heinich on locally compact Hadamard spaces, Bini-Meini-Poloni for positive definite matrices.

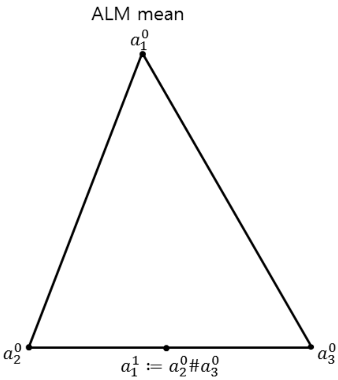
• $\mathrm{Alm}_3 =$ the Birkhoff shortening.

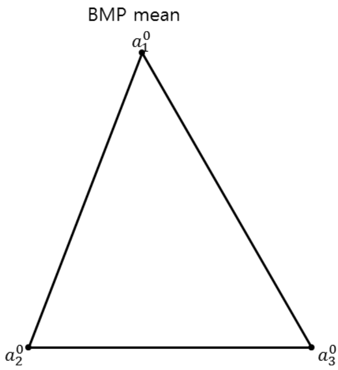
• There are infinitely many symmetric and contractive means on on a complete metric space with midpoint selector satisfying the Busemann NPC-inequality.

ALM and BMP Means

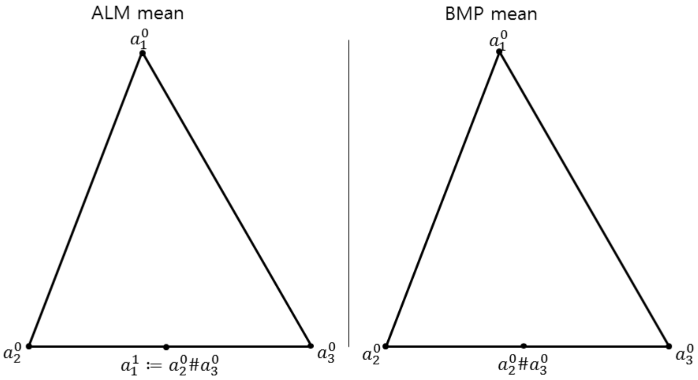


ALM and BMP Means

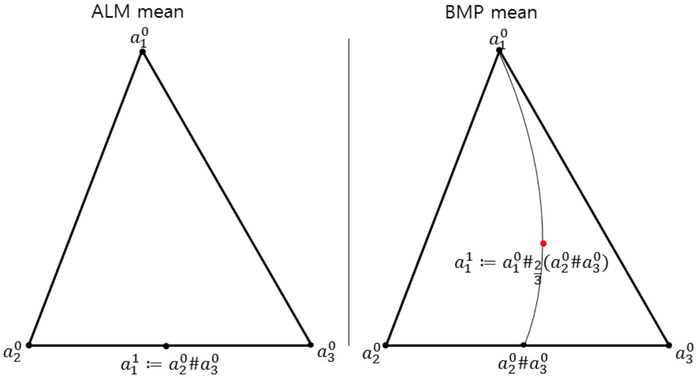




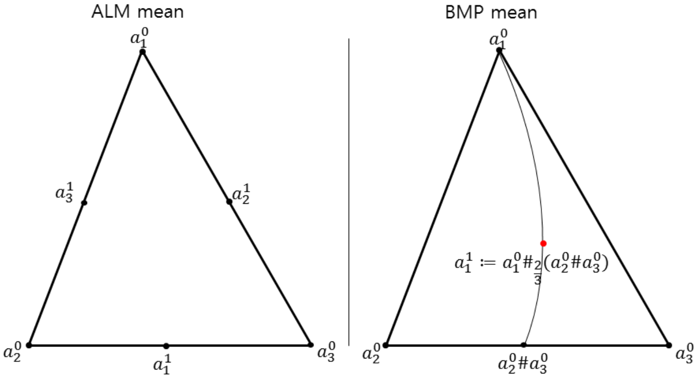
ALM and BMP Means



ALM and BMP Means



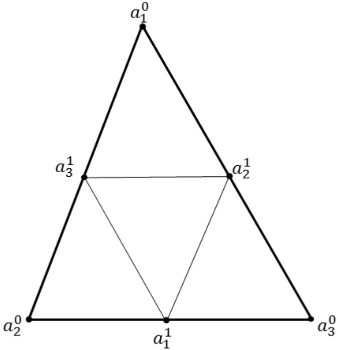
ALM and BMP Means



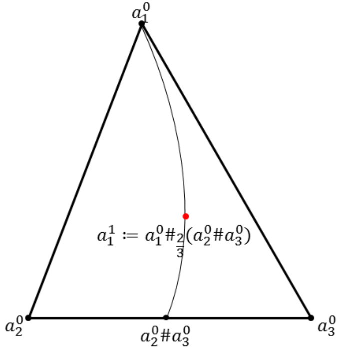
$$a_1^1 := (a_2^0 \# a_3^0)$$
$$a_2^1 := (a_3^0 \# a_1^0)$$
$$a_3^1 := (a_1^0 \# a_2^0)$$

ALM and BMP Means

ALM mean

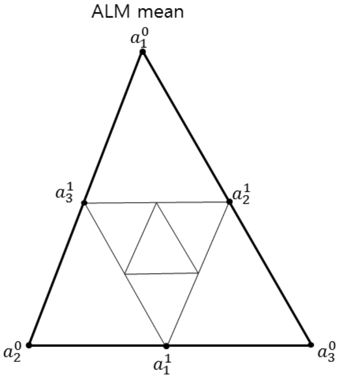


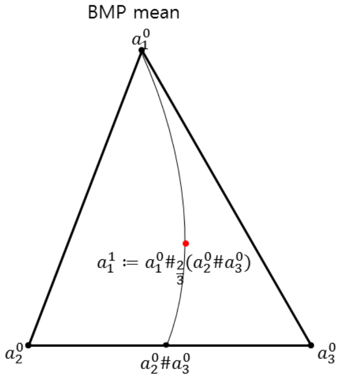
BMP mean



$$a_1^1 := (a_2^0 \# a_3^0)$$
$$a_2^1 := (a_3^0 \# a_1^0)$$
$$a_3^1 := (a_1^0 \# a_2^0)$$

