Enhancements to the von Neumann trace inequality

N. Komaroff

School of Information Technology and Electrical Engineering, The University of Queensland, QLD 4072, Australia

Received 9 February 2007; accepted 30 July 2007
Available online 1 October 2007
Submitted by R.A. Brualdi

Abstract

Upper trace bounds for the product of two \( n \times n \) complex matrices are presented. The real component of the trace inequality is tighter than von Neumann’s inequality, and the imaginary component is new. © 2007 Elsevier Inc. All rights reserved.

AMS classification: 15A18; 15A42

Keywords: Eigenvalues; Singular values; Trace; Inequality

1. Introduction

Let \( X \) be an \( n \times n \) complex matrix. Denote the eigenvalues of \( X \) by \( \lambda_i(X) \), and the singular values of \( X \) by \( \sigma_i(X) \), \( i = 1, 2, \ldots, n \). All real numbers are arranged in decreasing order.

Let \( X, Y \) be \( n \times n \) complex matrices, and let \( V, W \) be \( n \times n \) unitary matrices. von Neumann’s 1937 trace inequality for singular values [1,3, p. 514] states that

\[
\text{Re} \ tr(V X W Y) \leq \sum_{i=1}^{n} \sigma_i(X) \sigma_i(Y),
\]

where \( \text{Re} \) denotes the real part, and \( tr \) denotes the trace.

The purpose of this paper is to present an inequality that is tighter than (1) and to show its complement, an upper bound for \( \text{Im} \ tr(X Y) \), where \( \text{Im} \) indicates the imaginary part of the trace.

The derivations are based on the singular value decompositions of matrices \( X \) and \( Y \).
2. Preliminaries

For a complex number \( x = a + \sqrt{-1}b \), where \( a, b \) are real numbers, we write \( \text{Re} \, x = a \), \( \text{Im} \, x = b \), and the magnitude of \( x \) as \( |x| \). The complex conjugate transpose of a matrix \( X \) is \( X^* \).

The following theorems shall be used in the derivations.

**Theorem 2.1.** Let \( X, Y \) be \( n \times n \) Hermitian matrices. Then for \( j = 1, 2, \ldots, n \)
\[
\delta_j (XY - Y^*X^*) = 0,
\]
where \( \delta_j (\cdot) \) is the diagonal element of matrix \( (\cdot) \) in location \( j \).

**Proof.** A Hermitian matrix \( X \) has real diagonal elements, whilst off-diagonal elements are complex conjugates of each other. Direct multiplication of the \( j \)th row of \( X \) with the \( j \)th column of \( Y \), and similarly for the product of \( Y^*X^* \), show that (2) is valid. \( \Box \)

**Theorem 2.2** [2, 3, p. 228]. Let \( X \) be an \( n \times n \) complex matrix with diagonal elements \( x_{i,i} = 1, 2, \ldots, n \). Then for \( k = 1, 2, \ldots, n \)
\[
\sum_{1}^{k} \text{Re} \, x_i \leq \sum_{1}^{k} |x_i| \leq \sum_{1}^{k} \sigma_i(X).
\]

**Theorem 2.3** [2, 3, p. 243]. Let \( X, Y \) be \( n \times n \) complex matrices. Then for \( k = 1, 2, \ldots, n \)
\[
\sum_{1}^{k} \sigma_i(X + Y) \leq \sum_{1}^{k} (\sigma_i(X) + \sigma_i(Y)).
\]

**Theorem 2.4** [3, p. 250]. Let the \( m \) matrices \( X_1, \ldots, X_m \) be \( n \times n \) complex matrices. Then for \( k = 1, 2, \ldots, n \)
\[
\sum_{1}^{k} \sigma_i(X_1 \cdots X_m) \leq \sum_{1}^{k} \sigma_i(X_1) \cdots \sigma_i(X_m).
\]

**Singular value decomposition.** Let \( A, B \) be an \( n \times n \) complex matrices. Then there exist unitary matrices \( P, Q \) and \( R, S \) such that
\[
A = PA_1Q \quad \text{and} \quad B = RB_1S,
\]
where \( A_1, B_1 \) are diagonal real matrices and \( \lambda_i(A_1) = \sigma_i(A) \) and \( \lambda_i(B_1) = \sigma_i(B) \).

