Interpolating the arithmetic–geometric mean inequality and its operator version☆

Rajendra Bhatia

Indian Statistical Institute, 7 S.J.S. Sansanwal Marg, New Delhi 110 016, India

Abstract

Two families of means (called Heinz means and Heron means) that interpolate between the geometric and the arithmetic mean are considered. Comparison inequalities between them are established. Operator versions of these inequalities are obtained. Failure of such extensions in some cases is illustrated by a simple example.

Keywords: Inequalities for means; Operator inequalities; Positive definite matrix; Unitarily invariant norm

1. Introduction

The arithmetic–geometric mean inequality

$$\sqrt{ab} \leqslant \frac{a+b}{2} \tag{1}$$

for positive numbers a, b, has been generalised, extended and strengthened in various directions. A matrix version proved in [2] says that if A, B and X are $n \times n$ matrices with A and B positive definite, then for every unitarily invariant norm $|||\cdot|||$

$$|||A^{1/2}XB^{1/2}||| \leqslant \frac{1}{2}|||AX + XB|||. \tag{2}$$

☆ Results presented in this paper formed a part of the author's talk at the International Linear Algebra Society Conference held at Coimbra in July, 2004. The conference was dedicated to Richard Brualdi.

E-mail address: rbh@isid.ac.in

There are several (parameterised families of) means that interpolate between the geometric and the arithmetic mean. One such family, that we call the *Heinz means* are defined as

$$H_\nu(a,b) = \frac{a^{1-\nu}b^\nu + a^\nu b^{1-\nu}}{2}, \tag{3}$$

$0 \leqslant \nu \leqslant 1$. For $\nu = 0, 1$ this is equal to the arithmetic mean, and for $\nu = 1/2$ to the geometric mean. It is easy to see that as a function of ν, $H_\nu(a,b)$ is convex and attains its minimum at $\nu = 1/2$. Thus

$$\sqrt{ab} \leqslant H_\nu(a,b) \leqslant \frac{a+b}{2}, \quad 0 \leqslant \nu \leqslant 1. \tag{4}$$

The matrix version proved in [2] says that if A, B, X are matrices with $A, B \geqslant 0$ (i.e. A, B are positive definite), then for every unitarily invariant norm the function

$$g(\nu) = |||A^\nu X B^{1-\nu} + A^{1-\nu} X B^\nu|||$$

is convex on [0, 1] and attains its minimum at $\nu = 1/2$. Thus we have for $0 \leqslant \nu \leqslant 1$

$$2|||A^{1/2} X B^{1/2}||| \leqslant |||A^{1-\nu} X B^\nu + A^\nu X B^{1-\nu}||| \leqslant |||AX + XB|||. \tag{5}$$

(The special case of (5) in which the norm is the operator bound norm is an old inequality of Heinz [5], who used it to derive several inequalities in the perturbation theory of operators. For this norm the inequality (2) was proved by McIntosh [16] who derived from it the Heinz inequality.)

The inequality (2) aroused much interest and several alternate proofs were given. Of these the one germane to our discussion occurs in the papers of Horn [10] and Mathias [15]. (Another interesting proof was given by Kittaneh [12].)

A familiar trick with 2×2 block matrices [1, p. 264] shows that inequalities like (2) and (5) follow from their special case with $A = B$. In this case we may assume (because of unitary invariance) that A is diagonal, $A = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$. Then the inequality (2) asserts that the norm of the matrix $[\sqrt{\lambda_i \lambda_j} x_{ij}]$ is not bigger than half the norm of the matrix $[(\lambda_i + \lambda_j) x_{ij}]$.

