Linear Algebra and its Applications 435 (2011) 1370–1398

Contents lists available at ScienceDirect

Linear Algebra and its Applications

journal homepage: www.elsevier.com/locate/laa

Toward solution of matrix equation $X = Af(X)B + C$

Bin Zhou [a,*], James Lam [b], Guang-Ren Duan [a]

[a] Center for Control Theory and Guidance Technology, Harbin Institute of Technology, Harbin 150001, PR China
[b] Department of Mechanical Engineering, University of Hong Kong, Hong Kong

ARTICLE INFO

Article history:
Received 20 May 2010
Accepted 10 March 2011
Available online 22 April 2011

Submitted by C.K. Li

AMS classification:
15A24
65F10
65F05

Keywords:
Closed-form solutions
Iteration
Matrix equations
Numerical solutions
Stein equations
Conjugated and transpose

ABSTRACT

This paper studies the solvability, existence of unique solution, closed-form solution and numerical solution of matrix equation $X = Af(X)B + C$ with $f(X) = X^{\mathrm{T}}$, $f(X) = \overline{X}$ and $f(X) = X^{\mathrm{H}}$, where X is the unknown. It is proven that the solvability of these equations is equivalent to the solvability of some auxiliary standard Stein equations in the form of $W = \mathcal{A}W\mathcal{B} + \mathcal{C}$ where the dimensions of the coefficient matrices $\mathcal{A}$, $\mathcal{B}$ and $\mathcal{C}$ are the same as those of the original equation. Closed-form solutions of equation $X = Af(X)B + C$ can then be obtained by utilizing standard results on the standard Stein equation. On the other hand, some generalized Stein iterations and accelerated Stein iterations are proposed to obtain numerical solutions of equation $X = Af(X)B + C$. Necessary and sufficient conditions are established to guarantee the convergence of the iterations.

1. Introduction

Matrix equations play a very important role in linear system theory. For example, the continuous-time Lyapunov matrix equation $AX + XA^{\mathrm{T}} = -Q$ and the discrete-time Lyapunov matrix equation $AXA^{\mathrm{T}} - X = -Q$ (which is also known as the Stein equation) have very important applications in the stability analysis of continuous-time and discrete-time linear systems, respectively [26]. A more general class of equations in the form of $AX+XB = C$, which is called the Sylvester matrix equation, can be used to solve many control problems such as pole assignment, eigenstructure assignment, robust pole assignment and observer design (see, for example [2,25]). Other matrix equations such as the

* Corresponding author.
E-mail addresses: binzhoulee@163.com, binzhou@hit.edu.cn. (B. Zhou), james.lam@hku.hk (J. Lam), g.r.duan@hit.edu.cn (G. Duan).

0024-3795/$ - see front matter
doi:10.1016/j.laa.2011.03.003

coupled Sylvester matrix equations and the Riccati equations have also found numerous applications in control theory. For more related introduction, see [8,15,17,25] and the references therein.

Because of the wide applications of linear equations in control theory, many approaches have been established in the literature to provide both closed-form and numerical solutions. Take the Sylvester matrix equation $AX+XB=C$ for example. The controllability and observability matrices are used in [2] to construct the closed-form solutions, and such a solution is also obtained by the inversion of a square matrix in [12]. Regarding the numerical methods for solving the Sylvester matrix equation, one of the most efficient approaches is the Hessenberg–Schur form based algorithm proposed in [9]. Besides, the iterative algorithm by using the hierarchical identification principle proposed in [4,5] is also shown to be effective in solving this class of equations. A more general form of this class of equations is recently considered in [27]. For more references on this topic, see [3,7,11,30,16,21,18,23,24,29] and the references given there.

As a generalization of the continuous-time Lyapunov matrix equation, the following linear matrix equation

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

has received much attention in the literature over the past several decades. For example, special cases of (1) in the form of $AX\pm X^{\mathrm{T}}A^{\mathrm{T}}=B$ were thoroughly studied in [28,6,10] where necessary and sufficient conditions for the existence of solutions were obtained. Very recently, the solvability and closed-form solutions of Eq. (1) is studied in [19] by using the Moore–Penrose generalized inverse. As a more general case, the linear equation

$$AX^{\mathrm{T}}B+EXF=-C, \tag{2}$$

is investigated in [17,22] where iterative approaches are developed to obtain the numerical solutions.

In this paper, we study matrix equation

$$X=Af\left(X\right)B+C, \tag{3}$$

with $f(X)=X^{\mathrm{T}}, f(X)=\overline{X}$ and $f(X)=X^{\mathrm{H}}$, and A,B,C being known matrices. If $f(X)=X^{\mathrm{T}}$, namely, $X=AX^{\mathrm{T}}B+C$, then Eq. (3) can be regarded as a special case of (2). If $f(X)=\overline{X}$, Eq. (3) becomes $X=A\overline{X}B+C$ which was first studied in [13]. Several problems will be studied for Eq. (3). These problems include solvability, existence of unique solution, closed-form solution and numerical solution. The key idea is to transform this class of equations into the standard Stein equation

$$W=\mathcal{A}W\mathcal{B}+\mathcal{C} \tag{4}$$

where $\mathcal{A}$, $\mathcal{B}$ and $\mathcal{C}$ are functions of A,B,C, and W is unknown. It is shown that the solvability and the existence of unique solution of the original equation (3) is equivalent to the solvability and the existence of unique solution of Eq. (4). Moreover, we have proven that if the general solution (denoted by $\boldsymbol{W}$) of Eq. (4) has been obtained, then the general solution (denoted by $\boldsymbol{X}$) of the original equation (3) can be obtained via the formulation

$$\boldsymbol{X}=\frac{1}{2}\left(\boldsymbol{W}+Af\left(\boldsymbol{W}\right)B+C\right).$$

Regarding numerical solutions, we extend our early results on Smith iteration for standard Stein equation given in [24] to Eq. (3) and necessary and sufficient conditions are obtained to guarantee the convergence of the iterates. Different from the real representation based approach for solving equation $X=A\overline{X}B+C$ developed in [13] where the matrix equation is transformed into a standard Stein equation whose coefficient matrices have dimensions twice of the original coefficient matrices, here the dimensions of the coefficient matrices of the auxiliary matrix equation (4) is the same as the dimensions of coefficient matrices of the original equation (3).

The rest of this paper is organized as follows: some preliminary results in linear algebra that will be used in this paper is recalled in Section 2. A basic result regarding the solution of linear equation is developed in Section 3, which plays an important role in the development of this paper. Then, for $f(X) = X^{\mathrm{T}}, f(X) = \overline{X}$ and $f(X) = X^{\mathrm{H}}$, the matrix equation in (3) is, respectively, considered in Sections 4–6. A more general case is treated in Section 7. The paper is finally concluded in Section 8.

2. Notations and preliminaries

2.1. Notations and some standard results

In this paper, for a matrix A, we use $A^{\mathrm{T}}, A^{\mathrm{H}}, \overline{A}, \lambda\{A\}, \rho(A)$, rank (A) and det (A) to denote, respectively, the transpose, the conjugated transpose, the conjugate, the spectrum, the spectral radius, the rank and the determinant of A. The symbol $\otimes$ denotes the Kronecker product. For an $m \times n$ matrix $R = [r_{ij}]$, the so-called column stretching function vec $(\cdot)$ is defined as

$$\operatorname{vec}(R) = \begin{bmatrix} r_{11} & \cdots & r_{m1} & \cdots & r_{1n} & \cdots & r_{mn} \end{bmatrix}^{\mathrm{T}}.$$

For three matrices A, X and B with appropriate dimensions, we have the following well-known results related to the column stretching function:

$$\operatorname{vec}(AXB) = \left(B^{\mathrm{T}} \otimes A\right) \operatorname{vec}(X). \tag{5}$$

We then can introduce the following standard result.

Lemma 1 [1,17]. *Let $X \in \mathbf{C}^{m\times n}$ be any matrix. Then*

$$vec\left(X^T\right) = P_{(m,n)} vec(X),$$

where $P_{(m,n)}$ is uniquely determined by the integers m and n. Moreover, the matrix $P_{(m,n)}$ has the following properties.

1. *For two arbitrary integers m and n, $P_{(m,n)}$ has the following explicit form*

$$P_{(m,n)} = \begin{bmatrix} E_{11}^T & E_{12}^T & \cdots & E_{1n}^T \\ E_{21}^T & E_{22}^T & \cdots & E_{2n}^T \\ \vdots & \vdots & \ddots & \vdots \\ E_{m1}^T & E_{m2}^T & \cdots & E_{mn}^T \end{bmatrix} \in \mathbf{R}^{mn\times mn},$$

   where $E_{ij}, i = 1, 2, \ldots, m, j = 1, 2, \ldots, n$ is an $m \times n$ matrix with the element at position (i, j) being 1 and the others being 0.
2. *For two arbitrary integers m and n, $P(m, n)$ is a unitary matrix, that is*

$$P_{(m,n)}P_{(m,n)}^T = P_{(m,n)}^T P_{(m,n)} = I_{mn}.$$

3. *For two arbitrary integers m and n, there holds $P_{(m,n)} = P_{(n,m)}^T = P_{(n,m)}^{-1}$.*
4. *Let m, n, p and q be four integers and $A \in \mathbf{C}^{m\times n}, B \in \mathbf{C}^{p\times q}$. Then*

$$P_{(m,p)}(B \otimes A) = (A \otimes B) P_{(n,q)}. \tag{6}$$

Let $A \in \mathbf{C}^{m\times n}$ be any given matrix, the generalized inverse of A, denoted by A^{-}, is such that

$$AA^{-}A = A.$$

Then we can recall the following standard result regarding solutions of linear equation.

Lemma 2. *Let $A \in \mathbf{C}^{m\times n}$ and $b \in \mathbf{C}^{m}$ be given. Then the linear equation*

$$Ax = b \tag{7}$$

is solvable if and only if there exists a A^{-} such that $AA^{-}b = b$. Moreover, if it is solvable and denote $x^{\dagger}$ as one of its solutions, then the general solution of Eq. (7) is given by

$$x = x^{\dagger} + \left(I_n - A^{-}A\right)t,$$

where $t \in \mathbf{C}^{n}$ is an arbitrary vector.

The following result is the well-known Schur complement.

Lemma 3 [1]**.** *Let X_1, X_2, X_{12} and X_{21} be some matrices with appropriate dimensions. Assume that X_1 and X_2 are nonsingular. Then,*

$$\begin{aligned}\det\begin{bmatrix} X_1 & X_{12} \\ X_{21} & X_2 \end{bmatrix} &= \det(X_1)\det\left(X_2 - X_{21}X_1^{-1}X_{12}\right) \\ &= \det(X_2)\det\left(X_1 - X_{12}X_2^{-1}X_{21}\right).\end{aligned}$$

Finally, we recall some facts about linear iteration.

Definition 1 [24]**.** For a linear iteration

$$x_{k+1} = Mx_k + g, \quad g \in \mathbf{R}^{n}, \tag{8}$$

where M and g are, respectively, constant matrix and vector and $\rho(M) < 1$, the asymptotic exponential convergence rate is $-\ln(\rho(M))$.

Basically, the smaller the number $\rho(M)$, the faster convergence of the iteration (8). See [24] for explanations.

2.2. Smith iteration and closed-form solutions of standard stein equation

Consider the following standard Stein matrix equation

$$X = AXB + C, \tag{9}$$

where $A \in \mathbf{C}^{m\times m}$, $B \in \mathbf{C}^{n\times n}$ and $C \in \mathbf{C}^{m\times n}$ are known and $X \in \mathbf{C}^{m\times n}$ is to be determined. The associated Smith iteration is defined as

$$X_{k+1} = AX_kB + C, \quad \forall X_0 \in \mathbf{C}^{m\times n}. \tag{10}$$

It is known that the above Smith iteration converges to the unique solution of Eq. (9) for arbitrarily initial condition if and only if $\rho(A)\rho(B) < 1$. Moreover, the asymptotic exponential convergence rate is $-\ln(\rho(A)\rho(B))$. Therefore, if $\rho(A)\rho(B)$ is very close to 1, the convergenceis very slow. To improve the convergence rate, the so-called Smith (l) iteration can be applied. The idea of Smith (l) iteration is to combine l steps of Smith iterations into one step. Notice that

$$\begin{aligned}X_{k+l} &= AX_{k+l-1}B + C\\ &= A(AX_{k+l-2}B + C)B + C\\ &= A^2X_{k+l-2}B^2 + ACB + C\\ &= \cdots\\ &= A^lX_kB^l + \sum_{i=0}^{l-1} A^iCB^i.\end{aligned} \tag{11}$$

By denoting $X_{kl} = Y_k$, $k = 0, 1, \ldots,$ we can rewrite (11) as

$$Y_{k+1} = A^lY_kB^l + \sum_{i=0}^{l-1} A^iCB^i, \quad \forall Y_0 = X_0 \in \mathbf{C}^{m\times n}. \tag{12}$$

It is shown in [24] that the above iteration converges to the unique solution of (9) for arbitrarily initial condition if and only if $\rho(A)\rho(B) < 1$. Moreover, the asymptotic exponential convergence rate is $-l\ln(\rho(A)\rho(B))$. Therefore, the convergence rate of (12) increases as l increases.

