Linear Algebra and its Applications 432 (2010) 956–963

Contents lists available at ScienceDirect

Linear Algebra and its Applications

journal homepage: www.elsevier.com/locate/laa

The optimal perturbation bounds of the Moore–Penrose inverse under the Frobenius norm ☆

Lingsheng Meng, Bing Zheng *

School of Mathematics and Statistics, Lanzhou University, Lanzhou 730000, PR China

ARTICLE INFO

Article history:
Received 27 July 2009
Accepted 24 September 2009
Available online 30 October 2009

Submitted by R.A. Brualdi

AMS classification:
15A09
15A24

Keywords:
Singular Value Decomposition
Frobenius norm
Moore–Penrose inverse
Additive perturbation
Multiplicative perturbation

ABSTRACT

We obtain the optimal perturbation bounds of the Moore–Penrose inverse under the Frobenius norm by using Singular Value Decomposition, which improved the results in the earlier paper [P.-Å. Wedin, Perturbation theory for pseudo-inverses, BIT 13 (1973) 217–232]. In addition, a perturbation bound of the Moore–Penrose inverse under the Frobenius norm in the case of the multiplicative perturbation model is also given.

© 2009 Elsevier Inc. All rights reserved.

1. Introduction

Let $C^{m\times n}$ and $C_r^{m\times n}$ be the set of all $m \times n$ complex matrices and its subset with rank r, respectively. Without loss of generality, we always assume that $m \geqslant n$. For a given matrix $A \in C^{m\times n}$, the symbols A^*, $\|A\|_2$ and $\|A\|_F$ will stand for the conjugate transpose, the spectral norm (2-norm) and the Frobenius norm (F-norm) of A, respectively. I_m denotes the identity matrix of order m.

We recall that the Moore–Penrose inverse $A^\dagger$ of a matrix $A \in C^{m\times n}$ is the unique solution X to the following four Moore–Penrose equations [1,2]:

☆ Supported by the Fundamental Research Fund for Physics and Mathematic of Lanzhou University and Gansu Project of Science and Technology (096RJZA049), China.

* Corresponding author.
E-mail address: bzheng@lzu.edu.cn (B. Zheng).

0024-3795/$ - see front matter © 2009 Elsevier Inc. All rights reserved.
doi:10.1016/j.laa.2009.10.009

$$(1)\ AXA = A, \quad (2)\ XAX = X, \quad (3)\ (AX)^* = AX, \quad (4)\ (XA)^* = XA.$$

The Moore–Penrose inverse plays an important role in numerical computations, statistics and other engineering problems. However, in most numerical applications the elements of A will seldom be known exactly, so it is necessary to have bounds on the effects of the uncertainties in A. Motivated by this, much effort has been made for estimating the perturbation bounds of the Moore–Penrose inverse, see [3,4,5,6,7,8,9,10]. In an old paper of Wedin [8], he presented some perturbation bounds of the Moore–Penrose inverse under general unitarily invariant norm, the spectral norm and the Frobenius norm, respectively. Here we restate them below:

Theorem 1.1 [8]. *Let $A \in C^{m\times n}$ and $B = A + E$. Then*

$$\|B^\dagger - A^\dagger\| \leqslant \mu \max\left\{\|A^\dagger\|_2^2, \|B^\dagger\|_2^2\right\}\|E\|, \tag{1.1}$$

where μ is listed in the following table:

$\|\ \|$	*Unitarily invariant norm*	*2-Norm*	*F-Norm*
μ	*3*	$\frac{1+\sqrt{5}}{2}$	$\sqrt{2}$

In particular, if $\operatorname{rank}(A) = \operatorname{rank}(B)$, *then*

$$\|B^\dagger - A^\dagger\| \leqslant \mu\|A^\dagger\|_2\|B^\dagger\|_2\|E\|, \tag{1.2}$$

where μ is listed in the following table:

rank	$\|\ \|$		
	Unitarily invariant norm	*2-Norm*	*F-Norm*
$\operatorname{rank}(A) < n \leqslant m$	*3*	$\frac{1+\sqrt{5}}{2}$	$\sqrt{2}$
$\operatorname{rank}(A) = n < m$	*2*	$\sqrt{2}$	*1*
$\operatorname{rank}(A) = n = m$	*1*	*1*	*1*

In an earlier report Wedin [5] considers the sharpness of the constants μ in (1.2) and shows that for the 2-norm μ cannot be made smaller. To our knowledge, the sharpness of the constant μ for the F-norm has not been studied yet in literature. Hence the purpose of this paper is to develop the optimal constants μ in (1.1) and (1.2) for the F-norm. The main tool used here is the Singular Value Decomposition (SVD), which is a much different approach from that considered in [8] or widely cited in other literature (see [11]). The Singular Value Decomposition of a matrix can be stated as follows:

Lemma 1.1 [1,2]. *Let $A \in C^{m\times n}$ and $\operatorname{rank}(A) = r$. Then there exist unitary matrices $U \in C^{m\times m}$, $V \in C^{n\times n}$ such that*

$$A = U\begin{pmatrix} \Sigma & 0 \\ 0 & 0 \end{pmatrix}V^*,$$

where Σ is the diagonal matrix given by $\Sigma = \operatorname{diag}(\sigma_1, \ldots, \sigma_r)$, $\sigma_1 \geqslant \sigma_2 \geqslant \cdots \geqslant \sigma_r > 0$ are the singular values of A, i.e., the positive square roots of the positive eigenvalues of A^*A.

