AN EXTENSION OF A THEOREM OF G. SZEGÖ AND ITS APPLICATION TO THE STUDY OF STOCHASTIC PROCESSES(1)

BY

ULF GRENANDER AND MURRAY ROSENBLATT

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

$$Q_n = c'M^{(n)}c \tag{1.1}$$

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

$$m_{p,q}^{(n)} = \int_{-\pi}^{\pi} e^{i(p-q)\lambda} f(\lambda) d\lambda, \qquad p, q = 0, 1, \cdots, n.$$

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 2π 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 μ_n of Q_n subject to the restraint

$$P_n(\alpha) = 1$$

where

$$P_n(w) = \sum_{\nu=0}^{n} c_\nu w^\nu.$$

He has shown that if $|\alpha| < 1$, the limit μ of μ_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

$$\log f(\lambda) \sim k_0 + 2\sum_{\nu=1}^{\infty} (k_\nu \cos \nu\lambda + l_\nu \sin \nu\lambda).$$

Putting

Received by the editors November 22, 1952.

(1) This work has been carried out with the support of the Office of Naval Research.

$$g(w) = \frac{k_0}{2} + \sum_{\nu=1}^{\infty} (k_\nu - il_\nu)w^\nu$$

and

$$D(w) = e^{g(w)},$$

Szegö has shown that

$$\mu = 2\pi \mid D(\alpha) \mid^2 (1 - \mid \alpha \mid^2).$$

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

(1.2) $$C: \quad P^{(k)}(\alpha_j) = \beta_j^k, \qquad k = 0, 1, \cdots, n_j, j = 1, \cdots, m,$$

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

$$C^*: \quad P^{(k)}(\alpha_j) = \beta_j^k, \qquad k \in S_j, j = 1, \cdots, m,$$

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, $r = 1, 2, \cdots, N = \sum_{j=1}^{m} (n_j + 1)$. Defining the inner product of two polynomials $g(w)$, $h(w)$ as

$$(g, h) = \int_{-\pi}^{\pi} g(e^{i\lambda}) [h(e^{i\lambda})]^* f(\lambda) d\lambda (^2),$$

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

$$P_n(w) = \sum_{\nu=0}^{n} d_\nu \phi_\nu(w)$$

so that

$$Q_n = \sum_{\nu=0}^{n} \mid d_\nu \mid^2.$$

THEOREM 1(3).

(2.1) $$\mu_n = \beta'(H_n')^{-1}\beta, \qquad n > N,$$

(2) $[\cdots]^*$ denotes the conjugate of $[\cdots]$.

(3) We thank the referee for suggesting the simple proof of Theorem 1 given above.

where β is a column vector with the N components $\beta_r = \beta_j^k$ and H_n is a nonsingular $N \times N$ matrix with elements

$$h_{r,s} = \sum_{\nu=0}^{n} \phi_\nu^{(k)}(\alpha_j)\phi_\nu^{(k')}(\alpha_{j'}),$$

$r \leftrightarrow (j, k)$, $s \leftrightarrow (j', k')$.

If $\log f(\lambda)$ *is integrable and all the restraints* (1.2) *are at points α_j inside the unit circle (we call such a set of restraints C_i), then*

$$\mu = \lim_{n\to\infty} \mu_n = \beta'(H')^{-1}\beta \tag{2.2}$$

where H is a nonsingular $N \times N$ matrix with elements

$$h_{r,s} = \frac{1}{2\pi}\left(\frac{\partial^k}{\partial x^k}\frac{\partial^{k'}}{\partial y^{*k'}}\frac{1}{1 - xy^*}\frac{1}{D(x)[D(y)]^*}\right), \qquad x = \alpha_j,\ y = \alpha_{j'}.$$

Proof. We have to minimize $Q_n = \sum_{\nu=0}^{n} |d_\nu|^2$, $n \geqq N$, with the restraint

$$\sum_{\nu=0}^{n} d_\nu \phi_\nu^{(k)}(\alpha_j) = \beta_j^k, \quad \text{or} \quad \sum_{\nu=0}^{n} d_\nu l_{\nu r} = \beta_r, \qquad r = 1, 2, \cdots, N.$$

