On singular value inequalities for matrix means

Raluca Dumitru a,*

aDepartment of Mathematics and Statistics, University of North Florida; Institute of Mathematics of the Romanian Academy, Bucharest, Romania

Rachel Levanger b

bDepartment of Mathematics, Rutgers University

Bogdan Visinescu c

cDepartment of Mathematics and Statistics, University of North Florida; Institute of Mathematics of the Romanian Academy, Bucharest, Romania

Abstract

For positive semidefinite $n \times n$ matrices $A$ and $B$, the singular value inequality $(2 + t)s_j(A^rB^{2-r} + A^{2-r}B^r) \leq 2s_j(A^2 + tAB + B^2)$ is shown to hold for $r = \frac{1}{2}, 1, \frac{3}{2}$ and all $-2 < t \leq 2$.

Key words: Singular values, Matrix means, Singular value inequalities

1991 MSC: 15A18, 15A42
1 Introduction

The arithmetic-geometric mean inequality for positive real numbers $a$ and $b$, $ab \leq \frac{(a^2 + b^2)}{2}$ has been proved by Bhatia and Kittaneh [5] to hold for singular values of arbitrary $n \times n$ matrices $A$ and $B$:

$$2s_j(AB^*) \leq s_j(A^*A + B^*B)$$

(1)

for all $j = 1, 2, ..., n$. Using an additional $n \times n$ matrix $X$, Bhatia and Davis [3] obtained the following operator norm inequality:

$$2\|AXB^*\| \leq \|A^*AX + XB^*B\|$$

for all unitarily invariant norms, an inequality which was generalized by X. Zhan [8] as stated below.

Let $A, B, X \in M_n(C)$ with $A$ and $B$ positive semidefinite, $1 \leq 2r \leq 3$, and $-2 < t \leq 2$. Then for any unitarily invariant norm,

$$(2 + t)\|A^rXB^2-r + A^{2-r}XB^r\| \leq 2\|A^2X + tAXB + XB^2\|.$$  

These inequalities have been proved to hold for $t = 0$ and all $0 \leq r \leq 2$, a result due to Audenaert [1]. In this paper, we prove the inequalities in equation (2) for $r = 1, 2, 3$ and all $-2 < t \leq 2$.

2 Preliminaries

With $A$ a positive semidefinite matrix (we use the standard notation $A \geq 0$), denote by $\lambda_j(A)$ and $s_j(A)$ its eigenvalues and singular values, respectively, arranged in non-increasing order. Denote by $u_j(A)$ the eigenvectors of $A$ corresponding to the eigenvalues $\lambda_j(A)$.

Let $M_0$ and $M_1$ be two positive semidefinite $n \times n$ matrices. Let $M(t) = M_0 + tM_1$, where $t$ is a scalar parameter. Then ( [7], Chapter 2, Section 1) the eigenvalues $\lambda_j(M(t))$, $j = 1, ..., s$, of $M(t)$ are branches of analytic functions of $t$ with only algebraic singularities and the number $s$ of eigenvalues of $M(t)$
is independent of $t$, with the exception of a finite number of points $t$, called exceptional points. More precisely, there are exactly two possibilities. If $s = n$ (when these analytic functions are all distinct), then $M(t)$ has simple spectrum for all non-exceptional points $t$. If, on the other hand, some of these analytic functions are identical, then $s < n$, and $M(t)$ is called permanently degenerate.

In conclusion, $\lambda_j(M(t))$ are everywhere continuous functions of $t$ and they are differentiable everywhere, except maybe at a finite number of points (the exceptional points).

Let us consider next the derivative with respect to $t$ of these functions, whenever the derivative exists.

In what follows, we will use the notations $u_j(t) = u_j(M(t))$ and $\lambda_j(t) = \lambda_j(M(t))$. For simplicity, we will look at the derivative of $\lambda_j(M(t))$ at $t = 0$. We will consider three cases.

Case 1. We assume that the eigenvalues of $M_0$ are all simple.

Then the eigenvalues of $M(t)$ are also simple for small enough values of $t$, say $t \in (-a, a)$, for some $a$. This follows from Weyl’s inequalities,

$$\lambda_j(M_0) + \lambda_n(tM_1) \leq \lambda_j(M_0 + tM_1) \leq \lambda_j(M_0) + \lambda_1(tM_1)$$

Hence, if $t \|M_1\|$ is small enough, namely strictly less than one half the minimum distance between all pairs of eigenvalues of $M_0$, then the eigenvalues of $M_0 + tM_1$ will be simple.

