

## LINEAR ESTIMATION OF REGRESSION COEFFICIENTS FOR VECTOR-VALUED PROCESSES\*

By

ROLF K. ADENSTEDT (*Lehigh University*)

**1. Introduction.** In this paper we generalize to vector-valued random processes some results obtained in [1] for linearly estimating coefficients in regression models. Of special importance is the problem of estimating the mean for a stationary vector-valued random sequence or process.

Let  $T$  be an arbitrary index set and  $X(t)$ ,  $t \in T$ , be an  $r \times 1$  vector-valued random process with known  $r \times r$  matrix-valued covariance function  $R(s, t)$ . We allow the components of  $X(t)$  to be complex-valued.  $X(t)$  has mean value  $\beta(t)c$ , where  $\beta(t)$  is a known  $r \times m$  matrix-valued regression function and  $c$  is an unknown  $m \times 1$  vector of regression constants  $c_1, c_2, \dots, c_m$ . The process is observed for  $t$  in a fixed subset  $S$  of  $T$ . To be considered is the problem of forming unbiased estimators  $\hat{c}_j$  for the  $c_j$  from linear combinations of the components of the observations. Of particular interest for each  $j$  is the best linear unbiased estimator (BLUE)  $\hat{c}_{j, BLUE}$  for  $c_j$ , i.e. that  $\hat{c}_j$  having least variance.

A partial list of related papers in regression analysis was given in [1]. These papers deal primarily with scalar processes, although Rosenblatt [6] considered stationary vector processes in generalizing results of Grenander [3] concerning asymptotic efficiency of the least squares estimator (LSE) for certain regressions in discrete time. Most authors consider the problem of finding the BLUE for special cases of regression models. Instead, we employ a general Hilbert space setting as in [1], without initially restricting the model. The notation and methods of this paper are similar to those in [1], with some minor modifications.

In Sec. 2 we formalize the model and define what is meant by linear unbiased estimators. By a reparametrization we find in Sec. 3 that the model is equivalent to one already studied in [1], and obtain a useful representation of the  $\hat{c}_{j, BLUE}$  and their covariance matrix (Theorem 1). We note also how lower bounds for this covariance matrix may be obtained (Theorem 2). In Secs. 4 and 5 we apply these results to large-sample estimation of the mean for a stationary vector-valued process in discrete and continuous time, respectively. Theorems 3 and 5 show that the covariance matrices of the BLUE and LSE (or sample mean) approach the same limit, while Theorems 4 and 6 show that the LSE is asymptotically efficient relative to the BLUE under fairly general conditions. In Sec. 6 we discuss a more general situation for stationary processes that may be studied by our methods, and cite the results for the simplest case.

**2. Definitions and notation.** We regard our vector process  $X(t)$  as defined on a space  $\Omega$ , which could be taken as sample path space. Letting  $\mu_c$  be a probability measure

\* Received October 22, 1973.

corresponding to the regression vector  $c$ , we describe our model by

$$\int_{\Omega} X(t) d\mu_c = \beta(t)c \quad (1)$$

and

$$\int_{\Omega} [X(s) - \beta(s)c][X(t) - \beta(t)c]' d\mu_c = R(s, t), \quad (2)$$

where the prime denotes Hermitian transposition. As noted,  $X(t)$  is  $r \times 1$ ,  $\beta(t)$  is  $r \times m$ ,  $c$  is  $m \times 1$  and  $R(s, t)$  is  $r \times r$ . We denote the components of these vectors and matrices respectively by  $X_i(t)$ ,  $\beta_{ik}(t)$ ,  $c_i$  and  $R_{ik}(s, t)$  in the usual way.

Of special importance is the measure  $\mu_0$  corresponding to  $c = 0$ , and  $E$  will always represent expectation with respect to  $\mu_0$ . Equalities between and claims of uniqueness for random variables are understood to be modulo  $\mu_0$ . We denote by  $H_X(S)$  the subspace of  $L^2(\mu_0)$  spanned by the  $X_i(t)$  for  $1 \leq j \leq r$  and  $t \in S$ .  $H_X(T)$  is similarly defined.