We introduce the column vectors $d = \{d_\nu^*\}$, $l_r = \{l_{\nu r}\} = \{\phi_\nu^{(k)}(\alpha_j)\}$ in n-space so that the restraints have the form $d'l_r = \beta_r$ and $|d|^2$ has to be minimized The vectors l_r are linearly independent since

$$\sum t_j^k \phi_\nu^{(k)}(\alpha_j) = 0, \qquad \nu = 0, 1, \cdots, n,$$

would mean that

$$\sum_{j,k} t_j^k f^{(k)}(\alpha_j) = 0$$

for *any* polynomial $f(z)$ of degree n. By proper choice of f we find $t_j^k = 0$.

Now the vectors l_r' span a linear manifold

$$\sum_{\nu=1}^{N} \lambda_r l_r'$$

and the *projection* of d onto this furnishes the minimum of $|d|$. So assume d has the latter form. The restraints are now

$$\sum_{r=1}^{N} \lambda_r (l_s, l_r) = \beta_s, \qquad s = 1, 2, \cdots, N.$$

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

$$\left|\sum_{r=1}^{N} \lambda_r l_r'\right|^2 = \sum_{r,s} \lambda_r \bar{\lambda}_s (l_s, l_r) = \lambda' H_n \lambda = \beta'(H_n')^{-1} H_n H_n^{-1} \beta = \beta'(H_n')^{-1}\beta.$$

Substituting we obtain for the elements of H_n:

$$(l_r, l_s) = \sum_{\nu=0}^{n} \phi_\nu^{(k)}(\alpha_j)[\phi^{(k')}(\alpha_{j'})]^*.$$

But

$$\lim_{n\to\infty} \sum_{\nu=0}^{n} \phi_\nu(x)[\phi_\nu(y)]^* = \frac{1}{2\pi}\,\frac{1}{1 - xy^*}\,\frac{1}{D(x)[D(y)]^*}$$

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

$$\lim_{n\to\infty} \sum_{\nu=0}^{n} \phi_\nu^{(k)}(\alpha_j)[\phi_\nu^{(k')}(\alpha_{j'})]^*$$

$$= \frac{1}{2\pi}\left(\frac{\partial^k}{\partial x^k}\,\frac{\partial^{k'}}{\partial y^{*k'}}\,\frac{1}{1 - xy^*}\,\frac{1}{D(x)[D(y)]^*}\right), \quad x = \alpha_j,\ y = \alpha_{j'},$$

and we have (2.2).

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

THEOREM 2. *Let*

$$f(\lambda) = g(\lambda)\prod_{\nu=1}^{s} |e^{i\lambda} - e^{i\theta_\nu}|^{2l_\nu}, \qquad -\pi < \theta_\nu \leqq \pi, \tag{3.1}$$

where $g(\lambda)$ is positive and continuous and the l_ν are positive integers. Then

$$\begin{aligned}\mu_n = \frac{2\pi}{n^{2\rho+1}} \sum_{r_j+t_j=\rho} \binom{r_j + t_j}{r_j}^2 \cdot |\beta_j^{r_j}|^2 \cdot d(\rho, \nu_j)\cdot g(\lambda_j)\cdot \prod_{e^{i\theta_\nu}\neq\alpha_j} |\alpha_j - e^{i\theta_\nu}|^{2l_\nu} \\ + O\left(\frac{1}{n^{2\rho+2}}\right), \qquad \alpha_j = e^{i\lambda_j},\end{aligned} \tag{3.2}$$

where ρ, r_j, t_j, ν_j and $d(\rho, \nu_j)$ are defined below.

