AD 606275

OPTIMUM LINEAR ESTIMATION FOR RANDOM PROCESSES AS THE LIMIT OF ESTIMATES BASED ON SAMPLED DATA

Peter Swerling

P-1206

October 24, 1957

The RAND Corporation

1700 MAIN ST. • SANTA MONICA • CALIFORNIA

(24)

CLEARINGHOUSE FOR FEDERAL SCIENTIFIC AND TECHNICAL INFORMATION, CFSTI
DOCUMENT MANAGEMENT BRANCH 410.11

LIMITATIONS IN REPRODUCTION QUALITY

Accession # AD 606275

☐ 1. We regret that legibility of this document is in part unsatisfactory. Reproduction has been made from best available copy.

☑ 2. A portion of the original document contains fine detail which may make reading of photocopy difficult.

☐ 3. The original document contains color, but distribution copies are available in black-and-white reproduction only.

☐ 4. The initial distribution copies contain color which will be shown in black-and-white when it is necessary to reprint.

☐ 5. Limited supply on hand; when exhausted, document will be available in Microfiche only.

☐ 6. Limited supply on hand; when exhausted document will not be available.

☐ 7. Document is available in Microfiche only.

☐ 8. Document available on loan from CFSTI (TT documents only).

☐ 9.

Processor: ral

NBS____ 9/64

ACKNOWLEDGMENT

A portion of this work was done while the author was a member of the staff of the Control Systems Laboratory of the University of Illinois.

Also, the author wishes to thank Professor Edgar Reich of the University of Minnesota for many invaluable discussions of this subject.

SUMMARY

We consider a generalized form of the problem of optimum linear filtering and prediction for random processes. It is shown that, under very general conditions, the optimum linear estimation based on the received signal, observed continuously for a finite interval $a \leq t \leq b$, is the limit of optimum linear estimation based on sampled data - i.e. based on the received signal observed at only a finite number of points in the interval $a \leq t \leq b$ - as the points of observation become denser and denser in the interval.

This yields a method for obtaining the optimum linear estimation in cases where the conventional generalized Wiener-Hopf integral equation technique has not been shown to yield a solution. The relationship between the sampled-data solution and the Wiener-Hopf integral equation solution (when the latter exists) is discussed.

Also, a problem is posed concerning the rate at which the error variance of optimum sampled-data estimates approaches the error variance of the optimum estimate based on continuous observation, as the sampled points become denser in the observation interval. This problem is solved in one case.

I. INTRODUCTION

A. Statement of the Problem

A large class of linear estimation problems can be cast in the following form: suppose one is given a random process $z(t)$ in an interval $a \leq t \leq b$. A sample function of this process will be called an observed sample. It is assumed that

$$E\left[z(t)\right] = 0 \qquad \text{for } a \leq t \leq b \tag{1}$$

$$\phi_z(s,t) = E\left[z(s)z(t)\right] < \infty \quad \text{for } a \leq s \leq b, \; a \leq t \leq b \tag{2}$$

Here and in the following, the notation $E\left[\;\right]$ denotes the <u>ensemble</u> expected value of the quantity in brackets. The function ϕ_z is called the covariance function of the random process $z(t)$. We do not assume that $z(t)$ is necessarily stationary.

Also suppose there is a random variable, q, with

$$E\left[q\right] = 0 \tag{3}$$

$$E\left[q^2\right] < \infty \tag{4}$$

and

$$\rho(t) = E\left[qz(t)\right] < \infty \qquad \text{for } a \leq t \leq b \tag{5}$$

The problem is to form the optimum linear estimate $\tilde{q}$ of q in the following sense:

Find an estimate $\tilde{q}$ of q such that

(a) $\tilde{q}$ is formed by a linear operation on $z(t)$, in the sense to be defined in Section I.B; and

(b) $E\left[(\tilde{q} - q)^2\right]$ is a minimum for all estimates satisfying (a).

In Section I.B below we define the class of linear operations on $z(t)$ within which the optimum is to be found. This class contains all the types of linear operations which are usually considered. Also, the existence of a unique $\tilde{q}$ which is optimum within this class is guaranteed (see for example Ref. 1 or Ref. 6).