This can be interpreted in another way. Let $X \circ Y$ be the entrywise product of two matrices X and Y (also called the Schur product or the Hadamard product, this is the matrix with entries $x_{ij} y_{ij}$). Then the inequality (2) says that for all positive numbers $\lambda_1, \ldots, \lambda_n$ and for all X

$$\left|\left|\left| \left[\frac{2\sqrt{\lambda_i \lambda_j}}{\lambda_i + \lambda_j} \right] \circ X \right|\right|\right| \leqslant |||X|||. \tag{6}$$

Finding the "Schur multiplier norm" of a matrix is, in general, a hard problem. However, one special case is easy. If $Y \geqslant 0$, then for all X

$$|||Y \circ X||| \leqslant \max_i y_{ii} |||X|||. \tag{7}$$

See Theorems 5.5.18 and 5.5.19 in [11]. It is not very difficult to see that the matrix

$$Y = \left[\frac{2\sqrt{\lambda_i \lambda_j}}{\lambda_i + \lambda_j} \right] \tag{8}$$

is positive definite, and thus the inequality (6) follows from (7).

It is natural to ask whether stronger inequalities like (5) might be proved using this argument. This was a part of the motivation for our study [4]. The matrices whose positive definiteness is now to be established are more complicated than (8). For example, to prove the second inequality in (5) one needs to show that the matrix

$$Y = \left[\frac{\lambda_i^{1-\nu}\lambda_j^{\nu} + \lambda_i^{\nu}\lambda_j^{1-\nu}}{\lambda_i + \lambda_j} \right] \tag{9}$$

is positive definite for $0 \leqslant \nu \leqslant 1$. The main idea in [4] was to show that such matrices are congruent to others whose positivity follows from the positive definiteness of certain *functions*. A well-developed theory exists for the latter. This technique turned out to be very useful and was applied to many examples in [4].

Independently, and a little before the work in [4] was completed, Kosaki studied similar questions in [13] and developed a general technique to solve them. The scope of these methods has been explored in all kinds of directions in the subsequent work by Hiai and Kosaki [7,8] and expounded by them in a very interesting monograph [9].

One of the questions that arises from this work is the following. Does every inequality between means of positive numbers lead to a corresponding inequality for positive matrices? More precisely, let $M(a, b)$ be a mean on positive numbers (see Section 2 for a precise definition). If A is a positive definite matrix with eigenvalues $\lambda_1, \ldots, \lambda_n$, let $M(A, A)$ be the matrix whose ij entry is $M(\lambda_i, \lambda_j)$. Let M and L be two means such that

$$M(a, b) \leqslant L(a, b) \text{ for all } a, b. \tag{10}$$

Then must we have for all positive matrices A and for all X

$$|||M(A, A) \circ X||| \leqslant |||L(A, A) \circ X||| \tag{11}$$

for every unitarily invariant norm?

If the means M and L satisfy the condition (10) we say that $M \leqslant L$, and if they satisfy (11) we say that $M \ll L$. Our question is whether $M \leqslant L$ implies that $M \ll L$. There are many examples in [4,7,8,13] for which this is the case. However, Hiai and Kosaki [8] show that this need not always be true.

In this paper we study a simple class of means not included in the studies mentioned above. These means, that we call *Heron means*, are defined as

$$F_\alpha(a, b) = (1 - \alpha)\sqrt{ab} + \alpha \frac{a + b}{2}, \tag{12}$$

$0 \leqslant \alpha \leqslant 1$. This family is the linear interpolant between the geometric and the arithmetic mean. Perhaps because of its naivete, it has received less attention than other families of interpolating means. In our context it reveals some interesting phenomena. Clearly, $F_\alpha \leqslant F_\beta$ whenever $\alpha \leqslant \beta$. However, we will see that $F_\alpha \ll F_\beta$ only when $\beta \geqslant 1/2$. This gives a simpler, and more dramatic, example than the one in [8]. We prove other inequalities for this family, including comparisons with the Heinz means.

2. Heron means

A *mean* $M(a, b)$ is a positive continuous function on $(0, \infty) \times (0, \infty)$ that satisfies the following conditions:

(i) $M(a, b) = M(b, a)$,
(ii) $M(\alpha a, \alpha b) = \alpha M(a, b)$ for all $\alpha > 0$,
(iii) $M(a, b)$ is monotone increasing in a and b,
(iv) $\min(a, b) \leqslant M(a, b) \leqslant \max(a, b)$.

