C. R. Acad. Sci. Paris, Ser. I 353 (2015) 1053–1059

Contents lists available at ScienceDirect

C. R. Acad. Sci. Paris, Ser. I

www.sciencedirect.com

Numerical analysis

Solving the mixed Sylvester matrix equations by matrix decompositions

Résolution d'équations matricielles de Sylvester mixtes par décompositions de matrices

Yongxin Yuan

School of Mathematics and Statistics, Hubei Normal University, Huangshi 435002, PR China

ARTICLE INFO

Article history:
Received 21 March 2014
Accepted after revision 29 August 2015
Available online 23 October 2015

Presented by Olivier Pironneau

ABSTRACT

By applying the generalized singular-value decompositions (GSVDs) of matrix pairs, a necessary and sufficient solvability condition for mixed Sylvester equations is established, the explicit representation of the general solution is given. Also, the minimum-norm solution of the matrix equations is discussed.

RÉSUMÉ

En utilisant les décompositions en valeurs singulières généralisées (GSVDs) de couples de matrices, on établit une condition nécessaire et suffisante de résolubilité d'équations de Sylvester mixtes et on donne une représentation explicite de la solution générale. On étudie également la solution de norme minimale d'équations matricielles.

1. Introduction

The purpose of this work is to study the so-called mixed Sylvester matrix equations

$$A_1X - YB_1 = C_1,\ A_2Z - YB_2 = C_2, \tag{1}$$

where $A_1 \in \mathbf{C}^{m\times n}$, $B_1 \in \mathbf{C}^{l\times q}$, $C_1 \in \mathbf{C}^{m\times q}$, $A_2 \in \mathbf{C}^{m\times p}$, $B_2 \in \mathbf{C}^{l\times d}$ and $C_2 \in \mathbf{C}^{m\times d}$.

Eqs. (1) can also be equivalently written as (in matlab notation):

$$\mathrm{blkdiag}(A_1, A_2)\mathrm{blkdiag}(X, Z) - \mathrm{kron}(I_2, Y)\mathrm{blkdiag}(B_1, B_2) = \mathrm{blkdiag}(C_1, C_2).$$

There are some valuable works on formulating solutions to the mixed Sylvester matrix Eqs. (1). For example, Liu [5] derived a solvability condition of (1) by using the ranks of matrices. By applying the equivalence of matrices, Lee and Vu [4] showed that Eqs. (1) are consistent if and only if there exist invertible matrices R_1, R_2 and S such that

E-mail address: yuanyx_703@163.com.

http://dx.doi.org/10.1016/j.crma.2015.08.010
1631-073X/

$$\begin{bmatrix} A_1 & C_1 \\ 0 & B_1 \end{bmatrix} R_1 = S \begin{bmatrix} A_1 & 0 \\ 0 & B_1 \end{bmatrix}, \quad \begin{bmatrix} A_2 & C_2 \\ 0 & B_2 \end{bmatrix} R_2 = S \begin{bmatrix} A_2 & 0 \\ 0 & B_2 \end{bmatrix}.$$

Recently, Wang and He [8] provided some necessary and sufficient solvability conditions and the expression of the general solution of (1) by virtue of the ranks and generalized inverses of matrices.

The strategy adopted here is to use the generalized singular value decompositions (GSVDs) of matrix pairs to decouple the equations of (1) to obtain some profound results. Beginning in Section 2, we first consider some special cases where some constraints are imposed on coefficient matrices, then, for the general situation, we formulate the necessary and sufficient conditions for the existence of the solution of (1) directly by means of the generalized singular value decompositions of the matrix pairs (A_1, A_2) and (B_1, B_2), and construct the explicit representation of the general solution when it is solvable. Furthermore, we will provide the minimum-norm solution of (1) by using the expression of the general solution.

Throughout this paper, we denote the complex $m \times n$ matrix space by $\mathbf{C}^{m\times n}$, the set of all unitary matrices in $\mathbf{C}^{n\times n}$ by $\mathbf{UC}^{n\times n}$. A^{H} and A^{+} stand for the conjugate transpose and the Moore–Penrose generalized inverse of a complex matrix A, respectively. I_n represents the identity matrix of size n. We define an inner product: $\langle A, B\rangle = \operatorname{trace}(B^{\mathrm{H}}A)$ for all $A, B \in \mathbf{C}^{m\times n}$, then $\mathbf{C}^{m\times n}$ is a Hilbert inner product space and the norm of a matrix generated by this inner product is Frobenius norm. For $A = [\alpha_{ij}]_{m\times n}$ and $B = [\beta_{ij}]_{m\times n}$, $A * B$ represents the Hadamard product of the matrices A and B, that is, $A * B = [\alpha_{ij}\beta_{ij}]_{m\times n}$.

2. The solution to the mixed Sylvester matrix Eqs. (1)

The following lemmata are needed in what follows.

