AN EXTENSION OF A THEOREM OF G. SZEGÖ AND ITS APPLICATION TO THE STUDY OF STOCHASTIC PROCESSES

BY

ULF GRENANDER AND MURRAY ROSENBLATT

1. Introduction. In this paper we study minimum problems associated with quadratic forms

\[ Q_n = c'M^{(n)}c \]

where \( c \) is a column vector with components \( c_0, c_1, \ldots, c_n \) and \( M^{(n)} \) is a Hermitian matrix with the elements

\[ m_{p,q}^{(n)} = \int_{-\pi}^{\pi} e^{i(p-q)\lambda} f(\lambda) d\lambda, \quad p, q = 0, 1, \ldots, n. \]

We denote the conjugate of the transpose of a matrix \( A \) by \( A' \). Here \( f(\lambda) \) is a nonnegative integrable function in \( (-\pi, \pi] \). We shall define \( f(\lambda) \) with period \( 2\pi \) on the real axis. Some of these minimum problems arise in the theory of stationary stochastic processes. These applications will be discussed in §5 [3].

Szegö [6] has studied the minimum \( \mu_n \) of \( Q_n \) subject to the restraint

\[ P_n(\alpha) = 1 \]

where

\[ P_n(\omega) = \sum_{p=0}^{n} c_p \omega^p. \]

He has shown that if \( |\alpha| < 1 \), the limit \( \mu \) of \( \mu_n \) as \( n \to \infty \) is positive if and only if

\[ \int_{-\pi}^{\pi} \log f(\lambda) d\lambda > -\infty. \]

Then we can write the formal Fourier expansion

\[ \log f(\lambda) \sim k_0 + 2 \sum_{l=1}^{\infty} (k, \cos \nu \lambda + l, \sin \nu \lambda). \]

Putting

Received by the editors November 22, 1952.

(!) This work has been carried out with the support of the Office of Naval Research.
EXTENSION OF A THEOREM OF G. SZEGÖ

\[ g(w) = \frac{k_0}{2} + \sum_{r=1}^{\infty} (k_r - i\lambda_r) w^r \]

and

\[ D(w) = e^{\psi(w)} \]

Szegö has shown that

\[ \mu = 2\pi |D(\alpha)|^2 (1 - |\alpha|^2) \]

We are going to study the minimization of the quadratic form \( Q_n \) with restraints

\[ C: \quad P^{(k)}(\alpha_j) = \beta_j^k, \quad k = 0, 1, \ldots, n, \; j = 1, \ldots, m, \]

where the \( \alpha_j \) are different points in the closed unit circle and where the \( \beta_j^k \) do not all vanish. The results of the paper are valid with appropriate modification when the restraints are of the form

\[ C^*: \quad P^{(k)}(\alpha_j) = \beta_j^k, \quad k \in S_j, \; j = 1, \ldots, m, \]

where \( S_j \) is a finite set of nonnegative integers. We have restricted ourselves to restraints of the form (1.2) in order to avoid excessive notation.

2. Conditions inside the unit circle. We order the pairs \((j, k)\) according to increasing \(j\) and for fixed \(j\) according to increasing \(k\); let \(r\) be the numbering index of these pairs, \(r = 1, 2, \ldots, N = \sum_{j=1}^{m} (n_j+1)\). Defining the inner product of two polynomials \(g(w), h(w)\) as

\[ (g, h) = \int g(e^{i\lambda}) [h(e^{i\lambda})]^* f(\lambda) d\lambda(\cap), \]

we introduce the orthonormal polynomials \(\phi_r(w)\), \(r = 0, 1, \ldots\), obtained by the Gram-Schmidt procedure from \(1, w, w^2, \ldots\) [6]. Then we can write

\[ P_n(w) = \sum_{r=0}^{n} d_r \phi_r(w) \]

so that

\[ Q_n = \sum_{r=0}^{n} |d_r|^2. \]

**Theorem 1**

\[ (2.1) \quad \mu_n = \beta'(H_n)^{-1} \beta, \quad n > N, \]

\(^{(\cdot \cdot \cdot)}^* \) denotes the conjugate of \((\cdot \cdot \cdot)\).