The geometric mean $G(a, b)$, the arithmetic mean $A(a, b)$, the Heinz means $H_\nu(a, b)$ and the Heron means $F_\alpha(a, b)$ defined by (3) and (12), respectively, all are examples of means.

The quantity

$$F_{2/3}(a, b) = \frac{a + b + \sqrt{ab}}{3} \tag{13}$$

is usually called the *Heronian mean*, and occurs in the formula for the volume of a frustum (a body obtained by slicing a pyramid, or a cone, by a plane parallel to its base). If the frustum has height h and its base and top have areas a and b, respectively, then its volume is

$$V = \frac{1}{3}h\left(a + b + \sqrt{ab}\right).$$

The quantity

$$F_{1/2}(a, b) = \frac{a + b + 2\sqrt{ab}}{4} = \left(\frac{a^{1/2} + b^{1/2}}{2}\right)^2 \tag{14}$$

is one of the family of *power means*, or the *binomial means* defined as

$$B_\alpha(a, b) = \left(\frac{a^\alpha + b^\alpha}{2}\right)^{1/\alpha}, \quad -\infty \leqslant \alpha \leqslant \infty. \tag{15}$$

Another mean of interest in geometry, statistics, and thermodynamics is the *logarithmic mean* defined as

$$L(a, b) = \frac{a - b}{\log a - \log b} = \int_0^1 a^t b^{1-t}\, \mathrm{d}t. \tag{16}$$

The inequality

$$G(a, b) \leqslant L(a, b) \leqslant A(a, b) \tag{17}$$

is well-known.

The next few statements give more comparisons between some of the means.

For $0 \leqslant \nu \leqslant 1$ let

$$\alpha(\nu) = 1 - 4(\nu - \nu^2). \tag{18}$$

This is a convex function, its minimum value is $\alpha(1/2) = 0$, and its maximum value is $\alpha(0) = \alpha(1) = 1$.

Lemma 1. *The Heinz and the Heron means satisfy the inequality*

$$H_\nu(a,b) \leqslant F_{\alpha(\nu)}(a,b). \tag{19}$$

for all $0 \leqslant \nu \leqslant 1$.

Proof. The inequality (19), in expanded form, says

$$\frac{a^{1-\nu}b^\nu + a^\nu b^{1-\nu}}{2} \leqslant 4(\nu-\nu^2)a^{1/2}b^{1/2} + \left(1-4(\nu-\nu^2)\right)\frac{a+b}{2}.$$

Put $a = e^x, b = e^y$. A small calculation reduces this inequality to

$$\cosh\left((1-2\nu)\left(\frac{x-y}{2}\right)\right) \leqslant 4(\nu-\nu^2) + \left(1-4(\nu-\nu^2)\right)\cosh\left(\frac{x-y}{2}\right).$$

Now put $\beta = 1-2\nu$. The inequality to be proved is

$$\cosh\beta x \leqslant (1-\beta^2) + \beta^2\cosh x \tag{20}$$

for all x and $-1 \leqslant \beta \leqslant 1$. The series expansion for $\cosh x$ reduces this to

$$\frac{\beta^4x^4}{4!} + \frac{\beta^6x^6}{6!} + \cdots \leqslant \beta^2\left(\frac{x^4}{4!} + \frac{x^6}{6!} + \cdots\right),$$

and this is plainly true. □

When $\nu = 0$ or 1, then $\alpha(\nu) = 1$ and the two sides of (19) are equal to the arithmetic mean of a and b. When $\nu = 1/2$, then $\alpha(\nu) = 0$ and the two sides of (19) are equal to the geometric mean of a and b.