Lemma 1. *(See [2].) If $A \in \mathbf{C}^{m\times k}$, $B \in \mathbf{C}^{l\times n}$ and $C \in \mathbf{C}^{m\times n}$, then the equation $AXB = C$ has a solution $X \in \mathbf{C}^{k\times l}$ if and only if $AA^{+}CB^{+}B = C$. In this case, the general solution of the matrix equation $AXB = C$ can be described as $X = A^{+}CB^{+} + (I_k - A^{+}A)W + T(I_l - BB^{+})$, where $W, T \in \mathbf{C}^{k\times l}$ are arbitrary matrices.*

Lemma 2. *(See [1].) Let $A \in \mathbf{C}^{m\times k}$, $B \in \mathbf{C}^{l\times n}$ and $C \in \mathbf{C}^{m\times n}$. The equation*

$$AX - YB = C \tag{2}$$

has a solution $X \in \mathbf{C}^{k\times n}$, $Y \in \mathbf{C}^{m\times l}$ if and only if $(I_m - AA^{+})C(I_n - B^{+}B) = 0$. If this is the case, the general solution of (2) has the form

$$X = A^{+}C + A^{+}TB + (I_k - A^{+}A)W, \quad Y = -(I_m - AA^{+})CB^{+} + T - (I_m - AA^{+})TBB^{+},$$

where $W \in \mathbf{C}^{k\times n}$ and $T \in \mathbf{C}^{m\times l}$ are arbitrary matrices.

We consider some special cases.

Case 1. If A_1 is square and nonsingular, then from the first equation of (1), we can get

$$X = A_1^{-1}C_1 + A_1^{-1}YB_1. \tag{3}$$

By Lemma 2, the second equation of (1) has a solution $Z \in \mathbf{C}^{p\times d}$, $Y \in \mathbf{C}^{m\times l}$ if and only if

$$(I_m - A_2A_2^{+})C_2(I_d - B_2^{+}B_2) = 0.$$

If this is the case, the general solution of the matrix equation $A_2Z - YB_2 = C_2$ has the form

$$Z = A_2^{+}C_2 + A_2^{+}TB_2 + (I_p - A_2^{+}A_2)W, \tag{4}$$

$$Y = -(I_m - A_2A_2^{+})C_2B_2^{+} + T - (I_m - A_2A_2^{+})TB_2B_2^{+}, \tag{5}$$

where $W \in \mathbf{C}^{p\times d}$ and $T \in \mathbf{C}^{m\times l}$ are arbitrary matrices. Substituting (5) into (3), we have obtained the following result.

Theorem 1. *Suppose that $A_1 \in \mathbf{C}^{m\times m}$, $B_1 \in \mathbf{C}^{l\times q}$, $C_1 \in \mathbf{C}^{m\times q}$, $A_2 \in \mathbf{C}^{m\times p}$, $B_2 \in \mathbf{C}^{l\times d}$ and $C_2 \in \mathbf{C}^{m\times d}$. If A_1 is nonsingular, then the equation of (1) has a solution $X \in \mathbf{C}^{m\times q}$, $Y \in \mathbf{C}^{m\times l}$, $Z \in \mathbf{C}^{p\times d}$ if and only if $(I_m - A_2A_2^{+})C_2(I_d - B_2^{+}B_2) = 0$. In this case, the general solution of (1) can be expressed as*

$$X = A_1^{-1}C_1 - A_1^{-1}(I_m - A_2A_2^{+})C_2B_2^{+}B_1 + A_1^{-1}TB_1 - A_1^{-1}(I_m - A_2A_2^{+})TB_2B_2^{+}B_1,$$

$$Y = -(I_m - A_2A_2^{+})C_2B_2^{+} + T - (I_m - A_2A_2^{+})TB_2B_2^{+}, \quad Z = A_2^{+}C_2 + A_2^{+}TB_2 + (I_p - A_2^{+}A_2)W,$$

where $W \in \mathbf{C}^{p\times d}$ and $T \in \mathbf{C}^{m\times l}$ are arbitrary matrices.

If A_2 is square and nonsingular, the approach is similar.

Case 2. If B_1 is square and nonsingular, then from the first equation of (1), we can get

$$Y = A_1 X B_1^{-1} - C_1 B_1^{-1}. \tag{6}$$

Substituting (6) into the second equation of (1), we have

$$A_2 Z - A_1 X B_1^{-1} B_2 = C_2 - C_1 B_1^{-1} B_2. \tag{7}$$

Applying Lemma 1, the equation of (7) with respect to Z has a solution if and only if

$$\tilde{A} X \tilde{B} = \tilde{C}, \tag{8}$$

where $\tilde{A} = (I_m - A_2 A_2^+)A_1$, $\tilde{B} = B_1^{-1} B_2$ and $\tilde{C} = (I_m - A_2 A_2^+)(C_1 B_1^{-1} B_2 - C_2)$. Using Lemma 1 again, the equation of (8) has a solution $X \in \mathbf{C}^{n\times l}$ if and only if $\tilde{A}\tilde{A}^+\tilde{C}\tilde{B}^+\tilde{B} = \tilde{C}$. In this case, the general solution of (8) can be described as

$$X = \tilde{A}^+\tilde{C}\tilde{B}^+ + (I_n - \tilde{A}^+\tilde{A})W + T(I_l - \tilde{B}\tilde{B}^+), \quad \text{where } W, T \in \mathbf{C}^{n\times l} \text{ are arbitrary matrices.}$$

Summing up above discussion, we have proved the following result.

