Geometries on positive-definite matrices and their relation to the power means

Bruno **Iannazzo**, Università di **Perugia**, **Italy**

with Nadia **Chouaieb** and Maher **Moakher**
Univeristy of **Tunis El Manar, Tunisia**

Roma, April 7, 2022

Outline

- A new **family of geometries** on $\mathcal{P}_n$ (positive definite matrices of size n);

Outline

- A new **family of geometries** on $\mathcal{P}_n$ (positive definite matrices of size n);
- **Explicit** geodesics and distances;

Outline

- A new **family of geometries** on $\mathcal{P}_n$ (positive definite matrices of size n);
- **Explicit** geodesics and distances;
- A new **power mean** of matrices.

Outline

- A new **family of geometries** on $\mathcal{P}_n$ (positive definite matrices of size n);
- **Explicit** geodesics and distances;
- A new **power mean** of matrices.

Why a new geometry on $\mathcal{P}_n$?

Outline

- A new **family of geometries** on $\mathcal{P}_n$ (positive definite matrices of size n);
- **Explicit** geodesics and distances;
- A new **power mean** of matrices.

Why a new geometry on $\mathcal{P}_n$?

Why on $\mathcal{P}_n$? How is this related to scientific computing?

Geometry of $\mathcal{P}_n$

The set of real symmetric matrices, $\mathbb{H}^n$, is a Euclidean space

$$\langle H, K \rangle = \mathrm{trace}(HK).$$

$\mathcal{P}_n$ is an **open cone** in $\mathbb{H}^n$

Geometry of $\mathcal{P}_n$

The set of real symmetric matrices, $\mathbb{H}^n$, is a Euclidean space

$$\langle H, K \rangle = \mathrm{trace}(HK).$$

$\mathcal{P}_n$ is an **open cone** in $\mathbb{H}^n$ and it is **convex**.

We can treat it as an open subset of $\mathbb{R}^n$.

Why a new geometry on $\mathcal{P}_n$?

Geometry of $\mathcal{P}_n$

The set of real symmetric matrices, $\mathbb{H}^n$, is a Euclidean space

$$\langle H, K \rangle = \mathrm{trace}(HK).$$

$\mathcal{P}_n$ is an **open cone** in $\mathbb{H}^n$ and it is **convex**.

We can treat it as an open subset of $\mathbb{R}^n$.

Why a new geometry on $\mathcal{P}_n$?

Euclidean geometry is not always the best model.

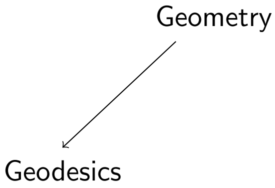
Why a new geometry on $\mathcal{P}_n$?

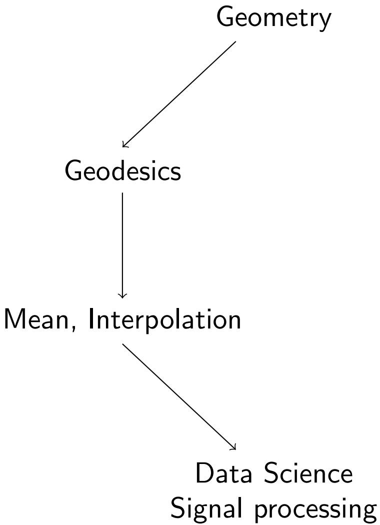
Euclidean geometry is not always the best model.

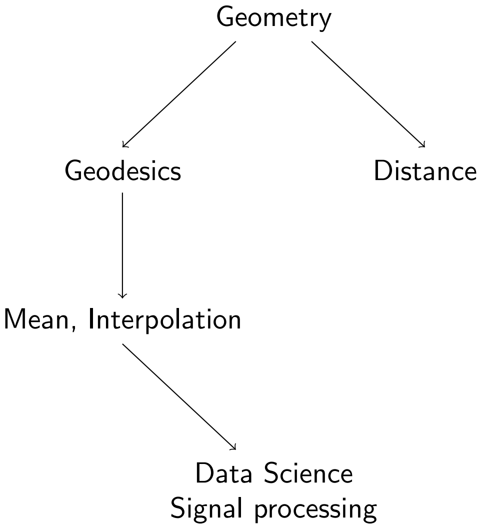
But first...

