Journal of Computational Mathematics, Vol.21, No.2, 2003, 175–182.

LEAST-SQUARES SOLUTION OF $AXB = D$ OVER SYMMETRIC POSITIVE SEMIDEFINITE MATRICES X [*1)]

Anping Liao
(*Department of Mathematics, Hunan Normal University, Changsha 410081*)
(*Department of Mathematics and Information Science Changsha University, Changsha 410003, China*)

Zhongzhi Bai[†]
(*State Key Laboratory of Scientific/Engineering Computing, Institute of Computational Mathematics and Scientific/Engineering Computing, Academy of Mathematics and System Sciences, Chinese Academy of Sciences, P.O. Box 2719, Beijing 100080, China*)

Abstract

Least-squares solution of $AXB = D$ with respect to symmetric positive semidefinite matrix X is considered. By making use of the generalized singular value decomposition, we derive general analytic formulas, and present necessary and sufficient conditions for guaranteeing the existence of the solution. By applying MATLAB 5.2, we give some numerical examples to show the feasibility and accuracy of this construction technique in the finite precision arithmetic.

Key words: Least-squares solution, Matrix equation, Symmetric positive semidefinite matrix, Generalized singular value decomposition.

1. Introduction

Denote by $R^{n\times m}$ the set of all real $n\times m$ matrices, I_k the identity matrix in $R^{k\times k}$, $\mathrm{O}R^{n\times n}$ the set of all orthogonal matrices in $R^{n\times n}$, $\mathrm{S}R^{n\times n}$ the set of all symmetric matrices in $R^{n\times n}$, and $\mathrm{S}R_0^{n\times n}(\mathrm{S}R_+^{n\times n})$ the set of all symmetric positive semidefinite (definite) matrices in $R^{n\times n}$. The notations $A\geq O$ ($A>O$) represents that the matrix A is a symmetric positive semidefinite (definite) matrix, A^+ represents the Moor-Penrose generalized inverse of a matrix A, and $\|\cdot\|$ denotes the Frobenius norm. We use both O and 0 to denote the zero matrix.

The purpose of this paper is to study the least-squares problem of the matrix equation $AXB = D$ with respect to $X\in \mathrm{S}R_0^{n\times n}$, i.e.,

Problem I. For given matrices $A\in R^{m\times n}$, $B\in R^{n\times p}$, $D\in R^{m\times p}$, find a matrix $\hat{X}\in \mathrm{S}R_0^{n\times n}$ such that

$$\|A\hat{X}B-D\| = \min_{X\in \mathrm{S}R_0^{n\times n}} \|AXB-D\|.$$

Allwright[1] first investigated a special case of Problem I, i.e.,

Problem A. For given matrices $B, D\in R^{n\times m}$, find a matrix $\hat{X}\in \mathrm{S}R_0^{n\times n}$ such that

$$\|\hat{X}B-D\| = \min_{X\in \mathrm{S}R_0^{n\times n}} \|XB-D\|.$$

The solution of Problem A can be used as estimates of the inverse Hessian of a nonlinear differentiable function $f: R^n\rightarrow R^1$, which is to be minimized with respect to a parameter

* Received August 11, 2000.
1) Subsidized by The Special Funds For Major State Basic Research Project G1999032803.
† Email address: bzz@lsec.cc.ac.cn.

vector $x \in R^n$ by a quasi-Newton-type algorithm. Allwright, Woodgate[2,3] and Liao[4] gave some necessary and sufficient conditions for the existence of solution of Problem A as well as explicit formulas of the solution for some special cases.

Dai and Lancaster[5] studied in detail another special case of Problem I, i.e.,

Problem B. For given matrices $A \in R^{n\times m}$, $D \in R^{m\times m}$, find a matrix $\hat{X} \in \mathrm{S}R_0^{n\times n}$ such that

$$\|A^T\hat{X}A - D\| = \min_{X\in \mathrm{S}R_0^{n\times n}} \|A^TXA - D\|.$$

An inverse problem[6,7] arising in structural modification of the dynamic behaviour of a structure calls for solution of Problem B. Dai and Lancaster[5] successfully solved Problem B by using singular value decomposition(SVD).

Obviously, Problem I is a nontrivial generalization of Problems A and B, and the approachs adopted for solving Problems A and B in [1-5] are not suitable for Problem I.

In this paper, by applying the generalized singular value decomposition (GSVD) we will present necessary and sufficient conditions for the existence of the solution of Problem I, and give analytic expression of it, too.

