Sylvester matrix equation under the semi-tensor product of matrices

Patrawut Chansangiam · Sorin V. Sabau

Abstract We investigate the Sylvester matrix equation in which the product is given by the semi-tensor product, and all involved matrices are matrices over an arbitrary field. We discuss necessary/sufficient condition(s) for the matrix equation to have a solution or a unique solution, or infinitely many solutions. These conditions concern ranks and linear independence. Moreover, we apply a certain kind of vectorization and matrix partitioning to transform the Sylvester equation into an equivalent linear system with respect to the conventional matrix product. Our study includes the Lyapunov equation and the equation $A \bowtie X = C$ as special cases.

Keywords Sylvester matrix equation · matrices over a field · semi-tensor product · Kronecker product

Mathematics Subject Classification (2010) 15A24 · 15A69

1 Introduction

In mathematics and computer science, matrices are two-dimensional rectangular arrays to represent certain data. Matrix operations can be viewed as procedures to produce new data from existing ones. One of familiar matrix operations is the usual matrix multiplication, which is defined for two matrices $A \in F^{m \times n}$ and $B \in F^{p \times q}$ such that $n = p$. Here, $F$ is an arbitrary field. In order to solve decoupling problems of a linear system and certain problems with multidimension data, D. Cheng (e.g. [2,3]) defined the semi-tensor product (STP) for matrices $A \in F^{m \times n}$ and $B \in F^{p \times q}$ with arbitrary dimensions, namely:

$$A \bowtie B = (A \otimes I_\frac{n}{\alpha}) (B \otimes I_\frac{p}{\alpha}),$$

(1.1)

where the integer $\alpha$ is the least common multiple of $n$ and $p$. Here, the operation $\otimes$ is the Kronecker (tensor) multiplication; see Section 2 for details. In the case $n = p$, the STP reduces to the usual matrix product $AB$. It turns
Pattrawut Chansangiam, Sorin V. Sabau

out that the STP for matrices over a field possesses rich algebraic properties like the usual matrix product, such as the associativity, the left/right distribution over the matrix addition, certain identity-like properties, and the compatibility with the scalar multiplication, the transposition, and the inversion; see e.g. [3–5]. The STP provides a convenient way to convert a multilinear function or a set of higher-dimensional data into a matrix-vector expression; see e.g. [3]. It turns out that the STP can be applied in several areas such as Boolean networks [6], classical logic and fuzzy mathematics [7,10], game theory [8], dynamic systems, [9], and others.

In recent years, many authors investigated linear matrix equations in which the matrix product is the STP. In 2015, Yao et al. [18] studied equivalent conditions for the matrix equation

\[ A \bowtie X = C \] (1.2)

to have a solution or a unique solution. These conditions involve linear dependence and the ranks of associated augmented matrices. Moreover, we can transform the matrix equation (1.2) into a linear system with respect to the usual product, so that we can solve it via an ordinary method. Recently, Wang [19] extended the study to the $MM - 2$ semi-tensor product. Li et al. [13] investigated the solvability of the system of matrix equations $A \bowtie X = B$ and $X \bowtie C = D$. The matrix equation

\[ A \bowtie X \bowtie B = C \] (1.3)

has been extensively used in non-linear programming, power science, parameter identification, etc. Ji [11] investigated the solvability condition of Eq. (1.3) where $A, B, C$ are given complex matrices. There are also interests in nonlinear matrix equations, e.g. the equation $A \bowtie X \bowtie X = B$ [20].

The previous discussions motivated us to investigate another linear matrix equations that are arised naturally in mathematics. The famous Sylvester matrix equation

\[ AX + XB = C \] (1.4)

and the famous Lyapunov equation $AX + XA^T = C$ are significantly used in control theory and differential equations; see e.g. [14]. A usual algebraic method to solve the Sylvester/Lyapunov equation is to transform the matrix equation into a linear system by means of certain kind of vectorizations, namely, the column/row vector operator; see e.g. [17]. See more recent information in [12].

In the present work, we investigate the Sylvester matrix equation

\[ A \bowtie X + X \bowtie B = C \] (1.5)

with respect to the STP. Here, the coefficient matrices $A, B, C$ are studied in a full generality, i.e., they are matrices over an arbitrary field. The matrix equation (1.5) includes Eqs. (1.2), (1.3) and the Lyapunov equation:

\[ A \bowtie X + X \bowtie A^T = C \] (1.6)
as special cases. We investigate the solvability and the unique solvability for Eq. (1.5) according to the matrix dimensions, ranks, and linear independence. Moreover, we show that Eq. (1.5) can be transformed to a linear system with respect to the usual product, so that we can solve it via elementary methods.

The rest of this paper is organized as follows. In Section 2, we setup basic notation and recall related background. We investigate the Sylvester matrix equation when the unknown is a vector in Section 3. We then discuss the case when the unknown is a matrix in Section 4. In particular, we investigate the Lyapunov equation (1.6) in Section 5. We conclude the whole work in the last section.

2 Preliminaries

Throughout, let $F$ be a field. For any positive integers $m, n$, we denote the set of $m \times n$ matrices over $F$ by $F^{m \times n}$. When $n = 1$, we set $F^m := F^{m \times 1}$.

The studies of matrix equations often concern real/complex matrices. In fact, we can generalize the theory to that for matrices over a suitable algebraic framework, such as fields (e.g. [16]), commutative rings (e.g. [1]), or principal ideal domains (e.g. [15]). Recall that theory of linear systems over a field works in the same way as that for matrices over the real field. An important result is the following Kronecker-Capelli theorem:

**Theorem 2.1 (see, e.g., [16])** Let $A \in F^{m \times n}$ and $b \in F^m$. Then the linear system $Ax = b$ has a solution $x \in F^n$ if and only if $\text{rank } [Ab] = \text{rank } A$. The system has a unique solution if and only if $\text{rank } [Ab] = \text{rank } A = n$. The system has infinitely many solutions if and only if $\text{rank } [Ab] = \text{rank } A < n$.