Therefore, on the interval $t \in (-a, a)$, every eigenvalue $\lambda_j(t)$ has a unique eigenvector $u_j(t)$, up to a constant multiple. Furthermore, since the zeros of the characteristic polynomial of $M(t)$ are simple (hence the polynomial has nonzero derivative at these zeros) then by applying the implicit function theorem, we get that the eigenvalues $\lambda_j(t)$ are smooth in $(-a, a)$. Since the eigenvectors $u_j(t)$ are determined up to a scalar by the equations $(M(t) - \lambda_j(t)I_n)u_j(t) = 0$ and $u_j(t)^*u_j(t) = 1$, the implicit function theorem allows us to locally select the functions $u_j(t)$ to also be smooth.

Let now $t \in (-a, a)$. By differentiating the equation

$$M(t)u_j(t) = \lambda_j(t)u_j(t)$$

we obtain the relation

$$M'(t)u_j(t) + M(t)u'_j(t) = \lambda'_j(t)u_j(t) + \lambda_j(t)u'_j(t) \quad (3)$$
and, by taking the inner product with $u_j(t)$ in equation (3), we obtain

$$u_j^*(t)M'(t)u_j(t) + u_j^*(t)M(t)u_j'(t) = \lambda_j(t)u_j^*(t)u_j(t) + \lambda_j(t)u_j^*(t)u_j'(t).$$

Since $M(t)$ is Hermitian, we have $u_j^*(t)M(t) = \lambda_j(t)u_j^*(t)$ and, using $u_j^*(t)u_j(t) = 1$, we get

$$\frac{d}{dt} \lambda_j(t) = u_j^*(t)M'(t)u_j(t).$$

(4)

In particular, the derivative at 0 of the eigenvalue function in the case when $M_0$ has simple eigenvalues is given by

$$\frac{d}{dt} \bigg|_{t=0} \lambda_j(t) = u_j^*(0)M_1u_j(0).$$

(5)

We assume next that $M_0$ has degenerate eigenvalues. Since the degeneracy of the eigenvalues of $M(t)$ is either accidental (for isolated values of $t$, such as $t = 0$ here) or permanent (for all values of $t$), there are two cases to consider.

We will first consider the case when $M_1$ is such that it removes the degeneracy of $M_0$ for small enough values of $t$, so $M(t)$ has an exceptional point at $t = 0$. The second case will be when $M(t)$ is permanently degenerate.

The problem is, in both these two cases, that the eigenvectors $u_j(t)$ are no longer unique.

Case 2. Assume that $t = 0$ is an exceptional point.

Then for $t$ small enough, $t \in (-a, a) \setminus 0$, the eigenvalues of $M(t)$ are simple, and therefore the corresponding eigenvectors are unique. Hence $\lambda_j(t)$ is differentiable for all values of $t$ in $(-a, a) \setminus 0$ and equation (4) above still holds. Note that $\lambda_j(t)$ might not be differentiable at $t = 0$, but it is continuous.

Case 3. Assume that $M(t)$ is permanently degenerate.

Let now $a$ be the largest positive value for which $M(t)$ has no exceptional point in the interval $(-a, a)$. Let $\lambda_j(t)$ be an eigenvalue function such that $\lambda_j(0)$ has multiplicity $m$.

Using [7](Chapter 2, equations 2.3, 2.21 and 2.34), we get

$$\frac{d}{dt} \lambda_j(t) = \sum_{n=1}^{\infty} nt^{n-1} \frac{1}{mn} Tr(M'(0)P_j^{(n-1)}),$$

(6)
where \( M'(0) = \frac{d}{dt} M(t) \bigg|_{t=0} \), \( M(t) = M_1 \) and \( P_j(t) = \sum_{n=0}^{\infty} t^n P_j^{(n)} \) denotes the projection onto the eigenspace generated by \( \lambda_j(t) \).

Note that, in this setting, there is no splitting of \( \lambda \) in the interval \((-a, a)\), so that the \( \lambda \)-group consists of a single eigenvalue of multiplicity \( m \). Hence the weighted mean of the \( \lambda \)-group, \( \hat{\lambda}_j(t) \) (used in equations 2.21 and 2.34) is the same as \( \lambda_j(t) \).

