JOURNAL OF MULTIVARIATE ANALYSIS 3, 408-419 (1973)

On the Parametrization of Autoregressive Models by Partial Autocorrelations

O. BARNDORFF-NIELSEN AND G. SCHOU

Department of Theoretical Statistics, Institute of Mathematics
University of Aarhus, 8000 Aarhus, Denmark

Communicated by G. Kallianpur

One of the difficulties that arise in the statistical analysis of autoregressive schemes is the very complex nature of the domain of the regression parameters. In the present paper we study an alternative parametrization of autoregressive models of finite order, namely the parametrization by the partial autocorrelations. These are shown to vary freely from -1 to $+1$ and to be in a one-to-one, continuously differentiable correspondence with the regression parameters. Properties of the asymptotic normal distribution of the maximum likelihood estimates are discussed, and we present a new deduction of Quenouille's result on the asymptotic independence of some of the estimated partial autocorrelations.

I. INTRODUCTION

Denote by $\{z_t\}$ an autoregressive process of finite order p, i.e., $\{z_t\}$ is the stationary solution to the stochastic difference equation

$$(z_t - \mu) - \varphi_1(z_{t-1} - \mu) - \cdots - \varphi_p(z_{t-p} - \mu) = a_t \qquad t = 0, \pm 1,.., \quad (1)$$

where $\{a_t\}$ denotes a set of independent and identically normally distributed random variables with mean value 0 and variance σ_a^2. $Ez_t = \mu$ and $\varphi_1, \ldots, \varphi_p$ are real constants. A necessary and sufficient condition for the stationarity of the solution of (1) is that no one of the roots of the characteristic polynomial

$$\varphi(z) = 1 - \varphi_1 z - \cdots - \varphi_p z^p, \qquad z \in \mathbb{C},$$

Received May 29, 1973.

AMS 1970 subject classification numbers: Primary 62M10; Secondary 62F99, 60G10, 60G15.

Keywords and phrases: Autoregressive processes, partial autocorrelations, variation independent parametrization, domain of regression parameters, asymptotic distribution of estimates, asymptotic independence of estimates, conditional correlations.

lies on the unit circle. Furthermore, we will require the z-process to be representable as a backward moving average of the a-process,

$$z_t = \sum_{j=0}^{\infty} \psi_j a_{t-j} + \mu \qquad \text{(convergence in square mean).}$$

It is well known that this is equivalent to assuming that $\boldsymbol{\varphi} = (\varphi_1, \ldots, \varphi_p)$ belongs to

$$\Phi_p = \{\boldsymbol{\varphi} \in \mathbb{R}^p \mid \boldsymbol{\varphi}(z) \neq 0, z \in \mathbb{C}, \mid z \mid \leqslant 1\},$$

(see e.g., Anderson [1]).

In general Φ_p is very complicated, but in the cases $p = 1, 2, 3$, it is possible to give rather simple criteria which are expressed directly in the φ's. Thus Φ_3 is determined by the equations

$$\begin{aligned} \varphi_1 + \varphi_2 + \varphi_3 &< 1, \\ -\varphi_1 + \varphi_2 - \varphi_3 &< 1, \\ \varphi_3(\varphi_3 - \varphi_1) - \varphi_2 &< 1, \\ \mid \varphi_3 \mid &< 1. \end{aligned} \tag{2}$$

In Wise [9] the root criterion is converted into a set of conditions on the φ's, well suited for practical purposes, but apparently without any simple statistical interpretation and difficult to handle in theoretical studies.

Now, let π_k denote k-th partial autocorrelation, i.e., the conditional correlation between z_t and z_{t+k} given the intervening z's, $z_{t+1}, \ldots, z_{t+k-1}$. Moreover, let Π be the mapping which transforms $\boldsymbol{\varphi}$ to $\boldsymbol{\pi} = (\pi_1, \ldots, \pi_p)$ and set $\Pi_p = \Pi(\Phi_p)$. ($\{z_t\}$ is a p-th order Markov process and hence $\pi_k = 0$ for $k > p$). It will be shown in Section III that Π is one-to-one and both ways continuously differentiable. Thus the class of p-th order autoregressive models may be smoothly parametrized by $\boldsymbol{\pi}$. This parametrization has the advantages that Π_p, the variation domain for $\boldsymbol{\pi}$, is the simple product set $]-1, 1[^p$ and that the asymptotic normal distribution of efficient estimates of $\boldsymbol{\pi}$ has some useful properties of independence and constant variance (confer, respectively, Sections III and IV). The functional expression for the likelihood function based on a sample $z_1, \ldots, z_n$ is, roughly speaking, equally complex in terms of $\boldsymbol{\pi}$ and in terms of $\boldsymbol{\varphi}$. For $\mu = 0$, $\sigma_a^2 = 1$ the expression has, respectively, the form

