1 Introduction

1.1 State Variables and Communication Theory

A state variable description of systems and random processes offers several advantages from both theoretical and practical viewpoints. From a theoretical aspect, such a description provides a very general characterization in terms of which a large class of systems, possibly time varying and nonlinear, can be modeled. Many powerful and elegant statements can be made with regard to systems described in this manner. From a practical aspect, they often provide a more representative physical description of the actual dynamics of the systems involved. More importantly, a state variable approach often leads to solution techniques that are readily implemented on a computer. This is highly desirable when specific numerical results are required.

The essential feature of a state variable approach is that the systems and processes of interest are described in terms of differential equations and their excitation, which is usually a white noise process. This is in contrast to the impulse response and covariance function description of systems and processes commonly used in the analysis of communication problems. Since a computer is ideally suited for integrating differential equations, one can easily see how a state variable formulation leads to effective computational solution methods.

In the area of automatic control, state variable concepts have been used extensively, so much so that they are the approach used in the majority of problems now studied. In communication theory, by contrast, they are not employed nearly as extensively. While a state variable description is certainly not appropriate in many situations, there are a large number of problems in communication theory where these concepts can be used advantageously. This monograph is directed to those people in communication theory who want to exploit some of the concepts and methods of state variables in the analysis of their problems.

The use of state variables is not novel in that they have already provided effective solutions to several important problems in communication theory. Undoubtedly, the most significant of these is the original work of Kalman and Bucy on optimal linear minimum mean square error realizable filtering.[35] In the classical approach, as used by Wiener, the random processes are represented in terms of their covariances. The impulse response of the optimum filter then is determined in terms of these covariances, or their associated spectra. Consequently, the optimal estimate is the result of the explicit operation of this impulse response upon the observed signal. In contrast, Kalman and Bucy represented the random processes in terms of state variables. They then found a structure for the optimum filter in which the desired estimate is specified implicitly as the solution to a set of differential equations. The principal advantage here is that it is usually more convenient, especially computationally, to implement solving the differential equations than it is to realize the operation implied by the impulse response.

Starting with the concepts introduced in their papers, several people have used state variable techniques in analyzing problems concerned with the detection and estimation of random processes. Particularly noteworthy contributions have been made by Schweppe in the detection of Gaussian random signals in Gaussian noise,[52] Kushner in the general theory of nonlinear filtering,[40] and Snyder in the application of state variable, nonlinear filtering to communication systems.[61] Certainly, many other results published in the control literature are also relevant to communication problems.

Here, we are principally interested in how state variables can be used effectively to solve several of the integral equations that frequently appear in communications theory. These equations and their associated theory assume an important role in communications. One often encounters situations where a fundamental result of a particular analysis is succinctly stated, or formulated, in terms of some appropriate integral equation that needs to be solved.

There are several prominent examples of this. The Karhunen-Loève theorem describes an orthonormal expansion of a random process where the set of orthonormal functions $\{\phi_i(t)\}$ is chosen such that the coefficients are uncorrelated.* These functions are specified by the solution of a homogeneous integral equation

$$\int_{T_0}^{T_f} \mathbf{K}_y(t, \tau)\phi_i(\tau)\, d\tau = \lambda_i \phi_i(t), \qquad T_0 \leqq t \leqq T_f, \tag{1.1}$$

where the kernel $\mathbf{K}_y(t, \tau)$ is the covariance of the expanded random process, $\phi_i(t)$ is an eigenfunction solution, and λ_i is its associated eigenvalue, which is equal to the mean square value of the ith generalized Fourier coefficient in the expansion. In many applications this expansion simply is done conceptually in the course of an analysis. There are several problems, however, where one is interested in the actual expansion, especially the eigenvalues. In these situations obtaining specific solutions reduces to solving this homogeneous integral equation.

A problem often encountered is the detection of a known signal in the presence of a colored noise.† Typically, on one hypothesis only a noise process $\mathbf{n}(t)$ with a covariance $\mathbf{K}_n(t, \tau)$ is present, while on the other hypothesis there is a known signal $\mathbf{s}(t)$ present in addition to the noise. The optimal receiver and its performance are specified by an inhomogeneous integral equation

$$\int_{T_0}^{T_f} \mathbf{K}_n(t, \tau)\mathbf{g}(\tau)\, d\tau = \mathbf{s}(t), \qquad T_0 \leqq t \leqq T_f. \tag{1.2}$$

The receiver correlates the observed signal with the solution $\mathbf{g}(t)$ of this integral equation, and the detector performance can be related to the integrated product of $\mathbf{g}(t)$ and $\mathbf{s}(t)$. Several problems in communications essentially reduce to solving this integral equation; therefore, obtaining a solution can often be of significant practical interest.

The Wiener-Hopf equation has a fundamental importance in much of communication theory.‡ In its general form, it specifies the optimum, linear, minimum mean square error estimator of a random signal in noise. This estimator appears frequently in both detection and estimation theory problems. For example, it is the estimator in the estimator-correlator for the detection of Gaussian random signals in noise, or it is used to compute the likelihood function for estimating the parameters of a random process imbedded in noise. This equation has the form

$$\int_{T_0}^{T_f} \mathbf{h}_0(t, u)\mathbf{K}_r(u, \tau)\, du = \mathbf{K}_{dr}(t, \tau), \qquad T_0 \leqq \tau \leqq T_f, \tag{1.3}$$

* Ref. 22, pp. 96–101, Ref. 67, pp. 178–198.
† Ref. 67, pp. 287–325, or Ref. 29, pp. 95–121.
‡ Ref. 67, Chap. 6 and Ref. 68, Chaps. 3–5.

where $\mathbf{K}_r(u, \tau)$ is the covariance of an observed signal, $\mathbf{K}_{dr}(t, T)$ is the cross covariance of the desired signal and the observed signal, and $\mathbf{h}_0(t, u)$ is the impulse response of the optimum estimator at time t. Usually we are interested in the estimate rather than this impulse response; consequently, in a state variable approach we derive a differential equation structure specifying the estimate and from which the impulse response can be obtained if desired.

