Solving the generalized Sylvester matrix equation

\(AV + BW = VF\) via Kronecker map

Aiguo WU 1,2, Siming ZHAO 2, Guangren DUAN 1,2

(1.Institute for Information and Control, Harbin Institute of Technology Shenzhen Graduate School, Shenzhen Guangdong 518055, China;
2.Center for Control Theory and Guidance Technology, Harbin Institute of Technology, Harbin Heilongjiang 150001, China)

Abstract: This note considers the solution to the generalized Sylvester matrix equation \(AV + BW = VF\) with \(F\) being an arbitrary matrix, where \(V\) and \(W\) are the matrices to be determined. With the help of the Kronecker map, an explicit parametric solution to this matrix equation is established. The proposed solution possesses a very simple and neat form, and allows the matrix \(F\) to be undetermined.

Keywords: Kronecker map; Sylvester matrix equation; Parametric solution

1 Introduction

When dealing with many problems for linear systems, such as eigenstructure assignment [1, 2], pole assignment [3, 4], output regulation [5], observers design and fault detection [6], the following generalized Sylvester matrix equation is often encountered:

\[AV + BW = VF,\]  \hfill (1)

where \(A \in \mathbb{R}^{n \times n}, B \in \mathbb{R}^{n \times p}\) and \(F \in \mathbb{R}^{p \times p}\) are known matrices, and \(V \in \mathbb{R}^{n \times p}\) and \(W \in \mathbb{R}^{r \times p}\) need to be determined.

Regarding the solution to (1), there exist several numerical solutions, such as the SVD-based block algorithm [7, 8], the large-scale algorithm [9, 10], the parallel algorithm [11], and so on. However, for several applications it is important to obtain symbolic solutions to such an equation. For example, in a robust pole assignment problem, one encounters optimization problems in which the criterion functions can be expressed in terms of the solution to the generalized Sylvester matrix equation (1) [12]. When \(F\) is in Jordan form, an attractive analytical and restriction-free solution for (1) is presented in [13]. Reference [1] proposes two solutions to matrix equation (1), also for the case that the matrix \(F\) is in Jordan form. These approaches can be found in related references [14]. The first one is in an iterative form, while the second one is in an explicit parametric form. To obtain the explicit solution given in [1], one needs to carry out a right coprime factorization or a series of singular value decompositions. An explicit solution expressed by a Hankel matrix, a symmetric operator matrix and a controllability matrix when the matrix \(F\) is in companion form is established in [15]. However, this solution has a restriction that the matrices \(A\) and \(F\) must have no common eigenvalues.

In this note, we aim to derive, for the generalized Sylvester matrix equation (1) with \(F\) being an arbitrary matrix, a general complete parametric solution in direct closed form with the help of the so-called Kronecker map. The concept of the Kronecker map was first introduced in [16] based on the underlying idea of [17]. In [16], the Kronecker map is utilized to prove the completeness of the proposed solution. Different from the idea in [16], in this note the Kronecker map is directly applied to obtaining the solution to the generalized Sylvester matrix equation (1).

Throughout this note, the symbol ‘\(\otimes\)’ denotes the Kronecker product of two matrices, \(\sigma(A)\) denotes the set of eigenvalues of matrix \(A\). For a matrix \(A \in \mathbb{C}^{m \times n}\), \(\text{vec}(A)\) is defined as

\[\text{vec}(A) = [a_1^T a_2^T \cdots a_n^T]^T,\]

where \(a_i\) is the \(i\)-th column of the matrix \(A\).

2 Kronecker map

Now we introduce the concepts of Sylvester sum and Kronecker map which were first proposed in [16].

Definition 1 Let \(T(s) = \sum_{i=0}^{t} T_i s^i \in \mathbb{R}^{m \times q}[s], F \in \mathbb{R}^{p \times p}\) and \(Z \in \mathbb{R}^{q \times p}\). The following matrix sum

\[\text{Syl}(T(s), F, Z) = \sum_{i=0}^{t} T_i Z F^i\]

is called the Sylvester sum associated with \(T(s), F\) and \(Z\).

Definition 2 Let \(T(s) = \sum_{i=0}^{t} T_i s^i \in \mathbb{R}^{m \times q}[s], F \in \mathbb{R}^{p \times p}\). The following map

\[\mathcal{F}[T(s)] = \sum_{i=0}^{t} (F^i) \otimes T_i\]

is called the \(F\)-Kronecker map of \(T(s)\). Based on the defined Kronecker map, we have the following relation:

\[\text{vec} \left( \text{Syl}(T(s), F, Z) \right) = \mathcal{F}[T(s)] \text{vec}(Z).\]

The following lemma gives the important property of the Kronecker map, which can be found in [16].

Lemma 1 Let \(X(s) \in \mathbb{R}^{q \times r}[s], Y(s) \in \mathbb{R}^{r \times m}[s]\), and
Let $F$ be a square matrix. Then
\[ F [X(s)] Y(s) = F[X(s)] F[Y(s)]. \]

Based on the above property of the Kronecker map, the following conclusions can be obtained. These conclusions can be found in [16].

