R.A. Davis et al. (eds.), *Selected Works of Murray Rosenblatt*, Selected Works in Probability and Statistics, DOI 10.1007/978-1-4419-8339-8_14, © Springer Science+Business Media, LLC 2011

SOME REGRESSION PROBLEMS IN TIME SERIES ANALYSIS

MURRAY ROSENBLATT[1]
COLUMBIA UNIVERSITY AND NEW YORK UNIVERSITY

1. Introduction

Estimates of the regression coefficients which are unbiased and linear in the observations are discussed in this paper. The residual is assumed to be a stationary process. Two specific estimates are discussed, the least-squares estimate and the Markov estimate. I call the estimate which is computed under the assumption that the residual is an orthogonal process the least-squares estimate. The Markov estimate is the linear unbiased estimate with minimal covariance matrix. The basic assumptions made in the paper are discussed in section 2 and are held to throughout the paper. In section 3 some remarks about the approximation of a continuous positive definite matrix-valued function by finite trigonometric forms are made. These remarks are used in section 4 to obtain the main results about the asymptotic behavior of the covariance matrices of the least-squares and Markov estimates. The next section discusses the many interesting cases in which the least-squares estimate is asymptotically as good as the Markov estimate. The first really systematic discussion of some of these problems was given by U. Grenander [1]. Further work was carried out by U. Grenander and M. Rosenblatt in [2], [3], and [4]. The author considers some of these problems in the case of a vector-valued time series in [5]. Some of the results of this paper are a generalization of some of those obtained in [5].

A few cases in which the least-squares estimate is not asymptotically efficient in the class of linear unbiased estimates are discussed in sections 5 and 7. Some small sample computations for a linear regression with a residual which is a first order autoregressive scheme are carried out in section 6 to test the asymptotic theory.

2. Assumptions and notation

I assume that the observed process y_t is a vector-valued process (a k-vector)

$$y_t = x_t + m_t, \qquad t = \cdots, -1, 0, 1, \cdots, \tag{2.1}$$

where $m_t = Ey_t$ is the mean value sequence and x_t, $Ex_t \equiv 0$, is the sequence of residuals. The residual x_t is assumed to be weakly stationary, that is, the covariances

$$r_{t\ \tau} = r_{t-\tau} = Ex_t x'_\tau = E(y_t - m_t)(y_\tau - m_\tau)'\ ^{2} \tag{2.2}$$

depend only on the difference $t - \tau$. For mathematical convenience, in sections 3 and 4, I assume that the components of the vector observations are complex valued. The real-

Based in part on research supported by the Office of Naval Research at the Statistical Research Center, University of Chicago.

[1] Now at Indiana University.

[2] x_t is column vector. Given a matrix A, A' denotes the conjugated transpose of A.

valued case is the one of statistical interest and will be examined in detail later on. The mean value m_t is assumed to be of a regression form

$$m_t = \beta_1 \varphi_t^{(1)} + \beta_2 \varphi_t^{(2)} + \cdots + \beta_s \varphi_t^{(s)}, \tag{2.3}$$

where the regression vector sequences

$$\varphi_1^{(j)}, \varphi_2^{(j)}, \cdots, \varphi_n^{(j)}, \qquad j = 1, \cdots, s, \tag{2.4}$$

are assumed known and the regression coefficients $\beta_1, \cdots, \beta_s$ are unknown. I shall discuss the problem of estimating the regression coefficients by unbiased estimates linear in the observations

$$y_1, \cdots, y_n. \tag{2.5}$$

Two specific linear unbiased estimates will be discussed in some detail, the "least-squares" estimate and the Markov estimate. The least-squares estimate is obtained by treating the residuals as if they were orthogonal, that is,

$$Ex_t x_\tau' = \delta_{t,\tau} I \tag{2.6}$$

where $\delta_{t,\tau}$ is the Kronecker delta and I is the identity matrix. The Markov estimate is the optimal linear unbiased estimate in the sense of minimal covariance matrix of the estimate.

The covariance sequence r_t of a weakly stationary process has the representation

$$r_t = \int_{-\pi}^{\pi} e^{it\lambda} dF(\lambda) \tag{2.7}$$

where $F(\lambda)$ is a nondecreasing matrix-valued ($k \times k$) function, that is, $\Delta F(\lambda) \geqq 0$.[3] I shall assume that $F(\lambda)$ is absolutely continuous, that is,

$$F(\lambda) = \int_{-\pi}^{\lambda} f(\mu)\, d\mu \tag{2.8}$$

so that

$$r_t = \int_{-\pi}^{\pi} e^{it\lambda} f(\lambda)\, d\lambda. \tag{2.9}$$

The function $F(\lambda)$ is called the spectral distribution function of the process while $f(\lambda)$ is called the spectral density of the process. The spectral density is a nonnegative function of λ since

$$f(\lambda) = \frac{dF(\lambda)}{d\lambda} \geqq 0. \tag{2.10}$$

I shall assume that $f(\lambda)$ is a continuous function of λ [each element of $f(\lambda)$ is a continuous function of λ] and that $f(\lambda)$ is nonsingular for all λ.

For convenience I introduce the following notation. Let

$$y = \begin{pmatrix} y_1 \\ \cdot \\ \vdots \\ y_n \end{pmatrix}, \qquad x = \begin{pmatrix} x_1 \\ \cdot \\ \vdots \\ x_n \end{pmatrix}. \tag{2.11}$$

[3] Given a square matrix A, $A \geqq 0$ means that the corresponding quadratic form is positive semidefinite. $A > 0$ means that the corresponding quadratic form is positive definite.

Let

(2.12) $$\varphi_t = (\varphi_t^{(1)}, \cdots, \varphi_t^{(s)})$$

and

(2.13) $$\varphi = \begin{pmatrix} \varphi_1 \\ \cdot \\ \cdot \\ \cdot \\ \varphi_n \end{pmatrix}.$$

Equation (2.1) can then be rewritten in the form

(2.14) $$y = \varphi\beta + x .$$

The matrix R is the covariance matrix of y (or x). Here

(2.15) $$\beta = \begin{pmatrix} \beta_1 \\ \cdot \\ \cdot \\ \cdot \\ \beta_s \end{pmatrix}.$$

The matrix R is nonsingular since $f(\lambda)$ is continuous and nonsingular for all λ. The matrix $\varphi'\varphi$ is also assumed to be nonsingular.

The "least-squares" estimate β_L is the vector β that minimizes the quadratic form

(2.16) $$(y - \varphi\beta)'(y - \varphi\beta)$$

and is given by

(2.17) $$\beta_L = [\varphi'\varphi]^{-1}\varphi'y .$$

It is clearly an unbiased estimate and has covariance matrix

(2.18) $$E(\beta_L - \beta)(\beta_L - \beta)' = [\varphi'\varphi]^{-1}\varphi'R\varphi[\varphi'\varphi]^{-1} .$$

The linear unbiased estimate with minimal covariance matrix or Markov estimate is given by

(2.19) $$\beta_M = [\varphi'R^{-1}\varphi]^{-1}\varphi'R^{-1}y .$$

Its covariance matrix is given by

(2.20) $$E(\beta_M - \beta)(\beta_M - \beta)' = [\varphi'R^{-1}\varphi]^{-1} .$$

These remarks on the least-squares and Markov estimates are well known.

The techniques used in the paper can be considered a sort of generalized harmonic analysis. In order to carry out the analysis, various assumptions on the asymptotic behavior of the regression vectors are introduced. These assumptions are broad enough to include most of the usual types of regression, such as polynomial and trigonometric regression. They do not include the case of exponential regression.

