Properties of sub-matrices of Sylvester matrices and triangular toeplitz matrices

Yousong Luo, Robin Hill and Uwe Schwerdtfeger
School of Mathematical and Geospatial Sciences,
RMIT University, GPO Box 2476V
Melbourne, Vic. 3001, AUSTRALIA
e-mail: yluo@rmit.edu.au, r.hill@rmit.edu.au, u.schwerdtfeger@rmit.edu.au

Abstract

In this note we discover and prove some interesting and important relations among sub-matrices of Sylvester matrices and triangular toeplitz matrices. The main result is Hill’s identity discovered by R. D. Hill which has an important application in optimal control problems.

1 Introduction

When studying the optimal state evolution of the dual state in an optimal control problem, R. Hill discovered an interesting relation (see Theorem 1.1) among the sub-matrices of Sylvester matrices and triangular toeplitz matrices, see [2] and [3] for details. If these relations holds then we can formulate the exact pattern how the modified states evolve. In such a sense, the result here is not only an interesting result in linear algebra but also has a direct significant impact in control theory.

We would like also to announce that we have an alternative proof for Theorem 1.1 using the tools given in [1] which is an entirely different approach.

We formulate the problems first. Define the following $m \times m$ lower and upper triangular matrices:

\[
D_L := \begin{pmatrix}
  d_1 & 0 & \cdots & \cdots & 0 \\
  d_2 & d_1 & 0 & \cdots & 0 \\
  \vdots & \ddots & \ddots & \ddots & \vdots \\
  d_{m-1} & \ddots & \ddots & 0 & \vdots \\
  d_m & d_{m-1} & \cdots & d_2 & d_1
\end{pmatrix}

D_U := \begin{pmatrix}
  d_{m+1} & d_m & \cdots & d_3 & d_2 \\
  0 & d_{m+1} & d_m & \cdots & d_3 \\
  \vdots & \ddots & \ddots & \ddots & \vdots \\
  \vdots & \ddots & \ddots & \ddots & d_m \\
  0 & \cdots & \cdots & 0 & d_{m+1}
\end{pmatrix}
\]
Consider the Sylvester matrix
\[ S := \begin{pmatrix} D & N_L \\ D_U & N_U \end{pmatrix} \]
and the lower triangular matrix
\[ D := \begin{pmatrix} D_L & 0 \\ D_U & D_L \end{pmatrix}. \]

The entries \( d_1, d_2, \ldots, d_m, d_{m+1} \) and \( n_1, n_2, \ldots, n_m, n_{m+1} \) are assumed to be nonzero real numbers such that both \( S \) and \( D \) are invertible. Under such an assumption we define
\[ A := D^{-1} \quad B := S^{-1}. \]

If we use \( A_T \) and \( B_T \) to denote the matrices consisting of the first \( m \) rows of \( A \) and \( B \), \( A_B \) and \( B_B \) the last \( m \) rows of \( A \) and \( B \) respectively, then we can write
\[ A = \begin{pmatrix} A_T \\ A_B \end{pmatrix} \quad \text{and} \quad B = \begin{pmatrix} B_T \\ B_B \end{pmatrix}. \]

The \( m \times m \) sub-matrices of \( A_B \) consisting of the \( m \) consecutive columns of it and starting from the \( i \)th column is denoted by \( A_i \). There are \( m + 1 \) of them:
\[ A_1, A_2, \ldots, A_m, A_{m+1}. \tag{1} \]

Similarly, the sub-matrices of \( B_B \) consisting of \( m \) consecutive columns of it and starting from the \( i \)th column is denoted by \( B_i \):
\[ B_1, B_2, \ldots, B_m, B_{m+1}. \tag{2} \]

Our objective of this paper is to prove these relations, as well as discover and prove some other new relations among those sub-matrices. The main result is the following Hill’s identity.

**Theorem 1.1** For \( 1 \leq i < j \leq m+1 \) we have
\[ A_i B_j = A_j B_i. \tag{3} \]
Theorem 1.2 Assume that both $S$ and $D$ be invertible. Let $A_i$ and $B_j$ be the sub matrices defined in (1) and (2). Then, for all $i, j = 1, \ldots, m+1$, $A_i$ and $B_j$ are invertible and the following identities hold

$$A_i^{-1}A_j = B_i^{-1}B_j$$

(4)

or equivalently

$$A_jB_j^{-1} = A_iB_i^{-1}.$$  

(5)

and

Theorem 1.3 For $1 \leq i < j \leq m + 1$ we have

$$B_i^{-1}B_j = B_jB_i^{-1}.$$  

(6)

As we can easily see that Theorem 1.1 is a consequence of the combination of Theorem 1.2 and 1.3.