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

$$\frac{1}{2\pi}|a(e^{i\lambda})|^2 \leqq g(\lambda) \leqq \frac{1}{2\pi}|b(e^{i\lambda})|^2, \qquad |b(e^{i\lambda})| - |a(e^{i\lambda})| < \epsilon.$$

Let us now consider the case when $g(\lambda)$ is exactly equal to

$$\frac{1}{2\pi}\left| a(e^{i\lambda}) \right|^2 = \frac{1}{2\pi}\left| \sum_{\nu=0}^{p} a_\nu e^{i\nu\lambda} \right|^2.$$

Then we should minimize under conditions C_b

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} \left| P_n(e^{i\lambda}) p(e^{i\lambda}) \right|^2 d\lambda$$

where

$$p(w) = a(w) \prod_\nu (w - z_\nu)^{l_\nu}, \qquad z_\nu = e^{i\theta_\nu}.$$

Let $q(w) = P_n(w)p(w)$. 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

$$q^{(k)}(w_j) = \eta_j^k, \qquad k = 0, 1, \cdots, \nu_j, \; j = 1, 2, \cdots, M,$$

where w_j is one of the α_j's or z_ν'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)

$$\mu_n = \eta'(H_n')^{-1}\eta. \tag{3.3}$$

To compute H_n we observe that in the present case $f(\lambda) = 1/2\pi$ so that $\phi_\nu(w) = w^\nu$. Hence if $|x| = |y| = 1$,

$$\sum_{\nu=0}^{n} \phi_\nu^{(k)}(x)[\phi_\nu^{(k')}(y)]^*$$

$$= x^{-k}y^{k'} \sum_{\nu=0}^{n} \nu(\nu-1)\cdots(\nu-k+1)\nu(\nu-1)\cdots(\nu-k'+1)x^\nu y^{-\nu}.$$

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

$$H_n = \mathcal{D}_n \Lambda_n \mathcal{D}_n \tag{3.4}$$

where

$$\mathcal{D}_n = \begin{Bmatrix} n^{1/2} & & & & & & & 0 \\ & n^{1+1/2} & & & & & & \\ & & \ddots & & & & & \\ & & & n^{\nu_1+1/2} & & & & \\ & & & & n^{1/2} & & & \\ 0 & & & & & \ddots & & \\ & & & & & & & n^{\nu_M+1/2} \end{Bmatrix} \tag{3.5}$$

and

$$\Lambda_n = \Lambda + \frac{1}{n} R_n$$

where R_n has bounded elements. Here we have put

$$\Lambda = \begin{Bmatrix} \gamma_1 & & & 0 \\ & \gamma_2 & & \\ & & \ddots & \\ 0 & & & \gamma_M \end{Bmatrix}$$

where

$$\gamma_j = \left\{\frac{w_j^{-\nu+\mu}}{\nu+\mu+1}; \nu, \mu = 0, 1, \cdots, \nu_j\right\}.$$

Now note that

$$q^{(k)}(w) = \sum_{s=0}^{k} \binom{k}{s} P_n^{(s)}(w) p^{(k-s)}(w).$$

If z_ν does not coincide with any α_j, the conditions induced at z_ν are

$$q^{(k)}(z_\nu) = 0, \qquad k = 0, \cdots, l_\nu - 1.$$

If α_j does not coincide with any z_ν, the condition at α_j become

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

and if $\beta_j^{r_j}$ is the first nonzero β corresponding to α_j

$$q^{(r_j)}(\alpha_j) = \beta_j^{r_j} a(\alpha_j) \prod_\nu (\alpha_j - z_\nu)^{l_\nu}$$

and

$$q^{(k)}(\alpha_j) = 0, \qquad k < r_j.$$

Now let $\alpha_j = z_\nu$. Let r_j be defined as above. Then

$$q^{(k)}(z_\nu) = 0, \qquad k < r_j + l_\nu,$$

$$q^{(r_j+l_\nu)}(z_\nu) = \binom{r_j + l_\nu}{r_j} \beta_j^{r_j} a(z_\nu) \prod_{\mu \neq \nu} (z_\nu - z_\mu)^{l_\mu} .$$

