

An iterative method for solving a kind of constrained linear matrix equations system

JING CAI^{1,2} and GUOLIANG CHEN¹

¹Department of Mathematics, East China Normal University, Shanghai 200062, P.R. China

²School of Science, Huzhou Teachers College, Huzhou Zhejiang 313000, P.R. China

E-mails: glchen@math.ecnu.edu.cn / caijing@hutc.zj.cn

Abstract. In this paper, an iterative method is constructed to solve the following constrained linear matrix equations system: $[\mathcal{A}_1(X), \mathcal{A}_2(X), \dots, \mathcal{A}_r(X)] = [E_1, E_2, \dots, E_r]$, $X \in \mathcal{S} = \{X | X = \mathcal{U}(X)\}$, where \mathcal{A}_i is a linear operator from $C^{m \times n}$ onto $C^{p_i \times q_i}$, $E_i \in C^{p_i \times q_i}$, $i = 1, 2, \dots, r$, and \mathcal{U} is a linear self-conjugate involution operator. When the above constrained matrix equations system is consistent, for any initial matrix $X_0 \in \mathcal{S}$, a solution can be obtained by the proposed iterative method in finite iteration steps in the absence of roundoff errors, and the least Frobenius norm solution can be derived when a special kind of initial matrix is chosen. Furthermore, the optimal approximation solution to a given matrix can be derived. Several numerical examples are given to show the efficiency of the presented iterative method.

Mathematical subject classification: 15A24, 65D99, 65F30.

Key words: iterative method, linear matrix equations system, linear operator, least Frobenius norm solution, optimal approximation.

1 Introduction

Throughout this paper, we need the following notations. $C^{m \times n}$ denotes the set of $m \times n$ complex matrices. For a matrix $A \in C^{m \times n}$, we denote its transpose, conjugate transpose, trace, Frobenius norm and column space by A^T , A^H , $\text{tr}(A)$, $\|A\|$ and $R(A)$, respectively. Let I_n and S_n denote the $n \times n$ unit matrix and reverse unit matrix respectively. The symbol $\text{vec}(\cdot)$ stands for the vec operator,

i.e., for $A = (a_1, a_2, \dots, a_n) \in C^{m \times n}$, where $a_i (i = 1, 2, \dots, n)$ denotes the i th column of A , $\text{vec}(A) = (a_1^T, a_2^T, \dots, a_n^T)^T$. In the vector space $C^{m \times n}$, we define inner product as: $\langle X, Y \rangle = \text{tr}(Y^H X)$ for all $X, Y \in C^{m \times n}$. Two matrices X and Y are said to be orthogonal if $\langle X, Y \rangle = 0$.

Let $LC^{m \times n, p \times q}$ denote the set of linear operators from $C^{m \times n}$ onto $C^{p \times q}$. Particularly, when $p = m$ and $q = n$, $LC^{m \times n, p \times q}$ is denoted by $LC^{m \times n}$. Let \mathcal{I} stand for the identity operator on $C^{m \times n}$. For $\mathcal{A} \in LC^{m \times n, p \times q}$, its conjugate operator \mathcal{A}^* is a linear operator satisfying the following equality:

$$\langle \mathcal{A}(X), Y \rangle = \langle X, \mathcal{A}^*(Y) \rangle, \quad \text{for all } X \in C^{m \times n}, Y \in C^{p \times q}.$$

For example, if \mathcal{A} is defined as: $\mathcal{A} : X \rightarrow AXB$, then we have $\mathcal{A}^* : Y \rightarrow A^H Y B^H$.

We note that the linear matrix equations, such as the well-known Lyapunov matrix equation $XA + A^*X = H$, Sylvester equation $AX \pm XB = C$, Stein equation $X \pm AXB = D$, $AXB + CX^T D = E$ and $[A_1 X B_1, A_2 X B_2] = [C_1, C_2]$, which have many applications in system theory (see [7–11, 15, 18] for instance), can all be rewritten as the following linear matrix equations system:

$$[\mathcal{A}_1(X), \mathcal{A}_2(X), \dots, \mathcal{A}_r(X)] = [E_1, E_2, \dots, E_r], \quad (1.1)$$

where $\mathcal{A}_i \in LC^{m \times n, p_i \times q_i}$ and $E_i \in C^{p_i \times q_i}, i = 1, 2, \dots, r$.

Recently, there are many papers considering the matrix equations system (1.1) with various constraints on solutions. For instance, Chu [4], Dai [5] and Chang and Wang [2] considered the symmetric solutions of the matrix equations $AX = B, AXB = C$ and $(A^T X A, B^T X B) = (C, D)$ respectively. Peng et al. [14] and Qiu et al. [16] discussed the reflexive and anti-reflexive solution of the matrix equations $A^H X B = C$ and $AX = B, X C = D$, respectively. Li et al. [10], Yuan and Dai [19] and Zhang et al. [20] considered the generalized reflexive and anti-reflexive solutions of the matrix equations $(B X = C, X D = E), AXB = D$ and $AX = B$. In these literatures, the necessary and sufficient conditions for the consistency of the constrained matrix equations and the explicit expressions of the constrained solutions were derived by using the singular value decomposition (SVD), generalized singular value decomposition (GSVD), canonical correlation decomposition (CCD) or the generalized

inverses of matrices. On the other hand, Peng et al. [13] considered an iterative method for symmetric solutions of the system of matrix equations $A_1XB_1 = C_1$, $A_2XB_2 = C_2$. In Huang [9], an iterative method was constructed to obtain the skew-symmetric solution of the matrix equation $AXB = C$. Dehghan and Hajarian [6] proposed an iterative algorithm to obtain the reflexive solution (namely the generalized centro-symmetric solution) of the matrix equations $AYB = E, CYD = F$.

