Journal of Computational Mathematics, Vol.25, No.2, 2007, 211–220.

MINIMIZATION PROBLEM FOR SYMMETRIC ORTHOGONAL ANTI-SYMMETRIC MATRICES *1)

Yuan Lei
(*College of Mathematics and Econometrics, Hunan University, Changsha 410082, China
Email: yleimath@yahoo.com*)

Anping Liao
(*Department of Information and Computing Science, Changsha University, Changsha 410003, China
College of Mathematics and Econometrics, Hunan University, Changsha 410082, China
Email: liaoap@hnu.cn*)

Lei Zhang
(*College of Mathematics and Econometrics, Hunan University, Changsha 410082, China*)

Abstract

By applying the generalized singular value decomposition and the canonical correlation decomposition simultaneously, we derive an analytical expression of the optimal approximate solution $\widehat{X}$, which is both a least-squares symmetric orthogonal anti-symmetric solution of the matrix equation $A^TXA = B$ and a best approximation to a given matrix X^*. Moreover, a numerical algorithm for finding this optimal approximate solution is described in detail, and a numerical example is presented to show the validity of our algorithm.

Mathematics subject classification: 15A24, 65F20, 65F22, 65K10.
Key words: Symmetric orthogonal anti-symmetric matrix, Generalized singular value decomposition, Canonical correlation decomposition.

1. Introduction

Denote the set of all symmetric (anti-symmetric) matrices in $\mathcal{R}^{n\times n}$ by $\mathcal{S}^{n\times n}(\mathcal{A}^{n\times n})$, the set of all orthogonal matrices in $\mathcal{R}^{n\times n}$ by $\mathcal{O}^{n\times n}$, the $n\times n$ identity matrix by I_n, the transpose and the Frobenius norm of a real matrix A by A^T and $\|A\|$, respectively. For $A=(a_{ij})\in\mathcal{R}^{n\times m}$, $B=(b_{ij})\in\mathcal{R}^{n\times m}$, $A*B$ represents the Hadamard product of the matrices A and B, that is, $A*B=(a_{ij}b_{ij})\in\mathcal{R}^{n\times m}$. Let $\mathcal{SOR}^{n\times n}$ be the set of all real $n\times n$ symmetric orthogonal matrices, i.e., $\mathcal{SOR}^{n\times n}=\{P|P=P^T, P^2=I_n, P\in\mathcal{R}^{n\times n}\}$.

Definition 1.1. *Given $P\in\mathcal{SOR}^{n\times n}$ and let $X\in\mathcal{S}^{n\times n}$.*

(1) The matrix X is called symmetric orthogonal symmetric with respect to P if $(PX)^T=PX$. The set of all $n\times n$ symmetric orthogonal symmetric matrices is denoted by $\mathcal{SS}_P^{n\times n}$;

(2) The matrix X is called symmetric orthogonal anti-symmetric matrix with respect to P if $(PX)^T=-PX$. The set of all symmetric orthogonal anti-symmetric matrices is denoted by $\mathcal{SA}_P^{n\times n}$.

---

* Received January 24, 2005; final revised March 1, 2006; accepted June 29, 2006.
1) The Project supported by Scientific Research Fund of Hunan Provincial Education Department, by National Natural Science Foundation of China (10171031), and by Natural Science Fundation of Hunan Province (03JJY6028).

The symmetric orthogonal (anti-)symmetric matrices play an important role in numerical analysis and matrix theory. For example, the matrix $X \in \mathcal{SS}_P^{n\times n}(\mathcal{SA}_P^{n\times n})$ can preserve the symmetric (anti-symmetric) structure after applying a Householder transformation, because the Householder matrix is symmetric and orthogonal. Let $J_n = [e_n, e_{n-1}, \ldots, e_1]$, where e_i denote the ith column of I_n. It is easy to verify that $J_n \in \mathcal{SOR}^{n\times n}$. If $P = J_n$, then $\mathcal{SS}_P^{n\times n}$ and $\mathcal{SA}_P^{n\times n}$ are a bi-symmetric matrix set [5, 12, 19] and a symmetric and skew anti-symmetric matrix set [18, 22], respectively, which have been applied in various areas [18, 24], such as information theory, linear system theory and numerical analysis. If $P = I_n$, then $\mathcal{SS}_P^{n\times n}$ is a symmetric matrix set and $\mathcal{SA}_P^{n\times n}$ is trivial due to the fact that $\mathcal{S}^{n\times n} \cap \mathcal{A}^{n\times n} = \{0\}$.

The symmetric orthogonal (anti-)symmetric matrices were initially considered by Zhou, Hu and Zhang, associated with matrix equations and inverse eigenvalue problems, see [27]. Peng [17] has investigated the symmetric orthogonal symmetric solution to the matrix equation

$$A^T XA = B, \tag{1.1}$$

which arose in an inverse problem of structural modification or the dynamic behaviour of a structure [2, 3, 4, 7, 13, 14]. The symmetric skew anti-symmetric solution of (1.1) and its optimal approximation were also obtained in [18] by using the generalized singular value decomposition (GSVD). However, it may happen that the matrix equation (1.1) is inconsistent due to the inaccuracies in the measured data. In this case, we may consider the solution of (1.1) in the least-squares sense [6, 12, 21]. The purpose of this paper is to extend the results in [18] to the least-squares problem with a symmetric orthogonal anti-symmetric constraint, which can be described as follows:

Problem 1.1. *Given matrices $A \in \mathcal{R}^{n\times m}$, $B \in \mathcal{S}^{m\times m}$, $P \in \mathcal{SOR}^{n\times n}$ and $X^* \in \mathcal{S}^{n\times n}$. Let*

$$\mathcal{S}_E = \left\{ X \mid X \in \mathcal{SA}_P^{n\times n}, \|A^T XA - B\| = \min_{Y\in\mathcal{SA}_P^{n\times n}} \|A^T YA - B\| \right\}. \tag{1.2}$$

Then find $\widehat{X} \in \mathcal{S}_E$ such that

$$\|\widehat{X} - X^*\| = \min_{X\in\mathcal{S}_E} \|X - X^*\|. \tag{1.3}$$

The minimization problem (1.3) arises in the structural modification and model updating [8]. The initial analytical matrix X^* is experimentally obtained from a practical measurement, but it may not satisfy the structural requirement or the minimum residual requirement. Hence, it is necessary to find the updated matrix $\widehat{X}$, which is not only a least-squares solution of matrix equation (1.1) with given structural requirement, but also a best approximation to the initial matrix X^*.

Similar to [12], the solution $\widehat{X}$ of Problem 1.1 can not be obtained by means of the canonical correlation decomposition (CCD) of a matrix pair, and the difficulty lies in the fact that the invariance of the Frobenius norm does not hold for general nonsingular matrices in CCD (see, for instance, (8) and (19) in [12]). In order to overcome this difficulty, a method, based on the projection theorem, GSVD and CCD, is adopted to solve Problem 1.1, and this approach has been applied successfully to find the least-squares solution of the matrix equations $(AXB, GXH) = (C, D)$ with minimum norm [15].

The outline of this paper is as follows. First, in Section 2, we will introduce several lemmas which will be used in the latter sections. Then, we will discuss Problem 1.1 and give the expression of its solution in Section 3. Finally, in Section 4, we will give the numerical algorithm to compute the solution of Problem 1.1 and report our numerical experiments.

2. Some Lemmas

As a preliminary, we briefly state the concepts of the GSVD and CCD, which are essential tools for deriving the solution of Problem 1.1. We refer to [9, 10, 16, 20] for details.

The GSVD can be described as follows: Given $A_1 \in \mathcal{R}^{k\times m}$ and $A_2 \in \mathcal{R}^{(n-k)\times m}$, then there exist orthogonal matrices $U \in \mathcal{O}^{k\times k}$, $V \in \mathcal{O}^{(n-k)\times(n-k)}$ and a nonsingular matrix $M \in \mathcal{R}^{m\times m}$ such that

$$A_1 = U(\Sigma_1, 0)M \quad \text{and} \quad A_2 = V(\Sigma_2, 0)M, \tag{2.1}$$

where $\Sigma_1 = \mathrm{diag}(I_r, S_1, 0)$ and $\Sigma_2 = \mathrm{diag}(0, S_2, I_{t-r-s})$ are block diagonal matrices (may not be square) with the same column partitioning, and the diagonal matrices S_1 and S_2 are given by

$$\begin{cases} S_1 = \mathrm{diag}(\lambda_1, \lambda_2, \ldots, \lambda_s), & 1 > \lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_s > 0, \\ S_2 = \mathrm{diag}(\mu_1, \mu_2, \ldots, \mu_s), & 0 < \mu_1 \leq \mu_2 \leq \ldots \leq \mu_s < 1, \quad \lambda_i^2 + \mu_i^2 = 1 \ (1 \leq i \leq s). \end{cases}$$

Here

$$t = \mathrm{rank}(A_1^T, A_2^T), \quad r = t - \mathrm{rank}(A_2), \quad s = \mathrm{rank}(A_1) + \mathrm{rank}(A_2) - t.$$

We further partition the nonsingular matrix

$$M^{-1} = \begin{pmatrix} M_1 & M_2 & M_3 & M_4 \end{pmatrix} \in \mathcal{R}^{m\times m} \ . \tag{2.2}$$

compatibly with the block column partitioning of $(\Sigma_1, 0)$, e.g., $M_1 \in \mathcal{R}^{m\times r}$, $M_2 \in \mathcal{R}^{m\times s}$.

The CCD can be described as follows: Given $A_1 \in \mathcal{R}^{k\times m}$ and $A_2 \in \mathcal{R}^{(n-k)\times m}$, then there exist nonsingular matrices $E_{A_1} \in \mathcal{R}^{k\times k}$, $E_{A_2} \in \mathcal{R}^{(n-k)\times(n-k)}$ and an orthogonal matrix $Q \in \mathcal{O}^{m\times m}$ such that

$$A_1^T = Q(\Pi_1, 0)E_{A_1}^{-1} \quad \text{and} \quad A_2^T = Q(\Pi_2, 0)E_{A_2}^{-1}, \tag{2.3}$$

where

$$\Pi_1 = \begin{pmatrix} \Xi_1 \\ \Xi_2 \end{pmatrix} \quad \text{and} \quad \Pi_2 = \begin{pmatrix} I_h \\ 0 \end{pmatrix}$$

are block matrices with the same row partitioning, $\Xi_1 = \mathrm{diag}(I_{r'}, C_A, 0)$ and $\Xi_2 = \mathrm{diag}(0, S_A, I_{t'})$ are block diagonal matrices (may not be square), and the diagonal matrices C_A and S_A are given by

$$\begin{cases} C_A = \mathrm{diag}(\alpha_1, \alpha_2, \ldots, \alpha_{s'}), & 1 > \alpha_1 \geq \alpha_2 \geq \ldots \geq \alpha_{s'} > 0, \\ S_A = \mathrm{diag}(\beta_1, \beta_2, \ldots, \beta_{s'}), & 0 < \beta_1 \leq \beta_2 \leq \ldots \leq \beta_{s'} < 1, \quad \alpha_i^2 + \beta_i^2 = 1 \ (1 \leq i \leq s'). \end{cases}$$