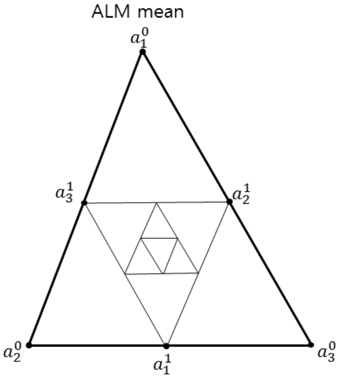
ALM and BMP Means

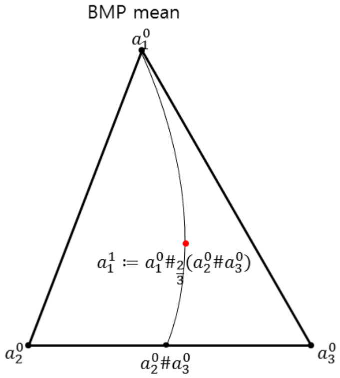




$$a_1^1 := (a_2^0 \# a_3^0)$$
$$a_2^1 := (a_3^0 \# a_1^0)$$
$$a_3^1 := (a_1^0 \# a_2^0)$$

ALM and BMP Means

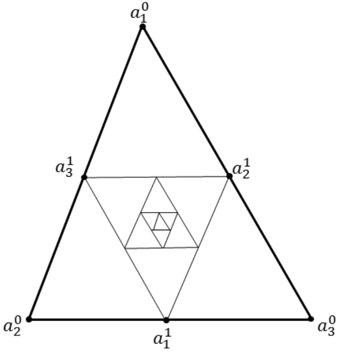




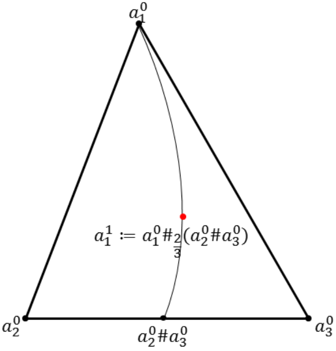
$$a_1^1 := (a_2^0 \# a_3^0)$$
$$a_2^1 := (a_3^0 \# a_1^0)$$
$$a_3^1 := (a_1^0 \# a_2^0)$$

ALM and BMP Means

ALM mean



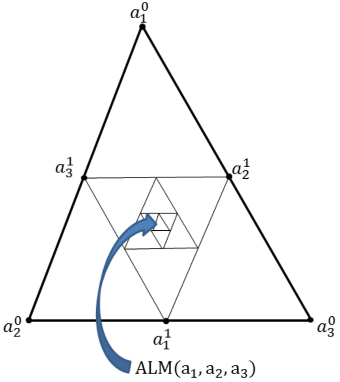
BMP mean



$$a_1^1 := (a_2^0 \# a_3^0)$$
$$a_2^1 := (a_3^0 \# a_1^0)$$
$$a_3^1 := (a_1^0 \# a_2^0)$$

ALM and BMP Means

ALM mean

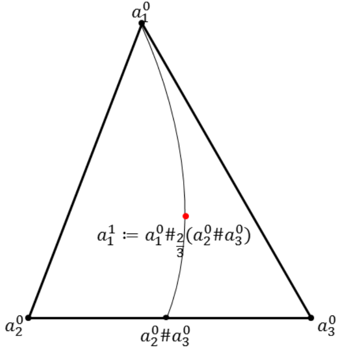


$$a_1^1 := (a_2^0 \# a_3^0)$$

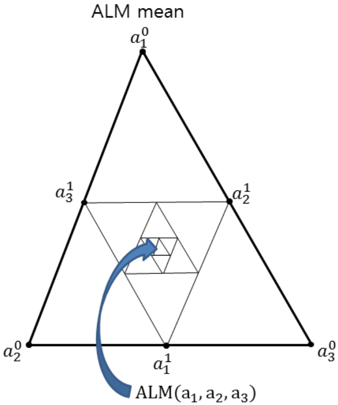
$$a_2^1 := (a_3^0 \# a_1^0)$$

$$a_3^1 := (a_1^0 \# a_2^0)$$

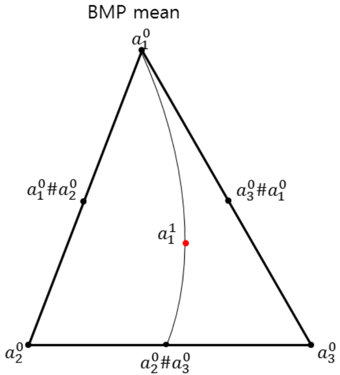
BMP mean



ALM and BMP Means

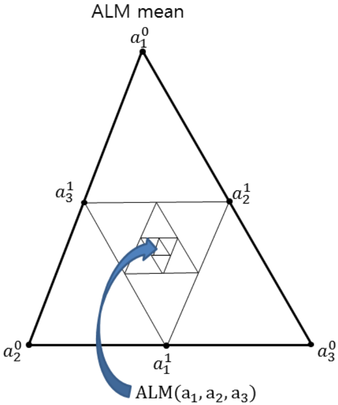


$$a_1^1 := (a_2^0 \# a_3^0)$$
$$a_2^1 := (a_3^0 \# a_1^0)$$
$$a_3^1 := (a_1^0 \# a_2^0)$$

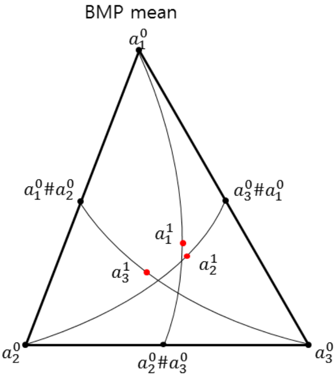


$$a_1^1 := a_1^0 \#_{\frac{2}{3}} (a_2^0 \# a_3^0)$$
$$a_2^1 := a_2^0 \#_{\frac{2}{3}} (a_3^0 \# a_1^0)$$
$$a_3^1 := a_3^0 \#_{\frac{2}{3}} (a_1^0 \# a_2^0)$$

