



A key inequality for functions of matrices

Rajendra Bhatia ^a, Chandler Davis ^{b,*}

^aIndian Statistical Institute, 7, S.J.S. Sansanwal Marg, New Delhi 110016, India

^bDepartment of Mathematics, University of Toronto, Toronto, Canada M5S 3G3

Received 4 February 2000; accepted 27 April 2000

Submitted by T. Ando

Abstract

An inequality is proved for convex functions applied to self-adjoint matrices. Several known inequalities are shown to be consequences, but properly weaker. © 2001 Elsevier Science Inc. All rights reserved.

Keywords: Functions of matrices; Positive linear maps

We derive several previously studied matrix inequalities from one; and that one turns out surprisingly (at least to us) to be simple, and to be sharp where most of its predecessors were not.

We deal with self-adjoint matrices or operators A , for which we assume $m \leq A \leq M$. Here m and M are real numbers, and the order between operators is that in which $A \leq B$ means $B - A$ positive (semi-)definite. The inequalities we consider are of the form

$$\Phi(f(A)) \leq \dots \tag{1}$$

Here f is a real function defined on $[m, M]$, $f(A)$ is defined by the usual functional calculus, and Φ is a unital positive linear mapping. That is, Φ takes the identity to the identity (perhaps of a different algebra of operators), and $B \geq 0$ implies $\Phi(B) \geq 0$. The right-hand member of (1) will involve m, M and ΦA , and of course it will be different for different f .

The assumptions just made will be in force throughout the paper. Typical of these unital positive mappings Φ is the mapping $\Phi B = \sum_{j=1}^m P_j B P_j$, where $P_j = P_j^*$

* Corresponding author.

E-mail addresses: rbh@isid.ac.in. (R. Bhatia), davis@math.toronto.edu (C. Davis).

and $P_j P_k = P_j \delta_{kj}$. In case $\sum_{j=1}^m P_j = E$ is the identity, this is called a “pinching”; if not, Φ must be thought of as mapping to the algebra of operators living on range (E) , so $\Phi(1) = E$.

Several inequalities of the form (1) have proved their worth, with f assumed convex. These are sometimes called “matrix converses to Jensen’s inequality” [12] and we pause to explain why. As Φ is a sort of averaging, for f convex we expect $\Phi(f(A))$ to be bounded below by $f(\Phi A)$, and a special case is known as Jensen’s inequality; thus (1) is, if not really converse, at any rate complementary to the expected inequality. Now

$$f(\Phi A) \leq \Phi(f(A)) \quad (2)$$

does indeed hold for these Φ , provided f is not merely convex but matrix-convex [5,6]. This means that for all matrices A, B and all $0 < \lambda < 1$, $f(\lambda A + (1 - \lambda)B) \leq \lambda f(A) + (1 - \lambda)f(B)$. For such f , then, adding (1) along with (2) means $\Phi(f(A))$ is bounded both above and below.

Many of the important f are matrix-convex and so satisfy (2), but some like $f(t) = e^t$ are not; still we will obtain the other bound (1). Functions that are not even convex can be brought in by comparing them to convex ones, as pointed out by Li and Mathias [8].

Our comprehensive result is Theorem 1. We state two known corollaries first.

Corollary 1.1.

$$\Phi(A^2) \leq (M + m)\Phi A - Mm.$$

Corollary 1.2. Assume $0 < m$. Then

$$\Phi(A^{-1}) \leq -\frac{1}{Mm}\Phi A + \frac{1}{M} + \frac{1}{m}.$$

Theorem 1. If f is convex, then letting L be the linear interpolant

$$L(t) = \frac{1}{M - m}(f(M)(t - m) + f(m)(M - t)), \quad (3)$$

we have

$$\Phi(f(A)) \leq L(\Phi A). \quad (4)$$

Proof. As f is convex and L linear on (m, M) , with agreement at the end points, $f(A) \leq L(A)$. Because Φ is positive, we deduce $\Phi(f(A)) \leq \Phi(L(A))$. But since L is linear and Φ unital, $\Phi(L(A)) = L(\Phi(A))$. \square

Theorem 2. Equality holds for every Φ in the conclusion of Theorem 1 if and only if either f is linear or $\text{spectrum}(A) \subseteq \{m, M\}$.

Proof. If f is linear, then $f = L$ and equality holds. If $\text{spectrum}(A) \subseteq \{m, M\}$, then for a spectral projector P , $A = mP + M(1 - P)$,

$$\begin{aligned} f(A) &= f(m)P + f(M)(1 - P) \\ &= L(m)P + L(M)(1 - P) = L(A), \end{aligned}$$

and equality holds.

In any other case, choose the spectral projector P of A belonging to an interval where A has spectrum (so P is non-zero) and $L(t) - f(t) \geq \delta$ for suitable $\delta > 0$. Choose Φ which does not annihilate P_j ; for example, any pinching. Then equality cannot hold in (4), for

$$\begin{aligned} L(\Phi A) - \Phi(f(A)) &= \Phi(L(A) - f(A)) \\ &\geq \Phi(P(L(A) - F(A))) \\ &\geq \Phi(P\delta) = \delta\Phi P \stackrel{>}{\neq} 0. \quad \square \end{aligned}$$

Now for the function $f(t) = t^2$, we get Corollary 1.1, for $L(t) = (M + m)t - Mm$ fits the prescription in Theorem 1. This is our result [4, Theorem 2]; it appeared different there because a further quadratic term $(\Phi A)^2$ was subtracted from both sides. By Theorem 2, the inequality is sharp.

There is more to say about the function $f(t) = t^{-1}$ (for $0 < m$). Corollary 1.2 is obtained using

$$L(t) = -\frac{1}{Mm}t + \frac{1}{M} + \frac{1}{m};$$

the result was in the proof of our [4, Theorem 3]. (We ought to have pointed out, and we point out now, that the reasoning was earlier known [9].) By Theorem 2, this inequality is sharp.