Here

$$\begin{aligned} &h = \mathrm{rank}(A_2), \quad r' = \mathrm{rank}(A_1) + \mathrm{rank}(A_2) - \mathrm{rank}(A_1^T, A_2^T), \\ &s' = \mathrm{rank}(A_2A_1^T) - r', \quad t' = \mathrm{rank}(A_1) - r' - s'. \end{aligned}$$

We further partition the orthogonal matrix

$$Q = \begin{pmatrix} Q_1 & Q_2 & Q_3 & Q_4 & Q_5 & Q_6 \end{pmatrix} \in \mathcal{O}^{m\times m} \tag{2.4}$$

compatibly with the row partitioning of Π_1, e.g., $Q_1 \in \mathcal{R}^{m\times r'}, Q_2 \in \mathcal{R}^{m\times s'}, Q_3 \in \mathcal{R}^{m\times(h-r'-s')}$.

Since $P \in \mathcal{SOR}^{n\times n}$, there exists an orthogonal matrix $H \in \mathcal{O}^{n\times n}$ such that

$$P = H\begin{pmatrix} I_k & 0 \\ 0 & -I_{n-k} \end{pmatrix}H^T, \quad k = \mathrm{rank}(I + P). \tag{2.5}$$

In fact, the representation (2.5) is a spectral decomposition of the matrix P (see [11]). By Definition 1.1 and the spectral decomposition of P, it is easy to prove that the structure of $\mathcal{SA}_P^{n\times n}$ has special form described in the following lemma (see also [26]).

Lemma 2.1.

$$\mathcal{SA}_P^{n\times n} = \left\{ X | X = H \begin{pmatrix} 0 & Y \\ Y^T & 0 \end{pmatrix} H^T, Y \in \mathcal{R}^{k\times(n-k)} \right\}. \tag{2.6}$$

Lemma 2.2. *Given matrices $D \in \mathcal{S}^{s\times s}$ and $E \in \mathcal{R}^{s\times s}$. Let $\Gamma = \mathrm{diag}(\gamma_1, \gamma_2, \cdots, \gamma_s)$ and $\Delta = \mathrm{diag}(\delta_1, \delta_2, \cdots, \delta_s)$ be given diagonal matrices of positive diagonal entries, satisfying $\gamma_j^2 + \delta_j^2 = 1$ $(j = 1, 2, \cdots, s)$. Then there exists a unique matrix $\bar{Y} \in \mathcal{R}^{s\times s}$ such that*

$$F(\bar{Y}) = \|\Gamma\bar{Y} + \bar{Y}^T\Gamma - D\|^2 + \|\Delta\bar{Y} - E^T\|^2 + \|\bar{Y}^T\Delta - E\|^2 = \min.$$

Moreover, the matrix $\bar{Y}$ possesses the analytical expression

$$\bar{Y} = \Phi * [\Gamma D + \Delta E^T - \Gamma(D\Gamma + E\Delta)\Gamma], \tag{2.7}$$

with

$$\Phi = (\phi_{ij}) \in \mathcal{R}^{s\times s}, \quad \phi_{ij} = \frac{1}{1 - \gamma_i^2\gamma_j^2}, \quad i, j = 1, 2, ..., s.$$

Proof. For matrices $D = (d_{ij}) \in \mathcal{S}^{s\times s}$, $E = (e_{ij}) \in \mathcal{R}^{s\times s}$ and $Y = (y_{ij}) \in \mathcal{R}^{s\times s}$, we have

$$F(Y) = \sum_{i,j} [(\gamma_i y_{ij} + y_{ji}\gamma_j - d_{ij})^2 + (\delta_i y_{ij} - e_{ji})^2 + (y_{ji}\delta_j - e_{ij})^2].$$

It then follows from straightforward operations that $F(\bar{Y}) = \min$ if and only if

$$\bar{y}_{ij} + \gamma_i\gamma_j\bar{y}_{ji} = \gamma_i d_{ij} + \delta_i e_{ji}. \tag{2.8}$$

Then (2.8) implies

$$\begin{cases} \bar{y}_{ij} + \gamma_i\gamma_j\bar{y}_{ji} = \gamma_i d_{ij} + \delta_i e_{ji}, \\ \bar{y}_{ji} + \gamma_i\gamma_j\bar{y}_{ij} = \gamma_j d_{ji} + \delta_j e_{ij}. \end{cases}$$

After suitable manipulations, we obtain

$$\bar{y}_{ij} = \frac{\gamma_i d_{ij} + \delta_i e_{ji} - \gamma_i(\gamma_j d_{ji} + \delta_j e_{ij})\gamma_j}{1 - \gamma_i^2\gamma_j^2}. \tag{2.9}$$

From (2.9) we immediately get (2.7).

3. The Solution of Problem 1.1

In order to find the solution $\widehat{X}$ of Problem 1.1, we first transform the least-squares problem (1.2) with respect to the inconsistent matrix equation (1.1) to a consistent one by applying the projection theorem.