Under the Singular Value Decomposition, we know that the 2-norm of matrix A is $\|A\|_2 = \sigma_1$. Furthermore, if we partition the unitary matrices U, V as $U = (U_1, U_2)$ and $V = (V_1, V_2)$ with $U_1 \in C^{m\times r}$ and $V_1 \in C^{n\times r}$, then $A = U_1\Sigma V_1^*$ and $A^\dagger = V_1\Sigma^{-1}U_1^*$. Naturally, we have $\|A^\dagger\|_2 = \frac{1}{\sigma_r}$.

There are in general two perturbation models: additive perturbation model and multiplicative perturbation model. In this paper we first consider the optimal perturbation bounds of the Moore–

Penrose inverse under the F-norm for additive perturbation and then as a special additive perturbation, the multiplicative perturbation bound is also developed. The multiplicative perturbation bounds are much sharper than the optimally additive perturbation bounds for the Moore–Penrose inverse in some cases. Hence multiplicative perturbation bound possesses its particular significance. The following two lemmas also play important roles when we estimate the perturbation bounds.

Lemma 1.2 [12]. *Let $U = (U_1, U_2) \in C^{m\times m}$ and $V = (V_1, V_2) \in C^{n\times n}$ be unitary matrices, where $U_1 \in C^{m\times r}$, $V_1 \in C^{n\times s}$, $r \leqslant m$ and $s \leqslant n$. Then for any matrix $E \in C^{m\times n}$ we have*

$$\|E\|_F^2 = \|U_1^*EV_1\|_F^2 + \|U_1^*EV_2\|_F^2 + \|U_2^*EV_1\|_F^2 + \|U_2^*EV_2\|_F^2.$$

Lemma 1.3 [13]. *Let $W \in C^{n\times n}$ be a unitary matrix, and rewritten W in the block form*

$$W = \begin{pmatrix} W_{11} & W_{12} \\ W_{21} & W_{22} \end{pmatrix},$$

where $W_{11} \in C^{r\times r}$, $W_{22} \in C^{(n-r)\times(n-r)}$, $1 \leqslant r < n$. Then $\|W_{12}\| = \|W_{21}\|$ for any unitarily invariant norm.

This paper is organized as follows. Section 2 investigates the optimal additive perturbation bounds of the Moore–Penrose inverse under the F-norm by using the Singular Value Decomposition. That is, the coefficient μ in (1.1) and (1.2) can be uniformly reduced into 1 when the case of F-norm. A perturbation bound of the Moore–Penrose inverse under the F-norm when the case of the multiplicative perturbation model is presented in Section 3. Numerical example shows that as a special additive perturbation, the multiplicative perturbation bound is much better, in some cases, than the additive perturbation bound.

2. Additive perturbation bounds

In [8] the author studied the additive perturbation of the Moore–Penrose inverse and obtained perturbation bounds under the 2-norm, F-norm and generally unitary invariant norm, respectively. In this section, we further investigate the additive perturbation bounds of the Moore–Penrose inverse under the F-norm by means of SVD. We improve the perturbation bounds for the general additive perturbation and a special additive perturbation, i.e. $\operatorname{rank}(A+E) = \operatorname{rank}(A)$. Two numerical examples are given to confirm the sharpness of the bounds.

For the general case, we have the following result:

Theorem 2.1. *Let $A \in C_r^{m\times n}$ and $B = A + E \in C^{m\times n}$. Then*

$$\|B^\dagger - A^\dagger\|_F \leqslant \max\left\{\|A^\dagger\|_2^2, \|B^\dagger\|_2^2\right\} \|E\|_F. \tag{2.1}$$

Proof. Let $B = A + E \in C_s^{m\times n}$, A and B with the following SVDs

$$A = U\begin{pmatrix} \Sigma_1 & 0 \\ 0 & 0 \end{pmatrix} V^* = U_1\Sigma_1V_1^* \quad \text{and} \quad B = \widetilde{U}\begin{pmatrix} \widetilde{\Sigma}_1 & 0 \\ 0 & 0 \end{pmatrix}\widetilde{V}^* = \widetilde{U}_1\widetilde{\Sigma}_1\widetilde{V}_1^*, \tag{2.2}$$

