On the perturbation of the Moore–Penrose inverse of a matrix

Xuefeng Xu[*]

January 3, 2020

Abstract

The Moore–Penrose inverse of a matrix has been extensively investigated and widely applied in many fields over the past decades. One reason for the interest is that the Moore–Penrose inverse can succinctly express some important geometric constructions in finite-dimensional spaces, such as the orthogonal projection onto a subspace and the linear least squares problem. In this paper, we establish new perturbation bounds for the Moore–Penrose inverse under the Frobenius norm, some of which are sharper than the existing ones.

Keywords: Moore–Penrose inverse, perturbation, singular value decomposition

AMS subject classifications: 15A09, 15A18, 47A55, 65F35

1 Introduction

Let $\mathbb{C}^{m \times n}$, $\mathbb{C}_r^{m \times n}$, and $\mathscr{U}_n$ be the set of all $m \times n$ complex matrices, the set of all $m \times n$ complex matrices of rank r, and the set of all $n \times n$ unitary matrices, respectively. For any $M \in \mathbb{C}^{m \times n}$, the symbols M^*, $\mathrm{rank}(M)$, $\|M\|_{\mathscr{U}}$, $\|M\|_2$, and $\|M\|_F$ denote the conjugate transpose, the rank, the unitarily invariant norm, the spectral norm, and the Frobenius norm of M, respectively.

The Moore–Penrose (MP) inverse of $M \in \mathbb{C}^{m \times n}$ is denoted by $M^\dagger$, which is defined as the unique matrix $X \in \mathbb{C}^{n \times m}$ satisfying the following equations [10, 11]:

$$\text{(i) } MXM = M, \quad \text{(ii) } XMX = X, \quad \text{(iii) } (MX)^* = MX, \quad \text{(iv) } (XM)^* = XM.$$

In particular, if M is a square and nonsingular matrix, then $M^\dagger$ will coincide with the usual inverse M^{-1}. The MP inverse can concisely express some important geometric constructions in finite-dimensional spaces, such as the orthogonal projection onto a subspace and the linear least squares problem. More specifically, the orthogonal projection onto the column space of A can be expressed as $P_A = AA^\dagger$; see [18] for the perturbation analysis of P_A. Recall that the linear least squares problem can be described as follows: Find $\mathbf{x}_\star \in \mathbb{C}^n$ such that

$$\mathbf{x}_\star \in \underset{\mathbf{x} \in \mathbb{C}^n}{\arg\min} \|A\mathbf{x} - \mathbf{b}\|_2, \tag{1.1}$$

---

[*]Department of Mathematics, Purdue University, West Lafayette, IN 47907, USA (`xuxuefeng@lsec.cc.ac.cn`; `xu1412@purdue.edu`).

where $A \in \mathbb{C}^{m \times n}$ and $\mathbf{b} \in \mathbb{C}^m$. It is well known that the solutions of (1.1) can be formulated as

$$\mathbf{x}_\star = A^\dagger \mathbf{b} + (I_n - A^\dagger A)\mathbf{z},$$

where I_n denotes the $n \times n$ identity matrix and $\mathbf{z} \in \mathbb{C}^n$ is an arbitrary vector. Furthermore, the minimum 2-norm solution of (1.1) is $\mathbf{x}_\star = A^\dagger \mathbf{b}$. The MP inverse has been widely applied in many fields such as matrix computation, algorithm analysis, statistics, and engineering; see, e.g., [2, 6, 7, 3]. Over the past decades, many researchers have investigated the perturbation analysis of MP inverse, which can be found, e.g., in [12, 16, 1, 13, 14, 9, 4, 17, 8].

Let $A \in \mathbb{C}_r^{m \times n}$, $B \in \mathbb{C}_s^{m \times n}$, and $E = B - A$. Wedin [16] established the estimate (see also [14, 9])

$$\|B^\dagger - A^\dagger\| \le \mu_1 \max \left\{\|A^\dagger\|_2^2, \|B^\dagger\|_2^2\right\}\|E\|. \tag{1.2}$$

In particular, if $s = r$, then

$$\|B^\dagger - A^\dagger\| \le \mu_2 \|A^\dagger\|_2 \|B^\dagger\|_2 \|E\|. \tag{1.3}$$

The above parameters μ_1 and μ_2 are listed in Table 1. In 2010, Meng and Zheng [9, Theorems 2.1

$\| \cdot \|$	μ_1	$\mu_2 \ (r < \min\{m,n\})$	$\mu_2 \left(r = \min\limits_{m \ne n}\{m,n\}\right)$	$\mu_2 \ (r = m = n)$
$\| \cdot \|_{\mathscr{U}}$	3	3	2	1
$\| \cdot \|_2$	$\frac{1+\sqrt{5}}{2}$	$\frac{1+\sqrt{5}}{2}$	$\sqrt{2}$	1
$\| \cdot \|_F$	$\sqrt{2}$	$\sqrt{2}$	1	1

Table 1: The values of μ_1 and μ_2.

and 2.2] improved the estimates (1.2) and (1.3) under the Frobenius norm. More specifically, they derived that

$$\|B^\dagger - A^\dagger\|_F \le \max \left\{\|A^\dagger\|_2^2, \|B^\dagger\|_2^2\right\}\|E\|_F. \tag{1.4}$$

In particular, if $s = r$, then

$$\|B^\dagger - A^\dagger\|_F \le \|A^\dagger\|_2 \|B^\dagger\|_2 \|E\|_F. \tag{1.5}$$

Recently, Li et al. [8, Theorem 3.1] further refined the estimate (1.4). They obtained that

$$\|B^\dagger - A^\dagger\|_F^2 \le \max \left\{\|A^\dagger\|_2^4, \|B^\dagger\|_2^4\right\}\|E\|_F^2 - \frac{\|A^\dagger E B^\dagger\|_F^2 + \|B^\dagger E A^\dagger\|_F^2}{2} \left(\max \left\{\frac{\|A^\dagger\|_2^2}{\|B^\dagger\|_2^2}, \frac{\|B^\dagger\|_2^2}{\|A^\dagger\|_2^2}\right\} - 1\right). \tag{1.6}$$

