isid/ms/2014/9
July 8, 2014
http://www.isid.ac.in/~statmath/eprints

The geometric mean of exponentials of Pauli matrices

Rajendra Bhatia and Tanvi Jain

Indian Statistical Institute, Delhi Centre
7, SJSS Marg, New Delhi–110 016, India

THE GEOMETRIC MEAN OF EXPONENTIALS OF PAULI MATRICES

RAJENDRA BHATIA AND TANVI JAIN

Abstract. An explicit formula is found for the geometric mean of exponentials of the Pauli matrices.

1. Introduction

Let $\mathbb{P}(n)$ be the space of $n \times n$ complex positive definite matrices. This space is equipped with a natural Riemannian metric $ds = (\mathrm{tr}(A^{-1}\, dA)^2)^{1/2}$. The associated metric distance between A and B is given by

$$\delta_2(A, B) = \| \log\, A^{-1/2}\, B\, A^{-1/2} \|_2, \tag{1}$$

where $\|X\|_2 = (\mathrm{tr}\, X^* X)^{1/2}$ is the Frobenius norm. With this $\mathbb{P}(n)$ is a complete Riemannian manifold of nonpositive curvature.

Let $A_1, \ldots, A_m$ be any given points in $\mathbb{P}(n)$. The Riemannian barycentre of these points is defined as

$$G(A_1, \ldots, A_m) = \underset{X \in \mathbb{P}(n)}{\mathrm{argmin}} \sum_{j=1}^{m} \delta_2^2(X, A_j). \tag{2}$$

It is a classical theorem of E. Cartan that the minimiser in (2) exists and is unique. This is also the unique positive definite solution of the equation.

$$\sum_{j=1}^{m} \log\left(X^{1/2}\, A_j^{-1}\, X^{1/2}\right) = 0. \tag{3}$$

In recent years $G(A_1, \ldots, A_m)$ has been presented as the "geometric mean" of $A_1, \ldots, A_m$ (See [Bh], [BH], [M1]) and has been variously called the Riemannian mean, the Karcher mean, or the least squares mean. Its operator theoretic properties have been studied in several papers such as [BH], [LL], [LP1]. At the same time this mean has been adopted as an appropriate notion of averaging (or of smoothing of data) in areas such as diffusion tensor imaging, radar signal processing, elasticity, and statistics on manifolds. See, for example, [Ba], [FJ], [M2].

For two positive definite matrices A and B, an explicit formula for $G(A,B)$ has long been known. In this case $G(A,B)$ is denoted by $A\#B$, and is given by the well-known formula of Pusz and Woronowicz [PW]

$$A\#B = A^{1/2}\left(A^{-1/2}\,B\,A^{-1/2}\right)^{1/2}\ A^{1/2}. \tag{4}$$

It is also known that

$$A\#B = A(A^{-1}B)^{1/2} = (AB^{-1})^{1/2}B. \tag{5}$$

See Chapters 4 and 6 of [Bh] for a discussion of these topics and for references.

In the case of three or more matrices no such expression has been found. However, good numerical algorithms for its computation have been developed [BI], [JVV].

To advance our understanding of the multivariate geometric mean it would be useful to have "by-hand" computations, at least for some special examples. In this note we present such a computation.

The simplest situation is when $n = 2$ and $m = 3$. The most famous triple of 2×2 Hermitian matrices is the Pauli matrices

$$\sigma_1 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \sigma_2 = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}, \sigma_3 = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}. \tag{6}$$

The exponentials of these matrices are

$$A_1 = \begin{bmatrix} \cosh(1) & \sinh(1) \\ \sinh(1) & \cosh(1) \end{bmatrix}, \qquad A_2 = \begin{bmatrix} \cosh(1) & -i\sinh(1) \\ i\sinh(1) & \cosh(1) \end{bmatrix},$$
$$A_3 = \begin{bmatrix} e & 0 \\ 0 & e^{-1} \end{bmatrix}. \tag{7}$$

A little more generally, we consider the triple

$$A_1 = \begin{bmatrix} \cosh x & \sinh x \\ \sinh x & \cosh x \end{bmatrix}, \qquad A_2 = \begin{bmatrix} \cosh x & -i\sinh x \\ i\sinh x & \cosh x \end{bmatrix},$$
$$A_3 = \begin{bmatrix} e^x & 0 \\ 0 & e^{-x} \end{bmatrix}. \tag{8}$$

where x is any real number. Then A_1, A_2, A_3 are positive definite matrices, each having its determinant equal to one. We have the following theorem.

