Hindawi Publishing Corporation
Abstract and Applied Analysis
Volume 2013, Article ID 824641, 8 pages
http://dx.doi.org/10.1155/2013/824641

Research Article

A Note on the $\top$-Stein Matrix Equation

Chun-Yueh Chiang

Center for General Education, National Formosa University, Huwei 632, Taiwan

Correspondence should be addressed to Chun-Yueh Chiang; chiang@nfu.edu.tw

Received 16 May 2013; Revised 16 July 2013; Accepted 16 July 2013

Academic Editor: Masoud Hajarian

Copyright © 2013 Chun-Yueh Chiang. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This note is concerned with the linear matrix equation $X = AX^\top B + C$, where the operator $(\cdot)^\top$ denotes the transpose ($\top$) of a matrix. The first part of this paper sets forth the necessary and sufficient conditions for the unique solvability of the solution X. The second part of this paper aims to provide a comprehensive treatment of the relationship between the theory of the generalized eigenvalue problem and the theory of the linear matrix equation. The final part of this paper starts with a brief review of numerical methods for solving the linear matrix equation. In relation to the computed methods, knowledge of the residual is discussed. An expression related to the backward error of an approximate solution is obtained; it shows that a small backward error implies a small residual. Just like the discussion of linear matrix equations, perturbation bounds for solving the linear matrix equation are also proposed in this work.

1. Introduction

Our purpose of this work is to study the so-called $\top$-Stein matrix equation

$$X = AX^\top B + C, \tag{1}$$

where $A, B, C \in \mathbb{R}^{n\times n}$ are known matrices and $X \in \mathbb{R}^{n\times n}$ is an unknown matrix to be determined. Our interest in the $\top$-Stein equation originates from the study of completely integrable mechanical systems, that is, the analysis of the $\top$-Sylvester equation

$$AX + X^\top B = C, \tag{2}$$

where A, B, and C are matrices in $\mathbb{R}^{n\times n}$ [1, 2]. By means of the generalized inverses or QZ decomposition [3], the solvability conditions of (2) are studied in [1, 2, 4]. Suppose that the matrix pencil $A-\lambda B^\top$ is regular; that is, $aA+bB^\top$ is invertible for some scalars a and b. The $\top$-Sylvester equation (2) can be written as

$$\left(aA + bB^\top\right) X + X^\top \left(aB + bA^\top\right) = aC + bC^\top. \tag{3}$$

Premultiplying both sides of (3) by $(aA+bB^\top)^{-1}$, we have

$$X + UX^\top V = D, \tag{4}$$

where $U = (aA + bB^\top)^{-1}$, $V = aB + bA^\top$, and $D = (aA + bB^\top)^{-1}(aC + bC^\top)$. This is of the form (1). In other words, numerical approaches for solving (2) can be obtained by transforming (2) into the form of (1) and then applying numerical methods to (1) for the solution [4–6]. With this in mind, in this note we are interested in the study of $\top$-Stein matrix equation (1).

Our major purpose in this work can be divided into three parts. First, we determine necessary and sufficient conditions for the unique solvability of the solution to (1). In doing so, Zhou et al. [7] transform (1) to the standard Stein equation

$$W = AB^\top WA^\top B + AC^\top B + C \tag{5}$$

with respect to the unknown matrix $W \in \mathbb{R}^{m\times n}$ and give the following necessary condition:

$$\mu\nu \neq 1, \quad \forall \mu, \nu \in \sigma\left(A^\top B\right). \tag{6}$$

Here, $\sigma(A^\top B)$ is the set of all eigenvalues of $A^\top B$. Zhou and his coauthors show that if (5) has a unique solution, then (1) has a unique solution. However, a counterexample is provided in [7] to show that the relation (6) is only a necessary condition for the unique solvability of (1).

In [4, 8], the periodic QZ (PQZ) decomposition [3] is applied to consider the necessary and sufficient conditions of the unique solvability of (1); conditions given in [8] ignore the possibility of the existence of the unique solution, while 1 is a simple root of $\sigma(A^\top B)$. This condition is included in our subsequent discussion and the following remark is provided to support our observation.

Remark 1. Let $A = -1$ and let $B = 1$; that is, $\sigma(AB^\top) = \{-1\}$. It is clear that the scalar equation $X = -X^\top + C$ has a unique solution $X = C/2$. But condition (6) is not satisfied by choosing $\mu = \nu = -1$.

It can also be observed from Remark 1 that even if (1) is uniquely solvable, it does not imply that (5) (namely, $X = X + C - C$) is uniquely solvable. Conditions in [4, Equation(4.6)] provided those conditions for the unique solvability of the solution to (1) via a structured algorithm. In our work, through a complete analysis of square coefficient matrices in terms of the analysis of the spectra of the matrix $A^\top B$, the new approach to the condition of unique solvability of the $\top$-Stein equation (1) can be obtained.

Second, we present the invariant subspace method and, more generally, the deflating subspace method to solve the $\top$-Stein equation. Our methods are based on the analysis of the eigeninformations for a matrix pencil. We carry out a thorough discussion to address the various eigeninformation encountered in the subspace methods. These ideas can be implemented in algorithms easily.

Finally, we take full account of the error analysis of (1). Expressions and implications such as the residual, the backward error, and perturbation bounds are derived in this work. Note that for an approximate solution Y of (1), the backward error tells us how much the matrices A, B, and C must be perturbed. An important point found in Section 5 is that a small backward error indicates a small value for the residual $\mathscr{R} = Y - AY^\top B - C$, but the reverse is not usually true.

Beginning in Section 2, we formulate the necessary and sufficient conditions for the existence of the solution of (1) directly by means of the spectrum analysis. In Section 3 we provide a deflating subspace method for computing the solution of (1). Numerical methods for solving (1) and the related residual analysis are discussed in Section 4. The associated error analysis of (1) is given in Section 5 and concluding remarks are given in Section 6.

2. Solvability Conditions of the Matrix Equation (1)

In order to formalize our discussion, let $A \otimes B$ be the Kronecker product of matrices A and B, let I_n be the $n \times n$ identity matrix, and let $\|\cdot\|_F$ denote the Frobenius norm.

