The Matrix Equation $XA - BX = R$ and Its Applications

Karabi Datta*
Department of Mathematical Sciences
Northern Illinois University
DeKalb, Illinois 60115

Submitted by David H. Carlson

ABSTRACT

We study the well-known Sylvester equation $XA - BX = R$ in the case when A and B are given and R is known up to its first $n-1$ rows. We prove new results on the existence and uniqueness of X. Our results essentially state that, in case A is a nonderogatory matrix, there always exists a solution to this equation; a solution is uniquely determined by its first row x_1; and there is an interesting relationship between x_1 and the rows of R. We also give a complete characterization of the nonsingularity of X in this case. As applications of our results we develop direct methods for constructing symmetrizers and commuting matrices, computing the characteristic polynomial of a matrix, and finding the numbers of common eigenvalues between A and B. Some well-known important results on symmetrizers, Bezoutians, and inertia are recovered as special cases.

I. INTRODUCTION

The Sylvester matrix equation

$$XA - BX = R, \tag{1.1}$$

*This work was partially supported by Air Force Office of Scientific Research Grant AFOSR-83-0334.

655 Avenue of the Americas, New York, NY 10010 0024-3795/88/$3.50

where A, B, and R are given complex matrices of appropriate dimensions and X is the unknown matrix, has been widely studied in the literatures of linear algebra and control theory. It is known that a solution X to this matrix equation may or may not exist, and when a solution exists, it is unique iff A and B do not have an eigenvalue in common. No complete characterization of nonsingular solutions is available.

In this paper, we study the Sylvester equation in the case when A and B are given and R is known up to its first $n-1$ rows. The equation (1.1), in the special case when $B = -\bar{A}$ is a normalized lower Hessenberg matrix and the first $n-1$ rows of R are zero, was used earlier by Carlson and Datta [3] in connection with developing an algorithm for computing the inertia of A, and recently by Datta [7] and Datta and Datta [8] for solving the well-known eigenvalue and canonical form assignment problems in control theory. Here, with B as a normalized lower Hessenberg matrix and A as an arbitrary matrix, we prove a new result (Theorem 1) on the existence and uniqueness of a solution X to (1.1). Our results essentially state that, in contrast to the known results on the Sylvester equation, the equation (1.1) in this case always admits a solution, a solution is uniquely determined by its first row x_1, and there is an interesting relationship between x_1 and the rows of R. We also give a complete characterization of the nonsingularity of X in this case. As applications of our results we develop direct methods for constructing symmetrizers and commuting matrices, computing the characteristic polynomial of a matrix, and finding the number of common eigenvalues between A and B. Some well-known important results on symmetrizers, commuting matrices, bezoutians, and inertia are recovered as special cases.

In conclusion here, we remark that, though we assume (for the sake of simplicity only) that B is a normalized lower Hessenberg matrix, the results of this paper are valid (with trivial modifications) for any nonderogatory matrix B. Note that a nonderogatory matrix B is orthogonally similar to an unreduced Hessenberg matrix, that is, a Hessenberg matrix with nonzero codiagonal. An unreduced Hessenberg matrix can further be reduced to a normalized Hessenberg matrix by diagonal similarity [3].

II. A THEOREM ON THE EXISTENCE AND UNIQUENESS OF THE SOLUTION OF $XA - BX = R$

THEOREM 1. *Let $B = (b_{ij})$ be an $n \times n$ normalized lower Hessenberg matrix—that is, $b_{ij} = 0$ if $j > i+1$ and $b_{i\,i+1} = 1$, $i = 1, \dots, n-1$—and let A be arbitrary. Let $r_1, r_2, \dots, r_{n-1}$ be $n-1$ vectors in row n-space.*

(i) *there always exists an* X *such that* $XA - BX = R$ *has as its first* $n - 1$ *rows* r_1 *through* r_{n-1}.

(ii) X *is uniquely determined by* x_1, *the first row of* X.

(iii) *Let* r_n *be the nth row of* R. *Then*

$$x_1\phi(A) = r_1 D_{11} + \cdots + r_{n-1} D_{n-1,1} + (-1)^{n+1} r_n,$$

where D_{ij} *is the* $n \times n$ *matrix which is the cofactor of the element* d_{ij} *in the matrix*

$$D = \begin{pmatrix} b_{11}I - A & I & 0 & \cdot & \cdot & 0 \\ b_{21}I & b_{22}I - A & I & 0 & \cdots & 0 \\ \cdot & & \cdot & \ddots & \ddots & \vdots \\ \cdot & & & \cdot & I & 0 \\ \cdot & & & & \cdot & I \\ b_{n1}I & \cdot & \cdot & \cdot & \cdot & b_{nn}I - A \end{pmatrix},$$

and $\phi(x)$ *is the characteristic polynomial of* B.