Theorem 1. *Let A_1, A_2, A_3 be the three matrices in* (8)*. Then*

$$G(A_1, A_2, A_3) = \frac{A_1 + A_2 + A_3}{[\det(A_1 + A_2 + A_3)]^{1/2}}. \tag{9}$$

Apart from the Riemannian mean G defined in (2), other candidates for a geometric mean have been proposed [ALM]. The first one called the Ando-Li-Mathias (ALM, for short) mean is given by an inductive procedure, which in the case of three matrices can be described as follows. Given A_1, A_2, A_3 define a sequence of triples as follows

$$\begin{aligned} \mathbf{A^0} &= (A_1, A_2, A_3), \\ \mathbf{A^{(k+1)}} &= \left(A_1^{(k)} \# A_2^{(k)}, A_2^{(k)} \# A_3^{(k)}, A_3^{(k)} \# A_1^{(k)}\right), \end{aligned}$$

$k = 1, 2, 3, \ldots$. Then as $k \to \infty$, the sequence $A_i^{(k)} \# A_j^{(k)}$, $1 \le i, j \le 3$, converges. The limit of this sequence is called the ALM mean of A_1, A_2, A_3. A variant of this construction is used to define the BMP mean given by Bini, Meini and Poloni [BMP]. These means share several important properties (the 10 "ALM properties") but are generally different from each other. We will see that in the example considered in Theorem 1, the Riemannian, the ALM and the BMP means coincide.

For two 2×2 matrices of determinant one, a formula similar to (9) is known to be true. However, for three matrices it holds rarely. After the proof of Theorem 1 in the next section, we identify some other triples for which the same formula is valid.

2. Proof

We split the proof of Theorem 1 into two parts. The first is a proposition about the geometric mean of all 2×2 positive matrices, while the second part exploits the special structure of Pauli matrices.

Proposition 2. *Let $A_1, \ldots, A_m$ be 2×2 positive definite matrices. Then $G(A_1, \ldots, A_m)$ is in the linear span of $A_1, \ldots, A_m$.*

Proof. The geodesic segment joining any two points A, B in the manifold $\mathbb{P}(n)$ can be naturally parametrised as

$$A \#_t B = A^{1/2} \left(A^{-1/2} B A^{-1/2}\right)^t A^{1/2}, \quad 0 \le t \le 1. \tag{10}$$

It is evident that $A \# B$ is the midpoint $A \#_{1/2} B$ of this geodesic.

The Riemannian mean $G(A_1, \ldots, A_m)$ can be obtained as a limit of "asymmetric averages" constructed from $A_1, \ldots, A_m$ as follows. Let $\overline{k} = k(\operatorname{mod} m)$, and define the sequence S_k as follows

$$S_1 = A_1, \;\; S_{k+1} = S_k \#_{1/(k+1)} A_{\overline{k+1}}. \tag{11}$$

Then

$$G(A_1, \ldots, A_m) = \lim_{k \to \infty} S_{k+1}. \tag{12}$$

This is a major theorem in the subject, established in a series of papers [LL], [BK], [H], [LP2].

Now we restrict ourselves to the case $n = 2$. Then, it is known that if A and B have determinant one, then

$$A \#_{1/2} B = \frac{A + B}{(\det\, A + \det\, B)^{1/2}}$$

See [M1], and Proposition 4.1.12 in [Bh]. This shows that $A\#_{1/2}B$ is in the linear span of A and B. (This is true for all A, B in $\mathbb{P}(2)$.) Taking the geometric mean of $A\#_{1/2}B$ with A and B, we see that $A\#_{1/4}B$ and $A\#_{3/4}B$ are in the linear span of A and B. This argument can be repeated to show that $A\#_t B$ is in the linear span of A and B for all dyadic rationals t in $[0, 1]$. By continuity this is true for all t in $[0, 1]$.

This argument shows that if $A_1, \ldots, A_m$ are 2×2 positive matrices, then all terms of the sequence S_k defined by (11) are in the linear span of $A_1, \ldots, A_m$. Hence, so is their limit $G(A_1, \ldots, A_m)$. ■

Remark The ALM and BMP means are also defined as limits of sequences of binary means. So Proposition 2 is equally valid for these means.

Now we come to the proof of Theorem 1. Let U be the unitary matrix

$$U = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ i & -i \end{bmatrix}.$$

The crux of our argument lies in the observation that the three matrices in (8) satisfy the relations

$$U^* A_1 U = A_2, \quad U^* A_2 U = A_3, \quad U^* A_3 U = A_1, \tag{13}$$

which can be readily verified. So, using the symmetry and congruence-invariance properties of the geometric mean, we have

$$\begin{aligned} G(A_1, A_2, A_3) = G(A_2, A_3, A_1) &= G(U^*A_1U, U^*A_2U, U^*A_3U) \\ &= U^*G(A_1, A_2, A_3)U. \end{aligned} \tag{14}$$

By Proposition 2, there exist real numbers $\alpha_1, \alpha_2, \alpha_3$ such that

$$G(A_1, A_2, A_3) = \alpha_1 A_1 + \alpha_2 A_2 + \alpha_3 A_3. \tag{15}$$

From the relations (13), (14) and (15) one can see that

$$\alpha_1 A_1 + \alpha_2 A_2 + \alpha_3 A_3 = \alpha_1 A_2 + \alpha_2 A_3 + \alpha_3 A_1. \tag{16}$$

