Which matrix geometric mean do you mean?

Frank Nielsen

2023

Abstract

We explain how to generalize the geometric mean between two positive reals to symmetric positive-definite matrices by considering different properties of the geometric mean. Namely, we present the viewpoint of Riemannian centroids, inductive means, and limits of power means.

The geometric mean between two positive reals x and y is defined by

$$G(x,y) = \sqrt{xy}.$$

To extend the scalar geometric mean to symmetric positive-definite matrices X and Y, we may generalize different properties of the scalar geometric mean as follows:

Limit of quasi-arithmetic power means. First, let us consider the scalar geometric mean as the limit of power means $M_p(x,y) = (x^p + y^p)^{\frac{1}{p}}$ when $p \to 0$:

$$G(x,y) = \lim_{p\to 0} M_p(x,y).$$

A quasi-arithmetic mean induced by strictly increasing and differentiable real-valued functional generators $f(u)$ is defined by

$$M_f(x,y) = f^{-1}\left(\frac{f(x)+f(y)}{2}\right).$$

Quasi-arithmetic means are also called Kolmogorov-Nagumo means [8, 10]. Since power means $M_p(x,y) = M_{f_p}(x,y)$ are quasi-arithmetic means for the smooth family of generators

$$f_p(u) = \begin{cases} \frac{u^p-1}{p}, & p \in \mathbb{R}\backslash\{0\}, \\ \log(u), & p = 0. \end{cases}, \quad f_p^{-1}(u) = \begin{cases} (1+up)^{\frac{1}{p}}, & p \in \mathbb{R}\backslash\{0\}, \\ \exp(u), & p = 0. \end{cases},$$

with $G(x,y) = M_0(x,y)$, we get by extension the following matrix geometric mean by taking the corresponding matrix quasi-arithmetic mean:

$$\mathrm{LE}(X,Y) = M_{\log}(X,Y) = \exp\left(\frac{\log X + \log Y}{2}\right),$$

where $\exp(X)$ and $\log(X)$ denote the matrix exponential and the matrix logarithm of X, respectively. This matrix mean is called the **log-Euclidean matrix mean** [2]. One drawback of this matrix geometric mean (MGM) is that it is not operator monotone where a matrix mean $M(X,Y)$ is said operator monotone [3] if for $X' \preceq X$ and $Y' \preceq Y$, we have $M(X',Y') \preceq M(X,Y)$ where $\preceq$ denotes Loewner partial ordering on the cone $\mathbb{P}$: $P \preceq Q$ if and only if $Q - P$ is positive semi-definite. The scalar geometric mean satisfies the operator monotone property: if $x' \leq x$ and $y' \leq y$ then we have $G(x',y') \leq G(x,y)$. Ando-Li-Mathias [1] defined a set of 10 properties that a matrix geometric mean should satisfy, including the operator monotone property.

Riemannian centroid. Second, the scalar geometric mean can be interpreted as the unique centroid with respect to the distance $\rho(x,y) = \left\|\log \frac{x}{y}\right|$:

$$G(x,y) = \arg\min_{c\in\mathbb{R}_{>0}} \frac{1}{2}\rho^2(x,c) + \frac{1}{2}\rho^2(c,y).$$

Let $\mathbb{P}$ denote the set of symmetric positive-definite $d\times d$ matrices. Consider the Riemannian manifold $(\mathbb{P}, g)$ where g is the so-called trace metric, i.e., a collection of smoothly varying inner products g_P for $P\in\mathbb{P}$ defined by

$$g_P(S_1,S_2) = \mathrm{tr}\left(P^{-1}S_1P^{-1}S_2\right),$$

where S_1 and S_2 are matrices belonging to the vector space of symmetric $d\times d$ matrices. S_1 and S_2 are geometrically interpreted as vectors of the tangent plane T_P of $P\in\mathbb{P}$. The Riemannian geodesic distance is

$$\rho(P_1,P_2) = \left\|\log\left(P_1^{-\frac{1}{2}}\,P_2\,P_1^{-\frac{1}{2}}\right)\right\|_F = \sqrt{\sum_{i=1}^d \log^2\lambda_i\left(P_1^{-\frac{1}{2}}\,P_2\,P_1^{-\frac{1}{2}}\right)},$$

where $\lambda_i(M)$ denotes the i-th largest real eigenvalue of a symmetric matrix M, $\|\cdot\|_F$ denotes the Frobenius norm, and $\log P$ is the unique matrix logarithm of a SPD matrix P. It follows that the Riemannian matrix geometric mean is

$$\mathrm{RG}(X,Y) = X^{\frac{1}{2}}\,(X^{-\frac{1}{2}}\,Y\,X^{-\frac{1}{2}})^{\frac{1}{2}}\,X^{\frac{1}{2}}.$$

This mean is proven to be the unique solution to the matrix Ricatti equation $CX^{-1}C = Y$, is invariant under inversion (i.e., $\mathrm{RG}(X,Y) = G(X^{-1},Y^{-1})^{-1}$), and satisfies the determinant property $\det(\mathrm{RG}(X,Y)) = \sqrt{\det(X)\det(Y)}$. Furthermore, the matrix mean RG = AHM is operator monotone [4].

Inductive means. Third, the scalar geometric scalar mean can be defined as limits of sequences of iterations:

$$\begin{aligned} a_{t+1} &= A(a_t,g_t) := \frac{a_t+g_t}{2},\\ h_{t+1} &= H(a_t,g_t) = \frac{2a_tg_t}{a_t+g_t}, \end{aligned}$$

initialized with $a_0 = x > 0$ and $g_0 = y > 0$. We have

$$\mathrm{AHM}(x,y) = \lim_{t\to\infty} a_t = \lim_{t\to\infty} h_t = \sqrt{xy} = G(x,y).$$

The AHM iterations enjoys fast quadratic convergence [7]. These kinds of means defined as limits of sequences have been termed **inductive means** [13].

