Journal of Applied Analysis and Computation
Volume 6, Number 1, February 2016, 216–226
Website:http://jaac-online.com/
doi:10.11948/2016018

NORM ESTIMATIONS FOR PERTURBATIONS OF THE WEIGHTED MOORE-PENROSE INVERSE*

Xiaobo Zhang[1], Qingxiang Xu[1,†] and Yimin Wei[2]

Abstract For a complex matrix $A \in \mathbb{C}^{m\times n}$, the relationship between the weighted Moore-Penrose inverse $A^\dagger_{M_1N_1}$ and $A^\dagger_{M_2N_2}$ is studied, and an important formula is derived, where $M_1 \in \mathbb{C}^{m\times m}, N_1 \in \mathbb{C}^{n\times n}$ and $M_2 \in \mathbb{C}^{m\times m}, N_2 \in \mathbb{C}^{n\times n}$ are different pair of positive definite hermitian matrices. Based on this formula, this paper initiates the study of the perturbation estimations for $A^\dagger_{MN}$ in the case that A is fixed, whereas both M and N are variable. The obtained norm upper bounds are then applied to the perturbation estimations for the solutions to the weighted linear least squares problems.

Keywords Weighted Moore-Penrose inverse, norm upper bound, weighted linear least squares problem.

MSC(2010) 15A09, 15A60, 65F35.

1. Introduction

Throughout this paper $\mathbb{C}^{m\times n}$ is the set of $m \times n$ complex matrices and $\|A\|$ denotes the 2-norm or spectral-norm of $A \in \mathbb{C}^{m\times n}$. When $m = n$ a positive definite matrix of $\mathbb{C}^{n\times n}$ is always assumed to be hermitian, and the identity matrix of $\mathbb{C}^{n\times n}$ is denoted by I_n or simply by I. For any $A \in \mathbb{C}^{m\times n}$, the range, the null space and the conjugate transpose of A are denoted by $\mathcal{R}(A), \mathcal{N}(A)$ and A^* respectively. Let $M \in \mathbb{C}^{m\times m}$ and $N \in \mathbb{C}^{n\times n}$ be two positive definite matrices, the weighted Moore-Penrose inverse $A^\dagger_{MN}$ is the unique element X of $\mathbb{C}^{n\times m}$ which satisfies

$$AXA = A, XAX = X, (MAX)^* = MAX \text{ and } (NXA)^* = NXA. \tag{1.1}$$

The weighted Moore-Penrose inverse has many applications in the weighted linear least squares problem [2–5, 7, 15–17], statistics [6], analytical dynamics [12], two-point boundary value problems [8] and so on. In this paper we study the perturbation estimation for the weighted Moore-Penrose inverse $A^\dagger_{MN}$. Some literatures [13, 18] are focused on the case that the weights M and N are fixed, whereas A is variable. Some others [1, 3–5, 9–11, 14] studied another case that A is fixed, N is the identity matrix, while M is a variable positive definite diagonal matrix. For a motivation to the study of the later case, the reader is referred to [3, Section

†the corresponding author. Email address:qxxu@shnu.edu.cn(Q. Xu)
1Department of Mathematics, Shanghai Normal University, Shanghai 200234, PR China
2Institute of Mathematics, School of Mathematical Sciences, Fudan University, Shanghai, 200433, PR China
*The authors were supported by National Natural Science Foundation of China under grant 11171222 and National Science Foundation of China under grant 11271084.

8], [4, Section 1.1] and [5, Section 1.1] for the interior methods in linear programming and convex quadratic programming. The key point of this paper is the characterization of the relationship between the weighted Moore-Penrose inverses $A^\dagger_{MN}$, where A is fixed, while both M and N are variable. Based on this formula, this paper initiates the study of the perturbation estimations for $A^\dagger_{MN}$ in the case that A is fixed, whereas both M and N are variable. The obtained norm upper bounds are then applied to the perturbation estimations for the solutions to the weighted linear least squares problems.

The paper is organized as follows. In Section 2, the relationship between the weighted Moore-Penrose inverses $A^\dagger_{M_1N_1}$ and $A^\dagger_{M_2N_2}$ is studied, and an important formula (2.7) is derived. This formula is applied in Section 3 to the study of the perturbation estimations for $A^\dagger_{MN}$, where A is fixed, while M and N are variable. The obtained norm upper bounds are then applied in Section 4 to the study of the perturbation estimations for the solutions to the weighted linear least squares problems. Finally, two numerical examples are provided in Section 5 to illustrate the upper bounds obtained in Sections 3 and 4.

2. Relationship between the weighted Moore-Penrose inverse

Throughout this section $A\in\mathbb{C}^{m\times n}$ is arbitrary, and $M, M_1, M_2\in\mathbb{C}^{m\times m}$, $N, N_1, N_2\in\mathbb{C}^{n\times n}$ are all positive definite.

Lemma 2.1. [13, Theorem 1.4.4] *It holds that*

$$\mathcal{R}(A^\dagger_{MN}) = N^{-1}\mathcal{R}(A^*) \text{ and } \mathcal{N}(A^\dagger_{MN}) = M^{-1}\mathcal{N}(A^*).$$

Lemma 2.2. *It holds that* $AA^\dagger_{MN_1} = AA^\dagger_{MN_2}$ *and* $A^\dagger_{M_1N}A = A^\dagger_{M_2N}A$.