With the Kronecker product, (1) can be written as the enlarged linear system:

$$\left(I_{n^2} - \left(B^\top \otimes A\right)\mathscr{P}\right)\operatorname{vec}(X) = \operatorname{vec}(C), \tag{7}$$

where $\operatorname{vec}(X)$ stacks the columns of X into a column vector and $\mathscr{P}$ is the Kronecker permutation matrix [9] which maps $\operatorname{vec}(X)$ into $\operatorname{vec}(X^\top)$; that is,

$$\mathscr{P} = \sum_{1\le i,j\le n^2} e_j e_i^\top \otimes e_i e_j^\top, \tag{8}$$

where e_i denotes the ith column of the $n^2 \times n^2$ identity matrix I_{n^2}. Due to the specific structure of $\mathscr{P}$, it has been shown in [10, Corollary 4.3.10] that

$$\mathscr{P}^\top\left(B^\top \otimes A\right)\mathscr{P} = A \otimes B^\top. \tag{9}$$

It then follows that

$$\left(\left(B^\top \otimes A\right)\mathscr{P}\right)^2 = \left(B^\top \otimes A\right)\mathscr{P}\mathscr{P}^\top\left(A \otimes B^\top\right) = B^\top A \otimes AB^\top, \tag{10}$$

since $\mathscr{P}^2 = I_{n^2}$ and $\mathscr{P} = \mathscr{P}^\top$. Note that eigenvalues of matrices $A^\top B$ and $AB^\top$ are the same. By (10) and the property of the Kronecker product [11, Theorem 4.8], we know that

$$\begin{aligned}&\sigma\left(\left(\left(B^\top \otimes A\right)\mathscr{P}\right)^2\right)\\ &\quad = \left\{\lambda_i\lambda_j \mid \lambda_i, \lambda_j \in \sigma\left(A^\top B\right) = \{\lambda_1, \ldots, \lambda_n\},\ 1 \le i, j \le n\right\}.\end{aligned} \tag{11}$$

That is, the eigenvalues of $(B^\top \otimes A)\mathscr{P}$ are related to the square roots of the eigenvalues of $\sigma(A^\top B)$, but from (10), no more information can be used to decide the positivity or nonnegativity of the eigenvalues of $(B^\top \otimes A)\mathscr{P}$. A question immediately arises as to whether it is possible to obtain the explicit expression of the eigenvalues of $(B^\top \otimes A)\mathscr{P}$, provided the eigenvalues of $A^\top B$ are given. In the following two lemmas, we first review the periodic QZ decomposition for two matrices and then apply it to discuss the eigenvalues of $(B^\top \otimes A)\mathscr{P}$.

Lemma 2 (see [3]). *Let A and B be two matrices in $\mathbb{R}^{n\times n}$. Then, there exist unitary matrices $P, Q \in \mathbb{C}^{n\times n}$ such that $U_A := PAQ$ and $U_B := Q^H B^\top P^H$ are two upper triangular matrices.*

Lemma 3. *Let A and B be two matrices in $\mathbb{R}^{m\times n}$. Then*

(1) $(B^\top \otimes A)\mathscr{P} = (Q \otimes P^H)(U_A \otimes U_B)\mathscr{P}(Q^H \otimes P)$,

(2) $\sigma((B^\top \otimes A)\mathscr{P}) = \{\lambda_i, \pm\sqrt{\lambda_i\lambda_j} \mid \lambda_i, \lambda_j \in \sigma(A^\top B) = \{\lambda_1, \ldots, \lambda_n\},\ 1 \le i < j \le n\}$.

Here, $\sqrt{z}$ denotes the principal square root of a complex number z.

Proof. Part 1 follows immediately from Lemma 2 since $U_A = PAQ$ and $U_B = Q^H B^\top P^H$ for some unitary matrices P and Q; that is,

$$\begin{aligned}\left(B^\top \otimes A\right)\mathscr{P} &= \left(Q \otimes P^H\right)\left(U_B \otimes U_A\right)\left(P \otimes Q^H\right)\mathscr{P}\\ &= \left(Q \otimes P^H\right)\left(U_A \otimes U_B\right)\mathscr{P}\left(Q^H \otimes P\right).\end{aligned} \tag{12}$$

Let the diagonal entries of U_A and U_B be denoted by $\{a_{ii}\}$ and $\{b_{jj}\}$, respectively. Then, $U_A \otimes U_B$ is an upper triangular matrix with given diagonal entries, specified by a_{ii} and b_{jj}. After multiplying $U_A \otimes U_B$ by $\mathscr{P}$ from the right, the position of the entry $a_{ii}b_{jj}$ is changed to be in the $j+n(i-1)$ th row and the $i+n(j-1)$ th column of the matrix $(U_A \otimes U_B)\mathscr{P}$. They are then reshuffled by a sequence of permutation matrices to form a block upper triangular matrix with diagonal entries arranged in the following order:

$$\begin{aligned}&\left\{a_{11}b_{11}, \begin{bmatrix} 0 & a_{11}b_{22} \\ a_{22}b_{11} & 0 \end{bmatrix}, \ldots, \begin{bmatrix} 0 & a_{11}b_{nn} \\ a_{nn}b_{11} & 0 \end{bmatrix},\right. \\ &\quad a_{22}b_{22}, \begin{bmatrix} 0 & a_{22}b_{33} \\ a_{33}b_{22} & 0 \end{bmatrix}, \ldots, \begin{bmatrix} 0 & a_{nn}b_{22} \\ a_{22}b_{nn} & 0 \end{bmatrix}, \ldots, \\ &\quad \left.\begin{bmatrix} 0 & a_{n-1,n-1}b_{nn} \\ a_{nn}b_{n-1,n-1} & 0 \end{bmatrix}, a_{nn}b_{nn}\right\}.\end{aligned} \tag{13}$$

Note that the reshuffling process is not hard to see, when $n = 2$, $U_A = \left[\begin{smallmatrix} a_{11} & a_{12} \\ 0 & a_{22} \end{smallmatrix}\right]$, and $U_B = \left[\begin{smallmatrix} b_{11} & b_{12} \\ 0 & b_{22} \end{smallmatrix}\right]$, we have

$$(U_A \otimes U_B)\mathscr{P} = \begin{bmatrix} a_{11}b_{11} & a_{12}b_{11} & a_{11}b_{12} & a_{12}b_{12} \\ 0 & 0 & a_{11}b_{22} & a_{12}b_{22} \\ 0 & a_{22}b_{11} & 0 & a_{22}b_{12} \\ 0 & 0 & 0 & a_{22}b_{22} \end{bmatrix}. \tag{14}$$

However, it is conceptually simple and regular but operationally tedious to reorder $(U_A \otimes U_B)\mathscr{P}$ to show this result even for $n = 3$ and that will be left as an exercise.

By (13), it can be seen that

$$\begin{aligned}\sigma\left((B^\top \otimes A)\mathscr{P}\right) &= \left\{a_{ii}b_{ii}, \pm\sqrt{a_{ii}a_{jj}b_{ii}b_{jj}},\ 1 \le i,\ j \le n\right\} \\ &= \left\{\lambda_i, \pm\sqrt{\lambda_i\lambda_j},\ 1 \le i,\ j \le n\right\},\end{aligned} \tag{15}$$

where $\lambda_i = a_{ii}b_{ii} \in \sigma(A^\top B)$ for $1 \le i \le n$.

