On Ando–Li–Mathias geometric mean equations

Yongdo Lim *

Department of Mathematics, Kyungpook National University, Taegu 702-701, Republic of Korea

Received 1 August 2007; accepted 15 October 2007

Submitted by Chi-Kwong Li

Abstract

In this paper we consider a family of non-linear matrix equations based on the higher-order geometric means of positive definite matrices that proposed by Ando–Li–Mathias. We prove that the geometric mean equation

$$X = B + G(A_1, A_2, \ldots, A_m, X, X, \ldots, X)$$

has a unique positive definite solution depending continuously on the parameters of positive definite $A_i$ and positive semidefinite $B$. It is shown that the unique positive definite solutions $G_n(A_1, A_2, \ldots, A_m)$ for $B = 0$ satisfy the minimum properties of geometric means, yielding a sequence of higher-order geometric means of positive definite matrices.

© 2007 Elsevier Inc. All rights reserved.

Keywords: Positive definite matrix; Geometric mean; Non-linear matrix equation; Contraction; Fixed point; Thompson metric

1. Introduction

The matrix equation $X = Q - A^*X^{-1}A$ with $Q$ positive definite has been studied recently by several authors (see [1,6–8,10,19–22]). For the application areas in which the equations arise, see the references therein. As a special case, the non-linear matrix equation

$$X = T - BX^{-1}B$$

(1.1)
where $T$ is positive definite and $B$ is positive semidefinite, is solved by the author [18] via Anderson–Morley–Trapp [1] and Engwerda’s results (Theorem 11 of [7,8]): It has a positive definite solution if and only if $2B \leq T$, and the maximal and minimal positive definite solutions are explicitly described in terms of geometric mean of positive definite matrices:

$$X_+ = \frac{1}{2}(T + (T + 2B)\#(T - 2B)), \quad (1.2)$$

$$X_- = \frac{1}{2}(T - (T + 2B)\#(T - 2B)), \quad (1.3)$$

respectively, where $A\#B = A^{1/2}(A^{-1/2}BA^{-1/2})^{1/2}A^{1/2}$ denotes the geometric mean of positive definite matrices $A$ and $B$. Realizing the geometric mean $A\#B$ as a unique positive definite solution of the Riccati equation $XA^{-1}X = B$, Eq. (1.1) under the side conditions $T > 2B$ and $X > B$ is equivalent to the following geometric mean equation:

$$X = B + C\#X \quad (T = 2B + C).$$

Recently Ando–Li–Mathias [2] proposed a successful definition of geometric mean $G(A_1, A_2, \ldots, A_n)$ of $n$-positive definite matrices $A_i$ via symmetrization procedure. The main concern of this paper is the extended geometric mean equations based on the Ando–Li–Mathias’s geometric mean of several positive definite matrices:

$$X = B + G(A_1, A_2, \ldots, A_m, X, X, \ldots, X) \quad (n = 0, 1, 2, \ldots). \quad (1.4)$$

We show that Eq. (1.4) has a unique positive definite solution depending continuously on the parameters of positive definite $A_i$ and positive semidefinite $B$. For $B = 0$, the unique positive definite solution, denoted by $G_n(A_1, A_2, \ldots, A_m)$, is viewed as a matrix mean and satisfies all properties of the geometric mean of Ando–Li–Mathias presented in [2]. This provides a sequence of higher-order geometric means of positive definite matrices and yields a problem to distinguish these geometric means with that of Ando–Li–Mathias.

Throughout this paper, we assume that $\Omega = \Omega(k)$ is the convex cone of positive definite $k \times k$ Hermitian matrices. For Hermitian matrices $X$ and $Y$, we write that $X \preceq Y$ if $Y - X$ is positive semidefinite, and $X < Y$ if $Y - X$ is positive definite (positive semidefinite and invertible).