Proof. From

$$XA - BX = R$$

we have

$$XA = R + BX.$$

Let

$$X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}.$$

Then

$$
\begin{pmatrix} x_1 \\ x_2 \\ \cdot \\ \cdot \\ \cdot \\ x_n \end{pmatrix} A
$$

$$
= \begin{pmatrix}
b_{11} & 1 & 0 & 0 & \cdots & 0 & 0 \\
b_{21} & b_{22} & 1 & 0 & \cdots & 0 & 0 \\
\cdot & & & \cdot & & & \cdot \\
\cdot & & & & \cdot & & \cdot \\
\cdot & & & & & \cdot & \cdot \\
b_{n-1,1} & \cdot & \cdot & \cdot & \cdot & b_{n-1,n-1} & 1 \\
b_{n1} & \cdot & \cdot & \cdot & \cdot & b_{n-1,n} & b_{nn}
\end{pmatrix}
\begin{pmatrix} x_1 \\ x_2 \\ \cdot \\ \cdot \\ \cdot \\ x_{n-1} \\ x_n \end{pmatrix}
+
\begin{pmatrix} r_1 \\ r_2 \\ \cdot \\ \cdot \\ \cdot \\ r_{n-1} \\ r_n \end{pmatrix}.
$$

Comparing the two sides, we have

$$
\begin{aligned}
x_1(b_{11}I - A) + x_2 I &= -r_1, \\
x_1 b_{21} I + x_2(b_{22}I - A) + x_3 I &= -r_2, \\
&\;\vdots \\
x_1 b_{n-1,1} I + \cdots + x_{n-1}(b_{n-1,n-1}I - A) + x_n I &= -r_{n-1}, \\
x_1 b_{n1} I + \cdots + x_n(b_{nn}I - A) &= -r_n.
\end{aligned} \tag{2.1}
$$

From the first $n-1$ equations of (2.1) it is obvious that given x_1, the unknowns x_2 through x_n can be determined uniquely by rows $r_1, r_2, \dots, r_{n-1}$ of R.

Multiplying the 1st, 2nd,$\dots$, nth equations of (2.1) on the right by the cofactors of $d_{11}, d_{21}, \dots, d_{n1}$ respectively and adding, we have

$$
-x_1\phi(A) = r_1 D_{11} + r_2 D_{21}, \dots, + r_{n-1} D_{n-1,1} + (-1)^{n+1} r_n. \quad \blacksquare \tag{2.2}
$$

NOTE. Given r_n, the system (2.2) has a unique solution x_1 iff $\phi(A)$ is nonsingular, that is, iff A and B do not have an eigenvalue in common. The matrix X is also uniquely determined in this case. We thus recover a well-known result on the uniqueness of solution of Sylvester's equation [10, 15].

Nonsingularity of X

We now give a characterization for nonsingularity of solutions X of the Sylvester equation

$$XA - BX = R,$$

where $A, B, R, X \in F^{n,n}$, the set of $n \times n$ matrices over a complex field F. We shall assume further that B is normalized lower Hessenberg. Partial solutions have been obtained previously by Hearon [12] and Bhattacharyya and DeSouza [1]. Carlson and Datta [4] obtained a characterization of nonsingularity of solutions to a special type of quadratic matrix equation: $XA + A^*X = X^*B^*BX$, under the assumption that (A, B^*) is controllable.

Let X and R have rows x_j, r_j respectively. Let $\langle x_1, \dots, x_i \rangle$ be the space spanned by $x_1, \dots, x_i$. Let $Y \in F^{n,n}$ with rows y_j such that

$$\begin{aligned} &y_1 = x_1, \\ &y_{i+1} = x_{i+1} \bmod \langle x_1, \dots, x_i \rangle, \qquad i = 1, 2, \dots, n-1, \end{aligned} \tag{2.3}$$

i.e., $y_{i+1} = x_{i+1} + a_{i1}x_1 + \cdots + a_{ii}x_i$ for some $a_{i1}, a_{i2}, \dots, a_{ii} \in F$.