Let

(2.21) $$\Phi_n^{(j)} = \sum_{t=1}^{n} \varphi_t^{(j)\prime}\varphi_t^{(j)}, \qquad j = 1, \cdots, s .$$

It is assumed that

$$\Phi_n^{(j)} \to \infty \tag{2.22}$$

as $n \to \infty$, $j = 1, \cdots, s$. It is also assumed that

$$\lim_{n\to\infty} \frac{\Phi_{n+h}^{(j)}}{\Phi_n^{(j)}} = 1 \tag{2.23}$$

for every fixed h. Consider the k-vectors

$$\varphi_t^{(j)} = \begin{pmatrix} {}_1\varphi_t^{(j)} \\ \cdot \\ \cdot \\ \cdot \\ {}_k\varphi_t^{(j)} \end{pmatrix} \qquad \begin{matrix} j = 1, \cdots, s \\ t = 1, \cdots, n . \end{matrix} \tag{2.24}$$

Let the limits

$$\lim_{n\to\infty} \sum_{t=1}^{n} \frac{{}_p\varphi_{t+h}^{(j)}\, {}_q\varphi_t^{(l)}}{\sqrt{\Phi_n^{(j)}\Phi_n^{(l)}}} = {}_{jl}m_h^{(p,\, q)} \tag{2.25}$$

exist, $j, l = 1, \cdots, s$ and $p, q = 1, \cdots, k$ [if $t < 0$ set ${}_q\varphi_t^{(l)} = 0$]. Set

$${}_{jl}M_h = \{ {}_{jl}m_h^{(p,\, q)} ; \quad p, q = 1, \cdots, k \} \tag{2.26}$$

and

$$M_h = \{ {}_{jl}M_h ; \quad j, l = 1, \cdots, s \} . \tag{2.27}$$

The matrices M_h, $h = \cdots, -1, 0, 1, \cdots$ form a positive definite sequence, that is, given any finite collection of ks-vectors $\{a_\nu\}$

$$\sum_{\nu,\, \mu} a'_\nu M_{\nu-\mu}\, a_\mu \geqq 0 . \tag{2.28}$$

It then follows that the matrices M_h have the representation

$$M_h = \int_{-\pi}^{\pi} e^{ih\lambda} dM(\lambda) \tag{2.29}$$

where $M(\lambda)$ is a nondecreasing matrix-valued ($ks \times ks$) function of λ. In accordance with the notation introduced in (2.26) and (2.27) I write

$$M(\lambda) = \{ {}_{jl}M(\lambda) ; \quad j, l = 1, \cdots, s \} \tag{2.30}$$

and

$$\begin{aligned} {}_{jl}M(\lambda) &= \{ {}_{jl}M_{pq}(\lambda) ; \quad p, q = 1, \cdots, k \} , \\ M_{pq}(\lambda) &= \{ {}_{jl}M_{pq}(\lambda) ; \quad j, l = 1, \cdots, s \} . \end{aligned} \tag{2.31}$$

Note that

$$M_0 = \int_{-\pi}^{\pi} dM(\lambda) = M(\pi) - M(-\pi) . \tag{2.32}$$

It will be convenient to introduce some additional notation. Given a $k \times k$ matrix f and a form $M = \{ {}_{ij}m_{uv} ; i, j = 1, \cdots, s ; u, v = 1, \cdots, k \}$, let

$$(f \cdot M) = \left\{ \sum_{i,\, j} f_{ij}\, {}_{uv}m_{ij} ; \quad u, v = 1, \cdots, s \right\} . \tag{2.33}$$

By the integral

$$\int_{-\pi}^{\pi} [f(\lambda) \cdot dM(\lambda)] \tag{2.34}$$

I shall mean

$$\left\{\int_{-\pi}^{\pi} \sum_{i,j} f_{ij}(\lambda)\, d_{pq}M_{ij}(\lambda); \qquad p, q = 1, \cdots, s\right\}. \tag{2.35}$$

The matrix

$$T = (I \cdot M_0) \tag{2.36}$$

is assumed to be nonsingular. This means that the vector sequences

$$\varphi_1^{(j)}, \cdots, \varphi_n^{(j)}, \qquad j = 1, \cdots, s, \tag{2.37}$$

are asymptotically linearly independent in a sense which is relevant in this context. Conditions (2.22) and (2.23) are introduced to ensure that the estimates of the regression coefficients converge to the true regression coefficients in the mean square.

It will also be convenient to introduce the matrix

$$D_n = \begin{bmatrix} \Phi_n^{(1)\,1/2} & & 0 \\ & \ddots & \\ 0 & & \Phi_n^{(s)\,1/2} \end{bmatrix}. \tag{2.38}$$

3. Remarks on approximation

The arguments used to obtain results on the asymptotic behavior of the covariance matrices of the least-squares and Markov estimates are approximation arguments. They make use of finite trigonometric polynomials that uniformly approximate the matrix-valued spectral density $f(\lambda)$.

LEMMA 1. *Let $f(\lambda)$ be a continuous positive definite matrix-valued ($k \times k$) function of λ. Given any $\epsilon > 0$, there is a positive definite matrix-valued trigonometric polynomial*

$$g(\lambda) = \sum_{u=-p}^{p} g_u e^{iu\lambda} \tag{3.1}$$

with coefficients g_u Hermitian $k \times k$ matrices such that

$$\epsilon z'z > z'[f(\lambda) - g(\lambda)]z > -\epsilon z'z \tag{3.2}$$

for every k-vector z.

There are finite trigonometric polynomials $g_{ij}(\lambda)$ such that

$$g_{ij}(\lambda) = \overline{g_{ji}(\lambda)} \tag{3.3}$$

and

$$|f_{ij}(\lambda) - g_{ij}(\lambda)| < \delta, \qquad \delta < 0, \tag{3.4}$$

where $i, j = 1, \cdots, k$. But then on setting $g(\lambda) = \{g_{ij}(\lambda); i, j = 1, \cdots, k\}$,

$$z'[f(\lambda) - g(\lambda)]z \leqq \sum \delta |z_i| |z_j| \leqq k\delta z'z. \tag{3.5}$$

If $\delta > 0$ is chosen sufficiently small

$$z'[f(\lambda) - g(\lambda)]z < \epsilon z'z. \tag{3.6}$$

One can show similarly that

$$-\epsilon z\, z < z' [f(\lambda) - g(\lambda)]\, z\,. \tag{3.7}$$

Note that p is the maximal order of the polynomials $p_{ij}(\lambda)$.

LEMMA 2. *Let $f(\lambda)$ be a continuous positive definite matrix-valued $(k \times k)$ function of λ. Given any $\epsilon > 0$ sufficiently small, there are positive definite matrix-valued trigonometric polynomials*

$$\begin{aligned} g(\lambda) &= \sum_{u=-p}^{p} g_u e^{iu\lambda} \\ h(\lambda) &= \sum_{u=-p}^{p} h_u e^{iu\lambda} \end{aligned} \tag{3.8}$$

with coefficients g_u, h_u Hermitian $k \times k$ matrices such that

$$0 < f(\lambda) - \epsilon I < g(\lambda) < f(\lambda) < h(\lambda) < f(\lambda) + \epsilon I \tag{3.9}$$

where I is the identity matrix $(k \times k)$.