### 2. Higher order geometric mean

Let $(X, d)$ be a metric space. A $k$-mean on $X$ is a $k$-ary operation $\mu : X^k \to X$ that satisfies a generalized idempotency law: $\mu(x, \ldots, x) = x$ for all $x \in X$. We need some preliminaries: A $k$-mean on $X$ is called non-expansive if it satisfies for all $x = (x_1, \ldots, x_k), y = (y_1, \ldots, y_k) \in X^k$

$$d(\mu(x), \mu(y)) \leq d_\kappa(x, y) := \max_{1 \leq j \leq k} d(x_j, y_j). \quad (2.5)$$

For $0 < \rho < 1$, a $k$-mean $\mu$ on $X$ is called coordinatewise $\rho$-contractive if for any $x, y \in X^k$ that differ only in one coordinate, say $x_j \neq y_j$,

$$d(\mu(x), \mu(y)) \leq \rho d(x_j, y_j).$$

Moreover, the barycentric operator $\beta : X^{k+1} \to X^{k+1}$ is defined by

$$\beta(x) = (\mu(\pi_{\neq 1}x), \ldots, \mu(\pi_{\neq k+1}x)).$$
where \( x = (x_1, \ldots, x_{k+1}) \) and \( \pi_{\neq j} x = (x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_{k+1}) \in X^k \).

Then Lawson and Lim [17] showed as follows.

**Theorem 2.1.** Let \((X, d)\) be a complete metric space equipped with a non-expansive coordinate-wise \(\rho\)-contractive \(k\)-mean \(\mu : X^k \to X, k \geq 2\) and \(0 < \rho < 1\). Then the barycentric operator \(\beta\) is power convergent in the sense that for each \(x = (x_1, \ldots, x_{k+1}) \in X^{k+1}\) there exists \(x^* \in X\) such that \(\lim_{n \to \infty} \beta^n(x) = (x^*, \ldots, x^*)\).

Moreover, there exists a (unique) continuous \((k+1)\)-mean \(\tilde{\mu} : X^{k+1} \to X\) that \(\beta\)-extends \(\mu\) and \(\tilde{\mu}\) is non-expansive and coordinate \(\rho\)-contractive.

Let \(\Omega = \Omega(k)\) be the convex cone of positive definite \(k \times k\) Hermitian matrices. Let \(A\) and \(B\) be positive definite Hermitian matrices. The Thompson metric \(d\) is defined by

\[
d(A, B) = \| \log A^{-1/2} BA^{-1/2} \|.
\]

Then \((\Omega, d)\) is a complete metric space [24]. The geometric mean of \(A\) and \(B\) in the sense of Kubo–Ando [12] is defined by

\[
A \# B = A^{1/2}(A^{-1/2} BA^{-1/2})^{1/2} A^{1/2}.
\]

It follows that the geometric mean \(\#\) is a 2-mean on \(\Omega\). Moreover, Corach et al. [5] showed as follows:

\[
d(A_1 \# B_1, B_1 \# B_2) \leq \frac{1}{2} d(A_1, B_1) + \frac{1}{2} d(A_2, B_2). \quad (2.6)
\]

See also [3,4,15,16,17] for more general setting.

Therefore, it easily follows that the geometric mean \(\#\) is non-expansive and coordinatewise \(\frac{1}{2}\)-contractive.

In [2], Ando–Li–Mathias proposed a successful definition of geometric mean of several positive definite matrices via the symmetrization procedure. Let \(A_1, A_2\) and \(A_3\) be positive definite matrices. By Theorem 2.1, the symmetrization procedure of the geometric mean

\[
\beta(A_1, A_2, A_3) = (A_2 \# A_3, A_1 \# A_3, A_1 \# A_2)
\]

is power convergent. There is, there exists a positive definite matrix \(G(A_1, A_2, A_3)\) such that

\[
\lim_{n \to \infty} \beta^n(A_1, A_2, A_3) = (G(A_1, A_2, A_3), G(A_1, A_2, A_3), G(A_1, A_2, A_3)).
\]

By Theorem 2.1, inductively the geometric mean is extended to all orders, defining \(m\)-geometric mean \(G(A_1, A_2, \ldots, A_m)\).

