

## LINEAR REGRESSION IN CONTINUOUS TIME

E. J. HANNAN

(Received 11 February 1971)

Communicated by P. D. Finch

### 1. Introduction

We consider a regression relation of the form

$$(1) \quad y(t) = B'x(t) + e(t), \quad t \in (-\infty, T)$$

wherein  $y(t)$  and  $x(t)$  are real (column) vectors of  $q$  and  $p$  components and  $e(t)$  is real and is generated by a stationary generalised vector process of  $q$  components with zero mean and covariance function (a  $q$  rowed matrix)  $\Gamma(t-s) = E\{x(s)x(t)'\}$ . (See Hannan (1970; pages 23–26, 91–94) and references therein for definitions of terms used.) We assume  $e(t)$  to be independent of  $x(s)$  for all  $s, t$ . Thus we may regard  $x(t)$  as a fixed time function and not stochastic and we shall henceforth do that. We take  $\Gamma(t)$  to be continuous and to correspond to an absolutely continuous spectral function with spectral density which is uniformly bounded and continuous. Then we have

$$\Gamma(t) = \int_{-\infty}^{\infty} e^{it\lambda} f(\lambda) d\lambda, \quad -\infty < t < \infty.$$

We do not exclude the possibility that for the  $j$ th diagonal element,  $f_{jj}$ , of  $f$  we have

$$\int_{-\infty}^{\infty} f_{jj}(\lambda) d\lambda = \infty.$$

We may write

$$e(t) = \int_{-\infty}^{\infty} e^{-it\lambda} z(d\lambda)$$

where  $z(\lambda)$  is a vector process of orthogonal increments with  $E\{z(d\lambda)z(d\lambda)^*\} = f(\lambda)d\lambda$ . Putting  $\|A\| = [\text{tr}(A^*A)]^{\frac{1}{2}}$  for any matrix  $A$  we assume that

$$\int_{-\infty}^T \|x(t)\|^2 dt < \infty.$$

We take  $t \in (-\infty, T)$  for convenience of exposition but later shall show that we have also covered the case  $t \in [0, T)$ , for example.

We shall be substantially concerned with improving some results in Kholevo (1969) and shall repeatedly refer to that work. In Kholevo (1969; Theorem 2) an asymptotic ( $T \rightarrow \infty$ ) lower bound to the (normalised) covariance matrix of the best linear unbiased estimate (BLUE) of  $B$  is obtained for the case  $q = 1$ . We here wish to investigate conditions under which the lower bound becomes the asymptotic value of the (normalised) covariance matrix. Kholevo uses his lower bound to find sufficient conditions for the least squares estimate (LSE) to be asymptotically efficient. Our improvement of his result will allow us to show that these conditions are also necessary. We generalise to the vector case (for  $y(t)$ ) partly because in this case we can obtain an interesting characterisation of the nature of any  $x(t)$  which makes LSE efficient for a sufficiently wide class. of  $f(\lambda)$ . Finally we shall relate these results to the problem of smoothing a continuous time record discussed in Blackman (1965).

The results in Kholevo (1969) and in the present paper are extensions to continuous time of results obtained for discrete time by Grenander (1964), Hannan (1968). For further references see Kholevo (1969) and Hannan (1970). The first attempt at a general solution of the continuous time problem is found in Heble (1961). However, Theorem 2 is incorrect, as is pointed out in Kholevo (1969). Heble's method of proof is based upon a discrete time approximation to the continuous time phenomenon and differs from that in Kholevo (1969) and the present paper. The key to this extension is the observation, first made in this connection in Rozanov (1964), that  $x(t)$  must be of the special form if it is not to be true that  $B$ , no matter what it may be, can almost surely be distinguished from the null matrix from observations on  $(-\infty, T]$ . This special form is, in case  $q = 1$ ,

$$(2) \quad x(t) = \int_{-\infty}^{\infty} e^{it\lambda} \phi^{(T)}(\lambda) f(\lambda) d\lambda, \quad t < T.$$

wherein  $\phi^{(T)}(\lambda)$  is of the form

$$\int_{-\infty}^{\infty} e^{-it\lambda} a^{(T)}(t) dt, \quad \int_{-\infty}^T \|a^{(T)}(t)\|^2 dt < \infty$$

or is the limit in mean square (with weight function  $f(\lambda)$ ) of a sequence of such (vectors of) functions. For completeness a proof of these assertions relating to (2), and its generalisation (2)', below, is given in an appendix to this paper. For  $q > 1$  it is convenient to introduce the notation  $\beta = \text{vec } B$  for the column vector which has  $\beta_{ij}$  in row  $(j-1)q + 1$  (i.e. which has the columns of  $B$  placed successively down it). By  $A \otimes C$  we mean the Kronecker product of these matrices, having  $a_{ij}c_{kl}$  in row  $(i-1)p + k$ , column  $(j-1)q + l$ , where  $C$  is a  $p \times x$  matrix. Then (1) may be rewritten as