Theorem 2. *Let $A_1 \in \mathbf{C}^{m\times n}$, $B_1 \in \mathbf{C}^{l\times l}$, $C_1 \in \mathbf{C}^{m\times l}$, $A_2 \in \mathbf{C}^{m\times p}$, $B_2 \in \mathbf{C}^{l\times d}$ and $C_2 \in \mathbf{C}^{m\times d}$, and let $\tilde{A} = (I_m - A_2 A_2^+)A_1$, $\tilde{B} = B_1^{-1}B_2$ and $\tilde{C} = (I_m - A_2 A_2^+)(C_1 B_1^{-1} B_2 - C_2)$. If B_1 is nonsingular, then the equation of (1) has a solution $X \in \mathbf{C}^{n\times l}$, $Y \in \mathbf{C}^{m\times l}$, $Z \in \mathbf{C}^{p\times d}$ if and only if $\tilde{A}\tilde{A}^+\tilde{C}\tilde{B}^+\tilde{B} = \tilde{C}$. In this case, the general solution of (1) can be expressed as*

$$\begin{aligned}
X &= \tilde{A}^+\tilde{C}\tilde{B}^+ + (I_n - \tilde{A}^+\tilde{A})W + T(I_l - \tilde{B}\tilde{B}^+),\\
Y &= A_1\tilde{A}^+\tilde{C}\tilde{B}^+B_1^{-1} + A_1(I_n - \tilde{A}^+\tilde{A})WB_1^{-1} + A_1T(I_l - \tilde{B}\tilde{B}^+)B_1^{-1} - C_1B_1^{-1},\\
Z &= A_2^+A_1\tilde{A}^+\tilde{C}\tilde{B}^+\tilde{B} + A_2^+A_1(I_n - \tilde{A}^+\tilde{A})W\tilde{B} + A_2^+(C_2 - C_1B_1^{-1}B_2) + (I_p - A_2^+A_2)L,
\end{aligned}$$

where $W \in \mathbf{C}^{n\times l}$, $T \in \mathbf{C}^{n\times l}$ and $L \in \mathbf{C}^{p\times d}$ are arbitrary matrices.

If B_2 is square and nonsingular, the approach is similar.

Now, assume that $A_1 \in \mathbf{C}^{m\times n}$, $B_1 \in \mathbf{C}^{l\times q}$, $A_2 \in \mathbf{C}^{m\times p}$ and $B_2 \in \mathbf{C}^{l\times d}$ are arbitrary complex matrices. The GSVD (see, e.g., [3, 6,7]) of the matrix pair (A_1, A_2) is of the form

$$A_1 = M\Sigma_1 U^{\mathrm{H}}, \;\; A_2 = M\Sigma_2 V^{\mathrm{H}}, \tag{9}$$

where $U \in \mathbf{UC}^{n\times n}$, $V \in \mathbf{UC}^{p\times p}$ and $M \in \mathbf{C}^{m\times m}$ is a nonsingular matrix, and

$$\Sigma_1 = \begin{bmatrix} I & 0 & 0 \\ 0 & S_1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{matrix} r-s \\ s \\ k-r \\ m-k \end{matrix}, \qquad \Sigma_2 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & S_2 & 0 \\ 0 & 0 & I \\ 0 & 0 & 0 \end{bmatrix} \begin{matrix} r-s \\ s \\ k-r \\ m-k \end{matrix},$$
$$\qquad \begin{matrix} r-s & s & n-r \end{matrix} \qquad\qquad\qquad \begin{matrix} g & s & k-r \end{matrix}$$

$$g = p + r - k - s,\; r = \operatorname{rank}(A_1),\; k = \operatorname{rank}[A_1, A_2], \quad \text{and} \quad S_1 = \operatorname{diag}\{\alpha_1, \cdots, \alpha_s\}, \;\; S_2 = \operatorname{diag}\{\beta_1, \cdots, \beta_s\}$$
$$\text{with } 1 > \alpha_1 \ge \alpha_2 \ge \cdots \ge \alpha_s > 0, \;\; 0 < \beta_1 \le \beta_2 \le \cdots \le \beta_s < 1, \;\; \alpha_i^2 + \beta_i^2 = 1 \; (i = 1, \cdots, s).$$

