SOME PROBLEMS IN ESTIMATING THE SPECTRUM OF A TIME SERIES

ULF GRENANDER
UNIVERSITY OF STOCKHOLM
AND
MURRAY ROSENBLATT*
THE UNIVERSITY OF CHICAGO

1. Introduction

We consider a probability model that has proved to be useful in various applied fields. Statistical problems that arise in the analysis of data obtained in these fields are discussed. The basic model is that of a stochastic process $\{y_t\}$

$$y_t = x_t + m_t \tag{1.1}$$

where $m_t = Ey_t$ is the mean value of y_t and x_t, $Ex_t \equiv 0$, is the residual. The residual x_t is assumed to be stationary with respect to the parameter t, that is,

$$x_{t_1}, \cdots, x_{t_n} \tag{1.2}$$

have the same probability distribution as

$$x_{t_1+h}, \cdots, x_{t_n+h} \tag{1.3}$$

for all possible values of $t_1, \cdots, t_n, h$. In other words, the probability distribution of x_t is invariant under t displacement. This implies that the set of possible values of t, which we shall call T, is a group or semigroup under addition. Typical examples of the parameter set T are the set of all points in Euclidean k-space or the set of lattice points in Euclidean k-space. These are in fact the examples of greatest interest and they will be discussed in some detail.

The process $\{y_t\}$ may be vector valued. An example of interest in which the vector-valued case is appropriate will be described.

A usual situation is that in which t is thought of as time. The parameter t will then be a point of the form kh if the observations are taken at discrete time points with h seconds between each observation. If the observation is continuous, t will be any real number.

The case in which $m_t \equiv 0$ is of considerable importance. Such a model is appropriate where the phenomenon studied consists of random fluctuations which are of a stable character.

Some of the fields in which such a model has been used will be discussed in section 2. These fields are in the physical sciences. They are discussed to give some motivation to

Research carried out in part at the Statistical Research Center, University of Chicago, under sponsorship of the Statistics Branch, Office of Naval Research.

* Now at Indiana University.

the development and because the methods considered have been successful in some degree in their application. Before going on to consider the statistical questions of interest, we shall develop some of the probability theory required. We want to give a survey of some of the results that have been obtained in the past five years and indicate that there are many problems of interest that still remain open. We shall therefore sketch out some of the older results and try to motivate them. Proofs will be given only when they are relevant to the discussion or when they relate to new results.

The paper discusses a spectral representation of the residual $\{x_t\}$. We are basically concerned with the statistical problem of estimating the spectrum from a time series or partial realization of the process. The estimation of the spectrum is of great interest for various reasons. First of all, knowledge of the spectrum gives us information about the structure of the process. Knowledge of the spectrum is also essential in the linear problems of prediction, interpolation and filtering.

2. Some fields of application

One of the earliest fields of application is in the study of random noise. In electrical or electronic circuits noise of random character sometimes arises. This may be due to the random drift of electrons in the circuit or perhaps from shot noise due to tubes in the circuit. Another example is clutter on a radar scope or a television screen due to reflection from surrounding buildings or snow. In the first example t is time and y_t is a number. The residual x_t can be thought of as the noise which masks the message m_t which one would want to estimate. A detailed discussion of some of the problems that arise in this context can be found in Lawson and Uhlenbeck [10].

Another field of application is in the study of turbulence. Consider a fluid forced through a rectangular grid. Let the kinematic viscosity of the fluid be small. The velocity field $v(\tau, u)$ of the fluid behind the grid seems to be random. Here τ is the time and u is the point at which the observation is taken. The velocity field of the fluid fluctuates even though the macroscopic conditions are the same. It then seems reasonable to consider the velocity field as a stochastic process. Assume that energy is fed in at the same rate at which it is dissipated. At an intermediate distance from the grid the distribution of the velocity field appears to be invariant under space displacement. Homogeneous turbulence is an idealization in which one imagines all of space filled with a turbulent fluid whose velocity field has a probability distribution invariant under space displacement. Let us also assume that the turbulence is also stationary in time since the energy fed in balances out that dissipated. Here $t = (\tau, u)$ is 4-dimensional while $x_t = v(\tau, u)$ is 3-dimensional. One then looks for a stationary process satisfying the equations of motion, that is, the continuity equation (assuming incompressibility) and the Navier-Stokes equation. A detailed discussion of homogeneous turbulence can be found in Batchelor's book [3].

Various meteorologists are now studying the atmosphere, considering it as a turbulent fluid. The assumption of local homogeneity and stationarity which they make may not be a bad one (see Panofsky [12]).

Still another field is the study of storm generated ocean waves. Consider a storm on the ocean surface. Let $h(\tau, u)$ be the vertical displacement of the ocean surface at time τ and position u with respect to the undisturbed surface. Well within the storm area $h(\tau, u)$ can be considered stationary with respect to displacement in u along the sea surface. If the period of observation is small with respect to the duration of the storm, the disturbance can be considered stationary with respect to time. Here $t = (\tau, u)$ is 3-di-

mensional and $x_t = h(\tau, u)$ is 1-dimensional. See Pierson [13] for a discussion of this application.

3. Spectral representation of the process

Assume that $m_t \equiv 0$. We shall now consider a basic representation of the process x_t. In effect, we are carrying out a random Fourier analysis of the process.

It is reasonable to assume that the process is real valued or has real-valued components. The covariance matrices[1]

$$R_{t,\tau} = R_{t-\tau} = Ex_t x_\tau' \tag{3.1}$$

depend only on the difference $t - \tau$ of the parameters because of the stationarity. Consider the case in which t is integral and x_t possibly vector valued. The results cited in this section are still valid with appropriate modification if t is continuous or vector valued. If x_t is k-dimensional and $\{c_t\}$ is any finite sequence of k-vectors

$$\sum_{t,\tau} c_t' R_{t-\tau} c_\tau \geqq 0\,, \tag{3.2}$$

since $\{R_t\}$ is a sequence of covariance matrices. The process x_t has the Fourier representation

$$x_t = \int_{-\pi}^{\pi} e^{it\lambda} dZ(\lambda) \tag{3.3}$$

where $Z(\lambda)$ is an orthogonal process, that is,

$$\begin{aligned} E[dZ(\lambda)] &\equiv 0 \\ E[dZ(\lambda)\, dZ(\mu)'] &= \delta_{\lambda\mu} dF(\lambda) \end{aligned} \tag{3.4}$$

where $\delta_{\lambda\mu}$ is the Kronecker delta and $F(\lambda)$ is a nondecreasing matrix-valued function (see Cramér [4]). The differential notation occasionally used in the discussion is to be understood in the usual way. Here $dF(\lambda)$ denotes the increment of the function $F(\lambda)$ over a small λ-interval. In (3.3) x_t is expressed as a superposition of harmonics $\exp(i\lambda)$ with corresponding random amplitudes $dZ(\lambda)$. The random amplitudes $dZ(\lambda)$ are orthogonal and the covariance matrix of the random vector weight $dZ(\lambda)$ is $dF(\lambda) \geqq 0$. The process $Z(\lambda)$ with orthogonal increments is given by