$$(1)' \quad y(t) = \{I_q \otimes x(t)'\} \beta + e(t).$$

The condition (2) now becomes

$$(2)' \quad \{I_q \otimes x(t)'\} = \int_{-\infty}^{\infty} e^{it\lambda} f(\lambda) \Phi^{(T)}(\lambda) d\lambda, \quad t < T$$

where  $\Phi^{(T)}$  is a  $q \times pq$  matrix of functions which is of the form

$$\Phi^{(T)}(\lambda) = \int_{-\infty}^T e^{-it\lambda} A(t) dt, \quad \int_{-\infty}^T \|A(t)\|^2 dt < \infty$$

or is in the closure of this linear set with respect to the norm

$$\|\Phi^{(T)}\| = \left\{ \text{tr} \left[ \int_{-\infty}^{\infty} \Phi^{(T)}(\lambda) * f(\lambda) \Phi^{(T)}(\lambda) d\lambda \right] \right\}^{\frac{1}{2}}.$$

As mentioned above we give a proof that (2)' must hold if the inference problem is not to become, in a sense, trivial in the appendix to this paper. We henceforth assume that (2)' holds.

In the discrete time case the analogous theorems are proved under the, not unreasonable, condition that  $f(\lambda) > 0, \lambda \in (-\pi, \pi]$ . (Here and below we mean the notation  $f(\lambda) > 0$  to mean that  $f(\lambda)$  is positive definite.) In the continuous time situation the following condition must be imposed, namely,

$$(3) \quad \inf_{\lambda} \text{tr} \{f(\lambda)(1 + \lambda^2)^n\} > 0, \text{ some } n \geq 0.$$

There is no intrinsic reason why the function  $(1 + \lambda^2)^n$  should be used and it could be replaced, under suitable conditions, by other functions corresponding to a filter having response (namely  $(1 + i\lambda)^n$  in the present case) with all of its zeros in one half plane (so that it is one-sided and invertible). However the fact that  $i\lambda$  is the response function of the operator  $d/dt$  makes (3) a natural choice.

The theorem we prove below expresses the covariance matrix of the BLUE,  $\hat{B}$ , of  $B$  and the LSE,  $\hat{B}$ , of  $B$  in terms of a generalised spectrum for  $x(t)$ . As in Kholevo (1969) we henceforth assume that  $x(t)$  has  $n$  derivatives (where  $n$  is determined by (3)) that are square integrable on  $(-\infty, T]$ . We introduce a polynomial  $P(z)$  of degree  $n$ , with all of its zeros in the left half plane (which, as in Kholevo (1969), we call a stable polynomial), and put

$$\hat{x}(t) = P(d/dt)x(t).$$

We put

$$\hat{q}_k(T) = \left\{ \int_{-\infty}^T \hat{x}_k(t)^2 dt \right\}^{\frac{1}{2}}$$

and shall call  $\hat{Q}(T)$  the diagonal matrix with  $\hat{q}_k(t)$  in the  $k$ th place in the main diagonal. We assume that the following limits exist and have the values shown

$$(4) \quad \lim_{T \rightarrow \infty} \hat{x}_k(T)/\hat{q}_k(T) = 0$$

$$(5) \quad \lim_{T \rightarrow \infty} \int_{-\infty}^T \hat{x}_k(s) \hat{x}_k'(s+t) ds / \{\hat{q}_k(T) \hat{q}_k'(T)\} = \hat{p}_{ki}(t).$$

We call  $\hat{R}(t)$  the matrix with entries  $\hat{p}_{ki}(t)$  and assume that  $\hat{R}(t)$  is continuous at  $t = 0$  and that  $\hat{R}(0)$  is nonsingular. Then since  $\alpha' \hat{R}(t) \alpha$  is a non-negative definite function for all  $\alpha$  it follows from Bochner's Theorem that

$$\hat{R}(t) = \int_{-\infty}^{\infty} e^{it\lambda} \hat{M}(d\lambda)$$

where  $\hat{M}(\lambda)$  is an Hermitian matrix which is nondecreasing as a function of  $\lambda$ . When all of this is true we say that  $\hat{x}(t)$  has a spectral measure.

In Kholevo (1969; page 91) the condition (4) is omitted. However some such condition is necessary as the example  $x(t) = \exp t$  shows, for it satisfies (5), with  $R(t) = \exp t$ , which is continuous and nonzero, but evidently is *not* non-negative definite. In Kholevo (1969)  $R(t)$  is defined differently but this has the effect, in the present example, only of changing  $R(t)$  to  $\exp -t$ . If we use (2)' we see that

$$(6) \quad \{I_q \otimes \hat{x}(t)'\} = \int_{-\infty}^{\infty} e^{it\lambda} \hat{f}(\lambda) \hat{\Phi}^{(T)}(\lambda) d\lambda$$

wherein  $\hat{f}(\lambda) = |P(i\lambda)|^2 f(\lambda)$ ,  $\hat{\Phi}^{(T)}(\lambda) = \overline{P(i\lambda)}^{-1} \Phi^{(T)}(\lambda)$ . The fact that  $\hat{\Phi}^{(T)}(\lambda)$  is of the required form follows from the location of the zeros of  $P(z)$ .