Likewise, the GSVD of the matrix pair (B_1, B_2) is of the form

$$B_1 = N\Omega_1 P^{\mathrm{H}}, \;\; B_2 = N\Omega_2 Q^{\mathrm{H}}, \tag{10}$$

where $P \in \mathbf{UC}^{q\times q}$, $Q \in \mathbf{UC}^{d\times d}$ and $N \in \mathbf{C}^{l\times l}$ is a nonsingular matrix, and

$$\Omega_1 = \begin{bmatrix} I & 0 & 0 \\ 0 & D_1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{matrix} h-t \\ t \\ e-h \\ l-e \end{matrix}, \qquad \Omega_2 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & D_2 & 0 \\ 0 & 0 & I \\ 0 & 0 & 0 \end{bmatrix} \begin{matrix} h-t \\ t \\ e-h \\ l-e \end{matrix},$$
$$\qquad \begin{matrix} h-t & t & q-h \end{matrix} \qquad\qquad\qquad \begin{matrix} u & t & e-h \end{matrix}$$

$$u = d + h - e - t,\; h = \operatorname{rank}(B_1),\; e = \operatorname{rank}[B_1, B_2], \quad \text{and} \quad D_1 = \operatorname{diag}\left\{\gamma_1, \cdots, \gamma_t\right\}, \;\; D_2 = \operatorname{diag}\{\delta_1, \cdots, \delta_t\}$$
$$\text{with} \;\; 1 > \gamma_1 \ge \gamma_2 \ge \cdots \ge \gamma_t > 0, \;\; 0 < \delta_1 \le \delta_2 \le \cdots \le \delta_t < 1, \;\; \gamma_i^2 + \delta_i^2 = 1 \; (i = 1, \cdots, t).$$

Lemma 3. *Suppose that* $E_{22}, F_{22} \in \mathbf{C}^{s\times t}$, $S_1 = \operatorname{diag}\{\alpha_1, \cdots, \alpha_s\}$, $S_2 = \operatorname{diag}\{\beta_1, \cdots, \beta_s\}$ $(\alpha_i > 0, \beta_i > 0, i = 1, \cdots, s)$, $D_1 = \operatorname{diag}\{\gamma_1, \cdots, \gamma_t\}$ *and* $D_2 = \operatorname{diag}\{\delta_1, \cdots, \delta_t\}$ $(\gamma_j > 0, \delta_j > 0, j = 1, \cdots, t)$. *Then there exists a unique* Y_{22} *such that*

$$f = \left\| S_1^{-1}E_{22} + S_1^{-1}Y_{22}D_1 \right\|^2 + \left\| S_2^{-1}F_{22} + S_2^{-1}Y_{22}D_2 \right\|^2 = \min \tag{11}$$

holds, and Y_{22} *can be expressed as*

$$Y_{22} = -K * (S_1^2F_{22}D_2 + S_2^2E_{22}D_1),\ K = [k_{ij}]_{s\times t},\ k_{ij} = \frac{1}{\alpha_i^2\delta_j^2 + \beta_i^2\gamma_j^2},\ i = 1, \cdots, s;\ j = 1, \cdots, t. \tag{12}$$

Proof. If we write $E_{22} = [e_{ij}]_{s\times t}$, $F_{22} = [f_{ij}]_{s\times t}$, $Y_{22} = [y_{ij}]_{s\times t}$, then the minimization problem of (11) is equivalent to

$$f = \sum_{i=1}^{s}\sum_{j=1}^{t}\left(\left| \frac{1}{\alpha_i}e_{ij} + \frac{\gamma_j}{\alpha_i}y_{ij} \right|^2 + \left| \frac{1}{\beta_i}f_{ij} + \frac{\delta_j}{\beta_i}y_{ij} \right|^2 \right) = \min .$$

Clearly, f is a continuously differentiable function of $2st$ variables of $\mathrm{Re}(y_{ij})$ and $\mathrm{Im}(y_{ij})$, $i = 1, \cdots, s$; $j = 1, \cdots, t$, and the function of concerning variables y_{ij} in f is

$$f(y_{ij}) = \left| \frac{1}{\alpha_i}e_{ij} + \frac{\gamma_j}{\alpha_i}y_{ij} \right|^2 + \left| \frac{1}{\beta_i}f_{ij} + \frac{\delta_j}{\beta_i}y_{ij} \right|^2,\ i = 1, \cdots, s;\ j = 1, \cdots, t.$$

It is easy to verify that the function f attains the smallest value at

$$\frac{\partial f(y_{ij})}{\partial \mathrm{Re}(y_{ij})} = 0,\ \frac{\partial f(y_{ij})}{\partial \mathrm{Im}(y_{ij})} = 0,\ i = 1, \cdots, s;\ j = 1, \cdots, t,$$

$$\text{which yields}\quad y_{ij} = \frac{-\alpha_i^2 f_{ij}\delta_j - \beta_i^2 e_{ij}\gamma_j}{\alpha_i^2\delta_j^2 + \beta_i^2\gamma_j^2},\ i = 1, \cdots, s;\ j = 1, \cdots, t. \tag{13}$$

