On Fisher's Information Matrix of an ARMA Process

André Klein and Peter Spreij

Abstract

In this paper we study the Fisher information matrix for a stationary ARMA process with the aid of Sylvester's resultant matrix. Some properties are explained via realizations in state space form of the derivates of the white noise process with respect to the parameters.

1. Introduction

The Cramér-Rao bound is of considerable importance for evaluating the performance of (stationary) autoregressive moving average (ARMA) models, where the focus is on the error covariance matrix of the estimated parameters. See Cramér [Cr] and Rao [Ra]. For computing the Cramér-Rao bound and the asymptotic distribution of a Wald test statistic (Klein [Kl]) the inverse of Fisher's information matrix is needed. The latter is singular in the presence of common roots of the AR and the MA polynomial and vice versa. This fact is considered to be well-known in time series analysis, see [Po1] or [McL] for a proof along different lines than the ones we follow below.

In this paper we give an elementary proof of this equivalence by linking Fisher's information matrix to Sylvester's resultant matrix and an interpretation in terms of a state space realization.

In the literature the resultant matrix has been used in various studies in the fields of time series and systems theory. For instance, in [AS] this matrix shows up in a convergence analysis of maximum likelihood estimators of the ARMA parameters (more precisely in the study of the convergence of the criterion function), in Barnett [Ba1] a relationship between Sylvester's resultant matrix and the companion matrix of a polynomial is given. Kalman [Ka] has investigated the concept of observability and controllability in function of Sylvester's resultant matrix. Similar results can be found in Barnett [Ba2] which contains discussions on these topics and a number of further references.

Fisher's information matrix is studied in [Ro] for problems of local and global identifiability in a static context, whereas identifiability problems for parameterizations of linear (stochastic) systems are studied by Glover and Willems in [GW].

I. Csiszár et al., *Stochastic Differential and Difference Equations*

Furthermore in Söderström & Stoica [So] (pages 162 ff.) a discussion on overparameterization in terms of the transfer function of a system can be found.

We now introduce some notation. Consider the following two scalar polynomials in the variable z.

$$A(z) = z^p + a_1 z^{p-1} + \ldots + a_p \tag{1.1}$$

$$C(z) = z^q + c_1 z^{q-1} + \ldots + c_q \tag{1.2}$$

By A^* and C^* we denote the reciprocal polynomials, so $A^*(z) = z^p A(z^{-1})$ and $C^*(z) = z^q C(z^{-1})$, and also write $a^T = [a_1, \ldots, a_p]$ and $c^T = [c_1, \ldots, c_q]$.

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

$$S(a,c) = \begin{array}{c} q\left\{ \vphantom{\begin{matrix} 1 \\ \ddots \\ 0 \end{matrix}} \right. \\ \\ p\left\{ \vphantom{\begin{matrix} 1 \\ \ddots \\ 0 \end{matrix}} \right. \end{array} \left[\begin{array}{cccccc} 1 & a_1 & \cdots & \cdots & a_p & 0 \\ & \ddots & \ddots & & & \ddots \\ 0 & & 1 & a_1 & \cdots \cdots & a_p \\ \hdashline 1 & c_1 & \cdots & \cdots & c_q & 0 \\ & \ddots & \ddots & & & \ddots \\ 0 & & 1 & c_1 & \cdots \cdots & c_q \end{array} \right] \tag{1.3}$$

In the presence of common roots of A and C the matrix $S(a,c)$ becomes singular. Moreover it is known (see e.g. [VW, page 106]) that

$$\det S(a,c) = \prod_{i=1}^{p} \prod_{j=1}^{q} (\gamma_j - \alpha_i) \tag{1.4}$$

where the α_i and the γ_j are the roots of A and C respectively.

Remark. The origin of Sylvester's matrix lies in the following problem, see [VW]. Find monic polynomials $K(z) = z^p + \sum_{i=1}^{p} k_i z^{p-i}$ and $L(z) = z^q + \sum_{i=1}^{q} l_i z^{q-i}$ such that $A(z)L(z) + C(z)K(z) = 0$. Writing the coefficients of K and L in column vectors k and l, one can cast this problem as solving a set of linear equations in k and l with $S(a,c)^T$ as coefficient matrix. Clearly, the solution of the problem is then given by the affine subspace $\begin{bmatrix} -a \\ c \end{bmatrix} + \ker S(a,c)^T$. Notice that for any solution K and L of the problem the rational function K/L coincides with A/C in the points where both are defined.