Therefore,

\[
\frac{d}{dt} \lambda_j(t) = \frac{1}{m} \sum_{n=1}^{\infty} Tr(M'(0)t^{n-1}P_j^{(n-1)}) = \frac{1}{m} Tr(M'(0)P_j(t)),
\]

for all \( t \) in the interval \((-a, a)\).

3 The Behavior of an Eigenvalue Function

**Theorem 3.1** For \( A, B \in M_n(C) \), \( A, B \geq 0 \), \( j = 1, 2, ..., n \) and \( t \in (-2, \infty) \), the function

\[
f(t) = \frac{1}{2 + t} \lambda_j(A^2 + B^2 + \frac{t}{2}AB + \frac{t}{2}BA)
\]

is non-increasing.

**Proof.** Let \( A, B \geq 0 \). Note that \( f(t) \) is continuous everywhere on \((-2, \infty)\) and it is differentiable everywhere except maybe at a finite number of points (the exceptional points). Let \( t_0 \in (-2, \infty) \). We prove next that there exists an interval centered at \( t_0 \) where the function \( f \) is non-increasing.

Let \( M(t) = A^2 + B^2 + \frac{t}{2}AB + \frac{t}{2}BA \). We will consider three cases, according to whether \( M(t) \) is (permanently) degenerate or not.

**Case 1.** Assume that all eigenvalues of \( M(t_0) \) are simple. Then the eigenvalues of \( M(t) \) are simple for values of \( t \) in a small enough neighborhood of \( t_0 \), say \( t \in (t_0 - a, t_0 + a) \).

Then \( f \) is differentiable everywhere on \((t_0 - a, t_0 + a)\) and using equation 4 above, we obtain
\[ f'(t) = \frac{d}{dt} \lambda_j(t)(2 + t) - \lambda_j(t) \]
\[ = \frac{u_j^*(t)AB + BA}{2}u_j(t)(2 + t) - \lambda_j(t) \]
\[ = \frac{u_j^*(t)(A - B)^2u_j(t)}{(2 + t)^2} \]
\[
= - \frac{u_j^*(t)(A - B)^2u_j(t)}{(2 + t)^2},
\]

for all \( t \in (t_0 - a, t_0 + a) \).

Since \((A - B)^2 \geq 0\), then for all \( u \in C^n \) we have \( \langle (A - B)^2u, u \rangle \geq 0 \).

In particular, this implies that \( u_j^*(t)(A - B)^2u_j(t) \geq 0 \), and therefore \( f'(t) \leq 0 \), so \( f \) is non-increasing on \((t_0 - a, t_0 + a)\).

Case 2. Assume that \( t_0 \) is an exceptional point for \( M(t) \), hence \( f \) might not be differentiable at \( t_0 \).

There is however a small interval centered at \( t_0 \), say \((t_0 - a, t_0 + a)\), such that \( f \) is differentiable everywhere on \((t_0 - a, t_0 + a)\)\( \{t_0\} \). Then the derivative of \( f \) will be computed as in equation 8 and hence \( f'(t) \leq 0 \) on \((t_0 - a, t_0 + a)\)\( \{t_0\} \). Since \( f \) is continuous everywhere, we conclude that \( f \) is non-increasing on \((t_0 - a, t_0 + a)\).

Case 3. Assume that \( M(t) \) is permanently degenerate and let \( a \) be the largest positive value for which \( M(t) \) has no exceptional point in the interval \((t_0 - a, t_0 + a)\). Let \( \lambda_j(t) \) be an eigenvalue function such that \( \lambda_j(t_0) \) has multiplicity \( m \).

Then \( f(t) \) is everywhere differentiable on \((t_0 - a, t_0 + a)\), however its derivative cannot be computed in the same way as in equation 8 since the corresponding eigenvectors are not unique anymore.
Using equation [7] we obtain
\[
f'(t) = \frac{(2 + t) \frac{d}{dt} \lambda_j(t) - \lambda_j(t)}{(2 + t)^2} - \frac{1}{m} Tr(M(t)P_j(t))
\]
\[
= \frac{1}{m} Tr(((2 + t)M'(0) - M(t))P_j(t))
\]
\[
= \frac{-1}{m} Tr((A - B)^2 P_j(t))
\]
Hence \( f'(t) \leq 0 \) on \((t_0-a, t_0+a)\) and we conclude again that \( f \) is non-increasing on \((t_0 - a, t_0 + a)\). □

Using Theorem 3.1 for \( t \in (-2, 2] \), we obtain in particular the following corollary.

