NORTH-HOLLAND

On Fisher's Information Matrix of an ARMAX Process and Sylvester's Resultant Matrices

André Klein

Department of Economic Statistics
University of Amsterdam
Roetersstraat 11
1018 WB Amsterdam, The Netherlands

and

Peter Spreij

Department of Econometrics
Vrije Universiteit
De Boelelaan 1105
1081 HV Amsterdam, The Netherlands

Submitted by George P. H. Styan

ABSTRACT

We establish a relation between Fisher's information matrix of a stationary autoregressive moving average process, with an exogenous component, and two Sylvester's resultant matrices.

1. INTRODUCTION

The Cramér-Rao bound is of paramount importance for evaluating the performance of (stationary) autoregressive moving average models with an exogeneous component (ARMAX), where the focus is on the error covariance matrix of the estimated parameters. See Cramér [3] and Rao [8]. For computing the Cramér-Rao bound the inverse of Fisher's information matrix is needed. The latter is singular in the presence of common roots of the AR, the MA polynomial, and the polynomial describing the influence of the exogenous part.

The purpose of the present paper is to study the link between Fisher's information matrix and Sylvester's resultant matrices involving the three

LINEAR ALGEBRA AND ITS APPLICATIONS 237/238:579–590 (1996)

0024-3795/96/$15.00
655 Avenue of the Americas, New York, NY 10010 SSDI 0024-3795(95)00552-3

polynomials that model the ARMAX process. In a previous paper [7] we investigated the simpler case for ARMA processes.

The link between statistical considerations and algebraic results is of independent interest, since one is based on the Wald test statistic for testing common roots (see Klein [6] for the ARMA case), and the other one is deduced from the structure of the Fisher information matrix of a stationary ARMAX process.

In Barnett [1] a relationship between Sylvester's resultant matrix and the companion matrix of a polynomial is given. Kalman [4] has investigated the concept of observability and controllability in a function of Sylvester's resultant matrix. Similar results can be found in Barnett [2], which contains discussions on these topics and a number of further references. Furthermore, in Söderström and Stoica [9, p. 162 ff.] a discussion on overparametrization in terms of the transfer function of a system can be found.

2. PRELIMINARY ALGEBRAIC RESULTS AND NOTATION

Consider the following two scalar polynomials in the variable z:

$$A(z) = z^p + a_1 z^{p-1} + \cdots + a_p, \tag{2.1}$$

$$B(z) = z^q + b_1 z^{q-1} + \cdots + b_q. \tag{2.2}$$

The Sylvester resultant matrix of A and B is defined as the $(p+q) \times (p+q)$ matrix

$$S(A,B) = \begin{array}{c} q\left\{\begin{array}{c}\\ \\ \\ \end{array}\right. \\ p\left\{\begin{array}{c}\\ \\ \\ \end{array}\right. \end{array} \left[\begin{array}{cccccccc} 1 & a_1 & \cdots & \cdots & a_p & & 0 \\ & \ddots & \ddots & & & \ddots & \\ 0 & & 1 & a_1 & \cdots & \cdots & a_p \\ \hline 1 & b_1 & \cdots & \cdots & b_q & & 0 \\ & \ddots & \ddots & & & \ddots & \\ 0 & & 1 & b_1 & \cdots & \cdots & b_q \end{array}\right]. \tag{2.3}$$

In the presence of common roots of A and B the matrix $S(A,B)$ becomes singular. Moreover it is known that

$$\det S(A,B) = \prod_{i=1}^{p} \prod_{j=1}^{q} (\beta_j - \alpha_i), \tag{2.4}$$

where the α_i and the β_j are the roots of A and B respectively.

In the context of ARMA(X) processes it is more natural to work not with monic polynomials, but instead with polynomials with a constant term equal to 1. These are linked with the monic polynomials as follows. Given a monic polynomial A as in Equation (2.1), we define $a(z) = z^p A(z^{-1}) = 1 + a_1 z + \cdots + a_p z^p$. Similarly we have a qth order polynomial b given by $b(z) = 1 + b_1 z + \cdots + b_q z^q$ and an rth order polynomial c given by $c(z) = 1 + c_1 z + \cdots + c_r z^r$. In the sequel we will use the notation $S(a, b)$ to denote the Sylvester matrix in Equation (2.3) and not $S(A, B)$. Similar notation will be used for all other Sylvester matrices appearing in the rest of the paper.

