MATH. SCAND. 10 (1962), 137–144

AN ASYMPTOTIC RESULT FOR THE FINITE PREDICTOR

GLEN BAXTER[1]

1. Introduction.

In this note we will demonstrate two asymptotic results for the finite linear predictor in the case that there exists a spectral density with special properties. Our results can be contrasted with those in Grenander and Szegö [3, § 10.10]. The method used will involve an application of an inequality and a convergence equivalence for Szegö polynomials recently investigated by the author in a different connection [1].

Let $\{x_k\}_{-\infty}^{\infty} \in L_2(\Omega)$ be a stationary stochastic process with spectral density function $f(\theta)$ ($\not\equiv 0$). We assume that $x_0, x_{-1}, \ldots, x_{-n+1}$ are known and that it is desired to determine constants $a_k(n)$, $0 \leqq k \leqq n-1$, to obtain the best linear predictor

$$(1) \qquad x_1^* = \sum_{m=0}^{n-1} a_m(n)\, x_{-m}$$

of x_1. Finding $a_k(n)$ is equivalent to obtaining the minimum value of

$$(2) \left\| x_1 - \sum_{m=0}^{n-1} a_m(n) x_{-m} \right\|^2 = \frac{1}{2\pi} \int_{-\pi}^{\pi} |p_0 + p_1 z + \ldots + p_n z^n|^2 f(\theta)\, d\theta, \qquad z = e^{i\theta},$$

where $p_0 = 1$, $p_k = -a_{k-1}(n)$, and where $\|\cdot\|$ is the norm in $L_2(\Omega)$. This minimum is obtained for the polynomial

$$p(z) = \frac{D_n(f)}{D_{n-1}(f)}\, u_n(z)\,,$$

with minimum value

$$\mu_n = \frac{D_n(f)}{D_{n-1}(f)},$$

if $u_n(z)$ is the unique polynomial of degree n satisfying

$$(4) \qquad \frac{1}{2\pi} \int_{-\pi}^{\pi} u_n(z) f(\theta) e^{-ik\theta}\, d\theta = \delta_{k0}, \qquad 0 \leqq k \leqq n,\ z = e^{i\theta}\,,$$

Received March 13, 1962.

[1] Fulbright Research Fellow at the University of Aarhus, Denmark.

and where $D_n(f)$ is the Toeplitz determinant of order $n+1$ associated with $f(\theta)$. Two questions arise in this connection which will now be considered.

First, if $\log f(\theta)$ is integrable, then μ_n decreases as n tends to infinity to a limit $\mu > 0$ which represents a lower bound for the error of the finite predictor x_1^*. The rate of decrease of $\delta_n = \mu_n - \mu$ to zero is important in the asymptotic behavior of x_1^*. Results in this direction have been given in [3, § 10.10] and [2] and may be compared with the following special case of our Theorem 3.1: *If $f(\theta)$ is almost everywhere equal to a positive function with k integrable derivatives* ($k \geqq 2$), *then* $\delta_n = o(n^{\varepsilon+2-2k})$, $\varepsilon > 0$. Our result is an especially good improvement of the known results in case k is large.

The second question concerns the computation of the predictor itself, that is, the computation of $a_k(n)$. If $f(\theta)$ is sufficiently nice, $a_k(n)$ approaches a limit a_k as n tends to infinity, or equivalently, u_{nk} approaches a limit U_k. It can happen that $a_k(n)$ is difficult to compute exactly while the computation of a_k is relatively easy. For example, if $f(\theta)$ had the form $\exp(a \cos\theta)$, then the limiting values of $a_k(n)$ are easily computed from the unique solution of (4) with n set equal to infinity. In fact, the factorization $f(\theta) = \exp(\frac{1}{2}ae^{i\theta}) \exp(\frac{1}{2}ae^{-i\theta})$ in this case yields

$$\lim_{n\to\infty} u_n(z) = U(z) = \exp(-\tfrac{1}{2}az) .$$

Thus $u_{n0} = 1/\mu_n \to 1$ and $a_{k-1}(n) = -p_k(n) \to -(-\frac{1}{2}a)^k/k!$ The question is simply what additional error is introduced by using a_k instead of $a_k(n)$ in (1)? This additional error is in norm

