J. Math. Anal. Appl. 404 (2013) 115–128

Contents lists available at SciVerse ScienceDirect

Journal of Mathematical Analysis and Applications

journal homepage: www.elsevier.com/locate/jmaa

Convex geometric means

Yongdo Lim

Department of Mathematics, Sungkyunkwan University, Suwon 440-746, Republic of Korea

ARTICLE INFO

Article history:
Received 4 September 2012
Available online 13 March 2013
Submitted by Gustavo Corach

Keywords:
Positive definite matrix
Geometric mean
Geometric mean majorization
Convex function
Karcher mean

ABSTRACT

A class of multivariable weighted geometric means of positive definite matrices admitting Jensen-type inequalities for geodesically convex functions is considered. It is shown that there are infinitely many such geometric means including the weighted inductive, Bini–Meini–Poloni and Karcher means and each of these means provides a geometric mean majorization on the space of positive definite matrices. Some connections between our geometric mean majorizations and classical results of the standard majorization of real numbers are discussed. In particular, we establish the Hardy–Littlewood–Pólya majorization theorem and also Rado's theorem and Schur's convexity theorem for the weighted Karcher mean.

© 2013 Elsevier Inc. All rights reserved.

1. Introduction

The Hardy–Littlewood–Pólya–Rado majorization theorem (cf. [3]) says in particular that for an $n \times n$ doubly stochastic matrix $W = (w_{ij})$ and $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$, $y = Wx$ is a convex combination of the $n!$ vectors $x_\sigma = (x_{\sigma(1)}, \ldots, x_{\sigma(n)})$ where σ varies over the permutation group $S^n = \{\sigma_1, \ldots, \sigma_{n!}\}$ of n-letters (Rado's theorem), that is,

$$Wx = \sum_{k=1}^{n!} \mu_k x_{\sigma_k} \tag{1.1}$$

for some probability vector $\mu = (\mu_1, \ldots, \mu_{n!}) \in \mathbb{R}^{n!}$, and equivalently (Hardy–Littlewood–Pólya theorem), for every continuous convex function f defined on an interval I containing x_i and y_i, $i = 1, \ldots, n$,

$$\sum_{i=1}^{n} f(y_i) \le \sum_{i=1}^{n} f(x_i). \tag{1.2}$$

Furthermore for every continuous convex function $f : I^n \to \mathbb{R}$ invariant under the permutation of coordinates (Schur's convexity),

$$f(y_1, \ldots, y_n) \le f(x_1, \ldots, x_n). \tag{1.3}$$

Embedding $\mathbb{R}^n$ into the space of $n \times n$ diagonal matrices and applying the exponential function, we may restate these beautiful results equivalently for the $n \times n$ positive diagonal matrices, where we replace the arithmetic mean by the geometric mean and a convex function by a geodesically convex function, $f(a^{1-t}b^t) \le (1-t)f(a) + tf(b)$, $t \in [0, 1]$. For instance, Rado's theorem is equivalent to the statement that for a positive diagonal matrix $\mathrm{diag}(a_1, \ldots, a_n)$, the diagonal matrix whose ii-th entry is the $\omega^i := (w_{i1}, \ldots, w_{in})$-weighted geometric mean of positive reals $a_1, \ldots, a_n$ is the diagonal matrix whose ii-th entry is the μ-weighted geometric mean of the $n!$ positive real numbers $a_{\sigma_1(i)}, \ldots, a_{\sigma_{n!}(i)}$.

E-mail addresses: ylim@skku.edu, ylim@knu.ac.kr.

0022-247X/$ – see front matter © 2013 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.jmaa.2013.03.006

The main purpose of this paper is to extend these results on diagonal matrices to the non-commutative setting of positive definite matrices by taking a multivariable weighted geometric mean of positive definite matrices, a multivariable extension of the weighted geometric mean $A\#_tB = A^{1/2}(A^{-1/2}BA^{-1/2})^tA^{1/2}$ of two positive definite matrices A and B, and a geodesically convex function which satisfies $f(A\#_tB) \le (1-t)f(A) + tf(B)$. There is no formal definition of a weighted geometric mean of finite number of positive definite matrices and defining multivariable geometric means is a non-trivial task that has been a recent topic of interest in core linear algebra. However, the open convex cone $\mathbb{P}_m$ of $m \times m$ positive definite matrices has various Finsler structures where the curve $t \mapsto A\#_tB$ on $[0, 1]$ acts as a minimal geodesic between A and B, which provides various types of geodesically convex sets and functions.

There are infinitely many multivariable weighted geometric means of positive definite matrices that are independent of matrix size and number of variables that satisfy the weighted version of the ten Ando–Li–Mathias properties [1]. Specifically, a weighted geometric mean G consists of maps

$$G_n^m : \Delta_n \times \mathbb{P}_m^n \to \mathbb{P}_m$$

such that each G_n^m satisfies the weighted Ando–Li–Mathias properties, where Δ_n is the simplex of positive probability vectors convexly spanned by the unit coordinate vectors and $\mathbb{P}_m$ is the convex cone of $m\times m$ positive definite matrices. The weighted Karcher mean based on the Riemannian trace metric and the weighted Bini–Meini–Poloni mean defined by induction using a symmetrization procedure are standard examples of weighted geometric means. In fact, a currently interesting problem for the Karcher mean is to find properties that distinguish it from other geometric means [6,24,29,23,25].

Let G be a weighted geometric mean and let $W = (w_{ij})_{n\times n}$ be a doubly stochastic matrix. A variant of Rado's theorem for positive definite matrices would be stated in the following from: for an n-tuple $\mathbb{A} = (A_1, \dots, A_n)$ of $m \times m$ positive definite matrices, the block diagonal matrix whose ii-entry is $G_n^m(\omega^i; \mathbb{A})$, $\omega^i = (w_{i1}, \dots, w_{in})$ can be realized as a $G_{n!}^{mn}$-weighted mean of the $n!$-block diagonal matrices $\mathbb{A}^{(i)} = \mathrm{diag}(A_{\sigma_i(1)}, \dots, A_{\sigma_i(n)})$, that is, there is a probability vector $\mu \in \mathbb{R}^{n!}$ satisfying

$$\begin{pmatrix} G_n^m(\omega^1; \mathbb{A}) & & \\ & \ddots & \\ & & G_n^m(\omega^n; \mathbb{A}) \end{pmatrix} = G_{n!}^{mn}(\mu; \mathbb{A}^{(1)}, \dots, \mathbb{A}^{(n!)}).$$

Rado's problem in the preceding form gives an important notion of geometric mean majorization for positive definite matrices. Let G be a weighted matrix geometric mean. For $\mathbb{A} = (A_1, \dots, A_n), \mathbb{B} = (B_1, \dots, B_n) \in \mathbb{P}_m^n$, we define $\mathbb{A} \prec^G \mathbb{B}$ if there exists a doubly stochastic matrix $W = (w_{ij})_{n\times n}$ such that for each $i = 1, \dots, n$,

$$A_i = G(w_{i1}, \dots, w_{in}; \mathbb{B}).$$

We construct infinitely many weighted geometric means such that if $\mathbb{A} \prec^G \mathbb{B}$, then $\{A_1, \dots, A_n\} \in [B_1, \dots, B_n]$, the geodesic convex hull of $\{B_1, \dots, B_n\}$ in the Riemannian manifold of positive definite matrices and $\sum_{i=1}^n f(A_i^{\pm 1}) \le \sum_{i=1}^n f(B_i^{\pm 1})$ for any continuous geodesically convex function f (Jensen inequality). We further establish Rado's theorem and Schur's convexity theorem for the Karcher mean, among other weighted geometric means.

2. Invariant metrics

Let $\mathbb{H} = \mathbb{H}_m$ be the space of Hermitian matrices of a fixed size m, and $\mathbb{P} = \mathbb{P}_m$ the corresponding convex cone of positive definite Hermitian matrices. For $X, Y \in \mathbb{H}$, we write $X \le Y$ if $Y - X$ is positive semidefinite, and $X < Y$ if $Y - X$ is positive definite. For $A \in \mathbb{H}$, $\lambda_j(A)$'s are denoted by the eigenvalues of A in non-increasing order: $\lambda_1(A) \ge \lambda_2(A) \ge \cdots \ge \lambda_m(A)$.

For $A, B \in \mathbb{P}$ and $t \in \mathbb{R}$, the t-weighted geometric mean of A and B is defined by

$$A\#_tB = A^{1/2}(A^{-1/2}BA^{-1/2})^tA^{1/2}.$$

The following properties for the weighted geometric mean are well-known [17].

Lemma 2.1. *Let $A, B, C, D \in \mathbb{P}$ and let $t \in \mathbb{R}$. Then*

(i) $A\#_tB = A^{1-t}B^t$ *for* $AB = BA$;
(ii) $(aA)\#_t(bB) = a^{1-t}b^t(A\#_tB)$ *for* $a, b > 0$;
(iii) (Löwner–Heinz inequality) $A\#_tB \le C\#_tD$ *for* $A \le C$, $B \le D$ *and* $t \in [0, 1]$;
(iv) $M(A\#_tB)M^* = (MAM^*)\#_t(MBM^*)$ *for non-singular* M;
(v) $A\#_tB = B\#_{1-t}A$, $(A\#_tB)^{-1} = A^{-1}\#_tB^{-1}$;
(vi) $(\lambda A + (1-\lambda)B)\#_t(\lambda C + (1-\lambda)D) \ge \lambda(A\#_tC) + (1-\lambda)(B\#_tD)$ *for* $\lambda, t \in [0, 1]$;
(vii) $\det(A\#_tB) = \det(A)^{1-t}\det(B)^t$;
(viii) $((1-t)A^{-1} + tB^{-1})^{-1} \le A\#_tB \le (1-t)A + tB$ *for* $t \in [0, 1]$;
(ix) $(A\#_tB)\#_s(A\#_uB) = A\#_{(1-s)t+su}B$ *for any* $s, t, u \in \mathbb{R}$.