where $U = (U_1, U_2)$, $\widetilde{U} = (\widetilde{U}_1, \widetilde{U}_2) \in C^{m\times m}$ and $V = (V_1, V_2)$, $\widetilde{V} = (\widetilde{V}_1, \widetilde{V}_2) \in C^{n\times n}$ are unitary matrices, $U_1 \in C^{m\times r}$, $\widetilde{U}_1 \in C^{m\times s}$, $V_1 \in C^{n\times r}$, $\widetilde{V}_1 \in C^{n\times s}$, $\Sigma_1 = diag(\sigma_1, \ldots, \sigma_r)$, $\widetilde{\Sigma}_1 = diag(\widetilde{\sigma}_1, \ldots, \widetilde{\sigma}_s)$, $\sigma_1 \geqslant \cdots \geqslant \sigma_r > 0$ and $\widetilde{\sigma}_1 \geqslant \cdots \geqslant \widetilde{\sigma}_s > 0$. By (2.2), the perturbation matrix E can be written as

$$E = B - A = \widetilde{U}_1\widetilde{\Sigma}_1\widetilde{V}_1^* - U_1\Sigma_1V_1^*. \tag{2.3}$$

Note that

$$U^*U = I_m \Longrightarrow U_1^*U_1 = I_r \quad \text{and} \quad U_1^*U_2 = 0;$$
$$V^*V = I_n \Longrightarrow V_1^*V_1 = I_r \quad \text{and} \quad V_1^*V_2 = 0;$$

$$\widetilde{U}^*\widetilde{U} = I_m \Longrightarrow \widetilde{U}_1^*\widetilde{U}_1 = I_s \quad \text{and} \quad \widetilde{U}_1^*\widetilde{U}_2 = 0; \qquad (2.4)$$
$$\widetilde{V}^*\widetilde{V} = I_n \Longrightarrow \widetilde{V}_1^*\widetilde{V}_1 = I_s \quad \text{and} \quad \widetilde{V}_1^*\widetilde{V}_2 = 0,$$

which together with (2.3) give the following equalities:

$$\widetilde{\Sigma}_1\widetilde{V}_1^*V_1 - \widetilde{U}_1^*U_1\Sigma_1 = \widetilde{U}_1^*EV_1, \qquad (2.5)$$
$$U_1^*\widetilde{U}_1\widetilde{\Sigma}_1 - \Sigma_1V_1^*\widetilde{V}_1 = U_1^*E\widetilde{V}_1, \qquad (2.6)$$
$$U_2^*\widetilde{U}_1\widetilde{\Sigma}_1 = U_2^*E\widetilde{V}_1, \quad \widetilde{U}_2^*U_1\Sigma_1 = -\widetilde{U}_2^*EV_1, \qquad (2.7)$$
$$\widetilde{\Sigma}_1\widetilde{V}_1^*V_2 = \widetilde{U}_1^*EV_2, \quad \Sigma_1V_1^*\widetilde{V}_2 = -U_1^*E\widetilde{V}_2. \qquad (2.8)$$

Noting the fact that both $\widetilde{\Sigma}_1$ and Σ_1 are nonsingular, it follows from (2.5) and (2.6) that

$$\widetilde{V}_1^*V_1\Sigma_1^{-1} - \widetilde{\Sigma}_1^{-1}\widetilde{U}_1^*U_1 = \widetilde{\Sigma}_1^{-1}\widetilde{U}_1^*EV_1\Sigma_1^{-1} \qquad (2.9)$$

and

$$\Sigma_1^{-1}U_1^*\widetilde{U}_1 - V_1^*\widetilde{V}_1\widetilde{\Sigma}_1^{-1} = \Sigma_1^{-1}U_1^*E\widetilde{V}_1\widetilde{\Sigma}_1^{-1}. \qquad (2.10)$$

Since $A^\dagger = V_1\Sigma_1^{-1}U_1^*$ and $B^\dagger = \widetilde{V}_1\widetilde{\Sigma}_1^{-1}\widetilde{U}_1^*$, we can obtain that

$$\begin{pmatrix}\widetilde{V}_1^*\\ \widetilde{V}_2^*\end{pmatrix}(B^\dagger - A^\dagger)(U_1, U_2) = \begin{pmatrix}\widetilde{\Sigma}_1^{-1}\widetilde{U}_1^*U_1 - \widetilde{V}_1^*V_1\Sigma_1^{-1} & \widetilde{\Sigma}_1^{-1}\widetilde{U}_1^*U_2\\ -\widetilde{V}_2^*V_1\Sigma_1^{-1} & 0\end{pmatrix} \qquad (2.11)$$

and

$$\begin{pmatrix}V_1^*\\ V_2^*\end{pmatrix}(B^\dagger - A^\dagger)(\widetilde{U}_1, \widetilde{U}_2) = \begin{pmatrix}V_1^*\widetilde{V}_1\widetilde{\Sigma}_1^{-1} - \Sigma_1^{-1}U_1^*\widetilde{U}_1 & -\Sigma_1^{-1}U_1^*\widetilde{U}_2\\ V_2^*\widetilde{V}_1\widetilde{\Sigma}_1^{-1} & 0\end{pmatrix}. \qquad (2.12)$$

