J. Appl. Math. & Informatics Vol. 26(2008), No. 5 - 6, pp. 1233 - 1245
Website: http://www.kcam.biz

AN ITERATIVE ALGORITHM FOR SOLVING THE LEAST-SQUARES PROBLEM OF MATRIX EQUATION AXB+CYD=E

KAI-JUAN SHEN*, CHUAN-HUA YOU AND YU-XIA DU

ABSTRACT. In this paper, an iterative method is proposed to solve the least-squares problem of matrix equation $AXB + CYD = E$ over unknown matrix pair $[X, Y]$. By this iterative method, for any initial matrix pair $[X_1, Y_1]$, a solution pair or the least-norm least-squares solution pair of which can be obtained within finite iterative steps in the absence of roundoff errors. In addition, we also consider the optimal approximation problem for the given matrix pair $[X_0, Y_0]$ in Frobenius norm. Given numerical examples show that the algorithm is efficient.

AMS Mathematics Subject Classification: 65F10
Key words and phrases: Matrix equation; iterative algorithm; least-squares problem; least-norm least-squares solution; optimal approximation.

1. Introduction

We firstly introduce some symbols. Denoted by $R^{m\times n}$ be the set of all $m \times n$ real matrices. The superscripts T and + respect the transpose and Moore-Penrose generalized inverse of matrices. In space $R^{m\times n}$, we define inner product by $< A, B >= trace(B^T A)$ for all $A, B \in R^{m\times n}$, which generates the Frobenius norm, i.e., $\|A\| = \sqrt{< A, A >}$. Let $R(\cdot)$ be the column space of a matrix, and $vec(.)$ be the vec operator, that is, $vec(A) = (a_1^T, a_2^T, ..., a_n^T)^T$, where $A = (a_1, a_2, ..., a_n) \in R^{m\times n}$, $a_i \in R^m$, i=1,2,...,n. $A \otimes B$ stands for the kronecker product[1] of matrices A and B. Furthermore, we say matrices F and G is orthogonal each other, if $trace(G^T F) = 0$.

The well-known linear matrix equation

$$AXB + CYD = E \tag{1}$$

Received June 25, 2007. Revised April 14, 2008. Accepted April 17, 2008.
*Corresponding author.

has been widely considered over unknown matrices $X$and Y, such as references [2-5]. In these literatures, the solvability conditions are obtained by using g-inverse or matrices decomposition (i.e., singular value decomposition (SVD), generalized SVD, or canonical correlation decomposition (CCD)). However, the matrices A, B, C, D and E, are experimentally occurring in practice, which may not satisfy these solvability conditions because of the influence of experimental errors etc. Therefore, the matrix equation $AXB+CYD = E$ is always inconsistent, we need consider its least-squares problem. The least-squares problem of matrix equation (1), which has been studied in [6-7] by matrix decomposition, can be expressed as follows.

Problem I. it Given matrices $A \in R^{p\times n_1}$, $B \in R^{n_2\times q}$, $C \in R^{p\times m_1}$, $D \in R^{m_2\times q}$, $E \in R^{p\times q}$, find $X \in R^{n_1\times n_2}$,$Y \in R^{m_1\times m_2}$, such that

$$\|AXB + CYD - E\| = \min \quad for\, X \in R^{n_1\times n_2}, Y \in R^{m_1\times m_2}.$$

In addition, the problem that finds a nearest matrix in the least-squares solution set of a matrix equation to given matrix is so-called matrix nearness problem (see [8] for details). The matrix nearness problem associated with the matrix equation (1) can be described as follows.

Problem II. *For given matrices $X_0 \in R^{n_1\times n_2}, Y_0 \in R^{m_1\times m_2}$, find $[\hat{X}, \hat{Y}] \in S_E$, satisfying*

$$\|\hat{X} - X_0\|^2 + \|\hat{Y} - Y_0\|^2 = \min_{[X,Y]\in S_E} \{\|X - X_0\|^2 + \|Y - Y_0\|^2\},$$

where S_E is the solution set of Problem I.

In fact, Problem II is to find the least-norm solution of Problem I, when $X_0 = 0$, $Y_0 = 0$.

The methods using matrix decomposition for matrix equation problem are feasible but complicated. Peng [9] has constructed an iterative algorithm to find the symmetric solution of matrix equation $AXB = C$, the method make it easier to solve matrix equation problems. Whereafter, Peng [10] has obtained the symmetric least-squares solution of matrix equation $AXB = C$ by similar iterative method. Meanwhile, according to the fundamental idea of the classical conjugate methods, using the orthogonal direction method, Deng and Bai [11] has obtained the Hermitian minimum F-norm solution of matrix equation $AXB = C$ and $(AX, XB) = (C, D)$. Furthermore, by constructing iterative method, the real solution pair of matrix equation $AXB + CYD = E$ has been derived by Peng [12]. They have showed that, For any initial matrix, the solutions or least-norm solution can be obtained within finite iterative steps.

The matrix nearness Problem II occurs frequently in experimental design. Here the matrix pair $[X, Y]$ may be obtained from experiments, but it may not be a solution pair of Problem I. The best estimate $[\hat{X}, \hat{Y}]$ is the matrix pair not only is a solution pair of Problem I, but also is the best approximation of the matrix pair $[X_0, Y_0]$ (see [7, 13, 14] for more details).