Let t_j be the order of the zero of $q(w)$ at α_j. If there is no zero we set $t_j = 0$ and if $\alpha_j = z_\nu$, then $t_j = l_\nu$. Let the least k such that there is an $\eta_j^k \neq 0$ be called ρ. It follows from (3.3), (3.4), and (3.5) that the leading term of μ_n will consist only of contributions from the conditions with $k = \rho$. It is clear from the discussion above that we then get

$$\mu_n = \frac{1}{n^{2\rho+1}} \sum_{r_j + t_j = \rho} \binom{r_j + t_j}{r_j}^2 \cdot | \beta_j^{r_j} |^2 \cdot \left| a(\alpha_j) \prod_{z_\mu \neq \alpha_j} (\alpha_j - z_\mu)^{l_\mu} \right|^2 \cdot d(\rho, \nu_j)$$
$$+ O\left(\frac{1}{n^{2\rho+2}}\right)$$

where

$$d(\rho, n) = \left\{ \frac{1}{\nu + \mu + 1} ;\ \nu, \mu = 0, \cdots, n \right\}^{-1}_{\rho,\rho}$$
$$= (2\rho + 1) \prod_{j=0, j\neq \rho}^{n} \left(\frac{j + \rho + 1}{j - \rho}\right)^2$$

(see [1, p. 177]).

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

REMARK 1. If none of the zeros of $f(\lambda)$ coincide with any α_j, then ρ is simply the smallest k such that there is a nonzero β_j^k. We then have

$$\mu_n = \frac{2\pi}{n^{2\rho+1}} \sum_j | \beta_j^\rho |^2 \cdot f\left(\frac{\log \alpha_j}{i}\right) \cdot d(\rho, n_j) + O\left(\frac{1}{n^{2\rho+2}}\right).$$

REMARK 2. The case when there are conditions both inside and on the unit circle is now easily handled. If some β in C_i 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 β's in C_i vanish, $\mu_n \to 0$ and we get its principal term from (3.2) as if C_b were the only conditions.

Theorem 2 suggests that the error μ_n' computed with the minimizing polynomial corresponding to a uniform weight function is of the same order as the error μ_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 α_j of the conditions C_b. Let*

$$\mu_n' = \int_{-\pi}^{\pi} | P_n(e^{i\lambda}) |^2 f(\lambda) d\lambda$$

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

$$\lim_{n\to\infty} \frac{\mu_n'}{\mu_n} = 1.$$

Proof. Let $P_n(w) = \sum_0^n \gamma_\nu w^\nu$ be the minimizing polynomial under conditions C_b and the assumption that the spectral density is uniform ($f(\lambda)=1$). Let

$$\frac{d^k}{dw^k} w^\nu = \psi_{\nu,k}(w).$$

Then

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} | P_n(e^{i\lambda}) |^2 d\lambda = \sum_0^n | \gamma_\nu |^2$$

where

$$\gamma_\nu = \sum_{j,k} \lambda_j^k [\psi_{\nu,k}(\alpha_j)]^*$$

or

$$\gamma = \psi\lambda$$

where

$$\psi = \begin{Bmatrix} [\psi_{0,0}(\alpha_1)]^* & [\psi_{0,1}(\alpha_1)]^* \cdots & [\psi_{0,n_1}(\alpha_1)]^* & [\psi_{0,0}(\alpha_2)]^* \cdots \\ [\psi_{1,0}(\alpha_1)]^* & [\psi_{1,1}(\alpha_1)]^* \cdots & [\psi_{1,n_1}(\alpha_1)]^* & [\psi_{1,0}(\alpha_2)]^* \cdots \\ \cdots & \cdots & \cdots & \cdots \\ [\psi_{n,0}(\alpha_1)]^* & [\psi_{n,1}(\alpha_1)]^* \cdots & [\psi_{n,n_1}(\alpha_1)]^* & [\psi_{n,0}(\alpha_2)]^* \cdots \end{Bmatrix}$$

and $\lambda = H_n^{-1}\beta$. As before we factor H_n so that $H_n = \mathfrak{D}_n\Lambda_n\mathfrak{D}_n$. Then

$$\begin{aligned} \mu_n' &= \sum \gamma_\nu m_{\nu,\mu}^{(n)} \gamma_\mu^* = \gamma' M^{(n)} \gamma = \lambda'\psi' M^{(n)} \psi\lambda = \beta'(H_n')^{-1}\psi' M^{(n)} \psi H_n^{-1}\beta \\ &= \beta'\mathfrak{D}_n'^{-1}(\Lambda_n')^{-1}\mathfrak{D}_n'^{-1}\psi' M^{(n)} \mathfrak{D}_n^{-1}\Lambda_n^{-1}\mathfrak{D}_n^{-1}\beta. \end{aligned}$$