When T_f is fixed, this estimator corresponds to the optimum unrealizable filter, often referred to as the optimal smoother. In contrast, in the problem solved by Kalman and Bucy, T_f increases in time and t equals or exceeds T_f. Although their work has an important place in much of our discussion, we are principally concerned with the cases of fixed T_f and when t is less than an increasing T_f by a fixed amount corresponding to the optimal smoother and filter with delay, respectively.

The commonly used impulse description of systems specifies a linear integral operator, and often one is led naturally to an integral equation in the analysis of many problems. Consequently, either in the abovementioned problems or in the context of some other, these three integral equations frequently appear in communication theory. By demonstrating how state variable concepts can be used to solve them, we can provide a useful approach to many problems that appear in communication theory.

1.2 Organization

We essentially divide the monograph into two sections. In the first, we develop from first principles the state variable solution techniques for homogeneous and inhomogeneous Fredholm integral equations. We make three essential assumptions. First, the kernel of the integral equation is a covariance function of a random process. This is the common situation in many communication theory problems. Second, a random process with this covariance function can be generated by exciting a linear system with white noise. This is analogous to generating a process with a specified spectrum by driving a system having an appropriate transfer function with white noise. Finally, we assume that the system for generating this process with the specified covariance has a known, finite dimensional state variable description of its input-output relationship. This assumption relates the description of the random processes by their covariances to their description as being generated via state variable methods.

Under these assumptions we can analyze problems involving a large class of kernels. Many kernels corresponding to covariances of the output of time-varying systems can be considered in addition to the important special case of stationary kernels, or covariances, with rational Fourier transforms corresponding to the outputs of time-invariant, or constant-parameter, systems.

To solve these integral equations, we first need to discuss how random processes propagate through linear systems described by state variables. In particular we need to develop the properties of the covariance functions of these processes. This is done in Chapter 2.

With this preliminary discussion, we study the homogeneous and inhomogeneous integral equations in Chapters 3 and 4, respectively. Both of these equations are reduced to two linear differential equations and an associated set of boundary conditions. The coefficients of these differential equations and the boundary conditions are specified directly by the matrices describing the system that generates the random process with the specified kernel.

The eigenvalues of the homogeneous equation are found to be solutions of a transcendental equation involving the transition matrix of the differential equations mentioned above. The eigenfunctions also follow directly. By using this same transcendental equation we can derive an effective method for calculating the Fredholm determinant function.

We then derive the differential equations and boundary conditions for the inhomogeneous integral equations. Since the resulting differential equations are identical with those that specify the structure of optimal smoother, we can exploit the solution techniques that have been developed in the literature for this problem. Throughout our analysis, we place our discussion in the context of the problem of detecting a known signal in the presence of colored noise.

In the second section of this monograph, we discuss two specific applications of our integral equation theory. We can observe the utility of both the actual results of the theory and the approaches used in deriving it in this context. In Chapter 5 we consider the design of optimal signals for detection in the presence of colored noise using modern optimal control theory. We focus our attention on additive signal-independent noise channels. When energy and mean square bandwidth constraints are imposed, we are able to solve the signal design problem completely. We present some specific examples, indicating both the optimal signals and performance gain over more conventional signals. While we focus on this problem, however, our approach is not limited to this class of channels. We can consider different constraints and some signaldependent channels, including some reverberation models. To date, however, the nonlinear equations resulting from the application of the optimal control theory have not proved at all easy to solve.

In Chapter 6, we consider the estimation problems of optimum linear minimum mean square error smoothing and filtering with delay of random processes that are observed corrupted by additive noise. One of the central issues is the solution of the Wiener-Hopf equation. In the filtering problem as studied by Kalman and Bucy, the estimate is made at the end point of an observation interval using all the available past data. Since the estimate is optimum only at a single point for any particular observation interval, the filter generates what is often termed a point estimate. This observation interval increases in time as more data are received, and the filter generates a sequence of estimates, each of which is an optimum realizable, or causal, estimate of the signal process at the end point of the observation interval defined at that specific instant of time. In contrast, the smoother is a noncausal interval estimator, roughly analogous to the unrealizable filter. For a fixed observation interval, it generates an optimum estimate of the process over the entire interval. Like the Kalman-Bucy filter, the filter with delay is a point estimator with an evolving end point. However, we make an estimate at an interior point within the observation interval rather than at the end point. By allowing the delay, we can improve our estimator performance over that of the Kalman filter, and still use a realizable filter whose output evolves in time as more data are received. In both of these problems we derive the estimator structure and its associated performance. The results of our integral equation theory are the starting point for our approach.

In Chapter 7 we briefly consider some aspects of nonlinear estimation theory which can be approached using our methods. To do this we need to change our estimation criterion to one of maximum *a posteriori* probability and restrict ourselves to Gaussian processes that have been observed by means of a nonlinear modulation. Here the solution methods become rather difficult, and we are quickly led to approximate techniques.

Throughout the monograph we present many examples. We do this for two reasons. We work a number of analytic examples to illustrate the use of methods we derive. We also present a number of examples analyzed by numerical methods. In the course of the monograph we emphasize the numerical aspects of our methods. We feel this is where the major application of much of the material lies. Finding effective numerical procedures is a very relevant problem, since most problems are two complex to be analyzed analytically.

1.3 Notation

We also indicate our notational conventions. Generally, scalars are symbols in italic type, vectors are lower case symbols in boldface type, and matrices are upper case symbols in boldface type.