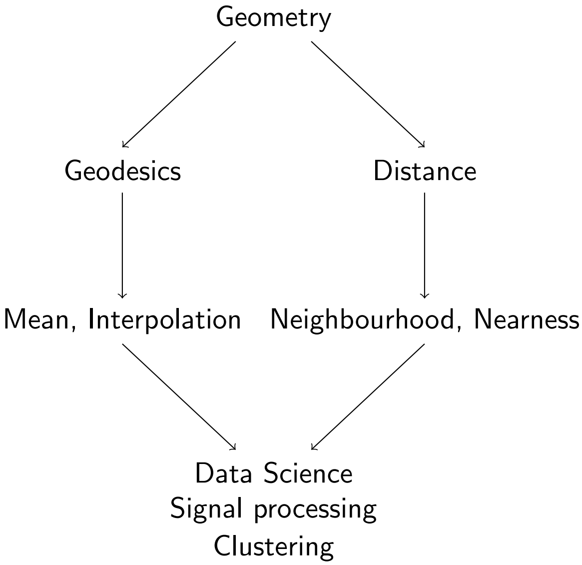
Why on $\mathcal{P}_n$? How is this related to scientific computing?

Geometry


Geometry
Geodesics


Geometry
Geodesics
Mean, Interpolation
Data Science
Signal processing


Geometry
Geodesics
Distance
Mean, Interpolation
Data Science
Signal processing



A new geometry with explicit geodesics and distances related to a common mean may be useful

A non-Euclidean geometry on $\mathcal{P}_n$

The most celebrated non-Euclidean geometry, the **affine-invariant** geometry, on $\mathcal{P}_n$ arises in

- Geometry [Lang, Fundamentals of Differential Geometry, Ch. XII, '99];
- Optimization [Nesterov-Todd, '02];
- Information Geometry [Ohara-Suda-Amari, '96];
- Matrix Analysis [Lawson-Lim] (with an eye to functional analysis).

A non-Euclidean geometry on $\mathcal{P}_n$

Define the (convex) self-concordant barrier for $\mathcal{P}_n$

$$\varphi(X) = -\log\det(X).$$

The Hessian

$$D^2\varphi(X)[H, K] = \mathrm{trace}(X^{-1}HX^{-1}K),$$

is a **scalar product** on $\mathbb{H}^n$.

Defines a **Riemannian geometry** on $\mathbb{P}_n$.

Riemannian geometry

Riemannian geometry on $\mathcal{M}$:

A scalar product g_X on the tangent space $T_x\mathcal{M}$ that smoothly varies as X.

For our case we **do not need** abstraction.

Riemannian geometry on $\mathcal{P}_n$

- The tangent space to $\mathcal{P}_n$ is $\mathbb{H}^n$ ($T\mathbb{H}^n_x \cong \mathbb{H}^n \cong \mathbb{R}^{n(n+1)/2}$);
- $\mathcal{P}_n \subset \mathbb{H}^n$ and the inclusion is a chart (we need only one chart);
- A Riemannian geometry on $\mathcal{P}_n$ is a smooth function $f : \mathcal{P}_n \to \mathcal{P}_{n^2}$.

Examples:

Riemannian geometry on $\mathcal{P}_n$

- The tangent space to $\mathcal{P}_n$ is $\mathbb{H}^n$ ($T\mathbb{H}^n_x \cong \mathbb{H}^n \cong \mathbb{R}^{n(n+1)/2}$);
- $\mathcal{P}_n \subset \mathbb{H}^n$ and the inclusion is a chart (we need only one chart);
- A Riemannian geometry on $\mathcal{P}_n$ is a smooth function $f : \mathcal{P}_n \to \mathcal{P}_{n^2}$.

Examples:

- The affine-invariant geometry $X \to X^{-1} \otimes X^{-1}$;

Riemannian geometry on $\mathcal{P}_n$

- The tangent space to $\mathcal{P}_n$ is $\mathbb{H}^n$ ($T\mathbb{H}^n_x \cong \mathbb{H}^n \cong \mathbb{R}^{n(n+1)/2}$);
- $\mathcal{P}_n \subset \mathbb{H}^n$ and the inclusion is a chart (we need only one chart);
- A Riemannian geometry on $\mathcal{P}_n$ is a smooth function $f : \mathcal{P}_n \to \mathcal{P}_{n^2}$.

Examples:

- The affine-invariant geometry $X \to X^{-1} \otimes X^{-1}$;
- The Euclidean geometry $X \to I$;
- Bures-Wasserstein geometry $X \to (I \otimes X^{1/2} + X \otimes X^{-1/2})^2$ (optimal transport).