$$(1/2\pi)^{n/2}(1 - \pi_1^2)^{1/2}(1 - \pi_2^2) \cdots (1 - \pi_p^2)^{p/2}\, e^{-(1/2)P(\boldsymbol{\pi})}$$

where $P(\boldsymbol{\pi})$ is a polynomial in the π_k's, and the form

$$(1/2\pi)^{n/2} \mid \boldsymbol{\Gamma}_p \mid^{-1/2} e^{-(1/2)Q(\boldsymbol{\varphi})}$$

where the determinant $|\boldsymbol{\Gamma}_p|^{-1}$ and $Q(\boldsymbol{\varphi})$ are polynomials in the φ_k's, $Q(\boldsymbol{\varphi})$ being of the second degree. For example, with $p = 2$ we have

$$P(\boldsymbol{\pi}) = \sum_1^n z_i^2 + \pi_1^2(1-\pi_2)^2 \sum_2^{n-1} z_i^2 + \pi_2^2 \sum_3^{n-2} z_i^2 - 2\pi_1(1-\pi_2) \sum_2^n z_{i-1}z_i$$

$$- 2\pi_2 \sum_3^n z_{i-2}z_i + 2\pi_1\pi_2(1-\pi_2) \sum_3^{n-1} z_{i-1}z_i$$

and

$$|\boldsymbol{\Gamma}_2|^{-1} = (1+\varphi_2)^2[(1-\varphi_2)^2 - \varphi_1^2].$$

Thus, shifting from $\boldsymbol{\varphi}$ to $\boldsymbol{\pi}$ simplifies the factor in front of the exponential term but makes the exponent more complicated.

Partial autocorrelations play an important role in the methodology for time series analysis developed in recent years by Box and Jenkins [3].

II. PREREQUISITES

In addition to the quantities introduced in Section I let us consider the variance, covariance, and autocorrelations of $\{z_t\}$,

$$\begin{aligned} \gamma_0 &= E(z_t - \mu)^2, & & \\ \gamma_k &= E(z_t - \mu)(z_{t+k} - \mu), & k &= 1, 2,\ldots, \\ \rho_k &= \gamma_k/\gamma_0\,, & k &= 1, 2,\ldots, \end{aligned}$$

and, moreover, the auxiliary variables $\varphi_{k,j}$ defined by

$$\begin{bmatrix} 1 & \rho_1 & \cdots & \rho_{k-1} \\ \rho_1 & 1 & \cdots & \cdot \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{k-1} & \cdot & \cdots & 1 \end{bmatrix} \begin{bmatrix} \varphi_{k,1} \\ \varphi_{k,2} \\ \vdots \\ \varphi_{k,k} \end{bmatrix} = \begin{bmatrix} \rho_1 \\ \rho_2 \\ \vdots \\ \rho_k \end{bmatrix}, \qquad k = 1, 2,\ldots . \tag{3}$$

The coefficients $\varphi_1 ,\ldots, \varphi_p$ satisfy the Yule–Walker equations

$$\begin{bmatrix} 1 & \rho_1 & \cdots & \rho_{p-1} \\ \rho_1 & 1 & \cdots & \cdot \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{p-1} & \cdot & \cdots & 1 \end{bmatrix} \begin{bmatrix} \varphi_1 \\ \varphi_2 \\ \vdots \\ \varphi_p \end{bmatrix} = \begin{bmatrix} \rho_1 \\ \rho_2 \\ \vdots \\ \rho_p \end{bmatrix}$$

and, therefore,

$$\varphi_{p,j} = \varphi_j, \qquad j = 1,\dots, p.$$

Furthermore,

$$\varphi_{k,k} = \pi_k, \qquad k = 1, 2, \dots .$$

Hence (3) connects $\boldsymbol{\varphi}$ and $\boldsymbol{\pi}$. Particularly, we have

$$\pi_p = \varphi_p .$$

As noted by Durbin [6] the following recursive formulas hold:

$$\begin{aligned} \varphi_{k+1,j} &= \varphi_{k,j} - \varphi_{k+1,k+1}\varphi_{k,k+1-j}, \qquad j = 1,\dots, k \\ \varphi_{k+1,k+1} &= \left(\rho_{k+1} - \sum_{j=1}^{k} \varphi_{k,j}\rho_{k+1-j}\right)\Big/\left(1 - \sum_{j=1}^{k} \varphi_{k,j}\rho_j\right), \end{aligned} \tag{4}$$

and these will prove useful in the sequel.

