Relation between the T-congruence Sylvester equation and the generalized Sylvester equation

Yuki Satake*, Masaya Oozawa*, Tomohiro Sogabe*
Yuto Miyatake†, Tomoya Kemmochi*, Shao-Liang Zhang*

Abstract

The T-congruence Sylvester equation is the matrix equation $AX + X^T B = C$, where $A \in \mathbb{R}^{m \times n}$, $B \in \mathbb{R}^{n \times m}$, and $C \in \mathbb{R}^{m \times m}$ are given, and $X \in \mathbb{R}^{n \times m}$ is to be determined. Recently, Oozawa et al. discovered a transformation that the matrix equation is equivalent to one of the well-studied matrix equations (the Lyapunov equation); however, the condition of the transformation seems to be too limited because matrices $A$ and $B$ are assumed to be square matrices ($m = n$). In this paper, two transformations are provided for rectangular matrices $A$ and $B$. One of them is an extension of the result of Oozawa et al. for the case $m \geq n$, and the other is a novel transformation for the case $m \leq n$.

1 Introduction

We consider the T-congruence Sylvester equation of the form

$$AX + X^T B = C,$$

where $A \in \mathbb{R}^{m \times n}$, $B \in \mathbb{R}^{n \times m}$, and $C \in \mathbb{R}^{m \times m}$ are given, and $X \in \mathbb{R}^{n \times m}$ is to be determined. The matrix equation (1) appears in the industrial and applied mathematical field: palindromic eigenvalue problems for studying the vibration of rail tracks under the excitation arising from high-speed trains [1, 2] and also in an abstract mathematical field: the tangent spaces of the Stiefel manifold [3].

In the past decade, the numerical solvers for solving equation (1) in view of direct methods and iterative methods have been well developed, (see e.g., [4, 5, 6]); however, the theoretical features of the matrix equation itself such as the uniqueness of the solution have not been explored extensively. In 2018, Terán et al. provided an important theoretical contribution regarding the necessary and sufficient condition for the existence of exactly one solution of the matrix equation (1) in [7], whose theoretical approach is based on transforming the matrix equation into large linear systems and then analyzing the linear systems.

Though irrelevant to the above, Oozawa et al., in 2018, developed a novel transformation that can be used to transform the matrix equation (1) into a more well-researched matrix equation, i.e., $X^T$ does not appear in its transformed matrix equation, as shown in the following theorem:

**Theorem 1.** (§ Theorem 3.1) Let $A, B, C \in \mathbb{R}^{n \times n}$ and assume that $A$ is nonsingular. Moreover, let $S := B^T A^{-1}$ and suppose $\Lambda(S)$ is reciprocal free. Then the T-congruence Sylvester equation (1) is equivalent to the Lyapunov equation

$$\bar{X} - S \bar{X} S^T = Q,$$

where $\bar{X} := AX$ and $Q := C - (SC)^T$. 

*Graduate School of Engineering, Nagoya University, Furo-cho, Chikusa-ku, Nagoya 464-8603, Japan, Email: {y-satake,m-oozawa,sogabe,kemmochi,zhang}@na.nuap.nagoya-u.ac.jp
†Cybermedia Center, Osaka University, 1-32 Machikaneyama, Toyonaka, Osaka 560-0043, Japan, Email: miyatake@cas.cmc.osaka-u.ac.jp
The significance of Theorem 1 is that finding mathematical features and solvers of the matrix equation (1) reduces to only finding them of the Lyapunov equation whose features are studied extensively in control theory [9] and whose solvers have been well developed (see e.g., [10]). Furthermore, very recently, Miyajima utilized Theorem 1 to design a fast verified algorithm for the solution of the matrix equation (1) in [11]. From the viewpoint of the complexities of the numerical error verification, utilizing Theorem 1 successfully reduces the complexities of $O(n^6)$ to $O(n^3)$.

Although Theorem 1 is of interest, the class of matrices that satisfies the required condition is limited. In fact, Theorem 1 does not hold if matrices $A$ and $B$ are rectangular. This motivates us to find transformations for the case where matrices $A$ and $B$ are rectangular. We will see in the following sections that a transformation for the case $m \geq n$ can be regarded as an extension of Theorem 1 and a transformation for the case $m \leq n$ is a novel transformation in that a similar mathematical approach as used in Theorem 1 cannot be employed mainly due to under-determined linear systems underlying the matrix equation.

