

isid/ms/2002/14

May 10, 2002

<http://www.isid.ac.in/~statmath/eprints>

On the exponential metric increasing property

RAJENDRA BHATIA

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

On the exponential metric increasing property

Rajendra Bhatia
Indian Statistical Institute
7, S. J. S. Sansanwal Marg
New Delhi - 110 016 (INDIA)
Email:rbh@isid.ac.in

Abstract

The metric increasing property of the exponential map is known to be equivalent to the fact that the set of positive definite matrices is a Riemannian manifold of nonpositive curvature. We show that this property is an easy consequence of the logarithmic-geometric mean inequality for positive numbers. Operator versions of this inequality lead to a generalisation of the exponential metric increasing property to all Schatten-von Neumann norms.

1 Introduction

For a fixed n , let \mathbb{M} be the space of $n \times n$ complex matrices, \mathbb{S} the (real vector) subspace of \mathbb{M} consisting of Hermitian matrices, and \mathbb{P} the subset consisting of positive (definite) matrices.

Let $s_1(A) \geq \cdots \geq s_n(A)$ be the singular values of A (these are the positive square roots of the eigenvalues of A^*A). For each $A \in \mathbb{M}$ let $\|A\|_2 := \left[\sum s_j^2(A) \right]^{\frac{1}{2}} = (\text{tr } A^*A)^{\frac{1}{2}}$. This is the norm associated with the inner product $\langle A, B \rangle = \text{tr } A^*B$ on \mathbb{M} . We denote by $\lambda_i(A)$, $1 \leq i \leq n$, the eigenvalues of A , and by $\text{Eig}A$ a diagonal matrix with diagonal entries $\lambda_i(A)$.

If A, B are positive, then the product AB has positive eigenvalues. For $A, B \in \mathbb{P}$ let

$$\delta_2(A, B) := \|\log \text{Eig}(AB^{-1})\|_2 = \left[\sum_{i=1}^n \log^2 \lambda_i(AB^{-1}) \right]^{\frac{1}{2}}. \quad (1)$$

This defines a metric on the manifold \mathbb{P} . The tangent space to \mathbb{P} at any of its points A is the space $T_A\mathbb{P} = \{A\} \times \mathbb{S}$. The metric (1) is the distance associated with the arc length with respect to the Riemannian metric $ds^2 = \text{tr}(A^{-1}dA)^2$.

The exponential map is a bijection of \mathbb{S} onto \mathbb{P} . The *exponential metric increasing property* (EMI) says that for all $A, B \in \mathbb{S}$

$$\|A - B\|_2 \leq \delta_2(\exp A, \exp B) \quad (2)$$

and if A, B are on the same line through the origin, then (2) is an equality.

This property is equivalent to another important property of \mathbb{P} : it is a Riemannian manifold with nonpositive curvature; see Chapters XI and XII of S. Lang [10]. (The discussion there is confined to real matrices.) The standard general reference for the subject is [1].

In this note we present a proof of (2) that is much shorter than that of Lang [10] which, in turn, is based on the exposition in Mostow [13]. Our proof may provide a little more insight into this inequality as it reduces it to the classical logarithmic-geometric mean inequality for positive numbers. After that we show that the EMI remains true when the norm $\|\cdot\|_2$ is replaced by any of the Schatten-von Neumann norms (also called unitarily invariant norms).

Every Schatten-von Neumann norm on \mathbb{M} arises as a *symmetric gauge function* of the singular values [2]. We use the notation $\|\cdot\|_\Phi$ for the norm on \mathbb{M} corresponding to the symmetric gauge function Φ on \mathbb{R}^n . To show that each such function when $\|\cdot\|_2$, in the definition (1) is replaced by $\|\cdot\|_\Phi$, leads to a metric δ_Φ on \mathbb{P} we use a theorem of Lidskii [12]. This is one from the circle of ideas that have recently come into prominence because of the solution of Horn's problem [3,8]. To prove a general version of (2) we use an operator analogue of the logarithmic-geometric mean inequality proved in [9], and later in [5]. The inequalities obtained in this note have an independent interest as operator inequalities. For example, the generalized EMI is a strengthening in several ways of the famous Golden-Thompson inequality of mathematical physics.