Since $f(\lambda)$ is a positive continuous function of λ, for all sufficiently small $\epsilon > 0$

$$0 < f(\lambda) - \epsilon I \tag{3.10}$$

and clearly

$$f(\lambda) - \epsilon I < f(\lambda)\,. \tag{3.11}$$

By lemma 1, there is a trigonometric polynomial $g(\lambda)$ such that

$$-\tfrac{1}{2}\epsilon I < f(\lambda) - \tfrac{1}{2}\epsilon I - g(\lambda) < \tfrac{1}{2}\epsilon I\,. \tag{3.12}$$

But then

$$0 < f(\lambda) - \epsilon I < g(\lambda) < f(\lambda)\,. \tag{3.13}$$

One can similarly show that there is a trigonometric polynomial $h(\lambda)$ such that

$$f(\lambda) < h(\lambda) < f(\lambda) + \epsilon I\,. \tag{3.14}$$

LEMMA 3. *Let*

$$f(\lambda) = a_0 + a_1 \cos \lambda + \cdots + a_p \cos p\lambda + b_1 \sin \lambda + \cdots + b_p \sin p\lambda \tag{3.15}$$

be a positive definite matrix-valued $(k \times k)$ function of λ with the coefficients a_ν, b_ν $(k \times k)$ matrices. Then $f(\lambda)$ can be written in the form

$$f(\lambda) = \frac{1}{2\pi} \left(\sum_{j=0}^{p} c_j e^{-ij\lambda} \right) \left(\sum_{j=0}^{p} c_j e^{-ij\lambda} \right) \tag{3.16}$$

where c_0 is nonsingular and

$$\int_{-\pi}^{\pi} e^{-il\lambda} \left(\sum_{j=0}^{p} c_j e^{-ij\lambda} \right)^{-1} d\lambda = 0 \tag{3.17}$$

where $l > 0$.

Since $f(\lambda)$ is a positive definite matrix-valued $(k \times k)$ function of λ, there is a k-dimensional weakly stationary process x_t, $Ex_t \equiv 0$, having $f(\lambda)$ as its spectral density. The covariance matrices

$$r_t = Ex_{t+\tau}x'_\tau = \int_{-\pi}^{\pi} e^{it\lambda} f(\lambda)\, d\lambda \tag{3.18}$$

are the null matrix if $|t| > p$. Let Px_t be the projection of x_t on the closed linear manifold spanned by $x_{t-1}, x_{t-2}, \cdots$. Consider $\eta_t = x_t - Px_t$. The η_t are orthogonal to each other, that is, $E\eta_t\eta'_\tau = 0$ when $t \neq \tau$. Since x_t and η_t are stationary I can write

$$x_t = \int_{-\pi}^{\pi} e^{it\lambda} dZ_x(\lambda), \qquad \eta_t = \int_{-\pi}^{\pi} e^{it\lambda} dZ_\eta(\lambda) \tag{3.19}$$

where $z_x(\lambda)$, $z_\eta(\lambda)$ are processes with orthogonal increments

$$\begin{aligned} EdZ_x(\lambda)\, dZ_x(\mu)' &= \delta_{\lambda\mu} f(\lambda)\, d\lambda \\ EdZ_\eta(\lambda)\, dZ_\eta(\mu)' &= \frac{1}{2\pi}\, \delta_{\lambda\mu} N d\lambda . \end{aligned} \tag{3.20}$$

Here $N = E\eta_t\eta'_t$. Because $r_t = 0$ for $t < -p$ it is clear that

$$x_t = \int_{-\pi}^{\pi} e^{it\lambda} dZ_x(\lambda) = \sum_{j=0}^{p} h_j \eta_{t-j} = \int_{-\pi}^{\pi} e^{it\lambda} \left(\sum_{j=0}^{p} h_j e^{-ij\lambda} \right) dZ_\eta(\lambda), \tag{3.21}$$

and $h_0 = I$. On approximating the characteristic function of the set $[-\pi, \lambda]$ in the mean square by linear combinations of the exponentials exp $it\lambda$, equation

$$\int_{-\pi}^{\lambda} dZ_x(\lambda) = \int_{-\pi}^{\lambda} \sum_{j=0}^{p} h_j e^{-ij\lambda} dZ_\eta(\lambda) \tag{3.22}$$

is obtained. On taking the covariance matrix of both sides of equation (3.22)

$$F_x(\lambda) = \frac{1}{2\pi} \int_{-\pi}^{\lambda} \left(\sum_{j=0}^{p} h_j e^{-ij\lambda} \right) N \left(\sum_{j=0}^{p} h_j e^{-ij\lambda} \right)' d\lambda \tag{3.23}$$

or

$$f(\lambda) = \frac{1}{2\pi} \left(\sum_{j=0}^{p} h_j e^{-ij\lambda} \right) N \left(\sum_{j=0}^{p} h_j e^{-ij\lambda} \right)' \tag{3.24}$$

is obtained. Since $f(\lambda)$ is nonsingular, the matrix N must be nonsingular. Norm the η_t's so as to get

$$\xi_t = N^{-1/2} \eta_t . \tag{3.25}$$

The ξ_t's are an orthonormal process, that is,

$$E\xi_t\, \xi'_\tau = \delta_{t,\tau} I . \tag{3.26}$$

I can now write

$$x_t = \sum_{j=0}^{p} c_j\, \xi_{t-j} \tag{3.27}$$

where $c_0 = N^{1/2}$. By using the argument that led to equation (3.22) one can see that

$$f(\lambda) = \frac{1}{2\pi} \left(\sum_{j=0}^{p} c_j e^{-ij\lambda} \right) \left(\sum_{j=0}^{p} c_j e^{-ij\lambda} \right)' . \tag{3.28}$$

Since ξ_t is weakly stationary

$$\xi_t = \int_{-\pi}^{\pi} e^{it\lambda} dZ_\xi(\lambda) = \int_{-\pi}^{\pi} e^{it\lambda} \Big(\sum_{j=0}^{p} c_j e^{-ij\lambda} \Big)^{-1} dZ_x(\lambda) \tag{3.29}$$

$$E dZ_\xi(\lambda)\, dZ_\xi(\mu)' = \delta_{\lambda\mu} \frac{1}{2\pi} I\, d\lambda . \tag{3.30}$$

By our construction ξ_t is in the linear manifold spanned by $x_t, x_{t-1}, \cdots$ so that

$$\int_{-\pi}^{\pi} e^{-il\lambda} \Big(\sum_{j=0}^{p} c_j e^{-ij\lambda} \Big)^{-1} d\lambda = 0 \tag{3.31}$$

when $l > 0$.

4. The asymptotic covariance matrices

In obtaining the asymptotic form of the covariance matrices of the least-squares and Markov estimates, it will be convenient to deal with

$$D_n E(\beta_L - \beta)(\beta_L - \beta)' D_n = D_n [\varphi'\varphi]^{-1} D_n D_n^{-1} \varphi' R \varphi D_n^{-1} D_n [\varphi'\varphi]^{-1} D_n \tag{4.1}$$

and

$$D_n E(\beta_M - \beta)(\beta_M - \beta)' D_n = D_n [\varphi' R^{-1} \varphi]^{-1} D_n . \tag{4.2}$$

THEOREM 1. *Under the conditions on the spectrum of the process x_t given in section 2 and the conditions on the regression vectors specified there*

$$\lim_{n\to\infty} D_n E(\beta_L - \beta)(\beta_L - \beta)' D_n = 2\pi T^{-1} \int_{-\pi}^{\pi} [f(-\lambda) \cdot dM(\lambda)] T^{-1} . \tag{4.3}$$

The conditions on the regression vectors specified in section 2 imply that

$$D_n^{-1} [\varphi'\varphi] D_n^{-1} \to T \tag{4.4}$$

as $n \to \infty$ and the limiting $s \times s$ matrix T is nonsingular. I therefore need only consider the asymptotic behavior of

$$D_n^{-1} \varphi' R \varphi D_n^{-1} . \tag{4.5}$$

Given any sufficiently small $\epsilon > 0$, by lemma 2 there are finite matrix-valued ($k \times k$) trigonometric polynomials

$$\begin{aligned} g(\lambda) &= \frac{1}{2\pi} \sum_{u=-p}^{p} g_u e^{iu\lambda} \\ h(\lambda) &= \frac{1}{2\pi} \sum_{u=-1}^{p} h_u e^{iu\lambda} \end{aligned} \tag{4.6}$$

such that

$$0 < f(\lambda) - \epsilon I < g(\lambda) < f(\lambda) < h(\lambda) < f(\lambda) + \epsilon I . \tag{4.7}$$

