A. V. BALAKRISHNAN

KALMAN FILTERING THEORY

A. V. BALAKRISHNAN

KALMAN FILTERING THEORY

OPTIMIZATION SOFTWARE, INC.
PUBLICATIONS DIVISION, NEW YORK

Author

A.V. Balakrishnan
Electrical Engineering Department
School of Engineering and Applied Science
University of California
Los Angeles, California 90024

Library of Congress Cataloging-in-Publication Data

Balakrishnan, A. V.
Kalman filtering theory / A.V. Balakrishnan.
p. cm. — Series in communication and control systems)
Bibliography: p.
Includes index.
ISBN 0-911575-49-9
1. Control theory. 2. Estimation theory. 3. Kalman filtering.
I. Title. II. Series.
QA402.3.B287 1987
003 — dc19 87-31312
CIP

SERIES IN COMMUNICATION AND CONTROL SYSTEMS

Translations Series in Mathematics and Engineering

M. I. Yadrenko
Spectral Theory of Random Fields
1983, 267 pp., ISBN 0-911575-00-6

G. I. Marchuk
Mathematical Models in Immunology
1983, 378 pp., ISBN 0-911575-01-4

A. A. Borovkov, ed.
Advances in Probability Theory: Limit Theorems and Related Problems
1984, 392 pp., ISBN 0-911575-03-0

V. A. Dubovitskij
The Ulam Problem of Optimal Motion of Line Segments
1985, 128 pp., ISBN 0-911575-04-9

N. V. Krylov, R. Sh. Liptser, and A. A. Novikov, eds.
Statistics and Control of Stochastic Processes
1985, 521 pp., ISBN 0-911575-18-9

Yu. G. Evtushenko
Numerical Optimization Techniques
1985, 575 pp., ISBN 0-911575-07-3

V. F. Dem'yanov and L. V. Vasil'ev
Nondifferentiable Optimization
1985, 472 pp., ISBN 0-911575-09-X

A. A. Borovkov, ed.
Advances in Probability Theory: Limit Theorems for Sums of Random Variables
1985, 313 pp., ISBN 0-911575-17-0

A. V. Balakrishnan, A. A. Dorodnitsyn, and J. L. Lions, eds.
Vistas in Applied Mathematics: Numerical Analysis, Atmospheric Sciences, Immunology
1986, 396 pp., ISBN 0-911575-38-3

V. F. Kolchin
Random Mappings
1986, 224 pp., ISBN 0-911575-16-2

L. Telksnys, ed.
Detection of Changes in Random Processes
1986, 240 pp., ISBN 0-911575-20-0

V. F. Dem'yanov and A. M. Rubinov
Quasidifferential Calculus
1986, 301 pp., ISBN 0-911575-35-9

V. P. Chistyakov, B. A. Sevast'yanov, and V. K. Zakharov
Probability Theory for Engineers
1987, 175 pp., ISBN 0-911575-13-8

B. T. Polyak
Introduction to Optimization
1987, 464 pp., ISBN 0-911575-14-6

V. V. Ivanishchev and A. D. Krasnoshchekov
Control of Variable Structure Networks
1987, 154 pp., ISBN 0-911575-05-7

Series in Communication and Control Systems

R. E. Kalman, G. I. Marchuk, A. E. Ruberti, & A. J. Viterbi, eds.
Recent Advances in Communication and Control Theory
1987, 500 pp., ISBN 0-911575-46-4

Scientific Applications Software

Kalman Filtering Software
Diskette & User's Guide
1985, ISBN 0-911575-31-6
Updated version in Quick-Basic, 1987

ABOUT THE AUTHOR

A.V. Balakrishnan is Professor of Electrical Engineering and Professor of Mathematics at UCLA. He is a Fellow of the IEEE, and past Chairman of the IEEE Information Theory Group. He received a Certificate of Recognition from the National Aeronautics and Space Administration for his scientific contributions to the processing of Flight Test data. He was awarded the Guillemin Prize for his many contributions to Communication and Control Systems Engineering. He is also the recipient of the Silver Core Award from the International Federation of Information Processing, and more recently, of a Group Achievement Award from NASA Langley FRC for his leadership role in the Space Craft Control Laboratory Experiment (SCOLE). He is the author of several books and monographs in Applied Mathematics and Engineering, and serves on the Editorial Boards of numerous journals and monograph series.

PREFACE TO THE SECOND EDITION

The warm reception by the industrial community of the first edition intended primarily as a text for graduate students has emboldened me in this second edition to strive to make it palatable to both audiences. The new edition is considerably revised and enlarged with this end in mind. Two new chapters have been added and much of the earlier material expanded with more emphasis on applications.