2 A simple proof of the EMI

Let $\exp'(A)$ denote the derivative, at a point A , of the exponential map \exp from \mathbb{S} onto \mathbb{P} . By standard calculus arguments the EMI (2) is a consequence of the inequality

$$\|B\|_2 \leq \|\exp(-A)\exp'(A)(B)\|_2 \quad (3)$$

valid for all $A, B \in \mathbb{S}$.

We have the well-known formula [2, p.311]

$$\exp'(A)(B) = \int_0^1 e^{tA} B e^{(1-t)A} dt. \quad (4)$$

The *logarithmic mean* of two positive numbers a, b is the quantity

$$L(a, b) := \frac{a - b}{\log a - \log b} = \int_0^1 a^t b^{1-t} dt. \quad (5)$$

It is easy to see that this number lies in between the geometric and the arithmetic means of a and b . Here is a simple proof of the part that we need:

$$\sqrt{ab} \leq \frac{a - b}{\log a - \log b}. \quad (6)$$

To see this assume $b < a$, divide both sides by b , and then replace a/b by x^2 . The inequality (6) then reduces to

$$2 \log x \leq \frac{x^2 - 1}{x} \text{ for } x \geq 1.$$

When $x = 1$ the two sides are zero, and for $x > 1$ the derivative of the left-hand side is smaller than that of the right-hand side.

Now let A be any positive matrix. Then for any matrix X we have

$$\|A^{\frac{1}{2}} X A^{\frac{1}{2}}\|_2 \leq \left\| \int_0^1 A^t X A^{1-t} dt \right\|_2 \quad (7)$$

To see this choose an orthonormal basis in which A is diagonal with diagonal entries λ_i . Then the matrix on the left-hand side of (7) has entries $\sqrt{\lambda_i \lambda_j} x_{ij}$, that on the right has entries $\left[\int_0^1 \lambda_i^t \lambda_j^{1-t} dt \right] x_{ij}$. So, the inequality (7) follows from the logarithmic-geometric mean inequality.

Now to prove (3) let A, B be any two Hermitian matrices. Write $B = e^{A/2} \left(e^{-\frac{A}{2}} B e^{-\frac{A}{2}} \right) e^{\frac{A}{2}}$. Then using (7) we have

$$\begin{aligned} \|B\|_2 &\leq \left\| \int_0^1 e^{tA} \left(e^{-A/2} B e^{-A/2} \right) e^{(1-t)A} dt \right\|_2 \\ &= \|e^{-A/2} \left[\int_0^1 e^{tA} B e^{(1-t)A} dt \right] e^{-A/2}\|_2. \end{aligned}$$

Using (4) this gives

$$\|B\|_2 \leq \|e^{-\frac{A}{2}} [\exp'(A)(B)] e^{-\frac{A}{2}}\|_2. \quad (8)$$

To get the inequality (3) from this, we use the fact that if a matrix product XY is Hermitian, then

$$\|XY\|_2 \leq \|YX\|_2. \quad (9)$$

To see this note that the singular values of a Hermitian matrix are the absolute values of its eigenvalues. So,

$$\begin{aligned} \|XY\|_2^2 &= \sum s_j^2(XY) = \sum \lambda_j^2(XY) = \sum \lambda_j^2(YX) \\ &\leq \sum s_j^2(YX) = \|YX\|_2^2. \end{aligned}$$

(The inequality used in the above chain can be derived from Schur's theorem that says every operator can be reduced to an upper triangular form in some basis.)