If $A \in \mathbb{C}_n^{m \times n} \ (m \ge n)$ and $B = A + E \in \mathbb{C}_s^{m \times n}$, Li et al. [8, Theorem 3.2] also proved that

$$\|B^\dagger - A^\dagger\|_F^2 \le \frac{\|A^\dagger\|_2^2 \|B^\dagger\|_2^2}{\|A^\dagger\|_2^2 + \|B^\dagger\|_2^2} \left(\|E A^\dagger\|_F^2 + \|E B^\dagger\|_F^2 + (n - s)\frac{\|A^\dagger\|_2^2}{\|B^\dagger\|_2^2}\right). \tag{1.7}$$

In particular, if $s = n$, then

$$\|B^\dagger - A^\dagger\|_F^2 \le \min \left\{\|B^\dagger\|_2^2 \|E A^\dagger\|_F^2, \|A^\dagger\|_2^2 \|E B^\dagger\|_F^2\right\}. \tag{1.8}$$

2

Although the estimate (1.6) has sharpened (1.4), the upper bound in (1.6) is still too large in certain cases. We now give a simple example:

$$A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad B = \begin{pmatrix} \frac{1}{1+2\tau} & 0 \\ 0 & \tau \end{pmatrix}, \tag{1.9}$$

where $0 < \tau < \frac{1}{2}$. In this example, it holds that

$$\|B^\dagger - A^\dagger\|_F^2 = 4\tau^2 + \frac{1}{\tau^2}.$$

Direct computation yields that the upper bound in (1.6) is

$$4\tau^2 + \frac{1}{\tau^2} + \frac{4}{\tau^2(1+2\tau)^2} - 4 =: u(\tau).$$

It is easy to see that

$$u(\tau) - \|B^\dagger - A^\dagger\|_F^2 = \frac{4}{\tau^2(1+2\tau)^2} - 4,$$

which will be very large if $0 < \tau \ll \frac{1}{2}$. Moreover, if τ is sufficiently small, then

$$u(\tau) \simeq \frac{5}{\tau^2} \simeq 5\|B^\dagger - A^\dagger\|_F^2.$$

Motivated by the above observation, we revisit the perturbation of MP inverse under the Frobenius norm. Some new upper bounds for $\|B^\dagger - A^\dagger\|_F^2$ are presented. Theoretical analysis shows that the new bounds are sharper than the existing ones.

The rest of this paper is organized as follows. In Section 2, we introduce a trace inequality and several auxiliary results on $\|B^\dagger - A^\dagger\|_F^2$. In Section 3, we present some new upper bounds for $\|B^\dagger - A^\dagger\|_F^2$, and compare the new bounds with the existing ones theoretically.

2 Preliminaries

Let $\{\sigma_i(M)\}_{i=1}^t$ and $\{\sigma_i(N)\}_{i=1}^t$ be the singular values of $M \in \mathbb{C}^{m \times n}$ and $N \in \mathbb{C}^{m \times n}$, respectively, where $t = \min\{m, n\}$. Assume that $\{\sigma_i(M)\}_{i=1}^t$ and $\{\sigma_i(N)\}_{i=1}^t$ are arranged in the same (increasing or decreasing) order. The celebrated von Neumann's trace inequality [15] reads

$$\operatorname{Re} \operatorname{tr}(UMVN^*) \leq \sum_{i=1}^t \sigma_i(M)\sigma_i(N),$$

where $\operatorname{Re} \operatorname{tr}(\cdot)$ denotes the real part of the trace of a matrix, and both $U \in \mathscr{U}_m$ and $V \in \mathscr{U}_n$ are arbitrary. Indeed, the following more accurate characterization for $\operatorname{Re} \operatorname{tr}(UMVN^*)$ [15] holds.

Lemma 2.1. *Let $M \in \mathbb{C}^{m \times n}$, $N \in \mathbb{C}^{m \times n}$, and $t = \min\{m, n\}$. Let $\{\sigma_i(M)\}_{i=1}^t$ and $\{\sigma_i(N)\}_{i=1}^t$ be the singular values of M and N, respectively, which are arranged in the same (increasing or decreasing) order. Then*

$$\max_{\substack{U \in \mathscr{U}_m \\ V \in \mathscr{U}_n}} \operatorname{Re} \operatorname{tr}(UMVN^*) = \sum_{i=1}^t \sigma_i(M)\sigma_i(N). \tag{2.1}$$

Based on the singular value decomposition (SVD) of a matrix, we can derive two characterizations of $\|B^\dagger - A^\dagger\|_F^2$ (see Lemma 2.2), which play a fundamental role in our analysis. Let $A \in \mathbb{C}_r^{m \times n}$ and $B \in \mathbb{C}_s^{m \times n}$ (*throughout this paper, we only consider the nontrivial case that $r \geq 1$ and $s \geq 1$*) have the following SVDs:

$$A = U \begin{pmatrix} \Sigma_1 & 0 \\ 0 & 0 \end{pmatrix} V^* = U_1 \Sigma_1 V_1^*, \tag{2.2a}$$

$$B = \widetilde{U} \begin{pmatrix} \widetilde{\Sigma}_1 & 0 \\ 0 & 0 \end{pmatrix} \widetilde{V}^* = \widetilde{U}_1 \widetilde{\Sigma}_1 \widetilde{V}_1^*, \tag{2.2b}$$

where $U = (U_1, U_2) \in \mathscr{U}_m$, $V = (V_1, V_2) \in \mathscr{U}_n$, $\widetilde{U} = (\widetilde{U}_1, \widetilde{U}_2) \in \mathscr{U}_m$, $\widetilde{V} = (\widetilde{V}_1, \widetilde{V}_2) \in \mathscr{U}_n$, $U_1 \in \mathbb{C}^{m \times r}$, $V_1 \in \mathbb{C}^{n \times r}$, $\widetilde{U}_1 \in \mathbb{C}^{m \times s}$, $\widetilde{V}_1 \in \mathbb{C}^{n \times s}$, $\Sigma_1 = \mathrm{diag}(\sigma_1, \ldots, \sigma_r)$, $\widetilde{\Sigma}_1 = \mathrm{diag}(\widetilde{\sigma}_1, \ldots, \widetilde{\sigma}_s)$, $\sigma_1 \geq \cdots \geq \sigma_r > 0$, and $\widetilde{\sigma}_1 \geq \cdots \geq \widetilde{\sigma}_s > 0$. In view of (2.2a) and (2.2b), the MP inverses $A^\dagger$ and $B^\dagger$ can be explicitly expressed as follows:

$$A^\dagger = V \begin{pmatrix} \Sigma_1^{-1} & 0 \\ 0 & 0 \end{pmatrix} U^* = V_1 \Sigma_1^{-1} U_1^*, \tag{2.3a}$$

$$B^\dagger = \widetilde{V} \begin{pmatrix} \widetilde{\Sigma}_1^{-1} & 0 \\ 0 & 0 \end{pmatrix} \widetilde{U}^* = \widetilde{V}_1 \widetilde{\Sigma}_1^{-1} \widetilde{U}_1^*. \tag{2.3b}$$

Using (2.2a), (2.2b), (2.3a), and (2.3b), we can obtain the following identities.