However, there is no iterative algorithm constructed to solve the least squares problem of matrix equation $AXB + CYD = E$. In this paper, we establish an iterative method for Problem I. Meanwhile, the matrix nearness Problem II is also considered. For any initial matrix pair $[X_1, Y_1]$, we will show that the least-squares solution pair and the least-norm least-squares solution pair of matrix equation (1) can be obtained within finite iterative steps by the iteration method.

This paper is organized as follows: In section 2, we will give the iterative method for Problem I. In section 3, for given matrices X_0, Y_0, we will show that the unique solution of Problem II can be found by the least-norm least-squares solution pair $[\widetilde{X}^*, \widetilde{Y}^*]$ of the new least-squares problem of matrix equation $A\widetilde{X}B+C\widetilde{Y}D = \widetilde{E}$, where $\widetilde{X} = X-X_0, \widetilde{Y} = Y-Y_0$, and $\widetilde{E} = E-AX_0B-CY_0D$. Finally, we will offer some numerical examples to illustrate our results.

2. Iterative algorithm for solving Problem I

In this section, we will give the iterative algorithm for solving Problem I, meanwhile, we will show that the iterative method is feasible. As we all know that, the least-squares problem of matrix equation can be transformed equally into finding the solutions of its normal matrix equation, which is always consistent. Therefore, we have the following assertion.

Lemma 1. *The solvability of Problem I is equivalent to that of the linear matrix equations*

$$\begin{cases} A^TAXBB^T + A^TCYDB^T = A^TEB^T, \\ C^TAXBD^T + C^TCYDD^T = C^TED^T. \end{cases} \tag{2}$$

Now, according to matrix equations (2), the iterative algorithm for Problem I can be expressed as follows:

Algorithm 1

Step 1 : Given matrices $A \in R^{p\times n_1}$, $B \in R^{n_2\times q}$, $C \in R^{p\times m_1}$, $D \in R^{m_2\times q}$, $E \in R^{p\times q}$ and $X_1 \in R^{n_1\times n_2}$, $Y_1 \in R^{m_1\times m_2}$;

Step 2 : Compute

$$M_1 = A^TEB^T - A^TAX_1BB^T - A^TCY_1DB^T,$$

$$N_1 = C^TED^T - C^TAX_1BD^T - C^TCY_1DD^T,$$

$$R_1 = \begin{pmatrix} M_1 & 0 \\ 0 & N_1 \end{pmatrix},$$

$$P_1 = A^TAM_1BB^T + A^TCN_1DB^T,$$

$$Q_1 = C^TAM_1BD^T + C^TCN_1DD^T,$$

$$k := 1;$$

Step 3 : Compute

$$X_k = X_{k-1} + \frac{\| R_{k-1} \|^2}{\| P_{k-1} \|^2 + \| Q_{k-1} \|^2} P_{k-1},$$

$$Y_k = Y_{k-1} + \frac{\| R_{k-1} \|^2}{\| P_{k-1} \|^2 + \| Q_{k-1} \|^2} Q_{k-1};$$

Step 4 : Compute

$$M_k = A^T E B^T - A^T A X_k B B^T - A^T C Y_k D B^T,$$

$$N_k = C^T E D^T - C^T A X_k B D^T - C^T C Y_k D D^T,$$

$$R_k = \begin{pmatrix} M_k & 0 \\ 0 & N_k \end{pmatrix},$$

$$P_k = A^T A M_k B B^T + A^T C N_k D B^T + \frac{\| R_k \|^2}{\| R_{k-1} \|^2} P_{k-1},$$

$$Q_k = C^T A M_k B D^T + C^T C N_k D D^T + \frac{\| R_k \|^2}{\| R_{k-1} \|^2} Q_{k-1};$$

Step 5 : If $R_k = 0$, that is, $[X_k, Y_k]$ is a solution pair of Problem I, stop iteration; Otherwise, go to step 3.

Next, we will give some lemmas to analyze the properties of Algorithm 1, and show that the solution can be obtained within finite steps.

Lemma 2. *Suppose that* $[\bar{X}, \bar{Y}]$ *is an arbitrary solution pair of Problem I. Then*

$$< P_k, \bar{X} - X_k > + < Q_k, \bar{Y} - Y_k > = \| R_k \|^2, \; k = 1, 2, ... \tag{3}$$

Proof. We will complete the proof of (3) by the principle of induction.

Algorithm 1 implies that $\| R_k \|^2 = \| M_k \|^2 + \| N_k \|^2$, then, we have

$$\begin{aligned} & < P_1, \bar{X} - X_1 > + < Q_1, \bar{Y} - Y_1 > \\ = & < A^T A M_1 B B^T + A^T C N_1 D B^T, \bar{X} - X_1 > \\ & + < C^T A M_1 B D^T + C^T C N_1 D D^T, \bar{Y} - Y_1 > \\ = & < M_1, A^T A(\bar{X} - X_1) B B^T + A^T C(\bar{Y} - Y_1) D B^T > \\ & + < N_1, C^T A(\bar{X} - X_1) B D^T + C^T C(\bar{Y} - Y_1) D D^T > \\ = & < M_1, A^T E B^T - A^T A X_1 B B^T - A^T C Y_1 D B^T > \\ & + < N_1, C^T E D^T - C^T A X_1 B D^T - C^T C Y_1 D D^T > \\ = & \| M_1 \|^2 + \| N_1 \|^2 \\ = & \| R_1 \|^2 . \end{aligned}$$

that is, when $k = 1$, the conclusion (3) holds.