ALM and BMP Means

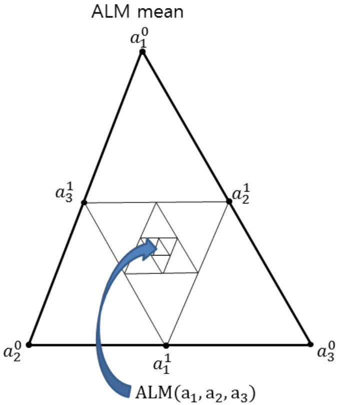


$$a_1^1 := (a_2^0 \# a_3^0)$$
$$a_2^1 := (a_3^0 \# a_1^0)$$
$$a_3^1 := (a_1^0 \# a_2^0)$$

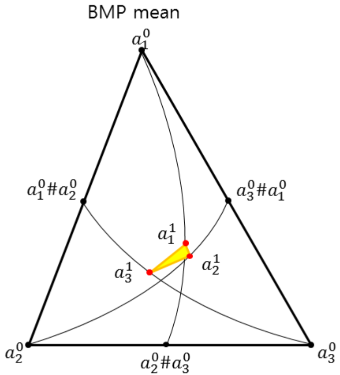


$$a_1^1 := a_1^0 \#_{\frac{2}{3}} (a_2^0 \# a_3^0)$$
$$a_2^1 := a_2^0 \#_{\frac{2}{3}} (a_3^0 \# a_1^0)$$
$$a_3^1 := a_3^0 \#_{\frac{2}{3}} (a_1^0 \# a_2^0)$$

ALM and BMP Means



$$a_1^1 := (a_2^0 \# a_3^0)$$
$$a_2^1 := (a_3^0 \# a_1^0)$$
$$a_3^1 := (a_1^0 \# a_2^0)$$



$$a_1^1 := a_1^0 \#_{\frac{2}{3}} (a_2^0 \# a_3^0)$$
$$a_2^1 := a_2^0 \#_{\frac{2}{3}} (a_3^0 \# a_1^0)$$
$$a_3^1 := a_3^0 \#_{\frac{2}{3}} (a_1^0 \# a_2^0)$$

Hadamard Space

A complete metric space (M, δ) is called a *Hadamard space* if it satisfies the semiparallelogram law; for each $x, y \in M$, there exists an $m \in M$ satisfying

$$\delta^2(m, z) \leqslant \frac{1}{2}\delta^2(x, z) + \frac{1}{2}\delta^2(y, z) - \frac{1}{4}\delta^2(x, y), \forall z \in M.$$

- m is the unique metric midpoint between x and y.
- unique minimal geodesic $\gamma_{a,b}$; $\quad a\#_t b := \gamma_{a,b}(t)$.
- Hadamard space $\Longrightarrow$ Busemann NPC space
- Inductive, Birkhoff shortening, ALM and BMP means exist.
- The Karcher mean exists and unique: $\Lambda(a_1, \ldots, a_n) = \underset{x \in M}{\operatorname{argmin}} \sum_{i=1}^{n} \delta^2(x, a_i)$

Examples of Hadamard Spaces

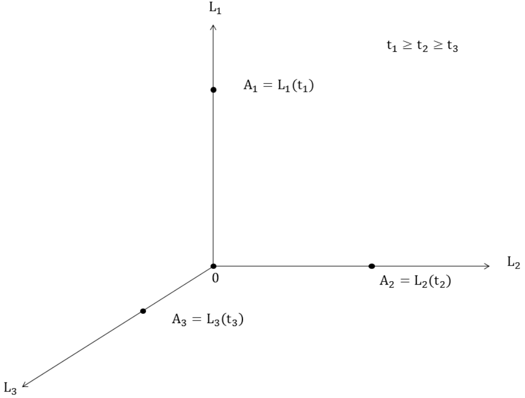
- Cartan-Hadamard Riemannian manifolds; e.g., the convex cone of positive definite matrices with trace metric

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

  In this case $A \#_t B = A^{1/2}(A^{-1/2} B A^{-1/2})^t A^{1/2}$, the t-weighted matrix geometric mean of A and B.
- (Infinite dim.) Lorentz cones
- Phylogenetic Trees (DNA sequences and Genomes)
- Booklets, metric trees,
- Spiders equipped with Taxicab metric
- Subsets, images, products and Gromov-Hausdorff limits of Hadamard spaces

The 3-spider

The 3-spider is defined as the set of 3-distinct half-rays in the plane with a common point at the origin.



Problem: Compute the Karcher, ALM and BMP of $A_i = L(t_i), i = 1, 2, 3$.

Formulae

Let $A_i = L_i(t_i), i = 1, 2, 3$. We may assume by the "permutation invariancy" that $t_1 \geqslant t_2 \geqslant t_3 \geqslant 0$.

- [T. Sturm] $\Lambda = \begin{cases} 0, & \text{if } \; t_1 \leqslant t_2 + t_3, \\ L_1(\frac{t_1 - t_2 - t_3}{3}), & \text{if } \; t_1 \geqslant t_2 + t_3. \end{cases}$
- $\mathrm{Alm}_3 = L_1\left(\frac{t_1 - t_2}{3}\right)$.
- $\mathrm{Bmp}_3 = \begin{cases} 0, & \text{if } \; t_1 = t_2, \\ L_1\left(\frac{(2^{n+2}-1)(t_1 - t_2) - t_3}{3^{n+2}}\right), & \text{if } \; t_1 \neq t_2, \end{cases}$ where

$$n = \left\lfloor \log_2\left(\frac{t_1 - t_2 + t_3}{t_1 - t_2}\right)\right\rfloor .$$

The least squares mean, ALM and BMP means are indeed distinct. The ALM mean is computationally cumbersome and BMP mean exhibits more rapid convergence properties for positive definite matrices. Not true on the 3-spider.

Least Squares Mean

Let (M, δ) be a Hadamard space. The Karcher mean is $\Lambda(a_1, \ldots, a_n) = \underset{x \in M}{\operatorname{argmin}} \sum_{i=1}^{n} \delta^2(x, a_i)$.