An example of a more familiar-looking type of problem which can be represented in the above form is as follows:

Let the random process $x(t)$ be called the 'signal,' with

$$E\left[x(t)\right] = 0 \tag{6}$$

$$\phi_x(s,t) = E\left[x(s)\,x(t)\right] < \infty \tag{7}$$

Let the random process $n(t)$ be called the 'noise,' with

$$E\left[n(t)\right] = 0 \tag{8}$$

$$\phi_n(s,t) = E\left[n(s)\,n(t)\right] < \infty \tag{9}$$

and

$$\phi_{xn}(s,t) = E\left[x(s)\,n(t)\right] < \infty \tag{10}$$

The observed process $z(t)$ is

$$z(t) = x(t) + n(t) \qquad \text{for } a \leq t \leq b \tag{11}$$

Thus

$$\phi_z(s,t) = E\left[z(s)\,z(t)\right] = \phi_x(s,t) + \phi_n(s,t) + \phi_{xn}(s,t) + \phi_{xn}(t,s) \tag{12}$$

A typical interpolation or extrapolation problem is to estimate the signal value $x(t_o)$ at some time t_o by a linear operation on the observed sample $z(t)$ over the interval $\left[a,b\right]$. If t_o is in the observation interval

$[a,b]$ this is called an interpolation (filtering or smoothing) problem; if t_0 is outside $[a,b]$, an extrapolation (prediction) problem.

Stated in the formulation of the first few paragraphs above, we would have

$$q = x(t_0) \tag{13}$$

$$\rho(t) = E\left[z(t)\,x(t_0)\right] = \phi_x(t_0,t) + \phi_{xn}(t_0,t) \tag{14}$$

and $\phi_z(s,t)$ given by Eq. (12). We wish to find a linear estimate $\check{x}(t_0)$ for which $E\left[\{\check{x}(t_0) - x(t_0)\}^2\right]$ is a minimum.

Frequently the time t_0 and the observation interval $[a,b]$ are regarded as varying, and one is interested in the manner in which the optimum linear operator on $z(t)$ and the minimum error variance vary with t_0, a, or b. The formulas in the remainder of this paper will not explicitly reflect the variation with t_0, since they will be expressed in terms of the formulation of the first few paragraphs. However, the dependence of the optimum estimator and the minimum error variance on t_0 in this case can always be made explicit by use of Eqs. (12), (13), and (14). In a similar manner, many other problems such as estimation of the derivative of the signal, etc., can be translated into the formulation of the first few paragraphs.

In another problem which is frequently studied, it is assumed that the signal $x(t)$ has a non-random component: $E\left[x(t)\right] = m(t)$, where $m(t)$ is known except for a finite number of unknown but non-random parameters, upon which $m(t)$ depends linearly, i.e. $m(t) = \sum_{k=1}^{N} a_k P_k(t)$ where $P_k(t)$ are known functions. In this case, the objective is to find, say, the optimum linear estimate $\check{x}(t_0)$ of $x(t_0)$ subject to the additional restriction that $E\left[x(t_0)\right] = m(t_0)$ identically in the unknown parameters a_k. The main results below can be extended to this case in a fairly straightforward way. This is left as an exercise for the interested reader.

B. Definition of Linear Operation on the Observed Process

Definition. Given the random process $z(t)$ in the interval $a \leq t \leq b$, assumed to have zero ensemble mean and finite covariance function, a (real) linear operation on $z(t)$ we define to be any random variable of the form

$$\underset{n \to \infty}{\mathrm{l.i.m.}} \sum_{i=1}^{n} c_i^{(n)} z(t_i^{(n)}) \tag{15}$$

where l.i.m. stands for 'limit in the mean,' $c_i^{(n)}$ are real constants, and $t_i^{(n)}$ are points of $[a,b]$. (Here and in the following, the notation $[a,b]$ signifies the closed interval $a \leq t \leq b$.)

All the types of linear operations which are usually considered are special cases of this. Also, the existence of a unique linear estimate γ which is optimum within this class of estimates is guaranteed.

For processes which are continuous in the mean over a finite interval, an alternative (and equivalent) definition is as follows (Ref. 1): a theorem of Karhunen states that processes $z(t)$ which are continuous in the mean over the finite interval $[a,b]$ can be represented in the form

$$z(t) = \sum_{\nu=1}^{\infty} z_\nu \frac{\phi_\nu(t)}{\sqrt{\lambda_\nu}} \tag{16}$$

with convergence in the mean for t in $[a,b]$. $\phi_\nu(t)$ and λ_ν are the eigenfunctions and eigenvalues of the integral equation

$$\phi(s) = \lambda \int_a^b \mathcal{C}_z(s,t)\, \phi(t)\,dt \tag{17}$$

and z_ν are random variables given by

$$z_\nu = \sqrt{\lambda_\nu} \int_a^b z(t)\, \phi_\nu(t)dt \tag{18}$$

with

$$E\left[z_\nu\right] = 0; \qquad E\left[z_\mu z_\nu\right] = \delta_{\mu\nu} \quad \text{(Kronecker delta)} \tag{19}$$