Finally we assume

$$(7) \quad \lim_{T \rightarrow \infty} \int_{-\infty}^T x_k(t)^2 dt = \infty$$

for without this condition the LSE cannot be consistent. This is easily seen to imply

$$(7)' \quad \lim_{T \rightarrow \infty} \int_{-\infty}^T \hat{x}_k(t)^2 dt = \infty.$$

We henceforth assume all of the conditions of this section to hold and refer to them collectively as Conditions A.

## 2. The Asymptotic Variance of the BLUE

The LSE is, by definition

$$\hat{\beta} = \left\{ \int_{-\infty}^T \{I_q \otimes x(t)x(t)'\} dt \right\}^{-1} \cdot \int_{-\infty}^T \{I_q \otimes x(t)\} y(t) dt$$

i.e.  $\text{vec } \hat{B} = \hat{\beta}$  where

$$\hat{B} = \left\{ \int_{-\infty}^T x(t)x(t)' dt \right\}^{-1} \cdot \int_{-\infty}^T x(t)y(t)' dt.$$

Let  $\Omega(T)$  be the covariance matrix of  $\hat{\beta}$ . Then if  $x(t)$  has a spectral measure it is not difficult to show that

$$\begin{aligned} \lim_{T \rightarrow \infty} \{I_q \otimes Q(T)\} \Omega(T) \{I_q \otimes Q(T)\} \\ = \{I_q \otimes R(0)^{-1}\} \int_{-\infty}^{\infty} \{2\pi f(\lambda)\} \otimes M'(d\lambda) \{I_q \otimes R(0)^{-1}\}. \end{aligned}$$

We omit the proof, which is of the same nature as that given in Hannan (1970; pages 216–219).

The BLUE is, assuming  $f(\lambda)$  to be known,

$$(8) \quad \hat{\beta} = \left\{ \int_{-\infty}^{\infty} \Phi^{(T)}(\lambda)^* f(\lambda) \Phi^{(T)}(\lambda) d\lambda \right\}^{-1} \int_{-\infty}^{\infty} \Phi^{(T)}(\lambda)^* \{(dy) + f(\lambda) \Phi^{(T)}(\lambda) \beta\} d\lambda,$$

with covariance matrix  $\Omega_0(T)$  which is the first factor. (See Kholevo (1969; page 82). Of course we also have

$$\Omega_0(T) = \left\{ \int_{-\infty}^{\infty} \hat{\Phi}^{(T)}(\lambda)^* f(\lambda) \hat{\Phi}^{(T)}(\lambda) d\lambda \right\}^{-1}.$$

We now establish the following theorem, assuming also

$$(9) \quad \sup_{\lambda} (1 + \lambda^2)^n \text{tr}\{f(\lambda)\} < \infty.$$

**THEOREM 1.** *Let Conditions A and (9) be satisfied. Then*

$$\lim_{T \rightarrow \infty} \{I_q \otimes \hat{Q}(T)\} \Omega_0(T) \{I_q \otimes \hat{Q}(T)\} = \left[ \int_{-\infty}^{\infty} \{2\pi \hat{f}(\lambda)\}^{-1} \otimes M'(d\lambda) \right]^{-1}.$$

We may of course compute  $\hat{\beta}$  in terms of  $\hat{x}(t)$  (i.e. putting hats on  $\phi^{(T)}$ ,  $f$  and  $z$ ).

We observe first that we may find matrices of rational functions of  $z$ ,  $R_j(z)$ ,  $j = 1, 2$ , whose determinants have all zeros and poles located inside the left half plane, for which

$$\sup_{\lambda} R_j(i\lambda) < \infty, \quad \inf_{\lambda} R_j(i\lambda) > 0,$$

and which are such that (in the usual partial ordering of Hermitian matrices)

$$\{R_1(i\lambda)R_1(i\lambda)^*\}^{-1} \leq \hat{f}(\lambda)^{-1} \leq \{R_2(i\lambda)R_2(i\lambda)^*\}^{-1}$$

and

$$(10) \quad \{R_2(i\lambda)R_2(i\lambda)^*\}^{-1} - \{R_1(i\lambda)R_1(i\lambda)^*\}^{-1} \leq \varepsilon I, \quad \varepsilon > 0.$$

This may be established by mapping the closed left half plane onto the closed unit disc by means of  $z = (1 + w)/(1 - w)$ , which maps  $z = i\lambda$  onto  $w = \tan \frac{1}{2}\theta$ , approximating to  $\hat{f}(\tan \frac{1}{2}\theta)^{-1}$  uniformly above and below by positive definite

matrices of trigonometric polynomials and reversing the transformation. If we can establish the theorem with  $R_f(i\lambda)R_f(i\lambda)^*$  replacing  $\hat{f}(\lambda)$  we shall evidently have proved what we wish in virtue of (10). Thus we now take  $\hat{f}(\lambda)$  to be a "rational spectrum" of this form. Of course a representation (6) still holds since this is now true for any  $\hat{x}(t)$  which is square integrable on  $(-\infty, T]$  because of the fact that  $\hat{f}$  is bounded and bounded away from zero. We shall call the expression on the right of (6)  $X^{(T)}(t)$  it being understood that this is  $I_q \otimes \hat{x}(t)'$  for  $t \leq T$ . We first establish the following result.