- [T. Sturm] S.L.L.N holds for the inductive mean;

$$\lim_{k \to \infty} S(a_{\omega(1)}, \ldots, a_{\omega(k)}) = \Lambda(a_1, \ldots, a_n), \quad a.e.\ \omega \in P_n.$$

Least Squares Mean

Let (M, δ) be a Hadamard space. The Karcher mean is $\Lambda(a_1, \ldots, a_n) = \underset{x \in M}{\operatorname{argmin}} \sum_{i=1}^{n} \delta^2(x, a_i)$.

- [T. Sturm] S.L.L.N holds for the inductive mean;

$$\lim_{k\to\infty} S(a_{\omega(1)}, \ldots, a_{\omega(k)}) = \Lambda(a_1, \ldots, a_n), \quad a.e.\ \omega \in P_n.$$

- [J. Holbrook] For positive definite matrices

$$\lim_{k\to\infty} S(A_{\omega^*(1)}, \ldots, A_{\omega^*(k)}) = \Lambda(A_1, \ldots, A_n).$$

The proof depends heavily on matrix analysis and the Riemannian trace metric. Very slow in convergence. The "No Dice" conjecture remains open for general Hadamard spaces (no differential structures).
[cf. Es-Sahib and Heinich's S.L.L.N and no dice via ALM mean]

Geometric power means

Let (M, δ) be a Hadamard space.

Theorem

Let $G : M^n \to M$ be a contractive mean and let $t \in (0, 1]$. Then for each $(a_1, \dots, a_n) \in M^n$, the following equation has a unique solution in M:

$$x = G(x\#_t a_1, \dots, x\#_t a_n). \tag{1}$$

Define by $G_t(a_1, \dots, a_n)$ the unique solution of (1).

- $G_1 = G$ and each G_t is a contractive mean.
- The map $x \to G(x\#_t a_1, \dots, x\#_t a_n)$ is a strict contraction with the least contraction coefficient less than or equal to $1 - t$.
- The preceding hold on any complete metric space with midpoint selector satisfying the *Busemann NPC-inequality*.

Geometric power means

Definition

A map $G : M^n \to M$ is said to satisfy the *extended metric inequality*, EMI for short, if

$$\delta^2(x, G(\mathbf{a})) \leqslant \frac{1}{n} \sum_{i=1}^{n} \delta^2(x, a_i)$$

for all $x \in M, \mathbf{a} = (a_1, \ldots, a_n) \in M^n$. We denote $\mathfrak{E}_n$ by the set of all contractive n-means satisfying EMI.

Definition

A map $G : M^n \to M$ is said to satisfy the *extended metric inequality*, EMI for short, if

$$\delta^2(x, G(\mathbf{a})) \leqslant \frac{1}{n} \sum_{i=1}^{n} \delta^2(x, a_i)$$

for all $x \in M, \mathbf{a} = (a_1, \ldots, a_n) \in M^n$. We denote $\mathfrak{E}_n$ by the set of all contractive n-means satisfying EMI.

- $\mathfrak{E}_2 = \{\#\}$.

Geometric power means

Definition

A map $G : M^n \to M$ is said to satisfy the *extended metric inequality*, EMI for short, if

$$\delta^2(x, G(\mathbf{a})) \leqslant \frac{1}{n} \sum_{i=1}^{n} \delta^2(x, a_i)$$

for all $x \in M, \mathbf{a} = (a_1, \ldots, a_n) \in M^n$. We denote $\mathfrak{E}_n$ by the set of all contractive n-means satisfying EMI.

- $\mathfrak{E}_2 = \{\#\}$.
- S, Alm_n, Bmp_n and Birkhoff shortening belong to $\mathfrak{E}_n$.

Geometric power means

Definition

A map $G : M^n \to M$ is said to satisfy the *extended metric inequality*, EMI for short, if

$$\delta^2(x, G(\mathbf{a})) \leqslant \frac{1}{n} \sum_{i=1}^{n} \delta^2(x, a_i)$$

for all $x \in M, \mathbf{a} = (a_1, \ldots, a_n) \in M^n$. We denote $\mathfrak{E}_n$ by the set of all contractive n-means satisfying EMI.

- $\mathfrak{E}_2 = \{\#\}$.
- S, Alm_n, Bmp_n and Birkhoff shortening belong to $\mathfrak{E}_n$.
- $G_t \in \mathfrak{E}_n$ for all $t \in (0, 1]$ if $G \in \mathfrak{E}_n$.

Geometric power means

Definition

A map $G : M^n \to M$ is said to satisfy the *extended metric inequality*, EMI for short, if

$$\delta^2(x, G(\mathbf{a})) \leqslant \frac{1}{n} \sum_{i=1}^{n} \delta^2(x, a_i)$$

for all $x \in M, \mathbf{a} = (a_1, \ldots, a_n) \in M^n$. We denote $\mathfrak{E}_n$ by the set of all contractive n-means satisfying EMI.

- $\mathfrak{E}_2 = \{\#\}$.
- $S, \mathrm{Alm}_n, \mathrm{Bmp}_n$ and Birkhoff shortening belong to $\mathfrak{E}_n$.
- $G_t \in \mathfrak{E}_n$ for all $t \in (0, 1]$ if $G \in \mathfrak{E}_n$.
- $G\#_t H \in \mathfrak{E}_n$ if $G, H \in \mathfrak{E}_n$, where $(G\#_t H)(\mathbf{a}) = G(\mathbf{a})\#_t H(\mathbf{a})$.

Geometric power means

Definition

A map $G : M^n \to M$ is said to satisfy the *extended metric inequality*, EMI for short, if

$$\delta^2(x, G(\mathbf{a})) \leqslant \frac{1}{n} \sum_{i=1}^{n} \delta^2(x, a_i)$$

for all $x \in M, \mathbf{a} = (a_1, \ldots, a_n) \in M^n$. We denote $\mathfrak{E}_n$ by the set of all contractive n-means satisfying EMI.

