The Annals of Statistics
1982, Vol. 10, No. 1, 121–131

THE EVALUATION OF CERTAIN QUADRATIC FORMS OCCURRING IN AUTOREGRESSIVE MODEL FITTING

BY R. J. BHANSALI

University of Liverpool

Let $\mathbf{R}$ be an infinite dimensional stationary covariance matrix, let $\mathbf{R}(k)$ and $\mathbf{W}(k)$ denote the top $k \times k$ left hand corners of $\mathbf{R}$ and $\mathbf{R}^{-1}$ respectively and let $\mathbf{\Sigma}(k)$ and $\mathbf{\Gamma}(k)$ denote the approximations for $\mathbf{R}(k)^{-1}$ suggested by Whittle (1951) and Shaman (1976) respectively. We consider quadratic forms of the type $Q(k) = \boldsymbol{\beta}(k)' \mathbf{R}(k)^{-1}\boldsymbol{\alpha}(k)$, when the vectors $\boldsymbol{\beta}(k)$ and $\boldsymbol{\alpha}(k)$ constitute the first k elements of the infinite absolutely summable sequences $\{\beta_j\}$ and $\{\alpha_j\}$. If $\chi_1(k) = \boldsymbol{\beta}(k)' \mathbf{W}(k)\, \boldsymbol{\alpha}(k)$ and $\chi_2(k) = \boldsymbol{\beta}(k)' \mathbf{\Sigma}(k)\, \boldsymbol{\alpha}(k)$, then, as $k \to \infty$, $Q(k)$ and $\chi_1(k)$ converge to the same limiting value for all such $\boldsymbol{\alpha}(k)$ and $\boldsymbol{\beta}(k)$, but $\chi_2(k)$ does not necessarily do so. Further, if $\tilde{\boldsymbol{\alpha}}(k) = (\alpha_k, \cdots, \alpha_1)'$ and $\tilde{\boldsymbol{\beta}}(k) = (\beta_k, \cdots, \beta_1)'$ then $\chi_1(k) \equiv \tilde{\boldsymbol{\beta}}(k)'\mathbf{\Gamma}(k)\, \tilde{\boldsymbol{\alpha}}(k)$. We discuss the use of $\mathbf{W}(k)$ for evaluating the asymptotic covariance structure of the autoregressive estimates of the inverse covariance function and the moving average parameters.

1. Introduction. Consider a second-order stationary process $\{x_t : t = 0, \pm 1, \cdots\}$ with zero mean, covariance function $R(u) = Ex_t x_{t+u}$ and the spectral density function

$$f(\lambda) = \frac{1}{2\pi} \sum_{u=-\infty}^{\infty} R(u) \exp(-iu\lambda).$$

We will assume that the covariance function is absolutely summable, i.e.

$$\sum_{u=-\infty}^{\infty} |R(u)| < \infty, \tag{1.1}$$

and that $f(\lambda)$ is non-vanishing, i.e.

$$f(\lambda) \neq 0, \qquad -\pi < \lambda < \pi. \tag{1.2}$$

Then x_t has (Brillinger, 1975, page 78) the infinite autoregressive representation

$$\sum_{u=0}^{\infty} a(u) x_{t-u} = \varepsilon_t, \qquad a(0) = 1, \tag{1.3}$$

in which ε_t is a sequence of uncorrelated random variables with 0 mean and finite variance, σ^2, and the $\{a(u)\}$ are real coefficients satisfying

$$\sum_{u=0}^{\infty} |a(u)| < \infty. \tag{1.4}$$

Following Parzen (1974), we will call $fi(\lambda) = (2\pi)^{-2}\{f(\lambda)\}^{-1}$ the inverse spectral density function, its Fourier coefficients

$$Ri(u) = \int_{-\pi}^{\pi} e^{iu\lambda} fi(\lambda)\, d\lambda$$

the inverse covariance function, and $ri(u) = Ri(u)/Ri(0)$ the inverse correlation function, of x_t.

Let $\mathbf{R} = [R(u - v)] (u, v = 1, 2, \cdots)$ denote the infinite dimensional covariance matrix of the semi-infinite vector $\mathbf{x}' = (x_{-1}, x_{-2}, \cdots)$ and $\mathbf{R}(k) = [R(u - v)]\ (u, v = 1, \cdots, k)$ be

Received August 1979; revised July 1981.

AMS 1970 subject classifications. Primary 62M20, secondary 60G10.

Key words and phrases. Stationary process, inverse of covariance matrix, convergence of a sequence of matrices, autoregressive model fitting, inverse covariance function, moving average process.

the corresponding $k \times k$ covariance matrix of the vector $\mathbf{x}(k)' = (x_{-1}, \cdots, x_{-k})$, and let $\mathbf{R}^{-1}$ and $\mathbf{R}(k)^{-1}$ respectively denote the inverses of $\mathbf{R}$ and $\mathbf{R}(k)$. An exact expression for $\mathbf{R}(k)^{-1}$, valid for each $k = 1, 2, \cdots$, is given by Berk (1974). However, in analytical work, a difficulty in using this exact expression is that it is not written directly in terms of the coefficients $a(u)$ of the autoregressive representation (1.3). However, let $\mathbf{W}(k)$ denote the submatrix in the top $k \times k$ left hand corner of $\mathbf{R}^{-1}$. An explicit expression in terms of $a(u)$ for $\mathbf{W}(k)$ and $\mathbf{R}^{-1}$ is given in Section 2, by using the results of Wise (1955). The main objective of this paper is to show that $\mathbf{W}(k)$ may be used as an approximation to $\mathbf{R}(k)^{-1}$ for evaluating the limiting value, as $k \to \infty$, of a class of quadratic forms of the general type

$$Q(k) = \boldsymbol{\beta}(k)' \mathbf{R}(k)^{-1} \boldsymbol{\alpha}(k) \tag{1.5}$$

where $\boldsymbol{\alpha}(k) = (\alpha_1, \cdots, \alpha_k)'$ and $\boldsymbol{\beta}(k) = (\beta_1, \cdots, \beta_k)'$, respectively, constitute the first k elements of infinite dimensional vectors $\boldsymbol{\alpha} = (\alpha_1, \alpha_2, \cdots)'$ and $\boldsymbol{\beta} = (\beta_1, \beta_2, \cdots)'$ and $\boldsymbol{\alpha}$, $\boldsymbol{\beta}$ belong to the Banach space, ℓ^1, of absolutely summable sequences, i.e. $\Sigma |\alpha_j| < \infty$ and $\Sigma |\beta_j| < \infty$.

Examples of applications where quadratic forms of this type arise are given in Section 4.

Let $\{\mathbf{R}_k(\infty)^{-1}\}$ and $\{\mathbf{W}_k(\infty)\}$ $(k = 1, 2, \cdots)$ denote two sequences of infinite dimensional matrices, obtained by placing the matrices $\mathbf{R}(k)^{-1}$ and $\mathbf{W}(k)$ in their top $k \times k$ left hand corners respectively, and 0's everywhere else. Also let $\chi_1(k) = \boldsymbol{\beta}(k)'\mathbf{W}(k)\boldsymbol{\alpha}(k)$ be the corresponding quadratic form obtained from $Q(k)$ by replacing $\mathbf{R}(k)^{-1}$ by $\mathbf{W}(k)$. We have

$$Q(k) = \boldsymbol{\beta}' \mathbf{R}_k(\infty)^{-1} \boldsymbol{\alpha}, \qquad \chi_1(k) = \boldsymbol{\beta}' \mathbf{W}_k(\infty) \boldsymbol{\alpha}. \tag{1.6}$$

As discussed later in Section 2, the question of whether $\chi_1(k)$ and $Q(k)$ converge to the same limiting value as $k \to \infty$ is closely related to the question: What is the mode of convergence of $\{\mathbf{W}_k(\infty)^{-1}\}$ to $\mathbf{R}^{-1}$? This question is examined in Section 3.