3 The generalized EMI

In this section we replace the norm $\|\cdot\|_2$ by a more general class of norms. A norm Φ on \mathbb{R}^n is called a *symmetric gauge function* if it is invariant under permutations and sign changes of coordinates. It is customary to assume a normalisation condition $\Phi(1, 0, \dots, 0) = 1$. The norms $\|x\|_p = (\sum |x_j|^p)^{\frac{1}{p}}$, $1 \leq p \leq \infty$, are examples of such norms. For $A \in \mathbb{M}$, let

$\|A\|_\Phi = \Phi(\{s_j(A)\})$ where $\{s_j(A)\}$ are the singular values of A . Then $\|A\|_\Phi$ is a norm on \mathbb{M} that is unitarily invariant, i.e. $\|UAV\|_\Phi = \|A\|_\Phi$ for all unitary U, V . By a theorem of von Neumann all unitarily invariant norms on \mathbb{M} arise in this way. The ones corresponding to $\|\cdot\|_p$ are called the Schatten p -norms, $1 \leq p \leq \infty$.

For any vector $x = (x_1, \dots, x_n)$ in \mathbb{R}^n we write $x^\downarrow = (x_1^\downarrow, \dots, x_n^\downarrow)$ for the vector whose coordinates are obtained by arranging the x_j in decreasing order and $x^\uparrow = (x_1^\uparrow, \dots, x_n^\uparrow)$ for the vector obtained by arranging them in increasing order. We say $x \prec_w y$ if $\sum_{j=1}^k x_j^\downarrow \leq \sum_{j=1}^k y_j^\downarrow$ for all $1 \leq k \leq n$; and $x \prec y$ if, in addition, we have $\sum_{j=1}^n x_j^\downarrow = \sum_{j=1}^n y_j^\downarrow$.

Lemma 1. [2,p.45] If $x, y \in \mathbb{R}_+^n$ and $x \prec_w y$, then $\Phi(x) \leq \Phi(y)$ for every symmetric gauge function Φ .

The following theorem is a corollary of a theorem of Gel'fand and Naimark, and is sometimes called Lidskii's theorem.

Theorem 2.[2, p.73] Let $A, B \in \mathbb{P}$. Then

$$\{\log \lambda_i^\downarrow(A) + \log \lambda_i^\uparrow(B)\} \prec \{\log \lambda_i(AB)\} \prec \{\log \lambda_i^\downarrow(A) + \log \lambda_i^\downarrow(B)\}. \quad (10)$$

Let $A, B \in \mathbb{P}$ and let Φ be any symmetric gauge function. Define

$$\delta_\Phi(A, B) = \Phi(\{\log \lambda_i(AB^{-1})\}). \quad (11)$$

Proposition 3. δ_Φ is a metric on \mathbb{P} .

Proof Only the triangle inequality is nontrivial. Let A, B, C be three positive matrices. Then

$$\begin{aligned} \{\log \lambda_i(AC^{-1})\} &= \{\log \lambda_i(B^{-\frac{1}{2}}AC^{-1}B^{\frac{1}{2}})\} \\ &= \{\log \lambda_i(B^{-\frac{1}{2}}AB^{-\frac{1}{2}} \cdot B^{\frac{1}{2}}C^{-1}B^{\frac{1}{2}})\}. \end{aligned}$$

So, by Theorem 2

$$\begin{aligned} \{\log \lambda_i(AC^{-1})\} &\prec \{\log \lambda_i^\downarrow(B^{-\frac{1}{2}}AB^{-\frac{1}{2}}) + \log \lambda_i^\downarrow(B^{\frac{1}{2}}C^{-1}B^{\frac{1}{2}})\} \\ &= \{\log \lambda_i^\downarrow(AB^{-1}) + \log \lambda_i^\downarrow(BC^{-1})\}. \end{aligned}$$

Using Lemma 1, we get from this

$$\delta_\Phi(A, C) \leq \delta_\Phi(A, B) + \delta_\Phi(B, C). \quad \blacksquare$$