It follows from (2.11), (2.12) and some basic properties of the F-norm that

$$\begin{aligned}2\|B^\dagger - A^\dagger\|_F^2 &= \|\widetilde{\Sigma}_1^{-1}\widetilde{U}_1^*U_1 - \widetilde{V}_1^*V_1\Sigma_1^{-1}\|_F^2 + \|\widetilde{\Sigma}_1^{-1}\widetilde{U}_1^*U_2\|_F^2 + \|\widetilde{V}_2^*V_1\Sigma_1^{-1}\|_F^2\\ &\quad + \|V_1^*\widetilde{V}_1\widetilde{\Sigma}_1^{-1} - \Sigma_1^{-1}U_1^*\widetilde{U}_1\|_F^2 + \|\Sigma_1^{-1}U_1^*\widetilde{U}_2\|_F^2 + \|V_2^*\widetilde{V}_1\widetilde{\Sigma}_1^{-1}\|_F^2.\end{aligned} \qquad (2.13)$$

Hence from (2.7)–(2.10), (2.13) and Lemma 1.2, we have

$$\begin{aligned}2\|B^\dagger - A^\dagger\|_F^2 &= \|\widetilde{\Sigma}_1^{-1}\widetilde{U}_1^*EV_1\Sigma_1^{-1}\|_F^2 + \|\widetilde{\Sigma}_1^{-2}\widetilde{V}_1^*E^*U_2\|_F^2 + \|\widetilde{V}_2^*E^*U_1\Sigma_1^{-2}\|_F^2\\ &\quad + \|\Sigma_1^{-1}U_1^*E\widetilde{V}_1\widetilde{\Sigma}_1^{-1}\|_F^2 + \|\Sigma_1^{-2}V_1^*E^*\widetilde{U}_2\|_F^2 + \|V_2^*E^*\widetilde{U}_1\widetilde{\Sigma}_1^{-2}\|_F^2\\ &\leqslant \frac{1}{\tilde{\sigma}_s^2\sigma_r^2}(\|\widetilde{U}_1^*EV_1\|_F^2 + \|U_1^*E\widetilde{V}_1\|_F^2) + \frac{1}{\tilde{\sigma}_s^4}(\|\widetilde{V}_1^*E^*U_2\|_F^2 + \|V_2^*E^*\widetilde{U}_1\|_F^2)\\ &\quad + \frac{1}{\sigma_r^4}(\|\widetilde{V}_2^*E^*U_1\|_F^2 + \|V_1^*E^*\widetilde{U}_2\|_F^2)\\ &\leqslant \max\left\{\frac{1}{\tilde{\sigma}_s^4}, \frac{1}{\sigma_r^4}\right\}\Big(\|\widetilde{U}_1^*EV_1\|_F^2 + \|\widetilde{U}_1^*EV_2\|_F^2 + \|\widetilde{U}_2^*EV_1\|_F^2\\ &\quad + \|U_1^*E\widetilde{V}_1\|_F^2 + \|U_2^*E\widetilde{V}_1\|_F^2 + \|U_1^*E\widetilde{V}_2\|_F^2\Big)\\ &\leqslant 2\max\left\{\frac{1}{\tilde{\sigma}_s^4}, \frac{1}{\sigma_r^4}\right\}\|E\|_F^2.\end{aligned} \qquad (2.14)$$

Notice that $\|A^\dagger\|_2 = \frac{1}{\sigma_r}$ and $\|B^\dagger\|_2 = \frac{1}{\tilde{\sigma}_s}$, we immediately have

$$\|B^\dagger - A^\dagger\|_F \leqslant \max\left\{\|A^\dagger\|_2^2, \|B^\dagger\|_2^2\right\}\|E\|_F.$$

The proof is completed. □

The following example shows that the perturbation bound in Theorem 2.1 is optimal.

Example 1. Let $A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$, $E = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 10^{-6} & 0 \\ 0 & 0 & 0 \end{pmatrix}$. Then $B = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 10^{-6} & 0 \\ 0 & 0 & 0 \end{pmatrix}$, $A^\dagger = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$ and $B^\dagger = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 10^{6} & 0 \\ 0 & 0 & 0 \end{pmatrix}$. So, we have

$$\|B^\dagger - A^\dagger\|_F = 10^6 \quad \text{and} \quad \max\left\{\|A^\dagger\|_2^2, \|B^\dagger\|_2^2\right\} \|E\|_F = 10^6,$$

which means that the bound (2.1) is optimal.

However, when $\operatorname{rank}(B) = \operatorname{rank}(A)$, we can get a sharper bound than (2.1) by replacing the term $\max\left\{\|A^\dagger\|_2^2, \|B^\dagger\|_2^2\right\}$ with the product $\|A^\dagger\|_2\|B^\dagger\|_2$.