Two other approximations to $\mathbf{R}(k)^{-1}$ have been suggested previously in the literature: the first $\boldsymbol{\Sigma}(k)$, say, by Whittle (1951, 1952) and the second, $\boldsymbol{\Gamma}(k)$, say, by Shaman (1976). We examine the mode of convergence of the sequence of infinite dimensional matrices $\{\boldsymbol{\Sigma}_k(\infty)\}$, $(k = 1, 2, \cdots)$ to $\mathbf{R}^{-1}$, where $\boldsymbol{\Sigma}_k(\infty)$ has $\boldsymbol{\Sigma}(k)$ in its top $k \times k$ left hand corner and 0's everywhere else. Hence, we show that $Q(k)$ and the quadratic form, $\chi_2(k)$, say, obtained from $Q(k)$ by replacing $\mathbf{R}(k)^{-1}$ by $\boldsymbol{\Sigma}(k)$ do not necessarily converge to the same limiting value for all $\boldsymbol{\alpha}, \boldsymbol{\beta} \in \ell^1$. Hence $\mathbf{R}(k)^{-1}$ may not in general be replaced by $\boldsymbol{\Sigma}(k)$ for the evaluation of $Q(k)$, as $k \to \infty$.

Let $\mathbf{J}(k)$ be a $k \times k$ matrix with unity in the transverse diagonal and 0's everywhere else, and let $\tilde{\boldsymbol{\alpha}}(k) = \mathbf{J}(k)\boldsymbol{\alpha}(k) = (\alpha_k, \cdots, \alpha_1)'$, and $\tilde{\boldsymbol{\beta}}(k) = \mathbf{J}(k)\boldsymbol{\beta}(k) = (\beta_k, \cdots, \beta_1)'$ be $k \times 1$ vectors. From the results given later in Section 2, it follows that

$$\chi_1(k) = \tilde{\boldsymbol{\beta}}(k)' \boldsymbol{\Gamma}(k) \tilde{\boldsymbol{\alpha}}(k), \qquad k = 1, 2, \cdots. \tag{1.7}$$

Thus the limiting value of $Q(k)$ may also be evaluated by replacing $\mathbf{R}(k)^{-1}$ by $\boldsymbol{\Gamma}(k)$, provided the direction of the vectors $\boldsymbol{\alpha}(k)$ and $\boldsymbol{\beta}(k)$ is re-oriented.

A related reference is Huzii (1977) who, on the assumption that $a(u) = d^u$, $0 < d < 1/2$, obtains bounds for the error in approximating the elements of $\mathbf{R}(k)^{-1}$ by those of $\mathbf{W}(k)$, as $k \to \infty$.

2. Approximations to $\mathbf{R}(k)^{-1}$. Let $\boldsymbol{\varepsilon} = (\varepsilon_{-1}, \varepsilon_{-2}, \cdots)'$ denote a semi-infinite vector and $\mathbf{U} = [U_{nm}]$ $(n, m = 1, 2, \cdots)$ denote an infinite dimensional auxiliary matrix with 1's in the diagonal immediately below the main diagonal and 0's everywhere else; i.e., $U_{nm} = 1$, $n = m + 1$, and $U_{nm} = 0$, otherwise. By convention, we set $\mathbf{U}^0 = \mathbf{I}$, where $\mathbf{I} = [I_{nm}]$ $(n, m = 1, 2, \cdots)$, with $I_{nm} = 1$ $(n = m)$ and $I_{nm} = 0$ $(n \neq m)$, denotes an infinite dimensional identity matrix. The matrix $\mathbf{U}$ may also be employed for obtaining an explicit expression for $\mathbf{R}^{-1}$ in terms of $a(u)$. We have, as in Wise (1955)

$$[\textstyle\sum_{j=0}^{\infty} a(j)\mathbf{U}'^{j}]\mathbf{x} = \boldsymbol{\varepsilon}.$$

Hence,

$$\mathbf{R} = E[\mathbf{x}\mathbf{x}'] = \sigma^2[\sum_{j=0}^{\infty} a(j)\mathbf{U}'^j]^{-1}[\sum_{j=0}^{\infty} a(j)\mathbf{U}^j]^{-1} \tag{2.1}$$

and

$$\mathbf{R}^{-1} = \frac{1}{\sigma^2}[\sum_{j=0}^{\infty} a(j)\mathbf{U}^j][\sum_{j=0}^{\infty} a(j)\mathbf{U}'^j] = \frac{1}{\sigma^2}\mathbf{S}\mathbf{S}', \qquad \text{say,} \tag{2.2}$$

where $\mathbf{S}$ is an infinite dimensional lower triangular matrix

$$\mathbf{S} = \sum_{j=0}^{\infty} a(j)\mathbf{U}^j.$$

The inverse of the matrices appearing in (2.1) exist because (1.4) holds.

Since $\mathbf{R}(k)$ is the top $k \times k$ sub-matrix of $\mathbf{R}$, for sufficiently large k an approximation to $\mathbf{R}(k)^{-1}$ is provided by the corresponding matrix appearing in the upper left hand $k \times k$ corner of $\mathbf{R}^{-1}$. This gives

$$\mathbf{R}(k)^{-1} \approx \mathbf{W}(k), \tag{2.3}$$

where

$$\mathbf{W}(k) = \frac{1}{\sigma^2}\mathbf{S}(k)\mathbf{S}(k)', \qquad \mathbf{S}(k) = \sum_{j=0}^{k-1} a(j)\mathbf{L}^j, \tag{2.4}$$

and $\mathbf{L}$ is the k-dimensional analogue of $\mathbf{U}$; see, e.g., Anderson (1977).

The approximation to $\mathbf{R}(k)^{-1}$ suggested by Whittle (1951, 1952) is of the form

$$\mathbf{R}(k)^{-1} \approx \mathbf{\Sigma}(k), \tag{2.5}$$

where $\mathbf{\Sigma}(k)$ has $Ri(u - v)$ in its uth row and vth column. For a derivation of this approximation, see Anderson (1977) and Shaman (1975).

Hannan (1970, page 397) modified the approximation (2.5) by replacing the integral occurring in the definition of $Ri(u)$ by a finite sum. The corresponding approximation suggested for $\mathbf{R}(k)^{-1}$ is

$$\mathbf{R}(k)^{-1} \approx \phi(k) = \mathbf{P}(k)\mathbf{D}(k)^{-1}\mathbf{P}(k)^*, \tag{2.6}$$

where $\mathbf{P}(k)$ is a $k \times k$ unitary matrix with $k^{-1/2}\exp(iu(2\pi v)/k)$ in its uth row and vth column, $\mathbf{P}(k)^*$ is the conjugate transpose of $\mathbf{P}(k)$, and $\mathbf{D}(k)$ is a diagonal matrix with $2\pi f(2\pi j/k)$ in the jth place in the main diagonal.