Let G, H be the covariance matrices of y if x_t has the spectral densities $g(\lambda)$, $h(\lambda)$, respectively. Then

$$\varphi' G \varphi < \varphi' R \varphi < \varphi' H \varphi . \tag{4.8}$$

I shall obtain the limit of (4.5) as $n \to \infty$ when the covariance matrix of x is G. The matrix

$$\varphi' G \varphi = \sum_{t,\,\tau=1}^{n} \varphi_t' \, g_{t-\tau} \, \varphi_\tau . \tag{4.9}$$

A typical element of (4.9) is therefore of the form

$$\frac{\sum_{t,\,\tau=1}^{n} \varphi_t^{(p)\prime} g_{t-\tau} \varphi_\tau^{(q)}}{[\Phi_n^{(p)} \Phi_n^{(q)}]^{1/2}} . \tag{4.10}$$

Since there are only a finite number of nonzero g_u's, the limit of expression (4.10) is

$$2\pi \sum_{i,\,j=1}^{k} \int_{-\pi}^{\pi} g_{ij}(-\lambda) \, d_{pq} M_{ij}(\lambda) . \tag{4.11}$$

Thus

$$\lim_{n\to\infty} D_n^{-1} \varphi' G \varphi D_n^{-1} = 2\pi \int_{-\pi}^{\pi} [g(-\lambda) \cdot dM(\lambda)] . \tag{4.12}$$

Similarly

$$\lim_{n\to\infty} D_n^{-1} \varphi' H \varphi D_n^{-1} = 2\pi \int_{-\pi}^{\pi} [h(-\lambda) \cdot dM(\lambda)] . \tag{4.13}$$

On letting $\epsilon \to 0$, it is clear that

$$\lim_{n\to\infty} D_n^{-1} \varphi' R \varphi D_n^{-1} = 2\pi \int_{-\pi}^{\pi} [f(-\lambda) \cdot dM(\lambda)] . \tag{4.14}$$

Note that the expression

$$\int_{-\pi}^{\pi} [f(-\lambda) \cdot dM(\lambda)] \tag{4.15}$$

is nonsingular for $f(\lambda) > \epsilon I$ if $\epsilon > 0$ is sufficiently small. But then

$$\int_{-\pi}^{\pi} [f(-\lambda) \cdot dM(\lambda)] > \epsilon \int_{-\pi}^{\pi} [I \cdot dM(\lambda)] = \epsilon (I \cdot M_0) \tag{4.16}$$

which is nonsingular.

THEOREM 2. *Under the conditions on the spectrum of x_t and the regression vectors assumed in section* 2

$$\lim_{n\to\infty} D_n E(\beta_M - \beta)(\beta_M - \beta)' D_n = 2\pi \left(\int_{-\pi}^{\pi} [f^{-1}(-\lambda) \cdot dM(\lambda)] \right)^{-1} . \tag{4.17}$$

By lemmas 2 and 3, for every sufficiently small $\epsilon > 0$ there are finite trigonometric polynomials

$$\begin{aligned} g(\lambda) &= \sum_{u=0}^{p} g_u e^{-iu\lambda} \\ h(\lambda) &= \sum_{u=0}^{p} h_u e^{-iu\lambda} \end{aligned} \tag{4.18}$$

with g_0, h_0 nonsingular and satisfying relation (3.17) and such that

$$0 < f(\lambda) - \epsilon I < \frac{1}{2\pi} g^{-1}(\lambda) g^{-1}(\lambda)' < f(\lambda) < \frac{1}{2\pi} h^{-1}(\lambda) h^{-1}(\lambda)' < f(\lambda) + \epsilon I .$$

Let G, H be the covariance matrices of the vector x if the process x_t has the spectral densities $(1/2\pi)g^{-1}(\lambda)g^{-1}(\lambda)'$, $(1/2\pi)h^{-1}(\lambda)h^{-1}(\lambda)'$, respectively.

In obtaining our result it will be sufficient to consider the asymptotic behavior of $D_n^{-1}\varphi' R^{-1}\varphi D_n^{-1}$. Now

$$0 < \varphi' H^{-1}\varphi < \varphi' R^{-1}\varphi < \varphi' G^{-1}\varphi \, . \tag{4.19}$$

Assume for the moment that x_t has the spectral density $(1/2\pi)g^{-1}(\lambda)g^{-1}(\lambda)'$ and consider the asymptotic behavior of $D_n^{-1}\varphi' G^{-1}\varphi D_n^{-1}$ as $n \to \infty$. Now carry out a Gramm-Schmidt orthogonalization procedure with respect to $x_1, \cdots, x_n$ starting with x_1. The first p equations are

$$\begin{aligned} d_{11}x_1 &= \xi_1 \\ d_{21}x_1 + d_{22}x_2 &= \xi_2 \\ &\cdots \\ d_{p1}x_1 + \cdots + d_{pp}x_p &= \xi_p \, . \end{aligned} \tag{4.20}$$

The residual ξ's are orthonormal. At the next step I have

$$g_p x_1 + \cdots + g_0 x_{p+1} = \xi_{p+1} = \eta_{p+1} \tag{4.21}$$

since x_t satisfies the difference equation

$$\sum_{u=0}^{p} g_u \, x_{t-u} = \eta_t \tag{4.22}$$

with the η_t's orthonormal. From then on we have the equation

$$g_p x_k + \cdots + g_0 \, x_{k+p} = \xi_{k+p} = \eta_{k+p} \, , \qquad k = 1, \cdots, n-p \tag{4.23}$$

Let the matrix

$$\Delta = \begin{pmatrix} d_{11} & & & & & \\ d_{21} d_{22} & & & 0 & & \\ \cdot \;\cdot\; \cdot & & & & & \\ d_{p1} d_{p2} & \cdots d_{pp} & & & & \\ g_p \; g_{p-1} & \cdots g_1 \; g_0 & & & & \\ & g_p & \cdots g_2 \; g_1 \; g_0 & & & \\ & & \cdot & & \cdot & \\ & 0 & & \cdot & & \cdot \\ & & & & \cdot & \cdot \end{pmatrix} . \tag{4.24}$$

Then

$$\Delta x = \xi = \begin{pmatrix} \xi_1 \\ \cdot \\ \cdot \\ \xi_n \end{pmatrix} \tag{4.25}$$

where ξ is a vector of orthonormal random vectors. Δ is a nonsingular transformation taking $x_1, \cdots, x_n$ into $\xi_1, \cdots, \xi_n$. On taking the covariance matrix of both sides of equation (4.25) I have

$$\Delta G \Delta' = I \tag{4.26}$$

so that

$$G^{-1} = \Delta'\Delta\,. \tag{4.27}$$

Let $\{g_{u,v}^{-1};\ u, v = 1,\cdots, n\} = G^{-1}$. Then

$$g_{\nu\mu}^{-1} = \sum_{u=-\infty}^{\infty} g'_{\nu+u}\, g_{\mu+u} \tag{4.28}$$

if ν or μ is greater than p. Here g_ν is understood to be zero if $\nu < 0$ or $\nu > p$. But

$$\sum_{u=-\infty}^{\infty} g'_{\nu+u}\, g_{\mu+u} = \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{i(\nu-\mu)\lambda} g(\lambda)'g(\lambda)\, d\lambda\,. \tag{4.29}$$

Now the (p, q)th element of $D_n^{-1}\varphi' G^{-1}\varphi D_n^{-1}$ is

$$\frac{\sum_{t,\tau=1}^{n} \varphi_t^{(p)\prime} g_{t,\tau}^{-1} \varphi_\tau^{(q)}}{(\Phi_n^{(p)}\Phi_n^{(q)})^{1/2}} = \sum_{m=0}^{p}\sum_{v=1}^{n-m} \frac{\phi_{m+v}^{(p)\prime}\gamma_m \varphi_v^{(q)}}{\sqrt{\Phi_n^{(p)}\Phi_n^{(q)}}} + \sum_{m=-1}^{-p}\sum_{u=m+1}^{n} \frac{\varphi_u^{(p)\prime}\gamma_m \varphi_{u-m}^{(q)}}{\sqrt{\Phi_n^{(p)}\Phi_n^{(q)}}} + \delta_n\,. \tag{4.30}$$

Here $\gamma_m = (1/2\pi)\int_{-\pi}^{\pi} e^{im\lambda} g(\lambda)'g(\lambda)d\lambda$.