It should be remarked that the following common constraints on the solutions of matrix equations (see [3, 12, 14, 19, 20] for more details):

- (1) Symmetric (Skew-symmetric) constraint:
 $X = X^T (X = -X^T)$;
- (2) Centrosymmetric (Centroskew symmetric) constraint:
 $X = S_nXS_n (-S_nXS_n)$;
- (3) Reflexive (Anti-reflexive) constraint:
 $X = PXP (-PXP)$;
- (4) Generalized reflexive (anti-reflexive) constraint:
 $X = PXQ (-PXQ)$;
- (5) P orthogonal-symmetric (orthogonal-anti-symmetric) constraint:
 $(PX)^T = PX(-PX)$, here $P^T = P = P^{-1} \neq I_n$ and $Q^T = Q = Q^{-1} \neq I_n$, are all special cases of the following constraint:

$$X = \mathcal{U}(X) \tag{1.2}$$

where $\mathcal{U} \in LC^{m \times n}$ is a self-conjugate involution operator, i.e., $\mathcal{U}^* = \mathcal{U}, \mathcal{U}^2 = \mathcal{I}$.

Hence it motivates us to consider solving the more general matrix equations system (1.1) with the constraint (1.2) and its associated optimal approximate problem (see the following two problems) by iterative methods, which generalize the main results of [6, 9, 13].

Problem I. Let \mathcal{S} denote the set of constrained matrices defined by (1.2). For given $\mathcal{A}_i \in LC^{m \times n, p_i \times q_i}, E_i \in C^{p_i \times q_i}, i = 1, 2, \dots, r$, find $X \in \mathcal{S}$ such that

$$[\mathcal{A}_1(X), \mathcal{A}_2(X), \dots, \mathcal{A}_r(X)] = [E_1, E_2, \dots, E_r]. \tag{1.3}$$

Problem II. Let S denote the solution set of Problem I, for given $\widehat{X} \in C^{m \times n}$, find $\overline{X} \in S$, such that

$$\|\overline{X} - \widehat{X}\| = \min_{X \in S} \|X - \widehat{X}\|. \quad (1.4)$$

The rest of the paper is organized as follows. In Section 2, we propose an iterative algorithm to obtain a solution or the least Frobenius norm solution of Problem I and present some basic properties of the algorithm. In Section 3, we consider the iterative method for solving Problem II. Some numerical examples are given in Section 4 to show the efficiency of the proposed iterative method. Finally, we put some conclusions in Section 5.

2 Iterative method for solving Problem I

In this section, we first propose an iterative algorithm to solve Problem I, then present some basic properties of the algorithm. We also consider finding the least Frobenius norm solution of Problem I. In the sequel, the least norm solution always means the least Frobenius norm solution.

Algorithm 2.1.

Step 1. Input $\mathcal{A}_i \in LC^{m \times n, p_i \times q_i}$, $E_i \in C^{p_i \times q_i}$, $i = 1, 2, \dots, r$, and an arbitrary $X_0 \in \mathcal{S}$;

Step 2. Compute

$$R_0^{(l)} = E_l - \mathcal{A}_l(X_0), l = 1, 2, \dots, r;$$

$$Z_0 = \frac{1}{2} \sum_{l=1}^r [\mathcal{A}_l^*(R_0^{(l)}) + \mathcal{U} \mathcal{A}_l^*(R_0^{(l)})];$$

$$P_0 = Z_0;$$

$$k := 0;$$

Step 3. If $\sum_{l=1}^r \|R_k^{(l)}\|^2 = 0$ or $\sum_{l=1}^r \|R_k^{(l)}\|^2 \neq 0$ but $P_k = 0$, then stop; else, $k := k + 1$;

Step 4. Compute

$$\begin{aligned} \alpha_{k-1} &= \frac{\sum_{l=1}^r \|R_{k-1}^{(l)}\|^2}{\|P_{k-1}\|^2}; \\ X_k &= X_{k-1} + \alpha_{k-1}P_{k-1}; \\ R_k^{(l)} &= E_l - \mathcal{A}_l(X_k) = R_{k-1}^{(l)} - \alpha_{k-1}\mathcal{A}_l(P_{k-1}), l = 1, 2, \dots, r; \\ Z_k &= \frac{1}{2} \sum_{l=1}^r [\mathcal{A}_l^*(R_k^{(l)}) + \mathcal{U} \mathcal{A}_l^*(R_k^{(l)})], \beta_{k-1} = \frac{\text{tr}(Z_k^H P_{k-1})}{\|P_{k-1}\|^2}; \\ P_k &= Z_k - \beta_{k-1}P_{k-1}; \end{aligned}$$

Step 5. Go to Step 3.

Some basic properties of Algorithm 2.1 are listed in the following lemmas.

Lemma 2.1. *The sequences $\{X_i\}$ and $\{P_i\}$ generated by Algorithm 2.1 are contained in the constraint set \mathcal{S} .*

Proof. We prove the conclusion by induction. For $i = 0$, by $\mathcal{U}^2 = I$ and Algorithm 2.1, we have

$$\begin{aligned} \mathcal{U}(P_0) &= \mathcal{U} \left(\frac{1}{2} \sum_{l=1}^r [\mathcal{A}_l^*(R_0^{(l)}) + \mathcal{U} \mathcal{A}_l^*(R_0^{(l)})] \right) \\ &= \frac{1}{2} \sum_{l=1}^r [\mathcal{U} \mathcal{A}_l^*(R_0^{(l)}) + \mathcal{A}_l^*(R_0^{(l)})] = P_0, \end{aligned}$$

i.e., $P_0 \in \mathcal{S}$.