The Kalman Filter is many things to many people—the Numerical Analysts see it as Least Squares, the Communication and Circuit engineers view it as Low-pass Filter. For the Control engineers it is part of Stochastic Control, while for the Astrodynamic community it is Differential Correction. I have chosen to follow in the main the historical development, emphasizing the Statistical Estimation (Maximum Likelihood) framework and the System theoretic aspects, essential for proofs of asymptotic properties. The presentation underscores what can be proved—as an antidote to the tendency to view it as a piece of hardware, to be tweeked to make it work: to remind users that simulation must be backed by solid theory, even though, of course, each application has its own unique quirks and the modelling part—of wedding the application to the theory—can only be learned by experience.

The structure of the first five chapters has been retained, even as the presentation is beefed up and more application-oriented examples are incorporated. Chapters 1 and 2 review relevant state-space theory and stochastic process theory, respectively. Chapter 3 treats statistical estimation theory, and now includes an introduction to computational aspects: Local Linearization and the Newton-Raphson algorithm, as well as new examples drawn from Navigation and Tracking. Kalman Filtering Theory per se is still covered in Chapter 4, but expanded, featuring additional examples—filtering of accelerometer data, to mention one. Chapter 5 develops Likelihood Ratio formulas using Kalman Filtering Theory. Chapters 6 and 7 are totally new. Chapter 6 deals with EKF—the Extended Kalman Filter, because of its wide-spread use in industry. In essence, this is an ad hoc linearization procedure employed in many applications—such as orbit determination. We show that the Maximum Likelihood Principle provides at once the rationale, and the limitations. We develop a modified Newton-Raphson algorithm for nonlinear problems, which we show can be instrumented as a Kalman Smoother in each iteration. The final chapter, Chapter 7, treats Stochastic Control Theory—specifically the Linear Quadratic Regulator problem—because of its "duality" with Kalman Filtering. The Bellman dynamic programming argument provides a compact way to establish optimality of the Feedback Controls. The steady-state, or asymptotic theory, exploits the results of Kalman Filtering Theory in an essential way.

I hope that the book, while still not a compendium, will prove useful to a variety of readers—to communication and control engineers, numerical analysts, astrodynamicists, management scientists, econometricians, and other specialists both in academe and industry, who need to know Kalman Filtering theory.

PREFACE TO THE FIRST EDITION

This is a textbook intended for a one-quarter (or one-semester, depending on the pace) course at the graduate level in Engineering. The prerequisites are Elementary State-Space theory and Elementary (second-order Gaussian) Stochastic Process theory. As a textbook, it does not purport to be a compendium of all known work on the subject. Neither is it a "trade book." Rather it attempts a logically sequenced set of topics of proven pedagogical value, emphasizing theory while not devoid of practical utility. The organization is based on experience gained over a period of 10 years of classroom teaching. It develops those aspects of Kalman Filtering lore which can be given a firm mathematical basis, avoiding the industry syndrome manifest in professional short courses: "Here is the recipe. Use it, it will 'work'!"

The first two chapters cover review material on State-Space theory and Signal (Random Process) theory—necessary but not sufficient for the sequel. The third chapter deals with Statistical Estimation theory, the mathematical framework on which Kalman Filtering rests. The main chapter is the fourth chapter dealing with the subject matter per se. It begins in Section 4.1, with the basic theory and formulas,making a compromise in generality between too many obscuring details and too little practical application. Thus we consider only the case where the observation noise is white and is independent of the signal, although we allow the system to be time-variant. Because of the uncertainty in the initial covariances, in practice no Kalman filter can be optimal except in the steady state—and this is by far its important use. Hence Section 4.2 specializes to time-invariant systems and considers asymptotic behavior of the filter. Section 4.3 examines the steady-state results from the frequency-domain point of view, relating them to the more classical transfer-function approach. In Section 4.4 we study a canonical application of Kalman filtering: to System Identification. In Section 4.5 we study the "Kalman smoother": the on-line version of two-sided interpolation. In Sections 4.6 and 4.7 we study generalizations of the basic theory of Section 4.1; thus we allow the signal and noise to be correlated in Section 4.6, and allow the observation noise to be non-white in Section 4.7. Section 4.8 features a simple example which illustrates some of the theory and techniques discussed in the chapter.

The book concludes with a chapter on Likelihood Ratios in which the Kalman filter formulation plays an essential role.

We only consider discrete-time models throughout, since all Kalman filter implementation envisaged involves digital computation.

The problems accompanying each chapter serve the traditional role of testing the student's comprehension of the text, with an occasional foray into areas of contiguous interest.

NOTATION

Square matrices

I denotes Identity matrix

Tr. denotes Trace

$|A|$ denotes Determinant of A

Rectangular matrices

A^* denotes (Conjugate) Transpose of A.

