J. Appl. Math. & Informatics Vol. **33**(*2015*), *No. 3 - 4, pp. 379 - 386*
http://dx.doi.org/10.14317/jami.2015.379

A NEW PROOF TO CONSTRUCT MULTIVARIABLE GEOMETRIC MEANS BY SYMMETRIZATION†

SEJONG KIM* AND DÉNES PETZ

Abstract. The original geometric mean of two positive definite operators A and B is given by

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

In this article we provide a new proof to construct from the two-variable geometric mean to the multivariable mean via symmetrization introduced by Lawson and Lim [5]. Finally we provide an algorithm to find three-variable geometric mean via symmetrization, which plays an important role to construct higher-order geometric means.

AMS Mathematics Subject Classification : 47B65; 47H07.
Key words and phrases : positive definite operator; operator mean.

1. Introduction

As a typical value of a finite number of positive real numbers, average or mean plays an important role in probability theory, statistics, and economics. For instance, the arithmetic, geometric, and harmonic means have been commonly used:

$$A(x_1, \dots, x_n) = \frac{1}{n}\sum_{j=1}^{n} x_j,$$

$$G(x_1, \dots, x_n) = \left[\prod_{j=1}^{n} x_j\right]^{1/n},$$

$$H(x_1, \dots, x_n) = \left[\frac{1}{n}\sum_{j=1}^{n} x_j^{-1}\right]^{-1}.$$

Received January 13, 2015. Revised March 18, 2015. Accepted March 20, 2015. *Corresponding author. †This work was supported by the research grant of the Chungbuk National University in 2013.

One can construct a definition of means for positive real numbers as following.

Definition 1.1. Let $\mathbb{R}_+$ be the set of all positive real numbers. A function $M : \mathbb{R}_+^n \to \mathbb{R}_+$ is called a *mean* of positive real numbers if

(1) $M(x, \dots, x) = x$ for any $x \in \mathbb{R}_+$;
(2) $M(\alpha x_1, \dots, \alpha x_n) = \alpha M(x_1, \dots, x_n)$ for any $\alpha \in \mathbb{R}_+$;
(3) $M(x_1, \dots, x_n) = M(x_{\sigma(1)}, \dots, x_{\sigma(n)})$ for any permutation σ on $\{1, \dots, n\}$;
(4) $M(x_1, \dots, x_n) \leq M(y_1, \dots, y_n)$ whenever $x_j \leq y_j$ for all $j = 1, \dots, n$; and
(5) M is continuous in each variable.

The three of the most familiar means listed above satisfy the axioms of means and hold the inequality:

$$H(x_1, \dots, x_n) \leq G(x_1, \dots, x_n) \leq A(x_1, \dots, x_n).$$

This n-variable mean can be naturally defined for positive definite bounded operators. The arithmetic and harmonic means of positive definite operators can be defined as the same as those of positive real numbers, but it is not the case of the geometric mean because of non-commutativity. The purpose of this paper is to suggest a new method to construct $(n+1)$-variable geometric mean from n-variable geometric mean.

Let $\mathbb{P}$ be the open convex cone of positive definite bounded operators. For self-adjoint operators X and Y we define $X \leq Y$ if and only if $Y - X$ is positive semidefinite, and $X < Y$ if and only if $Y - X$ is positive definite. This relation, known as the *Löewner order*, gives a partial order on $\mathbb{P}$.

2. Two-variable geometric mean

The original *geometric mean* of positive definite operators A and B

$$A\#B := A^{1/2}(A^{-1/2}BA^{-1/2})^{1/2}A^{1/2} \tag{1}$$

was introduced by Kubo and Ando in [4], and its several properties have been found: see the references [1] and [3]. One can naturally define the weighted geometric mean of positive definite operators A and B such as

$$A\#_t B := A^{1/2}(A^{-1/2}BA^{-1/2})^{t}A^{1/2}, \tag{2}$$

where $t \in [0, 1]$. If A and B are not invertible, then we can take

$$A\#_t B := \lim_{\epsilon \to +0} (A + \epsilon I)\#_t(B + \epsilon I).$$

We list some properties of the weighted geometric mean.

