Matrix means in several variables

Miklós Pálfia

Department of Automation and Applied Informatics
Budapest University of Technology and Economics

August 10, 2010

palfia.miklos@aut.bme.hu

Introduction

Let $P(\mathrm{n})$ denote the space of hermitian, positive definite $n \times n$ matrices.

Definition (Matrix mean)

A two-variable function $M : P(n) \times P(n) \mapsto P(n)$ is called a mean function if

(i) $M(X, X) = X$ for every $X \in P(n)$,

(ii) If $X < Y$, then $X < M(X, Y) < Y$,

(iii) If $X < X'$ and $Y < Y'$, then $M(X, Y) < M(X', Y')$,

(iv) $M(X, Y)$ is continuous,

(v) $M(CXC^*, CYC^*) = CM(X, Y)C^*$ for all invertible C.

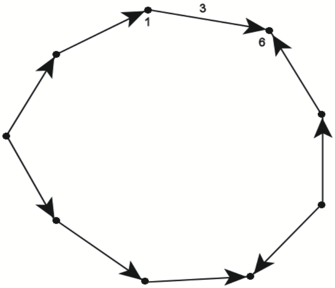
by Kubo-Ando theory

$$M(A, B) = A^{1/2} f\left(A^{-1/2} B A^{-1/2}\right) A^{1/2}$$

Iterative mean

Definition

Let G be a directed graph with n vertices and n edges such that the undirected version of the graph contains a cycle which is Hamiltonian and Eulerian at the same time. Suppose we have a labeling for edges from 1 to n and for the vertices from 1 to n.



Definition (Iterative mean)

Let $X = (X_1^0 \ldots X_n^0)$ where $X_i^0 \in P(r)$ and G^l be an infinite sequence of arbirtarily directed graphs given in as above for $l = 0, 1, 2, \ldots$.
Let $1, \ldots, n$ label the vertices of G^l, and for each edge $i = e(j, s)$ in G^l, define

$$X_i^{l+1} = M(X_j^l, X_s^l), \tag{2.1}$$

where j be the tail vertex and s the head vertex of the edge labelled i and $M(X, Y)$ is a mean satisfying the above properties. Computing these means, we will get another n tuple of matrices as a result. We apply this procedure recursively. This procedure yields a sequence of n-tuple of matrices.

Theorem

The sequences given in the above definition for any matrix mean $M(A,B) \leq \left(\frac{A^2+B^2}{2}\right)^{1/2}$ are convergent for all n and have the same limit point.

note that every symmetric matrix mean

$$M(A,B) \leq \frac{A+B}{2} \leq \left(\frac{A^2+B^2}{2}\right)^{1/2} \tag{2.2}$$

this already implies the above for every symmetric mean

Proof of the convergence of the Iterative mean iteration

Lemma (First Lemma)

The matrix sequences X_i^l are bounded for all n.

Euclidean distance on the space of complex squared matrices

$$d(A,B)^2 = Tr\left\{(A-B)^*(A-B)\right\} \tag{2.3}$$

Lemma (Second Lemma)

For the above distance and for any matrix mean $M(A,B) \leq \left[\left(\frac{A+B}{2}\right)^2 + \epsilon\left(\frac{A-B}{2}\right)^2\right]^{1/2}$ where $\epsilon < 1$ the following holds

$$d(0, M(A,B))^2 \leq \frac{d(0,A)^2 + d(0,B)^2}{2} - \frac{k}{8} d(A,B)^2, \tag{2.4}$$

for $k = 2 - 2\epsilon$.

Note that if $\epsilon = 0$ then the above means that $M(A, B) \leq \frac{A+B}{2}$ which is true for all symmetric means.

