Numerical solutions of $AXB = C$ for centrosymmetric matrix $X$ under a specified submatrix constraint

Jiao-fen Li$^1$,*,†, Xi-yan Hu$^2$ and Lei Zhang$^2$

$^1$School of Mathematics and Computational Science, Guilin University of Electronic Technology, Guilin 541004, People’s Republic of China
$^2$College of Mathematics and Econometrics, Hunan University, Changsha, Hunan 410082, People’s Republic of China

SUMMARY

We say that $X = [x_{ij}]$ is centrosymmetric if $x_{ij} = x_{n-j+1,n-i+1}$, $1 \leq i, j \leq n$. In this paper, we present an efficient algorithm for minimizing $\|AXB - C\|$ where $\|\cdot\|$ is the Frobenius norm, $A \in \mathbb{R}^{m \times n}$, $B \in \mathbb{R}^{n \times s}$, $C \in \mathbb{R}^{m \times s}$ and $X \in \mathbb{R}^{n \times n}$ is centrosymmetric with a specified central submatrix $[x_{ij}]_{p \leq i, j \leq n-p}$. Our algorithm produces a suitable $X$ such that $AXB = C$ in finitely many steps, if such an $X$ exists. We show that the algorithm is stable in any case, and we give results of numerical experiments that support this claim. Copyright © 2011 John Wiley & Sons, Ltd.

1. INTRODUCTION

Let $J_n$ be the $n \times n$ order flip matrix with ones on the secondary diagonal and zeros elsewhere. A matrix $X \in \mathbb{R}^{n \times n}$ is said to be centrosymmetric if $J_nXJ_n = X$, i.e.

$$x_{n-j+1,n-i+1} = x_{i,j}, \quad 1 \leq i, j \leq n,$$

or centroskew if $J_nXJ_n = -X$, i.e.

$$x_{n-j+1,n-i+1} = -x_{i,j}, \quad 1 \leq i, j \leq n.$$

Throughout this paper, $\mathbb{C}\mathbb{S}^{n \times n}$ and $\mathbb{C}\mathbb{A}\mathbb{S}^{n \times n}$ denote, respectively, the set of all $n \times n$ real centrosymmetric matrices and the set of all real centroskew matrices, whereas $\mathbb{R}^{m \times n}$ denotes the set of all $m \times n$ real matrices. Let $\|Z\| = (\langle Z, Z \rangle)^{1/2}$ be the Frobenius matrix norm of a matrix $Z$, $\langle Z, Y \rangle = \text{tr}(Z^TY)$ is the associated inner product of $Z$ with a matrix $Y$, and $\text{tr}(W)$ denotes the trace of a square matrix $W$. By using properties of the trace operator, we have that for any matrices $W, Y, Z$, $\langle W, YZ \rangle = \langle Y^TW, Z \rangle = \langle WZ^T, Y \rangle$. $\mathbb{R}(A)$ represents the column space of $A$. ver(.) represents the vec operator. $A \otimes B$ stands for the Kronecker product of matrices $A$ and $B$.

Centrosymmetric and centroskew matrices play an important role in many areas such as pattern recognition [1], the numerical solution of differential equations [2], Markov processes [3] and

*Correspondence to: Jiao-fen Li, School of Mathematics and Computational Science, Guilin University of Electronic Technology, Guilin 541004, People’s Republic of China.
†E-mail: lijiaofen603@yahoo.cn

Copyright © 2011 John Wiley & Sons, Ltd.
various physics and engineering problems [4, 5]. Symmetric centrosymmetric matrices form an important subclass of centrosymmetric matrices, which naturally occur in digital signal processing and other areas [6]. Important examples of symmetric centrosymmetric matrices are the symmetric Toeplitz matrices, the centrosymmetric Hankel matrices and the symmetric circulant matrices.

Recently there has been interest in the submatrix constraint problem of centrosymmetric and symmetric centrosymmetric matrices. For example, Bai [7] have considered the inverse eigenproblem of centrosymmetric matrices with a submatric constraint and the associated approximation problem. Peng et al. [8] have studied the inverse problem of symmetric centrosymmetric matrices with a submatrix constraint. While Liao and Lei [9] have studied the least-squares solution of the inverse problem of symmetric centrosymmetric matrices with a submatrix constraint. However, because of the specified centrosymmetric structure, it is unfit for discussing centrosymmetric matrices under their leading principal submatrices (i.e. the principal submatrices lie in the left-top of a given matrix) constraint, for they destroy the special centrosymmetry. Therefore, we present a concept of central principal submatrix, which was originally proposed by Yin [10]. The definition is as follows:

Definition 1
Given \( M = [m_{ij}] \in \mathbb{R}^{n \times n} \), if \( n - q \) is even, then a \( q \)-square central principal submatrix of \( M \), denoted as \( M_c(q) \), is a \( q \)-square submatrix obtained by deleting the first and last \( (n - q)/2 \) rows and columns of \( M \), that is \( M_c(q) = [m_{ij}]_{(n-q)/2 < i,j < n-(n-q)/2} \).

For example, if

\[
M = \begin{pmatrix}
a & b & c & d \\
e & f & g & h \\
i & j & k & l \\
m & n & o & p \\
\end{pmatrix} \in \mathbb{R}^{4 \times 4},
\]

then the 2-square central principal submatrix of \( M \) is

\[
M_c(2) = \begin{pmatrix}
f & g \\
j & k \\
\end{pmatrix}.
\]

It is intuitive and obvious that a matrix of odd (even, respectively) order has only central principal submatrices of odd (even, respectively) order.

The problem of finding various constraint solutions \( X \) of the matrix equation \( AXB = C \) has been considered by several authors [11–14] because of its applications in the design and analysis of the vibrating structures. However, the submatrix constraint problem that we are considering has not been discussed. We also should point out that the matrices \( A, B, \) and \( C \) occurring in practice are usually obtained from experiments and they may not satisfy the solvability conditions. Therefore, we study the least-squares problem, which can be mathematically formulated as follows.

Problem 1
Given \( A \in \mathbb{R}^{m \times n}, B \in \mathbb{R}^{n \times s}, C \in \mathbb{R}^{m \times s}, \) and \( X_q \in \mathbb{C}^{S_q \times q} \). Let

\[
\mathcal{S} = \{ X | X \in \mathbb{C}^{n \times n} \text{ with } X_c(q) = X_q \}.
\]

Find \( \tilde{X} \in \mathcal{S} \) such that

\[
\| A \tilde{X} B - C \| = \min_{X \in \mathcal{S}} \| A X B - C \|.
\]

We also consider the optimal approximation problem, which occurs frequently in structural identification [15].
Problem II
Let \( X_s \in \mathbb{R}^{n \times n} \) be given and \( X_{se}(q) = X_q \) (be given in Problem I) and let \( S_E \) denote the solution set of Problem I. Find \( \tilde{X} \in S_E \) such that
\[
\| \tilde{X} - X_s \| = \min_{\tilde{X} \in S_E} \| \tilde{X} - X_s \|.
\]

Our results are natured extension of results obtain in [8, 7, 9]. That work was motivated by results of serval investigators including [16–20]. These references describe application in which such problems arise. We are particularly influenced by Zhao et al. [21], who considered the case \( AX = B \) for symmetric centrosymmetric matrix \( X \) under central principal submatrix constraint. In these papers, inevitably, Moore–Penrose generalized inverses and some complicated matrix decompositions such as canonical correlation decomposition (CCD) [22] and general singular value decomposition (GSVD) [23] are involved. Because of the obvious difficulties in numerical instability and computational complexity, those constructional solutions narrow down their applications. Indeed, it is impractical to find a solution by those formulas if the matrix size is large. In the present paper, we extend and develop the above research, however, in a totally different way.

