Information Geometry of Matrices and Mean

Attila Andai

RIKEN, BSI, Amari Research Unit

May 29, 2008

Outline

– Generalization of means.

  general concept of the mean

  extension to more variables

  extension to matrices

  difficulties with combining these ideas together

– Means in quantum information geometry.

– More geometry related to the means.

  examples

For strictly positive numbers x, y

arithmetic mean

$$M_a(x,y) = \frac{x+y}{2}$$

geometric mean

$$M_g(x,y) = \sqrt{xy}$$

harmonic mean

$$M_h(x,y) = \frac{2}{\frac{1}{x} + \frac{1}{y}}$$

Well-known inequality

$$M_h(x,y) \leq M_g(x,y) \leq M_a(x,y)$$

Generalizations

Some natural questions related to means:

Is the function

$$M_{\log}(x,y) = \frac{x-y}{\log x - \log y}$$

is a mean? (*logarithmic mean*)

For more variables we have the intuition

$$M_a(x,y,z) = \frac{x+y+z}{3} \qquad M_g(x,y,z) = \sqrt[3]{xyz}$$

How to generalize the logarithmic mean to more variables?

For matrices we have the intuition

$$M_a(X,Y) = \frac{1}{2}(X+Y).$$

But to compute the geometric mean of matrices?

What is the logarithmic mean of three matrices???

Generalizations

Some natural questions related to means:
Is the function

$$M_{\log}(x,y) = \frac{x-y}{\log x - \log y}$$

is a mean? (*logarithmic mean*)
For more variables we have the intuition

$$M_a(x,y,z) = \frac{x+y+z}{3}, \quad M_g(x,y,z) = \sqrt[3]{xyz}$$

How to generalize the logarithmic mean to more variables?
For matrices we have the intuition

$$M_a(X,Y) = \frac{1}{2}(X+Y).$$

But to compute the geometric mean of matrices?
What is the logarithmic mean of three matrices???

Generalizations

Some natural questions related to means:
Is the function

$$M_{\log}(x,y) = \frac{x-y}{\log x - \log y}$$

is a mean? (*logarithmic mean*)
For more variables we have the intuition

$$M_a(x,y,z) = \frac{x+y+z}{3} \quad M_g(x,y,z) = \sqrt[3]{xyz}.$$

How to generalize the logarithmic mean to more variables?
For matrices we have the intuition

$$M_a(X,Y) = \frac{1}{2}(X+Y).$$

But to compute the geometric mean of matrices?
What is the logarithmic mean of three matrices???

Generalizations

Some natural questions related to means:
Is the function

$$M_{\log}(x,y) = \frac{x-y}{\log x - \log y}$$

is a mean? (*logarithmic mean*)
For more variables we have the intuition

$$M_a(x,y,z) = \frac{x+y+z}{3} \qquad M_g(x,y,z) = \sqrt[3]{xyz}.$$

How to generalize the logarithmic mean to more variables?
For matrices we have the intuition

$$M_a(X,Y) = \frac{1}{2}(X+Y).$$

But to compute the geometric mean of matrices?
What is the logarithmic mean of three matrices???

Generalizations

Some natural questions related to means:
Is the function

$$M_{\log}(x,y) = \frac{x-y}{\log x - \log y}$$

is a mean? (*logarithmic mean*)
For more variables we have the intuition

$$M_a(x,y,z) = \frac{x+y+z}{3} \quad M_g(x,y,z) = \sqrt[3]{xyz}.$$

How to generalize the logarithmic mean to more variables?
For matrices we have the intuition

$$M_a(X,Y) = \frac{1}{2}(X+Y).$$

But to compute the geometric mean of matrices?
What is the logarithmic mean of three matrices???

What is a mean?

A function $M : \mathbb{R}^+ \times \mathbb{R}^+ \to \mathbb{R}^+$ is a *mean* if $(\forall x, y, x_0, y_0, t \in \mathbb{R}^+)$

What is a mean?

A function $M : \mathbb{R}^+ \times \mathbb{R}^+ \to \mathbb{R}^+$ is a *mean* if
$(\forall x, y, x_0, y_0, t \in \mathbb{R}^+)$

$$M(x, x) = x$$

What is a mean?

A function $M : \mathbb{R}^+ \times \mathbb{R}^+ \to \mathbb{R}^+$ is a *mean* if $(\forall x, y, x_0, y_0, t \in \mathbb{R}^+)$

$$M(x,x) = x$$

$$M(x,y) = M(y,x)$$

What is a mean?

A function $M:\mathbb{R}^+\times\mathbb{R}^+\to\mathbb{R}^+$ is a *mean* if $(\forall x,y,x_0,y_0,t\in\mathbb{R}^+)$

$$M(x,x)=x$$

$$M(x,y)=M(y,x)$$

$$x<y\ \Rightarrow\ x<M(x,y)<y$$

What is a mean?

A function $M : \mathbb{R}^+ \times \mathbb{R}^+ \to \mathbb{R}^+$ is a *mean* if $(\forall x, y, x_0, y_0, t \in \mathbb{R}^+)$

$$M(x,x) = x$$

$$M(x,y) = M(y,x)$$

$$x < y \;\Rightarrow\; x < M(x,y) < y$$

$$x < x_0, y < y_0 \Rightarrow M(x,y) < M(x_0,y_0)$$

What is a mean?

A function $M : \mathbb{R}^+ \times \mathbb{R}^+ \to \mathbb{R}^+$ is a *mean* if $(\forall x, y, x_0, y_0, t \in \mathbb{R}^+)$

$$M(x, x) = x$$

$$M(x, y) = M(y, x)$$

$$x < y \;\Rightarrow\; x < M(x, y) < y$$

$$x < x_0, y < y_0 \Rightarrow M(x, y) < M(x_0, y_0)$$

$M(x, y)$ is continuous

What is a mean?

A function $M : \mathbb{R}^+ \times \mathbb{R}^+ \to \mathbb{R}^+$ is a *mean* if $(\forall x, y, x_0, y_0, t \in \mathbb{R}^+)$

$$M(x,x) = x$$

$$M(x,y) = M(y,x)$$

$$x < y \ \Rightarrow \ x < M(x,y) < y$$

$$x < x_0, y < y_0 \Rightarrow M(x,y) < M(x_0,y_0)$$

$M(x,y)$ is continuous

$$M(tx,ty) = tM(x,y)$$

What is a mean?

A function $M : \mathbb{R}^+ \times \mathbb{R}^+ \to \mathbb{R}^+$ is a *mean* if $(\forall x, y, x_0, y_0, t \in \mathbb{R}^+)$

$M(x,x) = x$

$M(x,y) = M(y,x)$

$x < y \ \Rightarrow \ x < M(x,y) < y$

$x < x_0, y < y_0 \Rightarrow M(x,y) < M(x_0,y_0)$

$M(x,y)$ is continuous

$M(tx,ty) = tM(x,y)$

$$M(x,y) = xf\left(\frac{y}{x}\right)$$

What is a mean?

A function $M : \mathbb{R}^+ \times \mathbb{R}^+ \to \mathbb{R}^+$ is a *mean* if $(\forall x, y, x_0, y_0, t \in \mathbb{R}^+)$

$M(x,x) = x$ $\qquad f(1) = 1$

$M(x,y) = M(y,x)$

$x < y \ \Rightarrow \ x < M(x,y) < y$

$x < x_0, y < y_0 \Rightarrow M(x,y) < M(x_0,y_0)$

$M(x,y)$ is continuous

$M(tx,ty) = tM(x,y)$

$$M(x,y) = xf\left(\frac{y}{x}\right)$$

What is a mean?

A function $M : \mathbb{R}^+ \times \mathbb{R}^+ \to \mathbb{R}^+$ is a *mean* if $(\forall x, y, x_0, y_0, t \in \mathbb{R}^+)$

$M(x,x) = x$ $\qquad f(1) = 1$

$M(x,y) = M(y,x)$ $\qquad f(t) = tf(t^{-1})$

$x < y \ \Rightarrow \ x < M(x,y) < y$

$x < x_0, y < y_0 \Rightarrow M(x,y) < M(x_0,y_0)$

$M(x,y)$ is continuous

$M(tx,ty) = tM(x,y)$

$$M(x,y) = xf\left(\frac{y}{x}\right)$$

What is a mean?

A function $M : \mathbb{R}^+ \times \mathbb{R}^+ \to \mathbb{R}^+$ is a *mean* if $(\forall x, y, x_0, y_0, t \in \mathbb{R}^+)$

$M(x,x) = x$ $\qquad$ $f(1) = 1$

$M(x,y) = M(y,x)$ $\qquad$ $f(t) = tf(t^{-1})$

$x < y \;\Rightarrow\; x < M(x,y) < y$ $\qquad$ $f(\cdot > 1) > 1,\; f(0 < \cdot < 1) < 1$

$x < x_0, y < y_0 \Rightarrow M(x,y) < M(x_0,y_0)$

$M(x,y)$ is continuous

$M(tx,ty) = tM(x,y)$

$$M(x,y) = xf\left(\frac{y}{x}\right)$$

What is a mean?

A function $M:\mathbb{R}^+\times\mathbb{R}^+\to\mathbb{R}^+$ is a *mean* if $(\forall x,y,x_0,y_0,t\in\mathbb{R}^+)$

$M(x,x)=x$	$f(1)=1$
$M(x,y)=M(y,x)$	$f(t)=tf(t^{-1})$
$x<y \;\Rightarrow\; x<M(x,y)<y$	$f(\cdot>1)>1,\; f(0<\cdot<1)<1$
$x<x_0, y<y_0 \Rightarrow M(x,y)<M(x_0,y_0)$	f increasing
$M(x,y)$ is continuous	
$M(tx,ty)=tM(x,y)$	

$$M(x,y)=xf\left(\frac{y}{x}\right)$$

What is a mean?

A function $M : \mathbb{R}^+ \times \mathbb{R}^+ \to \mathbb{R}^+$ is a *mean* if $(\forall x, y, x_0, y_0, t \in \mathbb{R}^+)$

$M(x,x) = x$	$f(1) = 1$
$M(x,y) = M(y,x)$	$f(t) = tf(t^{-1})$
$x < y \;\Rightarrow\; x < M(x,y) < y$	$f(\cdot > 1) > 1,\ f(0 < \cdot < 1) < 1$
$x < x_0, y < y_0 \Rightarrow M(x,y) < M(x_0,y_0)$	f increasing
$M(x,y)$ is continuous	f continuous
$M(tx,ty) = tM(x,y)$	

$$M(x,y) = xf\left(\frac{y}{x}\right)$$

We have

$$\text{means} = \left\{ f \in C(\mathbb{R}^+, \mathbb{R}^+) \;\middle|\; \begin{array}{c} f \text{ increasing} \\ f(1) = 1 \\ \forall t \in \mathbb{R}^+ : \; f(t) = t f(t^{-1}) \end{array} \right\}$$

$$M(x,y) = x f\left(\frac{y}{x}\right)$$

We have

$$\text{means} = \left\{ f \in C(\mathbb{R}^+, \mathbb{R}^+) \,\middle|\, \begin{array}{c} f \text{ increasing} \\ f(1) = 1 \\ \forall t \in \mathbb{R}^+ : \; f(t) = t f(t^{-1}) \end{array} \right\}$$

$$M(x,y) = x f\left(\frac{y}{x}\right)$$

arithmetic mean: $f(t) = \dfrac{1+t}{2}$

geometric mean: $f(t) = \sqrt{t}$

logarithmic mean: $f(t) = \dfrac{t-1}{\log t}$

Definition: A mean m of three variables is said to be of *type 1 invariant* with respect to M if

$$m(M(x,y), M(y,z), M(z,x)) = m(x,y,z)\ .$$

Definition: A mean m of three variables is said to be of *type 1 invariant* with respect to M if

$$m(M(x,y), M(y,z), M(z,x)) = m(x,y,z) \ .$$

Theorem: To each M there exists a unique m which is type 1 invariant with respect to M.

Proof:

Define $x_0 := x$, $y_0 := y$, $z_0 := z$ and iterate

$$x_{n+1} := M(y_n, z_n) \quad y_{n+1} := M(z_n, x_n) \quad z_{n+1} := M(x_n, y_n)$$

Definition: A mean m of three variables is said to be of *type 1 invariant* with respect to M if

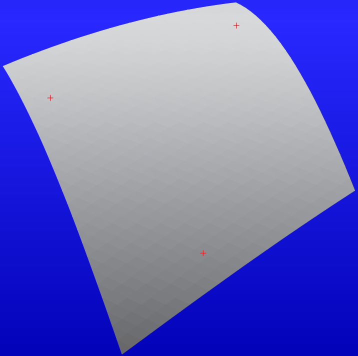
$$m(M(x,y), M(y,z), M(z,x)) = m(x,y,z)\ .$$

Theorem: To each M there exists a unique m which is type 1 invariant with respect to M.