A norm Φ on $\mathbb{R}^m$ is called a *symmetric gauge function* if it is invariant under permutations and sign changes of coordinates. Every symmetric gauge function Φ induces a unitarily invariant norm on $\mathbb{H}$

$$\|A\|_\Phi := \Phi(\lambda_1(A), \lambda_2(A), \ldots, \lambda_m(A)),$$

and conversely all unitarily invariant norms arise in this way by a theorem of von Neumann. Let Φ be a symmetric gauge norm. For $A \in \mathbb{P}$, we define a norm on the tangent space $T_A(\mathbb{P}) = \{A\} \times \mathbb{H} \equiv \mathbb{H}$ by $\|X\|_A = \|A^{-1/2}XA^{-1/2}\|_\Phi$. This yields a Finsler metric on $\mathbb{P}$. For a path $\gamma : [0, 1] \to \mathbb{P}$, we define its length as

$$L_\Phi(\gamma) = \int_0^1 \|\gamma^{-1/2}(t)\gamma'(t)\gamma^{-1/2}(t)\|_\Phi dt \tag{2.4}$$

and for $A, B \in \mathbb{P}$, its distance

$$d_\Phi(A, B) = \inf\{L_\Phi(\gamma) : \gamma \text{ is a path from } A \text{ to } B\}. \tag{2.5}$$

Theorem 2.2 ([4]). *We have $d_\Phi(A, B) = \|\log(A^{-1/2}BA^{-1/2})\|_\Phi$ and d_Φ is a complete metric distance on $\mathbb{P}$ such that for $A, B \in \mathbb{P}$ and for invertible matrix M,*

(i) $d_\Phi(A, B) = d_\Phi(A^{-1}, B^{-1}) = d_\Phi(MAM^*, MBM^*)$;
(ii) $d_\Phi(A\#B, A) = d_\Phi(A\#B, B) = \frac{1}{2}d_\Phi(A, B)$, *where* $A\#B = A\#_{\frac{1}{2}}B$;
(iii) $d_\Phi(A\#_tB, A\#_sB) = |s - t|d_\Phi(A, B)$ *for all* $t, s \in [0, 1]$;
(iv) $d_\Phi(A\#_tB, C\#_tD) \le (1 - t)d_\Phi(A, C) + td_\Phi(B, D)$ *for all* $t \in [0, 1]$.

By the triangular inequality,

$$d_\Phi(A\#_tB, C\#_sD) \le (1 - t)d_\Phi(A, C) + td_\Phi(B, D) + |t - s|d_\Phi(C, D) \tag{2.6}$$

for all $s, t \in [0, 1]$. Indeed,

$$\begin{aligned} d_\Phi(A\#_tB, C\#_sD) &\le d_\Phi(A\#_tB, C\#_tD) + d_\Phi(C\#_tD, C\#_sD) \\ &\le (1 - t)d_\Phi(A, C) + td_\Phi(B, D) + |t - s|d_\Phi(C, D). \end{aligned}$$

Example 2.3 (*Schatten p-Norms*). For $1 < p < \infty$, let Φ_p be the l_p-norm, which is a symmetric gauge function. The corresponding unitarily invariant norm on $\mathbb{H}$ is known as the Schatten p-norm and is defined by

$$\|X\|_p = \left[\sum_{i=1}^m \lambda_i(|X|)^p\right]^{\frac{1}{p}},$$

where $|X| := (X^2)^{\frac{1}{2}}$. The corresponding Finsler distance metric is given by $d_p(A, B) = \|\log(A^{-1/2}BA^{-1/2})\|_p$ on $\mathbb{P}$. In [10] C. Conde has proved that

$$d_p^2(X, A\#B) \le \frac{1}{2}d_p^2(X, A) + \frac{1}{2}d_p^2(X, B) - \frac{p-1}{4}d_p^2(A, B) \quad (p \le 2)$$
$$d_p^p(X, A\#B) \le \frac{1}{2}d_p^p(X, A) + \frac{1}{2}d_p^p(X, B) - \frac{1}{2^p}d_p^p(A, B) \quad (p \ge 2)$$

for all $X > 0$. (The formulas actually extend to the setting of compact operators on a separable Hilbert space.)

Example 2.4 (*Riemannian Trace Metric*). The Frobenius norm $\|\cdot\|_2$ gives rise to the Riemannian structure

$$\langle X, Y\rangle_A = \mathrm{Tr}(A^{-1}XA^{-1}Y),$$

where $A \in \mathbb{P}$ and $X, Y \in T_A(\mathbb{P}) = \mathbb{H}$. In this case, the curve $t \mapsto A\#_tB$ is the unique minimal geodesic (up to parametrization) from A to B and $A\#B$ is a unique midpoint between A and B. One important property of the metric $\delta(A, B) = \|\log A^{-1/2}BA^{-1/2}\|_2$ is the semiparallelogram law

$$\delta^2(Z, X\#Y) \le \frac{1}{2}\delta^2(Z, X) + \frac{1}{2}\delta^2(Z, Y) - \frac{1}{4}\delta^2(X, Y)$$

and its general form for any $t \in [0, 1]$

$$\delta^2(Z, X\#_tY) \le (1 - t)\delta^2(Z, X) + t\delta^2(Z, Y) - t(1 - t)\delta^2(X, Y). \tag{2.7}$$

The metric space $(\mathbb{P}, \delta)$ is an important example of a Hadamard space, a complete metric space satisfying the semiparallelogram law.

Example 2.5 (*Thompson Metric*)**.** For $p = \infty$, the metric on $\mathbb{P}$ arising from the spectral norm $\|\cdot\|_\infty$ is given by $d_\infty(A, B) = \|\log(A^{-1/2}BA^{-1/2})\|_\infty$ and coincides with the Thompson metric

$$d_\infty(A, B) = \max\{\log M(B/A), \log M(A/B)\}, \tag{2.8}$$

where $M(B/A) = \inf\{\alpha > 0 : B \leq \alpha A\} = \lambda_1(A^{-1/2}BA^{-1/2}) = \lambda_1(A^{-1}B)$. We note that $d_\infty(A, B) = \max\{\log \lambda_1(A^{-1}B), \log \lambda_1(B^{-1}A)\}$.

3. Convex sets and functions

Definition 3.1. A subset $C \subset \mathbb{P}$ is called geodesically convex (occasionally, in context, convex) if $A\#_t B \in C$ for all $t \in [0, 1]$ whenever $A, B \in C$. A function $f : C \to \mathbb{R}$ on a geodesically convex set C is called geodesically convex if for any $A, B \in C$ and $t \in [0, 1]$,

$$f(A\#_t B) \leq (1 - t)f(A) + tf(B).$$

Remark 3.2. Alternatively $f : \mathbb{P} \to \mathbb{R}$ is geodesically convex if and only if for all $A, B \in \mathbb{P}$, the composition $f\gamma_{A,B} : [0, 1] \to \mathbb{R}$ is convex in the usual sense, where $\gamma_{A,B}(t) = A\#_t B$.

Proof. Since by Lemma 2.1 (ix)

$$\gamma_{A,B}((1 - t)s + tu) = A\#_{(1-t)s+tu}B = (A\#_s B)\#_t(A\#_u B) = \gamma_{A,B}(s)\#_t\gamma_{A,B}(u),$$

we have

$$f\gamma_{A,B}((1 - t)s + tu) = f(\gamma_{A,B}(s)\#_t\gamma_{A,B}(u)) \leq (1 - t)f\gamma_{A,B}(s) + tf\gamma_{A,B}(u)$$

if f is convex. Conversely letting $s = 0$ and $u = 1$ in the first equation yields

$$f(A\#_t B) = f\gamma_{A,B}(0)\#_t f\gamma_{A,B}(1) \leq (1 - t)f\gamma_{A,B}(0) + tf\gamma_{A,B}(0) = (1 - t)f(A) + tf(B)$$

if $f\gamma_{A,B}$ is convex in the usual sense. □

Remark 3.3. By the Hermite–Hadamard inequality of convex functions on real intervals, which says that for a convex function g on $[a, b]$,

$$g\left(\frac{a+b}{2}\right) \leq \frac{1}{b-a}\int_a^b g(x)dx \leq \frac{g(a) + g(b)}{2},$$

we have

$$f(A\#B) \leq \int_0^1 f(A\#_t B)dt \leq \frac{f(A) + f(B)}{2} \tag{3.9}$$

for any geodesically convex function f on $\mathbb{P}$. See [11].

For $C \subseteq \mathbb{P}$, let $M(C) = \{X\#Y : X, Y \in C\}$. Note $C \subseteq M(C)$ since $X\#X = X$. Inductively set $M^{n+1}(C) = M(M^n(C))$. Then it is easy to see that $M^\infty(C) := \bigcup_n M^n(C)$ is the smallest midpoint-convex set containing C. It turns out [7,15] that the closure of $M^\infty(C)$ is the smallest closed, convex set containing C. We call it the *closed convex hull* of C and denote by $[C]$. It is straightforward to see that $[[C_1] \cup [C_2]] = [C_1 \cup C_2]$. We denote the closed convex hull of a finite set $\{X_1, \ldots, X_n\}$ by $[X_1, \ldots, X_n]$.