3. MAIN RESULT

First we specify Fisher's information matrix of a Gaussian ARMAX(p, q, r) process. Let a, b, and c be the same polynomials as in the previous section. Consider then the stationary ARMAX process y that satisfies

$$a(z)y = b(z)u + c(z)\varepsilon. \tag{3.1}$$

Notice that we use z invariably as a complex number and as the lag operator. In Equation (3.1) u is a given stationary process with spectral density $(2\pi)^{-1}R_u$, and ε a white noise sequence with variance σ^2. The processes u and ε are supposed to be independent, or at least uncorrelated. Assume also that b has no zeros on the unit circle. Let θ denote the parameter vector, $\theta = (a_1, \ldots, a_p, b_1, \ldots, b_q, c_1, \ldots, c_r)^T$ (here and elsewhere superscript T stands for transposition). Denote by $\varepsilon_t^{\theta_i}$ the derivative of ε_t with respect to θ_i. Then we have

$$\begin{aligned}
\varepsilon_t^{a_j} &= \frac{b(z)}{a(z)c(z)} u_{t-j} + \frac{1}{a(z)} \varepsilon_{t-j}, \\
\varepsilon_t^{b_k} &= -\frac{1}{c(z)} u_{t-k}, \\
\varepsilon_t^{c_l} &= -\frac{1}{c(z)} \varepsilon_{t-l}.
\end{aligned}$$

We introduce some more convenient notation. For each positive integer k we write $u_k(z)^T = [1, z, \ldots, z^{k-1}]$. Then we have for the Fisher information matrix (we write ε_t^θ as the row vector of all the derivatives) $F(\theta) = \sigma^{-2} E_\theta \varepsilon_t^{\theta T} \varepsilon_t^\theta$. See also Klein and Mélard [5] for an expression in the more general case of SISO models.

For $G(\theta) = \sigma^2 F(\theta)$ we have the following block decomposition:

$$G(\theta) = \begin{bmatrix} G_{aa} & G_{ab} & G_{ac} \\ G_{ab}^T & G_{bb} & G_{bc} \\ G_{ac}^T & G_{bc}^T & G_{cc} \end{bmatrix}, \tag{3.2}$$

where the matrices appearing are given by

$$G_{aa} = \frac{1}{2\pi i} \oint_{|z|=1} \frac{b(z)b(z^{-1})}{a(z)a(z^{-1})c(z)c(z^{-1})} R_u(z) u_p(z) u_p(z^{-1})^T \frac{dz}{z}$$

$$+ \frac{\sigma^2}{2\pi i} \oint_{|z|=1} \frac{1}{a(z)a(z^{-1})} u_p(z) u_p(z^{-1})^T \frac{dz}{z},$$

$$G_{ab} = -\frac{1}{2\pi i} \oint_{|z|=1} R_u(z) \frac{b(z)}{a(z)c(z)c(z^{-1})} u_p(z) u_q(z^{-1})^T \frac{dz}{z},$$

$$G_{ac} = \frac{-\sigma^2}{2\pi i} \oint_{|z|=1} \frac{1}{a(z)c(z^{-1})} u_p(z) u_r(z^{-1})^T \frac{dz}{z},$$

$$G_{bb} = \frac{1}{2\pi i} \oint_{|z|=1} R_u(z) \frac{1}{c(z)c(z^{-1})} u_q(z) u_q(z^{-1})^T \frac{dz}{z},$$

$$G_{bc} = 0,$$

$$G_{cc} = \frac{\sigma^2}{2\pi i} \oint_{|z|=1} \frac{1}{c(z)c(z^{-1})} u_r(z) u_r(z^{-1})^T \frac{dz}{z}.$$

