doi:10.3969/j.issn.1001-4616.2016.01.004

Some Inequalities for Commutators

Yang Xingdong,Liu Shihui,Tu Yuanyuan

(College of Mathematics & Statistics,Nanjing University of Information Science and Technology,Nanjing 210044,China)

Abstract: If $\boldsymbol{\Gamma}$ and $\tilde{\boldsymbol{\Gamma}}$ are $m\times n$ diagonal matrces and $\boldsymbol{A}$ and $\boldsymbol{B}$ are normal matrices, then the upper bound of Frobenius norm of $\boldsymbol{A}\boldsymbol{I}^{(r)}_{m\times n}-\boldsymbol{I}^{(r)}_{m\times n}\boldsymbol{B}$ should be $\frac{1}{\sigma}\|\boldsymbol{A}\boldsymbol{\Gamma}-\tilde{\boldsymbol{\Gamma}}\boldsymbol{B}\|_F$. Some related inequalities are proved, and the results are generalized.

Key words: eigenvalues, singular values, frobenius norm, inequality, commutator

CLC number: 15A57, 15A42 **Document code**: A **Article ID**: 1001-4616(2016)01-0025-04

若干交换算子不等式

杨兴东,刘诗卉,涂媛媛

(南京信息工程大学数学与统计学院,江苏 南京 210044)

[摘要] 本文中设 $\boldsymbol{\Gamma}$ 和 $\tilde{\boldsymbol{\Gamma}}$ 为 $m\times n$ 矩阵,$\boldsymbol{A}$ 和 $\boldsymbol{B}$ 分别为 m 及 n 阶正规矩阵,利用矩阵特征值与奇异值性质,证明如下不等式:$\sigma\|\boldsymbol{A}\boldsymbol{I}^{(r)}_{m\times n}-\boldsymbol{I}^{(r)}_{m\times n}\boldsymbol{B}\|_F\le\|\boldsymbol{A}\boldsymbol{\Gamma}-\tilde{\boldsymbol{\Gamma}}\boldsymbol{B}\|_F$. 同时,推广了相关文献的结论.

[关键词] 特征值,奇异值,F-范数,不等式,交换算子

1 Introduction

Throughout the paper, we use the following notations: $M(\boldsymbol{C})$ is a set of complex matrices, and $\boldsymbol{C}^{m\times n}$ is the set of m by n complex matrices; $\boldsymbol{C}_r^{m\times n}$ is the set of $m\times n$ complex matrices having rank r. $R(\boldsymbol{A})$ denotes the column space of $\boldsymbol{A}$. $\boldsymbol{A}^H$ and $\boldsymbol{A}^+$ denote the conjugate transpose and Moore-Penrose inverse of $\boldsymbol{A}$, respectively. $\boldsymbol{I}_n$ is the unit matrix of order n and $m\times n$ matrix $\boldsymbol{I}^{(r)}_{m\times n}$ is defined by

$$\boldsymbol{I}^{(r)}_{m\times n}=\begin{pmatrix}\boldsymbol{I}_r & O\\ O & O\end{pmatrix}.$$

Let $\boldsymbol{A}=(a_{ij})\in\boldsymbol{C}^{n\times n}$, $d_i(\boldsymbol{A})$ denotes the i th diagonal element of $\boldsymbol{A}$, that is, $d_i(\boldsymbol{A})=a_{ii}(i=1,\dots,n)$. $\|\cdot\|_2$ denotes the Euclidean norm of a vector or the spectral norm of a matrix and $\|\cdot\|_F$ the Frobenius norm of matrix.

The main problem considered in this paper is comparing the Frobenius norm of the matrix $\boldsymbol{A}\boldsymbol{I}^{(r)}_{m\times n}-\boldsymbol{I}^{(r)}_{m\times n}\boldsymbol{B}$ with that of a commutator $\boldsymbol{A}\boldsymbol{\Gamma}-\tilde{\boldsymbol{\Gamma}}\boldsymbol{B}$, where $\boldsymbol{A}$, $\boldsymbol{B}$, $\boldsymbol{\Gamma}$ and $\tilde{\boldsymbol{\Gamma}}$ are elements of $M(\boldsymbol{C})$ satisfying some additional conditions. This kind of problems have been studied by the authors and the results have been found useful in the field of numerical analysis and physics[1-7].

It is well known that the diagonal elements a_{ii}'s of Hermitian semi-definite matrix $\boldsymbol{A}$ are nonnegative real numbers. Hence, if $\boldsymbol{A}$ is an Hermitian semi-definite matrix, and $\boldsymbol{\Lambda}=\mathrm{diag}(a_1,\cdots,a_n)$ with positive diagonal elements, then

Received data: 2015-02-17.

Foundation item: National Natural Science Foundation of China(40975037).

Corresponding author: Yang Xingdong, professor, majored in numerical algebra. E-mail: xdyandnuist@163.com

$$\mathrm{tr}(\boldsymbol{A\Lambda})=\mathrm{tr}(\boldsymbol{\Lambda A})\geqslant 0. \tag{1}$$

We need the inequality(1)in the following proof of the result.