Lemma 2.1. *Let $A, B, C, D \in \mathbb{P}$ and let $s, t, u \in [0, 1]$. Then the following are satisfied.*

(1) $A\#_t B = A^{1-t}B^t$ *if A and B commute.*
(2) $(aA)\#_t(bB) = a^{1-t}b^t(A\#_t B)$ *for any $a, b > 0$.*
(3) $A\#_t B \leq C\#_t D$ *whenever $A \leq C$ and $B \leq D$.*

(4) $P(A\#_t B)P^\dagger = (PAP^\dagger)\#_t(PBP^\dagger)$ *for any invertible operator* P.
(5) $A\#_t B = B\#_{1-t}A$.
(6) $(A\#_t B)^{-1} = A^{-1}\#_t B^{-1}$.
(7) $(A\#_s B)\#_t(A\#_u B) = A\#_{(1-t)s+tu}B$.
(8) $[(1-\lambda)A+\lambda B]\#_s[(1-\lambda)C+\lambda D] \geq (1-\lambda)(A\#_s C)+\lambda(B\#_s D)$ *for any* $\lambda \in [0,1]$.
(9) $[(1-t)A^{-1}+tB^{-1}]^{-1} \leq A\#_t B \leq (1-t)A+tB$ *for any* $t \in [0,1]$.

3. A new proof of extension

We now present a new extension of two-variable geometric mean to multivariable geometric mean. Let Δ_n be a set of all positive probability vectors in $\mathbb{R}^n$, that is, $\omega = (w_1, \ldots, w_n) \in \Delta_n$ means that $w_j > 0$ for all $j = 1, \ldots, n$ and $\sum_{j=1}^{n} w_j = 1$.

Let $\mathbb{A} = (A_1, \ldots, A_n), \mathbb{B} = (B_1, \ldots, B_n) \in \mathbb{P}^n$ and $\omega = (w_1, \ldots, w_n) \in \Delta_n$. We consider an operator geometric mean $G : \Delta_n \times \mathbb{P}^n \to \mathbb{P}$ satisfying

(G1) G is *idempotent*: for any $A \in \mathbb{P}$

$$G(\omega; A, \ldots, A) = A,$$

(G2) G is *jointly homogeneous*: for all $a_j > 0$

$$G(\omega; a_1A_1, \ldots, a_nA_n) = \left(\prod_{j=1}^{n} a_j^{w_j}\right) G(\omega; \mathbb{A}),$$

(G3) G is *permutation invariant*: for any permutation σ on $\{1, \ldots, n\}$

$$G(\omega; \mathbb{A}) = G(\omega_\sigma; \mathbb{A}_\sigma),$$

where $\mathbb{A}_\sigma = (A_{\sigma(1)}, \ldots, A_{\sigma(n)})$,

(G4) G is *monotone*: if $A_j \leq B_j$ for all $j = 1, \ldots, n$, then

$$G(\omega; \mathbb{A}) \leq G(\omega; \mathbb{B}).$$

For a uniform probability vector $\omega = (1/n, \ldots, 1/n)$ we simply write $G(\mathbb{A}) = G(\omega; \mathbb{A})$.

Lemma 3.1. [6, Proposition 2.5] *Let* $G : \Delta_n \times \mathbb{P}^n \to \mathbb{P}$ *be jointly homogeneous and monotone. Then the following contractive property for the Thompson metric is satisfied:*

$$d(G(\omega; \mathbb{A}), G(\omega; \mathbb{B})) \leq \sum_{j=1}^{n} w_j d(A_j, B_j),$$

where $d(A,B) := \| \log(A^{-1/2}BA^{-1/2})\|$ *for the operator norm* $\| \cdot \|$.

Let $\mathbb{A} = (A_1, \ldots, A_n) \in \mathbb{P}^n$. For convenience, we use the notation

$$\mathbb{A}_{\neq j} = (A_1, \ldots, A_{j-1}, A_{j+1}, \ldots, A_n) \in \mathbb{P}^{n-1}$$

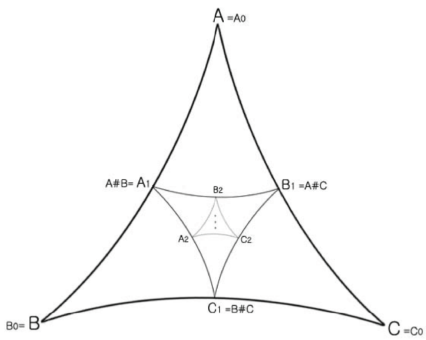
for any $j \in \{1, \ldots, n\}$.