The rest of this paper is organized as follows. Section 2 describes some mathematical preliminaries on the Kronecker product. Section 3 presents the main results that under some conditions the $T$-congruence Sylvester equation is equivalent to a generalized Sylvester equation. The conclusion is provided in Section 4.

Throughout this paper, $\Lambda(S)$ denotes the spectrum of $S$, i.e., the set of eigenvalues of $S \in \mathbb{R}^{m \times m}$. Let $I_m$ denote the $m \times m$ identity matrix and $A^\dagger$ the Moore–Penrose inverse of $A$.

## 2 Preliminaries

In this section, some mathematical preliminaries are collected on the Kronecker product and the vectorization of a matrix, referred to as the vec operator. Given $A = [a_{ij}] \in \mathbb{R}^{m \times n}$ and $B \in \mathbb{R}^{p \times d}$, the Kronecker product $\otimes$ is defined by

$$A \otimes B := \begin{bmatrix} a_{11}B & a_{12}B & \cdots & a_{1n}B \\ a_{21}B & a_{22}B & \cdots & a_{2n}B \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1}B & a_{m2}B & \cdots & a_{mn}B \end{bmatrix} \in \mathbb{R}^{mp \times nq}.$$ 

In addition, for $C \in \mathbb{R}^{p \times q}$ and $D \in \mathbb{R}^{q \times r}$, it is known that

$$(A \otimes B)(C \otimes D) = (AC) \otimes (BD).$$

Any eigenvalue of $A \otimes B$ is a product of an eigenvalue of $A$ and an eigenvalue of $B$ as shown below.

**Lemma 2.** ([12], p.412) Let $\lambda_1, \ldots, \lambda_m$ be the eigenvalues of $A \in \mathbb{R}^{m \times m}$ and $\mu_1, \ldots, \mu_n$ be the eigenvalues of $B \in \mathbb{R}^{n \times n}$. Then the eigenvalues of $A \otimes B$ can be written as

$$\lambda_i \mu_j \quad (1 \leq i \leq m, 1 \leq j \leq n).$$

For $A = [a_1, a_2, \ldots, a_n] \in \mathbb{R}^{m \times n}$, the vec operator, $\text{vec}: \mathbb{R}^{m \times n} \to \mathbb{R}^{mn}$ is defined by

$$\text{vec}(A) := \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix} \in \mathbb{R}^{mn},$$

and the inverse vec operator, $\text{vec}^{-1}: \mathbb{R}^{mn} \to \mathbb{R}^{m \times n}$ is defined by

$$\text{vec}^{-1}(\text{vec}(A)) = A.$$ 

For the vec operator, the following fact holds:
Lemma 3. ([13] Theorem 10) Let $A \in \mathbb{R}^{m \times n}$, $B \in \mathbb{R}^{n \times p}$, and $C \in \mathbb{R}^{p \times q}$. Then, it follows that

$$
\text{vec}(ABC) = (C^T \otimes A)\text{vec}(B).
$$

We collect some properties on permutation matrices related to the vec operator.

Lemma 4. ([13]) Let $e_i$ denote an $n$-dimensional column vector which has 1 in the $i$th position and 0 for otherwise, i.e.,

$$
e_i = [0, 0, \ldots, 0, 1, 0, \ldots, 0]^T \in \mathbb{R}^n.
$$

Let $A \in \mathbb{R}^{m \times n}$ and $B \in \mathbb{R}^{p \times q}$. Then, for the permutation matrix

$$P_{mn} := \begin{bmatrix}
I_m \otimes e_1^T \\
I_m \otimes e_2^T \\
\vdots \\
I_m \otimes e_n^T
\end{bmatrix} \in \mathbb{R}^{mn \times mn},
$$

the following properties hold:

$$
P_{mn}^T = P_{nm},
$$

$$
P_{mn}^T P_{mn} = P_{mn} P_{mn}^T = I_{mn},
$$

$$
\text{vec}(A) = P_{mn} \text{vec}(A^T),
$$

$$
P_{mp}(A \otimes B)P_{nq}^T = B \otimes A.
$$

It is known that T-congruence Sylvester equation (1) can be transformed into a linear system whose coefficient matrix has the Kronecker product structure with the permutation matrix multiplication ($P_{mn}$). To see this, applying the vec operator to (1) with Lemma 3 yields

$$
(I_m \otimes A)\text{vec}(X) + (B^T \otimes I_m)\text{vec}(X^T) = \text{vec}(C).
$$