Recall that the Kronecker product (see e.g. [17]) of $A = [a_{ij}] \in F^{m \times n}$ and $B \in F^{h \times k}$, denoted by $A \otimes B$, is defined to be the block matrix

$$A \otimes B = \begin{bmatrix} a_{11}B & a_{12}B & \cdots & a_{1n}B \\ a_{21}B & a_{22}B & \cdots & a_{2n}B \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1}B & a_{m2}B & \cdots & a_{mn}B \end{bmatrix} \in F^{mh \times nk}.$$  

The vector operator (see e.g. [17]) $\text{Vec}(\cdot)$ assigns to each matrix $A = [a_{ij}] \in F^{m \times n}$ the column vector

$$\text{Vec}(A) = [a_{11} \ldots a_{m1} a_{12} \ldots a_{m2} \ldots a_{1n} \ldots a_{mn}]^T \in F^{mn \times 1}.$$  

Clearly, the vector operator is linear and bijective.

**Lemma 2.2** For any matrices $A \in F^{m \times n}, B \in F^{n \times p}$ and $C \in F^{p \times q}$, we have

$$\text{Vec}(AXB) = (B^T \otimes A) \text{Vec}(X).$$  

265
Proof. The proof is essentially the same as that for real matrices; see e.g. [17]. ⊓ ⊔

**Lemma 2.3** Let $A \in F^{m\times n}$ and $B \in F^{p\times q}$. Let us partition $B$ into its columns as $B = [B_1 \cdots B_q]$. Then

$$\text{Vec}(A \otimes B) = \begin{pmatrix} I_n \otimes [I_m \otimes B_1] \\ \vdots \\ I_n \otimes [I_m \otimes B_q] \end{pmatrix} \text{Vec}(A).$$

Proof. The proof is essentially the same as that for real matrices; see e.g. [17, Theorem 3.5]. ⊓ ⊔

### 3 The Sylvester equation in an unknown vector

In this section, we investigate the following problem:

**Problem 3.1** Given $A \in F^{m\times n}$, $B \in F^{h\times k}$, and $C \in F^{u\times v}$, find $X = [x_1 x_2 \cdots x_p]^T \in F^{p\times 1}$ such that

$$A \bowtie X + X \bowtie B = C. \quad (3.1)$$

Let us say that Problem 3.1 is well-defined if all matrix dimensions are compatible. We propose suitable matrix partitionings in order to transform Eq. (3.1) into a simple linear system with respect to the usual matrix product. Then, we investigate criterions for existence and uniqueness of solutions for the equation.

Let us denote $t = \text{lcm } \{n,p\}$. By definition of the semi-tensor product, we have

$$A \bowtie X = (A \otimes I_{\frac{mt}{n}})(X \otimes I_{\frac{t}{p}}) \in F^{\frac{mt}{n}\times \frac{t}{p}} \quad \text{and} \quad X \bowtie B = (X \otimes I_{t})B \in F^{ph\times k}.$$ 

Since $C \in F^{u\times v}$, we conclude that Problem 3.1 is well-defined if and only if $u = \frac{mt}{n} = ph$ and $v = \frac{t}{p} = k$.

From now on, assume that Problem 3.1 is well-defined. Next, we seek for appropriate partitionings of $A, B, C$ that fit with solution finding. Since $\frac{t}{p} = k$, Eq. (3.1) becomes

$$C = (A \otimes I_{\frac{t}{n}})(X \otimes I_{k}) + (X \otimes I_{t})B. \quad (3.2)$$

We split $A \otimes I_{t/n}$ into $p^2$ blocks $\hat{A}_{11}, \hat{A}_{12}, \ldots, \hat{A}_{pp} \in F^{h\times k}$, and split $C$ into $p$ blocks $\hat{C}_1, \hat{C}_2, \ldots, \hat{C}_p \in F^{h\times k}$. We split $X \otimes I_{k}$ into $p$ blocks $x_1I_k, \ldots, x_pI_k$. 

---

266
So, we can write Eq. (3.2) as
\[
\begin{bmatrix}
\dot{C}_1 \\
\dot{C}_2 \\
\vdots \\
\dot{C}_p
\end{bmatrix}
= \begin{bmatrix}
\hat{A}_{11} & \hat{A}_{12} & \cdots & \hat{A}_{1p} \\
\hat{A}_{21} & \hat{A}_{22} & \cdots & \hat{A}_{2p} \\
\vdots & \vdots & \ddots & \vdots \\
\hat{A}_{p1} & \hat{A}_{p2} & \cdots & \hat{A}_{pp}
\end{bmatrix}
\begin{bmatrix}
x_{1}I_{k} \\
x_{2}I_{k} \\
\vdots \\
x_{p}I_{k}
\end{bmatrix}
+ \begin{bmatrix}
x_{1}B \\
x_{2}B \\
\vdots \\
x_{p}B
\end{bmatrix}.
\]
By considering each \(i\)-th row block of the equation, we have
\[
\dot{C}_i = \hat{A}_{1i}(x_1I_k) + \hat{A}_{2i}(x_2I_k) + \cdots + \hat{A}_{pi}(x_pI_k) + x_iB
= x_1\hat{A}_{1i} + x_2\hat{A}_{2i} + \cdots + x_i(\hat{A}_{ii} + B) + \cdots + x_p\hat{A}_{pi}.
\] (3.3)

Thus, the solvability of Eq. (3.1) implies that, for each \(i\), we can write \(\dot{C}_i\) in terms of a linear combination of \(\hat{A}_{1i}, \hat{A}_{2i}, \ldots, \hat{A}_{ii} + B, \ldots, \hat{A}_{pi}\), or equivalently, the set \(\{\hat{A}_{1i}, \hat{A}_{2i}, \ldots, \hat{A}_{ii} + B, \ldots, \hat{A}_{pi}, \dot{C}_i\}\) is linearly dependent. The latter condition can be stated in terms of ranks and augment matrices as follows:

**Theorem 3.2** Assume that Problem 3.1 is well-defined. If Eq. (3.1) has a solution, then for each \(i = 1, \ldots, p\),

(i) the set \(\{\hat{A}_{1i}, \hat{A}_{2i}, \ldots, \hat{A}_{ii} + B, \ldots, \hat{A}_{pi}, \dot{C}_i\}\) is linearly dependent in the vector space \(F^{h\times k}\).