The geometric mean defined above has the following properties [2]:

- **(P1)** Consistency with scalars. If \(A_i\)’s are mutually commute then
  \[G(A_1, A_2, \ldots, A_m) = (A_1 \cdots A_m)^{1/m} \]

- **(P2)** Permutation invariance. For any permutation \(\sigma\),
  \[G(A_1, A_2, \ldots, A_m) = G(A_{\sigma(1)}, A_{\sigma(2)}, \ldots, A_{\sigma(m)}) \]

- **(P3)** Congruence invariance. For an invertible operator \(M\),
  \[G(MA_1 M^*, MA_2 M^*, \ldots, MA_m M^*) = MG(A_1, A_2, \ldots, A_m)M^* \]
(P4) Monotonicity. If \( A_i \leq B_i \) for all \( i \), then
\[
G(A_1, A_2, \ldots, A_m) \leq G(B_1, B_2, \ldots, B_m).
\]

(P5) Self-duality. \( G(A_1, A_2, \ldots, A_m)^{-1} = G(A_1^{-1}, A_2^{-1}, \ldots, A_m^{-1}) \).

(P6) Continuity. The function \( (A_1, A_2, \ldots, A_m) \mapsto G(A_1, A_2, \ldots, A_m) \) is continuous from \( \Omega^m \rightarrow \Omega \).

(P7) Determinant identity. For positive definite matrices \( A_i \),
\[
\det G(A_1, A_2, \ldots, A_m) = (\det A_1 \det A_2 \cdots \det A_m)^{1/m}.
\]

(P8) Joint homogeneity. For positive scalars \( s_i \),
\[
G(s_1 A_1, s_2 A_2, \ldots, s_m A_m) = (s_1 s_2 \cdots s_m)^{1/m} G(A_1, A_2, \ldots, A_m).
\]

(P9) Joint concavity. For \( 0 < \lambda < 1 \),
\[
G(\lambda A_1 + (1-\lambda) B_1, \lambda A_2 + (1-\lambda) B_2, \ldots, \lambda A_m + (1-\lambda) B_m) \geq \lambda G(A_1, A_2, \ldots, A_m) + (1-\lambda) G(B_1, B_2, \ldots, B_m).
\]

(P10) Arithmetic–geometric–harmonic mean inequalities.
\[
m(A_1^{-1} + A_2^{-1} + \cdots + A_m^{-1})^{-1} \leq G(A_1, A_2, \ldots, A_m) \leq \frac{A_1 + A_2 + \cdots + A_m}{m}.
\]

For positive semidefinite matrices \( A_1, A_2, \ldots, A_m \), their geometric mean can be determined by
\[
G(A_1, A_2, \ldots, A_m) = \lim_{\epsilon \downarrow 0} G(A_1 + \epsilon I, A_2 + \epsilon I, \ldots, A_m + \epsilon I).
\]

The following result is new and will be useful for our purpose.

**Proposition 2.2.** Let \( A, B \) be positive definite matrices and let \( t_1, t_2, \ldots, t_m \in [0, 1] \). Then
\[
G(A^\# t_1 B, A^\# t_2 B, \ldots, A^\# t_m B) = A^\# \frac{1}{m} \sum_{i=1}^{m} t_i B.
\]

In particular,
\[
G(A, A, \ldots, A, B, B, \ldots, B) = A^{\# p/m} B, \quad (2.7)
\]
\[
G(A, B, A^{\# B}) = A^{\# B}. \quad (2.8)
\]

**Proof.** This follows from the affine change of parameter \( (A^\# t B)^\# (A^\# t B) = A^\# (t(t+t)/2) B \) (cf. [9,14]) and by induction. In this case, the symmetrization procedure is same as that of arithmetic means of real numbers \( t_i \). (2.7) follows from \( A = A^{\# 0} B, B = A^{\# 1} B \) (cf. [11]). \( \square \)

3. **Geometric mean equations**

The following results will play a crucial role for the continuity of solution maps on geometric mean equations.