Set

$$\mathbf{P}_k = \begin{bmatrix} 1 & \rho_1 & \cdots & \rho_{k-1} \\ \rho_1 & 1 & \cdots & \cdot \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{k-1} & \cdot & \cdots & 1 \end{bmatrix} \qquad \text{and} \qquad \boldsymbol{\Gamma}_k = \gamma_0 \mathbf{P}_k .$$

It seems rather well known, and may be proved by elementary means that

$$|\mathbf{P}_{k+1}| = (1 - \pi_1^2)^k (1 - \pi_2^2)^{k-1} \cdots (1 - \pi_k^2), \qquad k = 1, 2, \dots . \tag{5}$$

Using (4) we easily find

$$1 - \rho_1\varphi_{k,1} - \cdots - \rho_k\varphi_{k,k} = (1 - \pi_1^2)(1 - \pi_2^2) \cdots (1 - \pi_k^2), \qquad k = 1, 2, \dots, \tag{6}$$

and thus, in particular,

$$1 - \rho_1\varphi_1 - \cdots - \rho_p\varphi_p = (1 - \pi_1^2)(1 - \pi_2^2) \cdots (1 - \pi_p^2). \tag{7}$$

From (5), (7), and the equation

$$\gamma_0 = \sigma_a^2/(1 - \rho_1\varphi_1 - \cdots - \rho_p\varphi_p)$$

we finally obtain

$$\gamma_0 = \sigma_a^2(1 - \pi_1^2)^{-1} \cdots (1 - \pi_p^2)^{-1}. \tag{8}$$

To prove the relation $\Pi_p = \,]-1, 1[^p$, we need the following characterization (due to Duffin [5]) of Schur polynomials, which are the polynomials $f(z)$ with th property that $f(z) = 0$ implies $|z| < 1$.

CRITERION 1. *Let $f(z)$ be the polynomial*

$$f(z) = a_0 + a_1 z + \cdots + a_n z^n,$$

where $a_0 \neq 0$, $a_n \neq 0$ and $n \neq 0$. Let $\check{f}(z)$ be the reduced polynomial

$$\check{f}(z) = (\bar{a}_n a_1 - a_0 \bar{a}_{n-1}) + (\bar{a}_n a_2 - a_0 \bar{a}_{n-2}) z + \cdots + (\bar{a}_n a_n - a_0 \bar{a}_0) z^{n-1}$$

of degree $n - 1$. Then $f(z)$ is a Schur polynomial if and only if

$$|a_0| < |a_n|,$$

$\check{f}(z)$ is a Schur polynomial.

III. THE CORRESPONDENCE BETWEEN $\boldsymbol{\varphi}$ AND $\boldsymbol{\pi}$

Let Π denote the mapping of $\boldsymbol{\varphi} = (\varphi_1, \ldots, \varphi_p) \in \Phi_p$ into $\boldsymbol{\pi} = (\pi_1, \ldots, \pi_p) \in \Pi_p$ as it is defined in I. Since π_k is a partial correlation in a regular normal distribution, we have

$$\Pi_p \subseteq]-1, 1[^p.$$

From (4) we see that

$$\begin{aligned} \varphi_{p,j} &= \varphi_{p-1,j} - \varphi_{p,p}\varphi_{p-1,p-j} && (9) \\ &= \varphi_{p-2,j} - \varphi_{p-1,p-1}\varphi_{p-2,p-1-j} - \varphi_{p,p}\varphi_{p-1,p-j} \\ &= \ldots \\ &= \varphi_{j,j} - \varphi_{j+1,j+1}\varphi_{j,1} - \varphi_{j+2,j+2}\varphi_{j+1,2} - \cdots - \varphi_{p,p}\varphi_{p-1,p-j}, && (10) \\ & \qquad j = 1, \ldots, p-1. \end{aligned}$$

The right-hand side of (10) contains the π_k's and $\varphi_{l,m}$'s with $m < l < p$ and we can continue using (10), ending up with a polynomial in the π_k's. Equation (10) therefore defines a mapping $\tilde{\Phi}$ from $\mathbb{R}^p$ into $\mathbb{R}^p$, the restriction of which to Π_p is the inverse Φ of Π (recall that $\varphi_{p,j} = \varphi_j$).

