Research Article

An Iterative Method for the Least-Squares Problems of a General Matrix Equation Subjects to Submatrix Constraints

Li-fang Dai, Mao-lin Liang, and Yong-hong Shen

School of Mathematics and Statistics, Tianshui Normal University, Tianshui, Gansu 741001, China

Correspondence should be addressed to Li-fang Dai; tsnclml2012@hotmail.com

Received 26 July 2013; Accepted 22 October 2013

1. Introduction

Throughout this paper, we denote the set of all $m \times n$ real matrices by $\mathbb{R}^{m \times n}$. The symbol $A^T$ represents the transpose of matrix $A$. $I_n$ and $I_p$ stand for the reverse unit matrix, and identity matrix, respectively. For $A, B \in \mathbb{R}^{m \times n}$, the inner product of matrices $A$ and $B$ is defined by $\langle A, B \rangle = \text{trace}(B^TA)$, which leads to the Frobenius norm, that is, $\|A\| = \sqrt{\langle A, A \rangle}$.

A matrix $A = (a_{ij}) \in \mathbb{R}^{m \times n}$, is called centro-symmetric (centro-skew symmetric) if and only if

$$a_{ij} = a_{n-i+1,n-j+1}, \quad i, j = 1, 2, \ldots, n,$$

which can also be characterized equivalently by $J_n A J_n = A$ ($J_n A J_n = -A$). The set of all centro-symmetric (centro-skew symmetric) matrices is denoted by $\text{CSR}^{m \times n}$ ($\text{CCSR}^{m \times n}$). This kind of matrices plays an important role in many applications (see, e.g., [1–4]), and has been frequently and widely investigated (see, e.g., [5–7]) by using generalized inverse, generalized singular value decomposition (GSVD) [8], and so forth. For more, we refer the readers to [9–16] and therein.

We firstly introduce the concept of the central principal submatrix which is originally put forward by Yin [17].

Definition 1. Let $A \in \mathbb{R}^{m \times n}$, if $n - p$ is even, a $p \times p$ central principal submatrix of $A$, denoted by $A[p]$, is obtained by deleting the first and last $(n - p)/2$ rows and columns of $A$, namely, $A[p] = (a_{ij})_{(n-p)/2 \times (n-p)/2}$.

Evidently, a matrix with odd (even) order only has central principal submatrices of odd (even) order.

Now, the first problem to be studied here can be stated as follows.

Problem 2. Given $M_i \in \mathbb{R}^{m \times p_i}, N_j \in \mathbb{R}^{p_i \times n}$ and $X_i \in \text{CSR}^{p_i \times q_i}$ ($i = 1, 2, \ldots, t$), $F \in \mathbb{R}^{m \times n}$. Find the least-squares solution $Z_i \in \Pi_i$ of matrix equation

$$M_i Z_i N_j + M_j Z_j N_i + \cdots + M_z Z_z N_z = F,$$

in which $\Pi_i = \{Z_i \mid Z_i \in \text{CSR}^{p_i \times q_i}, \text{with } Z_i[q_i] = X_i\}$, $p_i > q_i$, $i \in \Gamma = \{1, 2, \ldots, t\}$, and $Z_i[q_i]$ represents the $q_i \times q_i$ central principal submatrix of $Z_i$.

Problem 2 is the submatrix constrained problem of matrix equation (2), which originally arises from a practical subsystem expansion problem, and has been deeply investigated (see, e.g., [7, 18–22]). In these literatures, the generalized inverses or some complicated matrix decompositions such as canonical correlation decomposition (CCD) [23] and GSVD...
are employed. However, it is almost impossible to solve (2) by the above methods. The iterative method is an efficient approach. Recently, kinds of iteration methods have been constructed: Zhou and Duan [24] studied the generalized Sylvester matrix equation

$$\sum_{i=0}^{\phi} A_i X F_i + \sum_{k=0}^{v} B_k Y F_k = R$$  

(3)

by so-called generalized Sylvester mapping that has pretty properties. Wu et al. [25] presented an finite iterative method for a class of complex matrix equations including conjugate and transpose of unknown solution. Motivated by the well-known Jacobi and Gauss-Seidel iterations methods, Ding and Chen, in [26], proposed a general family of iterative methods to solve linear matrix equations; meanwhile, these methods were also extended to solve the following coupled Sylvester matrix equations

$$\sum_{j=1}^{p} A_{ij} X B_{ij} = C_{ij}, \quad i = 1, 2, \ldots, n$$  

(4)

Although these iterative algorithms are efficient, there still exist some handicaps when meeting the constrained matrix equation problem (i.e., to find the solution of matrix equation in some matrices sets with specifical structure, for instance, symmetric matrices, centro-symmetric matrices, and bi-symmetric matrices sets) and the submatrix constrained problem, since these methods cannot keep the special properties of the unknown matrix in the iterative process. Based on the classical conjugate gradient (CG) method, Peng et al. [27] gave an iterative method to find the bisymmetric solution of matrix equation (2). Similar method was constructed to solve matrix equations (4) with generalized bisymmetric $X_i$ in [28]. In particular, Li et al. [29] proposed an elegant algorithm for solving the generalized Sylvester (Lyapunov) matrix equation $AXB + CYD = E$ with bisymmetric $X$ and symmetric $Y$, the two unknown matrices include the given central principal submatrices and leading principal submatrix, respectively. This method shunned the difficulties in numerical instability and computational complexity, and solved the problem, completely. By borrowing the thinking of this iterative algorithm, we will solve Problem 2 by iteration method.

The second problem to be considered is the optimal approximation problem.