Though Smith (l) iteration (12) can improve the convergence rate of the standard Smith iteration (10), it still converges exponentially. In order to obtain super-exponential convergence rate and by noting that the closed-form solution of Eq. (9) is

$$X^\star = \sum_{i=0}^{\infty} A^iCB^i, \quad \rho(A)\rho(B) < 1,$$

the so-called r-Smith iteration can be constructed.

Lemma 4 [24]. *Let $r \geqslant 2$ be a prescribed positive integer. Assume that $\rho(A)\rho(B) < 1$. Construct the following iteration:*

$$X_{k+1} = \sum_{i=0}^{r-1} A_k^iX_kB_k^i, \quad A_{k+1} = A_k^r,\ B_{k+1} = B_k^r,\ k \geqslant 0, \tag{13}$$

with initial condition $X_0 = C$, $A_0 = A$ and $B_0 = B$. Denote $E_k = X^\star - X_k$ where $X^\star$ is the unique solution of (9). Then $\lim_{k\to\infty} X_k = X^\star$ and

$$X_k = \sum_{i=0}^{r^k-1} A^iCB^i, \quad E_k = A^{r^k}X^\star B^{r^k},\ k \geqslant 0. \tag{14}$$

If $r = 2$, the r-Smith iteration reduces to the well-known Smith accelerative iteration ([20]). It is clear that the r-Smith iteration (13) has a super-exponential convergence rate and is very effective in obtaining numerical solutions of Eq. (9) ([24]).

In the following, we consider closed-form solutions of the Stein equation (9). Let the characteristic polynomial of A be

$$\alpha_A(s) = \det(sI_m - A) = s^m + \sum_{i=0}^{m-1} \alpha_i s^i.$$

Then it is easy to verify that

$$h_A(s) \triangleq \det(I_m - sA) = \alpha_m + \sum_{i=1}^{m} \alpha_{m-i}s^i, \quad \alpha_m = 1.$$

We then can cite the following result regarding closed-form solutions of Eq. (9).

Lemma 5 [13]. *Let $A \in \mathbf{C}^{m\times m}$, $B \in \mathbf{C}^{n\times n}$ and $C \in \mathbf{C}^{m\times n}$ be given. Then*

1. *If Eq. (9) has a solution X, then*

$$Xh_A(B) = \sum_{k=1}^{m}\sum_{s=1}^{k} \alpha_k A^{k-s} C B^{m-s}.$$

2. *If Eq. (9) has a unique solution $X^\star$, then*

$$X^\star = \left(\sum_{k=1}^{m}\sum_{s=1}^{k} \alpha_k A^{k-s} C B^{m-s}\right)(h_A(B))^{-1}.$$

3. A basic result on linear equation

Before giving detailed discussion to Eq. (3), we present a basic result on linear equation. This result is essential in the development of this paper.

Proposition 1. *Let $A \in \mathbf{C}^{n\times n}$ and $c \in \mathbf{C}^n$ be given. Then the linear equation*

$$\left(I_n - A^2\right)\omega = (I_n + A)\,c, \tag{15}$$

is solvable with unknown ω if and only if the linear equation

$$(I_n - A)\,x = c, \tag{16}$$

is solvable with unknown x. More specifically,

1. *If x is a solution of Eq. (16), then $\omega = x$ is a solution of Eq. (15).*
2. *If $\boldsymbol{\omega}$ is the general solution of (15), then the general solution of Eq. (16) is*

$$\boldsymbol{x} = \frac{1}{2}(I_n + A)\,\boldsymbol{\omega} + \frac{1}{2}c. \tag{17}$$

3. *If Eq. (15) has a unique solution $\omega^\star$, then Eq. (16) also has a unique solution $x^\star = \omega^\star$.*

Proof. *Proof of Item 1.* This is trivial since Eq. (15) can be written as $(I_n + A)\left((I_n - A)\,\omega - c\right) = 0$.

Proof of Item 2. Let P be some matrix such that

$$A = P\begin{bmatrix} R & 0 & 0 \\ 0 & J & 0 \\ 0 & 0 & K \end{bmatrix}P^{-1}, \quad R \in \mathbf{C}^{n_0\times n_0}, \quad J \in \mathbf{C}^{n_1\times n_1},\ K \in \mathbf{C}^{n_2\times n_2}, \tag{18}$$

where all the eigenvalues of J are 1, all the eigenvalues of K are -1 and $\{-1, 1\} \cap \lambda\{R\} = \varnothing$. Clearly, $n_0 + n_1 + n_2 = n$. Then we can compute

$$I_n \pm A = P\begin{bmatrix} R_\pm & 0 & 0 \\ 0 & J_\pm & 0 \\ 0 & 0 & K_\pm \end{bmatrix}P^{-1}, \tag{19}$$

$$I_n - A^2 = P\begin{bmatrix} R_+R_- & 0 & 0 \\ 0 & J_-J_+ & 0 \\ 0 & 0 & K_+K_- \end{bmatrix}P^{-1}, \tag{20}$$

where $R_{\pm} = I_{n_0} \pm R, J_{\pm} = I_{n_1} \pm J$ and $K_{\pm} = I_{n_2} \pm K$. According to Lemma 2 and (20), the general closed-form solution of (15) can be written as

$$\boldsymbol{\omega} = P \begin{bmatrix} 0 \\ J_{+}^{-1}\left(I_{n_1} - J_{-}^{-}J_{-}\right)s \\ K_{-}^{-1}\left(I_{n_1} - K_{+}^{-}K_{+}\right)t \end{bmatrix} + \omega^{\dagger}, \quad \forall s \in \mathbf{R}^{n_1},\ \forall t \in \mathbf{R}^{n_2}, \tag{21}$$

where J_{-}^{-} and K_{+}^{-} are any fixed generalized inversion of J_{-} and K_{+}, respectively, and $\omega^{\dagger}$ is a particular solution of Eq. (15), namely,

$$\left(I_n - A^2\right)\omega^{\dagger} = \left(I_n + A\right)c. \tag{22}$$

Let

$$x^{\dagger} \triangleq \frac{1}{2}\left(I_n + A\right)\omega^{\dagger} + \frac{1}{2}c.$$

Then, it follows from (22) that

$$\begin{aligned}
\left(I_n - A\right)x^{\dagger} &= \left(I_n - A\right)\left(\frac{1}{2}\left(I_n + A\right)\omega^{\dagger} + \frac{1}{2}c\right) \\
&= \frac{1}{2}\left(I_n - A^2\right)\omega^{\dagger} + \frac{1}{2}\left(I_n - A\right)c \\
&= \frac{1}{2}\left(I_n + A\right)c + \frac{1}{2}\left(I_n - A\right)c \\
&= c,
\end{aligned} \tag{23}$$

namely, $x^{\dagger}$ is a particular solution of Eq. (16). Hence, by using (18)–(21), we have

$$\begin{aligned}
\frac{1}{2}\left(I_n + A\right)\boldsymbol{\omega} + \frac{1}{2}c &= \frac{1}{2}P\begin{bmatrix} R_{+} & 0 & 0 \\ 0 & J_{+} & 0 \\ 0 & 0 & K_{+} \end{bmatrix} P^{-1}P\begin{bmatrix} 0 \\ J_{+}^{-1}\left(I_{n_1} - J_{-}^{-}J_{-}\right)s \\ K_{-}^{-1}\left(I_{n_2} - K_{+}^{-}K_{+}\right)t \end{bmatrix} + \frac{1}{2}\left(I_n + A\right)\omega^{\dagger} + \frac{1}{2}c \\
&= \frac{1}{2}P\begin{bmatrix} 0 \\ \left(I_{n_1} - J_{-}^{-}J_{-}\right)s \\ K_{+}K_{-}^{-1}\left(I_{n_2} - K_{+}^{-}K_{+}\right)t \end{bmatrix} + x^{\dagger} \\
&= \frac{1}{2}P\begin{bmatrix} 0 \\ \left(I_{n_1} - J_{-}^{-}J_{-}\right)s \\ K_{-}^{-1}K_{+}\left(I_{n_2} - K_{+}^{-}K_{+}\right)t \end{bmatrix} + x^{\dagger} \\
&= \frac{1}{2}P\begin{bmatrix} 0 \\ \left(I_{n_1} - J_{-}^{-}J_{-}\right)s \\ K_{-}^{-1}\left(K_{+} - K_{+}K_{+}^{-}K_{+}\right)t \end{bmatrix} + x^{\dagger} \\
&= P\begin{bmatrix} 0 \\ \left(I_{n_1} - J_{-}^{-}J_{-}\right)s' \\ 0 \end{bmatrix} + x^{\dagger},
\end{aligned} \tag{24}$$

where we have used the fact that $K_+ = K_+K_+^- K_+$ and denoted $s' = \frac{1}{2}s$. On the other hand, in view of (23) and by using Lemma 2 again, the general solution of Eq. (16) can be expressed as

$$\boldsymbol{x} = P \begin{bmatrix} 0 \\ \left(I_{n_1} - J_-^- J_-\right) s' \\ 0 \end{bmatrix} + x^\dagger, \ \forall s' \in \mathbf{C}^{n_1}. \tag{25}$$

Comparing (24) with (25), we clearly see that $\boldsymbol{x}$ in (17) is the general solution of Eq. (16).

Proof of Item 3. Eq. (15) has a unique solution

$$\omega^\star = \left(I_n - A^2\right)^{-1} (I_n + A)\, c = (I_n - A)^{-1} c,$$

if and only if $1 \notin \lambda\{A^2\}$ which implies that $1 \notin \lambda\{A\}$, namely, Eq. (16) has a unique solution $x^\star = (I_n - A)^{-1} c$. $\square$

We should point out that there is an important difference between Item 2 and Item 1 of Proposition 1: even though $\boldsymbol{x}$ is the general solution of Eq. (16), $\boldsymbol{\omega} = \boldsymbol{x}$ is not the general solution of Eq. (15). This can be observed from (21) and (25). For illustration, we give an example.

Example 1. Let A and c in Eqs. (15) and (16) be chosen as

$$A = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, \quad c = \begin{bmatrix} 0 \\ 1 \end{bmatrix}.$$

Then it is easy to see that the general solutions of Eqs. (16) and (15) are, respectively,

$$\boldsymbol{x} = \begin{bmatrix} t \\ \frac{1}{2} \end{bmatrix}, \quad \boldsymbol{\omega} = \begin{bmatrix} t \\ s \end{bmatrix}, \ \forall t,\ s \in \mathbf{C}.$$

It follows that there are two free parameters in $\boldsymbol{\omega}$ while there is only one free parameter in $\boldsymbol{x}$ and thus $\boldsymbol{\omega} = \boldsymbol{x}$ is not the general solution of Eq. (15).

Also, if Eq. (16) has a unique solution, we cannot in general conclude that Eq. (15) has a unique solution. The simple counterexample can be constructed by letting $-1 \in \lambda\{A\}$.

4. Equation $X = AX^{\mathrm{T}}B + C$

In this section, we consider Eq. (3) with $f(X) = X^{\mathrm{T}}$, namely, the following matrix equation

$$X = AX^{\mathrm{T}}B + C, \tag{26}$$

where $A \in \mathbf{C}^{m\times n}$, $B \in \mathbf{C}^{m\times n}$ and $C \in \mathbf{C}^{m\times n}$ are some given matrices and $X \in \mathbf{C}^{m\times n}$ is a matrix to be determined.

4.1. Solvability and closed-form solutions

By taking vec $(\cdot)$ on both sides of (26) and using Lemma 1, we have

$$\begin{aligned} \mathrm{vec}\,(C) &= \left(I_{mn} - \left(B^{\mathrm{T}} \otimes A\right) P_{(m,n)}\right) \mathrm{vec}\,(X) \\ &= \left(P_{(n,m)} - B^{\mathrm{T}} \otimes A\right) P_{(m,n)} \mathrm{vec}\,(X)\,. \end{aligned} \tag{27}$$

Therefore, the following simple result can be obtained.

Lemma 6. *Eq. (26) is solvable if and only if*

$$rank\left[\begin{array}{cc} P_{(n,m)} - B^T \otimes A & vec\,(C) \end{array}\right] = rank\left[\begin{array}{c} P_{(n,m)} - B^T \otimes A \end{array}\right]. \tag{28}$$

Moreover, it has a unique solution for arbitrary C if and only if

$$1 \notin \lambda\left\{\left(B^T \otimes A\right) P_{(m,n)}\right\}. \tag{29}$$

Notice that for test condition (29), we need to compute the eigenvalues of a matrix having dimension mn, which is demanding in practice. In the following, we will provide an alternative condition for the solvability of this equation.

Theorem 1. *Matrix equation (26) is solvable if and only if the following standard Stein equation*

$$W = AB^TWA^TB + AC^TB + C, \tag{30}$$

is solvable with unknown W. More specifically,

1. *If X is a solution of Eq. (26), then $W = X$ is a solution of Eq. (30).*
2. *If $\boldsymbol{W}$ is the general solution of Eq. (30), then the general solution of Eq. (26) is*

$$\boldsymbol{X} = \frac{1}{2}\left(\boldsymbol{W} + A\boldsymbol{W}^TB + C\right). \tag{31}$$

3. *Eq. (26) has a unique solution $X^\star$ if Eq. (30) has a unique solution $W^\star$, namely,*

$$\eta\gamma \neq 1, \quad \forall \eta,\ \gamma \in \lambda\left\{AB^T\right\}. \tag{32}$$

Moreover, $X^\star = W^\star$.

