EFFICIENT ESTIMATION OF A REGRESSION PARAMETER FOR CERTAIN SECOND ORDER PROCESSES[1]

By Charlotte T. Striebel

Lockheed Missiles and Space Division, Sunnyvale, California

0. Summary. The problem of estimation of a single regression parameter for a process with fixed known regression function and unknown covariance is attacked using a Hilbert space representation of the process. Some general results are obtained which characterize efficiency classes of covariances—that is, classes for each of which there exists a single estimate that is efficient for all members. These results are applied to both the discrete parameter and the continuous parameter stationary process with rational spectral density. Some special results are also obtained concerning the efficiency of the least square estimate.

1. Introduction. Let $x(t)$ be a second order complex-valued process with mean value function zero and covariance

$$E[x(t)\overline{x(s)}] = R(t, s), \tag{1.1}$$

and suppose that the process

$$y(t) = k\varphi(t) + x(t) \tag{1.2}$$

is observed for the parameter t in a subset C^T of the real line. The function $\varphi(t)$ is known, and the parameter k is to be estimated. The subsets of interest will be the intervals $(-\infty < t \leqq T)$ and $(0 \leqq t \leqq T)$ for the continuous parameter process and the integers $(t = T, T - 1, \cdots)$ and $(t = T, T - 1, \cdots, 0)$ for the discrete parameter process.

A linear unbiased estimate with finite variance will be represented as a linear functional

$$\bar{k}^T = \bar{k}^T[y(t), t \varepsilon C^T], \tag{1.3}$$

which is the limit in quadratic mean of unbiased finite linear combinations of the $y(t)$ process, that is,

$$\sum_{i=1}^{M_m} k_{im}^T y(t_{im}^T) \xrightarrow{\text{q.m.}} \bar{k}^T \qquad \text{as } m \to \infty, \tag{1.4}$$

where

$$t_{im}^T \varepsilon C^T \tag{1.5}$$

and

$$\sum_{i=1}^{M_m} k_{im}^T \varphi(t_{im}^T) = 1. \tag{1.6}$$

Received October 31, 1960; revised May 30, 1961.

[1] This paper was prepared under the sponsorship of the Lockheed General Research Program

The limit $\bar{k}^T$ of (1.4) is a random variable with finite variance. It can be thought of as an element of the L_2 space over the underlying probability space, it can be made to correspond to an element in the reproducing kernel Hilbert space defined by the kernel $R(s, t)$ (see Parzen [10]), or a correspondence can be set up with elements of another L_2 space as will be done in Section 2. However, it seems more appropriate to use the notation of a linear functional (1.3), since an estimate must finally be reduced to this form so that it can be applied to elements $y(t)$ of the sample space. Thus the notation $\bar{k}^T$ will always refer to a particular sequence of coefficients $\{k_{im}^T\}$ and time points $\{t_{im}^T\}$ satisfying (1.4)–(1.6), and the expression $\bar{k}^T[f(t), t \varepsilon C^T]$ will indicate the limit in the topology of the range space of $f(t)$ of the sums $\sum_i k_{im}^T f(t_{im}^T)$ provided this limit exists.

Since only linear unbiased estimates will be considered, and the criterion by which an estimate will be judged is its variance, it is clear that only second order properties are involved, so that for these purposes the estimation problem is completely determined by the pair (R, φ). An estimate $\bar{k}^T$ is said to be *asymptotically efficient* or simply *efficient* for the problem (R, φ) provided

$$E(T) = \frac{\text{variance } \hat{k}^T}{\text{variance } \bar{k}^T} \to 1 \qquad \text{as } T \to \infty, \tag{1.7}$$

where $\hat{k}^T$ is the minimum variance unbiased estimate of k for the process (1.2) with $t \varepsilon C^T$. $E(T)$ will be called the efficiency for the problem (R, φ).

Interest in efficient estimates arises from the fact that the "best" estimate $\hat{k}^T$ may be very inconvenient. This estimate is determined by the linear equation

$$\hat{k}^T[R(t, s), t \varepsilon C^T] = M^T\overline{\varphi(s)}, s \varepsilon C^T, \tag{1.8}$$

where M^T is a constant. For many problems of interest, the solution to this equation is difficult to exhibit explicitly, and provided it can be computed at all, it will depend on complete knowledge of $R(t, s)$. Thus, if the function $\varphi(t)$, which will be called the *regression function*, is known, but information concerning the covariance is limited or can be obtained only at considerable expense, it is desirable to find an estimate that is economical of information concerning $R(t, s)$ in that it is efficient for as wide a class of covariance functions as possible.

The principal estimate that has been proposed is the least square estimate given, for example, by

$$k_L^T = \int_0^T \overline{\varphi(t)} y(t)\, dt \Big/ \int_0^T |\varphi(t)|^2\, dt \tag{1.9}$$

for the case $C^T = (0 \leqq t \leqq T)$. This estimate has the advantages that it is easy to compute and requires no knowledge whatever of the covariance. Previous work on the problem of efficient estimates has been restricted to stationary processes, that is, $R(t, s) = R(t - s)$, and has been primarily devoted to determining those combinations (R, φ) for which the least square estimate is efficient.

For the continuous parameter Ornstein-Uhlenbeck process,

$$R(\tau) = e^{-\beta|\tau|}, \tag{1.10}$$

and for regression functions

$$\varphi(t) = t^r \quad \text{or} \quad e^{i\lambda_0 t}, \tag{1.11}$$

where r is a non-negative integer and λ_0 is a real frequency, Mann and Moranda [9] proved that the least square estimate is efficient. The author in [13] extended this result to include regression functions of the form

$$\varphi(t) = t^r e^{i\lambda_0 t} \tag{1.12}$$

and showed further that for the more general function,

$$\varphi(t) = \sum_{\alpha=1}^{n} \varphi_\alpha t^r e^{i\lambda_\alpha t}, \tag{1.13}$$

where the φ_α are non-zero constants, the λ_α are real and distinct, and $n > 1$, the least square estimate is not efficient.