**Problem 3.** Let $S_B$ be the solutions set of Problem 2. For given matrices $\bar{Z}_i \in R^{p_i \times p}$, find $Z_i$ such that

$$\min_{Z_i \in S_B} \| Z_i - \bar{Z}_i \|^2.$$  

(5)

This problem occurs frequently in experimental design (see for instance [30]). Here, the preliminary estimation $\bar{Z}_i$ of the unknown matrix $Z_i$ can be obtained from experiments, but it may not satisfy the structural requirement and/or spectral requirement. The best estimation of $Z_i$, is the matrix $\bar{Z}_i$ that satisfies both requirements, which is the optimal approximation of $Z_i$ (see, e.g., [31, 32]). About this problem, we also refer the authors to [9–11, 13, 15, 16, 20–23, 27–29, 33–36] and therein.

The rest of this paper is outlined as follows. In Section 2, an iterative algorithm will be proposed to solve Problem 2, and the properties of which will be investigated. In Section 3, we will consider the optimal approximation Problem 3 by using the iterative algorithm. In Section 4, some numerical examples will be given to verify the efficiency of this algorithm.

2. The Algorithm for Problem 2 and Its Properties

According to the definition of centro-symmetric matrix, when $n - q$ is even, a centro-symmetric matrix $Z \in CSR^{p \times p}$ can be divided into smaller submatrices, namely,

$$Z = \begin{pmatrix} Z_{11} & Z_{12} & Z_{13} \\ Z_{21} & Z_{22} & Z_{23} \\ Z_{31} & Z_{32} & Z_{33} \end{pmatrix}.$$  

(6)

where $Z_{11} \in R^{(n-q)/2 \times (n-q)/2}$, $Z_{12} \in R^{(n-q)/2 \times q}$, $Z_{13} \in R^{(n-q)/2 \times (n-q)/2}$, $Z_{23} \in R^{q \times (n-q)/2}$, and $Z_{21} \in CSR^{q \times q}$.

Now, for some fixed positive integer $i \in \Gamma$, we define two matrix sets.

$$CSR_{\times A_i}^{p \times p_i} = \{ Z_i \mid Z_i \in CSR^{p \times p} \text{ with } Z_i[q_i] = 0 \},$$

$$CSR_{\circ A_i}^{p \times p_i} = \left\{ Z_i \mid Z_i = \begin{pmatrix} 0 & 0 & 0 \\ 0 & X_i & 0 \\ 0 & 0 & 0 \end{pmatrix} \in CSR^{p \times p} \right\}.$$  

(7)

It is clear that both $CSR_{\times A_i}^{p \times p_i}$ and $CSR_{\circ A_i}^{p \times p_i}$ are linear subspaces of $R^{p \times p_i}$.

In addition, for any matrix $X \in R^{n \times n}$, it has uniquely decomposition in direct sum, that is, $X = X_1 \oplus X_2$, here $X_1 = (X + J_n X J_n)/2$, $X_2 = (X - J_n X J_n)/2$. Furthermore, $X_1 = X_{11} + X_{12}$ is also the direct sum decomposition of $X_1$, if $X_{11} \in CSR_{\times A_i}^{p \times p_i}$, $Z_{12} \in CSR_{\circ A_i}^{p \times p_i}$, since $\langle X_{11}, X_{12} \rangle = 0$. Hence, we obtain the following.

**Lemma 4.** Consider $R^{p \times p_i} = CSR_{\times A_i}^{p \times p_i} \oplus CSR_{\circ A_i}^{p \times p_i} \oplus CSR^{R \times R \times p}$.

Lemma 4 reveals that any matrix $W \in R^{p \times p_i}$ can be uniquely written as $W = W_1 + W_2 + W_3$, where $W_1 \in CSR_{\times A_i}^{p \times p_i}$, $W_2 \in CSR_{\circ A_i}^{p \times p_i}$, $W_3 \in CSR^{R \times R \times p}$. Then, we can define the following linear projection operators:

$$\mathcal{L}_i : R^{p \times p_i} \rightarrow CSR_{\times A_i}^{p \times p_i},$$

$$W \rightarrow W_1$$  

(8)

for $i \in \Gamma$. 

According to the definition of $\mathcal{L}_i$, if $W \in R^{p_i \times p_i}$ and $Y \in CSR_{p_i}^{p_i}$, we have

$$\langle W, Y \rangle = \langle W_i, Y \rangle = \langle \mathcal{L}_i (W), Y \rangle .$$

(9)

This property will be employed frequently in the residual context.

The following theorem is essential for solving Problem 2, which transforms equivalently Problem 2 into solving the least-square problem of another matrix equation.

**Theorem 5.** Any solution group of Problem 2 can be obtained by

$$(Z_1, Z_2, \ldots, Z_t) = (Y_1 + Z_1^\circ, Y_2 + Z_2^\circ, \ldots, Y_t + Z_t^\circ),$$

(10)

where $Y_i \in CSR_{p_i}^{p_i}$ is the least-squares solution of matrix equation

$$M_1Y_1N_1 + M_2Y_2N_2 + \cdots + M_tY_tN_t = G,$$

$$G = F - (M_1Z_1^\circ + M_2Z_2^\circ + \cdots + M_tZ_t^\circ),$$

$$Z_i^\circ = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \in CSR_{p_i}^{p_i},$$

$X_i \in CSR_{p_i}^{p_i}$ is the given central principal submatrix of $Z_i$ in Problem 2.

**Proof.** Noting that the definition of $CSR_{p_i}^{p_i}$, we have

$$\min_{Z_i \in p_i} \left\| \sum_{i=1}^{t} M_iZ_iN_i - F \right\|$$

$$\iff \min_{Y_i \in CSR_{p_i}^{p_i}} \left\| \sum_{i=1}^{t} M_i(Y_i + Z_i)N_i - F \right\|$$

(12)

$$\iff \min_{Y_i \in CSR_{p_i}^{p_i}} \left\| \sum_{i=1}^{t} M_iY_iN_i - G \right\| .$$

The proof is completed. $\square$

**Remark 6.** It follows, from Theorem 5, that Problem 2 can be solved completely by finding the least-squares solution of matrix equations (11) in subspaces $CSR_{p_i}^{p_i}$.