Remark. Let J_{p+q} the matrix in $\mathbf{R}^{(p+q)\times(p+q)}$ with ij-entries $\delta_{i,j+1}$ (a shifted identity matrix), $\alpha^T = (1, a_1, \ldots, a_p, \ldots, 0) \in \mathbf{R}^{p+q}$ and $\gamma^T = (1, c_1, \ldots, c_q, \ldots, 0) \in \mathbf{R}^{p+q}$. Then up to a permutation of its columns, the matrix $S(a,c)^T$ consists of the first $p+q$ columns of the controllability matrix of the pair $(J_{p+q}, (\alpha, \gamma))$,

2. A key result

First we specify Fisher's information matrix of an ARMA(p,q) process. Let A and C be the same monic polynomials as in the previous section. Consider then the stationary ARMA process y that satisfies

$$A^*(L)y = C^*(L)\varepsilon \tag{2.1}$$

with L the lag operator and ε a white noise sequence. We make the assumption (to give the expressions that we use below the correct meaning) that both A and C have zeros only inside the unit circle.
Let $\theta = (a_1, \ldots, a_p, c_1, \ldots . c_q)$ and denote by $\varepsilon_t^{\theta_i}$ the derivative of ε_t with respect to θ_i. Then we have

$$\begin{aligned} \varepsilon_t^{a_j} &= \frac{1}{a(z)}\varepsilon_{t-j} \\ \varepsilon_t^{c_l} &= -\frac{1}{c(z)}\varepsilon_{t-l} \end{aligned}$$

With ε_t^{θ} the column vector with elements $\varepsilon_t^{\theta_i}$ the Fisher information matrix $F(\theta)$ is equal to $E\varepsilon_t^{\theta}\varepsilon_t^{\theta T}$. As can be found in for instance Klein & Mélard [KM] $F(\theta)$ then has the following block decomposition

$$F(\theta) = \begin{bmatrix} F_{aa} & F_{ac} \\ F_{ac}^T & F_{cc} \end{bmatrix} \tag{2.2}$$

where the matrices appearing here have the following elements

$$\begin{aligned} F_{aa}^{jk} &= \frac{1}{2\pi i}\oint_{|z|=1} \frac{z^{j-k+p-1}}{A(z)A^*(z)}dz, (j,k = 1,\ldots,p) \\ F_{ac}^{jk} &= \frac{-1}{2\pi i}\oint_{|z|=1} \frac{z^{j-k+q-1}}{C(z)A^*(z)}dz, (j = 1,\ldots,p, k = 1,\ldots,q) \\ F_{cc}^{jk} &= \frac{1}{2\pi i}\oint_{|z|=1} \frac{z^{j-k+q-1}}{C(z)C^*(z)}dz, (j,k = 1,\ldots,q) \end{aligned}$$

The key result of this paper is the easy to prove lemma 2.1 below. First we have to introduce some auxiliary notation. Write for each positive integer k

$$\begin{aligned} u_k(z) &= [1, z, \ldots, z^{k-1}]^T \\ u_k^*(z) &= [z^{k-1}, \ldots, 1]^T = z^{k-1}u_k(z^{-1}) \end{aligned}$$

and let

$$K(z) = A(z)A^*(z)C(z)C^*(z).$$

Define moreover

$$P(\theta) = \frac{1}{2\pi i}\oint_{|z|=1} \frac{u_{p+q}(z)u_{p+q}^*(z)^T}{K(z)} dz \tag{2.3}$$

Notice that we can alternatively write

$$P(\theta) = \frac{1}{2\pi i}\oint_{|z|=1} \frac{u_{p+q}^*(z)u_{p+q}^*(\frac{1}{z})^T}{A(z)A(\frac{1}{z})C(z)C(\frac{1}{z})} \frac{dz}{z} \tag{2.4}$$