Assume that (3) holds for $k = s$, i.e., $< P_s, \bar{X} - X_s > + < Q_s, \bar{Y} - Y_s > = \| R_s \|^2$, then

$$
\begin{aligned}
&< P_{s+1}, \bar{X} - X_{s+1} > + < Q_{s+1}, \bar{Y} - Y_{s+1} > \\
&= \left\langle A^TAM_{s+1}BB^T + A^TCN_{s+1}DB^T + \frac{\| R_{s+1} \|^2}{\| R_s \|^2} P_s, \bar{X} - X_{s+1} \right\rangle \\
&\quad + \left\langle C^TAM_{s+1}BD^T + C^TCN_{s+1}DD^T + \frac{\| R_{s+1} \|^2}{\| R_s \|^2} Q_s, \bar{Y} - Y_{s+1} \right\rangle \\
&=< A^TAM_{s+1}BB^T + A^TCN_{s+1}DB^T, \bar{X} - X_{s+1} > \\
&\quad + < C^TAM_{s+1}BD^T + C^TCN_{s+1}DD^T, \bar{Y} - Y_{s+1} > \\
&\quad + \frac{\| R_{s+1} \|^2}{\| R_s \|^2} < P_s, \bar{X} - X_{s+1} > + \frac{\| R_{s+1} \|^2}{\| R_s \|^2} < Q_s, \bar{Y} - Y_{s+1} > \\
&=\| M_{s+1} \|^2 + \| N_{s+1} \|^2 + \frac{\| R_{s+1} \|^2}{\| R_s \|^2} < P_s, \bar{X} - X_s - \frac{\| R_s \|^2}{\| P_s \|^2 + \| Q_s \|^2} P_s > \\
&\quad + \frac{\| R_{s+1} \|^2}{\| R_s \|^2} < Q_s, \bar{Y} - Y_s - \frac{\| R_s \|^2}{\| P_s \|^2 + \| Q_s \|^2} Q_s > \\
&=\| R_{s+1} \|^2 + \frac{\| R_{s+1} \|^2}{\| R_s \|^2} [< P_s, \bar{X} - X_s > + < Q_s, \bar{Y} - Y_s > - \| R_s \|^2] \\
&=\| R_{s+1} \|^2 .
\end{aligned}
$$

Hence, for all positive integer number k, (3) holds. □

Remark 1. Lemma 2 implies that if $R_i \neq 0$, then $\| P_i \|^2 + \| Q_i \|^2 \neq 0$. This result implies that if $R_i \neq 0$, then Algorithm 1 can't be terminated.

Lemma 3. *For the sequences $\{R_i\}, \{P_i\}, \{Q_i\}$ generated by Algorithm 1, we have that*

$$
< R_i, R_j >= 0, < P_i, P_j > + < Q_i, Q_j >= 0, \ \ i, j = 1, 2, ..., k (k \geq 2), i \neq j. \tag{4}
$$

Proof. From the iterative Algorithm 1, we can obtain

$$
\begin{aligned}
< R_i, R_j >&= trace \left[\begin{pmatrix} M_j^T M_i & 0 \\ 0 & N_j^T N_i \end{pmatrix} \right] \\
&= trace(M_j^T M_i) + trace(N_j^T N_i) \\
&=< M_i, M_j > + < N_i, N_j >.
\end{aligned}
$$

Since the property of trace $tr(R_i^T R_j) = tr(R_j^T R_i)$, it is enough to prove the conclusion for $i > j$.

For $k = 2$, by Lemma 2 and Algorithm 1, it follows that

$$
\begin{aligned}
&< R_2, R_1 > \\
&=< M_2, M_1 > + < N_2, N_1 > \\
&=< M_1 - \frac{\| R_1 \|^2}{\| P_1 \|^2 + \| Q_1 \|^2} (A^TAP_1BB^T + A^TCQ_1DB^T), M_1 >
\end{aligned}
$$

$$
\begin{aligned}
&+ < N_1 - \frac{\| R_1 \|^2}{\| P_1 \|^2 + \| Q_1 \|^2}(C^T A P_1 B D^T + C^T C Q_1 D D^T), N_1 > \\
=& \| M_1 \|^2 + \| N_1 \|^2 - \frac{\| R_1 \|^2}{\| P_1 \|^2 + \| Q_1 \|^2} < P_1, A^T A M_1 B B^T + A^T C N_1 D B^T > \\
&- \frac{\| R_1 \|^2}{\| P_1 \|^2 + \| Q_1 \|^2} < Q_1, C^T A M_1 B D^T + C^T C N_1 D D^T > \\
=& \| M_1 \|^2 + \| N_1 \|^2 - \frac{\| R_1 \|^2}{\| P_1 \|^2 + \| Q_1 \|^2}(\| P_1 \|^2 + \| Q_1 \|^2) \\
=& \| M_1 \|^2 + \| N_1 \|^2 - \| R_1 \|^2 = 0, \qquad (5)
\end{aligned}
$$