Let $K(z) = a(z)a(z^{-1})c(z)c(z^{-1})$. Putting all these expressions together, we can write $G(\theta)$ as the sum of the following two matrices:

$$\frac{1}{2\pi i} \oint_{|z|=1} R_u(z) \frac{1}{K(z)} \begin{bmatrix} b(z)u_p(z) \\ -a(z)u_q(z) \\ 0 \end{bmatrix}$$

$$\times \left[b(z^{-1})u_p(z^{-1})^T, -a(z^{-1})u_q(z^{-1})^T, 0 \right] \frac{dz}{z}, \tag{3.3}$$

$$\frac{\sigma^2}{2\pi i} \oint_{|z|=1} \frac{1}{K(z)} \begin{bmatrix} c(z)u_p(z) \\ 0 \\ -a(z)u_r(z) \end{bmatrix} \left[c(z^{-1})u_p(z^{-1})^T, 0, -a(z^{-1})u_r(z^{-1})^T \right] \frac{dz}{z}. \tag{3.4}$$

These matrices can be written in a more compact form. In order to do so, we introduce some notation. Furthermore we split the Sylvester matrix

$$S(-b,a) = \begin{bmatrix} -S_p(b) \\ S_q(a) \end{bmatrix}.$$

Here $S_p(b)$ is formed by the top p rows of $S(-b,a)$. In a similar way we decompose

$$S(-c,a) = \begin{bmatrix} -S_p(c) \\ S_r(a) \end{bmatrix}.$$

Then we can write the matrices in the expressions (3.3) and (3.4) respectively as

$$\begin{bmatrix} -S_p(b) \\ S_q(a) \\ 0 \end{bmatrix} \frac{1}{2\pi i} \oint_{|z|=1} R_u(z) \frac{u_{p+q}(z) u_{p+q}(z^{-1})^T}{K(z)} \frac{dz}{z} [-S_p(b)^T, S_q(a)^T, 0] \tag{3.5}$$

and

$$\begin{bmatrix} -S_p(c) \\ 0 \\ S_r(a) \end{bmatrix} \frac{\sigma^2}{2\pi i} \oint_{|z|=1} \frac{u_{p+r}(z) u_{p+r}(z^{-1})^T}{K(z)} \frac{dz}{z} [-S_p(c)^T, 0, S_r(a)^T]. \tag{3.6}$$

The main theorem is now the following

THEOREM 3.1. *The Fisher information matrix of an* ARMAX(p,q,r) *process with polynomials* $a(z)$, $b(z)$, *and* $c(z)$ *of order* p, q, r *respectively becomes singular iff these three polynomials have at least one common root.*

Proof. Because $G(\theta) = \sigma^2 F(\theta)$ is the sum of the two nonnegative definite matrices in Equations (3.5) and (3.6), and the two integrals appearing there are both strictly positve (which is shown in Theorem 3.1 of Klein and Spreij [7]), we see that a vector

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} \quad \text{with} \quad x \in \mathbf{R}^p, \quad y \in \mathbf{R}^q, \text{ and } z \in \mathbf{R}^r$$

belongs to $\ker G(\theta)$ iff

$$\begin{bmatrix} x \\ y \end{bmatrix} \in \ker S(-b,a)^T \quad \text{and} \quad \begin{bmatrix} x \\ z \end{bmatrix} \in \ker S(-c,a)^T.$$

Assume that there exists a common zero ϕ of a, b, and c. It follows from the appendix (Lemma A.1) that the vector

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} \quad \text{with} \quad x = \alpha(\phi), \quad y = \beta(\phi), \text{ and } z = \gamma(\phi)$$

belongs to the kernel of $G(\theta)$, which then must be singular.