Proof:

Define $x_0 := x$, $y_0 := y$, $z_0 := z$ and iterate

$$x_{n+1} := M(y_n, z_n) \quad y_{n+1} := M(z_n, x_n) \quad z_{n+1} := M(x_n, y_n)$$



Definition: A mean m of three variables is said to be of *type 1 invariant* with respect to M if

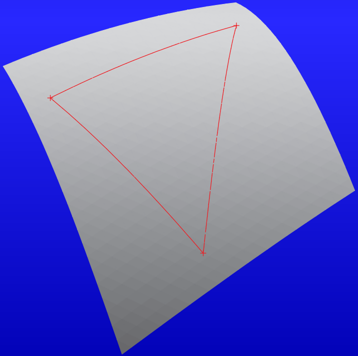
$$m(M(x,y), M(y,z), M(z,x)) = m(x,y,z) \ .$$

Theorem: To each M there exists a unique m which is type 1 invariant with respect to M.

Proof:

Define $x_0 := x$, $y_0 := y$, $z_0 := z$ and iterate

$$x_{n+1} := M(y_n, z_n) \quad y_{n+1} := M(z_n, x_n) \quad z_{n+1} := M(x_n, y_n)$$



Definition: A mean m of three variables is said to be of *type 1 invariant* with respect to M if

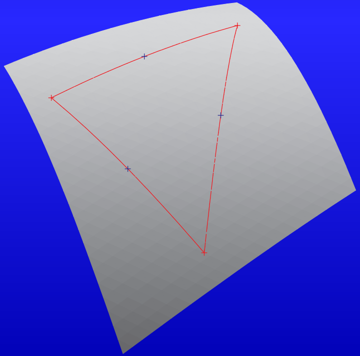
$$m(M(x,y), M(y,z), M(z,x)) = m(x,y,z)\ .$$

Theorem: To each M there exists a unique m which is type 1 invariant with respect to M.

Proof:

Define $x_0 := x$, $y_0 := y$, $z_0 := z$ and iterate

$$x_{n+1} := M(y_n, z_n) \quad y_{n+1} := M(z_n, x_n) \quad z_{n+1} := M(x_n, y_n)$$



Definition: A mean m of three variables is said to be of *type 1 invariant* with respect to M if

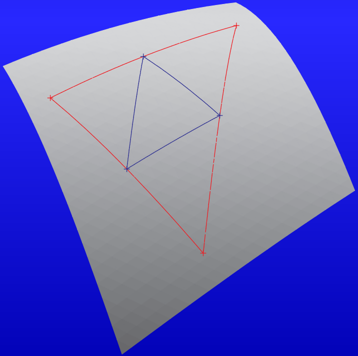
$$m(M(x,y),M(y,z),M(z,x)) = m(x,y,z)\ .$$

Theorem: To each M there exists a unique m which is type 1 invariant with respect to M.

Proof:

Define $x_0 := x$, $y_0 := y$, $z_0 := z$ and iterate

$$x_{n+1} := M(y_n, z_n) \quad y_{n+1} := M(z_n, x_n) \quad z_{n+1} := M(x_n, y_n)$$



Definition: A mean m of three variables is said to be of *type 1 invariant* with respect to M if

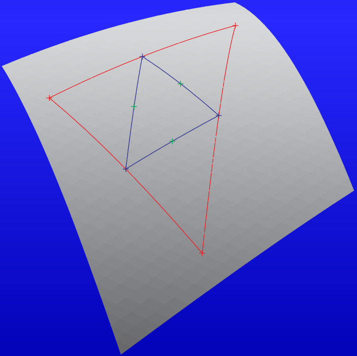
$$m(M(x,y),M(y,z),M(z,x)) = m(x,y,z)\ .$$

Theorem: To each M there exists a unique m which is type 1 invariant with respect to M.

Proof:

Define $x_0 := x$, $y_0 := y$, $z_0 := z$ and iterate

$$x_{n+1} := M(y_n, z_n) \quad y_{n+1} := M(z_n, x_n) \quad z_{n+1} := M(x_n, y_n)$$



Definition: A mean m of three variables is said to be of *type 1 invariant* with respect to M if

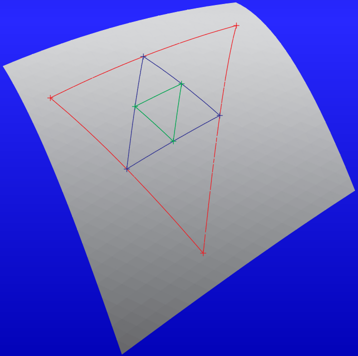
$$m(M(x,y), M(y,z), M(z,x)) = m(x,y,z)\ .$$

Theorem: To each M there exists a unique m which is type 1 invariant with respect to M.

Proof:

Define $x_0 := x$, $y_0 := y$, $z_0 := z$ and iterate

$$x_{n+1} := M(y_n, z_n) \quad y_{n+1} := M(z_n, x_n) \quad z_{n+1} := M(x_n, y_n)$$



Definition: A mean m of three variables is said to be of *type 1 invariant* with respect to M if

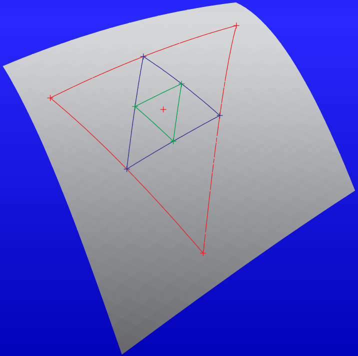
$$m(M(x,y),M(y,z),M(z,x)) = m(x,y,z)\ .$$

Theorem: To each M there exists a unique m which is type 1 invariant with respect to M.

Proof:

Define $x_0 := x$, $y_0 := y$, $z_0 := z$ and iterate

$$x_{n+1} := M(y_n,z_n) \quad y_{n+1} := M(z_n,x_n) \quad z_{n+1} := M(x_n,y_n)$$



Definition: A mean m of three variables is said to be of *type 1 invariant* with respect to M if

$$m(M(x,y), M(y,z), M(z,x)) = m(x,y,z) \ .$$

Theorem: To each M there exists a unique m which is type 1 invariant with respect to M.

Proof:

Define $x_0 := x$, $y_0 := y$, $z_0 := z$ and iterate

$$x_{n+1} := M(y_n, z_n) \quad y_{n+1} := M(z_n, x_n) \quad z_{n+1} := M(x_n, y_n)$$

Check $\lim\limits_{n\to\infty} |x_n - y_n| = 0$.

Define

$$m(x,y,z) := \lim_{n\to\infty} x_n \ .$$

Definition: A mean m of three variables is said to be of *type 1 invariant* with respect to M if

$$m(M(x,y), M(y,z), M(z,x)) = m(x,y,z) \ .$$

Theorem: To each M there exists a unique m which is type 1 invariant with respect to M.

Proof:

Define $x_0 := x$, $y_0 := y$, $z_0 := z$ and iterate

$$x_{n+1} := M(y_n, z_n) \quad y_{n+1} := M(z_n, x_n) \quad z_{n+1} := M(x_n, y_n)$$

Check $\lim_{n\to\infty} |x_n - y_n| = 0$.

Define

$$m(x,y,z) := \lim_{n\to\infty} x_n \ .$$

Example:

$$\frac{\frac{x+y}{2} + \frac{y+z}{2} + \frac{z+x}{2}}{3} = \frac{x+y+z}{3}$$

$$\sqrt[3]{\sqrt{xy} \times \sqrt{yz} \times \sqrt{zx}} = \sqrt[3]{xyz}$$

Logarithmic mean $L(x,y) = \dfrac{x-y}{\log x - \log y}$ with three variables:

$$U_0(x,y,z) = \sqrt{\frac{1}{2} \times \frac{(x-z)(y-z)(x-y)}{x \log \frac{y}{z} + y \log \frac{z}{x} + z \log \frac{x}{y}}}$$

$$U_1(x,y,z) = \frac{1}{2} \times \frac{(y-z)(x-z)(x-y)}{x(y-z)\log x + y(z-x)\log y + z(x-y)\log z}$$

Conjecture:

$$U_0(x,y,z) \leq L_3(x,y,z) \leq U_1(x,y,z)$$

Numerical example $x = 1$, $y = 2$, $z = 3$:

$$\sim 1.8644 \leq \sim 1.8791 \leq \sim 1.9111$$

$L_3(x,y,z) = ?$

Logarithmic mean $L(x,y) = \dfrac{x-y}{\log x - \log y}$ with three variables:

$$U_0(x,y,z) = \sqrt{\frac{1}{2} \times \frac{(x-z)(y-z)(x-y)}{x \log \frac{y}{z} + y \log \frac{z}{x} + z \log \frac{x}{y}}}$$

$$U_1(x,y,z) = \frac{1}{2} \times \frac{(y-z)(x-z)(x-y)}{x(y-z)\log x + y(z-x)\log y + z(x-y)\log z}$$

Conjecture:

$$U_0(x,y,z) \leq L_3(x,y,z) \leq U_1(x,y,z)$$

Numerical example $x=1$, $y=2$, $z=3$:

$$\sim 1.8644 \leq \sim 1.8791 \leq \sim 1.9111$$

$L_3(x,y,z) = ?$

Define means on $n \times n$, positive definite matrices $\mathcal{M}_n$:

$$X \in \mathcal{M}_n \iff X = X^*, \begin{cases} \langle v, Xv \rangle > 0 \ \forall 0 \neq v \in \mathbb{R}^n, \mathbb{C}^n \\ \text{every eigenvalue of } X \text{ is positive} \end{cases}$$

We write $X \leq Y$ if $Y - X \in \mathcal{M}_n$.

How to compute $f(X)$:

– $X \in \mathcal{M}_n$ can be diagonalized by some unitary matrix U, that is $X = UDU^*$

– X can be written as $X = \sum_{i=1}^{n} \lambda_i E_i$, where $(\lambda_i)_{i=1,\ldots,n}$ are the eigenvalues and $(E_i)_{i=1,\ldots,n}$ are the corresponding projections

$$f(X) = \sum_{i=1}^{n} f(\lambda_i) E_i$$

f is operator monotone if $X \leq Y$ then $f(X) \leq f(Y)$.

Define means on $n \times n$, positive definite matrices $\mathcal{M}_n$:

$$X \in \mathcal{M}_n \iff X = X^*, \begin{cases} \langle v, Xv \rangle > 0 \ \forall 0 \neq v \in \mathbb{R}^n, \mathbb{C}^n \\ \text{every eigenvalue of } X \text{ is positive} \end{cases}$$

We write $X \leq Y$ if $Y - X \in \mathcal{M}_n$.

How to compute $f(X)$:

– $X \in \mathcal{M}_n$ can be diagonalized by some unitary matrix U, that is $X = UDU^*$

– X can be written as $X = \sum_{i=1}^{n} \lambda_i E_i$, where $(\lambda_i)_{i=1,\ldots,n}$ are the eigenvalues and $(E_i)_{i=1,\ldots,n}$ are the corresponding projections

$$f(X) = \sum_{i=1}^{n} f(\lambda_i) E_i$$

f is operator monotone if $X \leq Y$ then $f(X) \leq f(Y)$.

Define means on $n \times n$, positive definite matrices $\mathcal{M}_n$:

$$X \in \mathcal{M}_n \iff X = X^*, \begin{cases} \langle v, Xv \rangle > 0 \ \forall 0 \neq v \in \mathbb{R}^n, \mathbb{C}^n \\ \text{every eigenvalue of } X \text{ is positive} \end{cases}$$

We write $X \leq Y$ if $Y - X \in \mathcal{M}_n$.

How to compute $f(X)$:

– $X \in \mathcal{M}_n$ can be diagonalized by some unitary matrix U, that is $X = UDU^*$

$$f(X) := Uf(D)U^*$$

– X can be written as $X = \sum_{i=1}^{n} \lambda_i E_i$, where $(\lambda_i)_{i=1,\dots,n}$ are the eigenvalues and $(E_i)_{i=1,\dots,n}$ are the corresponding projections

$$f(X) = \sum_{i=1}^{n} f(\lambda_i) E_i$$

f is operator monotone if $X \leq Y$ then $f(X) \leq f(Y)$.

