Polynomial Solution of Sylvester Matrix Equation

M. E. Charkani

Department of Mathematics
Faculty of Sciences, Dhar-Mahraz
P. O. Box 1796, Atlas-Fez, Morocco

S. Bouarga

Department of Mathematics
Faculty of Science and Technology, FST Fez saiss
Fez, Morocco

Copyright © 2015 M. E. Charkani and S. Bouarga. This article is distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Abstract

Let $\mathcal{M}_n(K)$ (resp. $\mathcal{M}_{(n,m)}(K)$) be the ring of all square $n \times n$ (resp. of all rectangular $n \times m$) matrices with entries from a field $K$. Let $A \in \mathcal{M}_n(K)$, $B \in \mathcal{M}_m(K)$ and $C \in \mathcal{M}_{(n,m)}(K)$. The Sylvester matrix equation (SME)

$$AX - XB = C$$

(1)

where $X \in \mathcal{M}_{(n,m)}(K)$, play a central role in many areas of applied mathematics and in particular in systems and control theory. It is well known that if $K$ is an algebraically closed field then the matrix equation (1) possesses a unique solution if and only if the matrices $A$ and $B$ have no common eigenvalues (see [3] and [11]). In this work we give a brief survey of methods used to solve the (SME) and we study the solvability of the Sylvester matrix equation (1) over an arbitrary field $K$. By using some new results on Sylvester operator (see [6]) we show that the Sylvester matrix equation admits a unique solution if and only if the characteristic polynomial of $A$ and the characteristic polynomial of $B$ have no common prime factors. In this case, a wonderful polynomial solution of the Sylvester matrix equation is described.

Keywords: Sylvester operator, Sylvester matrix equation, centralizer space
1 Introduction

The Sylvester matrix equations (1), containing the Lyapunov matrix equation as a special case \( B = -A^T \), has numerous applications in control theory, signal processing, filtering, model reduction, image restoration, decoupling techniques for ordinary and partial differential equations, implementation of implicit numerical methods for ordinary differential equations, and block-diagonalization of matrices (for more details, see, for example \([3],[5],[7],[2],[12]\) as a few references). In \([9]\) the authors describe some methods to solve the (SME):

- To solve the homogeneous equation of (1) one can reduce A and B to their normal Jordan form and convert then to a set simple matrix equation.

- Stewart method and the Hassenberg-Schur method. These methods are based on transforming the coefficient matrices into Schur or Hassenberg form and then solving linear equations directly by a back-substitution process.

- By using the Kronecker Product [See 2].

In this work we give an independently proof using Taylor formula of the polynomial solution of (SME) (the third presented method in \([9]\)).

2 The Kronecker Product

2.1 The Vec Operator

The Vec operator transforms a matrix to a vector by stacking its columns on top of each other. If \( X \) is a matrix of order \( (n \times m) \) and \( x_i \) is the \( i^{th} \) column of \( X \) the Vec operator is defined by \( Vec(X) = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{pmatrix} \) from it follows that \( Vec(X) \) is an \( mn \) column vector. \( Vec(X + Y) = Vec(X) + Vec(Y) \).

2.2 The Kronecker Product

**Definition 2.1** (see[1] p. 193) The Kronecker Product between a matrix \( X \) of order \( (n \times m) \) and a matrix \( Y \) of order \( (o \times p) \) is defined as \( X \otimes Y = \)
Polynomial solution of Sylvester matrix equation

\[
\begin{pmatrix}
  x_{11}Y & x_{12}Y & \cdots & x_{1m}Y \\
x_{21}Y & x_{22}Y & \cdots & x_{2m}Y \\
  \vdots & \vdots & \ddots & \vdots \\
x_{n1}Y & x_{n2}Y & \cdots & x_{nm}Y
\end{pmatrix}
\]
where \( x_{ij} \) denotes the \( i^{\text{th}}, j^{\text{th}} \) element of \( X \).

Hence \( X \otimes Y \) is a matrix order \((n \times mp)\).

**Theorem 2.2** For the matrices \( X, Y \) and \( Z \) such that the product \( XYZ \) is defined, the following property holds:

\[
\text{Vec}(XYZ) = (Z^T \otimes X)\text{Vec}(Y)
\]

where \( Z^T \) denotes the transpose matrix of the matrix \( Z \).

