Geometric Means

Chi-Kwong Li
Department of Mathematics
College of William and Mary
Williamsburg, VA 23185

Joint work with

T. Ando (Hokkaido University)
and R. Mathias (William and Mary)

1. Motivations and basic requirements

Motivated by the study of

operator inequalities, electrical network theory, etc.,

researchers have defined and studied the geometric means of positive (semi-)definite matrices.

If A and B are diagonal matrices

$$\operatorname{diag}(a_1, \ldots, a_n) \quad \text{and} \quad \operatorname{diag}(b_1, \ldots, b_n),$$

then their geometric mean can be naturally defined as

$$(AB)^{1/2} = \operatorname{diag}\left(\sqrt{a_1 b_1}, \ldots, \sqrt{a_n b_n}\right).$$

In general, there exists an invertible S such that $A = S^* D_A S$ and $B = S^* D_B S$. One can then define the geometric mean of A and B as

$$S^*(D_A D_B)^{1/2} S = A^{1/2}(A^{-1/2} B A^{-1/2})^{1/2} A^{1/2},$$

which can also be defined as

$$\max \left\{ X \geq 0 : \begin{pmatrix} A & X \\ X & B \end{pmatrix} \geq 0 \right\}.$$

Question How to define the geometric mean of 3 or more positive (semi-)definite matrices?

Reasonable requirements for a geometric mean $G(A,B,C)$ of 3 positive definite matrices A,B,C:

P1 Consistency with scalars.

If $A,\ B,\ C$ commute then $G(A,B,C) = (ABC)^{1/3}$.

In particular, $G(A,A,A) = A$.

P2 Joint Homogeneity.

$G(\alpha A, \beta B, \gamma C) = (\alpha\beta\gamma)^{1/3} G(A,B,C)$ for any $\alpha,\beta,\gamma > 0$.

So, $G(\alpha A, \alpha B, \alpha C) = \alpha G(A,B,C)$ for any $\alpha > 0$.

P3 Permutation invariance.

For any permutation $\pi(A,B,C)$ of (A,B,C), we have

$$G(A,B,C) \;=\; G\Big(\pi(A,B,C)\Big).$$

P4 Monotonicity.

If $A \geq A_0$, $B \geq B_0$, and $C \geq C_0$, then

$$G(A, B, C) \geq G(A_0, B_0, C_0).$$

P5 Continuity from above.

If $\{A_n\}, \{B_n\}, \{C_n\}$ are monotonic decreasing sequences converging to A, B, C, then $\{G(A_n, B_n, C_n)\}$ converges to $G(A, B, C)$.

P6 Congruence invariance.

If S is invertible, then

$$G(S^*AS, S^*BS, S^*CS) = S^*G(A, B, C)S.$$

P7 Joint concavity.

For any $\lambda \in (0,1)$,

$$G\Big(\lambda(A_1, A_2, A_3) + (1-\lambda)(B_1, B_2, B_3)\Big)$$
$$\geq \lambda G(A_1, A_2, A_3) + (1-\lambda) G(B_1, B_2, B_3).$$

P8 Self-duality.

$$G(A, B, C) = G(A^{-1}, B^{-1}, C^{-1})^{-1}.$$

P9 Determinant identity.

$$\det G(A, B, C) = \Big(\det A \cdot \det B \cdot \det C\Big)^{1/3}.$$

By P1, P3, P7, and P8, we have

P10 The arithmetic-geometric-harmonic mean inequality.

$$\frac{A+B+C}{3} \ \geq \ G(A, B, C) \geq \left(\frac{A^{-1}+B^{-1}+C^{-1}}{3}\right)^{-1}.$$

Remarks

* Any geometric mean should satisfy properties P1–P6 at a bare minimum.

  [Actually, P2 and P4 imply P5.]

* With P1–P6, one can uniquely extend the definition to positive semidefinite matrices by setting

$$G(A, B, C) = \lim_{\epsilon \downarrow 0} G(A + \epsilon I, B + \epsilon I, C + \epsilon I).$$

  This definition satisfies P1-P7, and the following stronger form of P6.

P6′ $\quad G(S^* AS, S^* BS, S^* CS) \geq S^* G(A, B, C)S \quad$ for all S.

2. Geometric means of two or more matrices

Let $G(A_1, A_2)$ be the usual geometric mean.

Suppose $G(X_1, \ldots, X_{k-1})$ is defined for $k-1$ positive definite matrices $X_1, \ldots, X_{k-1}$.

For $A = (A_1, \ldots, A_k)$ define

$$T(A) = (G(A_2, \ldots, A_k), \ldots, G(A_1, \ldots, A_{k-1}))\,.$$

Theorem 1 Let $A_1, \ldots A_k$ be given. The limit of this sequence $\{T^m(A_1, \ldots, A_k)\}_{m=1}^{\infty}$ exists and has the form $(\tilde{A}, \ldots, \tilde{A})$. If we define $G(A_1, \ldots, A_{k+1})$ to be $\tilde{A}$, then it satisfies P1–P10.