For $p = 1, 2, 3$, and 4 the mapping Φ is given by

$p = 1$: $\varphi_1 = \pi_1$

$p = 2$: $\varphi_1 = \pi_1(1 - \pi_2)$
$\varphi_2 = \pi_2$

$$
\begin{aligned}
p = 3:\quad & \varphi_1 = \pi_1 - \pi_1\pi_2 - \pi_2\pi_3 \\
& \varphi_2 = \pi_2 - \pi_1\pi_3 + \pi_1\pi_2\pi_3 \\
& \varphi_3 = \pi_3 \\
p = 4:\quad & \varphi_1 = \pi_1 - \pi_1\pi_2 - \pi_2\pi_3 - \pi_3\pi_4 \\
& \varphi_2 = \pi_2 - \pi_1\pi_3 - \pi_2\pi_4 + \pi_1\pi_2\pi_3 + \pi_1\pi_3\pi_4 - \pi_1\pi_2\pi_3\pi_4 \\
& \varphi_3 = \pi_3 - \pi_1\pi_4 + \pi_1\pi_2\pi_4 + \pi_2\pi_3\pi_4 \\
& \varphi_4 = \pi_4 .
\end{aligned}
$$

THEOREM 2. *The mapping Π, which transforms $\boldsymbol{\varphi} = (\varphi_1 ,..., \varphi_p)$ to $\boldsymbol{\pi} = (\pi_1 ,..., \pi_p)$ is one-to-one and onto $]-1, 1[^p$. Furthermore, both Π and its inverse Φ are continuously differentiable.*

Proof. Consider the polynomial

$$\boldsymbol{\varphi}(z) = 1 - \varphi_1 z - \cdots - \varphi_p z^p.$$

It is easily seen that $\boldsymbol{\varphi} \in \Phi_p$ if and only if

$$\boldsymbol{\psi}(z) = -\varphi_p - \varphi_{p-1}z - \cdots - \varphi_1 z^{p-1} + z^p = z^p \boldsymbol{\varphi}(1/z)$$

is a Schur polynomial, and, according to Criterion 1, this is equivalent to the conditions

$$
\begin{aligned}
& |\varphi_p| < 1, \\
& \tilde{\boldsymbol{\psi}}(z) = (1 - \varphi_p{}^2)(\alpha_0 + \alpha_1 z + \cdots + \alpha_{p-2}z^{p-2} + z^{p-1}) \text{ is a Schur polynomial,}
\end{aligned}
$$

where

$$\alpha_i = (\varphi_{p-1-i} + \varphi_{i+1}\varphi_p)/(1 - \varphi_p{}^2), \qquad i = 0, 1,..., p-2.$$

Next, it will be proved that

$$\alpha_i = \varphi_{p-1,p-1-i} , \qquad i = 0, 1,..., p-2. \tag{11}$$

The equations

$$\varphi_{p,j} = \varphi_{p-1,j} - \varphi_{p,p}\varphi_{p-1,p-j} , \qquad j = 1,..., p-1 \tag{12}$$

determine, for $\varphi_{p,p}$ fixed, a transformation of $(\varphi_{p,1} ,..., \varphi_{p,p-1})$ to $(\varphi_{p-1,1} ,..., \varphi_{p-1,p-1})$ with the Jacobian matrix

$$
\mathbf{J}_{p-1} = \begin{bmatrix}
1 & 0 & 0 & \cdots & 0 & -\pi_p \\
0 & 1 & 0 & \cdots & -\pi_p & 0 \\
\vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\
-\pi_p & 0 & 0 & \cdots & 0 & 1
\end{bmatrix}
$$