LEMMA.

$$\lim_{T \rightarrow \infty} \{I_q \otimes \hat{Q}(T)^{-1}\} \int_T^\infty X^{(T)}(t)^* X^{(T)}(t) dt \{I_q \otimes \hat{Q}(T)^{-1}\} = 0.$$

We prove the lemma in case  $p = q = 1$ , for simplicity, but the general case is only notationally more complicated. We first consider  $\hat{f}(\lambda)^{-1}\theta(\lambda)$  where

$$\theta(\lambda) = \frac{1}{2\pi} \int_{-\infty}^T e^{-it\lambda} \hat{x}(t) dt.$$

Of course

$$\begin{aligned} \int_{-\infty}^\infty e^{it\lambda} \hat{f}(\lambda) \{\hat{f}(\lambda)^{-1}\theta(\lambda)\} d\lambda &= \hat{x}(t), \quad t < T \\ &= 0, \quad t \geq T. \end{aligned}$$

However,  $\theta(\lambda)/\hat{f}(\lambda)$  is not of the required form, namely such that it is the Fourier transform which is zero for  $t \geq T$ . We may decompose  $\hat{f}(\lambda)^{-1}$  into partial fractions as

$$\hat{f}(\lambda)^{-1} = a_0 + \sum_1^r a_k (b_k - i\lambda)^{-1} + \sum_0^r \bar{a}_k (\bar{b}_k + i\lambda)^{-1}, \quad \Re(b_k) < 0.$$

(It is possible that terms involving  $(b_k + i\lambda)^{-2}$ , and so on, might occur but it will be seen that the proof is not materially altered. In any case since our  $\hat{f}_j(\lambda)$  are, to an extent, arbitrary we may make a small modification of  $\hat{f}(\lambda)$  which will avoid this situation.) Now  $\{a_0 + \sum \bar{a}_k (\bar{b}_k + i\lambda)^{-1}\}\theta(\lambda)$  is of the required form. Also

$$\begin{aligned} \sum_1^r a_k (b_k - i\lambda)^{-1} \theta(\lambda) &= \sum_1^r a_k \frac{1}{2\pi} \int_{-\infty}^T e^{-it\lambda} \left\{ \int_0^\infty e^{b_k s} \hat{x}(t-s) ds \right\} dt \\ &\quad + \sum_1^r a_k \frac{1}{2\pi} \int_T^\infty e^{-it\lambda} \left\{ \int_{t-T}^\infty e^{b_k s} \hat{x}(t-s) ds \right\} dt. \end{aligned}$$

The first term on the right is again of the required form but the second is not. Rearranging that second term we get

$$\sum_1^r \left\{ a_k \int_0^\infty e^{b_k s} \hat{x}(T-s) ds e^{-iT\lambda} \frac{1}{2\pi(i\lambda - b_k)} \right\}.$$

Let  $\zeta_k(\lambda)$  be of the form

$$\zeta_k(\lambda) = \frac{1}{2\pi} \int_{-\infty}^0 e^{-is\lambda} a_k(s) ds, \int_{-\infty}^0 |a_k(s)|^2 ds < \infty,$$

$$\int_{-\infty}^{\infty} e^{i\lambda\hat{f}(\lambda)} \left( \frac{1}{2\pi(i\lambda - b_k)} \right) - \zeta_k(\lambda) d\lambda = 0, \quad t < 0.$$

(It is not difficult to see that  $\zeta_k(\lambda)$  may be chosen to be a linear combination of the  $(i\lambda + \bar{b}_j)$ ). Now we have shown that we may choose

$$X^{(T)}(t) = \int_{-\infty}^{\infty} e^{it\lambda\hat{f}(\lambda)} \left\{ \hat{f}(\lambda)^{-1}\theta(\lambda) - \sum_1^r a_k \left( \int_0^{\infty} e^{b_k s} \hat{x}(t-s) ds \right) e^{-iT\lambda} \cdot \left( \frac{1}{2\pi(i\lambda - b_k)} - \zeta_k(\lambda) \right) \right\} d\lambda,$$

and have reduced the synthesis of  $X^{(T)}(t)$ , i.e. of  $\hat{\Phi}^{(T)}(\lambda)$ , to that of  $\zeta_k(\lambda)$ , in the case of rational  $\hat{f}(\lambda)$ . Moreover it is now evident that

$$\int_T^{\infty} |X^{(T)}(t)|^2 dt \leq \text{constant} \left\{ \sum_0^r |a_k \int_0^{\infty} e^{b_k s} \hat{x}(T-s) ds| \right\}^2,$$

where the constant is independent of  $\hat{x}(t)$  and of  $T$ , since the only component contributing to  $X^{(T)}(t)$  for  $t \geq T$  is

$$\sum_1^r \left\{ a_k \int_0^{\infty} e^{b_k s} \hat{x}(T-s) ds \int_{-\infty}^{\infty} e^{i(t-T)\lambda\hat{f}(\lambda)} \zeta_k(\lambda) d\lambda \right\}.$$

However it is easily seen that, because of (4),

$$\lim_{T \rightarrow \infty} \left| \int_0^{\infty} e^{b_k s} \hat{x}(T-s) ds \right|^2 / \int_{-\infty}^T |\hat{x}(t)|^2 dt = 0, \quad k = 1, \dots, r$$

and the lemma is proved.

