On a question of Bhatia and Kittaneh

S.W. Drury

Department of Mathematics and Statistics, McGill University, Burnside Hall, 805 ouest, rue Sherbrooke Street West, Montréal, Québec, Canada H3A 0B9

Abstract

We settle in the affirmative a question of Bhatia and Kittaneh. For $P$ and $Q$ positive semidefinite $n \times n$ matrices, the inequality

$$\sqrt{\sigma_r(PQ)} \leq \frac{1}{2} \lambda_r(P + Q)$$

holds for $r = 1, 2, \ldots, n.$

© 2012 Elsevier Inc. All rights reserved.

1. Introduction

There is an enormous literature on generalizations of the two variable arithmetic geometric mean inequality

$$\sqrt{ab} \leq \frac{1}{2} (a + b)$$

for $a, b \geq 0$ to the matrix setting, much of it associated with Bhatia and Kittaneh. In their 1990 paper [2], they establish the inequality

$$\sigma_j(A^*B) \leq \frac{1}{2} \lambda_j(AA^* + BB^*)$$

© 2012 Elsevier Inc. All rights reserved.

0024-3795/$ - see front matter © 2012 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.laa.2012.04.040
for general compact operators $A$ and $B$ on a Hilbert space. For a matrix $M$ or indeed a compact operator on Hilbert space, $\sigma_k(M)$ denotes the $k$th singular value of $M$ and for a hermitian matrix or hermitian compact operator $H$, $\lambda_k(H)$ denotes the $k$th eigenvalue both with the decreasing ordering. They studied many possible variations in 2000 [3] and revisited the subject in 2008 [4]. In that paper, they put a lot of emphasis on what they describe as level three inequalities. The following theorem settles positively their key question in this area.

**Theorem 1.** For $P$ and $Q$ positive semidefinite $n \times n$ matrices, the inequality

$$\sqrt{\sigma_r(PQ)} \leq \frac{1}{2} \lambda_r(P + Q)$$

holds for $r = 1, 2, \ldots, n$.

The analogous result is also true in the compact operator setting but we leave the verification of this to the interested reader.

Bhatia and Kittaneh [4] point out that a consequence of (1) is that

$$||AB||^\frac{1}{2} \leq \frac{1}{2} ||A + B||$$

holds for all unitarily invariant norms $|| \cdot ||$ and all $A$ and $B$ positive semidefinite.

In [3] they have already established (1) in the cases $r = 1$ and $r = n$.

### 2. An eigenvalue estimate for certain block matrices

Before attempting the proof of Theorem 1 we need to develop some preliminary material.

Let $B$ and $X$ be two positive definite matrices of the same size. Then the geometric mean of $B$ and $X$ denoted by $B\#X$ is the unique positive definite matrix such that $B = (B\#X)X^{-\frac{1}{2}}(B\#X)$ or equivalently $X = (B\#X)B^{-\frac{1}{2}}(B\#X)$. The geometric mean is given by the explicit formula $B\#X = B^{\frac{1}{2}}(B^{-\frac{1}{2}}XB^{-\frac{1}{2}})^{\frac{1}{2}}B^{\frac{1}{2}}$ or equivalently $B\#X = X^{\frac{1}{2}}(X^{-\frac{1}{2}}XB^{-\frac{1}{2}})X^{\frac{1}{2}}$. It is symmetric, i.e. $B\#X = X\#B$. We refer the reader to Bhatia’s book [1] for all details relating to the theory of this topic.

**Proposition 2.** Let $B$ and $X$ be positive definite $r \times r$ matrices. Let

$$R = \begin{pmatrix} B & (B\#X)^{-1} \\ (B\#X)^{-1} & X \end{pmatrix}$$

a $2r \times 2r$ matrix. Then $\lambda_r(R) \geq 2$.