Theorem 3.2. *Let* $\mathbb{A} = (A_1, A_2, \ldots, A_{n+1}) \in \mathbb{P}^{n+1}$*. Assume that the* n*-variable geometric mean* G *satisfying (G1) through (G4) exists. Consider the recursive sequences*

$$A_j^{(0)} = A_j,\ A_j^{(r)} = G(\mathbb{A}^{(r-1)}_{\neq n+2-j}) \tag{1}$$

for any $j \in \{1, \ldots, n+1\}$*, where* $\mathbb{A}^{(r-1)} = (A_1^{(r-1)}, A_2^{(r-1)}, \ldots, A_{n+1}^{(r-1)})$ *for* $r = 1, 2, \ldots$*. Then sequences* $A_j^{(r)}$ *converge as* $r \to \infty$ *and their limits equal to* $G(\mathbb{A})$*.*

For an $(n+1)$-tuple $\mathbb{A}$ of positive definite operators, the jth sequence in the first construction is made by the operator mean G of the n-tuple obtained by removing $(n+2-j)$th component of $\mathbb{A}$ for any $j \in \{1, \ldots, n+1\}$. This construction gives us a new $(n+1)$-tuple $\mathbb{A}^{(1)}$ of positive definite operators. We continue this process to get the operator means of $(n+1)$ variables from the same mean of n variables. The following shows how to construct three-variable geometric mean via symmetrization process.



Via Section 2 and Section 3 of [5] Lawson and Lim have introduced a way to extend higher-order means from nonexpansive and coordinatewise contractive means in a complete metric space X. See [5, Definition 3.7] for coordinatewise contractivity and [5, Definition 3.11] for nonexpansivity. Note that $\mathbb{P}$ is the complete metric space with respect to the Thompson metric, and the mean G satisfies the nonexpansive and contractive properties by Lemma 3.1. So the map

$$\beta : \mathbb{P}^{n+1} \to \mathbb{P}^{n+1},\ \beta(\mathbb{A}) = (G(\mathbb{A}_{\neq 1}), \ldots, G(\mathbb{A}_{\neq n+1}))$$

is power convergent, which means that

$$\lim_{k \to \infty} \beta^k(\mathbb{A}) = (M, \ldots, M)$$

for some $M \in \mathbb{P}$. Since the mean G is permutation invariant by (G3), our limit $G(\mathbb{A})$ in Theorem 3.2 is the equal to M (see Remark 2.2 in [5]). While Lawson and Lim have shown it using (locally) convex hull in the geometric sense (see Proposition 3.13 in [5]), we provide a different proof of power convergence in this article.

Proof. We follow two steps.

Step 1. Assume that $A_1 \le A_2 \le \cdots \le A_{n+1}$. Then by the monotonicity (G4)

$$A_1^{(1)} = G(A_1, \ldots, A_{n-1}, A_n) \le G(A_1, \ldots, A_{n-1}, A_{n+1}) = A_2^{(1)},$$

so similarly $A_1^{(1)} \le A_2^{(1)} \le \cdots \le A_{n+1}^{(1)}$. Therefore, inductively we have

$$A_1^{(r)} \le A_2^{(r)} \le \cdots \le A_{n+1}^{(r)}$$

for all $r \ge 1$.

Moreover, $\{A_1^{(r)}\}_{r=0}^{\infty}$ is increasing. Indeed, by the idempotency (G1) and the monotonicity (G4)

$$A_1^{(r-1)} = G(A_1^{(r-1)}, \ldots, A_1^{(r-1)}) \le G(A_1^{(r-1)}, \ldots, A_n^{(r-1)}) = A_1^{(r)}.$$

Similarly, we can prove that $\{A_{n+1}^{(r)}\}_{r=0}^{\infty}$ is decreasing.

