WHAT IS...?

an Inductive Mean?

Frank Nielsen

Notions of means. The notion of means [10] is central to mathematics and statistics, and plays a key role in machine learning and data analytics. The three classical Pythagorean means of two positive reals x and y are the arithmetic (A), geometric (G), and harmonic (H) means, given respectively by

$$A(x,y)=\frac{x+y}{2},\quad G(x,y)=\sqrt{xy},\quad H(x,y)=\frac{2xy}{x+y}.$$

These Pythagorean means were originally geometrically studied to define proportions, and the harmonic mean led to a beautiful connection between mathematics and music. The Pythagorean means enjoy the following inequalities:

$$\min(x,y)\leq H(x,y)\leq G(x,y)\leq A(x,y)\leq \max(x,y),$$

with equality if and only if $x=y$. These Pythagorean means belong to a broader parametric family of means, the power means $M_p(x,y)=(x^p+y^p)^{\frac{1}{p}}$ defined for $p\in\mathbb{R}\backslash\{0\}$. We have $A(x,y)=M_1(x,y)$, $H(x,y)=M_{-1}(x,y)$ and in the limits: $G(x,y)=\lim_{p\to 0} M_p(x,y)$, $\max(x,y)=\lim_{p\to+\infty} M_p(x,y)$, and $\min(x,y)=\lim_{p\to-\infty} M_p(x,y)$. Power means are also called binomial, Minkowski, or Hölder means in the literature.

There are many ways to define and axiomatize means with a rich literature [8]. An important class of means are the quasi-arithmetic means induced by strictly increasing and differentiable real-valued functional generators $f(u)$:

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

Quasi-arithmetic means satisfy the in-betweenness property of means: $\min(x,y)\leq M_f(x,y)\leq \max(x,y)$, and are called so because $f(M_f(x,y))=\frac{f(x)+f(y)}{2}=A(f(x),f(y))$ is the arithmetic mean on the f-representation of numbers.

The power means are quasi-arithmetic means, $M_p=M_{f_p}$, obtained for the following continuous family of generators:

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

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

Power means are the only homogeneous quasi-arithmetic means, where a mean $M(x,y)$ is said to be homogeneous when $M(\lambda x,\lambda y)=\lambda M(x,y)$ for any $\lambda>0$.

Quasi-arithmetic means can also be defined for n-variable means (i.e., $M_f(x_1,\ldots,x_n)=f^{-1}(\frac{1}{n}\sum_{i=1}^n f(x_i))$), and more generally for calculating expected values of random variables [10]: We denote by $\mathbb{E}_f[X]=f^{-1}(\mathbb{E}[f(X)])$ the quasi-arithmetic expected value of a random variable X induced by a strictly monotone and differentiable function $f(u)$. For example, the geometric and harmonic expected values of X are defined by $\mathbb{E}^G[X]=\mathbb{E}_{\log x}[X]=\exp(\mathbb{E}[\log X])$ and $\mathbb{E}^H[X]=\mathbb{E}_{x^{-1}}[X]=\frac{1}{\mathbb{E}[1/X]}$, respectively. The ordinary expectation is recovered for $f(u)=u$: $\mathbb{E}^A[X]=\mathbb{E}_x[X]=\mathbb{E}[X]$. The quasi-arithmetic expected values satisfy a strong law of large numbers and a central limit theorem ([10], Theorem 1): Let $X_1,\ldots,X_n$ be independent and identically distributed (i.i.d.) with finite variance $\mathbb{V}[f(X)]<\infty$ and derivative $f'(\mathbb{E}_f[X])\neq 0$ at $x=\mathbb{E}_f[X]$.

Frank Nielsen is a fellow at Sony Computer Science Laboratories Inc., Tokyo, Japan. His email address is frank.nielsen.x@gmail.com.

Additional material is available online at https://franknielsen.github.io/MatrixGeometricMean/.

Communicated by Notices *Associate Editor Richard Levine.*

DOI: https://doi.org/10.1090/noti2832

Then we have

$$M_f(X_1,\dots,X_n) \xrightarrow{a.s.} \mathbb{E}_f[X]$$
$$\sqrt{n}\left(M_f(X_1,\dots,X_n) - \mathbb{E}_f[X]\right) \xrightarrow{d} N\left(0, \frac{\mathbb{V}[f(X)]}{\left(f'(\mathbb{E}_f[X])\right)^2}\right)$$

as $n \to \infty$, where $N(\mu, \sigma^2)$ denotes a normal distribution of expectation μ and variance σ^2.