It follows that $\det Y = \det X$, so that $\det Y \neq 0$ is a necessary and sufficient condition for the nonsingularity of X.

We shall first define vectors

$$Z_1^1, Z_2^1, \dots, Z_{n-1}^1; Z_1^2; \dots; Z_{n-2}^2; \dots; Z_1^{n-1}$$

recursively by

$$Z_j^1 = r_j, \qquad j = 1, \dots, n-1,$$

and then for each $p = 1, \dots, n-1$,

$$Z_j^{p+1} = Z_{j+1}^p + \sum_{k=1}^{j} b_{jk} Z_j^p + r_j A^p, \qquad j = 1, \dots, n-1. \tag{2.4}$$

Let us consider $y_1 = x_1$, and define

$$y_{i+1} = -\sum_{l=1}^{i} r_l A^{i-l} - \sum_{l=1}^{i-1} \sum_{k=1}^{l} b_{lk} Z_k^{i-l} + x_1 A^i, \qquad i = 1,\dots,n-1. \quad (2.5)$$

As an illustration, for $n = 3$, we first define

$$Z_1^1 = r_1, \qquad Z_2^1 = r_2, \qquad Z_1^2 = Z_2^1 + b_{11} Z_1^1 + r_1 A = r_2 + r_1(b_{11} I + A),$$

and then

$$y_1 = x_1, \qquad y_2 = -r_1 + x_1 A,$$

$$y_3 = -r_2 - r_1(b_{11} I + A) + x_1 A^2.$$

From (2.1) we have

$$x_{j+1} = -r_j - \sum_{k=1}^{j} b_{jk} x_k + x_j A, \qquad j = 1,\dots,n-1. \quad (2.6)$$

For all $p = 0, 1, \dots, n-1$, we have

$$x_{j+1} A^p = -r_j A^p - \sum_{k=1}^{j} b_{jk} x_k A^p + x_j A^{p+1}, \qquad j = 1,\dots,n-1. \quad (2.7)$$

We must prove that (2.3) holds. Fix the index (i, i). From now on all calculations are made mod $\langle x_1,\dots,x_i \rangle$. From (2.7) for $p = 0$ we have

$$x_j A = r_j = Z_j^1, \qquad j = 1,\dots,i-1.$$

Also, by induction, for $p = 1,\dots,n-1$ we have

$$x_j A^p = Z_j^p, \qquad j = 1,\dots,i-p. \quad (2.8)$$

Assuming (2.8) and (2.4),

$$x_j A^{p+1} = x_{j+1}A^p + r_j A^p + \sum_{k=1}^{j} b_{jk} x_k A^p = Z_j^{p+1}, \qquad j = 1,\dots, i-p-1.$$

Now from (2.7) for $(j, p) = (i,0), (i-1,1),\dots,(1, i-1)$,

$$\begin{aligned}
x_{i+1} &= -r_i + x_i A,\\
x_i A &= -r_{i-1}A - \sum_{k=1}^{i-1} b_{i-1,k} Z_k^1 + x_{i-1}A^2,\\
x_{i-1}A^2 &= -r_{i-2}A^2 - \sum_{k=1}^{i-2} b_{i-2,k} Z_k^2 + x_{i-2}A^3,\\
&\vdots\\
x_2 A^{i-1} &= -r_1 A^{i-1} - b_{11} Z_1^{i-1} + x_1 A^i.
\end{aligned}$$

Adding and canceling $x_i A, \dots, x_2 A^{i-1}$, we obtain

$$x_{i+1} = -\sum_{l=1}^{i} r_l A^{i-l} - \sum_{l=1}^{i-1} \sum_{k=1}^{l} b_{lk} Z_k^{i-l} + x_1 A^i = y_{i+1}.$$

We have proved the following:

THEOREM 2. *Let B be a normalized lower Hessenberg matrix. If X is a solution of* (1.1), *then X is nonsingular iff*

$$\det\begin{pmatrix} x_1 \\ -r_1 + x_1 A \\ -r_2 - r_1 A - b_{11} Z_1 + x_1 A^2 \\ -r_{n-1} - r_{n-2}A - \cdots - r_1 A^{n-2} - \sum_{l=1}^{n-2} \sum_{k=1}^{l} b_{lk} Z_k^{n-1-l} + x_1 A^{n-1} \end{pmatrix} \neq 0.$$

A special case: If the first $n-1$ rows of R are zero, then a solution X of the equation (1.1) is nonsingular iff

$$\operatorname{rank}\begin{pmatrix} x_1 \\ x_1 A \\ \vdots \\ x_1 A^{n-1} \end{pmatrix} = n,$$

that is, iff (A^T, x_1^T) is controllable.