Proposition 3.4 ([15])**.** *We have $\Delta_\Phi[C] = \Delta_\Phi C$ for any subset C of $\mathbb{P}$ and any symmetric gauge function Φ, where $\Delta_\Phi C$ denotes the d_Φ-diameter of the set C.*

We list some examples of geodesically convex sets and convex functions.

Proposition 3.5. (1) *Every d_Φ-ball is geodesically convex;*
(2) *Every Löwner order interval $[X, Y] := \{Z \in \mathbb{P} : 0 < X \leq Z \leq Y\}$ is geodesically convex;*
(3) *Every closed subset of $\mathbb{P}_m$ which is stable under the arithmetic $\frac{A+B}{2}$ and harmonic mean $2(A^{-1}+B^{-1})^{-1}$ is geodesically convex;*
(4) *For a non-singular M, the sets*

$$\{X > 0 : MXM^* = X\}, \qquad \{X > 0 : MXM^* = X^{-1}\}$$

are geodesically convex;
(5) *The map $d_\Phi^\alpha(\cdot, Z)$ is geodesically convex for any $\alpha \geq 1$;*
(6) *The trace and determinant functions are geodesically convex;*

(7) *The log-determinant function* $X \mapsto \log\det(X)$ *is geodesically convex;*
(8) *If* f *is operator monotone and convex and* $g : \mathbb{P} \to \mathbb{R}$ *is monotone and convex, then* $g \circ f$ *is geodesically convex; and*
(9) *If* $\Psi : \mathbb{M}_m \to \mathbb{M}_k$ *is a strictly positive linear map and* $g : \mathbb{P}_k \to \mathbb{R}$ *is monotone and convex, then* $g \circ \Psi$ *is geodesically convex.*

Proof. (1) and (5): From Theorem 2.2.

(2), (6) and (7): From Lemma 2.1.

(3) and (4): See [21,22].

(8) Let f be an operator monotone and convex function and let $g : \mathbb{P} \to \mathbb{R}$ be monotone and convex. Then $f(A\#_t B) \leq f((1-t)A + tB) \leq (1-t)f(A) + tf(B)$ for all $t \in [0, 1]$. Since g is monotone and convex,

$$(g \circ f)(A\#_t B) \leq g((1-t)f(A) + tf(B)) \leq (1-t)(g \circ f)(A) + t(g \circ f)(B),$$

which implies the convexity of $g \circ f$.

(9) Follows from $\Psi(A\#_t B) \leq \Psi(A)\#_t\Psi(B)$ (cf. Theorem 4.1.5 of [5]). □

See Example 2.2.1 of [5] for a family of positive linear maps and also [22] for a one-parameter family of geodesically convex sets which covers the space $\mathbb{P}$. We note that any map satisfying $f(A\#_t B) \leq f(A)\#_t f(B) = f(A)^{1-t}f(B)^t$ for all $t \in [0, 1]$ is geodesically convex.

Remark 3.6. We have by Example 2.3 that $d_p^2(\cdot, Z)$ is uniformly convex for $1 < p \leq 2$ and $d_p^p(\cdot, Z)$ for $p > 2$ is uniformly convex. Here a map $f : \mathbb{P} \to \mathbb{R}$ is called uniformly convex with respect to on a complete metric d on $\mathbb{P}$ if there is a strictly increasing function $\phi : [0, \infty) \to [0, \infty)$ such that

$$f(A\#B) \leq \frac{1}{2}(f(A) + f(B)) - \phi(d(A, B))$$

for all $A, B \in \mathbb{P}$. A lower semicontinuous uniformly convex function $f : \mathbb{P} \to \mathbb{R}$ that is bounded below has a unique minimizer $\arg\min_{X\in\mathbb{P}} f(X)$ (see [28, Proposition 1.7]). We further note that $(\mathbb{P}, \delta)$ is a p-uniformly convex Cartan Hadamard manifold for any $p > 1$ (see [12]).

4. Geometric means

Once one realizes that the matrix geometric mean $G_2(A, B) = A\#B$ is a metric midpoint of A and B for the metric d_Φ, it is natural to use an averaging technique over this metric to extend this mean to a larger number of variables. M. Moakher [26] and Bhatia and Holbrook [7] suggested extending the geometric mean to n-points by taking the mean to be the unique minimizer of the sum of the squares of the distances:

$$\Lambda_n(A_1, \ldots, A_n) = \underset{X\in\mathbb{P}}{\arg\min} \sum_{i=1}^{n} \delta^2(X, A_i).$$

Another approach, independent of metric notions, was suggested by Ando, Li, and Mathias [1] via a "symmetrization procedure" and induction. The Ando–Li–Mathias paper was also important for listing, and deriving for their mean, ten desirable properties for extended geometric means that one might anticipate from properties of the two-variable geometric mean. The Ando–Li–Mathias mean proved to be computationally cumbersome, and Bini, Meini, and Poloni [9] suggested an alternative with more rapid convergence properties, which also satisfied the ten axioms. One notes in particular that while the axioms characterize the two-variable case, this is no longer true in the n-variable case, $n > 2$.

The ten properties originating from [1] may be generalized to the setting of weighted geometric means of n-positive definite matrices, where the weights $\omega = (w_1, \ldots, w_n)$ vary over Δ_n, the simplex of positive probability vectors convexly spanned by the unit coordinate vectors. We define a *symmetric weighted geometric mean* of n positive definite matrices to be a map $G : \Delta_n \times \mathbb{P}^n \to \mathbb{P}$ that satisfies the following properties: For $\mathbb{A} = (A_1, \ldots, A_n), \mathbb{B} = (B_1, \ldots, B_n) \in \mathbb{P}^n, \sigma \in S^n$ a permutation on n-letters, $\mathbf{a} = (a_1, \ldots, a_n) \in \mathbb{R}^n_{++}$ ($\mathbb{R}_{++} = (0, \infty)$), these are

(P1) (Consistency with scalars) $G(\omega; \mathbb{A}) = A_1^{w_1} \cdots A_n^{w_n}$ if the A_i's commute;
(P2) (Joint homogeneity) $G(\omega; a_1A_1, \ldots, a_nA_n) = a_1^{w_1} \cdots a_n^{w_n} G(\omega; \mathbb{A})$;
(P3) (Permutation invariance) $G(\omega_\sigma; \mathbb{A}_\sigma) = G(\omega; \mathbb{A})$, where $\omega_\sigma = (w_{\sigma(1)}, \ldots, w_{\sigma(n)})$ and $\mathbb{A}_\sigma = (A_{\sigma(1)}, \ldots, A_{\sigma(n)})$;
(P4) (Monotonicity) If $B_i \leq A_i$ for all $1 \leq i \leq n$, then $G(\omega; \mathbb{B}) \leq G(\omega; \mathbb{A})$;
(P5) (Continuity) The map $G(\omega; \cdot)$ is continuous;
(P6) (Congruence invariance) $G(\omega; M^*\mathbb{A}M) = M^*G(\omega; \mathbb{A})M$ for any invertible matrix M, where $M(A_1, \ldots, A_n)M^* = (MA_1M^*, \ldots, MA_nM^*)$;
(P7) (Joint concavity) $G(\omega; \lambda\mathbb{A} + (1-\lambda)\mathbb{B}) \geq \lambda G(\omega; \mathbb{A}) + (1-\lambda)G(\omega; \mathbb{B})$ for $0 \leq \lambda \leq 1$;
(P8) (Self-duality) $G(\omega; \mathbb{A}^{-1})^{-1} = G(\omega; \mathbb{A})$, where $\mathbb{A}^{-1} = (A_1^{-1}, \ldots, A_n^{-1})$;
(P9) (Determinantal identity) $\mathrm{Det}G(\omega; A_1, \ldots, A_n) = \prod_{i=1}^n (\mathrm{Det}A_i)^{w_i}$; and
(P10) (AGH weighted mean inequalities) $(\sum_{i=1}^n w_iA_i^{-1})^{-1} \leq G(\omega; A_1, \ldots, A_n) \leq \sum_{i=1}^n w_iA_i$.

We call a map G satisfying (P1)–(P10) except for (P3) a weighted geometric mean. We note that the two-variable weighted geometric mean $G_2(w_1, w_2; A, B) = A\#_{w_2}B$ satisfies (P1)–(P10) and is uniquely determined by (P1) and (P6).

The following result is valid for any map G satisfying (P2) and (P5); see [18].