The remainder of the proof of the theorem is now easily accomplished. Again we give the proof for  $q = 1$  for notational convenience. Put

$$\hat{Q}(T)^{-1} \int_{-\infty}^{\infty} \hat{\Phi}^{(T)}(\lambda) * \hat{\Phi}^{(T)}(\lambda) \hat{f}(\lambda) d\lambda \hat{\Phi}(T)^{-1}$$

$$= \int_{-\infty}^{\infty} \{ \hat{f}(\lambda) \hat{Q}(T)^{-1} \hat{\Phi}^{(T)}(\lambda) \} * \hat{f}(\lambda)^{-1} \{ \hat{f}(\lambda) \hat{Q}(T)^{-1} \hat{\Phi}^{(T)}(\lambda) \} d\lambda.$$

Now

$$\lim_{T \rightarrow \infty} \int_{-\infty}^{\infty} e^{it\lambda} \{ \hat{f}(\lambda) \hat{Q}(T)^{-1} \hat{\Phi}^{(T)}(\lambda) \} * \{ \hat{f}(\lambda) \hat{Q}(T)^{-1} \hat{\Phi}^{(T)}(\lambda) \} d\lambda$$

$$= \lim_{T \rightarrow \infty} \frac{1}{2\pi} \int_{-\infty}^{\infty} \hat{Q}(T)^{-1} X^{(T)}(s) X^{(T)}(s+t)' \hat{Q}(T)^{-1} ds$$

and because of the lemma this is

$$\begin{aligned} \lim_{T \rightarrow \infty} \frac{1}{2\pi} \int_{-\infty}^T \hat{Q}(T)^{-1} \hat{x}(s) \hat{x}(s+t)' \hat{Q}(T)^{-1} ds &= \frac{1}{2\pi} \hat{R}(t) \\ &= \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{it\lambda} \hat{M}(d\lambda). \end{aligned}$$

Since  $\hat{f}(\lambda)^{-1}$  is a bounded continuous function the result now follows for  $q = 1$ . (For  $q = 1$  we may replace  $\hat{M}'(d\lambda)$  by  $\hat{M}(d\lambda)$  since  $\hat{f}(\lambda) = \hat{f}(-\lambda)$  otherwise it is necessary to prime, or conjugate  $\hat{M}(\lambda)$  with the definitions which we have adopted.)

The following is an almost immediate corollary. It constitutes Kholevo (1969; Theorem 2). The proof given there is probably more direct but our Theorem enables other results to be proved later.

**COROLLARY 1.** *Under Conditions A we have*

$$\lim_{T \rightarrow \infty} (I_q \otimes \hat{Q}(T)) \Omega_0(T) (I_q \otimes \hat{Q}(T)) \geq \left\{ \int_{-\infty}^{\infty} 2\pi f(\lambda)^{-1} \otimes \hat{M}'(d\lambda) \right\}^{-1}.$$

**PROOF.** Though now  $\hat{f}(\lambda)$  is not bounded we may still bound it below by a rational spectrum,  $R(i\lambda)R(i\lambda)^*$  of the required kind and so that

$$0 \leq \{R(i\lambda)R(i\lambda)^*\}^{-1} - \hat{f}(\lambda)^{-1} \leq \varepsilon I_q, \quad \varepsilon > 0.$$

Allowing  $\varepsilon$  to tend to zero the corollary follows.

As Kholevo (1969; Lemma 3) shows, provided

$$(11) \quad \lim_{T \rightarrow \infty} x_k^{(l)}(T) / \hat{q}_k(T) = 0; \quad k = 1, \dots, p; \quad l = 0, 1, \dots, (n-1),$$

then the existence of a spectral measure,  $\hat{M}(\lambda)$ , for  $\hat{x}(t)$  implies the existence of a spectral measure,  $M(\lambda)$ , for  $x(t)$ . Moreover if we put

$$q_k(T) = \left\{ \int_{-\infty}^T x_k(t)^2 dt \right\}^{\frac{1}{2}},$$

and define  $Q(T)$  in terms of the  $q_k(T)$  then

$$A = \lim_{T \rightarrow \infty} Q^{-1}(T) \hat{Q}(T)$$

exists and for any Borel set  $E$

$$M(E) = A \int_E |P(i\lambda)|^{-2} \hat{M}(d\lambda) A.$$

Now we have the following corollary.