(ii) \(\text{rank } [\hat{A}_i \; \dot{C}_i] = \text{rank } \hat{A}_i\), where \(\hat{A}_i = [\hat{A}_{i1} \; \hat{A}_{i2} \; \cdots \; \hat{A}_{ii} + B \; \cdots \; \hat{A}_{ip}]\).

**Remark 3.1** The converse of Theorem 3.2 is false in general, even in the case of real matrices. Consider the Sylvester equation when we are given
\[
A = \begin{bmatrix}
0 & 2 & 2 & 0 \\
2 & 1 & 1 & 1 \\
1 & 3 & 1 & 1 \\
1 & 0 & 1 & 0
\end{bmatrix},
B = \begin{bmatrix}
1 & 1 \\
2 & 1 \\
1 & 0 \\
1 & 0
\end{bmatrix}
\text{ and } C = \begin{bmatrix}
1 & 1 \\
0 & 1 \\
1 & 0 \\
0 & 1
\end{bmatrix},
\]
and we will solve for \(X \in \mathbb{R}^{2\times 1}\). We split \(A\) into 4 equal-size blocks as
\[
\hat{A}_{11} = \begin{bmatrix}
0 & 2 \\
2 & 1
\end{bmatrix},
\hat{A}_{12} = \begin{bmatrix}
2 & 0 \\
1 & 1
\end{bmatrix},
\hat{A}_{21} = \begin{bmatrix}
1 & 3 \\
1 & 0
\end{bmatrix},
\hat{A}_{22} = \begin{bmatrix}
1 & 1 \\
1 & 0
\end{bmatrix},
\]
and split \(C\) into 2 equal-size blocks as
\[
\dot{C}_1 = \begin{bmatrix}
1 & 1 \\
0 & 1
\end{bmatrix},
\dot{C}_2 = \begin{bmatrix}
1 & 0 \\
0 & 1
\end{bmatrix}.
\]
Letting \(\hat{A}_i = [\hat{A}_{1i} + B \; \hat{A}_{12}]\) and \(\hat{A}_2 = [\hat{A}_{21} \; \hat{A}_{22} + B]\), we compute
\[
\text{rank } [\hat{A}_1 \; \dot{C}_1] = \text{rank } \hat{A}_1 = 2 \text{ and } \text{rank } [\hat{A}_2 \; \dot{C}_2] = \text{rank } \hat{A}_2 = 2.
\]
However, the equation has no solution.
The next theorem suggests how to transform the Sylvester equation into an equivalent linear system.

**Theorem 3.3** Assume that Problem 3.1 is well-defined. Then the Sylvester equation $A \times X + X \times B = C$ is equivalent to the following linear system

$$\tilde{A}X = \tilde{C}, \quad (3.4)$$

where

$$\tilde{A} = \begin{bmatrix}
\text{Vec}(\hat{A}_{11} + B) & \text{Vec}(\hat{A}_{12}) & \cdots & \text{Vec}(\hat{A}_{1p}) \\
\text{Vec}(\hat{A}_{21}) & \text{Vec}(\hat{A}_{22} + B) & \cdots & \text{Vec}(\hat{A}_{2p}) \\
\vdots & \vdots & \ddots & \vdots \\
\text{Vec}(\hat{A}_{p1}) & \text{Vec}(\hat{A}_{p2}) & \cdots & \text{Vec}(\hat{A}_{pp} + B)
\end{bmatrix}, \quad \tilde{C} = \begin{bmatrix}
\text{Vec}(\hat{C}_1) \\
\text{Vec}(\hat{C}_2) \\
\vdots \\
\text{Vec}(\hat{C}_p)
\end{bmatrix}.$$

**Proof.** We apply the vector operator to Eq. (3.3) and, thus, obtain

$$x_1 \text{Vec}(\hat{A}_{11}) + x_2 \text{Vec}(\hat{A}_{12}) + \cdots + x_i \text{Vec}(\hat{A}_{ii} + B) + \cdots + x_p \text{Vec}(\hat{A}_{ip}) = \text{Vec}(\hat{C}_i),$$

for each $i = 1, 2, \ldots, p$. Note that the left hand side of this equation can be written as the usual product between $[\text{Vec}(\hat{A}_{11}) \cdots \text{Vec}(\hat{A}_{ii} + B) \cdots \text{Vec}(\hat{A}_{ip})]$ and $X = [x_1 \cdots x_p]^T$. Now, we can put all equations into the linear system (3.4) as desire. Since the vector operator is injective, the original matrix equation and the resulting linear system are equivalent.

As immediate consequences of Theorem 3.3 together with Kronecker-Capelli theorem (Theorem 2.1), we conclude:

**Corollary 3.4** Assume the assumption and notation in Theorem 3.3. Then the Sylvester equation (3.1) is consistent if and only if $\text{rank}[\tilde{A}] = \text{rank}[\tilde{A} \tilde{C}]$. It has a unique solution if and only if $\text{rank}[\tilde{A}] = \text{rank}[\tilde{A} \tilde{C}] = p$. It has infinitely many solutions if and only if $\text{rank}[\tilde{A}] = \text{rank}[\tilde{A} \tilde{C}] < p$.

