Journal of Computational Mathematics, Vol.25, No.5, 2007, 543–552.

OPTIMAL APPROXIMATE SOLUTION OF THE MATRIX EQUATION $AXB = C$ OVER SYMMETRIC MATRICES *

An-Ping Liao and Yuan Lei
(*College of Mathematics and Econometrics, Hunan University, Changsha 410082, P.R. China
liaoap@hnu.cn*)

Abstract

Let S_E denote the least-squares symmetric solution set of the matrix equation $AXB = C$, where A, B and C are given matrices of suitable size. To find the optimal approximate solution in the set S_E to a given matrix, we give a new feasible method based on the projection theorem, the generalized SVD and the canonical correction decomposition.

Mathematics subject classification: 15A24, 65F20, 65F22, 65K10.
Key words: Least-squares solution, Optimal approximate solution, Generalized singular value decomposition, Canonical correlation decomposition.

1. Introduction

Denote by $R^{m\times n}$ the set of real $m \times n$ matrices, and $SR^{n\times n}$ the set of symmetric matrices in $R^{n\times n}$. In this paper, we consider the following problem:

Problem 1.1. *Given $A \in R^{m\times n}$, $B \in R^{n\times p}$, $C \in R^{m\times p}$ and $X^* \in SR^{n\times n}$. Let*

$$S_E = \{X | X \in SR^{n\times n},\ \|AXB - C\| = \min_{Y \in SR^{n\times n}} \|AYB - C\|\}.$$

Find $\widehat{X} \in S_E$ such that

$$\|\widehat{X} - X^*\| = \min_{X \in S_E} \|X - X^*\|,$$

where $\|\cdot\|$ denotes the Frobenius norm.

In other word, S_E is the least-squares symmetric solution set of the matrix equation

$$AXB = C, \tag{1.1}$$

and $\widehat{X}$ is the optimal approximate least-squares symmetric solution of the matrix equation (1.1) to the given matrix X^*.

The consistency conditions of the matrix equation (1.1) with the symmetric solution were given by Chu [1] (see also Dai [3]), and the symmetric solutions can also be obtained by using the generalized singular value decomposition (GSVD) when the matrix equation is consistent. For the matrix equation (1.1), Wang and Chang [17] gave the least-squares symmetric solution by using GSVD; Liao and Bai [12] and Deng [5] considered the least-squares solution over the symmetric positive semi-definite matrices and positive semi-definite matrices, respectively; and Yuan [19] also gave the minimum-norm least-squares symmetric solution for the consistent matrix equation (1.1) by using the canonical correlation decomposition (CCD).

* Received December 15, 2005; final revised September 1, 2006; accepted October 1, 2006.

The problem of finding a nearest matrix in the least-squares symmetric solution set of a matrix equation to a given matrix in the sense of the Frobenius norm, that is, Problem 1.1 in this paper, is called the matrix nearness problem. The matrix nearness problem is initially proposed in the processes of test or recovery of linear systems due to incomplete dates or revising dates. A preliminary estimate X^* of the unknown matrix X can be obtained by the experimental observation values and the information of statical distribution. There are many important results on the discussions of the matrix nearness problem associated with other matrix equations, we refer the reader to [2, 4, 8, 9, 10, 15] and references therein.

In this paper, we develop an efficient method to solve Problem 1.1. Our approach is based on the projection theorem in Hilbert space, GSVD and CCD of matrix pairs. It can be essentially divided into three parts: First, we find a least-squares solution X_0 of the matrix equation (1.1) by using GSVD; then utilizing the solution X_0 and the projection theorem, we transfer Problem 1.1 to a problem of finding the optimal approximate symmetric solution of a consistent matrix equation; finally, we find the optimal approximate symmetric solution of the consistent matrix equation by using CCD.

The paper is organized as follows. After introducing some necessary notations and several useful lemmas in Section 2, we will discuss Problem 1.1 in Section 3, and give the expression of its solution. Then, in Section 4, we give the numerical algorithm to compute the solution of Problem 1.1. Numerical experiments will be carried out in Section 4.