Proof. Clearly, $\mathcal{R}(AA^\dagger_{MN_1}) = \mathcal{R}(A) = \mathcal{R}(AA^\dagger_{MN_2})$, and by Lemma 2.1 we have

$$\mathcal{N}(AA^\dagger_{MN_1}) = \mathcal{N}(A^\dagger_{MN_1}) = M^{-1}\mathcal{N}(A^*) = \mathcal{N}(AA^\dagger_{MN_2}).$$

This completes the proof that $AA^\dagger_{MN_1}$ and $AA^\dagger_{MN_2}$ have the same range and the same null space. Since both of them are idempotent, they must be equal. The proof of $A^\dagger_{M_1N}A = A^\dagger_{M_2N}A$ is similar. □

Lemma 2.3. *It holds that*

$$(I - A^\dagger_{MN_1}A)N_1^{-1}N_2A^\dagger_{MN_2}A = 0, \tag{2.1}$$

$$AA^\dagger_{M_2N}M_2^{-1}M_1(I - AA^\dagger_{M_1N}) = 0. \tag{2.2}$$

Proof. By (1.1), we have

$$N_2A^\dagger_{MN_2}A = (A^\dagger_{MN_2}A)^*N_2 \text{ and } A^\dagger_{MN_1}AN_1^{-1} = N_1^{-1}(A^\dagger_{MN_1}A)^*.$$

It follows that

$$\begin{aligned}&A^\dagger_{MN_1}AN_1^{-1}N_2A^\dagger_{MN_2}A = N_1^{-1}(A^\dagger_{MN_1}A)^*(A^\dagger_{MN_2}A)^*N_2\\ =&N_1^{-1}\Big(A^\dagger_{MN_2}(AA^\dagger_{MN_1}A)\Big)^*N_2 = N_1^{-1}(A^\dagger_{MN_2}A)^*N_2 = N_1^{-1}N_2A^\dagger_{MN_2}A.\end{aligned}$$

This completes the proof of (2.1). The proof of (2.2) is similar. □

Lemma 2.4. *It holds that* $A^{\dagger}_{MN_2} = R^{-1}_{M;N_1,N_2} \cdot A^{\dagger}_{MN_1}$, *where*

$$R_{M;N_1,N_2} = A^{\dagger}_{MN_1}A + (I - A^{\dagger}_{MN_1}A)N_1^{-1}N_2. \tag{2.3}$$

Proof. First, we prove that $R_{M;N_1,N_2}$ is nonsingular. Let $x \in \mathbb{C}^n$ be given such that $R_{M;N_1,N_2}x = 0$. Then it is obvious from (2.3) that

$$A^{\dagger}_{MN_1}Ax = 0 \text{ and } (I - A^{\dagger}_{MN_1}A)N_1^{-1}N_2x = 0,$$

which in turn implies $Ax = 0$, and $N_1^{-1}N_2x \in \mathcal{R}(A^{\dagger}_{MN_1}A) = \mathcal{R}(A^{\dagger}_{MN_1}) = N_1^{-1}\mathcal{R}(A^*)$. Thus, $x = N_2^{-1}A^*u$ for some $u \in \mathbb{C}^m$. It follows that

$$\langle N_2^{-1}A^*u, A^*u\rangle = \langle x, A^*u\rangle = \langle Ax, u\rangle = 0 \Longrightarrow A^*u = 0 \Longrightarrow x = N_2^{-1}(A^*u) = 0.$$

Next, we prove that $A^{\dagger}_{MN_1} = R_{M;N_1,N_2} \cdot A^{\dagger}_{MN_2}$. In fact, by (2.3) and (2.1), we have

$$R_{M;N_1,N_2} \cdot A^{\dagger}_{MN_2}A = A^{\dagger}_{MN_1}(AA^{\dagger}_{MN_2}A) = A^{\dagger}_{MN_1}A,$$

which is combined with Lemma 2.2 to conclude that

$$\begin{aligned} R_{M;N_1,N_2} \cdot A^{\dagger}_{MN_2} &= R_{M;N_1,N_2} \cdot A^{\dagger}_{MN_2}AA^{\dagger}_{MN_2} \\ =& R_{M;N_1,N_2} \cdot A^{\dagger}_{MN_2}AA^{\dagger}_{MN_1} = A^{\dagger}_{MN_1}AA^{\dagger}_{MN_1} = A^{\dagger}_{MN_1}. \end{aligned}$$

□

Lemma 2.5. *It holds that* $A^{\dagger}_{M_2N} = A^{\dagger}_{M_1N} \cdot L^{-1}_{M_1,M_2;N}$, *where*

$$L_{M_1,M_2;N} = AA^{\dagger}_{M_1N} + M_2^{-1}M_1(I - AA^{\dagger}_{M_1N}). \tag{2.4}$$

Proof. First, we prove that $L_{M_1,M_2;N}$ is nonsingular. For any $x \in \mathbb{C}^m$, if

$$L_{M_1,M_2;N}x = AA^{\dagger}_{M_1N}x + M_2^{-1}M_1(I - AA^{\dagger}_{M_1N})x = 0, \tag{2.5}$$

then by (2.5) and (2.2) we get

$$AA^{\dagger}_{M_1N}x = (AA^{\dagger}_{M_2N})AA^{\dagger}_{M_1N}x = AA^{\dagger}_{M_2N}L_{M_1,M_2;N}x = 0.$$