Proof. *Proofs of Item 1 and Item 2.* By using (6), we know that

$$P_{(m,n)}\left(B^{\mathrm{T}} \otimes A\right) P_{(m,n)} = A \otimes B^{\mathrm{T}}. \tag{33}$$

With this, by taking $\mathrm{vec}\,(\cdot)$ on both sides of (30) and denoting $\omega = \mathrm{vec}\,(W)$, we get

$$\begin{aligned} \omega &= \left(B^{\mathrm{T}} \otimes A\right) \mathrm{vec}\left(B^{\mathrm{T}}WA^{\mathrm{T}}\right) + \mathrm{vec}\left(AC^{\mathrm{T}}B\right) + \mathrm{vec}\,(C) \\ &= \left(B^{\mathrm{T}} \otimes A\right)\left(A \otimes B^{\mathrm{T}}\right)\omega + \left(I_{mn} + \left(B^{\mathrm{T}} \otimes A\right) P_{(m,n)}\right) \mathrm{vec}\,(C) \\ &= \left(B^{\mathrm{T}} \otimes A\right) P_{(m,n)} \left(B^{\mathrm{T}} \otimes A\right) P_{(m,n)}\omega + \left(I_{mn} + \left(B^{\mathrm{T}} \otimes A\right) P_{(m,n)}\right) \mathrm{vec}\,(C) \\ &= \Upsilon^2\omega + (I_{mn} + \Upsilon)\,\mathrm{vec}\,(C), \end{aligned}$$

where $\Upsilon = \left(B^{\mathrm{T}} \otimes A\right) P_{(m,n)}$. The above equation can also be written as

$$\left(I_{mn} - \Upsilon^2\right)\omega = (I_{mn} + \Upsilon)\,\mathrm{vec}\,(C). \tag{34}$$

Similarly, taking $\mathrm{vec}\,(\cdot)$ on both sides of (26) and denoting $x = \mathrm{vec}\,(X)$ produce

$$(I_{mn} - \Upsilon)\,x = \mathrm{vec}\,(C), \tag{35}$$

and taking $\mathrm{vec}\,(\cdot)$ on both sides of expression (31) yields

$$\mathrm{vec}\,(\boldsymbol{X}) = \frac{1}{2}(I_{mn} + \Upsilon)\,\mathrm{vec}\,(\boldsymbol{W}) + \frac{1}{2}\mathrm{vec}\,(C). \tag{36}$$

We notice that Eq. (34) is in the form of (15), Eq. (35) is in the form of (16) and expression (36) is similar to (17). Then Item 1 and Item 2 follow from Proposition 1 directly.

Proof of Item 3. From the above proof and Proposition 1, we conclude that Eq. (26) has a unique solution if Eq. (30) has a unique solution which is equivalent to

$$1 \notin \lambda \left\{ \left(A^{\mathrm{T}}B\right)^{\mathrm{T}} \otimes AB^{\mathrm{T}} \right\} = \lambda \left\{ B^{\mathrm{T}}A \otimes AB^{\mathrm{T}} \right\}. \tag{37}$$

The remaining is to show that (32) and (37) are equivalent. Without loss of generality, we assume that $n \geqslant m$. Let $\lambda \left\{AB^{\mathrm{T}}\right\} = \{\eta_1, \eta_2, \ldots, \eta_m\}$. Then

$$\lambda \left\{B^{\mathrm{T}}A\right\} = \{\eta_1, \eta_2, \ldots, \eta_m, 0, 0, \ldots, 0\}.$$

As a result, we have

$$\lambda \left\{B^{\mathrm{T}}A \otimes AB^{\mathrm{T}}\right\} = \left\{\eta_i\eta_j, 0, 0, \ldots, 0\right\}, \quad i, j \in \{1, 2, \ldots, m\}.$$

Therefore, (37) is true if and only if (32) is satisfied. □

One may ask whether the condition in (32) is necessary for the uniqueness of solution of Eq. (26). The answer is negative, as the following example indicates.

Example 2. Consider a linear equation in the form of (26) with

$$A = \begin{bmatrix} 2 & 0 \\ 1 & \alpha \end{bmatrix}, \quad B = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \tag{38}$$

where α is to be determined. By direct computation, the characteristic polynomial of $\left(B^{\mathrm{T}} \otimes A\right) P_{(m,n)}$ is $(s-2)(s-a)(s^2-2\alpha) = 0$. Therefore, according to (29), the equation in (26) has a unique solution for arbitrary C if and only if $\alpha \neq 1$ and $\alpha \neq \frac{1}{2}$. On the other hand, as $\lambda \left\{B^{\mathrm{T}}A\right\} = \{2, \alpha\}$, the condition in (32) is equivalent to $\alpha \neq \pm 1$ and $\alpha \neq \frac{1}{2}$. Hence, if $\alpha = -1$, condition (32) is not satisfied while the Eq. (26) has a unique solution. This clearly implies that (32) is only a sufficient but not a necessary condition for the existence of unique solution of Eq. (26) for arbitrary C.

Though Theorem 1 only provides a sufficient condition on the existence of the unique solution to Eq. (26), the advantage is that the condition in (32) is easier to test than (29) because only the eigenvalues of a matrix of dimension m is required.

Remark 1. If $n \leqslant m$, we need only to compute the eigenvalues of a matrix of dimension n because condition (32) can be replaced by $\eta\gamma \neq 1,\ \forall \eta, \gamma \in \lambda \left\{BA^{\mathrm{T}}\right\}$.

By combining Lemma 5 and Theorem 1, we have the following result regarding the closed-form solutions of Eq. (26).

Theorem 2. *Let $A \in \mathbf{C}^{m\times n}$, $B \in \mathbf{C}^{m\times n}$, $C \in \mathbf{C}^{m\times n}$ and*

$$h_{AB^T}(s) = \det\left(I_m - sAB^T\right) = \alpha_m + \sum_{i=1}^{m} \alpha_{m-i}s^i, \quad \alpha_m = 1.$$

1. *If Eq. (26) has a solution X, then*

$$Xh_{AB^T}\left(A^TB\right) = \sum_{k=1}^{m}\sum_{s=1}^{k} \alpha_k \left(AB^T\right)^{k-s} \left(AC^TB + C\right) \left(A^TB\right)^{m-s}.$$

2. *If (32) is satisfied, then Eq. (26) has a unique solution $X^\star$ given by*

$$X^\star = \left(\sum_{k=1}^{m}\sum_{s=1}^{k}\alpha_k \left(AB^T\right)^{k-s}\left(AC^TB+C\right)\left(A^TB\right)^{m-s}\right)\left(h_{AB^T}\left(A^TB\right)\right)^{-1}.$$

4.2. Smith iteration for equation $X = AX^TB + C$

By recognizing the Smith iteration (10) for the standard Stein equation (9), we can construct the following Smith iteration for matrix equation (26):

$$X_{k+1} = AX_k^{\mathrm{T}}B + C, \quad \forall X_0 \in \mathbf{C}^{m\times n}. \tag{39}$$

In the following, we will study the convergence of this iteration. First, we introduce a lemma.

Lemma 7. *Let $A, B \in \mathbf{C}^{m\times n}$ be given and $\Upsilon = \left(B^T \otimes A\right)P_{(m,n)}$. Then*

$$\rho\left(\Upsilon\right) = \rho\left(B^TA\right).$$

Proof. By definition of spectral radius and using (33), we have

$$\begin{aligned}
\rho^2\left(\Upsilon\right) &= \rho\left(\Upsilon^2\right)\\
&= \rho\left(\left(\left(B^{\mathrm{T}}\otimes A\right)P_{(m,n)}\right)\left(\left(B^{\mathrm{T}}\otimes A\right)P_{(m,n)}\right)\right)\\
&= \rho\left(\left(B^{\mathrm{T}}\otimes A\right)\left(P_{(m,n)}\left(B^{\mathrm{T}}\otimes A\right)P_{(m,n)}\right)\right)\\
&= \rho\left(\left(B^{\mathrm{T}}\otimes A\right)\left(A\otimes B^{\mathrm{T}}\right)\right)\\
&= \rho\left(B^{\mathrm{T}}A\otimes AB^{\mathrm{T}}\right)\\
&= \rho\left(B^{\mathrm{T}}A\right)\rho\left(AB^{\mathrm{T}}\right)\\
&= \rho^2\left(B^{\mathrm{T}}A\right). \quad \square
\end{aligned}$$

We then can prove the following result.

Theorem 3. *Assume that Eq. (26) has a unique solution $X^\star$. Then the Smith iteration (39) converges to $X^\star$ for arbitrary initial condition X_0 if and only if*

$$\rho\left(B^TA\right) < 1. \tag{40}$$

Moreover, the asymptotic exponential convergence rate is $-\ln\left(\rho\left(B^TA\right)\right)$.

Proof. Taking $\mathrm{vec}(\cdot)$ on both sides of (39) gives

$$\mathrm{vec}\left(X_{k+1}\right) = \Upsilon\mathrm{vec}\left(X_k\right) + \mathrm{vec}\left(C\right),$$

which converges to a finite vector $\mathrm{vec}\left(X_\infty\right)$ independent of the initial condition if and only if $\rho\left(\Upsilon\right) < 1$. This is equivalent to (40) in view of Lemma 7. Moreover, we have

$$X_\infty = AX_\infty^{\mathrm{T}}B + C.$$

The proof is completed by noting the fact that Eq. (26) has only a unique solution under the condition (40). $\square$

Notice that Eq. (30) is a standard Stein equation whose associated Smith iteration is

$$W_{k+1} = AB^{\mathrm{T}}W_kA^{\mathrm{T}}B + AC^{\mathrm{T}}B + C, \quad \forall W_0 \in \mathbf{C}^{m\times n}. \tag{41}$$

It is trivial to show that this iteration converges to a constant matrix independent of the initial condition W_0 if and only if (40) is satisfied. As (40) implies (32), we conclude from Theorem 1 that the iteration (41) also converges to the unique solution $X^\star = W^\star$ to Eq. (26).

Remark 2. In view of (11), it is easy to see that iteration (41) can be regarded as Smith (2) iteration associated with the Smith iteration (39). Therefore, by using a similar technique as that used in (11), for arbitrary integer $l > 1$, we can construct Smith (l) iteration for Eq. (26). The details are omitted for brevity.

Remark 3. Since condition (40) implies condition (32), it follows from Item 3 of Theorem 1 that the unique solution of Eq. (26) under condition (40) is also the unique solution of Eq. (30), namely,

$$X^\star = W^\star = \sum_{i=0}^{\infty} \left(AB^{\mathrm{T}}\right)^i \left(AC^{\mathrm{T}}B + C\right) \left(A^{\mathrm{T}}B\right)^i.$$

Therefore, in view of Lemma 4, for any integer $r \geqslant 1$, we can construct the following r-Smith iteration

$$X_{k+1} = \sum_{i=0}^{r-1} \mathcal{A}_k^i X_k \mathcal{B}_k^i, \quad \mathcal{A}_{k+1} = \mathcal{A}_k^r,\ \mathcal{B}_{k+1} = \mathcal{B}_k^r,\ k \geqslant 0,$$

with $X_0 = AC^{\mathrm{T}}B + C$, $\mathcal{A}_0 = AB^{\mathrm{T}}$ and $\mathcal{B}_0 = A^{\mathrm{T}}B$. Consequently, we have $\lim_{k\to\infty} X_k = X^\star$. Actually, we have

$$X_k = \sum_{i=0}^{r^k-1} \left(AB^{\mathrm{T}}\right)^i \left(AC^{\mathrm{T}}B + C\right) \left(A^{\mathrm{T}}B\right)^i.$$

5. Revisit of equation $X = A\overline{X}B + C$

To study matrix equation $X = AX^{\mathrm{H}}B + C$, we should first revisit the following matrix equation

$$X = A\overline{X}B + C, \tag{42}$$

where $A \in \mathbf{C}^{m\times m}$, $B \in \mathbf{C}^{n\times n}$ and $C \in \mathbf{C}^{m\times n}$ are given, and $X \in \mathbf{C}^{m\times n}$ is to be determined. This equation was first studied in [13]. To introduce the results obtained in that paper, we need some preliminaries. For a complex matrix $A = A_1 + \mathrm{i}A_2$ where $\mathrm{i} = \sqrt{-1}$, the associated real representation is

$$A_\sigma = \begin{bmatrix} A_1 & A_2 \\ A_2 & -A_1 \end{bmatrix}. \tag{43}$$

The real representation has many good properties, one of which is the following.

Lemma 8 Proposition 3.2 in [13]**.** *Let $A \in \mathbf{C}^{m\times m}$. If $\alpha \in \lambda\{A_\sigma\}$, then $\{\pm\alpha, \pm\overline{\alpha}\} \subset \lambda\{A_\sigma\}$.*

Based on the real representation, the following elegant result was established in [13] regarding solutions of Eq. (42).