In the next part of this section, we will establish an iterative algorithm for (11) and analysis its properties. For the convenience of expression, we define a matrix function

$$\mathcal{F} (Z_1, Z_2, \ldots, Z_t) = M_1Z_1N_1 + M_2Z_2N_2 + \cdots + M_tZ_tN_t,$$

(13)

then matrix equation (11) can be simplified as

$$\mathcal{F} (Z_1, Z_2, \ldots, Z_t) = G .$$

(14)

Moreover, we can easily verify that

$$\langle X, \mathcal{F} (Z_1, Z_2, \ldots, Z_t) \rangle = \sum_{i=1}^{t} \langle M_i^T XN_i^T, Z_i \rangle$$

holds for arbitrary $X \in R^{m \times n}$.

The iterative algorithm for the least squares problem of matrix equations (II) can be expressed as follows.

**Algorithm 7.** Consider the following.

Step 1. Let $M_i \in R^{m \times p_i}, N_i \in R^{p_i \times n}, F \in R^{m \times n}$ and $X_i \in CSR_{p_i}^{p_i}$, for $i \in \Gamma$.

Input arbitrary matrices $Y_i^{(0)} \in CSR_{p_i}^{p_i}$.

Step 2. Calculate

$$R_0 = G - \mathcal{F} (Y_1^{(0)}, Y_2^{(0)}, \ldots, Y_t^{(0)}),$$

$$P_0 = \mathcal{L}_i (M_i^T R_0 N_i^T), \quad Q_i^{(0)} = P_i^{(0)},$$

$$k = 0 .$$

Step 3. Calculate

$$Y_i^{(k+1)} = Y_i^{(k)} + \alpha_k Q_i^{(k)},$$

$$\alpha_k = \frac{\sum_{i=1}^{t} \| P_i^{(k)} \|^2}{\| \mathcal{F} (Q_1^{(k)}, Q_2^{(k)}, \ldots, Q_t^{(k)}) \|^2} .$$

(17)

Step 4. Calculate

$$R_{k+1} = G - \mathcal{F} (Y_1^{(k+1)}, Y_2^{(k+1)}, \ldots, Y_t^{(k+1)}),$$

$$= R_k - \alpha_k \mathcal{F} (Q_1^{(k)}, Q_2^{(k)}, \ldots, Q_t^{(k)}),$$

$$P_i^{(k+1)} = \mathcal{L}_i (M_i^T R_{k+1} N_i^T),$$

$$Q_i^{(k+1)} = P_i^{(k+1)} + \beta_k Q_i^{(k)},$$

$$\beta_k = \frac{\sum_{i=1}^{t} \| P_i^{(k+1)} \|^2}{\sum_{i=1}^{t} \| P_i^{(k)} \|^2} .$$

(18)

Step 5. If $\sum_{i=1}^{t} \| P_i^{(k)} \|^2 = 0$, stop. Otherwise, $k := k + 1$, go to Step 3.

From Algorithm 7, we can see that $Y_i^{(k)}, P_i^{(k)}, Q_i^{(k)} \in CSR_{p_i}^{p_i}$. In particular, $(Y_1^{(1)}, Y_2^{(1)}, \ldots, Y_t^{(1)})$ is a least-squares solution group belonging to matrix equation (11) if $P_i^{(k)} = 0$ for all $i \in \Gamma$. The following lemma gives voice to the reason.

**Lemma 8.** If $L_i(M_i^T R_k N_i^T) = 0$ ($i = 1, 2, \ldots, t$) hold simultaneously for some positive $k$, then $(Y_1^{(k)}, Y_2^{(k)}, \ldots, Y_t^{(k)})$ generated by Algorithm 7 is a solution group of matrix equation (11).

**Proof.** Let $L = \{ L \mid L = \mathcal{F} (Y_i), Y_i \in CSR_{p_i}^{p_i} \}$ and $\widetilde{G} = \mathcal{F} (Y_i^{(k)})$. Obviously, $\widetilde{G} \in L$. Then, from the Project Theorem, $(Y_1^{(k)}, Y_2^{(k)}, \ldots, Y_t^{(k)})$ is a least-square solution group of matrix equation (11) if and only if $G - \widetilde{G} \perp L$. That is to say, for
any matrices \( Y_{i} \in \text{CSR}^{R_{i} \times P_{i}}, \) noting that Lemma 4, we have
\[
\langle G - \tilde{G}, \mathcal{F}(Y_1, Y_2, \ldots, Y_t) \rangle = 0,
\]
that is,
\[
\langle G - \tilde{G}, \mathcal{F}(Y_1, Y_2, \ldots, Y_t) \rangle = \langle R_{k+1}, \mathcal{F}(Y_1, Y_2, \ldots, Y_t) \rangle
\]
\[
= \sum_{i=1}^{t} \langle M_i^T R_{k+1} N_i^T, Y_i \rangle
\]
\[
= \sum_{i=1}^{t} \langle \mathcal{L}_i (M_i^T R_{k+1} N_i^T), Y_i \rangle = 0,
\]
which also deduces that
\[
\sum_{i=1}^{t} \langle Q_i^0, P_i^{(1)} \rangle = 0,
\]
which completes the proof. \( \square \)

In addition, the sequences \( \{Y_{i}^{(0)}\}, \{P_{i}^{(0)}\}, \{Q_{i}^{(0)}\} \) generated by Algorithm 7 are self-orthogonal, that is, as follows.