- $\mathfrak{E}_2 = \{\#\}$.
- S, Alm_n, Bmp_n and Birkhoff shortening belong to $\mathfrak{E}_n$.
- $G_t \in \mathfrak{E}_n$ for all $t \in (0, 1]$ if $G \in \mathfrak{E}_n$.
- $G\#_t H \in \mathfrak{E}_n$ if $G, H \in \mathfrak{E}_n$, where $(G\#_t H)(\mathbf{a}) = G(\mathbf{a})\#_t H(\mathbf{a})$.
- $[G_t]_s = G_{st}$, that is, $\{G_t\}_{t\in(0,1]}$ forms a one-parameter semigroup.

A limit theorem

Theorem

Let $G \in \mathfrak{E}_n$. *Then*

$$\delta(G_t(\mathbf{a}), G_s(\mathbf{a})) \leqslant \sqrt{\frac{s+t}{2}}\Delta(\mathbf{a}), \quad \forall \mathbf{a} \in M^n.$$

Furthermore, $\lim_{t\to 0^+} G_t(\mathbf{a}) = \Lambda(\mathbf{a})$.

A limit theorem

Theorem

Let $G \in \mathfrak{E}_n$. *Then*

$$\delta(G_t(\mathbf{a}), G_s(\mathbf{a})) \leqslant \sqrt{\frac{s+t}{2}\Delta(\mathbf{a})}, \ \ \forall \mathbf{a} \in M^n.$$

Furthermore, $\lim_{t\to 0^+} G_t(\mathbf{a}) = \Lambda(\mathbf{a})$.

Key idea: $\delta^2(x, G(\mathbf{a})) \leqslant \frac{1}{n}\sum_{i=1}^n \delta^2(x, a_i)$ implies that

$$\delta^2(x, G_t(\mathbf{a})) \leqslant \frac{1}{n}\sum_{i=1}^n \delta^2(x, a_i) - \frac{1-t}{n}\sum_{i=1}^n \delta^2(G_t(\mathbf{a}), a_i).$$

A limit theorem

Theorem

Let $G \in \mathfrak{E}_n$. *Then*

$$\delta(G_t(\mathbf{a}), G_s(\mathbf{a})) \leqslant \sqrt{\frac{s+t}{2}\Delta(\mathbf{a})}, \ \forall \mathbf{a} \in M^n.$$

Furthermore, $\lim_{t \to 0^+} G_t(\mathbf{a}) = \Lambda(\mathbf{a})$.

Key idea: $\delta^2(x, G(\mathbf{a})) \leqslant \frac{1}{n}\sum_{i=1}^n \delta^2(x, a_i)$ implies that

$$\delta^2(x, G_t(\mathbf{a})) \leqslant \frac{1}{n}\sum_{i=1}^n \delta^2(x, a_i) - \frac{1-t}{n}\sum_{i=1}^n \delta^2(G_t(\mathbf{a}), a_i).$$

As $t \to 0^+$,

$$\sum_{i=1}^n \delta^2(G_0(\mathbf{a}), a_i) \leqslant \sum_{i=1}^n \delta^2(x, a_i) - n\delta^2(x, G_0(\mathbf{a})).$$

A limit theorem

Theorem
Let $G \in \mathfrak{E}_n$. *Then*

$$\delta(G_t(\mathbf{a}), G_s(\mathbf{a})) \leqslant \sqrt{\frac{s+t}{2}}\Delta(\mathbf{a}), \ \forall \mathbf{a} \in M^n.$$

Furthermore, $\lim_{t\to 0^+} G_t(\mathbf{a}) = \Lambda(\mathbf{a})$.

• A deterministic approach; $\lim_{t\to 0^+} S_t(\mathbf{a}) = \Lambda(\mathbf{a})$. Almost all properties of the "inductive mean" are preserved in the limit.

A limit theorem

Theorem

Let $G \in \mathfrak{E}_n$. *Then*

$$\delta(G_t(\mathbf{a}), G_s(\mathbf{a})) \leqslant \sqrt{\frac{s+t}{2}\Delta(\mathbf{a})}, \ \forall \mathbf{a} \in M^n.$$

Furthermore, $\lim_{t\to 0^+} G_t(\mathbf{a}) = \Lambda(\mathbf{a})$.

- A deterministic approach; $\lim_{t\to 0^+} S_t(\mathbf{a}) = \Lambda(\mathbf{a})$. Almost all properties of the "inductive mean" are preserved in the limit.
- $\Lambda \in \mathfrak{E}_n$.

A limit theorem

Theorem

Let $G \in \mathfrak{E}_n$. *Then*

$$\delta(G_t(\mathbf{a}), G_s(\mathbf{a})) \leqslant \sqrt{\frac{s+t}{2}}\Delta(\mathbf{a}),\ \ \forall \mathbf{a} \in M^n.$$

Furthermore, $\lim_{t\to 0^+} G_t(\mathbf{a}) = \Lambda(\mathbf{a})$.

- A deterministic approach; $\lim_{t\to 0^+} S_t(\mathbf{a}) = \Lambda(\mathbf{a})$. Almost all properties of the "inductive mean" are preserved in the limit.
- $\Lambda \in \mathfrak{E}_n$.
- Every closed ball (convex set) is stable for the Karcher mean.

A limit theorem

Theorem

Let $G \in \mathfrak{E}_n$. *Then*

$$\delta(G_t(\mathbf{a}), G_s(\mathbf{a})) \leqslant \sqrt{\frac{s+t}{2}}\Delta(\mathbf{a}), \ \forall \mathbf{a} \in M^n.$$

Furthermore, $\lim_{t \to 0^+} G_t(\mathbf{a}) = \Lambda(\mathbf{a})$.

- A deterministic approach; $\lim_{t \to 0^+} S_t(\mathbf{a}) = \Lambda(\mathbf{a})$. Almost all properties of the "inductive mean" are preserved in the limit.
- $\Lambda \in \mathfrak{E}_n$.
- Every closed ball (convex set) is stable for the Karcher mean.
- $\Lambda_t = \Lambda$ for all $t \in (0, 1]$.