The expression (12) of the matrix Y_{22} follows from (13) straightforwardly. □

Theorem 3. *Suppose that* $A_1 \in \mathbf{C}^{m\times n}$, $B_1 \in \mathbf{C}^{l\times q}$, $C_1 \in \mathbf{C}^{m\times q}$, $A_2 \in \mathbf{C}^{m\times p}$, $B_2 \in \mathbf{C}^{l\times d}$ *and* $C_2 \in \mathbf{C}^{m\times d}$, *and the GSVDs of* (A_1, A_2) *and* (B_1, B_2) *are given by (9) and (10), respectively. Let* $M^{-1}C_1P = [E_{ij}]_{4\times 3}$, $M^{-1}C_2Q = [F_{ij}]_{4\times 3}$, *where the row partitions of the matrices* $M^{-1}C_1P$ *and* $M^{-1}C_2Q$ *are, respectively, compatible with those of* Σ_1 *and* Σ_2, *and the column partitions of the matrices* $M^{-1}C_1P$ *and* $M^{-1}C_2Q$ *are, respectively, compatible with those of* Ω_1 *and* Ω_2. *Then the mixed Sylvester matrix Eqs. (1) is consistent if and only if*

$$E_{33} = 0,\ E_{43} = 0,\ F_{11} = 0,\ F_{41} = 0,\ E_{42}D_1^{-1} = F_{42}D_2^{-1}. \tag{14}$$

In this case, the general solution of (1) can be expressed as

$$X = U\begin{bmatrix} Y_{11} + E_{11} & E_{12} - F_{12}D_2^{-1}D_1 & E_{13} \\ S_1^{-1}E_{21} + S_1^{-1}Y_{21} & S_1^{-1}E_{22} + S_1^{-1}Y_{22}D_1 & S_1^{-1}E_{23} \\ X_{31} & X_{32} & X_{33} \end{bmatrix}P^{\mathrm{H}}, \tag{15}$$

$$Y = M\begin{bmatrix} Y_{11} & -F_{12}D_2^{-1} & -F_{13} & Y_{14} \\ Y_{21} & Y_{22} & Y_{23} & Y_{24} \\ -E_{31} & -E_{32}D_1^{-1} & Y_{33} & Y_{34} \\ -E_{41} & -F_{42}D_2^{-1} & -F_{43} & Y_{44} \end{bmatrix}N^{-1}, \tag{16}$$

$$Z = V\begin{bmatrix} Z_{11} & Z_{12} & Z_{13} \\ S_2^{-1}F_{21} & S_2^{-1}F_{22} + S_2^{-1}Y_{22}D_2 & S_2^{-1}F_{23} + S_2^{-1}Y_{23} \\ F_{31} & F_{32} - E_{32}D_1^{-1}D_2 & F_{33} + Y_{33} \end{bmatrix}Q^{\mathrm{H}}, \tag{17}$$

where $Y_{11}, Y_{33}, Y_{i4}, Y_{2j}, X_{3j}, Z_{1j}$ $(i = 1, 2, 3, 4;\ j = 1, 2, 3)$ *are all arbitrary matrices.*

Proof. By using (9) and (10), Eqs. (1) can be equivalently written as

$$M\Sigma_1U^{\mathrm{H}}X - YN\Omega_1P^{\mathrm{H}} = C_1, \tag{18}$$

$$M\Sigma_2V^{\mathrm{H}}Z - YN\Omega_2Q^{\mathrm{H}} = C_2. \tag{19}$$

$$\text{Write}\quad M^{-1}YN = [Y_{ij}]_{4\times 4},\ U^{\mathrm{H}}XP = [X_{ij}]_{3\times 3},\ V^{\mathrm{H}}ZQ = [Z_{ij}]_{3\times 3}, \tag{20}$$

then the equations of (18) and (19) are respectively equivalent to

$$\begin{bmatrix} X_{11}-Y_{11} & X_{12}-Y_{12}D_1 & X_{13} \\ S_1X_{21}-Y_{21} & S_1X_{22}-Y_{22}D_1 & S_1X_{23} \\ -Y_{31} & -Y_{32}D_1 & 0 \\ -Y_{41} & -Y_{42}D_1 & 0 \end{bmatrix} = \begin{bmatrix} E_{11} & E_{12} & E_{13} \\ E_{21} & E_{22} & E_{23} \\ E_{31} & E_{32} & E_{33} \\ E_{41} & E_{42} & E_{43} \end{bmatrix}$$

and

$$\begin{bmatrix} 0 & -Y_{12}D_2 & -Y_{13} \\ S_2Z_{21} & S_2Z_{22}-Y_{22}D_2 & S_2Z_{23}-Y_{23} \\ Z_{31} & Z_{32}-Y_{32}D_2 & Z_{33}-Y_{33} \\ 0 & -Y_{42}D_2 & -Y_{43} \end{bmatrix} = \begin{bmatrix} F_{11} & F_{12} & F_{13} \\ F_{21} & F_{22} & F_{23} \\ F_{31} & F_{32} & F_{33} \\ F_{41} & F_{42} & F_{43} \end{bmatrix}.$$