Before demonstrating the unique solvability conditions, we need to define that a subset $\Lambda = \{\lambda_1, \ldots, \lambda_n\}$ of complex numbers is said to be $\top$-*reciprocal free* if and only if whenever $i \ne j$, $\lambda_i \ne 1/\lambda_j$. This definition also regards 0 and ∞ as reciprocals of each other. Then, we have the following solvability conditions of (1). □

Theorem 4. *The $\top$-Stein matrix equation* (1) *is uniquely solvable if and only if the following conditions are satisfied:*

(1) *the set of $\sigma(A^\top B) \setminus \{-1\}$ is $\top$-reciprocal free;*

(2) *-1 can be an eigenvalue of the matrix $A^\top B$ but must be simple.*

Proof. From (7), we know that the $\top$-Stein matrix equation (1) is uniquely solvable if and only if

$$1 \notin \sigma\left((B^\top \otimes A)\mathscr{P}\right). \tag{16}$$

By Lemma 3, if $\lambda \in \sigma(A^\top B)$, then $1/\lambda \notin \sigma(A^\top B)$. Otherwise, $1 = \sqrt{\lambda \cdot (1/\lambda)} \in ((B^\top \otimes A)\mathscr{P})$. On the other hand, if $-1 \in \sigma(A^\top B)$ and -1 is not a simple eigenvalue, then $1 \in \sigma((B^\top \otimes A)\mathscr{P})$. This verifies (16) and the proof of the theorem is complete. □

It is worthy noting that the condition (1) of Theorem 4 is contained in the condition (6) (also appear in [7, Theorem 1]), which is the necessary and sufficient conditions for the solvability of the solution to the Stein equation (5). However, as mentioned before in Remark 1 or [7, Example 2], condition (1) is just a necessary condition for unique solvability of the solution to (1). The $\top$-Stein matrix equation (1) is uniquely solvable provided that both conditions (1) and (2) of Theorem 4 are satisfying.

3. The Connection between Deflating Subspace and (1)

The relationship between solution of matrix equations and the matrix eigenvalue problems has been widely studied in many applications. It is famous that solution of Riccati and polynomial matrix equations can be found by computing invariant subspaces of matrices and deflating subspaces of matrix pencils [12]. This reality leads us to find some algorithms for computing solution of (1) based on the numerical computation of invariant or deflating subspaces.

Given a pair of $n \times n$ matrices A and B, recall that the function $A-\lambda B$ in the variable λ is said to be the matrix pencil related to the pair (A, B). For a k-dimensional subspace $\mathscr{X} \in \mathbb{C}^n$ is called a deflating subspace for the pencil $A - \lambda B$ if there exists a k-dimensional subspace $\mathscr{Y} \in \mathbb{C}^n$ such that

$$A\mathscr{X} \subseteq \mathscr{Y}, \quad B\mathscr{X} \subseteq \mathscr{Y}; \tag{17}$$

that is,

$$AX = YT_1, \quad BX = YT_2, \tag{18}$$

where $X, Y \in \mathbb{C}^{n\times k}$ are two full rank matrices whose columns span the spaces $\mathscr{X}$ and $\mathscr{Y}$, respectively, and matrices $T_1, T_2 \in \mathbb{C}^{k\times k}$. In particular, if in (18), $X = Y$ and $B = T_2 = I$ for an $n \times n$ identity matrix I, then we have the simplified formula

$$AX = XT_1. \tag{19}$$

Here, the space $\mathscr{X}$ spanned by the columns of the matrix X is called an invariant subspace for A and satisfies

$$A\mathscr{X} \subseteq \mathscr{X}. \tag{20}$$

One strategy to analyze the eigeninformation is to transform one matrix pencil to its simplified and equivalent form. That is, two matrix pencils $A - \lambda B$ and $\widetilde{A} - \lambda\widetilde{B}$ are said to be equivalent if and only if there exist two nonsingular matrices P and Q such that

$$P(A - \lambda B)Q = \widetilde{A} - \lambda\widetilde{B}. \tag{21}$$

In the subsequent discussion, we will use the notion $\sim$ to describe this equivalence relation; that is, $A - \lambda B \sim \widetilde{A} - \lambda\widetilde{B}$.

Our task in this section is to identify eigenvectors of problem (18) and then associate these eigenvectors (left and right) with the solution of (1). We begin this analysis by studying the eigeninformation of two matrices A and B, where $A - \lambda B$ is a regular matrix pencil.

Note that for the ordinary eigenvalue problem, if the eigenvalues are different, then the eigenvectors are linearly independent. This property is also true for every regular matrix pencil and is demonstrated as follows. For a detailed proof, the reader is referred to [13, Theorem 7.3] and [14, Theorem 4.2].

Theorem 5. *Given a pair of $n \times n$ matrix A and B, if the matrix pencil $A - \lambda B$ is regular, then its Jordan chains corresponding to all finite and infinite eigenvalues carry the full spectral information about the matrix pencil and consist of n linearly independent vectors.*

Lemma 6. *Let $A - \lambda B \in \mathbb{C}^{n\times n}$ be a regular matrix pencil. Assume that matrices $X_i, Y_i \in \mathbb{C}^{n\times n_i}$, $i = 1, 2$, are full rank and satisfy the following equations:*

$$AX_i = Y_i R_i, \tag{22a}$$

$$BX_i = Y_i S_i, \tag{22b}$$

where R_i and S_i, $i = 1, 2$, are square matrices of size $n_i \times n_i$. Then

(i) *$R_i - \lambda S_i \in \mathbb{C}^{n_i\times n_i}$ are regular matrix pencils for $i = 1, 2$,*

(ii) *if $\sigma(R_1 - \lambda S_1) \cap \sigma(R_2 - \lambda S_2) = \phi$, then the matrix $\begin{bmatrix} X_1 & X_2 \end{bmatrix} \in \mathbb{R}^{n\times(n_1+n_2)}$ is full rank.*

We also need the following useful lemma.

Lemma 7. *Given two regular matrix pencils $A_i - \lambda B_i \in \mathbb{C}^{n_i\times n_i}$, $1 \le i \le 2$, consider the following equations with respect to $U, V \in \mathbb{C}^{n_1\times n_2}$*

$$A_1 U = V A_2, \tag{23a}$$

$$B_1 U = V B_2. \tag{23b}$$

Then, if $\sigma(A_1 - \lambda B_1) \cap \sigma(A_2 - \lambda B_2) = \phi$, (23a) has the unique solution $U = V = 0$.

Proof. For $n_2 = 1$, we get

$$\begin{aligned} A_1 u &= a_2 v, \\ B_1 u &= b_2 v, \end{aligned} \tag{24}$$

where $a_2, b_2 \in \mathbb{C}$, $u, v \in \mathbb{C}^{n_1\times 1}$. We may without loss of generality assume that $b_2 \neq 0$; then $A_1 u = (a_2/b_2) B_1 u$ and thus $u = v = 0$. Now, for any $n_2 > 1$, consider the generalized Schur decomposition of $A_2 - \lambda B_2$. We can assume that $A_2 = [a_{ij}]$ and $B_2 = [b_{ij}]$ are upper triangular matrices (i.e., $a_{ij} = b_{ij} = 0, 1 \le j < i \le n_2$). Denote that the ith columns of U and V are u_i and v_i, respectively. Thus,

$$A_1 u_i = \sum_{k=1}^{i} a_{ki} v_k, \tag{25a}$$

$$B_1 u_i = \sum_{k=1}^{i} b_{ki} v_k, \tag{25b}$$

for $i = 1, 2, \ldots, n_2$.