**Definition 3.1.** Let \( (X, d) \) be a metric space. A mapping \( f : X \rightarrow X \) is a strict contraction if there exists \( 0 \leq \lambda < 1 \) such that \( d(f(x), f(y)) \leq \lambda d(x, y) \) for all \( x, y \in X \). The least contraction coefficient (Lipschitz constant) of \( f \) is defined by
Proposition 3.2 [23, Proposition II.6]. Let \((X, d)\) be a complete metric space, \(0 \leq \lambda < 1\), and \(\mathcal{C}_\lambda(X) = \{ f : X \to X : L(f) \leq \lambda \}\). For \(f \in \mathcal{C}_\lambda(X)\) let \(p(f) \in X\) denote the unique fixed point of \(f\). If we endow \(\mathcal{C}_\lambda(X)\) with the topology of pointwise convergence, then the fixed point map \(p : \mathcal{C}_\lambda(X) \to X\) is continuous.

Proof. Pick a convergence subsequence \(L(f_{n_\alpha}) \to L_0 \leq L\). From
\[
d(f(x), f(y)) = \lim_{n_\alpha \to \infty} d(f_{n_\alpha}(x), f_{n_\alpha}(y)) \leq \lim_{n_\alpha \to \infty} L(f_{n_\alpha}) d(x, y) \leq L_0 d(x, y),
\]
we have \(L(f) \leq L_0\) and hence \(f\) is a strict contraction. The convergence of \(p(f_{n_\alpha})\) to \(p(f)\) follows from Proposition 3.2.

□

Theorem 3.4. Let \((A_1, A_2, \ldots, A_m) \in \Omega^m\). Then for each non-negative integer \(n\), the equation
\[
G(A_1, A_2, \ldots, A_m, X, X, \ldots, X)^n = X
\]
has a unique positive definite solution depending continuously on the parameters \(A_1, A_2, \ldots, A_m\).

Proof. We may assume that \(n \geq 1\). We will show that
\[
d(G(\underbrace{A, X, X, \ldots, X}_n), G(\underbrace{A, Y, Y, \ldots, Y}_n)) \leq \frac{n}{n + 1} d(X, Y)
\]
for any \(X, Y > 0\) and \(\underbrace{A, (A_1, \ldots, A_m) \in \Omega^m, \text{for the Thompson metric } d(A, B). \text{This implies}}_{\text{for all } \underbrace{A, \in \Omega^m \text{ and therefore the fixed point of } f_{\lambda} \text{ varies continuously on } A, \in \Omega^m}}\)
a unique positive definite fixed point by completeness of the metric. Then the uniqueness and existence of the positive definite solution of (3.9) are immediate. Furthermore, \(f_{\lambda} \in \mathcal{C}_{\frac{n}{n+1}}(\Omega)\) for all \(A, \in \Omega^m\) and therefore the fixed point of \(f_{\lambda}\) varies continuously on \(A, \in \Omega^m\) (Proposition 3.2).

The proof proceeds by induction on \(m\).

(1) \(m = 1\). In this case by Proposition 2.2, \(G(A_1, X, X, \ldots, X) = A_1 \#_{\frac{n}{n+1}} X\) and therefore
\[
d(G(\underbrace{A_1, X, X, \ldots, X}_n), G(\underbrace{A_1, Y, Y, \ldots, Y}_n)) = d(A_1 \#_{\frac{n}{n+1}} X, A_1 \#_{\frac{n}{n+1}} Y)
\]
\[
\leq \frac{n}{n + 1} d(X, Y),
\]
where the inequality follows from (2.6).