A linear operation on the process z(t) is then defined to be a random variable of the form

$$\sum_{\nu=1}^{\infty} \alpha_\nu z_\nu \tag{20}$$

where

$$\sum_{\nu=1}^{\infty} \alpha_\nu^2 < \infty$$

There are useful applications for both of these definitions of linear operation to the linear interpolation and extrapolation problem. The optimum linear estimate $\check{q}$ and the minimum error variance $\check{\mu}$ can be found, for processes which are continuous in the mean, by utilizing the representation (16); (see Ref. 1, section 6 and the Appendix to the present paper). However, for purposes of this paper the first definition above of linear operation is of main interest.

C. The variational or 'Integral Equation' Method of Solution

What might be called the conventional method of solution of the problem posed in Section I.A is as follows: one considers linear operators of the form

$$\hat{q} = \int_a^b z(\tau)\, k(\tau)d\tau \tag{21}$$

Even in the simplest cases, when dealing with finite observation intervals, it is necessary to assume that k may contain delta functions of various orders -- that is, that $\hat{q}$ may give finite weight to the values of z(t) or its derivatives at individual points.

For estimates $\hat{q}$ of the form (21),

$$E\left[(\hat{q} - q)^2\right] = \int_a^b k(\tau)\left[\int_a^b \phi_z(\sigma,\tau)\, k(\sigma)d\sigma - 2\rho(\tau)\right] d\tau + E\left[q^2\right] \quad (22)$$

Denoting the estimate which minimizes the expression in (22) by $\check{q}$, and the corresponding weighing function by $\check{k}$, an application of the usual variational technique yields the following integral equation for $\check{k}$:

$$\int_a^b \check{k}(\tau)\, \phi_z(\tau,t)\, d\tau = \rho(t) \quad (a \leq t \leq b) \quad (23)$$

where ϕ_z and ρ are given by (2) and (5). The minimum error variance $E\left[(\check{q} - q)^2\right]$ is given by

$$E\left[(\check{q} - q)^2\right] = E\left[q^2\right] - \int_a^b \check{k}(\tau)\, \rho(\tau)\, d\tau \quad (24)$$

Even in the simplest cases, when the observation interval is finite the integral Eq. (23) does not have a solution unless $\check{k}$ is permitted to contain delta functions.

Zadeh and Ragazinni (Ref. 2) have given the solution to this equation when ϕ_z is a stationary kernel (i.e. depends only on $|s - t|$) and corresponds to a rational spectral density function. It turns out that for such cases $\check{k}$ need only contain delta functions at the endpoints a and b of the observation interval.* This is a fairly remarkable fact. There are important stationary random processes with rational spectral densities for which the sample functions are differentiable to a certain order, say ν, with probability one. In such cases, one might expect that the optimum linear operator might be of the form

$$\sum_{i=0}^{\nu} \int_a^b z^{(i)}(\tau)\, dK_i(\tau) \quad (25)$$

* Actually some further conditions on ϕ_z and ρ are needed.

where K_i are functions of bounded variation and $z^{(i)}(\tau)$ is the $i^{\underline{th}}$ derivative of $z(\tau)$. The fact that the optimum operators in these cases actually involve delta functions only at the points a and b is fairly remarkable.

It is of course necessary to justify the variational derivation of the integral Eq. (23) when $\hat{k}$ is permitted to contain delta functions. That is, even if (23) has a solution, it is in such cases necessary to give a separate proof that the satisfaction of (23) is in fact a necessary and sufficient condition for the minimization of the expression in (22); see for example Ref. 4. Also, one should prove that the optimum estimate derived in this way is actually the optimum of all possible linear estimates and not just the optimum among estimates of the form (21). This can be proved in many cases.

For general kernels $\phi_z(s,t)$, however, one cannot assume that (23) has a solution even if $\hat{k}$ is permitted to contain delta functions; nor can one assume that the optimum linear estimate is of the form (25), say.

It would be very desirable to exhibit a method for obtaining the optimum linear estimate $\hat{q}$, without making any special assumptions as to the form of ϕ_z. The main purpose of this paper is to exhibit such a method, subject only to the very general restriction that the process $z(t)$ be continuous in the mean over $[a,b]$. (The Appendix briefly discusses an alternative method for obtaining $\hat{q}$.)