If $i = 1$, we obtained $u_1 = v_1 = 0$ from the above discussion. Given an integer i such that $1 \le i < n_2$ and assume that $u_k = v_k = 0$ for $1 \le k \le i$, we claim $u_{i+1} = v_{i+1} = 0$, indeed, from (25a) and (25b), we have

$$\begin{aligned} A_1 u_{i+1} &= a_{i+1,i+1} v_{i+1}, \\ B_1 u_{i+1} &= b_{i+1,i+1} v_{i+1}. \end{aligned} \tag{26}$$

Again, the case $k = i + 1$ is following the special case $n_2 = 1$. By using mathematical induction, we prove this lemma. □

Corollary 8. *Given $A \in \mathbb{C}^{n\times n}$ and $\Lambda \in \mathbb{C}^{k\times k}$, if $\sigma(A) \cap \sigma(\Lambda) = \phi$, then the equation with respect to $U \in \mathbb{C}^{n\times k}$*

$$AU = U\Lambda \tag{27}$$

has the unique solution $U = 0$.

Now we have enough tools to analyze the solution of (1) associated with some deflating spaces. We first establish an important matrix pencil; let the matrix pencil $\mathscr{M} - \lambda\mathscr{L}$ be defined as

$$\mathscr{M} - \lambda\mathscr{L} := \begin{bmatrix} BA^\top & 0 \\ -CA^\top & I_n \end{bmatrix} - \lambda \begin{bmatrix} I_n & 0 \\ AC^\top & AB^\top \end{bmatrix} \in \mathbb{R}^{2n\times 2n}; \tag{28}$$

it is clear that

$$\sigma(\mathscr{M} - \lambda\mathscr{L}) = \sigma(BA^\top) \cup \sigma(I_n - \lambda AB^\top); \tag{29}$$

a direct calculation shows that X is a solution of the (1) if and only if

$$\begin{aligned} \mathscr{M} \begin{bmatrix} I_n \\ XA^\top \end{bmatrix} &= \begin{bmatrix} I_n \\ AX^\top \end{bmatrix} BA^\top, \\ \mathscr{L} \begin{bmatrix} I_n \\ XA^\top \end{bmatrix} &= \begin{bmatrix} I_n \\ AX^\top \end{bmatrix} \end{aligned} \tag{30}$$

or if and only if its dual form is

$$\begin{aligned} \begin{bmatrix} -AX^\top & I_n \end{bmatrix} \mathscr{M} &= \begin{bmatrix} -XA^\top & I_n \end{bmatrix}, \\ \begin{bmatrix} -AX^\top & I_n \end{bmatrix} \mathscr{L} &= AB^\top \begin{bmatrix} -XA^\top & I_n \end{bmatrix}. \end{aligned} \tag{31}$$

Armed with the property given in Theorem 5 and Lemma 7, we can now tackle the problem of determining how the deflating subspace is related to the solution of (1).

Theorem 9. *Let A, B, and $C \in \mathbb{R}^{n\times n}$ be given in (1) and let*

$$\mathscr{M} \begin{bmatrix} U_1 \\ V_1 \end{bmatrix} = \begin{bmatrix} U_2 \\ V_2 \end{bmatrix} T_1, \tag{32a}$$

$$\mathscr{L} \begin{bmatrix} U_1 \\ V_1 \end{bmatrix} = \begin{bmatrix} U_2 \\ V_2 \end{bmatrix} T_2, \tag{32b}$$

where $\begin{bmatrix} U_i \\ V_i \end{bmatrix}$ is full rank, $i = 1, 2$. Assume that the set of $\sigma(BA^\top)$ is $\top$-reciprocal free. Then, one has

(1) $U_1 = U_2 = 0$ *if* $\sigma(T_1 - \lambda T_2) = \sigma(I_n - \lambda AB^\top)$,

(2) U_1 *and* U_2 *are nonsingular if* $T_1 - \lambda T_2 \sim BA^\top - \lambda I_n$. *Moreover, if* A *is nonsingular, then* $X = V_1 U_1^{-1} A^{-\top} = U_2^{-\top} V_2^\top A^{-\top}$ *is the unique solution of* (1).

Proof. From (32a) and (32b) we get

$$BA^\top U_1 = U_2 T_1, \tag{33a}$$

$$-CA^\top U_1 + V_1 = V_2 T_1, \tag{33b}$$

$$U_1 = U_2 T_2, \tag{33c}$$

$$AC^\top U_1 + AB^\top V_1 = V_2 T_2. \tag{33d}$$

(i) It follows from (33a) and (33c) that since $\sigma(BA^\top - \lambda I_n) \cap \sigma(T_1 - \lambda T_2) = \phi$, we have $U_1 = U_2 = 0$ by Lemma 7.

(ii) It can be seen that there exist two nonsingular matrices U and V such that

$$\begin{aligned} \mathscr{M}\begin{bmatrix} 0 \\ U \end{bmatrix} &= \begin{bmatrix} 0 \\ V \end{bmatrix} T_2, \\ \mathscr{L}\begin{bmatrix} 0 \\ U \end{bmatrix} &= \begin{bmatrix} 0 \\ V \end{bmatrix} T_1. \end{aligned} \tag{34}$$

Hence, together with (32a) and (32b) we have

$$\begin{aligned} \mathscr{M}\begin{bmatrix} 0 & U_1 \\ U & V_1 \end{bmatrix} &= \begin{bmatrix} 0 & U_2 \\ V & V_2 \end{bmatrix}\begin{bmatrix} T_2 & 0 \\ 0 & T_1 \end{bmatrix}, \\ \mathscr{L}\begin{bmatrix} 0 & U_1 \\ U & V_1 \end{bmatrix} &= \begin{bmatrix} 0 & U_2 \\ V & V_2 \end{bmatrix}\begin{bmatrix} T_1 & 0 \\ 0 & T_2 \end{bmatrix}. \end{aligned} \tag{35}$$

Since $\sigma(\mathscr{M} - \lambda\mathscr{L}) = \sigma(BA^\top - \lambda I_n) \cup \sigma(I_n - \lambda AB^\top)$ and $\sigma(BA^\top - \lambda I_n) \cap \sigma(I_n - \lambda AB^\top) = \phi$, by Theorem 5 and Lemma 6, the matrix $\left[\begin{smallmatrix} 0 & U_1 \\ U & V_1 \end{smallmatrix}\right]$ is nonsingular. Together with (33c), U_1 and U_2 are nonsingular.