A *finite* linear estimator has the form

$$\hat{c}_i = \sum_{t=1}^n a_t X(t), \quad t, \in S, \quad (3)$$

where the  $a_t$  are constant  $1 \times r$  row vectors. It is unbiased for  $c_i$  if  $\int \hat{c}_i d\mu_c = c_i$  for all  $c$ , i.e. if

$$\sum_t a_t \beta(t) = e_i, \quad (4)$$

the  $1 \times m$  row vector with entries  $\delta_{ik}$ . From (1) and (2) it is apparent that the variance of such a  $\hat{c}_i$  (whether unbiased or not) is the same under all  $\mu_c$  and is given by  $E |\hat{c}_i|^2$ . Similarly, the covariance between two estimators of the form (3) does not depend on  $c$ .

Letting  $U_i(S)$  be the closure in  $L^2(\mu_0)$  of the set of all finite linear combinations (3) that satisfy (4), we define a linear unbiased estimator for  $c_i$  as any element  $\hat{c}_i$  in  $U_i(S)$ , and its variance by  $E |\hat{c}_i|^2$ . A linear unbiased estimator for  $c$  is an  $m \times 1$  random vector  $\hat{c}$  whose components  $\hat{c}_i \in U_i(S)$  for  $1 \leq j \leq m$ , and we denote its covariance matrix by  $\sigma(\hat{c}) = E \hat{c} \hat{c}'$ .

While some interpretational questions may arise from these definitions, they are easily resolved. The reader is referred to Kuk and Petunin [4] for a discussion of some technical problems that arise in defining estimators.

To guarantee existence of linear unbiased estimators, we assume throughout the sequel that the columns of  $\beta(t)$ , considered as functions on  $S$ , are linearly independent. Then each  $U_i(S)$  is a non-empty closed convex subset of  $H_X(S)$  and contains a unique element  $\hat{c}_{i,BLU}$  of minimum norm  $E |\hat{c}_{i,BLU}|^2$ . The BLUE  $\hat{c}_{BLU}$  for  $c$  is defined as the  $m \times 1$  vector having components  $\hat{c}_{i,BLU}$ , and its covariance matrix is  $\sigma(\hat{c}_{BLU})$ .

**3. Representation of the BLUE.** We may reformulate the model in terms of the scalar process  $X(t, j) = X_j(t)$  depending on the parameter  $(t, j)$ . Under  $\mu_c$  this scalar process has mean value  $\sum_{k=1}^m c_k \beta_k(t, j)$ , where  $\beta_k(t, j) = \beta_{jk}(t)$ , and covariance function  $R(s, j; t, k) = R_{jk}(s, t)$ . The problem of estimating  $c_1, \dots, c_m$  from linear combinations of the  $X(t, j)$  for  $t \in S$  and  $1 \leq j \leq r$  is then equivalent to the one encountered in Sec. 8 of [1]. A restatement of Theorem 5 in [1], with appropriate notational changes, yields

**THEOREM 1.** Let the  $m \times 1$  random vector  $\psi$  have components  $\psi_i$  in  $H_X(S)$  and satisfy

$$EX(t)\psi' = \beta(t), \quad t \in S. \tag{5}$$

Then  $\psi$  has nonsingular covariance matrix  $\sigma(\psi) = E\psi\psi'$  and

$$\hat{c}_{BLU} = \sigma(\psi)^{-1}\psi, \quad \sigma(\hat{c}_{BLU}) = \sigma(\psi)^{-1}.$$

Moreover,  $\sigma(\hat{c}) - \sigma(\hat{c}_{BLU})$  is positive semidefinite for any linear unbiased estimator  $\hat{c}$  for  $c$ . If (5) has no solution as specified, then  $\sigma(\hat{c}_{BLU})$  is singular.

We note that the results in [1] were proved only for real-valued quantities. However, extensions to the complex case are obvious. The statement in the theorem that  $\sigma(\hat{c}) - \sigma(\hat{c}_{BLU})$  is positive semidefinite is true whether (5) has a solution or not, as is well known. In addition,  $\sigma(\hat{c}) = \sigma(\hat{c}_{BLU})$  only if  $\hat{c} = \hat{c}_{BLU}$ .

Also immediately applicable to our model is the result in Sec. 8 of [1] giving a lower bound for  $\sigma(\hat{c}_{BLU})$ . Employing the standard notation for Hermitian matrices of writing  $A \geq B$  to mean that  $A - B$  is positive semidefinite, we have

**THEOREM 2.** If (5) has a solution  $\psi$  with components in  $H_X(T)$ , then  $\sigma(\psi)$  is nonsingular and  $\sigma(\hat{c}_{BLU}) \geq \sigma(\psi)^{-1}$ . If (5) has no such solution,  $\sigma(\hat{c}_{BLU})$  is singular.