For $i = 1$, since $X_0 \in \mathcal{S}$, by Algorithm 2.1, we have

$$\mathcal{U}(X_1) = \mathcal{U}(X_0) + \alpha_0 \mathcal{U}(P_0) = X_0 + \alpha_0 P_0 = X_1, \quad \text{i.e., } X_1 \in \mathcal{S}.$$

Moreover, we have

$$\begin{aligned} \mathcal{U}(P_1) &= \mathcal{U}(Z_1) - \beta_0 \mathcal{U}(P_0) \\ &= \mathcal{U} \left(\frac{1}{2} \sum_{l=1}^r [\mathcal{A}_l^*(R_1^{(l)}) + \mathcal{U} \mathcal{A}_l^*(R_1^{(l)})] \right) - \beta_0 P_0 \\ &= \frac{1}{2} \sum_{l=1}^r [\mathcal{U} \mathcal{A}_l^*(R_1^{(l)}) + \mathcal{A}_l^*(R_1^{(l)})] - \beta_0 P_0 \\ &= Z_1 - \beta_0 P_0 = P_1, \end{aligned}$$

i.e., $P_1 \in \mathcal{S}$. Assume that the conclusion holds for $i = s$ ($s > 1$), i.e., $X_s, P_s \in \mathcal{S}$. Then

$$\mathcal{U}(X_{s+1}) = \mathcal{U}(X_s) + \alpha_s \mathcal{U}(P_s) = X_s + \alpha_s P_s = X_{s+1}$$

and

$$\begin{aligned} \mathcal{U}(P_{s+1}) &= \mathcal{U}(Z_{s+1}) - \beta_s \mathcal{U}(P_s) \\ &= \mathcal{U} \left(\frac{1}{2} \sum_{l=1}^r [\mathcal{A}_l^*(R_{s+1}^{(l)}) + \mathcal{U} \mathcal{A}_l^*(R_{s+1}^{(l)})] \right) - \beta_s \mathcal{U} P_s \\ &= \frac{1}{2} \sum_{l=1}^r [\mathcal{U} \mathcal{A}_l^*(R_{s+1}^{(l)}) + \mathcal{A}_l^*(R_{s+1}^{(l)})] - \beta_s P_s \\ &= Z_{s+1} - \beta_s P_s = P_{s+1}, \end{aligned}$$

from which we get $X_{s+1}, P_{s+1} \in \mathcal{S}$.

By the principle of induction, we draw the conclusion. □

Lemma 2.2. *Assume that X^* is a solution of Problem 1. Then the sequences $\{X_i\}, \{R_i^{(l)}\}$ and $\{P_i\}$ generated by Algorithm 2.1 satisfy the following equality:*

$$\langle X^* - X_i, P_i \rangle = \sum_{l=1}^r \|R_i^{(l)}\|^2, i = 0, 1, 2, \dots \tag{2.1}$$

Proof. We prove the conclusion by induction. For $i = 0$, it follows from $X^*, X_0 \in \mathcal{S}$ that $X^* - X_0 \in \mathcal{S}$. Then by Algorithm 2.1, we have

$$\begin{aligned} \langle X^* - X_0, P_0 \rangle &= \langle X^* - X_0, \frac{1}{2} \sum_{l=1}^r [\mathcal{A}_l^*(R_0^{(l)}) + \mathcal{U} \mathcal{A}_l^*(R_0^{(l)})] \rangle \\ &= \frac{1}{2} \sum_{l=1}^r \langle X^* - X_0, \mathcal{A}_l^*(R_0^{(l)}) + \mathcal{U} \mathcal{A}_l^*(R_0^{(l)}) \rangle \\ &= \frac{1}{2} \sum_{l=1}^r \langle \mathcal{A}_l(X^* - X_0) + \mathcal{A}_l \mathcal{U}(X^* - X_0), R_0^{(l)} \rangle \\ &= \sum_{l=1}^r \langle \mathcal{A}_l(X^* - X_0), R_0^{(l)} \rangle = \sum_{l=1}^r \langle E_l - \mathcal{A}_l(X_0), R_0^{(l)} \rangle \\ &= \sum_{l=1}^r \langle R_0^{(l)}, R_0^{(l)} \rangle = \sum_{l=1}^r \|R_0^{(l)}\|^2. \end{aligned}$$

For $i = 1$, by Algorithm 2.1 and Lemma 2.1, we have

$$\begin{aligned}
 \langle X^* - X_1, P_1 \rangle &= \langle X^* - X_1, Z_1 - \beta_0 P_0 \rangle \\
 &= \langle X^* - X_1, Z_1 \rangle - \beta_0 \langle X^* - X_1, P_0 \rangle \\
 &= \langle X^* - X_1, \frac{1}{2} \sum_{l=1}^r [\mathcal{A}_l^*(R_1^{(l)}) + \mathcal{U} \mathcal{A}_l^*(R_1^{(l)})] \rangle \\
 &\quad - \beta_0 \langle X^* - X_0 - \alpha_0 P_0, P_0 \rangle \\
 &= \frac{1}{2} \sum_{l=1}^r \langle X^* - X_1, \mathcal{A}_l^*(R_1^{(l)}) + \mathcal{U} \mathcal{A}_l^*(R_1^{(l)}) \rangle \\
 &\quad - \beta_0 \langle X^* - X_0, P_0 \rangle + \beta_0 \alpha_0 \langle P_0, P_0 \rangle \\
 &= \frac{1}{2} \sum_{l=1}^r \langle \mathcal{A}_l(X^* - X_1) + \mathcal{A}_l \mathcal{U}(X^* - X_1), R_1^{(l)} \rangle \\
 &\quad - \beta_0 \sum_{l=1}^r \|R_0^{(l)}\|^2 + \beta_0 \frac{\sum_{l=1}^r \|R_0^{(l)}\|^2}{\|P_0\|^2} \|P_0\|^2 \\
 &= \sum_{l=1}^r \langle \mathcal{A}_l(X^* - X_1), R_1^{(l)} \rangle = \sum_{l=1}^r \langle E_l - \mathcal{A}_l(X_1), R_1^{(l)} \rangle \\
 &= \sum_{l=1}^r \|R_1^{(l)}\|^2.
 \end{aligned}$$