Proposition 2 [13]**.** *The Eq. (42) is solvable if and only if the following equation*

$$Y = A_\sigma Y B_\sigma + C_\sigma, \tag{44}$$

is solvable with unknown Y. Moreover, let Y be any real[1] *solution of (44), then the solution of Eq. (42) is*

$$X=\frac{1}{4}\begin{bmatrix} I_m & iI_m \end{bmatrix}(Y+Q_mYQ_n)\begin{bmatrix} I_n \\ iI_n \end{bmatrix},$$

where

$$Q_s=\begin{bmatrix} 0 & I_s \\ -I_s & 0 \end{bmatrix}. \tag{45}$$

To use this result, we need to solve linear equation (44) whose dimension is twice of the original matrix equation (42). In this section, by using the idea used in Section 4, we will give an alternative criterion for the solvability and computation methods for Eq. (42).

5.1. Some useful results

In this subsection, we give some useful results that will be used later.

Lemma 9. *Let $A=A_1+iA_2\in\mathbf{C}^{m\times n}$, where $A_1,A_2\in\mathbf{R}^{m\times n}$, be any given complex matrix. Then*

$$A_\phi\triangleq\begin{bmatrix} A_1 & -A_2 \\ A_2 & A_1 \end{bmatrix}=Z_m\begin{bmatrix} \overline{A} & 0 \\ 0 & A \end{bmatrix}Z_n^H, \tag{46}$$

where Z_n is a unitary matrix defined as

$$Z_n=\frac{\sqrt{2}}{2}\begin{bmatrix} iI_n & I_n \\ I_n & iI_n \end{bmatrix}.$$

Moreover, for two arbitrary matrices $U\in\mathbf{C}^{m\times n}$ and $V\in\mathbf{C}^{n\times q}$, there holds

$$U_\sigma V_\sigma=\left(U\overline{V}\right)_\phi. \tag{47}$$

Proof. Eq. (46) can be shown by direct manipulation. We next prove (47). By letting $U=U_1+\mathrm{i}U_2$ and $V=V_1+\mathrm{i}V_2$, where U_i and V_i are real matrices, we have

$$\begin{aligned} U_\sigma V_\sigma &= \begin{bmatrix} U_1 & U_2 \\ U_2 & -U_1 \end{bmatrix}\begin{bmatrix} V_1 & V_2 \\ V_2 & -V_1 \end{bmatrix} \\ &= \begin{bmatrix} U_1V_1+U_2V_2 & U_1V_2-U_2V_1 \\ U_2V_1-U_1V_2 & U_1V_1+U_2V_2 \end{bmatrix} \\ &= \left(U_1V_1+U_2V_2+\mathrm{i}\left(U_2V_1-U_1V_2\right)\right)_\phi \\ &= \left(\left(U_1+\mathrm{i}U_2\right)\left(V_1-\mathrm{i}V_2\right)\right)_\phi \\ &= \left(U\overline{V}\right)_\phi, \end{aligned}$$

which is (47). □

[1] Though it is not mentioned in [13] that Y is real solution, it indeed should be so.

Let $A = A_1 + iA_2 \in \mathbf{C}^{m\times n}$ be any given matrix, where A_1 and A_2 are real matrices. Then the linear mapping $\varphi(A) : \mathbf{C}^{m\times n} \to \mathbf{R}^{2mn}$ is defined as

$$\varphi(A) = \begin{bmatrix} \operatorname{vec}(A_1) \\ \operatorname{vec}(A_2) \end{bmatrix}.$$

We notice that mapping $\varphi(\cdot)$ is bijective and denote

$$\varphi^{-1}(c) = \operatorname{vec}^{-1}(c_1) + \operatorname{vec}^{-1}(c_2)\,\mathrm{i}, \quad c = \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} \in \mathbf{R}^{2n\times 1},\ c_1 \in \mathbf{R}^{n\times 1}.$$

Here we have assumed that $\operatorname{vec}^{-1}(\cdot)$ is defined associated with $\operatorname{vec}(\cdot)$ in an obvious way. The following lemma gives further properties of $\varphi(A)$.

Lemma 10. *Let $A \in \mathbf{C}^{p\times m}$ and $B \in \mathbf{C}^{n\times q}$ be given. Then for any $X \in \mathbf{C}^{m\times n}$, there hold*

$$\varphi(A\overline{X}B) = \left(B^T \otimes A\right)_\sigma \varphi(X), \tag{48}$$

$$\varphi(AXB) = \left(B^T \otimes A\right)_\phi \varphi(X). \tag{49}$$

Proof. Let $A = A_1 + iA_2, B = B_1 + iB_2, X = X_1 + iX_2$ where A_i, B_i and $X_i, i = 1, 2$ are real matrices. Then $Y = A\overline{X}B$ where $Y = Y_1 + iY_2 \in \mathbf{C}^{m\times n}$ with Y_1 and Y_2 being real matrices, is equivalent to

$$\begin{cases} Y_1 = (A_1X_1 + A_2X_2)\,B_1 - (A_2X_1 - A_1X_2)\,B_2, \\ Y_2 = (A_2X_1 - A_1X_2)\,B_1 + (A_1X_1 + A_2X_2)\,B_2. \end{cases}$$

Taking vec$(\cdot)$ on both sides of the above two equations, we get $\varphi(Y) = T(A, B)\,\varphi(X)$, where

$$T(A, B) = \begin{bmatrix} B_1^{\mathrm{T}} \otimes A_1 - B_2^{\mathrm{T}} \otimes A_2 & B_1^{\mathrm{T}} \otimes A_2 + B_2^{\mathrm{T}} \otimes A_1 \\ B_1^{\mathrm{T}} \otimes A_2 + B_2^{\mathrm{T}} \otimes A_1 & B_2^{\mathrm{T}} \otimes A_2 - B_1^{\mathrm{T}} \otimes A_1 \end{bmatrix}.$$

In view of the real representation in (43), we can see that

$$\begin{aligned} T(A, B) &= \left(\left(B_1^{\mathrm{T}} \otimes A_1 - B_2^{\mathrm{T}} \otimes A_2\right) + \mathrm{i}\left(B_1^{\mathrm{T}} \otimes A_2 + B_2^{\mathrm{T}} \otimes A_1\right)\right)_\sigma \\ &= \left(\left(B_1^{\mathrm{T}} + \mathrm{i}B_2^{\mathrm{T}}\right) \otimes (A_1 + \mathrm{i}A_2)\right)_\sigma \\ &= \left(B^{\mathrm{T}} \otimes A\right)_\sigma, \end{aligned} \tag{50}$$

which is (48).

Next we show (49). By using (47), (48) and by definition of $\varphi(\cdot)$, we have

$$\begin{aligned} \varphi(AXB) &= \left(B^{\mathrm{T}} \otimes A\right)_\sigma \varphi(\overline{X}) \\ &= \left(B^{\mathrm{T}} \otimes A\right)_\sigma \begin{bmatrix} I_{mn} & 0 \\ 0 & -I_{mn} \end{bmatrix} \varphi(X) \\ &= \left(B^{\mathrm{T}} \otimes A\right)_\sigma (I_{mn})_\sigma\, \varphi(X) \\ &= \left(\left(B^{\mathrm{T}} \otimes A\right) \bar{I}_{mn}\right)_\phi \varphi(X) \\ &= \left(B^{\mathrm{T}} \otimes A\right)_\phi \varphi(X). \quad \square \end{aligned}$$

The following lemma regarding properties of real representation is important in our further development.

Lemma 11. *Let $A = A_1 + iA_2 \in \mathbf{C}^{m\times m}$ where A_1 and A_2 are real matrices. Then*

$$\mathit{rank}\,(A_\sigma) = 2\mathit{rank}\,(A)\,, \tag{51}$$

$$\mathit{rank}\,(I_{2m} - A_\sigma) = m + \mathit{rank}\left(I_m - A\overline{A}\right). \tag{52}$$

Proof. We first show Eq. (51). By using Lemma 9, we have

$$\begin{aligned}\operatorname{rank}(A_\sigma) &= \operatorname{rank}(A_\sigma\,(I_m)_\sigma) = \operatorname{rank}(A_\phi)\\ &= \operatorname{rank}\left(\begin{bmatrix}\overline{A} & 0\\ 0 & A\end{bmatrix}\right)\\ &= \operatorname{rank}(A) + \operatorname{rank}\left(\overline{A}\right)\\ &= 2\operatorname{rank}(A)\,.\end{aligned}$$

We next show (52). Invoking of (47) and (46) we obtain

$$\begin{aligned}\operatorname{rank}(I_{2m} - A_\sigma) &= \operatorname{rank}((I_m)_\sigma - A_\sigma\,(I_m)_\sigma)\\ &= \operatorname{rank}((I_m)_\sigma - A_\phi)\\ &= \operatorname{rank}\left(Z_m^{\mathrm{H}}\,((I_m)_\sigma - A_\phi)\,Z_m\right)\\ &= \operatorname{rank}\left(Z_m^{\mathrm{H}}\,(I_m)_\sigma\,Z_m - \begin{bmatrix}\overline{A} & 0\\ 0 & A\end{bmatrix}\right)\\ &= \operatorname{rank}\left(\begin{bmatrix}-\overline{A} & -\mathrm{i}I_m\\ \mathrm{i}I_m & -A\end{bmatrix}\right)\\ &= \operatorname{rank}\left(\begin{bmatrix}-\overline{A} & -\mathrm{i}I_m\\ \mathrm{i}I_m & -A\end{bmatrix}\begin{bmatrix}0 & -\mathrm{i}I_m\\ I_m & \overline{A}\end{bmatrix}\right)\\ &= \operatorname{rank}\left(\begin{bmatrix}-\mathrm{i}I_m & 0\\ -A & I_m - A\overline{A}\end{bmatrix}\right)\\ &= m + \operatorname{rank}\left(I_m - A\overline{A}\right). \quad \square\end{aligned}$$

5.2. *Solvability of equation $X = A\overline{X}B + C$*

Based on linear equation theory, to study the inhomogeneous matrix equation (42), we should study the homogeneous matrix equation $X = A\overline{X}B$. Associated with this equation, we define a set $\mathscr{S}' \subset \mathbf{C}^{m\times n}$ as

$$\mathscr{S}' = \left\{X \in \mathbf{C}^{m\times n} : X = A\overline{X}B\right\}.$$

For any $X_1 \in \mathscr{S}'$ and $X_2 \in \mathscr{S}'$ and any $\alpha \in \mathbf{C}$, we have

$$(X_1 + \alpha X_2) - A\left(\overline{X_1 + \alpha X_2}\right)B = \alpha X_2 - \overline{\alpha}A\overline{X_2}B,$$

which is not zero in general, namely, $X_1 + \alpha X_2 \notin \mathscr{S}'$ in general. Therefore, $\mathscr{S}'$ is *not* a linear subspace contained in $\mathbf{C}^{m\times n}$ over the field $\mathbf{C}$. Alternatively, if we define a real set $\mathscr{S} \subset \mathbf{R}^{mn}$ as

$$\mathscr{S} = \left\{\varphi(X) : X = A\overline{X}B, X \in \mathbf{C}^{m\times n}\right\}.$$

then by using the fact that mapping $\varphi(\cdot)$ is bijective and (48), we have

$$\begin{aligned}\mathscr{S} &= \left\{\varphi(X) : \varphi(X) = \varphi\left(A\overline{X}B\right),\ X \in \mathbf{C}^{m\times n}\right\} \\ &= \left\{\varphi(X) : \varphi(X) = \left(B^{\mathrm{T}} \otimes A\right)_\sigma \varphi(X),\ X \in \mathbf{C}^{m\times n}\right\} \\ &= \left\{x : x = \left(B^{\mathrm{T}} \otimes A\right)_\sigma x,\ x \in \mathbf{R}^{2mn}\right\} \\ &= \left\{x : \left(I_{2mn} - \left(B^{\mathrm{T}} \otimes A\right)_\sigma\right) x = 0,\ x \in \mathbf{R}^{2mn}\right\},\end{aligned} \tag{53}$$

which is clearly a linear subspace contained in $\mathbf{R}^{2mn}$ over the field $\mathbf{R}$. Moreover, we can see from (53) that

$$\dim(\mathscr{S}) = 2mn - \operatorname{rank}\left(I_{2mn} - \left(B^{\mathrm{T}} \otimes A\right)_\sigma\right). \tag{54}$$

where $\dim(\mathscr{S})$ denotes the degree of freedom (measured by the maximal number of real parameters) in the solution. Therefore, $\mathscr{S}$ is more suitable for characterizing the general solutions of Eq. (42). In fact, if we let

$$\mathscr{X} = \left\{X \in \mathbf{C}^{m\times n} : X = A\overline{X}B + C\right\},$$

and $X^{\dagger}$ be any particular solution of Eq. (42), then

$$\mathscr{X} = \left\{X^{\dagger} + \varphi^{-1}(x) : \forall x \in \mathscr{S}\right\}.$$

To present our main result, we need an auxiliary result.