**4. Application to stationary sequences.** Let  $X(t)$ ,  $t = 0, \pm 1, \dots$  be an  $r \times 1$  vector-valued random sequence that is wide-sense stationary with mean value  $c$  and covariance function  $R(s, t) = R(s - t)$  under the measure  $\mu_c$ . This corresponds to  $T = \{0, \pm 1, \dots\}$ ,  $m = r$  and  $\beta(t) = I$  (the  $r \times r$  identity matrix) in the general case.  $c$  is to be estimated from the observations  $X(t)$  for  $t = 1, 2, \dots, N$ . We shall compare performance as  $N \rightarrow \infty$  of the LSE, or sample mean,  $\hat{c}_{LS} = N^{-1} \sum_{t=1}^N X(t)$  with that of the BLUE. The two theorems we shall prove are the direct vector analogues of Theorem 3 in [1]. The proofs are however somewhat complicated by the fact that we are dealing with covariance matrices, not scalar variances.

We make use of well-known spectral representations (see, e.g., Rozanov [5]). Thus

$$R(t) = \int_{-\pi}^{\pi} \exp(it\lambda) dF(\lambda),$$

where the  $r \times r$  matrix-valued spectral distribution function  $F(\lambda)$  has positive semidefinite increments. We employ the decomposition

$$dF(\lambda) = dF_s(\lambda) + (2\pi)^{-1}f(\lambda) d\lambda, \tag{6}$$

where the real and imaginary parts of all components of  $dF_s(\lambda)$  are singular with respect to  $d\lambda$  and the  $r \times r$  Hermitian matrix  $f(\lambda) \geq 0$  for all  $\lambda$ .  $F_s(\lambda)$  includes both the discrete and singular parts of the spectrum.

Making note of the spectral decomposition of each component  $X_i(t)$  of  $X(t)$ , we may also write

$$X(t) = \int_{-\pi}^{\pi} \exp(it\lambda) dZ(\lambda).$$

Here each component  $Z_i(\lambda)$  of the  $r \times 1$  random vector  $Z(\lambda)$  is in the subspace of  $L^2(\mu_0)$  spanned by the  $X_i(t)$ ; hence is in  $H_X(T)$ . Furthermore (symbolically),

$$\begin{aligned} E dZ(\lambda_1) dZ(\lambda_2)' &= dF(\lambda_1) \quad \text{if } \lambda_1 = \lambda_2, \\ &= 0 \quad \text{if } \lambda_1 \neq \lambda_2. \end{aligned} \tag{7}$$

**THEOREM 3.** As  $N \rightarrow \infty$ ,  $\lim \sigma(\hat{c}_{BLU}) = \lim \sigma(\hat{c}_{LS}) = dF(0)$ .

*Proof.* By dominated convergence,

$$\sigma(\hat{c}_{LS}) = \int_{-\pi}^{\pi} N^{-2} \left| \sum_{t=1}^N \exp(it\lambda) \right|^2 dF(\lambda) \rightarrow dF(0)$$

as  $N \rightarrow \infty$ . Since  $\sigma(\hat{c}_{BLU}) \leq \sigma(\hat{c}_{LS})$ , the proof will be complete if we establish that  $\sigma(\hat{c}_{BLU}) \geq dF(0)$ . In fact, it will be enough to show that

$$\alpha \sigma(\hat{c}) \alpha' = E |\alpha \hat{c}|^2 \geq \alpha dF(0) \alpha' \quad (8)$$

for any  $1 \times r$  row vector  $\alpha$  and linear unbiased estimator  $\hat{c}$ .

Since (8) is certainly true if the quantity on the right vanishes, we assume that  $\alpha dF(0) \alpha' > 0$  and set  $\gamma = [\alpha dF(0) \alpha']^{-1} \alpha$ . The random variable  $\phi = \gamma dZ(0)$  is in  $H_X(T)$  and is seen to satisfy  $EX(t)\phi = dF(0)\gamma'$  for all  $t$ . It follows that  $E\hat{c}\phi = dF(0)\gamma'$  and hence that  $E\alpha\hat{c}\phi = \alpha dF(0)\gamma' = 1$ . But the Cauchy-Schwartz inequality then yields

$$E |\alpha \hat{c}|^2 \geq 1/E |\phi|^2 = 1/\gamma dF(0)\gamma' = \alpha dF(0) \alpha',$$

as desired.