Now δ_n is the sum of at most $4p^2$ terms of the form

$$\frac{\varphi_u^{(p)\prime}\, g_{u,v}^{-1}\, \varphi_v^{(q)}}{\sqrt{\Phi_n^{(p)}\Phi_n^{(q)}}}\,. \tag{4.31}$$

Since every element of $g_{u,v}^{-1}$ is smaller in absolute value than δ^{-1} for some small but fixed $\delta > 0$, the elements of

$$\frac{\varphi_u^{(p)\prime}\, g_{u,v}^{-1}\, \varphi_v^{(q)}}{\sqrt{\Phi_n^{(p)}\Phi_n^{(q)}}} \tag{4.32}$$

converge to zero as $n \to \infty$. But then

$$\lim_{n\to\infty} \frac{\sum_{t,\tau=1}^{n} \varphi_t^{(p)\prime} g_{t,\tau}^{-1} \varphi_\tau^{(q)}}{(\Phi_n^{(p)}\Phi_n^{(q)})^{1/2}} = \frac{1}{2\pi}\int_{-\pi}^{\pi} \sum_{i,j=1}^{k} \{g(-\lambda)'g(-\lambda)\}_{ij} d_{pq} M_{ij}(\lambda)\,. \tag{4.33}$$

I have now shown that

$$\lim_{n\to\infty} D_n^{-1}\varphi' G^{-1}\varphi\, D_n^{-1} = \int_{-\pi}^{\pi} [g(-\lambda)'g(-\lambda) \cdot dM(\lambda)]\,. \tag{4.34}$$

In the same way one can show that

$$\lim_{n\to\infty} D_n^{-1}\varphi' H^{-1}\varphi\, D_n^{-1} = \int_{-\pi}^{\pi} [h(-\lambda)'h(-\lambda) \cdot dM(\lambda)]\,. \tag{4.35}$$

On letting $\epsilon \downarrow 0$ the desired result

$$\lim_{n\to\infty} D_n^{-1}\varphi' R^{-1}\varphi\, D_n^{-1} = \frac{1}{2\pi}\int_{-\pi}^{\pi} [f^{-1}(-\lambda) \cdot dM(\lambda)] \tag{4.36}$$

is obtained. The matrix (4.36) can be seen to be nonsingular by using the argument used at the end of the proof of theorem **1**.

5. Asymptotic efficiency

It is interesting to investigate those types of regression for which the least-squares estimate is asymptotically efficient in the class of linear unbiased estimates for all admissible spectra $f(\lambda)$. In most cases the covariance matrix R is unknown and a reasonably large sample size is required to get adequate precision in estimating it. For this reason it would be convenient if one could use the least-squares estimate instead of the Markov estimate since the least-squares estimate does not require knowledge of R. Even if R is known, it may be difficult to compute R^{-1} which is required for the Markov estimate. In view of the results already obtained, the least-squares estimate will be asymptotically efficient if

$$T^{-1}\int_{-\pi}^{\pi}[f(-\lambda)\cdot dM(\lambda)]\,T^{-1}\int_{-\pi}^{\pi}[f^{-1}(-\lambda)\cdot dM(\lambda)] = I \tag{5.1}$$

for all admissible $f(\lambda)$. The case of interest is that in which the process x_t and the regression vectors have real components. Because of this $f(\lambda)$ and $M(\lambda)$ must satisfy the additional restraints

$$\begin{aligned} f(\lambda) &= \overline{f(-\lambda)} \\ dM(\lambda) &= \overline{dM(-\lambda)}\,. \end{aligned} \tag{5.2}$$

When $k = 1$, asymptotic efficiency of the least-squares estimate has been discussed in U. Grenander [1] and U. Grenander and M. Rosenblatt [2], [3].

In the one-dimensional case it is convenient to set

$$T(\lambda) = M(\lambda+) - M(-\lambda-) \tag{5.3}$$

and rewrite (5.1) in the form

$$T^{-1}\int_{0-}^{\pi} f(\lambda)\,dT(\lambda)\,T^{-1}\int_{0-}^{\pi} f^{-1}(\lambda)\,dT(\lambda) = I\,. \tag{5.4}$$

Equation (5.4) is satisfied for all positive continuous $f(\lambda)$ if and only if $T(\lambda)$ increases only at a finite number of points $0 \leqq \lambda_1 < \cdots < \lambda_q \leqq \pi$, $q < s$, and the jumps

$$T_i = \Delta T(\lambda_i) \tag{5.5}$$

satisfy the relations

$$T_i T^{-1} T_j = \delta_{ij} T_i\,. \tag{5.6}$$

These conditions are satisfied if one has a polynomial regression

$$m_t = \beta_0 + \cdots + \beta_{s-1} t^{s-1}\,, \tag{5.7}$$

a trigonometric regression

$$m_t = \beta_1 \cos t\lambda_1 + \cdots + \beta_{s_1} \cos t\lambda_{s_1} + \beta_{s_1+1} \sin t\lambda_1 + \cdots + \beta_{s_1+s_2} \sin t\lambda_{s_2} \tag{5.8}$$

(with the points λ_i distinct), or more generally a mixed polynomial and trigonometric regression

$$m_t = \sum_{u=1}^{q} \sum_{v=1}^{{}_1s_u} {}_1\beta_{u,v}\, t^{v-1} \cos t\lambda_u + \sum_{u=1}^{q} \sum_{v=1}^{{}_2s_u} {}_2\beta_{u,v}\, t^{v-1} \sin t\lambda_u \tag{5.9}$$

(with points λ_i distinct) [3]. Obviously the sine terms in this last regression form disappear if $\lambda_i = 0$. Notice that these regression sequences include most of those used in standard statistical work. It is easy to construct a regression sequence where the least-squares estimate is not asymptotically efficient. Consider

$$m_t = \beta(a_0 + a_1 \cos t\lambda_1 + \cdots + a_k \cos t\lambda_k), \quad 0 < \lambda_1 < \cdots < \lambda_k, \tag{5.10}$$

where the constants a_i are known. Such a regression has a form similar to the pulse trains encountered in communication theory. In the case of such a regression $T(\lambda)$ increases only at the points $0, \lambda_1, \cdots, \lambda_k$. The jump of $T(\lambda)$ at 0 is

$$\frac{a_0^2}{a_0^2 + \frac{1}{2}\sum_{j=1}^{k} a_j^2} \tag{5.11}$$

and the jump at λ_j, $j = 1, \cdots, k$, is

$$\frac{\frac{1}{2}a_j^2}{a_0^2 + \frac{1}{2}\sum_{j=1}^{k} a_j^2}. \tag{5.12}$$

The asymptotic efficiency of the least-squares estimate β_L in the class of linear unbiased estimates is

$$\frac{\left(a_0^2 + \frac{1}{2}\sum_{j=1}^{k} a_j^2\right)^2}{\left[a_0^2 f(0) + \frac{1}{2}\sum_{j=1}^{k} a_j^2 f(\lambda_j)\right]\left[\frac{a_0^2}{f(0)} + \frac{1}{2}\sum_{j=1}^{k}\frac{a_j^2}{f(\lambda_j)}\right]} \leq 1. \tag{5.13}$$

In section 7 I shall discuss the question of how much additional information about the spectrum $f(\lambda)$ is required to construct an estimate with the same asymptotic mean square error as the Markov estimate.