(6)

It follows from (3), (4), and (5) of Lemma 4 that the second term of the left-hand side of equation (6) is

$$
(B^T \otimes I_m)\text{vec}(X^T) = (B^T \otimes I_m)P_{mn} \text{vec}(X)
$$

$$
= P_{mn}(I_m \otimes B^T)P_{mn}^T \text{vec}(X)
$$

$$
= P_{mn}(I_m \otimes B^T)\text{vec}(X).
$$

Thus, T-congruence Sylvester equation (1) can be transformed into the following linear system:

$$
\{I_m \otimes A + P_{mn}(I_m \otimes B^T)\}x = c,
$$

(7)

where $x := \text{vec}(X)$ and $c := \text{vec}(C)$.

3 Main results

In this section, a transformation of the T-congruence Sylvester equation into a generalized Sylvester equation is shown. A transformation for the case $m \geq n$ is provided in Section 3.1, which is an extension of Theorem 1 and a novel transformation for the case $m \leq n$ is provided in Section 3.2. Before stating the main results, the term reciprocal free is described. The following definition admits 0 and $\infty$ as elements of $\lambda_1, \ldots, \lambda_n$.

Definition 5. (reciprocal free [1, 2]) A set $\{\lambda_1, \ldots, \lambda_n\} \subset \mathbb{C} \cup \{\infty\}$ is said to be reciprocal free if $\lambda_i \neq \frac{1}{\lambda_j}$ for any $1 \leq i, j \leq n$. 

3
3.1 The case $m \geq n$

In this subsection, a transformation from the T-congruence Sylvester equation (1) to the generalized Sylvester equation for the case $m \geq n$ is described.

**Theorem 6.** Let $m \geq n$, $A \in \mathbb{R}^{m\times n}$, $B \in \mathbb{R}^{n\times m}$, and $C \in \mathbb{R}^{m\times m}$. Assume that there exists a matrix $S \in \mathbb{R}^{m\times n}$ such that $B^T = SA$ and $\Lambda(S)$ is reciprocal free. Then, the T-congruence Sylvester equation (1) is equivalent to the generalized Sylvester equation

$$AX - B^TS^T = C - (SC)^T.$$  

(8)

**Proof.** First, we show that $G := I_m \otimes I_m - P_{mn}(I_m \otimes S)$ is nonsingular. Then we prove that the T-congruence Sylvester equation (1) and the generalized Sylvester equation (8) are equivalent.

Let $(\lambda(G), v)$ be an eigenpair of $G$. Then it follows that

$$P_{mn}(I_m \otimes S)v = (1 - \lambda(G))v.$$ 

Thus, $G$ is nonsingular if and only if $K := P_{mn}(I_m \otimes S)$ has no eigenvalue equal to 1. From (2) and (5), it follows that

$$K^2 = P_{mn}(I_m \otimes S)P_{mn}(I_m \otimes S)$$

$$= (S \otimes I_m)(I_m \otimes S)$$

$$= S \otimes S.$$  

(9)

As a result of (9) and Lemma 2, any eigenvalue of $K^2$ is a product of two eigenvalues of $S$. Therefore, thanks to the assumption that $\Lambda(S)$ is reciprocal free, $G$ is nonsingular.

Now, we show the statement of Theorem 6. By multiplying (7) by $G$ and using (2) and (5), we have

$$\{I_m \otimes A - P_{mn}(I_m \otimes SA) + P_{mn}(I_m \otimes B^T) - P_{mn}(I_m \otimes S)P_{mn}(I_m \otimes B^T)\}x = Ge,$$

which implies

$$(I_m \otimes A - S \otimes B^T)x = Ge,$$  

(10)

since $B^T = SA$. From (4) and (5), the right-hand side of (10) becomes

$$Ge = c - P_{mn}(I_m \otimes S)e$$

$$= c - (S \otimes I_m)P_{mn}e$$

$$= \text{vec}(C) - (S \otimes I_m)\text{vec}(C^T).$$

Therefore, applying the inverse vec operator to (10), we obtain (8) by using Lemma 2. Hence, we complete the proof of Theorem 6. \qed

In Theorem 6, we have not mentioned the existence of the matrix $S$ satisfying the condition $B^T = SA$. If the matrix $A$ has full column rank, there always exist infinitely many matrices $S$ satisfying $B^T = SA$. The simplest one is expressed by using the Moore–Penrose inverse of $A$.