**COROLLARY 2.** *Under conditions A and (11)*

$$\lim_{T \rightarrow \infty} (I_q \otimes Q(T)) \Omega_0(T) (I_q \otimes Q(T)) = \left\{ \int_{-\infty}^{\infty} (2\pi f(\lambda))^{-1} \otimes M'(d\lambda) \right\}^{-1}$$

PROOF. Using (11) the result of Corollary 1 may be written as

$$\overline{\lim}_{T \rightarrow \infty} (I_q \otimes Q(T))\Omega_0(T)(I_q \otimes Q(T)) \geq \left\{ \int_{-\infty}^{\infty} (2\pi f(\lambda))^{-1} \otimes M'(d\lambda) \right\}^{-1}.$$

However consider also the spectrum  $f(\lambda) + \varepsilon I_q$ ,  $\varepsilon > 0$ . Since this satisfies the conditions of the theorem with  $n = 0$  in (3) and (9) we have

$$\overline{\lim}_{T \rightarrow \infty} (I_q \otimes Q(T))\Omega_0(T)(I_q \otimes P(T)) \leq \left\{ \int_{-\infty}^{\infty} [2\pi(f(\lambda) + \varepsilon I_q)]^{-1} \otimes M'(d\lambda) \right\}^{-1}$$

Allowing  $\varepsilon$  to tend to zero the corollary is obtained.

The example in Kholevo (1969; page 99) shows that some such condition as (11) is needed if Corollary 2 is to hold. However the result still appears to be a very general one.

### 3. The Asymptotic Efficiency of the LSE

Precisely as in Grenander (1954) and following Kholevo (1969) we may introduce a unique (maximal) decomposition of  $(-\infty, \infty)$  into sets  $S_j$ ,  $j = 1, \dots, r \leq p$  so that, assuming  $x(t)$  to have a spectral measure, then

$$N_j = R(0)^{-\frac{1}{2}} \int_{S_j} M(d\lambda)R(0)^{-\frac{1}{2}}$$

are a set of mutually annihilating idempotents summing to  $I_p$ . We call the  $S_j$  the elements of the spectrum (of  $z(t)$ ). For a physical interpretation of the  $N_j$  see Hannan (1970; page 432).

We say that the LSE is efficient if  $(I_q \otimes Q(T))(\Omega(T) - \Omega_0(T))(I_q \otimes Q(T))$  converges to zero. The proof of the following theorem is now of exactly the same form as in Hannan (1970; Theorem VII, 8, page 433].

**THEOREM 2.** *Let Conditions A and (11) hold. Then the necessary and sufficient condition that the LSE be efficient is the condition that  $f(\lambda)$  be constant on the elements of the spectrum.*

This theorem improves Theorem 4 in Kholevo (1969) by adding the necessity of the condition. Kholevo also compares the BLUE (for  $q = 1$ ) with the LSE got from  $P_1(d/dt)x(t)$ ,  $P_1(d/dt)y(t)$  when  $P_1(z)$  is a stable polynomial of degree  $m$ ,  $0 \leq m \leq n$  and  $\sup f(\lambda)(1 + \lambda^2)^m < \infty$ . This result of Kholevo's is evidently an immediate corollary of Theorem 2, and necessity again may be inserted. In the case  $q > 1$  the extension is more elaborate and we forego the details.

The following is evidently an immediate corollary of Theorem 2.

**COROLLARY 3.** *Let Conditions A and (11) hold. Then the necessary and sufficient condition that the LSE be efficient for all  $f(\lambda)$  in some class which se-*

parates the points of  $(-\infty, \infty)$  for  $q > 1$  (separates the points of  $[0, \infty)$  for  $q = 1$ ) is that the elements of the spectrum be points for  $q > 1$  (be points or pairs of points symmetrically placed with respect to the origin for  $q = 1$ ).

Again this completes a result of Kholevo (1969) (the corollary to Theorem 4). The spectrum  $(1 + \lambda^2)^{-1}$  separates the points of  $[0, \infty)$ , for example.

We may also characterise the nature of an  $x(t)$  for which Corollary 3 holds exactly as in Corollary (VII.1) in Hannan (1970; page 436), the proof being almost precisely the same. Thus put  $N_j = U_j + iV_j$  and consider the real vector space,  $\mathcal{X}$  let us say, in which  $U_j, V_j$  act. Then  $\mathcal{X}$  may be decomposed into mutually perpendicular subspaces,  $\mathcal{X}_j$ , one for each  $\theta_j \geq 0$  in the regression spectrum, so that the projection on  $W_j$  is  $U_j$  if  $\theta_j = 0$  and otherwise is  $2U_j$ . The matrix  $V_j$  is skew symmetric and if  $S_j$  is a single point (as must be so for  $q > 1$ ) then in a suitable basis for  $\mathcal{X}_j$  it is

$$V_j = \frac{1}{2} \begin{bmatrix} 0 & 1 & & & \\ -1 & 0 & & & \\ & & 0 & 1 & \\ & & -1 & 0 & \\ & & & & \ddots \end{bmatrix}$$

where omitted elements are zero. Thus  $\mathcal{X}_j$  is even dimensional.