To see the motivation for the approximation suggested by Shaman (1976), which is also implicit in Anderson (1975), define the random variables

$$\tilde{\mathbf{x}}_1 = \varepsilon_1, \qquad \tilde{\mathbf{x}}_t = -\sum_{u=1}^{t-1} a(u)\mathbf{x}_{t-u} + \varepsilon_t, \qquad t = 2, \cdots, k,$$

which are obtained from x_t by setting $x_p = 0$, if $p \le 0$. Also let $\xi(k) = (x_1, x_2, \cdots, x_k)'$, and $\boldsymbol{\varepsilon}(k) = (\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_k)'$ denote two $k \times 1$ vectors, and let $\mathbf{V}(k) = E[\xi(k)\xi(k)']$ denote the covariance matrix of $\xi(k)$. Then,

$$[\sum_{j=0}^{k-1} a(j)\mathbf{L}^j]\xi(k) = \boldsymbol{\varepsilon}(k), \qquad \mathbf{V}(k) = \sigma^2[\sum_{j=0}^{k-1} a(j)\mathbf{L}^j]^{-1}[\sum_{j=0}^{k-1} a(j)\mathbf{L}'^j]^{-1}$$

and

$$\mathbf{V}(k)^{-1} = \mathbf{\Gamma}(k) = \frac{1}{\sigma^2}[\sum_{j=0}^{k-1} a(j)\mathbf{L}'^j][\sum_{j=0}^{k-1} a(j)\mathbf{L}^j]. \tag{2.7}$$

For sufficiently large k, it may be supposed that the effect of setting the x's that have non-positive arguments equal to zero is small. This leads to the approximation

$$\mathbf{R}(k)^{-1} \approx \mathbf{\Gamma}(k). \tag{2.8}$$

On comparing (2.7) with (2.4), it will be noticed that $\mathbf{\Gamma}(k)$ is the "upside down" image of $\mathbf{W}(k)$, i.e.,

$$\Gamma(k) = \mathbf{J}(k)\mathbf{W}(k)\mathbf{J}(k). \tag{2.9}$$

Because $\mathbf{W}(k)$ and $\Gamma(k)$ are related to each other in such an elementary manner, in the sequel we do not explicitly distinguish between them as approximations for $\mathbf{R}(k)^{-1}$. Next, for the purpose of examining the efficacy of the approximations $\mathbf{W}(k)$ and $\boldsymbol{\Sigma}(k)$ to $\mathbf{R}(k)^{-1}$ for evaluating the limiting value of $Q(k)$ as $k \to \infty$, we bring together the definitions of four different modes of convergence of a sequence of infinite dimensional matrices $\{\mathbf{B}^{(n)}: n = 1, 2, \cdots\}$ to a fixed infinite dimensional matrix $\mathbf{B}$.

Let ℓ^1 be the Banach space of all vectors $\boldsymbol{\alpha}$ whose elements $\alpha_1, \alpha_2, \cdots$ are real or complex numbers and satisfy

$$\|\boldsymbol{\alpha}\| = \textstyle\sum_{j=1}^{\infty} |\alpha_j| < \infty. \tag{2.10}$$

The quantity $\|\boldsymbol{\alpha}\|$ is called the norm $\boldsymbol{\alpha}$.

Let $\gamma^{(n)} = \mathbf{B}^{(n)}\boldsymbol{\alpha}$ and $\gamma = \mathbf{B}\boldsymbol{\alpha}$. We will only be interested in the matrices defined on ℓ^1, i.e. in those matrices which transform a ℓ^1 vector into another ℓ^1 vector. Hence we assume that $\gamma^{(n)}, \gamma \in \ell^1$.

A matrix, $\mathbf{B}$, on ℓ^1 is said to be bounded if there is a constant C such that $\|\mathbf{B}\boldsymbol{\alpha}\| \le C\|\boldsymbol{\alpha}\|$, for all $\boldsymbol{\alpha} \in \ell^1$. Then, the quantity

$$\|\mathbf{B}\| = \sup_{\|\alpha\|=1} \|\mathbf{B}\boldsymbol{\alpha}\| \tag{2.11}$$

defines the corresponding norm of the matrix $\mathbf{B}$. A consequence of the definition (2.11) is that (e.g. Köthe, 1969, page 130)

$$\|\mathbf{B}\boldsymbol{\alpha}\| \le \|\mathbf{B}\|\|\boldsymbol{\alpha}\|, \tag{2.12}$$

and if $\mathbf{A}$ is another matrix on ℓ^1 then

$$\|\mathbf{AB}\| \le \|\mathbf{A}\|\|\mathbf{B}\|, \qquad \|\mathbf{A} + \mathbf{B}\| \le \|\mathbf{A}\| + \|\mathbf{B}\|.$$

Let b_{uv} denote the element in the uth row and the vth column of $\mathbf{B}$ and $b_{uv}^{(n)}$ that of $\mathbf{B}^{(n)}$. An explicit expression for $\|\mathbf{B}\|$ in terms of b_{uv} can also be obtained, provided that the sum to the right in the following definition (2.13) is finite. We have (e.g. Taylor, 1958, page 220)

$$\|\mathbf{B}\| = \sup_v \Sigma_{u=1}^{\infty} |b_{uv}|. \tag{2.13}$$

The Banach space, ℓ^∞, of all bounded sequences $\zeta = (\zeta_1, \zeta_2, \cdots)$, with norm

$$\|\zeta\|_\infty = \sup_k |\zeta_k|, \tag{2.14}$$

will also be needed. In the present context, the importance of ℓ^∞ arises from the fact that it is the "dual", or the "conjugate", space of ℓ^1. Thus, if $\Psi = \Psi(\alpha)$ denotes a linear functional on ℓ^1 then it can be shown that the space formed by Ψ is isometric to ℓ^∞ (see Taylor, 1958, page 194; Kothe, 1969, page 132).

The sequence $\{\mathbf{B}^{(n)}\}$ of matrices on ℓ^1 is said to converge to a fixed matrix $\mathbf{B}$ on ℓ^1(e.g., Riesz and Nagy, 1956) as follows.

(i) *uniformly* (*or in the norm*): if, as $n \to \infty$, $\|\mathbf{B}^{(n)} - \mathbf{B}\| \to 0$;
(ii) *strongly*: if, for every fixed $\boldsymbol{\alpha} \in \ell^1$, $\|\mathbf{B}^{(n)}\boldsymbol{\alpha} - \mathbf{B}\boldsymbol{\alpha}\| \to 0$, as $n \to \infty$;
(iii) *weakly*: if, for every fixed $\boldsymbol{\alpha} \in \ell^1$ and for every $\zeta \in \ell^\infty$, $|\zeta'\mathbf{B}^{(n)}\boldsymbol{\alpha} - \zeta'\mathbf{B}\boldsymbol{\alpha}| \to 0$, as $n \to \infty$;
(iv) *pointwise*: if, as $n \to \infty$, $b_{uv}^{(n)} \to b_{uv}$, for all u and v.