In this paper we are only concerned with iteration method, and the main idea is based on the classical Conjugate Gradient Least-Squares method (CGLS) as well as the minimal residual iteration idea proposed in [14]. We first transform Problem I to an equivalent least-squares problem over a linear subspace. It provides a way to construct an algorithm for solving the equivalent problem. With the proposed algorithm, the required submatrix constraint condition is automatically satisfied if the initial guess \( X_0 \) is chosen within a certain set, and a solution can be obtained with finitely many steps. The algorithms require little work and low storage requirements per iteration. In fact, we need only to compute a residual matrix and update the iterative solution in each iteration. We have also verified that the algorithm satisfies a minimization property, which ensures that this algorithm possesses a smooth convergence. In addition, the related optimal approximation problem is also solved. Some numerical results display the efficiency of these algorithms. Moreover, combined with numerical examples, we give some perturbation analysis on the approximation problem, and show that our algorithms are numerically stable associated with the approximation problem.

2. SOME PRELIMINARY RESULTS

We first make the following splitting of centrosymmetric matrix \( X \) into smaller submatrices:

\[
X = \begin{pmatrix}
X_{11} & X_{12} & X_{13} \\
X_{21} & X_{22} & J_q X_{21} J_{\frac{n-q}{2}} \\
J_{\frac{n+q}{2}} X_{13} & J_{\frac{n+q}{2}} X_{12} J_q & J_{\frac{n+q}{2}} X_{11} J_{\frac{n+q}{2}}
\end{pmatrix}
\]  

(2)

where \( X_{11} \in \mathbb{R}^{((n-q)/2) \times ((n-q)/2)} \), \( X_{12} \in \mathbb{R}^{((n-q)/2) \times (n-q)/2} \), \( X_{13} \in \mathbb{R}^{((n-q)/2) \times ((n-q)/2)} \), and \( X_{22} \in \mathbb{C}^{q \times q} \) (the \( q \)-square central principal submatrix of \( X \)). Actually, we can partition \( X \) into the following form:

\[
X = \begin{pmatrix}
X_{11} & X_{12} & X_{13} \\
X_{21} & X_{22} & X_{23} \\
X_{31} & X_{32} & X_{33}
\end{pmatrix}.
\]  

(3)

Comparing the two side of \( J_q X J_q = X \), we obtain \( X_{33} = J_{(n-q)/2} X_{11} J_{(n-q)/2} \), \( X_{22} = J_q X_{22} J_q \), \( X_{23} = J_q X_{21} J_{(n-q)/2} \), \( X_{31} = J_{(n-q)/2} X_{13} J_{(n-q)/2} \), and \( X_{32} = J_{(n-q)/2} X_{12} J_q \). Substituting \( X_{13}, X_{23}, X_{33}, \) and \( X_{32} \) into (3) and noticing that \( X_{22} = J_q X_{22} J_q \), we have (2).
\textbf{Definition 2}

\[ \mathbb{CS}_a^{n \times n} = \{ X | X \in \mathbb{CS}^{n \times n}, \text{ with } \lambda_X(q) = 0 \}, \]
\[ \mathbb{CS}_o^{n \times n} = \left\{ X \right\} X = \left( \begin{array}{ccc} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{array} \right), X \in \mathbb{CS}^{n \times n}, \forall \tilde{X}_q = \lambda_X(q) \in \mathbb{CS}^{q' \times q} \right\}. \]

Obviously, \( \mathbb{CS}_a^{n \times n} \) and \( \mathbb{CS}_o^{n \times n} \) are linear subspace of \( \mathbb{R}^{n \times n} \). For any arbitrary \( n \times n \) centrosymmetric matrix \( X \), suppose that it is partitioned as (2), then

\[ X = \begin{pmatrix} X_{11} & X_{12} & X_{13} \\ X_{21} & 0 & J_q \end{pmatrix} \begin{pmatrix} 0 & 0 & 0 \\ 0 & X_{22} & 0 \end{pmatrix}. \]

We remark that the right two block matrices partitioned in this way are unique, and their inner product is zero, which implies that the two linear subspaces \( \mathbb{CS}_a^{n \times n} \) and \( \mathbb{CS}_o^{n \times n} \) are orthogonal to each other. Combining with \( \mathbb{R}^{n \times n} = \mathbb{CS}_a^{n \times n} \oplus \mathbb{CS}_o^{n \times n} \), we can derive the following lemma.

\textbf{Lemma 1}

For any arbitrary \( Z \in \mathbb{R}^{n \times n} \), there exist unique \( Z_1 \in \mathbb{CS}_a^{n \times n} \), \( Z_2 \in \mathbb{CS}_o^{n \times n} \), and \( Z_3 \in \mathbb{CS}_o^{n \times n} \) such that

\[ Z = Z_1 + Z_2 + Z_3 \quad \text{where } \langle Z_i, Z_j \rangle = 0, \quad i \neq j, \quad i, j = 1, 2, 3. \]

From Lemma 1, we can define the following linear projection operator:

\[ \mathcal{A} : \mathbb{R}^{n \times n} \longrightarrow \mathbb{CS}_a^{n \times n}, \]
\[ Z \longrightarrow Z_1. \]

Then for any \( Z \in \mathbb{R}^{n \times n} \), \( W \in \mathbb{CS}_a^{n \times n} \)

\[ \langle Z, W \rangle = \langle Z_1 + Z_2 + Z_3, W \rangle = \langle Z_1, W \rangle = \langle \mathcal{A}(Z), W \rangle. \quad (4) \]

The following lemma-Projection Theorem from [24] is directly useful for stating our main results.

\textbf{Lemma 2}

Let \( \mathbb{X} \) be a finite-dimensional inner product space, \( \mathbb{M} \) be a subspace of \( \mathbb{X} \), and \( \mathbb{M}^\perp \) be the orthogonal complement of \( \mathbb{M} \). For a given \( x \in \mathbb{X} \), there always exists an \( m_0 \in \mathbb{M} \) such that \( \| x - m_0 \| \leq \| x - m \|, \forall m \in \mathbb{M} \), where \( \| \cdot \| \) is the norm associated with the inner product defined in \( \mathbb{X} \). Moreover, \( m_0 \in \mathbb{M} \) is the unique minimization vector in \( \mathbb{M} \) if and only if \( (x - m_0) \perp \mathbb{M} \) i.e. \( (x - m_0) \in \mathbb{M}^\perp \).