Thus we have

$$E_{33}=0,\ \ E_{43}=0,\ \ F_{11}=0,\ \ F_{41}=0, \tag{21}$$

$$-Y_{42}D_1=E_{42},\ \ -Y_{42}D_2=F_{42}, \tag{22}$$

$$X_{12}-Y_{12}D_1=E_{12},\ \ -Y_{12}D_2=F_{12}, \tag{23}$$

$$S_1X_{22}-Y_{22}D_1=E_{22},\ \ S_2Z_{22}-Y_{22}D_2=F_{22}, \tag{24}$$

$$-Y_{32}D_1=E_{32},\ \ Z_{32}-Y_{32}D_2=F_{32}, \tag{25}$$

$$Y_{31}=-E_{31},\ \ Y_{41}=-E_{41},\ \ Y_{13}=-F_{13},\ \ Y_{43}=-F_{43}, \tag{26}$$

$$X_{13}=E_{13},\ \ Z_{31}=F_{31},\ \ S_1X_{23}=E_{23},\ \ S_2Z_{21}=F_{21}, \tag{27}$$

$$X_{11}-Y_{11}=E_{11},\ \ Z_{33}-Y_{33}=F_{33},\ \ S_1X_{21}-Y_{21}=E_{21},\ \ S_2Z_{23}-Y_{23}=F_{23}. \tag{28}$$

Using the relations of (20)–(28), we can easily obtain the expressions (14)–(17) □

Corollary 4. *Under the same assumptions as in Theorem 3, if the condition of (14) is satisfied, then a solution (X, Y, Z) of (1) satisfies $\|X\|^2+\|Z\|^2=\min$ if and only if*

$$Y_{11}=-E_{11},\ Y_{21}=-E_{21},\ Y_{23}=-F_{23},\ Y_{33}=-F_{33},\ X_{3j}=0,\ Z_{1j}=0\ (j=1,2,3).$$

In this case, such solution can be expressed as

$$X=U\begin{bmatrix} 0 & E_{12}-F_{12}D_2^{-1}D_1 & E_{13} \\ 0 & S_1^{-1}E_{22}+S_1^{-1}Y_{22}D_1 & S_1^{-1}E_{23} \\ 0 & 0 & 0 \end{bmatrix}P^{\mathrm{H}}, \tag{29}$$

$$Y=M\begin{bmatrix} -E_{11} & -F_{12}D_2^{-1} & -F_{13} & Y_{14} \\ -E_{21} & Y_{22} & -F_{23} & Y_{24} \\ -E_{31} & -E_{32}D_1^{-1} & -F_{33} & Y_{34} \\ -E_{41} & -F_{42}D_2^{-1} & -F_{43} & Y_{44} \end{bmatrix}N^{-1}, \tag{30}$$

$$Z=V\begin{bmatrix} 0 & 0 & 0 \\ S_2^{-1}F_{21} & S_2^{-1}F_{22}+S_2^{-1}Y_{22}D_2 & 0 \\ F_{31} & F_{32}-E_{32}D_1^{-1}D_2 & 0 \end{bmatrix}Q^{\mathrm{H}}, \tag{31}$$

$$\text{where}\quad Y_{22}=-K*(S_1^2F_{22}D_2+S_2^2E_{22}D_1),\ K=[k_{ij}]_{s\times t},\ k_{ij}=\frac{1}{\alpha_i^2\delta_j^2+\beta_i^2\gamma_j^2}, i=1,\cdots,s;\ j=1,\cdots,t, \tag{32}$$

and Y_{i4} $(i=1,2,3,4)$ are arbitrary matrices.

Proof. It follows from (15) and (17) that $\|X\|^2+\|Z\|^2=\min$ if and only if

$$Y_{11}=-E_{11},\ Y_{21}=-E_{21},\ Y_{23}=-F_{23},\ Y_{33}=-F_{33}, \tag{33}$$

$$X_{3j}=0,\ Z_{1j}=0\ (j=1,2,3), \tag{34}$$

and

$$\left\|S_1^{-1}E_{22}+S_1^{-1}Y_{22}D_1\right\|^2+\left\|S_2^{-1}F_{22}+S_2^{-1}Y_{22}D_2\right\|^2=\min. \tag{35}$$

Applying Lemma 3, the expression (32) of the matrix Y_{22} follows from (35) straightforwardly. Now, substituting (32), (33) and (34) into the expressions of X, Y and Z in (15), (16) and (17), we can get the expressions (29), (30) and (31). □