In the case of multidimensional time series, new phenomena arise. Consider first the case of a polynomial regression. If each component of the time series has a polynomial regression, that is,

$$_i m_t = \sum_{k=1}^{s} {}_i\beta_k t^{k-1}, \qquad i = 1, \cdots, k, \tag{5.14}$$

the least-squares estimate β_L is still asymptotically efficient. However, if the different components have polynomial regressions of different orders, the least-squares estimate is no longer asymptotically efficient [5]. A simple example is that in which the mean value of the first coordinate of a two-dimensional time series is unknown while the mean value of the second coordinate is known to be zero. Then

$$m_t = \beta\varphi_t, \qquad \varphi_t = \begin{pmatrix}1\\0\end{pmatrix}. \tag{5.15}$$

The function $M(\lambda)$ increases only at zero and the jump at zero

$$\Delta M(0) = \begin{pmatrix}\Delta M_{11}(0) & \Delta M_{12}(0)\\ \Delta M_{21}(0) & \Delta M_{22}(0)\end{pmatrix} = \begin{pmatrix}1 & 0\\ 0 & 0\end{pmatrix}. \tag{5.16}$$

Thus

$$
\begin{aligned}
&\int_{-\pi}^{\pi} [f(-\lambda) \cdot dM(\lambda)] = f_{11}(0) \\
(5.17) \qquad &\int_{-\pi}^{\pi} [f^{-1}(-\lambda) \cdot dM(\lambda)] = \frac{f_{22}(0)}{f_{11}(0)\, f_{22}(0) - f_{12}^2(0)} .
\end{aligned}
$$

The asymptotic efficiency of the least-squares estimate is

$$
(5.18) \qquad 1 - \frac{f_{12}^2(0)}{f_{11}(0)\, f_{22}(0)} .
$$

If there is a mixed polynomial and trigonometric regression (5.9) not only must the same regression form occur in each component, but one must also have ${}_1s_u = {}_2s_u$ if $\lambda_u \neq 0$. Thus, the least-squares estimate will be asymptotically efficient in the case of a regression

$$
\begin{aligned}
(5.19) \qquad m_t &= \beta_1 \varphi_t^{(1)} + \beta_2 \varphi_t^{(2)} + \beta_3 \varphi_t^{(3)} + \beta_4 \varphi_t^{(4)} \\
&= \beta_1 \begin{pmatrix} \cos t\lambda \\ 0 \end{pmatrix} + \beta_2 \begin{pmatrix} \sin t\lambda \\ 0 \end{pmatrix} + \beta_3 \begin{pmatrix} 0 \\ \cos t\lambda \end{pmatrix} + \beta_4 \begin{pmatrix} 0 \\ \sin t\lambda \end{pmatrix}, \qquad \lambda \neq 0 ,
\end{aligned}
$$

but not in the case of a regression

$$
(5.20) \qquad m_t = \beta_1 \varphi_t^{(1)} + \beta_2 \varphi_t^{(2)} = \beta_1 \begin{pmatrix} \cos t\lambda \\ 0 \end{pmatrix} + \beta_2 \begin{pmatrix} 0 \\ \cos t\lambda \end{pmatrix}, \qquad \lambda \neq 0 .
$$

It is worthwhile examining this last regression in a little more detail. The function $M(\lambda)$ increases at two points, λ and $-\lambda$. The jumps

$$
\begin{aligned}
&\Delta M_{11}(\lambda) = \Delta M_{11}(-\lambda) = \begin{pmatrix} \frac{1}{2} & 0 \\ 0 & 0 \end{pmatrix} \\
(5.21) \qquad & \\
&\Delta M_{12}(\lambda) = \Delta M_{12}(-\lambda) = \begin{pmatrix} 0 & \frac{1}{2} \\ 0 & 0 \end{pmatrix} \\
&\Delta M_{21}(\lambda) = \Delta M_{21}(-\lambda) = \begin{pmatrix} 0 & 0 \\ \frac{1}{2} & 0 \end{pmatrix} \\
(5.22) \qquad & \\
&\Delta M_{22}(\lambda) = \Delta M_{22}(-\lambda) = \begin{pmatrix} 0 & 0 \\ 0 & \frac{1}{2} \end{pmatrix} .
\end{aligned}
$$

Thus

$$
(5.23) \qquad \int_{-\pi}^{\pi} [f(-\lambda) \cdot dM(\lambda)] = \begin{pmatrix} f_{11}(\lambda) & \operatorname{Re} f_{12}(\lambda) \\ \operatorname{Re} f_{12}(\lambda) & f_{22}(\lambda) \end{pmatrix}
$$

and

$$
\begin{aligned}
(5.24) \qquad &\int_{-\pi}^{\pi} [f^{-1}(-\lambda) \cdot dM(\lambda)] \\
&= \frac{1}{f_{11}(\lambda)\, f_{22}(\lambda) - |f_{12}(\lambda)|^2} \begin{pmatrix} f_{22}(\lambda) & -\operatorname{Re} f_{12}(\lambda) \\ -\operatorname{Re} f_{12}(\lambda) & f_{11}(\lambda) \end{pmatrix} .
\end{aligned}
$$

It is worthwhile noting that one does have asymptotic efficiency of the least-squares estimate if $\operatorname{Im} f_{12}(\lambda) = 0$.

In the multidimensional case $k \geq 2$ set

$$
(5.25) \qquad N(\lambda) = \{T^{-1/2}\, {}_{jl}M(\lambda)\, T^{-1/2} ; \qquad j, l = 1, \cdots, s\} .
$$

Then equation (5.1) can be rewritten

$$(5.26)\qquad \int_{-\pi}^{\pi}[f(-\lambda)\cdot dN(\lambda)]\int_{-\pi}^{\pi}[f^{-1}(-\lambda)\cdot dN(\lambda)] = I.$$

I have not been able to get simple necessary and sufficient conditions on $N(\lambda)$ for equation (5.26) to be satisfied for all admissible $f(\lambda)$. However, one can get simple conditions of this type when $N(\lambda)$ is known to increase only at zero. This corresponds to the interesting case of a polynomial regression. Equation (5.26) can then be written as

$$(5.27)\qquad (f\cdot N)(f^{-1}\cdot N) = I,$$

where

$$(5.28)\qquad f = f(0) > 0$$

and

$$(5.29)\qquad N = \Delta N \geqq 0.$$

Set

$$(5.30)\qquad N_{ij} = \{{}_{pq}N_{ij};\quad p, q = 1, \cdots, s\}$$

and

$$(5.31)\qquad W_{ij} = N_{ij} + N_{ji}, \qquad i \neq j.$$

Since $f = f(0)$ and $N = \Delta N(0)$, the elements of f and N are all real.

THEOREM 3. *Let f and N be positive definite and positive semidefinite symmetric matrices, respectively. The equation*

$$(5.32)\qquad (f\cdot N)(f^{-1}\cdot N) = I$$

is valid for all positive definite symmetric f if and only if

$$(5.33)\qquad \begin{aligned} &N_{ii}N_{jj} = \delta_{ij}N_{ii}; && i, j = 1, \cdots, k,\\ &\sum N_{ii} = I && \\ &W_{ij}W_{ik} = N_{jj}W_{jk} = W_{jk}N_{kk}, && j \neq k,\\ &W_{ij}^2 = N_{ii} + N_{jj} && \\ &W_{ij}W_{kl} = 0, && i \neq k, l; \quad j \neq k, l;\\ &W_{ij}N_{kk} = N_{kk}W_{ij} = 0, && i, j \neq k. \end{aligned}$$

Consider first the case in which f is a diagonal matrix

$$(5.34)\qquad f = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_k \end{pmatrix}, \qquad \lambda_i > 0.$$