For a much broader class of covariance function and essentially the same regression functions, this problem was first discussed by Grenander in [2]. Further work was carried out by Grenander and Rosenblatt in [3] and [4]. Rosenblatt considered some of the same problems in the case of vector-valued time series in [11] and extended his results in [12]. Most of these results, together with some examples, appear in Chapter 7 of [5]. In this work only the discrete parameter case is considered, and the regression functions considered are slightly more general than those of the form (1.13). All restrictions on the class of covariances are imposed on the equivalent class of spectral densities $f(\lambda)$, which by assumption exist and satisfy the relation

$$R(t) = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{i\lambda t} f(\lambda)\, d\lambda \tag{1.14}$$

for a discrete parameter process and

$$R(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{i\lambda t} f(\lambda)\, d\lambda \tag{1.15}$$

for a continuous parameter process. In the discrete parameter case for positive continuous spectral density and "slowly increasing" regression function, a necessary and sufficient condition is given in [5] for the least square estimate to be efficient. The same theorem is obtained in [13] for the continuous parameter Ornstein-Uhlenbeck process and regression function of the form (1.13). Theorem 4 in Section 3 extends this result to the continuous parameter processes with rational spectral density.

In Chapter 1.3 of [6] Grenander and Szegö reproduce a few of the results of [5] using the methods of Toeplitz forms. In Chapter 1.4, under certain regularity conditions on $f(\lambda)$, he extends his results to the continuous parameter case for the single example

$$\varphi(t) = 1. \tag{1.16}$$

With the exception of those in [6], all the above-mentioned results are derived for the more general problem

$$E[y(t)] = \sum_{i=1}^{p} k_i \varphi_i(t), \tag{1.17}$$

where the k_i are unknown parameters and the $\varphi_i(t)$ are known functions. For $p > 1$, the definition of efficiency used by Mann, Moranda, and Striebel is different from that used by Rosenblatt and Grenander. For the case $p = 1$, both agree with definition (1.7) made above. In the present paper only the case $p = 1$ will be considered though it is believed that the results obtained could be generalized to larger values of p.

In Section 2, for a rather broad class of processes, necessary and sufficient conditions are given for the existence of an estimate that is efficient for two problems (R_1, φ) and (R_2, φ). When such an estimate exists, it will be said that R_1 and R_2 are *efficiency equivalent*. In Section 3 these results are applied to the problem of a stationary process with rational spectral density and regression function (1.13) where the λ_α are complex with $\mathscr{I}\lambda_\alpha = -a \leqq 0$. Both the continuous and discrete cases are considered.

2. Efficiency equivalence. It will be assumed that $\varphi(t)$ and $R(t, s)$ can be represented as follows:

$$R(t, s) = \int_\Lambda \xi(t, \lambda)\overline{\xi(s, \lambda)}\, dF(\lambda), \tag{2.1}$$

$$\varphi(t) = \int_\Lambda \xi(t, \lambda)\overline{\Phi^T(\lambda)}\, dF(\lambda), \qquad t \in C^T, \tag{2.2}$$

where $\xi(t, \lambda)$ is a complex-valued measurable function on $R \times R$, the set

$$\Lambda = \bigcup_T \bigcup_{t \in C^T} (\lambda \mid \xi(t, \lambda) \neq 0) \tag{2.3}$$

is measurable, F is a measure on the subspace $(\Lambda, \mathcal{B})$ of the reals, and $\Phi^T(\lambda)$ is in the linear span $L^T(F)$ of $\{\xi(t, \lambda),\ t \in C^T\}$ in the Hilbert space $L_2(\Lambda, \mathcal{B}, F)$. Under these assumptions it follows that to each unbiased linear estimate $\bar{k}^T$ with finite variance there corresponds an element $n^T(\lambda)$ in the subspace $L^T(F)$ such that

$$n^T(\lambda) = \bar{k}^T[\xi(\lambda, t),\ t \in C^T], \tag{2.4}$$

$$(\Phi^T, n^T) = 1, \tag{2.5}$$

and

$$\text{variance } \bar{k}^T = (n^T, n^T). \tag{2.6}$$

The function $n^T(\lambda)$ corresponding to $\bar{k}^T$ is unique a.s. F. A minimum variance unbiased estimate $\hat{k}^T$ exists,

$$\frac{\Phi^T(\lambda)}{(\Phi^T, \Phi^T)} = \hat{k}^T[\xi(t, \lambda),\ t \in C^T] \tag{2.7}$$

and

$$\text{variance } \hat{k}^T = 1/(\Phi^T, \Phi^T). \tag{2.8}$$

These results are fairly standard and can be obtained for example, from more general results by Parzen [10].

The cases which will be considered in the next section are

$$\xi(t, \lambda) = (2\pi)^{-\frac{1}{2}} e^{it\lambda},$$

$\Lambda = [-\pi, \pi]$ for the discrete and $\Lambda = (-\infty, \infty)$ for the continuous parameter stationary process. The solution $\Phi^T(\lambda)$ of the equation (2.2) will be found by the Wiener-Hopf technique for C^T half-infinite.