\section{3. Iterative Method for Problem I}

In this section we first briefly review the algorithm CGLS, then transform Problem I to an equivalent least-squares problem—Problem A, which is crucial for solving the problem completely and is a special feature of this paper. Then we construct an algorithm for Problem A based on the CGLS method. We call the resulting algorithm the Extended CGLS Algorithm. And then describe some of its properties. Finally, we show that the iteration terminates within finitely many steps in exact arithmetic and the minimum norm solution of Problem A can be obtained by choosing a special kind of initial matrices.

Copyright © 2011 John Wiley & Sons, Ltd.

\textbf{Numer. Linear Algebra Appl.} 2011; \textbf{18}:857–873

DOI: 10.1002/nla
3.1. The CGLS algorithms

The CGLS algorithm was originally proposed by Stiefel [25] for solving the following least-squares problem:

\[ \min_x \| Mx - f \|_2 \]

with given \( M \in \mathbb{R}^{m \times n} \) and \( f \in \mathbb{R}^m \). The CGLS algorithm is the Conjugate Gradient (CG) method applied on the normal equation

\[ M^T Mx = M^T f. \]

Starting at an initial guess \( x_0 \), CGLS finds an optimal solution \( x_k \) in the affine Krylov subspace \( x_0 + \mathcal{K}_k \) that minimizes the residual error \( \| f - Mx \|_2 \),

\[ x = \arg \min_{x \in x_0 + \mathcal{K}_k} \| f - Mx \|_2, \]

where \( \mathcal{K}_k \) is the Krylov subspace

\[ \mathcal{K}_k \equiv \mathcal{K}_k(M^T M, r_0) = \text{span}\{r_0, (M^T M)r_0, \ldots, (M^T M)^{k-1}r_0\}, \]

and \( r_0 = M^T(f - Mx_0) \) is the residual vector of the initial guess \( x_0 \) corresponding to the normal equations. Let \( r_k = M^T(f - Mx_k) \) be the residual vector of \( x_k \) corresponding to the normal equation and \( p_1, p_2, \ldots, p_k \) an \( M^T M \)-orthogonal basis vector of \( \mathcal{K}_k \), i.e.

\[ p_i^T(M^T M)p_j = 0 \quad \text{if} \ i \neq j. \]

The optimal solution \( x_k \) defined above can be constructed iteratively by

\[ x_{k+1} = x_k + a_k p_k, \quad a_k = \rho_k / \| Mp_k \|_2^2, \]
\[ r_{k+1} = r_k - a_k M^T M p_k, \quad \rho_{k+1} = \| r_{k+1} \|_2^2, \]
\[ p_{k+1} = r_{k+1} + \beta_k p_k, \quad \beta_k = \rho_{k+1} / \rho_k. \]

Initially, \( p_0 = r_0 \) and \( \rho_0 = \| p_0 \|_2^2 \). It can be verified that the residual vectors are orthogonal to each other, i.e. \( r_i^T r_j = 0 \) for \( i \neq j \).

Theoretically, CGLS converges within at most \( n \) iterations if exact arithmetic could be performed, where \( n \) is the order of \( M \). In practice, the iteration numbers may be larger than \( n \) because of the computational errors.

3.2. An equivalent least-squares problem

First, the technique of transformation is precisely described in the following lemma.

**Lemma 3**

Let \( A \) and \( B \) be as in Problem I. The general solution of Problem I can be expressed as \( X = Y + \tilde{X} \), where \( Y \) is the general solution of the following minimization problem:

\[ \text{Problem A: Find } Y \in \mathbb{C}S_n^{n \times n} \text{ such that } \| AYB - D \| = \min, \text{ where} \]

\[ D = C - A\tilde{X}B \quad \text{and} \quad \tilde{X} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & X_q & 0 \\ 0 & 0 & 0 \end{pmatrix} \in \mathbb{C}S_n^{n \times n}. \]
Then there exist $Z^* \in \text{gradient } \nabla X$, which implies that $\mathcal{S}$ is a linear manifold. And we have

$$\min_{X \in \mathcal{S}} \|AXB - C\| \Leftrightarrow \min_{Y \in CS^*_{n \times n}} \|A(\bar{X} + Y)B - C\| \Leftrightarrow \min_{Y \in CS^*_{n \times n}} \|AYB - (C - A\bar{X}B)\|
\Leftrightarrow \min_{Y \in CS^*_{n \times n}} \|AYB - D\|.$$ 

**Remark 1**

Since $\mathcal{S}$ is a linear manifold, we cannot construct the iterative algorithm directly, since it is hard to ensure that the updated matrix satisfies the constraints of the symmetric centrosymmetric and the submatrix requirement. Hence, we convert the problem over linear manifold to an equivalent problem over a linear subspace. This provides a way to construct an iterative algorithm for solving Problem I completely.

Now we must find the least-squares solution over $CS^*_{n \times n}$ of matrix equation

$$AYB = D. \quad (6)$$

Then the solution of Problem I can be expressed as $\bar{X} = Y + \bar{X}$.

**Lemma 4**

Let $\bar{R}$ be the residual of Equation (6) corresponding to matrix $\bar{Y}$, i.e. $\bar{R} = D - A\bar{Y}B$. $\bar{Y}$ is a least-squares solution of Equation (6) over $CS^*_{n \times n}$ if and only if $\mathcal{A}(A^T\bar{R}B^T) = 0$.

**Proof**

Let

$$\mathcal{L} = \{L | L = AYB, Y \in CS^*_{n \times n}\}.$$ 

Obviously, $\mathcal{L}$ is a linear subspace of $\mathbb{R}^{m \times m}$. For $\bar{Y} \in CS^*_{n \times n}$, denote $\bar{D} = A\bar{Y}B$, then $\bar{D} \in \mathcal{L}$. From the Projection Theorem, $\bar{Y}$ is a least-squares solution of Equation (6) if and only if $(D - \bar{D}) \perp \mathcal{L}$, i.e. $\langle D - \bar{D}, AYB \rangle = \langle \bar{R}, AYB \rangle = 0$ for any arbitrary $Y \in CS^*_{n \times n}$. From Lemma 1 and (4)

$$\langle \bar{R}, AYB \rangle = \langle A^T\bar{R}B^T, Y \rangle = \langle \mathcal{A}(A^T\bar{R}B^T), Y \rangle.$$

Hence $\bar{Y}$ is a least-squares solution of Equation (6) over $CS^*_{n \times n}$ if and only if $\mathcal{A}(A^T\bar{R}B^T) = 0$. \hfill \Box

**Remark 2**

Lemma 4 is in accordance with the theory established in linear system. Actually, if we define function $f: \mathbb{R}^{m \times n} \to \mathbb{R}$, given by $f(Z) = \|AZB - D\|^2$, then $f$ is differentiable and its gradient can be obtained as follows. Consider the auxiliary function $g: \mathbb{R} \to \mathbb{R}$, given by $g(t) = f(Z + tP)$, for any arbitrary matrix $P$. From basic calculus we know that $g'(0) = \langle \nabla f(Z), P \rangle_F$. After simple algebraic manipulations, it follows that

$$g'(0) = 2\langle AZB - D, APB \rangle_F = 2\langle A^T(AZB - D)B^T, P \rangle_F,$$

and hence the gradient of $f$ is given by

$$\nabla f(Z) = 2A^T(AZB - D)B^T = -2A^T\bar{R}B^T.$$

According to the necessary and sufficient condition of minimization problem established in Kelly [26]. Let $f(Z)$ be a continuous, differential and convex function on subspace $\mathcal{P}$.

