Generalized Hankel Matrices of Markov Parameters and Their Applications to Control Problems

B. N. Datta
Mathematical Sciences Department
Northern Illinois University
DeKalb, Illinois 60115

Submitted by Richard A. Brualdi

ABSTRACT

The concept of Hankel matrices of Markov parameters associated with two polynomials is generalized for matrices. The generalized Hankel matrices of Markov parameters are then used to develop methods for testing the relative primeness of two matrices A and B, for determining stability and inertia of a matrix, and for constructing a class of matrices C such that $A+C$ has a desired spectrum. Neither the method of construction of the generalized Hankel matrices nor the methods developed using these matrices require explicit computation of the characteristic polynomial of A (or of B).

I. INTRODUCTION

Given two polynomials $f(x)$ and $g(x)$, the degree of $g(x)$ being less than or equal to that of $f(x)$, the quantities s_i, $i=-1,0,1,2,\ldots,$ defined by

$$R(x)=\frac{g(x)}{f(x)}=S-1+\frac{s_0}{x}+\frac{s_1}{x^2}+\cdots$$

are called Markov parameters associated with $R(x)$, and the symmetric matrices $H_{fg}=(s_{i+j})$ are known as Hankel matrices of Markov parameters. The use of Hankel matrices of Markov parameters in computing the Cauchy index $\mathrm{I}_{-\infty}^{+\infty}R(x)$ and determining the criterion of stability are well known [1,2,8]. In a recent paper [7], the author has shown how these matrices can be employed to obtain information on the location of zeros of a polynomial

52 Vanderbilt Ave., New York, NY 10017 0024-3795/84/$3.00

inside a half plane and the unit circle and has given a criterion for aperiodicity of a polynomial using these matrices.

Since these problems are basically the eigenvalue or eigenvalue related problems for companion matrices associated with $f(x)$ and $g(x)$, and the reduction of an arbitrary matrix to its companion form is numerically an unstable process [13], it is natural to investigate how, given two matrices A and B, Hankel matrices of Markov parameters can be constructed without actually computing the characteristic polynomials of A and B.

In this paper, given an arbitrary matrix A of order m and a Hessenberg matrix B of order n $(n \leqslant m)$, a family of symmetric matrices H_{AB} is constructed, which can be considered as generalized Hankel matrices of Markov parameters, in the sense that included as a special case in the family is the Hankel matrix of Markov parameters associated with $f(x) = \det(xI - A)$ and a suitably chosen polynomial $g(x)$ constructed from the characteristic polynomial of B.

The matrix H_{AB} is then used to develop methods for testing the relative primeness of A and B, for determining the stability (in fact the inertia) of A (or B), and for constructing matrices C such that $A + C$ has a desired spectrum Ω. A recent theorem of the author (Theorem 4 in [7]) is derived as a special case of the inertia method. The paper also contains a simple algorithm for constructing solutions of the matrix equation $AX = XA^T$, and a result on controllability which might be of independent interests.

Neither our method of construction of the matrices H_{AB} nor the methods developed using H_{AB} require the explicit computation of the characteristic polynomial of A (or of B).

Since our interest in Hankel matrices is mainly in their applications to the eigenvalue related problems, the assumption that one of the matrices, (namely, the matrix B) is a Hessenberg matrix is not unrealistic at all, because an arbitrary matrix can be transformed to a Hessenberg matrix by similarity, and there exist efficient and numerically stable algorithms (e.g., Householder's method [13], Givens's method [9]) for doing this. Furthermore, one can assume that the transformed Hessenberg matrix has nonzero codiagonal. Indeed, the appearance of a zero element on the codiagonal reduces the problem of the original matrix to problems of lower order, each involving a Hessenberg matrix with nonzero codiagonal (for more precise statements, see Section V on applications). A Hessenberg matrix with nonzero codiagonal is called an unreduced Hessenberg matrix. An unreduced Hessenberg matrix can further be reduced to one having 1's along the codiagonal by a diagonal similarity. Such a matrix is called a normalized Hessenberg matrix. A method for reducing an unreduced lower Hessenberg matrix A to a normalized one without explicitly computing the transforming diagonal matrix appears in [6].

II. SOME LEMMAS

In this section, we establish a few lemmas which will be used later. In the following e_i stands for the ith column of an identity matrix.

A pair of matrices (A, B), where A is of order n and B is of order $n \times m$, is controllable if the $n \times nm$ matrix $C(A, B) = (B, AB, A^2B, \ldots, A^{n-1}B)$ has rank n.

