Semigroups of operator means and generalized Karcher equations

Miklós Pálfia

Department of Mathematics
Sungkyunkwan University
Republic of Korea

July 8, 2013

palfia.miklos@aut.bme.hu

Introduction

Let E be a Hilbert space (possibly infinite dimensional).

- $S(E)$: the space of self-adjoint operators
- $\mathbb{P} \subseteq S(E)$: the cone of positive definite operators

If E is finite dimensional then $\mathbb{P}$ has a Riemannian structure with tangent space $S(E)$:

$$\begin{aligned}
\langle X, Y\rangle_A &= Tr\left\{A^{-1}XA^{-1}Y\right\} \\
d^2(A,B) &= \langle \log_A(B), \log_A(B)\rangle_A = Tr\left\{\log^2(A^{-1/2}BA^{-1/2})\right\} \\
\exp_A(X) &= A^{1/2}\exp(A^{-1/2}XA^{-1/2})A^{1/2} \\
\log_A(B) &= A^{1/2}\log(A^{-1/2}BA^{-1/2})A^{1/2}
\end{aligned}$$

for $X \in S(E)$, $A, B \in \mathbb{P}$.

Geometric structures and operator means

Matrix means as unique minimizers of sum of least squares:
$\omega = (w_1, \dots, w_n) \in \Delta_n$ probability vector, $w_i > 0$, $\sum_{i=1}^n w_i = 1$,
$\mathbb{A} = (A_1, \dots, A_n)$, $A_i \in \mathbb{P}$

$$M(\omega; \mathbb{A}) = \arg\min_{X \in \mathbb{P}} \sum_{i=1}^{n} w_i d^2(X, A_i)$$

1. Karcher mean: $d(A, B) = \sqrt{Tr \left\{ \log^2(A^{-1/2} B A^{-1/2}) \right\}}$, sysmmetric space $GL(E)/U(E)$
2. Arithmetic mean: $d(A, B) = \sqrt{Tr \left\{ (A - B)^2 \right\}}$, subset of the vector space $S(E)$
3. Harmonic mean: $d(A, B) = \sqrt{Tr \left\{ (A^{-1} - B^{-1})^2 \right\}}$, isometric to the above $f(X) = X^{-1}$

In the above cases we have corresponding operator equations of the gradient of $C(X) = \sum_{i=1}^{n} w_i d^2(X, A_i)$:

$$\nabla C(X) = -2 \sum_{i=1}^{n} w_i \log_X(A_i),$$

where $\log_X$ is the logarithm map, the inverse of the exponential map $\exp_X$.
$M(\omega; \mathbb{A})$ is the unique solution of the corresponding gradient equation $\nabla C(X) = 0$, (Karcher equation)

1. Karcher mean $\Lambda(\omega; \mathbb{A})$:
   $\log_X(Y) = X^{1/2} \log(X^{-1/2} Y X^{-1/2}) X^{1/2}$, unknown closed expression for the solution (the Karcher equation)
2. Arithmetic mean: $\log_X(Y) = X^{1/2}(X^{-1/2} Y X^{-1/2} - I) X^{1/2}$, the solution is $\sum_{i=1}^{n} w_i A_i$
3. Harmonic mean: $\log_X(Y) = X^{1/2}(I - X^{1/2} Y^{-1} X^{1/2}) X^{1/2}$, the solution is $\left(\sum_{i=1}^{n} w_i A_i^{-1}\right)^{-1}$

in all above cases

$$\log_X(Y) = X^{1/2} \log_I(X^{-1/2} Y X^{-1/2}) X^{1/2}$$

where $\log_I$ is an operator monotone function on $(0, \infty)$ with $\log_I(1) = 0$, $\log_I'(1) = 1$.

1. the arithmetic mean: $\log_I(x) = x - 1$
2. the harmonic mean: $\log_I(x) = 1 - x^{-1}$
3. the geometric mean: $\log_I(x) = \log(x)$

General class of geometric structures on $\mathbb{P}$

Affinely connected structures on $\mathbb{P}$ with $t \in \mathbb{R}$:

$$\nabla_{X_p} Y_p = DY[p][X_p] - \frac{1-t}{2}\left(X_p p^{-1} Y_p + Y_p p^{-1} X_p\right)$$

$S(E)$ is the tangent space at every $p \in \mathbb{P}$, i.e. $X_p, Y_p \in S(E)$

Theorem

The exponential and logarithm map of the affine connections are given in the form

$$\exp_p(Y) = p^{1/2} \exp_I\left(p^{-1/2} Y p^{-1/2}\right) p^{1/2}$$
$$\log_p(X) = p^{1/2} \log_I\left(p^{-1/2} X p^{-1/2}\right) p^{1/2},$$

where

$$\exp_I(X) = \begin{cases} [tX + 1]^{\frac{1}{t}} & \text{if } t \neq 0, \\ \exp(X) & \text{else,} \end{cases}$$

$$\log_I(X) = \begin{cases} \frac{X^t - 1}{t} & \text{if } t \neq 0, \\ \log(X) & \text{else.} \end{cases}$$

Corresponding gradient equations can be defined:

$$\sum_{i=1}^{n} w_i \log_X(A_i) = 0$$

1. case $t = -1$: harmonic mean
2. case $t = 0$: Karcher mean
3. case $t = 1$: arithmetic mean

Matrix power means $P_t(\omega; \mathbb{A})$

Theorem

$\sum_{i=1}^{n} w_i \log_X(A_i) = 0$ *is equivalent to* $X = \sum_{i=1}^{n} w_i (X \#_t A_i)$*, with* $A \#_t B = A^{1/2} \left(A^{-1/2} B A^{-1/2}\right)^t A^{1/2}$.

Definition (Matrix power means)

For $t \in (0, 1]$, denote $P_t(\omega; \mathbb{A})$ the unique solution of

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

For $t \in [-1, 0)$, we define $P_t(\omega; \mathbb{A}) = P_{-t}(\omega; \mathbb{A}^{-1})^{-1}$, where $\mathbb{A}^{-1} = (A_1^{-1}, \ldots, A_n^{-1})$.