Then there exist $Z_0 \in \mathcal{P}$ such that $f(Z_0) = \min_{Z \in \mathcal{P}} f(Z)$, if and only if the projection of the gradient $\nabla f(Z_0)$ onto $\mathcal{P}$ is 0, we know that $\bar{Y}$ is a solution of the Problem I if and only if $\mathcal{A}(A^T\bar{R}B^T) = 0$. 

DOI: 10.1002/nla
3.3. The extended CGLS algorithm and its basic properties

In this subsection, we will construct an extended version of CGLS algorithm, and describe its some basic properties. Finally, we will show that it terminates within finitely many steps.

Algorithm 1. For an arbitrary initial matrix \( Y_0 \in \mathbb{C}^{n \times n}_s \), compute

\[
\text{Step1:} \quad R_0 = D - AY_0B; \quad P_0 = \mathcal{A}(A^T R_0 B^T); \quad Q_0 = P_0;
\]

\[
\text{Step2:} \quad \text{If} \quad \|P_k\|^2 = 0, \text{then stop; else compute}
\]

\[
\text{Step3:} \quad (a) \quad \alpha_k = \frac{\|P_k\|^2}{\|A Q_k B\|^2}; \quad Y_{k+1} = Y_k + \alpha_k Q_k;
\]

\[
(b) \quad R_{k+1} = R_k - \alpha_k A Q_k B;
\]

\[
(c) \quad P_{k+1} = \mathcal{A}(A^T R_{k+1} B^T) = P_k - \alpha_k \mathcal{A}(A^T A Q_k B B^T);
\]

\[
(d) \quad \beta_k = \frac{\|P_{k+1}\|^2}{\|P_k\|^2}; \quad Q_{k+1} = P_{k+1} + \beta_k Q_k;
\]

\[
\text{Step4:} \quad \text{Let} \quad k := k + 1 \text{ and go to step 2.}
\]

Remark 3

\( P_{k+1} = \mathcal{A}(A^T R_{k+1} B^T) \) implies \( P_{k+1} \in \mathbb{C}^{n \times n}_s \) for all \( k \). Since \( Q_0 = P_0 \) by assumption and \( P_{k+1} \in \mathbb{C}^{n \times n}_s \) for all \( k \), the second equation in substep (d) and induction imply that \( Q_{k+1} \in \mathbb{C}^{n \times n}_s \). Since \( Y_0 \in \mathbb{C}^{n \times n}_s \) by assumption and \( Q_k \in \mathbb{C}^{n \times n}_s \) for all \( k \), the second equation in substep (a) and induction imply that \( Y_{k+1} \in \mathbb{C}^{n \times n}_s \). \( R_k \) is the residual of Equation (6) where \( k = 0, 1, 2, \ldots \).

Now we consider the computational complexity of our algorithm. Note that the dominant costs of Algorithm 1 are to do Step 3. Other costs are negligible. We first consider the cost of the linear projection operator \( \mathcal{A} \). Define matrix \( G \in \mathbb{C}^{n \times n}_s \), in which \( \tilde{X}_q = I_q \), where \( I_q \) denote the \( q \times q \) order identity matrix. Then from Lemma 1, for any arbitrary matrix \( Z \in \mathbb{C}^{n \times n}_s \),

\[ \mathcal{A}(Z) = \frac{Z + J_n Z J_n}{2} - G \frac{Z + J_n Z J_n}{2} G. \quad (7) \]

Since \( J_n \) has only 1 nonzero entries per row and per column, \( G \) has only \( p \) nonzero entries, by the conventional algorithm for computing matrix–matrix multiplication, it requires \( O(n^2) \) operations to compute \( \mathcal{A}(Z) \). Substep (a) requires \( O(n^2 m + mns) \) operations. Substep (b) obviously requires \( O(n^2 m + mns) \) operations. Substep (c) requires \( O(n^2 s + mns) \) operations if choosing iterative formula \( P_{k+1} = \mathcal{A}(A^T R_{k+1} B^T) \). While substep (c) requires \( O(n^2 (m+s) + n^3) \) operations if choosing \( P_{k+1} = P_k - \alpha_k \mathcal{A}(A^T A Q_k B B^T) \) in the first cycle, it only requires \( O(n^3) \) in the following cycles since \( A^T A \) and \( B^T B \) could be saved. Substep (d) obviously requires \( O(n^2) \) operations. From above, we know that the total cost of the algorithm is \( O(n^3 + n^2 (m+s) + mns) \) operations per cycle.

Next we will show the properties of the iteration method by induction.

Lemma 5

For matrices \( P_i \) and \( Q_i \) generated by Algorithm 1, if there exists a positive integer \( k \) such that \( \|P_i\|^2 \neq 0, \alpha_i \neq 0, \) and \( \alpha_i \neq \infty \) for all \( i = 0, 1, 2, \ldots, k \), then

\[
(1) \quad \langle P_i, P_j \rangle = 0, \quad i, j = 0, 1, 2, \ldots, k, \quad i \neq j,
\]

\[
(2) \quad \langle A Q_i B, A Q_j B \rangle = 0, \quad i, j = 0, 1, 2, \ldots, k, \quad i \neq j,
\]

\[
(3) \quad \langle Q_i, P_j \rangle = 0, \quad 0 \leq i < j \leq k.
\]
Proof

We know that \( \langle M, N \rangle = \langle N, M \rangle \) for an arbitrary \( M, N \in \mathbb{R}^{n \times n} \). Therefore, we need to only prove that \( \langle P_i, P_j \rangle = 0 \), \( \langle AQ_i, B, AQ_j B \rangle = 0 \), and \( \langle Q_i, P_j \rangle = 0 \) for \( 0 \leq i < j \leq k \). This requires a two-step induction. First, for convenience of discussion later, we introduce the following equality from (4) and use it in a similar way in several other instances.

\[
\langle P_0, \mathcal{A}(A^T A Q_0 B B^T) \rangle = \langle P_0, A^T A Q_0 B B^T \rangle = \langle A P_0 B, A Q_0 B \rangle.
\]