Lemma 12. *Eq. (42) has a unique solution for arbitrary C if and only if*

$$\eta\gamma \neq 1, \quad \forall \eta \in \lambda\left\{A\overline{A}\right\},\ \forall \gamma \in \lambda\left\{\overline{B}B\right\}. \tag{55}$$

Proof. Taking $\varphi(\cdot)$ on both sides of (42) and using (48) give

$$\left(I_{2mn} - \left(B^{\mathrm{T}} \otimes A\right)_\sigma\right)\varphi(X) = \varphi(C). \tag{56}$$

Therefore, Eq. (42) has a unique solution for arbitrary C if and only if

$$\operatorname{rank}\left(I_{2mn} - \left(B^{\mathrm{T}} \otimes A\right)_\sigma\right) = 2mn. \tag{57}$$

On the other hand, in view of Lemma 11, we can see that

$$\begin{aligned}\operatorname{rank}\left(I_{2mn} - \left(B^{\mathrm{T}} \otimes A\right)_\sigma\right) &= mn + \operatorname{rank}\left(I_{mn} - \left(B^{\mathrm{T}} \otimes A\right)\left(\overline{B^{\mathrm{T}} \otimes A}\right)\right) \\ &= mn + \operatorname{rank}\left(I_{mn} - \left(\overline{B}B\right)^{\mathrm{T}} \otimes A\overline{A}\right).\end{aligned} \tag{58}$$

Hence, (57) is true if and only if

$$\operatorname{rank}\left(I_{mn} - \left(\overline{B}B\right)^{\mathrm{T}} \otimes A\overline{A}\right) = mn,$$

or, equivalent, $1 \notin \lambda\left\{\left(\overline{B}B\right)^{\mathrm{T}} \otimes A\overline{A}\right\}$. This is further equivalent to (55). □

We then can give the following alternative conditions regarding solvability of Eq. (42), which can be regarded as the generalization of Theorem 1 to Eq. (42).

Theorem 4. *Matrix equation (42) is solvable if and only if the following standard Stein equation*

$$W = A\overline{A}W\overline{B}B + A\overline{C}B + C, \tag{59}$$

is solvable with unknown W. More specifically:

1. *If X is a solution of Eq. (42), then $W = X$ is a solution of Eq. (59).*
2. *If $\boldsymbol{W}$ is the general closed-form solution of Eq. (59), then the general closed-form solution of Eq. (42) is given by*

$$\boldsymbol{X} = \frac{1}{2}\left(\boldsymbol{W} + A\overline{\boldsymbol{W}}B + C\right). \tag{60}$$

3. *Eq. (42) has a unique solution $X^{\star}$ if and only if Eq. (59) has a unique solution $W^{\star}$. Moreover, $X^{\star} = W^{\star}$.*

Proof. *Proofs of Item 1 and Item 2.* Taking $\varphi(\cdot)$ on both sides of (59), using Lemma 10 and denoting $E = B^{\mathrm{T}} \otimes A$, we get

$$\begin{aligned}
\varphi(W) &= \varphi\left(A\overline{A}W\overline{B}B\right) + \varphi\left(A\overline{C}B\right) + \varphi(C) \\
&= E_{\sigma}\varphi\left(A\overline{W}B\right) + \left(I_{2mn} + E_{\sigma}\right)\varphi(C) \\
&= E_{\sigma}^{2}\varphi(W) + \left(I_{2mn} + E_{\sigma}\right)\varphi(C),
\end{aligned}$$

which is further equivalent to

$$\left(I_{2mn} - E_{\sigma}^{2}\right)\varphi(W) = \left(I_{2mn} + E_{\sigma}\right)\varphi(C). \tag{61}$$

Similarly, Eq. (42) is equivalent to (56) which can be rewritten as

$$\left(I_{2mn} - E_{\sigma}\right)\varphi(X) = \varphi(C), \tag{62}$$

and the expression in (60) is equivalent to

$$\varphi(\boldsymbol{X}) = \frac{1}{2}\left(I_{2mn} + E_{\sigma}\right)\varphi(\boldsymbol{W}) + \frac{1}{2}\varphi(C). \tag{63}$$

Since Eq. (61) is in the form of (15), Eq. (62) is in the form of (16) and expression (63) is in the form of (17), Item 1 and Item 2 follow from Proposition 1 directly.

Proof of Item 3. Notice that, by using Lemma 12, Eq. (42) has a unique solution $X^{\star}$ if and only if (55) is satisfied, which is also the necessary and sufficient condition for the existence of unique solution of Eq. (59). □

Remark 4. The difference between Theorem 1 and Theorem 4 is that the condition for the existence of a unique solution to Eq. (42) is the same as the condition for the existence of a unique solution to the standard Stein Eq. (59).

Remark 5. Different from Proposition 2, the solvability condition for Eq. (42) in Theorem 4 is related with solvability of Eq. (59) which has the same dimension as (42). Also, different from Proposition 2, we have given in Theorem 4 a necessary and sufficient condition for the existence of unique solution of Eq. (42).

The following corollary is then a consequence of Theorem 4.

Corollary 1. *Eq. (42) is solvable if and only if*

$$rank\left(I_{mn}-\left(\overline{B}B\right)^{T}\otimes A\overline{A}\right)=rank\left(\left[\begin{array}{cc} I_{mn}-\left(\overline{B}B\right)^{T}\otimes A\overline{A} & \left(B^{T}\otimes A\right)vec\left(\overline{C}\right)+vec\left(C\right)\end{array}\right]\right). \qquad (64)$$

If the above relation is satisfied, then the degree of freedom (measured by the maximal number of free real parameters), namely, $\dim(\mathscr{S})$, *in the solution is given by*

$$\dim(\mathscr{S})=mn-rank\left(I_{mn}-\left(\overline{B}B\right)^{T}\otimes A\overline{A}\right). \qquad (65)$$

5.3. Relationship between equations $X=A\overline{X}B+C$ and $Y=A_{\sigma}YB_{\sigma}+C_{\sigma}$

From Proposition 2 we can see easily that if Eq. (44) has a unique solution, then Eq. (42) must also have a unique solution. But the converse is not clear. In this subsection, by using the results obtained in the above two subsections, we are able to close this gap. Our result reveals some deep relationships between these two equation (42) and (44).

Theorem 5. *Eq. (42) has a unique solution $X^{\star}$ if and only if Eq. (44) has a unique solution $Y^{\star}$. Moreover, there holds $Y^{\star}=X_{\sigma}^{\star}$.*

Proof. Obviously, Eq. (44) has a unique solution for arbitrary C_{σ} if and only if

$$\alpha\beta\neq 1,\quad \forall\alpha\in\lambda\left\{A_{\sigma}\right\},\ \forall\beta\in\lambda\left\{B_{\sigma}\right\}. \qquad (66)$$

For an arbitrary matrix $A=A_1+\mathrm{i}A_2\in\mathbf{C}^{m\times m}$ where $A_1,A_2\in\mathbf{R}^{m\times m}$, we can compute

$$\begin{aligned}
\det\left(sI_{2m}-A_{\sigma}\right)&=\mathrm{i}^{m}\det\left(\begin{bmatrix} I_m & 0\\ -I_m & \mathrm{i}I_m\end{bmatrix}\right)\det\left(\begin{bmatrix} sI_m-A_1 & -A_2\\ -A_2 & sI_m+A_1\end{bmatrix}\right)\det\left(\begin{bmatrix} \frac{\mathrm{i}}{2}I_m & I_m\\ \frac{1}{2}I_m & \mathrm{i}I_m\end{bmatrix}\right)\\
&=\mathrm{i}^{m}\det\left(\begin{bmatrix} I_m & 0\\ -I_m & \mathrm{i}I_m\end{bmatrix}\begin{bmatrix} sI_m-A_1 & -A_2\\ -A_2 & sI_m+A_1\end{bmatrix}\begin{bmatrix} \frac{\mathrm{i}}{2}I_m & I_m\\ \frac{1}{2}I_m & \mathrm{i}I_m\end{bmatrix}\right)\\
&=\mathrm{i}^{m}\det\left(\begin{bmatrix} \frac{\mathrm{i}}{2}sI_m-\frac{\mathrm{i}}{2}\overline{A} & sI_m-A\\ \mathrm{i}\overline{A} & -2sI_m\end{bmatrix}\right)\\
&=\mathrm{i}^{m}\det\left(-2sI_m\right)\det\left(\frac{\mathrm{i}}{2}sI_m-\frac{\mathrm{i}}{2}\overline{A}-\left(sI_m-A\right)\left(-2sI_m\right)^{-1}\mathrm{i}\overline{A}\right)\\
&=\mathrm{i}^{m}\det\left(\left(-2sI_m\right)\left(\frac{\mathrm{i}}{2}sI_m-\frac{\mathrm{i}}{2}\overline{A}\right)-\left(sI_m-A\right)\mathrm{i}\overline{A}\right)\\
&=\mathrm{i}^{m}\det\left(-\mathrm{i}\left(s^2I_m-A\overline{A}\right)\right)\\
&=\det\left(s^2I_m-A\overline{A}\right)\\
&=\det\left(s^2I_m-\overline{A}A\right).
\end{aligned} \qquad (67)$$

where we have used Lemma 3.

Let α and β be two arbitrary elements in $\lambda\left\{A_{\sigma}\right\}$ and $\lambda\left\{B_{\sigma}\right\}$, respectively. Then it follows from Lemma 8 that $\left\{\pm\alpha,\pm\overline{\alpha}\right\}\subset\lambda\left\{A_{\sigma}\right\}$, $\left\{\pm\beta,\pm\overline{\beta}\right\}\subset\lambda\left\{B_{\sigma}\right\}$ and it follows from (67) that $\left\{\alpha^2,\overline{\alpha}^2\right\}\subset\lambda\left\{A\overline{A}\right\}$, $\left\{\beta^2,\overline{\beta}^2\right\}\subset\lambda\left\{\overline{B}B\right\}$. Direct manipulation shows that

$$st\neq 1,\quad \forall s\in\left\{\pm\alpha,\pm\overline{\alpha}\right\}\subset\lambda\left\{A_{\sigma}\right\},\ \forall t\in\left\{\pm\beta,\pm\overline{\beta}\right\}\subset\lambda\left\{B_{\sigma}\right\},$$

if and only if $\alpha\beta \neq \pm 1$ and $\alpha\overline{\beta} \neq \pm 1$. Also, direct computation indicates that

$$st \neq 1, \quad \forall s \in \left\{\alpha^2, \overline{\alpha}^2\right\} \subset \lambda\left\{A\overline{A}\right\}, \ \forall t \in \{\beta^2, \overline{\beta}^2\} \subset \lambda\{\overline{B}B\},$$

if and only if $\alpha\beta \neq \pm 1$ and $\alpha\overline{\beta} \neq \pm 1$. According to the arbitrariness of α and β, we conclude from the above that (55) and (66) are equivalent.

Finally, it is trivial to show that $Y^\star = X_\sigma^\star$. $\square$

Proposition 2 and Theorem 5 are still unable to answer the following question: if $\boldsymbol{Y}$ is the general real solution of equation (44), then is the following

$$\boldsymbol{X}' \triangleq \frac{1}{4}\begin{bmatrix} I_m & \mathrm{i}I_m \end{bmatrix}(\boldsymbol{Y} + Q_m\boldsymbol{Y}Q_n)\begin{bmatrix} I_n \\ \mathrm{i}I_n \end{bmatrix} \tag{68}$$

the general solution of Eq. (42)? The following example may indicate that the answer is positive.

Example 3. Consider a linear equation in the form of (42) with

$$A = \begin{bmatrix} 2 & 0 \\ 0 & \mathrm{i} \end{bmatrix}, \quad B = 1, \quad C = \begin{bmatrix} C_1 \\ C_2 \end{bmatrix} = \begin{bmatrix} C_{11} + \mathrm{i}C_{12} \\ C_{21} + \mathrm{i}C_{22} \end{bmatrix},$$

where $C_{ij}, i, j = 1, 2$ are real scalars. Then Eq. (59) can be written as

$$W = \begin{bmatrix} w_1 \\ w_2 \end{bmatrix} = A\overline{A}W\overline{B}B + A\overline{C}B + C = \begin{bmatrix} 4w_1 + 2\overline{C_1} + C_1 \\ w_2 + \mathrm{i}\overline{C_2} + C_2 \end{bmatrix},$$

which is solvable if and only if $\mathrm{i}\overline{C_2} + C_2 = 0$. Consequently, the general solution of Eq. (59) is

$$\boldsymbol{W} = \begin{bmatrix} -\frac{1}{3}\left(2\overline{C_1} + C_1\right) \\ t \end{bmatrix}, \quad \forall t = t_1 + \mathrm{i}t_2 \in \mathbf{C},$$

where t_1 and t_2 are real scalars. Hence, according to Theorem 4, the general solution of Eq. (42) is given by

$$\boldsymbol{X} = \frac{1}{2}\left(\boldsymbol{W} + A\overline{\boldsymbol{W}}B + C\right) = \begin{bmatrix} -\frac{1}{3}\left(C_1 + 2\overline{C_1}\right) \\ \frac{1}{2}(1 + \mathrm{i})(t_1 + t_2) + \frac{1}{2}C_2 \end{bmatrix}.$$

On the other hand, direct computation shows that the general real solution of Eq. (44) is

$$\boldsymbol{Y} = \begin{bmatrix} 0 & 0 \\ \frac{1}{2}s & \frac{1}{2}s \\ 0 & 0 \\ \frac{1}{2}s & -\frac{1}{2}s \end{bmatrix} + \begin{bmatrix} -C_{11} & \frac{1}{3}C_{12} \\ \frac{1}{2}C_{21} & \frac{1}{2}C_{22} \\ \frac{1}{3}C_{12} & C_{11} \\ \frac{1}{2}C_{22} & -\frac{1}{2}C_{21} \end{bmatrix}, \quad \forall s \in \mathbf{R}.$$

Then $\boldsymbol{X}'$ defined in (68) can be computed as

$$\boldsymbol{X}' = \begin{bmatrix} -\frac{1}{3}\left(C_1 + 2\overline{C_1}\right) \\ \frac{1}{2}C_2 \end{bmatrix} + \begin{bmatrix} 0 \\ \frac{1}{2}(1 + \mathrm{i})s \end{bmatrix}.$$

It is clear that $\boldsymbol{X}' = \boldsymbol{X}$ if we set $t_1 + t_2 = s$ in $\boldsymbol{X}$.