For any $X \in GL_n$, the transformation $\Gamma_X(A) := X^*AX$ is called a congruence on \mathbb{M} . Any element of \mathbb{P} is congruent to the identity I . So the group of congruences acts transitively on \mathbb{P} . It is easy to see that every congruence is an isometry of δ_Φ and so is the inversion map; i.e.

$$\delta_\Phi(X^*AX, X^*BX) = \delta_\Phi(A, B) = \delta_\Phi(A^{-1}, B^{-1}). \quad (12)$$

Next we show that the inequality (3) remains true for all norms $\|\cdot\|_{\Phi}$. The proof uses the same ideas as in Section 2; the tools needed are harder. Instead of (7) we need the inequality

$$\|A^{\frac{1}{2}}XA^{\frac{1}{2}}\|_{\Phi} \leq \left\| \int_0^1 A^t X A^{1-t} dt \right\|_{\Phi} \quad (13)$$

for positive A and for all Φ . This is true [5,9], though harder to prove. The more general version of (9)

$$\|XY\|_{\Phi} \leq \|YX\|_{\Phi} \quad (14)$$

whenever XY is Hermitian is also known to be true [2,p.253]. Combining these we can prove the general version of (3), and as a corollary the following.

Theorem 4. For any $A, B \in \mathbb{S}$ and for any symmetric gauge function Φ we have

$$\|A - B\|_{\Phi} \leq \delta_{\Phi}(\exp A, \exp B). \quad (15)$$

The two sides of (15) are equal if A, B lie on the same line through the origin.

(The second statement of the theorem is easy to prove.) So the EMI is true for all Schatten-von Neumann norms.

One special case of this, that of the norm $\|A\|_{\infty} = \max s_j(A)$ corresponding to the symmetric gauge function $\|x\|_{\infty} = \max |x_j|$, has been studied earlier by Corach, Porta, and Recht [7].

Let us discuss the inequality (15) in the context of known matrix inequalities. We can write it as

$$\|A - B\|_{\Phi} \leq \|\log(e^{-B/2}e^Ae^{-B/2})\|_{\Phi} \quad (16)$$

for Hermitian A, B and for all unitarily invariant norms. Using the properties of the exponential function with respect to the order \prec [2, Chapter II] one gets the weaker inequality

$$\|e^{A-B}\|_{\Phi} \leq \|e^{-B/2}e^Ae^{-B/2}\|_{\Phi}. \quad (17)$$

The special case $\|\cdot\|_{\Phi} = \|\cdot\|_{\infty}$ is known as Segal's inequality [15, p.260]. Changing signs and using (14) we get from (17) another known result [2, p.261]. We have

$$\|e^{A+B}\|_{\Phi} \leq \|e^Ae^B\|_{\Phi}. \quad (18)$$

Choosing the special norm $\|A\|_{\Phi} = \|A\|_{\text{tr}} = \sum_{j=1}^n s_j(A)$, reduces (18) to the famous Golden-Thompson inequality

$$\text{tr}(e^{A+B}) \leq \text{tr} e^A e^B.$$

Generalisations of this inequality have been sought and proved by several mathematicians and physicists [2,p.285], [15,p.333]. The general version of the EMI we have obtained continues this tradition.

From the first Lidskii inequality in (10) one gets, using standard arguments [2],

$$\|\text{Eig}^\perp A - \text{Eig}^\perp B\|_\Phi \leq \delta_\Phi(\exp A, \exp B) \quad (19)$$

for any two Hermitian matrices A, B . The inequality (15) is stronger than this. Here it may be appropriate to remark that non-Riemannian differential geometry involving explicit computation of geodesic length has been used earlier [4] in obtaining tight bounds for spectral variation of unitary matrices.

Let us put the generalized EMI in the context of geodesics in metric spaces [6]. A geodesic segment in a metric space X is a distance-preserving map from a compact interval into X . The space X is said to be a *geodesic space* if the distance between any two points is equal to the length of a geodesic segment joining them. We have shown that the space \mathbb{P} with the metric δ_Φ is a geodesic space. The curve e^{tA} is a geodesic segment joining I and e^A .