Define means on $n \times n$, positive definite matrices $\mathcal{M}_n$:

$$X \in \mathcal{M}_n \iff X = X^*, \begin{cases} \langle v, Xv \rangle > 0 \ \forall 0 \neq v \in \mathbb{R}^n, \mathbb{C}^n \\ \text{every eigenvalue of } X \text{ is positive} \end{cases}$$

We write $X \leq Y$ if $Y - X \in \mathcal{M}_n$.
How to compute $f(X)$:
– $X \in \mathcal{M}_n$ can be diagonalized by some unitary matrix U, that is $X = UDU^*$

$$f(X) := U f(D) U^*$$

– X can be written as $X = \sum_{i=1}^{n} \lambda_i E_i$, where $(\lambda_i)_{i=1,\dots,n}$ are the eigenvalues and $(E_i)_{i=1,\dots,n}$ are the corresponding projections

$$f(X) = \sum_{i=1}^{n} f(\lambda_i) E_i$$

f is operator monotone if $X \leq Y$ then $f(X) \leq f(Y)$.

Define means on $n \times n$, positive definite matrices $\mathcal{M}_n$:

$$X \in \mathcal{M}_n \iff X = X^*, \begin{cases} \langle v, Xv \rangle > 0 \ \forall 0 \neq v \in \mathbb{R}^n, \mathbb{C}^n \\ \text{every eigenvalue of } X \text{ is positive} \end{cases}$$

We write $X \leq Y$ if $Y - X \in \mathcal{M}_n$.
How to compute $f(X)$:
– $X \in \mathcal{M}_n$ can be diagonalized by some unitary matrix U, that is $X = UDU^*$

$$f(X) := Uf(D)U^*$$

– X can be written as $X = \sum_{i=1}^{n} \lambda_i E_i$, where $(\lambda_i)_{i=1,\dots,n}$ are the eigenvalues and $(E_i)_{i=1,\dots,n}$ are the corresponding projections

$$f(X) = \sum_{i=1}^{n} f(\lambda_i) E_i$$

f is operator monotone if $X \leq Y$ then $f(X) \leq f(Y)$.

M is a *mean of matrices* if for every $X, Y \in \mathcal{M}_n$

– $X \leq X_0,\ Y \leq Y_0:\ M(X,Y) \leq M(X_0,Y_0)$

– $(X_n)_{n\in\mathbb{N}}$ and $(Y_n)_{n\in\mathbb{N}}$ are decreasing sequences $(X_{n+1} \leq X_n,\ Y_{n+1} \leq Y_n)$ in $\mathcal{M}_n$ with limits X and Y then $M(X_n, Y_n)$ is decreasing and

$$\lim_{n\to\infty} M(X_n, Y_n) = M(X, Y)$$

– $T^* M(X,Y) T \leq M(T^* X T, T^* Y T)$ for all T

– $M(X,X) = X$

Theorem: there exists an operator monotone function f with properties $f(t) = t f(t^{-1})$ and $f(1) = 1$ such that for every $X, Y \in \mathcal{M}_n$

$$M(X,Y) = X^{1/2} f(X^{-1/2} Y X^{-1/2}) X^{1/2}$$

For real numbers we had:

$$M(x,y) = x f\left(\frac{y}{x}\right).$$

M is a *mean of matrices* if for every $X, Y \in \mathcal{M}_n$

– $X \le X_0,\ Y \le Y_0:\ M(X,Y) \le M(X_0,Y_0)$

– $(X_n)_{n\in\mathbb{N}}$ and $(Y_n)_{n\in\mathbb{N}}$ are decreasing sequences $(X_{n+1} \le X_n,\ Y_{n+1} \le Y_n)$ in $\mathcal{M}_n$ with limits X and Y then $M(X_n,Y_n)$ is decreasing and

$$\lim_{n\to\infty} M(X_n,Y_n) = M(X,Y)$$

– $T^*M(X,Y)T \le M(T^*XT, T^*YT)$ for all T

– $M(X,X) = X$

Theorem: there exists an operator monotone function f with properties $f(t) = tf(t^{-1})$ and $f(1) = 1$ such that for every $X, Y \in \mathcal{M}_n$

$$M(X,Y) = X^{1/2} f(X^{-1/2} Y X^{-1/2}) X^{1/2}$$

For real numbers we had:

$$M(x,y) = xf\left(\frac{y}{x}\right).$$

M is a *mean of matrices* if for every $X, Y \in \mathcal{M}_n$
– $X \leq X_0,\ Y \leq Y_0:\ M(X,Y) \leq M(X_0, Y_0)$
– $(X_n)_{n\in\mathbb{N}}$ and $(Y_n)_{n\in\mathbb{N}}$ are decreasing sequences $(X_{n+1} \leq X_n,\ Y_{n+1} \leq Y_n)$ in $\mathcal{M}_n$ with limits X and Y then $M(X_n, Y_n)$ is decreasing and

$$\lim_{n\to\infty} M(X_n, Y_n) = M(X,Y)$$

– $T^* M(X,Y) T \leq M(T^*XT, T^*YT)$ for all T
– $M(X,X) = X$

Theorem: (Kubo, Ando) If M is a matrix mean, then there exists an operator monotone function f with properties $f(t) = t f(t^{-1})$ and $f(1) = 1$ such that for every $X, Y \in \mathcal{M}_n$

$$M(X,Y) = X^{1/2} f(X^{-1/2} Y X^{-1/2}) X^{1/2}$$

For real numbers we had:

$$M(x,y) = x f\left(\frac{y}{x}\right).$$

M is a *mean of matrices* if for every $X, Y \in \mathcal{M}_n$

– $X \leq X_0,\ Y \leq Y_0:\ M(X,Y) \leq M(X_0, Y_0)$

– $(X_n)_{n\in\mathbb{N}}$ and $(Y_n)_{n\in\mathbb{N}}$ are decreasing sequences $(X_{n+1} \leq X_n,\ Y_{n+1} \leq Y_n)$ in $\mathcal{M}_n$ with limits X and Y then $M(X_n, Y_n)$ is decreasing and

$$\lim_{n\to\infty} M(X_n, Y_n) = M(X,Y)$$

– $T^* M(X,Y) T \leq M(T^*XT, T^*YT)$ for all T

– $M(X,X) = X$

Theorem: (Kubo–Ando) If M is a mean of matrices then there exists an operator monotone function f with properties $f(t) = tf(t^{-1})$ and $f(1) = 1$ such that for every $X, Y \in \mathcal{M}_n$

$$M(X,Y) = X^{1/2} f(X^{-1/2} Y X^{-1/2}) X^{1/2}$$

For real numbers we had:

$$M(x,y) = xf\left(\frac{y}{x}\right).$$

M is a *mean of matrices* if for every $X, Y \in \mathcal{M}_n$

– $X \leq X_0,\ Y \leq Y_0:\ M(X,Y) \leq M(X_0, Y_0)$

– $(X_n)_{n\in\mathbb{N}}$ and $(Y_n)_{n\in\mathbb{N}}$ are decreasing sequences $(X_{n+1} \leq X_n,\ Y_{n+1} \leq Y_n)$ in $\mathcal{M}_n$ with limits X and Y then $M(X_n, Y_n)$ is decreasing and

$$\lim_{n\to\infty} M(X_n, Y_n) = M(X, Y)$$

– $T^* M(X,Y) T \leq M(T^* X T, T^* Y T)$ for all T

– $M(X,X) = X$

Theorem:

there exists an operator monotone function f with properties $f(t) = t f(t^{-1})$ and $f(1) = 1$ such that for every $X, Y \in \mathcal{M}_n$

$$M(X,Y) = X^{1/2} f(X^{-1/2} Y X^{-1/2}) X^{1/2}$$

For real numbers we had:

$$M(x,y) = x f\left(\frac{y}{x}\right).$$

M is a *mean of matrices* if for every $X, Y \in \mathcal{M}_n$

– $X \leq X_0,\ Y \leq Y_0:\ M(X,Y) \leq M(X_0,Y_0)$

– $(X_n)_{n\in\mathbb{N}}$ and $(Y_n)_{n\in\mathbb{N}}$ are decreasing sequences $(X_{n+1} \leq X_n,\ Y_{n+1} \leq Y_n)$ in $\mathcal{M}_n$ with limits X and Y then $M(X_n,Y_n)$ is decreasing and

$$\lim_{n\to\infty} M(X_n,Y_n) = M(X,Y)$$

– $T^*M(X,Y)T \leq M(T^*XT, T^*YT)$ for all T

– $M(X,X) = X$

Theorem: (*Kubo–Ando*) If M is a matrix mean, then there exists an operator monotone function f with properties $f(t) = tf(t^{-1})$ and $f(1) = 1$ such that for every $X, Y \in \mathcal{M}_n$

$$M(X,Y) = X^{1/2} f(X^{-1/2} Y X^{-1/2}) X^{1/2}$$

For real numbers we had:

$$M(x,y) = xf\left(\frac{y}{x}\right).$$

M is a *mean of matrices* if for every $X, Y \in \mathcal{M}_n$

– $X \leq X_0,\ Y \leq Y_0:\ M(X,Y) \leq M(X_0, Y_0)$

– $(X_n)_{n\in\mathbb{N}}$ and $(Y_n)_{n\in\mathbb{N}}$ are decreasing sequences $(X_{n+1} \leq X_n,\ Y_{n+1} \leq Y_n)$ in $\mathcal{M}_n$ with limits X and Y then $M(X_n, Y_n)$ is decreasing and

$$\lim_{n\to\infty} M(X_n, Y_n) = M(X,Y)$$

– $T^*M(X,Y)T \leq M(T^*XT, T^*YT)$ for all T

– $M(X,X) = X$

Theorem: (*Kubo–Ando*) If M is a matrix mean, then there exists an operator monotone function f with properties $f(t) = tf(t^{-1})$ and $f(1) = 1$ such that for every $X, Y \in \mathcal{M}_n$

$$M(X,Y) = X^{1/2} f(X^{-1/2} Y X^{-1/2}) X^{1/2}$$

For real numbers we had:

$$M(x,y) = xf\left(\frac{y}{x}\right).$$

Problems with means

– General mean

Problems with means

– General mean ✓

Problems with means

– General mean ✓
– More variables

Problems with means

– General mean ✓

– More variables : if more $= 3$ ✓

Problems with means

– General mean ✓

– More variables : if more $= 3$ ✓

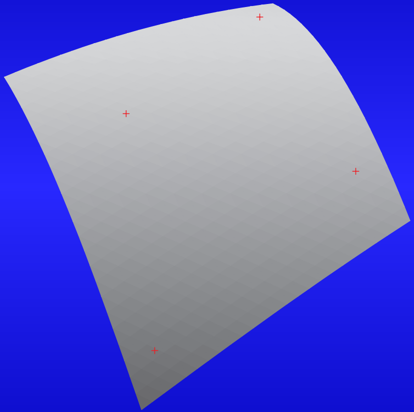
if more $= 4$:

Problems with means

– General mean ✓

– More variables : if more = 3 ✓

if more = 4 :

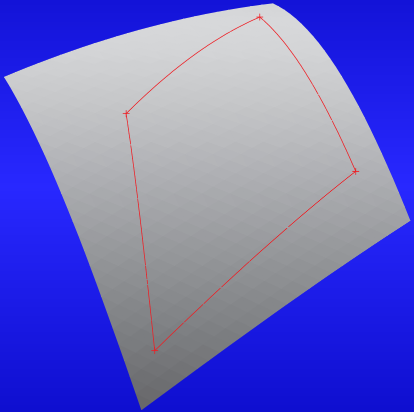


Problems with means

– General mean ✓

– More variables : if more = 3 ✓

if more = 4 :

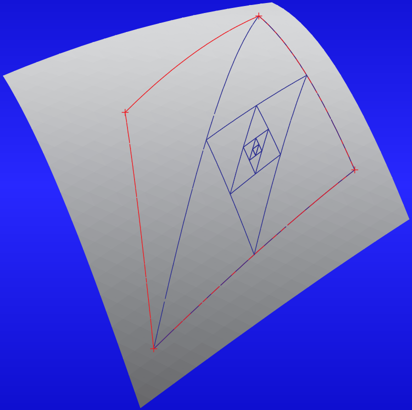


Problems with means

– General mean ✓

– More variables : if more $= 3$ ✓

if more $= 4$:

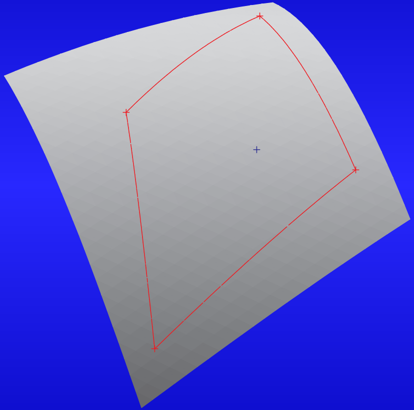


Problems with means

– General mean ✓

– More variables : if more = 3 ✓

if more = 4 :

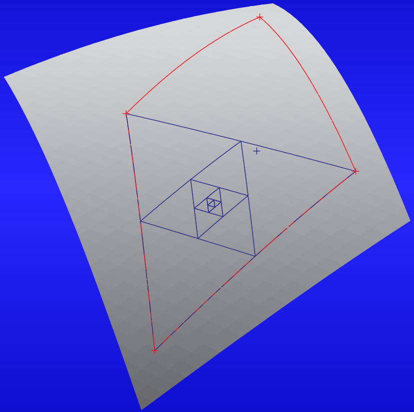


Problems with means

– General mean ✓

– More variables : if more $= 3$ ✓

if more $= 4$:

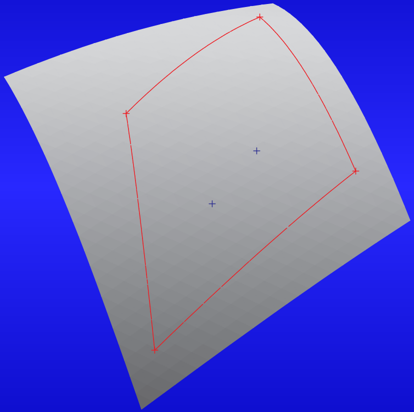


Problems with means

– General mean ✓

– More variables : if more = 3 ✓

if more = 4 :

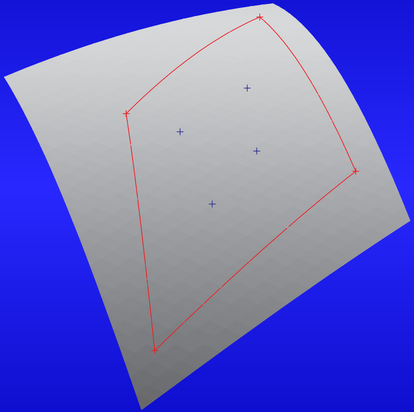


Problems with means

– General mean ✓

– More variables : if more $= 3$ ✓

if more $= 4$:

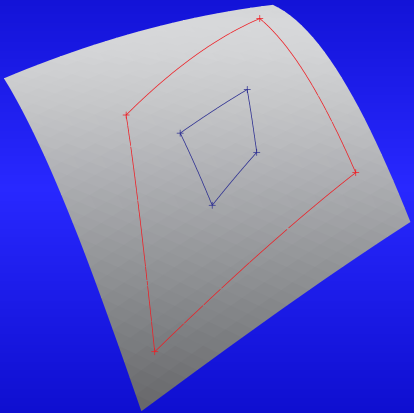


Problems with means

– General mean ✓

– More variables : if more = 3 ✓

if more = 4 :

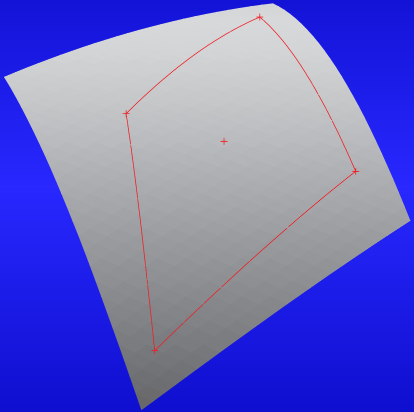


Problems with means

– General mean ✓

– More variables : if more = 3 ✓

if more = 4 :



Problems with means

– General mean ✓

– More variables : if more = 3 ✓

if more = 4 : ✓

Problems with means

– General mean ✓
– More variables ✓

Problems with means

- General mean ✓
- More variables ✓ Explicit form???

Problems with means

– General mean ✓

– More variables ✓ Explicit form???

– Matrices

Problems with means

– General mean ✓
– More variables ✓ Explicit form???
– Matrices ✓

Problems with means

– General mean ✓
– More variables ✓ Explicit form???
– Matrices ✓

General mean of more matrices:

Problems with means

– General mean ✓
– More variables ✓   Explicit form???
– Matrices ✓

General mean of more matrices: ....

Problems with means

- General mean ✓
- More variables ✓ Explicit form???
- Matrices ✓

General mean of more matrices: ....

General mean of 3 matrices:

Problems with means

– General mean ✓
– More variables ✓ Explicit form???
– Matrices ✓

General mean of more matrices: ....

General mean of 3 matrices:
$+$: The symmetrization method is convergent for the arithmetic, geometric and harmonic means.

Problems with means

– General mean ✓
– More variables ✓ Explicit form???
– Matrices ✓

General mean of more matrices: ....

General mean of 3 matrices:
+ : The symmetrization method is convergent for the arithmetic, geometric and harmonic means.
- : The convergence is unknown in the other cases!

Problems with means

– General mean ✓
– More variables ✓ Explicit form???
– Matrices ✓

General mean of more matrices: ....

General mean of 3 matrices:
+ : The symmetrization method is convergent for the arithmetic, geometric and harmonic means.
- : The convergence is unknown in the other cases!
+? : Conjecture: $\|x_{n+1} - y_{n+1}\| \leq \|x_n - y_n\|$. $(Petz)$

Problems with means

– General mean ✓
– More variables ✓ Explicit form???
– Matrices ✓

General mean of more matrices: ....

General mean of 3 matrices:
+ : The symmetrization method is convergent for the arithmetic, geometric and harmonic means.
- : The convergence is unknown in the other cases!
+? : Conjecture: $\|x_{n+1} - y_{n+1}\| \leq \|x_n - y_n\|$. $(Petz)$
- : $x_0 := 0.01$, $y_0 := 0.02$, $z_0 := 1$

$$f(x) = \frac{500x}{999x+1} + \frac{500x}{x+999}$$

$$\|x_{n+1} - y_{n+1}\| \approx 0.02669 > 0.02 = \|x_n - y_n\|.$$

Contradiction!

Means in Quantum Info. Geometry

In the classical case: uniqueness of the Fisher information.

Means in Quantum Info. Geometry

In the classical case: uniqueness of the Fisher information.
Open set of distributions on $X_n = \{1, \ldots, n\}$

$$\mathcal{P}_n = \left\{ (p_1, \ldots, p_n) \,\middle|\, 0 < p_i < 1, \sum_{i=1}^{n} p_i = 1 \right\}.$$

Means in Quantum Info. Geometry

In the classical case: uniqueness of the Fisher information. Open set of distributions on $X_n = \{1, \dots, n\}$

$$\mathcal{P}_n = \left\{ (p_1, \dots, p_n) \,\middle|\, 0 < p_i < 1, \sum_{i=1}^{n} p_i = 1 \right\}.$$

Theorem (*Čencov:*) Assume that for every $n \in \mathbb{N}$ the pair $(\mathcal{P}_n, g_n)$ is a Riemannian-manifold. If for every Markovian map $\kappa : X_n \times X_m \to \mathbb{R}$
the following monotonicity property holds

$$g_{\tilde{\kappa}(p)}(\kappa^*(X), \kappa^*(X)) \leq g_p(X, X) \qquad \forall p \in \mathcal{P}_n, \forall X \in \mathrm{T}_p\mathcal{P}_n \ ,$$

then the family of metrics $(g_n)_{n \in \mathbb{N}}$ is unique up to a positive real number.

Means in Quantum Info. Geometry

In the classical case: uniqueness of the Fisher information.
... set of distributions

$$\mathcal{M}_n = \left\{ D \in \mathrm{Mat}(n, \mathbb{C}) \;\middle|\; D = D^*, D > 0, \mathrm{Tr}\, D = 1 \right\}.$$

Theorem (*Čencov:*) Assume that for every $n \in \mathbb{N}$ the pair $(\mathcal{P}_n, g_n)$ is a Riemannian-manifold. If for every Markovian map $\kappa : X_n \times X_m \to \mathbb{R}$ the following monotonicity property holds

$$g_{\tilde{\kappa}(p)}(\kappa^*(X), \kappa^*(X)) \leq g_p(X, X) \qquad \forall p \in \mathcal{P}_n, \forall X \in \mathrm{T}_p\mathcal{P}_n \ ,$$

then the family of metrics $(g_n)_{n \in \mathbb{N}}$ is unique up to a positive real number.

Means in Quantum Info. Geometry

In the classical case: uniqueness of the Fisher information.
... set of distributions

$$\mathcal{M}_n = \left\{ D \in \mathrm{Mat}(n, \mathbb{C}) \;\middle|\; D = D^*, D > 0, \mathrm{Tr}\, D = 1 \right\}.$$

(*Petz*:) Assume that for every $n \in \mathbb{N}$ the pair $(\mathcal{P}_n, g_n)$ is a Riemannian-manifold. If for every Markovian map $\kappa : X_n \times X_m \to \mathbb{R}$
the following monotonicity property holds

$$g_{\tilde{\kappa}(p)}(\kappa^*(X), \kappa^*(X)) \leq g_p(X, X) \qquad \forall p \in \mathcal{P}_n, \forall X \in \mathrm{T}_p\mathcal{P}_n \ ,$$

then the family of metrics $(g_n)_{n \in \mathbb{N}}$ is unique up to a positive real number.

Means in Quantum Info. Geometry

In the classical case: uniqueness of the Fisher information.
... set of distributions

$$\mathcal{M}_n = \left\{ D \in \mathrm{Mat}(n, \mathbb{C}) \,\middle|\, D = D^*, D > 0, \mathrm{Tr}\, D = 1 \right\}.$$

(*Petz*:) Assume that for every $n \in \mathbb{N}$ the pair $(\mathcal{M}_n, g_n)$ is a Riemannian-manifold. If for every Markovian map $\kappa : X_n \times X_m \to \mathbb{R}$ the following monotonicity property holds

$$g_{\tilde{\kappa}(p)}(\kappa^*(X), \kappa^*(X)) \leq g_p(X, X) \qquad \forall p \in \mathcal{P}_n, \forall X \in \mathrm{T}_p\mathcal{P}_n \ ,$$

then the family of metrics $(g_n)_{n \in \mathbb{N}}$ is unique up to a positive real number.

Means in Quantum Info. Geometry

In the classical case: uniqueness of the Fisher information.
... set of distributions

$$\mathcal{M}_n = \left\{ D \in \mathrm{Mat}(n, \mathbb{C}) \,\middle|\, D = D^*, D > 0, \mathrm{Tr}\, D = 1 \right\}.$$

($Petz$:) Assume that for every $n \in \mathbb{N}$ the pair $(\mathcal{M}_n, g_n)$ is a Riemannian-manifold. If for every stochastic map T (trace preserving CP. map) the following monotonicity property holds

$$g_{\tilde{\kappa}(p)}(\kappa^*(X), \kappa^*(X)) \leq g_p(X, X) \qquad \forall p \in \mathcal{P}_n, \forall X \in \mathrm{T}_p\mathcal{P}_n \ ,$$

then the family of metrics $(g_n)_{n\in\mathbb{N}}$ is unique up to a positive real number.

Means in Quantum Info. Geometry

In the classical case: uniqueness of the Fisher information.
... set of distributions

$$\mathcal{M}_n = \left\{ D \in \mathrm{Mat}(n, \mathbb{C}) \,\middle|\, D = D^*, D > 0, \mathrm{Tr}\, D = 1 \right\}.$$

(*Petz:*) Assume that for every $n \in \mathbb{N}$ the pair $(\mathcal{M}_n, g_n)$ is a Riemannian-manifold. If for every stochastic map T (trace preserving CP. map) the following monotonicity property holds

$$g_{T(D)}(T(X), T(X)) \leq g_D(X, X) \ \forall D \in \mathcal{M}_n, \forall X \in \mathrm{T}_p\mathcal{M}_n \ ,$$

then the family of metrics $(g_n)_{n \in \mathbb{N}}$ is unique up to a positive real number.

Means in Quantum Info. Geometry

In the classical case: uniqueness of the Fisher information.
... set of distributions

$$\mathcal{M}_n = \left\{ D \in \mathrm{Mat}(n, \mathbb{C}) \,\middle|\, D = D^*, D > 0, \mathrm{Tr}\, D = 1 \right\}.$$

Theorem (*Petz*:) Assume that for every $n \in \mathbb{N}$ the pair $(\mathcal{M}_n, g_n)$ is a Riemannian-manifold. If for every stochastic map T (trace preserving CP. map) the following monotonicity property holds

$$g_{T(D)}(T(X), T(X)) \leq g_D(X, X) \ \forall D \in \mathcal{M}_n, \forall X \in \mathrm{T}_p\mathcal{M}_n \ ,$$

then the family of metrics $(g_n)_{n\in\mathbb{N}}$ is can be expressed as...