Lemma 2.1 *The following factorization holds.*

$$F(\theta) = S(c, -a)P(\theta)S(c, -a)^T \tag{2.5}$$

Proof. A simple computation shows that we can write $F(\theta)$ in matrix form as

$$F(\theta) = \frac{1}{2\pi i}\oint_{|z|=1} \frac{1}{K(z)} \begin{bmatrix} C^*(z)u_p(z) \\ -A^*(z)u_q(z) \end{bmatrix} \begin{bmatrix} C(z)u_p^*(z)^T & -A(z)u_q^*(z)^T \end{bmatrix} dz \tag{2.6}$$

It also straightforward to verify that the following identities hold.

$$S(c, -a)u_{p+q}(z) = \begin{bmatrix} C^*(z)u_p(z) \\ -A^*(z)u_q(z) \end{bmatrix} \tag{2.7}$$

$$S(c, -a)u_{p+q}^*(z) = \begin{bmatrix} C(z)u_p^*(z) \\ -A(z)u_q^*(z) \end{bmatrix} \tag{2.8}$$

Hence equation (2.5) follows now immediately from equations (2.3), (2.6), (2.7) and (2.8). ∎

Remark. It follows that with probability one ε_t^θ belongs to the image space of $S(c, -a)$.

Corollary 2.2 *The Fisher information matrix of an ARMA(p,q) process with polynomials $A^*(z)$ and $C^*(z)$ of order p, q respectively becomes singular iff the polynomials A and C have at least one common root.*

Proof. Clearly the matrix $F(\theta)$ becomes singular if A and C have at least one common root in view of equation (1.4) and lemma 2.1. In order to prove the converse, we only have to prove that $P(\theta)$ is strictly positive definite, (again) because of (1.4) and (2.5). This can be shown via a straight forward computation (see also the next section for an alternative consideration):

Rewrite $P(\theta)$ as

$$P(\theta) = \frac{1}{2\pi i} \oint_{|z|=1} \frac{u_{p+q}(z)u_{p+q}(z^{-1})^T}{A(z)A(z^{-1})C(z)C(z^{-1})} z^{-1} dz$$

Take now $z = e^{i\phi}$, then we get

$$P(\theta) = \frac{1}{2\pi} \int_0^{2\pi} \frac{u_{p+q}(e^{i\phi})u_{p+q}(e^{-i\phi})^T}{A(e^{i\phi})A(e^{-i\phi})C(e^{i\phi})C(e^{-i\phi})} d\phi$$

which in turn can be rewritten as

$$P(\theta) = \frac{1}{2\pi} \int_0^{2\pi} \frac{u_{p+q}(e^{i\phi})}{A(e^{i\phi})C(e^{i\phi})} \overline{\frac{u_{p+q}(e^{i\phi})}{A(e^{i\phi})C(e^{i\phi})}}^T d\phi$$

Let now $x \in \mathbf{R}^{p+q}$ such that $x^T P(\theta) = 0$. Then it follows that $x^T \frac{u_{p+q}(e^{i\phi})}{A(e^{i\phi})C(e^{i\phi})} = 0$ for almost all ϕ. But this is clearly only possible if $x = 0$. So $P(\theta) > 0$. ■

Remark. In view of lemma 2.1, the nonsingularity of $P(\theta)$ and the first remark in the introduction we observe that for any $\begin{bmatrix} k \\ l \end{bmatrix}$ in the affine subspace $\begin{bmatrix} a \\ c \end{bmatrix} + \ker F(\theta)$ we find that the rational function $L(z)/K(z)$, where K and L depend on k and l as before, equals the transfer function $C(z)/A(z)$. This fact has also been noticed in [Pol], although the link with Sylvester's matrix is absent. A fairly explicit characterization of $\ker F(\theta)$ is given in [KS].

As a side remark we notice that the matrix $P(\theta)$ can be calculated by means of Cauchy's integral formula in the presence of common roots as follows. Let δ be a common root of A and C that appears as a zero of AC of order $l \geq 2$. Then

$$P(\theta) = \frac{1}{2\pi i} \oint_{|z|=1} \frac{f(z)}{(z-\delta)^l} dz$$

with

$$f(z) = \frac{u_{p+q}(z)u_{p+q}(z^{-1})^T}{A(z)A(z^{-1})C(z)C(z^{-1})} \frac{(z-\delta)^l}{z},$$

which is analytic in a disk of radius ρ around δ for sufficiently small ρ. Cauchy's theorem states that $P(\theta)$ is the sum of residuals, of which in particular the residual in δ can be computed as

$$f^{(l-1)}(\delta) = \frac{(l-1)!}{2\pi i} \oint_{|z-\delta|=\rho} \frac{f(z)}{(z-\delta)^l} dz. \tag{2.9}$$

It then follows that the more common roots A and C have, the less residuals are needed for the computation of $P(\theta)$.