Substituting the above equation into (2.5) yields

$$\begin{aligned} & M_2^{-1}M_1(I - AA^{\dagger}_{M_1N_1})\,x = 0 \\ \Longrightarrow & (I - AA^{\dagger}_{M_1N_1})\,x = 0 \\ \Longrightarrow & x = AA^{\dagger}_{M_1N_1}x + (I - AA^{\dagger}_{M_1N_1})x = 0. \end{aligned}$$

Next, we prove that $A^{\dagger}_{M_2N} \cdot L_{M_1,M_2;N} = A^{\dagger}_{M_1N}$. In fact, from (2.2) we have

$$A^{\dagger}_{M_2N}M_2^{-1}M_1(I - AA^{\dagger}_{M_1N}) = 0. \tag{2.6}$$

Therefore, by (2.4), (2.6) and Lemma 2.2, we obtain

$$A^{\dagger}_{M_2N} \cdot L_{M_1,M_2;N} = (A^{\dagger}_{M_2N}A)A^{\dagger}_{M_1N} = (A^{\dagger}_{M_1N}A)A^{\dagger}_{M_1N} = A^{\dagger}_{M_1N}.$$

□

Now we state the main result of this section as follows:

Theorem 2.1. *It holds that*

$$A^\dagger_{M_2N_2} = R^{-1}_{M_1;N_1,N_2} \cdot A^\dagger_{M_1N_1} \cdot L^{-1}_{M_1,M_2;N_1}, \tag{2.7}$$

where $R_{M_1;N_1,N_2}$ and $L_{M_1,M_2;N_1}$ are defined by (2.3) and (2.4), respectively.

Proof. First note from (2.3) and Lemma 2.2 that $R_{M_2;N_1,N_2} = R_{M_1;N_1,N_2}$. Thus, we may apply Lemmas 2.4 and 2.5 to conclude that

$$A^\dagger_{M_2N_2} = R^{-1}_{M_1;N_1,N_2} \cdot A^\dagger_{M_2N_1} = R^{-1}_{M_1;N_1,N_2} \cdot A^\dagger_{M_1N_1} \cdot L^{-1}_{M_1,M_2;N_1}.$$

□

3. Norm estimations for the weighted Moore-Penrose inverse

Throughout this section, $A \in \mathbb{C}^{m\times n}$ is fixed, $M \in \mathbb{C}^{m\times m}$ and $N \in \mathbb{C}^{n\times n}$ are two positive definite matrices. Let $\widehat{M}$ and $\widehat{N}$ be perturbations of M and N defined by

$$\widehat{M} = M + \delta_M \text{ and } \widehat{N} = N + \delta_N, \tag{3.1}$$

such that

$$\delta_M \in \mathbb{C}^{m\times m} \text{ and } \delta_N \in \mathbb{C}^{n\times n} \text{ are hermitian,} \tag{3.2}$$

$$\|\delta_M\| < \frac{1}{\|M^{-1}\|} \text{ and } \|\delta_N\| < \frac{1}{\|N^{-1}\|}. \tag{3.3}$$

It follows from (3.2) and (3.3) that both $\widehat{M} = M + \delta_M$ and $\widehat{N} = N + \delta_N$ are also positive definite. Based on the formula (2.7), we study norm estimations associated with $A^\dagger_{MN}$ and $A^\dagger_{\widehat{M}\widehat{N}}$.

Lemma 3.1. *The matrices $M(I - AA^\dagger_{MN}) \in \mathbb{C}^{m\times m}$ and $(I - A^\dagger_{MN}A)N^{-1} \in \mathbb{C}^{n\times n}$ are both positive semi-definite.*

Proof. For simplicity, we put

$$T = M(I - AA^\dagger_{MN}) \text{ and } S = (I - A^\dagger_{MN}A)N^{-1}. \tag{3.4}$$

By (1.1), we have

$$(AA^\dagger_{MN})^*T = (MAA^\dagger_{MN})^*(I - AA^\dagger_{MN}) = MAA^\dagger_{MN}(I - AA^\dagger_{MN}) = 0,$$

so $T = (I - AA^\dagger_{MN})^*T = (I - AA^\dagger_{MN})^*M(I - AA^\dagger_{MN})$, which is positive semi-definite.

Similarly, $S = (I - A^\dagger_{MN}A)N^{-1}(I - A^\dagger_{MN}A)^*$ is also positive semi-definite. □

Remark 3.1. By Lemma 3.1 we know that

$$\left\|M(I - AA^\dagger_{MN})\right\| = r_1 \text{ and } \left\|(I - A^\dagger_{MN}A)N^{-1}\right\| = r_2, \tag{3.5}$$

where r_1 and r_2 are the largest eigenvalues of $M(I - AA^\dagger_{MN})$ and $(I - A^\dagger_{MN}A)N^{-1}$ respectively.

