BOOK REVIEWS

Correspondence concerning reviews should be addressed to the Book Review Editor, Professor William Kruskal, Department of Statistics, University of Chicago, Chicago 37, Illinois.

P. WHITTLE, *Prediction and Regulation by Linear Least-squares Methods.* English Universities Press, London, and Van Nostrand, Princeton, 1963. 25s or $3.75 x + 147 pp.

Review by FREDERICK J. BEUTLER[1]

University of Michigan

The theory of linear least square error prediction (l.l.s.p.) for wide-sense stationary sequences and processes was developed, almost simultaneously, by Wiener and Kolmogorov. The former emphasized the role of spectral factorization, while the latter placed l.l.s.p. in its natural context of Hilbert spaces and functions of Hardy classes. With this beginning, l.l.s.p. achieved maturity almost at its very inception.

Those wishing to master the theory may turn to the books of Doob [2] or Grenander and Rosenblatt [3], both of which present l.l.s.p. in rigorous fashion. These references, however, lack examples, contain little material on combined filtering and prediction, and make no concessions to weaknesses in the reader's mathematical training. Consequently, many people interested principally in applications of l.l.s.p. would prefer to avail themselves of other sources.

The engineer, in particular, has recourse to many texts purporting to treat systems with random inputs. More often than not, he will find that the analysis consists of a series of implausible computational steps that finally yield—mirabile dictu—Wiener's formula (see e.g. [4]).

Between the two approaches mentioned above there lies a middle ground that has been incompletely exploited. If certain assumptions are made on the spectra of the processes involved, much of l.l.s.p. and filtering theory can be presented satisfactorily to those of modest mathematical attainment. An expository treatment along these lines, supplemented with more recent contributions, and illustrated by examples drawn from a diversity of fields, should find a large and interested audience.

Whittle's book is one of few (cf. [6]) that attempts to meet this need, and more. The classical l.l.s.p. and filtering theory is presented, of course, but there is also material on multivariate prediction, interpolation, unbiased l.l.s.p. of processes containing deterministic components with unknown parameters, l.l.s.p. for some non-stationary processes, and finally, linear least square error regulation (feedback control).

The text—for such it is meant to be—is a sound and sober account of those facts on the above topics that can be presented at an intermediate level, and

[1] Now a Visiting Scholar, University of California, Berkeley.

that lead to applications in technology, economics, geology, population statistics, etc. The derivations are leavened with commentary that strengthens intuitive insight, relates apparently disparate parts of the theory, connects theory with practice, and furnishes historical background. There are many well chosen examples, and a large number of nontrivial exercises; the scope of the latter reaches from extensions of the theory to a determination of explicit formulas useful for computation.

Nevertheless, the book is seriously flawed. The author has ignored the existence of most of the recent l.l.s.p. literature (see [7] for a 1960–1963 review), some of which improves on the text material both in generality and and felicity of presentation. Continuous parameter processes are often sketchily and unlovingly treated by reference to an analogy with the corresponding result for sequences. The notation is confusing, especially on distinctions between functions on the unit circle and in the complex plane, and on the factorization of spectral densities of processes with complex covariance. There are many typographical errors that impede the progress of the careful reader in following Whittle's sometimes fragmentary arguments. The normalization of orthogonal random variables (generally denoted by $\{\epsilon_n\}$) is inconsistent, and must occasionally be deduced from the context. Last, but far from least, Whittle derives functional equations for the l.l.s.p. by ad hoc methods (cf. Equations 3.4.4, 4.1.5, 4.3.8, 5.1.3, 6.1.4, 7.4.2, 7.5.2), instead of invoking the orthogonality condition as a basic unifying notion. (Briefly, this condition may be stated as follows: If M is a subspace of a Hilbert space, $\inf_{x \varepsilon M} \|y - x\|$ is uniquely attained by that $x_0 \varepsilon M$ which renders $y - x_0$ orthogonal to every $x \varepsilon M$.) Indeed, the latter appears only as an exercise limited to proving necessity in a narrow context.