The Thompson metric on $\mathbb{P}$ is defined by

$$d_\infty(A, B) = ||\log(A^{-1/2}BA^{-1/2})||_\infty,$$

where $||X||_\infty$ denotes the spectral norm of X. $(\mathbb{P}, d_\infty)$ is complete. Let $f(X) = \sum_{i=1}^n w_i(X\#_t A_i)$. Then

$$d_\infty(f(X), f(Y)) \leq (1-t)d_\infty(X, Y)$$

for $0 < t \leq 1$, which follows from that $g(X) = X\#_t B$ is a strict contraction.

Some notations:
for $\mathbb{A} = (A_1, \ldots, A_n) \in \mathbb{P}^n$,
$M \in \mathrm{GL}(E), \mathbf{a} = (a_1, \ldots, a_n) \in (0, \infty)^n, \omega = (w_1, \ldots, w_n) \in \Delta_n$,
and for a permutation σ on n-letters, let

$$\begin{aligned}
M\mathbb{A}M^* &= (MA_1M^*, \ldots, MA_nM^*), \quad \mathbb{A}_\sigma = (A_{\sigma(1)}, \ldots, A_{\sigma(n)}), \\
\mathbb{A}^{(k)} &= (\underbrace{\mathbb{A}, \ldots, \mathbb{A}}_{k}) \in \mathbb{P}^{nk}, \quad \omega^{(k)} = \frac{1}{k}(\underbrace{\omega, \ldots, \omega}_{k}) \in \Delta_{nk}, \\
\mathbf{a}^t &= (a_1^t, \ldots, a_n^t), \quad \omega \odot \mathbf{a} = \frac{1}{\sum_{i=1}^n w_i a_i}(w_1 a_1, \ldots, w_n a_n) \in \Delta_n, \\
\hat{\omega} &= \frac{1}{1 - w_n}(w_1, \ldots, w_{n-1}) \in \Delta_{n-1}, \quad \mathbf{a} \cdot \mathbb{A} = (a_1 A_1, \ldots, a_n A_n)
\end{aligned}$$

Proposition

Let $\mathbb{A} = (A_1, \ldots, A_n), \mathbb{B} = (B_1, \ldots, B_n) \in \mathbb{P}^n, \omega \in \Delta_n, \mathbf{a} = (a_1, \ldots, a_n) \in (0, \infty)^n$ *and let* $s, t \in [-1, 1] \setminus \{0\}$.

(1) $P_t(\omega; \mathbb{A}) = \left(\sum_{i=1}^n w_i A_i^t\right)^{\frac{1}{t}}$ *if the* A_i*'s commute*;

(2) $P_t(\omega; \mathbf{a} \cdot \mathbb{A}) = \left(\sum_{i=1}^n w_i a_i^t\right)^{\frac{1}{t}} P_t(\omega \odot \mathbf{a}^t; \mathbb{A})$;

(3) $P_t(\omega_\sigma; \mathbb{A}_\sigma) = P_t(\omega; \mathbb{A})$ *for any permutation* σ;

(4) $P_t(\omega; \mathbb{A}) \leq P_t(\omega; \mathbb{B})$ *if* $A_i \leq B_i$ *for all* $i = 1, 2, \ldots, n$;

(5) $d_\infty(P_t(\omega; \mathbb{A}), P_t(\omega; \mathbb{B})) \leq \max_{1 \leq i \leq n}\{d_\infty(A_i, B_i)\}$;

(6) $(1-u)P_{|t|}(\omega; \mathbb{A}) + uP_{|t|}(\omega; \mathbb{B}) \leq P_{|t|}(\omega; (1-u)\mathbb{A} + u\mathbb{B})$ *for any* $u \in [0, 1]$;

(7) $P_t(\omega; M\mathbb{A}M^*) = MP_t(\omega; \mathbb{A})M^*$ *for any invertible matrix* M;

(8) $P_t(\omega; \mathbb{A}^{-1})^{-1} = P_{-t}(\omega; \mathbb{A})$;

(9) $\mathrm{Det}(P_{-|t|}(\omega; \mathbb{A})) \leq \prod_{i=1}^n \mathrm{Det}(A_i)^{w_i} \leq \mathrm{Det}(P_{|t|}(\omega; \mathbb{A}))$;

(10) $(\sum_{i=1}^{n} w_i A_i^{-1})^{-1} \leq P_t(\omega; \mathbb{A}) \leq \sum_{i=1}^{n} w_i A_i$;

(11) $P_t(\omega^{(k)}; \mathbb{A}^{(k)}) = P_t(\omega; \mathbb{A})$ for any $k \in \mathbb{N}$;

(12) $P_t(\omega; A_1, \ldots, A_{n-1}, X) = X$ if and only if $X = P_t(\hat{\omega}; A_1, \ldots, A_{n-1})$. In particular, $P_t(A_1, \ldots, A_n, X) = X$ if and only if $X = P_t(A_1, \ldots, A_n)$;

(13) For $s \in (0, 1]$, $P_t(\omega; X\#_s A_1, \ldots, X\#_s A_n) = X$ if and only if $X = P_{st}(\omega; \mathbb{A})$;

(14) If $t \in (0, 1]$, then $\Phi(P_t(\omega; \mathbb{A})) \leq P_t(\omega; \Phi(\mathbb{A}))$ for any positive unital linear map Φ, where $\Phi(\mathbb{A}) = (\Phi(A_1), \ldots, \Phi(A_n))$. If $t \in [-1, 0)$, then $\Phi(P_t(\omega; \mathbb{A})) \geq P_t(\omega; \Phi(\mathbb{A}))$ for any strictly positive unital linear map Φ; and

(15) For any unitarily invariant norm $|||\cdot|||$ and $t \in (0,1]$,

$$|||P_t(\omega;\mathbb{A})||| \leq \left[\sum_{i=1}^{n} w_i |||A_i|||^t\right]^{\frac{1}{t}}$$

and

$$|||P_{-t}(\omega;\mathbb{A})||| \geq \left[\sum_{i=1}^{n} w_i |||A_i^{-1}|||^t\right]^{-\frac{1}{t}}.$$

The problem of finding an explicit form of $P_t(\omega; \mathbb{A})$ is non-trivial, except for $n = 2$.