Let $X_i = V_i U_i^{-1}$, $i = 1, 2$; then form (33b) and (33d)

$$\begin{aligned} AC^\top + AB^\top X_1 &= V_2 T_2 U_1^{-1} = X_2, \\ -CA^\top + X_1 &= V_2 T_1 U_1^{-1} = X_2 BA^\top, \end{aligned} \tag{36}$$

or

$$\begin{aligned} AC^\top + AB^\top X_1 &= X_2, \\ AC^\top + AB^\top X_2^\top &= X_1^\top. \end{aligned} \tag{37}$$

Since the set of $\sigma(AB^\top) = \sigma(BA^\top)$ is $\top$-reciprocal free, together with

$$X_1^\top - X_2 - AB^\top\left(X_1^\top - X_2\right)^\top = 0, \tag{38}$$

we get $X_1 = X_2^\top$. If A is nonsingular, it is easy to verify that two matrices $X_1 A^{-\top}$ and $X_2^\top A^{-\top}$ are both satisfying $\top$-Stein equation (1). The proof of part (ii) is complete. □

Remark 10. (1) It is easily seen that $\left[\begin{smallmatrix} I_n \\ XA^\top \end{smallmatrix}\right]$ and $\left[\begin{smallmatrix} U_1 \\ V_2 \end{smallmatrix}\right]$ both span the unique deflating subspace of $\mathscr{M} - \lambda\mathscr{L}$ corresponding to the set of $\sigma(BA^\top)$. Otherwise, in part (ii) we know that T_2 is nonsingular. We then are able to transform the formulae defined in (32a) and (32b) into the generalized eigenvalue problem as follows:

$$\mathscr{M}\begin{bmatrix} U_1 \\ V_1 \end{bmatrix} = \mathscr{L}\begin{bmatrix} U_1 \\ V_1 \end{bmatrix} BA^\top. \tag{39}$$

That is, some numerical methods for the computation of the eigenspace of $\mathscr{M} - \lambda\mathscr{L}$ corresponding to the set of $\sigma(BA^\top)$ can be designed and solved (1).

(2) Since the transport of the unique solution X of (1) is equal to the unique solution Y of the following matrix equation

$$Y = B^\top Y A^\top + C^\top, \tag{40}$$

analogous to the consequences of Theorem 9, the similar results can be obtained with respect to (40) if B is nonsingular. However, we point out that (1) can be solved by computing deflating subspaces of other matrix pencils. For instance we let

$$\mathscr{M}_1 - \lambda\mathscr{L}_1 := \begin{bmatrix} A^\top B & 0 \\ -C - AC^\top B & I_n \end{bmatrix} - \lambda \begin{bmatrix} I_n & 0 \\ 0 & AB^\top \end{bmatrix}. \tag{41}$$

Assume that the set of $\sigma(BA^\top)$ is $\top$-reciprocal free; it can be shown that $\mathscr{M}_1 \left[\begin{smallmatrix} I_n \\ X \end{smallmatrix}\right] = \mathscr{L}_1 \left[\begin{smallmatrix} I_n \\ X \end{smallmatrix}\right] A^\top B$ and it has similar results to the conclusion of Theorem 9. The unique solution X of (1) can be found by computing deflating subspaces of the matrix pencil $\mathscr{M}_1 - \lambda\mathscr{L}_1$ without the assumption of the singularity of A and B.

4. Computational Methods for Solving (1)

Numerical methods for solving (1) have received great attention in theory and in practice and can be found in [5, 6] for the Krylov subspace methods and in [15–17] for the Smith-type iterative methods. In particular, Smith-type iterative methods are only workable in the case $\rho(AB^\top) < 1$, where $\rho(AB^\top)$ denotes the spectral radius of $AB^\top$. In the recent years, a structure algorithm has been studied for (1) [4] via PQZ decomposition, which consists of transforming into Schur form by a PQZ decomposition and then solving the resulting triangular system by way of back-substitution. In this section, we revisit these numerical methods and point out the advantages and drawbacks of all algorithms.

4.1. Krylov Subspace Methods. Since the $\top$-Stein equation is essentially a linear system (7), we certainly can use the Krylov subspace methods to solve (7). See, for example, [5, 6] and the reference cited therein. The general idea for applying the Krylov subspace methods is defining the $\star$-Stein operator $\mathscr{T}$ as $\mathscr{T} : X \to X - AX^\top B$ and its adjoint liner operator $\mathscr{T}$ as $\mathscr{T}^* : Y \to Y - BY^\top A$ such that $\langle \mathscr{T}(X), Y\rangle = \langle X, \mathscr{T}^*(Y)\rangle$. Here, $X, Y \in \mathbb{R}^{m\times n}$ and the notion $\langle\cdot,\cdot\rangle$ is denoted as the Frobenius inner product. Then, the iterative method based on the Krylov subspaces for (1) is as follows.

(i) The conjugate gradient (CG) method [5]:

$$X_{k+1} = X_k + \frac{\|R_k\|^2}{\|P_k\|^2} P_k,$$
$$R_{k+1} = C - \mathscr{T}(X_{k+1}) = R_k - \frac{\|R_k\|^2}{\|P_k\|^2} \mathscr{T}(P_k), \tag{42}$$
$$D_{k+1} = \mathscr{T}^*(R_{k+1}) + \frac{\|R_{k+1}\|^2}{\|R_k\|^2} D_k,$$

with an initial matrix X_0 and the corresponding initial conditions

$$R_0 = C - \mathscr{T}(X_0), \qquad D_0 = \mathscr{T}^*(R_0). \tag{43}$$

Note that when the solvability conditions of Theorem 4 are met, the CG method is guaranteed to converge in a finite number of iterations for any initial matrix X_0.

4.2. The Bartels-Stewart-Like Algorithm [18]. In this section, we focus on the discussion of the Bartels-Stewart algorithm, which is known to be a numerical stable algorithm, to solve $\top$-Stein equations. This method is to solve (1) by means of the PQZ decomposition [18]. Its approach has been discussed in [4] and can be summarized as follows. From Lemma 3, we know that there exist two unitary matrices P and Q (see [3] for the computation procedure) such that

$$PX\overline{Q} - PAQ \cdot Q^H X^\top P^\top \cdot \overline{P}B\overline{Q} = PC\overline{Q}. \tag{44}$$