III. APPLICATIONS

A. *Symmetrizers*

A symmetric matrix X is called a right symmetrizer of B if $BX = XB^T$. Existence of a symmetrizer for an arbitrary matrix B was proved by Taussky and Zassenhaus [17] in 1959.

In our discussion of Section II, if we choose r_1 through r_{n-1} to be zero and then $A = B^T$, then from (2.2) we have $r_n = 0$, because by the Cayley-Hamilton theorem $\phi(A) = 0$. This gives $R = 0$, and the matrix equation (1.1) then becomes $XB^T = BX$. If the first row x_1 of a solution X of this equation is chosen so that (B, x_1^T) is controllable, then according to Theorem 2 X is nonsingular and therefore, B being nonderogatory, by a well-known theorem due to Taussky and Zassenhaus [17, Theorem 2], X is a right symmetrizer of B.

Combining the above arguments with the results of Theorem 1, we can now state:

THEOREM 3. *Given a normalized lower Hessenberg matrix B and an arbitrary n-vector x_1 such that (B, x_1^T) is controllable, there always exists a nonsingular symmetrizer X of B with x_1 as its first row; and X is uniquely determined by such a x_1.*

REMARKS.

(1) Since (B, e_n^T) is controllable, it follows from the above result that a right symmetrizer X of B with e_n as its first row is nonsingular. In fact, taking $x_1 = e_n = (0,0,,\ldots,1)$ and computing x_2 through x_n, using the recur-

sive relation mentioned above, it is easy to see that the symmetrizer X has $x_{i,n+1-i} = 1$ and $x_{ij} = 0$ if $i + j < n + 1$.

(2) A symmetric matrix X is called a left symmetrizer of B_1 if $XB_1 = B_1^T X$. The case of left symmetrizes can similarly be disposed of. However, it is to be noted that a left symmetrizer of B_1 is uniquely determined by its last row x_n chosen so that (B_1^T, x_n^T) is controllable. Such a symmetrizer is, of course, nonsingular.

The Bezout Matrix as a Symmetrizer and Derivation of an Important Property. Let

$$A = \begin{pmatrix} 0 & 1 & 0 & \cdot & \cdot & \cdot & \cdot & 0 \\ 0 & 0 & 1 & 0 & \cdot & \cdot & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & 1 \\ a_1 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & a_n \end{pmatrix}$$

and

$$B = \begin{pmatrix} 0 & 1 & 0 & 0 & \cdot & \cdot & \cdot & 0 \\ 0 & 0 & 1 & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & 1 \\ b_1 & b_2 & \cdot & \cdot & \cdot & \cdot & \cdot & b_n \end{pmatrix}$$

be two companion matrices (in normalized lower Hessenberg forms) of the polynomials

$$f(x) = x^n - \left(a_n x^{n-1} + \cdots + a_1\right),$$

$$g(x) = x^n - \left(b_n x^{n-1} + \cdots + b_1\right).$$

Then it is easy to see that

$$X = \operatorname{diag}\left(1, -1, \ldots, (-1)^{n-1}\right)$$

is a solution of the matrix equation $XA - (-B)X = R$, where the first $n-1$ rows of R are zero and

$$r_n = \left[(-1)^{n-1} a_1 + b_1, (-1)^{n-1} a_2 - b_2, \ldots, (-1)^{n-1}(a_n + b_n)\right].$$

Construct now a left symmetrizer of B with r_n as its last row. This symmetrizer can be easily recognized as the Bezout matrix associated with $f(x)$ and $g(-x)$ [6]. A well-known classical result [13, 14] about Bezout matrices is: the Bezout matrix associated with $f(x)$ and $g(-x)$ is nonsingular iff $f(x)$ and $g(-x)$ are relatively prime. Many diferent proofs of this result are available [6, 13, 14]. Using the results of our Theorem 1, we can still give an alternative proof of this result as follows:

Let $c_1, c_2, \ldots, c_n$ be the n rows of the Bezout matrix. Then by Theorem 1, we have

$$c_n = r_n = (-1)^n x_1 \phi(-B) = (-1)^n e_1 \phi(-B),$$

where

$$e_1 = (1,0,\ldots,0).$$

Then from the construction of the left symmetrizer of B we have

$$\begin{aligned} c_i &= c_{i+1}(B) \\ &= c_n(B)^{n-i} \\ &= (-1)^n e_1 \phi(-B)(B)^{n-i} \\ &= (-1)^n e_1 (B)^{n-i} \phi(-B) \qquad \text{for} \quad i = 1,2,\ldots,n-1. \end{aligned}$$

Thus, the Bezout matrix associated with $f(x)$ and $g(x)$ is

$$\begin{pmatrix} c_1 \\ c_2 \\ \cdot \\ \cdot \\ \cdot \\ c_n \end{pmatrix} = \begin{pmatrix} (-1)^n e_1 B^{n-1} \phi(-B) \\ (-1)^n e_1 B^{n-2} \phi(-B) \\ \cdot \\ \cdot \\ \cdot \\ (-1)^n e_1 \phi(-B) \end{pmatrix}$$

$$= (-1)^n \begin{pmatrix} 0 & 0 & \cdots & 1 \\ 0 & 0 & 1 & \cdot \\ \vdots & & & \vdots \\ 1 & 0 & \cdots & 0 \end{pmatrix} \begin{pmatrix} e_1 \\ e_1 B \\ \vdots \\ e_1 B^{n-1} \end{pmatrix} \phi(-B).$$

Since (B^T, e_1^T) is controllable, it follows that the Bezout matrix is nonsingular if $\phi(-B)$ is nonsingular. But from the nonsingularity of $\phi(-B)$, it follows that $f(x)$ and $g(-x)$ are relatively prime, and conversely.

B. Commuting Matrices

In (1.1) choose $A = B$ and r_1 through r_n as zero vectors. Then $R = 0$; and the solutions X to (1.1) are then matrices that commute with A.

COROLLARY to Theorem 1. *Associated with a normalized lower Hessenberg matrix A and an n-vector x_1, there is a matrix X with x_1 as its first row such that X commutes with A. X is uniquely determined by x_1. X is nonsingular iff (A^T, x_1^T) is controllable.*

As before, our procedure yields a recursive formula for generating a n-parameter family of matrices which commute with A.

Derivation of a Known Result on Commuting Matrices. It is shown in [5] that the vectors x_2 through x_n defined recursively are the second through nth rows of a polynomial matrix $P(A)$ in a normalized Hessenberg matrix A with first row x_1. We therefore immediately have the following well-known result on commuting matrices [10]: A matrix X that commutes with a normalized lower Hessenberg matrix A is a polynomial in A.

C. Computing the Characteristic Polynomial of a Matrix

In this subsection, we show that how the results of Theorem 1 can be used to develop a method for computing the characteristic polynomial of a normalized lower Hessenberg matrix.

THEOREM 3. *Let B be a normalized lower Hessenberg matrix, and let*

$$A = \begin{pmatrix} 0 & 1 & 0 & \cdot & \cdot & \cdot & 0 \\ 0 & 0 & 1 & \cdot & \cdot & \cdot & 0 \\ 0 & 0 & 0 & 1 & \cdot & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & 1 \\ 0 & 0 & \cdot & \cdot & \cdot & \cdot & 0 \end{pmatrix}.$$

Let X be the solution of the equation (1.1), *with $x_1 = e_1 = (1,0,\ldots,0)$. Let $\phi(x) = (-1)^n x^n + b_{n-1}x^{n-1} + \cdots + b_0$ be the characteristic polynomial of B. Let the first $n-1$ rows of R be zero. Then r_n, the last row of R, is*

$$(-1)^n r_n = (b_0, b_1, b_2, \ldots, b_{n-1})$$

Proof. From Theorem 1, we have

$$(-1)^n r_n = e_1 \phi(A).$$

Write $\phi(x) = (-1)^n x^n + P(x)$. Since $A^n = 0$, we have $\phi(A) = P(A)$. Thus

$$\begin{aligned}(-1)^n r_n &= e_1 P(A) = b_0 e_1 + b_1 e_1 A + \cdots + b_{n-1} e_1 A^{n-1}\\ &= b_0 e_1 + b_1 e_2 + \cdots + b_{n-1} e_{n-1}\\ &= (b_0, b_1, \dots, b_{n-1}).\end{aligned}$$

■

REMARKS. Note that the above theorem gives us a constructive procedure for computing the characteristic polynomial of a normalized Hessenberg matrix B.