(recall that $\pi_p = \varphi_{p,p}$) where the central element is $1 - \pi_p$ if p is even. Since the value of $|\mathbf{J}_{p-1}|$ is

$$(1 - \pi_p)^{[p/2]}(1 + \pi_p)^{[(p-1)/2]} = \begin{cases} (1 - \pi_p^2)^{(p-1)/2} & \text{if } p \text{ is odd,} \\ (1 - \pi_p^2)^{(p-2)/2}(1 - \pi_p) & \text{if } p \text{ is even,} \end{cases}$$

which is $\neq 0$ if and only if $\pi_p^2 \neq 1$, it suffices to prove that the α's satisfy (12):

$$\begin{aligned} \alpha_{p-1-j} - \varphi_p\alpha_{j-1} &= (1 - \varphi_p^2)^{-1}((\varphi_j + \varphi_{p-j}\varphi_p) - \varphi_p(\varphi_{p-j} + \varphi_j\varphi_p)) \\ &= \varphi_j . \end{aligned}$$

Thus (11) is established.

Observing that

$$\alpha_0 = \varphi_{p-1,p-1} = \pi_{p-1} ,$$

a repetitive use of Criterion 1 immediately shows that $\psi(z)$ is a Schur polynomial if and only if

$$|\pi_1| < 1 \wedge |\pi_2| < 1 \wedge \cdots \wedge |\pi_p| < 1$$

and thus

$$\Pi_p =]-1, 1[^p.$$

Since $\tilde{\Phi}$ (defined by (10)) is a polynomial in $\boldsymbol{\pi}$, $\tilde{\Phi}$ is continuously differentiable. $\Pi_p =]-1, 1[^p$ is open, $\Phi = \Pi^{-1}$ is the restriction of $\tilde{\Phi}$ to Π_p, and hence Π^{-1} is continuously differentiable. To prove that Π itself is continuously differentiable, it is therefore, by the inverse function theorem, sufficient to show that

$$\left| \frac{\partial \Phi}{\partial \boldsymbol{\pi}^*} \right| \neq 0, \qquad \boldsymbol{\pi} \in \Pi_p \tag{13}$$

($\boldsymbol{\pi}^*$ denotes the transposed of $\boldsymbol{\pi}$). As was pointed out by Daniels [4], the Jacobian (13) can be found by repetitive use of the transformation (12), yielding

$$\begin{aligned} |\partial\Phi/\partial\boldsymbol{\pi}^*| &= |\mathbf{J}_{p-1}| \cdot |\mathbf{J}_{p-2}| \cdot \cdots \cdot |\mathbf{J}_1| \\ &= \prod_{k=2}^{p} (1 - \pi_k)^{[k/2]}(1 + \pi_k)^{[(k-1)/2]} \neq 0, \qquad \boldsymbol{\pi} \in \Pi_p . \quad \blacksquare \end{aligned}$$

IV. Asymptotic Distribution of Estimates

Let $\hat{\boldsymbol{\varphi}}_n$ denote the maximum likelihood estimate of $\boldsymbol{\varphi} = (\varphi_1, \ldots, \varphi_p)$ based on a sample $z_1, \ldots, z_n$. It is well known that, as $n \to \infty$, $\hat{\boldsymbol{\varphi}}_n$ is asymptotically normally distributed with mean value $\boldsymbol{\varphi}$ and variance $(1/n)\,\mathbf{j}(\boldsymbol{\varphi})^{-1}$ where $\mathbf{j}(\boldsymbol{\varphi})$, the information matrix for $\boldsymbol{\varphi}$, is given by

$$\mathbf{j}(\boldsymbol{\varphi}) = (1/(1 - \varphi_1\rho_1 - \cdots - \varphi_p\rho_p))\,\mathbf{P}_p \tag{14}$$

(a detailed proof of this asymptotic result can be found in [2]). It follows, by Theorem 2, that the maximum likelihood estimate $\hat{\boldsymbol{\pi}}_n (= \Pi(\hat{\boldsymbol{\varphi}}_n))$ of $\boldsymbol{\pi}$ has an asymptotic normal distribution whose mean value is $\boldsymbol{\pi}$ and whose variance is $(1/n)\,\mathbf{i}(\boldsymbol{\pi})^{-1}$ where

$$\mathbf{i}(\boldsymbol{\pi}) = (\partial\Phi/\partial\boldsymbol{\pi}^*)\,\mathbf{j}(\Phi(\boldsymbol{\pi}))\,(\partial\Phi^*/\partial\boldsymbol{\pi}). \tag{15}$$

Two important properties of $\mathbf{i}(\boldsymbol{\pi})$ are described in Theorems 3 and 4 below.