**Proposition 7.** Let $m \geq n$, $A \in \mathbb{R}^{m\times n}$ have full column rank and $B \in \mathbb{R}^{n\times m}$. Then $S := B^TA^\dagger$ yields $B^T = SA$.

**Proof.** When $A$ has full column rank, the Moore–Penrose inverse of $A$ can be written as

$$A^\dagger = (A^TA)^{-1}A^T,$$

which implies $B^T = SA$.

Notice that if $m = n$, Theorem 6 with Proposition 7 is equivalent to Theorem 1. Thus it can be regarded as an extension of Theorem 1. Theorem 6 also holds true for the case $m < n$; however, $B^T = SA$ holds true only for limited cases, i.e., an additional assumption is required for $B$ and $A$. Thus, in the next subsection, we consider another transformation.
3.2 The case \( m \leq n \)

In this subsection, we describe a novel transformation from the T-congruence Sylvester equation (1) to the generalized Sylvester equation in the case \( m \leq n \).

**Theorem 8.** Let \( m \leq n \), \( A \in \mathbb{R}^{m \times n} \), \( B \in \mathbb{R}^{n \times m} \), and \( C \in \mathbb{R}^{n \times m} \). Assume that there exists a matrix \( D \in \mathbb{R}^{n \times m} \) such that \( I_m = AD \), and for \( S := B^T D \), \( \Lambda(S) \) is reciprocal free. Then the T-congruence Sylvester equation (1) is equivalent to the generalized Sylvester equation

\[
A\tilde{X} - B^T \tilde{X}S^T = C,
\]

where \( \tilde{X} \in \mathbb{R}^{n \times m} \) satisfies \( X = \tilde{X} - D\tilde{X}^T B \).

**Proof.** First, we show that \( G := I_m \otimes I_n - P_{nm}(D \otimes B^T) \) is nonsingular. Then we prove that the T-congruence Sylvester equation (1) and the generalized Sylvester equation (11) are equivalent.

Let \( (\lambda^{(G)}, v) \) be an eigenpair of \( G \). Then, it follows that

\[
P_{nm}(D \otimes B^T)v = (1 - \lambda^{(G)})v.
\]

Thus, \( G \) is nonsingular if and only if \( K := P_{nm}(D \otimes B^T) \) has no eigenvalue equal to 1. From (2) and (5), it follows that

\[
K^2 = P_{nm}(D \otimes B^T)P_{nm}(D \otimes B^T) = (B^T \otimes D)(D \otimes B^T) = B^TD \otimes DB^T.
\]

(12)

Since \( B^T D \) and \( DB^T \) have the same set of nonzero eigenvalues, then we have from (12) and Lemma 2 that a square of any eigenvalue of \( K \) is 0 or a product of two eigenvalues of \( S \). Therefore, from the assumption that \( \Lambda(S) \) is reciprocal free, \( G \) is nonsingular.

Now, we prove the statement of Theorem 8. By using the nonsingular matrix \( G \), (2), (3), and (5), the equation (7) becomes

\[
\{I_m \otimes A + P_{nm}(I_m \otimes B^T)\}\{I_m \otimes I_n - P_{nm}(D \otimes B^T)\} \tilde{x} = c,
\]

which implies

\[
(I_m \otimes A - S \otimes B^T) \tilde{x} = c,
\]

(13)

where \( \tilde{x} := G^{-1}x \). Since \( \text{vec}(\tilde{X}) = \tilde{x} \), applying the inverse vec operator to (13), we obtain (11) by using Lemma 2. Hence we complete the proof of Theorem 8. \( \square \)

In Theorem 8 there exist infinitely many matrices \( D \). The simplest one is the Moore–Penrose inverse of \( A \).

**Proposition 9.** Let \( m \leq n \), \( A \in \mathbb{R}^{m \times n} \) have full row rank and \( B \in \mathbb{R}^{n \times m} \). Then \( D := A^+ \) yields \( I_m = AD \).