Idea of the proof:

Idea of the proof:
monotonicity:

$$g_{T(D)}(T(X), T(X)) \leq g_D(X, X) \ \forall D \in \mathcal{M}_n, \forall X \in \mathrm{T}_p\mathcal{M}_n \ ,$$

Idea of the proof:
monotonicity:

$$g_{T(D)}(T(X), T(X)) \leq g_D(X, X) \ \forall D \in \mathcal{M}_n, \forall X \in \mathrm{T}_p\mathcal{M}_n \ ,$$

$$g_D(X, Y) = \langle X, \mathbf{J}_D^{-1}(Y) \rangle = \mathrm{Tr}(X\mathbf{J}_D^{-1}(Y)),$$ where
$\mathbf{J}_D : \mathrm{Mat}(n, \mathbb{C}) \to \mathrm{Mat}(n, \mathbb{C})$ linear map.

Idea of the proof:
monotonicity:

$$g_{T(D)}(T(X), T(X)) \leq g_D(X, X) \ \forall D \in \mathcal{M}_n, \forall X \in \mathrm{T}_p\mathcal{M}_n \ ,$$

$g_D(X, Y) = \left\langle X, \mathbf{J}_D^{-1}(Y) \right\rangle = \mathrm{Tr}(X\mathbf{J}_D^{-1}(Y))$, where
$\mathbf{J}_D : \mathrm{Mat}(n, \mathbb{C}) \to \mathrm{Mat}(n, \mathbb{C})$ linear map.

$$g_{T(D)}(T(X), T(X)) = \left\langle T(X), \mathbf{J}_{T(D)}^{-1}(T(X)) \right\rangle$$

Idea of the proof:
monotonicity:

$$g_{T(D)}(T(X), T(X)) \leq g_D(X, X) \ \forall D \in \mathcal{M}_n, \forall X \in \mathrm{T}_p\mathcal{M}_n \ ,$$

$$g_D(X,Y) = \left\langle X, \mathbf{J}_D^{-1}(Y) \right\rangle = \mathrm{Tr}(X\mathbf{J}_D^{-1}(Y)),$$ where $\mathbf{J}_D : \mathrm{Mat}(n, \mathbb{C}) \to \mathrm{Mat}(n, \mathbb{C})$ linear map.

$$g_{T(D)}(T(X), T(X)) = \left\langle T(X), \mathbf{J}_{T(D)}^{-1}(T(X)) \right\rangle$$

$$g_{T(D)}(T(X), T(X)) = \left\langle X, T^*\mathbf{J}_{T(D)}^{-1} T(X) \right\rangle$$

Idea of the proof:
monotonicity:

$$g_{T(D)}(T(X),T(X)) \le g_D(X,X) \ \forall D \in \mathcal{M}_n, \forall X \in \mathrm{T}_p\mathcal{M}_n \ ,$$

$g_D(X,Y) = \left\langle X, \mathbf{J}_D^{-1}(Y) \right\rangle = \mathrm{Tr}(X\mathbf{J}_D^{-1}(Y))$, where
$\mathbf{J}_D : \mathrm{Mat}(n,\mathbb{C}) \to \mathrm{Mat}(n,\mathbb{C})$ linear map.

$$g_{T(D)}(T(X),T(X)) = \left\langle T(X), \mathbf{J}_{T(D)}^{-1}(T(X)) \right\rangle$$

$$g_{T(D)}(T(X),T(X)) = \left\langle X, T^*\mathbf{J}_{T(D)}^{-1}T(X) \right\rangle$$

$$g_D(X,X) = \left\langle X, \mathbf{J}_D^{-1}(X) \right\rangle = \left\langle X, T^*\mathbf{J}_D^{-1}T(X) \right\rangle$$

Idea of the proof:
monotonicity:

$$g_{T(D)}(T(X),T(X)) \leq g_D(X,X) \ \forall D \in \mathcal{M}_n, \forall X \in \mathrm{T}_p\mathcal{M}_n \ ,$$

$g_D(X,Y) = \langle X, \mathbf{J}_D^{-1}(Y)\rangle = \mathrm{Tr}(X\mathbf{J}_D^{-1}(Y))$, where
$\mathbf{J}_D : \mathrm{Mat}(n,\mathbb{C}) \to \mathrm{Mat}(n,\mathbb{C})$ linear map.

$$g_{T(D)}(T(X),T(X)) = \left\langle T(X), \mathbf{J}_{T(D)}^{-1}(T(X))\right\rangle$$

$$g_{T(D)}(T(X),T(X)) = \left\langle X, T^*\mathbf{J}_{T(D)}^{-1}T(X)\right\rangle$$

$$g_D(X,X) = \langle X, \mathbf{J}_D^{-1}(X)\rangle = \langle X, T^*\mathbf{J}_D^{-1}T(X)\rangle$$

monotonicity:

Idea of the proof:
monotonicity:

$$g_{T(D)}(T(X),T(X)) \leq g_D(X,X) \ \forall D \in \mathcal{M}_n, \forall X \in \mathrm{T}_p\mathcal{M}_n \ ,$$

$g_D(X,Y) = \langle X, \mathbf{J}_D^{-1}(Y)\rangle = \mathrm{Tr}(X\mathbf{J}_D^{-1}(Y))$, where
$\mathbf{J}_D : \mathrm{Mat}(n,\mathbb{C}) \to \mathrm{Mat}(n,\mathbb{C})$ linear map.

$$g_{T(D)}(T(X),T(X)) = \left\langle T(X), \mathbf{J}_{T(D)}^{-1}(T(X))\right\rangle$$

$$g_{T(D)}(T(X),T(X)) = \left\langle X, T^*\mathbf{J}_{T(D)}^{-1}T(X)\right\rangle$$

$$g_D(X,X) = \langle X, \mathbf{J}_D^{-1}(X)\rangle = \langle X, T^*\mathbf{J}_D^{-1}T(X)\rangle$$

monotonicity:

$$T^*\mathbf{J}_{T(D)}^{-1}T \leq \mathbf{J}_D^{-1}$$

Idea of the proof:
monotonicity:

$$g_{T(D)}(T(X), T(X)) \leq g_D(X, X) \ \forall D \in \mathcal{M}_n, \forall X \in \mathrm{T}_p\mathcal{M}_n \ ,$$

$g_D(X, Y) = \left\langle X, \mathbf{J}_D^{-1}(Y) \right\rangle = \mathrm{Tr}(X\mathbf{J}_D^{-1}(Y))$, where
$\mathbf{J}_D : \mathrm{Mat}(n, \mathbb{C}) \to \mathrm{Mat}(n, \mathbb{C})$ linear map.

$$g_{T(D)}(T(X), T(X)) = \left\langle T(X), \mathbf{J}_{T(D)}^{-1}(T(X)) \right\rangle$$

$$g_{T(D)}(T(X), T(X)) = \left\langle X, T^*\mathbf{J}_{T(D)}^{-1} T(X) \right\rangle$$

$$g_D(X, X) = \left\langle X, \mathbf{J}_D^{-1}(X) \right\rangle = \left\langle X, T^*\mathbf{J}_D^{-1} T(X) \right\rangle$$

monotonicity:

$$T^*\mathbf{J}_{T(D)}^{-1} T \leq \mathbf{J}_D^{-1}$$

$$T\mathbf{J}_D T^* \leq \mathbf{J}_{T(D)}$$

What can $\mathbf{J}_D(X)$ be?

What can $\mathbf{J}_D(X)$ be?

"D can act on left $\varphi_1(D)X$ and on the right $X\varphi_1(D)$"

What can $\mathbf{J}_D(X)$ be?
"D can act on left $\varphi_1(D)X$ and on the right $X\varphi_1(D)$"
in general $\varphi_1(D)X\varphi_2(D)$ gives the idea:

$$\mathbf{J}_D(X) = M(L_D, R_D)(X).$$

Where $L_D(X) = DX$ and $R_D(X) = XD$.

What can $\mathbf{J}_D(X)$ be?
"D can act on left $\varphi_1(D)X$ and on the right $X\varphi_1(D)$"
in general $\varphi_1(D)X\varphi_2(D)$ gives the idea:

$$\mathbf{J}_D(X) = M(L_D, R_D)(X).$$

Where $L_D(X) = DX$ and $R_D(X) = XD$.
We have $M(L_D, R_D) = M(R_D, L_D)$

What can $\mathbf{J}_D(X)$ be?
"D can act on left $\varphi_1(D)X$ and on the right $X\varphi_1(D)$"
in general $\varphi_1(D)X\varphi_2(D)$ gives the idea:

$$\mathbf{J}_D(X) = M(L_D, R_D)(X).$$

Where $L_D(X) = DX$ and $R_D(X) = XD$.
We have $M(L_D, R_D) = M(R_D, L_D)$
and the monotonicity

$$T\mathbf{J}_D T^* \leq \mathbf{J}_{T(D)}$$

gives

$$TM(L_D, R_D)T^* \leq M(TL_DT^*, TR_DT^*).$$

What can $\mathbf{J}_D(X)$ be?
"D can act on left $\varphi_1(D)X$ and on the right $X\varphi_1(D)$"
in general $\varphi_1(D)X\varphi_2(D)$ gives the idea:

$$\mathbf{J}_D(X) = M(L_D, R_D)(X).$$

Where $L_D(X) = DX$ and $R_D(X) = XD$.
We have $M(L_D, R_D) = M(R_D, L_D)$
and the monotonicity

$$T\mathbf{J}_D T^* \leq \mathbf{J}_{T(D)}$$

gives

$$TM(L_D, R_D)T^* \leq M(TL_D T^*, TR_D T^*).$$

M is a mean!

Theorem (*Petz:*) Assume that for every $n \in \mathbb{N}$ the pair $(\mathcal{M}_n, g_n)$ is a Riemannian-manifold. If for every stochastic map T
the following monotonicity property holds

$$g_{T(D)}(T(X), T(X)) \leq g_D(X, X) \; \forall D \in \mathcal{M}_n, \forall X \in \mathrm{T}_p\mathcal{M}_n \; ,$$

then there exists an operator monotone function $f : \mathbb{R}^+ \to \mathbb{R}$ with the property $f(x) = x f(x^{-1})$, such that

$$g_D(X, Y) = \mathrm{Tr}\left(X \left(R_{n,D}^{\frac{1}{2}} f(L_{n,D} R_{n,D}^{-1}) R_{n,D}^{\frac{1}{2}} \right)^{-1}(Y) \right) .$$

Examples:

$$f(x)=\frac{2x}{1+x}:\ g_D^{(\mathrm{LA})}(X,Y)=\frac{1}{2}\,\mathrm{Tr}(XD^{-1}Y+YD^{-1}X)$$

$$f(x)=\frac{x-1}{\log x}:\ g_D^{(\mathrm{KM})}(X,Y)=\mathrm{Tr}\int_0^\infty X(D+t)^{-1}Y(D+t)^{-1}\,\mathrm{d}\,t$$

$$f(x)=\frac{1+x}{2}:\ g_D^{(\mathrm{SM})}(X,Y)=\mathrm{Tr}(XZ_{D,Y}),$$

where $Z_{D,Y}$ is the solution of the equation

$$2Y=DZ_{D,Y}+Z_{D,Y}D.$$

We have the inequality

$$g_D^{(\mathrm{SM})}(X,X)\le g_D^{(f)}(X,X)\le g_D^{(\mathrm{LA})}(X,X).$$

Geometry of Means

Geometrical point of view:
Assume that $(\mathcal{M}, g)$ is a Riemannian manifold. Let us define the *mean* of two arbitrary points $X, Y \in \mathcal{M}$:
– Connect X and Y with a geodesic line γ, such that $\gamma(0) = X$ and $\gamma(1) = Y$.
– Then the mean of X and Y is the point $\gamma(1/2)$.

Then we have:

$$\tilde{M}(X,Y) = \tilde{M}(Y,X)$$

$$\tilde{M}(X,X) = X$$

$$X \prec \tilde{M}(X,Y) \prec Y.$$

Geometry of Means

Geometrical point of view:
Assume that $(\mathcal{M}, g)$ is a Riemannian manifold. Let us define the *mean* of two arbitrary points $X, Y \in \mathcal{M}$:
– Connect X and Y with a geodesic line γ, such that $\gamma(0) = X$ and $\gamma(1) = Y$.
– Then the mean of X and Y is the point $\gamma(1/2)$.

