Stochastic models, estimation, and control

VOLUME 1

PETER S. MAYBECK

Stochastic models, estimation, and control
VOLUME 1

PETER S. MAYBECK
DEPARTMENT OF ELECTRICAL ENGINEERING
AIR FORCE INSTITUTE OF TECHNOLOGY
WRIGHT-PATTERSON AIR FORCE BASE
OHIO

ACADEMIC PRESS New York San Francisco London 1979
A Subsidiary of Harcourt Brace Jovanovich, Publishers

DR36/4

ACADEMIC PRESS, INC.
111 Fifth Avenue, New York, New York 10003

United Kingdom Edition published by
ACADEMIC PRESS, INC. (LONDON) LTD.
24/28 Oval Road, London NW1 7DX

Library of Congress Cataloging in Publication Data

Maybeck, Peter S
Stochastic models, estimation and control.

(Mathematics in science and engineering ; v.)
Includes bibliographies.
1. System analysis. 2. Control theory. 3. Estimation theory. I. Title. II. Series.
QA402.M37 519.2 78-8836
ISBN 0-12-480701-1 (v. 1)

PRINTED IN THE UNITED STATES OF AMERICA

82 9 8 7 6 5 4 3 2

Preface

The purpose of this book is twofold. First, it attempts to develop a thorough understanding of the *fundamental concepts* incorporated in stochastic processes, estimation, and control. Furthermore, it provides some experience and insights into *applying* the theory to realistic practical problems.

The approach taken is oriented toward an *engineer* or an engineering student. We shall be interested not only in mathematical results, but also in a *physical interpretation* of what the mathematics means. In this regard, considerable effort will be expended to generate graphical representations and to exploit geometric insights where possible. Moreover, our attention will be concentrated upon eventual implementation of estimation and control algorithms, rather than solely upon rigorous derivation of mathematical results in their most general form. For example, all assumptions will be described thoroughly in order to yield precise results, but these assumptions will further be *justified* and their *practical implications* pointed out explicitly. Topics where additional generality or rigor can be incorporated will also be delineated, so that such advanced disciplines as functional analysis can be exploited by, but are not required of, the reader.

Because this book is written for engineers, we shall avoid measure theory, functional analysis, and other disciplines that may not be in an engineer's background. Although these fields can make the theoretical developments more rigorous and complete, they are not essential to the practicing engineer who wants to *use* optimal estimation theory results. Furthermore, the book can serve as a text for a first-year graduate course in estimation and stochastic control, and these advanced disciplines are not generally studied prior to such a course. However, the places where these disciplines do contribute will be pointed out for those interested in pursuing more rigorous developments. The developments in the text will also be motivated in part by the concepts of analysis and functional analysis, but without requiring the reader to be previously familiar with these fields. In this way, the reader will become aware of the kinds of questions that have to be answered in a completely rigorous derivation and will be introduced to the concepts required to resolve them properly.

This work is intended to be a text from which a reader can *learn* about estimation and stochastic control, and this intent has dictated a format of presentation. Rather than strive for the mathematical precision of a theorem–proof structure, fundamentals are first motivated conceptually and physically, and then the mathematics developed to serve the purpose. Practical aspects and eventual implementation of algorithms are kept at the forefront of concern. Finally, the progression of topics is selected to maximize learning: a firm foundation in linear system applications is laid before nonlinear applications are considered, conditional probability density functions are discussed before conditional expectations, and so forth. Although a reference book might be organized from the most general concepts progressively to simpler and simpler special cases, it has been our experience that people grasp basic ideas and understand complexities of the general case better if they build up from the simpler problems. As generalizations are made in the text, care is taken to point out all ramifications—what changes are made in the previous simpler case, what concepts generalize and how, what concepts no longer apply, and so forth.

With an eye to practicality and eventual implementations, we shall emphasize the case of continuous-time dynamic systems with *discrete-time* data sampling. Most applications will in fact concern continuous-time systems, while the actual estimator or controller implementations will be in the form of software algorithms for a digital computer, which inherently involves data samples. These *algorithms* will be developed in detail, with special emphasis on the various *design tradeoffs* involved in achieving an efficient, practical configuration.