With $\widehat{A} = PAQ$ and $\widehat{B}^\top = Q^H B^\top P^H$ being uppertriangular, the transformed equation looks like

$$\begin{bmatrix} \widehat{X}_{11} & \widehat{x}_{12} \\ \widehat{x}_{21} & \widehat{x}_{22} \end{bmatrix} - \begin{bmatrix} \widehat{A}_{11} & \widehat{a}_{12} \\ 0 & \widehat{a}_{22} \end{bmatrix} \begin{bmatrix} \widehat{X}_{11}^\top & \widehat{x}_{21}^\top \\ \widehat{x}_{12}^\top & \widehat{x}_{22}^\top \end{bmatrix} \begin{bmatrix} \widehat{B}_{11} & 0 \\ \widehat{b}_{21} & \widehat{b}_{22} \end{bmatrix} = \begin{bmatrix} \widehat{C}_{11} & \widehat{c}_{12} \\ \widehat{c}_{21} & \widehat{c}_{22} \end{bmatrix} \tag{45}$$

with $\widehat{X} = \begin{bmatrix} \widehat{x}_{11} & \widehat{x}_{12} \\ \widehat{x}_{21} & \widehat{x}_{22} \end{bmatrix}$. We then have

$$\widehat{x}_{22} - \widehat{a}_{22}\widehat{x}_{22}^\top \widehat{b}_{22} = \widehat{c}_{22}, \tag{46}$$
$$\widehat{x}_{21} - \widehat{a}_{22}\widehat{x}_{12}^\top \widehat{B}_{11} = \widehat{c}_{21} + \widehat{a}_{22}\widehat{x}_{22}^\top \widehat{b}_{21}, \tag{47}$$
$$\widehat{x}_{12} - \widehat{A}_{11}\widehat{x}_{21}^\top \widehat{b}_{22} = \widehat{c}_{12} + \widehat{a}_{12}\widehat{x}_{22}^\top \widehat{b}_{22}, \tag{48}$$
$$\widehat{X}_{11} - \widehat{A}_{11}\widehat{X}_{11}^\top \widehat{B}_{11} = \widehat{C}_{11} + \widehat{a}_{12}\widehat{x}_{12}^\top \widehat{B}_{11} + \widehat{A}_{11}\widehat{x}_{21}^\top \widehat{b}_{21} + \widehat{a}_{12}\widehat{x}_{22}^\top \widehat{b}_{21}. \tag{49}$$

Thus, the Bartels-Stewart algorithm can easily be constructed by first solving $\widehat{x}_{22}$ from (46), using $\widehat{x}_{22}$ to obtain $\widehat{x}_{12}$ and $\widehat{x}_{21}$ from (47) and (48), and then repeating the same discussion as (46)–(48) by taking advantage of the property of $\widehat{A}_{11}$ and $\widehat{B}_{11}$ being lower triangular matrices from (49).

4.3. Smith-Type Iterative Methods. Recently, a class method referred to as the Smith accelerative iteration has been studied for solving the Sylvester matrix equation [15]. The Smith accelerative iteration has attracted much interests in many papers (see [7, 17] and the references therein) because of its nice numerical behavior, a quadratic convergence rate, low computational cost, and high numerical reliability, despite the lack of a rigorous error analysis. Since the Sylvester matrix equation can be transformed into the Stein matrix equation with a suitable transformation, Zhou et al. proposed Smith-type iterative methods (including the Smith accelerative iteration, Smith (ℓ) iteration, and r-Smith iteration) for solving the Stein matrix equation [17] and applying Smith-type iterative methods to (1) [7]. In this section, we try to explain the Smith accelerative iteration based on the invariant subspace method and summarize the recent results from [7].

Originally, the Smith-type iterative methods are developed to solve the standard Stein equation

$$X = \mathscr{A}X\mathscr{B} + \mathscr{C}, \quad \mathscr{A}, \mathscr{B}, \mathscr{C} \in \mathbb{R}^{n\times n}. \tag{50}$$

As mentioned before, the unknown X is highly related to the generalized eigenspace problems

$$\begin{bmatrix} \mathscr{B} & 0 \\ -\mathscr{C} & I \end{bmatrix} \begin{bmatrix} I \\ X \end{bmatrix} = \begin{bmatrix} I & 0 \\ 0 & \mathscr{A} \end{bmatrix} \begin{bmatrix} I \\ X \end{bmatrix} \mathscr{B}, \tag{51a}$$

or

$$\mathscr{A} \begin{bmatrix} X & I \end{bmatrix} \begin{bmatrix} \mathscr{B} & 0 \\ 0 & I \end{bmatrix} = \begin{bmatrix} X & I \end{bmatrix} \begin{bmatrix} I & 0 \\ -\mathscr{C} & \mathscr{A} \end{bmatrix}. \tag{51b}$$

Premultiplying (51a) by the matrix $\left[\begin{smallmatrix} \mathscr{B} & 0 \\ -\mathscr{A}\mathscr{C} & I \end{smallmatrix}\right]$ and postmultiplying (51b) by the matrix $\left[\begin{smallmatrix} I & 0 \\ -\mathscr{C}\mathscr{B} & \mathscr{A} \end{smallmatrix}\right]$, we get

$$\begin{aligned} \begin{bmatrix} \mathscr{B}^2 & 0 \\ -\mathscr{C} - \mathscr{A}\mathscr{C}\mathscr{B} & I_n \end{bmatrix} \begin{bmatrix} I_n \\ X \end{bmatrix} &= \begin{bmatrix} I_n & 0 \\ 0 & \mathscr{A}^2 \end{bmatrix} \begin{bmatrix} I_n \\ X \end{bmatrix} \mathscr{B}^2, \\ \mathscr{A}^2 \begin{bmatrix} X & I_n \end{bmatrix} \begin{bmatrix} \mathscr{B}^2 & 0 \\ 0 & I_n \end{bmatrix} &= \begin{bmatrix} X & I_n \end{bmatrix} \begin{bmatrix} I_n & 0 \\ -\mathscr{C} - \mathscr{A}\mathscr{C}\mathscr{B} & \mathscr{A}^2 \end{bmatrix}. \end{aligned} \tag{52}$$

Then, for any positive integer $k > 0$, we obtain

$$\begin{aligned} \begin{bmatrix} \mathscr{B}^{2^{k-1}} & 0 \\ -C_k & I_n \end{bmatrix} \begin{bmatrix} I_n \\ X \end{bmatrix} &= \begin{bmatrix} I_n & 0 \\ 0 & \mathscr{A}^{2^{k-1}} \end{bmatrix} \begin{bmatrix} I_n \\ X \end{bmatrix} \mathscr{B}^{2^{k-1}}, \\ \mathscr{A}^{2^{k-1}} \begin{bmatrix} X & I_n \end{bmatrix} \begin{bmatrix} \mathscr{B}^{2^{k-1}} & 0 \\ 0 & I_n \end{bmatrix} &= \begin{bmatrix} X & I_n \end{bmatrix} \begin{bmatrix} I_n & 0 \\ -C_k & \mathscr{A}^{2^{k-1}} \end{bmatrix}, \end{aligned} \tag{53}$$

where the sequence $\{C_k\}$ is defined by

$$C_k = C_{k-1} + \mathscr{A}^{2^{k-1}} C_{k-1} \mathscr{B}^{2^{k-1}}, \quad k \geq 1, \tag{54a}$$
$$C_0 = \mathscr{C}. \tag{54b}$$

The explicit expression of C_k is given as follows:

$$C_k = \sum_{i=1}^{2^k-1} \mathscr{A}^i \mathscr{C} \mathscr{B}^i. \tag{55}$$

Under the condition $\rho(\mathscr{A})\rho(\mathscr{B}) < 1$, it is easy to see that $\{C_k\}$ is convergent, and

$$\limsup_{k\to\infty} \sqrt[2^k]{\|X - C_k\|} \le \rho(\mathscr{A})\,\rho(\mathscr{B}); \tag{56}$$

that is, C_k converges quadratically to X as $k \to \infty$. This iterative method (54a) and (54b) is called the Smith accelerative iteration [15]. In recent years, some modified iterative methods are so-called Smith-type iteration, which are based on Smith iteration and improve its speed of convergence. See, for example, [17] and the references cited therein.