LEMMA 1. *Let A be a $n \times n$ matrix and b be a column vector such that (A^T, b) is controllable. If X satisfies the equation*

$$AX = XA^T, \tag{2.1}$$

then X is nonsingular iff (A, Xb) is controllable.

Proof. Assume first that X is nonsingular. Then

$$\begin{aligned}
&\operatorname{rank}(Xb, AXb, A^2Xb, \ldots, A^{n-1}Xb) \\
&\quad = \operatorname{rank}\big(Xb, XA^Tb, X(A^T)^2b, \ldots, X(A^T)^{n-1}b\big) \\
&\quad = \operatorname{rank}\big[X\big(b, A^Tb, (A^T)^2b, \ldots, (A^T)^{n-1}b\big)\big].
\end{aligned}$$

Since X is nonsingular and (A^T, b) is controllable, it follows that (A, Xb) is controllable.

Next, let (A, Xb) be controllable. Then

$$\begin{aligned}
n &= \operatorname{rank}(Xb, AXb, A^2Xb, \ldots, A^{n-1}Xb) \\
&= \operatorname{rank}\big(Xb, XA^Tb, X(A^T)^2b, \ldots, X(A^T)^{n-1}b\big) \\
&= \operatorname{rank}\big[X\big(b, A^Tb, (A^T)^2b, \ldots, (A^T)^{n-1}b\big)\big].
\end{aligned}$$

Since (A^T, b) is controllable, it follows that X is nonsingular. ∎

LEMMA 2. *Let $A = (a_{ij})$ be a normalized lower Hessenberg matrix, i.e., $a_{i\,i+1} = 1$ for all $i = 1, 2, \ldots, n$, and $a_{ij} = 0$ whenever $j > i + 1$. If $x_1, x_2, \ldots, x_n$*

are the n successive columns of a matrix X satisfying (2.1), *then*

(i) x_1 *can be chosen arbitrarily,*
(ii) x_2 *through* x_n *satisfy the recursive relations*

$$x_{i+1} = Ax_i - \sum_{j=1}^{i} a_{ij}x_j, \qquad i = 1,2,\dots,n-1. \tag{2.2}$$

Proof. The equation (2.1) is equivalent to the systems of equations

$$Ax_i = x_{i+1} + \sum_{j=1}^{i} a_{ij}x_j, \qquad i = 1,2,\dots,n-1, \tag{2.3}$$

and

$$Ax_n = a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n. \tag{2.4}$$

The recursive relations (2.2) and (2.3) are the same. Also, substituting x_2 through x_n successively from (2.3) in (2.4), it is easy to see that

$$\Psi(A)x_1 = 0,$$

where $\Psi(x)$ is the characteristic polynomial of A. In fact, it can be shown [4] that

$$x_{i+1} = \Psi_i(A)x_i,$$

where $\Psi_i(x)$ is the characteristic polynomial of the submatrix of A consisting of the first i rows and i columns and $\Psi_n(x) = \Psi(x)$. Since by the Cayley-Hamilton theorem $\Psi(A) = 0$, x_1 can be chosen arbitrarily. ∎

REMARK. Since (A^T, e_1) is controllable, by Lemma 1 we conclude that if A is a normalized lower Hessenberg matrix, then a solution of (2.1) is nonsingular iff $(A, Xe_1 = x_1)$ is controllable.

LEMMA 3. *Let $M=(m_{ij})$ be a normalized lower Hessenberg matrix, and $\phi_i(x)$ be the characteristic polynomial of the submatrix of M consisting of the last i rows and i columns. Then*

$$\phi_{k+1}(\lambda)=(\lambda-m_{n-k,n-k})\phi_k(\lambda)-m_{n-k+1,n-k}\phi_{k-1}(\lambda)$$
$$-m_{n-k+2,n-k}\phi_{k-2}(\lambda)-\cdots-m_{n,n-k}\phi_0(\lambda),$$

where $\phi_0(\lambda)=1$.