\(^{(\cdot \cdot \cdot)}\) We thank the referee for suggesting the simple proof of Theorem 1 given above.
where $\beta$ is a column vector with the $N$ components $\beta_r = \beta_j$ and $H_n$ is a nonsingular $N \times N$ matrix with elements
\[
h_{r,s} = \sum_{r=0}^{n} \phi_r^{(k)}(\alpha_j) \phi_r^{(k')}(\alpha_j'),
\]
$r \rightarrow (j, k), s \rightarrow (j', k').$

If $\log f(\lambda)$ is integrable and all the restraints (1.2) are at points $\alpha_j$ inside the unit circle (we call such a set of restraints $C_i$), then
\[
(2.2) \quad \mu = \lim_{n \to \infty} \mu_n = \beta'(H')^{-1}\beta
\]
where $H$ is a nonsingular $N \times N$ matrix with elements
\[
h_{r,s} = \frac{1}{2\pi} \left( \frac{\partial^k}{\partial x^k} \frac{\partial^{k'}}{\partial y^{k'}} 1 - xy^* D(x)[D(y)]^* \right), \quad x = \alpha_j, y = \alpha_j'.
\]

**Proof.** We have to minimize $Q_n = \sum_{n=0}^{N} |d_n|^2$, $n \geq N$, with the restraint
\[
\sum_{r=0}^{n} d_r \phi_r^{(k)}(\alpha_j) = \beta_j^k, \quad \text{or} \quad \sum_{r=0}^{n} d_r l_{rr} = \beta_r,
\]
where $l_{rr} = \{\phi_r^{(k)}(\alpha_j)\}$ in $n$-space so that the restraints have the form $d'l_r = \beta$, and $|d|^2$ has to be minimized. The vectors $l_r$ are linearly independent since
\[
\sum_{j,k} d_r \phi_r^{(k)}(\alpha_j) = 0,
\]
would mean that
\[
\sum_{j,k} l_{jk} \phi_r^{(k)}(\alpha_j) = 0
\]
for any polynomial $f(z)$ of degree $n$. By proper choice of $f$ we find $l_r^* = 0$.

Now the vectors $l_r^*$ span a linear manifold
\[
\sum_{r=1}^{N} \lambda_r l_r^* = 0,
\]
and the projection of $d$ onto this furnishes the minimum of $|d|$. So assume $d$ has the latter form. The restraints are now
\[
\sum_{r=1}^{N} \lambda_r (l_s, l_r) = \beta_s, \quad s = 1, 2, \ldots, N.
\]

The Hermitian matrix $H_n = [(l_s, l_r)]$ is positive definite. Introducing the column vector $\lambda = \{\lambda_r\}$, we can write these equations as follows: $H_n \lambda = \beta$ so that $\lambda = H_n^{-1}\beta$. The minimum in question is
1954] EXTENSION OF A THEOREM OF G. SZEGÖ 115

\[ \left| \sum_{r,s=0}^{N} \lambda_r \lambda_s (l_r, l_s) \right|^2 = \sum_{r,s} \lambda_r \lambda_s = \lambda' H_n \lambda = \beta' (H'_{n})^{-1} H_n \beta = \beta' (H'_{n})^{-1} \beta. \]

Substituting we obtain for the elements of \( H_n \):

\[(l_r, l_s) = \sum_{r=0}^{n} \phi_r (\alpha_j) [\phi_r (\alpha_j')]^*.

But

\[ \lim_{n \to \infty} \sum_{r=0}^{n} \phi_r (x) [\phi_r (y)]^* = \frac{1}{2\pi} \frac{1}{1 - \alpha_j^* y} \frac{1}{D(x) [D(y)]^*} \]

uniformly for \(|x|, |y| \leq r < 1\) [6, Satz XXXI]. From this it easily follows that

\[ \lim_{n \to \infty} \sum_{r=0}^{n} \phi_r (\alpha_j) [\phi_r (\alpha_j')]^* = \frac{1}{2\pi} \left( \frac{\partial^k}{\partial x^k} \frac{\partial^{k'}}{\partial y^{k'}} \frac{1}{1 - \alpha_j^{*} y} \frac{1}{D(x) [D(y)]^*} \right), \quad x = \alpha_j, \ y = \alpha_j', \]

and we have (2.2).

3. Conditions on the unit circle. We now consider the restraints (1.2) at points \( \alpha_j \) on the unit circle and call such a set of restraints \( C_0 \). This case differs considerably from that just treated. Here we get \( \mu = 0 \) and are mainly interested in the principal term of \( \mu_n \) as \( n \to \infty \). To study this we have to introduce certain regularity conditions on \( f(\lambda) \).