**Proof.** See [10] and [1] \( \blacksquare \)

### 2.3 Application to solve Sylvester matrix equation

Let \( I_n \) stands for the identity matrix of order \( n \). We can rewrite the Sylvester matrix equation as \((I_m \otimes A - B^T \otimes I_n)\text{Vec}(X) = \text{Vec}(C)\) indeed \( AX - XB = C \) implies \( \text{Vec}(AX - XB) = \text{Vec}(C) \), hence \( \text{Vec}(AX) - \text{Vec}(XB) = \text{Vec}(C) \), hence \( \text{Vec}(AXI_m) - \text{Vec}(I_nXB) = \text{Vec}(C) \), so \((I_m \otimes A)\text{Vec}(X) - (B^T \otimes I_n)\text{Vec}(X) = \text{Vec}(C)\) i.e., \((I_m \otimes A - B^T \otimes I_n)\text{Vec}(X) = \text{Vec}(C)\).

### 3 Sylvester operator

Throughout this paper, we use the following notations. \( K \) will always denote a field. If \( P, Q \in K[X] \) are two polynomials, then \( \gcd(P, Q) \) denote the greater common divisor of the polynomials \( P \) and \( Q \) and \( \text{Res}(P, Q) \) is the resultant of \( P \) and \( Q \).

Let \( A \) be a square matrix, then \( C_A \) stands for the characteristic polynomial of \( A \) and \( m_A \) stands for the minimal polynomial of \( A \).

#### 3.1 Properties of Sylvester operator

**Definition 3.1** Let \( A \in \mathcal{M}_n(K) \) and \( B \in \mathcal{M}_m(K) \). The linear transformation

\[
\psi_{(A,B)} : \mathcal{M}_{n,m}(K) \rightarrow \mathcal{M}_{n,m}(K)
\]

\[
T \mapsto AT - TB
\]

is called the Sylvester operator associated to \( A \) and \( B \).
Recall that \( \psi(A, B) = \phi_A - \phi'_B \) where

\[
\phi_A : \mathcal{M}_{n,m}(K) \rightarrow \mathcal{M}_{n,m}(K) \quad T \mapsto AT
\]

and

\[
\phi'_B : \mathcal{M}_{n,m}(K) \rightarrow \mathcal{M}_{n,m}(K) \quad T \mapsto TB
\]

are commuting linear transformations associated to \( A \) and \( B \).

**Remark 3.2** Let \( \mathcal{C}(A, B) = \{ T \in \mathcal{M}_{n,m}(K) \mid AT = TB \} \) be the kernel of \( \psi(A, B) \). The subspace \( \mathcal{C}(A, B) \) of the \( K \)-vector space \( \mathcal{M}_{n,m}(K) \) is called the Sylvester space (or the centralizer space) of \( A \) and \( B \). In the sequel of the paper we will use the notation \( \psi \) instead of \( \psi(A, B) \).

**Proposition 3.3** Let \( K \) be a field. Let \( P \in K[X] \). Let \( A \in \mathcal{M}_n(K) \), \( B \in \mathcal{M}_m(K) \) and let \( \psi = \phi_A - \phi'_B \) be the Sylvester operator associated to \( A \) and \( B \). Then for any polynomial \( P \in K[X] \), the following results hold

- \( P(\phi_A) = \phi_P(A) \)
- \( P(\phi'_B) = \phi'_P(B) \)
- \( \psi^k(P(\phi'_B)(T)) = \psi^k(T)P(B) \) for all positive integer \( k \) and for all \( T \in \mathcal{M}_{n,m}(K) \).

**Proof.** Indeed, it suffice to notice that for all positive integer \( k \) and for all \( T \in \mathcal{M}_{n,m}(K) \):

- \( \phi_A^k(T) = A^kT \)
- \( \phi'_B(T) = TB^k \)
- \( \psi(TP(B)) = \psi(T)P(B) \).