$$Z(\lambda) = \frac{1}{2\pi} \sum_{t \neq 0} x_t \frac{e^{-it\lambda} - e^{it\pi}}{-it} + \frac{1}{2\pi}(\lambda + \pi)\,, \tag{3.5}$$

that is, $Z(\lambda)$ is the integral of the formal Fourier series with Fourier coefficients x_t. The process $Z(\lambda)$ is introduced because the formal Fourier series referred to does not exist. The function $F(\lambda)$ is usually called the *spectral distribution function* of the process x_t. The representation (3.3) of x_t implies that

$$R_t = \int_{-\pi}^{\pi} e^{it\lambda} dF(\lambda)\,. \tag{3.6}$$

From this one can see that knowledge of the covariance sequence R_t and knowledge of the spectral distribution function $F(\lambda)$ are equivalent. Nonetheless, in many fields, especially those referred to in section 2, it is much more natural to think of the process in

[1] x_t is a column vector. If A is a matrix, A' denotes the conjugated transpose of A.

terms of its spectrum rather than its covariance sequence. Moreover, the statistical problems that arise are answered much more naturally and elegantly when looked at from the point of view of the spectrum of the process rather than the covariance sequence. Note that $dF(\lambda) = \overline{dF(-\lambda)}$ since the process x_t has real-valued components.

In the representation (3.3) of the process x_t, $Z(\lambda)$ has complex-valued components in general. Because of this unpleasantness it is convenient to introduce the auxiliary processes

$$\begin{aligned} Z_1(\lambda) &= Z(\lambda) - Z(-\lambda) \\ Z_2(\lambda) &= i[Z(\lambda) + Z(-\lambda) - 2Z(0)] \end{aligned} \tag{3.7}$$

with real-valued components. We then have the following real representation for x_t

$$x_t = \int_0^\pi \cos t\lambda \, dZ_1(\lambda) + \int_0^\pi \sin t\lambda \, dZ_2(\lambda). \tag{3.8}$$

The case of greatest practical interest is that in which $F(\lambda)$ has no singular part, that is, $F(\lambda)$ has an absolutely continuous part and jumps. In the remainder of this paper we shall assume that $F(\lambda)$ is absolutely continuous, that is,

$$F(\lambda) = \int_{-\pi}^{\lambda} f(u) \, du. \tag{3.9}$$

The matrix-valued function $f(\lambda)$ is called the *spectral density* of the process. Let us consider the case of a two-dimensional process $x_t = \begin{pmatrix} {}_1x_t \\ {}_2x_t \end{pmatrix}$ in some detail.[2] Then

$$\frac{dF(\lambda)}{d\lambda} = f(\lambda) = \begin{pmatrix} f_{11}(\lambda) & f_{12}(\lambda) \\ f_{21}(\lambda) & f_{22}(\lambda) \end{pmatrix} \geqq 0. \tag{3.10}$$

The functions $f_{11}(\lambda), f_{22}(\lambda) \geqq 0$ are the spectral densities of ${}_1x_t$, ${}_2x_t$, respectively, while $f_{12}(\lambda)$ is the cross-spectral density of ${}_1x_t$ and ${}_2x_t$. Now $f_{11}(\lambda) = f_{11}(-\lambda), f_{22}(\lambda) = f_{22}(-\lambda)$, and $f_{12}(\lambda) = f_{21}(-\lambda) = \overline{f_{21}(\lambda)}$ since ${}_1x_t$ and ${}_2x_t$ are real valued. In the real representation (3.7) of x_t let

$$Z_1(\lambda) = \begin{pmatrix} Z_{11}(\lambda) \\ Z_{21}(\lambda) \end{pmatrix}, \qquad Z_2(\lambda) = \begin{pmatrix} Z_{12}(\lambda) \\ Z_{22}(\lambda) \end{pmatrix}. \tag{3.11}$$

Then

$$\begin{cases} E\, dZ_{11}(\lambda)\, dZ_{11}(\mu) = E\, dZ_{12}(\lambda)\, dZ_{12}(\mu) = 2\delta_{\lambda\mu} f_{11}(\lambda)\, d\lambda \\ E\, dZ_{21}(\lambda)\, dZ_{21}(\mu) = E\, dZ_{22}(\lambda)\, dZ_{22}(\mu) = 2\delta_{\lambda\mu} f_{22}(\lambda)\, d\lambda \\ E\, dZ_{11}(\lambda)\, dZ_{12}(\mu) = E\, dZ_{21}(\lambda)\, dZ_{22}(\mu) = 0 \\ E\, dZ_{11}(\lambda)\, dZ_{21}(\mu) = E\, dZ_{12}(\lambda)\, dZ_{22}(\mu) = 2\delta_{\lambda\mu} \operatorname{Re} f_{12}(\lambda)\, d\lambda \\ E\, dZ_{11}(\lambda)\, dZ_{22}(\mu) = -E\, dZ_{21}(\lambda)\, dZ_{12}(\mu) = 2\delta_{\lambda\mu} \operatorname{Im} f_{12}(\lambda)\, d\lambda. \end{cases} \tag{3.12}$$

The real part of the cross-spectral density $\operatorname{Re} f_{12}(\lambda)$ is often called the cospectrum of ${}_1x_t$ and ${}_2x_t$ while $\operatorname{Im} f_{12}(\lambda)$ is called the quadrature spectrum of ${}_1x_t$ and ${}_2x_t$. The cospectrum measures the dependence of the in-phase harmonics of the two processes ${}_1x_t$

[2] The discussion in the remainder of this section and in section 8 was suggested by conversations one of the authors had with W. J. Pierson, Jr. of the Department of Meteorology and L. J. Tick of the Research Division of N.Y.U.

and ${}_2x_t$, that is, the dependence between $\cos t\lambda\, dZ_{11}(\lambda)$ and $\cos t\lambda\, dZ_{21}(\lambda)$ or $\sin t\lambda\, dZ_{12}(\lambda)$ and $\sin t\lambda\, dZ_{22}(\lambda)$. The quadrature spectrum measures the dependence of the out-of-phase components of the two processes $\cos t\lambda\, dZ_{11}(\lambda)$ and $\sin t\lambda\, dZ_{22}(\lambda)$.

If the class of admissible distributions in a nonparametric problem is parametrized or labeled in a natural way, one is led to an infinite-dimensional parameter space. It seems natural to think of a statistical problem characterized by such an infinite-dimensional parameter space as a nonparametric problem. It would be natural to parametrize the class of processes we deal with by their covariance sequences. Since this parameter space is infinite-dimensional, the techniques we employ in the statistical analysis of time series would be nonparametric techniques in the sense described above.