Since the condition $\rho(\mathscr{A})\rho(\mathscr{B}) < 1$ implies that the assumptions of Theorem 4 hold, (1) is equivalent to (5). We can apply the Smith iteration to the (1) with the substitution $(\mathscr{A}, \mathscr{B}, \mathscr{C}) = (AB^\top, A^\top B, C + AC^\top B)$. One possible drawback of the Smith-type iterative methods is that they cannot always handle the case when there exist eigenvalues $\lambda, \mu \in \sigma(A^\top B)$ such that $\lambda\mu = -1$ even if the unique solution X exists. Based on the solvable conditions given in this work, it is possible to develop a specific technique working on the particular case and it is a subject currently under investigation.

5. Error Analysis

Error analysis is a way for testing the stability of a numerical algorithm and evaluating the accuracy of an approximated solution. In the subsequent discussion, we want to consider the backward error and perturbation bounds for solving (1).

As indicated in (44), matrices $\widehat{A}$ and $\widehat{B}^\top$ are both upper-triangular. We can then apply the error analysis for triangular linear systems in [19, Section 3.1] and [20] to obtain

$$\left\|\widehat{C} - \left(\widehat{X} - \widehat{A}\widehat{X}^\top\widehat{B}\right)\right\|_F \le c_{m,n}\mathbf{u}\left(1 + \left\|\widehat{A}\right\|_F\left\|\widehat{B}\right\|_F\right)\left\|\widehat{X}\right\|_F, \tag{57}$$

where $c_{m,n}$ is a content depending on the dimensions m and n and $\mathbf{u}$ is the unit roundoff. Since the PQZ decomposition is a stable process, it is true that

$$\left\|C - (X - AX^\top B)\right\|_F \le c'_{m,n}\mathbf{u}\left(1 + \|A\|_F\|B\|_F\right)\|X\|_F; \tag{58}$$

with a modest multiple $c'_{m,n}$.

Note that the inequality of the form (58) can serve as a stopping criterion for terminating iterations generated from the Krylov subspace methods [5, 6] and Smith-type iterative methods [15–17]. In what follows, we will derive the error associated with numerical algorithms, following the development in [20, 21].

5.1. Backward Error. Like the discussion of the ordinary Sylvester equations [20], the normwise backward error of an approximate solution Y of (1) is defined by

$$\begin{aligned}\eta(Y) \equiv \min\{\epsilon : Y = (C + \delta C) + (A + \delta A)\,Y^\top\,(B + \delta B),\\ \|\delta A\|_F \le \epsilon\alpha, \|\delta B\|_F \le \epsilon\beta, \|\delta C\|_F \le \epsilon\gamma\},\end{aligned} \tag{59}$$

where $\alpha \equiv \|A\|_F$, $\beta \equiv \|B\|_F$, and $\gamma \equiv \|C\|_F$. Let $\mathscr{R} \equiv \delta C + \delta AY^\top B + AY^\top\delta B + \delta AY^\top\delta B$, which implies that $\mathscr{R} = Y - AY^\top B - C$. It can be seen that the residual $\mathscr{R}$ satisfies

$$\|\mathscr{R}\|_F \le \eta(Y)\left(\gamma + \|Y\|_F\alpha\beta\left(2 + \eta(Y)\right)\right). \tag{60}$$

From (60), we know that a small backward error indeed implies a small relative residual $\mathscr{R}$. Since the coefficient matrices in (1) include nonlinearity, it appears to be an open problem to obtain the theoretical backward error with respect to the residual. Again, similar to the Sylvester equation discussed in [20, Section 16.2], the conditions under which a $\top$-Stein equation has a well-conditioned solution remain unknown.

5.2. Perturbation Bounds. Consider the perturbed equation

$$X + \delta X = (A + \delta A)\,(X + \delta X)^\top\,(B + \delta B) + (C + \delta C). \tag{61}$$

Let $S(X) = X - AX^\top B$ be the corresponding $\top$-Stein operator. We then have $S(\delta X) = \delta C + A(X+\delta X)^\top\delta B + \delta A(X+\delta X)^\top(B+\delta B)$. With the application of norm, it follows that

$$\|\delta X\|_F \le \left\|S^{-1}\right\|_F\left\{\|\delta C\|_F + \|\delta S\|_F\left(\|X\|_F + \|\delta X\|_F\right)\right\}, \tag{62}$$

where $\|\delta S\|_F \equiv \|A\|_F\|\delta B\|_F + \|\delta A\|_F(\|B\|_F + \|\delta B\|_F)$. When $\|\delta S\|_F$ is small enough so that $1 \ge \|S^{-1}\|_F \cdot \|\delta S\|_F$, we can rearrange the previous result to be

$$\frac{\|\delta X\|_F}{\|X\|_F} \le \frac{\left\|S^{-1}\right\|_F}{1 - \left\|S^{-1}\right\|_F \cdot \|\delta S\|_F}\left(\frac{\|\delta C\|_F}{\|X\|_F} + \|\delta S\|_F\right). \tag{63}$$

With $\|C\|_F = \|S(X)\|_F \le \|S\|_F \cdot \|X\|_F$ and the condition number $\kappa(S) \equiv \|S\|_F \cdot \|S^{-1}\|_F$, we arrive at the standard perturbation result

$$\frac{\|\delta X\|_F}{\|X\|_F} \le \frac{\kappa(S)}{1 - \kappa(S)\cdot\|\delta S\|_F/\|S\|_F}\left(\frac{\|\delta C\|_F}{\|C\|_F} + \frac{\|\delta S\|_F}{\|S\|_F}\right). \tag{64}$$

Thus the relative error in X is controlled by those in A, B, and C, magnified by the condition number $\kappa(S)$.