Lemma (Third Lemma)

If $M(A, B) \leq \left(\frac{A^2+B^2}{2}\right)^{1/2}$, then

$$\left[\frac{\sum_{i=1}^{n}\left(X_i^{l+1}\right)^2}{n}\right]^{1/2} \leq \left[\frac{\sum_{i=1}^{n}\left(X_i^{l}\right)^2}{n}\right]^{1/2}. \qquad (2.5)$$

The rest of the proof of the assertion

Let us consider one iteration step, which maps pairs of X_i^l to a X_s^{l+1} by taking the mean $M(X_i^l, X_j^l)$ of the matrices corresponding to the two vertices of an edge $s = e(i,j)$. Using the Second Lemma

$$d\left(0, X_i^1\right)^2 \leq \frac{d\left(0, X_{j_i}^0\right)^2 + d\left(0, X_{s_i}^0\right)^2}{2} - \frac{k}{8} d\left(X_{j_i}^0, X_{s_i^0}\right)^2, \quad (2.6)$$

where $X_i^1 = M(X_{j_i}^0, X_{s_i}^0)$. The undirected version of the graph has one cycle which is Hamiltonian and Eulerian, so every vertex has two edges connected to it. So if we sum up the equations above for every edge we get

$$\underbrace{\sum_{i=1}^{n} d\left(0, X_i^{l+1}\right)^2}_{a_{l+1}} \leq \underbrace{\sum_{i=1}^{n} d\left(0, X_i^{l}\right)^2}_{a_l} - \frac{k}{8} \underbrace{\sum_{i=1}^{n} d\left(X_{j_i}^l, X_{s_i}^l\right)^2}_{e_l}. \quad (2.7)$$

$a_l \geq 0$ measures the sum of the squared distances from 0 and we have

$$a_{l+1} \leq a_l - \frac{k}{8} e_l, \tag{2.8}$$

so $\lim_{l\to\infty} a_l$ exists, hence $e_l = \sum_{i=1}^{n} d\left(X_{j_i}^l, X_{s_i}^l\right)^2 \to 0$.
From the First Lemma we know that X_i^l are bounded, therefore have a convergent subsequence $X_i^{a_l}$ with the common limit A, by the above argument. Suppose we have aother subsequence $X_i^{b_l}$ with common limit B. Then by the Third Lemma

$$\left[\frac{\sum_{i=1}^{n}\left(X_i^{l+1}\right)^2}{n}\right]^{1/2} \leq \left[\frac{\sum_{i=1}^{n}\left(X_i^{l}\right)^2}{n}\right]^{1/2} \tag{2.9}$$

so we have $A \leq B$ and also $B \leq A$ therefore $A = B$. $\square$

Properties of the Iterative extension

The limit point may depend on the sequence of graphs $G = \{G^l\}_{l=1,2,\dots}$ so we denote the limit with $M_G(X_1, \cdots X_n)$.

Definition (Extended matrix mean)

An n-variable function $M_n : P(r)^n \to P(r)$ is a *mean* function if

(I) $M_n(X, \dots, X) = X$ for every $X \in P(r)$,

(II) $\min(X_1, \dots, X_n) \leq M_n(X_1, \dots, X_n) \leq \max(X_1, \dots, X_n)$ if min and max exists,

(III) If $a_i \leq A_i$, then $M_n(a_1, \dots, a_n) \leq M_n(A_1, \dots, A_n)$,

(IV) M_n is continuous,

(V) $M_n(CX_1C^*, \dots, CX_nC^*) = CM(X_1, \dots, X_n)C^*$ for all invertible C.

Proposition

The limit point $M_G(X_1, \ldots, X_n)$ satisfies (I), (II) and (III) with respect to G.

Proposition

If $M(A,B) \le N(A,B) \le \left[(A^2+B^2)/2\right]^{1/2}$ are matrix means, then $M_G(X_1, \ldots, X_n) \le N_G(X_1, \ldots, X_n)$ with respect to G.

Proposition

The limit point $M_G(X_1, \ldots, X_n)$ satisfies property (V) with respect to G.

Proposition

The limit point $M_G(X_1, \ldots, X_n)$ is continuous with respect to G.

Proof.