II. OPTIMUM LINEAR ESTIMATES AS LIMITS OF OPTIMUM LINEAR SAMPLED DATA ESTIMATES

Suppose we select n distinct points $t_1^{(n)}, \ldots, t_n^{(n)}$ from the interval $[a,b]$. We will call this a subdivision γ_n of the interval $[a,b]$.* We need not assume that these points are numbered in order of magnitude.

An estimate which depends only on the values of $z(t)$ at a finite number of points will be referred to as a sampled-data estimate. The most general linear operation which can be performed on the set of random variables $z(t_i^{(n)})$, $i = 1,\ldots,n$, is of the form

$$\hat{q}_n = \sum_{i=1}^{n} c_i^{(n)} z(t_i^{(n)}) \qquad (26)$$

where $c_i^{(n)}$ are real constants.

Definition: We define the optimum linear estimate $\tilde{q}_n$ of q, based on the observed values of $z(t)$ at just the points $t_i^{(n)}$, $i = 1,\ldots,n$, to be that random variable of the form given by (26) which minimizes $E(\hat{q}_n - q)^2$ among all random variables $\hat{q}_n$ of the form given by (26). (See Ref. 3 for a detailed discussion.)

(It is of course understood that q_n and $\tilde{q}_n$ depend on the specific points $t_i^{(n)}$ and not just on the number n of points, but for convenience this will not be explicitly indicated in the notation.)

Later on we will give explicit formulas for $\tilde{q}_n$. At this point, we wish to bring out the fact that the overall optimum linear estimate $\tilde{q}$ is a limit of optimum estimates based on sampled data.

---

* γ_n actually depends, of course, on the specific points $t_i^{(n)}$ selected, not just on the number of points.

For reference, we first give the following well known definition:
The process $z(t)$ is said to be continuous in the mean for $a \leq t \leq b$ provided $E\left[|z(t) - z(s)|^2\right]$ approaches zero as $s \to t$, for each t in $[a,b]$.

Theorem 1. Let $z(t)$ be continuous in the mean over the finite interval $[a,b]$. Let $\{\mathcal{S}_n\}$ be a sequence of subdivisions of $[a,b]$ and define

$$\Delta_n = \text{length of maximum interval between neighboring points of the point set consisting of } a,\ b,\ t_1^{(n)}, \ldots\ldots, t_n^{(n)} \tag{27}$$

where $t_i^{(n)}$, $i = 1, \ldots, n$ are the points corresponding to the subdivision $\mathcal{S}_n$. Suppose that $\Delta_n \to 0$ as $n \to \infty$. Then the overall optimum linear estimate $\check{q}$ is given by

$$\check{q} = \underset{n \to \infty}{\text{l.i.m}}\ \check{q}_n \tag{28}$$

Also, let

$$\check{\mu} = E\left[(\check{q} - q)^2\right]$$

$$\check{\mu}_n = E\left[(\check{q}_n - q)^2\right]$$

Then

$$\check{\mu} = \lim_{n \to \infty} \check{\mu}_n \tag{29}$$

Proof: Since $\check{q}$ is a linear operation on $z(t)$, $\check{q}$ is of the form

$$\check{q} = \underset{n \to \infty}{\text{l.i.m.}} \sum_{i=1}^{n} c_i^{(n)}\, z(\tau_i^{(n)}) \tag{30}$$

where $\tau_i^{(n)}$ are points of $[a,b]$, possibly different from $t_i^{(n)}$.

(Thus, by definition of 'linear operation on $z(t)$,' $\check{q}$ is the limit in the mean of <u>some</u> sequence of linear operations based on sampled data. The point of the theorem is to exhibit sequences of specific linear operations on sampled data which converge in mean to $\tilde{q}$.)

By continuity in the mean of $z(t)$, each estimate of the form

$\sum_{i=1}^{n} c_i^{(n)} \; z(\tau_i^{(n)})$ can be approximated in the mean arbitrarily closely

by linear operations on the variables $z(t_i^{(m)})$, for

sufficiently large m. Thus, one can without loss of generality say that

$$\tilde{q} = \underset{n \to \infty}{\text{l.i.m.}} \sum_{i=1}^{n} c_i^{(n)} \; z(t_i^{(n)}) \qquad (31)$$

for some set of constants $c_i^{(n)}$. For this set of constants for which (31) holds, let

$$\hat{q}_n = \sum_{i=1}^{n} c_i^{(n)} \; z(t_i^{(n)}) \qquad (32)$$

so that $\tilde{q}$ = l.i.m. $\hat{q}_n$.