**Example 3.1** We would like to find an unknown matrix $X \in \mathbb{R}^{2 \times 1}$ satisfying the Sylvester equation when

$$A = \begin{bmatrix} 1 & 3 & 1 \\ 2 & 1 & 3 \end{bmatrix}, \quad B = \begin{bmatrix} 3 & 1 & 2 \\ 2 & 1 & 1 \end{bmatrix} \quad \text{and} \quad C = \begin{bmatrix} 8 & 5 & 10 \\ 13 & 4 & 5 \\ 13 & 12 & 8 \\ 9 & 7 & 12 \end{bmatrix}.$$

By Theorem 3.2, it suffices to solve the linear system $\tilde{A}X = \tilde{C}$ where

$$\tilde{A} = \begin{bmatrix} 4 & 2 & 1 & 2 & 5 & 1 & 2 & 0 & 0 & 2 & 1 & 0 \\ 0 & 3 & 1 & 0 & 0 & 1 & 3 & 3 & 4 & 1 & 2 & 4 \end{bmatrix}^T,$$

$$\tilde{C} = \begin{bmatrix} 8 & 13 & 5 & 4 & 10 & 5 & 13 & 9 & 12 & 7 & 8 & 12 \end{bmatrix}^T.$$

Thus, $X = [2 \ 3]^T$ is a unique solution of the matrix equation.
4 The Sylvester equation in an unknown matrix

In this section, we investigate the following problem:

**Problem 4.1** Given \( A \in F^{m \times n}, B \in F^{h \times k} \), and \( C \in F^{u \times v} \), find \( X \in F^{p \times q} \) such that

\[
A \Join X + X \Join B = C. \tag{4.1}
\]

We say that Problem 4.1 is well-defined if all matrix dimensions are well-defined. We investigate criterions for existence and uniqueness of solutions for Eq. (4.1). Moreover, we transform the matrix equation into a linear system concerning the usual matrix product. We start the discussion with a simple case and then the general case. Let us denote \( t = \text{lcm} \{n, p\} \) and \( s = \text{lcm} \{q, h\} \).

4.1 The case \( m = u \) and \( q \mid h \)

In this subsection, we discuss Problem 4.1 when \( m = u \) and \( q \mid h \).

By the definition of the semi-tensor product, we have

\[
A \Join X = (A \otimes I_{\frac{m}{t}}) (X \otimes I_{\frac{p}{t}}) \in F^{\frac{m}{t} \times \frac{p}{t}} \quad \text{and} \quad X \Join B = (X \otimes I_{\frac{q}{s}}) B \in F^{\frac{q}{s} \times k}.
\]

Since \( C \in F^{u \times v} \), we have \( u = mt/n = ph/q \) and \( v = qt/p = k \). Since \( u = m \), we have \( t = n \) and \( p \mid n \). Thus, Problem 4.1 is well-defined if and only if (i) \( m = \frac{ph}{q} \) (in particular, \( p \mid m \)), (ii) \( v = \frac{qt}{p} = k \), and (iii) \( p \mid n \).

From now on, assume that Problem 4.1 is well-defined. We split \( A \) into \( p^2 \) equal-size blocks \( \hat{A}_{11}, \hat{A}_{12}, \ldots, \hat{A}_{pp} \in F^{\frac{p}{t} \times \frac{p}{t}} \), split \( B \) into \( q^2 \) equal-size blocks \( \hat{B}_{11}, \hat{B}_{12}, \ldots, \hat{B}_{qq} \in F^{\frac{q}{s} \times \frac{q}{s}} \), and split \( C \) into \( pq \) equal-size blocks.
\(\tilde{C}_{11}, \tilde{C}_{12}, \ldots, \tilde{C}_{pq} \in \mathbb{F}_{\frac{m}{p} \times \frac{n}{q}}\). So, we can write Eq. (4.1) as

\[A \otimes X + X \otimes B = A(X \otimes I_{\frac{n}{q}}) + (X \otimes I_{\frac{m}{p}})B\]

\[
\begin{bmatrix}
\tilde{A}_{11} & \tilde{A}_{12} & \cdots & \tilde{A}_{1p} \\
\tilde{A}_{21} & \tilde{A}_{22} & \cdots & \tilde{A}_{2p} \\
\vdots & \vdots & \ddots & \vdots \\
\tilde{A}_{p1} & \tilde{A}_{p2} & \cdots & \tilde{A}_{pp}
\end{bmatrix}
\begin{bmatrix}
x_{11} \otimes I_{\frac{n}{q}} & x_{12} \otimes I_{\frac{n}{q}} & \cdots & x_{1q} \otimes I_{\frac{n}{q}} \\
x_{21} \otimes I_{\frac{n}{q}} & x_{22} \otimes I_{\frac{n}{q}} & \cdots & x_{2q} \otimes I_{\frac{n}{q}} \\
\vdots & \vdots & \ddots & \vdots \\
x_{p1} \otimes I_{\frac{n}{q}} & x_{p2} \otimes I_{\frac{n}{q}} & \cdots & x_{pq} \otimes I_{\frac{n}{q}}
\end{bmatrix}
\begin{bmatrix}
\hat{B}_{11} & \hat{B}_{12} & \cdots & \hat{B}_{1q} \\
\hat{B}_{21} & \hat{B}_{22} & \cdots & \hat{B}_{2q} \\
\vdots & \vdots & \ddots & \vdots \\
\hat{B}_{q1} & \hat{B}_{q2} & \cdots & \hat{B}_{qq}
\end{bmatrix}
\]

\[
\begin{bmatrix}
\tilde{C}_{11} & \tilde{C}_{12} & \cdots & \tilde{C}_{1q} \\
\tilde{C}_{21} & \tilde{C}_{22} & \cdots & \tilde{C}_{2q} \\
\vdots & \vdots & \ddots & \vdots \\
\tilde{C}_{p1} & \tilde{C}_{p2} & \cdots & \tilde{C}_{pq}
\end{bmatrix}
\begin{bmatrix}
x_{11} & x_{12} & \cdots & x_{1q} \\
x_{21} & x_{22} & \cdots & x_{2q} \\
\vdots & \vdots & \ddots & \vdots \\
x_{p1} & x_{p2} & \cdots & x_{pq}
\end{bmatrix}
\]