As a corollary to lemma 2.1 we mention the following. Consider an AR process of order m, with AR polynomial $\tilde{A}^*(z)$ of order m. According to equation (2.5) and the fact that the Sylvester matrix is now the m-dimensional unit matrix, the Fisher information matrix $\tilde{F}$ becomes in this case

$$\tilde{F} = \frac{1}{2\pi i} \oint_{|z|=1} \frac{u_m(z) u_m^*(z)^T}{\tilde{A}(z)\tilde{A}^*(z)} dz \tag{2.10}$$

Take now in particular $\tilde{A}^*(z) = A^*(z)C^*(z)$ (so $m = p+q$), then it follows again from equation (2.10) and the fact that now $\tilde{A}(z)\tilde{A}^*(z) = K(z)$, that one has $P(\theta) = \tilde{F}$ and hence equation (2.5) reads

$$F(\theta) = S(c,-a)\tilde{F}S(c,-a)^T \tag{2.11}$$

So equation (2.11) gives a relationship between the Fisher information matrix of an ARMA(p,q) process and that of an appropriate AR($p+q$) process. See the next section for an explanation in state space terms of this phenomenon.

3. Computations in state space

We start with a realization of the ARMA process in state space form. Let $n = \max\{p,q\}$, $\alpha_n, \gamma_n \in \mathbf{R}^n$, $\alpha_n = [a_1, \ldots, a_n]^T$ and $\gamma_n = [c_1, \ldots, c_n]^T$, with zero entries for possibly previously undefined a_k or c_k (which happens only if $p \neq q$). We choose the following controllable realization.

$$X_{t+1} = AX_t + e\varepsilon_t \tag{3.1}$$
$$y_t = (\gamma_n - \alpha_n)^T X_t + \varepsilon_t \tag{3.2}$$

Here e is the first basis vector of the Euclidean space $\mathbf{R}^n$. A is then the matrix $J - e\alpha_n^T$, with J the shifted identity matrix, $J_{ij} = 1$ if $i = j+1$ and zero else. Below we also use the matrix F given by $F = J - e\gamma_n^T$.

Later on we will also use the notation e for the first basis vector in Euclidean spaces of possibly different dimension. Similarly, we denote by I the identity matrix of the appropriate size and 0 stands for the zero vector or matrix of appropriate dimensions. Occasionally these matrices and vectors will have a subscript, when it is necessary to indicate the sizes. Furthermore we will also use non-square 'identity' matrices, like I_{pn} which is the matrix in $\mathbf{R}^{p\times n}$ having ij-element δ_{ij}.
As before we denote by superscript partial derivatives with respect to a parameter. Let $Z_t = \text{vec}\,(X_t, X_t^{a_1}, \ldots, X_t^{a_p}, X_t^{c_1}, \ldots, X_t^{c_q})$. Then we can represent ε^θ after elementary computations as the output of the following system

$$Z_{t+1} = \mathcal{A}Z_t + e\varepsilon_t \tag{3.3}$$

$$\varepsilon_t^\theta = \mathcal{C}Z_t \tag{3.4}$$

where

$$\mathcal{A} = \begin{bmatrix} A & 0 & 0 \\ 0 & I_p \otimes F & 0 \\ -I_{qn} \otimes e_n & 0 & I_q \otimes F \end{bmatrix}$$

and

$$\mathcal{C} = \begin{bmatrix} I_{pn} & -I_p \otimes (\gamma_n - \alpha_n)^T & 0 \\ -I_{qn} & 0 & -I_q \otimes (\gamma_n - \alpha_n)^T \end{bmatrix}$$

If one computes the controllability matrix of this system, it is immediately seen that it contains a (middle) row of zero matrices. Therefore we replace it by the following system, using the same notation for the state variable and the coefficients.