Inductive means. An inductive mean is a mean defined as a limit of a convergence sequence of other means [15]. The notion of inductive means defined as limits of sequences was pioneered independently by Lagrange and Gauss [7] who studied the following double sequence of iterations:

$$a_{t+1} = A(a_t, g_t) = \frac{a_t + g_t}{2},$$
$$g_{t+1} = G(a_t, g_t) = \sqrt{a_t g_t},$$

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

$$g_0 \leq \dots \leq g_t \leq \mathrm{AGM}(x, y) \leq a_t \leq \dots \leq a_0,$$

where the homogeneous arithmetic-geometric mean (AGM) is obtained in the limit:

$$\mathrm{AGM}(x, y) = \lim_{t\to\infty} a_t = \lim_{t\to\infty} g_t.$$

There is no closed-form formula for the AGM in terms of elementary functions as this induced mean is related to the complete elliptic integral of the first kind $K(\cdot)$ [7]:

$$\mathrm{AGM}(x, y) = \frac{\pi}{4} \frac{x + y}{K\left(\frac{x-y}{x+y}\right)},$$

where $K(u) = \int_0^{\frac{\pi}{2}} \frac{\mathrm{d}\theta}{\sqrt{1-u^2\sin^2(\theta)}}$ is the elliptic integral. The fast quadratic convergence [11] of the AGM iterations makes it computationally attractive, and the AGM iterations have been used to numerically calculate digits of π or approximate the perimeters of ellipses among others [7].

Some inductive means admit closed-form formulas: For example, the arithmetic-harmonic mean $\mathrm{AHM}(x, y)$ obtained as the limit of the double sequence

$$a_{t+1} = A(a_t, h_t) = \frac{a_t + g_t}{2},$$
$$h_{t+1} = H(a_t, h_t) = \frac{2a_t h_t}{a_t + h_t},$$

initialized with $a_0 = x > 0$ and $h_0 = y > 0$ converges to the geometric mean:

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

In general, inductive means defined as the limits of double sequences with respect to two smooth symmetric means M_1 and M_2:

$$a_{t+1} = M_1(a_t, b_t),$$
$$b_{t+1} = M_2(a_t, b_t),$$

are proven to converge quadratically [11] to $\mathrm{DS}_{M_1,M_2}(a_0, b_0) = \lim_{t\to\infty} a_t = \lim_{t\to\infty} b_t$ (order-2 convergence).

Inductive means and matrix means. We have obtained so far three ways to get the geometric scalar mean $G(x, y) = \sqrt{xy}$ between positive reals x and y:

1. As an inductive mean with the arithmetic-harmonic double sequence: $G(x, y) = \mathrm{AHM}(x, y)$,
2. As a quasi-arithmetic mean obtained for the generator $f(u) = \log u$: $G(x, y) = M_{\log}(x, y)$, and
3. As the limit of power means: $G(x, y) = \lim_{p\to 0} M_p(x, y)$.

Let us now consider the geometric mean $G(X, Y)$ of two symmetric positive-definite (SPD) matrices X and Y of size $d \times d$. SPD matrices generalize positive reals. We shall investigate the three generalizations of the above approaches of the scalar geometric mean, and show that they yield different notions of matrix geometric means when $d > 1$.

First, the AHM iterations can be extended to SPD matrices instead of reals:

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

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 [3, 14] (AHM):

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

Remarkably, the matrix AHM enjoys quadratic convergence to the following SPD matrix:

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

When $X = x$ and $Y = y$ are positive reals, we recover $G(X, Y) = \sqrt{xy}$. When $X = I$, the identity matrix, we get $G(I, Y) = Y^{\frac{1}{2}} = \sqrt{Y}$, the positive square root of SPD matrix Y. Thus the matrix AHM iterations provide a fast method in practice to numerically approximate matrix square roots by bypassing the matrix eigendecomposition. When matrices X and Y commute (i.e., $XY = YX$), we have $G(X, Y) = \sqrt{XY}$. The geometric mean $G(A, B)$ is proven to be the unique solution to the matrix Ricatti equation $XA^{-1}X = B$, is invariant under inversion (i.e., $G(A, B) = G(A^{-1}, B^{-1})^{-1}$), and satisfies the determinant property $\det(G(A, B)) = \sqrt{\det(A)\det(B)}$.