For \( k = 1 \),

\[
\langle P_0, P_1 \rangle = \langle P_0, P_0 - \alpha_0 \mathcal{A}(A^T A Q_0 B B^T) \rangle = \langle P_0, P_0 \rangle - \alpha_0 \langle A P_0 B, A Q_0 B \rangle = \|P_0\|^2 - \alpha_0 \|A Q_0 B\|^2 = 0.
\]

\[
\langle AQ_0 B, AQ_1 B \rangle = \langle AQ_0 B, A(Q_1 + \beta_0 Q_0)B \rangle = \beta_0 \|A Q_0 B\|^2 + \langle AQ_0 B, A P_1 B \rangle
\]

\[
= \beta_0 \|A Q_0 B\|^2 + \frac{1}{\alpha_0} \langle R_0 - R_1, A P_1 B \rangle
\]

\[
= \beta_0 \|A Q_0 B\|^2 + \frac{1}{\alpha_0} \langle \mathcal{A}(A^T(R_0 - R_1)B^T), P_1 \rangle
\]

\[
= \beta_0 \|A Q_0 B\|^2 + \frac{1}{\alpha_0} (\langle P_0, P_1 \rangle - \langle P_1, P_1 \rangle)
\]

\[
= \beta_0 \|A Q_0 B\|^2 - \frac{1}{\alpha_0} \langle P_1, P_1 \rangle = 0.
\]

\[
\langle Q_0, P_1 \rangle = \langle P_0, P_1 \rangle = 0.
\]

If the conclusions \( \langle P_i, P_s \rangle = 0 \), \( \langle AQ_i B, AQ_s B \rangle = 0 \), \( \langle Q_i, P_s \rangle = 0 \) hold for all \( i \leq s - 1 \), \( 0 < s < k \), then

\[
\langle P_i, P_{s+1} \rangle = \langle P_i, P_s - \alpha_s \mathcal{A}(A^T A Q_s B B^T) \rangle = -\alpha_s \langle A P_i B, A Q_s B \rangle
\]

\[
= -\alpha_s \langle A(Q_i - \beta_{i-1} Q_{i-1})B, A Q_s B \rangle
\]

\[
= -\alpha_s \langle A Q_i B, A Q_s B \rangle + \alpha_s \beta_{i-1} \langle A Q_{i-1} B, A Q_s B \rangle = 0.
\]

\[
\langle AQ_i B, AQ_{s+1} B \rangle = \langle AQ_i B, A(P_{s+1} + \beta_s Q_s)B \rangle = \langle AQ_i B, A P_{s+1} B \rangle
\]

\[
= \frac{1}{\alpha_i} \langle R_i - R_{i+1}, A P_{s+1} B \rangle = \frac{1}{\alpha_i} \langle \mathcal{A}(A^T(R_i - R_{i+1})B^T), P_{s+1} \rangle
\]

\[
= \frac{1}{\alpha_i} \langle P_i - P_{i+1}, P_{s+1} \rangle = -\frac{1}{\alpha_i} \langle P_{i+1}, P_{s+1} \rangle.
\]

\[
\langle Q_i, P_{s+1} \rangle = \langle Q_i, P_s - \alpha_s \mathcal{A}(A^T A Q_s B B^T) \rangle = -\alpha_s \langle AQ_i B, AQ_s B \rangle = 0.
\]

Algorithm 1 shows that the matrix \( Q_i (i \leq s) \) can be expressed as

\[
Q_i = P_i + \beta_{i-1} Q_{i-1} = P_i + \beta_{i-1} P_{i-1} + \beta_{i-1} \beta_{i-2} P_{i-2} + \cdots + \beta_{i-1} \cdots \beta_1 \beta_0 P_0,
\]

which implies that

\[
\langle P_i, Q_i \rangle = \langle P_i, P_i + \beta_{i-1} P_{i-1} + \beta_{i-1} \beta_{i-2} P_{i-2} + \cdots + \beta_{i-1} \cdots \beta_1 \beta_0 P_0 \rangle = \|P_i\|^2.
\]
Then for $i = s$

$$
\langle P_s, P_{s+1} \rangle = \langle P_s, P_s - \alpha_s \langle A^T A Q_s B B^T \rangle \rangle = \langle P_s, P_s \rangle - \alpha_s \langle A P_s B, A Q_s B \rangle = \| P_s \|^2 - \alpha_s \langle A (Q_s - \beta_{s-1} Q_{s-1}) B, A Q_s B \rangle = \| P_s \|^2 - \alpha_s \| A Q_s B \|^2 = 0.
$$

$$
\langle A Q_s B, A Q_{s+1} B \rangle = \langle A Q_s B, A (P_{s+1} + \beta_s Q_s) B \rangle = \beta_s \| A Q_s B \|^2 + \langle A Q_s B, A P_{s+1} B \rangle = \beta_s \| A Q_s B \|^2 + \frac{1}{\alpha_s} \langle R_s - R_{s+1}, A P_{s+1} B \rangle = \beta_s \| A Q_s B \|^2 + \frac{1}{\alpha_s} \langle A^T (R_s - R_{s+1}) B^T, P_{s+1} \rangle = \beta_s \| A Q_s B \|^2 + \frac{1}{\alpha_s} \langle \mathcal{A} (A^T (R_s - R_{s+1}) B^T), P_{s+1} \rangle = \beta_s \| A Q_s B \|^2 + \frac{1}{\alpha_s} \langle P_s, P_{s+1} \rangle = \langle P_s, P_{s+1} \rangle = \| P_s \|^2 = \| P_s \|^2.
$$

Then the conclusion $\langle P_s, P_{s+1} \rangle = 0$ and the assumption $\langle A Q_s B, A Q_s B \rangle = 0$ show that $\langle A Q_s B, A Q_{s+1} B \rangle = 0$ for all $i \leq s - 1$. By induction, we know that Equation (3) holds for all $0 \leq i < j \leq k$, and Equations (1) and (2) hold for all $i, j = 0, 1, 2, \ldots, k, i \neq j$.  

Lemma 5 shows that $P_0, P_1, \ldots$ generated by Algorithm 1 are orthogonal to each other in the finite-dimensional space $\mathbb{C}^{\mathbb{S}_{n \times n}}$, and it is easy to show that the sequences $Q_0, Q_1, \ldots$ are linearly independent. Noting that the dimension of the subspace $\mathbb{C}^{\mathbb{S}_{n \times n}}$ is $(n + 1)/2$ when $n$ is odd or $n/2$ when $n$ is even, hence the dimension of the subspace $\mathbb{C}^{\mathbb{S}_{n \times n}}$ is $r = (n + 1)/2 - (q + 1)/2$ when $n, q$ are odd or $r = n/2 - q/2$ when $n, q$ are even. Hence, there exists a positive integer $t \leq r$, such that $P_t = 0$, i.e. the iteration will be terminated at most $r$ steps. Moreover, for all $k \leq t$, we have

$$
\text{span}[P_0, P_1, \ldots P_k] = \text{span}[Q_0, Q_1, \ldots Q_k].
$$

It is worthwhile to note that the conclusions of Lemma 5 may be false without the assumptions $\alpha_i \neq 0$ and $\alpha_i \neq \infty$. In particular, $\alpha_i = \infty$ means that the iteration breaks down before $P_i = 0$ for $i < t$. Hence, it is necessary to consider the cases that $\alpha_i = 0$ or $\alpha_i = \infty$.

If $\alpha_i = 0$, which implies $\| P_i \|^2 = 0$, then we have $P_i = 0$.