The corresponding results for the case of continuously available measurements will be presented, and its range of applicability discussed. However, only a formal derivation of the results will be provided; a rigorous derivation, though mathematically enriching, does not seem warranted because of this limited applicability. Rather, we shall try to develop physical insights and an engineering appreciation for these results.

Throughout the development, we shall regard the digital computer not only as the means for eventual implementation of on-line algorithms, but also as a *design tool* for generating the final "tuned" algorithms themselves. We shall develop means of synthesizing estimators or controllers, fully evaluating their performance capabilities in a real-world environment, and iterating upon the design until performance is as desired, all facilitated by software tools.

Because the orientation is toward engineering applications, *examples* will be exploited whenever possible. Unfortunately, even under our early restrictions of a linear system model driven by white Gaussian noises (these assumptions will be explained later), simple estimation or control examples are difficult to generate—either they are simple enough to work manually and are of little value, or are useful, enlightening, and virtually impossible to do by hand. At first, we shall try to gain *insights* into algorithm structure and behavior by solving relatively simple problems. Later, more complex and realistic problems will be considered in order to appreciate the *practical aspects* of estimator or controller implementation.

This book is the outgrowth of the first course of a two-quarter sequence taught at the Air Force Institute of Technology. Students had previously taken a course in applied probability theory, taught from the excellent Chapters 1–7 of Davenport's "Probability and Random Processes." Many had also been exposed to a first control theory course, linear algebra, linear system theory, deterministic optimal control, and random processes. However, considerable attention is paid to those fundamentals in Chapters 2–4, before estimation and stochastic control are developed at all. This has been done out of the conviction that system modeling is a critical aspect, and typically the "weak link," in applying theory to practice.

Thus the book has been designed to be self-contained. The reader is assumed to have been exposed to advanced calculus, differential equations, and some vector and matrix analysis on an engineering level. Any more advanced mathematical concepts will be developed within the text, requiring only a willingness on the part of the reader to deal with new means of conceiving a problem and its solution. Although the mathematics becomes relatively sophisticated at times, efforts are made to motivate the need for, and stress the underlying basis of, this sophistication. The objective is to investigate the theory and derive from it the tools required to reach the ultimate objective of generating practical designs for estimators and stochastic controllers.

The author wishes to express his gratitude to the students who have contributed significantly to the writing of this book through their helpful suggestions and encouragement. The stimulation of technical discussions and association with Professors John Deyst, Wallace Vander Velde, and William Widnall of the Massachusetts Institute of Technology and Professors Jurgen Gobien, James Negro, and J. B. Peterson of the Air Force Institute of Technology has also had a profound effect on this work. Appreciation is expressed to Dr. Robert Fontana, Head of the Department of Electrical Engineering, Air Force Institute of Technology, for his support throughout this endeavor, and to those who carefully and thoroughly reviewed the manuscript. I am also deeply grateful to my wife, Beverly, whose understanding and constant supportiveness made the fruition of this effort possible.

Notation

Vectors, Matrices

Scalars are denoted by upper or lower case letters in italic type.

Vectors are denoted by lower case letters in boldface type, as the vector $\mathbf{x}$ made up of components x_i.

Matrices are denoted by upper case letters in boldface type, as the matrix $\mathbf{A}$ made up of elements A_{ij} (ith row, jth column).

Random Vectors (Stochastic Processes), Realizations (Samples), and Dummy Variables

Random vectors are set in boldface sans serif type, as x made up of scalar components x_i.

Realizations of the random vector are set in boldface roman type, as $\mathbf{x}$: $\mathsf{x}(\omega_i) = \mathbf{x}$.

Dummy variables (for arguments of density or distribution functions, integrations, etc.) are denoted by the equivalent Greek letter, such as $\boldsymbol{\xi}$ being associated with x: e.g., $f_{\mathsf{x}}(\boldsymbol{\xi})$. The correspondences are $(\mathsf{x}, \boldsymbol{\xi})$, $(\mathsf{y}, \boldsymbol{\rho})$, $(\mathsf{z}, \boldsymbol{\zeta})$, $(\mathsf{Z}, \mathscr{Z})$.