D. Common Eigenvalue Problem

The problem of finding the number of common eigenvalues between two given matrices A and B and, in particular, of finding if they are relatively prime, is an important problem in mathematics and arises in control theory. It is a classical result in matrix theory [10, 14] that the resultant of the characteristic polynomials of A and B (equivalently, the determinant of $A \otimes I - I \otimes B$) is nonsingular iff A and B do not have an eigenvalue in common. However, this approach is not computationally attractive. Another approach is to solve the matrix equation $XA - BX = R$. It is well known [15] that X is the unique solution of this equation iff A and B are relatively prime. The most effective method available for solving this equation, namely the Hessenberg-Schur method of Golub, Nash, and Van Loan [11] requires however the transformation of one of these two matrices into real Schur form, which is equivalent to finding the spectrum of that matrix. Using our Theorem 1, we can derive direct methods for solving this problem. These methods will not require computation of the characteristic polynomials or the spectrum of any of the given matrices.

Let the first $n-1$ rows of R be zero. Then from (2.2) we have $r_n = (-1)^n x_1 \phi(A)$. Thus choosing x_1 successively as the 1st through the nth row of the identity matrix, we can construct $\phi(A)$ [without computing $\phi(x)$ explicitly]. Once $\phi(A)$ is known, the relative primeness of A and B (in fact the number of common eigenvalues) can be determined by finding the rank and nullity of $\phi(A)$. However, this approach will yield an $O(n^4)$ method, since we have to repeat the recursion n times. Certainly an $O(n^4)$ method is unpractical.

A more practical approach will be as follows:

1. Transform A also into lower Hessenberg Form. (Assume, for the sake of simplicity, that it has unit superdiagonal. If not, the problem can be decomposed into several smaller problems of lower dimensions.)
2. With $x_1 = e_1 = (1,0,0,\dots,0)$, construct a solution X of the equation (1.1) using the recursion (2.1).
3. Compute r_n.
4. With r_n as p_1, construct p_2 through p_n recursively using:

$$p_{i+1} = p_i(A - a_{ii}I) - \sum_{j=1}^{i-1} a_{ij}p_j, \qquad i = 1,\dots,n-1.$$

Let

$$P = \begin{pmatrix} p_1 \\ p_2 \\ \cdot \\ \cdot \\ \cdot \\ p_n \end{pmatrix}.$$

THEOREM [5]. *The number of common eigenvalues between A and B is equal to the nullity of P; in particular, A and B are relatively prime iff P is nonsingular.*

Proof. We have $p_1 = r_n = (-1)^n e_1 \phi(A)$; as noted earlier, the vectors p_2 through p_n define the rows of a polynomial matrix in A; and a polynomial matrix in normalized lower Hessenberg matrix is uniquely determined by its first row [5]. It follows that $P = (-1)^n \phi(A)$. ∎

E. Inertia Method

The inertia of an $n \times n$ square matrix A is defined as $\mathrm{In}(A) = (\pi(A), \nu(A), \delta(A))$, where $\pi(A), \nu(A), \delta(A)$ are respectively the numbers of eigenvalues of A with positive, negative, and zero real parts.

In [3], Carlson and Datta gave a direct method for computing the inertia of a nonhermitian matrix in terms of the inertia of a hermitian matrix. Here we give an alternative derivation of the Carlson-Datta method.

THEOREM [3]. *Let A be a normalized lower Hessenberg matrix. Let L be a nonsingular lower triangular matrix with $l_1 = (1,0,\dots,0)$ as its first row* *such that the first n − 1 rows of the matrix* $R = LA + \bar{A}L$ *are zero. Let S be a left symmetrizer of A with* r_n, *the last row of R, as its last row. Define* $H = L^*S$. *If H is nonsingular, then H is hermitian and* $\mathrm{In}(A) = \mathrm{In}(H)$.