In the rest of this section, we always assume that (3.2) and (3.3) are satisfied, and furthermore, the following inequalities hold:

$$r_2\|\delta_N\| < 1,\ \|M^{-1}\delta_M\| \cdot \big(1 + r_1\|M^{-1}\|\big) < 1. \tag{3.6}$$

In such case, from (3.3) we have

$$\big\|\widehat{M}^{-1} - M^{-1}\big\| = \big\|\big((I + M^{-1}\delta_M)^{-1} - I\big)M^{-1}\big\| \le \frac{\|M^{-1}\| \cdot \|M^{-1}\delta_M\|}{1 - \|M^{-1}\delta_M\|},$$

so

$$\big\|(\widehat{M}^{-1} - M^{-1})M(I - AA^\dagger_{MN})\big\| \le \frac{r_1\,\|M^{-1}\| \cdot \|M^{-1}\delta_M\|}{1 - \|M^{-1}\delta_M\|}. \tag{3.7}$$

Now, let $R_{M;N,\widehat{N}}$ and $L_{M,\widehat{M};N}$ be defined by (2.3) and (2.4), respectively. Then

$$\begin{aligned} R_{M;N,\widehat{N}} &= A^\dagger_{MN}A + (I - A^\dagger_{MN}A)N^{-1}\widehat{N} \\ &= I + (I - A^\dagger_{MN}A)N^{-1}\delta_N, \end{aligned} \tag{3.8}$$

$$\begin{aligned} L_{M,\widehat{M};N} &= AA^\dagger_{MN} + \widehat{M}^{-1}M(I - AA^\dagger_{MN}) \\ &= I + (\widehat{M}^{-1} - M^{-1})M(I - AA^\dagger_{MN}). \end{aligned} \tag{3.9}$$

We may combine (3.8), the second equation of (3.5) with the first inequality of (3.6) to conclude that

$$\Big\|R^{-1}_{M;N,\widehat{N}}\Big\| \le \frac{1}{1 - \big\|\big[(I - A^\dagger_{MN}A)N^{-1}\big]\delta_N\big\|} \le \frac{1}{1 - r_2\|\delta_N\|}, \tag{3.10}$$

$$\Big\|I - R^{-1}_{M;N,\widehat{N}}\Big\| \le \frac{\big\|\big[(I - A^\dagger_{MN}A)N^{-1}\big]\delta_N\big\|}{1 - \big\|\big[(I - A^\dagger_{MN}A)N^{-1}\big]\delta_N\big\|} \le \frac{r_2\|\delta_N\|}{1 - r_2\|\delta_N\|}. \tag{3.11}$$

Similarly, we may apply (3.9), (3.7) and the second inequality of (3.6) to get

$$\Big\|L^{-1}_{M,\widehat{M};N}\Big\| \le \frac{1}{1 - \frac{r_1\,\|M^{-1}\|\cdot\|M^{-1}\delta_M\|}{1 - \|M^{-1}\delta_M\|}} = \frac{1 - \|M^{-1}\delta_M\|}{1 - \|M^{-1}\delta_M\| \cdot \big(1 + r_1\|M^{-1}\|\big)}, \tag{3.12}$$

$$\Big\|I - L^{-1}_{M,\widehat{M};N}\Big\| \le \frac{r_1\,\|M^{-1}\| \cdot \|M^{-1}\delta_M\|}{1 - \|M^{-1}\delta_M\| \cdot \big(1 + r_1\|M^{-1}\|\big)}. \tag{3.13}$$

Theorem 3.1. *Under the conditions of (3.2), (3.3) and (3.6), we have*

$$\|A^\dagger_{\widehat{M}\widehat{N}}\| \le \frac{(1 - \|M^{-1}\delta_M\|) \cdot \|A^\dagger_{MN}\|}{(1 - r_2\|\delta_N\|)\Big[1 - \|M^{-1}\delta_M\|\,\big(1 + r_1\|M^{-1}\|\big)\Big]}, \tag{3.14}$$

$$\|A^\dagger_{\widehat{M}\widehat{N}} - A^\dagger_{MN}\| \le \Lambda \cdot \|A^\dagger_{MN}\|, \tag{3.15}$$

$$\|A^\dagger_{\widehat{M}\widehat{N}}A - A^\dagger_{MN}A\| \le \frac{r_2\|\delta_N\|}{1 - r_2\|\delta_N\|} \cdot \|A^\dagger_{MN}A\|, \tag{3.16}$$

$$\|AA^\dagger_{\widehat{M}\widehat{N}} - AA^\dagger_{MN}\| \le \frac{r_1\|M^{-1}\delta_M\| \cdot \|M^{-1}\|}{1 - \|M^{-1}\delta_M\|\big(1 + r_1\|M^{-1}\|\big)}\|AA^\dagger_{MN}\|, \tag{3.17}$$

where

$$\Lambda = \frac{r_2\|\delta_N\|\cdot(1-\|M^{-1}\delta_M\|)+r_1(1-r_2\|\delta_N\|)\|M^{-1}\delta_M\|\cdot\|M^{-1}\|}{(1-r_2\|\delta_N\|)\Big[1-\|M^{-1}\delta_M\|\cdot\big(1+r_1\|M^{-1}\|\big)\Big]}$$
$$= \frac{r_1\|M^{-1}\delta_M\|\cdot\|M^{-1}\|+r_2\|\delta_N\|\Big[1-\|M^{-1}\delta_M\|\cdot\big(1+r_1\|M^{-1}\|\big)\Big]}{(1-r_2\|\delta_N\|)\Big[1-\|M^{-1}\delta_M\|\cdot\big(1+r_1\|M^{-1}\|\big)\Big]}. \tag{3.18}$$