Lemma 2.2. *Let $A \in \mathbb{C}_r^{m \times n}$ and $B \in \mathbb{C}_s^{m \times n}$ have the SVDs (2.2a) and (2.2b), respectively, and let $E = B - A$. Then*

$$\|B^\dagger - A^\dagger\|_F^2 = \|\widetilde{\Sigma}_1^{-1} \widetilde{U}_1^* U_2\|_F^2 + \|\widetilde{V}_2^* V_1 \Sigma_1^{-1}\|_F^2 + \|B^\dagger E A^\dagger\|_F^2, \tag{2.4a}$$

$$\|B^\dagger - A^\dagger\|_F^2 = \|\widetilde{U}_2^* U_1 \Sigma_1^{-1}\|_F^2 + \|\widetilde{\Sigma}_1^{-1} \widetilde{V}_1^* V_2\|_F^2 + \|A^\dagger E B^\dagger\|_F^2. \tag{2.4b}$$

Proof. By (2.3a) and (2.3b), we have

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

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

Then

$$\|B^\dagger - A^\dagger\|_F^2 = \|\widetilde{\Sigma}_1^{-1} \widetilde{U}_1^* U_2\|_F^2 + \|\widetilde{V}_2^* V_1 \Sigma_1^{-1}\|_F^2 + \|\widetilde{\Sigma}_1^{-1} \widetilde{U}_1^* U_1 - \widetilde{V}_1^* V_1 \Sigma_1^{-1}\|_F^2, \tag{2.5a}$$

$$\|B^\dagger - A^\dagger\|_F^2 = \|\Sigma_1^{-1} U_1^* \widetilde{U}_2\|_F^2 + \|V_2^* \widetilde{V}_1 \widetilde{\Sigma}_1^{-1}\|_F^2 + \|V_1^* \widetilde{V}_1 \widetilde{\Sigma}_1^{-1} - \Sigma_1^{-1} U_1^* \widetilde{U}_1\|_F^2. \tag{2.5b}$$

In addition, using (2.2a), (2.2b), (2.3a), and (2.3b), we get

$$\widetilde{V}^* B^\dagger E A^\dagger U = \begin{pmatrix} \widetilde{V}_1^* V_1 \Sigma_1^{-1} - \widetilde{\Sigma}_1^{-1} \widetilde{U}_1^* U_1 & 0 \\ 0 & 0 \end{pmatrix},$$

$$V^* A^\dagger E B^\dagger \widetilde{U} = \begin{pmatrix} \Sigma_1^{-1} U_1^* \widetilde{U}_1 - V_1^* \widetilde{V}_1 \widetilde{\Sigma}_1^{-1} & 0 \\ 0 & 0 \end{pmatrix}.$$

Hence,

$$\|B^\dagger E A^\dagger\|_F^2 = \|\widetilde{V}_1^* V_1 \Sigma_1^{-1} - \widetilde{\Sigma}_1^{-1} \widetilde{U}_1^* U_1\|_F^2, \tag{2.6a}$$

$$\|A^\dagger E B^\dagger\|_F^2 = \|\Sigma_1^{-1} U_1^* \widetilde{U}_1 - V_1^* \widetilde{V}_1 \widetilde{\Sigma}_1^{-1}\|_F^2. \tag{2.6b}$$

Combining (2.5a) and (2.6a), we can arrive at the identity (2.4a). Similarly, the identity (2.4b) follows immediately from (2.5b) and (2.6b). $\qquad\square$

The following corollary can be directly deduced from Lemma 2.2.

Corollary 2.1. *Under the assumptions of Lemma* 2.2, *we have*

$$\|B^\dagger - A^\dagger\|_F^2 \leq \|B^\dagger\|_2^2 \|\widetilde{U}_1^* U_2\|_F^2 + \|A^\dagger\|_2^2 \|\widetilde{V}_2^* V_1\|_F^2 + \|B^\dagger E A^\dagger\|_F^2, \tag{2.7a}$$

$$\|B^\dagger - A^\dagger\|_F^2 \leq \|A^\dagger\|_2^2 \|\widetilde{U}_2^* U_1\|_F^2 + \|B^\dagger\|_2^2 \|\widetilde{V}_1^* V_2\|_F^2 + \|A^\dagger E B^\dagger\|_F^2. \tag{2.7b}$$

If $\mathrm{rank}(B) = \mathrm{rank}(A)$, then $\|\widetilde{U}_1^* U_2\|_F$, $\|\widetilde{U}_2^* U_1\|_F$, $\|\widetilde{V}_1^* V_2\|_F$, and $\|\widetilde{V}_2^* V_1\|_F$ have the following relations [5, Lemma 2.2].

Lemma 2.3. *Let* $A \in \mathbb{C}_r^{m \times n}$ *and* $B \in \mathbb{C}_s^{m \times n}$ *have the SVDs* (2.2a) *and* (2.2b), *respectively. If* $s = r$, *then*

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

3 Main results

In this section, we develop some new perturbation bounds for $\|B^\dagger - A^\dagger\|_F^2$. The first estimate depends only on the positive singular values of A and B.