The affine-invariant geometry and the geometric mean

The resulting geometry has explicit geodesics

$$\gamma(t) = A(A^{-1}B)^t =: A\#_t B, \qquad t \in [0,1],$$

the (weighted) **matrix geometric mean**.

The affine-invariant geometry and the geometric mean

The resulting geometry has explicit geodesics

$$\gamma(t) = A(A^{-1}B)^t =: A\#_t B, \qquad t \in [0,1],$$

the (weighted) **matrix geometric mean**.

Explicit distance, the trace metric

$$\delta(A,B) = \|\log(A^{-1/2}BA^{-1/2})\|_F = \Big(\sum_{\mu\in\sigma(A-\mu B)} \log^2(\mu)\Big)^{1/2}.$$

More geometric properties (worth of a chapter in Lang's book)

- Cartan-Hadamard (compare the Poincaré disk);
- Symmetric space.

Computational problems

- Compute $A \#_t B$ (Cholesky factorization, Schur form);

Computational problems

- Compute $A \#_t B$ (Cholesky factorization, Schur form);
- Compute $(A \#_t B)v$ (Rational Krylov subspaces) [Fasi, I., 16];

Computational problems

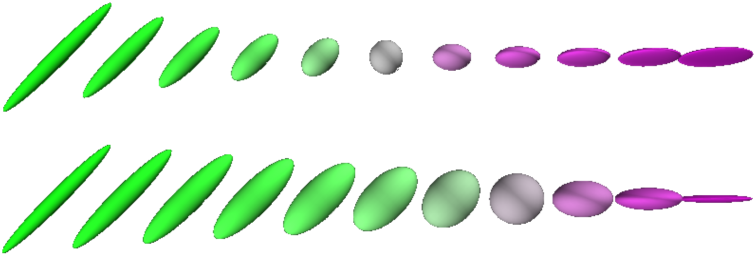
- Compute $A \#_t B$ (Cholesky factorization, Schur form);
- Compute $(A \#_t B)v$ (Rational Krylov subspaces) [Fasi, I., 16];
- Compute the geometric mean of $A_1, \ldots, A_m$

$$\operatorname{argmin}_X \sum_{i=1}^{m} \delta(A_i, X),$$

  (Riemannian Barzilai-Borwein [Porcelli, I., 18]), (Riemannian LBFGS [Absil et al., 21]);
- Compute means with further structures or with (quasi-)Toeplitz operators.

An application of the geometry

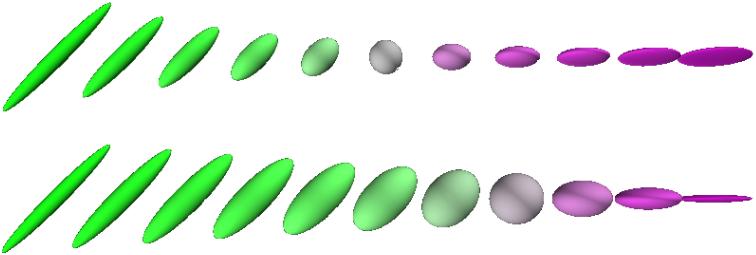
DTI problems: affine-invariant vs. Euclidean interpolation



[Moakher et al., 04], [I., Jeuris, Pompili, 19].

An application of the geometry

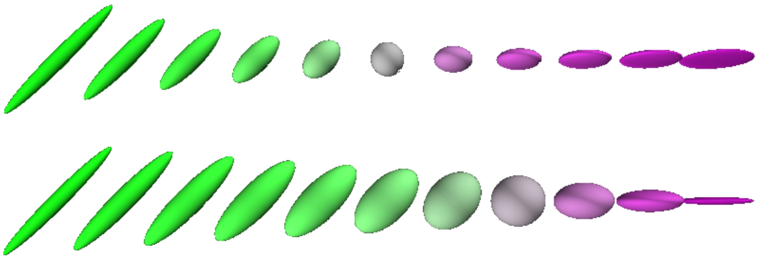
DTI problems: affine-invariant vs. Euclidean interpolation



Euclidean interpolation (below), $(1-t)A + tB$ (swelling effect).

An application of the geometry

DTI problems: affine-invariant vs. Euclidean interpolation



Affine-invariant interpolation (above),
$A\#_t B \to \det(A\#_t B) = (1-t)\det(A) + t\det(B)$.

Applications

Why new means and new distances?

Applications

Why new means and new distances?