The proof is only based on monotonicity (III) and the transformer equality (V). □

Numerical properties of the Iterative mean

- Iterative mean is directly based on the two-mean $M(A, B)$ so it is not recursive in n.
- Convergence rate of the Iterative mean is linear.
- Convergence rate can significantly be improved if we maximize e_l at every iteration step. This is equvalent to the travelling salesman problem (STM).
- Iterative mean appears to depend on the infinite sequence of graphs G so it is not permutation invariant.
- In the trivial cases Iterative mean iteration always converges to the known n-means (arithmetic, harmonic, in the scalar case geometric a well)

Ando-Li-Mathias mean

Definition (ALM iteration)

Let $X = (X_1^0 \ldots X_n^0)$ where $X_i^0 \in P(r)$ and define the iteration

$$X_i^{l+1} = M\left(Z_{\neq i}\left(X_1^l, \cdots, X_n^l\right)\right), \tag{2.10}$$

where $Z_{\neq i}(X_1^l, \cdots, X_n^l) = X_1^l, \cdots, X_{i-1}^l, X_{i+1}^l, \cdots, X_n^l$ and $M(X_1, \cdots, X_{n-1})$ is a well defined operation.

Theorem

In the above definition starting with any matrix mean $M(A, B) \leq \left(\frac{A^2+B^2}{2}\right)^{1/2}$ for $n = 3$, the sequences X_i^l have the same limit point which is denoted by $M(X_1, \cdots, X_3)$. This holds for all n therefore recursively defining $M(X_1, \cdots, X_n)$.

Sketch of the proof and properties

The proof is similar to the Iterative Mean case, although we have to use induction.

(I) $M_n(X, \ldots, X) = X$ for every $X \in P(r)$,

(II) $\min(X_1, \ldots, X_n) \leq M_n(X_1, \ldots, X_n) \leq \max(X_1, \ldots, X_n)$ if min and max exists,

(III) If $a_i \leq A_i$, then $M_n(a_1, \ldots, a_n) \leq M_n(A_1, \ldots, A_n)$,

(IV) M_n is continuous,

(V) $M_n(CX_1C^*, \ldots, CX_nC^*) = CM(X_1, \ldots, X_n)C^*$ for all invertible C

We recall the lemmas which need inductive arguments

Lemma (First Lemma)

The matrix sequences X_i^l are bounded for all n.

Lemma (Third Lemma)

If $M(A,B) \leq \left(\frac{A^2+B^2}{2}\right)^{1/2}$, then

$$\left[\frac{\sum_{i=1}^{n}\left(X_i^{l+1}\right)^2}{n}\right]^{1/2} \leq \left[\frac{\sum_{i=1}^{n}\left(X_i^{l}\right)^2}{n}\right]^{1/2}. \qquad (2.11)$$

Lemma (Fourth Lemma)

$M(X_1, \cdots, X_n) \leq \left[\frac{\sum_{i=1}^{n}(X_i)^2}{n}\right]^{1/2}$

The proofs of all the above goes by induction.
We will use the following notations

$$a_l^n = \sum_{i=1}^{n} d\left(0, X_i^l\right)^2$$
$$e_l^n = \sum_{i=1}^{n-1}\sum_{j=i+1}^{n} d\left(X_i^l, X_j^l\right)^2 \tag{2.12}$$

Lemma (Fifth Lemma)

$$a_{l+1}^n \leq a_l^n - \frac{k}{8} z_n e_l^n, \tag{2.13}$$

where $z_n = \frac{2}{n-1}$.

Now a similar argument as in the case of the Iterative mean proves the convergence to a common limit point for $n+1$. This argument is based on the First Lemma, the Third Lemma and the Fifth Lemma.

- Property (V) has again a similar proof to the case of the Iterative mean modified with induction.
- The proof of continuity is the same as in the case of the Iterative mean.