Let F_i be measures for which there exists Φ_i^T satisfying (2.2). Consider $n_{ij}^T(\lambda)$ in $L^T(F_i)$ which corresponds to an unbiased estimate $\bar{k}_j$ for the problem (R_i, φ). The following measures can then be defined:

$$\mu_i^T(B) = \int_B |\Phi_i^T(\lambda)|^2 \, dF_i(\lambda) \Big/ \int |\Phi_i^T(\lambda)|^2 \, dF_i(\lambda), \tag{2.9}$$

$$\nu_{ij}^T(B) = \int_B |n_{ij}^T(\lambda)|^2 \, dF_i(\lambda) \Big/ \int |n_{ij}^T(\lambda)|^2 \, dF_i(\lambda). \tag{2.10}$$

The first subscript on n_{ij}^T will be omitted when it is clear what problem (F_i, φ) is intended. The efficiency $E_{ij}(T)$ for the estimate $n_j^T(\lambda)$ for the problem (F_i, φ) is given by

$$\frac{1}{E_{ij}(T)} = \int |n_j^T(\lambda)|^2 \, dF_i(\lambda) \int |\Phi_i^T(\lambda)|^2 \, dF_i(\lambda). \tag{2.11}$$

LEMMA 1. *If $n_j^T(\lambda)$ is unbiased and efficient for (F_i, φ), then*

$$|\mu_i^T(B) - \nu_{ij}^T(B)| \to 0 \quad \text{as} \quad T \to \infty$$

uniformly for $B \in \mathcal{B}$.

PROOF. The subscripts will be omitted in the proof. Let

$$a_T = \left[\int |n^T(\lambda)|^2 \, dF(\lambda)\right]^{\frac{1}{2}}, \qquad b_T = [|\Phi^T(\lambda)|^2 \, dF(\lambda)]^{\frac{1}{2}};$$

then

$$\begin{aligned}
|\mu^T(B) - \nu^T(B)| &\leqq \int \left| \frac{|\Phi^T(\lambda)|^2}{b_T^2} - \frac{|n^T(\lambda)|^2}{a_T^2} \right| dF(\lambda) \\
&\leqq \int \left| \frac{\Phi^T(\lambda)}{b_T} - \frac{n^T(\lambda)}{a_T} \right| \left| \frac{\Phi^T(\lambda)}{b_T} + \frac{n^T(\lambda)}{a_T} \right| dF(\lambda) \\
&\leqq \left\{ \int \left| \frac{\Phi^T(\lambda)}{b_T} - \frac{n^T(\lambda)}{a_T} \right|^2 dF(\lambda) \int \left| \frac{\Phi^T(\lambda)}{b_T} + \frac{n^T(\lambda)}{a_T} \right|^2 dF(\lambda) \right\}^{\frac{1}{2}} \\
&= \left\{ \left(\frac{b_T^2}{b_T^2} + \frac{a_T^2}{a_T^2} - \frac{2\mathcal{R}(\Phi^T, n^T)}{a_T b_T} \right) \left(\frac{b_T^2}{b_T^2} + \frac{a_T^2}{a_T^2} + 2\mathcal{R}(\Phi^T, n^T) \right) \right\}^{\frac{1}{2}}.
\end{aligned}$$

The first inequality takes the absolute value under the integral; the second uses the elementary inequality

$$||a|^2 - |b|^2| \leqq |a - b|\,|a + b|;$$

and the third is the Schwarz inequality. Since the estimate is unbiased,

$$(\Phi^T, n^T) = 1,$$

and, since it is efficient,

$$1/E(T) = (n^T, n^T)(\Phi^T, \Phi^T) = a_T^2 b_T^2 \to 1.$$

LEMMA 2. *Let n_0^T and n_1^T be unbiased and efficient for (R, φ). If ν_{11}^T converges weakly to a measure N_{11},*

$$\nu_{11}^T \xrightarrow{w} N_{11},$$

then ν_{10}^T also converges weakly to that measure,

$$\nu_{10}^T \xrightarrow{w} N_{11}.$$

Complete convergence of ν_{11}^T

$$\nu_{11}^T \xrightarrow{c} N_{11}$$

implies complete convergence of ν_{10}^T

$$\nu_{10}^T \xrightarrow{c} N_{11}.$$

The terms weak and strong convergence are according to Loève [8]. This lemma is immediate from Lemma 1.

When it is said that $\bar{k}_0^T$ or n_0^T is an estimate for two measures F_1 and F_2, the following is intended: there are sequences $\{k_{im}^T\}$ and $\{t_{im}^T\}$ satisfying (1.4)–(1.6) where convergence is quadratic mean in (1.4) holds for both R_1 and R_2 or equivalently the sequence of functions

$$n_m^T(\lambda) = \sum_i k_{im}^T \xi(t_{im}, \lambda) \tag{2.12}$$

converges to $n_0^T(\lambda)$ in $L_2(F_1)$ norm and in $L_2(F_2)$ norm.

THEOREM 1. (i) *Let A be a countable union of intervals on which $dF_2(\lambda)/dF_1$ exists and is continuous except for a countable number of discontinuities. Consider a sequence T for which the following are satisfied.* (ii) *There exist estimates $n_i^T(\lambda)$ unbiased and efficient for (F_i, φ) $i = 1, 2$ for which*

$$\nu_{ii}^T \xrightarrow{w} N_{ii}, \quad i = 1, 2 \quad \text{and} \quad N_{22}(A) \neq 0.$$

Then if there exists an estimate $n_0^T(\lambda)$ that is unbiased and efficient for F_1 and F_2, it follows that N_{11} and N_{22} must satisfy the following condition: (iii) *For all $B \in \mathfrak{B}$*

$$\int_{B \cap A} \frac{dF_2(\lambda)}{dF_1}\, dN_{11}(\lambda) = cN_{22}(B \cap A) \tag{2.13}$$

where

$$(2.14)\qquad c = \lim_{T\to\infty} \frac{\int |n_2^T(\lambda)|^2 \, dF_2(\lambda)}{\int |n_1^T(\lambda)|^2 \, dF_1(\lambda)}.$$