Note that the normal processes are determined by their spectra since they are determined by their first and second moments. The representation (3.3) of the process x_t is a linear representation and thus is especially natural in the case of a normal process. Many of the statistical techniques employed are linear techniques since they are based on this linear representation. Nonetheless, they are still quite useful in obtaining information about the linear structure of nonnormal processes.

If the time parameter t is continuous, the range of the λ integration in (3.3) is from $-\infty$ to ∞. If the parameter t ranges over the lattice points in k-dimensional Euclidean space, the analogue of representation (3.3) is

$$x_t = \underbrace{\int_{-\pi}^{\pi} \cdots \int}_{k} e^{it\cdot\lambda} dZ(\lambda) \tag{3.13}$$

and the spectral distribution function $F(\lambda)$ is a function of the k-vector λ. When t is a continuous parameter, the λ integration in (3.12) ranges over all k-space.

The representations of the process x_t given here are valid if the weaker assumption of weak stationarity of x_t is made (see Doob [5]). Strong stationarity of x_t has been assumed because it is really made use of later on.

4. Moving averages and linear processes

We shall now motivate an assumption on the distribution of the stationary processes x_t dealt with. Consider the representation

$$x_t = \int_{-\pi}^{\pi} e^{it\lambda} dZ(\lambda) \tag{4.1}$$

$$E\, dZ(\lambda)\, dZ(\mu)' = \delta_{\lambda\mu} f(\lambda)\, d\lambda .$$

Note that $f(-\lambda) = f(\lambda)$. Assume that $f(\lambda) \geqq 0$ is a nonsingular matrix for almost all λ. We can then write

$$f(\lambda) = \frac{1}{2\pi} a(\lambda)\, a(\lambda)', \qquad a(-\lambda) = \overline{a(\lambda)} \tag{4.2}$$

where $a(\lambda)$ is nonsingular almost everywhere. Let

$$\xi_t = \int_{-\pi}^{\pi} e^{it\lambda}\, a^{-1}(\lambda)\, dZ(\lambda) . \tag{4.3}$$

Then

$$E\xi_t \equiv 0 \tag{4.4}$$

and

$$E\xi_t\xi_\tau' = \int_{-\pi}^{\pi} e^{i(t-\tau)\lambda}\, a^{-1}(\lambda) E[dZ(\lambda)\, dZ(\lambda)'][a^{-1}(\lambda)]' \tag{4.5}$$

$$= \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{i(t-\tau)\lambda} I\, d\lambda = I\,\delta_{t,\tau}.$$

The process $\{x_t\}$ can be expressed in terms of the process $\{\xi_t\}$,

$$x_t = \int_{-\pi}^{\pi} e^{i\lambda}\, a(\lambda)[a^{-1}(\lambda)]\, dZ(\lambda) \tag{4.6}$$

$$= \sum a_j \int_{-\pi}^{\pi} e^{i(t-j)\lambda} a^{-1}(\lambda)\, dZ(\lambda) = \sum_{j=-\infty}^{\infty} a_j \xi_{t-j},$$

where $a_j = (1/2\pi)\int_{-\pi}^{\pi} \exp(ij\lambda) a(\lambda)\, d\lambda$. This representation is *a moving average representation* of the process x_t. The assumption that $f(\lambda)$ be nonsingular for almost all λ is not really necessary. It was assumed so that a simple and intuitively appealing derivation of the moving average representation could be used.

We shall assume that the processes x_t we deal with have such a moving average representation where the ξ_t's are independent vectors with independent components. Such a process is called a *linear process*. We shall also assume that $f(\lambda)$ is continuous and nonsingular for all λ. These assumptions are made so that the process x_t will satisfy the central limit theorem, that is,

$$\frac{1}{\sqrt{n}}\sum_{t=1}^{n} x_t \tag{4.7}$$

is asymptotically normally distributed (see Moran [11]). Some of the results cited will depend on this asymptotic normality. Some such restriction has to be made since not all stationary processes do satisfy the central limit theorem. The linear processes are a fairly broad class of processes including most of those processes one usually encounters, for example, the normal processes, and the autoregressive schemes and moving averages of finite order with independent residuals. The results cited probably hold for almost all stationary processes satisfying the central limit theorem. However, most of the theorems on the central limit theorem for dependent processes are not very clear or well adapted for use in this context. It would, in fact, be of some interest to get a result indicating the class of stationary processes that satisfy the central limit theorem that is useful and meaningful in statistical work. Nonetheless, it is clear that any physicist would believe that a process whose correlation time is small compared to the time interval of observation satisfies the central limit theorem.

5. Estimation of the spectral density

First consider the problem of estimating the spectral density $f(\lambda)$ when x_t and t are 1-dimensional. Let t be integral and assume that one has observed the time series $x_1, x_2, \cdots, x_N$. The statistics

$$r_\nu^* = \frac{C_\nu}{N} = \frac{1}{N}\sum_{t=1}^{N-\nu} x_t x_{t+\nu} \tag{5.1}$$

$$r_\nu^* = r_{-\nu}^*, \qquad 0 < \nu < N,$$

are good estimates of the covariances r_ν. Since $f(\lambda)$ is formally given by

$$f(\lambda) \sim \frac{1}{2\pi} \sum_{\nu=-\infty}^{\infty} r_\nu e^{-i\nu\lambda} \tag{5.2}$$

one might then be tempted to estimate $f(\lambda)$ by

$$I_N(\lambda) = \frac{1}{2\pi} \sum_{\nu=-N}^{N} r_\nu^* e^{-i\nu\lambda} = \frac{1}{2\pi N} \left| \sum_{\nu=1}^{N} x_\nu e^{-i\nu\lambda} \right|^2, \tag{5.3}$$

the periodogram. The periodogram is an asymptotically unbiased estimate

$$\lim_{N\to\infty} EI_N(\lambda) = f(\lambda) \tag{5.4}$$

but unfortunately it is not consistent. In fact, the variance

$$D^2[I_N(\lambda)] \to f^2(\lambda) > 0 \tag{5.5}$$

as $N \to \infty$. The periodogram $I_N(\lambda)$ is asymptotically normally distributed with mean $f(\lambda)$ and variance $f^2(\lambda)$. Moreover, $I_N(\lambda)$ and $I_N(\mu)$, $\lambda \neq \mu$, are asymptotically independent. This suggests that one might obtain a consistent estimate of $f(\lambda)$ by smoothing $I_N(\lambda)$ by the law of large numbers. It should be noted that this is a statistical context in which there are in general no unbiased estimates of $f(\lambda)$.