It is easy to verify that (i) $\Rightarrow$ (ii) $\Rightarrow$ (iii) $\Rightarrow$ (iv). However, the converse implications are not necessarily correct. But, (ii) and (iii) coincide on ℓ^1, i.e. $\{\mathbf{B}^{(n)}\}$ converges weakly to $\mathbf{B}$ on ℓ^1 if and only if it converges strongly; see Kothe (1969, page 281) and Taylor (1961, page 210).

For $\boldsymbol{\alpha}, \boldsymbol{\beta} \in \ell^1$, let $\psi^{(n)} = \boldsymbol{\beta}'\mathbf{B}^{(n)}\boldsymbol{\alpha}$ and $\psi = \boldsymbol{\beta}'\mathbf{B}\boldsymbol{\alpha}$. Then the requirement that $\psi^{(n)}$ converge to ψ as $n \to \infty$ for all $\boldsymbol{\alpha}, \boldsymbol{\beta} \in \ell^1$, may be termed *weak*(2) *convergence* of $\mathbf{B}^{(n)}$ to $\mathbf{B}$. Since

$\ell^1 \subset \ell^\infty$ the concept of weak(2) convergence is seen to be weaker than that of weak convergence. Indeed, by adapting the results of Akhiezer and Glazman (1961, pages 36–37, 44–45) to the present context, we may show that the sequence $\{\mathbf{B}^{(n)}\}$ converges to $\mathbf{B}$ in the weak(2) sense if and only if (i) the convergence is pointwise; and (ii) there is a constant M such that $\|\mathbf{B}^{(n)}\boldsymbol{\alpha}\|_\infty \le M$ for all n and all $\boldsymbol{\alpha} \in \ell^1$.

3. Strong convergence of $\mathbf{W}_k(\infty)$. We will assume that in equation (1.3) not all $a(u)(u = 1, 2, \cdots)$ are identically equal to zero, because if $x_t = \varepsilon_t$ then $\mathbf{R}(k) = \sigma^2\mathbf{I}(k)$, where $\mathbf{I}(k)$ denotes the k-dimensional identity matrix, and

$$\mathbf{R}(k)^{-1} = \sigma^{-2}\mathbf{I}(k) = \mathbf{W}(k) = \boldsymbol{\Sigma}(k) = \boldsymbol{\Gamma}(k).$$

Partition the matrix $\mathbf{S}$ appearing in equation (2.2) as

$$\mathbf{S} = \begin{bmatrix} \mathbf{S}(k) & \mathbf{0} \\ \mathbf{A}_{-k}(\infty) & \mathbf{E}_{-k}(\infty) \end{bmatrix}, \tag{3.1}$$

where $\mathbf{S}(k)$ is defined by equation (2.4), $\mathbf{E}_{-k}(\infty) = \mathbf{S}$ and $\mathbf{A}_{-k}(\infty)$ has $a(k+u-v)$ in its vth column and uth row $(v = 1, \cdots, k;\quad u = 1, 2, \cdots)$. Also partition the vector $\boldsymbol{\alpha} \in \ell^1$ as

$$\boldsymbol{\alpha}' = (\boldsymbol{\alpha}(k)', \quad \boldsymbol{\alpha}_{-k}(\infty)'), \tag{3.2}$$

where $\boldsymbol{\alpha}_{-k}(\infty)' = (\alpha_{k+1}, \alpha_{k+2}, \cdots)$. Given $\boldsymbol{\alpha}$, define the vectors $\mathbf{h} = (h_1, h_2, \cdots)'$ and $\mathbf{h}^*(k) = (h_1^*, \cdots, h_k^*)'$ by

$$\mathbf{h} = \mathbf{R}^{-1}\boldsymbol{\alpha}, \tag{3.3}$$

$$\mathbf{h}^*(k) = \mathbf{R}(k)^{-1}\boldsymbol{\alpha}(k) \tag{3.4}$$

and partition $\mathbf{h}$ in the same way as $\boldsymbol{\alpha}$. Thus, let

$$\mathbf{h}' = (\mathbf{h}(k)', \quad \mathbf{h}_{-k}(\infty)'), \tag{3.5}$$

where $\mathbf{h}(k) = (h_1, \cdots, h_k)'$ and $\mathbf{h}_{-k}(\infty) = (h_{k+1}, h_{k+2}, \cdots)'$. We have

$$\mathbf{h}(k) = \mathbf{W}(k)\boldsymbol{\alpha}(k) + \frac{1}{\sigma^2}\mathbf{S}(k)\mathbf{A}_{-k}(\infty)'\boldsymbol{\alpha}_{-k}(\infty), \tag{3.6}$$

$$\mathbf{h}_{-k}(\infty) = \frac{1}{\sigma^2}\{\mathbf{A}_{-k}(\infty)\mathbf{S}(k)'\boldsymbol{\alpha}(k) + \mathbf{A}_{-k}(\infty)\mathbf{A}_{-k}(\infty)'\boldsymbol{\alpha}_{-k}(\infty)\} + \mathbf{R}^{-1}\boldsymbol{\alpha}_{-k}(\infty). \tag{3.7}$$

We note that

$$\|\mathbf{R}^{-1}\| \le \frac{1}{\sigma^2}\{\textstyle\sum_{u=0}^{\infty} |a(u)|\}^2 < \infty.$$

Hence, on using (2.12), we have

$$\|\mathbf{h}\| \le \|\mathbf{R}^{-1}\|\,\|\boldsymbol{\alpha}\| < \infty.$$

A useful inequality due to Baxter (1963) is

$$\|\mathbf{h}(k) - \mathbf{h}^*(k)\| \le \text{const}\,\|\mathbf{h}_{-k}(\infty)\|. \tag{3.8}$$

We also have the following lemma

LEMMA 1. *If* $\boldsymbol{\alpha} \in \ell^1$ *and* $\mathbf{h} = \mathbf{R}^{-1}\boldsymbol{\alpha}$ *is partitioned as shown in equation* (3.5) *then*

$$\lim_{k\to\infty} \|\mathbf{h}_{-k}(\infty)\| = 0. \tag{3.9}$$

PROOF. This follows immediately from the result that $\|\mathbf{h}\| < \infty$.

Theorem 1 given below establishes the strong convergence of $\{\mathbf{R}_k(\infty)^{-1}: k = 1, 2, \cdots\}$ to $\mathbf{R}^{-1}$ as $k \to \infty$.

THEOREM 1. *For every* $\boldsymbol{\alpha} \in \ell^1$, $\lim_{k\to\infty} \| \mathbf{R}_k(\infty)^{-1}\boldsymbol{\alpha} - \mathbf{R}^{-1}\boldsymbol{\alpha} \| = 0$.

PROOF. By direct computation

$$\| \mathbf{R}_k(\infty)^{-1}\boldsymbol{\alpha} - \mathbf{R}^{-1}\boldsymbol{\alpha} \| = \| \mathbf{h}^*(k) - \mathbf{h}(k) \| + \| \mathbf{h}_{-k}(\infty) \|.$$

The result follows by using the inequality (3.8) and Lemma 1.

The following theorem provides a slightly different interpretation of the manner in which $\mathbf{W}(k)$ converges to $\mathbf{R}(k)^{-1}$ as $k \to \infty$.