**Lemma 9.** Suppose that the sequences \( \{Y_{i}^{(0)}\}, \{P_{i}^{(0)}\}, \{Q_{i}^{(0)}\} \) generated by Algorithm 7 not equal null for \( l \leq k \), then

\[ (1) \sum_{i=1}^{t} \langle P_{i}^{(j)}, P_{i}^{(l)} \rangle = 0, \quad (20) \]
\[ (2) \sum_{i=1}^{t} \langle Q_{i}^{(j)}, P_{i}^{(l)} \rangle = 0, \quad (21) \]
\[ (3) \langle \mathcal{F}(Q_{1}^{(j)}, Q_{2}^{(j)}, \ldots, Q_{t}^{(j)}), \mathcal{F}(Q_{1}^{(0)}, Q_{2}^{(0)}, \ldots, Q_{t}^{(0)}) \rangle = 0, \quad (22) \]

where \( j, l = 1, 2, \ldots, k, j \neq l \leq k \).

**Proof.** In view of the symmetry of the inner product, we only prove (20)–(22) when \( j < l \). According to Algorithm 7, when \( k = 1 \), we have
\[
\sum_{i=1}^{t} \langle P_{i}^{(0)}, P_{i}^{(1)} \rangle
\]
\[
= \sum_{i=1}^{t} \langle P_{i}^{(0)}, P_{i}^{(0)} \rangle - \alpha_0 \mathcal{L}_i M_i^T \left( \mathcal{F}(Q_{1}^{(0)}, Q_{2}^{(0)}, \ldots, Q_{t}^{(0)}) N_i^T \right)
\]
\[
= \sum_{i=1}^{t} \left[ \langle P_{i}^{(0)}, P_{i}^{(0)} \rangle - \alpha_0 \langle P_{i}^{(0)}, \mathcal{L}_i M_i^T \left( \mathcal{F}(Q_{1}^{(0)}, Q_{2}^{(0)}, \ldots, Q_{t}^{(0)}) N_i^T \right) \rangle \right]
\]
\[
= \sum_{i=1}^{t} \| P_{i}^{(0)} \|^2 - \alpha_0 \sum_{i=1}^{t} \| M_i P_{i}^{(0)} \|^2 - \alpha_0 \sum_{i=1}^{t} \left( M_i P_{i}^{(0)} N_i \right) \mathcal{F}(Q_{1}^{(0)}, Q_{2}^{(0)}, \ldots, Q_{t}^{(0)})
\]
\[
= \sum_{i=1}^{t} \| P_{i}^{(0)} \|^2 - \alpha_0 \sum_{i=1}^{t} \| M_i P_{i}^{(0)} \|^2 - \alpha_0 \sum_{i=1}^{t} \left( M_i P_{i}^{(0)} N_i \right) \mathcal{F}(Q_{1}^{(0)}, Q_{2}^{(0)}, \ldots, Q_{t}^{(0)})
\]
\[
= 0,
\]
(23)
Assume that (20), (21), and (22) hold for positive integer \( s ( < k) \), that is, for \( j = 1, 2, \ldots, s - 1 \),

\[
\sum_{i=1}^{t} \langle p_i^{(j)}, p_i^{(s+1)} \rangle = 0,
\]

\[
\sum_{i=1}^{t} \langle q_i^{(j)}, p_i^{(s+1)} \rangle = 0,
\]

\[
\langle \mathcal{F} (Q_1^{(j)}, Q_2^{(j)}, \ldots, Q_t^{(j)}), \mathcal{F} (Q_1^{(s)}, Q_2^{(s)}, \ldots, Q_t^{(s)}) \rangle = 0.
\]

Then, similar to the above proof, noting that the assumptions, we have

\[
\sum_{i=1}^{t} \langle p_i^{(j)}, p_i^{(s+1)} \rangle = \sum_{i=1}^{t} \langle p_i^{(j)}, \mathcal{L}_1 (M_i^T \mathcal{F} (Q_1^{(s)}, Q_2^{(s)}, \ldots, Q_t^{(s)}) N_i^T) \rangle
\]

\[
= \sum_{i=1}^{t} \langle p_i^{(j)} - \alpha_i X_i (M_i^T \mathcal{F} (Q_1^{(s)}, Q_2^{(s)}, \ldots, Q_t^{(s)}) N_i^T) \rangle
\]

\[
= -\alpha_i \sum_{i=1}^{t} \langle M_i p_i^{(s+1)} N_i, \mathcal{F} (Q_1^{(s)}, Q_2^{(s)}, \ldots, Q_t^{(s)}) \rangle
\]

\[
= -\alpha_i \sum_{i=1}^{t} \langle M_i (Q_i^{(s)} - \beta_{j+1} Q_{j+1}^{(s+1)} N_i, \mathcal{F} (Q_1^{(s)}, Q_2^{(s)}, \ldots, Q_t^{(s)}) \rangle
\]

\[
= -\alpha_i \langle \mathcal{F} (Q_1^{(s)}, Q_2^{(s)}, \ldots, Q_t^{(s)}), \mathcal{F} (Q_1^{(s)}, Q_2^{(s)}, \ldots, Q_t^{(s)}) \rangle
\]

\[
+ \alpha \beta_{j+1} \langle \mathcal{F} (Q_1^{(s+1)}, Q_2^{(s+1)}, \ldots, Q_t^{(s+1)}), \mathcal{F} (Q_1^{(s)}, Q_2^{(s)}, \ldots, Q_t^{(s)}) \rangle
\]

\[
= 0.
\]