Assume that the conclusion holds for $i = s (s > 1)$, i.e., $\langle X^* - X_s, P_s \rangle = \sum_{l=1}^r \|R_s^{(l)}\|^2$, then for $i = s + 1$, we have

$$\begin{aligned}
 \langle X^* - X_{s+1}, P_{s+1} \rangle &= \langle X^* - X_{s+1}, Z_{s+1} \rangle - \beta_s \langle X^* - X_{s+1}, P_s \rangle \\
 &= \langle X^* - X_{s+1}, \frac{1}{2} \sum_{l=1}^r [\mathcal{A}_l^*(R_{s+1}^{(l)}) + \mathcal{U} \mathcal{A}_l^*(R_{s+1}^{(l)})] \rangle \\
 &\quad - \beta_s \langle X^* - X_s - \alpha_s P_s, P_s \rangle \\
 &= \frac{1}{2} \sum_{l=1}^r \langle X^* - X_{s+1}, \mathcal{A}_l^*(R_{s+1}^{(l)}) + \mathcal{U} \mathcal{A}_l^*(R_{s+1}^{(l)}) \rangle \\
 &\quad - \beta_s \langle X^* - X_s, P_s \rangle + \beta_s \alpha_s \langle P_s, P_s \rangle
 \end{aligned}$$

$$\begin{aligned}
 &= \sum_{l=1}^r \langle \mathcal{A}_l(X^* - X_{s+1}), R_{s+1}^{(l)} \rangle - \beta_s \sum_{l=1}^r \|R_s^{(l)}\|^2 \\
 &\quad + \beta_s \frac{\sum_{l=1}^r \|R_s^{(l)}\|^2}{\|P_s\|^2} \|P_s\|^2 \\
 &= \sum_{l=1}^r \langle E_l - \mathcal{A}_l(X_{s+1}), R_{s+1}^{(l)} \rangle = \sum_{l=1}^r \|R_{s+1}^{(l)}\|^2.
 \end{aligned}$$

This completes the proof by the principle of induction. □

Remark 2.1. Lemma 2.2 implies that if Problem I is consistent, then $\sum_{l=1}^r \|R_i^{(l)}\|^2 \neq 0$ implies that $P_i \neq 0$. Else if there exists a positive number k such that $\sum_{l=1}^r \|R_k^{(l)}\|^2 \neq 0$ but $P_k = 0$, then Problem I must be inconsistent. Hence the solvability of Problem I can be determined by Algorithm 2.1 in the absence of roundoff errors.

Lemma 2.3. For the sequences $\{R_i^{(l)}\}$, $\{P_i\}$ and $\{Z_i\}$ generated by Algorithm 2.1, it follows that

$$\sum_{l=1}^r \langle R_{i+1}^{(l)}, R_j^{(l)} \rangle = \sum_{l=1}^r \langle R_i^{(l)}, R_j^{(l)} \rangle - \alpha_i \langle P_i, Z_j \rangle. \tag{2.2}$$

Proof. By Algorithm 2.1, Lemma 2.1 and $\mathcal{U}^* = \mathcal{U}$, we have

$$\begin{aligned}
 \sum_{l=1}^r \langle R_{i+1}^{(l)}, R_j^{(l)} \rangle &= \sum_{l=1}^r \langle R_i^{(l)} - \alpha_i \mathcal{A}_l(P_i), R_j^{(l)} \rangle \\
 &= \sum_{l=1}^r \langle R_i^{(l)}, R_j^{(l)} \rangle - \alpha_i \sum_{l=1}^r \langle \mathcal{A}_l(P_i), R_j^{(l)} \rangle \\
 &= \sum_{l=1}^r \langle R_i^{(l)}, R_j^{(l)} \rangle - \alpha_i \sum_{l=1}^r \langle P_i, \mathcal{A}_l^*(R_j^{(l)}) \rangle \\
 &= \sum_{l=1}^r \langle R_i^{(l)}, R_j^{(l)} \rangle - \alpha_i \langle P_i, \sum_{l=1}^r \mathcal{A}_l^*(R_j^{(l)}) \rangle \\
 &= \sum_{l=1}^r \langle R_i^{(l)}, R_j^{(l)} \rangle - \frac{1}{2} \alpha_i \langle P_i + \mathcal{U}(P_i), \sum_{l=1}^r \mathcal{A}_l^*(R_j^{(l)}) \rangle
 \end{aligned}$$

$$\begin{aligned}
 &= \sum_{l=1}^r \langle R_i^{(l)}, R_j^{(l)} \rangle - \alpha_i \langle P_i, \frac{1}{2} \sum_{l=1}^r [\mathcal{A}_l^*(R_j^{(l)}) + \mathcal{U} \mathcal{A}_l^*(R_j^{(l)})] \rangle \\
 &= \sum_{l=1}^r \langle R_i^{(l)}, R_j^{(l)} \rangle - \alpha_i \langle P_i, Z_j \rangle.
 \end{aligned}$$

This completes the proof. □

Lemma 2.4. *Assume that Problem 1 is consistent. If there exists a positive integer k such that $\sum_{l=1}^r \|R_i^{(l)}\|^2 \neq 0$ for all $i = 0, 1, 2, \dots, k$, then the sequences $\{X_i\}$, $\{R_i^{(l)}\}$ and $\{P_i\}$ generated by Algorithm 2.1 satisfy*

$$\sum_{l=1}^r \langle R_i^{(l)}, R_j^{(l)} \rangle = 0 \quad \text{and} \quad \langle P_i, P_j \rangle = 0, \quad (i, j = 0, 1, 2, \dots, k, i \neq j). \quad (2.3)$$