Thus, for each \(i = 1, \ldots, p\) and \(j = 1, \ldots, q\), we get

\[\tilde{C}_{ij} = \tilde{A}_{ii}(x_{ij} \otimes I_{\frac{n}{q}}) + \cdots + \tilde{A}_{ip}(x_{pj} \otimes I_{\frac{n}{q}}) + (x_{i1} \otimes I_{\frac{m}{p}})\hat{B}_{1j} + \cdots + (x_{iq} \otimes I_{\frac{m}{p}})\hat{B}_{qj}
\]

\[= x_{ij}\hat{A}_{ii} + \cdots + x_{pj}\hat{A}_{ip} + x_{i1}\hat{B}_{1j} + \cdots + x_{iq}\hat{B}_{qj}. \quad (4.2)\]

The above equation says that we can write \(\tilde{C}_{ij}\) in terms of a linear combination of the block matrices \(\hat{A}_{ii}, \hat{A}_{ip}, \hat{B}_{1j}, \hat{B}_{2j}, \ldots, \hat{B}_{qj}\). Hence, we can deduce the following theorem.

**Theorem 4.2** Assume that \(m = u\) and \(q \mid h\), and Problem 4.1 is well-defined. If Eq. (4.1) has a solution, then for each \(i = 1, \ldots, p\) and \(j = 1, \ldots, q\),

(i) the set \(\{\tilde{A}_{i1}, \tilde{A}_{i2}, \ldots, \tilde{A}_{ip}, \hat{B}_{1j}, \hat{B}_{2j}, \ldots, \hat{B}_{qj}, \tilde{C}_{ij}\}\) is linearly dependent in the vector space \(\mathbb{F}_{\frac{m}{p} \times \frac{n}{q}}\).

(ii) \(\text{rank}[\hat{A}^t \hat{B}_j \hat{C}_{ij}] = \text{rank}[\hat{A}^t \hat{B}_j]\), where \(\hat{A}^t = [\hat{A}_{i1} \hat{A}_{i2} \cdots \hat{A}_{ip}] \) and \(\hat{B}_j = [\hat{B}_{1j} \hat{B}_{2j} \cdots \hat{B}_{qj}]\).
Remark 4.1 The converse of Theorem 4.2 does not hold in general. Consider the Sylvester matrix equation when

\[
A = \begin{bmatrix} 1 & 1 \\ 3 & 0 \end{bmatrix}, B = \begin{bmatrix} 1 & 0 & 2 \\ 2 & 1 & 3 \\ 1 & 1 & 2 \\ 1 & 2 & 3 \end{bmatrix}, \quad \text{and} \quad C = \begin{bmatrix} 0 & 1 & 0 \\ 3 & 1 & 1 \end{bmatrix}.
\]

We split \(B\) and \(C\) as follows:

\[
\hat{B}_{11} = \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}, \quad \hat{B}_{12} = \begin{bmatrix} 2 \\ 3 \\ 1 \end{bmatrix}, \quad \hat{B}_{21} = \begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix}, \quad \hat{B}_{22} = \begin{bmatrix} 2 \\ 3 \\ 1 \end{bmatrix},
\]

\[
\hat{C}_{11} = \begin{bmatrix} 0 \\ 1 \\ 3 \\ 1 \end{bmatrix}, \quad \hat{C}_{12} = \begin{bmatrix} 0 \\ 2 \\ 1 \\ 3 \end{bmatrix}.
\]

We get \(\text{rank}\left[\hat{A} \hat{B}_{11} \hat{C}_{11}\right] = \text{rank}\left[\hat{A} \hat{B}_{12}\right] = \text{rank}\left[\hat{A} \hat{B}_{21}\right] = \text{rank}\left[\hat{A} \hat{B}_{22}\right] = 2\).

However, the Sylvester equation has no solution.

Now for each \(i, j\), applying the vector operator to Eq. (4.2) yields

\[
\text{Vec}(\hat{C}_{ij}) = x_{1j} \text{Vec}(\hat{A}_{11}) + \cdots + x_{pj} \text{Vec}(\hat{A}_{1p}) + x_{i1} \text{Vec}(\hat{B}_{1j}) + \cdots + x_{iq} \text{Vec}(\hat{B}_{qj})
\]

\[
= \begin{bmatrix} \text{Vec}(\hat{A}_{11}) & \cdots & \text{Vec}(\hat{A}_{1p}) \\ \vdots & \ddots & \vdots \\ \text{Vec}(\hat{A}_{p1}) & \cdots & \text{Vec}(\hat{A}_{pp}) \end{bmatrix} \begin{bmatrix} x_{1j} \\ \vdots \\ x_{pj} \end{bmatrix} + [x_{i1} \cdots x_{iq}] \otimes \begin{bmatrix} \text{Vec}(\hat{B}_{1j}) \\ \vdots \\ \text{Vec}(\hat{B}_{qj}) \end{bmatrix}. \quad (4.3)
\]

Let us denote

\[
\hat{A} = \begin{bmatrix} \text{Vec}(\hat{A}_{11}) & \cdots & \text{Vec}(\hat{A}_{1p}) \\ \vdots & \ddots & \vdots \\ \text{Vec}(\hat{A}_{p1}) & \cdots & \text{Vec}(\hat{A}_{pp}) \end{bmatrix}, \quad \hat{B} = \begin{bmatrix} \text{Vec}(\hat{B}_{11}) & \cdots & \text{Vec}(\hat{B}_{1q}) \\ \vdots & \ddots & \vdots \\ \text{Vec}(\hat{B}_{q1}) & \cdots & \text{Vec}(\hat{B}_{qq}) \end{bmatrix},
\]

\[
\hat{C} = \begin{bmatrix} \text{Vec}(\hat{C}_{11}) & \cdots & \text{Vec}(\hat{C}_{1q}) \\ \vdots & \ddots & \vdots \\ \text{Vec}(\hat{C}_{p1}) & \cdots & \text{Vec}(\hat{C}_{pq}) \end{bmatrix}.
\]