**Corollary 3.4** Let \( K \) be a field. Let \( A \in \mathcal{M}_n(K) \), \( B \in \mathcal{M}_m(K) \) and let \( \psi \) be the Sylvester operator associated to \( A \) and \( B \). Then for any matrix \( T \in \mathcal{M}_{n,m}(K) \) and for all positive integer \( k \),

\[
\psi^k(T) = \sum_{i=0}^{k} C_i^k A^i T (-B)^{k-i}.
\]

**Proof.** Indeed \( \psi = \phi_A - \phi'_B \). As \( \phi_A \) and \( \phi'_B \) commute then it suffice to use the binomial formula.
4 Solvability of the (SME)

Let \( K \) be an algebraically closed field. Let \( A \in \mathcal{M}_n(K) \) and \( B \in \mathcal{M}_m(K) \). Let \( \sigma(A) \) denote the spectrum of the matrix \( A \) (i.e., the set of all eigenvalues of the matrix \( A \)). The Sylvester-Rosenblum Theorem on the solvability of the (SME) (see \([3], \S 2, p.2\)) states that \( \mathcal{C}(A, B) = 0 \) if and only if \( \sigma(A) \cap \sigma(B) = \emptyset \). This important theorem could be restated over any arbitrary field \( K \) as in the following result.

**Proposition 4.1** Let \( K \) be a field. Let \( A \in \mathcal{M}_n(K) \) and \( B \in \mathcal{M}_m(K) \). Then Sylvester matrix equation \( AX - XB = C \) admits a unique solution for any \( C \in \mathcal{M}_{(n,m)}(K) \) if and only if \( C_A \) and \( C_B \) have no common prime factors in \( K[X] \).

**Proof.** It is clear that (SME) admits a unique solution for any \( C \) if and only if the Sylvester operator associated to \( A \) and \( B \) is an isomorphism. It suffice then to use the following lemma.

**Lemma 4.2** Let \( K \) be a field. Let \( A \in \mathcal{M}_n(K) \) and \( B \in \mathcal{M}_m(K) \). Let \( \text{Res}(P, Q) \) denote the resultant of the polynomials \( P \) and \( Q \). The following are equivalent:

1. the Sylvester operator associated to \( A \) and \( B \) is an isomorphism
2. \( \mathcal{C}(A, B) = 0 \)
3. \( C_A \) and \( C_B \) have no common prime factors in \( K[X] \)
4. \( \text{Res}(C_A, C_B) \neq 0 \)
5. \( \text{Res}(m_A, m_B) \neq 0 \)

**Proof.** See \([6], Proposition 4.5, p.7\)

**Remark 4.3** Let \( K \) be a field. Let \( A \in \mathcal{M}_n(K) \) and \( B \in \mathcal{M}_m(K) \). Then \( C_A \) and \( C_B \) have no common prime factors in \( K[X] \) if and only if \( m_A \) and \( m_B \) have no common prime factors in \( K[X] \).

**Lemma 4.4** Let \( K \) be a field. Let \( A \in \mathcal{M}_n(K), \ B \in \mathcal{M}_m(K) \) and \( C \in \mathcal{M}_{(n,m)}(K) \). Any solution of the Sylvester matrix equations is the sum of a fixed particular solution of equation (1) and a solutions of the homogenous part of the matrix equation (SME) (equation (1) when \( C = 0 \)).
Proof. Indeed, if $X$ is the general solution of (1) and $X_0$ is a particular solution of (1) then $AX - XB = C$ and $AX_0 - X_0B = C$ so

$$A(X - X_0) - (X - X_0)B = 0$$

hence $(X - X_0)$ is the general solution of the homogenous part of the matrix equation (SME) (i.e., equation (1) when $C = 0$).

The following results are straightforward.

**Corollary 4.5** The set of the solutions of equation (1) is the affine space associated to the vector space of the solutions of the homogenous part of the matrix equation (1).

**Corollary 4.6** The dimension of the affine space of the solutions of equation (1) is equal to $\dim_K \mathcal{C}(A, B)$.

**Lemma 4.7** Let $K$ be field. Let $A \in M_n(K), B \in M_m(K)$ and $C \in M_{(n,m)}(K)$. The Sylvester matrix equation

$$AX - XB = C$$

where $X \in M_{(n,m)}(K)$, has a unique solution if and only if the unique solution of the homogenous equation $AX - XB = 0$ is $X = 0$.

