The Annals of Probability
1977, Vol. 5, No. 6, 1047-1051

BOOK REVIEW

H. DYM AND H. P. MCKEAN, *Gaussian Processes, Function Theory and the Inverse Spectral Problem*. Prob. and Math. Statist. **31** Academic Press, New York, San Francisco, London, 1976, xi+333 pp., $35.00.

Review by MICHAEL B. MARCUS
Northwestern University

Let X, Y be two normal random variables with mean zero, variance 1, and covariance ρ. Given the information $X = c$, how can it be used to improve our knowledge of Y? This problem is easily solved. $Z = Y - \rho X$ is a normal random variable with mean zero and variance $1 - \rho^2$ and is independent of X. Therefore

$$P[a < Y < b \mid X = c] = P[a < Z + \rho c < b]$$
$$= (2\pi(1 - \rho^2))^{-\frac{1}{2}} \int_a^b \exp\left\{-\frac{(u - \rho c)^2}{2(1 - \rho^2)}\right\} du \,.$$

Note that whatever the value c, $EZ^2 = 1 - \rho^2 \leqq EY^2$ so that given X the distribution of Y becomes more concentrated depending upon the degree of correlation ρ. We note also that $\rho X = E(Y \mid X)$.

Similarly let $X_1, \cdots, X_n$ be normal random variables with mean zero and covariance Γ_{ij}, $i, j = 1, \cdots, n$. The distribution of X_n given $X_1, \cdots, X_{n-1}$ is normal with mean $m = E(X_n \mid X_1, \cdots, X_{n-1})$ and variance $E((X_n - m)^2 \mid X_1, \cdots, X_{n-1}) = E(X_n - m)^2$. To see this consider the complex Hilbert space generated by sums $\eta = c_1X_1 + \cdots + c_nX_n$ with norm $||\eta|| = (E|\eta|^2)^{\frac{1}{2}}$. Because of the equivalence between orthogonality of zero mean Gaussian random variables with respect to this norm and their statistical independence, m is the orthogonal projection of X_n onto the subspace spanned by $X_1, \cdots, X_{n-1}$; and $X_n - m$, being orthogonal to this subspace, is independent of the Borel field generated by $X_1, \cdots, X_{n-1}$. Utilizing the observations $X_1, \cdots, X_{n-1}$ to obtain the probability distribution of X_n is what is meant by predicting X_n given $X_1, \cdots, X_{n-1}$.

When the number of observations involved is finite (and the covariance function is known) the prediction problem is easily solved, but for an infinite number of observations it is more difficult. The earliest work was done independently by Kolmogorov in 1939 and 1941 (references not included here can be found in Dym and McKean's book) and by Wiener in 1942. Wiener's 1942 report, the so-called "yellow peril," was classified as a military secret and was finally published openly as *Extrapolation, Interpolation and Smoothing of Stationary Time Series* in 1949. Wiener was unaware of Kolmogorov's work. The problem that motivated Wiener [3] and perhaps also Kolmogorov was that of automatic fire control for anti-aircraft batteries. Roughly speaking, one observes some parameter, say $X(t)$, of an airplane, such as its distance from an automatic gun, for time $-\tau \leqq t \leqq 0$ and then predicts $X(T)$ for a fixed $T > 0$. The a priori assumption is that the possible paths $X(T)$, after suitable normalization, form a zero-mean stationary Gaussian process with known covariance function. From a practical standpoint this idealization seems to leave a great deal to be desired; nevertheless, even as a theoretical problem it is very difficult. The problem initially solved, in the discrete case by Kolmogorov and in the continuous case by Wiener, was that of predicting $X(T)$, $T > 0$ when the whole past, $t \leqq 0$, is given. The prediction problem given only a finite segment of the past was solved by M. G. Krein in 1954, and this work also was unknown for many years by workers in the United States.

Let $X(t)$ be a real valued, zero-mean, stationary Gaussian process. As above, predicting $X(T)$ for $T \geqq 0$, given the past, i.e., $X(t)$, $-\tau \leqq t \leqq 0$ (where possibly $\tau = \infty$) means obtaining the projection m of $X(T)$ onto the closed subspace spanned by finite linear combinations of $X(t)$, $-\tau \leqq t \leqq 0$, relative to the norm $(E|\ |^2)^{\frac{1}{2}}$. The mean square error is $E(X(T) - m)^2$. This problem is further transformed as follows: associated with the Gaussian process is a covariance function R (always assumed known) and a unique symmetric measure Δ corresponding to R called the spectral function of the process. The relationship is given by Bochner's theorem

$$R(t_1 - t_2) = E(X(t_1)X(t_2)) = \int_{-\infty}^{\infty} e^{i\gamma t_1} e^{-i\gamma t_2}\, d\Delta(\gamma) .$$

Consequently the correspondence $e^{i\gamma t} \leftrightarrow X(t)$ is an isometric isomorphism between two Hilbert spaces, the span of $X(t)$, $-\infty < t < \infty$, relative to the norm $(E|\ |^2)^{\frac{1}{2}}$ and the space $Z(\Delta) = L^2(R, d\Delta)$. The problem of predicting $X(T)$ given the whole past is equivalent to finding the projection of $e^{i\gamma T}$ onto Z^-, the span of $e^{i\gamma t}$, $t \leqq 0$ in $Z(\Delta)$.