2. Solutions of Problem I

We first study the solution of Problem I when matrices $A, B, D \in R^{n\times n}$ and $\mathrm{rank}(A) = rank(B) = n$.

Lemma 2.1[1]**.** *If rank$(B) = n$, then Problem A has a unique solution.*

Theorem 2.1. *If $A, B, D \in R^{n\times n}$ and rank(A) = rank$(B) = n$, then Problem I has a unique solution.*

Proof. Because $A, B \in R^{n\times n}$ and $\mathrm{rank}(A) = \mathrm{rank}(B) = n$, we easily know that $\mathrm{rank}(\bar{B}) = n$ and

$$\|AXB - D\| = \|(AXA^T)(A^{-T}B) - D\| = \|\bar{X}\bar{B} - D\|, \tag{2.1}$$

where $\bar{B} = A^{-T}B$, and $\bar{X} = AXA^T$. Now, it is obvious that $\bar{X} \geq 0$ if and only if $X \geq 0$. Therefore, by Lemma 2.1 and (2.1) Problem I has a unique solution.

To study the solvability of Problem I in general case, we decompose the given matrix pair $[A^T, B]$ by GSVD[8] as follows:

$$A^T = M\Sigma_A U^T, \qquad B = M\Sigma_B V^T, \tag{2.2}$$

where M is an $n \times n$ nonsingular matrix, and

$$\Sigma_A = \begin{array}{cc} \begin{pmatrix} I_A & O & O \\ O & S_A & O \\ O & O & O_A \\ O & O & O \end{pmatrix} & \begin{array}{c} r \\ s \\ k-r-s \\ n-k \end{array} \\ \begin{array}{ccc} r & s & m-r-s \end{array} & \end{array}, \qquad \Sigma_B = \begin{array}{cc} \begin{pmatrix} O_B & O & O \\ O & S_B & O \\ O & O & I_B \\ O & O & O \end{pmatrix} & \begin{array}{c} r \\ s \\ k-r-s \\ n-k \end{array} \\ \begin{array}{ccc} p+k-r & s & k-r-s \end{array} & \end{array},$$

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

I_A and I_B are identity matrices, O_A and O_B are zero matrices, and

$$S_A = \mathrm{diag}(\alpha_1, \alpha_2, \ldots, \alpha_s), \quad S_B = \mathrm{diag}(\beta_1, \beta_2, \ldots, \beta_s)$$

with

$$1 > \alpha_1 \geq \alpha_2 \geq \ldots \geq \alpha_s > 0, \quad 0 < \beta_1 \leq \beta_2 \leq \ldots \leq \beta_s < 1, \quad \alpha_i^2 + \beta_i^2 = 1 (i = 1, 2, \ldots, s),$$

$$U = (\underset{r}{U_1}, \ \underset{s}{U_2}, \ \underset{m-r-s}{U_3}) \in \mathrm{O}R^{m\times m}, \quad V = (\underset{p+r-k}{V_1}, \ \underset{s}{V_2}, \ \underset{k-r-s}{V_3}) \in \mathrm{O}R^{p\times p}.$$

Theorem 2.2. *Let $A \in R^{m\times n}$, $B \in R^{n\times p}$, and $D \in R^{m\times p}$. Assume that the matrix pair $[A^T, B]$ has GSVD (2.2). Denote by D_{ij} the matrix $U_i^T D V_j (i, j = 1, 2, 3)$. Then Problem I exists a solution if and only if*

$$rank(\hat{X}_{22}) = rank(S_B^{-1} D_{12}^T, \hat{X}_{22}, S_A^{-1} D_{23}), \tag{2.3}$$

where $\hat{X}_{22}$ is a unique minimizer of $\|S_A X_{22} S_B - D_{22}\|$ with respect to $X_{22} \in SR_0^{s\times s}$.