**Proof.** For short, let $S = B\#X$, then it is well-known and easy to check that

$$R_1 = \begin{pmatrix} B & -S \\ -S & X \end{pmatrix}$$

is positive semidefinite. In fact, this matrix has rank $r$. Then

$$R - R_1 = \begin{pmatrix} 0 & S + S^{-1} \\ S + S^{-1} & 0 \end{pmatrix}.$$

The eigenvalues of $S + S^{-1}$ are clearly all $\geq 2$ since $S$ is positive definite. The eigenvalues of $R - R_1$ are the eigenvalues of $S + S^{-1}$ and their negatives. Hence $R - R_1$ has exactly $r$ eigenvalues $\geq 2$. It follows that $R$ has at least $r$ eigenvalues $\geq 2$. □
3. Calculations with characteristic polynomials

The following proposition is well-known and easily proved using Schur complements.

**Proposition 3.** Let $M_{11}, M_{12}, M_{21}$ and $M_{22}$ be $r \times r$ matrices and assume that $M_{12}$ is invertible. Let

$$M = \begin{pmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{pmatrix}.$$ 

Then $\det(M) = \det(M_{12}M_{22}M_{12}^{-1}M_{11} - M_{12}M_{21})$.

Its corollary is

**Corollary 4.** Let $M_{11}, M_{12}, M_{21}, M_{22}$ and $M$ be as above. Then

$$\det(\lambda I - M) = \det\left(\lambda^2 I - \lambda(M_{11} + M_{12}M_{22}M_{12}^{-1}) + (M_{12}M_{22}M_{12}^{-1}M_{11} - M_{12}M_{21})\right).$$

We can now prove the following proposition, whose significance will only be apparent later.

**Proposition 5.** Let $A$ and $B$ be $r \times r$ positive definite matrices, and let $Z$ be an $r \times r$ matrix such that $BA(I + ZZ^*)AB = I$. Let

$$T = \begin{pmatrix} A + B & AZ \\ Z^*A & Z^*AZ \end{pmatrix}.$$  

Then $\det(\lambda I - T) = \det\left(\lambda^2 I - \lambda(B + B^{-1}A^{-1}B^{-1}) + (B^{-1}A^{-1} - BA)\right)$.

**Proof.** Note that the hypotheses imply that $BA^2B \leq I$ or equivalently that $A^{-1}B^{-2}A^{-1} \geq I$.

We first assume that $Z$ is nonsingular. Then we will apply Corollary 4 with $M = T$. We note that

$$M_{12}M_{22}M_{12}^{-1} = AZZ^*AZ(AZ)^{-1} = AZ^* = B^{-2}A^{-1} - A.$$ 

Then

$$\det(\lambda I - T) = \det\left(\lambda^2 I - \lambda(A + B + AZ^*) + (AZ^*(A + B) - AZ^*A)\right)$$

$$= \det\left(\lambda^2 I - \lambda(A + B + AZ^*) + AZ^*B\right)$$

$$= \det\left(\lambda^2 I - \lambda(B + B^{-2}A^{-1}) + (B^{-2}A^{-1}B - AB)\right)$$

$$= \det\left(\lambda^2 I - \lambda(B + B^{-1}A^{-1}B^{-1}) + (B^{-1}A^{-1} - BA)\right)$$

using a similarity at the last step. This completes the proof in the case that $Z$ is nonsingular.

For the general case, we first observe that without loss of generality $Z$ may be replaced by the positive semidefinite matrix $W = (A^{-1}B^{-2}A^{-1} - I)^{1/2}$. In fact, the polar decomposition of $Z$ is $Z = WU$ where $U$ is a unitary. It is now easy to check that

$$T = \begin{pmatrix} A + B & AW \\ U^*W^*A & U^*W^*AW \end{pmatrix}$$

and

$$A + B \begin{pmatrix} AW \\ W^*A & W^*AW \end{pmatrix}$$

are unitarily similar and therefore have the same characteristic polynomial.

Next, we approximate $B$ by $B_k = \mu_k B$ where $0 < \mu_k < 1$ and where $\mu_k$ increases to 1. The corresponding $W_k = (A^{-1}B_k^{-2}A^{-1} - I)^{1/2}$ is now definitely invertible. One applies the previous argument to the approximating sequence to obtain
\[ \det(\lambda I - T_k) = \det\left(\lambda^2 I - \lambda(B_k + B_k^{-1}A^{-1}B_k^{-1}) + (B_k^{-1}A^{-1} - B_k A)\right) \] (3)

where

\[ T_k = \begin{pmatrix} A + B & AW_k \\ W_k^*A & W_k^*AW_k \end{pmatrix} \]

Finally one passes to the limit as \( k \to \infty \) on both sides of (3) to obtain the desired result. □

Again, using Corollary 4 we have

**Proposition 6.** Let \( B \) and \( S \) be \( r \times r \) positive definite matrices. Let \( X = SB^{-1}S \) (so that \( S = B^hX \)) and

\[ R = \begin{pmatrix} B & S^{-1} \\ S^{-1} & X \end{pmatrix}. \]

Then \( \det(\lambda I - R) = \det\left(\lambda^2 I - \lambda(B + B^{-1}S^2) + (B^{-1}S^2B - S^{-2})\right) \).