Let $w_N(u) \geqq 0$ be a nonnegative weight function of total mass 1,

$$\int_{-\pi}^{\pi} w_N(u)\, du = 1, \tag{5.6}$$

that is highly peaked in the neighborhood of $u = 0$. As $N \to \infty$, we want $w_N(u)$ to concentrate its mass more and more highly in the vicinity of $u = 0$. This will follow if for every $\epsilon > 0$,

$$w_N(u) \to 0 \tag{5.7}$$

uniformly for $|u| \geqq \epsilon$ as $N \to \infty$. Consider

$$\begin{aligned} f_N^*(\lambda) &= \int_{-\pi}^{\pi} w_N(u-\lambda)\, I_N(x)\, dx \\ &= \frac{1}{2\pi} \sum_{\nu=-N}^{N} r_\nu^* w_\nu^{(N)} e^{-i\nu\lambda} \end{aligned} \tag{5.8}$$

as an estimate of $f(\lambda)$. Here

$$w_\nu^{(N)} = \int_{-\pi}^{\pi} w_N(u)\, e^{i\nu u} dx. \tag{5.9}$$

Let x_t be a linear scheme

$$x_t = \sum_{\nu=-\infty}^{\infty} a_{t-\nu} \xi_\nu \tag{5.10}$$

with

$$\begin{aligned} &E\xi_t^2 = 1 \\ &E\xi_t^4 < \infty \end{aligned} \tag{5.11}$$

and

$$a_k = O\left(\frac{1}{|k|^{2+\delta}}\right), \qquad \delta > 0. \tag{5.12}$$

If x_t is a normal process, the condition on the sequence a_k will surely be satisfied if $f(\lambda)$ has a third-order derivative. Let $v_N(u) = w_N(u) * w_N(u)$ where the $*$ denotes convolution. We assume that

$$\left|\frac{v_N(u)}{v_N(0)} - 1\right| \to 0 \tag{5.13}$$

as $N \to \infty$ for $|u| \leqq A/N$ where A is any positive constant. One can then show that $f_N^*(\lambda)$ is asymptotically normally distributed with bias

$$b_N(\lambda) = Ef_N^*(\lambda) - f(\lambda) \sim \int_{-\pi}^{\pi} w_N(u)\,[f(u) - f(\lambda)]\,du + O\left(\frac{\log N}{N}\right) \tag{5.14}$$

and variance

$$D^2[f_N^*(\lambda)] \sim \frac{2\pi}{N}\int_{-\pi}^{\pi} w_N^2(u-\lambda)\,f^2(u)\,du \sim \frac{2\pi f^2(\lambda)}{N}\int_{-\pi}^{\pi} w_N^2(u)\,du. \tag{5.15}$$

The estimate $f_N^*(\lambda)$ is clearly consistent if

$$\int_{-\pi}^{\pi} w_N^2(u)\,du = o(N). \tag{5.16}$$

This result is also true for a variety of weight functions $w_N(u)$ that are not necessarily nonnegative everywhere (see Grenander and Rosenblatt [9]). The estimates $f_N^*(\lambda)$, $f_N^*(\mu)$, $\lambda, \mu > 0$, $\lambda \neq \mu$, of $f(\lambda)$ and $f(\mu)$, respectively, are asymptotically independent as $N \to \infty$. It is clear that there is an infinite class of weight functions $w_N(u)$ that lead to consistent estimates of the spectral density $f(\lambda)$. Bartlett [1], [2] and Tukey [14] appear to be the first statisticians who proposed specific estimates $f_N^*(\lambda)$ of $f(\lambda)$ of this type.

Typical choices of $w_N(u)$ are the Dirichlet kernel

$$w_N(u) = \frac{1}{2\pi}\frac{\sin\left(h_N + \frac{1}{2}\right)u}{\sin\frac{1}{2}u}, \qquad w_\nu^{(N)} = \begin{cases} 1, & \text{if } |\nu| \leqq h_N \\ 0, & \text{otherwise,} \end{cases} \tag{5.17}$$

and the Fejer kernel

$$w_N(u) = \frac{1}{2\pi h_N}\frac{\sin^2\frac{h_N}{2}u}{\sin^2\frac{1}{2}u}, \qquad w_\nu^{(N)} = \begin{cases} 1 - \frac{|\nu|}{h_N}, & \text{if } |\nu| \leqq h_N \\ 0, & \text{otherwise,} \end{cases} \tag{5.18}$$

where $h_N \to \infty$ as $N \to \infty$ and $h_N = o(N)$. In many cases one prefers an estimate with a nonnegative weight function like the Fejer kernel since such estimates are nonnegative with probability one. This seems plausible since $f(\lambda)$ is nonnegative. Note that the Dirichlet weight function is not nonnegative.

There is an "uncertainty principle" of sorts that holds here. As a typical example take the family of estimates based on the Fejer kernel. If the variance of the estimate is made

smaller, the bias of the estimate is increased and vice versa. It then seems reasonable to minimize the mean-square error

$$D^2[f_N^*(\lambda)] + b_N^2(\lambda) \tag{5.19}$$

of the estimate $f_N^*(\lambda)$ of $f(\lambda)$. The asymptotic mean-square error of a number of specific estimates has been examined in Grenander and Rosenblatt [9]. However, there is much work that still ought to be carried out in this direction. Each estimate has to be examined for a detailed evaluation of the mean-square error because the bias in general depends on the shape of the weight function $w_N(u)$.

Note that the specific estimates mentioned above do not make use of all the sample covariances r_ν^*, $\nu = 0, 1, \cdots, N$, but only of the sample covariances r_ν^*, $\nu = 0, 1, \cdots, h_N$. A typical choice of h_N is $h_N = KN^a$, $0 < a < 1$. In the case of the estimate using the Fejer kernel the best choice of a is $a = 1/3$. The mean-square error is then $O(N^{-2/3})$ as $N \to \infty$. In effect one loses little by neglecting the sample covariances r_ν^* of high order ν since they contain little information. The class of estimates $f_N^*(\lambda)$ of $f(\lambda)$ discussed above are consistent because they give little weight to the covariances r_ν^* of high order ν. The periodogram $I_N(\lambda)$ is not consistent because it gives too much weight to these high-order covariances. It should also be noted that the truncation at $\nu = h_N$ of the specific estimates mentioned above is convenient for computation because one does not have to compute many of the r_ν^*. In general, since $w_N(x)$ behaves more and more like a delta function at $x = 0$ as $N \to \infty$, $w_\nu^{(N)}$ is close to one when $|\nu|$ is small compared to N. However, since $\int_{-\pi}^{\pi} w_N^2(x)\,dx = o(N)$ as $N \to \infty$, $|w_\nu^{(N)}|$ is very small, that is, practically zero, when $|\nu|$ is large and comparable to N.