(2) Suppose that (3.10) holds true for all \(m - 1\) tuples of positive definite operators and for all \(n\). Let \(\underbrace{A, = (A_1, A_2, \ldots, A_m) \in \Omega^m\). We will show that by induction on \(n\)
\[ d(G(\mathbb{A}, X \cdot 1_n), G(\mathbb{A}, X \cdot 1_n)) \leq \frac{n}{n+1} d(X, Y) \]

for all \( X, Y > 0 \) and \( n = 1, 2, \ldots \) Here we denote \( X \cdot 1_n = (X, X, \ldots, X) \in \Omega^n \). By the coordinative \( 1/2 \)-contractive property of the symmetrization, \( d(G(\mathbb{A}, X), G(\mathbb{A}, Y)) \leq \frac{1}{2} d(X, Y) \) and hence our assertion holds true for \( n = 1 \). Suppose that

\[ d(G(\mathbb{A}, X \cdot 1_{n-1}), G(\mathbb{A}, X \cdot 1_{n-1})) \leq \frac{n-1}{n} d(X, Y) \]

for all \( X, Y \in \Omega \). We will show that

\[ d(G(\mathbb{A}, X \cdot 1_n), G(\mathbb{A}, X \cdot 1_n)) \leq \frac{n}{n+1} d(X, Y). \]

Consider the symmetrization on \( \Omega^{m+n} \):

\[ \beta(\mathbb{A}, X \cdot 1_n) = \left( G(\pi_{\neq 1} \mathbb{A}, X \cdot 1_n), \ldots, G(\pi_{\neq m} \mathbb{A}, X \cdot 1_n), \right. \]

\[ \left. G(\mathbb{A}, X \cdot 1_{n-1}), \ldots, G(\mathbb{A}, X \cdot 1_{n-1}) \right). \]

where \( \pi_{\neq i} \mathbb{A} = (A_1, \ldots, A_{i-1}, A_{i+1}, \ldots, A_m) \in \Omega^{m-1} \). By induction,

\[ d(G(\pi_{\neq i} \mathbb{A}, X \cdot 1_n), G(\pi_{\neq i} \mathbb{A}, Y \cdot 1_n)) \leq \frac{n}{n+1} d(X, Y), \]

\[ d(G(\mathbb{A}, X \cdot 1_{n-1}), G(\mathbb{A}, Y \cdot 1_{n-1})) \leq \frac{n-1}{n} d(X, Y) \leq \frac{n}{n+1} d(X, Y). \]

This implies that \( d_s(\beta(\mathbb{A}, X \cdot 1_n), \beta(\mathbb{A}, Y \cdot 1_n)) \leq \frac{n}{n+1} d(X, Y) \). The non-expansive property of the symmetrization \( \beta \) for the sup metric (2.5) implies that

\[ d_s(\beta^k(\mathbb{A}, X \cdot 1_n), \beta^k(\mathbb{A}, Y \cdot 1_n)) \leq d_s(\beta^{k-1}(\mathbb{A}, X \cdot 1_n), \beta^{k-1}(\mathbb{A}, Y \cdot 1_n)) \leq \cdots \]

\[ \leq d_s(\beta(\mathbb{A}, X \cdot 1_n), \beta(\mathbb{A}, Y \cdot 1_n)) \leq \frac{n}{n+1} d(X, Y) \]

for all \( k \) and hence limiting and projecting into the first coordinate yield

\[ d(G((\mathbb{A}, X \cdot 1_n), G(\mathbb{A}, Y \cdot 1_n)) \leq \frac{n}{n+1} d(X, Y). \]

This completes the proof. \( \square \)

**Remark 3.5.** The inequality (3.10) follows from Theorem 3.2 of [2]: In fact

\[ d(G(\mathbb{A}, X, X, \ldots, X), G(\mathbb{A}, Y, Y, \ldots, Y)) \leq \frac{n}{n+m} d(X, Y) \leq \frac{n}{n+1} d(X, Y) \]

for any \( \mathbb{A} \in \Omega^m \).

**Corollary 3.6.** Suppose that either of the following conditions holds:

1. \( B \) is positive semidefinite and \( A_i \)'s are positive definite for \( i = 1, 2, \ldots, m \),
2. \( B \) is positive definite and \( A_i \)'s are positive semidefinite \( i = 1, 2, \ldots, m \).
Then the following non-linear matrix equation

\[ X = B + G(A_1, A_2, \ldots, A_m, X, X, \ldots, X) \]

for \( n = 0, 1, 2, \ldots \) (3.11)

has the unique positive definite solution depending continuously on the parameters \( B, A_i \).