2 Main Results

Theorem 1 Let $\boldsymbol{A}\in\boldsymbol{C}^{m\times m}$, $\boldsymbol{B}\in\boldsymbol{C}^{n\times n}$ be normal matrices and

$$\boldsymbol{\Gamma}=\begin{pmatrix}\boldsymbol{\Omega} & O\\ O & O\end{pmatrix}\in\boldsymbol{C}_r^{m\times n}\ ,\ \tilde{\boldsymbol{\Gamma}}=\begin{pmatrix}\tilde{\boldsymbol{\Omega}} & O\\ O & O\end{pmatrix}\in\boldsymbol{C}_r^{m\times n}\ ,$$

$$\boldsymbol{\Omega}=\mathrm{diag}(\sigma_1,\cdots,\sigma_r)\geqslant \boldsymbol{O},\tilde{\boldsymbol{\Omega}}=\mathrm{diag}(\tilde{\sigma}_1,\cdots,\tilde{\sigma}_r)\geqslant 0.$$

And let the matrices $\boldsymbol{A},\boldsymbol{B}$ satisfy the following conditions:

$$d_i(\boldsymbol{B}^{\mathrm{H}}\boldsymbol{B})-d_i(\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A})\geqslant 0\ \ (i=1,2,\cdots,r)\ . \tag{2}$$

Then

$$\|\boldsymbol{A\Gamma}-\tilde{\boldsymbol{\Gamma}}B\|_{\mathrm{F}}\geqslant\sigma\|\boldsymbol{A}\boldsymbol{I}_{m\times n}^{(r)}-\boldsymbol{I}_{m\times n}^{(r)}\boldsymbol{B}\|_F\ , \tag{3}$$

Where $\sigma=\min\limits_{1\leqslant i\leqslant r}\sigma_i$.

Proof Firstly, we prove the inequality(3)holds for $\tilde{\boldsymbol{\Gamma}}-\boldsymbol{\Gamma}\geqslant 0$. Let $\boldsymbol{A}=\begin{pmatrix}\boldsymbol{A}_{11} & \boldsymbol{A}_{12}\\ \boldsymbol{A}_{21} & \boldsymbol{A}_{22}\end{pmatrix}$, $\boldsymbol{B}=\begin{pmatrix}\boldsymbol{B}_{11} & \boldsymbol{B}_{12}\\ \boldsymbol{B}_{21} & \boldsymbol{B}_{22}\end{pmatrix}$; where $\boldsymbol{A}_{11},\boldsymbol{B}_{11}\in\boldsymbol{C}^{r\times r}$. By (2) and the assumption that $\boldsymbol{A}$ and $\boldsymbol{B}$ are normal matrices, i.e., $\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}=\boldsymbol{A}\boldsymbol{A}^{\mathrm{H}}$, $\boldsymbol{B}^{\mathrm{H}}\boldsymbol{B}=\boldsymbol{B}\boldsymbol{B}^{\mathrm{H}}$,we get

$$\boldsymbol{A}_{11}\boldsymbol{A}_{11}^{\mathrm{H}}+\boldsymbol{A}_{12}\boldsymbol{A}_{12}^{\mathrm{H}}=\boldsymbol{A}_{11}^{\mathrm{H}}\boldsymbol{A}_{11}+\boldsymbol{A}_{21}^{\mathrm{H}}\boldsymbol{A}_{21},\boldsymbol{B}_{11}\boldsymbol{B}_{11}^{\mathrm{H}}+\boldsymbol{B}_{12}\boldsymbol{B}_{12}^{\mathrm{H}}=\boldsymbol{B}_{11}^{\mathrm{H}}\boldsymbol{B}_{11}+\boldsymbol{B}_{21}^{\mathrm{H}}\boldsymbol{B}_{21}, \tag{4}$$

and

$$d_i(\boldsymbol{B}_{11}\boldsymbol{B}_{11}^{\mathrm{H}}+\boldsymbol{B}_{12}\boldsymbol{B}_{12}^{\mathrm{H}})=d_i(\boldsymbol{B}_{11}^{\mathrm{H}}\boldsymbol{B}_{11}+\boldsymbol{B}_{21}^{\mathrm{H}}\boldsymbol{B}_{21})\geqslant d_i(\boldsymbol{A}_{11}^{\mathrm{H}}\boldsymbol{A}_{11}+\boldsymbol{A}_{21}^{\mathrm{H}}\boldsymbol{A}_{21})=d_i(\boldsymbol{A}_{11}\boldsymbol{A}_{11}^{\mathrm{H}}+\boldsymbol{A}_{12}\boldsymbol{A}_{12}^{\mathrm{H}})\ \ (i=1,\cdots,r)\ . \tag{5}$$

Assume that $\boldsymbol{\Sigma}=\boldsymbol{\Gamma}-\sigma\boldsymbol{I}_{m\times n}^{(r)}=\begin{pmatrix}\boldsymbol{\Omega}-\sigma\boldsymbol{I}_r & O\\ O & O\end{pmatrix}$, and $\tilde{\boldsymbol{\Sigma}}=\tilde{\boldsymbol{\Gamma}}-\sigma\boldsymbol{I}_{m\times n}^{(r)}=\begin{pmatrix}\tilde{\boldsymbol{\Omega}}-\sigma\boldsymbol{I}_r & O\\ O & O\end{pmatrix}$. Then $\boldsymbol{\Sigma}$ and $\tilde{\boldsymbol{\Sigma}}$ are two diagonal matrices with diagonal elements nonnegative. By direct computation, we get

$$\boldsymbol{A}\boldsymbol{I}_{m\times n}^{(r)}-\boldsymbol{I}_{m\times n}^{(r)}\boldsymbol{B}=\begin{pmatrix}\boldsymbol{A}_{11}-\boldsymbol{B}_{11} & -\boldsymbol{B}_{12}\\ \boldsymbol{A}_{21} & \boldsymbol{O}\end{pmatrix},$$

$$\boldsymbol{A\Sigma}-\tilde{\boldsymbol{\Sigma}}\boldsymbol{B}=\begin{pmatrix}\boldsymbol{A}_{11}(\Omega-\sigma\boldsymbol{I}_r)-(\tilde{\boldsymbol{\Omega}}-\sigma\boldsymbol{I}_r)\boldsymbol{B}_{11} & -(\tilde{\boldsymbol{\Omega}}-\sigma\boldsymbol{I}_r)\boldsymbol{B}_{12}\\ \boldsymbol{A}_{21}(\boldsymbol{\Omega}-\sigma\boldsymbol{I}_r) & O\end{pmatrix}.$$