We have thus far assumed that the mean value of the process $m_t = Ey_t \equiv 0$. In many cases m_t cannot be assumed to be identically zero. Consider the case in which the mean value m_t has the form of a regression

$$m_t = d_1\varphi_t^{(1)} + d_2\varphi_t^{(2)} + \cdots + d_p\varphi_t^{(p)} \tag{5.20}$$

where $\varphi_t^{(1)}, \cdots, \varphi_t^{(p)}$ are known regression sequences and the regression coefficients $d\nu$ are unknown. To avoid unnecessary complication we will confine ourselves to a brief discussion of the case $p = 2$. The following condition is introduced to prevent the two regression sequences $\varphi_t^{(1)}$ and $\varphi_t^{(2)}$ from becoming linearly dependent in the limit

$$R = \overline{\lim_{N\to\infty}} \frac{\left[\sum_1^N \varphi_t^{(1)}\varphi_t^{(2)}\right]^2}{\sum_1^N \varphi_t^{(1)2} \sum_1^N \varphi_t^{(2)2}} < 1. \tag{5.21}$$

The intuitively plausible procedure in estimating the spectrum of the residual x_t would be to estimate the mean value m_t, subtract it from the observation y_t and apply the techniques discussed above to the differences obtained. This intuitively plausible procedure is reasonable. Assume that $y_1, \cdots, y_N$ have been observed. Let d_1^*, d_2^* be the "least-squares" estimate of d_1 and d_2. By "least-squares" estimates we mean those estimates computed under the assumption that the residuals x_t are orthogonal. Then $m_t^* = d_1^*\varphi_t^{(1)} + d_2^*\varphi_t^{(2)}$ is an estimate of m_t. Compute

$$y_t - m_t^* = y_t - d_1^*\varphi_t^{(1)} - d_2^*\varphi_t^{(2)}, \qquad t = 1, \cdots, N. \tag{5.22}$$

Now compute an estimate $f_N^*(\lambda)$ of $f(\lambda)$ with $y_t - m_t^*, t = 1, \cdots, N$ in place of $x_1, \cdots, x_N$. The asymptotic expression (5.15) for the variance of $f_N^*(\lambda)$ is still valid (see Grenander and Rosenblatt [9]) if

$$\max_{-\pi \leqq u \leqq \pi} w_N(u) = O\left[\int_{-\pi}^{\pi} w_N^2(u)\, du\right]. \tag{5.23}$$

The integral of such an estimate $f_N^*(\lambda)$ of $f(\lambda)$

$$F_N^*(\lambda) = \int_0^{\lambda} f_N^*(u)\, du \tag{5.24}$$

can be used as an estimate of $F(\lambda) = \int_0^{\lambda} f(u)du$. Grenander and Rosenblatt [8] have obtained the asymptotic distribution of

$$\sqrt{N} \max_{0 \leqq \lambda \leqq \pi} |F_N^*(\lambda) - F(\lambda)| \tag{5.25}$$

for a large class of these estimates. One and two-sample tests for the spectral distribution function that are asymptotically nonparametric (with respect to the spectrum) are set up for normal processes. The form of this statistic looks like the analogue of the Kolmogorov statistic. However, this analogy should not be pushed too far as the asymptotic distribution is not the same as the Kolmogorov limiting distribution. The context and the proof are different. The limiting distribution is

$$\lim_{N \to \infty} P\{\max_{0 \leqq \lambda \leqq \pi} \sqrt{N} |F_N^*(\lambda) - F(\lambda)| < a\} = P\{\max_{0 \leqq \lambda \leqq \pi} |\eta(\lambda)| < a\}, \tag{5.26}$$

where $\eta(\lambda)$ is the normal process with mean zero and covariance

$$E\eta(\lambda)\eta(\mu) = e F(\lambda) F(\mu) + 2\pi \int_0^{\pi} f^2(u)\, du. \tag{5.27}$$

Here e is the fourth cumulant

$$e = E\xi_t^4 - 3. \tag{5.28}$$

It is curious and interesting that the parameter e which enters into the asymptotic results here does not enter into the asymptotic results on estimation of the spectral density. However, it is clear that the magnitude of e will determine the sample size at which the asymptotic results on estimation of the spectral density are valid. A heuristic derivation of these results on the distribution of

$$\sqrt{N} \max_{0 \leqq \lambda \leqq \pi} |F_N^*(\lambda) - F(\lambda)| \tag{5.29}$$

is given in Grenander and Rosenblatt [6], [7] together with tables of the limiting distribution.

It would be interesting to get a rigorous derivation of the asymptotic distribution of

$$\max_{0 \leqq \lambda \leqq \pi} |f_N^*(\lambda) - f(\lambda)| \tag{5.30}$$

or

$$\max_{0 \leqq \lambda \leqq \pi} \left| \log \frac{f_N^*(\lambda)}{f(\lambda)} \right| \tag{5.31}$$

since one could then set up confidence bands for the spectral density $f(\lambda)$. There are other open problems in extending the asymptotic distribution theory of the spectral estimates discussed above to the case of a continuous time parameter process. It is clear that analogues of the results obtained above will be valid in the case of a continuous time parameter but the tools required for a rigorous derivation may be heavier.

In the usual solution of the linear prediction problem, the spectrum is assumed known (see Doob [5]). A discussion of the prediction error when the spectral density $f(\lambda)$ is estimated is given in Grenander and Rosenblatt [9].

6. Another class of estimates of the spectral density

The estimates of the spectral density discussed in section 5 are particularly well suited for computation on a digital computer. If great accuracy is desired in the computation, the time series will be discretized if it is not already in discrete form. The data can then be analyzed by using an estimate of the class discussed in section 5.

In many contexts one has a continuous parameter time series and such great accuracy in computation is not required. It is then convenient to set up an analogue computer to compute the estimate. The following class of estimates are convenient to build into an analogue computer.

Let x_t, $0 < t < T$, be the observed time series. Here x_t and t are real valued. Assume that x_t is normally distributed. First mix the time series with a harmonic of frequency λ so as to get $x_t \exp(it\lambda)$. The complex form is used here for convenience in exposition. Then filter the resulting series through the linear filter $\mathcal{A}$ so as to get

$$z_t = \mathcal{A} x_t e^{it\lambda}. \tag{6.1}$$

Let the transient response function of $\mathcal{A}$ be $g_T(u)$ where $g_T(u) = 0$ when $u < 0$. Then

$$z_t = \int_0^t x_s e^{is\lambda} g_T(t-s)\, ds. \tag{6.2}$$

Let

$$\varphi_T(u) = \int_{-\infty}^{\infty} e^{-iut} g_T(t)\, dt, \qquad \int_{-\infty}^{\infty} |\varphi_T(u)|^2 du = 1, \tag{6.3}$$

be the frequency characteristic of the filter $\mathcal{A}$. The function $|\varphi_T(u)|$ is assumed to be highly peaked at $u = 0$. The mixing of x_t with $\exp(it\lambda)$ shifts the peak of $|\varphi_T(u)|$ to λ, the point at which one wishes to estimate $f(\lambda)$. The estimate of $f(\lambda)$ is the average power of the filtered message

$$f_T^*(\lambda) = \frac{1}{T}\int_0^T |z_t|^2 dt. \tag{6.4}$$

Assume that $|\varphi_T(u)|$ is more highly peaked at $u = 0$ with larger sample length T. Under reasonable regularity conditions on the transient response function (see Grenander and Rosenblatt [9]) the variance of the estimate

$$D^2[f_T^*(\lambda)] \sim \frac{2\pi f^2(\lambda)}{T}\int_{-\infty}^{\infty} |\varphi_T(u)|^4 du, \qquad \lambda \neq 0 \tag{6.5}$$

and the bias

$$b_T(\lambda) = Ef_T^*(\lambda) - f(\lambda) \sim \int_{-\infty}^{\infty} [f(u) - f(\lambda)]\, |\varphi_T(u-\lambda)|^2 du. \tag{6.6}$$