**Theorem 2.** Let

\[ f(\lambda) = g(\lambda) \prod_{\rho = 1}^{l} | e^{\rho} - e^{\theta_{\rho}^{*}} |^{2l_{\rho}}, \quad -\pi < \theta_{\rho} \leq \pi, \]

where \( g(\lambda) \) is positive and continuous and the \( l_{\rho} \) are positive integers. Then

\[ \mu_n = \frac{2\pi}{n^{2p+1}} \sum_{r_{j} + l_{j} = p} \left( \frac{r_{j} + l_{j}}{r_{j}} \right)^{2} \left| \beta_{j}^{l_{j}} \right|^{2} d(\rho, \nu_{j}) \cdot \prod_{e^{\theta_{\rho}^{*} \alpha_{j}}} | \alpha_{j} - e^{\theta_{\rho}^{*}} |^{2l_{\rho}} + O \left( \frac{1}{n^{2p+1}} \right), \quad \alpha_{j} = e^{\theta_{j}}, \]

where \( \rho, r_{j}, t_{j}, \nu_{j} \) and \( d(\rho, \nu_{j}) \) are defined below.

**Proof.** Choose two trigonometric polynomials \( a(w), b(w) \) in \( w = e^{\alpha} \) of order \( p \) so that

\[ \frac{1}{2\pi} \left| a(e^{\alpha}) \right|^{2} \leq g(\lambda) \leq \frac{1}{2\pi} \left| b(e^{\alpha}) \right|^{2}, \quad |b(e^{\alpha})| - |a(e^{\alpha})| < \epsilon. \]
Let us now consider the case when \( g(\lambda) \) is exactly equal to
\[
\frac{1}{2\pi} \left| a(e^{i\lambda}) \right|^2 = \frac{1}{2\pi} \left| \sum_{r=0}^{p} a_r e^{ir\lambda} \right|^2.
\]

Then we should minimize under conditions \( C_b \)
\[
\frac{1}{2\pi} \int_{-\pi}^{\pi} \left| P_n(e^{i\lambda}) \rho(e^{i\lambda}) \right|^2 d\lambda
\]
where
\[
\rho(w) = a(w) \prod_{j} (w - z_r)^{\nu_j}, \quad z_r = e^{i\theta_r}.
\]

Let \( q(w) = P_n(w) \rho(w) \). The problem can then be rephrased in the following manner. We minimize the integral
\[
\frac{1}{2\pi} \int_{-\pi}^{\pi} \left| q(e^{i\lambda}) \right|^2 d\lambda
\]
under the conditions, say \( C_b^* \), that the conditions \( C_b \) induce on \( q(w) \). It is clear that the \( C_b^* \) are of a similar form
\[
q^{(k)}(w_j) = \eta_j, \quad k = 0, 1, \ldots, \nu_j, \quad j = 1, 2, \ldots, M,
\]
where \( w_j \) is one of the \( \alpha_j \)'s or \( z_r \)'s. The range of \( (j, k) \) is not necessarily the same as in \( C_b \) but we carry out in a similar mapping of \( (j, k) \) onto a single index \( r \). Then we get from (2.1)

\[
\mu_n = \eta' (H_n')^{-1} \eta.
\]

To compute \( H_n \) we observe that in the present case \( f(\lambda) = 1/2\pi \) so that \( \phi(w) = w \). Hence if \( |x| = |y| = 1 \),
\[
\sum_{r=0}^{n} \phi^{(k)}(x) [\phi^{(k')}](y)^* = x^{-k} y^{-k'} \sum_{r=0}^{n} \nu(\nu - 1) \cdots (\nu - k + 1) \nu(\nu - 1) \cdots (\nu - k' + 1) x^{r} y^{-r}.
\]

If \( x = y \) we get
\[
x^{-k-k'} \frac{n^{k+k'+1}}{k + k' + 1} + O(n^{k+k'}).\]

If \( x \neq y \) we use the Abel summation formula and find that the expression is \( O(n^{k+k'}) \). Hence
where

\[ H_n = D_n \Lambda_n D_n \]

(3.4)

\[ D_n = \begin{pmatrix}
    n^{1/2} & 0 & & \\
    n^{1+1/2} & & & \\
    & \ddots & \ddots & \\
    n^{\gamma_1+1/2} & & & \\
    0 & n^{1/2} & & \\
    & \ddots & \ddots & \\
    & & 0 & n^{\gamma_M+1/2}
\end{pmatrix} \]

(3.5)

and

\[ \Lambda_n = \Lambda + \frac{1}{n} R_n \]

where \( R_n \) has bounded elements. Here we have put

\[ \Lambda = \begin{pmatrix}
    \gamma_1 & 0 & & \\
    & \gamma_2 & & \\
    & & \ddots & \\
    & & & \gamma_M
\end{pmatrix} \]

where

\[ \gamma_j = \left\{ \frac{w_j^{v+\mu}}{v + \mu + 1} ; v, \mu = 0, 1, \ldots, \nu_j \right\}. \]

