## BOOK REVIEWS

*Toeplitz forms and their applications.* By Ulf Grenander and Gabor Szegö. California Monographs in Mathematical Sciences, University of California Press, 1958. 7+245 pp. $6.00.

This book owes its timeliness, and much of its importance and unique charm to one particular quality which sets it apart from other research monographs. Its two authors have accomplished a successful synthesis of two important mathematical developments. One of these is the theory of Toeplitz forms, the other, more recent one, the theory of (wide sense) stationary stochastic processes.

The theory of Toeplitz forms has its roots in the work of Toeplitz, Féjèr, Carathéodory, F. Riesz on trigonometric series and harmonic functions. In two important papers (Math. Zeitschrift, 1920) G. Szegö unified and extended much of their work by creating a theory the central ideas and results of which also form the core of the present book. In other words, concepts and methods created forty years ago have now gained new interest as the analytical techniques in a branch of mathematics (prediction and estimation theory for stationary stochastic processes) which did not then exist. This remark could not have been made in 1939, when a brief account of the theory of Toeplitz forms first appeared in Szegö's book *Orthogonal polynomials*, but in any event the present treatment goes further and deeper. We mention in particular the new topics of the trigonometric moment problem (Chapter 4) and a chapter on applications to analytic functions (Chapter 5). To avoid confusion, it should be said at once that the book is divided into two parts; Part I (Chapters 1 to 8) deals with the theory of Toeplitz forms, and Part II (Chapters 9, 10, 11) is devoted to probability and statistics, with the exception of the above mentioned Chapter 9.

In the theory of Toeplitz forms the Fourier Stieltjes coefficients $c_k$ of a distribution function $\alpha(\theta)$ on $[0, 2\pi]$ are used to define the Toeplitz matrices $T_n=(c_{j-i})$, $i, j=0, 1, \cdots, n$. Since $\alpha(\theta)$ is real and nondecreasing, the quadratic forms associated with these matrices are Hermitean positive definite (this restriction, and its regrettable consequences from the point of view of certain applications will form the subject of later remarks). The principal problems of the theory are

(a) *The minimum problem*, which in its simplest form asks for the minimum $\mu_n$ of the quadratic form $uT_nu$ where the vector $u=(u_0, u_1, \cdots, u_n)$ is subject to the restriction $u_0=1$. The answer to this famous problem (in Chapter 3) is that the sequence $\mu_n$ converges to the geometric mean of the (almost everywhere existing) derivative of $\alpha(\theta)$.

(b) *The eigenvalue problem*, which seeks to describe the distribution of the eigenvalues of $T_n$ for large $n$. The simplest and best known, but by no means the sharpest of the answers obtained in Chapter 5, is that the fraction of eigenvalues located in an interval tends to the $\alpha$-measure of the interval.

Both problems, as well as extensions of them, which motivate Chapter 6 and parts of Chapters 7 and 8, are answered within the framework of a unified and extremely elegant theory. It turns out that, if the components of the minimizing vector $u$ in problem (a) are taken as the coefficients of a trigonometric polynomial $\phi_n(z)$, $z=e^{i\theta}$, these polynomials are orthogonal with respect to the weight function $d\alpha(\theta)$. Chapter 2 is devoted to the elementary theory of such polynomials which are orthogonal on the unit circle, and Chapter 3 contains very deep results concerning their asymptotic behavior as $n$ tends to infinity. This is the most important part of the book, as the degree of refinement of the solutions to problems (a) and (b) and, as the reader will find, also of the statistical theory in the second part of the book, stands in direct proportion to the extent of one's knowledge about the polynomials $\phi_n(z)$.

