AN ITERATIVE ALGORITHM FOR THE BEST APPROXIMATE (P,Q)-ORTHOGONAL SYMMETRIC AND SKEW-SYMMETRIC SOLUTION PAIR OF COUPLED MATRIX EQUATIONS

Fatemeh Panjeh Ali Beik[*,§] **and Davod Khojasteh Salkuyeh** [†]

[§]*Department of Mathematics, Vali-e-Asr University of Rafsanjan,*
P. O. Box 518, Rafsanjan, Iran
e-mail: f.beik@vru.ac.ir, beik.fatemeh@gmail.com

[†]*Faculty of Mathematical Sciences, University of Guilan, Rasht, Iran*
e-mail: salkuyeh@gmail.com

Abstract. This paper deals with developing a robust iterative algorithm to find the least-squares (P,Q)-orthogonal symmetric and skew-symmetric solution sets of the generalized coupled matrix equations. To this end, first, some properties of these type of matrices are established. Furthermore, an approach is offered to determine the optimal approximate (P,Q)-orthogonal (skew-)symmetric solution pair corresponding to a given arbitrary matrix pair. Some numerical experiments are reported to confirm the validity of the theoretical results and to illustrate the effectiveness of the proposed algorithm.

Keywords: Matrix equation; Least-squares problem; (P,Q)-orthogonal symmetric matrix; (P,Q)-orthogonal skew-symmetric matrix; Iterative algorithm.

2010 AMS Subject Classification: 65F10, 15A24.

1. Introduction

Let us first introduce some symbols exploited throughout the paper. The notation $\mathbb{R}^{m\times n}$ stands for the set of all $m\times n$ real matrices. The trace and the transpose of a given matrix A are respectively denoted by $\mathrm{tr}(A)$ and A^T. The symbol $\mathbb{SOR}^{n\times n}$ refers to the set of all symmetric orthogonal matrices in $\mathbb{R}^{n\times n}$, i.e.,

$$\mathbb{SOR}^{n\times n} = \{P \in \mathbb{R}^{n\times n} \mid P^T = P,\ P^2 = I\}.$$

For two given matrices $P \in \mathbb{SOR}^{n\times n}$ and $Q \in \mathbb{SOR}^{n\times n}$, the matrix $X \in \mathbb{R}^{n\times n}$ is called (P,Q)-orthogonal symmetric if $(PXQ)^T = PXQ$ and is said to be (P,Q)-orthogonal skew-symmetric if $(PXQ)^T = -PXQ$. The set of all (P,Q)-orthogonal symmetric and (P,Q)-orthogonal skew-symmetric matrices of order n are respectively represented by $\mathbb{SR}^{PQ}_{n\times n}$ and $\mathbb{SSR}^{PQ}_{n\times n}$. A special kind of permutation matrix with ones along the secondary diagonal and zeros elsewhere is called the *exchange* matrix.

Definition 1.1. (Golub and Loan, 1996) *Let $A = (a_{ij})$ be an $n\times n$ matrix. The matrix A is said to be a persymmetric matrix if*

$$a_{ij} = a_{n-j+1,n-i+1}$$

[*]Corresponding author.

for all $1 \leq i, j \leq n$. Or equivalently, A is called persymmetric if $AJ = JA^T$ where J is the exchange matrix.

Assume that $A \in \mathbb{R}^{n\times n}$ and J is the exchange matrix of order n. It can be verified that JA^T, JAJ and A^TJ are versions of the matrix A that have been rotated anti-clockwise by $90, 180$ and 270 degrees. Moreover, JA, JA^TJ, AJ and A^T are versions of the matrix A that have been reflected in lines at 0, 45, 90 and 135 degrees to the horizontal measured anti-clockwise.

Remark 1.2. *Assume that $A \in \mathbb{R}^{n\times n}$ and $J \in \mathbb{R}^{n\times n}$ is the exchange matrix. Note that A is persymmetric if $(AJ)^T = AJ$. In view of the fact that $J \in \mathbb{SOR}^{n\times n}$, it can be concluded that every persymmetric matrix is a (I, J)-orthogonal symmetric matrix where I stands for the identity matrix of order n.*

Definition 1.3. *A $n \times n$ matrix H is said to be (skew-)Hamiltonian if $\hat{J}H$ is (skew-)symmetric where $n = 2k$ and*

$$\hat{J} = \begin{pmatrix} 0 & I_k \\ -I_k & 0 \end{pmatrix}. \tag{1.1}$$

Hamiltonian and skew-Hamiltonian matrices have a key role in engineering, such as in linear–quadratic optimal control, H_∞ optimization and solving algebraic Riccati equations; see (Laub, 1991) and (Zhou et al., 1995). Eigenvalues and eigenvectors of a Hamiltonian matrix are important for describing a physical system, in fact the real eigenvalues correspond to the possible energy levels of the system and the related eigenvectors express the corresponding state; e.g., see (Reichl, 2013). In view of the importance of constructing (skew-)Hamiltonian matrices specially in inverse problems, solving matrix equations and the structured inverse eigenvalue problem over Hamiltonian matrices have been investigated by some researchers; see (Qian and Tan, 2013), (Zhang et al., 2005) and the references therein.

Remark 1.4. *Note that $\hat{J}^T = \hat{J}^{-1} = -\hat{J}$ and $\hat{J}H$ is (skew-)symmetric if and only if $H\hat{J}$ is (skew-)symmetric.*

Without loss of generality, we say H is a (skew-)Hamiltonian matrix if $(H\hat{J})^T = H\hat{J}$ $((H\hat{J})^T = -H\hat{J}\ )$.

Remark 1.5. *We would like to comment here that for a given matrix H, it is not difficult to see that H is a (skew-)Hamiltonian matrix if HJ is $(I, \tilde{J})$-orthogonal (skew-)symmetric matrix where*

$$\tilde{J} = \begin{pmatrix} -I_k & 0 \\ 0 & I_k \end{pmatrix} \qquad \text{and} \qquad J = \begin{pmatrix} 0 & I_k \\ I_k & 0 \end{pmatrix} \qquad n = 2k.$$

This follows from the fact that $J\hat{J} = \tilde{J}$. Notice that $J, \tilde{J} \in \mathbb{SOR}^{n\times n}$.

Linear matrix equations play a cardinal role in various fields, such as control and system theory, stability theory and some other fields of applied and pure mathematics; see Ding and Chen (2006), Ding et al. (2008), Hajarian and Dehghan (2011), Wu et al. (2013) and references therein. In recent years, several research works have discussed the way of obtaining explicit forms of the solution for different kinds of matrix equations. For example, Wu et al. (2011b) have presented an approach to resolve a general class of Sylvester-polynomial-conjugate matrix equations which include the Yakubovich-conjugate matrix equation as a special case. In

(Wu et al., 2012), several explicit parametric solutions have been derived for the generalized Sylvester-conjugate matrix equation. In (Zhao et al., 2010), an analytical expression is derived for the optimal approximate solution in the least-squares (P, Q)-orthogonal symmetric solution set of the matrix equation $A^T XB = Z$ with respect to a given matrix. In addition, the authors have obtained an explicit expression of the minimum-norm least-square (P, Q)-orthogonal symmetric solution for the matrix equation $A^T XB = Z$. In (Song et al., 2014) by using of the coefficients of the characteristic polynomial of matrix Formula and the Leverrier algorithm, the explicit solutions to the Sylvester-conjugate matrix equation $AX - \bar{X}B = C$ have been constructed.

Recently, Li et al. (2014) have derived the general expression of the least-squares solution of the matrix equation $AXB+CYD = E$ with the least-norm over symmetric arrowhead matrices. Nevertheless, in practice, applying the iterative algorithms is superior to computing explicit forms due to the required high computation costs of the explicit forms. As a matter of fact, for obtaining solutions by means of the derived explicit forms, Kronecker products of matrices are required to be explicitly formed. In addition, we need to compute the Moore-Penrose pseudoinverse which would be expensive for large-scale least-squares problems.

In the literature, the applications of iterative algorithms have been successfully examined for solving several types of matrix equations so far; for further details see Beik and Salkuyeh (2011), Beik et al. (2014), Ding and Chen (2006), Ding et al. (2008), Hajarian (2014a), Salkuyeh and Totounian (2006), Salkuyeh and Beik (2014), Wu et al. (2011a) and references therein.

The extension of the LSQR method for solving constrained matrix equations have been widely studied so far. For instance, Li and Huang (2012) have presented matrix form of the LSQR method to solve the following coupled matrix equations. Recently, Hajarian (2014b) have proposed matrix form of the LSQR algorithm for determining (least-squares) solutions of linear operator equation $\mathcal{A}(X) = B$. More recently, Karimi and Dehghan (2015) have developed the LSQR method to solve general constrained coupled matrix equations containing the transpose of the unknowns. The proposed strategy for constructing the algorithm in (Karimi and Dehghan, 2015) incorporates those given in (Li and Huang, 2012) and (Hajarian, 2014b). In addition, the performance of the proposed algorithm have been numerically compared with some of previously examined iterative schemes.

There is also a growing interest for using the idea of conjugate gradient method to construct iterative algorithms for determining the approximate least-squares solutions of various kinds of matrix equations; see for instance Beik and Salkuyeh (2013), Cai and Chen (2009), Cai et al. (2010), Hajarian (2014), Li et al. (2014), Peng et al. (2007), Peng and Liu (2009), Peng (2012), Peng (2013), Peng and Xin (2013), Peng (2015), Sheng and Chen (2007), Sheng and Chen (2010), Su and Chen (2010). Lately, Peng (2015) has developed the conjugate gradient least-squares (CGLS) method to determine the (R, S)-symmetric solutions of the following least-squares problem,

$$\sum\nolimits_{i=1}^{p} \left\| \sum\nolimits_{j=1}^{q} A_{ij} X_j B_{ij} - C_i \right\|^2 = \min .$$

More recently, Li et al. (2015) have employed the CGLS method to construct a hybrid algorithm for solving the following minimization problem

$$\frac{1}{2}\|AX-B\|^2=\min,$$

subject to the matrix inequality constraint $CXD \geq E$ over (R,S)-symmetric solutions.

Before stating the main goal of the current work, we recollect some definitions and properties which are required throughout this paper. The inner product of $X,Y \in \mathbb{R}^{m\times n}$ is defined by $\langle X,Y\rangle_F = \mathrm{tr}(Y^TX)$. The induced norm is the well-known Frobenius norm.

Suppose that $A=[a_{ij}]_{m\times s}$ and $B=[b_{ij}]_{n\times q}$ are two real matrices, the Kronecker product of the matrices A and B is defined as the $mn\times sq$ matrix $A\otimes B=[a_{ij}B]$. The "vec" operator transmutes a matrix A of size $m\times s$ to a vector $a=\mathrm{vec}(A)$ of size $ms\times 1$ by stacking its columns. In (Bernstein, 2009), it is demonstrated that the following relation holds

$$\mathrm{vec}(ABC)=(C^T\otimes A)\mathrm{vec}(B), \tag{1.2}$$

for arbitrary given matrices A,B and C with suitable dimensions.

1.1. **Motivation and main contribution.** More recently, Sarduvan et al. (2014) have focused on the solution set of the following least-squares problem

$$\|AXB-C\|^2=\min, \tag{1.3}$$

over (P,Q)-orthogonal (skew)-symmetric matrices and studied the matrix nearness problem associated with (1.3). More precisely, the explicit form for $\hat{X}\in S_{E_X}$ is determined such that

$$\|\hat{X}-X_0\|^2=\min_{X\in S_{E_X}}\|X-X_0\|^2, \tag{1.4}$$

where $X_0\in\mathbb{R}^{n\times n}$ is given and S_{E_X} stands for the solution set of (1.3). By invoking (1.2) and establishing some theoretical results, Algorithm 1 is presented to obtain the explicit form of $\hat{X}$.

As seen the computation costs of Algorithm 1 is high due to using the Kronecker product and vec(.) operator. As a matter of fact, the dimension of the matrix $\mathcal{A}$ (defined in Step 3 of Algorithm 1) can become large even in the case that the size of coefficient matrices A and B are moderate. Consequently, Steps 4, 5 and 6 consume high CPU–time (sec) to be performed. In addition, the extension of the presented approach for general types of matrix equations would be onerous. These drawbacks inspirit us to examine an iterative algorithm for solving coupled matrix equations and their corresponding least-squares problems over (P,Q)-orthogonal symmetric and (P,Q)-orthogonal skew-symmetric matrices. To this end, we first investigate some properties of the (P,Q)-orthogonal symmetric and (P,Q)-orthogonal skew-symmetric matrices which are utilized for constructing and analyzing the proposed iterative scheme. In order to elaborate our results for more general cases, we focus on the following least-squares problem

$$\left\|\begin{pmatrix} M \\ N \end{pmatrix}-\begin{pmatrix} A_1XB_1+A_2X^TB_2+C_1YD_1+C_2Y^TD_2 \\ E_1XF_1+E_2X^TF_2+G_1YH_1+G_2Y^TH_2 \end{pmatrix}\right\|=\min. \tag{1.5}$$

Algorithm 1: Proposed scheme to find the best approximation (P,Q)-orthogonal (skew)-symmetric solutions of $AXB = C$ (Sarduvan et al., 2014)

Step 1. Input the matrices $A \in \mathbb{R}^{m\times n}$, $B \in \mathbb{R}^{n\times r}$ $C \in \mathbb{R}^{m\times r}$, $P, Q \in \mathbb{SOR}^{n\times n}$ and $X_0 \in \mathbb{R}^{n\times n}$. Set $\Delta =$ '$+$' ($\Delta =$ '$-$') for finding the (P,Q)-orthogonal (skew-)symmetric solution.
Step 2. Compute the vectors

$$c_1 = \mathrm{vec}(C), \quad c_2 = \mathrm{vec}(C^T), \quad y_0 = \mathrm{vec}\left(\frac{1}{2}((PX_0Q)\Delta(PX_0Q)^T)\right).$$

Step 3. Compute

$$\mathcal{A} = \begin{pmatrix} (QB)^T \otimes (AP) \\ (AP) \otimes (QB)^T \end{pmatrix}, \quad g = \begin{pmatrix} \mathrm{vec}(C) \\ \mathrm{vec}(\Delta C^T) \end{pmatrix}, \quad A_1 = \mathcal{A}^T\mathcal{A}, \quad g_1 = \mathcal{A}^T g.$$

Step 4. Form the spectral decomposition A_1 as follows:

$$A_1 = V_{A_1}\begin{pmatrix} D & 0 \\ 0 & 0 \end{pmatrix} V_{A_1}^T,$$

where D is a diagonal matrix with diagonal entries consisting of positive eigenvalues of A_1 and V_{A_1} is an $n \times n$ orthogonal matrix formed by the corresponding linear independent eigenvectors of A_1.
Step 5. Compute $s = \mathrm{rank}(\mathcal{A}^T\mathcal{A})$, $k = \dim(\mathcal{A}^T\mathcal{A})$ and

$$M = \begin{pmatrix} \mathrm{zeros}(s) & \mathrm{zeros}(s, k-s) \\ \mathrm{zeros}(k-s, s) & \mathrm{eye}(k-s) \end{pmatrix}.$$

Step 6. Compute

$$\hat{y} = \mathcal{A}^\dagger g + (V_{\mathcal{A}^T\mathcal{A}}M)^\dagger (V_{\mathcal{A}^T\mathcal{A}}M)\left(y_0 - \mathcal{A}^\dagger g\right),$$

where the symbol † refers to the well-known Moore-Penrose generalized inverse.
Step 7. Compute $\hat{Y}$ such that $\hat{y} = \mathrm{vec}(\hat{Y})$.
Step 8. Compute $\hat{X} = P\hat{Y}Q$.