PROOF. Let $(a, b) = B^*$ be an interval contained in A on which $dF_2(\lambda)/dF_1$ is continuous, then

$$\int_{B^*} \frac{dF_2(\lambda)}{dF_1} \, d\nu_{10}^T(\lambda) = c(T)\nu_{20}^T(B^*),$$

where

$$c(T) = \frac{\int |n_0^T(\lambda)|^2 \, dF_2}{\int |n_0^T(\lambda)|^2 \, dF_1} = \frac{E_{10}(T)E_{22}(T)}{E_{20}(T)E_{11}(T)} \frac{\int |n_2^T(\lambda)|^2 \, dF_2(\lambda)}{\int |n_1^T(\lambda)|^2 \, dF_1(\lambda)}.$$

From Lemma 2 and (ii), since n_0^T is also efficient for F_i,

$$\nu_{i0}^T \xrightarrow{w} N_{ii}, \qquad i = 1, 2.$$

By the Helly-Bray Lemma

$$\int_{B^*} \frac{dF_2(\lambda)}{dF_1} \, d\nu_{10}^T(\lambda) \to \int_{B^*} \frac{dF_2(\lambda)}{dF_1} \, dN_{11}(\lambda).$$

Since $N_{22}(A) \neq 0$, there exists an interval B^* in A such that $N_{22}(B^*) \neq 0$, thus

$$\frac{\int |n_2^T(\lambda)|^2 \, dF_1}{\int |n_1^T(\lambda)|^2 \, dF_2} = \frac{c(T)E_{10}(T)E_{22}(T)}{E_{20}(T)E_{11}(T)} \to c = \frac{\int_{B^*} \frac{dF_2(\lambda)}{dF_1} \, dN_{11}(\lambda)}{N_{22}(B^*)}.$$

The measurable sets in A are generated by intervals of this type, so (2.13) must also hold for all $B \subset \mathcal{B}$.

THEOREM 2. *If in addition to assumptions* (i)–(iii), $c \neq 0$, $A = \Lambda$, dF_2/dF_1 *is bounded and*

$$\nu_{11}^T \xrightarrow{c} N_{11},$$

then $n_0^T(\lambda)$ *efficient and unbiased for* F_1 *implies* $n_0^T(\lambda)$ *is also an efficient unbiased estimate for* F_2.

PROOF. Let $\{\bar{k}_{im}^T\}_0$, $\{t_{im}^T\}_0$ be a sequence of simple estimates (1.4) which converges to $\bar{k}_0^T$ in quadratic mean, with respect to F_1. Then for the corresponding $\{n_m^T(\lambda)\}$ given by (2.12)

$$\int |n_0^T(\lambda) - n_m^T(\lambda)|^2 \, dF_2 = \int \frac{dF_2(\lambda)}{dF_1} |n_0^T(\lambda) - n_m^T(\lambda)|^2 \, dF_1(\lambda)$$
$$\leqq M \int |n_0^T(\lambda) - n_m^T(\lambda)|^2 \, dF_1(\lambda) \to 0,$$

where M is the bound of dF_2/dF_1. Thus (1.4) also converges to an estimate which corresponds to $n_0^T(\lambda)$ with respect to F_2.

$$\frac{1}{E_{20}(T)} = \frac{E_{11}(T)}{E_{10}(T)E_{22}(T)} \frac{\int |n_1^T(\lambda)|^2 \, dF_1}{\int |n_2^T(\lambda)|^2 \, dF_2} \int \frac{dF_2}{dF_1} \, d\nu_{10}$$
$$\to \frac{1}{c} \int \frac{dF_2(\lambda)}{dF_1} \, dN_{11}(\lambda) = N_{22}(\Lambda) \leqq 1.$$

This depends on Lemma 2 and the Helly-Bray Theorem in Section 11.3 of [8].

3. Rational spectral density and regression function. In this section the discrete and the continuous parameter stationary process will be considered. Thus

$$\xi(t, \lambda) = (2\pi)^{-\frac{1}{2}} e^{i\lambda t}, \tag{3.1}$$

$$\Lambda = [-\pi, \pi] \tag{3.2}$$

for the discrete parameter process, and

$$\Lambda = (-\infty, \infty) \tag{3.3}$$

for the continuous parameter process, and the representation (2.1) is given by (1.14) and (1.15), respectively. The case of C^T half-infinite will be considered first. For the discrete parameter t and T are integers, and

$$C^T = (T, T - 1, \cdots); \tag{3.4}$$

for the continuous parameter

$$C^T = (-\infty, T). \tag{3.5}$$

It will be assumed that the spectral densities $f(z)$ and $f(\lambda)$ are positive rational functions where for convenience in the discrete parameter case the density will be treated as a function of $z = e^{i\lambda}$. The densities can be factored

$$f(z) = |F(z)|^2, \tag{3.6}$$

$$f(\lambda) = |F(\lambda)|^2. \tag{3.7}$$

For the discrete process $F(z)$ is a quotient of two polynominals each of the same degree and having zeros inside the unit circle ($|z| < 1$); for the continuous process $F(\lambda)$ is a proper rational function and has poles and zeros in the upper half-plane ($\mathscr{I}\lambda > 0$). (See Doob [1], p. 502 and p. 542.)