2. Notations and Lemmas

The notation used in this paper can be summarized as follows: the set of all $n \times n$ orthogonal matrices in $R^{n\times n}$ is denoted by $OR^{n\times n}$. Denote by I the unit matrix. A^T, $tr(A)$ and $rank(A)$ respectively denote the transpose, the trace and the rank of the matrix A. For $A = (a_{ij}) \in R^{m\times n}$, $B = (b_{ij}) \in R^{m\times n}$, $A * B$ represents the Hadamard product of the matrices A and B, that is, $A * B = (a_{ij}b_{ij})_{m\times n}$. Let $\langle A, B\rangle$ represent the inner product of the matrices A and B, that is, $\langle A, B\rangle = tr(B^TA)$. Then $R^{m\times n}$ is a Hilbert inner product space, and the norm of a matrix produced by the inner product is the Frobenius norm.

We first state the concepts of the GSVD and CCD, which are essential tools for deriving the solution of Problem 1.1. See [6, 7, 11, 13, **?**, 16] for details.

Let $A \in R^{m\times n}$ and $B \in R^{n\times p}$. Then the GSVD of the matrix pair (A, B^T) is given by

$$A = U\Sigma_A M \quad \text{and} \quad B^T = V\Sigma_B M, \tag{2.1}$$

where $U \in OR^{m\times m}$ and $V \in OR^{p\times p}$; $M \in R^{n\times n}$ is a nonsingular matrix; and

$$\Sigma_A = \begin{pmatrix} I_r & 0 & 0 & 0 \\ 0 & S_A & 0 & 0 \\ 0 & 0 & 0_{(m-r-s)\times(k-r-s)} & 0 \end{pmatrix} \quad \text{and} \quad \Sigma_B = \begin{pmatrix} 0_{(p+r-k)\times r} & 0 & 0 & 0 \\ 0 & S_B & 0 & 0 \\ 0 & 0 & I_{(k-r-s)} & 0 \end{pmatrix}$$

are block matrices, with the diagonal matrices S_A and S_B being given by

$$S_A = diag(\alpha_1, \alpha_2, \cdots, \alpha_s) > 0 \quad \text{and} \quad S_B = diag(\beta_1, \beta_2, \cdots, \beta_s) > 0.$$

Here

$$k = rank(A^T, B), \quad r = k - rank(B), \quad s = rank(A) + rank(B) - k.$$

We further partition the orthogonal matrices

$$U = \begin{pmatrix} U_1 & U_2 & U_3 \end{pmatrix} \text{ and } V = \begin{pmatrix} V_1 & V_2 & V_3 \end{pmatrix} \tag{2.2}$$
$$\qquad\quad r \quad s \quad m-r-s \qquad\qquad\qquad p+r-k \quad s \quad k-r-s$$

compatibly with the block row partitioning of Σ_A and Σ_B, respectively.

The CCD of the matrix pair (A^T, B) is given by

$$A^T = Q(\Xi_A, 0)E_A^{-1} \quad \text{and} \quad B = Q(\Xi_B, 0)E_B^{-1}, \tag{2.3}$$

where $Q \in OR^{n\times n}$; $E_A \in R^{m\times m}$ and $E_B \in R^{p\times p}$ are nonsingular matrices; and

$$\Xi_A = \begin{pmatrix} I_{r'} & 0 & 0 \\ 0 & C_A & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & D_A & 0 \\ 0 & 0 & I_{t'} \end{pmatrix} \quad \text{and} \quad \Xi_B = \begin{pmatrix} I_{r'} & 0 & 0 \\ 0 & I_{s'} & 0 \\ 0 & 0 & I_{h-r'-s'} \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

are block matrices, with the diagonal matrices C_A and D_A being given by

$$C_A = diag(\mu_1, \mu_2, \cdots, \mu_{s'}) > 0 \quad \text{and} \quad D_A = diag(\lambda_1, \lambda_2, \cdots, \lambda_{s'}) > 0.$$

Here,

$$g = rank(A) = r' + s' + t', \quad h = rank(B),$$
$$r' = rank(A) + rank(B) - rank(A^T, B), \quad s' = rank(AB) - r'.$$

We further partition the nonsingular matrices

$$E_A = \begin{pmatrix} A_1 & A_2 & A_3 & A_4 \end{pmatrix} \text{ and } E_B = \begin{pmatrix} B_1 & B_2 & B_3 & B_4 \end{pmatrix} \tag{2.4}$$
$$\qquad r' \quad s' \quad t' \quad m-g \qquad\qquad\qquad r' \quad s' \quad h-r'-s' \quad p-h$$

compatibly with the block column partitioning of $(\Xi_A, 0)$ and $(\Xi_B, 0)$, respectively.