New tools for engineers and scientists

The power potential

The **power potential**

$$\varphi_\beta(X) = \frac{1 - \det(X)^\beta}{\beta}$$

is such that

$$\lim_{\beta \to 0} \varphi_\beta(X) = -\log \det(X).$$

Also known in Tsallis statistics as q-logarithm, with $q = 1 - \beta \Rightarrow$ potential application

Riemannian geometry from the power potential

For $H, K \in \mathbb{H}^n$ the derivative $D^2\varphi_\beta(X)[H, K]$ is

$$\underbrace{\det(X)^\beta \big(\operatorname{trace}(X^{-1}HX^{-1}K) - \beta \operatorname{trace}(X^{-1}H) \operatorname{trace}(X^{-1}K)\big)}_{g_X^\beta(H,K)}$$

that is positive definite for $\beta \in (-\infty, 0) \cup (0, 1/n)$.

In our notation, we get **a family of Riemannian geometries** on $\mathcal{P}_n$

$$X \to \det(X)^\beta (X^{-1} \otimes X^{-1} - \beta \operatorname{vec}(X^{-1}) \operatorname{vec}(X^{-1})^T).$$

Conformal to a rank-one modification of the affine invariant geometry $(X^{-1} \otimes X^{-1})$.

First problem: find geodesics

A **geodesic** between A and B is a (smooth) curve $\gamma : [0,1] \to \mathcal{P}_n$ such that $\gamma(0) = A$, $\gamma(1) = B$ and

$$\mathcal{L}(\gamma) = \int_0^1 \sqrt{g^{\beta}_{\gamma(t)}(\gamma'(t), \gamma'(t))} dt$$

is minimum.

- Reduce the problem to $A = I$, $B = D$;
- Prove that the geodesic between diagonal matrices is diagonal;
- Reduce the variational problem to a BVP;
- Solve the BVP.

Reduce the problem to diagonal matrices

There exists M such that $M^T AM = I$ and $M^T BM = D$.

If $\gamma(t)$ is such that $\gamma(0) = A$ and $\gamma(1) = B$, then the curve $\varphi(t) := M^T \gamma(t) M$ joins I with D and

$$\mathcal{L}(\varphi(t)) = |\det(M)|^{\beta} \mathcal{L}(\gamma(t)).$$

$\gamma(t)$ is a geodesic from A to $B \iff \varphi(t)$ is a geodesic from I to D

Reduce the problem to diagonal matrices

An isometry on the Riemannian manifold $\mathcal{M}$ is a function $f : \mathcal{M} \to \mathcal{M}$ such that

$$g_{f(X)}(df(X)[H], df(X)[K]) = g_X(H, K)$$

Fixed points of any set of isometries are **totally geodesic** submanifolds of $\mathcal{M}$

Reduce the problem to diagonal matrices

An isometry on the Riemannian manifold $\mathcal{M}$ is a function $f : \mathcal{M} \to \mathcal{M}$ such that

$$g_{f(X)}(df(X)[H], df(X)[K]) = g_X(H, K)$$

Fixed points of any set of isometries are **totally geodesic** submanifolds of $\mathcal{M}$

If M is such that $\det(M) = \pm 1$, then $f : X \to MXM^T$ is an **isometry** of $\mathcal{P}_n$ with the metric g_X^β ($X \in \mathcal{P}_n$, $A, B \in \mathcal{H}_n$)

$$g^\beta_{MXM^T}(MAM^T, MBM^T) = (\det(M)^2)^\beta g_X^\beta(A, B)$$

Reduce the problem to diagonal matrices

An isometry on the Riemannian manifold $\mathcal{M}$ is a function $f : \mathcal{M} \to \mathcal{M}$ such that

$$g_{f(X)}(df(X)[H], df(X)[K]) = g_X(H, K)$$

Fixed points of any set of isometries are **totally geodesic** submanifolds of $\mathcal{M}$

If M is such that $\det(M) = \pm 1$, then $f : X \to MXM^T$ is an **isometry** of $\mathcal{P}_n$ with the metric g_X^β ($X \in \mathcal{P}_n$, $A, B \in \mathcal{H}_n$)

$$g^\beta_{MXM^T}(MAM^T, MBM^T) = (\det(M)^2)^\beta g_X^\beta(A, B)$$

Reduce the problem to diagonal matrices

An isometry on the Riemannian manifold $\mathcal{M}$ is a function $f : \mathcal{M} \to \mathcal{M}$ such that

$$g_{f(X)}(df(X)[H], df(X)[K]) = g_X(H, K)$$

Fixed points of any set of isometries are **totally geodesic** submanifolds of $\mathcal{M}$

If M is such that $\det(M) = \pm 1$, then $f : X \to MXM^T$ is an **isometry** of $\mathcal{P}_n$ with the metric g_X^β ($X \in \mathcal{P}_n$, $A, B \in \mathcal{H}_n$)

$$g^\beta_{MXM^T}(MAM^T, MBM^T) = (\det(M)^2)^\beta g_X^\beta(A, B)$$

Positive-definite **diagonal** matrices are a totally geodesic submanifold $\Rightarrow$ the geodesic between I and D is made of diagonal matrices

We can find in this way other totally geodesic submanifold such as the positive multiples of a matrix. (Similar results for g_X^0).