$$Z_{t+1} = \mathcal{A}Z_t + e\varepsilon_t \tag{3.5}$$

$$\varepsilon_t^\theta = \mathcal{C}Z_t \tag{3.6}$$

where now

$$\mathcal{A} = \begin{bmatrix} A & 0 \\ -I_{qn} \otimes e_n & I_q \otimes F \end{bmatrix}$$

and

$$\mathcal{C} = \begin{bmatrix} I_{pn} & 0 \\ -I_{qn} & -I_q \otimes (\gamma_n - \alpha_n)^T \end{bmatrix}$$

Also this system is not controllable and we reduce it to one of lower order that is controllable by the same procedure that is used to decompose the state space a given system as a direct sum of the controllable subspace and its complement and the corresponding partitioning of the system matrices.

This is a well known procedure that we therefore only briefly sketch, leaving some computational details aside. We will work with the assumption that $c_q \neq 0$, although propositions 3.3 and 3.4 below and their consequences can also be proved if we drop this assumption by a little more complicated analysis.

First we compute the controllable subspace which is the linear span of the set $\{(I - \mathcal{A}z)^{-1}e : z \in (-\delta, \delta)\}$, where δ is sufficiently small. Then $(I - \mathcal{A}z)^{-1}e$ after a computation turns out to be equal to $(((I - Az)^{-1}e)^T, (-I_{qn}(I - Az)^{-1}e \otimes (I - Fz)^{-1}e)^T)^T$. Then the controllability matrix is computed by evaluating the derivatives of all orders of this function at $z = 0$. It is easily shown that the first $n + q$ columns of the controllability matrix span the controllable subspace and form a basis in the generic case where $c_q \neq 0$. In this case we form a non-singular matrix S consisting of these vectors augmented with a set of arbitrary independent vectors. Use S as a state space transformation to get a system described by the matrices $S^{-1}\mathcal{A}S$, $S^{-1}e = e$ and $\mathcal{C}S$. Then we first restrict the system to the controllable subspace and then by also restricting it further if $p < q$ in an appropriate way to a smaller subspace we get a new system (we use Z again to denote the state variable) that is given by

$$Z_{t+1} = \bar{A}Z_t + e\varepsilon_t \tag{3.7}$$

$$\varepsilon_t^\theta = \bar{C}Z_t \tag{3.8}$$

Here $\bar{A}$ is a companion matrix $\bar{A} = J - \hat{g}[0, \ldots, 0, 1]$, where $\hat{g}$ is the vector $\hat{g} = [g_{p+q}, \ldots, g_1]^T$ and the entries g_i are given by $z^{p+q} + \sum_{i=1}^{p+q} g_i z^{p+q-i} = A(z)C(z)$.

Finally we transform this system with the aid of the matrix $T \in \mathbb{R}^{(p+q)\times(p+q)}$ defined by its entries $T_{ij} = g_{j-i}$, with the convention that $g_0 = 1$ and $g_k = 0$ if $k < 0$. We then arrive at

Proposition 3.3 *The process ε^θ can be realized by the following stable and controllable system*

$$Z_{t+1} = \hat{A}Z_t + e\varepsilon_t \tag{3.9}$$

$$\varepsilon_t^\theta = \hat{C}Z_t, \tag{3.10}$$

where $\hat{A} = T^{-1}\bar{A}T = J - eg^T$ with $g^T = [g_1, \ldots, g_{p+q}]$, and $\hat{C} = S(c, -a)$. This system is observable iff the polynomials A and C have no common zeros.

Proof. The procedure outlined above has already shown the validity of equations (3.9) and (3.10). Controllability is obvious. If one computes the observability matrix one immediately sees that it has full rank iff the Sylvester matrix is invertible, so iff the A and C polynomials have no common factors. Furthermore the matrix $\hat{A}$ is stable, because its characteristic polynomial is $A(z)C(z)$ and we had assumed that all zeros of both A and C lie in the open unit disk. We skip the computations leading to $\hat{C} = S(c, -a)$. ∎