Proof. Similar to the one given in Wilkinson [13, p. 411]. ∎

III. AN INERTIA THEOREM

The inertia of a matrix A is defined to be an integer triple $\mathrm{In}(A)=(\pi(A),\nu(A),\delta(A))$, where $\pi(A)$, $\nu(A)$, and $\delta(A)$ are respectively the numbers of eigenvalues of A with positive, negative, and zero real parts. A matrix A of order n is called a stable matrix iff $\mathrm{In}(A)=(0,n,0)$. A direct method for computing the inertia of a normalized Hessenberg matrix appears in [3]. A similar method using a generalized Hankel matrix of Markov parameters is presented in this paper. The proposed method is as efficient as the method in [3]. The following inertia theorem will be needed later.

THEOREM 1 (Carlson and Schneider [5]). *Let A be an $n\times n$ complex matrix with $\delta(A)=0$, and let X be a nonsingular hermitian matrix such that $XA+A^*X$ is positive semidefinite. Then* $\mathrm{In}(A)=\mathrm{In}(X)$.

IV. CONSTRUCTIONS OF GENERALIZED HANKEL MATRICES OF MARKOV PARAMETERS

Let A be a matrix of order m, and $B=(b_{ij})$ be a normalized lower Hessenberg matrix of order n $(n\leqslant m)$.

Step 1. Choose column vectors b and c such that (A,b) and (A^T,c) are controllable.

Step 2. Construct a matrix V having the columns $v_1, v_2, \dots, v_n$ defined by

$$
\begin{aligned}
v_n &= b,\\
v_{n-1} &= -(A + b_{nn}I)v_n,\\
v_{n-2} &= -(A + b_{n-1,n-1}I)v_{n-1} - b_{n,n-1}v_n,\\
&\vdots\\
v_1 &= -(A + b_{22}I)v_2 - b_{32}v_3 - \cdots - b_{n2}v_n.
\end{aligned}
$$

Step 3. Compute

$$r = Av_1 + b_{11}v_1 + b_{21}v_2 + \cdots + b_{n1}v_n. \tag{4.0}$$

Step 4. Construct H_{AB} having the following properties:

$$AH_{AB} = H_{AB}A^T, \tag{4.1}$$

$$H_{AB}c = r. \tag{4.2}$$

Different choices of the vectors b and c will yield different matrices H_{AB}. These matrices will be called generalized Hankel matrices of Markov parameters (see the discussion for a special case). Since A is nonderogatory [the existence of the vector b such that (A, b) is controllable implies that A is nonderogatory], by a result of Taussky and Zassenhaus [12] the matrices H_{AB} are symmetric. In particular, if A is a normalized or unreduced lower Hessenberg matrix, then one member of the family H_{AB} can be obtained by choosing the first column as r_1 [since (A^T, e_1) is controllable and $H_{AB}e_1 = r$] and generating the remaining columns using the recursive relations (2.2) of Lemma 2.

A SPECIAL CASE. *Let*

$$A = \begin{pmatrix} 0 & 1 & 0 & 0 & \cdots & 0\\ 0 & 0 & 1 & 0 & \cdots & 0\\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots\\ 0 & 0 & 0 & 0 & \cdots & 1\\ a_1 & a_2 & a_3 & a_4 & \cdots & a_n \end{pmatrix} \tag{4.3}$$

and

$$B=\begin{pmatrix} 0 & 1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 1 & 0 & \cdots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & 0 & \cdots & 0 & 1 \\ b_1 & -b_2 & b_3 & -b_4 & \cdots & (-1)^{n-2}b_{n-1} & (-1)^{n-1}b_n \end{pmatrix}$$

be the companion matrices of two polynomials $f(x)$ and $g(x)$ respectively. Choose

$$b=\begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ (-1)^{n-1}S-1 \end{pmatrix} \quad \textit{and} \quad c=e_1=\begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$$

Then H_{AB} is the Hankel matrix of Markov parameters associated with

$$f(x)=\det(xI-A) \quad \textit{and} \quad g(x)=(-1)^n\det(xI+B).$$

Proof. First, note that (A,b) and (A^T,c) are controllable. In this special case, it is an easy computation to see that

$$v_{n-1}=-\left[A+(-1)^{n-1}b_nI\right]b$$

$$=(-1)^n\begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ s_{-1} \\ s_0 \end{pmatrix},$$

$$v_{n-2}=-Av_{n-1}+(-1)^{n-2}b_{n-1}v_n$$

$$=(-1)^{n-1}\begin{pmatrix} 0 \\ 0 \\ \vdots \\ s_{-1} \\ s_0 \\ s_1 \end{pmatrix}.$$

In general,

$$v_{n-i} = (-1)^{n-i+1} \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ s_{-1} \\ s_0 \\ \vdots \\ s_{i-1} \end{pmatrix}.$$