Theorem 3.1. *Given matrices $A \in \mathcal{R}^{n\times m}$, $B \in \mathcal{S}^{m\times m}$ and $P \in \mathcal{SOR}^{n\times n}$. Let $X_0 \in \mathcal{S}_E$, and define*

$$B_0 = A^T X_0 A. \tag{3.1}$$

Then the matrix equation

$$A^T X A = B_0 \tag{3.2}$$

is consistent, and its symmetric orthogonal anti-symmetric solution set is the same as the least-squares symmetric orthogonal anti-symmetric solution set $\mathcal{S}_E$ of the matrix equation (1.1).

Proof. Let

$$\mathcal{L} = \{Z \mid Z = A^T X A, \quad X \in \mathcal{SA}_P^{n\times n}\}.$$

Then $\mathcal{L}$ is obviously a linear subspace of $\mathcal{S}^{n\times n}$. Because X_0 is a least-squares solution of the matrix equation (1.1) over $\mathcal{SA}_P^{n\times n}$, from (3.1) we see that $B_0 \in \mathcal{L}$ and

$$\|B_0 - B\| = \|A^T X_0 A - B\| = \min_{X\in\mathcal{SA}_p^{n\times n}} \|A^T X A - B\| = \min_{Z\in\mathcal{L}} \|Z - B\|.$$

Now, by applying the projection theorem we have

$$(B_0 - B) \perp \mathcal{L} \quad \text{or} \quad (B_0 - B) \in \mathcal{L}^{\perp}.$$

For every $X \in \mathcal{SA}_P^{n\times n}$, we know that $(A^T XA - B_0) \in \mathcal{L}$. It then follows that

$$\begin{aligned} \|A^T XA - B\|^2 &= \|(A^T XA - B_0) + (B_0 - B)\|^2 \\ &= \|A^T XA - B_0\|^2 + \|B_0 - B\|^2, \end{aligned}$$

which implies that the conclusion of this theorem holds.

From Theorem 3.1, we easily see that the optimal approximate solution $\widehat{X}$ of the consistent matrix equation (3.2) to a given matrix X^* over $\mathcal{SA}_P^{n\times n}$ is nothing but the solution of Problem 1.1. Therefore, solving Problem 1.1 essentially reduces to find B_0, and the crux of finding B_0 is to derive a least-squares solution X_0 of the matrix equation (1.1) over $\mathcal{SA}_P^{n\times n}$. The following theorem give the general expression of these least-squares solutions.

Theorem 3.2. *Given matrices $A \in \mathcal{R}^{n\times m}$, $B \in \mathcal{S}^{m\times m}$ and $P \in \mathcal{SOR}^{n\times n}$. Let the spectral decomposition of P be* (2.5)*, and partition the matrix $A^T H$ into*

$$A^T H = (A_1^T, A_2^T) \in \mathcal{R}^{m\times n}, \quad A_1 \in \mathcal{R}^{k\times m}, \quad A_2 \in \mathcal{R}^{(n-k)\times m}. \tag{3.3}$$

Decompose the matrix pair (A_1^T, A_2^T) by using CCD as (2.3)*, and denote*

$$B' = (B'_{ij})_{6\times 6}, \quad B'_{ij} = Q_i^T B Q_j, \quad i,j = 1,2,\cdots,6, \tag{3.4}$$

where the matrices Q_i $(i = 1,2,\ldots,6)$ are given by (2.4)*. Then the set $\mathcal{S}_E$ in Problem* 1.1 *can be expressed as*

$$\mathcal{S}_E = \left\{ X | X = H \begin{pmatrix} 0 & Y \\ Y^T & 0 \end{pmatrix} H^T \right\}, \tag{3.5}$$

where

$$Y = E_{A_1} \begin{pmatrix} \frac{1}{2}B'_{11} + R & B'_{12} - (B'_{15})^T C_A S_A^{-1} & B'_{13} & Y_{14} \\ S_A^{-1}(B'_{15})^T & Y_{22} & Y_{23} & Y_{24} \\ (B'_{16})^T & (B'_{26})^T & (B'_{36})^T & Y_{34} \\ Y_{41} & Y_{42} & Y_{43} & Y_{44} \end{pmatrix} E_{A_2}^T, \tag{3.6}$$

with

$$\begin{aligned} &Y_{22} = K * [C_A B'_{22} + S_A (B'_{25})^T - C_A (B'_{22} C_A + B'_{25} S_A) C_A], \\ &Y_{23} = C_A B'_{23} + S_A (B'_{35})^T. \end{aligned} \tag{3.7}$$

Here

$$K = (k_{ij}) \in \mathcal{R}^{s'\times s'}, \quad k_{ij} = \frac{1}{1 - \alpha_i^2 \alpha_j^2}, \quad i,j = 1,2,...,s',$$

$R \in \mathcal{A}^{s'\times s'}$ is an arbitrary anti-symmetric matrix and the other unknown matrix blocks are arbitrary.