In order to prove the converse we proceed as follows. Assume that

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} \in \ker G(\theta), \qquad \text{with} \quad x \in \mathbf{R}^p, \quad y \in \mathbf{R}^q, \quad z \in \mathbf{R}^r.$$

Let $\{\phi_1, \ldots, \phi_k\}$ be the set of all common zeros of A and B, and let $\{\psi_1, \ldots, \psi_l\}$ be the set of all common zeros of A and C (counting multiplicities). The other $p-k-l$ zeros of A are denoted by $\rho_1, \ldots, \rho_{p-k-l}$. Then we can write $A(z) = \prod_{i=1}^k (z-\phi_i) \prod_{i=1}^l (z-\psi_i) \prod_{i=1}^{p-k-l} (z-\rho_i)$. As in the appendix, we denote by $A(\Phi)$ the $p \times k$ matrix with ith column equal to $J_p^{i-1} T_{\Phi,p} \alpha$. Recall that $T_{\Phi,p}$ is short for $\prod_{i=1}^k T_{\phi_i,p}$, $T_{\Psi,p} = \prod_{i=1}^l T_{\psi_i,p}$, and $T_{R,p} = \prod_{i=1}^{p-k-l} T_{\rho_i,p}$. The matrices $B(\Phi)$, $A(\Psi)$, and $C(\Psi)$ have a similar meaning to $A(\Phi)$. From Lemma A.2 of the appendix we find that $\ker S(-b,a)^T$ is given by the image of the map with matrix

$$\begin{bmatrix} A(\Phi) \\ B(\Phi) \end{bmatrix}.$$

So there exists a vector $v \in \mathbf{R}^k$ such that

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} A(\Phi) \\ B(\Phi) \end{bmatrix} v.$$

Similarly, there exists a vector $w \in \mathbf{R}^l$ such that

$$\begin{bmatrix} x \\ z \end{bmatrix} = \begin{bmatrix} A(\Psi) \\ C(\Psi) \end{bmatrix} w.$$

Combining the two expressions for x, we get $A(\Phi)v - A(\Psi)w = 0$, and from this

$$T_{R,p}A(\Phi)v - T_{R,p}A(\Psi)w = [T_{R,p}A(\Phi), T_{R,p}A(\Psi)] \begin{bmatrix} v \\ -w \end{bmatrix} = 0. \qquad (3.7)$$

Recall that the ith column of $A(\Phi)$ is given by $J_p^{i-1}T_{\Phi,p}\alpha$ and therefore the ith column of $T_{R,p}A(\Phi)$ is given by $J_p^{i-1}T_{R,p}T_{\Phi,p}\alpha$. Similarly to the observation we make in the appendix [premultiplying of α with $T_{\phi,p}$ if ϕ is a zero of a monic polynomial A gives the coefficients of the polynomial $A(z)/(z-\phi)$], we notice that the first $l+1$ elements of the first column are exactly the coefficients of the polynomial $\prod_{i=1}^{l}(z-\psi_i)$, the last $p-l-1$ coefficients being zero. Likewise the first column of the matrix $T_{R,p}A(\Psi)$ has the coefficients of the polynomial $\prod_{i=1}^{k}(z-\phi_i)$ on its first $k+1$ places, the other $p-k-1$ elements being also zero. Therefore the matrix $[T_{R,p}A(\Phi), T_{R,p}A(\Psi)]$ in Equation (3.7) is nothing else but the transpose of the Sylvester matrix $S(\Phi,\Psi)$ extended with a $(p-k-l)\times(k+l)$ lower block of zeros. By assumption there are no zeros shared by a, b, and c, so in particular the set of common zeros of Φ and Ψ is empty, and therefore the Sylvester matrix $S(\Phi,\Psi)$ is nonsingular, and then the matrix $T_{R,p}[A(\Phi), A(\Psi)]$ as well as $[A(\Phi), A(\Psi)]$ has full column rank. We conclude that v and w are zero and so are x, y, and z. This proves the theorem. ■

APPENDIX

In this appendix we give some results on the kernel of a Sylvester matrix. Contrary to what we wrote in the main text, we usually denote monic polynomials by lowercase letters. Let $S(a,b)$ be the Sylvester matrix associated with the real polynomials a and b, where $a(z) = \sum_{k=0}^{p} a_k z^{p-k}$, $a_0 = 1$, and $b(z) = \sum_{k=0}^{q} b_k z^{q-k}$, $b_0 = 1$. So $S(a,b) \in \mathbf{R}^{(p+q)\times(p+q)}$ and

$$S(a,b)^T = \begin{bmatrix} 1 & & & & 1 & & & \\ a_1 & 1 & & & b_1 & 1 & & \\ \vdots & a_1 & \ddots & & \vdots & b_1 & \ddots & \\ a_p & \vdots & & 1 & b_q & & & 1 \\ & a_p & & a_1 & & b_q & & b_1 \\ & & \ddots & \vdots & & & \ddots & \vdots \\ & & & a_p & & & & b_q \end{bmatrix}.$$