Here we point out that matrix equations which include both unknown matrix and its transpose, for instance, arise in solving time-invariant Hamiltonian systems; for further information see (Lancaster and Rozsa, 1983). They also appear in reducing a block anti-triangular matrix to a block anti-diagonal one, a reduction which is exploited in palindromic generalized eigenvalue problems (Kressner et al., 2009).

The generalized inverse eigenvalue problem for persymmetric matrices consists of determining nontrivial persymmetric matrices A and B such that

$$AX = BX\Lambda, \tag{1.6}$$

where X and the diagonal matrix Λ are known such that the columns of X are given eigenvectors and the diagonal entries of Λ are the given associated eigenvalues. As expressed in Remark 1.2, a persymmetric matrix belongs to the set of (P,Q)-orthogonal symmetric matrices. Note that the least-squares problem (1.5) incorporates the least-squares problem associated with the matrix equation (1.6). Therefore, our proposed algorithm can be applied to solve generalized persymmetric inverse eigenvalue problems (GPIEPs). Recently, Julio and Soto (2015) derived sufficient conditions for persymmetric nonnegative inverse eigenvalue problem (PNIEP) to have a solution. For a comprehensive survey on the applications of structured inverse eigenvalue problems in control theory over different kinds of matrices including persymmetric matrices, one may refer to (Chu and Golub, 2002) and the references therein.

A matrix is said to be a *generalized* (skew-)Hamiltonian matrix if HJ_1 is a (skew-)symmetric matrices where $J_1 \in \mathbb{R}^{2n\times 2n}$ is an arbitrary orthogonal skew-symmetric matrix; i.e., $J_1^{-1} = J_1^T = -J_1$.

Here it should be noticed that, without any changing, our algorithm can compute the generalized (skew-)Hamiltonian solutions of the mentioned matrix equations so that P_1 and P_2 are identity matrices and Q_1 and Q_2 are given real orthogonal skew-symmetric matrices. This follows from the fact that for an arbitrary given matrix X, we may write $X = H_1 + S_1$ such that $\langle H_1, S_1\rangle_F = 0$ where

$$H_1 = \frac{1}{2}\left(X + J_1X^TJ_1\right) \qquad \text{and} \qquad S_1 = \frac{1}{2}\left(X - J_1X^TJ_1\right).$$

It is not difficult to see that H_1 and S_1 are generalized Hamiltonian and generalized skew-Hamiltonian matrices; respectively. The convergence properties of the algorithm for determining least-squares generalized (skwe-)Hamiltonian solutions can be established with similar strategies used throughout this work, we do not discuss the details and leave them to reader.

The ensuing subsection is pertained to stating the main problems which the current paper is concerned with.

1.2. **Problem reformulation.** For simplicity, throughout the paper, we exploit the linear operators $\mathcal{K} : \mathbb{R}^{p\times p} \times \mathbb{R}^{r\times r} \to \mathbb{R}^{m\times n}$, $\mathcal{L} : \mathbb{R}^{p\times p} \times \mathbb{R}^{r\times r} \to \mathbb{R}^{\nu\times n}$, $\bar{\mathcal{K}} : \mathbb{R}^{m\times n} \times \mathbb{R}^{\nu\times n} \to \mathbb{R}^{p\times p}$ and $\bar{\mathcal{L}} : \mathbb{R}^{m\times n} \times \mathbb{R}^{\nu\times n} \to \mathbb{R}^{r\times r}$ defined as follows:

$$\begin{aligned}
\mathcal{K}(X,Y) &= A_1XB_1 + A_2X^TB_2 + C_1YD_1 + C_2Y^TD_2,\\
\mathcal{L}(X,Y) &= E_1XF_1 + E_2X^TF_2 + G_1YH_1 + G_2Y^TH_2,\\
\bar{\mathcal{K}}(Z,W) &= A_1^TZB_1^T + E_1^TWF_1^T + B_2Z^TA_2 + F_2W^TE_2,\\
\bar{\mathcal{L}}(Z,W) &= C_1^TZD_1^T + G_1^TWH_1^T + D_2Z^TC_2 + H_2W^TG_2.
\end{aligned}$$

In this paper, our goal is to propose an iterative algorithm to solve the following problems.

Problem 1. *Let the linear operators $\mathcal{K}$ and $\mathcal{L}$ be defined as before. Assume that $P_1 \in \mathbb{SOR}^{p\times p}$, $Q_1 \in \mathbb{SOR}^{p\times p}$, $P_2 \in \mathbb{SOR}^{r\times r}$ and $Q_2 \in \mathbb{SOR}^{r\times r}$ are given. Find $\tilde{X} \in \mathbb{SR}_{p\times p}^{P_1Q_1}$ and $\tilde{Y} \in \mathbb{SR}_{r\times r}^{P_2Q_2}$ such that*

$$\left\| \begin{pmatrix} M \\ N \end{pmatrix} - \begin{pmatrix} \mathcal{K}(\tilde{X},\tilde{Y}) \\ \mathcal{L}(\tilde{X},\tilde{Y}) \end{pmatrix} \right\|_F = \min_{\{(X,Y)\ |\ (X,Y)\in \mathbb{SR}_{p\times p}^{P_1Q_1}\times \mathbb{SR}_{r\times r}^{P_2Q_2}\}} \left\| \begin{pmatrix} M \\ N \end{pmatrix} - \begin{pmatrix} \mathcal{K}(X,Y) \\ \mathcal{L}(X,Y) \end{pmatrix} \right\|_F.$$

Problem 2. *Let $\mathcal{S}_s$ be the solution set of Problem 1. For two given matrices $X_0 \in \mathbb{R}^{p\times p}$ and $Y_0 \in \mathbb{R}^{r\times r}$, find $X^*, Y^* \in \mathcal{S}_s$ such that*

$$\|X^* - X_0\|_F^2 + \|Y^* - Y_0\|_F^2 = \min_{(X,Y)\in\mathcal{S}_s} \left\{ \|X - X_0\|_F^2 + \|Y - Y_0\|_F^2 \right\}.$$

Problem 3. *Let the linear operators $\mathcal{K}$ and $\mathcal{L}$ be defined as before. Assume that $P_1 \in \mathbb{SOR}^{p\times p}$, $Q_1 \in \mathbb{SOR}^{p\times p}$, $P_2 \in \mathbb{SOR}^{r\times r}$ and $Q_2 \in \mathbb{SOR}^{r\times r}$ are given. Find*

$\tilde{X} \in \mathbb{SSR}_{p\times p}^{P_1Q_1}$ *and* $\tilde{Y} \in \mathbb{SSR}_{r\times r}^{P_2Q_2}$ *such that*

$$\left\| \begin{pmatrix} M \\ N \end{pmatrix} - \begin{pmatrix} \mathcal{K}(\tilde{X},\tilde{Y}) \\ \mathcal{L}(\tilde{X},\tilde{Y}) \end{pmatrix} \right\|_F = \min_{\{(X,Y)\ |\ (X,Y)\in \mathbb{SSR}_{p\times p}^{P_1Q_1}\times \mathbb{SSR}_{r\times r}^{P_2Q_2}\}} \left\| \begin{pmatrix} M \\ N \end{pmatrix} - \begin{pmatrix} \mathcal{K}(X,Y) \\ \mathcal{L}(X,Y) \end{pmatrix} \right\|_F .$$

Problem 4. *Let $\mathcal{S}_{ss}$ be the solution set of Problem 3. For two given matrices $X_0 \in \mathbb{R}^{p\times p}$ and $Y_0 \in \mathbb{R}^{r\times r}$, find $X^*, Y^* \in \mathcal{S}_{ss}$ such that*

$$\|X^* - X_0\|_F^2 + \|Y^* - Y_0\|_F^2 = \min_{(X,Y)\in\mathcal{S}_{ss}} \left\{ \|X - X_0\|_F^2 + \|Y - Y_0\|_F^2 \right\}.$$

The remainder of this paper is organized as follows. In Section 2, we prove some properties of the (P,Q)-orthogonal symmetric and (P,Q)-orthogonal skew-symmetric matrices. Using the established characteristics, Section 3 is devoted to constructing and analyzing an iterative algorithm to solve Problems 1, 2, 3 and 4. Numerical results are reported in Section 4 which reveal the effectiveness of the algorithm to solve our mentioned problems. Finally the paper is ended with a brief conclusion in Section 5.

2. Useful theoretical results

This section is concerned with establishing some properties and theorems which provide fundamental tools for presenting our algorithm to solve Problems 1-4 and analyzing its convergence properties.

From the following theorem, it turns out that $\mathbb{R}^{m\times m}$ can be written as a direct sum of the subspaces $\mathbb{SR}_{m\times m}^{PQ}$ and $\mathbb{SSR}_{m\times m}^{PQ}$ where $P \in \mathbb{SOR}^{m\times m}$ and $Q \in \mathbb{SOR}^{m\times m}$.

Theorem 2.1. *Suppose that the matrices $P \in \mathbb{SOR}^{m\times m}$ and $Q \in \mathbb{SOR}^{m\times m}$ are given and the subspaces $\mathbb{SR}_{m\times m}^{PQ}$ and $\mathbb{SSR}_{m\times m}^{PQ}$ are signified as before. Then*

$$\mathbb{R}^{m\times m} = \mathbb{SR}_{m\times m}^{PQ} \oplus \mathbb{SSR}_{m\times m}^{PQ},$$

where $\oplus$ stands for the orthogonal direct sum with respect to the Frobenius inner product $\langle .,. \rangle_F$.

Proof. Before proving the validity of the assertion, we point out that straightforward computations reveals that $X \in \mathbb{SR}_{m\times m}^{PQ}$ if and only if $X = PQX^TPQ$ ($X \in \mathbb{SSR}_{m\times m}^{PQ}$ if and only if $X = -PQX^TPQ$). Here it should be noted that PQ is in general a nonsymmetric matrix satisfies $(PQ)^TPQ = PQ(PQ)^T = I_m$.

For simplicity, the proof is separated into the following two steps:

Step 1. It should be noticed that for an arbitrary given $X \in \mathbb{R}^{m\times m}$, we may write $X = W_1 + W_2$ where

$$W_1 = \frac{1}{2}\left(X + PQX^TPQ\right) \quad \text{and} \quad W_2 = \frac{1}{2}\left(X - PQX^TPQ\right),$$

and it is not difficult to see that $W_1 \in \mathbb{SR}_{m\times m}^{PQ}$ and $W_2 \in \mathbb{SSR}_{m\times m}^{PQ}$. Therefore, an arbitrary $m \times m$ matrix X can be written as the summation of two matrices W_1 and W_2 which are respectively (P,Q)-orthogonal symmetric and (P,Q)-orthogonal skew-symmetric matrices.

Step 2. In this step, it is shown that for two arbitrary given matrices $Z_1 \in \mathbb{SR}_{m\times m}^{PQ}$ and $Z_2 \in \mathbb{SSR}_{m\times m}^{PQ}$, we have $\langle Z_1, Z_2\rangle_F = 0$. Note that PZ_1Q and PZ_2Q are respectively symmetric and skew-symmetric matrices. It is well-known that symmetric and skew-symmetric matrices are orthogonal to each other, therefore the result follows from the following equality immediately,

$$\mathrm{tr}(Z_1^T Z_2) = \mathrm{tr}(Q^T Z_1^T P^T P Z_2 Q).$$

Now the result follows immediately from the given explanations in Steps 1 and 2. $\square$

It is not difficult to establish the following proposition by using the well-known properties of the trace of matrices.

Proposition 2.2. *Assume that the linear operators $\mathcal{K}$, $\mathcal{L}$, $\bar{\mathcal{K}}$ and $\bar{\mathcal{L}}$ are defined as before, then*

$$\langle \mathcal{K}(X,Y), Z\rangle_F + \langle \mathcal{L}(X,Y), W\rangle_F = \left\langle X, \bar{\mathcal{K}}(Z,W)\right\rangle_F + \left\langle Y, \bar{\mathcal{L}}(Z,W)\right\rangle_F, \tag{2.1}$$

for all $X \in \mathbb{R}^{p\times p}$, $Y \in \mathbb{R}^{r\times r}$, $Z \in \mathbb{R}^{m\times n}$ and $W \in \mathbb{R}^{\nu\times n}$.

Proof. It is well-known that

$$\begin{aligned}\langle \mathcal{K}(X,Y), Z\rangle_F + \langle \mathcal{L}(X,Y), W\rangle_F &= \mathrm{tr}\left(Z^T\left(A_1XB_1 + A_2X^TB_2 + C_1YD_1 + C_2Y^TD_2\right)\right)\\ &\quad +\mathrm{tr}\left(W^T\left(E_1XF_1 + E_2X^TF_2 + G_1YH_1 + G_2Y^TH_2\right)\right).\end{aligned}$$

Using the commutative property, we have

$$\begin{aligned}\langle \mathcal{K}(X,Y), Z\rangle_F + \langle \mathcal{L}(X,Y), W\rangle_F &= \mathrm{tr}\left(B_1Z^TA_1X\right) + \mathrm{tr}\left(X^TB_2Z^TA_2\right)\\ &\quad +\mathrm{tr}\left(D_1Z^TC_1Y\right) + \mathrm{tr}\left(Y^TD_2Z^TC_2\right)\\ &\quad +\mathrm{tr}\left(F_1W^TE_1X\right) + \mathrm{tr}\left(X^TF_2W^TE_2\right)\\ &\quad +\mathrm{tr}\left(H_1W^TG_1Y\right) + \mathrm{tr}\left(Y^TH_2W^TG_2\right).\end{aligned} \tag{2.2}$$

Invoking the fact that the trace of a matrix is equal to the trace of its transpose, it can be derived that

$$\mathrm{tr}\left(B_1Z^TA_1X\right) = \mathrm{tr}\left(X^TA_1^TZB_1^T\right), \qquad \mathrm{tr}\left(D_1Z^TC_1Y\right) = \mathrm{tr}\left(Y^TC_1^TZD_1^T\right),$$

$$\mathrm{tr}\left(F_1W^TE_1X\right) = \mathrm{tr}\left(X^TE_1^TWF_1^T\right) \quad \text{and} \quad \mathrm{tr}\left(H_1W^TG_1Y\right) = \mathrm{tr}\left(Y^TG_1^TWH_1^T\right).$$

The results follows immediately by substituting the above four equalities into (2.2). $\square$

In view of Theorem 2.1, the following two propositions can be concluded from Proposition 2.2 immediately. We only present the proof of the first proposition, the second one can be established in a similar manner.