Proof. Partition the matrix $E_{A_1}^{-1} Y E_{A_2}^{-T}$ compatibly to the block column partitioning of $(\Sigma_{A_1}, 0)$ and $(\Sigma_{A_2}, 0)$, respectively, into

$$E_{A_1}^{-1} Y E_{A_2}^{-T} = (Y_{ij})_{4\times 4}. \tag{3.8}$$

By Lemma 2.1 we know the matrix $X \in \mathcal{SA}_P^n$ is of the form (2.6), and from (3.3) and (2.3) we have

$$\begin{aligned} &\|A^T XA - B\|^2 = \left\| (A_1^T, A_2^T) \begin{pmatrix} 0 & Y \\ Y^T & 0 \end{pmatrix} \begin{pmatrix} A_1 \\ A_2 \end{pmatrix} - B \right\|^2 \\ = \;& \|(\Pi_1, 0) E_{A_1}^{-1} Y E_{A_2}^{-T} (\Pi_2, 0)^T + (\Pi_2, 0) E_{A_2}^{-1} Y^T E_{A_1}^{-T} (\Pi_1, 0)^T - Q^T B Q\|^2. \end{aligned}$$

After substituting (2.3), (3.4) and (3.8) into the equation above, we know that the minimization problem $\min_{X\in\mathcal{SA}_P^{n\times n}} \|A^TXA-B\|$ is reduced to solving the following six minimization problems:

$$\|Y_{12}+Y_{21}^TC_A-B'_{12}\|^2+\|Y_{21}^TS_A-B'_{15}\|^2=\min, \tag{3.9}$$

$$\|Y_{13}-B'_{13}\|^2=\min, \quad \|Y_{3i}-(B'_{i6})^T\|^2=\min \quad i=1,2,3, \tag{3.10}$$

$$\|Y_{11}+Y_{11}^T-B'_{11}\|^2=\min, \quad \|Y_{23}^TS_A-B'_{35}\|^2+\|C_AY_{23}-B'_{23}\|^2=\min, \tag{3.11}$$

$$\|C_AY_{22}+Y_{22}^TC_A-B'_{22}\|^2+\|Y_{22}^TS_A-B'_{25}\|^2+\|S_AY_{22}-(B'_{25})^T\|^2=\min. \tag{3.12}$$

It follows from (3.9) and (3.10) that

$$\begin{aligned} &Y_{21}=S_A^{-1}(B'_{15})^T; \quad Y_{12}=B'_{12}-B'_{15}C_AS_A^{-1};\\ &Y_{13}=B'_{13}; \quad Y_{3i}=(B'_{i6})^T,\ i=1,2,3. \end{aligned} \tag{3.13}$$

Following (3.11) and Lemma 2.1 in [23], we have

$$Y_{11}=\frac{1}{2}B'_{11}+R, \quad \forall R\in\mathcal{A}^{s'\times s'}; \quad Y_{23}=C_AB'_{23}+S_A(B'_{35})^T. \tag{3.14}$$

From Lemma 2.2, we know that the solution of (3.12) can be expressed as

$$Y_{22}=K*[C_AB'_{22}+S_A(B'_{25})^T-C_A(B'_{22}C_A+B'_{25}S_A)C_A]. \tag{3.15}$$

By combining (3.13)-(3.15), we arrive at (3.6), and from Lemma 2.1 we obtain (3.5).

By Theorem 3.2, we know that the matrix $X_0\in\mathcal{S}_E$ can be expressed as

$$X_0=H\begin{pmatrix} 0 & Y\\ Y^T & 0\end{pmatrix}H^T, \tag{3.16}$$

where Y is given by (3.6). Following (3.1), (3.3) and (3.16), we have

$$\begin{aligned} B_0 &= A^TH\begin{pmatrix} 0 & Y\\ Y^T & 0\end{pmatrix}H^TA=(A_1^T,A_2^T)\begin{pmatrix} 0 & Y\\ Y^T & 0\end{pmatrix}\begin{pmatrix} A_1\\ A_2\end{pmatrix}\\ &= A_1^TYA_2+A_2^TY^TA_1. \end{aligned} \tag{3.17}$$

Substituting the matrices A_1, A_2 in (2.3) and Y in (3.6) into (3.17), after obvious manipulations we can immediately get the expression of the matrix B_0 defined in (3.1) as follows:

$$B_0=Q\begin{pmatrix} B'_{11} & B'_{12} & B'_{13} & 0 & B'_{15} & B'_{16}\\ (B'_{12})^T & C_AY_{22}+Y_{22}^TC_A & C_AY_{23} & 0 & Y_{22}^TS_A & B'_{26}\\ (B'_{13})^T & Y_{23}^TC_A & 0 & 0 & Y_{23}^TS_A & B'_{36}\\ 0 & 0 & 0 & 0 & 0 & 0\\ (B'_{15})^T & S_AY_{22} & S_AY_{23} & 0 & 0 & 0\\ (B'_{16})^T & (B'_{26})^T & (B'_{36})^T & 0 & 0 & 0 \end{pmatrix}Q^T, \tag{3.18}$$

where Y_{22} and Y_{23} are given by (3.7).

Remark 3.1. The result (3.18) shows that the matrix B_0 given in Theorem 3.1 is unique. Furthermore, we can conclude that

$$\|B_0-B\|=\min_{X\in\mathcal{SA}_P^{n\times n}}\|A^TXA-B\|.$$

Obviously, we can not obtain the optimal approximate solution $\widehat{X}$ to a given matrix X^* by using (3.5) in the sense of the Frobenius norm due to that the matrices E_{A_1} and E_{A_2} in (3.6) may not be orthogonal. But Problem 1.1 can be transformed to the problem of finding the optimal approximate solution of the consistent matrix equation (3.2) over $\mathcal{SA}_P^{n\times n}$, where B_0 is determined by (3.18). Hence, similar to [18], we can obtain the solution $\widehat{X}$ of Problem 1.1 by using GSVD.