Moreover, if Problem I exists a solution X, then this solution has the following expression:

$$X = M^{-T} \begin{pmatrix} Y & YZ \\ (YZ)^T & Z^T Y Z + G_3 \end{pmatrix} M^{-1}, \tag{2.4}$$

where $\forall Z \in R^{k\times(n-k)}$ and

$$Y = \begin{pmatrix} X_{11} & D_{12} S_B^{-1} & D_{13} \\ S_B^{-1} D_{12}^T & \hat{X}_{22} & S_A^{-1} D_{23} \\ D_{13}^T & D_{23}^T S_A^{-1} & X_{33} \end{pmatrix}, \tag{2.5}$$

$$X_{11} = D_{12} S_B^{-1} \hat{X}_{22}^+ S_B^{-1} D_{12}^T + G_1, \tag{2.6}$$

$$X_{33} = D_{23}^T S_A^{-1} \hat{X}_{22}^+ S_A^{-1} D_{23} + (D_{13} - D_{12} S_B^{-1} \hat{X}_{22}^+ S_A^{-1} D_{23})^T G_1^+ (D_{13} - D_{12} S_B^{-1} \hat{X}_{22}^+ S_A^{-1} D_{23}) + G_2, \tag{2.7}$$

$$\forall G_1 \in SR_0^{r\times r}, \quad \forall G_2 \in SR_0^{(k-r-s)\times(k-r-s)}, \quad \forall G_3 \in SR_0^{(n-k)\times(n-k)},$$

and G_1 satisfies the following condition:

$$rank(G_1) = rank(G_1, D_{13} - D_{12} S_B^{-1} \hat{X}_{22}^+ S_A^{-1} D_{23}).$$

The following lemmas are necessary for proving Theorem 2.2.

Lemma 2.2. *Suppose that $A \in R^{m\times n}$, $B \in R^{n\times p}$ and $D \in R^{m\times p}$, and there exists a minimizer $\hat{X}$ to Problem I. Then*

$$\|A\hat{X}B - D\| = min_{X\in SR_0^{n\times n}} \|AXB - D\| = inf_{X\in SR_+^{n\times n}} \|AXB - D\|. \tag{2.8}$$

Proof. It is obvious that

$$\|A\hat{X}B - D\| = min_{X\in SR_0^{n\times n}} \|AXB - D\| \le inf_{X\in SR_+^{n\times n}} \|AXB - D\|.$$

On the other hand, if $AB = 0$, then $\hat{X} + \epsilon I > 0$ for any $\epsilon > 0$, and

$$\|A(\hat{X} + \epsilon I)B - D\| \le \|A\hat{X}B - D\| + \|\epsilon AB\| < \|A\hat{X}B - D\| + \epsilon;$$

if $AB \ne 0$, then by letting $\delta = (2\|AB\|)^{-1}\epsilon$ for any $\epsilon > 0$, we have $\hat{X} + \delta I > 0$, and

$$\|A(\hat{X} + \delta I)B - D\| \le \|A\hat{X}B - D\| + \|\delta AB\| < \|A\hat{X}B - D\| + \epsilon.$$

That is to say,

$$inf_{X\in SR_+^{n\times n}} \|AXB - D\| \le \|A\hat{X}B - D\| = min_{X\in SR_0^{n\times n}} \|AXB - D\|,$$

and it follows straightforwardly that (2.8) holds.

Lemma 2.3 (See [9], Theorem 1). *Let $Q = (Q_{ij})_{2\times 2}$ be a 2×2 real block symmetric matrix, with Q_{11} and Q_{22} square submatrices. Then Q is a symmetric positive semidefinite matrix if and only if*

$$Q_{11} \ge 0, \quad Q_{22} - Q_{12}^T Q_{11}^+ Q_{12} \ge 0 \quad and \quad rank(Q_{11}) = rank(Q_{11}, Q_{12}).$$

Lemma 2.3 directly results in the following Lemmas 2.4 and 2.5.

Lemma 2.4. *Let $Q = (Q_{ij})_{2\times 2} \in SR^{n\times n}$ be a 2×2 real block symmetric matrix, with $Q_{11} \in SR^{r\times r}$ the known submatrix, and Q_{12} and Q_{22} two unknown submatrices. Then there* *exist matrices* Q_{12} *and* Q_{22} *such that* $Q \geq 0$ *if and only if* $Q_{11} \geq 0$. *Furthermore, all submatrices* Q_{12} *and* Q_{22} *such that* $Q \geq 0$ *can be expressed as*

$$Q_{12} = Q_{11}Z, \quad Q_{22} = Z^T Q_{11} Z + G,$$

where $\forall Z \in R^{r\times(n-r)}$ *and* $\forall G \in SR_0^{(n-r)\times(n-r)}$.