The following lemmas are important for deriving an analytical formula of the solution of Problem 1.1.

Lemma 2.1. (The Projection Theorem [18]) *Let X be an inner product space, M be a subspace of X, and $M^\perp$ be the orthogonal complement subspace of M. For a given $x \in X$, if there exists an $m_0 \in M$ such that $\|x - m_0\| \le \|x - m\|$ holds for any $m \in M$, then m_0 is unique and $m_0 \in M$ is the unique minimization vector in M if and only if $(x - m_0)\perp M$, i.e., $(x - m_0) \in M^\perp$.*

Lemma 2.2. *Given matrices $F = (f_{ij}) \in R^{t'\times s'}$, $E = (e_{ij}) \in R^{s'\times t'}$ and $G = (g_{ij}) \in R^{s'\times t'}$. Let*

$$C_A = diag(\mu_1, \mu_2, \cdots, \mu_{s'}) \quad \text{and} \quad D_A = diag(\lambda_1, \lambda_2, \cdots, \lambda_{s'})$$

be given diagonal matrices of positive diagonal entries, satisfying $\mu_i^2 + \lambda_i^2 = 1$ $(i = 1, \ldots, s')$, and define

$$\varphi(Y) = \|Y - F\|^2 + \|C_A^{-1}(E - D_A Y^T) - G\|^2.$$

Then there exists a unique matrix $\widetilde{Y} = (\tilde{y}_{ij}) \in R^{t' \times s'}$ *such that*

$$\varphi(\widetilde{Y}) = \min_{Y \in R^{t' \times s'}} \varphi(Y).$$

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

$$\widetilde{Y} = [FC_A^2 + E^T D_A - G^T D_A C_A]. \tag{2.5}$$

Proof. For the given matrices, we have

$$\varphi(\widetilde{Y}) = \sum_{i,j} [(\tilde{y}_{ij} - f_{ij})^2 + (\frac{1}{\mu_j} e_{ji} - \frac{\lambda_j}{\mu_j} \tilde{y}_{ij} - g_{ji})^2].$$

Because $\varphi(\widetilde{Y})$ is a convex, continuous and differentiable function with respect to the $t's'$ variables $\tilde{y}_{ij}$ (i=$1, \ldots, t'$, j=$1, \ldots, s'$), we easily know that $\varphi(\widetilde{Y})$=min if and only if $\frac{\partial \varphi(\widetilde{Y})}{\partial \tilde{y_{ij}}}$=0. It then follows from direct computations that

$$\tilde{y}_{ij} = f_{ij} \mu_j^2 + e_{ji} \lambda_j - g_{ij} \lambda_j \mu_j, \qquad i = 1, \ldots, t', \quad j = 1, \ldots, s'. \tag{2.6}$$

By rewriting (2.6) in matrix form, we immediately obtain (2.5). □

Lemma 2.3. *Given matrices* $F = (f_{ij}) \in SR^{s' \times s'}$, $E = (e_{ij}) \in R^{s' \times s'}$ *and* $G = (g_{ij}) \in R^{s' \times s'}$. *Let*

$$C_A = diag(\mu_1, \mu_2, \cdots, \mu_{s'}) \quad and \quad D_A = diag(\lambda_1, \lambda_2, \cdots, \lambda_{s'})$$

be given diagonal matrices of positive diagonal entries, satisfying $\mu_i^2 + \lambda_i^2 = 1$ $(i = 1, \ldots, s')$, *and define*

$$\psi(Y) = \|Y - F\|^2 + 2\|E^T D_A^{-1} - Y C_A D_A^{-1} - G\|^2.$$

Then there exists a unique matrix $\widehat{Y} = (\hat{y}_{ij}) \in SR^{s' \times s'}$ *such that*

$$\psi(\widehat{Y}) = \min_{Y \in SR^{s' \times s'}} \psi(Y).$$

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

$$\widetilde{Y} = K * [D_A F D_A + D_A^2 E^T C_A + C_A E D_A^2 - D_A^2 G D_A C_A - D_A C_A G^T D_A^2], \tag{2.7}$$

where $K = (k_{ij}) \in R^{s' \times s'}$ *is defined by*

$$k_{ij} = \frac{1}{\lambda_j^2 + \lambda_i^2 \mu_j^2}, \qquad i, j = 1, \ldots, s'.$$