Solve the variational equation

The **Euler-Lagrange equation** gives the equivalent **BVP**

$$\begin{cases} \alpha' = -n\beta\left(\frac{1}{2}\alpha^2 - \frac{1}{n(1-n\beta)}\sum_{i=1}^n \nu_i^2\right), \\ \nu' = -n\beta\alpha\nu_i, \quad i = 1, \dots, n, \\ \nu_1 + \cdots + \nu_n = 0, \\ \lambda_i'/\lambda_i = \nu_i + \alpha, \quad i = 1, \dots, n, \\ \lambda_i(0) = 1, \quad \lambda_i(1) = d_i, \quad i = 1, \dots, n, \end{cases}$$

with $D = \mathrm{diag}(d_1, \dots, d_n)$

This is a **Riccati** (differential not algebraic) equation

Solve the variational equation

The **Euler-Lagrange equation** gives the equivalent **BVP**

$$\begin{cases} \alpha' = -n\beta\left(\frac{1}{2}\alpha^2 - \frac{1}{n(1-n\beta)}\sum_{i=1}^{n}\nu_i^2\right), \\ \nu' = -n\beta\alpha\nu_i, \quad i = 1, \ldots, n, \\ \nu_1 + \cdots + \nu_n = 0, \\ \lambda_i'/\lambda_i = \nu_i + \alpha, \quad i = 1, \ldots, n, \\ \lambda_i(0) = 1, \quad \lambda_i(1) = d_i, \quad i = 1, \ldots, n, \end{cases}$$

with $D = \operatorname{diag}(d_1, \ldots, d_n)$

This is a **Riccati** (differential not algebraic) equation

Commercial programs did not find the **explicit solution**,

Solve the variational equation

The **Euler-Lagrange equation** gives the equivalent **BVP**

$$\begin{cases} \alpha' = -n\beta\left(\frac{1}{2}\alpha^2 - \frac{1}{n(1-n\beta)}\sum_{i=1}^n \nu_i^2\right), \\ \nu' = -n\beta\alpha\nu_i, \quad i = 1, \ldots, n, \\ \nu_1 + \cdots + \nu_n = 0, \\ \lambda_i'/\lambda_i = \nu_i + \alpha, \quad i = 1, \ldots, n, \\ \lambda_i(0) = 1, \quad \lambda_i(1) = d_i, \quad i = 1, \ldots, n, \end{cases}$$

with $D = \mathrm{diag}(d_1, \ldots, d_n)$

This is a **Riccati** (differential not algebraic) equation

Commercial programs did not find the **explicit solution**,

but we were able to find it.

Special cases: positive numbers

The geodesic (with arc length parametrization) joining $a, b \in \mathcal{P}_1$ is

$$G_\beta(a, b; t) = \left((1-t)a^{\beta/2} + tb^{\beta/2}\right)^{2/\beta}, \qquad t \in [0, 1],$$

- It is the **weighted power mean** of a and b;

Special cases: positive numbers

The geodesic (with arc length parametrization) joining $a, b \in \mathcal{P}_1$ is

$$G_\beta(a, b; t) = \left((1-t)a^{\beta/2} + tb^{\beta/2}\right)^{2/\beta}, \qquad t \in [0, 1],$$

- It is the **weighted power mean** of a and b;
- Suggests that for matrices it might be a **power mean of matrices** in $\mathcal{P}_n$;
- Mathematical **curiosity**: interesting per se.

Special cases: a ray

Let A and B be **linearly dependent** in $\mathcal{P}_n$ then

$$G_\beta(A, B; t) = \big((1-t)A^{n\beta/2} + tB^{n\beta/2}\big)^{2/(n\beta)}.$$

Still a **"power mean"** with parameter $n\beta/2$.

