# Averaging operations on matrices

Tanvi Jain

Indian Statistical Institute

July 3, 2014

# Three Classical Averages on Positive Numbers

Let $a_1, a_2, \ldots, a_m$ be $m$ positive numbers.

## Arithmetic Mean

$$A(a_1, a_2, \cdots, a_m) = \frac{a_1 + a_2 + \cdots + a_m}{m}.$$

## Geometric Mean

$$G(a_1, a_2, \cdots, a_m) = a_1^{1/m} a_2^{1/m} \cdots a_m^{1/m}.$$

## Harmonic Mean

$$H(a_1, a_2, \cdots, a_m) = \left[\frac{a_1^{-1} + a_2^{-1} + \cdots + a_m^{-1}}{m}\right]^{-1}.$$

# Positive Matrices

An $n \times n$ complex matrix $A$ is called *positive semidefinite (psd)* if it is Hermitian and $x^*Ax \geq 0$ for all vectors $x$.

A psd matrix $A$ is *positive definite (pd)* if $x^*Ax > 0$ for all nonzero vectors $x$.

The space of $n \times n$ pd matrices is denoted by $\mathbb{P}_n$.

# Role of Positive Definite Matrices

- *Diffusion Tensor Imaging*: $3 \times 3$ pd matrices model water flow at each voxel of brain scan.
- *Elasticity*: $6 \times 6$ pd matrices model stress tensors.
- *Machine Learning*: $n \times n$ pd matrices occur as kernel matrices.

# Extension of Averages to Positive matrices

Let $A_1, \ldots, A_m \in \mathbb{P}_n$.

*Arithmetic Mean*

$$\mathcal{A}(A_1, \cdots, A_m) = \frac{A_1 + \cdots + A_m}{m}.$$

*Harmonic Mean*

$$\mathcal{H}(A_1, \cdots, A_m) = \left(\frac{A_1^{-1} + \cdots + A_m^{-1}}{m}\right)^{-1}.$$

# What About Geometric Mean

*No straightforward extension of Geometric Mean*

If $A$ and $B$ do not commute, then the expected extension of geometric mean $A^{1/2}B^{1/2}$ is not even self-adjoint, leave alone positive definite.

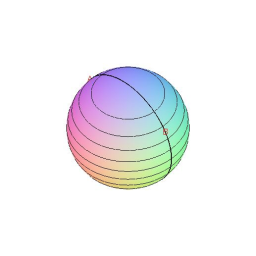
# Arithmetic Mean: Not always the best choice

**Arithmetic mean of two points on sphere does not lie on it.**

- Mean: Mid-point of the minor arc of the great circle through the points.
- Antipodal points do not have a unique mean.
- In spaces of positive curvature, mean need not be unique.

**Arithmetic mean on a surface need not lie on it.**

$$S^2 = \{(x, y, z) : x^2 + y^2 + z^2 = 1\}.$$

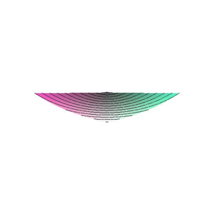


## Arithmetic mean of objects in a space need not lie in the space.

[Frechet; 1948] Finding mean of right-angled triangles.

$$\mathcal{S} = \{(x,y,z) \in \mathbb{R}^{+3} : x^2 + y^2 = z^2\}$$
$$= \left\{ \begin{bmatrix} z & x - \iota y \\ x + \iota y & z \end{bmatrix} : x, y, z > 0, z^2 = x^2 + y^2 \right\}.$$

## Surface of right triangles : Arithmetic mean not on $\mathcal{S}$.



**The determinant of arithmetic mean is not a mean value of the determinants.**

The determinant of the arithmetic mean is often more than that of input values which is not physically plausible in diffusion tensor imaging.

**Arithmetic mean is not self dual.**

In elasticity, it is desired that the mean of the inverses is the inverse of the mean. This is not satisfied by the arithmetic mean.

# Characteristic Property of Geometric Mean on Positive Numbers

If $\mathcal{M}$ is a mean on positive numbers satisfying

$$\mathcal{M}(a^{-1}, b^{-1}) = (\mathcal{M}(a, b))^{-1}$$

then it is always the *geometric mean*.

# Matrix Geometric Mean

Let $\mathcal{G}$ be a mapping from $\mathbb{P}_n^m$ to $\mathbb{P}_n$. Then $\mathcal{G}$ is a geometric mean if it satisfies the following conditions.

1) If $A_i$s commute, then
$\mathcal{G}(A_1, A_2, \cdots, A_m) = A_1^{1/m} A_2^{1/m} \cdots A_m^{1/m}$.

2) $\mathcal{G}(\alpha_1 A_1, \cdots, \alpha_m A_m) = (\alpha_1 \cdots \alpha_m)^{1/m} \mathcal{G}(A_1, \cdots, A_m)$.