A limit theorem

Theorem

Let $G \in \mathfrak{E}_n$. *Then*

$$\delta(G_t(\mathbf{a}), G_s(\mathbf{a})) \leqslant \sqrt{\frac{s+t}{2}\Delta(\mathbf{a})}, \ \forall \mathbf{a} \in M^n.$$

Furthermore, $\lim_{t \to 0^+} G_t(\mathbf{a}) = \Lambda(\mathbf{a})$.

- A deterministic approach; $\lim_{t \to 0^+} S_t(\mathbf{a}) = \Lambda(\mathbf{a})$. Almost all properties of the "inductive mean" are preserved in the limit.
- $\Lambda \in \mathfrak{E}_n$.
- Every closed ball (convex set) is stable for the Karcher mean.
- $\Lambda_t = \Lambda$ for all $t \in (0, 1]$.
- $G_t = G$ for some $t \in (0, 1)$ if and and only if $G = \Lambda$.

A limit theorem

Theorem

Let $\mathrm{G} \in \mathfrak{E}_n$. *Then*

$$\delta(\mathrm{G}_t(\mathbf{a}), \mathrm{G}_s(\mathbf{a})) \leqslant \sqrt{\frac{s+t}{2}} \Delta(\mathbf{a}), \;\; \forall \mathbf{a} \in M^n.$$

Furthermore, $\lim_{t \to 0^+} \mathrm{G}_t(\mathbf{a}) = \Lambda(\mathbf{a})$.

• A deterministic approach; $\lim_{t \to 0^+} S_t(\mathbf{a}) = \Lambda(\mathbf{a})$. Almost all properties of the "inductive mean" are preserved in the limit.

• $\Lambda \in \mathfrak{E}_n$.

• Every closed ball (convex set) is stable for the Karcher mean.

• $\Lambda_t = \Lambda$ for all $t \in (0, 1]$.

• $\mathrm{G}_t = \mathrm{G}$ for some $t \in (0, 1)$ if and and only if $\mathrm{G} = \Lambda$.

• More effective computationally than the "no dice"/ approach as $\Delta(\mathbf{a}) \to 0$.

Metric structures of $\mathfrak{E}_n$

$$\delta(G,H) = \sup_{\substack{\mathbf{a}\in M^n \\ \Delta(\mathbf{a})\neq 0}} \frac{\delta(G(\mathbf{a}),H(\mathbf{a}))}{\Delta(\mathbf{a})}.$$

• $(\mathfrak{E}_n,\delta)$ is a metric space with $\mathrm{Diam}(\mathfrak{E}_n)\leqslant 1$.
For any $G,H,G',H'\in\mathfrak{E}_n$,

• $\delta(G_t,H_t)\leqslant\sqrt{t}$.

• $\delta(G_t,\Lambda)\leqslant\sqrt{\frac{t}{2}}$. In particular, $\delta(G,\Lambda)\leqslant\frac{1}{\sqrt{2}}$.

• Upper bounds are independent on n.

• $\delta(G\#_s H, G\#_t H) = |s-t|\delta(G,H)$.

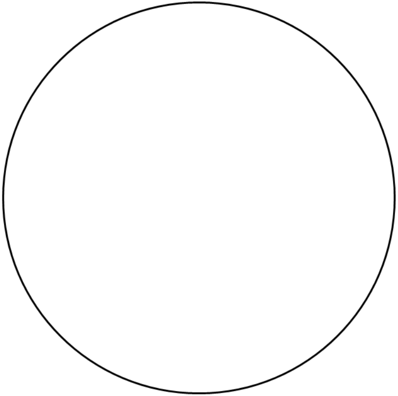
• $\delta(G\#_t H, G'\#_t H')\leqslant(1-t)\delta(G,G')+t\delta(H,H')$.

• Completeness of $(\mathfrak{E}_n,\delta)$? True for any bounded Hadamard spaces.

Metric structures of $\mathfrak{E}_n$

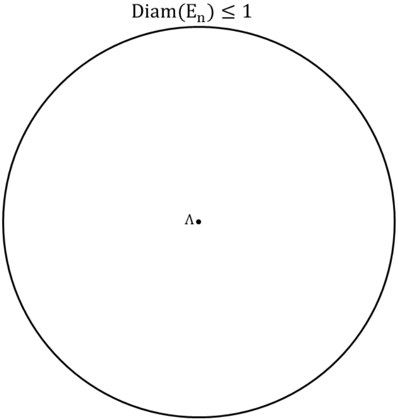
$$\delta(G, H) = \sup_{\substack{\mathbf{a} \in M^n \\ \Delta(\mathbf{a}) \neq 0}} \frac{\delta(G(\mathbf{a}), H(\mathbf{a}))}{\Delta(\mathbf{a})}.$$

$\mathrm{Diam}(E_n) \leq 1$



Metric structures of $\mathfrak{E}_n$

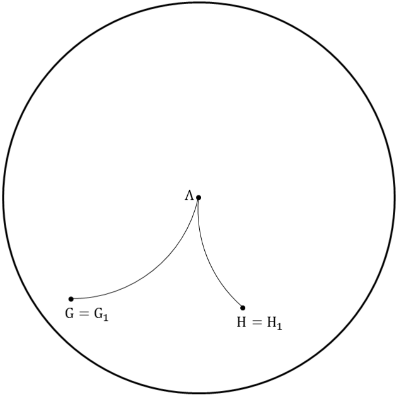
$$\delta(G,H) = \sup_{\substack{\mathbf{a}\in M^n \\ \Delta(\mathbf{a})\neq 0}} \frac{\delta(G(\mathbf{a}),H(\mathbf{a}))}{\Delta(\mathbf{a})}.$$



Metric structures of $\mathfrak{E}_n$

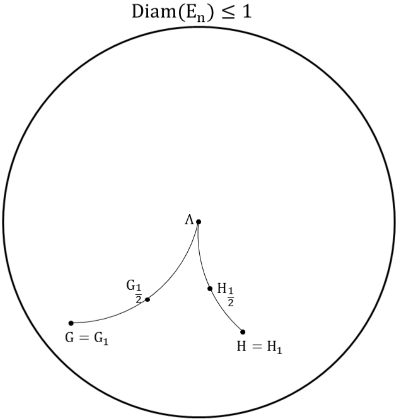
$$\delta(G, H) = \sup_{\substack{\mathbf{a} \in M^n \\ \Delta(\mathbf{a}) \neq 0}} \frac{\delta(G(\mathbf{a}), H(\mathbf{a}))}{\Delta(\mathbf{a})}.$$