Theorem 3.3. *Assume the conditions of Theorem* 3.2. *Decompose the matrix pair* (A_1, A_2) *by GSVD as* (2.1) *and define*

$$M^{-T}B_0M^{-1} = (\widetilde{B}_{ij})_{4\times 4}, \quad \widetilde{B}_{ij} = M_i^T B_0 M_j \quad i,j = 1,2,3,4, \tag{3.19}$$

where the matrices M_i $(i = 1,2,3,4)$ *are given by* (2.2). *Partition the matrices* $H^T X^* H$ *and* $U^T X_{12}^* V$ *into*

$$H^T X^* H = \begin{array}{cc} \begin{pmatrix} X_{11}^* & X_{12}^* \\ X_{12}^{*T} & X_{22}^* \end{pmatrix} & \begin{array}{c} k \\ n-k \end{array} \\ \begin{array}{cc} n & \quad n-k \end{array} & \end{array},$$

$$U^T X_{12}^* V = \begin{array}{cc} \begin{pmatrix} Z_{11}^* & Z_{12}^* & Z_{13}^* \\ Z_{21}^* & Z_{22}^* & Z_{23}^* \\ Z_{31}^* & Z_{32}^* & Z_{33}^* \end{pmatrix} & \begin{array}{c} r \\ s \\ k-r-s \end{array} \\ \begin{array}{ccc} n-k+r-t & s & t-r-s \end{array} & \end{array}. \tag{3.20}$$

Then the solution $\widehat{X}$ *of Problem* 1.1 *can be expressed as*

$$\widehat{X} = H \begin{pmatrix} 0 & \widehat{Y} \\ \widehat{Y}^T & 0 \end{pmatrix} H^T \quad \textit{with} \quad \widehat{Y} = U \begin{pmatrix} Z_{11}^* & \widetilde{B}_{12}S_2^{-1} & \widetilde{B}_{13} \\ Z_{21}^* & \widehat{Y}_{22} & S_1^{-1}\widetilde{B}_{23} \\ Z_{31}^* & Z_{32}^* & Z_{33}^* \end{pmatrix} V^T, \tag{3.21}$$

where

$$\begin{aligned} \widehat{Y}_{22} = {} & \frac{1}{2}S_1^{-1}\widetilde{B}_{22}S_2^{-1} + W * [S_1S_2^{-1}(\frac{1}{2}S_2^{-1}\widetilde{B}_{22}S_1^{-1} - Z_{22}^{*T}) \\ & -(\frac{1}{2}S_1^{-1}\widetilde{B}_{22}S_2^{-1} - Z_{22}^*)S_1S_2^{-1}], \end{aligned} \tag{3.22}$$

with

$$W = (w_{ij}) \in \mathcal{R}^{s\times s}, \quad w_{ij} = \frac{\lambda_j \mu_i^2 \mu_j}{\lambda_i^2\mu_j^2 + \lambda_j^2\mu_i^2}, \quad i,j = 1,2,...,s.$$

Proof. From (3.17), we know that the problem of finding the symmetric orthogonal anti-symmetric solution of the consistent matrix equation (3.2) is reduced to finding $\widetilde{Y} \in \mathcal{R}^{k\times(n-k)}$ such that

$$A_1^T \widetilde{Y} A_2 + A_2^T \widetilde{Y}^T A_1 = B_0. \tag{3.23}$$

Therefore, from Theorem 2.1 in [25] and Lemma 2.1, we can obtain a new expression of the set $\mathcal{S}_E$ in Problem 1.1 as follow:

$$\mathcal{S}_E = \left\{ X | X = H \begin{pmatrix} 0 & \widetilde{Y} \\ \widetilde{Y}^T & 0 \end{pmatrix} H^T \right\} \text{ with } \widetilde{Y} = U \begin{pmatrix} Y_{11} & \widetilde{B}_{12}S_2^{-1} & \widetilde{B}_{13} \\ Y_{21} & \widetilde{Y}_{22} & S_1^{-1}\widetilde{B}_{23} \\ Y_{31} & Y_{32} & Y_{33} \end{pmatrix} V^T, \tag{3.24}$$

where $\widetilde{Y}_{22} = \frac{1}{2}S_1^{-1}\widetilde{B}_{22}S_2^{-1} + G^T - S_2S_1^{-1}GS_1S_2^{-1}$, and $G, Y_{i1}, Y_{3j}, (i,j = 1,2,3)$ are arbitrary.

Obviously, the set $\mathcal{S}_E$ is nonempty and is a closed convex subset of the Hilbert space $\mathcal{R}^{n\times n}$. It follows from the optimal approximation theorem [1] that there exists a unique matrix $\widehat{X} \in \mathcal{S}_E$

satisfying (1.3). For all $X \in \mathcal{S}_E$, it follows from (3.20) and (3.24) that

$$\begin{aligned}
&\|X - X^*\|^2 = \|H^T X H - H^T X^* H\|^2 \\
= & \left\| \begin{pmatrix} 0 & \widetilde{Y} \\ \widetilde{Y}^T & 0 \end{pmatrix} - \begin{pmatrix} X_{11}^* & X_{12}^* \\ X_{12}^{*T} & X_{22}^* \end{pmatrix} \right\|^2 \\
= & \|-X_{11}^*\|^2 + \|U^T\widetilde{Y}V - U^T X_{12}^* V\|^2 + \|V^T\widetilde{Y}^T U - V^T X_{12}^{*T} U\|^2 + \|-X_{22}^*\|^2 \\
= & \|X_{11}^*\|^2 + \left\| \begin{pmatrix} Y_{11} - Z_{11}^* & \widetilde{B}_{12}S_2^{-1} - Z_{12}^* & \widetilde{B}_{13} - Z_{13}^* \\ Y_{21} - Z_{21}^* & \widetilde{Y}_{22} - Z_{22}^* & S_1^{-1}\widetilde{B}_{23} - Z_{23}^* \\ Y_{31} - Z_{31}^* & Y_{32} - Z_{32}^* & Y_{33} - Z_{33}^* \end{pmatrix} \right\|^2 \\
& + \left\| \begin{pmatrix} Y_{11}^T - Z_{11}^{*T} & Y_{21}^T - Z_{21}^{*T} & Y_{31}^T - Z_{31}^{*T} \\ S_2^{-1}\widetilde{B}_{12}^T - Z_{12}^{*T} & \widetilde{Y}_{22}^T - Z_{22}^{*T} & Y_{32}^T - Z_{32}^{*T} \\ \widetilde{B}_{13}^T - Z_{13}^{*T} & \widetilde{B}_{23}^T S_1^{-1} - Z_{23}^{*T} & Y_{33}^T - Z_{33}^{*T} \end{pmatrix} \right\|^2 + \|X_{22}^*\|^2.
\end{aligned}$$