Proof. Note that $\langle A, B \rangle = \langle B, A \rangle$ holds for arbitrary matrices A and B . We only need to prove the conclusion holds for all $0 \leq j < i \leq k$. For $k = 1$, by Lemma 2.3, we have

$$\begin{aligned}
 \sum_{l=1}^r \langle R_1^{(l)}, R_0^{(l)} \rangle &= \sum_{l=1}^r \langle R_0^{(l)}, R_0^{(l)} \rangle - \alpha_0 \langle P_0, Z_0 \rangle \\
 &= \sum_{l=1}^r \|R_0^{(l)}\|^2 - \frac{\sum_{l=1}^r \|R_0^{(l)}\|^2}{\|P_0\|^2} \langle P_0, P_0 \rangle = 0
 \end{aligned}$$

and

$$\langle P_1, P_0 \rangle = \langle Z_1 - \frac{\text{tr}(Z_1^H P_0)}{\|P_0\|^2} P_0, P_0 \rangle = \langle Z_1, P_0 \rangle - \frac{\text{tr}(Z_1^H P_0)}{\|P_0\|^2} \langle P_0, P_0 \rangle = 0.$$

Assume that $\sum_{l=1}^r \langle R_s^{(l)}, R_j^{(l)} \rangle = 0$ and $\langle P_s, P_j \rangle = 0$ for all $0 \leq j < s$, $0 < s \leq k$, we shall show that $\sum_{l=1}^r \langle R_{s+1}^{(l)}, R_j^{(l)} \rangle = 0$ and $\langle P_{s+1}, P_j \rangle = 0$ hold for all $0 \leq j < s + 1, 0 < s + 1 \leq k$.

By the hypothesis and Lemma 2.3, for the case where $0 \leq j < s$, we have

$$\begin{aligned}
 \sum_{l=1}^r \langle R_{s+1}^{(l)}, R_j^{(l)} \rangle &= \sum_{l=1}^r \langle R_s^{(l)}, R_j^{(l)} \rangle - \alpha_s \langle P_s, Z_j \rangle = -\alpha_s \langle P_s, P_j + \beta_{j-1} P_{j-1} \rangle \\
 &= -\alpha_s \langle P_s, P_j \rangle - \alpha_s \beta_{j-1} \langle P_s, P_{j-1} \rangle = 0,
 \end{aligned}$$

and

$$\begin{aligned} \langle P_{s+1}, P_j \rangle &= \langle Z_{s+1}, P_j \rangle - \beta_s \langle P_s, P_j \rangle = \langle Z_{s+1}, P_j \rangle \\ &= \frac{\sum_{l=1}^r \langle R_j^{(l)}, R_{s+1}^{(l)} \rangle - \sum_{l=1}^r \langle R_{j+1}^{(l)}, R_{s+1}^{(l)} \rangle}{\alpha_j} = 0. \end{aligned}$$

For the case $j = s$, we have

$$\begin{aligned} \sum_{l=1}^r \langle R_{s+1}^{(l)}, R_s^{(l)} \rangle &= \sum_{l=1}^r \langle R_s^{(l)}, R_s^{(l)} \rangle - \alpha_s \langle P_s, Z_s \rangle \\ &= \sum_{l=1}^r \|R_s^{(l)}\|^2 - \alpha_s \langle P_s, P_s + \beta_{s-1} P_{s-1} \rangle \\ &= \sum_{l=1}^r \|R_s^{(l)}\|^2 - \alpha_s \langle P_s, P_s \rangle \\ &= \sum_{l=1}^r \|R_s^{(l)}\|^2 - \frac{\sum_{l=1}^r \|R_s^{(l)}\|^2}{\|P_s\|^2} \|P_s\|^2 = 0, \end{aligned}$$

and

$$\begin{aligned} \langle P_{s+1}, P_s \rangle &= \langle Z_{s+1} - \beta_s P_s, P_s \rangle = \langle Z_{s+1}, P_s \rangle - \beta_s \langle P_s, P_s \rangle \\ &= \langle Z_{s+1}, P_s \rangle - \frac{\text{tr}(Z_{s+1}^H P_s)}{\|P_s\|^2} \|P_s\|^2 = 0. \end{aligned}$$

Then by the principle of induction, we complete the proof. □

Remark 2.2. Lemma 2.4 implies that when Problem I is consistent, the sequences $P_0, P_1, P_2 \dots$ are orthogonal to each other in the finite dimension matrix space \mathcal{S} . Hence there exists a positive integer $t \leq mn$ such that $P_t = 0$. According to Lemma 2.2, we have $\sum_{l=1}^r \|R_t^{(l)}\|^2 = 0$. Hence the iteration will be terminated in finite steps in the absence of roundoff errors.

Next we consider finding the least norm solution of Problem I. The following lemmas are needed for our derivation.

Lemma 2.5 [1]. *Suppose that the consistent system of linear equations $Ax = b$ has a solution $x^* \in R(A^H)$, then x^* is the unique least norm solution of the system of linear equations.*

Lemma 2.6 [8]. For $\mathcal{A} \in LC^{m \times n, p \times q}$, there exists a unique matrix $M \in C^{pq \times mn}$, such that $\text{vec}(\mathcal{A}(X)) = M \text{vec}(X)$ for all $X \in C^{m \times n}$.

According to Lemma 2.6 and the definition of conjugate operator, one can easily obtain the following corollary.

Corollary 2.1. Let \mathcal{A} and M be the same as those in Lemma 2.4, and \mathcal{A}^* be the conjugate operator of \mathcal{A} , then $\text{vec}(\mathcal{A}^*(Y)) = M^H \text{vec}(Y)$ for all $Y \in C^{p \times q}$.

Theorem 2.1. Assume that Problem I is consistent. If we choose the initial matrix $X_0 = \frac{1}{2} \sum_{l=1}^r [\mathcal{A}_l^*(H) + \mathcal{U} \mathcal{A}_l^*(H)]$, where H is an arbitrary matrix in $C^{p_l \times q_l}$, or more especially, let $X_0 = 0$, then the solution X^* obtained by Algorithm 2.1 is the least norm solution.