Diam(E_n) ≤ 1



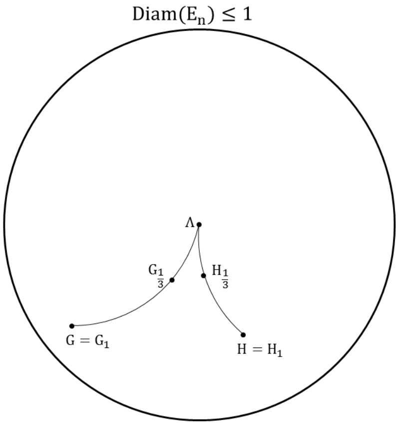
Metric structures of $\mathfrak{E}_n$

$$\delta(G,H) = \sup_{\substack{\mathbf{a}\in M^n \\ \Delta(\mathbf{a})\neq 0}} \frac{\delta(G(\mathbf{a}),H(\mathbf{a}))}{\Delta(\mathbf{a})}.$$



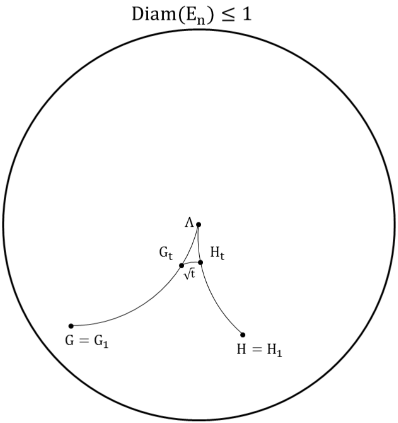
Metric structures of $\mathfrak{E}_n$

$$\delta(G, H) = \sup_{\substack{\mathbf{a}\in M^n \\ \Delta(\mathbf{a})\neq 0}} \frac{\delta(G(\mathbf{a}), H(\mathbf{a}))}{\Delta(\mathbf{a})}.$$



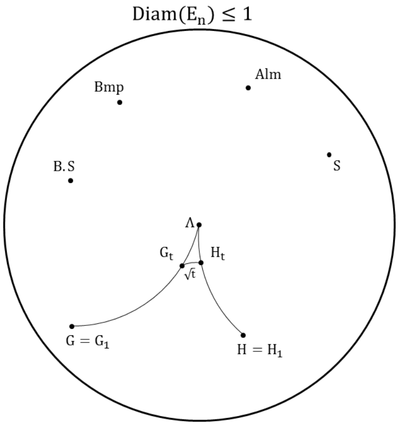
Metric structures of $\mathfrak{E}_n$

$$\delta(G, H) = \sup_{\substack{\mathbf{a}\in M^n \\ \Delta(\mathbf{a})\neq 0}} \frac{\delta(G(\mathbf{a}), H(\mathbf{a}))}{\Delta(\mathbf{a})}.$$



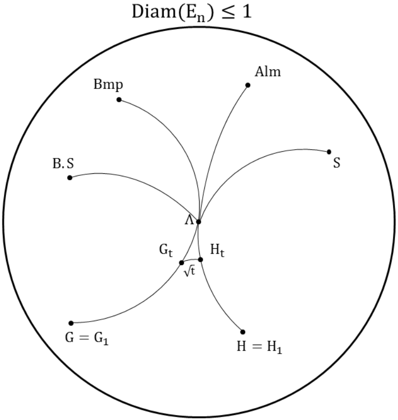
Metric structures of $\mathfrak{E}_n$

$$\delta(G, H) = \sup_{\substack{\mathbf{a}\in M^n \\ \Delta(\mathbf{a})\neq 0}} \frac{\delta(G(\mathbf{a}), H(\mathbf{a}))}{\Delta(\mathbf{a})}.$$



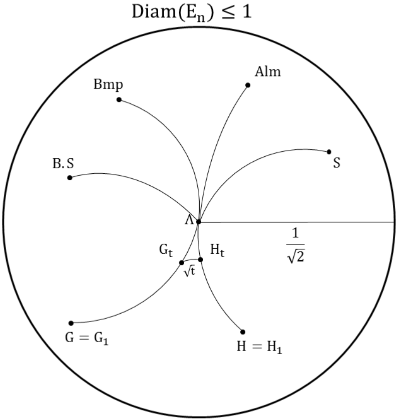
Metric structures of $\mathfrak{E}_n$

$$\delta(G, H) = \sup_{\substack{\mathbf{a}\in M^n \\ \Delta(\mathbf{a})\neq 0}} \frac{\delta(G(\mathbf{a}), H(\mathbf{a}))}{\Delta(\mathbf{a})}.$$



Metric structures of $\mathfrak{E}_n$

$$\delta(G,H) = \sup_{\substack{\mathbf{a}\in M^n \\ \Delta(\mathbf{a})\neq 0}} \frac{\delta(G(\mathbf{a}),H(\mathbf{a}))}{\Delta(\mathbf{a})}.$$



Metric structures of $\mathfrak{E}_n$

$$\delta(G, H) = \sup_{\substack{\mathbf{a} \in M^n \\ \Delta(\mathbf{a}) \neq 0}} \frac{\delta(G(\mathbf{a}), H(\mathbf{a}))}{\Delta(\mathbf{a})}.$$

Define $l_n : \mathfrak{E}_n \to [0, \frac{1}{\sqrt{2}}], \quad l_n(G) = \delta(G, \Lambda).$

Metric structures of $\mathfrak{E}_n$

$$\delta(G, H) = \sup_{\substack{\mathbf{a} \in M^n \\ \Delta(\mathbf{a}) \neq 0}} \frac{\delta(G(\mathbf{a}), H(\mathbf{a}))}{\Delta(\mathbf{a})}.$$

Define $l_n : \mathfrak{E}_n \to [0, \frac{1}{\sqrt{2}}], \quad l_n(G) = \delta(G, \Lambda).$

- Find an explicit formula for $l_n(G)$; e.g., $l_n(S)$.