So $A_1^{(r)}$ and $A_{n+1}^{(r)}$ converge as $r \to \infty$. Say $X = \lim_{r\to\infty} A_1^{(r)}$ and $Y = \lim_{r\to\infty} A_{n+1}^{(r)}$, respectively. Obviously, $X \le Y$.

By the permutation invariancy (G3) and Lemma 3.1 we have

$$\begin{aligned} &d(A_1^{(r)}, A_{n+1}^{(r)}) \\ &= d(G(A_1^{(r-1)}, \ldots, A_n^{(r-1)}), G(A_2^{(r-1)}, \ldots, A_{n+1}^{(r-1)})) \\ &= d(G(A_1^{(r-1)}, A_2^{(r-1)}, \ldots, A_n^{(r-1)}), G(A_{n+1}^{(r-1)}, A_2^{(r-1)}, \ldots, A_n^{(r-1)})) \\ &\le \frac{1}{n} d(A_1^{(r-1)}, A_{n+1}^{(r-1)}). \end{aligned}$$

Taking the limit as $r \to \infty$ we have $\left(1 - \frac{1}{n}\right) d(X,Y) \le 0$, or $d(X,Y) = 0$. So $X = Y$, and thus, it implies that

$$\lim_{r\to\infty} A_1^{(r)} = \cdots = \lim_{r\to\infty} A_{n+1}^{(r)}.$$

Step 2. For arbitrary positive definite operators $A_1, \ldots, A_{n+1}$, there exist positive constants $\alpha_1, \alpha_2, \ldots, \alpha_{n+1}$ such that

$$\alpha_1 A_1 \le \alpha_2 A_2 \le \cdots \le \alpha_{n+1} A_{n+1}.$$

Set $\hat{A}_1 = \alpha_1 A_1, \hat{A}_2 = \alpha_2 A_2, \ldots$, and $\hat{A}_{n+1} = \alpha_{n+1} A_{n+1}$. By Step 1, each sequence $\hat{A}_j^{(r)}$ converges as $r \to \infty$ and

$$\lim_{r\to\infty} \hat{A}_1^{(r)} = \cdots = \lim_{r\to\infty} \hat{A}_{n+1}^{(r)}.$$

On the other hand, for the positive numbers $\alpha_1, \ldots, \alpha_{n+1}$, it is easy to see that the recursion provides convergent sequences $\alpha_1^{(r)}, \ldots, \alpha_{n+1}^{(r)}$ with

$$\lim_{r\to\infty} \alpha_1^{(r)} = \cdots = \lim_{r\to\infty} \alpha_{n+1}^{(r)} = \sqrt[n+1]{\alpha_1 \cdots \alpha_{n+1}}.$$

Since $\hat{A}_j^{(r)} = \alpha_j^{(r)} A_j^{(r)}$ for all j and all $r \geq 1$, we have that $A_j^{(r)}$ converges as $r \to \infty$ and

$$\lim_{r\to\infty} A_j^{(r)} = \frac{1}{\sqrt[n+1]{\alpha_1 \cdots \alpha_{n+1}}} \lim_{r\to\infty} \hat{A}_j^{(r)}.$$

By the equation (3) we conclude all limits of $A_j^{(r)}$ are equal.

□

Proposition 3.3. *The* $(n+1)$*-variable geometric mean* G *obtained by the recursive sequence* (1) *also satisfies the following. Let* $\mathbb{A} = (A_1, \dots, A_{n+1})$ *and* $\mathbb{B} = (B_1, \dots, B_{n+1}) \in \mathbb{P}^{n+1}$.