Proof. By Theorem 2.1 we have

$$A^{\dagger}_{\widehat{M}\widehat{N}} = R^{-1}_{M;N,\widehat{N}}\cdot A^{\dagger}_{MN}\cdot L^{-1}_{M,\widehat{M};N}, \tag{3.19}$$

so

$$A^{\dagger}_{\widehat{M}\widehat{N}} - A^{\dagger}_{MN} = \big(R^{-1}_{M;N,\widehat{N}}-I\big)A^{\dagger}_{MN}L^{-1}_{M,\widehat{M};N} + A^{\dagger}_{MN}\big(L^{-1}_{M,\widehat{M};N}-I\big). \tag{3.20}$$

It is noticed by (3.8) and (3.9) that $AR_{M;N,\widehat{N}} = A = L_{M,\widehat{M};N}A$, and thus

$$AR^{-1}_{M;N,\widehat{N}} = A = L^{-1}_{M,\widehat{M};N}A. \tag{3.21}$$

It follows from (3.19) and (3.21) that

$$A^{\dagger}_{\widehat{M}\widehat{N}}A - A^{\dagger}_{MN}A = \big(R^{-1}_{M;N,\widehat{N}}-I\big)A^{\dagger}_{MN}A, \tag{3.22}$$
$$AA^{\dagger}_{\widehat{M}\widehat{N}} - AA^{\dagger}_{MN} = AA^{\dagger}_{MN}\big(L^{-1}_{M,\widehat{M};N}-I\big). \tag{3.23}$$

Norm upper bounds (3.14)–(3.17) then follows from (3.19), (3.20), (3.22), (3.23) and (3.10)–(3.13). □

Remark 3.2. The upper bound for $\|A^{\dagger}_{\widehat{M}\widehat{N}} - A^{\dagger}_{MN}\|$ given by (3.15) and (3.18) is somehow complicated, so it is meaningful to replace this upper bound with a simpler one. To this end, we need an elementary result as follows:

Lemma 3.2. *Suppose that $a>0$ and $r_1\geq 0$. Let $I=\big[0,\frac{1}{1+ar_1}\big)$, and*

$$f(x,y)=\frac{ar_1x+y-(1+ar_1)xy}{(1-y)\big(1-(1+ar_1)x\big)},\ \text{for } x\in I, y\in[0,1). \tag{3.24}$$

Then for any $x_1,x_2\in I$ and $y_1,y_2\in[0,1)$, we have

$$f(x_2,y_2)\geq f(x_1,y_1) \text{ whenever } x_1\leq x_2 \text{ and } y_1\leq y_2.$$

Proof. Let $x\in I$ and $y\in[0,1)$. Direct computation yields

$$\frac{\partial f}{\partial x}(x,y) = \frac{ar_1}{(1-y)\big(1-(1+ar_1)x\big)^2}\geq 0,$$
$$\frac{\partial f}{\partial y}(x,y) = \frac{1-x}{(1-y)^2\big(1-(1+ar_1)x\big)} > 0,$$

so the conclusion holds. □

Corollary 3.1. *Suppose that (3.2) and (3.3) are satisfied, and furthermore*

$$\varepsilon\,(1+r_1\|M^{-1}\|)<1,\ \textit{where}\ \varepsilon=\max\{\|M^{-1}\delta_M\|,\ r_2\|\delta_N\|\}. \tag{3.25}$$

Then

$$\big\|A^{\dagger}_{\widehat{M}\widehat{N}}-A^{\dagger}_{MN}\big\|\le\frac{(1+r_1\|M^{-1}\|)\,\varepsilon}{1-(1+r_1\|M^{-1}\|)\,\varepsilon}\,\big\|A^{\dagger}_{MN}\big\|. \tag{3.26}$$

Proof. By assumption $\varepsilon\,(1+r_1\|M^{-1}\|)<1$, so (3.6) is satisfied. Let $a=\|M^{-1}\|$, $x_0=\|M^{-1}\delta_M\|$ and $y_0=r_2\|\delta_N\|$. Then we may combine (3.15), (3.18) and (3.24) to get

$$\big\|A^{\dagger}_{\widehat{M}\widehat{N}}-A^{\dagger}_{MN}\big\|\le f(x_0,y_0)\cdot\|A^{\dagger}_{MN}\|. \tag{3.27}$$

By Lemma 3.2 we have

$$f(x_0,y_0)\le f(\varepsilon,\varepsilon)=\frac{(1+ar_1)\,\varepsilon}{1-(1+ar_1)\,\varepsilon}. \tag{3.28}$$