We will show that also in the general case this is a power mean with parameter $n\beta/2$

The general case

Theorem

Let $A, B \in \mathcal{P}_n$ linearly independent and $\beta \in (\beta_1, 0) \cup (0, \beta_2)$. There exists a unique geodesic joining A and B given by

$$G_\beta(A, B; t) = \eta(t)(A \#_{\alpha(t)} B) = \eta(t) A (A^{-1} B)^{\alpha(t)}, \quad t \in [0, 1],$$

where

$$\alpha(t) = \frac{1}{\gamma} \arctan\Big(\frac{t\sigma \sin\gamma}{1 - t + t\sigma\cos\gamma} \Big),$$

$$\eta(t) = \Big(\frac{(1-t)^2 + 2t(1-t)\sigma\cos\gamma + t^2\sigma^2}{\sigma^{2\alpha(t)}} \Big)^{1/(n\beta)},$$

with $\sigma = \det(A^{-1}B)^{\beta/2}$ *and* $\gamma = \frac{|\beta| \delta(A/\det(A)^{1/n}, B/\det(B)^{1/n})}{2\sqrt{1/n - \beta}}$

The general case

We introduce a **measure of linear independence**

$$\gamma_\beta(A, B) := \frac{|\beta|\delta(\widetilde{A}, \widetilde{B})}{2\sqrt{1/n - \beta}},$$

where $\widetilde{A} = A/\det(A)^{1/n}$ and $\widetilde{B} = B/\det(B)^{1/n}$.

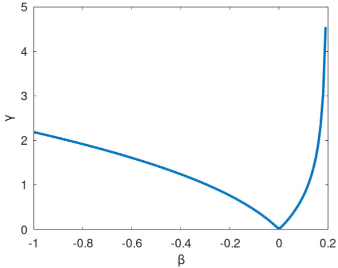
The general case

We introduce a **measure of linear independence**

$$\gamma_\beta(A,B) := \frac{|\beta|\delta(\widetilde{A},\widetilde{B})}{2\sqrt{1/n-\beta}},$$

where $\widetilde{A} = A/\det(A)^{1/n}$ and $\widetilde{B} = B/\det(B)^{1/n}$.

γ_β is 0 if and only if and only if A and B are linearly dependent.



$\beta \in (\beta_1, 0) \cup (0, \beta_2) \iff 0 < \gamma < \pi/2.$

The geodesic

For $0 < \gamma < \pi/2$

$$G_\beta(A, B; t) = \eta(t)(A \#_{\alpha(t)} B) = \eta(t) A (A^{-1} B)^{\alpha(t)}, \quad t \in [0, 1],$$

Can be extended to $\gamma < \pi$, but **not further**.

$$A = \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix}, \qquad B = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}.$$

The geodesic

For $0 < \gamma < \pi/2$

$$G_\beta(A, B; t) = \eta(t)(A \#_{\alpha(t)} B) = \eta(t) A (A^{-1} B)^{\alpha(t)}, \quad t \in [0, 1],$$

Can be extended to $\gamma < \pi$, but **not further**.

$$A = \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix}, \qquad B = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}.$$

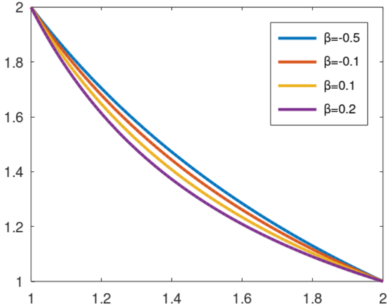
The geodesic

For $0 < \gamma < \pi/2$

$$G_\beta(A, B; t) = \eta(t)(A\#_{\alpha(t)}B) = \eta(t)A(A^{-1}B)^{\alpha(t)}, \quad t \in [0, 1],$$

Can be extended to $\gamma < \pi$, but **not further**.

$$A = \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix}, \qquad B = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}.$$



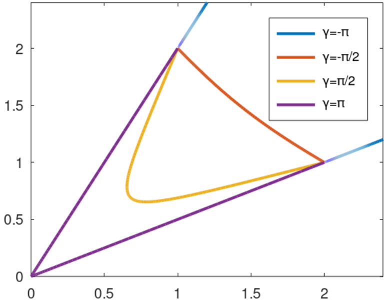
The geodesic

For $0 < \gamma < \pi/2$

$$G_\beta(A, B; t) = \eta(t)(A\#_{\alpha(t)}B) = \eta(t)A(A^{-1}B)^{\alpha(t)}, \quad t \in [0, 1],$$

Can be extended to $\gamma < \pi$, but **not further**.

$$A = \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix}, \qquad B = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}.$$



Properties

For a given couple there exist $\beta_1 < 0 < \beta_2$ such that

$$G_\beta(A, B; t), \qquad t \in [0, 1]$$

exists.