Proof. Consider the following matrix equations system:

$$\left\{ \begin{array}{l} \mathcal{A}_1(X) = E_1, \\ \mathcal{A}_2(X) = E_2, \\ \vdots \\ \mathcal{A}_r(X) = E_r, \\ \mathcal{A}_1(\mathcal{U}(X)) = E_1, \\ \mathcal{A}_2(\mathcal{U}(X)) = E_2, \\ \vdots \\ \mathcal{A}_r(\mathcal{U}(X)) = E_r. \end{array} \right. \quad (2.4)$$

If Problem I has a solution X , then X must be a solution of the system (2.4). Conversely, if the system (2.4) has a solution X , let $\tilde{X} = \frac{X + \mathcal{U}(X)}{2}$, then it is easy to verify that \tilde{X} is a solution of Problem I. Therefore, the solvability of Problem I is equivalent to the system (2.4).

If we choose the initial matrix $X_0 = \frac{1}{2} \sum_{l=1}^r [\mathcal{A}_l^*(H) + \mathcal{U} \mathcal{A}_l^*(H)]$, where H is an arbitrary matrix in $C^{p_l \times q_l}$, by Algorithm 2.1 and Remark 2.2, we can obtain the solution X^* of Problem I within finite iteration steps, which can be represented as $X^* = \frac{1}{2} \sum_{l=1}^r [\mathcal{A}_l^*(Y) + \mathcal{U} \mathcal{A}_l^*(Y)]$. Since the solution set of Problem I is a subset of that of the system (2.4), next we shall prove that X^* is

the least norm solution of the system (2.4), which implies that X^* is the least norm solution of Problem I.

Let M_l and N be the matrices such that $\text{vec}(\mathcal{A}_l(X)) = M_l \text{vec}(X)$ and $\text{vec}(\mathcal{U}(X)) = N \text{vec}(X)$ for all $X \in C^{m \times n}, l = 1, 2, \dots, r$. Then the system (2.4) is equivalent to the following system of linear equations:

$$\left\{ \begin{array}{l} M_1 \text{vec}(X) = \text{vec}(E_1), \\ M_2 \text{vec}(X) = \text{vec}(E_2), \\ \vdots \\ M_r \text{vec}(X) = \text{vec}(E_r), \\ M_1 N \text{vec}(X) = \text{vec}(E_1), \\ M_2 N \text{vec}(X) = \text{vec}(E_2), \\ \vdots \\ M_r N \text{vec}(X) = \text{vec}(E_r). \end{array} \right. \quad (2.5)$$

Note that

$$\begin{aligned} \text{vec}(X^*) &= \frac{1}{2} \sum_{l=1}^r [M_l^H \text{vec}(Y) + N^H M_l^H \text{vec}(Y)] \\ &\in R(M_1^H, M_2^H, \dots, M_r^H, (M_1 N)^H, (M_2 N)^H, \dots, (M_r N)^H). \end{aligned}$$

Then it follows from Lemma 2.5 that X^* is the least norm solution of the system (2.4), which is also the least norm solution of Problem I. □

3 Iterative method for solving Problem II

In this section, we consider iterative method for solving the matrix nearness Problem II. Let S denote the solution set of Problem I, for given $\hat{X} \in C^{m \times n}$ and arbitrary $X \in S$, we have

$$\begin{aligned} \|X - \hat{X}\|^2 &= \|X - \frac{\hat{X} + \mathcal{U}(\hat{X})}{2} - \frac{\hat{X} - \mathcal{U}(\hat{X})}{2}\|^2 \\ &= \|X - \frac{\hat{X} + \mathcal{U}(\hat{X})}{2}\|^2 + \|\frac{\hat{X} - \mathcal{U}(\hat{X})}{2}\|^2 \\ &\quad - 2\langle X - \frac{\hat{X} + \mathcal{U}(\hat{X})}{2}, \frac{\hat{X} - \mathcal{U}(\hat{X})}{2} \rangle. \end{aligned} \quad (3.1)$$

Note that

$$\begin{aligned} \langle X - \frac{\widehat{X} + \mathcal{U}(\widehat{X})}{2}, \frac{\widehat{X} - \mathcal{U}(\widehat{X})}{2} \rangle &= \langle \mathcal{U}(X - \frac{\widehat{X} + \mathcal{U}(\widehat{X})}{2}), \frac{\widehat{X} - \mathcal{U}(\widehat{X})}{2} \rangle \\ &= \langle X - \frac{\widehat{X} + \mathcal{U}(\widehat{X})}{2}, \frac{\mathcal{U}(\widehat{X}) - \widehat{X}}{2} \rangle \\ &= - \langle X - \frac{\widehat{X} + \mathcal{U}(\widehat{X})}{2}, \frac{\widehat{X} - \mathcal{U}(\widehat{X})}{2} \rangle, \end{aligned}$$

which implies $\langle X - \frac{\widehat{X} + \mathcal{U}(\widehat{X})}{2}, \frac{\widehat{X} - \mathcal{U}(\widehat{X})}{2} \rangle = 0$. Then (3.1) reduces to

$$\|X - \widehat{X}\|^2 = \|X - \frac{\widehat{X} + \mathcal{U}(\widehat{X})}{2}\|^2 + \|\frac{\widehat{X} - \mathcal{U}(\widehat{X})}{2}\|^2.$$

Hence $\min_{X \in \mathcal{S}} \|X - \widehat{X}\|$ is equivalent to $\min_{X \in \mathcal{S}} \|X - \frac{\widehat{X} + \mathcal{U}(\widehat{X})}{2}\|$, where

$$X - \frac{\widehat{X} + \mathcal{U}(\widehat{X})}{2} \in \mathcal{S}.$$

Let $\widetilde{X} = X - \frac{\widehat{X} + \mathcal{U}(\widehat{X})}{2}$ and $\widetilde{E}_l = E_l - \mathcal{A}_l(\frac{\widehat{X} + \mathcal{U}(\widehat{X})}{2})$. Then \widetilde{X} is a solution of the following constrained matrix equations system:

$$[\mathcal{A}_1(\widetilde{X}), \mathcal{A}_2(\widetilde{X}), \dots, \mathcal{A}_r(\widetilde{X})] = [\widetilde{E}_1, \widetilde{E}_2, \dots, \widetilde{E}_r], \quad \widetilde{X} \in \mathcal{S}.$$

Hence Problem II is equivalent to finding the least norm solution of above constrained matrix equations system.