Proposition 2.3. *Given the matrices $P_1 \in \mathbb{SOR}^{p\times p}$, $Q_1 \in \mathbb{SOR}^{p\times p}$, $P_2 \in \mathbb{SOR}^{r\times r}$ and $Q_2 \in \mathbb{SOR}^{r\times r}$, assume that the linear operators $\mathcal{K}$, $\mathcal{L}$, $\bar{\mathcal{K}}$ and $\bar{\mathcal{L}}$ are defined as before, then*

$$\begin{aligned}\langle \mathcal{K}(X,Y), Z\rangle_F + \langle \mathcal{L}(X,Y), W\rangle_F &= \frac{1}{2}\left\langle X, \bar{\mathcal{K}}(Z,W) + P_1Q_1(\bar{\mathcal{K}}(Z,W))^TP_1Q_1\right\rangle_F\\ &\quad +\frac{1}{2}\left\langle Y, \bar{\mathcal{L}}(Z,W) + P_2Q_2(\bar{\mathcal{L}}(Z,W))^TP_2Q_2\right\rangle_F,\end{aligned}$$

for all $X \in \mathbb{SR}_{p\times p}^{P_1Q_1}$, $Y \in \mathbb{SR}_{r\times r}^{P_2Q_2}$, $Z \in \mathbb{R}^{m\times n}$ and $W \in \mathbb{R}^{\nu\times n}$.

Proof. It is obvious that

$$\begin{aligned}\bar{\mathcal{K}}(Z,W) &= \frac{1}{2}\left(\bar{\mathcal{K}}(Z,W) + P_1Q_1(\bar{\mathcal{K}}(Z,W))^T P_1Q_1\right)\\ &\quad+\frac{1}{2}\left(\bar{\mathcal{K}}(Z,W) - P_1Q_1(\bar{\mathcal{K}}(Z,W))^T P_1Q_1\right)\end{aligned}$$

and

$$\begin{aligned}\bar{\mathcal{L}}(Z,W) &= \frac{1}{2}\left(\bar{\mathcal{L}}(Z,W) + P_2Q_2(\bar{\mathcal{L}}(Z,W))^T P_2Q_2\right)\\ &\quad+\frac{1}{2}\left(\bar{\mathcal{L}}(Z,W) - P_2Q_2(\bar{\mathcal{L}}(Z,W))^T P_2Q_2\right).\end{aligned}$$

Substitute two preceding equations in (2.1). By the assumption, $X \in \mathbb{SR}_{p\times p}^{P_1Q_1}$ and $Y \in \mathbb{SR}_{r\times r}^{P_2Q_2}$. Consequently, the assertion holds from the second step in the proof of Theorem 2.1. $\square$

Proposition 2.4. *Given the matrices $P_1 \in \mathbb{SOR}^{p\times p}$, $Q_1 \in \mathbb{SOR}^{p\times p}$, $P_2 \in \mathbb{SOR}^{r\times r}$ and $Q_2 \in \mathbb{SOR}^{r\times r}$, assume that the linear operators $\mathcal{K}$, $\mathcal{L}$, $\bar{\mathcal{K}}$ and $\bar{\mathcal{L}}$ are defined as before, then*

$$\begin{aligned}\langle \mathcal{K}(X,Y), Z\rangle_F + \langle \mathcal{L}(X,Y), W\rangle_F &= \frac{1}{2}\left\langle X, \bar{\mathcal{K}}(Z,W) - P_1Q_1(\bar{\mathcal{K}}(Z,W))^T P_1Q_1\right\rangle_F\\ &\quad+\frac{1}{2}\left\langle Y, \bar{\mathcal{L}}(Z,W) - P_2Q_2(\bar{\mathcal{L}}(Z,W))^T P_2Q_2\right\rangle_F,\end{aligned}$$

for all $X \in \mathbb{SSR}_{p\times p}^{P_1Q_1}$, $Y \in \mathbb{SSR}_{r\times r}^{P_2Q_2}$, $Z \in \mathbb{R}^{m\times n}$ and $W \in \mathbb{R}^{\nu\times n}$.

The following lemma is called `Projection Theorem` which is utilized to present some of our succeeding results, we refer the reader to (Wang, 2003) for its proof.

Lemma 2.5. *Let W be a finite dimensional space, U be a subspace of W and $U^\perp$ be the orthogonal complement subspace of U. For a given $w \in W$, always, there exists $u_0 \in U$ such that $\|w - u_0\| \le \|w - u\|$ for any $u \in U$. More precisely, $u_0 \in U$ is the unique minimization vector in U if and only if $(w - u_0) \perp U$. Here $\|.\|$ is the norm corresponding to the inner product defined on W.*

The following two useful propositions can be established by Lemma 2.5.

Proposition 2.6. *Given the matrices $P_1, Q_1 \in \mathbb{SOR}^{p\times p}$ and $P_2, Q_2 \in \mathbb{SOR}^{r\times r}$, let $X \in \mathbb{R}^{p\times p}$ and $Y \in \mathbb{R}^{r\times r}$ be arbitrary (P_1,Q_1)-orthogonal symmetric and (P_2,Q_2)-orthogonal symmetric matrices, respectively. Then the matrix pair $(X^*,Y^*) \in \mathbb{SR}_{p\times p}^{P_1Q_1} \times \mathbb{SR}_{r\times r}^{P_2Q_2}$ is the solution of Problem 1 if and only if for all $(X,Y) \in \mathbb{SR}_{p\times p}^{P_1Q_1} \times \mathbb{SR}_{r\times r}^{P_2Q_2}$,*

$$\begin{aligned}\big\langle X,\ \bar{\mathcal{K}}(R_1^*,R_2^*) &+ P_1Q_1(\bar{\mathcal{K}}(R_1^*,R_2^*))^T P_1Q_1\big\rangle_F\\ &+\left\langle Y, \bar{\mathcal{L}}(R_1^*,R_2^*) + P_2Q_2(\bar{\mathcal{L}}(R_1^*,R_2^*))^T P_2Q_2\right\rangle_F = 0,\end{aligned}$$

in which $R_1^* = M - \mathcal{K}(X^*,Y^*)$ and $R_2^* = N - \mathcal{L}(X^*,Y^*)$ where the linear operators $\mathcal{K}$, $\mathcal{L}$, $\bar{\mathcal{K}}$ and $\bar{\mathcal{L}}$ are defined as before.

Proof. Presume that

$$W = \left\{ \begin{pmatrix} V_1 \\ V_2 \end{pmatrix} \mid V_1 \in \mathbb{R}^{m\times n} \quad \text{and} \quad V_2 \in \mathbb{R}^{\nu\times n} \right\}.$$

Consider the following subspaces of W,

$$U = \left\{ \begin{pmatrix} \mathcal{K}(X,Y) \\ \mathcal{L}(X,Y) \end{pmatrix} \mid X \in \mathbb{SR}_{p\times p}^{P_1Q_1} \quad \text{and} \quad Y \in \mathbb{SR}_{r\times r}^{P_2Q_2} \right\}.$$

By Lemma 2.5, for

$$\begin{pmatrix} M \\ N \end{pmatrix} \in W,$$

we have

$$\left\| \begin{pmatrix} M \\ N \end{pmatrix} - \begin{pmatrix} \mathcal{K}(X^*,Y^*) \\ \mathcal{L}(X^*,Y^*) \end{pmatrix} \right\|_F = \min_L \left\| \begin{pmatrix} M \\ N \end{pmatrix} - \begin{pmatrix} \mathcal{K}(X,Y) \\ \mathcal{L}(X,Y) \end{pmatrix} \right\|_F,$$

if and only if

$$\left\langle \begin{pmatrix} \mathcal{K}(X,Y) \\ \mathcal{L}(X,Y) \end{pmatrix}, \begin{pmatrix} M - \mathcal{K}(X^*,Y^*) \\ N - \mathcal{L}(X^*,Y^*) \end{pmatrix} \right\rangle_F = 0,$$

for $(X,Y) \in L := \mathbb{SR}_{p\times p}^{P_1Q_1} \times \mathbb{SR}_{r\times r}^{P_2Q_2}$. Or equivalently,

$$\langle \mathcal{K}(X,Y), R_1^* \rangle_F + \langle \mathcal{L}(X,Y), R_2^* \rangle_F = 0, \qquad \forall\ (X,Y) \in \mathbb{SR}_{p\times p}^{P_1Q_1} \times \mathbb{SR}_{r\times r}^{P_2Q_2},$$

which completes the proof in view of Proposition 2.3. □

In a similar manner used in the proof of Proposition 2.6, the following proposition can be established.

Proposition 2.7. *Given the matrices $P_1, Q_1 \in \mathbb{SOR}^{p\times p}$ and $P_2, Q_2 \in \mathbb{SOR}^{r\times r}$, let $X \in \mathbb{R}^{p\times p}$ and $Y \in \mathbb{R}^{r\times r}$ be arbitrary (P_1,Q_1)-orthogonal skew-symmetric and (P_2,Q_2)-orthogonal skew-symmetric matrices, respectively. Then the matrix pair $(X^*,Y^*) \in \mathbb{SSR}_{p\times p}^{P_1Q_1} \times \in \mathbb{SSR}_{r\times r}^{P_2Q_2}$ is the solution of Problem 3 if and only if for all $(X,Y) \in \mathbb{SSR}_{p\times p}^{P_1Q_1} \times \mathbb{SSR}_{r\times r}^{P_2Q_2}$,*

$$\begin{aligned} \big\langle X,\ \bar{\mathcal{K}}(R_1^*,R_2^*) \quad - \quad & P_1Q_1(\bar{\mathcal{K}}(R_1^*,R_2^*))^T P_1Q_1 \big\rangle_F \\ & + \big\langle Y, \bar{\mathcal{L}}(R_1^*,R_2^*) - P_2Q_2(\bar{\mathcal{L}}(R_1^*,R_2^*))^T P_2Q_2 \big\rangle_F = 0, \end{aligned}$$

in which $R_1^* = M - \mathcal{K}(X^*,Y^*)$ and $R_2^* = N - \mathcal{L}(X^*,Y^*)$ where the linear operators $\mathcal{K}$, $\mathcal{L}$, $\bar{\mathcal{K}}$ and $\bar{\mathcal{L}}$ are defined as before.

3. Proposed algorithm and its convergence analysis

It is known that the conjugate gradient least-squares (CGLS) method is an effective algorithm to solve the following least-squares problem

$$\min_{x\in\mathbb{R}^n} \|b - Ax\|_2, \tag{3.1}$$

where $\|.\|_2$ is the Euclidean vector norm, A is a given large and sparse $m \times n$ matrix and $b \in \mathbb{R}^m$; for further details one may refer to Björck (1996) and Saad (1995).

In this section, the main goal is to examine the idea of the CGLS method to construct an iterative algorithm for solving Problems 1-4 by exploiting the elaborated results in the previous section. More precisely the current section is divided into three subsections. In the first part we propose an iterative scheme to solve Problems 1 and 3. The second subsection is devoted to proving the fact that the approximate solution pairs computed by the handled algorithm satisfy an optimality property at each iterate. In the last subsection it is momentarily described that how the presented algorithm can be employed to solve Problems 2 and 4.

3.1. **Solving Problems 1 and 3.** Utilizing the established facts in Section 2, we may present an extended form of the CGLS method for solving Problems 1 and 3 in Algorithm 2. In the sequel, the convergence properties of the algorithm are discussed.

Algorithm 2: The CGLS method for solving Problems 1 and 3.

Data: Input $A_i, B_i, C_i, D_i, E_i, F_i, G_i, H_i, M, N$ and the symmetric orthogonal matrices P_i and Q_i for $i = 1, 2$. Set $\eta = 0$ ($\eta = 1$) for solving Problem 1 (Problem 3).

Initialization: For solving Problem 1 (Problem 3), choose the initial guess $X(0) \in \mathbb{SR}_{p\times p}^{P_1Q_1}$ and $Y(0) \in \mathbb{SR}_{r\times r}^{P_2Q_2}$ ($X(0) \in \mathbb{SSR}_{p\times p}^{P_1Q_1}$ and $Y(0) \in \mathbb{SSR}_{r\times r}^{P_2Q_2}$).

begin

  Compute

  $R_1(0) = M - \mathcal{K}(X(0), Y(0));$

  $R_2(0) = N - \mathcal{L}(X(0), Y(0));$

  $P_x(0) = \frac{1}{2}\left(\bar{\mathcal{K}}(R_1(0), R_2(0)) + (-1)^{\eta} P_1Q_1(\bar{\mathcal{K}}(R_1(0), R_2(0)))^T P_1Q_1\right);$

  $P_y(0) = \frac{1}{2}\left(\bar{\mathcal{L}}(R_1(0), R_2(0)) + (-1)^{\eta} P_2Q_2(\bar{\mathcal{L}}(R_1(0), R_2(0)))^T P_2Q_2\right);$

  Set

  $Q_x(0) = P_x(0);$

  $Q_y(0) = P_y(0);$

  for $k = 0, 1, \ldots,$ *until the convergence* **do**

    $K = \mathcal{K}(Q_x(k), Q_y(k));$

    $L = \mathcal{L}(Q_x(k), Q_y(k));$

    $X(k+1) = X(k) + \frac{\|P_x(k)\|_F^2 + \|P_y(k)\|_F^2}{\|K\|_F^2 + \|L\|_F^2} Q_x(k);$

    $Y(k+1) = Y(k) + \frac{\|P_x(k)\|_F^2 + \|P_y(k)\|_F^2}{\|K\|_F^2 + \|L\|_F^2} Q_y(k);$

    $R_1(k+1) = R_1(k) - \frac{\|P_x(k)\|_F^2 + \|P_y(k)\|_F^2}{\|K\|_F^2 + \|L\|_F^2} K;$

    $R_2(k+1) = R_2(k) - \frac{\|P_x(k)\|_F^2 + \|P_y(k)\|_F^2}{\|K\|_F^2 + \|L\|_F^2} L;$

$$\begin{aligned} P_x(k+1) &= \frac{1}{2}\left(\bar{\mathcal{K}}(R_1(k+1), R_2(k+1))\right. \\ &\quad \left. + (-1)^{\eta} P_1Q_1 \left(\bar{\mathcal{K}}(R_1(k+1), R_2(k+1))\right)^T P_1Q_1\right); \end{aligned}$$

$$\begin{aligned} P_y(k+1) &= \frac{1}{2}\left(\bar{\mathcal{L}}(R_1(k+1), R_2(k+1))\right. \\ &\quad \left. + (-1)^{\eta} P_2Q_2 \left(\bar{\mathcal{L}}(R_1(k+1), R_2(k+1))\right)^T P_2Q_2\right); \end{aligned}$$

    $Q_x(k+1) = P_x(k+1) + \frac{\|P_x(k+1)\|_F^2 + \|P_y(k+1)\|_F^2}{\|P_x(k)\|_F^2 + \|P_y(k)\|_F^2} Q_x(k);$

    $Q_y(k+1) = P_y(k+1) + \frac{\|P_x(k+1)\|_F^2 + \|P_y(k+1)\|_F^2}{\|P_x(k)\|_F^2 + \|P_y(k)\|_F^2} Q_y(k);$

  end

end

The following theorem presents some properties of the sequences produced by Algorithm 2. These properties have a basic role for studying the convergence behavior of the algorithm to resolve Problems 1-4.