The next theorem shows that, under fairly general conditions, both  $\sigma(\hat{c}_{BLU})$  and  $\sigma(\hat{c}_{LS})$  approach their common limit at the same rate. In the proof we require a spectral representation of the elements in  $H_X(T)$ , already implied in the proof of Theorem 3. Let  $L^2(dF)$  be the class of all  $1 \times r$  row vector-valued functions  $h(\lambda)$  for which  $\int_{-\pi}^{\pi} h(\lambda) dF(\lambda) h(\lambda)' < \infty$ . As shown by Rozanov [5],  $L^2(dF)$  is a Hilbert space with inner product  $\int_{-\pi}^{\pi} h_1(\lambda) dF(\lambda) h_2(\lambda)'$  that is isometric to  $H_X(T)$ . The isometry is given by  $\phi = \int_{-\pi}^{\pi} h(\lambda) dZ(\lambda)$ , where  $\phi \in H_X(T)$  and  $h \in L^2(dF)$ . In particular,  $X_i(t)$  is identified with  $\exp(it\lambda)e_i$ , where  $e_i$  is the  $1 \times r$  row vector with entries  $\delta_{ik}$ .

**THEOREM 4.** Assume that

- (i)  $f(\lambda)$  is continuous and nonsingular at  $\lambda = 0$ ,
- (ii)  $f(\lambda)$  has an inverse  $f(\lambda)^{-1}$  almost everywhere on  $(-\pi, \pi)$  and  $\int_{-\pi}^{\pi} |p(\lambda)|^2 f(\lambda)^{-1} d\lambda$  exists for some trigonometric polynomial  $p(\lambda)$ , and
- (iii) for some  $\delta > 0$ ,  $F_s(\lambda) - F_s(-\lambda) = dF(0)$  when  $0 < \lambda < \delta$ . Then, as  $N \rightarrow \infty$ ,

$$\lim N[\sigma(\hat{c}_{BLU}) - dF(0)] = \lim N[\sigma(\hat{c}_{LS}) - dF(0)] = f(0).$$

*Proof.* By assumptions (i) and (iii),

$$\begin{aligned} N[\sigma(\hat{c}_{LS}) - dF(0)] &= \int_{\delta \leq |\lambda| \leq \pi} N^{-1} \left| \sum_{t=1}^N \exp(it\lambda) \right|^2 dF_s(\lambda) \\ &\quad + \int_{-\pi}^{\pi} (2\pi N)^{-1} \left| \sum \exp(it\lambda) \right|^2 f(\lambda) d\lambda \rightarrow f(0) \end{aligned}$$

as  $N \rightarrow \infty$ . Here we have used dominated convergence and properties of the Féjer kernel. Since  $\sigma(\hat{c}_{BLU}) \leq \sigma(\hat{c}_{LS})$ , the proof will be complete if we show that  $N[\sigma(\hat{c}_{BLU}) - dF(0)]$  is bounded below by a matrix that approaches  $f(0)$ .

Now if  $\hat{c}_{BLU} = \sum_{i=1}^N B_i X(t)$ , where the  $r \times r$  matrices  $B_i$  satisfy

$$B_1 + \cdots + B_N = I,$$

then

$$\begin{aligned} \sigma(\hat{c}_{BLU}) - dF(0) &= \int_{-\pi}^{\pi} H(\lambda) dF(\lambda)H(\lambda)' - dF(0) \\ &\geq (2\pi)^{-1} \int_{-\pi}^{\pi} H(\lambda)f(\lambda)H(\lambda)' d\lambda, \end{aligned} \tag{9}$$

where  $H(\lambda) = \sum \exp(it\lambda)B_i$ . The right side of (9) is in turn bounded below by  $\sigma(\hat{c}_{BLU})$  calculated under the hypothesis that  $F_s(\lambda) \equiv 0$ . To complete the proof it will thus suffice to assume that  $F_s$  vanishes and to show that  $N\sigma(\hat{c}_{BLU})$  is bounded below by a matrix that approaches  $f(0)$ .