If $\alpha_i = \infty$, which implies $A Q_i B = 0$, making inner product with both side by $R_i$, yields

$$
0 = \langle 0, R_i \rangle = \langle A Q_i B, R_i \rangle = \langle Q_i, \mathcal{A} A^T R_i B^T \rangle = \langle Q_i, P_i \rangle = \| P_i \|^2,
$$

which also means $P_t = 0$.

This discussion shows that if there is a positive integer $i$ such that $\alpha_i = 0$ or $\alpha_i = \infty$, then the corresponding matrix $Y_i$ is just the solution of Problem A.

From Lemma 5 and the discussion about the coefficient $\alpha_i$, we can conclude the following theorem.
\textbf{Theorem 1}
For any initial matrix $Y_0 \in \mathbb{C}\mathbb{S}_*^{n \times n}$, the matrix sequence $Y_k$ generated by Algorithm 3.3 yields a solution of Problem $A$ within finitely many steps in exact arithmetic.

Next, we will show that if we choose a special kind of initial matrix, we can obtain the unique least-norm solution of Problem $A$. To this end, we first define the following matrix set:

\[ \mathcal{F} = \{ Y | Y = \mathcal{A}(A^T H B^T), \forall H \in \mathbb{R}^{n \times n} \}. \]

Clearly, $\mathcal{F}$ is a linear subspace of $\mathbb{C}\mathbb{S}_*^{n \times n}$.

\textbf{Lemma 6}
Suppose $\tilde{Y}$ is a solution of Problem $A$, $\tilde{R}$ is the corresponding residual. Then any solution of Problem $A$ can be expressed as $Y + \tilde{Y}$, where $\tilde{Y} \in \mathbb{C}\mathbb{S}_*^{n \times n}$ and

\[ A\tilde{Y}B = 0. \]  

\[ \| A\tilde{Y}B - D \|^2 = \| A(\tilde{Y} + \tilde{Y})B - D \|^2 = \| A\tilde{Y}B - \tilde{R} \|^2 = \| A\tilde{Y}B \|^2 + \| \tilde{R} \|^2, \]  

which implies that $A\tilde{Y}B = 0$.

Conversely, if $\tilde{Y} \in \mathbb{C}\mathbb{S}_*^{n \times n}$ and $A\tilde{Y}A^T = 0$, then

\[ \| A(\tilde{Y} + \tilde{Y})B - D \|^2 = \| A\tilde{Y}B - D + A\tilde{Y}B \|^2 = \| A\tilde{Y}B - D \|^2, \]

which means that $\tilde{Y} + \tilde{Y}$ a solution of Problem $A$. \hfill \Box

\textbf{Theorem 2}
If we choose the initial matrix $Y_0 \in \mathcal{F}$, especially, let $Y_0 = 0 \in \mathbb{R}^{n \times n}$, by Algorithm 1 we can obtain the least-norm solution of Problem $A$.

\textbf{Proof}
From Algorithm 1 and Theorem 1, if $Y_0 \in \mathcal{F}$, we can obtain a solution $\tilde{Y}^*$ of Problem $A$ with finitely many steps and there exists a matrix $\tilde{H} \in \mathbb{R}^{n \times n}$ such that the solution $\tilde{Y}^*$ can be represented as $\tilde{Y}^* = \mathcal{A}(A^T \tilde{H} A)$. By Lemma 6 we know that arbitrary solution of Problem $A$ can be expressed as $Y^* + \tilde{Y}$, where $\tilde{Y} \in \mathbb{C}\mathbb{S}_*^{n \times n}$ and satisfies (10). Then

\[ \langle \tilde{Y}^* - \tilde{Y} \rangle = \langle \mathcal{A}(A^T \tilde{H} A) \tilde{Y} \rangle - \langle A^T \tilde{H} A, Y \rangle = \langle \tilde{H}, A\tilde{Y}A^T \rangle = 0. \]

Therefore

\[ \| \tilde{Y}^* + \tilde{Y} \|^2 = \| \tilde{Y}^* \|^2 + \| \tilde{Y} \|^2 \geq \| \tilde{Y}^* \|^2, \]

which implies $\tilde{Y}^*$ is the least-norm solution of Problem $A$. \hfill \Box

\textbf{Remark 4}
Since the solution set of Problem $A$ is nonempty, and from Lemma 6 this solution set is a closed convex cone, hence it is certain that there exists a unique least-norm solution. If matrix $\tilde{Y} \in \mathcal{F}$ is a solution of Problem $A$, then it is just the unique least-norm solution.

\subsection*{3.4. An alternative proof of Theorem 2}
In this subsection, we will give an alternative proof of Theorem 2. This proof is based on the theory of linear system and vec operators. First we will give the following lemma, which we need in this subsection.

\textbf{Lemma 7}
Suppose that the minimum residual problem $\| Mx - b \|_2 = \min$ has a solution $x^* \in \mathcal{A}(M^T)$, then $x^*$ is the unique least Frobenius norm solution of the problem.
Proof
We decompose the matrix $M$ by singular value decomposition (SVD):

$$M = U \Sigma V^T = U_1 \Sigma V_1^T,$$

where $U = (U_1, U_2), V = (V_1, V_2)$ are orthogonal matrices. The Moore–Penrose generalized inverse of the matrix $M$ is $M^+ = V_1 \Sigma^{-1} U_1^T$, and the general solution of the minimum residual problem

$$\|Mx - b\|_2 = \text{min},$$

where $z$ is an arbitrary vector of suitable dimension.

Since $M^+ = V_1 \Sigma^{-1} U_1^T \in \mathbb{R}(V_1), (I - M^+ M)z = V_2 V_2^T \in \mathbb{R}(V_2)$, and $V_1, V_2$ are orthogonal to each other, then $M^+ b$ is the unique least Frobenius norm solution of the minimum residual problem

$$\|Mx - b\|_2 = \text{min}.$$

On the other hand, since $M^T = V_1 \Sigma U_1^T$, and the solution $x^* \in \mathbb{R}(M^T)$, then $x^* \in R(V_1)$. Therefore, $x^*$ is the unique least Frobenius norm solution of the minimum residual problem

$$\|Mx - b\|_2 = \text{min}.$$

An alternative proof of Theorem 2: From Algorithm 1 and Theorem 1, if we let $Y_0 = A^T H B^T$, where $H$ is an arbitrary matrix in $\mathbb{R}^{p \times n}$, we can obtain the solution $Y^*$ of Problem A with finitely many steps, and the solution $Y^*$ can be represented as

$$Y^* = A^T H B^T.$$

In the sequel, we will prove that $Y^*$ is the least-norm solution of Problem A.

Consider the following minimum residual problem:

$$\min_{Z \in \mathbb{R}^{p \times n}} \left\| \begin{pmatrix} AZB - AGZGB \\ AJ_n ZJ_n B - AGJ_n ZJ_n GB \end{pmatrix} - \begin{pmatrix} D \\ D \end{pmatrix} \right\|.$$

(13)

Obviously, the solvability of Problem A is equivalent to that of the minimum residual problem (13). In order to prove that $Y^*$ is the least-norm solution of Problem A, it is enough to prove that $Y^*$ is the least-norm solution of the minimum residual problem (13).