The regression function that will be considered has the form

$$\varphi(t) = \sum_{\gamma=1}^{m} \varphi_\gamma t^{r_\gamma} e^{i\lambda_\gamma t}, \qquad t \geqq 0 \tag{3.8}$$

where λ_γ is complex and

$$\max_\gamma \mathcal{R}(i\lambda_\gamma) = a \geqq 0 \quad \text{and} \quad \varphi_\gamma \neq 0. \tag{3.9}$$

The exact form of $\varphi(t)$ for $t < 0$ will be seen to be immaterial for questions of efficiency as $T \to \infty$. For the discrete parameter case $\varphi(t)$ for $t < 0$ must be such that the sum

$$\Phi(z) = \sum_{t=-\infty}^{\infty} z^{-t}\varphi(t) \tag{3.10}$$

converges to a rational function in a ring

$$a < |z| < b. \tag{3.11}$$

Similarly, in the continuous case the integral

$$\Phi(\lambda) = \int_{-\infty}^{\infty} e^{-i\lambda t}\varphi(t)\, dt \tag{3.12}$$

must converge to a rational function in the strip

$$-b < \mathcal{I}\lambda < -a. \tag{3.13}$$

In this case it will also be assumed that $\Phi(\lambda)/F(\lambda)$ is a proper rational function. For any given degree e and $\varphi(t)$ given by (3.8) for $t \geqq 0$ it is always possible to define $\varphi(t)$ for $t < 0$ so that the degree of denominator of $\Phi(\lambda)$ exceeds that of the numerator by e and hence $\Phi(\lambda)/F(\lambda)$ is proper if the net degree of $1/F(\lambda)$ is less than e. In each case the terms of importance in (3.8) are those for which $\mathcal{I}(i\lambda_\gamma) = a$ and among these the ones for which r_γ is a maximum. The index of these terms will be indicated by $\alpha = 1, \cdots, n$. The functions $\Phi(z)$ and $\Phi(\lambda)$ can then be expanded as follows:

$$\Phi(z) = \sum_{\alpha=1}^{m}\sum_{j=0}^{r} \frac{\Phi_{\alpha j}\, z^{j+1}}{(z - z_\alpha)^{j+1}}, \tag{3.14}$$

where

$$\begin{aligned} z_\alpha &= e^{i\lambda\alpha} & & \\ |z_\alpha| &= a, & \alpha &= 1, \cdots, n, \\ |z_\alpha| &< a \quad \text{or} \quad \geqq b, & \alpha &= n+1, \cdots, m, \end{aligned} \tag{3.15}$$

and

$$\Phi_{\alpha n} = \Phi_\alpha = r!\varphi_\alpha; \tag{3.16}$$

$$\Phi(\lambda) = \sum_{\alpha=1}^{m}\sum_{j=0}^{r} \frac{\Phi_{\alpha j}}{(\lambda - \lambda_\alpha)^{j+1}}, \tag{3.17}$$

$$\begin{aligned} \mathcal{I}\lambda_\alpha &= -a, & \alpha &= 1, \cdots, n, \\ \mathcal{I}\lambda_\alpha &> -a \quad \text{or} \quad \leqq -b, & \alpha &= n+1, \cdots, m, \end{aligned} \tag{3.18}$$

$$\Phi_{\alpha r} = \Phi_\alpha = r!\varphi_\alpha(-i)^{r+1}. \tag{3.19}$$

Equation (2.2) can be written

$$\frac{1}{(2\pi)^{\frac{1}{2}} i} \oint_{|z|=1} z^{t-1} \dot{\Phi}^T(z) |F(z)|^2 \, dz = \varphi(t), \qquad t = T, T-1, \cdots, \tag{3.20}$$

where

$$\overline{\Phi^T(z)} = \sum_{t=-\infty}^{T} z^{-t} k_t , \tag{3.21}$$

and

$$(2\pi)^{-\frac{1}{2}} \int_{-\infty}^{\infty} e^{i\lambda t} \overline{\Phi^T(\lambda)} |F(\lambda)|^2 \, d\lambda = \varphi(t), \qquad -\infty < t \leqq T, \tag{3.22}$$

where

$$\overline{\Phi^T(\lambda)} = k^T[e^{-i\lambda t}, \; -\infty < t \leqq T]. \tag{3.23}$$

Under the assumptions made these equations can easily be solved by the Wiener-Hopf technique. (See, for example, [14] p. 313.) Solutions are given by

$$\overline{\Phi^T(z)} = \frac{1}{i(2\pi)^{\frac{3}{2}} \overline{F(z)}} \sum_{t=-\infty}^{T} z^{-t} \oint_{|w|=e^c} w^{t-1} \frac{\Phi(w)}{F(w)} \, dw \tag{3.24}$$

and

$$\overline{\Phi^T(\lambda)} = \frac{1}{(2\pi)^{\frac{3}{2}} \overline{F(\lambda)}} \int_{-\infty}^{T} e^{-i\lambda t} \int_{-\infty - ic}^{\infty - ic} e^{i\omega t} \frac{\Phi(\omega)}{F(\omega)} \, d\omega \, dt, \tag{3.25}$$

where

$$0 \leqq a < c < b. \tag{3.26}$$

Equation (3.23) can be written as an integral

$$\overline{\Phi^T(\lambda)} = \int_{-\infty}^{T} e^{-i\lambda t} K^T(t) \, dt \tag{3.27}$$

if $K^T(t)$ is permitted to include delta functions and their derivatives. Formulas for the estimates themselves will be given later.

For the case of C^T half-infinite the "best" estimate $\hat{k}^T$, which is clearly efficient, will be considered. For this estimate $n^T(\lambda)$ is given by (2.7). For a given spectrum f_i, the measure N_{ii} and an asymptotic expression for $\int |n_{ii}^T(\lambda)|^2 f_i(\lambda) \, d\lambda$ must be obtained in order to apply the theorems of the previous section. These can be obtained by a straightforward but somewhat lengthy calculation and will be given without proof.