$$\begin{aligned} \left\| x_1^* - \sum_{m=0}^{n-1} a_m x_{-m} \right\| &= \left\| x_0^* - \sum_{m=1}^{n} a_{m-1} x_{-m} \right\| \\ &= \left\| \sum_{m=0}^{n} (\mu_n u_{nm} - \mu U_m) x_{-m} \right\| \qquad (\mu_n u_{n0} = \mu U_0 = 1) \\ &\leqq \mu \left\| \sum_{m=0}^{n} (u_{nm} - U_m) x_{-m} \right\| + (\mu_n - \mu) \left\| \sum_{m=0}^{n} u_{nm} x_{-m} \right\| . \end{aligned} \tag{5}$$

The second term on the right in (5) has the order of $\delta_n = \mu_n - \mu$, but an investigation is still needed of the rate of convergence to zero of

$$\sigma_n^2 = \left\| \sum_{m=0}^{n} (u_{nm} - U_m) x_{-m} \right\|^2 = \frac{1}{2\pi} \int_{-\pi}^{\pi} \left| \sum_{m=0}^{n} (u_{nm} - U_m) z^m \right|^2 f(\theta) d\theta . \tag{6}$$

In this connection we prove as a special case: *If $f(\theta)$ is almost everywhere equal to a positive function with k integrable derivatives* ($k \geqq 2$), *then*

$\sigma_n = o(n^{\varepsilon+1-k})$, $\varepsilon > 0$. The fact that the estimate for σ_n is the square root of the estimate for δ_n seems to be intimately connected with the quantities themselves and is not a result of our methods.

As a final consideration in § 3 we will consider the case in which σ_n goes exponentially to zero and show that this is necessary and sufficient for $f(\theta)$ to be almost everywhere equal to a function which is analytic for real θ and has no real zeros.

2. Preliminary results.

The method hinges on an inequality associated with equations like (4). Let $f(\theta)$ be an integrable function on $-\pi \leqq \theta \leqq \pi$ with (complex) Fourier coefficients $\{c_k\}$. Two polynomials $g_n(\theta) = \sum_0^n g_{nm} e^{im\theta}$ and $h_n(\theta) = \sum_0^n h_{nm} e^{im\theta}$ are assumed to be related to each other by means of the equation

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} h_n(\theta) f(\theta) e^{-ik\theta} d\theta = g_{nk}, \quad 0 \leqq k \leqq n. \tag{7}$$

THEOREM 2.1. *If $f(\theta)$ is a positive continuous function whose Fourier coefficients $\{c_k\}$ have λ moments ($\lambda \geqq 0$), i.e. if $\sum_0^\infty m^\lambda |c_m| < \infty$, then there exists an integer N and a constant M, both depending only on $f(\theta)$, such that for every pair $g_n(\theta)$ and $h_n(\theta)$ of degree $n \geqq N$ satisfying* (7),

$$\sum_{m=0}^{n} (2^\lambda + m^\lambda)|h_{nm}| \leqq M \sum_{m=0}^{n} (2^\lambda + m^\lambda)|g_{nm}| . \tag{8}$$

For a proof of Theorem 2.1 the reader is referred to [1]. In [1] the inequality (8) is proved more generally with a "norm" function $\nu(m)$ replacing $2^\lambda + |m|^\lambda$. It is easy to see, however, that $\nu(m) = 2^\lambda + |m|^\lambda$ satisfies the norm condition $\nu(n) \leqq \nu(m)\nu(n-m)$ for all n, m.