Proof. For the given matrices, we have

$$\begin{aligned} \psi(\widehat{Y}) \;=\; & \sum_i \left[(\hat{y}_{ii} - f_{ii})^2 + 2(\tfrac{1}{\lambda_i} e_{ii} - \tfrac{\mu_i}{\lambda_i} \hat{y}_{ii} - g_{ii})^2 \right] \\ & + \sum_{i<j} \left[2(\hat{y}_{ij} - f_{ij})^2 + 2(\tfrac{1}{\lambda_j} e_{ji} - \tfrac{\mu_j}{\lambda_j} \hat{y}_{ij} - g_{ij})^2 + 2(\tfrac{1}{\lambda_i} e_{ij} - \tfrac{\mu_i}{\lambda_i} \hat{y}_{ij} - g_{ji})^2 \right]. \end{aligned}$$

Since $\psi(\widehat{Y})$ is a convex, continuous and differentiable function with respect to the $\frac{s'(s'+1)}{2}$ variables $\hat{y}_{ij}$ $(i, j = 1, \ldots, s')$, we easily know that $\psi(\widehat{Y}) = \min$ if and only if $\frac{\partial \psi(\widehat{Y})}{\partial \hat{y_{ij}}} = 0$. It then follows from direct computations that

$$\hat{y}_{ij} = \frac{\lambda_i f_{ij} \lambda_j + \lambda_i^2 e_{ji} \mu_j + \mu_i e_{ij} \lambda_j^2 - \lambda_i^2 g_{ij} \lambda_j \mu_j - \lambda_i \mu_i g_{ji} \lambda_j^2}{\lambda_j^2 + \lambda_i^2 \mu_j^2}, \qquad i, j = 1, \ldots, s'. \tag{2.8}$$

By rewriting (2.8) in matrix form, we immediately obtain (2.7). □

3. The Solution of Problem 1.1

In this section, we derive an analytical expression for the solution of Problem 1.1. To this end, we first transform the least-squares problem with respect to the matrix equation (1.1) to a consistent matrix equation. This technique is precisely described in the following theorem.

Theorem 3.1. *Suppose that the matrices A, B and C are given in Problem* 1.1. *Let X_0 be one of the least-squares solutions of the matrix equation* (1.1)*, and define*

$$C_0 = AX_0B. \tag{3.1}$$

Then the matrix equation

$$AXB = C_0 \tag{3.2}$$

is consistent over symmetric matrices, and its symmetric solution set is the same as the least-squares symmetric solution set S_E of the matrix equation (1.1).

Proof. Let

$$S = \{Y | Y = AXB, X \in SR^{n\times n}\}.$$

Then S is obviously a linear subspace of $R^{m\times p}$. Because X_0 is a least-squares symmetric solution of the matrix equation (1.1), from (3.1) we see that $C_0 \in S$ and

$$\begin{aligned} \|C_0 - C\| &= \|AX_0B - C\| \\ &= \min_{X\in SR^{n\times n}} \|AXB - C\| \\ &= \min_{Y\in S} \|Y - C\|. \end{aligned}$$

Now, by Lemma 2.1 we have

$$(C_0 - C) \perp S, \quad \text{or} \quad (C_0 - C) \in S^{\perp}.$$

For $X \in SR^{n\times n}$, we know that $(AXB - C_0) \in S$. It then follows that

$$\begin{aligned} \|AXB - C\|_F^2 &= \|(AXB - C_0) + (C_0 - C)\|_F^2 \\ &= \|AXB - C_0\|_F^2 + \|C_0 - C\|_F^2. \end{aligned}$$

Hence, the conclusion of this theorem holds true. $\square$

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^* is just the solution of Problem 1.1. Therefore, solving Problem 1.1 essentially reduces to find C_0, or a least-squares solution X_0 of the matrix equation (1.1). Based on the GSVD (2.1) of the matrix pair (A, B^T), the following theorem gives such a matrix C_0.