Proposition 4.1. *Every weighted geometric mean is "contractive" for the Thompson metric d_∞;*

$$d_\infty(G(\omega; \mathbb{A}), G(\omega; \mathbb{B})) \leq \sum_{i=1}^{n} w_i d_\infty(A_i, B_i).$$

Lemma 4.2. *If G and H are weighted geometric means, then $G\#_tH$ is a weighted geometric mean for all $t \in [0, 1]$, where*

$$(G\#_tH)(\omega; \mathbb{A}) = G(\omega; \mathbb{A})\#_tH(\omega; \mathbb{A}).$$

Proof. One can directly show from Lemma 2.1 that $G\#_tH$ satisfies (Pk) whenever G and H satisfy (Pk); that is, if G and H are weighted geometric means then $G\#_tH$ is also. $\square$

Definition 4.3 (*Weighted Inductive Mean*)**.** For $\omega = (w_1, \ldots, w_n) \in \Delta_n$ and $\mathbb{A} = (A_1, \ldots, A_n) \in \mathbb{P}^n$, define

$$S_1(1; A_1) = A_1,$$
$$S_n(\omega; \mathbb{A}) = S_{n-1}(\hat{\omega}; \hat{\mathbb{A}})\#_{w_n}A_n, \quad (n \geq 2),$$

where $\hat{\omega} = \frac{1}{1-w_n}(w_1, w_2, \ldots, w_{n-1}) \in \Delta_{n-1}$ and $\hat{\mathbb{A}} = (A_1, \ldots, A_{n-1})$. We call $S_n(\omega; \mathbb{A})$ the ω-weighted inductive mean (convex combination) of $A_1, \ldots, A_n$.

Proposition 4.4. *The weighted inductive mean is indeed a (non-symmetric) weighted geometric mean.*

Proof. By induction and Lemma 2.1. $\square$

The ω-weighted *Karcher mean*, also called the *least squares mean*, of n positive definite matrices $A_1, \ldots, A_n$ and $\omega = (w_1, \ldots, w_n) \in \Delta_n$ is defined as the unique minimizer of the sum of squares of the Riemannian trace metric distances to each of the A_i, i.e.,

$$\Lambda(\omega; A_1, \ldots, A_n) = \underset{X \in \mathbb{P}}{\arg\min} \sum_{i=1}^{n} w_i \delta^2(X, A_i). \tag{4.10}$$

From [7,26] the Karcher mean coincides with the unique positive definite solution of the *Karcher equation*

$$\sum_{i=1}^{n} w_i \log(X^{1/2}A_i^{-1}X^{1/2}) = 0. \tag{4.11}$$

Theorem 4.5 (*[17]*)**.** *The Karcher mean is a symmetric weighted geometric mean.*

We note that the weighted Karcher mean exists on any Hadamard space. (cf. [28]).

Remark 4.6 (*Fixed Point Means*)**.** Let $G : \Delta_{n+k} \times \mathbb{P}^{n+k} \to \mathbb{P}$ be a weighted geometric $(n+k)$-mean. Let $\omega = (w_1, \ldots, w_n) \in \Delta_n$ and $\mathbb{A} = (A_1, \ldots, A_n) \in \mathbb{P}^n$. Then for each $t > 1$, the equation

$$G(\omega_{t,k}; A_1, \ldots, A_n, \underbrace{X, \ldots, X}_{k}) = X$$

has a unique positive definite solution, where $\omega_{t,k} = (\frac{w_1}{t}, \ldots, \frac{w_n}{t}, \frac{t-1}{tk}, \ldots, \frac{t-1}{tk}) \in \Delta_{n+k}$, denoted by $G_{t,k}(\omega; \mathbb{A})$. Indeed, one can see by using Proposition 4.1 that the map

$$X \mapsto G(\omega_{t,k}; A_1, \ldots, A_n, \underbrace{X, \ldots, X}_{k})$$

is a strict contraction with the Lipschitz constant less than equal to $\frac{t-1}{t}$ with respect to the Thompson metric. One can show (cf. [20]) that $G_{t,k}$ is a weighted geometric mean and that the Karcher mean and the weighted inductive mean coincide with their respective fixed point means.

Remark 4.7 (*Geometric Power Means, [23]*)**.** Let G be a weighted geometric mean. For $t \in (0, 1]$, the following equation

$$X = G(\omega; X\#_tA_1, \ldots, X\#_tA_n)$$

has a unique positive definite solution. In fact, the map $X \mapsto G(\omega; X\#_tA_1, \ldots, X\#_tA_n)$ is a strict contraction on $\mathbb{P}$ with least contraction coefficient less than equal to $1 - t$ for the Thompson metric. Define by $G_t(\omega; A_1, \ldots, A_n)$ the unique solution. Then each G_t is a weighted geometric mean. Moreover,

$$\lim_{t \to 0^+} G_t(\omega; A_1, \ldots, A_n) = \Lambda(\omega; A_1, \ldots, A_n). \tag{4.12}$$

Remark 4.8 (*Weighted BMP Mean*)**.** In [19] Lee, Lim and Yamazaki have found via the induction argument and appropriate symmetrization procedures a continuous map $\Gamma : [0,1]^n \times \Delta_n \to \Delta_n$ satisfying

- $\Gamma(\mathbf{1}_n; \omega) = \omega$ for all $\omega \in \Delta_n$; and
- $\Gamma(\mathbf{t}; \frac{1}{n}\mathbf{1}_n) = \frac{1}{n}\mathbf{1}_n, \;\; \mathbf{1}_n := (1, 1, \ldots, 1) \in \mathbb{R}^n$

for all $\mathbf{t} \in [0,1]^n$, and a map $\mathfrak{G} : [0,1]^n \times \Delta_n \times \mathbb{P}^n \to \mathbb{P}$ satisfying

(Γ1) $\mathfrak{G}(\mathbf{t}; \omega; \mathbb{A}) = \prod_{i=1}^n A_i^{\Gamma(\mathbf{t};\omega)_i}$ for commuting A_i's;
(Γ2) (*Joint homogeneity*) $\mathfrak{G}(\mathbf{t}; \omega; a_1A_1, \ldots, a_nA_n) = \prod_{i=1}^n a_i^{\Gamma(\mathbf{t};\omega)_i}\mathfrak{G}(\mathbf{t}; \omega; \mathbb{A})$;
(Γ3) (*Permutation invariance*) $\mathfrak{G}(\mathbf{t}; \omega_\sigma; \mathbb{A}_\sigma) = \mathfrak{G}(\mathbf{t}; \omega; \mathbb{A})$;
(Γ4) (*Monotonicity*) If $B_i \le A_i$ for all $1 \le i \le n$, then $\mathfrak{G}(\mathbf{t}; \omega; \mathbb{B}) \le \mathfrak{G}(\mathbf{t}; \omega; \mathbb{A})$;
(Γ5) (*Continuity*) The map $\mathfrak{G}(\mathbf{t}; \omega; \cdot)$ is continuous;
(Γ6) (*Congruence invariance*) $\mathfrak{G}(\mathbf{t}; \omega; M^*\mathbb{A}M) = M^*\mathfrak{G}(\mathbf{t}; \omega; \mathbb{A})M$;
(Γ7) (*Joint concavity*) $\mathfrak{G}(\mathbf{t}; \omega; \cdot)$ is jointly concave;
(Γ8) (*Self-duality*) $\mathfrak{G}(\mathbf{t}; \omega; \mathbb{A}^{-1})^{-1} = G(\omega; \mathbb{A})$;
(Γ9) (*Determinantal identity*) $\det\mathfrak{G}(\mathbf{t}; \omega; \mathbb{A}) = \prod_{i=1}^n (\det A_i)^{\Gamma(\mathbf{t};\omega)_i}$; and
(Γ10) (*AGH mean inequalities*) $\left(\sum_{i=1}^n \Gamma(\mathbf{t}; \omega)_i A_i^{-1}\right)^{-1} \le \mathfrak{G}(\mathbf{t}; \omega; \mathbb{A}) \le \sum_{i=1}^n \Gamma(\mathbf{t}; \omega)_i A_i$.

For $\mathbf{t} \in [0,1]^n$, the map $\mathfrak{G}(\mathbf{t}; \cdot; \cdot) : \Delta_n \times \mathbb{P}^n \to \mathbb{P}$ satisfies the weighted geometric mean properties and each of these means interpolates between the weighted ALM ($\mathbf{t} = \mathbf{0}_n = (0, \ldots, 0)$) and BMP ($\mathbf{t} = \mathbf{1}_n$) means [1,9]. We note that $\mathfrak{G}(\mathbf{t}; \frac{1}{n}\mathbf{1}_n; \cdot)$ is a (un-weighted) geometric mean for all $\mathbf{t} \in [0,1]^n$ with

$$\mathfrak{G}\left(\mathbf{0}; \frac{1}{n}\mathbf{1}_n; \cdot\right) = \mathrm{Alm}_n, \qquad \mathfrak{G}\left(\mathbf{1}; \frac{1}{n}\mathbf{1}_n; \cdot\right) = \mathrm{Bmp}_n.$$

We note from $\Gamma(\mathbf{1}_n; \omega) = \omega$ that the weighted BMP mean is indeed a symmetric weighted geometric mean; satisfying all the properties (P1)–(P10) and that the weighted BMP mean is constructed by induction and the following symmetrization procedure:

(1) For $n = 2$, $\mathrm{Bmp}_2(w_1, w_2; A_1, A_2) = A_1\#_{w_2}A_2$.
(2) Assume that $\mathrm{Bmp}_{n-1}(\cdot; \cdot) : \Delta_{n-1} \times \mathbb{P}^{n-1} \to \mathbb{P}$ is defined. Let $\{A_i^{(r)}\}_{r=0}^\infty$ be the sequence defined by; $A_i^{(0)} = A_i$ and

$$A_i^{(r+1)} = A_i^{(r)}\#_{1-w_i}\mathrm{Bmp}_{n-1}\left(\left(\frac{w_j}{1-w_i}\right)_{j\neq i}; (A_j^{(r)})_{j\neq i}\right), \quad 1 \le i \le n, \tag{4.13}$$

where $(a_j)_{j\neq i} := (a_1, \ldots, a_{i-1}, a_{i+1}, \ldots, a_n)$. Then $\lim_{r\to\infty} A_i^{(r)}$ exists and has the same value for every i; we denote the common limit by $\lim_{r\to\infty} A_i^{(r)} = \mathrm{Bmp}_n(\omega; A_1, \ldots, A_n)$.