The upper bound (3.26) then follows from (3.27) and (3.28). □

4. The weighted linear least squares problem

We apply the obtained norm upper bounds to study the weighted linear least squares problem [15]. Let $A\in\mathbb{C}^{m\times n}$ be arbitrary, $M\in\mathbb{C}^{m\times m}$ and $N\in\mathbb{C}^{n\times n}$ be two positive definite matrices. For any $b\in\mathbb{C}^m$, let $x_0=A^{\dagger}_{MN}b$. It is known [15] that for any $x\in\mathbb{C}^n\setminus\{x_0\}$,

$$\|b-Ax_0\|_M\le\|b-Ax\|_M,$$

and

$$\|b-Ax_0\|_M=\|b-Ax\|_M\Longrightarrow\|x_0\|_N<\|x\|_N,$$

which means that $x_0=A^{\dagger}_{MN}b$ is the unique minimum N-norm M-least squares solution to the weighted linear squares problem

$$\|b-Ax\|_M=\min\{\|b-Az\|_M\,\big|\,z\in\mathbb{C}^n\}.$$

When M,N and b admit some errors, it is meaningful to provide norm estimations for $\widehat{x_0}-x_0$, where $\widehat{b}=b+\delta_b$ is a perturbation of b, and $\widehat{x_0}=A^{\dagger}_{\widehat{M}\widehat{N}}\widehat{b}$ is the minimum $\widehat{N}$-norm $\widehat{M}$-least squares solution to the associated perturbation problem. Since

$$\|\widehat{x_0}-x_0\|\le\|\big(A^{\dagger}_{\widehat{M}\widehat{N}}-A^{\dagger}_{MN}\big)b\|+\|A^{\dagger}_{\widehat{M}\widehat{N}}\,\delta_b\|\le\|A^{\dagger}_{\widehat{M}\widehat{N}}-A^{\dagger}_{MN}\|\,\|b\|+\|A^{\dagger}_{\widehat{M}\widehat{N}}\|\,\|\delta_b\|,$$

an upper bound for $\|\widehat{x_0}-x_0\|$ can be derived directly from (3.14), (3.15) and (3.18). Another upper bound for $\|\widehat{x_0}-x_0\|$ can also be given as follows:

Theorem 4.1. *Under the conditions of (3.2), (3.3) and (3.6), we have*

$$\begin{aligned}\|\widehat{x_0}-x_0\|\le&\frac{\big(1-\|M^{-1}\delta_M\|\big)\Big(X\|\delta_b\|+Y\|AA^{\dagger}_{MN}\|\Big)\,\|A^{\dagger}_{MN}\|}{\big(1-r_2\|\delta_N\|\big)\,X^2}\\&+\frac{r_2\|\delta_N\|\,\|x_0\|\,\|A^{\dagger}_{MN}A\|}{1-r_2\|\delta_N\|},\end{aligned} \tag{4.1}$$

where $x_0 = A^{\dagger}_{MN} b$, $\widehat{x_0} = A^{\dagger}_{\widehat{M}\widehat{N}}(b+\delta_b)$, $r = b - Ax_0$, r_1, r_2 *are defined by (3.5), and*

$$X = 1 - \|M^{-1}\delta_M\| \left(1 + r_1\|M^{-1}\|\right), \; Y = r_1 \, \|r\| \, \|M^{-1}\delta_M\| \, \|M^{-1}\|.$$

Proof. Direct computation yields

$$\widehat{x_0} - x_0 = A^{\dagger}_{\widehat{M}\widehat{N}}\,\delta_b + A^{\dagger}_{\widehat{M}\widehat{N}}(AA^{\dagger}_{\widehat{M}\widehat{N}} - AA^{\dagger}_{MN})\, r - (\, A^{\dagger}_{MN}A - A^{\dagger}_{\widehat{M}\widehat{N}}A\,)\, x_0,$$

so

$$\begin{aligned} \|\widehat{x_0} - x_0\| \le & \|A^{\dagger}_{\widehat{M}\widehat{N}}\| \cdot \left[\|\delta_b\| + \|r\| \, \|AA^{\dagger}_{\widehat{M}\widehat{N}} - AA^{\dagger}_{MN}\|\right] \\ & + \|A^{\dagger}_{MN}A - A^{\dagger}_{\widehat{M}\widehat{N}}A\| \, \|x_0\|. \end{aligned} \tag{4.2}$$

The upper bound (4.1) then follows from (4.2), (3.14), (3.17) and (3.16). □

5. Numerical examples

In this section, we provide two numerical examples to illustrate the upper bounds obtained in Sections 3 and 4.

Example 5.1. Let $A = \begin{pmatrix} 1 & 2 \\ 0 & 0 \end{pmatrix}$, $\widehat{M} = M = diag\,(2,1)$, $N = diag\,(1,4)$ and $\widehat{N} = N + \delta_N = \begin{pmatrix} 1-\varepsilon & 0 \\ 0 & 4+\varepsilon \end{pmatrix}$ for ε small enough, where $\delta_N = diag\,(-\varepsilon, \varepsilon)$. It is easy to verify that

$$A^{\dagger}_{\widehat{M}\widehat{N}} = \begin{pmatrix} (4+\varepsilon)/(8-3\varepsilon) & 0 \\ 2(1-\varepsilon)/(8-3\varepsilon) & 0 \end{pmatrix}, \; A^{\dagger}_{MN} = \begin{pmatrix} 0.5 & 0 \\ 0.25 & 0 \end{pmatrix},$$

and since in this case $\delta_M = 0$, the upper bounds given by (3.14)–(3.18), and (4.1) are reduced respectively to