Similarly, we may define the inductive matrix arithmetic-harmonic mean by the following sequence:

$$\begin{aligned} A_{t+1} &= \frac{A_t+H_t}{2} = A(A_t,H_t),\\ H_{t+1} &= 2\,(A_t^{-1}+H_t^{-1})^{-1} = H(A_t,H_t), \end{aligned}$$

where the matrix arithmetic mean is $A(X,Y) = \frac{X+Y}{2}$ and the matrix harmonic mean is $H(X,Y) = 2(X^{-1}+Y^{-1})^{-1}$. The AHM iterations initialized with $A_0 = X$ and $H_0 = Y$ yield in the limit $t\to\infty$, the matrix arithmetic-harmonic mean [11] (AHM):

$$\mathrm{AHM}(X,Y) = \lim_{t\to+\infty} A_t = \lim_{t\to+\infty} H_t = \mathrm{RG}(X,Y).$$

The matrix AHM iterations enjoys quadratic convergence.

Limits of power mean functional equation Fourth, we can define the scalar power means $M_p(x,y)$ as the solutions of the equation [9]:

$$m = \frac{1}{2}m^{1-p}x^p + \frac{1}{2}m^{1-p}y^p$$

which is solved as $m = \left(\frac{1}{2}x^p + \frac{1}{2}y^p\right)^{\frac{1}{p}} = M_p(x,y)$. Similarly, we can define the matrix power means $M_p(X,Y)$ for $p \in (0,1]$ by uniquely solving the matrix equation [9]:

$$M = \frac{1}{2}M\#_p X + \frac{1}{2}M\#_p Y, \tag{1}$$

where

$$X\#_t Y = X^{\frac{1}{2}}\left(X^{-\frac{1}{2}}\,Y\,X^{-\frac{1}{2}}\right)^t X^{\frac{1}{2}},$$

is the Riemannian barycenter minimizing

$$(1-t)\rho^2(X,P) + t\rho^2(Y,P).$$

In the limit case $p \to 0$, this matrix power mean M_p yields the Riemannian matrix geometric mean [9]:

$$\lim_{p\to 0^+} M_p(X,Y) = \mathrm{RG}(X,Y).$$

We have described several properties of the scalar geometric means which when extended to matrices yield different ways to define matrix geometric means. Ando-Li-Mathias [1] (ALM) listed 10 properties that such matrix geometric means should satisfy. There are infinitely many ways to construct matrix geometric means satisfying those ALM properties [5]. The Riemannian matrix geometric mean is often used as it can be obtained from many different generalizations (e.g., Riemannian centroid [4], arithmetic-harmonic mean [11], limit of power means [9]).

There are also several extensions of geometric means to n-parameter geometric means. In particular, a recursive matrix mean which converges with cubic order is given in [5].

This note is based on a paper entitled "What is... an inductive mean?" [12].

So which matrix geometric mean do you mean[1]?

References

[1] Tsuyoshi Ando, Chi-Kwong Li, and Roy Mathias. Geometric means. *Linear algebra and its applications*, 385:305–334, 2004.

[2] Vincent Arsigny, Pierre Fillard, Xavier Pennec, and Nicholas Ayache. Geometric means in a novel vector space structure on symmetric positive-definite matrices. *SIAM journal on matrix analysis and applications*, 29(1):328–347, 2007.

[3] Rajendra Bhatia and John Holbrook. Riemannian geometry and matrix geometric means. *Linear algebra and its applications*, 413(2-3):594–618, 2006.

[4] Rajendra Bhatia and Rajeeva L Karandikar. Monotonicity of the matrix geometric mean. *Mathematische Annalen*, 353:1453–1467, 2012.

[5] Dario Bini, Beatrice Meini, and Federico Poloni. An effective matrix geometric mean satisfying the Ando–Li–Mathias properties. *Mathematics of Computation*, 79(269):437–452, 2010.

[6] Miguel de Carvalho. Mean, what do you mean? *The American Statistician*, 70(3):270–274, 2016.

[1] Closing sentence by analogy to the paper entitled "Mean, what do you Mean?" [6]

[7] D. M. E. Foster and G. M. Phillips. A generalization of the Archimedean double sequence. *Journal of mathematical analysis and applications*, 101(2):575–581, 1984.

[8] A. N. Kolmogorov. Sur la notion de la moyenne. *Atti della Accademia nazionale dei Lincei Rendicoti*, 12:388–391, 1930.

[9] Yongdo Lim and Miklós Pálfia. Matrix power means and the Karcher mean. *Journal of Functional Analysis*, 262(4):1498–1514, 2012.

[10] Mitio Nagumo. Über eine klasse der mittelwerte. In *Japanese journal of mathematics: transactions and abstracts*, volume 7, pages 71–79. The Mathematical Society of Japan, 1930.

[11] Yoshimasa Nakamura. Algorithms associated with arithmetic, geometric and harmonic means and integrable systems. *Journal of computational and applied mathematics*, 131(1-2):161–174, 2001.

[12] Frank Nielsen. What is an inductive mean? *Notices of the AMS*, 70(11), 2023.

[13] Karl-Theodor Sturm. Probability measures on metric spaces of nonpositive curvature. *Heat Kernels and Analysis on Manifolds, Graphs, and Metric Spaces: Lecture Notes from a Quarter Program on Heat Kernels, Random Walks, and Analysis on Manifolds and Graphs: April 16-July 13, 2002, Emile Borel Centre of the Henri Poincaré Institute, Paris, France*, 338:357, 2003.