For $\beta \to 0$ **converges** to the weighted geometric mean

$$\lim_{\beta \to 0} G_\beta(A, B; t) = A \#_t B$$

Properties

For a given couple there exist $\beta_1 < 0 < \beta_2$ such that

$$G_\beta(A, B; t), \qquad t \in [0, 1]$$

exists.

For $\beta \to 0$ **converges** to the weighted geometric mean

$$\lim_{\beta \to 0} G_\beta(A, B; t) = A \#_t B$$

For a given β, the mean exists for $\gamma < \pi$, for matrices not "greatly independent".

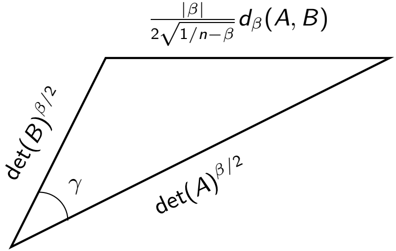
Analysis shows that the space is **not complete**.

Experiments suggest **negative curvature**.

Distance

The distance associated with g_β between A and B with $\gamma < \pi$

$$\underbrace{\frac{2\sqrt{1/n-\beta}}{|\beta|}\left((\det(A)^{\frac{\beta}{2}} - det(B)^{\frac{\beta}{2}})^2 + 4(\det(A)\det(B))^{\frac{\beta}{2}}\sin^2\frac{\gamma}{2}\right)^{\frac{1}{2}}}_{d_\beta(A,B)}$$



Distance

When $\det(A) = \det(B) = \Delta$,

$$d_\beta(A, B) = \frac{4\sqrt{1/n - \beta}}{|\beta|} \Delta^{\beta/2} \sin \frac{\gamma}{2}.$$

Moreover,

$$\lim_{\beta \to 0} d_\beta(A, B) = \delta(A, B).$$

It generalizes the geometric mean distance.

The power mean

The geodesic can be seen as a weighted power mean of positive definite matrices with parameter $p = n\beta/2$.

Euclidean power mean

$$R_p(A, B; t) := \big((1-t)A^p + tB^p\big)^{1/p}, \qquad p = n\beta/2;$$

Lim-Pálfia power mean [Lim-Palfia '12]

$$Q_p(A, B; t) := Af(A^{-1}B), \quad f(z) = \Big((1-t)+tz^p\Big)^{1/p}, \quad p = n\beta/2.$$

are different from our mean for linearly independent matrices.

Properties

- $G_\beta(M^T AM, M^T BM; t) = M^T G_\beta(A, B; t)M$, with M invertible (**commutativity with congruences**);
- $G_\beta(A, B; t) = G_\beta(B, A; 1 - t)$ (**symmetry**);
- $G_\beta(aA, bB; t) = \left((1-t)a^{n\beta/2} + tb^{n\beta/2}\right)^{2/(n\beta)} G_\beta\left(A, B; \frac{tb^{n\beta/2}}{(1-t)a^{n\beta/2}+tb^{n\beta/2}}\right)$, for $a, b > 0$ (**homogeneity**);
- $G_\beta(A^{-1}, B^{-1}, t) = A^{-1} G_\beta(A, B, 1-t) B^{-1}$ (**inversion**).

Comparison

With $p = n\beta/2$; P_p Euclidean power mean; Q_p Lim-Palfia mean

Property	P_p	Q_p	G_β

Comparison

With $p = n\beta/2$; P_p Euclidean power mean; Q_p Lim-Palfia mean

Property	P_p	Q_p	G_β
commutativity with congruences	no	yes	yes

$$G_\beta(M^T AM, M^T BM; t) = M^T G_\beta(A, B; t)M$$

Comparison

With $p = n\beta/2$; P_p Euclidean power mean; Q_p Lim-Palfia mean

Property	P_p	Q_p	G_β
commutativity with congruences	no	yes	yes
inversion	no	yes	yes

$$G_\beta(A^{-1}, B^{-1}, t) = A^{-1} G_\beta(A, B, 1-t) B^{-1}$$

Comparison

With $p = n\beta/2$; P_p Euclidean power mean; Q_p Lim-Palfia mean

Property	P_p	Q_p	G_β
commutativity with congruences	no	yes	yes
inversion	no	yes	yes
symmetry	yes	yes	yes