Lemma 3. *For the discrete case, $a > 0$ and a sequence $T \to \infty$ such that $e^{iT\Re\lambda_\alpha} \to l_\alpha$,*

$$\frac{(\Phi^T, \Phi^T)}{e^{2aT} T^{2r}} \to c(F) = \sum_{\alpha=1}^{n} \sum_{\beta=1}^{n} \frac{\varphi_\alpha \bar{\varphi}_\beta l_\alpha \bar{l}_\beta z_\alpha \bar{z}_\beta}{F(z_\alpha) \overline{F(z_\beta)} (z_\alpha \bar{z}_\beta - 1)} > 0, \tag{3.28}$$

and

$$\frac{|\Phi^T(z)F(z)|^2}{(\Phi^T, \Phi^T)} \xrightarrow{c} \frac{1}{(2\pi)^{\frac{1}{2}}c(F)} \left| \sum_{\alpha=1}^{n} \frac{\varphi_\alpha l_\alpha z_\alpha}{F(z_\alpha)(z_\alpha - z)} \right|^2. \tag{3.29}$$

For the discrete case, $a = 0$ and all sequences $T \to \infty$

$$\frac{(\Phi^T, \Phi^T)}{T^{2r+1}} \to c(F) = \frac{1}{2r+1} \sum_{\alpha=1}^{n} \left| \frac{\varphi_\alpha}{F(z_\alpha)} \right|^2 > 0, \tag{3.30}$$

and

$$\frac{1}{2\pi i z} \frac{|\Phi^T(z)F(z)|^2}{(\Phi^T, \Phi^T)} \xrightarrow{c} \frac{1}{c(F)(2r+1)} \sum_{\alpha=1}^{n} \left| \frac{\varphi_\alpha}{F(z_\alpha)} \right|^2 \delta(z - a_\alpha) \tag{3.31}$$

where δ is the Dirac delta function. For the continuous parameter case, $a > 0$ and a sequence $T \to \infty$ for which $e^{iT\Re\lambda_\alpha} \to l_\alpha$

$$\frac{(\Phi^T, \Phi^T)}{e^{2aT}T^{2r}} \to c(F) = \sum_{\alpha=1}^{n} \sum_{\beta=1}^{n} \frac{\varphi_\alpha \bar{\varphi}_\beta l_\alpha \bar{l}_\beta}{F(\lambda_\alpha)\overline{F(\lambda_\beta)}(i\lambda_\alpha - i\bar{\lambda}_\beta)} > 0, \tag{3.32}$$

and

$$\frac{|\Phi^T(\lambda)F(\lambda)|^2}{(\Phi^T, \Phi^T)} \xrightarrow{c} \frac{1}{(2\pi)^{\frac{1}{2}}c(F)} \left| \sum_{\alpha=1}^{n} \frac{\varphi_\alpha l_\alpha}{F(\lambda_\alpha)(\lambda_\alpha - \lambda)} \right|^2. \tag{3.33}$$

For the continuous case, $a = 0$ and all sequences $T \to \infty$

$$\frac{(\Phi^T, \Phi^T)}{T^{2r+1}} \to c(F) = \frac{1}{(2r+1)} \sum_{\alpha=1}^{n} \left| \frac{\varphi_\alpha}{F(\lambda_\alpha)} \right|^2 > 0, \tag{3.34}$$

and

$$\frac{|\Phi^T(\lambda)F(\lambda)|^2}{(\Phi^T, \Phi^T)} \to \frac{1}{c(F)(2r+1)} \sum_{\alpha=1}^{n} \left| \frac{\varphi_\alpha}{F(\lambda_\alpha)} \right|^2 \delta(\lambda - \lambda_\alpha), \tag{3.35}$$

where δ is a delta function.

All integrals involved here can be evaluated by contour integration in the complex plane. Simplifications occur due to the fact that terms contributed by poles at the z_α and λ_α for $\alpha = 1, \cdots, n$ dominate all others of $\Phi(z)$ and $\Phi(\lambda)$ as well as those of $1/F(z)$ and $1/F(\lambda)$.

THEOREM 3.

(i) *For $a > 0$ and a sequence $T \to \infty$ for which $e^{iT\Re\lambda_\alpha} \to l_\alpha$, there exists an estimate efficient for spectral densities f_1 and f_2 if and only if*

$$\left| \frac{F_2(z)}{F_1(z)} \right| = \frac{c(F_1)}{c(F_2)} \left| \frac{\sum_{\alpha=1}^{n} \frac{\varphi_\alpha l_\alpha z_\alpha}{F_2(z_\alpha)(z_\alpha - z)}}{\sum_{\alpha=1}^{n} \frac{\varphi_\alpha l_\alpha z_\alpha}{F_1(z_\alpha)(z_\alpha - z)}} \right| \tag{3.36}$$

for the discrete parameter process, and

$$\left|\frac{F_2(\lambda)}{F_1(\lambda)}\right| = \frac{c(F_1)}{c(F_2)} \left| \frac{\sum_{\alpha=1}^{n} \frac{\varphi_\alpha l_\alpha}{F_2(\lambda_\alpha)(\lambda_\alpha - \lambda)}}{\sum_{\alpha=1}^{n} \frac{\varphi_\alpha l_\alpha}{F_1(\lambda_\alpha)(\lambda_\alpha - \lambda)}} \right| \tag{3.37}$$

for the continuous parameter process.