It is easy to verify that

$$(\boldsymbol{A}\boldsymbol{I}_{m\times n}^{(r)}-\boldsymbol{I}_{m\times n}^{(r)}\boldsymbol{B})^{\mathrm{H}}(\boldsymbol{A\Sigma}-\tilde{\boldsymbol{\Sigma}}\boldsymbol{B})=\begin{pmatrix}(\boldsymbol{A}_{11}^{\mathrm{H}}-\boldsymbol{B}_{11}^{\mathrm{H}})(\boldsymbol{A}_{11}(\boldsymbol{\Omega}-\sigma\boldsymbol{I}_r)-(\tilde{\boldsymbol{\Omega}}-\sigma\boldsymbol{I}_r)\boldsymbol{B}_{11})+\boldsymbol{A}_{21}^{\mathrm{H}}\boldsymbol{A}_{21}(\boldsymbol{\Omega}-\sigma\boldsymbol{I}_r) & *\\ * & \boldsymbol{B}_{12}^{\mathrm{H}}(\tilde{\boldsymbol{\Omega}}-\sigma\boldsymbol{I}_r)\boldsymbol{B}_{12}\end{pmatrix}, \tag{6}$$

$$[(\boldsymbol{A\Sigma}-\tilde{\boldsymbol{\Sigma}}\boldsymbol{B})^{\mathrm{H}}(\boldsymbol{A}\boldsymbol{I}_{m\times n}^{(r)}-\boldsymbol{I}_{m\times n}^{(r)}\boldsymbol{B})]=\begin{pmatrix}((\boldsymbol{\Omega}-\sigma\boldsymbol{I}_r)\boldsymbol{A}_{11}^{\mathrm{H}}-\boldsymbol{B}_{11}^{\mathrm{H}}(\tilde{\boldsymbol{\Omega}}-\sigma\boldsymbol{I}_r))(\boldsymbol{A}_{11}-\boldsymbol{B}_{11})+(\boldsymbol{\Omega}-\sigma\boldsymbol{I}_r)\boldsymbol{A}_{21}^{\mathrm{H}}\boldsymbol{A}_{21} & *\\ * & \boldsymbol{B}_{12}^{\mathrm{H}}(\tilde{\boldsymbol{\Omega}}-\sigma\boldsymbol{I}_r)\boldsymbol{B}_{12}\end{pmatrix}, \tag{7}$$

where the star(*)parts in(6)and(7)will not be used in our derivations.

Hence,

$$\begin{aligned}\|\boldsymbol{A\Gamma}-\tilde{\boldsymbol{\Gamma}}\boldsymbol{B}\|_F^2&=\|\boldsymbol{A}(\boldsymbol{\Sigma}+\sigma\boldsymbol{I}_{m\times n}^{(r)})-(\tilde{\boldsymbol{\Sigma}}+\sigma\boldsymbol{I}_{m\times n}^{(r)})\boldsymbol{B}\|_F^2=\|\sigma(\boldsymbol{A}\boldsymbol{I}_{m\times n}^{(r)}-\boldsymbol{I}_{m\times n}^{(r)}\boldsymbol{B})+(\boldsymbol{A\Sigma}-\tilde{\boldsymbol{\Sigma}}\boldsymbol{B})\|_F^2=\sigma^2\|\boldsymbol{A}\boldsymbol{I}_{m\times n}^{(r)}-\boldsymbol{I}_{m\times n}^{(r)}\boldsymbol{B}\|_F^2+\\&\|\boldsymbol{A\Sigma}-\tilde{\boldsymbol{\Sigma}}\boldsymbol{B}\|_F^2+2\sigma\,\mathrm{Re}\,\mathrm{tr}[(\boldsymbol{A}\boldsymbol{I}_{m\times n}^{(r)}-\boldsymbol{I}_{m\times n}^{(r)}\boldsymbol{B})^{\mathrm{H}}(\boldsymbol{A\Sigma}-\tilde{\boldsymbol{\Sigma}}\boldsymbol{B})]\ .\end{aligned} \tag{8}$$

Here tr denotes the trace of a matrix, Re the real part of a complex number. We claim that

$$2\sigma\,\mathrm{Re}\,\mathrm{tr}[(\boldsymbol{A}\boldsymbol{I}_{m\times n}^{(r)}-\boldsymbol{I}_{m\times n}^{(r)}\boldsymbol{B})^{\mathrm{H}}(\boldsymbol{A\Sigma}-\tilde{\boldsymbol{\Sigma}}\boldsymbol{B})]\geqslant 0\ . \tag{9}$$

So,(9)together with(8)leads to

$$\|\boldsymbol{A\Gamma}-\tilde{\boldsymbol{\Gamma}}\boldsymbol{B}\|_F^2\geqslant\sigma^2\|\boldsymbol{A}\boldsymbol{I}_{m\times n}^{(r)}-\boldsymbol{I}_{m\times n}^{(r)}\boldsymbol{B}\|_F^2\ , \tag{10}$$

which proves(3). Now we have to prove(9).