### 4.1 The polynomial solution of (SME)

Let’s give a description to the polynomial solution of the Sylvester matrix equation. Indeed if $K$ is an arbitrary field and $P \in K[T]$, we denote by $\Delta^kP$ the coefficient of $Y^k$ in $P(X + Y)$. Then

$$P(X + Y) = \sum_{k=0}^{p} (\Delta^kP)(X)Y^k.$$ 

The following lemma is very useful to descrip the polynomial solution of (SME):

**Lemma 4.8** Let $K$ be a field. Let $P \in K[T]$ of degree $p$ and let $X, Y$ be two indeterminates over $K$. Then

- $\Delta^kP(X) \in K[X]$ for all $0 \leq k \leq p$
- $k!\Delta^kP(X) = P^{(k)}(X), \ P^{(k)}$ is the $k^{th}$ derivative of $P$.

**Proof.** See [[4], Proposition 5, p. 6]
Proposition 4.9  Let $K$ be a field. Let $P \in K[T]$ of degree $p$. Let $f, g \in \mathcal{L}_K(E)$ such that $f \circ g = g \circ f$. If $h = f - g$ then

$$P(f) = \sum_{k=0}^{p} \Delta^k P(h) g^k$$

Proof. $h = f - g$ implies $f = g + h$ and we deduce the result by substituting $X$ by $h$ and $Y$ by $g$ in the expression of $P(X + Y)$ above.

Remark 4.10 As $g = f - h$ then $P(f) = \sum_{k=0}^{p} \Delta^k P(g) h^k$.

Corollary 4.11 Let $K$ be a field. Let $P \in K\lbrack T \rbrack$ of degree $p$. Let $A \in M_n(K)$, $B \in M_m(K)$ and let $\psi = \phi_A - \phi_B$ be the Sylvester operator associated to $A$ and $B$. Then $P(\phi_A) - P(\phi_B) = (\sum_{k=1}^{p} \psi_{k-1} \circ \Delta^k P(\phi_B)) \circ \psi$

Proof. $\psi = \phi_A - \phi_B$ and $\phi_A, \phi_B$ are commuting linear transformations an by the proposition 4.9 and the remark 4.10. we have the result.

Theorem 4.12 Let $K$ be a field. $A \in M_n(K)$, $B \in M_m(K)$ and let $\psi = \phi_A - \phi_B$ be the Sylvester operator associated to $A$ and $B$. If the characteristic polynomial of $A$ and the characteristic polynomial of $B$ have no common prime factors then

i) $\psi$ is an isomorphism.

ii) There exists a polynomial $P \in K\lbrack X \rbrack$ of degree $p$ such that the inverse of $\psi$ is

$$\Phi = \left( \sum_{k=1}^{p} \psi_{k-1} \circ \Delta^k P(\phi_B) \right)$$

Proof. For i) see [6], Proposition 4.5, p.7]. For ii) since $\gcd(C_A, C_B) = 1$ there exists two polynomials $U \in K\lbrack X \rbrack, V \in K\lbrack X \rbrack$ such that $UC_A + VC_B = 1$. It suffice to choose $P = VC_B$. Such P verifies $P(A)=I$ and $P(B)=0$ so $P(\phi_A) = \phi_P(\phi_A) = I$ and $P(\phi_B) = \phi_P(\phi_B) = 0$. Hence by corollary 4.11 we have $\Phi \circ \psi = id_V$, where $V = M_{(n,m)}(K)$.

