Journal of Computational and Applied Mathematics 226 (2009) 155–165

Contents lists available at ScienceDirect

Journal of Computational and Applied Mathematics

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

On perturbations of some constrained subspaces ☆

Musheng Wei [a,*], Jinbing Zhang [b]

[a] Department of Mathematics, Shanghai Normal University, Scientific Computing Key Laboratory of Shanghai Universities, Shanghai 200234, China

[b] Department of Mathematics, Yancheng Teachers College, Jiangsu 224002, China

ARTICLE INFO

Article history:
Received 9 March 2007
Received in revised form 14 January 2008

MSC:
65F20
65F35

Keywords:
Constrained subspace
Perturbation
Singular value

ABSTRACT

Perturbation bounds of subspaces, such as eigen-spaces, singular subspaces, and canonical subspaces, have been extensively studied in the literature. In this paper, we study perturbations of some constrained subspaces of 1×2, 2×1, and 2×2 block matrices, in which only one of the sub-matrices can be changed. Such problems rise from the least squares–total least squares problem, the constrained least squares problem, and the constrained total least squares problem.

1. Introduction

In this paper we use the following conventional notation.

$\mathbb{C}^{m\times n}$ denotes the set of all $m \times n$ matrices, and $\mathbb{C}_r^{m\times n}$ is the subset of $\mathbb{C}^{m\times n}$, such that any matrix $A \in \mathbb{C}_r^{m\times n}$ has rank r. $\mathbb{U}_n$ is the set of $n \times n$ unitary matrices.

For any matrix $A \in \mathbb{C}^{m\times n}$, $\mathrm{rank}(A)$, $R(A)$, A^H and $A^\dagger$ are the rank, the range, the conjugate transpose and the Moore–Penrose pseudo-inverse of A, respectively. $\|\cdot\|$ is the Euclidian vector norm, or the corresponding subordinate matrix norm, $\|\cdot\|_F$ is the Frobenius matrix norm.

For a matrix $A \in \mathbb{C}^{m\times n}$, denote

$$P_A^\perp = I - AA^\dagger, \qquad P_{A^H}^\perp = I - A^\dagger A.$$

In the literature, there have been many articles studying perturbations of subspaces, such as eigen-spaces, singular subspaces, and canonical subspaces, see, e.g., in [5,1,10,11,18–21,13–15,6,7,12,23,26].

In some generalized least squares problems, such as the least squares–total least squares problem (LS–TLS) [4,16,17,9,24], the equality constrained least squares problem (LSE) [3,22], and the constrained total least squares problem (CTLS) [2,23], the situation becomes more complicated.

The idea of the LS–TLS problem [4,16,17,9,24] is as follows. For a given 1×2 block matrix $G_1 = (A, B)$ and an integer r with $\mathrm{rank}(A) \le r \le \mathrm{rank}(G_1)$, find a matrix $\overline{B}$ replacing B in G_1, such that $\mathrm{rank}(A, \overline{B}) = r$ and

$$\|(A, B) - (A, \overline{B})\|_F = \min_{\mathrm{rank}(A,\widetilde{B})=r} \|(A, B) - (A, \widetilde{B})\|_F,$$

☆ This work was supported by the National Natural Science Foundation of China under grant 10771073.

* Corresponding author.
E-mail addresses: mwei@shnu.edu.cn, mwei08@gmail.com (M. Wei).

0377-0427/$ – see front matter © 2009 Published by Elsevier B.V.
doi:10.1016/j.cam.2008.05.020

which is equivalent to that, for the integer $p_2 = r - \text{rank}(A)$, find a matrix $\overline{B}$ replacing B, such that $\text{rank}(P_A^{\perp}\overline{B}) = p_2$ and

$$\|P_A^{\perp}B - P_A^{\perp}\overline{B}\|_F = \min_{\text{rank}(P_A^{\perp}\widetilde{B})=p_2} \|P_A^{\perp}B - P_A^{\perp}\widetilde{B}\|_F. \tag{1.1}$$

Therefore, to obtain the matrix $\overline{B}$, we need to use the singular value decomposition (SVD) of the matrix $P_A^{\perp}B$, only retain p_2 largest singular values and set all other small singular values zero to obtain $P_A^{\perp}\overline{B}$. To analyze the perturbation bounds of the LS–TLS problem, we need to study the perturbation bounds of the matrix $P_A^{\perp}B$, and the column and row subspaces of $P_A^{\perp}\overline{B}$ which we call the constrained subspaces related to the matrix G_1. For detailed derivation of the LS–TLS problem, we refer to [4,16,9].

Similarly, the idea of the rank deficient LSE problem [22] is as follows. For a given 2×1 block matrix $G_2 = \begin{pmatrix} A \\ C \end{pmatrix}$ and an integer p_3, find a matrix $\overline{C}$ replacing C in G_2, such that $\text{rank}(\overline{C}P_{A^H}^{\perp}) = p_3$ and

$$\|CP_{A^H}^{\perp} - \overline{C}P_{A^H}^{\perp}\|_F = \min_{\text{rank}(\widetilde{C}P_{A^H}^{\perp})=p_3} \|CP_{A^H}^{\perp} - \widetilde{C}P_{A^H}^{\perp}\|_F. \tag{1.2}$$

Therefore, to obtain the matrix $\overline{C}$, we need to use the SVD of $CP_{A^H}^{\perp}$, only retain p_3 largest singular values and set all other small singular values zero to obtain $\overline{C}P_{A^H}^{\perp}$. To analyze the perturbation bounds of the rank deficient LSE problem, we need to study the perturbation bounds of the matrix $CP_{A^H}^{\perp}$, and the column and row subspaces of $\overline{C}P_{A^H}^{\perp}$ which we call the constrained subspaces related to the matrix G_2. For detailed derivation of the rank deficient LSE problem, we refer to [22].

The idea of the CTLS problem [2,23] is as follows. For a given 2×2 block matrix $G_3 = \begin{pmatrix} A & B \\ C & D \end{pmatrix}$ and an integer p, find a matrix $\overline{D}$ replacing D in G_3, such that $\text{rank}(P_N^{\perp}\overline{S}_A P_{M^H}^{\perp}) = p$ and

$$\|P_N^{\perp}S_A P_{M^H}^{\perp} - P_N^{\perp}\overline{S}_A P_{M^H}^{\perp}\|_F = \min_{\text{rank}(P_N^{\perp}\widetilde{S}_A P_{M^H}^{\perp})=p} \|P_N^{\perp}S_A P_{M^H}^{\perp} - P_N^{\perp}\widetilde{S}_A P_{M^H}^{\perp}\|_F, \tag{1.3}$$

where

$$\begin{aligned} &M = P_A^{\perp}B, \qquad N = CP_{A^H}^{\perp}, \qquad P_N^{\perp} = I - NN^{\dagger}, \qquad P_{M^H}^{\perp} = I - M^{\dagger}M, \\ &S_A = D - CA^{\dagger}B, \qquad \overline{S}_A = \overline{D} - CA^{\dagger}B, \qquad \widetilde{S}_A = \widetilde{D} - CA^{\dagger}B. \end{aligned}$$

Therefore, to obtain the matrix $\overline{D}$, we need to use the SVD of $P_N^{\perp}S_A P_{M^H}^{\perp}$, only retain p largest singular values and set all other small singular values zero to obtain $P_N^{\perp}\overline{S}_A P_{M^H}^{\perp}$. To analyze the perturbation bounds of the CTLS problem, we need to study the perturbation bounds of the matrix $P_N^{\perp}S_A P_{M^H}^{\perp}$, and the column and row subspaces of $P_N^{\perp}\overline{S}_A P_{M^H}^{\perp}$ which we call the constrained subspaces related to the matrix G_3. For detailed derivation of the CTLS problem, we refer to [2,23].

Perturbation estimates for the above generalized LS problems have been discussed, but to our knowledge, in the literature there has been no article discussing perturbations for the constrained subspaces of the above mentioned problems.

In this paper we will study the perturbations for the constrained subspaces. The paper is arranged as follows. In Section 2, we provide some preliminary results which are needed for our analysis; Section 3, we derive the perturbation bounds of constrained subspaces for 1×2 and 2×1 block matrices; in Section 4, we derive the perturbation bounds of constrained subspaces for an 2×2 block matrix; finally in Section 5, we conclude the paper with some remarks.

2. Preliminaries

In this section we mention the following results which are needed for our further discussion.