Theorem 3.2. *Suppose that the matrices A, B and C are given in Problem* 1.1. *Denote by*

$$U^TCV = (C_{ij})_{3\times 3}, \quad \textit{with} \quad C_{ij} = U_i^TCV_j, \quad i, j = 1, 2, 3, \tag{3.3}$$

where the matrices U_i and V_i (i=1, 2, 3) are given by (2.2). *Then the following matrix C_0 corresponds a least-squares solution X_0 of the matrix equation* (1.1) *and satisfies* (3.1)*:*

$$C_0 = U \begin{pmatrix} 0 & C_{12} & C_{13} \\ 0 & S_A\bar{X}_{22}S_B & C_{23} \\ 0 & 0 & 0 \end{pmatrix} V^T, \tag{3.4}$$

where

$$\bar{X}_{22} = \Phi * (S_A C_{22} S_B + S_B C_{22}^T S_A) \quad with \quad \Phi = (\frac{1}{\alpha_i^2\beta_j^2 + \beta_i^2\alpha_j^2}) \in SR^{s\times s}, \quad i,j = 1,\dots,s. \tag{3.5}$$

Proof. From [17] we know that the least-squares solutions of the matrix equation (1.1) can be given by using the GSVD of the matrix pair (A, B^T), and are of the following form

$$X = M^{-1}\begin{pmatrix} X_{11} & C_{12}S_B^{-1} & C_{13} & X_{14} \\ S_B^{-1}C_{12}^T & \bar{X}_{22} & S_A^{-1}C_{23} & X_{24} \\ C_{13}^T & C_{23}^T S_A^{-1} & X_{33} & X_{34} \\ X_{14}^T & X_{24}^T & X_{34}^T & X_{44} \end{pmatrix} M^{-T}, \tag{3.6}$$

where the block matrix $\bar{X}_{22}$ is defined by (3.5); X_{ii} $(i = 1, 3, 4)$ are arbitrary symmetric matrix blocks; X_{i4} $(i = 1, 2, 3)$ are arbitrary matrix blocks. By substituting (2.1) and (3.6) into (3.1), after concrete manipulations we can obtain (3.4). □

Evidently, (3.4) shows that the matrix C_0 given in Theorem 3.2 is unique, and only dependent on the matrices A, B, and C, but independent of the least-squares solution X_0 of the matrix equation (1.1). Therefore, we can conclude that

$$\|C_0 - C\| = \min_{X\in SR^{n\times n}} \|AXB - C\|.$$

Based on Theorems 3.1 and 3.2, we can obtain the analytical expression of the solutions of Problem 1.1 by using the CCD (2.3) of the matrix pair (A^T, B). To state the results, we denote

$$E_A^T C_0 E_B = (E_{ij})_{4\times 4}, \quad \text{with} \quad E_{ij} = A_i^T C_0 B_j \quad i,j = 1,2,3,4, \tag{3.7}$$

where C_0 is defined by (3.4), and the matrices A_i and B_i $(i$=$1, 2, 3, 4)$ are given by (2.4).

Theorem 3.3. *Suppose that the matrices A, B, C and X^* are given in Problem 1.1. Partition the matrix $Q^T X^* Q$ compatibly to the block row partitioning of Ξ_A and Ξ_B into*

$$Q^T X^* Q = (X_{ij}^*)_{6\times 6}, \quad with \quad X_{ij}^{*T} = X_{ji}^*, \quad i,j = 1,2,\dots,6, \tag{3.8}$$

where the matrix Q is defined in (2.3). *Then the unique solution $\widehat{X}$ of Problem* 1.1 *can be expressed as*

$$\widehat{X} = Q\begin{pmatrix} E_{11} & E_{12} & E_{13} & X_{14}^* & \hat{Y}_{15} & E_{31}^T \\ E_{12}^T & \hat{Y}_{22} & C_A^{-1}(E_{23} - D_A\widehat{Y}_{35}^T) & X_{24}^* & (E_{22}^T - \widehat{Y}_{22}C_A)D_A^{-1} & E_{32}^T \\ E_{13}^T & (E_{23}^T - \widehat{Y}_{35}D_A)C_A^{-1} & X_{33}^* & X_{34}^* & \widehat{Y}_{35} & E_{33}^T \\ X_{14}^{*T} & X_{24}^{*T} & X_{34}^{*T} & X_{44}^* & X_{45}^* & X_{46}^* \\ \widehat{Y}_{15}^T & D_A^{-1}(E_{22} - C_A\widehat{Y}_{22}^T) & \widehat{Y}_{35}^T & X_{45}^{*T} & X_{55}^* & X_{56}^* \\ E_{31} & E_{32} & E_{33} & X_{46}^{*T} & X_{56}^{*T} & X_{66}^* \end{pmatrix} Q^T, \tag{3.9}$$