Based on the analysis above, we summarize that once the unique least norm solution \widetilde{X}^* of above constrained consistent matrix equations system is obtained by applying Algorithm 2.1, the unique solution \overline{X} of Problem II can be obtained, where $\overline{X} = \widetilde{X}^* + \frac{\widehat{X} + \mathcal{U}(\widehat{X})}{2}$.

4 Numerical examples

In this section, we shall give some numerical examples to illustrate the efficiency of Algorithm 2.1. All the tests are performed by MATLAB 7.6 with machine precision around 10^{-16} . Because of the influence of roundoff errors, we regard a matrix X as zero matrix if $\|X\| < 1.0e - 010$.

Example 4.1. Find the least norm symmetric solution of the following matrix equations system:

$$\begin{cases} A^T X + X^T A = C \\ B X B^T = D \end{cases} \tag{4.1}$$

where

$$A = \begin{pmatrix} 1 & 4 & -3 & 2 & 6 \\ -2 & -2 & 6 & -4 & -3 \\ 0 & 8 & 0 & 0 & 12 \\ 4 & 2 & -12 & 8 & 3 \\ 1 & -10 & -3 & 2 & -15 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & 3 & -2 & -2 & -1 \\ -3 & 0 & 6 & 0 & 3 \\ 5 & -9 & -10 & 6 & -5 \\ -3 & -3 & 6 & 2 & 3 \\ 2 & -6 & -4 & 4 & -2 \end{pmatrix},$$

$$C = \begin{pmatrix} 8 & 6 & -8 & 12 & 7 \\ 6 & 4 & -10 & 10 & 5 \\ -8 & -10 & -24 & -4 & -9 \\ 12 & 10 & -4 & 16 & 11 \\ 7 & 5 & -9 & 11 & 6 \end{pmatrix}, \quad D = \begin{pmatrix} 1 & -6 & 13 & -5 & 6 \\ -6 & 36 & -78 & 30 & -36 \\ 13 & -78 & 169 & -65 & 78 \\ -5 & 30 & -65 & 25 & -30 \\ 6 & -36 & 78 & -30 & 36 \end{pmatrix}.$$

If we define $\mathcal{A}_1 : X \rightarrow A^T X + X^T A$ and $\mathcal{A}_2 : X \rightarrow B X B^T$, then we have $\mathcal{A}_1^* : Y \rightarrow AY + AY^T$ and $\mathcal{A}_2^* : Y \rightarrow B^T Y B$. Here we define $\mathcal{U} : X \rightarrow X^T$.

Let $X_0 = 0$, then by using Algorithm 2.1 and iterating 16 steps, we obtain the least norm symmetric solution of the system (4.1) as follows:

$$X_{16} = \begin{pmatrix} 0.4892 & 0.4573 & 0.5406 & 0.9735 & 0.5314 \\ 0.4573 & 1.4950 & 1.0086 & 1.4386 & 0.7785 \\ 0.5406 & 1.0086 & 0.8170 & 1.1925 & 0.7066 \\ 0.9735 & 1.4386 & 1.1925 & 1.2019 & 1.0961 \\ 0.5314 & 0.7785 & 0.7066 & 1.0961 & 0.6413 \end{pmatrix}$$

with $\|R_{16}^{(1)}\|^2 + \|R_{16}^{(2)}\|^2 = 1.4901e - 018$.

Example 4.2. Let S_E denote the solution set of the matrix equations system (4.1) with symmetric constraint, for given

$$\widehat{X} = \begin{pmatrix} 1 & 1 & 0 & 1 & 1 \\ -1 & 2 & 1 & 2 & 1 \\ 1 & 0 & 1 & 0 & -1 \\ -1 & 1 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 & -1 \end{pmatrix},$$

find $\bar{X} \in S_E$, such that $\|\bar{X} - \hat{X}\| = \min_{X \in S_E} \|X - \hat{X}\|$.

In order to find the optimal approximation solution to the matrix \hat{X} , let $\tilde{X} = X - \frac{\hat{X} + \hat{X}^T}{2}$, $\tilde{C} = C - A^T(\frac{\hat{X} + \hat{X}^T}{2}) - (\frac{\hat{X} + \hat{X}^T}{2})A$, and $\tilde{D} = D - B(\frac{\hat{X} + \hat{X}^T}{2})B^T$. By applying Algorithm 2.1 to the new matrix equations system:

$$\begin{cases} A^T \tilde{X} + \tilde{X}^T A = \tilde{C} \\ B \tilde{X} B^T = \tilde{D} \end{cases}$$

with symmetric constraint, and choosing the initial matrix $\tilde{X}_0 = 0$, we obtain its unique least norm symmetric solution as follows:

$$\tilde{X}_{17} = \begin{pmatrix} 0.2387 & -0.0927 & -0.1176 & 0.4657 & -0.2869 \\ -0.0927 & -0.1921 & 0.5102 & 0.2855 & 0.0665 \\ -0.1176 & 0.5102 & -0.2941 & 0.7672 & 0.5691 \\ 0.4657 & 0.2855 & 0.7672 & 0.5387 & -0.0493 \\ -0.2869 & 0.0665 & 0.5691 & -0.0493 & 1.6174 \end{pmatrix}$$

with $\|\tilde{R}_{17}^{(1)}\|^2 + \|\tilde{R}_{17}^{(2)}\|^2 = 1.1340e - 017$.