Also let

$$\hat{\mu}_n = E\left[(\hat{q}_n - q)^2\right] \qquad (33)$$

Since q_n is defined as the optimum linear estimate based on the points of $\mathcal{T}_n$,

$$\tilde{\mu} \leq \mu_n \leq \hat{\mu}_n \quad , \quad \text{all } n \qquad (34)$$

But, since $\tilde{q}$ = l.i.m. $\hat{q}_n$, it follows that $\tilde{q} - q$ = l.i.m. $(\hat{q}_n - q)$ and therefore $\tilde{\mu} = \lim \hat{\mu}_n$. Therefore, from (34),

$$\tilde{\mu} = \lim_{n \to \infty} \hat{\mu}_n = \lim_{n \to \infty} \mu_n \qquad (35)$$

Now, we wish also to show that $\check{q} = \text{l.i.m.}\ \check{q}_n$. This will be accomplished if we can show that $E\left[(\check{q}_n - \hat{q}_n)^2\right] \to 0$ as $n \to \infty$. To this end, let

$$\alpha' = (1 - k)\check{q}_n + k\hat{q}_n \qquad (36)$$

where k is any real number.

Then,

$$\alpha' - q = (1 - k)(\check{q}_n - q) + k(\hat{q}_n - q) \qquad (37)$$

Therefore,

$$E\left[(\alpha' - q)^2\right] = \tilde{\mu}_n \left[(1 - k)^2 + \frac{2\rho k(1 - k)}{\sqrt{e}} + \frac{k^2}{e}\right] \qquad (38)$$

where

$$\rho = \frac{1}{\sqrt{\hat{\mu}_n \check{\mu}_n}}\, E\left[(\hat{q}_n - q)(\check{q}_n - q)\right] \qquad (39)$$

and

$$e = \frac{\check{\mu}_n}{\hat{\mu}_n} \qquad (40)$$

Then

$$\frac{E\left[(\alpha' - q)^2\right]}{\tilde{\mu}_n} = 1 + 2k\left(\frac{\rho - \sqrt{e}}{\sqrt{e}}\right) + k^2\left(\frac{1 - 2\rho\sqrt{e} + e}{e}\right) \qquad (41)$$

Therefore, it must be true that

$$\rho = \sqrt{e} = \sqrt{\frac{\check{\mu}_n}{\hat{\mu}_n}} \qquad (42)$$

since otherwise it would be possible to choose k so that $E\left[(\alpha' - q)^2\right] < \tilde{\mu}_n$, which is impossible, since α' is a linear estimate based just on the points of $\mathcal{S}_n$.

By (42) and (39),

$$E\left[(\hat{q}_n - q)(\hat{\hat{q}}_n - q)\right] = \mu_n \tag{43}$$

Therefore,

$$E\left[(\hat{q}_n - \hat{\hat{q}}_n)^2\right] = \hat{\mu}_n - \mu_n \tag{44}$$

Therefore, by (35), $E\left[(\hat{q}_n - \hat{\hat{q}}_n)^2\right] \to 0$ as $n \to \infty$. This proves Theorem 1.

Corollary 1: Let C by any class of linear estimates of q which contains all estimates of the form

$$\hat{q} = \int_a^b z(\tau)\, dK(\tau)$$

where K is of bounded variation. If there exists in C an estimate, say $\hat{\hat{q}}$, which is optimum among all estimates belonging to C, then q is the overall optimum linear estimate $\hat{q}$.

Proof: This is an immediate consequence of

(a) the uniqueness of q,

(b) the fact that all sampled-data estimates belong to any such class C, and

(c) Theorem 1.

Specific formulas for the optimum sampled-data estimates q_n are as follows (see also Ref. 3):

Consider the subdivision π_n of $[a,b]$ with points $t_1^{(n)}, \ldots, t_n^{(n)}$. Define

$$\phi_{zij} = \phi_z\left(t_i^{(n)}, t_j^{(n)}\right)$$

$$\rho_i = \rho\left(t_i^{(n)}\right)$$

$$z_i = z\left(t_i^{(n)}\right)$$

It is desired to find constants $\hat{c}_i^{(n)}$ such that $E\left|\sum_{i=1}^{n} c_i^{(n)} z(t_i^{(n)}) - q\right|^2$ is minimized. In the following, in order to avoid circumlocutions, it will be assumed that the matrix (ϕ_{zij}) is nonsingular. Of course, this assumption is not necessary for the existence of a unique optimum $\hat{q}_n$ (although if (ϕ_{zij}) is singular, the set of constants $\hat{c}_i^{(n)}$ giving $\hat{q}_n$ will not be unique). However, assuming nonsingularity of (ϕ_{zij}) for convenience, the process of finding $\hat{c}_i^{(n)}$ is a simple differentiation problem, the result being