and

$$
\begin{aligned}
& < P_2, P_1 > + < Q_2, Q_1 > \\
=& < A^T A M_2 B B^T + A^T C N_2 D B^T + \frac{\| R_2 \|^2}{\| R_1 \|^2} P_1, P_1 > \\
& + < C^T A M_2 B D^T + C^T C N_2 D D^T + \frac{\| R_2 \|^2}{\| R_1 \|^2} Q_1, Q_1 > \\
=& < M_2, A^T A P_1 B B^T + A^T C Q_1 D B^T > + < N_2, C^T A P_1 B D^T \\
& + C^T C Q_1 D D^T > + \frac{\| R_2 \|^2}{\| R_1 \|^2}(\| P_1 \|^2 + \| Q_1 \|^2) \\
=& - \frac{\| P_1 \|^2 + \| Q_1 \|^2}{\| R_1 \|^2}\Big[< M_2, M_1 - M_2 > + < N_2, N_1 - N_2 > \Big] \\
& + \frac{\| R_2 \|^2}{\| R_1 \|^2}(\| P_1 \|^2 + \| Q_1 \|^2) \\
=& - \frac{\| P_1 \|^2 + \| Q_1 \|^2}{\| R_1 \|^2}(\| M_2 \|^2 + \| N_2 \|^2) + \frac{\| R_2 \|^2}{\| R_1 \|^2}(\| P_1 \|^2 + \| Q_1 \|^2) \\
=& 0. \qquad (6)
\end{aligned}
$$

Assume that (4) holds for $k = s$, similar to the proofs of (5) and (6), we get $< R_{s+1}, R_s >= 0$ and $< P_{s+1}, P_s > + < Q_{s+1}, Q_s >= 0$. Therefore, $< R_{i+1}, R_i >= 0$ and $< P_{i+1}, P_i > + < Q_{i+1}, Q_i >= 0$.

Furthermore, when $j = 1$, noting that $< R_s, R_1 >= 0$, $< P_s, P_1 > + < Q_s, Q_1) = 0$, we have that

$$
\begin{aligned}
& < R_{s+1}, R_1 > \\
=& < M_{s+1}, M_1 > + < N_{s+1}, N_1 > \\
=& < M_s - \frac{\| R_s \|^2}{\| P_s \|^2 + \| Q_s \|^2}(A^T A P_s B B^T + A^T C Q_s D B^T), M_1 > \\
& + < N_s - \frac{\| R_s \|^2}{\| P_s \|^2 + \| Q_s \|^2}(C^T A P_s B D^T + C^T C Q_s D D^T), N_1 >
\end{aligned}
$$

$$
\begin{aligned}
&= -\frac{\| R_s \|^2}{\| P_s \|^2 + \| Q_s \|^2} < P_s, A^T A M_1 B B^T + A^T C N_1 D B^T > \\
&\quad -\frac{\| R_s \|^2}{\| P_s \|^2 + \| Q_s \|^2} < Q_s, C^T A M_1 B D^T + C^T C N_1 D D^T > \\
&= -\frac{\| R_s \|^2}{\| P_s \|^2 + \| Q_s \|^2} \Big(< P_s, P_1 > + < Q_s, Q_1 > \Big) = 0.
\end{aligned}
$$

When $j = 2, ..., s-1$, according to the assumptions, we have

$$
\begin{aligned}
& < R_{s+1}, R_j > \\
= & < M_{s+1}, M_j > + < N_{s+1}, N_j > \\
= & < M_s - \frac{\| R_s \|^2}{\| P_s \|^2 + \| Q_s \|^2}(A^T A P_s B B^T + A^T C Q_s D B^T), M_j > \\
& + < N_s - \frac{\| R_s \|^2}{\| P_s \|^2 + \| Q_s \|^2}(C^T A P_s B D^T + C^T C Q_s D D^T), N_j > \\
= & < M_s, M_j > + < N_s, N_j > - \frac{\| R_s \|^2}{\| P_s \|^2 + \| Q_s \|^2} \Big[< P_s, A^T A M_j B B^T \\
& + A^T C N_j D B^T > + < Q_s, C^T A M_j B D^T + C^T C N_j D D^T > \Big] \\
= & -\frac{\| R_s \|^2}{\| P_s \|^2 + \| Q_s \|^2} \Big[< P_s, P_j - \frac{\| R_j \|^2}{\| R_{j-1} \|^2} P_{j-1} > + < Q_s, Q_j \\
& - \frac{\| R_j \|^2}{\| R_{j-1} \|^2} Q_{j-1} > \Big] \\
= & -\frac{\| R_s \|^2}{\| P_s \|^2 + \| Q_s \|^2} \Big(< P_s, P_j > + < Q_s, Q_j > \Big) \\
& + \frac{\| R_s \|^2}{\| P_s \|^2 + \| Q_s \|^2} \times \frac{\| R_j \|^2}{\| R_{j-1} \|^2} \Big(< P_s, P_{j-1} > + < Q_s, Q_{j-1} > \Big) = 0,
\end{aligned}
$$

and

$$
\begin{aligned}
& < P_{s+1}, P_j > + < Q_{s+1}, Q_j > \\
& = < A^T A M_{s+1} B B^T + A^T C N_{s+1} D B^T + \frac{\| R_{s+1} \|^2}{\| R_s \|^2} P_s, P_j > \\
& + < C^T A M_{s+1} B D^T + C^T C N_{s+1} D D^T + \frac{\| R_{s+1} \|^2}{\| R_s \|^2} Q_s, Q_j > \\
& = \frac{\| R_{s+1} \|^2}{\| R_s \|^2} (< P_s, P_j > + < Q_s, Q_j >) \\
& + \frac{\| P_j \|^2 + \| Q_j \|^2}{\| R_j \|^2} (< M_{s+1}, M_j - M_{j+1} > + < N_{s+1}, N_j - N_{j+1} >) = 0.
\end{aligned}
$$