where

$$\begin{cases} \widehat{Y}_{15} = (E_{21}^T - E_{12}C_A)D_A^{-1}, \\ \widehat{Y}_{35} = X_{35}^* C_A^2 + E_{23}^T D_A - X_{23}^* C_A D_A, \\ \widehat{Y}_{22} = K * [D_A X_{22}^* D_A + D_A^2 E_{22}^T C_A + C_A E_{22} D_A^2 - D_A^2 X_{25}^* D_A C_A - D_A C_A X_{25}^{*T} D_A^2], \end{cases} \tag{3.10}$$

with

$$K = (k_{ij}) \in R^{s'\times s'}, \quad k_{ij} = \frac{1}{\lambda_j^2 + \lambda_i^2\mu_j^2} \qquad i,j = 1,\dots,s'.$$

Proof. From Theorems 3.1 and 3.2 we know that the least-squares symmetric solution set of the matrix equation (1.1) is the same as the symmetric solution set of the consistent matrix equation (3.2), with the matrix E_0 being given by (3.4). From [19] we see that the solutions of the consistent matrix equation (3.2) can be expressed as

$$X = Q\begin{pmatrix} E_{11} & E_{12} & E_{13} & Y_{14} & \widehat{Y}_{15} & E_{31}^T \\ E_{12}^T & Y_{22} & \widetilde{X}_{23} & Y_{24} & \widetilde{X}_{25} & E_{32}^T \\ E_{13}^T & \widetilde{X}_{23}^T & Y_{33} & Y_{34} & Y_{35} & E_{33}^T \\ Y_{14}^T & Y_{24}^T & Y_{34}^T & Y_{44} & Y_{45} & Y_{46} \\ \widehat{Y}_{15}^T & \widetilde{X}_{25}^T & Y_{35}^T & Y_{45}^T & Y_{55} & Y_{56} \\ E_{31} & E_{32} & E_{33} & Y_{46}^T & Y_{56}^T & Y_{66} \end{pmatrix} Q^T, \tag{3.11}$$

where

$$\begin{cases} \widetilde{X}_{23} & = & C_A^{-1}(E_{23} - D_A Y_{35}^T), \\ \widetilde{X}_{25} & = & E_{22}^T D_A^{-1} - Y_{22} C_A D_A^{-1}; \end{cases} \tag{3.12}$$

the matrix block $\widehat{Y}_{15}$ is defined by (3.10); Y_{ii} $(i = 2, 3, 6)$ are arbitrary symmetric matrix blocks; and the other matrix blocks Y_{ij} $(i, j = 1, 2, 6)$ are arbitrary.

It follows from (3.8) and (3.11) that

$$\|X - X^*\|^2 = \|Q^T XQ - Q^T X^* Q\|^2$$
$$= \left\| \begin{pmatrix} E_{11} - X_{11}^* & E_{12} - X_{12}^* & E_{13} - X_{13}^* & Y_{14} - X_{14}^* & \widehat{Y}_{15} - X_{15}^* & E_{31}^T - X_{16}^* \\ E_{12}^T - X_{12}^{*T} & Y_{22} - X_{22}^* & \tilde{X}_{23} - X_{23}^* & Y_{24} - X_{24}^* & \tilde{X}_{25} - X_{25}^* & E_{32}^T - X_{26}^* \\ E_{13}^T - X_{13}^{*T} & \tilde{X}_{23}^T - X_{23}^{*T} & Y_{33} - X_{33}^* & Y_{34} - X_{34}^* & Y_{35} - X_{35}^* & E_{33}^T - X_{36}^* \\ Y_{14}^T - X_{14}^{*T} & Y_{24}^T - X_{24}^{*T} & Y_{34}^T - X_{34}^{*T} & Y_{44} - X_{44}^* & Y_{45} - X_{45}^* & Y_{46} - X_{46}^* \\ \widehat{Y}_{15}^T - X_{15}^{*T} & \tilde{X}_{25}^T - X_{25}^{*T} & Y_{35}^T - X_{35}^{*T} & Y_{45}^T - X_{45}^{*T} & Y_{55} - X_{55}^* & Y_{56} - X_{56}^* \\ E_{31} - X_{16}^{*T} & E_{32} - X_{26}^{*T} & E_{33} - X_{36}^{*T} & Y_{46}^T - X_{46}^{*T} & Y_{56}^T - X_{56}^{*T} & Y_{66} - X_{66}^* \end{pmatrix} \right\|^2.$$