Metric structures of $\mathfrak{E}_n$

$$\delta(G, H) = \sup_{\substack{\mathbf{a} \in M^n \\ \Delta(\mathbf{a}) \neq 0}} \frac{\delta(G(\mathbf{a}), H(\mathbf{a}))}{\Delta(\mathbf{a})}.$$

Define $l_n : \mathfrak{E}_n \to [0, \frac{1}{\sqrt{2}}]$, $\quad l_n(G) = \delta(G, \Lambda)$.

- Find an explicit formula for $l_n(G)$; e.g., $l_n(S)$.
- On the 3-spider, $l_3(S_t) = \frac{t}{2+t}$, $\quad l_3(S) = l_3(\mathrm{Alm}) = \frac{1}{3}$.

For $A = L_1(a), B = L_2(b), C = L_3(c)$ with $a \geqslant b \geqslant c$,

$$S_t(A, B, C) = \begin{cases} L_1\left(\frac{a-b-c}{3}\right) = \Lambda(A, B, C), & \text{if } \ a \geqslant b + c, \\ L_3\left(\frac{b+c-a}{2+t} t\right), & \text{if } \ a < b + c. \end{cases}$$

Monotonicity

The monotonicity of the Karcher mean of positive definite matrices equipped with the Löwner partial order $A \leqslant B$ if $B - A$ is positive semidefinite,

$$\Lambda(A_1, \ldots, A_n) \leqslant \Lambda(B_1, \ldots, B_n) \quad \text{if } A_i \leqslant B_i, \forall i,$$

was conjectured by Bhatia and Holbrook and is one of key axiomatic properties of matrix geometric means. It was recently established by Lawson and L. via S.L.L.N and by Bhatia and Karandikar via some probabilistic counting arguments, both arguments depending heavily on basic inequalities for the Riemannian metric. Also by Holbrook via a "No dice" approach and by L. and Pálfia by arithmetic power mean approach.
Our approach to the Karcher mean yields a simple, structured, and deterministic proof of the monotonicity extending to all finite dim. symmetric cones and infinite dim. Lorentz cones; the inductive mean S and its geometric power mean S_t are monotonic and hence the Karcher mean, the limit of S_t, also is.

Arithmetic power means

Let $A_1, \ldots, A_n$ be $m \times m$ positive definite matrices. For $t \in (0, 1]$, the following equation has a unique positive definite solution:

$$X = \frac{1}{n} \sum_{i=1}^{n} (X \#_t A_i). \tag{2}$$

Define by $P_t(A_1, \ldots, A_n)$ the unique solution of (2). Then

$$\lim_{t \to 0^+} P_t(A_1, \ldots, A_n) = \Lambda(A_1, \ldots, A_n).$$

• $X \to \frac{1}{n} \sum_{i=1}^{n} X \#_t A_i$ is a strict contraction for the "Thompson metric" with least contraction coefficient $\leqslant 1 - t$. For positive reals,

$P_t(a_1, \ldots, a_n) = \left[\frac{1}{n} \sum_{i=1}^{n} a_i^t\right]^{\frac{1}{t}}$, the usual power mean.

• [The Karcher equation]

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

Further work

Extend the previous work to the case of Infinite dimensional symmetric cones equipped with Thompson metric (Finsler metric). There are infinitely many midpoints between points, but any symmetric cone is a complete metric space with midpoint selector (the geometric mean defined via the Riccati Lemma) satisfying the *Busemann NPC-inequality* with respect to the Thompson metric. Lorentz cones (second order cones) are Cartan-Hadamard Riemannian manifolds.

• Karcher means and Karcher equations on symmetric cones (cf. Lawson and L., for the special case of positive definite operators on a Hilbert space).

• Establish limit theorems for geometric and arithmetic power means on symmetric cones.

• Extend S.L.L.N and “no dice” conjecture for contractive means (e.g., inductive, Birkhoff shortening, ALM and BMP) to symmetric cones.

References-matrices

[1] T. Ando, C.-K. Li and R. Mathias, Geometric means, LAA (2004).

[2] M. Moakher, A differential geometric approach to the geometric mean of symmetric positive-definite matrices, SIAM J. Matrix Anal. Appl. (2005).

[3] R. Bhatia and J. Holbrook, Riemannian geometry and matrix geometric means, LAA (2006).

[4] D. Bini, B. Meini and F. Poloni, An effective matrix geometric mean satisfying the Ando-Li-Mathias properties, Math. Comp. (2010).

[5] D. Bini and B. Iannazzo, Computing the Karcher mean of symmetric positive definite matrices, LAA (2012).

References-monotonicity

[6] J. Lawson and Y. Lim, Monotonic properties of the least squares mean, Math. Ann. (2011).

[7] R. Bhatia and R. Karandikar, Monotonicity of the matrix geometric mean, Math. Ann. (2012).

[8] J. Holbrook, No dice: a determinic approach to the Cartan centroid, J. Ramanujan Math. Soc. (2012)

[9] Y. Lim and M. Pálfia, The matrix power means and the Karcher mean, J. Functional Analysis (2012).

References-metric spaces

[10] H. Karcher, Riemannian center of mass and mollifier smoothing, Comm. Pure Appl. Math. (1977).

[11] A. Es-Sahib and H. Heinich, Barycentre canonique pour un espace métrique à courbure négative, Séminaire de probabilités (Strasbourg) (1999).

[12] K.-T. Sturm, Probability measures on metric spaces of nonpositive curvature, Contemp. Math. (2003).

[13] J. Lawson and Y. Lim, A general framework for extending means to higher orders, Colloq. Math. (2008).

[14] J. Lawson, H. Lee and Y. Lim, Weighted geometric means, Forum Math. (2012).

[15] Y. Lim and M. Pálfia, A deterministic approach for the Karcher mean on Hadamard spaces, submitted.

[16] Y. Lim, Y. Cho et al, Means on the 3-spider, in preparation.