Let $\mathbb{P}$ denote the set of symmetric positive-definite $d \times d$ matrices. The matrix geometric mean can be interpreted using a Riemannian geometry [5] of the cone $\mathbb{P}$: Equip $\mathbb{P}$ with the trace metric tensor, i.e., a collection of smoothly varying inner products g_P for $P \in \mathbb{P}$ defined by

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

where S_1 and S_2 are matrices belonging to the vector space of symmetric $d \times d$ matrices (i.e., S_1 and S_2 are geometrically vectors of the tangent plane T_P of $P \in \mathbb{P}$). The geodesic length distance on the Riemannian manifold $(\mathbb{P}, g)$ is

$$\rho(P_1, P_2) = \left\|\log\left(P_1^{-\frac{1}{2}}P_2P_1^{-\frac{1}{2}}\right)\right\|_F = \sqrt{\sum_{i=1}^{d}\log^2\lambda_i\left(P_1^{-\frac{1}{2}}P_2P_1^{-\frac{1}{2}}\right)},$$

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. Interestingly, the matrix geometric mean $G(X, Y) = \mathrm{AHM}(X, Y)$ can also be interpreted as the Riemannian center of mass of X and Y:

$$G(X, Y) = \arg\min_{P\in\mathbb{P}} \frac{1}{2}\rho^2(X, P) + \frac{1}{2}\rho^2(Y, P).$$

This Riemannian least squares mean is also called the Cartan, Kärcher, or Fréchet mean in the literature. More generally, the Riemannian geodesic $\gamma(X, Y; t) = X\#_t Y$ between X and Y of $(\mathbb{P}, g)$ for $t \in [0, 1]$ is expressed using the weighted matrix geometric mean $G(X, Y; 1-t, t) = X\#_t Y$ minimizing

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

This Riemannian barycenter can be solved as

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

with $G(X, Y) = X\#_{\frac{1}{2}}Y$, $X\#_t Y = Y\#_{1-t}X$, and $\rho(X\#_t Y, X) = t\rho(X, Y)$, i.e., t is the arc length parameterization of the constant speed geodesic $\gamma(X, Y; t)$. When matrices X and Y commute, we have $X\#_t Y = X^{1-t}Y^t$. We thus interpret the matrix geometric mean $G(X, Y) = X\#Y = X\#_{\frac{1}{2}}Y$ as the Riemannian geodesic midpoint.

Second, let us consider the matrix geometric mean as the limit of matrix quasi-arithmetic power means which can be defined [13] as $Q_p(X, Y) = (X^p + Y^p)^{\frac{1}{p}}$ for $p \in \mathbb{R}, p \neq 0$, with $Q_1(X, Y) = A(X, Y)$ and $Q_{-1}(X, Y) = H(X, Y)$. We get $\lim_{p\to 0} Q_p(X, Y) = \mathrm{LE}(X, Y)$, the log-Euclidean matrix mean defined by

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

where exp and log denote the matrix exponential and the matrix logarithm, respectively. We have $\mathrm{LE}(X, Y) \neq G(X, Y)$. Consider the Loewner partial order $\preceq$ on the cone $\mathbb{P}$: $P \preceq Q$ if and only if $Q - P$ is positive semi-definite. A mean $M(X, Y)$ is said operator monotone [5] if for $X' \preceq X$ and $Y' \preceq Y$, we have $M(X', Y') \preceq M(X, Y)$. The log-Euclidean mean $\mathrm{LE}(X, Y)$ is not operator monotone but the Riemannian geometric matrix mean $G(X, Y)$ is operator monotone.