Lemma 2.5. *Let* $Q = (Q_{ij})_{3\times 3} \in R^{n\times n}$ *be a* 3×3 *real block symmetric matrix, with* $Q_{11} \in R^{r\times r}$ *and* $Q_{33} \in R^{s\times s}$ *two unknown submatrices, and the other blocks the known submatrices. Then there exist* Q_{11} *and* Q_{33} *such that* $Q \geq 0$ *if and only if*

$$Q_{22} \geq 0 \quad \textit{and} \quad rank(Q_{22}) = rank(Q_{21}, Q_{22}, Q_{23}).$$

Furthermore, all submatrices Q_{11} *and* Q_{33} *such that* $Q \geq 0$ *can be expressed as*

$$\begin{cases} Q_{11} = Q_{12}Q_{22}^{+}Q_{21} + G_1, \\ Q_{33} = Q_{32}Q_{22}^{+}Q_{23} + (Q_{31} - Q_{32}Q_{22}^{+}Q_{21})G_1^{+}(Q_{13} - Q_{12}Q_{22}^{+}Q_{23}) + G_2, \end{cases} \tag{2.9}$$

where $\forall G_1 \in SR_0^{r\times r}, \forall G_2 \in SR_0^{s\times s}$ *and* G_1 *satisfies*

$$rank(G_1) = rank(G_1, Q_{13} - Q_{12}Q_{22}^{+}Q_{23}).$$

Proof. It follows from Lemma 2.3 that

$$Q \geq 0 \Longleftrightarrow \begin{pmatrix} Q_{11} & Q_{12} & Q_{13} \\ Q_{21} & Q_{22} & Q_{23} \\ Q_{31} & Q_{32} & Q_{33} \end{pmatrix} \geq 0 \Longleftrightarrow \begin{pmatrix} Q_{22} & Q_{21} & Q_{23} \\ Q_{12} & Q_{11} & Q_{13} \\ Q_{32} & Q_{31} & Q_{33} \end{pmatrix} \geq 0$$

$$\Longleftrightarrow Q_{22} \geq 0, \begin{pmatrix} Q_{11} & Q_{13} \\ Q_{31} & Q_{33} \end{pmatrix} - \begin{pmatrix} Q_{12} \\ Q_{32} \end{pmatrix} Q_{22}^{+}(Q_{21}, Q_{23}) \geq 0 \text{ and } \mathrm{rank}(Q_{22}) = \mathrm{rank}(Q_{22}, Q_{21}, Q_{23})$$

$$\Longleftrightarrow Q_{22} \geq 0, \quad \mathrm{rank}(Q_{22}) = \mathrm{rank}(Q_{21}, Q_{22}, Q_{23}) \quad \text{and}$$

$$\begin{pmatrix} Q_{11} - Q_{12}Q_{22}^{+}Q_{21} & Q_{13} - Q_{12}Q_{22}^{+}Q_{23} \\ Q_{31} - Q_{32}Q_{22}^{+}Q_{21} & Q_{33} - Q_{32}Q_{22}^{+}Q_{23} \end{pmatrix} \geq 0. \tag{2.10}$$

By Lemma 2.3 we know that there exist submatrices Q_{11} and Q_{33} such that (2.10) holds, and all submatrices Q_{11} and Q_{33} such that (2.10) holds can be expressed by (2.9) when $Q_{22} \geq 0$. This completes the proof of this lemma.

Proof of Theorem 2.2. For any $X \in SR_0^{n\times n}$, partition the matrix $(M^T X M)$ as

$$M^T X M = \begin{array}{c} \begin{pmatrix} X_{11} & X_{12} & X_{13} & X_{14} \\ X_{12}^T & X_{22} & X_{23} & X_{24} \\ X_{13}^T & X_{23}^T & X_{33} & X_{34} \\ X_{14}^T & X_{24}^T & X_{34}^T & X_{44} \end{pmatrix} \\ \begin{matrix} \scriptstyle r \quad & \scriptstyle s \quad & \scriptstyle k-r-s & \scriptstyle n-k \end{matrix} \end{array} \begin{array}{c} r \\ s \\ k-r-s \\ n-k \end{array}. \tag{2.11}$$

It follows from (2.2) and (2.11) that

$$\begin{aligned} \|AXB - D\|^2 &= \|\Sigma_A^T M^T X M \Sigma_B - U^T D V\|^2 \\ &= \left\| \begin{pmatrix} -D_{11} & X_{12}S_B - D_{12} & X_{13} - D_{13} \\ -D_{21} & S_A X_{22} S_B - D_{22} & S_A X_{23} - D_{23} \\ -D_{31} & -D_{32} & -D_{33} \end{pmatrix} \right\|^2 \\ &= \|S_A X_{22} S_B - D_{22}\|^2 + \|X_{12}S_B - D_{12}\|^2 \\ &\quad + \|S_A X_{23} - D_{23}\|^2 + \|X_{13} - D_{13}\|^2 + C^*, \end{aligned} \tag{2.12}$$

where

$$C^* = \|D_{11}\|^2 + \|D_{21}\|^2 + \|D_{31}\|^2 + \|D_{32}\|^2 + \|D_{33}\|^2. \tag{2.13}$$