Equation (5.32) then becomes

$$(5.35)\qquad \sum \lambda_i N_{ii} \sum \lambda_j^{-1} N_{jj} = I.$$

This equation is valid for all positive λ_i if and only if

$$N_{ii} N_{jj} = \delta_{ij} N_{ii} \tag{5.36}$$

and $\sum N_{ii} = I$. Since N is positive semidefinite and $N_{ii}N_{jj} = \delta_{ij}N_{ii}$, it follows that

$$W_{ij} N_{kk} = N_{kk} W_{ij} = 0 \tag{5.37}$$

if $i, j \neq k$ and

$$W_{ij} W_{kl} = 0 \tag{5.38}$$

if $i \neq k, l$ and $j \neq k, l$. Now consider the case in which

$$f = \begin{pmatrix} f_1 & f_2 & 0 \\ f_2 & f_3 & \\ 0 & & I \end{pmatrix} \tag{5.39}$$

and is positive definite. On differentiating equation (5.32) with respect to f_2 twice, equation

$$W_{12}^2 = N_{11} + N_{22} \tag{5.40}$$

is obtained. On differentiating equation (5.32) first with respect to f_1 and then with respect to f_2, equation

$$N_{11} W_{12} = W_{12} N_{22} \tag{5.41}$$

is obtained. Now let

$$f = \begin{pmatrix} f_1 & f_2 & f_3 & \\ f_2 & f_4 & f_5 & 0 \\ f_3 & f_5 & f_6 & \\ & 0 & & I \end{pmatrix} \tag{5.42}$$

and be positive definite. Differentiate equation (5.32) with respect to f_1, f_2, and f_5. I now get

$$M_{11} T_{13} = T_{12} T_{23}. \tag{5.43}$$

All the other equations (5.33) are obtained by taking some subscripts other than 1, 2, 3 or interchanging subscripts.

By using the conditions (5.33), one can readily verify that equation (5.32) is satisfied with any positive definite symmetric f.

6. Some computations

It is worthwhile looking at a process with a stationary residual of a special and simple form to see how good the asymptotic theory considered is for finite samples. Consider a process

$$y_t = x_t + \beta_1 + \beta_{2t} \tag{6.1}$$

where the residual x_t is a first order autoregressive scheme with covariances

$$r_n = \frac{[\text{illegible}]\,\rho^{|n|}}{1 - \rho^2}, \qquad -1 < \rho < 1. \tag{6.2}$$

Here ρ is the correlation coefficient of the scheme. The regression coefficients β_1, β_2 of the scheme are unknown and to be estimated. The covariance matrices of the least-squares estimate β_L and the Markov estimate β_M are considered. I have already noted that

$$E(\beta_L-\beta)(\beta_L-\beta)'=(\varphi'\varphi)^{-1}\varphi' R\varphi(\varphi'\varphi)^{-1} \tag{6.3}$$

and

$$E(\beta_M-\beta)(\beta_M-\beta)'=(\varphi' R^{-1}\varphi)^{-1}. \tag{6.4}$$

Here when the sample size is n

$$\varphi'=\begin{pmatrix}1 & 1\cdots 1\\ 1 & 2\cdots n\end{pmatrix} \tag{6.5}$$

so that

$$(\varphi'\varphi)^{-1}=\begin{pmatrix}\dfrac{2(2n+1)}{n(n-1)} & \dfrac{-6}{n(n-1)}\\ \dfrac{-6}{n(n-1)} & \dfrac{12}{n(n^2-1)}\end{pmatrix}. \tag{6.6}$$

Straightforward but tedious manipulations lead to

$$\varphi' R^{-1}\varphi=\begin{pmatrix}(n-2)(1-\rho)^2+2(1-\rho) & \dfrac{(n-1)n}{2}(1-\rho)^2+(n+\rho)(1-\rho)\\ \dfrac{(n-1)n}{2}(1-\rho)^2+(n+\rho)(1-\rho) & \dfrac{(n-1)n(2n-1)}{6}(1-\rho)^2+n^2(1-\rho)-\rho^2+n\rho\end{pmatrix}. \tag{6.7}$$

One can similarly show that

$$\varphi' R\varphi=\begin{pmatrix}\dfrac{n}{(1-\rho)^2}-\dfrac{2\rho(1-\rho^n)}{(1-\rho^2)(1-\rho)^2} & \dfrac{n(n+1)}{2(1-\rho)^2}-\dfrac{(n+1)\rho(1-\rho^n)}{(1-\rho^2)(1-\rho)^2}\\ \dfrac{n(n+1)}{2(1-\rho)^2}-\dfrac{(n+1)\rho(1-\rho^n)}{(1-\rho^2)(1-\rho)^2} & \dfrac{n(n+1)(2n+1)}{6(1-\rho)^2}-\dfrac{n(n+1)\rho}{(1+\rho)(1-\rho)^3}-\dfrac{2\rho^{n+2}(n+1)}{(1+\rho)(1-\rho)^4}+\dfrac{2\rho^2(1-\rho^{n+1})}{(1+\rho)(1-\rho)^5}\end{pmatrix}. \tag{6.8}$$

The covariance matrix as given by the asymptotic theory is

$$\frac{1}{(1-\rho)^2}\begin{pmatrix}\dfrac{4}{n} & -\dfrac{6}{n^2}\\ \dfrac{-6}{n^2} & \dfrac{12}{n^3}\end{pmatrix} \tag{6.9}$$

to the first order.

The (i,j)th elements of the covariance matrices of both the least-squares and Markov estimates, $i,j=1,2$, are given in Table I for the sample sizes $n=10, 15, 20, 50$ and correlation coefficients $\rho=-.8, -.6, \cdots, .8$. The approximation suggested by asymptotic theory is also given.

TABLE I

COVARIANCE MATRICES OF THE LEAST-SQUARES AND MARKOV ESTIMATES (AND AN ASYMPTOTIC APPROXIMATION OF THE COVARIANCE MATRICES) OF A LINEAR REGRESSION, RESIDUAL FIRST-ORDER AUTOREGRESSIVE

ρ		Matrix Elements (1, 1)	(1, 2)=(2, 1)	(2, 2)
		$n=10$		
+.2	(a)	0.65196	−0.091312	.016602
	(b)	0.64406	−0.090046	.016372
	(c)	0.62500	−0.093750	.018750
−.2	(a)	0.35898	−0.052116	.0094753
	(b)	0.34109	−0.049425	.0089863
	(c)	0.27778	−0.041666	.0083333
+.4	(a)	0.99189	−0.13464	.024880
	(b)	0.94826	−0.12785	.023245
	(c)	1.11111	−0.16667	.033333
−.4	(a)	0.29685	−0.043812	.0079656
	(b)	0.27749	−0.040614	.0073844
	(c)	0.20408	−0.030612	.0061224
+.6	(a)	1.69108	−0.21501	.039090
	(b)	1.54892	−0.19421	.035311
	(c)	2.50000	−0.37500	.075000
−.6	(a)	0.27142	−0.040923	.0074402
	(b)	0.22345	−0.032950	.0059909
	(c)	0.15625	−0.023438	.0046876
+.8	(a)	3.48150	−0.35878	.065229
	(b)	3.17040	−0.32391	.058893
	(c)	10.00000	−1.50000	.30000
−.8	(a)	0.32541	−0.051326	.0093318
	(b)	0.18379	−0.027257	.0049559
	(c)	0.12346	−0.018518	.0037037
		$n=15$		
+.2	(a)	0.42883	− .040942	.0051178
	(b)	0.42456	− .040469	.0050587
	(c)	0.41667	− .041667	.0055556
−.2	(a)	0.21958	− .021498	.0026873
	(b)	0.21716	− .021227	.0026534
	(c)	0.18519	− .018519	.0024692
+.4	(a)	0.68983	− .064546	.0080684
	(b)	0.66268	− .061576	.0076970
	(c)	0.74075	− .074075	.0098767
−.4	(a)	0.17508	− .017363	.0021704
	(b)	0.16642	− .016383	.0020478
	(c)	0.13606	− .013606	.0018141
+.6	(a)	1.29297	− .11604	.014505
	(b)	1.18142	− .10428	.013034
	(c)	1.66669	− .16669	.022223

(a) Least-squares, (b) Markov, and (c) Asymptotic.