Proposition

For $t \in (0, 1]$, *we have*

$$\begin{aligned} P_t(w_1, w_2; A, B) &= A \#_{\frac{1}{t}} [w_1 A + w_2 (A \#_t B)] \\ &= A^{1/2} \left(w_1 I + w_2 (A^{-1/2} B A^{-1/2})^t \right)^{\frac{1}{t}} A^{1/2}. \end{aligned}$$

In particular, $P_{\frac{1}{n}}(w_1, w_2; A, B) = \sum_{k=0}^{n} \binom{n}{k} w_1^k w_2^{n-k} (B \#_{\frac{k}{n}} A)$.

Note: $P_t(A, B)$ coincides with the operator mean arising from $f_t : (0, \infty) \to (0, \infty)$ defined by $f_t(x) = \left(\frac{x^t + 1}{2} \right)^{\frac{1}{t}}$.

Theorem

$P_t(A, B)$ is the geodesic connecting A, B with respect to the affine connections

$$\nabla_{X_p} Y_p = DY[p][X_p] - \frac{1-t}{2}\left(X_p p^{-1} Y_p + Y_p p^{-1} X_p\right)$$

given on $\mathbb{P}$ where $S(E)$ is the tangent space at every $p \in \mathbb{P}$ ($X_p, Y_p \in S(E)$).

Plausible question: What happens when $t \to 0$ with $P_t(A, B)$?

In fact $P_t(\omega; \mathbb{A})$ is continuous for all $t \in [-1, 1]$, moreover

Theorem

$$\lim_{t \to 0} P_t(\omega; \mathbb{A}) = \Lambda(\omega; \mathbb{A})$$

Corollary

(P1) (*Consistency with scalars*) $\Lambda(\omega; \mathbb{A}) = A_1^{w_1} \cdots A_n^{w_n}$ *if the* A_i*'s commute;*

(P2) (*Joint homogeneity*) $\Lambda(\omega; a_1 A_1, \dots, a_n A_n) = a_1^{w_1} \cdots a_n^{w_n} \Lambda(\omega; \mathbb{A})$;

(P3) (*Permutation invariance*) $\Lambda(\omega_\sigma; \mathbb{A}_\sigma) = \Lambda(\omega; \mathbb{A})$, *where* $\omega_\sigma = (w_{\sigma(1)}, \dots, w_{\sigma(n)})$;

(P4) (*Monotonicity*) *If* $B_i \leq A_i$ *for all* $1 \leq i \leq n$, *then* $\Lambda(\omega; \mathbb{B}) \leq \Lambda(\omega; \mathbb{A})$;

(P5) $d_\infty(\Lambda(\omega;\mathbb{A}),\Lambda(\omega;\mathbb{B}))\leq \max_{1\leq i\leq n}\{d_\infty(A_i,B_i)\}$;

(P6) (Invariancy) $\Lambda(\omega;M^*\mathbb{A}M)=M^*\Lambda(\omega;\mathbb{A})M$ for any invertible M;

(P7) (Joint concavity) $\Lambda(\omega;(1-u)\mathbb{A}+u\mathbb{B})\geq(1-u)\Lambda(\omega;\mathbb{A})+u\Lambda(\omega;\mathbb{B})$ for $0\leq u\leq 1$;

(P8) (Self-duality) $\Lambda(\omega;A_1^{-1},\dots,A_n^{-1})^{-1}=\Lambda(\omega;A_1,\dots,A_n)$;

(P9) (Determinant identity) $\mathrm{Det}\Lambda(\omega;\mathbb{A})=\prod_{i=1}^n(\mathrm{Det}A_i)^{w_i}$; and

(P10) (AGH weighted mean inequalities) $(\sum_{i=1}^n w_iA_i^{-1})^{-1}\leq\Lambda(\omega;\mathbb{A})\leq\sum_{i=1}^n w_iA_i$.

(P11) $\Lambda(\omega^{(k)};\mathbb{A}^{(k)})=\Lambda(\omega;\mathbb{A})$ for any $k\in\mathbb{N}$;

(P12) $\Lambda(\omega;A_1,\dots,A_{n-1},X)=X$ if and only if $X=\Lambda(\hat{\omega};A_1,\dots,A_{n-1})$. In particular, $\Lambda(A_1,\dots,A_n,X)=X$ if and only if $X=\Lambda(A_1,\dots,A_n)$;

(P13) for any $t\in(0,1]$, $X=\Lambda(\omega;X\#_tA_1,\dots,X\#_tA_n)$ if and only if $X=\Lambda(\omega;\mathbb{A})$;

(P14) $\Phi(\Lambda(\omega; \mathbb{A})) \leq \Lambda(\omega; \Phi(\mathbb{A}))$ for any positive unital linear map Φ. If Φ is strictly positive, then $\Phi(\Lambda(\omega; \mathbb{A})) = \Lambda(\omega; \Phi(\mathbb{A}))$.

(P15) $\prod_{i=1}^{n} |||A_i^{-1}|||^{-w_i} \leq |||\Lambda(\omega; \mathbb{A})||| \leq \prod_{i=1}^{n} |||A_i|||^{w_i}$ for any unitarily invariant norm $|||\cdot|||$.

About the power means $P_t(\omega; \mathbb{A})$ refer to:

Y. Lim and M. Pálfia, *Matrix power means and the Karcher mean*, Journal of Functional Analysis, 262:4 (2012), pp. 1498-1514.

