Journal of Computational Mathematics, Vol.23, No.5, 2005, 527–536.

ON STABLE PERTURBATIONS OF THE STIFFLY WEIGHTED PSEUDOINVERSE AND WEIGHTED LEAST SQUARES PROBLEM [*1)]

Mu-sheng Wei
(*Department of Mathematics, East China Normal University,
Shanghai 200062, China*)

Abstract

In this paper we study perturbations of the stiffly weighted pseudoinverse $(W^{\frac{1}{2}}A)^{\dagger}W^{\frac{1}{2}}$ and the related stiffly weighted least squares problem, where both the matrices A and W are given with W positive diagonal and severely stiff. We show that the perturbations to the stiffly weighted pseudoinverse and the related stiffly weighted least squares problem are stable, if and only if the perturbed matrices $\widehat{A} = A+\delta A$ satisfy several row rank preserving conditions.

Mathematics subject classification: 65F20, 65F35.
Key words: Stiffly, Weighted pseudoinverse, Weighted least squares, Perturbation, Stability.

1. Introduction

Consider the following stiffly weighted least squares (stiffly WLS) problem

$$\min_{x\in C^n} \|W^{\frac{1}{2}}(Ax-b)\| = \min_{x\in C^n} \|D(Ax-b)\| \tag{1}$$

and related weighted pseudoinverse [12]

$$A_W^{\dagger} \equiv (W^{\frac{1}{2}}A)^{\dagger}W^{\frac{1}{2}} \text{ with } A_W = WA(WA)^{\dagger}A, \tag{2}$$

where $A \in C^{m\times n}$ with $\text{rank}(A) = r$, $b \in C^m$ are known coefficient matrix and observation vector, respectively, $\|\cdot\| \equiv \|\cdot\|_2$ is the Euclidian vector norm or subordinate matrix norm,

$$D = \text{diag}(d_1, d_2, \cdots, d_m) = \text{diag}(w_1^{\frac{1}{2}}, w_2^{\frac{1}{2}}, \cdots, w_m^{\frac{1}{2}}) = W^{\frac{1}{2}} \tag{3}$$

is the weight matrix. The stiffly WLS problem Eq. (1) with extremely ill-conditioned weight matrix W, where the scalar factors $w_1, \cdots, w_m$ vary widely in size, arises from many areas of applied science, such as in electronic network, certain classes of finite element problems, interior point methods for constrained optimization (e.g., see [8]), and for solving the equality constrained least squares problem by the method of weighting [9, 1, 11], etc.

The stability conditions of the stiffly weighted pseudoinverse and the stiffly WLS problem are important subjects in both theoretical and computational point of view. Wei [11, 12, 13] studied the stability of weighted pseudoinverses and constrained weighted pseudoinverses when the weight matrix W ranges over a set $\mathcal{D}$ of positive diagonal matrices, and obtained necessary and sufficient stability conditions:

$$\textit{if and only if any } r \textit{ rows of the matrix } A \textit{ are linearly independent.} \tag{4}$$

* Received June 15, 2003; final revised November 22, 2004.
1) This work is supported by NSFC under grant No. 10371044, Science and Technology Commission of Shanghai Municipality grant No. 04JC14031, and Shanghai Priority Academic Discipline, China.

Based on these results, Wei and De Pieero [16] obtained stability conditions and upper perturbation bounds of WLS and equality constrained least squares problems when weight matrices W range over $\mathcal{D}$.

In practical scientific computations, however, the above condition is too restrictive, and the weight matrix W is usually fixed and severely stiff. In [14], the author found that in this case, the stiffly weighted pseudoinverse is close to a related multi-level constrained pseudoinverse $A_C^\dagger$ and the solution set of Eq. (1) is close to a related multi-level constrained least squares (MCLS) problem. Based on the findings in [14], in this paper we will derive the stability conditions of the stiffly weighted pseudoinverse and the stiffly WLS problem.

Without loss of generality, we make the following notation and assumptions for the matrices A and W.

Assumption 1.1. *The matrices A and W in Eq. (1) satisfy the following conditions: $\|A(i,:)\|$ have the same order for $i = 1, \cdots, m$, $w_1 > w_2 > \cdots > w_k > 0$, $m_1 + m_2 + \cdots + m_k = m$, and we denote*

$$A = \begin{pmatrix} A_1 \\ \vdots \\ A_k \end{pmatrix} \begin{matrix} m_1 \\ \vdots \\ m_k \end{matrix}, \quad C_j = \begin{pmatrix} A_1 \\ \vdots \\ A_j \end{pmatrix}, \quad j = 1, \cdots, k, \tag{5}$$

$$\begin{aligned} &W = \operatorname{diag}(w_1 I_{m_1}, w_2 I_{m_2}, \cdots, w_k I_{m_k}), \\ &0 < \epsilon_{ij} \equiv \tfrac{w_i}{w_j} \ll 1, \text{ for } 1 \le j < i \le k \text{ so } \epsilon = \max_{1 \le j < k} \{\epsilon_{j+1,j}\} \ll 1. \end{aligned} \tag{6}$$

We also set

$$P_0 = I_n, \quad P_j = I - C_j^\dagger C_j, \quad \operatorname{rank}(C_j) = r_j, \quad j = 1, \cdots, k. \tag{7}$$

The paper is organized as follows. In §2 we will review some preliminary results related to the weighted pseudoinverse; in §3 we will study stability conditions for the stiffly weighted pseudoinverse; in §4 we will deduce perturbation bounds for the solutions of the stiffly WLS problem Eq. (1); in §5 we will provide several numerical examples to verify our findings; finally in §6 we will conclude the paper with some remarks.

2. Preliminaries

In this section we provide some preliminary results which are necessary for our further discussion.

