

Research Article

A Bauer-Hausdorff Matrix Inequality

Zizong Yan¹ and Shanhe Wu²

¹ Department of Information and Mathematics, Yangtze University, Jingzhou, Hubei 434024, China

² Department of Mathematics and Computer Science, Longyan University, Longyan, Fujian 364012, China

Correspondence should be addressed to Shanhe Wu; shanhewu@gmail.com

Received 1 September 2013; Revised 9 November 2013; Accepted 12 November 2013

Academic Editor: Beong In Yun

Copyright © 2013 Z. Yan and S. Wu. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

We present a biorthogonal process for two subspaces of \mathbb{C}^n . Applying this process, we derive a matrix inequality, which generalizes the Bauer-Hausdorff inequality for vectors and includes the Wang-IP inequality for matrices. Meanwhile, we obtain its equivalent matrix inequality.

1. Introduction

The Cauchy-Bunyakovsky-Schwarz, or for short the C.B.S.-inequality, plays an important role in different branches of Modern Mathematics including Hilbert Space Theory, Probability & Statistics, Classical Real and Complex Analysis, Numerical Analysis, and Qualitative Theory of Differential Equations and their applications.

Given an n -dimensional complex space \mathbb{C}^n and two linear subspaces U and V such that

$$U \cap V = \{0\}, \quad (1)$$

there exists

$$\gamma = \gamma(U, V) \in [0, 1] \quad (2)$$

such that for all $x \in U$ and $y \in V$ the following strengthened C.B.S.-inequality holds (see [1]):

$$|x^* y| \leq \gamma \|x\| \cdot \|y\|, \quad (3)$$

where $\|\cdot\|$ denotes the standard Euclidian norm. The smallest quantity γ may be called the cosine of the angle ϕ between the spaces U and V , or be called the C.B.S.-ratio of U and V .

Strengthened C.B.S.-inequality has a long history and there exist various versions. The earliest result of this kind is due to the Wielandt [2] and was later generalized by many researchers. Among them two important extensions of the Wielandt inequality were given by Bauer and

Householder [3], and Wang and Ip [4]. The aim of this paper is to present a matrix version of the Bauer-Hausdorff inequality like one of the Wielandt inequality given by Wang and Ip in [4].

On the practical side, this has been used in the analysis of two-level methods. The survey by Eijkhout and Vassilevski [1] attributes the basic theory of this inequality and its applications in multilevel methods for the solution of linear systems arising from finite element or finite difference discretisation of elliptic partial differential equations. Auzinger and Kirlinger [5] proposed another extension of this inequality for the resolvent conditions in the Kreiss matrix theorem [6].

Throughout the paper, we denote by $R(X)$ the range of the matrix $X \in \mathbb{C}^{n \times p}$,

$$R(X) = \{x \mid x = Xz, z \in \mathbb{C}^p\}, \quad (4)$$

and by $W(A)$ the (closed) numerical range of an operator A on the space \mathbb{C}^n ,

$$W(A) = \{x^* Ax \mid x \in R(A), \|x\| = 1\}, \quad (5)$$

and by $L(a_1, a_2, \dots, a_p)$ the linear subspace spanned by $\{a_1, a_2, \dots, a_p\}$ with $a_i \in \mathbb{C}^n, i = 1, 2, \dots, p$.

Definition 1. For $A \in \mathbb{C}^{n \times n}$, $P_A = AA^+$ denotes the orthogonal projector onto the column space (range) of A , where A^+ is the Moore-Penrose inverse of A .

Definition 2. For two $n \times n$ positive semidefinite Hermitian matrices A , and B , we say that A is below B with respect to

the Löwner partial ordering, and we write $A \leq_L B$, if $B - A$ is positive semidefinite.

2. A Matrix Inequality

Let A be an $n \times n$ positive definite Hermitian matrix. For any two nonzero complex vectors $x \in U$ and $y \in V$, Bauer and Householder [3] asserted that

$$\frac{(x^*Ay)^2}{x^*Ax \cdot y^*Ay} \leq \cos^2\theta, \tag{6}$$

where θ satisfies

$$\cot^2\frac{\theta}{2} = \frac{\lambda_1}{\lambda_n} \cot^2\frac{\phi}{2}, \tag{7}$$

with $\phi \in (0, \pi/2]$ being the angle between two vector subspaces U and V mentioned above, and with λ_1 and λ_n being the largest and smallest, necessarily real and positive, eigenvalues of A .