**Proof.** We first observe that each translation \( X \mapsto B + X \) for positive semidefinite \( B \) is non-expansive on \( \mathbb{H} \) with respect to the Thompson metric, that is, \( d(B + A, B + C) \leq d(A, C) \) (cf. [17]). Indeed, the Thompson metric is alternatively expressed by

\[ d(A, B) = \max\{ \log M(A/B), \log M(B/A) \} \]

where \( M(A/B) := \inf \{ \lambda > 0 : A \leq \lambda B \} \), the largest eigenvalue of \( B^{-1/2}AB^{-1/2} \). Then there exists \( r \geq 1 \) such that \( \log r = d(A, C) \). Then \( A \leq rC \), and thus \( B + A \leq B + rC \leq r(B + C) \), and similarly \( C \leq rA \) implies \( B + C \leq r(B + A) \). Hence \( d(B + A, B + C) \leq \log r = d(A, C) \).

Let \( f(X) := G(A_1, A_2, \ldots, A_m, X, X, \ldots, X) \).

Case 1: In this case \( f \) is a strict contraction for the Thompson part metric from (3.10) and the first paragraph:

\[ d(B + f(X), B + f(Y)) \leq d(f(X), f(Y)) \leq \frac{n}{n+1} d(X, Y) \]

for all \( X, Y > 0 \). Therefore, the mapping \( X \mapsto B + f(X) \) is a strict contraction and hence it has a unique positive definite fixed point which coincides with the unique positive definite solution of Eq. (3.11).

Case 2: In this case the mapping \( X \mapsto B + f(X) \) is a self-map on \( \mathbb{H} \). Let \( f_k(X) = G(A_1 + I/k, A_2 + I/k, \ldots, A_m + I/k, X, X, \ldots, X) \). Then \( f_k \) is a strict contraction on \( \mathbb{H} \) with \( L(f_k) \leq \frac{n}{n+1} \). Since \( B + f_k(X) \to B + f(X) \) (pointwise), we can apply Proposition 3.3: the map \( X \mapsto B + f(X) \) is a strict contraction.

The continuity of fixed point for both cases follows from Propositions 3.2 and 3.3. \( \square \)

**Example 3.7** (\( m = n = 1 \)). The non-linear matrix equation

\[ X = B + A\#X \]

has the unique positive definite solution \( X = \frac{1}{2}(A + 2B + (A + 4B)\#A) \) for either \( A > 0 \) and \( B \geq 0 \) or \( B > 0 \) and \( A \geq 0 \). Indeed, from the Riccati Lemma (cf. [13]) which says that \( X = A\#B \) is a unique positive definite solution of the Riccati equation \( XA^{-1}X = B \), the equation is equivalent to the non-linear equation (1.1)

\[ X = 2B + A - BX^{-1}B \]

with a side condition \( X > B \). Suppose that \( A > 0 \). The maximal solution of (3.13) which is given by \( X_+ = \frac{1}{2}(2B + A + (A + 4B)\#A) \) satisfies the side condition \( X_+ > B \), and therefore it coincides with the unique solution of (3.12). If \( A \geq 0 \) (but \( B > 0 \)) then by taking a sequence of positive definite matrices \( A_n \) converging to \( A \) and by the continuity of fixed points we have

\[ X = \lim_{n \to \infty} \frac{1}{2}(2B + A_n + (A_n + 4B)\#A_n) = \frac{1}{2}(2B + A + (A + 4B)\#A). \]
Remark 3.8. When $A$ is positive definite, Eq. (3.12) is equivalent to an algebraic Riccati equation
\[ Y^2 - CY - YC + D = 0. \] (3.14)
To see this, set $Y = A^{-1/2}XA^{-1/2}$, $C = A^{-1/2}BA^{-1/2} + \frac{1}{2}I$ and $D = (A^{-1/2}BA^{-1/2})^2$. Then by the Riccati Lemma, $(X - B)A^{-1}(X - B) = X$ and hence
\[ XA^{-1}X - BA^{-1}X - XA^{-1}B + BA^{-1}B = X. \]
Taking the congruence transformation by $A^{-1/2}$ both sides yields (3.14).