Third, we can define matrix power means $M_p(X, Y)$ for $p \in (0, 1]$ by uniquely solving the following matrix equation [13]:

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

Let $M_p(X, Y) = M$ denote the unique solution of Eq. 2. This equation is the matrix analogue of the scalar equation $m = \frac{1}{2}m^{1-p}x^p + \frac{1}{2}m^{1-p}y^p$ which can be solved as $m = \left(\frac{1}{2}x^p + \frac{1}{2}y^p\right)^{\frac{1}{p}} = M_p(x, y)$, i.e., the scalar p-power mean. In the limit case $p \to 0$, this matrix power mean M_p yields the matrix geometric/Riemannian mean [13]:

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

In general, we get the following closed-form expression [13] of this matrix power mean for $p \in (0, 1)$:

$$M_p(X, Y) = X\#_{\frac{1}{p}}\left(\frac{1}{2}X + \frac{1}{2}(X\#_p Y)\right).$$

Inductive means, circumcenters, and medians of several matrices. To extend these various binary matrix means of two matrices to matrix means of n matrices $P_1, \dots, P_n$ of $\mathbb{P}$, we can use induction sequences [9]. First, the n-variable matrix geometric mean $G(P_1, \dots, P_n)$ can be defined as the unique Riemannian center of mass:

$$G(P_1, \dots, P_n) = \arg\min_{P\in\mathbb{P}} \sum_{i=1}^{n} \frac{1}{n}\rho^2(P, P_i).$$

This geometric matrix mean $G = G(P_1, \dots, P_n)$ can be characterized as the unique solution of $\sum_{i=1}^{n}\log\left(G^{-\frac{1}{2}}P_iG^{-\frac{1}{2}}\right) = 0$ (called the Kärcher equation), and is proven to satisfy the ten Ando–Li–Mathias properties [1] defining what should be a good matrix generalization of the scalar geometric mean.

Holbrook [12] proposed the following sequence of iterations to approximate $G(P_1, \dots, P_n)$:

$$M_{t+1} = M_t\#_{\frac{1}{t+1}}P_t \mod n \tag{3}$$

with M_1 initialized to P_1. In the limit $t \to \infty$, we get the n-variable geometric mean: $\lim_{t\to\infty} M_t = G(P_1, \dots, P_n)$. This deterministic inductive definition of the matrix geometric mean by Eq. 3 allows to prove that the geometric mean $G(P_1, \dots, P_n)$ is monotone [12]: That is, if $P_1' \preceq P_1, \dots, P_n' \preceq P_n$ then we have $G(P_1', \dots, P_n') \preceq G(P_1, \dots, P_n)$. The following matrix arithmetic-geometric-harmonic mean inequalities extends the scalar case:

$$\begin{aligned}&H(X, Y; 1-t, t)\\&= ((1-t)X^{-1} + tY^{-1})^{-1}\\&\preceq G(X, Y; 1-t, t) \preceq A(X, Y; 1-t, t) = (1-t)X + tY.\end{aligned}$$

Now, if instead of taking cyclically the input matrices $P_1, \ldots, P_n, P_1, \ldots, P_n, \ldots$, we choose at iteration t the farthest matrix in $P_1, \ldots, P_n$ to M_t with respect to the Riemannian distance ρ, we get the Riemannian circumcenter [2] $C(P_1, \ldots, P_n)$ which is the minimax minimizer:

$$C(P_1, \ldots, P_n) = \arg\min_{C \in \mathbb{P}} \max_{i \in \{1, \ldots, n\}} \rho(P_i, C).$$

The sequence of iterations

$$C_{t+1} = C_t \#_{\frac{1}{t+1}} P_{\text{farthest}(t)}, \tag{4}$$

where

$$\text{farthest}(t) = \arg\max_{i \in \{1, \ldots, n\}} \rho(C_t, P_i),$$

initialized with $C_1 = P_1$ is such that

$$C(P_1, \ldots, P_n) = \lim_{t \to \infty} C_t.$$

The uniqueness of the smallest enclosing ball and the proof of convergence of the iterations of Eq. 4 relies on the fact that the cone $\mathbb{P}$ is of nonpositive sectional curvatures [2]: $\mathbb{P}$ is a nonpositive curvature space or NPC space for short.

The Riemannian median minimizing $\arg\min_{P \in \mathbb{P}} \sum_{i=1}^n \frac{1}{n} \rho(P, P_i)$ is proven to be unique in Riemannian NPC spaces, and can be obtained as the limit of the following cyclic order sequence [4]:

$$\begin{aligned} X_{kn+1} &= X_{kn} \#_{t_{k,1}} P_1, \\ X_{kn+2} &= X_{kn+1} \#_{t_{k,2}} P_2, \\ \vdots &= \vdots \\ X_{kn+n} &= X_{kn+n-1} \#_{t_{k,n}} P_n, \end{aligned}$$

where $t_{k,n} = \min\left\{1, \frac{\lambda_k}{n\rho(P_n, X_{kn+n-1})}\right\}$ with the positive real sequence (λ_k) such that $\sum_{k=0}^{\infty} \lambda_k = \infty$ and $\sum_{k=0}^{\infty} \lambda_k^2 < \infty$ (e.g., $\lambda_k = \frac{1}{k+1}$).