Theorem 3.1. *Let* $A \in \mathbb{C}_r^{m \times n}$ *and* $B \in \mathbb{C}_s^{m \times n}$ *have the positive singular values* $\{\sigma_i\}_{i=1}^r$ *and* $\{\widetilde{\sigma}_i\}_{i=1}^s$, *respectively, where* $\sigma_1 \geq \cdots \geq \sigma_r$ *and* $\widetilde{\sigma}_1 \geq \cdots \geq \widetilde{\sigma}_s$.
  (i) *If* $s \leq r$, *then*

$$\sum_{i=1}^s \left(\frac{1}{\sigma_i} - \frac{1}{\widetilde{\sigma}_i}\right)^2 + \sum_{i=s+1}^r \frac{1}{\sigma_i^2} \leq \|B^\dagger - A^\dagger\|_F^2 \leq \sum_{i=1}^s \left(\frac{1}{\sigma_i} + \frac{1}{\widetilde{\sigma}_i}\right)^2 + \sum_{i=s+1}^r \frac{1}{\sigma_i^2}. \tag{3.1}$$

  (ii) *If* $s > r$, *then*

$$\sum_{i=1}^r \left(\frac{1}{\sigma_i} - \frac{1}{\widetilde{\sigma}_i}\right)^2 + \sum_{i=r+1}^s \frac{1}{\widetilde{\sigma}_i^2} \leq \|B^\dagger - A^\dagger\|_F^2 \leq \sum_{i=1}^r \left(\frac{1}{\sigma_i} + \frac{1}{\widetilde{\sigma}_i}\right)^2 + \sum_{i=r+1}^s \frac{1}{\widetilde{\sigma}_i^2}. \tag{3.2}$$

Proof. Observe first that

$$\|B^\dagger - A^\dagger\|_F^2 = \mathrm{tr}\left((B^\dagger - A^\dagger)^*(B^\dagger - A^\dagger)\right) = \sum_{i=1}^r \frac{1}{\sigma_i^2} + \sum_{i=1}^s \frac{1}{\widetilde{\sigma}_i^2} - 2\,\mathrm{Re}\,\mathrm{tr}\left(A^\dagger(B^\dagger)^*\right). \tag{3.3}$$

From (2.1), we deduce that

$$-\sum_{i=1}^{\min\{s,r\}} \frac{1}{\sigma_i \widetilde{\sigma}_i} \leq \mathrm{Re}\,\mathrm{tr}\left(A^\dagger(B^\dagger)^*\right) \leq \sum_{i=1}^{\min\{s,r\}} \frac{1}{\sigma_i \widetilde{\sigma}_i}. \tag{3.4}$$

Combining (3.3) and (3.4), we can obtain the inequalities (3.1) and (3.2). $\qquad\square$

5

Remark 3.1. According to the lower bounds in (3.1) and (3.2), we conclude that a necessary condition for $\lim_{B \to A} B^\dagger = A^\dagger$ (B is viewed as a variable) is that $\operatorname{rank}(B) = \operatorname{rank}(A)$ always holds as B tends to A. In fact, it is also a sufficient condition for $\lim_{B \to A} B^\dagger = A^\dagger$ [12].

Remark 3.2. Under the setting of (1.9), the lower and upper bounds in (3.2) are $4\tau^2 + \frac{1}{\tau^2}$ and $4(1 + \tau)^2 + \frac{1}{\tau^2}$, respectively. Clearly, the lower bound has attained $\|B^\dagger - A^\dagger\|_F^2$, and the upper bound will be very tight when τ is small.

In what follows, we establish some upper bounds involving the perturbation $E = B - A$.

Theorem 3.2. *Let $A \in \mathbb{C}_r^{m \times n}$, $B \in \mathbb{C}_s^{m \times n}$, and $E = B - A$. Then*

$$\|B^\dagger - A^\dagger\|_F^2 \leq \min \left\{ \alpha_1 + \|B^\dagger E A^\dagger\|_F^2, \alpha_2 + \|A^\dagger E B^\dagger\|_F^2 \right\}, \tag{3.5}$$

where

$$\alpha_1 := \|A^\dagger\|_2^2 \left(\|A^\dagger E\|_F^2 - \|A^\dagger E B^\dagger B\|_F^2 \right) + \|B^\dagger\|_2^2 \left(\|E B^\dagger\|_F^2 - \|A A^\dagger E B^\dagger\|_F^2 \right),$$
$$\alpha_2 := \|A^\dagger\|_2^2 \left(\|E A^\dagger\|_F^2 - \|B B^\dagger E A^\dagger\|_F^2 \right) + \|B^\dagger\|_2^2 \left(\|B^\dagger E\|_F^2 - \|B^\dagger E A^\dagger A\|_F^2 \right).$$

Proof. By (2.2a), (2.2b), (2.3a), and (2.3b), we have

$$U^* E B^\dagger \widetilde{U} = \begin{pmatrix} U_1^* \widetilde{U}_1 - \Sigma_1 V_1^* \widetilde{V}_1 \widetilde{\Sigma}_1^{-1} & 0 \\ U_2^* \widetilde{U}_1 & 0 \end{pmatrix},$$

$$U^* A A^\dagger E B^\dagger \widetilde{U} = \begin{pmatrix} U_1^* \widetilde{U}_1 - \Sigma_1 V_1^* \widetilde{V}_1 \widetilde{\Sigma}_1^{-1} & 0 \\ 0 & 0 \end{pmatrix}.$$

Hence,

$$\|\widetilde{U}_1^* U_2\|_F^2 = \|E B^\dagger\|_F^2 - \|A A^\dagger E B^\dagger\|_F^2. \tag{3.6}$$

Similarly, we have

$$V^* A^\dagger E \widetilde{V} = \begin{pmatrix} \Sigma_1^{-1} U_1^* \widetilde{U}_1 \widetilde{\Sigma}_1 - V_1^* \widetilde{V}_1 & -V_1^* \widetilde{V}_2 \\ 0 & 0 \end{pmatrix},$$

$$V^* A^\dagger E B^\dagger B \widetilde{V} = \begin{pmatrix} \Sigma_1^{-1} U_1^* \widetilde{U}_1 \widetilde{\Sigma}_1 - V_1^* \widetilde{V}_1 & 0 \\ 0 & 0 \end{pmatrix}.$$

Thus,

$$\|\widetilde{V}_2^* V_1\|_F^2 = \|A^\dagger E\|_F^2 - \|A^\dagger E B^\dagger B\|_F^2. \tag{3.7}$$