Without loss of generality, we may write the polynomial  $p(\lambda)$  in assumption (ii) as  $p(\lambda) = \sum_{-\alpha}^{\alpha} b_\nu \exp(i\nu\lambda)$ . Because of condition (i), we may also assume  $p(0) \neq 0$ . Then the  $r \times r$  matrix

$$G(\lambda) = \frac{p(\lambda)}{p(0)} \sum_{i=1}^{N+\alpha} \exp(it\lambda)f(\lambda)^{-1}$$

has rows in  $L^2(f d\lambda)$ , so  $\psi = \int_{-\pi}^{\pi} G(\lambda) dZ(\lambda)$  is an  $r \times 1$  vector with entries in  $H_X(T)$ . By a straightforward calculation,

$$EX(t)\psi' = (2\pi)^{-1} \int_{-\pi}^{\pi} \exp(it\lambda)f(\lambda)G(\lambda)' d\lambda = I, \quad 1 \leq t \leq N,$$

so that  $N\sigma(\hat{c}_{BLU}) \geq N\sigma(\psi)^{-1}$  by Theorem 2. But

$$\begin{aligned} (N + \alpha)^{-1}\sigma(\psi) &= \frac{1}{2\pi(N + \alpha)} \int_{-\pi}^{\pi} G(\lambda)f(\lambda)G(\lambda)' d\lambda \\ &= \frac{1}{2\pi(N + \alpha)} \int_{-\pi}^{\pi} \left| \sum_1^{N+\alpha} \exp(it\lambda) \right|^2 \left| \frac{p(\lambda)}{p(0)} \right|^2 f(\lambda)^{-1} d\lambda \rightarrow f(0)^{-1} \end{aligned}$$

as  $N \rightarrow \infty$ ; hence  $N\sigma(\psi)^{-1} \rightarrow f(0)$  and the proof is complete.

The theorem states in a sense that  $\hat{c}_{LS}$  is asymptotically efficient relative to  $\hat{c}_{BLU}$ . For  $F_s(\lambda) \equiv 0$ , the result is essentially a slightly strengthened version of a special case in Rosenblatt [6]. The point of condition (ii) is to allow  $f(\lambda)$  to be singular at points with, however, these singularities "removable" by  $|p(\lambda)|^2$ .

**5. Application to stationary processes.** We pass now to the continuous-time analogue of the preceding section. Thus  $T = (-\infty, \infty)$  and  $X(t)$ ,  $-\infty < t < \infty$ , is an  $r \times 1$  vector-valued process that is wide-sense stationary under each  $\mu_c$ , with mean value  $c$  and covariance function  $R(s - t)$ .  $c$  is to be estimated from observations  $X(t)$  for  $0 \leq t \leq A$ . The LSE or sample mean is now  $\hat{c}_{LS} = A^{-1} \int_0^A X(t) dt$ , and its performance is to be compared to that of the BLUE as  $A \rightarrow \infty$ .

We assume that (each component of)  $X(t)$  is continuous in the mean-square sense, so that  $R(t)$  admits a spectral representation

$$R(t) = \int_{-\infty}^{\infty} \exp(it\lambda) dF(\lambda).$$

The spectral distribution function  $F(\lambda)$  is decomposed as in (6), with a similar interpretation as therein. Similarly, we write

$$X(t) = \int_{-\infty}^{\infty} \exp(it\lambda) dZ(\lambda),$$

where  $Z(\lambda)$  has properties as in Sec. 4 and satisfies (7). The Hilbert space  $L^2(dF)$  isometric to  $H_X(T)$  is defined as in the preceding section, except that all integrals are now over  $(-\infty, \infty)$ .

**THEOREM 5.** As  $A \rightarrow \infty$ ,

$$\lim \sigma(\hat{c}_{BLU}) = \lim \sigma(\hat{c}_{LS}) = dF(0).$$

*Proof.* By dominated convergence,

$$\sigma(\hat{c}_{LS}) = \int_{-\infty}^{\infty} A^{-2} \left| \int_0^A \exp(it\lambda) dt \right|^2 dF(\lambda) \rightarrow dF(0)$$

as  $A \rightarrow \infty$ . The remainder of the proof is identical to that of Theorem 3.