Furthermore,

\[
\sum_{i=1}^{t} \langle q_i^{(j)}, p_i^{(s+1)} \rangle
\]

\[
= \sum_{i=1}^{t} \langle q_i^{(j)} - \alpha_i \mathcal{L}_1 (M_i^T \mathcal{F} (Q_1^{(s)}, Q_2^{(s)}, \ldots, Q_t^{(s)}) N_i^T) \rangle
\]

\[
= \sum_{i=1}^{t} \langle q_i^{(j)} - \alpha_i \mathcal{L}_1 (M_i^T \mathcal{F} (Q_1^{(s)}, Q_2^{(s)}, \ldots, Q_t^{(s)}) N_i^T) \rangle
\]

\[
= \sum_{i=1}^{t} \langle p_i^{(j)} - \alpha_i \mathcal{L}_1 (M_i^T \mathcal{F} (Q_1^{(s)}, Q_2^{(s)}, \ldots, Q_t^{(s)}) N_i^T) \rangle
\]

\[
= 0.
\]

Moreover,

\[
\sum_{i=1}^{t} \langle q_i^{(s)}, p_i^{(s)} \rangle
\]

\[
= \sum_{i=1}^{t} \langle q_i^{(s)} + \beta_{s+1} Q_{i+1}^{(s+1)} - \beta_{s+1} Q_{i+1}^{(s+1)} N_i, \mathcal{F} (Q_1^{(s)}, Q_2^{(s)}, \ldots, Q_t^{(s)}) \rangle
\]

\[
= 0.
\]
\[
\sum_{i=1}^{t} \langle Q^{(s)}_i, p_{(s+1)}^i \rangle = \sum_{i=1}^{t} \langle Q^{(s)}_i, p_{(s)}^i \rangle - \alpha_i \mathcal{F} \left( M_i^T \mathcal{F} \left( Q^{(s)}_1, Q^{(s)}_2, \ldots, Q^{(s)}_t \right) N_i^T \right) + \beta \| \mathcal{F} \left( Q^{(s)}_1, Q^{(s)}_2, \ldots, Q^{(s)}_t \right) \|^2
\]

It follows from (29) that

\[
\sum_{i=1}^{t} \langle Q^{(s)}_i, p_{(s+1)}^i \rangle = \sum_{i=1}^{t} \langle Q^{(s)}_i, p_{(s)}^i \rangle - \alpha_i \mathcal{F} \left( M_i^T \mathcal{F} \left( Q^{(s)}_1, Q^{(s)}_2, \ldots, Q^{(s)}_t \right) N_i^T \right) + \beta \| \mathcal{F} \left( Q^{(s)}_1, Q^{(s)}_2, \ldots, Q^{(s)}_t \right) \|^2
\]

\[
\sum_{i=1}^{t} \| p_{(s+1)}^i \|^2 - \alpha_i \| \mathcal{F} \left( Q^{(s+1)}_1, Q^{(s+1)}_2, \ldots, Q^{(s+1)}_t \right) \|^2 - \sum_{i=1}^{t} \| p_{(s)}^i \|^2 - \alpha_i \| \mathcal{F} \left( Q^{(s)}_1, Q^{(s)}_2, \ldots, Q^{(s)}_t \right) \|^2 = 0
\]

Remark 10. We know from Lemma 4 that the matrices sequences

\[
\left( \begin{array}{cccc}
P_{0,1} & P_{0,2} & \cdots & P_{0,t} \\
\vdots & \vdots & & \vdots \\
P_{k,1} & P_{k,2} & \cdots & P_{k,t}
\end{array} \right)
\]

are orthogonal to each other. Hence, it can be regarded as an orthogonal basis of matrix space \( R^{\sum_{i=1}^{t} p_{(s)}^i} \). Hence, the iteration will be terminated at most \( \sum_{i=1}^{t} (p_{i} + q_{i}) (p_{i} - q_{i}) / 2 \) steps in the absence of roundoff errors. Therefore, there exists a positive integer \( k \leq \sum_{i=1}^{t} (p_{i} + q_{i}) (p_{i} - q_{i}) / 2 \) such that \( \sum_{i=1}^{t} \| p_{(i)}^k \|^2 = 0 \), in this case, \( (Y_1^{(k)}, Y_2^{(k)}, \ldots, Y_t^{(k)}) \) can be regarded as a least-squares solution group of matrix equation (II).

In addition, we should point out that if \( \alpha_k = 0 \) or \( \infty \), the conclusions may not be true, and the iteration will break down before \( p_{(i)}^k = 0 \) for \( k < \sum_{i=1}^{t} (p_{i} + q_{i}) (p_{i} - q_{i}) / 2 \).

Actually, \( \alpha_k = 0 \) implies that \( \sum_{i=1}^{t} \| p_{(i)}^k \|^2 = 0 \) for \( i \in \Gamma \). While \( \alpha_k = \infty \) leads to \( \mathcal{F} (Q^{(k)}_1, Q^{(k)}_2, \ldots, Q^{(k)}_t) = 0 \), making inner product with \( R_j \) by both sides, it follows from Algorithm 7 that

\[
\langle R_j, \mathcal{F} \left( Q^{(k)}_1, Q^{(k)}_2, \ldots, Q^{(k)}_t \right) \rangle = \sum_{i=1}^{t} \langle p_{(i)}^k, p_{(i)}^k \rangle + \beta_{k-1} \langle q_{(k-1)}^i \rangle
\]

\[
\sum_{i=1}^{t} \| p_{(i)}^k \|^2 + \beta_{k-1} \sum_{i=1}^{t} \| q_{(k-1)}^i \|^2 = 0,
\]

which also implies the same situation as \( \alpha_k = 0 \). Hence, if there exists a positive integer \( s \) such that the coefficient \( \alpha_s = 0 \) or \( \alpha_s = \infty \), then the corresponding matrix group \( Y_1^{(s)}, Y_2^{(s)}, \ldots, Y_t^{(s)} \) is just the solution of matrix equation (II).

Together with the above analysis and Lemma 9, we can conclude the following theorem.