THEOREM 2. *Let* $\boldsymbol{\alpha} \in \ell^1$ *be partitioned as shown in equation* (3.2). *Then for every such* $\boldsymbol{\alpha}$

$$\lim_{k\to\infty} \| \mathbf{W}(k)\boldsymbol{\alpha}(k) - \mathbf{R}(k)^{-1}\boldsymbol{\alpha}(k) \| = 0. \tag{3.10}$$

PROOF. We have, using (3.4),

$$\| \mathbf{W}(k)\boldsymbol{\alpha}(k) - \mathbf{R}(k)^{-1}\boldsymbol{\alpha}(k) \| \le \| \mathbf{W}(k)\boldsymbol{\alpha}(k) - \mathbf{h}(k) \| + \| \mathbf{h}(k) - \mathbf{h}^*(k) \|. \tag{3.11}$$

The second term to the right in the inequality (3.11) tends to 0 as $k \to \infty$ by inequality (3.8) and Lemma 1, while the first term is bounded above by

$$\frac{1}{\sigma^2} \| \mathbf{S}(k) \| \| \mathbf{A}_{-k}(\infty)' \| \| \boldsymbol{\alpha}_{-k}(\infty) \|,$$

and, as $k \to \infty$, it converges to 0 because $\| \boldsymbol{\alpha}_{-k}(\infty) \| \to 0$, while σ^{-2}, $\| \mathbf{A}_{-k}(\infty)' \|$ and $\| \mathbf{S}(k) \|$ remain bounded.

The strong convergence of $[\mathbf{W}_k(\infty) : k = 1, 2, \cdots\}$ to $\mathbf{R}^{-1}$ as $k \to \infty$ is established in Theorem 3.

THEOREM 3. *For every* $\boldsymbol{\alpha} \in \ell^1$, $\lim_{k\to\infty} \| \mathbf{W}_k(\infty)\boldsymbol{\alpha} - \mathbf{R}^{-1}\boldsymbol{\alpha} \| = 0$.

PROOF.

$$\| \mathbf{W}_k(\infty)\boldsymbol{\alpha} - \mathbf{R}^{-1}\boldsymbol{\alpha} \| \le \| \mathbf{W}_k(\infty)\boldsymbol{\alpha} - \mathbf{R}_k(\infty)^{-1}\boldsymbol{\alpha} \| + \| \mathbf{R}_k(\infty)^{-1}\boldsymbol{\alpha} - \mathbf{R}^{-1}\boldsymbol{\alpha} \|.$$

The second term and the first term to the right of the above inequality converge to zero by Theorem 1 and Theorem 2, respectively.

COROLLARY 3.1. *Let* $\boldsymbol{\zeta} = (\zeta_i, \zeta_2, \cdots)' \in \ell^\infty$ *be partitioned as*

$$\boldsymbol{\zeta}' = [\boldsymbol{\zeta}(k)' : \boldsymbol{\zeta}_{-k}(\infty)'] \tag{3.12}$$

where $\boldsymbol{\zeta}(k) = (\zeta_1, \cdots, \zeta_k)'$ *and* $\boldsymbol{\zeta}_{-k}(\infty) = (\zeta_{k+1}, \zeta_{k+2}, \cdots)'$, *and let* $\boldsymbol{\alpha} \in \ell^1$ *be as in equation* (3.2). *Then for every such* $\boldsymbol{\alpha}$ *and* $\boldsymbol{\zeta}$

$$\lim_{k\to\infty} | \boldsymbol{\zeta}(k)'\mathbf{R}(k)^{-1}\boldsymbol{\alpha}(k) - \boldsymbol{\zeta}(k)'\mathbf{W}(k)\boldsymbol{\alpha}(k) | = 0. \tag{3.13}$$

PROOF. Let $S_{1k} = | \boldsymbol{\zeta}'\{\mathbf{R}_k(\infty)^{-1} - \mathbf{R}^{-1}\}\boldsymbol{\alpha} |$ and $S_{2k} = | \boldsymbol{\zeta}'\{\mathbf{W}_k(\infty) - \mathbf{R}^{-1}\}\boldsymbol{\alpha} |$. Then

$$| \boldsymbol{\zeta}(k)'\{\mathbf{R}(k)^{-1} - \mathbf{W}(k)\}\boldsymbol{\alpha}(k) | \le S_{1k} + S_{2k}.$$

Now, as $k \to \infty$, $S_{1k} \to 0$ by Theorem 1 and $S_{2k} \to 0$ by Theorem 3. Hence the result follows.

Theorem 3 may also be extended to the case when the elements α_j of $\boldsymbol{\alpha}$ are not fixed but depend upon k. Suppose that $\{\boldsymbol{\alpha}_k : k = 1, 2, \cdots\}$ is a sequence of vectors in ℓ^1 and $\boldsymbol{\alpha} \in \ell^1$ is a fixed vector such that $\| \boldsymbol{\alpha}_k - \boldsymbol{\alpha} \| \to 0$ as $k \to \infty$. It immediately follows from Theorem 3 that $\| \mathbf{W}_k(\infty)\boldsymbol{\alpha}_k - \mathbf{R}^{-1}\boldsymbol{\alpha} \| \to 0$ as $k \to \infty$. Analogously, Corollary 3.1 may also be extended to this case.

Consider now the mode of convergence of $\mathbf{\Sigma}_k(\infty)$ to $\mathbf{R}^{-1}$ as $k \to \infty$. Hannan (1970, page 378) has shown that if the terms of $O(k^{-1})$ are ignored then the matrix $\mathbf{P}(k)^*\mathbf{R}(k)\mathbf{P}(k)$ is diagonal with $2\pi f(2\pi j/k)$ in the jth place in the main diagonal. However, as discussed below, this result does not imply that as $k \to \infty$ the difference between the elements in the uth row and the vth column of $\mathbf{\Sigma}(k)$ and $\mathbf{R}(k)^{-1}$ converges to 0, for all u and v.

A stronger result on the nature of approximation provided by $\mathbf{\Sigma}(k)$ for $\mathbf{R}(k)^{-1}$, as $k \to \infty$, is stated by Whittle (1952), who has claimed that the approximation (2.5) is to be understood as implying

$$\lim_{k\to\infty} \frac{\mathbf{x}(k)'\mathbf{R}(k)^{-1}\mathbf{x}(k)}{\mathbf{x}(k)'\mathbf{\Sigma}(k)\mathbf{x}(k)} = 1 \tag{3.14}$$

almost certainly, when the elements of $\mathbf{x}(k)$ constitute an arbitrary, purely non-deterministic stationary process. A similar result, with a minor change of wording, is also claimed to be true by Wold (1953, Chapter 11), though neither of these two authors gives a proof in support of this claim. Nevertheless, a partial justification of this claim can be obtained from the results of Shaman (1975).

If x_t is a Gaussian autoregressive process, or a Gaussian moving average process, of order q where q is finite, and the ratio on the left hand side of (3.14) is denoted by τ, then Shaman has shown that as $k \to \infty$, the random variable $k(\tau - 1)$ is distributed like a weighted sum of q independent χ_2^2 variates, where the weights multiplying the χ_2^2's are all negative. This result implies that $\tau = 1 + O_p(k^{-1})$ and, as $k \to \infty$, $\tau \to 1$ in probability, for at least these two important special classes of stationary, purely non-deterministic processes.