Lemma 2.1. [4] *Suppose that $D, E \in C^{m\times n}$ and $\operatorname{rank}(D) = \operatorname{rank}(E)$. Then*

$$\begin{aligned} &\|DD^\dagger - EE^\dagger\| \le \min\{\|(D-E)D^\dagger\|, \|(D-E)E^\dagger\|, 1\}, \\ &\|D^\dagger D - E^\dagger E\| \le \min\{\|D^\dagger(D-E)\|, \|E^\dagger(D-E)\|, 1\}. \end{aligned} \tag{8}$$

Lemma 2.2. [14] *Under the notation of Assumption 1.1,*

$$\begin{aligned} &(A_j P_{j-1})^\dagger A_j P_{j-1} = C_j^\dagger C_j - C_{j-1}^\dagger C_{j-1}, \\ &\operatorname{rank}(A_j P_{j-1}) = \operatorname{rank}(C_j) - \operatorname{rank}(C_{j-1}) = r_j - r_{j-1} \end{aligned} \tag{9}$$

for $j = 2, \cdots, k$. Denote $(A_j P_{j-1})^H = Q_j R_j$ the unitary decomposition of $(A_j P_{j-1})^H$ (A_j^H is the conjugate transpose of the matrix A_j), where $Q_j^H Q_j = I_{r_j - r_{j-1}}$ and R_j has full row rank $r_j - r_{j-1}$. Then for $j = 1, \cdots, k$,

$$(Q_1, \cdots, Q_j)^H (Q_1, \cdots, Q_j) = I_{r_j}, \ C_j^\dagger C_j = \sum_{l=1}^{j} Q_l Q_l^H, \tag{10}$$

$$A_j P_{j-1} = A_j Q_j Q_j^H, \ (A_j P_{j-1})^\dagger = Q_j (A_j Q_j)^\dagger. \tag{11}$$

Lemma 2.3. [14] *Under the notation in Assumption 1.1,*

$$\begin{array}{rl} A_W &= B_\epsilon B_\epsilon^\dagger A = A_\epsilon A_\epsilon^\dagger A = (B_\epsilon^\dagger)^H B_\epsilon^H B_1 Q^H, \\ A_W^\dagger &= (W^{\frac{1}{2}}A)^\dagger W^{\frac{1}{2}} = Q(B_\epsilon^H B_1)^{-1} B_\epsilon^H, \\ B_\epsilon &= \begin{pmatrix} A_1Q_1 & 0 & \cdots & 0 \\ \epsilon_{21}A_2Q_1 & A_2Q_2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ \epsilon_{k1}A_kQ_1 & \epsilon_{k2}A_kQ_2 & \cdots & A_kQ_k \end{pmatrix}, \end{array} \tag{12}$$

in which B_ϵ has full column rank $r_k = \text{rank}(A)$ and B_1 is obtained from B_ϵ by replacing all ϵ_{ij} in B_ϵ with ones.

Lemma 2.4. [6, 10] *Let $A \in C_r^{m\times n}$ and $\widehat{A} = A + \delta A \in C^{m\times n}$. Then we have the following results.*

1. *If $\|\delta A\| \cdot \|A^\dagger\| < 1$, then $\text{rank}(\widehat{A}) \geq \text{rank}(A)$.*
2. *If $\|\delta A\| \cdot \|A^\dagger\| < 1$ and $\text{rank}(\widehat{A}) > \text{rank}(A)$, then $\|\widehat{A}^\dagger\| \geq \frac{1}{\|\delta A\|}$.*
3. *If $\|\delta A\| \cdot \|A^\dagger\| < 1$ and $\text{rank}(\widehat{A}) = \text{rank}(A)$, then*

$$\frac{\|A^\dagger\|}{1 + \|\delta A\| \cdot \|A^\dagger\|} \leq \|\widehat{A}^\dagger\| \leq \frac{\|A^\dagger\|}{1 - \|\delta A\| \cdot \|A^\dagger\|}.$$

So $\|\widehat{A}^\dagger\|$ is bounded for all small perturbations δA with

$$\|\delta A\| \cdot \|A^\dagger\| \leq \eta < 1, \text{ if and only if } \text{rank}(\widehat{A}) = \text{rank}(A),$$

where $0 \leq \eta < 1$ is a constant.

3. Stability Conditions for the Stiffly Weighted Pseudoinverse

In this section we will study the stability conditions of perturbations of the stiffly weighted pseudoinverse $A_W^\dagger$. For $j = 1, 2, \cdots, k$, let

$$\begin{array}{ll} \widehat{A}_j = A_j + \delta A_j, & \widehat{C}_j = C_j + \delta C_j, \\ \widehat{P}_j = I - \widehat{C}_j^\dagger \widehat{C}_j, & \widehat{w}_j = w_j + \delta w_j, \end{array} \tag{13}$$

be perturbed versions of A_j, C_j, P_j, w_j, respectively, and denote

$$\eta = \max_{1\leq j<i\leq k} \frac{|\widehat{\epsilon}_{ij} - \epsilon_{ij}|}{\epsilon_{ij}}, \ \widehat{\epsilon} = \max_{1\leq j<k} \{\widehat{\epsilon}_{j+1,j}\}. \tag{14}$$

We first define the stability of the stiffly weighted pseudoinverse.

Definition 3.1. *We say that the perturbations to the stiffly weighted pseudoinverse $A_W^\dagger$ are stable, if when $\eta \to 0$, $\epsilon_{ij} \to 0$ for $1 \leq j < i \leq k$, and perturbation $\delta A_j \to 0$ for $j = 1, \cdots, k$, all perturbed stiffly weighted pseudoinverses $\|\widehat{A}_{\widehat{W}}^\dagger\|$ are uniformly bounded, and $\widehat{A}_{\widehat{W}}^\dagger \to A_W^\dagger$ uniformly.*

In the remaining of this section we will study the stability conditions of the stiffly weighted pseudoinverse. We will show that the perturbed stiffly weighted pseudoinverse is stable, if and only if the perturbation for the matrix A satisfies the following assumption.