There is no inequality reverse to (19) in the following sense: we cannot have (for all positive numbers a and b)

$$F_\alpha(a,b) \leqslant H_\nu(a,b)$$

for any pair of indices $0 < \alpha < 1$ and $0 < \nu < 1/2$. Arguing as above the validity of this inequality can be shown to be equivalent to that of the inequality

$$(1-\alpha) + \alpha\cosh x \leqslant \cosh(1-2\nu)x.$$

Another step shows this to be equivalent to

$$\alpha\sum_{m=1}^{\infty}\frac{x^{2m}}{(2m)!} \leqslant \sum_{m=1}^{\infty}(1-2\nu)^{2m}\frac{x^{2m}}{(2m)!}.$$

For $0 < \nu < 1/2$, the coefficients $(1-2\nu)^{2m}$ decrease to 0. So this last inequality is violated for some x, unless $\alpha = 0$.

Lemma 2. *The inequality*

$$L(a,b) \leqslant F_\alpha(a,b) \tag{21}$$

is true for all a and b if and only if $\alpha \geqslant 1/3$.

Proof. Again, a substitution $a = e^x$, $b = e^y$ shows that the inequality (21) is equivalent to

$$\frac{\sinh x}{x} \leqslant (1-\alpha) + \alpha \cosh x$$

and thence to

$$\frac{x^2}{3!} + \frac{x^4}{5!} + \cdots \leqslant \alpha \left(\frac{x^2}{2!} + \frac{x^4}{4!} + \cdots \right).$$

This is true for all x if and only if $\alpha \geqslant 1/3$. □

3. Norm inequalities

Let M and L be two means on $(0,\infty) \times (0,\infty)$. In Section 1 we observed that if for all n and for all positive numbers $\lambda_1, \ldots, \lambda_n$ the matrices

$$Y = \left[\frac{M(\lambda_i, \lambda_j)}{L(\lambda_i, \lambda_j)} \right] \tag{22}$$

are positive definite, then the inequality (11) is true for all positive matrices A and all X, and thus $M \ll L$. Hiai and Kosaki [8] have observed that the positivity of matrices (22) is also a necessary condition for $M \ll L$. To see this observe that the matrix Y in (22) is a Hermitian matrix with all its diagonal entries equal to 1. Let X be the matrix all whose entries are 1. Then the trace norm $\|X\|_1 = n$. (By definition $\|X\|_1$ is the sum of all singular values of X.) For this X, the inequality (11) tells us that $\|Y\|_1 \leqslant n$. Since $\operatorname{tr} Y = n$, this is possible only if all eigenvalues of Y are positive. In other words, Y is positive.

In all the statements below, A, B and X are matrices of any order n with A and B positive definite and $|||\cdot|||$ is any unitarily invariant norm.

Theorem 3. *Let* $0 \leqslant \alpha \leqslant \beta \leqslant 1$. *If* $\beta \geqslant 1/2$, *the inequality*

$$\begin{aligned} &\left|\left|\left|(1-\alpha)A^{1/2}XB^{1/2} + \alpha\left(\frac{AX+XB}{2}\right)\right|\right|\right| \\ &\quad \leqslant \left|\left|\left|(1-\beta)A^{1/2}XB^{1/2} + \beta\left(\frac{AX+XB}{2}\right)\right|\right|\right| \end{aligned} \tag{23}$$

is always true. This restriction on β *is necessary.*

Proof. The assertion of the theorem is that $F_\alpha \ll F_\beta$ under the given conditions. This is equivalent to the statement that for all n and positive numbers $\lambda_1, \dots, \lambda_n$ the matrix Y with entries

$$y_{ij} = \frac{F_\alpha(\lambda_i, \lambda_j)}{F_\beta(\lambda_i, \lambda_j)}$$

is positive definite. Put

$$a = \frac{2(1-\alpha)}{\alpha}, \quad b = \frac{2(1-\beta)}{\beta}. \tag{24}$$

Note that $\alpha \leqslant \beta$ if and only if $a \geqslant b$. A small calculation shows that the statement to be proved is equivalent to saying that for all $a \geqslant b$ the matrix V with entries