However, at present we cannot prove the correctness the above result in general. This would be a conjecture that needs further study.

Conjecture 1. *If $\boldsymbol{Y}$ is the general real solution of Eq. (42), then the general solution of Eq. (42) is $\boldsymbol{X}'$ defined in (68).*

We should point out that even though the above conjecture is true, compared with the results in Theorem 4, using solutions of Eq. (44) to construct solution of Eq. (42) has no advantage since the dimensions of Eq. (44) has doubled the dimensions of equation (59) used in Theorem 4.

5.4. Closed-form and iterative solutions of equation $X = A\overline{X}B + C$

By combining Lemma 5 and Theorem 4 we can derive the following corollary regarding closed-from solutions of Eq. (42). We notice that this result is exactly Theorem 4.4 in [13] where the real representation method is developed to derive such result. In contrast, in this paper such result is only a consequence of Theorem 4 and Lemma 5 without using the intricate properties of the real representation.

Corollary 2. *Let $A \in \mathbf{C}^{m\times m}$, $B \in \mathbf{C}^{n\times n}$, $C \in \mathbf{C}^{m\times n}$ be given and*

$$h_{A\overline{A}}(s) = \det\left(I_m - sA\overline{A}\right) = \alpha_m + \sum_{i=1}^{m} \alpha_{m-i} s^i, \quad \alpha_m = 1.$$

1. *If Eq. (42) has a solution X, then*

$$Xh_{A\overline{A}}\left(\overline{B}B\right) = \sum_{k=1}^{m}\sum_{s=1}^{k} \alpha_k \left(A\overline{A}\right)^{k-s}\left(A\overline{C}B + C\right)\left(\overline{B}B\right)^{m-s}.$$

2. *If Eq. (42) has a unique solution $X^\star$, then*

$$X^\star = \left(\sum_{k=1}^{m}\sum_{s=1}^{k} \alpha_k \left(A\overline{A}\right)^{k-s}\left(A\overline{C}B + C\right)\left(\overline{B}B\right)^{m-s}\right)\left(h_{A\overline{A}}\left(\overline{B}B\right)\right)^{-1}.$$

Similar to the Smith Iteration for Eq. (26), we can also construct the Smith iteration for Eq. (42) as follows:

$$X_{k+1} = A\overline{X_k}B + C, \quad \forall X_0 \in \mathbf{C}^{m\times n}. \tag{69}$$

The following theorem is concerned with the convergence of this iteration.

Theorem 6. *The Smith iteration (69) converges to the unique solution $X^\star$ of Eq. (42) for arbitrary initial condition X_0 if and only if*

$$\rho\left(A\overline{A}\right)\rho\left(\overline{B}B\right) < 1. \tag{70}$$

Moreover, the asymptotic exponential convergence rate is $-\ln\left(\rho\left(A\overline{A}\right)\rho\left(\overline{B}B\right)\right)$.

Proof. We first show that if the iteration (69) converges, then (70) should be satisfied. As $\{X_k\}_{k=0}^{\infty}$ converges, we know that the subsequence $\{X_{2i}\}_{i=0}^{\infty}$ also converges. It is easy to verify that $\{X_{2i}\}_{i=0}^{\infty}$ is generated by the following iteration

$$X_{2(i+1)} = A\bar{A}X_{2i}\bar{B}B + A\bar{C}B + C, \quad \forall X_0 \in \mathbf{C}^{m\times n}.$$

Notice that the above iteration is a standard Smith iteration which converges to a constant matrix independent of the initial condition if and only if (70) is true.

We next show that if (70) is satisfied, then the iteration (69) must converge. Notice that the iteration (69) can be equivalently rewritten as

$$(X_{k+1})_\sigma = A_\sigma \, (X_k)_\sigma \, B_\sigma + C_\sigma, \quad \forall \, (X_0)_\sigma \in \mathbf{R}^{2m\times 2n}. \tag{71}$$

From (67) we can see that

$$\rho\,(A_\sigma)\,\rho\,(B_\sigma) = \sqrt{\rho\left(\bar{A}A\right)}\sqrt{\rho\,(\bar{B}B)} < 1,$$

which indicates that (71) converges to a constant matrix $(X_\infty)_\sigma$ independent of initial condition, namely, $(X_\infty)_\sigma = A_\sigma \, (X_\infty)_\sigma \, B_\sigma + C_\sigma$, or equivalently, $X_\infty = A\overline{X_\infty}B + C$. The proof is completed by observing that Eq. (42) has a unique solution under the condition (70). □

Remark 6. As condition (70) implies condition (55), it follows from Item 3 of Theorem 4 that the unique solution $X^\star$ to Eq. (42) under condition (70) is also the unique solution of Eq. (59), namely,

$$X^\star = W^\star = \sum_{i=0}^{\infty}\left(A\bar{A}\right)^i\left(A\bar{C}B + C\right)(\bar{B}B)^i. \tag{72}$$

Therefore, similar to Remarks 2 and 3, we can also construct Smith (l) iteration and r-Smith iteration for solving equation (42). The details are omitted for brevity.

6. Equation $X = AX^{\mathrm{H}}B + C$

In this section, we study the following matrix equation:

$$X = AX^{\mathrm{H}}B + C, \tag{73}$$

where $A \in \mathbf{C}^{m\times n}$, $B \in \mathbf{C}^{m\times n}$ and $C \in \mathbf{C}^{m\times n}$ are given, and $X \in \mathbf{C}^{m\times n}$ is unknown. In this section, we will generalize the results for Eqs. (26) and (42) to this equation.

6.1. Solvability of equation $X = AX^H B + C$

We first show a result parallel to Lemma 12.

Lemma 13. *Eq. (73) has a unique solution for arbitrary C if and only if*

$$\bar{\eta}\gamma \neq 1, \quad \forall \eta, \ \gamma \in \lambda\left\{AB^H\right\}. \tag{74}$$

Proof. Notice that, for any complex matrix $X \in \mathbf{C}^{m\times n}$, we have

$$\varphi\left(X^{\mathrm{T}}\right) = \begin{bmatrix} \mathrm{vec}\left(X_1^{\mathrm{T}}\right) \\ \mathrm{vec}\left(X_2^{\mathrm{T}}\right) \end{bmatrix} = \begin{bmatrix} P_{(m,n)}\mathrm{vec}\,(X_1) \\ P_{(m,n)}\mathrm{vec}\,(X_2) \end{bmatrix} = \mathcal{P}_{(m,n)}\varphi\,(X)\,,$$

where

$$\mathcal{P}_{(m,n)} = \begin{bmatrix} P_{(m,n)} & 0 \\ 0 & P_{(m,n)} \end{bmatrix}. \tag{75}$$

With this, by taking $\varphi(\cdot)$ on both sides of (73), using Lemma 10, and denoting $E = B^{\mathrm{T}} \otimes A$, we have

$$\begin{aligned} \varphi(X) &= \varphi\left(AX^{\mathrm{H}}B\right) + \varphi(C) \\ &= \left(B^{\mathrm{T}} \otimes A\right)_{\sigma} \varphi\left(X^{\mathrm{T}}\right) + \varphi(C) \\ &= E_{\sigma}\mathcal{P}_{(m,n)}\varphi(X) + \varphi(C). \end{aligned}$$

Therefore, Eq. (73) has a unique solution for arbitrary C if and only if

$$\operatorname{rank}\left(I_{2mn} - E_{\sigma}\mathcal{P}_{(m,n)}\right) = 2mn. \tag{76}$$

Similar to the proof of relation (52), by using Lemma 9, we can compute

$$\begin{aligned} \operatorname{rank}\left(I_{2mn} - E_{\sigma}\mathcal{P}_{(m,n)}\right) &= \operatorname{rank}\left(\mathcal{P}_{(n,m)} - E_{\sigma}\right) \\ &= \operatorname{rank}\left(\mathcal{P}_{(n,m)}\left(I_{mn}\right)_{\sigma} - E_{\sigma}\left(I_{mn}\right)_{\sigma}\right) \\ &= \operatorname{rank}\left(\mathcal{P}_{(n,m)}\left(I_{mn}\right)_{\sigma} - E_{\phi}\right) \\ &= \operatorname{rank}\left(Z_{mn}^{\mathrm{H}}\left(\mathcal{P}_{(n,m)}\left(I_{mn}\right)_{\sigma} - E_{\phi}\right)Z_{mn}\right), \\ &= \operatorname{rank}\left(Z_{mn}^{\mathrm{H}}\mathcal{P}_{(n,m)}\left(I_{mn}\right)_{\sigma}Z_{mn} - \begin{bmatrix} \overline{E} & 0 \\ 0 & E \end{bmatrix}\right) \\ &= \operatorname{rank}\left(\Omega - \begin{bmatrix} P_{(m,n)}\overline{E}P_{(m,n)} & 0 \\ 0 & E \end{bmatrix}\right), \end{aligned} \tag{77}$$

where

$$\begin{aligned} \Omega &\triangleq \begin{bmatrix} P_{(m,n)} & 0 \\ 0 & I_{mn} \end{bmatrix} Z_{mn}^{\mathrm{H}}\mathcal{P}_{(n,m)}\left(I_{mn}\right)_{\sigma}Z_{mn} \begin{bmatrix} P_{(m,n)} & 0 \\ 0 & I_{mn} \end{bmatrix} \\ &= \begin{bmatrix} 0 & -\mathrm{i}I_{mn} \\ \mathrm{i}I_{mn} & 0 \end{bmatrix}. \end{aligned}$$

Hence, we get from (77) that

$$\begin{aligned} \operatorname{rank}\left(I_{2mn} - E_{\sigma}\mathcal{P}_{(m,n)}\right) &= \operatorname{rank}\left(\begin{bmatrix} -P_{(m,n)}\overline{E}P_{(m,n)} & -\mathrm{i}I_{mn} \\ \mathrm{i}I_{mn} & -E \end{bmatrix}\right) \\ &= \operatorname{rank}\left(\begin{bmatrix} -P_{(m,n)}\overline{E}P_{(m,n)} & -\mathrm{i}I_{mn} \\ \mathrm{i}I_{mn} & -E \end{bmatrix}\begin{bmatrix} 0 & -\mathrm{i}I_{mn} \\ I_{mn} & P_{(m,n)}\overline{E}P_{(m,n)} \end{bmatrix}\right) \\ &= \operatorname{rank}\left(\begin{bmatrix} -\mathrm{i}I_{mn} & 0 \\ -E & I_{mn} - EP_{(m,n)}\overline{E}P_{(m,n)} \end{bmatrix}\right) \\ &= mn + \operatorname{rank}\left(I_{mn} - EP_{(m,n)}\overline{E}P_{(m,n)}\right). \end{aligned} \tag{78}$$

Notice that

$$\begin{aligned} EP_{(m,n)}\overline{E}P_{(m,n)} &= EP_{(m,n)}\overline{\left(B^{\mathrm{T}}\otimes A\right)}P_{(m,n)} \\ &= EP_{(m,n)}\left(B^{\mathrm{H}}\otimes \overline{A}\right)P_{(m,n)} \\ &= \left(B^{\mathrm{T}}\otimes A\right)\left(\overline{A}\otimes B^{\mathrm{H}}\right) \\ &= B^{\mathrm{T}}\overline{A}\otimes AB^{\mathrm{H}} \\ &= \left(A^{\mathrm{H}}B\right)^{\mathrm{T}}\otimes AB^{\mathrm{H}}. \end{aligned}$$

Hence it follows from (78) that condition (76) is true if and only if

$$1 \notin \lambda\left\{\left(A^{\mathrm{H}}B\right)^{\mathrm{T}}\otimes AB^{\mathrm{H}}\right\} \tag{79}$$

which is equivalent to (74) because

$$\begin{aligned} \lambda\left\{\left(A^{\mathrm{H}}B\right)^{\mathrm{T}}\otimes AB^{\mathrm{H}}\right\} &= \lambda\left\{B^{\mathrm{T}}\overline{A}\otimes AB^{\mathrm{H}}\right\} \\ &= \lambda\left\{\overline{AB^{\mathrm{H}}}\otimes AB^{\mathrm{H}}\right\} \\ &= \left\{\overline{\eta}\gamma,\ \ \forall \eta,\gamma\in\lambda\left\{AB^{\mathrm{H}}\right\}\right\}. \quad \square \end{aligned}$$

We are now able to present the results regarding the solvability of Eq. (73).