But

$$\mathcal{D}_n^{-1}\psi = \left\{\frac{[\psi_{\nu,k}(\alpha_j)]^*}{n^{k+1/2}}\right\}.$$

Hence a typical element of $\mathcal{D}_n^{-1}\psi' M^{(n)}\psi\mathcal{D}_n^{-1}$ is

$$\sum_{\nu,\mu=0}^{n}\frac{\psi_{\nu,k}(\alpha_j)}{n^{k+1/2}}\, m_{\nu,\mu}^{(n)}\,\frac{[\psi_{\mu,k'}(\alpha_{j'})]^*}{n^{k'+1/2}}$$

$$(3.6)\qquad = \frac{1}{n^{k+k'+1}}\int_{-\pi}^{\pi}\sum_{\nu,\mu=0}^{n} e^{i(\nu-\mu)\lambda}\nu(\nu-1)\cdots(\nu-k+1)\mu(\mu-1)\cdots (\mu-k'+1)\alpha_j^{\nu-k}\,\alpha_{j'}^{*\mu-k'} f(\lambda)d\lambda.$$

Let $\Lambda=\lim_{n\to\infty}\Lambda_n$ as before. We shall show that

$$(3.7)\qquad \mathcal{D}_n^{-1}\psi' M^{(n)}\psi\mathcal{D}_n^{-1} = f\Lambda + O(1)$$

where

$$f = \begin{pmatrix} \overbrace{f(\lambda_1) \quad \cdot \quad \cdot \quad \cdot \quad f(\lambda_1)}^{n_1} & & 0 \\ & \overbrace{f(\lambda_2) \quad \cdot \quad \cdot \quad \cdot \quad f(\lambda_2)}^{n_2} & \\ 0 & & \ddots \end{pmatrix}, \qquad \alpha_j = e^{i\lambda_j}.$$

First suppose $\alpha_j \neq \alpha_{j'}$. Then (3.6) can be rewritten as

$$\alpha_j^{-k}[\alpha_{j'}]^{*-k'}\int_{-\pi}^{\pi} e^{i(k'-k)\lambda}\eta_{n,k}(e^{i(\lambda-\lambda_j)})\,[\eta_{n,k'}(e^{i(\lambda-\lambda_{j'})})]^* f(\lambda)d\lambda$$

where

$$\eta_{n,k}(x) = \frac{1}{n^{k+1/2}}\,\frac{d^k}{dw^k}\left(\frac{1-w^{n+1}}{1-w}\right) \qquad w = x.$$

If $|\lambda-\lambda_j| \geqq \epsilon$,

$$|\eta_{n,k}(e^{i(\lambda-\lambda_j)})| < \frac{c(\epsilon)}{n^{1/2}}$$

while if $|\lambda-\lambda_j| < \epsilon$

$$|\eta_{n,k}(e^{i(\lambda-\lambda_j)})| \leqq n^{1/2}.$$

Hence if $\lambda_j \neq \lambda_{j'}$ expression (3.6) converges to zero. Now consider $\alpha_j = \alpha_{j'} = e^{i\lambda_j}$. Then (3.6) can be rewritten as

$$1/n^{(k+k'+1)}e^{i(k'-k)\lambda_j}\int_{-\pi}^{\pi}\sum_{\nu,\mu=0}^{n}\nu(\nu-1)\cdots(\nu-k+1)\mu(\mu-1)\cdots (\mu-k'+1)e^{i(\nu-\mu)(\lambda-\lambda_j)}f(\lambda)d\lambda.$$

Let

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

Consider

$$K_n(\lambda) = \sum_{l,k=1}^{n_j} y_l s_l(\lambda)[y_k s_k(\lambda)]^* = |\sum y_k s_k(\lambda)|^2 \geqq 0.$$

One can verify that

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

Hence (see [5, p. 49])