Theorem 2.2. *Let A, $B = A + E \in C^{m\times n}$, and $\operatorname{rank}(B) = \operatorname{rank}(A) = r$. Then we have*

$$\|B^\dagger - A^\dagger\|_F \leqslant \|A^\dagger\|_2\|B^\dagger\|_2\|E\|_F. \tag{2.15}$$

Proof. Let A and B have the SVDs (2.2). Then $\widetilde{U}^*U$ is an unitary matrix and can be expressed as the following block form:

$$\widetilde{U}^*U = \begin{pmatrix} \widetilde{U}_1^*U_1 & \widetilde{U}_1^*U_2 \\ \widetilde{U}_2^*U_1 & \widetilde{U}_2^*U_2 \end{pmatrix}.$$

Thus by Lemma 1.3, we have

$$\|\widetilde{U}_1^*U_2\|_F = \|\widetilde{U}_2^*U_1\|_F. \tag{2.16}$$

Similarly, by unitarity of the matrices $U^*\widetilde{U}$, $\widetilde{V}^*V$ and $V^*\widetilde{V}$, we easily know that

$$\|U_2^*\widetilde{U}_1\|_F = \|U_1^*\widetilde{U}_2\|_F, \quad \|\widetilde{V}_1^*V_2\|_F = \|\widetilde{V}_2^*V_1\|_F \text{and} \quad \|V_2^*\widetilde{V}_1\|_F = \|V_1^*\widetilde{V}_2\|_F. \tag{2.17}$$

From (2.7), (2.8), (2.13), (2.16), (2.17) and Lemma 1.2, we have

$$\begin{aligned}
2\|B^\dagger - A^\dagger\|_F^2 \leqslant{} & \|\widetilde{\Sigma}_1^{-1}\widetilde{U}_1^*EV_1\Sigma_1^{-1}\|_F^2 + \|\Sigma_1^{-1}U_1^*E\widetilde{V}_1\widetilde{\Sigma}_1^{-1}\|_F^2 + \frac{1}{\tilde{\sigma}_r^2}\left(\|\widetilde{U}_1^*U_2\|_F^2 + \|V_2^*\widetilde{V}_1\|_F^2\right) \\
& + \frac{1}{\sigma_r^2}\left(\|\widetilde{V}_2^*V_1\|_F^2 + \|U_1^*\widetilde{U}_2\|_F^2\right) \\
\leqslant{} & \frac{1}{\tilde{\sigma}_r^2\sigma_r^2}\left(\|\widetilde{U}_1^*EV_1\|_F^2 + \|U_1^*E\widetilde{V}_1\|_F^2\right) + \frac{1}{\tilde{\sigma}_r^2}\left(\|\widetilde{U}_2^*U_1\|_F^2 + \|V_1^*\widetilde{V}_2\|_F^2\right) \\
& + \frac{1}{\sigma_r^2}\left(\|\widetilde{V}_1^*V_2\|_F^2 + \|U_2^*\widetilde{U}_1\|_F^2\right) \\
={} & \frac{1}{\tilde{\sigma}_r^2\sigma_r^2}\left(\|\widetilde{U}_1^*EV_1\|_F^2 + \|U_1^*E\widetilde{V}_1\|_F^2\right) + \frac{1}{\tilde{\sigma}_r^2}\left(\|\widetilde{U}_2^*EV_1\Sigma_1^{-1}\|_F^2 + \|\Sigma_1^{-1}U_1^*E\widetilde{V}_2\|_F^2\right) \\
& + \frac{1}{\sigma_r^2}\left(\|\widetilde{\Sigma}_1^{-1}\widetilde{U}_1^*EV_2\|_F^2 + \|U_2^*E\widetilde{V}_1\widetilde{\Sigma}_1^{-1}\|_F^2\right) \\
\leqslant{} & \frac{1}{\tilde{\sigma}_r^2\sigma_r^2}\left(\|\widetilde{U}_1^*EV_1\|_F^2 + \|\widetilde{U}_2^*EV_1\|_F^2 + \|\widetilde{U}_1^*EV_2\|_F^2 + \|U_1^*E\widetilde{V}_1\|_F^2\right. \\
& \left. + \|U_1^*E\widetilde{V}_2\|_F^2 + \|U_2^*E\widetilde{V}_1\|_F^2\right) \\
\leqslant{} & \frac{2}{\tilde{\sigma}_r^2\sigma_r^2}\|E\|_F^2 = 2\|A^\dagger\|_2^2\|B^\dagger\|_2^2\|E\|_F^2.
\end{aligned}$$

Hence the inequality (2.15) is true. $\square$

Next example shows that the perturbation bound (2.15) is also optimal for the case $\mathrm{rank}(B) = \mathrm{rank}(A)$.

Example 2. Let $A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$, $E = \begin{pmatrix} 0.001 & 0 \\ 0 & 0 \end{pmatrix}$. Then $B = \begin{pmatrix} 1.001 & 0 \\ 0 & 0 \end{pmatrix}$, $A^\dagger = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$ and $B^\dagger = \begin{pmatrix} \frac{1000}{1001} & 0 \\ 0 & 0 \end{pmatrix}$. By simple computations, we know that

$$\|B^\dagger - A^\dagger\|_F = \frac{1}{1001} \quad \text{and} \quad \|A^\dagger\|_2 \|B^\dagger\|_2 \|E\|_F = \frac{1}{1001},$$

which implies that the bound (2.15) is optimal.