A very interesting and important special case of (6) is the Wielandt inequality

$$\frac{(x^*Ay)^2}{x^*Ax \cdot y^*Ay} \leq \left(\frac{\lambda_1 - \lambda_n}{\lambda_1 + \lambda_n}\right)^2 \tag{8}$$

established by Wielandt [2] when $\phi = \pi/2$. The upper bound in (8) is called the Wielandt ratio (see [7]). Wang and Ip [4] generalized this inequality as follows.

Let X and Y be complex $n \times p$ and $n \times q$ matrices, respectively. If $X^*Y = 0$, then for all generalized inverses $(Y^*AY)^-$

$$X^*AY(Y^*AY)^-Y^*AX \leq_L \left(\frac{\lambda_1 - \lambda_n}{\lambda_1 + \lambda_n}\right)^2 X^*AX \tag{9}$$

in the Löwner partial ordering.

The first statistical application of the Wielandt inequality seems to be Eaton [8]. There are various equivalent versions of (8) in the literature. The most important one of them is the famous Kantorovich inequality [9, 10] which was used in estimating convergence rate of the steepest descent method for minimizing quadratic problems. Historical and biographical remark on them can be found in [7].

Lemma 3 (see Zhan, [11]). *Let $A \in \mathbb{C}^{n \times n}$ be a Hermitian matrix with 2×2 block form*

$$A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}. \tag{10}$$

Then A is positive semidefinite if and only if A_{11} and A_{22} are positive semidefinite and if there is a contraction matrix W such that $A_{12} = A_{11}^{1/2}WA_{22}^{1/2}$.

If A can be partitioned as (10) and

$$X = \begin{pmatrix} I_p \\ 0 \end{pmatrix} \quad Y = \begin{pmatrix} 0 \\ I_{n-p} \end{pmatrix}, \tag{11}$$

then (9) becomes

$$A_{12}A_{11}^{-1}A_{21} \leq_L \left(\frac{\lambda_1 - \lambda_n}{\lambda_1 + \lambda_n}\right)^2 A_{22} \tag{12}$$

such that

$$W^*W \leq_L \left(\frac{\lambda_1 - \lambda_n}{\lambda_1 + \lambda_n}\right)^2 I. \tag{13}$$

In other words, if A is positive definite with the 2×2 block (10), then there is a contraction matrix W , with the maximum singular value being equal to or less than $(\lambda_1 - \lambda_n)/(\lambda_1 + \lambda_n)$, such that $A_{12} = A_{11}^{1/2}WA_{22}^{1/2}$.

The main result of this paper is stated in the following theorem.

Theorem 4. *Let $A \in \mathbb{C}^{n \times n}$ be a positive semidefinite Hermitian matrix with $\text{rank } r \leq n$ and eigenvalues $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_r > 0$. For any $X \in \mathbb{C}^{n \times p}$ and $Y \in \mathbb{C}^{n \times q}$, if $\phi \in (0, \pi/2]$ is the angle between the two vector spaces $R(P_A X)$ and $R(P_A Y)$ and*

$$\cot^2\frac{\theta}{2} = \frac{\lambda_1}{\lambda_r} \cot^2\frac{\phi}{2}, \tag{14}$$

then

$$X^*AY(Y^*AY)^-Y^*AX \leq_L \cos^2\theta \cdot X^*AX. \tag{15}$$

Theorem 4 will be proved in the next section. This states a very general form of a (strengthened) C.B.S.-inequality and covers various types of C.B.S.-inequalities and their matrix forms. For instance, the inequality (15) reduces to (6) when $p = q = 1$ and reduces to (9) when $\phi = \pi/2$.

Theorem 4 can be rewritten as the following equivalent form.

Theorem 5. *Under the assumptions of Theorem 4, there is a contraction matrix W (from Lemma 3), with the maximum singular value being equal to or less than $\cos \theta$, such that*

$$X^*AY = (X^*AX)^{1/2}W(Y^*AY)^{1/2}. \tag{16}$$

The maximum singular value of the contraction matrix W in Theorem 5 might be strictly less than $\cos \theta$. For example, if

$$A = \begin{pmatrix} \cos \psi & -\sin \psi \\ \sin \psi & \cos \psi \end{pmatrix} \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} \begin{pmatrix} \cos \psi & \sin \psi \\ -\sin \psi & \cos \psi \end{pmatrix} \tag{17}$$

$$= \begin{pmatrix} \lambda_1 \cos^2 \psi + \lambda_2 \sin^2 \psi & (\lambda_1 - \lambda_2) \sin \psi \cos \psi \\ (\lambda_1 - \lambda_2) \sin \psi \cos \psi & \lambda_1 \sin^2 \psi + \lambda_2 \cos^2 \psi \end{pmatrix},$$