Let for each ϕ the matrix $T_\phi \in C^{(p+q)\times(p+q)}$ be defined by having ij entry ϕ^{i-j} for $i \geq j$ and 0 otherwise. So

$$T_\phi = \begin{bmatrix} 1 & & & & & \\ \phi & 1 & & & & \\ \phi^2 & \phi & 1 & & & \\ \vdots & & \ddots & \ddots & & \\ \vdots & & & \ddots & 1 & \\ \phi^{p+q-1} & \cdots & \cdots & \cdots & \phi & 1 \end{bmatrix}.$$

Define also the polynomials a_k by the recursion $a_k(z) = za_{k-1}(z) + a_k$, with initial condition $a_{-1} = 0$ and the convention that $a_k = 0$ for $k > p$. It then follows that $a_k(z) = z^{k-p}a(z)$ for $k \geq p$, and consequently $a_k(z) = 0$ for $k \geq p$ if z is a zero of a. Similarly we define the polynomials b_k. With the aid of these polynomials we also define the polynomial $a(\cdot;\phi)$ by $a(z;\phi) = \sum_{k=0}^{p-1} a_k(\phi)z^{p-1-k}$. An easy calculation shows that $A(z;\phi) \equiv (z-\phi)a(z;\phi) = a(z) - a(\phi)$, whence the conclusion $A(\cdot;\phi) = a$ iff ϕ is a zero of a. Consider now the product

$$T_\phi S(a,b)^T = \begin{bmatrix} 1 & & & & 1 & & & \\ a_1(\phi) & 1 & & & b_1(\phi) & 1 & & \\ \vdots & a_1(\phi) & \ddots & & \vdots & b_1(\phi) & \ddots & \\ \vdots & & \ddots & 1 & \vdots & & \ddots & 1 \\ \vdots & & & a_1(\phi) & \vdots & & & b_1(\phi) \\ \vdots & & & \vdots & \vdots & & & \vdots \\ a_{p+q-1}(\phi) & \cdots & \cdots & a_p(\phi) & b_{p+q-1}(\phi) & \cdots & \cdots & b_q(\phi) \end{bmatrix}. \tag{A.1}$$

In order to prepare for the lemma below we introduce some notation. Let $\alpha^T = [1,\ldots,a_{p-1}] \in \mathbf{R}^p$ and $\beta^T = [1,\ldots,b_{q-1}] \in \mathbf{R}^q$. With the notation introduced at the beginning of this appendix we write $\alpha(\phi)^T = [1, a_1(\phi),\ldots,a_{p-1}(\phi)] \in \mathbf{R}^p$ and $\beta(\phi) \in \mathbf{R}^q$ likewise. In the main text of this paper we often encountered Sylvester matrices of the type $S(-b,a)$. In the rest of this appendix we will state some results for these matrices.

For pth order polynomials a and qth order polynomials b we sometimes partition the Sylvester matrix $S(-b,a)$ into its upper block $-S_p(b) \in$

$\mathbf{R}^{p\times(p+q)}$ and its lower block $S_q(a) \in \mathbf{R}^{q\times(p+q)}$. We also use 0_p, 0_q, 0_{p-1}, and 0_{q-1} to denote the zero vectors in $\mathbf{R}^p$, $\mathbf{R}^q$, $\mathbf{R}^{p-1}$, and $\mathbf{R}^{q-1}$ respectively. I_{p-1} and I_{q-1} are the identity matrices of orders $p-1$ and $q-1$.