Obviously, the perturbation bounds (2.1) in Theorem 2.1 and (2.15) in Theorem 2.2 are, respectively, sharper than the known bounds (1.1) and (1.2) given in Theorem 1.1, for the case of the F-norm. That is to say, the relevant μ in (1.1) and (1.2) can be uniformly reduced into 1 instead of $\sqrt{2}$ and in this case, we get the optimal perturbation bounds of Moore–Penrose inverse under the F-norm.

3. Multiplicative perturbation bound

As a special additive perturbation model, the multiplicative perturbation model has been considered by many authors, see [14,15,16,17,18]. However, it seems that the perturbation bound of the Moore–Penrose inverse with respect to the multiplicative perturbation model has not been discussed yet in the literature. In this section, we will present a multiplicative perturbation bound of the Moore–Penrose inverse under the F-norm.

Theorem 3.1. *Suppose A is an $m \times n$ matrix and $B = D_1^* A D_2$, where D_1 and D_2 are respectively $m \times m$ and $n \times n$ nonsingular matrices. Then*

$$\|B^\dagger - A^\dagger\|_F \leqslant \max\left\{\|A^\dagger\|_2, \|B^\dagger\|_2\right\} \Phi(D_1, D_2), \tag{3.1}$$

where $\Phi(D_1, D_2) = \sqrt{\|I_m - D_1\|_F^2 + \|I_m - D_1^{-1}\|_F^2 + \|I_n - D_2\|_F^2 + \|I_n - D_2^{-1}\|_F^2}$.

Proof. Since $B = D_1^* A D_2$, it is easy to know that

$$B - A = B(I_n - D_2^{-1}) + (D_1^* - I_m)A = (I_m - D_1^{-*})B + A(D_2 - I_n), \tag{3.2}$$

where D_1^{-*} denotes the inverse of the conjugate transpose of D_1. If A and B have the SVDs (2.2), then the identities (3.2) can be expresses as follows:

$$\begin{aligned} \widetilde{U}_1 \widetilde{\Sigma}_1 \widetilde{V}_1^* - U_1 \Sigma_1 V_1^* &= \widetilde{U}_1 \widetilde{\Sigma}_1 \widetilde{V}_1^* (I_n - D_2^{-1}) + (D_1^* - I_m) U_1 \Sigma_1 V_1^* \\ &= (I_m - D_1^{-*}) \widetilde{U}_1 \widetilde{\Sigma}_1 \widetilde{V}_1^* + U_1 \Sigma_1 V_1^* (D_2 - I_n). \end{aligned} \tag{3.3}$$

Using the identities in (2.4) and (3.3), we can get the following equalities:

$$\widetilde{\Sigma}_1 \widetilde{V}_1^* V_1 - \widetilde{U}_1^* U_1 \Sigma_1 = \widetilde{\Sigma}_1 \widetilde{V}_1^* (I_n - D_2^{-1}) V_1 + \widetilde{U}_1^* (D_1^* - I_m) U_1 \Sigma_1, \tag{3.4}$$

$$U_1^* \widetilde{U}_1 \widetilde{\Sigma}_1 - \Sigma_1 V_1^* \widetilde{V}_1 = U_1^* (I_m - D_1^{-*}) \widetilde{U}_1 \widetilde{\Sigma}_1 + \Sigma_1 V_1^* (D_2 - I_n) \widetilde{V}_1, \tag{3.5}$$

$$\widetilde{U}_2^* U_1 \Sigma_1 = -\widetilde{U}_2^* (D_1^* - I_m) U_1 \Sigma_1, U_2^* \widetilde{U}_1 \widetilde{\Sigma}_1 = U_2^* (I_m - D_1^{-*}) \widetilde{U}_1 \widetilde{\Sigma}_1 \tag{3.6}$$

and

$$\widetilde{\Sigma}_1 \widetilde{V}_1^* V_2 = \widetilde{\Sigma}_1 \widetilde{V}_1^* (I_n - D_2^{-1}) V_2, \quad \Sigma_1 V_1^* \widetilde{V}_2 = -\Sigma_1 V_1^* (D_2 - I_n) \widetilde{V}_2. \tag{3.7}$$

The identities (3.4) and (3.5) can be, respectively, reduced into

$$\widetilde{V}_1^* V_1 \Sigma_1^{-1} - \widetilde{\Sigma}_1^{-1} \widetilde{U}_1^* U_1 = \widetilde{V}_1^* \left(I_n - D_2^{-1}\right) V_1 \Sigma_1^{-1} + \widetilde{\Sigma}_1^{-1} \widetilde{U}_1^* (D_1^* - I_m) U_1 \tag{3.8}$$

and

$$\Sigma_1^{-1}U_1^*\widetilde{U}_1 - V_1^*\widetilde{V}_1\widetilde{\Sigma}_1^{-1} = \Sigma_1^{-1}U_1^*(I_m - D_1^{-*})\widetilde{U}_1 + V_1^*(D_2 - I_n)\widetilde{V}_1\widetilde{\Sigma}_1^{-1}. \tag{3.9}$$