(1) *(Idempotency) For any* $A \in \mathbb{P}$

$$G(A, \dots, A) = A.$$

(2) *(Joint homogeneity) For all* $a_j > 0$

$$G(a_1 A_1, \dots, a_{n+1} A_{n+1}) = \left[\prod_{j=1}^{n+1} a_j\right]^{\frac{1}{n+1}} G(A_1, \dots, A_{n+1}).$$

(3) *(Permutation invariancy) For any permutation* σ *on* $\{1, \dots, n+1\}$

$$G(\mathbb{A}) = G(\mathbb{A}_\sigma).$$

(4) *(Monotonicity) If* $A_j \leq B_j$ *for all* $j = 1, \dots, n+1$,

$$G(\mathbb{A}) \leq G(\mathbb{B}).$$

Proof. These properties can be easily seen from the proof of Theorem 3.2. □

The formula of $G(A_1, \dots, A_{n+1})$ is rather complicated, but it is nice with only two variables. We give an interesting property for $G(A_1, \dots, A_{n+1})$ constructed by only two variables.

Corollary 3.4. *Assume that* $A_1 = \cdots = A_k = A$ *and* $A_{k+1} = \cdots = A_n = B$ *for some* $1 < k < n+1$. *Then*

$$G(A_1, \dots, A_{n+1}) = A^{1/2}(A^{-1/2} B A^{-1/2})^t A^{1/2},$$

where $t = \dfrac{n-k}{n+1}$.

4. Numerical experiments

We give an algorithm to find three-variable geometric mean constructed by Theorem 3.2. We consider positive definite matrices instead of positive definite operators to be able to compute and show two examples using *MATLAB*.

Algorithm

Require: Points A_0, B_0 and C_0 which are positive definite
$\epsilon > 0$ which is small enough
$m[X, Y]$ which is a two-variable mean to be considered
$i = 0$

Step 1: If $\max\{\|A_i - B_i\|, \|B_i - C_i\|, \|C_i - A_i\|\} \geq \epsilon$, then compute
$A_{i+1} = m[A_i, B_i]$, $B_{i+1} = m[A_i, C_i]$, and $C_{i+1} = m[B_i, C_i]$;
Set $i = i + 1$
Continue **Step 1**

Step 2: If $\max\{\|A_i - B_i\|, \|B_i - C_i\|, \|C_i - A_i\|\} < \epsilon$,
then STOP

In this article we are interested in the geometric mean so that we set

$$m[X, Y] = X^{1/2}(X^{-1/2}YX^{-1/2})^{1/2}X^{1/2}.$$

Example 4.1. It has been known that any 2×2 density matrix $\rho_{\mathbf{v}}$, which is a 2-by-2 positive semidefinite Hermitian matrix with trace 1, can be parameterized by a Bloch vector $\mathbf{v}$ in the unit ball of $\mathbb{R}^3$. Here,

$$\rho_{\mathbf{v}} = \frac{1}{2}\begin{pmatrix} 1 + v_3 & v_1 - iv_2 \\ v_1 + iv_2 & 1 - v_3 \end{pmatrix}, \quad \mathbf{v} = \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} \in \mathbb{R}^3.$$

Moreover, the 2×2 invertible density matrix described by a Bloch vector in the open unit ball $\mathbf{B}$ of $\mathbb{R}^3$ plays an important role in quantum information theory. We give an example of three-variable geometric mean of 2×2 invertible density matrices.

Let $\rho_{\mathbf{u}}$, $\rho_{\mathbf{v}}$, and $\rho_{\mathbf{w}}$ be 2×2 density matrices parameterized by

$$\mathbf{u} = \begin{pmatrix} 3/13 \\ 4/13 \\ 0 \end{pmatrix}, \quad \mathbf{v} = \begin{pmatrix} 3/13 \\ 0 \\ 4/13 \end{pmatrix}, \quad \mathbf{w} = \begin{pmatrix} 0 \\ 3/13 \\ 4/13 \end{pmatrix}.$$

Set $\epsilon = 10^{-4}$. Then by the 12th iteration we obtain

$$G(\rho_{\mathbf{u}}, \rho_{\mathbf{v}}, \rho_{\mathbf{w}}) \approx \begin{pmatrix} 0.5857 & 0.0746 - 0.0871i \\ 0.0746 + 0.0871i & 0.3861 \end{pmatrix}.$$