Thus

$$v_1 = \begin{pmatrix} s_{-1} \\ s_0 \\ s_1 \\ \vdots \\ s_{n-2} \end{pmatrix},$$

and

$$r = b_1 v_n + A v_1 = b_1 \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ (-1)^{n-1} s_{-1} \end{pmatrix} + \begin{pmatrix} s_0 \\ s_1 \\ s_2 \\ \vdots \\ a_1 s_{-1} + a_2 s_0 + \cdots + a_n s_{n-2} \end{pmatrix}$$

$$= \begin{pmatrix} s_0 \\ s_1 \\ s_2 \\ \vdots \\ s_{n-1} \end{pmatrix} = H_{fg} e_1$$

(using the recursive relations between the coefficients of H_{fg}; see [8, p. 214]).

Since a Hankel matrix of Markov parameters H_{fg} satisfies $H_{fg} A = A^T H_{fg}$ [7], where A is the companion matrix of $f(x)$ defined by (4.3), and from Lemma 2 it follows that a solution of (2.1) is uniquely determined by its first column, the matrix H_{AB} in this special case must be the Hankel matrix H_{fg}. ■

V. APPLICATIONS

A. *Determination of Relative Primeness of Two Matrices*

Let A, B, H_{AB}, and r be the same as in Section IV.

THEOREM 2. *The following are equivalent:*

(i) (A, r) *is controllable;*
(ii) H_{AB} *is nonsingular; and*
(iii) A *and* $-B$ *are relatively prime (i.e.,* A *and* $-B$ *do not have an eigenvalue in common).*

Proof. The equivalence of (i) and (ii) follows from Lemma 1. To prove the equivalence of (i) and (iii), we proceed as follows: We first prove that

$$v_{n-i} = \phi_i(-A)v_n,$$

where $\phi_i(x)$ is the characteristic polynomial of the submatrix of B consisting of the last i rows and i columns. The proof is by induction.

The relation is obvious for $i = 1$ and $i = 2$. For $i = 1$,

$$\begin{aligned} v_{n-1} &= -(A + b_{nn}I)v_n \\ &= \phi_1(-A)v_n. \end{aligned}$$

For $i = 2$,

$$\begin{aligned} v_{n-2} &= (A + b_{n-1,n-1}I)(A + b_{nn}I)v_n - b_{n,n-1}v_n \\ &= \phi_2(-A)v_n. \end{aligned}$$

Assume that the relation is true for all values of $i = 1, 2, \dots, k$; then

$$\begin{aligned} v_{n-k-1} &= -(A + b_{n-k,n-k}I)v_{n-k} - b_{n-k+1,n-k}v_{n-k+1} - \cdots - b_{n,n-k}v_n \\ &= -(A + b_{n-k,n-k}I)\phi_k(-A)v_n - b_{n-k+1,n-k}\phi_{k-1}(-A)v_n \\ &\quad - \cdots - b_{n,n-k}v_n \\ &= \phi_{k+1}(-A)v_n \qquad \text{(by Lemma 3).} \end{aligned}$$

The induction is now complete.

Thus

$$\begin{aligned} r &= Av_1 + b_{11}v_1 + b_{21}v_2 + \cdots + b_{n1}v_n \\ &= b_{n1}v_n + b_{n-1,1}\phi_1(-A)v_n + b_{n-2,1}\phi_2(-A)v_n \\ &\quad + \cdots + (A + b_{11}I)\phi_{n-1}(-A)v_n \\ &= \phi_n(-A)v_n \qquad \text{(by Lemma 3)}. \end{aligned}$$

So

$$\begin{aligned} &\operatorname{rank}(r, Ar, A^2r, \ldots, A^{n-1}r) \\ &\quad = \operatorname{rank}(\phi_n(-A)v_n, A\phi_n(-A)v_n, A^2\phi_n(-A)v_n, \ldots, A^{n-1}\phi_n(-A)v_n) \\ &\quad = \operatorname{rank}[\phi_n(-A)(v_n, Av_n, A^2v_n, \ldots, A^{n-1}v_n)] \end{aligned}$$

[observe that $\phi_n(-A)$, being a polynomial in A, commutes with A and its various powers].