Now note that

\[ q^{(k)}(w) = \sum_{s=0}^{k} \binom{k}{s} P_{\nu}^{(s)}(w) \beta^{(k-s)}(w). \]

If \( z \) does not coincide with any \( \alpha_j \), the conditions induced at \( z \) are

\[ q^{(k)}(z) = 0, \quad k = 0, \ldots, l - 1. \]

If \( \alpha_j \) does not coincide with any \( z \), the condition at \( \alpha_j \) become

\[ q^{(k)}(\alpha_j) = \sum_{s=0}^{k} \binom{k}{s} \beta_j^{(k-s)}(\alpha_j), \]

and if \( \beta_j^{(k)} \) is the first nonzero \( \beta \) corresponding to \( \alpha_j \)

\[ q^{(r_j)}(\alpha_j) = \beta_j^{(r_j)} a(\alpha_j) \prod_{j \neq j} (\alpha_j - z), \]
and
\[ q^{(k)}(\alpha_j) = 0, \quad k < r_j. \]
Now let \( \alpha_j = z_\mu \). Let \( r_j \) be defined as above. Then
\[ q^{(k)}(z_\mu) = 0, \quad k < r_j + l, \]
\[ q_{(r_j + l)}^{(k)}(z_\mu) = \left( r_j + l \right) \beta_j a(\alpha_j) \prod_{\mu \neq \nu} (z_\mu - z_\nu)^{l_\mu}. \]

Let \( t_j \) be the order of the zero of \( q(w) \) at \( \alpha_j \). If there is no zero we set \( t_j = 0 \) and if \( \alpha_j = z_\nu \), then \( t_j = t_\nu \). Let the least \( k \) such that there is a nonzero \( \beta_j \) be called \( p \). It follows from (3.3), (3.4), and (3.5) that the leading term of \( \mu_n \) will consist only of contributions from the conditions with \( k = p \). It is clear from the discussion above that we then get
\[ \mu_n = \frac{1}{n^{2p+1}} \sum_{r_j + t_j = p} \left( r_j + l \right)^2 \left| \beta_j \right|^2 a(\alpha_j) \prod_{\mu \neq \nu} (\alpha_j - z_\nu)^{l_\mu} \cdot d(\rho, \nu_j) + O\left( \frac{1}{n^{2p+2}} \right) \]
where
\[ d(\rho, n) = \left\{ \frac{1}{\nu + \mu + 1}; \nu, \mu = 0, \ldots, n \right\}_{\rho, \rho}^{-1} \]
\[ = (2p + 1) \prod_{j=0, j \neq \rho}^{n} \left( \frac{j + \rho + 1}{j - \rho} \right)^2 \]
(see [1, p. 177]).

The true value of \( \mu_n \) if \( f(\lambda) \) is given by (3.1) will be between the two values computed for \( a(\lambda), b(\lambda) \). Letting \( \epsilon \) tend to zero we get the desired result.

**Remark 1.** If none of the zeros of \( f(\lambda) \) coincide with any \( \alpha_j \), then \( \rho \) is simply the smallest \( k \) such that there is a nonzero \( \beta_j \). We then have
\[ \mu_n = \frac{2\pi}{n^{2p+1}} \sum_{j} \left| \beta_j \right|^2 \cdot d(\rho, n_j) + O\left( \frac{1}{n^{2p+2}} \right). \]

**Remark 2.** The case when there are conditions both inside and on the unit circle is now easily handled. If some \( \beta \) in \( C_\iota \) does not vanish, then \( \mu > 0 \) and its value can be computed from (2.2) as if no conditions on the circle had been present, as is easily verified. On the other hand, if all \( \beta \)'s in \( C_\iota \) vanish, \( \mu_n \to 0 \) and we get its principal term from (3.2) as if \( C_\beta \) were the only conditions.

Theorem 2 suggests that the error \( \mu_n' \) computed with the minimizing poly-
nominal corresponding to a uniform weight function is of the same order as
the error \( \mu_n \) computed with the minimizing polynomial corresponding to
weight function \( f(\lambda) \). The following theorem indicates that this conjecture is
essentially true [4].