Theorem 3.1. *Suppose that k steps of Algorithm 2 have been performed. The sequences $P_x(l)$, $P_y(l)$, $Q_x(l)$ and $Q_y(l)$ ($l = 0, 1, \ldots, k$) produced by Algorithm 2 satisfy*

$$\langle P_x(i), P_x(j)\rangle_F + \langle P_y(i), P_y(j)\rangle_F = 0 \tag{3.2}$$

$$\begin{aligned}&\langle \mathcal{K}(Q_x(i), Q_y(i)), \mathcal{K}(Q_x(j), Q_y(j))\rangle_F \\ &\qquad + \langle \mathcal{L}(Q_x(i), Q_y(i)), \mathcal{L}(Q_x(j), Q_y(j))\rangle_F = 0,\end{aligned} \tag{3.3}$$

$$\langle Q_x(i), P_x(j)\rangle_F + \langle Q_y(i), P_y(j)\rangle_F = 0, \tag{3.4}$$

for $i, j = 0, 1, 2, \ldots, k$ ($i \neq j$).

Note that the previous theorem can be established by mathematical induction and its proof is given in Appendix.

Remark 3.2. *Without loss of generality, assume that we apply the proposed algorithm to solve Problem 1. Suppose that k ($k > 1$) steps of Algorithm 2 have been performed. To continue the algorithm at $(k+1)$-th step, we must have $P_x(k) \neq 0$ or $P_y(k) \neq 0$ in Lines 21 and 22. Note that if $P_x(k) = 0$ and $P_y(k) = 0$ then Proposition 2.6 implies that $(X(k), Y(k))$ is a solution pair of Problem 1. Considering Lines 15-18, it is required to presume that $\mathcal{K}(Q_x(k), Q_y(k)) \neq 0$ or $\mathcal{L}(Q_x(k), Q_y(k)) \neq 0$. Otherwise, using Proposition 2.3 and Eq.* (3.4)*, we have*

$$\begin{aligned}0 &= \langle \mathcal{K}(Q_x(k), Q_y(k)), R_1(k)\rangle_F + \langle \mathcal{L}(Q_x(k), Q_y(k)), R_2(k)\rangle_F \\ &= \langle Q_x(k), P_x(k)\rangle_F + \langle Q_y(k), P_y(k)\rangle_F \\ &= \langle P_x(k), P_x(k)\rangle_F + \langle P_y(k), P_y(k)\rangle_F,\end{aligned}$$

which indicates that $P_x(k) = 0$ and $P_y(k) = 0$, as discussed earlier, this ensures that $(X(k), Y(k))$ is a solution pair of Problem 1.

Remark 3.3. *In view of the orthogonality relation* (3.2) *presented in Theorem 3.1, it reveals that Algorithm 2 can solve Problems 1 and 3 within finite number of steps.*

Theorem 3.4. *Assume that $P_1 \in \mathbb{SOR}^{p\times p}$, $Q_1 \in \mathbb{SOR}^{p\times p}$, $P_2 \in \mathbb{SOR}^{r\times r}$ and $Q_2 \in \mathbb{SOR}^{r\times r}$ are given. Suppose that $(\tilde{X}, \tilde{Y})$ is a solution pair of Problem 1 (Problem 3). If $(\hat{X}, \hat{Y})$ is an alternative solution pair of Problem 1 (Problem 3), then there exist $X_1 \in \mathbb{SR}^{P_1Q_1}_{p\times p}$ and $Y_1 \in \mathbb{SR}^{P_2Q_2}_{r\times r}$ ($X_1 \in \mathbb{SSR}^{P_1Q_1}_{p\times p}$ and $Y_1 \in \mathbb{SSR}^{P_2Q_2}_{r\times r}$) such that*

$$\tilde{X} = \hat{X} + X_1 \quad \text{and} \quad \tilde{Y} = \hat{Y} + Y_1,$$

and

$$\mathcal{K}(X_1, Y_1) = 0 \quad \text{and} \quad \mathcal{L}(X_1, Y_1) = 0. \tag{3.5}$$

Proof. We prove the assertion only for the solutions of Problem 1, the same strategy can be handled to show that the conclusion is also valid for the solutions of Problem 3. Presume that $X_1 = \tilde{X} - \hat{X}$ and $Y_1 = \tilde{Y} - \hat{X}$, it is not difficult to see that $X_1 \in \mathbb{SR}^{P_1Q_1}_{p\times p}$ and $Y_1 \in \mathbb{SR}^{P_2Q_2}_{r\times r}$. By Propositions 2.3 and 2.6 and using some straightforward computations, it turns out that

$$\left\langle M - \mathcal{K}(\hat{X}, \hat{Y}), \mathcal{K}(X_1, Y_1)\right\rangle_F + \left\langle N - \mathcal{L}(\hat{X}, \hat{Y}), \mathcal{L}(X_1, Y_1)\right\rangle_F = 0. \tag{3.6}$$

By the assumption, $(\tilde{X}, \tilde{Y})$ and $(\hat{X},\hat{Y})$ are solutions of Problem 1. This hypophysis guarantees that

$$\left\| \begin{pmatrix} M - \mathcal{K}(\tilde{X}, \tilde{Y}) \\ N - \mathcal{L}(\tilde{X}, \tilde{Y}) \end{pmatrix} \right\|_F = \left\| \begin{pmatrix} M - \mathcal{K}(\hat{X}, \hat{Y}) \\ N - \mathcal{L}(\hat{X}, \hat{Y}) \end{pmatrix} \right\|_F,$$

which is equivalent to say that

$$\left\| M - \mathcal{K}(\tilde{X}, \tilde{Y}) \right\|_F^2 + \left\| N - \mathcal{L}(\tilde{X}, \tilde{Y}) \right\|_F^2 = \left\| M - \mathcal{K}(\hat{X}, \hat{Y}) \right\|_F^2 + \left\| N - \mathcal{L}(\hat{X}, \hat{Y}) \right\|_F^2.$$

Considering (3.6), we derive the following relation

$$\begin{aligned} \|\mathcal{K}(X_1, Y_1)\|_F^2 + \|\mathcal{L}(X_1, Y_1)\|_F^2 &= \left(\left\| M - \mathcal{K}(\tilde{X}, \tilde{Y}) \right\|_F^2 + \left\| N - \mathcal{L}(\tilde{X}, \tilde{Y}) \right\|_F^2 \right) \\ &\quad - \left(\left\| M - \mathcal{K}(\hat{X}, \hat{Y}) \right\|_F^2 + \left\| N - \mathcal{L}(\hat{X}, \hat{Y}) \right\|_F^2 \right), \end{aligned}$$

which completes the proof. $\square$

By Theorem 3.4, in the rest of this subsection, it is demonstrated that Algorithm 2 converges to the least-norm solution pair (X^*, Y^*) of Problem 1 (Problem 3). For two arbitrary given matrices $Z \in \mathbb{R}^{m\times n}$ and $W \in \mathbb{R}^{\nu\times n}$, we set

$$X(0) = \frac{1}{2}\left(\bar{\mathcal{K}}(Z, W) + (-1)^\eta P_1 Q_1 (\bar{\mathcal{K}}(Z, W))^T P_1 Q_1\right), \tag{3.7}$$

$$Y(0) = \frac{1}{2}\left(\bar{\mathcal{L}}(Z, W) + (-1)^\eta P_2 Q_2 (\bar{\mathcal{L}}(Z, W))^T P_2 Q_2\right), \tag{3.8}$$

where the value of η is respectively taken to be `zero` and `one` for solving Problem 1 and Problem 3. By applying Algorithm 2 with $X(0)$ and $Y(0)$ as shown by (3.7) and (3.8), Remark 3.3 implies that there exist $Z^* \in \mathbb{R}^{m\times n}$ and $W^* \in \mathbb{R}^{\nu\times n}$ such that the solution pair (X^*, Y^*) solves Problem 1 (for $\eta = 0$) and Problem 3 (for $\eta = 1$) where

$$X^* = \frac{1}{2}\left(\bar{\mathcal{K}}(Z^*, W^*) + (-1)^\eta P_1 Q_1 (\bar{\mathcal{K}}(Z^*, W^*))^T P_1 Q_1\right), \tag{3.9}$$

$$Y^* = \frac{1}{2}\left(\bar{\mathcal{L}}(Z^*, W^*) + (-1)^\eta P_2 Q_2 (\bar{\mathcal{L}}(Z^*, W^*))^T P_2 Q_2\right). \tag{3.10}$$

Now if $(\tilde{X}, \tilde{Y})$ is a solution pair of Problem 1 (Problem 3) then Theorem 3.4 indicates the existence of $(X_1, Y_1) \in \mathbb{SR}_{p\times p}^{P_1Q_1} \times \mathbb{SR}_{r\times r}^{P_2Q_2}$ ($(X_1, Y_1) \in \mathbb{SSR}_{p\times p}^{P_1Q_1} \times \mathbb{SSR}_{r\times r}^{P_2Q_2}$) where

$$\tilde{X} = X^* + X_1 \quad \text{and} \quad \tilde{Y} = Y^* + Y_1,$$

and the relations given in (3.5) are satisfied. Evidently, we have

$$\begin{aligned} \langle X^* + X_1, X^* + X_1\rangle_F + \langle Y^* + Y_1, Y^* + Y_1\rangle_F &= \langle X^*, X^*\rangle_F + \langle Y^*, Y^*\rangle_F \\ &\quad + 2\left(\langle X^*, X_1\rangle_F + \langle Y^*, Y_1\rangle_F\right) \\ &\quad + \langle X_1, X_1\rangle_F + \langle Y_1, Y_1\rangle_F. \end{aligned}$$

In view of (3.5), (3.9) and (3.10), Proposition 2.3 (Proposition 2.4) implies that

$$\langle X^*, X_1\rangle_F + \langle Y^*, Y_1\rangle_F = 0.$$

Consequently, it can be deduced that

$$\begin{aligned} \left\langle \tilde{X}, \tilde{X} \right\rangle_F + \left\langle \tilde{Y}, \tilde{Y} \right\rangle_F &= \langle X^* + X_1, X^* + X_1\rangle_F + \langle Y^* + Y_1, Y^* + Y_1\rangle_F \\ &= \left(\langle X^*, X^*\rangle_F + \langle Y^*, Y^*\rangle_F\right) + \left(\langle X_1, X_1\rangle_F + \langle Y_1, Y_1\rangle_F\right). \end{aligned}$$

Therefore, we obtain

$$\|X^*\|_F^2 + \|Y^*\|_F^2 \leq \left\|\tilde{X}\right\|_F^2 + \left\|\tilde{Y}\right\|_F^2.$$

It should be noticed here that the inequality in the above relation holds strictly as long as $X_1 \neq 0$ or $Y_1 \neq 0$. As a result (X^*, Y^*) is the unique least-norm solution pair of Problem 1 (Problem 3) which can be reached via Algorithm 2 by choosing $X(0)$ and $Y(0)$ as shown by of (3.7) and (3.8) with $\eta = 0$ ($\eta = 1$).

3.2. **Optimality property of the proposed algorithm.** The minimization property of Algorithm 2 can be deduced from the following proposition.

Proposition 3.5. *Suppose that k steps of Algorithm 2 have been performed. Then the sequence $\{R(l)\}_{l=1}^k$ satisfies*

$$\langle R_1(k), \mathcal{K}(Q_x(i), Q_y(i))\rangle_F + \langle R_2(k), \mathcal{L}(Q_x(i), Q_y(i))\rangle_F = 0, \tag{3.11}$$

for $i = 1, 2, \ldots, k-1$, where $R_1(k) = M - \mathcal{K}(X(k), Y(k))$, $R_2(k) = N - \mathcal{L}(X(k), Y(k))$ and $(X(k), Y(k))$ is the k-th approximate solution pair computed by Algorithm 2.

Proof. Assume that $\eta = 0$ ($\eta = 1$). In view of Proposition 2.3 (Proposition 2.4), it is deduced that

$$\begin{aligned}\langle R_1(k), \mathcal{K}(Q_x(i), Q_y(i))\rangle_F &+ \langle R_2(k), \mathcal{L}(Q_x(i), Q_y(i))\rangle_F \\ &= \langle P_x(k), Q_x(i)\rangle_F + \langle P_y(k), Q_y(i)\rangle_F.\end{aligned}$$

Now the result follows immediately by the property (3.4) given in Theorem 3.1. □

Considering that m-th step of Algorithm 2 has been performed, we elucidate the subspaces K_m and U_m as follows:

$$K_m = \operatorname{span}\left\{\begin{pmatrix} Q_x(0) & 0 \\ 0 & Q_y(0)\end{pmatrix}, \ldots, \begin{pmatrix} Q_x(m-1) & 0 \\ 0 & Q_y(m-1)\end{pmatrix}\right\}, \tag{3.12}$$

$$U_m = \operatorname{span}\left\{\begin{pmatrix} \mathcal{K}(Q_x(0), Q_y(0)) \\ \mathcal{L}(Q_x(0), Q_y(0))\end{pmatrix}, \ldots, \begin{pmatrix} \mathcal{K}(Q_x(m-1), Q_y(m-1)) \\ \mathcal{L}(Q_x(m-1), Q_y(m-1))\end{pmatrix}\right\}. \tag{3.13}$$

Evidently we can see that

$$\begin{pmatrix} X(m) & 0 \\ 0 & Y(m)\end{pmatrix} \in \begin{pmatrix} X(0) & 0 \\ 0 & Y(0)\end{pmatrix} + K_m.$$

By Proposition 3.5, it indicates that

$$\left(\begin{pmatrix} M \\ N\end{pmatrix} - \begin{pmatrix} \mathcal{K}(X(m), Y(m)) \\ \mathcal{L}(X(m), Y(m))\end{pmatrix}\right) \perp U_m.$$

The following theorem reveals that the approximate solutions produced by Algorithm 2 satisfy an optimality property which demonstrates the minimization property of the algorithm. The result is a direct consequence of Lemma 2.5, hence we omit its proof.