□

Numerical properties of the ALM mean

- Ando-Li-Mathias mean is recursive in n so computationally difficult.
- Convergence rate of the Ando-Li-Mathias mean is linear with possibly similar constant to the Iterative mean.
- Ando-Li-Mathias mean is permutation invariant.
- In the trivial cases Ando-Li-Mathias mean iteration always converges to the known n-means (arithmetic, harmonic, in the scalar case geometric a well)

Weighted two-means

$t \in [0, 1]$

Example (Weighted artihmetic mean)

$$A_t(X, Y) = (1 - t)X + tY \tag{2.14}$$

$$d_A(X, Y)^2 = Tr\left\{(X - Y)^*(X - Y)\right\} \tag{2.15}$$

Example (Weighted harmonic mean)

$$H_t(X, Y) = \left[(1 - t)X^{-1} + tY^{-1}\right]^{-1} \tag{2.16}$$

$$d_H(X, Y)^2 = Tr\left\{\left(X^{-1} - Y^{-1}\right)^*\left(X^{-1} - Y^{-1}\right)\right\} \tag{2.17}$$

Example (Weighted geometric mean)

$$G_t(X, Y) = X^{1/2} \left(X^{-1/2} Y X^{-1/2} \right)^t X^{1/2} \tag{2.18}$$

$$d_G(X, Y)^2 = Tr \left\{ \log^2 \left(X^{-1/2} Y X^{-1/2} \right) \right\} \tag{2.19}$$

- $d_A(X, Y)$ is Euclidean
- $d_H(X, Y)$ is also Euclidean
- $d_G(X, Y)$ is nonpositively curved

Uniquely geodesic metric structures automatically imply weighted means as points of geodesics.

Definition (Weighted mean $M_t(X, Y)$)

Let $M(\cdot,\cdot)$ be a symmetric matrix mean, $X, Y \in P(r)$ and $t \in [0, 1]$ and $a = 0, b = 1$.
Algorithm:
$A := X, B := Y$

1. If $a = t$ then $M_t(X, Y) := A, B := A$, return.
2. If $b = t$ then $M_t(X, Y) := B, A := B$, return.
3. If $\frac{a+b}{2} \leq t$ then $a := \frac{a+b}{2}, A := M(A, B)$ go to step 1.
4. If $\frac{a+b}{2} > t$ then $b := \frac{a+b}{2}, B := M(A, B)$ go to step 1.

We denote the limit of this procedure with $M_t(X, Y)$.

Properties of $M_t(A, B)$

$t \in [0, 1]$

(i) $M_t(A, A) = A$ for every $A \in P(n)$,

(ii) If $A < B$, then $A < M_t(A, B) < B$ for $t \in (0, 1)$,

(iii) If $A < A'$ and $B < B'$, then $M_t(A, B) < M_t(A', B')$,

(iv) $M_{1/2}(A, B) = M(A, B)$.

(v) In the case of arithmetic, geometric, harmonic mean we get back the corresponding weighted means.

(vi) If $N(A, B) \leq M(A, B)$ then $N_t(A, B) \leq M_t(A, B)$.

(vii) $M_t(A, B)$ is continuous in A and B,

(viii) $M_t(CAC^*, CBC^*) = CM_t(A, B)C^*$ for all invertible C.

Bini-Meini-Poloni mean

Definition (BMP iteration)

Let $X = (X_1^0 \dots X_n^0)$ where $X_i^0 \in P(r)$ and define the iteration

$$X_i^{l+1} = M_{\frac{n-1}{n}} \left(X_i^l, M \left(Z_{\neq i} \left(X_1^l, \cdots, X_n^l \right) \right) \right), \qquad (2.20)$$

where $Z_{\neq i}(X_1^l, \cdots, X_n^l) = X_1^l, \cdots, X_{i-1}^l, X_{i+1}^l, \cdots, X_n^l$ and $M(X_1, \cdots, X_{n-1})$ is a well defined operation.

Theorem

In the above definition starting with any matrix mean $M(A, B) \leq \left(\frac{A^2+B^2}{2}\right)^{1/2}$ for $n = 3$, the sequences X_i^l have the same limit point which is denoted by $M(X_1, \cdots, X_3)$. This holds for all n therefore recursively defining $M(X_1, \cdots, X_n)$.

BMP for the quasi-arithmetic mean

Lemma (Invariance Lemma)

$\left[\sum_{i=1}^{n} \frac{X_i^2}{n}\right]^{1/2}$ *is invariant under one BMP iteration step.*

Proof.