Corollary 5. *Under the same assumptions as in Theorem 3, if the condition (14) is satisfied, then a solution (X, Y, Z) of (1) satisfies $\|X\|^2 + \|M^{-1}YN\|^2 + \|Z\|^2 = \min$ if and only if*

$$Y_{i4} = 0,\ X_{3j} = 0,\ Z_{1j} = 0\ (i = 1, 2, 3, 4;\ j = 1, 2, 3). \tag{36}$$

$$Y_{11} = -\frac{1}{2} E_{11}, \tag{37}$$

$$Y_{21} = -L * E_{21},\ L = [l_{ij}]_{s\times(h-t)},\ l_{ij} = \frac{1}{1+\alpha_i^2},\ i = 1, 2, \cdots, s;\ j = 1, 2, \cdots, h-t, \tag{38}$$

$$Y_{33} = -\frac{1}{2} F_{33}, \tag{39}$$

$$Y_{23} = -W * F_{23},\ W = [w_{ij}]_{s\times(e-h)},\ w_{ij} = \frac{1}{1+\beta_i^2},\ i = 1, 2, \cdots, s;\ j = 1, 2, \cdots, e-h, \tag{40}$$

$$Y_{22} = -H * (S_1^2 F_{22} D_2 + S_2^2 E_{22} D_1), \tag{41}$$

$$\text{where}\quad H = [h_{ij}]_{s\times t},\ h_{ij} = \frac{1}{\alpha_i^2\beta_i^2 + \alpha_i^2\delta_j^2 + \beta_i^2\gamma_j^2},\ i = 1, \cdots, s;\ j = 1, \cdots, t.$$

In this case, the unique solution can be expressed as

$$X = U \begin{bmatrix} \frac{1}{2}E_{11} & E_{12} - F_{12}D_2^{-1}D_1 & E_{13} \\ S_1^{-1}E_{21} - S_1^{-1}(L * E_{21}) & S_1^{-1}E_{22} + S_1^{-1}Y_{22}D_1 & S_1^{-1}E_{23} \\ 0 & 0 & 0 \end{bmatrix} P^{\mathrm{H}}, \tag{42}$$

$$Y = M \begin{bmatrix} -\frac{1}{2}E_{11} & -F_{12}D_2^{-1} & -F_{13} & 0 \\ -L * E_{21} & Y_{22} & -W * F_{23} & 0 \\ -E_{31} & -E_{32}D_1^{-1} & -\frac{1}{2}F_{33} & 0 \\ -E_{41} & -F_{42}D_2^{-1} & -F_{43} & 0 \end{bmatrix} N^{-1}, \tag{43}$$

$$Z = V \begin{bmatrix} 0 & 0 & 0 \\ S_2^{-1}F_{21} & S_2^{-1}F_{22} + S_2^{-1}Y_{22}D_2 & S_2^{-1}F_{23} - S_2^{-1}(W * F_{23}) \\ F_{31} & F_{32} - E_{32}D_1^{-1}D_2 & \frac{1}{2}F_{33} \end{bmatrix} Q^{\mathrm{H}}, \tag{44}$$

where Y_{22} is given by (41).

Proof. It follows from (15), (16) and (17) that $\|X\|^2 + \|M^{-1}YN\|^2 + \|Z\|^2 = \min$ if and only if

$$Y_{i4} = 0,\ X_{3j} = 0,\ Z_{1j} = 0\ (i = 1, 2, 3, 4;\ j = 1, 2, 3), \quad \text{and}$$

$$\|Y_{11} + E_{11}\|^2 + \|Y_{11}\|^2 = \min, \tag{45}$$

$$\|S_1^{-1}E_{21} + S_1^{-1}Y_{21}\|^2 + \|Y_{21}\|^2 = \min, \tag{46}$$

$$\|Y_{33} + F_{33}\|^2 + \|Y_{33}\|^2 = \min, \tag{47}$$

$$\|S_2^{-1}F_{23} + S_2^{-1}Y_{23}\|^2 + \|Y_{23}\|^2 = \min, \tag{48}$$

$$\left\|S_1^{-1}E_{22} + S_1^{-1}Y_{22}D_1\right\|^2 + \left\|S_2^{-1}F_{22} + S_2^{-1}Y_{22}D_2\right\|^2 + \|Y_{22}\|^2 = \min. \tag{49}$$

Applying a similar approach as in Lemma 3, we can obtain the expressions (37)–(41) from the minimization problems (45)–(49). Substituting (36)–(41) into the expressions of X, Y and Z in (15), (16) and (17), we can get the expressions (42), (43) and (44). □

3. A numerical example

Example 1. Assume that $m = n = 4, l = 5, q = 3$ and $p = d = 3$. Given

$$A_1=\begin{bmatrix}0.42284 & 0.46013 & 0.16048 & 0.2385\\ 0.56442 & 0.339 & 0.64263 & 0.82737\\ 0.34074 & -0.056611 & 0.79468 & 0.9827\\ 0.88458 & 1.2214 & -0.067213 & 0.020226\end{bmatrix},\quad A_2=\begin{bmatrix}-0.093993 & 0.35871 & -0.18154\\ -0.040158 & -1.2639 & 0.93174\\ 0.94621 & 0.40563 & -1.0332\\ -0.78129 & 0.78852 & 0.05297\end{bmatrix},$$