From Theorem 2.1 we know that there exists a unique matrix $\hat{X}_{22}$ which minimizes $\|S_A X_{22} S_B - D_{22}\|$ with respect to $X_{22} \in SR_0^{s\times s}$. Obviously, if there exist submatrices $X_{11}, X_{33}, X_{14}, X_{24}, X_{34}$ and X_{44} such that the following matrix X:

$$X = M^{-T}\begin{pmatrix} X_{11} & D_{12}S_B^{-1} & D_{13} & X_{14} \\ S_B^{-1}D_{12}^T & \hat{X}_{22} & S_A^{-1}D_{23} & X_{24} \\ D_{13}^T & D_{23}^T S_A^{-1} & X_{33} & X_{34} \\ X_{14}^T & X_{24}^T & X_{34}^T & X_{44} \end{pmatrix} M^{-1} \tag{2.14}$$

is a symmetric positive semidefinite matrix, then by (2.12) we know that the matrix X is certainly a solution of Problem I, where the sizes of submatrices $X_{11}, X_{33}, X_{14}, X_{24}, X_{34}$ and X_{44} are the same as (2.11).

Thus, if (2.3) holds, i.e.,

$$\mathrm{rank}(\hat{X}_{22}) = \mathrm{rank}(S_B^{-1}D_{12}^T, \hat{X}_{22}, S_A^{-1}D_{23}),$$

then by Lemma 2.5 we know that there exist submatrices X_{11} and X_{33} such that

$$\begin{pmatrix} X_{11} & D_{12}S_B^{-1} & D_{13} \\ S_B^{-1}D_{12}^T & \hat{X}_{22} & S_A^{-1}D_{23} \\ D_{13}^T & D_{23}^T S_A^{-1} & X_{33} \end{pmatrix} \geq 0, \tag{2.15}$$

and that all submatrices X_{11} and X_{33} such that (2.15) holds can be expressed as (2.5), (2.6) and (2.7). Therefore, by Lemma 2.4 we know that Problem I has a solution when

$$\mathrm{rank}(\hat{X}_{22}) = \mathrm{rank}(S_B^{-1}D_{12}^T, \hat{X}_{22}, S_A^{-1}D_{23}),$$

and in this case the solution X of Problem I can be expressed by (2.4).

Conversely, we only need to show that there is no minimum point to Problem I when

$$\mathrm{rank}(\hat{X}_{22}) \neq \mathrm{rank}(S_B^{-1}D_{12}^T, \hat{X}_{22}, S_A^{-1}D_{23}).$$

Suppose that

$$\mathrm{rank}(\hat{X}_{22}) \neq \mathrm{rank}(S_B^{-1}D_{12}^T, \hat{X}_{22}, S_A^{-1}D_{23}),$$

and that there is a minimum point, say $X_0 \in SR_0^{n\times n}$. Then by partitioning the matrix $(M^T X_0 M)$ into

$$M^T X_0 M = \begin{array}{cc} \begin{pmatrix} \bar{X}_{11} & \bar{X}_{12} & \bar{X}_{13} & \bar{X}_{14} \\ \bar{X}_{12}^T & \bar{X}_{22} & \bar{X}_{23} & \bar{X}_{24} \\ \bar{X}_{13}^T & \bar{X}_{23}^T & \bar{X}_{33} & \bar{X}_{34} \\ \bar{X}_{14}^T & \bar{X}_{24}^T & \bar{X}_{34}^T & \bar{X}_{44} \end{pmatrix} & \begin{matrix} r \\ s \\ k-r-s \\ n-k \end{matrix} \\ \begin{matrix} \;\; r \quad & \quad s \quad & k-r-s & n-k \end{matrix} & \end{array}, \tag{2.16}$$

we obtain from Lemma 2.5 that $\bar{X}_{22} \geq 0$ and $\mathrm{rank}(\bar{X}_{22}) = \mathrm{rank}(\bar{X}_{12}^T, \bar{X}_{22}, \bar{X}_{23})$.