$G_\beta(A, B; t) = G_\beta(B, A; 1 - t)$

Comparison

With $p = n\beta/2$; P_p Euclidean power mean; Q_p Lim-Palfia mean

Property	P_p	Q_p	G_β
commutativity with congruences	no	yes	yes
inversion	no	yes	yes
symmetry	yes	yes	yes
homogeneity	yes	yes	yes

$$G_\beta(aA, bB; t) =$$
$$\left((1-t)a^{n\beta/2} + tb^{n\beta/2}\right)^{2/(n\beta)} G_\beta\left(A, B; \frac{tb^{n\beta/2}}{(1-t)a^{n\beta/2}+tb^{n\beta/2}}\right)$$

Comparison

With $p = n\beta/2$; P_p Euclidean power mean; Q_p Lim-Palfia mean

Property	P_p	Q_p	G_β
commutativity with congruences	no	yes	yes
inversion	no	yes	yes
symmetry	yes	yes	yes
homogeneity	yes	yes	yes
consistency with scalars	yes	yes	no

The mean of two diagonal matrices is the diagonal matrix with the mean of the corresponding entries in the diagonal.
Our mean **mixes** the components.

Comparison

With $p = n\beta/2$; P_p Euclidean power mean; Q_p Lim-Palfia mean

Property	P_p	Q_p	G_β
commutativity with congruences	no	yes	yes
inversion	no	yes	yes
symmetry	yes	yes	yes
homogeneity	yes	yes	yes
consistency with scalars	yes	yes	no
global	yes	yes	no

Our mean has a **restriction** on the parameter / matrices.

Comparison

With $p = n\beta/2$; P_p Euclidean power mean; Q_p Lim-Palfia mean

Property	P_p	Q_p	G_β
commutativity with congruences	no	yes	yes
inversion	no	yes	yes
symmetry	yes	yes	yes
homogeneity	yes	yes	yes
consistency with scalars	yes	yes	no
global	yes	yes	no
Riemannian geodesic	?	?	yes

Our mean is **ennobled** by a Riemannian structure.

What's next

- A new family of geometries on $\mathcal{P}_n$
- Explicit geodesics and distances
- A new power mean of positive definite matrices

What's next

- A new family of geometries on $\mathcal{P}_n$
- Explicit geodesics and distances
- A new power mean of positive definite matrices

The power mean is flexible because of a free parameter ([Mercado,Tudisco,Hein,18-19],[Fasi,I.,18]).

- Try it on problems from applications (statistics, network theory,...).

Extra time: a conformal geometry

We can consider the geometry **conformal** to the one that defines the power mean

$$\langle H, K\rangle_X = \mathrm{trace}(X^{-1}HX^{-1}K) - \beta\,\mathrm{trace}(X^{-1}H)\,\mathrm{trace}(X^{-1}K),$$

For $\beta = 0$ is the affine-invariant geometry with distance δ.

The Karcher mean

The Karcher mean is the barycenter of $A_1, \ldots, A_m \in \mathcal{P}_n$ with the affine-invariant geometry. It minimizes

$$f(X) = \sum_{i=1}^{n} \delta^2(X, A_i) = \sum_{i=1}^{n} \| \log(A_i^{-1/2} X A_i^{-1/2}) \|_F^2$$

over $\mathcal{P}_n$.

It is computed with **Riemannian optimization**

- Riemannian gradient descent [Bini-I., '13];
- Riemannian Barzilai-Borwein [I.-Porcelli, '16];
- Riemannian L-BFGS [Yuan-Huang-Absil-Gallivan, '20]

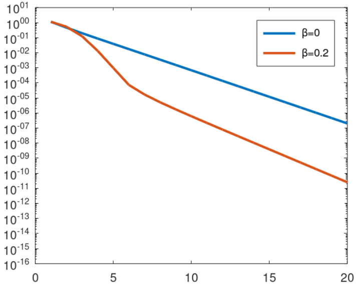
The Karcher mean computation

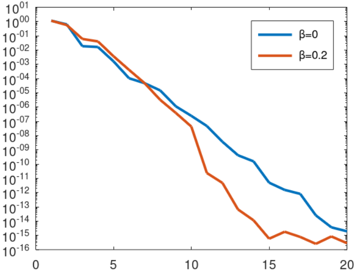
Lemma

The barycenter with respect to the conformal geometry is the Karcher mean for $\beta \in (-\infty, 0) \cup (0, 1/n)$.

A new parameter to set.

The Karcher mean computation





Left: Riemannian gradient descend
Right: Riemannian Barzilai-Borwein method