**Corollary 3.2** For \( A, B \in M_n(C) \), \( A, B \geq 0 \), \( j = 1, 2, \ldots, n \) and \( t \in (-2, 2] \), we have

(1) \( A^2 + B^2 + \frac{t}{2} AB + \frac{t}{2} BA \geq 0 \) and

(2) \( \frac{1}{2 + t} s_j(A^2 + B^2 + \frac{t}{2} AB + \frac{t}{2} BA) \geq \frac{1}{4} s_j(A + B)^2 \).

**4 On Zhan’s Conjecture**

We prove first that X. Zhan’s conjecture (equation [2]) holds for \( r = \frac{1}{2}, \frac{3}{2} \) and all \(-2 < t \leq 2\).

**Proposition 4.1** For \( A, B \in M_n(C) \), \( A, B \geq 0 \), \( j = 1, 2, \ldots, n \) and \( t \in (-2, 2] \), we have

\[
(2 + t)s_j(A^\frac{1}{2} B^\frac{3}{2} + A^\frac{3}{2} B^\frac{1}{2}) \leq 2s_j(A^2 + tAB + B^2).
\]

**PROOF.** Since \( A^2 + B^2 + \frac{t}{2} AB + \frac{t}{2} BA = Re(A^2 + tAB + B^2) \), by using [2] Proposition III.5.1 we get

\[
\lambda_j(A^2 + B^2 + \frac{t}{2} AB + \frac{t}{2} BA) \leq s_j(A^2 + tAB + B^2)
\]

which, by Corollary 3.2 (1), is the same as

\[
s_j(A^2 + B^2 + \frac{t}{2} AB + \frac{t}{2} BA) \leq s_j(A^2 + tAB + B^2).
\]

(9)
Using Corollary 3.2 (2), we conclude that
\[
\frac{2 + t}{4} s_j((A + B)^2) \leq s_j(A^2 + tAB + B^2).
\] (10)

Bhatia and Kittaneh [4] proved that
\[
2s_j(A^{1/2}B^{3/2} + A^{3/2}B^{1/2}) \leq s_j((A + B)^2),
\]
which, combined with equation (10), proves the desired result. □

Our next result shows that Zhan’s conjecture (equation (2)) holds for \( r = 1 \) and all \(-2 < t \leq 2\).

**Proposition 4.2** For \( A, B \in M_n(C) \), \( A, B \geq 0 \), \( j = 1, 2, \ldots, n \) and \( t \in (-2, 2] \), we have
\[
(2 + t)s_j(AB) \leq s_j(A^2 + tAB + B^2).
\]

**Proof.** Note that \( 4s_j(AB) \leq s_j(A + B)^2 \), an inequality recently proved by Drury in [6]. Using this inequality together with Corollary 3.2 (2) and Equation (9) we obtain
\[
(2 + t)s_j(AB) \leq \frac{2 + t}{4} s_j(A + B)^2
\]
\[
\leq s_j(A^2 + B^2 + \frac{t}{2}AB + \frac{t}{2}BA)
\]
\[
\leq s_j(A^2 + tAB + B^2).
\]
□

**References**

[1] K. Audenaert, *A singular value inequality for Heinz means*, Linear Algebra Appl., 422(2007), 279-283.

[2] R. Bhatia, *Matrix Analysis*, Springer, 1996.

[3] R. Bhatia, C. Davis, *More matrix forms of the arithmetic-geometric mean inequality*, SIAM J. Matrix Anal. Appl., 14(1993), 132-136.

[4] R. Bhatia, F. Kittaneh, *Notes on matrix arithmetic-geometric mean inequality*, Linear Algebra Appl., 308(2000), 203-211.
[5] R. Bhatia, F. Kittaneh, *On the singular values of a product of operators*, SIAM J. Matrix Anal. Appl., 11(1990), 272-277.

[6] S. W. Drury, *On a question of Bhatia and Kittaneh*, Linear Algebra Appl., 437(2012), 1955-1960.

[7] T. Kato, *Perturbation Theory for linear operators*, Springer-Verlag, Berlin, Heidelberg, New York, 1980.

[8] X. Zhan, *Inequalities for unitarily invariant norms*, SIAM J. Matrix Anal. Appl., 20(1999), 466-470.

[9] X. Zhan, *Some research problems on the Hadamard product and singular values of matrices*, Linear and Multilinear Algebra, Vol. 47, 191-194.