Therefore, by (2.13), (3.6)–(3.9) and Lemma 1.2, we have

$$\begin{aligned}
&2\|B^\dagger - A^\dagger\|_F^2 \\
&= \|\widetilde{V}_1^*(I_n - D_2^{-1})V_1\Sigma_1^{-1} + \widetilde{\Sigma}_1^{-1}\widetilde{U}_1^*(D_1^* - I_m)U_1\|_F^2 + \|\widetilde{\Sigma}_1^{-1}\widetilde{U}_1^*(I_m - D_1^{-1})U_2\|_F^2 \\
&\quad + \|\widetilde{V}_2^*(D_2^* - I_n)V_1\Sigma_1^{-1}\|_F^2 + \|\Sigma_1^{-1}U_1^*(I_m - D_1^{-*})\widetilde{U}_1 + V_1^*(D_2 - I_n)\widetilde{V}_1\widetilde{\Sigma}_1^{-1}\|_F^2 \\
&\quad + \|\Sigma_1^{-1}U_1^*(D_1 - I_m)\widetilde{U}_2\|_F^2 + \|V_2^*(I_n - D_2^{-*})\widetilde{V}_1\widetilde{\Sigma}_1^{-1}\|_F^2 \\
&\leqslant \left(\frac{1}{\sigma_r}\|\widetilde{V}_1^*(I_n - D_2^{-1})V_1\|_F + \frac{1}{\tilde{\sigma}_r}\|\widetilde{U}_1^*(D_1^* - I_m)U_1\|_F\right)^2 + \left(\frac{1}{\sigma_r}\|U_1^*(I_m - D_1^{-*})\widetilde{U}_1\|_F\right. \\
&\quad \left. + \frac{1}{\tilde{\sigma}_r}\|V_1^*(D_2 - I_n)\widetilde{V}_1\|_F\right)^2 + \frac{1}{\tilde{\sigma}_r^2}\left(\|\widetilde{U}_1^*(I_m - D_1^{-1})U_2\|_F^2 + \|V_2^*(I_n - D_2^{-*})\widetilde{V}_1\|_F^2\right) \\
&\quad + \frac{1}{\sigma_r^2}\left(\|\widetilde{V}_2^*(D_2^* - I_n)V_1\|_F^2 + \|U_1^*(D_1 - I_m)\widetilde{U}_2\|_F^2\right) \\
&\leqslant \left(\frac{2}{\sigma_r^2}\|\widetilde{V}_1^*(I_n - D_2^{-1})V_1\|_F^2 + \frac{2}{\tilde{\sigma}_r^2}\|\widetilde{U}_1^*(D_1^* - I_m)U_1\|_F^2\right) \\
&\quad + \left(\frac{2}{\sigma_r^2}\|U_1^*(I_m - D_1^{-*})\widetilde{U}_1\|_F^2 + \frac{2}{\tilde{\sigma}_r^2}\|V_1^*(D_2 - I_n)\widetilde{V}_1\|_F^2\right) \\
&\quad + \frac{1}{\tilde{\sigma}_r^2}\left(\|\widetilde{U}_1^*(I_m - D_1^{-1})U_2\|_F^2 + \|V_2^*(I_n - D_2^{-*})\widetilde{V}_1\|_F^2\right) \\
&\quad + \frac{1}{\sigma_r^2}\left(\|\widetilde{V}_2^*(D_2^* - I_n)V_1\|_F^2 + \|U_1^*(D_1 - I_m)\widetilde{U}_2\|_F^2\right) \\
&\leqslant \max\left\{\frac{1}{\tilde{\sigma}_r^2}, \frac{1}{\sigma_r^2}\right\}\Big[(2\|\widetilde{V}_1^*(I_n - D_2^{-1})V_1\|_F^2 + \|\widetilde{V}_1^*(I_n - D_2^{-1})V_2\|_F^2) + (2\|\widetilde{U}_1^*(D_1^* - I_m)U_1\|_F^2 \\
&\quad + \|\widetilde{U}_2^*(D_1^* - I_m)U_1\|_F^2) + (2\|U_1^*(I_m - D_1^{-*})\widetilde{U}_1\|_F^2 + \|U_2^*(I_m - D_1^{-*})\widetilde{U}_1\|_F^2) \\
&\quad + \left(2\|V_1^*(D_2 - I_n)\widetilde{V}_1\|_F^2 + \|V_1^*(D_2 - I_n)\widetilde{V}_2\|_F^2\right)\Big] \\
&\leqslant 2\max\left\{\frac{1}{\tilde{\sigma}_r^2}, \frac{1}{\sigma_r^2}\right\}\left(\|I_m - D_1\|_F^2 + \|I_m - D_1^{-1}\|_F^2 + \|I_n - D_2\|_F^2 + \|I_n - D_2^{-1}\|_F^2\right) \\
&= 2\max\left\{\|A^\dagger\|_2^2, \|B^\dagger\|_2^2\right\}\Phi^2(D_1, D_2).
\end{aligned} \tag{3.10}$$