Lemma 2.1 (*CSD [8]*). *Suppose $W \in \mathbb{U}_n$. Partition W as*

$$W = \begin{array}{cc} \begin{pmatrix} W_{11} & W_{12} \\ W_{21} & W_{22} \end{pmatrix} & \begin{array}{c} r_1 \\ r_2 \end{array} \\ \begin{array}{cc} c_1 & c_2 \end{array} & \end{array} \tag{2.1}$$

with $r_1 + r_2 = c_1 + c_2 = n$. Then

$$W = \begin{pmatrix} U_1 & 0 \\ 0 & U_2 \end{pmatrix} \begin{pmatrix} D_{11} & D_{12} \\ D_{21} & D_{22} \end{pmatrix} \begin{pmatrix} V_1^H & 0 \\ 0 & V_2^H \end{pmatrix}, \tag{2.2}$$

$U_1 \in \mathbb{U}_{r_1}, U_2 \in \mathbb{U}_{r_2}, V_1 \in \mathbb{U}_{c_1}, V_2 \in \mathbb{U}_{c_2}$,

$$\left(\begin{array}{c|c} D_{11} & D_{12} \\ \hline D_{21} & D_{22} \end{array}\right) = \left(\begin{array}{ccc|ccc} I & & & 0_S^H & & \\ & C & & & S & \\ & & 0_C & & & I \\ \hline 0_S & & & I & & \\ & S & & & -C & \\ & & I & & & 0_C^H \end{array}\right), \tag{2.3}$$

where 0_C *and* 0_S *are zero matrices with appropriate sizes,*

$$C = \mathrm{diag}(c_1, c_2, \ldots, c_l), \quad 1 > c_1 \geqslant c_2 \geqslant \cdots \geqslant c_l > 0,$$
$$S = \mathrm{diag}(s_1, s_2, \ldots, s_l), \quad 0 < s_1 \leqslant s_2 \leqslant \cdots \leqslant s_l < 1,$$

and C, S *satisfy*

$$C^2 + S^2 = I_l.$$

By applying the CSD in Lemma 2.1, one can easily have the following assertions [11,20,25].

Lemma 2.2. *Suppose that* $A, \widehat{A} \in \mathbb{C}^{m\times n}$.

(1) *If* $\mathrm{rank}(A) = \mathrm{rank}(\widehat{A})$, *then* $P_A P_{\widehat{A}}^{\perp}$ *and* $P_A^{\perp} P_{\widehat{A}}$ *have the same singular values, thus*

$$\|P_A P_{\widehat{A}}^{\perp}\| = \|P_A^{\perp} P_{\widehat{A}}\|. \tag{2.4}$$

(2) *If* $\|P_A - P_{\widehat{A}}\| < 1$, *then* $\mathrm{rank}(A) = \mathrm{rank}(\widehat{A})$, *and*

$$\|P_A - P_{\widehat{A}}\| = \|P_A P_{\widehat{A}}^{\perp}\| = \|P_A^{\perp} P_{\widehat{A}}\|. \tag{2.5}$$

(3) *If* $\mathrm{rank}(A) > \mathrm{rank}(\widehat{A})$, *then*

$$\|P_A P_{\widehat{A}}^{\perp}\| \geqslant \|P_A^{\perp} P_{\widehat{A}}\|. \tag{2.6}$$

3. Perturbation analysis for the constrained subspaces of 1 × 2 and 2 × 1 block matrices

Consider the perturbation between constrained subspaces of the 1×2 block matrices

$$G_1 = (A, B), \qquad \widehat{G}_1 = (\widehat{A}, \widehat{B}), \tag{3.1}$$

where $A, \widehat{A} = A + \Delta A \in \mathbb{C}_{p_1}^{m\times n_1}$, $B, \widehat{B} = B + \Delta B \in \mathbb{C}^{m\times n_2}$. We first take the SVD of A, then take the SVD of $M = P_A^{\perp} B$. For a chosen integer $p_2 \le \mathrm{rank}(M)$ we obtain the decomposition of G_1 as

$$G_1 = U_1 \left(\begin{array}{cc|cc} A_{11} & 0 & B_{11} & B_{12} \\ 0 & 0 & B_{21} & 0 \\ 0 & 0 & 0 & B_{32} \end{array}\right) \begin{pmatrix} V_1^H & 0 \\ 0 & V_2^H \end{pmatrix}, \tag{3.2}$$

where $U_1 \in \mathbb{U}_m, V_1 \in \mathbb{U}_{n_1}, V_2 \in \mathbb{U}_{n_2}, A_{11} = \mathrm{diag}(\sigma_1(A), \sigma_2(A), \ldots, \sigma_{p_1}(A))$, with $\sigma_1(A) \geqslant \sigma_2(A) \geqslant \cdots \geqslant \sigma_{p_1}(A)$ the nonzero singular values of A, $B_{21} = \mathrm{diag}(\sigma_1(M), \ldots, \sigma_{p_2}(M))$, $B_{32} = \mathrm{diag}(\sigma_{p_2+1}(M), \ldots, \sigma_l(M))$, with $\sigma_1(M) \geqslant \cdots \geqslant \sigma_{p_2}(M) > \sigma_{p_2+1}(M) \geqslant \cdots \geqslant \sigma_l(M)$ ($l = \min\{m - p_1, n_2\}$) the singular values of M.

Similarly, we have

$$\widehat{G}_1 = \widehat{U}_1 \left(\begin{array}{cc|cc} \widehat{A}_{11} & 0 & \widehat{B}_{11} & \widehat{B}_{12} \\ 0 & 0 & \widehat{B}_{21} & 0 \\ 0 & 0 & 0 & \widehat{B}_{32} \end{array}\right) \begin{pmatrix} \widehat{V}_1^H & 0 \\ 0 & \widehat{V}_2^H \end{pmatrix}, \tag{3.3}$$

where $\widehat{U}_1, \widehat{V}_1, \widehat{V}_2$ are unitary matrices, $\widehat{A}_{11} = \mathrm{diag}(\sigma_1(\widehat{A}), \sigma_2(\widehat{A}), \ldots, \sigma_{p_1}(\widehat{A}))$, with $\sigma_1(\widehat{A}) \geqslant \sigma_2(\widehat{A}) \geqslant \cdots \geqslant \sigma_{p_1}(\widehat{A})$ the nonzero singular values of $\widehat{A}$, $\widehat{B}_{21} = \mathrm{diag}(\sigma_1(\widehat{M}), \ldots, \sigma_{p_2}(\widehat{M}))$, $\widehat{B}_{32} = \mathrm{diag}(\sigma_{p_2+1}(\widehat{M}), \ldots, \sigma_l(\widehat{M}))$, with $\sigma_1(\widehat{M}) \geqslant \cdots \geqslant \sigma_l(\widehat{M})$ the singular values of $\widehat{M}$.

For $i = 1, 2$, partition U_1, $\widehat{U}_1$ and V_i $\widehat{V}_i$ as follows,

$$\begin{aligned} U_1 &= (U_{11}, U_{12}, U_{13}), & V_1 &= (V_{11}, V_{12}), & V_2 &= (V_{21}, V_{22}), \\ \widehat{U}_1 &= (\widehat{U}_{11}, \widehat{U}_{12}, \widehat{U}_{13}), & \widehat{V}_1 &= (\widehat{V}_{11}, \widehat{V}_{12}), & \widehat{V}_2 &= (\widehat{V}_{21}, \widehat{V}_{22}). \\ & \quad p_1, p_2, \overline{m} & & \quad p_1, \overline{n}_1, & & \quad p_2, \overline{n}_2 \end{aligned} \tag{3.4}$$

where $\overline{m} = m - p_1 - p_2$, $\overline{n}_1 = n_1 - p_1$, and $\overline{n}_2 = n_2 - p_2$. Then

$$A = U_{11}A_{11}V_{11}^H, \qquad M = (U_{12}, U_{13})\mathrm{diag}(B_{21}, B_{32})V_2^H,$$
$$\widehat{M} = (\widehat{U}_{12}, \widehat{U}_{13})\mathrm{diag}(\widehat{B}_{21}, \widehat{B}_{32})\widehat{V}_2^H,$$

and

$$\begin{aligned} &\overline{B} = AA^{\dagger}B + U_{12}B_{21}V_{21}^H, \qquad \overline{\widehat{B}} = \widehat{A}\widehat{A}^{\dagger}\widehat{B} + \widehat{U}_{12}\widehat{B}_{21}\widehat{V}_{21}^H, \\ &P_A^{\perp}\overline{B} = U_{12}B_{21}V_{21}^H, \qquad P_{\widehat{A}}^{\perp}\overline{\widehat{B}} = \widehat{U}_{12}\widehat{B}_{21}\widehat{V}_{21}^H. \end{aligned} \tag{3.5}$$

In order to derive the perturbation bounds for the constrained subspaces, $\|U_{12}U_{12}^H - \widehat{U}_{12}\widehat{U}_{12}^H\|$ and $\|V_{21}V_{21}^H - \widehat{V}_{21}\widehat{V}_{21}^H\|$, we first need to derive the perturbation bounds of the subspaces for the matrix A, as mentioned below.