More about these affine connections

$$\nabla_{X_p} Y_p = DY[p][X_p] - \frac{1-t}{2}\left(X_p p^{-1} Y_p + Y_p p^{-1} X_p\right)$$

Plausible question: Are they metric connections?
After studying the holonomy groups we obtain:

Theorem

The smooth manifolds with the above affine connections are metric iff

1. $\dim E = 1, 2$, *t arbitrary: the metric is symplectic over $\mathbb{R}$ or $\mathbb{C}$ accordingly.*
2. $\dim E \geq 3$, $t = -1, 0, 1$*: the metric is Riemannian.*

Furthermore another "no go" result:

Theorem

There are no other operator means that are solutions of gradient equations corresponding to affinely connected geometric structures.

So the only matrix means which are midpoint operations on Riemannian manifolds are the arithmetic, harmonic and geometric means.

- If E is infinite dimensional, even no Riemannian metrics exist at all!

From now on we **no longer** assume that E is finite dimensional.

Representing functions and logarithm maps

Definition (Operator mean in the sense of Kubo-Ando)

A two-variable function $M:\mathbb{P}^2 \mapsto \mathbb{P}$ is an operator mean if

1. $M(I, I) = I$,
2. if $A \leq A'$ and $B \leq B'$, then $M(A, B) \leq M(A', B')$,
3. $CM(A, B)C \leq M(CAC, CBC)$ for all hermitian C,
4. if $A_n \downarrow A$ and $B_n \downarrow B$ then $M(A_n, B_n) \downarrow M(A, B)$.

M is uniquely represented as

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

where $f(1) = 1$ (normalized) and $f : (0, \infty) \to (0, \infty)$ operator monotone
the set of these functions are denoted by $\mathfrak{M}$

Each such f has representation

$$f(x) = \int_{[0,\infty]} \frac{(1+t)x}{x+t} d\mu(t)$$

where μ is a probability measure on $[0, \infty]$.
By change of variables $t = \frac{s}{1-s}$ we can write equivalently

$$f(x) = \int_{[0,1]} (1 - s + sx^{-1})^{-1} d\nu(s)$$

where $d\nu(s) = d\mu\left(\frac{s}{1-s}\right)$.

I.e.

Theorem

Each operator mean $M(A, B)$ can be uniquely written as

$$M(A, B) = \int_{[0,1]} [(1-s)A^{-1} + sB^{-1}]^{-1} d\nu(s),$$

i.e. as the closure of convex combinations of weighted harmonic means, also $f'(1) = \int_{[0,1]} s d\nu(s)$, so $0 \leq f'(1) \leq 1$ and $f'(1) = 0$ iff $f(x) = 1$ and $f'(1) = 1$ iff $f(x) = x$.

1. weighted arithmetic mean: $f_s(x) = 1 - s + sx$, $f_s'(1) = s$
2. weighted harmonic mean: $f_s(x) = (1 - s + sx^{-1})^{-1}$, $f_s'(1) = s$
3. weighted geometric mean: $f_s(x) = x^s$, $f_s'(1) = s$
4. weighted power means $t \in [-1, 1]$: $f_s(x) = (1 - s + sx^t)^{1/t}$, $f_s'(1) = s$

$f'(1)$ acts as a "weight", $\mathfrak{P}(t) = \{f \in \mathfrak{M} : f'(1) = t\}$.

$\mathfrak{L}$ denotes the convex cone of operator monotone functions g on $(0, \infty)$, s.t. $g(1) = 0$, $g'(1) = 1$.

Theorem

Let $f \in \mathfrak{M}$ and $f'(1) = t_0 \in (0, 1)$. Then there exists a unique function $\log_I \in \mathfrak{L}$ which fulfills the functional equation (Schröder's equation)

$$\log_I(f(z)) = t_0 \log_I(z)$$

for all $z \in \mathbb{C} \backslash [-\infty, 0]$, moreover

$$\log_I(z) = \lim_{n \to \infty} \frac{f^{\circ n}(z) - 1}{f'(1)^n}$$

(Königs' function for Schröder's equation) where $f^{\circ n}(z) = f(f^{\circ(n-1)}(z))$.

Given $f \in \mathfrak{M}$ with $f'(1) = t_0 \in (0,1)$ we can uniquely write it as

$$f(z) = \exp_I(t_0 \log_I(z))$$

where $\log_I$ is called the *logarithm map* corresponding to f, similarly its inverse $\exp_I$ is called the *exponential map*.

Lemma

For all $\log_I \in \mathfrak{L}$ *we have*

$$1 - x^{-1} \leq \log_I(x) \leq x - 1.$$

One parameter family of functions in $\mathfrak{M}$:

$$f_t(z) = \exp_I(t \log_I(z))$$

Theorem

Let $\log_I \in \mathfrak{L}$. *Then* $f_t \in \mathfrak{P}(t)$ *for all* $t \in (0,1)$ *if and only if* $\log_I(z)$ *has no ramification point in* $\mathbb{H}^+ = \{z \in \mathbb{C} : \Im z > 0\}$ *(i.e.* $\log_I'(z) \neq 0$*).*

Proposition

Let $f_{t_0} \in \mathfrak{P}(t_0)$ *with* $f_{t_0}(z) = \exp_I(t_0 \log_I(z))$, $\log_I \in \mathfrak{L}$. *Then* $f_t \in \mathfrak{P}(t)$ *for all* $0 < t = t_0^j$, *where* $j = 1, 2, 3, \ldots$.

- In general one can assure that, if for a given $\log_I \in \mathfrak{L}$ with ramification points, $t \log_I(\mathbb{H}^+)$ avoids the image of the ramification points (of $\log_I$) under $\log_I$ in $\mathbb{H}^+$, then $f_t \in \mathfrak{P}(t)$.

Proposition

Let $f_t \in \mathfrak{P}(t)$ *and* $\log_I \in \mathfrak{L}$ *its corresponding logarithm map. Then* z_0 *is a ramification point of* f_t *iff it is a ramification point of* $\log_I$.