Five variations of the prediction problems are considered in this book:

(1) predicting $X(T)$, $T > 0$, given $X(t)$, $t \leqq 0$;
(2) predicting $X(T)$, $T > 0$, given $X(t)$, $-\tau \leqq t \leqq 0$;
(3) the interpolation problem of "predicting" $X(t)$ for $|t| < T$ from $X(t)$ for $|t| \geqq T$;
(4) the degree of dependence of the future upon the past as reflected in the projection of $X(t)$, $t \geqq 0$ upon $X(t)$, $t \leqq 0$; and
(5) the degree of mixing of the process as reflected in the projection of $X(t)$, $t \geqq T$, upon $X(t)$, $t \leqq 0$, as T goes to infinity.

Like problem (1), problems (2)—(5) also have obvious analogues in terms of projections of certain elements in subspaces of $Z(\Delta)$ onto other subspaces of $Z(\Delta)$. It is in this way that the solutions are obtained, followed by an inverse isomorphism to the stochastic formulation. Problems (1) and (2) are the problems of Kolmogorov, Wiener and Krein. Problem (3) receives its first full solution in this book; (4) has been worked on by Levinson and McKean and others, and new results are given here; (5) is also the subject of current activity.

One of the pleasing characteristics of this exceptional book is the clarity with which the authors explain what it is they are doing. The preface and introduction combined set out the program of the book in a cogent and stimulating fashion. The reader will see that many of my remarks are taken from this material, in which a much more extensive list of credits is given. The book is essentially divided into two parts according to the distinct mathematical approaches needed to study the prediction problem with infinite or with limited memory. When the entire past is known the problem is solved employing the Hardy space H^{2+}. This is the space of functions h analytic in the open upper half plane with norm $\|h\| = \sup_{b>0} (\int |h(a+ib)|^2 \, da)^{\frac{1}{2}} < \infty$. The solution depends upon being able to factor $\Delta'(\gamma) = |h(\gamma)|^2$, where $\Delta'(\gamma)$ is the derivative of $\Delta(\gamma)$ and h is an outer function of class H^{2+}. This can be accomplished when the prediction is imperfect, i.e., $e^{i\gamma T}$ is not contained in the infinite past, or equivalently when $\int (1+\gamma^2)^{-1} \log \Delta'(\gamma)\, d\gamma > -\infty$. This is classical material much of which appears elsewhere (e.g., [1], [3]), although not in such a thorough or illuminating fashion.

As the authors show, there are many interesting ramifications of this problem. For example, given a formula for the inverse isomorphism of the projection of $e^{i\gamma T}$ on Z^- onto the probability space, how can this be related to the information at hand, namely, $X(t)$, $t \leqq 0$? Pursuing this question leads to the consideration of Wiener filters. Further simplifications appear when Δ' is a rational function. The apparatus of Hardy functions is also used in pursuing problems (4) and (5).

The major portion of the book is devoted to material that has not been systematically treated elsewhere; the more "realistic" problem of prediction when only a finite portion of the past is known. In this case one computes the projection of $e^{i\gamma t}$, $t > T$, onto the subspace Z^T, the span of $e^{i\gamma t}$, $|t| \leqq T$, in $Z(\Delta) = L^2(R^1, d\Delta)$. Using the Paley–Wiener theorem one can obtain this projection for functions $f \in L^2(R^1, d\gamma)$, where γ is Lebesgue measure. Krein generalized this result to $Z(\Delta)$ where Δ is an odd nondecreasing function satisfying $\int (\gamma^2 + 1)^{-1} d\Delta(\gamma) < \infty$. If $\int (\gamma^2+1)^{-1} \log \Delta'(\gamma)\, d\gamma > -\infty$, then Z^T is a proper subspace of $Z(\Delta)$, identifiable as the span of integral functions $f \in Z$ of exponential type less than T.