Hence, (4) holds for $k = s + 1$. We complete the proof by the induction. □

Remark 2. From Lemma 3, we can see that, R_i $\Big(i = 1, 2, \ldots, (n_1 + n_2)(m_1 + m_2)\Big)$ are orthogonal each other, that is, they consists of an orthogonal basis of space $R^{(n_1+n_2)\times(m_1+m_2)}$. Therefore, it is certain that there exists a positive integer number k such that $R_k = 0$ in the iterative process.

The following lemma from [8] is a directly use for stating our mainly results.

Lemma 4. *Suppose that the consistent linear system of $My = b$ has a solution $y_0 \in R(M^T)$, then y_0 is the least-norm solution of the system of linear equations.*

According to the properties of Kronecker product, then the systems of matrices equations(2) is equivalent to the linear equations

$$\begin{pmatrix} BB^T \otimes AA^T & BD^T \otimes A^TC \\ DB^T \otimes C^TA & DD^T \otimes CC^T \end{pmatrix}\begin{pmatrix} vec(X) \\ vec(Y) \end{pmatrix}=\begin{pmatrix} B^T \otimes A \\ D^T \otimes C \end{pmatrix} vec(E).$$

Noting that

$$\begin{aligned} &\begin{pmatrix} vec(A^TAHBB^T + A^TC\widetilde{H}DB^T) \\ vec(C^TAHBD^T + C^TC\widetilde{H}DD^T) \end{pmatrix} \\ &= \begin{pmatrix} BB^T \otimes AA^T & BD^T \otimes A^TC \\ DB^T \otimes C^TA & DD^T \otimes CC^T \end{pmatrix}\begin{pmatrix} vec(H) \\ vec(\widetilde{H}) \end{pmatrix} \\ &= \begin{pmatrix} BB^T \otimes AA^T & BD^T \otimes A^TC \\ DB^T \otimes C^TA & DD^T \otimes CC^T \end{pmatrix}^T\begin{pmatrix} vec(H) \\ vec(\widetilde{H}) \end{pmatrix} \\ &\in R\left(\begin{pmatrix} BB^T \otimes AA^T & BD^T \otimes A^TC \\ DB^T \otimes C^TA & DD^T \otimes CC^T \end{pmatrix}^T\right). \end{aligned}$$

We know that, if we take initial matrices $X_1 = A^TAHBB^T + A^TC\widetilde{H}DB^T$, $Y_1 = C^TAHBD^T + C^TC\widetilde{H}DD^T$, where $H, \widetilde{H}$ are arbitrary, then all X_k and Y_k, generated by Algorithm 1, satisfy that

$$\begin{pmatrix} vec(X_k) \\ vec(Y_k) \end{pmatrix} \in R\left(\begin{pmatrix} BB^T \otimes AA^T & BD^T \otimes A^TC \\ DB^T \otimes C^TA & DD^T \otimes CC^T \end{pmatrix}^T\right).$$

Hence, from Lemma 4, we have that $[X^*, Y^*]$ generated by the iterative method, is the solution pair of Problem I, then it is the least-norm solution pair of which.

Above conclusions on the solution of Problem I can be stated as following theorem, and its proof is omitted.

Theorem 1. *The least-squares Problem I is always consistent, then for arbitrary initial matrix pair $[X_1, Y_1]$, the sequence $[X_k, Y_k]$ generated by Algorithm 1, converges to its solution pair within at most $(n_1 + n_2)(m_1 + m_2) + 1$ iteration steps. Furthermore, if we choose the initial matrices $X_1 = A^TAHBB^T +$*

$A^TC\widetilde{H}DB^T, Y_1 = C^TAHBD^T + C^TC\widetilde{H}DD^T$ ($H, \widetilde{H}$ are arbitrary), or especially, let $X_1 = 0 \in R^{n_1\times n_2}, Y_1 = 0 \in R^{m_1\times m_2}$, then the solution pair $[X^*, Y^*]$ obtained by Algorithm 1 is the least-norm solution pair of Problem I.