The special case of Theorem 2.1 of interest to us is that in which $g_{nm} = \delta_{m0}$ for all n, m. Equation (7) reduces in this special case to equation (4) and (8) becomes

$$\sum_{m=0}^{n} (2^\lambda + m^\lambda)|u_{nm}| \leqq M$$

for all n sufficiently large. Now, according to the Wiener–Lévy Theorem [4, p. 245], if $f(\theta) > 0$ is a continuous function whose Fourier coefficients have λ moments, then $\log f(\theta)$ is an integrable function whose Fourier coefficients have λ moments. Let $\{d_k\}$ be the Fourier coefficients of $\log f(\theta)$. Under the conditions of Theorem 2.1, it is known that

(9) $$U(z) = \exp\left\{-\sum_{m=0}^{\infty} d_m z^m\right\}, \qquad z = e^{i\theta}$$

is the unique solution of

(10) $$\frac{1}{2\pi}\int_{-\pi}^{\pi} U(z) f(\theta) e^{-ik\theta} d\theta = \delta_{k0}, \qquad 0 \leqq k < \infty, \quad z = e^{i\theta},$$

and that $\{U_m\}$ has λ moments. A comparison of (4) and (10) indicates that as n tends to infinity $u_n(z)$ approaches $U(z)$. The exact rate of approach of $u_n(z)$ to $U(z) = \sum_0^\infty U_m z^m$ is crucial in our methods and it is exactly in this direction that Theorem 2.1 is useful.

THEOREM 2.2. *If $f(\theta)$ is a positive continuous function whose Fourier coefficients have λ moments and if $\{u_{nm}\}$ and $\{U_m\}$, $n, m \geqq 0$, are the Fourier coefficients of the functions $u_n(z)$ and $U(z)$ determined by* (4) *and* (10), *then there exists an integer N and a constant $\hat{M}$, both depending only on $f(\theta)$, such that for all $n \geqq N$*

(11) $$\sum_{m=0}^{n} (2^\lambda + m^\lambda) |u_{nm} - U_m| \leqq \hat{M} \sum_{m=n+1}^{\infty} (2^\lambda + m^\lambda) |U_m| .$$

PROOF. It is easy to see from (4) and (10) that $u_n(z)f(\theta)$ and $U(z)f(\theta)$ have the same Fourier coefficients δ_{k0} for $0 \leqq k \leqq n$. Thus for $0 \leqq k \leqq n$

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} \left[\sum_{m=0}^{n} (u_{nm} - U_m) e^{im\theta}\right] f(\theta) e^{-ik\theta} d\theta$$

$$= \frac{1}{2\pi}\int_{-\pi}^{\pi} \left[\sum_{m=n+1}^{\infty} U_m e^{im\theta}\right] f(\theta) e^{-ik\theta} d\theta = \hat{g}_{nk} \text{ (say)} .$$

According to (8) applied to the special case of (7) which appears just above ($n \geqq N$),

$$\sum_{m=0}^{n} (2^\lambda + m^\lambda) |u_{nm} - U_m| \leqq M \sum_{m=0}^{n} (2^\lambda + m^\lambda) |\hat{g}_{nm}| .$$

We finish the proof by deriving, by means of the norm condition $\nu(n) \leqq \nu(m)\nu(n-m)$ for $\nu(m) = 2^\lambda + |m|^\lambda$ and the explicit expression for $\hat{g}_{nk}$ above, the inequality

$$\sum_{m=0}^{n} (2^\lambda + m^\lambda) |\hat{g}_{nm}| \leqq \sum_{m=0}^{\infty} (2^\lambda + m^\lambda) |c_m| \sum_{m=n+1}^{\infty} (2^\lambda + m^\lambda) |U_m| .$$

This gives (11) with

$$\hat{M} = M \sum_{m=0}^{\infty} (2^\lambda + m^\lambda)|c_m| .$$

The polynomial $u_n(z)$ in (4) is closely related to the Szegö polynomial $\varphi_n(z)$ associated with $f(\theta)$. In the notation of [3, Chapter II], $u_n(z) = k_n \varphi_n{}^*(z)$, where $k_n{}^2 = D_{n-1}(f)/D_n(f)$. Our third and final preliminary result will employ this connection together with Theorem 2.2 to give a useful convergence result for the sequence of constant terms $\{\varphi_n(0)\}$ in the Szegö polynomials.