If

$$(\bar{X}_{12}^T, \bar{X}_{23}) = (S_B^{-1}D_{12}^T, S_A^{-1}D_{23}),$$

then $\bar{X}_{22} \neq \hat{X}_{22}$. Otherewise, we have

$$\mathrm{rank}(\hat{X}_{22}) = \mathrm{rank}(\bar{X}_{22}) = rank(\bar{X}_{12}^T, \bar{X}_{22}, \bar{X}_{23}) = \mathrm{rank}(S_B^{-1}D_{12}^T, \hat{X}_{22}, S_A^{-1}D_{23}),$$

which obviously contradicts the initial assumption that

$$\mathrm{rank}(\hat{X}_{22}) \neq \mathrm{rank}(S_B^{-1}D_{12}^T, \hat{X}_{22}, S_A^{-1}D_{23}).$$

Hence, in both cases

$$(\bar{X}_{12}^T, \bar{X}_{23}) = (S_B^{-1}D_{12}^T, S_A^{-1}D_{23})$$

and

$$(\bar{X}_{12}^T, \bar{X}_{23}) \neq (S_B^{-1}D_{12}^T, S_A^{-1}D_{23}),$$

it follows from (2.2), (2.16), (2.13)-(2.14) and Lemma 2.2 that

$$\begin{aligned}\|AX_0B-D\|^2 &= \|S_A\bar{X}_{22}S_B-D_{22}\|^2+\|\bar{X}_{12}S_B-D_{12}\|^2\\ &\quad+\|S_A\bar{X}_{23}-D_{23}\|^2+\|\bar{X}_{13}-D_{13}\|^2+C^*\\ &> \|S_A\hat{X}_{22}S_B-D_{22}\|^2+C^*\\ &= min_{X_{22}\in SR_0^{s\times s}}\|S_AX_{22}S_B-D_{22}\|^2+C^*\\ &= inf_{X_{22}\in SR_+^{s\times s}}\|S_AX_{22}S_B-D_{22}\|^2+C^*\\ &= inf_{X\in SR_+^{n\times n}}\|AXB-D\|^2\\ &= min_{X\in SR_0^{n\times n}}\|AXB-D\|^2.\end{aligned}$$

This contradicts the optimality of X_0, and therefore, contradicts the existence of a minimum for Problem I when

$$rank(\hat{X}_{22})\neq rank(S_B^{-1}D_{12}^T,\hat{X}_{22},S_A^{-1}D_{23}).$$

This completes the proof.

Theorems 2.1 and 2.2 directly result in the following Corollary 2.1.

Corollary 2.1. Let $A\in R^{m\times n}$, $B\in R^{n\times p}$ and $D\in R^{m\times p}$. Assume that the matrix pair $[A^T,B]$ has GSVD (2.2). Denote by D_{ij} the matrix $(U_i^TDV_j)(i,j=1,2,3)$. Then Problem I exists a unique solution if and only if

$$rank(A)=rank(B)=n.$$

Moreover, if Problem I exists a unique solution $\hat{X}$, then this solution has the expression

$$\hat{X}=M^{-T}\hat{X}_{22}M^{-1},$$

where $\hat{X}_{22}$ is a unique minimizer of $\|S_AX_{22}S_B-D_{22}\|$ with respect to $X_{22}\in SR_0^{s\times s}$.

3. Numerical examples

Without loss of generality, we take numerical examples only when rank(A) = rank(B) = n.

Based on Corollary 2.1, we formulate the following algorithm to find the solution $\hat{X}$ of Problem I.

Algorithm I. Let $A\in R^{m\times n}$, $B\in R^{n\times p}$ and $D\in R^{m\times p}$ satisfying rank(A) = rank(B) = n. Then Problem I can be solved in the following steps:

Step 1. Make the GSVD of the matrix pair $[A^T,B]$ as (2.2), and determine the nonsingular matrix M, two orthogonal matrices

$$U=(\underset{n}{U_2},\quad \underset{m-n}{U_3}),\qquad V=(\underset{p-n}{V_1},\quad \underset{n}{V_2}),$$

and two diagonal matrices S_A and S_B.

Step 2. Compute $D_{22}=U_2^TDV_2$.

Step 3. Find $\hat{X}_{22}$ such that $\|S_A\hat{X}_{22}S_B-D_{22}\|=min_{X_{22}\in SR_0^{n\times n}}\|S_AX_{22}S_B-D_{22}\|$ by Steps 3.1-3.3:

Step 3.1. Let $X_{22}=R^TR$, where R is a real $n\times n$ upper triangular matrix.