Assumption 3.1.

$$\text{rank}(\widehat{C}_j) = \text{rank}(C_j) = r_j, \quad j = 1, 2, \cdots, k, \tag{15}$$

or equivalently,

$$\text{rank}(\widehat{A}_j \widehat{P}_{j-1}) = \text{rank}(A_j P_{j-1}) = r_j - r_{j-1}, \quad j = 1, 2, \cdots, k. \tag{16}$$

Theorem 3.1. *Suppose that A and W are given matrices satisfying the notation and conditions in Assumption 1.1, the perturbed matrices $\widehat{C}_j$ for $j = 1, \cdots, k$ satisfy Assumption 3.1, and*

$$E \cdot \|A_W^\dagger\| < 1 \text{ with}$$
$$E \equiv \|\delta A\| + \|A\| \cdot \|B_\epsilon^\dagger\| \cdot \left(\tfrac{\eta\epsilon}{1-\epsilon} \max_{1\le j<i\le k} \|A_i Q_j\| + \tfrac{1}{1-\epsilon(1+\eta)} \max_{1\le j\le i\le k} (\|\delta A_i\| + 2\sqrt{2}\|A_i\| \cdot \|C_j^\dagger \delta C_j\|)\right). \tag{17}$$

Then we have the following estimates

$$\|\widehat{A}_{\widehat{W}}^\dagger\| \le \frac{\|A_W^\dagger\|}{1 - E\cdot\|A_W^\dagger\|}, \quad \|\widehat{A}_{\widehat{W}}^\dagger - A_W^\dagger\| \le \frac{\sqrt{5}+1}{2} \cdot E \cdot \frac{\|A_W^\dagger\|^2}{1 - E \cdot \|A_W^\dagger\|}. \tag{18}$$

Proof. Let $\widehat{A}_{\widehat{W}} = \widehat{B}_{\widehat{\epsilon}} \widehat{B}_{\widehat{\epsilon}}^\dagger \widehat{A}$ and $\widehat{Q}_j$ be perturbed versions of $A_W = B_\epsilon B_\epsilon^\dagger A$ and Q_j, respectively. According to Lemma 2.1 of [17], there exist unitary matrices U_j of order $r_j - r_{j-1}$, such that

$$\|Q_j - \widehat{Q}_j U_j\| \le \sqrt{2}\|Q_j Q_j^H - \widehat{Q}_j \widehat{Q}_j^H\| \text{ for } j = 1, \cdots k.$$

Without loss of generality we can set $\widehat{Q}_j := \widehat{Q}_j U_j$ and so

$$\|Q_j - \widehat{Q}_j\| \le \sqrt{2}\|Q_j Q_j^H - \widehat{Q}_j \widehat{Q}_j^H\| \text{ for } j = 1, \cdots k. \tag{19}$$

From the conditions of the theorem we have by applying Lemma 2.1,

$$\begin{aligned} \|\widehat{A}_{\widehat{W}} - A_W\| = & \|\widehat{B}_{\widehat{\epsilon}} \widehat{B}_{\widehat{\epsilon}}^\dagger \widehat{A} - B_\epsilon B_\epsilon^\dagger A\| \\ \le & \|\widehat{B}_{\widehat{\epsilon}} \widehat{B}_{\widehat{\epsilon}} \delta A\| + \|(\widehat{B}_{\widehat{\epsilon}} \widehat{B}_{\widehat{\epsilon}}^\dagger - B_\epsilon B_\epsilon^\dagger) A\| \\ \le & \|\delta A\| + \|\widehat{B}_{\widehat{\epsilon}} - B_\epsilon\| \cdot \|B_\epsilon^\dagger\| \cdot \|A\|. \end{aligned} \tag{20}$$

Now

$$\|\widehat{B}_{\widehat{\epsilon}} - B_\epsilon\| \le E_1 + E_2,$$

in which

$$E_1 = \left\| \begin{pmatrix} 0 & 0 & \cdots & 0 & 0 \\ (\widehat{\epsilon}_{21} - \epsilon_{21}) A_2 Q_1 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & & \vdots & \vdots \\ (\widehat{\epsilon}_{k1} - \epsilon_{k1}) A_k Q_1 & (\widehat{\epsilon}_{k2} - \epsilon_{k2}) A_k Q_2 & \cdots & (\widehat{\epsilon}_{k,k-1} - \epsilon_{k,k-1}) A_k Q_{k-1} & 0 \end{pmatrix} \right\|$$
$$\le \|\mathrm{diag}((\widehat{\epsilon}_{21} - \epsilon_{21}) A_2 Q_1, \cdots, (\widehat{\epsilon}_{k,k-1} - \epsilon_{k,k-1}) A_k Q_{k-1})\|$$
$$+ \|\mathrm{diag}((\widehat{\epsilon}_{31} - \epsilon_{31}) A_3 Q_1, \cdots, (\widehat{\epsilon}_{k,k-2} - \epsilon_{k,k-2}) A_k Q_{k-2})\| + \cdots + \|(\widehat{\epsilon}_{k1} - \epsilon_{k1}) A_k Q_1\|$$
$$\le \eta(\epsilon \max_{1\le j<k} \|A_{j+1} Q_j\| + \epsilon^2 \max_{1\le j<k-1} \|A_{j+2} Q_j\| + \cdots + \epsilon^{k-1} \|A_k Q_1\|)$$
$$\le \eta(\epsilon + \epsilon^2 + \cdots + \epsilon^{k-1}) \max_{1\le j<i\le k} \|A_i Q_j\| \le \tfrac{\eta\epsilon}{1-\epsilon} \max_{1\le j<i\le k} \|A_i Q_j\|,$$