By using (6) and (7) and the property of the matrix trace: $\mathrm{tr}(\boldsymbol{MN})=\mathrm{tr}(\boldsymbol{NM})$ for two matrices $\boldsymbol{M}$ and $\boldsymbol{N}$ with suitable dimensions, we get

$$\begin{gathered}2\,\mathrm{Re}\,\mathrm{tr}[(\boldsymbol{A}\boldsymbol{I}_{m\times n}^{(r)}-\boldsymbol{I}_{m\times n}^{(r)}\boldsymbol{B})^{\mathrm{H}}(\boldsymbol{A}\boldsymbol{\Sigma}-\tilde{\boldsymbol{\Sigma}}\boldsymbol{B})]=\mathrm{tr}[(\boldsymbol{A}\boldsymbol{I}_{m\times n}^{(r)}-\boldsymbol{I}_{m\times n}^{(r)}\boldsymbol{B})^{\mathrm{H}}(\boldsymbol{A}\boldsymbol{\Sigma}-\tilde{\boldsymbol{\Sigma}}\boldsymbol{B})+(\boldsymbol{A}\boldsymbol{\Sigma}-\tilde{\boldsymbol{\Sigma}}\boldsymbol{B})^{\mathrm{H}}(\boldsymbol{A}\boldsymbol{I}_{m\times n}^{(r)}-\boldsymbol{I}_{m\times n}^{(r)}\boldsymbol{B})]=\\
\mathrm{tr}\{(\boldsymbol{A}_{11}^{\mathrm{H}}-\boldsymbol{B}_{11}^{\mathrm{H}})[\boldsymbol{A}_{11}(\boldsymbol{\Omega}-\sigma\boldsymbol{I}_r)-(\tilde{\boldsymbol{\Omega}}-\sigma\boldsymbol{I}_r)\boldsymbol{B}_{11}]\}+\mathrm{tr}\{[(\boldsymbol{\Omega}-\sigma\boldsymbol{I}_r)\boldsymbol{A}_{11}^{\mathrm{H}}-\boldsymbol{B}_{11}^{\mathrm{H}}(\tilde{\boldsymbol{\Omega}}-\sigma\boldsymbol{I}_r)](\boldsymbol{A}_{11}-\boldsymbol{B}_{11})\}+2\mathrm{tr}[(\tilde{\boldsymbol{\Omega}}-\sigma\boldsymbol{I}_r)\boldsymbol{B}_{12}\boldsymbol{B}_{12}^{\mathrm{H}}]+\\
2\mathrm{tr}[(\boldsymbol{\Omega}-\sigma\boldsymbol{I}_r)\boldsymbol{A}_{21}^{\mathrm{H}}\boldsymbol{A}_{21}]=\mathrm{tr}[(\boldsymbol{\Omega}-\sigma\boldsymbol{I}_r)\boldsymbol{A}_{11}^{\mathrm{H}}\boldsymbol{A}_{11}-(\boldsymbol{\Omega}-\sigma\boldsymbol{I}_r)\boldsymbol{A}_{11}^{\mathrm{H}}\boldsymbol{B}_{11}-(\boldsymbol{\Omega}-\sigma\boldsymbol{I}_r)\boldsymbol{B}_{11}^{\mathrm{H}}\boldsymbol{A}_{11}+(\boldsymbol{\Omega}-\sigma\boldsymbol{I}_r)\boldsymbol{B}_{11}^{\mathrm{H}}\boldsymbol{B}_{11}]+\\
\mathrm{tr}[(\tilde{\boldsymbol{\Omega}}-\sigma\boldsymbol{I}_r)\boldsymbol{A}_{11}\boldsymbol{A}_{11}^{\mathrm{H}}-(\tilde{\boldsymbol{\Omega}}-\sigma\boldsymbol{I}_r)\boldsymbol{B}_{11}\boldsymbol{A}_{11}^{\mathrm{H}}-(\tilde{\boldsymbol{\Omega}}-\sigma\boldsymbol{I}_r)\boldsymbol{A}_{11}\boldsymbol{B}_{11}^{\mathrm{H}}+(\tilde{\boldsymbol{\Omega}}-\sigma\boldsymbol{I}_r)\boldsymbol{B}_{11}\boldsymbol{B}_{11}^{\mathrm{H}}]+\mathrm{tr}[(\tilde{\boldsymbol{\Omega}}-\sigma\boldsymbol{I}_r)\boldsymbol{B}_{11}\boldsymbol{B}_{11}^{\mathrm{H}}]+\\
\mathrm{tr}[(\boldsymbol{\Omega}-\sigma\boldsymbol{I}_r)\boldsymbol{A}_{11}^{\mathrm{H}}\boldsymbol{A}_{11}]-\mathrm{tr}[(\boldsymbol{\Omega}-\sigma\boldsymbol{I}_r)\boldsymbol{B}_{11}^{\mathrm{H}}\boldsymbol{B}_{11}]-\mathrm{tr}[(\tilde{\boldsymbol{\Omega}}-\sigma\boldsymbol{I}_r)\boldsymbol{A}_{11}\boldsymbol{A}_{11}^{\mathrm{H}}]+2\mathrm{tr}[(\tilde{\boldsymbol{\Omega}}-\sigma\boldsymbol{I}_r)\boldsymbol{B}_{12}\boldsymbol{B}_{12}^{\mathrm{H}}]+2\mathrm{tr}[(\boldsymbol{\Omega}-\sigma\boldsymbol{I}_r)\boldsymbol{A}_{21}^{\mathrm{H}}\boldsymbol{A}_{21}]=\\
\mathrm{tr}[(\boldsymbol{\Omega}-\sigma\boldsymbol{I}_r)(\boldsymbol{A}_{11}-\boldsymbol{B}_{11})^{\mathrm{H}}(\boldsymbol{A}_{11}-\boldsymbol{B}_{11})]+\mathrm{tr}[(\tilde{\boldsymbol{\Omega}}-\sigma\boldsymbol{I}_r)(\boldsymbol{A}_{11}-\boldsymbol{B}_{11})(\boldsymbol{A}_{11}-\boldsymbol{B}_{11})^{\mathrm{H}}]+\mathrm{tr}[(\tilde{\boldsymbol{\Omega}}-\sigma\boldsymbol{I}_r)(\boldsymbol{B}_{11}\boldsymbol{B}_{11}^{\mathrm{H}}+\boldsymbol{B}_{12}\boldsymbol{B}_{12}^{\mathrm{H}})]-\\
\mathrm{tr}[(\boldsymbol{\Omega}-\sigma\boldsymbol{I}_r)\boldsymbol{B}_{11}^{\mathrm{H}}\boldsymbol{B}_{11}]+\mathrm{tr}[(\tilde{\boldsymbol{\Omega}}-\sigma\boldsymbol{I}_r)\boldsymbol{B}_{12}\boldsymbol{B}_{12}^{\mathrm{H}}]+\mathrm{tr}[(\boldsymbol{\Omega}-\sigma\boldsymbol{I}_r)(\boldsymbol{A}_{11}^{\mathrm{H}}\boldsymbol{A}_{11}+\boldsymbol{A}_{21}^{\mathrm{H}}\boldsymbol{A}_{21})]+\mathrm{tr}[(\boldsymbol{\Omega}-\sigma\boldsymbol{I}_r)\boldsymbol{A}_{21}^{\mathrm{H}}\boldsymbol{A}_{21}]-\mathrm{tr}[(\tilde{\boldsymbol{\Omega}}-\sigma\boldsymbol{I}_r)\boldsymbol{A}_{11}\boldsymbol{A}_{11}^{\mathrm{H}}].\end{gathered} \tag{11}$$