2 Proofs of the results

Now we introduce an $m \times 3m$ matrix

$$T := ( - D_U D_L^{-1} \ | \ I_m \ | \ - D_L D_U^{-1} )$$

(7)

where the symbol $\mid$ stands for an augmentation bar. This matrix $T$ plays a very important role in the following argument throughout this paper, so we call it “kernel”. The $m \times 2m$ sub-matrices of $T$ consisting of the $2m$ consecutive columns of it starting from the $i$th column is denoted by $T_i$ and we have $m + 1$ such matrices:

$$T_1, T_2, \ldots, T_m, T_{m+1}.$$  

Obviously $T_1 = ( - D_U D_L^{-1}, I_m )$ and $T_{m+1} = ( I_m, - D_L D_U^{-1} )$. Also, For each $i, j = 1, 2, \ldots, m + 1$, the $m \times m$ sub-matrices of $T_i$ consisting of the $m$ consecutive columns of $T_i$ starting from the $j$th column is denoted by $T_{ij}$.

Lemma 2.1 If $K = \begin{pmatrix} D_L & 0 \\ D_U & D_L \\ 0 & D_U \end{pmatrix}$, then

$$TK = 0.$$  

(8)

If $D_l = \begin{pmatrix} D_L \\ D_U \end{pmatrix}$, then for $i = 1, 2, \ldots, m + 1$ we have

$$T_iD_l = 0.$$  

(9)
Proof. Obviously
\[
TK = \begin{pmatrix} -DU & I_m & -DL & DU \end{pmatrix} \begin{pmatrix} DL & 0 \\ DU & DL \\ 0 & DU \end{pmatrix} = \begin{pmatrix} -DL & DL + DU & DL & DU \end{pmatrix} = \begin{pmatrix} 0 & 0 \end{pmatrix}.
\]
This immediately implies, by considering the first \(m\) columns and the last \(m\) columns of \(TK\), that
\[
T_1D_l = 0 \quad \text{and} \quad T_{m+1}D_l = 0.
\]
For \(1 < i < m+1\) let \(K_i\) be the \(m\) consecutive columns of \(K\) starting from the \(i\)th column. Then \(K_i\) is in the form
\[
K_i = \begin{pmatrix} O_i \\ D_l \\ O_{m-i} \end{pmatrix}
\]
where \(O_i\) is an \(i \times m\) zero matrix and \(O_{m-i}\) is an \((m-1)i \times m\) zero matrix. Therefore
\[
T_iD_l = TK_i = 0.
\]
Prove of Theorem 1.2. We define
\[
D_r := \begin{pmatrix} 0 \\ DL \end{pmatrix}
\]
and hence
\[
D = \begin{pmatrix} D_l & D_r \end{pmatrix}.
\]
By Lemma 2.1, \(T_iD_l = 0\). Then, for \(i, j = 1, \ldots, m+1\), we have
\[
T_i = T_iDA = T_i \begin{pmatrix} D_l & D_r \end{pmatrix} A = \begin{pmatrix} 0 & T_iD_r \end{pmatrix} \begin{pmatrix} A_T \\ AB \end{pmatrix} = T_iD_rA_B
\]
which implies
\[
T_{ij} = T_iD_rA_j.
\]
From the definition of \(T\) we can see that \(T_{m-i+2,i} = I\). Then we have
\[
I = T_{m-j+2,m+1}D_rA_j,
\]
that is \(A_j\) is invertible and
\[
A_j^{-1} = T_{m-j+2,m+1}D_r
\]
or
\[
T_iD_r = (A_{m-i+2})^{-1}.
\]
By substituting (15) into (13) we obtain
\[ T_i = (A_{m-i+2})^{-1}A_B \quad \text{or} \quad A_i^{-1}A_B = T_{m-i+2}. \] (16)
This implies that
\[ A_i^{-1}A_j = T_{m-i+2,j}. \] (17)
On the other hand we perform the same process to \( B \) as follows. We define
\[ N := \begin{pmatrix} N_L & N_U \\ N_L & N_U \end{pmatrix}. \] (18)
By Lemma 2.1 we have, for \( i, j = 1, \ldots, m + 1, \)
\[ T_i = T_iSB = T_i \begin{pmatrix} D_i & N \end{pmatrix} B = \begin{pmatrix} 0 & T_iN \end{pmatrix} \begin{pmatrix} B_T \\ B_B \end{pmatrix} = T_iNB_B \] (19)
which implies
\[ T_{ij} = T_iNB_j. \]
From the definition of \( T \) we know that \( T_{m-i+2,i} = I. \) Then we have
\[ I = T_{m-j+2}NB_j, \]
that is
\[ T_{m-j+2}N = B_j^{-1} \] (20)
or
\[ T_iN = (B_{m-i+2})^{-1}. \] (21)
By substituting (21) into (19) we obtain
\[ T_i = (B_{m-i+2})^{-1}B_B \quad \text{or} \quad B_i^{-1}B_B = T_{m-i+2}. \] (22)
This implies that
\[ B_i^{-1}B_j = T_{m-i+2,j}. \] (23)
Equations (17) and (23) show that
\[ A_i^{-1}A_j = B_i^{-1}B_j \]
for each \( i, j = 1, 2, \ldots, m + 1. \) This completes the proof. QED