and

$$E_2 = \left\| \begin{pmatrix} \widehat{A}_1 \widehat{Q}_1 - A_1 Q_1 & 0 & \cdots & 0 \\ \widehat{\epsilon}_{21}(\widehat{A}_2 \widehat{Q}_1 - A_2 Q_1) & \widehat{A}_2 \widehat{Q}_2 - A_2 Q_2 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ \widehat{\epsilon}_{k1}(\widehat{A}_k \widehat{Q}_1 - A_k Q_1) & \widehat{\epsilon}_{k2}(\widehat{A}_k \widehat{Q}_2 - A_k Q_2) & \cdots & \widehat{A}_k \widehat{Q}_k - A_k Q_k \end{pmatrix} \right\|$$
$$\le \|\mathrm{diag}(\widehat{A}_1 \widehat{Q}_1 - A_1 Q_1, \cdots, \widehat{A}_k \widehat{Q}_k - A_k Q_k)\|$$
$$+ \|\mathrm{diag}(\widehat{\epsilon}_{21}(\widehat{A}_2 \widehat{Q}_1 - A_2 Q_1), \cdots, \widehat{\epsilon}_{k,k-1}(\widehat{A}_k \widehat{Q}_{k-1} - A_k Q_{k-1}))\| + \cdots + \|\widehat{\epsilon}_{k1}(\widehat{A}_k \widehat{Q}_1 - A_k Q_1)\|$$
$$\le \max_{1\le j\le k} \|\widehat{A}_j \widehat{Q}_j - A_j Q_j\| + \widehat{\epsilon} \max_{1\le j<k} \|\widehat{A}_{j+1} \widehat{Q}_j - A_{j+1} Q_j\| + \cdots + \widehat{\epsilon}^{k-1} \|\widehat{A}_k \widehat{Q}_1 - A_k Q_1\|$$
$$\le (1 + \widehat{\epsilon} + \cdots + \widehat{\epsilon}^{k1}) \max_{1\le j\le i\le k} \|\widehat{A}_i \widehat{Q}_j - A_i Q_j\| \le \tfrac{1}{1-\widehat{\epsilon}} \max_{1\le j\le i\le k} \|\widehat{A}_i \widehat{Q}_j - A_i Q_j\|.$$

Now by applying Eq. (19) we have

$$\begin{aligned}\|\widehat{A}_i\widehat{Q}_j - A_iQ_j\| &\le \|\delta A_i\widehat{Q}_j\| + \|A_i(\widehat{Q}_j - Q_j)\|\\ &\le \|\delta A_i\widehat{Q}_j\| + \sqrt{2}\|A_i\|\|\widehat{Q}_j\widehat{Q}_j^H - Q_jQ_j^H\|\end{aligned}$$

for $1 \le j \le i \le k$, and then by applying Lemmas 2.1–2.2 we obtain

$$\begin{aligned}\|\widehat{Q}_j\widehat{Q}_j^H - Q_jQ_j^H\| &= \|(\widehat{C}_j^\dagger\widehat{C}_j - \widehat{C}_{j-1}^\dagger\widehat{C}_{j-1}) - (C_j^\dagger C_j - C_{j-1}^\dagger C_{j-1})\|\\ &\le \|\widehat{C}_j^\dagger\widehat{C}_j - C_j^\dagger C_j\| + \|\widehat{C}_{j-1}^\dagger\widehat{C}_{j-1} - C_{j-1}^\dagger C_{j-1}\|\\ &\le \|C_j^\dagger\delta C_j\| + \|C_{j-1}^\dagger\delta C_{j-1}\|.\end{aligned}$$

By substituting the above inequalities into that for E_2, and then substituting the bounds for E_1 and E_2 into Eq. (20), we deduce that $\|\widehat{A}_{\widehat{W}} - A_W\| \le E$, and by applying Lemma 2.4 we prove the first inequality of Eq. (18).

From the following decomposition

$$\begin{aligned}\widehat{A}_{\widehat{W}}^\dagger - A_W^\dagger &= -\widehat{A}_{\widehat{W}}^\dagger(\widehat{A}_{\widehat{W}} - A_W)A_W^\dagger + \widehat{A}_{\widehat{W}}^\dagger(I - A_WA_W^\dagger)\\ &\quad -(I - \widehat{A}_{\widehat{W}}^\dagger\widehat{A}_{\widehat{W}})A_W^\dagger A_WA_W^\dagger,\end{aligned}$$

we have for any $0 \ne x \in C^n$ (similar to the derivation in [10]),

$$\begin{aligned}\|(\widehat{A}_{\widehat{W}}^\dagger - A_W^\dagger)x\|^2 &\le (\|\widehat{A}_{\widehat{W}}^\dagger\|\|\widehat{A}_{\widehat{W}} - A_W\|\|A_W^\dagger\|)^2\\ &\quad\times((\|A_WA_W^\dagger x\| + \|(I - A_WA_W^\dagger)x\|)^2 + \|A_WA_W^\dagger x\|^2)\\ &\le (E\cdot\|\widehat{A}_{\widehat{W}}^\dagger\|\cdot\|A_W^\dagger\|)^2\|x\|^2(\tfrac{\sqrt{5}+1}{2})^2,\end{aligned}$$

so we have

$$\|\widehat{A}_{\widehat{W}}^\dagger - A_W^\dagger\| \le \frac{\sqrt{5}+1}{2}\cdot E\cdot\|\widehat{A}_{\widehat{W}}^\dagger\|\cdot\|A_W^\dagger\|,$$

obtaining the second inequality of Eq. (18).

We now consider some special cases that can guarantee the conditions in Assumption 3.1 and so the stability of the perturbations. To simplify the notation we set $C_0 = \emptyset$, $r_0 = m_0 = 0$ and $M_i = \sum_{l=0}^{i} m_l$ for $i = 0, 1, \cdots, k$.

Corollary 3.1. *Suppose that A and W are given matrices satisfying the notation and conditions in Assumption 1.1, and A satisfies*

$$\operatorname{rank}(C_i) = \min\{M_i,\ n\}\ \text{for}\ i = 1, \cdots, k. \tag{21}$$

If the perturbations satisfy Eq. (17) of Theorem 3.1, then the estimate in Eq. (18) holds.

Proof. If the conditions of the corollary hold, then by applying Lemma 2.4 we observe that, for sufficiently small perturbations δC_i , $\widehat{C}_i$ and C_i should have the same rank $\min\{M_i,\ n\}$ for $i = 1, \cdots, k$, so the perturbations satisfy Assumption 3.1 and the perturbations are stable.