See [15] for the weighted BMP mean in a general setting of metric spaces.

Definition 4.9. A map $G : \Delta_n \times \mathbb{P}^n \to \mathbb{P}$ is called *stable* on a subset $\Omega \subset \mathbb{P}$ if $G(\omega; A_1, \ldots, A_n) \in \Omega$ for all $\omega \in \Delta_n$ and $A_i \in \Omega, i = 1, \ldots, n$.

Lemma 4.10. *Let $\Omega \subset \mathbb{P}$ be a closed geodesically convex set and let $G : \Delta_n \times \mathbb{P}^n \to \mathbb{P}$ be a weighted geometric mean. If G is stable on Ω, then its geometric power mean and fixed point mean are stable on Ω.*

Proof. Let $\omega \in \Delta_n$ and let $\mathbb{A} = (A_1, \ldots, A_n) \in \Omega^n$. By Remark 4.7, the map $f(X) = G(\omega; X\#_tA_1, \ldots, X\#_tA_n)$ is a strict contraction for the Thompson metric whose unique fixed point is $G_t(\omega; A_1, \ldots, A_n)$. By the Ω-stability of f and $X\#_tA_i \in \Omega$ for all $i = 1, \ldots, n$, f maps Ω into itself. Pick $X_0 \in \Omega$. Since Ω is closed,

$$G_t(\omega; A_1, \ldots, A_n) = \lim_{k\to\infty} f^k(X_0) \in \Omega.$$

This shows that G_t is stable on Ω.

A similar argument holds for the fixed point means $G_{t,k}$. $\square$

Theorem 4.11 (*Stability*)**.** *The weighted inductive, BMP and Karcher means are stable on any closed geodesically convex set.*

Proof. It is easy to see that the weighted inductive and BMP means are stable on a geodesically convex set Ω from the closedness of Ω and their symmetrization procedures and induction.

For the case of the Karcher mean, we let G be either the inductive mean or BMP mean. Let $X_t = G_t(\omega; A_1, \ldots, A_n)$, where $A_i \in \Omega, i = 1, \ldots, n$ and $t \in (0,1]$. Applying the previous lemma and the first paragraph yields $X_t \in \Omega$ for all $t \in (0,1]$. It then follows from (4.12) and closedness of Ω that

$$\Lambda_n(\omega; A_1, \ldots, A_n) = \lim_{t\to 0^+} X_t \in \Omega.$$

That is, the weighted Karcher mean is stable on Ω. $\square$

Remark 4.12. The stability of the weighted Karcher mean, in particular the result that

$$\Lambda_n(\omega; A_1, \ldots, A_n) \in [A_1, \ldots, A_n] \tag{4.14}$$

for all $\omega \in \Delta_n$ and $(A_1, \ldots, A_n) \in \mathbb{P}^n$, which was shown by Bhatia and Holbrook [7], can be proved alternatively by Sturm's Strong Law of Large Number [28,16,8] or by Holbrook's no dice theorem [13] for un-weighted case $\omega = (1/n, \ldots, 1/n)$. The key step in the proof of Bhatia and Holbrook [7] is the so called Pythagorean inequality for Hadamard metric spaces (see [5, Theorem 6.2.7]); see also [14, Theorem 2.3.3] for a nice proof of this result.

It is typically non-trivial to show that the stability theorem, equivalently (4.14), holds for a given weighted geometric mean. Alternatively, this property provides a new class of weighted geometric means, those satisfying the stability theorem or (4.14).

Remark 4.13 (*Block Diagonal Matrices*)**.** Let $p_i \in \mathbb{N}$ with $\sum_{i=1}^k p_i = m$. For $P_i \in \mathbb{P}_{p_i}$, $i = 1, \ldots, k$, the block diagonal matrices $\mathrm{diag}(P_1, \ldots, P_k) \in \mathbb{P}_m$. Obviously the set $\mathbb{P}(p_1, \ldots, p_k)$ of such positive definite block diagonal matrices forms a closed geodesically convex subset of $\mathbb{P}_m$ and hence the Karcher mean is stable on it. Let $A_i = \mathrm{diag}(P_{i1}, \ldots, P_{ik}) \in \mathbb{P}(p_1, \ldots, p_k) \subset \mathbb{P}_m$, $i = 1, \ldots, n$. One can see by the Karcher equation that

$$\Lambda_n(\omega; A_1, \ldots, A_n) = \mathrm{diag}(\Lambda_n(\omega; P_{11}, \ldots, P_{n1}), \ldots, \Lambda_n(\omega; P_{1k}, \ldots, P_{nk})).$$

5. Convex geometric means

Definition 5.1. A weighted geometric mean $G : \Delta_n \times \mathbb{P}^n \to \mathbb{P}$ is called *convex* if

$$f(G(\omega; A_1, \ldots, A_n)) \le \sum_{i=1}^n w_i f(A_i)$$

for all $\omega = (w_1, \ldots, w_n) \in \Delta_n$, $(A_1, \ldots, A_n) \in \mathbb{P}^n$ and continuous geodesically convex functions $f : \mathbb{P} \to \mathbb{R}$. We denote by $\mathcal{C}_n$ the set of all weighted convex geometric n-means.

Proposition 5.2. *Let $G : \Delta_n \times \mathbb{P}^n \to \mathbb{P}$ be a convex weighted geometric mean. Then for any $\alpha \ge 1$, $X \in \mathbb{P}$, $\omega = (w_1, \ldots, w_n) \in \Delta_n$ and $\mathbb{A} = (A_1, \ldots, A_n) \in \mathbb{P}^n$,*

$$d_\Phi^\alpha(X, G(\omega; A_1, \ldots, A_n)) \le \sum_{i=1}^n w_i d_\Phi^\alpha(X, A_i) \tag{5.15}$$

for all symmetric gauge functions Φ. In particular for all $X \in \mathbb{P}$,

$$d_\Phi^2(X, G(\omega; A_1, \ldots, A_n)) \le \sum_{i=1}^n w_i d_\Phi^2(X, A_i). \tag{5.16}$$

Furthermore, G is stable on any d_Φ-balls.

Proof. The first part of proof follows by the convexity of the map $X \mapsto d_\Phi^\alpha(X, Z)$ and the observation that sums and nonnegative scalar multiples of convex functions are again convex. Let $A_1, \ldots, A_n \in \mathcal{B}_\Phi(A_0, r)$, the d_Φ-ball of radius r and centered at A_0. Then by (5.15),

$$d_\Phi(A_0, G(\omega; A_1, \ldots, A_n)) \le \sum_{i=1}^n w_i d_\Phi(A_0, A_i) \le r,$$

which shows that $G(\omega; A_1, \ldots, A_n)$ lies in the ball. □

Remark 5.3. By (5.15), one can derive that

$$d_\Phi(G(\omega; A_1, \ldots, A_n), G(\omega; B_1, \ldots, B_n)) \le \sum_{i=1}^n w_i \left[\sum_{j=1}^n w_j d_\Phi(A_i, B_j) \right].$$

This is weaker than the corresponding result for the Thompson metric (Proposition 4.1).

Theorem 5.4. *The set $\mathcal{C}_n$ is closed under the geometric mean $G \#_t H$, the fixed point mean and geometric power mean operations. Moreover the weighted inductive mean S_n, the weighted BMP mean Bmp_n and the Karcher mean Λ_n belong to $\mathcal{C}_n$.*

Proof. Let $G, H \in \mathcal{C}_n$. Then $G\#_t H$ is a weighted geometric mean by Lemma 4.2. Let f be a continuous geodesically convex function. Then

$$\begin{aligned} f((G\#_t H)(\omega; \mathbb{A})) &= f(G(\omega; \mathbb{A})\#_t H(\omega; \mathbb{A})) \\ &\leq (1-t)f(G(\omega; \mathbb{A})) + tf(H(\omega; \mathbb{A})) \\ &\leq \sum_{i=1}^{n} w_i f(A_i). \end{aligned}$$

Therefore $G\#_t H \in \mathcal{C}_n$.

Next, we will show that S_n and Bmp_n are convex. Suppose that S_{n-1} is convex. Then for a continuous geodesically convex function f,

$$\begin{aligned} f(S_n(\omega; \mathbb{A})) &= f(S_{n-1}(\hat{\omega}; \hat{\mathbb{A}})\#_{w_n} A_n) \\ &\leq (1-w_n)f(S_{n-1}(\hat{\omega}; \hat{\mathbb{A}})) + w_n f(A_n) \\ &\leq (1-w_n)\sum_{i=1}^{n-1} \frac{w_i}{1-w_n} f(A_i) + w_n f(A_n) \\ &= \sum_{i=1}^{n} f(A_i). \end{aligned}$$

By induction, S_n is convex. Similarly one can show that weighted BMP mean is convex from the fact that it is defined by induction and a symmetrization procedure.