After substitution we get

$$\left[\sum_{i=1}^{n} \frac{\frac{X_i^2}{n} + \frac{n-1}{n} \sum_{j=1, j\neq i}^{n} \frac{X_j^2}{n-1}}{n}\right]^{1/2} = \left[\sum_{i=1}^{n} \frac{X_i^2}{n}\right]^{1/2}. \qquad (2.21)$$

□

$$d(A,B)^2 = Tr\left\{(A-B)^*(A-B)\right\} \tag{2.22}$$

Lemma (Second Lemma)

For the above distance and for any matrix mean $M_t(A,B) \leq \left[((1-t)A+tB)^2 + \epsilon t(1-t)(A-B)^2\right]^{1/2}$ *where* $\epsilon < 1$ *the following holds*

$$d(0, M_t(A,B))^2 \leq (1-t)d(0,A)^2 + td(0,B)^2 - \frac{k}{2}t(1-t)d(A,B)^2, \tag{2.23}$$

for $k = 2-2\epsilon$*.*

Proof.

By similar computation. □

It is also enough if

$$M(A, B) \leq \left[\left(\frac{A+B}{2} \right)^2 + \epsilon \left(\frac{A-B}{2} \right)^2 \right]^{1/2}. \tag{2.24}$$

Lemma (First Lemma)

The matrix sequences X_i^l are bounded for all n.

Lemma (Third Lemma)

If $M(A, B) \leq \left(\frac{A^2+B^2}{2} \right)^{1/2}$, then

$$\left[\frac{\sum_{i=1}^n \left(X_i^{l+1} \right)^2}{n} \right]^{1/2} \leq \left[\frac{\sum_{i=1}^n \left(X_i^l \right)^2}{n} \right]^{1/2}. \tag{2.25}$$

Lemma (Fourth Lemma)

$$M(X_1, \cdots, X_n) \leq \left[\frac{\sum_{i=1}^{n} (X_i)^2}{n}\right]^{1/2}$$

Lemma (Properties)

(I) *$M_n(X, \ldots, X) = X$ for every $X \in P(r)$,*

(II) $\min(X_1, \ldots, X_n) \leq M_n(X_1, \ldots, X_n) \leq \max(X_1, \ldots, X_n)$ *if* min *and* max *exists,*

(III) *If $a_i \leq A_i$, then $M_n(a_1, \ldots, a_n) \leq M_n(A_1, \ldots, A_n)$,*

(IV) *M_n is continuous,*

(V) *$M_n(CX_1C^*, \ldots, CX_nC^*) = CM(X_1, \ldots, X_n)C^*$ for all invertible C.*

We will use the following notations

$$a_l^n = \sum_{i=1}^{n} d\left(0, X_i^l\right)^2$$
$$e_l^n = \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} d\left(X_i^l, X_j^l\right)^2 \tag{2.26}$$

Lemma (Fifth Lemma)

$$a_{l+1}^n \leq a_l^n - \frac{k}{8} z_n e_l^n, \tag{2.27}$$

where $z_n = \frac{4}{(n-1)n}$.

The proofs of all the above and the assertion are by induction.

Now a similar argument as in the case of the Iterative mean proves the convergence to a common limit point for $n+1$. This argument is based on the First Lemma, the Third Lemma and the Fifth Lemma again.

- Property (V) has again a similar proof to the case of the Iterative mean modified with induction.
- The proof of continuity is the same as in the case of the Iterative mean.

□

Numerical properties of the BMP mean

- Bini-Meini-Poloni mean is recursive in n so computationally also difficult.
- Convergence rate of the Bini-Meini-Poloni mean is cubic in a neighborhood of the limit point.
- Bini-Meini-Poloni mean is permutation invariant.
- In the trivial cases Bini-Meini-Poloni mean iteration always converges to the known n-means (arithmetic, harmonic, in the scalar case geometric a well)

Metric geometry and C_k-domains

Let (X, d) denote a complete metric space.

Definition

We call a curve $\gamma : I \mapsto X$, where I is a real interval, a minimizing geodesic or geodesic if $d(\gamma(t_1), \gamma(t_2)) = |t_1 - t_2|$ for all $t_1, t_2 \in I$.