Theorem 3.6. *Suppose that m steps of Algorithm 2 have been performed. Presume that the subspaces K_m and U_m are defined by (3.12) and (3.13), respectively. Then*

$$\left\|\begin{pmatrix} M \\ N\end{pmatrix} - \begin{pmatrix} \mathcal{K}(X(m), Y(m)) \\ \mathcal{L}(X(m), Y(m))\end{pmatrix}\right\|_F = \min_{L}\left\|\begin{pmatrix} M \\ N\end{pmatrix} - \begin{pmatrix} \mathcal{K}(X, Y) \\ \mathcal{L}(X, Y)\end{pmatrix}\right\|_F,$$

where

$$L := \left\{ \begin{pmatrix} X & 0 \\ 0 & Y \end{pmatrix} \mid \begin{pmatrix} X & 0 \\ 0 & Y \end{pmatrix} \in \begin{pmatrix} X(0) & 0 \\ 0 & Y(0) \end{pmatrix} + K_m \right\}.$$

3.3. **On the solution of Problems 2 and 4.** In the current subsection we discuss an approach to solve Problems 2 and 4. Assume that $P_1 \in \mathbb{SOR}^{p\times p}$, $Q_1 \in \mathbb{SOR}^{p\times p}$, $P_2 \in \mathbb{SOR}^{r\times r}$ and $Q_2 \in \mathbb{SOR}^{r\times r}$ are given. Let us denote the solution set of Problems 2 and 4 by $\mathcal{S}_s$ and $\mathcal{S}_{ss}$, respectively. Evidently the sets $\mathcal{S}_s$ and $\mathcal{S}_{ss}$ are both nonempty. For two given $X_0 \in \mathbb{R}^{p\times p}$ and $Y_0 \in \mathbb{R}^{r\times r}$, Theorem 2.1 ensures the existence of Z_i and W_i $(i = 1, 2)$ where $X_0 = Z_1 + Z_2$ and $Y_0 = W_1 + W_2$ such that $Z_1 \in \mathbb{SR}^{P_1Q_1}_{p\times p}$, $Z_2 \in \mathbb{SSR}^{P_1Q_1}_{p\times p}$, $W_1 \in \mathbb{SR}^{P_2Q_2}_{r\times r}$ and $W_2 \in \mathbb{SSR}^{P_2Q_2}_{r\times r}$.

Remark 3.7. *It is not difficult to see that the following statements hold*

- *If* $(X, Y) \in \mathcal{S}_s$ *then*

$$\begin{aligned} \langle X_0 - X, X_0 - X\rangle_F + \langle Y_0 - Y, Y_0 - Y\rangle_F &= \langle Z_1 - X, Z_1 - X\rangle_F + \langle W_1 - Y, W_1 - Y\rangle_F \\ &\quad + \langle Z_2, Z_2\rangle_F + \langle W_2, W_2\rangle_F \end{aligned} \tag{3.14}$$

- *If* $(X, Y) \in \mathcal{S}_{ss}$ *then*

$$\begin{aligned} \langle X_0 - X, X_0 - X\rangle_F + \langle Y_0 - Y, Y_0 - Y\rangle_F &= \langle Z_2 - X, Z_2 - X\rangle_F + \langle W_2 - Y, W_2 - Y\rangle_F \\ &\quad + \langle Z_1, Z_1\rangle_F + \langle W_1, W_1\rangle_F \end{aligned} \tag{3.15}$$

Consequently, from (3.14)*, it can be concluded that*

$$\min_{(X,Y)\in\mathcal{S}_s} \left\{ \|X - X_0\|_F^2 + \|Y - Y_0\|_F^2 \right\} = \min_{(X,Y)\in\mathcal{S}_s} \left\{ \|Z_1 - X\|_F^2 + \|W_1 - Y\|_F^2 \right\}. \tag{3.16}$$

In addition, (3.15) *implies that*

$$\min_{(X,Y)\in\mathcal{S}_{ss}} \left\{ \|X - X_0\|_F^2 + \|Y - Y_0\|_F^2 \right\} = \min_{(X,Y)\in\mathcal{S}_{ss}} \left\{ \|Z_2 - X\|_F^2 + \|W_2 - Y\|_F^2 \right\}. \tag{3.17}$$

From Remark 3.7, we infer that for solving Problem 2 and Problem 4, it is sufficient to evaluate the least-norm solutions of the following new problems, respectively.

Problem 5. *Let the linear operators $\mathcal{K}$ and $\mathcal{L}$ be defined as before. Assume that $P_1 \in \mathbb{SOR}^{p\times p}$, $Q_1 \in \mathbb{SOR}^{p\times p}$, $P_2 \in \mathbb{SOR}^{r\times r}$ and $Q_2 \in \mathbb{SOR}^{r\times r}$, $X_0 \in \mathbb{R}^{p\times p}$ and $Y_0 \in \mathbb{R}^{r\times r}$ are given. Find $\tilde{X} \in \mathbb{SR}^{P_1Q_1}_{p\times p}$ and $\tilde{Y} \in \mathbb{SR}^{P_2Q_2}_{r\times r}$ such that*

$$\left\| \begin{pmatrix} \tilde{M} \\ \tilde{N} \end{pmatrix} - \begin{pmatrix} \mathcal{K}(\tilde{X}, \tilde{Y}) \\ \mathcal{L}(\tilde{X}, \tilde{Y}) \end{pmatrix} \right\|_F = \min_{\{(X,Y) \mid (X,Y)\in\mathbb{SR}^{P_1Q_1}_{p\times p}\times\mathbb{SR}^{P_2Q_2}_{r\times r}\}} \left\| \begin{pmatrix} \tilde{M} \\ \tilde{N} \end{pmatrix} - \begin{pmatrix} \mathcal{K}(X, Y) \\ \mathcal{L}(X, Y) \end{pmatrix} \right\|_F,$$

in which $\tilde{M} = M - \mathcal{K}(Z_1, W_1)$ and $\tilde{N} = N - \mathcal{L}(Z_1, W_1)$ where

$$Z_1 = \frac{1}{2}\left(X_0 + P_1Q_1X_0^TP_1Q_1\right) \quad \textit{and} \quad W_1 = \frac{1}{2}\left(Y_0 + P_2Q_2Y_0^TP_2Q_2\right).$$

Problem 6. *Let the linear operators $\mathcal{K}$ and $\mathcal{L}$ be defined as before. Assume that $P_1 \in \mathbb{SOR}^{p\times p}$, $Q_1 \in \mathbb{SOR}^{p\times p}$, $P_2 \in \mathbb{SOR}^{r\times r}$ and $Q_2 \in \mathbb{SOR}^{r\times r}$, $X_0 \in \mathbb{R}^{p\times p}$ and $Y_0 \in \mathbb{R}^{r\times r}$ are given. Find $\tilde{X} \in \mathbb{SSR}^{P_1Q_1}_{p\times p}$ and $\tilde{Y} \in \mathbb{SSR}^{P_2Q_2}_{r\times r}$ such that*

$$\left\| \begin{pmatrix} \tilde{M} \\ \tilde{N} \end{pmatrix} - \begin{pmatrix} \mathcal{K}(\tilde{X}, \tilde{Y}) \\ \mathcal{L}(\tilde{X}, \tilde{Y}) \end{pmatrix} \right\|_F = \min_{\{(X,Y) \mid (X,Y)\in\mathbb{SSR}^{P_1Q_1}_{p\times p}\times\mathbb{SSR}^{P_2Q_2}_{r\times r}\}} \left\| \begin{pmatrix} \tilde{M} \\ \tilde{N} \end{pmatrix} - \begin{pmatrix} \mathcal{K}(X, Y) \\ \mathcal{L}(X, Y) \end{pmatrix} \right\|_F,$$

in which $\tilde{M} = M - \mathcal{K}(Z_2, W_2)$ *and* $\tilde{N} = N - \mathcal{L}(Z_2, W_2)$ *where*

$$Z_2 = \frac{1}{2}\left(X_0 - P_1Q_1X_0^TP_1Q_1\right) \quad \text{and} \quad W_2 = \frac{1}{2}\left(Y_0 - P_2Q_2Y_0^TP_2Q_2\right).$$

To determine the solution of Problem 2 (Problem 4), we first compute $(\tilde{X}, \tilde{Y})$ as the least-norm solution pair of Problem 5 (Problem 6) using the approach described at the end of Subsection 3.1. Then the solution pair (X^*, Y^*) of Problem 2 (Problem 4) is obtained by $X^* = \tilde{X} + Z_1$ and $Y^* = \tilde{Y} + W_1$ ($X^* = \tilde{X} + Z_2$ and $Y^* = \tilde{Y} + W_2$).

4. Numerical experiments

In this section, some numerical instances are examined to demonstrate the validity of the elaborated theoretical results and to illustrate the practicability of the proposed algorithm for solving Problems 1-4. More precisely, we consider matrix equations in the cases that they are consistent or inconsistent over (P_1, Q_1)-orthogonal symmetric and skew-symmetric and Hamiltonian matrices. All of reported experiments were performed on a 64-bit 2.45 GHz core i7 processor and 8.00GB RAM using some MATLAB codes on MATLAB version 8.3.0532.

Example 4.1. Consider the following least-squares problem

$$\|M - \mathcal{K}(\tilde{X})\|_F = \min_{X \in \mathbb{SR}^{P_1Q_1}_{5\times 5}} \|M - \mathcal{K}(X)\|_F, \tag{4.1}$$

where $\mathcal{K}(X) = A_1XB_1$ in which

$$A_1 = \begin{pmatrix} 1 & 2 & 3 & -1 & 2 \\ 3 & 1 & 0 & 2 & 1 \\ 2 & 3 & 4 & 1 & 0 \\ 1 & 2 & 3 & -1 & 1 \\ 3 & 2 & 1 & 1 & -1 \end{pmatrix}, \quad B_1 = \begin{pmatrix} 1 & 2 & 3 & -1 & 3 \\ -1 & 2 & 0 & 2 & 2 \\ 1 & 1 & 1 & 2 & 0 \\ 1 & 2 & 2 & -1 & 1 \\ 0 & 0 & -1 & 1 & 1 \end{pmatrix},$$

$$M = \begin{pmatrix} 28 & 83 & 74 & 29 & 55 \\ -18 & -5 & -14 & -4 & -44 \\ 5 & 46 & 27 & 56 & 6 \\ 23 & 69 & 62 & 33 & 47 \\ -19 & -1 & -10 & 19 & -27 \end{pmatrix},$$

and the matrices $P_1, Q_1 \in \mathbb{SOR}^{5\times 5}$ are given by

$$P_1 = \begin{pmatrix} 0 & 0 & 0 & 0 & -1 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ -1 & 0 & 0 & 0 & 0 \end{pmatrix}, \quad Q_1 = \begin{pmatrix} 0 & 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}.$$

In order to solve (4.1) by Algorithm 2, we exploit a zero matrix as an initial guess and the following stopping criterion,

$$\delta_k = \|X(k) - X(k-1)\|_F < 10^{-10}.$$

The algorithm converges in 22 iterations and the approximate solution is obtained by

$$X(22) = \begin{pmatrix} -1.0000 & 0.0000 & -0.0000 & -0.0000 & -6.0000 \\ -0.0000 & 6.0000 & -1.0000 & 1.0000 & -0.0000 \\ 1.0000 & -1.0000 & 6.0000 & 0.0000 & 0.0000 \\ -6.0000 & 0.0000 & -1.0000 & 0.0000 & -1.0000 \\ -0.0000 & 1.0000 & 0.0000 & 6.0000 & 0.0000 \end{pmatrix},$$

with the residual norm $\eta_{22} = 2.86e-11$, where $\eta_k = \|R(k)\|_F = \|M - A_1 X(k) B_1\|_F$. It is noted that the matrix

$$\tilde{X} = \begin{pmatrix} -1 & 0 & 0 & 0 & -6 \\ 0 & 6 & -1 & 1 & 0 \\ 1 & -1 & 6 & 0 & 0 \\ -6 & 0 & -1 & 0 & -1 \\ 0 & 1 & 0 & 6 & 0 \end{pmatrix},$$

is a solution of $\mathcal{K}(X) = M$ which is (P_1, Q_1)-orthogonal symmetric. This reveals that the algorithm has provided an approximate solution to $\mathcal{K}(X) = M$ which is also an approximate solution to (4.1). Convergence history of the method is plotted in Figure 1 for more clarification.

We now utilize Algorithm 2 to compute matrix X^* such that

$$\|X^* - X_0\|_F = \min_{X \in \mathcal{S}_s} \|X - X_0\|_F, \tag{4.2}$$

where $\mathcal{S}_s$ is the solution set of (4.1) and

$$X_0 = \begin{pmatrix} 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 \end{pmatrix}. \tag{4.3}$$

According to Problem 5, we set

$$Z_1 = \frac{1}{2}(X_0 + P_1 Q_1 X_0^T P_1 Q_1) = \frac{1}{2} \begin{pmatrix} 2 & 1 & 0 & -1 & 0 \\ 1 & 2 & 2 & 0 & -1 \\ -1 & 2 & 2 & 1 & 0 \\ 0 & -1 & 1 & 1 & 2 \\ -1 & 0 & 1 & 2 & 1 \end{pmatrix},$$

and

$$\tilde{M} = M - \mathcal{K}(Z_1) = \frac{1}{2} \begin{pmatrix} 53 & 123 & 128 & 18 & 86 \\ -45 & -35 & -49 & -21 & -117 \\ 4 & 35 & 23 & 59 & -29 \\ 45 & 98 & 105 & 28 & 70 \\ -41 & -34 & -45 & 16 & -90 \end{pmatrix}.$$

Applying Algorithm 2 to solve the ensuing least-squares problem,

$$\|\tilde{M} - \mathcal{K}(\tilde{X})\|_F = \min_{X \in \mathbb{SR}_{5\times 5}^{P_1 Q_1}} \|\tilde{M} - \mathcal{K}(X)\|_F, \tag{4.4}$$

we derive

$$X(23) = \begin{pmatrix} -2.0000 & -0.5000 & -0.0000 & 0.5000 & -6.0000 \\ -0.5000 & 5.0000 & -2.0000 & 1.0000 & 0.5000 \\ 1.5000 & -2.0000 & 5.0000 & -0.5000 & 0.0000 \\ -6.0000 & 0.5000 & -1.5000 & -0.5000 & -2.0000 \\ 0.5000 & 1.0000 & -0.5000 & 5.0000 & -0.5000 \end{pmatrix},$$

as an approximation of $\tilde{X}$. Hence, the approximate solution of (4.2) is determined by

$$X^* \approx Z_1 + X(23) = \begin{pmatrix} -1.0000 & 0.0000 & -0.0000 & 0.0000 & -6.0000 \\ -0.0000 & 6.0000 & -1.0000 & 1.0000 & -0.0000 \\ 1.0000 & -1.0000 & 6.0000 & -0.0000 & 0.0000 \\ -6.0000 & 0.0000 & -1.0000 & 0.0000 & -1.0000 \\ -0.0000 & 1.0000 & -0.0000 & 6.0000 & -0.0000 \end{pmatrix}.$$

For more details, $\log_{10}\delta_k$ has been displayed in Figure 1.