Theorem 11. For any initial iteration matrices \( Y_i^{(0)} \in \text{CSR}_+^{d_i} \), \( i = 1, 2, \ldots, t \), the least-squares solution of matrix equation (II) can be obtained within finite iteration steps. Moreover, suppose that \( (\bar{Y}_1, \bar{Y}_2, \ldots, \bar{Y}_t) \) is a least-squares solution group of (II), then the general solution to Problem 2 can be expressed as

\[
Z_i = \bar{Y}_i + X_i + Z_i^{(t)} (i = 1, 2, \ldots, t), \quad \text{where} \quad X_i \in \text{CSR}_+^{d_i} \text{ satisfy homogeneous equation}
\]

\[
\mathcal{F} (X_1, X_2, \ldots, X_t) = 0,
\]

\( Z_i^{(t)} \) as in Theorem 5.
In order to show the validity of Theorem 11, it is adequate to prove the following conclusion.

**Proposition 12.** The least-squares solution group \((Y_1, Y_2, \ldots, Y_t)\) of matrix equation (11) can be expressed as \(\bar{Y}_1 + X_i\), where \(X_i\) satisfy equality (33).

**Proof.** According to the assumptions, we obtain

\[
\min_{Y_i \in \mathbb{C}^{m \times n}} \| \mathcal{F}(Y_1, Y_2, \ldots, Y_t) - G \| \\
= \| \mathcal{F}(\bar{Y}_1, \bar{Y}_2, \ldots, \bar{Y}_t) - G \|. 
\]

(34)

On the other hand, noting that \(\mathcal{F}(X_1, X_2, \ldots, X_i) = 0\), then

\[
\| \mathcal{F}(\bar{Y}_1 + X_i, \bar{Y}_2 + X_i, \ldots, \bar{Y}_t + X_i) - G \|^2 \\
= \| \mathcal{F}(\bar{Y}_1, \bar{Y}_2, \ldots, \bar{Y}_t) - G + \mathcal{F}(X_1, X_2, \ldots, X_i) \|^2 \\
= \| \mathcal{F}(\bar{Y}_1, \bar{Y}_2, \ldots, \bar{Y}_t) - G \|^2. 
\]

(35)

The proof is completed. \(\square\)

Next, we will show that the unique least norm solution of matrix equation (11) can be derived by choosing a special kind of initial iteration matrices.

**Theorem 13.** Let the initial iteration matrices \(Y_i^{(0)} = \mathcal{L}_i(M_i^T H_i N_i^T)\) with arbitrary \(H \in \mathbb{R}^{m \times n}\), \(i = 1, 2, \ldots, t\), and then \((Y_1^*, Y_2^*, \ldots, Y_t^*)\) generated by Algorithm 7 is the least-norm least-squares solution group of matrix equation (11). Furthermore, the least-norm solution group to Problem 2 can be expressed by

\[
(Z_1^*, Z_2^*, \ldots, Z_t^*) = (Y_1^* + Z_1^*, Y_2^* + Z_2^*, \ldots, Y_t^* + Z_t^*). 
\]

(36)

**Proof.** From Algorithm 7 and Theorem 11, for initial iteration matrices \(Y_i^{(0)} = \mathcal{L}_i(M_i^T H_i N_i^T)\), we can obtain a least-squares solution \(Y_i^*\) of matrix equation (11) and there exists a matrix \(H^*\) such that \(Y_i^* = \mathcal{L}_i(M_i^T H^* N_i^T)\). Hence, it is enough to prove that the \(Y_i^*\) is the least-norm solution. In fact, noting that (33) and Proposition 12, we have

\[
\sum_{i=1}^t \|Y_i^* + X_i\|^2 \\
= \sum_{i=1}^t (\|Y_i^*\|^2 + \|X_i\|^2 + 2 \langle Y_i^*, X_i \rangle) \\
= \sum_{i=1}^t \|Y_i^*\|^2 + \sum_{i=1}^t \|X_i\|^2 + 2 \sum_{i=1}^t \langle \mathcal{L}_i(M_i^T H^* N_i^T), X_i \rangle \\
\geq \sum_{i=1}^t \|Y_i^*\|^2, 
\]

(37)

as required. \(\square\)

Theorems 11 and 13 display the efficiency of Algorithm 7. Actually, the iteration sequence \(\{Y_i^{(k)}\}\) converges smoothly to the solution \(Y_i\), that is the minimization property of Algorithm 7.

**Theorem 14.** For any initial iteration matrices \(Y_i^{(0)}\), the \(Y_i^{(k)}\) generated by Algorithm 7 satisfy the minimization problem

\[
\| \mathcal{F}(Y_i^{(k)}), Y_2^{(k)}, \ldots, Y_t^{(k)} \| - G \\
= \min_{Y_i \in \mathbb{C}^{m \times n}} \| \mathcal{F}(Y_1, Y_2, \ldots, Y_t) - G \|. 
\]

(38)

where \(\mathbb{L}_i = Y_i^{(0)} + \text{span}\{Q_1^{(i)}, Q_2^{(i)}, \ldots, Q_{t-1}^{(i)}\}, i = 1, 2, \ldots, t\).

**Proof.** From the definition of \(\mathbb{L}_i\), there exist a series of real numbers \(a_0, a_1, \ldots, a_{k-1}\) such that \(Y_i = Y_i^{(0)} + \sum_{i=0}^{k-1} a_i Q_i^{(i)}\).

