J. Soc. Indust. Appl. Math.
Vol. 8, No. 1, March, 1960
Printed in U.S.A.

MEAN SQUARE TRUNCATION ERROR IN SERIES EXPANSIONS OF RANDOM FUNCTIONS[1]

J. L. BROWN, Jr.[2]

1. Introduction. The purpose of this note is to give a simple proof of a result obtained by Jordan [1] concerning the mean integral square error between a random function $x(t)$ and a truncated series representation of $x(t)$ in orthonormal functions. When the integral in the error expression is extended over a finite range of integration $a \leqq t \leqq b$, and the series is truncated after the Nth term, Jordan has shown that the optimum choice of orthonormal functions in the expansion (with respect to minimizing the mean square error) is a set of N eigenfunctions of a related integral equation. The result is of interest since this integral equation is identical to the equation involved in the Karhunen-Loève expansion [2, pp. 96–98], where the criterion is that of obtaining expansion coefficients which are uncorrelated.

The method used in this note will be a reduction of the problem to a classical problem in the theory of quadratic integral forms.

2. Statement of problem. Given a sequence of functions, $\{\phi_n(t)\}$, defined, continuous, and orthonormal on $a \leqq t \leqq b$, we may associate with each square integrable random function $x(t)$ a formal Fourier series of the form

$$x(t) \sim \textstyle\sum_{n=1}^{\infty} a_n \phi_n(t) \qquad (a \leqq t \leqq b), \tag{1}$$

where the $\{a_n\}$ are random variables given by

$$a_n = \int_a^b x(t)\phi_n(t)\, dt. \tag{2}$$

As a tractable measure of the error involved in using an N-term truncation of series (1) to represent $x(t)$, we define

$$\epsilon_N[\{\phi_n(t)\}] = E\left[\int_a^b \left\{x(t) - \sum_{n=1}^{N} a_n\phi_n(t)\right\}^2 dt\right], \tag{3}$$

where $E[\quad]$ denotes the mean or expectation of the quantity in brackets. For a given random function $x(t)$ with continuous autocorrelation function $R(s, t) = E[x(s)x(t)]$ $(a \leqq s, t \leqq b)$ and fixed N, the problem is to determine the functions $\phi_1(t), \cdots, \phi_N(t)$ so as to minimize $\epsilon_N[\{\phi_n(t)\}]$.

---

[1] Received by the editors April 7, 1959.

[2] Ordnance Research Laboratory, the Pennsylvania State University, University Park, Pennsylvania.

3. The theorem.

THEOREM. *The mean square error* (3) *is minimized if* $\phi_1(t), \cdots, \phi_N(t)$ *are taken as the orthonormalized eigenfunctions of the Fredholm integral equation,*

$$\lambda \int_a^b R(s, t)\phi(t)\, dt = \phi(s) \qquad (a \leqq s \leqq b) \tag{4}$$

corresponding to the eigenvalues $\lambda_1, \cdots, \lambda_N$, *respectively where the eigenvalues of* (4) *are numbered according to increasing magnitude; that is,* $0 < |\lambda_1| \leqq |\lambda_2| \leqq \cdots$.

Proof. Expanding the square in (3) and making use of the orthonormality of the $\{\phi_n(t)\}$, it follows that

$$\int_a^b \left\{x(t) - \sum_{n=1}^N a_n \phi_n(t)\right\}^2 dt = \int_a^b x^2(t)\, dt - \sum_{n=1}^N a_n^2. \tag{5}$$

From (3), and linearity of the expectation,

$$\epsilon_N \left[\left\{\phi_n(t)\right\}\right] = \int_a^b E\left[x^2(t)\right] dt - \sum_{n=1}^N E\left[a_n^2\right]. \tag{6}$$

Since the first term in (6) is independent of the choice of the $\{\phi_n(t)\}$, the error is clearly minimized when the nonnegative quantity

$$\textstyle\sum_{n=1}^N E[a_n^2]$$

is maximized.

But,

$$\begin{aligned} E[a_n^2] &= E\left[\int_a^b x(t)\phi_n(t)\, dt \int_a^b x(s)\phi_n(s)\, ds\right] \\ &= \int_a^b \int_a^b E\left[x(s)x(t)\right] \phi_n(s)\phi_n(t)\, ds\, dt \\ &= \int_a^b \int_a^b R(s, t)\phi_n(s)\phi_n(t)\, ds\, dt \end{aligned}$$

and thus the quantity to be maximized is

$$\sum_{n=1}^N \int_a^b \int_a^b R(s, t)\phi_n(s)\phi_n(t)\, ds\, dt \tag{7}$$

subject to $\int_a^b \phi_m(s)\phi_n(s)\, ds = \delta_{mn}$.

The solution to the problem of choosing $\phi(t)$ to maximize the quadratic integral form

$$J(\phi, \phi) = \int_a^b \int_a^b K(s, t)\phi(s)\phi(t)\, ds\, dt$$

where $K(s, t)$ is a given continuous symmetric kernel and $\int_a^b \phi^2(t)\, dt = 1$ is well-known [3, pp. 122–125]; the maximum value of $J(\phi, \phi)$ is $1/\lambda_1$, where λ_1 is the eigenvalue of smallest magnitude of the Fredholm integral equation

$$\lambda \int_a^b K(s, t)\phi(t)\, dt = \phi(s) ,$$

and the function for which this maximum is attained is the corresponding normalized eigenfunction, $\phi_1(t)$.