Lemma 3.1. *Suppose that the matrices $G_1 = (A, B)$, $\widehat{G}_1 = (\widehat{A}, \widehat{B}) \in \mathbb{C}^{m\times(n_1+n_2)}$ are defined in (3.1) with $\mathrm{rank}(A) = \mathrm{rank}(\widehat{A}) = p_1$, the decompositions of G_1, $\widehat{G}_1$ are in (3.2)–(3.3). Then*

$$\begin{aligned} &\mathrm{dist}(R(A), R(\widehat{A})) \leqslant \min\left\{\frac{\|(\widehat{U}_{12}, \widehat{U}_{13})^H \Delta A V_{11}\|}{\sigma_{p_1}(A)}, \frac{\|(U_{12}, U_{13})^H \Delta A\widehat{V}_{11}\|}{\sigma_{p_1}(\widehat{A})}\right\}, \\ &\mathrm{dist}(R(A^H), R(\widehat{A}^H)) \leqslant \min\left\{\frac{\|U_{11}^H\Delta A\widehat{V}_{12}\|}{\sigma_{p_1}(A)}, \frac{\|\widehat{U}_{11}^H\Delta A V_{12}\|}{\sigma_{p_1}(\widehat{A})}\right\}. \end{aligned} \tag{3.6}$$

Proof. From (3.2)–(3.3), the SVD of A and $\widehat{A}$ have the following forms,

$$A = U_1\mathrm{diag}(A_{11}, 0, 0)V_1^H = U_{11}A_{11}V_{11}^H,$$
$$\widehat{A} = \widehat{U}_1\mathrm{diag}(\widehat{A}_{11}, 0, 0)\widehat{V}_1^H = \widehat{U}_{11}\widehat{A}_{11}\widehat{V}_{11}^H,$$

so $\mathrm{dist}(R(A),\ R(\widehat{A})) = \|U_{11}U_{11}^H - \widehat{U}_{11}\widehat{U}_{11}^H\|$. Therefore,

$$\begin{aligned} \mathrm{dist}(R(A), R(\widehat{A})) &= \|U_1^H(U_{11}U_{11}^H - \widehat{U}_{11}\widehat{U}_{11}^H)\widehat{U}_1\| \\ &= \left\|\begin{pmatrix} 0 & U_{11}^H\widehat{U}_{12} & U_{11}^H\widehat{U}_{13} \\ -U_{12}^H\widehat{U}_{11} & 0 & 0 \\ -U_{13}^H\widehat{U}_{11} & 0 & 0 \end{pmatrix}\right\| \\ &= \max\left\{\|U_{11}^H(\widehat{U}_{12}, \widehat{U}_{13})\|, \|(U_{12}, U_{13})^H\widehat{U}_{11}\|\right\} \\ &= \|U_{11}^H(\widehat{U}_{12}, \widehat{U}_{13})\| = \|(U_{12}, U_{13})^H\widehat{U}_{11}\|. \end{aligned} \tag{3.7}$$

From the identities

$$\Delta A = \widehat{A} - A = \widehat{U}_{11}\widehat{A}_{11}\widehat{V}_{11}^H - U_{11}A_{11}V_{11}^H,$$
$$(\widehat{U}_{12}, \widehat{U}_{13})^H\Delta A V_{11} = -(\widehat{U}_{12}, \widehat{U}_{13})^H U_{11}A_{11},$$
$$(U_{12}, U_{13})^H\Delta A\widehat{V}_{11} = (U_{12}, U_{13})^H\widehat{U}_{11}\widehat{A}_{11},$$

we observe that

$$\begin{aligned} &(\widehat{U}_{12}, \widehat{U}_{13})^H U_{11} = -(\widehat{U}_{12}, \widehat{U}_{13})^H\Delta A V_{11}A_{11}^{-1}, \\ &(U_{12}, U_{13})^H\widehat{U}_{11} = (U_{12}, U_{13})^H\Delta A\widehat{V}_{11}\widehat{A}_{11}^{-1}. \end{aligned} \tag{3.8}$$

From the above equalities, we obtain the first estimate in (3.6). The second inequality in (3.6) can be derived similarly. □

We now derive perturbation bounds for the constrained subspaces of the matrix $P_A^{\perp}\overline{B}$ defined in (3.5).

Theorem 3.2. *Under the conditions of Lemma 3.1, if furthermore, for a chosen integer $p_2 \le \mathrm{rank}(M)$, $\sigma_{p_2}(M) - \sigma_{p_2+1}(M) > 2\|\Delta M\|$ ($\Delta M = \widehat{M} - M$), then for the matrices $\overline{B}$ and $\overline{\widehat{B}}$ defined in (3.5), $\mathrm{rank}(P_A^{\perp}\overline{B}) = \mathrm{rank}(P_{\widehat{A}}^{\perp}\overline{\widehat{B}}) = p_2$, and we have the following estimates,*

$$\begin{aligned} &\|U_{12}U_{12}^H - \widehat{U}_{12}\widehat{U}_{12}^H\| \leqslant \min\left\{\frac{\|\widehat{U}_{12}^H\Delta A V_{11}\|}{\sigma_{p_1}(A)} + \frac{\eta_1}{\sigma_{p_2}(\widehat{M}) - \sigma_{p_2+1}(M)}, \frac{\|U_{12}^H\Delta A\widehat{V}_{11}\|}{\sigma_{p_1}(\widehat{A})} + \frac{\eta_2}{\sigma_{p_2}(M) - \sigma_{p_2+1}(\widehat{M})}\right\}, \\ &\|V_{21}V_{21}^H - \widehat{V}_{21}\widehat{V}_{21}^H\| \leqslant \min\left\{\frac{\eta_1}{\sigma_{p_2}(\widehat{M}) - \sigma_{p_2+1}(M)}, \frac{\eta_2}{\sigma_{p_2}(M) - \sigma_{p_2+1}(\widehat{M})}\right\}, \end{aligned} \tag{3.9}$$

where

$$\begin{aligned}\eta_1 &= \max\left\{\left(\|U_{13}^H \Delta B \widehat{V}_{21}\| + \frac{\|\widehat{B}_{11}\|\|U_{13}^H \Delta A \widehat{V}_{11}\|}{\sigma_{p_1}(\widehat{A})}\right), \left(\|\widehat{U}_{12}^H \Delta B V_{22}\| + \frac{\|B_{12}\|\|\widehat{U}_{12}^H \Delta A V_{11}\|}{\sigma_{p_1}(A)}\right)\right\},\\ \eta_2 &= \max\left\{\left(\|\widehat{U}_{13}^H \Delta B V_{21}\| + \frac{\|B_{11}\|\|\widehat{U}_{13}^H \Delta A V_{11}\|}{\sigma_{p_1}(A)}\right), \left(\|U_{12}^H \Delta B \widehat{V}_{22}\| + \frac{\|\widehat{B}_{12}\|\|U_{12}^H \Delta A \widehat{V}_{11}\|}{\sigma_{p_1}(\widehat{A})}\right)\right\}.\end{aligned} \tag{3.10}$$

Proof. By the perturbation analysis of the singular values,

$$\begin{aligned}&\sigma_{p_2}(\widehat{M}) - \sigma_{p_2+1}(M) \geqslant \sigma_{p_2}(M) - \sigma_{p_2+1}(M) - \|\Delta M\| > 0,\\ &\sigma_{p_2}(M) - \sigma_{p_2+1}(\widehat{M}) \geqslant \sigma_{p_2}(M) - \sigma_{p_2+1}(M) - \|\Delta M\| > 0,\\ &\sigma_{p_2}(\widehat{M}) - \sigma_{p_2+1}(\widehat{M}) \geqslant \sigma_{p_2}(M) - \sigma_{p_2+1}(M) - 2\|\Delta M\| > 0.\end{aligned}$$

Therefore, both B_{21} and $\widehat{B}_{21}$ are nonsingular, $\operatorname{rank}(P_A^{\perp}\overline{B}) = \operatorname{rank}(P_{\widehat{A}}^{\perp}\overline{\widehat{B}}) = p_2$. Notice that

$$\|U_{12}U_{12}^H - \widehat{U}_{12}\widehat{U}_{12}^H\| = \|U_1^H(U_{12}U_{12}^H - \widehat{U}_{12}\widehat{U}_{12}^H)\widehat{U}_1\|,$$

therefore, by applying Lemma 2.2 we observe that

$$\begin{aligned}\|U_{12}U_{12}^H - \widehat{U}_{12}\widehat{U}_{12}^H\| &= \left\|\begin{pmatrix}-U_{11}^H\widehat{U}_{12} & 0 & 0\\ -U_{13}^H\widehat{U}_{12} & 0 & 0\\ 0 & U_{12}^H\widehat{U}_{11} & U_{12}^H\widehat{U}_{13}\end{pmatrix}\right\|\\ &= \|(U_{11}, U_{13})^H\widehat{U}_{12}\| = \|U_{12}^H(\widehat{U}_{11}, \widehat{U}_{13})\|.\end{aligned} \tag{3.11}$$