**Theorem 3.** Let \( f(\lambda) \) be a nonnegative continuous function having no zeros
in common with the points \( \alpha_j \) of the conditions \( C_b \). Let

\[
\mu'_n = \int_{-\pi}^{\pi} \left| P_n(e^{i\theta}) \right|^2 f(\lambda) d\lambda
\]

where \( P_n(w) \) minimizes \( \int_{-\pi}^{\pi} \left| P_n(e^{i\theta}) \right|^2 d\lambda \) under conditions \( C_b \). Then

\[
\lim_{n \to \infty} \frac{\mu'_n}{\mu_n} = 1.
\]

**Proof.** Let \( P_n(w) = \sum_{0}^{n} \gamma_j w^j \) be the minimizing polynomial under condi-
tions \( C_b \) and the assumption that the spectral density is uniform \( (f(\lambda) = 1) \). Let

\[
\frac{d^k}{dw^k} w^r = \psi_{r,k}(w).
\]

Then

\[
\frac{1}{2\pi} \int_{-\pi}^{\pi} \left| P_n(e^{i\theta}) \right|^2 d\lambda = \sum_{0}^{n} |\gamma_r|^2
\]

where

\[
\gamma_r = \sum_{j,k} \lambda_j^k [\psi_{r,k}(\alpha_j)]^*
\]
or

\[
\gamma = \psi \lambda
\]

where

\[
\psi = \begin{bmatrix}
[\psi_{0,0}(\alpha_1)]^* & [\psi_{0,1}(\alpha_1)]^* & \cdots & [\psi_{0,n}(\alpha_1)]^* & [\psi_{0,0}(\alpha_2)]^* & \cdots \\
[\psi_{1,0}(\alpha_1)]^* & [\psi_{1,1}(\alpha_1)]^* & \cdots & [\psi_{1,n}(\alpha_1)]^* & [\psi_{1,0}(\alpha_2)]^* & \cdots \\
\cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\
[\psi_{n,0}(\alpha_1)]^* & [\psi_{n,1}(\alpha_1)]^* & \cdots & [\psi_{n,n}(\alpha_1)]^* & [\psi_{n,0}(\alpha_2)]^* & \cdots
\end{bmatrix}
\]

and \( \lambda = H_n^{-1} \beta \). As before we factor \( H_n \) so that \( H_n = D_n \Lambda_n D_n \). Then

\[
\mu'_n = \sum \gamma M^{(n)} \gamma = \lambda \psi M^{(n)} \psi \lambda = \beta'(H_n)^{-1} \psi M^{(n)} \psi H_n^{-1} \beta
\]

\[
= \beta' D_n^{-1} (\Lambda_n)^{-1} D_n^{-1} \psi M^{(n)} D_n^{-1} \Lambda_n D_n^{-1} \beta.
\]
But
\[
D_n^{-1} \psi = \left\{ \left[ \psi_{\nu, k}(\alpha_j) \right]^* \right\}.
\]

Hence a typical element of \(D_n^{-1} \psi M^{(n)} \psi D_n^{-1}\) is
\[
\sum_{\nu, k=0}^{n} \frac{\psi_{\nu, k}(\alpha_j)}{n^{b+1/2}} \left[ \psi_{\mu, k'}(\alpha_{j'}) \right]^* = \frac{1}{n^{b+k+1}} \int_{-\pi}^{\pi} \sum_{\nu, k=0}^{n} e^{i(\nu-k)\lambda}(\nu - 1) \cdots (\mu - k - 1) \mu(\mu - 1) \cdots (\nu - k + 1) \alpha_j^{n-k} \alpha_{j'}^{n-k'} f(\lambda) d\lambda.
\]

(3.6)

Let \(\Lambda = \lim_{n \to \infty} \Lambda_n\) as before. We shall show that
\[
D_n^{-1} \psi M^{(n)} \psi D_n^{-1} = f\Lambda + O(1)
\]
where
\[
f = \begin{pmatrix}
f(\lambda_1) & 0 \\
\vdots & \ddots & 0 \\
f(\lambda_n) & \ddots & f(\lambda_2) \\
0 & \ddots & \ddots & \ddots \\
\end{pmatrix}, \quad \alpha_j = e^{\lambda_j}.
\]

First suppose \(\alpha_j \neq \alpha_{j'}\). Then (3.6) can be rewritten as
\[
\alpha_j^{n-k} \alpha_{j'}^{n-k'} \int_{-\pi}^{\pi} e^{i(k' - k)\lambda} \eta_{n, k} \eta_{n, k'} (e^{i(\lambda - \lambda_j)}) \eta_{n, k'} (e^{i(\lambda - \lambda_{j'}))}) f(\lambda) d\lambda
\]
where
\[
\eta_{n, k}(x) = \frac{1}{n^{b+1/2}} \frac{d^k}{d^w} \left( \frac{1 - w^{n+1}}{1 - w} \right), \quad w = x.
\]