The proof is left to the reader.

**Theorem 7.** Let \( A \) and \( B \) be \( r \times r \) positive definite matrices, and let \( Z \) be an \( r \times r \) matrix such that \( BA(I + ZZ^*)AB = I \). Then

\[ \lambda_r \begin{pmatrix} A + B & AZ \\ Z^*A & Z^*AZ \end{pmatrix} \geq 2. \]

**Proof.** Let \( S = (B^{-\frac{1}{2}}A^{-\frac{1}{2}}B^{-\frac{1}{2}}) \frac{1}{2} \) and let \( X = SB^{-1}S \) as in Proposition 6 and \( R \) also as defined there. The matrix \( T \) is as defined in Proposition 5. Then

\[
\begin{align*}
\det(\lambda I - R) &= \det\left(\lambda^2 I - \lambda(B + B^{-\frac{1}{2}}A^{-\frac{1}{2}}B^{-\frac{1}{2}}) + (B^{-\frac{1}{2}}A^{-\frac{1}{2}}B^{-\frac{1}{2}}B - B^{\frac{1}{2}}A^{-\frac{1}{2}}B^{\frac{1}{2}})\right) \\
&= \det\left(\lambda^2 I - \lambda(B + B^{-1}A^{-1}B^{-1}) + (B^{-1}A^{-1} - BA)\right) \\
&= \det(\lambda I - T)
\end{align*}
\]

using a similarity to get from (4) to (5). So \( R \) and \( T \) have the same eigenvalues. It now follows from Proposition 2 that \( \lambda_r(T) \geq 2. \) □

It will not escape the reader that since in the above proof, \( R \) and \( T \) are hermitian matrices with the same eigenvalues, then they must be unitarily similar. However this does not appear to be easy to prove directly.

### 4. Resolution of the question

**Proof of Theorem 1.** First of all, we may assume without loss of generality that \( P \) is positive definite, since the general case can be obtained by approximating with such matrices. Let us fix \( r \) in the range \( 1 \leq r \leq n \) and normalize so that \( \sigma_r(PQ) = 1 \). Our objective is then to show that \( \lambda_r(P + Q) \geq 2 \). We restate \( \sigma_r(PQ) = 1 \) as \( \lambda_r(PQ^2P) = 1 \). Let the spectral decomposition of \( PQ^2P \) be given by

\[ PQ^2P = \sum_{k=1}^n \lambda_k(PQ^2P)e_k \otimes e_k^* \]
where \((e_k)\) is an orthonormal basis. Then \(\lambda_k(PQ^2P) \geq 1\) for \(k = 1, 2, \ldots, r\) and \(\lambda_k(PQ^2P) \leq 1\) for \(k = r, \ldots, n\). We now define a positive semidefinite matrix \(Q_1\) by

\[
Q_1 = \left( P^{-1} \left( \sum_{k=1}^{r} e_k \otimes e_k^* \right) P^{-1} \right)^{1/2}.
\]

It follows that \(Q_1^2 \leq Q^2\). Next we use the fact that the square root is a matrix monotone function to assert that \(Q_1 \leq Q\). This is a special case of the Löwner–Heinz inequality see Zhan [6, Theorem 1.1] or Donoghue [5] for matrix monotonicity issues. It follows that \(P + Q_1 \leq P + Q\) and if the statement \(\lambda_r(P + Q_1) \geq 2\) is true then a fortiori \(\lambda_r(P + Q) \geq 2\). Hence we may always assume without loss of generality that \(PQ^2P\) is an orthogonal projection of rank \(r\).