Step 3.2. Apply function **leastsq** in MATLAB 5.2 (in Optimization Toolbox) to find an $n\times n$ upper triangular matrix $\hat{R}$ such that

$$\|S_A\hat{R}^T\hat{R}S_B-D_{22}\|=min_{R\in UR^{n\times n}}\|S_AR^TRS_B-D_{22}\|,$$

where $UR^{n\times n}$ represents the set of all real $n\times n$ upper triangular matrix.

Step 3.3. Compute $\hat{X}_{22}=\hat{R}^T\hat{R}$.

Step 4. Compute $\hat{X}=M^{-T}\hat{X}_{22}M^{-1}$.

Example 1. Let

$$A=\begin{pmatrix}1&1&1&1\\0&1&1&1\\0&0&1&1\\0&0&0&1\end{pmatrix},\qquad B=\begin{pmatrix}1&2&3&4&5\\1&2&3&4&0\\1&2&3&0&0\\1&2&0&0&0\end{pmatrix},\qquad D=\begin{pmatrix}4&4&4&4&1\\4&4&4&4&2\\4&4&4&4&3\\4&4&4&4&4\end{pmatrix}.$$

By Algorithm I we can obtain

$S_A = diag(0.67900045105510, 0.29237120107268, 0.22052424039119, 0.11580293242550)$,

$S_B = diag(0.73413785317675, 0.95630490994416, 0.97538149428820, 0.99327220883384)$,

$$D_{22} = \begin{pmatrix} 7.17868562330402 & -4.34925845431121 & 5.23827163251109 & 13.06724226129175 \\ -0.40736482721515 & 0.40355766917324 & -1.04429834846536 & 0.56545475670779 \\ 0.25850324718914 & -0.28398319841221 & 0.79562977702294 & -0.59141348223689 \\ 0.35408102236058 & -0.37013855799996 & 0.99999987562708 & -0.65296851547864 \end{pmatrix},$$

$$M^{-1} = \begin{pmatrix} -0.01555475779938 & -0.00943920405869 & -0.02234168930958 & 0.36443137182595 \\ 0.03405753947243 & 0.05781738134286 & -0.37863876928710 & 0.24123666845791 \\ -0.12095689494276 & 0.30552896086328 & -0.12231369029055 & -0.02864203963210 \\ 0.15032348727376 & -0.02350051290568 & -0.01374893154067 & -0.01383273238163 \end{pmatrix}$$

and

$$\hat{X}_{22} = \begin{pmatrix} 16.92191115413709 & -5.00339698440652 & 6.71455609224128 & 17.15393787459080 \\ -5.00339698440652 & 3.26222671723772 & -4.10376434898835 & -3.05896814967633 \\ 6.71455609224128 & -4.10376434898835 & 5.18170828291030 & 4.41458237110209 \\ 17.15393787459080 & -3.05896814967633 & 4.41458237110209 & 19.66212838748324 \end{pmatrix}.$$

By Step 4 we obtain the solution $\hat{X}$ of Problem I as

$$\hat{X} = \begin{pmatrix} 0.32029627401864 & -0.11109026289899 & -0.11798376068749 & 0.51769660209808 \\ -0.11109026289899 & 0.28124971516215 & 0.17971787222005 & 0.16750828998783 \\ -0.11798376068749 & 0.17971787222005 & 0.12285189590644 & 0.00777278233746 \\ 0.51769660209808 & 0.16750828998783 & 0.00777278233746 & 1.33302305422671 \end{pmatrix}$$

and the residual error of $\hat{X}$ as

$$\|A\hat{X}B - D\| = 5.90033297964392.$$

Example 2[1]**.** Let

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & 6 & 0 \\ 4 & 3 & 0 \\ 0 & 0 & -0.5 \end{pmatrix}, \quad D = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -2 & 3 \\ 0 & 2 & 4 \end{pmatrix}.$$