On the other hand, A. Ungar has shown in [7, Theorem 6.93] that the gyrocentroid of Bloch vectors $\mathbf{u}$, $\mathbf{v}$, and $\mathbf{w}$ in $\mathbf{B}$ are given by

$$C(\mathbf{u}, \mathbf{v}, \mathbf{w}) = \frac{\gamma_{\mathbf{u}}\mathbf{u} + \gamma_{\mathbf{v}}\mathbf{v} + \gamma_{\mathbf{w}}\mathbf{w}}{\gamma_{\mathbf{u}} + \gamma_{\mathbf{v}} + \gamma_{\mathbf{w}}},$$

where $\gamma_{\mathbf{v}} = \dfrac{1}{\sqrt{1 - \|\mathbf{v}\|^2}}$ is the Lorentz factor. In this example we have $C(\mathbf{u}, \mathbf{v}, \mathbf{w}) = \dfrac{\mathbf{u} + \mathbf{v} + \mathbf{w}}{3}$ since $\|\mathbf{u}\| = \|\mathbf{v}\| = \|\mathbf{w}\|$, and so

$$\rho_C \approx \begin{pmatrix} 0.6026 & 0.0769 - 0.0897i \\ 0.0769 + 0.0897i & 0.3974 \end{pmatrix}.$$

We easily check that

$$\frac{1}{\mathrm{tr}G(\rho_{\mathbf{u}},\rho_{\mathbf{v}},\rho_{\mathbf{w}})}G(\rho_{\mathbf{u}},\rho_{\mathbf{v}},\rho_{\mathbf{w}}) \approx \begin{pmatrix} 0.6027 & 0.0768-0.0897i \\ 0.0768+0.0897i & 0.3973 \end{pmatrix} \neq \rho_C.$$

Example 4.2. Let $A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, B = \begin{pmatrix} 1 & 2 \\ 2 & 5 \end{pmatrix}, C = \begin{pmatrix} 1 & 3 \\ 3 & 10 \end{pmatrix}$. These are positive definite matrices whose determinants are all 1. In this case we can use the following formula of two-variable geometric mean for 2×2 positive definite matrices A and B whose determinants are 1 (see [1, Proposition 4.1.12]):

$$A\#B = \frac{A+B}{\sqrt{\det(A+B)}}.$$

This may reduce computing time because the geometric mean is calculated by matrix sum instead of matrix power and multiplication. Set $\epsilon = 10^{-5}$. Then by the 20th iteration we obtain

$$G(A,B,C) \approx \begin{pmatrix} 0.63770 & 1.10936 \\ 1.10936 & 3.49801 \end{pmatrix},$$

and we can easily verify that $G(A,B,C)$ also has determinant 1.

References

1. R. Bhatia, *Positive Definite Matrices*, Princeton University Press, NJ, 2007.
2. S. K. Chakraborty and G. Panda, *A higher order iterative algorithm for multivariate optimization problem*, J. Appl. Math. & Informatics **32** (2014), 747–760.
3. F. Hiai and D. Petz, *Introduction to Matrix Analysis and Applications*, Hindustan Book Agancy and Springer, 2014.
4. F. Kubo and T. Ando, *Means of positive linear operators*, Math. Ann. **246** (1980), 205–224.
5. J. Lawson and Y. Lim, *A general framework for extending means to higher orders*, Colloq. Math. **113** (2008), 191–221.
6. J. Lawson and Y. Lim, *Karhcer means and Karcher equations of positive definite operators*, Trans. Amer. Math. Soc. Series B **1** (2014), 1–22.
7. A. Ungar, *Analytic Hyperbolic Geometry and Albert Einstein's Special Theory of Relativity*, World Scientific, 2008.

Sejong Kim received Master of Sciences from Kyungpook National University and Ph.D at Louisiana State University. Since 2013 he has been at Chungbuk National University. His research interests include matrix geometry, operator mean, and quantum information theory.

Department of Mathematics, Chungbuk National University, Cheongju 362-763, Korea.
`e-mail: skim@chungbuk.ac.kr`

Dénes Petz received Master of Sciences and Ph.D. from Eötvös University of Budapest. He is currently a professor at Budapest University of Technology and Economics. His research interests are operator algebras, quantum probability, and quantum information theory.

Department of Mathematics, Budapest University of Technology and Economics, Budapest, Hungary.
`e-mail: petz@math.bme.hu`