(ii) *If this condition is satisfied, then any estimate that is efficient for one is also efficient for the other.*

PROOF. Under the assumptions of this section $dF_2(\lambda)/\,dF_1 = f_2(\lambda)/f_1(\lambda)$ is continuous and $A = \Lambda$. For the discrete process $f_2(z)/f_1(z)$ is always bounded; for the continuous process condition (3.37) implies that $f_2(\lambda)/f_1(\lambda)$ is bounded above and away from zero. In both cases $c = c(F_1)/c(F_2) \neq 0$. Thus Theorems 1 and 2 apply. Expressions (3.36) and (3.37) can be obtained directly from (2.13) by substituting the appropriate forms from Lemma 3.

THEOREM 4.

(i) *For $a = 0$ and any sequence $T \to \infty$, there exists an estimate efficient for f_1 and f_2 if and only if*

$$f_2(\lambda_\alpha) = \frac{c(F_1)}{c(F_2)} f_1(\lambda_\alpha) \qquad \alpha = 1, \cdots, n \tag{3.38}$$

for both the discrete and the continuous parameter process.

(ii) *For the discrete parameter process if* (3.38) *is satisfied, then any estimate that is efficient for one is efficient for the other.*

(iii) *For the continuous parameter process if* (3.38) *is satisfied, $\bar{k}_0^T$ is an efficient estimate for f_1, and $f_2(\lambda)/f_1(\lambda)$ is bounded; then $\bar{k}_0^T$ is also efficient for f_2.*

PROOF. As before (3.38) is obtained from (2.13) using Lemma 3, and Theorems 1 and 2 apply.

The stronger result of Theorem 3 (ii) is not true in the case $a = 0$ for the continuous parameter processes, since it is possible to find an efficient estimate for f_1 that depends on derivatives of $y(t)$ which will not exist for f_2 if the degree of $1/f_2$ is less than that of $1/f_1$. This is, of course, the case when f_2/f_1 is unbounded. However, it is possible to find an estimate that is efficient for all f_2 satisfying (3.38). Such an estimate is given by (3.46).

The case of $C^T = (0, 1, \cdots, T)$ and $C^T = (0, T)$ can now be treated easily. Under the assumptions made on f and φ, a solution to the equation (2.2) does exist for both the discrete and the continuous parameter process. (See, for example, Laning and Battin [7], Chapter 8.4.) However, it will not be convenient to use this as the efficient estimates n_{ii}^T required in the theorems of the previous section. Instead, the "best" estimates for the half-infinite interval will be computed and truncated. The estimates obtained in this way are of some interest and will be given explicitly. For $a > 0$ and the discrete case let

$$\frac{1}{(z - e^{i\lambda_\alpha})F(z)} = \sum_{s=1}^{\infty} z^{-s} m_{s-1}^{\alpha}, \tag{3.39}$$

then

$$\bar{M}^T \bar{k}^T = \sum_{t=0}^{T} y(T - t) \sum_{\alpha=1}^{n} \frac{\bar{\varphi}_\alpha e^{-iT\Re\lambda_\alpha} m_t^\alpha}{\overline{F(z_\alpha)}}. \tag{3.40}$$

For the continuous case, let

$$\frac{1}{(\lambda - \bar{\lambda}_\alpha)F(\lambda)} = E_\alpha(\lambda) + M_\alpha(\lambda) \tag{3.41}$$

where $E_\alpha(\lambda)$ is a polynominal

$$E_\alpha(\lambda) = \sum_{j=0}^{e-1} e_j^\alpha \lambda^j \tag{3.42}$$

and $M_\alpha(\lambda)$ is a proper rational function. Let

$$m^\alpha(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{i\lambda t} M_\alpha(\lambda)\, d\lambda, \tag{3.43}$$

then

$$\begin{aligned} \bar{M}^T \bar{k}^T = {} & \sum_{j=0}^{e-1} y^{(j)}(T) \sum_{\alpha=1}^{n} \frac{\bar{\varphi}_\alpha e^{-iT\Re\lambda_\alpha}(-i)^j e_j^\alpha}{\overline{F(\lambda_\alpha)}} \\ & + \int_0^T y(T - t) \sum_{\alpha=1}^{n} \frac{\bar{\varphi}_\alpha e^{-iT\Re\lambda_\alpha} m^\alpha(t)}{\overline{F(\lambda_\alpha)}}\, dt. \end{aligned} \tag{3.44}$$

For $a = 0$

$$\bar{M}^T \bar{k}^T = \sum_{t=0}^{T} y(t) \sum_{\alpha=1}^{n} \frac{\bar{\varphi}_\alpha t^r e^{-i\lambda_\alpha t}}{f(\lambda_\alpha)} \tag{3.45}$$

and

$$\bar{M}^T \bar{k}^T = \int_0^T y(t) \sum_{\alpha=1}^{n} \frac{\bar{\varphi}_\alpha t^r e^{-i\lambda_\alpha t}}{f(\lambda_\alpha)}\, dt. \tag{3.46}$$

In all cases $\bar{M}^T$ is a constant to be determined so that the estimate is unbiased; that is, $\bar{M}^T$ is given by the right side of the expression with $\varphi(t)$ substituted for $y(t)$. A straightforward computation of their variances shows that these estimates are efficient for the half-infinite problem discussed above for all sequences $T \to \infty$. Thus by Lemma 2

$$\nu_{i0}^T \xrightarrow{c} N_{ii}$$

where n_{i0}^T indicates the estimates (3.40), (3.44), (3.45), and (3.46) for f_i, and N_{ii} are the limit measures given in Lemma 3. The asymptotic forms $c(F_i)$ also hold for the $1/\int |n_{i0}^T(\lambda)|^2\, dF_i(\lambda)$ since

$$(\Phi_i^T, \Phi_i^T) = 1/E_{i0}(T)(n_{i0}^T, n_{i0}^T).$$

Thus Theorems 3 and 4 also hold for $t = 0, 1, \cdots, T$ in the discrete case and $0 \leqq t \leqq T$ in the continuous parameter case.