A set of  $x_k(t)$  satisfying these conditions is

$$\left. \begin{matrix} t^u \cos \theta_j t \\ t^u \sin \theta_j t \end{matrix} \right\} u = 0, 1, \dots, p_j - 1; j = 1, \dots, s.$$

In particular, as was first pointed out by Rosenblatt (1959), if  $s = 1$  and  $p_1 = 1$  (for example) efficiency cannot obtain when  $q > 1$  for all  $f(\lambda)$  in a suitably wide class if  $\cos \theta t$  occurs in the regression without  $\sin \theta t$  (or the reverse). In other words the knowledge of the nature of the polarisation of a sinusoidal signal provides information which the LSE does not use.

#### 4. Observations on $[0, T]$

We have considered the interval of observation to be  $(-\infty, T)$  because it is slightly more convenient. A more relevant interval would probably be  $[0, T)$  or  $(-T, T)$ . It is fairly obvious that all of the theorems given above continue to hold. Indeed the only point at which the proof essentially varies is in the Lemma which now must read, for  $[0, T]$  for example,

$$\lim_{T \rightarrow \infty} (I_q \otimes \hat{Q}(T)^{-1}) \left[ \int_{-\infty}^0 + \int_T^{\infty} \{X^{(T)}(t) * X^{(T)}(t)\} dt \right] (I_q \otimes \hat{Q}(T)^{-1}) = 0$$

The proof is however much the same.

### 5. An Application to Filtering

In Blackman (1965) the following problem is considered. We are to construct a linear filter of length  $2T$  which will pass any polynomial,  $p(s)$ , of prescribed degree  $p-1$  and will produce an output with minimum variance if the input has spectrum  $f(\lambda)$  (We require  $f(\lambda)$  to be scalar and to satisfy Conditions A.) We take the filter to be ‘‘centred’’ so that it is to be of the form, for input  $z(s)$

$$\int_{-T}^T z(t+s) a(t) dt$$

or is to be a limit on mean square of such expressions. It is intuitively seen (see Hannan (1970; page 191) that the filter is obtained as follows. We take  $t \in [-T, T]$ ,  $y(t) = z(s+t)$  and  $x_k(t) = t^{k-1}$ ,  $k = 1, \dots, p$  in our model (for  $q = 1$ ). If  $\hat{\beta}_k(s)$  is the BLUE (it does of course depend upon  $s$ ) then the output of the optimal filter at time point  $s$  is

$$\sum \hat{\beta}_k(s) t^{k-1}.$$

Of course this begs the question of the actual construction of the  $\hat{\beta}_k(s)$ , which Blackman extensively discusses for rational  $f(\lambda)$ . Of course the problem was effectively solved in the rational case during the proof of the lemma, above. Blackman also considers

$$\sum \hat{\beta}_k(s) t^{k-1}$$

where  $\hat{\beta}_k(s)$  is the LSE from the same data, and observes that, for example for  $p = 2$  and  $f(\lambda) = (1 + \rho^2 \lambda^2)^{-1}$ ,  $\rho > 0$ , the LSE performs almost equally as well as the BLUE. Indeed the ratio of the variance of output of the latter to that of the former reaches a maximum, of about 1.07 and declines to unity as  $T \rightarrow \infty$ . (The ratio depends only upon  $\rho T$  and the maximum is at  $\rho T$  near to unity.) Needless to say these results are aspects of our Theorems for here  $x_k(t)$  evidently satisfies all conditions required of it in Conditions A and (11) and are of the form required in Corollary 3 so that as  $T \rightarrow \infty$  the two procedures must become equivalent. Of course Blackman’s results are of independent interest since they show that in the special cases he discusses the LSE is never much worse than the BLUE and, for example in the case described above, when the autocorrelation function of the noise,  $\exp(-\rho |t|)$ , declines quickly to zero the two are almost equivalent even for quite small values of  $T$ .

**Appendix. Proof of Formula (2)'**

Rewrite (1) as  $y(t) = \alpha m(t) + e(t)$ ,  $t \in (-\infty, T)$ , where  $\alpha$  is a scalar. Let  $L^2(-\infty, T)$  be the space of vectors,  $a(t)$ , of  $q$  components, for which

$$\int_{-\infty}^T a(t)^* a(t) dt < \infty .$$

Then  $L^2 \subset \mathcal{H}$  where  $\mathcal{H}$  is the space of vectors for which

$$\int \int_{-\infty}^T a^*(s) \Gamma(s-t) a(t) ds dt < \infty .$$

Indeed putting

$$\begin{aligned} \hat{a}(\lambda) &= \int_{-\infty}^T e^{-i\lambda t} a(t) dt \\ \int \int_{-\infty}^T a^*(s) \Gamma(s-t) a(t) ds dt &= \int_{-\infty}^{\infty} \hat{a}(\lambda)^* f(\lambda) \hat{a}(\lambda) d\lambda \\ &\leq c \int_{-\infty}^{\infty} \hat{a}(\lambda)^* \hat{a}(\lambda) d\lambda = 2\pi c \int_{-\infty}^T a^*(t) a(t) dt. \end{aligned}$$