Define a function of \(k\) variables \(f(a_0, a_1, \ldots, a_{k-1})\), that is,

\[
f(a_0, a_1, \ldots, a_{k-1}) \\
= \sum_{i=1}^t \left[ M_i \left( Y_i^{(0)} + \sum_{i=0}^{k-1} a_i Q_i^{(i)} \right) N_i - G \right]^2. 
\]

(39)

In addition, from Algorithm 7, we know that

\[
R_0 = R_1 + a_{k-1} \mathcal{F}(Q_1^{(k-1)}, Q_2^{(k-1)}, \ldots, Q_{t-1}^{(k-1)}) \\
+ \cdots + a_0 \mathcal{F}(Q_1^{(i)}, Q_2^{(i)}, \ldots, Q_{t-1}^{(i)}). 
\]

(40)

Noting that (22) and making the inner product with \(\mathcal{F}(Q_1^{(i)}, Q_2^{(i)}, \ldots, Q_{t-1}^{(i)})\) on both sides of (40) yield

\[
\langle \mathcal{F}(Q_1^{(i)}, Q_2^{(i)}, \ldots, Q_{t-1}^{(i)}), R_0 \rangle = \langle \mathcal{F}(Q_1^{(i)}, Q_2^{(i)}, \ldots, Q_{t-1}^{(i)}), R_i \rangle. 
\]

(41)

Hence, by simple calculation, (40) and (41), the function can be rewritten as

\[
f(a_0, a_1, \ldots, a_{k-1}) \\
= \|R_0\|^2 + \sum_{i=0}^{k-1} a_i^2 \| \mathcal{F}(Q_1^{(i)}, Q_2^{(i)}, \ldots, Q_{t-1}^{(i)}) \|^2 \\
- 2 \sum_{i=1}^{k-1} a_i \langle \mathcal{F}(Q_1^{(i)}, Q_2^{(i)}, \ldots, Q_{t-1}^{(i)}), R_i \rangle. 
\]

(42)

Then,

\[
\min_{a_i} f(a_0, a_1, \ldots, a_{k-1}) \\
\iff \min_{Y_i \in \mathbb{C}^{m \times n}} \| \mathcal{F}(Y_1, Y_2, \ldots, Y_t) - G \|. 
\]

(43)
Since \( f(a_0, a_1, \ldots, a_{k-1}) = \min \) only if \( \frac{\partial f}{\partial a_l} = 0 \), it follows from (29) that
\[
a_l = \frac{\langle \mathcal{F}(Q^{(l)}_1, Q^{(l)}_2, \ldots, Q^{(l)}_r), R_l \rangle}{\| \mathcal{F}(Q^{(l)}_1, Q^{(l)}_2, \ldots, Q^{(l)}_r) \|^2}
= \frac{\sum_{i=1}^r \langle Q^{(l)}_i, P_{ij} \rangle}{\| \mathcal{F}(Q^{(l)}_1, Q^{(l)}_2, \ldots, Q^{(l)}_r) \|^2}
= \frac{\sum_{i=1}^r \| P_{ij} \|^2}{\| \mathcal{F}(Q^{(l)}_1, Q^{(l)}_2, \ldots, Q^{(l)}_r) \|^2} = \alpha_l.
\]
Combined with (43), we complete the proof. \(\square\)

Theorem 14 reveals that the sequence
\[
\| \mathcal{F}(Y_1^{(l)}, Y_2^{(l)}, \ldots, Y_t^{(l)}) - G \|
\]
monotonically decreases with respect to increasing integer \(k\). The descent property of the residual norm of matrix equation (11) leads to the smoothly convergence of Algorithm 7.

3. The Solution of Problem 3

In this section, we discuss the optimal approximation Problem 3. Since the least squares problem is always consistent, it is easy to verify that the solution set \(Z_0\) of Problem 2 is a nonempty convex cone, so the optimal approximation solution is unique.

Without loss of generality, we can assume that the given matrices \(\tilde{Z}_i \in \text{CSR}_{x_j}^{P, x_P}\). In fact, from Lemma 4, arbitrary \(\tilde{Z}_i \in \mathbb{R}^{n \times m}\) can be divided into
\[
\tilde{Z}_i = \tilde{Z}_{i1} + \tilde{Z}_{i2} + \tilde{Z}_{i3}, \quad \text{with} \quad \tilde{Z}_{i1} \in \text{CSR}_{x_j}^{P, x_P},
\]
\[
\tilde{Z}_{i2} \in \text{CSR}_{x_0, x_0}^{P, x_P}, \quad \tilde{Z}_{i3} \in \text{CASR}_{x_0, x_0}^{P, x_P}.
\]
Furthermore, if \(Z_i \in \text{CSR}_{x_j}^{P, x_P}\), then
\[
\| Z_i - \tilde{Z}_i \|^2 = \| Z_i - \tilde{Z}_{i1} \|^2 + \| \tilde{Z}_{i2} \|^2 + \| \tilde{Z}_{i3} \|^2,
\]
which meets the claim.

Denote \(Z_i = Z_i - \tilde{Z}_i, F = F - \mathcal{F}(\tilde{Z}_i, \tilde{Z}_2, \ldots, \tilde{Z}_t)\), then to solve Problem 3 is equivalent to find the least-norm solution of the new matrix equation
\[
\mathcal{F}(\tilde{Z}_1, \tilde{Z}_2, \ldots, \tilde{Z}_t) = F.
\]
Furthermore, similar to the construction of (11), Problem 2 is transformed equivalently into finding the least-norm least-squares solution of matrix equation
\[
\mathcal{F}(\bar{Y}_1, \bar{Y}_2, \ldots, \bar{Y}_t) = \bar{G}, \quad \text{with} \quad \bar{Y}_i \in \text{CSR}_{x_j}^{P, x_P},
\]
in which \(\bar{G} = F - \mathcal{F}(Z_{i1}^0, Z_{i2}^0, \ldots, Z_{it}^0)\).