**Lemma 2** Let $D(s) \in \mathbb{R}^{(n+r) \times r}[s], F \in \mathbb{R}^{p \times p}$. Then $\text{rank} F[D(s)] = rp$ if and only if $\text{rank} D(s) = r$ for any $s \in \sigma(F)$.

**Lemma 3** Let $T(s) \in \mathbb{R}^{n \times (n+r)}[s], F \in \mathbb{R}^{p \times p}$. Then $\text{rank} F[T(s)] = np$ if and only if $\text{rank} T(s) = n$ for any $s \in \sigma(F)$.

**3 Main result**

In this section, we discuss the solution to the matrix equation (1) with the help of the above so-called Kronecker map. Firstly, we show the degrees of freedom of the matrix equation (1) with the help of the Kronecker map.

**Lemma 4** Let $A \in \mathbb{R}^{n \times n}, B \in \mathbb{R}^{n \times r}, F \in \mathbb{R}^{p \times p}$. Then the matrix equation (1) has $rp$ degrees of freedom if and only if
\[ \text{rank} [A - sI B] = n, \forall s \in \sigma(F). \quad (2) \]

**Proof** Let
\[ T(s) = [A - sI B], X = \begin{bmatrix} V \\ W \end{bmatrix}, \]
then by applying the notations defined in the previous section, the matrix equation (1) can be compactly written as
\[ \text{Syl}(T(s), F, X) = 0. \quad (4) \]
This equation is equivalent to
\[ F [T(s)] \text{vec}(X) = \text{vec}(0). \quad (5) \]
It follows from linear algebra theory that the degrees of freedom of the matrix equation (1) is
\[ d = (n + r)p - \text{rank} F[T(s)]. \]
Combining this with Lemma 3 gives the conclusion immediately.

Regarding the solution to the matrix equation (1), we have the following result.

**Theorem 1** Let $A \in \mathbb{R}^{n \times n}, B \in \mathbb{R}^{n \times r}, F \in \mathbb{R}^{p \times p}$ satisfy
\[ \text{rank} [A - sI B] = n, \forall s \in \sigma(F). \]
Further, let $N(s) \in \mathbb{R}^{(n+r) \times r}[s]$ be a polynomial matrix satisfying
\[ [A - sI B] N(s) = 0_{n \times r}. \]
Then, 1) the matrices $V \in \mathbb{R}^{n \times p}$ and $W \in \mathbb{R}^{r \times p}$ given by
\[ \begin{bmatrix} V \\ W \end{bmatrix} = \text{Syl}(N(s), F, Z) \quad (6) \]
satisfy matrix equation (1); 2) when $\text{rank} N(s) = r$ for any $s \in \sigma(F)$, all the matrices $V$ and $W$ satisfying matrix equation (1) can be explicitly expressed by (6).

**Proof** Let $T(s)$ and $X$ be defined by (3), then matrix equation (1) is equivalent to
\[ \text{vec}(X) = F[N(s)] \text{vec}(Z). \quad (7) \]
Then we have
\[ F [T(s)] \text{vec}(X) = F[T(s)] F[N(s)] \text{vec}(Z) = F[T(s)] N(s) \text{vec}(Z) = F \text{vec}(Z) = \text{vec}(0). \]
This implies that the matrices $V$ and $W$ given by (6) satisfy matrix equation (1).

On the other hand, according to Lemma 2, we have $\text{rank} F[N(s)] = rp$ when $\text{rank} N(s) = r$ for any $s \in \sigma(F)$. In view of the equivalence between (6) and (7), this implies that the solution given in (6) can provide $rp$ degrees of freedom. Combining this fact with Lemma 4 gives the second conclusion.

**Remark 1** To obtain the explicit solution to the generalized Sylvester matrix equation (1), one should solve the so-called Diophantine polynomial equations. For numerical solutions to such a kind of equations, one can refer to [18].

**Remark 2** In this section, we provide a general complete parametric solution for the matrix equation (1). The presented solution is in an explicit form with respect to the matrix $F$. Therefore, this matrix $F$, together with the parameter matrix $Z$, can be further utilized to achieve some system performances in some applications. This will give great convenience and advantages to the practical applications.

**Remark 3** From the discussion in this paper, it has been seen that the Kronecker map is crucial and suitable to serve as the theoretical basis of the considered matrix equation. It is our hope that the further insight into the intrinsic property of the matrix equation mentioned in this paper can be developed with the help of the Kronecker map.

**4 A numerical example**

Consider a generalized Sylvester matrix in the form of (1) with the following parameters:
\[ A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, B = \begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix}. \]
By simple computations, we have
\[ N(s) = \begin{bmatrix} 1 & s & 0 & 0 & s^2 \\ 0 & 0 & 1 & s - 1 & 1 \end{bmatrix}^T. \]
It is easy to check that $\text{rank} N(s) = 2$ for any $s \in \mathbb{C}$. Thus for any matrix $F$, all the solutions to the matrix equation can be expressed by
\[ \begin{bmatrix} V \\ W \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} Z F^2 + \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} Z F + \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix} Z. \]
For the case that
\[ F = \begin{bmatrix} 1 & -1 \\ 2 & 1 \end{bmatrix}, \]
when specially choosing
\[ Z = \begin{bmatrix} -1 & 0 \\ 1 & 1 \end{bmatrix}, \]