**Corollary 2.2** We define
\[ M := \begin{pmatrix} M_1 & M_2 \end{pmatrix} = \begin{pmatrix} N_L & 0 \\ N_U & N_L \\ 0 & N_U \end{pmatrix}. \] (24)
Let \( H = TM \) and \( H_i \) be the sub-matrix of \( H \) consisting the \( m \) consecutive columns of \( H \) starting from the \( i \)th column. Then
\[ H_i = (B_{m-i+2})^{-1} \quad \text{or} \quad H_{m-i+2} = B_i^{-1}. \]
Proof  Consider

\[ H = TM = T \begin{pmatrix} N_L & 0 \\ N_U & N_L \\ 0 & N_U \end{pmatrix} = \begin{pmatrix} T_1N & T_{m+1}N \end{pmatrix}. \] (25)

This gives immediately

\[ H_1 = T_1N \quad \text{and} \quad H_{m+1} = T_{m+1}N. \] (26)

Equations (21) then implies \( H_1 = (B_{m+1})^{-1} \) and \( H_{m+1} = B_1^{-1} \). For \( 1 < i < m + 1 \) let \( M_i \) be the sub-matrix of \( M \) consisting the \( m \) consecutive columns of \( M \) starting from the \( i \)th column. Then \( M_i \) is in the form

\[ M_i = \begin{pmatrix} O_i \\ N \\ O_{m-i} \end{pmatrix} \]

where \( O_i \) is an \( i \times m \) zero matrix and \( O_i \) is an \((m - 1)i \times m \) zero matrix. Therefore

\[ H_i = TM_i = T_iN. \] (27)

Again, equations (21) shows \( H_i = (B_{m-i+2})^{-1} \).

QED

Remark 2.3 This theorem reveals two remarkable features of \( A_i \)'s and \( B_i \)'s. First, equation (3) demonstrates the invariance of \( A_i B_i^{-1} \) with respect to \( i \). More precisely we have

\[ A_i B_i^{-1} = A_B N. \]

Secondly, equation (4) shows that \( B_i^{-1} B_j \) is independent of \( n_h \)'s which are the elements defining \( S \). This is quite significant as \( B_i \)'s are sub-matrices of \( B \), which is the inverse of \( S \) and therefore depends on \( n_h \)'s.

Remark 2.4 The proof of this theorem also demonstrates an interesting feature of those \( A_i \)'s and \( B_i \)'s. By the definition of \( T \) we can see that, for \( i, j = 1, 2, \ldots, m + 1 \) and \( 1 \leq k \leq \max\{m - i + 1, j\} \) we have

\[ T_{i+k,j-k} = T_{i,j}. \]

This, together with (17) and (23), shows that

\[ A_i^{-1} A_j = (A_{i+k})^{-1} A_{j+k} \quad \text{and} \quad B_i^{-1} B_j = (B_{i+k})^{-1} B_{j+k} \] (28)

for such \( k \)'s that the right hand sides of the above equations are defined. For example,

\[ B_1^{-1} B_2 = B_2^{-1} B_3 = \cdots = B_{m-1}^{-1} B_{m+1}. \]
Proof of Theorem 1.3  It is well known that $B$ can be represented by

$$B = \begin{pmatrix} N_U B_z & -N_L B_z \\ -D_U B_z & D_L B_z \end{pmatrix}$$

(29)

where $B_z = B_T(D, N)^{-1}$ where $B_T(D, N)$ is the Bezoutian matrix generated by $D$ and $N$ in the following manner:

$$B_T(D, N) = D_L N_U - N_L D_U = N_U D_L - D_U N_L.$$  

(30)

For detailed properties of Bezoutian matrices we refer to the comprehensive article [1]. Using this representation we have $B_1 = -D_U B_z$ and $B_{m+1} = D_L B_z$.