Using (2.7a), (3.6), and (3.7), we obtain

$$\|B^\dagger - A^\dagger\|_F^2 \leq \alpha_1 + \|B^\dagger E A^\dagger\|_F^2. \tag{3.8}$$

Interchanging the roles of A and B yields

$$\|B^\dagger - A^\dagger\|_F^2 \leq \alpha_2 + \|A^\dagger E B^\dagger\|_F^2. \tag{3.9}$$

The desired result (3.5) follows immediately by combining (3.8) and (3.9). □

6

Remark 3.3. If $A \in \mathbb{C}_n^{m \times n}$ $(m \geq n)$, then $V_1 = V$ and V_2 vanishes. In this case, we have

$$\alpha_1 = \|B^\dagger\|_2^2 \big(\|EB^\dagger\|_F^2 - \|AA^\dagger EB^\dagger\|_F^2 \big) + (n - s)\|A^\dagger\|_2^2,$$
$$\alpha_2 = \|A^\dagger\|_2^2 \big(\|EA^\dagger\|_F^2 - \|BB^\dagger EA^\dagger\|_F^2 \big).$$

Note that

$$\|AA^\dagger EB^\dagger\|_F^2 \geq \frac{\|A^\dagger EB^\dagger\|_F^2}{\|A^\dagger\|_2^2} \quad \text{and} \quad \|BB^\dagger EA^\dagger\|_F^2 \geq \frac{\|B^\dagger EA^\dagger\|_F^2}{\|B^\dagger\|_2^2}.$$

In light of (3.5), we have

$$\|B^\dagger - A^\dagger\|_F^2 \leq \frac{\|A^\dagger\|_2^2 \|B^\dagger\|_2^2}{\|A^\dagger\|_2^2 + \|B^\dagger\|_2^2} \left(\frac{\alpha_1 + \|B^\dagger EA^\dagger\|_F^2}{\|B^\dagger\|_2^2} + \frac{\alpha_2 + \|A^\dagger EB^\dagger\|_F^2}{\|A^\dagger\|_2^2} \right)$$
$$\leq \frac{\|A^\dagger\|_2^2 \|B^\dagger\|_2^2}{\|A^\dagger\|_2^2 + \|B^\dagger\|_2^2} \left(\|EA^\dagger\|_F^2 + \|EB^\dagger\|_F^2 + (n - s) \frac{\|A^\dagger\|_2^2}{\|B^\dagger\|_2^2} \right).$$

Therefore, (3.5) has improved the estimate (1.7).

Remark 3.4. If $A \in \mathbb{C}_n^{m \times n}$ and $B \in \mathbb{C}_n^{m \times n}$ $(m \geq n)$, then

$$\alpha_1 = \|B^\dagger\|_2^2 \big(\|EB^\dagger\|_F^2 - \|AA^\dagger EB^\dagger\|_F^2 \big),$$
$$\alpha_2 = \|A^\dagger\|_2^2 \big(\|EA^\dagger\|_F^2 - \|BB^\dagger EA^\dagger\|_F^2 \big).$$

Due to $\text{rank}(B) = \text{rank}(A)$, it follows that

$$\|EB^\dagger\|_F^2 - \|AA^\dagger EB^\dagger\|_F^2 = \|\widetilde{U}_1^* U_2\|_F^2 = \|\widetilde{U}_2^* U_1\|_F^2 = \|EA^\dagger\|_F^2 - \|BB^\dagger EA^\dagger\|_F^2,$$

where we have used Lemma 2.3. Then

$$\alpha_1 = \|B^\dagger\|_2^2 \big(\|EA^\dagger\|_F^2 - \|BB^\dagger EA^\dagger\|_F^2 \big) \leq \|B^\dagger\|_2^2 \|EA^\dagger\|_F^2 - \|B^\dagger EA^\dagger\|_F^2,$$
$$\alpha_2 = \|A^\dagger\|_2^2 \big(\|EB^\dagger\|_F^2 - \|AA^\dagger EB^\dagger\|_F^2 \big) \leq \|A^\dagger\|_2^2 \|EB^\dagger\|_F^2 - \|A^\dagger EB^\dagger\|_F^2,$$

because

$$\|BB^\dagger EA^\dagger\|_F^2 \geq \frac{\|B^\dagger EA^\dagger\|_F^2}{\|B^\dagger\|_2^2} \quad \text{and} \quad \|AA^\dagger EB^\dagger\|_F^2 \geq \frac{\|A^\dagger EB^\dagger\|_F^2}{\|A^\dagger\|_2^2}.$$

By (3.5), we have

$$\|B^\dagger - A^\dagger\|_F^2 \leq \min \big\{ \|B^\dagger\|_2^2 \|EA^\dagger\|_F^2, \|A^\dagger\|_2^2 \|EB^\dagger\|_F^2 \big\},$$

which is exactly (1.8). Thus, (3.5) has also improved the estimate (1.8).

The following two corollaries are based on Theorem 3.2.

Corollary 3.1. *Under the assumptions of Theorem 3.2, it holds that*

$$\|B^\dagger - A^\dagger\|_F^2 \leq \min \big\{ \beta_1 + \|B^\dagger EA^\dagger\|_F^2, \beta_2 + \|A^\dagger EB^\dagger\|_F^2 \big\}, \tag{3.10}$$

where

$$\beta_1 := \|A^\dagger\|_2^4 \big(\|E\|_F^2 - \|EB^\dagger B\|_F^2 \big) + \|B^\dagger\|_2^4 \big(\|E\|_F^2 - \|AA^\dagger E\|_F^2 \big),$$
$$\beta_2 := \|A^\dagger\|_2^4 \big(\|E\|_F^2 - \|BB^\dagger E\|_F^2 \big) + \|B^\dagger\|_2^4 \big(\|E\|_F^2 - \|EA^\dagger A\|_F^2 \big).$$