THEOREM 2.3. *If $f(\theta)$ is a positive continuous function whose Fourier coefficients have λ moments, then $\sum_0^\infty m^\lambda |\varphi_m(0)| < \infty$.*

PROOF. For the purpose of *this proof only* we will use the notation $\|F\| = \sum (2^\lambda + |m|^\lambda)|F_m|$ where F is an arbitrary integrable function, with Fourier coefficients F_m. We begin with the known identity (see [3, page 41])

$$k_n z \varphi_n(z) = k_{n+1} \varphi_{n+1}(z) - l_{n+1} \varphi^*{}_{n+1}(z) \tag{12}$$

in which $l_n = \varphi_n(0)$. Using $u_n(z) = k_n \varphi_n{}^*(z)$ and performing some manipulations, it follows from (12) that

$$u_n(z) - u_k(z) = \sum_{m=k+1}^{n} \alpha_m z^m \bar{u}_m(z^{-1}), \qquad \alpha_m = \frac{l_m}{k_m}, \tag{13}$$

where $\bar{u}_m(z^{-1})$ is the polynomial in $1/z$ with coefficients $\bar{u}_{nk}$. Using $U(z)$ as defined in (9) and at the same time suppressing dependence on z, let us now write (13) in the form

$$u_n - u_k + \sum_{m=k+1}^{n} \alpha_m z^m (\bar{U} - \bar{u}_m) = \bar{U} \sum_{m=k+1}^{n} \alpha_m z^m . \tag{14}$$

Due to the special form of $U(z)$ in (9) and the assumptions of the theorem, which imply that $\{d_k\}$ has λ moments, the reciprocal function $V(z) = 1/\bar{U}(z^{-1})$ $(z = e^{i\theta})$ is such that $\|V\| < \infty$. Dividing through (14) by $\bar{U}$ and taking norms yields

$$\left\| \sum_{k+1}^{n} \alpha_m e^{im\theta} \right\| \leqq \|V\| \cdot \|u_n - u_k\| + \|V\| \cdot \left\| \sum_{k+1}^{n} \alpha_m e^{im\theta} (\bar{U} - \bar{u}_m) \right\| .$$

Now, the first term on the right above is bounded according to Theorem 2.2. Moreover, if m is sufficiently large, then $\|V\| \cdot \|\bar{U} - \bar{u}_m\| < \varepsilon < 1$. Thus, for all k sufficiently large,

$$\left\| \sum_{k+1}^{n} \alpha_m e^{im\theta} \right\| \leqq \|V\| \cdot \|u_n - u_k\| + \varepsilon \left\| \sum_{k+1}^{n} \alpha_m e^{im\theta} \right\| .$$

This shows that $\sum_0^\infty (2^\lambda + m^\lambda)|\alpha_m| < \infty$. To finish the proof we remark that $k_n^{-2} \to \mu > 0$ and recall the relationship $\alpha_m = l_m/k_m$.

3. Application to the finite predictor.

In view of the preliminary results our asymptotic study of the finite predictor is now rather straightforward. To obtain the fullest power from the theorem of the last section, we will state our results in terms of moment conditions on the spectral density function $f(\theta)$.

THEOREM 3.1. *If $f(\theta)$ is a positive continuous function whose Fourier coefficients have λ moments ($\lambda > 0$), then $\delta_n = o(n^{-2\lambda})$ and $\sigma_n = o(n^{-\lambda})$.*

PROOF. We show first that $\delta_n = o(n^{-2\lambda})$. According to Theorem 2.3, $\sum m^\lambda |\varphi_m(0)| < \infty$. Thus [3, p. 188],

$$\begin{aligned} \delta_n &= \mu_n \mu \sum_{m=n+1}^{\infty} |\varphi_m(0)|^2 \\ &\leqq \mu_n \mu n^{-2\lambda} \sum_{m=n+1}^{\infty} m^{2\lambda} |\varphi_m(0)|^2 \qquad (15) \\ &= o(n^{-2\lambda}) \,. \end{aligned}$$