On the other hand, we can also drop the high-order terms in the perturbation to obtain

$$\delta X - A\delta X^\top B = AX^\top\delta B + \delta AX^\top B + \delta C. \tag{65}$$

We then rewrite the system in terms of

$$\mathscr{Q}\,\mathrm{vec}(\delta X) = \left[\left(X^\top B\right)^\top \otimes I_m I_n \otimes \left(AX^\top\right) I_{mn}\right]\begin{bmatrix}\mathrm{vec}(\delta A)\\ \mathrm{vec}(\delta B)\\ \mathrm{vec}(\delta C)\end{bmatrix}, \tag{66}$$

where $\mathscr{Q} = I_{mn} - (B^\top \otimes A)\mathscr{P}$. Let $\zeta = \max\{\|\delta A\|_F/\|A\|_F, \|\delta B\|_F/\|B\|_F, \|\delta C\|_F/\|C\|_F\}$. It can be shown that

$$\frac{\|\delta X\|_F}{\|X\|_F} \le \sqrt{3}\Psi\zeta, \tag{67}$$

where $\Psi = \|\mathscr{Q}^{-1}[\alpha(X^\top B)^\top \otimes I_m\ \beta I_n \otimes (AX^\top)\gamma I_{mn}]\|_2/\|X\|_F$.

A possible disadvantage of the perturbation bound (67), which ignores the consideration of the underlying structure of the problem, is that it overestimates the effect of the perturbation on the data. But this "universal" perturbation bound is accessible to any given matrices A, B, and C of (1).

Unlike the perturbation bound (67), it is desirable to obtain a posteriori error bound by assuming $\delta A = \delta B = 0$ and $\delta C = \widehat{X} - A\widehat{X}^\top B - C$ in (61). This assumption gives rise to

$$\frac{\|\delta X\|_F}{\|X\|_F} \leq \frac{\left\|P^{-1}\right\|_2 \|R\|_F}{\|X\|_F}. \tag{68}$$

It is true that while doing numerical computation, this bound given in (68) provides a simpler way for estimating the error of the solution of (1).

6. Conclusion

In this note, we propose a novel approach to the necessary and sufficient conditions for the unique solvability of the solution X of the $\top$-Stein equation for square coefficient matrices in terms of the analysis of the spectra $\sigma(A^\top B)$. Solvability conditions have been derived and algorithms have been proposed in [4, 8] by using PQZ decomposition. On the other hand, one common procedure to solve the Stein-type equations is by means of the invariant subspace method. We believe that our discussion is the first which implements the techniques of the deflating subspace for solving $\top$-Stein matrix equation and might also give rise to the possibility of developing an advanced and effective solver in the future. Also, we obtain the theoretical residual analysis, backward error analysis, and perturbation bounds for measuring accurately the error in the computed solution of (1).

Acknowledgments

The author wishes to thank Proffessor Eric King-wah Chu (Monash University, Australia), Proffessor Matthew M. Lin (National Chung Cheng University, Taiwan), and two anonymous referees for many interesting and valuable suggestions on the paper. This research work is partially supported by the National Science Council (NSC101-2115-M-150-002) and the National Center for Theoretical Sciences in Taiwan.

References

[1] H. W. Braden, "The equations $A^\top X \pm X^\top A = B$," *SIAM Journal on Matrix Analysis and Applications*, vol. 20, no. 2, pp. 295–302, 1998.

[2] X. Ma, J. Han, and F. Pu, "A study of solution existence for matrix equation $AX + A^\top C = B$," *Journal of Shenyang Institute of Aeronautical Engineering*, vol. 4, pp. 64–66, 2003.

[3] A. W. Bojanczyk, G. H. Golub, and P. Van Dooren, "Periodic schur decomposition: algorithms and applications," in *Advanced Signal Processing Algorithms, Architectures, and Implementations*, vol. 1770 of *Proceedings of SPIE*, pp. 31–42, San Diego, Calif, USA, July 1992.

[4] C.-Y. Chiang, E. K.-W. Chu, and W.-W. Lin, "On the $*$-Sylvester equation $AX \pm X^* B^*$," *Applied Mathematics and Computation*, vol. 218, no. 17, pp. 8393–8407, 2012.

[5] Y. Su and G. Chen, "Iterative methods for solving linear matrix equation and linear matrix system," *International Journal of Computer Mathematics*, vol. 87, no. 4, pp. 763–774, 2010.

[6] M. Wang, X. Cheng, and M. Wei, "Iterative algorithms for solving the matrix equation $AXB + CX^\top D = E$," *Applied Mathematics and Computation*, vol. 187, no. 2, pp. 622–629, 2007.

[7] B. Zhou, J. Lam, and G.-R. Duan, "Toward solution of matrix equation $Af(X)B + C$," *Linear Algebra and its Applications*, vol. 435, no. 6, pp. 1370–1398, 2011.

[8] K. D. Ikramov and O. Yu. Vorontsov, "The matrix equation $X + AX^\top B = C$: conditions for unique solvability and a numerical algorithm for its solution," *Doklady Mathematics*, vol. 85, pp. 265–267, 2012.

[9] D. S. Bernstein, *Matrix Mathematics*, Theory, Facts, and Formulas, Princeton University Press, Princeton, NJ, USA, 2nd edition, 2009.

[10] R. A. Horn and C. R. Johnson, *Topics in Matrix Analysis*, Cambridge University Press, Cambridge, UK, 1994, Corrected reprint of the 1991 original.

[11] F. Zhang, *Matrix Theory: Basic Results and Techniques*, Universitext, Springer, 1999.

[12] D. A. Bini, B. Iannazzo, and B. Meini, *Numerical Solution of Algebraic Riccati Equations*, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, Pa, USA, 2012.

[13] I. Gohberg, P. Lancaster, and L. Rodman, *Matrix Polynomials*, Computer Science and Applied Mathematics, Academic Press, New York, NY, USA, 1982.

[14] C.-Y. Chiang and M. M. Lin, "Deflating subspaces and Sylvester equations," Tech. Rep., NCTS, Taiwan, 2013.

[15] R. A. Smith, "Matrix equation $XA + BX = C$," *SIAM Journal on Applied Mathematics*, vol. 16, pp. 198–201, 1968.

[16] T. Penzl, "A cyclic low-rank Smith method for large sparse Lyapunov equations," *SIAM Journal on Scientific Computing*, vol. 21, no. 4, pp. 1401–1418, 1999.

[17] B. Zhou, J. Lam, and G.-R. Duan, "On Smith-type iterative algorithms for the Stein matrix equation," *Applied Mathematics Letters*, vol. 22, no. 7, pp. 1038–1044, 2009.

[18] R. H. Bartels and G. W. Stewart, "Solution of the matrix equation $AX + XB = C$: algorithm 432," *Communications of the ACM*, vol. 15, no. 9, pp. 820–826, 1972.

[19] G. H. Golub and C. F. Van Loan, *Matrix Computations*, Johns Hopkins Studies in the Mathematical Sciences, Johns Hopkins University Press, Baltimore, Md, USA, 3rd edition, 1996.

[20] N. J. Higham, *Accuracy and Stability of Numerical Algorithms*, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, Pa, USA, 2nd edition, 2002.

[21] A. R. Ghavimi and A. J. Laub, "Backward error, sensitivity, and refinement of computed solutions of algebraic Riccati equations," *Numerical Linear Algebra with Applications*, vol. 2, no. 1, pp. 29–49, 1995.