Some symmetric gauge functions Φ on \mathbb{R}^n have the property that all geodesic segments (in the metric induced by Φ) are straight lines. The Schatten p -norms have this property for $1 < p < \infty$. In these cases, for any two points x, y in \mathbb{R}^n their algebraic mid-point $z = (x + y)/2$ is the only *metric mid-point* (a point equidistant from x and y). The norms $\|\cdot\|_\Phi$ on \mathbb{M} inherit this property from Φ . For all such norms any two points in \mathbb{P} have a unique δ_Φ geodesic segment joining them. (Compare with Theorem 3.6 of Lang [10,p.313].)

There is a very interesting description of the mid-point on this segment.

The *geometric mean* of two elements A, B of \mathbb{P} is the matrix

$$A\#B = A^{\frac{1}{2}}(A^{-\frac{1}{2}}BA^{-\frac{1}{2}})^{\frac{1}{2}}A^{\frac{1}{2}}, \quad (20)$$

a definition introduced by Pusz and Woronowicz [14]. This object has many interesting properties, including symmetry not so obvious from the above definition. It is easy to see that

$$\{\lambda_i(AB^{-1})\} = \{\lambda_i^2(A(A\#B)^{-1})\}$$

Thus

$$\delta_\Phi(A, B) = 2\delta_\Phi(A, A\#B).$$

This shows that $A\#B$ is a metric midpoint between A and B for every metric δ_Φ . For certain metrics, such as δ_p induced by the Schatten norms for $1 < p < \infty$, the geometric mean is the unique midpoint between A and B .

For an interesting recent discussion of the geometric mean see [11].

References

1. W. Ballmann, M. Gromov and V. Schroeder, *Manifolds of Nonpositive Curvature*, Birkhäuser, 1995.
2. R. Bhatia, *Matrix Analysis*, Springer, 1997.
3. R. Bhatia, Linear algebra to quantum cohomology: the story of Alfred Horn's inequalities, *Amer. Math. Monthly*, 108 (2001) 289-318.
4. R. Bhatia and J.A.R. Holbrook, Short normal paths and spectral variation, *Proc. Amer. Math. Soc.*, 94 (1985) 377-382.
5. R. Bhatia and K. R. Parthasarathy, Positive definite functions and operator inequalities, *Bull. London Math. Soc.*, 32 (2000) 214-228.
6. M. R. Bridson and A. Haefliger, *Metric spaces of non-positive curvature*, Springer, 1999.
7. G. Corach, H. Porta, and L. Recht, Geodesics and operator means in the space of positive operators, *International J. Math.*, 4(1993)193-202.
8. W. Fulton, Eigenvalues, invariant factors, highest weights, and Schubert calculus, *Bull. Amer. Math. Soc.*, 37 (2000) 209-249.
9. F. Hiai and H. Kosaki, Comparison of various means for operators, *J. Funct. Anal.*, 163 (1999) 300-323.
10. S. Lang, *Fundamentals of Differential Geometry*, Springer, 1999.
11. J. D. Lawson and Y. Lim, The geometric mean, matrices, metrics, and more, *Amer. Math. Monthly*, 108 (2001) 797-812.
12. V. B. Lidskii, On the proper values of a sum and product of symmetric matrices, *Dokl. Akad. Nauk SSSR*, 75 (1950) 769-772.
13. G. D. Mostow, Some new decomposition theorems for semisimple groups, *Memoirs Amer. Math. Soc.*, 14 (1955) 31-54.
14. W. Pusz and S. L. Woronowicz, Functional calculus for sesquilinear forms and the purification map, *Rep. Math. Phys.*, 8 (1975) 159-170.
15. M. Reed and B. Simon, *Methods of Modern Mathematical Physics, Vol II*, Academic Press, 1975.