To demonstrate the effectiveness of the proposed algorithm to compute a (P_1, Q_1)-orthogonal skew-symmetric solution of the investigated problem, let us consider the following problem

$$\|M - \mathcal{K}(\tilde{X})\|_F = \min_{X \in \mathbb{SSR}^{P_1 Q_1}_{5\times 5}} \|M - \mathcal{K}(X)\|_F, \tag{4.5}$$

with

$$M = \begin{pmatrix} 18 & -33 & -12 & -9 & -63 \\ -6 & -21 & -30 & 0 & -24 \\ 9 & -36 & -33 & 0 & -66 \\ 15 & -27 & -12 & -3 & -57 \\ -3 & -9 & -24 & 15 & -27 \end{pmatrix},$$

and

$$\|X^* - X_0\|_F = \min_{X \in \mathcal{S}_{ss}} \|X - X_0\|_F, \tag{4.6}$$

where $\mathcal{S}_{ss}$ is the set of solutions of the problem (4.5). With a similar manner used for finding the solution in the orthogonal symmetric case, Algorithm 2 is implemented to compute approximate solutions to the problems (4.5) and (4.6). All the other assumptions are as before. For both of the problems, the algorithm calculates an approximate solution of the following form

$$X(13) = \begin{pmatrix} -3.0000 & -0.0000 & -0.0000 & 0.0000 & 0 \\ -0.0000 & 0 & 3.0000 & 3.0000 & -0.0000 \\ -3.0000 & -3.0000 & 0 & -0.0000 & -0.0000 \\ 0 & -0.0000 & -3.0000 & 0.0000 & 3.0000 \\ 0.0000 & -3.0000 & 0.0000 & 0 & 0.0000 \end{pmatrix},$$

for $\tilde{X}$ and X^* in 13 iterations. Here, we have $\eta_{13} = 7.11e-12$ for Problem 3. It is mentioned that

$$\tilde{X} = \begin{pmatrix} -3 & 0 & 0 & 0 & 0 \\ 0 & 0 & 3 & 3 & 0 \\ -3 & -3 & 0 & 0 & 0 \\ 0 & 0 & -3 & 0 & 3 \\ 0 & -3 & 0 & 0 & 0 \end{pmatrix},$$

is (P_1, Q_1)-orthogonal skew-symmetric matrix and $\tilde{X} \approx X(13)$. Convergence behavior of the method to solve problems (4.5) and (4.6) is depicted in Figure 1. Subsequently, we consider Algorithm 2 to solve (4.6) where X_0 is given by (4.3). Having in mind Problem 6, we set

$$Z_2 = \frac{1}{2}(X_0 - P_1Q_1X_0^TP_1Q_1) = \frac{1}{2}\begin{pmatrix} 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 1 & 1 & 0 \\ 1 & 0 & -1 & 0 & 1 \end{pmatrix},$$

and

$$\tilde{M} = M - \mathcal{K}(Z_2) = \frac{1}{2}\begin{pmatrix} 29 & -83 & -44 & -6 & -154 \\ -15 & -67 & -75 & -7 & -71 \\ 6 & -107 & -99 & 3 & -169 \\ 23 & -70 & -43 & 4 & -138 \\ -11 & -42 & -65 & 24 & -74 \end{pmatrix}.$$

If we apply Algorithm 2 to solve the least-squares problem,

$$\|\tilde{M} - \mathcal{K}(\tilde{X})\|_F = \min_{X \in \mathbb{SR}_{5\times 5}^{P_1Q_1}} \|\tilde{M} - \mathcal{K}(X)\|_F, \tag{4.7}$$

we get

$$X(13) = \begin{pmatrix} -3.0000 & -0.5000 & -0.0000 & -0.5000 & 0 \\ -0.5000 & 0 & 3.0000 & 3.0000 & -0.5000 \\ -3.5000 & -3.0000 & 0 & -0.5000 & -0.0000 \\ 0 & -0.5000 & -3.5000 & -0.5000 & 3.0000 \\ -0.5000 & -3.0000 & 0.5000 & 0 & -0.5000 \end{pmatrix},$$

as an approximation of $\tilde{X}$. Hence, the approximate solution of (4.6) is determined by

$$X^* \approx Z_2 + X(13) = \begin{pmatrix} -3.0000 & -0.0000 & -0.0000 & 0.0000 & 0 \\ -0.0000 & 0 & 3.0000 & 3.0000 & -0.0000 \\ -3.0000 & -3.0000 & 0 & -0.0000 & -0.0000 \\ 0 & -0.0000 & -3.0000 & 0.0000 & 3.0000 \\ 0.0000 & -3.0000 & 0.0000 & 0 & 0.0000 \end{pmatrix}.$$

For more details, $\log_{10} \delta_k$ has been displayed in Figure 1.

We now set $M = I$, where I is the identity matrix of order n. In this case it not difficult to see that the equation $A_1XB_1 = M$ is inconsistent over $\mathbb{SR}_{m\times n}^{P_1Q_1}$ and $\mathbb{SSR}_{m\times n}^{P_1Q_1}$. We employ Algorithm 2 to solve Problems 1, 2, 3 and 4 and the results are given in Table 1. In this table, "Iters" stands for the number of the required iterations for the convergence. Convergence history of the methods are presented in Figure 2. Apparently, the numerical results show that the proposed algorithm is effective.

TABLE 1. Numerical results for Example 4.1 in the inconsistent case.

	Problem 1	Problem 2	Problem 3	Problem 4
δ_k	8.17e-11	8.45e-11	2.95e-13	1.93e-12
Iters	22	23	13	13
η_k	1.29	75.42	1.71	49.04

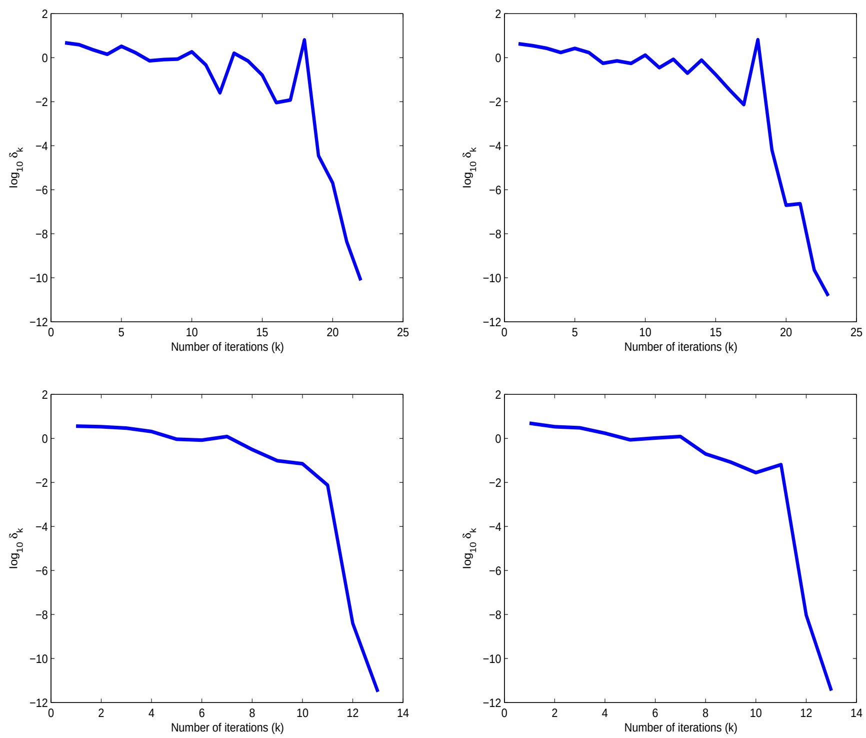

FIGURE 1. Convergence history for Example 4.1 (consistent case). Problem 1: top-left; Problem 2: top-right; Problem 3: down-left; Problem 4: down-right.

Example 4.2. In this example, we consider Problem 1 with

$$\begin{aligned}\mathcal{K}(X,Y) &= A_1XB_1 + X^T + C_1YD_1 + Y^T,\\ \mathcal{L}(X,Y) &= X + E_2X^TF_2 + Y + G_2Y^TH_2,\end{aligned}$$

where

$$\begin{aligned}A_1 &= \text{pentadiag}_n(2,-2,-6,1,1), && B_1 = \text{pentadiag}_n(-1,-2,0,1,1),\\ C_1 &= \text{tridiag}_n(-1,2,7), && D_1 = \text{tridiag}_n(2,-1,4),\\ E_2 &= \text{tridiag}_n(1,3,-1), && F_2 = \text{tridiag}_n(-1,6,3),\\ G_2 &= \text{pentadiag}_n(2,-1,-3,1,3), && H_2 = \text{pentadiag}_n(-2,0,2,3,2).\end{aligned}$$

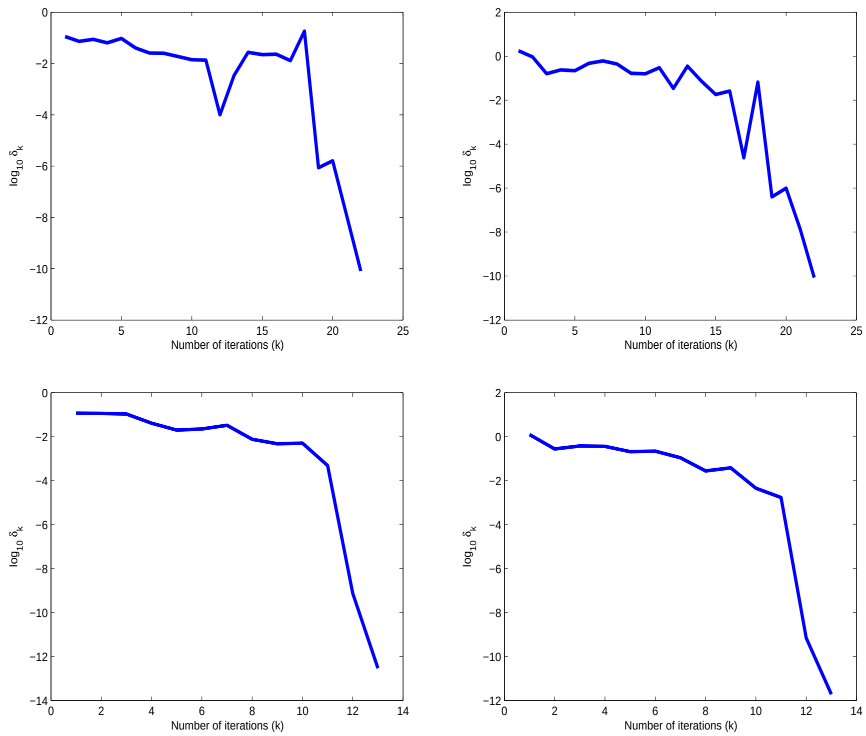

FIGURE 2. Convergence history for Example 4.1 (inconsistent case). Problem 1: top-left; Problem 2: top-right; Problem 3: down-left; Problem 4: down-right.

Presume that $p_1 = (1, 1, \ldots, 1)^T$, $p_2 = (n, n-1, \ldots, 1)^T$, $q_1 = (1, 2, \ldots, n)^T$, $q_2 = (1, 0, \ldots, 0)^T$, the matrices P_1, P_2, Q_1 and Q_2 are specified as follows:

$$P_i = I - 2\frac{p_i p_i^T}{p_i^T p_i}, \quad Q_i = I - 2\frac{q_i q_i^T}{q_i^T q_i}, \quad i = 1, 2,$$

where I signifies the identity matrix of order n. It is not difficult to verify that P_1, P_2, Q_1 and Q_2 are $n \times n$ symmetric orthogonal matrices. Suppose that $\tilde{X} = P_1(W_x + W_x^T)Q_1$ and $\tilde{Y} = P_2(W_y + W_y^T)Q_2$ where $W_x = \text{tridiag}_n(-1, 2, 1)$ and $W_y = \text{tridiag}_n(-1, 1, 2)$. Evidently, we have $\tilde{X} \in \mathbb{SR}_{n\times n}^{P_1Q_1}$ and $\tilde{Y} \in \mathbb{SR}_{n\times n}^{P_2Q_2}$. The matrices M and N are selected so that $M = \mathcal{K}(\tilde{X}, \tilde{Y})$ and $N = \mathcal{L}(\tilde{X}, \tilde{Y})$. Note that in this case the system

$$\left\{ \begin{array}{l} \mathcal{K}(X, Y) = M, \\ \mathcal{L}(X, Y) = N, \end{array} \right. \tag{4.8}$$

is consistent over $\mathbb{SR}_{m\times n}^{P_1Q_1} \times \mathbb{SR}_{m\times n}^{P_2Q_2}$. Now, let us focus on the solution of Problem 1 with the above data and $n = 500$. In order to handle Algorithm 2 for resolving Problem 1, we utilize $(X(0), Y(0)) = (0, 0)$ as an initial guess and the following stopping criterion

$$\delta_k = \max\{\|X(k) - X(k-1)\|_F, \|Y(k) - Y(k-1)\|_F\} < 10^{-10}.$$

TABLE 2. Numerical results for Example 4.2 (inconsistent case).

	Problem 1	Problem 2	Problem 3	Problem 4
δ_k	9.89e-11	7.14e-11	8.51e-11	8.27e-11
Iters	73	81	71	70
η_k	22.05	721.72	44.32	53.93

In this case the algorithm converges in 85 iterations and we have $\eta(85) = 2.27e-9$, where

$$\eta(k) = \max\{\|M - \mathcal{K}(X(k), Y(k))\|_F, \|N - \mathcal{L}(X(k), Y(k))\|_F\}.$$

The graph of $\log_{10} \delta_k$ is illustrated in Figure 3. To investigate the applicability of the algorithm to solve Problem 2, we choose $X_0 = Y_0 = I$, where I is the identity matrix of order n. All of the other assumptions are the same as before. In this case, Algorithm 2 converges in 85 iterations. The convergence performance of the method is illustrated in Figure 3.

To demonstrate the effectiveness of the algorithm for finding the solution pairs of Problems 3 and 4, we set $\tilde{X} = P_1(W_x - W_x^T)Q_1$ and $\tilde{Y} = P_2(W_y - W_y^T)Q_2$ where $W_x = \text{tridiag}_n(1, 0, 5)$ and $W_y = \text{tridiag}_n(1, 3, 2)$. Straightforward computations reveal that $\tilde{X} \in \mathbb{SSR}^{P_1Q_1}_{n\times n}$ and $\tilde{Y} \in \mathbb{SSR}^{P_2Q_2}_{n\times n}$. The matrices M and N are chosen such that $M = \mathcal{K}(\tilde{X}, \tilde{Y})$ and $N = \mathcal{L}(\tilde{X}, \tilde{Y})$. Therefore, the system (4.8) is consistent over $\mathbb{SSR}^{P_1Q_1}_{m\times n} \times \mathbb{SSR}^{P_2Q_2}_{m\times n}$. All of the other hypophyses are assumed to be those of the orthogonal symmetric case. In this case, Algorithm 2 for both of Problem 3 and Problem 4 converges in 80 iterations. For Problem 3, we have $\eta(80) = 2.43e-9$. In Figure 3, the convergence history of the method for the problems are displayed for further information.

Finally, we choose $M = \text{tridiag}_n(1, 1, 1)$ and $N = \text{pentadiag}_n(1, 1, 0, 1, 1)$. In this case, it can be verified that the system (4.8) is inconsistent over $\mathbb{SR}^{P_1Q_1}_{m\times n} \times \mathbb{SR}^{P_2Q_2}_{m\times n}$ and $\mathbb{SSR}^{P_1Q_1}_{m\times n} \times \mathbb{SSR}^{P_2Q_2}_{m\times n}$. Algorithm 2 was handled to solve our mentioned main problems. The obtained results are disclosed in Table 2 and convergence actions are delineated in Figure 4.