In our case, $R(s, t) = R(t, s)$ by a basic property of the autocorrelation function; additionally, $R(s, t)$ is a positive kernel [4, p. 124] and thus all the eigenvalues are positive. [Note that $R(s, t)$ need not be positive definite.]

If the eigenvalues of (4) (not necessarily distinct) are numbered in nondecreasing order so that

$$0 < \lambda_1 \leqq \lambda_2 \leqq \lambda_3 \leqq \cdots , \tag{8}$$

then it is clear that (7) will be maximized by choosing $\phi_1(t), \cdots, \phi_N(t)$ as N orthonormalized solutions of (4) corresponding to eigenvalues $\lambda_1, \cdots, \lambda_N$ respectively. When these eigenvalues are distinct, a unique solution is obtained. If, however, λ_{n_0} is an eigenvalue of index k, the k corresponding orthonormal eigenfunctions are not uniquely determined since there is an inherent nonuniqueness in the Schmidt process of orthonormalization. It is easily shown that if $\psi_1(t), \cdots, \psi_k(t)$ and $\phi_1(t), \cdots, \phi_k(t)$ are two different orthonormal sets of eigenfunctions for λ_{n_0}, then there exists a $k \times k$ matrix $A = \{a_{ij}\}$ such that

$$\phi_i(t) = \textstyle\sum_{j=1}^k a_{ij}\psi_j(t)$$

and

$$\textstyle\sum_{j=1}^k a_{ij}a_{mj} = \delta_{im} ;$$

that is, any two competing sets of orthonormal solutions for a given eigenvalue are related by an orthogonal transformation.

This proves the theorem and the minimum mean integral square truncation error is given explicitly as

$$\epsilon_N[\{\phi_n(t)\}]_{\min} = \int_a^b E[x^2(t)]dt - \sum_{n=1}^N \frac{1}{\lambda_n} , \tag{9}$$

where $\{\lambda_n\}$ are the eigenvalues of (4) arranged in nondecreasing order.

4. Conclusion. The above reduction of the problem to the application of a classical result provides a considerable shortening of Jordan's original proof. It also avoids the necessity of assuming that the eigenfunctions of (4) form a complete set, although this will usually be the case in practice as noted by Jordan.

Since the integral equation (4), which generates the optimum orthonormal functions, is the same as the one involved in the Karhunen-Loève expansion, the theorem of this note may be restated in the following form: *the Karhunen-Loève expansion*

$$x(t) \sim \sum_{n=1}^{\infty} a_n \phi_n(t),$$

[*where the* a_n *are given by* (2) *and the* $\{\phi_n(t)\}$ *are orthonormal solutions of* (4) *corresponding to eigenvalues* $\{\lambda_n\}$ *ordered as in* (8)] *is an expansion for which the minimum integral mean square truncation error is achieved for any fixed choice of* N. In addition, by the result of Karhunen and Loève, the coefficients $\{a_n\}$ form a set of uncorrelated random variables.

Since the original writing of this paper, Jordan [5] has presented an alternate proof which makes use of a theorem in functional analysis. It is believed that the proof given in the present note has the advantage of being more direct than either of Jordan's proofs.

The author is indebted to the referee for calling attention to [6] where a more general version of this theorem is formulated in an abstract setting. In [6], it is stated that the minimum error is given by the series

$$\sum_{n=N+1}^{\infty} \frac{1}{\lambda_n}.$$

That this is equivalent to equation (9) can be shown by applying Mercer's Theorem ([2], p. 374) to the kernel $R(s, t)$. Term-by-term integration of the resulting identity

$$R(s, s) = \sum_{n=1}^{\infty} \frac{\phi_n^{\,2}(s)}{\lambda_n}$$

then yields

$$\int_a^b R(s, s)\, ds = \int_a^b E[x^2(s)]\, ds = \sum_{n=1}^{\infty} \frac{1}{\lambda_n}$$

or

$$\int_a^b E[x^2(s)]\, ds - \sum_{n=1}^{N} \frac{1}{\lambda_n} = \sum_{n=N+1}^{\infty} \frac{1}{\lambda_n}$$

as required. In practice, of course, equation (9) is preferable since only the first N eigenvalues are involved.

REFERENCES

1. K. L. JORDAN, JR., *Minimization of truncation error in series expansions of random processes*, Quarterly Progress Report, Research Laboratory of Electronics, Mass. Inst. of Technology, No. 52 (1959), pp. 70–78.
2. W. L. ROOT AND W. B. DAVENPORT, JR., *An Introduction to the Theory of Random Signals and Noise*, McGraw-Hill, New York, 1958.
3. R. COURANT AND D. HILBERT, *Methods of Mathematical Physics*, Vol. I, Interscience, New York, 1953.
4. F. G. TRICOMI, *Integral Equations*, Interscience, New York, 1957.
5. K. L. JORDAN, JR., *Minimization of truncation error in series expansions of random processes*, Quarterly Progress Report, Research Laboratory of Electronics, Mass. Inst. of Technology, No. 53 (1959), pp. 84–86.
6. H. P. KRAMER, *Best finite linear approximation to a second order stochastic process* (Abstract), Amer. Math. Soc. Notices, 5 (1958), Abstract No. 542–121, p. 201.