$$B_1=\begin{bmatrix}-0.35425 & -0.3446 & -0.23877\\ 0.043436 & -0.49913 & -0.30311\\ 0.062814 & -0.34911 & -0.20951\\ -0.45534 & -0.21019 & -0.16401\\ -0.088286 & -0.32693 & -0.2075\end{bmatrix},\quad B_2=\begin{bmatrix}0.23712 & -0.31609 & 0.11599\\ 0.35782 & 0.28332 & -0.32045\\ 0.089122 & 0.27873 & -0.21547\\ -0.028823 & -0.47285 & 0.31909\\ -0.84508 & 0.27845 & 0.1393\end{bmatrix},$$

$$C_1=\begin{bmatrix}9.0075 & 16.66 & 12.474\\ 18.039 & 18.224 & 15.765\\ 16.563 & 16.555 & 14.424\\ 15.379 & 24.087 & 18.169\end{bmatrix},\quad C_2=\begin{bmatrix}3.2791 & -3.694 & 3.9634\\ -5.5945 & -3.933 & -6.4169\\ 4.4899 & -2.6044 & 6.6608\\ 2.0442 & 1.6027 & 2.4247\end{bmatrix}.$$

We can easily see that the condition (14) is satisfied:

$$E_{33}=5.4498e-006,\ E_{43}=3.9681e-006,\ F_{11}=1.6941e-005,\ F_{41}=1.4966e-005,$$
$$E_{42}D_1^{-1}=F_{42}D_2^{-1}=-2.0595.$$

Thus, by Corollary 4 and choosing $Y_{i4}=0, i=1,2,3,4$, we can get

$$X=\begin{bmatrix}2.7946 & -1.6182 & 0.50824\\ 1.3163 & -0.90325 & 0.45661\\ 3.7457 & -1.9494 & 0.34305\\ 4.7665 & -2.4991 & 0.46507\end{bmatrix},\quad Y=\begin{bmatrix}8.8372 & 13.783 & 10.22 & 5.5065 & 12.291\\ 15.196 & 17.043 & 11.11 & 10.866 & 8.743\\ 11.586 & 14.259 & 10.404 & 8.7471 & 13.705\\ 16.631 & 18.127 & 12.513 & 12.894 & 14.328\end{bmatrix},$$

$$Z=\begin{bmatrix}-0.084378 & -1.3282 & 0.9686\\ 1.385 & 2.3561 & 3.6709\\ -0.91998 & -0.63325 & -3.3763\end{bmatrix}.$$

Also, we can figure out $\|C_1-(A_1X-YB_1)\|=4.1311e-004$, $\|C_2-(A_2Z-YB_2)\|=1.8800e-004$.

According to Corollary 5, we can get

$$X=\begin{bmatrix}5.7406 & 4.7506 & 4.6446\\ 4.1062 & 5.0946 & 4.3533\\ 5.5112 & 1.9198 & 2.8541\\ 7.1973 & 2.8184 & 3.9164\end{bmatrix},\quad Y=\begin{bmatrix}4.0847 & 8.9837 & 6.9075 & 1.6322 & 8.1986\\ 6.3885 & 7.8982 & 5.2555 & 4.2831 & 4.196\\ 5.3045 & 7.9134 & 5.8688 & 3.4672 & 7.2639\\ 8.2264 & 9.6558 & 6.7268 & 6.0972 & 7.4948\end{bmatrix},$$

$$Z=\begin{bmatrix}0.17363 & -1.4295 & 0.93668\\ 2.2413 & 2.185 & 3.4573\\ -1.7328 & -0.43177 & -3.199\end{bmatrix}.$$

Furthermore, we can figure out

$$\|C_1-(A_1X-YB_1)\|=5.9548e-004,\quad \|C_2-(A_2Z-YB_2)\|=1.9650e-004.$$

Acknowledgement

The author would like to express his special thanks to an anonymous referee for very valuable comments and suggestions that substantially improved the quality of this paper.

References

[1] J.K. Baksalary, R. Kala, The matrix equation $AX-YB=C$, Linear Algebra Appl. 25 (1979) 41–43.
[2] A. Ben-Israel, T.N.E. Greville, Generalized Inverses: Theory and Applications, second ed., Springer, New York, 2003.
[3] G.H. Golub, C.F. Van Loan, Matrix Computations, The Johns Hopkins University Press, Baltimore, MD, USA, 1983.
[4] S.G. Lee, Q.P. Vu, Simultaneous solutions of matrix equations and simultaneous equivalence of matrices, Linear Algebra Appl. 437 (2012) 2325–2339.
[5] Y.-H. Liu, Ranks of solutions of the linear matrix equation $AX+YB=C$, Comput. Math. Appl. 52 (2006) 861–872.
[6] C.C. Paige, M.A. Saunders, Towards a generalized singular value decomposition, SIAM J. Numer. Anal. 18 (1981) 398–405.
[7] G.W. Stewart, Computing the CS-decomposition of a partitioned orthogonal matrix, Numer. Math. 40 (1982) 297–306.
[8] Q.-W. Wang, Z.-H. He, Solvability conditions and general solution for mixed Sylvester equations, Automatica 49 (2013) 2713–2719.