4. Geometric mean properties

Let $A_1, A_2, \ldots, A_m$ be positive definite matrices. By Theorem 3.4, the non-linear matrix equation
\[ G(A_1, A_2, \ldots, A_m, X, X, \ldots, X) = X \]
has a unique positive definite solution, which is denoted by $G_n(A_1, A_2, \ldots, A_m)$.

We observe from Proposition 2.2 and $A = A\#_0 B$, $B = A\#_1 B$ that if $X = A\#_p/m B$, then
\[ G(A, A, \ldots, A, B, B, \ldots, B, X, X, \ldots, X) = A\#_\frac{1}{m+p} (p+np/m) B = A\#_p/m B = X. \]
Therefore,
\[ G_n(A, A, \ldots, A, B, B, \ldots, B) = A\#_p/m B = G(A, A, \ldots, A, B, B, \ldots, B). \] (4.15)
In particular, $G_n(A, A, \ldots, A) = A$ and $G_n(A, B) = A\# B = G(A, B)$ for all non-negative integer $n$ by (4.15). Furthermore from $G_0(A_1, A_2, \ldots, A_m) = G(A_1, A_2, \ldots, A_m)$, it can be viewed as a geometric mean of several positive definite matrices. Indeed, the following shows that the matrix mean $G_n(A_1, A_2, \ldots, A_m)$ satisfies the minimum properties of geometric means.

Proposition 4.1. The matrix mean $G_n(A_1, A_2, \ldots, A_m)$ satisfies all the properties (P1)–(P10) of the Ando–Li–Mathias’s geometric mean.

Proof

(P1′) Suppose that $A_i$’s are mutually commute. Let $X = (A_1A_2 \cdots A_m)^{1/m}$. Then
\[ G(A_1, A_2, \ldots, A_m, X, X, \ldots, X) = (A_1A_2 \cdots A_m X^n)^{\frac{1}{m+n}} \]
\[ = (A_1A_2 \cdots A_m)^{\frac{1}{m+n}} X^{\frac{n}{m+n}} \]
\[ = (A_1A_2 \cdots A_m)^{\frac{1}{m+n} + \frac{n}{m(m+n)}} \]
\[ = (A_1A_2 \cdots A_m)^{1/m} = X \]
and therefore $G_n(A_1, A_2, \ldots, A_m) = (A_1A_2 \cdots A_m)^{1/m}$. 
(P2') This follows from the permutation invariancy of the geometric mean:
\[ G(A_1, A_2, \ldots, A_m, X, X, \ldots, X) = G(A_{\sigma(1)}, A_{\sigma(2)}, \ldots, A_{\sigma(m)}, X, X, \ldots, X). \]

(P3') Follows from the congruence transformation invariancy of the geometric mean:
\[ X = G(A_1, A_2, \ldots, A_m, X, X, \ldots, X) \text{ if and only if } \]
\[ M X M^* = G(M A_1 M^*, M A_2 M^*, \ldots, M A_m M^*, M X M^*, M X M^*, \ldots, M X M^*). \]

(P4') Suppose that \( A_i \leq B_i, i = 1, 2, \ldots, m \). Let \( g(X) = G(B_1, B_2, \ldots, B_m, X, X, \ldots, X) \).
Then \( g \) is monotone increasing, i.e., \( X \leq Y \) then \( g(X) \leq g(Y) \) by the monotonicity of the geometric mean. Let \( X = G_n(A_1, A_2, \ldots, A_m) \). Then again by the monotonicity of the geometric mean,
\[ X = G(A_1, A_2, \ldots, A_m, X, X, \ldots, X) \leq G(B_1, B_2, \ldots, B_m, X, X, \ldots, X) = g(X) \]
and therefore \( X \leq g(X) \leq g^2(X) \leq \cdots \leq g^k(X) \to G_n(B_1, B_2, \ldots, B_m). \)

(P5') By the self-duality of the geometric mean, \( X = G_n(A_1, A_2, \ldots, A_m) \) is the unique fixed point of a strict contraction (Theorem 3.4) with least contraction coefficient \( \frac{n}{n+1} \). It then follows by Proposition (3.2).