$[A, B]$ denotes Tr. AB^*.

$\|A\|_0$ denotes Operator norm.

$\{a_{ij}\}$ denotes Matrix with entries a_{ij}.

'Column' vector $\nu \sim \nu^*\nu$ is 1×1.

'Row' vector $c \sim cc^*$ is 1×1.

R^n denotes Euclidean space of dimension $n \equiv$ Linear space of $n \times 1$ column vectors.

Self-adjoint matrices

A: Self-adjoint $A = A^*$.

Nonnegative Definite: $[Ax, x] \geqslant 0$ for every x.

$A \geqslant B \leftrightarrow (A - B)$ is self-adjoint nonnegative definite.

Gradient of a function

If $g(\theta)$ is a scalar function of θ:

Gradient of $g(\theta) = \nabla_\theta g(\theta)$,

where:

$$(\nabla_\theta g(\theta))h = (\frac{d}{d\lambda} g + \lambda h)|_{\lambda=0},$$

$$\nabla_\theta g(\theta) \text{ is } 1 \times m \text{ if } \theta \text{ is } m \times 1.$$

If $g(\cdot)$ maps $R^m \to R^n$, then $\nabla_\theta g(\theta)$ is $n \times m$.

Random Variables

$\mathbf{E}(\cdot)$ denotes Expected value.

$\mathbf{E}(\cdot|\cdot)$ denotes Conditional expectation.

$p(\cdot)$ denotes Probability density function.

Integral

$d|\nu|$ denotes Volume element in R^n as in $\int_{R^n} f(\nu)d|\nu|$.

Kronecker delta

$\delta_n^m = 0,\ m \neq n,$

$\quad = 1,\ m = n.$

TABLE OF CONTENTS

CHAPTER 1

REVIEW OF LINEAR SYSTEM THEORY

A Kalman filter is a linear system. This chapter presents a brief review of Linear Systems Theory from the "state-space" point of view, since the Kalman filter is best described in that way. For a succinct introductory treatment of State-Space theory, the reader is referred to [3]. More elaborate treatment may be found in [11], [17], [27], among other texts.

A system is characterized by its "input," its "state" and the "output." These are functions of time. Time may be continuous or discrete; in the latter case time is indexed by the integers. We shall only be concerned with the discrete case in this book.

Let $\{u_n\}$ denote the input and $\{v_n\}$ the output. A linear system for our purposes is then completely characterized by a "state-input" equation

$$x_{n+1} = A_n x_n + B_n u_n \tag{1.1}$$

and by an "output-state-input" equation

$$v_n = C_n x_n + D_n u_n, \tag{1.2}$$

where A_n is a square matrix and B_n, C_n, D_n are rectangular matrices. If the state-space dimension is p, then A_n will be $p \times p$. If the input sequence is such that each u_n is $q \times 1$, then B_n will be $p \times q$. If the output sequence is such that each v_n is $m \times 1$, then C_n will be $m \times p$ and D_n will be $m \times q$. We can "solve" (1.1), (1.2) or express the output in terms of the state at some initial time and the input from then on. Thus we have, taking the initial (or starting) time to be k:

$$x_n = \psi_{n,k} x_k + \sum_{i=k}^{n-1} \psi_{n,i+1} B_i u_i, \tag{1.3}$$

where $\psi_{n,k}$, called the State-Transition Matrix, is defined by

$$\psi_{n,k} = A_{n-1} \cdots A_k, \qquad k \leqslant n-1 \cdot \tag{1.4}$$
$$\psi_{n,n} = I \qquad \text{(Identity Matrix)}\dagger.$$

Note that it has the "transition" property:

$$\psi_{n,k}\psi_{k,m} = \psi_{n,m}. \tag{1.5}$$

From (1.3), which specifies the state at any time n, $n \geqslant k$, the output is readily expressed explicitly in terms of the "initial" state and the current input as

$$v_n = C_n \psi_{n,k} x_k + \sum_{i=k}^{n-1} C_n \psi_{n,i+1} B_i u_i + D_n u_n. \tag{1.6}$$

Here the first term is the "initial state" (or initial condition) response and the second term is the "input response." The function

$$W_{n,i} = C_n \psi_{n,i+1} B_i, \qquad i < n, \tag{1.7}$$
$$= D_n, \qquad i = n,$$

is referred to as the "weighting matrix" or "weighting pattern" of the system.

Time-Invariant Systems

We are most concerned with the case where the system is "time-invariant," where the system matrices are all independent of time:

$$A_n = A,$$
$$B_n = B,$$
$$C_n = C,$$
$$D_n = D,$$

---

† Here and throughout, the letter I will always denote the identity matrix regardless of dimension.