Theorem 3. *The information matrix for $\boldsymbol{\pi}$ is of the form*

$$\mathbf{i}(\boldsymbol{\pi}) = \left[\begin{array}{c|c} & 0 \\ \mathbf{A}_{11} & \vdots \\ & 0 \\ \hline 0 \cdots 0 & (1 - \pi_p{}^2)^{-1} \end{array}\right]$$

where $\mathbf{A}_{11} = \mathbf{A}_{11}(\boldsymbol{\pi})$ denotes a $(p - 1) \times (p - 1)$ matrix.

Proof. In view of formulas (14) and (15) it is enough to prove the following results (a)–(c).

(a)
$$\frac{\partial\Phi^*}{\partial\boldsymbol{\pi}} = \begin{bmatrix} \mathbf{B}_{11} & \mathbf{B}_{12} \\ \mathbf{B}_{21} & \mathbf{B}_{22} \end{bmatrix}$$

where

$\mathbf{B}_{11}$ is $(p - 1) \times (p - 1)$

$\mathbf{B}_{12}$ and $\mathbf{B}_{21}^*$ are $(p - 1) \times 1$

$\mathbf{B}_{21} = (0, \ldots, 0)$

$\mathbf{B}_{22}$ is 1×1.

(b)
$$\frac{\partial\Phi}{\partial\boldsymbol{\pi}^*}\mathbf{P}_p = \begin{bmatrix} \mathbf{C}_{11} & \mathbf{C}_{12} \\ \mathbf{C}_{21} & \mathbf{C}_{22} \end{bmatrix}$$

where

$$\begin{aligned}&\mathbf{C}_{11} \text{ is } (p-1)\times(p-1)\\ &\mathbf{C}_{12} \text{ and } \mathbf{C}_{21}^{*} \text{ are } (p-1)\times 1\\ &\mathbf{C}_{21} = (0,\dots,0)\\ &\mathbf{C}_{22} = [c_{22}] \text{ is } 1\times 1.\end{aligned}$$

(c) $(1/(1-\varphi_1\rho_1-\cdots-\varphi_p\rho_p))(\mathbf{C}_{21}\mathbf{B}_{12}+\mathbf{C}_{22}\mathbf{B}_{22}) = (1-\pi_p{}^2)^{-1}$.

Re(a). The elements of $\mathbf{B}_{21}$ are

$$\frac{\partial\varphi_p}{\partial\pi_k} = \frac{\partial\pi_p}{\partial\pi_k} = 0, \qquad k = 1,\dots,p-1.$$

Re(b). From (10) we obtain

$$\partial\varphi_{p,j}/\partial\varphi_{p,p} = -\varphi_{p-1,p-j}\,, \qquad 1\leqslant j<p \tag{16}$$

and so we have to prove that

$$[-\varphi_{p-1,p-1}\,, -\varphi_{p-1,p-2}\,,\dots, -\varphi_{p-1,1}\,, 1]\mathbf{P}_p = [0,\dots,0,c_{22}], \tag{17}$$

but this is equivalent to (3) with $k = p-1$.

Re(c). Using (6), (7) and (17) we obtain (c). ▮

From (10) it is easy to find that

$$\left.\frac{\partial\varphi_{p,j}}{\partial\varphi_{k,k}}\right|_{\varphi_{p,p}=0} = \frac{\partial\varphi_{p-1,j}}{\partial\varphi_{k,k}}, \qquad k\leqslant p-1,\ \ j\leqslant p-1. \tag{18}$$

$\varphi_{p-1,1}\,,\dots,\varphi_{p-1,p-1}$ are the φ-parameters in the $(p-1)$-th order autoregression defined by $\pi_1\,,\dots,\pi_{p-1}$. We will denote by Φ^1 the mapping $\boldsymbol{\pi}\to\boldsymbol{\varphi}$ for this autoregression, and by $\boldsymbol{\pi}^1$ the vector $(\pi_1\,,\dots,\pi_{p-1})$. Using (6), (16), (17), and (18) we obtain