Since $(A, v_n = b)$ is controllable, it follows that (A, r) is controllable iff $\phi_n(-A)$ is nonsingular. The nonsingularity of $\phi_n(-A)$ again implies that A and $-B$ do not have an eigenvalue in common, because the eigenvalues of $\phi_n(-A)$ are $(-1)^n\prod_{j=1}^{n}(\lambda_i + \mu_j)$, $i = 1, \ldots, n$, where $\lambda_1, \lambda_2, \ldots, \lambda_n$ are the eigenvalues of A, and $\mu_1, \mu_2, \ldots, \mu_n$ are the eigenvalues of B. ■

Theorem 2 suggests two tests of relative primeness of A and $-B$:

Test 1. Compute the vector r given by (4.0) and then test the controllability of the pair (A, r).

Test 2. Form the symmetric matrix H_{AB} given by (4.1) and (4.2) and test for its nonsingularity.

Since our algorithm for constructing solutions of (2.1) is restricted to Hessenberg matrices, implementation of Test 2 requires that both matrices A and B be decomposed into lower Hessenberg forms, whereas to implement test 1, only the matrix B needs to be so decomposed. Furthermore, construction of H_{AB} requires computation of r anyway. Thus Test 1 is more direct and efficient than Test 2.

In Test 2, if either A or B or both have zeros on their superdiagonals, they can always be partitioned in the form

$$A=\begin{pmatrix} A_{11} & 0 & 0 & \cdots & 0 \\ A_{21} & A_{22} & 0 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ A_{p1} & A_{p2} & A_{p3} & \cdots & A_{pp} \end{pmatrix}, \tag{5.1}$$

$$B=\begin{pmatrix} B_{11} & 0 & \cdots & 0 \\ B_{21} & B_{22} & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ B_{q1} & B_{q2} & \cdots & B_{qq} \end{pmatrix}, \tag{5.2}$$

where A_{ii} $(i=1,\dots,p)$ and B_{jj} $(j=1,\dots,q)$ are unreduced lower Hessenberg matrices. Each matrix A_{ii} and B_{jj} are then further reduced to normalized forms, and the test is applied successively to the pairs $(A_{11},B_{11}),\dots,(A_{11},B_{qq});(A_{22},B_{11}),\dots,(A_{22},B_{qq});\dots;(A_{pp},B_{11}),(A_{pp},B_{22}),\dots,(\mathrm{A}_{\mathrm{pp}},\mathrm{B}_{\mathrm{qq}})$.

In Test 1, if B has one or more zeros on its superdiagonal, it is then partitioned in the form (5.1) and the test is applied to (A,B_{11}), $(A,B_{22}),\dots(A,B_{qq})$.

B. *Computation of the Inertia and Stability of a Matrix*

Let A be a normalized lower Hessenberg matrix.

Step 1. Construct H_{AA} choosing $c=e_1$.
Step 2. Form

$$S=VH_{AA}.$$

Then

THEOREM 3. *Whenever S is nonsingular, it is symmetric. Furthermore,* $\delta(A)=0$ *and* $\mathrm{In}(A)=\mathrm{In}(S)$. *In particular, A is stable iff S is negative definite.*

Proof.

$$\begin{aligned} AS+SA^T &= AVH_{AA}+VH_{AA}A^T \\ &= AVH_{AA}+VAH_{AA} \\ &= (AV+VA)H_{AA}. \end{aligned} \tag{5.3}$$

It readily follows from the construction procedure that $AV+VA$ is a matrix whose first column is r and all other columns are zero. So, from (5.3), we get

$$AS+SA^{T}=re_1^{T}H_{AA}=H_{AA}e_1e_1^{T}H_{AA}=N. \tag{5.4}$$

The matrix N is clearly symmetric and positive semidefinite. The nonsingularity of S implies that H_{AA} is nonsingular, and by Theorem 2, we then have that A and $-A$ do not have an eigenvalue in common. This implies that

(1) S is an unique solution of (5.4) and therefore symmetric; and
(2) $\delta(A)=0$.

Application of the inertia theorem of Carlson and Schneider to (5.4) yields the desired result. ■

REMARKS.

(i) For different choices of the vector b, the above procedure will yield different H_{AA}, thus yielding a family of symmetric matrices S_{AA}, the inertia of each of which will determine the inertia of A.

(ii) After decomposition into lower Hessenberg form, if there are one or more zero elements on the superdiagonal, then A should be partitioned in the form (5.1), each A_{ii} should be reduced further to normalized form, and the above procedure should be applied successively to each A_{ii}, $i=1,\dots,p$.