Finally, let us mention that Bini, Meini, and Poloni [6] proposed a class of recursive geometric matrix means $G_{s_1, \ldots, s_{n-1}}(P_1, \ldots, P_n)$ parameterized by $(n-1)$-tuple of scalar parameters, and defined recursively as the common limit of the following sequences:

$$\begin{aligned} &P_i^{(r+1)} \\ &= P_i^{(r)} \#_{s_1} G_{s_2, \ldots, s_{n-1}}(P_1, \ldots, P_{i-1}, P_{i+1}, \ldots, P_n), \quad i \in \{1, \ldots, n\}. \end{aligned}$$

In particular, these matrix means exhibit a unique $(n-1)$-tuple for which the recursive mean $G_{\frac{n-1}{n}, \frac{n-2}{n-1}, \ldots, \frac{1}{2}}(P_1, \ldots, P_n)$ converges fast in cubic order (order-3 convergence). This geometric matrix mean is called the BMP mean in the literature. Furthermore, the mean $G_{1,1,\ldots,1,\frac{1}{2}}(P_1, \ldots, P_n)$ coincides with the Ando–Li–Mathias geometric mean [1] (ALM) which converges linearly.

Random variables, expectations, and the law of large numbers. Although inductive means as limits of sequences have been considered since the 18th century (AGM by Lagrange and Gauss), this term was only recently coined by Karl-Theodor Sturm in 2003 (see Definition 4.6 in [15]), who considered inductive sequences to calculate probability expectations of random variables on nonpositive curvature complete metric spaces. For example, let $\mathcal{P}(\mathbb{P})$ denote the set of probability measures on $\mathbb{P}$ with bounded support [15]. Let $X : \Omega \to \mathbb{P}$ be a SPD-valued random variable with probability density function p_X expressed with respect to the canonical Riemannian volume measure $\mathrm{d}\omega(P) = \sqrt{\det(g_P)}$. The expectation $\mathbb{E}[X]$ and the variance $\mathbb{V}[X]$ of a random variable $X \sim p_X$ are defined respectively as the unique minimizer of $C \mapsto \mathbb{E}[\rho^2(X, C)] = \int_{\mathbb{P}} \rho^2(C, P) p_X(P) \mathrm{d}\omega(P)$ and $\inf_{P \in \mathbb{P}} \mathbb{E}[\rho^2(X, P)]$. Consider $(X_i)_{i \in \mathbb{N}}$ to be an independent sequence of measurable maps $X_i : \Omega \to \mathbb{P}$ with identical distributions $p_{X_i} = p_X$, and let $p_n = \frac{1}{n} \sum_{i=1}^n \delta_{X_i} \in \mathcal{P}(\mathbb{P})$ denote the empirical distribution. Then the following empirical law of large numbers holds as $n \to \infty$:

$$G(X_1, \ldots, X_n) \to \mathbb{E}[X].$$

Several proofs are reported in the literature (e.g., Proposition 6.6 of [15], Theorem 1 of [9], or Theorem 5.1 of [4]). Thus the expectation $\mathbb{E}[X]$ of a SPD-valued random variable can be estimated incrementally by considering increasing sequences $(X_i)_{i \in \mathbb{N}}$ of i.i.d. random vectors, and incrementally computing their Riemannian means. Experiments demonstrating convergence to various probability law expectations p_X are reported in [9].

Closing remarks. The AHM double sequence yielding the matrix geometric mean can further be generalized to define self-dual operators on convex functionals in Hilbert spaces [3] based on the Legendre–Fenchel transformation (called convex geometric mean functionals). For example, the AHM iterations initialized on a pair of nonzero complex numbers $z_1 = r_1 e^{i\theta_1}$ and $z_2 = r_2 e^{i\theta_2}$ expressed in polar forms is proven to converge to $\mathrm{AHM}(z_1, z_2) = \sqrt{r_1 r_2} e^{i\frac{\theta_1+\theta_2}{2}}$ which involves both the scalar arithmetic mean $A(\theta_1, \theta_2)$ and the scalar geometric mean $G(r_1, r_2)$.