Also,

$$\|V_{21}V_{21}^H - \widehat{V}_{21}\widehat{V}_{21}^H\| = \|V_{21}^H\widehat{V}_{22}\| = \|V_{22}^H\widehat{V}_{21}\|. \tag{3.12}$$

From (3.8) we can derive

$$\|\widehat{U}_{1i}^H U_{11}\| \leqslant \frac{\|\widehat{U}_{1i}^H \Delta A V_{11}\|}{\sigma_{p_1}(A)}, \qquad \|U_{1i}^H\widehat{U}_{11}\| \leqslant \frac{\|U_{1i}^H \Delta A \widehat{V}_{11}\|}{\sigma_{p_1}(\widehat{A})}, \quad i = 2, 3. \tag{3.13}$$

From the identity $\Delta B = \widehat{B} - B$ and (3.2)–(3.5),

$$\begin{aligned}B &= U_{11}B_{11}V_{21}^H + U_{12}B_{21}V_{21}^H + U_{11}B_{12}V_{22}^H + U_{13}B_{32}V_{22}^H,\\ \widehat{B} &= \widehat{U}_{11}\widehat{B}_{11}\widehat{V}_{21}^H + \widehat{U}_{12}\widehat{B}_{21}\widehat{V}_{21}^H + \widehat{U}_{11}\widehat{B}_{12}\widehat{V}_{22}^H + \widehat{U}_{13}\widehat{B}_{32}\widehat{V}_{22}^H,\end{aligned}$$

we observe that

$$\begin{aligned}U_{13}^H \Delta B \widehat{V}_{21} &= U_{13}^H\widehat{U}_{11}\widehat{B}_{11} + U_{13}^H\widehat{U}_{12}\widehat{B}_{21} - B_{32}V_{22}^H\widehat{V}_{21},\\ \widehat{U}_{12}^H \Delta B V_{22} &= \widehat{B}_{21}\widehat{V}_{21}^H V_{22} - \widehat{U}_{12}^H U_{11}B_{12} - \widehat{U}_{12}^H U_{13}B_{32}.\end{aligned}$$

Combining the above equalities and (3.8), we obtain

$$\begin{aligned}U_{13}^H\widehat{U}_{12}\widehat{B}_{21} &= U_{13}^H \Delta B \widehat{V}_{21} - U_{13}^H \Delta A \widehat{V}_{11}\widehat{A}_{11}^{-1}\widehat{B}_{11} + B_{32}V_{22}^H\widehat{V}_{21},\\ \widehat{B}_{21}\widehat{V}_{21}^H V_{22} &= \widehat{U}_{12}^H \Delta B V_{22} - \widehat{U}_{12}^H \Delta A V_{11}A_{11}^{-1}B_{12} + \widehat{U}_{12}^H U_{13}B_{32}.\end{aligned}$$

Therefore, we observe that

$$\begin{aligned}&\sigma_{p_2}(\widehat{M})\|U_{13}^H\widehat{U}_{12}\| \le \eta_1 + \sigma_{p_2+1}(M)\|V_{22}^H\widehat{V}_{21}\|,\\ &\sigma_{p_2}(\widehat{M})\|\widehat{V}_{21}^H V_{22}\| \le \eta_1 + \sigma_{p_2+1}(M)\|\widehat{U}_{12}^H U_{13}\|.\end{aligned} \tag{3.14}$$

So

$$\begin{aligned}\|U_{13}^H\widehat{U}_{12}\| &\le \frac{1}{\sigma_{p_2}(\widehat{M})}\left(\eta_1 + \frac{\sigma_{p_2+1}(M)}{\sigma_{p_2}(\widehat{M})}\left(\eta_1 + \sigma_{p_2+1}(M)\|U_{13}^H\widehat{U}_{12}\|\right)\right),\\ \|\widehat{V}_{21}^H V_{22}\| &\le \frac{1}{\sigma_{p_2}(\widehat{M})}\left(\eta_1 + \frac{\sigma_{p_2+1}(M)}{\sigma_{p_2}(\widehat{M})}\left(\eta_1 + \sigma_{p_2+1}(M)\|\widehat{V}_{21}^H V_{22}\|\right)\right),\end{aligned}$$

and from above inequalities we have that

$$\begin{aligned} \|U_{13}^H \widehat{U}_{12}\| &\le \frac{\eta_1}{\sigma_{p_2}(\widehat{M}) - \sigma_{p_2+1}(M)}, \\ \|\widehat{V}_{21}^H V_{22}\| &\le \frac{\eta_1}{\sigma_{p_2}(\widehat{M}) - \sigma_{p_2+1}(M)}. \end{aligned} \tag{3.15}$$

In a similar manner, we have

$$\begin{aligned} \|\widehat{U}_{13}^H U_{12}\| &\le \frac{\eta_2}{\sigma_{p_2}(M) - \sigma_{p_2+1}(\widehat{M})}, \\ \|V_{21}^H \widehat{V}_{22}\| &\le \frac{\eta_2}{\sigma_{p_2}(M) - \sigma_{p_2+1}(\widehat{M})}. \end{aligned} \tag{3.16}$$

From the inequalities derived in (3.11)–(3.13) and (3.15)–(3.16), we obtain the desired estimates of the theorem. □

Remark 3.1. From (3.9)–(3.10) we observe that, the perturbation bounds for the constrained subspaces of the matrix $P_A^{\perp}\overline{B}$ are more complicated than those for the matrix A. Also, these bounds are realistic in the sense that, one can find a example that the true perturbations are close to these perturbation bounds.

The following corollary is the direct conclusion of Theorem 3.2.

Corollary 3.3. *Under the notation and the conditions of Theorem 3.2, if furthermore,* $\mathrm{rank}(\widehat{M}) = \mathrm{rank}(M) = p_2$*, then*

$$\begin{aligned} &\mathrm{dist}(R(M), R(\widehat{M})) \leqslant \min\left\{ \frac{\|\widehat{U}_{12}^H \Delta A V_{11}\|}{\sigma_{p_1}(A)} + \frac{\eta_1}{\sigma_{p_2}(\widehat{M})}, \frac{\|U_{12}^H \Delta A \widehat{V}_{11}\|}{\sigma_{p_1}(\widehat{A})} + \frac{\eta_2}{\sigma_{p_2}(M)} \right\}, \\ &\mathrm{dist}(R(M^H), R(\widehat{M}^H)) \leqslant \min\left\{ \frac{\eta_1}{\sigma_{p_2}(\widehat{M})}, \frac{\eta_2}{\sigma_{p_2}(M)} \right\}. \end{aligned} \tag{3.17}$$

Proof. We only need to set $\sigma_{p_2+1}(M) = \sigma_{p_2+1}(\widehat{M}) = 0$ in (3.9). □

Remark 3.2. When considering the LS–TLS problem, we suppose $\widehat{A}$ is the perturbation of A. If $\mathrm{rank}(\widehat{A}) \neq \mathrm{rank}(A)$, then small perturbation in A can cause very large errors of the LS–TLS solutions. So we need to enforce the condition $\mathrm{rank}(\widehat{A}) = \mathrm{rank}(A)$. The perturbation bounds derived in Theorem 3.2 and Corollary 3.3 can be applied to estimate the perturbations in the LS–TLS problem and the CTLS problem.

Now we consider the perturbation bounds of the constrained subspaces of

$$G_2 = \begin{pmatrix} A \\ C \end{pmatrix}, \qquad \widehat{G}_2 = \begin{pmatrix} \widehat{A} \\ \widehat{C} \end{pmatrix}, \tag{3.18}$$

where $A, \widehat{A} = A + \Delta A \in \mathbb{C}_{p_1}^{m_1 \times n}$, $C, \widehat{C} = C + \Delta C \in \mathbb{C}^{m_2 \times n}$, $N = C(I - A^{\dagger}A)$, $\widehat{N} = \widehat{C}(I - \widehat{A}^{\dagger}\widehat{A})$. Then similar to the analysis for G_1 and $\widehat{G}_1$, we have the following decompositions of G_2 and $\widehat{G}_2$,

$$G_2 = \begin{pmatrix} U_1 & 0 \\ 0 & U_2 \end{pmatrix} \left(\begin{array}{ccc} A_{11} & 0 & 0 \\ 0 & 0 & 0 \\ \hline C_{11} & C_{12} & 0 \\ C_{21} & 0 & C_{23} \end{array}\right) V_1^H, \tag{3.19}$$