From (4.3), we can put all \(i, j\) equations to get

\[
\hat{A}X + X \otimes \hat{B} = \hat{C}.
\]

Now taking the vector operator and utilizing Lemma 2.2 yields

\[
\text{Vec}(\hat{A}X) = (I_q \otimes \hat{A}) \text{Vec}(X).
\]
Let us denote each $j$-th column of $I_{m/p}$ by $e_{m/p}^{j}$, and write

$$K_{m,p,q} = I_q \otimes \begin{bmatrix} I_p \otimes e_{1/p}^{1/m} \\ \vdots \\ I_p \otimes e_{m/p}^{m/p} \end{bmatrix}.$$ 

By Lemmas 2.2 and 2.3, we obtain

$$\text{Vec}(X \circledast \tilde{B}) = \text{Vec} \left( (X \otimes I_{m/p}) \tilde{B} \right) = (\tilde{B} \otimes I_{m}) \text{Vec}(X \otimes I_{m/p}).$$

Hence, we obtain the following theorem.

**Theorem 4.3** Assume that $m = u$ and $q \mid h$, and Problem 4.1 is well-defined. Then, Eq. (4.1) is equivalent to the following linear system:

$$\left(I_q \otimes \tilde{A} + (\tilde{B}^T \otimes I_p)K_{m,p,q}\right)\text{Vec}(X) = \text{Vec}(\tilde{C}). \tag{4.4}$$

This theorem asserts that we can solve Eq. (4.1) by solving the linear system (4.4) instead.

**Remark 4.2** If $p = m$, then $K_{m,p,q} = I_{pq}$ and thus Eq. (4.4) reduces to

$$(I_q \otimes \tilde{A} + \tilde{B}^T \otimes I_p)\text{Vec}(X) = \text{Vec}(\tilde{C}).$$

**Corollary 4.4** Let $A \in F^{p \times p}, B \in F^{q \times q}$ and $C \in F^{p \times q}$. Then, the Sylvester matrix equation $AX + XB = C$ in unknown $X \in F^{p \times q}$ is equivalent to the following linear system

$$(I_q \otimes A + B^T \otimes I_p)\text{Vec}(X) = \text{Vec}(C).$$

**Proof.** This is a special case of Theorem 4.3 when $p = m = n = u$ and $q = h = k = v$. In this case, we have $\tilde{A} = A, \tilde{B} = B, \tilde{C} = C$, and $K_{m,p,q} = I_{pq}$. \hfill \Box

### 4.2 The general case

In this subsection, we investigate Problem 4.1 in a full generality.

By definition of the semi-tensor product, we have

$$A \circledast X = (A \otimes I_{\frac{m}{n}})(X \otimes I_{\frac{p}{n}}) \in F^{\frac{mp}{n} \times \frac{mn}{n}}$$

$$X \circledast B = (X \otimes I_{\frac{q}{n}})(B \otimes I_{\frac{p}{n}}) \in F^{\frac{mp}{n} \times \frac{qn}{n}}.$$

Thus, Problem 4.1 is well-defined if and only if
(i) \( u = \frac{ml}{n} = \frac{ps}{q}, \)
(ii) \( v = \frac{nt}{p} = \frac{ks}{h}. \)

From now on, assume that Problem 4.1 is well-defined. In order to reduce the Sylvester matrix equation (4.1) to a linear system, we make the following matrix partitions:

- split \( A \otimes I_q \) into \( p^2 \) equal-sizes blocks \( \tilde{A}_{11}, \tilde{A}_{12}, \ldots, \tilde{A}_{ip}, \tilde{B}_{1j}, \tilde{B}_{2j}, \ldots, \tilde{B}_{pq}, \) \( \tilde{C}_{ij}, \) where \( \tilde{A}_{ij} = x_{ij}I_q, \)
- split \( X \otimes I_q \) into \( pq \) blocks \( \tilde{X}_{11}, \tilde{X}_{12}, \ldots, \tilde{X}_{pq}, \) where \( \tilde{X}_{ij} = x_{ij}I_q, \)
- split \( B \otimes I_q \) into \( q^2 \) blocks \( \tilde{B}_{11}, \tilde{B}_{12}, \ldots, \tilde{B}_{pq}, \)
- split \( C \) into \( pq \) blocks \( \tilde{C}_{11}, \tilde{C}_{12}, \ldots, \tilde{C}_{pq} \in F_n^\times, \)

Now we can write Eq. (4.1) as

\[
A \times X + X \times B = (A \otimes I_q)(X \otimes I_q) + (X \otimes I_q)(B \otimes I_q)
\]

\[
= \begin{bmatrix}
\tilde{A}_{11} & \tilde{A}_{12} & \cdots & \tilde{A}_{1p} \\
\tilde{A}_{21} & \tilde{A}_{22} & \cdots & \tilde{A}_{2p} \\
\vdots & \vdots & \ddots & \vdots \\
\tilde{A}_{p1} & \tilde{A}_{p2} & \cdots & \tilde{A}_{pp}
\end{bmatrix}
\begin{bmatrix}
x_{11} \otimes I_q^p & x_{12} \otimes I_q^p & \cdots & x_{1q} \otimes I_q^p \\
x_{21} \otimes I_q^p & x_{22} \otimes I_q^p & \cdots & x_{2q} \otimes I_q^p \\
\vdots & \vdots & \ddots & \vdots \\
x_{p1} \otimes I_q^p & x_{p2} \otimes I_q^p & \cdots & x_{pq} \otimes I_q^p
\end{bmatrix}
\begin{bmatrix}
\tilde{B}_{11} & \tilde{B}_{12} & \cdots & \tilde{B}_{1q} \\
\tilde{B}_{21} & \tilde{B}_{22} & \cdots & \tilde{B}_{2q} \\
\vdots & \vdots & \ddots & \vdots \\
\tilde{B}_{q1} & \tilde{B}_{q2} & \cdots & \tilde{B}_{pq}
\end{bmatrix}
\]

\[
= \begin{bmatrix}
\tilde{C}_{11} & \tilde{C}_{12} & \cdots & \tilde{C}_{1q} \\
\tilde{C}_{21} & \tilde{C}_{22} & \cdots & \tilde{C}_{2q} \\
\vdots & \vdots & \ddots & \vdots \\
\tilde{C}_{p1} & \tilde{C}_{p2} & \cdots & \tilde{C}_{pq}
\end{bmatrix}
\]