<u>**Theorem 2:**</u> Let (ϕ_{zij}) be nonsingular, and let

$$(\xi_{zij}) = (\phi_{zij})^{-1} \quad \text{(matrix inverse)} \tag{46}$$

Then

$$\hat{q}_n = \sum_{i,j=1}^{n} \xi_{zij}\, z_i\, \rho_j \tag{47}$$

$$\mu_n = E\, q^2 - \sum_{i,j=1}^{n} \xi_{zij}\, \rho_i\, \rho_j \tag{48}$$

It is interesting to note what happens heuristically when n is allowed to approach infinity (and Δ_n defined by (27) approaches zero) in (47) and (48). We have

$$\hat{q}_n = \sum_{i=1}^{n} k_i\, z_i \quad \text{and} \quad \mu_n = E\, q^2 - \sum_{i=1}^{n} k_i\, \rho_i$$

where

$$k_i = \sum_{j=1}^{n} \xi_{zij}\, \rho_j$$

That is, (k_i) is the solution of the equation $\rho_i = \sum_{j=1}^{n} \phi_{zij}\, k_j$

In the limit, then (of course, on a purely heuristic basis),

$$\tilde{q}_n \longrightarrow \int_a^b \tilde{k}(\tau)\, z(\tau)\ d\tau$$

$$\tilde{\mu}_n \longrightarrow E\left[q^2\right] - \int_a^b \tilde{k}(\tau)\, \rho(\tau)\, d\tau$$

where $\tilde{k}(\tau)$ is the solution of $\rho(t) = \int_a^b \phi_z\ (t,\tau)\, k(\tau)\, d\tau$.

This is seen to be the result expressed in (23) and (24) derived by the 'conventional' variational techniques. Of course, this passage to the limit is unjustified if the resulting integral equation does not have a solution. If the integral equation has a solution, and if it can be proved that this solution actually solves the problem of minimizing the expression in (22), than this passage to the limit is actually justified. For a discussion of the 'integral equation' method of solution, see for example Ref. 4.

An interesting problem which now arises is the following: given ϕ_z, ρ, $E\left[q^2\right]$, and a subdivision $\mathcal{S}_n$ of $\left[a,b\right]$, establish an upper bound for $\tilde{\mu}_n - \tilde{\mu}$. If a sampled-data estimate $\tilde{q}_n$ is used as an approximation of $\tilde{q}$, it would be interesting to know how far short this falls of the optimum estimate $\tilde{q}$. It would be desirable to obtain such a bound directly in terms of the properties of ϕ_z and ρ, without having to know explicit expressions in closed form for $\tilde{q}$ and $\tilde{\mu}$. This problem seems difficult to treat in general. An example of this type of result is the following:

<u>Theorem 3</u>: Let

$$\phi_z\ (s,t) = \beta e^{-\alpha\ |s-t|}$$

$$t_i^{(n)} = a + \frac{i(b-a)}{n}$$

and let $\rho(t)$ be at least three times differentiable and satisfy, for t in the finite interval $\left[a,b\right]$,

$$
\begin{aligned}
\left|\rho^2(t+h) - \rho^2(t)\right| &\leq K_1 |h| \\
\left|\rho(t+h)\,\dot{\rho}(t+h) - \rho(t)\,\dot{\rho}(t)\right| &\leq K_2 |h| \qquad (49) \\
\left|\rho(t+h)\,\ddot{\rho}(t+h) - \rho(t)\,\ddot{\rho}(t)\right| &\leq K_3 |h|
\end{aligned}
$$

Then

$$
\begin{aligned}
\tilde{\mu}_n - \tilde{\mu} \leq \left(\frac{b-a}{n}\right) \Bigg\{ & \frac{b-a}{2\beta\alpha}\left[\alpha^2 K_1 + K_3\right] \\
& + \frac{1}{\alpha\beta}\left[\alpha^2 K_1 + K_2\right] \qquad (50) \\
& + \frac{1}{4\alpha\beta}\Big[\alpha^2\left\{\rho^2(a) + \rho^2(b)\right\} \\
& + \rho(a)\,\ddot{\rho}(a) + \rho(b)\,\ddot{\rho}(b)\Big]\Bigg\} + o\left(\frac{b-a}{n}\right)
\end{aligned}
$$