LEMMA A.1. *Let $S = S(-b, a)$ be the Sylvester matrix associated with the polynomials a and b as above, and suppose that ϕ is a common zero of a and b. Then*

$$\ker S^T = \begin{bmatrix} 0_{p-1}^T \\ I_{p-1} \\ 0_{q-1}^T \\ I_{q-1} \end{bmatrix} \ker S(-b(\cdot;\phi), a(\cdot;\phi))^T \oplus \left\langle \begin{bmatrix} \alpha(\phi) \\ \beta(\phi) \end{bmatrix} \right\rangle. \tag{A.2}$$

In particular we have that $\dim\ker S(-b,a)^T = \dim\ker S(-b(\cdot;\phi), a(\cdot\phi))^T + 1$.

Proof. Obviously for $x_1 \in \mathbf{R}^{p-1}$ and $x_2 \in \mathbf{R}^{q-1}$ we have

$$\begin{bmatrix} 0 \\ x_1 \\ 0 \\ x_2 \end{bmatrix} \in \ker S^T \qquad \text{if} \qquad \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \in \ker S^T(-b(\cdot,\phi), a(\cdot,\phi)).$$

Write $\alpha_1^T = [a_1, \ldots, a_p]$, $\beta_1^T = [b_1, \ldots, b_q]$. It is easy to see that

$$S(-b,a)^T \begin{bmatrix} \alpha_1 \\ \beta_1 \end{bmatrix} = \begin{bmatrix} \beta_1 \\ 0_p \end{bmatrix} - \begin{bmatrix} \alpha_1 \\ 0_q \end{bmatrix}. \tag{A.3}$$

A similar result holds if we replace a and b with $a(\cdot;\phi)$ and $b(\cdot;\phi)$, and α_1 and β_1 with $\alpha(\phi)_1 = [a_1(\phi), \ldots, a_{p-1}(\phi)]^T$ and $\beta(\phi)_1 = [b_1(\phi), \ldots, b_{q-1}(\phi)]^T$. So

$$S(-b(\cdot,\phi), a(\cdot,\phi))^T \begin{bmatrix} \alpha(\phi)_1 \\ \beta(\phi)_1 \end{bmatrix} = \begin{bmatrix} \beta(\phi)_1 \\ 0_{p-1} \end{bmatrix} - \begin{bmatrix} \alpha(\phi)_1 \\ 0_{q-1} \end{bmatrix}. \tag{A.4}$$

In parallel notation to the one introduced above we have

$$T_\phi S(-b,a)^T = \begin{bmatrix} 1 & 0 \cdots 0 & 1 & 0 \cdots 0 \\ \begin{bmatrix} -\beta(\phi)_1 \\ 0_{p-1} \end{bmatrix} & -S_{p-1}^T(b(\cdot;\phi)) & \begin{bmatrix} \alpha(\phi)_1 \\ 0_{q-1} \end{bmatrix} & S_{q-1}^T(a(\cdot;\phi)) \\ 0 & 0 \cdots 0 & 0 & 0 \cdots 0 \end{bmatrix}. \tag{A.5}$$

Hence we see from Equations (A.4) and (A.5) that

$$\begin{bmatrix} \alpha(\phi) \\ \beta(\phi) \end{bmatrix} = \begin{bmatrix} 1 \\ \alpha(\phi)_1 \\ 1 \\ \beta(\phi)_1 \end{bmatrix}$$

belongs to the kernel of $S(-b,a)^T$. Since the two subspaces in (A.2) have a trivial intersection, we get that their sum is actually a direct sum and that it belongs to $\ker S(-b,a)^T$.

Now we have to prove the converse. So take a vector $v \in \ker S(-b,a)^T$, which we can always partition as

$$\begin{bmatrix} z \\ x \\ z \\ y \end{bmatrix} \quad \text{for some} \quad x \in \mathbf{R}^{p-1}, \quad y \in \mathbf{R}^{q-1}, \text{ and } z \in \mathbf{R}.$$

We have to show that

$$\begin{bmatrix} x - z\alpha(\phi)_1 \\ y - z\beta(\phi)_1 \end{bmatrix} \in \ker S(-b(\cdot;\phi), a(\cdot;\phi)).$$