Subscripts

a:	augmented	c:	continuous-time
d:	discrete-time	t:	true, truth model

Superscripts

$^{\mathrm{T}}$:	transpose (matrix)	$\bar{\ }$:	Fourier transform
$^{-1}$:	inverse (matrix)	$\hat{\ }$:	estimate
$*$:	complement (set) or complex conjugate		

Matrix and Vector Relationships

$\mathbf{A} > \mathbf{0}$: $\mathbf{A}$ is positive definite.

$\mathbf{A} \geq \mathbf{0}$: $\mathbf{A}$ is positive semidefinite.

$\mathbf{x} \leq \mathbf{a}$: componentwise, $x_1 \leq a_1, x_2 \leq a_2, \ldots,$ and $x_n \leq a_n$.

List of symbols and pages where they are defined or first used

Contents

Chapter 6 Design and performance analysis of Kalman filters

Chapter 7 Square root filtering

Contents of Volume 2

CHAPTER 1

Introduction

1.1 WHY STOCHASTIC MODELS, ESTIMATION, AND CONTROL?

When considering system analysis or controller design, the engineer has at his disposal a wealth of knowledge derived from *deterministic* system and control theories. One would then naturally ask, why do we have to go beyond these results and propose *stochastic* system models, with ensuing concepts of estimation and control based upon these stochastic models? To answer this question, let us examine what the deterministic theories provide and determine where the shortcomings might be.

Given a physical system, whether it be an aircraft, a chemical process, or the national economy, an engineer first attempts to develop a mathematical model that adequately represents some aspects of the behavior of that system. Through physical insights, fundamental "laws," and empirical testing, he tries to establish the interrelationships among certain variables of interest, inputs to the system, and outputs from the system.

With such a mathematical model and the tools provided by system and control theories, he is able to investigate the system structure and modes of response. If desired, he can design compensators that alter these characteristics and controllers that provide appropriate inputs to generate desired system responses.

In order to observe the actual system behavior, measurement devices are constructed to output data signals proportional to certain variables of interest. These output signals and the known inputs to the system are the only information that is directly discernible about the system behavior. Moreover, if a feedback controller is being designed, the measurement device outputs are the only signals directly available for inputs to the controller.

There are three basic reasons why deterministic system and control theories do not provide a totally sufficient means of performing this analysis and design. First of all, *no mathematical system model is perfect*. Any such model depicts only those characteristics of direct interest to the engineer's purpose. For instance, although an endless number of bending modes would be required to depict vehicle bending precisely, only a finite number of modes would be included in a useful model. The objective of the model is to represent the dominant or critical modes of system response, so many effects are knowingly left unmodeled. In fact, models used for generating online data processors or controllers must be pared to only the basic essentials in order to generate a computationally feasible algorithm.

Even effects which are modeled are necessarily *approximated* by a mathematical model. The "laws" of Newtonian physics are adequate approximations to what is actually observed, partially due to our being unaccustomed to speeds near that of light. It is often the case that such "laws" provide adequate system *structures*, but various *parameters* within that structure are not determined absolutely. Thus, there are many sources of uncertainity in any mathematical model of a system.

A second shortcoming of deterministic models is that dynamic systems are driven not only by our own control inputs, but also by *disturbances which we can neither control nor model deterministically*. If a pilot tries to command a certain angular orientation of his aircraft, the actual response will differ from his expectation due to wind buffeting, imprecision of control surface actuator responses, and even his inability to generate exactly the desired response from his own arms and hands on the control stick.

A final shortcoming is that *sensors do not provide perfect and complete data* about a system. First, they generally do not provide all the information we would like to know: either a device cannot be devised to generate a measurement of a desired variable or the cost (volume, weight, monetary, etc.) of including such a measurement is prohibitive. In other situations, a number of different devices yield functionally related signals, and one must then ask how to generate a best estimate of the variables of interest based on partially redundant data. Sensors do not provide exact readings of desired quantities, but introduce their own system dynamics and distortions as well. Furthermore, these devices are also always noise corrupted.

As can be seen from the preceding discussion, to assume perfect knowledge of all quantities necessary to describe a system completely and/or to assume perfect control over the system is a naive, and often inadequate, approach. This motivates us to ask the following four questions:

(1) How do you develop system models that account for these uncertainties in a direct and proper, yet practical, fashion?

(2) Equipped with such models and incomplete, noise-corrupted data from available sensors, how do you optimally estimate the quantities of interest to you?