We will now show that the sequence $\{\mathbf{\Sigma}_k(\infty)\}$ does not even converge pointwise to $\mathbf{R}^{-1}$ as $k \to \infty$ and hence we show that a result similar to (3.14) is not true if the components of $\mathbf{x}(k)$ are real or complex numbers such that $\mathbf{x}(k)$ constitutes the first k elements of a vector $\mathbf{x} \in \ell^1$.

THEOREM 4. *The sequence* $\{\mathbf{\Sigma}_k(\infty) : k = 1, \cdots\}$ *does not converge pointwise to* $\mathbf{R}^{-1}$ *as* $k \to \infty$.

PROOF. Let $\mathbf{e} = (1, 0, 0, \cdots)'$ be an infinite dimensional vector. Then it will be sufficient to show that as $k \to \infty$, $|\mathbf{e}'\{\mathbf{\Sigma}_k(\infty) - \mathbf{R}^{-1}\}\mathbf{e}| \nrightarrow 0$. We have, when all the $a(j)$ are not equal to 0,

$$|\mathbf{e}'\{\mathbf{\Sigma}_k(\infty) - \mathbf{R}^{-1}\}\mathbf{e}| = \textstyle\sum_{j=1}^{\infty} a^2(j) > 0$$

for all k; hence the result follows.

THEOREM 5. *Let* $\boldsymbol{\alpha} \in \ell^1$ *be partitioned as shown in equation* (3.2). *Then there exists an integer* K_0, *a positive constant* $M(k)$ *and at least one choice of* $\boldsymbol{\alpha}$ *such that for all* $k > K_0$

$$|\boldsymbol{\alpha}(k)'\{\mathbf{\Sigma}(k) - \mathbf{R}(k)^{-1}\}\boldsymbol{\alpha}(k)| \geq M(k) > 0. \tag{3.15}$$

PROOF. Let $N_{1k} = |\boldsymbol{\alpha}'\{\mathbf{\Sigma}_k(\infty) - \mathbf{R}^{-1}\}\boldsymbol{\alpha}|$ and $N_{2k} = |\alpha'\{\mathbf{R}_k(\infty) - \mathbf{R}^{-1}\}\alpha|$. Then

$$|\boldsymbol{\alpha}(k)'\{\mathbf{\Sigma}(k) - \mathbf{R}(k)^{-1}\}\boldsymbol{\alpha}(k)| \geq N_{1k} - N_{2k}.$$

Theorem 1 shows that for all $\boldsymbol{\alpha} \in \ell^1$, $N_{2k} \to 0$ as $k \to \infty$. Hence given $\epsilon > 0$ we may find an integer K_1 such that $N_{2k} \leq \epsilon$ for all $k > K_1$. On the other hand, if $\boldsymbol{\alpha} = \mathbf{e}$ then N_{1k} is bounded away from 0 as $k \to \infty$ by Theorem 7. Since ϵ is arbitrary, we may choose ϵ and K_2 such that $N_{1k} > \epsilon$ for all $k > K_2$, and then set $M(k) = N_{1k} - \epsilon$ and $K_0 = \max(K_1, K_2)$.

However, it should be emphasized that although the matrix sequence $\{\mathbf{W}_k(\infty)\}$ converges strongly to $\mathbf{R}^{-1}$, it does not do so uniformly, i.e. $\|\mathbf{W}_k(\infty) - \mathbf{R}^{-1}\| \nrightarrow 0$ as $k \to \infty$. To illustrate this last point and some of the earlier results, consider the special case of a first order autoregressive process

$$x_t + a\, x_{t-1} = \epsilon_t, \qquad |a| < 1. \tag{3.16}$$

We have

$$\mathbf{R}(k)^{-1} = \begin{bmatrix} 1 & a & 0 & \cdots & 0 & 0 \\ a & 1+a^2 & a & \cdots & 0 & 0 \\ 0 & a & 1+a^2 & \cdots & 0 & 0 \\ \vdots & \vdots & & & \vdots & \vdots \\ 0 & 0 & & \cdots & 1+a^2 & a \\ 0 & 0 & & \cdots & a & 1 \end{bmatrix}$$

The matrices $\mathbf{W}(k)$, $\mathbf{\Sigma}(k)$ and $\mathbf{\Gamma}(k)$ are of the same form as $\mathbf{R}(k)^{-1}$; but $\mathbf{W}(k)$ has $1 + a^2$ in its kth diagonal element, $\mathbf{\Gamma}(k)$ has $1 + a^2$ in its 1st diagonal element and $\mathbf{\Sigma}(k)$ has $1 + a^2$ in its 1st and kth diagonal elements. We have

$$\|\mathbf{W}(k) - \mathbf{R}(k)^{-1}\| = \|\mathbf{\Gamma}(k) - \mathbf{R}(k)^{-1}\| = \|\mathbf{\Sigma}(k) - \mathbf{R}(k)^{-1}\| = a^2.$$

Also, using (2.2), we get

$$\mathbf{R}^{-1} = \frac{1}{\sigma^2}(\mathbf{I} + a\mathbf{U})(\mathbf{I} + a\mathbf{U}')$$

and it is easy to verify that, as $k \to \infty$,

$$\|\mathbf{W}_k(\infty) - \mathbf{R}^{-1}\| = (1 + |a|)^2 \not\to 0,$$

though

$$\|\mathbf{W}_k(\infty)\boldsymbol{\alpha} - \mathbf{R}^{-1}\boldsymbol{\alpha}\| = |a\alpha_{k+1}| + \textstyle\sum_{j=k}^{\infty} |a(\alpha_j + \alpha_{j+2}) + (1 + a^2)\alpha_{j+1}|$$

does tend to zero for all $\boldsymbol{\alpha} \in \ell^1$.

Also, if $\mathbf{e}(k) = [1, 0, \cdots 0]'$ is a $k \times 1$ vector then for $k = 1, 2, \cdots$

$$|\mathbf{e}(k)'\{\mathbf{\Sigma}(k) - \mathbf{R}(k)^{-1}\}\mathbf{e}(k)| = a^2 > 0, \qquad |\mathbf{e}(k)'\{\mathbf{\Gamma}(k) - \mathbf{R}(k)^{-1}\}\mathbf{e}(k)| = a^2 > 0$$

and so the difference between the element in the 1st row and 1st column of $\mathbf{\Sigma}(k)$ and $\mathbf{R}(k)^{-1}$ does not vanish as $k \to \infty$. The same holds also for $\mathbf{\Gamma}(k)$ and $\mathbf{R}(k)^{-1}$.

4. Applications. The applications described next relate to the case when T consecutive observations $X_1, \cdots, X_T$ of x_t are available and kth order autoregressive estimates $C_1 = \hat{a}_{k1}, \cdots, C_k = \hat{a}_{kk}$ are obtained by minimizing

$$(T-k)^{-1} \textstyle\sum_{j=0}^{T-1-k} (X_{k+j+1} + C_1 X_{k+j} + \cdots + C_k X_{1+j})^2$$

with minimum $\hat{\sigma}^2(k)$. An autoregressive estimate of the spectral density function is given by

$$\hat{f}_k(\lambda) = \frac{\hat{\sigma}^2(k)}{2\pi} |1 + \textstyle\sum_{u=1}^{k} \hat{a}_{ku} e^{-iu\lambda}|^{-2}, \qquad -\pi \le \lambda \le \pi; \tag{4.1}$$

and that of the inverse covariance and correlation functions are given by

$$\hat{R}i_k(u) = \frac{1}{2\pi Q} \sum_{j=0}^{Q-1} \left\{ \hat{f}_k\left(\frac{2\pi j}{Q}\right) \right\}^{-1} \exp\left(iu\frac{2\pi d}{Q}\right), \tag{4.2}$$

$$\hat{r}i_k(u) = \hat{R}i_k(u)/\hat{R}i_k(0), \qquad u = 1, \cdots, k, \tag{4.3}$$

where $Q \ge 2k$ is an integer suitable for applying a Fast Fourier Transform algorithm.