Hence,

$$\|X - X^*\| = \min, \qquad \forall X \in S_E$$

(see the definition of the matrix set S_E in Problem 1.1) if and only if

$$\begin{cases} Y_{14} = X_{14}^*, & Y_{24} = X_{24}^*, & Y_{34} = X_{34}^*, & Y_{45} = X_{45}^*, & Y_{46} = X_{46}^*, \\ Y_{56} = X_{56}^*, & Y_{33} = X_{33}^*, & Y_{44} = X_{44}^*, & Y_{55} = X_{55}^*, & Y_{66} = X_{66}^*, \end{cases} \tag{3.13}$$

$$\|Y_{35} - X_{35}^*\|^2 + \|C_A^{-1}(E_{23} - D_A Y_{35}^T) - X_{23}^*\|^2 = \min, \quad \forall\, Y_{35} \in R^{(h-r'-s')\times s'} \tag{3.14}$$

and

$$\|Y_{22} - X_{22}^*\|^2 + 2\|(E_{22}^T - Y_{22}C_A)D_A^{-1} - X_{25}^*\|^2 = \min, \quad \forall\, Y_{22} \in SR^{s'\times s'}. \tag{3.15}$$

By making use of Lemmas 2.2 and 2.3 we know that the solutions of (3.14) and (3.15) are of the form

$$\begin{cases} \widehat{Y}_{35} = X_{35}^* C_A^2 + E_{23}^T D_A - X_{23}^* C_A D_A, \\ \widehat{Y}_{22} = K * [D_A X_{22}^* D_A + D_A^2 E_{22}^T C_A + C_A E_{22} D_A^2 - D_A^2 X_{25}^* D_A C_A - D_A C_A X_{25}^{*T} D_A^2]. \end{cases}$$

Substituting these $\widehat{Y}_{35}$, $\widehat{Y}_{22}$ and (3.13) into (3.11) yields (3.9). $\square$

Remark 3.1. In Problem 1.1, if the matrix X^* is not symmetric, then from

$$\|X - X^*\|^2 = \|X - \frac{1}{2}(X^* + X^{*T})\|^2 + \|\frac{1}{2}(X^* - X^{*T})\|^2, \quad \forall X \in S_E,$$

we know that the minimization problem

$$\|X - X^*\| = \min \quad \forall X \in S_E$$

is equivalent to the following minimization problem

$$\|X - \frac{1}{2}(X^* + X^{*T})\| = \min \quad \forall X \in S_E.$$

Therefore, without loss of generality, in following discussion we suppose that the matrix X^* is symmetric in Problem 1.1.

4. A Numerical Algorithm for Solving Problem 1.1

Based on Theorem 3.3, we can establish an algorithm for finding the solution of Problem 1.1.

Algorithm for solving Problem 1.1

1. Input matrices A, B, C and X^*.
2. Make the GSVD of the matrix pair (A, B^T) according to (2.1).
3. Partition the matrix $U^T CV = (C_{ij})_{3\times 3}$ according to (3.3).
4. Compute $\bar{X}_{22}$ according to (3.5).
5. Compute C_0 according to (3.4).
6. Make the CCD of the matrix pair (A^T, B) according to (2.3).
7. Partition the matrices E_A and E_B according to (2.4).
8. Compute the matrix $\widehat{X}$ according to (3.9).

Example 1. Let

$$A = \begin{pmatrix} ones(5,5) & zeros(5,4) \\ zeros(4,5) & pascal(4) \end{pmatrix}, \qquad B = \begin{pmatrix} hankel(1:4) & zeros(4,5) \\ zeros(5,4) & zeros(5,5) \end{pmatrix},$$
$$C = \begin{pmatrix} toeplitz(1:4) & zeros(4,5) \\ zeros(5,4) & hilb(5) \end{pmatrix}, \quad X^* = \begin{pmatrix} eye(4) & \frac{1}{2}ones(4,5) \\ \frac{1}{2}ones(5,4) & eye(5) \end{pmatrix},$$

where $hilb(n)$ and $pascal(n)$ denote the n-th order Hilbert matrix and Pascal matrix, respectively, and $toeplitz(1:n)$ and $hankel(1:n)$ denote the n-th order Toeplitz matrix and Hankel matrix whose first row is $(1, 2, \cdots, n)$, respectively.