The result of the cascade of state space transformations leading to (3.9) and (3.10) can be summarized as follows. Start with the system described by equations (3.5) and (3.6). Apply a transformation with a nonsingular matrix $\mathcal{M} \in \mathbb{R}^{(n+n^2)\times(n+n^2)}$, that is such that its upperleft block of size $(p+q) \times (p+q)$ is given by

$$\begin{bmatrix} S_p(c) \\ 0_{q\times 1} \; I_q \; 0_{q\times(q-1)} \end{bmatrix}$$

This matrix is nonsingular under the condition that $c_q \neq 0$. Then it is a straightforward calculation to show that $\mathcal{M}^{-1}\mathcal{A}\mathcal{M} = \hat{\mathcal{A}}$ which has upperleft block of size $(p+q) \times (p+q)$ equal to $\hat{A}$.

Next we turn to the Lyapunov equation

$$P = \hat{A}P\hat{A}^T + ee^T, \tag{3.11}$$

because we see from proposition 3.3 that $E\varepsilon_t^\theta \varepsilon_t^{\theta T} = S(c, -a)PS(c, -a)^T$, with P the unique strictly positive solution of this Lyapunov equation, which exists since the pair $(\hat{A}, e)$ is controllable and $\hat{A}$ is stable. This solution is given (see [LR]) by

$$P = \frac{1}{2\pi i} \oint (z - \hat{A})^{-1} ee^T (\frac{1}{z} - \hat{A})^{-1} \frac{dz}{z}. \tag{3.12}$$

Because of the companion form of the matrix $\hat{A}$ we have $(z - \hat{A})^{-1}e = \frac{1}{A(z)C(z)} u_{p+q}^*(z)$. So we get from (2.4) that P is nothing else but the matrix P_θ from equation (2.3).

Notice also that the state process in (3.9) is equal to the ε^θ for an AR-process with AR-polynomial equal to $A^*(z)C^*(z)$, which gives an alternative explanation of (2.11).

The realization of the ε^θ process of proposition 3.3 is in an alternative way explained if we work with transfer functions. Consider first the system (3.5) and (3.6), and its transfer function ϕ. We had already computed the transfer function $(I - \mathcal{A}z)^{-1}e$ from ε to Z as $(((I - Az)^{-1}e)^T, (-I_{qn}(I -$

$Az)^{-1}e \otimes (I - Fz)^{-1}e)^T)^T$. Premultiplying it with C gives

$$\phi(z) = \begin{bmatrix} \frac{u_p(z)}{A^*(z)} \\ \frac{-u_q(z)}{C^*(z)} \end{bmatrix} \tag{3.13}$$

On the other hand the transfer function $\hat{C}(I - \hat{A}z)^{-1}e$ of the system of proposition 3.3 is computed as $S(c, -a)\frac{u_{p+q}(z)}{A^*(z)C^*(z)}$. Using again the relations (2.7) and (2.8) we get the $\phi(z)$ of (3.13) back.

We got the system in proposition 3.3 starting from the controllable realization of the ARMA process y. Alternatively we could have started from its observable realization. Following a similar procedure one then arrives at another system that realizes the ε^θ process as its state process. Specifically, we have

Proposition 3.4 *The process ε^θ is the state process of the stable system given by*

$$\varepsilon^\theta_{t+1} = \tilde{A}\varepsilon^\theta_t + B\varepsilon_t, \tag{3.14}$$

where $\tilde{A} = \begin{bmatrix} A_0 & 0 \\ 0 & F_0 \end{bmatrix}$, $A_0 = J - e[a_1, \ldots, a_p] \in \mathbf{R}^{p\times p}$, $F_0 = J - e[c_1, \ldots, c_q] \in \mathbf{R}^{q\times q}$ *and* $B = \begin{bmatrix} e_p \\ -e_q \end{bmatrix}$. *This system is controllable iff A and C have no common zeros. Moreover we have the relations $\tilde{A} = S(c, -a)\hat{A}$ with $\hat{A}$ as in proposition 3.3 and $B = S(c, -a)e$.*