But A_1, A_2, A_3 are linearly independent. So, it follows from (16) that $\alpha_1 = \alpha_2 = \alpha_3$. Hence $G(A_1, A_2, A_3) = \alpha(A_1 + A_2 + A_3)$ for some positive real number α. Since A_1, A_2, A_3 all have determinant one, $\det G(A_1, A_2, A_3) = 1$. Hence

$$\begin{aligned} G(A_1, A_2, A_3) &= \frac{\alpha(A_1 + A_2 + A_3)}{[\det(\alpha(A_1 + A_2 + A_3))]^{1/2}} \\ &= \frac{A_1 + A_2 + A_3}{[\det(A_1 + A_2 + A_3)]^{1/2}}. \end{aligned}$$

■

The properties of G that we have invoked in the proof are possessed by the ALM and BMP means. So the formula (9) is valid for these means too.

3. Remarks

The crucial properties of the triple (8) used in the proof of (9) are that A_1, A_2, A_3 are linearly independent over $\mathbb{R}$, have determinant one, and there exists a matrix X with $|\det X| = 1$ such that $X^*(A_1, A_2, A_3)X = (A_2, A_3, A_1)$. There are other triples with these properties.

Let X be any 2×2 nonnormal matrix whose spectrum consists of two different cube roots of unity. Then X is unitarily similar to one of the matrices

$$\begin{bmatrix} \omega & 0 \\ c & 1 \end{bmatrix}, \begin{bmatrix} \omega^2 & 0 \\ c & 1 \end{bmatrix}, \begin{bmatrix} \omega & 0 \\ c & \omega^2 \end{bmatrix},$$

where $c \neq 0$. Then the matrices

$$A_1 = X^*X, \quad A_2 = X^{*2}X^2, \quad A_3 = X^{*3}X^3 = I$$

have the properties mentioned above. So the formula (9) is valid for them.

More generally, if we choose any 2×2 matrix X with $|\det X| = 1$, and $(\operatorname{tr} X)^2 = \det X$, and let D be any positive diagonal matrix, then the triple

$$A_1 = X^* D X, \quad A_2 = X^* A_1 X, \quad A_3 = X^* A_2 X = D,$$

has the required properties.

References

[ALM] T. Ando, C.-K. Li and R. Mathias, Geometric means, Linear Algebra Appl., 385 (2004) 305-334.

[Ba] F. Barbaresco, Innovative tools for radar signal processing based on Cartan's geometry of SPD matrices and information geometry, IEEE Radar Confererence, Rome, May 2008.

[Bh] R. Bhatia, Positive Definite Matrices, Princeton University Press, 2007.

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

[BK] R. Bhatia and R. L. Karandikar, Monotonicity of the matrix geometric mean, Math. Ann., 353 (2012) 1453-1467.

[BI] D. A. Bini and B. Iannazzo, Computing the Karcher mean of positive definite matrices, Linear Algebra Appl., 438 (2013), 1700-1710.

[BMP] D. A. Bini, B. Meini and F. Poloni, An effective geometric mean satisfying the Ando-Li-Mathias properties, Math. Comp., 79 (2010) 437-452.

[FJ] P. Fletcher and S. Joshi, Riemannian geometry for the statistical analysis of diffusion tensor data, Signal Processing, 87 (2007) 250-262.

[JVV] B. Juris, R. Vanderbil and B. Vandereycken, A survey and comparison of contemporary algorithms for computing the matrix geometric mean, Electron. Trans. Numer Anal., 39 (2012) 379-402.

[H] J. Holbrook, No dice: a deterministic approach to the Cartan centroid, J. Ramanujan Math. Soc., 27 (2012) 509-521.

[LL] J. Lawson and Y. Lim, Monotonic properties of the least squares mean, Math. Ann., 351 (2011) 267-279.

[LP1] Y. Lim and M. Palfia, Matrix power means and the Karcher mean, J. Funct. Anal., 262 (2012) 1498-1514.

[LP2] Y. Lim and M. Palfia, Weighted deterministic walks for the least squares mean on Hadamard spaces, Bull. London Math. Soc., to appear.

[M1] M. Moakher, A differential geometric approach to the geometric mean of symmetric positive-definite matrices, SIAM J. Matrix Anal. Appl., 26 (2005) 735-747.

[M2] M. Moakher, On the averaging of symmetric positive-definite tensors, J. Elasticity, 82 (2006) 273-296.

[PW] W. Pusz and S. L. Woronowicz, Functional calculus for sesquilinear forms and the purification map, Rep. Math. Phys., 8 (1975) 159-170.

Indian Statistical Institute, New Delhi
E-mail address: rbh@isid.ac.in

Indian Statistical Institute, New Delhi
E-mail address: tanvi@isid.ac.in