**THEOREM 6.** Assume that

- (i)  $f(\lambda)$  is continuous and nonsingular at  $\lambda = 0$ ,
- (ii)  $f(\lambda)$  has an inverse  $f(\lambda)^{-1}$  almost everywhere on  $(-\infty, \infty)$ , and  $\int_{-\infty}^{\infty} |p(\lambda)|^2 (1 + \lambda^2)^{-n} f(\lambda)^{-1} d\lambda$  exists for some positive integer  $n$  and function of the form  $p(\lambda) = \sum_{r=1}^{\infty} b_r \exp(-it_r \lambda)$ , and
- (iii) for some  $\delta > 0$ ,  $F_S(\lambda) - F_S(-\lambda) = dF(0)$  when  $0 < \lambda < \delta$ . Then, as  $A \rightarrow \infty$ ,

$$\lim A[\sigma(\hat{c}_{BLU}) - dF(0)] = \lim A[\sigma(\hat{c}_{LS}) - dF(0)] = f(0).$$

*Proof.* By dominated convergence and properties of the continuous Féjér kernel,

$$\begin{aligned} A[\sigma(\hat{c}_{LS}) - dF(0)] &= \int_{|\lambda| \geq \delta} A^{-1} \left| \int_0^A \exp(it\lambda) dt \right|^2 dF_S(\lambda) \\ &\quad + \int_{-\infty}^{\infty} (2\pi A)^{-1} \left| \int_0^A \exp(it\lambda) dt \right|^2 f(\lambda) d\lambda \rightarrow f(0) \end{aligned}$$

as  $A \rightarrow \infty$ . We have used conditions (i) and (iii). It remains to be shown that  $A[\sigma(\hat{c}_{BLU}) - dF(0)]$  is bounded below by a matrix that approaches  $f(0)$ .

We may choose finite linear unbiased estimators  $\hat{c}_N$  for  $c$  such that  $\hat{c}_N \rightarrow \hat{c}_{BLU}$  (componentwise) as  $N \rightarrow \infty$ , hence also  $\sigma(\hat{c}_N) \rightarrow \sigma(\hat{c}_{BLU})$ . By an argument as in Theorem 4, each  $\sigma(\hat{c}_N) - dF(0)$  is bounded below by the covariance matrix of the BLUE calculated under the hypothesis that  $F_S(\lambda) \equiv 0$ . Therefore  $\sigma(\hat{c}_{BLU}) - dF(0)$  is similarly so bounded below. It suffices thus to assume that  $F_S$  vanishes and to show that  $A\sigma(\hat{c}_{BLU})$  is then bounded below by a matrix that approaches  $f(0)$  as  $A \rightarrow \infty$ .

Without loss of generality we may take the constants  $t_r$  of the function  $p(\lambda)$  in condition (ii) to be non-negative, and assume that  $p(0) \neq 0$ . We let  $\gamma = \max\{t_1, \dots, t_n\}$ . Defining

$$g_A(\lambda) = \int_0^A \exp(it\lambda) dt + \sum_{k=0}^{n-1} \binom{n}{k+1} (i\lambda)^k [(-1)^k + \exp(iA\lambda)],$$

where  $n$  is the integer in condition (ii), we see that the  $r \times r$  matrix

$$G(\lambda) = \frac{p(\lambda)}{p(0)} g_{A+\gamma}(\lambda) (1 + \lambda^2)^{-n} f(\lambda)^{-1}$$

has rows in  $L^2(f d\lambda)$ . Therefore the  $r \times 1$  vector  $\psi = \int_{-\infty}^{\infty} G(\lambda) dZ(\lambda)$  has components in  $H_X(T)$ . As noted in [1] (Theorem 4),

$$(2\pi)^{-1} \int_{-\infty}^{\infty} \exp(it\lambda) \overline{g_A(\lambda)} (1 + \lambda^2)^{-n} d\lambda = 1, \quad 0 \leq t \leq A.$$