A summary of the contents of Whittle's book follows, together with reviewer's comments applicable to specific sections. After an introduction replete with suggestive examples, there is a review of wide-sense stationary processes. Moving average and autoregressive representations are introduced, and used to solve the l.l.s.p. problem (for that part of the sequence corresponding to the absolutely continuous spectral component) under the assumption that the spectral density is analytic and zero-free in an annulus containing the unit circle. There is a fine section on approximations by rational spectra; a useful approximating algorithm is presented, and the tempting pitfalls (see [7], p. 238, for a reference and discussion) are avoided.

Whittle next treats linear least square estimation of vector random variables, including also unbiased estimation of variables whose means are linear combinations with unknown regression coefficients. This chapter is routine, and can hardly be termed modern in treatment. For instance, the author uses the inverse of the covariance matrix, giving a lame argument for its existence; instead, he could have utilized the pseudoinverse without apology, and at no cost in difficulty (the construction of the pseudo-inverse is accomplished by a trivial variation on the method Whittle proposes for the inverse).

By minimizing with respect to the coefficients of the l.l.s.p. (a not very elegant method), the author derives the matrix-vector equation whose solution yields the l.l.s.p. for sequences. There is an extended discussion of projection on a finite sample (finite memory filter, in engineering terminology) and on interpolation, based on finite autoregression assumptions, and equivalently, on rational spectral densities for the continuous parameter case. Explicit solution methods are developed; although the same equations are treated in several engineering texts (see e.g. Appendix 2 of [1]), Whittle's exposition is the most lucid the reviewer has seen.

Then comes a chapter on the extrapolation of trends, and the fitting of simple types of deterministic components. L.l.s.p. of accumulated processes (which generalize autoregressive sequences) is discussed and solved. Especially interesting are a series of theorems giving conditions under which a l.l.s.p. for a stationary process is unbiased with respect to the addition of certain polynomial components.

The l.l.s.p. for multivariate stochastic processes is the subject of Chapter 9. The principal problem—not yet wholly solved—is that of spectral factorization. Whittle presents an algorithm based on the Yule-Walker relations, and applicable to reational spectral density matrices. It is not clear that Whittle's algorithm is more convenient than those of others (see again [7], p. 238, and also [5]), none of which he mentions or references.

The final chapter, occupying the last third of the book is concerned with linear least square error regulation (feedback control in American terms). It is at once the most rewarding and the most irritating part of the book. The treatment is generally heuristic, and a number of results admit of counter-examples unless additional conditions are imposed on the fixed element $\alpha(\cdot)$ of the control system. The system models principally analyzed will be barely recognized by control theorists; the usually applicable model is relegated to a series of exercises. In spite of these objections, the basic techniques of the chapter (while not original with the author) are valuable, and deserve to be more widely disseminated. The most stimulating subject, however, is the principle of certainty equivalence, which substitutes for stochastic optimization the optimization of the same quadratic form (without taking its expectation) with least square estimates substituted for all variables not yet observed. This principle, which is new to the reviewer, provides an alternative (sometimes more convenient) method for obtaining the equations of optimal control.

REFERENCES

[1] Davenport, W. and Root, W. (1958). *Introduction to the Theory of Random Signals and Noise*. McGraw-Hill, New York.

[2] Doob, J. L. (1953). *Stochastic Processes*. Wiley, New York.

[3] Grenander, U. and Rosenblatt, M. (1957). *Statistical Analysis of Stationary Time Series*. Wiley, New York.

[4] Tsien, H. S. (1954). *Engineering Cybernetics*. McGraw-Hill, New York.

[5] Wong, E. and Thomas, J. (1961). On the multidimensional prediction and filtering problem and the factorization of spectral matrices. *J. Franklin Inst.* **272** 87–99.

[6] Yaglom, A. M. (1962). *Introduction to the Theory of Stationary Random Functions*. Prentice-Hall, Englewood Cliffs, N.J.

[7] Zadeh, L. A., et al., (1963). Report on progress in information theory in the U.S.A. 1960–63. *IEEE Trans. of the Prof. Tech. Group on Information Th., IT-9*. 221–264.