Let denote this mean with $\tilde{M}(X, Y)$.
Then we have:

$$\tilde{M}(X,Y) = \tilde{M}(Y,X)$$

$$\tilde{M}(X,X) = X$$

$$X \prec \tilde{M}(X,Y) \prec Y.$$

Geometry of Means

Geometrical point of view:
Assume that $(\mathcal{M}, g)$ is a Riemannian manifold. Let us define the *mean* of two arbitrary points $X, Y \in \mathcal{M}$:
– Connect X and Y with a geodesic line γ, such that $\gamma(0) = X$ and $\gamma(1) = Y$.
– Then the mean of X and Y is the point $\gamma(1/2)$.

Let denote this mean with $\tilde{M}(X, Y)$.
Then we have:

$$\tilde{M}(X, Y) = \tilde{M}(Y, X)$$

$$\tilde{M}(X, X) = X$$

$$X \prec \tilde{M}(X, Y) \prec Y.$$

Geometry of Means

Geometrical point of view:
Assume that $(\mathcal{M}, g)$ is a Riemannian manifold. Let us define the *mean* of two arbitrary points $X, Y \in \mathcal{M}$:
– Connect X and Y with a geodesic line γ, such that $\gamma(0) = X$ and $\gamma(1) = Y$.
– Then the mean of X and Y is the point $\gamma(1/2)$.

Let denote this mean with $\tilde{M}(X, Y)$.
Then we have:

$$\tilde{M}(X, Y) = \tilde{M}(Y, X)$$

$$\tilde{M}(X, X) = X$$

$$X \prec \tilde{M}(X, Y) \prec Y.$$

Geometry of Means

Geometrical point of view:
Assume that $(\mathcal{M}, g)$ is a Riemannian manifold. Let us define the *mean* of two arbitrary points $X, Y \in \mathcal{M}$:
– Connect X and Y with a geodesic line γ, such that $\gamma(0) = X$ and $\gamma(1) = Y$.
– Then the mean of X and Y is the point $\gamma(1/2)$.

Let denote this mean with $\tilde{M}(X, Y)$.
Then we have:

$$\tilde{M}(X, Y) = \tilde{M}(Y, X)$$

$$\tilde{M}(X, X) = X$$

$$X \prec \tilde{M}(X, Y) \prec Y.$$

Example:

1. $\mathcal{M} := \mathbb{R}^+$, and $g : \mathcal{M} \to \mathbb{R}^+$ smooth function.

At $p \in \mathcal{M}$ the "scalar product" of the "vectors" $x, y \in \mathbb{R}$ is

$$g_p(x, y) = xyg(p).$$

The equation of the geodesic line $\gamma(t)$

$$\ddot{\gamma}(t) + \frac{g'(\gamma(t))}{2g(\gamma(t))} \left(\dot{\gamma}(t)\right)^2 = 0.$$

$$\ddot{\gamma}(t) = 0$$

and its solution $\gamma(t) = C_1 + C_2 t$
which satisfies $\gamma(0) = x$, $\gamma(1) = y$

$$\gamma(t) = x + (y - x)t$$

in this case

$$\tilde{M}(x, y) = \gamma\left(\frac{1}{2}\right) = \frac{x + y}{2}\ .$$

Example:

1. $\mathcal{M} := \mathbb{R}^+$, and $g : \mathcal{M} \to \mathbb{R}^+$ smooth function.
At $p \in \mathcal{M}$ the "scalar product" of the "vectors" $x, y \in \mathbb{R}$ is

$$g_p(x, y) = xyg(p).$$

The equation of the geodesic line $\gamma(t)$

$$\ddot{\gamma}(t) + \frac{g'(\gamma(t))}{2g(\gamma(t))} \big(\dot{\gamma}(t)\big)^2 = 0.$$

Consider the metric $g(t) = 1$. The differential equation:

$$\ddot{\gamma}(t) = 0$$

and its solution $\gamma(t) = C_1 + C_2 t$
which satisfies $\gamma(0) = x$, $\gamma(1) = y$

$$\gamma(t) = x + (y - x)t$$

in this case

$$\tilde{M}(x, y) = \gamma\left(\frac{1}{2}\right) = \frac{x + y}{2} \ .$$

2. Consider the following metric $g(t) = t^{2(p-1)}$: $(p \neq 1)$
The differential equation:

$$\ddot{\gamma}(t) + (p-1)\frac{1}{\gamma(t)}\left(\dot{\gamma}(t)\right)^2 = 0$$

and its solution

$$\begin{cases} \gamma(t) = (C_1 + C_2 t)^{1/p} & \text{if} \quad p \neq 0 \\ \gamma(t) = C_1 C_2^t & \text{if} \quad p = 0 \end{cases}$$

which satisfies $\gamma(0) = x$, $\gamma(1) = y$

$$\begin{cases} \gamma(t) = \sqrt[p]{x^p + (y^p - x^p)t} & \text{if} \quad p \neq 0 \\ \gamma(t) = x\left(\dfrac{y}{x}\right)^t & \text{if} \quad p = 0 \end{cases}$$

in this case

$$\tilde{M}(x,y) = \gamma\left(\frac{1}{2}\right) = \begin{cases} \left(\dfrac{x^p + y^p}{2}\right)^{\frac{1}{p}} & \text{if} \quad p \neq 0 \\ \sqrt{xy} & \text{if} \quad p = 0\ . \end{cases}$$

(*Power-mean* and Geometric Mean.)

Example:

3. $\mathcal{M} := \mathcal{M}_n$, and

$$g_D(X,Y) = \frac{1}{2} \operatorname{Tr} D^{-1} X D^{-1} Y.$$

(Fisher information metric on the space of Gaussian distributions.)

The equation of the geodesic line $\gamma(t) : \mathbb{R} \to \mathcal{M}_n$

$$\ddot{\gamma}(t) - \dot{\gamma}(t)\gamma(t)^{-1}\dot{\gamma}(t) = 0$$

and its solution $\gamma(t) = C_1 C_2^t$

which satisfies $\gamma(0) = X$, $\gamma(1) = Y$

$$\gamma(t) = X^{1/2}(X^{-1/2} Y X^{-1/2})^t X^{1/2}$$

in this case

$$\tilde{M}(X,Y) = X^{1/2}(X^{-1/2} Y X^{-1/2})^{1/2} X^{1/2}\ .$$

Example:

3. $\mathcal{M} := \mathcal{M}_n$, and

$$g_D(X,Y) = \frac{1}{2}\operatorname{Tr} D^{-1} X D^{-1} Y.$$

(Fisher information metric on the space of Gaussian distributions.)

The equation of the geodesic line $\gamma(t) : \mathbb{R} \to \mathcal{M}_n$

$$\ddot{\gamma}(t) - \dot{\gamma}(t)\gamma(t)^{-1}\dot{\gamma}(t) = 0$$

and its solution $\gamma(t) = C_1 C_2^t$

which satisfies $\gamma(0) = X$, $\gamma(1) = Y$

$$\gamma(t) = X^{1/2}(X^{-1/2} Y X^{-1/2})^t X^{1/2}$$

in this case

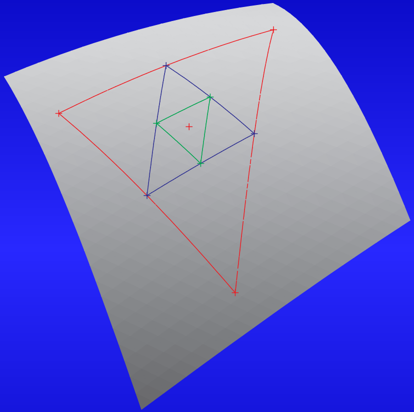
$$\tilde{M}(X,Y) = X^{1/2}(X^{-1/2} Y X^{-1/2})^{1/2} X^{1/2} \ .$$

This gives the geometric mean.

There are two candidates for the geometric mean of three matrices:

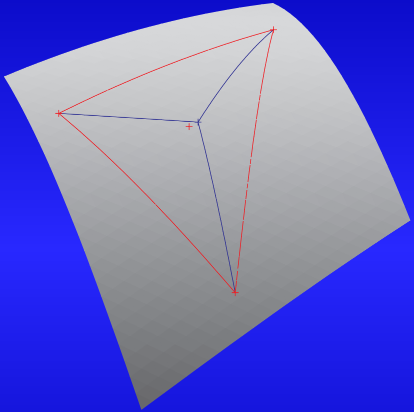
This gives the geometric mean.

There are two candidates for the geometric mean of three matrices:



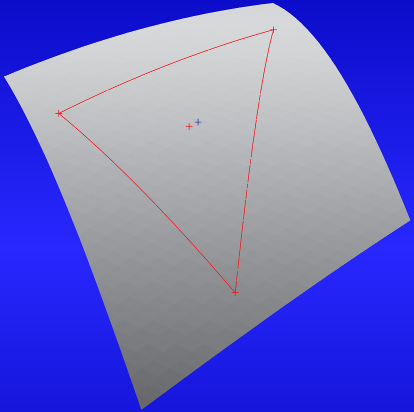
This gives the geometric mean.

There are two candidates for the geometric mean of three matrices:



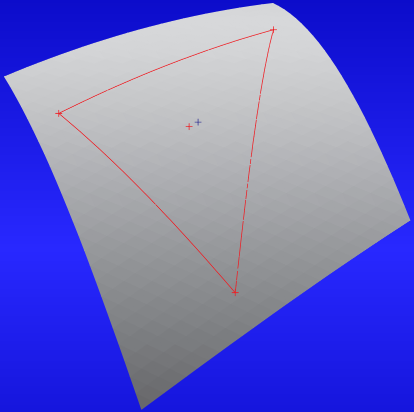
This gives the geometric mean.

There are two candidates for the geometric mean of three matrices:



This gives the geometric mean.

There are two candidates for the geometric mean of three matrices:



$$\log C^{-1}X + \log C^{-1}Y + \log C^{-1}Z = 0$$

Example:

4. $\mathcal{M} := (\mathcal{M}_n, \mathbb{R}^n)$, and

$$g_{D,\underline{u}}((X,\underline{x}),(Y,\underline{y})) = \frac{1}{2}\operatorname{Tr} D^{-1}XD^{-1}Y + \langle \underline{x}, D\underline{y}\rangle.$$

(Fisher information metric on the space of Gaussian distributions.)

The equation of the geodesic line $\gamma_1(t) : \mathbb{R} \to \mathcal{M}_n$, $\gamma_2(t) : \mathbb{R} \to \mathbb{R}^n$

$$\ddot\gamma_1(t) - \dot\gamma_1(t)\gamma_1(t)^{-1}\dot\gamma_1(t) = 0$$
$$\ddot\gamma_2(t) + \gamma_1(t)^{-1}\dot\gamma_1(t)\dot\gamma_2(t) = 0$$

...skip the details...

$$\tilde M((X,\underline{x}),(Y,\underline{y})) = \Bigg(X^{1/2}(X^{-1/2}YX^{-1/2})^{1/2}X^{1/2},$$
$$\underline{x} + \left[\exp\left(\frac{1}{2}X^{-1/2}\log(X^{-1/2}YX^{-1/2})X^{1/2}\right) + I_n\right]^{-1}(\underline{y}-\underline{x})\Bigg).$$

Example:

4. $\mathcal{M} := (\mathcal{M}_n, \mathbb{R}^n)$, and

$$g_{D,\underline{u}}((X,\underline{x}),(Y,\underline{y})) = \frac{1}{2}\operatorname{Tr} D^{-1}XD^{-1}Y + \langle \underline{x}, D\underline{y}\rangle .$$

(Fisher information metric on the space of Gaussian distributions.)

The equation of the geodesic line $\gamma_1(t) : \mathbb{R} \to \mathcal{M}_n$, $\gamma_2(t) : \mathbb{R} \to \mathbb{R}^n$

$$\ddot{\gamma}_1(t) - \dot{\gamma}_1(t)\gamma_1(t)^{-1}\dot{\gamma}_1(t) = 0$$
$$\ddot{\gamma}_2(t) + \gamma_1(t)^{-1}\dot{\gamma}_1(t)\dot{\gamma}_2(t) = 0$$

...skip the details...

$$\tilde{M}((X,\underline{x}),(Y,\underline{y})) = \Bigg(X^{1/2}(X^{-1/2}YX^{-1/2})^{1/2}X^{1/2},$$
$$\underline{x} + \left[\exp\left(\frac{1}{2}X^{-1/2}\log(X^{-1/2}YX^{-1/2})X^{1/2}\right) + I_n\right]^{-1}(\underline{y}-\underline{x})\Bigg).$$

Example:

4. $\mathcal{M} := (\mathcal{M}_n, \mathbb{R}^n)$, and

$$g_{D,\underline{u}}((X,\underline{x}),(Y,\underline{y})) = \frac{1}{2}\operatorname{Tr} D^{-1}XD^{-1}Y + \langle \underline{x}, D\underline{y}\rangle.$$

(Fisher information metric on the space of Gaussian distributions.)