Hence, $\|X - X^*\| = \min$ if and only if

$$Y_{i1} = Z_{i1}^*, \quad Y_{3j} = Z_{3j}^*, \quad i, j = 1, 2, 3, \tag{3.25}$$

and

$$\begin{aligned}
&\|\widetilde{Y}_{22} - Z_{22}^*\| \\
= & \|\frac{1}{2}S_1^{-1}\widetilde{B}_{22}S_2^{-1} + G^T - S_2S_1^{-1}GS_1S_2^{-1} - Z_{22}^*\| = \min, \quad \forall\, G \in \mathcal{R}^{s\times s}.
\end{aligned} \tag{3.26}$$

By making use of Lemma 2.2 in [25] we know that the solution of (3.26) is

$$\begin{aligned}
\widehat{Y}_{22} \;=\; & \frac{1}{2}S_1^{-1}\widetilde{B}_{22}S_2^{-1} + W * [S_1S_2^{-1}(\frac{1}{2}S_2^{-1}\widetilde{B}_{22}S_1^{-1} - Z_{22}^{*T}) \\
& -(\frac{1}{2}S_1^{-1}\widetilde{B}_{22}S_2^{-1} - Z_{22}^*)S_1S_2^{-1}].
\end{aligned}$$

Now, after substituting this $\widehat{Y}_{22}$ and Y_{i1}, Y_{3j} $(i, j = 1, 2, 3)$ in (3.25) into (3.24), we get (3.21).

4. Numerical Algorithm and Example

Based on Theorem 3.3, we establish a direct algorithm for finding the solution of Problem 1.1 as follows:

1) Input matrices A, B, X^* and P;
2) Form the spectral decomposition of P as (2.5), and determine the matrices A_1 and A_2 by (3.3);
3) Find the CCD of the matrix pair (A_1^T, A_2^T) as (2.3), and determine the matrices B'_{ij} by (3.4);
4) Compute the matrix B_0 by (3.18);
5) Find the GSVD of the matrix pair (A_1, A_2) as (2.1), and determine the matrices $\widetilde{B}_{ij}$ by (3.19);
6) Compute the matrix $\widehat{Y}_{22}$ by (3.22), and determine the matrix blocks Z_{ij}^* by (3.20);
7) Compute the solution $\widehat{X}$ of Problem 1.1 by (3.21).

Example 4.1. *Let*

$$A = \begin{pmatrix} toeplitz(1:k) & I_k \\ ones(k) & I_k \end{pmatrix}, \quad P = \frac{1}{\sqrt{2}}\begin{pmatrix} -I_k & -J_k \\ -J_k & I_k \end{pmatrix},$$

where $toeplitz(1:k)$ denotes Toeplitz matrix of order k with its first rows being $(1, 2, \ldots, k)$; and $ones(k)$ denotes the matrix of order k whose all elements are one.

For convenience, we construct the matrix $X \in \mathcal{SA}_P^{n\times n}$ by Lemma 2.1 with $Y = ones(k)$, and $B = A^T[X + \varepsilon \cdot ones(2k)]A$, where ε is an arbitrary nonnegative number. If we take $X^* = X$, then the matrix X is exactly the unique solution of Problem 1.1 when $\varepsilon = 0$ due to the nonsingularity of the matrix A. Moreover, we can theoretically show that the solution $\widehat{X}$ of Problem 1.1 approaches to X as ε goes to zero. Our numerical results are listed in Table 1.

Table 1: Numerical results for Example 4.1

k	ε	$\Vert\widehat{X}-X\Vert$	$\Vert A^TXA-B\Vert$	$\Vert A^T\widehat{X}A-B\Vert$	$\Vert B-B_0\Vert$
10	1	74.0568	2.8658e+4	2.8097e+4	2.8097e+4
	10e-2	0.7406	286.5800	280.9744	280.9744
	10e-4	0.0074	2.8658	2.8097	2.8097
	10e-6	7.4059e-5	0.0287	0.0281	0.0281
50	10e-2	0.5556	4.5257e+5	4.5256e+5	4.5256e+5
	10e-4	0.0056	4.5257e+3	4.5256e+3	4.5256e+3
	10e-6	5.5565e-5	45.2573	45.2563	45.2563
	10e-8	5.5565e-7	0.4526	0.4526	0.4526
100	10e-4	0.0069	1.3037e+5	1.3037e+5	1.3037e+5
	10e-6	6.8580e-5	1.3037e+3	1.3037e+3	1.3037e+3
	10e-8	6.8581e-7	13.0374	13.0374	13.0374
	10e-10	6.8581e-9	0.1304	0.1304	0.1304

The example shows that the distance between B and B_0 goes to zero as $\widehat{X}$ approaches X, and from the results in Table 1, we can also conclude

$$\Vert A^TXA - B\Vert \geq \Vert A^T\widehat{X}A - B\Vert = \Vert B - B_0\Vert.$$

These features are in accordance with the theory established in this paper. Therefore, the above-described algorithm is valid for solving Problem 1.1.