3. The solution of problem II

For the matrix nearness Problem II, there certainly exists an unique solution since the solution set of Problem I is a nonempty closed convex cone, then the linear matrix equations (2) is equivalent to the linear matrix equations

$$\begin{cases} A^TA(X-X_0)BB^T + A^TC(Y-Y_0)DB^T = A^T(E-AX_0B-CY_0D)B^T, \\ C^TA(X-X_0)BD^T + C^TC(Y-Y_0)DD^T = C^T(E-AX_0B-CY_0D)D^T. \end{cases}$$

which is the normal matrix equation of $A\widetilde{X}B + C\widetilde{Y}D = \widetilde{E}$, here $\widetilde{X} = X - X_0, \widetilde{Y} = Y - Y_0, \widetilde{E} = E - AX_0B - CY_0D$, then finding the solution of the Problem II is equivalent to find the least-norm solution pair of the matrix equations

$$\begin{cases} A^TA\widetilde{X}BB^T + A^TC\widetilde{Y}DB^T = A^T\widetilde{E}B^T, \\ C^TA\widetilde{X}BD^T + C^TC\widetilde{Y}DD^T = C^T\widetilde{E}D^T. \end{cases} \tag{7}$$

From Theorem 1, by the Algorithm 1, if the initial iterative matrices $\widetilde{X}_1 = A^TAHBB^T + A^TC\widetilde{H}DB^T, \widetilde{Y}_1 = C^TAHBD^T + C^TC\widetilde{H}DD^T(H, \widetilde{H}$ are arbitrary), or especially, let $\widetilde{X}_1 = 0 \in R^{n_1\times n_2}, \widetilde{Y}_1 = 0 \in R^{m_1\times m_2}$, we can obtain the unique least-norm solution pair $[\widetilde{X}^*, \widetilde{Y}^*]$ of the matrix equations (7), hence the solution of Problem II can be obtained by $\hat{X} = \widetilde{X}^* + X_0, \hat{Y} = \widetilde{Y}^* + Y_0$.

4. Numerical examples

In this section, will give some numerical examples to illustrate our results.

Let matrices A, B, C, D, E as follows, $A = \begin{pmatrix} 3 & 0 & -4 & 0 & 3 \\ 0 & -2 & 9 & 0 & -5 \\ -1 & 6 & 2 & 0 & 0 \\ 0 & 0 & -5 & 6 & -8 \\ 0 & 0 & 3 & 0 & 4 \\ 2 & 5 & 7 & 0 & -4 \end{pmatrix}$,

$$B = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}, D = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix},$$

$$C = \begin{pmatrix} 1 & -2 & 1 & 0 \\ -5 & 4 & -5 & -2 \\ -3 & 5 & -3 & 5 \\ 0 & -7 & 0 & -1 \\ 2 & 7 & 2 & 2 \\ -6 & 9 & -6 & -1 \end{pmatrix}, \quad E = \begin{pmatrix} 2 & -1 & 4 & 6 & 4 \\ 3 & 8 & -2 & -1 & 2 \\ -2 & 7 & 3 & 2 & 0 \\ 8 & 0 & 3 & 1 & 3 \\ 0 & 2 & 2 & 1 & 2 \\ 9 & 0 & 3 & 0 & 2 \end{pmatrix}.$$

First, we will find the general solution pair and the least-norm solution pair of Problem I. However, because of the influence of the errors of calculation, R_i is usually unequal to zero in the iterative process , for arbitrary small enough positive number ε, e.g., $\varepsilon = 1.0e - 010$, whenever $\| R_k \| < \varepsilon$, we stop the iteration, and $[X_k, Y_k]$ is regarded as a solution pair of Problem I.

If arbitrary initial iterative matrix pair $[X_1, Y_1]$ is

$$X_1 = \begin{pmatrix} 6 & 0 & 3 & -1 & 2 \\ -1 & 8 & -3 & 5 & 0 \\ 0 & -3 & -2 & 0 & 8 \\ -1 & 6 & 0 & 1 & -1 \\ 2 & 0 & 0 & -2 & 3 \end{pmatrix}, Y_1 = \begin{pmatrix} 4 & 6 & -3 & -4 & 6 \\ -1 & 8 & 3 & 2 & 0 \\ 0 & 7 & 4 & 1 & 2 \\ -3 & 2 & 0 & 9 & 0 \end{pmatrix},$$

by Algorithm 1, we have that

$$X_9 = \begin{pmatrix} 2.5615 & 1.0286 & 1.0403 & 3.6974 & 3.9065 \\ 1.2904 & -0.2463 & -1.5594 & 0.0244 & 2.7556 \\ 1.2736 & 2.0998 & -2.1307 & 1.0392 & 4.5622 \\ 0.3037 & 1.1790 & -0.5670 & -1.3276 & 1.1474 \\ -1.1793 & 0.5853 & -1.1278 & -1.4348 & 0.2333 \end{pmatrix},$$

$$Y_9 = \begin{pmatrix} 3.0665 & -0.7079 & -3.9220 & 2.4346 & 3.8117 \\ -0.4075 & -1.4866 & 1.7081 & -0.7498 & -3.1029 \\ -0.9335 & 0.2921 & 3.0780 & -0.5654 & -0.1883 \\ -0.2583 & 2.2985 & 1.3171 & 2.5659 & 0.9266 \end{pmatrix},$$

In this case, $\| X_9 \|^2 + \| Y_9 \|^2 = 170.4124$, $\| R_9 \| = 2.2845e - 011 < \epsilon$.

Hence, we obtain a general solution pair $[X_9, Y_9]$ of Problem I .