The equation of the geodesic line $\gamma_1(t) : \mathbb{R} \to \mathcal{M}_n$, $\gamma_2(t) : \mathbb{R} \to \mathbb{R}^n$

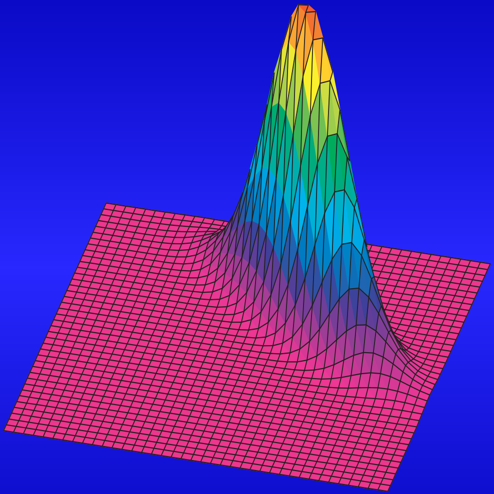
$$\ddot{\gamma}_1(t) - \dot{\gamma}_1(t)\gamma_1(t)^{-1}\dot{\gamma}_1(t) = 0$$
$$\ddot{\gamma}_2(t) + \gamma_1(t)^{-1}\dot{\gamma}_1(t)\dot{\gamma}_2(t) = 0$$

...skip the details...

$$\tilde{M}((X,\underline{x}),(Y,\underline{y})) = \Bigg(X^{1/2}(X^{-1/2}YX^{-1/2})^{1/2}X^{1/2},$$
$$\underline{x} + \left[\exp\left(\frac{1}{2}X^{-1/2}\log(X^{-1/2}YX^{-1/2})X^{1/2}\right) + I_n\right]^{-1}(\underline{y}-\underline{x})\Bigg).$$

Geodesic lines between Gaussian distributions

Geodesic lines between Gaussian distributions

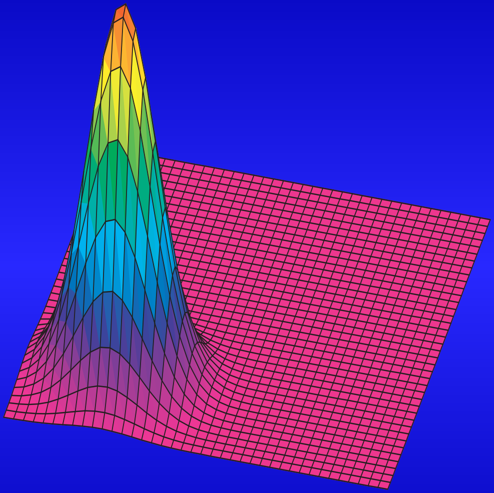


Information Geometry of Matrices and Mean

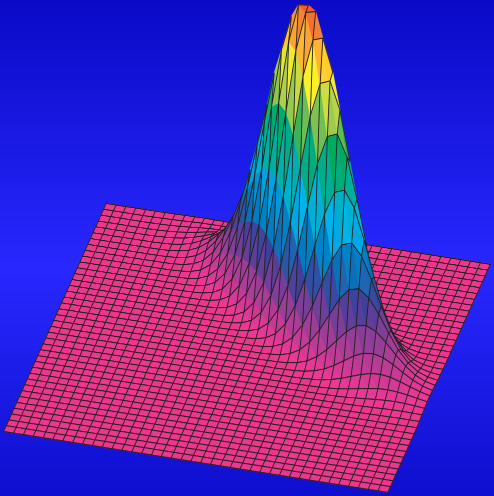
Attila Andai

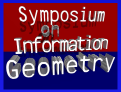
Symposium on Information Geometry

Geodesic lines between Gaussian distributions

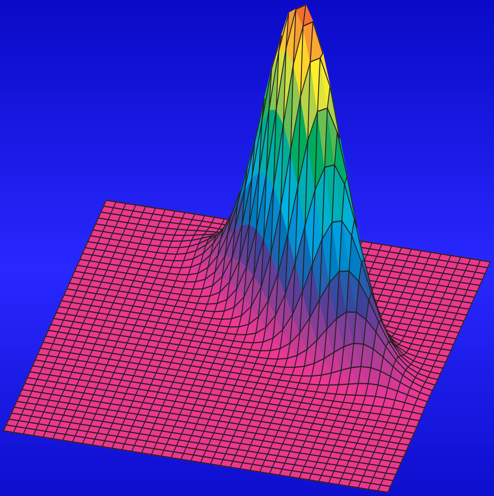


Geodesic lines between Gaussian distributions

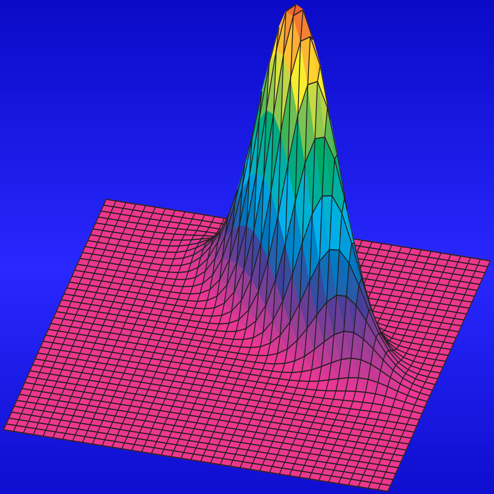




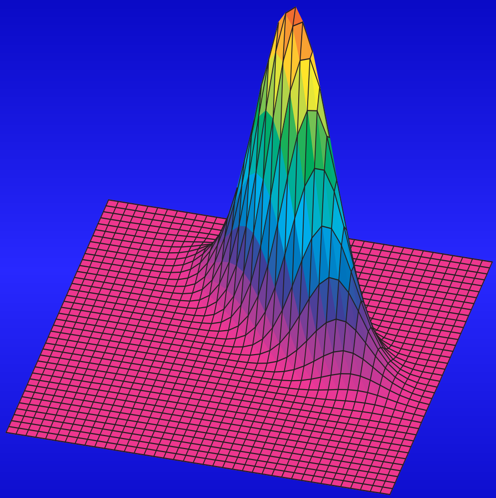
Geodesic lines between Gaussian distributions



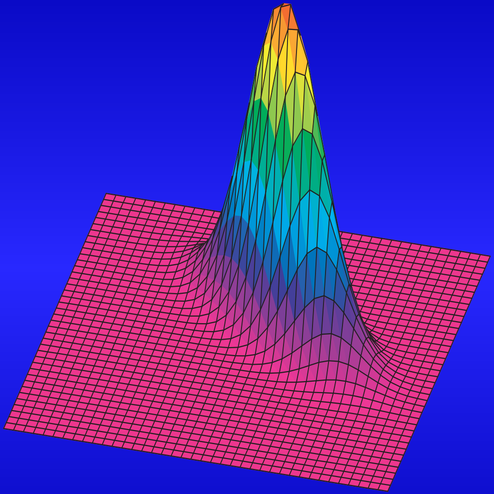
Geodesic lines between Gaussian distributions



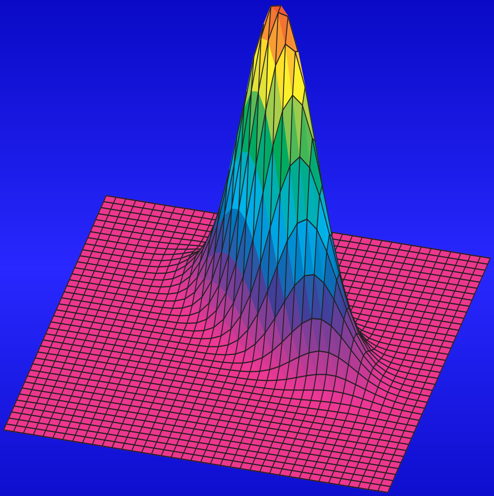
Geodesic lines between Gaussian distributions



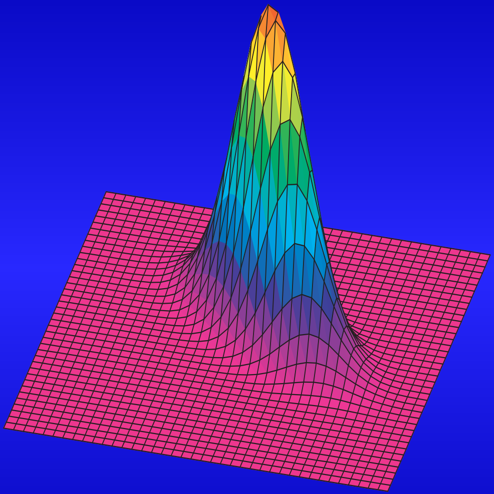
Geodesic lines between Gaussian distributions



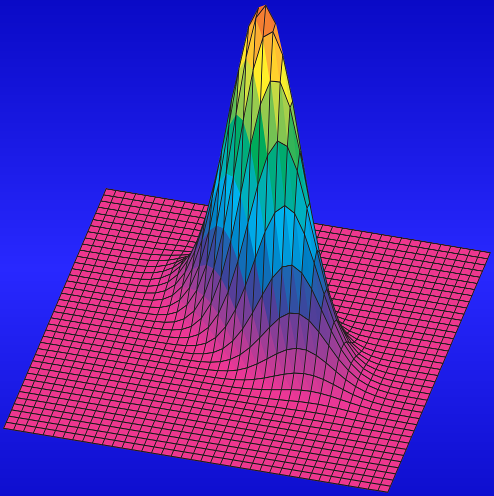
Geodesic lines between Gaussian distributions



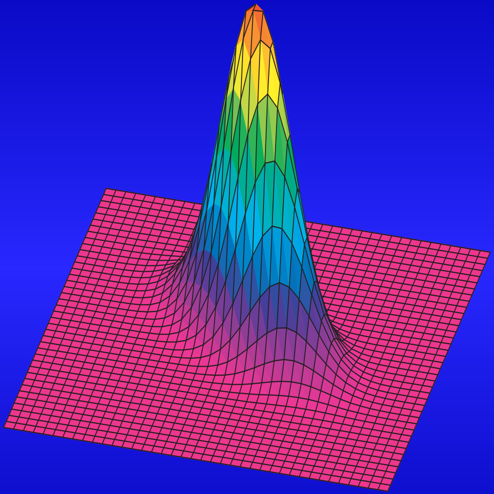
Geodesic lines between Gaussian distributions



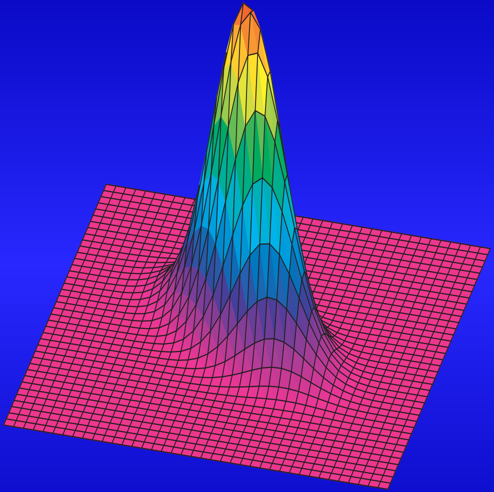
Geodesic lines between Gaussian distributions



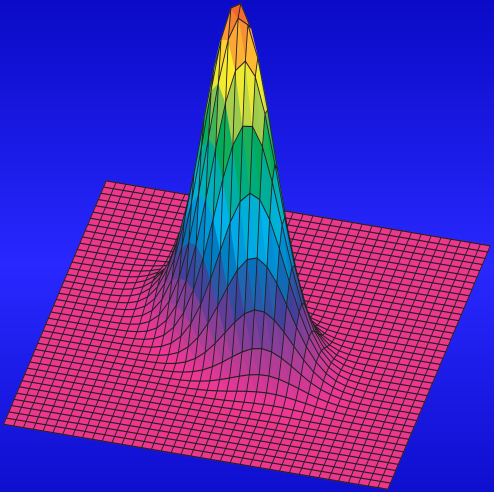
Geodesic lines between Gaussian distributions