References

[1] J.P. Aubin, Appliced Functional Analysis, John Wiley & Sons, Inc, 1979.

[2] A. Berman, Mass matrix correction using an incomplete set of measure model, *AIAA J.*, **17** (1979), 1147-1148.

[3] A. Berman and E. J. Nagy, Improvement of a large analytical model using test data, *AIAA J.*, **21** (1983), 1168-1173.

[4] H. Dai and P. Lancaster, Linear matrix equations from an inverse problem of vibration theory, *Linear Algebra Appl.*, **246** (1996), 31-47.

[5] H. Dai, Computing a nearest p-symmetric nonnegative definite matrix under linear restriction, *J. Comput. Math.*, **22** (2004), 671-680.

[6] Y.-B. Deng, X.-Y. Hu and L. Zhang, The symmetric and symmetric positive semidefinite solutions of linear matrix equation $B^TXB = D$ on linear manifolds, *Numer. Math. J. Chineses Univ.*, **12** (2003), 186-192.

[7] Y.-B. Deng and X.-Y. Hu, On solution of the linear matrix equation $AXA^T + BYB^T = C$, *J. Comput. Math.*, **23** (2005), 17-26.

[8] M.I. Friswell and J.E. Mottershead, Finite Element Model Updating in Strctual Dynamics, Kluwer Academic Publication, 1995.

[9] G.H. Golub and H. Zha, Perturbation analysis of the canonical correlations of matrix pairs, *Linear Algebra Appl.*, **210** (1994), 3-28.

[10] G.H. Golub and C.F. Van Loan, Matrix Computations, 3rd ed., Johns Hopkins University Press, Baltimore and London, 1996.

[11] P. Lancaster and M. Tismenetsky, The Theory of Matrices, 2rd ed., Academic Press, New York, 1985.

[12] A.-P. Liao and Z.-Z. Bai, Least-squares solutions of the matrix equation $A^TXA = D$ in bisymmetric matrix set, *Math. Numer. Sinica*, **24** (2002), 9-20 (in Chinese).

[13] A.-P. Liao and Z.-Z. Bai, The constrained solutions of two matrix equations, *Acta Math. Sinica*, **18** (2002), 671-678.

[14] A.-P. Liao and Z.-Z. Bai, Least-squares solution of $AXB = D$ over symmetric positive semidefinite matrices X, *J. Comput. Math.*, **21** (2003), 11-18.

[15] A.-P. Liao and Y. Lei, Least-squares solution with the minimum-norm for the matrix equation $(AXB, GXH) = (C, D)$, *J. Comput. Math. Appl.*, **50** (2005), 539-549.

[16] C.C. Paige and M.A. Saunders, Towards a generalized singular value decomposition, *SIAM J. Numer. Anal.*, **18** (1981), 398-405.

[17] Y.-X. Peng, X.-Y. Hu and L. Zhang, The symmetric and ortho-symmetric solutions of linear matrix equation $A^TXA = B$ and its optimal approximation, *Numer. Math. J. Chinese Univ.*, **25** (2003), 372-377 (in Chinese).

[18] Z.-Y. Peng, X.-Y. Hu and L. Zhang, One kind of inverse problems for symmetric and skew anti-symmetric matrices, *Numer. Math. J. Chinese Univ.*, **25** (2002), 144-152 (in Chinese).

[19] Z.-Y. Peng, X.-Y. Hu and L. Zhang, The nearest bisymmetric solutions of linear matrix equations, *J. Comput. Math.*, **22** (2004), 873-880.

[20] G.W. Stewart and J.-G. Sun, Matrix Perturbation Theory, Academic Press, New York, 1990.

[21] D.-X. Xie, Least-squares solutions of $X^TAX = B$ over positive semidefinite matrices, *J. Comput. Math.*, **21** (2003), 167-174.

[22] D.-X. Xie, Y.-P. Sheng and X.-Y. Hu, The least-squares solutions of inconsistent matrix equation over symmetric and antipersymmetric matrices, *Appl. Math. Letters*, **16** (2003), 589-598.

[23] G.-P. Xu, M.-S. Wei and D.-S. Zheng, On solutions of matrix equation $AXB+CYD = F$, *Linear Algebra Appl.*, **279** (1998), 93-109.

[24] Z. Xu, K.-Y. Zhang and Q. Lu, Fast Algorithms of matrices of Toeplitz Form, Northwest Industry Univ. Press, 1999 (in Chinese).

[25] Y.-X. Yuan, The minimum norm solutions of two classes of matrix equations *Numer. Math. J. Chinese Univ.*, **24** (2002), 127-134 (in Chinese).

[26] F.-Z. Zhou, X.-Y. Hu and L. Zhang, The inverse problems for symmetric orthogonal anti-symmetric matrices, *Acta Math. Sci. Ser. A*, **24** (2004), 543-550 (in Chinese).

[27] F.-Z. Zhou, X.-Y. Hu and L. Zhang, The solvability conditions for the inverse problems of symmetric ortho-symmetric matrices, *Appl. Math. Comput.*, **154** (2004), 153-166.