By the above any operator mean $M(A, B)$ is given in the form

$$M(A, B) = A^{1/2} \exp_I(t_0 \log_I(A^{-1/2} B A^{-1/2})) A^{1/2}$$

and has a corresponding one parameter family of operator means for $0 < t = t_0^j$, $j = 1, 2, \ldots$:

$$M_t(A, B) = A^{1/2} \exp_I(t \log_I(A^{-1/2} B A^{-1/2})) A^{1/2}.$$

- Moreover if $\log_I$ has no ramification point, then the above for $t \in [0, 1]$ is also a (Kubo-Ando) operator mean.

The logarithm maps of some means:

1. the arithmetic mean: $\log_I(x) = x - 1$
2. the harmonic mean: $\log_I(x) = 1 - x^{-1}$
3. the geometric mean: $\log_I(x) = \log(x)$
4. matrix power means: $\log_I(x) = \frac{x^t - 1}{t}$

Contraction property of operator means

The operator equation defining the matrix power means is

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

it has unique solution in $\mathbb{P}$ since $f(X) = X \#_t A$ is a strict contraction with respect to $d_\infty(\cdot, \cdot)$.
The analogue equation for an operator mean M is

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

The key result:

Theorem (Contraction Theorem for Kubo-Ando means)

Let M be an operator mean with representing function $g(x) \in \mathfrak{M}$ and $f(X) = M(A, X)$. If M is not the right trivial mean (i.e. $g(x) \neq x$) then the mapping $f(X)$ is a strict contraction on $\overline{B}_A(r)$ for all $r < \infty$, i.e. there exists $0 < \rho_r < 1$ such that

$$d_\infty(f(X), f(Y)) \leq \rho_r d_\infty(X, Y)$$

for all $X, Y \in \overline{B}_A(r)$.
If M is the right trivial mean (i.e. $g(x) = x$) then $f(X)$ is nonexpansive on $\mathbb{P}$, that is

$$d_\infty(f(X), f(Y)) \leq d_\infty(X, Y)$$

for all $A, X, Y \in \mathbb{P}$.

Some background of the theorem and the sketch of its proof:

Lemma (Lawson-Lim 2008)

1. $d_\infty(rA, rB) = d_\infty(A, B)$ *for any* $r > 0$,
2. $d_\infty(A^{-1}, B^{-1}) = d_\infty(A, B)$,
3. $d_\infty(MAM^*, MBM^*) = d_\infty(A, B)$ *for all* $M \in \mathrm{GL}(E)$ *where* $\mathrm{GL}(E)$ *denotes the Banach-Lie group of all invertible bounded linear operators on* E,
4. $d_\infty(\sum_{i=1}^k t_i A_i, \sum_{i=1}^k t_i B_i) \leq \max_{1\leq i\leq k} d_\infty(A_i, B_i)$ *where* $t_i > 0$,
5. $e^{-d_\infty(A,B)} B \leq A \leq e^{d_\infty(A,B)} B$ *and* $e^{-d_\infty(A,B)} A \leq B \leq e^{d_\infty(A,B)} A$.

A new weighted version of property 4.:

Proposition

Let $A_i, B_i \in \mathbb{P}$, $1 \leq i \leq k$ *and suppose that* $d_\infty(A_m, B_m) \geq d_\infty(A_i, B_i)$. *Then we have*

$$e^{d_\infty\left(\sum_{i=1}^k A_i, \sum_{i=1}^k B_i\right)} \leq \max\left\{ \frac{\sum_{i=1}^k e^{d_\infty(A_i,B_i)} e^{-d_\infty(A_m,A_i)}}{\sum_{i=1}^k e^{-d_\infty(A_m,A_i)}}, \frac{\sum_{i=1}^k e^{d_\infty(A_i,B_i)} e^{-d_\infty(B_m,B_i)}}{\sum_{i=1}^k e^{-d_\infty(B_m,B_i)}} \right\}.$$

Notation: $\overline{B}_A(r) = \{X \in \mathbb{P} : d_\infty(A, X) \leq r\}$

The new weighted property 4. leads to

Lemma

Let $a, b > 0$ be real numbers. Then the mappings

$$h^{+}_{a,b,A}(B) = aA + bB \text{ and } h^{-}_{a,b,A}(B) = (aA^{-1} + bB^{-1})^{-1}$$

are strict contractions on every $\overline{B}_A(r)$ for all $r < \infty$, i.e. for all $X, Y \in \overline{B}_A(r)$

$$d_\infty(h^{\pm}_{a,b,A}(X), h^{\pm}_{a,b,A}(Y)) \leq \rho d_\infty(X, Y)$$

where

$$\rho = \frac{\log \frac{e^{-r-|\log a - \log b|} + e^{2r}}{e^{-r-|\log a - \log b|} + 1}}{2r} < 1.$$

Putting it together:
Each operator mean $M(A, B)$ can be uniquely written as

$$M(A, B) = \int_{[0,1]} [(1-s)A^{-1} + sB^{-1}]^{-1} d\nu(s),$$

i.e. as the closure of convex combinations of weighted harmonic means.
We split the integral to the sum of integrals over mutually disjoint intervals $I_1 = [0, a)$, $I_2 = [a, 1-a]$, $I_3 = (1-a, 1]$ for some $a \in (0, 1/2)$ such that ν has nonzero mass on I_2.
Notation:

$$f_i(X) = \int_{I_i} [(1-s)I + sX^{-1}]^{-1} d\nu(s)$$

By Krein-Milman and weak $*$-compactness of the set of probability measures

$$f_{i,k}(X) = \int_{I_i} [(1-s)I + sX^{-1}]^{-1} d\nu_k(s)$$

for $i = 1, 2, 3$, i.e.