**Proof.** When \( A \) has full row rank, the Moore–Penrose inverse of \( A \) can be written as

\[
A^+ = A^T(AA^T)^{-1}.
\]

Thus, letting \( D = A^+ \), we have \( AD = I_m \).

In the square case, by setting \( \hat{X} = A\tilde{X} \), Theorem 8 reduces to the following result:

**Corollary 10.** Let \( A, B, C \in \mathbb{R}^{n \times n} \). Assume that \( A \) is nonsingular, and for \( S := B^T A^{-1} \), \( \Lambda(S) \) is reciprocal free. Then, the T-congruence Sylvester equation (1) is equivalent to the Lyapunov equation

\[
\hat{X} - S\hat{X}S^T = C,
\]

where \( \hat{X} := A\hat{X} \) and \( \hat{X} \) satisfies \( X = \hat{X} - A^{-1}\hat{X}^T B \).

This is a different result from that given in Theorem 1.
4 Conclusion

Our results are summarized in Table 1. The key result is that under certain conditions $T$-congruence Sylvester equation is equivalent to the generalized Sylvester equation. This may provide flexibility in research for finding further mathematical features and solvers of $T$-congruence Sylvester equation (or the generalized Sylvester equation) because mathematical features and solvers of the generalized Sylvester equation can be utilized for $T$-congruence Sylvester equation, and vice versa.

| Theorem | $m \geq n$ | $m \leq n$ |
|---------|-------------|-------------|
|         | generalized Sylvester eq. (Th. 3.2) | generalized Sylvester eq. (Th. 3.5) |
| $m = n$ | $AX - B^T X S^T = C - (SC)^T$ | $A\tilde{X} - B^T \tilde{X} S^T = C$ |
| Lyapunov eq. [8, Th. 3.1] | Lyapunov eq. (Cor. [10]) |
| $\tilde{X} - S \tilde{X} S^T = Q$ | $\tilde{X} - S \tilde{X} S^T = C$ |

References

[1] R. Byers, D. Kressner, Structured condition numbers for invariant subspaces, SIAM J. Matrix Anal. Appl. 28 (2) (2006) 326–347.

[2] D. Kressner, C. Schröder, D.S.Watkins, Implicit QR algorithms for palindromic and even eigenvalue problems, Numer. Algorithms 51 (2) (2009) 209–238.

[3] P.-A. Absil, R. Mahony, R. Sepulchre, Optimization Algorithms on Matrix Manifolds, Princeton University Press, Princeton, 2008.

[4] C.-Y. Chiang, E. K.-W. Chu, W.-W. Lin, On the ⋆-Sylvester equation $AX \pm X\star B\star = C$, Appl. Math. Comput. 218 (17) (2012) 8393–8407.

[5] F. De Terán, F. M. Dopico, Consistency and efficient solution of the Sylvester equation for ⋆-congruence, Electron. J. Linear Algebra 22 (2011) 849–863.

[6] M. Hajarian, Developing Bi-CG and Bi-CR methods to solve generalized Sylvester-transpose matrix equations, Int. J. Autom. Comput. 11 (1) (2014) 25–29.

[7] F. De Terán, B. Iannazzo, F. Poloni, L. Robol, Solvability and uniqueness criteria for generalized Sylvester-type equations, Linear Algebra Appl. 542 (2018) 501–521.

[8] M. Oozawa, T. Sogabe, Y. Miyatake, S.-L. Zhang, On a relationship between the $T$-congruence Sylvester equation and the Lyapunov equation, J. Comput. Appl. Math. 329 (2018) 51–56.

[9] Z. Gajic, M. T. J. Qureshi, Lyapunov Matrix Equation in System Stability and Control, Mathematics in Science and Engineering, Academic Press, Inc., San Diego, 1995.

[10] B. N. Datta, Numerical Methods for Linear Control Systems, Elsevier Academic Press, San Diego, 2004.

[11] S. Miyajima, Fast verified computation for the solution of the $T$-congruence Sylvester equation, Jpn. J. Ind. Appl. Math. 35 (2) (2018) 541–551.

[12] P. Lancaster, M. Tismenetsky, The Theory of Matrices, Computer Science and Applied Mathematics, Academic Press, Inc., Orlando, 1985.

[13] H. Zhang, F. Ding, On the Kronecker products and their applications, J. Appl. Math. (2013) 1–8.