Proof.

$$\begin{aligned} HA + A^*H &= L^*SA + A^*L^*S \\ &= L^*A^TS + A^*L^*S \\ &= (LA + \bar{A}L)^*S \\ &= r_n^*r_n \geqslant 0. \end{aligned} \tag{2.9}$$

Since

$$H = L^*S$$

and L is nonsingular by construction, the nonsingularity of H implies that S is nonsingular. From the special case of Theorem 2, it therefore follows that (A^T, r_n^T) is controllable, where

$$r_n = (-1)^{n+1} l_1 \phi(A).$$

Thus,

$$K = \begin{pmatrix} r_n A^{n-1} \\ \vdots \\ r_n A \\ r_n \end{pmatrix} = \begin{pmatrix} (-1)^{n+1} l_1 & \phi(A) & A^{n-1} \\ (-1)^{n+1} l_1 & \phi(A) & A^{n-2} \\ \vdots & \vdots & \vdots \\ (-1)^{n-1} l_1 & \phi(A) & A \end{pmatrix}$$

$$= \left(\begin{pmatrix} (-1)^{n-1} l_1 & A^{n-1} \\ \vdots & \vdots \\ (-1)^{n+1} l_1 & A \end{pmatrix} \phi(A) \right).$$

The controllability of (A^T, r_n^T) implies that K is nonsingular. However, K is nonsingular iff $\phi(A)$ is nonsingular, and the nonsingularity of $\phi(A)$ implies A and $\bar{A}$ do not have any common eigenvalues. This means that $\delta(A) = 0$. Now, applying the well-known inertia theorem of Carlson and Schneider [2] to the equation (2.9), we conclude that $\mathrm{In}(A) = \mathrm{In}(H)$. ∎

I would like to express my sincere thanks and gratitude to Professor David Carlson for making some constructive suggestions on this paper. In particular, he suggested the proof of Theorem 2.

REFERENCES

1 S. P. Bhattacharyya and E. DeSouze, Controllability, observability, and the solution of $AX - XB = C$. *Linear Algebra Appl.* 39:167–188 (1981).
2 D. Carlson and H. Schneider, Inertia theorem for matrices; the semidefinite case, *J. Math. Anal. Appl* 6:430–446 (1963).
3 D. Carlson and B. N. Datta, On the effective computation of the inertia of a nonhermitian matrix, *Numer. Math.* 33:315–322 (1979).
4 D. Carlson and B. N. Datta, The matrix equation $SA + A^*S^* = S^*B^*BS$, *Linear Algebra Appl.* 28:43–52 (1979).
5 B. N. Datta and Karabi Datta, An algorithm for computing the powers of a Hessenberg matrix and its applications, *Linear Algebra Appl.* 14:273–284 (1976).
6 B. N. Datta, Controllability, Bezoutian and relative primeness, *Internal. J. Math. Math. Sci.* 3:185–188 (1980).
7 B. N. Datta, An algorithm for assigning eigenvalues in a Hessenberg matrix: Single input case, *IEEE Trans. Automat. Control* AC-08, Apr. 1986.
8 B. N. Datta and Karabi Datta, On eigenvalue and canonical form assignments, submitted for publication.
9 K. Datta, Produtos de Kronecker, simetrizadoras e algorithmos paralelos e sequenciais na algebra linear, Ph.D. Thesis, State Univ. of Campinas, Campinas, S. P. Brazil, 1982.
10 F. R. Gantmacher, *The Theory of Matrices*, Vol. I, Chelsea, New York, 1959.
11 G. H. Golub, S. Nash, and C. Van Loan, A Hessenberg-Schur method for the problem $AX - XB = C$, *IEEE Trans. Automat. Control* AC-24:909–912 (1979).
12 J. Z. Hearon, Nonsingular solutions of $TA - BT = C$, *Linear Algebra Appl.* 16:57–63 (1977).
13 A. S. Householder, Bezoutians, elimination and location, *SIAM Rev.* 12:73–78 (1970).
14 M. G. Krein and M. A. Naimark, The method of symmetric and hermitian forms in the theory of the separation of the roots of algebraic equations (translated from Russian by O. Boshko and J. L. Howland), *Linear and Multilinear Algebra* 10:265–308 (1981).
15 P. Lancaster, Explicit solutions of linear matrix equations. *SIAM Rev.* 12(4):544–566 (1970).
16 V. Pták, Lyapunov, Bezout, and Hankel, *Linear Algebra Appl.* 58:363–390 (1984).
17 O. Taussky and A. Zassenhaus, On the similarity transformation between a matrix and its transpose, *Pacific J. Math.* 17:893–896 (1959).

Received October 1987; final manuscript accepted 3 November 1987