A *special case* is the derivation of the Routh-Hurwitz-Markov theorem (Theorem 4 in [7]). In case A is the companion matrix defined by (4.3) associated with a polynomial $f(x)=x^n-a_nx^{n-1}-\cdots-a_2x-a_1$, and $b=v_n$ is chosen to be

$$b=\begin{pmatrix}0\\0\\\vdots\\0\\(-1)^{n-1}s_{-1}\end{pmatrix},$$

where

$$\frac{f(-x)}{f(x)}=s_{-1}+\frac{s_0}{x}+\frac{s_1}{x^2}+\cdots,$$

then H is the Hankel matrix of Markov parameters associated with $R(x) = f(-x)/f(x)$, and the matrix S can be easily recognized as the matrix H'_{nn} defined by (13) in [7]. Therefore, Theorem 4 of [7] follows as a special case of our theorem and we have a very short and simple proof of this theorem.

CONJECTURE. Since the choice of $B = A$ in the construction procedure of H_{AB} has produced an algorithm for computing the inertia of A, it is reasonable to conjecture that other suitable choices of B (mainly as functions of A) might lead to procedures for eigenvalue separation of A in other regions of the complex plane. For example, in case A is invertible, the choice of $B = A^{-1}$ should give a constructive procedure for solving the unit circle problem (the problem of counting the numbers of eigenvalues inside and outside the unit circle). The recent works of Gutman and Jury [10] and of Jury and Ahn [11] are relevant to a solution of this conjecture and might very well be the hunting grounds for a way to choose B for a specified region.

C. *Constructions of Matrices C Such That A + C Has a Desired Spectrum*

Given a matrix A of order n and a set of n numbers Ω, here we consider the problem of finding a matrix C such that $A + C$ has the spectrum Ω. This problem is very closely related to a well-known important problem known as the pole assignment problem in mathematical control theory. We present a solution of the problem using the inverse of a generalized Hankel matrix, whenever the inverse exists.

Let A be a normalized lower Hessenberg matrix.

Step 1. Choose a matrix B, any normalized lower Hessenberg matrix such that the spectrum of $-B$ is the set Ω; in particular, B can be chosen as an upper triangular matrix with its diagonal as the set $-\Omega$, 1's along the superdiagonal, and the rest of the entries zeros.

Step 2. Form H_{AB}, and assume that it is nonsingular.

Step 3. Compute $S = H_{AB}^{-1}V$.

Step 4. Form $F = RS^{-1}$, where $R = AV + VB$ and $C = H_{AB}^{-1}F$.

Then

THEOREM 4. *$A^T - C$ has the spectrum Ω.*

Proof. First of all, we prove that S is nonsingular. Since $S = H_{AB}^{-1}V$, all we have to show then is that V is nonsingular. But the controllability of (A, b)

implies this. For, the equations in Step 2 of Section IV can be written as

$$\begin{aligned}
v_{n-1} &= -Av_n - b_{nn}v_n = -Av_n + t_{12}v_n,\\
v_{n-2} &= (A + b_{n-1,n-1}I)(A + b_{nn}I)v_n - b_{n,n-1}v_n\\
&= A^2v_n + (b_{n-1,n-1} + b_{nn})Av_n + (b_{n-1,n-1}b_{nn} - b_{n,n-1})v_n\\
&= A^2v_n + t_{23}Av_n + t_{13}v_n,\\
&\ \vdots\\
v_1 &= (-1)^{n-1}A^{n-1}v_n + \cdots + t_{2n}Av_n + t_{1n}v_n.
\end{aligned}$$

Thus, we have

$$(v_n, v_{n-1}, \dots, v_2, v_1) = (v_n, Av_n, A^2v_n, \dots, A^{n-1}v_n)$$

$$\times \begin{pmatrix} 1 & t_{12} & t_{13} & \cdots & t_{1n} \\ 0 & -1 & t_{23} & \cdots & t_{2n} \\ 0 & 0 & 1 & \cdots & t_{3n} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & 0 & (-1)^{n-1} \end{pmatrix}$$

Since $(A, b = v_n)$ is controllable, it follows that V is nonsingular.