Let $G \in \mathcal{C}_n$ and let $t \in (0, 1]$. Let $X = G_t(\omega; \mathbb{A})$. Then for a continuous geodesically convex function f,

$$\begin{aligned} f(X) &= f(G(\omega; X\#_t A_1, \ldots, X\#_t A_n)) \\ &\leq \sum_{i=1}^{n} w_i f(X\#_t A_i) \\ &\leq \sum_{i=1}^{n} w_i[(1-t)f(X) + tf(A_i)] \\ &= (1-t)f(X) + t\sum_{i=1}^{n} w_i f(A_i) \end{aligned}$$

which ensures that $f(X) \leq \sum_{i=1}^{n} w_i f(A_i)$. That is, the geometric power mean G_t is convex whenever G is.

Similarly each fixed point mean $G_{t,k}$ belongs to $\mathcal{C}_n$. □

Remark 5.5. The convexity of the Karcher mean holds on any Hadamard space (Sturm, [28]).

Definition 5.6. A weighted geometric mean $G : \Delta_n \times \mathbb{P}^n \to \mathbb{P}$ is called *Schur-convex* if $G \in \mathcal{C}_n$ and

$$G(\omega; A_1, \ldots, A_n) \in [A_1, \ldots, A_n] \tag{5.17}$$

for all $\omega \in \Delta_n$ and $(A_1, \ldots, A_n) \in \mathbb{P}^n$. We denote by $\mathcal{SC}_n$ the set of all Schur-convex weighted geometric n-means.

Theorem 5.7. *The set $\mathcal{SC}_n$ is closed under the geometric mean $G\#_t H$, the fixed point mean and geometric power mean operations. Moreover the weighted inductive mean S_n, the weighted BMP mean Bmp_n and the Karcher mean Λ_n belong to $\mathcal{SC}_n$.*

Proof. Let $G, H \in \mathcal{SC}_n$. For $t \in [0, 1]$,

$$(G\#_t H)(\omega; \mathbb{A}) = G(\omega; \mathbb{A})\#_t H(\omega; \mathbb{A}) \in [A_1, \ldots, A_n]$$

by the convexity of $[A_1, \ldots, A_n]$. That is, $G\#_t H \in \mathcal{SC}_n$. By Lemma 4.10, the geometric power mean G_t and the fixed point mean $G_{t,k}$ are stable on $[A_1, \ldots, A_n]$.

The remaining part of proof follows by Theorem 4.11. □

Corollary 5.8. *Let $G \in \mathcal{C}_n$ and $H \in \mathcal{SC}_n$. Then*

$$d_\Phi(H(\omega; \mathbb{A}), G(\omega; \mathbb{A})) \leq \Delta_\Phi(\mathbb{A}) := \max_{1\leq i,j\leq n} d_\Phi(A_i, A_j)$$

for all symmetric gauge functions Φ.

Proof. Let $G \in \mathcal{C}_n$ and let $\mathbb{A} = (A_1, \ldots, A_n) \in \mathbb{P}^n$. By Proposition 3.4, the d_Φ-diameter of $\{A_i\}_{i=1}^n$ is $\Delta_\Phi(\mathbb{A}) = \max_{1\leq i,j\leq n} d_\Phi(A_i, A_j)$. Applying (5.16) with $X = H(\omega; \mathbb{A})$ leads to

$$d_\Phi^2(X, G(\omega; \mathbb{A})) \leq \sum_{i=1}^n w_i d_\Phi^2(X, A_i).$$

By Theorem 5.7 and (5.17), $d_\Phi(X, A_i) \leq \Delta_\Phi(\mathbb{A})$ for all $i = 1, \ldots, n$. □

Remark 5.9. We note that the weighted inductive mean S_n, the Bmp_n and the Karcher mean Λ_n can be extended continuously into nonnegative weights $\omega \in \overline{\Delta_n}$ from the symmetrization procedures and the Karcher equation.

Remark 5.10 (*Un-Weighted Case*)**.** We consider the geometric mean, convex function and Schur-convex function for the un-weighted case $\omega = (1/n, \ldots, 1/n)$. Then the previous results hold and the un-weighted ALM mean is Schur-convex.

Problem 1. Is any weighted geometric mean (Schur) convex?

Problem 2. Extend (3.9) to (un-weighted) Schur-convex geometric means. We note that the following multivariable Hermite–Hadamard inequality holds true [2]:

$$f\left(\frac{p_1 + \cdots + p_n}{n}\right) \leq \frac{1}{\mathrm{Vol}(S)} \int_S f \leq \frac{f(p_1) + \cdots + f(p_n)}{n}$$

for any convex function $f : S \to \mathbb{R}$, where $S = [p_1, \ldots, p_n] \subset \mathbb{R}^n$. An appropriate Hermite–Hadamard inequality for a Schur-convex geometric mean would be the following:

$$f(G(\mathbb{A})) \leq \frac{1}{(n-1)!} \int_{\omega\in\Delta_n} f(G(\omega; \mathbb{A}))d\omega \leq \frac{f(A_1) + \cdots + f(A_n)}{n}$$

for any continuous geodesically convex function $f : \mathbb{P} \to \mathbb{R}$.

6. Geometric mean majorizations

Definition 6.1. Let $G \in \mathcal{C}_n$ and let $\mathbb{A} = (A_1, \ldots, A_n), \mathbb{B} = (B_1, \ldots, B_n) \in \mathbb{P}^n$. We say that $\mathbb{A}$ is G-majorized by $\mathbb{B}$ (abbreviated, $\mathbb{A} \prec^G \mathbb{B}$) if there exists a positive doubly stochastic matrix $W = (w_{ij})_{n\times n}$ such that for all $i = 1, \ldots, n$,

$$A_i = G(w_{i1}, \ldots, w_{in}; \mathbb{B}).$$

Remark 6.2. If $G \in \mathcal{C}_n$ can be defined on $\overline{\Delta_n}$ (e.g., S_n, Λ_n, and Bmp_n, see Remark 5.9), then the positivity condition of G-majorization can be excluded.

Remark 6.3 ($n = 2$)**.** Since the map $(t, A, B) \mapsto A\#_t B$ is the unique weighted geometric mean for $n = 2$, $(A_1, A_2) \prec (B_1, B_2)$ if and only if $A_1 = B_1\#_t B_2$ and $A_2 = B_1\#_{1-t} B_2 = B_2\#_t B_1$ for some $t \in [0, 1]$ if and only if A_1 and A_2 lie in the geodesic segment between B_1 and B_2 and $d_\Phi(A_1, B_1) = d_\Phi(A_2, B_2)$ for some (all) symmetric gauge function Φ.

Example 6.4. We have $(G(\mathbb{A}), \ldots, G(\mathbb{A})) \prec^G \mathbb{A}$ for all $\mathbb{A} \in \mathbb{P}^n$ and $G \in \mathcal{C}_n$. Use

$$W = \frac{1}{n}\begin{pmatrix} 1 & 1 & \cdots & 1 \\ 1 & 1 & \cdots & 1 \\ & & \ddots & \\ 1 & 1 & \cdots & 1 \end{pmatrix}.$$

Proposition 6.5. *Suppose that* $\mathbb{A} \prec^G \mathbb{B}$.

(1) $M\mathbb{A}M^* \prec^G M\mathbb{B}M^*$ *for any non-singular* M.
(2) $\mathbb{A}^{-1} \prec^G \mathbb{B}^{-1}$.
(3) $\log\det\mathbb{A} \prec \log\det\mathbb{B}$, *where* $\det(\mathbb{A}) = (\det A_1, \ldots, \det A_n)$.
(4) *If* G *is symmetric, then* $\mathbb{A}_\sigma \prec^G \mathbb{B}_\sigma$ *for any permutation* σ.
(5) *If* $G \in \{\Lambda_n, S_n, \mathrm{Bmp}_n\}$, *then for any strictly positive unital linear map* Φ,

$$\Phi(\mathbb{A}) \prec^\Lambda \Phi(\mathbb{B})$$

where $\Phi(\mathbb{A}) = (\Phi(A_1), \ldots, \Phi(A_n))$.

Proof. (1)–(3) follow by the congruence invariancy, self-duality and determinantal identity of G.

(4) Let σ be a permutation. Let $W_\sigma = (w_{\sigma(i)\sigma(j)})_{n\times n}$. Then W_σ is a positive doubly stochastic matrix and

$$A_{\sigma(i)} = G(w_{\sigma(i)1,\ldots,w_{\sigma(i)}n}; \mathbb{B}) = G(w_{\sigma(i)\sigma(1)}, \ldots, w_{\sigma(i)\sigma(n)}; \mathbb{B}_\sigma)$$

where the second equality follows from the permutation invariancy of G. This shows that $\mathbb{A}_\sigma \prec^G \mathbb{B}_\sigma$.

(5) Use $\Phi(\Lambda_n(\omega; \mathbb{A})) = \Lambda_n(\omega; \Phi(A_1), \ldots, \Phi(A_n))$. $\square$

The next theorem is a partial extension of Hardy–Littlewood–Pólya Theorem to convex geometric means of positive definite matrices.

Theorem 6.6. *If $\mathbb{A} \prec^G \mathbb{B}$, then $\sum_{i=1}^n f(A_i^{\pm 1}) \le \sum_{i=1}^n f(B_i^{\pm 1})$ for any continuous geodesically convex function f.*

Proof. This follows from that

$$\begin{aligned}\sum_{i=1}^n f(A_i) &= \sum_{i=1}^n f(G(w_{i1}, \ldots, w_{in}; B_1, \ldots, B_n)) \\ &\le \sum_{i=1}^n \sum_{j=1}^n w_{ij} f(B_j) = \sum_{j=1}^n \sum_{i=1}^n w_{ij} f(B_j) = \sum_{j=1}^n f(B_j).\end{aligned}$$

The proof follows from $\mathbb{A}^{-1} \prec^G \mathbb{B}^{-1}$. $\square$

Corollary 6.7. *Let $G \in \mathcal{C}_n$, $\mathbb{A} \prec^G \mathbb{B}$, $\alpha \ge 1$ and let Φ be symmetric gauge function on $\mathbb{R}^m$.*

(1) $\sum_{1\le i<j\le n} d_\Phi^\alpha(A_i, A_j) \le \sum_{1\le i<j\le n} d_\Phi^\alpha(B_i, B_j)$.