where U_1, U_2, V_1 are unitary matrices, $A_{11} = \mathrm{diag}(\sigma_1(A), \sigma_2(A), \ldots, \sigma_{p_1}(A))$, with $\sigma_1(A) \geqslant \sigma_2(A) \geqslant \cdots \geqslant \sigma_{p_1}(A)$ the nonzero singular values of A, $C_{12} = \mathrm{diag}(\sigma_1(N), \ldots, \sigma_{p_3}(N))$, $C_{23} = \mathrm{diag}(\sigma_{p_3+1}(N), \ldots, \sigma_l(N))$, with $\sigma_1(N) \geqslant \cdots \geqslant \sigma_l(N)$ the singular values of N,

$$\widehat{G}_2 = \begin{pmatrix} \widehat{U}_1 & 0 \\ 0 & \widehat{U}_2 \end{pmatrix} \left(\begin{array}{ccc} \widehat{A}_{11} & 0 & 0 \\ 0 & 0 & 0 \\ \hline \widehat{C}_{11} & \widehat{C}_{12} & 0 \\ \widehat{C}_{21} & 0 & \widehat{C}_{23} \end{array}\right) \widehat{V}_1^H, \tag{3.20}$$

where $\widehat{U}_1, \widehat{U}_2, \widehat{V}_1$ are unitary matrices, $\widehat{A}_{11} = \mathrm{diag}(\sigma_1(\widehat{A}), \sigma_2(\widehat{A}), \ldots, \sigma_{p_1}(\widehat{A}))$, with $\sigma_1(\widehat{A}) \geqslant \sigma_2(\widehat{A}) \geqslant \cdots \geqslant \sigma_{p_1}(\widehat{A})$ the nonzero singular values of $\widehat{A}$, $\widehat{C}_{12} = \mathrm{diag}(\sigma_1(\widehat{N}), \ldots, \sigma_{p_3}(\widehat{N}))$, $\widehat{C}_{23} = \mathrm{diag}(\sigma_{p_3+1}(\widehat{N}), \ldots, \sigma_l(\widehat{N}))$, with $\sigma_1(\widehat{N}) \geqslant \cdots \geqslant \sigma_l(\widehat{N})$ the singular values of $\widehat{N}$. For $i = 1$, partition U_i, $\widehat{U}_i$ and V_1 $\widehat{V}_1$ as follows,

$$\begin{aligned} U_1 &= (U_{11}, U_{12}),\ U_2 = (U_{21}, U_{22}),\ V_1 = (V_{11}, V_{12}, V_{13}), \\ \widehat{U}_1 &= (\widehat{U}_{11}, \widehat{U}_{12}),\ \widehat{U}_2 = (\widehat{U}_{21}, \widehat{U}_{22}),\ \widehat{V}_1 = (\widehat{V}_{11}, \widehat{V}_{12}, \widehat{V}_{13}), \\ &\quad p_1, \overline{m}_1 \qquad\qquad p_3, \overline{m}_2 \qquad\qquad p_1, p_3, \overline{m} \end{aligned} \tag{3.21}$$

where $\overline{m}_1 = m_1 - p_1$, $\overline{m}_2 = m_2 - p_3$, and $\overline{m} = m - p_1 - p_3$.

Then

$$\begin{aligned} &\mathrm{rank}(G_2) = \mathrm{rank}(A) + \mathrm{rank}(N), \qquad \mathrm{rank}(\widehat{G}_2) = \mathrm{rank}(\widehat{A}) + \mathrm{rank}(\widehat{N}), \\ &A = U_{11}A_{11}V_{11}^H, \qquad N = U_2\mathrm{diag}(C_{12}, C_{23})(V_{12}, V_{13})^H, \\ &\widehat{N} = \widehat{U}_2\mathrm{diag}(\widehat{C}_{12}, \widehat{C}_{23})(\widehat{V}_{12}, \widehat{V}_{13})^H, \end{aligned}$$

and

$$\begin{aligned} &\overline{C} = CA^\dagger A + U_{21}C_{12}V_{12}^H, \qquad \overline{\widehat{C}} = \widehat{C}\widehat{A}^\dagger\widehat{A} + \widehat{U}_{21}\widehat{C}_{12}\widehat{V}_{12}^H, \\ &\overline{C}P_{A^H}^\perp = U_{21}C_{12}V_{12}^H, \qquad \overline{\widehat{C}}P_{\widehat{A}^H}^\perp = \widehat{U}_{21}\widehat{C}_{12}\widehat{V}_{12}^H. \end{aligned} \tag{3.22}$$

Notice that G_2^H and $\widehat{G}_2^H$ are 1×2 block matrices. Therefore, by using exactly the same procedure, we have the following results.

Theorem 3.4. *Suppose that the matrices G_2, $\widehat{G}_2$ are defined in (3.18), the decompositions of G_1, $\widehat{G}_1$ are in (3.19)–(3.21). Suppose that* $\mathrm{rank}(A) = \mathrm{rank}(\widehat{A}) = p_1$, *and a chosen integer p_3 satisfying $p_3 \leq \mathrm{rank}(N)$. If $\sigma_{p_3}(N) - \sigma_{p_3+1}(N) > 2\|\Delta N\|$ ($\Delta N = \widehat{N} - N$), then* $\mathrm{rank}(\overline{C}P_{A^H}^\perp) = \mathrm{rank}(\overline{\widehat{C}}P_{\widehat{A}^H}^\perp) = p_3$, *and we have the following estimates,*

$$\begin{aligned} &\|V_{12}V_{12}^H - \widehat{V}_{12}\widehat{V}_{12}^H\| \leqslant \min\left\{\frac{\|U_{11}^H\Delta A\widehat{V}_{12}\|}{\sigma_{p_1}(A)} + \frac{\eta_3}{\sigma_{p_3}(\widehat{N}) - \sigma_{p_3+1}(N)}, \frac{\|\widehat{U}_{11}^H\Delta AV_{12}\|}{\sigma_{p_1}(\widehat{A})} + \frac{\eta_4}{\sigma_{p_3}(N) - \sigma_{p_3+1}(\widehat{N})}\right\}, \\ &\|U_{21}U_{21}^H - \widehat{U}_{21}\widehat{U}_{21}^H\| \leqslant \min\left\{\frac{\eta_3}{\sigma_{p_3}(\widehat{N}) - \sigma_{p_3+1}(N)}, \frac{\eta_4}{\sigma_{p_3}(N) - \sigma_{p_3+1}(\widehat{N})}\right\}, \end{aligned} \tag{3.23}$$

where

$$\begin{aligned} \eta_3 &= \max\left\{\left(\|\widehat{U}_{21}^H\Delta CV_{13}\| + \frac{\|\widehat{C}_{11}\|\|\widehat{U}_{11}^H\Delta AV_{13}\|}{\sigma_{p_1}(\widehat{A})}\right), \left(\|U_{22}^H\Delta C\widehat{V}_{12}\| + \frac{\|C_{21}\|\|U_{11}^H\Delta A\widehat{V}_{12}\|}{\sigma_{p_1}(A)}\right)\right\}, \\ \eta_4 &= \max\left\{\left(\|U_{21}^H\Delta C\widehat{V}_{13}\| + \frac{\|C_{11}\|\|U_{11}^H\Delta A\widehat{V}_{13}\|}{\sigma_{p_1}(A)}\right), \left(\|\widehat{U}_{22}^H\Delta CV_{12}\| + \frac{\|\widehat{C}_{21}\|\|\widehat{U}_{11}^H\Delta AV_{12}\|}{\sigma_{p_1}(\widehat{A})}\right)\right\}. \end{aligned} \tag{3.24}$$

Furthermore, when $\mathrm{rank}(\widehat{N}) = \mathrm{rank}(N) = p_3$, *then*

$$\begin{aligned} &\mathrm{dist}(R(N), R(\widehat{N})) \leqslant \min\left\{\frac{\eta_3}{\sigma_{p_3}(\widehat{N})}, \frac{\eta_4}{\sigma_{p_3}(N)}\right\}, \\ &\mathrm{dist}(R(N^H), R(\widehat{N}^H)) \leqslant \min\left\{\frac{\|U_{11}^H\Delta A\widehat{V}_{12}\|}{\sigma_{p_1}(A)} + \frac{\eta_3}{\sigma_{p_3}(\widehat{N})}, \frac{\|\widehat{U}_{11}^H\Delta AV_{12}\|}{\sigma_{p_1}(\widehat{A})} + \frac{\eta_4}{\sigma_{p_3}(N)}\right\}. \end{aligned} \tag{3.25}$$

Remark 3.3. For the rank deficient LSE problem, we suppose that $\widehat{A}$ is the perturbation of A. If $\mathrm{rank}(\widehat{A}) \neq \mathrm{rank}(A)$, even when the perturbations are very small, the perturbation of the LSE solution will be very large [22,25]. Therefore, we need to enforce the condition $\mathrm{rank}(\widehat{A}) = \mathrm{rank}(A)$.