Moreover, if the initial matrices $X_1 = A^T AHBB^T + A^T C\widetilde{H}DB^T$, $Y_1 = C^T AHBD^T + C^T C\widetilde{H}DD^T$, let

$$H = \begin{pmatrix} 2 & 1 & 0 & -1 & 3 \\ 0 & 2 & 9 & 3 & 0 \\ -3 & 5 & -2 & 7 & 1 \\ 4 & 0 & 2 & -1 & 3 \\ 6 & 1 & 0 & -2 & 2 \end{pmatrix}, \widetilde{H} = \begin{pmatrix} 1 & 0 & -3 & 2 & 4 \\ 0 & 4 & -1 & -3 & 1 \\ 2 & -1 & 5 & 0 & 3 \\ 6 & 0 & 3 & 2 & 1 \end{pmatrix},$$

then, by the iterative method, we have

$$X_9 = \begin{pmatrix} 1.1707 & 0.4166 & 1.0397 & 1.4249 & 1.3448 \\ 0.1675 & -1.2378 & 0.0028 & -0.5554 & -0.6638 \\ -0.3424 & 0.8080 & -0.3068 & -0.1361 & 0.0080 \\ 0.9334 & 1.0033 & 0.8562 & 0.8437 & 1.0770 \\ -0.5640 & 0.4457 & 0.1621 & 0.6061 & 0.2516 \end{pmatrix}, \quad (8)$$

$$Y_9 = \begin{pmatrix} 0.0675 & -0.7355 & -0.1460 & -0.4756 & -0.2674 \\ 0.6623 & -0.5581 & 0.2703 & -0.1568 & 0.1113 \\ 0.0675 & -0.7355 & -0.1460 & -0.4756 & -0.2674 \\ -0.8112 & 2.3209 & 0.4818 & 0.9919 & 0.6302 \end{pmatrix}, \quad (9)$$

and $\| X_9 \|^2 + \| Y_9 \|^2 = 25.3593$, $\| R_9 \| = 2.1856e - 011 < \epsilon$.

According to Theorem 1, the iterative solution pair $[X_9, Y_9]$ in (8)-(9) is the least-norm least squares solution of Problem I. Especially, if $X_1 = 0 \in R^{5\times5}, Y_1 = 0 \in R^{4\times5}$, by Algorithm 1, we can also find the least-norm least-squares solution of Problem I as (8) and (9).

In addition, suppose that the given matrix pair $[X_0, Y_0]$ are

$$X_0 = \begin{pmatrix} 6 & -2 & 0 & 3 & 2 \\ 4 & 1 & -2 & 5 & -1 \\ 0 & -4 & 1 & 7 & 6 \\ 3 & -1 & 8 & -5 & 3 \\ 2 & 0 & 9 & 4 & -8 \end{pmatrix}, Y_0 = \begin{pmatrix} 4 & 0 & -8 & 1 & -2 \\ -2 & 6 & 0 & 3 & 5 \\ 7 & 2 & 4 & 6 & -1 \\ 9 & 3 & 0 & -9 & 4 \end{pmatrix},$$

let $\widetilde{X} = X - X_0, \widetilde{Y} = Y - Y_0, \widetilde{E} = E - AX_0B - CY_0D$, then, by the Algorithm 1 and iteration 9 steps, the least-norm solution pair $[\widetilde{X}^*, \widetilde{Y}^*]$ of matrix equations (7) is

$$\widetilde{X}^* = \widetilde{X}_9 = \begin{pmatrix} -0.8837 & 1.8123 & -2.8077 & 0.3850 & 1.2355 \\ -0.8581 & -4.7108 & 3.1170 & -1.7986 & 0.0525 \\ 3.9957 & 1.7881 & -0.8497 & -2.3199 & -5.9086 \\ -4.0451 & 0.4695 & -1.5596 & 6.9361 & -4.4270 \\ -4.4837 & -0.9251 & -3.6530 & -2.4667 & 5.9214 \end{pmatrix},$$

$$\widetilde{Y}^* = \widetilde{Y}_9 = \begin{pmatrix} -2.6358 & -2.5208 & -0.1628 & -2.1830 & 2.2704 \\ -0.1785 & -4.2662 & -0.6545 & -6.6657 & -4.6771 \\ -2.6358 & -2.5208 & -0.1628 & -2.1830 & 2.2704 \\ -8.1298 & 0.1414 & -3.3029 & 9.6093 & -1.6509 \end{pmatrix},$$

and $\| \widetilde{R}_9 \| = 4.4103e - 011 < \epsilon$.

Therefore, the solution pair $[\hat{X}, \hat{Y}]$ of Problem II can be obtained by

$$\hat{X} = \widetilde{X}^* + X_0 = \begin{pmatrix} 5.1163 & -0.1877 & -2.8077 & 3.3850 & 3.2355 \\ 3.1419 & -3.7108 & 1.1170 & 3.2014 & -0.9475 \\ 3.9957 & -2.2119 & 0.1503 & 4.6801 & 0.0914 \\ -1.0451 & -0.5305 & 6.4404 & 1.9361 & -1.4270 \\ -2.4837 & -0.9251 & 5.3470 & 1.5333 & -2.0786 \end{pmatrix},$$

$$\hat{Y} = \widetilde{Y}^* + Y_0 = \begin{pmatrix} 1.3642 & -2.5208 & -8.1628 & -1.1830 & 0.2704 \\ -2.1785 & 1.7338 & -0.6545 & -3.6657 & 0.3229 \\ 4.3642 & -0.5208 & 3.8372 & 3.8170 & 1.2704 \\ 0.8702 & 3.1414 & -3.3029 & 0.6093 & 2.3491 \end{pmatrix}.$$