It is easily seen that this is a rather unelegant result. An outline of the proof is as follows: For the particular case assumed in the theorem, it is possible to explicitly invert the matrix (ϕ_{zij}) as in Ref. 5, and thus get an explicit expression for

$$\tilde{\mu}_n = \sum_{i,j=1}^{n} \xi_{zij}\,\rho_i\,\rho_j$$

We then evaluate $\tilde{\mu} = \lim \tilde{\mu}_n$ directly as in Ref. 5, except retaining the first order terms in $\delta = \frac{b-a}{n}$. The result comes out as stated in the theorem. Unfortunately this method of proof is of very special applicability, in that in general one cannot expect to be able to invert (ϕ_{zij}); also a knowledge of $\tilde{\mu}$ in closed form is involved in the proof.

If the process $z(t)$ and the random variable q are jointly Gaussian, then $\tilde{q}$, the optimum linear estimate of q, is actually optimum among a much wider class of estimates. We define the class F of operations on $z(t)$ as follows:

<u>Definition</u>: F contains all random variables of the form

$$\text{l.i.m.} \left\{ f_{(n)} \left[z(t_1^{(n)}), \ldots, z(t_n^{(n)}) \right] \right\} \tag{51}$$

where $t_1^{(n)}, \ldots, t_n^{(n)}$ belong to $\left[a,b\right]$, and $f_{(n)}$ is a random variable measurable on the sample space of $z(t_1^{(n)}), \ldots, z(t_n^{(n)})$ and for which $E\left[f_{(n)}^2\right] < \infty$.

<u>Theorem 4</u>: Let the process $z(t)$ and the random variable q be jointly Gaussian. Also, as before, let $\tilde{q}$ be the optimum linear estimate of q and let $\tilde{\mu} = E\left[(\tilde{q} - q)^2\right]$. Then

$$\tilde{\mu} \leqq E\left[(\hat{q} - q)^2\right]$$

for all estimates $\hat{q}$ belonging to the class F.

<u>Proof</u>:

An estimate $\hat{q}$ belonging to F is of the form (51). Let us write this

$$\hat{q} = \text{l.i.m.}\ \hat{q}_n \tag{52}$$

where now

$$\hat{q}_n = f_{(n)} \left[z(t_1^{(n)}), \ldots, z(t_n^{(n)}) \right]$$

However, according to Ref. 6, pp. 561-562, if $z(t)$ and q are Gaussian,

$$E\left[(\tilde{q}_n - q)^2\right] \leqq E\left[(\hat{q}_n - q)^2\right] \tag{54}$$

where $\tilde{q}_n$ is the optimum <u>linear</u> estimate based on just the variables $z(t_1^{(n)})$, $\ldots$, $z(t_n^{(n)})$.

Now, since $\check{q} = \text{l.i.m.}\ \check{q}_n$, then $E[(\check{q} - q)^2] = \lim E[(\check{q}_n - q)^2]$.
But from (54), $\inf E[(\check{q}_n - q)^2] = \lim E[(\hat{q}_n - q)^2]$. Also
$\mu = E[(\check{q} - q)^2] \geq \inf E[(\check{q}_n - q)^2]$. This proves the theorem.

We can also state

<u>Theorem 5</u>: Let the process $z(t)$ and the random variable q be jointly Gaussian. Let $\{\pi_n\}$ be a sequence of subdivisions of the finite interval $[a,b]$ satisfying the conditions of Theorem 1 and, in addition, having the property that π_m is a refinement of π_n for $m > n$. Also, let $\check{q}_n$ be the optimum linear estimate of q based on the points of π_n. Then

$$\check{q} = \lim \check{q}_n$$

with probability one.

<u>Proof</u>: This follows directly from a theorem of Doob (Ref. 7, pp. 232-233).

APPENDIX

The Karhunen representation of processes which are continuous in the mean can be used to obtain q as follows:

We have, for z(t) continuous in the mean over $[a,b]$, (a, b both finite),

$$z(t) = \sum_{\nu=1}^{\infty} z_\nu \frac{\phi_\nu(t)}{\sqrt{\lambda_\nu}}, \quad t \in [a,b] \qquad (A1)$$

with convergence in the mean. The ϕ_ν and λ_ν are eigenfunctions and eigenvalues of the integral equation

$$\phi(t) = \lambda \int_a^b \phi_z(s,t)\, \phi(s)\, ds \qquad (A2)$$

The random variables z_ν are given by

$$z_\nu = \sqrt{\lambda_\nu} \int_a^b z(t)\, \phi_\nu(t)\, dt \qquad (A3)$$

Also

$$E(z_\nu) = 0 \qquad (A4)$$

$$E(z_\mu z_\nu) = \delta_{\mu\nu}$$

We now let

$$\hat{q} = \sum_{\nu=1}^{\infty} \alpha_\nu z_\nu \quad \text{with} \quad \sum_{1}^{\infty} \alpha_\nu^2 < \infty \qquad (A5)$$

and determine α_ν so that $E\,(q - \hat{q})^2$ is minimized. The resulting constants will be denoted by $\{\tilde{\alpha}_\nu\}$