If \(|\lambda - \lambda_j| \geq \epsilon]\),
\[
| \eta_{n, k}(e^{i(\lambda - \lambda_j)}) | < \frac{c(\epsilon)}{n^{1/2}}
\]
while if \(|\lambda - \lambda_j| < \epsilon]\).
Hence if \( \lambda_j \neq \lambda_j' \), expression (3.6) converges to zero. Now consider \( \alpha_j = \alpha_j' = e^{bi} \). Then (3.6) can be rewritten as

\[
\frac{1}{n(k+k'+1)e^{i(k'-k)\lambda}} \sum_{r,m=0}^{n} \nu(\nu - 1) \cdots (\nu - k + 1) \mu(\mu - 1) \cdots (\mu - k' + 1) e^{i(\nu-k)\lambda} f(\lambda) d\lambda.
\]

Let

\[
s_k(\lambda) = \frac{1}{n^{k+1/2}} \sum_{r=0}^{n} e^{i(\nu-k)\lambda} \nu(\nu - 1) \cdots (\nu - k + 1).
\]

Consider

\[
K_n(\lambda) = \sum_{i,k=1}^{n} y_i s_i(\lambda) \left[y_k s_k(\lambda)\right]^* = \left| \sum_{i,k=1}^{n} y_i s_i(\lambda) \right|^2 \geq 0.
\]

One can verify that

1. \( K_n(\lambda) \to 0 \) uniformly in \( \lambda \) if \( |\lambda - \lambda_j| > \epsilon \),
2. \( \lim_{n \to \infty} \int_{-\pi}^{\pi} K_n(\lambda) d\lambda = 2\pi \sum_{k+l+1} y_k y_l^* \).

Hence (see [5, p. 49])

\[
\lim_{n \to \infty} \int_{-\pi}^{\pi} s_i(\lambda) s_k(\lambda) f(\lambda) d\lambda = \frac{2\pi f(\lambda_j)}{l + k + 1}.
\]

But (3.7) then follows immediately. Now

\[
\mu' = \beta' D_n^{-1}((\Lambda')^{-1} + o(1))(fA + o(1))((\Lambda^{-1} + o(1))D_n^{-1} \beta
\]

\[
= \beta' D_n^{-1}((\Lambda')^{-1} + o(1))fD_n^{-1} \beta.
\]

The theorem then follows.

4. The approach of \( \mu_n - \mu \) to zero under the restraint \( P_n(0) = 1 \). This problem is of interest in the theory of stochastic processes. Moreover, it does give some insight into the more general problem where the restraints are of the type \( C_i \). Let \( \delta_n = \mu_n - \mu \).

**Theorem 4.** The decrease of \( \delta_n \) to zero is at least exponential if and only if

1. \( f(\lambda) \) coincides almost everywhere with a function \( g(\lambda) \) that is analytic for all real \( \lambda \) and
2. \( g(\lambda) \) has no zeros.

**Proof.** Assume that (1) and (2) are satisfied. Then the function

\[
\phi(w) = f(\lambda), \quad w = e^{i\lambda}, \quad -\infty < \lambda < \infty,
\]
where we have chosen one determination of the logarithmic function, can be analytically extended to an annular region $\rho_1 < |w| < \rho_2$, $\rho_1 < 1 < \rho_2$. In this region we can then represent $\log \phi(w)$ as a convergent Laurent series

$$
\log \phi(w) = \sum_{\gamma_n} \gamma_n w^n.
$$

We then note that

$$
D(w) = \exp \left\{ \gamma_0 + \sum_{n=1}^{\infty} \gamma_n w^n \right\}
$$

and hence $D(w)$ is analytic in the closed region $|w| \leq 1$ and has no zeros in this region. One can define the inner product of two functions $g(w), h(w)$ such that $g(w)D(w), h(w)D(w) \in H_2$ as in 2. Then $\|g\|^2 = (g, g)$ and the set of functions $g(w)$ such that $g(w)D(w) \in H_2$ is a Hilbert space. Now

$$
\beta_n - \beta \leq \|s_n(w)\|^2 - \|D(0)/D(w)\|^2
$$

where $s_n(w) = \sum_0^n d_n w^n$ is the nth partial sum of the Taylor expansion of $D(0)/D(w)$. But $|d_n| < d < 1$ so that $\delta_n \leq K d^{n/2}, 0 < d < 1$.