$$\left\|A^{\dagger}_{\widehat{M}\widehat{N}}\right\| \le \frac{\|A^{\dagger}_{MN}\|}{1 - r_2\|\delta_N\|}, \tag{5.1}$$

$$\left\|A^{\dagger}_{\widehat{M}\widehat{N}} - A^{\dagger}_{MN}\right\| \le \frac{r_2\|\delta_N\|}{1 - r_2\|\delta_N\|} \, \|A^{\dagger}_{MN}\|, \tag{5.2}$$

$$\left\|A^{\dagger}_{\widehat{M}\widehat{N}}A - A^{\dagger}_{MN}A\right\| \le \frac{r_2\|\delta_N\|}{1 - r_2\|\delta_N\|} \, \left\|A^{\dagger}_{MN}A\right\|, \tag{5.3}$$

$$\left\|AA^{\dagger}_{\widehat{M}\widehat{N}} - AA^{\dagger}_{MN}\right\| = 0, \tag{5.4}$$

$$\|\widehat{x_0} - x_0\| \le \frac{\|\delta_b\| \, \|A^{\dagger}_{MN}\| + r_2\|\delta_N\| \, \|x_0\| \, \|A^{\dagger}_{MN}A\|}{1 - r_2\|\delta_N\|}. \tag{5.5}$$

Table 1. Numerical values of the upper bound (5.1)

ε	$\|A^{\dagger}_{\widehat{M}\widehat{N}}\|$	upper bound (5.1)	relative error
10^{-1}	0.58152242018800	0.59628479399994	2.5386 %
10^{-2}	0.56112821284186	0.56253282452825	0.2503 %
10^{-3}	0.55922677429243	0.55936659849901	0.0250 %

Table 2. Numerical values of the upper bound (5.2)

ε	$\|A^{\dagger}_{\widehat{M}\widehat{N}} - A^{\dagger}_{MN}\|$	upper bound (5.2)	relative error
10^{-1}	0.03629980482954	0.03726779962500	2.6667 %
10^{-2}	0.00350700749294	0.00351583015330	0.2516 %
10^{-3}	3.4952×10^{-4}	3.4960×10^{-4}	0.0250 %

Table 3. Numerical values of the upper bound (5.3)

ε	$\|A^{\dagger}_{\widehat{M}\widehat{N}}A - A^{\dagger}_{MN}A\|$	upper bound (5.3)	relative error
10^{-1}	0.08116883116883	0.08333333333333	2.6667 %
10^{-2}	0.00784190715182	0.00786163522013	0.2516 %
10^{-3}	7.8154×10^{-4}	7.8174×10^{-4}	0.0250 %

Table 4. Numerical values of the upper bound (5.5) $b = (1/25, 4)^T, \delta_b = (2\varepsilon, 0)^T$

ε	$\|\widehat{x_0} - x_0\|$	upper bound (5.5)	relative error
10^{-1}	0.11723791748695	0.12112034878124	3.3116 %
10^{-2}	0.01130784521208	0.01142644799823	1.0489 %
10^{-3}	0.00112690300346	0.00113621340320	0.8262 %

Example 5.2. Let $A = \begin{pmatrix} 1 & 0 \\ 2 & 0 \end{pmatrix}$, $a = b = 8$, $M = diag\,(a, b)$, $\widehat{N} = N = diag\,(1, 4)$ and $\delta_M = \begin{pmatrix} -\varepsilon & 0 \\ 0 & \varepsilon \end{pmatrix}$ for ε small enough. It is easy to verify that

$$A^{\dagger}_{\widehat{M}\widehat{N}} = \begin{pmatrix} \frac{a-\varepsilon}{a+4b+3\varepsilon} & \frac{2(b+\varepsilon)}{a+4b+3\varepsilon} \\ 0 & 0 \end{pmatrix}, \ A^{\dagger}_{MN} = \begin{pmatrix} \frac{a}{a+4b} & \frac{2b}{a+4b} \\ 0 & 0 \end{pmatrix},$$

and since in this case $\delta_N = 0$, the upper bounds given by (3.14)–(3.18), and (4.1) are reduced respectively to

$$\|A^{\dagger}_{\widehat{M}\widehat{N}}\| \leq \frac{(1 - \|M^{-1}\delta_M\|) \cdot \|A^{\dagger}_{MN}\|}{1 - \|M^{-1}\delta_M\| \left(1 + r_1\|M^{-1}\|\right)}, \tag{5.6}$$

$$\|A^{\dagger}_{\widehat{M}\widehat{N}} - A^{\dagger}_{MN}\| \leq \frac{r_1\|M^{-1}\delta_M\| \cdot \|M^{-1}\|}{1 - \|M^{-1}\delta_M\| \cdot \left(1 + r_1\|M^{-1}\|\right)} \|A^{\dagger}_{MN}\|, \tag{5.7}$$

$$\|A^{\dagger}_{\widehat{M}\widehat{N}}A - A^{\dagger}_{MN}A\| = 0, \tag{5.8}$$

$$\|AA^{\dagger}_{\widehat{M}\widehat{N}} - AA^{\dagger}_{MN}\| \leq \frac{r_1\|M^{-1}\delta_M\| \cdot \|M^{-1}\|}{1 - \|M^{-1}\delta_M\|\left(1 + r_1\|M^{-1}\|\right)} \|AA^{\dagger}_{MN}\|, \tag{5.9}$$

$$\|\widehat{x_0} - x_0\| \leq \frac{\left(1 - \|M^{-1}\delta_M\|\right)\left(X\|\delta_b\| + Y\|AA^{\dagger}_{MN}\|\right)\|A^{\dagger}_{MN}\|}{X^2}. \tag{5.10}$$