A straightforward calculation then establishes

$$EX(t)\psi' = (2\pi)^{-1} \int_{-\infty}^{\infty} \exp(it\lambda)f(\lambda)G(\lambda)' d\lambda = I, \quad 0 \leq t \leq A,$$

so that  $\sigma(\hat{c}_{BLV}) \geq \sigma(\psi)^{-1}$  by Theorem 2. As shown in [1], also

$$\lim_{A \rightarrow \infty} (2\pi A)^{-1} \int_{-\infty}^{\infty} |g_A(\lambda)|^2 u(\lambda)(1 + \lambda^2)^{-n} d\lambda = u(0)$$

for any function  $u(\lambda)$  integrable over  $(-\infty, \infty)$  and continuous at  $\lambda = 0$ . This implies that, as  $A \rightarrow \infty$ ,

$$\begin{aligned} (A + \gamma)^{-1}\sigma(\psi) &= \frac{1}{2\pi(A + \gamma)} \int_{-\infty}^{\infty} G(\lambda)f(\lambda)G(\lambda)' d\lambda \\ &= \frac{1}{2\pi(A + \gamma)} \int_{-\infty}^{\infty} |g_{A+\gamma}(\lambda)|^2 \left| \frac{p(\lambda)}{p(0)} \right|^2 (1 + \lambda^2)^{-2n} f(\lambda)^{-1} d\lambda \rightarrow f(0)^{-1}. \end{aligned}$$

Therefore  $A\sigma(\hat{c}_{BLV}) \geq A\sigma(\psi)^{-1} \rightarrow f(0)$  and the proof is complete.

**6. Constant regression for stationary processes.** The conclusions of the preceding two sections need not be valid unless the means  $c_1, \dots, c_r$  of the components of  $X(t)$  are independently estimated. If there are known linear constraints among the  $c_i$ , then the LSE need not be asymptotically as good as the BLUE. These remarks are illustrated in a simple example by Freiberger, Rosenblatt and Van Ness [2].

The case of  $r - m$  independent constraints may be transformed into the model  $\int X(t) d\mu_c = \beta c$ , where  $X(t)$  is stationary and  $\beta$  is a constant  $r \times m$  matrix ( $m \leq r$ ) of rank  $m$ . The LSE is now  $\hat{c}_{LS} = (\beta'\beta)^{-1}\beta'\bar{X}$ , where  $\bar{X}$  is the sample mean. While  $\hat{c}_{LS}$  is not necessarily asymptotically efficient, one might conjecture that there is an asymptotically efficient estimator based on the sample mean. Such an estimator would have the form  $\hat{c} = B\bar{X}$ , where  $B$  is some  $m \times r$  left inverse of  $\beta$ . A study of this possibility could be based on the theorems in Sec. 3 and methods of Secs. 4 and 5.

We restrict our attention in the present paper to the stationary case when  $m = r$  and  $\int X(t) d\mu_c = \beta c$ , where  $\beta$  is  $r \times r$  and non-singular. According to the above remarks, this corresponds in a sense to no linear constraints among the means, and should lead to results of the type given in Secs. 4 and 5.

Such results are now easily obtainable by a simple transformation and without detailed analysis. We set  $Y(t) = \beta^{-1}X(t)$  and note that the  $Y$ -process has spectral measure  $\beta^{-1} dF(\lambda)\beta$  if the  $X$ -process has spectral measure  $dF(\lambda)$ . As  $\int Y(t) d\mu_c = c$  and  $\hat{c}_{LS} = \beta^{-1}\bar{X} = \bar{Y}$ , we immediately find that Theorems 3 through 6 remain valid for this case as long as  $dF(0)$  and  $f(0)$  in the conclusions therein are replaced, respectively, by  $\beta^{-1} dF(0)\beta$  and  $\beta^{-1}f(0)\beta$ .

REFERENCES

- [1] R. K. Adenstedt and B. Eisenberg, *Linear estimation of regression coefficients*, Quart. Appl. Math. **32**, 317-327 (1973)
- [2] W. Freiberger, M. Rosenblatt and J. Van Ness, *Regression analysis of vector-valued random processes*, J. SIAM **10**, 89-102 (1962)

- [3] U. Grenander, *On the estimation of regression coefficients in the case of an autocorrelated disturbance*, Ann. Math. Stat. **25**, 252–272 (1954)
- [4] J. V. Kuk and J. I. Petunin, *Observable linear estimates of the mathematical expectation of a random process*, Soviet Math. Dokl. **14**, 323–326 (1973)
- [5] Y. A. Rozanov, *Stationary random processes*, Holden-Day, San Francisco (1967)
- [6] M. Rosenblatt, *On the estimation of regression coefficients of a vector-valued time series with a stationary residual*, Ann. Math. Stat. **26**, 99–121 (1955)