For an exposition of the reasons for taking interest in $ri(u)$, see Cleveland (1972), Parzen

(1974) and Bhansali (1980). In particular, the $\hat{r}i_k(u)$ may be employed for estimating the order of a finite moving average process. Thus, suppose that x_t is a moving average process of order q defined by

$$x_t = \textstyle\sum_{j=0}^{q} \nu(j)\epsilon_{t-j}, \qquad \nu(0) = 1, \tag{4.4}$$

where ϵ_t is a sequence of independent, identically distributed random variables with 0 mean, variance σ^2 and finite fourth cumulant κ_4 and the $\nu(j)$ are real parameters such that the polynomial $\nu_q(z) = 1 + \nu(1)z + \cdots + \nu(q)z^q$ is bounded away from 0, $|z| \leq 1$. If q is known, the method suggested by Durbin (1959) for estimating the $\nu(j)$ may be viewed (see Bhansali, 1980) as providing the estimates $\hat{\nu}_k(1), \cdots, \hat{\nu}_k(q)$ by solving the equations

$$\textstyle\sum_{j=0}^{q} \hat{\nu}_k(j)\hat{r}i_k(u-j) = 0, \qquad u = 1, \cdots, q, \tag{4.5}$$

with $\hat{\nu}_k(0) = 1$. But q is invariably unknown. The equations (4.5) may be solved recursively and an estimate of q may be obtained by using an order determination criterion, such as Akaike's information criterion. Although the final estimates of the moving average parameters obtained by solving the equations (4.5) do not compare favourably with those obtained by maximizing the Gaussian likelihood, a "two-step" procedure recommended by Bhansali (1980) may be employed.

In the following applications $k = k(T)$ is a function of T such that as $T \to \infty$, $k \to \infty$ but $k^3/T \to 0$.

EXAMPLE 1. For evaluating the covariance term in the joint asymptotic distribution of $\sqrt{T}\{\hat{R}i_k(u_1) - Ri(u_1)\}, \cdots, \sqrt{T}\{\hat{R}i_k(u_G) - Ri(u_G)\}$, when $u_1, \cdots, u_G$ are fixed non-negative integers, it is necessary to evaluate (Bhansali, 1980) the limiting value as $k \to \infty$ of the quantity

$$Q_1(k) = q_{11}(k) + q_{12}(k) + q_{21}(k) + q_{22}(k) + p(k),$$

where

$$q_{ij}(k) = \sigma^{-2}\gamma_i(k)'\,\mathbf{R}(k)^{-1}\boldsymbol{\delta}_j(k)(i,j = 1, 2),$$
$$\lim_{k\to\infty} p(k) = 2\,Ri(u)Ri(v) + \frac{\kappa_4}{\sigma^4}\,Ri(u)Ri(v),$$

u and v are fixed non-negative integers, $\gamma_1(k)' = (a(u+1), \cdots, a(u+k))$, $\boldsymbol{\delta}_1(k)' = (a(v+1), \cdots, a(v+k))$, $\gamma_2(k)' = (0, \cdots, 0, 1, a(1), \cdots a(k-u))$ and $\boldsymbol{\delta}_2(k)' = (0, \cdots, 0, 1, \cdots, a(k-v))$ are all $k \times 1$ vectors.

Consider $q_{11}(k)$. We have, using Corollary 3.1,

$$\lim_{k\to\infty} q_{11}(k) = \lim_{k\to\infty} \sigma^{-2}\gamma_1(k)'\,\mathbf{W}(k)\boldsymbol{\delta}_1(k)$$
$$= \lim_{k\to\infty}[\{\textstyle\sum_{j=1}^{k} Ri(u+j)Ri(v+j)\} - \phi_k(j)],$$

where

$$\phi_k(j) = \textstyle\sum_{j=1}^{k}\{\eta_1(j) + \eta_2(j) - \eta_3(j)\},$$
$$\eta_1(j) = \sigma^{-2}\textstyle\sum_{s=k-j+1}^{\infty} a(s)a(u+j+s)Ri(v+j),$$
$$\eta_2(j) = \sigma^{-2}\textstyle\sum_{s=k-j+1}^{\infty} a(s)a(v+j+s)Ri(u+j),$$
$$\eta_3(j) = \sigma^{-4}\textstyle\sum_{s=k-j+1}^{\infty}\sum_{t=k-j+1}^{\infty} a(s)a(u+j+s)a(t)a(v+j+t).$$

Now, the $R(u)$ and the $a(u)$ are absolutely summable. Hence, by appealing to the Lebesgue Dominated Convergence theorem, we may show that

$$\lim_{k\to\infty} q_{11}(k) = \textstyle\sum_{j=1}^{\infty} Ri(u+j)Ri(v+j). \tag{4.6}$$

In the same way, we may evaluate the limiting values as $k \to \infty$ of $q_{12}(k)$, $q_{21}(k)$ and $q_{22}(k)$ to finally obtain

(4.7) $\lim_{k\to\infty} Q_1(k) = \sum_{j=-\infty}^{\infty}\{Ri(u+j)Ri(v+j) + Ri(u+j)Ri(v-j)\} + \frac{\kappa_4}{\sigma^4} Ri(u)Ri(v).$

We note that (4.7) is in accordance with a remark of Parzen (1974) concerning the asymptotic covariance structure of the autoregressive estimate of the inverse covariance function.

It is easy to verify that the use of $\mathbf{\Sigma}(k)$ as an approximation to $\mathbf{R}(k)^{-1}$ does not produce the desired result, which, in view of Theorem 5 of Section 3, is not surprising.

EXAMPLE 2. Suppose that x_t satisfies (4.4) where q is known. Put

$$\mu(j) = \frac{1}{2\pi}\int_{-\pi}^{\pi} \overline{\{\nu_q(\lambda)\}^{-1}}\nu_q(\lambda)\exp(-ij\lambda)\,d\lambda,$$

where $\nu_q(\lambda) = 1 + \nu(1)e^{-i\lambda} + \cdots + \nu(q)e^{-iq\lambda}$ and the overbar denotes complex conjugate.

For evaluating the covariance term, $c(u, v)$ say, in the joint asymptotic distribution of $\sqrt{T}\{\hat{\nu}_k(1) - \nu(1)\}, \cdots, \sqrt{T}\{\hat{\nu}_k(q) - \nu(q)\}$, it is necessary to evaluate the limiting value, $d(m, n)$, say, as $k \to \infty$ of the quadratic form

$$Q_2(k) = \sigma^{-2}\mu_m(k)'\mathbf{R}(k)^{-1}\mu_n(k), \tag{4.8}$$

where m and n are fixed integers such that $1 \le m, n \le q$, $\mu_m(k) = (\mu(1-m), \cdots, \mu(k-m))'$ and $\mu_n(k)$ is defined analogously; see Bhansali (1980). Indeed

$$c(u, v) = \textstyle\sum_{m=1}^{q}\sum_{n=1}^{q} Ri^{-1}(u, m)d(m, n)Ri^{-1}(v, n),$$

where $Ri^{-1}(u, v)$ denotes the term in the uth row and vth column of $\mathbf{\Sigma}(q)^{-1}$.