Then we have

$$\bar{X} = \tilde{X}_{17} + \frac{\hat{X} + \hat{X}^T}{2} = \begin{pmatrix} 1.2387 & -0.0927 & 0.3824 & 0.4657 & 0.7131 \\ -0.0927 & 1.8079 & 1.0102 & 1.7855 & 0.5665 \\ 0.3824 & 1.0102 & 0.7059 & 1.2672 & 0.5691 \\ 0.4657 & 1.7855 & 1.2672 & 1.5387 & 0.9507 \\ 0.7131 & 0.5665 & 0.5691 & 0.9507 & 0.6174 \end{pmatrix}$$

and $\|\bar{X} - \hat{X}\| = 3.8408$.

5 Conclusions

In this paper, an iterative algorithm is constructed to solve a kind of constrained matrix equations system, i.e., the matrix equations system (1.1) with the constraint (1.2), and its associated optimal approximation problem. By this algorithm, for any initial matrix X_0 satisfying the constraint (1.2), a solution X^* can be obtained in finite iteration steps in the absence of roundoff errors, and the least norm solution can be obtained by choosing a special kind of initial matrix.

In addition, using this iterative method, the optimal approximation solution \overline{X} to a given matrix \widehat{X} can be derived by first finding the least norm solution of a new corresponding constrained matrix equations system. The given numerical examples show that the proposed iterative algorithm is quite efficient.

Acknowledgement. Research supported by Shanghai Science and Technology Committee (No. 062112065) and Natural Science Foundation of Zhejiang Province (No. Y607136).

REFERENCES

- [1] A. Ben-Israel and T.N.E. Greville, *Generalized Inverse: Theory and Applications*, Springer, New York (2003).
- [2] X.W. Chang and J.S. Wang, *The symmetric solution of the matrix equations $AX + YA = C$, $AXA^T + BYB^T = C$, and $(A^T XA, B^T XB) = (C, D)$* , Linear Algebra Appl., **179** (1993), 171–189.
- [3] J.L. Chen and X.H. Chen, *Special Matrices*, Tsinghua University Press, Beijing (2001).
- [4] K.E. Chu, *Symmetric solutions of linear matrix equations by matrix decompositions*, Linear Algebra Appl., **119** (1989), 35–50.
- [5] H. Dai, *On the symmetric solutions of linear matrix equations*, Linear Algebra Appl., **131** (1990), 1–7.
- [6] M. Dehghan and M. Hajarian, *An iterative algorithm for solving a pair of matrix equations $AYB = E$, $CYD = F$ over generalized centro-symmetric matrices*, Comput. Math. Appl., **56** (2008), 3246–3260.
- [7] G.H. Golub, S. Nash and C. Vanloan, *A Hessenberg-Schur method for the problem $AX + XB = C$* , IEEE Trans. Auto. Control., **24** (1979), 909–913.
- [8] R.A. Horn and C.R. Johnson, *Topics in Matrix Analysis*, Cambridge University Press, New York (1991).
- [9] G.X. Huang, F. Yin and K. Guo, *An iterative method for the skew-symmetric solution and the optimal approximate solution of the matrix equation $AXB = C$* , J. Comput. Appl. Math., **212** (2008), 231–244.
- [10] F.L. Li, X.Y. Hu and L. Zhang, *The generalized anti-reflexive solutions for a class of matrix equations $BX = C$, $XD = E$* , Comput. Appl. Math., **27** (2008), 31–46.
- [11] S.K. Mitra, *Common solutions to a pair of linear matrix equations $A_1XB_1 = C_1$, $A_2XB_2 = C_2$* , Proc. Cambridge Philos. Soc., **74** (1973), 213–216.
- [12] X.Y. Peng, X.Y. Hu and L. Zhang, *The orthogonal-symmetric or orthogonal-anti-symmetric least-square solutions of the matrix equation*, Chinese J. Engi. Math., **23**(6) (2006), 1048–1052.

- [13] Y.X. Peng, X.Y. Hu and L. Zhang, *An iterative method for symmetric solutions and optimal approximation solution of the system of matrix equations $A_1XB_1 = C_1, A_2XB_2 = C_2$* , Appl. Math. Comput., **183** (2006), 1127–1137.
- [14] Y.X. Peng, X.Y. Hu and L. Zhang, *The reflexive and anti-reflexive solutions of the matrix equation $A^HXB = C$* , J. Comput. Appl. Math., **200** (2007), 749–760.
- [15] F.X. Piao, Q.L. Zhang and Z.F. Wang, *The solution to matrix equation $AX + X^TC = B$* , J. Fran. Inst., **344** (2007), 1056–1062.
- [16] Y.Y. Qiu, Z.Y. Zhang and J.F. Lu, *The matrix equations $AX = B, XC = D$ with $PX = sXP$ constraint*, Appl. Math. Comput., **189** (2007), 1428–1434.
- [17] X.P. Sheng and G.L. Chen, *A finite iterative method for solving a pair of linear matrix equations $(AXB, CXD) = (E, F)$* , Appl. Math. Comput., **189** (2007), 1350–1358.
- [18] M.H. Wang, X.H. Cheng and M.S. Wei, *Iterative algorithms for solving the matrix equation $AXB + CX^TD = E$* , Appl. Math. Comput., **187** (2007), 622–629.
- [19] Y.X. Yuan and H. Dai, *Generalized reflexive solutions of the matrix equation $AXB = D$ and associated optimal approximation problem*, Comput. Math. Appl., **56** (2008), 1643–1649.
- [20] J.C. Zhang, S.Z. Zhou and X.Y. Hu, *The (P, Q) generalized reflexive and anti-reflexive solutions of the matrix equation $AX = B$* , Appl. Math. Comput., **209** (2009), 254–258.