To show that $\sigma_n = o(n^{-\lambda})$ we write first from (6)

$$\sigma_n{}^2 \leqq c_0 \left(\sum_{m=0}^{n} |u_{nm} - U_m| \right)^2 , \qquad (16)$$

and then apply (11) with $\lambda = 0$, getting

$$\begin{aligned} \sigma_n{}^2 &\leqq c_0 \hat{M}^2 \left(\sum_{m=n+1}^{\infty} |U_m| \right)^2 \\ &\leqq c_0 \hat{M}^2 n^{-2\lambda} \left(\sum_{m=n+1}^{\infty} m^\lambda |U_m| \right)^2 \\ &= o(n^{-2\lambda}) \,. \end{aligned}$$

The statements in the introduction follow as special cases of the above, since if $f(\theta)$ is almost everywhere equal to a positive function with k integrable derivatives ($k \geqq 2$), then the Fourier coefficients of $f(\theta)$ have λ moments for any $\lambda < k - 1$.

The asymptotic estimate given for σ_n is just the square root of that given for δ_n. This fact seems to be intrinsic in the quantities under discussion and is not just a phenomenon of the method employed. Some light is shed in this direction by the example with spectral density $f(\theta) = 1 + \varrho^2 - 2\varrho \cos\theta$ $(0 < \varrho < 1)$. It is easy to show for this example that

$$\delta_n = \frac{\varrho^{2n}(\varrho^2-\varrho^4)}{1-\varrho^{2n+2}}.$$

On the other hand,

$$u_{nm} = \frac{\varrho^m - \varrho^{2n+2-m}}{1-\varrho^{2n+4}} \to \varrho^m ,$$

and a computation reveals that $c\varrho^n < \sigma_n < C\varrho^n$ for suitably chosen constants c and C, showing that in this example the "size" of σ_n is exactly the square root of the "size" of δ_n.

Following the lead of Grenander and Szegö [3], who find a necessary and sufficient condition that $\delta_n = \mu_n - \mu$ decreases exponentially to zero, we now take up the question of finding a necessary and sufficient condition that σ_n go exponentially to zero. Simply enough, the condition found for δ_n by Grenander and Szegö [3] works also for σ_n. Of course, there is always the basic assumption that $f(\theta) \geqq 0$ and $\log f(\theta)$ are integrable function on $-\pi \leqq \theta \leqq \pi$.

THEOREM 3.2. *A necessary and sufficient condition that σ_n goes at least exponentially to zero as n tends to infinity is that $f(\theta)$ coincide in $-\pi \leqq \theta \leqq \pi$ almost everywhere with a function which is analytic for real θ and has no real zeros.*

PROOF. To prove the necessity we combine the definition of $\sigma_n{}^2$ in (6) with the minimal property of μ_n mentioned in (2) (and after) to find first

$$\sigma_n{}^2 \geqq (u_{n0} - U_0)^2 \mu_n = \frac{(\mu_n - \mu)^2}{\mu_n \mu^2}.$$

Thus, if $\sigma_n{}^2$ goes at least exponentially to zero, so does $\delta_n = \mu_n - \mu$, and the necessity follows from the known result for δ_n. On the other hand, if the spectral density $f(\theta)$ satisfies the conditions of the theorem, then the Fourier coefficients $\{d_k\}$ of $\log f(\theta)$ go at least exponentially to zero. It follows from (9) that $|U_k| < K_1 a^k$ with $|a| < 1$. Finally, an application of (11) with $\lambda = 0$, gives in view of (16),

$$\begin{aligned}\sigma_n{}^2 &\leqq c_0 \left(\sum_0^n |u_{nm} - U_m| \right)^2 \\ &\leqq c_0 \hat{M}^2 \left(\sum_{m=n+1}^{\infty} |U_m| \right)^2 \\ &\leqq K_2 a^{2n} ,\end{aligned}$$

and the proof is complete.

REFERENCES

1. G. Baxter, *A norm inequality for a "finite-section" Wiener-Hopf equation*, Illinois J. Math. (to appear).
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. 76 (1954), 112–126.
3. U. Grenander and G. Szegö, *Toeplitz forms and their applications*, University of California Press, Berkeley, 1958.
4. A. Zygmund, *Trigonometrical series* (second edition), Cambridge Press, Cambridge, 1959.

UNIVERSITY OF MINNESOTA, MINNEAPOLIS, MINN., U.S.A.
AND
UNIVERSITY OF AARHUS, DENMARK