Let Δ be the spectral function of a stationary Gaussian process. This function can be uniquely identified as the principal spectral function of a string lmk. A string is an interval $0 \leqq x \leqq l$ of length $l \leqq \infty$ equipped with a "mass function" $m(x)$ and a "tying constant" k which is part of a boundary condition on the oscillation u of the string at l. The oscillations of the string are given by the wave equation

$$\frac{\partial^2 u}{\partial t^2} = \frac{\partial^2 u}{\partial m\, \partial x},$$

subject to certain boundary conditions. The differential operator $\mathscr{G} = d^2/dm\, dx$, also subject to boundary conditions involving both l and k, on its domain

$D(\mathscr{S}) \subset L^2([0, l], dm) \equiv M$, is nonpositive and self-adjoint. Consequently one can compute eigenvibrations of the string: $\mathscr{S}f_\gamma(x) = -\gamma^2 f_\gamma(x)$. Let $A(x, \gamma) = f_\gamma(x)$. The Green operator $G_\omega = (-\omega^2 - \mathscr{S})^{-1}$ exists for $-\omega^2$ outside $(-\infty, 0]$ and can be identified as the kernel of an integral operator on M. This kernel has the representation

$$G_\omega(x, y) = \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{A(x, \gamma)A(y, \gamma)}{\gamma^2 - \omega^2} \, d\Delta(\gamma)$$

for appropriate ω and for a unique measure Δ which is called the principal spectral function of $\mathscr{S}$. In this way the string is associated with a Gaussian process which has spectrum Δ.

When $\int (\gamma^2 + 1)^{-1} \log \Delta'(\gamma) \, d\gamma > -\infty$ and $f \in M$, the projection of f onto Z^T is given by a relationship "in perfect analogy with the classical recipe" of Paley and Wiener. Furthermore $\int_{0-}^{l} A(x, \gamma) f(x) \, dm(x)$ is a one-one map of M onto even functions in $Z(\Delta)$ and there exist analogous equations for the projection of $e^{i\gamma T}$ onto Z^T. This is the desired result.

These remarks are intended to suggest, however vaguely, how a stationary Gaussian process is associated with a string. The result requires a great deal of interesting mathematics and Dym and McKean present it with clarity and style. Their treatment includes a wealth of related results. A variation of the above argument is used to present the first full solution of the interpolation problem (3), for which a new class of strings is introduced. The method of strings is also used to give a solution of the Stieltjes moment problem.

The book consists of six chapters. The first three are devoted to background. Topics relating to entire functions are given in Chapter 1. The next chapter deals with Hardy spaces. The third chapter is probabilistic, covering some elements of the theory of stationary Gaussian processes, metric transitivity, degree of independence, and mutual information. Most of the results in these three chapters are proven, although, most likely, in too brief a fashion for a person learning this material for the first time. Nevertheless these chapters provide a valuable outline of the prerequisites for understanding the rest of this book. They comprise a quarter of the book.

Chapter Four treats the classical prediction problem and its many ramifications. Chapter Five explores the relationship between strings, their principal spectral functions, and the symmetric measures that characterize stationary Gaussian processes. In Chapter Six de Branges spaces are used to prove the uniqueness of the principal spectral function associated with certain strings. Finally the prediction problem with limited memory is solved as in the interpolation problem. There are many interesting exercises throughout the book.

There are still open problems in this area. Perhaps the most interesting one the authors describe concerns the lack of probabilistic content in the results obtained. The "prediction" is expressed as an inverse isomorphism of a certain projection. The form in which it is given seems to have no meaning probabilistically. In fact all the analysis in this book very quickly leaves the more intuitive areas of probability far behind, and it would be very useful if someone would give a more stochastic description of what is going on. This, most likely, is not an easy task. In other areas of harmonic analysis involving random functions it is not possible, at least at present, to explain analytical results in a way that makes sense probabilistically.

Who should read this book? Of course anyone interested in the question of prediction. But the book deserves a wider audience than this. It is really a very exciting work. Any analyst should enjoy seeing so much interesting mathematics of diverse origins brought to bear upon a difficult problem. The material in this book represents classical analysis at its best. However, it is not a book for browsing. It is the authors' style to include definitions and notations casually in the general discourse and it is difficult to understand the statement of a theorem without carefully reading much of the preceding material.

Wiener's student, Norman Levinson, helped explain his "yellow peril" and also worked on the germ and gap problem with McKean. This book, written by Henry McKean and his student Harry Dym, is dedicated to the memory of Professor Norman Levinson. It is a fitting tribute.

REFERENCES

[1] GRENANDER, U. and SZEGÖ, G. (1958). *Toeplitz Forms and Their Applications*. Univ. of Calif. Press.

[2] ROZANOV, YU. A. (1967). *Stationary random processes*. Holden-Day, San Francisco.

[3] WIENER, N. (1956). *I am a Mathematician*. M.I.T. Press, Cambridge.

DEPARTMENT OF MATHEMATICS
NORTHWESTERN UNIVERSITY
EVANSTON, ILLINOIS 60201