A reasonable choice of $g_T(t)$ is

$$g_T(t) = \begin{cases} \sqrt{\frac{\beta}{\pi}}\, e^{-\beta t}, & t \geqq 0, \\ 0, & t < 0, \end{cases} \tag{6.7}$$

where $\beta = KT^{-1/3}$. The mean-square error is then $O(T^{-2/3})$ as $T \to \infty$.

7. Estimation of the spectral density when the parameter is multidimensional

For convenience we discuss the case in which the process is real valued and the parameter is 2-dimensional. The parameter set T consists of the lattice points in 2-space. The process $x_{t,\tau}$ has the representation

$$x_{t,\tau} = \int_{-\pi}^{\pi} \int e^{it\lambda + i\tau\mu} dZ(\lambda, \mu), \tag{7.1}$$

where

$$E[dZ(\lambda_1, \mu_1)\, dZ(\lambda_2, \mu_2)] = \delta_{\lambda_1, \lambda_2}\, \delta_{\mu_1, \mu_2}\, f(\lambda_1, \mu_1)\, d\lambda_1\, d\mu_1 . \tag{7.2}$$

The spectral density $f(\lambda, \mu)$ is a function defined on the square $-\pi \leqq \lambda, \mu \leqq \pi$. The covariances

$$r_{t,\tau} = \int_{-\pi}^{\pi} \int e^{it\lambda + i\tau\mu} f(\lambda, \mu)\, d\lambda\, d\mu . \tag{7.3}$$

Since $x_{t,\tau}$ is real valued it follows that

$$r_{t,\tau} = r_{-t,-\tau} \tag{7.4}$$

and

$$f(\lambda, \mu) = f(-\lambda, -\mu). \tag{7.5}$$

Now assume that a sample $x_{t,\tau}$, $t = 1, \cdots, N$, $\tau = 1, \cdots, M$, has been observed. We want to estimate $f(\lambda, \mu)$ from this sample. Let us first construct a 2-dimensional analogue of the periodogram

$$I_{N,M}(\lambda, \mu) = \frac{1}{(2\pi)^2 NM} \left| \sum_{t=1}^{N} \sum_{\tau=1}^{M} x_{t,\tau} e^{-it\lambda - i\tau\mu} \right|^2 . \tag{7.6}$$

Throughout this discussion $x_{t,\tau}$ is assumed to be a normal process. The results obtained are valid for linear processes. We restrict ourselves to normal processes so as to simplify the derivation. One can show that $I_{N,M}(\lambda, \mu)$ is asymptotically unbiased

$$\lim_{N, M \to \infty} EI_{N,M}(\lambda, \mu) = f(\lambda, \mu). \tag{7.7}$$

However, it is not consistent just as in the 1-dimensional for

$$D^2[I_{N,M}(\lambda, \mu)] \to f^2(\lambda, \mu) \tag{7.8}$$

as $N, M \to \infty$.

Let $w_{N,M}(u, v)$ be a weight function

$$w_{N,M}(u, v) = \frac{1}{(2\pi)^2} \sum_{j,k} w_{j,k}^{(N,M)} e^{-iju - ikv} \tag{7.9}$$

highly peaked at $u, v = 0$. Let $w_{N,M}(u, v)$ be nonnegative and of total mass one

$$\int_{-\pi}^{\pi}\int w_{N,M}(u, v)\,du\,dv = 1\,. \tag{7.10}$$

We also assume that for every $\epsilon > 0$

$$w_{N,M}(u, v) \to 0 \tag{7.11}$$

uniformly for $|u|, |v| \geqq \epsilon$ as $N, M \to \infty$. Consider the following estimate

$$f^*_{N,M}(\lambda, \mu) = \int_{-\pi}^{\pi}\int w_{N,M}(u-\lambda, v-\mu)\, I_{N,M}(u, v)\,du\,dv \tag{7.12}$$

of $f(\lambda, \mu)$. Then

$$\begin{aligned} f^*_{N,M}(\lambda, \mu) &= \frac{1}{(2\pi)^2}\sum_{j=-N}^{N}\sum_{k=-M}^{M} r^*_{j,k}\, w^{(N,M)}_{j,k}\, e^{-ij\lambda-ik\mu} \\ &= \frac{1}{(2\pi)^2 NM}\sum_{j,k}\sum_{\substack{t_1-t_2=j\\ \tau_1-\tau_2=k}} x_{t_1,\tau_1}x_{t_2,\tau_2}\, w^{(N,M)}_{j,k}\, e^{-ij\lambda-ik\mu}\,. \end{aligned} \tag{7.13}$$

Here

$$r^*_{j,k} = \frac{1}{NM}\sum_{t=1}^{N-j}\sum_{\tau=1}^{M-k} x_{t,\tau}\, x_{t+j,\tau+k}\,, \qquad j, k \geqq 0\,, \tag{7.14}$$

is the sample covariance and is analogously defined for j, k with different sign. We shall obtain the asymptotic variance of the estimate $f^*_{N,M}(\lambda, \mu)$ under fairly mild assumptions on the weight function and the spectral density. Now

$$\begin{aligned} &NM\,D^2[f^*_{N,M}(\lambda, \mu)] \\ &= \frac{1}{(2\pi)^4 NM}\sum_{j,k,l,m}\sum_{t_1,t_2,\tau_1,\tau_2} \operatorname{cov}(x_{t_1,\tau_1}x_{t_1+j,\tau_1+k},\, x_{t_2,\tau_2}x_{t_2+l,\tau_2+m}) \\ &\qquad \cdot w^{(N,M)}_{j,k} w^{(N,M)}_{l,m}\, e^{-ij\lambda-ik\mu}\, e^{-il\lambda-im\mu} \\ &= \frac{1}{(2\pi)^4 NM}\sum_{j,k,l,m}\sum_{t_1,t_2,\tau_1,\tau_2} (r_{t_1-t_2,\,\tau_1-\tau_2}\, r_{t_1-t_2+j-l,\,\tau_1-\tau_2+k-m} \\ &\qquad + r_{t_1-t_2-l,\,\tau_1-\tau_2-m}\, r_{t_1-t_2+j,\,\tau_1-\tau_2+k})\, w^{(N,M)}_{j,k} w^{(N,M)}_{l,m}\, e^{-ij\lambda-ik\mu}\, e^{-il\lambda-im\mu} \\ &= S_1 + S_2\,. \end{aligned} \tag{7.15}$$