Example 4.3. In this example, we work on the inverse eigenvalue problem $HX = X\Lambda$ over Hamiltonian matrices, i.e, the unknown matrix H is determined so the H is a Hamiltonian matrix and minimizes $\|HX - X\Lambda\|_F$ over Hamiltonian matrices. This example has been previously examined in the literature, for instance see (Qian and Tan (2013), Table 1). For $m = 2\ell$, we choose ℓ eigenpairs $(\lambda_1, y_1), \ldots, (\lambda_\ell, y_\ell)$ of the $n \times n$ Hamiltonian matrix H of Example 16 from the benchmark examples for continuous algebraic Riccati equations (CARE) in Benner et al. (1995). We set $X = [Y_1,\ \hat{J}Y_1] \in \mathbb{R}^{n\times m}$ and $\Lambda = \text{diag}(\Lambda_1, -\Lambda_1) \in \mathbb{R}^{m\times m}$ where $Y_1 = [y_1, y_2, \ldots, y_\ell]$, $\Lambda_1 = \text{diag}\,(\lambda_1, \lambda_2, \ldots, \lambda_\ell)$ and $\hat{J}$ is defined by (1.1). Here our aim in Problem 1 is to determine least-squares Hamiltonian solution H of $HX = X\Lambda$. Note that Problem 2 is equivalent to Problem 1.2 in Qian and Tan (2013). To solve Problem 2, we choose $A = \text{rand}(n, n)$ and our goal is to find the least-squares Hamiltonian solution H^* of $HX = X\Lambda$ so that $\|A - H\|_F$ is minimized.

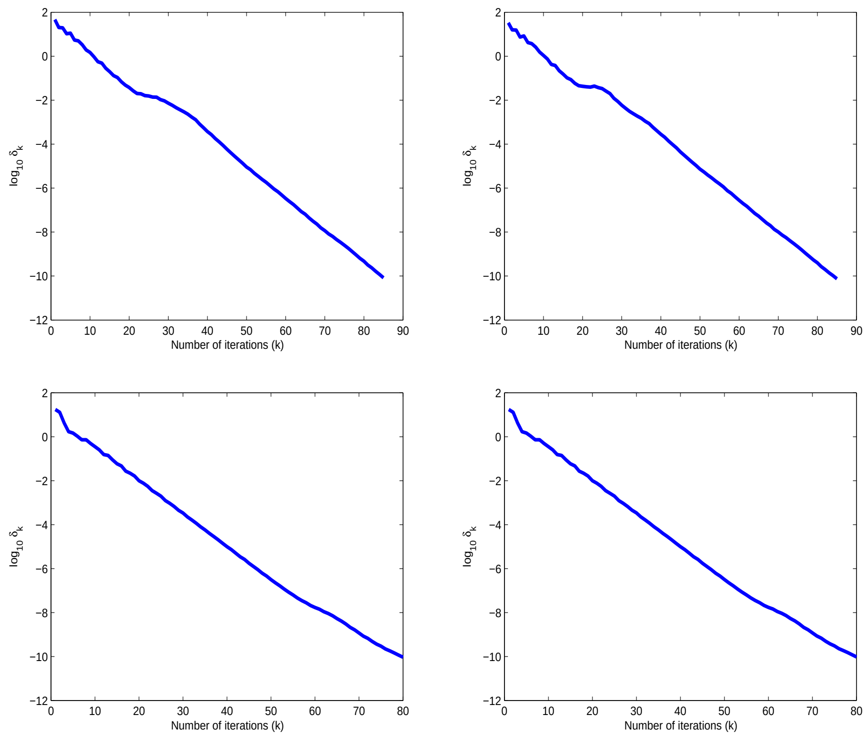

FIGURE 3. Convergence history for Example 4.2 (consistent case). Problem 1: top-left; Problem 2: top-right; Problem 3: down-left; Problem 4: down-right.

Now, we use Algorithm 2 to compute an approximate solutions of Problems 1 and 2. It is noted that in the algorithm, the matrices P_1 and Q_1 should be respectively replaced by I_n and $\hat{J}$. Numerical results for different values of n and m are presented in Tables 3 and 4. In these tables number of iterations (Iters) and CPU time (CPU) are presented and $\widetilde{H}(k)$ is the computed solution at iteration k.

From the reported numerical results in Tables 3 and 4, we can observe that Algorithm 2 computes good approximate solutions of mentioned problems in short times and small number of iterations.

5. Conclusion

It has been established that the set of all real square matrices can be written as a direct sum of the set of all (P, Q)-orthogonal symmetric and (P, Q)-orthogonal skew-symmetric matrices. Invoking the earlier pointed fact, we have scrutinized the extension of the well-known conjugate gradient least-squares (CGLS) method for finding the least-squares (P, Q)-orthogonal (skew-)symmetric solution pairs of general coupled matrix equations. Some numerical examples have been examined to illustrate the good performance of the proposed algorithm to solve the considered problems.

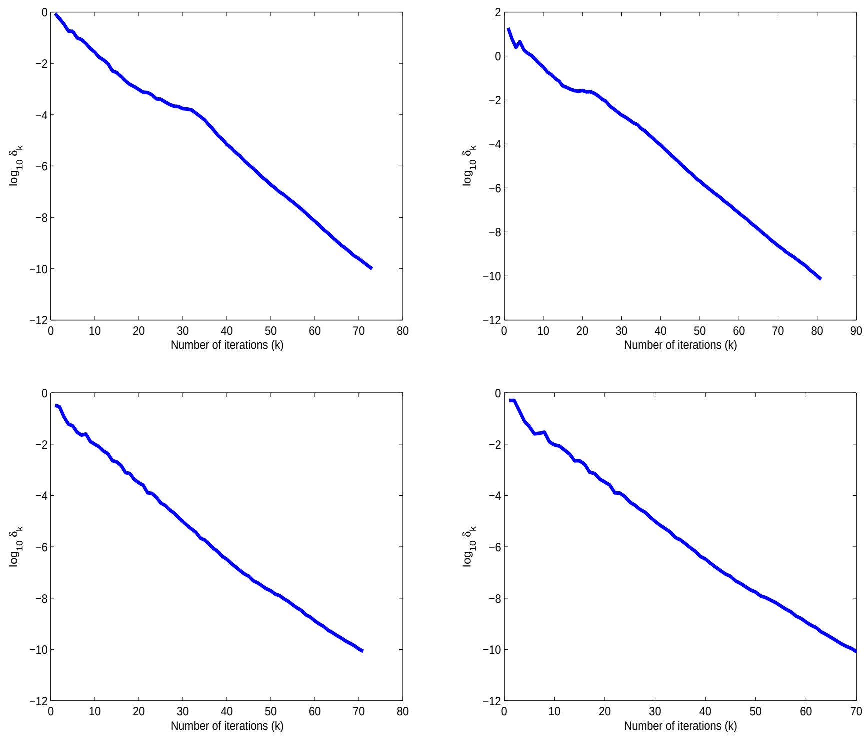

FIGURE 4. Convergence history for Example 4.2 (inconsistent case). Problem 1: top-left; Problem 2: top-right; Problem 3: down-left; Problem 4: down-right.

TABLE 3. Numerical results of Example 4.3 for solving Problem 1.

(n, m)	CPU	Iters	$\\|\widetilde{H}(k)X - X\Lambda\\|_F$	$\\|H - \widetilde{H}(k)\\|_F$
$(500, 120)$	0.021	2	$2.8962e-14$	40.3689
$(500, 250)$	0.025	2	$5.1415e-14$	21.8341
$(500, 300)$	0.027	2	$5.6861e-14$	16.8128
$(1000, 250)$	0.077	2	$5.1564e-14$	55.9557
$(1000, 500)$	0.114	2	$8.7826e-14$	30.8797
$(1000, 600)$	0.121	2	$9.7364e-14$	23.7755
$(1000, 800)$	0.136	2	$1.1231e-13$	14.3510
$(1500, 400)$	0.223	2	$7.2879e-14$	66.2312
$(1500, 750)$	0.286	2	$1.1569e-13$	37.8189
$(1500, 900)$	0.288	2	$1.2843e-13$	29.1186
$(1500, 1200)$	0.344	2	$1.4943e-13$	17.5762

Acknowledgments

The authors would like to express their heartfelt thanks to three anonymous referees for their valuable suggestions and constructive comments which have improved the quality of the paper.

TABLE 4. Numerical results of Example 4.3 for solving Problem 2.

(n, m)	CPU	Iters	$\\|\widetilde{H}(k)X - X\Lambda\\|_F$	$\\|A - \widetilde{H}(k)\\|_F$
$(500, 120)$	0.027	2	$4.6345e-13$	219.0886
$(500, 250)$	0.040	3	$1.6679e-13$	229.1630
$(500, 300)$	0.039	2	$3.9906e-13$	231.3716
$(1000, 250)$	0.119	2	$2.9313e-13$	434.7278
$(1000, 500)$	0.167	3	$3.8481e-13$	450.9314
$(1000, 600)$	0.208	3	$1.6326e-13$	461.7456
$(1000, 800)$	0.209	3	$6.7428e-13$	461.8669
$(1500, 400)$	0.348	3	$6.9589e-13$	651.2436
$(1500, 750)$	0.493	3	$1.0435e-12$	673.8064
$(1500, 900)$	0.630	3	$1.2253e-12$	680.7159
$(1500, 1200)$	0.781	3	$4.8160e-13$	689.2712

REFERENCES

[Beik and Salkuyeh, 2011] Beik FPA and Salkuyeh DK (2011) On the global Krylov subspace methods for solving general coupled matrix equation. *Computers and Mathematics with Applications* 62 (11): 4605–4613.

[Beik and Salkuyeh, 2013] Beik FPA and Salkuyeh DK (2013) The coupled Sylvester-transpose matrix equations over generalized centro-symmetric matrices. *International Journal of Computer Mathematics* 90 (7): 1546–1566.

[Beik et al., 2014] Beik FPA, Salkuyeh DK and Moghadam MM (2014) Gradient-based iterative algorithm for solving the generalized coupled Sylvester-transpose and conjugate matrix equations over reflexive (anti-reflexive) matrices. *Transactions of the Institute of Measurement and Control* 36 (1): 99–110.

[Benner et al., 1995] Benner P, Laub A, Mehrmann V (1995), *A collection of benchmark examples for the numerical solution of algebraic Riccati equations I: continuous-time case*. Tech. Report SPC 95_22, Fak. f. Mathematik, TU Chenmitz-Zwickau, 09107 Chemnitz, FRG.

[Bernstein, 2009] Bernstein DS (2009) *Matrix Mathematics: theory, facts, and formulas*. Second edition, Princeton University Press.

[Björck , 1996] Björck A (1996) *Numerical Methods for Least Squares Problems*. SIAM, Philadelphia.

[Cai and Chen, 2009] ai J and Chen G (2009) An iterative algorithm for the least squares bisymmetric solutions of the matrix equations $A_1XB_1 = C_1$, $A_2XB_2 = C_2$. *Mathematical and Computer Modelling* 50 (7–8): 1237–1244.

[Cai et al., 2010] Cai J, Chen G and Liu Q (2010) An iterative method for the bisymmetric solutions of the consistent matrix equations $A_1XB_1 = C_1$, $A_2XB_2 = C_2$. *International Journal of Computer Mathematics* 87 (12): 2706–2715.

[Chu and Golub, 2002] Chu MT and Golub GH (2002) Structured inverse eigenvalue problems. *Acta Numerica* 11: 1–71.

[Ding and Chen, 2006] Ding F and Chen T (2006) On iterative solutions of general coupled matrix equations. *SIAM Journal on Control and Optimization* 44 (6): 2269–2284.

[Ding et al., 2008] Ding F, Liu PX and Ding J (2008) Iterative solutions of the generalized Sylvester matrix equations by using the hierarchical identification principle. *Applied Mathematics and Computation* 197 (1): 41–50.

[Golub and Loan, 1996] Golub HG and Loan CFV (1996) *Matrix Computations*. (3rd ed.), Baltimore: Johns Hopkins

[Hajarian and Dehghan, 2011] Hajarian M and Dehghan M (2011) The generalized centro-symmetric and least squares generalized centro-symmetric solutions of the matrix equation $AYB + CY^TD = E$. *Mathematical Methods in the Applied Sciences* 34 (13): 1562–1579.

[Hajarian, 2014] Hajarian M (2014) Developing the CGLS algorithm for the least squares solutions of the general coupled matrix equations. *Mathematical Methods in the Applied Sciences* 37 (17): 2782–2798.

[Hajarian, 2014a] Hajarian M (2014a) A gradient-based iterative algorithm for generalized coupled Sylvester matrix equations over generalized centro-symmetric matrices. *Transactions of the Institute of Measurement and Control* 36 (2): 252–259.

[Hajarian, 2014b] Hajarian M (2014b) A matrix LSQR algorithm for solving constrained linear operator equations. *Bulletin of the Iranian Mathematical Society* 40 (1): 41–53.

[Julio and Soto, 2015] Julio AI and Soto RL (2015) Persymmetric and bisymmetric nonnegative inverse eigenvalue problem. *Linear Algebra and its Applications* 469: 130–152.

[Karimi and Dehghan, 2015] Karimi S and Dehghan M (2015) A general iterative approach for solving the general constrained linear matrix equations system. *Transactions of the Institute of Measurement and Control* DOI: 10.1177/0142331215581640.

[Kressner et al., 2009] Kressner D, Schrder C and Watkins DS (2009) Implicit QR algorithms for palindromic and even eigenvalue problems. *Numerical Algorithms* 51 (2): 209–238.

[Lancaster and Rozsa, 1983] Lancaster P and Rozsa P (1983) On the matrix equation $AX+X^*A^* = C$. *SIAM Journal on Algebraic Discrete Methods* 4 (4): 432–436.

[Laub, 1991] Laub AJ (1991) *Invariant subspace methods for the numerical solution of Riccati equations*. in: S. Bittanti, A.J. Laub, J.C. Willems (Eds.), The Riccati Equation, Springer, New York, pp. 163–196.

[Li et al., 2015] Li JF, Li W and Peng P (2015) A hybrid algorithm for solving minimization problem over (R,S)-symmetric matrices with the matrix inequality constraint. *Linear and Multilinear Algebra* 63 (5): 1049–1072.

[Li et al., 2014] Li H, Gao Z and Zhao D (2014) Least squares solutions of the matrix equation $AXB + CYD = E$ with the least norm for symmetric arrowhead matrices. *Applied Mathematics and Computation* 226: 719–724.

[Li and Huang, 2012] Li SK and Huang TZ (2012) LSQR iterative method for generalized coupled Sylvester matrix equations. *Applied Mathematical Modelling* 36 (8): 3545–3554.