To conclude, let us say that not only is it important to consider which mean we mean [10] but it is also essential to state which matrix geometric mean we mean!

References

[1] T. Ando, Chi-Kwong Li, and Roy Mathias, *Geometric means*, Linear Algebra Appl. **385** (2004), 305–334, DOI 10.1016/j.laa.2003.11.019. MR2063358

[2] Marc Arnaudon and Frank Nielsen, *On approximating the Riemannian 1-center*, Comput. Geom. **46** (2013), no. 1, 93–104, DOI 10.1016/j.comgeo.2012.04.007. MR2949613

[3] Marc Atteia and Mustapha Raïssouli, *Self dual operators on convex functionals; geometric mean and square root of convex functionals*, J. Convex Anal. **8** (2001), no. 1, 223–240. MR1829063

[4] Miroslav Bačák, *Computing medians and means in Hadamard spaces*, SIAM J. Optim. **24** (2014), no. 3, 1542–1566, DOI 10.1137/140953393. MR3264572

[5] Rajendra Bhatia and John Holbrook, *Riemannian geometry and matrix geometric means*, Linear Algebra Appl. **413** (2006), no. 2-3, 594–618, DOI 10.1016/j.laa.2005.08.025. MR2198952

[6] Dario A. Bini, Beatrice Meini, and Federico Poloni, *An effective matrix geometric mean satisfying the Ando-Li-Mathias properties*, Math. Comp. **79** (2010), no. 269, 437–452, DOI 10.1090/S0025-5718-09-02261-3. MR2552234

[7] Jonathan M. Borwein and Peter B. Borwein, *Pi and the AGM*, Canadian Mathematical Society Series of Monographs and Advanced Texts, John Wiley & Sons, Inc., New York, 1987. A study in analytic number theory and computational complexity; A Wiley-Interscience Publication. MR877728

[8] P. S. Bullen, *Handbook of means and their inequalities*, Mathematics and its Applications, vol. 560, Kluwer Academic Publishers Group, Dordrecht, 2003. Revised from the 1988 original [P. S. Bullen, D. S. Mitrinović and P. M. Vasić, *Means and their inequalities*, Reidel, Dordrecht; MR0947142], DOI 10.1007/978-94-017-0399-4. MR2024343

[9] Guang Cheng, Jeffrey Ho, Hesamoddin Salehian, and Baba C. Vemuri, *Recursive computation of the Fréchet mean on non-positively curved Riemannian manifolds with applications*, Riemannian computing in computer vision, Springer, Cham, 2016, pp. 21–43. MR3444345

[10] Miguel de Carvalho, *Mean, what do you mean?*, Amer. Statist. **70** (2016), no. 3, 270–274, DOI 10.1080/00031305.2016.1148632. MR3535513

[11] D. M. E. Foster and G. M. Phillips, *A generalization of the Archimedean double sequence*, J. Math. Anal. Appl. **101** (1984), no. 2, 575–581, DOI 10.1016/0022-247X(84)90121-5. MR748590

[12] John Holbrook, *No dice: a deterministic approach to the Cartan centroid*, J. Ramanujan Math. Soc. **27** (2012), no. 4, 509–521. MR3027448

[13] Yongdo Lim and Miklós Pálfia, *Matrix power means and the Karcher mean*, J. Funct. Anal. **262** (2012), no. 4, 1498–1514, DOI 10.1016/j.jfa.2011.11.012. MR2873848

[14] Yoshimasa Nakamura, *Algorithms associated with arithmetic, geometric and harmonic means and integrable systems*, J. Comput. Appl. Math. **131** (2001), no. 1-2, 161–174, DOI 10.1016/S0377-0427(00)00316-2. MR1835710

[15] Karl-Theodor Sturm, *Probability measures on metric spaces of nonpositive curvature*, Heat kernels and analysis on manifolds, graphs, and metric spaces (Paris, 2002), Contemp. Math., vol. 338, Amer. Math. Soc., Providence, RI, 2003, pp. 357–390, DOI 10.1090/conm/338/06080. MR2039961

Frank Nielsen

Credits

Photo of Frank Nielsen is courtesy of Maryse Beaumont.