4. Perturbation analysis for the constrained subspaces of 2 × 2 block matrix

Demmel [2], Wei [23] studied the perturbations of the 2×2 block matrix

$$G_3 = \begin{array}{cc} \begin{pmatrix} A & B \\ C & D \end{pmatrix} & \begin{array}{c} m_1 \\ m_2 \end{array} \\ \begin{array}{cc} n_1, & n_2 \end{array} & \end{array}, \tag{4.1}$$

in which only the sub-matrix D can be changed. let

$$M = P_A^{\perp} B, \qquad N = C P_{A^H}^{\perp}, \tag{4.2}$$

and suppose that

$$p_1 = \operatorname{rank}(A), \qquad p_2 = \operatorname{rank}(M) \quad \text{and} \quad p_3 = \operatorname{rank}(N). \tag{4.3}$$

It can be shown [2,23] that there exist unitary matrices U_1, U_2, V_1 and V_2, such that

$$\begin{aligned} G_3 &= \left(\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right) \\ &= \begin{pmatrix} U_1 & 0 \\ 0 & U_2 \end{pmatrix} \left(\begin{array}{ccc|cc} A_{11} & 0 & 0 & B_{11} & B_{12} \\ 0 & 0 & 0 & B_{21} & 0 \\ 0 & 0 & 0 & 0 & 0 \\ \hline C_{11} & C_{12} & 0 & D_{11} & D_{12} \\ C_{21} & 0 & 0 & D_{21} & D_{22} \end{array}\right) \begin{pmatrix} V_1^H & 0 \\ 0 & V_2^H \end{pmatrix}, \end{aligned} \tag{4.4}$$

where

$$\begin{aligned} A_{11} &= \operatorname{diag}(\sigma_1(A), \sigma_2(A), \ldots, \sigma_{p_1}(A)), \\ B_{21} &= \operatorname{diag}(\sigma_1(M), \sigma_2(M), \ldots, \sigma_{p_2}(M)), \\ C_{12} &= \operatorname{diag}(\sigma_1(N), \sigma_2(N), \ldots, \sigma_{p_3}(N)), \end{aligned}$$

and

$$\begin{aligned} &\sigma_1(A) \geqslant \sigma_2(A) \geqslant \cdots \geqslant \sigma_{p_1}(A) > 0, \\ &\sigma_1(M) \geqslant \sigma_2(M) \geqslant \cdots \geqslant \sigma_{p_2}(M) > 0, \\ &\sigma_1(N) \geqslant \sigma_2(N) \geqslant \cdots \geqslant \sigma_{p_3}(N) > 0, \end{aligned}$$

are respectively the nonzero singular values of A, M and N. Partition U_i, V_i as follows for $i = 1, 2$:

$$\begin{aligned} U_1 &= \underset{p_1,\, p_2,\, \overline{m}_1}{(U_{11}, U_{12}, U_{13})}, \quad U_2 = \underset{p_3,\, \overline{m}_2}{(U_{21}, U_{22})}, \\ V_1 &= \underset{p_1,\, p_3,\, \overline{n}_1}{(V_{11}, V_{12}, V_{13})}, \quad V_2 = \underset{p_2,\, \overline{n}_2}{(V_{21}, V_{22})}, \end{aligned} \tag{4.5}$$

where $\overline{m}_1 = m_1 - p_1 - p_2$, $\overline{m}_2 = m_2 - p_3$, $\overline{n}_1 = n_1 - p_1 - p_3$, $\overline{n}_2 = n_2 - p_2$. Then from (4.3)–(4.5), one has that [23]

$$\begin{aligned} &A = U_{11} A_{11} V_{11}^H, \qquad P_A^{\perp} = I - U_{11} U_{11}^H, \qquad P_{N(A^H)} = I - V_{11} V_{11}^H, \\ &M = P_A^{\perp} B = U_{12} B_{21} V_{21}^H, \qquad I - M^{\dagger} M = I - V_{21} V_{21}^H = V_{22} V_{22}^H, \\ &N = C P_{N(A^H)} = U_{21} C_{12} V_{22}^H, \qquad I - N N^{\dagger} = I - U_{21} U_{21}^H = U_{22} U_{22}^H. \end{aligned} \tag{4.6}$$

So one carries out from (4.3)–(4.6) that

$$B(I - M^{\dagger} M) = U_{11} B_{12} V_{22}^H, \qquad (I - N N^{\dagger}) C = U_{22} C_{21} V_{11}^H. \tag{4.7}$$

Let $\widehat{G}_3 = G_3 + \Delta G_3$ be the perturbed version of G_3, with $\widehat{A} = A + \Delta A$, $\widehat{B} = B + \Delta B$, $\widehat{C} = C + \Delta C$, $\widehat{D} = D + \Delta D$. We now enforce the conditions

$$p_1 = \operatorname{rank}(\widehat{A}), \qquad p_2 = \operatorname{rank}(\widehat{M}) \quad \text{and} \quad p_3 = \operatorname{rank}(\widehat{N}). \tag{4.8}$$

Then $\widehat{G}_3$ has the following decomposition

$$\begin{aligned} \widehat{G}_3 &= \left(\begin{array}{c|c} \widehat{A} & \widehat{B} \\ \hline \widehat{C} & \widehat{D} \end{array}\right) \\ &= \begin{pmatrix} \widehat{U}_1 & 0 \\ 0 & \widehat{U}_2 \end{pmatrix} \left(\begin{array}{ccc|cc} \widehat{A}_{11} & 0 & 0 & \widehat{B}_{11} & \widehat{B}_{12} \\ 0 & 0 & 0 & \widehat{B}_{21} & 0 \\ 0 & 0 & 0 & 0 & 0 \\ \hline \widehat{C}_{11} & \widehat{C}_{12} & 0 & \widehat{D}_{11} & \widehat{D}_{12} \\ \widehat{C}_{21} & 0 & 0 & \widehat{D}_{21} & \widehat{D}_{22} \end{array}\right) \begin{pmatrix} \widehat{V}_1^H & 0 \\ 0 & \widehat{V}_2^H \end{pmatrix}, \end{aligned} \tag{4.9}$$

where

$$\begin{aligned} \widehat{A}_{11} &= \operatorname{diag}(\sigma_1(\widehat{A}), \sigma_2(\widehat{A}), \ldots, \sigma_{p_1}(\widehat{A})), \\ \widehat{B}_{21} &= \operatorname{diag}(\sigma_1(\widehat{M}), \sigma_2(\widehat{M}), \ldots, \sigma_{p_2}(\widehat{M})), \\ \widehat{C}_{12} &= \operatorname{diag}(\sigma_1(\widehat{N}), \sigma_2(\widehat{N}), \ldots, \sigma_{p_3}(\widehat{N})), \end{aligned}$$

and

$$\sigma_1(\widehat{A}) \geqslant \sigma_2(\widehat{A}) \geqslant \cdots \geqslant \sigma_{p_1}(\widehat{A}) > 0,$$
$$\sigma_1(\widehat{M}) \geqslant \sigma_2(\widehat{M}) \geqslant \cdots \geqslant \sigma_{p_2}(\widehat{M}) > 0,$$
$$\sigma_1(\widehat{N}) \geqslant \sigma_2(\widehat{N}) \geqslant \cdots \geqslant \sigma_{p_3}(\widehat{N}) > 0,$$

are respectively the nonzero singular values of $\widehat{A}$, $\widehat{M}$ and $\widehat{N}$. For $i = 1, 2$, partition $\widehat{U}_i$, $\widehat{V}_i$ as follows,

$$\begin{aligned} \widehat{U}_1 &= \underset{p_1,\, p_2,\, \overline{m}_1}{(\widehat{U}_{11}, \widehat{U}_{12}, \widehat{U}_{13})}, & \widehat{U}_2 &= \underset{p_3,\, \overline{m}_2}{(\widehat{U}_{21}, \widehat{U}_{22})} \\ \widehat{V}_1 &= \underset{p_1,\, p_3,\, \overline{n}_1}{(\widehat{V}_{11}, \widehat{V}_{12}, \widehat{V}_{13})}, & \widehat{V}_2 &= \underset{p_2,\, \overline{n}_2}{(\widehat{V}_{21}, \widehat{V}_{22})}. \end{aligned} \tag{4.10}$$