[Peng et al., 2007] Peng ZH, Hu XY and Zhang L (2007) The bisymmetric solutions of the matrix equation $A_1X_1B_1 + A_2X_2B_2 + \cdots + A_\ell X_\ell B_\ell = C$ and its optimal approximation. *Linear Algebra and its Applications* 426 (2): 583–595.

[Peng and Liu, 2009] Peng ZH and Liu J (2009) The generalized bisymmetric solutions of the matrix equation $A_1X_1B_1 + A_2X_2B_2 + \cdots + A_\ell X_\ell B_\ell = C$ and its optimal approximation. *Numerical Algorithms* 50 (2): 127–144.

[Peng, 2012] Peng ZH and Zhou ZJ (2012) An efficient algorithm for the submatrix constraint of the matrix equation $A_1X_1B_1 + A_2X_2B_2 + \cdots + A_\ell X_\ell B_\ell = C$. *International Journal of Computer Mathematics* 89 (12): 1641–1662.

[Peng, 2013] Peng ZH (2013) The reflexive least squares solutions of the matrix equation $A_1X_1B_1 + A_2X_2B_2 + \cdots + A_\ell X_\ell B_\ell = C$ with a submatrix constraint. *Numerical Algorithms* 64 (3): 455–480.

[Peng and Xin, 2013] Peng ZH and Xin H(2013) The reflexive least squares solutions of the general coupled matrix equations with a submatrix constraint. *Applied Mathematics and Computation* 225: 425–445.

[Peng, 2015] Peng ZH (2015) The (R,S)-symmetric least squares solutions of the general coupled matrix equations. *Linear Multilinear Algebra* 63 (6): 1086–1105.

[Qian and Tan, 2013] Qian J and Tan RCE (2013) On some inverse eigenvalue problems for Hermitian and generalized Hamiltonian/skew-Hamiltonian matrices. *Journal of Computational and Applied Mathematics* 250: 28–38.

[Reichl, 2013] Reichl L (2013) *The transition to chaos: conservative classical systems and quantum manifestations.* Springer Science & Business Media.

[Saad, 1995] Saad Y (1995) *Iterative Methods for Sparse Linear Systems*. PWS Press, New York.

[Salkuyeh and Toutounian, 2006] Salkuyeh DK and Toutounian F (2006) New approaches for solving large Sylvester equations. *Applied Mathematics and Computation* 173 (1): 9–18.

[Salkuyeh and Beik, 2014] Salkuyeh DK and Beik FPA (2014) An iterative method to solve symmetric positive definite matrix equations. *Mathematical Reports* 16 (2): 271–283.

[Sarduvan et al., 2014] Sarduvan M, Şimşek S and Özdemir H (2014) On the best approximate (P, Q)-orthogonal symmetric and skew-symmetric solution of the matrix equation $AXB = C$. *Journal of Numerical Mathematics* 22 (3): 255-269.

[Sheng and Chen, 2007] Sheng X and Chen G (2007) A nite iterative method for solving a pair of linear matrix equations $(AXB, CXD) = (E, F)$. *Applied Mathematics and Computation* 189 (2): 1350–1358.

[Sheng and Chen, 2010] Sheng X and Chen G (2010) An iterative method for the symmetric and skew symmetric solutions of a linear matrix equation $AXB+CYD = E$. *Journal of Computational and Applied Mathematics* 233 (11): 3030–3040.

[Song et al., 2014] Song C, Feng J and Zhao J (2014) A new technique for solving continuous Sylvester-conjugate matrix equations. *Transactions of the Institute of Measurement and Control* 36 (11): 946–953.

[Su and Chen, 2010] Su Y and Chen G (2010) Iterative methods for solving linear matrix equation and linear matrix system. *International Journal of Computer Mathematics* 87 (4): 763–774.

[Wang, 2003] Wang RS (2003) *Functional Analysis and Optimization Theory.* Beijing Univ of Aeronautics Astronautics, Beijing.

[Wu et al., 2011a] Wu AG, Lv L and Duan GR (2011) Iterative algorithms for solving a class of complex conjugate and transpose matrix equations. *Applied Mathematics and Computation* 217 (21): 8343–8353.

[Wu et al., 2011b] Wu AG, Li B, Zhang Y and Duan GR (2011) Finite iterative solutions to coupled Sylvester-conjugate matrix equations. *Applied Mathematical Modelling* 35 (3): 1065–1080.

[Wu et al., 2012] Wu AG , Zhang E and Liu F (2012) On closed-form solutions to the generalized Sylvester-conjugate matrix equation. *Applied Mathematics and Computation* 218 (19): 9730–9741.

[Wu et al., 2013] Wu AG, Tong T and Duan GR (2013) Finite iterative algorithm for solving coupled Lyapunov equations appearing in continuous-time Markov jump linear systems. *International Journal of Systems Science* 44 (11): 2082–2093.

[Zhang et al., 2005] Zhang ZZ, Hu XY and Zhang L (2005) On the Hermitian-generalized Hamiltonian solutions of linear matrix equations. *SIAM Journal on Matrix Analysis and Applications* 27 (1): 294–303.

[Zhao et al., 2010] Zhao L, Chen G and Liu Q (2010) Least squares (P, Q)-orthogonal symmetric solutions of the matrix equation and its optimal approximation. *Electronic Journal of Linear Algebra* 20 (1): 537–551.

[Zhou et al., 1995] Zhou K, Doyle J and Glover K (1995) *Robust and Optimal Control.* Prentice-Hall, Upper Saddle River, New Jersey.

Appendix

In this part we prove Theorem 3.1 by mathematical induction. The theorem is established for (P, Q)-orthogonal symmetric matrices, for (P, Q)-orthogonal skew-symmetric matrices the proof follows from similar strategy.

In what follows, given $P_1, Q_1 \in \mathbb{SOR}^{p\times p}$ and $P_2, Q_2 \in \mathbb{SOR}^{r\times r}$, for simplicity, we use the linear operators $\mathcal{D}_x : \mathbb{R}^{m\times n} \times \mathbb{R}^{\nu\times n} \to \mathbb{R}^{p\times p}$ and $\mathcal{D}_y : \mathbb{R}^{m\times n} \times \mathbb{R}^{\nu\times n} \to \mathbb{R}^{r\times r}$ which are defined as follows:

$$\begin{aligned}
\mathcal{D}_x(Z, W) &= \frac{1}{2}\left(\bar{\mathcal{K}}(Z, W) + P_1Q_1\left(\bar{\mathcal{K}}(Z, W)\right)^T P_1Q_1\right),\\
\mathcal{D}_y(Z, W) &= \frac{1}{2}\left(\bar{\mathcal{L}}(Z, W) + P_1Q_1\left(\bar{\mathcal{L}}(Z, W)\right)^T P_1Q_1\right).
\end{aligned}$$

Proof of Theorem 3.1. For $k = 0, 1, 2, \ldots$, we set

$$\alpha_k = \frac{\|P_x(k)\|_F^2 + \|P_y(k)\|_F^2}{\|K_k\|_F^2 + \|L_k\|_F^2} \qquad \text{and} \qquad \beta_k = \frac{\|P_x(k+1)\|_F^2 + \|P_y(k+1)\|_F^2}{\|P_x(k)\|_F^2 + \|P_y(k)\|_F^2},$$

where $K_k = \mathcal{K}(Q_x(k), Q_y(k))$ and $L_k = \mathcal{L}(Q_x(k), Q_y(k))$. Because of commutative property of the inner product $\langle ., .\rangle$, we only need to prove the validity of (3.2), (3.3) and (3.4) for $0 \le i < j \le k$.

Step 1. For $k=1$, by Proposition 2.3, we have

$$\begin{aligned}
&\langle P_x(0), P_x(1)\rangle_F + \langle P_y(0), P_y(1)\rangle_F \\
&\quad = \langle P_x(0), \mathcal{D}_x(R_1(1), R_2(1))\rangle_F + \langle P_y(0), \mathcal{D}_y(R_1(1), R_2(1))\rangle_F \\
&\quad = \langle P_x(0), P_x(0)\rangle_F + \langle P_y(0), P_y(0)\rangle_F \\
&\qquad + \langle P_x(0), -\alpha_0 \mathcal{D}_x(K_0, L_0)\rangle_F + \langle P_y(0), -\alpha_0 \mathcal{D}_y(K_0, L_0)\rangle_F \\
&\quad = \langle P_x(0), P_x(0)\rangle_F + \langle P_y(0), P_y(0)\rangle_F \\
&\qquad - \alpha_0 \left(\langle K_0, L_0\rangle_F + \langle K_0, L_0\rangle_F\right) \\
&\quad = 0.
\end{aligned}$$

By Proposition 2.3 and using some straightforward computations, it reveals that

$$\begin{aligned}
\langle K_0, K_1\rangle_F + \langle L_0, L_1\rangle_F &= \beta_0 \left(\langle K_0, K_0\rangle_F + \langle L_0, L_0\rangle_F\right) \\
&\quad + \frac{1}{\alpha_0}\big(\langle R_1(0) - R_1(1), \mathcal{K}(P_x(1), P_y(1))\rangle_F \\
&\quad + \langle R_2(0) - R_2(1), \mathcal{L}(P_x(1), P_y(1))\rangle_F\big) \\
&= \beta_0 \left(\langle K_0, K_0\rangle_F + \langle L_0, L_0\rangle_F\right) \\
&\quad + \frac{1}{\alpha_0}\big(\langle P_x(0) - P_x(1), P_x(1)\rangle_F + \langle P_y(0) - P_y(1), P_y(1)\rangle_F\big) \\
&= 0.
\end{aligned}$$

It is not difficult to verify that

$$\langle P_x(1), Q_x(0)\rangle_F + \langle P_y(1), Q_y(0)\rangle_F = 0.$$

Consequently (3.2), (3.3) and (3.4) are valid for $k=1$.

Step 2. Let us assume that assertions (3.2), (3.3) and (3.4) are true for $k \le \nu$. Now, we show the validity of the conclusions for $k = \nu + 1$.

With a similar strategy used in the previous step and the assumption of the induction, it can be demonstrated that

$$\begin{aligned}
\langle P_x(\nu), P_x(\nu+1)\rangle_F + \langle P_y(\nu), P_y(\nu+1)\rangle_F &= \langle P_x(\nu), P_x(\nu)\rangle_F + \langle P_y(\nu), P_y(\nu)\rangle_F \\
&\quad - \alpha_\nu \left(\langle K_\nu, L_\nu\rangle_F + \langle K_\nu, L_\nu\rangle_F\right) = 0,
\end{aligned}$$

$$\begin{aligned}
\langle K_\nu, K_{\nu+1}\rangle_F + \langle L_\nu, L_{\nu+1}\rangle_F &= \beta_\nu \left(\langle K_\nu, K_\nu\rangle_F + \langle L_\nu, L_\nu\rangle_F\right) \\
&\quad + \frac{1}{\alpha_\nu}\big(\langle P_x(\nu) - P_x(\nu+1), P_x(\nu+1)\rangle_F \\
&\quad + \langle P_y(\nu) - P_y(\nu+1), P_y(\nu+1)\rangle_F\big) \\
&= 0.
\end{aligned}$$

and,

$$\begin{aligned}
&\langle P_x(\nu+1), Q_x(\nu)\rangle_F + \langle P_y(\nu+1), Q_y(\nu)\rangle_F \\
&\quad = \beta_{\nu-1}\big(\langle P_x(\nu+1), Q_x(\nu-1)\rangle_F + \langle P_y(\nu+1), Q_y(\nu-1)\rangle_F\big) \\
&\quad = \beta_{\nu-1}\big(\langle P_x(\nu) - \alpha_\nu \mathcal{D}_x(K_\nu, L_\nu), Q_x(\nu-1)\rangle_F \\
&\qquad + \langle P_y(\nu) - \alpha_\nu \mathcal{D}_y(K_\nu, L_\nu), Q_y(\nu-1)\rangle_F\big) \\
&\quad = -\alpha_\nu \beta_{\nu-1}\left(\langle K_\nu, K_{\nu-1}\rangle_F + \langle L_\nu, L_{\nu-1}\rangle_F\right) \\
&\quad = 0.
\end{aligned}$$

For $\ell = 0, 1, 2, \ldots, \nu - 1$, it can be deuced that

$$\begin{aligned}
\langle P_x(\nu+1), P_x(\ell)\rangle_F + \langle P_y(\nu+1), P_y(\ell)\rangle_F &= \langle P_x(\nu) - \alpha_\nu \mathcal{D}_x(K_\nu, L_\nu), P_x(\ell)\rangle_F \\
&\quad + \langle P_y(\nu) - \alpha_\nu \mathcal{D}_y(K_\nu, L_\nu), P_y(\ell)\rangle_F \\
&= -\alpha_\nu \left(\langle K_\nu, K_\ell\rangle_F + \langle L_\nu, L_\ell\rangle_F\right) \\
&\quad + \alpha_\nu \beta_{\ell-1} \left(\langle K_\nu, K_{\ell-1}\rangle_F + \langle L_\nu, L_{\ell-1}\rangle_F\right) \\
&= 0.
\end{aligned}$$

In addition, we can conclude that

$$\begin{aligned}
&\langle K_{\nu+1}, K_\ell\rangle_F + \langle L_{\nu+1}, L_\ell\rangle_F \\
&\quad = \langle \mathcal{K}(P_x(\nu+1), P_y(\nu+1)), K_\ell\rangle_F + \langle \mathcal{L}(P_x(\nu+1), P_y(\nu+1)), L_\ell\rangle_F \\
&\quad = \frac{1}{\alpha_\ell}\big(\langle \mathcal{K}(P_x(\nu+1), P_y(\nu+1)), R_1(\ell) - R_1(\ell+1)\rangle_F\big) \\
&\qquad + \frac{1}{\alpha_\ell}\big(\langle \mathcal{L}(P_x(\nu+1), P_y(\nu+1)), R_2(\ell) - R_2(\ell+1)\rangle_F\big) \\
&\quad = \frac{1}{\alpha_\ell}\left(\langle P_x(\nu+1), P_x(\ell) - P_x(\ell+1)\rangle_F\right) \\
&\qquad + \frac{1}{\alpha_\ell}\big(\langle P_y(\nu+1), P_y(\ell) - P_y(\ell+1)\rangle_F\big) \\
&\quad = 0,
\end{aligned}$$

and,

$$\begin{aligned}
&\langle P_x(\nu+1), Q_x(\ell)\rangle_F + \langle P_y(\nu+1), Q_y(\ell)\rangle_F \\
&\qquad = \langle P_x(\nu) - \alpha_\nu \mathcal{D}_x(K_\nu, L_\nu), Q_x(\ell)\rangle_F \\
&\qquad\quad + \langle P_y(\nu) - \alpha_\nu \mathcal{D}_y(K_\nu, L_\nu)), Q_y(\ell)\rangle_F \\
&\qquad = -\alpha_\nu \left(\langle K_\nu, K_\ell\rangle_F + \langle L_\nu, L_\ell\rangle_F\right) \\
&\qquad = 0.
\end{aligned}$$

From Steps 1 and 2, the proof is competed by the principle of the mathematical induction. $\square$