Definition

(X, d) is a geodesic length space if there is a minimizing geodesic $\gamma(t)$ for every $x, y \in X$ such that $d(x, y) = L(\gamma(t))$, where $L(\gamma(t))$ denotes the length of the curve $\gamma(t)$.

C_k-domains

Definition

Let $k \in (0, 2]$.

- An open set U in a geodesic metric space (X, d) is called a a C_k-domain for k if for any three points x, y, z, any minimal geodesic $\gamma : [0, 1] \mapsto X$ between x, y and for all $t \in [0, 1]$ we have

$$d(z, \gamma(t))^2 \leq (1 - t)d(z, x)^2 + td(z, y)^2 - \frac{k}{2}t(1 - t)d(x, y)^2. \tag{3.1}$$

- A geodesic metric space (X, d) is k-convex if it is itself a C_k-domain.
- A geodesic metric space (X, d) is locally k-convex if every point in X is contained in a C_k-domain.

Properties of C_k-domains

Lemma

If an open ball $B(x, r)$ is a C_k-domain then for any two points in $B(x, r)$ a minimal geodesic is unique between them. In particular any two points in a C_k-domain are connected by a unique minimal geodesic.

Lemma

Let S be a bounded subset of a C_k-domain. Then there exists a unique closed ball with minimal radius r containing S.

Lemma

Let $B(x, r) \subset D$ where D is a C_k-domain. Then $B(x, r)$ is a geodesically convex set which means that every geodesic which connects two points in $B(x, r)$ are also a subset of $B(x, r)$.

Definition (Geodesic convex hull)

The geodesic convex hull $GCH(S)$ is the intersection of all convex sets containing S.

Proposition

$GCH(S)$ can be obtained as $GCH(S) = \bigcup_{n\geq 0} F_n$, where $F_0 = S$ and for $n \geq 1$ the set F_n consists of all points which lie on geodesics with starting and ending points in F_{n-1}.

What follows is that the geodesic convex hull of a bounded set is contained in a convex metric ball.

Iterative mean in C_k-domains

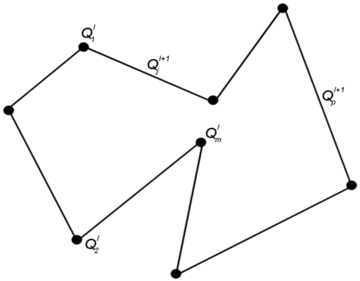
Theorem

Let (X, d) be a complete k-convex geodesic metric space. Let $Q_1^0, \dots, Q_n^0 \in X$. Set up the iteration of the Iterative mean on these points with respect to an infinite sequence of graphs $G = \{G^0, G^1, \dots\}$ where the mean $M(A, B) := \gamma_{A,B}(1/2)$. Then the sequences Q_i^l are converging to a common limit point.

Proof.

A similar inequality follows in this case, although it holds for all $x \in X$

$$\underbrace{\sum_{i=1}^{n} d\left(x, Q_i^{l+1}\right)^2}_{a_{l+1}(x)} \leq \underbrace{\sum_{i=1}^{n} d\left(x, Q_i^{l}\right)^2}_{a_l(x)} - \frac{k}{8} \underbrace{\sum_{i=1}^{n} d\left(Q_{j_i}^{l}, Q_{m_i}^{l}\right)^2}_{e_l}. \quad (3.2)$$



So we have again that

$$a_{l+1}(x) \leq a_l(x) - \frac{k}{8} e_l. \tag{3.3}$$

This again tells us that $e_l \to 0$ but in this case we do not need the subsequence argument to show convergence, it follows from completeness. $\square$

Theorem

Let (X, d) be a complete k-convex geodesic metric space. Let $Q_1^0, \ldots, Q_n^0 \in X$. Let R denote the Iterative mean of the points. Then

$$\frac{a_{l+1}(R)}{a_l(R)} \leq 1 - \frac{k}{2n^2}. \tag{3.4}$$

Proof.