Then from (4.8)–(4.10), one has that

$$\begin{aligned} &\widehat{A} = \widehat{U}_{11}\widehat{A}_{11}\widehat{V}_{11}^H, \qquad P_{\widehat{A}}^{\perp} = I - \widehat{U}_{11}\widehat{U}_{11}^H, \qquad P_{\widehat{A}^H}^{\perp} = I - \widehat{V}_{11}\widehat{V}_{11}^H, \\ &\widehat{M} = P_{\widehat{A}}^{\perp}\widehat{B} = \widehat{U}_{12}\widehat{B}_{21}\widehat{V}_{21}^H, \qquad I - \widehat{M}^{\dagger}\widehat{M} = I - \widehat{V}_{21}\widehat{V}_{21}^H = \widehat{V}_{22}\widehat{V}_{22}^H, \\ &\widehat{N} = \widehat{C}P_{\widehat{A}^H}^{\perp} = \widehat{U}_{21}\widehat{C}_{12}\widehat{V}_{22}^H, \qquad I - \widehat{N}\widehat{N}^{\dagger} = I - \widehat{U}_{21}\widehat{U}_{21}^H = \widehat{U}_{22}\widehat{U}_{22}^H. \end{aligned} \tag{4.11}$$

Therefore,

$$\widehat{B}(I - \widehat{M}^{\dagger}\widehat{M}) = \widehat{U}_{11}\widehat{B}_{12}\widehat{V}_{22}^H, \qquad (I - \widehat{N}\widehat{N}^{\dagger})\widehat{C} = \widehat{U}_{22}\widehat{C}_{21}\widehat{V}_{11}^H. \tag{4.12}$$

Theorem 4.1. *Let G_3 and $\widehat{G}_3$ and their decompositions be in (4.4) and (4.8), respectively, with $\widehat{A} = A + \Delta A$, $\widehat{B} = B + \Delta B$, $\widehat{C} = C + \Delta C$, $\widehat{D} = D + \Delta D$. Let*

$$\begin{aligned} D_1 &= (I - NN^{\dagger})(D - CA^{\dagger}B)(I - M^{\dagger}M), \\ \widehat{D}_1 &= (I - \widehat{N}\widehat{N}^{+})(\widehat{D} - \widehat{C}\widehat{A}^{+}\widehat{B})(I - \widehat{M}^{+}\widehat{M}). \end{aligned} \tag{4.13}$$

If

$$\begin{aligned} &\operatorname{rank}(A) = \operatorname{rank}(\widehat{A}), \qquad \operatorname{rank}(M) = \operatorname{rank}(\widehat{M}), \\ &\operatorname{rank}(N) = \operatorname{rank}(\widehat{N}), \end{aligned} \tag{4.14}$$

then

$$\begin{aligned} \|D_1 - \widehat{D}_1\| \leqslant\ & \|\widehat{U}_{22}^H \Delta D V_{22}\| + \|\widehat{U}_{22}^H \widehat{C}\widehat{A}^{\dagger} \Delta B V_{22}\| + \|\widehat{U}_{22}^H \widehat{C}\widehat{A}^{\dagger} \Delta A A^{\dagger} B V_{22}\| + \|\widehat{U}_{22}^H \Delta C A^{\dagger} B V_{22}\| \\ & + \omega_1 \|\widehat{U}_{22}^H(\widehat{D} - \widehat{C}\widehat{A}^{\dagger}\widehat{B})\| + \omega_2 \|(D - CA^{\dagger}B)V_{22}\|, \end{aligned} \tag{4.15}$$

where

$$\begin{aligned} \omega_1 &= \min\left\{\frac{\eta_1}{\sigma_{p_2}(\widehat{M})}, \frac{\eta_2}{\sigma_{p_2}(M)}\right\}, \\ \omega_2 &= \min\left\{\frac{\eta_3}{\sigma_{p_3}(\widehat{N})}, \frac{\eta_4}{\sigma_{p_3}(N)}\right\}, \end{aligned} \tag{4.16}$$

η_1, η_2 are defined in Theorem 3.2, and η_3, η_4 are defined in Theorem 3.4.

Proof. From the formulations of D_1 and $\widehat{D}_1$ in (4.13),

$$\begin{aligned} \|D_1 - \widehat{D}_1\| \leq\ & \|(I - \widehat{N}\widehat{N}^{\dagger})\left(\Delta D - \widehat{C}\widehat{A}^{\dagger}\Delta B - \widehat{C}(\widehat{A}^{\dagger} - A^{\dagger})B - \Delta C A^{\dagger} B\right)(I - M^{\dagger}M)\| \\ & + \|(I - \widehat{N}\widehat{N}^{\dagger})(\widehat{D} - \widehat{C}\widehat{A}^{\dagger}\widehat{B})(M^{\dagger}M - \widehat{M}^{\dagger}\widehat{M})\| + \|(NN^{\dagger} - \widehat{N}\widehat{N}^{\dagger})(D - CA^{\dagger}B)(I - M^{\dagger}M)\|. \end{aligned} \tag{4.17}$$

Notice that [19]

$$\widehat{A}^{\dagger} - A^{\dagger} = -\widehat{A}^{\dagger}\Delta A A^{\dagger} + \widehat{A}^{\dagger}(I - AA^{\dagger}) - (I - \widehat{A}^{\dagger}\widehat{A})A^{\dagger},$$

so

$$(I - \widehat{N}\widehat{N}^{\dagger})\widehat{C}(\widehat{A}^{\dagger} - A^{\dagger})B(I - M^{\dagger}M) = -(I - \widehat{N}\widehat{N}^{\dagger})\widehat{C}\widehat{A}^{\dagger}\Delta A A^{\dagger} B(I - M^{\dagger}M),$$

because $(I - AA^\dagger)B = M$ and $\widehat{C}(I - \widehat{A}^\dagger\widehat{A}) = \widehat{N}$. Notice that

$$I - M^\dagger M = I - V_{21}V_{21}^H = V_{22}V_{22}^H, \qquad I - NN^\dagger = I - U_{21}U_{21}^H = U_{22}U_{22}^H$$
$$I - \widehat{M}^\dagger\widehat{M} = I - \widehat{V}_{21}\widehat{V}_{21}^H = \widehat{V}_{22}\widehat{V}_{22}^H, \qquad I - \widehat{N}\widehat{N}^\dagger = I - \widehat{U}_{21}\widehat{U}_{21}^H = \widehat{U}_{22}\widehat{U}_{22}^H.$$

Therefore, in (4.17) by applying Corollary 3.3 and Theorem 3.4, we obtain the desired estimates of the theorem. □

Remark 4.1. Obviously, the estimates derived in Theorem 4.1 is sharper than that in Theorem 3.1 of [23]. In Remark 3.2 of [23] Wei mentioned that Demmel in formula (∗), p. 206 of [2] just considered the simplest case that both G_3 and $\widehat{G}_3$ can be transformed into the standard forms as in (4.4) and (4.9) with the same pairs of unitary matrices $\begin{pmatrix} U_1 & \\ & U_2 \end{pmatrix}$ and $\begin{pmatrix} V_1 & \\ & V_2 \end{pmatrix}$. In this case $M^\dagger M = \widehat{M}^\dagger\widehat{M}$ and $NN^\dagger = \widehat{N}\widehat{N}^\dagger$, and the estimate in Theorem 3.1 of [23] reduces to that obtained in p. 206 of [2]. In Example 3.1 of [23] it is obvious that the estimates in [2] is not valid in general case.

Theorem 4.2. *Under the notation and conditions in Theorem 4.1, furthermore, suppose that the SVD of D_1 and $\widehat{D}_1$ are respectively*

$$\begin{aligned} D_1 &= (Z_1, Z_2)\mathrm{diag}(T_1, T_2)(W_1, W_2)^H, \\ \widehat{D}_1 &= (\widehat{Z}_1, \widehat{Z}_2)\mathrm{diag}(\widehat{T}_1, \widehat{T}_2)(\widehat{W}_1, \widehat{W}_2)^H, \end{aligned} \tag{4.18}$$

where $l_1 = \min\{m_2 - p_3, n_2 - p_2\}$, $Z, \widehat{Z}, W, \widehat{W}$ are unitary matrices, $Z_1, \widehat{Z}_1, W_1, \widehat{W}_1$ are respectively the first p columns of $Z, \widehat{Z}, W, \widehat{W}$,

$$T_1 = \mathrm{diag}(\sigma_1(D_1), \ldots, \sigma_p(D_1)), \qquad T_2 = \mathrm{diag}(\sigma_{p+1}(D_1), \ldots, \sigma_{l_1}(D_1)),$$
$$\widehat{T}_1 = \mathrm{diag}(\sigma_1(\widehat{D}_1), \ldots, \sigma_p(\widehat{D}_1)), \qquad T_2 = \mathrm{diag}(\sigma_{p+1}(\widehat{D}_1), \ldots, \sigma_{l_1}(\widehat{D}_1)),$$

$$\sigma_1(D_1) \geqslant \cdots \geqslant \sigma_p(D_1) > \sigma_{p+1}(D_1) \geqslant \cdots \geqslant \sigma_{l_1}(D_1)$$
$$\sigma_1(\widehat{D}_1) \geqslant \cdots \geqslant \sigma_p(\widehat{D}_1) \geqslant \sigma_{p+1}(\widehat{D}_1) \geqslant \cdots \geqslant \sigma_{l_1}(\widehat{D}_1)$$