then $W = A_{11}^{-1/2}A_{12}A_{22}^{-1/2}$ such that

$$W^*W = \frac{(\lambda_1 - \lambda_2)^2 \sin^2 \psi \cos^2 \psi}{(\lambda_1 \cos^2 \psi + \lambda_2 \sin^2 \psi)(\lambda_1 \sin^2 \psi + \lambda_2 \cos^2 \psi)} \tag{18}$$

$$= \frac{(\lambda_1 - \lambda_2)^2 \sin^2 \psi \cos^2 \psi}{(\lambda_1 - \lambda_2)^2 \sin^2 \psi \cos^2 \psi + \lambda_1 \lambda_2}.$$

That is, the singular value of the contraction matrix W belongs to the interval $[0, (\lambda_1 - \lambda_2)/(\lambda_1 + \lambda_2)]$.

3. Proof of the Main Result

In this section we present an elementary proof of Theorem 4 by a biorthogonal procedure.

The C.B.S.-ratio γ in (3) can be redefined as

$$\gamma = \max \{x^* y \mid \|x\| = \|y\| = 1, x \in U, y \in V\}. \quad (19)$$

Since in finite dimensional spaces the unit sphere is compact, the maximum value of (19) is attained. The following result is obvious.

Lemma 6. *Let γ be the C.B.S.-ratio of two subspaces U and V of \mathbb{C}^n satisfying (1). Then there exist two unit vectors $u \in U$ and $v \in V$ satisfying $\gamma = u^* v$ such that*

$$u - \gamma v \perp V, \quad v - \gamma u \perp U. \quad (20)$$

One direct consequence of Lemma 6 is the following theorem.

Theorem 7. *Let U and V be p and q dimensional linear subspaces of \mathbb{C}^n satisfying (1) with $p, q \geq 2$ and $p + q \leq n$. Then there exist two standard orthogonal bases $\{u_1, u_2, \dots, u_p\}$ of U and $\{v_1, v_2, \dots, v_q\}$ of V such that for each $i = 1, 2, \dots, \min\{p, q\}$*

$$u_i \perp L(v_{i+1}, \dots, v_q), \quad v_i \perp L(u_{i+1}, \dots, u_p). \quad (21)$$

Proof. We shall achieve the desired result by the following biorthogonal process. Start with $U_1 = U$ and $V_1 = V, i = 1$. By Lemma 6, one finds two unit vectors $u_i \in U_i$ and $v_i \in V_i$ such that

$$\gamma_i = u_i^* v_i, \quad (22)$$

$$u_i - \gamma_i v_i \perp V_i \quad v_i - \gamma_i u_i \perp U_i, \quad (23)$$

where γ_i is the C.B.S.-ratio of U_i and V_i with $\gamma_i \leq \gamma$. If $\dim(U_i) = 1$ or $\dim(V_i) = 1$, then the procedure is completed; otherwise, update U_i and V_i by setting

$$U_{i+1} = \left\{ x - \frac{x^* v_i}{\gamma_i} u_i \mid x \in U_i \right\}, \quad (24)$$

$$V_{i+1} = \left\{ y - \frac{y^* u_i}{\gamma_i} v_i \mid y \in V_i \right\}.$$

It is easily proved that

$$u_i \perp V_{i+1}, \quad v_i \perp U_{i+1}. \quad (25)$$

Replace i by $i + 1$, and repeat the above procedure until $i = \min\{p, q\}$.

If p is not equal to q , one finds a standard orthogonal bases $\{u_{q+1}, \dots, u_p\}$ of U_p ($q < p$) or $\{v_{p+1}, \dots, v_q\}$ of V_p ($p < q$). This procedure generates two bases $\{u_1, u_2, \dots, u_p\}$ of U and $\{v_1, v_2, \dots, v_q\}$ of V such that $U_i = L(u_i, u_{i+1}, \dots, u_p)$ and $V_i = L(v_i, v_{i+1}, \dots, v_q)$ for each $i = 1, 2, \dots, \min\{p, q\}$. Equations (23) and (25) imply that

$$u_i \perp U_{i+1}, \quad v_i \perp V_{i+1}, \quad (26)$$

such that these two bases are standard orthogonal. Finally, (21) holds since it is equivalent to (25). \square

In order to acquire our main result, we need the following lemmas.