By triangle inequality again

$$\sum_{i=1}^{n} d\left(Q_{j_i}^l, Q_{m_i}^l\right) \geq 2 \max_{1 \leq p,q \leq n} d\left(Q_p^l, Q_q^l\right). \tag{3.5}$$

Using the Cauchy-Schwarz inequality

$$\sum_{i=1}^{n} d\left(Q_{j_i}^{l}, Q_{m_i}^{l}\right) \leq \sqrt{n}\sqrt{\sum_{i=1}^{n} d\left(Q_{j_i}^{l}, Q_{m_i}^{l}\right)^2} = \sqrt{n}\sqrt{e_l}, \qquad (3.6)$$

hence

$$e_l \geq \frac{4}{n} \max_{1\leq p,q\leq n} d\left(Q_p^l, Q_q^l\right)^2. \qquad (3.7)$$

According to the preceding proof, for arbitrary points x, y in the geodesic convex hull $A^l = GCH\left(\left\{Q_1^l, \dots, Q_n^l\right\}\right)$

$$d(x,y)^2 \leq \max_{1\leq p,q\leq n} d\left(Q_p^l, Q_q^l\right)^2. \qquad (3.8)$$

Thus

$$\frac{e_l}{a_l(R)} \geq \frac{4}{n^2}. \qquad (3.9)$$

□

Center of mass in C_k-domains

Theorem

Let (X, d) be a complete k-convex geodesic metric space. Let $Q_1, \ldots, Q_n \in X$. Then the center of mass

$$Y = \arg\min_{x \in X} \sum_{i=1}^{n} d\left(x, Q_i\right)^2 \tag{3.10}$$

exists and is a unique point in X.

Proof.

Follows from uniqueness of geodesics, strict geodesic convexity and continuity of the function. □

Theorem

Let (X, d) be a complete k-convex geodesic metric space. Let $Q_1, \ldots, Q_n \in X$. Then

$$Y = \arg\min_{x \in X} \sum_{i=1}^{n} d\left(x, Q_i\right)^2 \tag{3.11}$$

and the limit point R fullfill the following

$$d(R, Y) \leq \sqrt{\frac{\sum_{i=1}^{n} d\left(Y, Q_i\right)^2 - \frac{k}{8} \sum_{l=1}^{\infty} e_l}{n}}. \tag{3.12}$$

Proof.

$$\frac{k}{8} e_l \leq a_l(x) - a_{l+1}(x) \tag{3.13}$$

$$\lim_{m\to\infty} a_m(x) \leq a_0(x) - \frac{k}{8}\sum_{l=1}^{\infty} e_l \tag{3.14}$$

□

Proposition

If X is a Euclidean space then the limit point R is the center of mass.

Proof.

$$a_{l+1}(x) \leq a_l(x) - \frac{k}{8} e_l \tag{3.15}$$

turns into

$$a_{l+1}(x) = a_l(x) - \frac{1}{4} e_l. \tag{3.16}$$

Minimizing both sides gives us that the centroid is left invariant by the iteration. □

The ALM and BMP process in C_k-domains

Theorem

Let (X, d) be a complete k-convex geodesic metric space. Let $Q_1^0, \ldots, Q_n^0 \in X$. Let $M_t(A, B) := \gamma_{A,B}(t)$. Then the sequences Q_i^l obtained by the ALM or BMP iteration are converging to a common limit point.

Theorem

With similar assumptions let R denote the ALM mean of the points. Then

$$\frac{a_{l+1}(R)}{a_l(R)} \leq 1 - \frac{k}{2n^2}. \tag{3.17}$$

Similarly for the BMP iteration with limit point R

$$\frac{a_{l+1}(R)}{a_l(R)} \leq 1 - \frac{k}{n^3}. \tag{3.18}$$

The geometric mean

$$G(X,Y) = X^{1/2}\left(X^{-1/2}YX^{-1/2}\right)^{1/2}X^{1/2} \tag{4.1}$$

$$\langle U,V\rangle_p = Trp^{-1}Up^{-1}V \tag{4.2}$$

$$d(X,Y)^2 = Tr\left\{\log^2\left(X^{-1/2}YX^{-1/2}\right)\right\} \tag{4.3}$$

- The above space is a Riemannian symmetric space.
- The group of isometries is $G = GL(r,\mathbb{C})$.
- The isotropy group at I is $K = U(r,\mathbb{C})$.
- G acts on $P(r) \cong GL(r,\mathbb{C})/U(r,\mathbb{C})$ as gpg^* where $g \in G$ and $p \in P(r)$.