Table 5. Numerical values of the upper bound (5.6)

ε	$\|A^{\dagger}_{\widehat{M}\widehat{N}}\|$	upper bound (5.6)	relative error
10^{-1}	0.44723562396299	0.45294710313457	1.2771 %
10^{-2}	0.44721381877167	0.44777401353943	0.1253 %
10^{-3}	0.44721359773569	0.44726951117832	0.0125 %

Table 6. Numerical values of the upper bound (5.7)

ε	$\left\|A^{\dagger}_{\widehat{M}\widehat{N}} - A^{\dagger}_{MN}\right\|$	upper bound (5.7)	relative error
10^{-1}	0.00443884462035	0.00573350763461	29.1667 %
10^{-2}	4.4688×10^{-4}	5.6042×10^{-4}	25.4073 %
10^{-3}	4.4718×10^{-4}	5.5916×10^{-4}	25.0406 %

Table 7. Numerical values of the upper bound (5.9)

ε	$\left\|AA^{\dagger}_{\widehat{M}\widehat{N}} - AA^{\dagger}_{MN}\right\|$	upper bound (5.9)	relative error
10^{-1}	0.00992555831266	0.01282051282051	29.1667 %
10^{-2}	9.9925×10^{-4}	0.00125313283208	25.4073 %
10^{-3}	9.9992×10^{-4}	1.2503×10^{-4}	25.0406 %

Table 8. Numerical values of the upper bound (5.10)$b = (1/25, 25)^T$, $\delta_b = (0.1\varepsilon, 0)^T$

ε	$\|\widehat{x_0} - x_0\|$	upper bound (5.10)	relative error
10^{-1}	0.05142928039702	0.06924610482490	34.6433 %
10^{-2}	0.00517986510117	0.00670121463263	29.3704 %
10^{-3}	5.1836×10^{-4}	6.6796×10^{-4}	28.8607 %

References

[1] A. Ben-Tal and M. Teboulle, *A geometric property of the least squares solution of linear equations*, Linear Algebra Appl., 139(1990), 165–170.

[2] M. Benzi amd M. K. Ng, *Preconditioned iterative methods for weighted Toeplitz least squares problems*, SIAM J. Matrix Anal. Appl., 27(2006), 1106–1124.

[3] E. Bobrovnikova and S. Vavasis, *Accurate solution of weighted least squares by iterative methods*, SIAM J. Matrix Anal. Appl., 22(2001), 1153–1174.

[4] A. Forsgren, *On linear least-squares problems with diagonally domainant weight matrices*, SIAM J. Matrix Anal. Appl., 17(1996), 763–788.

[5] A. Forsgren and G. Sporre, *On weighted linear least-squares problems related to interior methods for convex quadratic programming*, SIAM J. Matrix Anal. Appl., 23(2001), 42–56.

[6] F. Graybill, *Matrices and Applications to Statistics*, 2nd Edition, Wadsworth, Belmont, CA, 1983.

[7] N. Karmarker, *A new polynomial time algorithm for linear programming*, Combinatorica, 4(1984), 373–395.

[8] K. N. Murthy and S. Sivasundaram, *Application of weighted Moore-Penrose inverse to two-point boundary value problems*, J. Math. Phys. Sci., 26(1992), 493–501.

[9] D. O'Leary, *On bounds for scaled projections and pseudoinverses*, Linear Algebra Appl., 132(1990), 115–117.

[10] G. W. Stewart, *On scaled projections and pseudoinverses*, Linear Algebra Appl., 112(1989), 189–193.

[11] M. J. Todd, *A Dantzig-Wolfe-like variant of Karmarkar's interior-point linear programming algorithm*, Oper. Res., 38(1990), 1006–1018.

[12] F. E. Udwadia and R. E. Kalaba, *Analytical Dynamics: a New Approach*, Cambridge University Press, Cambridge, England, 1996.

[13] G. Wang, Y. Wei and S. Qiao, *Generalized Inverses: Theory and Computations*, Science Press, Beijing, 2004.

[14] M. Wei, *Upper bound and stability of scaled pseudoinverses*, Numer. Math., 72(1995), 285–293.

[15] M. Wei, *Supremum and Stability of Weighted Pseudoinverses and Weighted Least Sqaures Problems: Analysis and Computations*, New York, Nova Science Publishers, 2001.

[16] M. Wei, *On stable perturbations of the stiffly weighted pseudoinverse and weighted least squares problem*, J. Comput. Math., 23(2005), 527–536.

[17] M. Wei and Q. Liu, *A numerically stable block modified Gram-Schmidt algorithm for solving stiff weighted least squares problems*, J. Comput. Math., 25(2007), 595–619.

[18] Z. Xu, J. Sun and C. Gu, *Perturbation for a pair of oblique projectors* $AA^{\dagger}_{MN}$ *and* $BB^{\dagger}_{MN}$, Appl. Math. Comput., 203(2008), 432–446.