Theorem 7. *Matrix equation (73) is solvable if and only if the following standard Stein equation*

$$W = AB^{H}WA^{H}B + AC^{H}B + C, \tag{80}$$

is solvable with unknown W. More specifically,

1. *If X is a solution of Eq. (73), then $W = X$ is also a solution of Eq. (80).*
2. *If $\boldsymbol{W}$ is the general solution of Eq. (80), then the general solution to Eq. (73) is*

$$\boldsymbol{X} = \frac{1}{2}\left(\boldsymbol{W} + A\boldsymbol{W}^{H}B + C\right). \tag{81}$$

3. *Eq. (73) has a unique solution $X^{\star}$ if and only if Eq. (80) has a unique solution $W^{\star}$, namely, (74) is satisfied. Moreover, $X^{\star} = W^{\star}$.*

Proof. *Proofs of Item 1 and Item 2.* For any matrices $A \in \mathbf{C}^{m\times n}$, $P \in \mathbf{R}^{p\times m}$ and $Q \in \mathbf{R}^{n\times q}$, it cam be readily verified that

$$(PAQ)_{\sigma} = \begin{bmatrix} P & 0 \\ 0 & P \end{bmatrix} A_{\sigma} \begin{bmatrix} Q & 0 \\ 0 & Q \end{bmatrix}.$$

By using this fact, Lemmas 1 and 10, we obtain

$$\begin{aligned} \varphi(W) &= \varphi\left(A\overline{B}^{\mathrm{T}}W\overline{A}^{\mathrm{T}}B\right) + \varphi\left(A\overline{C}^{\mathrm{T}}B\right) + \varphi(C) \\ &= E_{\sigma}\varphi\left(B^{\mathrm{T}}\overline{W}A^{\mathrm{T}}\right) + E_{\sigma}\varphi\left(C^{\mathrm{T}}\right) + \varphi(C) \\ &= E_{\sigma}\left(A\otimes B^{\mathrm{T}}\right)_{\sigma}\varphi(W) + \left(I_{2mn} + E_{\sigma}\mathcal{P}_{(m,n)}\right)\varphi(C) \\ &= E_{\sigma}\left(P_{(m,n)}\left(B^{\mathrm{T}}\otimes A\right)P_{(m,n)}\right)_{\sigma}\varphi(W) + \left(I_{2mn} + E_{\sigma}\mathcal{P}_{(m,n)}\right)\varphi(C) \\ &= E_{\sigma}\mathcal{P}_{(m,n)}E_{\sigma}\mathcal{P}_{(m,n)}\varphi(W) + \left(I_{2mn} + E_{\sigma}\mathcal{P}_{(m,n)}\right)\varphi(C), \end{aligned}$$

where $E = B^{\mathrm{T}} \otimes A$ and $\mathcal{P}_{(m,n)}$ is defined in (75). It follows that

$$\left(I_{2mn} - \left(E_\sigma \mathcal{P}_{(m,n)}\right)^2\right) \varphi\left(W\right) = \left(I_{2mn} + E_\sigma \mathcal{P}_{(m,n)}\right) \varphi\left(C\right).$$

Similarly, Eq. (73) and expression (81) are, respectively, equivalent to

$$\left(I_{2mn} - E_\sigma \mathcal{P}_{(m,n)}\right) = \varphi\left(C\right),$$

and

$$\varphi\left(\boldsymbol{X}\right) = \frac{1}{2}\left(I_{2mn} + E_\sigma \mathcal{P}_{(m,n)}\right) \varphi\left(\boldsymbol{W}\right) + \frac{1}{2}\varphi\left(C\right).$$

The remaining of the proof is similar to the proof of Theorem 1 and is omitted for brevity.

Proof of Item 3. Clearly, Eq. (80) has a unique solution if and only if (79) is satisfied. By using Lemma 13, (79) is equivalent to (74) which is just the condition for the existence of unique solution of Eq. (73). The remaining of the proof is similar to the proof of Item 3 of Theorem 1. □

The advantage of Theorem 7 is that to solve Eq. (73), we need only to consider the standard Stein Eq. (80) which is again in the form of (9).

Similarly to the set $\mathscr{S}$ defined in (53), if we let

$$\mathscr{T} = \left\{x : \left(I_{2mn} - E_\sigma \mathcal{P}_{(m,n)}\right) x = 0,\ x \in \mathbf{R}^{2mn}\right\},$$

which is a linear subspace contained in $\mathbf{R}^{2mn}$ over the field $\mathbf{R}$, then we can obtain the following corollary that parallels Corollary 1.

Corollary 3. *Eq. (73) is solvable if and only if*

$$\mathit{rank}\left(I_{mn} - \left(A^H B\right)^T \otimes AB^H\right) = \mathit{rank}\left(\left[\ I_{mn} - \left(A^H B\right)^T \otimes AB^H\ \ \left(B^T \otimes A\right) \mathit{vec}\left(C^H\right) + \mathit{vec}\left(C\right)\ \right]\right).$$

If the above relation is satisfied, then the degree of freedom (measured by the maximal number of free real parameters), namely, $\dim\left(\mathscr{T}\right)$, *in the solution is given by*

$$\dim\left(\mathscr{T}\right) = mn - \mathit{rank}\left(I_{mn} - \left(A^H B\right)^T \otimes AB^H\right).$$

6.2. Solvability conditions based on real representation

In this subsection, by using real representation of complex matrices, we generalize the results in [13] (namely, Proposition 2) to Eq. (73).

Theorem 8. *Eq. (73) is solvable if and only if equation*

$$Y = A_\sigma Y^T B_\sigma + C_\sigma, \tag{82}$$

is solvable with unknown Y. More specifically,

1. *If X is a solution of Eq. (73), then $Y = X_\sigma$ is a solution of Eq. (82).*
2. *If Y is a real solution of (82), then a solution of Eq. (73) is*

$$X = \frac{1}{4}\left[\ I_m\ \ iI_m\ \right]\left(Y + Q_m Y Q_n\right)\begin{bmatrix} I_n \\ iI_n \end{bmatrix}.$$

3. *Eq. (73) has a unique solution* $X^{\star}$ *if Eq. (82) has a unique solution* $Y^{\star}$, *namely* $1 \notin \lambda\left\{\left(B_{\sigma}^{T} \otimes A_{\sigma}\right) P_{(2m,2n)}\right\}$. *Moreover,* $Y^{\star} = X_{\sigma}^{\star}$.

Proof. *Proof of Item 1.* Let $A = A_1 + \mathrm{i}A_2$, $B = B_1 + \mathrm{i}B_2$ and $C = C_1 + \mathrm{i}C_2$ where A_i, B_i and C_i, $i = 1, 2$ are real matrices. Then by substituting $X = X_1 + \mathrm{i}X_2$ where $X_1, X_2 \in \mathbf{R}^{m\times n}$ into (73), we obtain

$$\begin{cases} X_1 = \left(A_1X_1^{\mathrm{T}} + A_2X_2^{\mathrm{T}}\right) B_1 - \left(A_2X_1^{\mathrm{T}} - A_1X_2^{\mathrm{T}}\right) B_2 + C_1, \\ X_2 = \left(A_2X_1^{\mathrm{T}} - A_1X_2^{\mathrm{T}}\right) B_1 + \left(A_1X_1^{\mathrm{T}} + A_2X_2^{\mathrm{T}}\right) B_2 + C_2, \end{cases}$$

which is also equivalent to

$$\begin{bmatrix} X_1 & X_2 \\ X_2 & -X_1 \end{bmatrix} = \begin{bmatrix} A_1 & A_2 \\ A_2 & -A_1 \end{bmatrix} \begin{bmatrix} X_1 & X_2 \\ X_2 & -X_1 \end{bmatrix}^{\mathrm{T}} \begin{bmatrix} B_1 & B_2 \\ B_2 & -B_1 \end{bmatrix} + \begin{bmatrix} C_1 & C_2 \\ C_2 & -C_1 \end{bmatrix}.$$

This indicates that $Y = X_{\sigma}$ satisfies (82) in view of the real representation (43).

Proof of Item 2. Let

$$Y = \begin{bmatrix} Y_1 & Y_2 \\ Y_3 & Y_4 \end{bmatrix}, \quad Y_i \in \mathbf{R}^{m\times n}, \ i = 1, 2, 3, 4,$$

be a real solution of Eq. (82). Note that $Q_m A_{\sigma} Q_n = A_{\sigma}$ [13] and

$$Q_m^{-1} = \begin{bmatrix} 0 & I_m \\ -I_m & 0 \end{bmatrix}^{-1} = \begin{bmatrix} 0 & -I_m \\ I_m & 0 \end{bmatrix} = Q_m^{\mathrm{T}}.$$

Then by multiplying the left hand side and right hand side of (82) by, respectively, Q_m and Q_n, we obtain

$$\begin{aligned} Q_mYQ_n &= Q_mA_{\sigma}Y^{\mathrm{T}}B_{\sigma}Q_n + Q_mC_{\sigma}Q_n \\ &= Q_mA_{\sigma}Q_nQ_n^{-1}Y^{\mathrm{T}}Q_m^{-1}Q_mB_{\sigma}Q_n + Q_mC_{\sigma}Q_n \\ &= A_{\sigma}Q_n^{\mathrm{T}}Y^{\mathrm{T}}Q_m^{\mathrm{T}}B_{\sigma} + C_{\sigma} \\ &= A_{\sigma}\left(Q_mYQ_n\right)^{\mathrm{T}}B_{\sigma} + C_{\sigma}. \end{aligned}$$

This indicates that Q_mYQ_n is also a solution of Eq. (82). Consequently,

$$\begin{aligned} \mathcal{Y} &= \frac{1}{2}\left(Y + Q_mYQ_n\right) \\ &= \frac{1}{2}\begin{bmatrix} Y_1 - Y_4 & Y_2 + Y_3 \\ Y_3 + Y_2 & -\left(Y_1 - Y_4\right) \end{bmatrix} \\ &\triangleq \begin{bmatrix} X_1 & X_2 \\ X_2 & -X_1 \end{bmatrix}, \end{aligned}$$

is also a solution of Eq. (82). Therefore, the following matrix

$$X \triangleq X_1 + \mathrm{i}X_2 = \frac{1}{2}\begin{bmatrix} I_m & \mathrm{i}I_m \end{bmatrix}\mathcal{Y}\begin{bmatrix} I_n \\ \mathrm{i}I_n \end{bmatrix},$$

is a solution of Eq. (73).

Proof of Item 3. We need only to show that if Eq. (82) has a unique solution then Eq. (73) also has a unique solution. We show this by contradiction. Assume that Eq. (73) has at least two solutions, say, X^1 and X^2 with $X^1 \neq X^2$. Then, according Item 1 of this theorem, both X_σ^1 and X_σ^2 are solutions of Eq. (82). Clearly, we have $X_\sigma^1 \neq X_\sigma^2$ which yields a contradiction. □

The following example indicates that the converse of Item 3 of Theorem 8 may also be true.

Example 4. Consider a linear equation in the form of (73) with A and B given in (38). Then Eq. (82) has a unique solution if and only if

$$\begin{aligned} 1 &\notin \lambda \left\{ \left(B_\sigma^{\mathrm{T}} \otimes A_\sigma \right) P_{(2m,2n)} \right\} \\ &= \left\{ \pm 2, \pm\sqrt{|\alpha|}, \alpha, \overline{\alpha}, \pm\sqrt{2\alpha}, \pm\sqrt{2\overline{\alpha}} \right\}, \end{aligned}$$

namely, $|\alpha| \neq 1$ and $\alpha \neq \frac{1}{2}$. On the other hand, since $\lambda \left\{ AB^{\mathrm{H}} \right\} = \{2, \alpha\}$, the condition in (74) is also equivalent to $|\alpha| \neq 1$ and $\alpha \neq \frac{1}{2}$. This implies, for A and B given in (38), that Eq. (82) has a unique solution if and only if Eq. (73) has a unique solution.

However, we also cannot prove the above statement in general. We state it as a conjecture which can be regarded as the generalization of Theorem 5 to Eqs. (73) and (82).

Conjecture 2. *If Eq. (73) has a unique solution $X^\star$, then Eq. (82) has a unique solution $Y^\star$. Moreover, there holds $Y^\star = X_\sigma^\star$.*

Even though the above conjecture is true, compared with the results in Subsection 6.1, there is no advantage by transforming Eq. (73) into Eq. (82) via its real representation.

6.3. Closed-form and iterative solutions of equation $X = AX^H B + C$

Based on Theorem 7, we can extend Corollary 2 to Eq. (73) without a proof.

Theorem 9. *Let $A \in \mathbf{C}^{m\times n}$, $B \in \mathbf{C}^{m\times n}$, $C \in \mathbf{C}^{m\times n}$ and*

$$h_{AB^H}(s) = \det\left(I_m - sAB^H \right) = \alpha_m + \sum_{i=1}^{m} \alpha_{m-i} s^i, \quad \alpha_m = 1.$$

1. *If Eq. (73) has a solution X, then*

$$X h_{AB^H}\left(A^H B \right) = \sum_{k=1}^{m} \sum_{s=1}^{k} \alpha_k \left(AB^H \right)^{k-s} \left(AC^H B + C \right) \left(A^H B \right)^{m-s}.$$

2. *If Eq. (73) has a unique solution $X^\star$, then*

$$X^\star = \left(\sum_{k=1}^{m} \sum_{s=1}^{k} \alpha_k \left(AB^H \right)^{k-s} \left(AC^H B + C \right) \left(A^H B \right)^{m-s} \right) \left(h_{AB^H}\left(A^H B \right) \right)^{-1}.$$

We next consider the Smith iteration associated with Eq. (73). Similar to (39), this iteration is in the form of

$$X_{k+1} = AX_k^{\mathrm{H}} B + C, \quad \forall X_0 \in \mathbf{C}^{m\times n}. \tag{83}$$

The following result is about the convergence of the above iteration.