Now,

$$E\left[(\hat{q} - q)^2\right] = E\left[\left(\sum \alpha_\nu z_\nu - q\right)^2\right]$$

$$= \sum \alpha_\nu{}^2 + E\left[q^2\right] - 2\sum \alpha_\nu E\left[q z_\nu\right]$$

Thus,

$$\tilde{\alpha}_\nu = E\left[q z_\nu\right] \qquad \text{(A6)}$$

Also,

$$\rho(t) = E\left[q z(t)\right] = \sum \frac{1}{\sqrt{\lambda_\nu}} E\left[q z_\nu\right] \phi_\nu(t)$$

so that

$$E\left[q z_\nu\right] = \sqrt{\lambda_\nu} \int_a^b \rho(t)\, \phi_\nu(t)\, dt \qquad \text{(A7)}$$

Therefore

$$\tilde{q} = \sum_1^\infty \tilde{\alpha}_\nu z_\nu \quad \text{(convergence in mean)} \qquad \text{(A8)}$$

with

$$\tilde{\alpha}_\nu = \sqrt{\lambda_\nu} \int_a^b \rho(t)\, \phi_\nu(t)\, dt \qquad \text{(A9)}$$

and z_ν given by (A3). Also,

$$E\left[(\tilde{q} - q)^2\right] = \tilde{\mu} = E\left[q^2\right] - \sum_1^\infty \tilde{\alpha}_\nu^2 \qquad \text{(A10)}$$

Equations (A8) and (A9) could be written

$$\tilde{q} = \text{l.i.m.} \int_a^b z(t)\, \tilde{k}_n(t)\, dt \tag{A11}$$

where

$$\tilde{k}_n(t) = \sum_1^n \tilde{\gamma}_\nu \sqrt{\lambda_\nu}\, \phi_\nu(t) \tag{A12}$$

This is very similar to the result that would be obtained if one tried to 'solve' the integral Eq. (23) by the method of eigenfunction expansions. The resulting 'solution' for $\tilde{k}(t)$ would be

$$\tilde{k}(t) = \sum_1 \sqrt{\lambda_\nu}\, \tilde{\gamma}_\nu\, \phi_\nu(t) \tag{A13}$$

with $\tilde{\gamma}_\nu$ given by (A9). The trouble with this is of course that the sum on the right side of (A13) usually does not converge in any ordinary sense to a weighting function which corresponds to $\tilde{q}$. In other words, although $\tilde{q}$ is the limit of operations on $z(t)$ which correspond to the weighting functions $\tilde{k}_n$, the limiting operation $\tilde{q}$ does not correspond to any weighting function $\tilde{k}$ which is the limit of the $\tilde{k}_n$ in the ordinary sense.

REFERENCES

1. Grenander, U, 'Stochastic Processes and Statistical Inference,' Ark. f. Mat. Vol. 1, pp. 195-277, October 1950.

2. Zadeh, L. A. and J. R. Ragazzini, 'Optimum Filters for the Detection of Signals in Noise,' Proc I.R.E. Vol. 40, pp. 1223-1231, October 1952.

3. Blum, Marvin, 'An Extension of Minimum Mean Square Prediction Error Theory for Sampled Data,' I.R.E. Transactions on Information Theory, Vol. IT-2, No. 3, September 1956, pp. 176-184.

4. Dolph, C L. and M. A. Woodbury, 'On the Relation Between Green's Functions and Covariances of Certain Stochastic Processes and Its Application to Unbiased Linear Prediction,' Trans. Amer. Math. Soc. Vol. 72, pp. 519-550, May 1952.

5. Reich, E. and P. Swerling, 'The Detection of a Sine Wave in Gaussian Noise,' J. Appl. Phys. Vol. 24 pp. 289-296, March 1953.

6. Doob, J. L., Stochastic Processes, John Wiley and Sons, 1953.

7. Doob, J. L., 'The Elementary Gaussian Processes,' Annals of Math. Stat. Vol. 15, 1944.