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

But (3.7) then follows immediately. Now

$$\begin{aligned}\mu_n' &= \beta'\mathcal{D}_n^{-1}((\Lambda')^{-1} + o(1))(f\Lambda + o(1))(\Lambda^{-1} + o(1))\mathcal{D}_n^{-1}\beta \\ &= \beta'\mathcal{D}_n^{-1}((\Lambda')^{-1} + o(1))f\mathcal{D}_n^{-1}\beta.\end{aligned}$$

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 δ_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 λ and* (2) *$g(\lambda)$ has no zeros.*

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

$$\phi(w) = f(\lambda), \qquad w = e^{i\lambda}, \qquad -\infty < \lambda < \infty,$$

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

$$\log \phi(w) = \sum_{\nu=-\infty}^{\infty} \gamma_\nu w^\nu.$$

We then note that

$$D(w) = \exp \left\{ \frac{\gamma_0}{2} + \sum_1^\infty \gamma_\nu^* w^\nu \right\}$$

and hence $D(w)$ is analytic in the closed region $|w| \leqq 1$ 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

$$\begin{aligned} \mu_n - \mu &\leqq \|s_n(w)\|^2 - \|D(0)/D(w)\|^2 \\ &= (\|s_n(w)\| + \|D(0)/D(w)\|)(\|s_n(w)\| - \|D(0)/D(w)\|) \\ &\leqq K_1 \left\| s_n(w) - \frac{D(0)}{D(w)} \right\| \leqq K_2 \left(\sum_{n+1}^\infty |d_\nu|^2 \right)^{1/2} \end{aligned}$$

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

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

$$|\phi_\nu(0)| < K_3 d^{n/2}.$$

However, $|\phi_\nu(w)| < K_4 |w|^\nu$ on $|w| = 1 + \epsilon$, $\epsilon > 0$ (see [6, Satz XXXII]). If $1 + \epsilon < 1/d^{1/2}$ we have uniform convergence of

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

so that $2\pi \sum_0^n [\phi_\nu(0)]^* \phi_\nu(w)$ 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

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

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 δ_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*

$$g(\lambda) \prod_{\nu=1}^{s} \left| e^{i\lambda} - e^{i\lambda_\nu} \right|^{2l_\nu},$$

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

$$\mu_n = \beta' (H'_{n+M})^{-1} \beta, \qquad M = \sum l_\nu,$$

where β has its first component equal to one and the remaining components are zero and

$$H_{n+M} = \mathcal{D}_n \left\{ \begin{pmatrix} \Lambda & & & & 0 \\ & \gamma_1 & & & \\ & & \cdot & & \\ & & & \cdot & \\ 0 & & & & \gamma_p \end{pmatrix} + \frac{1}{n} B_n \right\} \mathcal{D}_n$$

where B_n is bounded and

$$\mathcal{D}_n = \begin{Bmatrix} 1 & & & & & & 0 \\ & n^{1/2} & & & & & \\ & & \cdot & & & & \\ & & & \cdot & & & \\ & & & & n^{l_1+1/2} & & \\ & & & & & n^{1/2} & \\ 0 & & & & & & \cdot \\ & & & & & & & \cdot \end{Bmatrix}.$$

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 $m_t = Ex_t$. 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

$$\rho_{s,t} = Ey_s y_t^* = r_{s-t}.$$

Then we know that

$$r_\nu = \int_{-\pi}^{\pi} e^{i\nu\lambda} dF(\lambda)$$

where $F(\lambda)$ is bounded and nondecreasing in $(-\pi, \pi)$. The completely nondeterministic processes form an important subclass. They are completely characterized by

(1) $F(\lambda)$ absolutely continuous,

$$F(\lambda) = \int_{-\pi}^{\lambda} f(l) dl,$$

and

$$\int_{-\pi}^{\pi} \log f(\lambda) d\lambda > -\infty. \tag{2}$$

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 $\sum_0^n c_\nu x_\nu$ subject to some conditions on the c_ν's. But this variance is

$$\sum_{\nu,\mu=0}^{n} c_\nu c_\mu^* r_{\nu-\mu}$$

which is of the form (1.1). We shall consider some problems of this type.