ρ		Matrix Elements (1, 1)	(1, 2)=(2, 1)	(2, 2)
		n=15—*Continued*		
−.6	(a)	0.15221	− .015363	.0019204
	(b)	0.13159	− .013023	.0016278
	(c)	0.10417	− .010417	.0013889
+.8	(a)	3.16763	− .24715	.030894
	(b)	2.76763	− .21008	.026261
	(c)	6.66675	− .66675	.088890
−.8	(a)	0.15817	− .016409	.0020513
	(b)	0.10665	− .010598	.0013247
	(c)	0.082305	− .0082305	.0010974
		n=20		
+.2	(a)	0.31940	− .023130	.0022033
	(b)	0.31679	− .022912	.0021821
	(c)	0.31250	− .023438	.0023438
−.2	(a)	0.15777	− .011650	.0011097
	(b)	0.15640	− .011532	.0010983
	(c)	0.13889	− .010417	.0010417
+.4	(a)	0.52774	− .037659	.0035871
	(b)	0.50996	− .036167	.0034445
	(c)	0.55556	− .041667	.0041667
−.4	(a)	0.12341	− .0092078	.00087707
	(b)	0.11854	− .0087887	.00083702
	(c)	0.10204	− .0076530	.00076530
+.6	(a)	1.03889	− .071960	.0068545
	(b)	0.95601	− .065169	.0062066
	(c)	1.25000	− .093750	.0093750
−.6	(a)	0.10438	− .0079058	.00075303
	(b)	0.092934	− .0069183	.00065888
	(c)	0.078126	− .0058595	.00058595
+.8	(a)	2.84008	− .17755	.016913
	(b)	2.45270	− .14856	.014148
	(c)	5.00000	− .37500	.037500
−.8	(a)	0.10458	− .0081673	.00077794
	(b)	0.074810	− .0055867	.00053206
	(c)	0.061728	− .0046296	.00046296
		n=50		
+.2	(a)	0.12612	− .0037310	.000146311
	(b)	0.12565	− .0037140	.000145648
	(c)	0.12500	− .0037500	.000150000
−.2	(a)	0.058456	− .0017433	.000068360
	(b)	0.058229	− .0017352	.000068045
	(c)	0.055555	− .0016667	.000066666
+.4	(a)	0.21807	− .0064149	.000251554
	(b)	0.21442	− .0062866	.000246535
	(c)	0.2222	− .0066667	.000266667

TABLE I—*Continued*

ρ		Matrix Elements (1, 1)	(1, 2) = (2, 1)	(2, 2)
		$n=50$—*Continued*		
−.4	(a)	0.044081	−.0013209	.000051799
	(b)	0.043302	−.0012933	.000050719
	(c)	0.040816	−.0012245	.000048979
+.6	(a)	0.46697	−.013595	.00053312
	(b)	0.44575	−.012856	.00050415
	(c)	0.50000	−.015000	.00060000
−.6	(a)	0.035245	−.0010643	.000041737
	(b)	0.033457	−.0010010	.000039256
	(c)	0.031250	−.00093751	.000037500
+.8	(a)	1.61368	−.045419	.0017810
	(b)	1.43485	−.039365	.0015437
	(c)	2.00000	−.060000	.0024000
−.8	(a)	0.031137	−.00095752	.000037549
	(b)	0.026625	−.00079766	.000031281
	(c)	0.024691	−.00074074	.000029629

7. Some special examples

In section 5 a few special but interesting types of regression sequences were considered where the least-squares estimate of the regression coefficient was not asymptotically efficient in the class of linear unbiased estimates. I now consider two of these regression sequences to find out what information about the spectrum $f(\lambda)$ is required to construct an estimate of the regression coefficient with the same asymptotic mean square error as the Markov estimate.

The first example is that of a one-dimensional process

$$y_t = x_t + \beta\varphi_t \tag{7.1}$$

where x_t is stationary and

$$\varphi_t = a_0 + a_1 \cos t\lambda_1 + \cdots + a_k \cos t\lambda_k\,, \qquad 0 < \lambda_1 < \cdots < \lambda_k\,. \tag{7.2}$$

Note that

$$\Phi_n = \sum_{t=1}^{n} \varphi_t^2 \sim n\left(a_0^2 + \frac{1}{2}\sum_{j=1}^{k} a_j^2\right). \tag{7.3}$$

An estimate of β which has the same asymptotic mean square error as the Markov estimate is

$$\beta^* = \frac{1}{n}\left(\frac{a_0^2}{f(0)} + \frac{1}{2}\sum_{j=1}^{k}\frac{a_j^2}{f(\lambda_j)}\right)\sum_{t=1}^{n} y_t\left(\frac{a_0}{f(0)} + \sum_{j=1}^{k}\frac{a_j}{f(\lambda_j)}\cos t\lambda_j\right) \tag{7.4}$$

for

$$E\beta^* = \beta + O\left(\frac{1}{n}\right) \tag{7.5}$$

and

(7.6) $$\sigma^2(\beta^*) \sim \frac{1}{n}\left(\frac{a_0^2}{f(0)} + \frac{1}{2}\sum_{j=1}^{k} \frac{a_j^2}{f(\lambda_j)}\right)^{-1}.$$

Notice that the only information about the spectrum $f(\lambda)$ required for the construction of β^* is knowledge of the ratios

(7.7) $$\frac{f(0)}{f(\lambda_1)}, \ldots, \frac{f(0)}{f(\lambda_k)}.$$

The second example is a two-dimensional process

(7.8) $$y_t = x_t + \beta\varphi$$

with x_t stationary and

(7.9) $$\varphi_t = \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

An estimate of β which has the same asymptotic mean square error as the Markov estimate is

(7.10) $$\beta^* = \frac{1}{n}\sum_{t=1}^{n}\left({}_1y_t - \frac{f_{12}(0)}{f_{22}(0)}\,{}_2y_t\right)$$

where ${}_1y_t$, ${}_2y_t$ are the components of y_t. Note that

(7.11) $$E\beta^* = \beta$$

and

(7.12) $$\sigma^2(\beta^*) \sim \frac{1}{n}\left[f_{11}(0) - \frac{f_{12}^2(0)}{f_{22}(0)}\right].$$

8. Final remarks

There are many interesting open problems. It is clear that one ought to be able to obtain analogues of the results obtained thus far in the case of a continuous time parameter. It is likely that such a program would require heavier tools.

The results obtained thus far have an immediate implication for various types of nonstationary processes, specifically processes which are integrals or sums of stationary processes. Consider as an example

(8.1) $$Z_t = \sum_{\tau=1}^{t} x_\tau$$

where x_t is a stationary process. Results on estimation of regression coefficients with Z_t as a residual can be obtained from corresponding results with x_t as a residual.

A much more detailed investigation of specific types of regression sequences would be worthwhile pushing through. It is worthwhile noting that all the main results obtained can be derived for processes with a vector time parameter in the same way.

I wish to thank M. de Groot and the computing staff of the Statistical Research Center at the University of Chicago for carrying out the computations in section 6.

REFERENCES

[1] U. GRENANDER, "On the estimation of regression coefficients in the case of an autocorrelated disturbance," *Annals of Math. Stat.*, Vol. 25 (1954), pp. 252–272.
[2] U. GRENANDER and M. ROSENBLATT, "An extension of a theorem of G. Szegö and its application to the study of stochastic processes," *Trans. Amer. Math. Soc.*, Vol. 76 (1954), pp. 112–126.
[3] ———, "Regression analysis of time series with stationary residuals," *Proc. Nat. Acad. Sci.*, Vol. 40 (1954), pp. 812–816.
[4] ———, A monograph on time series analysis, to be published.
[5] M. ROSENBLATT, "On the estimation of regression coefficients of a vector-valued time series with a stationary residual," to be published in *Annals of Math. Stat.*