Corollary 1.2 implies [4, Theorem 3]

$$\Phi(A^{-1}) \leq \frac{(M + m)^2}{4Mm} (\Phi A)^{-1}. \tag{5}$$

(A survey of proofs of related inequalities is in [1].)

Or, Corollary 1.2 implies

$$\Phi(A^{-1}) \leq (\Phi A)^{-1} + \frac{(\sqrt{M} - \sqrt{m})^2}{Mm}; \tag{6}$$

this is the bound we would get by using an argument of Mond and Pečarić [12, Theorem 3], and it generalizes their inequality [11, (4)]. Indeed, Corollary 1.2 implies a continuum of inequalities (the parameter μ ranging over $[m, M]$):

$$\Phi(A^{-1}) \leq \frac{1}{Mm} \{ \mu^2 (\Phi A)^{-1} - 2\mu + M + m \}. \tag{7}$$

To prove (7) from Corollary 1.2, one simply calculates that

$$-\frac{1}{Mm}t + \frac{1}{M} + \frac{1}{m} \leq \frac{1}{Mm} \left\{ \frac{\mu^2}{t} - 2\mu + M + m \right\}.$$

The choice $\mu = (M + m)/2$ yields (5), and the choice $\mu = \sqrt{Mm}$ yields (6); neither of (5) and (6) implies the other. But (5)–(7) are not sharp; indeed, even for

$$A = \begin{bmatrix} M & 0 \\ 0 & m \end{bmatrix}$$

and Φ trivial, (7) does not become equality for any μ . These competitor inequalities are more natural, in that their right-hand member is in terms of an inverse, but they are weaker.

Like remarks apply to the bounds obtained by parallel reasoning for other convex f .

Corollary 1.3. *Under the conditions of Theorem 1,*

$$\Phi(f(A)) \leq f(\Phi A) + \max\{L(t) - f(t) : t \in [m, M]\}.$$

Proof. This can be proved as in [12, Theorem 3]: Let g denote the negative convex function $f - L$, with $g(m) = g(M) = 0$ and $\min\{g(t) : t \in [m, M]\} = -\beta$; we have to show that $\Phi(f(A)) - f(\Phi A) \leq \beta$. But $\Phi(f(A)) - f(\Phi A) = \Phi(g(A)) - g(\Phi A)$ because $g - f = L$ is linear, and $\Phi(g(A)) - g(\Phi A)$ is tractable: $g(A) \leq 0$, so $\Phi(g(A)) \leq 0$; and spectrum $(\Phi A) \subseteq [m, M]$, so $-g(\Phi A) \leq \beta$. \square

However, it is at least as edifying to see Corollary 1.3 as a corollary of Theorem 1. For that, we have to show that its bound is bigger than the right-hand member in (4) – that is, that

$$L(\Phi A) \leq f(\Phi A) + \max\{L(t) - f(t) : t \in [m, M]\}.$$

Well, of course; and this is one of the steps in the proof we just gave!

Corollary 1.3 in case $f(t) = t^2$ is the Popoviciu inequality (cf. [7] and [4, Corollary 1]). Corollary 1.3 in case $f(t) = t^{-1}$ is (6). All non-trivial f give non-sharp inequalities in Corollary 1.3 and must turn to Theorem 1 to get sharp ones.

After this work was completed, we learned of the earlier independent work [10], which uses linear interpolants in a similar way.

Acknowledgements

The second author thanks NSERC of Canada for funding and the Indian Statistical Institute for hosting his visit to Delhi where this work was completed.

References

- [1] G. Alpargu, The Kantorovich inequality, with some extensions and with some statistical applications. M.S. Thesis, McGill University, 1999.

- [2] T. Ando, Concavity of certain maps on positive definite matrices and applications to Hadamard products, *Linear Algebra Appl.* 26 (1979) 203–241.
- [3] R. Bhatia, *Matrix Analysis*, Springer, Berlin, 1997.
- [4] R. Bhatia, Ch. Davis, A better bound on the variance, *Amer. Math. Monthly* 107 (2000) 353–357.
- [5] M.-D. Choi, Some assorted inequalities for positive linear maps on C^* -algebras, *J. Operator Theory* 4 (1980) 271–285.
- [6] Ch. Davis, A Schwarz inequality for convex operator functions. *Proc. Amer. Math. Soc.* 8 (1957) 42–44.
- [7] S.T. Jensen, G.P.H. Styan, Some comments and a bibliography on the von Szökefalvi–Nagy–Popoviciu and Nair–Thomson inequalities with extensions and applications in statistics and matrix theory, in: T.M. Rassias (Ed.), *Classical Inequalities: Generalizations and Applications*, Kluwer, Dordrecht, to appear.
- [8] C.-K. Li, R. Mathias, Matrix inequalities involving a positive linear map, *Linear Multilinear Algebra* 41 (1996) 221–231.
- [9] A.W. Marshall, I. Olkin, Matrix versions of Cauchy and Kantorovich inequalities, *Aequationes Math.* 40 (1990) 89–93.
- [10] J. Mičić, J. Pečarić, Y. Seo, Complementary inequalities to inequalities of Jensen and Ando based on the Mond–Pečarić method, *Linear Algebra Appl.*, 318 (2000) 87–107.
- [11] B. Mond, J.E. Pečarić, A matrix version of the Ky Fan generalization of the Kantorovich inequality, *Linear Multilinear Algebra* 36 (1994) 217–221.
- [12] B. Mond, J.E. Pečarić, Matrix inequalities for convex functions, *J. Math. Anal. Appl.* 209 (1997) 147–153.