By using MATLAB 6.5, we obtain

$$C_0 = \begin{pmatrix} 2.0000 & 1.6000 & 1.6000 & 2.0000 & 0 & 0 & 0 & 0 & 0 \\ 2.0000 & 1.6000 & 1.6000 & 2.0000 & 0 & 0 & 0 & 0 & 0 \\ 2.0000 & 1.6000 & 1.6000 & 2.0000 & 0 & 0 & 0 & 0 & 0 \\ 2.0000 & 1.6000 & 1.6000 & 2.0000 & 0 & 0 & 0 & 0 & 0 \\ 2.0000 & 1.6000 & 1.6000 & 2.0000 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

and

$$\widehat{X} = \begin{pmatrix} 0.8258 & -0.2692 & -0.2480 & -0.2214 & 0.4129 & 0 & 0 & 0 & 0 \\ -0.2692 & 0.6358 & -0.3430 & -0.3164 & 0.3179 & 0 & 0 & 0 & 0 \\ -0.2480 & -0.3430 & 0.6783 & -0.2952 & 0.3391 & 0 & 0 & 0 & 0 \\ -0.2214 & -0.3164 & -0.2952 & 0.7314 & 0.3657 & 0 & 0 & 0 & 0 \\ 0.4129 & 0.3179 & 0.3391 & 0.3657 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}.$$

By concrete computations, we have

$$A\widehat{X}B = C_0, \quad \min_{X \in S_E} \|X - X^*\| = \|[\widehat{X} - X^*]\| = 4.4141.$$

In addition, we also have

$$\min_{X \in SR^{n\times n}} \|AXB - C\| = \|C_0 - C\| = 5.7358$$

and

$$\langle C_0, C_0 - C\rangle = tr(C_0^T(C_0 - C)) = -4.7073 \times 10^{-14}.$$

This demonstrates that the above-described algorithm is feasible for solving Problem 1.1.

Acknowledgments. The work of this author was supported in part by Natural Science Foundation of Hunan Province (No. 03JJY6028).

References

[1] K.E. Chu, Symmetric solution of linear matrix equations by matrix decomposition, *Linear Algebra Appl.*, **119** (1989), 35-50.

[2] M. Chu and G. Golub, Inverse Eigenvalue Problems, Theory Algorithms and Applications, Oxford University Press, 2005.

[3] H. Dai, On the symmetric solutions of linear matrix equations, *Linear Algebra Appl.*, **131** (1990), 1-7.

[4] Y.B. Deng and X. Hu, On solutions of matrix equation $Ax^T + Byb^T = C$, *J. Comput. Math.*, **23** (2005), 17-26.

[5] Y.B. Deng, X.Y. Hu and L. Zhang, Least squares solution of $BXA^T = T$ over symmetric, skew-symmetric, and positive semidefinite X^*, *SIAM J. Matrix Anal. Appl.*, **25** (2003), 486-494.

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

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

[8] N.J. Higham, Computing a nearest symmetric positive semidefinite matrix, *Linear Algebra Appl.*, **103** (1988), 103-118.

[9] X. Jin and Y. Wei, Numerical Linear Algebra and its Applications, Science Press, Beijing/New York, 2004.

[10] M. Konstaintinov, D. Gu, V. Mehrmann and P. Petkov, Perturbation Theory for Matrix Equations, Elsevier, 2003.

[11] W. Li and J.-X. Chen, The eigenvalue perturbation bound for arbitrary matrices, *J. Comput. Math.*, **24** (2006), 141-148.

[12] A.P. Liao and Z.Z. Bai, Least-squares solution of $AXB = D$ over symmetric positive semidefinite matrices X, *J. Comput. Math.*, **21** (2003), 175-182.

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

[14] M.-X. Pang, New estimates for singular values of a matrix, *J. Comput. Math.*, **23** (2005), 199-204.

[15] R. Penrose, On best approximate solutions of linear matrix equations, *Proc. Cambridge Philos. Soc.*, **52** (1956), 17-19.

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

[17] J.S. Wang and X.W. Chang, Symmetric solutions of a linear matrix equation, *Nanjing Univ. J. Math. Biquarterly*, **7** (1990), 125-129 (in Chinese).

[18] R.-C. Wang, Functional Analysis and Optimization Theory, Beijing University of Aeronautics and Astronautics Press, Beijing, 2003 (in Chinese).

[19] Y.-X. Yuan, The optimal solution of linear matrix equation by using matrix decompositions, *Math. Numer. Sinica*, **24** (2002), 165-176 (in Chinese).