$$f(X) = \lim_k f_{1,k}(X) + f_{2,k}(X) + f_{3,k}(X)$$

also all $f_{i,k}$ are nonexpansive:

$$d_\infty(f_{i,k}(X), f_{i,k}(Y)) \leq \rho_i d_\infty(X, Y)$$

by the preceding lemma:

$$\rho_2 = \frac{\log \frac{e^{-r-|\log a - \log(1-a)|} + e^{2r}}{e^{-r-|\log a - \log(1-a)|} + 1}}{2r}$$

Finally using the weighted property 4.:

$$e^{d_\infty(f(X),f(Y))} \leq \max\left\{ \frac{\sum_{i=1}^3 e^{\rho_i d_\infty(X,Y)} J_i e^{-d_\infty(f_m(X),f_i(X))}}{\sum_{i=1}^3 J_i e^{-d_\infty(f_m(X),f_i(X))}}, \frac{\sum_{i=1}^3 e^{\rho_i d_\infty(X,Y)} J_i e^{-d_\infty(f_m(Y),f_i(Y))}}{\sum_{i=1}^3 J_i e^{-d_\infty(f_m(Y),f_i(Y))}} \right\}$$
$$\leq \frac{e^{d_\infty(X,Y)}(J_1 + J_3) + e^{\rho_2 d_\infty(X,Y)} e^{-L}}{e^{-L} + J_1 + J_3}.$$

where

$$L \geq \max_{i,j=1,2,3} \sup_{X \in \overline{B}_I(r)} d_\infty(f_j(X), f_i(X))$$

and $J_i = 1$ if ν is supported on I_i, otherwise $J_i = 0$

so we see that

$$d_\infty(f(X), f(Y)) \leq \rho d_\infty(X, Y)$$

with

$$\rho = \frac{\log\left(\frac{e^{2r}(J_1+J_3)+e^{\rho_2 2r}e^{-L}}{e^{-L}+J_1+J_3}\right)}{2r}$$

and clearly $\rho < 1$. □

- The Contraction Theorem leads to the existence of the ALM mean for any Kubo-Ando mean in the full cone $\mathbb{P}$, by the paper of Lawson and Lim.

Induced operator means via contraction principle

Lemma

Let $\omega \in \Delta_k$ and $A_i \in \mathbb{P}$, $1 \leq i \leq k$ and M an operator mean with representing function $f \in \mathfrak{M}$, $f(x) \neq x$. Then the function

$$f_M(X) = \sum_{i=1}^{k} w_i M(X, A_i)$$

is a strict contraction with respect to $d_\infty(\cdot, \cdot)$ on every bounded $S \subseteq \mathbb{P}$ s.t. $A_i \in S$.

Proposition

Let $\omega \in \Delta_k$ and $A_i \in \mathbb{P}$, $1 \leq i \leq k$ and M an operator mean. Then the equation $X = \sum_{i=1}^{k} w_i M(X, A_i)$ has a unique solution in $\mathbb{P}$.

Definition (Induced Operator Mean)

Let M be an operator mean, $\mathbb{A} \in \mathbb{P}^k$ and $\omega \in \Delta_k$. We denote by $M(\omega; \mathbb{A})$ the unique solution of the equation

$$X = \sum_{i=1}^{k} w_i M(X, A_i)$$

and call the ω-weighted induced operator mean of M.

Proposition

Let $\mathbb{A} = (A_1, \ldots, A_k), \mathbb{B} = (B_1, \ldots, B_k) \in \mathbb{P}^k, \omega \in \Delta_k$ *and* $M(A, B), N(A, B)$ *operator means and* $M(\omega; \mathbb{A})$, $N(\omega; \mathbb{A})$ *the corresponding induced operator means. Then*

(1) $M(\omega; \mathbb{A}) = A$ *if* $A_i = A$ *for all* $1 \leq i \leq k$;

(2) $M(\omega_\sigma; \mathbb{A}_\sigma) = M(\omega; \mathbb{A})$ *for any permutation* σ;

(3) $M(\omega; \mathbb{A}) \leq M(\omega; \mathbb{B})$ *if* $A_i \leq B_i$ *for all* $i = 1, 2, \ldots, k$;

(4) *if* $M(A, B) \leq N(A, B)$ *for all* $A, B \in \mathbb{P}$ *then* $M(\omega; \mathbb{A}) \leq N(\omega; \mathbb{A})$;

(5) $M(\omega; X\mathbb{A}X^*) = XM(\omega; \mathbb{A})X^*$ *for any* $X \in \mathrm{GL}(E)$;

(6) $(1-u)M(\omega; \mathbb{A}) + uM(\omega; \mathbb{B}) \leq M(\omega; (1-u)\mathbb{A} + u\mathbb{B})$ *for any* $u \in [0, 1]$;

(7) $d_\infty(M(\omega; \mathbb{A}), M(\omega; \mathbb{B})) \leq \max_{1 \leq i \leq k}\{d_\infty(A_i, B_i)\}$;

(8) $M(\omega^{(n)}; \mathbb{A}^{(n)}) = M(\omega; \mathbb{A})$ *for any* $n \in \mathbb{N}$;

(9) $M(\omega; A_1, \ldots, A_{k-1}, X) = X$ if and only if $X = M(\hat{\omega}; A_1, \ldots, A_{k-1})$. In particular, $M(A_1, \ldots, A_k, X) = X$ if and only if $X = M(A_1, \ldots, A_k)$;

(10) $\Phi(M(\omega; \mathbb{A})) \leq M(\omega; \Phi(\mathbb{A}))$ for any positive unital linear map Φ, where $\Phi(\mathbb{A}) = (\Phi(A_1), \ldots, \Phi(A_k))$.

Proposition

Let $\omega \in \Delta_2$, $A, B \in \mathbb{P}$ and M an operator mean with representing function $f \in \mathfrak{M}$. Then

$$M(w_1, w_2; A, B) = A^{1/2} g\left(A^{-1/2} B A^{-1/2}\right) A^{1/2}$$

where

$$g^{-1}(x) = x f^{-1}\left(\frac{1 - w_1 f(x^{-1})}{w_2}\right).$$

Generalized Karcher equations and lambda extensions

Let $f \in \mathfrak{M}$. Then $f(x) = \exp_I(t_0 \log_I(x))$ where $\log_I \in \mathfrak{L}$.