Both of the terms S_1, S_2 can be treated in the same way. We shall consider the term S_2. Now

$$\begin{aligned} S_2 &= \frac{1}{(2\pi)^4}\sum_{j,k,l,m}\sum_{p,q}\left(1-\frac{|p|}{N}\right)\left(1-\frac{|q|}{M}\right) \\ &\qquad \cdot r_{p-l,\,q-m}\, r_{p+j,\,q+k}\, w^{(N,M)}_{j,k} w^{(N,M)}_{l,m}\, e^{-ij\lambda-ip\mu}\, e^{-il\lambda-im\mu} \\ &\sim (2\pi)^2\int_{-\pi}^{\pi}\int \frac{1}{2\pi N}\frac{\sin^2\frac{N}{2}u}{\sin^2\frac{1}{2}u}\,\frac{1}{2\pi M}\frac{\sin^2\frac{M}{2}v}{\sin^2\frac{1}{2}v}\,\{f(u, v) \\ &\qquad \cdot w_{N,M}(u-\lambda, v-\mu)\} * \{f(u, v)\, w_{N,M}(u-\lambda, v-\mu)\}\,du\,dv \end{aligned} \tag{7.16}$$

where the * denotes convolution with respect to u, v. Assume that $f(u, v)$ is continuous and positive. Let

$$v_{N,M}(u, v) = w_{N,M}(u, v) * w_{N,M}(u, v). \tag{7.17}$$

We also assume that

$$\left| \frac{v_{N,M}(u, v)}{v_{N,M}(0, 0)} - 1 \right| \to 0 \tag{7.18}$$

as $N, M \to \infty$ when $|u| < A/N$, $|v| < A/M$ where A is any positive constant. It then follows that

$$\begin{aligned} S_2 &\sim (2\pi)^2 \{ f(u, v) w_{N,M}(u - \lambda, v - \mu) \} \\ &\qquad * \{ f(u, v) w_{N,M}(u - \lambda, v - \mu) \}_{u,v=0} \\ &= (2\pi)^2 \int_{-\pi}^{\pi} \int f^2(u, v) w^2_{N,M}(u - \lambda, v - \mu)\, du\, dv \\ &\sim (2\pi)^2 f^2(\lambda, \mu) \int_{-\pi}^{\pi} \int w^2_{N,M}(u, v)\, du\, dv \end{aligned} \tag{7.19}$$

as $N, M \to \infty$. The same sort of argument shows that $S_1 = o(S_2)$ if $(\lambda, \mu) \neq (0, 0)$ and $S_1 \sim S_2$ if $(\lambda, \mu) = (0, 0)$. But then

$$D^2[f^*_{N,M}(\lambda, \mu)] \sim \frac{(2\pi)^2}{NM} f^2(\lambda, \mu) \int_{-\pi}^{\pi} \int w^2_{N,M}(u, v)\, du\, dv \tag{7.20}$$

if $(\lambda, \mu) \neq (0, 0)$ as $N, M \to \infty$. The asymptotic expression given in (7.20) should be doubled when $(\lambda, \mu) = (0, 0)$. Thus $f^*_{N,M}(\lambda, \mu)$ is a consistent estimate of $f(\lambda, \mu)$ if

$$\frac{1}{NM} \int_{-\pi}^{\pi} \int w^2_{N,M}(u, v)\, du\, dv \to 0 \tag{7.21}$$

as $N, M \to \infty$.

These estimates are well suited for computation on a digital machine. It is rather doubtful whether one could build useful analogue computers making use of the 2-dimensional counterparts of the estimates discussed in section 6.

The following problem is interesting and has not yet been answered satisfactorily. Suppose that $x_{t,\tau}$ is a continuous parameter process and $f(\lambda, \mu)$ is known to be circularly symmetric about zero. What would then be an efficient estimate of $f(\lambda, \mu)$ making use of the known symmetry? There are, of course, many higher dimensional analogues of this problem.

8. Estimation of the cospectrum and quadrature spectrum

Let $x_t = \begin{pmatrix} {}_1x_t \\ {}_2x_t \end{pmatrix}$ be a process with real-valued components and t integral. The spectrum is assumed to be absolutely continuous with a continuous and nonsingular spectral density. The sample x_t, $t = 1, \cdots, N$, has been observed and we want to estimate the cospectrum $\operatorname{Re} f_{12}(\lambda)$ and the quadrature spectrum $\operatorname{Im} f_{12}(\lambda)$ (see section 3).

We first discuss estimation of the cospectrum. Let a be any real number. The time series ${}_1x_t + a\, {}_2x_t$ then has the spectral density

$$f_{11}(\lambda) + 2a \operatorname{Re} f_{12}(\lambda) + a^2 f_{22}(\lambda). \tag{8.1}$$

Now

$$(8.2)\qquad \int_{-\pi}^{\pi} w_N(u-\lambda)\frac{1}{2\pi N}\Big|\sum ({}_1x_t + a\,{}_2x_t)\,e^{itu}\Big|^2 du$$

is a reasonable estimate of the spectral density (8.1) if the weight function $w_N(u)$ satisfies the conditions cited in section 3. This implies that

$$(8.3)\qquad \int_{-\pi}^{\pi} w_N(u-\lambda)\,\mathrm{Re}\left\{\frac{1}{2\pi N}\sum_{t=1}^{N} {}_1x_t e^{itu}\sum_{\tau=1}^{N} {}_2x_\tau e^{-iru}\right\} du = \frac{1}{2\pi}\sum_{\nu=-N}^{N} w_\nu^{(N)}\,{}_{12}r_\nu^{*}\cos\nu\lambda$$

is an estimate of $\mathrm{Re}\,f_{12}(\lambda)$, where

$$(8.4)\qquad {}_{12}r_\nu^{*} = \frac{1}{N}\sum_{t-\tau=\nu} {}_1x_t\,{}_2x_\tau .$$

But an argument similar to that of section 7 implies that

$$(8.5)\qquad \mathrm{cov}\left\{\int_{-\pi}^{\pi} w_N(u-\lambda)\frac{1}{2\pi N}\Big|\sum {}_1x_t e^{itu}\Big|^2 du,\ \int_{-\pi}^{\pi} w_N(u-\lambda)\frac{1}{2\pi N}\Big|\sum {}_2x_t e^{itu}\Big|^2 du\right\}$$
$$\sim \frac{2\pi}{N}|f_{12}(\lambda)|^2\int_{-\pi}^{\pi} w_N^2(u)\,du$$
$$= \frac{2\pi}{N}\{[\mathrm{Re}\,f_{12}(\lambda)]^2 + [\mathrm{Im}\,f_{12}(\lambda)]^2\}\int_{-\pi}^{\pi} w_N^2(u)\,du .$$