$$\begin{aligned}&\mathbf{i}(\pi_1\,,\dots,\pi_{p-1}\,,0)\\ &= \left.\left(\frac{\partial\Phi}{\partial\boldsymbol{\pi}^*}\,\mathbf{j}(\Phi(\boldsymbol{\pi}))\,\frac{\partial\Phi^*}{\partial\boldsymbol{\pi}}\right)\right|_{\pi_p=0}\\ &= \frac{1}{(1-\pi_1{}^2)\cdots(1-\pi_{p-1}^2)}\\ &\quad\times\left[\begin{array}{c:c} \dfrac{\partial\Phi^1}{\partial\boldsymbol{\pi}^{1*}} & \begin{matrix}0\\ \vdots\\ 0\end{matrix}\\ \hdashline -\varphi_{p-1,p-1}\cdots-\varphi_{p-1,1} & 1\end{array}\right]\left[\begin{array}{c:c} \mathbf{P}_{p-1} & \begin{matrix}\rho_{p-1}\\ \vdots\\ \rho_1\end{matrix}\\ \hdashline \rho_{p-1}\cdots\rho_1 & 1\end{array}\right]\left[\begin{array}{c:c} \dfrac{\partial\Phi^{1*}}{\partial\boldsymbol{\pi}^1} & \begin{matrix}-\varphi_{p-1,p-1}\\ \vdots\\ -\varphi_{p-1,1}\end{matrix}\\ \hdashline 0\cdots0 & 1\end{array}\right]\end{aligned}$$

$$= \frac{1}{(1-\pi_1{}^2)\cdots(1-\pi_{p-1}^2)}$$

$$\times \left[\begin{array}{c|c} \dfrac{\partial \Phi^1}{\partial \boldsymbol{\pi}^{1*}}\mathbf{P}_{p-1} & \vdots \\ \hline 0\cdots 0 & c_{22} \end{array}\right] \left[\begin{array}{c|c} \dfrac{\partial \Phi^{1*}}{\partial \boldsymbol{\pi}^1} & \begin{matrix} -\varphi_{p-1,p-1} \\ \vdots \\ -\varphi_{p-1,1} \end{matrix} \\ \hline 0\cdots 0 & 1 \end{array}\right]$$

$$= \left[\begin{array}{c|c} \mathbf{i}^1(\boldsymbol{\pi}^1) & \begin{matrix} 0 \\ \vdots \\ 0 \end{matrix} \\ \hline 0\cdots 0 & 1 \end{array}\right]$$

where $\mathbf{i}^1(\boldsymbol{\pi}^1)$ is the information matrix in the $(p-1)$-th order autoregression.

Repetition of this procedure proves the following theorem.

THEOREM 4. *To indicate the dependence of* $\mathbf{i}(\boldsymbol{\pi})$ *on* p, *the autoregressive order, we write* $\mathbf{i}_p(\boldsymbol{\pi})$. *With this notation we have for* $p < \tilde{p}$

$$\mathbf{i}_{\tilde{p}}(\pi_1, \ldots, \pi_p, 0, \ldots, 0) = \begin{bmatrix} \mathbf{i}_p(\pi_1, \ldots, \pi_p) & \mathbf{0} \\ \mathbf{0} & \mathbf{1} \end{bmatrix},$$

where $\mathbf{1}$ *is the unit matrix of order* $\tilde{p} - p$ *and* $\mathbf{0}$ *stands for a matrix consisting of zeroes.*

The theorem shows that if an autoregressive model of order $\tilde{p}$ is fitted to an observed autoregression which is actually of order $p < \tilde{p}$, then asymptotically the estimates $\hat{\pi}_{p+1}, \ldots, \hat{\pi}_{\tilde{p}}$ are independent and identically normally distributed with mean 0 and variance n^{-1}; furthermore $(\hat{\pi}_1, \ldots, \hat{\pi}_p)$ is asymptotically independent of $(\hat{\pi}_{p+1}, \ldots, \hat{\pi}_{\tilde{p}})$. These properties hold, of course, not only for the maximum likelihood estimates but for any asymptotically efficient estimates, such as, for instance, the often employed estimates which are obtained through estimating the autocorrelations ρ_k by

$$\check{\rho}_k = \left(\sum_{i=1}^{n-k}(z_i - \bar{z})(z_{i+k} - \bar{z})\right) \Big/ \sum_{i=1}^{n}(z_i - \bar{z})^2, \qquad k = 1, \ldots,$$

and then converting the $\check{\rho}_k$'s into estimates of the π_k's by means of (4).

The abovementioned result goes back to Quenouille [7, 8], but the derivation given here is different from Quenouille's.