By Algorithm I we can obtain the solution of Problem I as

$$\hat{X} = \begin{pmatrix} 0.22351194666855 & -0.11059428172018 & 0.24343268003258 \\ -0.11059428172018 & 0.05472233348565 & -0.12045111068973 \\ 0.24343268003258 & -0.12045111068973 & 0.26512887061414 \end{pmatrix},$$

and the residual error of $\hat{X}$ as

$$\|A\hat{X}B - D\| = 5.60099906971185.$$

Example 3. Let $\hat{X}$ represent the solution of Problem I, $\lambda_{min}(\hat{X})$ the smallest eigenvalue of $\hat{X}$. Denote by hilb(n) the n-th order Hilbert matrix, pascal(n) the n-th order Pascal matrix, magic(n) the n-th order Magic matrix, toeplitz$(1:n)$ the n-th order Toeplitz matrix whose first row is $(1, 2, ..., n)$, and hankel$(1:n)$ the n-th order Hankel matrix whose first row is $(1, 2, ..., n)$. For example,

$$\text{hilb}(4) = \begin{pmatrix} 1 & 1/2 & 1/3 & 1/4 \\ 1/2 & 1/3 & 1/4 & 1/5 \\ 1/3 & 1/4 & 1/5 & 1/6 \\ 1/4 & 1/5 & 1/6 & 1/7 \end{pmatrix}, \quad pascal(4) = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & 3 & 4 \\ 1 & 3 & 6 & 10 \\ 1 & 4 & 10 & 20 \end{pmatrix},$$

$$\mathrm{magic}(4) = \begin{pmatrix} 16 & 2 & 3 & 13 \\ 5 & 11 & 10 & 8 \\ 9 & 7 & 6 & 12 \\ 4 & 14 & 15 & 1 \end{pmatrix}, \quad \mathrm{toeplitz}(1:4) = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 1 & 2 & 3 \\ 3 & 2 & 1 & 2 \\ 4 & 3 & 2 & 1 \end{pmatrix},$$

$$\mathrm{hankel}(1:4) = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 4 & 0 \\ 3 & 4 & 0 & 0 \\ 4 & 0 & 0 & 0 \end{pmatrix}.$$

We run a variety of numerical experiments about Algorithm I for the above classes of matrices of different dimensions. Some numerical results are listed in the following table.

n	A	B	D	$\|A\hat{X}B - D\|$	$\lambda_{min}(\hat{X})$
5	pascal(n)	hilb(n)	magic(n)	27.65834202778205	0.00000000023753
8	toeplitz(1:n)	pascal(n)	magic(n)	1.555704103619826e+02	0.00027396920
10	hankel(1:n)	pascal(n)	hilb(n)	1.37750594595019	0.00000000000000
12	pascal(n)	hankel(1:n)	hilb(n)	1.44223367332441	0.00000000000000
15	toeplitz(1:n)	hankel(1:n)	pascal(n)	5.326721946822061e+07	0.00003917588353
18	pascal(n)	toeplitz(1:n)	hilb(n)	0.73324543583197	0.00000000001451
20	toeplitz(1:n)	hankel(1:n)	hilb(n)	0.76224214450855	0.00000000000000
22	toeplitz(1:n)	hankel(1:n)	hilb(n)	0.78857254740837	0.00000000000000
25	toeplitz(1:n)	hankel(1:n)	hilb(n)	0.82387824421197	0.00000000000002
28	hankel(1:n)	toeplitz(1:n)	hilb(n)	0.85505173764166	0.00000000000000
30	hankel(1:n)	toeplitz(1:n)	pascal(n)	4.001069977211434e+16	0.00000961119604
35	hankel(1:n)	toeplitz(1:n)	hilb(n)	0.91555674992815	0.00000000000024

From Examples 1-3, we clearly see that Algorithm I is feasible and accurate for solving Problem I.

Acknowledgements: Part of the first author's work was supported by State Key Laboratory of Scientific/Engineering Computing of Chinese Academy of Sciences.

References

[1] J.C. Allwright, Positive semidefinite matrices: Characterization via conical hulls and least-squares solution of a matrix equation, *SIAM J. Control Optim.*, **26**(1988), 537-556.

[2] J.C. Allwright, K.G. Woodgate, Errata and Addendum to: Positive semidefinite matrices: Characterization via conical hulls and least-squares solution of a matrix equation, *SIAM J. Control Optim.*, **28**(1990), 250-251.

[3] K.G. Woodgate, Least-squares solution of $F = PG$ over positive semidefinite matrices P, *Linear Algebra Appl.*, **245**(1996), 171-190.

[4] A.P. Liao, On the least squares problem of a matrix equation, *J. Comput. Math.*, **17**(1999), 589-584.

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

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

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

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

[9] A. Albert, Conditions for positive and nonnegative in terms of pseudoinverses, *SIAM J. Appl. Math.*, **17**(1969), 434-440.