7

Proof. In view of (2.2a), (2.2b), (2.3a), and (2.3b), we have

$$U^* E \widetilde{V} = \begin{pmatrix} U_1^* \widetilde{U}_1 \widetilde{\Sigma}_1 - \Sigma_1 V_1^* \widetilde{V}_1 & -\Sigma_1 V_1^* \widetilde{V}_2 \\ U_2^* \widetilde{U}_1 \widetilde{\Sigma}_1 & 0 \end{pmatrix}, \tag{3.11}$$

$$U^* E B^\dagger B \widetilde{V} = \begin{pmatrix} U_1^* \widetilde{U}_1 \widetilde{\Sigma}_1 - \Sigma_1 V_1^* \widetilde{V}_1 & 0 \\ U_2^* \widetilde{U}_1 \widetilde{\Sigma}_1 & 0 \end{pmatrix}, \tag{3.12}$$

$$U^* A A^\dagger E \widetilde{V} = \begin{pmatrix} U_1^* \widetilde{U}_1 \widetilde{\Sigma}_1 - \Sigma_1 V_1^* \widetilde{V}_1 & -\Sigma_1 V_1^* \widetilde{V}_2 \\ 0 & 0 \end{pmatrix}. \tag{3.13}$$

From (3.11), (3.12), and (3.13), we deduce that

$$\|\widetilde{V}_2^* V_1 \Sigma_1\|_F^2 = \|E\|_F^2 - \|E B^\dagger B\|_F^2,$$
$$\|\widetilde{\Sigma}_1 \widetilde{U}_1^* U_2\|_F^2 = \|E\|_F^2 - \|A A^\dagger E\|_F^2.$$

Then

$$\|A^\dagger E\|_F^2 - \|A^\dagger E B^\dagger B\|_F^2 = \|\widetilde{V}_2^* V_1\|_F^2 \leq \|A^\dagger\|_2^2 \|\widetilde{V}_2^* V_1 \Sigma_1\|_F^2 = \|A^\dagger\|_2^2 \big(\|E\|_F^2 - \|E B^\dagger B\|_F^2\big),$$
$$\|E B^\dagger\|_F^2 - \|A A^\dagger E B^\dagger\|_F^2 = \|\widetilde{U}_1^* U_2\|_F^2 \leq \|B^\dagger\|_2^2 \|\widetilde{\Sigma}_1 \widetilde{U}_1^* U_2\|_F^2 = \|B^\dagger\|_2^2 \big(\|E\|_F^2 - \|A A^\dagger E\|_F^2\big).$$

Hence,

$$\alpha_1 \leq \|A^\dagger\|_2^4 \big(\|E\|_F^2 - \|E B^\dagger B\|_F^2\big) + \|B^\dagger\|_2^4 \big(\|E\|_F^2 - \|A A^\dagger E\|_F^2\big).$$

Similarly,

$$\alpha_2 \leq \|A^\dagger\|_2^4 \big(\|E\|_F^2 - \|B B^\dagger E\|_F^2\big) + \|B^\dagger\|_2^4 \big(\|E\|_F^2 - \|E A^\dagger A\|_F^2\big).$$

Using (3.5), we can obtain the estimate (3.10) immediately. □

Corollary 3.2. *Under the assumptions of Theorem* 3.2, *it holds that*

$$\|B^\dagger - A^\dagger\|_F^2 \leq \min\big\{\gamma_1 + \|B^\dagger E A^\dagger\|_F^2, \gamma_2 + \|A^\dagger E B^\dagger\|_F^2\big\}, \tag{3.14}$$

where

$$\gamma_1 := \|A^\dagger\|_2^2 \bigg(\|A^\dagger E\|_F^2 - \frac{\|A^\dagger E B^\dagger\|_F^2}{\|B^\dagger\|_2^2}\bigg) + \|B^\dagger\|_2^2 \bigg(\|E B^\dagger\|_F^2 - \frac{\|A^\dagger E B^\dagger\|_F^2}{\|A^\dagger\|_2^2}\bigg),$$

$$\gamma_2 := \|A^\dagger\|_2^2 \bigg(\|E A^\dagger\|_F^2 - \frac{\|B^\dagger E A^\dagger\|_F^2}{\|B^\dagger\|_2^2}\bigg) + \|B^\dagger\|_2^2 \bigg(\|B^\dagger E\|_F^2 - \frac{\|B^\dagger E A^\dagger\|_F^2}{\|A^\dagger\|_2^2}\bigg).$$

Proof. It is easy to verify that

$$\|A^\dagger E B^\dagger B\|_F^2 \geq \frac{\|A^\dagger E B^\dagger\|_F^2}{\|B^\dagger\|_2^2} \quad \text{and} \quad \|A A^\dagger E B^\dagger\|_F^2 \geq \frac{\|A^\dagger E B^\dagger\|_F^2}{\|A^\dagger\|_2^2}.$$

The estimate (3.14) then follows from Theorem 3.2. □

The following theorem provides the sharper counterparts of (1.4), (1.5), and (1.6).

8

Theorem 3.3. *Let $A \in \mathbb{C}_r^{m \times n}$, $B \in \mathbb{C}_s^{m \times n}$, and $E = B - A$. Then*

$$\|B^\dagger - A^\dagger\|_F^2 \le \min\left\{\delta_1 + \|B^\dagger E A^\dagger\|_F^2, \delta_2 + \|A^\dagger E B^\dagger\|_F^2\right\}, \tag{3.15}$$

where

$$\delta_1 := \max\left\{\|A^\dagger\|_2^4, \|B^\dagger\|_2^4\right\}\left(\|E\|_F^2 - \max\left\{\frac{\|AA^\dagger E B^\dagger\|_F^2}{\|B^\dagger\|_2^2}, \frac{\|A^\dagger E B^\dagger B\|_F^2}{\|A^\dagger\|_2^2}\right\}\right),$$

$$\delta_2 := \max\left\{\|A^\dagger\|_2^4, \|B^\dagger\|_2^4\right\}\left(\|E\|_F^2 - \max\left\{\frac{\|BB^\dagger E A^\dagger\|_F^2}{\|A^\dagger\|_2^2}, \frac{\|B^\dagger E A^\dagger A\|_F^2}{\|B^\dagger\|_2^2}\right\}\right).$$