Corollary 4.13 Let $K$ be a field. Let $A \in M_n(K)$, $B \in M_m(K)$ and $C \in M_{(n,m)}(K)$. If $AT - TB = C$ and if $\gcd(C_A, C_B) = 1$ then there exists a polynomial $P \in K\lbrack X \rbrack$ of degree $p$ such that

$$T = \sum_{k=1}^{p} \sum_{i=0}^{k-1} A^i C \eta_k(B)$$

where $\eta_k(B) = C_{k-1}^i(-B)^{k-i-1} \frac{1}{k!} P^{(k)}(B)$ is a polynomial matrix of $B$. 
Proof. Indeed $AT - TB = C$ is equivalent to $\psi(T) = C$. If $gcd(C_A, C_B) = 1$ then $\psi(T) = C$ is equivalent to $T = \Phi(C)$. By the above theorem there exists a polynomial $P \in K[X]$ of degree $p$ such that $T = \Phi(C) = \left( \sum_{k=1}^{p} \psi^{p-1} \circ \Delta^k P(\phi_B) \right)(C)$. Hence by straightforward calculations we deduce that $T = \sum_{k=1}^{p} \frac{1}{k!} \psi^{p-1}(C) P(k)(B)$. By corollary 3.4 we got the result.

Corollary 4.14 With the same assumption as above if $P = \sum_{s=0}^{p} a_s X^s$ then

$$T = \sum_{k=1}^{p} \sum_{i=0}^{k-1} \sum_{s=k}^{p} f_{i,s,k} A^i C B^{s-i-1}$$

where $f_{i,s,k} = C_{k-1}^i C_k^s (-1)^{k-i-1} a_s$

Proof. Indeed, if $P = \sum_{s=0}^{p} a_s X^s$ then $P(k) = \sum_{s=k}^{p} a_s \frac{s!}{(s-k)!} X^{s-k}$.

$$\eta_k(B) = C_{k-1}^i (-B)^{k-i-1} \frac{1}{k!} \sum_{s=k}^{p} a_s \frac{s!}{(s-k)!} B^{s-k}$$

$$= \sum_{s=k}^{p} (-1)^{k-i-1} a_s C_{k-1}^i \frac{s!}{(s-k)!} \frac{1}{k!} B^{s-i-1}$$

$$= \sum_{s=k}^{p} (-1)^{k-i-1} a_s C_{k-1}^i C_k^s B^{s-i-1}$$

Hence

$$T = \sum_{k=1}^{p} \sum_{i=0}^{k-1} \sum_{s=k}^{p} f_{i,s,k} A^i C B^{s-i-1}$$

Remark 4.15 Recall that $C_n^p C_q^r = C_n^a C_{n-q}^{n-p}$ and $C_n^{p+1} = \frac{n-p}{p+1} C_n^p$. Then $\eta_k(B) = \sum_{s=k}^{p} (-1)^{k-i-1} a_s C_{s-k+1}^{s-i+1} C_{s-k+1}^{s-i+1} B^{s-i-1}$

4.2 Implementation of (SME)

We describe briefly the implementation of the solution of the (SME). With the last corollary and the last remark one can easily implement the solution using maple software. and the commands in maple: $gcd(C_A, C_B)$ to test if $gcd(C_A, C_B) = 1$ if this is the case we can use the command $gcdex(C_A, C_B, U', V')$ to determine $U, V$ such that $UC_A + VC_B = 1$ and hence to compute the polynomial $P$. And by using the command add we can compute the solution $T$ from the formula above.
Algorithm 1
\>
> with(LinearAlgebra);
> Insert the matrices A,B,C;
> \(C_A:=\text{characteristic}(A, x)\);
> \(C_B:=\text{characteristic}(B, x)\);
> \(\gcd(C_A, C_B)\);
> \(\gcdex(C_A, C_B, x, 'U', 'V')\);
> \(V\);
> \(P := \text{expand}(VC_B)\);
> \(p := \text{degree}(P)\);
> \(T := \text{add(add(add(evalm(((binomial(k-1,i)*(-1)^(k-i-1)*binomial(s, k)*\text{coeff}(P, x, s)*A^i)*C)*B^{(s-i-1)})), s = k..p), i = 0..k-1), k = 1..p)\);
> evalm(T);
> evalm((A*T)-(T*B));

5 Conclusion

In this work we study the solvability of the Sylvester matrix equation (1) over an arbitrary field \(K\). By using some new results on Sylvester operator (see [[6]]) we show that the Sylvester matrix equation admits a unique solution if and only if the characteristic polynomial of \(A\) and the characteristic polynomial of \(B\) have no common prime factors. In this case, a wonderful polynomial solution of the Sylvester matrix equation is described.

References


Received: March 16, 2015; Published: May 15, 2015