With the exception of a few pages Chapter 10 is an exposition of the theory of least squares prediction for stationary (both discrete and continuous time) stochastic processes. A misleadingly simple but true statement of the connection between the prediction problem and the theory of Toeplitz forms is that the prediction problem is equivalent to our earlier problem (a), where the matrix $T_n$ has to be replaced by the matrix $R_n=(r_{ij})$, $i, j=0, 1, \cdots, n$; $r_{ij}=r_{|i-j|}$ being the covariance function of the process. Actually this reformulation of the problem requires deep but well known results from functional analysis, and already in Doob's book *Stochastic processes* the prediction problem is formulated in the way described. The simplicity and lucidity of the present treatment, which goes beyond Doob's in a few respects, (for example, the predictor is given both in terms of the moving average and spectral representations of the process), is due to Grenander's observation, that the prediction problem had been solved by Szegö in 1920. In other words, for prediction one unit time ahead, the predictor turns out to be that linear function of the last $n$ observations, whose coefficients are essentially those of the orthogonal polynomial $\phi_n(z)$. The predictor for an infinite past, as well as the mean square error of the prediction, both emerge as corollaries of the theory of Toeplitz forms. Helson and Lowdenslager have recently

(Acta Math., 1958) extended the solution of problem (a), and thereby parts of prediction theory from one dimension to two.

Chapter 11 is devoted to a very condensed account of some topics in the modern theory of statistical inference for stationary processes. Much of the pioneer work in this field is Grenander's own (Arkiv för Matematik 1950 and 1952). A particularly elegant application of Toeplitz forms is the derivation of the asymptotic mean square error of the best linear estimate of the mean value of a process, in §11.3. Another area of application which is not quite of the simple type: autocorrelation matrix = Toeplitz matrix, is the theory of the distribution of quadratic forms of normal variates (§11.5), which has considerable importance as a means for estimating the spectral density of a process. It seems that some of the theoretical developments in Chapters 7 and 8 were inspired by the practical requirements of this statistical problem. However, the reader looking for additional motivation in §11.5 will be severely disappointed by the terse remark that "there are reasons why this function (the estimate of the spectral density) should be a quadratic form." Grenander and Rosenblatt's 1957 book *Statistical analysis of stationary time series* complements the theory on this and other points in Chapter 11.

The reviewer feels quite critical vis-à-vis one single aspect of this book. We refer specifically to §10.16, where a certain random walk problem is discussed, "in order to give the reader an idea of how and why Toeplitz forms are useful for certain other probability problems." Only a few years ago the authors could not have written this passage, as it is based on recent work of Kac (Duke Math. J., 1954). But by now, in 1958, enough evidence has mounted up to show that not only can the problem treated in §10.16 be carried much farther than it has been, but also that it demands a careful reinterpretation and extension of some of the principal theorems in the theory of Toeplitz forms.

Following Kac, the authors discuss the stochastic process $S_n = X_1 + \cdots + X_n$, where the $X_i$ are identically distributed, independent lattice random variables. It is then clear that if $c_k = \Pr[X_i = k]$, the matrices $T_n = (c_{j-i})$, $i, j = 0, 1, \cdots, n$, are of the Toeplitz type. The theory of Toeplitz forms can now be, and is invoked (by an argument the real point of which is to substantiate remark (i) below) to derive a simple expression for the conditional expectation

$$E[M_n | S_n = 0]$$

where the random variables $M_n$ are defined by $M_n = \max(0, S_1, \cdots, S_n)$. But the method used was never designed to obtain such results, and we shall indicate how one can now do much better.

It is easy to verify that the probability Pr $[M_n \leqq k]$ is the limit, as $m \to \infty$, of the sum of the $k$th row of the matrix $T_m$ raised to the $n$th power. But as the $n$th power of a finite Toeplitz matrix is not a Toeplitz matrix, one must take generating functions instead, so that Pr $[M_n \leqq k]$ is the coefficient of $\lambda^n$ in $\lim_{m\to\infty} \sum_{l=0}^{m} [I-\lambda T_m]_{kl}^{-1}$. As $I-\lambda T_m$ is a Toeplitz matrix, this explains why *a class of probability problems can be reduced to the problem of inverting a Toeplitz matrix.*