$$\begin{aligned} v_{ij} &= \frac{a\sqrt{\lambda_i\lambda_j} + (\lambda_i + \lambda_j)}{b\sqrt{\lambda_i\lambda_j} + (\lambda_i + \lambda_j)} \\ &= \frac{(a-b)\sqrt{\lambda_i\lambda_j}}{b\sqrt{\lambda_i\lambda_j} + (\lambda_i + \lambda_j)} + 1 \end{aligned}$$

is positive definite.

Making the substitution $\lambda_i = \mathrm{e}^{x_i}, \lambda_j = \mathrm{e}^{x_j}$ and arguing as in [4] one sees that such matrices V are positive definite if and only if the function

$$f(x) = \frac{a-b}{b + 2\cosh x} + 1 = g(x) + 1 \tag{25}$$

is a positive definite function whenever $a \geqslant b$.

Clearly f is positive definite whenever g is. The converse is also true. If f is positive definite, then by Bochner's Theorem there exists a finite positive measure μ on the real line such that $f = \hat{\mu}$, the Fourier transform of μ. (See [6] or [17].) Separate out the part of μ concentrated at 0; i.e., write μ as

$$\mu = \mu_1 + r\delta_0,$$

where δ_0 is the Dirac measure at 0, r is a nonnegative real number and μ_1 is a positive measure with zero mass at 0.

Since $g = f - 1$, we have $g = \hat{\nu}$, where

$$\nu = \mu - \delta_0 = \mu_1 + (r-1)\delta_0.$$

Since g is a rapidly decreasing C^∞ function, its Fourier transform measure ν cannot have a positive mass at 0. So ν is equal to the positive measure μ_1. Hence, by Bochner's Theorem, g is positive definite.

Thus the inequality (23) is always true if and only if the function g in (25) is positive definite. This function has been studied in [4, p. 225], and from the analysis there we conclude that g is positive definite if and only if $b \leqslant 2$. Using (24) this condition translates to $\beta \geqslant 1/2$. This proves the theorem. □

In a recent paper [3] a proof simpler than the ones known before [4,14] is given for the fact that the function g occurring in the proof above is positive definite if $-2 < b \leqslant 2$.

Corollary 4. *If* $1/2 \leqslant \alpha \leqslant 1$, *then*

$$\left|\left|\left|\int_0^1 A^t X B^{1-t}\,\mathrm{d}t\right|\right|\right| \leqslant \left|\left|\left|(1-\alpha)A^{1/2}XB^{1/2} + \alpha\left(\frac{AX+XB}{2}\right)\right|\right|\right|. \tag{26}$$

This restriction on α *is necessary.*

Proof. Hiai and Kosaki [8, p. 924] have proved (26) for $\alpha = 1/2$. The inequality (23) then shows that (26) is true for $1/2 \leqslant \alpha \leqslant 1$.

Further, we know that

$$|||A^{1/2}XB^{1/2}||| \leqslant \left|\left|\left|\int_0^1 A^t X B^{1-t}\,\mathrm{d}t\right|\right|\right| \leqslant \frac{1}{2}|||AX + XB|||. \tag{27}$$

See [4,7]. If $\alpha < 1/2$ the right-hand side of (26) does not always dominate $|||A^{1/2}XB^{1/2}|||$. So, in this case (26) is not always true. □

A few remarks are in order.