The least square estimate for C^T half-infinite is given by

$$\bar{M}^T \bar{k}_0^T = \sum_{t=-\infty}^{T} y(t)\bar{\varphi}(t) \tag{3.47}$$

and

$$\bar{M}^T \bar{k}_0^T = \int_{-\infty}^{T} y(t)\, \overline{\varphi(t)}\, dt. \tag{3.48}$$

For the discrete parameter case, $F(z) = 1$ provides a bona fide convariance for which $n_0^T(z)$ for the least square estimate is given by (3.24). Thus from Theorems 3 and 4 in this case the least square estimate is efficient for $F(z)$ if and only if (3.36) or (3.38) hold for $F(z) = F_2(z)$ and $F_1(z) = 1$. In the continuous case if the least square estimate is efficient then by Lemma 2 ν_{ii}^T and ν_{i0}^T must converge to the same limit. N_{ii} the limit of ν_{ii}^T is given by (3.33) and (3.35) in Lemma 3. N_{i0} the limit of ν_{i0}^T can be computed by use of Lemma 3 and the Helly-Bray Theorem, since $n_0^T(\lambda)$ is identical with $\overline{\Phi^T(\lambda)}$ except for a constant where $\overline{\Phi^T(\lambda)}$ is given by (3.25) with $F(\lambda) = 1$. For $a > 0$ this limit is

$$N_{i0}(B) = c \int_B f(\lambda) \left| \sum_{\alpha=1}^{n} \frac{\varphi_\alpha l_\alpha}{(\lambda_\alpha - \lambda)} \right|^2 d\lambda, \tag{3.49}$$

and for $a = 0$ by

$$N_{i0}(B) = \frac{1}{\sum_\alpha |\varphi_\alpha|^2 f(\lambda_\alpha)} \int_B \sum_\alpha |\varphi_\alpha|^2 f(\lambda_\alpha)\delta(\lambda - \lambda_\alpha)\, d\lambda. \tag{3.50}$$

Thus for $a = 0$ if the least square estimate is efficient for $f(\lambda)$ it follows that

$$f(\lambda_\alpha) = \text{constant} \qquad \alpha = 1, \cdots, n. \tag{3.51}$$

An asymptotic form for $\int |n_0^T(\lambda)|^2 f(\lambda)\, d\lambda$ can also be found.

$$T^{2r+1} \int |n_0^T(\lambda)|^2 f(\lambda)\, d\lambda \to (2r+1) \sum_{\alpha=1}^{n} |\varphi_\alpha|^2 f(\lambda_\alpha) / [\sum_\alpha |\varphi_\alpha|^2]^2. \tag{3.52}$$

From this and (3.34) of Lemma 3 it is clear that (3.51) is also sufficient. $N_{i0} = N_{ii}$ for $a > 0$ becomes

$$c' f(\lambda) \left| \sum_{\alpha=0}^{n} \frac{\varphi_\alpha l_\alpha}{(\lambda_\alpha - \lambda)} \right|^2 = \left| \sum_{\alpha=1}^{n} \frac{\varphi_\alpha l_\alpha}{F(\lambda_\alpha)(\lambda_\alpha - \lambda)} \right|^2, \tag{3.53}$$

but this is not possible since $f(\lambda)$ must be proper. Thus for $a > 0$ the least square estimate is never efficient.

REFERENCES

[1] J. L. Doob, *Stochastic Processes*, John Wiley and Sons, New York, 1953.

[2] Ulf Grenander, "On the estimation of regression coefficients in the case of an autocorrelated disturbance," *Ann. Math. Stat.*, Vol. 25 (1954), pp. 252–272.

[3] Ulf Grenander and Murray 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.

[4] ULF GRENANDER AND MURRAY ROSENBLATT, "Regression analysis of time series with stationary residuals," *Proc. Nat. Acad. Sci.*, Vol. 40 (1954), pp. 812–816.

[5] ULF GRENANDER AND MURRAY ROSENBLATT, *Statistical Analysis of Stationary Time Series*, John Wiley and Sons, New York, 1957.

[6] ULF GRENANDER AND GABOR SZEGÖ, *Toeplitz Forms and Their Application*, University of California Press, Berkeley, 1958.

[7] J. HALCOMBE LANING, JR. AND RICHARD H. BATTIN, *Random Processes in Automatic Control*, McGraw-Hill Book Co., New York, 1956.

[8] MICHEL LOÈVE, *Probability Theory*, D. van Nostrand, New York, 1955.

[9] HENRY B. MANN AND PAUL B. MORANDA, "On the efficiency of the least square estimates of parameters in the Ornstein Uhlenbeck process," *Sankhyā*, Vol. 13 (1954), pp. 351–358.

[10] EMANUEL PARZEN, "Statistical inference on time series by Hilbert space methods, I," Department of Statistics, Stanford University Technical Report No. 23, January 2, 1959.

[11] MURRAY ROSENBLATT, "On the estimation of regression coefficients of a vector-valued time series with stationary residuals," *Ann. Math. Stat.*, Vol. 27 (1956), pp. 99–121.

[12] MURRAY ROSENBLATT, "Some regression problems in time series analysis," *Proceedings of the Third Berkeley Symposium in Probability and Statistics*, Vol. 1, pp. 165–186, University of California Press, Berkeley, 1956.

[13] CHARLOTTE T. STRIEBEL, "On the efficiency of estimates of trend in the Ornstein Uhlenbeck process," *Ann. Math. Stat.*, Vol. 29 (1958), pp. 192–200.

[14] BALTHASAR VAN DER POL AND H. BREMMER, *Operational Calculus*, Cambridge University Press, Cambridge, 1955.