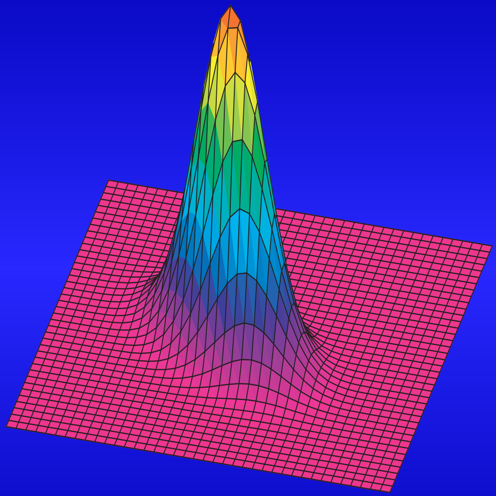
Geodesic lines between Gaussian distributions



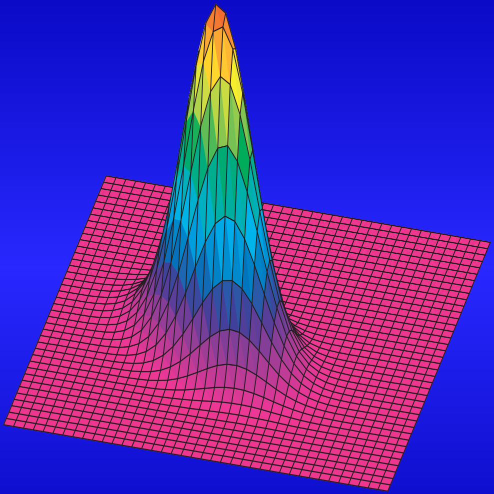
Geodesic lines between Gaussian distributions



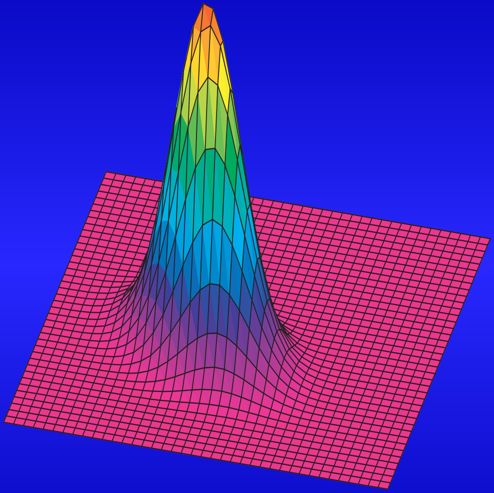
Geodesic lines between Gaussian distributions



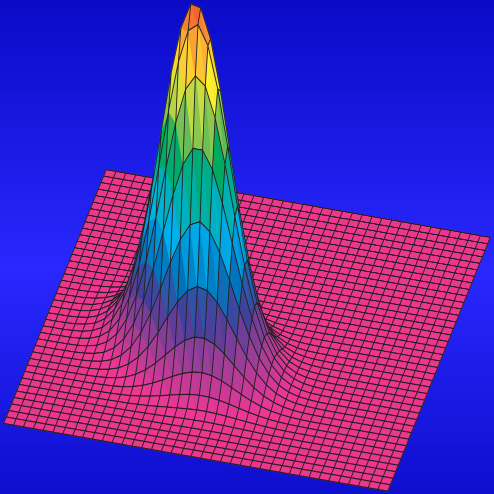
Geodesic lines between Gaussian distributions



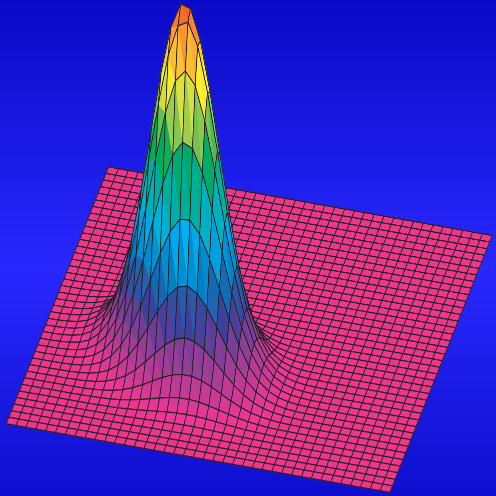
Geodesic lines between Gaussian distributions



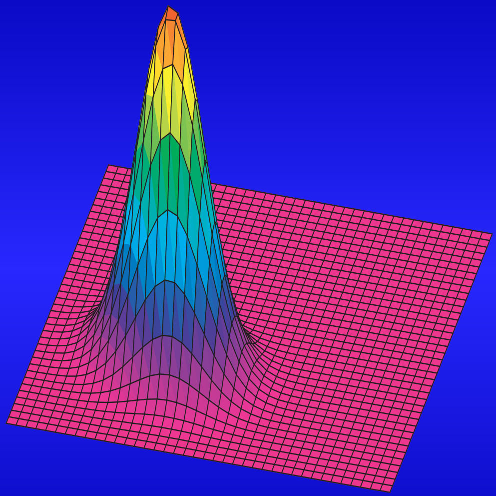
Geodesic lines between Gaussian distributions



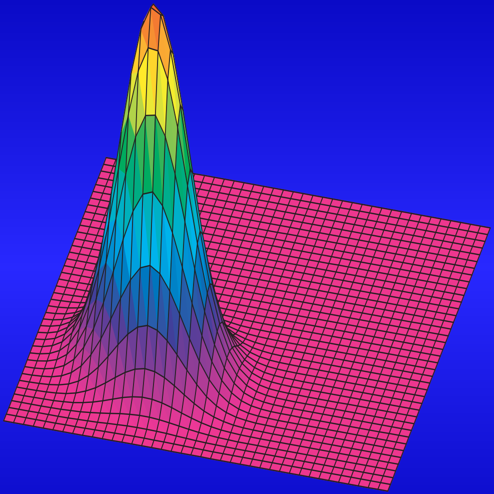
Geodesic lines between Gaussian distributions



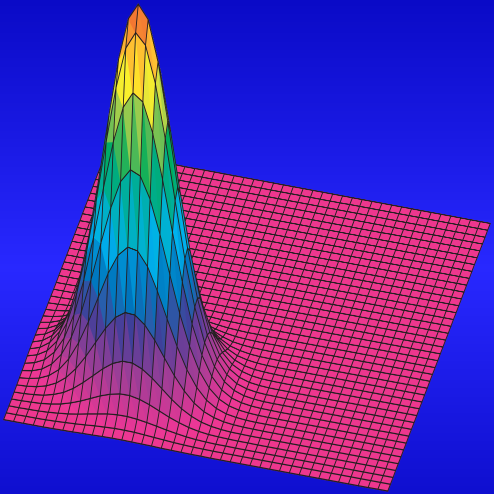
Geodesic lines between Gaussian distributions



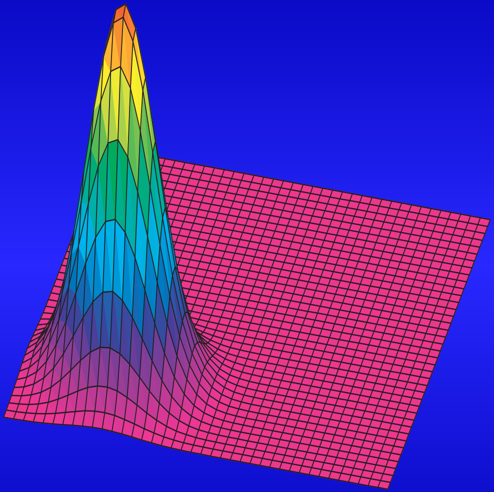
Geodesic lines between Gaussian distributions



Geodesic lines between Gaussian distributions



Geodesic lines between Gaussian distributions



Example:

5. $\mathcal{M} := \mathcal{M}_n$, and

$$g_D(X,Y) = \frac{1}{2}\operatorname{Tr} D^{-2} X D^{-2} Y.$$

The equation of the geodesic line $\gamma(t) : \mathbb{R} \to \mathcal{M}_n$

$$\ddot{\gamma}(t) - 2\dot{\gamma}(t)\gamma(t)^{-1}\dot{\gamma}(t) = 0$$

which satisfies $\gamma(0) = X$, $\gamma(1) = Y$

$$\gamma(t) = (X^{-1} + (Y^{-1} - X^{-1})t)^{-1}$$

in this case

$$\tilde{M}(X,Y) = 2(X^{-1} + Y^{-1})^{-1}\ .$$

Example:

5. $\mathcal{M} := \mathcal{M}_n$, and

$$g_D(X,Y) = \frac{1}{2}\,\mathrm{Tr}\,D^{-2}XD^{-2}Y.$$

The equation of the geodesic line $\gamma(t) : \mathbb{R} \to \mathcal{M}_n$

$$\ddot{\gamma}(t) - 2\dot{\gamma}(t)\gamma(t)^{-1}\dot{\gamma}(t) = 0$$

and its solution $\gamma(t) = (C_1 + C_2 t)^{-1}$

which satisfies $\gamma(0) = X$, $\gamma(1) = Y$

$$\gamma(t) = (X^{-1} + (Y^{-1} - X^{-1})t)^{-1}$$

in this case

$$\tilde{M}(X,Y) = 2(X^{-1} + Y^{-1})^{-1}\ .$$

Questions

1. Which Riemannian metrics guarantee the scaling property:

$$t\tilde{M}(X,Y) = \tilde{M}(tX,tY) \text{ ?}$$

2. How one can find a Riemannian metric for a given mean?

3. Can the geometrical background help to prove that the iteration

$$x_{n+1} := M(y_n, z_n), \; y_{n+1} := M(z_n, x_n), \; z_{n+1} := M(x_n, y_n)$$

is convergent in the space of matrices?

This work was supported by:

Japan Society for the Promotion of Science (JSPS)

Thank you for your attention!

Questions

1. Which Riemannian metrics guarantee the scaling property:

$$t\tilde{M}(X,Y) = \tilde{M}(tX,tY) \ ?$$

2. How one can find a Riemannian metric for a given mean?

3. Can the geometrical background help to prove that the iteration

$$x_{n+1} := M(y_n, z_n), \ y_{n+1} := M(z_n, x_n), \ z_{n+1} := M(x_n, y_n)$$

is convergent in the space of matrices?

This work was supported by:

Japan Society for the Promotion of Science (JSPS)

Thank you for your attention!

Questions

1. Which Riemannian metrics guarantee the scaling property:

$$t\tilde{M}(X,Y) = \tilde{M}(tX,tY) \ ?$$

2. How one can find a Riemannian metric for a given mean?

3. Can the geometrical background help to prove that the iteration

$$x_{n+1} := M(y_n, z_n), \ y_{n+1} := M(z_n, x_n), \ z_{n+1} := M(x_n, y_n)$$

is convergent in the space of matrices?

This work was supported by:

Japan Society for the Promotion of Science (JSPS)

Thank you for your attention!

Questions

1. Which Riemannian metrics guarantee the scaling property:

$$t\tilde{M}(X,Y) = \tilde{M}(tX,tY) \ ?$$

2. How one can find a Riemannian metric for a given mean?
3. Can the geometrical background help to prove that the iteration

$$x_{n+1} := M(y_n, z_n), \ y_{n+1} := M(z_n, x_n), \ z_{n+1} := M(x_n, y_n)$$

is convergent in the space of matrices?

This work was supported by:
Japan Society for the Promotion of Science (JSPS)

Thank you for your attention!

Questions

1. Which Riemannian metrics guarantee the scaling property:

$$t\tilde{M}(X,Y) = \tilde{M}(tX,tY) \ ?$$

2. How one can find a Riemannian metric for a given mean?
3. Can the geometrical background help to prove that the iteration

$$x_{n+1} := M(y_n, z_n),\ y_{n+1} := M(z_n, x_n),\ z_{n+1} := M(x_n, y_n)$$

is convergent in the space of matrices?

This work was supported by:

Japan Society for the Promotion of Science (JSPS)

Thank you for your attention!

Questions

1. Which Riemannian metrics guarantee the scaling property:

$$t\tilde{M}(X,Y) = \tilde{M}(tX,tY) \ ?$$

2. How one can find a Riemannian metric for a given mean?
3. Can the geometrical background help to prove that the iteration

$$x_{n+1} := M(y_n, z_n),\ y_{n+1} := M(z_n, x_n),\ z_{n+1} := M(x_n, y_n)$$

is convergent in the space of matrices?

This work was supported by:

Japan Society for the Promotion of Science (JSPS)

Thank you for your attention!