Now, by Corollary 2.2 we have

$$B_1 H = B_1 \begin{pmatrix} (B_{m+1})^{-1} & B_1^{-1} \\ (-D_U B_z B_z^{-1} D_L^{-1}) & I \end{pmatrix} = \begin{pmatrix} I & -D_U D_L \\ 0 & I \end{pmatrix} = T_1$$

and hence

$$B_1 B_i^{-1} = B_1 H_{m-i+2} = T_{1,m-i+2}.$$

This, together with equation (22), implies

$$B_1 B_{i}^{-1} = (B_{m+1})^{-1} B_{m-i+2}.$$  

Putting $k = m - i + 1$ in (28) gives

$$B_{i}^{-1} B_1 = (B_{i+k})^{-1} B_{1+k} = (B_{m+1})^{-1} B_{m-i+2}.$$  

Therefore $B_1 B_{i}^{-1} = B_1^{-1} B_1$ for each $i = 1, 2, \ldots, m + 1$.

Similarly

$$B_{m+1} H = B_{m+1} \begin{pmatrix} (B_{m+1})^{-1} & B_1^{-1} \\ (I D_L B_z (-B_z^{-1} D_U^{-1})) & (I -D_L D_U^{-1}) \end{pmatrix} = T_{m+1}.$$  

This, together with equation (22), proves

$$B_{m+1} B_{i}^{-1} = T_{m+1,m+2-i} = B_1^{-1} B_{m+2-i}.$$  

Equation (28) with $k = i - 1$ gives

$$B_{1}^{-1} B_{m+2-i} = (B_{1+i-1})^{-1} B_{m+2-i+i-1} = B_{i}^{-1} B_{m+1}.$$
and hence $B_{m+1}B_i^{-1} = B_i^{-1}B_{m+1}$ for each $i = 1, 2, \ldots, m + 1$. This is equivalent to
\begin{align*}
B_i(B_{m+1})^{-1}& = (B_{m+1})^{-1}B_i. \tag{31}
\end{align*}
Now for $1 < i < m + 1$, by equation (23)
\begin{align*}
B_iH &= B_i\left( (B_{m+1})^{-1} \ B_i^{-1} \right) = \left( B_i(B_{m+1})^{-1} \ B_iB_i^{-1} \right) \\
&= \left( (B_{m+1})^{-1}B_i \ B_i^{-1}B_i \right) \\
&= \left( T_{1,i} \ T_{m+1,i} \right).
\end{align*}
Let $t_j$ denote the $j$th column of $T$. The observation
\begin{equation}
T = (t_1, \ldots, t_{i-1}, \underbrace{T_{1,i}, \ldots, T_{m+1,i}}_{T_i}, \underbrace{t_{m+i}, \ldots, t_{2m+i-1}}_{T_{m+i}}, \ldots, t_{2m}, \ldots, t_{3m}) \tag{32}
\end{equation}
shows that
\begin{align*}
\left( T_{1,i} \ T_{m+1,i} \right) &= T_i,
\end{align*}
and hence
\begin{align*}
B_iH &= T_i. \tag{33}
\end{align*}
From this we obtain $B_jB_i^{-1} = B_i^{-1}B_j$. QED

**Corollary 2.5** For $i, j = 1, 2, \ldots, m + 1$ we have
\begin{align*}
B_iB_j &= B_jB_i, \tag{34}
\end{align*}
and, for all $l$ such that both $B_{i+l}$ and $B_{j-l}$ are meaningful,
\begin{align*}
B_iB_j &= B_{i+l}B_{j-l}. \tag{35}
\end{align*}

**Proof** The second equation follows from (28) by putting $k = i - j + l$:
\begin{align*}
B_{j-l}B_j^{-1} &= B_j^{-1}B_{j-l} = (B_{j+k})^{-1}B_{j-l+k} = (B_{i+l})^{-1}B_i.
\end{align*}

References

[1] Georg Heinig and Karla Rost, *Introduction to Bezoutians*, Advances and Applications, Vol. 199, 25 - 118, (2010)

[2] Robin Hill, Uwe Schwerdtfeger and Michael Baake, *Dynamic programming and duality applied to an optimal control problem*, Proceeding of Australian Control Conference, to appear, (2011)

[3] Robin D. Hill, *Dual periodicity in $l_1$-norm minimisation problems*, Systems & Control Letters, 57, 489 - 496, (2008)