By considering each block matrix in the above equation, we obtain

\[
\tilde{C}_{ij} = x_{1j}A_{i1} + \cdots + x_{pj}A_{ip} + x_{iq}B_{1j} + \cdots + x_{iq}B_{qj},
\]

(4.5)

From Eq. (4.5), we can conclude the following theorem.

**Theorem 4.5** Assume that Eq. (4.1) is well defined. If Eq. (4.1) has a solution, then for each \( i = 1, \ldots, p \) and \( j = 1, \ldots, q, \)

(i) the set \( \{A_{i1}, A_{i2}, \ldots, A_{ip}, B_{1j}, B_{2j}, \ldots, B_{qj}, C_{ij}\} \) is linearly dependent in \( F_n^\times. \)
(ii) \( \text{rank}[\hat{A}i \hat{B}j \hat{C}_{ij}] = \text{rank}[\hat{A}i \hat{B}j] \) where \( \hat{A}i = [\hat{A}_{i1} \hat{A}_{i2} \cdots \hat{A}_{ip}] \) and \( \hat{B}j = [\hat{B}_{1j} \hat{B}_{2j} \cdots \hat{B}_{qj}] \).

For each \( i, j \), we apply the vector operator to Eq. (4.5) and obtain the following equation

\[
\text{Vec}(\hat{C}_{ij}) = x_{1j} \text{Vec}(\hat{A}_{i1}) + \cdots + x_{pj} \text{Vec}(\hat{A}_{ip}) + x_{i1} \text{Vec}(\hat{B}_{1j}) + \cdots + x_{iq} \text{Vec}(\hat{B}_{qj})
\]

For convenience, denote

\[
\hat{A} = \begin{bmatrix}
\text{Vec}(\hat{A}_{11}) & \cdots & \text{Vec}(\hat{A}_{1p}) \\
\vdots & \ddots & \vdots \\
\text{Vec}(\hat{A}_{p1}) & \cdots & \text{Vec}(\hat{A}_{pp})
\end{bmatrix}, \quad \hat{B} = \begin{bmatrix}
\text{Vec}(\hat{B}_{11}) & \cdots & \text{Vec}(\hat{B}_{1q}) \\
\vdots & \ddots & \vdots \\
\text{Vec}(\hat{B}_{q1}) & \cdots & \text{Vec}(\hat{B}_{qq})
\end{bmatrix}, \quad \hat{C} = \begin{bmatrix}
\text{Vec}(\hat{C}_{11}) & \cdots & \text{Vec}(\hat{C}_{1q}) \\
\vdots & \ddots & \vdots \\
\text{Vec}(\hat{C}_{p1}) & \cdots & \text{Vec}(\hat{C}_{pq})
\end{bmatrix}.
\]

(4.7)

Putting Eq. (4.6) for all \( i, j \) together yields

\[
\hat{A}X + X \times \hat{B} = \hat{C}.
\]

Let us denote \( \gamma = \text{lcm}\{q, \frac{p}{d}\} \). Now, we can take the vector operator and apply Lemmas 2.2 and 2.3 to get an equivalent linear system as follows:

\[
\text{Vec}(\hat{C}) = \text{Vec}(\hat{A}X + X \times \hat{B}) = \text{Vec}(\hat{A}X) + \text{Vec}(X \times \hat{B}) = (I_q \otimes \hat{A}) \text{Vec}(X) + \text{Vec}((X \otimes I_{\frac{p}{d}})(\hat{B} \otimes I_{\frac{p}{d}})) = (I_q \otimes \hat{A}) \text{Vec}(X) + (\hat{B}^T \otimes I_{\frac{p}{d}} \otimes I_{\frac{p}{d}}) \text{Vec}(X \otimes I_{\frac{p}{d}}) = (I_q \otimes \hat{A}) \text{Vec}(X) + (\hat{B}^T \otimes I_{\frac{p}{d}} \otimes I_{\frac{p}{d}}) K_{\gamma,p,q} \text{Vec}(X),
\]

where

\[
K_{\gamma,p,q} = I_q \otimes \begin{bmatrix}
I_p \otimes e_1^{\frac{\gamma}{q}} \\
\vdots \\
I_p \otimes e_{\frac{\gamma}{q}}^{\frac{\gamma}{q}}
\end{bmatrix},
\]

\[
K_{\gamma,p,q} = I_q \otimes \begin{bmatrix}
I_p \otimes e_1^{\frac{\gamma}{q}} \\
\vdots \\
I_p \otimes e_{\frac{\gamma}{q}}^{\frac{\gamma}{q}}
\end{bmatrix}.
\]
Theorem 4.6 Assume that Problem 4.1 is well-defined. Let us denote 
\[ \beta = \gamma^2 p^2 / (stq) \]  
Then, Eq. (4.1) is equivalent to the following linear system:

\[
(I_q \otimes \hat{A} + (\hat{B}^T \otimes I_\beta) K_{\gamma,p,q}) \text{Vec}(X) = \text{Vec}(\hat{C}).
\]  

(4.8)

Here, \( \hat{A}, \hat{B}, \hat{C} \) are defined by Eq. (4.7).