On using Corollary 3.1, we have

$$d(m, n) = \lim_{k\to\infty} \sigma^{-2}\mu_m(k)'\mathbf{W}(k)\mu_n(k) = \sigma^{-2}Ri(m-n),$$

$$c(u, v) = \sigma^{-2}Ri^{-1}(u, v),$$

and thus the method suggested by Durbin (1959) provides asymptotically efficient estimates of the moving average parameters, relative to maximum likelihood in the Gaussian case.

It may again be verified that the use of $\mathbf{\Sigma}(k)$ in place of $\mathbf{R}(k)^{-1}$ for evaluating $Q_3(k)$ does not produce the desired result.

EXAMPLE 3. The covariance term $g(u, v)$, say, in the joint asymptotic distribution of $\sqrt{T}\{\hat{a}_{ku} - a(u)\}$ and $\sqrt{T}\{\hat{a}_{kv} - a(v)\}$, when u and v are held fixed, is given by the limiting value as $k \to \infty$ of the quantity (Bhansali, 1978)

$$Q_3(k) = \sigma^2\mathbf{e}_u(k)'\mathbf{R}(k)^{-1}\mathbf{e}_v(k),$$

where $\mathbf{e}_u(k)$ has unity in the uth place and 0 everywhere else, and $\mathbf{e}_v(k)$ is defined analogously. Using Corollary 3.1, we have

$$\lim_{k\to\infty} Q_3(k) = \lim_{k\to\infty} \sigma^2\mathbf{e}_u(k)'\mathbf{W}(k)\mathbf{e}_v(k) = \textstyle\sum_{j=0}^{\min(u,v)-1} a(j)a(j+|v-u|), \tag{4.9}$$

which is in accordance with the result of Bhansali (1978).

Parzen (1969, page 403) stated, without proof, that one may regard the estimated autoregressive coefficients $\hat{a}_{k1}, \cdots, \hat{a}_{kk}$ as a covariance stationary time series with means $a(1), \cdots, a(k)$ and spectral density function $(2\pi)^{-1}T^{-1}|A(\lambda)|^2$. Parzen thus appears to have in mind replacing $g(u, v)$ by $\sigma^{-2}Ri(u-v)$, i.e. to evaluate the limiting value as $k \to \infty$ of $Q_3(k)$ by replacing $\mathbf{R}(k)^{-1}$ by $\mathbf{\Sigma}(k)$. As shown earlier in Theorem 7, this procedure may not be justified. From (4.9), it is seen that if $u \ne v$ then the asymptotic variances of $\hat{a}_{ku}$ and $\hat{a}_{kv}$ are not the same and hence under the regularity conditions stated by Bhansali (1978), $\hat{a}_{ku}$ does not constitute a covariance stationary time series. Parzen's statement is thus in error. Nevertheless, note that if $\min(u, v)$ is not too small, then (4.9) and $\sigma^2 Ri(u - v)$ are close to each other.

Acknowledgements. The author would like to thank Professor S. J. Taylor of the Pure Mathematics Department, Liverpool University for a very helpful discussion on the subject of mode of convergence of $\mathbf{W}(k)$ to $\mathbf{R}^{-1}$ as $k \to \infty$.

A part of the research reported in this paper was carried out while the author was Visting Scientist at the University of Western Ontario (U.W.O.), London, Canada, and was supported by a grant from the National Science and Engineering Research Council of Canada. The author thanks Professor Ian B. MacNeill for making available the splendid research facilities of U.W.O. and for helpful discussions during his stay there. Thanks also are due to Dr. S. Z. Ditor and Dr. A. I. McLeod of U.W.O., and Professor C. T. G. Wall and Dr. M. Barlow of Liverpool University, for several useful comments. The helpful comments of the referees are also sincerely appreciated.

REFERENCES

AKHIEZER, N. I. and GLAZMAN, I. M. (1961). *Theory of Linear Operators in Hilbert Space, Volume I.* Ungar, New York.

ANDERSON, T. W. (1975). Maximum likelihood estimation of parameters of autoregressive processes with moving average residuals and other covariance matrices with linear structure. *Ann. Statist.* **3** 1283–1304.

ANDERSON, T. W. (1977). Estimation for autoregressive moving average models in the time and frequency domains. *Ann. Statist.* **5** 842–865.

BAXTER, G. (1963). A norm inequality for a "finite section" Wiener-Hopf equation. *Illinois J. Math.* **7** 97–103.

BERK, K. N. (1974). Consistent autoregressive spectral estimate. *Ann. Statist.* **2** 489–502.

BHANSALI, R. J. (1978). Linear prediction by autoregressive model fitting in the time domain. *Ann. Statist.* **6** 224–231.

BHANSALI, R. J. (1980). Autoregressive and window estimates of the inverse correlation function. *Biometrika* **67** 551–566.

BRILLINGER, D. R. (1975). *Time Series: Data Analysis and Theory.* Holt, Rinehart and Winston, New York.

CLEVELAND, W. S. (1972). The inverse correlations of a time series and their applications. *Technometrics* **14** 277–293.

DURBIN, J. (1959). Efficient estimation of parameters in moving average models. *Biometrika* **46** 306–316.

HANNAN, E. J. (1970). *Multiple Time Series.* Wiley, New York.

HUZII, M. (1977). On a spectral estimate obtained by an autoregressive model fitting. *Ann. Inst. Statist. Math.* **29** Part A 415–431.

KÖTHE, G. (1969). *Topological Vector Spaces I.* Springer-Verlag, Berlin.

PARZEN, E. (1969). Multiple time series modelling. In *Multivariate Analysis II* (ed. P. R. Krishnaiah). Academic, New York. 389–409.

PARZEN, E. (1974). Some recent advances in time series modelling. *IEEE Trans. Auto Control* **AC-19** 723–730.

RIESZ, F. and NAGY, B. (1956). *Functional Analysis.* Blackie, London.

SHAMAN, P. (1975). An approximate inverse for the covariance matrix of moving average and autoregressive processes. *Ann. Statist.* **3** 532–538.

SHAMAN, P. (1976). Approximations for stationary covariance matrices and their inverses with application to ARMA models. *Ann. Statist.* **4** 292–301.

TAYLOR, A. E. (1958). *Introduction to Functional Analysis.* Wiley, New York.

WHITTLE, P. (1951). *Hypothesis Testing in Time Series Analysis.* Almqvist and Wiksell, Uppsala.

WHITTLE, P. (1952). Some results in time series analysis. *Skand. Actuarietidskrift* **35** 48–60.

WISE, J. (1955). The autocorrelation function and the spectral density function. *Biometrika* **42** 151–159.

WOLD, H. (1953). *Demand Analysis.* Wiley, New York.

DEPARTMENT OF COMPUTATIONAL AND STATISTICAL SCIENCE
UNIVERSITY OF LIVERPOOL
VICTORIA BUILDING, BROWNLOW HILL
P.O. BOX 147
LIVERPOOL L69 3BX ENGLAND