(P6') We observe that \( X = G_n(A_1, A_2, \ldots, A_m) \) is the unique fixed point of a strict contraction (Theorem 3.4) with least contraction coefficient \( \frac{n}{n+1} \). Then \( X \geq g(X) \geq g^2(X) \to G_n(B_1, B_2, \ldots, B_m). \)

(P7') Follows from the determining equation of \( G_n \) and the determinant formula of the geometric mean.

(P8') Let \( X = G_n(A_1, A_2, \ldots, A_m) \). Then
\[ (s_1 s_2 \cdots s_m)^{1/m} X = (s_1 s_2 \cdots s_m)^{1/m} G(A_1, A_2, \ldots, A_m, X, X, \ldots, X) \]
\[ = G(s_1 A_1, s_2 A_2, \ldots, s_m A_m, (s_1 s_2 \cdots s_m)^{1/m} X \cdot 1_n). \]

(P9') Let \( X = G_n(A_1, A_2, \ldots, A_m), Y = G_n(B_1, B_2, \ldots, B_m) \), and let
\[ f(Z) = G(\lambda A_1 + (1 - \lambda) B_1, \lambda A_2 + (1 - \lambda) B_2, \ldots, \lambda A_m + (1 - \lambda) B_m, Z, Z, \ldots, Z). \]
Then the joint concavity of the geometric mean implies that
\[ \lambda X + (1 - \lambda) Y = \lambda G(A_1, A_2, \ldots, A_m, X, X, \ldots, X) \]
\[ + (1 - \lambda) G(B_1, B_2, \ldots, B_m, Y, Y, \ldots, Y) \leq f(\lambda X + (1 - \lambda) Y) \]
and hence by the monotone property of \( f \) we have
\[
\lambda X + (1 - \lambda)Y \leq f^k(\lambda X + (1 - \lambda)Y) \\
\rightarrow G_n(\lambda A_1 + (1 - \lambda)B_1, \lambda A_2 + (1 - \lambda)B_2, \ldots, \lambda A_m + (1 - \lambda)B_m).
\]

(P10') Follows from (P10) and the self-duality of the mean \( G_n. \)

\[
X = G(A_1, A_2, \ldots, A_m, X, X, \ldots, X) \leq \frac{1}{m+n} (A_1 + A_2 + \cdots + A_m + nX)
\]
implies that \( X \leq \frac{1}{m} (A_1 + A_2 + \cdots + A_m). \)

Remark 4.2. One can directly see that
\[
G_n(A_{i_1} B, A_{i_2} B, \ldots, A_{i_m} B) = G(A_{i_1} B, A_{i_2} B, \ldots, A_{i_m} B) = A + \sum_{i=1}^m t_i B.
\]

From \( G = G_0 \) and \( G_n(A, B) = G(A, B) \), one can expect that \( G = G_n \) for any positive integers. Computer simulations (programmed in MatLab) for \( 2 \times 2 \) matrices \( A, B \) of determinant 1 with \( \text{tr}(A) = \text{tr}(B) = \text{tr}(A^{-1} B) \) show that \( G_n(A, B, I_2) = G(A, B, I_2) = \frac{A + B + I_2}{\sqrt{\det(A + B + I_2)}}. \) We do not have a proof for this and general cases.

Acknowledgements

The author wishes to thank the referee for making useful suggestions, which lead to improvements in the presentation. This research was supported by Kyungpook National University Research Fund, 2007.

References