It then follows that the asymptotic variance of (8.3) is

$$(8.6)\qquad \frac{\pi}{N}\{f_{11}(\lambda)f_{22}(\lambda) + [\mathrm{Re}\,f_{12}(\lambda)]^2 - [\mathrm{Im}\,f_{12}(\lambda)]^2\}\int_{-\pi}^{\pi} w_N^2(u)\,du .$$

One can see that a reasonable estimate of $\mathrm{Im}\,f_{12}(\lambda)$ is

$$(8.7)\qquad \int_{-\pi}^{\pi} w_N(u-\lambda)\,\mathrm{Im}\left\{\frac{1}{2\pi N}\sum_{t=1}^{N} {}_1x_t e^{itu}\sum_{\tau=1}^{N} {}_2x_\tau e^{-iru}\right\} du = \frac{1}{2\pi}\sum_{\nu=-N}^{N} w_\nu^{(N)}\,{}_{12}r_\nu^{*}\sin\nu\lambda .$$

An argument similar to that given above indicates that the asymptotic variance of (8.7) is

$$(8.8)\qquad \frac{\pi}{N}\{f_{11}(\lambda)f_{22}(\lambda) + [\mathrm{Im}\,f_{12}(\lambda)]^2 - [\mathrm{Re}\,f_{12}(\lambda)]^2\}\int_{-\pi}^{\pi} w_N^2(u)\,du, \quad \lambda \neq 0 .$$

This expression for the asymptotic variance should not be taken seriously very close to $\lambda = 0$ as the estimate and $\mathrm{Im}\,f_{12}(\lambda)$ are zero at $\lambda = 0$.

9. Prefiltering of a time series

The results given above are asymptotic results. Here asymptotic must be interpreted not only in terms of the magnitude of N but also in terms of the variation of the spectral

density. If, for example, one has a rapidly changing spectral density, one must expect a certain amount of contamination in the estimation of the spectral density at such a point from the immediate neighborhood. In many cases one will not know this beforehand. However, one may be led to this belief from rapid change in the estimated spectral density. Tukey suggests that if one suspects this at a point one should prefilter the series so as to smooth out the spectrum in the neighborhood of the point and estimate the spectrum of the filtered process. Assume the weight function $w_N(u)$ is rectangular for convenience, that is,

$$w_N(u) = \begin{cases} \frac{1}{2h}, & \text{if} \quad |u| \leqq h, \\ 0, & \text{otherwise} \end{cases} \tag{9.1}$$

(see [9]). Then the bias is asymptotically proportional to $f''(\lambda)$. Let us see what effect such prefiltering has on an estimate of the spectral density and how it should most advantageously be set up.

Let $\varphi(\lambda)$ be the frequency characteristic of the prefiltering. The spectrum of the filtered process is then

$$g(\lambda) = |\varphi(\lambda)|^2 f(\lambda). \tag{9.2}$$

Let the estimate of the modified spectrum be $g^*(\lambda)$. Then $f^*(\lambda) = g^*(\lambda)/|\varphi(\lambda)|^2$ is the natural estimate of $f(\lambda)$. The variance of $f^*(\lambda)$ is still asymptotically

$$\frac{2\pi f^2(\lambda)}{N} \int_{-\pi}^{\pi} w_N^2(u)\, du \tag{9.3}$$

so that it has no effect on the variance.

Let us see what effect it has on the bias. The estimate $f^*(\lambda)$ now has bias proportional to

$$\frac{[|\varphi(\lambda)|^2 f(\lambda)]''}{|\varphi(\lambda)|^2} = f''(\lambda) + \frac{2(|\varphi(\lambda)|^2)'}{|\varphi(\lambda)|^2} f'(\lambda) + \frac{(|\varphi(\lambda)|^2)''}{|\varphi(\lambda)|^2} f(\lambda). \tag{9.4}$$

In the case of a minimum at λ

$$f''(\lambda) > 0, \; f'(\lambda) = 0, \tag{9.5}$$

and $f(\lambda)$ is small so that the bias is proportional to

$$f''(\lambda) + \frac{(|\varphi(\lambda)|^2)''}{|\varphi(\lambda)|^2} f(\lambda). \tag{9.6}$$

In smoothing the spectrum we would have $|\varphi(\lambda)|^2$ high at λ. If we have a maximum of $|\varphi(\lambda)|^2$ near λ then $[|\varphi(\lambda)|^2]'' < 0$ so that if $\varphi(\lambda)$ is expeditiously chosen the bias would be less. It would then be best to choose $|\varphi(\lambda)|^2$ so that

$$\frac{[|\varphi(\lambda)|^2]''}{|\varphi(\lambda)|^2} = -\frac{f''(\lambda)}{f(\lambda)}. \tag{9.7}$$

REFERENCES

[1] M. S. BARTLETT, "Smoothing periodograms from time series with continuous spectra," *Nature*, Vol. 161 (1948), pp. 686–687.

[2] ———, "Periodogram analysis and continuous spectra," *Biometrika*, Vol. 37 (1950), pp. 1–16.

[3] G. K. BATCHELOR, *The Theory of Homogeneous Turbulence* (Cambridge Monograph), Cambridge, The University Press, 1953.

4] H. CRAMÉR, "On the theory of stationary random processes," *Annals of Math.*, Vol. 41 (1940), pp. 215–230.
[5] J. L. DOOB, *Stochastic Processes*, New York, John Wiley and Sons, 1953.
[6] U. GRENANDER and M. ROSENBLATT, "On spectral analysis of stationary time series," *Proc. Nat. Acad. Sci.*, Vol. 38 (1952), pp. 519–521.
[7] ———, "Comments on statistical spectral analysis," *Skand. Aktuar.*, Vol. 36 (1953), pp. 182–202.
[8] ———, "Statistical spectral analysis of time series arising from stationary stochastic processes," *Annals of Math. Stat.*, Vol. 24 (1953), pp. 537–558.
[9] ———, monograph on time series analysis, to be published.
[10] J. L. LAWSON and G. E. UHLENBECK, *Threshold Signals* (M.I.T. Radiation Laboratory Series), Vol. 24, New York, McGraw-Hill Book Co., 1950.
[11] P. A. P. MORAN, "Some theorems on time series I," *Biometrika*, Vol. 34 (1947), pp. 281–291.
[12] H. A. PANOFSKY, "Power-spectrum analysis in meteorology," to be published in *The Jour. Amer. Stat. Assoc.*
[13] W. J. PIERSON, JR., *A Unified Mathematical Theory for the Analysis, Propagation and Refraction of Storm Generated Ocean Surface Waves*, Parts I and II, New York, Research Division, College of Engineering, New York University, 1952.
[14] J. TUKEY, "Measuring noise color," an unpublished manuscript.