Theorem 10. *The Smith iteration (83) converges to the unique solution $X^{\star}$ of Eq. (73) for arbitrary initial condition X_0 if and only if*

$$\rho\left(B^{H}A\right) < 1. \tag{84}$$

Moreover, the asymptotic exponential convergence rate is $-\ln\left(\rho\left(B^{H}A\right)\right)$.

Proof. Similar to the proof of Theorem 6, we can show that if (83) converges, we must have (84). Therefore, we need only to show that if (84) is satisfied, the iteration in (83) must converge to the unique solution of Eq. (73).

By using the real representation and Theorem 3, the iteration in (83) is equivalent to

$$(X_{k+1})_{\sigma} = A_{\sigma}\left((X_k)_{\sigma}\right)^{\mathrm{T}} B_{\sigma} + C_{\sigma}. \tag{85}$$

According to Theorem 3, the above iteration converges to a constant matrix independent of the initial condition if and only if $\rho\left(A_{\sigma}B_{\sigma}^{\mathrm{T}}\right) < 1$. According to Lemma 9, we have

$$\begin{aligned}\rho\left(A_{\sigma}B_{\sigma}^{\mathrm{T}}\right) &= \rho\left(\left(AB^{\mathrm{H}}\right)_{\phi}\right)\\ &= \max\left\{\rho\left(AB^{\mathrm{H}}\right), \rho\left(\overline{AB^{\mathrm{H}}}\right)\right\}\\ &= \rho\left(AB^{\mathrm{H}}\right).\end{aligned}$$

Hence, if (84) is satisfied, iteration (85) converges to a constant matrix X_{∞} satisfying $(X_{\infty})_{\sigma} = A_{\sigma}\left((X_{\infty})_{\sigma}\right)^{\mathrm{T}} B_{\sigma} + C_{\sigma}$, or equivalently, $X_{\infty} = AX_{\infty}^{\mathrm{H}}B + C$. The proof is completed by observing that the Eq. (73) has a unique solution under condition (84). □

Remark 7. As condition (84) implies condition (74), it follows from Item 3 of Theorem 7 that the unique solution of Eq. (73) under condition (84) is also the unique solution of Eq. (80), namely,

$$X^{\star} = W^{\star} = \sum_{i=0}^{\infty}\left(AA^{\mathrm{H}}\right)^{i}\left(AC^{\mathrm{H}}B + C\right)\left(B^{\mathrm{H}}B\right)^{i}.$$

Therefore, Remarks 2 and 3 are also applicable to Eq. (73).

We finally use a simple numerical example to end this section.

Example 5. Consider a linear equation in the form of (73) with

$$A = \begin{bmatrix} 1 & 1+\mathrm{i} & 1 \\ -2 & \mathrm{i} & -\mathrm{i} \\ 1-\mathrm{i} & 0 & -1 \end{bmatrix}, \quad B = \begin{bmatrix} \mathrm{i} & 1 & -1 \\ 0 & \mathrm{i} & 2+\mathrm{i} \\ 1+\mathrm{i} & 3 & -\mathrm{i} \end{bmatrix},$$

and

$$C = \begin{bmatrix} -5+\mathrm{i} & -4-\mathrm{i} & -5-12\mathrm{i} \\ 2-\mathrm{i} & -4-2\mathrm{i} & 6+8\mathrm{i} \\ 1+3\mathrm{i} & 15-5\mathrm{i} & -4-5\mathrm{i} \end{bmatrix}.$$

It is readily to verify that the eigenvalue set of AB^{H} satisfies the condition in (74). Then it follows from Theorem 7 that the matrix equation in (73) has a unique solution which can be obtained by using Theorem 9 as follows:

$$X = \begin{bmatrix} 1+3\mathrm{i} & -2 & 0 \\ 1 & 2-\mathrm{i} & 1 \\ -2 & 2 & 2+\mathrm{i} \end{bmatrix}.$$

7. A general equation

In this section, we point out that the results obtained above can be easily extended to a general class of matrix equations having more terms on the right hand side, say,

$$X = \sum_{i=0}^{N} A_i f(X) B_i + C, \tag{86}$$

where $N \geqslant 0$, $A_i, B_i, i \in \{0, 1, \ldots, N\}$ and C are given, and X is to be determined. In fact, by using Proposition 1 and a similar technique used in Section 5, we can extend Theorem 4 to Eq. (86), as stated in the following theorem.

Theorem 11. *Matrix equation (86) is solvable if and only if the following equation*

$$W = \sum_{i=0}^{N} A_i \left(\sum_{k=0}^{N} f(A_k f(W) B_k) \right) B_i + \sum_{i=0}^{N} A_i f(C) B_i + C, \tag{87}$$

is solvable with unknown W. More specifically:

1. *If X is a solution of Eq. (86), then $W = X$ is a solution of Eq. (87).*
2. *If $\boldsymbol{W}$ is the general closed-form solution of Eq. (87), then the general closed-form solution of Eq. (86) is given by*

$$\boldsymbol{X} = \frac{1}{2}\left(\boldsymbol{W} + \sum_{i=0}^{N} A_i f(\boldsymbol{W}) B_i + C \right).$$

3. *If $f(X) = \overline{X}$ or $f(X) = X^H$, then Eq. (86) has a unique solution $X^\star$ if and only if Eq. (87) has a unique solution $W^\star$. If $f(X) = X^T$, then Eq. (86) has a unique solution $X^\star$ if Eq. (87) has a unique solution $W^\star$. In both cases, $X^\star = W^\star$.*

8. Conclusion

This paper is concerned with solvability, existence of unique solution, closed-form solution and numerical solution of matrix equation $X = Af(X)B + C$ with $f(X) = X^{\mathrm{T}}$, $f(X) = \overline{X}$ and $f(X) = X^{\mathrm{H}}$, where X is a matrix to be determined. It is established that the solvability of these equations are equivalent to the solvability of some auxiliary standard Stein matrix equations in the form of $W = \mathcal{A}W\mathcal{B}+\mathcal{C}$ where the dimensions of the coefficient matrices $\mathcal{A}$, $\mathcal{B}$ and $\mathcal{C}$ are the same as dimensions of the coefficient matrices of the original equation. Based on the auxiliary standard Stein matrix equation, the conditions for solvability and the existence of unique solution are proposed. Closed-form solutions are also obtained by inversion of a square matrix, which is a generalization of the standard results on the standard Stein equation. Numerical solutions are approximated by iterations which are generalizations of the Smith iteration and accelerated Smith iteration associated with the standard Stein equations. We should point out that the idea in this paper can be readily adopted to study equations in the form of $AX - f(X)B = C$ with $f(X) = X^{\mathrm{T}}$, $f(X) = \overline{X}$ and $f(X) = X^{\mathrm{H}}$. Corresponding results on the solvability, existence of unique solution, closed-form solutions, and iterative solutions of these equations of the form AX–$f(X)B = C$ have been developed and will be reported elsewhere.

Acknowledgements

The authors thank the Editor and the anonymous reviewers for their helpful comments and suggestions which have helped to improve the quality of the paper.

This work is supported in part by the National Natural Science Foundation of China under Grant numbers 60904007 and 61074111, the China Postdoctoral Science Foundation under Grant number 20100480059, the Foundation for Innovative Research Group of the National Natural Science Foundation of China under Grant 601021002, the Development Program for Outstanding Young Teachers at Harbin Institute of Technology under Grant number HITQNJS.2009.054, the Heilongjiang Postdoctoral Foundation of China under Grant No. LRB10-194, and by HKU CRCG 201007176243.

References

[1] J.L. Chen, X.H. Chen, Special Matrices, Tsinghua University Press, 2002 (in Chinese).
[2] E. Desouza, S.P. Bhattacharyya, Controllability, observability and the solution of $AX - XB = C$, Linear Algebra Appl. 39 (1981) 167–188.
[3] M. Dehghan, M. Hajarian, The general coupled matrix equations over generalized bisymmetric matrices, Linear Algebra Appl. 432 (6) (2010) 1531–1552.
[4] F. Ding, T. Chen, Gradient based iterative algorithms for solving a class of matrix equations, IEEE Trans. Automat. Control 50 (8) (2005) 1216–1221.
[5] F. Ding, P.X. Liu, J. Ding, Iterative solutions of the generalized Sylvester matrix equations by using the hierarchical identification principle, Appl. Math. Comput. 197 (1) (2008) 41–50.
[6] D.S. Djordjević, Explicit solution of the operator equation $A^*X + X^*A = B$, J. Comput. Appl. Math. 200 (2) (2007) 701–704.
[7] X. Duan, A. Liao, On the existence of Hermitian positive definite solutions of the matrix equation $X^s + A^*X^tA = Q$, Linear Algebra Appl. 429 (4) (2008) 673–687.
[8] J. Feng, J. Lam, Y. Wei, Spectral properties of sums of certain Kronecker products, Linear Algebra Appl. 431 (9) (2009) 1691–1701.
[9] G.H. Golub, S. Nash, C.F. Van Loan, A Hessenberg–Schur method for the matrix problem $AX + XB = C$, IEEE Trans. Automat. Control 24 (1979) 909–913.
[10] J.H. Hodges, Some matrix equations over a finite field, Ann. Mat. Pura Appl. (4) 44 (1957) 245–250.
[11] L. Huang, The matrix equation $AXB - GXD = E$ over the quaternion field, Linear Algebra Appl. 234 (1996) 197–208.
[12] A. Jameson, Solution of the equation $AX - XB = C$ by inversion of an $M \times M$ or $N \times N$ matrix, SIAM J. Appl. Math. 16 (1968) 1020–1023.
[13] T. Jiang, M. Wei, On solution of the matrix equations $X - AXB = C$ and $X - A\overline{X}B = C$, Linear Algebra Appl. 367 (2003) 225–233.
[14] P. Lancaster, M. Tismenetsky, The Theory of Matrices with Applications, second ed., Academic Press, New York, 1985.
[15] C.-K. Li, L. Rodman, N.-K. Tsing, Linear operators preserving certain equivalence relations originating in system theory, Linear Algebra Appl. 161 (15) (1992) 165–225.
[16] C.-K. Li, L. Rodman, Preservers of spectral radius, numerical radius, or spectral norm of the sum on nonnegative matrices, Linear Algebra Appl. 430 (7) (2009) 1739–1761.
[17] Z.-Y. Li, Y. Wang, B. Zhou, G.-R. Duan, Least squares solution with the minimum-norm to general matrix equations via iteration, Appl. Math. Comput. 215 (10) (2010) 3547–3562.
[18] D.F. Miller, The iterative solution of the matrix equation $XA + BX + C = 0$, Linear Algebra Appl. 105 (1988) 131–137.
[19] F. Piao, Q. Zhang, Z. Wang, The solution to matrix equation $AX + X^{\mathrm{T}}C = B$, J. Franklin Inst. 344 (8) (2007) 1056–1062.
[20] R.A. Smith, Matrix equation $XA + BX = C$, SIAM J. Appl. Math. 16 (1) (1968) 198–201.
[21] T.Y. Tam, M.C. Thompson, Determinant and Pfaffian of sum of skew symmetric matrices, Linear Algebra Appl. 433 (2) (2010) 412–423.
[22] M.H. Wang, X.H. Cheng, M.H. Wei, Iterative algorithms for solving the matrix equation $AXB + CX^{\mathrm{T}}D = E$, Appl. Math. Comput. 187 (2) (2007) 622–629.
[23] H.K. Wimmer, Linear matrix equations: the module theoretic approach, Linear Algebra Appl. 120 (1989) 149–164.
[24] B. Zhou, J. Lam, G.R. Duan, On Smith-type iterative algorithms for the Stein matrix equation, Appl. Math. Lett. 22 (2009) 1038–1044.
[25] B. Zhou, G.R. Duan, An explicit solution to the matrix equation $AX - XF = BY$, Linear Algebra Appl. 402 (3) (2005) 345–366.
[26] K. Zhou, J. Doyle, K. Glover, Robust and Optimal Control, Prentice-Hall (1996).
[27] G. Bourgeois, How to solve the matrix equation $XA - AX = f(X)$, Linear Algebra Appl. 434 (3) (2011) 657–668.
[28] F. De Teran, F.M. Dopico, The solution of the equation $XA + AX^{\mathrm{T}} = 0$ and its application to the theory of orbits, Linear Algebra Appl. 434 (1) (2011) 44–67.
[29] M.L. Arias, M.C. Gonzalez, Positive solutions to operator equations $AXB = C$, Linear Algebra Appl. 433 (6) (2011) 1194–1202.
[30] K. Jbilou, ADI preconditioned Krylov methods for large Lyapunov matrix equations, Linear Algebra Appl. 432 (10) (2010) 2473–2485.