Therefore, we can apply Algorithm 7 to derive the required solution of matrix equation (49). Virtually, it follows from Theorem 13 that if let the initial iteration matrices \(\bar{Y}_i^{(0)} = \mathcal{F}(M_i Z_0^T H N_i)\) with arbitrary \(H \in \mathbb{R}^{m \times n}\), or especially \(\bar{Y}_i^{(0)} = 0 \in \mathbb{R}^{n \times m}\), then the iteration solutions \(\bar{Y}_i^k\) consist of least-norm least-squares solution of which. In this case, the unique optimal approximation solution to Problem 3 can be obtained by
\[
(\bar{Z}_1, \bar{Z}_2, \ldots, \bar{Z}_t)
= (Y_1^* + Z_{11}^0, Y_2^* + Z_{21}^0, \ldots, Y_t^* + Z_{t1}^0).
\]

4. Numerical Example

In this section, we illustrate the efficiency and reliability of Algorithm 7 by some numerical experiments. All the numerical experiments are performed by using Matlab 6.5. In addition, because of the influence of the round-off errors, \(P_i^{(k)}\) may not equal zero within finite iteration steps, so the iteration will be terminated if \(\sum_{i=1}^t \| P_i^{(k)} \|^2 < \epsilon\), for example, let \(\epsilon = 1.0 \times 10^{-08}\). At this time, \((Y_1^{(k)}, Y_2^{(k)}, \ldots, Y_t^{(k)})\) can be regarded as a solution of matrix equation (11), and \(Y_i^{(k)} + Z_i^0\) \((i = 1, 2, \ldots, t)\) consist of the solution group to Problem 2. In particular, let the initial iteration matrices \(Y_i^{(0)} = 0\), then we will obtain the least-norm solution by (36).

**Example 1.** Input matrices \(M_1, M_2, M_3, N_1, N_2, N_3, F\) as follows:

\[
M_1 = \begin{bmatrix} \text{hilb}(r) & \text{ones}(r) \\ \text{hankel}(1 : r) & \text{zeros}(r) \end{bmatrix},
\]
\[
M_2 = \begin{bmatrix} \text{toeplitz}(1 : r) & \text{hilb}(r) \\ \text{ones}(r) & \text{hankel}(1 : r) \end{bmatrix},
\]
\[
M_3 = \begin{bmatrix} \text{zeros}(r) & \text{hankel}(1 : r) \\ \text{hilb}(r) & \text{ones}(r) \end{bmatrix},
\]
\[
N_1 = \text{eye}(r), \quad N_2 = \text{ones}(r), \quad N_3 = \text{tridiag}([7, 1, -1], r),
\]
\[
F = \begin{bmatrix} 3 & -2 & -1 \\ \vdots & 3 & -2 \\ -1 & -2 & -1 \\ \vdots & -2 & 3 & -2 \\ -1 & -2 & 3 \end{bmatrix},
\]
where toeplitz(k), hilb(k), hankel(k), zeros(k), and eye(k)
denote the Toeplitz matrix, Hilbert matrix, Hankel matrix, null matrix, identity matrix with orders $k$, and the elements of matrix ones() are one, tridiag([7, 1, -1], $r$) represents $r \times r$ tri-diagonal matrix produced by vector [7, 1, -1].

Let the given central principal matrices

$$\mathcal{X}_1 = \text{zeros}(\frac{r}{2}),$$

$$\mathcal{X}_2 = \text{ones}(\frac{r}{2}) \ast 10, \quad \mathcal{X}_3 = \text{toeplitz}(1: \frac{r}{2}).$$ (52)

By using the Algorithm 7, we obtain the solution to Problem 2. To save space, we shall not report the explicit datum of the solution, but the bars graphs of the components for the solution matrices will be given. Let $r = 20$, Figure 1 shows the bars graphs of $Z_1$, $Z_2$, $Z_3$ when we choose the initial iterative matrices

$$Y_i = \begin{pmatrix}
\text{ones}(\frac{r-q}{2}) & \text{ones}(\frac{r-q}{2}, q) & \text{ones}(\frac{r-q}{2}) \\
\text{ones}(q, \frac{r-q}{2}) & \text{zeros}(\frac{r}{2}) & \text{ones}(q, \frac{r-q}{2}) \\
\text{ones}(\frac{r-q}{2}) & \text{ones}(\frac{r-q}{2}, q) & \text{ones}(\frac{r-q}{2})
\end{pmatrix},$$

$$i = 1, 2, 3,$$ (53)

and the terminal condition $\|P_1^{(k)}\| + \|P_2^{(k)}\| + \|P_3^{(k)}\| < \epsilon = 1.0 \times 10^{-12}$.

Moreover, when $r = 20$ and $r = 40$, the convergence curves for the Frobenius norm of the residual denoted by RES = \( \|R_k\| \) and the termination condition denoted by TC = \( \|P_1^{(k)}\| + \|P_2^{(k)}\| + \|P_3^{(k)}\| \) are plotted in Figures 2 and 3, respectively.
From Figure 2, we can see that the residual norm of Algorithm 7 is monotonically decreasing, which is in accordance with the theory established in Theorem 14, namely, this algorithm is numerically stable. While Figure 3 shows that the terminated condition \( \| P_1^{(k)} \| + \| P_2^{(k)} \| + \| P_3^{(k)} \| \) is oscillating back and forth and approaches to zero as iterative process. Hence, the iterative Algorithm 7 is efficient, but it lacks of smooth convergence. Of course, for a problem with large and sparse matrices, Algorithm 7 may not terminate in a finite number of steps because of roundoff errors. How to establish an efficient and smooth algorithm is an important problem which we should study in a future work.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

The authors would like to express their sincere gratitude to the editor and two anonymous reviewers for their valuable comments and suggestions which have helped immensely improving the quality of the paper. Mao-lin Liang acknowledges the support of the scientific foundation of Tian-shui Normal University (no. TSA1104). Young-hong Shen is supported by the “QingLan” Talent Engineering Funds of Tianshui Normal University.

References