Let  $\bar{L}^2$  be the closure of  $L^2$  in  $\mathcal{H}$ . The operator  $a \in \mathcal{H} \rightarrow \Gamma a \in \mathcal{H}$ ,

$$\Gamma a = \int_{-\infty}^T \Gamma(t-s) a(s) ds,$$

is well defined, maps  $\mathcal{H}$  into  $L^2$  and is bounded and Hermitian. Indeed

$$\Gamma a = \int_{-\infty}^{\infty} e^{i\lambda t} f(\lambda) \hat{a}(\lambda) d\lambda$$

where  $\hat{a}$  is now a limit with respect to  $f(\lambda)$  as weight function. Thus

$$\int_{-\infty}^{\infty} \hat{a}(\lambda)^* f^3(\lambda) \hat{a}(\lambda) d\lambda \leq \text{constant} \int_{-\infty}^{\infty} \hat{a}(\lambda)^* f(\lambda) \hat{a}(\lambda) d\lambda$$

since  $f(\lambda)$  is a continuous matrix function whose norm is uniformly bounded. Now we assert that if it is not possible to be, almost surely, to determine whether or not  $\alpha = 0$ , then it must be true that  $m(t) \in \Gamma \bar{\mathcal{H}}$  (i.e. the range of  $\Gamma$  as an operator in  $\bar{\mathcal{H}}$ ). If that is not so then  $m$  is not orthogonal to  $(\Gamma \bar{\mathcal{H}})^\perp$ , the orthogonal complements of  $\Gamma \bar{\mathcal{H}}$  in  $L^2$  (which is the same as  $(\Gamma \mathcal{H})^\perp$ ). Then there is a  $b(t) \in \Gamma$  such that

$$\begin{aligned} \int_{-\infty}^T b(t)^* m(t) dt &\neq 0 \\ \int \int_{-\infty}^T b(t)^* \Gamma(t-s) b(s) ds dt &= 0 \end{aligned}$$

(since the second expression merely requires that  $b(t)$  be orthogonal to  $\Gamma L^2$ ).  
But now, almost surely

$$\int_{-\infty}^T b(t)^* y(t) dt = \alpha \int_{-\infty}^T b(t)^* m(t) dt$$

since

$$\int_{-\infty}^T b(t)^* e(t) dt$$

has mean square

$$\int \int_{-\infty}^T b(t)^* \Gamma(t-s) b(s) ds dt = 0,$$

Thus  $m(t) \in \Gamma \bar{\mathcal{H}}$  i.e.

$$m(t) = \int_{-\infty}^{\infty} e^{it\lambda} f(\lambda) \hat{a}(\lambda) d\lambda$$

where  $\hat{a}(\lambda)$  is the Fourier transform of a function in  $L^2$  or is the limit in mean square with weight matrix  $f(\lambda)$ , of a sequence of such functions.

Now put  $\alpha = 1$  and

$$m(t) = \{I_q \otimes x(t)'\} \beta = \int_{-\infty}^{\infty} e^{it\lambda} f(\lambda) \hat{a}_\beta(\lambda) d\lambda$$

where evidently  $\hat{a}_\beta$  depends linearly on  $\beta$ . Since this relation is to hold for all  $\beta$  we must have

$$m(t) = \int_{-\infty}^{\infty} e^{it\lambda} f(\lambda) \Phi^{(T)}(\lambda) d\lambda,$$

where  $\Phi^{(T)}$  is as described below (2)'.

### References

- R. B. Blackman (1965), *Data Smoothing and Prediction*. (Addison-Wesley, U.S.A., 1965).  
 U. Grenander (1954), 'On the estimation of regression coefficients in the case of autocorrelated disturbance', *Ann. Math. Statist.* **25**, 252-272.  
 E. J. Hannan, (1968), 'Least squares efficiency for vector time series', *J. Roy. Statist. Soc. B* **30**, 490-498.  
 E. J. Hannan (1970), *Multiple Time Series*, (John Wiley, New York, 1970).  
 M. D. Heble (1961), 'A regression problem concerning stationary processes', *Trans. Amer. Math. Soc.* **99**, 350-371.  
 A. S. Kholevo (1969), 'On estimates of regression coefficients', *Theory of Probability and its Applications.* **15**, 79-104.  
 M. Rosenblatt (1959), 'Statistical analysis of stochastic processes with stationary residual' in "Probability and Statistics", the Harald Cramer volume (Ed. U. Grenander). Almqvist and Wiksells, Uppsala, 1959).

Yu. A. Rozanov (1964), Appendix to Russian translation of *Time Series Analysis* by E. J. Hannan. (Mir, Moscow, 1964).

A. M. Yaglom (1963), 'On the equivalence and perpendicularity of two Gaussian probability measures in function space', Chapter 22 in *Time Series Analysis* (Ed. M. Rosenblatt). (John Wiley, New York, 1963.)

Since this was written there has been much work on the topic by Russian writers. For a survey see A. S. Holevo (1973), 'On the general problem of mean estimation' *J. Multivariate Analysis*, **3**, 262–275.

Department of Statistics  
Institute of Advanced Studies  
The Australian National University  
P.O. Box 4  
Canberra, 2600, Australia