Note that $\boldsymbol{\Omega}-\sigma\boldsymbol{I}_r\geqslant 0$, $(\boldsymbol{A}_{11}-\boldsymbol{B}_{11})^{\mathrm{H}}(\boldsymbol{A}_{11}-\boldsymbol{B}_{11})\geqslant\boldsymbol{O}$, and $\tilde{\boldsymbol{\Omega}}-\sigma\boldsymbol{I}_r\geqslant 0$, $(\boldsymbol{A}_{11}-\boldsymbol{B}_{11})\ \ (\boldsymbol{A}_{11}-\boldsymbol{B}_{11})^{\mathrm{H}}\geqslant 0$.

So, by using (1), we readily have

$$\mathrm{tr}[(\boldsymbol{\Omega}-\sigma\boldsymbol{I}_r)(\boldsymbol{A}_{11}-\boldsymbol{B}_{11})^{\mathrm{H}}(\boldsymbol{A}_{11}-\boldsymbol{B}_{11})]\geqslant 0\,,$$

and

$$\mathrm{tr}[(\tilde{\boldsymbol{\Omega}}-\sigma\boldsymbol{I}_r)(\boldsymbol{A}_{11}-\boldsymbol{B}_{11})(\boldsymbol{A}_{11}-\boldsymbol{B}_{11})^{\mathrm{H}}]\geqslant 0\,.$$

In this case, the equality (11) becomes the following inequality:

$$\begin{gathered}(11)\geqslant\mathrm{tr}[(\boldsymbol{\Omega}-\sigma\boldsymbol{I}_r)(\boldsymbol{A}_{11}^{\mathrm{H}}\boldsymbol{A}_{11}+\boldsymbol{A}_{21}^{\mathrm{H}}\boldsymbol{A}_{21})]-\mathrm{tr}[(\tilde{\boldsymbol{\Omega}}-\sigma\boldsymbol{I}_r)\boldsymbol{A}_{11}\boldsymbol{A}_{11}^{\mathrm{H}}]+\mathrm{tr}[(\tilde{\boldsymbol{\Omega}}-\sigma\boldsymbol{I}_r)(\boldsymbol{B}_{11}\boldsymbol{B}_{11}^{\mathrm{H}}+\boldsymbol{B}_{12}\boldsymbol{B}_{12}^{\mathrm{H}})]-\\
\mathrm{tr}[(\boldsymbol{\Omega}-\sigma\boldsymbol{I}_r)\boldsymbol{B}_{11}^{\mathrm{H}}\boldsymbol{B}_{11}]+\mathrm{tr}[(\tilde{\boldsymbol{\Omega}}-\sigma\boldsymbol{I}_r)\boldsymbol{B}_{12}\boldsymbol{B}_{12}^{\mathrm{H}}]+\mathrm{tr}[(\boldsymbol{\Omega}-\sigma\boldsymbol{I}_r)\boldsymbol{A}_{21}^{H}\boldsymbol{A}_{21}].\end{gathered} \tag{12}$$