In particular, if $s = r$, then

$$\|B^\dagger - A^\dagger\|_F^2 \le \min\left\{\varepsilon_1 + \|B^\dagger E A^\dagger\|_F^2, \varepsilon_2 + \|A^\dagger E B^\dagger\|_F^2\right\}, \tag{3.16}$$

where

$$\varepsilon_1 := \|A^\dagger\|_2^2\|B^\dagger\|_2^2\left(\|E\|_F^2 - \max\left\{\frac{\|BB^\dagger E A^\dagger\|_F^2}{\|A^\dagger\|_2^2}, \frac{\|B^\dagger E A^\dagger A\|_F^2}{\|B^\dagger\|_2^2}\right\}\right),$$

$$\varepsilon_2 := \|A^\dagger\|_2^2\|B^\dagger\|_2^2\left(\|E\|_F^2 - \max\left\{\frac{\|AA^\dagger E B^\dagger\|_F^2}{\|B^\dagger\|_2^2}, \frac{\|A^\dagger E B^\dagger B\|_F^2}{\|A^\dagger\|_2^2}\right\}\right).$$

Proof. According to (3.11), we deduce that

$$\|E\|_F^2 = \|(U_1^* \widetilde{U}_1 - \Sigma_1 V_1^* \widetilde{V}_1 \widetilde{\Sigma}_1^{-1})\widetilde{\Sigma}_1\|_F^2 + \|\Sigma_1 V_1^* \widetilde{V}_2\|_F^2 + \|U_2^* \widetilde{U}_1 \widetilde{\Sigma}_1\|_F^2$$

$$\ge \frac{\|AA^\dagger E B^\dagger\|_F^2}{\|B^\dagger\|_2^2} + \frac{\|\widetilde{V}_2^* V_1\|_F^2}{\|A^\dagger\|_2^2} + \frac{\|\widetilde{U}_1^* U_2\|_F^2}{\|B^\dagger\|_2^2}$$

and

$$\|E\|_F^2 = \|\Sigma_1(\Sigma_1^{-1} U_1^* \widetilde{U}_1 \widetilde{\Sigma}_1 - V_1^* \widetilde{V}_1)\|_F^2 + \|\Sigma_1 V_1^* \widetilde{V}_2\|_F^2 + \|U_2^* \widetilde{U}_1 \widetilde{\Sigma}_1\|_F^2$$

$$\ge \frac{\|A^\dagger E B^\dagger B\|_F^2}{\|A^\dagger\|_2^2} + \frac{\|\widetilde{V}_2^* V_1\|_F^2}{\|A^\dagger\|_2^2} + \frac{\|\widetilde{U}_1^* U_2\|_F^2}{\|B^\dagger\|_2^2}.$$

Hence,

$$\frac{\|\widetilde{U}_1^* U_2\|_F^2}{\|B^\dagger\|_2^2} + \frac{\|\widetilde{V}_2^* V_1\|_F^2}{\|A^\dagger\|_2^2} \le \|E\|_F^2 - \max\left\{\frac{\|AA^\dagger E B^\dagger\|_F^2}{\|B^\dagger\|_2^2}, \frac{\|A^\dagger E B^\dagger B\|_F^2}{\|A^\dagger\|_2^2}\right\}. \tag{3.17}$$

Using (2.7a) and (3.17), we obtain

$$\|B^\dagger - A^\dagger\|_F^2 \le \max\left\{\|A^\dagger\|_2^4, \|B^\dagger\|_2^4\right\}\left(\frac{\|\widetilde{U}_1^* U_2\|_F^2}{\|B^\dagger\|_2^2} + \frac{\|\widetilde{V}_2^* V_1\|_F^2}{\|A^\dagger\|_2^2}\right) + \|B^\dagger E A^\dagger\|_F^2$$

$$\le \delta_1 + \|B^\dagger E A^\dagger\|_F^2.$$

Interchanging the roles of A and B, we arrive at

$$\|B^\dagger - A^\dagger\|_F^2 \le \delta_2 + \|A^\dagger E B^\dagger\|_F^2.$$

Thus, the estimate (3.15) is valid.

We next consider the special case $s = r$. Direct computation yields

$$\widetilde{U}^* E V = \begin{pmatrix} \widetilde{\Sigma}_1 \widetilde{V}_1^* V_1 - \widetilde{U}_1^* U_1 \Sigma_1 & \widetilde{\Sigma}_1 \widetilde{V}_1^* V_2 \\ -\widetilde{U}_2^* U_1 \Sigma_1 & 0 \end{pmatrix},$$

which leads to

$$\|E\|_F^2 = \|\widetilde{\Sigma}_1 \widetilde{V}_1^* V_1 - \widetilde{U}_1^* U_1 \Sigma_1\|_F^2 + \|\widetilde{\Sigma}_1 \widetilde{V}_1^* V_2\|_F^2 + \|\widetilde{U}_2^* U_1 \Sigma_1\|_F^2. \qquad (3.18)$$

If $s = r$, we get from (3.18) and Lemma 2.3 that

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

and

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

Hence,

$$\frac{\|\widetilde{U}_1^* U_2\|_F^2}{\|A^\dagger\|_2^2} + \frac{\|\widetilde{V}_2^* V_1\|_F^2}{\|B^\dagger\|_2^2} \leq \|E\|_F^2 - \max\left\{ \frac{\|B B^\dagger E A^\dagger\|_F^2}{\|A^\dagger\|_2^2}, \frac{\|B^\dagger E A^\dagger A\|_F^2}{\|B^\dagger\|_2^2} \right\}. \qquad (3.19)$$

By (2.7a) and (3.19), we have

$$\begin{aligned}
\|B^\dagger - A^\dagger\|_F^2 &\leq \|A^\dagger\|_2^2 \|B^\dagger\|_2^2 \left(\frac{\|\widetilde{U}_1^* U_2\|_F^2}{\|A^\dagger\|_2^2} + \frac{\|\widetilde{V}_2^* V_1\|_F^2}{\|B^\dagger\|_2^2} \right) + \|B^\dagger E A^\dagger\|_F^2 \\
&\leq \varepsilon_1 + \|B^\dagger E A^\dagger\|_F^2.
\end{aligned}$$

Interchanging the roles of A and B yields

$$\|B^\dagger - A^\dagger\|_F^2 \leq \varepsilon_2 + \|A^\dagger E B^\dagger\|_F^2.$$

Therefore, the estimate (3.16) is proved. This completes the proof. $\qquad \square$