3) For any permutation $\sigma$ on $m$ symbols,
$\mathcal{G}(A_1, \cdots, A_m) = \mathcal{G}(A_{\sigma(1)}, \cdots, A_{\sigma(m)})$.

4) For an invertible matrix $X$,
$\mathcal{G}(XA_1X^*, \cdots, XA_mX^*) = X\mathcal{G}(A_1, \cdots, A_m)X^*$.

5) $\mathcal{G}(A_1^{-1}, \cdots, A_m^{-1}) = (\mathcal{G}(A_1, \cdots, A_m))^{-1}$.

6) $\det(\mathcal{G})(A_1, \cdots, A_m)) = (\det(A_1) \cdots \det(A_m))^{1/m}$.

$+ \cdots$ (other conditions depending on the requirements in different areas.)

# Geometric Mean of Two Positive Definite Matrices

[Pusz,Woronowicz; 1975]

The Geometric mean of two pd matrices $A$ and $B$ is given by

$$\mathcal{G}(A, B) = A^{1/2}(A^{-1/2}BA^{-1/2})^{1/2}A^{1/2}$$

Further this is the unique geometric mean satisfying the mentioned properties.

# Arithmetic Mean as Centre of Mass in Euclidean Geometry

Arithmetic mean of two pd matrices is the mid-point of the straight line joining the two matrices.

$$\mathcal{A}(A_1, \cdots, A_m) = \arg \min_{X \in \mathbb{P}_n} \|A_i - X\|_2^2,$$

where $\|B\|_2 = \left( \sum_{i,j=1}^{n} |b_{ij}|^2 \right)^{1/2}$ denotes the Frobenius norm of $B$.

# Riemannian Geometry and Geometric Mean

## $\mathbb{P}_n$ as a Riemannian Manifold

$\mathbb{P}_n$ is a Riemannian manifold with the geodesic between two points $A$ and $B$ given by

$$\Gamma(t) = A^{1/2}(A^{-1/2}BA^{-1/2})^t A^{1/2},$$

and the geodesic distance $\delta_2$ given by

$$\begin{aligned} \delta_2(A, B) &= L(\Gamma) \\ &= \| \log(A^{-1/2}BA^{-1/2})\|_2. \end{aligned}$$

## Geometric mean: Mid-point of the geodesic and centre of mass.

$$\mathcal{G}(A, B) = \arg\min_{X \in \mathbb{P}_n} \delta_2^2(X, A) + \delta_2^2(X, B).$$

# Multivariate Matrix Geometric Mean

**[Moakher; 2005], [Bhatia, Holbrook; 2006]**

Proposed geometric mean of $A_1, \ldots, A_m$ to be

$$\mathcal{G}(A_1, \cdots, A_m) = \arg \min_{X \in \mathbb{P}_n} \sum_{i=1}^{m} \delta_2^2(X, A_i).$$

Taking the gradient, the geometric mean is the unique solution of the equation

$$\sum_{i=1}^{m} \log(XA_i^{-1}) = 0.$$

## No closed expression for multivariate geometric mean

There is no closed expression known for the multivariate geometric mean even in the case of $2 \times 2$ positive definite matrices. This can be computed numerically through iterative procedures, e.g. one can use the Matlab tool box provided by D. Bini and B. Iannazzo.

# Geometric Mean of the Exponentials of Triple of Pauli Matrices

## Pauli matrices and their exponentials

$$\sigma_1 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \qquad A_1 = \begin{bmatrix} \cosh(1) & \sinh(1) \\ \sinh(1) & \cosh(1) \end{bmatrix}$$

$$\sigma_2 = \begin{bmatrix} 0 & -\iota \\ \iota & 0 \end{bmatrix} \qquad A_2 = \begin{bmatrix} \cosh(1) & -\iota \sinh(1) \\ \iota \sinh(1) & \cosh(1) \end{bmatrix}$$

$$\sigma_3 = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \qquad A_3 = \begin{bmatrix} \mathrm{e} & 0 \\ 0 & \mathrm{e}^{-1} \end{bmatrix}$$

## Geometric Mean $\mathcal{G}$

$$\mathcal{G}(A_1, A_2, A_3) = \frac{A_1 + A_2 + A_3}{\sqrt{\det(A_1 + A_2 + A_3)}}.$$

# References


(1) R. Bhatia and J. Holbrook, *Riemannian geometry and matrix geometric means*, Linear Algebra Appl., 413 (2006) 594-618.

(2) R. Bhatia and T. Jain, *The geometric mean of the exponentials of Pauli matrices*, preprint.

(3) D. Bini and B. Iannazzo, *Matrix Computation Toolbox,* http://bezout.dm.unipi.it/software/mmtoolbox/.

(4) M. Moakher, *A differential approach to the geometric mean of symmetric positive definite matrices*, SIAM J. Matrix Anal. Appl., 26 (2005) 735-747.

(5) W. Pusz, S. L. Woronowicz, *Functional calculus for sesquilinear forms and the purification map*, Rep. Math. Phy., 8 (1975) 159-175.