Lemma 8. *Let $W(A)$ be the numerical range of a linear operator A on the space \mathbb{C}^n . Then*

$$W(X^* AX) \subseteq W(A) \cdot W(X^* P_A X) \quad (27)$$

for any operator $X : \mathbb{C}^k \rightarrow \mathbb{C}^n$ ($k \in \mathbb{N}$ arbitrary), where the multiplication is defined by $X \cdot Y = \{xy : x \in X, y \in Y\}, X, Y \in \mathbb{C}$.

Proof. For any $z \in \mathbb{C}^k$, if $P_A Xz = 0$, then $0 = (X^* AXz, z) \in W(A) \cdot W(X^* P_A X)$. If $P_A Xz \neq 0$ and $\|z\| = 1$, then $A = AA^+ AA^+ A = P_A A P_A$ such that

$$\begin{aligned} &(X^* AXz, z) \\ &= \|P_A Xz\|^2 \left(A \frac{P_A Xz}{\|P_A Xz\|}, \frac{P_A Xz}{\|P_A Xz\|} \right) \in W(A) \cdot W(X^* P_A X), \end{aligned} \quad (28)$$

which results in the desired assertion. \square

When A is invertible, Lemma 8 was established by Fujii [12].

Lemma 9. *If $p \leq q$, then the matrix*

$$C = \begin{pmatrix} I_p & B \\ B^T & I_q \end{pmatrix} \quad (29)$$

has single eigenvalues $1 \pm \gamma_1, \dots, 1 \pm \gamma_p$, where I_p and I_q are two $p \times p$ and $q \times q$ unit matrices, and $B = (\Lambda, 0) \in \mathbb{C}^{p \times q}$ with a $p \times p$ diagonal matrix $\Lambda = \text{diag}(\gamma_1, \gamma_2, \dots, \gamma_p)$. Furthermore, if $p < q$, then C has multiple eigenvalue 1.

Proof. If two vectors $\xi = (\xi_1, \dots, \xi_j, \dots, \xi_{p+q})^T$ and $\eta = (\eta_1, \dots, \eta_j, \dots, \eta_{p+q})^T$ are defined by

$$\begin{aligned} \xi_j &= \begin{cases} 1, & j = i, p + i, \\ 0, & j \neq i, p + i, \end{cases} \\ \eta_j &= \begin{cases} 1, & j = i, \\ -1, & j = p + i, \\ 0, & j \neq i, p + i \end{cases} \end{aligned} \quad (30)$$

then they are eigenvectors of C with the eigenvalues $1 + \gamma_i$ and $1 - \gamma_i$, respectively.

If $p < q$, for each $j = 2p + 1, \dots, p + q$, the j th column vector of the $(p + q) \times (p + q)$ unit matrix is the eigenvector of C with the multiple eigenvalue 1. The proof is completed. \square

Finally, we give the proof of Theorem 4.

Proof of Theorem 4. Let U and V be two standard orthogonal bases of ranges $R(P_A X)$ and $R(P_A Y)$, respectively, with the ranks p and q in Theorem 7. Without loss of generality, we may assume that $p \leq q$. Let $U^* V = (\Lambda, 0) \in \mathbb{C}^{p \times q}$, where

$\Lambda = \text{diag}(\gamma_1, \gamma_2, \dots, \gamma_p)$ is a $p \times p$ diagonal matrix with $\gamma_i \leq \gamma$ for each $i = 1, 2, \dots, p$ and γ the C.B.S.-ratio of $R(P_A X)$ and $R(P_A Y)$. There exist two matrices $X_1 \in \mathbb{C}^{p \times p_1}$ and $Y_1 \in \mathbb{C}^{q \times q_1}$ such that $P_A X = U X_1$ and $P_A Y = V Y_1$.

Letting $Z = (U, V)$, $C = Z^* P_A Z$ can be expressed as the form (29) such that $W(Z^* P_A Z) = [1 - \gamma, 1 + \gamma]$ from Lemma 9. Lemma 8 shows that the matrix

$$Q = Z^* A Z = \begin{pmatrix} U^* A U & U^* A V \\ V^* A U & V^* A V \end{pmatrix} \tag{31}$$

has the largest and smallest eigenvalues $\lambda_1(1 + \gamma)$ and $\lambda_r(1 - \gamma)$. Since (7) holds and

$$\frac{1 + \gamma}{1 - \gamma} = \frac{1 + \cos \phi}{1 - \cos \phi} = \cot^2 \frac{\phi}{2}, \tag{32}$$