Corollary 3.2. *Suppose that A and W are given matrices satisfying the notation and conditions in Assumption 1.1, and A satisfies*

$$\operatorname{rank}(A) < \min\{M_k,\ n\}\ \text{and}\ \operatorname{rank}(C_{k-1}) = M_{k-1}. \tag{22}$$

When the perturbations satisfy

$$\operatorname{rank}(\widehat{A}) = \operatorname{rank}(A), \tag{23}$$

and the inequality in Eq. (17), then the estimates in Eq. (18) hold.

Proof. Because C_{k-1} has full row rank M_{k-1}, so all $C_1, \cdots, C_{k-1}$ have full row ranks. Therefore, when the perturbations are sufficiently small, $\widehat{C}_1, \cdots, \widehat{C}_k$ satisfy Assumption 3.1 and Theorem 3.1 is applicable.

We now study the situation when the perturbations to the stiff weighted pseudoinverse are unstable. In terms of Corollaries 3.1-3.2, we can exclude the cases mentioned in Eqs. (21)–(23).

Case 1. $\operatorname{rank}(A) < \min\{m,\ n\}$, and we allow $\operatorname{rank}(\widehat{A}) > \operatorname{rank}(A)$.

Theorem 3.2. *Suppose that A and W are given matrices satisfying the notation and conditions in Assumption 1.1,* $\mathrm{rank}(A) < \min\{m,\ n\}$. *Then for any value $0 < \xi \ll 1$, there exists a perturbed matrix $\widehat{A} = A + \delta A$ satisfying $\|\delta A\| = \xi$,*

$$\mathrm{rank}(\widehat{A}) > \mathrm{rank}(A), \tag{24}$$

$$\|\widehat{A}_W^\dagger\| \geq \frac{1}{\xi} \ \textit{and}\ \|\widehat{A}_W^\dagger - A_W^\dagger\| \geq \frac{1}{\xi}. \tag{25}$$

Proof. From the condition of the theorem, $\mathcal{N}(A) \neq \{0\}$, $\mathcal{N}(A^H W) \neq \{0\}$. Therefore we can pick vectors $q \in \mathcal{N}(A)$, $f \in \mathcal{N}(A^H W)$ with $\|q\| = 1$, $\|f\| = 1$. Define

$$\widehat{A} = A + \delta A,\ \delta A = \xi f q^H,$$

then $\mathrm{rank}(\widehat{A}) = r + 1$, and from the facts $A_W^\dagger f = 0$, $Aq = 0$,

$$\begin{aligned} \widehat{A}_W^\dagger &= (\widehat{A}^H W \widehat{A})^\dagger \widehat{A}^H W \\ &= (A^H W A + \xi^2 (f^H W f) q q^H)^\dagger (A^H W + \xi q f^H W) \\ &= A_W^\dagger + (\xi f^H W f)^{-1} q f^H W, \end{aligned}$$

$$\begin{aligned} \|\widehat{A}_W^\dagger\| &\geq \|\widehat{A}_W^\dagger f\| = \|(\xi f^H W f)^{-1} q f^H W f\| = \tfrac{1}{\xi}, \\ \|\widehat{A}_W^\dagger - A_W^\dagger\| &\geq \|(\widehat{A}_W^\dagger - A_W^\dagger) f\| = \|(\xi f^H W f)^{-1} q f^H W f\| = \tfrac{1}{\xi}, \end{aligned}$$

proving the assertions of the theorem.

Case 2. The condition $\mathrm{rank}\widehat{A} = \mathrm{rank}(A) = r$ holds, but other conditions in Assumption 3.1 do not hold. Then we need the following result.

Lemma 3.1. *Suppose that $L \in C^{m\times m}$, $K \in C^{m\times n}$, $M \in C^{n\times m}$, $N \in C^{n\times n}$, and $D = \begin{pmatrix} L & K \\ M & N \end{pmatrix}$ such that L and D are nonsingular. Then $N - ML^{-1}K$ is also nonsingular, and*

$$D^{-1} = \begin{pmatrix} L^{-1} + L^{-1}K(N - ML^{-1}K)^{-1}ML^{-1} & -L^{-1}K(N - ML^{-1}K)^{-1} \\ -(N - ML^{-1}K)^{-1}ML^{-1} & (N - ML^{-1}K)^{-1} \end{pmatrix}. \tag{26}$$

Theorem 3.3. *Suppose that A and W are given matrices satisfying the notation and conditions in Assumption 1.1, and* $\mathrm{rank}(A) = r$. *Suppose that there exists an integer i with $1 \leq i < k$, such that*

$$\mathrm{rank}(C_{i-1}) = M_{i-1},\ \mathit{rank}(C_i) < \min\{M_i, n\} \leq n. \tag{27}$$

Let l be the largest integer satisfying $k \geq l > i$ and

$$\mathrm{rank}(C_{l-1}) < r,\ \mathrm{rank}(C_l) = r. \tag{28}$$

Then for any value $0 < \xi \ll 1$, there exists a perturbed matrix $\widehat{A} = A + \delta A$ satisfying $\|\delta A\| = \xi$,

$$\mathrm{rank}(\widehat{C}_i) > \mathrm{rank}(C_i),\ \mathrm{rank}(\widehat{A}) = \mathrm{rank}(A) = r, \tag{29}$$

$$\|\widehat{A}_W^\dagger\| \geq \tfrac{\xi}{\xi^2 + a\epsilon_{li}},\ \|\widehat{A}_W^\dagger - A_W^\dagger\| \geq \tfrac{\xi}{\xi^2 + a\epsilon_{li}}, \tag{30}$$

in which $a > 0$ is a constant which is independent of the parameter ξ.