Indeed the theory of Szegö leads to a very simple inversion method for finite Toeplitz matrices, which differs from, and is in some respects simpler than the standard method based on the spectral theorem for Hermitean matrices. Let $T_n$ be a sequence of Toeplitz matrices whose associated orthonormal polynomials are $\phi_n(z)$. It is very simple to derive from equation (7) in §2.2 the result that the $(k, j)$th element of the inverse of $T_n$ is the coefficient of $z^k\bar{a}^j$ in $\sum_{\nu=0}^{n} \phi_\nu(z)\overline{\phi_\nu(a)}$. The usefulness of this result is directly due to the theorems describing the asymptotic behavior of the polynomials $\phi_n(z)$ in §3.5. Indeed it turns out that the reviewer's theory giving the distribution of $M_n$ for arbitrary random variables (Trans. Amer. Math. Soc., 1956) when specialized to symmetric lattice random variables, is obtainable as a simple corollary of theorems in Chapter 3.

The problem discussed strongly suggests the desirability, and even the possibility, of extending parts of Szegö's theory to nonsymmetric Toeplitz matrices. As the above argument uses the method of hindsight, and could not possibly have been included in the book, we conclude with two more items of evidence to the same effect which could, and perhaps should have been included.

(i) The paper of Kac, quoted above, leads to a very startling and profound conclusion, quite different from the less interesting analytical details discussed in the book. It demonstrates that some of the deepest theorems concerning the distribution of eigenvalues (problem (b) above), can be interpreted as theorems in probability, and can be proved by methods outside the theory of Toeplitz forms which do not depend on the symmetry of the Toeplitz matrices involved. As a particular example we mention equation 11 in §5.2, which is one of the basic results in the equidistribution theory. It remains valid for all sequences of Toeplitz matrices $T_n$ (Hermitean or not) with the property that $\sum_{-\infty}^{\infty} |c_k| < \infty$.

(ii) The most famous result on inversion of Toeplitz matrices is Wiener's Tauberian theorem for trigonometric series. If viewed in the proper light, it gives necessary and sufficient conditions for the existence of an inverse of a doubly infinite Toeplitz matrix $T=(c_{i-j})$, $i, j=0, \pm 1, \pm 2, \cdots$, again with $\sum_{-\infty}^{\infty} |c_k| < \infty$, viewed as an operator on the space of bounded sequences. This theorem again has nothing to do with symmetry, and that seems important as we have seen that the nonsymmetric inversion problem arises naturally in probability theory. Unfortunately the analogous problem for one-sided infinite Toeplitz matrices was not yet solved when this book was written.[1]

The book is authoritatively documented by means of an appendix of 10 pages, providing references as well as remarks which clarify the mathematical or historical setting of important ideas in the text. The mathematical presentation is of the same high caliber as in Szegö's *Orthogonal polynomials*, but even more elegant because the subject matter here is so much more unified. Most of the background theory required in the book is relegated to an introductory chapter. Therefore there are no digressions from the natural development of the theory, and this has enabled the authors to write in a terse but at the same time pleasingly informal style.

Not only good students but also serious research workers may find this book difficult if they want to fully bridge the conceptual gap between the two fields which are unified here. But as the book offers so much more than would two separate monographs in the corresponding subjects of analysis and probability, this is precisely the challenge it offers to the reader. The content of the book is evidence enough that this challenge will contribute to the growth of mathematics.

F. SPITZER

*Contributions to the theory of games*, vol. III. Ed. by M. Dresher, A. W. Tucker and P. Wolfe. Annals of Mathematics Study, no. 39. Princeton, Princeton University Press, 1957. 8+435 pp. $5.00.

This is the third volume in a series on the theory of games, a series which can teach an interesting lesson in mathematical publication applicable to other branches of mathematics. The present volume, as the preceding volumes, is made up of a number (twenty-three in this instance) of individual papers on the theory of games grouped into five general classifications. The volume is prefaced by an introduction which briefly explains this classification and then gives brief summaries of the individual papers. The would-be reader can decide from these summaries what papers he wishes to read. No one not interested in game theory need enter these portals and waste his time.

[1] *Added in proof.* The continuous analogue of this problem is famous under the name of the Wiener-Hopf equation. It was recently solved by M. Krein (Uspehi Mat. Nauk, 1958).