Now assume $\delta_n \leq K d^n, 0 < d < 1$. Then

$$
|\phi_n(0)| < K d^n/2.
$$

However, $|\phi_n(w)| < K d^{|w|}$ on $|w| = 1 + \epsilon, \epsilon > 0$ (see [6, Satz XXXII]). If $1 + \epsilon < 1/d^{1/2}$ we have uniform convergence of

$$
2\pi \sum_0^n [\phi_n(0)]^* \phi_n(w)
$$

so that $2\pi \sum_0^n [\phi_n(0)]^* \phi_n(w)$ represents an analytic function in $|w| < 1 + \epsilon$. However it coincides with

$$
\frac{1}{D(0)^*D(w)}
$$

when $|w| < 1$ (see [6]). We can then extend $1/D(w)$ analytically into $|w| < 1 + \epsilon$. But then $D(w)$ is analytic and different from zero in $|w| < 1 + \epsilon$. But we have except on a set of measure zero

$$
f(\lambda) = |D(e^{i\lambda})|^2 = D(w)D^*(1/w), \quad w = e^{i\lambda},
$$

where $D^*(w)$ denotes the function obtained from $D(w)$ by taking the conjugates of its Taylor coefficients. From this the result follows.
Let us note that Theorem 4 is true more generally for conditions of the type $C_i$.

We have seen that if $\delta_n$ decreases exponentially, $f(\lambda)$ cannot have any essential zeros. In the following theorem we study what happens when $f(\lambda)$ has zeros.

**Theorem 5.** If $f(\lambda)$ coincides almost everywhere with

$$g(\lambda) \prod_{r=1}^{N} \left| e^{i\lambda} - e^{i\lambda_r} \right|^{2\gamma_r},$$

where $g(\lambda)$ is positive and has an integrable third derivative, then

$$\delta_n = O(1/n).$$

The order is attained for some such $f(\lambda)$.

**Proof.** If $g(\lambda)$ has an integrable third derivative, $D(w)$ has a bounded derivative on $|w| = 1$. Repeating an argument used in the proof of Theorem 2, we see that

$$\mu_n = \beta' (H_{n+M})^{-1} \beta, \quad M = \sum l_r,$$

where $\beta$ has its first component equal to one and the remaining components are zero and

$$H_{n+M} = D_n \left( \begin{array}{ccc} \Lambda & 0 & 0 \\ 0 & \gamma_1 & 0 \\ \vdots & \vdots & \gamma_p \end{array} \right) + \frac{1}{n} B_n D_n$$

where $B_n$ is bounded and

$$D_n = \left( \begin{array}{cccc} 1 & n^{1/2} & 0 & \cdots \\ n^{1/2} & \ddots & \ddots & \ddots \\ 0 & \cdots & \ddots & n^{1/2} \\ \end{array} \right).$$

The theorem follows immediately. The bound can be realized by $f(\lambda) = |1 - e^{i\lambda}|^2$.

5. **Applications to stochastic processes.** Consider a discrete stochastic process $x_t$, $-\infty < t < \infty$, with mean value $m_t = E x_t$. We assume that the second order moments $E |x_t|^2$ exist and that the reduced process $y_t = x_t - m_t$
is stationary in the wide sense, that is
\[ \rho_{t,s} = E y_t y_s^* = r_{t-s}. \]
Then we know that
\[ r_\ell = \int_{-\ell}^{\ell} e^{i\lambda} dF(\lambda) \]
where \( F(\lambda) \) is bounded and nondecreasing in \((-\pi, \pi)\). The completely non-deterministic processes form an important subclass. They are completely characterized by

1. \( F(\lambda) \) absolutely continuous,
   \[ F(\lambda) = \int_{-\ell}^{\ell} f(t) dt, \]
   and
   \[ \int_{-\ell}^{\ell} \log f(\lambda) d\lambda > -\infty. \]

The nonnegative function \( f(\lambda) \) is called the spectral density of the process.

In various problems one is interested in minimizing the variance of a linear form \( \sum c_\ell x_\ell \) subject to some conditions on the \( c_\ell \)'s. But this variance is
\[ \sum_{\ell,\mu=0}^n c_\ell c_\mu^* r_{\ell-\mu} \]
which is of the form (1.1). We shall consider some problems of this type.