This theorem asserts that we can solve Eq. (4.1) by solving the linear system (4.9) instead. Theorem 4.6 together with Kronecker-Capelli theorem (Theorem 2.1) allow us to derive equivalent conditions for the Sylvester equation (4.1) to have a solution or unique solution or infinitely many solutions, concerning ranks of the augmented matrix associated with the linear system (4.9).

Remark 4.3 Consider the case that \( n \mid p \) and \( h \mid q \), or equivalently, \( pq = st \). We have \( t = p, \gamma = q = s, \hat{C} = C, K_{\gamma,p,q} = I_{pq} \), and thus the linear system (4.9) becomes

\[
(I_q \otimes \hat{A} + \hat{B}^T \otimes I_p) \text{Vec}(X) = \text{Vec}(C).
\]

When \( B = 0 \), the Sylvester equation is reduced to the equation \( A \ltimes X = C \).

Corollary 4.7 Consider the matrix equation \( A \ltimes X = C \), where \( A, C \) are given matrices and \( X \) is an unknown. Then this equation is equivalent to the following linear system:

\[
(I_q \otimes \hat{A}) \text{Vec}(X) = \text{Vec}(\hat{C}),
\]  

(4.9)

where \( \hat{A} \) and \( \hat{C} \) are defined by Eq. (4.7).

5 The Lyapunov matrix equation

A special case \( B = A^T \) of the Sylvester equation is known as the Lyapunov equation.

Problem 5.1 Given, \( A \in F^{m \times n} \) and \( C \in F^{u \times v} \), find \( X \in F^{p \times q} \) such that

\[
A \ltimes X + X \ltimes A^T = C.
\]  

(5.1)

Let us denote \( t = \text{lcm}\{m,p\} \) and \( s = \text{lcm}\{n,q\} \). Then Problem 5.1 is well-defined if and only if \( m = n, u = t, v = s \), and \( p/q = t/s \).

From now on, suppose that Problem 5.1 is well-defined. Since \( q \mid s \), we can write \( s =fq \) for some \( f \in \mathbb{N} \). We get \( s/q = t/p = f \), and thus

\[
A \ltimes X + X \ltimes A^T = (A \otimes I_f)(X \otimes I_f) + (X \otimes I_f)(A^T \otimes I_{fn}).
\]

To transform Eq. (5.1) into a linear system, we make the following matrix partitions:
- split $A \otimes I_p$ into $p^2$ equal-sizes blocks $\hat{A}_{11}, \hat{A}_{12}, \ldots, \hat{A}_{pp} \in F^{f \times f},$
- split $X \otimes I_f$ into $pq$ blocks $\hat{X}_{11}, \hat{X}_{12}, \ldots, \hat{X}_{pq} \in F^{f \times f}$, where $\hat{X}_{ij} = x_{ij}I_f,$
- split $A^T \otimes I_s$ into $q^2$ blocks $\hat{B}_{11}, \hat{B}_{12}, \ldots, \hat{B}_{qq} \in F^{f \times f},$
- split $C$ into $pq$ blocks $\hat{C}_{11}, \hat{C}_{12}, \ldots, \hat{C}_{pq} \in F^{f \times f}.$

We now form the matrices $\tilde{A}, \tilde{B}, \tilde{C}$ as in Eq. (4.7). From Theorem 4.6, we obtain the following:

**Corollary 5.2** Assume that Problem 5.1 is well-defined. Then, the Lyapunov equation (5.1) is equivalent to the following linear system:

$$
(I_q \otimes \hat{A} + (\tilde{B}^T \otimes I_{f^2p})K_{f,p,q}) \text{Vec}(X) = \text{Vec}(\hat{C}),
$$

where

$$
K_{f,p,q} = I_q \otimes \begin{bmatrix} I_p \otimes e_1^T \\
\vdots \\
I_p \otimes e_{f^2}^T \end{bmatrix}.
$$

**Remark 5.1** Consider the Lyapunov equation (5.1) when the unknown $X$ is a square matrix, i.e. $p = q.$ Then $s = t,$ $f = 1,$ and $K_{f,p,q} = I_p^2.$ We also have $\hat{C} = C$ and $\hat{B}_{ij} = \hat{A}_{ji}^T$ for all $i,j.$ However, $\tilde{B}$ is not necessarily equal to $\tilde{A}.$ The linear system (5.2) is reduced to

$$
(I_p \otimes \tilde{A} + \tilde{B}^T \otimes I_p) \text{Vec}(X) = \text{Vec}(\tilde{C}).
$$

Our work also includes the study of the matrix equation $A \times X = C$ in [18] (by putting $B = 0$ in the Sylvester equation) and the equation $X \times B = C.$

**6 Conclusion**

We investigate the Sylvester equation $A \times X + X \times B = C$ and the Lyapunov equation $A \times X + X \times A^T = C$ where $A, B, C, X$ are matrices over an arbitrary field. Here, the product $\times$ is the semi-tensor product, which is a generalization of the usual matrix product. We consider the case that the unknown $X$ is a column vector and, in general, $X$ is a rectangular matrix. When all matrix dimensions are compatible, we find criteria for solvability and unique solvability for the matrix equations according to ranks and linear independence. Moreover, we show that the matrix equations under the semi-tensor product can be transformed into a linear system under the usual product, so that we can solve them via elementary methods. This work includes the studies of the equation $A \times X = C,$ the equation $X \times B = C,$ and the Sylverster/Lyapunov equations under the usual matrix product.
References


Received: 03.V.2022 / Revised: 07.VI.2022 / Accepted: 08.VI.2022
AUTHORS

PATTRAWUT CHANSANGIAM (Corresponding author),
Department of Mathematics,
School of Science,
King Mongkut’s Institute of Technology Ladkrabang,
Thailand,
E-mail: pattrawut.ch@kmitl.ac.th

SORIN V. SABAU,
Department of Biology,
Faculty of Biological Sciences,
Tokai University,
Japan,
E-mail: sorin@tokai.ac.jp