Proof. Let $Q_1, \cdots, Q_k$ be as in Lemma 2.2. We observe that $r - \mathrm{rank}(C_{l-1}) > 0$ and Q_l is an n by $r - \mathrm{rank}(C_{l-1})$ matrix. Define $Q_l \equiv (Q_{l_1}, q_{l_2})$ where q_{l_2} is the last column of Q_l. There exists a unit vector $f_i \in C^{M_i}$ such that $f_i^H W_i C_i = 0$. Then $f_i(M_{i-1}+1 : M_i) \neq 0$. (If $f_i(M_{i-1}+1 : M_i) = 0$, then we have $f_i(1 : M_{i-1})^H W_{i-1} C_{i-1} = 0$, so $f_i(1 : M_{i-1}) = 0$ because $W_{i-1}C_{i-1}$ has full row rank, a contradiction.) Define

$$f = \begin{pmatrix} f_i \\ 0 \end{pmatrix} \in C^m,\ \widetilde{Q} = (Q_1, \cdots, Q_{l-1}, Q_{l_1}, \cdots Q_k),\ \delta A = \xi f q_{l_2}^H. \tag{31}$$

Notice that $W^{\frac{1}{2}}AQ$ has full column rank r, $W^{\frac{1}{2}}A\widetilde{Q}$ has full column rank $r-1$, and $W^{\frac{1}{2}}Aq_{l_2}$ is a column of $W^{\frac{1}{2}}AQ$. Now suppose θ is the angle between $\mathcal{R}(W^{\frac{1}{2}}A\widetilde{Q})$ and $\mathcal{R}(W^{\frac{1}{2}}Aq_{l_2})$, then $0<\theta\le\pi/2$ and

$$\begin{aligned}\sin^2\theta\|W^{\frac{1}{2}}Aq_{l_2}\|^2 &= q_{l_2}^H A^H W^{\frac{1}{2}}(I-W^{\frac{1}{2}}A\widetilde{Q}(W^{\frac{1}{2}}A\widetilde{Q})^\dagger)W^{\frac{1}{2}}Aq_{l_2}\\ &\le w_l\sin^2\theta\|Aq_{l_2}\|^2\equiv aw_l,\end{aligned} \tag{32}$$

because $A_1q_{l_2}=0,\cdots,A_{l-1}q_{l_2}=0$ from Lemma 2.2. By choosing $\widehat{A}=A+\delta A$ and noticing $f^HWA=0$, we observe

$$\|\delta A\|=\xi,\ \operatorname{rank}(\widehat{C}_i)=\operatorname{rank}(C_i)+1,\ \operatorname{rank}(\widehat{A})=\operatorname{rank}(A).$$

From the identities $\widehat{A}_W^\dagger=(\widehat{A}^HW\widehat{A})^\dagger\widehat{A}^HW$ and

$$\begin{aligned}(\widehat{A}^HW\widehat{A})^\dagger &= (QQ^HA^HWAQQ^H+\xi^2(f^HWf)q_{l_2}q_{l_2}^H)^\dagger\\ &=(\widetilde{Q},q_{l_2})\begin{pmatrix}\widetilde{Q}^HA^HWA\widetilde{Q} & \widetilde{Q}^HA^HWAq_{l_2}\\ q_{l_2}^HA^HWA\widetilde{Q} & q_{l_2}^HA^HWAq_{12}+\xi^2(f_i^HW_if_i)\end{pmatrix}^{-1}(\widetilde{Q},q_{l_2})^H,\end{aligned} \tag{33}$$

we deduce by applying Lemma 3.1 and Eqs. (32)–(33),

$$\begin{aligned}\|\widehat{A}_W^\dagger\| &\ge \|q_{l_2}^H\widehat{A}_W^\dagger f\|\\ &= \frac{\xi(f_i^HW_if_i)}{\xi^2(f_i^HW_if_i)+q_{l_2}^HA^HW^{\frac{1}{2}}(I-W^{\frac{1}{2}}A\widetilde{Q}(W^{\frac{1}{2}}A\widetilde{Q})^\dagger)W^{\frac{1}{2}}Aq_{l_2}}\\ &\ge \frac{\xi(f_i^HW_if_i)}{\xi^2(f_i^HW_if_i)+aw_l}=\frac{\xi}{\xi^2+(aw_l/f_i^HW_if_i)}\ge\frac{\xi}{\xi^2+a\epsilon_{li}},\\ \|\widehat{A}_W^\dagger-A_W^\dagger\| &\ge \|q_{l_2}^H(\widehat{A}_W^\dagger-A_W^\dagger)f\|=\|q_{l_2}^H\widehat{A}_W^\dagger f\|\ge\frac{\xi}{\xi^2+a\epsilon_{li}},\end{aligned}$$

because $f_i^HW_i^2f_i\ge w_if_i^Hf_i=w_i$.

We now summarize the results of this section in the following theorem.

Theorem 3.4. *Suppose that A and W are given matrices satisfying the notation and conditions in Assumption 1.1. Then perturbed stiffly weighted pseudoinverse $\widehat{A}_{\widehat{W}}^\dagger$ is stable, if and only if the perturbations satisfy*

$$\operatorname{rank}(\widehat{C}_j)=\operatorname{rank}(C_j)\ \text{for}\ j=1,\cdots,k. \tag{34}$$

4. Perturbation Bounds for the Stiffly Weighted Least Squares Problem

Having analyzing the stability conditions for the perturbations of the stiffly weighted pseudoinverse, in this section we will provide a perturbation analysis for the stiffly WLS problem Eq. (1). We have

Theorem 4.1. *Consider the stiffly WLS problem Eq. (1), in which the matrices A and W satisfy the conditions and notation of Assumption 1.1. Let $\widehat{A}_j=A_j+\delta A_j,\widehat{C}_j=C_j+\delta C_j,\widehat{w}_j=w_j+\delta w_j$, for $j=1,\cdots,k$, and $\widehat{b}=b+\delta b$ are perturbed version of A_j,C_j,w_j and b, respectively. Suppose that the perturbations satisfy Assumption 3.1, and E defined in Theorem 3.1 satisfies $E\|A_W^\dagger\|<1$. Consider the following perturbed WLS problem*

$$\min_{x\in C^n}\{\|\widehat{W}^{\frac{1}{2}}(\widehat{A}x-\widehat{b})\|\}. \tag{35}$$

Then for the minimum norm solutions x_{WLS} of Eq. (1) and $\widehat{x}_{WLS}$ of Eq. (35) with $\delta x_{WLS}=\widehat{x}_{WLS}-x_{WLS}$,

$$\begin{aligned}\|\delta x_{WLS}\| &\le \frac{\|A_W^\dagger\|}{1-E\|A_W^\dagger\|}(\|\delta b\|+\|\delta A\|\|x_{WLS}\|+E\|A_W^\dagger\|\|r(x_{WLS})\|)\\ &\quad+\delta_{rn}\|\delta A\|\|A^\dagger\|\|x_{WLS}\|,\end{aligned} \tag{36}$$

in which $r(x_{WLS}) = b - Ax_{WLS}$ *is the residual vector,* $\delta_{rn} = 0$ *for* $r = n$ *and* $\delta_{rn} = 1$ *for* $r < n$.