1. Let us first assume $m_t \equiv 0$. Having observed $x_1, x_2, \cdots, x_n$ we want to form a linear combination $\sum_1^n c_\nu x_\nu$ such that

$$E\left| x_0 - \sum_1^n c_\nu x_\nu \right|^2 = \min.$$

This is a familiar problem of extrapolation. This is the type of problem treated in the previous sections since we can write it as

$$\mu_n = \int_{-\pi}^{\pi} | P_n(e^{i\lambda}) |^2 f(\lambda) d\lambda = \min,$$

$$P_n(0) = 1.$$

The only restraint is inside the unit circle and we then know that μ_n tends to a positive value μ as $n \to \infty$. From Theorem 1 we get

$$\mu = 2\pi | D(0) |^2 = 2\pi \exp \left\{ \frac{1}{2\pi} \int_{-\pi}^{\pi} \log f(\lambda) d\lambda \right\}. \tag{5.1}$$

The "predictor" is

$$\sum_{\nu=1}^{n} c_\nu x_\nu = \int_{-\pi}^{\pi} \sum_{1}^{n} c_\nu e^{i\nu\lambda} dZ(\lambda)$$

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

$$\int_{-\pi}^{\pi} \left[1 - \frac{D(0)}{D(e^{i\lambda})}\right] dZ(\lambda). \tag{5.2}$$

(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 $x_k, x_{k+1}, \cdots, x_n$ to predict x_0. We see that this corresponds to the conditions

$$\begin{aligned} P(0) &= 1, \\ P'(0) &= 0, \\ \cdot\ \cdot\ \cdot\ &\cdot\ \cdot\ \cdot , \\ P^{(k-1)}(0) &= 0. \end{aligned}$$

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

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

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

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

4. Let $m_t = mt(t-1) \cdots (t-k+1)e^{it\lambda_0}$. 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

$$E|m^* - m|^2 \sim \frac{2\pi f(e^{i\lambda_0})}{n^{2k+1}} (2k+1).$$

We can also apply Theorem 3.

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

$$\phi_t^{(k,j)} = t(t-1)\cdots(t-k+1)e^{it\lambda_j},$$

$$m_t = \sum_{k,j} c_{k,j}\phi_t^{(k,j)}$$

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

$$P^{(l)}(e^{i\lambda_s}) = 0 \qquad \text{if } (l, s) \neq (k, j),$$

$$P^{(k)}(e^{i\lambda_j}) = e^{-ik\lambda_j}.$$

All the conditions are on the unit circle.

6. If $m_t \equiv m$ is unknown and we wish to predict the value of x_0 from $x_k, x_{k+1}, \cdots, x_n$ it may be advantageous to use an unbiased predictor

$$\sum_{\nu=k}^{n} c_\nu x_\nu, \qquad \sum_{\nu=k}^{n} c_\nu = 1.$$

We then get the conditions

$$\begin{aligned} P(0) &= 1, \\ P'(0) &= 0, \\ \cdots&\cdots, \\ P^{(k-1)}(0) &= 0, \\ P(1) &= 0. \end{aligned}$$

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

References

1. A. Cauchy, *Oeuvres complètes*, 2d series, vol. 12, Paris, 1916.

2. H. Cramer, *On harmonic analysis in certain functional spaces*, Arkiv för Matemtatik, Astronomi och Fysik (1942).

3. U. Grenander, *On Toeplitz forms and stationary processes*, Arkiv för Matematik, Astronomi och Fysik vol. 1 (1951).

4. ———, *On the estimation of regression coefficients in the case of an uncorrelated disturbance*, unpublished.

5. E. Hille, *Functional analysis and semi-groups*, New York, 1948.

6. G. Szegö, *Beiträge zur Theorie der Toeplitzschen Formen*, Math. Zeit. vol. 6 (1919) pp. 167–202; vol. 9 (1921) pp. 167–190.

7. N. Wiener, *Extrapolation, interpolation and smoothing of stationary time series*, New York, 1949.

University of Chicago,
Chicago, Ill.