5. Conclusions

In this paper, we first introduce an iterative method, that is, Algorithm 1 for finding the least-squares solution of Problem I with unknown matrices X and Y, the sequence $[X_k, Y_k]$ generated by Algorithm 1 converges to its a solution pair within at most $(n_1+n_2)(m_1+m_2)+1$ iteration steps in the absence of roundoff errors. We also prove that if let the initial matrices $X_1 = A^TAHBB^T + A^TC\widetilde{H}DB^T, Y_1 = C^TAHBD^T + C^TC\widetilde{H}DD^T$ ($H, \widetilde{H}$ are arbitrary), or especially, let $X_1 = 0 \in R^{n_1\times n_2}, Y_1 = 0 \in R^{m_1\times m_2}$, then the solution pair $[X^*, Y^*]$ obtained by Algorithm 1 is the least-norm least-squares solution pair of Problem I. Moreover, the solution of Problem II is represented by the iterative method. At last, the given examples tested in MATLAB 6.5 verify that the iterative algorithm is feasible.

References

1. R. Horn, C. Johnson, *Topics in matrix analysis* (Vol 2), Cambridge Univ. Press, 1991.
2. C. N. He, *The general solution of the matrix equation $AXB+CYD=F$*, Acta Sci. Natu. Univ. Nor. Hunanensis, China **19** (1996), 17-20 (in Chinese).
3. G. Z. Xu, M. S. Wei, D. Zheng, *On solutions of matrix equation $AXB+CYD=F$*, Linear Algera Appl. **279** (1980), 93-109.
4. A. P. Liao and Z. Z. Bai, *The constrained solutions of two matrix equations*, Acta Math. Sinica (English Ser.) **18** (2002), 671-678.
5. A. P. Liao and Z. Z. Bai, *Least squares symmetric and skew-symmetric solutions of the matrix equation $AXA^T+BYB^T=C$ with the least norm*, Math. Numer. Sinica **27** (2005), 81-95.
6. S. Shim and Y. Chen, *Least squares solution of matrix equation $AXB^*+CYD^*=F$*, SIAM J. Matrix Analysis Appl. **3** (2003), 8002-8008.
7. A. P. Liao, Z. Z. Bai and Y. Lei, *Best approximate solution of matrix equation $AXB+CYD=E$*, SIAM J. Matrix Anal. Appl. **27** (2005), 675-688.
8. M. Baruch, *Optimization procedure to correct stiffness and flexibility matrices using vibration tests*, J. Amer. Inst. Aero. Astr (AIAA). **16** (1978), 1208-1210.
9. Y. X. Peng, *An iterative method for the symmetric solutions and the optimal approximation solution of the linear matrix equation $AXB=C$*, Appl. Math. Comput. **160** (2005), 763-777.
10. Z. Y. Peng, X. Y. Hu and L. Zhang, *An iterative method for the Least-squares symmetric solution of the linear matrix equation $AXB=C$*, Appl. Math. Comput. **170** (2005), 711-723.
11. Y. B. Deng, Z. Z. Bai and Y. H. Gao, *Iterative orthogonal direction methods for Hermitian minimum norm solutions of two consistent matrix equations*, Numer. Linear Algebra Appl. **13** (2006), 801-823.
12. Z. Y. Peng and Y. X. Peng, *An efficient iterative method for solving the matrix equation $AXB+CYD=E$*, Numer. Linear Algebra Appl. **13** (2006), 473-485.
13. A. P. Liao and Z. Z. Bai, *Least-squares solution of AXB=D over symmetric positive semidefinite matrices X*, J. Comput. Math. **21** (2003), 175-182.
14. A. P. Liao and Z. Z. Bai, *Bisymmetric least-squares solutions of the matrix equation $A^TXA=D$*, Math. Numer. Sinica **24** (2002), 9-20.

Kai-juan Shen received her B.Sc from Lanzhou University (China,2005) and now is studing at Lanzhou University in China for M.Sc. Her main research area is Numerical Linear Algebra.

School of Mathematics and Statistics, Lanzhou University, Lanzhou 730000, P. R. China
e-mail: shenkj05@lzu.cn

Chuan-hua You received his B.Sc from Lanzhou University (China,1975) and M.Sc from Fudan University (China,1978). Now he is working at School of Mathematics and Statistics of Lanzhou University in China as a associate professor. His main research area is Numerical Linear Algebra, including matrix computation and inverse eigenvalue problem. He has published some papers in various international and national journals .

School of Mathematics and Statistics, Lanzhou University, Lanzhou 730000, P. R. China
e-mail: chyou@lzu.edu.cn

Yu-xia Du received her B.Sc from Yantai Normal university (China,2005) and now is studing at Lanzhou University in China for M.Sc. Her main research area is Numerical Linear Algebra.

School of Mathematics and Statistics, Lanzhou University, Lanzhou 730000, P. R. China
e-mail: duyx05@lzu.cn