(2) *For any $X \in \mathbb{P}$,*

$$(d_\Phi^\alpha(A_1, X), \ldots, d_\Phi^\alpha(A_n, X)) \prec_w (d_\Phi^\alpha(B_1, X), \ldots, d_\Phi^\alpha(B_n, X)).$$

(3) *For any $X \in \mathbb{P}$ and symmetric gauge norm Ψ on $\mathbb{R}^n$,*

$$\Psi(d_\Phi^\alpha(A_1, X), \ldots, d_\Phi^\alpha(A_n, X)) \le \Psi(d_\Phi^\alpha(B_1, X), \ldots, d_\Phi^\alpha(B_n, X)).$$

(4) *If $G \in \mathcal{SC}_n$, then $\{A_1, \ldots, A_n\} \subset [B_1, \ldots, B_n]$.*

Proof. (1) Let $f_i(X) = d_\Phi(A_i, X)$ and $g_i(X) = d_\Phi(B_i, X)$. Set $f = \sum_{i=1}^n f_i^\alpha$ and $g = \sum_{i=1}^n g_i^\alpha$. Then f and g are continuous and geodesically convex functions. By Theorem 6.6,

$$\begin{aligned}\sum_{1\le i<j\le n} d_\Phi^\alpha(A_i, A_j) &= \sum_{j=1}^n \sum_{i=1}^n f_i^\alpha(A_j) = \sum_{j=1}^n f(A_j) \\ &\le \sum_{j=1}^n f(B_j) = \sum_{1\le i<j\le n} d_\Phi^\alpha(A_i, B_j) \\ &= \sum_{i=1}^n \sum_{j=1}^n g_j^\alpha(A_i) \\ &= \sum_{i=1}^n g(A_i) \le \sum_{i=1}^n g(B_i) \\ &= \sum_{1\le i<j\le n} d_\Phi^\alpha(B_i, B_j).\end{aligned}$$

(2) Let f be a continuous nondecreasing convex function defined on $[0, \infty)$. Then $f \circ d_\Phi^\alpha(\cdot, X)$ is a continuous geodesically convex function on $\mathbb{P}$. By Theorem 6.6,

$$\sum_{i=1}^n f(d_\Phi^\alpha(A_i, X)) \le \sum_{i=1}^n f(d_\Phi^\alpha(B_i, X)).$$

The conclusion then follows by the characterization of weak majorization on real numbers (M. Tomic and H. Weyl).

(3) This follows from (2).

(4) By Schur-convexity of G, $A_i = G(w_{i1}, \ldots, w_{in}; B_1, \ldots, B_n) \in [B_1, \ldots, B_n]$. $\square$

7. Schur's convexity for the Karcher mean

We note that $\mathbb{P}^n$ is a Hadamard space with the product metric

$$\delta(\mathbb{A}, \mathbb{B}) = \sqrt{\sum_{i=1}^{n} \delta^2(A_i, B_i)}.$$

We also note that a map $f : \mathbb{P}^n \to \mathbb{R}$ is geodesically convex if

$$f(A_1 \#_t B_1, \ldots, A_n \#_t B_n) \le (1-t) f(A_1, \ldots, A_n) + t f(B_1, \ldots, B_n)$$

for all $t \in [0, 1]$ and $(A_1, \ldots, A_n), (B_1, \ldots, B_n) \in \mathbb{P}^n$ and that the Karcher mean on the Hadamard space $\mathbb{P}^n$ is given by

$$\Lambda_k(\omega; \mathbb{A}_1, \ldots, \mathbb{A}_k) = (\Lambda_k(\omega; A_{11}, \ldots, A_{k1}), \ldots, \Lambda_k(\omega; A_{1n}, \ldots, A_{kn})) \tag{7.18}$$

where $\mathbb{A}_i = (A_{i1}, \ldots, A_{in}) \in \mathbb{P}^n$.

Let

$$S^n = \{\sigma_i : 1 \le i \le n!\},$$

the set of all permutations of n-letters. For $\sigma \in S^n$ and $\mathbb{A} = (A_1, \ldots, A_n) \in \mathbb{P}^n$, we denote

$$\mathbb{A}_\sigma = (A_{\sigma(1)}, \ldots, A_{\sigma(n)}) \in \mathbb{P}^n.$$

Theorem 7.1 (*Rado's Theorem*). *$\mathbb{A} \prec^{\Lambda_n} \mathbb{B}$ if and only if $\mathbb{A} = \Lambda_{n!}(\omega; \mathbb{B}_{\sigma_1}, \ldots, \mathbb{B}_{\sigma_{n!}})$ for some $\omega \in \overline{\Delta_{n!}}$.*

Proof. Let $\omega = (w_1, \ldots, w_{n!}) \in \overline{\Delta_{n!}}$. Suppose that $\mathbb{A} = \Lambda_{n!}(\omega; \mathbb{B}_{\sigma_1}, \ldots, \mathbb{B}_{\sigma_{n!}})$. We shall construct a doubly stochastic matrix $W = (p_{ij})_{n \times n}$ such that for all $i = 1, \ldots, n$,

$$A_i = \Lambda_n(p_{i1}, \ldots, p_{in}; B_1, \ldots, B_n)$$

which would imply $\mathbb{A} \prec^{\Lambda_n} \mathbb{B}$. Set

$$\alpha_{ij} = \{\sigma_k : \sigma_k(i) = j\}, \quad 1 \le i, j \le n$$

and $p_{ij} = \sum_{\sigma_k \in \alpha_{ij}} w_k = \sum_{\sigma_k(i)=j} w_k$. One can check that $W = (p_{ij})$ is a doubly stochastic matrix. By the permutation invariancy of the Karcher mean,

$$\begin{aligned} A_i &= \Lambda_{n!}(\omega; B_{\sigma_1(i)}, \ldots, B_{\sigma_{n!}(i)}) \\ &= \Lambda_n\left(\sum_{\sigma_k(i)=1} w_k, \ldots, \sum_{\sigma_k(i)=n} w_k; B_1, \ldots, B_n\right) \\ &= \Lambda_n(p_{i1}, \ldots, p_{in}; B_1, \ldots, B_n). \end{aligned}$$

Conversely, suppose that $\mathbb{A} \prec^{\Lambda_n} \mathbb{B}$. By Birkhoff's theorem on doubly stochastic matrices, the doubly stochastic matrix W is a linear combination of permutation matrices; there exists a probability vector $\mu = (u_1, \ldots, u_{n!})$ such that

$$W = (w_{ij})_{n \times n} = \sum_{k=1}^{n!} u_k P_k,$$

where P_k is the permutation matrix induced by the permutation σ_k. By hypothesis and the property of permutation matrices, we have

$$w_{ij} = \sum_{\sigma_k(i)=j} \mu_k$$

and

$$\begin{aligned} A_i &= \arg\min_{X \in \mathbb{P}} \sum_{i=1}^{n} w_{ij} \delta^2(X, B_j) \\ &= \arg\min_{X \in \mathbb{P}} \sum_{i=1}^{n} \left[\sum_{\sigma_k(i)=j} \mu_k\right] \delta^2(X, B_j) \\ &= \arg\min_{X \in \mathbb{P}} \sum_{i=1}^{n} \sum_{\sigma_k(i)=j} \mu_k \delta^2(X, B_j) \\ &= \arg\min_{X \in \mathbb{P}} \sum_{k=1}^{n!} \mu_k \delta^2(X, \mathbb{B}_{\sigma_k(i)}). \end{aligned}$$

By (7.18), $\mathbb{A} = (A_1, \ldots, A_n) \in \mathbb{P}^n$ is the μ-weighted Karcher mean of $\mathbb{B}_{\sigma_1}, \ldots, \mathbb{B}_{\sigma_{n!}}$. $\square$

Remark 7.2. We note that $\mathbb{A} \prec^{\Lambda_n} \mathbb{B}$ if and only if there exists $\omega \in \overline{\Delta_{n!}}$ such that

$$A_i = \Lambda_{n!}(\omega; B_{\sigma_1(i)}, \ldots, B_{\sigma_{n!}(i)}), \quad i = 1, \ldots, n,$$

that is,

$$\mathrm{diag}(A_1, \ldots, A_n) = \mathrm{diag}(\Lambda_{n!}(\omega; B_{\sigma_1(1)}, \ldots, B_{\sigma_{n!}(1)}), \ldots, \Lambda_{n!}(\omega; B_{\sigma_1(n)}, \ldots, B_{\sigma_{n!}(n)})).$$

By Remark 4.13, the right-hand of the equality coincides with

$$\Lambda_{n!}(\omega; Q_1, \ldots, Q_{n!}),$$

where $Q_i = \mathrm{diag}(B_{\sigma_i(1)}, \ldots, B_{\sigma_i(n)}),\ i = 1, \ldots, n!$ This shows that $\mathbb{A} \prec^{\Lambda_n} \mathbb{B}$ if and only if

$$\mathrm{diag}(A_1, \ldots, A_n) = \Lambda_{n!}(\omega; Q_1, \ldots, Q_{n!})$$

for some $\omega \in \overline{\Delta_{n!}}$, where $Q_i = \mathrm{diag}(B_{\sigma_i(1)}, \ldots, B_{\sigma_i(n)})$, which extends Rado's theorem on (positive) real numbers.