Proof. We can use the following decomposition of $\widehat{A}^{\dagger}_{\widehat{W}} - A^{\dagger}_{W}$ [16],

$$\widehat{A}^{\dagger}_{\widehat{W}} - A^{\dagger}_{W} = -\widehat{A}^{\dagger}_{\widehat{W}}\delta A A^{\dagger}_{W} + \widehat{A}^{\dagger}_{\widehat{W}}(I - AA^{\dagger}_{W}) - (I - \widehat{A}^{\dagger}\widehat{A})A^{\dagger}AA^{\dagger}_{W}, \tag{37}$$

and apply the identity $(I - A_W A^{\dagger}_{W})(I - AA^{\dagger}_{W}) = (I - AA^{\dagger}_{W})$, to obtain

$$\begin{aligned} \delta x_{WLS} &= \widehat{A}^{\dagger}_{\widehat{W}}\widehat{b} - A^{\dagger}_{W}b = \widehat{A}^{\dagger}_{\widehat{W}}\delta b + (\widehat{A}^{\dagger}_{\widehat{W}} - A^{\dagger}_{W})b \\ &= \widehat{A}^{\dagger}_{\widehat{W}}(\delta b - \delta A x_{WLS}) + \widehat{A}^{\dagger}_{\widehat{W}}(I - A_W A^{\dagger}_{W})r(x_{WLS}) \\ &-(I - \widehat{A}^{\dagger}\widehat{A})A^{\dagger}Ax_{WLS}. \end{aligned} \tag{38}$$

Furthermore, by applying Lemma 2.1 and Theorem 3.1 we have

$$\begin{aligned} &\|\widehat{A}_{\widehat{W}}\widehat{A}^{\dagger}_{\widehat{W}}(I - A_W A^{\dagger}_{W})\| \leq \|\widehat{A}_{\widehat{W}} - A_W\|\|A^{\dagger}_{W}\| \leq E\|A^{\dagger}_{W}\|, \\ &\|(I - \widehat{A}^{\dagger}\widehat{A})A^{\dagger}A\| = 0 \text{ for } r = n, \\ &\|(I - \widehat{A}^{\dagger}\widehat{A})A^{\dagger}A\| \leq \|\delta A\|\|A^{\dagger}\| \text{ for } r < n. \end{aligned}$$

By taking norms in both sides of Eq. (38) and substituting the above inequalities, we obtain the desired estimate in Eq. (36).

Theorem 4.2. *If in Theorem 4.1,* $r < n$, *and the perturbations satisfy Assumption 3.1, then for any WLS solution* x *of the WLS problem Eq. (1) of the form*

$$x = A^{\dagger}_{W}b + (I - A^{\dagger}A)z, \tag{39}$$

there exists a solution $\widehat{x}$ *of the perturbed WLS problem in Eq. (35), such that with* $\delta x = \widehat{x} - x$,

$$\begin{aligned} \|\delta x\| &\leq \frac{\|A^{\dagger}_{W}\|}{1 - E\|A^{\dagger}_{W}\|}(\|\delta b\| + \|\delta A\|\|x_{WLS}\| + E\|A^{\dagger}_{W}\|\|r(x_{WLS})\|) \\ &+\|\delta A\|\|A^{\dagger}\|\|x - x_{WLS}\|. \end{aligned} \tag{40}$$

and vice versa.

Proof. For any WLS solution x of Eq. (1) of the form in Eq. (39), let $\widehat{x}$ be of the form

$$\widehat{x} = \widehat{A}^{\dagger}_{\widehat{W}}\widehat{b} + (I - \widehat{A}^{\dagger}\widehat{A})(x_{WLS} + (I - A^{\dagger}A)z). \tag{41}$$

Then $\widehat{x}$ is a WLS solution of Eq. (35). By applying Eq. (38) we have that

$$\begin{aligned} \delta x &= \widehat{A}^{\dagger}_{\widehat{W}}(\delta b - \delta A x_{WLSE}) + \widehat{A}^{\dagger}_{\widehat{W}}(I - A_W A^{\dagger}_{W})r(x_{WLS}) \\ &-\widehat{A}^{\dagger}\widehat{A}(I - A^{\dagger}A)z. \end{aligned} \tag{42}$$

From this identity we obtain the desired estimate in Eq. (40). By interchanging the roles of x and $\widehat{x}$ the reverse is also true.

5. Numerical Examples

We now provide a numerical example to verify our analysis of this paper. For stiffly weighted LS problems, Powell and Reid [5] proposed the column pivoting and row interchanging Householder QRD method to make row-wise roundoff errors small; Björck [2] conjectured, and Cox and Higham [3] proved that the column pivoting and row sorting Householder QRD method can also make row-wise roundoff errors small.

However, from our analysis in the previous sections, small row-wise roundoff errors do not imply the computational solutions accurate. In [15], we propose a row block column pivoting and row interchanging/sorting Householder QRD algorithm, which can correctly determine the numerical ranks of C_j, and make row-wise roundoff errors small as well. We now provide a numerical example. We perform the numerical computations using MATLAB software so the machine precision is $u \sim 10^{-16}$. We denote

- M1: Column pivoting and row sorting Householder QRD [5].
- M2: Column pivoting and row interchanging Householder QRD [2, 3].
- M3: Row block column pivoting and row interchanging Householder QRD [15].