We have not found a simple general expression for $\mathbf{i}(\boldsymbol{\pi})$. The formula for $\mathbf{i}_4(\boldsymbol{\pi})$ is written out below. From this $\mathbf{i}_p(\boldsymbol{\pi})$ for $p = 1, 2, 3$ may be easily obtained by setting the appropriate π_k's equal to zero (cf. Theorem 4).

$$
\begin{aligned}
&(1-\pi_1^2)(1-\pi_2^2)(1-\pi_3^2)(1-\pi_4^2)\,\mathbf{i}_4(\boldsymbol{\pi}) \\
&= \left[\begin{array}{c:c:c:c}
(1-\pi_2)^2\left\{\begin{matrix} 1-2\pi_1\pi_3+4\pi_1\pi_3\pi_4-2\pi_1^2\pi_4 \\ -2\pi_2\pi_4+2\pi_1^2\pi_2\pi_4-2\pi_1\pi_3\pi_4^2 \\ \pi_4^2+\pi_3^2+\pi_3^2\pi_4^2-2\pi_3^2\pi_4 \end{matrix}\right\} & 2(1-\pi_1^2)(1-\pi_2)\left\{\begin{matrix} -\pi_3-\pi_1\pi_4 \\ +\pi_1\pi_2\pi_4+\pi_2\pi_3\pi_4 \\ -\pi_3\pi_4^2+\pi_3\pi_4 \end{matrix}\right\} & -2\pi_4(1-\pi_1^2)(1-\pi_2^2)(1-\pi_2) & 0 \\
\hdashline
2(1-\pi_1^2)(1-\pi_2)\left\{\begin{matrix} -\pi_3-\pi_1\pi_4 \\ +\pi_1\pi_2\pi_4+\pi_2\pi_3\pi_4 \\ -\pi_3\pi_4^2+\pi_3\pi_4 \end{matrix}\right\} & (1-\pi_1^2)\left\{\begin{matrix} 2\pi_1\pi_3+2\pi_1^2\pi_4-2\pi_1^2\pi_2\pi_4 \\ -4\pi_1\pi_2\pi_3\pi_4+\pi_3^2-2\pi_2\pi_3^2\pi_4 \\ +1+\pi_4^2-2\pi_4-\pi_3^2\pi_4^2-2\pi_1\pi_3\pi_4^2 \end{matrix}\right\} & 2\pi_4(\pi_1+\pi_3)(1-\pi_1^2)(1-\pi_2^2) & 0 \\
\hdashline
-2\pi_4(1-\pi_1^2)(1-\pi_2^2)(1-\pi_2) & 2\pi_4(\pi_1+\pi_3)(1-\pi_1^2)(1-\pi_2^2) & (1-\pi_1^2)(1-\pi_2^2)(1+2\pi_2\pi_4+\pi_4^2) & 0 \\
\hdashline
0 & 0 & 0 & (1-\pi_1^2)(1-\pi_2^2)(1-\pi_3^2)
\end{array}\right]
\end{aligned}
$$

ACKNOWLEDGMENTS

We are very grateful to T. W. Anderson, P. Blæsild and F. Poulsen for several helpful discussions and to O. Østerby for drawing our attention to the work of Duffin [5].

REFERENCES

[1] ANDERSON, T. W. (1971). *The Statistical Analysis of Time Series*. Wiley, New York.

[2] BLÆSILD, P. (1972). Om autoregressive processer (in Danish). Statistiske Interna No. 18, Matematisk Institut, Aarhus Universitet.

[3] BOX, G. E. P. AND JENKINS, G. M. (1970). *Time Series Analysis, Forecasting and Control*. Holden Day, San Francisco, Calif.

[4] DANIELS, H. E. (1956). The approximate distribution of serial correlation coefficients. *Biometrika* **43** 169–185.

[5] DUFFIN, E. J. (1969). Algorithms for classical stability problems. *SIAM Review* **11** 196–213.

[6] DURBIN, J. (1960). The fitting of time series models. *Rev. Int. Stat. Inst.* **28** 233–244.

[7] QUENOUILLE, M. H. (1947). A large-sample test for the goodness of fit of autoregressive schemes. *J. Roy. Statist. Soc. A* **110** 123–129.

[8] QUENOUILLE, M. H. (1949). Approximate tests of correlation in time-series. *J. Roy. Statist. Soc. B* **11** 68–84.

[9] WISE, J. (1956). Stationarity conditions for stochastic processes of the autoregressive and moving-average type. *Biometrika* **43** 215–219.