From the partitioning (A.5) of $T_\phi S(-b,a)$ and the fact that $v \in \ker S(-b,a)$ we readily obtain the equality

$$S^T(-b(\cdot;\phi), a(\cdot;\phi)) \begin{bmatrix} x \\ y \end{bmatrix} = z \left(\begin{bmatrix} \beta(\phi)_1 \\ 0_{p-1} \end{bmatrix} - \begin{bmatrix} \alpha(\phi)_1 \\ 0_{q-1} \end{bmatrix} \right).$$

The result then follows from Equation (A.4). ■

Let now $J_p \in \mathbf{R}^{p\times p}$ be the shifted unit matrix defined by its elements $J_{p,ij} = \delta_{i,j+1}$, and J_q likewise. Furthermore for complex numbers $\phi_1, \ldots, \phi_k$ we define the $p \times p$ matrices $T_{\phi_l,p}$ given by their ij entries $T^{ij}_{\phi_l,p} = \phi_l^{i-j}$ if $i \geq j$ and zero elsewhere. Note that the matrices T_{ϕ_l} commute. Then we define the matrix $T_{\Phi,p}$ as the product of the matrices $T_{\phi_l,p}$. Notice that commutativity makes the order of multiplication irrelevant. It is also easy to see that the matrices J_p and $T_{\phi_l,p}$ commute. The matrix $T_{\Phi,q}$ is defined as its $q \times q$ analog.

Denote by $A(\Phi)$ the matrix with $J_p^{j-1}T_{\Phi,p}\alpha$ as its jth column. Similarly we write $B(\Phi)$ for the matrix with $J_p^{j-1}T_{\Phi,q}\beta$ as its jth column.

The following result is a direct consequence of the previous lemma.

LEMMA A.2. *Let $\phi_1, \ldots, \phi_k$ be all the common zeros of the polynomials a and b (counting multiplicity). With the notation introduced before, we have that* $\ker S(-b,a)^T$ *is k-dimensional and spanned by the independent vectors*

$$\left\{ \begin{bmatrix} J_p^{i-1} T_{\Phi,p}\alpha \\ J_q^{i-1} T_{\Phi,q}\beta \end{bmatrix} \right\}_{i=1}^{k}.$$

Equivalently, $\ker S(-b,a)^T$ *is the image space of*

$$\begin{bmatrix} A(\Phi) \\ B(\Phi) \end{bmatrix}.$$

Proof. The idea is to iterate the procedure described in the proof of the previous lemma by decomposing $\ker S(-b(\cdot;\phi_1), a(\cdot;\phi_1))^T$ in the same way as we decomposed $\ker S(-b,a)^T$.

So we find that $\ker S(-b(\cdot;\phi_1), a(\cdot;\phi_1))^T$ is the direct sum of

$$\left\langle \begin{bmatrix} \alpha(\phi_1,\phi_2) \\ \beta(\phi_1,\phi_2) \end{bmatrix} \right\rangle \quad \text{and} \quad \ker S(-b(\cdot;\phi_1,\phi_2), a(\cdot;\phi_1,\phi_2))^T,$$

where e.g. $\alpha(\phi_1,\phi_2) = T_{\phi_2,p-1}[I_{p-1}, 0_{p-1}]\alpha(\phi_1)$ and $a(z;\phi_1,\phi_2) = a(z)/(z-\phi_1)(z-\phi_2)$. This shows that also

$$\begin{bmatrix} 0 \\ \alpha(\phi_1,\phi_2) \\ 0 \\ \beta(\phi_1,\phi_2) \end{bmatrix} = \begin{bmatrix} J_p T_{\phi_2,p}\alpha(\phi_1) \\ J_q T_{\phi_2,q}\beta(\phi_1) \end{bmatrix}$$

belongs to $\ker S(-b,a)^T$. The iterations give step by step Sylvester matrices of decreasing dimension, and they stop when we arrive at the Sylvester matrix $S(-b(\cdot;\phi_1,\ldots,\phi_k), a(\cdot;\phi_1,\ldots,\phi_k))^T$, where $a(z;\phi_1,\ldots,\phi_k) = a(z)/\prod_{i=1}^k (z-\phi_i)$ and $b(z;\phi_1,\ldots,\phi_k) = b(z)/\prod_{i=1}^k (z-\phi_i)$. Since these polynomials have no common zeros, the associated Sylvester matrix is nonsingular. In conclusion we can say that $\ker S(-b,a)^T$ is spanned by the k independent vectors

$$\begin{bmatrix} v_i \\ w_i \end{bmatrix} \quad \text{with} \quad v_i = J_p^{i-1} T_{\phi_i,p} \cdots T_{\phi_1,p}\alpha \text{ and } w_i = J_p^{i-1} T_{\phi_i,p} \cdots T_{\phi_1,p}\beta$$

with $i = 1, \ldots, k$.