Definition

For $\omega \in \Delta_k$ and $\mathbb{A} \in \mathbb{P}^k$ the generalized Karcher equation is:

$$\sum_{i=1}^{k} w_i \log_X(A_i) = 0$$

also $f_t = \exp_I(t \log_I(x)) \in \mathfrak{M}$ for $t = t_0^n$, $n = 1, 2, \ldots$

Proposition

If $\log_I$ has no ramification points, then the one-parameter family of induced operator means $M_t(\omega; \mathbb{A})$ induced by the $M_t(A, B)$ with representing function $f_t(x)$ is continuous for $t \in (0, 1]$ on any bounded set $S \subseteq \mathbb{P}$.

Theorem

There exists $X_0 \in \mathbb{P}$ such that

$$\lim_{n\to\infty} M_{t_0^n}(\omega; \mathbb{A}) = X_0$$

in the strong operator topology. Furthermore for $0 < t \leq s \leq t_0$ we have

$$X_0 \leq M_t(\omega; \mathbb{A}) \leq M_s(\omega; \mathbb{A}) \leq M_{t_0}(\omega; \mathbb{A}),$$

given $M_t(A, B)$, $M_s(A, B)$ are well defined operator means.

Definition

Let $\Lambda_M(\omega; \mathbb{A}) = \lim_{n\to\infty} M_{t_0^n}(\omega; \mathbb{A})$ and call it the ω-weighted lambda extension of $M(A, B)$.

Theorem

Let $\mathbb{A} = (A_1, \ldots, A_k), \mathbb{B} = (B_1, \ldots, B_k) \in \mathbb{P}^k, \omega \in \Delta_k$ *and* M, N *operator means and* $\Lambda_M(\omega; \mathbb{A})$, $\Lambda_N(\omega; \mathbb{A})$ *the corresponding lambda extensions. Then*

(1) $\Lambda_M(\omega; \mathbb{A}) = A$ *if* $A_i = A$ *for all* $1 \leq i \leq k$*;*

(2) $\Lambda_M(\omega_\sigma; \mathbb{A}_\sigma) = \Lambda_M(\omega; \mathbb{A})$ *for any permutation* σ;

(3) $\Lambda_M(\omega; \mathbb{A}) \leq \Lambda_M(\omega; \mathbb{B})$ *if* $A_i \leq B_i$ *for all* $i = 1, 2, \ldots, k$;

(4) *if* $M(A, B) \leq N(A, B)$ *for all* $A, B \in \mathbb{P}$ *then* $\Lambda_M(\omega; \mathbb{A}) \leq \Lambda_N(\omega; \mathbb{A})$*;*

(5) $\Lambda_M(\omega; X\mathbb{A}X^*) = X\Lambda_M(\omega; \mathbb{A})X^*$ *for any* $X \in \mathrm{GL}(E)$;

(6) $(1-u)\Lambda_M(\omega; \mathbb{A}) + u\Lambda_M(\omega; \mathbb{B}) \leq \Lambda_M(\omega; (1-u)\mathbb{A} + u\mathbb{B})$ *for any* $u \in [0, 1]$;

(7) $d_\infty(\Lambda_M(\omega; \mathbb{A}), \Lambda_M(\omega; \mathbb{B})) \leq \max_{1 \leq i \leq k}\{d_\infty(A_i, B_i)\}$;

(8) $\Lambda_M(\omega^{(n)}; \mathbb{A}^{(n)}) = \Lambda_M(\omega; \mathbb{A})$ *for any* $n \in \mathbb{N}$;

(9) $\Phi(\Lambda_M(\omega;\mathbb{A})) \leq \Lambda_M(\omega;\Phi(\mathbb{A}))$ for any positive unital linear map Φ, where $\Phi(\mathbb{A}) = (\Phi(A_1),\dots,\Phi(A_k))$;

(10) $\left(\sum_{i=1}^k w_i A_i^{-1}\right)^{-1} \leq \Lambda_M(\omega;\mathbb{A}) \leq \sum_{i=1}^k w_i A_i$.

(11) $\Lambda_M(\omega; A_1,\dots,A_{k-1},X) = X$ if and only if $X = \Lambda_M(\hat{\omega}; A_1,\dots,A_{k-1})$. In particular, $\Lambda_M(A_1,\dots,A_k,X) = X$ if and only if $X = \Lambda_M(A_1,\dots,A_k)$;

(12) $\Lambda_M(\omega;\mathbb{A})$ is the unique solution of the operator equation

$$0 = \sum_{i=1}^k w_i \log_X(A_i)$$

where $\log_X(A) = X^{1/2}\log_I\left(X^{-1/2}AX^{-1/2}\right)X^{1/2}$.

Further results about lambda extensions and induced means

Proposition

Let $\omega \in \Delta_2$, $A, B \in \mathbb{P}$ and M an operator mean with representing function $f(x) = \exp_I(t \log_I(x))$. Then

$$\Lambda_M(w_1, w_2; A, B) = A^{1/2} g\left(A^{-1/2} B A^{-1/2}\right) A^{1/2}$$

where

$$g^{-1}(x) = x \exp_I\left(-\frac{w_1}{w_2} \log_I(x^{-1})\right).$$

Elements in $\mathfrak{P}(t)$ generate elements of $\mathfrak{L}$:

Proposition

Let $f \in \mathfrak{P}(t)$. Then the function $\log_I(x) = \frac{f(x)-1}{f'(1)}$ is in $\mathfrak{L}$.