The ALM, the BMP and the Iterative mean procedure converges for this mean. The first have linear convergence rate (1/2). The second has cubic convergence, the third one has linear convergence again $(1 - 1/n^2)$.

Definition (Riemannian mean)

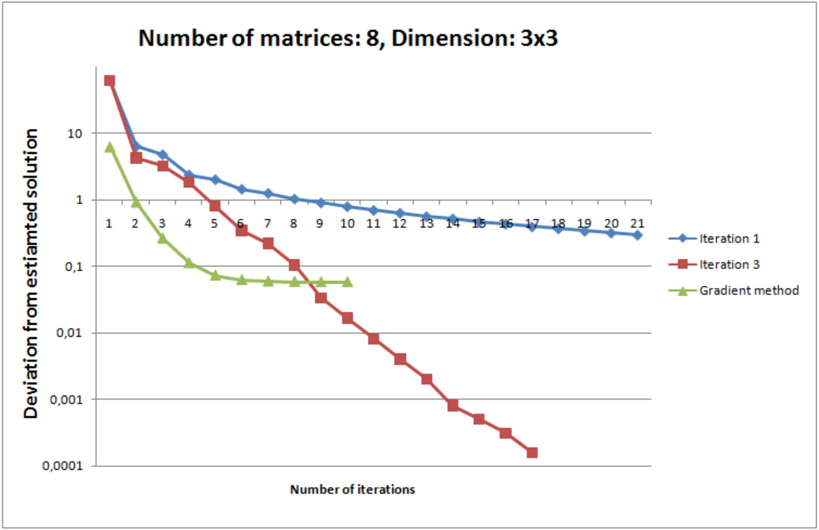
$\mathrm{R}(\mathrm{X}_1, \ldots, X_n) = \arg\min_{x \in P(r)} \sum_{i=1}^{n} d\left(x, X_i\right)^2$

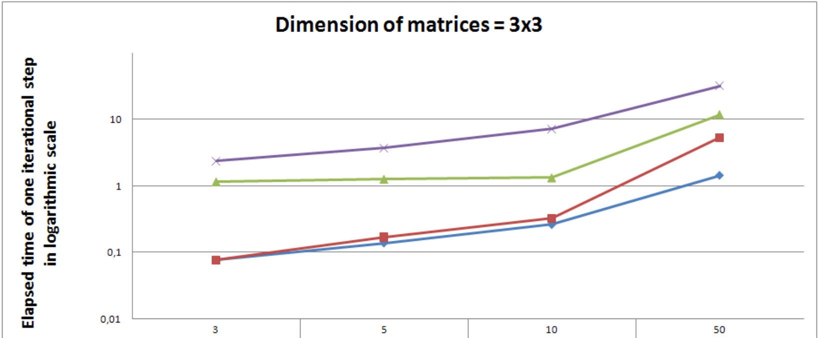
The above has unique positive definite solution characterized as

$$\sum_{i=1}^{n} \log\left(X^{-1/2} X_i X^{-1/2}\right) = 0. \tag{4.4}$$

The cost function is $f(X) = \sum_{i=1}^{n} d\left(x, X_i\right)^2$. Then the gradient is

$$\nabla f(X) = -\sum_{i=1}^{n} X^{1/2} \log\left(X^{-1/2} X_i X^{-1/2}\right) X^{1/2}. \tag{4.5}$$


Number of matrices: 8, Dimension: 3x3
Deviation from estiamted solution
10
1
0,1
0,01
0,001
0,0001
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21
Number of iterations
Iteration 1
Iteration 3
Gradient method



	3	5	10	50
Iteration 1	0,078	0,14	0,265	1,438
Iteration 3	0,078	0,172	0,328	5,329
Gradient method	1,156	1,266	1,328	11,75
Newton method	2,406	3,734	7,251	32,158