Hence, we have

$$\|B^\dagger - A^\dagger\|_F \leqslant \max\left\{\|A^\dagger\|_2, \|B^\dagger\|_2\right\}\Phi(D_1, D_2).$$

The proof is completed. □

Remark. Since $B = D_1^*AD_2$ can be rewritten as $B = A + E$ with $E = -(I_m - D_1^*)A - D_1^*A(I_n - D_2)$ or $E = -A(I_n - D_2) - (I_m - D_1^*)AD_2$, the multiplicative perturbation is a special additive perturbation. Hence the additive perturbation bound (2.15) can also be applied to the multiplicative perturbation model. But it is difficult to compare theoretically the multiplicative perturbation bound (3.1) and the additive perturbation bound (2.15). However, in some cases, for example, when the smallest singular values of A and B are less than 1, i.e., $\sigma_r < 1$ and $\tilde{\sigma}_r < 1$, the bound (3.1) is generally better than the bound (2.15). The reason is that the coefficient term $\max\left\{\|A^\dagger\|_2, \|B^\dagger\|_2\right\}$ in (3.1) is less than the coefficient term $\|A^\dagger\|_2\|B^\dagger\|_2$ in (2.15) in this case.

Example 3. Let $A = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 0.001 & 0 \\ 0 & 0 & 0 \end{pmatrix}$, $D_2 = \begin{pmatrix} 1.001 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ and $D_1 = I_3$. Then $B = \begin{pmatrix} 2.002 & 0 & 0 \\ 0 & 0.001 & 0 \\ 0 & 0 & 0 \end{pmatrix}$ and $E = B - A = \begin{pmatrix} 0.002 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$. By simple computations, we know that

$$\|A^\dagger\|_2 \|B^\dagger\|_2 \|E\|_F = 2000$$

and

$$\max\left\{\|A^\dagger\|_2, \|B^\dagger\|_2\right\} \Phi(D_1, D_2) \approx \sqrt{2}.$$

Thus the perturbation bounds (2.15) and (3.1) give the following inequalities:

$$\|B^\dagger - A^\dagger\|_F \leqslant 2000 \quad \text{(additive perturbation bound)}$$

and

$$\|B^\dagger - A^\dagger\|_F \leqslant \sqrt{2} \quad \text{(multiplicative perturbation bound)}.$$

Obviously, the multiplicative perturbation bound (3.1) is occasionally much better than the additive perturbation bound (2.15).

References

[1] A. Ben-Israel, T.N.E. Greville, Generalized Inverse: Theory and Applications, Wiley-Interscience, 1974, second ed., Springer-Verlag, New York, 2002.
[2] G. Wang, Y. Wei, S. Qiao, Generalized Inverses: Theory and Computations, Science Press, Beijing, 2004.
[3] A. Ben-Israel, On error bounds for generalized inverses, SIAM J. Numer. Anal. 3 (1966) 585–592.
[4] G.W. Stewart, On the continuity of the generalized inverses, SIAM J. Appl. Math. 17 (1969) 33–45.
[5] P.-Å. Wedin, On pseudoinverses of perturbed matrices, Lund Un. Comp. Sc. Tech. Rep. 1969.
[6] G.W. Stewart, On the perturbation of the pseudo-inverse, projections, and linear squares problems, SIAM J. Rev. 19 (1977) 634–662.
[7] P.-Å. Wedin, Perturbation bounds in connection with singular value decomposition, BIT 12 (1972) 99–111.
[8] P.-Å. Wedin, Perturbation theory for pseudo-inverses, BIT 13 (1973) 217–232.
[9] N.N. Abdelmalek, On the solution of the linear least squares problems and pseudo-inverses, Computing 13 (1974) 215–228.
[10] V. Pereyra, Stability of general systems of linear equations, Aequat. Math. 2 (1969) 194–206.
[11] G.W. Stewart, J.G. Sun, Matrix Perturbation Theory, Academic Press, Boston, 1990.
[12] X.S. Chen, W. Li, A note on the perturbation bounds of eigenspaces for Hermitian matrices, J. Comput. Appl. Math. 196 (2006) 338–346.
[13] X.S. Chen, W. Li, Relative perturbation bounds for the subunitary polar factor under unitarily invariant norms, Adv. Math. 35 (2006) 178–184. (in Chinese)
[14] R.C. Li, Relative perturbation bounds for the unitary polar factor, BIT 37 (1997) 67–75.
[15] R.C. Li, Relative perturbation theory: (II) Eigenspace and singular subspace variations, SIAM.J. Matrix Anal. Appl. 20 (1999) 471–492.
[16] R.C. Li, Relative perturbation theory: (IV) sin 2θ theorems, Linear Algebra Appl. 311 (2000) 45–60.
[17] X.S. Chen, W. Li, W.W. Sun, Some new perturbation bounds for the generalized polar decomposition, BIT 44 (2004) 237–244.
[18] W. Li, Multiplicative perturbation bounds for spectral and singular value decompositions, J. Comput. Appl. Math. 217 (2008) 243–251.