Define for $j = 1, \ldots, k$ polynomials $f_j(z) = \prod_{i=j+1}^k (z-\phi_i)$, and write these as $f_j(z) = \sum_{i=1}^k N_{ij} z^{k-i}$, so $N_{ij} = 0$ if $i < j$ and $N_{jj} = 1$. By f_j^* we denote the reversed polynomial, $f_j^*(z) = z^{k-j} f_j(z^{-1}) = \prod_{i=j+1}^k (1 - \phi_i z)$. Notice that for $j = k$ we get an empty product, which is 1 by convention.

It is easy to check that $T_{\phi_i,p}^{-1} = I - \phi_i J_p$ and hence $\sum_{i=1}^k N_{ij} J_p^{i-j} = f_j^*(J_p) = (T_{\phi_{j+1},p} \cdots T_{\phi_k,p})^{-1}$ for $j < k$ and $f_k^*(J_p) = I_p$.

Let N be the $k \times k$ matrix with entries N_{ij} as above. Consider now the jth column of the product $[\alpha, J_p\alpha, \dots, J_p^{k-1}\alpha]N$. It is given by $\sum_{i=1}^k J_p^{i-1} N_{ij} = f_j^*(J_p)\alpha$. Then we get

$$\begin{aligned} A(\Phi)N &= T_{\Phi,p}\big[\alpha, J_p\alpha, \dots, J_p^{k-1}\alpha\big]N \\ &= T_{\Phi,p}\big[f_1^*(J_p)\alpha, \dots, f_k^*(J_p)J_p^{k-1}\alpha\big] \\ &= \big[T_{\phi_1,p}\alpha, \dots, T_{\phi_1,p} \cdots T_{\phi_k,p}J_p^{k-1}\alpha\big] \\ &= [v_1, \dots, v_k]. \end{aligned}$$

Similarly we have the relation $B(\Phi)N = [w_1, \dots, w_k]$. So we finally have

$$\begin{bmatrix} v_1 & \cdots & v_k \\ w_1 & \cdots & w_k \end{bmatrix} = \begin{bmatrix} A(\Phi) \\ B(\Phi) \end{bmatrix} N,$$

which proves the lemma, since N is nonsingular. ■

REFERENCES

1 S. Barnett, A new formulation of the the theorems of Hurwitz, Routh and Sturm, *J. Inst. Math. Appl.* 8:240–250 (1971).

2 S. Barnett, Matrices, polynomials and linear time invariant systems, *IEEE Trans. Automat. Control* AC-18:1–10 (1973).

3 H. Cramér, *Mathematical Methods of Statistics*, Princeton U.P., Princeton, N.J., 1951.

4 R. E. Kalman, Mathematical description of linear dynamical systems, *SIAM J. Control* 1:152–192 (1963).

5 A. Klein and G. Mélard, Computation of the Fisher information matrix for SISO models, *IEEE Trans. Signal Process.* 42:684–688 (1994).

6 A. Klein, Hypothesis testing of common roots, *Statist. Probab. Lett.* 20:163–167 (1994).

7 A. Klein and P. J. C. Spreij, On Fisher's Information Matrix of an ARMA Process and Sylvester's Resultant Matrix, Research Memorandum 1993-33, Dept. of Econometrics, Vrije Univ., 1993.

8 C. R. Rao, *Linear Statistical Inference and Its Applications*, New York, Wiley, 1965.

9 T. Söderström and P. Stoica, *System Identification*, Prentice-Hall, 1989.

Received 23 June 1993; final manuscript accepted 14 June 1995