Hence, (12), together (4) and (11), lends to

$$\begin{gathered}2\,\mathrm{Re}\,\mathrm{tr}[(\boldsymbol{A}\boldsymbol{I}_{m\times n}^{(r)}-\boldsymbol{I}_{m\times n}^{(r)}\boldsymbol{B})^{\mathrm{H}}(\boldsymbol{A}\boldsymbol{\Sigma}-\tilde{\boldsymbol{\Sigma}}\boldsymbol{B})]\geqslant\mathrm{tr}[(\tilde{\boldsymbol{\Omega}}-\boldsymbol{\Omega})(\boldsymbol{B}_{11}^{\mathrm{H}}\boldsymbol{B}_{11}+\boldsymbol{B}_{21}^{\mathrm{H}}\boldsymbol{B}_{21})]-\mathrm{tr}[(\tilde{\boldsymbol{\Omega}}-\boldsymbol{\Omega})(\boldsymbol{A}_{11}\boldsymbol{A}_{11}^{\mathrm{H}}+\boldsymbol{A}_{12}\boldsymbol{A}_{12}^{\mathrm{H}})]+\\
\mathrm{tr}[(\tilde{\boldsymbol{\Omega}}-\sigma\boldsymbol{I}_r)\boldsymbol{B}_{12}\boldsymbol{B}_{12}^{\mathrm{H}}]+\mathrm{tr}[(\boldsymbol{\Omega}-\sigma\boldsymbol{I}_r)\boldsymbol{A}_{21}^{\mathrm{H}}\boldsymbol{A}_{21}]+\mathrm{tr}[(\tilde{\boldsymbol{\Omega}}-\sigma\boldsymbol{I}_r)\boldsymbol{A}_{12}\boldsymbol{A}_{12}^{\mathrm{H}}]+\mathrm{tr}[(\boldsymbol{\Omega}-\sigma\boldsymbol{I}_r)\boldsymbol{B}_{21}^{H}\boldsymbol{B}_{21}]\geqslant\\
\mathrm{tr}[(\tilde{\boldsymbol{\Omega}}-\boldsymbol{\Omega})(\boldsymbol{B}_{11}^{\mathrm{H}}\boldsymbol{B}_{11}+\boldsymbol{B}_{21}^{\mathrm{H}}\boldsymbol{B}_{21})]-\mathrm{tr}[(\tilde{\boldsymbol{\Omega}}-\boldsymbol{\Omega})(\boldsymbol{A}_{11}\boldsymbol{A}_{11}^{\mathrm{H}}+\boldsymbol{A}_{12}\boldsymbol{A}_{12}^{\mathrm{H}})].\end{gathered} \tag{13}$$

Let $\boldsymbol{A}_{11}\boldsymbol{A}_{11}^{\mathrm{H}}+\boldsymbol{A}_{12}\boldsymbol{A}_{12}^{\mathrm{H}}=(a_{ij})_{r\times r}$, and $\boldsymbol{B}_{11}^{\mathrm{H}}\boldsymbol{B}_{11}+\boldsymbol{B}_{21}^{\mathrm{H}}\boldsymbol{B}_{21}=(b_{ij})_{r\times r}$.

Obviously, $\tilde{\boldsymbol{\Gamma}}-\boldsymbol{\Gamma}\geqslant\boldsymbol{O}\Leftrightarrow\tilde{\sigma}_i-\sigma_i\geqslant 0$, and $d_i(\boldsymbol{B}^{\mathrm{H}}\boldsymbol{B})-d_i(\boldsymbol{B}^{\mathrm{H}}\boldsymbol{B})\geqslant\boldsymbol{O}\ \ d_i(\boldsymbol{B}^{\mathrm{H}}\boldsymbol{B})\geqslant\ \ d_i(\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A})\,(i=1,2,\cdots,r)$, we get

$$\mathrm{tr}[(\tilde{\boldsymbol{\Omega}}-\boldsymbol{\Omega})(\boldsymbol{B}_{11}^{\mathrm{H}}\boldsymbol{B}_{11}+\boldsymbol{B}_{21}^{\mathrm{H}}\boldsymbol{B}_{21})]=\sum_{i=1}^{r}(\tilde{\sigma}_i-\sigma_i)b_{ii}\geqslant\sum_{i=1}^{r}(\tilde{\sigma}_i-\sigma_i)a_{ii}=\mathrm{tr}[(\tilde{\boldsymbol{\Omega}}-\boldsymbol{\Omega})(\boldsymbol{A}_{11}\boldsymbol{A}_{11}^{\mathrm{H}}+\boldsymbol{A}_{12}\boldsymbol{A}_{12}^{\mathrm{H}})].$$

which proves (9). Hence (3) holds for $\tilde{\boldsymbol{\Gamma}}-\boldsymbol{\Gamma}\geqslant 0$.