1. Let us first assume \( m_\ell = 0 \). Having observed \( x_1, x_2, \ldots, x_n \) we want to form a linear combination \( \sum c_\ell x_\ell \) such that
   \[ E \left| x_0 - \sum_{\ell=1}^n c_\ell x_\ell \right|^2 = \min. \]
   This is a familiar problem of extrapolation. This is the type of problem treated in the previous sections since we can write it as
   \[ \mu_n = \int_{-\ell}^{\ell} | P_n(e^{i\lambda})|^2 f(\lambda) d\lambda = \min, \]
   \[ P_n(0) = 1. \]
   The only restraint is inside the unit circle and we then know that \( \mu_n \) tends to a positive value \( \mu \) as \( n \to \infty \). From Theorem 1 we get
   \[ \mu = 2\pi | D(0) |^2 = 2\pi \exp \left\{ \frac{1}{2\pi} \int_{-\ell}^{\ell} \log f(\lambda) d\lambda \right\}. \]
The "predictor" is

$$\sum_{r=1}^{n} c_r x_r = \int_{-\pi}^{\pi} \sum_{r=1}^{n} c_r e^{ir\lambda} dZ(\lambda)$$

where $Z(\lambda)$ is the orthogonal process corresponding to the stationary process $x_t$ (see [2]). The limit in the mean of this stochastic variable as $n \to \infty$ is

$$\int_{-\pi}^{\pi} \left[ 1 - \frac{D(0)}{D(e^{i\lambda})} \right] dZ(\lambda).$$

(5.1) and (5.2) can be found in [7].

2. A slightly more general case is extrapolation $k$ steps back, i.e., given a sample $x_k, x_{k+1}, \ldots, x_n$ to predict $x_0$. We see that this corresponds to the conditions

$$P(0) = 1,$$

$$P'(0) = 0,$$

$$\cdots$$

$$P^{(k-1)}(0) = 0.$$

This is again a context treated in Theorem 1. Analogues of expressions (5.1) and (5.2) in this case can be found in a similar way.

3. Suppose that $m_1$ is equal to an unknown constant $m$. From the sample $x_0, x_1, \ldots, x_n$ we want to construct a linear, unbiased estimate $m^*$ of minimum variance. It is immediately seen that this is equivalent to the minimization of $\int_{-\pi}^{\pi} \left| P_n(e^{i\lambda}) \right|^2 f(\lambda) d\lambda$ under the condition $P_n(1) = 1$. As this is a condition on the unit circle, we can apply Theorem 2 which gives us

$$E \left| m^* - m \right|^2 \sim \frac{2\pi f(0)}{n}.$$

Theorem 3 implies that if $f(0) \neq 0$ and $f(\lambda)$ is continuous, we get an asymptotically equivalent estimate by solving the same minimization problem for a uniform spectral density. But that would give us just the empirical mean

$$x^* = \frac{1}{n+1} \sum_{0}^{n} x_r.$$

4. Let $m_1 = mt(t-1) \cdots (t-k+1)e^{i\lambda_0}$. If we now want to get a linear unbiased estimate of minimum variance $m^*$ of $m$ we have the condition $P_n^{(k)}(e^{i\lambda_0}) = e^{-ik\lambda_0}$.

Again we can use Theorem 2 and get

$$E \left| m^* - m \right|^2 \sim \frac{2\pi f(e^{i\lambda_0})}{n^{2k+1}} (2k + 1).$$
We can also apply Theorem 3.

5. If we are interested in polynomial or trigonometric regression we put

\[ \phi_{i}^{(k,j)} = i(t - 1) \cdots (t - k + 1)e^{i\theta_{j}}, \]

\[ m_{t} = \sum_{k,j} c_{k,j} \phi_{i}^{(k,j)} \]

where the regression coefficients \( c_{k,j} \) are unknown. To get an unbiased minimum variance estimate of \( c_{k,j} \) we have the conditions

\[ P^{(l)}(e^{i\theta_{j}}) = 0 \quad \text{if} \quad (l, s) \neq (k, j), \]

\[ P^{(k)}(e^{i\theta_{j}}) = e^{-ik\theta_{j}}. \]

All the conditions are on the unit circle.

6. If \( m_{t} = m \) is unknown and we wish to predict the value of \( x_{0} \) from

\[ x_{k}, x_{k+1}, \ldots, x_{n} \]

it may be advantageous to use an unbiased predictor

\[ \sum_{r=k}^{n} c_{r}x_{r}, \quad \sum_{r=k}^{n} c_{r} = 1. \]

We then get the conditions

\[ P(0) = 1, \]

\[ P'(0) = 0, \]

\[ \ldots \ldots \ldots \ldots \ldots \]

\[ P^{(k-1)}(0) = 0, \]

\[ P(1) = 0. \]

It follows from Remark 2 that the limiting variance of this predictor is the same as that of the predictor in problem 2.

References


University of Chicago,
Chicago, Ill.