In this case

$$X = \sum_{i=1}^{k} w_i M(X, A_i) \Rightarrow 0 = \sum_{i=1}^{k} w_i [M(X, A_i) - X]$$

$$\begin{aligned} 0 &= \sum_{i=1}^{k} w_i \frac{M(X, A_i) - X}{f'(1)} \\ 0 &= \sum_{i=1}^{k} w_i X^{1/2} \log_I (X^{-1/2} A_i X^{-1/2}) X^{1/2} \\ 0 &= \sum_{i=1}^{k} w_i \log_X (A_i) \end{aligned}$$

Theorem

All induced operator means $M(\omega;\mathbb{A})$ are unique solutions of generalized Karcher equations.

The converse:

Proposition

Let $\log_I \in \mathfrak{L}$ and suppose that $\lim_{x\to 0+}\log_I(x) > -\infty$. Then the function

$$f_t(x) = t\log_I(x) + 1$$

is in $\mathfrak{P}(t)$ for $0 < t \leq \frac{1}{|\lim_{x\to 0+}\log_I(x)|}$ and the induced operator mean $M_t(\omega,\mathbb{A})$ is the unique solution of the generalized Karcher equation

$$\sum_{i=1}^{k} w_i \log_X(A_i) = 0$$

in $\mathbb{P}$ where $\log_X(A) = X^{1/2}\log_I\left(X^{-1/2}AX^{-1/2}\right)X^{1/2}$.

Proposition

Let M be an operator mean with representing function $f(x) = \exp_I(t \log_I(x))$. Suppose that the corresponding $\Lambda_M(w_1, w_2; A, B) = N(w_1, w_2; A, B)$, an induced mean for all $\omega \in \Delta_2$, $A, B \in \mathbb{P}$ where N is an operator mean with representing function $g(x)$. Then

$$\log_I(x) = \frac{g(x) - 1}{g'(1)}.$$

Not all $\log_I \in \mathfrak{L}$ can be given in the above form: a counterexample is $\log_I(x) = \log(x)$.

Some plausible questions:

- How big is the set of lambda extensions $\Lambda_M(w_1, w_2; A, B)$ in $\mathfrak{M}$?
- Is any element of $\mathfrak{M}$ a lambda extension of some $M \in \mathfrak{M}$?

Proposition

Let M be given with representing function $f(x) = \exp_I(t \log_I(x)) \in \mathfrak{M}$. Then the representing function $g(x)$ of the corresponding lambda extension $\Lambda_M(w_1, w_2; A, B)$ for all $\omega \in \Delta_2$, $A, B \in \mathbb{P}$ is in $\mathfrak{P}(w_2)$.

A (not so) well known previous result: $g(x) \in \mathfrak{M}$ if and only if the function $g^*(x) = \frac{x}{g(x)} \in \mathfrak{M}$

Theorem

Let $g(x) \in \mathfrak{M}$ be the representing function of the mean $M(A, B)$ such that $g'(1) \neq 0, 1/2, 1$. Let $g^*(x) = \frac{x}{g(x)}$ denote the conjugate pair. Define the function $h(x)$ s.t.

1. *if $g'(1) < 1/2$ then $h(x) := \frac{x}{g^{*-1}(x)}$*
2. *if $g'(1) > 1/2$ then $h(x) := xg^{-1}(x^{-1})$.*

Then M is a lambda extension if and only if there exists a positive integer n, such that the function $h^{\circ(2n)}(x)$ is in $\mathfrak{M}$. Moreover in this case the function $\log_I \in \mathfrak{L}$ is unique in the corresponding generalized Karcher equation.

I was unable to derive similar characterizations in the case when $g'(1) = 1/2$, however:

Proposition

Let $g(x) \in \mathfrak{M}$ with corresponding mean $M(A, B)$ such that $g'(1) = 1/2$. Define the function as $h(x) := xg^{-1}(x^{-1})$. Then M is a lambda extension if and only if $g(x) = xg(1/x)$ and there exists a holomorphic function $k(z)$ with $k(1) = 0$, $k'(1) = 1/2$ s.t.

$$\log_I(z) = -k(h(z)) + k(z)$$

where $\log_I \in \mathfrak{L}$ and there exists a $t \in (0, 1]$ such that the function $f_t(x) = \exp_I(t \log_I(x))$ is in $\mathfrak{P}(t)$ (where $\exp_I$ denotes the inverse of $\log_I$ as usual).

Open questions and problems

Problem

Is $\sum_{i=1}^{k} w_i \log_X(A_i)$ *the (Finslerian) gradient of some convex functional? Equivalently is* $\log_X(A)$ *itself a gradient?*

Problem

Can we determine a better characterization for lambda extensions in 2-variables? Can we do the same thing for induced operator means in 2-variables?

Problem

Is the (in fact symmetric) logarithmic mean with representing function $f(x) = \frac{x-1}{\log x}$ *a lambda extension, or induced operator mean?*

Problem

Give an example of a mean which is neither a lambda extension, nor an induced operator mean.

Problem (A possible wild guess)

Are all functions $M(A_1, \ldots, A_k) : \mathbb{P}^k \mapsto \mathbb{P}$ *with*

1. $M(A, \ldots, A) = A$,
2. $CM(A_1, \ldots, A_k)C^* = M(CA_1C^*, \ldots, CA_kC^*)$ *for all* $C \in \mathrm{GL}(E)$,
3. $M(A_1, \ldots, A_k)$ *is operator monotone,*
4. $M(a_1, \ldots, a_k)$ *is real for all* $0 < a_i \in \mathbb{R}$,

representable in the form

$$M(A_1, \ldots, A_k) = \int_{\Delta_k} \left(\sum_{i=1}^{k} w_i A_i^{-1} \right)^{-1} d\nu(w_1, \ldots, w_k)$$

where ν *is a probability measure over the simplex* Δ_k*?*

M. Pálfia, *Semigroups of operator means and generalized Karcher equations*, preprint (2013), arXiv:1208.5603.

Thank you for your kind attention!

Special thanks go to the organizers, who also invited me four years ago to Trieste, which has proven to be a crucial turning point for me.