Secondly, we prove the general situation. Let the positive number δ satisfy the following conditions:

$$\delta\,\mathrm{tr}(\boldsymbol{B}^{\mathrm{H}}\boldsymbol{I}_{n\times n}^{(r)}\boldsymbol{B})+\mathrm{tr}[\boldsymbol{B}^{\mathrm{H}}\boldsymbol{I}_{n\times m}^{(r)}(\boldsymbol{A}\boldsymbol{\Gamma}-\tilde{\boldsymbol{\Gamma}}\boldsymbol{B})+(\boldsymbol{A}\boldsymbol{\Gamma}-\tilde{\boldsymbol{\Gamma}}\boldsymbol{B})^{\mathrm{H}}\boldsymbol{I}_{m\times n}^{(r)}\boldsymbol{B}]\geqslant 0\text{, and }\tilde{\boldsymbol{\Omega}}+\delta\boldsymbol{I}_r\geqslant\boldsymbol{\Omega}.$$

Then

$$(\tilde{\boldsymbol{\Omega}}+\delta\boldsymbol{I}_r)-\boldsymbol{\Omega}\geqslant 0\,, \tag{14}$$

and

$$3\mathrm{tr}(\boldsymbol{B}^{\mathrm{H}}\delta^2\boldsymbol{I}_{n\times n}^{(r)}\boldsymbol{B}+\mathrm{tr}\{\boldsymbol{B}^{\mathrm{H}}\delta\boldsymbol{B}_{n\times m}^{(r)}[\boldsymbol{A}\boldsymbol{\Gamma}-(\tilde{\boldsymbol{\Gamma}}+\delta\boldsymbol{I}_{m\times n}^{(r)})\boldsymbol{B}]+[\boldsymbol{A}\boldsymbol{\Gamma}-(\tilde{\boldsymbol{\Gamma}}+\delta\boldsymbol{I}_{m\times n}^{(r)})\boldsymbol{B}]^{\mathrm{H}}\delta\boldsymbol{I}_{m\times n}^{(r)}\boldsymbol{B}\}\geqslant 0. \tag{15}$$

Let $\boldsymbol{T}=\tilde{\boldsymbol{\Gamma}}+\delta\boldsymbol{I}_{m\times n}^{(r)}$. Then, the inequality (14) implies that the following inequality holds: $\boldsymbol{T}-\tilde{\boldsymbol{\Gamma}}\geqslant\boldsymbol{O}$. By simple computation, we have

$$\begin{gathered}\|\boldsymbol{A}\boldsymbol{\Gamma}-\tilde{\boldsymbol{\Gamma}}\boldsymbol{B}\|_F^2=\|\boldsymbol{A}\boldsymbol{\Gamma}-(\boldsymbol{T}-\delta\boldsymbol{I}_{m\times n}^{(r)})\boldsymbol{B}\|_F^2=\|\boldsymbol{A}\boldsymbol{\Gamma}-\boldsymbol{T}\boldsymbol{B}\|_F^2+\|\delta\boldsymbol{I}_{m\times n}^{(r)}\boldsymbol{B}\|_F^2+2\delta\mathrm{Re}\,\mathrm{tr}[\boldsymbol{B}^{\mathrm{H}}\boldsymbol{I}_{n\times m}^{(r)}(\boldsymbol{A}\boldsymbol{\Gamma}-\boldsymbol{T}\boldsymbol{B})]=\\
\|\boldsymbol{A}\boldsymbol{\Gamma}-\boldsymbol{T}\boldsymbol{B}\|_F^2+3\mathrm{tr}(\boldsymbol{B}^{\mathrm{H}}\delta^2\boldsymbol{I}_{n\times n}^{(r)}\boldsymbol{B})-2\mathrm{tr}(\boldsymbol{B}^{\mathrm{H}}\delta^2\boldsymbol{I}_{n\times n}^{(r)}\boldsymbol{B})+\mathrm{tr}[\boldsymbol{B}^{\mathrm{H}}\delta\boldsymbol{I}_{n\times m}^{(r)}(\boldsymbol{A}\boldsymbol{\Gamma}-\boldsymbol{T}\boldsymbol{B})+(\boldsymbol{A}\boldsymbol{\Gamma}-\boldsymbol{T}\boldsymbol{B})^{\mathrm{H}}\delta\boldsymbol{I}_{m\times n}^{(r)}\boldsymbol{B}]=\\
\|\boldsymbol{A}\boldsymbol{\Gamma}-\boldsymbol{T}\boldsymbol{B}\|_F^2+3\mathrm{tr}(\boldsymbol{B}^{\mathrm{H}}\delta^2\boldsymbol{I}_{n\times n}^{(r)}\boldsymbol{B})+\mathrm{tr}\{[\boldsymbol{B}^{\mathrm{H}}\delta\boldsymbol{I}_{n\times m}^{(r)}(\boldsymbol{A}\boldsymbol{\Gamma}-(\tilde{\boldsymbol{\Gamma}}+\delta\boldsymbol{I}_{m\times n}^{(r)})\boldsymbol{B}]+[\boldsymbol{A}\boldsymbol{\Gamma}-(\tilde{\boldsymbol{\Gamma}}+\delta\boldsymbol{I}_{m\times n}^{(r)})\boldsymbol{B}]^{\mathrm{H}}\delta\boldsymbol{I}_{m\times n}^{(r)})\boldsymbol{B}\}.\end{gathered} \tag{16}$$

So,(16)together with(15),leads to $\|A\Gamma-\tilde{\Gamma}B\|_F^2 \geqslant \|A\Gamma-TB\|_F^2$.

From the inequality $\tilde{\Gamma}-\Gamma\geqslant 0$ and the specific case above we have proved, it may be concluded that $\|A\Gamma-TB\|_F^2 \geqslant \sigma^2\|AI_{m\times n}^{(r)}-I_{m\times n}^{(r)}B\|_F^2$. Hence,(3)holds. The proof is completed.