It follows from proposition 3.4 that $F(\theta)$ is also the solution to the Lyapunov equation

$$F = \tilde{A}F\tilde{A}^T + BB^T. \tag{3.15}$$

Again stability of the matrix $\tilde{A}$, which obviously has characteristic polynomial $A(z)C(z)$, ensures that this equation has a nonnegative definite solution, which is strictly positive definite if the pair $(\tilde{A}, B)$ is controllable. Let $R(\tilde{A}, B)$ be the controllability matrix of this pair and $R(\hat{A}, e)$ the nonsingular controllability matrix of the pair $(\hat{A}, e)$. Then we also have the relation $R(\tilde{A}, B) = S(c, -a)R(\hat{A}, e)$. Hence the pair $R(\tilde{A}, B)$ is controllable iff $S(c, -a)$ is nonsingular iff A and C have no common zeros, which is –of course– in perfect agreement with the previous results. Finally, we again recognize the factorization $F(\theta) = S(c, -a)P(\theta)S(c, -a)^T$ of lemma 2.1, because of the relations mentioned in proposition 3.4 and the facts that $P(\theta)$ solves equation (3.11) and $F(\theta)$ solves equation (3.15).

4. Final remarks

In a sense we considered in this paper an identifiability problem for ARMA processes with the emphasis on the role of Sylvester's resultant matrix related to Fisher's information matrix. Some additional considerations in terms of state space realization were given. The factorization given in lemma 2.1 can be extended to a more complicated one for ARMAX processes. This is discussed in [KS], where algebraic properties of Sylvester's matrix are stressed. Again non-singularity of Fisher's information matrix follows from the absence of common zeros of the three polynomials involved, although this result can alternatively be proved via a direct extension of the analysis in [Po1], communicated to us in [Po2].

References

[AS] K.J. Åström and T. Söderström, Uniqueness of the Maximum Likelihood Estimates of the Parameters of an ARMA model, *IEEE Transactions on Automatic Control*, **AC-19** (1974), 769-773.

[Ba1] S. Barnett, A new formulation of the the theorems of Hurwitz, Routh and Sturm, *J. Inst. Maths. Applics*, **8** (1971), 240 - 250.

[Ba2] S. Barnett, Matrices, polynomials and linear time invariant systems, *IEEE Trans. Automat. Control*, **AC-18** (1973),1 - 10.

[Cr] H. Cramér, *Mathematical methods of statistics*, Princeton NJ; Princeton University Press, 1951.

[GW] K. Glover and J.C. Willems, Parametrizations of linear dynamical systems: canonical forms and identifiability, *IEEE Transactions on Automatic Control*, **AC-19** (1974), 640-646.

[Ka] R.E. Kalman, Mathematical description of linear dynamical systems, *SIAM J. Contr.* **1** (1963), 152 - 192.

[KM] A. Klein and G. Mélard, Fisher's information matrix for seasonal ARMA models, *J. Time Series*, **11** (1990), 231 - 237.

[Kl] A. Klein, Hypothesis Testing of Common Roots, *Statistics and Probability Letters*, **20** (1994), 163-167.

[KS] A. Klein and P.J.C. Spreij, On Fisher's Information Matrix of an ARMAX Process and Sylvester's Resultant Matrices, *Linear Algebra Appl.*, **237/238** (1996), 579-590.

[LR] P. Lancaster and L. Rodman, *Algebraic Riccati Equations*, Oxford Science Publications, (1995).

[McL] A.I. McLeod, A note on ARMA model redundancy, *J. of Time Series Analysis*, **14/2** (1993), 207-208.

[Po1] B.M. Pötscher, The behaviour of the Lagrangian multiplier test in testing the orders of an ARMA model, *Metrika*, **32** (1985), 129-150.

[Po2] B.M. Pötscher, *private communication.* (1996).

[Ra] C.R. Rao (1965), *Linear statistical inference and its applications*, Wiley & Sons New York, 1965.

[Ro] T.J. Rothenberg, Identification in parametric models, *Econometrica*, **39** (1971), 577-591.

[So] T. Söderström and P. Stoica (1989), *System identification*, Prentice hall.

[VW] B.L. van der Waerden (1966), *Algebra I*, Springer.

André Klein
Department of Actuarial Sciences,
Econometrics and Quantitative Methods
University of Amsterdam
Roetersstraat 11
1018 WB Amsterdam

Peter Spreij
Department of Econometrics
Vrije Universiteit
De Boelelaan 1105
1081 HV Amsterdam

Received September 28, 1996