Corollary 7.3 (*Schur's Convexity*)**.** *If* $\mathbb{A} \prec^{\Lambda_n} \mathbb{B}$, *then*

(1) $f(\mathbb{A}) \le f(\mathbb{B})$ *for any continuous geodesically convex function* $f : \mathbb{P}^n \to \mathbb{R}$ *invariant under the permutation of coordinates;*
(2) $\mathbb{A}$ *lies in the convex hull of the* $n!$ *permutations of* $\mathbb{B}$ *in the product space* $\mathbb{P}^n$.

Proof. Follows from $\mathbb{A} = \Lambda_{n!}(\omega; \mathbb{B}_{\sigma_1}, \ldots, \mathbb{B}_{n!}) \in \mathbb{P}^n$ and

$$\begin{aligned} f(\mathbb{A}) &= f\left(\Lambda_{n!}(\omega; \mathbb{B}_{\sigma_1}, \ldots, \mathbb{B}_{\sigma_{n!}})\right) \\ &\le \sum_{k=1}^{n!} w_i f(\mathbb{B}_{\sigma_k}) = f(\mathbb{B}) \end{aligned}$$

where the inequality follows from the convexity of the Karcher mean on the Hadamard space $\mathbb{P}^n$ (Remark 5.5). □

Remark 7.4. Our Rado's theorem is new and holds on any Hadamard space by the same method of the proof of Theorem 7.1. In [27] Niculescu and Roventa have recently obtained several results along the same lines of this section for the setting of finite probability measures on Hadamard spaces. The Schur's convexity on a Hadamard space appears in [27] with a proof for $n = 3$.

Remark 7.5. The set of all weighted Karcher means $\Lambda_{n!}(\omega; \mathbb{B}_{\sigma_1}, \ldots, \mathbb{B}_{\sigma_{n!}})$ varying over $\omega \in \overline{\Delta_{n!}}$ is contained in the convex hull $[\mathbb{B}_{\sigma_1}, \ldots, \mathbb{B}_{\sigma_{n!}}] \subset \mathbb{P}^n$. It seems that the set of all weighted Karcher mean values in $\mathbb{P}^n$ is a proper subset of the convex hull $[\mathbb{B}_{\sigma_1}, \ldots, \mathbb{B}_{\sigma_{n!}}] \subset \mathbb{P}^n$.

Corollary 7.6. *Let* $\mathbb{A} \prec^{\Lambda_n} \mathbb{B}$. *Then*

$$\sum_{i=1}^{n} d_\Phi^\alpha(X, A_i) \le \sum_{i=1}^{n} d_\Phi^\alpha(X, B_i)$$

for all $\alpha \ge 1$, $X \in \mathbb{P}$ *and all symmetric gauge functions* Φ. *In particular,*

$$\sum_{i=1}^{n} \delta^2(\Lambda(\omega; \mathbb{A}), A_i) \le \sum_{i=1}^{n} \delta^2(\Lambda(\omega; \mathbb{B}), A_i) \le \sum_{i=1}^{n} \delta^2(\Lambda(\omega; \mathbb{B}), B_i) \tag{7.19}$$

for all $\omega \in \Delta_n$.

Proof. For a fixed $X \in \mathbb{P}$, the map $(Y_1, \ldots, Y_n) \mapsto \sum_{i=1}^{n} d_\Phi^\alpha(X, Y_i)$ is a geodesically convex. By the previous corollary, $\mathbb{A} \prec^{\Lambda_n} \mathbb{B}$ ensures the desired assertion. □

Remark 7.7. We know that the Karcher mean $\Lambda_n(\omega; \mathbb{A})$ is the unique minimizer of the objective function

$$f_{\mathbb{A}}(X) = \sum_{i=1}^{n} w_i \delta^2(X, A_i).$$

By the previous result, we have a nice relationship between the Karcher mean majorization $\mathbb{A} \prec^{\Lambda_n} \mathbb{B}$ and the minimum values of $f_{\mathbb{A}}$ and $f_{\mathbb{B}}$ as following: $\mathbb{A} \prec^{\Lambda_n} \mathbb{B}$ implies that $\min_{X\in\mathbb{P}} f_{\mathbb{A}}(X) \le \min_{X\in\mathbb{P}} f_{\mathbb{B}}(X)$.

Problem 3. Does Rado's (Schur's convexity) theorem hold for any $G \in \mathcal{SC}_n$ or for $G = \mathrm{Bmp}_n$? Alternatively, this property would provide a characteristic property of the Karcher mean among other symmetric weighted geometric means.

Acknowledgments

The author thanks Professor Jimmie Lawson for his insightful comments and suggestions. The author also thanks the referee for helpful comments. This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MEST) (No. 2012-005191).

References

[1] T. Ando, C.K. Li, R. Mathias, Geometric means, Linear Algebra Appl. 385 (2004) 305–334.
[2] M. Bessenyei, The Hermite–Hadamard inequality on simpleces, Amer. Math. Monthly 115 (2008) 339–345.
[3] R. Bhatia, Matrix Analysis, in: Graduate Texts in Mathematics, Springer, 1996.
[4] R. Bhatia, On the exponential metric increasing property, Linear Algebra Appl. 375 (2003) 211–220.
[5] R. Bhatia, Positive Definite Matrices, in: Princeton Series in Applied Mathematics, Princeton University Press, Princeton, NJ, 2007.
[6] R. Bhatia, The Riemannian mean of positive matrices, in: F. Nielsen, R. Bhatia (Eds.), Matrix Information Geometry, Springer, 2012, pp. 35–51.
[7] R. Bhatia, J. Holbrook, Riemannian geometry and matrix geometric means, Linear Algebra Appl. 413 (2006) 594–618.
[8] R. Bhatia, R. Karandikar, Monotonicity of the matrix geometric mean, Math. Ann. 353 (2012) 1453–1467.
[9] D. Bini, B. Meini, F. Poloni, An effective matrix geometric mean satisfying the Ando–Li–Mathias properties, Math. Comp. 79 (2010) 437–452.
[10] C. Conde, Nonpositive curvature in p-Schatten class, J. Math. Anal. Appl. 356 (2009) 664–673.
[11] C. Conde, A version of the Hermite–Hadamard inequality in a nonpositive curvature space, Banach J. Math. Anal. 6 (2012) 159–167.
[12] C. Conde, G. Larotonda, Manifolds of semi-negative curvature, Proc. Lond. Math. Soc. 100 (2010) 670–704.
[13] J. Holbrook, No dice: a determinstic approach to the Cartan centroid, J. Ramanujan Math. Soc. (in press).
[14] J. Jost, Nonpositive Curvature: Geometric and Analytic Aspects, in: Lectures in Mathematics, ETH Zürich, Birkhäuser, Basel, 1997.
[15] J. Lawson, H. Lee, Y. Lim, Weighted geometric means, Forum Math. 24 (2012) 1067–1090.
[16] J.D. Lawson, Y. Lim, The geometric mean, matrices, metrics, and more, Amer. Math. Monthly 108 (2001) 797–812.
[17] J. Lawson, Y. Lim, Monotonic properties of the least squares mean, Math. Ann. 351 (2011) 267–279.
[18] J. Lawson, Y. Lim, Karcher means and Karcher equations of positive definite operators, Trans. Amer. Math. Soc. (in press).
[19] H. Lee, Y. Lim, T. Yamazaki, Multi-variable weighted geometric means of positive definite matrices, Linear Algebra Appl. 435 (2011) 307–322.
[20] Y. Lim, On Ando–Li–Mathias geometric mean equations, Linear Algebra Appl. 428 (2008) 1767–1777.
[21] Y. Lim, Bushell's equations and polar decompositions, Math. Nachr. 282 (2009) 1567–1574.
[22] Y. Lim, Maximum-volume symmetric gauge ball problem on the convex cone of positive definite matrices and convexity of optimal sets, SIAM J. Optim. 21 (2011) 1275–1288.
[23] Y. Lim, A limit theorem for non-expansive matrix means (submitted for publication).
[24] Y. Lim, M. Pálfia, The matrix power means and the Karcher mean, J. Funct. Anal. 262 (2012) 1498–1514.
[25] Y. Lim, T. Yamazaki, On some inequalities for the matrix power and Karcher means, Linear Algebra Appl. 438 (2013) 1293–1304.
[26] M. Moakher, A differential geometric approach to the geometric mean of symmetric positive-definite matrices, SIAM J. Matrix Anal. Appl. 26 (2005) 735–747.
[27] C. Niculescu, I. Roventa, Majorization in spaces with a curved geometry, Preprint.
[28] K.-T. Sturm, Probability measures on metric spaces of nonpositive curvature, in: P. Auscher, et al. (Eds.), Heat Kernels and Analysis on Manifolds, Graphs, and Metric Spaces, in: Contemp. Math., vol. 338, Amer, Math. Soc. (AMS), Providence, 2003.
[29] T. Yamazaki, The Riemannian mean and matrix inequalities related to the Ando–Hiai inequality and Chaotic order, Oper. Matrices 6 (2012) 577–588.