are the singular values of D_1 and $\widehat{D}_1$, respectively. If $\sigma_p(D_1) - \sigma_{p+1}(D_1) > 2\|\Delta D_1\|$, then

$$\begin{aligned} \|Z_1Z_1^H - \widehat{Z}_1\widehat{Z}_1^H\| &\leqslant \min\left\{\frac{\eta^{(1)}}{\sigma_p(\widehat{D}_1) - \sigma_{p+1}(D_1)}, \frac{\eta^{(2)}}{\sigma_p(D_1) - \sigma_{p+1}(\widehat{D}_1)}\right\}, \\ \|W_1W_1^H - \widehat{W}_1\widehat{W}_1^H\| &\leqslant \min\left\{\frac{\eta^{(1)}}{\sigma_p(\widehat{D}_1) - \sigma_{p+1}(D_1)}, \frac{\eta^{(2)}}{\sigma_p(D_1) - \sigma_{p+1}(\widehat{D}_1)}\right\}, \end{aligned} \tag{4.19}$$

where

$$\begin{aligned} \eta^{(1)} &= \max\left\{\|\widehat{Z}_1^H\Delta D_1W_2\|, \|Z_2^H\Delta D_1\widehat{W}_1\|\right\}, \\ \eta^{(2)} &= \max\left\{\|Z_1^H\Delta D_1\widehat{W}_2\|, \|\widehat{Z}_2^H\Delta D_1W_1\|\right\}. \end{aligned} \tag{4.20}$$

Proof. From the perturbation analysis of the singular values, with the conditions of the theorem, we have $\sigma_p(D_1) - \sigma_{p+1}(\widehat{D}_1) > 0$ and $\sigma_p(\widehat{D}_1) - \sigma_{p+1}(D_1) > 0$. Furthermore, from the formulas in (4.18),

$$Z_2^H\Delta D_1\widehat{W}_1 = Z_2^H\widehat{Z}_1\widehat{T}_1 - T_2W_2^H\widehat{W}_1,$$
$$\widehat{Z}_1^H\Delta D_1W_2 = \widehat{T}_1\widehat{W}_1^HW_2 - \widehat{Z}_1^HZ_2T_2.$$

Therefore,

$$\|Z_2^H\widehat{Z}_1\| \le \frac{1}{\sigma_p(\widehat{D}_1)}\left[\|Z_2^H\Delta D_1\widehat{W}_1\| + \sigma_{p+1}(D_1)\|W_2^H\widehat{W}_1\|\right],$$
$$\|\widehat{W}_1^HW_2\| \le \frac{1}{\sigma_p(\widehat{D}_1)}\left[\|\widehat{Z}_1^H\Delta D_1W_2\| + \sigma_{p+1}(D_1)\|\widehat{Z}_1^HZ_2\|\right],$$

and

$$\|Z_2^H\widehat{Z}_1\| \le \frac{1}{\sigma_p(\widehat{D}_1)}\left\{\|Z_2^H\Delta D_1\widehat{W}_1\| + \frac{\sigma_{p+1}(D_1)}{\sigma_p(\widehat{D}_1)}\left[\|\widehat{Z}_1^H\Delta D_1W_2\| + \sigma_{p+1}(D_1)\|\widehat{Z}_1^HZ_2\|\right]\right\},$$
$$\frac{\sigma_p(\widehat{D}_1)^2 - \sigma_{p+1}(D_1)^2}{\sigma_p(\widehat{D}_1)^2}\|Z_2^H\widehat{Z}_1\| \le \frac{\sigma_p(\widehat{D}_1) + \sigma_{p+1}(D_1)}{\sigma_p(\widehat{D}_1)^2}\eta_1,$$

so $\|Z_2^H\widehat{Z}_1\| \le \frac{\eta_1}{\sigma_p(\widehat{D}_1) - \sigma_{p+1}(D_1)}$. In a similar manner, we also have $\|\widehat{Z}_2^HZ_1\| \le \frac{\eta_2}{\sigma_p(D_1) - \sigma_{p+1}(\widehat{D}_1)}$. Notice that from Lemma 2.2, $\|Z_2^H\widehat{Z}_1\| = \|\widehat{Z}_2^HZ_1\|$, we then obtain the first inequality in (4.19). The second one in (4.19) can be derived similarly. □

5. Concluding remarks

In this paper, we have deduced the perturbation bounds of some constrained subspaces that relate to the matrices G_1, G_2 and G_3. In a separate paper, we will study the perturbation analysis of the LS–TLS, LSE, and CTLS problems using the bounds obtained in this paper.

Acknowledgments

We are grateful to the editor and the anonymous referees for providing many useful comments and suggestions, which greatly improve the presentation of the article.

References

[1] C. Davis, W.M. Kahan, The rotation of eigenvectors by a perturbation III, SIAM J. Numer. Anal. 7 (1970) 1–46.
[2] J. Demmel, The smallest perturbation of a submatrix which lowers the rank and constrained total least squares problems, SIAM J. Numer. Anal. 24 (1987) 199–206.
[3] L. Eldén, Perturbation theory for the least squares problem with linear equality constraints, SIAM J. Numer. Anal. 17 (1980) 338–350.
[4] G.H. Golub, A. Hoffman, G.W. Stewart, A generalization of the Eckart–Young–Mirky matrix approximation theorem, Linear Algebra Appl. 88/89 (1987) 317–327.
[5] T. Kato, Perturbation Theory for Linear Operators, Springer-verlag, New York, 1966.
[6] R.-C. Li, Bounds on perturbations of generalized singular values and of accociated subspaces, SIAM J. Matrix Anal. Appl. 14 (1) (1993) 195–234.
[7] R.-C. Li, On perturbations of matrix pencils with real spectra, a revisit, Math. Comp. 72 (242) (2003) 715–728.
[8] C.C. Paige, M.A. Saunders, Toward a generalized singular value decomposition, SIAM J. Numer. Anal. 18 (1981) 398–405.
[9] C.C. Paige, M. Wei, Analysis of the generalized total least squares problem $AX \approx B$ when some columns are free of error, Numer. Math. 65 (1993) 177–202.
[10] G.W. Stewart, Error and perturbation bounds for subspaces associated with certain eigenvalue problems, SIAM Rev. 15 (1973) 727–764.
[11] G.W. Stewart, On the perturbation of pseudo-inverses, projections, and linear least squares problems, SIAM Rev. 19 (1977) 634–662.
[12] G.W. Stewart, J.-G. Sun, Matrix Perturbation Theory, Academic Press, Boston, 1990.
[13] J.-G. Sun, Orthogonal projections and the perturbation of the eigenvalues of singular pencils, J. Comput. Math. 1 (1983) 63–74.
[14] J.-G. Sun, Perturbation analysis for the generalized singular value problem, SIAM J. Numer. Anal. 20 (1983) 611–625.
[15] J.-G. Sun, The perturbation bounds for eigenspaces of a definite matrix-pair, Numer. Math. 41 (1983) 321–343.
[16] S. Van Huffel, J. Vandewalle, Analysis and properties of the generalized total least squares problem $AX \approx B$ when some or all columns in A are subject to error, SIAM J. Matrix Anal. Appl. 10 (1989) 294–315.
[17] S. Van Huffel, J. Vandewalle, The Total Least Squares Problem: Computational Aspects and Analysis, SIAM, Philadelphia, 1991.
[18] P.A. Wedin, Perturbation bounds in connection with the singular value decomposition, BIT 12 (1972) 99–111.
[19] P.A. Wedin, On the almost rank deficient case of the least squares problem, BIT 13 (1973) 344–354.
[20] P.A. Wedin, Perturbation theory for pseudoinverses, BIT 13 (1973) 217–232.
[21] P.A. Wedin, On angles between subspaces of a finite dimensional inner product space, in: B. Kågström, A. Ruhe (Eds.), Matrix Pencils, Springer, New York, 1983, pp. 263–285.
[22] M. Wei, Perturbation analysis for the rank deficient equality constrained least squares prolem, SIAM J. Numer Anal. 29 (1992) 1462–1481.
[23] M. Wei, Perturbation theory for the Eckart–Young–Mirsky theorem and the constrained total least squares problem, Linear Algebra Appl. 280 (1998) 267–287.
[24] M. Wei, On perturbation analysis for TLS and LS solution, Numer. Math. Sin. 20 (3) (1998) 267–278 (in Chinese).
[25] M. Wei, The Theory and Computation of Generalized Least Squares Problem (in Chinese), Science Press, Beijing, 2006.
[26] M. Wei, A.R. De Pierro, Perturbation analysis of the canonical subspaces, Linear Algebra Appl. 279 (1998) 135–151.