Denote $\text{vec}(Z) = z, \text{vec}(Y^*) = y^*, \text{vec}(D) = d, \text{vec}(H) = h$. Then the minimum residual problem (13) is equivalent to the following minimum residual problem:

$$\min_{z \in \mathbb{R}^n} \left\| \begin{pmatrix} B^T \otimes A - (GB)^T \otimes (AG) \\ (J_nB)^T \otimes (AJ_n) - (J_nGB)^T \otimes (AGJ_n) \end{pmatrix} z \right\|.$$

(14)

Noting that $G$ and $J_n$ are permutably, from (7)

$$\tilde{Y}^* = A^T H B^T = \frac{A^T H B^T + J_n A^T H B^T J_n}{2} - \frac{G A^T H B^T + J_n A^T H B^T J_n}{2} G,$$

then

$$2\tilde{Y}^* = [B \otimes A^T + (J_nB) \otimes (AJ_n)^T]h - [(GB) \otimes (AG)^T + (J_nGB) \otimes (AGJ_n)^T]h$$

$$= \left( B^T \otimes A - (GB)^T \otimes (AG) \right)^T \left( J_nB \otimes (J_nA) - (J_n GB)^T \otimes (AGJ_n) \right) h$$

$$\in \mathbb{R} \left( B^T \otimes A - (GB)^T \otimes (AG) \right)^T \left( J_nB \otimes (J_nA) - (J_n GB)^T \otimes (AGJ_n) \right).$$
Hence, \( \tilde{Y}^* \) is the least-norm solution of the minimum residual problem (14) from Lemma (7). Since vec operator is isomorphic, then \( Y^* \) is also the unique least-norm solution of the minimum residual problem (13); furthermore, \( \tilde{Y}^* \) is the unique least-norm solution of Problem A.

3.5. The minimization property of Algorithm 1

In this subsection, we will characterize the minimization property of Algorithm 1, which ensures that the Algorithm 1 converges smoothly.

**Theorem 3**

For any initial matrix \( Y_0 \in \mathbb{C}^{n \times n} \), the matrix \( Y_k \) generated by Algorithm 1 at the \( k \)th iteration step satisfies the following minimization problem:

\[
\|AY_k B - D\|^2 = \min_{Y \in Y_0 + \text{span}\{Q_0, Q_1, \ldots, Q_{k-1}\}} \|AY - D\|^2.
\]

**Proof**

For arbitrary matrix \( Y \in Y_0 + \text{span}\{Q_0, Q_1, \ldots, Q_{k-1}\} \), there exists a set of real number \( \{t_i\}_{i=1}^{k-1} \) such that \( Y = Y_0 + \sum_{i=1}^{k-1} t_i Q_i \). Denoting

\[
g(t_0, \ldots, t_{k-1}) = \left\| A \left( Y_0 + \sum_{i=0}^{k-1} t_i Q_i \right) B - D \right\|^2,
\]

by the conclusion (2) in Lemma 5, we have

\[
g(t_0, \ldots, t_{k-1}) = \left\| AY_0 B - D + \sum_{i=0}^{k-1} t_i A Q_i B \right\|^2
\]

\[
= \|R_0\|^2 + \sum_{i=0}^{k-1} \sum_{j=0}^{k-1} t_i t_j \langle AQ_i B, AQ_j B \rangle - 2 \sum_{i=0}^{k-1} t_i \langle AQ_i B, R_0 \rangle
\]

\[
= \|R_0\|^2 + \sum_{i=0}^{k-1} t_i^2 \|AQ_i B\|^2 - 2 \sum_{i=0}^{k-1} t_i \langle AQ_i B, R_0 \rangle.
\]

The third equation of step 3 in Algorithm 1 shows that the initial residual \( R_0 \) can be expressed as

\[
R_0 = R_1 + z_{i-1} AQ_{i-1} B + \cdots + z_0 AQ_0 B.
\]

Since \( g(t_0, \ldots, t_{k-1}) \) is a continuous and differentiable function with respect to the \( k \) variable \( t_i (i = 0, 1, \ldots, k-1) \), it follows that \( g(t_0, t_1, \ldots, t_{k-1}) = \min \) if and only if

\[
\frac{\partial g(t_0, t_1, \ldots, t_{k-1})}{\partial t_i} = 0.
\]

Therefore (2) in Lemma 5 and (8) implies that

\[
t_i = \frac{\langle AQ_i B, R_0 \rangle}{\|AQ_i B\|^2} = \frac{\langle AQ_i B, R_i \rangle}{\|AQ_i B\|^2} = \frac{\langle Q_i, \mathcal{A}(A^T R_i B^T) \rangle}{\|AQ_i B\|^2} = \frac{\langle Q_i, P_i \rangle}{\|AQ_i B\|^2} = \frac{\|P_i\|^2}{\|AQ_i B\|^2} = \gamma_i.
\]

Since

\[
\min_{t_i} g(t_0, t_1, \ldots, t_{k-1}) = \min_{Y \in Y_0 + \text{span}\{Q_0, Q_1, \ldots, Q_{k-1}\}} \|AY - D\|^2,
\]

this completes the proof. \( \square \)
Theorem 3 shows that the approximation solution \( Y_k \) minimizes the residual norm in the affine subspace \( Y_0 + \text{span}(Q_0, Q_1, \ldots, Q_{k-1}) \) for all initial matrix \( Y_0 \) within \( \mathbb{C}^{\mathcal{S}_{r \times n}} \). Furthermore, since \( Y_{k-1} \in Y_0 + \text{span}(Q_0, Q_1, \ldots, Q_{k-1}) \), we have
\[
\|AY_k B - D\| \leq \|AY_{k-1} B - D\|,
\]
herefore the sequence \( \|AY_0 B - D\|, \|AY_1 B - D\|, \ldots \) is monotonically decreasing. The descent property of the residual norm ensures that the Algorithm I possesses fast and smooth convergence.

4. THE SOLUTION OF PROBLEM II

The optimal approximation problem occurs frequently in experimental design, see for instance [15]. Here the matrix \( X_s \) may be obtained from experiments, but it may not satisfy the centrosymmetric requirement and minimum residual requirement. The optimal approximation \( \widehat{X} \) is the matrix that satisfies the centrosymmetric and the minimum residual restriction, and is closed to the given \( X \) in Frobenius norm (may be spectral norm or others). About the optimal approximation problem, we refer the reader to references [8, 7, 9, 21, 11, 27, 28].

For the optimal approximation problem—Problem II, there certainly exists a unique solution since the solution set of the minimum residual problem—Problem I is a nonempty closed convex cone. Noting that for arbitrary matrix \( X \) the solution set of the minimum residual problem—Problem I is a nonempty closed convex cone. All the tests are performed using Matlab 7.0, which has a machine precision of around \( 10^{-16} \). Because of calculation errors, the iteration will not stop within finitely many steps. Hence, we regard the approximation solution \( Y_k \) as the solution of Problem A if the corresponding \( P_k \) satisfies \( \|P_k\| \leq 10^{-10} \).