1. The inequality (26) is a strengthening of the second inequality in (27).
2. Compare Lemma 2 and Corollary 4. The first says $L \leqslant F_\alpha$ for $\alpha \geqslant 1/3$; the second says $L \ll F_\alpha$ only if $\alpha \geqslant 1/2$.
3. Hiai and Kosaki [8, p. 924] have shown that $L \ll F_{1/2} \ll F_{2/3}$.
4. If we use the first equality in (16) we can see, following the arguments of [8], that the statement $L \ll F_\alpha$ if and only if $\alpha \geqslant 1/2$ is equivalent to the statement that the function
$$f(x) = \frac{\sinh x}{x(\cosh x + a)} \tag{28}$$
is positive definite if and only if $0 \leqslant a \leqslant 1$.
5. Hiai and Kosaki [8, p. 924] have compared the logarithmic mean L with the binomial means B_α defined in (15). They observe that $L \leqslant B_\alpha$ for all $\alpha \geqslant 1/3$, but the domination $L \ll B_{1/3}$ is false while $L \ll B_{1/2}$ is true.

At this stage it is natural to raise the question of strong domination between the Heinz and the Heron means: do we have a good version of (19) with the order $\ll$ in place of $\leqslant$? To answer this we have to compare the two sides of (20) and to decide whether the function

$$f(x) = \frac{\cosh \beta x}{(1-\beta^2) + \beta^2 \cosh x}, \quad -1 \leqslant \beta \leqslant 1 \tag{29}$$

is positive definite. The referee of this paper has informed us that Kosaki has shown that this function is positive definite only in the trivial cases $\beta = 0$ or ± 1. This shows that $H_\nu \ll F_{\alpha(\nu)}$ only in the trivial cases $\nu = 0, 1/2$, or 1.

This gives one more example of a situation where the two orders between means are strikingly different.

Acknowledgement

The author is grateful to an anonymous referee who pointed out an egregious error in the first version of this paper, and informed us about H. Kosaki's preprint.

References

[1] R. Bhatia, Matrix Analysis, Springer, 1997.
[2] R. Bhatia, C. Davis, More matrix forms of the arithmetic–geometric mean inequality, SIAM J. Matrix Anal. Appl. 14 (1993) 132–136.
[3] R. Bhatia, D. Drissi, Generalized Lyapunov equations and positive definite functions, SIAM J. Matrix Anal. Appl., in press.
[4] R. Bhatia, K.R. Parthasarathy, Positive definite functions and operator inequalities, Bull. London Math. Soc. 32 (2000) 214–228.
[5] E. Heinz, Beiträge zur Störungstheorie der Spektralzerlengung, Math. Ann. 123 (1951) 415–438.
[6] H. Helson, Harmonic Analysis, second ed., Hindustan Book Agency, 1995.
[7] F. Hiai, H. Kosaki, Comparison of various means for operators, J. Funct. Anal. 163 (1999) 300–323.
[8] F. Hiai, H. Kosaki, Means of matrices and comparison of their norms, Indiana Univ. Math. J. 48 (1999) 899–936.
[9] F. Hiai, H. Kosaki, Means of Hilbert space operators, Lecture Notes in Mathematics 1820, Springer, 2003.
[10] R.A. Horn, Norm bounds for Hadamard products and an arithmetic–geometric mean inequality for unitarily invariant norms, Linear Algebra Appl. 223/224 (1995) 355–361.
[11] R.A. Horn, C.R. Johnson, Topics in Matrix Analysis, Cambridge Univ. Press, 1990.
[12] F. Kittaneh, A note on the arithmetic–geometric mean inequality for matrices, Linear Algebra Appl. 171 (1992) 1–8.
[13] H. Kosaki, Arithmetic–geometric mean and related inequalities for operators, J. Funct. Anal. 156 (1998) 429–451.
[14] M.K. Kwong, On the definiteness of the solutions of certain matrix equations, Linear Algebra Appl. 108 (1988) 177–197.
[15] R. Mathias, An arithmetic–geometric–harmonic mean inequality involving Hadamard products, Linear Algebra Appl. 184 (1993) 71–78.
[16] A. McIntosh, Heinz inequalities and perturbation of spectral families, Macquarie Math. Reports, 1979.
[17] M. Reed, B. Simon, Methods of Modern Mathematical Physics II: Fourier Analysis, Self-adjointness, Academic Press, 1975.