(3) In the face of uncertain system descriptions, incomplete and noise-corrupted data, and disturbances beyond your control, how do you optimally control a system to perform in a desirable manner?

(4) How do you evaluate the performance capabilities of such estimation and control systems, both before and after they are actually built?

This book has been organized specifically to answer these questions in a meaningful and useful manner.

1.2 OVERVIEW OF THE TEXT

Chapters 2–4 are devoted to the stochastic modeling problem. First Chapter 2 reviews the pertinent aspects of deterministic system models, to be exploited and generalized subsequently. Probability theory provides the basis of all of our stochastic models, and Chapter 3 develops both the general concepts and the natural result of static system models. In order to incorporate dynamics into the model, Chapter 4 investigates stochastic processes, concluding with practical linear dynamic system models. The basic form is a linear system driven by white Gaussian noise, from which are available linear measurements which are similarly corrupted by white Gaussian noise. This structure is justified extensively, and means of describing a large class of problems in this context are delineated.

Linear estimation is the subject of the remaining chapters. Optimal filtering for cases in which a linear system model adequately describes the problem dynamics is studied in Chapter 5. With this background, Chapter 6 describes the design and performance analysis of practical online Kalman filters. Square root filters have emerged as a means of solving some numerical precision difficulties encountered when optimal filters are implemented on restricted word-length online computers, and these are detailed in Chapter 7.

Volume 1 is a complete text in and of itself. Nevertheless, Volume 2 will extend the concepts of linear estimation to smoothing, compensation of model inadequacies, system identification, and adaptive filtering. Nonlinear stochastic system models and estimators based upon them will then be fully developed. Finally, the theory and practical design of stochastic controllers will be described.

1.3 THE KALMAN FILTER: AN INTRODUCTION TO CONCEPTS

Before we delve into the details of the text, it would be useful to see where we are going on a conceptual basis. Therefore, the rest of this chapter will provide an overview of the optimal linear estimator, the Kalman filter. This will be conducted at a very elementary level but will provide insights into the underlying concepts. As we progress through this overview, contemplate the ideas being presented: try to conceive of graphic *images* to portray the concepts involved (such as time propagation of density functions), and to generate a *logical structure* for the component pieces that are brought together to solve the estimation problem. If this basic conceptual framework makes sense to you, then you will better understand the need for the details to be developed later in the text. Should the idea of where we are going ever become blurred by the development of detail, refer back to this overview to regain sight of the overall objectives.

First one must ask, what is a Kalman filter? A Kalman filter is simply an *optimal recursive data processing algorithm*. There are many ways of defining *optimal*, dependent upon the criteria chosen to evaluate performance. It will be shown that, under the assumptions to be made in the next section, the Kalman filter is optimal with respect to virtually any criterion that makes sense. One aspect of this optimality is that the Kalman filter incorporates all information that can be provided to it. It processes all available measurements, regardless of their precision, to estimate the current value of the variables of interest, with use of (1) knowledge of the system and measurement device dynamics, (2) the statistical description of the system noises, measurement errors, and uncertainty in the dynamics models, and (3) any available information about initial conditions of the variables of interest. For example, to determine the velocity of an aircraft, one could use a Doppler radar, or the velocity indications of an inertial navigation system, or the pitot and static pressure and relative wind information in the air data system. Rather than ignore any of these outputs, a Kalman filter could be built to combine all of this data and knowledge of the various systems' dynamics to generate an overall best estimate of velocity.

The word *recursive* in the previous description means that, unlike certain data processing concepts, the Kalman filter does not require all previous data to be kept in storage and reprocessed every time a new measurement is taken. This will be of vital importance to the practicality of filter implementation.

The "filter" is actually a *data processing algorithm*. Despite the typical connotation of a filter as a "black box" containing electrical networks, the fact is that in most practical applications, the "filter" is just a computer program in a central processor. As such, it inherently incorporates discrete-time measurement samples rather than continuous time inputs.

Figure 1.1 depicts a typical situation in which a Kalman filter could be used advantageously. A system of some sort is driven by some known controls, and measuring devices provide the value of certain pertinent quantities. Knowledge of these system inputs and outputs is all that is explicitly available from the physical system for estimation purposes.

The *need* for a filter now becomes apparent. Often the variables of interest, some finite number of quantities to describe the "state" of the system, cannot