Remark 1 Let $U,\tilde{U}\in C^{m\times m}, V,\tilde{V}\in C^{n\times n}$ be unitary matrices, and let Γ, $\tilde{\Gamma}$, Ω, $\tilde{\Omega}$ be of the forms in Theorem 1. Then $\tilde{U}^{\mathrm{H}}U\in C^{m\times m}$, $\tilde{V}^{\mathrm{H}}V\in C^{n\times n}$ are unitary. Of course, they are normal matrices, e.g.,

$$(\tilde{U}^{\mathrm{H}}U)^{\mathrm{H}}(\tilde{U}^{\mathrm{H}}U)=(\tilde{U}^{\mathrm{H}}U)(\tilde{U}^{\mathrm{H}}U)^{\mathrm{H}}=\mathrm{I}_m,\ (\tilde{V}^{\mathrm{H}}V)^{\mathrm{H}}(\tilde{V}^{\mathrm{H}}V)=(\tilde{V}^{\mathrm{H}}V)(\tilde{V}^{\mathrm{H}}V)^{\mathrm{H}}=I_n.$$

Hence,

$$d_i[(\tilde{U}^{\mathrm{H}}U)^{\mathrm{H}}\tilde{U}^{\mathrm{H}}U]=d_i[(\tilde{V}^{\mathrm{H}}U)^{\mathrm{H}}\tilde{U}^{\mathrm{H}}U]=1,\ (i=1,\cdots,r,1\leqslant r\leqslant\min\{m,n\}).$$

For these reasons, it follows from the unitary invariance of $\|\cdot\|_F$ and Theorem 1 that

$$\|U\Gamma V^{\mathrm{H}}-\tilde{U}\tilde{\Gamma}\tilde{V}^{\mathrm{H}}\|_F=\|\tilde{U}^{\mathrm{H}}U\Gamma-\tilde{\Gamma}\tilde{V}^{\mathrm{H}}V\|_F\geqslant\sigma\|\tilde{U}^{\mathrm{H}}UI_{m\times n}^{(r)}-I_{m\times n}^{(r)}\tilde{V}^{\mathrm{H}}V\|_F=\sigma\|UI_{m\times n}^{(r)}V^{\mathrm{H}}-\tilde{U}I_{m\times n}^{(r)}\tilde{V}^{\mathrm{H}}\|_F,$$

where $\sigma=\min\limits_{1\leqslant i,j\leqslant r}\{\sigma_i,\tilde{\sigma}_j\}$.

Overall, we have the following inequality,

$$\|U\Gamma V^{\mathrm{H}}-\tilde{U}\tilde{\Gamma}\tilde{V}^{\mathrm{H}}\|_F\geqslant\sigma\|UI_{m\times n}^{(r)}V^{\mathrm{H}}-\tilde{U}I_{m\times n}^{(r)}\tilde{V}^{\mathrm{H}}\|_F. \tag{17}$$

Particularly, choose V^{H}, $\tilde{U}$ as $n\times n$ nuit matrix and $m\times m$ nuit matrix, respectively. And let $\tilde{V}^{\mathrm{H}}=V$. Then the inequality(17)becomes

$$\|U\Gamma-\tilde{\Gamma}V\|_F\geqslant\sigma\|UI_{m\times n}^{(r)}-I_{m\times n}^{(r)}V\|_F. \tag{18}$$

The inequality(17)is the result of Theorem 2 in[1],and the inequality(18)is Lemma 1 in[1].

Remark 2 We get the following result, which was proved by Sun in[2]: Let $A\in C^{m\times m}$, $B\in C^{n\times n}$ be normal matrices and

$$\Gamma=\begin{pmatrix}\Omega & O\\ O & O\end{pmatrix}\in C_r^{m\times n},$$

Where $\Omega=\mathrm{diag}(\sigma_1,\cdots,\sigma_n)\geqslant 0$. Then

$$\|A\Gamma-\Gamma B\|_F\geqslant\sigma_n\|AI_{m\times n}^{(r)}-I_{m\times n}^{(r)}B\|_F. \tag{19}$$

In fact, If $\tilde{\Gamma}=\Gamma$ in Theroem 1, then $\Omega=\tilde{\Omega}$. In this case, from the inequality(13)we see that

$$2\,\mathrm{Re}\,\mathrm{tr}[(AI_{m\times n}^{(r)}-I_{m\times n}^{(r)}B)^{\mathrm{H}}(A\Sigma-\tilde{\Sigma}B)]\geqslant 0,$$

which means that(19)holds.

[参考文献]

[1] LI R C. A perturbation bound for the generalized polar decomposition[J]. BIT Numerical Mathematic, 1993, 33: 304-308.

[2] SUN J G, CHEN C H. Generalized polar decomsition[J]. Journal of computational mathematics, 1989(3): 262-273.

[3] MARSHALL A W, OLKIN I. Inequalities: theory of majoration and its applications[J]. New York: Academic Press, 1979.

[4] WEISS G. The Fuglede commutativity theorem modulo the Hilbert-Schmidt class and generating functions for matrix operators, I, II[J]. Trans Amer Math Soc, 1978, 246: 193-209.

[5] LI RC C. New perturbation bounds for the unitary polar factor[J]. SIAM J Matrix Annl Appl, 1995, 16: 327-332.

[6] RAJENDRA R, CHANDLER D, FUAD K. Some inequalities for commutators and an applications to spectral variation[J]. Aequationes mathematicae, 1991, 41: 70-78.

[7] DAVIS C, KAHAN W M. The rotation of eigenvectors by a perturbation[J]. SIAM J Numer Anal, 1970(1): 1-46.

[责任编辑:陆炳新]