Next,

$$\begin{aligned}
A^TS + SB &= A^TH_{AB}^{-1}V + H_{AB}^{-1}VB\\
&= H_{AB}^{-1}AV + H_{AB}^{-1}VB\\
&\quad \left(\text{since } AH_{AB} = H_{AB}A^T, \text{ and } H_{AB} \text{ is nonsingular}\right)\\
&= H_{AB}^{-1}(AV + VB) = H_{AB}^{-1}R.
\end{aligned}$$

Then,

$$\begin{aligned}
A^T - C &= A^T - H_{AB}^{-1}F = A^T - H_{AB}^{-1}RS^{-1}\\
&= A^T - (A^TS + SB)S^{-1}\\
&= -SBS^{-1}.
\end{aligned}$$

The theorem is now proved. ■

REMARK. If H_{AB} turns out to be singular, by Theorem 2 we immediately conclude that there is an eigenvalue of A in Ω. In fact, it is easy to see that if the nullity of H_{AB} is p, then Ω will contain exactly p eigenvalues of A. Therefore, in practice, the above procedure for constructing a matrix C such that $A^T - C$ has the desired spectrum can always be made to work by choosing a suitable shift.

APPENDIX. SYSTEM THEORETIC INTERPRETATIONS OF SOME RESULTS

Let $p(x)$ and $q(x)$ be two relatively prime polynomials with real coefficients, the degree of $q(x)$ being greater than that of $p(x)$. Then the triple (A, b, c) is a minimal realization of $p(x)/q(x)$ if

$$R(x) = \frac{p(x)}{q(x)} = c^T(xI - A)^{-1}b.$$

A well-known result in linear systems theory on minimal realization is:

A triple (A, b, c) is a minimal realization of $R(x)$ iff (A, b) is controllable and (c^T, A) is observable (that is, (A^T, c) is controllable).

Theorem 2 in this paper asserts that the generalized Hankel matrix H_{AB} is nonsingular iff (A, r), where r is defined by (4.0), is controllable. Also, H_{AB} is constructed under the assumption that (c^T, A) is observable. Thus, it follows immediately from the above result that:

THEOREM 5. *(A, r, c) is a minimal realization of $c^T(xI - A)^{-1}r$ iff H_{AB} is nonsingular.*

Next, we assume that (A, r, c) is a minimal realization of $c^T(xI - A)^{-1}r$. Then, since (4.1) and (4.2) hold, it follows from a result of B. D. O. Anderson [1, Theorem 3] that:

THEOREM 6. *The signature of H_{AB} is the Cauchy index of $c^T(xI - A)^{-1}r$.*

REMARK. Theorem 6 was pointed out to the author by an anonymous referee, who also remarked that this result could possibly be used to clarify some of other conclusions of the paper. That is true. However, a better system theoretic interpretation of the vector r would be needed to do this.

REFERENCES

1 B. D. O. Anderson, Applications of the second method of Lyapunov to the proof of the Markov stability criterion, *Internat. J. Control* 5:473–482 (1967).

2 B. D. O. Anderson, On the computation of the Cauchy index, *Quart. Appl. Math.*, 1972, pp. 577–582.

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, On the solution of $AX = XB$, in preparation.

5 D. Carlson and H. Schneider, Inertia theorems for matrices: The semidefinite case, *J. Math. Anal. Appl.* 6:430–446 (1963).

6 B. N. Datta and Karabi Datta, The matrix equation $XA = A^TX$ and an associated algorithm for the inertia problem, submitted for publication.

7 B. N. Datta, Applications of Hankel matrices of Markov parameters to the solutions of the Routh-Hurwitz and the Schur-Cohn problems, *J. Math. Anal. Appl.* 68:276–290 (1979).

8 F. R. Gantmacher, *The Theory of Matrices*, Vol. II, Chelsea, New York, 1959.

9 W. Givens, Numerical computation of the characteristic values of a real symmetric matrix, Oak Ridge Natl. Lab., 1954.

10 S. Gutman and E. I. Jury, A general theory for matrix root-clustering in subregions of the complex plane, *IEEE Trans. Automat. Control* AC-26:853–863 (1981).

11 E. I. Jury and S. M. Ahn, Remarks on the root-clustering of a polynomial in a certain region in the complex plane, *Quart. Appl. Math.*, 1974, pp. 203–205.

12 O. Taussky and A. Zassenhaus, On the similarity transformation between a matrix and its transpose, *Pacific J. Math.* 7:893–896 (1959).

13 J. H. Wilkinson, *The Algebraic Eigenvalue Problem*, Clarendon, Oxford, 1965.

Received 17 March 1983; revised 16 July 1983