Since \(P\) is assumed invertible, we conclude that \(Q\) has rank \(r\). Splitting the underlying ambient space as the direct sum of the image and kernel of \(Q\), we can after applying a unitary similarity assume that

\[
Q = \begin{pmatrix} Q_{11} & 0 \\ 0 & 0 \end{pmatrix} \quad \text{and} \quad P = \begin{pmatrix} P_{11} & P_{12} \\ P_{12}^* & P_{22} \end{pmatrix}
\]

in block matrix form. Here the diagonal blocks are square, the first being of size \(r\) and the second of size \(n - r\). Note that \(P_{11}\) is necessarily invertible. Now since \(PQ^2P\) is an orthogonal projection of rank \(r\), the same is true of \(QP^2Q\) and we see that

\[
Q_{11}(P_{11}^2 + P_{12}P_{12}^*)Q_{11} = I.
\]

Now let

\[
P_1 = \begin{pmatrix} P_{11} & P_{12} \\ P_{12}^* & P_{12}^*P_{11}^{-1}P_{12} \end{pmatrix}
\]

Then \(P_1\) has rank \(r\), satisfies both \(P_1 \leq P\) and \(PQ^2Q = QP^2Q\), so that \(\lambda_r(P + Q) \geq \lambda_r(P_1 + Q)\). Hence, we can and do assume that \(P_{22} = P_{12}^*P_{11}^{-1}P_{12}\) at the expense of no longer being able to assert that \(P\) is necessarily invertible.

We now wish to obtain matrices \(A, B\) and \(Z\) for which Theorem 7 can be applied. The procedure depends on the relative sizes of \(n\) and \(r\).

- If \(n = 2r\), we set \(A = P_{11}, B = Q_{11}\) and \(Z = P_{11}^{-1}P_{12}\).
- If \(n < 2r\), we set \(A = P_{11}, B = Q_{11}\) and \(Z\) to be the matrix obtained by appending \(2r - n\) zero columns to \(P_{11}^{-1}P_{12}\). We set \(P\) and \(Q\) to be the matrices obtained by appending \(2r - n\) zero rows and \(2r - n\) zero columns to \(P\) and \(Q\) respectively. The eigenvalues of \(P + Q\) are then seen to be those of \(P + Q\) but with \(2r - n\) zeros appended. Thus \(\lambda_r(P + Q) = \lambda_r(P + Q)\) and \(\bar{P} + \bar{Q}\) has the desired form for \(T\) as in (2).
- If \(n > 2r\), then \(r < n - r\) and \(\text{rank}(P_{12}) \leq r\). Therefore, there exists a \((n - r) \times (n - r)\) unitary matrix \(U\) such that \(P_{12}U\) has its last \(n - 2r\) columns zero. Then the matrices \(Q\) and \(P\) are simultaneously unitarily similar to the matrices \(Q\) and

\[
\bar{P} = \begin{pmatrix} P_{11} & P_{12}U \\ U^*P_{12}^* & U^*P_{12}^*P_{11}^{-1}P_{12}U \end{pmatrix}
\]

respectively. Therefore the matrices \(P + Q\) and \(\bar{P} + \bar{Q}\) have the same eigenvalues. The matrices \(\bar{P}\) and \(\bar{Q}\) have their last \(n - 2r\) rows and columns zero and we define \(\bar{P}\) and \(\bar{Q}\) to be the corresponding matrices with these rows and columns deleted. We set \(A = P_{11}\) and \(B = Q_{11}\). The matrix \(Z\) is taken to be the matrix \(P_{11}^{-1}P_{12}U\) but with the last \(n - 2r\) columns deleted. We observe that the eigenvalues of \(P + Q\) are those of \(\bar{P} + \bar{Q}\) but with \(n - 2r\) zeros appended. In particular \(\lambda_r(P + Q) = \lambda_r(\bar{P} + \bar{Q})\). Furthermore \(\bar{P} + \bar{Q}\) has the required form for \(T\) as in (2).
We leave the reader to check (using (6)) that in all cases, the hypotheses of Theorem 7 are satisfied. Applying Theorem 7 yields \( \lambda_r(P + Q) = \lambda_r(T) \geq 2 \) as required. □

Acknowledgement

I would like to thank the referees for their helpful comments.

References