Example 1
Let matrices
\[
A = \begin{pmatrix} \text{hilb}(r/2) & \text{ones}(r/2) \\ \text{hankel}(1: r/2) & \text{zeros}(r/2) \end{pmatrix}, \quad B = A^T, \quad C = \begin{pmatrix} \text{toeplitz}(1: r/2) & \text{ones}(r/2) \\ \text{zeros}(r/2) & \text{ones}(r/2) \end{pmatrix},
\]
where \( r \) is a constant that determines the magnitudes of the dimension of the system, \( \text{hilb}(n) \) denotes the \( n \)th Hilbert matrix, \( \text{toeplitz}(1:n) \) and \( \text{hankel}(1:n) \) denote the \( n \)th order Toeplitz matrix and
Hankel matrix whose first row is (1, 2, \ldots, n), respectively, and ones(n) and zeros(n), respectively, denote the n × n matrices whose all elements are one and zero. We denote \( X_c(q) = \text{toeplitz}(1:8) \in \mathbb{C}^{8 \times 8} \), which is the given central principal submatrix.

We first compute \( D = C - \bar{A}X \), where \( \bar{X} \) is the matrix defined by (5) with compatible dimension. In Table I, we list our numerical results with dimensions from \( r = 12 \) to \( r = 192 \). If we set the initial iterative matrices be zeros\( (r) \), then the least Frobenius norm least-squares solution of the inconsistent matrix equation \( AYB = D \) over \( \mathbb{C}S_n^{n \times n} \) can be obtained by Algorithm 1. In this table, we list rank\( (A) \), cond\( (A) \), iteration numbers \( (k) \), CPU times, iterative terminate condition \( \| P_k \| \), and residual norm for different values of \( r \) with the stopping criteria \( \| P_k \| \leq 10^{-10} \).

It is interesting that the convergence of Algorithm 1 is very fast even if the condition numbers of \( A \) are very large. These phenomena often occur in our tests.

To save space, we shall not report the dates of \( Y_k \), but will make them available upon request. Once these \( Y_k \) are obtained, the corresponding least-squares solution \( \tilde{X} \) to \( AXB = C \) over \( \mathcal{S} \) (defined by (1)) can be expressed as \( \tilde{X} = Y_k + \bar{X} \). In Figure 1, we characterize the convergence curves for the Frobenius norm of \( P_k \) when \( r = 24, 48, 72, 96 \). The figure shows that all \( \| P_k \| \)s are oscillating back and forth and gradually approach to zero as iterative process. We also characterize the convergence curve (see Figure 2) for the Frobenius norm of the residual when \( r = 48 \). The result in this figure shows clearly that the residual norm of Algorithm 1 is monotonically decreasing, which is in accordance with the theory established in this paper.

Let \( r = 48 \), without loss of generality, let the preliminary estimate \( X_\ast \) be given by ones\( (48) \) but replace its 8 \times 8 central principal submatrix by toeplitz\( (1:8) \), which implies that \( X_\ast \in \mathcal{S} \). To find the optimal approximation matrix \( \hat{X} \) associated with \( X_\ast \), let \( Z = X - X_\ast \), \( D = C - A X_\ast B \), and the initial iterative matrix \( Z_0 = \text{zeros}(48) \). By using Algorithm 1 and iterating 1053 steps, we can obtain the least-squares least Frobenius norm solution \( Z_\ast \) of \( AZB = D \) over \( \mathbb{C}S_{48}^{48 \times 48} \). Then the
NUMERICAL SOLUTIONS OF $AXB = C$ FOR CENTROSYMMETRIC MATRIX

Figure 2. Convergence curve for the Frobenius norm of the residual when $r = 48$.

Figure 3. Y axis: $\log_{10} \| \hat{X}(\epsilon) - \hat{X} \|$ (marked by □), $\log_{10} \| \hat{X}(\epsilon) - \hat{X} \|$ (marked by △), $\log_{10} \| \hat{X}(\epsilon) - \hat{X}(\epsilon) \|$ (marked by +), $\log_{10} \| \hat{X} - \hat{X} \|$ (marked by −) versus $\log_{10} \epsilon(\log_{10}(\epsilon))$.

optimal approximation $\hat{X}$ can be expressed as $\hat{X} = Z_\ast + X_\ast$. To save space, we shall also not report the date of $\hat{X}$, but will make them available upon request.

Give the preliminary estimate $X_\ast$ a perturbation $X_\ast(\epsilon) = X_\ast + \epsilon F$, where $F \in \mathbb{C}^{48 \times 48}$ could be randomly constructed. By Algorithm 1, we also obtain $\hat{X}(\epsilon)$. In Figure 3, we plot the following four quantities for $\epsilon$ from $10^{-8}$ to $10^6$: $\log_{10} \| X_\ast(\epsilon) - X_\ast \|$ (marked by ‘□’); $\log_{10} \| \hat{X}(\epsilon) - \hat{X} \|$ (marked by ‘△’); $\log_{10} \| \hat{X} - X_\ast \|$ (marked by ‘+’); and $\log_{10} \| \hat{X}(\epsilon) - X_\ast(\epsilon) \|$ (marked by ‘−’).

We see from Figure 3 that:

(I) $\hat{X}(\epsilon)$ gradually approaches to $\hat{X}$ as $\epsilon$ goes to zeros, which implies that this algorithm is numerical stable for the approximation problem.

(II) $\log_{10} \| \hat{X}(\epsilon) - \hat{X} \| \leq \log_{10} \| X_\ast(\epsilon) - X_\ast \|$ as $\epsilon$ varies from $10^6$ to $10^{-8}$, and the two curves are proportional decline, which means that there exists a constant $\alpha(0 < \alpha < 1)$ such that

$$\log_{10} \| \hat{X}(\epsilon) - \hat{X} \| = \alpha \cdot \log_{10} \| X_\ast(\epsilon) - X_\ast \|.$$
(III) \(\|\tilde{X}(\epsilon) - X_\ast(\epsilon)\|\) approaches to \(\|\tilde{X} - X_\ast\|\) as \(\epsilon\) approaches zero, which implies that this algorithm is numerically stable for the right-hand side matrix \(D\). Actually, \(\tilde{X} - X_\ast\) is the unique least Frobenius norm solution of \(\min_{Z \in \mathbb{C}^{48 \times 48}} \|AZB - \hat{D}\|\), whereas \(\tilde{X}(\epsilon) - X_\ast(\epsilon)\) is the unique least Frobenius norm solution of \(\min_{Z \in \mathbb{C}^{48 \times 48}} \|AZB - \hat{D}(\epsilon)\|\), where \(\hat{D}(\epsilon) = \hat{D} - \epsilon AFB\).

ACKNOWLEDGEMENTS

This work is supported by National Natural Science Foundation of China under grant number 10861005 and Provincial Natural Science Foundation of Guangxi under grant number 0991238.

REFERENCES