Remark 3.5. From (3.15), we deduce that

$$\|B^\dagger - A^\dagger\|_F^2 \leq \frac{1}{2} \left(\delta_1 + \delta_2 + \|A^\dagger E B^\dagger\|_F^2 + \|B^\dagger E A^\dagger\|_F^2 \right). \qquad (3.20)$$

Using the inequalities

$$\|A A^\dagger E B^\dagger\|_F^2 \geq \frac{\|A^\dagger E B^\dagger\|_F^2}{\|A^\dagger\|_2^2} \quad \text{and} \quad \|B B^\dagger E A^\dagger\|_F^2 \geq \frac{\|B^\dagger E A^\dagger\|_F^2}{\|B^\dagger\|_2^2},$$

10

we obtain

$$\delta_1 + \delta_2 \leq \max\left\{\|A^\dagger\|_2^4, \|B^\dagger\|_2^4\right\}\left(2\|E\|_F^2 - \frac{\|A^\dagger E B^\dagger\|_F^2 + \|B^\dagger E A^\dagger\|_F^2}{\|A^\dagger\|_2^2\|B^\dagger\|_2^2}\right)$$

$$= 2\max\left\{\|A^\dagger\|_2^4, \|B^\dagger\|_2^4\right\}\|E\|_F^2 - \max\left\{\frac{\|A^\dagger\|_2^2}{\|B^\dagger\|_2^2}, \frac{\|B^\dagger\|_2^2}{\|A^\dagger\|_2^2}\right\}\left(\|A^\dagger E B^\dagger\|_F^2 + \|B^\dagger E A^\dagger\|_F^2\right)$$

$$\leq 2\max\left\{\|A^\dagger\|_2^4, \|B^\dagger\|_2^4\right\}\|E\|_F^2 - \left(\|A^\dagger E B^\dagger\|_F^2 + \|B^\dagger E A^\dagger\|_F^2\right).$$

Thus, the estimate (3.20) is sharper than both (1.4) and (1.6). Furthermore, since

$$\|B^\dagger\|_2^2\|BB^\dagger E A^\dagger\|_F^2 \geq \|B^\dagger E A^\dagger\|_F^2 \quad \text{and} \quad \|A^\dagger\|_2^2\|AA^\dagger E B^\dagger\|_F^2 \geq \|A^\dagger E B^\dagger\|_F^2,$$

it follows that the estimate (3.16) is sharper than (1.5).

Remark 3.6. Under the setting of (1.9), the upper bounds in (3.5), (3.10), (3.14), and (3.15) are listed in Table 2. Table 2 shows that the upper bounds in (3.5), (3.10), and (3.15) have attained

Estimate	Upper bound for $\|B^\dagger - A^\dagger\|_F^2$
(3.5)	$4\tau^2 + \frac{1}{\tau^2}$
(3.10)	$4\tau^2 + \frac{1}{\tau^2}$
(3.14)	$4\tau^2 + \frac{1}{\tau^2} + \frac{4\tau^2}{(1+2\tau)^2} - 4\tau^4$
(3.15)	$4\tau^2 + \frac{1}{\tau^2}$

Table 2: The upper bounds in (3.5), (3.10), (3.14), and (3.15).

the exact value $4\tau^2 + \frac{1}{\tau^2}$. In addition, if τ is sufficiently small (i.e., the perturbed matrix is very close to the original one), the upper bound in (3.14) will be very close to the exact value.

Acknowledgments

References

[1] N. N. Abdelmalek. On the solution of the linear least squares problems and pseudo-inverses. *Computing*, 13:215–228, 1974.

[2] A. Ben-Israel and T. N. E. Greville. *Generalized Inverses: Theory and Applications.* Springer-Verlag, New York, 2nd edition, 2003.

[3] T. Bodnar, H. Dette, and N. Parolya. Spectral analysis of the Moore–Penrose inverse of a large dimensional sample covariance matrix. *J. Multivariate Anal.*, 148:160–172, 2016.

[4] L.-X. Cai, W.-W. Xu, and W. Li. Additive and multiplicative perturbation bounds for the Moore–Penrose inverse. *Linear Algebra Appl.*, 434:480–489, 2011.

[5] Y. Chen, X. Chen, and W. Li. On perturbation bounds for orthogonal projections. *Numer. Algorithms*, 73:433–444, 2016.

[6] P. Drineas, M. W. Mahoney, S. Muthukrishnan, and T. Sarlós. Faster least squares approximation. *Numer. Math.*, 117:219–249, 2011.

[7] D. C. Hoyle. Accuracy of pseudo-inverse covariance learning–a random matrix theory analysis. *IEEE Trans. Pattern Anal. Mach. Intell.*, 33:1470–1481, 2011.

[8] W. Li, Y. Chen, S. Vong, and Q. Luo. Some refined bounds for the perturbation of the orthogonal projection and the generalized inverse. *Numer. Algorithms*, 79:657–677, 2018.

[9] L. Meng and B. Zheng. The optimal perturbation bounds of the Moore–Penrose inverse under the Frobenius norm. *Linear Algebra Appl.*, 432:956–963, 2010.

[10] R. Penrose. A generalized inverse for matrices. *Proc. Cambridge Philos. Soc.*, 51:406–413, 1955.

[11] R. Penrose. On best approximate solutions of linear matrix equations. *Proc. Cambridge Philos. Soc.*, 52:17–19, 1956.

[12] G. W. Stewart. On the continuity of the generalized inverse. *SIAM J. Appl. Math.*, 17:33–45, 1969.

[13] G. W. Stewart. On the perturbation of pseudo-inverses, projections and linear least squares problems. *SIAM Rev.*, 19:634–662, 1977.

[14] J.-G. Sun. *Matrix Perturbation Analysis*. Science Press, Beijing, 2nd edition, 2001 (in Chinese).

[15] J. von Neumann. Some matrix-inequalities and metrization of matrix-space. *Tomsk. Univ. Rev.*, 1:286–300, 1937.

[16] P.-Å. Wedin. Perturbation theory for pseudo-inverses. *BIT Numer. Math.*, 13:217–232, 1973.

[17] X. Xu. Generalization of the Sherman–Morrison–Woodbury formula involving the Schur complement. *Appl. Math. Comput.*, 309:183–191, 2017.

[18] X. Xu. On the perturbation of an L^2-orthogonal projection. *J. Comput. Appl. Math.*, 368:112327, 2020.