3. Results

**Theorem 3.1.** Let \( A, B \) be \( n \times n \) complex matrices with the singular value decompositions (6), and let \( V, W \) be \( n \times n \) unitary matrices. Write \( T = V P Q W R S \). Then
\[
\text{Re} \, \text{tr}(VAB) \leq \sum_{1}^{n} |\text{Re} \, \lambda_i(T)| \sigma_i(A)\sigma_i(B),
\]
where the \( |\text{Re} \, \lambda_i(T)| \) are in decreasing order, and
Proof. Using (6), the product \( VAWB \) becomes
\[
VAWB = V(PA_1Q)W(RB_1S) = VPA_1(VP)S^*VPQWRS^*B_1S = A_2TB_2,
\]
where \( A_2 = V(PA_1)(VP) \), \( T = VPQWRS \) and \( B_2 = S^*B_1S \). For the Hermitian matrices \( A_2, B_2, \lambda_i(A_2) = \sigma_i(A), \lambda_i(B_2) = \sigma_i(B) \).

In (9), unitary matrix \( T \) may be diagonalized by a unitary matrix \( U \) to produce
\[
U^*VAWBU = U^*A_2UU^*TU^*U^*B_2U = A_3AB_3,
\]
where \( A \) is diagonal with \( \lambda_i(A) = \lambda_i(T) \). Again, \( A_3 = U^*A_2U \) and \( B_3 = U^*B_2U \) are Hermitian matrices with \( \lambda_i(A_3) = \sigma_i(A) \) and \( \lambda_i(B_3) = \sigma_i(B) \).

Separating the diagonal matrix \( A \) into its real part \( A_R \) and its imaginary part \( A_I \), and taking traces, (10) yields
\[
\text{tr}(U^*VAWBU) = \text{tr}(VAWB) = \text{tr}(A_3A_RB_3) + \text{tr}(A_3A_IB_3). \tag{11}
\]

We shall now derive upper bounds for each of the terms in the extreme right-hand side of (11).

Consider first
\[
2\text{tr}(A_3A_RB_3) = 2\text{tr}(A_RB_3A_3) = \text{tr}(A_R(B_3A_3 + A_3^*B_3^*)) + \sigma_i(A_RB_3(A_3^*B_3^*)). \tag{12}
\]

In view of (2), because \( A_R \) is diagonal real
\[
\text{tr}(A_R(B_3A_3 - A_3^*B_3^*)) = 0. \tag{13}
\]

Taking (13) into account, \( \text{tr}(A_3A_RB_3) \) in (11) and (12) is the trace of the product of diagonal real matrix \( A_R \) with Hermitian matrix \( (A_3B_3 + B_3^*A_3^*) \). This means that
\[
\text{tr}(A_3A_RB_3) = \text{Re} \text{tr}(A_3A_RB_3). \tag{14}
\]

Following an application of (3)–(5) in turn to (12) yields
\[
2\text{Re} \text{tr}(A_3A_RB_3) \leq \sum_1^n 2\sigma_i(A_3A_RB_3) \leq \sum_1^n (\sigma_i(A_RA_RB_3) + \sigma_i(A_RB_3A_3^*)) \leq \sum_1^n 2\sigma_i(A)\sigma_i(B)|\lambda_i(A_R)| \tag{15}
\]
as \( \sigma_i(A_R) = |\lambda_i(A_R)| \). Because \( |\lambda_i(A_R)| = |\text{Re} \lambda_i(T)| \), (15) is a statement of (7).

Next consider \( \text{tr}(A_3A_IB_3) \) in (11). Notice that the product \( (A_1/\sqrt{-1}) \) is a diagonal real matrix. Then exactly as (15) was derived from \( \text{tr}(A_3A_RB_3) \), we obtain for \( \text{tr}(A_3(A_1/\sqrt{-1})B_3) \) the inequalities
\[
2\text{Re} \text{tr}(A_3(A_1/\sqrt{-1})B_3) \leq \sum_1^n 2\sigma_i(A_3(A_1/\sqrt{-1})B_3) \leq \sum_1^n (\sigma_i((A_1/\sqrt{-1})A_3B_3) + \sigma_i((A_1/\sqrt{-1})B_3^*A_3^*)) \leq \sum_1^n 2\sigma_i(A)\sigma_i(B)|\lambda_i(A_1)|. \tag{16}
\]
Because \( \text{tr}(A_3(A_1/\sqrt{-1})B_3) = \text{Im} \text{tr}(A_3A_IB_3) \) and \( |\lambda_i(A_1)| = |\text{Im} \lambda_i(T)| \), (16) expresses (8). \( \Box \)
Remark 3.1. Because $|\lambda_i(A_R)| = |\text{Re} \lambda_i(T)| \leq 1$, inequality (7) of Theorem 3.1 is tighter than (1).

References