Our proof uses the following multiplicative metric on the space of pairs of positive definite matrices:

$$R(A, B) = \max\{\rho(A^{-1}B), \rho(B^{-1}A)\};$$

The metric $R(\cdot, \cdot)$ has many nice properties, for example,

$$R(A, C) \leq R(A, B)R(B, C).$$

The properties we need in our proof are

$$R(A, B) \geq 1, \quad R(A, B) = 1 \Leftrightarrow A = B,$$

$$\text{and} \quad R(A, B)^{-1}A \leq B \leq R(A, B)A,$$

which implies the norm bound

$$\|A - B\| \leq (R(A, B) - 1)\|A\|.$$

Theorem 2 The geometric means defined in Theorem 1 satisfy

$$R\left(G(A_1,\dots,A_k),G(B_1,\dots,B_k)\right) \le \left\{\prod_{i=1}^{k} R(A_i,B_i)\right\}^{1/k}$$

for $k = 2, 3, \dots$

Extend the definition of the geometric mean to the case of positive semidefinite matrices. Then the extended geometric mean satisfies P1 –P7 as well as P6$'$. Moreover, we have

Theorem 3 For any positive semidefinite matrices $A_1, \dots, A_k$

$$\mathrm{range}(G(A_1,\dots,A_k)) = \bigcap_{i=1}^{k} \mathrm{range}(A_i), \quad k = 2, 3, \dots.$$

3. Some consequences

Theorem 4 Let $\phi : P_n \to P_m$ be monotone, continuous from above, and continuous in the interior of P_n. Suppose

$$G(\phi(X), \phi(Y)) - \phi(G(X, Y))$$

is positive semidefinite (respectively, negative semidefinite or zero) for any $X, Y \in P_n$. Then so is

$$G(\phi(A_1), \ldots, \phi(A_k)) - \phi(G(A_1, \ldots, A_k))$$

any $k \geq 2$.

Special cases

* If ϕ is positive linear such that $\phi(I)$ is positive definite, then

$$G(\phi(A_1), \ldots, \phi(A_k)) \geq \phi(G(A_1, \ldots, A_k)).$$

* If we take $\phi(X) = \prod_{i=1}^{r} \lambda_i(X)$, where λ_i denotes the ith largest eigenvalue, then for any $p \times p$ positive definite matrices $A_1, \ldots, A_k$ and any $1 \leq r \leq p$

$$\prod_{i=1}^{r} \lambda_i(G(A_1, \ldots, A_k)) \leq \prod_{i=1}^{r} \left(\prod_{l=1}^{k} \lambda_i(A_l) \right)^{1/k}$$

$$\text{and } \prod_{i=r}^{p} \lambda_i(G(A_1, \ldots, A_k)) \geq \prod_{i=r}^{p} \left(\prod_{l=1}^{k} \lambda_i(A_l) \right)^{1/k}.$$

* The Schur complement of a positive semidefinite matrix

$$X = \begin{pmatrix} X_{11} & X_{12} \\ X_{12}^* & X_{22} \end{pmatrix}$$

is defined by

$$\phi(X) = \begin{pmatrix} X_{11} - X_{12} X_{22}^{\dagger} X_{12}^* & 0 \\ 0 & 0 \end{pmatrix} \equiv \begin{pmatrix} S_X & 0 \\ 0 & 0 \end{pmatrix},$$

and can be characterized by

$$S_A = \max \left\{ X : A \geq \begin{pmatrix} X & 0 \\ 0 & 0 \end{pmatrix} \right\}.$$

Since $S_{G(A,B)} \geq G(S_A, S_B)$, we have

$$S_{G(A_1,\ldots,A_k)} \geq G(S_{A_1}, \ldots, S_{A_k})$$

for any positive semidefinite $A_1, \ldots, A_k$.

* Let $\phi(A) = C_q(A)$, be the q-th multiplicative compound of an $n \times n$ matrix A $(1 \leq q \leq n)$. Then

$$C_q(G(A_1, A_2)) = G(C_q(A_1), C_q(A_2)).$$

Hence, for any positive semidefinite matrices $A_1, \ldots, A_k$,

$$C_q(G(A_1, \ldots, A_k)) = G(C_q(A_1), \ldots, C_q(A_k)).$$

* The qth additive compound Δ_q is defined by

$$\Delta_q(A) = \lim_{t \to 0} \frac{C_q(A + tI) - C_q(A)}{t}.$$

Then

$$G(\Delta_q(A), \Delta_q(B)) \geq \Delta_q(G(A, B)).$$

Thus, for any positive semidefinite matrices $A_1, \ldots, A_k$,

$$G(\Delta_q(A_1), \ldots, \Delta_q(A_k)) \geq \Delta_q(G(A_1, \ldots, A_k)).$$

4. Functional Characterization

Our geometric mean satisfies the following functional characterization.

Theorem 5 The function G in Theorem 1 is the only family of functions $f_k : P_n^k \to P_n$, $k = 2, 3, \ldots$ that satisfies

1. $f_2(A, B) = G(A, B)$.
2. f_k maps any k-tuples of positive semidefinite matrices to a positive semidefinite matrix and it is monotone and continuous from above.
3. f_k maps any k-tuples of positive definite matrices to a positive definite matrix and it is continuous.
4. For $k \geq 3$,

$$f_k((A_i)_{i=1}^k) = f_k(f_{k-1}((A_i)_{i\neq 1}), \ldots, f_{k-1}((A_i)_{i\neq k})).$$

Remarks

* The fourth condition, which certainly is desirable, does not seem to be essential for a geometric mean.
* The properties P1-P9 are not sufficient to characterize the geometric mean of more than two matrices.

Questions

* Determine a formula for $G(A_1, \ldots, A_k)$ without iterations.
* Extend the results to infinite dimensional operators.

THANK YOU FOR YOUR ATTENTION!