Let

$$A = \begin{pmatrix} -4 & 2 & -3 \\ 4 & 2 & 2 \\ 2 & 1 & 1 \\ 1 & -1 & 1 \end{pmatrix}, \quad b = \begin{pmatrix} -9 \\ 4 \\ 1 \\ 4 \end{pmatrix},$$

$$D = \operatorname{diag}(d_1, d_1, d_2, d_3) = W^{\frac{1}{2}},$$

so

$$x_{WLS} = \begin{pmatrix} -3 \\ 0 \\ 7 \end{pmatrix} + \frac{1}{4+d_3^2} \begin{pmatrix} -4 \\ 4 \\ 8 \end{pmatrix}$$

$$\operatorname{rank}(A(1:2,:)) = \operatorname{rank}(A(1:3,:)) = 2, \quad \operatorname{rank}(A) = 3.$$

Choice 5.1. $d_1 = d_2 = 1 \geq d_3$. In this case, we set $A = C_1 = A_1$ if $d_2/d_3 < 10^2$; otherwise $C_1 = A(1:3,:)$, $A = C_2$. We list the computational results in Table 5.1. Notice that the matrix A satisfies the condition in Theorem 3.3, with $\epsilon_{li} = d_3^2$ and $\xi \sim u$. From Table 5.1 it is obvious that M1 and M2 are numerically unstable, M3 is numerically stable, and the numerical results are consistent with our analysis.

Table 5.1 $\|\delta x\|$ for Choice 5.1

d_3	1	e-2	e-4	e-6	e-8	e-12
M1	1.16e-14	2.29e-11	7.50e-8	2.22e-3	5.69e-1	5.69e-1
M2	9.57e-15	2.29e-11	7.50e-8	2.22e-3	5.69e-1	5.69e-1
M3	4.45e-15	6.75e-15	1.84e-15	1.11e-15	2.01e-15	4.44e-16

Choice 5.2. $d_1 > d_2 = 1 \geq d_3$. In this case, we set $A_1 = A(1:2,:)$ $A_2 = A(3:4,:)$ if $d_2/d_3 < 10^2$; otherwise we set $A_1 = A(1:2,:)$ $A_2 = A(3,:)$, $A_3 = A(4,:)$.

We list the computational results in Table 5.2. Notice that the matrix A satisfies the condition in Theorem 3.3, with $\epsilon_{li} = \left(\frac{d_3}{d_2}\right)^2$ and $\xi \sim u$. From Table 5.2 it is obvious that M1 and M2 are numerically unstable, M3 is numerically stable, and the numerical results are consistent with our analysis.

$d_1 = 1 > d_2 \geq d_3$.

Table 5.2 $\|\delta x\|$ for Choice 5.2

d_2	e-2	e-4	e-4	e-8	e-8	e-4
d_3	e-4	e-4	e-8	e-8	e-12	e-12
M1	2.81e-11	8.95e-16	1.04e-6	1.13e-14	3.19e-7	1.05e+2
M2	2.81e-11	8.95e-16	1.04e-6	1.13e-14	3.19e-7	1.05e+2
M3	1.16e-14	1.59e-14	1.77e-14	1.53e-14	1.33e-14	2.00e-14

6. Conclusion

In this paper we have analyzed the stability conditions for the perturbations of the stiffly weighted pseudoinverse and stiffly WLS problem. We have shown that, the perturbations of the stiff weighted pseudoinverse and stiffly WLS problem are stable, if and only if the perturbations satisfy Assumption 3.1. Numerical experiments also confirm our analysis.

References

[1] J.L. Barlow and S.L. Handy, The direct solution of weighted and equality constrained least squares problems, *SIAM J. Sci. Statist. Comput.*, **9** (1988), 704-716.

[2] Å. Björck, Numerical Methods For Least Squares problems, *SIAM*, Philadelphia, 1996.

[3] A.J. Cox and N.J. Higham, Stability of Householder QR factorization for weighted least squares problems, Numerical Analysis, 1997, Proceedings of the 17th Dundee Conference, D. F. Griffiths, D. J. Higham, and G. A. Watson Edit., Addison-Wesley Longman, Ltd, Harlow, Essex, UK, 1998, 57-73.

[4] G.H. Golub and C.F. Van Loan, Matrix Computations, 3rd Edit., The Johns Hopkins University Press, Baltimore, MD, 1996.

[5] M. Powell and J. Reid, On applying Householder transformations to linear least squares problems, Proc.of the IFIP Congress (1968), 122-126.

[6] G.W. Stewart, On the continuity of the generalized inverse, *SIAM J. Appl. Math.*, **17** (1969), 33-45.

[7] G.W. Stewart, On the weighting method for least squares problems with linear equality constraints, *BIT*, **37** (1997), 961-967.

[8] G. Strang, A framework for equilibrium equations, *SIAM Rev.*, **30** (1988), 283-297.

[9] C.F. Van Loan, On the method of weighting for equality- constrained least squares problems, *SIAM J. Numer. Anal.*, **22** (1985), 851-864.

[10] P.-Å. Wedin, Perturbation theory for pseudoinverses, *BIT*, **13** (1973), 217-232.

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

[12] M. Wei, Equivalent formulae for the supremum and stability of weighted pseudoinverses, *Math. Comp.*, **66**:220 (1997), 1487-1508.

[13] M. Wei, Supremum and Stability of Weighted Pseudoinverses and Weighted Least Squares Problems: Analysis and Computations, Nova Science, New York, 2001.

[14] M. Wei, Relationship between the stiffly weighted pseudoinverse and multi-level constrained pseudoinverse, *J. Comput. Math.*, **3** (2004), 427-436.

[15] M. Wei, A numerically stable row block Householder QRD for solving stiffly weighted least squares problems, preprint.

[16] M. Wei and A.R. De Pierro, Upper perturbation bounds of weighted projections, weighted and constrained least squares problems, *SIAM J. Matrix Anal. Appl.*, **21**:3 (2000), 931-951.

[17] M. Wei and A.R. De Pierro, Perturbation analysis of the canonical subspaces , *Linear Algebra Appl.*, **279**:1-3, (1998), 135-151.