the Wielandt ratio of Q is

$$\begin{aligned} \left(\frac{\lambda_1(1 + \gamma) - \lambda_r(1 - \gamma)}{\lambda_1(1 + \gamma) + \lambda_r(1 - \gamma)} \right)^2 &= \left(\frac{\lambda_1/\lambda_r \cot^2 \phi/2 - 1}{\lambda_1/\lambda_r \cot^2 \phi/2 + 1} \right)^2 \\ &= \left(\frac{\cot^2 \theta/2 - 1}{\cot^2 \theta/2 + 1} \right)^2 = \cos^2 \theta \end{aligned} \tag{33}$$

such that the matrix

$$\begin{pmatrix} U^* A U & U^* A V \\ V^* A U & \cos^2 \theta \cdot V^* A V \end{pmatrix} \tag{34}$$

being positive semidefinite by the use of the Schur complement theory (see [13, 14]) to the inequality (9), which leads to the matrix

$$\begin{aligned} \begin{pmatrix} X^* A Y & X^* A Y \\ Y^* A X & \cos^2 \theta \cdot Y^* A Y \end{pmatrix} &= \begin{pmatrix} X_1^* & 0 \\ 0 & Y_1^* \end{pmatrix} \\ &\quad \times \begin{pmatrix} U^* A U & U^* A V \\ V^* A U & \cos^2 \theta \cdot V^* A V \end{pmatrix} \\ &\quad \times \begin{pmatrix} X_1 & 0 \\ 0 & Y_1 \end{pmatrix} \end{aligned} \tag{35}$$

is positive semidefinite. Applying the Schur complement theory again, the desired result is proved. \square

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

The present investigation was supported, in part, by the China Postdoctoral Science Foundation under Grant 2012M520417 and, in part, by the Natural Science Foundation of Fujian province of China under Grant 2012J01014. The authors are grateful to the anonymous referees for their careful comments and valuable suggestions on this paper.

References

- [1] V. Eijkhout and P. Vassilevski, "The role of the strengthened Cauchy-Buniakowskii-Schwarz inequality in multilevel methods," *SIAM Review*, vol. 33, no. 3, pp. 405–419, 1991.
- [2] H. Wielandt, "Inclusion theorems for eigenvalues," in *Simultaneous Linear Equations and the Determination of Eigenvalues*, National Bureau of Standards Applied Mathematics Series, pp. 75–78, U.S. Government Printing Office, Washington, DC, USA, 1953.
- [3] F. L. Bauer and A. S. Householder, "Some inequalities involving the euclidean condition of a matrix," *Numerische Mathematik*, vol. 2, pp. 308–311, 1960.
- [4] S. G. Wang and W. C. Ip, "A matrix version of the Wielandt inequality and its applications to statistics," *Linear Algebra and Its Applications*, vol. 296, no. 1–3, pp. 171–181, 1999.
- [5] W. Auzinger and G. Kirlinger, "Kreiss resolvent conditions and strengthened Cauchy-Schwarz inequalities," *Applied Numerical Mathematics*, vol. 18, no. 1–3, pp. 57–67, 1995.
- [6] H. O. Kreiss, "Über die Stabilitätsdefinition für Differenzgleichungen die partielle Differentialgleichungen approximieren," *BIT Numerical Mathematics*, vol. 2, pp. 153–181, 1962.
- [7] S. W. Drury, S. Liu, C. Y. Lu, S. Puntanen, and G. P. H. Styan, "Some comments on several matrix inequalities with applications to canonical correlations: historical background and recent developments," *Sankhyā*, vol. 64, no. 2, pp. 453–507, 2002.
- [8] M. L. Eaton, "A maximization problem and its application to canonical correlation," *Journal of Multivariate Analysis*, vol. 6, no. 3, pp. 422–425, 1976.
- [9] L. V. Kantorovič, "Functional analysis and applied mathematics," *Uspekhi Matematicheskikh Nauk*, vol. 3, no. 6, pp. 89–185, 1948.
- [10] F. Zhang, "Equivalence of the Wielandt inequality and the Kantorovich inequality," *Linear and Multilinear Algebra*, vol. 48, no. 3, pp. 275–279, 2001.
- [11] X. Z. Zhan, *Matrix Theory*, Higher Education Press, Beijing, China, 2008 (Chinese).
- [12] M. Fujii, *Wielandt Theorem: Simple Proof and Its Generalizations*, Hokkaido University Technical Report Series in Mathematics, 1997.
- [13] R. A. Horn and C. R. Johnson, *Matrix Analysis*, Cambridge University Press, Cambridge, UK, 1985.
- [14] F. Zhang, *The Schur Complement and Its Applications*, vol. 4 of *Numerical Methods and Algorithms*, Springer, New York, NY, USA, 2005.